BOOK To READER

* Thank you for not tearing my pages.
* Grateful for not Writing Comments or putting unsightly markings.

CALLING URGENT ATTENTION

1. Tearing of pages causes permanent damage to the Book. Please think of the Reader who finds missing pages after reading that far. It is cruelty to the innocent.

2. Writing Comments and putting markings disfigure the Book. Please take care.

Secretary

1003

സാമ്പിളന സർവെ

കെ. രാമകൃഷ്ണപിള്ള

കേരള ഭാഷാ ഇൻസ്റ്റിററ്യൂട്ട്
തിരുവനന്തപുരം

ജനറൽ എഡിററർ

എൻ. വി. കൃഷ്ണവാരിയർ

ചീഫ് എഡിററർ

പ്രൊ. സി. കെ. മൂസ്സത്

സാങ്കേതിക എഡിററർ

സി. പി. നാരായണൻ

വിദഗ്ധോപദേശകൻ

പി. എസ്. ജനാർദനൻ നായർ

പ്രൂഫ്

കെ. അബ്ദുല്ല

ആമുഖം

സർവകലാശാലാ നിലവാരത്തിലുള്ള ഗ്രന്ഥങ്ങൾ മലയാളത്തിൽ നിർമിക്കുന്നതിനു് കേന്ദ്രഗവണ്മെണ്ടിന്റെ ധനസഹായത്തോടെ നടപ്പിൽ വരുത്തിക്കൊണ്ടിരിക്കുന്ന പരിപാടിയിൽ ഉൾപ്പെടുത്തിയാണു് "സാമ്പിളന സർവെ" പ്രസിദ്ധം ചെയ്യുന്നതു്. ഈ പരിപാടിയനുസരിച്ചു് സാംഖ്യികത്തിൽ ഇരുപത്തിമൂന്നു പുസ്തകങ്ങൾ തയ്യാറാക്കാൻ ഉദ്ദേശിക്കുന്നു. ഇവയിൽ "സാംഖ്യികം–1" ഇതിനകം പ്രസിദ്ധീകരിച്ചു കഴിഞ്ഞിട്ടുണ്ടു്.

സാംഖ്യികം ഐച്ഛികവിഷയമായി സ്വീകരിച്ചിട്ടുള്ള ബി. എസ്‌സി. വിദ്യാർഥികൾക്കു് പാഠപുസ്തകമായി ഉദ്ദേശിക്കപ്പെട്ടതാണു് ഈ ഗ്രന്ഥം. സാമ്പിളനസർവെ ഇന്നു് ജീവശാസ്ത്രം, ധനശാസ്ത്രം മുതലായ പല ശാസ്ത്രശാഖകളിലും പ്രയോഗിക്കപ്പെട്ടുവരുന്നുണ്ടു്. അതിനാൽ ആ ശാസ്ത്രങ്ങളിൽ പഠനം നടത്തുന്നവർക്കും പ്രവത്തിക്കുന്നവർക്കും സാമ്പിളനത്തിന്റെ അടിസ്ഥാനപ്രമാണങ്ങൾ മനസ്സിലാക്കുവാൻ "സാമ്പിളന സർവെ" ഉപകരിക്കും. ഈ ഉദ്ദേശം മുൻനിർത്തി, ഗണിതസംജ്ഞകൾ കഴിയുന്നത്ര ഒഴിവാക്കാനും പ്രതിപാദനം ലളിതമാക്കാനും ഈ ഗ്രന്ഥത്തിൽ ശ്രമിച്ചിട്ടുണ്ടു്.

പാലാ സെൻറ്തോമസ് കോളേജിലെ സാംഖ്യികപ്രഫസ്സർ ശ്രീ. കെ. രാമകൃഷ്ണപിള്ളയാണു് ഈ പുസ്തകം തയ്യാറാക്കിയിട്ടുള്ളതു്. കേരള സർവകലാശാലയിലെ സാംഖ്യിക റീഡർ ശ്രീ. പി. എസ്. ജനാർദനൻനായർ ഇതു് നിഷ്കൃഷ്ടമായി പരിശോധിച്ചിട്ടുണ്ടു്.

ഈ ഗ്രന്ഥത്തിലുള്ള ന്യൂനതകൾ ചൂണ്ടിക്കാണിക്കുന്നവരോടു് ഞങ്ങൾ കൃതജ്ഞരായിരിക്കും. കൂടുതൽ മെച്ചപ്പെട്ട രണ്ടാം പതിപ്പു് തയ്യാറാക്കുന്നതിനു് അവരുടെ വിമർശനങ്ങൾ ഉപകരിക്കുമല്ലോ.

കേരള ഭാഷാ ഇൻസ്റ്റിറ്റ്യൂട്ട്, എൻ. വി. കൃഷ്ണവാരിയർ
തിരുവനന്തപുരം–1
15 ഒക്ടോബർ, 1971

ഉള്ളടക്കം

അധ്യായം 5: ക്രമാനുസൃത സാമ്പിളനം

അധ്യായം 6: അനുപാത സമാശ്രയണാകലങ്ങൾ

1

ചില അടിസ്ഥാനതത്വങ്ങൾ

1 പ്രാരംഭം

സ്ഥിതിവിവരങ്ങളുടെ ആവശ്യവും പ്രാധാന്യവും അനുദിനം വർധിച്ചുവരികയാണ്. സാമ്പത്തികസാമൂഹ്യമണ്ഡലങ്ങളിൽ ഈ നൂററാണ്ടിലുണ്ടായ വിപ്ലവകരമായ പരിവർത്തനമത്രെ ഇതിനു പ്രധാന കാരണം. ഭരണകൂടങ്ങൾ മാത്രമാണ് ഒരു കാലത്ത് സ്ഥിതിവിവരങ്ങൾ ശേഖരിച്ചിരുന്നത്. അതും നികുതി ചുമത്തുവാനും പ്രതിരോധസജ്ജീകരണങ്ങൾക്കുള്ള വിഭവശേഷി നിർണയിക്കുവാനും ആവശ്യമായ വിവരങ്ങൾ മാത്രം. പക്ഷേ ഇന്നത്തെ സ്ഥിതി അതല്ല. രാഷ്ട്രത്തിന്റെ കർത്തവ്യങ്ങളെപ്പററിയുള്ള ധാരണകൾ തന്നെ മാറിയിരിക്കുന്നു. ജനതയുടെ എല്ലാ പ്രവർത്തനമണ്ഡലങ്ങളിലും ഇന്ന് ഭരണകൂടവും പങ്കാളിയാണ്. നിലവിലുള്ള സ്ഥിതിഗതികൾ മനസ്സിലാക്കാനും ഭാവിപരിപാടികൾ ആസൂത്രണം ചെയ്യാനും വൻതോതിലുള്ള സ്ഥിതിവിവരശേഖരണം ആവശ്യമായി തീർന്നിരിക്കുന്നു. മിക്കവാറും എല്ലാ പരിഷ്കൃത രാഷ്ട്രങ്ങളും ആസൂത്രണം ഒരു നയമായി സ്വീകരിച്ചതോടെ ശേഖരിക്കേണ്ട വിവരങ്ങളുടെ അളവും വ്യാപ്തിയും പതിന്മടങ്ങു വർധിച്ചു. ലഭ്യമായ വിഭവങ്ങൾ സമാഹരിച്ച് ഏററവും പ്രയോജനപ്രദമായി വിനിയോഗിക്കുക എന്നതാണല്ലോ ആസൂത്രണത്തിന്റെ ലക്ഷ്യം. ഇതിന് വിഭവങ്ങളെയും ആവശ്യങ്ങളെയും പററിയുള്ള വിവരങ്ങൾ ശേഖരിച്ചേ പററൂ. വിദ്യാഭ്യാസം, പൊതുജനാരോഗ്യം, ഭവനനിർമാണം തുടങ്ങിയവയെ പററിയുള്ള നയങ്ങൾക്ക് രൂപം കൊടുക്കുന്നത് ജനസംഖ്യ, സ്ത്രീപുരുഷഭേദം, പ്രായഭേദം, വിദ്യാഭ്യാസനിലവാരം, ജനസാന്ദ്രത തുടങ്ങിയവയെ പററിയുള്ള സ്ഥിതിവിവരങ്ങളെ അടിസ്ഥാനമാക്കിയാണ്. കാർഷികപദ്ധതികൾ ആസൂത്രണം ചെയ്യാൻ, ആവശ്യമുള്ള ഭക്ഷ്യപദാർഥങ്ങളുടെ അളവ്, നിലവിലുള്ള കൃഷിഭൂമിയുടെ ഉല്പാദനശേഷി, കൃഷിഭൂമിയായി മാററാവുന്ന തരിശുഭൂമിയുടെ അളവ്, കർഷകരുടെ എണ്ണം, അവരുടെ സാമ്പത്തികശേഷി, പരിശീലനം, കൈവശമുള്ള കൃഷിആയുധങ്ങൾ തുടങ്ങിയവയെപ്പററിയുള്ള സ്ഥിതിവിവരങ്ങൾ ശേഖരിക്കേണ്ടതുണ്ട്. ഇതു പോലെ തന്നെ വ്യവസായവികസന

പരിപാടികൾ ആവിഷ്ക്കരിക്കുന്നതെങ്കിൽ പ്രവർത്തിച്ചു കൊണ്ടിരിക്കുന്ന വ്യവസായങ്ങൾ, അവയുടെ ഉല്പന്നങ്ങളുടെ സ്വഭാവം, ഉല്പാദനശേഷി, കൂടുതലായി ആവശ്യമുള്ള വ്യാവസായികോല്പന്നങ്ങൾ, പുതുതായി ആരംഭിക്കണമെന്നു ഉദ്ദേശിക്കുന്ന വ്യവസായങ്ങൾക്ക് ആവശ്യമുള്ള അസംസ്കൃതപദാർത്ഥങ്ങൾ, യന്ത്രസാമഗ്രികൾ തുടങ്ങിയവ, അവയുടെ ലഭ്യത എന്നിവയെപ്പററിയെല്ലാമുള്ള സ്ഥിതിവിവരങ്ങൾ ആവശ്യമാണ്. തൊഴിൽനയം രൂപീകരിക്കാൻ ഓരോ തൊഴിലിലും ഏർപ്പെട്ടിട്ടുള്ള തൊഴിലാളികളുടെ എണ്ണം, സ്ത്രീപുരുഷഭേദം, വയസ്സ്, ജോലി ചെയ്യാനുള്ള കഴിവും പരിശീലനവും, ജീവിതനിലവാരം, തൊഴിലില്ലാത്തവരുടെ എണ്ണം, അവർക്കു ലഭിച്ചിട്ടുള്ള പരിശീലനത്തിൻെറ സ്വഭാവം തുടങ്ങിയവയെ പററിയുള്ള സ്ഥിതിവിവരങ്ങൾ ശേഖരിക്കണം. ഇതു പോലെ തന്നെ മററെല്ലാ ഭരണമണ്ഡലങ്ങളിലെയും നയരൂപീകരണത്തിന് അതിനോടു ബന്ധപ്പെട്ട സ്ഥിതിവിവരങ്ങൾ ശേഖരിച്ചേ തീരൂ. നടപ്പിലാക്കിയ പരിപാടികൾ എത്രമാത്രം വിജയപ്രദമായി എന്ന് അറിയാനും അവയിൽ വരുത്തേണ്ട ഭേദഗതികളെപ്പററി മനസ്സിലാക്കുവാനും സ്ഥിതിവിവരശേഖരണം ആവശ്യമാണ്. ഭരണകൂടത്തിൻെറ പ്രവർത്തനങ്ങളെ വിലയിരുത്തുകയും വിമർശിക്കുകയും ചെയ്യുന്നവർക്കും സ്ഥിതിവിവരങ്ങൾ ശേഖരിക്കേണ്ടതുണ്ട്. ചുരുക്കത്തിൽ ഒരു ആധുനികഭരണകൂടത്തിന് തൃപ്തികരമായി പ്രവർത്തിക്കാൻ വളരെയധികം സ്ഥിതിവിവരങ്ങൾ ആവശ്യമാണ്. മിക്കവാറും എല്ലാ രാജ്യങ്ങളിലും ഇതിനായി ഒരു പ്രത്യേകവകുപ്പും അതിൻെറ കീഴിൽ പ്രവർത്തിക്കുന്ന നിരവധി ഓഫീസുകളും ഉണ്ടായിരിക്കും. രാജ്യവ്യാപകമായ ഒരു പ്രവർത്തകശൃംഖല ഈ രംഗത്ത് അനവരതം പ്രവർത്തിച്ചുകൊണ്ടിരിക്കുകയും ചെയ്യും.

ഏററവുമധികം സ്ഥിതിവിവരങ്ങൾ ശേഖരിക്കാറുള്ളത് ഭരണകൂടങ്ങളാണെങ്കിലും സ്ഥിതിവിവരങ്ങൾ ശേഖരിക്കുകയും ഉപയോഗിക്കുകയും ചെയ്യുന്ന നിരവധി സ്ഥാപനങ്ങളുണ്ട്. അവ ഗവണ്മെൻറു പ്രസിദ്ധീകരണങ്ങളിൽ നിന്നു ലഭിക്കാവുന്ന സ്ഥിതിവിവരങ്ങൾ ഉപയോഗപ്പെടുത്തുകയും പോരാതെ വരുന്നതു് സ്വന്തമായി ശേഖരിക്കുകയും ചെയ്യുന്നു. വാണിജ്യവ്യവസായസ്ഥാപനങ്ങൾ, സാമൂഹ്യക്ഷേമസംഘടനകൾ തുടങ്ങിയവയാണ് ഈ വിഭാഗത്തിൽ പെടുന്നതു്. അവക്കും തങ്ങളുടെ പദ്ധതികൾ ആസൂത്രണം ചെയ്യാൻ സ്ഥിതിവിവരങ്ങൾ ആവശ്യമാണ്. വൻതോതിലുള്ള ഉല്പാദനവിതരണങ്ങൾ ആധുനികയുഗത്തിൻെറ ഒരു പ്രത്യേകതയാണല്ലോ. ഉപഭോക്താക്കളുടെ ആവശ്യങ്ങളെയും മനോഭാവത്തെയും അസംസ്കൃതസാധനങ്ങളുടെ ലഭ്യത, ഉല്പാദനച്ചെലവു് തുടങ്ങിയവയെയും മററും പററിയുള്ള സ്ഥിതിവിവരങ്ങളെ അടിസ്ഥാനമാക്കിയാണ് വൻകിടവാണിജ്യവ്യവസായസ്ഥാപനങ്ങൾ തങ്ങളുടെ പദ്ധതികൾക്കു് രൂപം നൽകുന്നതു്. ഉൽപാദനപ്രക്രിയ നടന്നുകൊണ്ടിരിക്കുമ്പോൾ തന്നെ, പല ഘട്ടങ്ങളിലായി അവയുടെ വിവിധ സ്വഭാവങ്ങളെ പററി സ്ഥിതിവിവരങ്ങൾ ശേഖരിക്കുകയും അങ്ങനെ ഉല്പന്നങ്ങളുടെ ഗുണം നിയന്ത്രിക്കുകയും ചെയ്യുന്നതും മിക്കവാറും എല്ലാ വ്യവസായശാലകളുടെയും പതിവാണ്. വിലനിലവാര

ത്തെപ്പറ്റിയും സാധനങ്ങളുടെ ലഭ്യത, ഉപഭോക്താക്കളുടെ ആവശ്യങ്ങൾ തുടങ്ങിയവയെപ്പററിയുമുള്ള സ്ഥിതിവിവരങ്ങൾ വാണിജ്യസ്ഥാപനങ്ങൾക്കു് ജീവൽപ്രധാനങ്ങളത്രെ. സാമൂഹ്യക്ഷേമസംഘടനകളും സ്വന്തം സേവനരംഗത്തെപ്പററി ചുരുങ്ങിയ തോതിലെങ്കിലും സ്ഥിതിവിവരങ്ങൾ ശേഖരിക്കാറുണ്ടു്. സേവനം അർപ്പിക്കുന്നവരെ പററി മനസ്സിലാക്കാനും സംഘടനയ്ക്കു് എന്തു ചെയ്യാൻ കഴിയുമെന്നു് നിർണയിക്കാനും ഇതുപകരിക്കും.

യുനസ്കോ പോലുള്ള അന്തർദേശീയസ്ഥാപനങ്ങളും, സാമൂഹ്യ- സാമ്പത്തിക പ്രശ്നങ്ങളെപ്പററി ഗവേഷണം നടത്തുന്ന വ്യക്തികൾ, സ്ഥാപനങ്ങൾ തുടങ്ങിയവയും മറ്റുമാണു് സ്ഥിതിവിവരശേഖരണത്തിൽ താല്പര്യമുള്ള മററു ഘടകങ്ങൾ. ഇവ താരതമ്യേന കുറച്ചു വിവരങ്ങളേ ശേഖരിക്കാറുള്ളു എങ്കിലും ശേഖരിക്കുന്ന വിവരങ്ങൾ സങ്കീർണസ്വഭാവമുള്ളവയും ശേഖരിക്കാനവലംബിക്കുന്ന മാർഗങ്ങൾ ഏററവുമധികം പരിഷ്കൃതങ്ങളും ആയിരിക്കുമെന്നുള്ളതു് പ്രത്യേകം ശ്രദ്ധേയമാണു്.

2. സ്ഥിതിവിവരശേഖരണം

സാധാരണ ശേഖരിക്കാറുള്ള സ്ഥിതിവിവരങ്ങൾ പൊതുവെ പറഞ്ഞാൽ രണ്ടു വിഭാഗത്തിൽ പെടുന്നവയാണു്. ഒന്നു്, നിലവിലുള്ളവ; രണ്ടു്, പരീക്ഷണങ്ങൾ നടത്തി ശേഖരിക്കുന്നവ. ജനസംഖ്യ, വിവിധ തൊഴിലുകളിൽ ഏർപ്പെട്ടിട്ടുള്ളവരുടെ എണ്ണം, ഒരു സാധനത്തിന്റെ പല വർഷങ്ങളിലെ വിലകൾ, ഒരു രാജ്യത്തിന്റെ ഭൂപ്രകൃതിയെപ്പററിയുള്ള വിവരങ്ങൾ, ഓരോ വർഷത്തെയും കയറ്റുമതിയും ഇറക്കുമതിയും സംബന്ധിച്ചുള്ള വിവരങ്ങൾ തുടങ്ങിയവയെല്ലാം ആദ്യത്തെ വിഭാഗത്തിൽപ്പെടും. സ്ഥിതിവിവരാന്വേഷകനു് ഈ വിവരങ്ങളെല്ലാം ശേഖരിച്ചെടുക്കുക എന്ന ജോലി മാത്രമേ ചെയ്യാനുള്ളു. പക്ഷേ, രണ്ടാം വിഭാഗത്തിൽപ്പെട്ട വിവരങ്ങൾ ശേഖരിക്കാൻ അന്വേഷകൻ സങ്കീർണങ്ങളായ പരീക്ഷണങ്ങൾ നടത്തേണ്ടതായി വരും. ഉദാഹരണമായി പല തരം വിത്തുകളുടെ ഉല്പാദനശേഷിയെപ്പററിയുള്ള വിവരങ്ങളാണു് ആവശ്യമുള്ളതെന്നിരിക്കട്ടെ, അത്തരത്തിൽ പെട്ട എല്ലാ വിത്തുകളും കൃഷി ചെയ്തിട്ടുള്ള സ്ഥലങ്ങൾ കണ്ടുപിടിച്ചു് അവിടെ നിന്നെല്ലാം വിവരം ശേഖരിക്കുക എളുപ്പമല്ല. പലപ്പോഴും എല്ലാ വിത്തുകളും കൃഷിക്കാരുടെ ഇടയിൽ പ്രചാരത്തിൽ വന്നിരിക്കയില്ല. അവ പ്രചാരത്തിലുണ്ടെങ്കിൽ തന്നെ, പല തരത്തിലുള്ള കൃഷിഭൂമികളിലും പല തരം വളങ്ങൾ ചേർത്തും പല കൃഷിസമ്പ്രദായങ്ങളുപയോഗിച്ചുമായിരിക്കും അവ കൃഷി ചെയ്തു വരുന്നതു്. വിത്തുകളുടെ ഉല്പാദനശേഷി താരതമ്യപ്പെടുത്താൻ ഇമ്മാതിരി കൃഷിസ്ഥലങ്ങളിൽ നിന്നു ലഭിക്കുന്ന വിവരങ്ങൾ ഒട്ടും ഉപകരിക്കുകയില്ല, ഒരേതരം കൃഷിഭൂമിയിൽ ഒരേ വളങ്ങൾ ചേർത്തു്, ഒരുപോലെ കൃഷിചെയ്തു നോക്കിയിട്ടു വേണം ഏതു വിത്തിനാണു് കൂടുതൽ ഉല്പാദനശേഷിയെന്നു നിശ്ചയിക്കാൻ. അങ്ങനെ ഈ ആവശ്യത്തിനുള്ള വിവരങ്ങൾ ശേഖരിക്കാൻ അന്വേഷകൻ തന്നെ പരീക്ഷണാടിസ്ഥാനത്തിൽ കൃഷി നടത്തി വിവരങ്ങൾ ശേഖരിക്കണം. അതായതു് പരീക്ഷണങ്ങൾ നടത്തുക, വിവരങ്ങൾ ശേഖരി

ക്കുക എന്നീ രണ്ടു ജോലികളും സ്ഥിതിവിവരാന്വേഷകൻെറ കർത്തവ്യപരിധിയിൽ ഉൾപ്പെടും. ഇങ്ങനെ ശേഖരിക്കുന്ന വിവരങ്ങളാണ് രണ്ടാം വിഭാഗത്തിൽ പെടുക.

ഇതിൽ ആദ്യത്തെ വിഭാഗത്തിൽ പെട്ട വിവരങ്ങൾ ശേഖരിക്കുവാനും അപഗ്രഥിക്കുവാനുമുള്ള പദ്ധതികളാണ് ഈ ഗ്രന്ഥത്തിൽ പ്രധാനമായും ചർച്ച ചെയ്യുന്നതു്. രണ്ടാമത്തെ വിഭാഗത്തിൽ പെട്ട വിവരങ്ങൾ ശേഖരിക്കാൻ ഇതിൽ നിന്നു് തികച്ചും വ്യത്യസ്തങ്ങളായ മാർഗങ്ങളാണ് അവലംബിക്കേണ്ടതു്. "പരീക്ഷണങ്ങളുടെ ഡിസൈനൽ" എന്ന സാംഖ്യികശാസ്ത്രശാഖ ഈ വിഷയം കൈകാര്യം ചെയ്യുന്നു. ചുരുങ്ങിയ കാലം കൊണ്ടു് അതിബൃഹത്തും അത്യന്തം പ്രയോജനപ്രദവുമായ ഒരു ശാസ്ത്രവിഭാഗമായി അതു് വളർന്നു കഴിഞ്ഞിട്ടുണ്ടു്.

നിലവിലുള്ള സ്ഥിതിഗതികളെപ്പററിയുള്ള വിവരങ്ങൾ എങ്ങനെയാണ് ശേഖരിക്കുക എന്നതാണല്ലോ ഇവിടുത്തെ ചർച്ചാവിഷയം. നിലവിലുള്ള സ്ഥിതിഗതികൾ തന്നെ പൊതുവെ രണ്ടു തരത്തിൽ പെട്ടവയാണ്. ചിലതെല്ലാം കാലം മാറുന്നതനുസരിച്ചു് മാറും. ഇങ്ങനെയുള്ളവയാണ് ആദ്യത്തെ വിഭാഗത്തിൽ പെടുന്നതു്. മററു ചിലവ കാലം മാറുന്നതു കൊണ്ടു് വലിയ മാററങ്ങളൊന്നും സംഭവിക്കാത്തവയാണ്. അവ രണ്ടാമത്തെ വിഭാഗത്തിലും പെടും.

ജനസംഖ്യ, വിലനിലവാരം, രോഗബാധിതരുടെ എണ്ണം, തൊഴിലില്ലാത്തവരുടെ സംഖ്യ തുടങ്ങിയവ ആദ്യത്തെ വിഭാഗത്തിൽ പെടും. ഇവയെല്ലാം കാലം മാറുന്നതനുസരിച്ചു് മാറുന്നവയാണല്ലോ. ഒരു നിശ്ചിത സമയത്തോടോ കാലഘട്ടത്തോടോ ബന്ധപ്പെടുത്തിയേ ഇവ പരിഗണിക്കാനാവൂ. പക്ഷേ, ഒരു രാജ്യത്തെ മലകളുടെ എണ്ണം, നദികളുടെ ദൈർഘ്യം തുടങ്ങിയവയെപ്പററിയുള്ള വിവരങ്ങൾ വളരെ കാലത്തേയ്ക്കു് മാററംകൂടാതെ ഇരിക്കുന്നവയാകയാൽ അവയെ രണ്ടാം വിഭാഗത്തിൽ പെടുത്താം. ഏതെങ്കിലും ഒരു പ്രത്യേകകാലഘട്ടത്തോടു് അവയെ ബന്ധപ്പെടുത്തേണ്ടതില്ല. അമ്മാതിരി വിവരങ്ങൾ ശേഖരിക്കാൻ സങ്കീർണങ്ങളായ സ്ഥിതിവിവരശേഖരണരീതികളൊന്നും ഉപയോഗിക്കണമെന്നില്ല. ഒരിക്കൽ ശേഖരിച്ചാൽ പിന്നീടു് അടുത്ത കാലത്തെങ്ങും ഇമ്മാതിരി വിവരങ്ങൾ ശേഖരിക്കേണ്ടതായി വരികയുമില്ല.

ഇതിൽ ആദ്യത്തെ വിഭാഗത്തിൽപ്പെട്ട വിവരങ്ങൾ അത്ര എളുപ്പത്തിൽ ശേഖരിക്കാവുന്നവയല്ല. വളരെ ചുരുങ്ങിയ ഒരു കാലയളവിനുള്ളിൽ എല്ലാ വിവരങ്ങളും ശേഖരിച്ചു കഴിയണം എന്നതാണ് ഈ ബുദ്ധിമുട്ടിൻെറ പ്രധാന കാരണം. 1970 ആരംഭത്തിൽ കേരളത്തിൽ തൊഴിലില്ലാത്ത എത്ര യുവാക്കന്മാരുണ്ടെന്നു് അറിയേണ്ടിയിരിക്കുന്നു എന്നിരിക്കട്ടെ. 1970 ജനുവരി ഒന്നാം തീയതിക്കു മുൻപോ പിൻപോ ശേഖരിച്ച വിവരങ്ങൾ ഈ ആവശ്യത്തിനുപകരിക്കുകയില്ല. ആ ഒരു ദിവസം കൊണ്ടു് കേരളത്തിലെ എല്ലാ ഭാഗത്തു നിന്നും ആവശ്യമായ വിവരങ്ങൾ ശേഖരിക്കുക എന്നതു് എത്രയോ ബുദ്ധിമുട്ടുള്ള ഒരു കാര്യമാണ്. സ്ഥിതിവിവരശേഖരണത്തിനു് ചെലവഴിക്കാവുന്ന പണം പരിമിതവും

കൂടിയാണെങ്കിൽ (മിക്കവാറും അങ്ങനെ ആയിരിക്കാനാണ് സാധ്യത) സ്ഥിതി എത്ര വിഷമകരമാണ്. ഇമ്മാതിരി സന്ദർഭങ്ങളിലാണ് ശാസ്ത്രീയമായ സാമ്പിളനം അന്വേഷകൻെറ സഹായത്തിനെത്തുന്നത്. നിശ്ചിതസമയത്തിനുള്ളിൽ കുറഞ്ഞ ചെലവിൽ ആവശ്യമായ സൂക്ഷ്മതയോടു കൂടി വിവരങ്ങൾ ശേഖരിക്കുവാനുള്ള ഒരേ ഒരു മാർഗം സാമ്പിളനം മാത്രമാണ്. വിവിധസാമ്പിളനരീതികളെപ്പറ്റിയുള്ള വിശദമായ ഒരു ചർച്ചയാണ് ഈ ഗ്രന്ഥത്തിൻെറ പ്രതിപാദ്യവിഷയം.

സ്ഥിതിവിവരങ്ങൾ ശേഖരിക്കുക എന്നതു് പ്രഥമദൃഷ്ട്യാ തോന്നുന്നത്ര ലളിതമായ ഒരു പ്രക്രിയയല്ല. സമഷ്ടി ചെറുതാണെങ്കിൽ എളുപ്പത്തിൽ വിവരങ്ങൾ ശേഖരിക്കാൻ കഴിയും. പക്ഷേ സമഷ്ടി വലുതാകുന്തോറും സങ്കീർണങ്ങളായ അനവധി പ്രശ്നങ്ങൾക്കു് പരിഹാരം കാണേണ്ടതായി വരും. സംഭവിക്കാവുന്ന പിശകുകളുടെ എണ്ണം വർധിക്കും. നിശ്ചിത സമയത്തിനുള്ളിൽ വിവരം ശേഖരിച്ചു തീരണമെന്നുള്ളതുകൊണ്ടു് വിപുലമായ ആസൂത്രണം ആവശ്യമായി വരും. സമഷ്ടിയിലെ എല്ലാ വ്യക്തികളിൽ നിന്നും വിവരം ശേഖരിക്കണമോ അതോ അതിലെ ഏതാനും വ്യക്തികളിൽ നിന്നു് ശേഖരിച്ചാൽ മതിയോ എന്നു തീരുമാനിക്കണം. അന്വേഷകരെ ശ്രദ്ധാപൂർവം തിരഞ്ഞെടുക്കുകയും പരിശീലിപ്പിക്കുകയും വേണം. അവരുടെ പ്രവർത്തനം വേണ്ട വിധം പരിശോധിക്കുകയും നിയന്ത്രിക്കുകയും വേണം. അന്വേഷകർ കടന്നു ചെല്ലേണ്ടതു് ദുർഗമങ്ങളായ വനാന്തരങ്ങളിലും ഉത്തുംഗങ്ങളായ പർവതശിഖരങ്ങളിലും ആണെന്നു വരും. വിവരങ്ങൾ നല്കേണ്ടവരിൽ പലരെയും കണ്ടെത്താൻ കഴിഞ്ഞില്ലെന്നും കണ്ടെത്തിയവരിൽ തന്നെ ചിലരെല്ലാം വിവരങ്ങൾ നല്കാൻ കഴിവില്ലാത്തവരോ വിസമ്മതിക്കുന്നവരോ തെററായ വിവരങ്ങൾ നല്കുന്നവരോ ആയിരുന്നെന്നും വരാവുന്നതാണ്. ഈ പ്രശ്നങ്ങൾക്കെല്ലാം പരിഹാരം കാണണം. വിവരങ്ങൾ ശേഖരിച്ചു കഴിഞ്ഞാൽ അടുത്ത ഘട്ടം അവ അപഗ്രഥനം ചെയ്തു് നിഗമനങ്ങളിലെത്തുകയാണ്. ഇതും താത്വികവും പ്രായോഗികവുമായ അനവധി പ്രശ്നങ്ങൾ സൃഷ്ടിച്ചേക്കാവുന്ന ഒന്നാണ്. അവയും പരിഹരിച്ചേ കഴിയൂ. ഇതു പോലെ തന്നെ പ്രധാനമാണ് സ്ഥിതിവിവരശേഖരണം ഉൾക്കൊള്ളുന്ന സാമ്പത്തികബാധ്യത ഏകദേശമായെങ്കിലും പ്രാരംഭത്തിൽ തന്നെ നിർണയിക്കുക എന്നതു്. സാമ്പിളനശാസ്ത്രത്തിൻെറ പരിധിയിൽ പെടുന്ന വിഷയങ്ങളാണ് ഇവയെല്ലാം. പക്ഷേ ഈ ശാസ്ത്രശാഖയുടെ പ്രധാന ചർച്ചാവിഷയങ്ങൾ ഇവയാണെന്നു പറഞ്ഞുകൂടാ. സമഷ്ടിയുടെ ഒരു ഭാഗത്തു നിന്നു മാത്രമേ വിവരം ശേഖരിക്കുന്നുള്ളു എങ്കിൽ ആ ഭാഗം എങ്ങനെ തിരഞ്ഞെടുക്കണം, അതിൽ നിന്നു സമഷ്ടിയെപ്പററി എന്തെല്ലാം മനസ്സിലാക്കാൻ കഴിയും, ഈ നിഗമനങ്ങളുടെ വിശ്വാസ്യത എത്ര മാത്രമാണ് എന്നു തുടങ്ങിയ വിഷയങ്ങളാണ് സാമ്പിളനശാസ്ത്രത്തിലെ പ്രധാന പരിഗണനാവിഷയങ്ങൾ. സാമ്പിളനത്തോടു് ബന്ധപ്പെട്ട താത്വികവും പ്രായോഗികവുമായ പ്രശ്നങ്ങളെപ്പററി, കഴിഞ്ഞ അര നൂററാണ്ടിനുള്ളിൽ പ്രഗത്ഭരായ അനവധി ശാസ്ത്രജ്ഞന്മാർ ഗവേഷണം നടത്തുകയുണ്ടായി. ഈ

ശാസ്ത്രശാഖയ്ക്ക് അത്ഭുതാവഹമായ വളർച്ചയാണ് ഈ ചുരുങ്ങിയ കാലം കൊണ്ട് ഉണ്ടായിട്ടുള്ളത്. സാമ്പിളനത്തിന്റെ പ്രാധാന്യവും പ്രായോഗികസൗകര്യങ്ങളും ഇന്നു പൊതുവെ അംഗീകരിക്കപ്പെട്ടു കഴിഞ്ഞു. എല്ലാ പരിഷ്കൃതജനപദങ്ങളിലും സാമ്പിളനരീതിയിൽ സ്ഥിതിവിവരങ്ങൾ ശേഖരിക്കുന്ന സമ്പ്രദായം സർവസാധാരണമായിത്തീരുകയും ചെയ്തു.

3. വ്യക്തി, സമഷ്ടി, അഭിലക്ഷണം

ഒരു സ്ഥിതിവിവരശേഖരണത്തിൽ ആരെ, അല്ലെങ്കിൽ, എന്തിനെപ്പററിയുള്ള വിവരങ്ങളാണോ ശേഖരിക്കുന്നതു്, അവയ്ക്കു് ആ സ്ഥിതിവിവരാന്വേഷണത്തിലെ പ്രാഥമികവ്യക്തികൾ എന്നു പേർ പറയും. പ്രാഥമികവ്യക്തികൾ മനുഷ്യരോ കുടുംബങ്ങളോ സ്ഥാപനങ്ങളോ കൃഷിഭൂമികളോ വൃക്ഷങ്ങളോ രാജ്യങ്ങളോ എന്തു വേണമെങ്കിലുമാവാം. പ്രാഥമികവ്യക്തികളെ സംശയാതീതമായി നിർവചിക്കേണ്ടതു് ഏതൊരു സ്ഥിതിവിവരാന്വേഷണത്തിന്റെയും വിജയത്തിനു് ആവശ്യമാണു്. കേരളത്തിലെ ഇടത്തരക്കാരുടെ കുടുംബച്ചെലവുകളെപ്പററിയുള്ള ഒരന്വേഷണത്തിൽ പ്രാഥമികവ്യക്തി കേരളത്തിലെ ഒരു ഇടത്തരം കുടുംബമാണു്. പക്ഷേ "ഇടത്തരം കുടുംബം" എന്നു വെറുതെ പറഞ്ഞതു കൊണ്ടായില്ല. അതു് വ്യക്തമായി നിർവചിക്കണം. അല്ലെങ്കിൽ അന്വേഷകനു് ഒരിടത്തരം കുടുംബം എങ്ങനെയാണു് കണ്ടെത്താൻ കഴിയുക ? 1968-ൽ അഞ്ഞൂറു രൂപയ്ക്കും ആയിരം രൂപയ്ക്കും മധ്യേ വാർഷികവരുമാനമുള്ള കുടുംബങ്ങൾ ഇടത്തരം കുടുംബങ്ങളായി പരിഗണിക്കപ്പെടും എന്നൊരു നിർവചനം കൂടിയുണ്ടെങ്കിൽ സ്ഥിതിഗതി കുറെയെല്ലാം മെച്ചപ്പെട്ടു. അന്വേഷകനു് സംശയാതീതമായി പ്രാഥമികവ്യക്തികളെ തിരിച്ചറിയുവാൻ കഴിയുക എന്നതായിരിക്കണം നിർവചനത്തിന്റെ ലക്ഷ്യം. ഈ നിർവചനത്തിൽ സംഭവിക്കാവുന്ന പോരായ്മകൾ അന്വേഷണത്തിന്റെ ലക്ഷ്യത്തെ ഒരു പക്ഷേ പരാജയപ്പെടുത്തിയേക്കാം എന്ന വസ്തുത ഓർമയിരിക്കണം. വിവരങ്ങൾ ആവശ്യമുള്ള വ്യക്തികളും സ്ഥാപനങ്ങളുമായി വിശദമായി ചർച്ച ചെയ്തിട്ടു വേണം ഈ നിർവചനത്തിനു് രൂപം കൊടുക്കുവാൻ. ഒരിക്കൽ നൽകിയ നിർവചനം അന്വേഷണം പൂർത്തിയാവുന്നതിനു് മുൻപു് യാതൊരു കാരണവശാലും വ്യത്യാസപ്പെടുത്തരുതെന്നുള്ളതും പ്രത്യേകം ഓർമിക്കേണ്ട ഒരു കാര്യമാണു്. അങ്ങനെ വരുത്തുന്ന വ്യത്യാസങ്ങൾ ആദ്യം ശേഖരിച്ച വിവരങ്ങളും പിന്നീടു ശേഖരിക്കുന്ന വിവരങ്ങളും വ്യത്യസ്തസ്വഭാവങ്ങളോടു കൂടിയവയും വിഭിന്ന സമഷ്ടികളോടു ബന്ധപ്പെട്ടവയും ആയിത്തീരാൻ ഇടയാക്കിയെന്നു വരാം.

പലപ്പോഴും അന്വേഷകനു് വിവരങ്ങൾ നൽകുന്നതു് പ്രാഥമികവ്യക്തിയായിരിക്കുകയില്ല. അന്വേഷകൻ ആരിൽ നിന്നു വിവരം ശേഖരിക്കുന്നുവോ ആ വ്യക്തിക്കു് 'വിവരം നൽകുന്ന വ്യക്തി' എന്നു പേർ പറയുന്നു. പ്രാഥമികവ്യക്തി കുടുംബമാണെങ്കിൽ വിവരം നൽകുന്ന വ്യക്തി കുടുംബനാഥനാവാം. കുരുമുളക്ചെടികളെപ്പററിയുള്ള ഒരന്വേഷണത്തിൽ പ്രാഥമികവ്യക്തി കുരുമുളക് ചെടിയാണു്. പക്ഷേ വിവരം നൽകുന്ന വ്യക്തി അതിന്റെ ഉടമസ്ഥ

നാവാം. അന്വേഷകൻ കുരുമുളക്‌ചെടിയിൽ നിന്ന് നേരിട്ട് വിവരങ്ങൾ ശേഖരിക്കുകയാണെങ്കിൽ പ്രാഥമികവ്യക്തിയും വിവരം നൽകുന്ന വ്യക്തിയും ഒന്നുതന്നെയാണ്. കഴിവതും അവ രണ്ടും ഒന്നായിരിക്കാൻ ശ്രദ്ധിക്കേണ്ടതാണ്. പക്ഷേ പലപ്പോഴും അതു സാധിച്ചെന്നു വരികയില്ല. എങ്കിലും വിവരം നൽകുന്ന വ്യക്തിക്ക് പ്രാഥമികവ്യക്തിയെപ്പററി പൂർണ്ണമായും അറിയാമായിരിക്കണം എന്ന കാര്യം പ്രത്യേകം ശ്രദ്ധിക്കേണ്ടതാണ്.

അപഗ്രഥനഘട്ടത്തിലെ വ്യക്തി പ്രാഥമികവ്യക്തിയോ പ്രാഥമികവ്യക്തികളുടെ സമുച്ചയമോ ആകാവുന്നതാണ്. എത്തിച്ചേരേണ്ട നിഗമനങ്ങളുടെ സ്വഭാവവും അപഗ്രഥനസൗകര്യവുമാണ് ഈ സന്ദർഭത്തിൽ പ്രത്യേകം പരിഗണിക്കാനുള്ളതു്.

ഒരന്വേഷണത്തിൽ പരിഗണിക്കുന്ന പ്രാഥമികവ്യക്തികളുടെ സമ്പൂർണ്ണസമുച്ചയത്തിന്, ആ അന്വേഷണത്തോടു ബന്ധപ്പെട്ട സമഷ്ടി എന്നു പേർ പറയും. സമഷ്ടിയുടെ നിർവചനം പൂർണമാകണമെങ്കിൽ അതിന്റെ ഉള്ളടക്കം, ഉപയോഗിച്ചിരിക്കുന്ന മാത്ര, അതിന്റെ വ്യാപ്തി, സമയം തുടങ്ങിയവയെല്ലാം വ്യക്തമായി നിർദേശിച്ചിരിക്കണം. 1969 ൽ കേരളത്തിലുള്ള ഇടത്തരം കുടുംബങ്ങൾ ഒരു സമഷ്ടിയാണ്. 1970 ജൂൺ 1-ാം തീയതി ഇൻഡ്യയിലെ അംഗീകൃത കോളേജുകളിൽ പഠനം നടത്തുന്ന വിദ്യാർഥികൾ മറെറാരു സമഷ്ടിയാണ്. സമഷ്ടിയെ നിർവചിക്കുമ്പോൾ ഓർമയിരിക്കേണ്ടതു്, ഒരു പ്രാഥമികവ്യക്തി സമഷ്ടിയിലെ അംഗമാണോ അല്ലയോ എന്ന് സംശയം വരാൻ ഇടയാകരുതു് എന്നതാണ്.

സമഷ്ടിയിലെ അംഗങ്ങളുടെ എണ്ണം അനന്തമാണെങ്കിൽ അതു് അനന്തസമഷ്ടി എന്നും അല്ലെങ്കിൽ Finite population സാന്തസമഷ്ടി എന്നും വ്യവഹരിക്കപ്പെടും. പ്രധാനമായും സാന്തസമഷ്ടികളാണ് സാമ്പിളനശാസ്ത്രത്തിൽ പരിഗണിക്കപ്പെടാറുള്ളതു്. സാന്തസമഷ്ടിയിൽ നിന്നുള്ള സാമ്പിളനത്തിൽ ഉപയോഗിക്കാൻ വേണ്ടി വ്യുൽപ്പാദിപ്പിക്കുന്ന സൂത്രങ്ങളുടെ സീമകളായിരിക്കും അനന്തസമഷ്ടിയിലെ സംഗതങ്ങളായ സൂത്രങ്ങൾ എന്നതാണ് ഇതിനു കാരണം.

ഓരോ പ്രാഥമികവ്യക്തിയുടെയും ഒന്നോ അതിലധികമോ അഭിലക്ഷണങ്ങളെപ്പററിയുള്ള വിവരങ്ങളാണ് ശേഖരിക്കുന്നതു്. കോളേജ് വിദ്യാർത്ഥികളെപ്പററിയുള്ള ഒരു അന്വേഷണത്തിൽ ഓരോ വിദ്യാർഥിയുടെയും വയസ്സ്, സ്ത്രീപുരുഷഭേദം, തൂക്കം, ഉയരം എന്നീ വിവരങ്ങൾ ശേഖരിക്കുന്നു എന്നിരിക്കട്ടെ ഇവിടെ നാല് അഭിലക്ഷണങ്ങളാണ് പരിഗണിക്കപ്പെടുന്നതു്. ഇവയിൽ വയസ്സ്, ഉയരം, തൂക്കം എന്നിവ ചരരൂപത്തിലുള്ളവയും സ്ത്രീപുരുഷഭേദം ഗുണരൂപത്തിലുള്ളതുമായ അഭിലക്ഷണങ്ങളാണ്. ചരരൂപത്തിലുള്ള അഭിലക്ഷണത്തിന്റെ മാനം സംഖ്യാരൂപത്തിലായിരിക്കുമല്ലോ. പക്ഷേ ഗുണരൂപത്തിലുള്ളതിന്റേതു് അങ്ങനെയല്ല. എങ്കിലും ഏതെങ്കിലും ഒരു സങ്കേതം അനുസരിച്ച് അവയേയും സംഖ്യാരൂപത്തിൽ നിർണയിക്കാൻ കഴിയും. സ്ത്രീത്വത്തെ "0" കൊണ്ടും, പുരുഷത്വത്തെ "1" കൊണ്ടും സൂചിപ്പിക്കുന്നു എന്നു സങ്കല്പിച്ചാൽ സ്ത്രീപു

രുഷഭേദം എന്ന അഭിലക്ഷണത്തെ സംബന്ധിച്ച വിവരങ്ങളും സംഖ്യാരൂപത്തിൽ ആയിത്തീരും. അങ്ങനെ പരിഗണിക്കപ്പെടുന്ന എല്ലാ അഭിലക്ഷണങ്ങളെയും സംബന്ധിക്കുന്ന വിവരങ്ങളും സംഖ്യാരൂപത്തിൽ രേഖപ്പെടുത്താൻ കഴിയുമെന്നു വന്നു ചേരുന്നു. ഇതു് ഗണിതശാസ്ത്രപരമായ പല സൗകര്യങ്ങളും ഉണ്ടാക്കിത്തരുന്ന ഒന്നാണ്. നാം ഒരു അഭിലക്ഷണമേ പരിഗണിക്കുന്നുള്ളു എന്നു വിചാരിക്കുക. അന്വേഷകൻ ചെയ്യുന്നതു് സമഷ്ടിയിലെ ഓരോ പ്രാഥമിക വ്യക്തിയിൽ നിന്നും ആ അഭിലക്ഷണത്തെ സംബന്ധിച്ച വിവരങ്ങൾ ശേഖരിക്കുകയാണ്. അങ്ങനെ ശേഖരിക്കുന്ന വിവരങ്ങൾ സംഖ്യാരൂപത്തിലാണെങ്കിൽ അവയെ ആ സമഷ്ടിയിലെ അംഗങ്ങളോടു ബന്ധപ്പെടുത്തി നിർവചിച്ചിരിക്കുന്ന ഒരു ചരത്തിൻെറ മൂല്യങ്ങളായി വിചാരിക്കാവുന്നതാണ്. നാം പരിഗണിക്കുന്ന അഭിലക്ഷണത്തിൻെറ മൂല്യം x എന്ന ചരം സൂചിപ്പിക്കുന്നു എന്നു സങ്കൽപ്പിച്ചാൽ, അന്വേഷകൻ രേഖപ്പെടുത്തുന്നതു് ഓരോ വ്യക്തിയോടും ബന്ധപ്പെട്ട x ൻെറ മൂല്യമാണ് എന്നു സാരം. സമഷ്ടിയിൽ എത്ര വ്യക്തികളുണ്ടോ അത്രയും മൂല്യങ്ങൾ x നും ഉണ്ടായിരിക്കുമെന്നു് ഇതിൽ നിന്നു സിദ്ധിക്കുന്നു. ഈ മൂല്യങ്ങളുടെ സമുച്ചയത്തിനു് നിർദിഷ്ട സമഷ്ടിയോടു ബന്ധപ്പെട്ട സാംഖ്യികസമഷ്ടി എന്നു പറയും. അപഗ്രഥനഘട്ടത്തിൽ നാം പരിഗണിക്കുന്നതു് ഈ സാംഖ്യികസമഷ്ടി മാത്രമാണ്. ഈ സമഷ്ടി സംഖ്യകളുടെ സമുച്ചയമായതു കൊണ്ടു് അതു ഗണിതശാസ്ത്രപരമായി കൈകാര്യം ചെയ്യാൻ കഴിയും. അപഗ്രഥനത്തെ സംബന്ധിച്ചിടത്തോളം അത്യന്തം ഉപകാരപ്രദമായ ഒരു വസ്തുതയാണ് ഇതു്. ഒന്നിലധികം അഭിലക്ഷണങ്ങൾ പരിഗണിക്കുമ്പോഴും സ്ഥിതിഗതികൾ ഇതിൽ നിന്നു വ്യത്യസ്തമല്ല. ഓരോ വ്യക്തിയിൽ നിന്നും രണ്ടു് അഭിലക്ഷണങ്ങളെ പററിയുള്ള വിവരങ്ങളാണ് ശേഖരിക്കുന്നതു് എന്നിരിക്കട്ടെ. ഉദാഹരണമായി ഒരു നഗരത്തിലെ പ്രായപൂർത്തിയായ പുരുഷന്മാരാണ് സമഷ്ടിയെന്നും അവരുടെ വയസ്സും തൂക്കവുമാണ് പരിഗണിക്കപ്പെടുന്ന അഭിലക്ഷണങ്ങളെന്നും കരുതുക. വയസ്സിനെ x_1 എന്ന ചരവും തൂക്കത്തെ x_2 എന്ന ചരവും സൂചിപ്പിക്കുന്നതായി സങ്കല്പിച്ചാൽ, ഓരോ വ്യക്തിയിൽ നിന്നും അന്വേഷിക്കുന്നതു് x_1, x_2 എന്നീ ചരങ്ങളുടെ മൂല്യങ്ങളാണ് എന്നു വരും. $x = (x_1, x_2)$ എന്നിരിക്കട്ടെ. ഇവിടെ x ഒരു ക്രമബദ്ധസമുച്ചയമാണ്: ഗണിതശാസ്ത്രത്തിന്റെ ഭാഷയിൽ പറഞ്ഞാൽ അതു് ഒരു സദിശമാണ്. രണ്ടു് അംഗങ്ങളുള്ള ഒരു സദിശം. ഓരോ വ്യക്തിയോടും ബന്ധപ്പെട്ട x ൻെറ മൂല്യം ഓരോന്നായിരിക്കുമല്ലോ. അങ്ങനെ x ഒരു സദിശചരമാണെന്നു സിദ്ധിക്കുന്നു. അഭിലക്ഷണങ്ങളുടെ എണ്ണത്തോളം അംഗങ്ങളാണ് ആ സദിശചരത്തിനു് ഉണ്ടായിരിക്കുക. അങ്ങനെ ഒന്നിലധികം അഭിലക്ഷണങ്ങളെ പരിഗണിക്കുമ്പോഴും സ്ഥിതി വ്യത്യസ്തമല്ല. അപ്പോഴും നിർദിഷ്ടസമഷ്ടിയോടു ബന്ധപ്പെട്ട ഒരു സാംഖ്യികസമഷ്ടി ഉണ്ടായിരിക്കും. അതിൻെറ അംഗങ്ങൾ അഭിലക്ഷണങ്ങളുടെ എണ്ണത്തോളം അംഗങ്ങളുള്ള സദിശങ്ങളായിരിക്കുമെന്നു മാത്രം. അഥവാ അമ്മാതിരി ഒരു സദിശചരത്തിൻെറ മൂല്യങ്ങളായിരിക്കും സാംഖ്യികസമഷ്ടിയുടെ അംഗങ്ങൾ.

ഒരു സമഷ്ടിയെപ്പററി പഠനം നടത്തുമ്പോൾ പ്രായോഗികവൈഷമ്യങ്ങൾ മൂലം അതിൻെറ ചില ഭാഗങ്ങൾ പഠനവിധേയമാക്കാതെ വിട്ടുകളയാറുണ്ട്. പക്ഷെ, പഠനത്തിൽ നിന്നു ലഭിക്കുന്ന വിവരങ്ങളെ സമഷ്ടിയുടെ മുഴുവൻ സ്വഭാവമായാണ് അംഗീകരിക്കാറുള്ളതു്. ഇതിൽ പഠനവിധേയമായ സമഷ്ടിക്ക് പഠിതസമഷ്ടി എന്നും പരിഗണനാവിധേയമായ സമഷ്ടിക്ക് ലക്ഷ്യസമഷ്ടി എന്നു പേർ പറയും. കേരളത്തിലെ ഹൈസ്കൂൾവിദ്യാർഥികളുടെ 1970 സ്കൂൾ വർഷത്തെ ശരാശരി ചെലവ് നിർണയിക്കുകയാണ് അന്വേഷണലക്ഷ്യമെന്നിരിക്കട്ടെ. 1970 സ്കൂൾവർഷത്തിൽ കേരളത്തിലെ ഹൈസ്കൂളുകളിൽ പഠിക്കുന്ന വിദ്യാർഥികളുടെ സമുച്ചയമാണ് ഇതിലെ ലക്ഷ്യസമഷ്ടി. പക്ഷെ, വിദൂരങ്ങളായ മലയോരങ്ങളിലുള്ള ചില സ്കൂളുകളിൽ നിന്ന് വിവരങ്ങൾ ശേഖരിക്കുക ബുദ്ധിമുട്ടാണെന്ന് കണ്ടു് മററു് സ്കൂളുകളെ മാത്രം അടിസ്ഥാനമാക്കിയാണ് വിവരങ്ങൾ ശേഖരിച്ചതു് എന്നു കരുതുക. ഇവിടെ വിവരങ്ങൾ ശേഖരിക്കുന്നതിൽ പരിഗണിക്കപ്പെട്ട വിദ്യാർഥികളുടെ സമുച്ചയമാണ് പഠിതസമഷ്ടി. പഠിതസമഷ്ടിയിലെ അഭിലക്ഷണമൂല്യങ്ങളുടെ സമാന്തരമാധ്യം ലക്ഷ്യസമഷ്ടിയിലെ അഭിലക്ഷണമൂല്യങ്ങളുടെ സമാന്തരമാധ്യമായി അംഗീകരിക്കുന്നതു് വളരെ കരുതി മാത്രം ചെയ്യേണ്ട ഒന്നാണ്. കഴിവതും പഠിതസമഷ്ടിയും ലക്ഷ്യസമഷ്ടിയും ഒന്നായിരിക്കാൻ ശ്രദ്ധിക്കണം.

4. സെൻസസ്സും സാമ്പിളനവും

സമഷ്ടിയിലെ എല്ലാ പ്രാഥമിക വ്യക്തികളിൽ നിന്നും വിവരം ശേഖരിക്കുന്ന രീതിക്ക് "സെൻസസ്" സമ്പ്രദായമെന്നും, സമഷ്ടിയുടെ ഒരു വിഭാഗത്തിൽ നിന്നു മാത്രം വിവരം ശേഖരിക്കുന്ന സമ്പ്രദായത്തിന്ന് സാമ്പിളനമെന്നും പേർ പറയുന്നു. പ്രാകൃതരൂപത്തിലുള്ള സാമ്പിളനം സർവസാധാരണമാണ്. ഒരു ഭക്ഷണപദാർഥത്തിൽ നിന്ന് അല്പമെടുത്തു രുചിച്ചുനോക്കി സ്വാദു നിശ്ചയിക്കുമ്പോഴും ഒരു പിടി ധാന്യമണികൾ പരിശോധിച്ചുനോക്കി വില്പനക്കിട്ടിരിക്കുന്ന ധാന്യ കൂമ്പാരങ്ങളുടെ ഗുണദോഷം നിശ്ചയിക്കുമ്പോഴും സാമ്പിളനരീതിയാണ് നാമംഗീകരിക്കുന്നതു്. പരിഗണിക്കപ്പെടുന്ന അഭിലക്ഷണത്തെ സംബന്ധിച്ചിടത്തോളം പ്രാഥമികവ്യക്തികൾ തമ്മിൽ കാര്യമായ വ്യത്യാസമില്ലാത്ത സന്ദർഭങ്ങളിൽ മാത്രമെ ഇമ്മാതിരി സാമ്പിളനരീതി സമഷ്ടിയെപ്പററി ശരിയായ വിവരങ്ങൾ നൽകയുള്ളു. മററു സന്ദർഭങ്ങളിൽ കൂടുതൽ സങ്കീർണങ്ങളായ സാമ്പിളനരീതികൾ അംഗീകരിക്കേണ്ടതുണ്ടു്.

സെൻസസ് രീതിയാണോ സാമ്പിളനരീതിയാണോ ഒരു പ്രത്യേക സന്ദർഭത്തിൽ സ്വീകരിക്കേണ്ടതു് എന്നതു് ശ്രദ്ധാപൂർവം തീരുമാനിക്കേണ്ട ഒരു കാര്യമാണ്. സാമ്പിളനത്തിന്ന് സെൻസസിനെ അപേക്ഷിച്ച് സുപ്രധാനങ്ങളായ ചില വൈശിഷ്ട്യങ്ങളും അതുപോലെ തന്നെ ചില പോരായ്മകളും ഉണ്ടു്.

ചുരുങ്ങിയ സമയത്തിനുള്ളിൽ അത്യാവശ്യമായ ചില വിവരങ്ങൾ ശേഖ

രിക്കാൻ കഴിയുമെന്നുള്ളതാണ് സാമ്പിളനത്തിന്റെ ഒരു പ്രധാന മേന്മ. പെട്ടെന്നു തീരുമാനമെടുക്കേണ്ടി വരുന്ന സന്ദർഭങ്ങളിൽ ഇതു പ്രത്യേകിച്ചും ഉപയോഗപ്രദമാണ്. ഒരു തിരഞ്ഞെടുപ്പിൽ സ്ഥാനാർഥിയായി നിൽക്കാൻ ആഗ്രഹിക്കുന്ന ഒരാൾ തനിക്ക് വിജയസാധ്യതയുണ്ടോ എന്ന് സാമാന്യമായി ഒന്നു പരിശോധിച്ചിട്ടാവാം നാമനിർദേശപത്രിക സമർപ്പിക്കുക എന്നു കരുതുന്നതായി സങ്കല്പിക്കുക. സമ്മതിദായകരുടെ മനോഭാവം മനസ്സിലാക്കുകയാണ് ഇവിടെ ആവശ്യം. സെൻസസ് സമ്പ്രദായമാണ് ഈ അന്വേഷണത്തിന് സ്വീകരിക്കുന്നതെങ്കിൽ ഓരോ സമ്മതിദായകനെയും സമീപിച്ച് അനുകൂലമായ വോട്ടു ലഭിക്കാനുള്ള സാധ്യത ആരായേണ്ടതുണ്ട്. സമ്മതിദായകരുടെ എണ്ണം വലുതാണെങ്കിൽ ചുരുങ്ങിയ സമയം കൊണ്ട് ഇതു് സാധിച്ചെന്ന് വരികയില്ല. സാമ്പിളനരീതി അവലംബിച്ചാൽ താരതമ്യേന എളുപ്പത്തിലും വേഗത്തിലും ആവശ്യമായ വിവരങ്ങൾ ശേഖരിക്കുവാൻ കഴിയും.

ചെലവ് കുറഞ്ഞിരിക്കുമെന്നതാണ് സാമ്പിളനത്തിന്റെ മറ്റൊരു മേന്മ. സമഷ്ടിയിലെ എല്ലാ വ്യക്തികളിൽ നിന്നും വിവരങ്ങൾ ശേഖരിക്കുന്നില്ലാത്തതുകൊണ്ട്, അന്വേഷകരുടെ എണ്ണം, അന്വേഷണം സംഘടിപ്പിക്കുന്നതിലുള്ള മററു ചെലവുകൾ, അപഗ്രഥന ഘട്ടത്തിലെ ചെലവുകൾ തുടങ്ങിയവയെല്ലാം കുറവായിരിക്കും. സാമ്പത്തികശേഷി കുറവുള്ള വ്യക്തികളും സംഘടനകളും വിവരം ശേഖരിക്കാൻ ശ്രമിക്കുമ്പോൾ ഇതൊരു പ്രധാന പരിഗണനയാണ്.

സെൻസസ് രീതി സ്വീകാര്യമല്ലാത്ത ചില സന്ദർഭങ്ങളുണ്ട്. സമഷ്ടി അനന്തമോ സാങ്കല്പികമോ ആവുമ്പോഴും, സ്ഥിതിവിവരശേഖരണരീതി വ്യക്തിയെ നശിപ്പിക്കുന്ന രൂപത്തിലുള്ളതാകുമ്പോഴും, സെൻസസ് രീതി സ്വീകാര്യമല്ല. ഒരു വ്യവസായസ്ഥാപനം അവരുടെ പ്രത്യേകസാങ്കേതികരീതിയനുസരിച്ച് ഉല്പാദിപ്പിക്കുന്ന ഇലക്ട്രിക്ക് ബൾബുകളുടെ ജീവിതദൈർഘ്യം നിർണയിക്കുകയാണ് അന്വേഷണലക്ഷ്യമെന്നു വിചാരിക്കുക. ഈ അന്വേഷണത്തിലെ സമഷ്ടി ഈ സാങ്കേതിക രീതിയനുസരിച്ച് ആ സ്ഥാപനം ഉൽപാദിപ്പിച്ചിട്ടുള്ളതും ഉൽപാദിപ്പിച്ചേക്കാവുന്നതുമായ ബൾബുകളാണ്. ഇതു് ഒരു സാങ്കൽപിക സമഷ്ടിയാണ്. ഒരു പക്ഷേ അനന്തവുമാകാം. ഓരോ ബൾബും ഫ്യൂസാകുന്നതു വരെ കത്തിച്ചു നോക്കിയാണല്ലോ അതിന്റെ ജീവിതദൈർഘ്യം നിർണയിക്കേണ്ടതു്. വിവരം ശേഖരിക്കുവാനുള്ള ഈ രീതി വ്യക്തിയെ നശിപ്പിക്കുന്ന രൂപത്തിലുള്ളതുമാണ്. ഇമ്മാതിരി സന്ദർഭങ്ങളിൽ സെൻസസ് രീതി സ്വീകാര്യമേയല്ല.

സെൻസസ് രീതിയിലുള്ള അന്വേഷണം കൊണ്ട് ശേഖരിക്കാൻ ബുദ്ധിമുട്ടുള്ള പല വിവരങ്ങളും സാമ്പിളന രീതി അവലംബിച്ച് ശേഖരിക്കാൻ കഴിയുമെന്നുള്ളതാണ് മറെറാരു പ്രത്യേകത. ചില വിവരങ്ങൾ ശേഖരിക്കാൻ പ്രത്യേകം പരിശീലനം സിദ്ധിച്ച അന്വേഷകരും, വില കൂടിയ ഉപകരണങ്ങളും ആവശ്യമായി വരും. ഉദാഹരണമായി കേരളത്തിലെ കൃഷിഭൂമികളെ മണ്ണു പരിശോധിച്ച് തരം തിരിക്കണമെന്ന് വിചാരിക്കുക. മണ്ണു പരിശോധിക്കാൻ വിദഗ്ദ്ധന്മാരും ഉപകരണങ്ങളും ആവശ്യമാണ്. സെൻസസ് രീതി

ഉപയോഗിച്ച് ഈ സംരംഭത്തിൽ ഏർപ്പെട്ടാൽ കേരളത്തിലുള്ള ഓരോ തുണ്ടു കൃഷിഭൂമിയും പരിശോധനാവിധേയമാക്കേണ്ടി വരും. ഇതിനു വേണ്ട വിദഗ്ധന്മാരുടേയും ഉപകരണങ്ങളുടേയും എണ്ണം വളരെ വലുതായിരിക്കും. ഇത്രയധികം വിദഗ്ദ്ധന്മാരും ഉപകരണങ്ങളും പലപ്പോഴും ലഭിച്ചില്ലെന്നു വരും. ഈ പരിതസ്ഥിതിയിൽ ഓരോ വില്ലേജിലേയും കൃഷിഭൂമിയെ പ്രതിനിധീകരിക്കുന്ന ഓരോ തുണ്ടു ഭൂമി വീതം തിരഞ്ഞെടുത്ത് അവയെല്ലാമടങ്ങുന്ന സാമ്പിളിനെ പരിശോധനാവിധേയമാക്കി മേൽപറഞ്ഞ പഠനം നടത്താവുന്നതാണ്. ഇതിനു് വേണ്ടി വരുന്ന വിദഗ്ധന്മാരും ഉപകരണങ്ങളും സെൻസസ് രീതി അവലംബിക്കുമ്പോൾ വേണ്ടി വരുന്നതിനെക്കാൾ വളരെ കുറവായിരിക്കും. അങ്ങനെ സാമ്പിളനരീതി ഇത്തരം സന്ദർഭങ്ങളിൽ വളരെ സഹായകരമായിത്തീരുന്നു.

സെൻസസ് രീതി അവലംബിച്ചാലാണ് ഏററവും ശരിയായ വിവരങ്ങൾ ലഭിക്കുക എന്ന ധാരണയും അത്ര ശരിയല്ല. പലപ്പോഴും പ്രായോഗിക വൈഷമ്യങ്ങൾ മൂലം എല്ലാ പ്രാഥമികവ്യക്തികളിൽ നിന്നും വിവരം ശേഖരിക്കാൻ സാധിച്ചില്ലെന്നു വരും. അന്വേഷണ പരിപാടി ആസൂത്രണം ചെയ്യുന്നതിലുള്ള പോരായ്മകൾ കൊണ്ടും അന്വേഷകരുടെ പരിശീലനക്കുറവു കൊണ്ടും പല തരം പിശകുകൾ സംഭവിക്കാം. സാമ്പിളനത്തിൽ വിവരം ശേഖരിക്കേണ്ട പ്രാഥമികവ്യക്തികളുടെ എണ്ണം താരതമ്യേന കുറവായതു കൊണ്ട് ഇത്തരം പോരായ്മകൾ ഒരു പരിധി വരെയെങ്കിലും ഒഴിവാക്കാൻ കഴിയും. അധികം അന്വേഷകരുടെ സേവനം ആവശ്യമില്ലാത്തതു കൊണ്ട് അവരെ ശ്രദ്ധാപൂർവം തിരഞ്ഞെടുക്കാനും പരിശീലിപ്പിക്കാനും സാധിക്കും. സാമ്പിളിൽ ഉൾപ്പെടുന്ന എല്ലാ വ്യക്തികളെയും കണ്ടുപിടിച്ച് വിവരങ്ങൾ ശേഖരിക്കാൻ പ്രത്യേകം നിഷ്കർഷിക്കാം. അന്വേഷണത്തിൻെറ എല്ലാ ഘട്ടങ്ങളിലും ആവശ്യമായ പരിശോധനയും നിയന്ത്രണവും ഏർപ്പെടുത്താം. ചുരുക്കത്തിൽ സെൻസസ് രീതിയിലുള്ള അന്വേഷണം കൊണ്ടു ലഭിക്കുന്നതിനേക്കാൾ സൂക്ഷ്മമായ വിവരങ്ങൾ തന്നെ സാമ്പിളനം കൊണ്ടു ലഭിക്കാൻ പാടില്ലായ്കയില്ല.

പിശകുകൾ സെൻസസ് രീതിയിലോ സാമ്പിളന രീതിയിലോ ശേഖരിക്കുന്ന വിവരങ്ങളിലും കടന്നു കൂടാവുന്നതാണ്. സെൻസസ് രീതിയിൽ ശേഖരിച്ച വിവരങ്ങളിൽ സംഭവിച്ചിരിക്കാവുന്ന പിശകുകളുടെ അളവിനെ പററി ഒന്നും പറയാനാവില്ല. പക്ഷെ ശാസ്ത്രീയരീതികൾ അവലംബിച്ച് നടത്തുന്ന സാമ്പിളനത്തിൽ സാമ്പിളനം മൂലം സംഭവിക്കാവുന്ന പിശകുകളുടെ വലുപ്പത്തെപ്പററി സംഭാവ്യതാടിസ്ഥാനത്തിൽ മുൻകൂട്ടി തന്നെ പറയാൻ കഴിയും. സാമ്പിളനത്തിലൂടെ ലഭിക്കാവുന്ന വിവരങ്ങളുടെ വിശ്വാസ്യത നിർണയിക്കാൻ ഇതു് ഉപകരിക്കുന്നു. സാമ്പിളനത്തിൻെറ അതിപ്രധാനമായ ഒരു മേന്മയാണിതു്.

ഇങ്ങനെ പല മേന്മകളും സാമ്പിളനരീതിക്കുണ്ടെങ്കിലും സെൻസസ് രീതി അവലംബിച്ചേ മതിയാവു എന്ന സന്ദർഭങ്ങളും ഇല്ലാതില്ല. സമഷ്ടിയിലെ ഓരോ വ്യക്തിയേയും പററിയുള്ള വിവരങ്ങൾ ലഭിക്കേണ്ടതു് ആവശ്യമായി വരുന്ന സന്ദർഭങ്ങളിൽ സാമ്പിളനരീതി സ്വീകാര്യമല്ല. ഉദാഹരണ

ത്തിനു് ഓരോ പൗരനും നൽകേണ്ട നികുതി നിർണയിക്കുന്നതിനായി ഓരോരുത്തർക്കും എന്തു് വാർഷിക വരുമാനമുണ്ടെന്നു് കണ്ടുപിടിക്കുകയാണു് അന്വേഷണ ലക്ഷ്യമെങ്കിൽ സെൻസസ് രീതി തന്നെ അവലംബിക്കണം. പക്ഷെ, നികുതിയായി എന്തു് തുക ഗവൺമെൻറിനു് ലഭിച്ചേക്കാമെന്നു് ഏകദേശമായി നിർണയിക്കുകയാണു് ലക്ഷ്യമെങ്കിൽ സാമ്പിളനം മതിയാവും.

ഇതിൽ നിന്നെല്ലാം ഒരു സംഗതി വ്യക്തമാണു്. അന്വേഷണത്തിൻെറ ഉദ്ദേശവും ഉപയോഗവും പരിഗണിച്ചു വേണം അന്വേഷണരീതി നിശ്ചയിക്കുവാൻ. എത്ര സമയത്തിനുള്ളിൽ വിവരങ്ങൾ ശേഖരിക്കണം, എന്തുചെലവ് അനുവദിക്കും, എത്രമാത്രം സൂക്ഷ്മത വേണം, എല്ലാ വ്യക്തികളെപ്പററിയുമുള്ള വിവരങ്ങൾ ആവശ്യമുണ്ടോ എന്നെല്ലാം പരിഗണിച്ചിട്ടു വേണം അന്വേഷണ രീതിയെപ്പററി തീരുമാനമെടുക്കുവാൻ.

5. സാമ്പിളന ശാസ്ത്രം

പ്രാകൃതരൂപത്തിലുള്ള സാമ്പിളനം സർവസാധാരണമാണെങ്കിലും അതിൻെറ പരിമിതികൾ അവഗണിക്കാവുന്നവയല്ല. വ്യക്തികൾ തമ്മിലുള്ള അന്തരം താരതമ്യേന നിസ്സാരമായിരിക്കുമ്പോൾ മാത്രമെ ഈ രീതിയിൽ സമഷ്ടിയുടെ സ്വഭാവത്തെപ്പററി മനസ്സിലാക്കാൻ സാധിക്കു. സമഷ്ടിക്കുള്ളിലെ വൈവിധ്യം വർധിക്കുംതോറും സാമ്പിളനം കൂടുതൽ ശ്രദ്ധാപൂർവം ആസൂത്രണം ചെയ്യേണ്ടതായി വരും. എങ്കിൽ മാത്രമെ സമഷ്ടിയെ സാമ്പിൾ പ്രതിനിധീകരിക്കു. താത്വികവും പ്രായോഗികവുമായ അനവധി പ്രശ്നങ്ങൾ ഇതോടനുബന്ധിച്ചു് പൊന്തിവരും. അവക്കെല്ലാം പരിഹാരം കാണണം. കയ്യിലൊതുങ്ങുന്ന ഒരു ചെറിയ സമഷ്ടിയിൽ നിന്നു് സാമ്പിളെടുക്കുന്നതു പോലെ ലളിതമല്ല വൻതോതിലുള്ള സാമ്പിളനം. സമഷ്ടി സങ്കീർണവും വ്യക്തികൾ തമ്മിലുള്ള അന്തരം വലുതുമായിരിക്കുമ്പോൾ സാമ്പിളിൽ ഉൾപ്പെടേണ്ട അംഗങ്ങൾ ഏതെല്ലാമെന്നു് നിർണയിക്കുക, അവയിൽ നിന്നെല്ലാം വിവരങ്ങൾ ശേഖരിക്കുക, ശേഖരിച്ച വിവരങ്ങൾ അപഗ്രഥനം ചെയ്തു് സമഷ്ടിയുടെ പൊതുസ്വഭാവം മനസ്സിലാക്കുക തുടങ്ങിയവയെല്ലാം ബുദ്ധിമുട്ടേറിയ കാര്യങ്ങളാണു്. സാമ്പിളനശാസ്ത്രം കൈകാര്യം ചെയ്യുന്ന വിഷയങ്ങളാണു് ഇവ. ഇതു പോലെ തന്നെ സമഷ്ടിയോടു ബന്ധപ്പെട്ട പ്രാചലങ്ങളുടെ ആകലനം, നിർദിഷ്ട സൂക്ഷ്മതയോടു കൂടിയ ആകലനത്തിനു് ആവശ്യമായ സാമ്പിളിൻെറ വലുപ്പം, വിവിധ ആകലങ്ങളുടെ സൂക്ഷ്മത നിർണയിക്കൽ തുടങ്ങിയവയും സാമ്പിളനശാസ്ത്രത്തിൻെറ ഭാഗങ്ങളത്രെ. വിവിധ സാമ്പിളനരീതികൾ വിശദീകരിക്കുന്നതിലും അവ ഓരോന്നിൽ നിന്നും ലഭിക്കാവുന്ന ആകലങ്ങളുടെ പ്രത്യേകതകൾ ചർച്ച ചെയ്യുന്നതിലും സാമ്പിളനശാസ്ത്രം പ്രത്യേകം ശ്രദ്ധ പതിപ്പിക്കുന്നു. മുൻസൂചിപ്പിച്ചതു പോലെ കഴിഞ്ഞ അര നൂററാണ്ടിനുള്ളിൽ സാമ്പിളന ശാസ്ത്രത്തിനുണ്ടായ വളർച്ച സാമ്പിളനത്തിൻെറ കാര്യക്ഷമത വളരെയേറെ വർധിപ്പിച്ചിട്ടുണ്ടു്. ഒരു പ്രത്യേക വിവരം നിർദിഷ്ട സൂക്ഷ്മതയോടു കൂടി നിർണയിക്കാൻ അംഗീകരിക്കേണ്ട സാമ്പിളന രീതിയും, സാമ്പിളിൻെറ വലുപ്പവും,

അപഗ്രഥിക്കേണ്ട വിധവും ആവശ്യമായി വന്നേക്കാവുന്ന ചെലവിനെപ്പററിയുള്ള സാമാന്യവിവരങ്ങളും നിർദേശിക്കുവാൻ ഇന്നു് സാമ്പിളന ശാസ്ത്രത്തിനു് കഴിവുണ്ടു്.

parameter sampling error

6. പ്രാചലം, സാമ്പിളജം, സാമ്പിളനപ്പിശക്

population

Characteristic

ഏതൊരു സ്ഥിതിവിവരാന്വേഷണത്തിലും ഒരു സമഷ്ടിയും അതിനോടു ബന്ധപ്പെട്ട ഏതാനും അഭിലക്ഷണങ്ങളുമായിരിക്കുമല്ലോ പരിഗണിക്കപ്പെടുക. ഒരു അഭിലക്ഷണം മാത്രമെ പരിഗണിക്കാനുള്ളുവെങ്കിൽ, സമഷ്ടിയിലെ ഓരോ അംഗത്തിനും സംഗതമായ ഓരോ അഭിലക്ഷണമൂല്യവും ഉണ്ടായിരിക്കും. ഒന്നിലധികം അഭിലക്ഷണങ്ങൾ പരിഗണനയിലുണ്ടെങ്കിൽ അഭിലക്ഷണങ്ങളുടെ അത്രയും അംഗങ്ങളുള്ള ഓരോ സദിശമായിരിക്കും സമഷ്ടിയിലെ ഓരോ അംഗത്തിനും സംഗതമായി ഉണ്ടായിരിക്കുക. ഈ അഭിലക്ഷണമൂല്യങ്ങളുടെ ഏതൊരു ഫലനത്തിനും ഒരു പ്രാചലം എന്നു പേർ പറയും. ഉദാഹരണമായി 1969 സ്കൂൾ വർഷത്തിൽ കേരളത്തിലെ കോളേജുകളിൽ പഠനം നടത്തുന്ന വിദ്യാർഥികളുടെ വയസ്സാണു് പരിഗണിക്കപ്പെടുന്ന അഭിലക്ഷണമെന്നു കരുതുക. വയസ്സുകളുടെ ശരാശരി, മാധ്യം, പ്രസരണം തുടങ്ങിയവയെല്ലാം അഭിലക്ഷണമൂല്യങ്ങളുടെ ഫലനങ്ങളാണല്ലോ. അതു കൊണ്ടു് അവയെല്ലാം പ്രാചലങ്ങളാണു്. സമഷ്ടിയോടു ബന്ധപ്പെട്ട അചരങ്ങളാണു് പ്രാചലങ്ങൾ. ഒരു സമഷ്ടിക്കു് എത്ര പ്രാചലങ്ങൾ വേണമെങ്കിലും ഉണ്ടാവാം. സമഷ്ടിയുടെ അഭിലക്ഷണമൂല്യങ്ങൾക്കു എത്ര ഫലനങ്ങൾ വേണമെങ്കിലും ആവാമല്ലോ. പക്ഷെ അഭിലക്ഷണമൂല്യങ്ങളുടെ സംഭാവ്യതാവിതരണം അറിയാമെങ്കിൽ അതുൾക്കൊള്ളുന്നവയും സമഷ്ടിയോടു ബന്ധപ്പെട്ടവയുമായ അചരങ്ങളെ മാത്രം പ്രാചലങ്ങളായി പരിഗണിച്ചാൽ മതി. മുൻനിർവചനമനുസരിച്ചുള്ള മറ്റെല്ലാ പ്രാചലങ്ങളും ഇവയുടെ ഫലനങ്ങൾ മാത്രമായിരിക്കും. ഉദാഹരണമായി അഭിലക്ഷണമൂല്യങ്ങൾ നോർമൽ വിതരണത്തെ അനുവർത്തിക്കുന്നു എന്നിരിക്കട്ടെ. അതായതു് അഭിലക്ഷണമൂല്യം x ആണെങ്കിൽ,

$$P(x) = \frac{1}{\sqrt{2\pi}\sigma} e^{-\frac{(x-a)^2}{2\sigma^2}}, \quad -\infty \leq x \leq +\infty$$

എന്നർഥം. ഇവിടെ a, σ എന്നിവയാണു് പ്രാചലങ്ങൾ. അഭിലക്ഷണമൂല്യങ്ങളുടെ ഏതൊരു ഫലനത്തേയും a, σ എന്നിവയുടെ ഫലനമായി എഴുതാൻ കഴിയും. പ്രാചലങ്ങളെ അതതു പ്രശ്നത്തിന്റെ സ്വഭാവമനുസരിച്ചു് ഇതിലേതു വിധത്തിൽ വേണമെങ്കിലും നിർവചിക്കാം. ഫലത്തിൽ രണ്ടു നിർവചനങ്ങളും സമാനങ്ങളാണെന്നുള്ളതു് വ്യക്തമാണല്ലോ.

ഒന്നോ അതിലധികമോ പ്രാചലങ്ങളുടെ മൂല്യങ്ങൾ നിർണയിക്കുകയായിരിക്കും മിക്കവാറും എല്ലാ സ്ഥിതിവിവരാന്വേഷണങ്ങളുടെയും പ്രത്യക്ഷമോ പരോക്ഷമോ ആയ ലക്ഷ്യം. സാമ്പിളനരീതിയാണു് അവലംബിക്കുന്നതെങ്കിൽ ഇവയെ ആകലനം ചെയ്യാൻ മാത്രമെ കഴിയുകയുള്ളൂ. സമഷ്ടിയിലെ അഭില

ക്ഷണ മൂല്യങ്ങളുടെ ശരാശരിയാണ് നിർണയിക്കേണ്ട പ്രാചലമെന്നിരിക്കട്ടെ. സാമ്പിളനത്തിൽ, സമഷ്ടിയിലെ ഏതാനും അംഗങ്ങളുടെ അഭിലക്ഷണമൂല്യങ്ങൾ മാത്രമെ നാം ശേഖരിക്കുന്നുള്ളു. അവ മാത്രം ഉപയോഗിച്ച് സമഷ്ടിയുടെ ശരാശരി നിർണയിക്കാനാവില്ല. പക്ഷെ സാമ്പിൾ ശരാശരിയെ സമഷ്ടിയുടെ ശരാശരിയുടെ ആകലമായി സ്വീകരിക്കാം. ഇങ്ങനെ ആകലനത്തിനും ആകലങ്ങളുടെ സൂക്ഷ്മത നിർണയിക്കുന്നതിനും മററുമായി സാമ്പിളിലെ അഭിലക്ഷണമൂല്യങ്ങളുടെ പല ഫലനങ്ങളും പരിഗണിക്കേണ്ടതായി വരും. സാമ്പിൾ മൂല്യങ്ങളുടെ ഏതൊരു ഫലനത്തിനും സാമ്പിളജം എന്നു പേർ പറയുന്നു. സാമ്പിളജങ്ങളും അവയുടെ പ്രത്യേകതകളും സാമ്പിളനശാസ്ത്രത്തിലെ സുപ്രധാനങ്ങളായ പഠനവിഷയങ്ങളാണ്. സാമ്പിൾ ശരാശരി, മാധ്യം, പ്രസരണം തുടങ്ങിയവയെല്ലാം സാമ്പിളജങ്ങളത്രെ.

സാമ്പിളജങ്ങളുടെ ഒരു പ്രത്യേകത അവയുടെ മൂല്യം സാമ്പിൾ മാറുന്നതനുസരിച്ച് മാറിക്കൊണ്ടിരിക്കുമെന്നതാണ്. അവ പ്രാചലങ്ങളെപ്പോലെ അചരങ്ങളല്ല. സമഷ്ടിയിലെ ഓരോ വ്യക്തിക്കും സാമ്പിളിൽ ഉൾപ്പെട്ടുവാനുള്ള സംഭാവ്യത നിശ്ചിതമായിരിക്കത്തക്കവണ്ണം എടുക്കുന്ന സാമ്പിളുകൾക്ക് സംഭാവ്യതാ സാമ്പിളുകൾ എന്നു പറയും. സംഭാവ്യതാസാമ്പിളുകളോടു ബന്ധപ്പെട്ട സാമ്പിളജങ്ങൾ യാദൃച്ഛികചരങ്ങളാണ്. അവയ്ക്ക് ഓരോന്നിനും ഓരോ സംഭാവ്യതാ വിതരണമുണ്ടായിരിക്കും. ഇതിനെ ആ സാമ്പിളജത്തിൻെറ സാമ്പിളനവിതരണം എന്നാണ് വിളിക്കുന്നതു്. പക്ഷേ അതു നിർണയിക്കുക വളരെ ബുദ്ധിമുട്ടുള്ള ഒരു കാര്യമാണ്. സമഷ്ടി ചില പൊതു നിബന്ധനകൾ അനുസരിക്കുന്നുണ്ടെങ്കിൽ, അതിൽ നിന്നു് എടുക്കുന്ന സാമ്പിളുകളോടു ബന്ധപ്പെട്ട സാമ്പിളജങ്ങളുടെ സാമ്പിൾ പരിമാണം വലുതാണെങ്കിൽ സാമ്പിളന വിതരണം നോർമൽ വിതരണത്തോടു് അടുത്തു വരുമെന്ന് സാംഖ്യിക ശാസ്ത്രം സിദ്ധാന്തിക്കുന്നു. ഈ പൊതു നിബന്ധനകൾ മിക്കവാറും എല്ലാ സമഷ്ടികളും അനുസരിക്കുമെന്നതുകൊണ്ടു് സാമ്പിളനശാസ്ത്രത്തിലെ സുപ്രധാന സിദ്ധാന്തങ്ങളിലൊന്നാണിതു്. നോർമൽ വിതരണത്തിൻെറ സ്വഭാവം നിർണയിക്കുന്നതു് പ്രധാനമായും അതിൻെറ പ്രതീക്ഷയും പ്രസരണവുമാണ്. അതുകൊണ്ടാണ് സാമ്പിളജങ്ങളുടെ പ്രതീക്ഷയും സാമ്പിളന പ്രസരണവും നിർണയിക്കുന്നതിനു് സാമ്പിളനശാസ്ത്രം വളരെയധികം ശ്രദ്ധിക്കുന്നതു്.

പ്രാചലങ്ങളുടെ ആകലനമാണല്ലൊ സാമ്പിളനത്തിൻെറ പ്രധാന ലക്ഷ്യം. സാമ്പിളജങ്ങളാണ് അതിനു് ഉപയോഗിക്കുന്നതു്. ഒരു പ്രാചലത്തിൻെറ ആകലനത്തിനു് നിർദേശിക്കപ്പെട്ടിരിക്കുന്ന സാമ്പിളജത്തിനു് അതിൻെറ ആകലം എന്നും ഒരു പ്രത്യേക സാമ്പിളിനെ അടിസ്ഥാനമാക്കി നിർണയിക്കുന്ന അതിൻെറ മൂല്യത്തിനു് ആ പ്രാചലത്തിൻെറ ആകലിതമെന്നും പേർ പറയുന്നു. ഇവിടെ ശ്രദ്ധിക്കേണ്ട ഒരു പ്രത്യേകത ആകലം ഒരു യാദൃച്ഛിക ചരവും ആകലിതം ഒരു അചരവും ആണെന്നുള്ളതാണ്. സമഷ്ടിയുടെ ശരാശരിയാണ് ആകലനം ചെയ്യേണ്ട പ്രാചലമെന്നും സാമ്പിളിൻെറ ശരാശരിയാണ് അതിനു് ഉപയോഗിക്കുന്നതെന്നും കരുതുക. സാമ്പിൾ ശരാശരി എന്ന സാമ്പിള

ജത്തിനു് "ആകലം" എന്നും ഒരു പ്രത്യേക സാമ്പിളിലെ അതിൻെറ മൂല്യത്തിനു് "ആകലിത"മെന്നും പേർ പറയുന്നു.

സമഷ്ടിയുടെ ഒരു ഭാഗം മാത്രമാണു് സാമ്പിൾ. അതിൻെറ അടിസ്ഥാനത്തിൽ സമഷ്ടിയെപ്പററിയുള്ള നിഗമനങ്ങളിൽ എത്തിച്ചേരാനാണു് നാം ശ്രമിക്കുന്നതു്. സ്വാഭാവികമായും ഇങ്ങനെ എത്തിച്ചേരുന്ന നിഗമനങ്ങളിൽ പിശകുകളുണ്ടാവാം. ഇത്തരം പിശകുകൾക്കു് സാമ്പിളനപ്പിശകുകൾ എന്നു പേർ പറയുന്നു. മററൊരു തരത്തിൽ പറഞ്ഞാൽ, സമഷ്ടിയുടെ ഒരംശം മാത്രമായ സാമ്പിളിനെ അടിസ്ഥാനമാക്കി സമഷ്ടിയെപ്പററിയുള്ള നിഗമനങ്ങൾ ആവിഷ്ക്കരിക്കുന്നതിൽ സംഭവിക്കാവുന്ന പിശകുകൾക്കു് സാമ്പിളനപ്പിശകുകൾ എന്നു പേർ പറയുന്നു. എന്നാൽ സാമ്പിളെടുക്കുന്നതിലോ അഭിലക്ഷണമൂല്യങ്ങൾ കണ്ടുപിടിക്കുന്നതിലോ, അവയുപയോഗിച്ചു് ഗണനങ്ങൾ നടത്തുന്നതിലോ സംഭവിച്ചേക്കാവുന്ന പിശകുകൾ ഇതിൽ ഉൾപ്പെടുന്നില്ല. അത്തരം പിശകുകൾക്കു് സാമ്പിളനേതര പിശകുകൾ എന്നു് പേർ പറയുന്നു. ഇമ്മാതിരി പിശകുകൾ പ്രത്യേകം ശ്രദ്ധിച്ചാൽ കുറെയെല്ലാം ഒഴിവാക്കാൻ സാധിക്കും. പക്ഷേ സാമ്പിളന പിശകുകളുടെ സ്ഥിതി അതല്ല. സാമ്പിൾ പരിമാണം സമഷ്ടിയുടെ പരിമാണത്തോടു തുല്യമാക്കുക എന്നതു മാത്രമാണു് ഇതു് ഒഴിവാക്കാനുള്ള ഒരേ ഒരു മാർഗം. അതായതു് സാമ്പിളനരീതി ഉപേക്ഷിക്കുക. പക്ഷേ ഇതു് അർഥശൂന്യമായ ഒരു പരിഹാരനിർദേശമാണല്ലോ.

സാമ്പിളനപ്പിശകു് ഒഴിവാക്കാനാവില്ലെങ്കിലും അതിൻെറ സ്വഭാവത്തെയും പരിമാണത്തെയും പററി സംഭാവ്യതാടിസ്ഥാനത്തിൽ വളരെയെല്ലാം മനസ്സിലാക്കുവാൻ ആധുനികസാമ്പിളനശാസ്ത്രത്തിനു് കഴിഞ്ഞിട്ടുണ്ടു്. സാമ്പിൾ പരിമാണം വർധിക്കുന്നതനുസരിച്ചു് സാമ്പിളനപ്പിശകിൻെറ പരിമാണം കുറയുമെന്നതാണു് സുപ്രധാനമായ ഒരു വസ്തുത. പക്ഷെ വളരെ ചെറിയ സംഖ്യയിൽ നിന്നു് സാമ്പിൾ പരിമാണം വർധിച്ചു തുടങ്ങുമ്പോൾ സാമ്പിളനപ്പിശകു് പെട്ടെന്നു് ഒന്നു കുറയുമെങ്കിലും പിന്നീടുള്ള കുറവു് അത്ര ഗണനീയമല്ല. എന്നാൽ സാമ്പിൾ പരിമാണത്തിൻെറ വർധനവു് സാമ്പിളനേതരപ്പിശകുകൾ അധികരിപ്പിക്കുകയും അന്വേഷണത്തിൻെറ ആകെ ചെലവു് വർധിപ്പിക്കുകയും ചെയ്യും. ഈ കാര്യങ്ങളെല്ലാം പരിഗണിച്ചിട്ടാണു് ഇഷ്ടതമമായ സാമ്പിൾ പരിമാണം നിർണയിക്കേണ്ടതു്.

7. നോർമൽ വിതരണം

സാംഖ്യികത്തിൻെറ മററു ശാഖകളിലെന്നതുപോലെ സാമ്പിളനത്തിലും നോർമൽ വിതരണത്തിനു് സുപ്രധാനമായ ഒരു സ്ഥാനമുണ്ടു്. സാമ്പിൾ പരിമാണം വലുതാണെങ്കിൽ സാമ്പിളജങ്ങളുടെ സംഭാവ്യതാവിതരണം നോർമൽ വിതരണത്തോടു് അടുത്തു വരുമെന്നു് മുൻപുതന്നെ സൂചിപ്പിച്ചിട്ടുണ്ടല്ലൊ. ഇക്കാരണം കൊണ്ടു് ആകലങ്ങൾ നിർണയിക്കുന്നതിലും അവയുടെ സൂക്ഷ്മത നിശ്ചയിക്കുന്നതിലുമെല്ലാം നോർമൽ വിതരണം വലിയൊരു പങ്കു വഹിക്കുന്നു.

നോർമൽ വിതരണത്തിന്റെ ചില പൊതുസ്വഭാവങ്ങൾ താഴെ കൊടുക്കുന്നു.

X എന്ന യാദൃച്ഛിക ചരം സന്തതവും അതിന്റെ ആവൃത്തിഫലനം

$$f(x) = \frac{1}{\sqrt{2\pi}\,\sigma} e^{-\frac{(x-a)^2}{2\sigma^2}}, -\infty \leq x \leq +\infty$$

എന്നതുമാണെങ്കിൽ x ഒരു നോർമൽ യാദൃച്ഛികചരവും അതോടു ബന്ധപ്പെട്ട സമഷ്ടിയുടെ സംഭാവ്യതാവിതരണം നോർമലും ആണെന്നു പറയും. ഇതിന്റെ പ്രാചലങ്ങൾ a-യും σ-യുമാണ്. ഒരു നോർമൽ സമഷ്ടിയുടെ എല്ലാ പ്രാചലങ്ങളും ഇവയുടെ ഫലനങ്ങളാണെന്നു കാണിക്കുവാൻ കഴിയും. a, x ന്റെ പ്രതീക്ഷയും, σ അതിന്റെ മാനകവിചലനവുമാണ്.

y=f(x)ന്റെ ലേഖയുടെ സാമാന്യരൂപം അനുബന്ധം 1-ൽ കൊടുത്തിരിക്കുന്നു. ഈ ലേഖ x = a എന്ന രേഖയ്ക്കു സമമിതമാണ്. σ യുടെ മൂല്യമാണ് ലേഖയുടെ പ്രകൃതി നിർണയിക്കുന്നത്. σ യുടെ മൂല്യം വർധിക്കുമ്പോൾ ലേഖയുടെ മധ്യഭാഗം കൂടുതൽ കൂർത്തതും കുറയുമ്പോൾ കൂടുതൽ പരന്നതും ആയിരിക്കും. x ന്റെ മൂല്യം X എന്ന സംഖ്യയെക്കാൾ കുറവായിരിക്കുവാനുള്ള സംഭാവ്യത $(-\infty, X)$ എന്ന അന്തരാളത്തിലെ f(x) ന്റെ സമാകലമായിരിക്കും.

$$P_r(x \leq X) = \frac{1}{\sqrt{2\pi}\,\sigma} \int_{-\infty}^{X} e^{-\frac{(x-a)^2}{2\sigma^2}} dx$$

ഇതു് ലേഖയിൽ x = X എന്ന രേഖയുടെ ഇടത്തു വശത്തു് ലേഖക്കും x-അക്ഷത്തിനും ഇടക്കുള്ള വിസ്തീർണമാണല്ലോ. x ന്റെ വിവിധ മൂല്യങ്ങൾക്ക് അനുഗുണമായി ഈ സമാകലം ഗണിച്ച് സാരണികൾ തയ്യാറാക്കാവുന്നതാണ്. സാരണി തയ്യാറാക്കുന്നതിൽ ലേഖയുടെ സമമിതി വളരെ ഉപയോഗപ്രദമത്രെ. x ന്റെ മൂല്യം (a_1, a_2) എന്ന അന്തരാളത്തിലായിരിക്കുവാനുള്ള സംഭാവ്യത ഈ സാരണി ഉപയോഗിച്ച് ഗണിച്ചെടുക്കാം. വിവിധ ആവശ്യങ്ങൾക്ക് ഉപയോഗപ്രദമാകുവാൻ വേണ്ടി നോർമൽ വിതരണത്തിന്റെ പല സാരണികളും തയ്യാറാക്കിയിട്ടുണ്ട്. x ന്റെ മൂല്യം $(a-k\sigma, a+k\sigma)$ എന്ന അന്തരാളത്തിലായിരിക്കുവാനുള്ള സംഭാവ്യത, സാമ്പിളനത്തിലും മററു പല സാംഖ്യികശാഖകളിലും ആവശ്യമായി വരാറുണ്ട്. ഇതിനായി k യുടെ വിവിധ മൂല്യങ്ങൾക്ക് അനുസൃതമായുള്ള ഈ സംഭാവ്യത നൽകുന്ന സാരണികളും തയ്യാറാക്കണം. $t = (x-a)/\sigma$ ആണെങ്കിൽ t ക്ക് മാനകനോർമൽ ചരം എന്നു പറയും. ഇതിന്റെ ആവൃത്തി ഫലനം

$$P(t) = \frac{1}{\sqrt{2\pi}} e^{-\frac{t^2}{2}}, -\infty \leq t \leq \infty$$

ആണ്. x ന്റെ മൂല്യം $(a-k\sigma, a+k\sigma)$ എന്ന അന്തരാളത്തിലായിരി

കുവാനുള്ള സംഭാവ്യത t യുടെ മൂല്യം (—k, k) എന്ന അന്തരാളത്തിലായിരിക്കു വാനുള്ള സംഭാവ്യത തന്നെയാണ്. കൂടുതൽ സൗകര്യത്തിനായി

$$P_r(t<T) = \frac{1}{\sqrt{2\pi}} \int_{-\infty}^{T} e^{-\frac{t^2}{2}} dt$$

എന്ന സമാകലത്തിൻെറ സാരണിയാണ് സാധാരണ തയ്യാറാക്കാറുള്ളത്. ഇമ്മാതിരി ഒരു സാരണി അനുബന്ധം (1) ആയി ചേർത്തിരിക്കുന്നു. t യുടെ മൂല്യം (—k, k) എന്ന അന്തരാളത്തിലായിരിക്കുവാനുള്ള സംഭാവ്യത ഈ സാരണിയിൽ നിന്നു കണ്ടുപിടിക്കാം. k യുടെ ഏതാനും മൂല്യങ്ങൾക്കു് അനുഗുണമായുള്ള ഈ സംഭാവ്യത ഒരു സാരണിയായി താഴെ കൊടുക്കുന്നു.

k യുടെ ഏതാനും മൂല്യങ്ങൾക്കു് അനുഗുണമായി മാനക നോർമൽ വിതരണം അനുവർത്തിക്കുന്ന ഒരു യാദൃച്ഛികചരം (—k, k) എന്ന അന്തരാളത്തിലായിരിക്കുവാനുള്ള സംഭാവ്യത.

k	2.576	1.960	1.645	0.674
സംഭാവ്യത	.99	.95	.90	.50

ഈ മൂല്യങ്ങൾ സാമ്പിളനത്തിൽ അത്യന്തം ഉപയോഗപ്രദമായതു കൊണ്ടാണ് അവ ഇവിടെ കൊടുത്തിരിക്കുന്നതു്. k യുടെ മറ്റു മൂല്യങ്ങൾക്കു് അനുഗുണമായ സംഭാവ്യതകൾ അനുബന്ധം (1) ൽ നിന്നും കണ്ടുപിടിക്കാവുന്നതാണ്.

8. സംഭാവ്യതാസാമ്പിളനവും ഐച്ഛിക സാമ്പിളനവും

സാമ്പിളനരീതികളെ പൊതുവെ രണ്ടായി തിരിക്കാം: സംഭാവ്യതാസാമ്പിളനവും ഐച്ഛിക സാമ്പിളനവും. സമഷ്ടിയിലെ ഓരോ വ്യക്തിക്കും സാമ്പിളിൽ ഉൾപ്പെടുവാനുള്ള സംഭാവ്യതക്കു് നിശ്ചിത മൂല്യങ്ങൾ ഉണ്ടായിരിക്കത്തക്കവണ്ണമുള്ള സാമ്പിളന രീതിക്കു് സംഭാവ്യതാസാമ്പിളനം എന്നു പേർ പറയുന്നു. മറെറാരു തരത്തിൽ പറഞ്ഞാൽ സമഷ്ടിയിലെ u_i എന്ന വ്യക്തി സാമ്പിളനത്തിൽ ഉൾപ്പെടുവാനുള്ള സംഭാവ്യത P_i എന്ന ഒരു നിശ്ചിത സംഖ്യയായിരിക്കത്തക്കവണ്ണമുള്ള സാമ്പിളനരീതിയാണ് സംഭാവ്യതാസാമ്പിളനം. ഒരു വിദ്യാലയത്തിലെ വിദ്യാർത്ഥികളിൽ നിന്നു് പത്തു് അംഗങ്ങളുള്ള ഒരു സാമ്പിൾ എടുക്കണമെന്നിരിക്കട്ടെ. ഓരോ വിദ്യാർത്ഥിയുടെയും പേരു് ഒരേ വലുപ്പമുള്ള കടലാസുതുണ്ടുകളിൽ എഴുതി ഒരുപോലെ ചുരുട്ടി ഒരു പാത്രത്തിലിട്ടു് നല്ലതുപോലെ ഇളക്കി അതിൽ നിന്നു് പത്തു് കടലാസുതുണ്ടുകൾ എടുത്തു് അവയിൽ എഴുതിയിരിക്കുന്ന പേരുകളുള്ള വിദ്യാർത്ഥികളെ തിരഞ്ഞെടുക്കുക. സമഷ്ടിയിലെ ഓരോ അംഗത്തിനും സാമ്പിളിൽ ഉൾപ്പെടുവാനുള്ള സംഭാവ്യത തുല്യമായിരിക്കത്തക്കവണ്ണമുള്ള ഒരു സാമ്പിളനരീതിയാണ് ഇതു്. അതുകൊണ്ടു തന്നെ അതു് ഒരു സംഭാവ്യതാസാമ്പിളനവുമാണ്.

ഐച്ഛികസാമ്പിളനത്തിൽ സാമ്പിളെടുക്കുന്ന ആളിന്റെ അഭിപ്രായമാണ് ഒരു വ്യക്തിയെ സാമ്പിളിൽ ഉൾപ്പെടുത്തണമോ വേണ്ടയോ എന്നു നിർണയിക്കുന്നത്. മുൻ ഉദാഹരണത്തിൽ സ്കൂളിലെ ഹെഡ്മാസ്റ്റർ അദ്ദേഹത്തിനു യുക്തമെന്നു തോന്നുന്ന മാനദണ്ഡങ്ങളുപയോഗിച്ച് പത്ത് വിദ്യാർഥികളെ തിരഞ്ഞെടുക്കുകയാണെങ്കിൽ അത് ഐച്ഛിക സാമ്പിളനമാണ്. ഐച്ഛിക സാമ്പിൾ സംഭാവ്യതാ സാമ്പിളിനെക്കാളധികം സമഷ്ടിയെ പ്രതിനിധീകരിച്ചു എന്നു വരാം. പക്ഷേ അതിൽ നിന്നു ലഭിക്കുന്ന വിവരങ്ങളുടെ സൂക്ഷ്മതയെപ്പററി സംഭാവ്യതാടിസ്ഥാനത്തിൽ ഒന്നും പറയാനാവില്ല. മറെറാരു തരത്തിൽ പറഞ്ഞാൽ, ഐച്ഛിക സാമ്പിളനത്തിലെ സാമ്പിളനപ്പിശകിന്റെ പരിമാണം എത്രയെന്ന് നിർണയിക്കുവാൻ മാർഗമൊന്നുമില്ല. സാമ്പിളനസിദ്ധാന്തങ്ങൾ സംഭാവ്യതാ സാമ്പിളനരീതിയിൽ ലഭിക്കുന്ന സാമ്പിളുകൾ ഉപയോഗിച്ച് സമഷ്ടിയെപ്പററി പഠനം നടത്താൻ മാത്രമെ ഉപകരിക്കു. ഐച്ഛികസാമ്പിളനത്തിൽ സാമ്പിളനത്തിനു് ഒരു പങ്കും വഹിക്കുവാനില്ല.

സംഭാവ്യതാസാമ്പിളനം പല തരത്തിലുണ്ട്. അവയിൽ പ്രധാനപ്പെട്ട ചിലതിനെപ്പററി മാത്രമാണ് ഈ ഗ്രന്ഥത്തിൽ ചർച്ച ചെയ്യുന്നത്.

9. ആകലങ്ങളും അഭിനതിയും

പ്രാചലങ്ങളുടെ ആകലനമാണ് സാമ്പിളനത്തിന്റെ ഏററവും പ്രധാന ലക്ഷ്യം. സാധാരണ ആകലനം ചെയ്യേണ്ടി വരാറുള്ള പ്രാചലങ്ങൾ സമഷ്ടിയുടെ സമാന്തരമാധ്യം, പ്രസരണം, ഏതെങ്കിലും പ്രത്യേകതയുള്ള അംഗങ്ങളുടെ അനുപാതം തുടങ്ങിയവയാണ്. ആകലനത്തിനു് ഉപയോഗിക്കപ്പെടുന്നത് സാമ്പിളജങ്ങളാണ്. ഒരേ പ്രാചലത്തിന്റെ ആകലങ്ങളായി ഒന്നിലധികം സാമ്പിളജങ്ങൾ നിർദേശിക്കപ്പെട്ടു എന്നു വരാം. താത്വികമായി പറഞ്ഞാൽ, ഏതൊരു സാമ്പിളനത്തെയും ഏതൊരു പ്രാചലത്തിന്റെയും ആകലമായി നിർദേശിക്കാം. ഉദാഹരണമായി സമഷ്ടിയുടെ സമാന്തരമാധ്യത്തിന്റെ ആകലമായി സാമ്പിൾ സമാന്തരമാധ്യമോ, സാമ്പിൾ മാധ്യകമോ, ബഹുലകമോ നിർദേശിക്കപ്പെടാം. ഇവയിൽ ഏതാണ് കൂടുതൽ മെച്ചപ്പെട്ട ആകലം എന്നു നിർണയിക്കേണ്ടതു് ഒരു പ്രശ്നമായി തീരുന്നു. ഒരു നല്ല ആകലത്തിനു് എന്തെല്ലാം ഗുണങ്ങളാണ് ഉണ്ടായിരിക്കേണ്ടതു് എന്നതിനെപ്പറ്റിയുള്ള ധാരണകളാണ് ഏററവും നല്ല ആകലം തിരഞ്ഞെടുക്കുന്നതിനു് ആധാരം. ആകലനശാസ്ത്രത്തിൽ ഒരു ആകലനത്തിനു് അഭിലഷണീയങ്ങളായ വിവിധ ഗുണങ്ങളെപ്പററി ചർച്ച ചെയ്യുന്നുണ്ട്. അതിൽ സുപ്രധാനമായ ഒന്നാണ് അനഭിനതത്വം—അഭിനതി ഇല്ലാതിരിക്കുക എന്നതു്. θ എന്ന പ്രാചലത്തിന്റെ ആകലമായി t എന്ന സാമ്പിളജം നിർദേശിക്കപ്പെടുന്നു എന്നിരിക്കട്ടെ. t ഒരു യാദൃച്ഛികചരമാണ്; t യുടെ പ്രതീക്ഷയെ $E(t)$ കുറിക്കുന്നു എന്നു സങ്കല്പിക്കുക. $E(t) = \theta$ ആണെങ്കിൽ, t ഒരു അനഭിനത ആകലമാണെന്നു പറയും. $B(t) = E(t) - \theta$ യാണ് 't' യുടെ 'അഭിനതി'. $B(t)$ യുടെ ഋണധനത്വം അനുസരിച്ച് അഭിനതി യഥാക്രമം ഋണമെന്നോ ധനമെന്നോ വ്യവഹരിക്കപ്പെടുകയും ചെയ്യും.

അനഭിനതത്വം അഭിലഷണീയമായ ഒരു ഗുണമാണെങ്കിലും അഭിനത ആകലങ്ങൾക്ക് പലപ്പോഴും മുൻഗണന നൽകാറുണ്ട്. അനഭിനതആകലങ്ങൾ ഇല്ലാതെ വരികയോ, ഗണനസൗകര്യം തുടങ്ങിയ മററു പരിഗണനകൾ അഭിനത ആകലത്തിനു് അനുകൂലമായി വരികയോ ചെയ്യുമ്പോഴാണ് ഇങ്ങനെ ചെയ്യാറുള്ളതു്.

അവിരോധിത്വം എന്ന ഗുണം അഭിനത ആകലങ്ങളിൽ നിന്നു് ഏററവും മെച്ചപ്പെട്ടതു് തെരഞ്ഞെടുക്കുന്നതിൽ പരിഗണിക്കപ്പെടാറുണ്ട്. 't' യുടെ മൂല്യം θ യോടു് അടുത്തടുത്തു വരുവാനുള്ള സംഭാവ്യത സാമ്പിൾ പരിമാണം വർധിക്കുന്നതനുസരിച്ച് വർധിച്ചു വരിക എന്നതാണ് ഈ ഗുണം. മറെറാരു തരത്തിൽ പറഞ്ഞാൽ, ε എത്ര വേണമെങ്കിലും ചെറുതാകാവുന്ന ഒരു ധന സംഖ്യയും n സാമ്പിൾ പരിമാണവുമാണെന്നിരിക്കട്ടെ. $Pr[|t-\theta|<\varepsilon]$, n വലുതാകുന്നതനുസരിച്ച്, 1 നോടു് അടുത്തടുത്തു വരുമെങ്കിൽ, t അവിരോധിത്വമുള്ള ഒരു ആകലമാണ്. വലിയ സാമ്പിളാണ് എടുക്കുന്നതെങ്കിൽ ആകലിതം പ്രാചലത്തിന്റെ പരിസരത്തിലായിരിക്കുവാൻ വളരെയധികം സംഭാവ്യതയുണ്ട് എന്നതാണ് ഇമ്മാതിരി ആകലങ്ങളുടെ മെച്ചം.

10. പിശകു വർഗമാധ്യവും സാമ്പിളനപ്രസരണവും

θ എന്ന പ്രാചലത്തിന്റെ ആകലമായി t എന്ന സാമ്പിളജം നിർദേശിക്കപ്പെട്ടിരിക്കുന്നു എന്നിരിക്കട്ടെ. സാമ്പിൾ മാറുന്നതനുസരിച്ച് t യുടെ മൂല്യവും മാറാവുന്നതാണ്. അതുകൊണ്ട് t യിലെ പിശകായ $t-\theta$ ഒരു യാദൃച്ഛികചരമാണെന്നു സിദ്ധിക്കുന്നു. $(t-\theta)^2$ നു് പിശകു വർഗം എന്നും അതിന്റെ പ്രതീക്ഷയായ $E(t-\theta)^2$ നു് പിശകുവർഗമാധ്യം എന്നും പേർ പറയുന്നു. ഇതിന്റെ വർഗമൂലത്തിനു് പിശകുവർഗമാധ്യമൂലം എന്നും പേർ പറയുന്നു. $M(t)$ എന്ന ചിഹ്നം കൊണ്ടാണ് സാധാരണയായി ഇതിനെ സൂചിപ്പിക്കാറുള്ളതു്. അതായതു് $M(t) = E(t-\theta)^2$. t യുടെ പ്രതീക്ഷ $E(t)$ ആണെന്നിരിക്കട്ടെ. $V(t) = E\{t - E(t)\}^2$ നു് t യുടെ സാമ്പിളനപ്രസരണം എന്നു പേർ പറയും. ഒരു സമഷ്ടിയിൽ നിന്നു് നിർദിഷ്ടമായ സാമ്പിളനരീതിയനുസരിച്ച് എടുക്കാവുന്ന നിശ്ചിത പരിമാണമുള്ള സാമ്പിളുകളിലെ 't' യുടെ മൂല്യങ്ങളുടെ പ്രസരണമാണിതു്. $M(t)$ യും $V(t)$ യും തമ്മിലുള്ള ബന്ധം താഴെ കൊടുക്കുന്നു.

$$\begin{aligned} M(t) = E[t-\theta]^2 &= E[t - E(t) + E(t) - \theta]^2 \\ &= E[t - E(t)]^2 + [E(t) - \theta]^2 \\ &= V(t) + [E(t) - \theta]^2 \end{aligned}$$

$E(t) = \theta$ യാണെങ്കിൽ, അതായതു് t ഒരു അനഭിനത ആകലമാണെങ്കിൽ $M(t)$ യും $V(t)$ യും തുല്യമായിരിക്കുമെന്നു് ഇതിൽ നിന്നു സിദ്ധിക്കുന്നു.

സാമ്പിളനപ്രസരണത്തിന്റെ വർഗമൂലത്തിനു് ആ ആകലത്തിന്റെ പ്രമാണപ്പിശകു് എന്നു പേർ പറയുന്നു. $\sigma(t)$ എന്ന ചിഹ്നമാണ് t യുടെ പ്രമാണപ്പിശകിനെ സൂചിപ്പിക്കാനായി ഉപയോഗിക്കാറുള്ളതു്. അതായതു്,

$\sigma(t) = \sqrt{V(t)}$. പ്രമാണപ്പിശകും ആകലത്തിൻെറ പ്രതീക്ഷയുമായുള്ള അനുപാതത്തിനു് ആപേക്ഷികപ്രമാണപ്പിശകു്, അല്ലെങ്കിൽ വിചരണഗുണാങ്കം എന്നു പറയും. $C(t)$ എന്ന ചിഹ്നം കൊണ്ടു സൂചിപ്പിക്കാറുള്ള ഇതു് ശതമാനമായി എഴുതുകയാണു് പതിവു്.

$$C(t) = \frac{\sigma(t)}{E(t)} \times 100$$

എന്നുസാരം. ഇതിൻെറ വർഗത്തിനു് ആകലത്തിൻെറ ആപേക്ഷികപ്രസരണമെന്നും പറയും.

11. സാംഖ്യികഅനുമാനം സാമ്പിളനത്തിൽ

സാമ്പിളനശാസ്ത്രത്തിൻെറ അസ്ഥിവാരം സാംഖ്യികഅനുമാനമാണു്. സാമ്പിളിൽ നിന്നു് സമഷ്ടിയെപ്പററിയുള്ള അനുമാനങ്ങൾ എത്തിച്ചേരാനുള്ള മാർഗങ്ങൾ അതായതു്, സമഷ്ടിയുടെ വിതരണം നിർണയിക്കുക, അതുൾക്കൊള്ളുന്ന പ്രാചലങ്ങൾ ആകലനം ചെയ്യുക, അവയുടെ വിശ്വാസ്യതാന്തരാളങ്ങൾ നിർദേശിക്കുക, സമഷ്ടിയുടെ വിതരണത്തെയും പ്രാചലങ്ങളുടെ മൂല്യങ്ങളെയും പററിയുള്ള സങ്കല്പങ്ങൾ പരീക്ഷിക്കുക തുടങ്ങിയവ ആണല്ലോ സാംഖ്യിക അനുമാനം കൈകാര്യം ചെയ്യുന്ന വിഷയങ്ങൾ. ഇവയെല്ലാം ഒരു രൂപത്തിലല്ലെങ്കിൽ മറെറാരു രൂപത്തിൽ സാമ്പിളനത്തിൽ ഉപയോഗിക്കേണ്ടതായി വരും. സാമ്പിളിൻെറ വലുപ്പം വർധിക്കുമ്പോൾ ചില പ്രാഥമിക നിബന്ധനകൾ അനുസരിക്കുന്ന എല്ലാ സാമ്പിളജങ്ങളുടെയും സാമ്പിളനവിതരണം നോർമൽ വിതരണത്തോടു് അടുത്തു വരുമെന്നുള്ള സാംഖ്യിക അനുമാനശാസ്ത്രതത്വം സാമ്പിളനത്തിൽ മൗലിക പ്രാധാന്യമുള്ള ഒന്നാണു്. സാമ്പിളജങ്ങളുടെ പ്രതീക്ഷയും പ്രസരണവും നിർണയിക്കുന്നതിനു് സാമ്പിളനത്തിൻെറ ഏറിയ ഭാഗവും വിനിയോഗിച്ചിരിക്കുന്നതു് ഇക്കാരണം കൊണ്ടാണു്. നോർമൽ വിതരണം അനുവർത്തിക്കുന്ന യാദൃച്ഛികചരങ്ങളോടു് ബന്ധപ്പെട്ട പ്രാചലങ്ങളുടെ ആകലനം, വിശ്വാസ്യതാന്തരാളനിർണയം എന്നിവയും അവയോടു ബന്ധപ്പെട്ട സങ്കല്പങ്ങളുടെ പരീക്ഷയും, അവയുടെ പ്രതീക്ഷയേയും പ്രസരണത്തേയും ആധാരമാക്കിയാണല്ലോ നിർവഹിക്കുക. അങ്ങനെ സാമ്പിളനത്തിൻെറ താത്വിക പശ്ചാത്തലം സാംഖ്യിക അനുമാനമാണു്.

12. സൂക്ഷ്മതയും യാഥാർഥ്യവും

പ്രാചലങ്ങളുടെ ആകലനവുമായി ബന്ധപ്പെട്ടത്തി സാമ്പിളനത്തിൽ ഉപയോഗിക്കാറുള്ള രണ്ടു പദങ്ങളാണു് ഇവ. ഒരു മീററർ നീളമുള്ളതെന്നു കരുതപ്പെടുന്ന ഒരു സ്കെയിൽ ഉപയോഗിച്ചു് കൃത്യം 50 മീററർ അകലത്തിലുള്ള രണ്ടു ബിന്ദുക്കൾ തമ്മിലുള്ള ദൂരം അളക്കുന്നു എന്നിരിക്കട്ടെ. ആദ്യത്തെ പ്രാവശ്യം അളക്കുമ്പോൾ കിട്ടുന്ന അകലവും രണ്ടാമതൊരു പ്രാവശ്യം അളന്നാൽ കിട്ടുന്ന അകലവും ഒന്നു തന്നെ ആയിരിക്കണമെന്നില്ല. ആവർത്തിച്ചാവർത്തിച്ചളന്നാൽ വ്യത്യസ്തങ്ങളായ അളവുകൾ കിട്ടി എന്നു വരാവുന്നതാണു്. എല്ലാ പ്രാവശ്യം അളക്കുമ്പോഴും ഒരേ അളവു തന്നെയാണു്

കിട്ടുന്നതെങ്കിൽ അളവു സൂക്ഷ്മമാണെന്നു നാം പറയും. അവ തമ്മിലുള്ള വ്യത്യാസം വർധിക്കുന്തോറും അളവിന്റെ സൂക്ഷ്മത കുറയുന്നതായി കരുതപ്പെടും. ഇതാണ് സൂക്ഷ്മതയെന്നതു കൊണ്ട് അർഥമാക്കുന്നത്. അളന്നു കിട്ടുന്ന ദൂരം അൻപതു മീററാണെങ്കിൽ അളവ് യഥാർഥമാണെന്നു പറയാം. ഉപയോഗിച്ച സ്കെയിലിന്റെ നീളം കൃത്യം ഒരു മീററല്ലെങ്കിൽ അളവ് യഥാർഥമായിരിക്കുകയില്ല. സൂക്ഷ്മമായിരുന്നെന്നു വരാം താനും. സാധാരണ ഭാഷയിൽ അംഗീകരിച്ചിരിക്കുന്ന സൂക്ഷ്മത, യാഥാർഥ്യം എന്നിവയുടെ അർഥം ഇതാണല്ലോ. സാമ്പിളനത്തിലും ഏതാണ്ട് ഇതോടു ബന്ധപ്പെട്ട ഒരു അർഥമാണ് ഈ വാക്കുകൾക്കുള്ളത്.

θ എന്ന പ്രാചലത്തിന്റെ ആകലമായി t എന്ന സാമ്പിളജം നിർദേശിക്കപ്പെട്ടിരിക്കുന്നു എന്നു കരുതുക. വിവിധ സാമ്പിളുകളിൽ നിന്നു ഗണിച്ചു കിട്ടുന്ന 't' യുടെ മൂല്യങ്ങൾ വളരെ വ്യത്യാസമില്ലാത്ത സംഖ്യകളാണെങ്കിൽ സൂക്ഷ്മതയുള്ള ഒരു ആകലമാണ് 't' എന്നു പറയും. മറെറാരു തരത്തിൽ പറഞ്ഞാൽ, t യുടെ മൂല്യങ്ങൾ $E(t)$ യുടെ പരിസരത്തിൽ എത്ര മാത്രം കേന്ദ്രീകരിക്കുന്നുവോ അത്രമാത്രം സൂക്ഷ്മതയുണ്ട് ആ ആകലത്തിന്. t യുടെ സാമ്പിളനപ്രസരണം നോർമലോ, നോർമലിനോട് അടുത്തതോ ആണെങ്കിൽ (മിക്കവാറും അങ്ങിനെയായിരിക്കുമല്ലൊ) ഈ കേന്ദ്രീകരണം $1/V(t)$ ക്ക് ആനുപാതികമായിരിക്കും. അങ്ങനെ t യുടെ സൂക്ഷ്മത അളക്കാനുള്ള മാനദണ്ഡം $1/V(t)$ ആണെന്നു സിദ്ധിക്കുന്നു. വിവിധ സാമ്പിളുകളിൽ നിന്നു ലഭിക്കുന്ന 't' യുടെ മൂല്യങ്ങൾ θ യുടെ യഥാർഥമൂല്യത്തിന്റെ പരിസരത്തിൽ എത്രമാത്രം കേന്ദ്രീകരിക്കുന്നുവോ അത്രമാത്രം യാഥാർഥ്യമുള്ള ആകലമാണ് t എന്നു പറയും. ഇതളക്കാനുള്ള മാനദണ്ഡം സ്വാഭാവികമായും $1/M(t)$ ആണ്. $E(t) = \theta$ യാണെങ്കിൽ സൂക്ഷ്മതയും യാഥാർഥ്യവും ഒന്നു തന്നെയാവും.

13. സംഗ്രഹം

1. സ്ഥിതിവിവരശേഖരണത്തിന്റെ പ്രാധാന്യം കൂടുതൽ കൂടുതൽ അംഗീകരിക്കപ്പെട്ടു വരികയാണ്. ഗവണ്മെൻറ് വ്യവസായ വാണിജ്യസ്ഥാപനങ്ങൾ, സാമൂഹ്യക്ഷേമസംഘടനകൾ, ഗവേഷണസ്ഥാപനങ്ങൾ മുതലായവയെല്ലാം സ്ഥിതിവിവരശേഖരണാപഗ്രഥനങ്ങളിൽ തല്പരരത്രെ.

2. സ്ഥിരമായിട്ടുള്ളവ, കാലമനുസരിച്ച് മാറുന്നവ എന്നീ രണ്ടു വിഭാഗങ്ങളായി സ്ഥിതിവിവരങ്ങളെ വിഭജിക്കാം. ഇതിൽ രണ്ടാമത്തെ തരത്തിൽ പെട്ട വിവരങ്ങൾ ശേഖരിക്കാനാണ് ബുദ്ധിമുട്ട്.

3. സ്ഥിതിവിവരശേഖരണം പ്രഥമദൃഷ്ട്യാ തോന്നുന്നത്ര ലളിതമായ ഒരു പ്രക്രിയയല്ല. ആസൂത്രണം, വിവരശേഖരണം, അപഗ്രഥനം എന്നീ മൂന്നു ഘടകങ്ങളിലും ഗൗരവതരമായ പല പ്രശ്നങ്ങളും പ്രതിബന്ധങ്ങളും അഭിമുഖീകരിക്കേണ്ടി വരും.

4. ഒന്നോ അതിലധികമോ അഭിലക്ഷണങ്ങളെപ്പററിയുള്ള വിവരങ്ങൾ ശേഖരിക്കേണ്ടതുണ്ടാവും. ആരെ, അല്ലെങ്കിൽ എന്തിനെ, പററിയാണോ വിവരം ശേഖരിക്കുന്നത് അതാണ് വ്യക്തി. വ്യക്തികളുടെ സമുച്ചയമാണ് സമഷ്ടി.

5. എല്ലാ വ്യക്തികളിൽ നിന്നും വിവരം ശേഖരിക്കുന്ന സമ്പ്രദായത്തിന് സെൻസസ് എന്നും പ്രാതിനിധ്യസ്വഭാവമുള്ള ഏതാനും വ്യക്തികളെ സമഷ്ടിയിൽ നിന്നു തിരഞ്ഞെടുത്ത് അവയിൽ നിന്നു മാത്രമാണ് വിവരങ്ങൾ ശേഖരിക്കുന്നതെങ്കിൽ അതിന് സാമ്പിളനം എന്നും പറയും. ഓരോന്നിനും അതിന്റേതായ ഗുണങ്ങളും ദോഷങ്ങളുമുണ്ട്.

6. ലഭ്യമായ വിഭവശേഷി ഏററവും പ്രയോജനപ്രദമായി വിനിയോഗിച്ച് സാമ്പിൾ തിരഞ്ഞെടുക്കൽ, വിവരശേഖരണം, അപഗ്രഥനം എന്നിവ എങ്ങിനെ നിർവഹിക്കാമെന്നതാണ് സാമ്പിളനശാസ്ത്രത്തിന്റെ ചർച്ചാവിഷയം.

7. പ്രാചലങ്ങളെ സാമ്പിളജങ്ങളുപയോഗിച്ചാണ് ആകലനം ചെയ്യുന്നത്. വലിയ സാമ്പിളുകളിൽ നിന്നു ലഭിക്കുന്ന സാമ്പിളജങ്ങൾ മിക്കവാറും നോർമൽ വിതരണം അനുവർത്തിക്കും. അവയുടെ സാമ്പിളന പ്രസരണം ആകലത്തെപ്പററി വിലപ്പെട്ട പല വിവരങ്ങളും നൽകും.

8. സാമ്പിളനം രണ്ടുവിധം— സംഭാവ്യതാ സാമ്പിളനം, ഐച്ഛിക സാമ്പിളനം.

9. ആകലങ്ങളെ താരതമ്യപ്പെടുത്താൻ അവയുടെ അഭിനതി, അവിരോധിതത്വം തുടങ്ങിയവ പരിഗണിക്കേണ്ടതാണ്.

10. t ആണ് θ യുടെ ആകലമെങ്കിൽ $M(t) = E(t - \theta)^2$ ത്തെ അതിന്റെ പിശകുവർഗ മാധ്യമെന്നും $V(t) = E[t - E(t)^2]$ ത്തെ അതിന്റെ സാമ്പിളൽ പ്രസരണമെന്നും പറയുന്നു. ഇതിൽ $M(t)$ ആകലത്തിന്റെ യാഥാർഥ്യത്തിന്റെയും $V(t)$ സൂക്ഷ്മതയുടെയും നിദർശകങ്ങളാണ്.

11. സാംഖ്യിക അനുമാനം സാമ്പിളനഫലങ്ങൾ അപഗ്രഥനം ചെയ്യാൻ പ്രയോജനപ്രദമായി ഉപയോഗിക്കാം. ആകലനം, പരീക്ഷ, വിശ്വാസ്യതാന്തരാളങ്ങളുടെ നിർണയം തുടങ്ങിയവ സാമ്പിളനത്തിൽ സുപ്രധാനങ്ങളാണല്ലോ.

ചില സഹായകഗ്രന്ഥങ്ങളും ലേഖനങ്ങളും

1. Sampling in Nutshell (1960) , Morris James Slonim Simon and Schuster, New York.
2. Sampling methods for Census and Surveys (1953) : Yates F, Charles Griffin & Co., London.
3. The History and uses of Modern Sampling procedures (1948) : Journal of the American Statistical Association 43 (12—39)
4. Sampling Theory (1968) : Des Raj (ch. 11) McGraw Hill Book Company.
5. Sampling Theory and Methods (1967) : M. N. Moorthy, Statistical Publishing House, Calcutta.

അഭ്യാസം-1

1. സ്ഥിതി വിവരശേഖരണത്തിൻെറ പ്രാധാന്യം ചർച്ച ചെയ്യുക.
2. സെൻസസ്സും സാമ്പിളനവും താരതമ്യപ്പെടുത്തി ഓരോന്നും കൂടുതൽ പ്രയോജനപ്രദമായേക്കാവുന്ന സന്ദർഭങ്ങൾ ഉദാഹരണ സഹിതം വിവരിക്കുക.
3. സെൻസസ് രീതി അവലംബിക്കാൻ ആവശ്യമായ പണവും മററു സാഹചര്യങ്ങളും ഇല്ലാതെ വന്നാൽ സാമ്പിളനത്തെ ശരണംപ്രാപിക്കുകയേ നിർവാഹമുള്ളൂ. ചർച്ച ചെയ്യുക.
4. സാമ്പിളനത്തിൽ നോർമൽ വിതരണത്തിൻെറ പ്രാധാന്യം വിശദമാക്കുക.
5. സാമ്പിളനത്തിൽ സംഭവിക്കാവുന്ന വിവിധതരം പിശകുകളെ പററി ഉപന്യസിക്കുക.
6. താഴെ പറയുന്ന ആവശ്യങ്ങൾക്ക് വേണ്ടിയുള്ള അന്വേഷണത്തിൽ സെൻസസാണോ സാമ്പിളനമാണോ മെച്ചം? എന്തു കൊണ്ട്?
   (a) ഒരു നഗരത്തിലെ ക്ഷയരോഗികളുടെ അനുപാതം നിർണയിക്കുക.
   (b) വിവിധ വർത്തമാനപത്രങ്ങളോട് വായനക്കാരുടെ മനോഭാവം പരിശോധിക്കുക.
   (c) നികുതി ചുമത്തുന്നതിനു വേണ്ടി 1000 രൂപയിലധികം വരുമാനമുള്ളവരുടെ ലിസ്റ്റ് തയ്യാറാക്കുക.

2

സാമ്പിളനത്തിൻെറ അണിയറ

1. പ്രാരംഭം

സ്ഥിതിവിവരശേഖരണം വളരെ പ്രയാസമേറിയ ഒരു ജോലിയാണ്. സെൻസസായാലും സാമ്പിളനമായാലും ബുദ്ധിമുട്ടിനു് കുറവൊന്നുമില്ല. ഓരോന്നിലും പ്രത്യേകം പരിഗണിക്കേണ്ട വിഷയങ്ങൾ ഓരോന്നായിരിക്കുമെന്നു മാത്രം. ശ്രദ്ധാപൂർവമുള്ള ആസൂത്രണം രണ്ടിനും ആവശ്യമാണ്. ശരിയായ ആസൂത്രണം കൂടാതെ നടത്തുന്ന ഏതൊരു സ്ഥിതിവിവരശേഖരണപരിശ്രമവും പരാജയപ്പെടാനാണ് സാധ്യത. പണവും അധ്വാനവും സമയവും വേണ്ടതിലധികം ചെലവഴിക്കേണ്ടതായി വരാം. ശേഖരിച്ച വിവരങ്ങൾ ആവശ്യത്തിനു് ഉതകിയില്ലെന്നും വന്നേക്കാം. ചില പ്രത്യേകപ്രശ്നങ്ങൾക്കു പരിഹാരം കാണുവാൻ വേണ്ടിയായിരിക്കുമല്ലൊ വ്യക്തികളോ സ്ഥാപനങ്ങളോ സ്ഥിതിവിവരശേഖരണത്തെപ്പററി ചിന്തിക്കുന്നതു്. സമർത്ഥനും പരിചയസമ്പന്നനുമായ ഒരു സാംഖ്യികജ്ഞൻെറ സേവനം ഉപയോഗപ്പെടുത്തിയാൽ കുറഞ്ഞ സമയം കൊണ്ടു് കുറഞ്ഞ ചെലവിൽ ആവശ്യമായ സൂക്ഷ്മതയോടു കൂടി വിവരങ്ങൾ ശേഖരിക്കാൻ കഴിയും.

ആസൂത്രണം, ശേഖരണം, അപഗ്രഥനം എന്നീ മൂന്നു ഘട്ടങ്ങൾ ഏതൊരു സ്ഥിതിവിവരാന്വേഷണത്തിലുമുണ്ടു്. ഇവ ഒററപ്പെട്ടവയല്ല. ശേഖരണരീതിയോടു ബന്ധപ്പെട്ടതായിരിക്കും അപഗ്രഥനപദ്ധതി. ഇവ രണ്ടും ആസൂത്രണത്തിൽ പരിഗണിക്കുകയും വേണം. ശേഖരണഘട്ടത്തിൽ പൊന്തിവന്നേക്കാവുന്ന പ്രശ്നങ്ങൾ പരിഗണിച്ച് അവ പരിഹരിക്കത്തക്കവണ്ണം ആസൂത്രണം നിർവഹിക്കണം. അതു പോലെ തന്നെ അപഗ്രഥനത്തിൽ അംഗീകരിക്കുന്ന പരിപാടികളും തത്വങ്ങളും ആസൂത്രണത്തിനു് മാർഗദർശകങ്ങളാണ്. ചുരുക്കത്തിൽ ഈ മൂന്നു ഘട്ടങ്ങളും ഒന്നിച്ചു് പരിഗണിക്കേണ്ടവയാണ്. ചർച്ച ചെയ്യാനുള്ള സൗകര്യത്തിനു വേണ്ടി മാത്രമാണ് ഈ വിഭജനം.

സാമ്പിളനത്തിൻെറ ആസൂത്രണവും നിർവഹണവുമാണ് ഈ അധ്യായത്തിലെ ചർച്ചാവിഷയം. ഇവിടെ പരിഗണിക്കുന്ന കാര്യങ്ങൾ സെൻസസി

ല്യം ഒരളവു വരെ പ്രസക്തങ്ങളത്രെ. പൊതുവായ ചില തത്വങ്ങൾ നിർദേശിക്കാനാണ് പ്രധാനമായും ശ്രദ്ധിച്ചിട്ടുള്ളത്. സന്ദർഭത്തിനനുസരിച്ചും യുക്തിപൂർവമായും ഇതു പ്രായോഗികമാക്കുന്നതിലാണ് സ്ഥിതിവിവരാന്വേഷകന്റെ സാമർഥ്യവും സ്ഥിതിവിവരാന്വേഷണത്തിന്റെ വിജയവും നിലകൊള്ളുന്നത്. പ്രായോഗികബുദ്ധിയും പരിചയസമ്പത്തും ഇമ്മാതിരി സന്ദർഭങ്ങളിൽ കണ്ടെപ്പെട്ട കൈമുതലുകളായിരിക്കും.

2. ലക്ഷ്യബോധം

ലക്ഷ്യബോധം ആസൂത്രണഘട്ടത്തിൽ അതിപ്രധാനമായ ഒന്നാണ്. സ്ഥിതിവിവരങ്ങൾ ആർക്കുവേണ്ടിയാണോ ശേഖരിക്കുന്നത് അവർക്ക് അവരുടെതായ ചില പ്രശ്നങ്ങളുണ്ടായിരിക്കും. ആ പ്രശ്നങ്ങൾ പരിഹരിക്കാൻ ചില സ്ഥിതിവിവരങ്ങൾ ശേഖരിച്ച് അപഗ്രഥിച്ചു നോക്കുന്നത് ഉപകാരപ്രദമായേക്കും എന്ന പ്രതീക്ഷയിലായിരിക്കും അവർ അതിനായി മുന്നിട്ടിറങ്ങുന്നത്. ആ പ്രശ്നങ്ങളുടെ സങ്കീർണത കാരണം പലപ്പോഴും സ്ഥിതിവിവരാന്വേഷണത്തിന്റെ മൗലികലക്ഷണങ്ങളെപ്പററിയുള്ള അവരുടെ ധാരണ അപൂർണവും അവ്യക്തവുമായിരിക്കാനാണ് സാധ്യത. ഈ വസ്തുത, അന്വേഷണം ആസൂത്രണം ചെയ്യുന്ന സാംഖ്യികജ്ഞൻ കണക്കിലെടുക്കണം. അവരുടെ പ്രശ്നങ്ങൾ ശരിക്കു മനസ്സിലാക്കാൻ ശ്രമിക്കണം. ആവശ്യമായി വന്നാൽ നീണ്ട ചർച്ചകളിൽ പങ്കെടുക്കണം. വിവിധതരം സ്ഥിതിവിവരങ്ങൾ ശേഖരിക്കുന്നതു കൊണ്ട് ഉണ്ടാകാവുന്ന പ്രയോജനങ്ങൾ വിശദീകരിച്ചു കൊടുക്കണം. അവർക്കു പറയാനുള്ള കാര്യങ്ങൾ ശ്രദ്ധിച്ചു കേൾക്കുകയും അതിലുള്ള അവ്യക്തതകൾ തുറന്നു കാണിക്കുകയും ചെയ്യണം. അന്വേഷണലക്ഷ്യങ്ങളെപ്പററി വ്യക്തവും സൂക്ഷ്മവുമായ ധാരണ ഉടലെടുക്കുന്നതു വരെ ഈ പ്രക്രിയ തുടരേണ്ടിവരും. അങ്ങനെ എത്തിച്ചേരുന്ന നിഗമനങ്ങളും അതോടു ബന്ധപ്പെട്ട മററു പ്രധാന കാര്യങ്ങളും ഒരു പ്രസ്താവനയുടെ രൂപത്തിൽ എഴുതിയുണ്ടാക്കി അതിനു് അംഗീകാരം നേടേണ്ടതു് അന്വേഷണം സംഘടിപ്പിക്കുന്ന സാംഖ്യികജ്ഞന്റെ സുഗമമായ പ്രവർത്തനത്തിനു് ആവശ്യമാണ്. അന്വേഷണം ആസൂത്രണം ചെയ്യുമ്പോൾ ആ പ്രസ്താവന അത്യന്തം ഉപകാരപ്രദമായിരിക്കുകയും ചെയ്യും. ഈ പ്രസ്താവനയിൽ താഴെപ്പറയുന്ന കാര്യങ്ങൾ തീർച്ചയായും ഉൾപ്പെട്ടിരിക്കണം.

1. പഠനം നടത്തേണ്ട സമഷ്ടിയുടെ വ്യക്തമായ നിർവചനം
2. ശേഖരിക്കേണ്ട വിവരങ്ങൾ ഉപയോഗിക്കേണ്ട ഏകകങ്ങൾ എന്നിവയ്ക്കുണ്ടായിരിക്കേണ്ട സൂക്ഷ്മത
3. വിവരശേഖരണത്തിനു് അനുവദിക്കാവുന്ന കാലയളവ്
4. സാമ്പത്തികപരിമിതി
5. അപഗ്രഥനം പൂർത്തിയാവുന്നതിനു മുൻപു തന്നെ ഏതെങ്കിലും പ്രത്യേക വിഷയങ്ങളെപ്പററി പ്രാഥമിക റിപ്പോർട്ടുകൾ ആവശ്യമുണ്ടോ? ഉണ്ടെങ്കിൽ അവ ഏതെല്ലാം?
6. സമഷ്ടിയുടെ ഏതെല്ലാം സ്വഭാവങ്ങളിലാണ് പ്രത്യേക താല്പര്യമുള്ളത്?

എങ്ങനെയുള്ള സാരണികൾ തയ്യാറാക്കുന്നതായിരിക്കും കൂടുതൽ ഉപകാരപ്രദം?

7. ആകലനം ചെയ്യേണ്ട പ്രാചലങ്ങളും ആകലങ്ങൾക്ക് ഉണ്ടായിരിക്കേണ്ട സൂക്ഷ്മതയും.

പ്രസ്താവനയിൽ ഉൾക്കൊള്ളിക്കേണ്ട അതിപ്രധാനങ്ങളായ ചില കാര്യങ്ങളാണ് മുകളിൽ കൊടുത്തിരിക്കുന്നതു്. ഓരോ അന്വേഷണത്തിനും അതിൻെറതായ ചില പ്രത്യേകതകൾ ഉണ്ടായിരിക്കും. എന്ന വസ്തുത കൂടി കണക്കിലെടുക്കണം. അന്വേഷണത്തെ സംബന്ധിച്ചിടത്തോളം പൂർണവും വ്യക്തവുമായ മാർഗദർശനം ലഭിക്കത്തക്കവണ്ണം വളരെ ശ്രദ്ധിച്ചു തയ്യാറാക്കേണ്ട ഒന്നാണ് ഈ പ്രസ്താവന. വിവരങ്ങൾ ആവശ്യമുള്ള വ്യക്തികളോ സ്ഥാപനങ്ങളോ വളരെ മനസ്സിരുത്തി പരിശോധിച്ച് അംഗീകരിക്കേണ്ട ഒരു സുപ്രധാന രേഖയുമാണിതു്. അന്വേഷണം കുറെ പുരോഗമിച്ചതിനു ശേഷം ഈ കാര്യങ്ങളിൽ ഏതെങ്കിലും മാററം വരുത്തുക എളുപ്പമായിരിക്കുകയില്ല എന്നതു് ഓർമയിരിക്കണം. അങ്ങനെ മാററം വരുത്തിയാൽ വളരെയേറെ സമയനഷ്ടവും പാഴ്ച്ചെലവുകളും വന്നുചേരാവുന്നതാണ്. അതു പോലെ തന്നെ ഈ പ്രസ്താവനയിൽ കടന്നുകൂടിയേക്കാവുന്ന അവ്യക്തതകളും അതു് തയ്യാറാക്കുന്നതിലെ അനവധാനതയും അന്വേഷണഘട്ടത്തിലും അപഗ്രഥനഘട്ടത്തിലും പലവിധ ബുദ്ധിമുട്ടുകൾക്കും കാരണമായി തീരാം. ആവശ്യമില്ലാത്ത പല വിവരങ്ങളും ശേഖരിക്കുകയും അതിപ്രധാനങ്ങളായ വിവരങ്ങൾ ശേഖരിക്കുന്നതിൽ വേണ്ടത്ര ശ്രദ്ധ പതിപ്പിക്കാതിരിക്കുകയും ചെയ്യാൻ ഇതു കാരണമാവാം. ആകലങ്ങളുടെ സൂക്ഷ്മത ആവശ്യമുള്ളത്ര ഉയർന്നതല്ലെന്ന് അവസാനഘട്ടത്തിൽ വ്യക്തമായേക്കാം. അപഗ്രഥനഫലമായി എത്തിച്ചേരുന്ന നിഗമനങ്ങൾ അന്വേഷണത്തിൻെറ മൗലികലക്ഷ്യങ്ങൾ നേടുന്നതിൽ വേണ്ടത്ര പ്രയോജനപ്രദമല്ലെന്നു വന്നു ചേരാവുന്നതാണ്. ഇതു പോലെ ഗൗരവതരങ്ങളായ പല തകരാറുകൾക്കും മുൻപറഞ്ഞ പ്രസ്താവനയിലെ അപാകതകൾ കാരണമാകാമെന്നതു കൊണ്ടു് അതു തയ്യാറാക്കുന്നതിൽ അങ്ങേയററത്തെ ശ്രദ്ധയും ജാഗ്രൂകതയും അനുപേക്ഷണീയമാണ്.

അന്വേഷണലക്ഷ്യങ്ങൾക്ക് രൂപം കൊടുക്കുന്നതിൽ പരിഗണിക്കേണ്ട മറ്റൊരു പ്രധാന വസ്തുത, അന്വേഷകനു് ആ ലക്ഷ്യങ്ങൾ നേടുവാനുള്ള കഴിവും സൗകര്യവുമുണ്ടോ എന്നതാണ്. ഒരു നഗരത്തിൽ പ്രതിദിനം വില്പന നടത്തപ്പെടുന്ന വിദേശമദ്യങ്ങളുടെ ശരാശരി അളവു് നിർണയിക്കുകയാണ് അന്വേഷണലക്ഷ്യമെന്നിരിക്കട്ടെ. വിദേശമദ്യങ്ങൾ വില്പന നടത്തുന്ന എല്ലാ വ്യാപാരസ്ഥാപനങ്ങളിൽ നിന്നോ, അവയിൽ നിന്നു് തിരഞ്ഞെടുത്ത ഏതാനും സ്ഥാപനങ്ങളിൽ നിന്നോ വിവരം ശേഖരിക്കേണ്ടതുണ്ടു്. വ്യാപാരത്തെ സംബന്ധിച്ച പൂർണവും സൂക്ഷ്മവുമായ കണക്കുകൾ സൂക്ഷിക്കുകയും അതു് അന്വേഷകനു് നൽകുകയും ചെയ്യുവാൻ വ്യാപാരസ്ഥാപനം തയ്യാറായെങ്കിലേ അന്വേഷണലക്ഷ്യം സാക്ഷാൽകരിക്കപ്പെടുകയുള്ളൂ. ഭരണകൂടമാണ് അന്വേഷണം സംഘടിപ്പിക്കുന്നതെങ്കിൽ അധികാരശക്തി ഉപയോഗിച്ചു് ഒരളവു വരെയെങ്കിലും വ്യാപാരസ്ഥാപനങ്ങളുടെ സഹകരണം നേടിയെടുക്കാൻ കഴിഞ്ഞേക്കും.

പക്ഷേ, അന്വേഷണം അർധ ഔദ്യോഗികം മാത്രമാണെങ്കിൽ സഹകരണത്തിനു വേണ്ടി "അപേക്ഷിക്കു"ക മാത്രമേ കരണീയമായിട്ടുള്ളു. സ്വകാര്യവ്യക്തികൾ സംഘടിപ്പിക്കുന്ന അന്വേഷണത്തിൻെറ സ്ഥിതി ഇതിലും ദയനീയമാണ്. സഹകരണത്തിനു വേണ്ടിയുള്ള "യാചന"യുമായി അവക്കു വ്യാപാരസ്ഥാപനങ്ങളെ സമീപിക്കേണ്ടി വരും. ഇമ്മാതിരി സാഹചര്യങ്ങൾ ലഭിക്കുന്ന വിവരങ്ങളുടെ സൂക്ഷ്മതയെയും ബാധിക്കുമെന്നുള്ളതു് വ്യക്തമാണല്ലോ. വിവരങ്ങൾക്കായി "ആജ്ഞാപിക്കുക"യും "അപേക്ഷിക്കുക"യും "യാചിക്കുക"യും ചെയ്യുന്നവരോടെല്ലാം ഒരേ പ്രതികരണമായിരിക്കുകയില്ല ദാതാവിനു് ഉണ്ടാവുക. ചുരുക്കത്തിൽ ഓരോ വിഭാഗത്തിൽപ്പെട്ട അന്വേഷകരും അവരവരുടെ പരിമിതികൾ പ്രാരംഭത്തിൽ തന്നെ കണക്കിലെടുക്കണം. അന്വേഷണലക്ഷ്യങ്ങളിൽ അതിനനുസരിച്ച മാററങ്ങൾ ആദ്യം തന്നെ വരുത്തുകയും വേണം.

3. സമഷ്ടിയുടെ നിർവചനം

പ്രാരംഭത്തിൽ തയ്യാറാക്കുന്ന പ്രസ്താവനയിലെ ആദ്യത്തെ ഇനം സമഷ്ടിയുടെ നിർവചനമാണല്ലോ. ഇതു് സമഗ്രമായ പഠനവും ചർച്ചയും അർഹിക്കുന്ന ഒന്നായതു കൊണ്ടു് അതു് പ്രത്യേകം പരിഗണിക്കുകയാണു്.

വ്യക്തികളുടെ സമുച്ചയമാണു് സമഷ്ടി. അതു കൊണ്ടു് വ്യക്തിയുടെ നിർവചനം സമഷ്ടിയെ സംബന്ധിച്ചിടത്തോളം നിർണായകമാണു്. കേരളത്തിലെ ഇടത്തരക്കാരുടെ ജീവിതച്ചെലവുകളെപ്പററി പഠനം നടത്താൻ ആഗ്രഹിക്കുന്നു എന്നിരിക്കട്ടെ. കേരളത്തിലെ ഇടത്തരക്കാരുടെ സമുച്ചയമാണു് പഠനം നടത്തേണ്ട സമഷ്ടി. "ഇടത്തരക്കാർ" എന്നതു കൊണ്ടു് എന്താണു് അർഥമാക്കുന്നതു് എന്നതാണു് ഇവിടത്തെ പ്രശ്നം. പഠനവിധേയമാക്കേണ്ടതു് കുടുംബങ്ങളെയാണോ വ്യക്തികളെയാണോ എന്നു് ആദ്യം തീരുമാനിക്കണം. കുടുംബങ്ങളെപ്പററിയാണു് പഠിക്കുന്നതു് എന്നു വിചാരിക്കുക. കുടുംബത്തിൻെറ വാർഷിക വരുമാനത്തെ അടിസ്ഥാനമാക്കി "ഇടത്തരം കുടുംബം" എന്താണെന്നു നിർവചിക്കണം. കേരളത്തിലെ ഏതെങ്കിലും ഒരു കുടുംബത്തെപ്പററി ആവശ്യമായ വിവരങ്ങൾ നൽകിയാൽ അതു് "ഇടത്തരം" കുടുംബമാണോ അല്ലയോ എന്നു് സംശയാതീതമായി നിർണയിക്കാൻ പര്യാപ്തമായിരിക്കണം ആ നിർവചനം. ഈ മൗലികലക്ഷ്യം സാക്ഷാൽക്കരിക്കത്തക്കവണ്ണം ഒരു നിർവചനം കണ്ടുപിടിക്കുക അത്ര എളുപ്പമായ ഒരു കാര്യമല്ല. വാർഷികവരുമാനം 1200 രൂപയ്ക്കും 4800രൂപയ്ക്കും ഇടയ്ക്കുള്ള കുടുംബങ്ങളെ ഇടത്തരം കുടുംബങ്ങളായി നിർവചിക്കുന്നു എന്നിരിക്കട്ടെ. പ്രഥമദൃഷ്ട്യാ വലിയ അപകടമൊന്നും തോന്നുകയില്ല. പക്ഷേ കൂടുതൽ ഗാഢമായി ചിന്തിക്കുമ്പോൾ പല അപാകതകളും ശ്രദ്ധയിൽപ്പെടും. കുടുംബത്തിൽപ്പെട്ട ആരുടെയെല്ലാം വരുമാനമാണു് കണക്കിലെടുക്കേണ്ടതു് എന്നതാണു് ഒരു പ്രശ്നം. ദൂരസ്ഥലത്തു് താമസിച്ചുകൊണ്ടു് ചെറിയ തുകകൾ വീട്ടിലേക്കു് അയച്ചുകൊടുക്കുന്ന ആളുകൾ കുടുംബത്തിലുണ്ടാവാം. അവരെ കുടുംബാംഗങ്ങളായി പരിഗണിക്കേണ്ടതുണ്ടോ എന്നു തീരുമാനിക്കണം. അവരെ കുടുംബാംഗങ്ങളായി സ്വീകരിച്ചാൽ അവർ അയച്ചുകൊടുക്കുന്ന തുക മാത്രം വരുമാനത്തിൽ ഉൾപ്പെടുത്തിയാൽ

മതിയോ, അതോ അവരുടെ ആകെ വരുമാനം ഉൾപ്പെടുത്തണമോ എന്നു നിർണ്ണയിക്കണം. കുടുംബത്തിൽ തന്നെ കഴിഞ്ഞുകൂടുന്ന ചിലർ സ്വന്തം വരുമാനത്തിൽ നിന്നു് ഒരു തുക ബന്ധുക്കൾക്കു് കൊടുക്കുന്നുണ്ടാവാം. ആ തുക വരുമാനത്തിൽ ഉൾപ്പെടുത്തണമോ എന്നതു് തീരുമാനമെടുക്കേണ്ട മറ്റൊരു കാര്യമാണു്. ഇങ്ങനെ നിരവധി കാര്യങ്ങൾ നിർവചനത്തിനു് രൂപം കൊടുക്കുമ്പോൾ പരിഗണിക്കേണ്ടതുണ്ടു്. അന്വേഷണലക്ഷ്യങ്ങൾ, അന്വേഷണത്തിനു് അവലംബിക്കുന്ന മാർഗം, അന്വേഷണഫലങ്ങൾ എന്താവശ്യത്തിനാണു് ഉപയോഗപ്പെടുത്തുന്നതു് എന്നു തുടങ്ങി ബന്ധപ്പെട്ട എല്ലാ കാര്യങ്ങളും കണക്കിലെടുത്തു വേണം വ്യക്തിയുടെ നിർവചനം തയ്യാറാക്കാൻ. അതിൽ സംഭവിക്കാവുന്ന അപാകതകൾക്കു് ദൂരവ്യാപകങ്ങളായ പ്രത്യാഘാതങ്ങളുണ്ടു് എന്ന വസ്തുത ഓർമയിലിരിക്കണം.

വ്യക്തികളുടെ നിർവചനം അംഗീകരിച്ചു കഴിയുന്നതോടെ സമഷ്ടിയും നിർവചിക്കപ്പെട്ടു കഴിഞ്ഞിരിക്കും. വ്യക്തികളുടെ ആകത്തുകയാണല്ലൊ സമഷ്ടി. പക്ഷേ പ്രായോഗിക ബുദ്ധിമുട്ടുകൾ മൂലം സമഷ്ടിയിൽ ചില നീക്കുപോക്കുകൾ വരുത്താറുണ്ടു്. കേരളത്തിലെ കൃഷിഭൂമികളുടെ ഉൽപാദനശേഷിയെപ്പററിയുള്ള ഒരു പഠനത്തിൽ എല്ലാ കൃഷിഭൂമികളും, അവ കേരളത്തിൻെറ നാലതിരുകൾക്കുള്ളിലാണെങ്കിൽ, സമഷ്ടിയിലെ അംഗങ്ങളാണു്. എന്നാൽ അന്വേഷണത്തിൻെറ സുഗമമായ നടത്തിപ്പിനു് അതിൽ ചില ഭേദഗതികൾ വരുത്തുന്നതു് സൗകര്യമായിരിക്കും. വനാന്തരങ്ങളിലുള്ള കൃഷിഭൂമികൾ അന്വേഷണവിധേയമാക്കുക അത്ര എളുപ്പമല്ല. അവ ഒഴിവാക്കിയാൽ, ഉൽപാദനശേഷിയെപ്പററി അന്വേഷണത്തിൽ നിന്നു ലഭിക്കുന്ന വിവരങ്ങൾ വനഭൂമികൾക്കു് ബാധകമല്ല എന്നു വന്നു ചേരുകയും ചെയ്യും. അതു് അത്ര അഭിലഷണീയമല്ലല്ലൊ. വനഭൂമികളുടെ ഉൽപാദനശേഷി മററു കൃഷിഭൂമികളുടെതിൽ നിന്നു് വളരെയൊന്നും വ്യത്യസ്തമല്ലെന്നു് കരുതാൻ ന്യായമുണ്ടെങ്കിൽ, വനഭൂമിയെ ഒഴിവാക്കിക്കൊണ്ടു് പഠനം നടത്തിയിട്ടു്, അതിൽ നിന്നു ലഭിക്കുന്ന വിവരങ്ങൾ എല്ലാ കൃഷിഭൂമികളെ സംബന്ധിച്ചും ശരിയാണു് എന്നു് അംഗീകരിക്കുന്നതിൽ വലിയ അപാകതയില്ല.

പഠനവിധേയമാക്കുന്ന സമഷ്ടിക്കു് പഠിതസമഷ്ടി എന്നും എത്തിച്ചേരുന്ന നിഗമനങ്ങൾ ഏതൊരു സമഷ്ടിയെ സംബന്ധിച്ചതായി അംഗീകരിക്കുന്നുവോ അതിനു് ലക്ഷ്യസമഷ്ടി എന്നും പേർ പറയുന്നു. പഠിതസമഷ്ടിയും ലക്ഷ്യസമഷ്ടിയും ഒന്നു തന്നെ ആയിരിക്കുന്നതാണു് അഭികാമ്യം. പഠിതസമഷ്ടിയിൽ ഉൾപ്പെടാത്ത വ്യക്തികൾക്കു് അതിൽ ഉൾപ്പെട്ടിട്ടുള്ള വ്യക്തികളുമായി പരിഗണിക്കപ്പെടുന്ന അഭിലക്ഷണത്തെ സംബന്ധിച്ചിടത്തോളം വലിയ മാററമില്ലെന്നു് വിശ്വസിക്കാവുന്ന സാഹചര്യങ്ങളിലാണു്, പഠിതസമഷ്ടി ലക്ഷ്യസമഷ്ടിയിൽ നിന്നു് വ്യത്യസ്തമായിരിക്കാൻ അനുവദിക്കാറുള്ളതു്. അതതു പ്രശ്നങ്ങളുടെ പ്രത്യേകതകളും സമഷ്ടിയിലെ വൈവിധ്യവുമെല്ലാം കണക്കിലെടുത്തു വേണം ഇക്കാര്യത്തിൽ തീരുമാനമെടുക്കാൻ. പ്രായോഗികസൗകര്യങ്ങളും സാമ്പത്തികപരിഗണനകളും നിർണായകഘടകങ്ങളായിരിക്കുമെന്നുള്ളതും വിസ്മരിക്കാവുന്നതല്ല.

4. പ്രാഥമികമോ ദ്വിതീയകമോ

സമഷ്ടി ഏതാണെന്നും ഏതെല്ലാം വിവരങ്ങളാണ് ശേഖരിക്കേണ്ടതെന്നും തീരുമാനിച്ചു എന്നിരിക്കട്ടെ. അടുത്ത പടി അന്വേഷണം ആസൂത്രണം ചെയ്യുകയും അതിനനുസരിച്ച് സ്ഥിതിവിവരങ്ങൾ ശേഖരിക്കുകയുമാണ്. ഇപ്രകാരം ശേഖരിക്കുന്ന വിവരങ്ങളാണ് പ്രാഥമികസ്ഥിതിവിവരങ്ങൾ. പലപ്പോഴും ഇതേ സമഷ്ടിയിൽ നിന്നു തന്നെ ഏതാണ്ട് ഇതേ സ്വഭാവത്തിലുള്ള വിവരങ്ങൾ മററു സ്ഥാപനങ്ങളോ വ്യക്തികളോ നേരത്തെ തന്നെ ശേഖരിച്ചിട്ടുണ്ടാവാം. അങ്ങനെ ആരെങ്കിലും ചെയ്തിട്ടുണ്ടോ എന്ന് അന്വേഷിക്കുകയും ഉണ്ടെങ്കിൽ ആ വിവരങ്ങൾ പരിശോധിച്ച് അതു് സ്വന്തം ആവശ്യത്തിനു് മതിയാവുമോ എന്നു് തീരുമാനിക്കുകയും ചെയ്യണം. ഇങ്ങനെ ലഭിക്കുന്ന വിവരങ്ങൾക്കു് ദ്വിതീയകസ്ഥിതിവിവരങ്ങൾ എന്നു പറയും. ദ്വിതീയക സ്ഥിതിവിവരങ്ങൾ ലഭ്യമാണെങ്കിൽ കഴിവതും അതു് ഉപയോഗപ്പെടുത്താൻ സ്ഥിതിവിവരാന്വേഷകർ ശ്രമിക്കാറുണ്ട്. വിവരശേഖരണപ്രക്രിയയുടെ ബുദ്ധിമുട്ടുകളും പണച്ചെലവും ഒഴിവാക്കാമല്ലോ എന്നതു കൊണ്ട്. പക്ഷേ ഇമ്മാതിരി സ്ഥിതിവിവരങ്ങൾ സമഗ്രമായ പരിശോധനയ്ക്കു ശേഷമേ സ്വീകരിക്കാവൂ എന്നതു് ഓർമയിരിക്കണം.

ദ്വിതീയകസ്ഥിതിവിവരങ്ങൾ പരിശോധിക്കുമ്പോൾ പ്രത്യേകം ശ്രദ്ധിക്കേണ്ട ചില കാര്യങ്ങളുണ്ട്. ഒന്നാമതായി ആരാണ് അഥവാ ഏതു സ്ഥാപനമാണ് അവ ശേഖരിച്ചതു് എന്നു മനസ്സിലാക്കണം. വിവരങ്ങളുടെ വിശ്വാസ്യതയെപ്പററി പൊതുവായ ഒരു ധാരണ ഉണ്ടാവാൻ അതു സഹായിക്കും. ഉദാഹരണമായി പറഞ്ഞാൽ ഗവേഷണപ്രവത്തനത്തിൽ ഏർപ്പെട്ടിട്ടുള്ള വ്യക്തികളോ സ്ഥാപനങ്ങളോ ശേഖരിച്ചിട്ടുള്ള വിവരങ്ങളോളം സൂക്ഷ്മമോ പൂർണമോ ആയിരിക്കുകയില്ല ഗവണ്മെന്റ് തലത്തിൽ ശേഖരിച്ചിട്ടുള്ള വിവരങ്ങൾ. സാമ്പത്തികസ്ഥിതിയെപ്പററിയും മററുമുള്ള വിവരങ്ങൾ ഗവണ്മെൻറുദ്യോഗസ്ഥന്മാരോട് തുറന്നു പറയാൻ പലരും മടി കാണിക്കുമല്ലോ. അതു പോലെ തന്നെ മററു ജോലിത്തിരക്കുകളുള്ള റവന്യൂജീവനക്കാരെയും മററും ഉപയോഗിച്ച് വിവരശേഖരണം നടത്തുന്നതു കൊണ്ടു് ഗവണ്മെൻറു ശേഖരിക്കുന്ന വിവരങ്ങളുടെ വിശ്വാസ്യത കുറവായിരിക്കാൻ സാധ്യതയുണ്ട താനും. ആരാണ് വിവരങ്ങൾ ശേഖരിച്ചതെന്ന് കൃത്യമായി അറിഞ്ഞുകൂടാതെ വരുന്ന സന്ദർഭങ്ങളിൽ സമഷ്ടിയുടെ ഏതാനും ഭാഗത്തു നിന്നു് സ്വയം വിവരം ശേഖരിച്ചു ലഭ്യമായ വിവരങ്ങളുമായി താരതമ്യപ്പെടുത്തി അതിൻെറ വിശ്വാസ്യത നിർണയിക്കാൻ കഴിയും. ദ്വിതീയക സ്ഥിതിവിവരങ്ങൾ പരിശോധിക്കുമ്പോൾ ശ്രദ്ധിക്കേണ്ട മറെറാരു പ്രധാന കാര്യം വിവരങ്ങൾ ശേഖരിച്ച ആളിനു് നിഷ്പക്ഷമായ ഒരു സമീപനരീതിയാണോ ഉണ്ടായിരുന്നതു് എന്നതാണ്. ഒരു വ്യവസായസ്ഥാപനം അവരുടെ ഉൽപ്പന്നങ്ങളുടെ പ്രചാരത്തെപ്പററി നടത്തിയ അന്വേഷണഫലങ്ങൾ പരസ്യത്തിനു് ഉപയോഗിക്കാൻ ഉദ്ദേശിക്കപ്പെട്ടിട്ടുള്ളവയാണെങ്കിൽ പ്രത്യേകിച്ചും. അങ്ങനെ മുൻവിധിയോടു കൂടിയോ പ്രചരണലക്ഷ്യങ്ങൾ മനസ്സിൽ വച്ചുകൊണ്ടോ നടത്തി

യിട്ടുള്ള അന്വേഷണങ്ങളിൽ നിന്നു ലഭിച്ചിട്ടുള്ള വിവരങ്ങൾ ഒട്ടും തന്നെ സ്വീകാര്യമല്ല. നാം പരിശോധിക്കുന്ന ദ്വിതീയകസ്ഥിതിവിവരങ്ങൾ നിഷ്പക്ഷമായ അന്വേഷണത്തിൽ നിന്നു ലഭിച്ചവയാണെന്നും അവ അപൂർണമോ അവിശ്വസനീയമോ ആണെന്നു കരുതാൻ ന്യായമില്ലെന്നും കണ്ടാൽ താഴെ പറയുന്ന സുപ്രധാനങ്ങളായ ചില കാര്യങ്ങൾ കൂടി പരിഗണിച്ചതിനു ശേഷം അവ സ്വീകരിക്കാവുന്നതാണ്.

i) *അന്വേഷണലക്ഷ്യം*

അന്വേഷണലക്ഷ്യങ്ങളിലുള്ള മാററം വളരെ ഗൗരവമായ ഒന്നാണ്. ഒരു ലക്ഷ്യം വെച്ചു കൊണ്ടു ശേഖരിച്ച വിവരങ്ങൾ മറെറാരു ലക്ഷ്യമുള്ള അന്വേഷണത്തിന് ഉപകരിച്ചു എന്നു വരികയില്ല. കേരളത്തിലെ കാർഷികോൽപന്നങ്ങളെ പററിയാണ് അന്വേഷണം നടത്തുന്നതെന്നിരിക്കട്ടെ. കയററുമതി സാധ്യതകൾ ആരായുവാൻ വേണ്ടി നടത്തിയ അന്വേഷണത്തിൽ നിന്നു ലഭിക്കുന്ന വിവരങ്ങൾ ഇറക്കുമതിയുടെ ആവശ്യവും അളവും നിർണയിക്കുവാൻ ഉപകരിക്കുകയില്ല. ചുരുക്കത്തിൽ നമ്മുടെ ലക്ഷ്യം എന്താണോ അതേ ലക്ഷ്യം വെച്ചു കൊണ്ടു തന്നെ ശേഖരിച്ച വിവരങ്ങൾ ലഭിച്ചെങ്കിലേ നമ്മുടെ ആവശ്യത്തിനു് ഉപകരിക്കു.

ii) *സമഷ്ടിയുടെ നിർവചനം*

ദ്വിതീയക സ്ഥിതിവിവരങ്ങൾ ശേഖരിക്കുമ്പോൾ പ്രത്യേകം ശ്രദ്ധിക്കാനുള്ള മറെറാരു കാര്യം സമഷ്ടിയെ എങ്ങനെയാണ് നിർവചിച്ചിരുന്നതു് എന്നതാണ്. ആ നിർവചനം നാം നൽകാൻ ആഗ്രഹിക്കുന്ന നിർവചനവുമായി വ്യത്യാസമുണ്ടെങ്കിൽ ആ സ്ഥിതി വിവരങ്ങൾ സ്വാഭാവികമായും സ്വീകാര്യമല്ലാതായി തീരും. ഹൈസ്കൂൾ വിദ്യാർത്ഥികളുടെ ചെലവുകളെ പററി നേരത്തെ ഒരു പഠനം നടത്തിയിട്ടുണ്ടെന്നിരിക്കട്ടെ. പന്ത്രണ്ടു വയസ്സിനും പതിനഞ്ചു വയസ്സിനും മധ്യെ പ്രായമുള്ള വിദ്യാർത്ഥികളുടെ ചെലവിനെ പററിയാണ് നാം അന്വേഷിക്കുന്നതെന്നും സങ്കല്പിക്കുക. സമഷ്ടിയുടെ നിർവചനം രണ്ടിലും ഒന്നല്ല. അതു കൊണ്ടു തന്നെ മുൻ അന്വേഷണത്തിൽ നിന്നു ലഭിച്ച വിവരങ്ങൾ നമ്മുടെ ആവശ്യത്തിനു് മതിയാവുകയില്ലെന്നു വരുന്നു. ഹൈസ്കൂൾ ക്ലാസ്സുകളിൽ പഠിക്കുന്ന വിദ്യാർത്ഥികളിൽ ഭൂരിപക്ഷവും പന്ത്രണ്ടു വയസ്സിനും പതിനഞ്ചു വയസ്സിനും മധ്യെ പ്രായമുള്ളവരായിരുന്നേക്കാമെങ്കിലും എല്ലാവരും അങ്ങനെ ആയിരിക്കണമെന്നില്ലല്ലൊ. ഈ പ്രായത്തിലുള്ളവർ മററു ക്ലാസ്സുകളിൽ പഠിക്കുന്നവരിലും കണ്ടേക്കാം.

iii) *അന്വേഷണരീതിയും സൂക്ഷ്മതയും*

നാം പരിശോധിക്കുന്ന ദ്വിതീയക സ്ഥിതിവിവരങ്ങൾ ശേഖരിക്കാൻ അവലംബിച്ച അന്വേഷണരീതിയും അതിനു് ഉണ്ടായിരിക്കുമെന്നു് സങ്കല്പിക്കപ്പെടുന്ന സൂക്ഷ്മതയും പ്രത്യേകം പരിഗണിക്കേണ്ട കാര്യങ്ങളാണ്. അവ നമുക്കു് അംഗീകാര്യങ്ങളാണെങ്കിൽ മാത്രമേ ആ സ്ഥിതിവിവരങ്ങൾ സ്വീകരിക്കാവൂ.

iv) *വിവരശേഖരണത്തിനു് ഉപയോഗിച്ച ചോദ്യാവലിയിലെ പദങ്ങൾക്കു് നൽകപ്പെട്ടിരുന്ന നിർവചനം*

ഒരേ ചോദ്യത്തിൻെറ തന്നെ ഉത്തരം ആ ചോദ്യത്തിൽ ഉപയോഗിച്ചിരിക്കുന്ന പദങ്ങളുടെ നിർവചനം മാറുന്നതനുസരിച്ച് മാറുമെന്നുള്ളതു് ശ്രദ്ധേയമായ മറെറാരു വസ്തുതയാണു്. 'സാക്ഷരതത്വം' എന്നതിനു് മുൻ അന്വേഷണത്തിൽ മാതൃഭാഷയിൽ എഴുതാനും വായിക്കാനും അറിയാമായിരിക്കുക എന്നാണു് നിർവചിച്ചിരുന്നതു് എന്നിരിക്കട്ടെ. പ്രൈമറി വിദ്യാഭ്യാസമെങ്കിലും ഉണ്ടായിരിക്കുക എന്നതാണു് സാക്ഷരതത്വത്തിനു് നാം നല്ലാനാഗ്രഹിക്കുന്ന നിർവചനമെങ്കിൽ മുൻ അന്വേഷണത്തിൽ ശേഖരിച്ച വിവരങ്ങൾ സ്വീകാര്യമല്ലാതായിത്തീരുന്നു. ഇതുപോലെ തന്നെ മററു പദങ്ങളുടെ നിർവചനത്തിന്റെ കാര്യവും. ചോദ്യാവലിയിലെ പദങ്ങൾക്കു് നാം നല്ലാനാഗ്രഹിക്കുന്ന നിർവചനങ്ങൾ തന്നെ നല്ലി ശേഖരിച്ച വിവരങ്ങളേ നമുക്കു് ഉപകരിക്കു.

v) *അന്വേഷണം നടത്തിയ കാലഘട്ടം*

സമയം മാറുന്നതനുസരിച്ച് ചില വിവരങ്ങൾക്കു് കാര്യമായ മാറ്റങ്ങളുണ്ടാവും. ഭക്ഷ്യധാന്യങ്ങളുടെ വിലനിലവാരത്തെപ്പററി കൊയ്ത്തുകാലത്തു ശേഖരിച്ച വിവരങ്ങൾ ലഭ്യമാണെന്നു വിചാരിക്കുക. നാം വിവരം ശേഖരിക്കാൻ ആഗ്രഹിക്കുന്ന കാലഘട്ടം കൊയ്ത്തുകാലമല്ലെങ്കിൽ മുൻ അന്വേഷണത്തിൽ ശേഖരിച്ച വിവരം കൊണ്ടു് നമുക്കു് വലിയ പ്രയോജനമൊന്നുമില്ല. അതു പോലെ തന്നെ ഇന്ത്യയുടെ കയററുമതിസാധനങ്ങളെപ്പററി 1940–ൽ ശേഖരിച്ച വിവരങ്ങൾ 1970 ൽ അപ്രസക്തങ്ങളാണു്. ദ്വിതീയക സ്ഥിതിവിവരങ്ങളുടെ സ്വീകാര്യത പരിശോധിക്കുമ്പോൾ പ്രത്യേകം ശ്രദ്ധിക്കേണ്ട ഒരു വിഷയമാണു്.

vi) *വിവരങ്ങൾ ശേഖരിച്ച വ്യക്തിയുടെ ആ വിഷയത്തിലുള്ള പരിജ്ഞാനവും വിവരം നല്കിയ വ്യക്തികളുടെ വിശ്വാസ്യതയും*

ആർ ആരിൽ നിന്നു് ശേഖരിച്ചതാണു് പരിഗണിക്കപ്പെടുന്ന ദ്വിതീയക സ്ഥിതിവിവരങ്ങൾ എന്നതും പ്രധാനപ്പെട്ട ഒരു കാര്യമാണു്. വിവരശേഖരണത്തിൻെറ സാങ്കേതികവശങ്ങൾ അറിയുന്നവർ വിവരങ്ങൾ നൽകാൻ കഴിവും സന്നദ്ധതയും ഉള്ളവരിൽ നിന്നു് ശേഖരിച്ചവയാണോ ആ വിവരങ്ങൾ എന്നു് പ്രത്യേകം ശ്രദ്ധിക്കണം.

ഈ എല്ലാ പരിഗണനകളും ദ്വിതീയക സ്ഥിതിവിവരങ്ങൾ സ്വീകരിക്കുന്ന കാര്യത്തിൽ ആവശ്യമാണു്. ശ്രദ്ധാപൂർവമുള്ള പരിശോധനയ്ക്കു ശേഷം സ്വീകാര്യമെന്നു തോന്നിയെങ്കിൽ മാത്രമേ അങ്ങനെയുള്ള സ്ഥിതിവിവരങ്ങൾ അംഗീകരിക്കാവൂ. ദ്വിതീയക സ്ഥിതിവിവരങ്ങളുടെ പ്രധാന ആകർഷണം പണം, അധ്വാനം, സമയം എന്നിവയിൽ ഉണ്ടാകാവുന്ന ലാഭമാണു്. പക്ഷേ അന്വേഷണലക്ഷ്യങ്ങൾ എത്ര കണ്ടു് സാക്ഷാത്ക്കരിക്കാൻ കഴിയും എന്നതിനെപ്പററിയും എത്രമാത്രം തെററുകൾ ഒഴിവാക്കാനാവും എന്നതിനെപ്പററിയും പ്രത്യേകം ചർച്ച ചെയ്തതിനു ശേഷം വേണം തീരുമാനമെടുക്കാൻ. കൂടുതൽ ശരിയായ വിവരങ്ങൾ ലഭിക്കുന്നതു കൊണ്ടുള്ള നേട്ടങ്ങളും അതിനു വേണ്ടിവരുന്ന ബുദ്ധിമുട്ട്, ധനനഷ്ടം, സമയനഷ്ടം എന്നിവയും താരതമ്യപ്പെടുത്തണം.

സാധാരണ ഗവണ്മെൻറ് പ്രസിദ്ധീകരണങ്ങൾ, വാണിജ്യവ്യവസായ സ്ഥാപനങ്ങളുടെ പ്രസിദ്ധീകരണങ്ങൾ, രാഷ്ട്രീയസ്വഭാവമുള്ള പത്രങ്ങൾ, മാസികകൾ തുടങ്ങിയവ ഗവേഷകന്മാരുടെയോ ഗവേഷണസ്ഥാപനങ്ങളുടെയോ അന്വേഷണറിപ്പോർട്ടുകൾ തുടങ്ങിയവയാണ് ദ്വിതീയക സ്ഥിതിവിവരങ്ങളുടെ ഖനികൾ.

5. സാമ്പിളന ഫ്രേം

സാമ്പിളനരീതിയാണ് വിവരശേഖരണത്തിന് അവലംബിക്കാൻ പോകുന്നത് എന്നു തീരുമാനിച്ചു കഴിഞ്ഞാൽ ആദ്യമായി സാമ്പിളന ഫ്രേം തയ്യാറാക്കാനാണ് ശ്രമിക്കേണ്ടത്. താത്വികമായി പറഞ്ഞാൽ, സാമ്പിളനവ്യക്തികളുടെ പട്ടികയാണ് സാമ്പിളന ഫ്രേം. ഒരു സമഷ്ടിയിൽ നിന്ന് സാമ്പിളെടുക്കുമ്പോൾ അതിലെ പ്രാഥമികവ്യക്തികളെയോ, വ്യക്തികളായി സങ്കൽപ്പിക്കാവുന്നതാണ്. അങ്ങനെ സാമ്പിളനത്തെ സംബന്ധിച്ചിടത്തോളം വ്യക്തികളായി അംഗീകരിക്കപ്പെടുന്നവയാണ് സാമ്പിളനവ്യക്തികൾ. അന്വേഷണലക്ഷ്യങ്ങളും അന്വേഷണത്തിന്റെ സ്വഭാവവും പ്രായോഗികസൗകര്യങ്ങളും എല്ലാം പരിഗണിച്ചു വേണം സാമ്പിളനവ്യക്തികളെ നിർവചിക്കുവാൻ. കേരളത്തിലെ വിദ്യാർത്ഥികളെപ്പററിയുള്ള ഒരു അന്വേഷണത്തിൽ വിദ്യാർത്ഥികളെയോ, ക്ലാസുകളെയോ, സ്കൂളുകളെയോ, വിദ്യാഭ്യാസജില്ലകളെയോ എന്തിനെ വേണമെങ്കിലും സാമ്പിളനവ്യക്തികളായി അംഗീകരിക്കാവുന്നതാണ്. ഇതിലെ പ്രാഥമികവ്യക്തി വിദ്യാർത്ഥിയാണ്. സാമ്പിളനവ്യക്തിയെ നിവർചിക്കുമ്പോൾ ഒരു പ്രാഥമികവ്യക്തി—ഇവിടെ ഒരു വിദ്യാർത്ഥി—ഒന്നിലധികം സാമ്പിളനവ്യക്തികളിൽ അംഗമായി വരുന്ന എന്ന കാര്യത്തിൽ പ്രത്യേകം ശ്രദ്ധിക്കണമെന്നു മാത്രം. അതായത് സാമ്പിളനവ്യക്തികൾ അന്യോന്യവ്യതിരിക്തങ്ങളായിരിക്കണം. സാധാരണഗതിയിൽ സാമ്പിളനവ്യക്തികളെ നിർവചിക്കുക അത്ര ബുദ്ധിമുട്ടുള്ള കാര്യമായിരിക്കുകയില്ല പക്ഷേ പ്രയാസം വരുന്ന സാഹചര്യങ്ങളും ഉണ്ട്. ഒരു പ്രത്യേകതരം നെൽച്ചെടിയുടെ ഉല്പാദനശേഷിയെപ്പറ്റിയാണ് അന്വേഷണം നടത്തുന്നത് എന്നിരിക്കട്ടെ. ഓരോ നെൽച്ചെടിയും ഓരോ പ്രാഥമികവ്യക്തിയാണ്. പക്ഷേ ഈ അന്വേഷണത്തിൽ പ്രാഥമികവ്യക്തിയെ സാമ്പിളനവ്യക്തിയായി അംഗീകരിച്ചാൽ പ്രായോഗികമായി വളരെ ബുദ്ധിമുട്ടുകളുണ്ടാവും. അതു കൊണ്ട് അവയുടെ ഏതെങ്കിലും സമുച്ചയത്തെ വേണം സാമ്പിളനവ്യക്തിയായി അംഗീകരിക്കാൻ. അവയുടെ സ്വയം വേർപെട്ടു നിൽക്കുന്ന സ്വന്തമായ വ്യക്തിത്വമുള്ള, സമുച്ചയങ്ങൾ നിലവിലില്ല. കൃത്രിമമായി സാമ്പിളനവ്യക്തികളെ നിർവചിക്കുകയേ നിർവാഹമുള്ളൂ. ആ നെൽവിത്ത് കൃഷി ചെയ്തിട്ടുള്ള വയലുകളെ സൗകര്യം പോലെ ബ്ലോക്കുകളായി വിഭജിച്ച് അവയോരോന്നിനെയും ഓരോ സാമ്പിളനവ്യക്തിയായി സ്വീകരിക്കുകയാണ് ചെയ്യാവുന്നത്. ഇത് അത്ര എളുപ്പമായ ഒരു കാര്യമല്ല എന്നത് വ്യക്തമാണല്ലോ.

സാമ്പിളനവ്യക്തികളെ നിർവചിച്ചിട്ട വേണം സാമ്പിളനഫ്രേം തയ്യാറാക്കുവാൻ. പക്ഷേ, പലപ്പോഴും സാമ്പിളനവ്യക്തികളുടെ പട്ടിക തയ്യാറാക്കുക എളുപ്പമായിരിക്കുകയില്ല. അങ്ങനെ വരുമ്പോൾ തത്തുല്യമായ മറെറന്തെങ്കിലും വിവരണത്തിന് സാമ്പിളനഫ്റേം എന്ന പേരു കൊടുക്കുകയാണ് പതിവ്. അതു് ഭൂപടങ്ങളോ വിവരണങ്ങളോ ആവാം. കേരളത്തിലെ പശുക്കളുടെ ക്ഷീരോല്പാദനശേഷി നിർണയിക്കുകയാണ് അന്വേഷണലക്ഷ്യമെന്നിരിക്കട്ടെ. ഏതാനും പശുക്കളെ തിരഞ്ഞെടുത്തു് അവയുടെ ക്ഷീരോൽപാദനശേഷി കണ്ടുപിടിച്ചാണ് ഇതു് സാധിക്കുവാനുദ്ദേശിക്കുന്നതു് എന്നും വിചാരിക്കുക ഇവിടെ സാമ്പിളനവ്യക്തികൾ കേരളത്തിലെ കറവപ്പശുക്കളാണ്. അവയുടെ പട്ടികയാണ് സാമ്പിളനഫ്രേം. പക്ഷേ അങ്ങനെ ഒരു പട്ടിക നിലവിലില്ല. ഒന്നു തയ്യാറാക്കുക എളുപ്പവുമല്ല. കറവപ്പശുക്കളെപ്പററി ലഭ്യമായ വിവരങ്ങളുടെ ആകത്തുകയാണ് ഫ്രേമായി സ്വീകരിക്കാവുന്നതു്. ഇതു് കറവപ്പശുക്കളുള്ള വീടുകളുടെ പട്ടികയാവാം; മേച്ചിൽസ്ഥലങ്ങളെപ്പററിയുള്ള വിവരങ്ങളാവാം; കറവപ്പശുക്കളുള്ള സ്ഥലങ്ങളിൽ എത്തിച്ചേരാനുള്ള മാർഗനിർദേശങ്ങളാവാം; ഫ്രേമിൻെറ സ്വഭാവമനുസരിച്ച് സാമ്പിളനപ്രക്രിയയിലും മാററങ്ങൾ വരും. സാമ്പിളനവ്യക്തികളുടെ പട്ടികയാണ് ഫ്രേമായി ഉപയോഗിക്കുന്നതു് എന്ന സങ്കല്പത്തിലാണ് ഇവിടെ നാമിതു ചർച്ച ചെയ്യുന്നതു്. മററു സന്ദർഭങ്ങളിൽ മൗലികതത്വങ്ങൾ ലംഘിക്കാതെ ആവശ്യമായ പ്രായോഗികക്രമീകരണങ്ങൾ വരുത്തേണ്ടതാണ്. ഒരു കമ്പനിയിലെ കണക്കു ബുക്കുകളാണ് സാമ്പിളനവ്യക്തികൾ എന്നിരിക്കട്ടെ. കണക്കുബുക്കുകൾക്ക് നമ്പരോ പേരോ ഉണ്ടെങ്കിൽ അവയുടെ പട്ടികയാണ് സാമ്പിളനഫ്രേം. അല്ലാത്ത പക്ഷം അവ അടുക്കി വെച്ചിരിക്കുന്ന അലമാരകളും ഓരോ അലമാരയിലും വെച്ചിരിക്കുന്ന ബുക്കുകളും എതെങ്കിലും ഒരു ക്രമത്തിൽ പരിഗണിച്ച് സാമ്പിളെടുക്കാവുന്നതാണ്. ബുക്കുകൾ വെച്ചിരിക്കുന്ന സ്ഥാനങ്ങളെപ്പററിയുള്ള അറിവാണ് സാമ്പിളനഫ്രേമിൻെറ ജോലി നിർവഹിക്കുന്നതു്. ഇങ്ങനെ സന്ദർഭാനുസരണം സാമ്പിളനവ്യക്തികളുടെ പട്ടികയ്ക്കു പകരം ലഭ്യമായ മററു വിവരങ്ങൾ സാമ്പിളനഫ്രേമായി ഉപയോഗിക്കാവുന്നതാണ്.

ഫ്രേമിൻെറ പൂർണത, തെററില്ലായ്മ, ഓരോ വ്യക്തിയെപ്പററിയും അതു നല്കുന്ന ആനുഷംഗികവിവരങ്ങൾ എന്നിവയെല്ലാം പ്രാധാന്യമർഹിക്കുന്ന കാര്യങ്ങളാണ്. അപൂർണമായ ഫ്രേം, അതായതു് സമഷ്ടിയിലെ എല്ലാ സാമ്പിളനവ്യക്തികളെയും ഉൾക്കൊള്ളാത്ത ഫ്രേം, ഉപയോഗിച്ച് എടുക്കുന്ന സാമ്പിളുകൾ സമഷ്ടിയെപ്പററി ശരിയായ വിവരങ്ങൾ നല്കാനിടയില്ല. അതുപോലെ തന്നെ ഫ്രേമിലെ തെററുകൾ അന്വേഷണഘട്ടത്തിൽ വളരെയേറെ ബുദ്ധിമുട്ടുകൾക്ക് ഇടയാക്കും. ഇല്ലാത്ത വ്യക്തികളെ തേടി നടക്കാനും വിവരം ശേഖരിക്കേണ്ട വ്യക്തികളെ തിരിച്ചറിയാതെ വരാനും മററും ഇതു് ഇടയാക്കിയേക്കാം. സാമ്പിളനവ്യക്തികളെപ്പററിയുള്ള ആനുഷംഗികവിവരങ്ങൾ ഫ്രേമിലുണ്ടായിരിക്കണമെന്നുള്ളതു് നിർബന്ധമാണ്. ഓരോ വ്യക്തിയെയും സംശയാതീത

മായി, ഏററവും വേഗത്തിൽ കണ്ടെത്തുവാൻ പര്യാപ്തമായ വിവരങ്ങളാണ് ആവശ്യം. അതു പോലെ തന്നെ വ്യക്തികളെപ്പററി പ്രാഥമികങ്ങളായ ചില വിവരങ്ങൾ ഫ്രേമിലുണ്ടായിരിക്കുന്നതു് അഭികാമ്യമാണ്. സമഷ്ടിയെ വിവിധ സ്തരങ്ങളായി വിഭജിക്കേണ്ടതായും മററും വരുമ്പോൾ ഇമ്മാതിരി വിവരങ്ങൾ അത്യന്തം ഉപയോഗപ്രദമായിരിക്കും. ഉദാഹരണമായി, കേരളത്തിലെ വിദ്യാലയങ്ങളുടെ പട്ടികയാണ് ഫ്രേമെന്നിരിക്കട്ടെ. ഓരോ വിദ്യാലയവും എവിടെയാണെന്നും പ്രൈമറി, സെക്കൻഡറി, കോളേജ് എന്നീ വിഭാഗങ്ങളിൽ ഏതിൽ പെട്ടതാണെന്നും, ഗവണ്മെന്റ് വകയാണോ, സ്വകാര്യഏജൻസികളുടെതാണോ എന്നും മററും ഫ്രേമിൽ തന്നെ രേഖപ്പെടുത്തിയിരുന്നാൽ അതു് അത്യന്തം ഉപകാരപ്രദമായിരിക്കും. ചുരുക്കത്തിൽ, സാമ്പിളനത്തിൻെറ വിജയം ഫ്രേമിൻെറ വൈശിഷ്ട്യത്തെ വളരെയേറെ ആശ്രയിച്ചാണിരിക്കുന്നതു്. സാമ്പിളന രീതി നിശ്ചയിക്കുന്നതിലും ഫ്രേമിനു് നിർണായകമായ ഒരു പങ്കു വഹിക്കുവാനുണ്ടു്.

പലപ്പോഴും ഔദ്യോഗികരേഖകളിൽ നിന്നോ, മുൻപു് തയ്യാറാക്കിയിട്ടുള്ള ഫ്റേമുകളിൽ ആവശ്യമായ ഭേദഗതികൾ വരുത്തിയോ ഫ്റേം തയ്യാറാക്കാൻ കഴിയും. ചുരുക്കം സന്ദർഭങ്ങളിൽ പുത്തനായി തന്നെ ഫ്റേം തയ്യാറാക്കേണ്ടതായും വരും. അങ്ങനെ വന്നാൽ ഒരു പ്രാഥമികാന്വേഷണം ആവശ്യമായി തീരും. പഴയ ഫ്റേമുകൾ ഉപയോഗിക്കുമ്പോൾ അതിൽ ആവശ്യമായ കൂട്ടിക്കുറയ്ക്കലുകൾ വരുത്താൻ പ്രത്യേകം ശ്രദ്ധിക്കണം. ഒരു രാജ്യത്തെ വീടുകളുടെ പട്ടികയാണ് ഫ്റേമെങ്കിൽ ആ പട്ടിക തയ്യാറാക്കിയതിനു ശേഷം പുതുതായി ഉണ്ടായിട്ടുള്ള വീടുകൾ കൂടി അതിൽ ഉൾപ്പെടുത്തുകയും പൊളിച്ചു കളഞ്ഞവയോ ഉപേക്ഷിക്കപ്പെട്ടവയോ ആയ വീടുകൾ അതിൽ നിന്നു് വിട്ടുകളയുകയും ചെയ്യണം. ഭൂപടമാണ് ഫ്റേമായി ഉപയോഗിക്കുന്നതെങ്കിൽ അതു് എന്നു തയ്യാറാക്കിയതാണെന്നു് പരിശോധിക്കണം. അതിനു ശേഷം വിശദാംശങ്ങളെ സംബന്ധിച്ചിടത്തോളം വളരെയേറെ മാററങ്ങൾ സംഭവിച്ചിട്ടുണ്ടാവും. അവ കൂടി രേഖപ്പെടുത്തി നവീകരിച്ചതിനു ശേഷം വേണം അതു് ഉപയോഗിക്കാൻ.

ഒരു ഫ്രേമിനുണ്ടായിരിക്കേണ്ട ഏററവും പ്രധാനപ്പെട്ട വൈശിഷ്ട്യം അതിൽ ഓരോ സാമ്പിളനവ്യക്തിയും ഒരു പ്രാവശ്യം ഉൾപ്പെട്ടിരിക്കണം, ഒന്നിലധികം പ്രാവശ്യം ഉൾപ്പെട്ടിരിക്കരുതു് എന്നതാണ്. പക്ഷേ ഇമ്മാതിരി ഫ്രേമുകൾ കിട്ടുക അത്ര എളുപ്പമല്ല. എത്ര ശ്രദ്ധിച്ചു തയ്യാറാക്കിയാലും, സമഷ്ടി താരതമ്യേന വലുതാണെങ്കിൽ കുറെയെല്ലാം പോരായ്മകൾ അതിനുണ്ടായിരിക്കുമെന്നതു് ഏതാണ്ടു് തീർച്ചയാണ്. ഫ്രേമിനു് പോരായ്മകളുണ്ടെന്നു ബോധ്യമായാൽ നമുക്കു് അവലംബിക്കാവുന്ന നടപടിക്രമങ്ങൾ താഴെ പറയുന്നവയാണ്.

a) പോരായ്മകൾ അവഗണിക്കുക

തെററുകൾ താരതമ്യേന കുറവായിരിക്കുമ്പോഴും, അവ കണ്ടു പിടിച്ചു് പരിഹരിക്കുവാൻ കൂടുതൽ സമയവും പണവും അധ്വാനവും ആവശ്യമായി

വരുമ്പോഴമാണ് അവയെ അവഗണിക്കുവാൻ തീരുമാനിക്കാറുള്ളതു്. പതിവായി ഹോട്ടലിൽ ഭക്ഷണം കഴിക്കാറുള്ളവരെ പററിയാണ് അന്വേഷണം നടത്തുന്നതു് എന്നിരിക്കട്ടെ. വിവിധ ഹോട്ടലുകളിൽ നിന്നു് അവിടെ പതിവായി ഭക്ഷണം കഴിക്കുന്നവരുടെ ലിസ്റ്റു് തയ്യാറാക്കി വാങ്ങിയാൽ നമുക്കു് ആവശ്യമായ ഫ്റേമായി. പക്ഷേ ഒന്നിലധികം ഹോട്ടലുകളിൽ നിന്നു് ഭക്ഷണം കഴിക്കുന്നവർ കണ്ടേക്കാം. അവരുടെ പേരു് ഫ്റേമിൽ ഒന്നിലധികം തവണ ഉണ്ടായിരിക്കും. അതുപോലെ തന്നെ ചിലരുടെ പേരുകൾ ലിസ്റ്റിൽ ഉൾപ്പെടുത്താൻ ഹോട്ടലുടമകൾ വിട്ടു പോയി എന്നും വരാവുന്നതാണ്. പക്ഷേ ഇങ്ങനെ ആവർത്തിക്കപ്പെടുകയോ വിട്ടു പോവുകയോ ചെയ്യുന്ന ആളുകളുടെ എണ്ണം താരതമ്യേന കുറവായിരിക്കുമെന്നതു കൊണ്ടു് അവയെ അവഗണിച്ചാലും വലിയ തകരാറുകൾ സംഭവിക്കാനില്ല.

ഇമ്മാതിരി സന്ദർഭങ്ങളിലാണു് പോരായ്മകൾ അവഗണിക്കാറുള്ളതു്. അന്വേഷകൻ നേരിട്ടു് ആവശ്യമായ വിവരങ്ങൾ ശേഖരിച്ചു് ഫ്റേമിൽ ഒന്നിലധികം തവണ ഉൾപ്പെട്ടിരിക്കാവുന്നവരുടെയും ഉൾപ്പെടാതെ പോയിരിക്കാവുന്നവരുടെയും എണ്ണം ആകലനം ചെയ്യാൻ ഒരു ശ്രമം നടത്തുന്നതും നല്ലതാണ്. നാം എത്തിച്ചേരുന്ന നിഗമനങ്ങളുടെ സ്വീകാര്യതയെപ്പററിയും സൂക്ഷ്മതയെപ്പറ്റിയും കുറെയെങ്കിലും മനസ്സിലാക്കുവാൻ ഈ ആകലങ്ങൾ സഹായിക്കും.

b) ലഭ്യമായ ഫ്രേമിന്റെ പോരായ്മകൾ പരിഗണിച്ചു് സമഷ്ടിയുടെ നിർവചനത്തിൽ ചില മാററങ്ങൾ വരുത്തുക

കഴിവതും ഒഴിവാക്കാൻ ശ്രമിക്കേണ്ട ഒരു നടപടിയാണു് ഇതു്. അസാധാരണഘട്ടങ്ങളിലോ മററു് ഗത്യന്തരമില്ലാതെ വരുമ്പോഴോ മാത്രമേ ഈ രീതി അംഗീകരിക്കാവൂ. ഗവണ്മെൻറുദ്യോഗസ്ഥന്മാരുടെ ജീവിതച്ചെലവുകളെപ്പററിയാണു് അന്വേഷണം നടത്തുന്നതു് എന്നിരിക്കട്ടെ. അന്വേഷണസമയത്തു് സർവീസിലുള്ള എല്ലാ ഉദ്യോഗസ്ഥന്മാരും സമഷ്ടിയിലെ അംഗങ്ങളാണ്. പക്ഷേ ലഭ്യമായ ഗവണ്മെൻറുറിക്കാർഡുകളിൽ താൽക്കാലികനിയമനം ലഭിച്ചിട്ടുള്ളവരുടെ പേരുകൾ ഉൾപ്പെടുത്തിയിട്ടില്ല എന്നു വരാവുന്നതാണ്. അങ്ങനെ നമുക്കു് റിക്കാർഡുകളെ അടിസ്ഥാനമാക്കി തയ്യാറാക്കാവുന്ന ഫ്റേം അപൂർണമായിരിക്കും. അതു് പൂർണമാക്കുക അത്ര എളുപ്പവുമല്ല. ഇങ്ങനെ ഒരു സന്ദർഭത്തിൽ സമഷ്ടിയുടെ നിർവചനത്തിൽ അൽപമൊരു മാററം വരുത്തിയാൽ പല ബുദ്ധിമുട്ടുകളും ഒഴിവാക്കാം. സ്ഥിരം സർവീസിലുള്ള ഗവണ്മെൻറുദ്യോഗസ്ഥന്മാരാണു് പരിഗണിക്കേണ്ട സമഷ്ടിയെന്നു സങ്കൽപിച്ചാൽ, ഫ്റേമുണ്ടാക്കുക എളുപ്പമായിത്തീരും. അതു് പൂർണമായിരിക്കുകയും ചെയ്യും. ഈ മാററം സാമ്പിളനലക്ഷ്യങ്ങളെ സംബന്ധിച്ചിടത്തോളം വളരെയൊന്നും അസ്വീകാര്യമല്ല താനും. പക്ഷേ മററു പല സന്ദർഭങ്ങളിലും സമഷ്ടിയുടെ നിർവചനത്തിൽ വരുത്തുന്ന മാററങ്ങൾ ഇത്ര തന്നെ നിരുപദ്രവങ്ങളായിരിക്കണമെന്നില്ല. ഓരോ സാഹചര്യത്തിലും സമഷ്ടിയുടെ നിർവചനത്തിൽ വരുത്തുന്ന ഏതെങ്കിലും മാററം സാമ്പിളനലക്ഷ്യങ്ങളെ എത്ര മാത്രം ബാധിക്കുമെന്നു് ശ്രദ്ധാപൂർവം

പരിശോധിച്ചിട്ട വേണം ആ മാററം വരുത്തണമോ എന്നു നിശ്ചയിക്കുവാൻ. വരുത്തിയിട്ടുള്ള മാററങ്ങൾ രേഖപ്പെടുത്തുകയും വേണം.

c) ഫ്രേം നവീകരിക്കുക

ഏററവും അഭികാമ്യമായ ഒരു നടപടിക്രമമാണിതു്. ലഭ്യമായ ഫ്രേം പരിശോധിച്ച് അതിലുള്ള പോരായ്മകൾ കണ്ടുപിടിച്ച് പരിഹരിക്കുക എന്നതാണു് നവീകരണം എന്നതു കൊണ്ടു് അർത്ഥമാക്കുന്നതു്. ഇതു് പലപ്പോഴും അത്ര ബുദ്ധിമുട്ടില്ലാതെ സാധിച്ചെന്നു വരും. അങ്ങനെയുള്ള സന്ദർഭങ്ങളിൽ തീർച്ചയായും ഈ നടപടിയാണു് അംഗീകരിക്കേണ്ടതു്. പക്ഷേ, മിക്കവാറും നവീകരണം അത്ര എളുപ്പമായിരിക്കുകയില്ല. പണച്ചെലവും ബുദ്ധിമുട്ടും സമയനഷ്ടവും ഉൾക്കൊള്ളുന്ന ഒന്നായിരിക്കും അതു്. അതു കൊണ്ടുള്ള നേട്ടങ്ങളും അതിനുള്ള ബുദ്ധിമുട്ടുകളും താരതമ്യപ്പെടുത്തി വേണം ഒരു തീരുമാനമെടുക്കുവാൻ. ഔചിത്യമാണു് ഇങ്ങനെയുള്ള സന്ദർഭങ്ങളിലെ പ്രധാന വഴികാട്ടി എന്നു മാത്രമേ പറയാനാവൂ.

6. സാമ്പിളെടുക്കൽ

ഫ്രേം തയ്യാറാക്കിക്കഴിഞ്ഞാൽ അടുത്തതു് സാമ്പിളെടുക്കലാണു്. സമഷ്ടിയെപ്പററിയാണല്ലോ പഠനം നടത്തേണ്ടതു്. സമഷ്ടിയിലെ എല്ലാ വ്യക്തികളിൽ നിന്നും വിവരം ശേഖരിക്കുക എന്ന പ്രയാസമേറിയ ജോലി ഒഴിവാക്കാനാണു് സാമ്പിളനരീതി അവലംബിക്കുന്നതു്. അതിൽ നിന്നു് തന്നെ ഒരു സംഗതി വ്യക്തമാണു്. സമഷ്ടിയിലെ ഏതെങ്കിലും ചില വ്യക്തികളുടെ വെറുമൊരു സമുച്ചയമല്ല സാമ്പിൾ. പരിഗണിക്കപ്പെടുന്ന അഭിലക്ഷണങ്ങളെ സംബന്ധിച്ചിടത്തോളം സമഷ്ടിയെ പ്രതിനിധീകരിക്കുന്നതായിരിക്കണം അതു്. സമഷ്ടിയുടെ എല്ലാ പൊതുസ്വഭാവങ്ങളും സാമ്പിളിനും ഉണ്ടായിരിക്കണം. സാമ്പിൾ ഒരു കൊച്ചു സമഷ്ടി തന്നെ ആയിരിക്കണം. പക്ഷേ, ഈ ലക്ഷ്യം അപ്രാപ്യമായ ഒന്നാണു്. സാമ്പിളും സമഷ്ടിയും ഒന്നു തന്നെയാകാത്തിടത്തോളം കാലം സാമ്പിളിനു് സമഷ്ടിയെ പൂർണ്ണമായി പ്രതിനിധീകരിക്കാനാവില്ല. എങ്കിലും പ്രായോഗികപരിഗണനകൾ വെച്ചു നോക്കുമ്പോൾ ഒരു സാമ്പിൾ സമഷ്ടിയെ പ്രതിനിധീകരിക്കണമെന്നല്ലാതെ പൂർണ്ണമായും പ്രതിനിധീകരിക്കണമെന്നു് ശാഠ്യം പിടിക്കുന്നതു് അത്ര ബുദ്ധിപൂർവകമല്ല എന്നു കാണുവാൻ കഴിയും. സാമ്പിളിൽ നിന്നും ലഭിക്കുന്ന സമഷ്ടിയെപ്പററിയുള്ള വിവരങ്ങൾ അപൂർണ്ണമാണെങ്കിൽ പോലും വേഗത്തിലും കുറഞ്ഞ ചെലവിലും ലഭിക്കുന്നു എന്നതു കൊണ്ടു് തന്നെ അത്യന്തം പ്രയോജനപ്രദമാണു്.

സാമ്പിളനശാസ്ത്രത്തിന്റെ മൗലികലക്ഷ്യങ്ങളിലൊന്നു് സമഷ്ടിയെ കഴിവതും പ്രതിനിധീകരിക്കുന്ന സാമ്പിളുകൾ എങ്ങനെ എടുക്കാം എന്നു് ആരായുകയാണു്. സ്വാഭാവികമായും വിവിധ സാമ്പിളനരീതികൾ നിർദേശിച്ചും അവയുടെ സൂക്ഷ്മത പരിശോധിച്ചുമാണു് ഇതു് നിർവഹിക്കുന്നതു്. വരുന്ന അധ്യായങ്ങളിൽ ഈ വിഷയം വിശദമായി ചർച്ച ചെയ്യുന്നതു കൊണ്ടു് അതിനെപ്പററി

കൂട്ടതലൊന്നും ഇവിടെ പറയുന്നില്ല. ഒന്നു മാത്രം സൂചിപ്പിക്കാം. എത്ര അംഗങ്ങളുള്ള സാമ്പിളാണ് എടുക്കേണ്ടതു്, അതു് എങ്ങനെ എടുക്കണം എന്നീ സുപ്രധാന പ്രശ്നങ്ങളെ ആവശ്യമായ സൂക്ഷ്മത, ചെലവാക്കാവുന്ന പണം, അനുവദിക്കാവുന്ന സമയം, സമഷ്ടിയുടെ പൊതുസ്വഭാവം എന്നിവയുമായി ബന്ധപ്പെടുത്തിയാണ് കൈകാര്യം ചെയ്യേണ്ടതു്. ഓരോ സാമ്പിളനരീതിയെപ്പററിയും ചർച്ച ചെയ്യുമ്പോൾ ഈ വിഷയങ്ങൾ പരിഗണിക്കുന്നതാണ്.

7. വിവരങ്ങൾ എങ്ങനെ ശേഖരിക്കണം

വിവരങ്ങൾ ശേഖരിക്കുന്ന രീതി പ്രത്യേക പരിഗണനയർഹിക്കുന്ന ഒന്നാണ്. സൂക്ഷ്മത, പ്രായോഗികത, ചെലവു കുറവ് എന്നീ മൂന്നു കാര്യങ്ങൾ വിവരശേഖരണരീതി നിർണയിക്കുന്നതിൽ ശ്രദ്ധിക്കപ്പെടണം.

പ്രധാനമായും മൂന്നു തരത്തിലാണ് വിവരങ്ങൾ ശേഖരിക്കാറുള്ളതു്.

1. ചോദ്യാവലി തപാൽ വഴി അയച്ചു് വിവരം ശേഖരിക്കുക.
2. പരിശീലനം നല്ലപ്പെട്ട അന്വേഷകരെ ഉപയോഗപ്പെടുത്തുക.
3. അന്വേഷണം സംഘടിപ്പിക്കുന്ന ആൾ തന്നെ നേരിട്ട് വിവരം ശേഖരിക്കുക.

ഈ ഓരോ സമ്പ്രദായത്തിനും അതിന്റേതായ വൈശിഷ്ട്യങ്ങളും പോരായ്മകളുമുണ്ട്. ആദ്യത്തെ രീതി താരതമ്യേന ചെലവുള്ള ഒന്നാണ്. ചോദ്യാവലിയും അതോടൊപ്പം അതു പൂരിപ്പിക്കാൻ ആവശ്യമായ നിർദേശങ്ങളും വിശദീകരണങ്ങളും ഉൾക്കൊള്ളുന്ന ഒരു നിർദേശപത്രികയും കൂടി വിവരങ്ങൾ നൽകേണ്ട ആളുകൾക്ക് അയച്ചു കൊടുക്കുന്നു. അവർ അതു് പൂരിപ്പിച്ചു് തിരിച്ചയക്കുന്നു. ഇതിനു് ചില പോരായ്മകളുണ്ട്. വിദ്യാഭ്യാസസമ്പന്നരായ ആളുകളിൽ നിന്നു് വിവരം ശേഖരിക്കുമ്പോൾ മാത്രമേ ഈ രീതി വിജയിക്കുകയുള്ളു എന്നതൊന്നു്. അശ്രദ്ധ കൊണ്ടോ ജോലിത്തിരക്കു കൊണ്ടോ പലരും മറുപടി അയച്ചില്ലെന്നു വരാവുന്നതാണ്. പൂരിപ്പിക്കുന്നതിൽ സംഭവിക്കാവുന്ന പാകപ്പിഴകൾ പരിഹരിക്കുക എളുപ്പമല്ല. അന്വേഷണവിധേയമായ പ്രശ്നത്തെപ്പറ്റി പ്രത്യേകം താൽപര്യമുള്ള ആളുകൾ ചോദ്യാവലി പൂരിപ്പിക്കുകയും മററുള്ളവർ അയക്കാതിരിക്കുകയും ചെയ്തു എന്നു വരാൻ സംഭാവ്യതയുള്ളതു കൊണ്ടു് പൂരിപ്പിച്ചു കിട്ടിയ ചോദ്യാവലികളെ മാത്രം അടിസ്ഥാനമാക്കി ഏതെങ്കിലും നിഗമനങ്ങളിൽ എത്തിച്ചേരുന്നതു് അഭികാമ്യമല്ല. മറുപടി അയക്കാത്തവരിൽ നിന്നു കൂടി വിവരങ്ങൾ ശേഖരിക്കാൻ ശ്രദ്ധിക്കേണ്ടതാണ്. അതു് പലപ്പോഴും അത്ര എളുപ്പമായിരിക്കുകയില്ല. രാഷ്ട്രീയസ്വഭാവമുള്ള വിഷയങ്ങളെ പററി വിവരം ശേഖരിക്കുമ്പോഴാണ് ഈ കുഴപ്പം ഗൗരവതരമാകുന്നതു്. ആ വിഷയത്തിൽ പ്രത്യേക അഭിപ്രായങ്ങളുള്ള കക്ഷികളോ വിഭാഗങ്ങളോ ഉണ്ടാവാം. അവർ ചോദ്യാവലി പൂരിപ്പിച്ചയക്കുന്നതിൽ താൽപര്യം കാണിക്കും. നിഷ്പക്ഷമതികളായവർ അത്രയൊന്നും ശ്രദ്ധിക്കുകയുമില്ല. ചുരുക്കത്തിൽ, അന്വേഷകനു ലഭിക്കുന്ന വിവരങ്ങൾ ഏതെങ്കിലും വശത്തേക്കു് ചായ്‌വുള്ളവയായിരിക്കും. അതിനെ അടിസ്ഥാനമാക്കി സമഷ്ടിയെപ്പററി ഏതെങ്കിലും നിഗമനത്തിൽ എത്തിച്ചേർന്നാൽ അതു് ശരിയായിരിക്കുകയുമില്ല. ഈ സമ്പ്രദായത്തിനുള്ള

മറെറാരു പോരായ്മ ചോദ്യാവലിയിൽ വന്നു പോകാവുന്ന ചെറിയ അവ്യക്തതകൾ പോലും തെററായ വിവരങ്ങൾ നൽകാൻ ഇടയാക്കിയേക്കാമെന്നുള്ളതാണ്. ഈ രീതിയാണ് അവലംബിക്കാനുദ്ദേശിക്കുന്നതെങ്കിൽ ചോദ്യാവലിയും അതിന്റെ വിശദീകരണപത്രികയും വളരെ അവധാനപൂർവം തയ്യാറാക്കണം. കഴിവതും ഒററ വാക്കിൽ ഉത്തരം പറയാവുന്ന ചോദ്യങ്ങൾ ഉൾപ്പെടുത്തണം. 'ശരി' അല്ലെങ്കിൽ 'തെററ്' എന്ന മറുപടി പ്രതീക്ഷിക്കുന്ന വിധത്തിൽ ചോദ്യങ്ങൾക്ക് രൂപം നൽകുന്നതു് ഏററ നന്ന്. ചോദ്യത്തിൽ ഉപയോഗിച്ചിരിക്കുന്ന പദങ്ങൾക്ക് പൂർണമായ വിശദീകരണം നൽകിയിരിക്കണം. ചോദ്യങ്ങളുടെ എണ്ണം എത്രയും കുറയ്ക്കാമോ അത്രയും കുറയ്ക്കണം. മറുപടി അയയ്ക്കാൻ മേൽവിലാസമെഴുതിയ, ആവശ്യത്തിനുള്ള സ്റ്റാമ്പൊട്ടിച്ച കവർ കൂടി ചോദ്യാവലിയോടൊപ്പം അയച്ചു കൊടുക്കുന്നതു് മറുപടി കിട്ടാൻ സഹായിക്കും. സാമ്പിളിൽ ഉൾപ്പെട്ടിട്ടുള്ള എല്ലാവരെയും ചോദ്യാവലി പൂരിപ്പിച്ചയക്കാൻ താൽപര്യമുള്ളവരാക്കിത്തീർക്കാനുള്ള മാർഗങ്ങൾ ആരായണം. ഉദ്യോഗസ്ഥന്മാരിൽ നിന്നാണ് വിവരങ്ങൾ ശേഖരിക്കുന്നതെങ്കിൽ അവരുടെ മേലുദ്യോഗസ്ഥന്മാരെക്കൊണ്ട് നിർദേശങ്ങളോ അഭ്യർഥനകളോ പുറപ്പെടുവിക്കാൻ ശ്രമിക്കണം. തൊഴിലാളികളിൽ നിന്നാണെങ്കിൽ തൊഴിൽസംഘടനകൾ മുഖേന സ്വാധീനം ചെലുത്തുന്നതു കൊള്ളാം. പത്രങ്ങൾ, റേഡിയോ തുടങ്ങിയ പ്രചരണോപാധികൾ സാധിക്കുമെങ്കിൽ ഉപയോഗപ്പെടുത്താം. അങ്ങനെ എല്ലാവരും ശ്രദ്ധയോടെ ചോദ്യാവലി പൂരിപ്പിച്ചയയ്ക്കുന്ന ഒരു അന്തരീക്ഷം സൃഷ്ടിച്ചാൽ ഈ വിവരശേഖരണരീതി വിജയിക്കും. ഏററവും ചെലവു കുറഞ്ഞ രീതിയുമായിരിക്കും. പക്ഷേ നല്ല ഒരു പങ്ക് ആളുകൾ ചോദ്യാവലി പൂരിപ്പിച്ചയക്കാതിരിക്കുകയോ, അപൂർണമായി മാത്രം പൂരിപ്പിക്കുകയോ, ചോദ്യങ്ങൾ തെററിദ്ധരിച്ച് ഉത്തരമെഴുതുകയോ ചെയ്താൽ ഈ രീതിയിൽ ലഭിച്ച എല്ലാ വിവരങ്ങളും നിരുപയോഗമായിത്തീരും.

രണ്ടാമത്തെ രീതിയാണ് സാധാരണയായി ഉപയോഗിക്കാറുള്ളതു്. ഇതു് താരതമ്യേന കൂടുതൽ ചെലവുള്ള ഒന്നാണ്. പക്ഷേ ഏററവും പ്രായോഗികമായ രീതി ഇതാണെന്നു പറയാതെ തരമില്ല. ആവശ്യമുള്ളിടത്തോളം അന്വേഷകരെ നേരത്തെ തന്നെ തിരഞ്ഞെടുത്ത് അവർക്ക് ആവശ്യമായ പരിശീലനം നല്കണം. സത്യസന്ധരും അധ്വാനതല്പരരും വിവരശേഖരണത്തിൽ താല്പര്യമുള്ളവരുമായ ആളുകളെ തിരഞ്ഞെടുക്കാൻ പ്രത്യേകം ശ്രദ്ധിക്കണം. കഴിവതും മററു ജോലികളിൽ ഏർപ്പെട്ടിരിക്കുന്നവരായിരിക്കരുതു് ഇങ്ങനെ തിരഞ്ഞെടുക്കപ്പെടുന്നവർ. തൃപ്തികരമായ പ്രതിഫലം കൊടുക്കുന്ന കാര്യം പ്രത്യേകം പരിഗണിക്കേണ്ട ഒന്നാണ്. അസംതൃപ്തരായ അന്വേഷകർ ശരിയായി ജോലി ചെയ്തെന്നു വരികയില്ല. സാമ്പിളനലക്ഷ്യങ്ങളെപ്പററിയും ചോദ്യാവലിയിലെ ഓരോ ചോദ്യത്തിന്റെയും വിശദാംശങ്ങളെപ്പററിയും അവർ ബോധവാന്മാരായിരിക്കണം. വിവരം നല്കുന്നവരിൽ നിന്നുണ്ടാകാവുന്ന സഹകരണക്കുറവ്, എതിർപ്പ് തുടങ്ങിയവ എങ്ങനെ കൈകാര്യം ചെയ്യണമെന്നുള്ളതിനെപ്പററി വേണ്ട

നിർദേശങ്ങൾ അവർക്ക് നൽകിയിരിക്കണം. വിവരം നല്ലേണ്ടയാളുകളോട്ട ചോദിച്ച് അന്വേഷകർ തന്നെ ചോദ്യാവലി പൂരിപ്പിക്കുകയോ വിവരം നൽകേണ്ടവരെക്കൊണ്ട തന്നെ വേണ്ട വിശദീകരണങ്ങൾ നൽകി എഴുതിക്കുകയോ ചെയ്യാവുന്നതാണ്. അന്വേഷകർ തന്നെ പൂരിപ്പിക്കുന്നതാണ് ഉത്തമം. തെററു വരാനുള്ള സാധ്യത കുറയും. ചിഹ്നങ്ങളോ സംഖ്യകളോ കൊണ്ട് ഉത്തരങ്ങൾ സൂചിപ്പിക്കുന്ന സമ്പ്രദായം അവലംബിക്കാനും കഴിയും. ഉദാഹരണമായി "ഏതു പത്രമാണ് പതിവായി വായിക്കുന്നതു" എന്നു് ഒരു ചോദ്യമുണ്ടെന്നിരിക്കട്ടെ. ആ നാട്ടിൽ പ്രചാരത്തിലുള്ള പത്രങ്ങളുടെ ഒരു പട്ടികയുണ്ടാക്കി ഓരോ പത്രത്തിനും ഓരോ ക്രമസംഖ്യ നല്കിയിരുന്നാൽ ചോദ്യത്തിന്റെ നേരെ വിവരം നല്കുന്ന ആൾ വായിക്കുന്ന പത്രത്തിൻെറ ക്രമസംഖ്യ എഴുതിയാൽ മതിയല്ലൊ. പത്രത്തിന്റെ മുഴുവൻ പേരും എഴുതുക എന്ന ബുദ്ധിമുട്ട് ഇങ്ങനെ ഒഴിവാക്കാൻ കഴിയും.

അന്വേഷകരെ ഉപയോഗിച്ച് വിവരം ശേഖരിക്കുമ്പോൾ അവരുടെ പ്രവർത്തനങ്ങൾ പരിശോധിക്കാൻ വേണ്ട ഏർപ്പാടുകൾ ചെയ്തിരിക്കണം. ഇതിനായി പരിശോധകരെ നിയമിക്കണം. അവർക്ക് രണ്ട വിധത്തിൽ അന്വേഷകരുടെ പ്രവർത്തനങ്ങൾ നിയന്ത്രിക്കാം: (1) അന്വേഷകർക്ക് വേണ്ട നിർദേശങ്ങൾ നൽകിയും അവർ വേണ്ടപോലെ പ്രവർത്തിക്കുന്നുണ്ടോ എന്നു നിരീക്ഷിച്ചും (2) അന്വേഷകർ വിവരം ശേഖരിച്ച ആളുകളിൽ ഏതാനും ചിലരിൽ നിന്നു് പരിശോധകനും അതേ വിവരങ്ങൾ തന്നെ ശേഖരിച്ച് ആദ്യം ശേഖരിച്ചവയുമായി താരതമ്യപ്പെടുത്തുക. വൻതോതിലുള്ള വിവരശേഖരണത്തിൽ പരിശോധകർക്കു് വഹിക്കാനുള്ള പങ്ക് വളരെ പ്രധാനപ്പെട്ട ഒന്നാണ്. അന്വേഷകരെ ശ്രദ്ധിച്ച് തിരഞ്ഞെടുക്കുകയും അവർക്ക് ആവശ്യമായ പരിശീലനം നൽകുകയും അവരുടെ പ്രവർത്തനങ്ങൾ വേണ്ട പോലെ നിയന്ത്രിക്കുകയും ചെയ്താൽ ഈ രീതി പൂർണവും സൂക്ഷ്മവുമായ വിവരങ്ങൾ നല്കുമെന്നുള്ളതിൽ സംശയമില്ല. അല്പം ചെലവു കൂടുമെന്നുള്ള ഒരേ ഒരു പോരായ്മയേ എടുത്തു പറയത്തക്കതായി ഈ സമ്പ്രദായത്തിനുള്ളു.

മൂന്നാമത്തെ സമ്പ്രദായം ഏററവും നല്ല ഒന്നാണ്. പക്ഷേ, സാമ്പിളിലെ അംഗങ്ങളുടെ എണ്ണം വളരെ ചെറുതായിരിക്കുമ്പോൾ മാത്രമേ അതു് പ്രായോഗികമാവുകയുള്ളു. അതുപോലെ തന്നെ വിവരം നൽകേണ്ട വ്യക്തികളെ കണ്ടുപിടിക്കുക അത്ര ബുദ്ധിമുട്ടായിരിക്കുകയുമരുതു്. അന്വേഷണം സംഘടിപ്പിക്കുന്ന ആൾ ആ വിഷയത്തിൽ ഒരു വിദഗ്ധനായിരിക്കുമെന്നതു കൊണ്ട് കൂടുതൽ നന്നായി വിവരങ്ങൾ ശേഖരിക്കാൻ കഴിയും. സാമ്പിളനലക്ഷ്യങ്ങളെപ്പററിയുള്ള വ്യക്തമായ ധാരണ, വിവരശേഖരണത്തിൽ വളരെ സഹായകമായിരിക്കും. ശരിയായ വിവരങ്ങൾ ശേഖരിക്കേണ്ടതിൻെറ ആവശ്യകതയെപ്പററി മററാരെയുംകാൾ അദ്ദേഹത്തിനു് ബോധമുണ്ടായിരിക്കുമല്ലൊ. ചോദ്യാവലി അല്പം നീണ്ടതോ, അവ്യക്തമോ ആയതുകൊണ്ട് വലിയ അപകടമൊന്നും സംഭവിക്കാനില്ല. വ്യക്തിപരമായ പ്രാധാന്യം മൂലം വിവരം നല്കേണ്ടവരുടെ ആദരവു നേടാനും അങ്ങനെ ശരിയായ വിവരങ്ങൾ ശേഖരിക്കാനും കഴിഞ്ഞെന്നു വരും; ഇങ്ങനെ

ഏതു പരിഗണന വെച്ചു നോക്കിയാലും സാമ്പിളനം സംഘടിപ്പിക്കുന്ന ആൾ തന്നെ നേരിട്ട് വിവരം ശേഖരിക്കുന്നിടത്തോളം നല്ലതല്ല മററൊരു രീതിയും. പക്ഷേ സാമ്പിൾ വലുതാകുമ്പോൾ, സാമ്പിളിലെ അംഗങ്ങൾ നാടിൻെറ നാനാ ഭാഗങ്ങളിലായി ചിതറിക്കിടക്കുമ്പോൾ ഈ രീതി തികച്ചും അപ്രായോഗികമായിത്തീരും.

ഓരോ സന്ദർഭത്തിലും ഏതു രീതിയായിരിക്കും കൂടുതൽ നല്ലതെന്ന്, അതിന്റേതായ പ്രത്യേകതകൾ കണക്കിലെടുത്തുകൊണ്ടു വേണം തീരുമാനിക്കുവാൻ. വിവരം ശേഖരിക്കാൻ എത്ര സമയം എടുക്കാമെന്നുള്ളതു് ഒരു പ്രധാന ഘടകമാണ്. പെട്ടെന്നു മാറ്റം വരുന്ന വിവരങ്ങളാണ് ശേഖരിക്കുന്നതെങ്കിൽ വളരെ കുറഞ്ഞ സമയം കൊണ്ടു് വിവരശേഖരണം പൂർത്തിയാക്കണം. വേഗത്തിൽ തീരുമാനങ്ങളെടുക്കുന്നതിനു വേണ്ടിയാണ് വിവരശേഖരണമെങ്കിലും, സ്ഥിതി ഇതു തന്നെ. ഇമ്മാതിരി സന്ദർഭങ്ങളിൽ നിശ്ചിതസമയപരിധിക്കുള്ളിൽ വിവരങ്ങൾ ലഭിക്കാൻ ഉപയുക്തമായ ശേഖരണസംപ്രദായം അവലംബിച്ചേ തീരൂ. കുരുമുളകിന്റെ വിവിധ വിപണികളിലെ വിലനിലവാരം അറിയുകയാണ് ലക്ഷ്യമെന്നിരിക്കട്ടെ. അന്വേഷകൻ തന്നെ ഓരോ വിപണികളിലും പോയി വിവരം ശേഖരിക്കുകയോ പോസ്റ്റൽ സംപ്രദായത്തിൽ വിവരം ശേഖരിക്കുകയോ ചെയ്യുക ഈ സാഹചര്യത്തിൽ സ്വീകാര്യമല്ല. ഓരോ നിമിഷത്തിലും വിലനിലവാരത്തിൽ മാറ്റം വന്നുകൊണ്ടിരിക്കുമല്ലോ. പരിശീലനം സിദ്ധിച്ച അന്വേഷകരെ ഓരോ വിപണിയിലും ഒരേ സമയത്തു് അയച്ചു് വിവരം ശേഖരിച്ചെങ്കിലേ അന്വേഷണലക്ഷ്യം സാക്ഷാൽക്കരിക്കപ്പെടുകയുള്ളു. നേരെ മറിച്ചു്, സാവകാശമായി ശേഖരിച്ചാൽ തകരാറില്ലാത്ത കാര്യങ്ങൾ സൗകര്യമുള്ള രീതി അവലംബിച്ചു ശേഖരിക്കാവുന്നതാണ്.

8. ചോദ്യാവലി

വിവരശേഖരണത്തിൽ ചോദ്യാവലിയുടെ പ്രാധാന്യം ഇതിനകം തന്നെ സൂചിപ്പിച്ചു കഴിഞ്ഞു. സാമ്പിളനമായാലും സെൻസസ്സായാലും ചോദ്യാവലി ആവശ്യമാണ്. വിദഗ്ധന്മാരും പരിചയസമ്പന്നരുമായ ഒന്നിലധികം ആളുകൾ ഒരുമിച്ചിരുന്നു് തയ്യാറാക്കേണ്ടതാണ് ഇതു്. വിവരം ആവശ്യമുള്ള വ്യക്തികളോ സ്ഥാപനങ്ങളോ ആയി ആലോചിക്കുകയും അവരുടെ അംഗീകാരം നേടുകയും ചെയ്തിട്ടു വേണം അതിനു് അവസാനരൂപം നൽകാൻ. വിവരശേഖരണത്തിനു് അംഗീകരിക്കുവാൻ പോകുന്ന സമ്പ്രദായം ചോദ്യാവലിയെ സംബന്ധിച്ചിടത്തോളം ഒരു നിർണായകഘടകമാണ്. അന്വേഷകൻ നേരിട്ടു വിവരം ശേഖരിക്കുകയാണെങ്കിൽ അല്പം സങ്കീർണ്ണമായ ചോദ്യാവലിയായാലും വലിയ തകരാറില്ല. പക്ഷേ അന്വേഷകരെ ഉപയോഗിച്ചോ തപാൽ മുഖേനയോ വിവരം ശേഖരിക്കാൻ ഉദ്ദേശിക്കുന്ന സന്ദർഭങ്ങളിൽ വളരെ കരുതലോടെ യാതൊരു തെററിദ്ധാരണയ്ക്കും ഇടയാക്കാത്ത വിധത്തിൽ ചോദ്യാവലി തയ്യാറാക്കണം. ചോദ്യാവലി തയ്യാറാക്കുന്നതിൽ പ്രത്യേകം ശ്രദ്ധിക്കേണ്ട ചില കാര്യങ്ങൾ താഴെ കൊടുക്കുന്നു.

(1) ചോദ്യങ്ങളുടെ എണ്ണം കഴിവതും കുറഞ്ഞിരിക്കണം

വിവരം നൽകേണ്ട ആളുകൾ നിരവധി ജോലിത്തിരക്കുകളുള്ളവരായിരിക്കും. കൃത്യാന്തരബാഹുല്യത്തിനിടയിൽ സ്ഥിതിവിവരാന്വേഷകനു വേണ്ടി അൽപം സമയം അനുവദിച്ചതു തന്നെ വെറും സൗമനസ്യത്തിന്റെ പേരിൽ ആയിരിക്കും. അവരെ വളരെയധികം ചോദ്യങ്ങൾ ചോദിച്ചു ബുദ്ധിമുട്ടിച്ചാൽ അതിന്റെ പ്രതികരണം അത്ര സന്തോഷപ്രദമായില്ല എന്നു വരാവുന്നതാണ്. ചിലരെങ്കിലും വിവരം നൽകാൻ വിസമ്മതിച്ചെന്നും പലരും അശ്രദ്ധമായി മറുപടി പറഞ്ഞെന്നും വരും. പക്ഷേ ആവശ്യമുള്ള എല്ലാ വിവരങ്ങളും ശേഖരിക്കാതെ തരമില്ലല്ലോ. ബുദ്ധിപൂർവമായി ചോദ്യാവലി തയ്യാറാക്കിയാൽ ഒരളവു വരെ ഇമ്മാതിരി തകരാറുകൾ ഒഴിവാക്കാം. ചോദ്യങ്ങളുടെ എണ്ണം ചുരുങ്ങിയിരിക്കുകയും ഓരോ ചോദ്യവും പരമാവധി വിവരങ്ങൾ നിഷ്കർഷണം ചെയ്യത്തക്കവണ്ണമുള്ളവയായിരിക്കുകയും ചെയ്യാൻ ശ്രദ്ധിക്കണം.

(2) ചോദ്യങ്ങൾ ഋജുവും വ്യക്തവുമായിരിക്കണം

തപാൽ മുഖേന വിവരം ശേഖരിക്കുമ്പോൾ ഇതു് സർവപ്രധാനമാണ്. ചോദ്യങ്ങളുടെ വളച്ചുകെട്ടലും അവ്യക്തതയും തെററായ ഉത്തരങ്ങൾ ലഭിക്കാൻ ഇടയാക്കും. ഒററ വാക്കിൽ ഉത്തരം പറയാവുന്ന ചോദ്യങ്ങൾ, "ശരി" അല്ലെങ്കിൽ "തെററു്" എന്നു് മറുപടി പറയാവുന്ന ചോദ്യങ്ങൾ തുടങ്ങിയവയ്ക്കു് മുൻഗണന നൽകണം. ചോദ്യത്തിനു് നൽകാവുന്ന വിവിധ ഉത്തരങ്ങൾ ചോദ്യത്തോടു കൂടിത്തന്നെ കൊടുക്കുകയും അവയിൽ നിന്നു് ശരിയായ ഉത്തരം തിരഞ്ഞെടുക്കാൻ വിവരം നൽകുന്ന ആളിനോടു് ആവശ്യപ്പെടുകയും ചെയ്യുന്നതു് ഒരു നല്ല രീതിയാണ്. ഉദാഹരണമായി 'വിദ്യാഭ്യാസനിലവാരം എന്താണ്' എന്ന ചോദ്യം അത്ര വ്യക്തമായ ഒന്നല്ല. അതിനു് (1) എഴുതാനും വായിക്കാനും അറിയാം. (2) പ്രൈമറി ക്ലാസുകളിൽ പഠിച്ചിട്ടുണ്ടു് (3) ഹൈസ്കൂളിൽ പഠിച്ചിട്ടുണ്ടു് (4) എസ്. എസ്. എൽ. സി പാസായിട്ടുണ്ടു്. (5) ഡിഗ്രി എടുത്തിട്ടുണ്ടു് (6) ബിരുദാനന്തരപഠനം നടത്തിയിട്ടുണ്ടു് എന്നീ ഉത്തരങ്ങളിൽ ഏതെങ്കിലുമാണു് നൽകാവുന്നതു് എന്നു് വിവരം നൽകുന്ന ആളോടു് പറഞ്ഞാൽ അദ്ദേഹത്തിനു് വളരെ വേഗം മറുപടി നൽകാൻ കഴിയും. മറുപടിയായി എന്താണു് ഉദ്ദേശിക്കുന്നതു് എന്നതിനെപ്പററിയുള്ള അവ്യക്തത ഒഴിവാക്കുകയും ചെയ്യാം.

(3) ചോദ്യങ്ങൾ ശ്രദ്ധാപൂർവം ക്രമീകരിക്കണം

ചോദ്യങ്ങളുടെ ക്രമീകരണവും പ്രധാനപ്പെട്ട ഒന്നാണ്. വിവരം നൽകുന്നവരുടെ താൽപര്യം വളർത്തിക്കൊണ്ടു വരത്തക്കവണ്ണം ചോദ്യങ്ങൾ ക്രമീകരിക്കണം. മറുപടി പറയാൻ ബുദ്ധിമുട്ടുള്ള ചോദ്യങ്ങൾ ആദ്യം ചോദിക്കരുതു്. അപഗ്രഥനഘട്ടത്തിൽ പട്ടികകളും മററും തയ്യാറാക്കാൻ എളുപ്പം കിട്ടണം എന്നതും ചോദ്യങ്ങളുടെ ക്രമം നിശ്ചയിക്കുന്നതിൽ ശ്രദ്ധിക്കേണ്ട ഒരു കാര്യമാണ്. ഉത്തരങ്ങൾ സംഖ്യകൾ കൊണ്ടോ, സൂചകചിഹ്നങ്ങൾ കൊണ്ടോ രേഖപ്പെടുത്തുന്നതിനുള്ള ക്രമീകരണങ്ങൾ ചെയ്യുന്നതും വിവരശേഖരണം എളുപ്പമാക്കി

തീക്കും. മുൻ ഉദാഹരണത്തിൽ തന്നെ വിദ്യാഭ്യാസനിലവാരത്തെപ്പററിയുള്ള ചോദ്യത്തിൻെറ ഉത്തരം 'ബിരുദാനന്തരപഠനം' നടത്തിയിട്ടുണ്ടു് എന്നു് ആണെങ്കിൽ അതു് ആ ഉത്തരത്തിൻെറ ക്രമസംഖ്യയായ '6' കൊണ്ടു സൂചിപ്പിക്കാവുന്നതാണു്.

(4) വിവരം നൽകുന്ന ആളിൻെറ സ്വകാര്യ ജീവിതത്തെ സ്പർശിക്കുന്ന ചോദ്യങ്ങൾ ഒഴിവാക്കണം. അയാൾ വെളിപ്പെടുത്താൻ ഇഷ്ടപ്പെടാത്തതൊ, അയാളുടെ മതവികാരങ്ങളെ വ്രണപ്പെടുത്തുന്നതൊ ആയ ചോദ്യങ്ങൾ ചോദിക്കരുതു്

വളരെ പ്രധാനപ്പെട്ട ഒരു കാര്യമാണു് ഇതു്. വിവരം നൽകുന്ന ആളിന്റെ അതൃപ്തി സമ്പാദിക്കത്തക്ക ചോദ്യങ്ങൾ ചോദിച്ചു തുടങ്ങിയാൽ ആ അന്വേഷണം പരാജയപ്പെട്ടുമെന്നുള്ളതു തീർച്ചയാണു്. അങ്ങനെയുള്ള ഏതെങ്കിലും വിവരങ്ങൾ അറിയേണ്ടതു് ആവശ്യമായി വരുമ്പോൾ വളരെ ശ്രദ്ധിച്ചു് ചോദ്യങ്ങൾ തയ്യാറാക്കണം.

(5) ചോദ്യങ്ങളിൽ ഉപയോഗിച്ചിരിക്കുന്ന പദങ്ങളുടെ അർഥവും വ്യാപ്തിയും വ്യക്തമായി വിശദീകരിക്കുന്ന കുറിപ്പുകൾ ചോദ്യാവലിയോടൊപ്പമുണ്ടായിരിക്കണം

തപാൽ മുഖേന വിവരം ശേഖരിക്കുമ്പോൾ ഇതു വളരെ അത്യാവശ്യമാണു്. മററു വിവരശേഖരണസമ്പ്രദായങ്ങൾ അവലംബിച്ചാലും ഇമ്മാതിരി വിശദീകരണക്കുറിപ്പുകൾ ആവശ്യമാണു്. ചോദ്യം തയ്യാറാക്കിയ ആൾ എന്താണു് ഉദ്ദേശിച്ചിരിക്കുന്നതെന്നു് വിവരം ശേഖരിക്കുന്നവക്കും നൽകുന്നവക്കും മനസ്സിലാകണമല്ലൊ. ഇമ്മാതിരി വിശദീകരണങ്ങൾ വളരെയധികമുണ്ടെങ്കിൽ അതു് ഒരു വിശദീകരണപത്രികയായി വേറെ നൽകുന്നതും കൊള്ളാം.

പൊതുവെ പറഞ്ഞാൽ, ആവശ്യമായ വിവരങ്ങൾ ഏററവും കുറച്ചു ചോദ്യങ്ങൾ ചോദിച്ചു്, വിവരം നൽകുന്ന ആളിനു ബുദ്ധിമുട്ടു തോന്നാത്ത വിധത്തിലും തെററുകൾ കഴിവതും ഒഴിവാക്കത്തക്ക വിധത്തിലും ശേഖരിക്കാൻ കഴിയണം എന്നതായിരിക്കണം ചോദ്യാവലിക്കു് രൂപം കൊടുക്കുമ്പോഴത്തെ മൗലികലക്ഷ്യം.

9. വിവരശേഖരണത്തിലെ ചില പ്രായോഗിക വൈഷമ്യങ്ങൾ

സാമ്പിളനത്തിൻെറ സാങ്കേതികാംശങ്ങളെല്ലാം കുററമററതായിരുന്നാൽ തന്നെയും പ്രായോഗികമായ ചില ബുദ്ധിമുട്ടുകൾ കാരണം പല പോരായ്മകളും സംഭവിക്കാം. ഒരുപക്ഷെ സാമ്പിളനം പരാജയപ്പെട്ടു എന്നു തന്നെ വരാം. സാമ്പിളിലെ എല്ലാ അംഗങ്ങളിൽ നിന്നും വിവരങ്ങൾ ശേഖരിക്കാൻ സാധിക്കാതെ വരിക എന്നതാണു് ഇതിൽ ഏററവും പ്രധാനപ്പെട്ടതു്. തപാൽ മുഖേനയുള്ള വിവരശേഖരണത്തിലാണു് ഇതു് കൂടുതലും സംഭവിക്കുന്നതു്. അന്വേഷകരെ ഉപയോഗിച്ചുള്ള വിവരശേഖരണസമ്പ്രദായം അവലംബിച്ചാൽ ഈ ബുദ്ധിമുട്ടു് ഒരു പരിധി വരെ ഒഴിവാക്കാൻ കഴിയും. എന്നിരുന്നാലും താഴെപ്പറയുന്ന കാരണങ്ങൾ കൊണ്ടു് ഇതു് സംഭവിക്കാവുന്നതാണു്.

(1) വിവരം നൽകേണ്ട ആളുകളെ അന്വേഷകർക്കു് കണ്ടെത്താൻ കഴിയാതെ വരിക

അന്വേഷകർ വിവരശേഖരണത്തിനായി എത്തിച്ചേരുന്ന സമയം വളരെ പ്രധാനപ്പെട്ട ഒന്നാണ്. വിദ്യാർഥികളിൽ നിന്നും ഉദ്യോഗസ്ഥന്മാരിൽ നിന്നും മററും വിവരം ശേഖരിക്കാൻ പത്തു മണിക്കും നാലു മണിക്കും മധ്യെ വീടുകളിൽ ചെന്നതു കൊണ്ടു പ്രയോജനമില്ല. സായാഹ്നങ്ങളിലൊ അവധി ദിവസങ്ങളിലൊ വേണം അങ്ങനെയുള്ളവരെ സമീപിക്കാൻ. പക്ഷേ, ഓരോരുത്തരുടെയും സൗകര്യമനുസരിച്ച് അവരുടെ താമസസ്ഥലങ്ങളിൽ ചെന്നു് വിവരം ശേഖരിക്കുന്നതിൽ അന്വേഷകൻ അനുഭവിക്കേണ്ടി വരുന്ന ബുദ്ധിമുട്ടു് അവഗണിക്കാവുന്നതല്ല. ഒരാളെ കാണാൻ തന്നെ പല പ്രാവശ്യം ഒരിടത്തു പോകേണ്ടതായി വരും. ചിലർ വീട്ടിലിരിക്കുന്ന സ്വഭാവക്കാരായിരിക്കുകയില്ല. മററു ചിലർക്കു് സാവകാശവും സൗകര്യവുമുള്ളപ്പോഴല്ലാതെ വിവരം നല്ലാൻ സന്മനസ്സുണ്ടായെന്നു വരികയില്ല. മററു ചിലർ യാത്രയിലായിരിക്കും. എന്നാണു് അവർ തിരിച്ചെത്തുകയെന്നു് നിശ്ചയമുണ്ടായിരിക്കുകയില്ല. ഇങ്ങനെ പല കാരണങ്ങൾ കൊണ്ടും അന്വേഷകക്കു് വിവരം നല്ലേണ്ടവരെ കണ്ടെത്താൻ കഴിയാതെ വരാം. അങ്ങനെ സംഭവിച്ചാൽ അതു് സാമ്പിളനലക്ഷ്യങ്ങളെ പരാജയപ്പെടുത്തിയേക്കും. ഈ അപാകടം ഒഴിവാക്കാൻ വിവരശേഖരണഘട്ടത്തിൽ കഴിവതും ശ്രമിക്കേണ്ടതാണു്. എഴുത്തു മുഖേനയൊ ഫോൺ മുഖേനയൊ സന്ദർശനം ഏർപ്പാടു ചെയ്തതിനു ശേഷം വിവരശേഖരണത്തിനായി സമീപിക്കുന്നതു്, പ്രത്യേകിച്ചും ജോലിത്തിരക്കുള്ളവരിൽ നിന്നു് വിവരം ശേഖരിക്കാൻ വളരെ ഉപകരിക്കും. വിദ്യാർഥികളെ സ്കൂൾസമയത്തു് സ്കൂളിൽ ചെന്നു് സന്ദർശിക്കുന്നതും, തൊഴിലാളികളെ അവരുടെ പണിസ്ഥലത്തു ചെന്നു സന്ദർശിക്കുന്നതും സൗകര്യപ്രദമായിരിക്കും. വിവരം നല്ലേണ്ട ആൾ ദീർഘകാലത്തേക്കു് സ്ഥലത്തു് ഇല്ലാതെ വന്നാൽ, അയാളെ അടുത്തറിയാവുന്ന മററാരിൽ നിന്നെങ്കിലും ആവശ്യമായ വിവരങ്ങൾ ചോദിച്ചറിയാം. പക്ഷേ ഇതു് ഒഴിവാക്കാൻ കഴിവതും ശ്രമിക്കണം. ഇങ്ങനെ കരുതലോടും ബുദ്ധിമുട്ടുകൾ കാര്യമാക്കാതെയും പരിശ്രമിച്ചാൽ ഈ പ്രായോഗികവൈഷമ്യം കുറെയെല്ലാം പരിഹരിക്കാം.

(2) വിവരം നൽകാൻ വിസമ്മതിക്കുക

ഇതു് പല കാരണങ്ങൾ കൊണ്ടു് ഉണ്ടാവാം. പല നിലകളിലുള്ളവരും പല സ്വഭാവക്കാരുമായ ആളുകളിൽ നിന്നു് വിവരങ്ങൾ ശേഖരിക്കേണ്ടതു് ആവശ്യമായി തീരുമല്ലൊ. അവരെയെല്ലാം ഒരുപോലെ സമീപിച്ചാൽ പ്രയോജനമുണ്ടാവുകയില്ല. ജോലി ചെയ്തു ക്ഷീണിച്ചിരിക്കുമ്പോഴൊ തിരക്കുള്ളപ്പോഴൊ ആണു് അന്വേഷകൻ എത്തിച്ചേരുന്നതെങ്കിൽ വിവരം നൽകാൻ അയാൾ വിസമ്മതിച്ചെന്നു വരും. സന്ദർഭം മനസ്സിലാക്കി അതിനനുസരിച്ച് പ്രവർത്തിക്കാൻ അന്വേഷകനുള്ള കഴിവാണു് ഇമ്മാതിരി സന്ദർഭങ്ങളിൽ പ്രയോജനപ്പെടുന്നതു്. മറെറാരു സന്ദർഭത്തിൽ ആ ആളെ സമീപിച്ചാൽ പ്രയാസം കൂടാതെ വിവരങ്ങൾ ലഭിച്ചെന്നു വരും. അന്വേഷകൻെറ പെരുമാററം വിവരം നല്ലേണ്ട ആൾക്കു

രസിച്ചില്ലെന്ന ഒററ കാരണം കൊണ്ടു് ഒരാൾ വിവരം നല്കാൻ വിസമ്മതിച്ചെന്നു വരാം. അങ്ങനെയുള്ള സന്ദർഭങ്ങളുണ്ടാകാതിരിക്കാൻ പ്രത്യേകം ശ്രദ്ധിക്കണം. അതിനു് ആവശ്യമായ നിർദേശങ്ങൾ അന്വേഷകർക്കു് നേരത്തെ തന്നെ കൊടുക്കണം. ചോദ്യങ്ങളുടെ സ്വഭാവമാവാം ചിലപ്പോൾ വിസമ്മതത്തിനു കാരണം. തൻെറ രഹസ്യങ്ങൾ ചോർത്തിയെടുക്കുവാനുള്ള ഒരു ശ്രമമായി ഈ അന്വേഷണത്തെ തെററിദ്ധരിക്കാനും അതുകൊണ്ടു തന്നെ വിവരം നൽകാൻ വിസമ്മതിക്കാനും ഇടയുണ്ടു്. ചോദ്യാവലി തയ്യാറാക്കുമ്പോൾ അമ്മാതിരി ചോദ്യങ്ങൾ കഴിവതും ഒഴിവാക്കാൻ ശ്രമിക്കണം. ചിലതെല്ലാം ഉൾപ്പെടുത്താൻ നിർബന്ധിതരായാൽ തന്നെ നല്കുന്ന വിവരങ്ങൾ നിയമപരമായൊ മറേറതെങ്കിലും തരത്തിലൊ വിവരം നല്കുന്ന ആളിനു് എതിരായി ഉപയോഗിക്കുകയില്ലെന്നു് ഉറപ്പു കൊടുക്കുക, നൽകപ്പെടുന്ന വിവരങ്ങൾ പരമരഹസ്യമായി സൂക്ഷിക്കുമെന്നു് ബോധ്യപ്പെടുത്തുക തുടങ്ങിയ പരിഹാരമാർഗങ്ങൾ കൂടി സ്വീകരിക്കണം. ചില കാര്യങ്ങൾ പരസ്യമായി പറയാനുള്ള മടിയാവാം പലപ്പോഴും വിസമ്മതത്തിനു കാരണം. കുടുംബാസൂത്രണപരിപാടികളെപ്പററിയും മററുമുള്ള അന്വേഷണങ്ങളിൽ ഇതു് സംഭവിക്കാറുണ്ടു്. സന്ദർശനം കഴിവതും രഹസ്യമാക്കാൻ ശ്രമിച്ചാൽ ഈ ബുദ്ധിമുട്ടു് ഒഴിവാക്കാം. വിവരം ശേഖരിക്കുന്ന സ്ഥാപനത്തേയൊ വ്യക്തിയേയൊ പററിയുള്ള മതിപ്പാണു് വിവരം നൽകുന്നവരുടെ മനോഭാവത്തിൽ സ്വാധീനം ചെലുത്തുന്ന മറെറാരു ഘടകം. പ്രശസ്തരായ വ്യക്തികളൊ സ്ഥാപനങ്ങളൊ നടത്തുന്ന അന്വേഷണത്തിൽ വിവരം നൽകാൻ ചുരുക്കം ആളുകളേ വിസമ്മതം കാണിക്കാറുള്ളു.

ഇവിടെ ചർച്ച ചെയ്ത കാര്യങ്ങളെല്ലാം മനസ്സിൽ വെച്ചുകൊണ്ടു് വിവരം ശേഖരിക്കാൻ ശ്രമിച്ചാൽ വിസമ്മതക്കാരുടെ എണ്ണം വളരെ കുറയ്ക്കാവുന്നതാണു്.

(3) വിവരം നൽകാനുള്ള കഴിവില്ലായ്മ

മാനസികമൊ ശാരീരികമൊ ആയ രോഗപീഡ കൊണ്ടു ചിലപ്പോൾ ഒരാൾക്കു് വിവരം നൽകാൻ കഴിഞ്ഞില്ലെന്നു വരും. അന്വേഷകർ മറേറതെങ്കിലും വിധത്തിൽ ആവശ്യമുള്ള വിവരങ്ങൾ ശേഖരിക്കുകയേ ഇമ്മാതിരി ചുററുപാടിൽ കരണീയമായിട്ടുള്ളു. ഭാഷാപരമായ ബുദ്ധിമുട്ടു് പലപ്പോഴും വിവരം ശേഖരിക്കാൻ തടസ്സമായി എന്നു വരാം. വിവരം നൽകേണ്ട ആളിനു് വശമുള്ള ഭാഷ അന്വേഷകനും അറിയാമായിരിക്കണമെന്നു് നിഷ്കർഷിച്ചാൽ ഈ ബുദ്ധിമുട്ടു് ഒഴിവാക്കാം. ചോദ്യാവലിയിലെ ചില ചോദ്യങ്ങളുടെ ഉത്തരം വിവരം നല്കുന്ന ആളിനു നിശ്ചയമില്ലെന്നും വരാവുന്നതാണു്. ഉദാഹരണമായി, കൃത്യമായി കണക്കു സൂക്ഷിക്കാത്തവർക്കു് വാർഷികവരുമാനം എന്തെന്നു് അത്ര നിശ്ചയം കാണുകയില്ല. ഇങ്ങനെയുള്ള അവസരങ്ങളിൽ അന്വേഷകൻ കൂടി സഹകരിച്ചാൽ ഏതാണ്ടു സൂക്ഷ്മമായി ആവശ്യമുള്ള ഉത്തരങ്ങൾ കണ്ടുപിടിക്കാൻ കഴിയും.

(4) ചോദ്യാവലി അപൂർണമായി പൂരിപ്പിക്കപ്പെടുകയൊ പൂരിപ്പിച്ച ഏതാനും ചോദ്യാവലികൾ നഷ്ടപ്പെടുകയൊ ചെയ്യുക

അന്വേഷകരും പരിശോധകരും അന്വേഷണവുമായി ബന്ധപ്പെട്ട മററു ആളുകളും ഈ തകരാറു് ഒഴിവാക്കാൻ പരമാവധി പരിശ്രമിക്കേണ്ടതാണു്.

10. ചെലവിൻെറ പ്രശ്നം

സാമ്പിളനത്തെ സംബന്ധിച്ചിടത്തോളം മർമപ്രധാനമായ ഒന്നാണു് ഇതു്. ഏററവും കുറഞ്ഞ ചെലവിൽ ആവശ്യമുള്ള വിവരം ശേഖരിക്കുവാനാണു് സാമ്പിളനശാസ്ത്രം ശ്രമിക്കുന്നതു്. സാമ്പിളനത്തെ സംബന്ധിക്കുന്ന പ്രധാനപ്പെട്ട എല്ലാ തീരുമാനങ്ങളും ചെലവു കൂടി പരിഗണിച്ചു വേണം സ്വീകരിക്കുവാൻ. സൂക്ഷ്മത വർധിപ്പിക്കാൻ ശ്രമിക്കുമ്പോൾ ചെലവും വർധിക്കുമെന്നുള്ളതു് ഒരു സാമാന്യതത്വമാണു്. സൂക്ഷ്മത നിർണയിച്ചിട്ടു് അതിനനുസരിച്ചു് ഏററവും കുറഞ്ഞ ചെലവു് നിശ്ചയിക്കുകയോ ചെലവു് നിശ്ചയിച്ചിട്ടു് അതു കൊണ്ടു് നേടാവുന്ന ഏററവും കൂടിയ സൂക്ഷ്മത നിർണയിക്കുകയോ ചെയ്യാവുന്നതാണു്.

11. സംഗ്രഹം

1. സാമ്പിളനത്തിലും സെൻസസ്സിലും അടിസ്ഥാനപ്രമാണങ്ങൾ ഒന്നു തന്നെയാണു്. രണ്ടിനും ശ്രദ്ധാപൂർവമുള്ള ആസൂത്രണം ആവശ്യമത്രെ. നിർവഹണഘട്ടത്തിൽ ശ്രദ്ധിക്കേണ്ട കാര്യങ്ങളും ഏതാണ്ടു് ഒന്നു തന്നെ.

2, ആസൂത്രണഘട്ടത്തിൽ ഏററവും പ്രധാനപ്പെട്ടതു് അന്വേഷകൻെറ ലക്ഷ്യബോധമാണു്. പലപ്പോഴും ലക്ഷ്യങ്ങളിൽ പരിഷ്ക്കാരങ്ങൾ നടത്തേണ്ടിവരും. ഉപയോഗം, സാമ്പത്തികശേഷി എന്നിവ പരിഗണിച്ചാണു് ഇതു് നിർവഹിക്കുന്നതു്.

3. സമഷ്ടിയുടെ നിർവചനം അതിവ്യാപ്തി—അവ്യാപ്തിദോഷങ്ങളില്ലാത്ത വിധത്തിലും വിവരശേഖരണക്കാർക്കു് വ്യക്തികളെ തിരിച്ചറിയാൻ കഴിയത്തക്ക വിധത്തിലും നിർണയിക്കണം.

4. പ്രാഥമികമായി വിവരങ്ങൾ ശേഖരിക്കണമോ അതോ ആവശ്യമുള്ളത്ര സൂക്ഷ്മതയുള്ള ദ്വിതീയക വിവരങ്ങൾ ലഭ്യമാണോ എന്നു പരിശോധിക്കണം. ദ്വിതീയകവിവരങ്ങൾ ലഭ്യമാണോ എന്നു പരിശോധിക്കണം. ദ്വിതീയക വിവരങ്ങൾ സ്വീകരിക്കുമ്പോൾ ആ വിവരങ്ങൾ എന്തു ലക്ഷ്യത്തോടു കൂടി ശേഖരിച്ചതാണെന്നും സമഷ്ടി എങ്ങനെയാണു് നിർവചിക്കപ്പെട്ടിരുന്നതെന്നും ചോദ്യാവലിയിലെ പദങ്ങളുടെ നിർവചനം എന്തായിരുന്നു എന്നും ഏതു കാലത്തു് ആരു് നടത്തിയ അന്വേഷണത്തിൽ നിന്നു് ലഭിച്ചതാണു് ആ വിവരങ്ങൾ എന്നും മററും പരിഗണിക്കണം.

6. സാമ്പിളന ഫ്റേം തയ്യാറാക്കലാണു് പ്രാഥമിക അന്വേഷണത്തിൽ ആദ്യം ചെയ്യേണ്ടതു്. തത്വത്തിൽ ഇതു് സമഷ്ടിയിലെ പ്രാഥമികവ്യക്തികളുടെ പട്ടികയാണു്. പക്ഷേ ഇതു് ഒരു ഭൂപടമോ വ്യക്തികളെപ്പറ്റിയുള്ള ചില വിവരങ്ങളൊ മറ്റേതെങ്കിലും സൂചനകളൊ ആവാം. ഫ്രേമിൻെറ പൂർണത, തെറ്റില്ലായ്മ ഓരോ വ്യക്തിയെപ്പററിയും അതു നൽകുന്ന ആനുഷംഗികവിവരങ്ങൾ എന്നിവ

സുപ്രധാനങ്ങളാണ്. പഴയ ഫ്രേമുകൾ ലഭ്യമാണെങ്കിൽ അവയിലെ തെററു തിരുത്തി നവീകരിച്ചിട്ടു വേണം സാമ്പിളെടുക്കാൻ.

7. അംഗീകരിക്കുന്ന സാമ്പിളനരീതിയനുസരിച്ച് ഈ ഫ്രേമിൽ നിന്ന് സാമ്പിളിൽ ഉൾപ്പെടുത്തേണ്ട അംഗങ്ങളെ തിരഞ്ഞെടുക്കണം.

8. വിവരശേഖരണം തപാൽ മുഖേനയൊ, പരശീലനം സിദ്ധിച്ച അന്വേഷകർ മുഖേനയൊ, അന്വേഷകൻ നേരിട്ടൊ ആവാം. ഓരോന്നിനും അതിൻെറതായ ഗുണങ്ങളും ദോഷങ്ങളുമുണ്ട്.

9. ചോദ്യാവലി തയ്യാറാക്കുന്നതിൽ പ്രത്യേകം ശ്രദ്ധിക്കണം. ചോദ്യങ്ങൾ വ്യക്തവും ശ്രദ്ധാപൂർവം ക്രമീകരിച്ചിട്ടുള്ളവയും വിവരം നല്കുന്ന ആളിൻെറ വികാരങ്ങളെ ഒരു തരത്തിലും വ്രണപ്പെടുത്താത്തവയും ആയിരിക്കണം. ചോദ്യങ്ങളിൽ ഉപയോഗിച്ചിരിക്കുന്ന പദങ്ങൾ ശരിക്കും നിർവചിക്കപ്പെട്ടിരിക്കണം.

10. വിവരം നൽകേണ്ട ആളുകളെ കണ്ടെത്താൻ കഴിയാതെ വരിക, കണ്ടെത്തിയവർ വിവരം നൽകാൻ വിസമ്മതിക്കുന്നവരോ കഴിവില്ലാത്തവരോ ആയിരിക്കുക തുടങ്ങി വിവരശേഖരണത്തിനു നേരിട്ടേക്കാവുന്ന പ്രതിബന്ധങ്ങൾക്ക് നേരത്തെ തന്നെ പരിഹാരങ്ങൾ നിർദേശിച്ചിരിക്കണം.

11. സാമ്പിളിനത്തിനു അനുവദിക്കാവുന്ന ചെലവും സുപ്രധാനമായ ഒന്നാണ്. അനുവദിക്കാവുന്ന ചെലവിൽ കിട്ടാവുന്ന ഏററവും കൂടിയ സൂക്ഷ്മത ഉന്നം വെച്ചാണ് സാമ്പിളനം സംഘടിപ്പിക്കേണ്ടതു്.

ചില സഹായകഗ്രന്ഥങ്ങളും ലേഖനങ്ങളും

1. Survey Sampling (1965) Ch. I, XIII, Leslie Kis E. John Wiley & Sons
2. Some Theory of Sampling (1950) Deming W, S. John Wiley & Sons
3. 'A Sample Survey with both postal and interview stage' — Gray P. G. (1957) Applied Statistics 6, 139—153
4. 'Influence of the Interviewer on accuracy of Survey results' Hanson R. H. (1958) J. A. S. 53' 635-655
5. 'Comparison of three imformation gathering Strategies in a population study of sociomedical variables' Hochstim J. R. (1962) Proceedings of the social statistics section, American Statistical Association 154—159

അഭ്യാസം 2

1. സാമ്പിളനത്തിനു മുൻപായി എന്തെല്ലാം കാര്യങ്ങളാണ് പരിഗണിക്കേണ്ടതു് ?
2. ചോദ്യാവലി തയ്യാറാക്കുന്നതിൽ പ്രത്യേകം ശ്രദ്ധിക്കേണ്ട കാര്യങ്ങൾ ഏതെല്ലാം ?
3. ഇടത്തരക്കാരുടെ ജീവിതച്ചെലവുകളെപ്പററി അന്വേഷണം നടത്തുവാൻ ആവശ്യമായ പ്രാരംഭപ്രവർത്തനങ്ങൾ വിവരിക്കുക.

4. താഴെപ്പറയുന്ന അന്വേഷണങ്ങളിലെ സമഷ്ടിയും വ്യക്തിയും നിർവചിക്കുക.

1. വിദ്യാർഥികളുടെ ഇടയിലെ അസ്വസ്ഥതയ്ക്കു കാരണം
2. റേഡിയൊവിൽ പുതുതായി ഏർപ്പെടുത്തിയ ഒരു പരിപാടിയെപ്പററി കേൾവിക്കാരുടെ അഭിപ്രായം
3. നെൽവിളവും പുതിയ കൃഷിരീതികളുമായുള്ള ബന്ധം
4. കുരുമുളകിൻെറ വാർഷികോല്പാദനം കേരളത്തിൽ
5. ബസ്സുകളിലെ ആൾത്തിരക്ക്.

5. മുൻപറഞ്ഞ അന്വേഷണങ്ങളിലെ സാമ്പിളനഫ്രേം എങ്ങനെ തയ്യാറാക്കും? ഓരോന്നിനും ഉപയുക്തമായ ചോദ്യാവലികൾ തയ്യാറാക്കുക.
6. സാമ്പിളനേതരപിശകുകൾ ഏതെല്ലാം? അവ ഒഴിവാക്കാനുള്ള മാർഗങ്ങൾ നിർദേശിക്കുക.
7. വിവരശേഖരണത്തിനു് അവലംബിക്കാവുന്ന മാർഗങ്ങൾ താരതമ്യപ്പെടുത്തുക.

3

ലഘുയാദൃച്ഛിക സാമ്പിളനം

1. പ്രാരംഭം

സാമ്പിളനത്തിന്റെ പ്രാധാന്യവും പ്രായോഗികസൗകര്യവും നാം കണ്ടു കഴിഞ്ഞു. വിവിധ സാമ്പിളനരീതികളാണ് ഇനി ചർച്ച ചെയ്യുവാനുള്ളത്. സമഷ്ടിയുടെ ഏതൊരംശവും ഒർത്ഥത്തിൽ സാമ്പിളാണ്. പക്ഷേ അതു് സമഷ്ടിയെ പ്രതിനിധീകരിക്കുന്നതാണെങ്കിൽ മാത്രമേ സാമ്പിളെന്നു് സാങ്കേതികമായി വ്യവഹരിക്കപ്പെടുകയുള്ളു. അങ്ങനെയുള്ള സാമ്പിളുകളെ എങ്ങനെ തിരഞ്ഞെടുക്കാം എന്നതാണ് പ്രശ്നം. സംഭാവ്യതാസാമ്പിളനം, ഐച്ഛികസാമ്പിളനം എന്ന രണ്ടു് സാമ്പിളനരീതികളെപ്പററി ഒന്നാമധ്യായത്തിൽ പരാമർശിക്കുകയുണ്ടായി. സംഭാവ്യതാസാമ്പിളനത്തിന്റെ മേന്മകൾ സൂചിപ്പിക്കുകയും ചെയ്തു. ലളിതവും പ്രാഥമികവുമായ ഒരു സംഭാവ്യതാസാമ്പിളനരീതിയാണ് ലഘുയാദൃച്ഛികസാമ്പിളനം. അതോടു ബന്ധപ്പെട്ട വിവിധ വിഷയങ്ങളാണ് ഈ അധ്യായത്തിൽ ചർച്ച ചെയ്യുന്നതു്.

ഈ ഗ്രന്ഥത്തിൽ പ്രധാനമായും സാന്തസമഷ്ടികളാണ് പരിഗണിക്കുന്നതെന്നു് മുൻപു തന്നെ സൂചിപ്പിച്ചിട്ടുണ്ടല്ലോ. ഒരു സമഷ്ടിയിൽ N അംഗങ്ങളാണ് ഉള്ളതു് എന്നിരിക്കട്ടെ അതിൽ നിന്നു് N അംഗങ്ങളുള്ള ഒരു സാമ്പിൾ എടുക്കുവാനാണ് ഉദ്ദേശിക്കുന്നതെന്നും കരുതുക. സമഷ്ടിയിൽ നിന്നു് ഓരോന്നായി N അംഗങ്ങളെ തിരഞ്ഞെടുക്കുകയാണു ചെയ്യേണ്ടതു്. ഇതിൽ ഓരോ ഘട്ടത്തിലും തിരഞ്ഞെടുത്ത അംഗത്തെ തിരികെ സമഷ്ടിയിൽ ഉൾപ്പെടുത്തിയിട്ടാണ് അടുത്ത അംഗത്തെ തിരഞ്ഞെടുക്കുന്നതെങ്കിൽ ആ സാമ്പിളിനസമ്പ്രദായത്തിനു് പ്രതിസ്ഥാപനത്തോടുകൂടിയ സാമ്പിളനം എന്നു പറയും. ഈ രീതിയിൽ ഒരേ അംഗം തന്നെ ഒന്നിലധികം പ്രാവശ്യം തിരഞ്ഞെടുക്കപ്പെട്ടു എന്നു വരാം. സാമ്പിളനപ്രക്രിയയുടെ ഓരോ ഘട്ടത്തിലും അതുവരെ തിരഞ്ഞെടുക്കപ്പെട്ട അംഗങ്ങളെ സമഷ്ടിയിൽ നിന്നു് ഒഴിവാക്കിക്കൊണ്ടാണ് അടുത്ത അംഗത്തെ തിരഞ്ഞെടുക്കുന്നതെങ്കിൽ അതിനു് പ്രതിസ്ഥാപനമില്ലാത്ത സാമ്പിളനം എന്നു പേർ. ഈ സമ്പ്രദായമാണു സ്വീകരി

ക്കുന്നതെങ്കിൽ സമഷ്ടിയിലെ ഒരംഗത്തിനു ഒരു സാമ്പിളിൽ ഒന്നിലധികം പ്രാവശ്യം ഉൾപ്പെടുവാനാവില്ല. ഉദാഹരണമായി A, B, C, D എന്നു് നാലു് അംഗങ്ങളാണു് സമഷ്ടിയിലുള്ളതെന്നു കരുതുക. രണ്ടു് അംഗങ്ങളുള്ള ഒരു സാമ്പിളാണു് എടുക്കേണ്ടതെന്നും സങ്കല്പിക്കുക. പ്രതിസ്ഥാപനത്തോടു കൂടിയ സാമ്പിളനരീതിയാണു് അവലംബിക്കുന്നതെങ്കിൽ നമുക്കു കിട്ടുന്നതു് താഴെ കൊടുത്തിരിക്കുന്ന പതിനാറു സാമ്പിളുകളിൽ ഏതെങ്കിലും ഒന്നായിരിക്കും.

AA, AB, AC, AD, BA, BB, BC, BD, CA, CB, CC, CD, DA, DB, DC, DD (ഇവിടെ സാമ്പിളിലെ അംഗങ്ങളുടെ ക്രമം കൂടി പരിഗണിച്ചിരിക്കുന്നു).

പ്രതിസ്ഥാപനമില്ലാത്ത സാമ്പിളനത്തിൽ ലഭിക്കാവുന്നതു് താഴെ കൊടുത്തിരിക്കുന്ന ആറു് സാമ്പിളുകളിൽ ഏതെങ്കിലും ഒന്നാണു്.

AB, AC, AD, BC, BD, CD (ഇവിടെ വ്യത്യസ്തങ്ങളായ സാമ്പിളുകൾ മാത്രമേ പരിഗണിച്ചിട്ടുള്ളൂ.) ഇതിൽ ഏതു രീതിയാണു് സാമ്പിളനത്തിൽ അവലംബിക്കേണ്ടതെന്നു് ആദ്യമായി തീരുമാനിക്കുക. അടുത്തതു് സാമ്പിളെടുക്കലാണു്. മുൻ ഉദാഹരണത്തിലെ സമഷ്ടി തന്നെ പരിഗണിക്കാം. പ്രതിസ്ഥാപനത്തോടു കൂടിയ സാമ്പിളാണു് എടുക്കുന്നതെന്നിരിക്കട്ടെ. നമുക്കു കിട്ടുന്ന സാമ്പിൾ മുൻപറഞ്ഞ പതിനാറു സാമ്പിളുകളിൽ ഏതെങ്കിലും ഒന്നായിരിക്കും. അവയിൽ ഏതു ലഭിക്കാനുമുള്ള സംഭാവ്യത തുല്യമായിരിക്കത്തക്ക വണ്ണം ആസൂത്രണം ചെയ്യുന്ന സാമ്പിളനസമ്പ്രദായത്തിനു് പ്രതിസ്ഥാപനത്തോടു കൂടിയ ലഘ്വയാദൃച്ഛിക സാമ്പിളനം എന്നു പേർ പറയും. അങ്ങനെ ലഭിക്കുന്ന സാമ്പിളിനു് പ്രതിസ്ഥാപനത്തോടു കൂടിയ ലഘ്വയാദൃച്ഛിക സാമ്പിൾ എന്നു പേർ. ഇതു പോലെ തന്നെ പ്രതിസ്ഥാപനമില്ലാത്ത ലഘ്വയാദൃച്ഛിക സാമ്പിളനവും നിർവചിക്കാവുന്നതാണു്. ചുരുക്കത്തിൽ നിർദിഷ്ട സമഷ്ടിയിൽ നിന്നു് ലഭിക്കാവുന്ന നിർദിഷ്ട വലുപ്പമുള്ള സാമ്പിളുകൾക്കെല്ലാം തുല്യ സംഭാവ്യത വരത്തക്കവണ്ണം തിരഞ്ഞെടുക്കുന്ന സാമ്പിൾ ലഘ്വയാദൃച്ഛിക സാമ്പിൾ എന്നു് വ്യവഹരിക്കപ്പെടും. ഇതിൽ പ്രതിസ്ഥാപനമുള്ളതു്, പ്രതിസ്ഥാപനമില്ലാത്തതു് എന്നു രണ്ടു് തരമുണ്ടു്. ഇതിൽ പ്രതിസ്ഥാപനമില്ലാത്ത ലഘ്വയാദൃച്ഛിക സാമ്പിളിനു് താരതമ്യേന സൂക്ഷ്മത കൂടുമെന്നു് വഴിയെ വ്യക്തമാവും. അതുകൊണ്ടു് അമ്മാതിരി സാമ്പിളനമാണു് ആദ്യമായി ചർച്ച ചെയ്യുന്നതു്.

ലഘ്വയാദൃച്ഛിക സാമ്പിൾ സാമ്പിളനശാസ്ത്രത്തിൽ അതിപ്രധാനമായ ഒന്നാണു്. മററു സാമ്പിളനരീതികളുടെ എല്ലാം അടിസ്ഥാനം ഇതാണെന്നു പോലും പറയാവുന്നതാണു്. എല്ലാ സംഭാവ്യതാസാമ്പിളനസമ്പ്രദായങ്ങളിലും ഒരു ഘട്ടത്തിലല്ലെങ്കിൽ മറെറാരു ഘട്ടത്തിൽ ലഘ്വയാദൃച്ഛിക സാമ്പിളനം ഉപയോഗിക്കപ്പെടുന്നുണ്ടു്. സാമ്പിളനശാസ്ത്രത്തോടു ബന്ധപ്പെട്ട ഗണിതസിദ്ധാന്തങ്ങൾ ആവിഷ്ക്കരിക്കുന്നതിലും ലഘ്വയാദൃച്ഛിക സാമ്പിളനത്തിനു മൗലികപ്രാധാന്യമുണ്ടു്. ചുരുക്കിപ്പറഞ്ഞാൽ ലഘ്വയാദൃച്ഛിക സാമ്പിളനം ആ

ദ്യമായി ചർച്ച ചെയ്യുന്നതിനു് പ്രേരകമായി മൂന്നു പ്രധാന കാരണങ്ങളാണു ള്ളതു്.

1. സാമ്പിളനശാസ്ത്രത്തിലെ പല അടിസ്ഥാനതത്വങ്ങളെയും ലഘുയാദൃച്ഛികസാമ്പിളനത്തോടു ബന്ധപ്പെടുത്തി അവതരിപ്പിക്കുകയും ചർച്ച ചെയ്യുകയും ചെയ്യാൻ എളുപ്പമാണു്.

2. ലഘു യാദൃച്ഛിക സാമ്പിളനത്തോടു ബന്ധപ്പെട്ട പ്രശ്നങ്ങളെപ്പററി ശരിയായ അവഗാഹം നേടിയാൽ മററു സാമ്പിളനസമ്പ്രദായങ്ങൾ മനസ്സിലാക്കാനും അവയുടെ സങ്കീർണത എത്രത്തോളമുണ്ടെന്നു ഗ്രഹിക്കാനും സാധിക്കും.

3. ഏററവും ലളിതമായ സംഭാവ്യതാസാമ്പിളനസമ്പ്രദായം ഇതാണു്.

പ്രായോഗികമായി പറഞ്ഞാൽ, വളരെ ചുരുക്കം സന്ദർഭങ്ങളിൽ മാത്രമേ ലഘുയാദൃച്ഛികസാമ്പിളനം ഉപയോഗിക്കാറുള്ളു. പരിഗണിക്കപ്പെടുന്ന അഭിലക്ഷണത്തെ സംബന്ധിച്ചിടത്തോളം സമഷ്ടിയിലെ അംഗങ്ങൾ വളരെയൊന്നും വിഭിന്നമല്ലാത്ത സാഹചര്യങ്ങളിൽ മാത്രമേ ഇമ്മാതിരി സാമ്പിൾ, സമഷ്ടിയുടെ ശരി പ്രതിരൂപമായിരിക്കു. അങ്ങനെയുള്ള സമഷ്ടികൾ താരതമ്യേന കുറവായതുകൊണ്ടാണു് മററു സാമ്പിളനരീതികൾ അംഗീകരിക്കാൻ നാം നിർബന്ധിതരാകുന്നതു്.

2. ലഘുയാദൃച്ഛിക സാമ്പിൾ എങ്ങനെ എടുക്കാം ?

സമഷ്ടിയുടെ പരിമാണം വളരെ ചെറുതാണെങ്കിൽ താഴെ കൊടുത്തിരിക്കുന്ന രീതിയിൽ സാമ്പിളെടുക്കാവുന്നതാണു്. വലിയ സമഷ്ടികളിൽ നിന്നു് സാമ്പിൾ എടുക്കുമ്പോഴും തത്വത്തിൽ ഈ രീതി തന്നെയാണു് നാം ഉപയോഗിക്കുന്നതു്. പ്രായോഗികസൗകര്യത്തിനു വേണ്ടി മററു ചില ക്രമീകരണങ്ങൾ ചെയ്തിരിക്കുന്നു എന്നു മാത്രം.

സമഷ്ടിയിൽ N അംഗങ്ങളാണു് ഉള്ളതെന്നിരിക്കട്ടെ. ഒരേ തരത്തിലും വലുപ്പത്തിലുമുള്ള N കടലാസുതുണ്ടുകൾ എടുക്കുക. അവയിൽ യഥാക്രമം 1 മുതൽ N വരെയുള്ള സംഖ്യകൾ എഴുതുക. സമഷ്ടിയിലെ അംഗങ്ങളുടെ ഒരു ലിസ്റ്റ് തയ്യാറാക്കി അവർക്കും 1 മുതൽ N വരെയുള്ള ക്രമസംഖ്യകൾ നല്കുക. കടലാസുതുണ്ടുകൾ ഒരുപോലെ ചുരുട്ടി ഒരു പാത്രത്തിലിട്ടു് നല്ലപോലെ ഇളക്കി അതിൽ നിന്നു് ഓരോന്നായി n കടലാസുതുണ്ടുകൾ എടുക്കുക. അങ്ങനെ എടുത്ത കടലാസുതുണ്ടുകളിലെ സംഖ്യകൾ ക്രമസംഖ്യകളായുള്ള സമഷ്ടിയിലെ അംഗങ്ങളെ സാമ്പിളായി സ്വീകരിക്കുക. ഓരോ കടലാസുതുണ്ടും എടുത്തു് അതിലെ സംഖ്യ എഴുതിയെടുത്തു് അതു് തിരികെ പാത്രത്തിലിട്ടതിനു ശേഷമാണു് അടുത്ത കടലാസ് എടുക്കുന്നതെങ്കിൽ നമുക്കു ലഭിക്കുന്ന സാമ്പിൾ പ്രതിസ്ഥാപനത്തോടു കൂടിയ ഒരു ലഘുയാദൃച്ഛികസാമ്പിളായിരിക്കും. നേരെ മറിച്ചു്, എടുത്ത കടലാസ് തിരിച്ചിടാതെയാണു് അടുത്തതു് എടുക്കുന്നതെങ്കിൽ പ്രതിസ്ഥാപനമില്ലാത്ത ലഘുയാദൃച്ഛികസാമ്പിളായിരിക്കും ലഭിക്കുക. ഓരോ കടലാസു തുണ്ടും എടുക്കുന്നതിനു മുൻപു് പാത്രത്തിലുള്ള കടലാസുതുണ്ടുകൾ നല്ലവണ്ണം ഇളക്കാൻ

പ്രത്യേകം ശ്രദ്ധിക്കണം. ഇപ്രകാരം സാമ്പിൾ എടുക്കുന്ന രീതിയെ 'ലോട്ടറി' സമ്പ്രദായം എന്നു പറയുന്നു.

സമഷ്ടിയിലെ അംഗസംഖ്യ വർധിക്കുംതോറും മേൽപറഞ്ഞ രീതി കൂടുതൽ കൂടുതൽ ബുദ്ധിമുട്ടുള്ളതായിത്തീരും. കടലാസുതുണ്ടുകൾ തയ്യാറാക്കുവാനുള്ള അധ്വാനഭാരം മാത്രമല്ല, പാത്രത്തിലിട്ട തുണ്ടുകൾ ശരിയായി ഇളക്കുക ഏതാണ്ട് അസാധ്യമായി തീരുമെന്നുള്ള അപകടം കൂടി നാം അഭിമുഖീകരിക്കേണ്ടിവരും. ശരിയായി ഇളക്കാതിരുന്നാൽ ചില അംഗങ്ങൾക്ക് സാമ്പിളിൽ ഉൾപ്പെടുവാനുള്ള സംഭാവ്യത മറ്റംഗങ്ങളെ അപേക്ഷിച്ച് കൂടുതലോ കുറവോ ആയിത്തീരുകയും അങ്ങനെ നമുക്കു ലഭിക്കുന്ന സാമ്പിൾ ഒരു ലഘു യാദൃച്ഛിക സാമ്പിൾ അല്ലാതാവുകയും ചെയ്യും. ഈ തകരാറുകൾ ഒഴിവാക്കുവാനുള്ള ഒരു പ്രായോഗികമാർഗം യാദൃച്ഛിക സാമ്പിളനസംഖ്യകളുടെ ഉപയോഗമാണ്. തത്വത്തിൽ മുൻപറഞ്ഞ രീതിയോട് തുല്യമായ യാന്ത്രിക സമ്പ്രദായങ്ങൾ അവലംബിച്ച് ക്രമീകരിച്ചിട്ടുള്ള സംഖ്യകളാണ് യാദൃച്ഛികസംഖ്യാപട്ടികകളിൽ കൊടുത്തിരിക്കുന്നത്. ഉദാഹരണമായി, നാലക്കങ്ങളുള്ള യാദൃച്ഛികസംഖ്യാപട്ടിക തയ്യാറാക്കിയിരിക്കുന്നത് തത്വത്തിൽ താഴെ പറയുന്നതു പോലെയാണ്. 1 മുതൽ 9999 വരെയുള്ള സംഖ്യകൾ ഒരേ തരത്തിലുള്ള 9999 കടലാസുതുണ്ടുകളിൽ എഴുതി ഒരുപോലെ ചുരുട്ടി ഒരു പാത്രത്തിലിടുന്നു. അതിൽ നിന്ന് കടലാസുതുണ്ടുകൾ നല്ലപോലെ ഇളക്കിയതിനു ശേഷം ഒരു ചുരുൾ എടുത്ത് അതിലെ സംഖ്യ യാദൃച്ഛികസംഖ്യാപട്ടികയിലെ ആദ്യസംഖ്യയായി രേഖപ്പെടുത്തുന്നു. ആ കടലാസുതുണ്ട് പാത്രത്തിൽ തിരികെ ഇട്ട് വീണ്ടും ഇളക്കിയിട്ട് രണ്ടാമതൊരു തുണ്ട് എടുക്കുന്നു. അതിലെ സംഖ്യ പട്ടികയിലെ രണ്ടാമത്തെ സംഖ്യയായി രേഖപ്പെടുത്തുന്നു. ഈ പ്രക്രിയ തുടർന്ന് ആവശ്യമുള്ളിടത്തോളം സംഖ്യകളുള്ള പട്ടിക തയ്യാറാക്കുന്നു.

യഥാർഥത്തിൽ യാദൃച്ഛികസംഖ്യാപട്ടികകൾ തയ്യാറാക്കുവാൻ ഉപയോഗിക്കുന്ന മാർഗങ്ങൾ കൂടുതൽ സങ്കീർണങ്ങളാണ്. എങ്കിലും ആ സമ്പ്രദായങ്ങളെല്ലാം മുൻപറഞ്ഞ സമ്പ്രദായത്തിന്റെ പരിഷ്കരിച്ച പതിപ്പുകൾ മാത്രമത്രെ.

യാദൃച്ഛിക സംഖ്യകൾ ഉപയോഗിച്ച് സാമ്പിളെടുക്കാൻ സാമ്പിളനഫ്രേമിന്റെ പ്രത്യേകതയനുസരിച്ച് പല നടപടിക്രമങ്ങളും സ്വീകരിക്കാറുണ്ട്. അതിൽ ഏററവും ലളിതമായ ഒന്നാണ് താഴെ കൊടുക്കുന്നത്. സമഷ്ടിയിൽ **N** അംഗങ്ങളാണ് ഉള്ളതെന്നിരിക്കട്ടെ. അതിൽ നിന്ന് **n** അംഗങ്ങളുള്ള, പ്രതിസ്ഥാപനമില്ലാത്ത ഒരു സാമ്പിളാണ് എടുക്കേണ്ടത് എന്നും സങ്കൽപ്പിക്കുക. ആദ്യമായി ചെയ്യേണ്ടത് സമഷ്ടിയിലെ അംഗങ്ങൾക്ക് 1 മുതൽ **N** വരെയുള്ള ക്രമസംഖ്യകൾ നൽകുക എന്നതാണ്. സമഷ്ടിയിലെ അംഗങ്ങളുടെ പട്ടികയാണ് ഫ്രേമെങ്കിൽ ഇതു് എളുപ്പത്തിൽ സാധിക്കും. അല്ലാത്ത സന്ദർഭങ്ങളിൽ സൗകര്യമുള്ള ഒരു രീതി അവലംബിച്ച് അവയെ ക്രമസംഖ്യകളുമായി ബന്ധിപ്പിക്കുക. ഒരു ഓഫീസിലുള്ള **N** ഫയലുകളാണ് സമഷ്ടി എന്നിരിക്കട്ടെ. അവ ഏതാനും അലമാരകളിലായി അടുക്കി വെച്ചിരിക്കുകയാണെന്നും സങ്കൽപ്പിക്കുക.

അലമാരകൾക്ക് അവ വെച്ചിരിക്കുന്ന സ്ഥാനമനുസരിച്ച് ഒരു ക്രമം നിശ്ചയിക്കുകയും അവയിൽ അടുക്കിയിരിക്കുന്ന ഫയലുകൾ ഒരു പ്രത്യേക ക്രമത്തിൽ പരിഗണിക്കാൻ തീരുമാനിക്കുകയും ചെയ്താൽ ഫയലുകളെ ഒന്നാമത്തേത്, രണ്ടാമത്തേത്...... N-ാമത്തേത് എന്ന് ക്രമസംഖ്യകൾ കൊണ്ട് സൂചിപ്പിക്കാൻ സാധിക്കും. മറ്റു സന്ദർഭങ്ങളിലും ഇതു പോലെയുള്ള ഏതെങ്കിലും ഒരു കൗശലം ഉപയോഗിച്ച് സമഷ്ടിയിലെ അംഗങ്ങൾക്ക് ക്രമസംഖ്യ നൽകുക. അടുത്തതു്, യാദൃച്ഛികസംഖ്യാപട്ടിക ഉപയോഗിക്കലാണ് (അനുബന്ധം 2 നോക്കുക). സൗകര്യത്തിനു വേണ്ടി യാദൃച്ഛികസംഖ്യാ പട്ടികയിലെ ഒരു പേജിൽ നിന്ന് കുറെ സംഖ്യകൾ താഴെ കൊടുക്കുന്നു.

7819	0447	5418
1299	1283	2649
6739	4201	6785
9380	2711	1440
1781	9757	9001
2103	6285	7758
4705	9488	9452
4193	4797	1510
7890	1333	2990
3992	1577	7617
		3778

യാദൃച്ഛികസംഖ്യാപട്ടികയിലെ ഏതെങ്കിലും ഒരു പേജ് എടുക്കുക. പേജ് തിരഞ്ഞെടുക്കുന്നതിൽ പ്രത്യേകമായി ഒന്നും ശ്രദ്ധിക്കേണ്ടതില്ല. ഏതു പേജ് എടുത്താലും കുഴപ്പമൊന്നുമില്ല. N=700 എന്നിരിക്കട്ടെ. എഴുനൂറോ അതിൽ കുറവോ ആയ യാദൃച്ഛികസംഖ്യകൾ പരിഗണിച്ചാൽ മതി. തിരഞ്ഞെടുത്ത പേജിലെ ഏതെങ്കിലും ഒരു സംഖ്യയിൽ ആരംഭിക്കുക. ഏതെങ്കിലും ഒരു കോളവും അതിലെ ഏതെങ്കിലും ഒരു വരിയും എടുക്കാവുന്നതാണ്. ഉദാഹരണമായി മുൻ കൊടുത്തിരിക്കുന്ന സംഖ്യകൾ രണ്ടാമത്തെ കോളത്തിൽ മൂന്നാമത്തെ വരിയിലുള്ള സംഖ്യയിലാണ് ആരംഭിക്കുന്നതെന്നിരിക്കട്ടെ. ആ സംഖ്യ കൊണ്ടു തുടങ്ങുന്ന മൂന്നക്കമുള്ള സംഖ്യ 7398. അതു മുതൽ താഴോട്ട് മൂന്നക്കമുള്ള സംഖ്യകൾ ഓരോന്നായി വായിക്കുക.

739
380
781
103
705
194
890

.... ഇവയിൽ 700 ൽ കൂടുതലുള്ള സംഖ്യകളെ അവഗണിക്കുക. ഒരിക്കൽ കിട്ടിയ സംഖ്യ ആവർത്തിച്ചു

വരികയാണെങ്കിൽ അതും അവഗണിക്കുക. അങ്ങനെ

380
103
149

എന്നിങ്ങനെ സംഖ്യകൾ കിട്ടും. ആ പേജിന്റെ അടിയിൽ എത്തിയാൽ ആദ്യത്തെ വരിയിൽ അഞ്ചാമത്തെ കോളം മുതൽ ആരംഭിച്ച് മുൻപറഞ്ഞതു പോലെ സംഖ്യകൾ എടുക്കുക. n സംഖ്യകൾ കിട്ടുന്നതു വരെ ഈ പ്രക്രിയ തുടരണം. അങ്ങനെ ലഭിക്കുന്ന n യാദൃച്ഛിക സംഖ്യകൾ ക്രമസംഖ്യകളായുള്ള സമഷ്ടിയിലെ അംഗങ്ങളെയാണ് സാമ്പിളിൽ ഉൾപ്പെടുത്തുന്നത്. ഇവിടെ സംഖ്യകൾ താഴോട്ടു താഴോട്ടാണ് എടുത്തത്. പക്ഷേ വലത്തോട്ടു വലത്തോട്ടു വേണമെങ്കിലും സംഖ്യകൾ എടുക്കാവുന്നതാണ്.

ഒരിക്കൽ കിട്ടിയ യാദൃച്ഛികസംഖ്യ വീണ്ടും വന്നാൽ അതു് അവഗണിക്കുന്നില്ലെങ്കിൽ നമുക്കു കിട്ടുന്നതു് പ്രതിസ്ഥാപനത്തോടു കൂടിയ ലഘു യാദൃച്ഛിക സാമ്പിളായിരിക്കും.

മുൻ വിവരിച്ച രണ്ടു സമ്പ്രദായങ്ങളിലും സമഷ്ടിയിലെ ഏതൊരംഗത്തിനും സാമ്പിളിൽ ഉൾപ്പെടുവാനുള്ള സംഭാവ്യത തുല്യമായിരിക്കത്തക്ക വണ്ണമാണ് സാമ്പിളനരീതി ആവിഷ്ക്കരിച്ചിരിക്കുന്നത്. പക്ഷേ ലഘുയാദൃച്ഛിക സാമ്പിൾ നിർവചിച്ചപ്പോൾ സമഷ്ടിയിൽ നിന്ന് എടുക്കാവുന്ന നിർദിഷ്ട വലുപ്പമുള്ള എല്ലാ സാമ്പിളുകളുടെയും സംഭാവ്യത തുല്യമായിരിക്കത്തവണ്ണം എടുക്കുന്ന സാമ്പിളാണ് അതെന്നു് പറയുകയുണ്ടായല്ലൊ. മുൻ വിവരിച്ച രീതിയിൽ എടുക്കുന്ന സാമ്പിൾ ലഘുയാദൃച്ഛികസാമ്പിളാണോ എന്നു പരിശോധിക്കുന്നതു് ഈ സന്ദർഭത്തിൽ സംഗതമാണ്.

പ്രതിസ്ഥാപനമില്ലാത്ത സാമ്പിളനരീതി പരിശോധിച്ചു നോക്കാം. N അംഗങ്ങളുള്ള ഒരു സമഷ്ടിയിൽ നിന്നു് n വലുപ്പമുള്ളതും പ്രതിസ്ഥാപനമില്ലാത്തതുമായ $\binom{N}{n}$ സാമ്പിളുകളാണല്ലൊ എടുക്കാവുന്നതു്. അവ ഓരോന്നിനും തിരഞ്ഞെടുക്കപ്പെടുവാനുള്ള സംഭാവ്യത തുല്യമായിരിക്കണം. അതായതു് ഏതൊരു സാമ്പിളിന്റെയും തിരഞ്ഞെടുക്കപ്പെടുവാനുള്ള സംഭാവ്യത $\frac{1}{\binom{N}{n}}$ ആയിരിക്കണം. അങ്ങനെ തിരഞ്ഞെടുക്കപ്പെടുന്ന സാമ്പിൾ ലഘുയാദൃച്ഛിക സാമ്പിളാണ്. അടുത്തതായി മുൻ വിവരിച്ച രീതിയിലാണ് സാമ്പിളെടുക്കുന്നതെങ്കിൽ ഏതെങ്കിലും ഒരു സാമ്പിൾ തിരഞ്ഞെടുക്കപ്പെടുവാനുള്ള സംഭാവ്യത എന്താണെന്നു പരിശോധിക്കാം. ഓരോ ഘട്ടത്തിലും പാത്രത്തിൽ അവശേഷിച്ചിട്ടുള്ള ഓരോ കടലാസുതുണ്ടും തിരഞ്ഞെടുക്കപ്പെടുവാനുള്ള ആദ്യഘട്ടത്തിൽ തുല്യസംഭാവ്യത $\frac{1}{N}$ ഉം രണ്ടാം ഘട്ടത്തിൽ, $\frac{1}{N-1}$ ഉം എന്ന ക്രമത്തിലായിരിക്കും. അപ്പോൾ ആദ്യമെടു

ക്കുന്ന കടലാസ്സുതുണ്ടിൽ എഴുതിയിരിക്കുന്നതു് നിർദിഷ്ടമായ ഒരു സാമ്പിളിലെ n അംഗങ്ങളിൽ ഏതെങ്കിലും ഒന്നു് ആകുവാനുള്ള സംഭാവ്യത $\frac{n}{N}$ ആണു് (സംഭാവ്യതാസങ്കലനനിയമപ്രകാരം) പാത്രത്തിൽ $N-1$ കടലാസ്സുതുണ്ടുകളാണു് ബാക്കിയുള്ളതു്. രണ്ടാമതു് എടുക്കുന്ന തുണ്ടു് അതിൽ ഏതെങ്കിലും ഒന്നായിരിക്കും. അതു് നിർദിഷ്ട സാമ്പിളിലെ ബാക്കിയുള്ള $n-1$ അംഗങ്ങളിൽ ഏതെങ്കിലും ഒന്നിനെ സൂചിപ്പിക്കുന്നതാകുവാനുള്ള സംഭാവ്യത $\frac{n-1}{N-1}$ അത്രെ. ഇങ്ങനെ n-ാമതു് എടുക്കുന്ന തുണ്ടു് ആ സാമ്പിളിൽ ബാക്കി ശേഷിച്ചിരിക്കുന്ന ഒരംഗത്തെ സൂചിപ്പിക്കുവാനുള്ള സംഭാവ്യത $\frac{1}{N-n+1}$ ആണു്. അതുകൊണ്ടു് നിർദിഷ്ടമായ സാമ്പിൾ ലഭിക്കുവാനുള്ള സംഭാവ്യത

$$= \frac{n}{N} \times \frac{n-1}{N-1} \times \cdots \times \frac{1}{N-n+1}$$

(സംഭാവ്യതാഗുണനനിയമപ്രകാരം)

$$= \frac{n!\ (N-n)!}{N!} = \frac{1}{(N)}$$

n അംഗങ്ങളുള്ള ഏതു സാമ്പിൾ ലഭിക്കുവാനുമുള്ള സംഭാവ്യത ഇതായതുകൊണ്ടു് നമുക്കു ലഭിക്കുന്ന സാമ്പിൾ ലഘുയാദൃച്ഛിക സാമ്പിളായിരിക്കുമെന്നു സിദ്ധിക്കുന്നു.

ലഘുയാദൃച്ഛിക സാമ്പിളനത്തിനു സുപ്രധാനമായ ഒരു പ്രത്യേകതയുണ്ടു്. സാമ്പിളനപ്രക്രിയയിലെ ഏതെങ്കിലുമൊരു ഘട്ടത്തിൽ സമഷ്ടിയിലെ നിർദിഷ്ടമായ ഏതെങ്കിലും ഒരംഗം സാമ്പിളിലേക്കു തിരഞ്ഞെടുക്കപ്പെടുവാനുള്ള സംഭാവ്യത, അതു് ആദ്യഘട്ടത്തിൽ തിരഞ്ഞെടുക്കപ്പെടുവാനുള്ള സംഭാവ്യതയായ $\frac{1}{N}$ തന്നെയാണു്. നാം നിർദേശിക്കുന്ന അംഗം A ആണെന്നിരിക്കട്ടെ. ആദ്യഘട്ടത്തിൽ സമഷ്ടിയിലെ N അംഗങ്ങളിൽ ഏതെങ്കിലും ഒന്നായിരിക്കുമല്ലൊ സാമ്പിളിലേക്കു തിരഞ്ഞെടുക്കപ്പെടുക. അതു് A ആയിരിക്കുവാനുള്ള സംഭാവ്യത സ്വാഭാവികമായും $\frac{1}{N}$ ആയിരിക്കും. 'r', n-ൽ കുറഞ്ഞ ഏതെങ്കിലും ഒരു സംഖ്യയാണെന്നിരിക്കട്ടെ. A എന്ന അംഗം r-ാമതായി സാമ്പിളിലേക്കു തിരഞ്ഞെടുക്കപ്പെടുവാനുള്ള സംഭാവ്യത എന്താണെന്നു പരിശോധിക്കാം. ആദ്യം തിരഞ്ഞെടുക്കപ്പെടുന്നതു് A ഒഴികെ മറേറതെങ്കിലും അംഗമായിരിക്കണം. അതിനുള്ള സംഭാവ്യത $\frac{N-1}{N}$. രണ്ടാമതു് തിരഞ്ഞെടുക്കപ്പെടുന്നതു്, A യും ആദ്യം തിരഞ്ഞെടുക്കപ്പെട്ട അംഗവും ഒഴികെയുള്ള ഏതെങ്കിലും ഒരു അംഗമായിരിക്ക

ണം. അതിനുള്ള സംഭാവ്യത $\frac{N-1}{N-1}$ അങ്ങനെ ക്രമേണ (r—1) –ാമതു് തിരഞ്ഞെടുക്കപ്പെടുന്നതു് A യും അതു വരെ തിരഞ്ഞെടുക്കപ്പെട്ട (r—2) അംഗങ്ങളും ഒഴികെയുള്ള ഒരംഗമായിരിക്കണം. അതിന്റെ സംഭാവ്യത $\frac{N-r+1}{N-r+2}$. r–ാമതു് തിരഞ്ഞെടുക്കപ്പെടുന്നതു് A ആയിരിക്കണം. അതിൻെറ സംഭാവ്യത

$$= \frac{1}{N-r+1}$$

∴ r–ാമതു് തിരഞ്ഞെടുക്കപ്പെടുന്നതു് A ആയിരിക്കുവാനുള്ള സംഭാവ്യത

$$= \frac{N-1}{N} \times \frac{N-2}{N-1} \times \cdots \times \frac{N-r+1}{N-r+2} \times \frac{1}{N-r+1} = \frac{1}{N}$$

ആദ്യം തന്നെ A തിരഞ്ഞെടുക്കപ്പെടുവാനുള്ള സംഭാവ്യതയും ഇതു തന്നെയാണല്ലൊ. സമഷ്ടിയിലെ ഏതംഗത്തെ സംബന്ധിച്ചും ഇതു ശരിയാണു്. മറ്റൊരു തരത്തിൽ പറഞ്ഞാൽ, ലഘു യാദൃച്ഛിക സാമ്പിളനത്തിൽ സമഷ്ടിയിലെ ഏതൊരംഗത്തിനും സാമ്പിളിൽ ഉൾപ്പെടുവാനുള്ള സംഭാവ്യത തുല്യമാണു്. ആദ്യത്തെതു മുതൽ n–ാമത്തെതുവരെയുള്ള ഏതു ഘട്ടത്തിൽ വേണമെങ്കിലും ഒരംഗത്തിനു് സാമ്പിളിലേക്കു് തിരഞ്ഞെടുക്കപ്പെടാമെന്നതു കൊണ്ടു്, ഏതൊരംഗത്തിനും സാമ്പിളിൽ ഉൾപ്പെടുവാനുള്ള സംഭാവ്യത $\frac{n}{N}$ ആണു്. ഏതൊരംഗത്തിനും സാമ്പിളിൽ ഒൾപ്പെടുവാനുള്ള സംഭാവ്യത തുല്യമായിരിക്കുക എന്നതു് ലഘു യാദൃച്ഛിക സാമ്പിളനത്തിൻെറ നിർവചനമായി ചിലർ സ്വീകരിച്ചു കാണുന്നുണ്ടു്. പക്ഷേ, ആ നിർവചനം മററുപല സാമ്പിളനരീതികളിലും അപൂർണമാണെന്നു പറയാതെ തരമില്ല. ഉദാഹരണത്തിനു് ക്രമാനുസൃതസാമ്പിളനത്തിൽ സമഷ്ടിയിലെ ഏതൊരംഗത്തിനും സാമ്പിളിൽ ഉൾപ്പെടുവാനുള്ള സംഭാവ്യത തുല്യമാണു്. അതുകൊണ്ടു് ലഘു യാദൃച്ഛികസാമ്പിളനത്തെ നേരത്തെ പറഞ്ഞതു് പോലെ തന്നെ നിർവചിക്കേണ്ടതാണു്.

യാദൃച്ഛിക സംഖ്യകളുപയോഗിച്ചു് സാമ്പിളെടുക്കുമ്പോൾ ചില കാര്യങ്ങൾ ഓർമിക്കണം. സമഷ്ടിയുടെ അംഗസംഖ്യയായ N- നോടു് തുല്യമോ, അതിൽ കുറവോ ആയ യാദൃച്ഛിക സംഖ്യകളാണു് നാം തിരഞ്ഞെടുക്കുന്നതു്. യാദൃച്ഛിക സംഖ്യാപട്ടികയിൽ നിന്നു് തുടർച്ചയായി സംഖ്യകളെടുക്കുമ്പോൾ N–നെ ക്കാൾ വലിയ സംഖ്യകൾ വന്നാൽ അവയെ അവഗണിക്കുകയാണല്ലൊ ചെയ്യുന്നതു്. സമഷ്ടിയുടെ അംഗസംഖ്യ 100-ൽ കുറവാണെങ്കിൽ രണ്ടക്കമുള്ള യാദൃച്ഛിക സംഖ്യകൾ, 100–ൽ കൂടുതലും 1000 ത്തിൽ കുറവുമാണെങ്കിൽ മൂന്നക്കമുള്ള യാദൃച്ഛികസംഖ്യകൾ, 1000ത്തിനും 10,000 ത്തിനു മധ്യേയാണു് അംഗസംഖ്യയെങ്കിൽ നാലു് അക്കമുള്ള സംഖ്യകൾ എന്നിങ്ങനെ Nലെ അക്കങ്ങളുടെ എണ്ണമനുസരിച്ചു് എത്ര അക്കമുള്ള യാദൃച്ഛികസംഖ്യകളാണു് പരിഗണിക്കേണ്ടതു് എന്നു

നാം തീരുമാനിക്കുന്നു. സമഷ്ടിയിലെ അംഗസംഖ്യ 300 ആണെന്നിരിക്കട്ടെ. മൂന്നക്കമുള്ള യാദൃച്ഛിക സംഖ്യകളാണ് പരിഗണിക്കുന്നത്. യാദൃച്ഛിക സംഖ്യാ പട്ടികയിൽ നിന്നു് മൂന്നക്കമുള്ള സംഖ്യകൾ തുടർച്ചയായി പരിഗണിച്ച് അവയിൽ 300-ൽ കൂടുതലുള്ള സംഖ്യകളെയെല്ലാം അവഗണിക്കുകയാണല്ലോ ചെയ്യുന്നതു്. ഇതിൻെറ ഫലമായി പട്ടികയിലുള്ള സംഖ്യകളിൽ ഏതാണ്ടു് മൂന്നിൽ രണ്ടു ഭാഗം സംഖ്യകളെയും അവഗണിക്കേണ്ടതായി വരും. ഇതു് ചില പ്രായോഗിക അസൗകര്യങ്ങൾക്കു് കാരണമായിത്തീരാം. സമയനഷ്ടമാണു് ഏററവും പ്രധാന അസൗകര്യം. ഈ ബുദ്ധിമുട്ടു് ഒഴിവാക്കാനായി മൂന്നക്കമുള്ള യാദൃച്ഛരികസംഖ്യകൾ എടുത്തു് അവയിൽ 300-ൽ കൂടുതലുള്ള സംഖ്യകളുണ്ടെങ്കിൽ അവയെ 300 കൊണ്ടു് ഹരിച്ചു് കിട്ടുന്ന ശിഷ്ടം അതിനു പകരമായി എടുക്കുകയും, അങ്ങനെ കിട്ടുന്ന സംഖ്യകൾ ക്രമസംഖ്യകളായുള്ള അംഗങ്ങളെ സാമ്പിളിൽ ഉൾപ്പെടുത്തുകയും ചെയ്യാറുണ്ടു്. ഈ സമ്പ്രദായം അവലംബിച്ചാൽ ആവർത്തിച്ചു വരുന്ന യാദൃച്ഛരികസംഖ്യകളൊഴികെ മറെറാന്നും അവഗണിക്കേണ്ടിവരികയില്ല. ഉദാഹരണമായി 468 ആണു് കിട്ടിയ യാദൃച്ഛരികസംഖ്യയെന്നിരിക്കട്ടെ. അതിനെ 300 കൊണ്ടു ഹരിച്ചാൽ ശിഷ്ടം 168. 168 ക്രമസംഖ്യയായുള്ള സമഷ്ടിയിലെ അംഗത്തെയാണു് 468 സൂചിപ്പിക്കുന്നതെന്നു് നാം അംഗീകരിക്കുന്നു. ഈ രീതിക്കു് ഗൗരവമായ ഒരു തകരാറുണ്ടു്. 900 ത്തിൽ ശേഷം 99 സംഖ്യകൾ മാത്രമാണല്ലോ ഉള്ളതു്. ഇതിൻെറ ഫലമായി 1 മുതൽ 99 വരെയുള്ള സംഖ്യകൾ ക്രമസംഖ്യകളായുള്ള സമഷ്ടിയിലെ അംഗങ്ങൾ സാമ്പിളിലേക്കു തിരഞ്ഞെടുക്കപ്പെടുവാനുള്ള സംഭാവ്യത മററംഗങ്ങളുടെതിനെക്കാൾ കൂടുതലായിരിക്കും. ഉദാഹരണമായി, 85 ക്രമസംഖ്യയായുള്ള ഒരംഗത്തെ എടുക്കുക. നമുക്കു കിട്ടുന്ന യാദൃച്ഛരിക സംഖ്യ 85, 385, 685, 985 എന്നിവയിൽ ഏതെങ്കിലും ഒന്നാണെങ്കിൽ ആ അംഗം തിരഞ്ഞെടുക്കപ്പെടും. പക്ഷേ 123 ക്രമസംഖ്യയായുള്ള ഒരംഗം തിരഞ്ഞെടുക്കപ്പെടണമെങ്കിൽ 123, 423, 723 എന്നീ സംഖ്യകളിൽ ഏതെങ്കിലും ഒന്നു കിട്ടണം. ഇതിൽ നിന്നു് എല്ലാ അംഗങ്ങളുടെയും സാമ്പിളിൽ ഉൾപ്പെടുവാനുള്ള സംഭാവ്യത തുല്യമല്ലെന്നു സിദ്ധിക്കുന്നു. അതു കൊണ്ടു തന്നെ അങ്ങനെ ലഭിക്കുന്ന സാമ്പിൾ ലഘുയാദൃച്ഛരികസാമ്പിൾ അല്ലാതായിത്തീരുകയും ചെയ്യുന്നു. ഈ തകരാറു് ഒഴിവാക്കാനുള്ള മാർഗം 900 ത്തിൽ കൂടിയ യാദൃച്ഛരിക സംഖ്യകൾ അവഗണിക്കുക എന്നതാണു്. മറെറാരു തരത്തിൽ പറഞ്ഞാൽ N നു് എത്ര അക്കങ്ങളുണ്ടോ അത്രയും അക്കങ്ങളുള്ള ഏററവും വലിയ സംഖ്യ പരിഗണിക്കുക. അതിനെക്കാൾ കുറവായിട്ടുള്ള N ൻെറ ഏററവും വലിയ ഗുണിതം ഏതെന്നു നോക്കുക. അതു വരെയുള്ള യാദൃച്ഛരികസംഖ്യകൾ സ്വീകരിക്കുക. ബാക്കി സംഖ്യകൾ അവഗണിക്കുക. ഈ ഭേദഗതിയോടു കൂടി മുൻ വിവരിച്ച രീതിയിൽ സാമ്പിളെടുത്താൽ അതു് ലഘു യാദൃച്ഛരിക സാമ്പിളായിരിക്കും.

ഫ്രേമിൻെറ സ്വഭാവവ്യത്യാസമനുസരിച്ചു് യാദൃച്ഛരിക സംഖ്യകളുപയോഗിക്കുന്നതിലും ചില ഭേദഗതികൾ വേണ്ടി വരുമെന്നു് സൂചിപ്പിച്ചല്ലോ.

മുൻ വിവരിച്ച സാമ്പിളനരീതി സമഷ്ടിയിലെ അംഗങ്ങളുടെ ഒരു ലിസ്റ്റ് ഉണ്ടെങ്കിൽ മാത്രമേ ഫലിക്കുകയുള്ളു. പലപ്പോഴും ഫ്രേമിന്റെ സ്വഭാവം അതായിരിക്കുകയില്ല. അങ്ങനെയുള്ള സാഹചര്യങ്ങളിൽ സമഷ്ടിയിൽ നിന്നു് എടുക്കാവുന്ന നിർദിഷ്ടവലുപ്പമുള്ള എല്ലാ സാമ്പിളുകൾക്കും തുല്യസംഭാവ്യത ഉണ്ടായിരിക്കണമെന്ന അടിസ്ഥാനതത്വം ലംഘിക്കാതിരിക്കാൻ പ്രത്യേകം ശ്രദ്ധിക്കണം. പ്രധാനമായും സമഷ്ടിയിലെ എല്ലാ അംഗങ്ങൾക്കും സാമ്പിളിൽ ഉൾപ്പെടുവാനുള്ള സംഭാവ്യത തുല്യമാണോ എന്നു പരിശോധിച്ചാൽ മതി ഇതിനു്. ഒരു ഗ്രാമത്തിലെ തെങ്ങുകൃഷിയെപ്പററിയാണു് പഠനം നടത്തുന്നതു് എന്നിരിക്കട്ടെ. ഗ്രാമത്തിലെ കൃഷിഭൂമി പലതായി തിരിച്ച് ഓരോന്നിനും ഓരോ സർവെ നമ്പർ വീതം കൊടുത്തിരിക്കുന്നു. ഓരോ സർവെ നമ്പരിലും പെട്ട ഭൂമിയെ ഓരോ വ്യക്തിയായി കണക്കാക്കുന്നു എന്നിരിക്കട്ടെ. 1 മുതൽ 16 വരെയും 25 മുതൽ 43 വരെയും 86 മുതൽ 93 വരെയുമുള്ള സർവെ നമ്പറുകളിൽ തെങ്ങുകൃഷിയും മററു സർവെ നമ്പറുകളിൽ നെൽകൃഷിയും ചെയ്യുന്നെന്നിരിക്കട്ടെ. തെങ്ങ് കൃഷി ചെയ്യുന്ന സർവെ നമ്പറുകളിൽ നിന്നാണു നമുക്കു് സാമ്പിളെടുക്കേണ്ടതു്. രണ്ടക്കമുള്ള ഒരു യാദൃച്ഛികസംഖ്യ എടുത്തു് അതു സർവെ നമ്പരായുള്ള ഭൂമി തെങ്ങു കൃഷി ചെയ്യുന്നതാണെങ്കിൽ അതു് സാമ്പിളിൽ ഉൾപ്പെടുത്തുക. അല്ലാത്ത പക്ഷം ആ സംഖ്യ കഴിഞ്ഞുള്ള സർവെ നമ്പറുകളിൽ ഏററവും ചെറിയ ഏതു സർവെ നമ്പരാണോ തെങ്ങുകൃഷി സ്ഥലത്തിന്റേതായി വരുന്നതു് ആ കൃഷിസ്ഥലം സാമ്പിളിൽ ഉൾപ്പെടുത്തുക. ഇതാണു് സാമ്പിളെടുക്കുന്നതിനു് അംഗീകരിക്കുന്ന നിയമമെന്നിരിക്കട്ടെ. ഇങ്ങനെ എടുക്കുന്ന ഒരു സാമ്പിൾ യാദൃച്ഛികസാമ്പിളായിരിക്കുകയില്ല എന്നു കാണാൻ പ്രയാസമില്ല. 17 മുതൽ 24 വരെയുള്ള ഏതു യാദൃച്ഛികസംഖ്യ കിട്ടിയാലും 25-ാം സർവെ നമ്പറാണു് സാമ്പിളിൽ ഉൾപ്പെടുത്തുന്നതു്. നേരെ മറിച്ചു് 26-ാം സർവെ നമ്പർ സാമ്പിളിൽ ഉൾപ്പെടുത്തണമെങ്കിൽ 26 എന്ന യാദൃച്ഛികസംഖ്യ കിട്ടണം. ഇതിൽ നിന്നു് 25, 26 എന്നീ സർവെ നമ്പറുകൾക്കു് സാമ്പിളിൽ ഉൾപ്പെടുവാനുള്ള സംഭാവ്യത തുല്യമല്ലെന്നു് സിദ്ധിക്കുന്നല്ലോ. ഇങ്ങനെയുള്ള ഒരു സന്ദർഭത്തിൽ തെങ്ങുകൃഷിസ്ഥലങ്ങളുടെ സർവെ നമ്പറുകൾ മാത്രമെടുത്തു് അവയ്ക്കു് ക്രമസംഖ്യ കൊടുത്തു് മുൻ വിവരിച്ചതു പോലെ സാമ്പിളെടുക്കണം. ചുരുക്കത്തിൽ യാദൃച്ഛികസംഖ്യകൾ ഉപയോഗിച്ചു് സാമ്പിളെടുക്കുവാൻ പുതിയ രീതികൾ നിർദേശിക്കുന്നതു് വളരെ കരുതി ചെയ്യേണ്ട ഒന്നാണു്.

3. സമാന്തരമാധ്യത്തിന്റെ ആകലനം — ചില പ്രാഥമികതത്വങ്ങൾ

നിർദിഷ്ടമായ ഒരു സാന്തസമഷ്ടിയിൽ നിന്നു് നിശ്ചിത വലുപ്പമുള്ള ഒരു ലഘു യാദൃച്ഛിക സാമ്പിൾ എങ്ങനെയാണു് എടുക്കുക എന്നു നാം കണ്ടു കഴിഞ്ഞു. അടുത്തതു് സമഷ്ടിയെപ്പററി നാം അറിയാനാഗ്രഹിക്കുന്ന വിവരങ്ങൾ ഇമ്മാതിരി ഒരു സാമ്പിൾ ഉപയോഗിച്ചു് എങ്ങനെ കണ്ടുപിടിക്കാം എന്നു പരിശോധിക്കുകയാണു്. സമഷ്ടിയിലെ അംഗങ്ങളോടു ബന്ധപ്പെട്ട ഒന്നോ അതി

ലധികമോ അഭിലക്ഷണങ്ങളിൽ അന്വേഷകനു താല്പര്യമുണ്ടാവാം. ഈ അഭിലക്ഷണങ്ങൾ ചരരൂപത്തിലോ ഗുണരൂപത്തിലോ ആവാം. അടിസ്ഥാനതത്വങ്ങൾ മനസ്സിലാക്കുന്നതിനുള്ള സൗകര്യത്തിനു വേണ്ടി വളരെ ലളിതമായ ഒരു സാഹചര്യം ആദ്യമായി പരിഗണിക്കാം. ചരരൂപത്തിലുള്ള ഒരേ ഒരു അഭിലക്ഷണത്തിൽ മാത്രമേ നമുക്ക് താല്പര്യമുള്ളു എന്നു സങ്കല്പിക്കുക. സാമ്പിളിലെ അംഗങ്ങളിൽ നിന്ന് ഈ അഭിലക്ഷണത്തിന്റെ മൂല്യങ്ങൾ ശേഖരിക്കുന്നു. ഉദാഹരണമായി 1970 ആഗസ്റ്റ് 1--ാം തിയതി കേരളത്തിലെ കോളേജുകളിൽ പഠനം നടത്തുന്ന വിദ്യാർഥികളാണ് സമഷ്ടി എന്നിരിക്കട്ടെ. അവരുടെ തൂക്കമാണ് നമുക്ക് താല്പര്യമുള്ള അഭിലക്ഷണമെന്നും സങ്കല്പിക്കുക. ചരരൂപത്തിലുള്ള ഒരു അഭിലക്ഷണമാണിത്. മുൻ വിവരിച്ച ഏതെങ്കിലും രീതിയിൽ ഒരു ലഘു യാദൃച്ഛിക സാമ്പിൾ എടുക്കുന്നു. സാമ്പിളിലെ അംഗങ്ങളായ വിദ്യാർഥികളുടെ തൂക്കം കണ്ടുപിടിച്ചു രേഖപ്പെടുത്തുന്നു. ഈ അഭിലക്ഷണമൂല്യങ്ങൾ ഉപയോഗിച്ച് സമഷ്ടിയിലെ അംഗങ്ങളുടെ അഭിലക്ഷണമൂല്യങ്ങളെപ്പററി ചിലതെല്ലാം മനസ്സിലാക്കാനാണ് നാം ശ്രമിക്കുന്നത്.

സൗകര്യത്തിനു വേണ്ടി ഇനിമേലുള്ള ചർച്ചയിൽ സമഷ്ടി എന്നതു കൊണ്ട് സമഷ്ടിയിലെ അംഗങ്ങളോടു ബന്ധപ്പെട്ട അഭിലക്ഷണമൂല്യങ്ങളുടെ സമുച്ചയമെന്നും സാമ്പിൾ എന്നതു കൊണ്ട് സാമ്പിളിലെ അംഗങ്ങളിൽ നിന്ന് ശേഖരിച്ച അഭിലക്ഷണമൂല്യങ്ങളെന്നും അർഥമാക്കാം. സമഷ്ടിയുടെ സമാന്തരമാധ്യം ആകലനം ചെയ്യുകയാണ് സാമ്പിളനത്തിന്റെ ലക്ഷ്യമെന്നിരിക്കട്ടെ. അത് എങ്ങനെ നിർവഹിക്കാമെന്നതും അങ്ങനെ ലഭിക്കുന്ന ആകലനങ്ങൾക്ക് എത്ര സൂക്ഷ്മതയുണ്ടായിരിക്കുമെന്നതുമാണ് നാം പരിഗണിക്കേണ്ട പ്രശ്നങ്ങൾ.

സമഷ്ടിയുടെ സമാന്തരമാധ്യം അതിനോടു ബന്ധപ്പെട്ട ഒരു പ്രാചലമാണ്. സാമ്പിളിൾമൂല്യങ്ങളുടെ ഏതു ഫലനത്തെ വേണമെങ്കിലും അതിന്റെ ആകലമായി നിർദേശിക്കാവുന്നതാണ്. ഏതു ഫലനമാണ് സ്വീകരിക്കുക എന്നത് അതിന് അഭിലക്ഷണീയങ്ങളായ ഏതെല്ലാം വൈശിഷ്ട്യങ്ങളാണ് ഉള്ളത് എന്നതനുസരിച്ചാണ് തീരുമാനിക്കുന്നത്. ഒരു ആകലത്തിനുണ്ടായിരിക്കേണ്ട അഭിലക്ഷണീയവൈശിഷ്ട്യങ്ങൾ ഏതെല്ലാമെന്നത് സന്ദർഭവും സാഹചര്യവും അനുസരിച്ചേ തീരുമാനിക്കാൻ കഴിയൂ താനും. എന്നിരുന്നാലും പൊതുവെ സ്വീകാര്യങ്ങളായ രണ്ടു വൈശിഷ്ട്യങ്ങളുണ്ട്. അവ (1) അഭിനതി ഇല്ലായ്മ (2) അവിരോധിത്വം എന്നിവയാണ്. ഇവയുടെ ഗണിതശാസ്ത്രപരമായ നിർവചനം ഒന്നാമധ്യായത്തിൽ (1.9) കൊടുത്തിട്ടുണ്ട്. സമഷ്ടിയുടെ സമാന്തരമാധ്യത്തിന്റെ ആകലമായി സാമ്പിൾ സമാന്തരമാധ്യം നിർദേശിക്കപ്പെട്ടു എന്നിരിക്കട്ടെ. പ്രസ്തുത ആകലത്തെ അടിസ്ഥാനമാക്കി ഈ വൈശിഷ്ട്യങ്ങളുടെ സാമാന്യരൂപം പരിശോധിച്ചു നോക്കാം.

പൊതുവായ ഒരു ചർച്ചയിലേക്കു കടക്കുന്നതിനു മുൻപായി വളരെ ചെറിയ ഒരു സമഷ്ടിയും അതിൽ നിന്ന് എടുക്കാവുന്ന പല വലുപ്പമുള്ള സാമ്പിളുകളും അവയിൽ നിന്ന് ലഭിക്കുന്ന ആകലങ്ങളുടെ പ്രത്യേകതകളും നമുക്കു പരിശോ

ധിക്കാം. ഈ പഠനം പൊതുസിദ്ധാന്തങ്ങൾ ആവിഷ്ക്കരിക്കുന്നതിനു ആവശ്യമായ പശ്ചാത്തലം സൃഷ്ടിക്കുകയും അതിലെ സൈദ്ധാന്തികസങ്കീർണതകളിലേക്ക് ഒരുൾക്കാഴ്ച നല്കുകയും ചെയ്യും.

നമ്മുടെ സമഷ്ടി A, B, C, D, E, F എന്ന ആറു തൊഴിലാളികൾ ചേർന്നതാണെന്നു സങ്കൽപ്പിക്കുക. അവരുടെ ദിവസക്കൂലിയാണ് പരിഗണിക്കപ്പെടുന്ന അഭിലക്ഷണമെന്നും കൂലികളുടെ സമാന്തരമാധ്യം ആകലനം ചെയ്യുകയാണ് സാമ്പിളനലക്ഷ്യമെന്നും സങ്കൽപ്പിക്കുക. ഈ ആറു തൊഴിലാളികളുടെയും ദിവസക്കൂലി താഴെ കൊടുക്കുന്നു.

പട്ടിക 3. 2

തൊഴിലാളിയുടെ പേർ	കൂലി (രൂപയിൽ)
A	3
B	1
C	4
D	2
E	6
F	5

സമഷ്ടിയുടെ സമാന്തരമാധ്യം $\dfrac{3+1+4+2+6+5}{6} = 3.5$ രൂപയാണ്. ഇതു് ലഘു യാദൃച്ഛിക സാമ്പിളുകൾ എടുത്തു് ആകലനം ചെയ്യാൻ ശ്രമിച്ചാൽ എന്തു സംഭവിക്കുമെന്നു പരിശോധിച്ചു നോക്കാം.

സാമ്പിൾസമാന്തരമാധ്യമെന്ന ആകലത്തിന്റെ പ്രതീക്ഷ സമഷ്ടിയുടെ സമാന്തരമാധ്യമായിരിക്കുമ്പോഴാണല്ലോ അതു് ഒരു അനഭിനത ആകലമെന്നു് വ്യവഹരിക്കപ്പെടുന്നതു്. ഗണിതശാസ്ത്രപരമായി പറഞ്ഞാൽ, ഈ ആകലത്തിനു് ഉണ്ടാകാവുന്ന മൂല്യങ്ങളെ അതതു് മൂല്യം കിട്ടാനുള്ള സംഭാവ്യത കൊണ്ടു ഗുണിച്ചു് ഒന്നിച്ചു കൂട്ടുമ്പോഴാണ് അതിന്റെ പ്രതീക്ഷ ലഭിക്കുന്നതു്. ഓരോ സാമ്പിളിനും ഓരോ സമാന്തരമാധ്യമുണ്ടായിരിക്കുമല്ലോ. അവയാണ് ഈ ആകലത്തിന്റെ മൂല്യങ്ങൾ. ലഘു യാദൃച്ഛിക സാമ്പിളനത്തിൽ ഏതു സാമ്പിൾ കിട്ടുവാനുമുള്ള സംഭാവ്യത തുല്യമായതു കൊണ്ടു് ആകലത്തിന്റെ എല്ലാ മൂല്യങ്ങളുടെയും സംഭാവ്യത തുല്യമാണ്. അങ്ങനെ വരുമ്പോൾ അതിന്റെ പ്രതീക്ഷ നിശ്ചിത വലുപ്പമുള്ളതും സാധ്യമായതുമായ എല്ലാ സാമ്പിളുകളെയും പരിഗണിച്ചു്, അവയുടെ എല്ലാം സമാന്തരമാധ്യം കണ്ടു് അങ്ങനെ കിട്ടിയ സമാന്തരമാധ്യങ്ങളുടെ തുകയെ സാമ്പിളുകളുടെ എണ്ണം കൊണ്ടു ഹരിച്ചതാണ്. ഈ ഉദാഹരണത്തിൽ 1 മുതൽ 6 വരെ അംഗങ്ങളുള്ള സാമ്പിളുകൾ എടുക്കാവുന്നതാണ്. ഇതിൽ 1 മുതൽ 3 വരെ അംഗങ്ങളുള്ളതും സ്വീകാര്യവുമായ സാമ്പിളുകളും അവയുടെ സമാന്തരമാധ്യങ്ങളും താഴെ കൊടുക്കുന്നു ഇവ ഓരോന്നിലും സാമ്പിൾ സമാന്തരമാധ്യങ്ങളുടെ പ്രതീക്ഷ, അതായതു് അവയുടെ സമാന്തരമാധ്യം, സമഷ്ടിയുടെ സമാന്തരമാധ്യമായ 3.5 തന്നെയാണെന്നു കാണാം.

4, 5, 6 അംഗങ്ങളുള്ള സാമ്പിളുകളുടെയും സ്ഥിതി ഇതു തന്നെ. ഇതിൽ നിന്നു് സാമ്പിൾസമാന്തരമാധ്യം സമഷ്ടിയുടെ സമാന്തരമാധ്യത്തിന്റെ ഒരു അനഭിനത ആകലമാണെന്നു സിദ്ധിക്കുന്നു.

പട്ടിക 3.3

ഒരംഗമുള്ള സാമ്പിളുകളും അവയുടെ സമാന്തരമാധ്യങ്ങളും

സാമ്പിൾ	സാമ്പിൾ മൂല്യങ്ങൾ	സമാന്തരമാധ്യം
A	3	3
B	1	1
C	4	4
D	2	2
E	6	6
F	5	5
		21

$$\left.\begin{array}{r}\text{സമാന്തരമാധ്യങ്ങളുടെ}\\ \text{സമാന്തരമാധ്യം}\end{array}\right\} = \frac{21}{6} = 3.5$$

പട്ടിക 3.4

രണ്ടംഗങ്ങളുള്ള സാമ്പിളുകളും അവയുടെ സമാന്തരമാധ്യങ്ങളും

സാമ്പിൾ	സാമ്പിൾ മൂല്യങ്ങൾ	സമാന്തരമാധ്യം
A,B	3,1	2.0
A,C	3,4	3.5
A,D	3,2	2.5
A,E	3,6	4.5
A,F	3,5	4.0
B,C	1,4	2.5
B,D	1,2	1.5
B,E	1,6	3.5
B,F	1,5	3.0
C,D	4,2	3.0
C,E	4,6	5.0
C,F	4,5	4.5
D,E	2,6	4.0
D,F	2,5	3.5
E,F	6,0	5.5
		52.5

$$\left.\begin{array}{r}\text{സമാന്തരമാധ്യങ്ങളുടെ}\\ \text{സമാന്തരമാധ്യം}\end{array}\right\} = \frac{52.5}{15} = 3.5$$

പട്ടിക 3.5

മൂന്നംഗങ്ങളുള്ള സാമ്പിളുകളും അവയുടെ സമാന്തരമാധ്യങ്ങളും

സാമ്പിളുകൾ	സാമ്പിൾ മൂല്യങ്ങൾ	സമാന്തരമാധ്യങ്ങൾ
A,B,C	3,1,4	2.67
A,B,D	3,1,2	2.00
A,B,E	3,1,6	3.33
A,B,F	3,1,5	3.00
A,C,D	3,4,2	3.00
A,C,E	3,4,6	4.33
A,C,F	3,4,5	4.00
A,D,E	3,2,6	3.67
A,D,F	3,2,5	3.33
A,E F	3,6,5	4.67
B,C,D	1,4,2	2.33
B,C,E	1,4,6	3.67
B,C,F	1,4,5	3.33
B,D,E	1,2,6	3.00
B,D,F	1,2,5	2.67
B,E,F	1,6,5	4.00
C,D,E	4,2,6	4.00
C,D,F	4,2,5	3.67
C,E,F	4,6,5	5.00
D,E,F	2,6,5	4.33
		70.00

$$\left.\begin{array}{r}\text{സമാന്തരമാധ്യങ്ങളുടെ}\\ \text{സമാന്തരമാധ്യം}\end{array}\right\} = \frac{70.00}{20} = 3.5$$

ഒരു ആകലം അനഭിനതമായതു കൊണ്ട് ഏതെങ്കിലും ഒരു സാമ്പിളിൽ നിന്നു ലഭിക്കുന്ന അതിന്റെ മൂല്യം പ്രാചലത്തിന്റെ യഥാർഥമൂല്യമായിരിക്കുമെന്ന് അർഥമാകുന്നില്ല. ഉദാഹരണമായി A, B എന്ന രണ്ട് അംഗങ്ങളുള്ള ഒരു സാമ്പിൾ ലഭിക്കുന്നു എന്നിരിക്കട്ടെ. അതിന്റെ സമാന്തരമാധ്യം 2 ആണ്. പ്രാചലത്തിന്റെ മൂല്യം 3.5 ഉം. ഇതിൽ നിന്ന് അഭിനതിയില്ലായ്മ മാത്രം കൊണ്ട് ഒരു ആകലം സ്വീകാര്യമാകുന്നില്ലെന്നു വ്യക്തമാകുന്നു. ആകലം യഥാർഥമൂല്യത്തോട് കഴിവതും അടുത്തു വരികയാണ് നമുക്ക് ആവശ്യം. മുൻ പട്ടി

കകൾ പരിശോധിച്ചാൽ സാമ്പിളിൻെറ വലുപ്പം വർധിക്കും തോറും സാമ്പിൾ സമാന്തരമാധ്യങ്ങളിൽ കൂടുതൽ കൂടുതലെണ്ണം യഥാർത്ഥമൂല്യമായ 3.5 നോട്ട് അടുത്തു വരുന്നതായി കാണുവാൻ കഴിയും. സാമ്പിൾസമാന്തരമാധ്യങ്ങൾ 1 മുതൽ 6 വരെ വിചരണം ചെയ്യുന്നുണ്ടല്ലൊ. ആ അന്തരാളത്തെ തുല്യ ദൈർഘ്യമുള്ള 6 ക്ലാസുകളായി തിരിച്ച് സാമ്പിൾ പരിമാണമനുസരിച്ച് ഓരോ ക്ലാസിലും എത്ര ശതമാനം സമാന്തരമാധ്യങ്ങൾ വീതമാണ് വരുന്നതെന്ന് താഴെക്കൊടുത്തിരിക്കുന്ന പട്ടികയിൽ രേഖപ്പെടുത്തിയിരിക്കുന്നു.

പട്ടിക 3.6

സാമ്പിൾ പരിമാണവും വിവിധ സാമാന്തരമാധ്യങ്ങളോടു കൂടിയ സാമ്പിളുകളുടെ ശതമാനവും

സമാന്തര മാധ്യം	സാമ്പിൾ പരിമാണം					
	1	2	3	4	5	6
1—1.9	16.67	6.67	0	0	0	
2—2.9	16.67	20.00	20 00	13.33	0	
3—3 9	16.67	33 34	45.00	60.00	83.3	100.00
4—4.9	16 67	26.66	30.00	26.67	16.7	
5—5.9	16 67	13.33	5.00	0	0	
6—6.9	16 67	0	0	0	0	

ഈ പട്ടിക പരിശോധിച്ചാൽ സാമ്പിൾ പരിമാണം വർധിക്കുന്നതനുസരിച്ച് ലഭിക്കാവുന്ന സാമ്പിളുകളിൽ കൂടുതൽ കൂടുതൽ ശതമാനത്തിൻെറയും സമാന്തരമാധ്യം സമഷ്ടിയുടെ സമാന്തരമാധ്യത്തോട് അടുത്തു വരുന്നതായി കാണാം. സാമ്പിൾ പരിമാണം സമഷ്ടിയുടെ പരിമാണത്തോട് തുല്യമാവുമ്പോൾ രണ്ടിൻെറയും സമാന്തരമാധ്യം തുല്യമാവും. അതായതു് 100 ശതമാനം സാമ്പിളുകളുടെയും സമാന്തരമാധ്യം സമഷ്ടിയുടെ സമാന്തരമാധ്യം തന്നെയാവും. ഇതു് ശ്രദ്ധേയമായ ഒരു വൈശിഷ്ട്യമാണ്. ഇതാണ് ഒന്നാമധ്യായത്തിൽ നാം പറഞ്ഞുവെച്ച 'അവിരോധിതം' എന്നതു്. അഭിനതിയില്ലായ്മയും അവിരോധിതവുമുള്ള ഒരു ആകലം, സാമ്പിൾ പരിമാണം സാമാന്യം വലുതാണെങ്കിൽ വലിയ അവിശ്വാസം കൂടാതെ സ്വീകരിക്കുവാൻ സാമാന്യബുദ്ധി നമ്മോട് ഉപദേശിക്കുമല്ലൊ.

മുൻ കൊടുത്ത പട്ടികകൾ പരിശോധിച്ചാൽ മറെറാരു സംഗതി കൂടെ വ്യ

ക്ഷമാവും. ഒരു സമഷ്ടിയിൽ നിന്ന് എടുക്കാവുന്ന നിശ്ചിത പരിമാണമുള്ള എല്ലാ സാമ്പിളുകളും പരിഗണിച്ച് അവയിൽ നിന്നെല്ലാം സമാന്തരമാധ്യങ്ങൾ കണ്ടുപിടിച്ചു എന്നിരിക്കട്ടെ. അവയുടെ സമാന്തരമാധ്യം സമഷ്ടിയുടെ സമാന്തരമാധ്യം തന്നെയാണെന്നു നാം കണ്ടു കഴിഞ്ഞു. ഈ സമാന്തരമാധ്യത്തിൽ നിന്നുള്ള അവയുടെ പ്രകീർണനം, സാമ്പിൾ പരിമാണം വർധിക്കുന്നതനുസരിച്ച് കുറഞ്ഞു കുറഞ്ഞു വരുന്നു. അതു കൊണ്ടാണല്ലോ കൂടുതൽ കൂടുതൽ ശതമാനം സമാന്തരമാധ്യങ്ങൾ അവയുടെ സമാന്തരമാധ്യത്തോട് (പ്രതീക്ഷയോട്) അടുത്തടുത്തു വരുന്നത്. ഈ പ്രകീർണനത്തെ സ്വാഭാവികമായും സാമ്പിൾസമാന്തരമാധ്യങ്ങളുടെ പ്രസരണം കൊണ്ട് അളക്കാം. പ്രസരണം കുറയും തോറും ആകലമെന്ന നിലയിലുള്ള അതിന്റെ വിശ്വാസ്യത വർധിക്കുന്നു. ഇത് എല്ലാ ആകലങ്ങളെ സംബന്ധിച്ചും ശരിയാണ്. ആകലത്തിന്റെ പ്രതീക്ഷ പ്രാചലത്തിന്റെ യഥാർഥമൂല്യത്തോടു തുല്യമാണെങ്കിൽ (ആകലം അനഭിനതമാണെങ്കിൽ അതിന്റെ സാമ്പിളനപ്രസരണം (കിട്ടാവുന്ന എല്ലാ സാമ്പിളുകളിൽ നിന്നും അതിന്റെ മൂല്യം കണ്ടുപിടിച്ചാൽ, ആ മൂല്യങ്ങളുടെ പ്രസരണം) എത്ര കുറയുന്നോ അത്രയ്ക്കു വിശിഷ്ടമാണ് ആ ആകലം. സാമ്പിൾ പരിമാണം സമഷ്ടിയുടെ പരിമാണത്തോട് തുല്യമാവുമ്പോൾ ഇതു് പൂജ്യമാവുമെങ്കിൽ ആ ആകലത്തിന് അവിരോധിത്വം എന്ന വൈശിഷ്ട്യമുണ്ട്.

മുൻ ഉദാഹരണത്തിൽ സാമ്പിൾ പരിമാണമനുസരിച്ചുള്ള സാമ്പിൾസമാന്തരമാധ്യങ്ങളുടെ സാമ്പിളനപ്രസരണം താഴെ കൊടുക്കുന്നു.

പട്ടിക 3.7

സാമ്പിൾപരിമാണവും സമാന്തരമാധ്യങ്ങളുടെ സാമ്പിളനപ്രസരണവും

സാമ്പിൾപരിമാണം	സമാന്തരമാധ്യങ്ങളുടെ സാമ്പിളനപ്രസരണം
1	2.92
2	1.16
3	0.58
4	0.29
5	0.12
6	0

സാമ്പിൾപരിമാണം വർധിക്കുമ്പോൾ സമാന്തരമാധ്യങ്ങളുടെ പ്രസരണം കുറയുന്നതു് ശ്രദ്ധിക്കുക.

ഈ ചർച്ചയിൽ നിന്ന് ഒരു സംഗതി വ്യക്തമാകുന്നു. ആകലങ്ങളെ തിരഞ്ഞെടുക്കുന്നതിൽ പ്രധാനമായി പരിഗണിക്കാനുള്ളതു് അവയുടെ പ്രതീക്ഷയും സാമ്പിളനപ്രസരണവുമാണ്. ഈ അധ്യായത്തിലെ അടുത്തു വരുന്ന ഖണ്ഡികകളിൽ ലഘുയാദൃച്ഛികസാമ്പിളനത്തോടു ബന്ധപ്പെട്ട ആകലങ്ങളുടെ പ്രതീക്ഷയും സാമ്പിളനപ്രസരണവും ഗണിതശാസ്ത്രരീത്യാ വ്യുൽപ്പാദിപ്പിച്ചിരിക്കുന്നു.

സൂക്ഷ്മതയെപ്പററി ഒന്നാമധ്യായത്തിൻെറ അവസാനഖണ്ഡികയിൽ ചർച്ചചെയ്തതു് ഓർമിക്കുക. ആവശ്യമുള്ളത്ര സൂക്ഷ്മതയോടു കൂടിയ ആകലങ്ങൾ ലഭിക്കുവാൻ വേണ്ട സാമ്പിൾപരിമാണം എന്തായിരിക്കണമെന്നു് നിർണയിക്കാൻ ഇതു സഹായിക്കും. ഈ വിഷയം വീണ്ടും ചർച്ച ചെയ്യുന്നതാണു്.

4. ചില നിർവചനങ്ങളും അങ്കനരീതികളും

സാമ്പിൾസമാന്തരമാധ്യങ്ങളുടെ പ്രതീക്ഷയും സാമ്പിളനപ്രസരണവും സൂചിപ്പിക്കുന്ന പൊതുവാക്യങ്ങൾ ഗണിതരീത്യാ വ്യുൽപ്പാദിപ്പിക്കുകയാണു് നമുക്കു് അടുത്തതായി ചെയ്യാനുള്ളതു്. ഇതിനായി ചില നിർവചനങ്ങളും അങ്കനരീതികളും സ്വീകരിക്കുന്നതു് സൗകര്യപ്രദമായിരിക്കും. സമഷ്ടിയിൽ N അംഗങ്ങളാണുള്ളതെന്നും അതിൽ നിന്നു് n അംഗങ്ങളുള്ള ലഘു യാദൃച്ഛികസാമ്പിളാണു് എടുക്കുന്നതെന്നും കരുതുക. സമഷ്ടിയിലെ അഭിലക്ഷണമൂല്യങ്ങൾ $X_1, X_2, \ldots X_N$ എന്നിവയും സാമ്പിൾ മൂല്യങ്ങൾ $x_1, x_2, \ldots x_n$ എന്നിവയുമാണെന്നും സങ്കല്പിക്കുക. സമഷ്ടിയുടെ സമാന്തരമാധ്യത്തെ $\bar{X}$ എന്നതു കൊണ്ടും സാമ്പിൾസമാന്തരമാധ്യത്തെ $\bar{x}$ എന്നതു കൊണ്ടും സൂചിപ്പിക്കുക. അതായതു്,

$$\bar{X} = \frac{1}{N} \sum_{1}^{N} x_i \quad , \quad \bar{X} = \frac{1}{n} \sum_{1}^{n} x_i$$

സമഷ്ടിയുടെ പ്രസരണത്തെ σ^2 കൊണ്ടു് സൂചിപ്പിക്കാം. അതായതു്,

$$\sigma^2 = \frac{1}{N} \sum_{1}^{N} (X_i - \bar{X})$$

പക്ഷേ സമഷ്ടിയിലെ വ്യക്തികളുടെ വിഭിന്നതയെ കുറെക്കൂടി വ്യക്തമായി സൂചിപ്പിക്കുന്നതു് സമഷ്ടിയുടെ വർഗമാധ്യം S^2 കൊണ്ടാണു്. ഇവിടെ

$$S^2 = \frac{1}{N-1} \sum_{1}^{N} (X_i - \bar{X})^2$$

എന്നു് നിർവചിച്ചിരിക്കുന്നു. സമഷ്ടിയുടെ പരിമാണം (N) വലുതാകുമ്പോൾ പ്രസരണവും (σ^2) വർഗമാധ്യവും (S^2) തമ്മിലുള്ള അന്തരം കുറയുന്നു. അനന്തസമഷ്ടിയിൽ ഇവ രണ്ടും ഒന്നു തന്നെയാണു്. σ^2 ൻെറയും S^2 ൻെറയും നിർവചനങ്ങളിൽ നിന്നു്

$$\sigma^2 = \frac{N-1}{N} S^2 \text{ എന്നും}$$

$$S^2 = \frac{N}{N-1} \sigma^2 \quad \text{എന്നും ഗണിക്കാമല്ലൊ.}$$

σ^2-ഉം S^2-ഉം സമഷ്ടിയിലെ രണ്ടു പ്രാചലങ്ങളാണല്ലോ. ഇതു പോലെ സാമ്പിൾപ്രസരണത്തെ

$$V(x) = \frac{1}{n} \sum_{1}^{n} (x_i - \bar{x})^2$$

കൊണ്ടും സാമ്പിൾ വർഗമാധ്യത്തെ

$$s^2 = \frac{1}{n-1} \sum_{1}^{n} (x_i - \bar{x})^2$$

കൊണ്ടും സൂചിപ്പിക്കാം. സാമ്പിൾ പ്രസരണവും (v), സാമ്പിൾ വർഗമാധ്യവും (s^2) യാദൃച്ഛിക ചരങ്ങളാണ്. യാദൃച്ഛരിക സാമ്പിളനത്തിൽ, സാമ്പിൾ വർഗമാധ്യം (s^2) സമഷ്ടിവർഗമാധ്യത്തിന്റെ (S^2) ഒരു അനഭിനത ആകലമാണെന്ന് വഴിയെ കാണാം. എന്നാൽ സാമ്പിൾപ്രസരണം (V) സമഷ്ടി പ്രസരണത്തിന്റെ ഒരു അഭിനത ആകലമേ ആകുകയുള്ളൂ.

സമഷ്ടിയിലെയും സാമ്പിളനത്തിലെയും വർഗമാധ്യങ്ങളുടെ നിർവചനം സാമ്പിളനശാസ്ത്രത്തിലെ പല വാക്യങ്ങളും അവയുടെ ആകലങ്ങളും ലഘൂകരിച്ച് എഴുതാൻ സഹായിക്കും.

സാമ്പിൾമൂല്യങ്ങളുടെ ഏതു ഫലനത്തിനും സാമ്പിളജം എന്നു പേർ പറയും. സാമ്പിളജങ്ങളുടെ മൂല്യം സാമ്പിൾ മാറുന്നതനുസരിച്ച് മാറുമല്ലൊ. ഒരേ പരിമാണമുള്ള വിവിധ സാമ്പിളുകളിലുള്ള ഒരു സാമ്പിളജത്തിന്റെ പ്രസരണത്തിനു് അതിന്റെ സാമ്പിളനപ്രസരണം എന്നും, സാമ്പിളനപ്രസരണത്തിന്റെ വർഗമൂലത്തിനു് മാനകപ്പിശകു് (standard error) എന്നുമാണു് പേർ. സാമ്പിൾസമാന്തരമാധ്യത്തിന്റെ സാമ്പിളനപ്രസരണത്തെ $s^2_{\bar{x}}$

$V(\bar{x})$ എന്നീ ചിഹ്നങ്ങൾ കൊണ്ടും അതിന്റെ മാനകപ്പിശകിനെ $S_{\bar{x}}$, $\sqrt{V(\bar{x})}$ എന്നീ ചിഹ്നങ്ങൾ കൊണ്ടും സൂചിപ്പിക്കുന്നു. സാമ്പിൾ പരിമാണവും സമഷ്ടിയുടെ പരിമാണവും തമ്മിലുള്ള അനുപാതത്തിനു് സാമ്പിളനഭിന്നിതമെന്നു പറയാം. f എന്ന അക്ഷരം കൊണ്ടാണു് അതു് സൂചിപ്പിക്കാറുള്ളതു്. അതായതു്,

$$f = \frac{n}{N}$$

5. സാമ്പിൾ സമാന്തരമാധ്യത്തിന്റെ പ്രതീക്ഷ

(a) പ്രതിസ്ഥാപനമില്ലാത്ത ലഘു യാദൃച്ഛിക സാമ്പിളനം

N അംഗങ്ങളുള്ള ഒരു സമഷ്ടിയിൽ നിന്നു് n അംഗങ്ങളുള്ള സാമ്പിളാണു് എടുക്കുന്നതു് എന്നിരിക്കട്ടെ. ആകെ എടുക്കാവുന്ന സാമ്പിളുകളുടെ എണ്ണം $\binom{N}{n}$ ആണു്. ഇവ ഓരോന്നിനും തുല്യസംഭാവ്യത വരത്തക്കവണ്ണമാണല്ലോ

സാമ്പിളനരീതി ആസൂത്രണം ചെയ്തിരിക്കുന്നതു്. അതുകൊണ്ടു് ഇതിൽ ഏതെങ്കിലും ഒരു സാമ്പിൾ കിട്ടുവാനുള്ള സംഭാവ്യത $1 / \binom{N}{n}$ ആണു്.

സാമ്പിൾ മൂല്യങ്ങൾ $x_1, x_2, \ldots x_n$ എന്നിവയാണെന്നും അവയുടെ സമാന്തര മാധ്യം $\bar{x}$ ആണെന്നുമാണല്ലൊ നാം സങ്കൽപിച്ചിരിക്കുന്നതു്.

$$\bar{x} = \frac{1}{n} \sum_{i=1}^{n} x_i \tag{3.5.1}$$

$\bar{x}$ ഒരു യാദൃച്ഛിക ചരമാണു്. അതിൻെറ മൂല്യം സാമ്പിൾ മൂല്യങ്ങളെ ആശ്രയിച്ചാണിരിക്കുന്നതു്. നിർദിഷ്ട സമഷ്ടിയിൽ നിന്നു് എടുക്കാവുന്ന $\binom{N}{n}$ സാമ്പിളുകളെ ഏതെങ്കിലും ഒരു ക്രമമനുസരിച്ചു് പരിഗണിക്കുന്നു എന്നു വിചാരിക്കുക. അതിൽ t -ാമത്തെ സാമ്പിളിൻെറ സമാന്തരമാധ്യം $\bar{x}_t$ യും അതിലെ മൂല്യങ്ങൾ $x_{1t}, x_{2t} \ldots\ldots x_{nt}$ യും ആണെന്നു വിചാരിക്കുക.

$$\bar{x}_t = \frac{1}{n} \sum_{i=1}^{n} x_{it}$$

$\bar{x}$ ൻെറ പ്രതീക്ഷയാണു് നമുക്കു് ആദ്യമായി കണ്ടുപിടിക്കുവാനുള്ളതു്. നിർവചനമനുസരിച്ചു് $\bar{x}$ നു് എടുക്കാവുന്ന മൂല്യങ്ങളെ അവയുടെ സംഭാവ്യതകൾ കൊണ്ടു് ഗുണിച്ചു് ഒന്നിച്ചു കൂട്ടുമ്പോഴാണു് അതിൻെറ പ്രതീക്ഷ ലഭിക്കുന്നതു്. $\bar{x}$ നു് എടുക്കാവുന്നവ $\bar{x}_1, \bar{x}_2 \ldots \bar{x}_{\binom{N}{n}}$ തുടങ്ങിയ $\binom{N}{n}$ മൂല്യങ്ങളും അവയുടെ ഓരോന്നിൻെറയും സംഭാവ്യത $\frac{1}{\binom{N}{n}}$ ഉം ആയതുകൊണ്ടു് $\bar{x}$ ൻെറ പ്രതീക്ഷ, $\bar{x}_1, \bar{x}_2 \ldots\ldots$ എന്നിവയുടെ സമാന്തരമാധ്യം തന്നെയായിത്തീരുന്നു.

$$\text{അതായതു്, } E(\bar{x}) = \sum_{t=1}^{\binom{N}{n}} \bar{x}_t \Big/ \binom{N}{n}$$

$$= \frac{1}{\binom{N}{n}} \left\{ \sum_{t=1}^{\binom{N}{n}} \left[\frac{1}{n} \sum_{i=1}^{n} x_{it} \right] \right\} = \frac{1}{\binom{N}{n} n} \left\{ \sum_{t=1}^{\binom{N}{n}} \sum_{i=1}^{n} x_{it} \right\}$$

ബ്രാക്കററിലുള്ള സംഖ്യ നിർണയിക്കാനുള്ള മാർഗം, സമഷ്ടിയിലെ ഏതെങ്കിലും ഒരംഗം എത്ര സാമ്പിളുകളിലെ അംഗമായിരിക്കുമെന്ന് കണ്ടുപിടിക്കുകയാണ്. അതിനായി നിർദിഷ്ടമായ അംഗത്തെ മാററിനിർത്തുക. ബാക്കിയുള്ള $(N-1)$ അംഗങ്ങളിൽ നിന്നു് $(n-1)$ അംഗങ്ങളെ എത്ര തരത്തിൽ എടുക്കാമെന്നു പരിശോധിക്കുക. വ്യക്തമായും അതു് $\binom{N-1}{n-1}$ ആണു് ഇങ്ങനെ കിട്ടുന്ന $(n-1)$ അംഗങ്ങളുള്ള ഓരോ സാമ്പിളിലും മാററി നിർത്തിയിരിക്കുന്ന അംഗത്തെ കൂടി ഉൾപ്പെടുത്തിയാൽ ആ വ്യക്തി അംഗമായുള്ള സാമ്പിളുകൾ കിട്ടും. ഇതിൽ നിന്നു് സമഷ്ടിയിലെ ഓരോ അംഗവും $\binom{N-1}{n-1}$ സാമ്പിളുകളിലെ അംഗമാണെന്നു സിദ്ധിക്കുന്നു. ബ്രാക്കററിലുള്ള തുക എല്ലാ സാമ്പിളുകളിലെയും മൂല്യങ്ങൾ ഒന്നിച്ചു കൂട്ടിയതാണു്. അതായതു്, സമഷ്ടിയിലെ ഓരോ മൂല്യവും $\binom{N-1}{n-1}$. പ്രാവശ്യം വീതം ഒന്നിച്ചുകൂട്ടിയതാണു്. ഇതിൽ നിന്നു്,

$$\sum_{t=1}^{\binom{N}{n}} \sum_{i=1}^{n} x_{it} = \binom{N-1}{n-1} \sum_{i=1}^{N} X_i$$ എന്നു വരുന്നു. അതായതു്

$$E(\bar{x}) = \frac{1}{n\binom{N}{n}} \times \binom{N-1}{n-1} \sum_{i=1}^{N} X_i$$

$$= \frac{n!\,(N-n)!}{n \times N!} \times \frac{(N-1)!}{(n-1)!\,(N-n)!} \times \sum_{i=1}^{N} x_i$$

$$= \frac{1}{N} \sum_{i=1}^{N} x_i = \bar{X}$$

അങ്ങനെ സാമ്പിൾസമാന്തരമാധ്യത്തിൻെറ പ്രതീക്ഷ സമഷ്ടിയുടെ സമാന്തരമാധ്യമാണെന്നു സിദ്ധിക്കുന്നു. അതായതു്, സാമ്പിൾസമാന്തരമാധ്യം സമഷ്ടിയുടെ സമാന്തരമാധ്യത്തിൻെറ ഒരു അനഭിനത ആകലമാണു്.

കുറിപ്പ്: താത്വികമായ വ്യക്തതയ്ക്കു വേണ്ടിയാണു് $E(\bar{x}) = \bar{X}$ എന്നു തെളിയിക്കുവാൻ ഈ രീതി അവലംബിച്ചതു്. ഇതു് മറെറാരു തരത്തിലും തെളിയിക്കുവാൻ കഴിയും.

$$\bar{x} = \frac{1}{n}\sum_{i=1}^{n} x_i$$

$$E(\bar{x}) = \frac{1}{n}\sum_{i=1}^{n} E(x_i)$$ (തുകയുടെ പ്രതീക്ഷ, പ്രതീക്ഷകളുടെ തുകയാണല്ലോ)

പക്ഷേ,

$$E(x_i) = X_1 P_1 + X_2 P_2 + \cdots + X_N P_N$$

[x_i ക്ക് എടുക്കാവുന്ന മൂല്യങ്ങൾ $x_1, x_2, \ldots x_N$ എന്നിവയാണ്. $P_1, P_2 \cdots\cdots P_N$ എന്നിവ, അവ ഓരോന്നുമാവാനുള്ള സംഭാവ്യതകളും.]

സമഷ്ടിയിലെ ഏതൊരംഗത്തിനും സാമ്പിളനത്തിൻെറ ഏതു ഘട്ടത്തിലും തിരഞ്ഞെടുക്കപ്പെടുവാനുള്ള സംഭാവ്യത $\frac{1}{N}$ ആണെന്നു് 3.2 ൽ തെളിയിച്ചിട്ടുണ്ടു്. അതുകൊണ്ടു്

$$E(x_i) = \left(X_1 + X_2 + \cdots + X_N\right)\frac{1}{N} = \bar{X}$$

അങ്ങനെ $$E(\bar{x}) = \frac{1}{n}\, n\bar{X} = \bar{X}.$$

എന്നു ലഭിക്കുന്നു. സാമ്പിളജങ്ങളുടെ പ്രതീക്ഷ നിർണയിക്കാൻ ഈ രീതികളിൽ ഏതു വേണമെങ്കിലും സൗകര്യം പോലെ ഉപയോഗിക്കാവുന്നതാണ്. രണ്ടാമത്തെ രീതി താരതമ്യേന കൂടുതൽ ലളിതമാണെന്നുള്ളതു് ശ്രദ്ധേയമാണ്.

(b) പ്രതിസ്ഥാപനത്തോടു കൂടിയ ലഘുയാദൃച്ഛികസാമ്പിളനം

പ്രതിസ്ഥാപനത്തോടു കൂടിയ സാമ്പിളനത്തിലും ഏതെങ്കിലും ഘട്ടത്തിൽ സമഷ്ടിയിലെ ഏതെങ്കിലും അംഗത്തിനു് സാമ്പിളിൽ ഉൾപ്പെടുവാനുള്ള സംഭാവ്യത $\frac{1}{N}$ തന്നെയാണ്. i-ാമതു് തിരഞ്ഞെടുക്കപ്പെട്ട അംഗം x_i ആണല്ലോ. അതു് $X_1, X_2 \ldots X_N$ എന്നിവയിൽ ഏതെങ്കിലും ആകാമെന്നുള്ളതുകൊണ്ടും അവ ഓരോന്നും ആകുന്നതിനുള്ള സംഭാവ്യത $\frac{1}{N}$ ആയതുകൊണ്ടും

$$E(x_i) = \frac{1}{n}(X_1 + X_2 + \cdots + X_n) = \bar{X}$$

അങ്ങനെ $$E(\bar{x}) = E\left(\frac{1}{n}\sum_{i=1}^{n} x_i\right)$$

$$= \frac{1}{n} \sum_{i=1}^{n} E(x_i) = \frac{1}{n} n\overline{X} = \overline{X}$$

അതായതു്, പ്രതിസ്ഥാപനത്തോടു കൂടിയ സാമ്പിളനത്തിലും സാമ്പിൾസമാന്തരമാധ്യം സമഷ്ടിയുടെ സമാന്തരമാധ്യത്തിൻെറ അനഭിനത ആകലമാണു്.

അടുത്തതായി അന്വേഷിക്കുവാനുള്ളതു് ഈ രണ്ടു് സാമ്പിളനരീതികളിലും സാമ്പിൾസമാന്തരമാധ്യങ്ങളുടെ പ്രസരണം എന്തായിരിക്കുമെന്നുള്ളതത്രേ. സാമ്പിളനരീതികളുടെ ദക്ഷത താരതമ്യപ്പെടുത്താൻ അതു് ഒരു പക്ഷേ ഉപകരിച്ചേക്കാം.

6. സാമ്പിൾസമാന്തരമാധ്യത്തിൻെറ പ്രസരണം

(a) പ്രതിസ്ഥാപനമില്ലാത്ത ലഘുയാദൃച്ഛിക സാമ്പിളനം.

സാമ്പിൾ സമാന്തരമാധ്യമായ $\overline{x}$ ഒരു യാദൃച്ഛികചരമാണു്. അതിൻെറ പ്രസരണമായ $V(\overline{x})$ ആണു് നമുക്കു് കണ്ടുപിടിക്കേണ്ടതു്. (സാമ്പിൾ മാറുന്നതനുസരിച്ചു് $\overline{x}$ മാറുന്നതുകൊണ്ടാണു് അതു് യാദൃച്ഛികചരമാവുന്നതു്.)

$V(\overline{x}) = E(\overline{x} - \overline{X})^2$ എന്നു് എഴുതാം.

അതായതു്, $$S^2_{\overline{x}} = E\left\{\frac{1}{n}\sum_{i=1}^{n}(x_i - \overline{X})\right\}^2$$

$$= \frac{1}{n^2} E\left\{\sum_{i=1}^{n}(x_i - \overline{X})^2 + \sum_{i\neq j=1}^{n}(x_i - \overline{X})(x_j - \overline{X})\right\}$$

$$= \frac{1}{n^2}\sum_{i=1}^{n} E(x_i - \overline{X})^2 + \frac{1}{n^2}\sum_{i=j=1}^{n} E(x_i - \overline{X})(x_j - \quad)$$

സാമ്പിളനത്തിന്റെ i-ാമത്തെ ഘട്ടത്തിൽ സമഷ്ടിയിലെ ഏതെങ്കിലും ഒരംഗം തിരഞ്ഞെടുക്കപ്പെടുവാനുള്ള സംഭാവ്യത $\frac{1}{N}$ ആയതുകൊണ്ടു്

$$E(x_i - \overline{X})^2 = \frac{1}{N}\sum_{i=1}^{N}(x_i - X)^2 = \sigma^2 = \frac{N-1}{N} S^2$$

എന്നു സിദ്ധിക്കുന്നു. i–ാമത്തെ ഘട്ടത്തിൽ സമഷ്ടിയിലെ x_1 എന്ന ഏതെങ്കിലും ഒരംഗവും j-ാമത്തെ ഘട്ടത്തിൽ x_m എന്ന മറേറതെങ്കിലും അംഗവും തിരഞ്ഞെടുക്കപ്പെടുവാനുള്ള സംഭാവ്യത $\frac{1}{N(N-1)}$ ആണു്.

അതുകൊണ്ടു്,

$$E(x_i-\bar{X})(x_j-\bar{X}) = \frac{1}{N(N-1)}\sum_{l\neq m=1}^{N}(x_l-\bar{X})(x_m-\bar{X})$$

$$= \frac{1}{N(N-1)}\left\{\left[\sum_{l=1}^{N}(x_l-\bar{X})\right]^2 - \sum_{l=1}^{N}(x_l-\bar{X})^2\right\}$$

$$= -\frac{1}{N(N-1)}\sum_{l=1}^{N}(x_l-\bar{X})^2 \quad = -\frac{S^2}{N-1}$$

$$= -\frac{1}{N-1}\sigma^2 = -\frac{1}{N}S^2, \because \sum_{l=1}^{N}(x_l-\bar{X})=0$$

ഇതിൽ നിന്നു്,

$$V(\bar{x}) = \frac{1}{n^2}\cdot n\sigma^2 + \frac{n(n-1)}{n^2}\left(-\frac{\sigma^2}{N-1}\right)$$

$$= \frac{1}{n}\frac{N-1}{N}S^2 - \frac{n-1}{n}\frac{1}{N}S^2$$

$$= \frac{N-n}{N}\left(\frac{S^2}{n}\right) \qquad (3.6.1)$$

ഇതിൽ നിന്നു് $\bar{x}$ ൻെറ മാനകപ്പിശക് ഇതിൻെറ വർഗമൂലമായ $\sqrt{(N-n)/N}\ \frac{S}{\sqrt{n}}$ ആണെന്നു സിദ്ധിക്കുന്നു. n വർധിക്കുമ്പോൾ ഇതിൻെറ മൂല്യം കുറഞ്ഞുവരുന്നു എന്നതു് പ്രത്യേകം ശ്രദ്ധിക്കേണ്ടതാണു്.

b) പ്രതിസ്ഥാപനത്തോടുകൂടിയ ലഘുയാദൃച്ഛിക സാമ്പിളനം

ഇവിടെയും സമീപനരീതി ഇതുതന്നെ.

$$V(x) = E(\bar{x}-\bar{X}^2) = E\left\{\frac{1}{n}\sum_{i=1}^{n}(x_i-X)\right\}^2$$

$$= \frac{1}{n^2}E\left\{\sum_{i=1}^{n}(x_i-\bar{X})^2 + \sum_{i\neq j=1}^{n}(x_i-X)(x_j-\bar{X})\right\}$$

$$= \frac{1}{n^2} \sum_{i=1}^{n} E(x_i - \bar{X})^2 + \frac{1}{n^2} \sum_{i \neq j = 1}^{n} E(x_i - \bar{X})(x_j - \bar{X})$$

ഇതിൽ $E(x_i - \bar{X})^2 = \frac{1}{N} \sum_{i=1}^{N} (x_i - \bar{X})^2 = \sigma^2$

$$E(x_i - \bar{X})(x_j - \bar{X}) = \frac{1}{N^2} \sum_{l, m=1}^{N} (x_l - \bar{X})(x_m - \bar{X})$$

[സാമ്പിളനം പ്രതിസ്ഥാപനത്തോട്ടുകൂടിയതായതു കൊണ്ട് i-ാമത്തെ ഘട്ടത്തിൽ x_l, j-ാമത്തെ ഘട്ടത്തിൽ x_m എന്നിവ കിട്ടുവാനുള്ള സംഭാവ്യത $\frac{1}{N^2}$ ആണ്. x_l, x_m എന്നിവ വ്യത്യസ്തങ്ങളായിരിക്കണമെന്നില്ലല്ലൊ.]

$$= \frac{1}{N^2} \left[\sum (x_l - \bar{X}) \right]^2 = 0$$

$$V(\bar{x}) = \frac{\sigma^2}{n} \text{ അല്ലെങ്കിൽ } \frac{N-1}{Nn} S^2 \qquad (3.6.2)$$

$\bar{x}$ ന്റെ മാനകപ്പിശക് $\frac{\sigma}{\sqrt{n}}$ അല്ലെങ്കിൽ $\sqrt{(N-1)/nN}\, S$ എന്നു വരുന്നു. n വർധിക്കുമ്പോൾ ഇതു കുറയുമെന്നു വ്യക്തം.

രണ്ടു തരം സാമ്പിളനത്തിലും $\bar{x}$ ന്റെ പ്രസരണം താരതമ്യപ്പെടുത്തുക. പ്രതിസ്ഥാപനത്തോട്ടു കൂടിയ സാമ്പിളനത്തിലെ $\bar{x}$ ന്റെ പ്രസരണത്തിന്റെ $\frac{N-n}{N-1}$ ഭാഗമാണ് പ്രതിസ്ഥാപനമില്ലാത്ത സാമ്പിളനത്തിലെ $\bar{x}$ ന്റെ പ്രസരണം. അതായത്, പ്രതിസ്ഥാപനമില്ലാത്ത സാമ്പിളനത്തിലെ $\bar{x}$ ന്റെ പ്രസരണം പ്രതിസ്ഥാപനത്തോട്ടു കൂടിയ സാമ്പിളനത്തിലെ $\bar{x}$ ന്റെ പ്രസരണത്തെക്കാൾ കുറവാണ്. ഇതാണ് പ്രതിസ്ഥാപനമില്ലാത്ത സാമ്പിളനത്തിന്റെ വൈശിഷ്ട്യവും.

7. $V(\bar{x})$ ന്റെ ഒരു അനഭിനത ആകലം

$\bar{x}$ ന്റെ പ്രസരണം സമഷ്ടിയുടെ വർഗമാധ്യമായ σ^2 ന്റെ ഒരു ഗുണിതമായിട്ടാണ് നമുക്കു ലഭിച്ചത്. പക്ഷേ S^2 നമ്മെ സംബന്ധിച്ചിടത്തോളം അജ്ഞാതമാണ്. ഈ പ്രസരണത്തിന്റെ പ്രാധാന്യം പരിഗണിക്കുമ്പോൾ

സാമ്പിൾ ഉപയോഗിച്ചു് അതു് ആകലനം ചെയ്യേണ്ടതു് ആവശ്യമാണു്. സ്വാഭാവികമായും S^2 ന്റെ ആകലമായി സാമ്പിൾ വർഗമാധ്യം s^2 സ്വീകരിക്കാൻ സഹജാവബോധം നമ്മെ പ്രേരിപ്പിക്കുന്നു. അതു് ഒരു അനഭിനത ആകലമാണോ എന്നു പരിശോധിച്ചുനോക്കാം.

$$s^2 = \frac{1}{n-1} \sum_{i=1}^{n} (x_i - \bar{x})^2$$

ആണു് സാമ്പിൾ വർഗമാധ്യം.

$$E(s^2) = E\left\{ \frac{1}{n-1} \sum_{i=1}^{n} (x_i - \bar{x})^2 \right\}$$

$$= \frac{1}{n-1} E\left\{ \sum_{i=1}^{n} x_i^2 - n\bar{x}^2 \right\}$$

$$= \frac{1}{n-1} \left\{ \sum_{i=1}^{n} E(x_i^2) - n\,E(\bar{x}^2) \right\}$$

$$= \frac{1}{n-1} \left\{ \sum_{i=1}^{n} \sum_{j=1}^{N} \frac{X_j^2}{N} - n\,E(\bar{x}^2) \right\}$$

$$= \frac{1}{n-1} \left\{ n \sum_{j=1}^{N} \frac{X_j^2}{N} - n\,E(\bar{x}^2) \right\}$$

ഇതിൽ,

$$\frac{\sum_{j=1} X_j^2}{N} = \frac{\sum_{j=1}^{N} X_j^2}{N} - \left(\frac{\sum_{j=1}^{N} X_j}{N} \right)^2 + \left(\frac{\sum_{j=1}^{N} X_j}{N} \right)^2$$

$$= \frac{N-1}{N} S^2 + \bar{X}^2$$

അതു പോലെ തന്നെ,

$$E(s^2) = \frac{n}{n-1}\left\{\frac{N-1}{N}S^2 + \bar{x}^2 - V(\bar{x}) - \bar{x}^2\right\}$$

$$= \frac{n}{n-1}\left\{\frac{N-1}{N}S^2 - V(\bar{x})\right\}$$

$$= \frac{n}{n-1}\left\{\frac{N-1}{N}S^2 \quad \frac{N-n}{N}\,\frac{S^2}{n}\right\}$$

$$= \frac{S^2}{N(N-1)}\left\{n(N-1) - (N-n)\right\}$$

$$= S^2$$

പ്രതിസ്ഥാപനമില്ലാത്ത യാദൃച്ഛിക സാമ്പിളനത്തിൽ അതായതു് s^2 എന്ന സാമ്പിളജം S^2 ന്റെ അനഭിനത ആകലമാണു്. ഈ ആകലം $V(\bar{x})$ ന്റെ വ്യംജകത്തിൽ പ്രതിസ്ഥാപിച്ചാൽ $V(\bar{x})$ ന്റെ ആകലം ലഭിക്കും. പ്രസരണത്തിന്റെ ആകലത്തെ v കൊണ്ടു് കുറിക്കുന്നു. പ്രതിസ്ഥാപനമില്ലാത്ത സാമ്പിളനത്തിൽ അതു് $V(\bar{x})$ ന്റെ ആകലം

$$v(\bar{x}) = \frac{N-n}{N}\cdot\frac{S^2}{n}$$ ആകുന്നു.

പ്രതിസ്ഥാപനത്തോടു കൂടിയ സാമ്പിളനത്തിൽ സാമ്പിൾവർഗമാധ്യം s^2, സമഷ്ടിയുടെ പ്രസരണമായ σ^2 ന്റെ അനഭിനതആകലമാണു്.

$$E(s^2) = E\left[\sum_{i=1}^{n}\frac{(x_i - \bar{x})^2}{n-1}\right]$$

$$= \frac{1}{n-1}\left[\sum_{i=1}^{n}E(x_i^2) - n\,E(\bar{x})^2\right]$$

$$= \frac{n}{n-1}\left[\sum_{i=1}^{N}\frac{x_i^2}{N} - V(\bar{x}) - \bar{x}^2\right]$$

$$= \frac{n}{n-1}\left[\sum^{N}\frac{x_i^2}{N} - \bar{x}^2 - \frac{\sigma^2}{n} - \bar{X}^2\right] = \frac{n}{n-1}\left[\sigma^2 - \frac{\sigma^2}{n}\right] = \sigma^2$$

ആയതിനാൽ പ്രതിസ്ഥാപനത്തോട്ടു കൂടിയ സാമ്പിളനത്തിൽ $V(\bar{x}) = \frac{\sigma^2}{n}$ന്റെ

അനഭിനത ആകലം $v(\bar{x}) = \frac{s^2}{n}$ ആകുന്നു.

കുറിപ്പ്: സമഷ്ടിയുടെ വലുപ്പം അനന്തമാണെങ്കിൽ (അതായതു് N അനന്തമാണെങ്കിൽ, $V(\bar{x}) = \frac{S^2}{n}$ എന്നു കാണുവാൻ പ്രയാസമില്ല. സമഷ്ടി സാന്തമാവുമ്പോൾ (പ്രതിസ്ഥാപനമില്ലാത്ത സാമ്പിളനത്തിൽ) അതു്

$\frac{N-n}{N}\frac{S^2}{n}$ ആണു്.

ഇവ താരതമ്യപ്പെടുത്തിയാൽ $\frac{N-n}{N}$ എന്ന ഘടകം കൊണ്ടാണു് ഇവ വ്യത്യാസപ്പെട്ടിരിക്കുന്നതെന്നു കാണാം. ഇതിനു് സാന്തസമഷ്ടിസംശോധനം എന്നു

പേർ പറയും. $\frac{N-n}{N} = 1 - \frac{n}{N}$ ആയതു കൊണ്ടു്

സാമ്പിളനഭിന്നം $\left(\frac{n}{N}\right)$ ചെറുതായിരുന്നാൽ ഈ സംശോധനം കാര്യമായ മാററമൊന്നും വരുത്തുകയില്ല. അതുകൊണ്ടു് സാധാരണയായി സാമ്പിളനഭിന്നം 0.1 ഓ അതിൽ കുറവോ ആയിരിക്കുമ്പോൾ ഈ സംശോധനം ഒഴിവാക്കാറുണ്ടു്. അതായതു് $\frac{S^2}{n}$ സാമ്പിൾസമാന്തരമാധ്യത്തിന്റെ പ്രസരണമായി സ്വീകരിച്ച് തുടർന്നുള്ള ഗണിതക്രിയകൾ നടത്തും.

8. വിശ്വാസ്യതാന്തരാളത്തിന്റെ ആകലനം

സാമ്പിൾസമാന്തരമാധ്യം സമഷ്ടിയുടെ സമാന്തരമാധ്യത്തിന്റെ ഒരു അനഭിനത ആകലമാണെന്നു നാം കണ്ടുകഴിഞ്ഞു. പക്ഷേ ഇതു് ഒരു ബിന്ദുആകലമാണു്. പലപ്പോഴും ആവശ്യമായി വരുന്നതു് വിശ്വാസ്യതാന്തരാളങ്ങളുടെ ആകലനമാവും. അതിനുള്ള മാർഗമാണു് ഇവിടെ ചർച്ച ചെയ്യുന്നതു്.

സാമ്പിൾസമാന്തരമാധ്യമായ $\bar{x}$ തന്നെ ഇതിനും ഉപയോഗിക്കാം. $\bar{x}$ ന്റെ സാമ്പിളനവിതരണത്തെ അടിസ്ഥാനമാക്കിയാണു് ഇതു് നിർണ്ണയിക്കുന്നതു്.

നമ്മുടെ ലക്ഷ്യം (a, b) എന്ന അന്തരാളം $\overline{X}$ നെ ഉൾക്കൊള്ളുവാനുള്ള സംഭാവ്യത ഒരു അഭിലഷിതസംഖ്യയായ a ആയിരിക്കത്തക്ക വണ്ണം, a, b എന്നീ സംഖ്യകൾ നിർണയിക്കുക എന്നതാണ്. ഇങ്ങനെ ലഭിക്കുന്ന അന്തരാളത്തിൻെറ വിശ്വാസ്യതാഗുണാങ്കം a യാണെന്നു പറയും. $\bar{x}$ ൻെറ സാമ്പിളനവിതരണം സമഷ്ടിയുടെ വിതരണത്തെ ആശ്രയിച്ചാണിരിക്കുന്നത്. സമഷ്ടി നോർമൽ വിതരണത്തോടുകൂടിയതും സമഷ്ടിയുടെ പ്രസരണമായ σ^2 നേരത്തെ അറിയാവുന്നതുമാണെങ്കിൽ,

$$t = \frac{\bar{x} - \overline{X}}{S_{\bar{x}}}$$

മാനകനോർമൽ വിതരണത്തോടുകൂടിയതായിരിക്കുമെന്നുള്ള സാംഖ്യികസിദ്ധാന്തം നമുക്ക് ഉപയോഗിക്കാം. 1.7 ൽ കൊടുത്തിരിക്കുന്ന പട്ടികയോ അതുപോലുള്ള മററു പട്ടികകളോ (അനുബന്ധം 2) ഉപയോഗിച്ചാൽ t യുടെ മൂല്യം (−k, k) എന്ന അന്തരാളത്തിൽ ആയിരിക്കുവാനുള്ള സംഭാവ്യത a ആയിരിക്കത്തക്കവണ്ണം k യുടെ മൂല്യം നിർണയിക്കാൻ കഴിയും. ഉദാഹരണമായി $a = .99$ എന്നിരിക്കട്ടെ. പട്ടികയിൽ നിന്നു്

$$P_r \left\{ -2.576 \leqslant \frac{\bar{x} - \overline{X}}{S_{\bar{x}}} \leqslant 2\ 576 \right\} = .99$$

എന്നു കിട്ടുന്നു. ഇതിൽ നിന്നു് $\overline{X}$ ൻെറ 0.99 വിശ്വാസ്യതാഗുണാങ്കത്തോടുകൂടിയ അന്തരാളം കണ്ടുപിടിക്കാം. നോക്കുക.

$$P_r \left\{ -2.576\, S_{\bar{x}} \leqslant \bar{x} - \overline{X} \leq 2.576 S_{\bar{x}} \right\} = .99$$

അതായതു്,

$$P_r \left\{ \bar{x} - 2.576\ S_{\bar{x}} \leq \overline{X} \leq \bar{x} + 2.576\ S_{\bar{x}} \right\} = .99$$

അങ്ങനെ

$$\left\{ \bar{x} - 2.576\ S_{\bar{x}} ,\ \ \bar{x} + 2.576\ S_{\bar{x}} \right\}$$

എന്ന അന്തരാളം $\overline{X}$ നെ ഉൾക്കൊള്ളുവാനുള്ള സംഭാവ്യത 0.99 ആണെന്നു വരുന്നു. പ്രതിസ്ഥാപനമില്ലാത്ത ലഘുയാദൃച്ഛികസാമ്പിളനത്തിൽ, ഇതു്

$$\left\{ \bar{x} - 2.576 \sqrt{(N-n)/\,nN}\, S, \bar{x} + 2.576 \sqrt{(N-n)/\,nN}\, S \right\}$$

എന്നായിത്തീരും. ഇതു് $\bar{x}$ ൻെറ 0.99 വിശ്വാസ്യതാഗുണാങ്കത്തോടുകൂടിയ വിശ്വാസ്യതാന്തരാളമാണ്. പൊതുവെ പറഞ്ഞാൽ, സമഷ്ടിയുടെ വിതരണം നോർമലായിരിക്കുകയും അതിൻെറ പ്രസരണമായ σ^2 മുൻകൂട്ടി അറിയാമായിരിക്കുക

യും ചെയ്താൽ സമാന്തരമാധ്യമായ $\overline{x}$ ന്റെ a വിശ്വാസ്യതാഗുണാങ്കത്തോടു കൂടിയ വിശ്വാസ്യതാന്തരാളം

$$\left\{\overline{x} - t_\alpha S_{\overline{x}} , \overline{x} + t_\alpha S_{\overline{x}}\right\}$$

എന്നതാണ്. ഇവിടെ t_α നിർവചിക്കപ്പെട്ടിരിക്കുന്നത് t എന്ന യാദൃച്ഛിക ചരം മാനകനോർമൽവിതരണത്തോടു കൂടിയതാണെങ്കിൽ,

$$P_r \left\{ - t_\alpha < t < t_\alpha \right\} = a$$

എന്ന സമീകരണത്താലാണ്. a യുടെ എതെങ്കിലും മൂല്യത്തിനു സംഗതമായ t_α യുടെ മൂല്യം മാനകനോർമൽസമാകലപട്ടികകളിൽ നിന്ന് കണ്ടുപിടിക്കാവുന്നതാണ്.

σ^2 ന്റെ മൂല്യം അറിഞ്ഞുകൂടെങ്കിൽ ഈ രീതിയിൽ വിശ്വാസ്യതാന്തരാളം നിർണയിക്കാനാവില്ല. അങ്ങിനെയുള്ള സന്ദർഭത്തിൽ

$$t = \frac{\overline{x} - \overline{X}}{S'_{\overline{x}}}$$

($S'_{\overline{x}}$ എന്നതു് $S_{\overline{x}}$ൽ s^2 ന്റെ ഒരു അനഭിനതആകലം പ്രതിസ്ഥാപിച്ചാൽ കിട്ടുന്നതാണ്) എന്ന യാദൃച്ഛികചരം $n-1$ സ്വാതന്ത്ര്യതാസംഖ്യയോടുകൂടിയ സ്റ്റുഡൻറു് t വിതരണം അനുവർത്തിക്കുന്നു എന്ന സാംഖ്യിക സിദ്ധാന്തം ഉപയോഗപ്പെടുത്താം.

t വിതരണത്തിന്റെ സമാകലപട്ടിക ഉപയോഗിച്ച്,

$$P_r \left[- t_\alpha \leq t \leq t_\alpha \right] = a$$

ആയിരിക്കത്തക്ക വണ്ണം t_α നിർണയിക്കാം. ആദ്യത്തേതു പോലെ തന്നെ വിശ്വാസ്യതാന്തരാളവും കണ്ടുപിടിക്കാം.

$$\left[\overline{x} - t_\alpha S'_{\overline{x}} , \overline{x} + t_\alpha S'_{\overline{x}} \right]$$

ആയിരിക്കും ആ അന്തരാളം.

സമഷ്ടിയുടെ വിതരണം നോർമലല്ലെങ്കിൽ മുൻപറഞ്ഞ രണ്ടു രീതികളും വിശ്വാസ്യതാന്തരാള നിർണയത്തിനു് ഉപകരിക്കുകയില്ല. പക്ഷെ വലിയ സാമ്പിളാണ് എടുക്കുന്നതെങ്കിൽ

$$\frac{\overline{x} - \overline{X}}{S'_{\overline{x}}}$$

ഏകദേശമായി മാനകനോർമൽവിതരണത്തോടു കൂടിയതായിരിക്കും. സാമ്പിളിനു് എത്ര വലുപ്പമുള്ളപ്പോഴാണ് ഈ ഏകദേശനം സ്വീകാര്യമാവുന്നതെന്ന് കൃത്യമായി പറയാനാവില്ല. സമഷ്ടിയുടെ വിതരണത്തെപ്പററി ലഭ്യമായ വി

വരങ്ങൾ ഉപയോഗിച്ച് ഒരു തീരുമാനമെടുക്കുകയേ നിർവാഹമുള്ളു. അങ്ങനെ വലിയ സാമ്പിളുകൾ ലഭിച്ചാൽ $\overline{X}$ ന്റെ വിശ്വാസ്യതാന്തരാളം ഏകദേശമായി നിർണയിക്കാം.

9. സാമ്പിൾപരിമാണം നിർണയിക്കൽ

സാമ്പിളന ആസൂത്രണം ചെയ്യുമ്പോൾ പ്രത്യേകം പരിഗണിക്കേണ്ടിവരുന്ന ഒരു പ്രശ്നമാണ് എന്തു വലുപ്പമുള്ള ഒരു സാമ്പിളാണ് എടുക്കേണ്ടതെന്നത്. വളരെ ചെറിയ സാമ്പിളാണെടുക്കുന്നതെങ്കിൽ അതിൽ നിന്നു ലഭിക്കുന്ന ആകലങ്ങൾ ആവശ്യമുള്ളത്ര സൂക്ഷ്മതയുള്ളതല്ലെന്നു വരാവുന്നതാണ്. അതു് പോലെ തന്നെ വളരെ വലിയ ഒരു സാമ്പിളെടുക്കുന്നതു് അനാവശ്യമായ പണച്ചെലവിനും ബുദ്ധിമുട്ടിനും സമയനഷ്ടത്തിനും ഇടയാക്കുകയും ചെയ്യും. ആവശ്യത്തിനുതകുന്ന ഏററവും ചെറിയ സാമ്പിൾ എടുക്കുകയാണ് നമ്മുടെ ലക്ഷ്യം. സമഷ്ടിയെപ്പററിയുള്ള നമ്മുടെ അജ്ഞത ഇങ്ങനെ ഒരു തീരുമാനമെടുക്കുന്നതിനു് വിലങ്ങുതടിയാണ്. ഇങ്ങനെയുള്ള സന്ദർഭങ്ങളിൽ തൃപ്തികരമായ ഒരു തീരുമാനമെടുക്കാൻ സഹായിക്കുന്ന ചില അടിസ്ഥാനപ്രമാണങ്ങൾ ആവിഷ്ക്കരിക്കാനാണ് ഇവിടെ ശ്രമിക്കുന്നതു്.

നമ്മുടെ ലക്ഷ്യം $\overline{X}$ ആകലനം ചെയ്യുക എന്നതാണെന്നിരിക്കട്ടെ. $\overline{x}$ ആണ് അതിനുപയോഗിക്കുന്നതു്. ആകലവും യഥാർഥമൂല്യവും തമ്മിലുള്ള വ്യത്യാസം d എന്ന ചെറിയ സംഖ്യയിൽ കുറഞ്ഞിരുന്നാൽ കൊള്ളാമെന്നു് നാം ആഗ്രഹിക്കുന്നു എന്നും സങ്കല്പിക്കുക. $\overline{X}$-ഉം $\overline{x}$-ഉം തമ്മിലുള്ള വ്യത്യാസം d യെക്കാൾ വലുതാവാനുള്ള സംഭാവ്യത എത്രയും കുറഞ്ഞിരിക്കണമെന്നു് സാരം. ആ സംഭാവ്യത β എന്ന സംഖ്യയിൽ കുറവായിരിക്കണം എന്നു് അഭിലഷിക്കുന്നു എന്നിരിക്കട്ടെ. $S^2_{\overline{x}}$ ന്റെ മൂല്യം സാമ്പിൾപരിമാണമായ n വർധിക്കുന്നതനുസരിച്ച് കുറയുമെന്നു നാം കണ്ടു കഴിഞ്ഞു. അതായതു്, $\overline{x}$ ന്റെ മൂല്യങ്ങൾ n വർധിക്കുന്നതനുസരിച്ച് $\overline{X}$ ന്റെ പരിസരത്തിൽ കേന്ദ്രീകരിക്കുന്നു. അപ്പോൾ നമ്മുടെ ലക്ഷ്യം നേടുവാനുള്ള മാർഗം

$$P_r\ [\,|\,\overline{x} - \overline{X}\,| \ \geq \ d] < \beta$$

ആയിരിക്കത്തക്കവണ്ണം n ന്റെ മൂല്യം നിർണയിക്കുകയാണ്.

$\overline{x}$ നോർമൽ വിതരണം അനുവർത്തിക്കുന്നു എന്നു സങ്കല്പിച്ചാൽ, $\dfrac{\overline{x} - \overline{X}}{S_{\overline{x}}}$ മാനകനോർമൽ വിതരണത്തോടു കൂടിയതാണെന്നു സിദ്ധിക്കും.

$$P_r \left\{ \left| \frac{\overline{x} - \overline{X}}{S_{\overline{x}}} \right| \geq t_\beta \right\} = \beta$$

ആയിരിക്കത്തക്കവണ്ണമുള്ള t_β നോർമൽപട്ടികയിൽ നിന്നും കണ്ടുപിടിക്കാം

അപ്പോൾ $t_\beta S_{\bar{x}} = a$ ആയിരിക്കത്തക്കവണ്ണം n നിർണയിക്കുകയാണു് നമുക്കു് കരണീയമായിട്ടുള്ളതു്.

അതായതു് (പ്രതിസ്ഥാപനമില്ലാത്ത യാദൃച്ഛികസാമ്പിളനത്തിൽ),

$$t_\beta \sqrt{(N-n)/N} \frac{S}{\sqrt{n}} = a$$

$$t_\beta^2 \frac{(N-n)S^2}{nN} = a^2$$

അതായതു്, $$n = \frac{\left(\frac{t_\beta S}{a}\right)^2}{1 + \frac{1}{N}\left(\frac{t_\beta S}{a}\right)^2}$$

N താരതമ്യേന വലുതാണെങ്കിൽ ആദ്യഏകദേശനമായി $n = \left(\frac{t_\beta S}{a}\right)$

എടുക്കാവുന്നതാണു്. $\frac{1}{N}\left(\frac{t_\beta S}{a}\right)^2$ അത്ര ചെറുതല്ലെങ്കിൽ n ന്റെ മൂല്യം ആദ്യം കൊടുത്തിരിക്കുന്നതു പോലെ തന്നെ കണ്ടുപിടിക്കണം. S ന്റെ മൂല്യം അറിയാമായിരിക്കുകയോ ലഭ്യമായ വിവരങ്ങൾ ഉപയോഗിച്ചു് അതു് ഏകദേശമായി നിർണയിക്കാൻ സാധിക്കുകയോ ചെയ്യണമെന്നുള്ളതു് പ്രത്യേകം ഓർമിക്കണം. പ്രതിസ്ഥാപനത്തോടു കൂടിയ സാമ്പിളിൽ

$$n = \left(\frac{t_\beta \sigma}{a}\right)^2$$

എന്നു് കാണാവുന്നതാണു്.

ഈ ചർച്ചകളിൽ നിന്നു് $\bar{x}$ ന്റെ സാമ്പിളനത്തിൽ പ്രസരണമായ $S^2_{\bar{x}}$ ന്റെ പ്രാധാന്യം വ്യക്തമായിരിക്കുമല്ലൊ. ചുരുക്കിപ്പറഞ്ഞാൽ, അതിന്റെ ഏററവും പ്രധാനപ്പെട്ട ഉപയോഗങ്ങൾ താഴെ കൊടുക്കുന്നവയാണു്.

1. ആകലത്തിന്റെ സൂക്ഷ്മത മനസ്സിലാക്കുക.
2. വിവിധ സാമ്പിളനരീതികളും ആകലനസമ്പ്രദായങ്ങളും താരതമ്യപ്പെടുത്തുക.
3. വിശ്വാസ്യതാന്തരാളങ്ങൾ നിർണയിക്കുക.
4. നിർദിഷ്ടസൂക്ഷ്മത ലഭിക്കാൻ ആവശ്യമായ സാമ്പിൾ പരിമാണം നിശ്ചയിക്കുക.

10 സമഷ്ടിയുടെ തുക ആകലനം ചെയ്യൽ

ഇതുവരെ നാം പരിഗണിച്ചതു് സമഷ്ടിയുടെ സമാന്തരമാധ്യം ആകലനം ചെയ്യാനുള്ള മാർഗങ്ങളാണ്. പക്ഷേ ചിലപ്പോഴെല്ലാം ആകലനം ചെയ്യേണ്ടി വരുന്നതു് സമഷ്ടിയുടെ തുകയാവും. ഉദാഹരണമായി ഒരു റബ്ബർതോട്ടത്തിൽ നിന്നു് ആണ്ടിൽ എത്ര കിലോഗ്രാം റബ്ബർ കിട്ടാൻ ഇടയുണ്ടെന്നു് ആകലനം ചെയ്യേണ്ടി വരാം. വ്യക്തികൾ റബ്ബർ മരങ്ങളാണ്. അവ ഓരോന്നിൽ നിന്നു ആണ്ടിൽ ആകെ കിട്ടുന്ന റബ്ബറിന്റെ തുകയാണു് പരിഗണിക്കപ്പെടുന്ന അഭിലക്ഷണം. ഒരു റബ്ബർമരത്തിന്റെ ശരാശരി ഉല്പാദനം കാണാനുള്ള മാർഗങ്ങളാണ് നാം ഇതു വരെ പരിഗണിച്ചതു്. ഇവിടെ ആവശ്യം ആകെ ഉല്പാദനത്തിന്റെ ആകലമാണ്. ഇതുപോലെയുള്ള സന്ദർഭങ്ങൾ പലപ്പോഴും ഉണ്ടാവാറുണ്ട്.

സമാന്തരമാധ്യത്തിന്റെ ആകലത്തിൽ നിന്നു് തുകയുടെ ആകലം എളുപ്പത്തിൽ കണ്ടുപിടിക്കാം. മുൻ ഉപയോഗിച്ച അങ്കനരീതിയനുസരിച്ച് $N\bar{X}$ ആണു് സമഷ്ടിയുടെ തുക. സാമ്പിളിന്റെ സമാന്തരമാധ്യമായ $\bar{x}$ നെ N കൊണ്ട് ഗുണിച്ചാൽ $N\bar{X}$ന്റെ ഒരു അനഭിനതആകലം ലഭിക്കുമെന്നുകാണാം. നോക്കുക: $E(N\bar{X}) = NE(\bar{X}) = N\bar{X}$ ആണല്ലൊ. അങ്ങനെ $N\bar{X}$ സമഷ്ടിയുടെ തുകയുടെ അനഭിനതആകലമാണ്.

ഈ ആകലത്തിന്റെ സാമ്പിളനപ്രസരണം, അതായതു്, $V(N\bar{x}) = N^2V(\bar{x})$ ആണല്ലൊ. പ്രതിസ്ഥാപനത്തോടുകൂടിയ ലഘുയാദൃച്ഛികസാമ്പിളനത്തിൽ ഇതിന്റെ മൂല്യം

$$N^2\left(\frac{\sigma^2}{n}\right) \text{ അല്ലെങ്കിൽ } \frac{N(N-1)S^2}{n} \text{ ഉം}$$

പ്രതിസ്ഥാപനമില്ലാത്ത ലഘുയാദൃച്ഛികസാമ്പിളനത്തിൽ

$$N^2\frac{(N-n)S^2}{Nn}, \text{ അതായതു് } \frac{N(N-n)}{n}S^2 \text{ - ഉം}$$

ആയിരിക്കുകയും ചെയ്യും. ഇവ ഉപയോഗിച്ചു് മുൻപു ചെയ്തതു പോലെ തന്നെ വിശ്വാസ്യതാന്തരാളങ്ങൾ ആകലനം ചെയ്കയും സാമ്പിൾപരിമാണം നിർണയിക്കുകയും ചെയ്യാവുന്നതാണ്.

11. അനുപാതങ്ങളുടെ ആകലനം

പലപ്പോഴും സമഷ്ടിയോടു ബന്ധപ്പെട്ട ഗുണരൂപത്തിലുള്ള ഒരു അഭിലക്ഷണത്തിലാവും അന്വേഷകന്റെ താല്പര്യം. സമഷ്ടിയിലെ ഓരോ അംഗത്തിനും ആ ഗുണം ഉണ്ടോ ഇല്ലയോ എന്നായിരിക്കും അയാൾ നിരീക്ഷിക്കുക. അന്തിമമായി ആകെയുള്ളതിൽ എന്തു് അനുപാതം അഥവാ ശതമാനം വ്യക്തികൾക്കു് ആ പ്രത്യേകത ഉണ്ടെന്നു കണ്ടുപിടിക്കുകയായിരിക്കും അന്വേഷണലക്ഷ്യവും. ഒരു രാജ്യത്തെ ആകെ ജനസംഖ്യയിൽ എന്തു് അനുപാതം അഥവാ എത്ര ശതമാനം ആളുകൾ തൊഴിലില്ലാത്തവരാണെന്നു കണ്ടുപിടിക്കുക. എന്തനുപാതം ആളുകൾ പുകവലിക്കാരാണെന്നു മനസ്സിലാക്കുക, ഒരു തോട്ടത്തിലെ എന്തനുപാതം വൃക്ഷ

ങ്ങൾ രോഗബാധിതങ്ങളാണെന്നു മനസ്സിലാക്കുക തുടങ്ങിയവ ഈ വിഭാഗത്തിൽ പെട്ട പ്രശ്നങ്ങളാണ്.

മറെറാരു തരത്തിൽ പറഞ്ഞാൽ, സമഷ്ടിയിലെ N അംഗങ്ങളിൽ M അംഗങ്ങൾക്ക് A എന്ന പ്രത്യേകതയുണ്ടെന്നും മററുള്ളവർക്ക് ആ പ്രത്യേകത ഇല്ലെന്നും വിചാരിക്കുക. അപ്പോൾ സമഷ്ടിയിലെ ആ പ്രത്യേകതയുള്ള അംഗങ്ങളുടെ അനുപാതം $P = \frac{M}{N}$ ആണ്. P യുടെ മൂല്യം ആകലനം ചെയ്യുകയാണ് അന്വേഷണത്തിൻെറ ലക്ഷ്യം. n അംഗങ്ങളുള്ള ഒരു ലഘുയാദൃച്ഛികസാമ്പിളെടുത്തെന്നും അതിൽ m അംഗങ്ങൾക്ക് ഈ പ്രത്യേകതയുണ്ടെന്നും കരുതുക. അപ്പോൾ സാമ്പിളിലെ അനുപാതം $p = \frac{m}{n}$ ആണ്. p, P യുടെ ഒരു അനഭിനത ആകലമാണെന്നു കാണിക്കാൻ കഴിയും.

യഥാർത്ഥത്തിൽ ഈ പ്രശ്നം മുമ്പ് ചർച്ചചെയ്ത പ്രശ്നങ്ങളിൽ നിന്ന് വ്യത്യസ്തമായ ഒന്നല്ല. സമഷ്ടിയിലെ അംഗങ്ങളോട് ബന്ധപ്പെടുത്തി X എന്ന ചരം നിർവചിക്കുക. i–ാമത്തെ അംഗത്തിന് നിർദിഷ്ടമായ പ്രത്യേകതയുണ്ടെങ്കിൽ അതിനോടു ബന്ധപ്പെട്ട X ൻെറ മൂല്യമായ x_i , 1 ആണെന്നും അല്ലെങ്കിൽ 0 ആണെന്നും സങ്കല്പിക്കുക. അങ്ങനെ സമഷ്ടിയുടെ മൂല്യങ്ങൾ X_1, X_2, X_N എന്നിവയാണെന്നു സങ്കല്പിച്ചാൽ അതിൽ M, 1 കളും $N-M$ പൂജ്യങ്ങളുമായിരിക്കും ഉണ്ടായിരിക്കുക. അതായത്,

$$\overline{X} = \frac{X_1 + X_2 + \cdots + X_N}{N} = \frac{M}{N} = P$$

സാമ്പിളിലെ $x_1, x_2 \cdots x_n$ എന്നീ മൂല്യങ്ങളിൽ m, 1–കളും ബാക്കി 0 ങ്ങളുമായിരിക്കും.

$$\therefore \overline{x} = \frac{x_1 + x_2 + \cdots + x_n}{n} = \frac{m}{n} = p$$

$E(\overline{x}) = \overline{X}$ എന്ന് നാം തെളിയിച്ചിട്ടുള്ളതുകൊണ്ട്, $E(p) = P$ എന്നു സിദ്ധിക്കുന്നു. അതായത് p, P യുടെ ഒരു അനഭിനത ആകലമാണ്.

അടുത്തതായി സമഷ്ടിയുടെ പ്രസരണമായ σ^2 എന്താണെന്നു പരിശോധിക്കാം.

$$s^2 = \frac{1}{N-1} \sum_{i=1}^{N} (X_i - \overline{X})^2 = \frac{1}{(N-1)} \sum_{i=1}^{N} (X_i - P)^2$$

$$= \frac{\sum_{i=1}^{N} X_i^2 - NP^2}{(N-1)} = \frac{M - NP^2}{N-1} = \frac{NP - NP^2}{N-1}$$

$$= \frac{N}{N-1} P(1-P) = \frac{N}{N-1} PQ \quad (\because Q = 1-P)$$

$$\therefore S^2 = \frac{NPQ}{N-1}.$$

സാമ്പിളിനെ സംബന്ധിച്ചാകുമ്പോൾ,

$$s^2 = \frac{1}{n-1} \sum_{i=1}^{n} (x_i - \bar{x})^2 = \frac{npq}{n-1}$$

(വ്യുൽപാദനം സമഷ്ടിയുടേതു പോലെ തന്നെ. $q = 1 - p$).

ഇവിടെ s^2, S^2 ൻെറ ഒരു അനഭിനത ആകലമാണെന്നുള്ളതു് ഓർമിക്കുക.

അടുത്തതായി നമുക്കു് കണ്ടുപിടിക്കുവാനുള്ളതു് p എന്ന സാമ്പിളജത്തിന്റെ സാമ്പിളനപ്രസരണമാണു്. $p = \bar{x}$ ആയതു കൊണ്ടു്,

$$V(p) = V(\bar{x}) = \frac{N-n}{Nn} S^2 = \frac{N-n}{N-1} \cdot \frac{PQ}{n}$$

ഇതിൻെറ അനഭിനത ആകലം

$$v(p) = \frac{N-n}{N} \cdot \frac{pq}{n-1}$$ എന്നും സിദ്ധിക്കുന്നു (S^2 നു് പകരം s^2 എഴുതുക).

നിർദിഷ്ടപ്രത്യേകതയുള്ള അംഗങ്ങളുടെ എണ്ണമാണു് ആകലനം ചെയ്യേണ്ടതെന്നിരിക്കട്ടെ. അതായതു്, $M = NP$ ആണു് ആകലനം ചെയ്യേണ്ടതു്. സ്വാഭാവികമായും Np അതിൻെറ അനഭിനതആകലമായിരിക്കും. അതിൻെറ സാമ്പിളനപ്രസരണം,

$$N^2 \left(\frac{N-n}{N-1}\right) \cdot \left(\frac{PQ}{n}\right)$$ ഉം

അതിൻെറ അനഭിനതആകലം

$$\frac{N(N-n)}{n-1} pq$$ ഉം

ആയിരിക്കുകയും ചെയ്യും.

അടുത്തതായി ചർച്ച ചെയ്യുവാനുള്ളതു് P യുടെ വിശ്വാസ്യതാന്തരാളങ്ങൾ എങ്ങനെ ആകലനം ചെയ്യാമെന്നുള്ളതാണു്. ഇവിടെയും മുൻ വിവരിച്ച രീതി തന്നെ അവലംബിക്കാം. സാമ്പിൾപരിമാണം സാമാന്യം വലുതാണെങ്കിൽ

$$\frac{p - P}{\sqrt{V(p)}},$$ മാനകനോർമൽ വിതരണത്തോടു കൂടി ആയിരിക്കും.

t എന്ന മാനകനോർമൽചരം $(-t_\alpha, t_\alpha)$ എന്ന അന്തരാളത്തിലായിരിക്കുവാനുള്ള സംഭാവ്യത നിർദിഷ്ടവിശ്വാസ്യതാഗുണാങ്കമായ α ആണെന്നിരിക്കട്ടെ.

$$P_r \left\{ -t_\alpha \leq \frac{p - P}{\sqrt{V(p)}} \leq t_\alpha \right\} = \alpha$$

അതായതു്, $P_r \left\{ p - t_\alpha \sqrt{V(p)} \leq P \leq p + t_\alpha \sqrt{V(p)} \right\} = \alpha.$

ഇവിടെ V (p) ക്കു പകരം അതിൻെറ അനഭിനത ആകലമായ v(p) എഴുതാവുന്നതാണു്. അങ്ങനെ P യുടെ, α വിശ്വാസ്യതാഗുണാങ്കത്തോടു കൂടിയ, വിശ്വാസ്യതാന്തരാളം

$$\left\{ p - t_\alpha \sqrt{\frac{N-1}{N} \cdot \frac{pq}{n-1}},\ p + t_\alpha \sqrt{\frac{N-n}{N} \cdot \frac{pq}{n-1}} \right\}$$

എന്നു സിദ്ധിക്കുന്നു.

നിർദിഷ്ടപരിധിയിൽ കവിയാത്ത പിശകു മാത്രമേ ഉണ്ടായിരിക്കു എന്നതിൻെറ സംഭാവ്യത ഒരു നിശ്ചിതസംഖ്യയായിരിക്കണമെങ്കിൽ എത്ര വലുപ്പമുള്ള സാമ്പിൾ എടുക്കണം എന്ന പ്രശ്നവും മുൻ ചെയ്തതു പോലെ തന്നെ കൈകാര്യം ചെയ്യാവുന്നതാണു്. ഉദാഹരണമായി അനുപാതത്തിൻെറ ആകലവും യഥാർത്ഥ മൂല്യവുമായുള്ള വ്യത്യാസം d യിൽ കുറവായിരിക്കുവാനുള്ള സംഭാവ്യത α യെങ്കിലുമായിരിക്കത്തക്ക വണ്ണം സാമ്പിൾ പരിമാണം നിർണയിക്കണമെന്നിരിക്കട്ടെ. അതായതു്,

$$P_r \left(| p - P | = d \right) = \alpha$$

p നോർമൽ വിതരണത്തോടുകൂടിയതാണെന്നു സങ്കല്പിച്ചാൽ,

$$P_r \left(| p - P | \leq t_\alpha \sqrt{V(p)} \right) = \alpha$$

ആയിരിക്കും. ഇവിടെ t_α യുടെ മൂല്യം പഴയതു പോലെ മാനകനോർമൽ പട്ടികയിൽ നിന്നു കണ്ടുപിടിക്കാവുന്നതാണു്.

$$\sqrt{V(p)} = \sqrt{\frac{N-n}{N-1} \cdot \frac{PQ}{n}}$$

എന്നു നാം കണ്ടു കഴിഞ്ഞു. അപ്പോൾ n ൻെറ മൂല്യം,

$$d = t_\alpha \sqrt{V(p)} = t_\alpha \sqrt{\frac{N-n}{N-1} \frac{PQ}{n}}$$

എന്ന സമീകരണത്തിൽ നിന്നു് നിർണയിക്കാം. അതായതു്,

$$n = \frac{\dfrac{t_\alpha^2 PQ}{d^2}}{1 + \dfrac{1}{N}\left(\dfrac{t_\alpha^2 PQ}{d^2} - 1\right)}.$$

N താരതമ്യേന വലുതാണെങ്കിൽ ഒരു ഏകദേശനമായി $n = \dfrac{t_\alpha^2 PQ}{d^2}$ എന്നു് എടുക്കാവുന്നതാണു്. ഇവിടെ P യുടെ മൂല്യം ലഭ്യമായ വിവരങ്ങൾ ഉപയോഗിച്ചു് ഏകദേശമായി നിർണയിക്കണം.

12. സംഗ്രഹം

1. നിർദിഷ്ട സമഷ്ടിയിൽ നിന്ന് തിരഞ്ഞെടുക്കാവുന്ന നിർദിഷ്ടവലുപ്പമുള്ള എല്ലാ സാമ്പിളുകൾക്കും തുല്യ സംഭാവ്യതയുള്ള സാമ്പിളന രീതിക്ക് ലഘുയാദൃച്ഛിക സാമ്പിൾ എന്നു പേർ പറയും. ഇതു പ്രതിസ്ഥാപനത്തോടു കൂടിയതോ പ്രതിസ്ഥാപനമില്ലാത്തതോ ആവാം. ഈ സാമ്പിളന രീതി സാമ്പിളന ശാസ്ത്രത്തിലെ അതിപ്രധാനമായ ഒന്നാണ്.

2. ലഘുയാദൃച്ഛിക സാമ്പിൾ നറുക്കെടുപ്പു മുഖേനയോ യാദൃച്ഛിക സംഖ്യാ പട്ടിക ഉപയോഗിച്ചോ തിരഞ്ഞെടുക്കാവുന്നതാണ്.

3. സമഷ്ടിയുടെ സമാന്തരമാധ്യമായ $\bar{X}$ ൻെറ ആകലമായി നിർദേശിക്കപ്പെട്ടിരിക്കുന്നതു് സാമ്പിൾ സമാന്തര മാധ്യമായ $\bar{x}$ ആണ്.

$$E(\bar{x}) = \bar{X}.$$

പ്രതിസ്ഥാപനമില്ലാത്ത ലഘു യാദൃച്ഛിക സാമ്പിളനത്തിൽ

$$V(\bar{x}) = \frac{N-n}{Nn}S^2; \quad \text{ഇവിടെ } S^2 = \frac{1}{N-1}\sum_{i=1}^{N}(x_i-\bar{X})^2.$$

പ്രതിസ്ഥാപനത്തോടുകൂടിയ സാമ്പിളനത്തിൽ

$$V(\bar{x}) = \frac{N-1}{nN}S^2 = \frac{\sigma^2}{n}$$

$V(\bar{x})$ ൻെറ ആകലമായ് $v(\bar{x})$, S^2 ൻെറ സ്ഥാനത്തു്

$$s^2 = \frac{1}{n-1}\sum_{i=1}^{n}(x_i-\bar{x})^2$$ പ്രതിസ്ഥാപിച്ചാൽ ലഭിക്കും.

4. സാമ്പിൾ പരിമാണം വലുതാണെങ്കിൽ,

$$t = \frac{\bar{x}-\bar{X}}{S_{\bar{x}}}$$ എന്ന സാമ്പിളജം മാനകനോർമൽ വിതരണം അനുവർത്തിക്കും. ഈ തത്വം ഉപയോഗപ്പെടുത്തി $\bar{X}$ ന്റെ നിർദിഷ്ടമായ വിശ്വാസ്യതാ ഗുണാങ്കത്തോടു കൂടിയ വിശ്വാസ്യതാന്തരാളം നിർണയിക്കാവുന്നതാണ്. മാനകനോർമൽചരമായ t $(-k, k)$ എന്ന അന്തരാളത്തിലായിരിക്കുവാനുള്ള സംഭാവ്യത, a എന്ന നിർദിഷ്ട സംഖ്യയായിരിക്കത്തക്കവണ്ണം k യുടെ മൂല്യം നോർമൽ പട്ടിക ഉപയോഗിച്ചു് നിർണയിച്ചാൽ, a വിശ്വാസ്യതാ ഗുണാങ്കമായിട്ടുള്ള $\bar{x}$ ൻെറ വിശ്വാസ്യതാന്തരാളം

$(\bar{x} - kS_{\bar{x}}, \bar{x} + kS_{\bar{x}})$ എന്നതാണ്. ഇവിടെ $S_{\bar{x}}$ ൽ S ൻെറ മൂല്യം അറിഞ്ഞുകൂടെങ്കിൽ അതിൻെറ ആകലം പ്രതിസ്ഥാപിക്കണം.

സാമ്പിൾ പരിമാണം ചെറുതാണെങ്കിൽ S ൻെറ മൂല്യം തരാത്തപ്പോൾ മാനകനോർമൽ വിതരണമല്ല, സ്റ്റുഡൻറ് t വിതരണമാണ് t അനുവർത്തിക്കുക. അതിൻെറ സ്വതന്ത്രതാസംഖ്യ $n-1$ ആയിരിക്കും. സ്റ്റുഡൻറ് t വിതരണ

ത്തിന്റെ (അനുബന്ധം-3) പട്ടിക ഉപയോഗിച്ച് $\overline{X}$ ന്റെ വിശ്വാസ്യതാന്തരാളം നിർണയിക്കാം.

5. $$n = \frac{\left(\frac{t_\beta S}{d}\right)^2}{1+\frac{1}{N}\left(\frac{t_\beta S}{d}\right)^2}$$

എന്ന വാക്യമുപയോഗിച്ച് $\bar{x}$, $\overline{X}$ ഇവ തമ്മിലുള്ള വ്യത്യാസം d എന്ന നിർദിഷ്ടസംഖ്യയിൽ കവിയാതിരിക്കത്തക്കവണ്ണമുള്ള സാമ്പിൾ പരിമാണം നിർണയിക്കാം. ഇവിടെ, t_β നിർണയിക്കുന്നത്,

$$P_r\left\{\left|\frac{x-\overline{X}}{S_{\bar{x}}}\right| \geqslant t_\beta\right\} = \beta$$ ആയിരിക്കത്തക്കവണ്ണമാണ്.

മാനകനോർമൽ പട്ടിക ഇതിനായി ഉപയോഗിക്കാം.

6. സമഷ്ടിയുടെ തുകയായ X ആണ് ആകലനം ചെയ്യേണ്ടതെങ്കിൽ ആകലമായി $N\bar{x}$ ഉപയോഗിക്കാം.

$$V(N\bar{x}) = N^2\, V(\bar{x}) = N^2 \frac{\sigma^2}{n}$$

$$= \frac{N(N-1)}{n} S^2 \quad \text{(പ്രതിസ്ഥാപനമുണ്ടെങ്കിൽ)}$$

$$= \frac{N(N-n)}{n} S^2 \quad \text{(പ്രതിസ്ഥാപനമില്ലെങ്കിൽ)}$$

7. P എന്ന അനുപാതമാണ് ആകലനം ചെയ്യേണ്ടതെങ്കിൽ അതിന് സാമ്പിൾ അനുപാതമായ p ഉപയോഗിക്കാം. p, P യുടെ അനഭിനത ആകലമാണ്.

$$S^2 = \frac{NPQ}{N-1}, \quad s^2 = \frac{npq}{n-1}$$

$$V(p) = \frac{N-n}{N-1} \cdot \frac{PQ}{n}$$

$$v(p) = \frac{N-n}{N} \cdot \frac{pq}{n-1}$$

8. നിർദിഷ്ട പ്രത്യേകതയുള്ള അംഗങ്ങളുടെ എണ്ണമാണ് ആകലനം ചെയ്യേണ്ടതെങ്കിൽ, അതിന് NP ഉപയോഗിക്കാം. അതിന്റെ പ്രസരണം

$\frac{N-n}{N-1} \cdot \frac{N^2PQ}{n}$ ആണ്. അതിന്റെ ആകലം

$\frac{N(N-n)}{n-1} \cdot pq$ അത്രെ.

9. സാമ്പിൾ പരിമാണം വലുതാണെങ്കിൽ $\frac{p-P}{\sqrt{V(p)}}$ എന്ന യാദൃച്ഛികചരം മാനക നോർമൽ വിതരണം അനുവർത്തിക്കും. അതുപയോഗപ്പെടുത്തി P യുടെ വിശ്വാസ്യതാന്തരാളം നിർണയിക്കാം.

$$\left\{ p - t_\alpha \sqrt{\frac{N-n}{N} \cdot \frac{pq}{n-1}},\ p + t_\alpha \sqrt{\frac{N-n}{N} \cdot \frac{pq}{n-1}} \right\}$$

എന്നതാണ് ആ അന്തരാളം. സാമ്പിൾ പരിമാണം ചെറുതാണെങ്കിൽ ഇതിനായി സ്റ്റുഡൻറ് t വിതരണം ഉപയോഗപ്പെടുത്തണം.

13. ചില ഉദാഹരണങ്ങൾ

ഉദാ: 3.1. കേരളത്തിലെ 12 ഗ്രാമങ്ങളിലെ കറവപ്പശുക്കളുടെ എണ്ണം താഴെ കൊടുക്കുന്നു. ഈ സമഷ്ടിയിൽ നിന്നു് പ്രതിസ്ഥാപനമില്ലാത്ത ലഘുയാദൃച്ഛിക സാമ്പിളന രീതിയുപയോഗിച്ച് കറവപ്പശുക്കളുടെ ഒരു സാമ്പിൾ എടുക്കുക.

ഗ്രാമം	കറവപ്പശുക്കളുടെ എണ്ണം
1	62
2	54
3	20
4	8
5	12
6	10
7	3
8	9
9	23
10	30
11	32
12	4
	267

ഇവിടെ നമുക്കു ലഭിച്ചിരിക്കുന്ന ഫ്രേം സമഷ്ടിയുടെ ലിസ്റ്റല്ല. 267 പശുക്കളാണു് സമഷ്ടിയിലുള്ളതു് എന്നു നമുക്കറിയാം. അതിൽ നിന്നു് 8 അംഗങ്ങളുള്ള ഒരു സാമ്പിളാണു് എടുക്കേണ്ടതു്. സൗകര്യത്തിനു വേണ്ടി ഈ 267 പശുക്കൾക്കു് 1 മുതൽ 267 വരെ ക്രമസംഖ്യകൾ നൽകി എന്നു വിചാരിക്കുക. ആദ്യത്തെ ഗ്രാമത്തിലെ പശുക്കൾക്കു് 1 മുതൽ 62 വരെയും രണ്ടാമത്തെ ഗ്രാമത്തിലെ പശുക്കൾക്കു് 63 മുതൽ 116 വരെയും, അങ്ങനെ തുടർച്ചയായി ക്രമസംഖ്യകൾ നൽകുക. ഓരോ ഗ്രാമത്തിലും ഉള്ള പശുക്കളുടെ ക്രമം അതിലെ ഉടമസ്ഥന്മാരുടെ വീട്ടു നംബറിൻെറ ക്രമമനുസരിച്ചോ അതു പോലെയുള്ള മറേറതെങ്കിലും പരിഗണനകൾ വെച്ചോ എടുക്കേണ്ടതാണു്. ആ ക്രമം എങ്ങനെയാണു് എടുക്കുന്നതെന്നു് മുൻകൂട്ടി തീരുമാനിക്കുകയും വേണം. താഴെ കൊടുത്തിരിക്കുന്ന പട്ടികയിൽ ഗ്രാമത്തിൻെറ നംബറും അവിടെയുള്ള ഒടുവിലത്തെ പശുവിൻെറ നംബറും കൊടുത്തിരിക്കുന്നു.

പട്ടിക 3.7

ഗ്രാമത്തിൻെറ നംബർ	ഒട്ടവിലത്തെ പശുവിൻെറ നംബർ
1	62
2	116
3	136
4	144
5	156
6	166
7	169
8	178
9	201
10	231
11	263
12	267

അടുത്തതായി യാദൃച്ഛികസംഖ്യാപട്ടികയിലെ ഒരു പേജെടുത്തു് മൂന്നക്കമുള്ള യാദൃച്ഛികസംഖ്യകൾ തുടർച്ചയായി വായിക്കുക. അവയിൽ പൂജ്യവും ആവർത്തിച്ചു വരുന്ന സംഖ്യകളും 801 നെക്കാൾ വലിയ സംഖ്യകളും വിട്ടുകളയുക. (999 നെക്കാൾ ചെറിയ 267 ൻെറ ഏറ്റവും വലിയ ഗുണിതം 801 ആണല്ലൊ). 267 നെക്കാൾ വലിയ സംഖ്യകളെ 267 കൊണ്ടു് ഹരിച്ചു വരുന്ന ശിഷ്ടം അതിനു പകരമായി എടുക്കുക. ഇങ്ങനെയുള്ള 8 സംഖ്യകൾ കിട്ടുന്നതു വരെ തുടരുക. കിട്ടിയ സംഖ്യകൾ 247, 129, 139, 178, 210, 68, 120, 46 എന്നിവയാണെന്നിരിക്കട്ടെ. മുൻ നിയമമനുസരിച്ചു് ഈ സംഖ്യകൾ ക്രമസംഖ്യകളായിവരുന്ന പശുക്കളാണു് സാമ്പിളിലെ അംഗങ്ങൾ. പട്ടിക 3.7-ൽ നിന്നു് ആദ്യത്തെ അംഗം 11-ാം ഗ്രാമത്തിലെ 16-ാമത്തെ പശുവാണെന്നു കാണാം. ഇതു പോലെ മററു പശുക്കളുടെ ഗ്രാമങ്ങളും അതാതു ഗ്രാമത്തിലെ പശുവിൻെറ ക്രമസംഖ്യയും കണ്ടുപിടിക്കാവുന്നതാണു്.

ഉദാ: 3.2.

നാലംഗങ്ങളുള്ള ഒരു സമഷ്ടിയിലെ X ൻെറ മൂല്യങ്ങൾ 7, 2, 1, 10 എന്നിവയാണു്. ഇതിൽ നിന്നു് എടുക്കാവുന്ന രണ്ടു് അംഗങ്ങൾ വീതമുള്ള പ്രതിസ്ഥാപനമില്ലാത്ത എല്ലാ ലഘുയാദൃച്ഛികസാമ്പിളുകളും പരിഗണിച്ചു് $\bar{x}$, $\bar{X}$ ൻെറ ഒരു അനഭിനത ആകലമാണെന്നും $S_{\bar{x}}^2$ ൻെറ മൂല്യം മുൻപു വ്യുൽപാദിപ്പിച്ച സൂത്രം ഉപയോഗിച്ചു് ഗണിച്ചെടുക്കുന്നതു തന്നെയാണെന്നും കാണിക്കുക.

ഇവിടെ സമഷ്ടിയുടെ സമാന്തരമാധ്യമായ $\bar{X} = \frac{7+2+1+10}{4} = 5.$ അതുപോലെ തന്നെ $\sigma^2 = \frac{1}{4}\left[(7-5)^2 + (2-5)^2 + (1-5)^2 + (10-5)^2\right]$

$$= \frac{54}{4} = 13.5$$

ഇതിൽ നിന്നു് എടുക്കാവുന്ന രണ്ടു് അംഗങ്ങളുള്ള സാമ്പിളുകൾ, അവയുടെ സമാന്തരമാധ്യം എന്നിവ പട്ടികയായി താഴെ കൊടുക്കുന്നു.

സാമ്പിളുകൾ	സമാന്തരമാധ്യം ($\bar{x}$)
7,2	4.5
7,1	4.0
7,10	8 5
2,1	1.5
2,10	6.0
1,10	5.5

സമാന്തരമാധ്യങ്ങളുടെ സമാന്തരമാധ്യം $= \dfrac{30.0}{6} = 5$

ഇതിൽ നിന്നു് $\bar{x}$, $\bar{X}$ ൻെറ അനഭിനത ആകലമാണെന്നു സിദ്ധിക്കുന്നു.

$S_{\bar{x}}^2$ = സമാന്തരമാധ്യങ്ങളുടെ പ്രസരണം

$$= \frac{1}{6}\Big[(4.5-5)^2 + (4.0-5)^2 + (8.5-5)^2 + (1.5-5)^2 + (6.0-5)^2 + (5.5-5)^2\Big]$$

$$= \frac{27}{6} = 4.5$$

$$S_{\bar{x}}^2 = \frac{N-n}{n(N-1)}\sigma^2 = \frac{4-2}{2(4-1)} \times 13.5 = 4.5$$

ഇതിൽ നിന്നു് $S_{\bar{x}}^2$ ൻെറ മൂല്യം സൂത്രമനുസരിച്ചു് ഗണിക്കുമ്പോൾ കിട്ടുന്നതു തന്നെയാണു് നേരിട്ടുള്ള ഗണനത്തിലും സിദ്ധിക്കുന്നതു്.

ഉദാ: 3.3.

ഉദാ: 3.1 ലെ പട്ടികയിൽ നിന്നു് മൂന്നു ഗ്രാമങ്ങളുടെ ഒരു സാമ്പിൾ എടുത്തു് അതിൽ നിന്നു് ഒരു ഗ്രാമത്തിലെ ശരാശരി കറവപ്പശുക്കളുടെ എണ്ണവും അതിൻെറ 0.95 വിശ്വാസ്യതാഗുണാങ്കത്തോടു കൂടിയ വിശ്വാസ്യതാന്തരാളവും ആകലനം ചെയ്യുക. (സമഷ്ടിയുടെ പ്രസരണം അറിഞ്ഞു കൂടെന്നു സങ്കൽപ്പിക്കുക.)

ആദ്യമായി രണ്ടു് അക്കമുള്ള യാദൃച്ഛികസംഖ്യകൾ തുടർച്ചയായി വായിക്കുക. അവയിൽ 0, 96 നെക്കാൾ വലിയ സംഖ്യകൾ, ആവർത്തിച്ചുവരുന്ന സംഖ്യകൾ എന്നിവ അവഗണിക്കുക. പന്ത്രണ്ടിൽ കൂടുതലുള്ളവയെ പന്ത്രണ്ടു കൊണ്ടു് ഹരിച്ചു് ശിഷ്ടം സ്വീകരിക്കുക. അങ്ങനെ കിട്ടിയ മൂന്നു സംഖ്യകൾ 1 6, 10 എന്നിവയാണെന്നിരിക്കട്ടെ. ഈ സംഖ്യകളോടു കൂടിയ ഗ്രാമങ്ങളാണു് സാമ്പിളിൽ ഉൾപ്പെടുന്നതു്. ഓരോ ഗ്രാമത്തെയും പററി നാം ശേഖരിക്കുന്ന വിവരം അവിടെയുള്ള കറവപ്പശുക്കളുടെ എണ്ണമാണു്. അങ്ങിനെ സാമ്പിൾമൂല്യങ്ങൾ 62, 10, 30 എന്നിവയാണു്.

$$\therefore \quad \bar{x} = \frac{62+10+30}{3} = 34$$

$\therefore$ ഒരു ഗ്രാമത്തിലെ ശരാശരി കറവപ്പശുക്കളുടെ എണ്ണത്തിൻെറ ആകലം $= 34$

$$s^2 = \frac{1}{2}\left\{(62-34)^2 + (10-34)^2 + (30-34)^2\right\} = 688$$

$$s = 26.23$$

സമഷ്ടിയുടെ പ്രസരണം അറിഞ്ഞുകൂടാത്തതു കൊണ്ട് സ്റ്റുഡന്റ് വിതരണം ഉപയോഗപ്പെടുത്തിയാണ് വിശ്വാസ്യതാന്തരാളങ്ങൾ നിർണയിക്കേണ്ടതു്. സ്വതന്ത്രതാസംഖ്യയായ t വിതരണത്തിൻെറ സമാകലപ്പട്ടികയിൽ നിന്നു്

$t_{.95} = 4.303$ എന്നു കിട്ടുന്നു.

അങ്ങനെ ആവശ്യമായ വിശ്വാസ്യതാന്തരാളം

$$\left\{34-4.303 \times \sqrt{\frac{12-3}{3\times 12}} \times 26.23,\ 34+4.33 \times \sqrt{\frac{12-3}{3\times 12}} \times 26.23\right\}$$

അതായതു് $[-22.4, 90.4]$ അല്ലെങ്കിൽ $[0,91]$ ആണെന്നു സിദ്ധിക്കുന്നു.

[എണ്ണമായതു കൊണ്ടു് അന്തരാളത്തിൻെറ അന്തങ്ങൾ ധനപൂർണ സംഖ്യകളായിരിക്കണം.]

ഉദാ: 3.4

ഒരു ഫാക്ടറിയിൽ 450 തൊഴിലാളികളുണ്ടു്. അവരുടെ കൂലികളുടെ പ്രസരണം 9 രൂപയാണെന്നു മാത്രമേ മുൻ അന്വേഷണങ്ങളിൽ നിന്നു് മനസ്സിലാക്കാൻ കഴിഞ്ഞിട്ടുള്ളൂ. എത്ര വലുപ്പമുള്ള ഒരു സാമ്പിളെടുത്തു് ശരാശരി കൂലി ആകലനം ചെയ്താലാണു് ശരാശരി കൂലിയുടെ യഥാർത്ഥമൂല്യവുമായി 0.5 രൂപയിലധികം വ്യത്യാസമില്ലാത്ത ഒരു ആകലം കിട്ടുമെന്നു് 0.95 സംഭാവ്യതയോടു കൂടി പ്രതീക്ഷിക്കാൻ കഴിയുക?

കൂലികളുടെ സമഷ്ടി നോർമൽ വിതരണത്തെ ഏകദേശമായി അനുസരിക്കുന്നു എന്നു സങ്കല്പിച്ചാൽ,

$$P_r\left\{\frac{|\bar{x}-\bar{X}|}{S_{\bar{x}}} \leqslant 1.96\right\} = .95$$

അതായതു്, $P_r\left\{|\bar{x}-\bar{X}| \leqslant 1.96\ S_{\bar{x}}\right\} = .95$

നമുക്കു് ആവശ്യമുള്ളതു്, $P_r\,[\,|\bar{x}-\bar{X} \leqslant |\ 0.5] = .95$ ആയിരിക്കത്തക്കവണ്ണമുള്ള n ൻെറ (സാമ്പിൾ പരിമാണത്തിൻെറ) മൂല്യമാണു്.

$$\therefore\ 1.96\ S_{\bar{x}} = 0.5$$

$$1.96 \times \sqrt{\frac{(N-n)}{Nn}}.\ S = 0.5$$

$$\therefore\ 1.96 \times \sqrt{\frac{450-n}{450\ n}}.\ 3 = 0.5$$

$$n = \frac{\left(\frac{1.96 \times 3}{0\ 5}\right)^2}{1 + \frac{1}{450}\left(\frac{1.96 \times 3}{0.5}\right)^2} \approx 106$$

കുറിപ്പ്: $S^2 = \frac{N}{N-1}.\sigma^2 = \frac{450}{449} \times 9 \approx 9;\ S \approx 3.$

ഉദാ: 3.5.

14848 വീടുകളുള്ള ഒരു നഗരത്തിൽ നിന്നു് 30 വീടുകളുടെ ഒരു സാമ്പിൾ എടുത്തു. ഓരോ വീട്ടിലും ഉണ്ടായിരുന്ന ആളുകളുടെ എണ്ണം താഴെ കൊടുക്കുന്നു. 5, 6, 3, 3, 2, 3, 3, 3, 4, 4, 3, 2, 7, 4, 3, 5, 4, 4, 3, 3, 4, 3, 3, 1, 2, 4, 3, 4, 2, 4. ആ നഗരത്തിലെ ആകെ ജനസംഖ്യ ആകലനം ചെയ്യുക. ജനസംഖ്യയുടെ 90% വിശ്വാസ്യതാഗുണാങ്കത്തോടു കൂടിയ വിശ്വാസ്യതാന്തരാളവും ആകലനം ചെയ്യുക.

സാമ്പിൾ സമാന്തരമാധ്യം, $\bar{x} = \frac{104}{30}$

ഇതു് ഒരു വീട്ടിൽ ശരാശരിയുള്ള ആളുകളുടെ ആകലമാണു്. ആകെ വീടുകളുടെ എണ്ണം കൊണ്ടു് ഇതിനെ ഗുണിച്ചാൽ ജനസംഖ്യയുടെ ആകലം കിട്ടും.

ജനസംഖ്യ = സമഷ്ടിയുടെ തുക $N\bar{x}$

$$= 14848 \times \frac{104}{30} = \underline{\underline{51473}}$$

വിശ്വാസ്യതാന്തരാളം നിർണയിക്കാൻ ഇതിൻെറ മാനകപ്പിശകു ആകലനം ചെയ്യണം. അതിനു് s ആവശ്യമുണ്ടു്.

$$s^2 = \frac{1}{29}\sum_{i=1}^{30}(x_i - \bar{x})^2 = \frac{1}{29}\sum_{i=1}^{30} x_i^2 - \frac{30}{29}\bar{x}^2$$

$$= \frac{1}{29}\left[404 - 360.53\right] = 1.5$$

$$s = 1.225$$

$$\text{മാനകപ്പിശക്} = \sqrt{v(N\bar{x})} = \sqrt{\frac{N-n}{N} \cdot \frac{NS}{\sqrt{n}}}$$

$$= \sqrt{14818/(14848 \times 30)} \times 14848 \times 1.225$$

$$= 3334\ 36$$

$$t_{.90} = 1.64 \quad (\text{നോർമൽ പട്ടികയിൽ നിന്നു്})$$

$$\therefore \quad t_{.90} \sqrt{v/(N\bar{x})} = 3334\ 36 \times 1.64 \approx 5468$$

$$90\% \text{ വിശ്വാസ്യതാന്തരാളം} = [\,51473 - 5468,\ 51473 + 5468\,]$$

$$= \underline{\underline{[\,46005,\ 56941\,]}}$$

ഉദാ: 3.6.

ഒരു ജില്ലയിൽ 5000 കിണറുകൾ കുഴിക്കാനും അവ കെട്ടി ശരിപ്പെടുത്താനുമുള്ള സാധനങ്ങൾ ഗവണ്മെൻറിൽ നിന്നു് കർഷകരുടെ ഇടയിൽ വിതരണം ചെയ്തു. കിണറിനായി സാധനങ്ങൾ ഏററുവാങ്ങിയ കർഷകരുടെ ലിസ്റ്റും കിണറു കുഴിക്കാൻ നിർദേശിക്കപ്പെട്ട സ്ഥാനങ്ങളുടെ ലിസ്റ്റും ലഭ്യമാണു്. ഏതാണ്ടു് പകുതിയോളമാളുകൾ കിട്ടിയ സാധനങ്ങൾ മററാവശ്യങ്ങൾക്കായി ഉപയോഗിച്ചെന്നു് സംശയിക്കപ്പെടുന്നു. എത്ര അനുപാതം കിണറുകൾ യഥാർത്ഥത്തിൽ നിർമിച്ചിട്ടുണ്ടെന്നു് ലഘുയാദൃച്ഛികസാമ്പിളന രീതിയിൽ പരിശോധിക്കണം. 95% വിശ്വാസത്തോടു കൂടി, 10% ത്തിലധികം പിശകില്ലാത്ത ആകലമാണു് ലഭിച്ചിരിക്കുന്നതെന്നു പറയാൻ കഴിയണമെങ്കിൽ എന്തു വലുപ്പമുള്ള സാമ്പിൾ എടുക്കണം?

സമഷ്ടിയിലെ അനുപാതം സംശയിക്കപ്പെടുന്ന 0.5 ആയി എടുക്കുക. അതിന്റെ പത്തു ശതമാനമാണല്ലോ അനുവദിക്കാവുന്ന പിശകു്. അതായതു്, $d = 0.10 \times .5 = 0.05;\ N = 5000,\ \alpha = 0.95,\ P = 5,\ Q = .5.$

നോർമൽ പട്ടികയിൽ നിന്നു് $t_\alpha = 1.96$. സാമ്പിൾ പരിമാണം നിർണയിക്കേണ്ടതു്,

$P_r\{|p-P| \leqslant d\} = \alpha$ ആയിരിക്കത്തക്ക വണ്ണമാണു്. $\frac{p-P}{V(p)}$ നോർമൽ വിതരണം അനുസരിക്കുന്നതായി സങ്കല്പിക്കുന്നു. $V(p) = \frac{N-n}{N-1} \cdot \frac{PQ}{n}$

P ക്കു് നൽകപ്പെട്ടിരിക്കുന്ന മൂല്യം 0.5 ആയതു കൊണ്ടു്

$$V(p) = \frac{5000 - n}{4999} \times \frac{.5 \times .5}{n}$$

$t_\alpha = d \sqrt{V(p)}$ ആയിരിക്കണമല്ലോ. അതായതു്,

$$1.96 = 0.05 \sqrt{\frac{5000 - n}{4999} \times \frac{.5 \times .5}{n}}$$

ഇതിൽ നിന്നു് $n = \underline{\underline{357.}}$

ചില സഹായകഗ്രന്ഥങ്ങളും ലേഖനങ്ങളും:

1. Sampling Techniques (Ch. II and III), W. G. Cochran, John Wiley & Sons.
2. Sampling Theory and Methods -- Chapter III --, M. N. Murthy, Statistical Publishing Society.
3. Sample Survey Methods, M. H. Hansen, W. H. Hurwitz and W. G. Madow, John Wiley & Sons.
4. Variance and Confidence Interval Estimations (1962), M. N. Murthy, Sankhya 24, 1—12
5. On samples from finite populations (1944), Cornfield. J, J. Amer, Stat. Assn. 39, 236—239.

അഭ്യാസം 3.

1. U_1, U_2, U_3, U_4, U_5 എന്നീ അംഗങ്ങളോടുകൂടിയ ഒരു സമഷ്ടിയിൽ നിന്നു് 3 അംഗങ്ങൾ വീതമുള്ള എല്ലാ സാമ്പിളുകളും എടുത്തു് സാമ്പിൾ സമാന്തരമാധ്യത്തിന്റെ പ്രതീക്ഷ, സമഷ്ടിയുടെ സമാന്തരമാധ്യമാണെന്നു് കാണിക്കുക.

2. 300 വീടുകളുള്ള ഒരു ഗ്രാമത്തിൽ നിന്നു് 10 വീടുകളുടെ ഒരു ലഘു യാദൃച്ഛിക സാമ്പിൾ പ്രതിസ്ഥാപനം കൂടാതെ എടുക്കുന്നു. അതിൽ 3 വീടുകളിൽ റെഫ്രിജറേറ്റർ ഉണ്ടായിരുന്നു. സാമ്പിളിൽ പെട്ട വീടുകളിലെ ആളുകളുടെ എണ്ണം യഥാക്രമം, 3, 5, 4, 4, 7, 3, 1, 5, 4,3 എന്നിവയായിരുന്നു. ആ ഗ്രാമത്തിലെ ജനസംഖ്യയും, റെഫ്രിജറേറ്ററുള്ള വീടുകളുടെ അനുപാതവും ആകലനം ചെയ്യുക. അവയുടെ .99 വിശ്വാസ്യതാഗുണാങ്കത്തോടുകൂടിയ വിശ്വാസ്യതാന്തരാളങ്ങൾ നിർണയിക്കുകയും ചെയ്യുക.

3. ഒരു ഗവേഷകൻ ഒരു ഗ്രാമത്തിലെ ക്ഷയരോഗബാധിതരുടെ അനുപാതം നിർണയിക്കാൻ തല്പരനാണു്. അനുപാതത്തിന്റെ യഥാർഥ മൂല്യവുമായി 5% ത്തിലധികം വ്യത്യാസമില്ലാത്ത ആകലം 95% വിശ്വാസ്യതയോടുകൂടി ലഭിക്കാനാണു് അയാൾ ശ്രമിക്കുന്നതു്. എത്ര വലുപ്പമുള്ള സാമ്പിൾ എടുക്കണം? (30% ആളുകൾ ക്ഷയരോഗബാധിതരാണെന്നു സംശയിക്കപ്പെടുന്നു എന്നു സങ്കല്പിക്കുക).

4. ഒരു ഹർജിയിൽ 676 പേജുകളിലായി ഒപ്പുകൾ ശേഖരിച്ചിരിക്കുന്നു. ഒരു പേജിൽ 42 ഒപ്പുകൾക്കുള്ള സ്ഥലമാണുള്ളതു്. പക്ഷേ എല്ലാ പേജിലും 42 ഒപ്പുകൾ വീതമില്ല. ലഘുയാദൃച്ഛിക സാമ്പിളന രീതിയനുസരിച്ചു് തിരഞ്ഞെടുത്ത 50 പേജുകളിൽ നിന്നു കിട്ടിയ വിവരങ്ങൾ താഴെ കൊടുക്കുന്നു.

ഒപ്പുകളുടെ എണ്ണം	42	41	36	32	29	27	23	19	16	15
അത്രയും ഒപ്പുകളുള്ള പേജുകൾ	23	4	1	1	1	2	1	1	2	2
	14	11	10	9	7	6	5	4	3	
	1	1	1	1	1	3	2	1	1	

ആകെ ഒപ്പുകളുടെ എണ്ണവും അതിന്റെ 80 % വിശ്വാസ്യതാന്തരാളവും ആകലനം ചെയ്യുക.

5. 2000 കോളേജുകളിൽ നിന്നു് 200 അംഗങ്ങളുള്ള ഒരു ലഘു യാദൃച്ഛിക സാമ്പിളെടുത്തു. അതിൽ 120 എണ്ണം അധ്യയനമാധ്യമം പ്രദേശിക ഭാഷയിലേക്കു് മാററുന്നതിനു് അനുകൂലമാണു്. ആകെ കോളേജുകളിൽ എന്തനുപാതം ഈ നിർദേശത്തിനു് അനുകൂലമാണു്? അതിന്റെ 90% വിശ്വാസ്യതാന്തരാളം നിർണയിക്കുക.

4

സ്തരിത യാദൃച്ഛിക സാമ്പിളനം

1. സ്തരണം എന്ത്, എന്തിന് ?

ലഘുയാദൃച്ഛിക സാമ്പിളനത്തോട് ബന്ധപ്പെട്ട ചർച്ചകളിൽ നിന്ന് ഒരു കാര്യം വ്യക്തമായി. ആകലങ്ങളുടെ പ്രസരണം സാമ്പിൾ പരിമാണത്തെയും സമഷ്ടിയുടെ പ്രസരണത്തെയും ആശ്രയിച്ചാണിരിക്കുന്നത്. ആകലങ്ങളുടെ സൂക്ഷ്മത അവയുടെ പ്രസരണം കുറയുന്നതനുസരിച്ച് വർധിക്കുമെന്ന് നേരത്തെ തന്നെ നാം കണ്ടതാണ്. അതുകൊണ്ട് ലഘു യാദൃച്ഛിക സാമ്പിളനം അവലംബിക്കുമ്പോൾ ആകലങ്ങളുടെ സൂക്ഷ്മത വർധിപ്പിക്കുവാനുള്ള ഒരു മാർഗം സാമ്പിൾ പരിമാണം വർധിപ്പിക്കുക മാത്രമാണ്. പക്ഷേ സാമ്പിൾ പരിമാണത്തിലുണ്ടാകുന്ന വർധനവ് സാമ്പിളനത്തിന് ആവശ്യമായി വരുന്ന പണച്ചെലവും വർധിപ്പിക്കും. സാമ്പത്തിക പരാധീനത മൂലം സാമ്പിൾ പരിമാണത്തിന് ഒരു പരിമിതി കല്പിക്കേണ്ടി വരുമെന്നു സാരം. അങ്ങനെ വരുമ്പോൾ സമഷ്ടിയുടെ പ്രസരണം കുറഞ്ഞിരുന്നാലേ ആകലങ്ങളുടെ പ്രസരണം കുറയുകയും, സൂക്ഷ്മത വർധിക്കുകയും ചെയ്യുകയുള്ളൂ. ഇങ്ങനെയുള്ള സന്ദർഭങ്ങളിൽ കൂടുതൽ സൂക്ഷ്മതയുള്ള ആകലങ്ങൾ ലഭിക്കുവാൻ സഹായിക്കുന്ന ഒരു സാമ്പിളന രീതിയാണ് സ്തരിത സാമ്പിളനം.

മിക്കവാറും എല്ലാ സമഷ്ടിയും ഏതു അഭിലക്ഷണമെടുത്താലും, ഭിന്നാത്മകമായിരിക്കും. പലപ്പോഴും അതിനെ ഒന്നോ കൂടുതലോ അഭിലക്ഷണങ്ങളിൽ താരതമ്യേന ഏകാത്മകങ്ങളായ ഉപസമുച്ചയങ്ങളായി വിഭജിക്കാൻ കഴിഞ്ഞെന്നു വരാം. ഇപ്രകാരമുള്ള സമുച്ചയങ്ങളിൽ ആ അഭിലക്ഷണങ്ങളുടെ പ്രസരണം വളരെ ചെറുതായിരിക്കും. ആയതിനാൽ അഭിലക്ഷണങ്ങളുടെ ആകലങ്ങളുടെ സൂക്ഷ്മത കൂടുതലായിരിക്കും. അങ്ങനെയുള്ള ഓരോ ഉപസമുച്ചയത്തിനും ഓരോ സ്തരം എന്നു പേർ പറയും. സമഷ്ടിയിൽ k സ്തരങ്ങളാണ് ഉള്ളതെന്നും അവയുടെ അംഗസംഖ്യ യഥാക്രമം $N_1, N_2, \ldots N_k$ എന്നിവയാണെന്നും സങ്കല്പിക്കുക. സമഷ്ടിയുടെ അംഗസംഖ്യ N എന്നിരിക്കട്ടെ. അപ്പോൾ,

$$N = N_1 + N_2 + \cdots + N_k \tag{4.1.1}$$

n അംഗങ്ങളുള്ള സാമ്പിളാണ് എടുക്കുന്നത് എന്നും ഓരോ സ്തരത്തിൽ നിന്നു് യഥാക്രമം $n_1, r_2, \cdots n_k$ അംഗങ്ങളെ സാമ്പിളിൽ ഉൾപ്പെടുത്തുന്നു എന്നുമിരിക്കട്ടെ.

$$n = n_1 + r_2 + \cdots + r_k \tag{4.1 2}$$

ആദ്യത്തെ സ്തരത്തിൽ നിന്നു് n_1 അംഗങ്ങളെയും രണ്ടാമത്തേതിൽനിന്നു് r_2 അംഗങ്ങളെയും, അങ്ങനെ യഥാക്രമം k-യാമത്തെ സ്തരത്തിൽ നിന്നു് r_k അംഗങ്ങളെയുമാണല്ലോ സാമ്പിളനം ഉൾപ്പെടുത്തുന്നതു്. ലഘുയാദൃച്ഛിക സാമ്പിളനം അനുസരിച്ചാണു് ഓരോ സ്തരത്തിൽ നിന്നും നിർദിഷ്ട എണ്ണം അംഗങ്ങളെ തിരഞ്ഞെടുക്കുന്നതെങ്കിൽ ഈ സാമ്പിളന സമ്പ്രദായത്തിനു് 'സ്തരിത യാദൃച്ഛിക സാമ്പിളനം' എന്നു പേർ പറയും.

സാമ്പിളന സമ്പ്രദായത്തിൽ വരുത്തുന്ന ഈ പരിഷ്കാരത്തിന്റെ പ്രധാന പ്രയോജനം ആകലങ്ങളുടെ സാമ്പിളന പ്രസരണം കുറച്ചു അവയുടെ സൂക്ഷ്മത വർധിപ്പിക്കുക എന്നതാണെന്നു് അടുത്തു തന്നെ വ്യക്തമാവും. ഇതോടൊപ്പം മറ്റു പല പരിഗണനകളും സ്തരിത സാമ്പിളന രീതി അംഗീകരിക്കുവാൻ പ്രേരകമാവാറുണ്ടു്. അവയിൽ ചിലതു് താഴെ കൊടുക്കുന്നു

സമഷ്ടിയെപ്പററി പൊതുവെയുള്ള വിവരങ്ങളോടൊപ്പം അതിലെ ചില പ്രത്യേക വിഭവങ്ങളെ സംബന്ധിച്ചുള്ള വിവരങ്ങളിലും അന്വേഷകനു് താല്പര്യമുണ്ടാകാവുന്നതാണു്, ഉദാഹരണമായി പുതുതായി ആവിഷ്കരിച്ച ഒരു സാമ്പത്തിക നയത്തെപ്പററി പൊതുജനങ്ങളുടെ പ്രതികരണം എന്തെന്നറിയാനുള്ള ഒരന്വേഷണത്തിൽ പൊതുവായ പ്രതികരണത്തോടൊപ്പം ധനികരുടെയും ഇടത്തരക്കാരുടെയും ദരിദ്രരുടെയും പ്രത്യേകം പ്രത്യേകമുള്ള പ്രതികരണം കൂടി മനസ്സിലാക്കുവാൻ ഗവണ്മെന്റിനു് താല്പര്യമുണ്ടാവാം. അതു കൂടി പരിഗണിച്ചിട്ടാവാം ഗവണ്മെൻറു് സുപ്രധാനങ്ങളായ പല തീരുമാനങ്ങളും എടുക്കുവാൻ ഉദ്ദേശിക്കുന്നതു്. ഇതിനു് ധനികർ, ഇടത്തരക്കാർ, ദരിദ്രർ എന്നിങ്ങനെ ജനതയെ മൂന്നു സ്തരങ്ങളായി വിഭജിച്ചു് വിവരം ശേഖരിക്കേണ്ടതു് ആവശ്യമാണു്. അതിനാൽ ഇവിടെ സ്തരിത സാമ്പിളന രീതി അവലംബിക്കണം.

സാമ്പിളന ഫ്രേമിന്റെ പ്രത്യേകതകളാണു് പലപ്പോഴും സ്തരിതസാമ്പിൾ എടുക്കാൻ പ്രേരകമാവാറുള്ളതു്. സംസ്ഥാനത്തെ സ്കൂളുകളുടെ ഒരു സാമ്പിളാണു് എടുക്കേണ്ടതെന്നിരിക്കട്ടെ. സ്കൂളുകൾ വിവിധ വിദ്യാഭാസ ജില്ലകളിലായി ചിതറിക്കിടക്കുന്നു. ഓരോ വിദ്യാഭ്യാസജില്ലയിലുമുള്ള സ്കൂളുകളെപ്പററിയുള്ള വിവരങ്ങൾ അതാതു ജില്ലാ തലസ്ഥാനത്തായിരിക്കും സൂക്ഷിച്ചിരിക്കുന്നതു്. ഈ രേഖകളെല്ലാം ശേഖരിച്ചു് ഒന്നിച്ചു ചേർത്തു് അതിൽ നിന്നു് സാമ്പിളെടുക്കുന്നതിനേക്കാൾ സൗകര്യം ഓരോ വിദ്യാഭ്യാസ ജില്ലയേയും ഓരോ സ്തരമായി കണക്കാക്കി ഓരോന്നിൽ നിന്നും ആവശ്യമുള്ളത്ര സ്കൂളുകൾ തിരഞ്ഞെടുക്കുന്നതായിരിക്കും. അന്വേഷണ ഉദ്യോഗസ്ഥനു് അതതു ജില്ലാതലസ്ഥാനത്തുള്ള തന്റെ അനുയായികൾക്കു് ആവശ്യമായ നിർദേശം നല്കി ഇതു നിർവഹിക്കാൻ കഴിയും.

സ്തരിത സാമ്പിളെടുക്കുന്നതിനു പ്രേരകമായ മററൊരു ഘടകം നിയന്ത്രണ സൗകര്യമാണ്. ഗവൺമെൻറ് നിലവാരത്തിലാണ് വിവരം ശേഖരിക്കുന്നതെന്നിരിക്കട്ടെ. ഭരണപരമായ ആവശ്യത്തിനായി വിവിധ ജില്ലകളിലായി സംസ്ഥാനത്തെ വിഭജിച്ചിട്ടുണ്ടാവും. അവ സ്തരങ്ങളായി സ്വീകരിച്ചാൽ സാമ്പിളനം ആസൂത്രണം ചെയ്യാനും നിയന്ത്രിക്കാനും വളരെയേറെ സൗകര്യമാവും. സമഷ്ടിയിലെ ചില വിഭാഗങ്ങൾക്ക് കൂടുതൽ പ്രാധാന്യം നൽകണമെന്ന് ഉദ്ദേശമുണ്ടെങ്കിൽ അതിനും സ്തരിത സാമ്പിളന രീതി അവലംബിക്കാതെ തരമില്ല. അതു പോലെ തന്നെ ഭൂമിശാസ്ത്ര പരമായ പ്രത്യേകതകളും സ്തരിത സാമ്പിളന സമ്പ്രദായം സ്വീകരിക്കാൻ നിർബന്ധിതരാക്കാറുണ്ട്.

ഇങ്ങനെയെല്ലാമുള്ള കാരണങ്ങൾ സ്തരിത സാമ്പിളന രീതി അംഗീകരിക്കാൻ പ്രേരകമാവാമെങ്കിലും, താത്വികമായി പറഞ്ഞാൽ വിഭിന്നാത്മകമായ സമഷ്ടിയെ ഏകാത്മകങ്ങളായ സമുച്ചയങ്ങളായി തരംതിരിച്ച് ഈ സമുച്ചയങ്ങളിൽ നിന്നുള്ള ആകലങ്ങളുടെ സൂക്ഷ്മത വർധിപ്പിക്കുകയും അപ്രകാരമുള്ള സമുച്ചയങ്ങളിലെ ആകലങ്ങളുപയോഗിച്ച് സമഷ്ടി സ്വഭാവങ്ങളുടെ ആകലങ്ങൾക്ക് സൂക്ഷ്മത കൂട്ടുകയുമാണ് ഉദ്ദേശ്യം.

ഇവിടെ ഒരു കാര്യം പ്രത്യേകം ശ്രദ്ധിക്കേണ്ടതായുണ്ട്. സ്തരത്തിന് ആധാരമായി സ്വീകരിക്കാവുന്ന ഏതെങ്കിലും ചില വിവരങ്ങൾ സമഷ്ടിയെപ്പററി അറിയാമെങ്കിൽ മാത്രമെ ഈ സാമ്പിളന രീതി ഉപയോഗിക്കാൻ സാധിക്കൂ. നിർദിഷ്ട സമഷ്ടിയെപ്പററി നേരത്തെ നടത്തിയിട്ടുള്ള പഠനങ്ങളോ പ്രാരംഭമായി നാം തന്നെ നടത്തുന്ന ചില പ്രാഥമിക അന്വേഷണങ്ങളോ സ്തരണത്തിന് ആവശ്യമായ വിശദാംശങ്ങൾ നല്കിയെന്നു വരാം. അങ്ങനെ ഏതെങ്കിലും വിവരം ലഭിക്കാതെ വന്നാൽ ലഘുയാദൃച്ഛിക സാമ്പിളന രീതിയോ മററോ അവലംബിക്കുകയേ നിർവാഹമുള്ളൂ.

സ്തരിത സാമ്പിളനത്തിൽ പ്രധാനമായി പരിഗണിക്കേണ്ടി വരുന്ന പ്രശ്നങ്ങൾ താഴെ പറയുന്നവയാണ്.

1. സമഷ്ടിയെ എത്ര സ്തരങ്ങളായി വിഭജിക്കണം?

2. ഓരോ സ്തരത്തിൽ നിന്നും എത്ര അംഗങ്ങളെ വീതം സാമ്പിളിൽ ഉൾപ്പെടുത്തണം?

3. ഓരോ സ്തരത്തിൽ നിന്നും സാമ്പിളിലേക്ക് അംഗങ്ങളെ എങ്ങനെ തിരഞ്ഞെടുക്കണം?

ഈ പ്രശ്നങ്ങൾ ഒററക്കൊററക്ക് പരിഗണിക്കാവുന്നവയല്ല; പ്രത്യേകിച്ച് ഒടുവിലത്തെ രണ്ട് പ്രശ്നങ്ങൾ. സാമ്പിളിൽ എത്ര അംഗങ്ങളെ ഉൾപ്പെടുത്തണമെന്നു തീരുമാനിക്കുന്നതു്, അവയെ എങ്ങനെയാണ് തിരഞ്ഞെടുക്കാൻ പോകുന്നതു് എന്നതു കൂടി പരിഗണിച്ചാണ്. എത്ര സ്തരങ്ങളായി വിഭജിക്കണമെന്നു തീരുമാനിക്കുന്നതിലും സ്വീകരിക്കാൻ പോകുന്ന സാമ്പിളന രീതി സ്വാധീനം ചെലുത്തും. സ്തരങ്ങളിൽ നിന്ന് ലഘു യാദൃച്ഛിക സാമ്പിളന രീതിയനുസരിച്ചാ

ണ് അംഗങ്ങളെ തിരഞ്ഞെടുക്കുന്നതു് എന്ന സങ്കല്പത്തിൽ ഓരോ സ്തരത്തിൽ നിന്നും എത്ര അംഗങ്ങളെ വീതം സാമ്പിളിൽ ഉൾപ്പെടുത്തണം എന്നതിനെപ്പറ്റിയാണ് അടുത്തു വരുന്ന ഖണ്ഡികകളിൽ ചർച്ചചെയ്യുന്നതു്. മററു സാമ്പിളന രീതികൾ അവലംബിക്കുമ്പോൾ അതിനനുസരിച്ചു ഈ സംഖ്യകളിൽ മാറ്റം വരുത്തേണ്ടി വരും.

സമഷ്ടിയെപ്പററി ചില ആനുഷംഗിക വിവരങ്ങൾ അറിയേണ്ടതു് സ്തരണത്തിന് ആവശ്യമാണെന്നു് മുൻപു് സൂചിപ്പിച്ചല്ലൊ. ഇതു് പലപ്പോഴും അന്വേഷണവിധേയമായ അഭിലക്ഷണത്തെ സംബന്ധിച്ച വിവരങ്ങളായിരിക്കുകയില്ല. അതിനോടു ബന്ധപ്പെട്ടതെന്നു കരുതാവുന്ന മറേറതെങ്കിലും അഭിലക്ഷണത്തെപ്പററിയുള്ളവയായിരിക്കും. ഉദാഹരണമായി, പഞ്ചസാരയുടെ ഉപയോഗത്തെപ്പററിയുള്ള അന്വേഷണമാണ് നടത്തുന്നതെന്നിരിക്കട്ടെ. സാമ്പത്തിക സ്ഥിതിയും പഞ്ചസാരയുടെ ഉപയോഗവും ബന്ധപ്പെട്ടതാണെന്നുള്ള സങ്കല്പത്തിൽ സാമ്പത്തികാടിസ്ഥാനത്തിൽ സമഷ്ടിയെ പലതായി വിഭജിക്കാറുണ്ട്. ഇങ്ങനെ സ്തരണത്തിനായി പരിഗണിക്കുന്ന അഭിലക്ഷണത്തിനു് സംപൂരക അഭിലക്ഷണം എന്നു പേർ പറയും. ഇവിടെ രണ്ടു കാര്യങ്ങൾ പ്രത്യേകം ശ്രദ്ധിക്കേണ്ടതുണ്ട്.

1. അംഗീകരിക്കുന്ന സംപൂരക അഭിലക്ഷണം പരിഗണിക്കപ്പെടുന്ന അഭിലക്ഷണവുമായി നേരിട്ടു ബന്ധമുള്ളതായിരിക്കണം.

2. സംപൂരക അഭിലക്ഷണത്തെപ്പററിയുള്ള വിവരങ്ങൾ നേരത്തെ ശേഖരിച്ചിട്ടുള്ളതോ വിഷമം കൂടാതെ ശേഖരിക്കാവുന്നതോ ആയിരിക്കണം.

സ്തരണത്തിൻെറ വിജയം വളരെയേറെ അനുയോജ്യമായ സംപൂരക അഭിലക്ഷണം തിരഞ്ഞെടുക്കുന്നതിനോട് ബന്ധപ്പെട്ടിരിക്കുന്നു. സാമാന്യബുദ്ധിയും സാമ്പിളനത്തിലുള്ള പ്രായോഗിക പരിചയവും മാത്രമാണു് ഇക്കാര്യത്തിൽ അന്വേഷകനു് സഹായകമാകുന്നതു്.

സ്തരിത സാമ്പിളനത്തിൻെറ വൈശിഷ്ട്യം ഒരു ഉദാഹരണം കൊണ്ടു വ്യക്തമാക്കാം. പട്ടിക 3.2 ൽ കൊടുത്തിരിക്കുന്ന സമഷ്ടി തന്നെ പരിഗണിക്കുക. A, B, C, D, E, F എന്നീ തൊഴിലാളികളുടെ കൂലി യഥാക്രമം 3, 1, 4, 2, 6, 5 രൂപ വീതമാണെന്നാണല്ലോ തന്നിരിക്കുന്നതു്. ഇതിൽ നിന്നു് രണ്ടു് അംഗങ്ങൾ വീതമുള്ള 15 സാമ്പിളുകൾ എടുക്കാമെന്നും അവയുടെ സമാന്തര മാധ്യങ്ങളുടെ പ്രസരണം 1.16 ആണെന്നും (പട്ടിക 3.7) നാം കണ്ടു. ഇതിൽ A, B, D എന്നിവർ പുതുതായി ജോലിയിൽ വന്നവരാണെന്നും മററുള്ളവർ പരിചയ സമ്പന്നരായ തൊഴിലാളികളാണെന്നുമുള്ള സംപൂരക വിവരം അറിയാമെന്നിരിക്കട്ടെ. സ്വാഭാവികമായും സമഷ്ടിയെ രണ്ടു് സ്തരങ്ങളായി വിഭജിക്കാം. A, B, D ഒരു സ്തരം; C, E, F രണ്ടാമത്തെ സ്തരം. ഇതിൽ ഓരോന്നിൽ നിന്നും ഓരോ അംഗത്തെ വീതം ലഘ്വയാദൃച്ഛിക സാമ്പിളന രീതിയനുസരിച്ചു് തിരഞ്ഞെടുക്കുന്നെന്നു് സങ്കല്പിക്കുക. അങ്ങനെ എടുക്കാവുന്ന രണ്ടു് അംഗങ്ങളുള്ള എല്ലാ സാമ്പിളുകളെയും പരിഗണിക്കാം. സാമ്പിളുകളും അവയുടെ സമാന്തര മാധ്യങ്ങളും സമാന്തരമാധ്യങ്ങളുടെ പ്രസരണവും താഴെ കൊടുക്കുന്നു.

പട്ടിക 4.1

സാമ്പിൾ	സാമ്പിൾ മൂല്യങ്ങൾ	സമാന്തരമാധ്യം
A,C	3,4	3.5
A,E	3,6	4.5
A,F	3,5	4.0
B,C	1,4	2.5
B,E	1,6	3.5
B,F	1,5	3.0
D,C	2,4	3.0
D,E	2,6	4.0
D,F	2,5	3.5

സമാന്തര മാധ്യങ്ങളുടെ പ്രസരണം = 0.33. ലഘുയാദൃച്ഛിക സാമ്പിളനത്തിൽ സാമ്പിൾ സമാന്തരമാധ്യത്തിന്റെ പ്രസരണം 1.16 ആയിരുന്നു. ഇതു് സ്തരണത്തിന്റെ ഫലമായി 0.33 ആയി കുറഞ്ഞതു നോക്കുക. അതായതു് സ്തരിത സാമ്പിളനത്തിൽ സാമ്പിൾ സമാന്തരമാധ്യങ്ങൾ സമഷ്ടിയുടെ സമാന്തരമാധ്യവുമായി കൂടുതൽ അടുത്തു വരുന്നു.

2. ചില നിർവചനങ്ങളും അങ്കനരീതികളും

സമഷ്ടിയുടെ സമാന്തരമാധ്യം ആകലനം ചെയ്യുകയാണു് അന്വേഷണലക്ഷ്യമെന്നിരിക്കട്ടെ. മുമ്പ് പറഞ്ഞതു പോലെ സമഷ്ടിയിൽ N അംഗങ്ങളുണ്ടെന്നും അതിനെ യഥാക്രമം $N_1, N_2 \dots N_k$ അംഗങ്ങൾ വീതമുള്ള k സ്തരങ്ങളായി വിഭജിച്ചിരിക്കുന്നെന്നും സങ്കല്പിക്കുക. h-ാമത്തെ സ്തരത്തിലെ i-ാമതു് അംഗത്തിന്റെ അഭിലക്ഷണമൂല്യം x_{hi} കൊണ്ടു് സൂചിപ്പിക്കാം. n അംഗങ്ങളുള്ള സാമ്പിളാണു് ആവശ്യം. യഥാക്രമം, $n_1, n_2, \dots n_k$ അംഗങ്ങളെയാണു് 1 മുതൽ k വരെയുള്ള സ്തരങ്ങളിൽ നിന്നു എടുക്കുന്നതു്. സ്വാഭാവികമായും,

$$N = N_1 + N_2 + \dots + N_k$$
$$n = n_1 + n_2 + \dots + n_k$$

h-ാമതു് സ്തരത്തിൽ നിന്നു് എടുത്ത i-ാമതു് അംഗത്തിന്റെ അഭിലക്ഷണമൂല്യം, x_{hi} എന്നു സൂചിപ്പിച്ചല്ലോ. അപ്പോൾ ഒന്നാമത്തെ സ്തരത്തിൽ നിന്നുമുള്ള സാമ്പിൾമൂല്യങ്ങൾ x_{11}, x_{12}, x_{1n_1},-ഉം രണ്ടാമത്തെ സ്തരത്തിൽ നിന്നും $x_{21}, x_{22}, \dots x_{2n_2}$-ഉം അങ്ങനെ k-ാമതു് സ്തരത്തിൽ നിന്നു് $x_{k1}, X_{k2}, \dots x_{kn_k}$-ഉം സാമ്പിൾമൂല്യങ്ങൾ ലഭിക്കുന്നു. ഓരോ സ്തരത്തിൽ നിന്നും പ്രതിസ്ഥാപനമില്ലാത്ത ലഘുയാദൃച്ഛികസാമ്പിളനരീതിയനുസരിച്ചാണു് അംഗങ്ങളെ തിരഞ്ഞെടുക്കുന്നതെന്നും സങ്കൽപിക്കുക.

h-ാമത്തെ സ്തരത്തിന്റെ സമാന്തരമാധ്യം $\overline{X}_h$ എന്നും, അതിൽ നിന്നു് സാമ്പിളിൽ ഉൾപ്പെടുത്തിയിട്ടുള്ള അംഗങ്ങളുടെ സമാന്തരമാധ്യം x_h എന്നും അങ്കനം

ചെയ്യുന്നു. അതു പോലെ തന്നെ h–ാമത്തെ സ്തരത്തിലെ അംഗങ്ങളുടെ വർഗമാധ്യം S^2_h എന്നും h–ാമതു സ്തരത്തിലെ സാമ്പിൾവർഗമാധ്യം s^2_h എന്നും സമഷ്ടിയുടെ സമാന്തരമാധ്യം $\bar{X}$ എന്നും, ആകെ സാമ്പിൾസമാന്തരമാധ്യം $\bar{x}$ എന്നും ഇരിക്കട്ടെ. എന്നാൽ,

$$\bar{X}_h = \frac{\sum_{i=1}^{N_h} X_{hi}}{N_h}, \qquad \bar{x}_h = \frac{\sum_{i=1}^{n_h} x_{hi}}{n_h}$$

$$S^2_h = \frac{\sum_{i=1}^{N_h} (X_{hi}-\bar{X}_h)^2}{N_h-1}, \qquad s^2_h = \frac{\sum_{i=1}^{n_h} (x_{hi}-\bar{x}_h)^2}{n_h-1}$$

$$\bar{X} = \frac{\sum_{h=1}^{k} N_h \bar{X}_h}{N}, \qquad \bar{x} = \frac{\sum_{h=1}^{k} n_h \bar{x}_h}{n}$$

എന്നിങ്ങനെ ഈ പ്രതീകങ്ങളെ നിർവചിച്ചിരിക്കുന്നു.

3. സമഷ്ടിയുടെ സമാന്തരമാധ്യത്തിന്റെ ഒരു അനഭിനതആകലം

ഇതിനായി നിർദ്ദേശിക്കപ്പെട്ടിരിക്കുന്ന സാമ്പിളജം,

$$\bar{x}_{st} = \frac{\sum_{h=i}^{k} N_h \bar{x}_h}{N} \qquad \ldots\ldots (4.3.1)$$

ഇതു് സമഷ്ടിയുടെ സമാന്തരമാധ്യത്തിന്റെ ഒരു അനഭിനത ആകലമാണെന്നു കാണിക്കുവാൻ കഴിയും. [$\bar{x}_{st}$ സാമ്പിൾ സമാന്തമാധ്യമല്ലെന്നുള്ളതു് പ്രത്യേകം ശ്രദ്ധിക്കുക. ഓരോ സ്തരത്തിൽ നിന്നും സാമ്പിളിലേക്കു തിരഞ്ഞെടുത്ത അംഗങ്ങളുടെ സമാന്തരമാധ്യത്തിന്റെ ഒരു ഭാരിത സമാന്തരമാധ്യമാണു് ഇതു്. ഭാരങ്ങളായി സ്വീകരിച്ചിരിക്കുന്നതു് $\frac{Nh}{N}$ ആണു്. അതായതു്, ഓരോ സ്തരത്തിലുമുള്ള അഗംങ്ങളുടെ അനുപാതം. ആകെ സാമ്പിൾസമാന്തരമാധ്യം

$$\bar{x} = \frac{\sum_{h=1}^{k} n_h \bar{x}_h}{n}$$ ആണല്ലൊ.

$\frac{Nh}{N} = \frac{nh}{n}$ ആണെങ്കിൽ $\bar{x}_{st} = \bar{x}$ എന്നു സിദ്ധിക്കും. അതായതു്, ഓരോ സ്തരത്തി

ല്യമുള്ള അംഗങ്ങളുടെ എണ്ണത്തിനു് ആനുപാതികമാണു് അതതു സ്തരത്തിൽ നിന്നു് സാമ്പിളിൽ ഉൾപ്പെടുത്തിയിട്ടുള്ള അംഗങ്ങളുടെ എണ്ണമെങ്കിൽ മുൻ നിർദേശിച്ച സാമ്പിളജം സാമ്പിൾ ആകെ സമാന്തരമാധ്യം തന്നെയായിരിക്കും. $n_1, n_2, \ldots n_k$ യുടെ മൂല്യങ്ങൾ ഈ രീതിയിൽ നിർണയിക്കുന്ന സമ്പ്രദായത്തിനു് ആനുപാതിക വിഭജന സമ്പ്രദായം എന്നു പേർ പറയും].

അടുത്തതായി $\bar{x}_{st}$ $\bar{X}$ ൻെറ ഒരു അനഭിനത ആകലമാണെന്നു തെളിയിക്കാം. ഓരോ സ്തരത്തിൽ നിന്നും പ്രതിസ്ഥാപനമില്ലാത്ത ലഘു യാദൃച്ഛിക സാമ്പിളന രീതി അവലംബിച്ചാണു് അംഗങ്ങളെ തിരഞ്ഞെടുക്കുന്നതെന്നാണല്ലോ സങ്കല്പിച്ചിരിക്കുന്നതു്. അതു കൊണ്ടു് $E(\bar{x}_h) = \bar{X}_h$ എന്നു വരുന്നു. അങ്ങനെ

$$E(\bar{x}_{st}) = E\left\{\frac{\sum_{h=1}^{k} N_h \bar{x}_h}{N}\right\}$$

$$= \frac{\sum_{h=1}^{k} N_h E(\bar{x}_h)}{N} = \frac{\sum_{h=1}^{k} N_h \bar{X}_h}{N} = \bar{X}$$

അതായതു് $\bar{x}_{st}$, $\bar{X}$ ൻെറ ഒരു അനഭിനത ആകലമാണു്. ഇവിടെ ഒരു സംഗതി ശ്രദ്ധിക്കേണ്ടതായുണ്ടു്. വിവിധ സ്തരങ്ങളിൽ നിന്നു് സാമ്പിളിലെ അംഗങ്ങളെ തിരഞ്ഞെടുക്കുന്നതു് ലഘുയാദൃച്ഛിക സാമ്പിളന രീതിയിൽ തന്നെ ആയിരിക്കണമെന്നു നിർബന്ധമില്ല. $E(\bar{x}_h) = \bar{X}_h$ ആയിട്ടുള്ള ഏതു സാമ്പിളന രീതി അവലംബിച്ചു് അംഗങ്ങളെ തിരഞ്ഞെടുത്താലും $\bar{x}_{st}$, $\bar{X}$ ൻെറ അനഭിനത ആകലമായിരിക്കും.

അടുത്തതായി $\bar{x}_{st}$ ൻെറ പ്രസരണം എന്താണെന്നു പരിശോധിക്കാം. $V(\bar{x}_{st})$ എന്നു് അതിനെ സാധാരണയായി അങ്കനം ചെയ്യുന്നു. സൗകര്യത്തിനു വേണ്ടി $W_h = \frac{N_h}{N}$ എന്നു് എഴുതാം.

$$\bar{x}_{st} \doteq \sum_{h=1}^{k} W_h \bar{x}_h \qquad (4.3.3)$$

വിവിധസ്തരങ്ങളിൽ നിന്നു് എടുത്തിരിക്കുന്ന സാമ്പിളുകൾ സ്വതന്ത്രങ്ങളാണല്ലോ. അതു കൊണ്ടു്, സ്വതന്ത്രങ്ങളായ യാദൃച്ഛികചരങ്ങളുടെ തുകയുടെ പ്രസരണം അവയുടെ പ്രസരണങ്ങളുടെ തുകയാണെന്നുള്ള സാംഖ്യിക ശാസ്ത്ര സിദ്ധാന്തം നമുക്കു് ഇവിടെ ഉപയോഗപ്പെടുത്താം. അങ്ങനെ

$$V(\bar{x}_{st}) = \sum_{h=1}^{k} W_h^2 V(\bar{x}_h)$$ എന്നു കിട്ടുന്നു.

ഓരോ സ്തരത്തെയും ഓരോ സമഷ്ടിയായി പരിഗണിക്കാം. h-ാമത്തെ സ്തരത്തിൽ N_h അംഗങ്ങളാണുള്ളതു്. അതിൽ നിന്നു് n_h അംഗങ്ങളെ പ്രതിസ്ഥാപനമില്ലാത്ത ലഘുയാദൃച്ഛികസാമ്പിളനരീതിയനുസരിച്ചു് തിരഞ്ഞെടുക്കുകയാണു് ചെയ്തതു്. അങ്ങനെ ലഭിച്ച സാമ്പിളിന്റെ സമാന്തരമാധ്യമാണു് $\bar{x}_h$. അതുകൊണ്ടു് 3.3 അനുസരിച്ചു്

$$V(\bar{x}_h) = \frac{N_h - n_h}{N_h} \cdot \frac{S_h^2}{n_h}$$

അങ്ങനെ $$V(\bar{x}_{st}) = \sum_{h=1}^{k} \frac{N_h^2}{N^2} \cdot \frac{N_h - n_h}{N_h} \cdot \frac{S_h^2}{n_h}$$

$$= \sum_{h=1}^{k} \left(\frac{1}{n_h} - \frac{1}{N_h} \right) W_h^2 S_h^2 \qquad (4.3.4)$$

എന്നു് സിദ്ധിക്കുന്നു. $\frac{n_h}{N_h}$ എല്ലാ സ്തരങ്ങളിലും താരതമ്യേന വളരെ ചെറുതാണെങ്കിൽ ഈ വ്യംജകത്തെ അൽപം കൂടി ലഘൂകരിക്കാൻ സാധിക്കും.

$$V(\bar{x}_{st}) \approx \sum_{h=1}^{k} W_h^2 \left(1 - \frac{n_h}{N_h} \right) \frac{S_h^2}{n_h}$$

$$\approx \sum_{h=1}^{k} W_h^2 \frac{S_h^2}{n_h} \qquad (4.3.5)$$

ഇവിടെ നാം ചെയ്തതു് ഓരോ സ്തരത്തിലും സാന്തസമഷ്ടീസംശോധനം $\left(1 - \frac{n_h}{N_h}\right)$ ഒഴിവാക്കുകയാണു്. സ്തരിതസാമ്പിളനത്തിൽ സംശോധനം ബുദ്ധിമുട്ടുള്ള ഒരു കാര്യമത്രെ. $\frac{n_h}{N_h}$ എല്ലാ സ്തരങ്ങളിലും ചെറുതാവണമെങ്കിൽ സ്തരങ്ങളുടെ പരിമാണം അതിൽ നിന്നു് സാമ്പിളനം ഉൾപ്പെടുത്തുന്ന അംഗങ്ങളുടെ എണ്ണത്തെക്കാൾ വളരെ വലുതായിരിക്കണം. സമഷ്ടിയുടെ പരിമാണം വലുതായിരിക്കുകയും സാമ്പിൾപരിമാണം താരതമ്യേന ചെറുതായിരിക്കുകയും ചെയ്യുമ്പോഴാണു് ഇതു സംഭവിക്കുക. അങ്ങനെയുള്ള സാഹചര്യത്തിലാണല്ലോ ലഘുയാദൃച്ഛികസാമ്പിളനത്തിലും സാന്തസമഷ്ടീസംശോധനം ഒഴിവാക്കുന്നതു്. ലഘുയാദൃച്ഛികസാമ്പിളനത്തിൽ ഈ സംശോധനത്തിനായി കുറക്കേണ്ടിവരുന്നതു് ഒരു നിശ്ചിത സംഖ്യയാണു്. പക്ഷേ ഇവിടെ സ്തരങ്ങളുടെ പരിമാണവും അതിൽ നിന്നു് സാമ്പിളിൽ ഉൾപ്പെടുത്തുന്ന അംഗങ്ങളുടെ എണ്ണവും പരി

ഗണിച്ച് അതു കണ്ടുപിടിക്കേണ്ടിയിരിക്കുന്നു. അതാണ് ഈ സംശോധനം സ്തരിതസാമ്പിളനത്തിൽ അല്പം ബുദ്ധിമുട്ടുണ്ടാക്കുന്നതാണെന്നു് മുമ്പു സൂചിപ്പിച്ചതു്.

ആനുപാതികവിഭജനരീതി അവലംബിക്കുമ്പോൾ $\frac{N_h}{N} = \frac{n_h}{n}$ ആയിരിക്കുമല്ലോ. അതു കൊണ്ടു്

$$V(\bar{x}_{st}) = \frac{1}{N^2} \sum_{h=1}^{k} N_h (N_h - n_h) \frac{S^2_h}{n_h}$$

$$= \frac{1}{N^2} \sum_{h=1}^{k} N_h \left(N_h - \frac{nN_h}{N}\right) \frac{NS^2_h}{nN_h}$$

$$= \sum_{h=1}^{k} \frac{N_h}{Nn} S^2_h \left(\frac{N-n}{N}\right)$$

$$= \frac{N-n}{Nn} \sum_{h=1}^{k} \frac{N_h}{N} S^2_h = \frac{N-n}{Nn} \sum_{h=1}^{k} W_h S^2_h \qquad (4.3.6)$$

ഈ പ്രത്യേക സന്ദർഭത്തിൽ സാന്തസമഷ്ടീസംശോധനം ഒരു നിശ്ചിതസംഖ്യയായ $\frac{N-n}{N}$ ആണെന്നുള്ളതു് ശ്രദ്ധേയമാണ്.

ഇവിടെ S_h^2, h–ാമത്തെ സ്തരത്തിൻെറ വർഗമാധ്യമാണല്ലൊ. എല്ലാ സ്തരങ്ങളിലെയും വർഗമാധ്യങ്ങൾ തുല്യമാണെങ്കിൽ, (ഉദാഹരണമായി $S^2_h = S^2$ എന്നിരിക്കട്ടെ), $V(\bar{x}_{st})$ ൻെറ വ്യഞ്ജകം ലളിതതരമായി തീരും.

$$V(\bar{x}_{st}) = \frac{N-n}{Nn} \sum_{h=1}^{k} W_h S_h^2 = \frac{N-n}{Nn} \sum_{h=1}^{k} W_h S^2$$

$$= \frac{N-n}{Nn} S^2 \sum_{h=1}^{k} \frac{N_h}{N} = \frac{N-n}{Nn} S^2 \qquad (4.3.7)$$

ഇതു് n പരിമാണമായുള്ള പുനസ്ഥാപനമില്ലാത്ത ഒരു ലഘുയാദൃച്ഛികസാമ്പിളനത്തിലെ സമാന്തരമാധ്യത്തിൻെറ പ്രസരണമാണെന്നു് കാണാമല്ലോ.

4. വിശ്വാസ്യതാന്തരാളങ്ങളുടെ ആകലനം

സമഷ്ടിയുടെ സമാന്തരമാധ്യത്തിൻെറ ഒരു അനഭിനത ആകലവും അതിൻെറ പ്രസരണവും നാം കണ്ടു കഴിഞ്ഞു. ഇവ ഉപയോഗപ്പെടുത്തി ലഘുയാദൃ

ച്ഛികസാമ്പിളനത്തിൽ എന്നതു പോലെ വിശ്വാസ്യതാന്തരാളങ്ങൾ ആകലനം ചെയ്യാൻ കഴിയും. ഇവിടെയും ആകലമായ $\bar{x}_{st}$, നോർമൽ വിതരണം അനുവർത്തിക്കുമെന്നു സങ്കൽപിച്ചാൽ, a വിശ്വാസ്യതാഗുണാങ്കമായ വിശ്വാസ്യതാന്തരാളം നിർണയിക്കാൻ,

$$P_r \left\{ -t_\alpha \leq \frac{\bar{x}_{st} - \bar{X}}{\sqrt{\bar{V}(\bar{x}_{st})}} \leq t_\alpha \right\} = a$$

ആയിരിക്കത്തക്ക വണ്ണം t_α കണ്ടു പിടിച്ചാൽ മതി. $\bar{x}_{st}$ നോർമൽ വിതരണം അനുവർത്തിക്കുന്നതു കൊണ്ടു് $\frac{\bar{x}_{st} - X}{\sqrt{V(\bar{x}_{st})}}$, മാനകനോർമൽ വിതരണത്തോടുകൂടിയതായിരിക്കും. നോർമൽ പട്ടികയിൽ നിന്നു് t_α കണ്ടുപിടിക്കുകയും ചെയ്യാം. അങ്ങനെ a വിശ്വാസ്യതാ ഗുണാങ്കത്തോടു കൂടിയ വിശ്വാസ്യതാന്തരാളം

$$\left\{ \bar{x}_{st} - t_\alpha \sqrt{V(\bar{x}_{st})}, \bar{x}_{st} + t_\alpha \sqrt{V(\bar{x}_{st})} \right\}$$

ആണെന്നു വരുന്നു. ഇവിടെ ഒരു ബുദ്ധിമുട്ടു് അനുഭവപ്പെടും.

$$V(\bar{x}_{st}) = \sum_{h=1}^{k} \left(\frac{1}{n_h} - \frac{1}{N_h} \right) W^2_h S^2_h$$

ആണല്ലൊ. ഇതിലെ S^2_h, അതായതു് സ്തരങ്ങളുടെ വർഗമാധ്യം, പലപ്പോഴും അജ്ഞാതമായിരിക്കും. സാമ്പിൾ ഉപയോഗിച്ചു തന്നെ അതു് ആകലനം ചെയ്യേണ്ടി വരും.

$$s^2_h = \frac{1}{(n_h - 1)} \sum_{i=1}^{n_h} (x_{hi} - \bar{x}_h)^2,$$ S^2_h, ഒരു അനഭിനത ആകലമാണെന്നു നമുക്കറിയാം. ഈ മൂല്യം S^2_h ന്റെ സ്ഥാനത്തു് പ്രതിസ്ഥാപനം ചെയ്തു് വിശ്വാസ്യതാന്തരാളം കണ്ടുപിടിക്കാവുന്നതാണു്. പക്ഷെ സാമ്പിൾ പരിമാണം വലുതാണെങ്കിൽ മാത്രമെ ഇങ്ങനെ ചെയ്യുന്നതു് താത്വികമായി ശരിയായിരിക്കുകയുള്ളു. ചെറിയ സാമ്പിളുകളെ സംബന്ധിച്ചിടത്തോളം

$$\frac{\bar{x}_{st} - \bar{X}}{\sqrt{\tilde{V}(\bar{x}_{st})}}$$

[$\tilde{V}(\bar{x}_{st})$ അങ്കനം ചെയ്തിരിക്കുന്നതു് $V(\bar{x}_{st})$ ന്റെ അനഭിനത ആകലത്തെയാണു്] മാനക നോർമൽ വിതരണമല്ല. സ്റ്റുഡന്റു് t വിതരണമാണു് അനുവർത്തിക്കുന്നതു്. പക്ഷെ അതിന്റെ സ്വാതന്ത്ര്യതാസംഖ്യയുടെ നിർണയവും മററും താരതമ്യേന സങ്കീർണമായതു കൊണ്ടു്, അതു് ഏകദേശമായി നോർമൽ വിതരണം അനുവർത്തിക്കുന്ന എന്ന സങ്കൽപത്തിൽ വിശ്വാസ്യതാന്തരാളങ്ങൾ നിർണയിക്കുകയാണു് പതിവു്.

$V(\bar{x}_{st})$ ന്റെ ആകലം കണ്ടുപിടിക്കുന്നതിൽ പ്രത്യേകം ശ്രദ്ധിക്കാനുള്ള ഒരു കാര്യമുണ്ട്. S^2_h ന്റെ ആകലമായി s^2_h ആണല്ലോ സ്വീകരിക്കുന്നത്. h–ാമത്തെ സ്തരത്തിൽ നിന്ന് സാമ്പിളിൽ ഒരംഗത്തെ മാത്രമെ ഉൾപ്പെടുത്തുന്നുള്ളു എങ്കിൽ s^2_h ന്റെ മൂല്യം '0' ആയിരിക്കും. മററു സ്തരങ്ങളിലെയും സ്ഥിതി ഇതു തന്നെ. ചുരുക്കത്തിൽ എല്ലാ സ്തരങ്ങളിൽ നിന്നും ഒരംഗത്തെ വീതമാണ് സാമ്പിളിൽ ഉൾപ്പെടുത്തുന്നതെങ്കിൽ $\tilde{v}(\bar{x}_{st}) = 0$ എന്ന നിഗമനത്തിലാണ് എത്തിച്ചേരുക. ഇത് അസ്വീകാര്യമായ ഒരു ആകലമാണല്ലൊ. അതു കൊണ്ടു എല്ലാ സ്തരങ്ങളിൽ നിന്നും ഒന്നിലധികം അംഗങ്ങളെ സാമ്പിളിൽ ഉൾപ്പെടുത്തിയിട്ടുണ്ടെങ്കിൽ മാത്രമെ $V(\bar{x}_{st})$ ന്റെ ആകലനം സാധ്യമാവൂ.

5. വിഭജനസമ്പ്രദായങ്ങൾ

N അംഗങ്ങളുള്ള ഒരു സമഷ്ടിയെ യഥാക്രമം $N_1, N_2, .. N_k$ അംഗങ്ങൾ വീതമുള്ള k സ്തരങ്ങളായി തിരിച്ചിരിക്കുന്നു. n അംഗങ്ങളുള്ള ഒരു സാമ്പിളാണ് എടുക്കേണ്ടിയിരിക്കുന്നത്. ആകെ n അംഗങ്ങളുള്ള സാമ്പിളിനെ ഓരോ സ്തരത്തിനും എത്ര അംഗങ്ങൾ വീതം വിഭജിക്കാം എന്നുള്ളതാണ് 'വിഭജനം' എന്നതുകൊണ്ട് ഇവിടെ അർത്ഥമാക്കുന്നത്. മറെറാരു തരത്തിൽ പറഞ്ഞാൽ, n_1 $n_2, \cdots n_k$ അംഗങ്ങളെയാണ് യഥാക്രമം ഓരോ സ്തരത്തിൽ നിന്നും സാമ്പിളിൽ ഉൾപ്പെടുത്തുന്നത് എന്നു സങ്കല്പിച്ചാൽ, അവയുടെ മൂല്യങ്ങൾ എങ്ങനെ നിർണയിക്കാമെന്നതാണ് ഇവിടുത്തെ ചർച്ചാ വിഷയം. ലഭ്യമായ വിവരങ്ങളും, വിവരം ശേഖരിക്കുന്നതിനുള്ള ചെലവ്, ഓരോ സ്തരത്തിന്റെയും പ്രസരണം മുതലായ പരിഗണനകളും വിഭജനരീതി തിരഞ്ഞെടുക്കുന്നതിൽ ശ്രദ്ധിക്കപ്പെടേണ്ടവയാണ്. സാധാരണ ഉപയോഗത്തിലുള്ള ചില വിഭജന രീതികളും അവയുടെ പ്രത്യേകതകളും താഴെ കൊടുക്കുന്നു.

(a) ആനുപാതിക വിഭജന രീതി.

ഏററവും ലളിതമായ വിഭജന രീതിയാണ് ഇത്. ഓരോ സ്തരത്തിലും എത്ര അംഗങ്ങൾ വീതമുണ്ടെന്നല്ലാതെ മറ്റൊന്നും അറിഞ്ഞുകൂടാതെ വരുമ്പോഴാണ് ഈ രീതി സ്വീകരിക്കാറുള്ളത്.

$$n_h = \frac{nN_h}{N} = n\,N_h \qquad \ldots 4.5.1$$

എന്ന നിയമമനുസരിച്ചാണ് ഓരോ സ്തരത്തിൽ നിന്നും സാമ്പിളിൽ ഉൾപ്പെടുത്തേണ്ട അംഗങ്ങളുടെ എണ്ണം നിശ്ചയിക്കുന്നത്. അതായത്, ഓരോ സ്തരങ്ങളിൽ നിന്നും സാമ്പിളിൽ ഉൾപ്പെടുത്തുന്ന അംഗങ്ങളുടെ എണ്ണം സ്തരത്തിന്റെ പരിമാണത്തിനു ആനുപാതികമായിരിക്കും. വിവിധ സ്തരങ്ങൾ തമ്മിൽ കാര്യമായ വ്യത്യാസമില്ലാത്ത സന്ദർഭത്തിലാണ് ഈ രീതി അംഗീകരിക്കപ്പെടുന്നത്. സമഷ്ടിയെ സ്തരങ്ങളായി വിഭജിച്ചിരിക്കുന്നതു് ഭരണ സൗകര്യത്തിനു വേണ്ടി മാത്രമാണെങ്കിൽ ഈ സങ്കല്പത്തിൽ വലിയ തെററുണ്ടാവില്ല. അങ്ങനെയുള്ള സന്ദർഭ

ങ്ങളിൽ ആനുപാതിക വിഭജന രീതി സ്വീകരിക്കുന്നതിൽ തെററില്ല. നേരെ മറിച്ചു്, സമഷ്ടിയുടെ വിവിധാത്മകത കാരണം, ഏകാത്മകങ്ങളായ സ്തരങ്ങളായി അതിനെ വിഭജിച്ചിരിക്കുകയാണെങ്കിൽ സ്തരങ്ങൾ തമ്മിലുള്ള വ്യത്യാസം ഗൗരവതരമായിരിക്കുമെന്നതു കൊണ്ടു് ഈ രീതി സ്വീകാര്യമല്ല തന്നെ. പ്രസരണം കൂടുതലുള്ള സ്തരങ്ങളിൽ നിന്നു് കൂടുതൽ അംഗങ്ങളെ സാമ്പിളിൽ ഉൾപ്പെടുത്തേണ്ടതു് സാമ്പിളിന്റെ പ്രാതിനിധ്യ സ്വഭാവം നിലനിത്താൻ ആവശ്യമാണു്. ആനുപാതിക വിഭജന രീതിയിൽ പരിമാണം കൂടുതലുള്ള സ്തരങ്ങളിൽ നിന്നാണു് കൂടുതലംഗങ്ങളെ തിരഞ്ഞെടുക്കുന്നതു്. കൂടുതൽ വലുപ്പമുള്ള സ്തരങ്ങളിൽ പ്രസരണം കൂടുതലും മററുള്ളവയിൽ കുറവും ആണെന്നു വിശ്വസിക്കാൻ ന്യായമുണ്ടെങ്കിൽ ഇത്തരം സാമ്പിളുകളിലും ഈ രീതിയുടെ ഉപയോഗം ന്യായീകരിക്കാൻ കഴിയും. ചുരുക്കത്തിൽ സ്തരങ്ങൾ തമ്മിൽ കാര്യമായ വ്യത്യാസമില്ലാതിരിക്കുകയും സ്തരങ്ങളിലെ പ്രസരണം അവയുടെ പരിമാണത്തിനു് ആനുപാതികമായിരിക്കുകയും ചെയ്യുന്ന സന്ദർഭങ്ങളിലാണു് ആനുപാതിക വിഭജനരീതി സ്വീകരിക്കേണ്ടതു്. എന്നിരുന്നാലും, ഈ സമ്പ്രദായത്തിൻെറ ലാളിത്യം കാരണം പലപ്പോഴും സ്തരങ്ങളുടെ പ്രത്യേകതകൾ പരിഗണിക്കാതെ തന്നെ ഈ വിഭജനസമ്പ്രദായം ഉപയോഗപ്പെടുത്താറുണ്ടു്.

സാമ്പിളെടുക്കുന്നതിനുള്ള ചെലവു് വിഭജനസമ്പ്രദായം പരിഗണിക്കുമ്പോൾ പ്രത്യേകം കണക്കിലെടുക്കേണ്ടതാണു്. സാമ്പിളിൽ ഉൾപ്പെടുത്തേണ്ട അംഗങ്ങളുടെ എണ്ണം വർധിക്കുമ്പോൾ ചെലവും വർധിക്കുക എന്നതു് സ്വാഭാവികമത്രെ. പക്ഷെ സാമ്പിൾ പരിമാണം എന്തു തന്നെയായിരുന്നാലും പറയത്തക്ക മാററമൊന്നും വരാത്ത ചില ചെലവുകളുണ്ടു്. സാമ്പിളനം ആസൂത്രണം ചെയ്യുന്ന ആളിൻെറ ശമ്പളം, ഓഫീസു് സ്ഥാപിക്കുന്നതിനുള്ള ചെലവുകൾ തുടങ്ങിയവ. ഇങ്ങനെയുള്ള ചെലവുകൾ C_o ആണെന്നിരിക്കട്ടെ. വിവരശേഖരണത്തിൻെറ ചെലവു് ഓരോ അംഗത്തിനും ഓരോന്നായിരിക്കും. അതു് ആ അംഗത്തെ കണ്ടെത്താനുള്ള ചെലവു്, വിവരശേഖരണത്തിനെടുക്കുന്ന സമയം തുടങ്ങിയവയെ ആശ്രയിച്ചിരിക്കുന്നു. ഒരു സ്തരത്തിൽ പെട്ട അംഗങ്ങളെ സംബന്ധിച്ചിടത്തോളം അതു് ഏതാണ്ടു് തുല്യമായിരിക്കുമെന്നു് സങ്കൽപിക്കുന്നതിൽ വലിയ അപാകതയില്ല. h–ാമത്തെ സ്തരത്തിലെ ഒരംഗത്തിൽ നിന്നു് വിവരം ശേഖരിക്കാനുള്ള ശരാശരി ചെലവു് C_h ആണെന്നു സങ്കൽപിക്കുക. അങ്ങനെയാണെങ്കിൽ ആകെ ചെലവു്

$$C = C_o + \sum_{h=1}^{k} r_h C_h \qquad 4.5.2$$

എന്നു സിദ്ധിക്കുന്നു.

നിർദിഷ്ടമായ സാമ്പിളനത്തിനു് ആകെ ചെലവാക്കാവുന്ന പണം $\bar{C}$ ആണെന്നിരിക്കട്ടെ. ആനുപാതികവിഭജനരീതിയിലാണു് സാമ്പിളെടുക്കുന്നതെങ്കിൽ സാമ്പിൾ പരിമാണമായ n താഴെ പറയും പ്രകാരം നിർണ്ണയിക്കാം.

$$n_h = nW_h$$

$$\bar{C} = C_o + \sum_{h=1}^{k} nW_hC_h$$

അതായതു്, $$n = \frac{C - C_o}{\sum_{h=1}^{k} W_h C_h} \qquad 4.5.3$$

അങ്ങനെ n ന്റെ മൂല്യം നിർണയിച്ചു കഴിഞ്ഞാൽ, $n_h = nW_h$ എന്ന നിയമമനുസരിച്ചു് $n_1, n_2, \dots n_k$ എന്നിവയും നിർണയിക്കാം.

(b) ഇഷ്ടതമവിഭജനരീതി

ആകെ ചെലവു ചെയ്യാവുന്ന പണം $\bar{C}$ ആണെന്നു തന്നിരിക്കുന്നു എന്നിരിക്കട്ടെ: സാമ്പിളനലക്ഷ്യം സമഷ്ടിയുടെ സമാന്തരമാധ്യം ആകലനം ചെയ്യുകയാണു്. $\bar{x}_{st}$ യെ ആണു് ആകലമായി സ്വീകരിക്കുന്നതു്. $V(\bar{x}_{st})$ ഏററവും കുറഞ്ഞിരിക്കത്തക്ക വണ്ണവും ആകെ ചെലവു്, $\bar{C}$ ൽ കവിയാത്ത വിധത്തിലും n_1, $n_2, \cdots n_k$ എന്നിവയുടെ മൂല്യങ്ങൾ നിർണയിക്കുന്ന രീതിക്കു് ഇഷ്ടതമവിഭജന രീതി എന്നു പറയും. ഇതു പോലെ തന്നെ $V(\bar{x}_{st})$ ന്റെ മൂല്യം V എന്നു് തന്നിരിക്കുന്ന ഒരു സംഖ്യയിൽ കവിയാതിരിക്കുകയും ആകെ ചെലവു് ഏററവും ചെറുതായിരിക്കുകയും ചെയ്യത്തക്ക വണ്ണം $n_1, n_2, \cdots n_k$ നിർണയിക്കുന്നതും ഇഷ്ടതമ വിഭജനരീതിയത്രെ. ഇങ്ങനെ ഇഷ്ടതമവിഭജനം രണ്ടു തരത്തിലുണ്ടു്. അവ ഓരോന്നും നമുക്കു് പരിശോധിക്കാം.

1. $C = \bar{C}$ ഉം $V(\bar{x}_{st})$ ഏററവും ചെറുതുമായിരിക്കത്തക്കവണ്ണമുള്ള വിഭജനം.
   സാന്തസമഷ്ടിശോധനം അവഗണിച്ചാൽ,

$$V(\bar{x}_{st}) = \sum_{h=1}^{k} W_h^2 \frac{S_h^2}{n_h}$$ ആണെന്നു് 4.3 ൽ തെളിയിച്ചിട്ടുണ്ടല്ലോ.

$$\bar{C} = C_o + \sum_{h=1}^{k} n_h C_h$$

എന്ന നിബന്ധനയ്ക്കു വിധേയമായി $V(\bar{x}_{st})$ നെ നിമ്നതമമാക്കിയാണു് n_h ന്റെ മൂല്യം നിർണയിക്കേണ്ടിയിരിക്കുന്നതു്. ലഗ്രാൻജെ സമ്പ്രദായമുപയോഗിച്ചു് ഇതു നിർവഹിക്കാവുന്നതാണു്. അതനുസരിച്ചു്,

$$L = V(\bar{x}_{st}) - \lambda\left(\bar{C} - C_o - \sum_{h=1}^{k} n_h C_h\right)$$

$$= \sum_{h=1}^{k} W^2_h \frac{S^2_h}{n_h} - \lambda (\bar{C} - C_o - \sum_{h=1}^{k} n_h C_h)$$

എന്ന വ്യംജകത്തെ നിമ്നതമമാക്കുന്ന n_h കളുടെ മൂല്യങ്ങളാണ് ആവശ്യം. ഇതു് n_h കളുടെ ഒരു ഫലനമാണ്. ഇതിൻെറ ഭാഗികഅവകലങ്ങൾ പൂജ്യത്തോടു തുല്യമാണെന്നു് എഴുതിയാൽ, n_h കൾ നിർണയിക്കാനുള്ള സമീകരണങ്ങൾ കിട്ടും. n_h ആധാരമാക്കി അവകലനം ചെയ്യുമ്പോൾ,

$$- \frac{W^2_h S^2_h}{n^2_h} + \lambda C_h = 0$$

അതായതു്, $n_h = \frac{W_h S_h}{\sqrt{\lambda C_h}}$. ഇതിൽ λ ഒരു അജ്ഞാതസംഖ്യയാണ്.

അതു നിർണയിക്കാനായി

$$\bar{C} = C_o + \sum_{h=1}^{k} n_h C_h$$ എന്ന സമീകരണത്തിൽ,

n_h ൻെറ മൂല്യം പ്രതിസ്ഥാപനം ചെയ്യുമ്പോൾ,

$$\bar{C} = C_o + \sum_{h=1}^{k} \frac{W_h S_h}{\sqrt{\lambda C_h}} C_h$$

$$\sqrt{\lambda} = \frac{\sum_{h=1}^{k} W_h S_h \sqrt{C_h}}{\bar{C} - C_o}$$

അങ്ങനെ,

$$n_h = \frac{(\bar{C} - C_o) \frac{W_h S_h}{\sqrt{C_h}}}{\sum_{h=1}^{k} W_h S_h \sqrt{C_h}} \qquad 4.5.4$$

മറെറാരു തരത്തിൽ പറഞ്ഞാൽ ഇറ് ഇഷ്ടതമവിഭജനത്തിൽ $\frac{W_h S_h}{\sqrt{C_h}}$ നു് ആനുപാതികമാണ് h-ാമത്തെ സ്തരത്തിൽ നിന്നു് സാമ്പിളിൽ ഉൾപ്പെടുത്തുന്ന അംഗങ്ങളുടെ എണ്ണം. ഈ വിഭജനരീതി സ്വീകരിച്ചാൽ ആ സാമ്പിളിൽ നിന്നു കിട്ടുന്നതു് നിർദിഷ്ടമായ ചെലവിൽ ലഭിക്കാവുന്ന ഏററവും സൂക്ഷ്മതയുള്ള ആകലമായിരിക്കും.

സാമ്പിൾ പരിമാണം $= n = \sum_{h=1}^{k} n_h$

$$= \frac{(\bar{C}-C_o) \sum_{h=1}^{k} \frac{W_h S_h}{\sqrt{C_h}}}{\sum_{h=1}^{k} W_h S_h \sqrt{C_h}} \qquad 4.5.5$$

ആകലത്തിന്റെ പ്രസരണം $= \dfrac{\sum_{h=1}^{k} W^2_h S^2_h \sqrt{C_h} \sum_{h=1}^{k} W_h S_h \sqrt{C_h}}{(\bar{C}-C_o)\, W_h S_h}$

$$= \frac{1}{(\bar{C}-C_o)} \left(\sum_{h=1}^{k} W_h S_h \sqrt{C_h} \right)^2 \qquad 4.5.6$$

ഇവിടെ വിവരശേഖരണത്തിനുള്ള ചെലവ് ഓരോ സ്തരത്തിനും ഓരോന്നാണെന്നാണല്ലോ സങ്കൽപിച്ചത്. അത് തുല്യമാണെന്നു കരുതിയാൽ, അതായത്, $C_1 = C_2 = \cdots = C_k = C$ എന്നു സങ്കൽപിച്ചാൽ,

$$n_h = \frac{(\bar{C}-C_o) \frac{W_h S_h}{\sqrt{C}}}{\sum_{h=1}^{k} W_h S_h \sqrt{C}}$$

$$= \frac{(\bar{C}-C_o)}{C} \cdot \frac{W_h S_h}{\sum_{h=1}^{k} W_h S_h}$$

$$= n \frac{W_h S_h}{\sum_{h=1}^{k} W_h S_h} = n \cdot \frac{N_h S_h}{\sum_{h=1}^{k} N_h S_h} \qquad 4.5.7$$

കുറിപ്പ്:— C_h തുല്യമാകുമ്പോൾ

$$\bar{C} = C_o + C \sum_{h=1}^{k} n_h$$

$$= C_o + nC$$

$$n = \frac{\bar{C} - C_o}{C}$$

അതായതു് h–ാമത്തെ സ്തരത്തിൽ നിന്നും എടുക്കേണ്ട അംഗങ്ങളുടെ എണ്ണം $W_h S_h$ നു് അല്ലെങ്കിൽ $N_h S_h$ നു് ആനുപാതികമാണു്.

അപ്പോഴത്തെ $V(\bar{x}_{st}) = \frac{1}{\bar{C} - C_o} \left(\sum_{h=1}^{k} W_h S_h \sqrt{C} \right)^2$

$$= \frac{C}{\bar{C} - C_o} \left(\sum_{h=1}^{k} W_h S_h \right)^2 = \frac{1}{n} \left(\sum_{h=1}^{k} W_h S_h \right)^2 \qquad 4.5.8$$

C_h തുല്യമാകുന്നതു കൂടാതെ വിവിധ സ്തരങ്ങളുടെ വർഗമാധ്യങ്ങൾ കൂടി തുല്യമാണെങ്കിൽ [$S_h = S$], ഈ വിഭജനസമ്പ്രദായം സ്വാഭാവികമായും ആനുപാതികവിഭജനസമ്പ്രദായം തന്നെയാകുമെന്നു് കാണാം.

$$n_h = n \frac{N_h S_h}{\sum_{h=1}^{k} N_h S_h} = n \frac{S_h N_h}{S \sum_{h=1}^{k} N_h} = n \left(\frac{N_h}{N} \right)$$

2. $V(\bar{x}_{st}) = \bar{V}$ ഉം ചെലവു് ഏററവും കുറവും ആയിരിക്കത്തക്ക വണ്ണമുള്ള വിഭജനം.

ഇവിടെയും സമീപനരീതി മുൻപിലത്തേതു തന്നെ. ലഗ്രാൻജെ സമ്പ്രദായം ഉപയോഗിച്ചു്

$C = C_o + \sum_{h=1}^{k} n_h C_h$ എന്ന ഫലനത്തെ

$$\bar{V} = \sum_{h=1}^{k} \frac{W_h^2 S_h^2}{n_h}$$

എന്ന ഉപാധിക്കു വിധേയമായി നിമ്നതമമാക്കുകയാണു് വേണ്ടതു്.

$$L = C_o + \sum_{h=1}^{k} n_h C_h + \lambda \left(\sum_{h=1}^{k} \frac{W_h^2 S_h^2}{n_h} - \bar{V} \right)$$

ഇതിന്റെ ഭാഗികഅവകലനങ്ങളെ പൂജ്യത്തോടു് സമീകരിക്കുമ്പോൾ,

$$C_h - \lambda \frac{W_h^2 S_h^2}{n_h^2} = 0$$

അതായത്, $n_h = \dfrac{\sqrt{\lambda}\, W_h S_h}{\sqrt{C_h}}$

λ യുടെ മൂല്യം കാണാനായി, $\bar{V} = \sum\limits_{h=1}^{k} \dfrac{W_h^2 S_h^2}{n_h}$

എന്ന സമീകരണത്തിൽ n_h ൻെറ മൂല്യം പ്രതിസ്ഥാപിക്കുമ്പോൾ,

$$\bar{V} = \sum_{h=1}^{k} \frac{W_h^2 S_h^2 \sqrt{C_h}}{\sqrt{\lambda}\, W_h S_h}$$

$$\sqrt{\lambda} = \frac{1}{\bar{V}} \sum_{h=1}^{k} W_h S_h \sqrt{C_h}$$

അതുകൊണ്ട്,

$$n_h = \frac{1}{\bar{V}} \left(\sum_{h=1}^{k} W_h S_h \sqrt{C_h} \right) \frac{W_h S_h}{\sqrt{C_h}} \qquad 4.5.9$$

ഈ വിഭജനരീതി സ്വീകരിച്ചാലത്തെ ആകെ ചെലവ്,

$$C = C_0 + \frac{1}{\bar{V}} \left(\sum_{h=1}^{k} W_h S_h \sqrt{C_h} \right)^2 \qquad 4.5.10$$

ഇവിടെയും $\dfrac{W_h S_h}{\sqrt{C_h}}$ ന് ആനുപാതികമാണ് h-ാമത്തെ സ്തരത്തിൽ നിന്നു് സാമ്പിളിൽ ഉൾപ്പെടുത്തുന്ന അംഗങ്ങളുടെ എണ്ണം.

ഇഷ്ടതമവിഭജനരീതിയാണ്, താത്വികമായി പറഞ്ഞാൽ ഏററവും നല്ല വിഭജനരീതി. പക്ഷെ പ്രായോഗികമായി ചില ബുദ്ധിമുട്ടുകൾ കാരണം ചുരുക്കമായേ ഈ രീതി സ്വീകരിക്കാറുള്ളൂ. ഈ രീതിയുടെ ഏറ്റവും വലിയ പ്രായോഗിക അസൗകര്യം, സ്തരങ്ങളുടെ പ്രസരണങ്ങൾ (വർഗമാധ്യങ്ങൾ) അറിയാതെ ഇതു് ഉപയോഗപ്പെടുത്താനാവില്ലെന്നുള്ളതാണ്. സ്തരങ്ങളുടെ പ്രസരണങ്ങൾ മിക്കവാറും അജ്ഞാതങ്ങളായിരിക്കുകയും ചെയ്യും. മുമ്പു നടത്തിയിട്ടുള്ള അന്വേഷണങ്ങളിൽ നിന്നോ പ്രത്യേകമായി ആസൂത്രണം ചെയ്യുന്ന പ്രാഥമിക അന്വേഷണത്തിൽ നിന്നോ ഇവയുടെ ഏകദേശമൂല്യങ്ങൾ നിർണയിക്കണം. അതു് ബുദ്ധിമുട്ടായി തോന്നിയാൽ നാം നടത്താനുദ്ദേശിക്കുന്ന അന്വേഷ

ണം തന്നെ പല ഘട്ടങ്ങളായി നടത്തി ഈ വിഭജനരീതിയിലുള്ള സാമ്പിൾ എടുക്കാം. ആദ്യഘട്ടത്തിൽ ലഭിക്കുന്ന വിവരങ്ങൾ കൊണ്ട് സ്തരങ്ങളുടെ പ്രസരണങ്ങൾ ഏകദേശമായി നിർണയിക്കുകയും ആ ഏകദേശപ്രസരണങ്ങളുടെ അടിസ്ഥാനത്തിൽ ഓരോ സ്തരത്തിൽ നിന്നും പോരാതെയുള്ള അംഗങ്ങളെ കൂടി സാമ്പിളിൽ ഉൾപ്പെടുത്തി വിവരങ്ങൾ ശേഖരിക്കുകയും അടുത്ത ഘട്ടത്തിലും ഇതേ നടപടി തന്നെ ആവർത്തിക്കുകയും അങ്ങനെ ആവശ്യമുള്ള പരിമാണത്തോടു കൂടിയ സാമ്പിളിൽ എത്തിച്ചേരുകയുമാണ് ഈ രീതി.

പൊതുവെ പറഞ്ഞാൽ വലിയ സ്തരങ്ങളിൽ നിന്നും കൂടുതൽ വിവിധാത്മകതയുള്ള സ്തരങ്ങളിൽ നിന്നും (പ്രസരണം വലുതായിട്ടുള്ള സ്തരങ്ങളിൽ നിന്നും) വിവരം ശേഖരിക്കാൻ അധികം ചെലവു വരാത്ത സ്തരങ്ങളിൽ നിന്നും, കൂടുതലംഗങ്ങളെ സാമ്പിളിൽ ഉൾപ്പെടുത്തുന്നതു് ചെലവു കുറവും സൂക്ഷ്മത കൂടുതലും ആക്കാൻ സഹായിക്കും. രണ്ടു തരത്തിലുള്ള ഇഷ്ടതമവിഭജനരീതികൾ ചർച്ച ചെയ്തതിൽ ആദ്യത്തേതാണ് കൂടുതൽ പ്രധാനപ്പെട്ടതു് എന്നും ഓർത്തിരിക്കേണ്ടതാണ്.

(c) സ്തരങ്ങളിലെ അഭിലക്ഷണമൂല്യങ്ങളുടെ തുകയ്ക്ക് ആനുപാതികമായ വിഭജനം

h–ാമത്തെ സ്തരത്തിലെ അഭിലക്ഷണമൂല്യങ്ങളുടെ തുക $N_h \bar{X}_h$ ആയിരിക്കുമല്ലോ. സമഷ്ടിയുടെ തുക $N\bar{X}$ ആണ്.

$$n_h = \frac{n\, N_h \bar{X}_h}{N\bar{X}} \qquad 4.4.11$$

എന്ന നിയമമനുസരിച്ച് n_h കൾ നിർണയിക്കുകയാണ് ഈ സമ്പ്രദായത്തിൽ ചെയ്യുന്നതു്. സ്തരങ്ങളിലെ പ്രസരണം തുല്യമാണെന്നു കരുതാൻ ന്യായമുള്ള സന്ദർഭങ്ങളിൽ ഇതു് ഇഷ്ടതമവിഭജനരീതി തന്നെയാണ്. മററു സന്ദർഭങ്ങളിൽ ആ രീതിയോടു് അടുത്തു വരികയും ചെയ്യും.

(d) പരാസത്തിൻെറ അടിസ്ഥാനത്തിലുള്ള വിഭജനം

R_h, h–ാമത്തെ സ്തരത്തിലെ അഭിലക്ഷണമൂല്യങ്ങളുടെ പരാസം (range) ആണെന്നിരിക്കട്ടെ.

$$n_h = \frac{n\, W_h R_h}{\sum_{h=1}^{k} W_h R_h}$$ എന്ന സമീകരണം അനുസരിച്ച്

n_h കളുടെ മൂല്യം നിർണയിക്കുകയാണ് ഈ സമ്പ്രദായം. R_h മിക്കവാറും എല്ലാ സമഷ്ടികളെ സംബന്ധിച്ചും S_h ൻെറ ഒരു ഏകദേശനമായിരിക്കും. വിവിധ സ്തരങ്ങളിൽ നിന്നു് വിവരം ശേഖരിക്കുന്നതിനുള്ള ചെലവു് തുല്യമാണെങ്കിൽ

ഇഷ്ടതമവിഭജനരീതിയിൽ,

$$n_h = \frac{nW_hS_h}{\sum_{h=1}^{k} W_h S_h}$$ ആണല്ലൊ.

ഈ സമീകരണത്തിൽ $S_h = R_h$ എന്നെഴുതിയാൽ മുൻകൊടുത്തിരിക്കുന്ന സമീകരണം കിട്ടും. അങ്ങനെ ഈ വിഭജനരീതിയും ഇഷ്ടതമവിഭജനരീതിയുടെ ഒരു ഏകദേശനം തന്നെയാണ്. വിവിധ സ്തരങ്ങളിൽ നിന്ന് വിവരം ശേഖരിക്കാനുള്ള ചെലവുകൾ വ്യത്യസ്തങ്ങളായിരിക്കുമ്പോൾ ഈ വിഭജനരീതിക്ക് ആധാരമായ സമീകരണത്തെ,

$$n_h = \frac{(\bar{C} - C_o)\,\dfrac{W_h R_h}{\sqrt{C_h}}}{\sum_{h=1}^{k} W_h R_h \sqrt{C_h}}$$

എന്ന് ഭേദഗതി ചെയ്യാവുന്നതാണ്. ഈ സമീകരണം ഇഷ്ടതമവിഭജനരീതിയിലെ സംഗതമായ സമീകരണത്തിൽ $S_h = R_h$ എന്ന് എഴുതിയാൽ കിട്ടുന്നതത്രെ. R_h, S_h ന്റെ ഏകദേശനമാണെന്നുള്ള സങ്കൽപം ചില സമഷ്ടികളെ സംബന്ധിച്ചിടത്തോളം അത്ര ശരിയല്ല. അങ്ങനെ വരുമ്പോൾ ഈ സമ്പ്രദായം കഴിവതും ഉപയോഗിക്കാതിരിക്കുന്നതാണ് നല്ലതു്.

6. സ്തരണം കൊണ്ടുള്ള നേട്ടം

നാമിപ്പോൾ രണ്ടു തരം സാമ്പിളനരീതികളാണ് ചർച്ച ചെയ്തതു്. ലഘുയാദൃച്ഛികസാമ്പിളനവും സ്തരിതസാമ്പിളനവും. ഇതിൽ ലഘുയാദൃച്ഛികസാമ്പിളനം താത്വികമായി താരതമ്യേന ലളിതവും എളുപ്പത്തിൽ എടുക്കാൻ കഴിയുന്നതുമാണ്. സ്തരിതസാമ്പിളനം അങ്ങനെയല്ല. കാര്യമായ എന്തെങ്കിലും പ്രയോജനമില്ലെങ്കിൽ സങ്കീർണമായ ഒരു സാമ്പിളനരീതി അംഗീകരിക്കുന്നതിൽ അർത്ഥമില്ല. അതുകൊണ്ട് സ്തരണം കൊണ്ടുള്ള നേട്ടമെന്താണെന്ന് പരിശോധിക്കേണ്ടതു് ആവശ്യമാണ്.

സ്തരിതസാമ്പിളനത്തിൽ പ്രധാനമായും രണ്ടു തരം വിഭജനരീതികളാണ് നിലവിലുള്ളതു്. ആനുപാതികവിഭജനരീതിയും ഇഷ്ടതമവിഭജനരീതിയും. മറ്റു വിഭജനരീതികൾ ഇവയുടെ പരിഷ്കരിച്ച പതിപ്പുകളത്രെ. അതുകൊണ്ട് ഈ രണ്ടു വിഭജനരീതികളെ അടിസ്ഥാനമാക്കിയാണ് സ്തരിതസാമ്പിളനവും ലഘുയാദൃച്ഛിക സാമ്പിളനവും താരതമ്യപ്പെടുത്തുന്നതു്.

ഒരു സാമ്പിളനരീതിയുടെ ഉൽക്കർഷം നിർണയിക്കുന്നതു് ആ രീതിയിലെ

ട്ടത്ത സാമ്പിളിൽ നിന്നു ലഭിക്കുന്ന ആകലങ്ങളുടെ സൂക്ഷ്മതയാണ്. ഒരു ആകലത്തിൻെറ സൂക്ഷ്മത അതിൻെറ പ്രസരണവുമായി വിപരീതാനുപാതത്തിലാണെന്നു് ഒന്നാമധ്യായത്തിൽ നാം കണ്ടു കഴിഞ്ഞു. അതായതു് പ്രസരണം എത്ര കുറയുന്നോ സൂക്ഷ്മത അത്രയും കൂടുതലാണ്. സമഷ്ടിയുടെ സമാന്തര മാധ്യത്തിൻെറ ആകലനമാണല്ലോ നാം പ്രധാനമായി പരിഗണിച്ചതു്. ഈ ആകലത്തിൻെറ ലഘുയാദൃച്ഛികസാമ്പിളനം, ആനുപാതികവിഭജനത്തോടു കൂടിയ സ്തരിത സാമ്പിളനം, ഇഷ്ടതമവിഭജനത്തോടു കൂടിയ ആദ്യത്തെ തരം സ്തരിതസാമ്പിളനം എന്നിവയിലെ പ്രസരണങ്ങൾ യഥാക്രമം V_R, V_P, V_O എന്നീ ചിഹ്നങ്ങൾ കൊണ്ടു സൂചിപ്പിക്കുന്നു എന്നിരിക്കട്ടെ. സാന്തസമഷ്ടീ സംശോധനം അവഗണിച്ചാൽ ഇവയുടെ മൂല്യങ്ങൾ,

$$V_R = \frac{S^2}{n}$$

$$V_P = \sum_{h=}^{k} \frac{N_h S^2_h}{nN}$$

$$V_O = \frac{\left(\sum_{h=1}^{k} N_h S_h\right)^2}{nN^2}$$ എന്നിവയാണ്.

ഇവ തമ്മിൽ താരതമ്യപ്പെടുത്തി നോക്കാം. S^2 ൻെറ നിർവചനത്തിൽ നിന്നു്,

$$(N-1)S^2 = \sum_{h=1}^{k} \sum_{i=1}^{n_h} (X_{hi}-\overline{X})^2$$

$$= \sum_h \sum_i (X_{hi}-\overline{X}_h+\overline{X}_h-\overline{X})^2$$

$$= \sum_h \sum_i (X_{hi}-\overline{X}_h)^2 + \sum_h N_h(\overline{X}_h-\overline{X})^2$$

$$= \sum_h (N_h-1)S^2_h + \sum_h N_h(\overline{X}_h-\overline{X})^2$$

$\frac{1}{N_h}$ വളരെ ചെറുതായതു കൊണ്ടു് അവഗണിക്കാമെന്നു സങ്കൽപിച്ചാൽ,

$$NS^2 = \sum_h N_h S^2_h + \sum_h N_h (X_h-\overline{X})^2$$ എന്നുവരും.

അതു കൊണ്ട്',

$$V_R = \frac{S^2}{n} = \frac{\sum_h N_h S^2_h}{nN} + \frac{\sum_h N_h (\bar{X}_h - \bar{X})^2}{nN}$$

$$= V_p + \sum_h \frac{N_h (\bar{X}_h - \bar{X})^2}{nN} \qquad 4.6.1$$

ഇതിൽ നിന്ന് $V_R \geqslant V_P$ എന്നു സിദ്ധിക്കുന്നു.
മാത്രമല്ല,

$$V_p - V_o = \frac{1}{nN} \left\{ \sum_h N_h S^2_h - \frac{\left(\sum_h N_h S_h\right)^2}{N} \right\}$$

$$= \frac{1}{nN} \sum_h N_h \left\{ S_h - \frac{\sum_h N_h S_h}{N} \right\}^2 \qquad 4.6.2$$

വലതു വശത്തെ വ്യംജകം എപ്പോഴും ധനമായിരിക്കുമല്ലോ. അതു കൊണ്ട് $V_P - V_O \geqslant 0$ എന്നും കിട്ടുന്നു അതായത്, $V_P \geqslant V_O$. അങ്ങനെ, $V_R \geqslant V_P \geqslant V_O$.

ചുരുക്കത്തിൽ സ്തരിതസാമ്പിളനം ലഘുയാദൃച്ഛിക സാമ്പിളനത്തെക്കാൾ സൂക്ഷ്മതയുള്ള ആകലങ്ങൾ നൽകും. സ്തരിതസാമ്പിളനത്തിൽ തന്നെ ആനുപാതികവിഭജന രീതിയിൽ ലഭിക്കുന്ന ആകലങ്ങളെക്കാൾ കൂടുതൽ സൂക്ഷ്മത ഇഷ്ടതമവിഭജനരീതിയിൽ ലഭിക്കുന്നവയുണ്ടാവും. മാത്രമല്ല,

$$V_R = V_O + \frac{\sum_h N_h \left(S_h - \frac{\sum_h N_h S_h}{N} \right)^2}{nN} + \frac{\sum N_h (\bar{X}_h - \bar{X})^2}{nN} \qquad 4.6.3$$

ആണെന്ന് ഈ ചർച്ചയിൽ നിന്ന് വെളിവാകുന്നതു കൊണ്ട് ലഘുയാദൃച്ഛിക സാമ്പളിനത്തിലൂടെ ലഭിക്കുന്ന ആകലത്തിന്റെ പ്രസരണത്തിൽ ഉൾപ്പെട്ടു വരുന്ന രണ്ട് അംശങ്ങൾ ഇഷ്ടതമവിഭജനത്തിലൂടെ ലഭിക്കുന്ന ആകലത്തിന്റെ പ്രസരണത്തിൽ ഉണ്ടായിരിക്കുകയില്ല. അവയിൽ ആദ്യത്തേതു് സ്തരങ്ങളുടെ മാനകവിചലനങ്ങൾ തമ്മിലുള്ള വ്യത്യാസത്തിൽ നിന്നുളവാകുന്നതും മറേറതു് സ്തരങ്ങളുടെ സമാന്തരമാധ്യങ്ങൾ തമ്മിലുള്ള വ്യത്യാസത്തിൽ നിന്നുണ്ടാകുന്നതുമാണെന്ന് മുകളിൽ കൊടുത്തിരിക്കുന്ന സമീകരണത്തിന്റെ വലതു വശത്തെ രണ്ടും മൂന്നും പദങ്ങൾ പരിശോധിച്ചാൽ മനസ്സിലാകും.

ഈ ചർച്ചയിൽ സാന്തസമഷ്ടീസംശോധനം നാം അവഗണിക്കുകയാണ് ചെയ്തതു്. അതു കൂടി കണക്കിലെടുത്താൽ,

$$V_R = V_P + \frac{N-n}{nN(N-1)} \left\{ \sum_{h=1}^{k} N_h (\bar{X}_h - \bar{X})^2 - \frac{1}{N} \sum_{h=1}^{k} (N-N_h) S^2_h \right\} \quad 4.6.4$$

എന്നു വരും. ഇവിടെ ബ്രാക്കറ്റിലുള്ള വ്യംജകം ധനസംഖ്യയല്ലെങ്കിൽ V_R, V_P യെക്കാൾ ചെറുതാണെന്നു വരാവുന്നതാണ്. എല്ലാ സ്തരങ്ങളുടെയും സമാന്തര മാധ്യങ്ങൾ തുല്യമായിരിക്കുമ്പോഴാണ് ഇതു് സംഭവിക്കാവുന്നതു്. മറെറല്ലാ സന്ദർഭങ്ങളിലും, സ്തരങ്ങൾ തീരെ ചെറുതല്ലെങ്കിൽ V_R, V_P യെക്കാൾ വലുതായിരിക്കും.

ഈ ചർച്ചയിൽ നിന്നു് ഇഷ്ടതമവിഭജനരീതിയാണ് നാം പരിഗണിച്ചതിൽ ഏററവും നല്ല സമ്പ്രദായമെന്നു വന്നു ചേരുന്നു. പക്ഷേ അതു സ്വീകരിക്കുന്നതിനു മുൻപായി പ്രായോഗികമായ പലതും പരിഗണിക്കേണ്ടതുണ്ട്. ആനുപാതികസമ്പ്രദായത്തെക്കാൾ സൂക്ഷ്മത ഇതിനു കൂടിയിരിക്കുമെന്നുള്ളതു് ശരിയാണ്. പക്ഷേ എത്ര കൂടുതലുണ്ടെന്നുള്ളതു് ഓരോ സമഷ്ടിയുടെയും സ്വഭാവത്തെ ആശ്രയിച്ചാണു് ഇരിക്കുന്നതു്. സൂക്ഷ്മതയിലുണ്ടാവുന്ന നേട്ടം സാമ്പിളന സമ്പ്രദായത്തിന്റെ സങ്കീർണ്ണതയുമായി താരതമ്യപ്പെടുത്തുമ്പോൾ നിസ്സാരമാണെങ്കിൽ ആനുപാതികവിഭജനരീതി തന്നെ സ്വീകരിക്കുന്നതാണ് ബുദ്ധിപൂർവകം. അതു പോലെ തന്നെ സ്തരങ്ങളുടെ പ്രസരണം മിക്കവാറും അജ്ഞാതമായിരിക്കുമെന്നതു കൊണ്ടു് അവയുടെ ഏകദേശനങ്ങൾ മറേറതെങ്കിലും വിധത്തിൽ കണ്ടുപിടിക്കേണ്ടതായി വരും. അങ്ങനെയുള്ള ഏകദേശനങ്ങളിൽ വരുന്ന പിശകുകൾ ആകലത്തിന്റെ സൂക്ഷ്മതയെ ബാധിച്ചെന്നു വരും. അതുകൊണ്ടു് ലഭ്യമായ ഏകദേശനങ്ങളുടെ സൂക്ഷ്മത സംശയാസ്പദമായിരിക്കുന്ന സാഹചര്യങ്ങളിൽ ഈ സമ്പ്രദായം (ആനുപാതികം) സ്വീകരിക്കാതിരിക്കുകയാണ് ഭേദം.

7. സാമ്പിൾ പരിമാണം

ഇതു വരെ നാം പരിഗണിച്ചതു് സാമ്പിൾപരിമാണം n ആണെങ്കിൽ അതിൽ എത്ര വീതം ഓരോ സ്തരത്തിൽ നിന്നു ആയിരിക്കണം എന്നു നിശ്ചയിക്കുവാനുള്ള മാർഗങ്ങളാണ്. അടുത്തതായി ചർച്ച ചെയ്യാനുള്ളതു് n എത്രയായിരിക്കണമെന്നുള്ള പ്രശ്നമത്രെ. പക്ഷേ ഈ പ്രശ്നം ഭംഗ്യന്തരേണ നേരത്തെ തന്നെ ചർച്ച ചെയ്തിട്ടുണ്ട്. ഉദാഹരണമായി S_h കളുടെയും $\bar{C}$ ന്റെയും മൂല്യം തന്നിരുന്നാൽ ഇഷ്ടതമവിഭജനസമ്പ്രദായത്തിൽ n_h കളുടെ മൂല്യം നിശ്ചയിക്കാൻ കഴിയും. $\sum_{h=1}^{k} n_h$ ആണല്ലോ n. അങ്ങനെ n ന്റെ മൂല്യവും നിർണയിക്കപ്പെടുന്നു. ഇതു

പോലെ മററു വിഭജനരീതികളോടു ബന്ധപ്പെട്ടും ഈ വിഷയം ചർച്ച ചെയ്യപ്പെട്ടുകയുണ്ടായി. പക്ഷേ ഇവിടെ അല്പം വ്യത്യസ്തമായ വീക്ഷണകോണത്തിലൂടെ ഈ പ്രശ്നത്തെ സമീപിക്കാനാണ് ശ്രമിക്കുന്നതു്.

$\bar{X}$ ആകലനം ചെയ്യുകയാണ് ലക്ഷ്യം. $\bar{x}_{st}$ ആണ് ആകലമായി സ്വീകരിക്കുന്നതു്. ഈ ആകലത്തിൻെറ മൂല്യം $\bar{X}$ ൻെറ യഥാർഥ മൂല്യത്തിൽ നിന്നു് വ്യത്യസ്തമായിരിക്കാൻ പാടില്ലായ്കയില്ല. മാറിമാറി സാമ്പിളെടുത്താൽ ഓരോന്നിൽ നിന്നും കിട്ടുന്ന മൂല്യങ്ങൾ ഓരോന്നായിരിക്കുകയും ചെയ്യും. ഈ പ്രകീർണനം അളക്കുന്നതു് $\bar{x}_{st}$ ൻെറ സാമ്പിളനപ്രസരണമാണ്, സാമ്പിൾ പരിമാണം വർധിക്കുമ്പോൾ $\bar{x}_{st}$ ൻെറ പ്രസരണം കുറയുമെന്നു് നാം കണ്ടുകഴിഞ്ഞു. അതായതു്, n വലുതാകുമ്പോൾ $\bar{x}_{st}$ ൻെറ മൂല്യങ്ങൾ $\bar{X}$ ൻെറ പരിസരത്തിൽ കേന്ദ്രീകരിക്കും. $\bar{x}_{st}$ ഉം $\bar{X}$ ഉം തമ്മിലുള്ള വ്യത്യാസം തന്നിരിക്കുന്ന (d എന്നിരിക്കട്ടെ) ഒരു സംഖ്യയെക്കാൾ കുറവായിരിക്കുവാനുള്ള സംഭാവ്യത നിർദിഷ്ടമായ ഒരു സംഖ്യയെക്കാൾ കുറവായിരിക്കത്തക്കവണ്ണം n ൻെറ മൂല്യം നിർണയിക്കുകയാണ് നമ്മുടെ പ്രശ്നം. അതായതു്,

$$P_r \{ | \bar{x}_{st} - \bar{X} | \leq d \} = \alpha \quad (\alpha \text{ തന്നിരിക്കുന്നു})$$

ആയിരിക്കത്തക്കവണ്ണം n ൻെറ മൂല്യം നിശ്ചയിക്കണം. $\dfrac{\bar{x}_{st} - \bar{X}}{\sqrt{V(\bar{x}_{st})}}$ മാനകനോർമൽ വിതരണം അനുവർത്തിക്കുന്നു എന്നു സങ്കല്പിച്ചാൽ,

$$P_r \left\{ \left| \frac{\bar{x}_{st} - \bar{X}}{\sqrt{V(\bar{x}_{st})}} \right| \geq t_\alpha \right\} = \alpha$$

ആയിരിക്കത്തക്കവണ്ണം t_α യുടെ മൂല്യം നോർമൽ സമാകലപട്ടികയിൽ നിന്നു് നിർണയിക്കാൻ കഴിയും. അപ്പോൾ നിർദിഷ്ടമായ നിബന്ധനയനുസരിച്ച്,

$d = t_\alpha \sqrt{V(\bar{x}_{st})}$ എന്നു വരുന്നു. അതായതു്, $V(\bar{x}_{st}) = \dfrac{d^2}{t_\alpha^2}$. സൗകര്യത്തിനായി $\dfrac{d^2}{t_\alpha^2} = V$ എന്നിരിക്കട്ടെ. 4.3.4 അനുസരിച്ച്

$$V(\bar{x}_{st}) = \sum_{h=1}^{k} \frac{N_h^2}{N^2} \left(\frac{N_h - n_h}{N_h} \right) \frac{S_h^2}{n_h}$$

$$= \frac{1}{N^2} \sum_{h=1}^{k} \frac{N_h^2}{n_h} S_h^2 - \frac{1}{N^2} \sum_{h=1}^{k} N_h S_h^2$$

അപ്പോൾ n ൻെറ മൂല്യം നിശ്ചയിക്കേണ്ടതു്,

$$V = \frac{1}{N^2} \sum_{h=1}^{k} \frac{N^2_h}{n_h} S^2_h - \frac{1}{N^2} \sum_{h=1}^{k} N_h S^2_h$$ എന്ന സമീകര

ണത്തിൽ നിന്നാണ്. ഇതിൽ $\frac{N_h}{N} = W_h$ എന്നും $\frac{n_h}{n} = w_h$ എന്നും എഴുതി ലഘൂകരിച്ചാൽ ഇതു്,

$$n = \frac{\sum_{h=1}^{k} \frac{W^2_h S^2_h}{w_h}}{V + \frac{1}{N} \sum_{h=1}^{k} W_h S^2_h} \qquad 4.7.1$$

എന്നാവും. ഇവിടെ $w_h = \frac{n_h}{n}$, നാം അവലംബിക്കാൻ പോകുന്ന വിഭജനരീതിയനുസരിച്ചു് നിർണയിക്കണം. S_h^2 ന്റെ മൂല്യം അറിഞ്ഞുകൂടാതെ വരുമ്പോൾ അതിൻെറ ആകലങ്ങൾ പ്രതിസ്ഥാപനം ചെയ്യുക. അങ്ങനെ ഇതിൽ നിന്നു് n ൻെറ മൂല്യം നിശ്ചയിക്കാം. സാന്തസമഷ്ടീസംശോധനം അവഗണിക്കുകയാണെങ്കിൽ,

$$n = \frac{1}{V} \sum_{h=1}^{k} \frac{W^2_h S^2_h}{w_h}$$

എന്നാവും. ഇതു് n ൻെറ ഒരു ഏകദേശനമാണ്. $\bar{n}$ എന്നു് ഇതിനെ അങ്കനം ചെയ്യുന്നു എന്നിരിക്കട്ടെ. അപ്പോൾ,

$$n = \frac{\bar{n}}{1 + \frac{1}{NV} \sum_{h=1}^{k} W_h S^2_h} \qquad 4.7.2$$

ആനുപാതികവിഭജനസമ്പ്രദായമാണ് സ്വീകരിക്കുന്നതെങ്കിൽ ഈ വാക്യം കൂടുതൽ ലളിതമാവും. $w_h = W_h = \frac{N_h}{N}$ ആണല്ലൊ. അതു കൊണ്ടു്

$$\bar{n} = \frac{1}{V} \sum_{h=1}^{k} W_h S^2_h$$ ആവും. അങ്ങനെ,

$$n = \frac{\frac{1}{V} \sum_{h=1}^{k} W_h S^2_h}{1 + \frac{1}{NV} \sum_{h=1}^{k} W_h S^2_h} = \frac{\bar{n}}{1 + \frac{\bar{n}}{N}} \qquad 4.7.3$$

8. അനുപാതത്തിൻെറ ആകലനം

ഒരു പ്രത്യേകതരത്തിൽ പെട്ട അംഗങ്ങളുടെ അനുപാതം പലപ്പോഴും ആകലനം ചെയ്യേണ്ടതായി വരാറുണ്ട്. അതിനും സ്തരിതസാമ്പിളനം ഉപയോഗിക്കാം. ഉദാഹരണമായി ഒരു രാജ്യത്തെ പ്രായപൂർത്തിയായ പൗരന്മാരിൽ എന്തനുപാതം പുകവലി ശീലിച്ചിട്ടുള്ളവരാണ് എന്ന് കണ്ടുപിടിക്കേണ്ടിയിരിക്കുന്നു എന്നിരിക്കട്ടെ. സ്വാഭാവികമായും വരുമാനവും പുകവലിയും ബന്ധപ്പെട്ടിരിക്കുമല്ലോ. അപ്പോൾ സമഷ്ടിയെ വരുമാനത്തിൻെറ അടിസ്ഥാനത്തിൽ വിവിധ സ്തരങ്ങളായി വിഭജിച്ച് സാമ്പിളെടുക്കുന്നതായിരിക്കും അഭികാമ്യം. ചുരുക്കത്തിൽ, അനുപാതങ്ങൾ ആകലനം ചെയ്യാനും സ്തരിതസാമ്പിളനം പ്രയോജനകരമായി ഉപയോഗിക്കാവുന്നതാണ്. മുൻപെന്നതു പോലെ സ്തരങ്ങളിലെ അംഗങ്ങൾ കഴിവതും ഏകാത്മകങ്ങളായിരിക്കാനാണ് ഇവിടെയും ശ്രദ്ധിക്കേണ്ടതു്.

A എന്ന പ്രത്യേകതയുള്ള അംഗങ്ങളുടെ അനുപാതമാണ് ആകലനം ചെയ്യേണ്ടതെന്നു കരുതുക. സമഷ്ടിയിൽ ഈ പ്രത്യേകതയുള്ളതായി M അംഗങ്ങളാണ് ഉള്ളതെന്നിരിക്കട്ടെ. $P=\frac{M}{N}$ ആണ് നമുക്ക് ആകലനം ചെയ്യേണ്ട അനുപാതം. h-ാം സ്തരത്തിൽ M_h അംഗങ്ങളും h-ാം സ്തരത്തിൽ നിന്ന് സാമ്പിളിൽ ഉൾപ്പെട്ടത്തിയവയിൽ m_h അംഗങ്ങളും ഈ പ്രത്യേകതയുള്ളവയാണെന്നു സങ്കൽപിക്കുക. $P_h=\frac{M_h}{N_h}$, $p_h=\frac{m_h}{n_h}$ എന്നു് അങ്കനം ചെയ്യാം.

$$p_{st} = \sum_{h=1}^{k} \frac{N_h p_h}{N} \tag{4.8.1}$$

ആണ് p യുടെ ആകലമായി സ്വീകരിക്കപ്പെടുന്നതു്. A എന്ന പ്രത്യേകതയുള്ള അംഗങ്ങളുടെ അഭിലക്ഷണമൂല്യം 1-ഉം മററുള്ളവയുടേതു് 0 വും ആണെന്നു സങ്കൽപിച്ചാൽ p, സമഷ്ടിയുടെ സമാന്തരമാധ്യമായ $\overline{X}$ ആണ്. p_{st} മുൻപ് സമഷ്ടിയുടെ സമാന്തരമാധ്യത്തിൻെറ ആകലമായി നിർദേശിച്ച $\bar{x}_{st}$ അത്രെ. അതു കൊണ്ട് അതോടനുബന്ധിച്ച് വ്യുൽപാദിപ്പിച്ച പ്രതീക്ഷ, പ്രസരണം തുടങ്ങിയവയെ ആവശ്യമായ പ്രതിസ്ഥാപനങ്ങളോടു കൂടി ഇവിടെയും സ്വീകരിക്കാം. p_{st}, P യുടെ അനഭിനത ആകലമാണെന്നു് ഇതിൽ നിന്നു സിദ്ധിക്കുന്നു. h-ാമത്തെ സ്തരത്തെ ഒരു സമഷ്ടിയായി പരിഗണിച്ചാൽ അതിൻെറ പ്രസരണമാണല്ലോ S^2_h. 3-ാമധ്യായത്തിൽ തെളിയിച്ചിട്ടുള്ളതു പോലെ,

$$S^2_h = \frac{N_h}{N_h-1} P_h Q_h \qquad (Q_h = 1 - P_h)$$

അതുകൊണ്ട്,

$$V(p_{st}) = V(\bar{x}_{st}) = \frac{1}{N^2} \sum_{h=1}^{k} \frac{N^2_h (N_h - n_h) P_h Q_h}{(N_h - 1)\, n_h} \tag{4.8.2}$$

സാന്തസമഷ്ടീസംശോധനം അവഗണിച്ചാൽ

$$V(p_{st}) = \frac{1}{N^2}\sum_{h=1}^{k}\frac{N^2_h P_h Q_h}{n_h} = \sum_{h=1}^{k}\frac{W^2_h P_h Q_h}{n_h} \qquad 4.8.3$$

സാന്തസമഷ്ടീസംശോധനം അവഗണിക്കാവതല്ലെങ്കിൽ തന്നെ $\frac{1}{N_h}$ ഉള്ള പദങ്ങൾ അവഗണിച്ചാൽ,

$$V(p_{st}) = \frac{1}{N^2}\sum_{h=1}^{k} N_h (N_h - n_h) \frac{P_h Q_h}{n_h}. \qquad 4.8.4$$

ആനുപാതികവിഭജനരീതിയാണ് അംഗീകരിക്കുന്നതെങ്കിൽ,

$$V(p_{st}) = \frac{N-n}{N} \cdot \frac{1}{nN}\sum_{h=1}^{k}\frac{N^2_h P_h Q_h}{N_h - 1}.$$

$$\approx \frac{N-n}{nN}\sum_{h=1}^{k} W_h P_h Q_h \qquad 4.8.5$$

P_h ന്റെ മൂല്യം മിക്കവാറും അജ്ഞാതമായിരിക്കും. അങ്ങനെയുള്ള ഘട്ടങ്ങളിൽ p_h നെ P_h ന്റെ ആകലമായി സ്വീകരിച്ച് മുൻ സമീകരണങ്ങളിൽ പ്രതിസ്ഥാപിക്കാവുന്നതാണ്.

വിവിധ വിഭജനരീതികൾ മുൻപു ചർച്ച ചെയ്തതു് ഇവിടെയും സംഗതമത്രെ. അവയോട്ട് ബന്ധപ്പെടുത്തി വ്യുൽപാദിപ്പിച്ച സമീകരണങ്ങളിൽ S_h^2 ന് പകരം $\frac{N_h}{N_h - 1} P_h Q_h$ എന്നു പ്രതിസ്ഥാപിക്കണമെന്നു മാത്രം. അതുപോലെ തന്നെ വിശ്വാസ്യതാന്തരാളങ്ങളെ സംബന്ധിച്ച സമീകരണങ്ങളും സാമ്പിൾ പരിമാണ വ്യംജകവും ഈ പ്രതിസ്ഥാപനത്തോടു കൂടി സ്വീകരിക്കാവുന്നതാണ്. ഉദാഹരണമായി ചെലവ് ക്ലിപ്തപ്പെടുത്തിക്കൊണ്ടുള്ള ഇഷ്ടതമവിഭജനത്തിൽ

$$n_h = \frac{(\bar{C} - C_o)\frac{W_h S_h}{\sqrt{C_h}}}{\sum_{h=1}^{k} W_h S_h \sqrt{C_h}}$$

എന്ന സമീകരണത്തിൽ S_h ന്റെ മൂല്യം പ്രതിസ്ഥാപനംചെയ്യുകയും,

$$n = \frac{(C - C_o) \sum_{h=1}^{k} \frac{W_h S_h}{\sqrt{C_h}}}{\sum_{h=1}^{k} W_h S_h \sqrt{C_h}}$$

എന്നതു് ഉപയോഗിക്കുകയും ചെയ്യുമ്പോൾ,

$$n_h \approx \frac{n N_h \sqrt{P_h Q_h / C_h}}{\sum_{h=1}^{k} N_h \sqrt{P_h Q_h / C_h}} \qquad 4.8.6$$

ഇവിടെ $\frac{N_h}{N_h - 1} \approx 1$ ആണെന്നു് കരുതുന്നു. ഇതു പോലെ തന്നെ മററു സമീകരണങ്ങളും വ്യുൽപാദിപ്പിക്കാവുന്നതാണു്.

ആനുപാതിക വിഭജനത്തിലും ഇഷ്ടതമവിഭജനത്തിലും ലഭിക്കുന്ന ആകലങ്ങളുടെ സൂക്ഷ്മത താരതമ്യപ്പെടുത്തിയിട്ടാണല്ലോ അവയിൽ ഏതാണു് സ്വീകരിക്കേണ്ടതെന്നു് തീരുമാനിക്കുന്നതു്. വിവരശേഖരണത്തിന്റെ ചെലവു് എല്ലാ സ്തരങ്ങളിലും തുല്യമാണെന്നു സങ്കല്പിക്കുക. P_h ന്റെ മൂല്യങ്ങൾ 0.1 നും 0.9 നും മധ്യേയാണെന്നും ഇരിക്കട്ടെ. സാന്തസമഷ്ടിസംശോധനം അവഗണിക്കാമെന്നു സങ്കല്പിച്ചാൽ,

$$V_p \approx \frac{\sum_{h=1}^{k} W_h P_h Q_h}{n}, \quad V_o \approx \frac{\left(\sum_{h=1}^{k} W_h \sqrt{P_h Q_h} \right)^2}{n}$$

എന്നിവയായിരിക്കുമല്ലോ. മുൻപറഞ്ഞ സാഹചര്യത്തിൽ ഇവ തമ്മിലുള്ള വ്യത്യാസം താരതമ്യേന നിസ്സാരമാണെന്നു കാണാം. അതുകൊണ്ടു് മിക്കവാറും സന്ദർഭങ്ങളിൽ ഇഷ്ടതമവിഭജനസമ്പ്രദായം സ്വീകരിച്ചാലുണ്ടാകുന്ന നേട്ടം ആനുപാതികവിഭജനം സ്വീകരിക്കുന്നതിലുള്ള പ്രായോഗിക സൗകര്യവുമായി താരതമ്യപ്പെടുത്തുമ്പോൾ നിസ്സാരമാണു്. അതുകൊണ്ടു് അനുപാതമാണു് ആകലനം ചെയ്യേണ്ടതെങ്കിൽ ആനുപാതികവിഭജനരീതി സ്വീകരിക്കുന്നതാണു് നല്ലതു്.

9. സ്തരങ്ങളുടെ എണ്ണവും അവയെ നിർവചിക്കുന്ന വിധവും

എത്ര സ്തരങ്ങൾ വേണമെന്നു നിശ്ചയിക്കുന്നതും ഓരോ സ്തരത്തിലും സമഷ്ടിയിലെ ഏതെല്ലാം അംഗങ്ങളെ ഉൾപ്പെടുത്തണമെന്നു തീരുമാനിക്കുന്നതും പ്രയാസമേറിയ പ്രശ്നങ്ങളാണു്. വ്യക്തമായ ഒരു നടപടിക്രമം ഇതിനായി നിർദേശിക്കാനാവില്ല. ചില പൊതു തത്വങ്ങളും മാർഗദർശകസിദ്ധാന്തങ്ങളും ആവി

ഷ്ക്കരിക്കാമെന്നു മാത്രം. സാമ്പിളനം ആസൂത്രണം ചെയ്യുന്ന ആളിൻെറ സഹജംവബോധവും പരിചയസമ്പത്തും തീരുമാനങ്ങളെടുക്കാനുള്ള കഴിവും എല്ലാം ഇതിൻെറ വിജയത്തിൽ കാര്യമായ സ്വാധീനം ചെലുത്തുന്നുണ്ട്.

സ്തരിതസാമ്പിളനത്തിൽ നിന്നു ലഭിക്കുന്ന ആകലത്തിൻെറ പ്രസരണം ലഘുയാദൃച്ഛിക സാമ്പിളനത്തിൽ നിന്നു് ലഭിക്കുന്നതിൻേറതിനെക്കാൾ ചെറുതായിരിക്കുമെന്നു് മുമ്പു തെളിയിച്ചിട്ടുണ്ടല്ലൊ. ഒരു സ്തരത്തെ വീണ്ടും വിഭജിക്കുന്നു എന്നിരിക്കട്ടെ. ആദ്യത്തെ സ്തരത്തിൽ നിന്നു് ലഘുയാദൃച്ഛിക സാമ്പിളന രീതിയനുസരിച്ചാണല്ലൊ സാമ്പിൾ എടുക്കുന്നതു്. അതുകൊണ്ടു് ഈ വിഭജനത്തിൻെറ ഫലമായി ലഭിക്കുന്ന രണ്ടു സ്തരങ്ങളിൽ നിന്നും ലഘുയാദൃച്ഛിക സാമ്പിളന രീതിയനുസരിച്ചു് അംഗങ്ങളെ തിരഞ്ഞെടുക്കുമ്പോൾ നമുക്കു കിട്ടുന്നതു് ഒരു സ്തരിതസാമ്പിളാണു്. അതിൽ നിന്നു് ഗണിക്കുന്ന ആകലത്തിൻെറ പ്രസരണം ആദ്യത്തേതിൽ നിന്നു് ഗണിക്കുന്ന ആകലത്തിൻെറ പ്രസരണത്തെക്കാൾ ചെറുതായിരിക്കും. ചുരുക്കത്തിൽ സ്തരങ്ങളുടെ എണ്ണം വർധിച്ചാൽ ആകലങ്ങളുടെ പ്രസരണം കുറയുകയും അങ്ങനെ സൂക്ഷ്മത വർധിക്കുകയും ചെയ്യും.

പക്ഷെ ഇതിൻെറ ഫലമായുണ്ടാകുന്ന അസൗകര്യങ്ങൾ കൂടി കണക്കിലെടുക്കേണ്ടതാണു്. ഏററവും പ്രധാന ബുദ്ധിമുട്ടു് ഓരോ സ്തരത്തിൻെറയും പ്രസരണം നിർണയിക്കണമെന്നുള്ളതാണു് സ്തരങ്ങളുടെ എണ്ണം വർധിക്കുന്തോറും ഇവയുടെ എണ്ണവും വർധിക്കും. അതുപോലെ തന്നെ വിഭജനത്തിലും ആകലനത്തിലും കൈകാര്യം ചെയ്യേണ്ട വ്യംജകങ്ങളുടെ സങ്കീർണതയും ഗണിതക്രിയകളുടെ ബുദ്ധിമുട്ടും വർധിക്കും. ഇതു കൂടി കണക്കിലെടുത്ത ശേഷം വേണം സ്തരങ്ങളുടെ എണ്ണം വർധിപ്പിക്കണമോ എന്നു തീരുമാനിക്കുവാൻ സൂക്ഷ്മതയിലുണ്ടാകുന്ന നേട്ടം ഗണനീയമല്ലാതായിത്തീരുന്ന ഘട്ടത്തിൽ സ്തരങ്ങളുടെ പുനർവിഭജനം അവസാനിപ്പിക്കുക എന്നതാണു് പൊതുവായി സ്വീകരിക്കാവുന്ന നടപടിക്രമം.

ഈ വിഷയത്തെപ്പററി ഗവേഷണങ്ങൾ ഇപ്പോഴും നടന്നുകൊണ്ടിരിക്കുന്നുണ്ടു്. പല സിദ്ധാന്തങ്ങളും ആവിഷ്ക്കരിക്കപ്പെട്ടിട്ടുമുണ്ടു്. ഈ ഗ്രന്ഥത്തിൻെറ പരിധിക്കപ്പുറമായതു കൊണ്ടു് ആ വിഷയം ഇവിടെ സ്പർശിക്കുന്നില്ല.

ഓരോ സ്തരത്തിലും ഉൾപ്പെടുത്തേണ്ട അംഗങ്ങൾ ഏതെല്ലാമെന്നു് തീരുമാനിക്കുക പ്രയാസമേറിയ മറെറാരു പ്രശ്നമാണു്. കാർഷികവും സാമൂഹ്യവുമായ അന്വേഷണങ്ങളിൽ ഭൂമിശാസ്ത്രപരമായ അട്ടപ്പം സ്തരങ്ങൾ നിർവചിക്കാനായി ഉപയോഗപ്പെടുത്താറുണ്ടു്. ഉദാഹരണമായി കൃഷിഭൂമികളുടെ ഉല്പാദനക്ഷമതയെപ്പററിയുള്ള അന്വേഷണത്തിൽ അടുത്തടുത്തുള്ള കൃഷിസ്ഥലങ്ങൾ ഒരേ സ്തരത്തിൽ ഉൾപ്പെടുത്തുന്നതിൽ വലിയ അപാകതയില്ല. പക്ഷെ സാധാരണയായി സ്തരണത്തിനു് അടിസ്ഥാനമായി സ്വീകരിക്കുന്നതു് ചില പൂരക അഭിലക്ഷണങ്ങളുടെ മൂല്യങ്ങളാണു്. ഉദാഹരണമായി വിദ്യാർഥികളുടെ രാഷ്ട്രീയചിന്താഗതിയെപ്പററിയുള്ള ഒരന്വേഷണത്തിൽ അവരെ വയസ്സിൻെറ അടിസ്ഥാനത്തിലോ പഠിക്കുന്ന ക്ലാസ്സിൻെറ അടിസ്ഥാനത്തിലോ സ്തരണം ചെയ്യുന്നതായിരിക്കും കൂടുതൽ

നല്ലതു്. ഇവിടെ വയസ്സ്, ക്ലാസ് തുടങ്ങിയവ പൂരക അഭിലക്ഷണങ്ങളാണ്. ഇവിടെയും അന്വേഷകന്റെ പരിചയസമ്പത്തും ലഭ്യമായ വിവരങ്ങളുടെ സ്വഭാവവുമാണ് നിർണായകഘടകങ്ങൾ.

10. സംഗ്രഹം

ഈ അധ്യായത്തിൽ ചർച്ച ചെയ്ത പ്രധാന വിഷയങ്ങളും വ്യുൽപാദിപ്പിച്ച വ്യംജകങ്ങളും സംഗ്രഹിച്ചു് താഴെ കൊടുക്കുന്നു.

1. വിവിധാത്മകമായ ഒരു സമഷ്ടിയിൽ നിന്നു് സാമ്പിളെടുക്കേണ്ടതു് ആവശ്യമായി വരുമ്പോൾ, സമഷ്ടിയെ കഴിവതും ഏകാത്മകങ്ങളായ സ്തരങ്ങളായി വിഭജിച്ചു്, ഓരോ സ്തരത്തിൽ നിന്നും ആവശ്യമുള്ളത്ര അംഗങ്ങളെ സാമ്പിളിൽ ഉൾപ്പെടുത്തുന്നതു് ആകലങ്ങളുടെ സൂക്ഷ്മത വർധിപ്പിക്കാൻ സഹായിക്കും. ഓരോ സ്തരത്തിൽ നിന്നും സാമ്പിളിൽ ഉൾപ്പെടുത്തേണ്ട അംഗങ്ങളെ ലഘുയാദൃച്ഛിക സാമ്പിളനമനുസരിച്ചാണ് തിരഞ്ഞെടുക്കുന്നതെങ്കിൽ, ആ സാമ്പിളന രീതിക്കു് സ്തരിത യാദൃച്ഛിക സാമ്പിളനം എന്നു പേർ പറയും.

2. സമഷ്ടിയെ k സ്തരങ്ങളായി വിഭജിച്ചിരിക്കുന്നു എന്നും ഓരോ സ്തരത്തിലും യഥാക്രമം $N_1, N_2, \cdots N_k$ അംഗങ്ങളാണ് ഉള്ളതെന്നും സങ്കല്പിക്കുക. n അംഗങ്ങളുള്ള സാമ്പിൾ എടുക്കുന്നു. സ്തരങ്ങളിൽ നിന്നു് യഥാക്രമം $n_1, n_2 \cdots n_k$ അംഗങ്ങളെ വീതമാണ് സാമ്പിളിൽ ഉൾപ്പെടുത്തിയതു് എന്നും ഇരിക്കട്ടെ. X_{hi} ആണ്, h–ാമത്തെ സ്തരത്തിലെ i–ാമതു് അംഗത്തിന്റെ അഭിലക്ഷണ മൂല്യം എന്നും x_{hi} ആണ് h–ാമതു് സ്തരത്തിൽ നിന്നു് സാമ്പിളിൽ ഉൾപ്പെടുത്തിയ i–ാമതു് അംഗത്തിന്റെ അഭിലക്ഷണ മൂല്യം എന്നും വിചാരിക്കുക.

$$\bar{X}_h = \frac{\sum_{i=1}^{N_h} X_{hi}}{N_h}$$

$$S^2_h = \frac{1}{N_h - 1} \sum_{i=1}^{N_h} (X_{hi} - \bar{X}_h)^2$$

$$= \frac{1}{N_h - 1} \left\{ \sum_{i=1}^{N_h} X_{hi}^2 - N_h \bar{X}^2_h \right\}$$

$$\bar{X} = \frac{1}{N} \sum_{h=1}^{k} N_h \bar{X}_h$$

$$\bar{x}_h = \frac{1}{n_h} \sum_{i=1}^{n_h} x_{hi}$$

$$\bar{x} = \frac{1}{n} \sum_{h=1}^{k} n_h \bar{x}_h$$

$$s^2_h = \frac{1}{n_h - 1} \sum_{h=1}^{n_h} (x_{hi} - \bar{x}_h)^2$$

$$= \frac{1}{n_h - 1} \left\{ \sum_{h=1}^{n_h} x^2_{hi} - n_h \bar{x}^2_h \right\}$$

എന്നീ അങ്കനങ്ങൾ ഉപയോഗിക്കുന്നു.

3. $\overline{X}$ ന്റെ അനഭിനത ആകലമായി നിർദ്ദേശിക്കപ്പെട്ടിരിക്കുന്നതു് $\bar{x}_{st}$ ആണു്.

$$\bar{x}_{st} = \frac{1}{N} \sum_{h=1}^{k} N_h \bar{x}_h$$

4. $$V(\bar{x}_{st}) = \frac{1}{N^2} \sum_{h=1}^{k} N_h (N_h - n_h) \frac{S^2_h}{n_h}$$

$$= \sum_{h=1}^{k} \left(\frac{1}{n_h} - \frac{1}{N_h} \right) W^2_h S^2_h$$

$$\hat{V}(\bar{x}_{st}) = \sum_{h=1}^{k} \left(\frac{1}{n_h} - \frac{1}{N_h} \right) W^2_h s^2_h$$

$\left[\hat{V}(\bar{x}_{st}), V(\bar{x}_{st}) \right.$ ന്റെ ആകലം $\left. W_h = \frac{N_h}{N} \right]$

5. സാധാരണ ഉപയോഗത്തിലുള്ള ഏതാനും വിഭജനരീതികൾ.

(a) ആനുപാതിക വിഭജന രീതി.

$$n_h = \frac{nN_h}{N}$$

(b) ഇഷ്ടതമ വിഭജന രീതി.

ചെലവു് നിശ്ചിതമായിട്ടുള്ള ഇഷ്ടതമവിഭജനത്തിൽ,

$$n_h = \frac{(\bar{C}-C_o)\,\dfrac{W_hS_h}{\sqrt{C_h}}}{\sum_{h=1}^{k} W_hS_h\sqrt{C_h}}$$

ഇവിടെ $\bar{C}$, അനുവദിക്കാവുന്ന ചെലവ്; C_o, സാമ്പിളനത്തിന്റെ ആസൂത്രണം തുടങ്ങിയ വസ്തുള്ള സാമ്പിൾ പരിമാണവുമായി ബന്ധമില്ലാത്ത ചെലവ്; C_h, h-ാമത്തെ സ്തരത്തിലെ ഒരംഗത്തിൽ നിന്ന് വിവരം ശേഖരിക്കുന്നതിനുള്ള ചെലവ്.

$$C = C_o + \sum_{h=1}^{k} n_h C_h$$

എന്നതു് ചെലവു സൂത്രം. എല്ലാ സ്തരങ്ങളിൽ നിന്നുമുള്ള വ്യക്തികളിൽ നിന്ന് വിവരം ശേഖരിക്കാനുള്ള ചെലവ് തുല്യമാണെങ്കിൽ, അതായതു് $C_1 = C_2 = \cdots = C_k = C$ ആണെങ്കിൽ,

$$n_h = \left(\frac{\bar{C}-C_o}{C}\right)\frac{W_hS_h}{\sum_{h=1}^{k} W_hS_h} = n\,\frac{W_hS_h}{\sum_{h=1}^{k} W_hS_h}$$

S_h ന്റെ മൂല്യം മററു മാർഗങ്ങളുപയോഗിച്ച് കണ്ടുപിടിക്കണം. n തന്നിരുന്നാൽ അതിനെ W_hS_h ന് ആനുപാതികമായി വിഭജിച്ചാൽ n_h-കൾ കിട്ടും.

ആകലത്തിന്റെ സാമ്പിളന പ്രസരണം നിശ്ചിതമാക്കിയുള്ള വിഭജനത്തിൽ,

$$n_h = \frac{1}{\bar{V}}\left(\sum_{h=1}^{k} W_hS_h\sqrt{C_h}\right)\frac{W_hS_h}{\sqrt{C_h}}$$

ഇവിടെ $\bar{V}$, ആകലത്തിന്റെ സാമ്പിളന പ്രസരണത്തിന്റെ ഉച്ചാവധി

ഇതിൽ ആദ്യത്തെ സമ്പ്രദായമാണ് സാധാരണ ഉപയോഗത്തിലുള്ളതു്.

(c) സ്തരങ്ങളിലെ അഭിലക്ഷണമൂല്യങ്ങൾക്ക് ആനുപാതികമായ വിഭജനം.

$$n_h = \frac{n\,N_h\bar{X}_h}{N\bar{X}}$$

(d) പരാസത്തിന്റെ അടിസ്ഥാനത്തിലുള്ള വിഭജനം.

$$n_h = \frac{n\,W_hR_h}{\sum_{h=1}^{k} W_h R_h}$$

ഇവിടെ R_h, h–ാമത്തെ സ്തരത്തിലെ അഭിലക്ഷ്യമൂല്യങ്ങളുടെ പരാസം.

6. $$V_R \geq V_P \geq V_O$$

ഇവിടെ V_R, ലഘുയാദൃച്ഛിക സാമ്പിളനമനുസരിച്ച് എടുത്ത സാമ്പിളിൽ നിന്നു ലഭിച്ച ആകലത്തിന്റെയും V_P, അനുപാത വിഭജന രീതിയിലെടുത്ത സാമ്പിളിൽ നിന്നു ലഭിച്ച ആകലത്തിന്റെയും, V_o, ഇഷ്ടതമവിഭജന രീതിയിലെടുത്ത സാമ്പിളിൽ നിന്നു ലഭിച്ച ആകലത്തിന്റെയും സാമ്പിളന പ്രസരണം.

7. അനുപാതങ്ങളുടെ ആകലനത്തിന് ആനുപാതീക വിഭജന രീതി അവലംബിച്ചാൽ മതിയാവും. സമഷ്ടിയിലെ അനുപാതം P യും h–ാമത് സ്തരത്തിൽ നിന്ന് സാമ്പിളിൽ ഉൾപ്പെടുത്തിയ വ്യക്തികളിലെ അനുപാതം p_h ഉം ആണെങ്കിൽ,

$$p_{st} = \sum_{h=1}^{k} \frac{N_h p_h}{N}$$ ആണ് P യുടെ ആകലം.

$$V(p_{st}) = \sum_{h=1}^{k} \frac{N_h^2}{N^2} \frac{P_h Q_h}{n_h}$$

ഇവിടെ P_h, h–ാമത്തെ സ്തരത്തിലെ അനുപാതം. $Q_h = 1 - P_h$.

$$\hat{V}(p_{st}) \approx \sum_{h=1}^{k} \frac{N_h^2}{N^2} \frac{p_h q_h}{n_h}.$$ ഇവിടെ $p_h = 1 - q_h$

11. ചില ഉദാഹരണങ്ങൾ

ഉദാഹരണം 4.1. താഴെ കൊടുത്തിരിക്കുന്ന പട്ടികയിൽ ഒരു കലാലയത്തിലെ വിവിധ വിഭാഗങ്ങളിൽ പഠിക്കുന്ന വിദ്യാർഥികളുടെ എണ്ണവും, ഓരോ വിഭാഗത്തിൽ പെട്ടവരുടെയും വയസ്സിന്റെ പ്രസരണത്തിന്റെ ആകലങ്ങളും കൊടുത്തിരിക്കുന്നു. വിദ്യാർഥികളുടെ ശരാശരി വയസ്സ് ആകലനം ചെയ്യുകയാണ് ലക്ഷ്യം. ഓരോ വിഭാഗത്തെയും ഓരോ സ്തരമായി പരിഗണിച്ച് ആനുപാതികം, ഇഷ്ടതമം എന്നീ വിഭജനരീതികളവലംബിച്ചാൽ 100 അംഗങ്ങളുള്ള ഒരു സാമ്പിളിൽ ഓരോ വിഭാഗത്തിൽ നിന്നും എത്ര അംഗങ്ങളെ വീതം ഉൾപ്പെടുത്തണമെന്ന് കണ്ടുപിടിക്കാം.

	വിഭാഗം	വിദ്യാർഥികളുടെ എണ്ണം (N_h)	വയസ്സിന്റെ മാനക വിചലനം (S_h)
1.	ബിരുദാനന്തര വിഭാഗം	220	2.3
2.	ബിരുദാർഥി വിഭാഗം	850	6.2
3.	ബിരുദപൂർവ വിഭാഗം	1350	3.4

ഇവിടെ $N = 2420, \quad N_1 = 220, \quad N_2 = 850, \quad N_3 = 1350$
$n = 100$

(1) ആനുപാതിക വിഭജനരീതി.

$$n_h = n \frac{N_h}{N}$$

ഓരോ വിഭാഗത്തിൽ നിന്നും സാമ്പിളിൽ ഉൾപ്പെടുത്തേണ്ട അംഗങ്ങളുടെഎണ്ണം ഗണിക്കുന്ന രീതി താഴെ കൊടുത്തിരിക്കുന്ന പട്ടികയിൽ കാണിച്ചിരിക്കുന്നു.

	വിഭാഗം (സ്തരം)	വിദ്യാർഥികളുടെ എണ്ണം (N_h)	$\frac{N_h}{N}$	സാമ്പിളിൽ ഉൾപ്പെടുത്തേണ്ട അംഗങ്ങൾ $n_h = n\frac{N_h}{N}$
1.	ബിരുദാനന്തരം	220	0.09	9
2.	ബിരുദാർഥി	850	0.35	35
3.	ബിരുദപൂർവം	1350	0.56	56
		2420		100

(2) ഇഷ്ടതമവിഭജനരീതി.

ഓരോ വിഭാഗത്തിൽ നിന്നും വിവരശേഖരണം നടത്തുന്നതിനുള്ള ചെലവ് വ്യത്യസ്തമല്ലാത്തതു കൊണ്ട് ഇവിടെ 100–നെ $W_h S_h$ ന്റെ അനുപാതത്തിൽ വിഭജിച്ചാൽ മതി. അതായത്,

$$n_h = \frac{n\, W_h S_h}{\sum_{h=1}^{k} W_h S_h}$$

ഗണനരീതി താഴെ കൊടുക്കുന്ന പട്ടികയിൽ കാണിച്ചിരിക്കുന്നു. ഇവിടെ $W_h = \frac{N_h}{N}$ ന്റെ മൂല്യം മുൻ പട്ടികയിൽ നിന്നു് എടുക്കാവുന്നതാണ്. ഇവിടെ നിശ്ചിത ചെലവുള്ള ഇഷ്ടതമ നിഭജന രീതിയാണ് പരിഗണിക്കുന്നതു്.

	വിഭാഗം (സ്തരം)	$W_h = \frac{N_h}{N}$	മാനകവിചലനം S_h	സാമ്പിളിൽ ഉൾപ്പെടുത്തേണ്ട അംഗങ്ങൾ $\frac{100 \times W_h S_h}{\Sigma W_h S_h}$
1.	ബിരുദാനന്തരം	0.09	2.3	5
2.	ബിരുദാർഥി	0.35	6.2	50
3.	ബിരുദപൂർവം	0.56	3.4	45

ഉദാ: 4.2.മുൻ ഉദാഹരണത്തിൽ എടുത്ത സാമ്പിളുകളിൽ ഓരോ സ്തരത്തിൽ

നിന്നും തിരഞ്ഞെടുത്ത വിദ്യാർഥികളുടെ വയസ്സിൻെറ സമാന്തരമാധ്യങ്ങളും പ്രസരണവും താഴെ കൊടുക്കുന്നു. വയസ്സുകളുടെ സമാന്തരമാധ്യം ആകലനം ചെയ്യുക. ആകലത്തിൻെറ സാമ്പിളന പ്രസരണവും ആകലനം ചെയ്യുക. വയസ്സുകളുടെ സമാന്തരമാധ്യത്തിൻെറ 95% വിശ്വാസ്യതാന്തരാളങ്ങളും നിർണയിക്കുക.

ആനുപാതികവിഭജനം

വിഭാഗം (സ്തരം)	സാമ്പിളിൽ ഉൾപ്പെട്ടത്തിയ അംഗങ്ങൾ (n_h)	വയസ്സിന്റെ സമാന്തര മാധ്യം ($\bar{x}_h$)	പ്രസരണം (s^2_h)
1. ബിരുദാനന്തരം	9	20.3	6.4
2. ബിരുദാർഥി	35	18.5	40.2
3. ബിരുദപൂർവം	56	16.2	12.1

ഇഷ്ടതമവിഭജനം

വിഭാഗം (സ്തരം)	സാമ്പിളിൽ ഉൾപ്പെട്ടത്തിയ അംഗങ്ങൾ (n_h)	വയസ്സിന്റെ സമാന്തര മാധ്യം ($\bar{x}_h$)	പ്രസരണം (s^2_h)
1. ബിരുദാനന്തരം	5	20.1	6.1
2. ബിരുദാർഥി	50	19.1	35.2
3. ബിരുദപൂർവം	45	16.3	10.5

ഇതിൽ ആദ്യത്തേതിൽ നിന്നു് സമാന്തരമാധ്യം ആകലനം ചെയ്യാൻ $\bar{x}_{st} \Sigma W_h \bar{x}_h$ എന്ന സൂത്രം ഉപയോഗിക്കാം. ഗണനരീതി പട്ടികരൂപത്തിൽ താഴെ കൊടുക്കുന്നു.

വിഭാഗം (സ്തരം)	സ്തരത്തിലെ അംഗസംഖ്യ (N_h)	$W_h = \frac{N_h}{N}$	സാമ്പിളിൽ ഉൾപ്പെട്ടത്തിയ അംഗങ്ങൾ (n_h)	വയസ്സിന്റെ സമാന്തര മാധ്യം ($\bar{x}_h$)	$W_h \bar{x}_h$
1. ബിരുദാനന്തരം	220	0.09	9	20.3	1.827
2. ബിരുദാർഥി	850	0.35	35	18.5	6.475
3. ബിരുദപൂർവം	1350	0.56	56	16.2	9.072
					17.374

ശരാശരി വയസ്സിൻെറ ആകലം = 17.37.

ഈ ആകലത്തിന്റെ പ്രസരണം ആകലനം ചെയ്യുവാൻ

$$\hat{V}(\bar{x}_{st}) = \sum_{h=1}^{3} W^2_h \left(\frac{1}{r_h} - \frac{1}{N_h}\right) s_h^2$$

എന്ന സൂത്രം ഉപയോഗിക്കാം. ഗണനരീതി താഴെ കൊടുത്തിരിക്കുന്ന പട്ടികയിൽ കാണിച്ചിരിക്കുന്നു.

വിഭാഗം	അംഗസംഖ്യ (N_h)	സാമ്പിളിൽപ്പെട്ടത്തിയ അംഗങ്ങൾ (r_h)	s^2_h	$W^2_h s^2_h$	$\left\lvert\frac{1}{n_h}-\frac{1}{N_h}\right\rvert$	$\left\{\frac{1}{r_h}-\frac{1}{n_h}\right\} W^2_h S^2_h$
1. ബിരുദാനന്തം	220	9	6.4	0.0518	0.107	0.00554
2. ബിരുദാർഥി	850	35	40.2	4.9245	0.027	0.13300
3. ബിരുദപൂർവം	1350	56	12.1	3.7946	0.017	0.06451

ആകലത്തിന്റെ പ്രസരണത്തിന്റെ ആകലം } = 0.20305

അടുത്തതായി കാണുവാനുള്ളതു് 95% വിശ്വാസ്യതാന്തരാളമാണു്. നോർമൽ പട്ടികയിൽ നിന്നു് $t_{.95} = 1.96$ എന്നു കാണാം. അതുകൊണ്ടു് 95% വിശ്വാസ്യതാന്തരാളം

$$\left(\bar{x}_{st} - 1.96\sqrt{\hat{V}(\bar{x}_{st})},\ \bar{x}_{st} + 1.96\sqrt{\hat{V}(\bar{x}_{st})}\right)$$

ആണു്. മൂല്യങ്ങൾ പ്രതിസ്ഥാപിക്കുമ്പോൾ ഇതു്, (16.49, 18.25) എന്നായിത്തീരുന്നു.

അടുത്തതായി ഇഷ്ടതമവിഭജനം പരിഗണിക്കാം. സമാന്തരമാധ്യത്തിന്റെ ഗണനം താഴെ കൊടുക്കുന്നു.

വിഭാഗം	$W_h = \frac{N_h}{N}$	സാമ്പിളിലെ അംഗങ്ങൾ (r_h)	സാമ്പിൾ സമാന്തരമാധ്യം ($\bar{x}_h$)	$W_h\bar{x}_h$
1. ബിരുദാനന്തരം	0.09	5	20.1	1.809
2. ബിരുദാർഥി	0.35	50	19.1	6.685
3. ബിരുദപൂർവം	0.56	45	16.3	9.128

ശരാശരി വയസ്സിന്റെ ആകലം = 17.62

ഈ ആകലത്തിൻെറ പ്രസരണവും മുൻസൂത്രമുപയോഗിച്ചു തന്നെ ആകലനം ചെയ്യാം. ആകലനരീതി താഴെ കൊടുക്കുന്നു.

വിഭാഗം	$W_h=\frac{N_h}{N}$	സാമ്പിളിലെ അംഗസംഖ്യ r_h	s^2_h	$W^2_h s^2_h$	$\frac{1}{r_h}-\frac{1}{N_h}$	$\left(\frac{1}{r_h}-\frac{1}{N_h}\right)\times W_h^2 S_h^2$
1. ബിരുദാനന്തരം	0.09	5	6.1	0.049	0.195	0.0096
2. ബിരുദാർഥി	0.35	50	35.2	4.312	0.019	0.0820
3. ബിരുദപൂർവം	0.56	45	10.5	3.293	0.021	0 0691

ആകലത്തിൻെറ പ്രസരണത്തിൻെറ ആകലം $= 0.1607$

വിശ്വാസ്യതാന്തരാളത്തിന്റെ നിർണ്ണയവും മുൻപിലത്തേതു പോലെ തന്നെ നടത്താവുന്നതാണ്. മാനകനോർമൽ പട്ടിക ഉപയോഗിച്ചാൽ 95 % വിശ്വാസ്യതാന്തരാളം, $(\bar{x}_{st}-1.96\sqrt{V(\bar{x}_{st})}\ ,\ \bar{x}_{st}+1.96\sqrt{V(\bar{x}_{st})})$ എന്നു കാണാം. മൂല്യങ്ങൾ പ്രതിസ്ഥാപിക്കുമ്പോൾ ഇതു് (16.84, 18.40) എന്നായിത്തീരും. ഇഷ്ടതമ വിഭജനരീതി അവലംബിച്ചതുകൊണ്ടു് പ്രസരണത്തിൽ വന്ന കുറവും വിശ്വാസ്യതാന്തരാളത്തിൻെറ ദൈർഘ്യക്കുറവും ശ്രദ്ധിക്കുക.

ഉദാ: 4.3. മുൻ ഉദാഹരണങ്ങളിലെ സമഷ്ടിയിൽ നിന്നു് കലാലയത്തിലെ വിദ്യാർഥിനികളുടെ അനുപാതം ആകലനം ചെയ്യുന്നു. അനുപാതവിഭജന രീതിയിൽ എടുത്ത സാമ്പിളിൽ നിന്നു് ലഭിച്ച വിവരങ്ങൾ താഴെ ചേർക്കുന്നു.

വിഭാഗം	സാമ്പിളിൽ ഉൾപ്പെട്ടത്തിയ അംഗങ്ങൾ	വിദ്യാർഥികളുടെ അനുപാതം
1. ബിരുദാനന്തരം	9	0.5
2. ബിരുദാർഥി	35	0.3
3. ബിരുദപൂർവം	56	0.4

സമഷ്ടിയിലെ വിദ്യാർഥിനികളുടെ അനുപാതം ആകലനം ചെയ്യുക. അതിൻെറ സാമ്പിളനപ്രസരണം ആകലനം ചെയ്യുക.

$p_{st}=\Sigma\frac{N_h}{N}p_h$ എന്ന സൂത്രമുപയോഗിച്ചു് വിദ്യാർഥിനികളുടെ എണ്ണം ആകലനം ചെയ്യാം. ആകലനരീതി താഴെ കൊടുക്കുന്നു.

വിഭാഗം	$\frac{N_h}{N}$	p_h	$\frac{N_h}{N}p_h$
1. ബിരുദാനന്തരം	.09	0.5	0.045
2. ബിരുദാർഥി	.35	0.3	0.105
3. ബിരുദപൂർവം	.56	0.4	0.224

വിദ്യാർത്ഥിനികളുടെ അനുപാതത്തിന്റെ ആകലം $= 0.374$

ഇതിന്റെ സാമ്പിളനപ്രസരണത്തിന്റെ ആകലം,

$$\hat{V}(p_{st}) \approx \Sigma \frac{N^2_h}{N^2} \frac{p_h q_h}{n_h}$$ എന്നതാണ്.

ഇതിന്റെ ഗണനം താഴെ കൊടുത്തിരിക്കുന്ന പട്ടികയിൽ കാണിച്ചിരിക്കുന്നു.

വിഭാഗം	$\frac{N^2_h}{N^2}$	$\frac{p_h q_h}{n_h}$	$\frac{N^2_h}{N^2} \frac{p_h q_h}{n_h}$
1. ബിരുദാനന്തരം	0.0081	0.028	0.000227
2. ബിരുദാർത്ഥി	0.1225	0.060	0.007350
3. ബിരുദപൂർവം	0.3136	0.043	0.013485

$$\hat{V}(p_{st}) = 0.021$$

12. ചില സഹായക ഗ്രന്ഥങ്ങളും ലേഖനങ്ങളും

1. Sampling Techniques (Ch. V), Cochran W. G, John Wiley.
2. Sampling Theory of Surveys with Applications (Ch. III), Sukhatme P. V, Indian Society of Agricultural Statistics.
3. Sampling Theory and Methods (Ch. VII), Murthy M. N, Statistical Publishing Society, Calcutta.
4. On two different aspects of the representative method (1934), Neyman J, J. R. S, 97, 558–625.
5. On Stratification and Optimum Allocations (1951), Evans W.D Journal of Amer. Stat. Assn., 46, 95–104
6. A Study of Stratified Random Sampling (1954), Aoyama H, Ann. Inst. Stat. Math., 6, 1–36.

അഭ്യാസം 4

1. ഒരു എസ്റ്റേററിലെ ജോലിക്കാരുടെ താമസസൗകര്യങ്ങളെപ്പററിയുള്ള അന്വേഷണത്തിൽ ഒരു ജോലിക്കാരനും കുടുംബത്തിനും കൂടി ശരാശരി എത്ര ചതുരശ്രയടി വിസ്തീർണമുള്ള താമസസ്ഥലങ്ങൾ ലഭിക്കുന്നുണ്ടെന്ന് ആകലനം ചെയ്യേണ്ടിരിക്കുന്നു. ജോലിക്കാരുടെ ശമ്പളമനുസരിച്ച് അവരെ നാലു വിഭാഗങ്ങളായി തിരിച്ചിരിക്കുന്നു. ഓരോ വിഭാഗത്തിൽ നിന്നും എട്ട് അംഗങ്ങളെ വീതം ലഘുയാദൃച്ഛിക സാമ്പിളനമനുസരിച്ച് തിരഞ്ഞെടുത്ത് അവർക്ക് നൽകപ്പെട്ടിരിക്കുന്ന താമസസ്ഥലത്തിന്റെ വിസ്തീർണം

അളന്നു കണ്ടുപിടിച്ചു് താഴെ കൊടുക്കുന്നു. ഒരാൾക്കു് നൽകപ്പെട്ടിരിക്കുന്ന താമസസ്ഥലത്തിൻെറ ശരാശരി വിസ്തീർണത്തിൻെറ 98% വിശ്വാസ്യ താന്തരാളം കണ്ടുപിടിക്കുക.

സ്തരം	അംഗസംഖ്യ	സാമ്പിളിൽ ഉൾപ്പെടുത്തിയ അംഗങ്ങളുടെ താമസസ്ഥലത്തിന്റെ വിസ്തീർണം (ച.മീ) 1	2	3	4	5	6	7	8
1	235	500	635	420	210	470	520	550	510
2	123	1350	1230	1000	1470	1200	950	1120	870
3	656	128	143	138	167	140	138	150	163
4	45	2570	1900	1870	2200	1700	1650	2100	200

2. മുൻ ഉദാഹരണത്തിൽ അനുപാതവിഭജനരീതിയനുസരിച്ചാണു് സാമ്പിളെടുക്കുന്നതെങ്കിൽ ഓരോ സ്തരത്തിൽ നിന്നും എത്ര അംഗങ്ങളെ വീതം സാമ്പിളിൽ ഉൾപ്പെടുത്തണം ? മുൻ ഉദാഹരണത്തിലെ സാമ്പിളിൽ നിന്നു് ഗണിച്ചെടുത്ത സ്തരങ്ങളിലെ പ്രസരണങ്ങൾ ഉപയോഗിച്ചു്, ഇഷ്ടതമവിഭജനരീതിയിൽ ഓരോ സ്തരത്തിൽ നിന്നും എത്ര അംഗങ്ങളെ വീതം സാമ്പിളിൽ ഉൾപ്പെടുത്തണമെന്നു് നിർണയിക്കുക.

3. ഒന്നാം ചോദ്യത്തിലെ സാമ്പിളിൽ നിന്നും ലഭിക്കുന്ന സ്തരങ്ങളുടെ പ്രസരണം ഉപയോഗിച്ചു്, ലഘുയാദൃച്ഛികസാമ്പിളനഅനുപാതവിഭജനരീതി, ഇഷ്ടതമവിഭജനരീതി എന്നിവയിലെ പ്രസരണങ്ങൾ നിർണയിച്ചു് താരതമ്യപ്പെടുത്തുക.

4. ഒരു ജില്ലയിലെ വിവിധ സമുദായത്തിൽപ്പെട്ട അംഗങ്ങളുടെ എണ്ണവും അവരിൽ എഴുതാനും വായിക്കാനും അറിയാവുന്നവരുടെ ശതമാനത്തെപ്പററിയുള്ള ഏകദേശമായ അറിവും താഴെ കൊടുക്കുന്നു.

സമുദായം	അംഗസംഖ്യ	സാക്ഷരതാ ശതമാനം (ഏകദേശം)
1	40,000	50
2	20,000	90
3	5,000	70

സമഷ്ടിയിലെ സാക്ഷരതാശതമാനം ആകലനം ചെയ്യാൻ ഉദ്ദേശിക്കപ്പെട്ടിട്ടുള്ള 1000 അംഗങ്ങളുള്ള ഒരു സാമ്പിളിൽ ഇഷ്ടതമവിഭജനരീതിയാണു് അവലംബിക്കുന്നതെങ്കിൽ ഓരോ സമുദായത്തിൽ നിന്നും എത്ര അംഗങ്ങളെ വീതം ഉൾപ്പെടുത്തണം ?.

5. അഭ്യാസം 5-ൽ സ്തരണം കൊണ്ടുള്ള നേട്ടം നിർണയിക്കുക.

5

ക്രമാനുസൃത സാമ്പിളനം

1. പ്രാരംഭം

രണ്ടു് സാമ്പിളനരീതികൾ നാം ചർച്ച ചെയ്തു കഴിഞ്ഞു. താത്വികമായി പറഞ്ഞാൽ അതിവിശിഷ്ടങ്ങളാണു് അവ രണ്ടും. പക്ഷെ ഇതിൽ ഏതു രീതി അവലംബിച്ചാലും സാമ്പിളെടുക്കുന്നിടത്തും വിവരം ശേഖരിക്കുന്നിടത്തും അന്വേഷകനു് പല പ്രായോഗികബുദ്ധിമുട്ടുകളും അഭിമുഖീകരിക്കേണ്ടി വരാറുണ്ടു്. സമഷ്ടിയുടെ സ്വഭാവമനുസരിച്ചു് ഈ ബുദ്ധിമുട്ടുകൾക്കു് ഏററക്കുറച്ചിൽ വരാമെന്നു മാത്രം. സാമ്പിളെടുക്കാൻ ശ്രമിക്കുമ്പോൾ അനുഭവപ്പെടുന്ന പ്രധാന ബുദ്ധിമുട്ടു് സമഷ്ടിയിലെ അംഗങ്ങളുടെ പൂർണ്ണമായ ഒരു ലിസ്റ്റ് തയ്യാറാക്കാതെ അതു് സാധ്യമാവുകയില്ല എന്നതാണു്. ഒരു വ്യാപാരസ്ഥാപനം ഒരു വർഷം എത്ര രൂപയുടെ സാധനങ്ങൾ വാങ്ങി എന്നു് സാമ്പിളനരീതിയിൽ ആകലനം ചെയ്യുകയാണു് അന്വേഷണലക്ഷ്യമെന്നിരിക്കട്ടെ. വാങ്ങൽ ബില്ലുകൾ നൂറെണ്ണം വീതമുള്ള ഫയലുകളായി സൂക്ഷിച്ചിരിക്കുന്നു. അങ്ങനെ 600 ഫയലുകളുണ്ടെന്നും സങ്കൽപിക്കുക. ലഘുയാദൃച്ഛികസാമ്പിളനരീതിയിൽ, 6,000 അംഗങ്ങളുള്ള ഒരു സാമ്പിളാണു് എടുക്കാൻ ഉദ്ദേശിക്കുന്നതെന്നിരിക്കട്ടെ. നാം ആദ്യമായി ചെയ്യേണ്ടതു്, ഈ 60,000 ബില്ലുകൾക്കു് ക്രമസംഖ്യകൾ കൊടുത്തു് അവയുടെ ഒരു ലിസ്ററു തയ്യാറാക്കുകയാണ. ഇതു് അത്യന്തം ശ്രമകരമായ ഒരു ജോലിയാണെന്നു പറയേണ്ടതില്ലല്ലോ. സ്തരിത സാമ്പിളാണു് എടുക്കുന്നതെങ്കിലും ഇതു കൂടാതെ കഴികയില്ല.

മററൊരു ബുദ്ധിമുട്ടു് അനുഭവപ്പെടുന്നതു് വിവരം ശേഖരിക്കാൻ ശ്രമിക്കുമ്പോഴാണു്. സാമ്പിളിലേയ്ക്കു തിരഞ്ഞെടുക്കപ്പെട്ട അംഗങ്ങൾ സമഷ്ടിയിൽ ഒരു ക്രമവുമില്ലാതെ ചിതറിക്കിടക്കുന്നവയായിരിക്കും. അവ ഓരോന്നും തേടിപ്പിടിക്കുക എന്നതു് പ്രത്യേകിച്ചും സമഷ്ടി വലുതായിരിക്കുമ്പോൾ, എത്ര സമർത്ഥനായ അന്വേഷകനെയും അടിയറവു പറയിപ്പിക്കുന്ന ഒരു കാര്യമാണു്. മുൻ ഉദാഹരണത്തിൽ തന്നെ ഓരോ ഫയലു പല പ്രാവശ്യം മറിച്ചു നോക്കേണ്ട ആവ

ശ്യം അന്വേഷകനുണ്ടാവും. ഇതുൾക്കൊള്ളുന്ന സമയനഷ്ടവും മനം മടുപ്പും ഊഹിക്കാവുന്നതാണല്ലോ.

ഈ രണ്ടു പ്രായോഗികവൈഷമ്യങ്ങളും ഭാഗികമായെങ്കിലും ലഘൂകരിക്കുന്നതും താത്വികമായി വലിയ പോരായ്മകളൊന്നുമില്ലാത്തതുമായ ഒരു സാമ്പിളനരീതിയാണ് ഈ അധ്യായത്തിൽ ചർച്ച ചെയ്യുന്ന ക്രമാനുസൃതസാമ്പിളനം. ഈ രീതി, ലഘുയാദൃച്ഛികസാമ്പിളനത്തെക്കാൾ കൂടുതൽ സൂക്ഷ്മതയുള്ള വിവരങ്ങൾ നൽകുന്ന സന്ദർഭങ്ങളും ഉണ്ടാകാവുന്നതാണെന്ന് വഴിയേ കാണാം.

2. സാമ്പിളെടുക്കൽ സമ്പ്രദായങ്ങൾ

താരതമ്യേന എളുപ്പത്തിൽ സാമ്പിളെടുക്കാവുന്ന ഒരു സാമ്പിളനരീതിയാണ് ഇത്. 'രേഖീയം' 'ചക്രീയം' എന്നീ രണ്ടു സമ്പ്രദായങ്ങൾ സാമ്പിളെടുക്കലിൽ ഉപയോഗത്തിലുണ്ട്. അവയ്ക്കോരോന്നിനും അതതിന്റേതായ വൈശിഷ്ട്യങ്ങളും പോരായ്മകളും ഉണ്ട് താനും.

1. രേഖീയസമ്പ്രദായം.

N അംഗങ്ങളുള്ള ഒരു സമഷ്ടിയിൽ നിന്ന് n അംശങ്ങളുള്ള സാമ്പിളാണ് എടുക്കേണ്ടതെന്നിരിക്കട്ടെ. N, n ന്റെ ഒരു ഗുണിതമാണെന്നു സങ്കൽപിക്കുക. ഉദാഹരണമായി $N=nk$ എന്നും വിചാരിക്കുക. ആദ്യമായി ചെയ്യേണ്ടതു് 1 മുതൽ k വരെയുള്ള സംഖ്യകളിൽ നിന്ന് യാദൃച്ഛികമായി ഒരു സംഖ്യ തിരഞ്ഞെടുക്കുകയാണ്. "r" എന്ന സംഖ്യയാണ് തിരഞ്ഞെടുക്കപ്പെട്ടതെന്നിരിക്കട്ടെ. അതിന് സാമ്പിളിന്റെ, യാദൃച്ഛികപ്രാരംഭം എന്ന പേർ പറയും. k-ക്ക് സാമ്പിളനഅന്തരാളം എന്നാണ് പേർ. സമഷ്ടിയിലെ അംഗങ്ങൾക്കു ക്രമസംഖ്യകൾ കൊടുത്തിട്ട് $r, r+k, r+2k, \cdots, r+(n-1)k$ എന്നീ ക്രമസംഖ്യകളോടു കൂടിയ അംഗങ്ങളെ സാമ്പിളിൽ ഉൾപ്പെടുത്തിയാൽ നമുക്കു ലഭിക്കുന്നതു് രേഖീയരീതിയിൽ തിരഞ്ഞെടുക്കപ്പെട്ട n അംഗങ്ങളുള്ള ഒരു ക്രമാനുസൃത സാമ്പിളായിരിക്കും. ഇവിടെ യാദൃച്ഛികമായി തിരഞ്ഞെടുക്കപ്പെട്ട r എന്ന സംഖ്യ, സാമ്പിളിനെ സംബന്ധിച്ചിടത്തോളം നിർണായകമാണെന്നുള്ള വസ്തുത പ്രത്യേകം ശ്രദ്ധേയമാണ്.

മുൻവിവരിച്ച വിധത്തിൽ തിരഞ്ഞെടുക്കപ്പെട്ട ഒരു ക്രമീകൃതസാമ്പിൾ സ്തരിതസാമ്പിളിന്റെയും ലഘുയാദൃച്ഛിക സാമ്പിളിന്റെയും ഏതാനും പ്രത്യേകതകൾ ഉൾക്കൊള്ളുന്നതാണെന്നുള്ളതു് പ്രസ്താവ്യമാണ്. സമഷ്ടിയെ k അംഗങ്ങൾ വീതമുള്ള n സ്തരങ്ങളായി വിഭജിച്ച് അവ ഓരോന്നിൽ നിന്നും ഓരോ അംഗത്തെ തിരഞ്ഞെടുക്കുകയാണല്ലോ ചെയ്യുന്നതു്. ഈ അർഥത്തിൽ ഇതു് ഒരു സ്തരിതസാമ്പിളാണ്. എന്നാൽ സമഷ്ടിയെ സ്തരങ്ങളായി വിഭജിക്കുമ്പോൾ പരിഗണിക്കേണ്ട ഏകാത്മകത എന്ന വൈശിഷ്ട്യം ഇവിടെ പരിഗണിക്കപ്പെട്ടിട്ടില്ല. അതായതു്, സ്തരങ്ങളിലെ അംഗങ്ങൾ താരതമ്യേന സാമ്യമുള്ളവയായിരിക്കത്തക്കവണ്ണമല്ല വിഭജനം നടത്തിയിരിക്കുന്നതു്. അതുപോലെ തന്നെ സ്തരിത സാമ്പിളനത്തിൽ ചെയ്യുന്നതു പോലെ ഓരോ സ്തരത്തിൽ നിന്നും യാദൃച്ഛികമായി

ട്ടല്ല തിരഞ്ഞെടുപ്പു നടത്തിയതു്. ആദ്യത്തെ സ്തരത്തിൽ നിന്നുള്ള തിരഞ്ഞെടുപ്പു മാത്രമേ അപ്രകാരം നടത്തിയുള്ളൂ. അങ്ങനെ ഇത്തരം ക്രമാനുസൃതസാമ്പിളനത്തിന്ന് സ്തരിതസാമ്പിളനവുമായുള്ള സാമ്യം ബാഹ്യരൂപത്തെ സംബന്ധിച്ചതു മാത്രമത്രെ. ഇതിന്ന് ലഘുയാദൃച്ഛികസാമ്പിളനവുമായുള്ള സാമ്യം കൂടുതൽ ദൃഢതരമാണെന്നു പറയാവുന്നതാണ്. സമഷ്ടിയിലെ അംഗങ്ങളെ 1, k+1, 2k+1 .....(n−1) k + 1, എന്നീ ക്രമസംഖ്യകളോടു കൂടിയവയുടെ സമുച്ചയം, 2, k+2, 2k+2, ... (n−1) k+2 എന്നീ ക്രമസംഖ്യകളോടു കൂടിയവയുടെ സമുച്ചയം എന്നിങ്ങനെ ക്രമത്തിൽ k സമുച്ചയങ്ങളായി വിഭജിക്കാം. അവയിൽ നിന്ന് ലഘുയാദൃച്ഛികസാമ്പിളനമനുസരിച്ച് ഒരു സമുച്ചയത്തെ തിരഞ്ഞെടുക്കുകയാണ്, ക്രമാനുസൃതസാമ്പിളനത്തിൽ നാം ചെയ്യുന്നതു്. ഇവിടെ n അംഗങ്ങൾ വീതമുള്ള മുൻപറഞ്ഞ k സമുച്ചയങ്ങളിൽ ഓരോന്നിനെയും ഓരോ വ്യക്തിയായി പരിഗണിച്ചാൽ അതിൽ നിന്ന് ഒരംഗമുള്ള ഒരു ലഘുയാദൃച്ഛിക സാമ്പിൾ എടുക്കുന്നതിന്ന് സമാനമാണ് ഈ സാമ്പിളനം.

ഒരു ചെറിയ ഉദാഹരണം കൊണ്ടു് സാമ്പിളെടുക്കുന്ന സമ്പ്രദായവും അതിനോടനുബന്ധിച്ചു് ചർച്ച ചെയ്ത കാര്യങ്ങളും വിശദമാക്കാം. താഴെ കൊടുത്തിരിക്കുന്ന സമഷ്ടി പരിഗണിക്കുക.

3, 8, 4, 6, 11, 5, 4, 9, 2

ഇതിൽ 9 അംഗങ്ങളാണുള്ളതു്. 3 അംഗങ്ങളുള്ള ഒരു സാമ്പിളാണ് എടുക്കേണ്ടതെന്നിരിക്കട്ടെ. ആദ്യമായി സമഷ്ടിയിലെ അംഗങ്ങൾക്കു് ക്രമസംഖ്യകൾ നൽകുക. സൗകര്യമായ ഏതു ക്രമം വേണമെങ്കിലും സ്വീകരിക്കാവുന്നതാണ്. വലിയ സമഷ്ടികൾ പരിഗണിക്കുമ്പോൾ അന്വേഷകന്ന് ഏററവും സൗകര്യമായി ഓരോ അംഗത്തേയും സമീപിക്കാൻ ഉപകരിക്കുമാറുള്ള ക്രമമാണ് സ്വീകരിക്കാറുള്ളതു്. വീടുകളുടെ സമഷ്ടിയാണു് പരിഗണിക്കുന്നതെങ്കിൽ അന്വേഷകൻ സഞ്ചരിക്കാനുദ്ദേശിക്കുന്ന വഴിയോടു ബന്ധപ്പെടുത്തി ക്രമേണ വീടുകൾക്കു് ക്രമസംഖ്യ കൊടുക്കാം. ഒരു ഓഫീസിൽ, അലമാരകളിൽ അടുക്കിവെച്ചിരിക്കുന്ന ഫയലുകളാണ് പരിഗണിക്കപ്പെടുന്ന സമഷ്ടിയെങ്കിൽ, അവ അടുക്കിയിരിക്കുന്ന ക്രമത്തിൽ തന്നെ ക്രമസംഖ്യകൾ കൊടുക്കുന്നതായിരിക്കും നല്ലതു്. ഇവിടെ തന്നിരിക്കുന്ന ക്രമത്തിൽ തന്നെ അംഗങ്ങൾക്കു് ക്രമസംഖ്യകൾ കൊടുക്കുന്നു എന്നിരിക്കട്ടെ.

പട്ടിക 5 1

ക്രമസംഖ്യ	അംഗം	ക്രമസംഖ്യ	അംഗം	ക്രമസംഖ്യ	അംഗം
1	3	4	6	7	4
2	8	5	11	8	9
3	4	6	5	9	2

3 അംഗങ്ങളുള്ള സാമ്പിളാണല്ലോ എടുക്കേണ്ടതു്. അതായതു് $N=9$, $n=3$. അതുകൊണ്ടു് $k=3$. എന്നു സിദ്ധിക്കുന്നു. 1 മുതൽ 3 വരെയുള്ള സംഖ്യകളിൽ നിന്ന് യാദൃച്ഛികമായി ഒരു സംഖ്യ തിരഞ്ഞെടുക്കണം. യാദൃച്ഛികസംഖ്യാപട്ടിക ഉപയോഗിച്ചു് ഇതു് ചെയ്യാവുന്നതാണ്. അങ്ങനെ കിട്ടിയ സംഖ്യ 2

ആണെന്നിരിക്കട്ടെ. $2, 2+3=5, 2+2\times3=8$ എന്നീ ക്രമസംഖ്യകളോടു കൂടിയ അംഗങ്ങളെയാണ് സാമ്പിളിൽ ഉൾപ്പെടുത്തേണ്ടതു്. അതായതു് 8, 11, 9, എന്നതാണ് സാമ്പിൾ. പട്ടിക 5.1 ലെ ഓരോ കോളത്തിലും എഴുതിയിരിക്കുന്ന മുമ്മൂന്ന് അംഗങ്ങളുടെ സമുച്ചയങ്ങൾ ഓരോ സ്തരങ്ങളാണെന്നു സങ്കല്പിച്ചാൽ അതിൽ ഓരോന്നിൽ നിന്നും ഓരോ അംഗത്തെ സാമ്പിളിലേക്കു തിരഞ്ഞെടുക്കുകയാണ് ചെയ്തിരിക്കുന്നതു്. ഈ അർത്ഥത്തിലാണ് ഇതൊരു സ്തരിതസാമ്പിളാണെന്ന് മുൻപു സൂചിപ്പിച്ചതു്. പക്ഷേ സ്തരണം സ്വേച്ഛാപരമായിട്ടാണ് ചെയ്തിരിക്കുന്നതു് എന്നതും, ആദ്യത്തേതൊഴികെയുള്ള സ്തരങ്ങളിൽ നിന്നു് സാമ്പിളംലേക്കു അംഗങ്ങളെ തിരഞ്ഞെടുത്തതു് യാദൃച്ഛികമായിട്ടല്ല എന്നതും പ്രത്യേകിശ്രദ്ധിക്കേണ്ട കാര്യങ്ങളത്രെ. പട്ടിക 5.1 ലെ ഓരോ വരിയിലെയും അംഗങ്ങളുടെ സമുച്ചയം ഓരോ വ്യക്തിയാണെന്നു സങ്കല്പിച്ചാൽ അതിൽ നിന്നു് യാദൃച്ഛികമായി ഒരു വ്യക്തിയെ, അതായതു് രണ്ടാമത്തെ വരിയിലെ അംഗങ്ങളുടെ സമുച്ചയത്തെ, തിരഞ്ഞെടുക്കുകയാണ് നാം ചെയ്തതു്. ഇതാണ് ലഘുയാദൃച്ഛികസാമ്പിളനവുമായി ഈ സാമ്പിളന രീതിക്കുള്ള സാമ്യം.

ഇവിടെ N, n ൻെറ ഒരു ഗുണിതമാണെന്നാണ് നാം സങ്കല്പിച്ചതു്. പക്ഷേ പലപ്പോഴും അതു് ശരിയായിരിക്കണമെന്നില്ല. അങ്ങനെ വരുമ്പോൾ $\frac{N}{n}$ നോടു ഏററവും അടുത്ത പൂർണസംഖ്യ k യുടെ മൂല്യമായി സ്വീകരിക്കുകയാണ് സാധാരണ പതിവു്. മുൻ ഉദാഹരണത്തിൽ 4 അംഗങ്ങളുള്ള സാമ്പിളാണ് എടുക്കേണ്ടിയിരിക്കുന്നതെന്നിരിക്കട്ടെ. ഇവിടെ $N=9$, $n=4$; $\frac{N}{n}=\frac{9}{4}=2.25$. k യുടെ മൂല്യം 2 എന്നു് എടുക്കുക. 2 ആണല്ലൊ 2.25 നോടു് ഏററവും അടുത്ത പൂർണ്ണസംഖ്യ. ഈ ഏകദേശനം മൂലം സംഭവിക്കാവുന്ന ഒരു കുഴപ്പം എടുത്തു പറയേണ്ടതത്രെ. n അംഗങ്ങളുള്ള സാമ്പിളായിരിക്കും നമുക്കു് ആവശ്യം. പക്ഷേ കിട്ടുന്ന സാമ്പിളിൻെറ പരിമാണം n തന്നെ ആയിരിക്കണമെന്നില്ല. അതു് n നേക്കാൾ കൂടുതലോ കുറവോ ആയിവരാവുന്നതാണ്. ഇവിടെത്തന്നെ 1,2 എന്നിവയിൽ നിന്നു് ഒരു സംഖ്യയാണല്ലൊ യാദൃച്ഛികപ്രാരംഭമായി തിരഞ്ഞെടുക്കുന്നതു്. 1 ആണ് തിരഞ്ഞെടുക്കുന്നതെങ്കിൽ സാമ്പിളിലെ അംഗങ്ങളുടെ ക്രമസംഖ്യകൾ 1, 3, 5, 7, 9 എന്നിവയാണ്. അതായതു്, സാമ്പിളിൽ 5 അംഗങ്ങളുണ്ടായിരിക്കും. നേരേ മറിച്ചു് 2 ആണ് യാദൃച്ഛികപ്രാരംഭമെങ്കിൽ 2, 4, 6, 8 എന്നീ ക്രമസംഖ്യകളുള്ള നാലു് അംഗങ്ങളാണ് സാമ്പിളിൽ ഉണ്ടായിരിരിക്കുക. മറെറാരുദാഹരണം എടുക്കാം. സമഷ്ടിയിൽ 189 അംഗങ്ങൾ ഉണ്ടെന്നും സാമ്പിൾ പരിമാണം 30 ആയിരിക്കണമെന്നാണ് ആഗ്രഹിക്കുന്നതെന്നും വിചാരിക്കുക. $N=189, n=30$, $k\ \frac{N}{n}=\frac{189}{30}=6.3$. മുൻ നിയമമനുസരിച്ചു് $k=6$. യാദൃച്ഛികപ്രാരംഭം 1,2,3. എന്നിവയിൽ ഏതെങ്കിലുമാണെങ്കിൽ സാമ്പി

ളിൽ 32-ഉം മറേറതെങ്കിലുമാണെങ്കിൽ 31 ഉം അംഗങ്ങളാണ് ഉണ്ടായിരിക്കുക. നേരേമറിച്ച് ഈ സമഷ്ടിയിൽ 198 അംഗങ്ങളുണ്ടെങ്കിൽ $k=7$ എന്നു വരും. 1,2 എന്നിവ യാദൃച്ഛികപ്രാരംഭങ്ങളായുള്ള സാമ്പിളുകളിൽ 29 അംഗങ്ങളും മററുള്ളവയിൽ 28 അംഗങ്ങളും മാത്രമെ ഉണ്ടായിരിക്കുകയുള്ളു. പൊതുവെ പറഞ്ഞാൽ N, n ൻെറ ഗുണിതമല്ലെങ്കിൽ $N= nq+r$ എന്നു് എഴുതാം. ഇവിടെ $r \geq \frac{n}{2}$ എങ്കിൽ k യുടെ മൂല്യം q ഉം, $r > \frac{n}{2}$ എങ്കിൽ, $q+1$ ഉം ആയിട്ടാണ് നാം സ്വീകരിക്കുന്നതു്. ഇങ്ങനെ ലഭിക്കുന്ന സാമ്പിളിൽ ഉണ്ടായിരിക്കാവുന്ന അംഗങ്ങളുടെ എണ്ണം $\left[\frac{N}{K}\right]$ അഥവാ $\left[\frac{N}{K}\right]+1$ ആയിരിക്കും. (ഇവിടെ ചതുര ബ്രാക്കററു് സംഖ്യയുടെ പൂർണാംശത്തെ സൂചിപ്പിക്കുന്നു.) സാമ്പിൾ പരിമാണം വലുതാണെങ്കിൽ ഈ വ്യതാസം അവഗണിക്കാവുന്നതാണ്. പക്ഷേ ചെറിയ സാമ്പിളാണ് എടുക്കുന്നതെങ്കിൽ ഇതു് അവഗണിക്കുന്നതു് ശരിയല്ല. പക്ഷേ ആകലങ്ങളുടെ പ്രതീക്ഷ, പ്രസരണം മുതലായവയുടെ വ്യുൽപാദനത്തിൽ ഇതു് സങ്കീർണതകൾ ഉളവാക്കും. അതുകൊണ്ടു് പ്രയോഗികസൗകര്യത്തെ മുൻനിർത്തി ഈ വ്യത്യാസം എപ്പോഴും അവഗണിക്കുക എന്ന രീതി നിലവിലുണ്ടു്. ഇതാണ് N,n ൻെറ ഗുണിതമല്ലാതെ വരുമ്പോൾ സാധാരണ സ്വീകരിക്കാറുള്ള ഒരു സമ്പ്രദായം. എന്നാൽ മററു ചില സമ്പ്രദായങ്ങളും ഉപയോഗത്തിലുണ്ടു്. അവയിൽ ചിലതു കൂടി താഴെ ചേർക്കുന്നു.

സാമ്പിൾപരിമാണം നിർദിഷ്ടമായ n ഓ അതിൽ കൂടുതലോ ആയിരിക്കത്തക്കവണ്ണം k ക്കു നൽകാവുന്ന ഏററവും വലിയ മൂല്യം നൽകുക. മുൻഉദാഹരണത്തിലെപ്പോലെ സമഷ്ടിയുടെ പരിമാണം 198–ഉം സാമ്പിൾപരിമാണമായി നിർദേശിക്കപ്പെട്ടിരിക്കുന്നതു് 30 ഉം ആണെങ്കിൽ k യുടെ മൂല്യം 6 എന്നു് എടുക്കുക. അപ്പോഴത്തെ സാമ്പിൾപരിമാണം 33 ആയിരിക്കും. ഇങ്ങനെ ലഭിക്കുന്ന സാമ്പിളിൽ അധികമുള്ള അംഗങ്ങളെ ഉപേക്ഷിച്ചു കളയുക എന്നതാണ് ഒരു രീതി. ഇവിടെ 3 അംഗങ്ങളെ ഉപേക്ഷിക്കണം. ഉപേക്ഷിക്കേണ്ട അംഗങ്ങളെ സാമ്പിളിൽ നിന്നു് യാദൃച്ഛികമായി തിരഞ്ഞെടുക്കുകയാണ് ചെയ്യുന്നതു്. ഇതു് താത്വികമായി ഒട്ടും അഭിലഷണീയമായ ഒരു സമ്പ്രദായമല്ല. സാമ്പിൾപരിമാണം വലുതാണെങ്കിൽ ഈ രീതി അവലംബിക്കുന്നതിൽ വലിതകരാറില്ലെന്നു മാത്രം.

മറെറാരു നിർദേശം സമഷ്ടിയുടെ പരിമാണത്തിൽ മാററം വരുത്തുക എന്നതാണ്. അതായതു് നിർദിഷ്ടമായ സാമ്പിൾപരിമാണത്തിൻെറ ഏററവും വലിയ ഗുണിതം കഴിഞ്ഞു കൂടുതൽ വരുന്ന അംഗങ്ങളെ സമഷ്ടിയിൽ നിന്നു് വിട്ടുകളയുക. സമഷ്ടിയുടെ പരിമാണം വലുതും വിട്ടുകളയേണ്ടി വരുന്ന അംഗങ്ങളുടെ എണ്ണം താരതമ്യേന ചെറുതുമാണെങ്കിൽ ഇതുകൊണ്ടു് സാരമായ തകരാറൊന്നും വരുന്നില്ല. പക്ഷേ ചെറിയ സമഷ്ടികൾ കൈകാര്യം ചെയ്യു

മ്പോൾ ഈ രീതി ഒട്ടും സ്വീകാര്യമല്ല. വിട്ടുകളയുന്ന അംഗങ്ങളെ യാദൃച്ഛികമായി തിരഞ്ഞെടുക്കുകയോ, യാദൃച്ഛികമായി തിരഞ്ഞെടുത്ത ഏതെങ്കിലും ഒരു അംഗം മുതൽ ക്രമത്തിൽ ആവശ്യമുള്ളത്ര അംഗങ്ങളെ വിട്ടുകളയുകയോ, ചെയ്യുക എന്ന രീതിയും നിലവിലുണ്ട്. വിട്ടുകളയുന്നതിനു പകരം സമഷ്ടിയുടെ പരിമാണം സാമ്പിൾ പരിമാണത്തിന്റെ ഗുണിതത്തോടു തുല്യമാക്കാൻ ആവശ്യമുള്ള ഏററവും കുറഞ്ഞ എണ്ണം അംഗങ്ങളെ സമഷ്ടിയിൽ നിന്നും യാദൃച്ഛികമായി തിരഞ്ഞെടുത്ത് അവയെ കൂടി സമഷ്ടിയിലെ അംഗങ്ങളുടെ ലിസ്റ്റിന്റെ അവസാനം കൂട്ടിച്ചേർക്കുക എന്ന രീതിയും ഉപയോഗത്തിലുണ്ടു. ഉദാഹരണമായി 128 അംഗങ്ങളുള്ള ഒരു സമഷ്ടിയിൽ നിന്നു് 30 അംഗങ്ങളുള്ള ഒരു സാമ്പിളാണു് എടുക്കേണ്ടതെന്നിരിക്കട്ടെ. 128, 30 ന്റെ ഒരു ഗുണിതമല്ല. 128 ൽ കുറവായ 30 ന്റെ ഏററവും വലിയ ഗുണിതം 120 ആണു്. ബാക്കിയുള്ള 8 അംഗങ്ങളെ സമഷ്ടിയിൽ നിന്നു് വിട്ടുകളയുകയാണു് ഒരു സമ്പ്രദായം. ആ എട്ട് അംഗങ്ങളെ യാദൃച്ഛികമായി തിരഞ്ഞെടുക്കുകയും വേണം. 30 ന്റെ അടുത്ത ഗുണിതം 150 ആണു്. 22 അംഗങ്ങൾ കൂടി ഉണ്ടെങ്കിലേ 150 ആകൂ. ഈ 22 അംഗങ്ങളെ സമഷ്ടിയിൽ നിന്നും യാദൃച്ഛികമായി തിരഞ്ഞെടുത്ത് ഫ്രേമിന്റെ അവസാനം കൂട്ടിച്ചേർക്കുകയാണു് രണ്ടാമതു് പറഞ്ഞ രീതി.

k യുടെ മൂല്യം പൂർണസംഖ്യ ആയിരിക്കണമെന്നുള്ള നിർബന്ധം വിട്ടുകളയുകയാണു് സ്വീകാര്യമായ മറെറാരു സമീപനം. 'അപൂർണസംഖ്യാന്തരാള' സമ്പ്രദായം എന്നു് ഇതിനു പേർ പറയുന്നു. $k = \frac{N}{n}$ എന്നു തന്നെ സ്വീകരിക്കുക. ഒരു ദാഹരണം കൊണ്ടു് ഈ സമ്പ്രദായം വിശദീകരിക്കാം. 835 അംഗങ്ങളാണു് സമഷ്ടിയിലുള്ളതെന്നും 100 അംഗങ്ങളുള്ള സാമ്പിളാണു് എടുക്കേണ്ടതെന്നും കരുതുക $k = \frac{835}{100} = 8.35$ ഒന്നു മുതൽ 835 വരെയുള്ള സംഖ്യകളിൽ നിന്നു് ഒരു സംഖ്യ യാദൃച്ഛികമായി തിരഞ്ഞെടുക്കുക. അങ്ങനെ കിട്ടിയ സംഖ്യയോടു് 835 വീണ്ടും വീണ്ടും കൂട്ടി പുതുതായി 99 സംഖ്യകൾ കൂടി കണ്ടുപിടിക്കുക. യാദൃച്ഛികമായി തിരഞ്ഞെടുത്ത സംഖ്യ 620 ആണെന്നിരിക്കട്ടെ. 620, 1455, 2290, ..., 83285 എന്നിവയാണു് കൂട്ടികിട്ടുന്ന സംഖ്യകൾ. ഇവയോരോന്നിന്റെയും അവസാനത്തെ രണ്ടു് അക്കങ്ങൾ ദശാംശസ്ഥാനങ്ങളാണു്. അവ സംശോധനം ചെയ്തു് ഇവയെ പൂർണസംഖ്യകളാക്കിയാൽ 6, 15, 23, 833 എന്നിങ്ങനെ നൂറു സംഖ്യകൾ കിട്ടും. ഈ സംഖ്യകൾ ക്രമസംഖ്യകളായുള്ള സമഷ്ടിയിലെ അംഗങ്ങളെ സാമ്പിളിൽ ഉൾപ്പെടുത്തുക. താത്വികമായി ഏററവും മെച്ചപ്പെട്ട രീതി ഇതാണു്.

2. ചക്രീസമ്പ്രദായം

രേഖീയസമ്പ്രദായത്തിനു് പ്രധാനമായി രണ്ടു പോരായ്മകളാണുള്ളതു്. സമഷ്ടിയുടെ പരിമാണമായ N സാമ്പിൾപരിമാണമായി നിർദേശിക്കപ്പെട്ടിരിക്കുന്ന n–ന്റെ ഗുണിതമല്ലാതെ വരുമ്പോഴാണു് ഈ പോരായ്മകൾ അനുഭവ

പ്പെടുന്നതു്. അവയിൽ ഒന്നു് നാം കണ്ടു കഴിഞ്ഞു–സാമ്പിൾപരിമാണം നിർദിഷ്ടമായ n ആയിരിക്കണമെന്നില്ല എന്ന സംഗതി. മറെറാരു തകരാറു് സാമ്പിൾ സമാന്തരമാധ്യം സമഷ്ടിയുടെ സമാന്തരമാധ്യത്തിൻെറ ഒരു അനഭിനത ആകലമായിരിക്കുകയില്ല എന്നതാണു്. അടുത്ത ഖണ്ഡികയിൽ ഈ വിഷയം ചർച്ച ചെയ്യുന്നുണ്ടു്. ഈ രണ്ടു തകരാറുകളും ഒഴിവാക്കാൻ ഉപയുക്തമായ ഒന്നാണു് ചക്രീയസമ്പ്രദായം. ഈ സമ്പ്രദായത്തിൽ 1 മുതൽ N വരെയുള്ള സംഖ്യകളിൽ നിന്നും യാദൃച്ഛികമായി തിരഞ്ഞെടുക്കുന്ന ഒരു സംഖ്യയെയാണു് യാദൃച്ഛികപ്രാരംഭമായി സ്വീകരിക്കുന്നതു്. $\frac{N}{n}$ നോടു് ഏററവും അടുത്ത പൂർണസംഖ്യയെയാണു് k യുടെ മൂല്യമായി സ്വീകരിക്കുക. r ആണു് യാദൃച്ഛിക പ്രാരംഭം എന്നിരിക്കട്ടെ. $r, \quad r+k, \quad \ldots\ldots \quad r+(n-1)k$ എന്നീ സംഖ്യകൾ പരിഗണിക്കുക. അവയിൽ N നെക്കാൾ കുറവുള്ളവയെ അതു പോലെ തന്നെ സ്വീകരിക്കുക. മററുള്ളവയിൽ നിന്നു് N കുറച്ചു് ശിഷ്ടവും എടുക്കുക. ഇങ്ങനെ ലഭിച്ച സംഖ്യകൾ ക്രമസംഖ്യകളായുള്ള അംഗങ്ങളെയാണു് സാമ്പിളിൽ ഉൾപ്പെടുത്തേണ്ടതു്. പ്രതീകാത്മകമായി പറഞ്ഞാൽ, സാമ്പിളിൽ ഉൾപ്പെടുത്തേണ്ട അംഗങ്ങളുടെ ക്രമസംഖ്യകൾ

$r+lk, \quad r+lk \leq N$ എങ്കിൽ

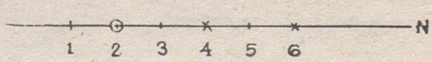
$r+lk-n, \quad r+lk > N$ എങ്കിൽ, $\quad l=0, 1, 2, \ldots n-1.$

എന്നിവയാണു്. ഈ സമ്പ്രദായം ആദ്യമായി നിർദേശിച്ചതു് ഡി.ബി. ലാഹിരി (1952) യാണു്. ഇന്നിതു് ധാരാളമായി ഉപയോഗിച്ചു വരുന്നുണ്ടു്.

ഈ രണ്ടു സാമ്പിളനരീതികളുടെയും ആരേഖചിത്രീകരണങ്ങൾ താഴെ കൊടുക്കുന്നു.

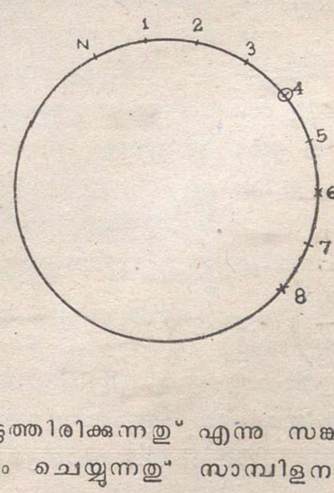
രേഖീയ സമ്പ്രദായം.

k = 2. O - യാദൃച്ഛികപ്രാരംഭം



ചക്രീയസമ്പ്രദായം.

k = 2. O - യാദൃച്ഛികപ്രാരംഭം.



3 സമാന്തരമാധ്യത്തിൻെറ ആകലനം

രേഖീയസമ്പ്രദായത്തിലാണു് സാമ്പിളെടുത്തിരിക്കുന്നതു് എന്നു സങ്കല്പിക്കുക. സമഷ്ടിയുടെ സമാന്തരമാധ്യം ആകലനം ചെയ്യുന്നതു് സാമ്പിളനലക്ഷ്യ

ങ്ങളിൽ പ്രധാനപ്പെട്ട ഒന്നാണ്. സാമ്പിൾ സമാന്തരമാധ്യമാണ് ആകലമായി ഉപയോഗിക്കാറുള്ളതു്. ലഘുയാദൃച്ഛിക സാമ്പിളനത്തിലും സ്തരിതസാമ്പിളനത്തിലും യഥാക്രമം സാമ്പിൾസമാന്തരമാധ്യവും ഭാരിതസാമ്പിൾസമാന്തരമാധ്യവും സമഷ്ടിയുടെ സമാന്തരമാധ്യത്തിന്റെ അനഭിനത ആകലങ്ങളാണെന്ന് കണ്ടു കഴിഞ്ഞു. ക്രമാനുസൃതസാമ്പിളനത്തിലെ സ്ഥിതി എന്താണെന്നു പരിശോധിച്ചു നോക്കാം. $N=nk$ ആണെന്നിരിക്കട്ടെ. സമഷ്ടിയിലെ അംഗങ്ങളെ താഴെ കൊടുത്തിരിക്കുന്നതു പോലെ ക്രമീകരിക്കാവുന്നതാണ്.

പട്ടിക 5.2

x_1	x_{k+1}	x_{2k+1}	...... $x_{(n-1)k+1}$
x_2	x_{k+2}	x_{2k+2}	...... $x_{(n-1)k+2}$
...	......	....	
...	......	....	
x_r	x_{k+r}	x_{2k+r}	...... $x_{(n-1)k+r}$
...	......	....	
...	......	....	
x_k	x_{2k}	x_{3k}	..... x_{nk}

ഈ പട്ടികയിലെ k വരികളിൽ ഏതെങ്കിലും ഒന്നിൽ എഴുതിയിരിക്കുന്ന അംഗങ്ങളാണ് സാമ്പിളായി ലഭിക്കുന്നതു്. അങ്ങനെ ഈ സമഷ്ടിയിൽ നിന്നു് എടുക്കാവുന്ന n പരിമാണമുള്ള ക്രമാനുസൃതസാമ്പിളുകളുടെ എണ്ണം k എന്നു സിദ്ധിക്കുന്നു. 1 മുതൽ k വരെയുള്ള സംഖ്യകളിൽ ഏതെങ്കിലും ഒന്ന് യാദൃച്ഛികമായി എടുത്താണല്ലോ യാദൃച്ഛികപ്രാരംഭം നിർണയിക്കുന്നതു്. അതുകൊണ്ടു് ഈ k സാമ്പിളുകളിൽ ഓരോന്നും ലഭിക്കുവാനുള്ള സംഭാവ്യത തുല്യമാണ്. ആ സംഭാവ്യത $1/k$ യുമത്രെ

ഇവിടെ ആനുഷംഗികമായി മറെറാരു കാര്യം കൂടി അനുസ്മരിക്കാവുന്നതാണ്. സമഷ്ടിയിലെ ഏതൊരംഗത്തിനും സാമ്പിളിൽ ഉൾപ്പെടുവാനുള്ള സംഭാവ്യത അതു് അംഗമായുള്ള സാമ്പിൾ തിരഞ്ഞെടുക്കപ്പെടുവാനുള്ള സംഭാവ്യതയായ $1/k$ തന്നെയാണ്. അതായതു്, സമഷ്ടിയിലെ ഓരോ അംഗത്തിനും സാമ്പിളിൽ ഉൾപ്പെടുവാനുള്ള സംഭാവ്യത തുല്യമാണ്. ലഘുയാദൃച്ഛികസാമ്പിളിനെ സംബന്ധിച്ചും നാം ഈ തത്വം തെളിയിക്കുകയുണ്ടായി. ഇവ തമ്മിലുള്ള വ്യത്യാസം ആ സംഭാവ്യതയുടെ പരിമാണത്തെ സംബന്ധിച്ചു മാത്രമാണ്. പ്രതിസ്ഥാപനമില്ലാത്ത ലഘുയാദൃച്ഛിക സാമ്പിളനത്തിൽ ഈ സംഭാവ്യത $1/\binom{N}{n}$ ആണെങ്കിൽ ഇവിടെ അതു് $1/k$ അത്രെ.

ക്രമാനുസൃതസാമ്പിളിലെ സമാന്തരമാധ്യത്തെ $\bar{x}_{sy}$ എന്ന ചിഹ്നം കൊണ്ടും സമഷ്ടിയുടെ സമാന്തരമാധ്യത്തെ X കൊണ്ടും, r യാദൃച്ഛിക പ്രാരംഭമായുള്ള സാമ്പിളിന്റെ സമാന്തരമാധ്യത്തെ $\bar{x}_r$ കൊണ്ടും സൂചിപ്പിക്കുന്നു എന്നിരിക്കട്ടെ.

$$\bar{X} = \frac{1}{nk} \sum_{r=1}^{k} \sum_{j=0}^{n-1} x_{jk+r}, \quad \bar{x}_r = \frac{1}{n} \sum_{j=0}^{n-1} x_{jk+r}$$

[x_{jk+r}, r–ാമത്തെ സാമ്പിളിലെ jാമത്തെ അംഗത്തിൻെറ x ൻെറ മൂല്യം]

$\bar{x}_{sy}$ ന് k മൂല്യങ്ങളാണ് എടുക്കാവുന്നത്. അവ $\bar{x}_1, \bar{x}_2 \ldots\ldots \bar{x}_k$ എന്നിവയാണ്. ഈ ഓരോ മൂല്യങ്ങളും എടുക്കുവാനുള്ള സംഭാവ്യത $\frac{1}{k}$ ആണെന്നും നാം കണ്ടു കഴിഞ്ഞു. അങ്ങിനെ,

$$E(\bar{x}_{sy}) = \frac{1}{k} \sum_{r=1}^{k} \bar{x}_r = \frac{1}{nk} \sum_{r=1}^{k} \sum_{j=0}^{n-1} x_{jk+r} = \bar{X} \qquad 5.3.1$$

അതായത്, ലഭിക്കാവുന്ന എല്ലാ സാമ്പിളുകളുടെയും സമാന്തരമാധ്യങ്ങളുടെ സമാന്തരമാധ്യം അഥവാ സാമ്പിൾസമാന്തരമാധ്യത്തിൻെറ പ്രതീക്ഷ സമഷ്ടിയുടെ സമാന്തരമാധ്യമാണെന്നു വരുന്നു. മററൊരു തരത്തിൽ പറഞ്ഞാൽ, സാമ്പിൾസമാന്തരമാധ്യം സമഷ്ടിയുടെ സമാന്തരമാധ്യത്തിൻെറ ഒരു അനഭിനത ആകലമാണ്.

ഇവിടെ N=nk എന്നു നാം സങ്കല്പിക്കുകയുണ്ടായി. അതായതു് സമഷ്ടിയുടെ പരിമാണം സാമ്പിൾപരിമാണത്തിൻെറ ഒരു ഗുണിതമാണെന്നു സങ്കല്പിച്ചു. അങ്ങനെയല്ലാതെ വരുമ്പോഴത്തെ സ്ഥിതി എന്താണെന്നുകൂടി പരിശോധിക്കേണ്ടിയിരിക്കുന്നു. കാര്യങ്ങൾ കൂടുതൽ വ്യക്തമാക്കുന്നതിനുവേണ്ടി ആദ്യമായി ഒരു സമഷ്ടി പരിഗണിക്കാം. സമഷ്ടിയിൽ 7 അംഗങ്ങളുണ്ടെന്നും അവയുടെ അഭിലക്ഷണമൂല്യങ്ങൾ $x_1, x_2, x_3, x_4, x_5, x_6, x_7$ എന്നിവയാണെന്നും കരുതുക. 2 അംഗങ്ങളുള്ള സാമ്പിളാണ് എടുക്കേണ്ടതു് എന്നിരിക്കട്ടെ. ഇവിടെ N=7, n=2, $k = \frac{7}{2}$. ഇവിടെ k യുടെ മൂല്യം 3 എന്നോ 4 എന്നോ എടുക്കാവുന്നതാണ്. k=4 എന്ന് നിശ്ചയിക്കുന്നു എന്നു സങ്കല്പിക്കുക. താഴെ കൊടുത്തിരിക്കുന്ന പട്ടിക പരിശോധിക്കുക.

പട്ടിക 5.3

x_1	x_5
x_2	x_6
x_3	x_7
x_4	$\bar{X}$

ഓരോ വരിയിലും എഴുതിയിരിക്കുന്നവയാണ് സാമ്പിളുകളായി ലഭിക്കാവുന്നതു് യാദൃച്ഛിക പ്രാരംഭം 1, 2, 3 എന്നിവയിൽ ഏതെങ്കിലുമാണെങ്കിൽ സാമ്പിളിൽ രണ്ടു് അംഗങ്ങൾ വീതമുണ്ടാവും. 4 ആണെങ്കിൽ ഒരംഗമേ ഉണ്ടാവൂ.

$$\bar{X} = \frac{x_1 + x_2 + x_3 + x_4 + x_5 + x_6 + x_7}{7}$$

$$\bar{x} = \frac{x_1+x_5}{2}, \bar{x}_2 = \frac{x_2+x_6}{2}, \bar{x}_3 = \frac{x_3+x_7}{2}$$

$$\bar{x}_y = x_y$$

ഇതിൽ നിന്നു്, $E(\bar{x}_{sy}) = \frac{1}{4}[\bar{x}_1+\bar{x}_2+\bar{x}_3+\bar{x}_4]$

$$= \frac{1}{4}\left[\frac{x_1+x_5}{2} + \frac{x_2+x_6}{2} + \frac{x_3+x_7}{2} + x_4\right]$$

$$= \frac{1}{8}\left[x_1+x_2+x_3+2x_4+x_5+x_6+x_7\right]$$

$$= \frac{7}{8}\bar{X} + \frac{x_4}{8}$$

അതായതു് $E(\bar{x}_{sy}) \neq \bar{X}$

ഈ തത്വം $N = nk$ അല്ലാത്ത, N അംഗങ്ങളുള്ള ഏതു സമഷ്ടിയെ സംബന്ധിച്ചും ശരിയാണെന്നു് സാമാന്യവൽക്കരിക്കാൻ പ്രയാസമില്ല. പട്ടിക 5.2ലെ അവസാനത്തെ കോളത്തിന്റെ ഒട്ടുവിലത്തെ ഏതാനും സംഖ്യകൾ ഇല്ലാതെ വരുമെന്നുള്ളതാണല്ലോ ഈ സ്ഥിതിയുടെ പ്രത്യേകത. ആ സങ്കല്പത്തിൽ ഇവിടെ കൊടുത്തിരിക്കുന്നതുപോലെ ക്രിയചെയ്താൽ $E(\bar{x}_{sy}) \neq \bar{X}$ എന്നു കാണാം. ചുരുക്കത്തിൽ സമഷ്ടിയുടെ പരിമാണം സാമ്പിൾ പരിമാണത്തിന്റെ ഗുണിതമല്ലെങ്കിൽ സാമ്പിൾ സമാന്തരമാധ്യം സമഷ്ടിയുടെ സമാന്തരമാധ്യത്തിന്റെ അനഭിനത ആകലമല്ല.

അസുഖകരമായ ഒരു സാഹചര്യത്തിലാണു് നാം. സാമ്പിൾ സമാന്തരമാധ്യം സമഷ്ടിയുടെ സമാന്തരമാധ്യത്തിന്റെ അനഭിനത ആകലമല്ലെന്നു വന്നാൽ അതിനെ ആകലമായി തന്നെ അംഗീകരിക്കാൻ നമുക്കു് വൈമനസ്യമുണ്ടാവുക സ്വാഭാവികമാണു്. ഈ ബുദ്ധിമുട്ടിൽനിന്നു് രക്ഷ നേടുവാൻ കുറുക്കുവഴികളുണ്ടോ എന്നു് അന്വേഷിക്കേണ്ടിയിരിക്കുന്നു. ഒരു നിർദേശം സാമ്പിളനരീതിയിൽ ചില പരിഷ്കാരങ്ങൾ വരുത്തുകയാണു്. 1 മുതൽ k വരെയുള്ള സംഖ്യകളിൽ നിന്നു് യാദൃച്ഛികമായി ഒന്നു തിരഞ്ഞെടുക്കുന്നതിനു് പകരം 1 മുതൽ N വരെയുള്ളതിൽ നിന്നു് ഒരു സംഖ്യ തിരഞ്ഞെടുത്തു് അതു് യാദൃച്ഛിക പ്രാരംഭമായി സ്വീകരിക്കുക; അതിൽ നിന്നു് മുൻപോട്ടും പിറകോട്ടും k യാമതു്, k യാമതു് അംഗത്തെ സാമ്പിളിൽ ഉൾപ്പെടുത്തുക— ഇതാണു് നിർദേശിക്കപ്പെടുന്ന ഒരു പരിഷ്കാരം. പട്ടിക 5.3 ലെ സമഷ്ടി തന്നെ പരിഗണിക്കാം. $k=4$ ആണല്ലോ. യാദൃച്ഛികപ്രാരംഭമായി 1 മുതൽ 7 വരെയുള്ള സംഖ്യകളിൽ ഏതെങ്കിലും ഒന്നിനെയാണു് തിരഞ്ഞെടുക്കുന്നതു്. അതു് 1 അഥവാ 5 ആണെങ്കിൽ സാമ്പിൾ x_1, x_5 ആയിരിക്കും. 2 അഥവാ 6 ആണെങ്കിൽ സാമ്പിൾ x_2, x_6. അതുപോലെ 3 അഥവാ 7 ആണെങ്കിൽ സാമ്പിൾ x_3, x_7. പക്ഷേ 4 ആണെങ്കിൽ

സാമ്പിൾ x_4. അങ്ങനെ ആദ്യത്തെ മൂന്നു സാമ്പിളുകൾ ലഭിക്കാനുള്ള സംഭാവ്യത $\frac{2}{7}$ ഉം നാലാമത്തേതിനുള്ളതു് $\frac{1}{7}$ ഉം. ഇതിലെ സാമ്പിൾസമാന്തര മാധ്യങ്ങളും അവയുടെ സംഭാവ്യതയും താഴെ കൊടുക്കുന്നു.

പട്ടിക 5.4

സാമ്പിൾസമാന്തരമാധ്യം	സംഭാവ്യത
$\frac{x_1+x_5}{2}$	$\frac{2}{7}$
$\frac{x_2+x_6}{2}$	$\frac{2}{7}$
$\frac{x_3+x_7}{2}$	$\frac{2}{7}$
x_4	$\frac{1}{7}$

$\bar{x}_{sy}$ നു് എടുക്കാവുന്ന മൂല്യങ്ങൾ ഈ സാമ്പിൾ സമാന്തരമാധ്യങ്ങളാണല്ലോ. അതുകൊണ്ടു്

$$E(\bar{x}_{sy}) = \left(\frac{x_1+x_5}{2}\right)\frac{2}{7} + \left(\frac{x_6+x_5}{2}\right)\frac{2}{7} + \left(\frac{x_2+x_7}{2}\right)\frac{2}{7} + x_4\frac{1}{7}$$

$$= \frac{1}{7}\left[x_1+x_5+X_2+x_6+x_3+x_7+x_4\right]$$

$$= \bar{X}$$

ഈ തത്വം N അംഗങ്ങളുള്ള സമഷ്ടിയിലേക്കു് സാമാന്യവൽക്കരിക്കാവുന്നതാണു്. അങ്ങനെ സാമ്പിളനരീതിയിൽ ഈ പരിഷ്കാരം വരുത്തിയാൽ സാമ്പിൾ സമാന്തരമാധ്യം ഒരു അനഭിനത ആകലമായിത്തീരുമെന്നു സിദ്ധിക്കുന്നു. ഇവിടെയും സാമ്പിൾ പരിമാണം ക്ലിപ്തമല്ല എന്ന വസ്തുത ഓർമിക്കേണ്ടതാണു്. മാത്രമല്ല, ഈ പരിഷ്കാരത്തിന്റെ ഫലമായി സമഷ്ടിയിലെ ഓരോ അംഗത്തിനും സാമ്പിളിൽ ഉൾപ്പെടുവാനുള്ള സംഭാവ്യത തുല്യമല്ലാതായി തീരുകയും ചെയ്യുന്നു. മുൻ ഉദാഹരണത്തിൽ x_1 നു് സാമ്പിളിൽ ഉൾപ്പെടുവാനുള്ള സംഭാവ്യത $\frac{2}{7}$ ആയിരിക്കുമ്പോൾ x_4 ന്റെതു $\frac{1}{7}$ മാത്രമാണല്ലോ. ഈ കാരണങ്ങളാൽ സാമ്പിളനരീതിയിൽ ഇങ്ങനെ ഒരു പരിഷ്കാരം വരുത്തുന്നതു് അത്ര അഭികാമ്യമാണെന്നു പറഞ്ഞുകൂട.

ചക്രീയരീതി അവലംബിക്കുകയാണ് ഈ കുഴപ്പം ഒഴിവാക്കാനുള്ള മറ്റൊരു പോംവഴി. സമഷ്ടിയുടെ പരിമാണം സാമ്പിൾപരിമാണത്തിൻെറ ഗുണിതമാണെങ്കിൽ ചക്രീയരീതിയും രേഖീയരീതിയും ഫലത്തിൽ ഒന്നുതന്നെയാണെന്നു കാണുവാൻ വിഷമമില്ല. $N=nk$ അല്ലാത്ത സന്ദർഭങ്ങളിലും ചക്രീയരീതി അവലംബിച്ചാൽ സമഷ്ടിയിലെ ഓരോ അംഗത്തിനും സാമ്പിളിൽ ഉൾപ്പെടുവാനുള്ള സംഭാവ്യത തുല്യമായിരിക്കുമെന്നു മാത്രമല്ല, സാമ്പിൾ സമാന്തരമാധ്യം സമഷ്ടിയുടെ സമാന്തരമാധ്യത്തിൻെറ ഒരു അനഭിനത ആകലമായിരിക്കുകയും ചെയ്യും. ഉദാഹരണമായി പട്ടിക 5.3 ൽ കൊടുത്തിരിക്കുന്ന 7 അംഗങ്ങളുള്ള സമഷ്ടി തന്നെ പരിഗണിക്കാം. രണ്ടു അംഗങ്ങളുള്ള സാമ്പിളാണ് എടുക്കുന്നതു് എന്നിരിക്കട്ടെ, $k=4$ ആണ്. യാദൃച്ഛികപ്രാരംഭം 1 മുതൽ 7 വരെയുള്ള ഏതെങ്കിലും സംഖ്യയാവാം. വിവിധ യാദൃച്ഛിക പ്രാരംഭങ്ങളും അതനുസരിച്ചുള്ള സാമ്പിളുകളും അവയുടെ സമാന്തരമാധ്യങ്ങളും. താഴെ കൊടുക്കുന്നു.

പട്ടിക 5.5

യാദൃച്ഛികപ്രാരംഭം	സാമ്പിൾ	സമാന്തരമാധ്യം
1	x_1, x_5	$\frac{x_1+x_5}{2}$
2	x_2, x_6	$\frac{x_2+x_6}{2}$
3	x_3, x_7	$\frac{x_3+x_7}{2}$
4	x_4, x_1	$\frac{x_4+x_1}{2}$
5	x_5, x_2	$\frac{x_5+x_2}{2}$
6	x_6, x_3	$\frac{x_6+x_3}{2}$
7	x_7, x_4	$\frac{x_7+x_4}{2}$

$$E(\bar{x}_{sy}) = \frac{1}{7}\left[\frac{x_1+x_5}{2}+\frac{x_2+x_6}{2}+\frac{x_3+x_7}{2}+\frac{x_4+x_1}{2}+\frac{x_5+x_2}{2}+\frac{x_6+x_3}{2}+\frac{x_7+x_4}{2}\right]$$

$$= \frac{1}{14}\left[2x_1+2x_2+2x_3+2x_4+2x_5+2x_6+2x_7\right] = \bar{X}$$

ഇവിടെ സമഷ്ടിയിലെ ഓരോ അംഗവും രണ്ടു സാമ്പിളുകളിൽ വീതം ഉൾപ്പെട്ടു വരുന്നു. അതായതു് ഏതൊരംഗത്തിനും സാമ്പിളിൽ ഉൾപ്പെടുവാനുള്ള സംഭാവ്യത $\frac{2}{7}$ ആണു്.

അടുത്തതായി N അംഗങ്ങളുള്ള സമഷ്ടിയിൽ നിന്നു് n അംഗങ്ങളുള്ള സാമ്പിൾ ചക്രീയരീതിയിൽ എടുക്കുമ്പോഴത്തെ അവസ്ഥ പരിഗണിക്കാം. 1 മുതൽ N വരെയുള്ള ഏതെങ്കിലും ഒരു സംഖ്യ യാദൃച്ഛികപ്രാരംഭമായി എടുക്കുന്നു. അങ്ങനെ N സാമ്പിളുകൾ ആണു് സമഷ്ടിയിൽ നിന്നു് ആകെ എടുക്കാൻ കഴിയുന്നതു്. ഓരോ സാമ്പിളും ലഭിക്കുവാനുള്ള സംഭാവ്യത $\frac{1}{N}$ ഒരു സാമ്പിളിൽ n അംഗങ്ങളാണല്ലോ ഉണ്ടായിരിക്കുക. സമഷ്ടിയിലെ ഒരംഗത്തിനു് സാമ്പിളിലെ ആദ്യത്തെ അംഗം, രണ്ടാമത്തെ അംഗം, മൂന്നാമത്തെ അംഗം,... n–ാമത്തെ അംഗം എന്നിങ്ങനെ n വിധങ്ങളിൽ സാമ്പിളിൽ ഉൾപ്പെടാം. അതുകൊണ്ടു് സമഷ്ടിയിലെ ഓരോ അംഗവും n സാമ്പിളുകളിൽ അംഗമായിരിക്കും. അങ്ങനെ സാമ്പിളുകളിലെ അംഗങ്ങളുടെ തുക സമഷ്ടിയിലെ അംഗങ്ങളുടെ തുകയുടെ n ഇരട്ടി ആയിരിക്കും. അതായതു് $Nn\,\overline{X}$.

$x_{i1}, x_{i2}, \cdots x_{in}$ എന്നിവയാണു് i-ാമത്തെ സാമ്പിളിലെ അംഗങ്ങളെന്നിരിക്കട്ടെ.

$$\overline{x}_i = \frac{x_{i1} + x_{i2} + \cdots \cdots + x_{in}}{n}$$

$$E(\overline{x}_{sy}) = \frac{1}{N} \sum_{i=1}^{N} \overline{x}_i = \frac{1}{Nn} \sum_{i=1}^{N} \sum_{j=1}^{n} x_{ij}$$

$$= \frac{nN\overline{X}}{Nn} = \overline{X} \left(\because \sum_{i=1}^{N} \sum_{j=1}^{n} x_{ij} = Nn\,\overline{X} \right) \qquad 5.3.2$$

അങ്ങനെ സാമ്പിൾ സമാന്തരമാധ്യം സമഷ്ടിയുടെ സമാന്തരമാധ്യത്തിൻെറ അനഭിനത ആകലമാണെന്നു വന്നുചേരുന്നു. മാത്രമല്ല, സമഷ്ടിയിലെ ഓരോ അംഗവും n സാമ്പിളുകളിൽ വീതം ഉൾപ്പെട്ടുവരുന്നതുകൊണ്ടു് അവ ഓരോന്നിനും സാമ്പിളിൽ അംഗമാകുവാനുള്ള സംഭാവ്യത $\frac{n}{N}$ ആണെന്നും സിദ്ധിക്കുന്നു. k യുടെ മൂല്യം സാധാരണയായി $\frac{N}{n}$ നോടു് ഏററവും അടുത്ത പൂർണസംഖ്യയായാണു് എടുക്കാറുള്ളതു്.

സാമ്പിൾ പരിമാണം n തന്നെ ആയിരിക്കുക, സാമ്പിൾസമാന്തരമാധ്യം സമഷ്ടിയുടെ സമാന്തരമാധ്യത്തിൻെറ അനഭിനത ആകലമായിരിക്കുക, സമഷ്ടിയിലെ ഓരോ അംഗത്തിനും സാമ്പിളിൽ ഉൾപ്പെടുവാനുള്ള സംഭാവ്യത തുല്യമായിരിക്കുക എന്നീ വൈശിഷ്ട്യങ്ങൾ ചക്രീയസമ്പ്രദായത്തിൽ തിരഞ്ഞെടുക്ക

പ്പെട്ട സാമ്പിളുകൾക്ക് ഉണ്ട് എന്നതാണ് ആ രീതി രേഖീയരീതിയെക്കാൾ ശ്രേഷ്ഠമാണെന്നു കരുതാനുള്ള കാരണം. പക്ഷേ പ്രായോഗിക സൗകര്യം രേഖീയരീതിക്കാണെന്നു പറയാതെ തരമില്ല.

സമാന്തരമാധ്യത്തിന്റെ ആകലമായി സ്വീകരിക്കപ്പെട്ട സാമ്പിൾ സമാന്തരമാധ്യത്തിന്റെ പ്രതീക്ഷയാണ് ഇതുവരെ ചർച്ചാവിഷയമായത്. അടുത്തതായി അതിന്റെ സാമ്പിളനപ്രസരണം എന്താണെന്നു പരിശോധിക്കാം. സാമ്പിൾ പരിമാണത്തിന്റെ ഗുണിതമാണ് സമഷ്ടിയുടെ പരിമാണമെന്ന്, അതായത് $N = nk$ ആണ് എന്ന് സങ്കല്പിച്ചുകൊണ്ടാണ് ഇവിടെ ചർച്ച ചെയ്യുന്നത്. അങ്ങനെ അല്ലാത്ത സന്ദർഭങ്ങളിലും ഇവിടെ വ്യുൽപാദിപ്പിക്കുന്ന വ്യംജകങ്ങൾ ഏകദേശനങ്ങൾ എന്ന നിലയിൽ സ്വീകാര്യങ്ങളത്രെ. കൂടുതൽ സങ്കീർണമാണ് എന്നതുകൊണ്ട് $N=nk$ അല്ലാത്തപ്പോഴത്തെ വ്യംജകങ്ങൾ ഇവിടെ വ്യുൽപ്പാദിപ്പിക്കുന്നില്ല,

$x_{i1}, x_{i2}, \ldots x_{in}$ ആണ് i–ാമത്തെ സാമ്പിൾ. $i = 1, 2, \ldots k$ $\bar{x}_i$, i–ാമത്തെ സാമ്പിളിന്റെ സമാന്തരമാധ്യം. സാമ്പിൾ സമാന്തരമാധ്യങ്ങളുടെ സാമ്പിളന പ്രസരണം $V(\bar{x}_{sy})$ എന്നിരിക്കട്ടെ. അവയുടെ പ്രതീക്ഷ $\bar{X}$ ആണെന്നു നാം കണ്ടു കഴിഞ്ഞു. നിർവചനമനുസരിച്ച്,

$$V(\bar{x}_{sy}) = \frac{1}{k} \sum_{i=1}^{k} (\bar{x}_i - \bar{X})^2 = \frac{k-1}{k} S^2w, \text{ ഇവിടെ}$$

$$S^2w = \sum_{i=1}^{k} \frac{(\bar{x}_i - \bar{X})^2}{k-1}, \text{ സാമ്പിൾസമാന്തരമാധ്യങ്ങളുടെ വർഗമാധ്യമാണ്.}$$

സമഷ്ടിയുടെ പ്രസരണം σ^2 എന്നിരിക്കട്ടെ.

$$\sigma^2 = \frac{1}{nk} \sum_{i=1}^{k} \sum_{j=1}^{n} (x_{ij} - \bar{X})^2$$

$$= \frac{1}{nk} \sum_{i=1}^{k} \sum_{j=1}^{n} (x_{ij} - \bar{x}_{ij} + \bar{x}_i - \bar{X})^2$$

$$= \frac{1}{nk} \sum_{i=1}^{k} \sum_{j=1}^{n} \left[(x_{ij} - \bar{x}_i)^2 + (\bar{x}_i - \bar{X})^2 + 2(x_{ij} - \bar{x}_i)(\bar{x}_i - X)\right]$$

$$= \frac{1}{nk} \sum_{i=1}^{k} \sum_{j=1}^{n} (x_{ij} - \bar{x}_i)^2 + \frac{1}{k} \sum_{i} (\bar{x}_i - \bar{X})^2 \qquad 5.3.3$$

ഇവിടെ വലതുവശത്തെ രണ്ടു പദങ്ങൾ ശ്രദ്ധിച്ച് പരിശോധിച്ചു നോക്കുക.

$\frac{1}{n} \sum_{j=1}^{n} (x_{ij} - \bar{x}_i)^2$ എന്നത് i–ാമത്തെ സാമ്പിളിന്റെ പ്രസരണമാണ്. σ^2_{wi}

എന്ന ചിഹ്നം കൊണ്ടു് അതിനെ സൂചിപ്പിക്കുന്നു എന്നിരിക്കട്ടെ.

$$\frac{1}{nk}\sum_{i=1}^{k}\sum_{j=1}^{n}(x_{ij}-\bar{x}_i)^2=\frac{1}{k}\sum_{i=1}^{k}\sigma^2_{wi}=\bar{\sigma}^2_w$$ എന്നിരിക്കട്ടെ.

വ്യക്തമായും $\bar{\sigma}^2_w$, സമഷ്ടിയിൽ നിന്നു് എടുക്കാവുന്ന n പരിമാണമുള്ള ക്രമാനുസൃത സാമ്പിളുകളുടെ പ്രസരണത്തിൻെറ സമാന്തരമാധ്യമത്രെ. സാമ്പിളുകൾക്കുള്ളിലെ പ്രസരണത്തിൻെറ മാപമായി അതിനെ സ്വീകരിക്കാം. രണ്ടാമത്തെ പദം $V(\bar{x}_{sy})$ ആണല്ലൊ.

അങ്ങനെ, $\bar{x}_{sy}$ യുടെ പ്രസരണം

$$V(\bar{x}_{sy})=\sigma^2-\bar{\sigma}^2_w \qquad 5.34$$

എന്നു വരുന്നു.

$V(\bar{x}_{sy})$ ൻെറ ഈ വ്യംജകം പരിശോധിച്ചു നോക്കുക. സാമ്പിൾ സമാന്തരമാധ്യങ്ങളുടെ പ്രസരണമാണല്ലൊ അതു്. അതായതു്, സാമ്പിളുകളുടെ വികീർണനത്തിൻെറ (സാമ്പിളുകൾ തമ്മിലുള്ള അന്തരത്തിൻെറ) മാപമായി അതിനെ കണക്കാക്കാം. σ^2_b എന്ന ചിഹ്നം കൊണ്ടു് അതു് സൂചിപ്പിക്കാറുണ്ടു്. σ_b^2 നെ സാമ്പിളുകളുടെ 'അന്തര'പ്രസരണമെന്നും $\bar{\sigma}_w^2$-നെ 'ആന്തര'പ്രസരണമെന്നും വിളിക്കുന്നു.

σ^2 സമഷ്ടിയുടെ പ്രസരണമാണു്. അതു് സാമ്പിൾ മാറുന്നതനുസരിച്ചു് മാറുന്നില്ല. അപ്പോൾ $V(\bar{x}_{sy})$ ചെറുതാക്കാനുള്ള മാർഗം (ആകലത്തിൻെറ സൂക്ഷ്മത വർധിപ്പിക്കുവാനുള്ള മാർഗം) $\bar{\sigma}_w^2$ വലുതാക്കുക എന്നതു മാത്രമാണു്. അതായതു്, പരിഗണിക്കപ്പെടുന്ന അഭിലക്ഷണത്തെ സംബന്ധിച്ചിടത്തോളം കഴിവുള്ളിടത്തോളം ഭിന്നാത്മകമായിരിക്കണം സാമ്പിൾ. സമഷ്ടിയെ k അംഗങ്ങൾ വീതമുള്ള സമുച്ചയങ്ങളായി വിഭജിച്ചു് അവ ഓരോന്നിൽ നിന്നും ഓരോ അംഗത്തെ വീതമാണല്ലൊ സാമ്പിളിൽ ഉൾപ്പെടുത്തുന്നതു്. ഈ സമുച്ചയങ്ങൾ കഴിവതും ഏകാത്മകങ്ങളാക്കിയാൽ അവ തമ്മിലുള്ള ഭിന്നാത്മകത്വം വർധിക്കും. സ്വാഭാവികമായും സാമ്പിളിൻെറ ആന്തരപ്രസരണം വർധിക്കുകയും ചെയ്യും.

ചുരുക്കത്തിൽ $V(\bar{x}_{sy})$ ചെറുതാക്കാനുള്ള മാർഗം സമഷ്ടിയിലെ അംഗങ്ങളുടെ ക്രമം ബോധപൂർവം വിന്യസിക്കുകയാണു്. കഴിവതും ഏകാത്മകങ്ങളായ k അംഗങ്ങൾക്കു വീതം 1 മുതൽ k വരെയും, $k+1$ മുതൽ $2k$ വരെയും അങ്ങനെ ക്രമേണ $(n-1)k+1$ മുതൽ nk വരെയുമുള്ള ക്രമസംഖ്യകൾ കൊടുത്താൽ ഇതു സാധിക്കാം.

ഇതിൽ നിന്നു് മറെറാരു വസ്തുത കൂടി വ്യക്തമാവുന്നു. സാമ്പിൾസമാന്തരമാധ്യത്തിൻെറ പ്രസരണം സമഷ്ടിയിലെ അംഗങ്ങൾ ഏതു ക്രമത്തിൽ പരിഗണിക്കപ്പെടുന്നു എന്നതിനെ ആശ്രയിച്ചാണിരിക്കുന്നതു്. ഇതു് ഒരു തരത്തിൽ ഗുണകരമാണെങ്കിലും മറ്റൊരു തരത്തിൽ നോക്കിയാൽ അത്യന്തം അപകടകരമാണെന്നു പറയാതെ തരമില്ല. ബോധപൂർവമുള്ള ക്രമീകരണം കൊണ്ടു് കൂടുതൽ സൂക്ഷ്മങ്ങളായ ആകലങ്ങൾ കണ്ടുപിടിക്കാമെന്നതാണു് ഇതിൻെറ ഗുണവശം. പക്ഷെ സമഷ്ടിയെപ്പററിയുള്ള നമ്മുടെ അറിവു് പരിമിതമായിരിക്കുകയോ

ഇഷ്ടംപോലെ ക്രമീകരിക്കാൻ പ്രായോഗിക വൈഷമ്യങ്ങളുണ്ടായിരിക്കുകയോ ചെയ്യുന്ന സന്ദർഭങ്ങളിൽ നമുക്ക് ലഭിക്കുന്ന ആകലം വളരെ മോശപ്പെട്ടതാവാൻ പാടില്ലായ്കയില്ല എന്നതാണ് ഇതിന്റെ മറുവശം.

രണ്ടു സാമ്പിളന രീതികൾ താരതമ്യപ്പെടുത്തുന്നത് പ്രധാനമായും അവയിൽ നിന്നു ലഭിക്കുന്ന ആകലങ്ങളുടെ സൂക്ഷ്മത പരിഗണിച്ചാണ്. മറ്റെല്ലാം തുല്യമായിരിക്കുമ്പോൾ, കൂടുതൽ സൂക്ഷ്മതയുള്ള ആകലങ്ങൾ ഏതു സാമ്പിളന രീതി അവലംബിച്ചാലാണോ കിട്ടുന്നത്, ആ രീതി ശ്രേഷ്ഠതരമായി പരിഗണിക്കപ്പെടും. ക്രമാനുസൃത സാമ്പിളനവും ലഘുയാദൃച്ഛികസാമ്പിളനവും താരതമ്യപ്പെടുത്തണമെന്നിരിക്കട്ടെ. ഒരേ സമഷ്ടിയിൽ നിന്ന് ഒരേ പരിമാണമുള്ള സാമ്പിളുകൾ രണ്ടു സമ്പ്രദായത്തിലും എടുക്കുക. സമഷ്ടിയുടെ പരിമാണം N-ഉം സാമ്പിൾ പരിമാണം n- ഉം ആണെന്നിരിക്കട്ടെ. സമഷ്ടിയുടെ സമാന്തരമാധ്യം സാമ്പിൾ സമാന്തരമാധ്യം ഉപയോഗിച്ചാണ് രണ്ടു സമ്പ്രദായത്തിലും ആകലനം ചെയ്യുന്നത്. $\bar{x}_R$, $\bar{x}_{sy}$ എന്നിവ യഥാക്രമം ലഘുയാദൃച്ഛിക സാമ്പിളിന്റെയും ക്രമാനുസൃതസാമ്പിളിന്റെയും സമാന്തരമാധ്യമങ്ങളാണെന്നും സങ്കല്പിക്കുക.

$$V(\bar{x}_R) = \frac{N-n}{N}\frac{s^2}{n} \quad \text{അഥവാ} \quad \frac{N-n}{N-1}\frac{\sigma^2}{n} \quad \text{എന്നും}$$

$$V(\bar{x}_{sy}) = \sigma^2 - \bar{\sigma}_w^2 \quad \text{എന്നും നാം കണ്ടുകഴിഞ്ഞു.}$$

ക്രമാനുസൃത സാമ്പിളിൽ നിന്നു ലഭിക്കുന്ന ആകലം ലഘുയാദൃച്ഛിക സാമ്പിളിൽ നിന്നു ലഭിക്കുന്നതിനെക്കാൾ സൂക്ഷ്മമായിരിക്കണമെങ്കിൽ,

$$V(\bar{x}_{sy}) < V(\bar{x}_R)$$

ആയിരിക്കണം. അതായത്,

$$\sigma^2 - \bar{\sigma}_w^2 < \frac{N-n}{N-1}\frac{\sigma^2}{n}$$

അതായത് $$\sigma^2\left[1 - \frac{N-n}{(N-1)n}\right] < \bar{\sigma}_w^2$$

$$\sigma^2\left[\frac{Nn-n-N+n}{(N-1)n}\right] < \bar{\sigma}_w^2$$

$$\frac{N(n-1)}{n(N-1)}\sigma^2 < \bar{\sigma}_w^2$$

$$\frac{N\sigma^2}{N-1} < \frac{n\bar{\sigma}_w^2}{n-1} \qquad 5.33$$

ഇങ്ങനെയുള്ള സന്ദർഭങ്ങളിൽ മാത്രമേ ക്രമാനുസൃത സാമ്പിളനം ലഘ്വയാദൃച്ഛിക സാമ്പിളനത്തെക്കാൾ മെച്ചപ്പെട്ടതു് എന്നു പറയാനാവൂ. ഇതു് പ്രത്യേകം ഓർമ്മിക്കേണ്ട ഒരു വസ്തുതയാണു്.

ഈ പ്രശ്നത്തെതന്നെ മറെറാരു വിധത്തിലും സമീപിക്കാവുന്നതാണു്.

$$V(\bar{x}_{sy}) = \frac{1}{k}\sum_{i=1}^{k}(\bar{x}_i-\bar{X})^2 = \frac{1}{k}\sum_{i=1}^{k}\left[\frac{1}{n}\sum_{j=1}^{n}x_{ij}-\bar{X}\right]^2$$

$$= \frac{1}{n^2k}\sum_{i=1}^{k}\left[\sum_{j=1}^{n}(x_{ij}-\bar{X})\right]^2$$

$$= \frac{1}{Nn}\sum_{i=1}^{k}\sum_{j=1}^{n}(x_{ij}-\bar{X})^2$$

$$+ \frac{1}{Nn}\sum_{i=1}^{k}\sum_{l\neq j=1}^{n}(x_{ij}-\bar{X})(x_{il}-\bar{X})$$

$$= \frac{\sigma^2}{n} + \frac{(n-1)}{n}\rho\sigma^2 \qquad 5.3.6$$

[ഇവിടെ $\dfrac{\sum_{i=1}^{k}\sum_{l\neq j=1}^{n}(x_{ij}-\bar{X})(x_{il}-\bar{X})}{kn(n-1)\sigma^2} = \rho w$

എന്നു സങ്കല്പിച്ചിരിക്കുന്നു. ρ_w നു അന്തരാവർഗസഹബന്ധം എന്നാണു് പേർ പറയുക. ഓരോ സാമ്പിളിലും ഉള്ള അംഗങ്ങളെ ജോഡികളായി എടുത്താൽ അവ തമ്മിലുള്ള സഹബന്ധമാണു് ഇതു് കുറിക്കുന്നതു്.]

അങ്ങനെ $V(\bar{x}_{sy}) = \dfrac{\sigma^2}{n}\left[1+(n-1)\rho_w\right]$

എന്നു സിദ്ധിക്കുന്നു.

$V(\bar{x}_{sy}) = \sigma^2 - \bar{\sigma}_w^2$ എന്നു് തെളിയിച്ചിട്ടുണ്ടല്ലോ.

അപ്പോൾ, $\dfrac{\sigma^2}{n}[1+(n-1)\rho_w] = \sigma^2 - \bar{\sigma}_w^2$ എന്നു വരുന്നു.

അതായതു്, $\rho_w = 1 - \dfrac{n}{n-1}\dfrac{\bar{\sigma}_w^2}{\sigma^2}$

$\bar{\sigma}^2_w$, σ^2 നോട്ട തുല്യമോ അതിനെക്കാൾ ചെറുതോ ആയിരിക്കാനേ തരമുള്ളു വല്ലൊ. അതുകൊണ്ടു് $\frac{\bar{\sigma}^2_w}{\sigma^2}$ ൻെറ ഏററവും വലിയ മൂല്യം 1 ആണ്. അങ്ങനെ ρ_w യുടെ ഏററവും ചെറിയ മൂല്യം $1 - \frac{n}{n-1} = - \frac{1}{n-1}$ ആണ് എന്നു വരുന്നു. ρ_w ഒരു സഹബന്ധഗുണാങ്കമായതു കൊണ്ടു് അതിൻെറ മൂല്യം 1 ൽ കൂടുതലാവാനും തരമില്ല. അതായതു്,

$$-\frac{1}{n-1} \leq \rho_w \leq 1 \qquad 5.3.7$$

പ്രതിസ്ഥാപനമില്ലാത്ത ലഘ്വയാദൃച്ഛികസാമ്പളനത്തെ സംബന്ധിച്ചിടത്തോളം

$$V(\bar{x}_R) = \frac{N-n}{N-1}\frac{\sigma^2}{n}.$$

ക്രമാനുസൃതസാമ്പിൾ, ലഘ്വയാദൃച്ഛികസാമ്പിളിനെക്കാൾ മെച്ചപ്പെട്ടതാകണമെങ്കിൽ,

$V(\bar{x}_{sy}) < V(\bar{x}_R)$ ആയിരിക്കണം. അതായതു്,

$\frac{\sigma^2}{n}\left[1+(n-1)\rho_w\right] < \frac{N-n}{N-1}\frac{\sigma^2}{n}$ ആയിരിക്കണം.

ഇതു് ലഘൂകരിക്കുമ്പോൾ, $\rho_w < -\left(\frac{1}{N-1}\right)$ എന്നാവും.

ലഘ്വയാദൃച്ഛികസാമ്പിളനവുമായി താരതമ്യപ്പെടുത്തുമ്പോഴത്തെ ക്രമാനുസൃതസാമ്പിളിൻെറ ക്ഷമത E ആണെങ്കിൽ,

$$E = \frac{V(\bar{x}_R)}{V(\bar{x}_{sy})} = \frac{N-n}{N-1} \cdot \frac{1}{1+(n-1)\rho_w} \qquad 5.3.8$$

ഇതിൽ നിന്നു് ρ_w യുടെ മൂല്യം വർധിക്കുന്നതനുസരിച്ചു് ക്ഷമത കുറയുന്നെന്നും ρ_w യുടെ മൂല്യം കുറയുന്നതനുസരിച്ചു് ക്ഷമത വർധിക്കുന്നെന്നും കാണാം. ρ_w യുടെ ഏതാനും മൂല്യങ്ങളും അതിനനുഗുണമായ ക്ഷമത E യുടെ മൂല്യങ്ങളും താഴെ കൊടുക്കുന്നു.

പട്ടിക 5 6

ρ_w	E
1	$\frac{N-n}{n(N-1)}$

.5	$\frac{2(N-n)}{(n+1)(N-1)}$
0	$\frac{N-n}{N-1}$
$-\frac{1}{N-1}$	1
$-\frac{1}{n-1}$	∞

ഈ പട്ടിക പരിശോധിച്ചുനോക്കുക. ρ_w യുടെ മൂല്യം $-\frac{1}{N-1}$ നെക്കാൾ കുറഞ്ഞു തുടങ്ങുമ്പോൾ മാത്രമെ ക്ഷമത 1 ൽ കൂടുതലാവൂ. ρ_w, ധനചിഹ്നത്തോടുകൂടിയതാണെങ്കിൽ അതിൽ വരുന്ന ചെറിയ വർധനവു പോലും (അതിനെ n—1 കൊണ്ടു ഗുണിക്കുന്നതു കൊണ്ട്) ക്ഷമതയിൽ സാരമായ കുറവുണ്ടാക്കും. അപ്പോൾ ക്ഷമത വർധിപ്പിക്കുവാനുള്ള മാർഗം ρ_w ഋണ ചിഹ്നത്തോടുകൂടിയതും അതിൻെറ നിരപേക്ഷമൂല്യം കഴിവതും വലുതും ആക്കുകയാണ്. സാമ്പിളിലെ അംശങ്ങളെ സാധ്യമായ എല്ലാ വിധത്തിലും ജോഡികളായെടുത്ത് നിർദേശിക്കുന്ന സഹബന്ധഗുണാങ്കമാണല്ലൊ ρ_w. അതുകൊണ്ട് അത് വലിയ ഋണ സംഖ്യയായിരിക്കണമെങ്കിൽ സാമ്പിളിലെ അംഗങ്ങൾ എത്രയും ഭിന്നാത്മകങ്ങളായിരിക്കണം. അതിനുള്ള ഒരു മാർഗം സമഷ്ടിയിലെ അംഗങ്ങളെ ആരോഹണക്രമത്തിലോ അവരോഹണക്രമത്തിലോ വിന്യസിക്കുക എന്നതാണ്. പരിഗണിക്കപ്പെടുന്ന അഭിലക്ഷണവുമായി നേരെ ബന്ധപ്പെട്ട സൗകര്യപ്രദമായ ഏതെങ്കിലും ഒന്നിനെ ആധാരമാക്കി ഇതു ചെയ്യാവുന്നതാണ്.

ലഘുയാദൃച്ഛികസാമ്പിളിലും സ്തരിതസാമ്പിളനത്തിലും, സാമ്പിളനപ്രസരണം സമഷ്ടിയുടെ പ്രസരണത്തേയും സാമ്പിൾപരിമാണത്തെയും മാത്രം ആശ്രയിച്ചിരിക്കുന്നു. എന്നാൽ ക്രമാനുസൃതസാമ്പിളനത്തിലെ സാമ്പിളനപ്രസരണം അതിനു പുറമെ അന്തരാവർഗസഹബന്ധവുമായി കൂടി ബന്ധപ്പെട്ടാണിരിക്കുന്നത്. മറെറാരു തരത്തിൽ പറഞ്ഞാൽ, സമഷ്ടിയിലെ അംഗങ്ങൾ ഏതു ക്രമത്തിലാണ് പരിഗണിക്കപ്പെടുന്നത് എന്നത് ക്രമാനുസൃതസാമ്പിളനത്തിലെ സാമ്പിളനപ്രസരണത്തിന്റെ പരിമാണത്തെ സംബന്ധിച്ചിടത്തോളം നിർണായകമത്രെ. ഒരുദാഹരണം കൊണ്ട് ഇതു വ്യക്തമാക്കാം. താഴെ കൊടുത്തിരിക്കുന്ന സമഷ്ടി പരിഗണിക്കുക.

പട്ടിക 5.7

1	2	2	3	4	6
8	9	8	14	16	18
20	22	22	24	28	30

ഇവിടെ 18 അംഗങ്ങളാണ് സമഷ്ടിയിലുള്ളത്. ഇതിൽ നിന്നു് 3 അംഗങ്ങൾ വീതമുള്ള സാമ്പിളാണ് എടുക്കുന്നതു് എന്നിരിക്കട്ടെ. സമഷ്ടിയിലെ അംഗങ്ങളെ മുകളിൽ എഴുതിയിരിക്കുന്ന ക്രമത്തിലാണ് പരിഗണിക്കുന്നതെങ്കിൽ ഓരോ കോളത്തിലും വരുന്നതു് ഇതിൽ നിന്നു് എടുക്കാവുന്ന ഒരോ സാമ്പിളാണ്. അവയുടെ സമാന്തരമാധ്യങ്ങൾ താഴെ കൊടുക്കുന്നു.

പട്ടിക 5.8

സാമ്പിൾ	സമാന്തരമാധ്യം	സാമ്പിൾ	സമാന്തരമാധ്യം
1, 8, 20	9.7	3, 14, 24	13.7
2, 9, 22	11.0	4, 16, 28	16.0
2, 8, 22	10.7	6, 18, 30	18.0

അടുത്തതായി സമഷ്ടിയിലെ അംഗങ്ങളെ മറെറാരു തരത്തിൽ പരിഗണിക്കാം.

പട്ടിക 5.9

1	2	2	3	4	6
18	16	14	8	9	8
20	22	22	24	28	30

ഇവിടെയും ഓരോ കോളത്തിൽ വരുന്നതു് ഇതിൽ നിന്നു് എടുക്കാവുന്ന ഓരോ സാമ്പിളാണ്. സാമ്പിളുകളും അവയുടെ സമാന്തരമാധ്യങ്ങളും താഴെ കൊടുക്കുന്നു.

പട്ടിക 5.10

സാമ്പിൾ	സമാന്തരമാധ്യം	സാമ്പിൾ	സമാന്തരമാധ്യം
1,18,20	13.0	3,8,24	11.7
2,16,22	13.3	4,9,28	13.7
2,14,22	12 7	6,8,30	14.7

പട്ടിക 5.8–ലെ സാമ്പിൾ സമാന്തരമാധ്യങ്ങളുടെ പ്രസരണം, അതായതു് പട്ടിക 5.7 ലെ ക്രമത്തിൽ സമഷ്ടിയിലെ അംഗങ്ങളെ പരിഗണിച്ചാലുള്ള,

$$V(\bar{x}_{sy}) = \frac{1}{6} \times 1097.27 - (13.2)^2 = 8.64$$

പട്ടിക 5.10–ലെ സാമ്പിൾ സമാന്തരമാധ്യങ്ങളുടെ പ്രസരണം അതായതു് പട്ടിക 5.9–ലെ ക്രമത്തിൽ സമഷ്ടിയിലെ അംഗങ്ങളെ പരിഗണിച്ചാലുള്ള,

$$V(\bar{x}_{sy}) = \frac{1}{6} \times 1047.85 - (13.2)^2 = 0.40.$$

നേരെ മറിച്ചു് പട്ടിക 5.7 ലെ ക്രമത്തിൽ പരിഗണിച്ചാലും പട്ടിക 5.9 ലെ ക്രമത്തിൽ പരിഗണിച്ചാലും മറേറതെങ്കിലും ക്രമത്തിൽ പരിഗണിച്ചാലും $\bar{x}_R$,

ലഘുയാദൃച്ഛിക സാമ്പിളനരീതിയിലും $\bar{x}_{st}$ സ്തരിതരീതിയിലുമുള്ള ആകലമെങ്കിൽ,

$$V(\bar{x}_R) = \frac{(18-3)}{17}\left\{\frac{1}{18} \times 4683 - (13.2)^2\right\} = 25.3$$

$$V(\bar{x}_{st}) = \frac{6^2 \times 5}{18^2 \times 6}\left[2.7 + 16\ 1 + 12.6\right] = 2.9$$

[ഇവിടെ പട്ടിക 5.7 ലെ ഓരോ വരിയേയും ഓരോ സ്തരമായി കണക്കാക്കുന്നു.]

പട്ടിക 5.7ലെ രണ്ടാമത്തെ വരിയിലെ അംഗങ്ങളുടെ ക്രമം തിരിച്ചിട്ടാൽ പട്ടിക 5.9 ലെ രണ്ടാമത്തെ വരിയായി. അതിൻെറ ഫലമായി സാമ്പിളനപ്രസരണത്തിൽ വന്ന കുറവ് ശ്രദ്ധിക്കുക.

സമഷ്ടിയിലെ അംഗങ്ങളെ ഏതു ക്രമത്തിലാണോ പരിഗണിക്കുന്നത് അതനുസരിച്ച് ക്രമാനുസൃതസാമ്പിളനത്തിലൂടെ ലഭിക്കുന്ന ആകലത്തിൻെറ സാമ്പിളനപ്രസരണം കൂടുകയോ കുറയുകയോ ചെയ്യും എന്ന വസ്തുത ഇതിൽ നിന്നെല്ലാം വ്യക്തമാവുന്നു. ഈ പ്രത്യേകത കണക്കിലെടുത്ത് ബോധപൂർവം പ്രവർത്തിച്ചാൽ പലപ്പോഴും വളരെ നല്ല ആകലങ്ങൾ ലഭിച്ചെന്നു വരും. അത് സാധ്യമാകണമെങ്കിൽ സമഷ്ടിയുടെ പൊതു സ്വഭാവത്തെ പററി സാമാന്യം സൂക്ഷ്മമായി തന്നെ മനസ്സിലാക്കണം. ചില പ്രത്യേക സ്വഭാവങ്ങളുള്ള സമഷ്ടികളിൽ നിന്ന് ക്രമാനുസൃതസാമ്പിൾ എടുത്താലത്തെ സ്ഥിതി അടുത്തതായി പരിഗണിക്കാം.

5.4 യാദൃച്ഛികക്രമത്തിലുള്ള സമഷ്ടികൾ

സമഷ്ടിയിലെ അംഗങ്ങൾക്ക് ക്രമസംഖ്യ കൊടുത്തിരിക്കുന്നത് യാദൃച്ഛികമായിട്ടാണെങ്കിൽ ആ സമഷ്ടി യാദൃച്ഛികക്രമത്തിലുള്ളതാണെന്നു നാം പറയും. ഉദാഹരണമായി ഒരു ഫാക്ടറിയിലെ തൊഴിലാളികളാണ് സമഷ്ടിയെന്നിരിക്കട്ടെ. ഒരു തൊഴിലാളിയുടെ ശരാശരി ശമ്പളമാണ് കണ്ടുപിടിക്കേണ്ടതെന്നു കരുതുക. അവരുടെ പേരുകൾ അക്ഷരമാലാക്രമത്തിൽ എഴുതിയിട്ട് ക്രമസംഖ്യ കൊടുക്കുകയാണെങ്കിൽ യാദൃച്ഛികക്രമത്തിലുള്ള ഒരു സമഷ്ടിയാണ് ലഭിക്കുക. നേരേ മറിച്ച് ഫാക്ടറിയിൽ സൂക്ഷിച്ചിട്ടുള്ള പ്രവേശനരജിസ്റ്ററിലെ ക്രമത്തിൽ അവരെ പരിഗണിച്ചാൽ സേവനകാലത്തിൻെറ അവരോഹക്രമത്തിൽ ക്രമീകരിച്ചിട്ടുള്ള ഒരു സമഷ്ടിയായിരിക്കും ലഭിക്കുക. അടുത്ത ഖണ്ഡത്തിൽ പരിഗണിക്കാൻ പോവുന്ന രേഖീയപ്രവണതയുള്ള സമഷ്ടിയുടെ ഒരുദാഹരണമാണ് ഇത്.

യാദൃച്ഛികക്രമത്തിലുള്ള സമഷ്ടിയിൽ നിന്ന് ഒരു ക്രമാനുസൃത സാമ്പിളെടുത്താൽ അത് അതിൽ നിന്ന് എടുത്ത ഒരു പ്രതിസ്ഥാപനമില്ലാത്ത ലഘു യാദൃച്ഛികസാമ്പിൾ തന്നെയായിരിക്കും. പ്രതിസ്ഥാപനമില്ലാത്ത ലഘുയാ

ദൃച്ഛികസാമ്പിളിന്റെ പ്രത്യേകത, N അംഗങ്ങളുള്ള സമഷ്ടിയിൽ നിന്നു് n അംഗങ്ങളുള്ള ഒരു സാമ്പിളാണെടുക്കുന്നതെങ്കിൽ അതിൽ നിന്നെടുക്കാവുന്ന ഓരോ സാമ്പിളും തിരഞ്ഞെടുക്കപ്പെട്ടുവാനുള്ള സംഭാവ്യത $\dfrac{1}{\binom{N}{n}}$ ആയിരിക്കും എന്നതാണല്ലോ. $N=nk$ എന്നിരിക്കട്ടെ രേഖീയരീതിയും ചക്രീയരീതിയും ഫലത്തിൽ തുല്യമാകുന്ന ഒരു സാഹചര്യമാണല്ലോ ഇതു്. ചക്രീയരീതിയിലാണു് സാമ്പിളെടുക്കുന്നതെന്നിരിക്കട്ടെ. N അംഗങ്ങളുള്ള ഒരു സമഷ്ടിയെ N! ക്രമങ്ങളിൽ പരിഗണിക്കാമല്ലോ. 'യാദൃച്ഛിക ക്രമത്തിലുള്ള സമഷ്ടി' എന്നു പറയുമ്പോൾ ഈ N! ക്രമങ്ങളിലുള്ള സമഷ്ടികളിൽ ഒന്നാണു് നമുക്കു ലഭിച്ചിരിക്കുന്ന സമഷ്ടി എന്നു വരുന്നു. യാദൃച്ഛികപ്രാരംഭമായി വരുന്നതു് സമഷ്ടിയിലെ N അംഗങ്ങളിൽ ഏതു വേണമെങ്കിലുമാവാം. അപ്പോൾ നമുക്കു ലഭിക്കാവുന്ന ക്രമാനുസൃതസാമ്പിളുകളുടെ എണ്ണം (N!) N ആണു്. . ഇവയെല്ലാം വ്യത്യസ്തങ്ങളായിരിക്കണമെന്നില്ല എന്ന വസ്തുത ഓർമിക്കുക. n അംഗങ്ങളുള്ള സാമ്പിളാണല്ലോ എടുക്കുന്നതു്. ആകെയുള്ള (N!)N ക്രമാനുസൃതസാമ്പിളുകളിൽ ഒരേ സാമ്പിൾ തന്നെ എത്ര പ്രാവശ്യം ആവർത്തിച്ചു് പരിഗണിക്കപ്പെട്ടുന്നുണ്ടെന്നു പരിശോധിക്കാം. ചക്രീയരീതിയിൽ സാമ്പിളെടുക്കുന്നു എന്നാണല്ലോ സങ്കല്പിച്ചിരിക്കുന്നതു്. സാമ്പിളിലെ അംഗങ്ങൾ k അകലത്തിലായിരിക്കുകയും വേണം. n നിർദിഷ്ട അംഗങ്ങൾ ഉൾക്കൊള്ളുന്ന ഒരു സാമ്പിൾ പരിഗണിക്കുക. k അകലത്തിലുള്ള n സ്ഥാനങ്ങളിലായിരിക്കുമല്ലൊ സാമ്പിളിലെ അംഗങ്ങൾ. അവയുടെ സ്ഥാനത്തിനു് മാററം വരുത്താതെ ബാക്കിയുള്ള (N−n) അംഗങ്ങളെ (N−n)! വിധത്തിൽ വിന്യസിക്കാൻ കഴിയും. അതുപോലെതന്നെ മററു (N−n) അംഗങ്ങളുടെ ഏതൊരു വിന്യാസക്രമത്തിനും അനുഗുണമായി സാമ്പിളിലെ അംഗങ്ങളെ n! വിധത്തിൽ വിന്യസിക്കാം. മാത്രമല്ല, യാദൃച്ഛിക പ്രാരംഭത്തിന്റെ സ്ഥാനഭേദമനുസരിച്ചു് നിർദിഷ്ടമായ സാമ്പിൾ N തരത്തിൽ വരികയും ചെയ്യാം. അങ്ങനെ ഓരോ സാമ്പിളും (N—n)! n! N പ്രാവശ്യം വീതം ആവർത്തിച്ചു് പരിഗണിക്കുമ്പോഴാണു് മുൻപറഞ്ഞ (N!) N സാമ്പിളുകൾ കിട്ടുന്നതു്. മററൊരുതരത്തിൽ പറഞ്ഞാൽ, സമഷ്ടിയിൽനിന്നു് ചക്രീയരീതിയിൽ എടുക്കാവുന്ന n അംഗങ്ങളുള്ള വ്യത്യസ്തങ്ങളായ സാമ്പിളുകളുടെ എണ്ണം $\dfrac{N!\,N}{(N-n)!\,n!\,N}=\binom{N}{n}$ ആണു്. അവ ഓരോന്നും തിരഞ്ഞെടുക്കപ്പെട്ടുവാനുള്ള സംഭാവ്യത തുല്യമായിരിക്കത്തക്കവണ്ണമാണു് സാമ്പിളനരീതി ആസൂത്രണം ചെയ്തിരിക്കുന്നതു്. അങ്ങനെ n അംഗങ്ങളുള്ള ക്രമാനുസൃത സാമ്പിളുകളുടെ എണ്ണം $\binom{N}{n}$ ഉം അവ ഓരോന്നും തിരഞ്ഞെടുക്കപ്പെട്ടുവാനുള്ള സംഭാവ്യത $\dfrac{1}{\binom{N}{n}}$ ഉം ആണെന്നു വരുന്നു. ഇതു തന്നെയാണല്ലോ പ്രതിസ്ഥാപനമില്ലാത്ത

ലഘുയാദൃച്ഛിക സാമ്പിളനത്തിലെയും സ്ഥിതി. ഇവിടെ N അംഗങ്ങളാണ് സമഷ്ടിയിലുള്ളത്. അത് N! തരത്തിൽ വിന്യസിക്കാനാവും. അങ്ങനെയുള്ള N! സമഷ്ടികളിൽ ഏതെങ്കിലും ഒന്നാണ് നമ്മുടെ സമഷ്ടി എന്ന സങ്കല്പത്തിലാണ് നാം മുന്നോട്ടു പോയതു്. യാദൃച്ഛിക ക്രമത്തിലുള്ള സമഷ്ടി ആയതുകൊണ്ട് പരിഗണിക്കപ്പെടുന്ന സമഷ്ടി ഇതിലേതെങ്കിലും ഒന്നാവാനുള്ള സംഭാവ്യത തുല്യമാണെന്നും സങ്കല്പിച്ചു ലഘുയാദൃച്ഛിക സാമ്പിളാണെടുക്കുന്നതെങ്കിൽ ഈ N! സമഷ്ടികളിൽ ഏതിൽ നിന്നു സാമ്പിളെടുത്താലും അതിൽ നിന്ന് ഗണിക്കുന്ന സമാന്തരമാധ്യത്തിന്റെ സാമ്പിളനപ്രസരണം $V(\bar{x}_R)$ തന്നെയായിരിക്കും. പക്ഷേ ഓരോന്നിൽ നിന്നും എടുക്കുന്ന ക്രമാനുസൃത സാമ്പിളിന്റെ സമാന്തരമാധ്യത്തിന്റെ സാമ്പിളനപ്രസരണം തുല്യമായിരിക്കണമെന്നില്ല. പക്ഷേ അവയുടെ പ്രതീക്ഷ, അതായത് $E[V(\bar{x}_{sy})]$, $V(\bar{x}_R)$ തന്നെയായിരിക്കുമെന്ന് മുൻപറഞ്ഞ വസ്തുതകളിൽ നിന്ന് സിദ്ധിക്കുകയും ചെയ്യുന്നു.

അങ്ങനെ യാദൃച്ഛിക ക്രമത്തിലുള്ള സമഷ്ടിയിൽ നിന്ന് ക്രമാനുസൃതസാമ്പിളെടുത്താൽ അതു് തത്വത്തിൽ ലഘുയാദൃച്ഛിക സാമ്പിളുമായി ഏതാണ്ട് സമാനമായിരിക്കും. പ്രായോഗികമായി നോക്കിയാൽ കൂടുതൽ സൗകര്യപ്രദമാണുതാനും. ഇമ്മാതിരി സമഷ്ടികളെ സംബന്ധിച്ചിടത്തോളം ക്രമാനുസൃതസാമ്പിളിൽ ലഘുയാദൃച്ഛിക സാമ്പിളിനെക്കാൾ പ്രായോഗികസൗകര്യം ഒഴിച്ചു മററു വൈശിഷ്ട്യങ്ങളൊന്നും എടുത്തു പറയാനില്ലെന്നുള്ളതാണ് പ്രസക്തമായ വസ്തുത.

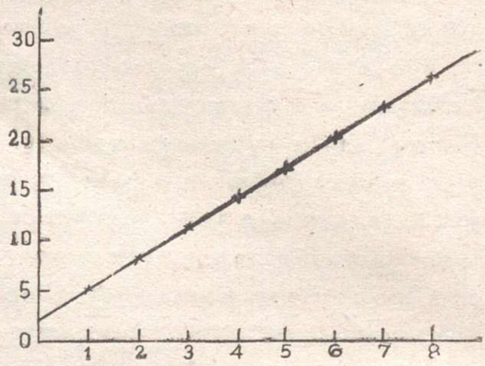
5 രേഖീയപ്രവണതയുള്ള സമഷ്ടികൾ

സമഷ്ടിയിലെ അംഗങ്ങളുടെ ക്രമസംഖ്യ x-അക്ഷത്തിലും അവയുടെ അഭിലക്ഷണമൂല്യങ്ങൾ y-അക്ഷത്തിലും എടുത്ത് ബിന്ദുക്കൾ അടയാളപ്പെടുത്തുന്നു എന്നിരിക്കട്ടെ. ഈ ബിന്ദുക്കൾ ഒരു ഋജുരേഖയുടെ പരിസരങ്ങളിലായി കേന്ദ്രീകരിക്കുകയാണെങ്കിൽ സമഷ്ടിക്ക് രേഖീയപ്രവണതയുണ്ടെന്നു പറയും. താഴെ കൊടുത്തിരിക്കുന്ന സമഷ്ടിയും ചിത്രവും പരിശോധിക്കുക. പൂർണമായ രേഖീയപ്രവണതയുള്ള ഒരു സമഷ്ടിയാണ് ഇതു്.

പട്ടിക 5.11

ക്രമസംഖ്യ	അംഗത്തിന്റെ അഭിലക്ഷണമൂല്യം
1	5
2	8
3	11
4	14
5	17
6	20
7	23
8	26

അഭിലക്ഷണമൂല്യങ്ങൾ ക്രമസംഖ്യ വർധിക്കുന്നതനുസരിച്ച് ഏതാണ്ട് ഒരേ ക്രമത്തിൽ വർധിക്കുകയോ കുറയുകയോ ചെയ്യുകയാണ് ഈ മാതിരി സമഷ്ടിയുടെ സ്വഭാവം. സർവീസനുസരിച്ച് ക്രമീകരിച്ചിട്ടുള്ള ഉദ്യോഗസ്ഥന്മാരുടെ ലിസ്റ്റ്, ശമ്പളം എന്ന അഭിലക്ഷണത്തെ സംബന്ധിച്ചിടത്തോളം ഈ സ്വഭാവമുള്ളതാണ്. രേഖീയപ്രവണത മേലോട്ടോ കീഴോട്ടോ ആവാം. ആ പ്രവണത സമ്പൂർണമാണെങ്കിൽ ബിന്ദുക്കളെല്ലാം ഒരു ഋജുരേഖയിൽ തന്നെ കിടക്കും. അങ്ങനെയുള്ള ഒരു സാങ്കല്പികസമഷ്ടി പരിഗണിച്ച് അതിൽ നിന്നെടുക്കുന്ന ക്രമാനുസൃത സാമ്പിളിന്റെ പ്രത്യേകതകൾ പഠിക്കാൻ ശ്രമിക്കാം. ശരിക്കും രേഖീയമായ ഒരു സമഷ്ടിയിൽ അംഗങ്ങളുടെ ക്രമ സംഖ്യയായ u-ഉം അഭിലക്ഷണമൂല്യമായ x-ഉം തമ്മിൽ



ചിത്രം 5.1

$$x = a + bu$$

എന്ന സമീകരണത്തിന്റെ അടിസ്ഥാനത്തിൽ ബന്ധപ്പെട്ടിരിക്കുമെന്ന് നമുക്കറിയാം. $u = 1, 2, \ldots i \ldots N$ ആവുമ്പോൾ യഥാക്രമം $x = x_1, x_2 \cdots x_i, \cdots x_N$ എന്നിരിക്കട്ടെ.

$$\bar{X} = \frac{1}{N} \Sigma x_i = \frac{1}{N} \sum_{i=1}^{N} (a+bi) = a+b \left(\frac{N(N+1)}{2N}\right)$$

$$= a + \frac{b(N+1)}{2}$$

$$\sigma^2 = \frac{1}{N} \sum_{i=1}^{N} (x_i - \bar{X})^2 = \frac{1}{N} \sum_{j=1}^{N} \left(a+bi-a - \frac{b(N+1)}{2N}\right)^2$$

$$= \frac{b^2}{N} \sum_{i=1}^{N} \left(i - \frac{N+1}{2}\right)^2 = \frac{b^2}{N} \left[\sum_{i=1}^{N} i^2 - \frac{N(N+1)^2}{4} \right]$$

$$= \frac{b^2 (N^2-1)}{12}$$

അതുകൊണ്ട് ഇതിൽ നിന്നെടുത്ത n അംഗങ്ങളുള്ള പ്രതിസ്ഥാപനത്തോടുകൂടിയ ഒരു ലഘുയാദൃച്ഛിക സാമ്പിളിന്റെ സമാന്തരമാധ്യത്തിന്റെ സാമ്പിളന പ്രസരണം,

$$\text{അതായതു്, } V'(\bar{x}_R) = \frac{N-n}{N-1}\frac{\sigma^2}{n}$$

$$= \frac{N-n}{N-1} \cdot \frac{b^2(N^2-1)}{12n}$$

$$= \frac{b^2}{12}(k-1)(N+1) \qquad 5.5.1$$

($N=nk$ എന്നു സങ്കല്പിച്ചാൽ)

പ്രതിസ്ഥാപനമില്ലാത്ത യാദൃച്ഛിക സാമ്പിളിന്റെ സമാന്തരമാധ്യത്തിന്റെ സാമ്പിളന പ്രസരണം

$$V''(\bar{x}_R) = \frac{\sigma^2}{n} = \frac{b^2}{12}\left(k - \frac{1}{n}\right)(N+1) \qquad 5.5.2$$

അടുത്തതായി ഇതിൽ നിന്നു് n അംഗങ്ങളുള്ള ഒരു ക്രമാനുസൃതസാമ്പിൾ എടുത്താലത്തെ സ്ഥിതി പരിശോധിച്ചു നോക്കാം. $N=nk$ എന്നു സങ്കല്പിക്കുക r ആണു് യാദൃച്ഛിക പ്രാരംഭമെന്നും വിചാരിക്കുക. സാമ്പിളിലെ അംഗങ്ങളുടെ ക്രമസംഖ്യകൾ

$$r+ik\ , \ i=0, 1, 2, \ldots n-1$$

ആയിരിക്കും. അവയുടെ അഭിലക്ഷണമൂല്യങ്ങൾ

$$x_{ik+r} = \{a+b(r+ik)\},\ i=0, 1, 2, \ldots n-1$$

എന്നിവയും ആയിരിക്കും. r യാദൃച്ഛികപ്രാരംഭമായുള്ള ഈ ക്രമാനുസൃതസാമ്പിളിന്റെ സമാന്തരമാധ്യത്തെ $\bar{x}_r$ കൊണ്ടു് സൂചിപ്പിക്കുന്നു എന്നിരിക്കട്ടെ.

$$\bar{x}_r = \frac{1}{n}\sum_{i=0}^{n-1} x_{ik+r} = \frac{1}{n}\sum_{i=0}^{n-1}\{a+b(r+ik)\}$$

$$= a+b\left(r+\frac{n-1}{2}k\right)$$

$$E(\bar{x}_r) = \frac{1}{k}\sum_{r=1}^{k}\bar{x}_r = a+b\left(\frac{k+1}{2}+\frac{(n-1)k}{2}\right)$$

$$= a+\frac{b(nk+1)}{2} = a+\frac{1}{2}b(N+1) = \bar{X}$$

$$V(\bar{x}_{sy}) = V(\bar{x}_r) = \frac{1}{k}\sum_{r=1}^{k}(\bar{x}_r-\bar{X})^2$$

$$= \frac{1}{k} \sum_{r=1}^{k} b^2 \left(r - \frac{k+1}{2} \right)^2$$

$$= \frac{b^2}{k} \left[\sum_{r=1}^{k} r^2 - k \left(\frac{k+1}{2} \right)^2 \right]$$

$$= \frac{b^2}{k} \left[\frac{k(k+1)(2k+1)}{6} - \frac{k(k+1)^2}{4} \right]$$

$$= \frac{b^2}{12} (k^2 - 1) = \frac{b^2}{12} (k - 1)(k+1) \qquad 5.5.3$$

$V' (\bar{x}_R) = \frac{b^2}{12} (k - 1)(N+1)$ എന്നും

$V'' (\bar{x}_R) = \frac{b^2}{12} \left(k - \frac{1}{n} \right) (N+1)$ എന്നും നാം കണ്ടുകഴിഞ്ഞു.

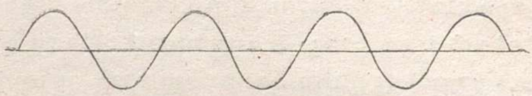
$V (\bar{x}_{sy})$ മായി ഈ രണ്ടു പ്രസരണങ്ങളെ താരതമ്യപ്പെടുത്തിയാൽ സാമ്പിൾപരിമാണം രണ്ടോ കൂടുതലോ ആകുമ്പോൾ $V(\bar{x}_{sy}) < V''(\bar{x}_R) < V'(\bar{x}_R)$ എന്നു കാണാവുന്നതാണ്. അതായതു് രേഖീയ പ്രവണതയുള്ള സമഷ്ടികളെ സംബന്ധിച്ചിടത്തോളം ക്രമാനുസൃതസാമ്പിളിന്റെ ക്ഷമത ലഘുയാദൃച്ഛിക സാമ്പിളിനെ അപേക്ഷിച്ച് വളരെ വലുതാണ്. രേഖീയപ്രവണതയുള്ള സമഷ്ടിയുടെ പരിമാണം സാമ്പിൾപരിമാണത്തെ അപേക്ഷിച്ച് വളരെ വലുതാകുമ്പോൾ, ക്രമാനുസൃതസാമ്പിളലിന്റെ ക്ഷമത, ലഘുയാദൃച്ഛിക സാമ്പിളലിനെക്കാൾ n മടങ്ങ് വർധിക്കുമെന്നു കാണാം. സമാനുപാതരീതിയിൽ ഇങ്ങനെ കുറിക്കാം.

$$V (\bar{x}_{sy}) : V'' (\bar{x}_R) : V' (\bar{x}_R) = \frac{1}{n} : 1 : 1$$

5.6 ആവർത്തികതയുള്ള സമഷ്ടികൾ

നിശ്ചിതഅകലത്തിലുള്ള അംഗങ്ങളുടെ അഭിലക്ഷണമൂല്യങ്ങൾ ഏതാണ്ടു തുല്യമായിട്ടുള്ള സമഷ്ടികൾക്കു് ആവർത്തികതയുള്ള സമഷ്ടികൾ എന്നു് പേർ പറയും. ക്രമേണയുള്ള അംഗങ്ങളുടെ അഭിലക്ഷണമൂല്യങ്ങൾ ആവർത്തന സ്വഭാവത്തോടു കൂടിയ ഒരു നിശ്ചിതക്രമത്തിൽ കൂടുകയും കുറയുകയും ചെയ്യുക എന്നതാണ് ഇമ്മാതിരി സമഷ്ടികളുടെ മൗലികസ്വഭാവം. മറെറാരു തരത്തിൽ പറഞ്ഞാൽ, അംഗങ്ങളുടെ ക്രമസംഖ്യ x അക്ഷത്തിലും അഭിലക്ഷണമൂല്യം y അക്ഷത്തിലും എടുത്തു് ബിന്ദുക്കൾ അടയാളപ്പെടുത്തി എന്നിരിക്കട്ടെ. ആ ബി

ന്ദുക്കൾ ചിത്രം 5.2 ലെ ലേഖയോട്ട് സാമ്യമുള്ള ഒരു ലേഖയുടെ പരിസരത്തിൽ കേന്ദ്രീകരിക്കുന്നതായി കണ്ടാൽ ആ സമഷ്ടി ആവർത്തികതയുള്ളതാണെന്നു പറയാം.



ചിത്രം 5.2

നിശ്ചിത അകലത്തിലുള്ള അംഗങ്ങളുടെ അഭിലക്ഷണമൂല്യങ്ങൾ തുല്യങ്ങളായിരിക്കുക എന്നതാണല്ലോ ഈ പ്രവണതയുള്ള സമഷ്ടികളുടെ പ്രത്യേകത. ഈ നിശ്ചിതഅകലത്തിനു് അതിന്റെ അവധി എന്നു പേർ പറയും. k അവധിയായുള്ള സമഷ്ടിയിൽ $x_r = x_{r+k} = x_{r+2k} = \dots\dots$ എന്നു സാരം. 5 വീടുകൾ വീതമുള്ള ബ്ലോക്കുകളായി നൂറു വീടുകൾ ഗവണ്മെന്റു് പണിയിച്ചിരിക്കുന്നു എന്നിരിക്കട്ടെ. ഓരോ ബ്ലോക്കിലേയും ആദ്യത്തെ വീടുകളുടെ വാടക തുല്യമാണെന്നു സങ്കൽപിക്കുക. അതു പോലെ തന്നെ രണ്ടാമത്തെ വീടുകളുടെയും, മൂന്നാമത്തെ വീടുകളുടെയും, നാലാമത്തെ വീടുകളുടെയും, അഞ്ചാമത്തെ വീടുകളുടെയും. ഓരോ ബ്ലോക്കിലെയും ആദ്യത്തെ വീടു മുതൽ വാടക വർധിച്ച് മൂന്നാമത്തേതിലെത്തുമ്പോൾ ഏററവും കൂടുകയും പിന്നീട്ട് ക്രമേണ കുറയുകയും ചെയ്യുമെന്നും സങ്കൽപിക്കുക. ഈ വീടുകളുടെ സമഷ്ടി 5 അവധിയായുള്ള ആവർത്തികതയുള്ള സമഷ്ടിയാണു്.

ഇമ്മാതിരി സമഷ്ടികളിൽ നിന്നു് ക്രമാനുസൃതസാമ്പിളെടുക്കുന്നതു് വളരെ കരുതി വേണം. അവധിയോ അതിന്റെ ഗുണിതമോ സാമ്പിളനഅന്തരാളമായുള്ള ക്രമാനുസൃത സാമ്പിളുകളിൽ നിന്നു ലഭിക്കുന്ന ആകലങ്ങൾ യാഥാർഥ്യവുമായി വളരെ വ്യത്യസ്തമായിരിക്കാൻ ഇടയുണ്ടു്. മറെറാരു തരത്തിൽ പറഞ്ഞാൽ അമ്മാതിരി ആകലങ്ങളുടെ സാമ്പിളനപ്രസരണം വളരെ വലുതായിരിക്കും. മുൻ ഉദാഹരണത്തിൽ 5 സാമ്പിളനഅന്തരാളമായ ക്രമാനുസൃതസാമ്പിൾ എടുക്കുന്നു എന്നിരിക്കട്ടെ. യാദൃച്ഛികപ്രാരംഭമായി 3 കിട്ടുന്നു എന്നും സങ്കൽപിക്കുക. സാമ്പിളിലെ 20 അംഗങ്ങളും ഏററവും വാടക കൂടിയ വീടുകളായിരിക്കും. അങ്ങനെ സാമ്പിൾസമാന്തരമാധ്യം സമഷ്ടിയുടെ സമാന്തരമാധ്യവുമായി അകന്നു പോവും. സാമ്പിളിലെ അംഗങ്ങളുടെ അഭിലക്ഷണമൂല്യങ്ങൾ ഏതാണ്ടു് തുല്യമായിരിക്കുമെന്നതു കൊണ്ടു്, സാമ്പിളുകളുടെ ആന്തരപ്രസരണം കുറവും അന്തരപ്രസരണം അതിനനുസരിച്ച് കൂടുതലുമാവും. ഈ ഉദാഹരണത്തിൽ ആന്തരപ്രസരണം "0" ആണു്. പൂർണമായ ആവർത്തികതയുള്ള സമഷ്ടികളിലെല്ലാം ഇതായിരിക്കും സ്ഥിതി. സാമ്പിളിന്റെ ക്ഷമത നിസ്സാരമാകുന്ന ഒരു സാഹചര്യമാണു് ഇതു് എന്നു പ്രത്യേകം പറയേണ്ടതില്ലല്ലോ. ആവർത്തികതയുള്ള സമഷ്ടികളിൽ നിന്നു് സാമ്പിളെടുക്കുമ്പോൾ അർധാവധിയുടെ $\left(\frac{k}{2}\right)$ ഒരു ഒറ്റഗുണിതം $\left(\frac{k}{2}, \frac{3k}{2}, \frac{5k}{2}\right.$ തുടങ്ങിയതിൽ ഏതെങ്കിലും ഒന്നു് $\left.\right)$

സാമ്പിളന അന്തരാളമായി സ്വീകരിച്ചാൽ സാമ്പിൾസമാന്തരമാധ്യവും സമഷ്ടിയുടെ സമാന്തരമാധ്യവും തുല്യമായി തീരുമെന്നു കാണാൻ പ്രയാസമില്ല. ഏകദേശമായി മാത്രം ആവർത്തികതയുള്ളതാണ് സമഷ്ടിയെങ്കിലും ഈ രീതിയിൽ സാമ്പിളെടുത്താൽ ആകലത്തിന്റെ സാമ്പിളനപ്രസരണം വളരെയധികം കുറയും. ഇമ്മാതിരി സമഷ്ടികൾ കൈകാര്യം ചെയ്യുമ്പോൾ സാമ്പിളനഅന്തരാളം നിരൂണയിക്കുന്നതു് വളരെ ശ്രദ്ധിച്ചു വേണമെന്നു് ഇതിൽ നിന്നെല്ലാം വ്യക്തമാവുന്നുണ്ടല്ലോ.

7. സാമ്പിളനപ്രസരണത്തിൻെറ ആകലനം

സമഷ്ടിയുടെ സമാന്തരമാധ്യമായ സാമ്പിൾസമാന്തരമാധ്യത്തിൻെറ സാമ്പിളനപ്രസരണം നാം വ്യുൽപ്പാദിപ്പിച്ചു. അതിൻെറ ഒരു അനഭിനതആകലം കണ്ടുപിടിക്കേണ്ടതു് ആവശ്യമാണ്. പക്ഷേ ക്രമാനുസൃതസാമ്പിളിനെ സംബന്ധിച്ചിടത്തോളം അതു് സാധ്യമല്ല. ലഘുയാദൃച്ഛിക സാമ്പിളനത്തിൽ സാമ്പിൾസമാന്തര മാധ്യത്തിൻെറ സാമ്പിളന പ്രസരണത്തിനു് അനഭിനത ആകലമുണ്ടു്. അതു്

$$\frac{1}{n-1}\sum_{i=1}^{n}(x_i-\bar{x})^2$$

എന്ന ഘടകം ഉൾക്കൊള്ളുന്നതാണെന്നും നാം കണ്ടു. $n=1$ ആണെങ്കിൽ ഈ വ്യംജകം അർഥശൂന്യമായി തീരുമെന്നതു കൊണ്ടു് സാമ്പിൾ പരിമാണം ഒന്നിൽ കൂടുതലായിരിക്കുമ്പോൾ മാത്രമേ സാമ്പിളന പ്രസരണത്തിൻെറ അനഭിനത ആകലം കണ്ടുപിടിക്കാനാവൂ. ക്രമാനുസൃതസാമ്പിളിനെ ഒററ അംഗമുള്ള ലഘുയാദൃച്ഛിക സാമ്പിളായി പരിഗണിക്കാമെന്നു് 5.2 ൽ കാണിക്കുകയുണ്ടായല്ലൊ. അതുകൊണ്ടാണു് ക്രമാനുസൃത സാമ്പിളിൽ നിന്നു ലഭിക്കുന്ന ആകലത്തിൻെറ സാമ്പിളന പ്രസരണത്തിനു് അനഭിനത ആകലം കണ്ടുപിടിക്കാനാവാത്തതു്. പക്ഷേ സ്ഥിതിഗതികൾ തീരെ നിരാശാജനകമല്ല. ഉപയോഗപ്രദങ്ങളായ ചില അഭിനത ആകലങ്ങൾ നിർദേശിക്കപ്പെട്ടിട്ടുണ്ടു്. അവയിൽ താരതമ്യേന മെച്ചപ്പെട്ടതും സാധാരണ ഉപയോഗത്തിലുള്ളതുമായ രണ്ടു് ആകലങ്ങൾ താഴെ കൊടുക്കുന്നു. ഉത്തരോത്തര അന്തര രീതിയിലുള്ള ആകലം എന്നാണു് ഇതു് വ്യവഹരിക്കപ്പെടുന്നതു്.

1) $x_1, x_2, \cdots x_n$ എന്നിവയാണു് സാമ്പിളിലെ അംഗങ്ങൾ എന്നിരിക്കട്ടെ.

$\bar{x}_{sy}$ ൻെറ സാമ്പിളന പ്രസരണത്തിൻെറ ആകലം $\hat{V}(\bar{x}_{sy})$ ആണെന്നു സങ്കല്പിക്കുക. ഈ രീതിയനുസരിച്ച്

$$\hat{V}(\bar{x}_{sy}) = \frac{N-n}{Nn}\sum_{i=1}^{n-1}\frac{(x_{i+1}-x_i)^2}{2(n-1)} \text{ ആണു്.}$$

ഇങ്ങനെ ഒരു വ്യംജകം നിർദേശിക്കപ്പെടുവാനുള്ള സാഹചര്യം കൂടി പ

രിശോധിക്കാം. സമഷ്ടി 2 k അംഗങ്ങൾ വീതമുള്ള $\frac{n}{2}$ സമുച്ചയങ്ങളായി വിഭജിക്കപ്പെട്ടിരിക്കുന്നു എന്നിരിക്കട്ടെ. (ഇവിടെ n ഇരട്ടസംഖ്യയാണെന്നു് സങ്കല്പിച്ചിരിക്കുന്നു. ഈ സങ്കല്പം താത്വികമായി തകരാറുള്ള ഒന്നല്ല. n ഒററസംഖ്യയാണെങ്കിൽ യാദൃച്ഛികമായി ഒരംഗത്തെ കൂടി ഉൾപ്പെടുത്തുക തുടങ്ങിയ ഏതെങ്കിലും മാർഗം ഉപയോഗിച്ചു് ആകലത്തെയോ അതിൻെറ പ്രസരണത്തെയോ കാര്യമായി ബാധിക്കാതെ n നെ ഇരട്ടസംഖ്യയാക്കാൻ കഴിയും.) ഈ ഓരോ സമുച്ചയങ്ങളിൽ നിന്നു് രണ്ടു് അംഗങ്ങളെ വീതമാണല്ലൊ സാമ്പിളിലേക്കു തിരഞ്ഞെടുക്കുന്നതു്. ഓരോ സമുച്ചയത്തിലെയും 2k അംഗങ്ങളെ നല്ലതുപോലെ മാററിമറിച്ചതിനു ശേഷമാണു് രണ്ടു് അംഗങ്ങളെ ക്രമാനുസൃതസമ്പ്രദായമനുസരിച്ചു് തിരഞ്ഞെടുത്തതു് എന്നു സങ്കല്പിച്ചാൽ ആ രണ്ടു അംഗങ്ങൾ 2k അംഗങ്ങളുള്ള ഒരു സമഷ്ടിയിൽ നിന്നു് എടുത്ത രണ്ടു് അംഗങ്ങളുള്ള ലഘുയാദൃച്ഛികസാമ്പിളാണെന്നു കരുതാം. i–ാമത്തെ സമുച്ചയത്തിൽ നിന്നു് എടുത്ത സാമ്പിളിലെ അംഗങ്ങൾ x_{2i}, x_{2i-1} എന്നിവയാണെന്നിരിക്കട്ടെ. അവയുടെ സമാന്തരമാധ്യം $\bar{x}_i = \frac{x_{2i} + x_{2i-1}}{2}$ ഉം

$$\text{പ്രസരണം, } \sigma_i^2 = \frac{1}{2}\left(x_{2i}^2 + x_{2i-1}^2\right) - \left(\frac{x_{2i}+x_{2i-1}}{2}\right)$$

$$= \left(\frac{x_{2i}-x_{2i-1}}{2}\right)^2$$

$\bar{x}_i$ ൻെറ സാമ്പിളനപ്രസരണത്തിൻെറ അനഭിനതആകലം

$\frac{k-1}{k}\left(\frac{x_{2i}-x_{2i-1}}{2}\right)^2$ ആയിരിക്കുമല്ലൊ. (പ്രതിസ്ഥാപനമില്ലാത്ത ലഘുയാദൃച്ഛിക സാമ്പിളിലെ $V(\bar{x})$ ൻെറ വ്യംജകത്തിൽ $N=2k$, $n=2$, $\sigma^2=\sigma_i^2$ എന്നിവ പ്രതിസ്ഥാപിച്ചാൽ കിട്ടുന്നതാണിതു്.) ലഘുയാദൃച്ഛിക സാമ്പിളനത്തിലെ നിയമമനുസരിച്ചു്

$$\bar{x}_{sy} = \frac{2}{n}\sum_{i=1}^{n/2}\left(\frac{x_{2i}+x_{2i-1}}{2}\right) \text{ ആയിരിക്കും.}$$

അതുകൊണ്ടു്, $V(\bar{x}_{sy})$ ൻെറ അനഭിനത ആകലമായ,

$$\hat{V}(\bar{x}_{sy}) = \frac{4}{n^2}\sum_{i=1}^{n/2}\hat{V}(\bar{x}_i) = \frac{k-1}{k}\cdot\frac{1}{n^2}\sum_{i=1}^{n/2}(x_{2i}-x_{2i-1})^2$$

$$= \frac{N-n}{Nn^2} \sum_{i=1}^{n/2} (x_{2i}-x_{2i-1})^2$$ ആണെന്നു് സിദ്ധിക്കുന്നു.

$$\left(\because \frac{k-1}{k} = \frac{nk-n}{nk} = \frac{N-n}{N} \right)$$

ഇവിടെ സാമ്പിളിലെ തുടർച്ചയായുള്ള ഈ രണ്ടു് അംഗങ്ങൾ വീതം യാദൃച്ഛികമായി തിരഞ്ഞെടുക്കപ്പെട്ടവയാണെന്നാണു് സങ്കല്പിച്ചിരിക്കുന്നതു്. ഈ സങ്കല്പം അൽപംകൂടി സാമാന്യവൽകരിച്ചു്, (1,2), അതുപോലെ (2, 3), (3, 4), (4, 5)... (n—1, n) എന്നീ ജോഡികൾ യാദൃച്ഛികമായി തിരഞ്ഞെടുത്തവയാണെന്നു സങ്കല്പിച്ചു്, മുൻ വിവരിച്ച സമീപനരീതി സ്വീകരിച്ചാൽ ഉത്തരോത്തര അന്തരരീതിയിൽ നിർദേശിക്കപ്പെട്ട ആകലമായ

$$\frac{N-n}{Nn} \sum_{i=1}^{n-1} \frac{(x_{i+1}-x_i)^2}{2(n-1)}$$ കിട്ടും.

ഒരു ഉദാഹരണം കാണിക്കാം. 100 അംഗങ്ങളുള്ള ഒരു സമഷ്ടിയിൽനിന്നു് 6 അംഗങ്ങളുള്ള ഒരു സാമ്പിൾ എടുത്തു എന്നിരിക്കട്ടെ. 2, 6, 5, 3, 8, 6 എന്നിവയാണു് സാമ്പിൾമൂല്യങ്ങൾ എന്നും സങ്കല്പിക്കുക.

$$\overline{x}_{sy} = \frac{2+6+5+3+8+6}{6} = 5$$

$$\hat{V}(\overline{x}_{sy}) = \frac{100-6}{100\times 6}\left[\frac{(2-6)^2+(6-5)^2+(5-3)^2+(3-8)^2+(8-6)^2}{2\times 5}\right]$$

$$= 0.78$$

സാമ്പിളന പ്രസരണത്തിന്റെ ഈ ആകലം അഭിനതമായിരിക്കണമെന്നില്ല എന്ന വസ്തുത പ്രത്യേകം ശ്രദ്ധിക്കേണ്ടതാണു്. ഇതിനെ അടിസ്ഥാനമാക്കി നിഗമനങ്ങളിൽ എത്തിച്ചേരാൻ ശ്രമിക്കുമ്പോൾ ആകലത്തിന്റെ അഭിനതിയുടെ അളവിനെപ്പററി സമാന്യജ്ഞാനമെങ്കിലും സമ്പാദിക്കാൻ ശ്രമിക്കേണ്ടതാണു്. ഇതേ സമഷ്ടിയെപ്പററി തന്നെ മുൻപു് ഏതെങ്കിലും പഠനങ്ങൾ നടത്തിയിട്ടുണ്ടെങ്കിൽ അതിൽ നിന്നു ലഭിച്ച വിവരങ്ങൾ ഇതിനു് ഉപകരിച്ചേക്കും.

(2) അന്തർവേധന ഉപ സാമ്പിൾ ഉപയോഗിച്ചുള്ള അനഭിനത ആകലം.

'n' പരിമാണമായുള്ള ഒരു സാമ്പിൾ എടുക്കുന്നതിനു പകരം സ്വതന്ത്രമായ $\left(\frac{n}{m}\right)$ പരിമാണമായുള്ള 'm' സാമ്പിളുകൾ എടുത്തു്, അവ ഓരോന്നിൽ നിന്നുമുള്ള ആകലങ്ങൾ ഉപയോഗിച്ചു് ആകെ സാമ്പിളിൽ (n പരിമാണം) നിന്നു ഒരു ആകലവും അതിന്റെ പ്രസരണത്തിന്റെ ഒരു അനഭിനത ആകലവും കണ്ടു

പിടിക്കുന്നതാണ് ഈ മാർഗം. m സ്വതന്ത്രയാദൃച്ഛരിക പ്രാരംഭങ്ങളിൽ നിന്നും 'k' സാമ്പിളന അന്തരാളമായി $\left(\frac{n}{m}\right)$ വലുപ്പമുള്ള 'm' ക്രമാനുസൃത ഉപസാമ്പിളുകൾ എടുക്കുന്നു. ഇങ്ങനെ എടുക്കുന്ന ഉപസാമ്പിളുകളിലെ സമാന്തരമാധ്യം സമഷ്ടിസമാന്തരമാധ്യത്തിന്റെ അനഭിനത ആകലമായിരിക്കുമല്ലോ. i–ാമത്തെ ഉപസാമ്പിളിലെ സമാന്തരമാധ്യം $\overline{y}_i$ എന്നിരിക്കട്ടെ. ഉപസാമ്പിളുകൾ കൂടി ചേർന്നുള്ള n പരിമാണമായുള്ള ആകെ സാമ്പിൾ ക്രമാനുസൃതസാമ്പിളുകളുടെ ഒരു സമുച്ചയമായതുകൊണ്ട്, ആ സാമ്പിളിനെ തന്നെ ഒരു ക്രമാനുസൃത സാമ്പിൾ ആയി പരിഗണിക്കാം. ആകെ സാമ്പിളിൽ നിന്നും കിട്ടുന്ന സമഷ്ടി സമാന്തര മാധ്യത്തിന്റെ ആകലത്തെ $\overline{x}$ എന്നു കുറിക്കാം.

$$\overline{x} = \frac{1}{m} \sum_{i=1}^{m} \overline{x}_i$$

ആയി എടുത്താൽ $\overline{x}$ ഉം സമഷ്ടിസമാന്തരമാധ്യത്തിന്റെ ഒരു അനഭിനത ആകലമാണെന്നു് കാണാം. കൂടാതെ

$$V(\overline{x})_{sy} = \frac{1}{m} V(x_i)_{sys}$$

എന്നും കാണാമല്ലോ. ഉപസാമ്പിളുകളുടെ സാമ്പിളന അന്തരാളം k ആയതുകൊണ്ട് സമഷ്ടിയിൽ നിന്നും $\left(\frac{n}{m}\right)$ പരിമാണമായുള്ള k ഉപസാമ്പിളുകൾ എടുക്കാമല്ലോ. ഇവയുടെ സമാന്തരമാധ്യങ്ങൾ $\overline{X}_1, \overline{X}_2, \cdots \overline{X}_k$ എന്നിരിക്കട്ടെ. നേരത്തെ പറഞ്ഞിട്ടുള്ള m ഉപസാമ്പിളുകൾ, ആകെയുള്ള k ഉപസാമ്പിളുകളിൽ നിന്നുള്ള ഒരു ലഘുയാദൃച്ഛികസാമ്പിൾ ആണ് ഉപസാമ്പിളുകൾ സ്വതന്ത്രയാദൃച്ഛിക പ്രാരംഭം സ്വീകരിക്കുന്നതുകൊണ്ട്, അവ മേൽപറഞ്ഞ k ഉപസാമ്പിളുകളുടെ സമഷ്ടിയിൽ നിന്ന് എടുക്കുന്ന പ്രതിസ്ഥാപനമില്ലാത്ത ഒരു ലഘുയാദൃച്ഛിക സാമ്പിൾ ആയി മാറുന്നു. ആയതിനാൽ $\overline{x}_1, \overline{x}_2, \ldots \overline{x}_m$ എന്നീ ഉപസാമ്പിൾസമാന്തരമാധ്യങ്ങൾ, $\overline{X}_1, \overline{X}_2 \ldots\ldots \overline{X}_k$ എന്നീ സമാന്തരമാധ്യങ്ങളുടെ സമഷ്ടിയിൽ നിന്നുള്ള ഒരു ലഘുയാദൃച്ഛരിക (പ്രതിസ്ഥാപനമില്ലാത്ത) സാമ്പിളത്രെ. ആകെ സാമ്പിളിൽ നിന്നുള്ള സമഷ്ടിസമാന്തരമാധ്യത്തിന്റെ ആകലം $\overline{x}$, $\overline{x}_1, \overline{x}_2, \ldots\ldots\overline{x}_m$ എന്നിവയുടെ സമാന്തരമാധ്യമാണ്. അതായത് $\overline{x}$ പ്രതിസ്ഥാപനമില്ലാത്ത ഒരു ലഘുയാദൃച്ഛരിക സാമ്പിളിന്റെ സാമ്പിളനസമാന്തരമാധ്യമത്രെ. മേൽ പറഞ്ഞ കാരണങ്ങൾ കൊണ്ട്,

$$V(\overline{x})_{sy} = \frac{\sigma'^2}{m}$$ എന്നു സിദ്ധിക്കുന്നു.

ഇവിടെ σ'^2 എന്നതു് $(\bar{X}_1, \bar{X}_2 \ldots\ldots \bar{X}_k)$ ൻെറ പ്രസരണമാണു്. ഈ പ്രസരണത്തിൻെറ അനഭിനത ആകലം

$$S^2 = \sum_{i=1}^{m} \frac{(\bar{x}_i - \bar{x})^2}{m-1}$$

ആണെന്നു് ഒന്നാം അധ്യായത്തിൽ കണ്ടിട്ടുണ്ടല്ലൊ. ആയതിനാൽ $V(\bar{x})_{sy}$ ന്റെ ഒരു അനഭിനത ആകലം

$$\hat{V}(\bar{x})_{sy} = \frac{\sum_{i=1}^{m} (\bar{x}_i - \bar{x})^2}{m(n-1)}$$

ആണെന്നു കാണാം.

സാമ്പിളനഅന്തരാളം (k) ചെറുതാകുമ്പോൾ സ്വതന്ത്രമായ m യാദൃച്ഛികപ്രാരംഭങ്ങളോടു കൂടിയ സാമ്പിളുകൾ എല്ലാം വ്യത്യസ്തമായിരിക്കണമെന്നില്ല. ഇങ്ങനെയുള്ള സന്ദർഭങ്ങളിൽ $\left(\frac{n}{m}\right)$ പരിമാണമുള്ള m ഉപസാമ്പിളുകളെ 1 മുതൽ $k' (= mk)$ വരെയുള്ള സംഖ്യകളിൽ യാദൃച്ഛികപ്രാരംഭമായും, പുനസ്ഥാപനമില്ലാത്തതുമായി എടുക്കുന്നതായിരിക്കും ഉത്തമം. ഇപ്രകാരമെടുക്കുന്ന ഉപസാമ്പിളുകളിൽ നിന്നു്

$$V(\bar{x})_{sys} = \frac{k'-m}{k'm} \; \frac{\sum_{i=1}^{k'} (\bar{X}_i - \bar{X})^2}{k-1}$$

എന്നും ഈ പ്രസരണത്തിൻെറ അനഭിനത ആകലത്തെ

$$\hat{V}(\bar{x})_{sy} = \frac{k'-m}{k'm} \; \frac{\sum_{i=1}^{m} (\bar{x}_i - \bar{x})^2}{m-1}$$

എന്നും കണക്കാക്കാം.

8 തുകയുടെയും അനുപാതത്തിൻെറയും ആകലനം

ഇതു വരെ നാം ചർച്ചചെയ്തതു് സമഷ്ടിയുടെ സമാന്തരമാധ്യം ആകലനം ചെയ്യാനുള്ള മാർഗങ്ങളാണു്. സാമ്പിളനലക്ഷ്യം എപ്പോഴും സമാന്തരമാധ്യത്തിൻെറ ആകലനമായിരിക്കണമെന്നില്ല. സമഷ്ടിയിലെ അംഗങ്ങളുടെ അഭിലക്ഷണമൂല്യങ്ങളുടെ തുക പലപ്പോഴും ആകലനം ചെയ്യേണ്ടതായിവരാം. അതുപോലെ തന്നെ നിർദിഷ്ടമായ ഏതെങ്കിലും പ്രത്യേകതയുള്ള അംഗങ്ങളുടെ അനുപാതം ആകലനം ചെയ്യേണ്ടതും സാമ്പിളനലക്ഷ്യമായി വരാവുന്നതാണു്. ഉദാ

ഹരണമായി ഒരു വ്യാപാരസ്ഥാപനത്തിലെ ബില്ലുകളിൽനിന്ന് ആ സ്ഥാപനം ആ വർഷം എത്ര രൂപ വിലക്കുള്ള സാധനങ്ങൾ വില്പന നടത്തി എന്നു കണ്ടുപിടിക്കേണ്ടി വരാം. സമഷ്ടിയിലെ അംഗങ്ങൾ ബില്ലുകളാണ്. ഓരോ ബില്ലിലേയും തുകയാണ് പരിഗണിക്കപ്പെടുന്ന അഭിലക്ഷണം. ഈ അഭിലക്ഷണമൂല്യങ്ങളുടെ തുകയാണല്ലോ നമുക്ക് ആകലനം ചെയ്യേണ്ടതു്. ബില്ലുകൾ തന്നെ റൊക്കം പണം കിട്ടിയവയും അല്ലാത്തവയുമാവാം. ആകെ ബില്ലുകളിൽ ഏതനുപാതം ബില്ലുകൾ റൊക്കം പണം കിട്ടിയവയാണ് എന്ന് ആകലനം ചെയ്യേണ്ടി വരാം. ഇവിടെ നിർദിഷ്ടമായ പ്രത്യേകതയുള്ള അംഗങ്ങളുടെ അനുപാതമാണ് ആകലനം ചെയ്യുന്നതു്.

സമഷ്ടിയുടെ തുകയെ ആകലനം ചെയ്യാൻ, അതായതു് $N\overline{X}$ ആകലനം ചെയ്യാൻ, $\overline{X}$–ന്റെ ആകലത്തെ N കൊണ്ടു ഗുണിച്ചാൽ മതി. ക്രമാനുസൃതസാമ്പിളനത്തിൽ $\overline{X}$ ന്റെ ആകലം $\bar{x}_{sy}$ ആണല്ലോ. അതുകൊണ്ട് $N\overline{X}$ ന്റെ ആകലം $N\bar{x}_{sy}$ ആയിരിക്കും.

$V(N\bar{x}_{sy}) = N^2\, V(\bar{x}_{sy})$ ആയതുകൊണ്ട് ഈ ആകലത്തിന്റെ സാമ്പിളനപ്രസരണം $\bar{x}_{sy}$ ന്റെ സാമ്പിളനപ്രസരണത്തെ N^2 കൊണ്ട് ഗുണിച്ചതായിരിക്കും.

നിർദിഷ്ടമായ പ്രത്യേകതയോടു കൂടിയ അംഗങ്ങളുടെ അനുപാതം ആകലം ചെയ്യാനും $\bar{x}_{sy}$ തന്നെ ഉപയോഗിക്കാം. ഇവിടെ അഭിലക്ഷണമൂല്യമായ x ഒരു പ്രത്യേകതരത്തിൽ നിർവചിക്കണമെന്നു മാത്രം. സാമ്പിൾ എടുത്തതിനു ശേഷം സാമ്പിളിലെ ഓരോ അംഗത്തെയും പരിശോധിക്കുക. ആ അംഗം നിർദിഷ്ടമായ പ്രത്യേകതകളുള്ളതാണെങ്കിൽ അതിന്റെ അഭിലക്ഷണമൂല്യം 1 എന്നും അല്ല എങ്കിൽ 0 എന്നും എടുക്കുക. അതായതു്, j–ാമത്തെ അംഗം നിർദിഷ്ടമായ പ്രത്യേകതകളുള്ളതാണെങ്കിൽ $x_i = 1$ അല്ലെങ്കിൽ $x_i = 0$. അഭിലക്ഷണമൂല്യം ഈ വിധത്തിൽ നിർവചിക്കപ്പെട്ടാൽ $\bar{x}_{sy}$, നിർദിഷ്ടമായ പ്രത്യേകതയോടു കൂടിയ അംഗങ്ങളുടെ അനുപാതത്തിന്റെ ആകലമായിരിക്കും. അതിന്റെ സാമ്പിളനപ്രസരണം $V(\bar{x}_{sy})$ ആണു താനും.

9 സാമ്പിൾപരിമാണം

ലഘുയാദൃച്ഛിക സാമ്പിളനത്തിലും സ്തരിത സാമ്പിളനത്തിലും സാമ്പിൾ പരിമാണം നിർണയിച്ചതു് ആകെ ചെലവ് ഒരു നിശ്ചിതപരിധിയിൽ കവിയരുതെന്നും ആകലത്തിന്റെ സാമ്പിളനപ്രസരണം എത്രയും ചെറുതായിരിക്കണമെന്നുമുള്ള മൗലികനിബന്ധനകളെ ആധാരമാക്കിയാണ്. ആ രണ്ടു സാമ്പിളനരീതികളിലും സാമ്പിൾപരിമാണം വർധിക്കുമ്പോൾ ആകലത്തിന്റെ സാമ്പിളനപ്രസരണം കുറയുമെന്നും നാം കണ്ടു. പക്ഷേ ക്രമാനുസൃതസാമ്പിളിന്റെ സ്ഥിതി ഇതിൽ നിന്നു വ്യത്യസ്തമാണ്. സമഷ്ടിയിലെ അംഗങ്ങൾ ഏതു ക്രമത്തിലാണ് പരിഗണിക്കപ്പെടുന്നതു് എന്നതു് സാമ്പിളനപ്രസരണത്തെ ബാധിക്കുന്ന ഒരു പ്രധാനഘടകമത്രെ. സമഷ്ടിയിലെ അംഗങ്ങളുടെ ക്രമം നിശ്ചിതമാണെങ്കിൽ സാമ്പിൾപരിമാണം വർധിക്കുന്നതനുസരിച്ച് സാമ്പിളനപ്രസരണം

കുറയും. പക്ഷേ സാമ്പിളനപ്രസരണവും സാമ്പിൾപരിമാണവും തമ്മിലുള്ള ബന്ധം ലഘുവായും വ്യക്തമായും നിർദേശിക്കുന്ന വാക്യങ്ങളൊന്നും തന്നെ നമ്മുടെ അറിവിലില്ല. അതുകൊണ്ട് ആകലത്തിനുണ്ടായിരിക്കേണ്ട സൂക്ഷ്മത തന്നിരുന്നാൽ അതായത് ആകലത്തിൻെറ സാമ്പിളനപ്രസരണം ഒരു നിർദിഷ്ട സംഖ്യയിൽ കവിയരുതെന്നു നിർദേശിച്ചിരുന്നാൽ അതു് ഉറപ്പു് വരുത്തുവാൻ ആവശ്യമായ സാമ്പിൾ പരിമാണം എത്രയായിരിക്കണമെന്നു് കൃത്യമായി കണ്ടുപിടിക്കാൻ മാർഗങ്ങളൊന്നുമില്ല. ദീർഘകാലപരിചയവും നിരവധി സാമ്പിളുകൾ പഠനം നടത്തിയതിൽ നിന്നു സിദ്ധിച്ചിട്ടുള്ള അനുഭവജ്ഞാനവും മാത്രമാണു് ഈ ഘട്ടത്തിൽ മാർഗനിർദേശം നൽകാൻ ഉപകരിക്കുക.

10. സ്തരിത ക്രമാനുസൃതസാമ്പിളനം

ക്രമാനുസൃതസാമ്പിളനം ഒരർഥത്തിൽ സ്തരിതസാമ്പിളനമാണെന്നു നാം കണ്ടുകഴിഞ്ഞു. എന്നാൽ അവ തമ്മിൽ മൗലികമായ ചില വ്യത്യാസങ്ങളുണ്ടുതാനും. ക്രമാനുസൃതസാമ്പിളനത്തിൻെറ ഏതാണ്ട് എല്ലാ വശങ്ങളും നാം ചർച്ചചെയ്തു. സ്തരിതസാമ്പിളനത്തിൻെറ പ്രത്യേകതകളും പഠിച്ചുകഴിഞ്ഞതാണു്. ഈ രണ്ടു സാമ്പിളനരീതികളെയും സമഞ്ജസമായി സമ്മേളിപ്പിച്ചുകൊണ്ടുള്ള ഒരു സാമ്പിളന രീതിയാണു് ഇവിടെ ചർച്ച ചെയ്യുന്നതു്.

സമഷ്ടിയിൽ N അംഗങ്ങളാണു് ഉള്ളതു്. അതിനെ യഥാക്രമം $N_1, N_2 \ldots N_h$ അംഗങ്ങൾ വീതമുള്ള h സ്തരങ്ങളായി വിഭജിച്ചിരിക്കുന്നു എന്നിരിക്കട്ടെ. n അംഗങ്ങളുള്ള സാമ്പിളാണു് എടുക്കേണ്ടതെന്നുംകരുതുക. സ്തരിതസാമ്പിളനത്തിൽ പരിഗണിച്ച ഏതെങ്കിലും തത്വമുപയോഗിച്ചു് ഓരോ സ്തരത്തിൽ നിന്നും എത്ര അംഗങ്ങളെ വീതമാണു് സാമ്പിളിൽ ഉൾപ്പെടുത്തേണ്ടതു് എന്നു തീരുമാനിക്കുക. j—ാമത്തെ സ്തരത്തിൽ നിന്നു് n_j അംഗങ്ങളെയാണു് സാമ്പിളിൽ ഉൾപ്പെടുത്തുന്നതെന്നിരിക്കട്ടെ. ഈ n_j അംഗങ്ങളെ j—ാമത്തെ സ്തരത്തിൽനിന്നു് ക്രമാനുസൃതസാമ്പിളനസമ്പ്രദായത്തിലാണു് തിരഞ്ഞെടുക്കുന്നതെങ്കിൽ, നമുക്കു കിട്ടുന്ന സാമ്പിൾ സ്തരിത ക്രമാനുസൃതസാമ്പിളനരീതിയിലുള്ളതാണു്. $\bar{x}_{jsy}$ യാണു് j—ാമത്തെ സ്തരത്തിൽ നിന്നു് തിരഞ്ഞെടുത്ത അംഗങ്ങളുടെ അഭിലക്ഷണ മൂല്യങ്ങളുടെ സമാന്തരമാധ്യമെങ്കിൽ X ൻെറ ആകലം,

$$x_{stsy} = \frac{\sum_{j=1}^{h} N_j \bar{x}_{jsy}}{N} \qquad 5.10.1$$

ആയിരിക്കും. ഇതു് ഒരു അനഭിനത ആകലമാണെന്നു കാണിക്കാൻ പ്രയാസമില്ല. മാത്രമല്ല,

$$V(\bar{x}_{stsy}) = \frac{\sum_{j=1}^{h} N_j^2 V(\bar{x}_{jsy})}{N^2}$$

ആണെന്നും എളുപ്പത്തിൽ കാണിക്കാൻ കഴിയും.

ഈ സാമ്പിളനപ്രസരണത്തിന്റെ ഒരു ആകലം കാണുന്നതിനു് വിവിധസ്തരങ്ങളിൽ നിന്നു് $V(\bar{x}_{jsy})$ ന്റെ ആകലങ്ങൾ കണ്ടു് ഈ വ്യംജകത്തിൽ പ്രതിസ്ഥാപിച്ചാൽ മതിയാവും. $V(\bar{x}_{jsy})$ കാണാൻ മുൻവിവരിച്ച ഉത്തരോത്തര അന്തര രീതി സ്വീകരിക്കാവുന്നതാണു്.

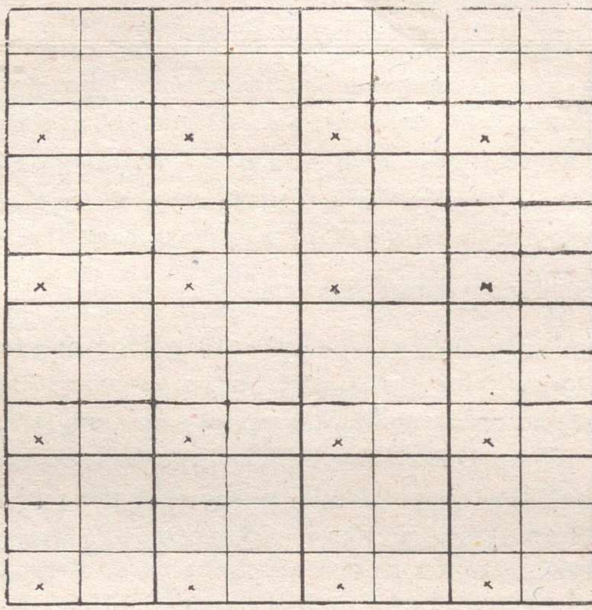
ഓരോ സ്തരങ്ങളിൽ നിന്നും ലഘുയാദൃച്ഛികസാമ്പിളനരീതിയിൽ ആവശ്യമുള്ള അംഗങ്ങളെ തിരഞ്ഞെടുക്കുക എന്നതാണല്ലോ സ്തരിതസാമ്പിളനത്തിൽ ചെയ്യുന്നതു്. സ്തരങ്ങളിലെ അംഗങ്ങളുടെ വിന്യാസരീതിയനുസരിച്ചു് ചിലപ്പോൾ ക്രമാനുസൃതസാമ്പിളനം ലഘുയാദൃച്ഛിക സാമ്പിളനത്തെക്കാൾ സൂക്ഷ്മതയുള്ളതായി വരാം. അങ്ങനെയുള്ള സന്ദർഭങ്ങളിൽ സ്തരിതക്രമാനുസൃതസാമ്പിളനം സ്തരിതസാമ്പിളനത്തിനെക്കാൾ മെച്ചപ്പെട്ടതായിരിക്കും.

11. ദ്വിമാന ക്രമാനുസൃതസാമ്പിളനം

ഇതുവരെ നാം പരിഗണിച്ച സമഷ്ടികൾ ഏകമാനങ്ങളായിരുന്നു. അതായതു് ഒരു ക്രമസംഖ്യ കൊണ്ടു് അവയിൽ ഓരോന്നിനെയും നമുക്കു് സൂചിപ്പിക്കാമായിരുന്നു. മറ്റൊരു തരത്തിൽ പറഞ്ഞാൽ, സമഷ്ടിയിലെ അംഗങ്ങളെ ഒരു ഋജുരേഖയിലെ ബിന്ദുക്കളെക്കൊണ്ടു് സൂചിപ്പിക്കാമായിരുന്നു. എല്ലാ സമഷ്ടികളും ഈ രീതിയിലുള്ളവയായിരിക്കണമെന്നില്ല. ഉദാഹരണമായി ഒരു നെൽപ്പാടത്തെ 10 സെന്റു് വിസ്തീർണവും ദീർഘചതുരാകൃതിയുമുള്ള ഭാഗങ്ങളായി തിരിച്ചിരിക്കുന്നു എന്നു വിചാരിക്കുക. നെൽപ്പാടം ദീർഘചതുരാകൃതിയാണെന്നും അതിന്റെ വശങ്ങളോടു് സമാന്തരമായ രേഖകൾ കൊണ്ടാണു് വിഭജനം നടത്തിയിരിക്കുന്നതെന്നും സങ്കല്പിക്കുക. l വരികളും m കോളങ്ങളുമാണു് ഉള്ളതെന്നും ഇരിക്കട്ടെ. ഓരോ ഭാഗത്തെയും (i,j) എന്ന ക്രമബദ്ധസംഖ്യാജോഡികൾ കൊണ്ടു് സൂചിപ്പിക്കാം. i–ാമത്തെ വരിയിൽ j–ാമത്തെ കോളത്തിൽ വരുന്ന ഭാഗമാണു് (i,j) സൂചിപ്പിക്കുന്നതു്. ഈ സമഷ്ടി ഒരു ദ്വിമാനസമഷ്ടിയാണു്. ഇമ്മാതിരി സമഷ്ടികളിൽ നിന്നു് സാമ്പിളെടുക്കാനാണു് ദ്വിമാനക്രമാനുസൃത സാമ്പിളനം ഉപയോഗിക്കുന്നതു്. ഈ ഓരോ ഭാഗത്തിനും ക്രമസംഖ്യകൾ കൊടുത്താൽ ഇതിനെ ഏകമാനസമഷ്ടിയാക്കാൻ കഴിയും. പക്ഷേ ദ്വിമാനസമഷ്ടിയായി തന്നെ പരിഗണിച്ചു് സാമ്പിളെടുക്കുന്നതാണു് കൂടുതൽ പ്രാതിനിധ്യമുള്ള സാമ്പിൾ ലഭിക്കാൻ ഉപകരിക്കുന്നതു്.

ഇനി ദ്വിമാനക്രമാനുസൃതസാമ്പിൾ എങ്ങനെയാണു് എടുക്കുക എന്നു പരിശോധിക്കാം. സമഷ്ടിയിൽ N അംഗങ്ങളാണു് ഉള്ളതെന്നിരിക്കട്ടെ. n അംഗങ്ങളുള്ള സാമ്പിളാണു് എടുക്കേണ്ടതെന്നും സങ്കല്പിക്കുക. $N=nk$ എന്നും $k=lm$ എന്നും സൗകര്യത്തിനു വേണ്ടി കരുതാം. അപ്പോൾ l വരികളും m കോളങ്ങളും ഉള്ള n ദീർഘചതുരങ്ങളായി, തന്നിരിക്കുന്ന ദ്വിമാനസമഷ്ടിയെ വിഭജിക്കാം. $i \leq l$, $j \leq m$ ആയിരിക്കത്തക്കവണ്ണം i, j എന്നു രണ്ടു് യാദൃച്ഛിക സംഖ്യകൾ കണ്ടുപിടിക്കുക. (i, j) സൂചിപ്പിക്കുന്ന അംഗങ്ങളെ സാമ്പിളിലെ അംഗങ്ങളായി സ്വീകരിക്കുക. l വരികളും m കോളങ്ങളുമുള്ള n വിഭാഗങ്ങ

ളാണല്ലോ ഉള്ളതു്. ഈ ഓരോ വിഭാഗത്തു നിന്നും (i,j) എന്ന ക്രമബദ്ധസംഖ്യാ ജോഡി സൂചിപ്പിക്കുന്ന അംഗത്തെ സാമ്പിളിൽ ഉൾപ്പെടുത്തുക എന്നു സാരം. ഇങ്ങനെ ലഭിക്കുന്ന സാമ്പിളാണു് ദ്വിമാനക്രമാനുസൃതസാമ്പിൾ. താഴെ കൊടുത്തിരിക്കുന്ന ചിത്രത്തിൽ 'x' ചിഹ്നം തിരഞ്ഞെടുക്കപ്പെട്ട അംഗങ്ങളെ സൂചിപ്പിക്കുന്നു.



ചിത്രം 5.3

$k=6, l=3, m=2, n=16, N=96$

ഈ ചിത്രത്തിൽ കാണിച്ചിരിക്കുന്നതു് 4 മീററർ നീളവും 4 മീററർ വീതിയുമുള്ള സമചതുരാകൃതിയിലുള്ള ഒരു കൃഷിസ്ഥലമാണു്. അതു് ഒരു മീററർ നീളവും വീതിയുമുള്ള 16 ചതുരങ്ങളായി വിഭജിച്ചു, ഓരോ ചതുരത്തെയും $\frac{1}{3}$ മീററർ വീതിയും $\frac{1}{2}$ മീററർ നീളവുമുള്ള ചെറുഭാഗങ്ങളായും ഭാഗിച്ചു. ഓരോ ചെറിയ ഭാഗത്തും തുല്യ എണ്ണം പയർവിത്തുകൾ പാകി. ഓരോ ഭാഗത്തും ശരാശരി എത്ര വിത്തുവീതം മുളക്കുന്നുണ്ടെന്നു കണ്ടുപിടിക്കുകയാണു് സാമ്പിളനലക്ഷ്യം. 16 അംഗങ്ങളുള്ള സാമ്പിളാണു് എടുക്കേണ്ടതു്. അതായതു് 1 മീററർ ചതുരങ്ങളിൽ ഓരോന്നിൽ നിന്നും ഓരോ ചെറിയ ഭാഗം തിരഞ്ഞെടുക്കണം. ഇതിനായി $1 \leq i \leq 3$, $1 \leq j \leq 2$ ആയിരിക്കത്തക്കവണ്ണം i,j എന്നു് രണ്ടു് യാദൃച്ഛിക സംഖ്യകളെടുത്തു. (1,1) എന്ന സംഖ്യകളാണു് കിട്ടിയതു്. ഓരോ ചതുരത്തിലും ഈ സംഖ്യാജോഡി സൂചിപ്പിക്കുന്ന ചെറുഭാഗങ്ങളെ, അതായതു് സമചതുരങ്ങളെ മുഴുവനായെടുത്തു (1,1), (1,3), (1,5), (1,7), ... (4,1) (4,3) ... (10,1) (10,3), (10,5) (10,7), എന്നീ സംഖ്യാജോഡികൾ സൂചിപ്പിക്കുന്ന ചെറുഭാഗങ്ങളെ സാമ്പിളിൽ ഉൾപ്പെടുത്തിയിരിക്കുന്നു.

ഈ ചിത്രം പരിശോധിച്ചാൽ സാമ്പിളിലെ അംഗങ്ങൾ സമഷ്ടിയുടെ എല്ലാ

ഭാഗത്തുമായി ചിതറിക്കിടക്കുന്നെന്നും അങ്ങനെ സമഷ്ടിയെ പ്രതിനിധീകരിക്കുന്നില്ലെന്നും വന്നും കാണാൻ വിഷമമില്ല. സാമ്പിളിലെ അംഗങ്ങൾക്ക് ക്രമസംഖ്യകൾ കൊടുത്ത് ഏകമാനസാമ്പിളെടുത്താൽ ഈ രീതിയിലുള്ള പ്രാതിനിധ്യസ്വഭാവം സാമ്പിളിന് ലഭിച്ചെന്നു വരികയില്ല.

ഇമ്മാതിരി സാമ്പിളുകളിൽ നിന്ന് സമഷ്ടിയുടെ സമാന്തരമാധ്യവും മററും ആകലനം ചെയ്യാനും അവയുടെ സാമ്പിളനപ്രസരണം കണ്ടുപിടിക്കുവാനും ഏകമാനസാമ്പിളനത്തിൽ ഉപയോഗിച്ച മാർഗങ്ങൾ തന്നെ അല്പസ്വല്പം ഭേദഗതികളോടെ ഉപയോഗപ്പെടുത്താവുന്നതാണ്.

12 ഒന്നിലധികം അഭിലക്ഷണങ്ങൾ

ഇതു വരെ നാം പരിഗണിച്ചതു് സമഷ്ടിയിലെ അംഗങ്ങളെ സംബന്ധിക്കുന്ന ഏതെങ്കിലും ഒരു അഭിലക്ഷണത്തെപ്പററി പഠിക്കുവാനുള്ള മാർഗങ്ങളാണ്. ക്രമാനുസൃതസാമ്പിളനത്തിന്റെ സൂക്ഷ്മത വർധിപ്പിക്കുവാനുള്ള ഒരു മാർഗം സമഷ്ടിയിലെ അംഗങ്ങളെ അഭിലക്ഷണമൂല്യങ്ങളുടെ അടിസ്ഥാനത്തിൽ ആരോഹണക്രമത്തിലോ അവരോഹണക്രമത്തിലോ വിന്യസിച്ചിട്ട് അതിൽ നിന്ന് സാമ്പിളെടുക്കുക എന്നതാണ്. അഭിലക്ഷണമൂല്യങ്ങൾ അജ്ഞാതങ്ങളാണെങ്കിൽ അതുമായി നേരെ ബന്ധപ്പെട്ടിട്ടുള്ള മറേറതെങ്കിലും സഹായകഅഭിലക്ഷണത്തെ അടിസ്ഥാനമാക്കി വേണം അംഗങ്ങളുടെ ക്രമം നിശ്ചയിക്കുവാൻ. പക്ഷേ ഒന്നിലധികം അഭിലക്ഷണങ്ങൾ പരിഗണിക്കേണ്ടിവരുമ്പോൾ ഇതു് അത്ര എളുപ്പമുള്ള ഒരു കാര്യമല്ലാതായി തീരുന്നു. ഉദാഹരണമായി രണ്ടു് അഭിലക്ഷണങ്ങളെപ്പററിയാണ് പഠനം നടത്തുന്നതെന്നിരിക്കട്ടെ. ഈ രണ്ടു് അഭിലക്ഷണങ്ങളോടും നേരെ ബന്ധപ്പെട്ടിട്ടുള്ള ഒരു സഹായക അഭിലക്ഷണമുണ്ടെങ്കിൽ സംഗതി എളുപ്പമാണ്. ഫാക്ടറികളാണ് പഠനവിധേയമാക്കപ്പെട്ടിരിക്കുന്ന സമഷ്ടി എന്നിരിക്കട്ടെ. അവയുടെ വാർഷിക ഉൽപ്പാദനവും, കൂലിയായി കൊടുക്കുന്ന തുകയുമാണ് പരിഗണിക്കപ്പെടുന്ന അഭിലക്ഷണങ്ങളെന്നും സങ്കല്പിക്കുക. ഇവ രണ്ടും തൊഴിലാളികളുടെ എണ്ണവുമായി നേരെ ബന്ധപ്പെട്ടിട്ടുള്ള അഭിലക്ഷണങ്ങളാണ്. അതുകൊണ്ടു് തൊഴിലാളികളുടെ എണ്ണത്തിന്റെ അടിസ്ഥാനത്തിൽ ആരോഹണക്രമത്തിലോ അവരോഹണക്രമത്തിലോ ഫാക്ടറികളെ പരിഗണിച്ച് സാമ്പിളെടുത്താൽ രണ്ടു് അഭിലക്ഷണങ്ങളെപ്പററിയും പഠനം നടത്താൻ പററിയ സൂക്ഷ്മതയുള്ള സാമ്പിളായിരിക്കും ലഭിക്കുക. പക്ഷേ പലപ്പോഴും അങ്ങനെയൊരു സഹായക അഭിലക്ഷണം ഉണ്ടായെന്നു വരികയില്ല. കർഷകകുടുംബങ്ങളാണ് പഠനവിധേയമാക്കപ്പെടുന്ന സമഷ്ടി എന്നിരിക്കട്ടെ. അംഗങ്ങളുടെ എണ്ണം, വരുമാനം എന്നീ രണ്ടു് അഭിലക്ഷണങ്ങളാണ് പരിഗണിക്കപ്പെടുന്നതു് എന്നും സങ്കല്പിക്കുക. ഈ രണ്ടു് അഭിലക്ഷണങ്ങളോടും നേരെ ബന്ധപ്പെട്ട ഒരു സഹായകഅഭിലക്ഷണം കണ്ടുപിടിക്കാൻ പ്രയാസമാണ്. മുൻസെൻസസ്സിൽ രേഖപ്പെടുത്തിയിട്ടുള്ള പ്രസക്തമായ വിവരങ്ങൾ അംഗസംഖ്യയോടു ബന്ധപ്പെട്ട ഒരു അഭിലക്ഷണമാണ്. പക്ഷേ അതും വരുമാനവുമായി വലിയ ബന്ധമൊന്നും ഉണ്ടായിരിക്കണമെന്നില്ല. ഓരോ കുടുംബത്തിന്റെയും അധീനതയിലുള്ള കൃഷിഭൂമിയുടെ അളവു് വരുമാനവുമായി ബന്ധപ്പെട്ട ഒരു സഹായകഅഭിലക്ഷണ

മാണ്. ഇങ്ങനെ വരുമ്പോൾ അംഗസംഖ്യയെപ്പററി പഠനം നടത്താൻ മുൻ സെൻസസ്സിൽ രേഖപ്പെടുത്തിയിട്ടുള്ള അംഗസംഖ്യയുടെ അടിസ്ഥാനത്തിൽ ആരോഹണ ക്രമത്തിലോ അവരോഹണക്രമത്തിലോ സമഷ്ടിയിലെ അംഗങ്ങളെ വിന്യസിച്ചിട്ടു വേണം സാമ്പിളെടുക്കുവാൻ. ഈ സാമ്പിൾ വരുമാനത്തെപ്പററി പഠനം നടത്താൻ ഉപകരിക്കുകയില്ല. അതിന്ന് ഓരോ കുടുംബവും കൈവശം വെച്ചിരിക്കുന്ന കൃഷിഭൂമിയുടെ അളവനുസരിച്ച് സമഷ്ടിയിലെ അംഗങ്ങളെ വിന്യസിച്ചിട്ട് അതിൽ നിന്ന് ഒരു സാമ്പിളെടുക്കണം. ചുരുക്കത്തിൽ രണ്ട് അഭിലക്ഷണങ്ങളെപ്പററി പഠിക്കുവാൻ രണ്ട് സാമ്പിളുകൾ ആവശ്യമാണ്. ഇതു് സാമ്പിളന ച്ചെലവ് ഇരട്ടിയാക്കുന്നു. കഴിവതും ഒഴിവാക്കേണ്ട ഒരു സാഹചര്യമാണ് ഇതു്. ഈ അപകടത്തിൽ നിന്ന് രക്ഷനേടാൻ മാർഗങ്ങളെന്തെങ്കിലുമുണ്ടോ എന്ന് പരിശോധിക്കേണ്ടതു് ആവശ്യമത്രെ.

ഒരു പരിഹാരനിർദ്ദേശം, പടിപ്പടിയായി ഈരണ്ട് അംഗങ്ങൾ വീതമുള്ള സാമ്പിളുകൾ എടുക്കുന്ന ഒരു സമ്പ്രദായമാണ്. ഇതു് രണ്ട് അഭിലക്ഷണങ്ങളെപ്പററി മാത്രം പഠനം നടത്തുമ്പോഴും സാമ്പിൾ പരിമാണം 2P എന്ന രൂപത്തിലുള്ളതായിരിക്കുമ്പോഴും മാത്രമേ ഉപയോഗിക്കാനാവൂ. ആദ്യമായി സമഷ്ടിയെ ഒരഭിലക്ഷണത്തോടു ബന്ധപ്പെട്ട സഹായക അഭിലക്ഷണത്തെ അടിസ്ഥാനമാക്കി ആരോഹണക്രമത്തിൽ വിന്യസിക്കുക. ഇതിൽ നിന്ന് 2 അംഗങ്ങളുള്ള എല്ലാ ക്രമാനുസൃത സാമ്പിളുകളും എടുക്കുക. അടുത്തതായി ഈ അംഗജോഡികളെ രണ്ടാമത്തെ അഭിലക്ഷണത്തോടു ബന്ധപ്പെട്ട സഹായക അഭിലക്ഷണത്തെ അടിസ്ഥാനമാക്കി ആരോഹണക്രമത്തിൽ വിന്യസിച്ച് അതിൽ നിന്ന് രണ്ട് അംഗജോഡികൾ വീതമുള്ള (4 അംഗങ്ങളുള്ള) എല്ലാ ക്രമാനുസൃത സാമ്പിളുകളും എടുക്കുക. ഈ സാമ്പിളുകളെ ആദ്യത്തെ സഹായക അഭിലക്ഷണത്തെ അടിസ്ഥാനമാക്കി ആരോഹണക്രമത്തിൽ വിന്യസിച്ച് 2 അംഗങ്ങളുള്ള (ഫലത്തിൽ 8 അംഗങ്ങളുള്ള) ക്രമാനുസൃതസാമ്പിളുകൾ പരിഗണിക്കുക. ഈ പരിപാടി സാമ്പിൾപരിമാണം നിർദിഷ്ടമായ 2P യാവുന്നതുവരെ തുടരുക. അങ്ങനെ ലഭിക്കുന്ന സാമ്പിളുകളിൽ നിന്ന് യാദൃച്ഛികമായി ഒന്നു തിരഞ്ഞെടുക്കുക. ഇതാണ് സമ്പ്രദായം. ഇവിടെ അവസാനം തിരഞ്ഞെടുക്കപ്പെടുന്ന സാമ്പിളിൽ നിന്ന് മാത്രമേ വിവരം ശേഖരിക്കുന്നുള്ളു എന്നതുകൊണ്ട് സാമ്പിളനച്ചെലവിൽ കാര്യമായ വർധനവു വരുന്നില്ല. സാമ്പിളെടുക്കുന്ന നടപടിക്രമം അല്പം സങ്കീർണ്ണമാണെന്നു മാത്രം. രണ്ടിലധികം അഭിലക്ഷണങ്ങളെപ്പററി പഠനം നടത്തേണ്ടിവരുമ്പോഴും സാമ്പിൾ പരിമാണം 2P യുടെ രൂപത്തിലല്ലാതെ വരുമ്പോഴും ഈ രീതിയിൽ ആവശ്യമായ മാററങ്ങൾ വരുത്തണം. ഈ സമ്പ്രദായത്തിൻെറ താത്വികപശ്ചാത്തലത്തിന് മാററമൊന്നും വരുത്താതെ അമ്മാതിരി സന്ദർഭങ്ങളിലേക്ക് ഉതകത്തക്കവണ്ണം രൂപം കൊടുത്തിട്ടുള്ള പല സാമ്പിളന സമ്പ്രദായങ്ങളും പ്രയോഗത്തിലുണ്ട്.

13. ക്രമാനുസൃതസാമ്പിളനത്തിൽ വരുത്താവുന്ന ചില പരിഷ്കാരങ്ങൾ

ക്രമാനുസൃതസാമ്പിളനം വളരെ കരുതലോടെ വേണം ഉപയോഗിക്കാൻ എന്ന് ഇതുവരെയുള്ള ചർച്ചകളിൽ നിന്ന് വ്യക്തമായിട്ടുണ്ടല്ലോ. പല ശാസ്ത്ര

ജ്ഞന്മാരും ക്രമാനുസൃതസാമ്പിളനത്തിൻെറ ഉപയോഗത്തെ നിരുത്സാഹപ്പെടുത്തുന്നവരാണ്. സമഷ്ടിയിൽ ആവർത്തകത ലീനമായിരുന്നേക്കാമെന്ന ഭയമാണ് അതിനു പ്രധാന കാരണം. സമഷ്ടിയെപ്പററി ആദ്യം തന്നെ സമാന്യമായ ഒരന്വേഷണം നടത്തിയാൽ ഈ അപകടം മിക്കവാറും ഒഴിവാക്കാൻ കഴിയും. പക്ഷേ സമഷ്ടിയെപ്പററിയുള്ള നമ്മുടെ അറിവ് അപൂർണമായിരിക്കുമ്പോൾ ക്രമാനുസൃതസാമ്പിളന രീതിയിൽ ചില പരിഷ്ക്കാരങ്ങൾ വരുത്തുന്നതു് അഭികാമ്യമാണ്. അങ്ങനെയുള്ള ചില പരിഷ്ക്കാരനിർദേശങ്ങൾ താഴെ കൊടുക്കുന്നു.

(a) സമഷ്ടിയിലെ അംഗങ്ങളുടെ ക്രമം യാദൃച്ഛികമാക്കിയതിനു ശേഷം സാമ്പിളെടുക്കുക.

സമഷ്ടിയിലെ അംഗങ്ങൾക്ക് ക്രമസംഖ്യകൾ കൊടുത്തതിനുശേഷം, ഒരു യാദൃച്ഛികസംഖ്യാ പട്ടിക ഉപയോഗിച്ച് അംഗങ്ങളെ യാദൃച്ഛികസംഖ്യകളുടെ ക്രമത്തിൽ പുനർവിന്യസിക്കുക. അങ്ങനെ ക്രമീകരിച്ച സമഷ്ടിയിൽ നിന്നു് ക്രമാനുസൃതസാമ്പിൾ എടുക്കുക. ഇതാണ് ഈ സമ്പ്രദായം. പ്രായോഗികമായി നോക്കിയാൽ പറയത്തക്ക യാതൊരു പ്രയോജനവുമില്ലാത്ത ഒരു പരിഷ്ക്കാരമാണ് ഇതു്. സമഷ്ടിയിൽ ഉണ്ടായിരിക്കാവുന്ന ആവർത്തകതയെ ഒഴിവാക്കാൻ ഈ പരിഷ്ക്കാരം ഉപകരിക്കുമെന്നതു് ശരിതന്നെ. പക്ഷേ ക്രമാനുസൃതസാമ്പിളനത്തിൻെറ ഏററവും വലിയ വൈശിഷ്ട്യമായ പ്രായോഗികസൗകര്യം നഷ്ടപ്പെട്ടുപോകുന്നു. മാത്രമല്ല രേഖീയത പോലുള്ള പ്രയോജനപ്രദമായ പ്രവണതകളുടെ ഗുണം അനുഭവപ്പെടാതെ വരികയും ചെയ്യും. ചുരുക്കത്തിൽ ഈ പരിഷ്ക്കാരം ക്രമാനുസൃതസാമ്പിളനത്തെ ലഘുയാദൃച്ഛിക സാമ്പിളനത്തെക്കാൾ ഒട്ടും മെച്ചമല്ലാത്ത ഒരു സമ്പ്രദായമാക്കിതീർക്കും.

(b) യാദൃച്ഛിക പ്രാരംഭത്തിൽ മാററം വരുത്തുക.

ക്രമാനുസൃതസാമ്പിളനത്തിൽ സമഷ്ടിയെ k അംഗങ്ങൾ വീതമുള്ള സമുച്ചയങ്ങളായി വിഭജിച്ചതിനു ശേഷം ഓരോ സമുച്ചയത്തിൽ നിന്നും ഓരോ അംഗത്തെ വീതം തിരഞ്ഞെടുക്കുകയാണല്ലോ ചെയ്യുന്നതു്. 1 മുതൽ k വരെയുള്ള സംഖ്യകളിൽ നിന്നു് ഒരു സംഖ്യ യാദൃച്ഛികമായി തിരഞ്ഞെടുത്തു് അതു് യാദൃച്ഛിക പ്രാരംഭമായി സ്വീകരിക്കുന്നു. അതു മുതൽ k അകലത്തിലുള്ള അംഗങ്ങളെ സാമ്പിളിൽ ഉൾപ്പെടുത്തുന്നു. ഇതിൻെറ ഫലമായി ആവർത്തകതയുള്ള ഒരു സമഷ്ടിയിൽ നിന്നാണ് സാമ്പിളെടുക്കുന്നതെങ്കിൽ ഒരേതരത്തിലുള്ള അംഗങ്ങളായിരിക്കും സാമ്പിളിൽ ഉൾപ്പെട്ടു വരിക. ഈ അപകടം ഒഴിവാക്കാനുള്ള ഒരു മാർഗമാണ് ഈ പരിഷ്ക്കാരം. ആദ്യം ഒരു യാദൃച്ഛികപ്രാരംഭത്തിൽ തുടങ്ങി ഏതാനും അംഗങ്ങളെ സാമ്പിളിലേക്കു തിരഞ്ഞെടുക്കുക. അടുത്ത സമുച്ചയത്തിൽ മറെറാരു യാദൃച്ഛികപ്രാരംഭം സ്വീകരിച്ച് തുടർന്നുള്ള സമുച്ചയങ്ങളിൽ നിന്നു് അംഗങ്ങളെ തിരഞ്ഞെടുക്കുക. ഈ രീതി ആവർത്തിക്കുക. ആവർത്തകതയിൽ നിന്നു് ഉളവാകാവുന്ന അപകടം ഒരളവു വരെ ഒഴിവാക്കാൻ ഈ രീതി സഹായിക്കും. ഒരുദാഹരണം കാണിക്കാം. സമഷ്ടിയിൽ 100 അംഗ

ങ്ങളാണ് ഉള്ളതെന്നിരിക്കട്ടെ. 20 അംഗങ്ങളുള്ള ഒരു സാമ്പിളാണ് എടുക്കേണ്ടതെന്നും സങ്കല്പിക്കുക. ഇവിടെ 5 ആണ് അന്തരാളം. 5 അംഗങ്ങൾ വീതമുള്ള 20 സമുച്ചയങ്ങളായി സമഷ്ടിയെ വിഭജിക്കുന്നു. 1 മുതൽ 5 വരെയുള്ള സംഖ്യകളിൽ നിന്ന് യാദൃച്ഛികമായി ഒന്നു തിരഞ്ഞെടുക്കുന്നു. 2 ആണ് കിട്ടിയ സംഖ്യ എന്നിരിക്കട്ടെ. 2, 7, 12, 17, 22 എന്നീ ക്രമസംഖ്യകളുള്ള അംഗങ്ങളെ ആദ്യത്തെ അഞ്ചു സമുച്ചയങ്ങളിൽ നിന്നു തിരഞ്ഞെടുക്കുന്നു. അടുത്തതായി യാദൃച്ഛികപ്രാരംഭം മാററുന്നു. 1 മുതൽ 5 വരെയുള്ള സംഖ്യകളിൽ നിന്ന് യാദൃച്ഛികമായി മറെറാരു സംഖ്യ തിരഞ്ഞെടുക്കുന്നു. 3 ആണ് കിട്ടിയ സംഖ്യ എന്നിരിക്കട്ടെ. 6–ാമത്തെ സമുച്ചയത്തിലെ 3–ാമത്തെ അംഗത്തെ അതായതു് 28 ക്രമസംഖ്യയായുള്ള അംഗത്തെ സാമ്പിളിൽ ഉൾപ്പെടുത്തുന്നു. തുടന്ന് 33, 38, 43, 48 എന്നീ ക്രമസംഖ്യകളുള്ള അംഗങ്ങളെ സാമ്പിളിൽ ഉൾപ്പെടുത്തുക വീണ്ടും യാദൃച്ഛികപ്രാരംഭം മാററുന്നു. അങ്ങനെ മുമ്പോട്ടു പോവുന്നു. ഇതാണ് ഈ രീതി.

(c) പല സാമ്പിളുകളെടുത്തു് ഒന്നിച്ചു ചേർക്കുക. (അന്തർവേധന ഉപസാമ്പിളുകൾ)

സാമ്പിളന അന്തരാളം k യാണല്ലോ. അതിനു പകരം mk സാമ്പിളനാന്തരാളമായി എടുക്കുക. ഇവിടെ സൗകര്യപ്രദമായ ഏതെങ്കിലും ഒരു സംഖ്യയെ m ൻെറ മൂല്യമായി സ്വീകരിക്കാവുന്നതാണ്. ഒരു യാദൃച്ഛികപ്രാരംഭം എടുക്കുന്നതിനു പകരം m പ്രാരംഭങ്ങൾ എടുക്കുക. അവ ഓരോന്നിനേയും അടിസ്ഥാനമാക്കി ഓരോ ക്രമാനുസൃതസാമ്പിൾ എടുക്കുക. അവയെയെല്ലാംകൂടി ഒന്നിച്ചുചേർത്താൽ നമുക്ക് ആവശ്യമുള്ള സാമ്പിൾ ലഭിക്കും. മുൻ ഉദാഹരണത്തിൽ 5 സാമ്പിളനഅന്തരാളമായി സ്വീകരിക്കുന്നതിനു പകരം $5\times4=20$ അന്താരാളമായി സ്വീകരിച്ചു എന്നിരിക്കട്ടെ. 4 യാദൃച്ഛികപ്രാരംഭങ്ങളെ അടിസ്ഥാനമാക്കി 5 അംഗങ്ങൾ വീതമുള്ള 4 ക്രമാനുസൃതസാമ്പിളുകൾ നമുക്ക് എടുക്കാൻ കഴിയും. ഈ എല്ലാ സാമ്പിളുകളും കൂട്ടിച്ചേർത്താൽ 20 അംഗങ്ങളുള്ള ഒരു സാമ്പിൾ ലഭിക്കും. ഈ പരിഷ്കാരവും സമഷ്ടിയിൽ ഉണ്ടായിരിക്കാവുന്ന ആവർത്തകതകളിൽ നിന്നുള്ള അപകടം ഒഴിവാക്കാൻ ഉദ്ദേശിച്ചു കൊണ്ടുള്ളതത്രെ. കൂടാതെ ഈ പരിഷ്ക്കാരം കൊണ്ടു് 5.7 ൽ കാണിച്ചതുപോലെ സാമ്പിൾ സമാന്തരമാധ്യത്തിൻെറ പ്രസരണത്തിന് ഒരു അനഭിനത ആകലവും കാണാൻ കഴിയും.

ഏതായാലും ഒരു കാര്യം ഓർമയിരിക്കണം. ക്രമാനുസൃത സാമ്പിളിനു പല പ്രായോഗികസൗകര്യങ്ങളും ഉണ്ടെങ്കിലും വളരെ കരുതി വേണം അതു കൈകാര്യം ചെയ്യാൻ. സമഷ്ടിയെപ്പററി പ്രാരംഭാന്വേഷണങ്ങളെല്ലാം നടത്തി ക്രമാനുസൃതസാമ്പിളനം ഉപയോഗിക്കുന്നതു കൊണ്ടു് തരക്കേടൊന്നുമില്ലെന്ന് ബോധ്യമായതിനുശേഷമേ അതു് ഉപയോഗിക്കാവൂ. അല്ലെങ്കിൽ വളരെ തെററായ നിഗമനങ്ങളിലേക്കു് അതു് നമ്മെ നയിച്ചു എന്നു വരാം.

14. സംഗ്രഹം

1. യാദൃച്ഛികസാമ്പിളനം സാമ്പിളെടുക്കൽ, വിവരം ശേഖരിക്കൽ എന്നീ രണ്ടു കാര്യങ്ങളെസംബന്ധിച്ചും മററു സാമ്പിളനരീതികളെക്കാൾ വളരെയേറെ സൗകര്യപ്രദമായ ഒന്നാണ്.

2. സമഷ്ടിയിൽ N അംഗങ്ങളുണ്ടെന്നിരിക്കട്ടെ. n അംഗങ്ങളുള്ള സാമ്പിളാണ് ആവശ്യമെന്നു സങ്കല്പിക്കുക. രേഖീയം, ചക്രീയം എന്നു് രണ്ടു സമ്പ്രദായങ്ങൾ സാമ്പിളെടുക്കാൻ ഉപയോഗിക്കാറുണ്ട്. $N=nk$ എന്നിരിക്കട്ടെ. സമഷ്ടിയിലെ അംഗങ്ങൾക്ക് ക്രമസംഖ്യ കൊടുക്കുക. 1 മുതൽ k വരെയുള്ള സംഖ്യകളിൽ നിന്നു് യാദൃച്ഛികമായി ഒന്നു് തിരഞ്ഞെടുക്കുക. r ആണ് കിട്ടിയതെന്നിരിക്കട്ടെ. $r, r+k, r+2k, \ldots\ldots r+(n-1)k$ എന്നീ ക്രമസംഖ്യകളുള്ള അംഗങ്ങളെ സാമ്പിളിൽ ഉൾപ്പെടുത്തുക. ഇതാണ് രേഖീയസമ്പ്രദായം. $N \neq nk$ ആയിരിക്കുമ്പോൾ ഈ രീതിയിൽ ആവശ്യമായ പരിഷ്കാരങ്ങൾ വരുത്തണം. ചക്രീയരീതിയിൽ 1 മുതൽ N വരെയുള്ള സംഖ്യകളിൽ നിന്നു് ഒന്നു് യാദൃച്ഛികമായി തിരഞ്ഞെടുക്കുന്നു. r ആണ് തിരഞ്ഞെടുത്ത സംഖ്യ എന്നിരിക്കട്ടെ. $r+k, r+2k, \ldots\ldots r+(n-1)k$ എന്നീ സംഖ്യകൾ ക്രമസംഖ്യകളായുള്ള അംഗങ്ങളെയാണ് സാമ്പിളിൽ ഉൾപ്പെടുത്തുക. ഇതിൽ ഏതെങ്കിലും സംഖ്യ N നെക്കാൾ വലുതായിവന്നാൽ അതിൽ നിന്നു് N ന്റെ സാധ്യമായ ഏററവും വലിയ ഗുണിതം കുറച്ചുവരുന്ന ശിഷ്ടമാണ് സൂചകസംഖ്യയായി സ്വീകരിക്കേണ്ടതു്.

3. സാമ്പിൾസമാന്തരമാധ്യമായ $\bar{x}_{sy}$, സമഷ്ടിയുടെ സമാന്തരമാധ്യമായ $\bar{X}$ ന്റെ അനഭിനതആകലമാണ്. $N=nk$ അല്ലെങ്കിൽ ഈ പ്രസ്താവന ശരിയാവണമെന്നില്ല. അങ്ങനെ വരുമ്പോൾ ആകലനരീതിയിൽ ചില മാററങ്ങൾ വരുത്തി അനഭിനതആകലം കണ്ടുപിടിക്കേണ്ടിവരും. സാമ്പിൾപരിമാണം വലുതാണെങ്കിൽ ഈ വ്യത്യാസം അവഗണിക്കാം.

4. $V(\bar{x}_{sy}) = \sigma^2 - \sigma^2_w$

ഇവിടെ s^2, സമഷ്ടിയുടെ പ്രസരണവും s^2_w, സമഷ്ടിയിൽ നിന്നു് എടുക്കാവുന്ന n പരിമാണമുള്ള എല്ലാ സാമ്പിളുകളുടെയും പ്രസരണങ്ങളുടെ സമാന്തരമാധ്യവുമാണ്. അതു പോലെ തന്നെ,

$$V(\bar{x}_{sy}) = \frac{\sigma^2}{n}\left[1+(n-1)\rho_w\right]$$

ഇവിടെ ρ_w അന്തരാവർഗസഹബന്ധമത്രെ. സമഷ്ടിയിൽ അംഗങ്ങൾ ഏതു ക്രമത്തിലാണ് പരിഗണിക്കപ്പെടുന്നതു് എന്നതിനെ ആശ്രയിച്ചാണ് $\bar{x}_{sy}$ ന്റെ സാമ്പിളനപ്രസരണം എന്നു ഇതിൽ നിന്നു സിദ്ധിക്കുന്നു. അംഗങ്ങളെ അഭിലക്ഷണമൂല്യങ്ങളുടെ ആരോഹണക്രമത്തിലോ അവരോഹണക്രമത്തിലോ പരിഗണിച്ചാൽ $V(\bar{x}_{sy})$ കുറഞ്ഞുവരും.

5. യാദൃച്ഛികരീതിയിൽ ക്രമീകരിച്ചിട്ടുള്ള സമഷ്ടിയിൽ നിന്നു് എടുക്കുന്ന ക്രമാനുസൃതസാമ്പിൾ ലഘുയാദൃച്ഛികസാമ്പിളിനെക്കാൾ ഒട്ടും മെച്ചപ്പെട്ടതായിരിക്കുകയില്ല. രേഖീയപ്രവണതയുള്ള സമഷ്ടികളിൽ നിന്നു ക്രമാനുസൃതസാമ്പിളെടുത്താൽ ആകലം അത്യന്തം മെച്ചപ്പെട്ടതായിരിക്കും. അവർത്തികതയുള്ള സമഷ്ടികളിൽ നിന്നു് ക്രമാനുസൃതസാമ്പിളെടുക്കുന്നതു് വളരെ കരുതിവേണം.

6. സാമ്പിളനപ്രസരണത്തിന്റെ അനഭിനതആകലങ്ങൾ മാത്രമേ നിർദേശിക്കപ്പെട്ടിട്ടുള്ളു.

7. ക്രമാനുസൃതസാമ്പിളനരീതിയെ ദ്വിമാനസമഷ്ടി, ത്രിമാനസമഷ്ടി തുടങ്ങിയവയിലേക്കു് സാമാന്യവൽക്കരിക്കാവുന്നതാണു്. ഒന്നിലധികം അഭിലക്ഷണങ്ങളെപ്പററി പഠനം നടത്താനും ഈ രീതി ഉപയോഗിക്കാം.

8. ക്രമാനുസൃതസാമ്പിളനത്തിൽ, യാദൃച്ഛികപ്രാരംഭത്തിൽ മാററം വരുത്തുക, സമഷ്ടിയിലെ അംഗങ്ങളെ യാദൃച്ഛികമായി ക്രമീകരിച്ചിട്ടു് സാമ്പിളെടുക്കുക, പല സാമ്പിളുകൾ എടുത്തു് ഒന്നിച്ചു് ചേർക്കുക തുടങ്ങി പല പരിഷ്കാരങ്ങളും നിർദേശിക്കപ്പെട്ടിട്ടുണ്ടു്. ഇവ ഓരോന്നും ഓരോ തരത്തിൽ ഗുണകരമാണെങ്കിലും മററു ചില വീക്ഷണകോണങ്ങളിൽ നിന്നു നോക്കുമ്പോൾ ദോഷകരമാണെന്നു പറയാതെ വയ്യ.

15. ചില സഹായക ഗ്രന്ഥങ്ങളും ലേഖനങ്ങളും

1. Sampling Techniques (Chapter-8) W. G. Cochran, Wiley
2. Survey Sampling (Chapter 4). Leslie Kish, Wiley
3. Sampling Theory and Methods (Chapter 5) M. N., Murthy, Statistical publishing society, Calcutta
4. Systematic Sampling (1948), Yates, Phil. Trans: Royal society London, A241, 345-377
5. Theory of Systematic Sampling (1944) Madow. W. G. Madow L. H, Ann: Math: Stat: 15, 1-24
6. A review of the Literature of Systematic Sampling (1951), Buckland, J. R S (B), 13, 208-215
7. Systematic Sampling and its relation to other Sampling Designs (1946) Madow L. H., J. A. S. A, 41, 204-217

അഭ്യാസം 5

1. പ്രസവത്തിനു ശേഷമുള്ള ഓരോ ആഴ്ചയുടെയും ആദ്യത്തെ ദിവസം ഒരു പശു നൽകിയ പാൽ താഴെ കൊടുക്കുന്നു.

ദിവസം	കിട്ടിയ പാൽ [ലിററിൽ]	ദിവസം	കിട്ടിയ പാൽ	ദിവസം	കിട്ടിയ പാൽ
1	17	9	21	17	16
2	18	10	21	18	14
3	19	11	18	19	14
4	19	12	20	20	15
5	21	13	19	21	14
6	16	14	18	22	15
7	17	15	17	23	15
8	26	16	17	24	13

ഇതിൽ നിന്നു് 8 അംഗങ്ങളുള്ള രേഖീയക്രമാനുസൃതസാമ്പിളെടുത്തു് ആ പശുവിന്റെ ശരാശരി പ്രതിദിനക്ഷീരോൽപാദനശേഷി ആകലനം ചെയ്യുക. 9 അംഗങ്ങളുള്ള സാമ്പിളാണെടുക്കേണ്ടതെങ്കിൽ എന്താണു് ചെയ്യുക. അങ്ങനെ ഒരു സിമ്പിളെടുത്തു് ശരാശരി ക്ഷീരോൽപാദനത്തിന്റെ അനഭിനതആകലം കാണുക.

2. അഭ്യാസം 1–ലെ സമഷ്ടിയിൽ നിന്നു് 9 അംഗങ്ങളുള്ള ചക്രീയക്രമാനുസൃതസാമ്പിളെടുത്തു് ശരാശരി ക്ഷീരോൽപാദനം ആകലനം ചെയ്യുക.

3. 1–ലെ ആകലനത്തിന്റെ സാമ്പിളനപ്രസരണം ആകലനം ചെയ്യുക.

4. 1–ലെ സമഷ്ടിയെ പാലിന്റെ അളവനുസരിച്ചു് ആരോഹണക്രമത്തിൽ ക്രമീകരിച്ചു് അതിൽ നിന്നു് 8 അംഗങ്ങളുള്ള സാമ്പിളെടുത്തു് ശരാശരി ക്ഷീരോൽ പാദനം ആകലനം ചെയ്യുക.

5. 1–ലെ സമഷ്ടിയുടെ സമാന്തരമാധ്യം കണ്ടു പിടിക്കുക. 1, 2, 4 എന്നീ അഭ്യാസങ്ങളിൽ ഗണിച്ചെടുത്ത ആകലങ്ങളും യഥാർഥമൂല്യവും താരതമ്യപ്പെടുത്തുക.

6. ഒരു കലാലയത്തിലെ വിദ്യാർഥികളുടെ ലിസ്റ്റു് സമുദായം തിരിച്ചു് തയ്യാറാക്കിയിരിക്കുന്നു. ആ ലിസ്റ്റു് ഉപയോഗിച്ച് ഒരു ക്രമാനുസൃതസാമ്പിളെടുത്താൽ താഴെ പറയുന്ന ഏതെല്ലാം കാര്യങ്ങളെപ്പറ്റി പഠനം നടത്താനാണു് അതു് ലഘുയാദൃച്ഛികസാമ്പിളിനെക്കാൾ മെച്ചമായിരിക്കുക? എന്തു കൊണ്ടു്?

എ. ആ കലാലയത്തിൽ പഠിക്കുന്ന ഓരോസമുദായത്തിലും പെട്ട കുട്ടികളുടെ അനുപാതം ആകലനം ചെയ്യുക.

ബി. അവരുടെ ശരാശരി ഭാരം ആകലനം ചെയ്യുക.

സി. അവരിൽ നേത്രരോഗികളുടെ അനുപാതം ആകലനം ചെയ്യുക.

7. റബ്ബർ വാങ്ങുന്ന ഒരു കടയിൽ രാവിലെ 6 മണി മുതൽ വൈകീട്ടു് 6 മണി വരെയുള്ള ഓരോ മണിക്കൂറിന്റെയും ആദ്യത്തെ 15 മിനിട്ടിനുള്ളിൽ വാങ്ങിച്ച റബ്ബറിന്റെ തൂക്കം താഴെ കൊടുക്കുന്നു.

സമയം.	6	7	8	9	10	11	12	1	2	3	4	5
വാങ്ങിയ റബ്ബർ (കി. ഗ്രാം)	12	46	48	63	55	70	38	40	60	53	18	14

ആകെ അന്നു വാങ്ങിയിരിക്കാവുന്ന റബ്ബറും അതിന്റെ സാമ്പിളനപ്രസരണവും ആകലനം ചെയ്യുക. 12 അംഗങ്ങളുടെ ലഘുയാദൃച്ഛികസാമ്പിളിനേക്കാൾ ഈ സാമ്പിൾ മെച്ചപ്പെട്ടതാണോ?

6

അനുപാത സമാശ്രയണാകലങ്ങൾ

(1) അനുപാതാകല

1. പ്രാരംഭം

സാമ്പിളനത്തിൻെറ പ്രധാനലക്ഷ്യങ്ങളിൽ ചിലതു് സമഷ്ടിയുടെ സമാന്തരമാധ്യം, തുക, ഏതെങ്കിലും പ്രത്യേകതയുള്ള അംഗങ്ങളുടെ അനുപാതം എന്നിവയുടെ ആകലനമാണ്. സാധാരണ ഉപയോഗത്തിലുള്ള മൂന്നു സാമ്പിളന രീതികളിൽ ഈ ആകലന പ്രക്രിയ എങ്ങനെ നിർവഹിക്കുന്നു എന്നും അങ്ങനെ ലഭിക്കുന്ന ആകലനങ്ങളുടെ സാമ്പിളന പ്രസരണം എന്തെന്നും മുൻ അധ്യായങ്ങളിൽ സാമാന്യമായി ചർച്ച ചെയ്തു കഴിഞ്ഞു. ഈ അധ്യായത്തിൽ പരിഗണിക്കുന്നതു് പുതിയ രണ്ടു് ആകലനസമ്പ്രദായങ്ങളാണ്. അവയിൽ ആദ്യത്തേതാണ് അനുപാതാകലനം. മററു ആകലന സമ്പ്രദായങ്ങളെ അപേക്ഷിച്ചു് പലപ്പോഴും കൂടുതൽ സൂക്ഷ്മതയുള്ളതാണ് ഈ സമ്പ്രദായമെന്ന് വഴിയെ കാണാം. പക്ഷേ സമഷ്ടിയെ സംബന്ധിക്കുന്ന ചില ഉപവിവരങ്ങൾ കൂടി അറിയാമെങ്കിൽ മാത്രമേ ഈ രീതി വിജയകരമായി ഉപയോഗിക്കാൻ കഴിയൂ. മാത്രമല്ല, സമഷ്ടിക്കു് ചില പ്രത്യേകതകളെല്ലാം ഉണ്ടായിരിക്കുകയും വേണം.

2. അനുപാതാകലനം

സമഷ്ടിയോടു ബന്ധപ്പെട്ട y എന്ന അഭിലക്ഷണത്തിലാണ് നമുക്കു് താല്പര്യമുള്ളതെന്നിരിക്കട്ടെ. അഭിലക്ഷണമൂല്യങ്ങളുടെ തുകയാണ് ആകലനം ചെയ്യേണ്ടതെന്നും സങ്കല്പിക്കുക. ലഘുയാദൃച്ഛിക സാമ്പിളന രീതിയാണ് അവലംബിക്കുന്നതെന്നും വിചാരിക്കുക. നിയമാനുസരണം സാമ്പിളെടുത്തു് ഓരോ അംഗത്തിൻെറയും അഭിലക്ഷണമൂല്യം രേഖപ്പെടുത്തി അവയുടെ സമാന്തരമാധ്യം കണ്ടുപിടിച്ചു് അതിനെ സമഷ്ടിയുടെ പരിമാണമായ N കൊണ്ടു ഗുണിച്ചാൽ അഭിലക്ഷണമൂല്യങ്ങളുടെ തുകയുടെ ആകലം ലഭിക്കുമെന്നാണ് നാം മനസ്സിലാക്കിയിരിക്കുന്നതു്. ഈ സമ്പ്രദായത്തെക്കാൾ അല്പം കൂടി സൂക്ഷ്മതയുള്ള

ഒരു രീതിയാണ് അനുപാതാകലനം. പരിഗണിക്കപ്പെടുന്ന y എന്ന അഭിലക്ഷണവുമായി ശക്തിയായ നേർ സഹബന്ധമുള്ള x എന്നൊരഭിലക്ഷണമുണ്ടെന്നിരിക്കട്ടെ. സമഷ്ടിയിലെ i-ാമത്തു് അംഗത്തിൻെറ y,x അഭിലക്ഷണമൂല്യങ്ങൾ യഥാക്രമം y_i x_i എന്നു സങ്കല്പിക്കുക. $\frac{y_i}{x_i}$ എന്ന അനുപാതം എല്ലാ അംഗങ്ങൾക്കും ഏതാണ്ടു് തുല്യമാണെന്നും കരുതുക.

$$X = \sum_{i=1}^{N} x_i, \; Y = \sum_{i=1}^{N} y_i \text{ എന്നിരിക്കട്ടെ.}$$

അതായതു് X, Y എന്നിവ യഥാക്രമം x,y എന്നീ അഭിലക്ഷണങ്ങളുടെ മൂല്യങ്ങളുടെ തുകയാണു്. $R = \frac{Y}{X}$ എന്നു നിർവചിക്കുക. R ൻെറ മൂല്യം സാമ്പിൾ ഉപയോഗിച്ചു് ആകലനം ചെയ്യുക. ആകലം $\hat{R}$ എന്നിരിക്കട്ടെ. X ൻെറ മൂല്യം അറിയാമെങ്കിൽ

$$\hat{Y} = RX \qquad 6.2.1$$

എന്ന സമീകരണമുപയോഗിച്ചു് Y യുടെ മൂല്യം ആകലനം ചെയ്യാം. Y_R കൊണ്ടു് സൂചിപ്പിച്ച ഈ ആകലത്തിനു് Y യുടെ അനുപാതാകലം എന്നു പേർ പറയുന്നു. ഇവിടെ $\hat{Y}$, $\hat{X}$ എന്നിവ Y യുടെയും X ൻെറയും അനഭിനത ആകലങ്ങളാണെങ്കിൽ

$$\hat{R} = \frac{\hat{Y}}{\hat{X}}$$ എന്നതു് R ൻെറ ഒരു ആകലമായി സ്വീകരിക്കാവുന്നതാണു്.

ഒരുദാഹരണം കാണിക്കാം. ഒരു ഓഫീസിൽ ജോലി ചെയ്യുന്ന ഉദ്യോഗസ്ഥന്മാർ പ്രതിമാസം എത്ര രൂപ മിച്ചം വെക്കുന്നുണ്ടെന്നു് ആകലനം ചെയ്യണമെന്നിരിക്കട്ടെ. എല്ലാവരും ശമ്പളത്തിൻെറ ഒരു നിശ്ചിതഅനുപാതം മിച്ചം വെക്കുന്നു എന്നു സങ്കല്പിച്ചാൽ മിച്ചം വെക്കുന്ന തുക എന്ന അഭിലക്ഷണവുമായി ശക്തിയായ നേർസഹബന്ധമുള്ള ഒരു അഭിലക്ഷണമാണു് ഓരോരുത്തരുടെയും ശമ്പളം. i-ാമത്തെ ഉദ്യോഗസ്ഥൻ മിച്ചം വെക്കുന്ന തുക y_i-യും അയാളുടെ ശമ്പളം x_i യും ആണെന്നു വിചാരിക്കുക. $\frac{y_i}{x_i}$ എല്ലാവർക്കും ഏതാണ്ടു് തുല്യമാണെന്നുള്ള സങ്കല്പമാണു് നമ്മുടെ ആകലന സമ്പ്രദായത്തിൻെറ അടിസ്ഥാനം. X ആകെ ശമ്പളവും Y ആകെ സമ്പാദ്യവുമാണു്. N ഉദ്യോഗസ്ഥന്മാരാണു് ഓഫീസിലുള്ളതെന്നിരിക്കട്ടെ. അവരിൽ നിന്നു് ലഘുയാദൃച്ഛിക സാമ്പിളനരീതിയിൽ n ഉദ്യോഗസ്ഥന്മാരെ തിരഞ്ഞെടുത്തു, അവരോരോരുത്തരുടെയും ശമ്പളവും അവർ മിച്ചം വെക്കുന്ന തുകയും ചോദിച്ചറിയുക. $\bar{x}$ സാമ്പിളിലെ അംഗങ്ങളുടെ ശമ്പ

ളത്തിൻെറ സമാന്തരമാധ്യവും $\bar{Y}$ മിച്ചം വെക്കുന്ന തുകകളുടെ സമാന്തരമാധ്യവുമാണെങ്കിൽ $N\bar{x}$, $N\bar{y}$ എന്നിവ ആകെ ശമ്പളത്തിൻെറയും ആകെ മിച്ചം വെക്കുന്ന തുകയുടെയും അനഭിനത ആകലങ്ങളായിരിക്കും. അതായതു്, $\hat{X} = N\bar{x}$, $\hat{Y} = N\bar{y}$

$$\text{അപ്പോൾ, } \hat{R} = \frac{N\bar{y}}{N\bar{x}} \; \frac{\bar{y}}{\bar{x}} \text{ എന്നെടുക്കാം.}$$

ആകെ ശമ്പളമായ X ഓഫീസിലെ കണക്കുകൾ പരിശോധിച്ചാൽ കണ്ടുപിടിക്കാം. നമുക്കു് ആവശ്യം Y യുടെ ആകലമാണു്. Y യുടെ അനുപാതാകലമായ Y താഴെ കൊടുത്തിരിക്കുന്ന സമീകരണമുപയോഗിച്ചു് കണ്ടുപിടിക്കാം.

$$y_R = \hat{R} X \quad \text{അതായതു്,} \quad Y_R = \frac{\bar{y}}{\bar{x}} X.$$

ഇവിടെ പ്രത്യേകം ശ്രദ്ധിക്കാനുള്ള ഒരു വസ്തുത X ൻെറ മൂല്യം മറേറതെങ്കിലും തരത്തിൽ കണ്ടുപിടിക്കാൻ സാധിച്ചെങ്കിൽ മാത്രമെ ഈ ആകലനരീതി ഉപയോഗിക്കാനാവൂ എന്നതാണു്.

പരിഗണിക്കപ്പെടുന്ന അഭിലക്ഷണവും അതോടു ബന്ധപ്പെട്ട മറെറാരഭിലക്ഷണവുമായുള്ള ശക്തിയായ സഹബന്ധം ഉപയോഗപ്പെടുത്തി ആകലത്തിൻെറ സൂക്ഷ്മത വർധിപ്പിക്കാനാണു് നാം ശ്രമം നടത്തുന്നതു്. $\frac{y_i}{x_i}$ എന്ന അനുപാതം സമഷ്ടിയിലെ എല്ലാ അംഗങ്ങൾക്കും ഏതാണ്ടു് തുല്യമാണെന്നുള്ള സങ്കല്പം എത്രത്തോളം ശരിയാണോ അത്രക്കു് അനുപാതാകലത്തിൻെറ സൂക്ഷ്മതയും വർധിക്കും. അതിൻെറ പ്രകീർണനം അനുപാതാകലത്തിൻെറ സൂക്ഷ്മതയിൽ കുറവു വരുത്തുകയും ചെയ്യും.

3. അനുപാതാകലത്തിൻെറ അഭിനതിയും സാമ്പിളനപ്രസരണവും

അനുപാതാകലം പൊതുവെ പറഞ്ഞാൽ അനഭിനതമല്ല. ചുരുക്കം ചില സമഷ്ടികളെ സംബന്ധിച്ചു മാത്രമേ അതു് അനഭിനതമായിരിക്കൂ. സാമ്പിൾ പരിമാണം വളരെ വലുതാണെങ്കിൽ ഈ അഭിനതി നിസ്സാരമാണു് എന്നു മാത്രം. ഇതു വരെ നാം പരിഗണിച്ച ആകലങ്ങളെല്ലാം അനഭിനതങ്ങളായിരുന്നു. ഇവിടെ ഒരു അഭിനത ആകലം എന്തിനാണു് സ്വീകരിക്കുന്നതു് എന്ന ചോദ്യം പ്രസക്തമത്രെ. സാമ്പിളനപ്രസരണം കുറവുള്ള അഭിനതആകലം അനഭിനതആകലത്തെക്കാൾ മെച്ചപ്പെട്ടതാണു് എന്നതാണു് ഇതിനു കാരണം.

അനുപാതാകലം നിർണയിക്കാൻ നാം രണ്ടു് അഭിലക്ഷണങ്ങളാണല്ലൊ പരിഗണിക്കുന്നതു്. പഠനവിധേയമാക്കേണ്ട അഭിലക്ഷണവും അതുമായി ശക്തിയായ നേർ സഹബന്ധമുള്ള മറെറാരഭിലക്ഷണവും. y, x എന്നിവയാണു് അവയെ സൂചിപ്പിക്കാൻ ഉപയോഗിച്ചതു്. ഈ രണ്ടു് അഭിലക്ഷണങ്ങളുടെയും മൂല്യങ്ങളുടെ തുകകൾ യഥാക്രമം Y, X എന്നിവയാണു്. അവയുടെ അനുപാത

മാണ് R. R ന്റെ ആകലമുപയോഗിച്ചാണ് Y യുടെ അനുപാതാകലം കണ്ടുപിടിച്ചത്. R ആകലനം ചെയ്യാൻ n അംഗങ്ങളുള്ള ഒരു ലഘുയാദൃച്ഛിക സാമ്പിൾ ഉപയോഗിക്കുന്നു എന്നിരിക്കട്ടെ. സാമ്പിളിലെ അഭിലക്ഷണ മൂല്യങ്ങളുടെ സമാന്തരമാധ്യങ്ങൾ യഥാക്രമം $\bar{y}, \bar{x}$ എന്നിവയാണെങ്കിൽ $N\bar{y}$, $N\bar{x}$ എന്നിവ Y X എന്നിവയുടെ അനഭിനത ആകലങ്ങളാണ്. $\hat{X}, \hat{Y}$ എന്നിവ യഥാക്രമം Y, X എന്നിവയുടെ അനഭിനത ആകലങ്ങളാണെങ്കിൽ $\hat{R} = \frac{\hat{Y}}{\hat{X}}$ എന്നാണല്ലോ നിർവചിച്ചിരിക്കുന്നത്. അതുകൊണ്ട്

$\hat{R} = \frac{N\bar{y}}{N\bar{x}} = \frac{\bar{y}}{\bar{x}}$ ആണ്. ഇതു് R ൻെറ ഒരു അനഭിനത ആകലമാണോ എന്നാണ് പരിശോധിക്കേണ്ടിയിരിക്കുന്നത്. ആണെങ്കിൽ Y യുടെ അനുപാതാകലമായ Y_R അനഭിനതമാണെന്നു കാണാൻ പ്രയാസമില്ല.

പ്രതിസ്ഥാപനമില്ലാത്ത ലഘുയാദൃച്ഛിക സാമ്പിളന രീതിയിലാണ് സാമ്പിളെടുത്തിരിക്കുന്നതെങ്കിൽ, ആകെ $\binom{N}{n}$ സാമ്പിളുകളാണ് എടുക്കാൻ കഴിയുന്നത്. $\bar{y}_1$, $\bar{x}_1$ എന്നിവയാണ് 1-ാമത്തു് സാമ്പിളിലെ അഭിലക്ഷണമൂല്യങ്ങളുടെ സമാന്തരമാധ്യങ്ങൾ എന്നിരിക്കട്ടെ.

$$E\left(\hat{R}\right) = \frac{1}{\binom{N}{n}} \sum_{l=1}^{\binom{N}{n}} \frac{\bar{y}_l}{\bar{x}_l}$$

ഇതു് സാധാരണഗതിയിൽ $\frac{Y}{X}$ നോട്ട് തുല്യമാവാൻ ഇടയില്ല. അതായതു്,

$E(\hat{R}) \neq R$. അങ്ങിനെ $\hat{R}$ ഒരു അനഭിനത ആകലമാണെന്നു വന്നുചേരുന്നു.

ഇതു മാത്രം കൊണ്ടു് $\hat{R}$ ൻെറ സ്വീകാര്യതയെപ്പററി ഒന്നും പറയാനാവില്ല. അഭിനതി വളരെ കുറവാണെങ്കിൽ അതു് അവഗണിക്കുന്നതുകൊണ്ട് വലിയ തകരാറില്ല. നേരേ മറിച്ച് അഭിനതിയുടെ പരിമാണം വർധിക്കുന്തോറും ശോധനകളുടെ ആവശ്യവും കൂടുന്നു.

$\hat{R}$ ൻെറ അഭിനതി എങ്ങനെ കണ്ടുപിടിക്കാമെന്നതാണ് അടുത്ത പ്രശ്നം. കൃത്യമായ ഒരു വ്യഞ്ജകം ഇതിനായി ഇതുവരെ നിർദേശിക്കപ്പെട്ടിട്ടില്ലാത്തതുകൊ

ണ്ടു് അഭിനതിയുടെ ചില ഏകദേശനങ്ങൾകൊണ്ടു് തൃപ്തിപ്പെടാനേ തരമുള്ളു ഏതു സാമ്പിളനരീതിയാണു് അവലംബിക്കുന്നതു് എന്നതിനെ ആശ്രയിച്ചാണു് ഈ ഏകദേശനങ്ങളുടെ യഥാർത്ഥ വ്യംജകങ്ങൾ രൂപമെടുക്കുന്നതു്. പൊതുവെ പറഞ്ഞാൽ $E\left(\hat{X}\right) = X$ ആയതുകൊണ്ടു് $\hat{X} = X\left(1 + \triangle\hat{X}\right)$

എന്നും $E(\hat{Y}) = Y$ ആയതുകൊണ്ടു് $\hat{Y} = Y(1 + \triangle\hat{Y})$ എന്നും എഴുതാം. അതായതു്,

$$\triangle\hat{X} = \frac{\hat{X} - X}{X}, \quad \triangle\hat{Y} = \frac{\hat{Y} - Y}{Y}$$

അതു് കൊണ്ടു്,

$$\hat{R} = \frac{\hat{Y}}{\hat{X}} = \frac{Y\left(1 + \triangle\hat{Y}\right)}{X\left(1 + \triangle\hat{Y}\right)} = \frac{Y}{X}(1 + \triangle\hat{Y})(1 + \triangle\hat{X})^{-1}$$

$$E(\hat{R}) = E\left\{\frac{Y}{X}(1 + \triangle\hat{Y})(1 + \triangle\hat{X} - 1\right\}$$

$$= \frac{Y}{X}E\left\{(1 + \triangle\hat{Y})(1 - \triangle\hat{X} + (\triangle\hat{X})^2 \ldots \ldots)\right\}$$

$$= \frac{Y}{X}E\left\{1 + \triangle\hat{Y} - \triangle\hat{X} + (\triangle\hat{X})^2 - (\triangle\hat{Y})(\triangle\hat{X}) + \cdots\right\}$$

$$\frac{Y}{X} + \frac{Y}{X}E\left\{\triangle\hat{Y} - \hat{X}) + \left[(\triangle\hat{X})^2 - (\triangle\hat{Y})(\triangle\hat{X})\right] + \cdots\right\}$$

ഇവിടെ $E(\triangle\hat{Y}) = E(\triangle\hat{X}) = 0$

$$E(\triangle\hat{X})^2 = E\left(\frac{\hat{X} - X}{X}\right)^2 = \frac{V(\hat{X})}{X^2}$$

$$E\left[(\triangle\hat{Y})(\triangle\hat{X})\right] = E\left[\frac{(\hat{Y} - Y)}{Y} \cdot \frac{(\hat{X} - X)}{X}\right]$$

$$= \frac{Cov(\hat{Y}\hat{X})}{XY}$$

അതുകൊണ്ട് $E(\hat{R}) = \frac{Y}{X} + \frac{Y}{X}\left[0 + \frac{V(\hat{X})}{X^2} - \frac{Cov(\hat{X}\hat{Y})}{XY} + \ldots\ldots\right]$

ആദ്യ ഏകദേശനമെന്ന നിലയിൽ,

$$E(\hat{R}) = \frac{Y}{X} = R \qquad 6.3.1$$

ഒരു പടി കൂടി മുന്നോട്ടു പോയാൽ

$$E(\hat{R}) = \frac{Y}{X} + \frac{Y}{X}\left[\frac{V(\hat{X})}{X^2} - \frac{Cov(\hat{X}\hat{Y})}{XY}\right]$$

$$= R + R\left[\frac{V(\hat{X})}{X^2} - \frac{Cov(\hat{X}\hat{Y})}{XY}\right] \qquad 6.3.2$$

പടിപടിയായി എത്ര സൂക്ഷ്മതയുള്ള വ്യഞ്ജകങ്ങൾ വേണമെങ്കിലും ഈ രീതിയിൽ വ്യുൽപ്പാദിപ്പിക്കാവുന്നതാണ്.

ആദ്യഏകദേശനമനുസരിച്ച് $\hat{R}$ ഒരു അനഭിനത ആകലമാണ്. ഇവിടെ ഫലത്തിൽ $\Delta\hat{Y}, \Delta\hat{X}$ എന്നിവ താരതമ്യേന വളരെ ചെറുതായിരിക്കുമെന്നും അവ അവഗണിക്കാമെന്നും നാം സങ്കല്പിക്കുകയാണ്. രണ്ടാമത്തെ ഏകദേശനത്തിൽ $\Delta\hat{Y}, \Delta\hat{X}$ എന്നിവയുടെ രണ്ടിൽ കവിഞ്ഞുള്ള ഘാതങ്ങൾ വളരെ ചെറുതായിരിക്കുമെന്നതുകൊണ്ട് അവ പരിഗണിക്കേണ്ടതില്ലെന്നു നിശ്ചയിക്കുന്നു. ഇതിൽ ആദ്യത്തെ ഏകദേശനം വളരെ സ്ഥൂലമാണെന്നു പറയാതെ തരമില്ല. രണ്ടാമത്തേതു് പ്രായോഗികമായ എല്ലാ സന്ദർഭങ്ങളിലേക്കും ഏതാണ്ട് ആവശ്യമുള്ളത്ര സൂക്ഷ്മതയുള്ളതാണ്. ഇതിൽ കൂടുതൽ സൂക്ഷ്മതയുള്ള ഏകദേശനങ്ങൾക്കുവേണ്ടിയുള്ള അന്വേഷണം അതീവസങ്കീർണങ്ങളായ വ്യഞ്ജകങ്ങളിലേക്ക് നമ്മെ നയിക്കുമെന്നതുകൊണ്ട് അത്ര അഭികാമ്യമല്ല.

$\hat{R}$ ന്റെ അഭിനതിയെ $b(\hat{R})$ എന്നതുകൊണ്ട് കുറിക്കുന്നു എന്നിരിക്കട്ടെ. രണ്ടാമത്തെ ഏകദേശനം സ്വീകരിച്ചാൽ,

$$b(\hat{R}) = E(\hat{R}) - R \simeq R\left[\frac{V(\hat{X})}{X^2} - \frac{Cov(\hat{X}\hat{Y})}{XY}\right]$$

$$\simeq \frac{R}{X^2}\left[V(\hat{X}) - \frac{X}{Y}Cov(\hat{X}\hat{Y})\right]$$

$$\approx \frac{R}{X^2}\left[V(\hat{X}) - \frac{Cov(\hat{X}\hat{Y})}{R}\right] \qquad 6.3.3$$

ഈ അഭിനതിയുടെ ഏറ്റവും കുറഞ്ഞ മൂല്യം പൂജ്യമാണ്. അങ്ങനെയാവണമെങ്കിൽ,

$$V(\hat{X}) = \frac{Cov(\hat{X}\hat{Y})}{R}$$ ആവണം.

അതായത്, $R = \frac{Cov(\hat{X}\hat{Y})}{V(\hat{X})}$ ആവണം.

ഇത് എപ്പോഴാണ് സംഭവിക്കുക എന്നു പരിശോധിച്ചുനോക്കാം. $\hat{Y}$ന് $\hat{X}$ ൻെറ മേലുള്ള സമാശ്രയണം രേഖീയമാണെങ്കിൽ, ആ രേഖ

$$\hat{Y} - Y = \frac{Cov(\hat{X}\hat{Y})}{V(\hat{X})}(\hat{X} - X) \qquad 6.3.4$$

എന്നതാണെന്നു നമുക്കറിയാം. $\hat{Y}$ ഉം $\hat{X}$ ഉം തമ്മിലുള്ള ബന്ധത്തെ സൂചിപ്പിക്കുന്ന ഈ രേഖ, മൂലബിന്ദുവിൽ കൂടി കടന്നുപോകുന്നെങ്കിൽ, അത്

$$\hat{Y} = \frac{Cov(\hat{X}\hat{Y})}{V(\hat{X})}\hat{X}$$

എന്നാവും. അതായത്,

$$\hat{R} = \frac{\hat{Y}}{\hat{X}} = \frac{Cov(\hat{X}\hat{Y})}{V(\hat{X})}.$$

അഭിനതി പൂജ്യമാവുന്നതിനുള്ള നിബന്ധന ഇവിടെ ഉപയോഗപ്പെടുത്തിയാൽ,

$\hat{R} = R$ എന്നാവും. അപ്പോൾ, $E(\hat{R}) = E(R) = R$.

അങ്ങനെ $\hat{Y}$ന് $\hat{X}$ ൻെറ മേലുള്ള സമാശ്രയണരേഖ, മൂലബിന്ദുവിലൂടെ

കടന്നുപോവുക എന്നതു് $\hat{R}$ അനഭിനത ആകലമാകാനുള്ള ഒരു സാഹചര്യമാണെന്നു വന്നു ചേരുന്നു.

ഈ ചർച്ചയിൽ $\triangle \hat{X}$, $\triangle \hat{Y}$ എന്നിവയുടെ രണ്ടിൽ കൂടിയ ഘാതങ്ങൾ അവഗണിക്കാമെന്നു നാം സങ്കല്പിക്കുകയുണ്ടായി. ഈ സങ്കല്പം ശരിയാവണമെങ്കിൽ $\triangle\hat{X}$, $\triangle\hat{Y}$ എന്നിവയുടെ മൂല്യം ഒന്നിൽ കുറവായിരിക്കണം. അതായതു്

$$\left|\frac{\hat{X}-X}{X}\right|, \left|\frac{\hat{Y}-Y}{Y}\right|$$

എന്നിവ ഒന്നിൽ കുറഞ്ഞിരിക്കണം. സാമ്പിൾ വളരെ ചെറുതാണെങ്കിൽ ഈ സങ്കല്പം ശരിയാവണമെന്നില്ല. സാമ്പിൾപരിമാണത്തിന്റെ വർധനവനുസരിച്ചു് ഇവയുടെ മൂല്യങ്ങളും കുറഞ്ഞുവരുമെന്നു് തെളിയിക്കാവുന്നതാണു്. ചുരുക്കിപ്പറഞ്ഞാൽ സാമാന്യം വലിയ സാമ്പിളുകൾ ഉപയോഗിക്കുന്ന സന്ദർഭങ്ങളിൽ മാത്രമേ ഇവിടെ പറഞ്ഞ ഏകദേശനങ്ങൾ പ്രസക്തമായിരിക്കുകയുള്ളു. സമഷ്ടിയിലെ അംഗങ്ങൾ ഏകാത്മകങ്ങളാണെങ്കിൽ ചെറിയ സാമ്പിളുപയോഗിക്കുമിടത്തും ഇതു് ശരിയായിരിക്കുകയും ചെയ്യാം.

$\hat{R}$ ന്റെ അഭിനതിയായ $b(\hat{R})$ എങ്ങനെ ആകലം ചെയ്യാൻ കഴിയും എന്നതു് ഇവിടെ പ്രസക്തമാണു്. $X, R, V(\hat{X})$, $Cov(\hat{X}, \hat{Y})$ എന്നിവയാണല്ലോ $b(\hat{R})$ ന്റെ വ്യംജകം ഉൾക്കൊള്ളുന്ന അജ്ഞാതരാശികൾ. സാമ്പിളിൽ നിന്നു് ഇവ ഓരോന്നിന്റെയും മെച്ചപ്പെട്ട ആകലങ്ങൾ സാംഖ്യികശാസ്ത്രസിദ്ധാന്തങ്ങളനുസരിച്ചു് കണ്ടുപിടിച്ചു് $b(\hat{R})$ ന്റെ വ്യംജകത്തിൽ പ്രതിസ്ഥാപിച്ചാൽ അതിന്റെ ആകലം ലഭിക്കും. ഈ ആകലം അനഭിനതമായിരിക്കണമെന്നില്ല എന്ന വസ്തുത പ്രത്യേകം ഓർമയിരിക്കേണ്ടതാണു്.

അടുത്തതായി പരിഗണിക്കാനുള്ളതു് $\hat{R}$ ന്റെ സാമ്പിളന പ്രസരണമാണു്. പക്ഷേ $\hat{R}$ അഭിനതമായതു കൊണ്ടു് അതിന്റെ പ്രസരണം കണ്ടുപിടിച്ചതു കൊണ്ടു് വലിയ പ്രയോജനമില്ല. ആകലമെന്ന നിലയിൽ അതിന്റെ സൂക്ഷ്മത നിർണയിക്കാൻ ഉതകണമെങ്കിൽ അതിന്റെ പിശകു വർഗമാധ്യം നിർണയിക്കേണ്ടിയിരിക്കുന്നു. $M(\hat{R})$ എന്നതു കൊണ്ടു് അതിനെ സൂചിപ്പിക്കുന്നു എന്നു സങ്കല്പിക്കുക.

$$M(\hat{R}) = E\,(\hat{R}-R)^2$$

$$\hat{R} = \frac{Y}{X}\left(1+\frac{\Delta\hat{X}}{X}\right)\left(1+\frac{\Delta\hat{Y}}{Y}\right)^{-1}$$

കുറിപ്പ്:–: $\hat{R}$ ന്റെ അഭിനതിയുടെ ശരിയായ രൂപം താഴെ കാണും പ്രകാരമാണ്.

$$E(\hat{X}) = X, E\,(\hat{Y}) = Y$$

$$\begin{aligned}
\mathrm{Cov}(\hat{R}\hat{X}) &= E(\hat{Y}) - E\,(\hat{R})\,E\,(\hat{X}) \\
&= E(\hat{X})\left[\frac{E(\hat{Y})}{E(\hat{X})} - E(\hat{R})\right] \\
&= -\,X\,[E(\hat{R}) - R] \\
&= -\,X\,b(\hat{R})
\end{aligned}$$

$$b(\hat{R}) = -\frac{\mathrm{Cov}(\hat{R}\hat{X})}{X} = -\frac{1}{X}\,\rho\,(\hat{R},\hat{\Delta})\,\sigma\,(\hat{R})\,\sigma\,\hat{X})$$

$$\frac{b(\hat{R})}{R} = |\rho(\hat{R}\hat{X})|\;C(\hat{R})\,C(\hat{X})$$

$$\left|\frac{b(\hat{R})}{R}\right| \leqslant C(\hat{R})\,C(\hat{X})$$

ഇതിൽ നിന്നും $(\hat{R})$ ന്റെ ആപേക്ഷിക അഭിനതി $\hat{R}$ ന്റെയും $\hat{X}$ ന്റെയും പ്രസരണ ഗുണാങ്കങ്ങളുടെ ഗുണനഫലത്തെക്കാൾ ചെറുതായിരിക്കുമെന്ന് കാണാം.

$\rho(\hat{R}\ \hat{X})$ എന്നതു് $\hat{R}$, $\hat{X}$, ഇവ തമ്മിലുള്ള സഹബന്ധ ഗുണാങ്കമാണ്.

എന്നു നാം കണ്ടു. $\hat{R}$ ന്റെ ഈ മൂല്യം $M(\hat{R})$ ന്റെ ഈ വ്യംജകത്തിൽ പ്രതിസ്ഥാപിച്ചാൽ

$$M\hat{R} = E\left\{\frac{Y}{X}\left(1+\frac{\triangle\hat{Y}}{Y}\right)\left(1+\frac{\triangle\hat{X}}{X}\right)^{-1}-R\right\}^2$$

$$=R^2 E\left\{\left(1+\frac{\triangle\hat{Y}}{Y}\right)\left(1+\frac{\triangle\hat{X}}{X}\right)^{-1}-1\right\}^2$$

$$=R^2 E\left\{\left(1+\frac{\triangle\hat{Y}}{Y}\right)\left(1-\frac{\triangle\hat{X}}{X}+\frac{(\triangle\hat{X})^2}{X^2}\cdots\right)-1\right\}^2$$

$$=R^2 E\left\{\frac{\triangle\hat{Y}}{Y}-\frac{\triangle\hat{X}}{X}+\frac{(\triangle\hat{X})^2}{X^2}-\cdots\cdots\right\}^2$$

ഇവിടെ $\triangle\hat{X}$ ന്റെയും $\triangle\hat{Y}$ ന്റെയും രണ്ടിൽ കൂടിയ ഘാതങ്ങൾ വളരെ ചെറുതായിരിക്കും. അവയെ അവഗണിക്കുകയാണെങ്കിൽ,

$$M(\hat{R})=R^2E\left\{\frac{(\triangle\hat{Y})^2}{Y^2}+\frac{(\triangle\hat{X})^2}{X^2}-\frac{2(\triangle\hat{X})(\triangle\hat{Y})}{XY}\right\}$$

$$=R^2\left\{\frac{V(\hat{Y})}{Y^2}+\frac{V(\hat{X})}{X^2}-\frac{2\,Cov(\hat{X}\hat{Y})}{XY}\right\} \qquad 6.3.5$$

$M(\hat{R})$ ന്റെ വ്യംജകത്തിലെ $\triangle\hat{X}$, $\triangle\hat{Y}$ എന്നിവയുടെ രണ്ടിൽ കൂടിയ ഘാതങ്ങൾ അവഗണിക്കുക എന്നു പറഞ്ഞാൽ അതിന്റെ അർത്ഥം $\hat{R}$ ന്റെ വ്യംജകത്തിൽ $\triangle\hat{X}$, $\triangle\hat{Y}$ എന്നിവയുടെ ഒന്നിൽ കൂടിയ ഘാതങ്ങൾ അവഗണിക്കുക എന്നാണല്ലൊ. ഇതിന്റെ ഫലം $\hat{R}$ ന്റെ അഭിനതിയെ സംബന്ധിച്ച് ഒന്നാം ഏകദേശനം നാമംഗീകരിക്കുന്നു എന്നത്രെ. അതായതു് $\hat{R}$ ഒരു അനഭിനതആകലമാണെന്നു സങ്കൽപിക്കുന്നു. അങ്ങനെ $\hat{R}$ ന്റെ പിശകുവർഗമാധ്യമായ $M(\hat{R})$, $\hat{R}$ ന്റെ സാമ്പിളനപ്രസരണമായ $V(\hat{R})$ തന്നെയാണെന്നു വന്നു ചേരുന്നു. അതു

കൊണ്ടു് $V(\hat{R})$ ൻെറ ഒരു ഏകദേശനം മുകളിൽ കൊടുത്തിരിക്കുന്ന വ്യംജക മാണെന്നു വരുന്നു. അതുൾക്കൊള്ളുന്ന $V(\hat{Y})$, $V(\hat{X})$, R, X, Y, $Cov(\hat{X},\hat{Y})$ എന്നിവയുടെ ആകലങ്ങൾ ആ വ്യംജകകത്തിൽ പ്രതിസ്ഥാപിച്ചാൽ അതിൻെറ ആകലം ലഭിക്കും. ഇവിടെയും സാമ്പിൾ വലുതാണെന്നുള്ള സങ്കൽപം ഒഴിച്ചു കൂടാൻ പാടില്ലാത്തതാണെന്നു് ഓർമയിരിക്കണം.

$\hat{R}$ ൻെറ അഭിനതിയും, സാമ്പിളന പ്രസരണവും നാം കണ്ടുകഴിഞ്ഞു. ഇതു പയോഗിച്ചു് അനുപാതാകലത്തിൻെറ അഭിനതിയും പ്രസരണവും നിർണയിക്കുകയാണു് അടുത്തപടി.

$Y_R = \hat{R}X$ എന്ന സമീകരണമുപയോഗിച്ചാണു് Y യുടെ അനുപാതാകലമായ Y_R നിർണയിക്കുന്നതു്. ഇതു് അഭിനതമാണോ എന്നു് ആദ്യമായി പരിശോധിക്കാം. X മററു തരത്തിൽ കണ്ടുപിടിച്ച x മൂല്യങ്ങളുടെ തുകയായതുകൊണ്ടു് അതു് ഒരു അചരമാണു്. അപ്പോൾ,

$$E(Y_R) = E(\hat{R})X$$

$E(\hat{R})$ ൻെറ ആദ്യ ഏകദേശനം (6.3.1) ഉപയോഗിച്ചാൽ

$$E(Y_R) = \frac{Y}{X} \cdot X = Y.$$

അതായതു്, Y_R, Y യുടെ ഒരു അനഭിനത ആകലമാണു്. പക്ഷേ $E(\hat{R})$ ൻെറ ആദ്യ ഏകദേശനം വളരെ സ്ഥൂലമാണെന്നു നമുക്കറിയാം. അതു കൊണ്ടു് രണ്ടാം ഏകദേശനം (6.3.2) ഉപയോഗിക്കുന്നതാണു് കൂടുതൽ നല്ലതു്. അങ്ങനെ വന്നാൽ,

$$E(Y_R) = \left[R + R\left\{\frac{V(\hat{X})}{X^2} - \frac{Cov(\hat{X}\hat{Y})}{XY}\right\}\right] X$$

$$= RX + \frac{1}{X}\left[RV(\hat{X}) - Cox(\hat{V}\hat{Y})\right]$$

$$= Y + \frac{1}{X}\left[RV(\hat{X}) - Cov(\hat{X}\hat{Y})\right] \qquad (6.3.6)$$

ഇതിൽ നിന്നു് Y_R ൻെറ അഭിനതി $b(Y_R)$ ആണെങ്കിൽ,

$$b(Y_R) = E(Y_R) - Y = \frac{1}{X}\left[RV(\hat{X}) - \text{Cov}(\hat{X}\hat{Y})\right] \qquad 6.3.7$$

എന്നു വരുന്നു. ഇതുപോലെതന്നെ,

$$V(Y_R) = X^2 V(\hat{R})$$

$$= X^2 R^2\left[\frac{V(\hat{Y})}{Y^2} + \frac{V(\hat{X})}{X^2} - \frac{2\,\text{Cov}(\hat{X}\hat{Y})}{XY}\right]$$

$$= V(\hat{Y}) + R^2 V(\hat{X}) - 2R\,\text{Cov}(\hat{X}\hat{Y}) \qquad 6.3.8$$

Y യുടെ അനഭിനത ആകലമായ $\hat{Y}$ നാം $\hat{R}$-ൽ ഉപയോഗിക്കുന്നുണ്ട്. ഈ ആകലവും Y_R എന്ന അനുപാതാകലവും താരതമ്യപ്പെടുത്തി നോക്കാം. അവയുടെ സാമ്പിളനപ്രസരണങ്ങൾ താരതമ്യപ്പെടുത്തുകയാണല്ലോ ഇതിനു ചെയ്യാവുന്നതു്.

Y_R കൂടുതൽ സൂക്ഷ്മതയുള്ളതാവണമെങ്കിൽ $V(Y_R)$, $V(\hat{Y})$ നെക്കാൾ ചെറുതായിരിക്കണം. അതായതു്,

$$V(\hat{Y}) > V(Y_R)$$

അതായതു്,

$$V(\hat{Y}) > V(\hat{Y}) + R^2V(\hat{X}) - 2R\text{Cov}(\hat{X}\hat{Y})$$

അതായതു്,

$$2R\text{Cov}(\hat{X}\hat{Y}) > R^2V(\hat{X}).$$

R ഒരു ധനസംഖ്യയാണെന്നു സങ്കല്പിച്ചാൽ ഇതു്,

$$\text{Cov}(\hat{X}\hat{Y}) > \tfrac{1}{2} R\, V(\hat{X})$$

അതായതു്,

$$\frac{\text{Cov}(\hat{X}\hat{Y})}{\sqrt{V(\hat{X})\,V(\hat{Y})}} > \tfrac{1}{2} R \sqrt{V(\hat{X})/V(\hat{Y})}$$

$$> \tfrac{1}{2}\,\frac{Y}{X}\,\frac{\sigma_{(\hat{X})}}{\sigma_{(\hat{Y})}}.$$

$$\text{അതായതു്,}\quad \rho\,(\hat{X}\hat{Y}) > \frac{1}{2}\,\frac{\dfrac{\sigma_{\hat{X}}}{\bar{X}}}{\dfrac{\sigma_{\hat{Y}}}{\bar{Y}}} > \frac{1}{2}\,\frac{C(\hat{X})}{C(\hat{Y})} \qquad 6.3.9.$$

ഇവിടെ $\rho\,(\hat{X}\hat{Y})$ എന്നതു് $\hat{X}$, $\hat{Y}$ എന്നിവ തമ്മിലുള്ള സഹബന്ധഗുണാങ്കവും $C(\hat{X})$ ഉം $C(\hat{Y})$ ഉം $\hat{X}$ ൻെറയും $\hat{Y}$ ൻെറയും പ്രസരണഗുണാങ്കവുമാണു്.

y എന്ന അഭിലക്ഷണവുമായി ശക്തിയായ നേർസഹബന്ധമുള്ള ഒരു അഭിലക്ഷണമാണല്ലൊ x. ആ സ്ഥിതിക്കു് $\rho\,(\hat{X}\hat{Y})$, ധനരൂപത്തിലുള്ളതും 1 നോടു് വളരെ അടുത്തതും ആയിരിക്കും. y യെക്കാൾ പ്രകീർണനം കുറവുള്ളതാണു് x എങ്കിൽ $C(\hat{X})$, $C(\hat{Y})$ യെക്കാൾ ചെറുതായിരിക്കും. അങ്ങനെ 6.3.9 ലെ നിബന്ധന അനുസരിക്കപ്പെടാൻ നല്ല സാധ്യതയുണ്ടു്. ഇതിൻെറ അർത്ഥം Y_R എന്ന ആകലം $\hat{Y}$ എന്ന അനഭിനത ആകലത്തെക്കാൾ സൂക്ഷ്മതയുള്ളതാണു് എന്നാണു്. അതുകൊണ്ടാണു് മുൻപറഞ്ഞ വൈശിഷ്ട്യങ്ങളുള്ള x എന്ന അഭിലക്ഷം ഉണ്ടായിരിക്കുകയും അഭിലക്ഷണമൂല്യങ്ങളുടെ തുകയായ X മററു തരത്തിൽ അറിയാമായിരിക്കുകയും ചെയ്താൽ, അനുപാതാകലത്തിനു് മുൻഗണന നൽകാൻ നാം പ്രേരിതരാകുന്നതു്.

Y_R ൻെറ അഭിനതിയും സാമ്പിളനപ്രസരണവും ആകലനം ചെയ്യാൻ $\hat{R}$ ൻെറ അഭിനതിയും പ്രസരണവും ആകലനം ചെയ്യാൻ അവലംബിച്ച രീതി തന്നെ സ്വീകരിക്കാം; രണ്ടിലെയും അജ്ഞാതരാശികളുടെ അനഭിനത ആകലങ്ങൾ അവയുടെ സ്ഥാനത്തു് പ്രതിസ്ഥാപിക്കുക എന്ന രീതി. ഈ ആകലങ്ങൾ അനഭിനതങ്ങളായിരിക്കണമെന്നില്ല എന്നതു് പ്രത്യേകം പ്രസ്താവ്യമത്രെ.

ഈ ചർച്ചയുടെ ആരംഭത്തിൽ R ഒരു ധനസംഖ്യയാണെന്നു് നാം സങ്കൽപിക്കുകയുണ്ടായി. പക്ഷേ ഇതു് എപ്പോഴും ശരിയായിരിക്കണമെന്നില്ല. X, Y എന്നിവയിൽ ഒന്നു് ധനസംഖ്യയും മറേറതു് ഋണസംഖ്യയുമായാൽ R ഋണസംഖ്യയാവും. അങ്ങനെയുള്ള സന്ദർഭങ്ങളിൽ അനുപാതാകലം അനഭിനതആകലത്തെക്കാൾ സൂക്ഷ്മതയുള്ളതായിരിക്കണമെങ്കിൽ,

$$\rho\,(\hat{X}\hat{Y}) < -\frac{1}{2}\,\frac{C(\hat{X})}{C(\hat{Y})} \qquad 6.3.10$$

എന്ന നിബന്ധന അനുസരിക്കപ്പെടണം.

4. അനുപാതാകലനം വിവിധസാമ്പിളന സമ്പ്രദായങ്ങളിൽ

അനുപാതാകലത്തെ നിർവചിക്കാനും അതിൻെറ അഭിനതിയും സാമ്പിളനപ്രസരണവും നൽകുന്ന വ്യംജകങ്ങൾ വ്യുൽപാദിപ്പിക്കുവാനുമാണല്ലോ ഇതുവരെ ശ്രമിച്ചതു്. ഏതെങ്കിലു പ്രത്യേക സാമ്പിളനരീതിയുമായി ബന്ധപ്പെട്ടത്തിയല്ല ഇവ വ്യുൽപാദിപ്പിച്ചതു്. ആ വ്യംജകങ്ങൾ പൊതു സ്വഭാവത്തോടു കൂടിയവയത്രെ. ലഘുയാദൃച്ഛികസാമ്പിളനം, സ്തരിത സാമ്പിളനം ക്രമാനുസൃത സാമ്പിളനം എന്നീ മൂന്നു സമ്പ്രദായങ്ങളാണു് നാം പരിഗണിച്ചതു്. അനുപാതാകലത്തിൻെറ അഭിനതിയും പ്രസരണവും കാണിക്കുന്ന വ്യംജകങ്ങൾ ഈ ഓരോ സാമ്പിളനരീതിയിലും ഏതു രൂപമാണു് എടുക്കുക എന്നു പരിശോധിക്കാം.

(1) ലഘുയാദൃച്ഛികസാമ്പിളനം

പ്രതിസ്ഥാപനമില്ലാത്ത ലഘുയാദൃച്ഛികസാമ്പിളനരീതിയാണു് അവലംബിക്കുന്നതെന്നിരിക്കട്ടെ. Y_R ൻെറ അഭിനതിയായ $b(Y_R)$, 6.3.7 ൽ നൽകിയിരിക്കുന്ന സമീകരണമനുസരിച്ചു്,

$$b(Y_R) = \frac{1}{X}\left[RV(\hat{X}) - Cov(\hat{X}\hat{Y})\right]$$

$\bar{x}$, $\bar{y}$ എന്നിവയാണു് സാമ്പിൾസമാന്തരമാധ്യങ്ങളെങ്കിൽ,

$$\hat{X} = N\bar{x},\ \hat{Y} = N\bar{y}\,.$$

പ്രതിസ്ഥാപനമില്ലാത്ത ലഘുയാദൃച്ഛികസാമ്പിളനത്തിൽ

$$V(\bar{x}) = \frac{N-n}{Nn} S_x^2 \qquad \left(S_x^2 = \frac{1}{N-1}\sum_{i=1}^{N}(x_i-\bar{x})^2\right)$$

$$V(\bar{y}) = \frac{N-n}{Nn} S_y^2 \qquad \left(S_y^2 = \frac{1}{N-1}\sum_{i=1}^{N}(y_i-\bar{y})^2\right)$$

അതുകൊണ്ടു്,

$$V(\hat{X}) = N^2V(\bar{x}) = \frac{N(N-n)}{n} S_x^2$$

$$V(\hat{Y}) = N^2V(\bar{y}) = \frac{N(N-n)}{n} S_y^2$$

$\hat{X}$, $\hat{Y}$ എന്നിവ തമ്മിലുള്ള സഹബന്ധഗുണാങ്കമാണു് ρ എങ്കിൽ,

$$Cov(\hat{X}\hat{Y}) = \rho\sqrt{V(\hat{X})\,V(\hat{Y})} = \rho\,\frac{N(N-n)}{n} S_xS_y\,.$$

അങ്ങനെ,

$$b(\hat{Y}_R) = \frac{1}{\bar{X}}\left(R \cdot \frac{N(N-n)}{n} S^2_x - \rho \frac{N(N-n)}{n} S_x S_y\right)$$

$$= \frac{N-n}{n\bar{X}} [R S^2_x - \rho S_x S_y] \qquad 6.4.1$$

ഇതിൽനിന്നു് സാമ്പിൾ പരിമാണമായ n വർധിക്കുന്നതനുസരിച്ചു് അഭിനതി കുറയുമെന്നു കാണാം. താരതമ്യേന വലിയ സാമ്പിളുകളെ സംബന്ധിച്ചിടത്തോളം ഈ അഭിനതി വളരെ നിസ്സാരമാണ്. ശ്രദ്ധിക്കേണ്ട മറ്റൊരു വസ്തുത അഭിനതി ധനമോ ഋണമോ ആകാമെന്നുള്ളതാണ്.

അടുത്തതായി Y_R ന്റെ സാമ്പിളന പ്രസരണം എന്താണെന്നു പരിശോധിക്കാം. 6.3.8 അനുസരിച്ചു്,

$$V(Y_R) = V(\hat{Y}) + R^2\ V(\hat{X}) - 2R\ \mathrm{Cov}\ (\hat{X}\hat{Y})$$

സാമ്പിളെടുത്തിരിക്കുന്നതു് പ്രതിസ്ഥാപനമില്ലാത്ത ലഘുയാദൃച്ഛികസാമ്പിളനമനുസരിച്ചായതു കൊണ്ടു്,

$$V(Y_R) = \frac{N(N-n)}{n} S^2_y + \frac{R^2 N(N-n)}{n} S^2_x - 2R\rho \frac{N(N-n)}{n} S_x S_y$$

$$= \frac{N(N-n)}{n} [S^2_y + R^2 S^2_x - 2\rho R S_x S_y] \qquad 6.4\ 2$$

ഇതിന്റെ വർഗമൂലമാണ് Y_R ന്റെ മാനകപ്പിശക്. ഇതും n വർധിക്കുന്നതനുസരിച്ചു് കുറയുന്നു. പക്ഷേ അഭിനതിയിൽ വരുന്ന ക്രമത്തിലുള്ള കുറവു് ഇതിൽ വരുന്നില്ലെന്നുള്ള വസ്തുത പ്രത്യേകം ശ്രദ്ധേയമാണ്. അഭിനതി $\frac{1}{n}$ ക്രമത്തിൽ കുറയുമ്പോൾ മാനകപ്പിശക് $\frac{1}{\sqrt{n}}$ ക്രമത്തിലേ കുറയുന്നുള്ളല്ലൊ.

അടുത്തതായി പരിഗണിക്കാനുള്ളതു് ഈ അഭിനതിയും പ്രസരണവും ആകലനം ചെയ്യാനുള്ള മാർഗങ്ങളാണ്. സാമ്പിൾ താരതമ്യേന വലുതാണെങ്കിൽ പ്രാചലങ്ങൾക്കു പകരം അവയുടെ ആകലങ്ങൾ പ്രതിസ്ഥാപിച്ചാൽ പ്രാചലഫലനങ്ങളുടെ ഏകദേശആകലങ്ങൾ ലഭിക്കും. ഈ തത്വം ഉപയോഗിച്ചു് 6.4.1 ലും, 6.4.2 ലും $\bar{X}$, R, S^2_x, S^2_y, ρ എന്നിവയ്ക്കു് പകരം അവയുടെ ആകലങ്ങളായ,

$$\bar{x},\ \hat{R} = \frac{\bar{y}}{\bar{x}},\ S^2_x, S^2_y, r$$

എന്നിവ പ്രതിസ്ഥാപിച്ചാൽ,

$$\hat{b}(Y_R) = \frac{N-n}{n\bar{x}}\left[\frac{\bar{y}}{\bar{x}} s^2_x - r\, s_x s_y\right]$$

$$= \frac{(N-n)\,\bar{y}}{n}\left[\frac{s^2_x}{\bar{x}^2} - \frac{r\, s_x s_y}{\bar{x}\,\bar{y}}\right] \qquad 6.4.3$$

$$\hat{V}(Y_R) = \frac{N(N-n)}{n}\left[s^2_y + \frac{\bar{y}^2}{\bar{x}} s^2_x - 2r\frac{\bar{y}}{\bar{x}} s_x s_y\right] \qquad 6.4.4$$

ഈ ആകലങ്ങൾ ഏകദേശനങ്ങളാണെന്നും അനഭിനതങ്ങളായിരിക്കണമെന്ന് നിർബന്ധമില്ലെന്നും ഓർമ്മയിരിക്കണം.

പ്രസരണത്തിൻെറ ഈ ആകലവ്യംജകത്തെ മററൊരു തരത്തിലും വ്യൽപാദിപ്പിക്കാറുണ്ട്.

$$Y_R = \hat{R}X = \frac{\bar{y}}{\bar{x}} X$$

ഇരുവശത്തുനിന്നും Y കുറച്ചാൽ

$$Y_R - Y = \frac{\bar{y}}{\bar{x}} X - Y$$

$$= X\left(\frac{\bar{y}}{\bar{x}} - \frac{Y}{X}\right)$$

$$= X\left(\frac{\bar{y}}{\bar{x}} - R\right)$$

$$= \frac{N\bar{X}}{\bar{x}}(\bar{y} - R\bar{x})$$

സാമ്പിൾ പരിമാണം വലുതാണെങ്കിൽ സാമ്പിൾസമാന്തരമാധ്യമായ $\bar{x}$ സമഷ്ടിയുടെ സമാന്തരമാധ്യമായ $\bar{X}$ മായി വളരെ വ്യത്യസ്തമായിരിക്കുകയില്ലെന്നു പ്രതീക്ഷിക്കാം. അങ്ങനെയെങ്കിൽ,

$$Y_R - Y = N(\bar{y} - R\bar{x})$$

$$V(Y_R) = N^2 V(\bar{w}), \qquad \text{ഇവിടെ } \bar{w} = \bar{y} - R\bar{x}$$

$w_i = y_i - Rx_i$ എന്നും $\bar{W} = \frac{1}{N}\sum_{i=1}^{N} w_i$ എന്നും സങ്കല്പിക്കുക.

അപ്പോൾ $S^2_w = \frac{1}{N-1}\sum_{i=1}^{N}(w_i - \bar{W})^2$ എന്നാവും.

പക്ഷേ, $\overline{W} = \frac{1}{N}\sum_{i=1}^{N} wi = \frac{1}{N}\sum_{j=1}^{N} (y_i - Rx_i)$

$$= \frac{1}{N}(Y-Rx) = \frac{1}{N}\left(Y - \frac{Y}{X}X\right) = 0$$

അതുകൊണ്ട്, $S^2_w = \frac{1}{N-1}\sum_{i=1}^{N} w_i^2 = \frac{1}{N-1}\sum_{i=1}^{N}(y_i - Rx_i)^2$

ലഘുയാദൃച്ഛിക സാമ്പിളനത്തിൽ (പ്രതിസ്ഥാപനമില്ലാത്തതു്.)

$$V(\overline{w}) = \frac{N-n}{Nn} S^2_w = \frac{N-n}{Nn}\frac{1}{N-1}\sum_{i=1}^{N}(y_i - Rx_i)^2$$

അങ്ങനെ, $V(Y_R) = N^2 V(\overline{w}) = \frac{N(N-n)}{n(N-1)}\sum_{i=1}^{N}(y_i - Rx_i)^2$

ഇതിൽ $\frac{1}{N-1}\sum_{i=1}^{N}(y_i - Rx_i)^2$ എന്നതിൻെറ ആകലം

$$\frac{1}{n-1}\sum_{i=1}^{n}(y_i - \hat{R}x_i)^2 \text{ ആണ്.}$$

അങ്ങനെ $\hat{V}(Y_R) = \frac{N(N-n)}{n(n-1)}\sum_{i=1}^{n}(y_i - \hat{R}x_i)^2$

$$= \frac{N(N-n)}{n(n-1)}\left[\sum_{i=1}^{n} y_i^2 - 2\hat{R}\sum_{i=1}^{n} y_i x_i + \hat{R}^2 \sum_{i=1}^{n} x_i^2\right] \qquad 6.4.5$$

താത്വികമായി അത്ര സൂക്ഷ്മമല്ലെങ്കിലും, പ്രായോഗികാവശ്യങ്ങൾക്കെല്ലാം ഉതകുന്ന ഒരു ആകലമാണ് ഇതു്. ഗണനത്തിനു് അത്യന്തം സൗകര്യപ്രദമായ ഒരു രൂപമാണ് ഇതെന്നു് പ്രത്യേകം പറയേണ്ടതുണ്ട്.

$\hat{R}$ ൻെറ സാമ്പിളനപ്രസരണം ആകലനം ചെയ്യാനും ഈ രീതി ഉപയോഗപ്പെടുത്താം.

$Y_R = \hat{R} \times$ അതുകൊണ്ട് $\frac{Y_R}{X} = \hat{R}$, അങ്ങനെ

$V(\hat{R}) = V\left(\frac{Y_R}{X}\right) = \frac{1}{X^2}V(Y_R)$ എന്നു സിദ്ധിക്കുന്നു.

$V(Y_R)$ ൻെറ ആകലനം നാം ചർച്ചചെയ്തു കഴിഞ്ഞു, അതിനെ X^2 കൊണ്ടു ഹരിച്ചാൽ $V(\hat{R})$ കിട്ടും. X^2 ൻെറ ആകലമായി $(N\overline{x})^2$ സ്വീകരിക്കാം.

അങ്ങനെ,

$$\hat{V}(\hat{R}) = \frac{N-n}{Nn\,(n-1)\,\bar{x}^2} \left[\sum_{i=1}^{n} y^2_i + \hat{R}^2 \sum_{i=1}^{n} x^2_i - 2\hat{R} \sum_{i=1}^{n} x_i y_i \right] \qquad 6.4.6$$

ഈ ആകലങ്ങളുടെ പ്രധാന പ്രയോജനം Y യുടെയും R–ൻെറയും വിശ്വാസ്യതാന്തരാളങ്ങൾ കണ്ടുപിടിക്കാമെന്നുള്ളതാണ്. സാമ്പിൾ പരിമാണം വല്യതാണെങ്കിൽ, $\dfrac{\hat{Y}_R - Y}{\sqrt{\hat{V}(Y_R)}}$ എന്നതു് "0" മാധ്യവും 1 മാനകവിചലനവുമായ, നോർമൽ വിതരണം അനുവർത്തിക്കും എന്നു് സങ്കൽപിക്കുന്നതിൽ വലിയ തകരാറില്ല. ഈ സങ്കൽപത്തിൻെറ അടിസ്ഥാനത്തിൽ Y യുടെ വിശ്വാസ്യതാന്തരാളം താഴെ പറയും പ്രകാരം നിർണയിക്കാം. വിശ്വാസ്യതാഗുണാങ്കം a ആയിരിക്കണമെന്നിരിക്കട്ടെ, നോർമൽ പട്ടിക ഉപയോഗിച്ചു് t മാനക നോർമൽ വിതരണം അനുവർത്തിക്കുന്നെങ്കിൽ $P(|t| > t_\alpha) = 1 - a$ ആയിരിക്കത്തക്ക വണ്ണം t_α കണ്ടുപിടിക്കുക.

അപ്പോൾ,

$$P\left\{ \left| \frac{\hat{Y}_R - Y}{\sqrt{\hat{V}(Y_R)}} \right| > t_\alpha \right\} = 1 - a$$

എന്നു വരും. അതായതു്,

$$\hat{Y}_R \pm t_\alpha \sqrt{\hat{V}(Y_R)}$$

എന്ന അന്തരാളം Y യെ ഉൾക്കൊള്ളുവാനുള്ള സംഭാവ്യത a യാണ്. മറ്റൊരു തരത്തിൽ പറഞ്ഞാൽ, a വിശ്വാസ്യതാഗുണാങ്കമായിട്ടുള്ള Y യുടെ വിശ്വാസ്യതാന്തരാളം $\hat{Y}_R \pm t_\alpha \sqrt{\hat{V}(Y_R)}$ എന്നതാണ്. ഇതു പോലെ തന്നെ R ൻെറ a വിശ്വാസ്യതാന്തരാളം

$\hat{R} \pm t_\alpha \sqrt{\hat{V}(\hat{R})}$ എന്നു കാണിക്കാവുന്നതത്രെ.

പലപ്പോഴും $\bar{Y}$ ആയിരിക്കും ആകലനം ചെയ്യേണ്ടതു്. അതു് Y യുടെ ആകലമായ Y_R നെ N കൊണ്ടു് ഹരിച്ചാൽ കിട്ടും. അതിൻെറ അഭിനതിയും പ്രസരണവും താഴെ കൊടുത്തിരിക്കുന്ന വിധം കണ്ടുപിടിക്കാം.

$$\bar{Y}_R = \frac{Y_R}{N} \qquad 6.4.7$$

$$b(\bar{Y}_R) = \frac{b(Y_R)}{N} = \frac{N-n}{Nn} [RS^2_x - \rho S_x S_y] \qquad 6.4.8$$

$$V(\bar{Y}_R) = \frac{1}{N^2} V(\bar{Y}_R) = \frac{N-n}{Nn} [S^2_y + R^2 S^2_x - 2\rho R S_x S_y] \qquad 6.4.9$$

ഇവയുടെ ആകലങ്ങൾ കാണുവാനും മുൻപു് അവലംബിച്ച മാർഗങ്ങൾ തന്നെ സ്വീകരിക്കാം. അതായതു് $\bar{X}$, R, S_x, S_y, ρ എന്നിവയ്ക്കു് പകരം അതിന്റെ സാമ്പിൾ മൂല്യങ്ങൾ പ്രതിസ്ഥാപിക്കുക. അങ്ങനെ,

$$\hat{b}(\bar{Y}_R) = \frac{N-n}{Nn\bar{x}} \left[\frac{\bar{y}}{\bar{x}} s^2_x - r\, s_x s_y \right] \qquad 6.4.10$$

$$\approx \frac{\bar{y}}{n} \left[\frac{s_x^2}{\bar{x}^2} - r \frac{s_x s_y}{\bar{x}\,\bar{y}} \right] \qquad 6.4.11$$

$$\hat{V}(\bar{Y}_R) = \frac{N-n}{Nn} \left\{ s_y^2 + \frac{\bar{y}^2}{\bar{x}^2} s^2_x - \frac{2r\bar{y}}{\bar{x}} s_x s_y \right\} \qquad 6.4.12$$

ഈ ആകലങ്ങൾ ഏകദേശനങ്ങളാണെന്നു് പ്രത്യേകം ഓർമയിരിക്കണം.

2. സ്തരിത സാമ്പിളനം

സമഷ്ടിയെ K സ്തരങ്ങളായി വിഭജിച്ചിരിക്കുന്നു. j–ാമതു് സ്തരത്തിൽ നിന്നു് n_j അംഗങ്ങളെ സാമ്പിളിൽ ഉൾപ്പെടുത്തുന്നു. $(j=1, 2 \ldots\ldots k)$. സാമ്പിളിലെ ഓരോ അംഗത്തിൽ നിന്നും y യുടെയും x ന്റെയും മൂല്യങ്ങൾ കണ്ടുപിടിക്കുന്നു. ഇത്രയും സങ്കല്പിച്ചു കൊണ്ടാണു് ആകലനത്തെപ്പററി ഇവിടെ ചർച്ച ചെയ്യുന്നതു്. സമഷ്ടിയിലെ y മൂല്യങ്ങളുടെ തുകയായ Y യെയാണു് ആകലനം ചെയ്യേണ്ടതു് എന്നിരിക്കട്ടെ. ഇതിനായി പല മാർഗങ്ങളും നിർദേശിക്കപ്പെട്ടിട്ടുണ്ടു്. സാധാരണ ഉപയോഗത്തിലുള്ള രണ്ടു രീതികൾ വിഭിന്ന അനുപാത ആകലനവും സംയുക്ത അനുപാത ആകലനവുമാണു്.

വിഭിന്ന അനുപാത രീതിയിൽ ഓരോ സ്തരത്തിലെ അംഗങ്ങളുടെയും y മൂല്യങ്ങളുടെ തുകയുടെ അനുപാതാകലങ്ങൾ ആദ്യം കണ്ടു പിടിക്കുന്നു. അങ്ങനെ ലഭിക്കുന്ന ആകലങ്ങളുടെ തുകയെ Y യുടെ ആകലമായി സ്വീകരിക്കുന്നു. ഓരോ സ്തരത്തിലെ അംഗങ്ങളുടെയും x മൂല്യങ്ങളുടെ തുക മററു തരത്തിൽ കണ്ടുപിടിക്കണം എന്നതാണു് ഈ രീതിയുടെ ഒരു ബുദ്ധിമുട്ടു്. $\bar{x}_j$, $\bar{y}_j$ എന്നിവ j–ാമതു് സ്തരത്തിൽ നിന്നു് സാമ്പിളിൽ ഉൾപ്പെടുത്തിയ അംഗങ്ങളുടെ x, y അഭിലക്ഷണ മൂല്യങ്ങളുടെ സമാന്തരമാധ്യങ്ങളും X_j, ആ സ്തരത്തിലെ അംഗങ്ങളുടെ x മൂല്യങ്ങളുടെ തുകയും ആണെന്നു സങ്കല്പിക്കുക. അതിലെ അംഗങ്ങളുടെ y മൂല്യങ്ങളുടെ തുകയുടെ അനുപാതാകലത്തെ Y_{Rj} കൊണ്ടു കുറിക്കാം. 6.2.1

അനുസരിച്ച്,

$$Y_{Rj} = \hat{R}_j X_j \quad \text{ഇവിടെ } \hat{R}_j = \frac{\bar{y}_j}{\bar{x}_j}$$

$$Y_R = \sum_{j=1}^{k} Y_{Rj} = \sum_{j=1}^{k} \frac{\bar{y}_j}{\bar{x}_j} X_j \qquad 6.4.13$$

ഓരോ സ്തരത്തിൽ നിന്നും സാമ്പിളിലേക്ക് അംഗങ്ങളെ തിരഞ്ഞെടുക്കുന്നത് ലഘുയാദൃച്ഛിക സാമ്പിളന രീതിയനുസരിച്ചാണ്. അതു കൊണ്ട് Y_{Rj} യുടെ അഭിനതിയും പ്രസരണവും ലഘുയാദൃച്ഛിക സാമ്പിളനത്തിന്റേതു തന്നെ. 6.4.1 അനുസരിച്ച്,

$$b(Y_{Rj}) = \frac{N_j - n_j}{n_j \bar{X}_j} \left[R_j S^2_{xj} - \rho_j S_{xj} S_{yj} \right]$$

ഇവിടെ N_j, j-ാമത് സ്തരത്തിലെ അംഗങ്ങളുടെ എണ്ണം. $\bar{X}_j$, j-ാമത് സ്തരത്തിലെ അംഗങ്ങളുടെ X മൂല്യങ്ങളുടെ സമാന്തരമാധ്യം. R_j ഈ സ്തരത്തിലെ അംഗങ്ങളുടെ y മൂല്യങ്ങളുടെയും x മൂല്യങ്ങളുടെയും തുകകളുടെ അനുപാതം. $\left(R_j = \frac{Y_j}{X_j} \right)$. S^2_{xj}, S^2_{yj} എന്നിവ യഥാക്രമം ഈ സ്തരത്തിലെ x മൂല്യങ്ങളുടെയും y മൂല്യങ്ങളുടെയും പ്രസരണങ്ങൾ. ρ_j അവയുടെ സഹബന്ധഗുണാങ്കം.

$$b(Y_R) = \sum_{j=1}^{k} b(Y_{Rj}) = \sum_{j=1}^{k} \frac{N_j - n_j}{n_j \bar{X}_j} \left[R_j S^2_{xj} - \rho_j S_{xj} S_{yj} \right] \qquad 6.4.14$$

ഇതു പോലെ തന്നെ, 6.4.2 അനുസരിച്ച്,

$$V(Y_{Rj}) = \frac{N_j (N_j - n_j)}{n_j} \left[S^2_{yj} + R^2_j S^2_{xj} - 2\rho_j R_j S_{xj} S_{yj} \right]$$

ഓരോ സ്തരത്തിൽ നിന്നും എടുത്ത സാമ്പിളുകൾ സ്വതന്ത്രങ്ങളായതുകൊണ്ട് അവ ഓരോന്നിൽ നിന്നും ലഭിച്ച അനുപാതാകലങ്ങളുടെ തുകയുടെ പ്രസരണം, ഓരോന്നിന്റെയും പ്രസരണത്തിന്റെ തുകയായിരിക്കും. അതുകൊണ്ട്,

$$V(Y_R) = \sum_{j=1}^{k} V(Y_{Rj})$$

$$= \sum^{k} \frac{N_j (N_j - n_j)}{n_j} \left[S^2_{yj} + R^2_j S^2_{xj} - 2\rho_j R_j S_{xj} S_{yj} \right] \qquad 6.4.15$$

ഇവയുടെ ആകലനത്തിനും യാദൃച്ഛിക സാമ്പിളനത്തിൽ നിർദേശിച്ച രീതി തന്നെ സ്വീകരിക്കാവുന്നതാണ്.

സംയുക്ത അനുപാത സമ്പ്രദായത്തിൽ, ഓരോ സ്തരത്തിൽ നിന്നും Y_{Rj} ആകലനം ചെയ്യുന്നില്ല. സ്തരിതസാമ്പിളനത്തിലെ രീതിയിൽ x മൂല്യങ്ങളുടെയും y മൂല്യങ്ങളുടെയും സമാന്തരമാധ്യങ്ങളുടെ ആകലങ്ങളായ $\bar{x}_{st}$, $\bar{y}_{st}$ എന്നിവ കണ്ടുപിടിക്കുക.

$$\bar{x}_{st} = \frac{1}{N} \sum_{j=1}^{k} N_j \bar{x}_j, \quad \bar{y}_{st} = \frac{1}{N} \sum_{j=1}^{k} N_j \bar{y}_j$$

അപ്പോൾ X, Y എന്നിവയുടെ അനഭിനത ആകലങ്ങളാണ് $N\bar{x}_{st}$, $N\bar{y}_{st}$ എന്നിവ. അപ്പോൾ,

$$\hat{R} = \frac{N\bar{y}_{st}}{N\bar{x}_{st}} = \frac{\bar{y}_{st}}{\bar{x}_{st}} \cdot$$

$Y_R = \hat{R} X$ എന്ന നിയമമനുസരിച്ച് Y യുടെ സംയുക്ത അനുപാത രീതിയിലുള്ള ആകലമായ Y_R,

$$Y_R = \frac{\bar{y}_{st}}{\bar{x}_{st}} X.$$ ഇവിടെ X സമഷ്ടിയിലെ x മൂല്യങ്ങളുടെ തുക.

ഇവിടെ ഉള്ള ഒരു സൗകര്യം ഓരോ സ്തരത്തിലെയും x മൂല്യങ്ങളുടെ തുക പ്രത്യേകം പ്രത്യേകം കണ്ടു പിടിക്കേണ്ടതില്ല എന്നതാണ്. വിഭിന്നഅനുപാത രീതിയിൽ ലഭിച്ച ആകലത്തിന്റെ അഭിനതിയും പ്രസരണവും കണ്ടുപിടിക്കാൻ ഉപയോഗിച്ച മാർഗങ്ങൾ തന്നെ ഉപയോഗിച്ച് ഇതിന്റെയും അഭിനതിയും പ്രസരണവും കാണാം. R_j യുടെ സ്ഥാനത്ത് R, X_j യുടെ സ്ഥാനത്ത് X എന്നീ മാററങ്ങളേ വരികയുള്ളു.

$$b(Y_R) = \frac{1}{\bar{X}} \sum_{j=1}^{k} \frac{N_j - n_j}{n_j} \left[R\, S^2_{xj} - \rho_j S_{xj} S_{yj} \right] \qquad 6.4.16$$

$$V(\bar{Y}_R) = \sum_{j=1}^{k} \frac{N_j (N_j - n_j)}{n_j} \left[S^2_{yj} + R^2 S^2_{xj} - 2\rho_j R\, S_{yj} S_{xj} \right] \qquad 6.4.17$$

എന്നു വരുന്നു.

ഈ രണ്ടു തരം ആകലങ്ങളെയും താരതമ്യപ്പെടുത്താൻ അവയുടെ പ്രസരണങ്ങൾ ഉപയോഗപ്പെടുത്താം. $V_s(Y_R)$, $V_c(Y_R)$ എന്നിവ യഥാക്രമം വിഭിന്ന അനുപാത രീതിയിലും സംയുക്ത അനുപാത രീതിയിലും കണ്ടു പിടിച്ച ആകലങ്ങളുടെ പ്രസരണങ്ങളാണെന്നിരിക്കട്ടെ.

(6.4.15), (6.4.17) എന്നിവയനുസരിച്ചു്,

$$V_e(Y_R) - V_s(Y_R) = \sum_{j=1}^{k} \frac{N_j(N_j - n_j)}{n_j} \Big[(R^2 - R_j^2) S^2_{xj} - 2(R - R_j) \rho_j S_{xj} S_{yj} \Big]$$

$$= \sum_{j=1}^{k} \frac{N_j(N_j - n_j)}{n_j} \Big[(R - R_j)^2 S^2_{xj} + 2(R_j - R)(\rho_j S_{yj} S_{xj} - R_j S^2_{xj}) \Big]$$

വലതുവശത്തെ രണ്ടാമത്തെ പദം വളരെ ചെറുതായിരിക്കാനാണു് സാധ്യത. ഒരോ സ്തരത്തിലും മൂലബിന്ദുവിലൂടെ കടന്നു പോകുന്ന സമാശ്രയണരേഖയാണു് x, y അഭിലക്ഷണങ്ങൾ തമ്മിലുള്ള ബന്ധം കുറിക്കുന്നതെങ്കിൽ ആ പദം പൂജ്യമായിരിക്കും. ഏകദേശമായെങ്കിലും അങ്ങനെ ഒരു ബന്ധം നിലവിലുണ്ടെങ്കിൽ മാത്രമേ അനുപാതാകലനത്തിനു് നാം മുതിരുകയുള്ളൂ. ഇതിൽ നിന്നെല്ലാം വന്നു ചേരുന്നതു്. $V_c(Y_R) - V_s(Y_R)$ ഒരു ധനസംഖ്യയായിരിക്കുമെന്നാണു്. അതായതു് വിഭിന്നഅനുപാതരീതിയിൽ കണ്ടു പിടിച്ച ആകലത്തിൻെറ പ്രസരണം സംയുക്തഅനുപാതരീതിയിൽ നിർണയിച്ച ആകലത്തിന്റെ പ്രസരണത്തെക്കാൾ ചെറുതാണു്. അങ്ങനെ വിഭിന്നഅനുപാതആകലനരീതിയാണു് താരതമ്യേന മെച്ചപ്പെട്ടതെന്നു് വന്നു ചേരുന്നു. പക്ഷേ ഓരോ സ്തരത്തിൽ നിന്നും സാമ്പിളിൽ ഉൾപ്പെടുത്തുന്ന അംഗങ്ങളുടെ എണ്ണം (n_j) വലുതാണെങ്കിൽ മാത്രമേ അഭിനതിയേയും പ്രസരണത്തെയും കുറിക്കുന്ന മുൻ പറഞ്ഞ ഏകദേശകവ്യംജകങ്ങൾക്കു് സാധുതയുണ്ടായിരിക്കു. അവയുടെ വ്യുല്പാദനത്തിൽ

$$\left| \frac{\hat{X}_j - X_j}{X_j} \right| < 1$$

എന്ന സങ്കല്പം ഉപയോഗപ്പെടുത്തിയിട്ടുണ്ടു്. വലിയ സാമ്പിളുകളെ സംബന്ധിച്ചു മാത്രമേ ഇതു ശരിയായിരിക്കുമെന്നു് പ്രതീക്ഷിക്കാനാവൂ. സംയുക്ത അനുപാതരീതിയാണു് അവലംബിക്കുന്നതെങ്കിൽ സാമ്പിൾ പരിമാണമായ n വലുതായിരിക്കണമെന്നു മാത്രമേയുള്ളൂ.

3. ക്രമാനുസൃതസാമ്പിളനം

ഇവിടെയും അഭിനതിയുടെയും പ്രസരണത്തിന്റെയും വ്യംജകങ്ങൾ കാണാൻ (6.3.7), (6.3.8) എന്നീവ്യംജകങ്ങളിൽ പ്രസക്തമായ പ്രതിസ്ഥാപനം നിർവഹിച്ചാൽ മതി. $N = nk$ അംഗങ്ങളുള്ള ഒരു സമഷ്ടിയിൽ നിന്നു് n അംഗങ്ങളുള്ള ക്രമാനുസൃത സാമ്പിളാണു് എടുക്കുന്നതു് എന്നു കരുതുക. അങ്ങനെ എടു

ക്കാവുന്ന സാമ്പിളുകളുടെ എണ്ണം k ആണ്. അതിൽ r യാദൃച്ഛികപ്രാരംഭമായുള്ള സാമ്പിളിലെ x, y അഭിലക്ഷണമൂല്യങ്ങളുടെ സമാന്തരമാധ്യങ്ങൾ യഥാക്രമം, $\bar{x}_r$, $\bar{y}_r$ എന്നിരിക്കട്ടെ.

$$Y_R = \hat{R}X$$

$$V(\bar{x}) = \frac{1}{k}\sum_{r=1}^{k} (\bar{x}_r - \bar{X})^2 = \frac{1}{k}\sum_{r=1}^{k} \bar{x}^2_r - \bar{X}^2$$

$$V(\bar{y}) = \frac{1}{k}\sum_{r=1}^{k} (\bar{y}_r - \bar{Y})^2 = \frac{1}{k}\sum_{r=1}^{k} \bar{y}_r^2 - \bar{Y}^2$$

$$Cov(\bar{x}, \bar{y}) = \frac{1}{k}\sum_{r=1}^{k} (\bar{x}_r - \bar{X})(\bar{y}_r - \bar{Y}) = \frac{1}{k}\sum_{r=1}^{k} \bar{x}_r \bar{y}_r - \bar{X}\bar{Y}$$

6.3.7 അനുസരിച്ച്,

$$b(Y_R) = \frac{1}{X}[R\,V(\hat{X}) - Cov(\hat{X}\hat{Y})]$$

$$= \frac{N^2}{X}[R\,V(\bar{x}) - Cov(\bar{x}\ \bar{y})]$$

$$= \frac{N^2}{X}\left[R\cdot\frac{1}{k}\sum_{r=1}^{k} \bar{x}^2_r - R\bar{X}^2 - \frac{1}{k}\sum_{r=1}^{k} \bar{x}_r\bar{y}_r + \bar{X}\bar{Y}\right]$$

$$= \frac{N^2}{X}\left[\frac{1}{k}\cdot R\sum_{r=1}^{k} \bar{x}^2_r - \frac{Y}{X}\cdot\frac{X^2}{N^2} - \frac{1}{k}\sum_{r=1}^{k} \bar{x}_r\bar{y}_r + \frac{XY}{N^2}\right]$$

$$= \frac{N^2}{X}\cdot\frac{1}{k}\sum_{r=1}^{k} [R\bar{x}^2_r - \bar{x}_r y_r]$$

$$= \frac{N}{\bar{X}}\cdot\frac{1}{k}\sum_{r=1}^{k} \bar{x}_r(R\bar{x}_r - \bar{y}_r) \qquad 6.4.18$$

ഇതുപോലെ തന്നെ, 6.3.8 അനുസരിച്ച്,

$$V(Y_R) = V(\hat{Y}) + R^2V(\hat{X}) - 2R\,Cov(\hat{X}\hat{Y})$$

$$= N^2[V(\bar{y}) + R^2\,V(\bar{x}) - 2R\,Cov(\bar{x}\,\bar{y})]$$

$$= N^2 \left[\frac{1}{k} \sum_{r=1}^{k} \bar{y}^2_r - \bar{y}^2 + R^2 \cdot \frac{1}{k} \sum_{r=1}^{k} \bar{x}^2_r - R^2 \bar{X}^2 - 2R \sum_{r=1}^{k} \bar{x}_r \bar{y}_r + 2R \bar{X}\bar{Y} \right]$$

$$= N^2 \left[\frac{1}{k} \sum_{r=1}^{k} \bar{y}^2_r - \frac{Y^2}{N^2} + \frac{R^2}{k} \sum_{r=1}^{k} \bar{x}^2_r - \frac{Y^2}{X^2} \frac{X^2}{N^2} - 2R \sum_{r=1}^{k} \bar{x}_r \bar{y}_r + 2 \frac{Y}{X} \cdot \frac{X}{N} \cdot \frac{Y}{N} \right]$$

$$= \frac{N^2}{k} \sum_{r=1}^{k} [\bar{y}^2_r + \bar{x}^2_r R^2 - 2R \bar{x}_r \bar{y}_r]$$

$$= N^2 \cdot \frac{1}{k} \sum_{r=1}^{k} (\bar{y}_r - R \bar{x}_r)^2 \qquad 6.4.19$$

ഇതിൽ നിന്നു് അനുപാതാകലത്തിൻെറ സൂക്ഷ്മത വർധിക്കുവാനുള്ള മാർഗം, അതായതു് V ചെറുതാകുവാനുള്ള മാർഗം $\frac{\bar{y}_r}{\bar{x}_r}$ ൻെറ പ്രകീർണനം കഴിവതും കുറഞ്ഞിരിക്കത്തക്കവണ്ണം സമഷ്ടിയിലെ അംഗങ്ങളെ ക്രമീകരിക്കുക എന്നതാണെന്നു വ്യക്തമാണു്.

5. അനഭിനത അനുപാതാകലനം

അനുപാതാകലം, ഇതു വരെ പഠിച്ച ഏതു സാമ്പിളന രീതിയിലും അഭിനതമായിരിക്കുമെന്നു നാം കണ്ടു. ഈ തകരാറു് ഒഴിവാക്കാനുള്ള ഒരു മാർഗം സാമ്പിളനരീതിയിൽ ചില പരിഷ്കാരങ്ങൾ വരുത്തുകയാണു്. പല പരിഷ്കരണരീതികളും നിർദേശിക്കപ്പെട്ടിട്ടുണ്ടു്. അവയിലൊന്നു് ഡി.ബി. ലാഹിരിയുടെയും എച്ച്. മിഡ്സുനോയുടേയും താഴെ കൊടുത്തിരിക്കുന്ന നിർദേശമാണു്. സാമ്പിളിലെ ആദ്യത്തെ അംഗത്തെ അതിൻെറ x മൂല്യത്തിനു ആനുപാതികമായ സംഭാവ്യതയോടു കൂടി തിരഞ്ഞെടുക്കുകയും ബാക്കിയുള്ള n—1 അംഗങ്ങളെ പ്രതിസ്ഥാപനമില്ലാത്ത ലഘുയാദൃച്ഛികസാമ്പിളനരീതിയാൽ തിരഞ്ഞെടുക്കുകയും ചെയ്യുക. ഇങ്ങനെ തിരഞ്ഞെടുക്കുന്ന സാമ്പിളിൽ നിന്നു ലഭിക്കുന്ന അനുപാതാകലം അനഭിനതമായിരിക്കും. ഇതു് തെളിയിക്കാൻ വിഷമമില്ല.

ആദ്യമായി തിരഞ്ഞെടുക്കുന്ന അംഗത്തിൻെറ x മൂല്യം $x_1, x_2, \dots x_N$ എന്നിവയിൽ ഏതെങ്കിലും ഒന്നാവാം. ഈ രീതിയനുസരിച്ചു് x_i, x മൂല്യമായിട്ടുള്ള

അംഗത്തെ ആദ്യത്തെ അംഗമായി തിരഞ്ഞെടുക്കുവാനുള്ള സംഭാവ്യതയായ p_i, x_i ക്കു് ആനുപാതികമാണു്. അതായതു് $p_i = kx_i$

$\sum_{i=1}^{N} p_i = 1$ ആയിരിക്കണമല്ലൊ. അതുകൊണ്ടു് $\sum_{i=1}^{N} kx_i = 1$ അതായതു്,

$$k = \frac{1}{\sum_{i=1}^{n} x_i} = \frac{1}{X}$$

അതിൽ നിന്നു് $p_i = \frac{x_i}{X}$ എന്നു സിദ്ധിക്കുന്നു.

അടുത്തതായി ഏതെങ്കിലും ഒരു പ്രത്യേകസാമ്പിൾ കിട്ടുവാനുള്ള സംഭാവ്യത എത്രയെന്നു് പരിശോധിക്കാം. (ഈ രീതിയിൽ എടുക്കാവുന്ന എല്ലാ സാമ്പിളുകളുടെയും സംഭാവ്യത തുല്യമായിരിക്കുമെന്നു് ഓർമിക്കുക.) $u=(u_1, u_2, \cdots u_n)$ എന്ന നിർദിഷ്ടമായ സാമ്പിൾ കിട്ടുവാനുള്ള സംഭാവ്യതയാണു് പരിഗണിക്കുന്നതെന്നിരിക്കട്ടെ. സമഷ്ടിയിലെ അംഗങ്ങളുടെ x മൂല്യങ്ങൾ $x_1, x_2, \cdots x_n$ എന്നിവയാണെന്നു സങ്കൽപിക്കുക. u_i യുടെ x മൂല്യമാണു് x_i എന്നിരിക്കട്ടെ. സാമ്പിളിലെ ആദ്യഅംഗമായി $u_1, u_2, \cdots u_n$ എന്നിവയിൽ ഏതെങ്കിലും കിട്ടാം അങ്ങനെ അന്യോന്യ അപവർജികളായ n വിധങ്ങളിൽ ഈ സാമ്പിൾ ലഭിക്കാം. ആ ഓരോന്നിൻെറയും സംഭാവ്യത ഒന്നിച്ചു കൂട്ടുമ്പോഴാണു് ഈ സാമ്പിളിൻെറ സാഭാവ്യത കിട്ടുക. ആദ്യഅംഗം u_i ആകാനുള്ള സംഭാവ്യത x_i/X ആണെന്നു നാം കണ്ടു കഴിഞ്ഞു. ബാക്കി n—1 അംഗങ്ങൾ സമഷ്ടിയിലെ ബാക്കിയുള്ള N—1 എണ്ണത്തിൽ നിന്നു് എടുക്കാവുന്ന, n—1 ൽ ഒന്നാണു്. അതായതു്, അതിൻെറ സംഭാവ്യത $1/\binom{N-1}{n-1}$ അത്രെ. ആദ്യത്തെ അംഗത്തിൻെറ തിരഞ്ഞെടുപ്പും മറ്റുള്ളവയുടെ തിരഞ്ഞെടുപ്പും സ്വതന്ത്രമാണു്. അങ്ങനെ u എന്ന സാമ്പിൾ ലഭിക്കുവാനുള്ള സംഭാവ്യത,

$$P(u) = \sum_{i=1}^{n} \frac{x_i}{X} \frac{1}{\binom{N-1}{n-1}} = \frac{1}{X\binom{N-1}{n-1}} \sum_{i=1}^{n} x_i$$

$$= \frac{n\bar{x}}{N\bar{X}\binom{N-1}{n-1}} = \frac{1}{\binom{N}{n}} \frac{\bar{x}}{\bar{X}} \qquad 6.5.1$$

ഇതിൽ നിന്നു ലഭിക്കുന്ന അനുപാതാകലത്തിൻെറ പ്രതീക്ഷ എന്താണെന്നു നോക്കാം.

$$Y_R = \hat{R}X = \frac{\bar{y}}{\bar{\bar{x}}} X$$

$$E(Y_R) = \sum_{r=1}^{\binom{N}{n}} \frac{\bar{y}_r}{\bar{\bar{x}}_r} X \cdot \frac{1}{\binom{N}{n}} \cdot \frac{\bar{\bar{x}}_r}{\bar{\bar{X}}}$$

$$= \sum_{r=1}^{\binom{N}{n}} \frac{N\bar{y}_r}{\binom{N}{n}} = \frac{N}{\binom{N}{n}} \sum_{r=1}^{\binom{N}{n}} \bar{y}_r$$

സമഷ്ടിയിലെ ഓരോ അംഗവും $\binom{N-1}{n-1}$ സാമ്പിളുകളിലെ അംഗമായിരിക്കുമെന്നതു കൊണ്ടു്,

$$\sum_{r=1}^{\binom{N}{n}} \bar{y}_r = \sum_{r=1}^{\binom{N}{n}} \sum_{j=1}^{n} \frac{y_{rj}}{n}$$

$$= \binom{N-1}{n-1} \frac{Y}{n}$$

അങ്ങനെ,

$$E(Y_R) = \frac{N}{\binom{N}{n}} \cdot \frac{\binom{N-1}{n-1}}{n} Y = Y$$

ചുരുക്കത്തിൽ Y_R, ഒരു അനഭിനത ആകലമാണു്.

6.5.1 പരിശോധിച്ചാൽ ഈ സാമ്പിളന രീതിയിൽ ഏതെങ്കിലും ഒരു സാമ്പിൾ ലഭിക്കുവാനുള്ള സംഭാവ്യത അതിന്റെ x മൂല്യങ്ങളുടെ സമാന്തരമാധ്യത്തിനു് ആനുപാതികമാണെന്നു കാണാം. ഈ പ്രത്യേകതയെ സാമ്പിളനരീതിയുടെ നിർവചനമായി കൊടുക്കാറുണ്ടു്.

ഈ അനഭിനത അനുപാതാകലത്തിന്റെ സാമ്പിളന പ്രസരണത്തിന്റെ അനഭിനത ആകലം,

$$\hat{V}(\hat{R}) = \hat{R}^2 - \frac{1}{N n \bar{x}\bar{X}} \left\{ \sum_{i=1}^{n} y_i^2 + \frac{N-1}{n-1} \sum_{i=1}^{n} \sum_{j=1}^{n} y_i\, y_j \right\} i \neq j \qquad 6.5.2$$

ആണെന്നു കാണിക്കാൻ കഴിയും. ഇതുപയോഗിച്ച് $\hat{V}(Y_R)$ കണ്ടുപിടിക്കാവുന്നതാണ്.

6. അനഭിനത അനുപാതപ്രരൂപ ആകലങ്ങൾ

മുൻ ഖണ്ഡികയിൽ സാമ്പിളനരീതിയിൽ മാററങ്ങൾ വരുത്തി അനുപാതാകലത്തെ അനഭിനതമാക്കുന്ന ഒരു സമ്പ്രദായമാണ് ചർച്ചചെയ്തതു്. ഇതു് പ്രായോഗികമായി പറഞ്ഞാൽ അത്ര സൗകര്യമുള്ള ഒരു മാർഗമല്ല. അതുകൊണ്ട് സാമ്പിളനരീതിയിൽ മാററമൊന്നും വരുത്താതെ ആകലനരീതിയിൽ മാത്രം ചില പരിഷ്കാരങ്ങൾ വരുത്തി ആകലത്തെ അനഭിനതമാക്കുന്ന സമ്പ്രദായങ്ങൾ ഗവേഷണവിധേയമായിട്ടുണ്ട്. പ്രതിസ്ഥാപനത്തോടുകൂടിയ യാദൃച്ഛിക സാമ്പിളനത്തിൽ,

$$\hat{R} = \bar{r} + \frac{(n-1)n}{N(n-1)} \cdot \frac{(\bar{y} - \bar{r}\bar{x})}{\bar{X}} \qquad \left[\text{ഇവിടെ } \bar{r} = \frac{1}{n} \sum \frac{y_i}{x_i} \right] \qquad 6.6.1$$

എന്ന സമീകരണമുപയോഗിച്ച് $\hat{R}$ ആകലനം ചെയ്യുക എന്നതാണ് ഒരു സമ്പ്രദായം. ഇങ്ങനെ ലഭിക്കുന്ന ആകലം അനഭിനതമാണെന്നു് കാണിക്കുവാൻ കഴിയും. തെളിവു് അല്പം സങ്കീർണമായതുകൊണ്ട് ഇവിടെ കൊടുക്കുന്നില്ല.

7. മററു അനുപാതാകലങ്ങൾ

$Y_R = \hat{R}\, X$ എന്ന സമീകരണമുപയോഗിച്ചുള്ള അനുപാതാകലനമാണ് നാമിതുവരെ പരിഗണിച്ചതു്. സ്തരിതസാമ്പിളനത്തെപ്പററി ചർച്ചചെയ്തപ്പോൾ മാത്രം അല്പം വ്യത്യസ്തമായ ഒരു ആകലനരീതി നിർദേശിക്കുകയുണ്ടായി. എങ്കിലും പൊതുവെ പറഞ്ഞാൽ ആദ്യം പറഞ്ഞ ആകലന സമ്പ്രദായമാണ് സമഗ്രമായ പഠനത്തിനു് വിധേയമായതു്. ഇവിടെ പുതിയ ചില അനുപാതാകലനരീതികൾ അവതരിപ്പിക്കുകയാണ്. അവയെപ്പററി കൂടുതലായി ഒന്നും തന്നെ ചർച്ച ചെയ്യുന്നില്ല.

i. ശൃംഖലിതഅനുപാതാകലം

ഒന്നിലധികം ഘട്ടങ്ങളായുള്ള സാമ്പിളനത്തിലാണ് ഈ രീതി സാധാരണ ഉപയോഗിക്കാറുള്ളതു്. ഉദാഹരണമായി രണ്ട് ഘട്ടങ്ങളാണ് ഉള്ളതെന്നിരിക്കട്ടെ. സമഷ്ടിയിൽ m ആദ്യഘട്ടവ്യക്തികളുണ്ടെന്നു് കരുതുക. 1-ാമത്തെ ആദ്യഘട്ടവ്യക്തിയിൽ N_1 പ്രാഥമികാംശങ്ങളുണ്ടെന്നു സങ്കല്പിക്കുക. ഉദാഹരണമായി കേരളത്തിൽ എത്ര ഏക്കർ സ്ഥലത്തു് നെൽകൃഷി നടത്തുന്നുണ്ടെന്നാണ്

കണ്ടുപിടിക്കേണ്ടിയിരിക്കുന്നതെന്നിരിക്കട്ടെ. വില്ലേജുകളെ ആദ്യഘട്ടവ്യക്തികളായും, കൃഷിസ്ഥലങ്ങളെ രണ്ടാംഘട്ട വ്യക്തികളായും പരിഗണിച്ച് ആ അന്വേഷണം സംഘടിപ്പിക്കാം. ആദ്യം ഒന്നാംഘട്ട വ്യക്തികളുടെ ഒരു സാമ്പിൾ എടുത്ത് അവ ഓരോന്നിൽ നിന്നും ആവശ്യമുള്ളത്ര രണ്ടാംഘട്ട വ്യക്തികളെ തിരഞ്ഞെടുക്കാം. I–ാമത്തെ വില്ലേജിൽ N_I കൃഷിസ്ഥലങ്ങളുണ്ടെന്ന് സങ്കല്പിക്കുക.

സാമ്പിളിലെ i– l ാമത് ആദ്യഘട്ടവ്യക്തിയിൽ നിന്നു തിരഞ്ഞെടുക്കപ്പെട്ട j—ാമത് പ്രാഥമികവ്യക്തിയുടെ x, y അഭിലക്ഷണമൂല്യങ്ങൾ യഥാക്രമം x_{ij} യും y_{ij} യും ആയിരുന്നിരിക്കട്ടെ. n_i ആദ്യഘട്ടവ്യക്തികളെയാണ് സാമ്പിളിൽ ഉൾപ്പെടുത്തിയിരിക്കുന്നതെന്നും, i–ാമത്തെ ആദ്യഘട്ടവ്യക്തിയിൽ നിന്ന് n_i രണ്ടാംഘട്ടവ്യക്തികളെ സാമ്പിളിൽ ഉൾപ്പെടുത്തുന്നു എന്നും കരുതുക. x_i, y_i എന്നിവ i–ാമത് ആദ്യഘട്ടവ്യക്തി ഉൾക്കൊള്ളുന്ന രണ്ടാംഘട്ടവ്യക്തികളുടെ x,y അഭിലക്ഷണമൂല്യങ്ങളുടെ തുകകളാണെന്നും, X,Y എന്നിവ യഥാക്രമം സമഷ്ടിയിലെ എല്ലാ അംഗങ്ങളുടെയും x,y അഭിലക്ഷണമൂല്യങ്ങളുടെ തുകയാണെന്നും സങ്കല്പിക്കുക. Y_{Re} ആണ് Y യുടെ ശൃംഖലിത അനുപാതാകലമെങ്കിൽ,

$$Y_{Rc} = X \frac{\sum_{i=1}^{n_1} x_i u_i}{\sum_{i=1}^{n_1} x_i} \qquad 6.7.1$$

ഇവിടെ $u_i = \dfrac{\sum_{j=1}^{n_i} y_{ij}}{\sum_{j=1}^{n_i} x_{ij}}$ ആണ്.

ഈ ആകലം അനഭിനതമല്ല. അഭിനതി ചെറുതായിരിക്കുമെന്നും പറയാനാവില്ല.

മുൻ ഉദാഹരണത്തോടു ബന്ധപ്പെടുത്തി ഈ ആകലനരീതി താഴെപറയുംപ്രകാരം വിശദീകരിക്കാം. ആദ്യഘട്ടവ്യക്തികളായ വില്ലേജുകളിൽ നിന്ന് n എണ്ണം സാമ്പിളിൽ ഉൾപ്പെടുത്തുന്നു. i–ാമത് വില്ലേജിൽ നിന്ന് n_i കൃഷിസ്ഥലങ്ങളാണ് തിരഞ്ഞെടുക്കുന്നത്. സാമ്പിളിൽ ഉൾപ്പെടുത്തിയ i–ാമത് വില്ലേജിൽ നിന്ന് സാമ്പിളിലേക്കു തിരഞ്ഞെടുക്കപ്പെട്ട j–ാമത് കൃഷിസ്ഥലത്തിന്റെ വിസ്തീർണം x_{ij} യും അതിൽ നെൽകൃഷി ചെയ്യുന്ന സ്ഥലം y_{ij} യുമാണ്. i-ാമത് വില്ലേജിൽ നിന്നു തിരഞ്ഞെടുക്കപ്പെട്ട എല്ലാ കൃഷിസ്ഥലങ്ങളിലെയും കൂടി നെൽകൃഷിസ്ഥലത്തിൻെറ വിസ്തീർണവും, ആകെ വിസ്തീർണവുമായുള്ള അനുപാതമാണ് u_i. i–ാമത് വില്ലേജിലെ ആകെ കൃഷിസ്ഥലത്തിൻെറ വിസ്തീർണം x_i യും നെൽകൃഷിസ്ഥലത്തിൻെറ വിസ്തീർണം y_i യുമത്രെ. നേരത്തെ

വിവരിച്ച അനുപാതാകലന സമ്പ്രദായമനുസരിച്ച്, i–ാമതു് വില്ലേജിലെ നെൽകൃഷിസ്ഥലത്തിൻെറ വിസ്തീർണത്തിൻെറ അനുപാതം R_i ആണെന്നു സങ്കല്പിച്ചാൽ അതിൻെറ ആകലമാണ് u_i; x_i, ആ വില്ലേജിലെ കൃഷിസ്ഥലത്തിന്റെ വിസ്തീർണവും. u_i കളുടെ, x_i കൾ ഭാരങ്ങളായുള്ള, ഭാരിതമാധ്യത്തെയാണ് അതായതു്,

$$\frac{\sum_{i=1}^{n_i} u_i x_i}{\sum_{i=1}^{n_1} x_i}$$

നെയാണ്, ഈ ആകലസമ്പ്രദായത്തിൽ R ൻെറ ആകലമായി സ്വീകരിക്കുന്നതു്. ഈ സമ്പ്രദായത്തിനു്, 6.4 ൽ സ്തരിതസാമ്പിളനത്തോടു ബന്ധപ്പെടുത്തി നിർദേശിച്ച 'വിഭിന്ന അനുപാത' ആകലവുമായുള്ള സാമ്യം ശ്രദ്ധിക്കുക. ആ രീതിയുമായി കൂടുതൽ സാമ്യമുള്ള മറെറാരു ആകലന സമ്പ്രദായമാണ് താഴെ കൊടുക്കുന്നതു്.

2. പശ്ചാൽ സ്തരിത അനുപാതാകലം

സമഷ്ടിയിലെ ചില വിഭാഗങ്ങളിൽ x, y അഭിലക്ഷണങ്ങൾ തമ്മിലുള്ള സഹബന്ധം കൂടുതലാണെന്നു കരുതാൻ ന്യായമുണ്ടെങ്കിലാണ് ഈ രീതി അവലംബിക്കാവുന്നതു്. തത്വത്തിൽ ഈ രീതി വിഭിന്നാനുപാത ആകലനരീതി തന്നെയാണ്. സ്തരണത്തിൻെറ അടിസ്ഥാനം X,Y അഭിലക്ഷണങ്ങൾ തമ്മിലുള്ള സഹബന്ധത്തിൻെറ തീക്ഷ്ണതയാണ് എന്നു മാത്രം. സമഷ്ടിയെ ഇത്തരത്തിൽ K ഭാഗങ്ങളായി (K പശ്ചാൽ സ്തരങ്ങളായി) വിഭജിക്കാമെന്നും, i–ാമത്തെ ഭാഗത്തിലെ x, y, അഭിലക്ഷണ മൂല്യങ്ങളുടെ തുകകളുടെ ആകലങ്ങൾ $\hat{x}_i, \hat{y}_i$. എന്നിവയാണെന്നും, അതിലെ മറേറതെങ്കിലും വിധത്തിൽ കൃത്യമായി കണ്ടുപിടിച്ച x–മൂല്യങ്ങളുടെ തുക X_i ആണെന്നും ഇരിക്കട്ടെ. Y_R ആണ് സമഷ്ടിയിലെ y–മൂല്യങ്ങളുടെ തുകയെങ്കിൽ,

$$Y_R = \sum_{i=1}^{k} \frac{\hat{y}_i}{\hat{x}_i} X_i \qquad 6.7.2$$

ഈ ആകലം കൂടുതൽ ദക്ഷതയുള്ളതായിരിക്കണമെങ്കിൽ, സമഷ്ടി മുഴുവനായി എടുക്കുമ്പോൾ ഉള്ളതിനെക്കാൾ ശക്തമായ സഹബന്ധം ഓരോ ഭാഗത്തിലും x, y കൾ തമ്മിൽ ഉണ്ടായിരിക്കുകയും, അവ തമ്മിലുള്ള സമാശ്രയണ രേഖ മൂലബിന്ദുവിലൂടെ കടന്നുപോവുകയും ചെയ്യണം.

3. ബഹുസഹായകചര ആകലനരീതി

x, y അഭിലക്ഷണങ്ങൾ തമ്മിലുള്ള ബന്ധം ഉപയോഗിച്ച് Y യുടെ ആകലനം കൂടുതൽ സൂക്ഷ്മതയുള്ളതാക്കാനുള്ള ശ്രമമാണല്ലോ ഇതുവരെ നടത്തിയതു്. y യുമായി നേർസഹബന്ധമുള്ള ഒന്നിലധികം, (m എന്നിരിക്കട്ടെ) അഭിലക്ഷണങ്ങളെപ്പററി അറിവുണ്ടെങ്കിൽ, ആ അറിവുകൾ ഉപയോഗപ്പെടുത്തി ആകലത്തെ സൂക്ഷ്മതരമാക്കാനുള്ള ഒരു മാർഗമാണ് ഇവിടെ ചർച്ചചെയ്യുന്നതു്.

$x_1, x_2 \cdots x_m$ എന്നീ m സഹായക അഭിലക്ഷണങ്ങളുടെ തുകകൾ യഥാക്രമം, $X_1, X_2, \cdots X_m$ എന്നിവയാണെന്നിരിക്കട്ടെ. Y യുടെ ആകലമായ Y_R,

$$Y_R = \sum_{i=1}^{m} w_i \frac{\hat{Y}}{\hat{X}_i} X_i \qquad 6.7.3$$

ഇവിടെ $\sum_{i=1}^{m} w_i = 1$ ആണ്.

മറ്റു ചില ഗവേഷകർ y യെ $y_1, y_2 \cdots y_m$ എന്ന m ഭാഗങ്ങളായി വിഭജിക്കണമെന്നു നിർദേശിക്കുന്നു. y_i, x_i യുമായി നേർസഹബന്ധമുള്ളതു് എന്നു് എങ്കല്പിച്ച്,

$$Y_R = \sum_{i=1}^{m} \frac{\hat{Y}_i}{\hat{X}_i} X_i \qquad 6.7.4$$

എന്നു് ആകലത്തെ നിർവചിക്കുകയും ചെയ്യുന്നു.

മറ്റു പലതരം അനുപാതാകലങ്ങളും നിർദേശിക്കപ്പെട്ടിട്ടുണ്ട്. മാതൃകയായി മൂന്നെണ്ണം ഇവിടെ കൊടുത്തെന്നേയുള്ളു.

8. ഗുണനഫലആകലം

ഇതുവരെ നാം പരിഗണിച്ചതു് y എന്ന അഭിലക്ഷണവുമായി ശക്തമായ നേർസഹബന്ധമുള്ള x എന്ന സഹായക അഭിലക്ഷണം നിലവിലുള്ളപ്പോഴത്തെ ആകലനരീതിയാണ്. ചിലപ്പോൾ പരിഗണിക്കപ്പെടുന്ന അഭിലക്ഷണവുമായി ശക്തമായ വിപരീത സഹബന്ധമുള്ള ഒരു അഭിലക്ഷണമാവാം അറിവിൽ പെടുന്നതു്. അതിനെ സംബന്ധിച്ച വിവരങ്ങൾ ഉപയോഗിച്ച് Y യുടെ ആകലനം നിർവ്വഹിക്കാൻ ഗുണനഫല ആകലനരീതി ഉപയോഗപ്പെടുത്താം. y മൂല്യങ്ങളുടെ തുകയുടെ ഈ വിധത്തിലുള്ള ആകലത്തെ Y_p കൊണ്ടു സൂചിപ്പിക്കുന്നു എന്നിരിക്കട്ടെ.

$$Y_p = \frac{\hat{Y}\hat{X}}{X} \qquad 6.8.1$$

ഇവിടെ $\hat{Y}$, $\hat{X}$ എന്നിവ Y,X എന്നിവയുടെ അനഭിനത ആകലങ്ങളും X, ഏതെങ്കിലും തരത്തിൽ കൃത്യമായി നിർണയിക്കുന്ന x മൂല്യങ്ങളുടെ തുകയുമത്രെ.

$$\hat{Y} = Y + \hat{Y} - Y = Y\,1 + \frac{\hat{Y}-Y}{Y}$$

$$\hat{X} = X + \hat{X} - X = X\,1 + \frac{\hat{X}-X}{X}$$

$$\hat{Y}\hat{X} = Y\,1 + \frac{\hat{Y}-Y}{Y}\,X\,1 + \frac{\hat{X}-X}{X}$$

$$= YX\left(1 + \frac{\hat{Y}-Y}{Y} + \frac{\hat{X}-X}{X} + \frac{(\hat{Y}-Y)(\hat{X}-X)}{XY}\right)$$

$$Y_p = Y\left(1 + \frac{\hat{Y}-Y}{Y} + \frac{\hat{X}-X}{X} + \frac{(\hat{Y}-Y)(\hat{X}-X)}{YX}\right)$$

ഇതിൽ നിന്നു്,

$$E(Y_P) = Y + \frac{1}{X}\,E(\hat{Y}-Y)(\hat{X}-X)$$

$$= Y + \frac{1}{X}\,\mathrm{Cov}(\hat{Y}\hat{X})$$ എന്നു സിദ്ധിക്കുന്നു.

അതായതു്, Y_p യുടെ അഭിനതിയായ $b(Y_p)$,

$$b(Y_P) = E(Y_P) - Y = \frac{1}{X}\,\mathrm{Cov}(\hat{Y}\hat{X}) \qquad 6.8.2$$

ഗുണനഫല ആകലത്തിന്റെ അഭിനതി മിക്കവാറും എല്ലാ സാമ്പിളനരീതികളിലും സാമ്പിൾ പരിമാണമായ n വർധിക്കുന്നതനുസരിച്ച് കുറയുമെന്നു് കാണിക്കാൻ കഴിയും.

അടുത്തതായി ഇതിന്റെ പ്രസരണമെന്താണെന്നു പരിശോധിക്കാം. ഒരു ആദ്യ ഏകദേശനമെന്ന നിലയിൽ $\frac{(\hat{Y}-Y)}{Y}\;\frac{(\hat{X}-X)}{X}$ എന്നതു് അവഗണിക്കുകയാണെങ്കിൽ, Y_p ഒരു അനഭിനത ആകലമാണെന്നു കാണാം. ആ സങ്കല്പത്തിൽ ചെയ്താൽ,

$$V(Y_P) = E(Y_P - Y)^2 E \approx \left\{ Y \frac{(\hat{Y} - Y}{Y} + \frac{\hat{X} - X)}{X} \right\}^2$$

$$= E \left\{ (\hat{Y} - Y) + \frac{Y}{X} (\hat{X} - X) \right\}^2$$

$$\approx E \left\{ (\hat{Y} - Y)^2 + \frac{y^2}{X^2} (\hat{X} - X)^2 + \frac{2Y}{X} (\hat{Y} - Y)(\hat{X} - X) \right\}$$

$$\approx V(\hat{Y}) + R^2 V(\hat{X}) + {}^2R \operatorname{Cov}(\hat{Y}, \hat{X}) \qquad 6.8.3$$

Y_p, $\hat{Y}$ നെക്കാൾ മെച്ചമായ ഒരു ആകലമാണെങ്കിൽ, $V(\hat{Y}) > (Y_p)$ ആയിരിക്കണം. അതായതു്, $V(\hat{Y}) > V(\hat{Y}) + R^2 V(\hat{X}) + {}^2R \operatorname{Cov}(\hat{Y} \hat{X})$

അതായതു്, $R^2 V(\hat{X}) + {}^2R \operatorname{Cov}(\hat{Y}\hat{X}) < O$

R ഒരു ധന സംഖ്യയാണെങ്കിൽ ഇതു്, $R\, V(\hat{X}) + {}^2 \operatorname{Cv}(\hat{Y}\hat{X}) < O$

അതായതു് $\frac{Y}{X} V(\hat{X}) + {}^2\rho(\hat{X}\hat{Y}) \sqrt{V(\hat{X}) V(\hat{Y})} < O$

അതു കൊണ്ടു്, $\rho(\hat{X}\hat{Y}) < -\frac{1}{2} \frac{\sqrt{\frac{V(\hat{X})}{X}}}{\sqrt{\frac{V(\hat{Y})}{Y}}} < -\frac{1}{2} \frac{\frac{\sigma_{\hat{X}}}{\bar{X}}}{\frac{\sigma_{\hat{Y}}}{\bar{Y}}}$

$$< -\frac{1}{2} \frac{C(\hat{X})}{C(\hat{Y})}$$

R ഒരു ഋണസംഖ്യയാണെങ്കിൽ ഇതു്,

$$\rho(\hat{X}\hat{Y}) > \frac{1}{2} \frac{\sqrt{\frac{V(\hat{X})}{X}}}{\sqrt{\frac{V(\hat{Y})}{-Y}}}$$

$$> \frac{1}{2} \frac{C(\hat{X})}{C(\hat{Y})} \qquad 6.8.5$$

എന്നായി തീരുമെന്നു കാണുവാൻ വിഷമമില്ല. ഈ നിബന്ധനകൾ ശരിയാവുന്ന സന്ദർഭങ്ങളിൽ മാത്രമേ ഗുണനഫലആകലനം കൂടുതൽ മെച്ചപ്പെട്ട ആകലങ്ങൾ നൽകുകയുള്ളൂ. 6.3 9, 6.3.10 എന്നീ നിബന്ധനകൾ അനുസരിക്കപ്പെട്ടുമ്പോൾ അനുപാതാകലമായിരിക്കും മെച്ചമെന്നു നാം കണ്ടു കഴിഞ്ഞു. ഈ രണ്ടു തരം നിബന്ധനകളും അനുസരിക്കപ്പെടാത്ത സാഹചര്യങ്ങളിൽ $\hat{Y}$ തന്നെ Y യുടെ ആകലമായി സ്വീകരിക്കുന്നതാണ് മെച്ചം. മറെറാരു തരത്തിൽ പറഞ്ഞാൽ R ഒരു ധനസംഖ്യയാണെങ്കിൽ, അനുപാതാകലം ഉപയോഗിക്കേണ്ടതു് $\hat{X}, \hat{Y}$ എന്നിവ തമ്മിലുള്ള സഹബന്ധം $\frac{1}{2}\, C(\hat{X})/C\,(\hat{Y})$ യെക്കാൾ വലുതായിരിക്കുമ്പോഴും, ഗുണനഫലആകലം ഉപയോഗിക്കേണ്ടതു് അതു് $-\frac{1}{2}\, C(\hat{X})/C(\hat{Y})$ യെക്കാൾകുറവായിരിക്കുമ്പോഴുമാണ്. മററു സന്ദർഭങ്ങളിൽ അനുപാതാകലം ഉപയോഗിക്കുന്നതു് സൂക്ഷ്മത കുറയുവാൻ മാത്രമേ ഇടയാക്കൂ. R ഋണസംഖ്യയായിരിക്കുമ്പോൾ നിബന്ധന ഇതിനു് നേരേ വിപരീതവുമത്രെ.

9. സമാശ്രയണാകലനം

y എന്ന അഭിലക്ഷണമാണ് പഠന വിധേയമാക്കിയിരിക്കുന്നതു്. അതിൻെറ മൂല്യങ്ങളുടെ തുകയായ Y യാണ് ആകലനം ചെയ്യേണ്ട പ്രാചലം. ഇതിനു് രണ്ടു മാർഗങ്ങൾ ഇതുവരെയായി നിർദേശിക്കപ്പെട്ടു. ഒന്നു്, സാമ്പിളിലെ y മൂല്യങ്ങളുടെ സമാന്തരമാധ്യത്തെ സമഷ്ടിയിലെ അംഗസംഖ്യയായ N കൊണ്ടു ഗുണിച്ചുകിട്ടുന്ന $\hat{Y}(=N\bar{y})$. മറെറാന്നു് y യുമായി ശക്തമായ നേർ സഹബന്ധമുള്ള x എന്നൊരഭിലക്ഷണം പരിഗണിച്ച് അതിൻെറ മൂല്യങ്ങൾ ഉപയോഗപ്പെടുത്തി കണ്ടുപിടിക്കുന്ന Y_R എന്ന അനുപാതാകലം. X എന്നതു് x മൂല്യങ്ങളുടെ തുകയും $\hat{X}$ അതിൻെറ അനഭിനത ആകലവുമെങ്കിൽ,

$Y_R = \frac{\hat{Y}}{\hat{X}} X$ എന്നാണ് നിർവചിക്കപ്പെട്ടതു്. ഇവിടെ ആദ്യത്തെ ആകലമായ $\hat{Y}$ നെ $\frac{X}{\hat{X}}$ കൊണ്ടു ഗുണിച്ചാണ് Y_R കണ്ടുപിടിച്ചതു്. അനുപാതാക

ലം ഉപയോഗപ്രദമാവണമെങ്കിൽ y യുടെ x ന്മേലുള്ള സമാശ്രയണം രേഖീയമായിരിക്കണമെന്നും ആ രേഖ, മൂലബിന്ദുവിലൂടെ കടന്നുപോകണമെന്നും നാം കണ്ടു. ഈ നിബന്ധന അല്പം ബുദ്ധിമുട്ടുണ്ടാക്കുന്ന ഒന്നാണെന്നു പറയാതെ തരമില്ല. y യുമായി ശക്തിയായ നേർസഹബന്ധമുള്ള x എന്ന ഒരു അഭിലക്ഷണം കണ്ടെത്തി എന്നിരിക്കട്ടെ. y ക്ക് x ന്മേലുള്ള സമാശ്രയണം രേഖീയമാണെന്നും സങ്കല്പിക്കുക. സ്വാഭാവികമായും x–മൂല്യങ്ങൾ Y യുടെ ആകലനത്തിനു് ഉപയോഗപ്പെട്ടുമെന്നു് നാം പ്രതിക്ഷിക്കുന്നു. പക്ഷേ ആ സമാശ്രയണരേഖ മൂലബിന്ദുവിൽകൂടി കടന്നു പോകണമെന്നില്ല. അങ്ങനെ വന്നാൽ x–നെപ്പററിയുള്ള അറിവു് പ്രയോജനപ്പെടുത്താൻ നിവൃത്തിയില്ലാത്ത ഒരു ചുറ്റുപാടിലാണു് നാം. ഈ സാഹചര്യത്തിൽ ലാഭകരമായി ഉപയോഗപ്പെടുത്താവുന്ന ഒരു ആകലനരീതിയാണു് സമാശ്രയണ ആകലനം. സമാശ്രയണരേഖ മൂലബിന്ദുവിൽ കൂടി കടന്നു പോകുമ്പോൾ ഈ ആകലവും അനുപാതാകലവും ഒന്നുതന്നെയായിരിക്കും. ചുരുക്കത്തിൽ അനുപാതാകലനത്തിൻെറ ഒരു സാമാന്യവൽക്കരണമാണു് ഈ ആകലനരീതി.

Y_r, Y യുടെ സമാശ്രയണാകലത്തെയും 'b', $\hat{Y}$ യുടെ $\hat{X}$ ന്മേലുള്ള സമാശ്രയണഗുണാങ്കത്തെയും സൂചിപ്പിക്കുന്നു എന്നിരിക്കട്ടെ.

$$Y_r = \hat{Y} + \hat{b}\,(X - \hat{X}) \qquad 6.9.1$$

എന്നു് നിർവചിക്കപ്പെട്ടിരിക്കുന്നു. നിമ്നതമപിശകുവർഗ സമ്പ്രദായമനുസരിച്ചു് b യുടെ മൂല്യം

$$b = \frac{\mathrm{Cov}\,(\hat{X}\hat{Y})}{V(\hat{X})}$$

ആണു്. ഇതു് അറിയാമെങ്കിൽ 6.9.1 ൽ ആ മൂല്യം പ്രതിസ്ഥാപിക്കാം. പക്ഷേ മിക്കവാറും സന്ദർഭങ്ങളിൽ ഇതു് ആകലനം ചെയ്യേണ്ടതകായിവരും. അതാണു് 6.9.1–ൽ '$\hat{b}$' ഉപയോഗിച്ചിരിക്കുന്നതു്.

6.10 സമാശ്രയണാകലത്തിൻെറ അഭിനതിയും പ്രസരണവും

സമാശ്രയണ ആകലം, അനുപാതാകലത്തെപ്പോലെ തന്നെ ഒരു അഭിനത ആകലമാണു്.

$$\Delta\hat{X} = \frac{\hat{X} - X}{X}, \quad \Delta\hat{Y} = \frac{\hat{Y} - Y}{Y}, \quad \Delta\hat{b} = \frac{\hat{b} - b}{b}$$

എന്നു കരുതുക.

$$\hat{X} = X+\hat{X}-X = X(1+\triangle\hat{X})$$

$$\hat{Y} = Y+\hat{Y}-Y = Y(1+\triangle\hat{Y})$$

$$\hat{b} = b+\hat{b}-b = b(1+\triangle\hat{b})$$

അതുകൊണ്ട്,

$$Y_r = Y(1+\triangle\hat{Y}) + b(1+\triangle\hat{b})(-X\triangle\hat{X})$$

$$= Y+Y\triangle\hat{Y} = bX\triangle\hat{X} - bX\triangle\hat{b}\triangle\hat{X}$$

$\hat{X}$, $\hat{Y}$ എന്നിവ x, y എന്നിവയുടെ അനഭിനത ആകലങ്ങളായതുകൊണ്ട്,

$$E(\triangle\hat{Y}) = E(\triangle\hat{X}) = O$$

അതുകൊണ്ട്,

$$E(Y_r) = Y - bXE(\triangle\hat{b}\cdot\triangle\hat{X})$$

$$= Y - bX\,Cov(\hat{b}, \hat{X})$$

Y_r ന്റെ അഭിനതി $b(Y_r)$ കൊണ്ട കുറിച്ചാൽ,

$$b(Y_r) = E(Y_r) - Y = -bX\,Cov(\hat{b}\hat{X}) \qquad 6.9.2$$

സാമ്പിൾപരിമാണം വലുതായിരിക്കുമ്പോൾ $(\triangle\hat{b}\cdot\triangle\hat{X})$ എന്ന പദം അവഗണിക്കാവുന്നതാണ്. അങ്ങനെയെങ്കിൽ Y_R ഒരു അനഭിനത ആകലമായിരിക്കും. ഈ സങ്കല്പം അംഗീകരിച്ചാൽ,

$$V(Y_r) = E(Y_r - Y)^2 \approx E(Y\triangle\hat{Y} - bX\triangle\hat{X})^2$$

$$\approx Y^2E(\triangle\hat{Y})^2 + b^2X^2\,E(\triangle\hat{X})^2 - 2bXY\,E(\triangle\hat{Y}\triangle\hat{X})$$

$$\approx Y^2\frac{V(\hat{Y})}{Y^2} + b^2X^2\frac{V(\hat{X})}{X^2} - 2bXY\frac{Cov(\hat{X}\hat{Y})}{XY}$$

$$\approx V(\hat{Y}) + b^2V(\hat{X}) - 2b\,Cov(\hat{X}\hat{Y}) \qquad 6.9.3$$

6.3.8 ൽ കൊടുത്തിരിക്കുന്ന അനുപാതാകലമായ Y_R ന്റെ പ്രസരണവുമായി ഇതു താരതമ്യപ്പെടുത്തുക.

$$V(Y_R) = V(\hat{Y}) + R^2V(\hat{X}) - 2R\,Cov(\hat{X}\hat{Y})$$

$$V(Y_R) - V(Y_r) = V(\hat{X})\,[R^2 - b^2] - 2\,Cov(\hat{X}\hat{Y})\,[R - b]$$

ഇവിടെ $b = \dfrac{Cov(\hat{X}\hat{Y})}{V(\hat{X})} = \dfrac{\rho\sqrt{V(\hat{Y})}}{\sqrt{V(\hat{X})}}$

എന്നതു് പ്രതിസ്ഥാപിച്ചു് ലഘൂകരിച്ചാൽ,

$$V(Y_R) - V(Y_r) = V(\hat{X})\left[R^2 - \frac{\rho^2V(\hat{Y})}{V(\hat{X})}\right] - 2\rho\sqrt{V(\hat{X})\,V(\hat{Y})}$$

$$\times\left[R - \rho\frac{\sqrt{V(\hat{Y})}}{\sqrt{V(\hat{X})}}\right]$$

$$= R^2\,V(\hat{X}) - \rho^2V(\hat{Y}) - 2\rho R\sqrt{V(\hat{X})V(\hat{Y})} + 2\rho^2V(\hat{Y})$$

$$= \left[R\sqrt{V(\hat{X})} - \rho\sqrt{V(\hat{Y})}\right]^2 \qquad 6.9.4$$

ഇതിൽ നിന്നു് $V(Y_R)$ എപ്പോഴും $V(Y_r)$ നെക്കാൾ വലുതായിരിക്കുമെന്നു് സിദ്ധിക്കുന്നു. അതായതു്, സമാശ്രയണാകലത്തിനു് അനുപാതാകലത്തെക്കാൾ സൂക്ഷ്മതയുണ്ടു്. പക്ഷേ അനുപാതാകലം ഗണിച്ചെടുക്കുന്നതിനെക്കാൾ വളരെയേറെ സങ്കീർണമാണു് സമാശ്രയണാകലം ഉൾക്കൊള്ളുന്ന ഗണനം. സാമ്പിൾ പരിമാണം വലുതാകുന്തോറും ഇതു് ഗൗരവമേറിയ ഒരു പ്രായോഗികബുദ്ധിമുട്ടായി തീരും. അതുകൊണ്ടു് ചുരുക്കം സന്ദർഭങ്ങളിൽ മാത്രമെ സമാശ്രയണാകലം ഉപയോഗിക്കാറുള്ളു.

സമാശ്രയണരേഖ, മൂലബിന്ദുവിലൂടെ കടന്നു പോകുന്നു എങ്കിൽ അതിന്റെ സമീകരണം, $\hat{Y} = b\,\hat{X}$ എന്നാവും. അതായതു്,

$b = \dfrac{\hat{Y}}{\hat{X}}$. 6.9.1 ൽ b യുടെ ഈ മൂല്യം $\hat{b}$ ന്റെ സ്ഥാനത്തു് പ്രതിസ്ഥാ

പിച്ചാൽ,

$$Y_r = \hat{Y} + \frac{\hat{Y}}{\hat{X}} \quad (X - \hat{X})$$

എന്നാവും. അതായതു്,

$Y_r = \dfrac{\hat{Y}}{\hat{X}}$ ഇതു തന്നെയാണെല്ലൊ അനുപാതാകലമായ Y_R. അങ്ങനെ

സമാശ്രയണരേഖ മൂലബിന്ദുവിലൂടെ കടന്നു പോകുമ്പോൾ അനുപാതാകലവും സമാശ്രയണാകലവും ഒന്നു തന്നെയാണെന്നു വന്നുചേരുന്നു.

x-മൂല്യങ്ങൾ നിർണയിക്കുക എളുപ്പവും y-മൂല്യങ്ങളുടെ നിർണയം വിഷമകരവുമാവുമ്പോഴാണു് സാധാരണയായി സമാശ്രയണാകലങ്ങൾ ഉപയോഗിക്കാറുള്ളതു്. ഉദാഹരണമായി ഒരു വനത്തിൽ നിന്നു് ലഭിക്കാവുന്ന ആകെ തടിയുടെ വ്യാപ്തം നിർണയിക്കുകയാണു് അന്വേഷണലക്ഷ്യമെന്നിരിക്കട്ടെ. വനത്തെ $\frac{1}{4}$ ഏക്കർ വീതമുള്ള ഭാഗങ്ങളായി വിഭജിക്കുന്നു. ഈ ഓരോ ഭാഗത്തെയുമാണു് വ്യക്തികളായി പരിഗണിക്കുന്നതു്. അതിലുള്ള തടിയുടെ വ്യാപ്തമാണു് Y–അഭിലക്ഷണം. പരിചയസമ്പന്നനായ ഒരു നിരീക്ഷകനു് ഒററ നോട്ടം കൊണ്ടുതന്നെ ഓരോ ഭാഗത്തും എത്ര വ്യാപ്തം തടിയുണ്ടെന്നു മതിപ്പ് പറയാൻ കഴിയും. അങ്ങനെ ലഭിക്കുന്ന വ്യാപ്തത്തിന്റെ ഏകദേശമൂല്യമാണു് x–അഭിലക്ഷണം.

വനവിഭാഗങ്ങളുടെ ഒരു സാമ്പിളെടുത്തു് $\hat{b}$, $\hat{x}$, $\hat{y}$ എന്നിവ ആകലനം ചെയ്യാം.

$Y_r = \hat{Y} + \hat{b}\,(\hat{X} - \hat{X})$ എന്ന സമീകരണമുപയോഗിച്ചു് ആകെ തടിയുടെ വ്യാപ്തം ആകലനം ചെയ്യാം. ഇവിടെ X, വനത്തിലെ ആകെ തടിയുടെ ഒററ നോട്ടം കൊണ്ടു് മതിച്ച വ്യാപ്തമാണു്. ഇതു് താരതമ്യേന എളുപ്പം ചെയ്യാവുന്നതായതുകൊണ്ടാണു് ഈ ആകലനസമ്പ്രദായം സ്വീകരിക്കുന്നതു്.

6.9.3–ൽ നൽകിയിരിക്കുന്ന $V(Y_r)$ ന്റെ വ്യംജകത്തെ അല്പംകൂടി ഗാഢമായി പരിശോധിച്ചുനോക്കുക.

$$V(Y_r) = V(\hat{Y}) + b^2 V(\hat{X}) - 2b\, Cov(\hat{X}\hat{Y})$$

$$= V(\hat{Y}) + \rho^2 \frac{V(\hat{Y})}{V(\hat{X})} V(\hat{X}) - 2\rho \sqrt{V(\hat{X})\ V(\hat{Y})} \cdot \frac{\rho \sqrt{V(\hat{Y})}}{\sqrt{V(\hat{X})}}$$

$$= V(\hat{Y})\ [1-\rho^2] \qquad 6.9.5$$

ഇതിൽ നിന്നു് ρ അതായതു് $\hat{X}, \hat{Y}$ എന്നിവ തമ്മിലുള്ള സഹബന്ധഗുണാങ്കം, എത്ര വലുതായിരിക്കുന്നോ അത്രക്കു് ചെറുതായിരിക്കും $V(\hat{Y}_r)$ എന്നു വന്നു ചേരുന്നു. അതായതു് സമാശ്രയണാകലത്തിൻെറ സൂക്ഷ്മത, $\hat{X}, \hat{Y}$ എന്നിവ തമ്മിലുള്ള സഹബന്ധം വർധിക്കുന്നതനുസരിച്ച് വർധിക്കുന്നു.

ഇതുപയോഗിച്ചു് $V(\hat{Y}_r)$ ആകലനം ചെയ്യാൻ എളുപ്പമാണു്. ഇതിൽ സാമ്പിൾ മൂല്യങ്ങൾ പ്രതിസ്ഥാപിച്ചാൽ,

$$\hat{V}(Y_r) = N \frac{(N-n)}{n}\ S^2_y\ (1-r^2) \qquad 6.9.6$$

6.10 സമാശ്രയണാകലനം ലഘുയാദൃച്ഛിക സാമ്പിളനത്തിൽ

അനുപാതാകലത്തിലെന്നതുപോലെ സമാശ്രണാകലനത്തിലും ആകലത്തിന്റെ അഭിനതിയും പ്രസരണവും നൽകുന്ന വ്യംജകങ്ങൾക്കു് സാമ്പിളന രീതിയനുസരിച്ചു് ചില മാററങ്ങൾ വരും. ഉദാഹരണമായി പുനസ്ഥാപനമില്ലാത്ത ലഘുയാദൃച്ഛിക സാമ്പിളനത്തിലെ വ്യംജകങ്ങൾ ഇവിടെ വ്യുൽപ്പാദിപ്പിക്കുന്നു.

സമഷ്ടിയിലെ അംഗങ്ങളുടെ എണ്ണം N ആണെന്നും n അംഗങ്ങളുള്ള സാമ്പിളാണു് എടുക്കുന്നതെന്നും ഇരിക്കട്ടെ. $\bar{x}, \bar{y}$ എന്നിവയാണു് x, y അഭിലക്ഷണ മൂല്യങ്ങളുടെ സാമ്പിൾ സമാന്തരമാധ്യങ്ങളെന്നു വിചാരിക്കുക. സാമ്പിളിലെ i-ാമതു് അംഗത്തിൻെറ x, y അഭിലക്ഷണമൂല്യങ്ങൾ യഥാക്രമം x_i, y_i എന്നിവയാണെന്നും സങ്കല്പിക്കുക.

$$\hat{X} = N\bar{x},\ \hat{Y} = N\bar{y}$$

$$\hat{b} = \frac{\frac{1}{n}\sum_{i=1}^{n}(x_i-\bar{x})(y_i-\bar{y})}{\frac{1}{n}\sum_{i=1}^{n}(x_i-\bar{x})^2}$$

ഈ മൂല്യങ്ങൾ പ്രതിസ്ഥാപിക്കുമ്പോൾ

$$Y_r = N\bar{y} + \hat{b}\,(X - N\bar{x})$$

$$= N\,[\bar{y} + \hat{b}\,(\bar{X} - \bar{x})] \qquad 6.10.1$$

$$V(\hat{Y}) = N^2\,V(\bar{y}) = N^2\,\frac{N-n}{Nn}\,S^2_y$$

ഈ മൂല്യം $V(Y_r)$-ന്റെ 6.9.5-ലെ ലഘൂകരിച്ച വ്യംജകത്തിൽ പ്രതിസ്ഥാപിച്ചാൽ,

$$V(Y_r) = \frac{N^2\,(N-n)}{Nn}\,S^2_y\,(1-\rho^2) = \frac{N\,(N-n)}{n}\,S^2_y\,(1-\rho^2)$$

$$= \frac{N\,(N-n)}{n}\,S'^2_y \qquad 6.10.2$$

Y_r ആദ്യ ഏകദേശനമെന്ന നിലയിൽ അനഭിനതമെന്നു സങ്കൽപിക്കുക. സാമ്പിൾ വലുതാണെങ്കിൽ 6.4 ൽ സ്വീകരിച്ചതിനു് സമാനമായ വ്യുൽപ്പാദന രീതി അവലംബിച്ചാൽ,

$$\hat{S}'^2_y = \frac{1}{n-1}\sum_{i=1}^{n}\left\{(y_i-\bar{y}) - \hat{b}\,(x_i-\bar{x})\right\}^2$$

എന്നു കാണിക്കുവാൻ കഴിയും. അതുകൊണ്ടു്, $V(Y_r)$ ന്റെ ആകലം,

$$\hat{V}(Y_r) = \frac{N\,(N-n)}{n\,(n-1)}\sum_{i=1}^{n}\left\{(y_i-\bar{y}) - \hat{b}\,(x_i-\bar{x})\right\}^2 \qquad 6.10.3$$

ഇതുപോലെതന്നെ മററു് സാമ്പിളന രീതികളിലും സമാശ്രയണാകലത്തിന്റെ അഭിനതിയും പ്രസരണവും നൽകുന്ന വ്യംജകങ്ങൾ വ്യുൽപ്പാദിപ്പിക്കാവുന്നതാണു്.

11. ബഹുചര സമാശ്രയണ ആകലം

y എന്ന അഭിലക്ഷണവുമായി ശക്തിയായ നേർസഹബന്ധമുള്ള $x_1, x_2 .. x_k$ എന്നു് k അഭിലക്ഷണങ്ങൾ ഉണ്ടെന്നിരിക്കട്ടെ. ഇവയെപ്പററിയുള്ള വിവര

ങ്ങൾകൂടി ശേഖരിച്ചാൽ Y യുടെ ആകലം കുറച്ചുകൂടി സൂക്ഷ്മതയുള്ളതാക്കാൻ സാധിക്കും.

$$Y_r = \hat{Y} + \sum_{j=1}^{k} \hat{b}_j (X_j - \hat{X}_j) \qquad 6.11.1$$

എന്ന സമീകരണമുപയോഗിച്ച് Y_r നിർണയിക്കാവുന്നതാണ്. ഇവിടെ

$_j$, x_j അഭിലക്ഷണത്തിന്റെ മൂല്യങ്ങളുടെ തുകയും $\hat{b}r$, $\hat{Y}$ യുടെ X_j $(j=1,2\ldots k$ യിലുള്ള സമാശ്രയണ ഗുണാങ്കത്തിന്റെ അനഭിനത ആകലവുമത്രെ. ഒരു സഹായക അഭിലക്ഷണം മാത്രമുണ്ടായിരുന്നപ്പോൾ വ്യുൽപ്പാദിപ്പിച്ച രീതിയിൽ തന്നെ ഇവിടെയും ആകലത്തിന്റെ പ്രസരണം കണ്ടുപിടിക്കാവുന്നതത്രെ.

12. സംഗ്രഹം

(1) സമഷ്ടിയോടു ബന്ധപ്പെട്ട ഒന്നോ അതിലധികമോ അഭിലക്ഷണങ്ങളെപ്പറ്റി പഠനം നടത്തുവാനാണ് സാമ്പിളനം ഉദ്ദേശിക്കപ്പെട്ടിട്ടുള്ളത്. പ്രാചലങ്ങളുടെ ആകലനം ഈ പഠനത്തിന്റെ ഒരു സുപ്രധാനഭാഗമാണ്. പരിഗണിക്കപ്പെട്ടുന്ന അഭിലക്ഷണത്തോട് നേർസഹബന്ധമുള്ള മറ്റൊരഭിലക്ഷണത്തെപ്പററി സാമ്പിളിൽ നിന്നു ശേഖരിച്ചതും മററുതരത്തിൽ ലഭിച്ചതുമായ വിവരങ്ങൾ ഉപയോഗപ്പെടുത്തി കൂടുതൽ സൂക്ഷ്മതയോടെ ആകലനം നിർവഹിക്കുവാനുള്ള രണ്ടു മാർഗങ്ങളാണ് അനുപാതാകലനവും സമാശ്രയണാകലനവും.

(2) y പരിഗണിക്കപ്പെടുന്ന അഭിലക്ഷണവും x അതുമായി നേർസഹബന്ധമുള്ള ഒരു സഹായകഅഭിലക്ഷണവും ആണെന്നിരിക്കട്ടെ. Y, X എന്നിവ യഥാക്രമം ഈ അഭിലക്ഷണമൂല്യങ്ങളുടെ തുകകളും, $R = \frac{Y}{X}$ ഉം $\hat{Y}$, $\hat{X}$ എന്നിവ

Y, X എന്നിവയുടെ അനഭിനത ആകലങ്ങളും, $\hat{R} = \hat{Y}/\hat{X}$ ഉം ആണെന്നു

സങ്കൽപിക്കുക. X-ന്റെ മൂല്യം മററു തരത്തിൽ അറിയാമെങ്കിൽ, $Y_R = \hat{R}\,X$ എന്ന സമീകരണം കൊണ്ട് നിർവചിച്ചിരിക്കുന്ന Y_R, Y-യുടെ അനുപാതാകലം എന്നു വ്യവഹരിക്കപ്പെടുന്നു. $\bar{y}$, $\bar{x}$ എന്നിവ ഈ അഭിലക്ഷണമൂല്യങ്ങളുടെ സാമ്പിളിലെ സമാന്തരമാധ്യങ്ങളാണെങ്കിൽ,

$$\hat{R} = \frac{\bar{y}}{\bar{x}}.$$

അതുകൊണ്ട് $Y_R = \frac{\bar{y}}{\bar{x}} X.$

(3) Y_R, Y യുടെ അനഭിനത ആകലമല്ല. കാരണം, $\hat{R}$, R ന്റെ അനഭിനത ആകലമല്ല എന്നതാണ്.

$$E(\hat{R}) = \frac{Y}{X} + \frac{Y}{X}\left[\frac{V(\hat{X})}{X^2} - \frac{Cov(\hat{X}\hat{Y})}{XY}\right] + \cdots\cdots$$

$$= R + R\left[\frac{V(\hat{X})}{X^2} - \frac{Cov(\hat{X}\hat{Y})}{XY}\right] +$$

ആദ്യഏകദേശനം $E(\hat{R}) = R$ ആണെന്നു സ്വീകരിച്ചാൽ $\hat{Y}$ നെ Y യുടെ അനഭിനത ആകലമായി അംഗീകരിക്കാം. രണ്ടാം ഏകദേശനമായി,

$$E(\hat{R}) = R + R\left[\frac{V(\hat{X})}{X^2} - \frac{Cov(\hat{X}\hat{R})}{XY}\right]$$

എന്നു സ്വീകരിക്കാം. അങ്ങനെയെങ്കിൽ $\hat{R}$ ന്റെ അഭിനതിയായ

$$b(\hat{R}) = R\left[\frac{V(\hat{X})}{X^2} - \frac{Cov(\hat{X}\hat{Y})}{XY}\right] = \frac{R}{X^2}\left[V(\hat{X}) - \frac{Cov(\hat{X}\hat{Y})}{R}\right]$$

ആണ്. $\hat{Y}$ ന് $\hat{X}$ ന്റെ മേലുള്ള സമാശ്രയണരേഖ മൂലബിന്ദുവിലൂടെ കടന്നു പോവുന്നെങ്കിൽ ഈ അഭിനതി '0' ആകും. അതായതു് $\hat{R}$, R ന്റെ അനഭിനത ആകലമാവും. $\hat{Y}$, Y യുടെ അനഭിനത ആകലവുമായിത്തീരും.

$$\hat{R} \text{ ന്റെ സാമ്പിളനപ്രസരണം } = R^2\left\{\frac{V(\hat{Y})}{Y^2} + \frac{V(\hat{X})}{X^2} - \frac{2\,Cov(\hat{X}\hat{Y})}{XY}\right\}$$

(ഏകദേശനം)

ഈ ഫലങ്ങൾ ഉപയോഗപ്പെടുത്തിയാൽ,

$$b(Y_R) = \frac{1}{X}\left[RV(\hat{X}) - Cov(\hat{X}\hat{Y})\right]$$

എന്നും,

$$V(Y_R) = V(\hat{Y}) + R^2\,V(\hat{X}) - 2R\,Cov(\hat{X}\hat{Y})$$

എന്നും സിദ്ധിക്കും.

(4) R ഒരു ധനസംഖ്യയാണെങ്കിൽ,

$$\rho(\hat{X}\hat{Y}) > \frac{1}{2}\frac{\sqrt{\dfrac{V(\hat{X})}{X}}}{\sqrt{\dfrac{V(\hat{Y})}{Y}}}$$

ആയിരിക്കുമ്പോഴും, R ഋണസംഖ്യയാണെങ്കിൽ,

$$\rho(\hat{X}\hat{Y}) < -\frac{1}{2}\frac{\sqrt{\dfrac{V(\hat{X})}{X}}}{\sqrt{\dfrac{V(\hat{Y})}{Y}}}$$

ആയിരിക്കുമ്പോഴും അനുപാതാകലമായ Y_R, $\hat{Y}$ എന്ന അനഭിനത ആകലത്തെക്കാൾ സൂക്ഷ്മതയുള്ളതായിരിക്കും.

(5) ലഘുയാദൃച്ഛികസാമ്പിളന രീതിയനുസരിച്ചാണ് സാമ്പിളെടുത്തിരിക്കുന്നതെങ്കിൽ.

$$b(Y_R) = \frac{N-n}{n\bar{X}}[RS^2_x - \rho S_x S_y]$$

$$b(Y_R) = \frac{N(N-n)}{n}[S^2_y + R^2S^2_x - 2\rho RS_xS_y]$$

ഇവയുടെ ആകലങ്ങൾ,

$$\hat{b}(Y_R) = \frac{N-n}{n}\bar{y}\left[\frac{S^2_x}{\bar{X}^2} - \frac{r\,S_x\,S_y}{\bar{x}\bar{y}}\right]$$

$$\hat{V}(Y_R) = \frac{N(N-n)}{n}\left[S^2_y + \frac{\bar{y}^2}{\bar{x}^2}S^2_x - 2r\frac{\bar{y}}{\bar{x}}S_xS_y\right]$$

(6) $\bar{y}$ ന്റെ അനുപാതാകലം $\bar{Y}_R$ ആണെങ്കിൽ $\bar{Y}_R = \dfrac{Y_R}{N}$.

$$b(\bar{Y}_R) = \frac{N-n}{Nn\bar{X}}\left[RS^2_x - \rho S_xS_y\right]$$

$$V(\bar{Y}_R) = \frac{N-n}{Nn}\left[S^2_y + R^2 S^2_x - 2\rho R\; S_x S_y\right]$$

$$\hat{b}(\bar{Y}_R) = \frac{\bar{y}}{n}\left[\frac{s^2_x}{\bar{x}^2} - \frac{r\, s_x s_y}{\bar{x}\,\bar{y}}\right]$$

$$\hat{V}(\bar{Y}_R) = \frac{N-n}{Nn}\left\{s^2_y + \frac{\bar{y}^2}{\bar{x}^2} s^2_x - \frac{2\,r\,\bar{y}}{\bar{x}} s_x s_y\right\}$$

(7) സ്തരിതസാമ്പിളനം അനുസരിച്ചാണ് സാമ്പിളെടുക്കുന്നതെങ്കിൽ Y യെ ആകലനം ചെയ്യാൻ രണ്ടു രീതികൾ സാധാരണ ഉപയോഗിക്കാറുണ്ട്. വിഭിന്ന അനുപാതരീതി, സംയുക്ത അനുപാതരീതി എന്നിങ്ങനെ അവ വ്യവഹരിക്കപ്പെടുന്നു. വിഭിന്ന അനുപാതരീതിയിൽ,

$$Y = \sum_{j=1}^{k} \frac{\bar{y}_j}{\bar{x}_j} X_j$$

$y_j, \bar{x}_j$ എന്നിവ j-ാമതു് സ്തരത്തിൽ നിന്നു് സാമ്പിളിൽ ഉൾപ്പെടുത്തിയ വ്യക്തികളുടെ y, x അഭിലക്ഷണമൂല്യങ്ങളുടെ സമാന്തരമാധ്യങ്ങളും X_j, ആ സ്തരത്തിലെ x മൂല്യങ്ങളുടെ തുകയുമാണ്).

$$b(Y_R) = \sum_{j=1}^{k} \frac{N_j - n_j}{n_j \bar{X}_j}\left[R_j S^2_{xj} - \rho_j S_{xj} S_{yj}\right]$$

[ഇവിടെ S_{xj}, S_{yj}, ρ_j എന്നിവ യഥാക്രമം j-ാമതു് സ്തരത്തിൽ നിന്നു് സാമ്പിളിൽ ഉൾപ്പെടുത്തിയ വ്യക്തികളുടെ x,y അഭിലക്ഷണ മൂല്യങ്ങളുടെ മാനകവിചലനങ്ങളും ρ_j, ആ സ്തരത്തിലെ അംഗങ്ങളുടെ x, y അഭിലക്ഷണമൂല്യങ്ങൾ തമ്മിലുള്ള സഹബന്ധഗുണാങ്കവുമാണ്.]

$$V(Y_R) = \sum_{j=1}^{k} \frac{N_j(N_j - n_j)}{n_j}\left[S^2_{yj} + R^2_j S^2_{xj} - 2\rho_j R_j S_{xj} S_{yj}\right]$$

$$\left[\text{ഇവിടെ } R_j = \frac{Y_j}{X_j}\right]$$

സംയുക്തഅനുപാതസമ്പ്രദായത്തിൽ,

$$Y_R = \frac{\bar{y}_{st}}{\bar{x}_{st}} X$$

[ഇവിടെ $\bar{y}_{st}$, $\bar{x}_{st}$ എന്നിവ $\bar{X}$, $\bar{Y}$ എന്നിവയുടെ സ്തരിതസാമ്പിളിൽനിന്നുള്ള ആകലങ്ങൾ]

$$b(Y_R) = \frac{1}{\bar{\bar{X}}} \sum_{j=1}^{k} \frac{N_j - n_j}{n_j} [RS^2_{xj} - \rho S_{xj} S_{yj}]$$

$$V(Y_R) = \sum_{j=1}^{k} \frac{N_j (N_j - n_j)}{n_j} [S^2_{yj} + R^2 S_{xj} - 2\rho_j R S_{yj} S_{xj}]$$

8. ക്രമാനുസൃത സാമ്പിളനത്തിലും $Y_R = \hat{R} X$. $\bar{x}_r$, $\bar{y}_r$ എന്നിവയാണ് r യാദൃച്ഛിക പ്രാരംഭമായുള്ള സാമ്പിളിന്റെ x, y അഭിലക്ഷണ മൂല്യങ്ങളുടെ സമാന്തരമാധ്യങ്ങളെങ്കിൽ,

$$b(Y_R) = \frac{N}{\bar{X}} \cdot \frac{1}{k} \sum_{r=1}^{k} \bar{x}_r (R\bar{x}_r - \bar{y}_r)$$

$$V(Y_R) = \frac{N^2}{k} \sum_{r=1}^{k} (\bar{y}_r - R\bar{x}_r)^2$$

9. സാമ്പിളിലെ ആദ്യത്തെ അംഗത്തെ അതിന്റെ x മൂല്യത്തിനു ആനുപാതികമായ സംഭാവ്യതയോടു കൂടിയും ബാക്കിയുള്ള n—1 അംഗങ്ങളെ പ്രതിസ്ഥാപനമില്ലാത്ത ലഘുയാദൃച്ഛിക സാമ്പിളനരീതിയിലും തിരഞ്ഞെടുത്താൽ ആ സാമ്പിളിൽ നിന്നു ലഭിക്കുന്ന അനുപാതാകലം അനഭിനതമായിരിക്കും. സാമ്പിളനരീതിയിൽ മാററം വരുത്താതെ ആകലനരീതിയിൽ മാററം വരുത്തിയും അനഭിനതആകലങ്ങൾ നിർദേശിക്കപ്പെട്ടിട്ടുണ്ട്.

10. ഇവിടെ ചർച്ചചെയ്തതിനു പുറമെ, ശൃംഖലിത അനുപാതാകലനം, പശ്ചാൽ സ്തരിത അനുപാതാകലനം, ബഹുസഹായകചര ആകലനം തുടങ്ങി മററു പല അനുപാതാകലന സമ്പ്രദായങ്ങളും നിലവിലുണ്ട്.

11. x അഭിലക്ഷണത്തിന് y യുമായി ശക്തിയായ വിപരീത സഹബന്ധമാണ് ഉള്ളതെങ്കിൽ ഗുണനഫലആകലനം ഉപയോഗിക്കാം. Y യുടെ ഗുണനഫല ആകലമാണ് y_p, എങ്കിൽ

$$Y_P = \frac{\hat{Y}\hat{X}}{X}$$

$$b(Y_P) = \frac{1}{X} \mathrm{Cov}(\hat{Y}\hat{X})$$

$$V(Y_P) = V(\hat{Y}) + R^2 V(\hat{X}) + 2R\,\mathrm{Cov}(\hat{Y}\hat{X})$$

12. R ധനസംഖ്യയാണെങ്കിൽ $\hat{X}$, $\hat{Y}$ എന്നിവ തമ്മിലുള്ള സഹബന്ധ

ണാങ്കം

$$\frac{1}{2}\;\frac{\sqrt{\dfrac{V(\hat{X})}{X}}}{\sqrt{\dfrac{V(\hat{Y})}{Y}}}$$

യെക്കാൾ കുറവായിരിക്കുമ്പോഴാണ് ഗുണനഫല ആകലം ഉപയോഗിക്കേണ്ടത് R ഋണമാണെങ്കിൽ മറിച്ചും.

13. സമാശ്രയണാകലനം y യുടെ സമാശ്രയണാകലം y_r ആണെങ്കിൽ

$$Y_r = \hat{Y} + \hat{b}\,(X - \hat{X})$$

$$b(Y_r) = -b \times Cov(\hat{b}, \hat{X})$$

$$V(Y_r) = V(\hat{Y}) + b^2\, V(\hat{X}) - 2b\, Cov(\hat{X}\hat{Y})$$

[ഇവിടെ 'b' y ക്ക് x മേലുള്ള സമാശ്രയണഗുണാങ്കം]

14. $V(Y_R) - V(Y_r) \leq 0$

അതുകൊണ്ട് സമാശ്രയണാകലം അനുപാതാകലത്തെക്കാൾ മെച്ചപ്പെട്ടതാണ്.

15. ലഘുയാദൃച്ഛിക സാമ്പിളനത്തിൽ

$$Y_r = N\,[\bar{y} + \hat{b}(\bar{X} - \bar{x})]$$

$$V(Y_r) = \frac{N(N-n)}{n}\, S'^2_y$$

$$\hat{V}(Y_r) = \frac{N(N-n)}{n(n-1)} \sum_{i=1}^{n} \left\{ (y_i - \bar{y}) - \hat{b}\,(x_i - \bar{x}) \right\}^2$$

15. y–അഭിലക്ഷണവുമായി ശക്തിയായ നേർസഹബന്ധമുള്ള $x_1, x_2, \dots x_k$ എന്ന k അഭിലക്ഷണങ്ങൾ ഉണ്ടെങ്കിൽ,

$$Y_r = \hat{Y} + \sum_{j=1}^{k} \hat{b}_j (X_j - \hat{X}_j)$$

13. ചില ഉദാഹരണങ്ങൾ

ഉദാഹരണം 6.1; യാദൃച്ഛികമായി തിരഞ്ഞെടുക്കപ്പെട്ട കേരളത്തിലെ 10 ഗ്രാമങ്ങളിലെ കർഷക കുടുംബങ്ങളുടെ എണ്ണവും കന്നുകാലികളുടെ എണ്ണവും താ

ഴെ കൊടുക്കുന്നു. ഒരു ഗ്രാമത്തിലെ ശരാശരി കർഷക കുടുംബങ്ങളുടെ എണ്ണം 320 ആണെങ്കിൽ കന്നുകാലികളുടെ ശരാശരി എണ്ണവും അതിന്റെ സാമ്പിളന പ്രസരണവും ആകലനം ചെയ്യുക.

ഗ്രാമം	കർഷക കുടുംബങ്ങളുടെ എണ്ണം	കന്നുകാലികളുടെ എണ്ണം
1	243	478
2	365	550
3	252	480
4	175	390
5	420	790
6	212	483
7	430	754
8	213	553
9	104	242
10	348	675

ഇവിടെ കർഷക കുടുംബങ്ങളുടെ എണ്ണം x എന്നും കന്നുകാലികളുടെ എണ്ണം y എന്നും വിചാരിക്കുക.

$$\bar{x} = 276.2$$
$$\bar{y} = 539.5$$

$$\hat{R} = \frac{539.5}{276.2} = 1.95$$

ശരാശരി കന്നുകാലികളുടെ എണ്ണം $= \hat{R} \cdot X = 1.95 \times 320 = 624$

$$\Sigma x_i^2 = 869936$$
$$\Sigma y_i^2 = 3159387$$
$$\Sigma x_i y_i = 1642387$$
$$S_x^2 = \tfrac{1}{9}\,[869936 - (276.2)^2 \cdot (10)] = 11830.2$$
$$S_y^2 = \tfrac{1}{9}\,[3159387 - (539.5)^2 (10)] = 27642.7$$
$$rS_xS_y = \tfrac{1}{9}\,[1642387 - (276.2)(539.5)(10)] = 16920.8$$

$$s^2(\bar{y}_a) = \frac{N-n}{Nn}\left\{s^2_y - 2r\,s_x s_y \frac{\bar{y}}{\bar{x}} + s^2_x \frac{\bar{y}^2}{\bar{x}^2}\right\}$$

$$\approx \frac{1}{n}\,\{s^2_y - 2rs_x s_y \hat{R} + s^2_x \hat{R}^2\}$$

(N വളരെ വലുതായതു കൊണ്ട് $\frac{n}{N}$ നെ അവഗണിക്കുന്നു)

$$= \tfrac{1}{9}\,\{27642.7 - 2 \times 16920.8 \times 1.95 + (1.95)^2 11830.2\}$$
$$= 737.32$$

ഈ ആകലത്തിന്റെ മാനക വിചലനം $\sqrt{737.32}$, അതായത് 27.1, ആ ത്രെ. ഇതിനു പകരം $\bar{y}$ ആണ് ആകലമായി സ്വീകരിച്ചിരുന്നതെങ്കിൽ അതിന്റെ സാമ്പിളന പ്രസരണം, N വലുതായതുകൊണ്ട് ഏകദേശം $\frac{S_y^2}{n} = 1692.08$ ആ രിക്കും. അനുപാതാകലത്തിന്റെ സാമ്പിളന പ്രസരണം ഇതിന്റെ പകുതിയിൽ കുറവാണെന്നുള്ളത് ശ്രദ്ധേയമത്രേ.

ഉദാഹരണം 6.2: മുൻഉദാഹരണത്തിൽ സമാശ്രയണാകലനരീതി ഉപയോഗിച്ച് കന്നുകാലികളുടെ ശരാശരി എണ്ണം ആകലനം ചെയ്യുക.

സമാശ്രയണാകലം $\bar{y}_r$ ആണെങ്കിൽ,

$$\bar{y}_r = \hat{a} + \hat{b}\bar{x}$$

ഇവിടെ,

$$\hat{a} = \bar{y} - b\,\bar{x}, \quad \hat{b} = \frac{r\, s_x s_y}{s^2_x}$$

ഉദാഹരണം 6.1 ൽ നടത്തിയ ഗണനങ്ങൾ ഉപയോഗപ്പെടുത്തിയാൽ,

$$\bar{x} = 276.2\ , \ \bar{y} = 539.5$$

$$s^2_x = 11830.2\ , \ s^2_y = 27642.7$$

$$r s_x s_y = 16920.8$$

$$\hat{b} = \frac{16920.8}{11830.2} = 1.43$$

$$\hat{a} = 539.5 - 1.43 \times 276.2 = 144.53$$

$$\bar{y}_r = 144.53 + 1.43 \times 320 = 602.13$$

സമാശ്രയണരീതിയിലുള്ള ശരാശരി കന്നുകാലികളുടെ ആകലം $= 602.13$

$$V(\bar{y}_r) \approx \frac{1}{n} s^2_y (1 - r^2) = \frac{1}{10} \times 27642.7\, (1 - .94^2)$$

സമാശ്രയണാകലത്തിന്റെ പ്രസരണം അനുപാതാകലത്തിന്റെ പ്രസരണത്തെക്കാൾ ചെറുതാണെന്നുള്ളത് ശ്രദ്ധിക്കുക.

14. ചില സഹായകഗ്രന്ഥങ്ങളും ലേഖനങ്ങളും

1. Sampling Theory and Methods (chapter 10), M. N. Murthy, Statistical Publishing Society, Calcutta.
2. Ratio Estimation in Sampling with equal and unequal probabilities (1954), Des Raj, J. Ind. Soc: Agr. Stat., 6, 127–138
3. Unbiased ratio estimation (1954), Hartley H. O. and Ross A, Nature 174, 270-271
4. A method of sample selection providing unbiased ratio estimates (1951), Lahiri D. B., Bull. Inter. Stat. Inst., 33, 133-140
5. On the estimation of ratio and product of population parameters (1965) Singh M. P., Sankhya, 27 (B), 321—328.

അഭ്യാസം 6

6.1 അഞ്ച് അംഗങ്ങളുള്ള ഒരു സമഷ്ടിയിലെ അംഗങ്ങളുടെ x, y അഭിലക്ഷണ മൂല്യങ്ങൾ താഴെ കൊടുക്കുന്നു.

അംഗം	1	2	3	4	5
x	1	3	2	5	11
y	2	7	5	12	21

y മൂല്യങ്ങളുടെ തുകയായ Y, മൂന്നംഗങ്ങളുള്ള പ്രതിസ്ഥാപനമില്ലാത്ത ഒരു ലഘുയാദൃച്ഛിക സാമ്പിളെടുത്ത് ആകലനം ചെയ്യുക. സാധ്യമായ എല്ലാ സാമ്പിളുകളുമെടുത്ത് ആവശ്യമായ ഗണനങ്ങൾ നടത്തി അനുപാതാകലത്തിൻെറ മേന്മ തെളിയിക്കുക.

6.2. മുൻഉദാഹരണത്തിൽ y യുടെ സമാശ്രയണാകലം നിർണയിക്കുക.

6.3. ഉത്തരേന്ത്യയിലെ ഒരു സംസ്ഥാനത്തെ തിരഞ്ഞെടുക്കപ്പെട്ട 15 ഗ്രാമങ്ങളിലെ കറവപ്പശുക്കളുടെ 1958 ലെയും 1956 ലെയും എണ്ണം താഴെ കൊടുക്കുന്നു. ആകെ ഗ്രാമങ്ങളുടെ എണ്ണം 9200 ഉം 1956 ലെ ആകെ കറവപ്പശുക്കളുടെ എണ്ണം 576,000 ഉം ആണെങ്കിൽ 1958 ലെ ആകെ കറവപ്പശുക്കളുടെ എണ്ണവും അതിൻെറ സാമ്പിളന പ്രസരണവും അനുപാതാകലനരീതിയിൽ നിർണയിക്കുക.

ഗ്രാമം	1958 ലെ പശുക്കളുടെ എണ്ണം	1956 ലെ എണ്ണം
1	35	47
2	38	46
3	71	253
4	4	119
5	63	121
6	4	4
7	14	5
8	7	7
9	66	50
10	44	162
11	8	9
12	229	256
13	27	74
14	30	28
15	29	41

6 4. മുൻ ഉദാഹരണത്തിലെ കറവപ്പശുക്കളുടെ എണ്ണം സമാശ്രയണരീതിയിൽ ആകലനം ചെയ്യുക.

6.5. അഭ്യാസം 6.3 ലെ പശുക്കളുടെ എണ്ണത്തിൻെറ 95% വിശ്വാസ്യതാന്തരാളം നിർണയിക്കുക.

7

മററു സാമ്പിളനരീതികൾ

1. സംഘസാമ്പിളനം

സംഘസാമ്പിളനം എന്തു്? എന്തിനു്?

ഇതുവരെ പരിഗണിച്ച സാമ്പിളന രീതികളിലെല്ലാം സമഷ്ടിയിലെ പ്രാഥമിക വ്യക്തികളെ തന്നെയാണു് സാമ്പിളിലേക്കു തിരഞ്ഞെടുക്കുന്നതു്. താത്വികമായി പറഞ്ഞാൽ ഇമ്മാതിരി സാമ്പിളുകൾക്കാണു് കൂടുതൽ സൂക്ഷ്മത യുണ്ടായിരിക്കുക. പക്ഷേ അനുവദിക്കാവുന്ന ചെലവും സമയവും പരിമിതമായിരിക്കുമ്പോൾ ഈ രീതി അത്ര അഭികാമ്യമല്ല. സമഷ്ടിയിലെ പ്രാഥമികാംഗങ്ങളുടെ എല്ലാം ഒരു ലിസ്റ്റ് തയ്യാറാക്കാതെ സാമ്പിളെടുക്കാനാവില്ലെന്നുള്ളതാണു് ഈ സമ്പ്രദായങ്ങളുടെ ഏററവും പ്രധാന തകരാറു്. ക്രമാനുസൃത സാമ്പിളനത്തിൽ ഈ ബുദ്ധിമുട്ട് ഭാഗികമായി ഒഴിവാക്കിയിട്ടുണ്ടെങ്കിലും വൻകിട സാമ്പിളന പ്രക്രിയകളിൽ അതുകൊണ്ടു് കാര്യമായ നേട്ടമൊന്നും ഉണ്ടാവുന്നില്ല 'സാമ്പിളനഫ്രേം' എന്നു വ്യവഹരിക്കാറുള്ള ഈ ലിസ്റ്റ് മിക്കവാറും ഓരോ സാമ്പിളനത്തിനും മുൻപായി പ്രത്യേകം തയ്യാറാക്കേണ്ടിവരും. ഒരു പക്ഷേ നേരത്തെ തയ്യാറാക്കിയ ഫ്രേമുകൾ ലഭ്യമായിരുന്നാൽ തന്നെയും അവ വേണ്ട പോലെ പരിശോധിച്ചു് തെററു തിരുത്താതെ ഉപയോഗിക്കുന്നതു് അപകടകരമാണു്. ഫ്രേം തയ്യാറാക്കൽ വളരെ സമയനഷ്ടവും ചെലവും ഉൾക്കൊള്ളുന്നതാണെന്നു് പറയേണ്ടതില്ലല്ലോ. മറെറാരു ബുദ്ധിമുട്ട് വിവരശേഖരണത്തോടനുബന്ധിച്ചാണു് ഉണ്ടാകുന്നതു്. ഇതുവരെ പരിഗണിച്ച എല്ലാ സാമ്പിളനരീതികളിലും സാമ്പിളിൽ ഉൾപ്പെടുത്തപ്പെടുന്ന വ്യക്തികൾ സമഷ്ടിയുടെ എല്ലാ ഭാഗങ്ങളിലുമായി ചിതറിക്കിടക്കുന്നവയായിരിക്കും. അവയെ കണ്ടുപിടിച്ചു് വിവരം ശേഖരിക്കുക എന്നതു് ബുദ്ധിമുട്ടുള്ള കാര്യമാണു്. അതിനു് കൂടുതൽ സമയം വേണ്ടിവരും. പണച്ചെലവും കുറച്ചൊന്നുമായിരിക്കുകയില്ല. ഒരുദാഹരണം കൊണ്ടു് ഈ ആശയം വിശദീകരിക്കാം. കേരളത്തിലെ ഒരു കുടുംബത്തിൻെറ അധീനതയിലുള്ള ശരാശരി കൃഷിഭൂമി എത്രയെന്നു കണ്ടുപിടിക്കുക

യാണ് അന്വേഷണ ലക്ഷ്യമെന്നിരിക്കട്ടെ. ഇവിടെ കുടുംബമാണ് വ്യക്തി. കേരളത്തിലെ എല്ലാ കുടുംബങ്ങളുടേയും ഒരു ലിസ്റ്റാണ് സാമ്പിളന ഫ്രേം. ഇതു തയ്യാറാക്കാനുള്ള ബുദ്ധിമുട്ട് പ്രത്യേകം വിവരിക്കേണ്ടതില്ലല്ലോ. 10,000 അംഗങ്ങളുള്ള ഒരു സാമ്പിളാണ് എടുക്കുന്നതു് എന്നിരിക്കട്ടെ. നാം ഇതുവരെ പഠിച്ച ഏതു സാമ്പിളന സമ്പ്രദായം അവലംബിച്ചാലും ഈ 10,000 കുടുംബങ്ങൾ കേരളത്തിൻെറ നാനാഭാഗങ്ങളിലായി ചിതറിക്കിടക്കും. അവ ഓരോന്നും കണ്ടുപിടിച്ചു് വിവരം ശേഖരിക്കുവാൻ അന്വേഷകനു് യാത്രച്ചെലവ്, താമസച്ചെലവ് തുടങ്ങിയ ഇനങ്ങളിൽ വലിയ ഒരു സംഖ്യ ചെലവാകും. സമയവും വളരെയേറെയാകും.

ഫ്രേം തയ്യാറാക്കുന്നതിനും വിവരം ശേഖരിക്കുന്നതിനും വേണ്ടിവരുന്ന ഈ അധികച്ചെലവും സമയനഷ്ടവും ഒരു പരിധിവരെ ഒഴിവാക്കാൻ പര്യാപ്തമായ ഒരു സാമ്പിളന രീതിയാണ് സംഘസാമ്പിളനം. ഇവിടെ സാമ്പിളന വ്യക്തികൾ പ്രാഥമികവ്യക്തികളുടെ സംഘങ്ങളാണ്. അതായതു്, അടുത്തടുത്തുള്ള ഏതാനും വ്യക്തികളുടെ കൂട്ടങ്ങൾ. സമഷ്ടിയെ ആദ്യം സംഘങ്ങളായി വിഭജിക്കുന്നു. അവയിൽ നിന്നു് ആവശ്യമുള്ളത്ര സംഘങ്ങളെ ഏതെങ്കിലും സാമ്പിളനരീതിയനുസരിച്ച് തിരഞ്ഞെടുക്കുന്നു. തിരഞ്ഞെടുത്ത സംഘങ്ങളിലെ എല്ലാ അംഗങ്ങളിൽ നിന്നും വിവരം ശേഖരിക്കുന്നു.

ഇവിടെ സാമ്പിളന ഫ്രേം പ്രാഥമിക വ്യക്തികളുടെ ലിസ്റ്റല്ല, നേരെമറിച്ച് സംഘങ്ങളുടെ ലിസ്റ്റാണ്. സംഘങ്ങളായുള്ള വിഭജനം ബുദ്ധിപൂർവം കരുതിക്കൂട്ടി നിർവഹിച്ചാൽ, അവയുടെ ലിസ്റ്റ തയ്യാറാക്കുക അത്ര വിഷമമുള്ള ഒരു കാര്യമായിരിക്കുകയില്ല. തിരഞ്ഞെടുത്ത സംഘങ്ങളിലെ എല്ലാ അംഗങ്ങളിൽ നിന്നും വിവരം ശേഖരിക്കുന്നതുകൊണ്ടു്, സാമ്പിളിൽ ഉൾപ്പെടുത്തിയിരിക്കുന്ന അംഗങ്ങളെ തേടിപ്പിടിക്കുന്ന പ്രശ്നവുമില്ല. മുൻ ഉദാഹരണത്തിൽ, കേരളത്തിലെ ഓരോ താലൂക്കും ഓരോ സംഘമായി പരിഗണിക്കുന്നു എന്നിരിക്കട്ടെ. താലൂക്കുകളുടെ ഒരു ലിസ്റ്റ് തയ്യാറാക്കുക ഒട്ടും ബുദ്ധിമുട്ടുള്ള ഒരു കാര്യമല്ല. എല്ലാ കുടുംബങ്ങളുടെയും ലിസ്റ്റ് തയ്യാറാക്കുന്നതുമായി താരതമ്യപ്പെടുത്തുമ്പോൾ എത്രയോ എളുപ്പമായ ഒന്നാണ് അതു്. തെരഞ്ഞെടുക്കപ്പെടുന്ന താലൂക്കിലെ എല്ലാ കുടുംബങ്ങളിൽ നിന്നും വിവരം ശേഖരിക്കുന്നതുകൊണ്ടു്, വിവരശേഖരണം താരതമ്യേന അനായാസമായും തീരുന്നു.

സംഘ സാമ്പിളനത്തിൻെറ സൂക്ഷ്മത മററുസാമ്പിളനരീതികളുമായി താരതമ്യപ്പെടുത്തുമ്പോൾ മോശമാണെന്നും സംഘങ്ങളിൽ ഉൾപ്പെടുത്തുന്ന അംഗങ്ങളുടെ എണ്ണം വർധിക്കുമ്പോൾ അതു് കുറഞ്ഞു വരുമെന്നും നാം വഴിയെ കാണും. പക്ഷേ സംഘങ്ങളുടെ പരിമാണം വർധിക്കുമ്പോൾ സാമ്പിളനത്തിനു് ആവശ്യമായിവരുന്ന ചെലവ് കൂടുമെന്നുള്ളതു് ഇതിൻെറ മറുവശമാണ്. മറെറാരു തരത്തിൽ പറഞ്ഞാൽ, ചെലവു കുറയുമ്പോൾ സൂക്ഷ്മതയും കുറയും. ഇവിടെ അവലംബിക്കാവുന്ന നയം അനുവദിക്കാവുന്ന ചെലവിൽ ലഭിക്കാവുന്ന ഏററവും സൂക്ഷ്മതയുള്ള സാമ്പിളിൽ ലക്ഷ്യം വെക്കുക എന്നതാണ്. അതായതു് സംഘത്തിന്റെ പരിമാണം അതിനനുസരിച്ച് നിർണയിക്കുക.

സമഷ്ടിയെ സംഘങ്ങളായി വിഭജിക്കുമ്പോൾ പ്രത്യേകം ശ്രദ്ധിക്കാനുള്ള കാര്യങ്ങൾ അടുത്തടുത്തുള്ള പ്രാഥമിക വ്യക്തികളെ വേണം ഓരോ സംഘത്തിലും ഉൾപ്പെടുത്താൻ എന്നതും ഒരു പ്രാഥമിക വ്യക്തി ഒന്നിലധികം സംഘങ്ങളിലെ അംഗമായിരിക്കരുത് എന്നതുമാണ്. അടുത്തടുത്തുള്ള വ്യക്തികളെ ചേർത്ത് സംഘങ്ങൾ രൂപീകരിച്ചില്ലെങ്കിൽ ഈ സമ്പ്രദായത്തിന് ഉണ്ടെന്നു നാം പറഞ്ഞുവെച്ച പ്രധാന ഗുണമായ വിവരശേഖരണത്തിന്റെ സൗകര്യം ഇല്ലാതാവും. ഒരേ വ്യക്തി ഒന്നിലധികം സംഘങ്ങളിൽ അംഗമായി വന്നാൽ, ഒരേ വ്യക്തി സാമ്പിളിൽ ഒന്നിലധികം പ്രാവശ്യം ഉൾപ്പെട്ടേക്കാം എന്ന താത്വികമായ തകരാറിന് അതു് വഴി തെളിയിക്കും.

പ്രകൃത്യാ ഉള്ള സംഘങ്ങളെ അതേ പടി സ്വീകരിക്കുന്നതു് പലതു കൊണ്ടും സൗകര്യമാണ്. പക്ഷേ, തുല്യപരിമാണമുള്ള സംഘങ്ങളാണ് ആവശ്യമുള്ളതെങ്കിൽ ഈ രീതി സ്വീകരിക്കാൻ ബുദ്ധിമുട്ടുണ്ടാവും. മുൻഉദാഹരണത്തിലെ താലൂക്ക് എന്ന സംഘം പ്രകൃത്യാ ഉള്ളതാണ്. പക്ഷെ ഓരോ താലൂക്കിലെയും വീടുകളുടെ എണ്ണം തുല്യമായിരിക്കണമെന്നില്ല.

മുൻവിവരിച്ച സംഘസാമ്പിളനരീതിക്ക് ഏകഘട്ട സംഘസാമ്പിളനം എന്നു പേർ പറയും. തിരഞ്ഞെടുക്കപ്പെടുന്ന സംഘങ്ങളിലെ എല്ലാ അംഗങ്ങളിൽ നിന്നും വിവരം ശേഖരിക്കാതെ അവയിൽ നിന്നു് എടുക്കുന്ന ഒരു സാമ്പിളിൽ ഉൾപ്പെട്ടു വരുന്ന അംഗങ്ങളിൽ നിന്നു മാത്രം വിവരശേഖണം നടത്തുകയാണെങ്കിൽ ആ സാമ്പിളനത്തിന് ദ്വിഘട്ട സംഘസാമ്പിളനം എന്നാണ് പറയുക. മുൻഉദാഹരണത്തിൽ തിരഞ്ഞെടുക്കപ്പെടുന്ന താലൂക്കുകളിലെ എല്ലാ വീടുകളിൽ നിന്നും വിവരം ശേഖരിക്കാതെ അവയുടെ ഒരു സാമ്പിളിൽ നിന്നു മാത്രം വിവരശേഖരണം നടത്തുക എന്നതാണ് ഈ രീതി.

ഇതുപോലെതന്നെ ത്രിഘട്ട സംഘസാമ്പിളനം, ബഹുഘട്ട സംഘസാമ്പിളനം തുടങ്ങിയവയും നിർവചിക്കാവുന്നതാണ്. സമഷ്ടിയെ ആദ്യം സംഘങ്ങളായും, സംഘങ്ങളെ ഉപസംഘങ്ങളായും, വിഭജിക്കുക. സംഘങ്ങളിൽ നിന്നു് ഒരു സാമ്പിളെടുക്കുക. തിരഞ്ഞെടുക്കപ്പെട്ട സംഘങ്ങളിലെ ഉപസംഘങ്ങളിൽ നിന്നു് മറെറാരു സാമ്പിളെടുക്കുക. അങ്ങനെ ലഭിക്കുന്ന ഉപസംഘങ്ങളിലെ അംഗങ്ങളിൽ നിന്നു് മൂന്നാമതൊരു സാമ്പിളെടുത്തു് അവയിൽ നിന്നു് വിവരം ശേഖരിക്കുക. ഇതാണ് ത്രിഘട്ട സംഘസാമ്പിളനം. ഇതുപോലെ ബഹുഘട്ടസാമ്പിളനത്തേയും നിർവചിക്കാം. താലൂക്കുകളാണ് സംഘങ്ങളെങ്കിൽ വില്ലേജുകളെ ഉപസംഘങ്ങളായി സ്വീകരിക്കാം. താലൂക്കുകളുടെ ഒരു സാമ്പിൾ ആദ്യമെടുക്കുക. തിരഞ്ഞെടുക്കപ്പെട്ട താലൂക്കുകളിലെ വില്ലേജുകളിൽ നിന്നു് മറെറാരു സാമ്പിളെടുക്കുക. ഇങ്ങനെ ലഭിച്ച വില്ലേജുകളിലെ വീടുകളിൽ നിന്നു് മൂന്നാമതൊരു സാമ്പിളെടുത്തു് ആ വീടുകളിൽ നിന്നു മാത്രം വിവരം ശേഖരിക്കുക. ഇതു് ത്രിഘട്ട സാമ്പിളനരീതിയാണ്. ഇത്തരം സാമ്പിളന രീതികൾ ബഹുഘട്ട സാമ്പിളനം എന്ന ശീർഷകത്തിൽ പ്രത്യേകം പരിഗണിക്കുന്നതാണ്.

2. ഏകഘട്ട സംഘസാമ്പിളനം

ഒരേയെണ്ണം പ്രാഥമികാംഗങ്ങളെ ഉൾക്കൊള്ളുന്ന സംഘങ്ങളായി സമഷ്ടിയെ വിഭജിക്കുമ്പോഴത്തെ അവസ്ഥയാണ് ആദ്യം ചർച്ച ചെയ്യുന്നതു്. സമഷ്ടിയിൽ M N അംഗങ്ങളെ ഉള്ളതെന്നിരിക്കട്ടെ. അതു് M അംഗങ്ങൾ വീതമുള്ള N സംഘങ്ങളായി വിഭജിക്കപ്പെട്ടിരിക്കുന്നു. ഇതിൽ നിന്നു് n സംഘങ്ങളുള്ള പ്രതിസ്ഥാപനമില്ലാത്ത ഒരു ലഘുയാദൃച്ഛിക സാമ്പിളെടുത്തു്, തിരഞ്ഞെടുക്കപ്പെട്ട ഓരോ സംഘത്തിലും ഉൾപ്പെട്ട എല്ലാ അംഗങ്ങളിൽ നിന്നും വിവരം ശേഖരിക്കുന്നു എന്നും സങ്കൽപിക്കുക. x എന്ന അഭിലക്ഷണമാണ് പരിഗണിക്കപ്പെടുന്നതെന്നും അതിൻെറ മൂല്യങ്ങളുടെ സമാന്തരമാധ്യമാണ് ആകലനം ചെയ്യേണ്ട പ്രാചലമെന്നും ഇരിക്കട്ടെ. i-ാമതു് സാമ്പിൾ സംഘത്തിലെ j-ാമതു് അംഗത്തിൻെറ അഭിലക്ഷണമൂല്യം x_{ij} എന്നും സാമ്പിളിൻെറ സമാന്തരമാധ്യം $\bar{x}$ എന്നും സമഷ്ടിയുടെ സമാന്തരമാധ്യം $\bar{X}$ എന്നും സങ്കൽപിക്കുക. i-ാമതു സാമ്പിൾ സംഘത്തിലെ അഭിലക്ഷണമൂല്യങ്ങളുടെ സമാന്തരമാധ്യം $\bar{x}_i$ എന്നും ഇരിക്കട്ടെ. $\bar{X}$ ൻെറ ആകലമായ $\bar{x}_c$, $\bar{x}$ ആണെന്നു നിർവചിക്കപ്പെട്ടിരിക്കുന്നു. അതായതു്,

$$\bar{x}_c = \frac{1}{n}\sum_{i=1}^{n} \bar{x}_i = \frac{1}{nM}\sum_{i=1}^{n}\sum_{j=1}^{M} x_{ij} \qquad 7.2.1$$

ആകെ N സംഘങ്ങളുള്ളതിൽ നിന്നു് n സംഘങ്ങളാണല്ലോ തിരഞ്ഞെടുക്കുന്നതു്. അതുകൊണ്ടു് $\binom{N}{n}$ സംഘസാമ്പിളുകളാണ് ഈ സമഷ്ടിയിൽ നിന്നു് എടുക്കാവുന്നതു്. ഏതു സംഘവും $\binom{N-1}{n-1}$ സാമ്പിളുകളിൽ ഉൾപ്പെട്ടുവരും. t-ാമതു് സാമ്പിളിലെ i-ാമതു് സംഘത്തിൻെറ സമാന്തരമാധ്യം $\bar{x}_{it}$ എന്നു സങ്കൽപിക്കുക.

$$E(\bar{x}_c) = \frac{1}{\binom{N}{n}} \sum_{t=1}^{\binom{N}{n}} \left[\frac{1}{n} \sum_{i=1}^{n} \bar{x}_{it} \right]$$

$$= \frac{1}{n\binom{N}{n}} \binom{N-1}{n-1} \sum_{i=1}^{N} \bar{X}_i$$

[ഇവിടെ $\bar{X}_i$, i-ാമതു് സംഘത്തിലെ അഭിലക്ഷണമൂല്യങ്ങളുടെ സമാന്തരമാധ്യമാണ്.]

$$= \frac{1}{N} \frac{\binom{N-1}{n-1}}{\binom{N-1}{n-1}} \sum_{i=1}^{N} \bar{X}_i = \frac{1}{N} \sum_{i=1}^{N} \bar{X}_i = \bar{X}$$

ഇതിൽ നിന്നു് $\bar{x}_c$ ഒരു അനഭിനത ആകലമാണെന്നു സിദ്ധിക്കുന്നു.

അടുത്തതായി അതിൻെറ സാമ്പിളനപ്രസരണം എന്താണെന്നു പരിശോധിക്കാം.

$$V(\bar{x}_c) = E(\bar{x}_c - \bar{X}^2)$$

$$= E\left[\frac{1}{n}\left(\sum_{i=1}^{n} \bar{x}_i\right) - \bar{X}\right]^2$$

$$= E\left[\frac{1}{n}\sum_{i=1}^{n} (\bar{x}_i - \bar{X})\right]^2$$

$$= \frac{1}{n^2}\sum_{j=1}^{n} E(\bar{x}_i - \bar{X})^2 + \frac{1}{n^2}\sum_{i\neq j=1}^{n} E(\bar{x}_i - \bar{X})(\bar{x}_j - \bar{X})$$

ഇവിടെ ആദ്യത്തെ പദത്തിലെ $E(\bar{x}_i - \bar{X})^2$ എന്നതു് പരിശോധിക്കുക. $\bar{x}_i$ എന്ന യാദൃച്ഛികചരത്തിനു് $\bar{X}_1\, \bar{X}_2 \cdots \bar{X}_N$ എന്ന N മൂല്യങ്ങൾ (N സംഘങ്ങളുടെ സമാന്തര മാധ്യങ്ങൾ) ആണു് എടുക്കാവുന്നതു്. ലഘുയാദൃച്ഛിക സാമ്പിളന രീതി അവലംബിക്കുന്നതു കൊണ്ടു് ഇവ ഓരോന്നിൻേറയും സംഭാവ്യത അത്രെ. അങ്ങനെ,

$$E(\bar{x}_i - \bar{X})^2 = \frac{1}{N}\sum_{i=1}^{N} (\bar{X}_i - \bar{X})^2 = \sigma_b^2$$

ഇതു് സംഘസമാന്തരമാധ്യങ്ങളുടെ പ്രസരണമാണല്ലൊ. ഇതിനെ സംഘാന്തരപ്രസരണം എന്നു വ്യവഹരിക്കുന്നു. σ_b^2 എന്ന ചിഹ്നം കൊണ്ടു് നമുക്കതിനെ കുറിക്കാം. രണ്ടാമത്തെ പദത്തിലെ $E(\bar{x}_i - \bar{X})(\bar{x}_j - \bar{X})$ $(i \neq j)$ എന്നതു് അടുത്തതായി പരിശോധിക്കാം. i-ാമതു് സംഘവും j-ാമതു സംഘവും കൂടി തിരഞ്ഞെടുക്കപ്പെടുവാനുള്ള സംഭാവ്യത $\frac{1}{N(N-1)}$ അത്രെ. $\bar{x}_i$ നു് $\bar{X}_1, \bar{X}_2 \ldots \bar{X}_N$ എന്നീ മൂല്യങ്ങളിൽ ഏതു വേണമെങ്കിലും എടുക്കാം. $\bar{X}_j$ നു് $\bar{X}_i$ എടുത്തതല്ലാത്ത എല്ലാ മൂല്യങ്ങളും എടുക്കാം. അതുകൊണ്ടു്,

$$E(\bar{x}_j - \bar{X})(\bar{x}_j - \bar{X}) = \frac{1}{N(N-1)} \sum_{l\neq m=1}^{N} (\bar{X}_l - \bar{X})(\bar{X}_m - \bar{X})$$

$$= \frac{1}{N(N-1)} \left\{ \left[\sum_{i=1}^{N} (\bar{X}_i - \bar{X}) \right]^2 - \sum_{i=1}^{N} (\bar{X}_i - \bar{X})^2 \right\}$$

$$= - \frac{1}{N(N-1)} \sum_{i=1}^{N} (\bar{X}_i - \bar{X})^2 = - \frac{1}{(N-1)} \sigma_b^2$$

അതുകൊണ്ടു്,

$$V(\bar{x}_c) = \frac{1}{n^2} \; n\sigma_b^2 - \frac{n(n-1)}{n^2} \cdot \frac{\sigma_b^2}{N-1}$$

$$= \left[\frac{1}{n} - \frac{n-1}{n(n-1)} \right] \sigma_b^2 = \frac{N-n}{N-1} \cdot \frac{\sigma_b^2}{n} \qquad 7.2.2$$

$V(\bar{x}_c)$ ന്റെ ആകലം ലഭിക്കണമെങ്കിൽ σ_b^2 ആകലനം ചെയ്യണം. സാമ്പിളിൽ നിന്നു് $\bar{x}_i$ ന്റെ മൂല്യങ്ങൾ കണ്ടതിനു ശേഷം $S_b^2 = \frac{1}{n-1} \sum_{i=1}^{n} (\bar{x}_i - \bar{x}_c)^2$ കണ്ടുപിടിച്ചാൽ അതു് σ_b^2 ന്റെ ഒരു അനഭിനത ആകലമായിരിക്കും. അങ്ങനെ,

$$\hat{V}(\bar{x}_c) = \frac{N-n}{N-1} \frac{S_b^2}{n} \qquad 7.2.3$$

σ_b^2 ൻെറ മൂല്യം അന്തരാവർഗ സഹബന്ധമുപയോഗിച്ചും എഴുതാവുന്നതാണു്.

$$\sigma_b^2 = \frac{1}{N} \sum_{i=1}^{N} (\bar{X}_i - \bar{X})^2 = \frac{1}{NM} \sum_{i=1}^{M} \left\{ \sum_{j=1}^{M} (X_{ij} - \bar{X}) \right\}^2$$

$$= \frac{1}{NM^2} \left[\sum_{i=1}^{N} \sum_{j=1}^{M} (X_{ij} - \bar{X})^2 + \sum_{i=1}^{N} \sum_{l \neq j=1}^{M} (X_{ij} - \bar{X})(X_{ij} - \bar{X}) \right]$$

$$= \frac{\sigma^2}{M} [1 + (M-1)\rho_c]$$

അങ്ങനെ, $$V(\bar{x}_c) = \frac{N-n}{N-1} \cdot \frac{1}{MN} [1 + (M-1)\rho_c] \sigma^2 \qquad 7.2\ 4$$

സംഘസാമ്പിളനത്തിനു പകരം പ്രതിസ്ഥാപനമില്ലാത്ത ലഘുയാദൃച്ഛിക സാമ്പിളനരീതിയാണു് അവലംബിച്ചിരുന്നതെങ്കിൽ, (സാമ്പിൾ പരിമാണം Mn)

$$V(\bar{x}) = \frac{NM-nM}{NM}\frac{S^2}{nM}$$

$$= \frac{NM-nM}{NM-1}\frac{\sigma^2}{nM}$$

ആണെന്നു നാം നേരത്തെ മനസ്സിലാക്കുകയുണ്ടായി.

അതുകൊണ്ടു് ലഘുയാദൃച്ഛികസാമ്പിളനവുമായി താരതമ്യപ്പെടുത്തിയാൽ സംഘസാമ്പിളനത്തിൻെറ ക്ഷമത,

$$\frac{V(\bar{x})}{V(\bar{x}_c)} = \frac{M(N-1)}{NM-1} \cdot \frac{1}{1+(M-1)\rho_c} \qquad 7.25$$

ഇതു് എപ്പോഴും ഒന്നിൽ കുറവായിരിക്കുമെന്നതു കൊണ്ടു്, സംഘസാമ്പിളനം ലഘുയാദൃച്ഛികസാമ്പിളനത്തെക്കാൾ ക്ഷമത കുറഞ്ഞതായിരിക്കുമെന്നു് വന്നു ചേരുന്നു. N ൻെറ മൂല്യം വർധിക്കുകയും (NM നിശ്ചിതമായതുകൊണ്ടു്) അതിനനുസരിച്ചു് M ൻെറ മൂല്യം കുറയുകയും ചെയ്യുമ്പോൾ ഈ ക്ഷമത 1 നോട്ടു് അടുത്തു വരുന്നതായി കാണാം. $N = NM$ ആകുമ്പോൾ, അതായതു്, $M = 1$ ആകുമ്പോൾ, ക്ഷമത 1 ആവും. അതായതു്, സംഘസാമ്പിളനത്തിൻെറ ക്ഷമത വർധിപ്പിക്കുവാനുള്ള മാർഗം സംഘങ്ങളുടെ അംഗസംഖ്യ കുറച്ചു് എണ്ണം വർധിപ്പിക്കുകയാണു്. പക്ഷെ ഇതിൻെറ മറെറാരു ഫലം സാമ്പിളനച്ചെലവിൻെറ വർധനവത്രെ. അതുകൊണ്ടു് അനുവദിക്കാവുന്ന ചെലവു് ആദ്യം നിശ്ചയിച്ചു് സാമ്പിളനപ്രസരണം ഏററവും കുറഞ്ഞിരിക്കത്തക്ക വണ്ണം സംഘങ്ങളുടെ എണ്ണം നിർണയിക്കുകയോ, സാമ്പിളനപ്രസരണം ഒരു നിശ്ചിതമൂല്യത്തിൽ കവിയാത്ത വിധവും ചെലവു് ഏററവും കുറഞ്ഞിരിക്കത്തക്ക വണ്ണവും സംഘങ്ങളുടെ എണ്ണം നിർണയിക്കുകയോ ചെയ്യുക എന്നതാണു്. സംഘങ്ങളുടെ എണ്ണം കൊണ്ടു് സമഷ്ടിയുടെ അംഗസംഖ്യയെ ഹരിച്ചാൽ സംഘപരിമാണം ലഭിക്കും.

സംഘസാമ്പിളനത്തൻെറ ചെലവു് സാമാന്യമായി പറഞ്ഞാൽ മൂന്നു് തരത്തിലുള്ള ചെലവുകൾ ചേർന്നതാണു്. ഒന്നു്, സാമ്പിളനത്തിനുള്ള ഒരുക്കങ്ങൾക്കു വേണ്ടിവരുന്നതു്. രണ്ടു്, ഓരോ സംഘങ്ങളിലും എത്തിചേരുവാനും അവയെപ്പററി പ്രാഥമിക പഠനങ്ങൾ നടത്തുവാനും മററും വേണ്ടിവരുന്ന ചെലവു്. മൂന്നു്, ഓരോ വ്യക്തിയിൽ നിന്നും വിവരം ശേഖരിക്കാൻ വേണ്ടിവരുന്ന ചെലവു്. ഇതിൽ ഒന്നാമത്തെ ചെലവു് C_0 എന്നിരിക്കട്ടെ. രണ്ടാമത്തേതു് എല്ലാ സംഘങ്ങൾക്കും തുല്യമാണെന്നും അതു് സംഘമൊന്നിനു് C_1 വീതമാണെന്നും സങ്കല്പിക്കുക. മൂന്നാമത്തേതു് അംഗമൊന്നിനു് C_2 വീതമാണെന്നിരിക്കട്ടെ. അങ്ങനെയായാൽ ആകെ ചെലവു്

$$C = C_0 + nC_1 + nM\,C_2 \qquad 7.26$$

ഏററവും ലളിതമായ ചെലവുഫലനമാണിതു്. അതതു സന്ദർഭത്തിൻെറ പ്രത്യേകതയനുസരിച്ചു് ഇതിൻെറ ഘടനയിൽ മാററങ്ങൾ വരുത്താവുന്നതും, വരുത്തേണ്ടതുമത്രെ. അങ്ങനെ ലഭിക്കുന്ന ചെലവു ഫലനവും 7.2.2–ലെ സാമ്പിളനപ്രസരണഫലനവും ഉപയോഗപ്പെടുത്തി, സംഘങ്ങളുടെ എണ്ണം മുൻപറഞ്ഞ രണ്ടു സമീപനരീതികളിൽ എതെങ്കിലും ഒന്നു് അവലംബിച്ചു് നിർണയിക്കാൻ കഴിയും. തത്വത്തിൽ ഇതു് സാധ്യമാണെങ്കിലും പ്രായോഗികമായി പറഞ്ഞാൽ നമുക്കു് കൈകാര്യം ചെയ്യേണ്ടിവരുന്ന ഫലനങ്ങളുടെ സങ്കീർണത കാരണം ഈ രീതി വിജയകരമായി ഉപയോഗപ്പെടുത്താൻ ബുദ്ധിമുട്ടുണ്ടു്. ഭാഗികമായി നിരീക്ഷണപരമായ സമീപനം കൂടി ആവശ്യമായി വരും.

സമഷ്ടിയുടെ സമാന്തരമാധ്യം ആകലനം ചെയ്യുന്നതിനുള്ള മാർഗവും അതിൻെറ സാമ്പിളന പ്രസരണത്തിൻെറ വ്യുൽപാദനവുമാണല്ലൊ ഇവിടെ ചർച്ച ചെയ്തതു്. സമഷ്ടിയുടെ അഭിലക്ഷണമൂല്യങ്ങളുടെ തുകയാണു് ആകലനം ചെയ്യേണ്ടതെങ്കിൽ അതു് ഈ ആകലത്തിൽ നിന്നു് കണ്ടുപിടിക്കാവുന്നതാണു്.

$$\hat{X}_c = N\bar{x}_c \qquad 7.2.7$$

$$V(\hat{x}_c) = N^2V(\bar{x}_c) \qquad 7.2.8$$

അതു പോലെ തന്നെ അനുപാതങ്ങൾ ആകലനം ചെയ്യാനും പ്രയാസമില്ല. സമഷ്ടിയിലെ ഏതെങ്കിലും പ്രത്യേകതയുള്ള അംഗങ്ങളുടെ അനുപാതമാണു് ആകലനം ചെയ്യേണ്ടതെന്നിരിക്കട്ടെ. സാമ്പിളിലെ ആ പ്രത്യേകതയുള്ള അംഗങ്ങളുടെ അഭിലക്ഷണമൂല്യം 1 എന്നും മററുള്ളവയുടേതു് 0 എന്നും സങ്കൽപിച്ചു് $\bar{\ }_c$ കണ്ടാൽ അതു് അനുപാതത്തിൻെറ ആകലമായിരിക്കും.

ഇവിടെ സാമ്പിൾ തിരഞ്ഞെടുക്കാൻ പ്രതിസ്ഥാപനമില്ലാത്ത ലഘുയാദൃച്ഛികം സാമ്പിളന രീതിയാണു് അവലംബിച്ചതു്. മററു സാമ്പിളനരീതികളും ഇതിനായി ഉപയോഗപ്പെടുത്താം. അതനുസരിച്ചു് ആകലത്തിൻെറയും പ്രസരണത്തിൻെറയും വ്യംജകങ്ങളിൽ മാററം വരുത്തണമെന്നു മാത്രം.

ഈ ചർച്ചയിൽ സംഘങ്ങളെല്ലാം ഒരേ വലുപ്പമുള്ളവയാണു് എന്നാണല്ലൊ സങ്കൽപിച്ചതു്. പക്ഷേ ഇതു് അത്യന്തം അസൗകര്യമായ ഒരു സങ്കൽപമാണെന്നു പറയാതെ തരമില്ല. പ്രകൃത്യാ ഉള്ള വിഭജനമനുസരിച്ചു് സംഘങ്ങൾ സ്വീകരിക്കുകയാണു് സൗകര്യപ്രദമായ രീതി. മുൻപു അവതരിപ്പിച്ച ഉദാഹരണത്തിൽ താലൂക്കുകൾ സംഘങ്ങളായി സ്വീകരിക്കുന്നതു് നല്ലതാണെന്നു നാം പറയുകയുണ്ടായി. പക്ഷേ എല്ലാ താലൂക്കുകളിലെയും വീടുകളുടെ എണ്ണം തുല്യമായിരിക്കണമെന്നില്ല. അതായതു് സംഘങ്ങൾ തുല്യപരിമാണമുള്ളവയല്ല. പ്രായോഗികസൗകര്യം പരിഗണിക്കുമ്പോൾ തുല്യപരിമാണമുള്ള സംഘങ്ങൾ വേണമെന്നു് നിർബന്ധം പിടിക്കാതെ, ആകലന സമ്പ്രദായത്തിൽ ആവശ്യമായ മാററം വരുത്തുന്നതാണു് കൂടുതൽ ബുദ്ധിപൂർവകം.

സമഷ്ടിയെ N അംഘങ്ങളായി വിഭജിച്ചിരിക്കുന്നു എന്നും i–ാമതു് സംഘത്തിൽ M_i അംഗങ്ങളാണു് ഉള്ളതെന്നും സങ്കല്പിക്കുക. പ്രതിസ്ഥാപനമില്ലാത്ത യാദൃച്ഛികസാമ്പിളനരീതിയിൽ n സംഘങ്ങൾ തിരഞ്ഞെടുത്തു് അവയിലെ എല്ലാ അംഗങ്ങളിൽ നിന്നും വിവരശേഖരണം നടത്തുന്നു. i–ാമതു് സാമ്പിൾ സംഘത്തിൽ M_i അംഗങ്ങളുണ്ടെന്നും അവയുടെ അഭിലക്ഷണമൂല്യങ്ങളുടെ സമാന്തരമാധ്യം $\bar{x}_i$ എന്നും ഇരിക്കട്ടെ.

$$\bar{x}_c = \frac{1}{n}\sum_{i=1}^{n} \bar{x}_i = \frac{1}{n}\sum_{i=1}^{n} \frac{X_i}{M_i} \qquad 7.2.9$$

എന്നു് നിർവചിച്ചിരിക്കുന്നു. തുല്യസംഘങ്ങൾ പരിഗണിച്ചപ്പോൾ അവലംബിച്ച രീതിയനുസരിച്ചു തന്നെ ഇതു് അനഭിനതമാണെന്നു കാണിക്കാൻ കഴിയും. അതു പോലെ തന്നെ.

$$V(\bar{x}_c) = \frac{N-n}{N-1}\frac{\sigma_b^2}{n} \qquad 7.2.10$$

എന്നും $$\hat{V}(\bar{x}_c) = \frac{N-n}{N-1}\frac{S_b^2}{n} \qquad 7.2.11$$

എന്നും തെളിയിക്കാവുന്നതാണു്.

ഇവിടെ എല്ലാ സംഘങ്ങൾക്കും തിരഞ്ഞെടുക്കപ്പെടുവാനുള്ള സംഭാവ്യത തുല്യമായിരിക്കത്തക്കവണ്ണമാണു് സാമ്പിളെടുത്തതു്. പക്ഷേ സംഘങ്ങൾ തുല്യ പരിമാണമുള്ളവയല്ലാത്തപ്പോൾ ഇതു് അത്ര നീതീകരിക്കത്തക്കതല്ല. ഒരു നിർദേശം ഓരോ സംഘവും തിരഞ്ഞെടുക്കപ്പെടുവാനുള്ള സംഭാവ്യത അതിന്റെ പരിമാണത്തിനു് ആനുപാതികമായിരിക്കണമെന്നതാണു്. അതായതു് i–ാമതു് സംഘം തിരഞ്ഞെടുക്കപ്പെടുവാനുള്ള സംഭാവ്യത $p_i = \frac{M_i}{\sum_{i=1}^{N} M_i}$ ആയിരിക്കണം.

ഇവിടെ $$\bar{x}_c = \frac{1}{n}\sum_{i=1}^{n} \bar{x}_i \qquad 7.2.12$$

$$E(\bar{x}_c) = \frac{1}{n}\sum_{i=1}^{n} E(\bar{x}_i) = \frac{1}{n}\cdot n\, E(\bar{x}_i) = \sum_{i=1}^{N} p_i \bar{x}_i$$

$$= \sum_{i=1}^{N} \bar{X}_i \frac{M_i}{\sum_{i=1}^{N} M_i} = \frac{\sum_{i=1}^{N} M_i \bar{X}_i}{\sum_{i=1}^{N} M_i} = \bar{X}$$

അതായതു് $\bar{x}_e$, $\bar{X}$ ൻെറ അനഭിനത ആകലമാണു്. മാത്രമല്ല,

$$V(\bar{x}_e) = \frac{1}{n} \frac{1}{\sum_{i=1}^{N} M_i} \sum_{i=1}^{N} (\bar{X}_i - \bar{X})^2 M_i \qquad 7.2.12$$

$$\hat{V}(\bar{x}_e) = \frac{1}{n} S_b^2 \qquad 7.2.13$$

എന്നും തെളിയിക്കാവുന്നതാണു്. ഈ രീതിയിൽ സാമ്പിളെടുക്കാൻ താഴെ കൊടുത്തിരിക്കുന്ന സമ്പ്രദായം അവലംബിക്കാവുന്നതാണു്.

പട്ടിക 7.1

സംഘസംഖ്യ	സംഘത്തിലെ അംഗങ്ങളുടെ എണ്ണം	സഞ്ചിതസംഘപരിമാണം
1	26	26
2	10	36
3	53	89

ഇതിൽ നിന്നു് ഒരു സംഘമാണു് തിരഞ്ഞെടുക്കേണ്ടതെന്നിരിക്കട്ടെ. രണ്ടു് അക്കമുള്ള ഒരു യാദൃച്ഛികസംഖ്യ എടുക്കുക. അതു് 89 ൽ കൂടുതലാണെങ്കിൽ അതു കളഞ്ഞു് മറെറാന്നെടുക്കുക. അങ്ങനെ 89 ൽ കുറവുള്ള ഒരു യാദൃച്ഛിക സംഖ്യ ലഭിച്ചാൽ അതു് 26 ഓ അതിൽ ചെറുതോ ആണെങ്കിൽ ആദ്യസംഘവും 26 നെക്കാൾ വലുതും 36 ഓ അതിൽ കുറവോ ആണെങ്കിൽ രണ്ടാമത്തെ സംഘവും അല്ലെങ്കിൽ മൂന്നാമത്തെ സംഘവും തിരഞ്ഞെടുക്കുക.

2. ബഹുഘട്ട സാമ്പിളനം

സംഘസാമ്പിളനത്തിൻെറ നിർവചനത്തോടു് അനുബന്ധിച്ചു് ബഹുഘട്ട സാമ്പിളനത്തെപ്പററി പരാമർശിക്കുകയുണ്ടായി. ഒരർഥത്തിൽ ബഹുഘട്ട സാമ്പിളനം സംഘസാമ്പിളനത്തിൻെറ സാമാന്യവൽക്കരണം തന്നെയാണു്. സംഘസാമ്പിളനത്തിൻെറ ചില പോരായ്മകൾ പരിഹരിക്കുവാനും അതിൻെറ വൈശിഷ്ട്യങ്ങൾ കഴിവതും നിലനിറുത്തുവാനും ഉദ്ദേശിച്ചു കൊണ്ടു് നിർദേശിക്കപ്പെട്ടിട്ടുള്ള ഒന്നാണു് ഇതു്.

സമഷ്ടിയെ ആദ്യം സംഘങ്ങളായി വിഭജിക്കുന്നു. ഇവയെ ആദ്യഘട്ട വ്യക്തികൾ എന്നും പറയുന്നു. അതിൽ നിന്നു് ഒരു സാമ്പിളെടുക്കുന്നു. ഇതിനു് ഇഷ്ട

മുള്ള സാമ്പിളനരീതി അവലംബിക്കാം. തിരഞ്ഞെടുക്കപ്പെട്ട സംഘങ്ങളെ ഉപസംഘങ്ങളായി വിഭജിക്കുന്നു. ഇവയ്ക്ക് രണ്ടാംഘട്ടവ്യക്തികൾ എന്നാണ് പേർ. ഇവയിൽ നിന്നാണ് അടുത്ത സാമ്പിൾ എടുക്കുന്നത്. തിരഞ്ഞെടുക്കപ്പെടുന്ന ഉപസംഘങ്ങളെ വീണ്ടും സംഘങ്ങളായി വിഭജിച്ചാൽ ലഭിക്കുന്നത് മൂന്നാംഘട്ടവ്യക്തികളാണ്. ഇവയിൽ നിന്നാണ് അടുത്ത സാമ്പിളനം. k ഘട്ടസാമ്പിളാണ് എടുക്കാൻ ഉദ്ദേശിക്കുന്നതെങ്കിൽ ഈ പ്രക്രിയ k—ാം ഘട്ടം വരെ തുടരുക. (k—1) –ാമതു് ഘട്ടസംഘങ്ങളിലെ പ്രാഥമികവ്യക്തികളിൽ നിന്നു് എടുക്കുന്ന സാമ്പിളാണ് വിവരശേഖരണത്തിനു് ഉപയോഗിക്കുക.

സാമ്പിളനഫ്രെയിമിന്റെ രൂപീകരണം ഉൾക്കൊള്ളുന്ന പണച്ചെലവു്, സമയനഷ്ടം തുടങ്ങിയവയും സാമ്പിളിലെ അംഗങ്ങൾ സമഷ്ടിയുടെ നാനാഭാഗത്തായി ചിതറിക്കിടക്കുമെന്നതുകൊണ്ടു് അവയെ കണ്ടുപിടിച്ച് വിവരം ശേഖരിക്കുന്നതിനുള്ള ബുദ്ധിമുട്ടുമാണല്ലോ. സംഘസാമ്പിളനരീതി സ്വീകരിക്കുവാൻ നമ്മെ പ്രേരിപ്പിച്ച പ്രധാനപരിഗണനകൾ. സമഷ്ടിയിലെ പ്രാഥമികവ്യക്തികളിൽ നിന്നു് എടുക്കുന്ന സാമ്പിളിനോളം ക്ഷമത സംഘസാമ്പിളിനു് ഉണ്ടായിരിക്കുകയില്ലെന്നും നാം കണ്ടു. ബഹുഘട്ടസാമ്പിളനത്തിന്റെ ക്ഷമത സാമ്പിളിലെ അംഗങ്ങൾ സമഷ്ടിയിൽ കുറെക്കൂടി ചിതറിക്കിടക്കുന്നതുകൊണ്ടു സംഘസാമ്പിളനത്തെക്കാൾ കൂടുതലായിരിക്കും. ഫ്രെം തയ്യാറാക്കൽ അത്ര ബുദ്ധിമുട്ടായിരിക്കുകയുമില്ല. ഓരോ ഘട്ടത്തിലും തിരഞ്ഞെടുക്കപ്പെടുന്ന സംഘങ്ങളുടെ ലിസ്റ്റ് തയ്യാറാക്കിയാൽ മതിയല്ലോ. അവസാനത്തേതിനു തൊട്ടു മുൻപുള്ള ഘട്ടത്തിൽ തിരഞ്ഞെടുക്കപ്പെടുന്ന ഉപസംഘങ്ങളിലെ പ്രാഥമികാംഗങ്ങളുടെ ലിസ്റ്റും തയ്യാറാക്കണം. സമഷ്ടിയിലെ പ്രാഥമികാംഗങ്ങളുടെ മുഴുവൻ പട്ടിക തയ്യാറാക്കുന്നതുമായി താരതമ്യപ്പെടുത്തുമ്പോൾ ഇതെത്ര നിസ്സാരമാണ്! സാമ്പിളിൽ ഉൾപ്പെടുത്തപ്പെടുന്ന അംഗങ്ങൾ അവസാനത്തേതിന്റെ മുൻപിലത്തെ ഘട്ടത്തിൽ തിരഞ്ഞെടുക്കപ്പെടുന്ന ഏതാനും ഉപസംഘങ്ങളിൽ മാത്രം ചിതറിക്കിടക്കുന്നവയായിരിക്കും. സമഷ്ടിയുടെ നാനാഭാഗങ്ങളിലായി ചിതറിക്കിടക്കുന്ന വ്യക്തികളിൽ നിന്നു് വിവരശേഖരണം നടത്തുന്നതുമായി തട്ടിച്ചുനോക്കുമ്പോൾ ഇതു് ഒട്ടും ആയാസകരമല്ല തന്നെ. അങ്ങനെ സാധാരണ സാമ്പിളനരീതികളെക്കാൾ പ്രായോഗികസൗകര്യമുള്ളതും ഗുച്ഛസാമ്പിളനത്തിനെക്കാൾ ക്ഷമതയുള്ളതുമായ ഒരു സാമ്പിളനരീതിയാണ് വിവിധഘട്ടസാമ്പിളനം.

പല സന്ദർഭങ്ങളിലും വിവിധഘട്ടസാമ്പിളനം അവലംബിക്കാൻ സാഹചര്യങ്ങളുടെ സമ്മർദം നമ്മെ നിർബന്ധിതരാക്കും. ഉദാഹരണത്തിനു് കേരളത്തിലെ ഒരു വീടിന്റെ ശരാശരി നിർമാണച്ചെലവാണ് അന്വേഷണ വിധേയമായ അഭിലക്ഷണമെന്നിരിക്കട്ടെ. കേരളത്തിലെ ഓരോ വീടും സമഷ്ടിയിലെ അംഗമാണ്. പക്ഷെ വീടുകളുടെ സമ്പൂർണമായൊരു പട്ടിക ലഭിക്കുക ഏതാണ്ട് അസാധ്യമത്രെ. ഇങ്ങനെ ഒരു സാഹചര്യത്തിൽ വിവിധ ഘട്ടസാമ്പിളനത്തിലേക്ക് അന്വേഷകൻ ദൃഷ്ടി തിരിക്കുന്നെങ്കിൽ അതിൽ അത്ഭുതപ്പെടുവാനൊന്നുമില്ല.

ജില്ലകളെ ആദ്യഘട്ട സംഘങ്ങളായി സ്വീകരിക്കാം. അവയുടെ പട്ടിക തയ്യാറാക്കാൻ പ്രയാസമില്ല. അതിൽ നിന്നു് ഒരു സാമ്പിളെടുക്കുന്നു. തിരഞ്ഞെടുക്കപ്പെടുന്ന ജില്ലകൾ മാത്രം ആദ്യഘട്ടത്തിൽ പരിഗണിച്ചാൽ മതി. താലൂക്കുകൾ രണ്ടാം ഘട്ടഉപസംഘങ്ങളായി എടുത്താൽ മുൻപറഞ്ഞ ജില്ലകളിലെ താലൂക്കുകളുടെ പട്ടിക തയ്യാറാക്കണം. അവയിൽ നിന്നു് ഒരു സാമ്പിളെടുക്കുന്നു. അങ്ങനെ തിരഞ്ഞെടുക്കപ്പെടുന്ന താലൂക്കുകളിലെ വില്ലേജുകളെ അടുത്ത ഘട്ടത്തിലെ ഉപസംഘങ്ങളായെടുക്കാം. അവയുടെ ലിസ്റ്റിൽ നിന്നു് മൂന്നാംഘട്ടസാമ്പിൾ എടുക്കുക. അങ്ങനെ ലഭിക്കുന്ന വില്ലേജുകളിലെ വീടുകളുടെ പട്ടിക തയ്യാറാക്കി അതിൽ നിന്നു് അവസാനഘട്ട സാമ്പിൾ എടുക്കണം. ഈ രീതി അവലംബിച്ചാൽ കേരളത്തിലുള്ള എല്ലാ വീടുകളുടെയും പട്ടിക തയ്യാറാക്കുക എന്ന ശ്രമകരമായ ജോലി ഒഴിവാക്കാം. പകരം തിരഞ്ഞെടുക്കപ്പെടുന്ന ഏതാനും വില്ലേജുകളിലെ വീടുകളുടെ ലിസ്റ്റു മാത്രം ഉണ്ടാക്കിയാൽ മതി. കൃഷിഭൂമികളുടെയും മററും സാമ്പിളെടുക്കുന്നിടത്തും വിവിധഘട്ട സാമ്പിളിനു് മുൻഗണന നൽകാൻ നാം പ്രേരിതരായിപ്പോവും. ബഹുഘട്ട സാമ്പിളനം തന്നെ പല തരത്തിലാവാം. ഘട്ടങ്ങളുടെ എണ്ണം, ഓരോ ഘട്ടത്തിലും സ്വീകരിക്കുന്ന സാമ്പിളനരീതി, ഓരോ ഘട്ടത്തിലെയും സാമ്പിളനവ്യക്തികളുടെ എണ്ണം തുടങ്ങിയവയാണു് ഈ വൈവിധ്യത്തിനു് ആധാരം. വിവിധഘട്ടസാമ്പിളനത്തെപ്പററി സമഗ്രമായ ഒരു ചർച്ച ഇവിടെ ഉദ്ദേശിക്കുന്നില്ല. മാതൃകയായി ഒരെണ്ണം മാത്രം പരിഗണിക്കാം.

4. ദ്വിഘട്ടസാമ്പിളനം—രണ്ടു ഘട്ടത്തിലും പ്രതിസ്ഥാപനമില്ലാത്ത ലഘ്യയാദൃച്ഛിക സാമ്പിളനം

സമഷ്ടിയിൽ N ആദ്യഘട്ടവ്യക്തികളാണു് ഉള്ളതെന്നിരിക്കട്ടെ. അതിൽ നിന്നു് n വ്യക്തികളെ പ്രതിസ്ഥാപനമില്ലാത്ത ലഘ്യയാദൃച്ഛിക സാമ്പിളനരീതിയിൽ തിരഞ്ഞെടുക്കുക. തിരഞ്ഞെടുക്കപ്പെട്ട i-ാമതു് വ്യക്തി M_i പ്രാഥമിക വ്യക്തികളെ ഉൾക്കൊള്ളുന്നു എന്നു സങ്കല്പിക്കുക. അതിൽ നിന്നു് മുൻസാമ്പിളനരീതിയനുസരിച്ചു് തന്നെ m_i അംഗങ്ങളെ തിരഞ്ഞെടുക്കുക. ഇങ്ങനെ ആദ്യഘട്ടത്തിൽ തിരഞ്ഞെടുക്കപ്പെട്ട എല്ലാ പ്രഥമഘട്ട വ്യക്തികളിൽ നിന്നുമായി തിരഞ്ഞെടുക്കപ്പെടുന്ന പ്രാഥമികാംഗങ്ങൾ ചേർന്നതാണു് സാമ്പിൾ. ഈ വ്യക്തികളിൽ നിന്നാണു് വിവരശേഖരണം നടത്തേണ്ടതു്. i-ാമതു് പ്രഥമഘട്ടവ്യക്തിയിൽ നിന്നു് തിരഞ്ഞെടുക്കപ്പെട്ട j-ാമതു് പ്രാഥമികവ്യക്തിയുടെ അഭിലക്ഷണമൂല്യം x_{ij} എന്നു സങ്കൽപിക്കാം.

ഒരുദാഹരണം കൊണ്ടു് ഈ രീതി വിശദീകരിക്കാം. കേരളത്തിലെ വീടുകളുടെ ശരാശരി വിസ്തീർണമാണു് പരിഗണിക്കപ്പെടുന്ന അഭിലക്ഷണമെന്നിരിക്കട്ടെ. വില്ലേജുകളാണു് ആദ്യഘട്ട വ്യക്തികൾ എന്നു സങ്കൽപിക്കുക. ഓരോ വില്ലേജിലുമുള്ള വീടുകളാണു് ദ്വിതീയഘട്ടവ്യക്തികൾ. ഈ അന്വേഷണത്തെ സംബന്ധിച്ചിടത്തോളം അവ പ്രാഥമികവ്യക്തികളായി പരിഗണി

ക്കപ്പെട്ടും. വില്ലേജുകളുടെ എണ്ണം N. ഇതിൽ നിന്നു വില്ലേജുകളെ പ്രതിസ്ഥാപനമില്ലാത്ത ലഘുയാദൃച്ഛികസാമ്പിളനരീതിയിൽ തിരഞ്ഞെടുക്കുന്നു. തിരഞ്ഞെടുക്കപ്പെട്ട i–ാമതു വില്ലേജിൽ M_i വീട്ടുകൾ ഉണ്ടെന്നു സങ്കൽപം. ഈ വീട്ടുകളിൽ m_i എണ്ണത്തെ പ്രതിസ്ഥാപനമില്ലാത്ത ലഘുയാദൃച്ഛിക സാമ്പിളനരീതിയിൽ തന്നെ തിരഞ്ഞെടുക്കുക. അതിൽ j-ാമതു് വീടിന്റെ വിസ്തീർണം x_{ij}. തിരഞ്ഞെടുക്കപ്പെട്ട എല്ലാ വില്ലേജുകളിൽ നിന്നുമായി $m_1+m_2\cdots+m_n=m$ വീട്ടുകൾ ലഭിക്കുന്നു. ഇതാണു് സാമ്പിൾ. അവ ഓരോന്നിന്റെയും വിസ്തീർണം കണ്ടു പിടിക്കുക.

സമഷ്ടിയിലെ അംഗങ്ങളുടെ അഭിലക്ഷണമൂല്യങ്ങളുടെ സമാന്തരമാധ്യം $\bar{X}$ എന്നിരിക്കട്ടെ. അതാണു് ആകലനം ചെയ്യേണ്ട പ്രാചലം.

$$\bar{X}_m = \frac{1}{n}\sum_{i=1}^{n}\frac{1}{m_i}\sum_{j=1}^{m_i} x_{ij} \qquad 7.4.1$$

എന്നതു് $\bar{X}$ ന്റെ ഒരു അഭിനതആകലമാണെന്നു കാണാം.

$$\left[\bar{X} = \frac{\sum_{i=1}^{N} M_i\bar{X}_i}{\sum_{i=1}^{N} M_i}\right]$$

ഇതു് അഭിനതമാണെന്നു കാണിക്കാൻ ശ്രമിക്കുമ്പോൾ ഒരു കാര്യ പ്രത്യേകം ഓർമയിരിക്കണം. സാമ്പിളനത്തിനു് രണ്ടു ഘട്ടമുണ്ടു് എന്ന വസ്തുത.

$$E(\bar{x}_M) = E_1\left[\frac{1}{n}\sum_{i=1}^{n} F_2\left(\frac{1}{m_i}\sum_{j=1}^{m_i} x_i\right)\right]$$

ഇവിടെ ബ്രാക്കറ്റിനകത്തെ E_2 നെ രണ്ടാം ഘട്ടത്തിലെ സാമ്പിളനത്തോടു ബന്ധപ്പെടുത്തി വേണം നിർവചിക്കുവാൻ. മറെറാരു തരത്തിൽ പറഞ്ഞാൽ, അതു് i നിശ്ചിതമാണു് എന്ന സങ്കൽപത്തിലുള്ള പ്രതീക്ഷയെ കുറിക്കുന്നു. ബ്രാക്കററിനു പുറത്തെ F_1, ആദ്യഘട്ട സാമ്പിളനത്തോടു ബന്ധപ്പെട്ട പ്രതീക്ഷയാണു്. (പ്രസക്തമായ, സാംഖ്യിക പ്രമേയങ്ങൾ അനുസ്മരിക്കുക.)

$$E_2\left(\frac{1}{m_i}\sum_{j=1}^{m_i} x_{ij}\right) = \frac{1}{m_i}\sum_{j=1}^{m_i} E_2(x_{ij}) = \frac{1}{M_i}\sum_{j=1}^{M_i} x_{ij} = \bar{X}_i$$

ഇവിടെ $\bar{X}_i$, i–ാമതു് ആദ്യഘട്ടവ്യക്തിയിലെ അംഗങ്ങളുടെ അഭിലക്ഷണമൂല്യങ്ങളുടെ സമാന്തരമാധ്യമാണു്. i–ാമതു് ആദ്യഘട്ടവ്യക്തിയാണു് സമഷ്ടിയെന്നു് സങ്കൽപിക്കുക. പ്രതിസ്ഥാപനമില്ലാത്ത ലഘുയാദൃച്ഛിക സാമ്പിളന

ത്തിൽ സാമ്പിൾ സമാന്തരമാധ്യത്തിൻെറ പ്രതീക്ഷ, സമഷ്ടിയുടെ സമാന്തരമാധ്യമാണെന്നു് മൂന്നാമധ്യായത്തിൽ തെളിയിച്ചിട്ടുണ്ടല്ലോ.

$$E_i\left(\frac{1}{n}\sum_{i=1}^{n}\bar{X}_i\right)=\frac{1}{n}\sum_{i=1}^{n}E_i(\bar{x}_i)=\frac{\sum_{i=1}^{N}\bar{X}_i}{N}\neq\bar{X}$$

ആയതിനാൽ $\bar{x}_M$, $\bar{X}$ ൻെറ അനഭിനതആകലമല്ല. ഇതിൻെറ അഭിനതി $[E(\bar{x}_M)-\bar{X}]$ ആകുന്നു. പ്രഥമഘട്ടവ്യക്തികളുടെ പരിമാണം സമമാകുമ്പോൾ മാത്രം $\bar{x}_M$, $\bar{X}$ ൻെറ അനഭിനത ആകലമാകും.

$E(\bar{x}_M)=X_a$ എന്നു കുറിച്ചാൽ,

$\bar{x}_M$ ന്റെ പ്രസരണം $V(\bar{x}_M)=E(x_M-\bar{X}_a)^2$

$$=\left(\frac{1}{n}-\frac{1}{N}\right)S_b^2+\frac{1}{nN}\sum_{i=1}^{N}\left(\frac{1}{m_i}-\frac{1}{M_i}\right)S_i^2 \qquad 7.4.2$$

എന്നു് പ്രസക്തമായ സാംഖ്യികനിയമങ്ങൾ ഉപയോഗിച്ച് കാണാം.

ഇവിടെ, $S_b^2=\frac{1}{N-1}\sum_{i=1}^{N}(\bar{X}_i-\bar{X}_a)^2$

$$S_i^2=\frac{1}{M_i-1}\sum_{j=1}^{M_i}(x_{ij}-\bar{X}_i)^2$$ ഉം ആകുന്നു.

പ്രഥമഘട്ടവ്യക്തികളുടെ പരിമാണങ്ങൾ (M_i) വ്യത്യസ്തങ്ങളായിരിക്കുമ്പോൾ $\bar{x}_M$ നു പകരം, താഴെ കൊടുത്തിട്ടുള്ള വേറെ രണ്ടു ആകലങ്ങൾ ഉപയോഗിക്കാറുണ്ട്.

$$\text{(i)}\ \bar{x}_M'=\frac{1}{N\bar{M}}\sum_{i=1}^{n}M_i\bar{x}_i \qquad \text{(ii)}\ \bar{x}_M''=\frac{\sum_{i=1}^{n}M_i\bar{x}_i}{\sum_{i=1}^{n}M_i}$$

$$\left[\bar{M}=\frac{1}{N}\sum_{i=1}^{N}M_i\right]$$

ഇതിൽ $\bar{x}_M'$, $\bar{X}$ ൻെറ ഒരു അനഭിനത ആകലവും, $\bar{x}_M''$ ഒരു അഭിനതആകലവുമാണു്. $\bar{x}_M''$ ഒരു അനുപാതആകലമായി കാണാവുന്നതു കൊണ്ടു്, ഈ ആകലത്തിൻെറ പ്രസരണം മററു രണ്ടു് ആകലങ്ങളുടെ പ്രസരണങ്ങളെക്കാൾ കുറവായിരിക്കുമെന്നു് തെളിയിക്കാവുന്നതാണു്.

മുമ്പ് ഉപയോഗിച്ചിട്ടുള്ള സാംഖ്യിക നിയമപ്രകാരം,

$$V(\bar{x}_M') = \left(\frac{1}{n} - \frac{1}{N}\right) S_b'^2 + \frac{1}{nN} \sum_{i=1}^{N} \left(\frac{1}{n_i} - \frac{1}{M_i}\right) S_i^2 \qquad 7.4.3$$

എന്നും

$$V(\bar{x}_M'') = \left(\frac{1}{n} - \frac{1}{N}\right) S_b''^2 + \frac{1}{nN} \Sigma \frac{M_i^2}{\bar{M}^2} \left(\frac{1}{n_i} - \frac{1}{M_i}\right) S_i^2 \qquad 7.4.4$$

എന്നും സിദ്ധിക്കും.

ഇവിടെ, $S_b'^2 = \frac{1}{N-1} \sum_{i=1}^{N} \frac{M_i}{\bar{M}} (\bar{X}_i - \bar{X})^2$ – ഉം

$$S_b''^2 = \frac{1}{N-1} \sum_{i=1}^{N} \frac{M_i^2}{\bar{M}^2} (\bar{X}_i - \bar{X})^2$$ – ഉം ആകുന്നു.

എല്ലാ ആദ്യഘട്ടവ്യക്തികളിലെയും അംഗസംഖ്യ തുല്യമാണെന്നിരിക്കട്ടെ. (M എന്നു വിചാരിക്കുക) അപ്പോൾ സ്വാഭാവികമായും തുല്യസംഖ്യ ദ്വിതീയഘട്ടവ്യക്തികളെയായിരിക്കും ഓരോ ആദ്യഘട്ടവ്യക്തിയിൽ നിന്നും സാമ്പിളിലേക്കു തിരഞ്ഞെടുക്കുക. (m ആണ് ഈ സംഖ്യ എന്നു വിചാരിക്കുക). അങ്ങനെ വന്നാൽ,

$E(\bar{X}_M) = \bar{X}$, അതായതു് $\bar{X}_M$, $\bar{X}$ ൻെറ ഒരു അനഭിനത ആകലമാണെന്നും

$$V(\bar{x}_M) = \left(\frac{1}{n} - \frac{1}{N}\right) S_b^2 + \left(\frac{1}{m} - \frac{1}{M}\right) \frac{\bar{S}_w^2}{n} \qquad 7.4.5$$

എന്നും സിദ്ധിക്കും. ഈ വ്യംജകത്തിൽ

$\bar{S}_w^2 = \frac{1}{N} \sum_{i=1}^{N} S_i^2$, $S_b^2 = \frac{1}{N-1} \Sigma (\bar{X}_i - \bar{X})^2$ എന്നാകും.

N, M എന്നിവ വലുതാണെന്നുള്ള സങ്കൽപം ഉപയോഗപ്പെടുത്തിയാൽ,

$$V(\bar{x}_M) = \frac{S_b^2}{n} + \frac{\bar{S}_w^2}{nm} \qquad 7.4.6$$

എന്നു സിദ്ധിക്കും.

$V(\bar{x}_M)$ ൻെറ ആകലനവും പ്രായോഗികപരിഗണനകൾ വെച്ചു നോക്കുമ്പോൾ സുപ്രധാനമാണ്.

$$\bar{x}_i = \frac{1}{m_i}\sum_{j=1}^{m_i} x_{ij},\ \bar{x} = \frac{1}{n}\sum_{i=1}^{n} \bar{x}_i$$

എന്നിരിക്കട്ടെ. ഇതിൽ $\bar{x} = \bar{x}_M$ എന്നു കാണുവാൻ വിഷമമില്ല.

$$s_i^2 = \frac{1}{m_i - 1}\sum_{j=1}^{m_i} (x_{ij} - \bar{x}_i)^2 \text{ എന്നും, } s_b^2 = \frac{1}{n-1}\sum_{i=1}^{n} (\bar{x}_i - \bar{x})^2$$

എന്നും സങ്കൽപിക്കുക. 7.4.3 ൽ കൊട്ടത്തിരിക്കുന്ന $V(\bar{x}_M)$ ന്റെ വ്യംജകത്തിൽ S_i^2 ൻെറ സ്ഥാനത്തു് s_i^2 ഉം, S_b^2 ൻെറ സ്ഥാനത്തു് s_b^2 ഉം പ്രതിസ്ഥാപിക്കുകയും $i = 1 \ldots n$ എന്നാക്കുകയും ചെയ്താൽ $V(\bar{x}_M)$ ൻെറ ഒരു ആകലം ലഭിക്കും. അതായതു്,

$$\hat{V}(\bar{x}_M) = \left(\frac{1}{n} - \frac{1}{N}\right) s_b^2 + \frac{1}{nN}\sum_{i=1}^{n}\left(\frac{1}{m_i} - \frac{1}{M_i}\right) s_i^2 \qquad 7.4.7$$

ഈ ആകലം അനഭിനതമല്ല. പക്ഷെ ആദ്യഘട്ട വ്യക്തികളെല്ലാം തുല്യഎണ്ണം പ്രാഥമികവ്യക്തികളെ ഉൾക്കൊള്ളുന്നവയാണെങ്കിൽ ഇതു് അനഭിനതമാവും.

$M_i = M$ ആകുമ്പോൾ

$$\hat{V}(\bar{x}_M) = \left(\frac{1}{n} - \frac{1}{N}\right) s_b^2 + \left(\frac{1}{m} - \frac{1}{M}\right) \bar{S}_w^2 \qquad 7.4.8$$

അംഗസംഖ്യയിലുള്ള വ്യത്യാസം, അതായതു് M_i കൾ തമ്മിലുള്ള വ്യത്യാസം വർധിക്കുന്തോറും അഭിനതിയും വർധിക്കും.

$V(\bar{x}_M)$ ൻെറ വ്യംജകം ശ്രദ്ധിച്ചു പരിശോധിച്ചു നോക്കുക. സൗകര്യത്തിനു വേണ്ടി 7.4.5 ലെ, ലഘൂകൃത വ്യംജകം തന്നെ എടുക്കാം. ഇതിനു് രണ്ടു ഭാഗങ്ങളുണ്ടു്. ആദ്യഭാഗം S_b^2 ൻെറയും രണ്ടാം ഭാഗം $\bar{S}_w^2$ ൻെറ യും ഗുണിതങ്ങളാണു്. S_b^2 ൻെറ നിർവചനം പരിഗണിച്ചാൽ അതു് ആദ്യഘട്ടവ്യക്തികളുടെ പ്രകീർണനത്തെ പ്രതിനിധീകരിക്കുന്നു എന്നു കാണാം. അതുപോലെ തന്നെ S_w^2 ആദ്യഘട്ടവ്യക്തികൾക്കുള്ളിലുള്ള പ്രകീർണനത്തെയും സൂചിപ്പിക്കുന്നു. $m = M$ ആണെങ്കിൽ, അതായതു് തിരഞ്ഞെടുക്കപ്പെടുന്ന ആദ്യഘട്ടവ്യക്തികളിലെ എല്ലാ പ്രാഥമികാംഗങ്ങളെയും സാമ്പിളിൽ ഉൾപ്പെടുത്തുകയാണെങ്കിൽ, ഇതിൽ രണ്ടാമത്തെ ഭാഗം അപ്രത്യക്ഷമാവും. അങ്ങനെ $\bar{x}_M$ ന്റെ സാമ്പിളനപ്രസരണത്തിന്റെ രണ്ടാമത്തെ ഭാഗം രണ്ടാംഘട്ട സാമ്പിളനത്തിൽ നിന്നു് ഉണ്ടാവുന്നതാണു്.

ഇവിടെ സാമ്പിളിൽ ആകെ m അംഗങ്ങളാണു് ഉണ്ടായിരിക്കുക. അവ തിരഞ്ഞെടുക്കപ്പെട്ട n പ്രാഥമികാംഗങ്ങളിൽ നിന്നാണു് തിരഞ്ഞെടുക്കപ്പെടുന്നതു്. ഇവിടെ ഒരു പ്രശ്നം ഉത്ഭവിക്കുന്നു. $n, m_1, m_2 \ldots\ldots m_n$ എന്നീ സംഖ്യ

കൾ എങ്ങനെയാണ് നിർണയിക്കുക? അഭിലഷണീയമെന്നു തോന്നുന്ന ഏതെങ്കിലും അടിസ്ഥാന നിബന്ധനയെ ആധാരമാക്കി ഇതു് നിർവഹിക്കാവുന്നതാണ്. സാധാരണ ചെയ്യാറുള്ളതു്, ആകെ ചെലവു സൂചിപ്പിക്കുന്ന വ്യഞ്ജകത്തിൻെറ മൂല്യം നിശ്ചിതമായിരിക്കണം എന്ന നിബന്ധനക്കു വിധേയമായി $V(\bar{x}_M)$ നിമ്നതമമാക്കുന്ന മൂല്യങ്ങൾ അവയ്ക്കു നൽകുക എന്നതാണ്.

മുൻപു സൂചിപ്പിച്ചതുപോലെ ചെലവു്, പ്രായോഗിക സൗകര്യം തുടങ്ങിയ പരിഗണനകൾ വെച്ചുനോക്കുമ്പോൾ സംഘസാമ്പിളനത്തോളം സൗകര്യപ്രദമല്ലെങ്കിലും സമഷ്ടിയിൽ നിന്നു് നേരിട്ടുള്ള സാമ്പിളനത്തെക്കാൾ വളരെയേറെ മെച്ചപ്പെട്ട ഒരു സാമ്പിളനസമ്പ്രദായമാണ് ഇതു്. സാമ്പിളനപ്രസരണത്തെ അടിസ്ഥാനമാക്കി വിലയിരുത്തിയാൽ സംഘസാമ്പിളനത്തെക്കാൾ വിശിഷ്ടവും, നേരിട്ടുള്ള സാമ്പിളനത്തെക്കാൾ മോശപ്പെട്ടതുമാണ് ഇതു്. ചുരുക്കത്തിൽ, സംഘസാമ്പിളനവും നേരിട്ടുള്ള സാമ്പിളനവുമായുള്ള ഒരു ഒത്തുതീർപ്പായി ഇതിനെ പരിഗണിക്കാം.

5. പരിമാണാനുപാതിക സംഭാവ്യതാസാമ്പിളനം

ലഘുയാദൃച്ഛികസാമ്പിളനം തുടങ്ങി നാം പരിഗണിച്ച പ്രധാനപ്പെട്ട മിക്ക സാമ്പിളനരീതികളിലും സമഷ്ടിയിലെ എല്ലാ അംഗങ്ങൾക്കും സാമ്പിളിൽ ഉൾപ്പെടുവാനുള്ള സംഭാവ്യത തുല്യമാണ്. പക്ഷെ, സമഷ്ടിയിലെ ഓരോ അംഗത്തിനും സാമ്പിളിൽ ഉൾപ്പെടുവാനുള്ള സംഭാവ്യത ഓരോന്നായിരിക്കത്തക്കവണ്ണം എടുക്കുന്ന സാമ്പിളുകൾ പലപ്പോഴും കൂടുതൽ മെച്ചപ്പെട്ട ആകലങ്ങൾ നൽകാറുണ്ട്. ഇമ്മാതിരി ഒരു സാമ്പിളന രീതിയാണ് പരിമാണാനുപാതികസംഭാവ്യതാസാമ്പിളനം. ഓരോ അംഗത്തിൻെറയും പരിമാണത്തിനു് ആനുപാതികമായിരിക്കണം അതിനു് സാമ്പിളിൽ ഉൾപ്പെടുവാനുള്ള സംഭാവ്യത എന്നതാണ് ഈ സാമ്പിളനരീതിയുടെ അടിസ്ഥാനപ്രമാണം. പരിമാണം നിർണയിക്കുന്നതു് പരിഗണിക്കപ്പെടുന്ന അഭിലക്ഷണവുമായി നേർസഹബന്ധമുള്ള ഒരു സഹായക അഭിലക്ഷണം ഉപയോഗപ്പെടുത്തിയാണ്. ഉദാഹരണമായി കേരളത്തിലെ ഒരു കലാലയത്തിൻെറ ശരാശരി വാർഷികച്ചെലവാണ് അന്വേഷണവിധേയമായ അഭിലക്ഷണമെന്നിരിക്കട്ടെ. ഇവിടെ കലാലയങ്ങളാണ് വ്യക്തികൾ. അവയുടെ സമുച്ചയമാണ് സമഷ്ടി. ഓരോ കലാലയത്തിന്റെയും വാർഷികച്ചെലവാണ് പരിഗണിക്കപ്പെടുന്ന അഭിലക്ഷണം. വിദ്യാർഥികളുടെ എണ്ണവുമായി ശക്തിയായ നേർസഹബന്ധമുള്ള ഒരു അഭിലക്ഷണമാണല്ലോ ഇതു്. അതുകൊണ്ടു് വിദ്യാർഥികളുടെ എണ്ണം കലാലയങ്ങളുടെ പരിമാണം നിർണയിക്കാൻ ഉപകരിക്കുന്ന ഒരു സഹായക അഭിലക്ഷണമാണ്. ഓരോ കലാലയത്തിനും സാമ്പിളിൽ ഉൾപ്പെടുവാനുള്ള സംഭാവ്യത അവിടെ അധ്യയനം ചെയ്യുന്ന വിദ്യാർഥികളുടെ എണ്ണത്തിനു് ആനുപാതികമായിരിക്കത്തക്കവണ്ണം എടുക്കുന്ന സാമ്പിൾ ഒരു പരിമാണാനുപാതിക സംഭാവ്യതാസാമ്പിളാണ്.

ഇങ്ങനെ എടുക്കുന്ന സാമ്പിളിനെ ആധാരമാക്കി ആകലനം നിർവഹിക്കുമ്പോൾ ഒരു കാര്യം പ്രത്യേകം ശ്രദ്ധിക്കേണ്ടതുണ്ട്. സമഷ്ടിയിലെ അഭിലക്ഷണ മൂല്യങ്ങളുടെ സമാന്തരമാധ്യമാണ് ആകലനം ചെയ്യേണ്ട അഭിലക്ഷണമെന്നിരിക്കട്ടെ. പതിവു പോലെ സാമ്പിൾ സമാന്തരമാധ്യം ആകലമായി സ്വീകരിച്ചാൽ അതു് അനഭിനതമായിരിക്കാൻ ഇടയില്ല. കാരണം, സമഷ്ടിയിലെ എല്ലാ വ്യക്തികൾക്കും സാമ്പിളിൽ ഉൾപ്പെടുവാനുള്ള സംഭാവ്യത തുല്യമല്ല. പ്രായോഗിക ആവശ്യങ്ങൾക്കു് അനഭിനത ആകലം ആവശ്യമാണുതാനും. അതുകൊണ്ടു് സമുചിതമായ ഒരു ഭാരിത സമാന്തരമാധ്യം വേണം ആകലമായി സ്വീകരിക്കുവാൻ.

6. പരിമാണാനുപാതിക സംഭാവ്യതാ സാമ്പിളന സമ്പ്രദായങ്ങൾ

സമഷ്ടിയിലെ അംഗങ്ങളുടെ പട്ടികയും ഓരോ വ്യക്തിയുടെയും സഹായക അഭിലക്ഷണത്തിൻെറ മൂല്യവും നൽകപ്പെട്ടിരിക്കുന്നു എന്നിരിക്കട്ടെ. മുൻഉദാഹരണത്തിലാണെങ്കിൽ കലാലയങ്ങളുടെ പട്ടികയും ഓരോ കലാലത്തിലും അധ്യയനം നടത്തുന്ന വിദ്യാർഥികളുടെ എണ്ണവും തന്നിരിക്കുന്നു. എങ്ങനെയാണു് n അംഗങ്ങളുള്ള ഒരു പരിമാണാനുപാതിക സാമ്പിൾ എടുക്കുക എന്നതാണു് പ്രശ്നം. ഇതിനായി പല സമ്പ്രദായങ്ങളും നിർദേശിക്കപ്പെട്ടിട്ടുണ്ടു്. അവയിൽ ചിലതു് താഴെ കൊടുക്കുന്നു.

1. സഞ്ചിത സംയുക്ത സമ്പ്രദായം

വളരെ ലളിതമായ ഒരു രീതിയാണു്. ഇതു്. സമഷ്ടിയിൽ N അംഗങ്ങളാണുള്ളതെന്നും അവയുടെ സഹായക അഭിലക്ഷണ മൂല്യങ്ങൾ $y_1, y_2 \cdots y_N$ എന്നിവയാണു് എന്നും ഇരിക്കട്ടെ. താഴെ കൊടുത്തിരിക്കുന്നതു പോലെയുള്ള ഒരു പട്ടിക തയ്യാറാക്കുക. ഇതിൽ ആദ്യത്തെ കോളം സമഷ്ടിയിലെ അംശങ്ങളുടെ ക്രമസംഖ്യകളും രണ്ടാമത്തേതു് സഹായക അഭിലക്ഷണത്തിൻെറ മൂല്യങ്ങളും മൂന്നാമത്തേതു് സഹായക അഭിലക്ഷണ മൂല്യങ്ങളുടെ സഞ്ചിത സംയുക്തങ്ങളുമത്രെ

പട്ടിക 7.1

ക്രമസംഖ്യ	സഹായകഅഭിലക്ഷണ മൂല്യങ്ങൾ	സഞ്ചിത സംയുക്തങ്ങൾ
1	y_1	y_1
2	y_2	y_1+y_2
3	y_3	$y_1+y_2+y_3$
⋮	⋮	⋮
N	y_N	$y_1+y_2+\cdots+y_N=Y$
	Y	

ഇവിടെ സഹായകഅഭിലക്ഷണമൂല്യങ്ങളുടെ തുക Y യാണു്. യാദൃച്ഛിക സംഖ്യാ പട്ടികയിലെ ഒരു പേജെടുത്തു് ക്രമത്തിൽ താഴോട്ടോ കുറുകെയോ

Y യിൽ എത്ര അക്കങ്ങളാണോ ഉള്ളതു് അത്രയും അക്കങ്ങളുള്ള യാദൃച്ഛിക സംഖ്യ വായിക്കുക. Y യെക്കാൾ വലിയ സംഖ്യകൾ അവഗണിക്കുക. സ്വീകരിക്കാവുന്ന ആദ്യത്തെ സംഖ്യ m ആണെന്നിരിക്കട്ടെ. അതു് (i—1) ൻെറ നേരെയുള്ള സഞ്ചിതസംയുക്തത്തെക്കാൾ വലുതും i-യുടെ നേരെയുള്ള സംയുക്തത്തോടു തുല്യമോ അതിൽ കുറവോ ആണെന്നും ഇരിക്കട്ടെ. അങ്ങനെ വന്നാൽ സമഷ്ടിയിലെ i ക്രമസംഖ്യയായുള്ള അംഗത്തെ സാമ്പിളിൽ ഉൾപ്പെടുത്താം. ഈ പരിപാടി തുടന്നാൽ സാമ്പിളിൽ ഉൾപ്പെടുത്തേണ്ട n അംഗങ്ങളെ തിരഞ്ഞെടുക്കാവുന്നതാണു്. പ്രതിസ്ഥാപനമില്ലാത്ത സാമ്പിളാണു് എടുക്കുന്നതെങ്കിൽ തിരഞ്ഞെടുക്കപ്പെട്ടു കഴിഞ്ഞ ഏതെങ്കിലും അംഗത്തെ സൂചിപ്പിക്കുന്ന യാദൃച്ഛികസംഖ്യകൾ ഓരോ ഘട്ടത്തിലും അവഗണിക്കണം. മാതൃകയായി കലാലയങ്ങളുടെ ഉദാഹരണം തന്നെ എടുക്കാം. അഞ്ചു് കലാലയങ്ങളാണുള്ളതെന്നും രണ്ടു് അംഗങ്ങളുള്ള പ്രതിസ്ഥാപനമില്ലാത്ത സാമ്പിളാണു് എടുക്കേണ്ടതെന്നും ഇരിക്കട്ടെ.

പട്ടിക 7.2

കോളേജുകൾ	വിദ്യാർത്ഥികളുടെ എണ്ണം	സഞ്ചിതസംയുക്തങ്ങൾ
1	2340	2340
2	1868	4208
3	1012	5220
4	520	5740
5	441	6181
	6181	

നാലു് അക്കങ്ങളുള്ള യാദൃച്ഛിക സംഖ്യകളാണു് എടുക്കേണ്ടതു്. 6181 നെക്കാൾ വലിയ സംഖ്യകൾ വന്നാൽ അവ അവഗണിക്കുക. ആദ്യം ലഭിച്ച സ്വീകാര്യങ്ങളായ രണ്ടു സംഖ്യകൾ 4436 ഉം 5807 ഉം ആണെന്നിരിക്കട്ടെ. 4436, 4208 നും 5220നും മദ്ധ്യേയാണു്. അതുകൊണ്ടു് 5220 നു് നേരെയുള്ള 3 ക്രമസംഖ്യയായുള്ള കോളേജു് ആണു് സാമ്പിളിലെ ആദ്യത്തെ അംഗം. 5807, 5 ക്രമസംഖ്യയായുള്ള കലാലയത്തെയും സൂചിപ്പിക്കുന്നു. അങ്ങനെ 3, 5 എന്നീ ക്രമസംഖ്യകളോടു കൂടിയ കലാലയങ്ങളെയാണു് സാമ്പിളിൽ ഉൾപ്പെടുത്തേണ്ടതു്.

ഈ സമ്പ്രദായത്തിൻെറ പ്രധാന ന്യൂനത സഞ്ചിത സംയുക്തങ്ങളുടെ പട്ടിക തയ്യാറാക്കണമെന്നുള്ളതാണു്. സമഷ്ടി വളരെ വലുതാകുമ്പോൾ ഇതു് അത്യന്തം വിഷമകരമായ ഒരു ജോലിയായി തീരുന്നു. അതിനു വേണ്ടി വരുന്ന ചെലവും സമയനഷ്ടവും ഗണനീയമായിരിക്കുമല്ലോ. ഈ തകരാറു് ഒരു പരിധിവരെ ലഘൂകരിക്കുന്ന മറെറാരു സമ്പ്രദായമാണു് 'ലാഹിരി സമ്പ്രദായം'

2. ലാഹിരി സമ്പ്രദായം

ഈ സമ്പ്രദായത്തിൽ രണ്ടു യാദൃച്ഛിക സംഖ്യകൾ ഉപയോഗിച്ചാണു് തിരഞ്ഞെടുപ്പു നടത്തുന്നതു്. സമഷ്ടിയിൽ N അംഗങ്ങളാണുള്ളതെന്നും അവയുടെ

പരിമാണങ്ങൾ സൂചിപ്പിക്കുന്ന അഭിലക്ഷണ മൂല്യങ്ങളിൽ ഏററവും വലുതു് M ആണെന്നും ഇരിക്കട്ടെ. ആദ്യമായി 1 മുതൽ N വരെയുള്ള സംഖ്യകളിൽ നിന്നു് യാദൃച്ഛികമായി ഒന്നു് തിരഞ്ഞെടുക്കുക. l ആണു് തിരഞ്ഞെടുക്കപ്പെട്ട സംഖ്യ എന്നിരിക്കട്ടെ. അടുത്തതായി 1 മുതൽ M വരെയുള്ള സംഖ്യകളിൽ നിന്നു് യാദൃച്ഛികമായി ഒരു സംഖ്യയെടുക്കുക. m ആണു് തിരഞ്ഞെടുക്കപ്പെട്ടതു് എന്നു സങ്കൽപ്പിക്കുക. l ക്രമസംഖ്യയായുള്ള അംഗത്തിൻെറ സഹായക അഭിലക്ഷണമൂല്യം y_l ആണെന്നിരിക്കട്ടെ. m, y_lനെക്കാൾ കുറവോ തുല്യമോ ആണെങ്കിൽ l ക്രമസംഖ്യയായുള്ള അംഗത്തെ സാമ്പിളിൽ ഉൾപ്പെടുത്തുക. അല്ലെങ്കിൽ ഈ രണ്ടു സംഖ്യകളും അവഗണിച്ചിട്ടു് ഇതുപോലെ തന്നെ പുതിയ രണ്ടു സംഖ്യകൾ എടുക്കുക. ആവശ്യമുള്ളത്ര അംഗങ്ങൾ തിരഞ്ഞെടുക്കപ്പെടുന്നതുവരെ ഈ പരിപാടി തുടരുക.

പട്ടിക 7.2 ലെ സമഷ്ടിയെ തന്നെ ഉദാഹരണമായി പരിഗണിക്കാം. സമഷ്ടിയിൽ 5 അംഗങ്ങളാണുള്ളതു്. അതുകൊണ്ടു് ആദ്യമായി 1 മുതൽ 5 വരെയുള്ള സംഖ്യകളിൽ നിന്നു് യാദൃച്ഛികമായി ഒന്നു് തിരഞ്ഞെടുക്കണം. ഒററ അക്കമുള്ള യാദൃച്ഛിക സംഖ്യാപട്ടിക ഉപയോഗിച്ചു് ഇതു ചെയ്യാവുന്നതാണു്. 3 ആണു് തിരഞ്ഞെടുക്കപ്പെട്ട സംഖ്യ എന്നിരിക്കട്ടെ. സഹായക അഭിലക്ഷണമൂല്യങ്ങളിൽ ഏററവും വലുതു് 2340 ആണു്. അപ്പോൾ 1 മുതൽ 2340 വരെയുള്ള സംഖ്യകളിൽ നിന്നു് ഒന്നു് യാദൃച്ഛികമായി തിരഞ്ഞെടുക്കണം. നാലക്കമുള്ള യാദൃച്ഛിക സംഖ്യാപട്ടിക ഇതിനായി ഉപയോഗിക്കാം. 2123 ആണു് ഇങ്ങനെ ലഭിച്ച സംഖ്യയെന്നു സങ്കൽപ്പിക്കുക. 3 ക്രമസംഖ്യയായുള്ള കലാലയത്തിലെ വിദ്യാർഥികളുടെ എണ്ണം 1012 ആണു്. 2123 ഇതിനെക്കാൾ ചെറുതല്ല. അതുകൊണ്ടു് 3 ക്രമസംഖ്യയായുള്ള കലാലയത്തെ സാമ്പിളിൽ ഉൾപ്പെടുത്താൻ നിർവാഹമില്ല. മുഴുവൻ പരിപാടിയും ആവർത്തിക്കുക. 2, 1148 എന്നീ സംഖ്യകളാണു് അടുത്തതായി കിട്ടുന്നതെന്നിരിക്കട്ടെ. 2 ക്രമസംഖ്യകളായുള്ള കോളേജിൽ 1868 വിദ്യാർഥികളുണ്ടു്. 1148 അതിൽ കുറവാണു്. അതുകൊണ്ടു് ആ കോളേജിനെ സാമ്പിളിൽ ഉൾപ്പെടുത്താം. അങ്ങനെ സാമ്പിളിലെ ആദ്യത്തെ അംഗം തിരഞ്ഞെടുക്കപ്പെട്ടു. തുടർന്നു് ബാക്കി ആവശ്യമുള്ള അംഗങ്ങളെയും ഇതു പോലെ തന്നെ തിരഞ്ഞെടുക്കുക.

സഞ്ചിതസംയുക്ത സമ്പ്രദായത്തിൽ സമഷ്ടിയിലെ ഓരോ അംഗത്തിനും സാമ്പിളിലേക്കു് തിരഞ്ഞെടുക്കപ്പെടുവാനുള്ള സംഭാവ്യത അതിൻെറ പരിമാണത്തിനു് (സഹായക അഭിലക്ഷണത്തിൻെറ മൂല്യത്തിനു്) ആനുപാതികമാണെന്നുള്ളതു് പ്രഥമദൃഷ്ട്യാ തന്നെ വ്യക്തമാണു്. പക്ഷേ ലാഹിരി സമ്പ്രദായത്തിൽ അതു് തെളിയിക്കേണ്ടിയിരിക്കുന്നു. l ക്രമസംഖ്യയോടു കൂടിയ അംഗം തിരഞ്ഞെടുക്കപ്പെടുവാനുള്ള സംഭാവ്യത എന്തെന്നു പരിശോധിച്ചു നോക്കാം.

ഓരോ ഘട്ടത്തിലും ഏതെങ്കിലും അംഗം തിരഞ്ഞെടുക്കപ്പെടുകയോ തിരഞ്ഞെടുക്കപ്പെടാതിരിക്കുകയോ ചെയ്യാം. തിരഞ്ഞെടുക്കപ്പെടുന്ന ഘട്ടങ്ങൾക്കു് സഫലഘട്ടങ്ങൾ എന്നും അല്ലാത്തവക്കു് വിഫലഘട്ടങ്ങൾ എന്നും പേർ പറ

യും. വിഫലഘട്ടങ്ങൾ അന്യോന്യം അപവർജ്ജികളായ N വിധങ്ങളിൽ ഉണ്ടാവാം. [ആദ്യം തിരഞ്ഞെടുക്കുന്ന സംഖ്യ i യും രണ്ടാമതു തിരഞ്ഞെടുക്കുന്നതു് Y_i യെക്കാൾ വലുതുമായിരുന്നാൽ ആ ഘട്ടം വിഫലമാവും i ക്കു് 1 മുതൽ N വരെയുള്ള ഏതു മൂല്യം വേണമെങ്കിലുമാവാമല്ലൊ] ഒരു ഘട്ട വിഫലമാവാനുള്ള സംഭാവ്യത ഇവയുടെ സംഭാവ്യതകളുടെ തുകയാണു്. ആദ്യസംഖ്യയായി i യും രണ്ടാം സംഖ്യയായി Y_i യെക്കാൾ വലിയ സംഖ്യയും തിരഞ്ഞെടുക്കപ്പെടുവാനുള്ള സംഭാവ്യത, ഈ രണ്ടു തിരഞ്ഞെടുപ്പുകളും സ്വതന്ത്രങ്ങളായതുകൊണ്ടു്,

$\frac{1}{N} \frac{(M-Y_i)}{M}$, ആണു്. അങ്ങനെ ഒരു ഘട്ടം വിഫലമാവാനുള്ള സംഭാവ്യത

$$\frac{1}{NM} \sum_{i=1}^{N} (M-Y_i) = \frac{NM-Y}{NM} = 1 - \frac{\bar{Y}}{M}$$

ആണു്. സമഷ്ടിയിലെ i ക്രമസംഖ്യയുള്ള അംഗം സാമ്പിളിലേയ്ക്കു് തിരഞ്ഞെടുക്കപ്പെടുന്നതു് ഒന്നാം ഘട്ടത്തിലോ രണ്ടാം ഘട്ടത്തിലോ അങ്ങനെ ഏതു ഘട്ടത്തിൽ വേണമെങ്കിലുമാവാം.

ഒന്നാം ഘട്ടത്തിൽ തിരഞ്ഞെടുക്കപ്പെടുവാനുള്ള സംഭാവ്യത

$$= \frac{1}{N} \frac{Y_i}{M}$$

ഒന്നാം ഘട്ടം വിഫലമാവുകയും രണ്ടാം ഘട്ടത്തിൽ അതു് തിരഞ്ഞെടുക്കപ്പെടുകയും ചെയ്യാനുള്ള സംഭാവ്യത

$$= \left(1 - \frac{\bar{Y}}{M}\right) \frac{1}{N} \cdot \frac{Y_i}{M}$$

ഒന്നും രണ്ടും ഘട്ടങ്ങൾ വിഫലമാവുകയും മൂന്നാം ഘട്ടത്തിൽ തിരഞ്ഞെടുക്കപ്പെടുകയും ചെയ്യാനുള്ള സംഭാവ്യത

$$= \left(1 - \frac{\bar{Y}}{M}\right)^2 \frac{1}{N} \frac{Y_i}{M}$$

അങ്ങനെ ഏതെങ്കിലും ഘട്ടത്തിൽ i ക്രമസംഖ്യയായുള്ള അംഗം തിരഞ്ഞെടുക്കപ്പെടുവാനുള്ള സംഭാവ്യത

$$= \frac{1}{N} \cdot \frac{Y_i}{M} + \left(1 - \frac{\bar{Y}}{M}\right) \frac{1}{N} \cdot \frac{Y_i}{M} + \left(1 - \frac{\bar{Y}}{M}\right)^2 \frac{1}{N} \cdot \frac{Y_i}{M} + \ldots\ldots$$

$$= \frac{1}{N} \frac{Y_i}{M} \sum_{j=0}^{N} \left(1 - \frac{\bar{Y}}{M}\right)^j$$

ഇവിടെ $1 - \frac{\bar{Y}}{M}$, ഒന്നിനെക്കാൾ ചെറുതായിരിക്കുമെന്നതുകൊണ്ടു്, ഇതിലെ

നന്ത ഗുണോത്തര ശ്രേണിയുടെ തുക,

$$\frac{1}{1-\left(1-\frac{\bar{Y}}{M}\right)}=\frac{M}{\bar{Y}}$$

അത്രെ. അങ്ങനെ ഈ സംഭാവ്യത,

$$\frac{1}{N}\cdot\frac{Y_i}{M}\cdot\frac{M}{\bar{Y}}=\frac{Y_i}{N\bar{Y}}=\frac{Y_i}{Y}$$

എന്നു വന്നുചേരുന്നു. അതായതു്, i ക്രമസംഖ്യയായുള്ള സമഷ്ടിയിലെ അംഗത്തിനു് സാമ്പിളിൽ ഉൾപ്പെട്ടവാനുള്ള സംഭാവ്യത അതിൻെറ പരിമാണത്തിനു് ആനുപാതികമാണു്. ഇതിൽ നിന്നു് ലാഹിരിസമ്പ്രദായത്തിൽ എടുക്കുന്ന സാമ്പിൾ പരിമാണാനുപാതിക സംഭാവ്യതാ സാമ്പിളാണെന്നു സിദ്ധിക്കുന്നു.

ലാഹിരിസമ്പ്രദായത്തില്പ് രണ്ടു് ഗുണങ്ങളാണുള്ളതു്. ഒന്നാമതു് സഞ്ചിതസംയുക്തങ്ങളുടെ പട്ടിക ആവശ്യമില്ല. രണ്ടാമതു്, എല്ലാ വ്യക്തികളുടെയും സഹായക അഭിലക്ഷണത്തിൻെറ മൂല്യം (പരിമാണം) അറിയണമെന്നു നിർബന്ധമില്ല. അവയിൽ ഏററവും വലുതും തിരഞ്ഞെടുക്കപ്പെടുന്ന ആദ്യസംഖ്യ സൂചിപ്പിക്കുന്ന വ്യക്തികളുടെ പരിമാണവും അറിഞ്ഞാൽമതി. ഇതു രണ്ടും പ്രായോഗികമായി വളരെ ഉപകാരപ്രദമത്രെ. പക്ഷേ ഇതിനു് ഗൗരവതരമായ ഒരു തകരാറുണ്ടെന്നുള്ള വസ്തുത വിസ്മരിക്കാനാവില്ല. y ഉം M ഉം ആയുള്ള വ്യത്യാസം വർധിക്കുന്തോറും ഒരു ഘട്ടം നിഷ്ഫലമാവാനുള്ള സംഭാവ്യതയും അങ്ങനെ നിഷ്ഫലഘട്ടങ്ങളുടെ എണ്ണവും വർധിക്കും. ഇതു് സമയനഷ്ടത്തിൽ കലാശിക്കും. ഇതു പരിഹരിക്കാൻ ചില മാർഗങ്ങളും നിർദേശിക്കപ്പെട്ടിട്ടുണ്ടു്. വിസ്തരഭയത്താൽ അവ ഇവിടെ ചർച്ചചെയ്യുന്നില്ല.

7. സമാന്തരമാധ്യത്തിൻെറ ആകലനം: പുനസ്ഥാപനത്തോടു കൂടിയുള്ള പരിമാണാനുപാതസംഭാവ്യതാ സാമ്പിളനത്തിൽ

സമഷ്ടിയുടെ സമാന്തരമാധ്യമായ $\bar{X}$ ആണു് ആകലനം ചെയ്യേണ്ടതെന്നിരിക്കട്ടെ. ഈ സാമ്പിളനരീതിയനുസരിച്ചു് പ്രതിസ്ഥാപനത്തോടു കൂടി, n അംഗങ്ങളുള്ള ഒരു സാമ്പിൾ എടുത്തിരിക്കുന്നു എന്നും അതിലെ അംഗങ്ങളുടെ അഭിലക്ഷണമൂല്യങ്ങൾ $x_1, x_2 \cdots x_n$ എന്നിവയാണെന്നും അവയുടെ സാമ്പിളിൽ ഉൾപ്പെടുവാനുള്ള സംഭാവ്യതകൾ യഥാക്രമം $p_1, p_2 \cdots p_n$ എന്നിവയാണെന്നും സങ്കൽപിക്കുക. സമഷ്ടിയിലെ i ക്രമസംഖ്യയായുള്ള അംഗത്തിനു് സാമ്പിളിൽ ഉൾപ്പെടുവാനുള്ള സംഭാവ്യത P_i ആണെന്നും ആ അംഗത്തിൻെറ അഭിലക്ഷണമൂല്യം X_i ആണെന്നും ഇരിക്കട്ടെ.

$$\bar{x}_p=\frac{1}{Nn}\sum_{i=1}^{n}\frac{x_i}{p_i} \qquad 7.7.1$$

$\bar{x}_p$ ആണ് സമഷ്ടിയുടെ സമാന്തരമാധ്യമായ $\bar{X}$ ന്റെ ആകലമായി നിർദേശിക്കപ്പെട്ടിരിക്കുന്നതു്. ഇതു് അനഭിനതമാണെന്നു കാണാൻ പ്രയാസമില്ല. $\frac{x_i}{p_i}$ ഒരു യാദൃച്ഛികചരമാണ്. അതിനു് എടുക്കാവുന്ന മൂല്യങ്ങൾ $\frac{X_1}{P_1}, \frac{X_2}{P_2}, \ldots$ $\frac{X_N}{P_N}$ എന്നിവയത്രെ. ഇവയുടെ സംഭാവ്യതകൾ യഥാക്രമം, $P_1, P_2, \ldots P_N$ എന്നിവയാണ്.

അങ്ങനെ,

$$E\left(\frac{x_i}{p_i}\right) = \sum_{i=1}^{N} \frac{X_i}{P_i} P_i = X$$

ആയതുകൊണ്ട്,

$$E(\bar{x}_p) = \frac{1}{Nn} E\left\{\sum_{i=1}^{n} \frac{x_i}{p_i}\right\} = \frac{1}{Nn} \sum_{i=1}^{n} X = \frac{n\bar{X}}{n} = \bar{X}$$

അതായതു്, $\bar{x}_p$ ഒരു അനഭിനതആകലമാണ്. അടുത്തതായി ഇതിന്റെ സാമ്പിളനപ്രസരണമാണ് പരിഗണിക്കാനുള്ളതു്. $\left\{\frac{x_i}{p_i}\right\}$ കൾ സ്വതന്ത്ര യാദൃച്ഛിക ചരങ്ങളായതുകൊണ്ട്, (പ്രതിസ്ഥാപനത്തോടുകൂടിയ സാമ്പിളനത്തിൽ)

$$V(\bar{x}_p) = \frac{1}{N^2n^2} \sum_{i=1}^{n} V\left(\frac{x_i}{p_i}\right) = \frac{1}{N^2n^2} \sum_{i=1}^{n} \left\{\sum_{j=1}^{N} \left(\frac{X_j}{P_j} - X\right)^2 P_j\right.$$

$$= \frac{1}{N^2n^2}\, n \sum_{j=1}^{N} \left(\frac{X_j}{P_j} - X\right)^2 P_j$$

$$= \frac{1}{N^2n}\left[\sum_{j=1}^{N} \frac{X_j^2}{P_j} - X^2\right] \qquad 7.7.2$$

സാമ്പിൾപരിമാണമായ n വർധിക്കുന്നതനുസരിച്ച് $\bar{x}_p$ ന്റെ സാമ്പിളന പ്രസരണം കുറയുമെന്നതു് ഇവിടെ പ്രത്യേകം ശ്രദ്ധിക്കേണ്ട ഒരു വസ്തുതയാണ്. $V(\bar{x}_p)$ ന്റെ ഒരു അനഭിനതആകലം ലഭിക്കുവാൻ $\frac{x_i}{Np_i}$, $(i=1,2\ldots n)$ കൾ $\bar{X}$ ന്റെ തുല്യപ്രസരണത്തോടുകൂടിയ n അനഭിനത, സ്വതന്ത്ര, ആകലങ്ങളാണെന്നുള്ള വസ്തുത ഉപയോഗപ്പെടുത്തിയാൽ മതി. അങ്ങനെ,

$$\hat{V}(\bar{x}_p) = \frac{1}{n(n-1)}\left(\sum_{i=1}^{n} \frac{x_i^2}{N^2p_i^2} - n\,\bar{x}_p^2\right)$$

ഇവിടെ വിവരിച്ച ആകലനരീതിയും സാമ്പിളനപ്രസരണത്തിൻെറ ഒരു ദാഹരണം കൊണ്ടു് വിശദമാക്കാം. പട്ടിക 7.2 ലെ സമഷ്ടിയിൽ നിന്നു് 3 അംഗങ്ങളുള്ള ഒരു സാമ്പിളാണു് എടുക്കുന്നതെന്നിരിക്കട്ടെ. 1,3,4 എന്നീ ക്രമസംഖ്യകളുള്ള അംഗങ്ങളാണു് തിരഞ്ഞെടുക്കപ്പെട്ടതെന്നും സങ്കല്പിക്കുക. അവയുടെ അഭിലക്ഷണമൂല്യങ്ങളും ഓരോന്നും തിരഞ്ഞെടുക്കപ്പെടുവാനുള്ള സംഭാവ്യതകളും താഴെ കൊടുക്കുന്നു.

പട്ടിക 7.3

സാമ്പിൾ കലാലയങ്ങൾ i	വാർഷികച്ചെലവു് (ലക്ഷങ്ങളിൽ) (x_i)	തിരഞ്ഞെടുക്കപ്പെടാനുള്ള സംഭാവ്യത (p_i)	$\frac{x_i}{p_i}$
1	12	$\frac{2340}{6181} = .378$	31.74
2	7	$\frac{1012}{6181} = .164$	42.68
3	3	$\frac{520}{6181} = 0841$	35.67

$$\bar{x}_p = \frac{110.09}{15} = 7.33$$

$$\hat{V}(\bar{x}_p) = \frac{1}{6}(164.06 - 161.19)$$

$$= .48$$

ഇവിടെ പ്രതിസ്ഥാപനത്തോടു കൂടിയ സാമ്പിളനമാണു് പരിഗണിച്ചതു്. പ്രതിസ്ഥാപനമില്ലാത്ത സാമ്പിളനരീതി അവലംബിച്ചാൽ ഓരോ ഘട്ടത്തിലും ഒരംഗത്തിനു് സാമ്പിളിൽ ഉൾപ്പെടുവാനുള്ള സംഭാവ്യത തുല്യമായിരിക്കുകയില്ല. സമാന്തരമാധ്യത്തിൻെറ ഒരു അനഭിനത ആകലം കണ്ടുപിടിക്കുന്നതും അതിൻെറ സാമ്പിളനപ്രസരണം നിർണയിക്കുന്നതും അത്യന്തം സങ്കീർണമായ ഒരു പ്രക്രിയയായിത്തീരും. അതുകൊണ്ടാണു് അതു് ഇവിടെ ചർച്ച ചെയ്യാത്തതു്.

ഈ സാമ്പിളനരീതി സാധാരണയായി സംഘസാമ്പിളനത്തിലും ബഹുഘട്ട സാമ്പിളനത്തിലും ഉപയോഗപ്പെടുത്താറുണ്ടു്. സംഘസാമ്പിളനത്തിൽ സംഘങ്ങളുടെ സാമ്പിളാണല്ലൊ ആദ്യമെടുക്കുന്നതു്. ഓരോ സംഘത്തിൻെറയും അംഗസംഖ്യയാണു് അതിൻെറ പരിമാണം. ഓരോ സംഘത്തിനും സാമ്പിളിൽ ഉൾപ്പെടുവാനുള്ള സംഭാവ്യത, അതിനു ആനുപാതികമാകത്തക്കവണ്ണം സാമ്പിളെടുത്താൽ, ആകലത്തിൻെറ സൂക്ഷ്മത വർധിക്കും. ഇതു പോലെ തന്നെ ബഹുഘട്ട സാമ്പിളനത്തിലും.

കൃഷിഭൂമികളിൽ നിന്നുള്ള സാമ്പിളനമാണ് ഈ സാമ്പിളനരീതി ലാഭകരമായി ഉപയോഗപ്പെടുത്താവുന്ന മററൊരു രംഗം. സാമ്പിളെടുക്കേണ്ട കൃഷിഭൂമിയുടെ ഒരു പ്ലാൻ തയ്യാറാക്കുക. പല വിസ്തീർണമുള്ള ചെറിയ ചെറിയ അംശങ്ങളായി ആകെ കൃഷിസ്ഥലത്തെ വിഭജിച്ചിരിക്കുമല്ലൊ. പ്ലാൻ മുഴുവൻ ആദ്യ ചതുർഥാംശത്തിൽ വരത്തക്ക വണ്ണം x.y അക്ഷങ്ങൾ സങ്കൽപിക്കുക. y അക്ഷത്തിൽ നിന്ന് l അകലത്തിലും x അക്ഷത്തിൽ നിന്ന് m അകലത്തിലും വരച്ചിരിക്കുന്ന സമാന്തരരേഖകൾക്കുള്ളിലാണ് പ്ലാൻ എന്നും സങ്കൽപിക്കുക. l,m എന്നിവയെക്കാൾ കുറവായ രണ്ട് യാദൃച്ഛികസംഖ്യകൾ എടുക്കുക. അവ നിർദേശാങ്കങ്ങളായുള്ള ബിന്ദു ഏതു അംശത്തിലാണോ വീഴുന്നത് ആ അംശം സാമ്പിളിൽ ഉൾപ്പെടുത്തുക. ഈ പരിപാടി ആവർത്തിച്ചാൽ സാമ്പിളിൽ ഉൾപ്പെടുത്തേണ്ടിടത്തോളം അംഗങ്ങളെ തിരഞ്ഞെടുക്കാവുന്നതാണ്. ഇവിടെ ഒരു കാര്യം പ്രത്യേകം ശ്രദ്ധിക്കണം. പ്രതിസ്ഥാപനത്തോടുകൂടിയ സാമ്പിളാണ് എടുക്കുന്നത്. അതുകൊണ്ട് ഒരംശത്തിൽ തന്നെ ഒന്നിലധികം ബിന്ദുക്കൾ വീണാൽ ആ അംശം അത്രയും പ്രാവശ്യം സാമ്പിളിൽ ഉൾപ്പെടുത്തണം. ആ അടിസ്ഥാനത്തിൽ വേണം ആകലനം നിർവഹിക്കുവാൻ.

8. ക്വോട്ടാ സാമ്പിളനം

8. ഇതു വരെ പരിഗണിച്ച എല്ലാ സാമ്പിളനരീതികളും സംഭാവ്യതാസാമ്പിളനങ്ങളാണ്. അതായതു്, ഓരോ അംഗത്തിനും സാമ്പിളിൽ ഉൾപ്പെടുവാനുള്ള സംഭാവ്യത നിശ്ചിതമായിരിക്കത്തക്കവണ്ണമാണ് സാമ്പിളെടുക്കുന്നതു്. ക്വോട്ടാസാമ്പിളനം ഈ വിഭാഗത്തിൽ പെടുതല്ല ഗവേഷകൻ സ്വന്തം അനുഭവജ്ഞാനത്തെ മാത്രം ആധാരമാക്കി സമഷ്ടിയെ വിവിധഭാഗങ്ങളായി വിഭജിക്കുന്നു. ഓരോ വിഭാഗത്തിലുമുള്ള എത്ര അംഗങ്ങളിൽ നിന്ന് വിവരശേഖരണം നടത്തണമെന്ന് സ്വയം നിശ്ചയിക്കുന്നു. ആ ജോലി നിർവഹിക്കാൻ അന്വേഷകരെ ചുമതലപ്പെടുത്തുന്നു. അവർ പ്രായോഗികസൗകര്യം പ്രധാനമായി കണക്കിലെടുത്തു കൊണ്ടു ആരിൽ നിന്നെല്ലാമാണ് വിവരം ശേഖരിക്കേണ്ടതെന്നു തീരുമാനിക്കുന്നു. ഇതാണ് ക്വോട്ടാസാമ്പിളനത്തിൻെറ തിരഞ്ഞെടുക്കൽ പരിപാടി. ഇങ്ങനെ തിരഞ്ഞെടുക്കപ്പെട്ട സാമ്പിളിൽ നിന്നു ലഭിക്കുന്ന വിവരങ്ങളെന്നു അടിസ്ഥാനമാക്കി ആകലനം നിർവഹിക്കുമ്പോൾ യുക്തമെന്നു തോന്നുന്ന ഭാരങ്ങൾ ഉപയോഗിച്ച് ഭാരിതആകലങ്ങൾ എടുക്കാറുണ്ട്. ആകലങ്ങളുടെ പ്രതീക്ഷയോ സാമ്പിളനപ്രസരണമോ കണ്ടു പിടിക്കാൻ മാർഗമൊന്നുമില്ല

പലപ്പോഴും ക്വോട്ടാസാമ്പിളനത്തിലൂടെ ലഭിക്കുന്ന വിവരങ്ങൾ യാഥാർഥ്യവുമായി വളരെ അടുത്തു വരാറുണ്ട്. മററു സാമ്പിളനരീതികൾ മുഖേന ലഭിക്കുന്നവയെക്കാൾ മെച്ചപ്പെട്ട ആകലങ്ങൾ ഈ രീതിയിൽ ലഭിച്ചെന്നു വരാം. പക്ഷേ ആകലങ്ങളുടെ സൂക്ഷ്മത കണ്ടു പിടിക്കാനാവില്ലെന്നുള്ളതാണ് ഈ രീതിയുടെ പ്രധാന ദൗർബല്യം.

ക്വോട്ടാസാമ്പിളനത്തിൻെറ എല്ലാ ഘട്ടത്തിലും ഗവേഷകൻ സ്വന്തം പരിചയസമ്പത്തിനെ ആധാരമാക്കി സ്വയം പരിപാടി ആവിഷ്ക്കരിക്കുകയാണ്

ചെയ്യുന്നതെന്നു പറഞ്ഞല്ലോ. എങ്കിലും ചില പൊതുതത്വങ്ങൾ സാധാരണയായി അംഗീകരിക്കാറുണ്ട്. സമഷ്ടിയെ വിവിധ ഭാഗങ്ങളായി തരംതിരിക്കുന്നതു് പരിഗണിക്കപ്പെടുന്ന അഭിലക്ഷണത്തോടു ബന്ധപ്പെട്ട ഏതെങ്കിലും സഹായകഅഭിലക്ഷണത്തെ ആധാരമാക്കിയാണ്. ഉദാഹരണമായി ഒരു തിരഞ്ഞെടുപ്പിൽ ഓരോ സ്ഥാനാർഥിക്കും അനുകൂലമായി ലഭിക്കാവുന്ന വോട്ടിൻെറ എണ്ണം ആകലനം ചെയ്കയാണ് അന്വേഷണലക്ഷ്യമെന്നിരിക്കട്ടെ. ആ നിയോജകമണ്ഡലത്തിലെ സമ്മതിദായകരാണ് സമഷ്ടി. അവരോരോരുത്തരും ഇഷ്ടപ്പെടുന്ന സ്ഥാനാർഥി ആരെന്നുള്ളതാണ് പരിഗണിക്കപ്പെടുന്ന അഭിലക്ഷണം. മുൻ തിരഞ്ഞെടുപ്പുകളിലെ അനുഭവവും, നിയോജകമണ്ഡലത്തെപ്പററി ഗവേഷകനുള്ള അറിവും മററും കണക്കിലെടുത്തു് ചിന്തിക്കുമ്പോൾ വോട്ടറുടെ സമുദായവും അയാൾ ഇഷ്ടപ്പെടുന്ന സ്ഥാനാർഥിയും തമ്മിൽ ബന്ധപ്പെട്ടിരിക്കുന്നു എന്ന് തോന്നുകയാണെങ്കിൽ സമുദായാടിസ്ഥാനത്തിൽ സമഷ്ടിയെ വിഭജിക്കുന്നതു് പ്രസക്തമാണ്. അതായതു്, സമഷ്ടിയെ വിവിധ ഭാഗങ്ങളായി വിഭജിക്കുന്നതിനു് സമുദായം എന്ന സഹായകഅഭിലക്ഷണം ഉപയോഗിക്കാം. നേരെ മറിച്ച്, സമ്മതിദായകൻെറ തൊഴിലും സ്ഥാനാർഥികളുടെ നേരെയുള്ള ചായ്‌വും ബന്ധപ്പെട്ടിരിക്കുന്നു എന്നു കരുതാനാണ് ന്യായമുള്ളതെങ്കിൽ തൊഴിൽ വേണം സഹായകഅഭിലക്ഷണമായി സ്വീകരിക്കുവാൻ.

അടുത്തതു് ഓരോ വിഭാഗത്തിൽ നിന്നു് എത്ര അംഗങ്ങളെ വീതം സാമ്പിളിൽ ഉൾപ്പെടുത്തണം എന്നതാണ്. ഓരോ വിഭാഗത്തിൻെറയും അംഗസംഖ്യക്കു് ആനുപാതികമായാണ് ഇതു് നിർവഹിക്കാറുള്ളതു്. മുൻ ഉദാഹരണത്തിൽ സമുദായത്തിൻെറ അടിസ്ഥാനത്തിലാണ് സമ്മതിദായകരെ വിഭജിച്ചിരിക്കുന്നതെങ്കിൽ, ഓരോ സമുദായത്തിൻെറയും അംഗസംഖ്യക്ക് ആനുപാതികമായി സാമ്പിളിൽ ഉൾപ്പെടുത്തേണ്ട അംഗങ്ങളുടെ എണ്ണം നിശ്ചയിക്കാം. രണ്ടു സമുദായങ്ങളിൽപ്പെട്ട സമ്മതിദായകരാണ് ഉള്ളതെന്നും അവയുടെ അംഗസംഖ്യ 40,000ഉം 60,000 ഉം ആണെന്നും സങ്കൽപിക്കുക. 100 അംഗങ്ങളുള്ള ഒരു സാമ്പിളാണ് എടുക്കേണ്ടതെന്നിരിക്കട്ടെ. ആദ്യത്തെ സമുദായത്തിൽ പെട്ട 40 പേരെയും രണ്ടാമത്തെ സമുദായത്തിൽപ്പെട്ട 60 പേരെയും സാമ്പിളിൽ ഉൾപ്പെടുത്തണമെന്നു് സാരം.

അടുത്ത ഘട്ടം സാമ്പിളിൽ ഉൾപ്പെടുത്തേണ്ട അംഗങ്ങളെ നിർണയിക്കലാണ്. ഈ ജോലി ചില നിബന്ധനകൾക്കു വിധേയമായി അന്വേഷകനു് വിട്ടുകൊടുക്കുകയാണ് ചെയ്യുന്നതു്. അന്വേഷകർ ബോധപൂർവം പ്രവർത്തിക്കുന്നതിലാണ് സാമ്പിളനത്തിൻെറ വിജയം നില കൊള്ളുന്നതു്. പരിഗണിക്കപ്പെടുന്ന അഭിലക്ഷണവുമായി ബന്ധപ്പെട്ട മററു് അഭിലക്ഷണങ്ങൾ കൂടി പരിഗണിച്ച് ഓരോ വിഭാഗത്തിൽ നിന്നും തിരഞ്ഞെടുക്കുന്ന അംഗങ്ങളെ ആ വിഭാഗത്തിൻെറ പ്രതിനിധികളാക്കി തീർക്കാൻ അന്വേഷകൻ ശ്രദ്ധിക്കണം. ഉദാഹരണമായി മുൻ സൂചിപ്പിച്ച രണ്ടു സമുദായങ്ങളിൽ നിന്നു് തിരഞ്ഞെടുക്കേണ്ട 40 ഉം 60 ഉം അംഗങ്ങളെ തിരഞ്ഞെടുക്കുവാൻ ഓരോ അന്വേഷകർ നിയോഗി

ക്കപ്പെടുന്നു എന്നിരിക്കട്ടെ. ഓരോ സമുദായത്തിലും സ്ത്രീകളും പുരുഷന്മാരുമുണ്ട്; ചെറുപ്പക്കാരും പ്രായമായവരുമുണ്ട്; ഗ്രാമങ്ങളിൽ താമസിക്കുന്നവരും പട്ടണങ്ങളിൽ താമസിക്കുന്നവരുമുണ്ട്; കൃഷിക്കാരും കച്ചവടക്കാരും ഉദ്യോഗസ്ഥന്മാരുമുണ്ട്. ഈ എല്ലാ വിഭാഗങ്ങൾക്കും അർഹമായ പ്രാതിനിധ്യം നൽകുവാൻ അന്വേഷകൻ ശ്രദ്ധിക്കണം. പക്ഷേ ഓരോ വിഭാഗത്തിൽ നിന്നും ഏതേതു വ്യക്തികളെയാണ് സാമ്പിളിൽ ഉൾപ്പെടുത്തേണ്ടതെന്നുള്ളത് അന്വേഷകൻെറ ഇഷ്ടാനുസരണം തീരുമാനിക്കാം. വിവരശേഖരണത്തിൻെറ സൗകര്യം മാത്രമാണ് ഇവിടത്തെ പ്രധാനപരിഗണന. വിവരം നൽകാൻ വിസമ്മതിക്കുന്നവരേയും കണ്ടുപിടിക്കാൻ ബുദ്ധിമുട്ടു വരുന്നവരെയും ഉപേക്ഷിച്ച് പകരം വ്യക്തികളെ ഉൾപ്പെടുത്താം.

താരതമ്യേന ഏററവും ചെലവു കുറഞ്ഞ ഒരു സാമ്പിളനരീതിയാണ് കോട്ടാസാമ്പിളനം. അതാണ് അതിൻെറ പ്രധാന ആകർഷണവും. സമർത്ഥന്മാരായ ഗവേഷകർ കൈകാര്യം ചെയ്താൽ വളരെ സൂക്ഷ്മമായ വിവരങ്ങൾ ഈ സമ്പ്രദായത്തിൽ ശേഖരിക്കാൻ കഴിഞ്ഞെന്നു വരാം.

സമഷ്ടിയെ വിവിധ ഭാഗങ്ങളായി ഭാഗിച്ച് അവയിൽ നിന്നും ഏതാനും എണ്ണം സംഭാവ്യതാസാമ്പിളനരീതിയിൽ തിരഞ്ഞെടുക്കുകയും അങ്ങനെ തിരഞ്ഞെടുത്ത ഭാഗങ്ങളിൽ നിന്ന് കോട്ടാസാമ്പിളനരീതിയിൽ വിവരം ശേഖരിക്കുകയും ചെയ്യുന്ന രീതിയും നിലവിലുണ്ട്. ഇതിന് സംഭാവ്യതാ–കോട്ടാ സാമ്പിളനം എന്നു പേർ പറയും. ഓരോ വിഭാഗത്തിലെയും വ്യക്തികൾ പരിഗണിക്കപ്പെടുന്ന അഭിലക്ഷണത്തെ സംബന്ധിച്ചിടത്തോളം ഏകാത്മകങ്ങളാണെങ്കിൽ ഈ സമ്പ്രദായം അത്യന്തം വിജയകരമായിരിക്കും. ആകലങ്ങളുടെ സാമ്പിളനപ്രസരണത്തെപ്പററി ഏകദേശമായെങ്കിലും മനസ്സിലാക്കാൻ കഴിയും എന്നതാണ് ഈ രീതിയുടെ മെച്ചം.

അഭിപ്രായങ്ങൾ, വിലനിലവാരം തുടങ്ങിയവയെപ്പററിയുള്ള അന്വേഷണങ്ങളിലാണ് കോട്ടാസാമ്പിളനം സാധാരണയായി ഉപയോഗിക്കാറുള്ളത്. ഓരോ അന്വേഷകനും ഓരോ കോട്ടാ (എത്ര വ്യക്തികളിൽ നിന്ന് വിവരം ശേഖരിക്കണമെന്ന നിർദ്ദേശം) നൽകുകയും, അയാൾ സ്വന്തം ഇഷ്ടമനുസരിച്ച് വ്യക്തികളെ തീരുമാനിച്ച് വിവരശേഖരണം നടത്തുകയും ചെയ്യുന്നതു കൊണ്ടാണ് ഈ സാമ്പിളനത്തിന് കോട്ടാ സാമ്പിളനം എന്നു പേർ പറയുന്നത്.

9. വിവിധ പ്രാവസ്ഥാസാമ്പിളനം

വിവിധഘട്ടസാമ്പിളനവുമായി സാമ്യമുള്ള ഒരു സാമ്പിളനസമ്പ്രദായമാണ് ഇത്. യുക്തമായ ഉപവിവരങ്ങൾ ശേഖരിക്കാനായി സാമാന്യം വലിയ ഒരു സാമ്പിൾ ഏതെങ്കിലും സംഭാവ്യതാസാമ്പിളനരീതിയിൽ തിരഞ്ഞെടുക്കുകയും ശേഖരിച്ച ഉപവിവരങ്ങളെ അടിസ്ഥാനമാക്കി ആദ്യസാമ്പിളിൽ ഉൾപ്പെടുത്തിയ അംഗങ്ങളിൽ നിന്ന് താരതമ്യേന ചെറുതായ ഒരു സാമ്പിളെടുത്ത് ആ അംഗങ്ങളിൽ നിന്ന് പ്രധാന വിവരങ്ങൾ ശേഖരിക്കുകയും ചെയ്യുന്ന സമ്പ്രദായ

ത്തിന് ദ്വിപ്രാവസ്ഥാസാമ്പിളനം എന്ന പേർ പറയും. ഇവിടെ രണ്ടു പ്രാവശ്യം സാമ്പിളെടുത്തല്ലൊ. ഇതേ രീതിയിൽ കൂടുതൽ പ്രാവശ്യം ഉപസാമ്പിളുകൾ എടുക്കുകയാണെങ്കിൽ ആ സമ്പ്രദായത്തിന് വിവിധ പ്രാവസ്ഥാസാമ്പിളനം എന്നു പേർ പറയും. വിവിധഘട്ടസാമ്പിളനത്തിൽ സമഷ്ടിയെ സംഘങ്ങളായി വിഭജിച്ച് അവയിൽ നിന്ന് സാമ്പിളെടുക്കുകയും, തിരഞ്ഞെടുക്കപ്പെട്ട സംഘങ്ങളെ ഉപസംഘങ്ങളായി വിഭജിച്ച് അവയിൽ നിന്ന് സാമ്പിളെടുക്കുകയും അങ്ങനെ മുന്നോട്ടു പോവുകയുമാണല്ലൊ ചെയ്യുന്നത്. നേരെ മറിച്ച് വിവിധപ്രാവസ്ഥാസാമ്പിളനത്തിൽ ആദ്യസാമ്പിളിലെയും ഉപസാമ്പിളുകളിലെയുമെല്ലാം അംഗങ്ങൾ സമഷ്ടിയിലെ പ്രാഥമികവ്യക്തികളാണ്. ഇതാണ് ഇവ തമ്മിലുള്ള മൗലികവ്യത്യാസം. ഈ കാരണത്താൽ സമഷ്ടിയിലെ എല്ലാ അംഗങ്ങളുടെയും ലിസ്റ്റ് — പൂർണമായ ഒരു സാമ്പിളനഫ്രേം — തയ്യാറാക്കേണ്ടതു് ആവശ്യമാണ്. ആദ്യസാമ്പിളിൽ നിന്ന് എളുപ്പത്തിൽ ശേഖരിക്കാവുന്ന തരത്തിലുള്ള പഠനവിധേയമായ അഭിലക്ഷണവുമായി ബന്ധപ്പെട്ടവയുമായ വിവരങ്ങളാണ് ശേഖരിക്കുക. ആ വിവരങ്ങളുപയോഗിച്ച് ആദ്യസാമ്പിളിനെ സ്തരങ്ങളായി വിഭജിച്ച് രണ്ടാം ഉപസാമ്പിൾ സ്തരിതസാമ്പിളനരീതിയിൽ തിരഞ്ഞെടുക്കുക. രണ്ടാം സാമ്പിളിൽ നിന്ന് പ്രധാന അഭിലക്ഷണത്തെപ്പററിയുള്ള വിവരങ്ങൾ ശേഖരിക്കുക. സാധാരണ ഉപയോഗത്തിലുള്ള ഒരു ദ്വിപ്രാവസ്ഥാസാമ്പിളനരീതിയാണ് ഇതു്. ഉദാഹരണമായി ഒരു നഗരത്തിലെ സ്ഥിരതാമസക്കാരുടെ ശരാശരി ജീവിതച്ചെലവ് നിർണയിക്കുകയാണ് അന്വേഷണലക്ഷ്യമെന്നിരിക്കട്ടെ. കുടുംബങ്ങളെ വേണം ഇവിടെ പ്രാഥമിക വ്യക്തികളായി പരിഗണിക്കാൻ. നഗരത്തിലെ കുടുംബങ്ങളുടെ ലിസ്റ്റിൽ നിന്ന് സാമാന്യം വലിയ ഒരു സാമ്പിൾ ലഘുയാദൃച്ഛികസാമ്പിളനരീതിയിൽ തിരഞ്ഞെടുത്തു എന്നു സങ്കൽപിക്കുക. ഓരോ കുടുംബത്തിൻെറയും പ്രതിമാസവരുമാനം എന്നൊരു വിവരം മാത്രം ആദ്യം ശേഖരിക്കുക. ഈ വിവരം ഉപയോഗിച്ച് ആദ്യസാമ്പിളിനെ വിവിധസ്തരങ്ങളായി വിഭജിക്കാം. അടുത്തതായി ആദ്യസാമ്പിളിലെ അംഗങ്ങളിൽ നിന്ന് ഒരു സ്തരിതസാമ്പിൾ എടുത്ത് അതിൽ ഉൾപ്പെട്ടു വരുന്ന കുടുംബങ്ങളിൽ നിന്ന് ജീവിതച്ചെലവിനെപ്പററിയുള്ള വിശദമായ വിവരങ്ങൾ ശേഖരിക്കാം. ഇതു പോലെ തന്നെ വിവിധ പ്രാവസ്ഥാസാമ്പിളുകളും എടുക്കാവുന്നതാണ്. വിശദമായ വിവരങ്ങൾ ചെറിയ ഒരു സാമ്പിളിൽ നിന്നു ശേഖരിച്ചാൽ മതിയാവുന്നതു കൊണ്ട് ആകെ ചെലവ് കുറഞ്ഞിരിക്കുമെന്നതാണ് ഈ സാമ്പിളനരീതിയുടെ വൈശിഷ്ട്യം.

വിവിധ പ്രാവസ്ഥാ സാമ്പിളനത്തിന്, ഘട്ടങ്ങളുടെ എണ്ണം, ഓരോ ഘട്ടത്തിലും അവലംബിക്കുന്ന സാമ്പിളന രീതി തുടങ്ങിയവയെ ആശ്രയിച്ച് നിരവധി രൂപങ്ങളുണ്ട്. മാതൃകയായി മുകളിൽപ്പറഞ്ഞ തരം ദ്വിപ്രാവസ്ഥാസാമ്പിൾ ഉപയോഗിച്ചുള്ള സമാന്തരമാധ്യത്തിൻെറ ആകലവും അതിൻെറ സാമ്പിളന പ്രസരണത്തിൻെറ ആകലവും മാത്രം താഴെ ചേർക്കുന്നു.

സമഷ്ടിയിൽ N അംഗങ്ങളാണുള്ളതെന്നിരിക്കട്ടെ. അതിൽ നിന്ന് പ്രതി

സ്ഥാപനമില്ലാത്ത ലഘ്വയാദൃച്ഛികസാമ്പിളനരീതിയിൽ m അംഗങ്ങളെ തിരഞ്ഞെടുക്കുന്നു. അവയിൽ നിന്നു് ആദ്യഘട്ടത്തിൽ ശേഖരിച്ച വിവരങ്ങളെ അടിസ്ഥാനമാക്കി അതിനെ k സ്തരങ്ങളായി വിഭജിക്കുന്നു. i–ാമതു് സ്തരത്തിൽ m_i അംഗങ്ങളാണു് ഉള്ളതു്. i–ാമത്തെ സ്തരത്തിൽ നിന്നു് n_i അംഗങ്ങൾ എന്ന ക്രമത്തിൽ n അംഗങ്ങളുള്ള ഒരു സ്തരിതയാദൃച്ഛികസാമ്പിൾ എടുക്കുന്നു. i–ാമതു് സ്തരത്തിൽ നിന്നെടുത്ത j–ാമതു് വ്യക്തിയുടെ അഭിലക്ഷണമൂല്യം x_{ij} എന്നിരിക്കട്ടെ.

$\bar{x}_i = \frac{1}{n_i} \sum_{j=1}^{n_i} x_{ij}$ എന്നു സങ്കൽപിക്കുക. $W_i = \frac{m_i}{m}$ എന്നും ഇരിക്കട്ടെ. $\bar{X}$ ൻെറ ആകലമായി നിർദേശിക്കപ്പെട്ടിരിക്കുന്ന $\bar{x}_p$, താഴെപ്പറയും പ്രകാരം നിർവചിക്കപ്പെട്ടിരിക്കുന്നു.

$$\bar{x}_p = \sum_{i=1}^{k} w_i \bar{x}_i \qquad 7.9.1$$

ഇതിൻെറ സാമ്പിളനപ്രസരണത്തിൻെറ ആകലത്തിൻെറ ഒരു ഏകദേശനം,

$$\hat{V}(\bar{x}_p) = \sum_{i=1}^{k} \left\{\frac{1}{n_i} - \frac{1}{m_i}\right\} w_i^2 s_i^2 + \left\{\frac{1}{m} - \frac{1}{N}\right\} \sum_{i=1}^{k} w_i (\bar{x}_i - \bar{x}_p)^2 \qquad 7.9.2$$

ഇവിടെ

$$s_i^2 = \frac{1}{n_i - 1} \sum_{j=1}^{n_j} \{x_{ij} - \bar{x}_i\}^2$$

മററുതരം വിവിധ പ്രാവസ്ഥാസാമ്പിളനങ്ങളിലെ ആകലങ്ങളും അവയുടെ സാമ്പിളന പ്രസരണങ്ങളുടെ ആകലങ്ങളും മററും അതതിൻെറ പ്രത്യേക സ്വഭാവം പരിഗണിച്ചു് നിർണയിക്കേണ്ടതാണു്.

ഈ അധ്യായത്തിൽ സാധാരണ ഉപയോഗത്തിലുള്ള ഏതാനും സാമ്പിളന സമ്പ്രദായങ്ങളെപ്പററി സാമാന്യമായ ചില വസ്തുതകൾ അവതരിപ്പിക്കാനാണു് ശ്രമിച്ചിട്ടുള്ളതു്. അവ തന്നെ സൂചിപ്പിക്കുക മാത്രമേ ചെയ്തിട്ടുള്ളു. കൂടുതൽ പഠനത്തിനു് 7.11ൽ കൊടുത്തിരിക്കുന്ന ആധാരഗ്രന്ഥങ്ങളും ലേഖനങ്ങളും ഉപയോഗിക്കാവുന്നതാണു്.

10. സംഗ്രഹം

1. സമഷ്ടിയെ വിവിധ സംഘങ്ങളായി വിഭജിച്ചു് അതിൽ നിന്നു് ഒരു സാമ്പിളെടുത്തു്, സാമ്പിളിൽ ഉൾപ്പെട്ടവരുന്ന എല്ലാ വ്യക്തികളിൽ നിന്നും വിവരം ശേഖരിക്കുന്ന സമ്പ്രദായത്തിനു് സംഘ സാമ്പിളനം എന്നു പേർ പറയും.

2. സമഷ്ടിയുടെ സമാന്തരമാധ്യമായ $\bar{X}$ ന്റെ ആകലമായ $\bar{x}_c$ താഴെ പറയും പ്രകാരം നിർവചിക്കപ്പെട്ടിരിക്കുന്നു.

$$\bar{x}_c = \frac{1}{n}\sum_{i=1}^{n} \bar{x}_i$$

ഇവിടെ $\bar{x}_i$, സാമ്പിളിൽ ഉൾപ്പെട്ട സംഘങ്ങളിൽ i-ാമത്തേതിന്റെ സമാന്തരമാധ്യം. n സാമ്പിൾപരിമാണം. ഈ ആകലം അനഭിനതമാണ്.

3. $$V(\bar{x}_c) = \frac{N-n}{N-1}\,\frac{\sigma_b^2}{n}.$$

ഇവിടെ $$\sigma_b^2 = \frac{1}{N}\sum_{i=1}^{N} (\bar{X}_i - \bar{X})^2$$

$$\hat{V}(\bar{x}_c) = \frac{N-n}{N-1}\,\frac{s_b^2}{n}$$. ഇവിടെ $$s_b^2 = \frac{1}{n-1}\sum_{i=1}^{n} (\bar{x}_i - \bar{x}_c)^2$$

4. സംഘങ്ങളെ ഉപസംഘങ്ങളായും ഉപസംഘങ്ങളെ വീണ്ടും ഉപസംഘങ്ങളായും അങ്ങനെ k-ാമത്തെ ഘട്ടത്തിൽ പ്രാഥമിക വ്യക്തികളിൽ എത്തിച്ചേരത്തക്കവണ്ണം പുനർവിഭജിക്കുന്നു എന്നിരിക്കട്ടെ. ആദ്യഘട്ടസംഘങ്ങളിൽ നിന്നു് ഒരു സാമ്പിളെടുത്തു് അതിൽ ഉൾപ്പെട്ടു വരുന്ന സംഘങ്ങളിൽ നിന്നു് ഉപസംഘങ്ങളെടുത്തു് അങ്ങനെ k-ാമതു ഘട്ടത്തിൽ പ്രാഥമികവ്യക്തികളിൽ നിന്നു് സാമ്പിളെടുക്കുന്ന സമ്പ്രദായത്തിനു് വിവിധഘട്ട സാമ്പിളനം എന്നു പറയും.

5. രണ്ടു ഘട്ടത്തിലും പ്രതിസ്ഥാപനമില്ലാത്ത ലഘ്വയാദൃച്ഛിക സാമ്പിളനരീതി അവലംബിക്കുന്ന സാഹചര്യത്തിൽ $\bar{X}$ ന്റെ ആകലങ്ങൾ

(i) $$\bar{x}_M = \frac{1}{n}\sum_{i=1}^{n} \frac{1}{m_i}\sum_{j=1}^{m_i} x_{ij}$$

(ii) $$\bar{x}_M' = \frac{1}{nM}\sum_{i=1}^{n} M_i\,\bar{x}_i$$

(iii) $$\bar{x}_M'' = \frac{\sum_{i=1}^{n} M_i\,\bar{x}_i}{\sum_{i=1}^{n} M_i}$$

എന്നിവയാണ്. ഇവിടെ m_i ആദ്യ സാമ്പിളിൽ ഉൾപ്പെട്ട i-ാമതു് സംഘത്തിൽ നിന്ന് സാമ്പിളിലേക്ക് തിരഞ്ഞെടുക്കപ്പെട്ട അംഗങ്ങളുടെ എണ്ണം.

$$V(\bar{x}_M) = \left(\frac{1}{n} - \frac{1}{N}\right) S_b^2 + \frac{1}{nN} \sum_{i=1}^{N} \left(\frac{1}{m_i} - \frac{1}{M_i}\right) S_i^2$$

$$V(\bar{x}_M') = \left(\frac{1}{n} - \frac{1}{N}\right) S_b'^2 + \frac{1}{nN} \sum_{i=1}^{N} \left(\frac{1}{m_i} - \frac{1}{M_i}\right) S_i^2$$

$$V(\bar{x}_M'') = \left(\frac{1}{n} - \frac{1}{N}\right) S_b''^2 + \frac{1}{nN} \sum \frac{M_i^2}{\bar{M}^2} \left(\frac{1}{m_i} - \frac{1}{M_i}\right) S_i^2$$

M_i മൂല്യങ്ങൾ സമമാകുമ്പോൾ

$$V(\bar{x}_M) = \left(\frac{1}{n} - \frac{1}{N}\right) S_b^2 + \left(\frac{1}{m} - \frac{1}{M}\right) \frac{\bar{S}_w^2}{n}$$

5. സമഷ്ടിയിലെ ഓരോ അംഗത്തിനും സാമ്പിളിൽ ഉൾപ്പെട്ടുവാനുള്ള സംഭാവ്യത അതിൻെറ പരിമാണസൂചകമായ ഒരു അഭിലക്ഷണത്തിൻെറ മൂല്യത്തിനു ആനുപാതികമായിരിക്കത്തക്കവണ്ണമുള്ള സാമ്പിളനരീതിക്ക് പരിമാണാനുപാതിക സാമ്പിളനം എന്നു പറയും. ഈ മാതിരി സാമ്പിൾ തിരഞ്ഞെടുക്കുവാൻ സഞ്ചിത സംയുക്തസമ്പ്രദായം, ലാഹിരിസമ്പ്രദായം എന്നു് രണ്ടു രീതികൾ ഉപയോഗത്തിലുണ്ടു് സാമ്പിളിലെ i-ാമതു് അംഗത്തിനു് സാമ്പിളിൽ ഉൾപ്പെട്ടുവാനുള്ള സംഭാവ്യത p_i ആണെങ്കിൽ, $\bar{X}$ ൻെറ ആകലം,

$$x_p = \frac{1}{Nn} \sum_{i=1}^{n} \frac{x_i}{p_i} \quad \text{ആണ്.}$$

$$V(\bar{x}_p) = \frac{1}{N^2 n} \left[\sum_{j=1}^{N} \frac{X_j^2}{P_j} - X^2 \right]$$

$$\hat{V}(\bar{x}_p) = \frac{1}{n(n-1)} \left(\sum_{i=1}^{n} \frac{x_i^2}{N^2 P_i^2} - n\,\bar{x}_p^2 \right)$$

6. അന്വേഷകൻ തന്റെ സാമാന്യ ബുദ്ധിമാത്രം ഉപയോഗിച്ച് സമഷ്ടിയെ പ്രതിനിധീകരിക്കത്തക്കവണ്ണം തിരഞ്ഞെടുക്കുന്ന സാമ്പിളാണ് ക്വോട്ടാസാമ്പിൾ.

7. ഒരു സാമ്പിളെടുത്തു് അതിൽ നിന്നു് ചില ഉപവിവരങ്ങൾ ശേഖരിച്ച് അതിൻെറ അടിസ്ഥാനത്തിൽ ആ സാമ്പിളിൽ നിന്നു് തന്നെ മറെറാരു സാമ്പിളെടുത്തു് അങ്ങനെ എത്രയെങ്കിലും ഘട്ടങ്ങൾ വരെ പോകുന്ന സാമ്പിളനരീതിക്ക് വിവിധ പ്രാവസ്ഥാ സാമ്പിളനം എന്നു പറയുന്നു. ആദ്യം ശേഖരിക്കുന്ന

വിവരങ്ങൾ സാമ്പിളിൻെറ സ്തരണത്തിന് ഉപയോഗിച്ച് രണ്ടാംഘട്ടത്തിൽ സ്തരിതയാദൃച്ഛികസാമ്പിളെടുക്കുന്ന രീതി സർവസാധാരണമത്രെ.

11. ചില സഹായകഗ്രന്ഥങ്ങളും ലേഖനങ്ങളും

1. Sampling Theory and Methods (Chapter 8&9) M. N. Murthy, Statistical Publishing Society, Calcutta.
2. A short manual of Sampling (1960) Vol I, United Nations, NewYork.
3. Sampling Theory of Surveys with Applications (chapter 6) Sukhathme. P. V. Indian Society of Agricultural Statistics.
4. On Large Scale Sample Surveys (1944), Mahalanobis P. C. Phil. Trans. Roy Soci, 231 (B), 329-451.
5. On double Sampling for p p s estimation (1964), Des Raj, Annah of Math. stat, 35, 900-902
6. Technical paper on some aspects of the developement of the Sample design (1954), Lahiri D. B., Sankya, 14, 264-316
7. On Sampling with varying probabilities in Sub-Sampling designs (1961), Rao J. N. K, J. Ind Soc. Agr: Stat, 13, 211-217.

അഭ്യാസം 7

7.1 ഒരു കർഷക കുടുംബത്തിൽ ശരാശരി എത്ര കന്നുകാലികളുണ്ടെന്ന് ആകലനം ചെയ്യാനായി ആകെയുള്ള 112 ഗ്രാമങ്ങളിൽ നിന്ന് 10 ഗ്രാമങ്ങൾ ലഘ്വയാദൃച്ഛിക സാമ്പിളനരീതിയിൽ തിരഞ്ഞെടുത്തു. ഓരോ ഗ്രാമത്തിലും 30 കർഷക കുടുംബങ്ങൾ വീതമാണ് ഉള്ളത്. സാമ്പിൾഗ്രാമങ്ങളും അവ ഓരോന്നിലുമുള്ള കന്നുകാലികളുടെ എണ്ണവും താഴെ കൊടുക്കുന്നു. ഒരു കർഷക കുടുംബത്തിലെ കന്നുകാലികളുടെ ശരാശരി എണ്ണവും, അതിൻെറ സാമ്പിളനപ്രസരണവും ആകലനം ചെയ്യുക.

ഗ്രാമം	കന്നുകാലികളുടെ എണ്ണം
1	83
2	98
3	64
4	85
5	130
6	73
7	48
8	69
9	53
10	92

7.2 ഒരു ഇംഗ്ലീഷ് മലയാള നിഘണ്ടുവിൽ എത്ര വാക്കുകളുണ്ടെന്ന് ആകലനം ചെയ്യണം. ആകെയുള്ള 56 അക്ഷരങ്ങളിൽ നിന്ന് 10 അക്ഷരങ്ങൾ, പരിമാണാനുപാതികസാമ്പിളനരീതിയിൽ തിരഞ്ഞെടുത്തു. പരിമാണം ഓരോ അക്ഷരം കൊണ്ട് ആരംഭിക്കുന്ന വാക്കുകൾ എത്ര പേജുകളിലായി അച്ചടിച്ചിരിക്കുന്നു എന്നതുകൊണ്ടാണ് നിർണയിച്ചത്. നിഘണ്ടുവിൽ 1120 പേജുകളുണ്ട്. ഓരോ അക്ഷരം കൊണ്ട് ആരംഭിക്കുന്ന വാക്കുകൾ അച്ചടിച്ചിരിക്കുന്ന പേജുകളിൽ നിന്ന് ഒരു പേജ് പ്രതിസ്ഥാപനമില്ലാത്ത ലഘ്വയാദൃച്ഛികസാമ്പിളന രീതിയിൽ തിരഞ്ഞെടുത്ത് അതിലുള്ള വാക്കുകൾ എണ്ണിനോക്കി, കിട്ടിയ വിവരങ്ങൾ താഴെ ചേർക്കുന്നു. ആകെ വാക്കുകളുടെ എണ്ണവും അതിൻെറ സാമ്പിളനപ്രസരണവും ആകലനം ചെയ്യുക.

സാമ്പിൾ അക്ഷരങ്ങൾ	പേജുകളുടെ എണ്ണം.	തിരഞ്ഞെടുത്ത പേജിലെ വാക്കുകളുടെ എണ്ണം.
N	18	
F	40	30
J	5	43
U	21	27
A	60	40
D	50	35
S	112	28
C	101	40
L	75	31
M	88	57

7.3 ലഘുയാദൃച്ഛിക സാമ്പിളന രീതിയിൽ 10 പേജുകൾ തിരഞ്ഞെടുത്ത് ആകെ വാക്കുകളുടെ എണ്ണം ആകലനം ചെയ്യുന്നതും 7.2–ൽ കൊടുത്തിരിക്കുന്ന വിധത്തിലുള്ള സാമ്പിളിൽ നിന്നുള്ള ആകലനവും താരതമ്യപ്പെടുത്തുക.

7.4 അഭ്യാസം 7.2–ലെ ലക്ഷ്യം നേടുവാൻ ഒരു വിവിധ പ്രാവസ്ഥാ സാമ്പിളന രീതിയും ആകലന സമ്പ്രദായവും നിർദേശിക്കുക.

8

സാമ്പിളനേതരപിശകുകൾ

1. പ്രാരംഭം

നിരവധി സാമ്പിളനസമ്പ്രദായങ്ങളെപ്പററി ഈ ഗ്രന്ഥത്തിൽ ചർച്ച ചെയ്യുകയുണ്ടായി. എങ്ങനെ സാമ്പിളെടുക്കണം, എങ്ങനെ ആകലനം നിർവഹിക്കണം തുടങ്ങിയ കാര്യങ്ങളാണ് പ്രധാനമായി പരിഗണിക്കപ്പെട്ടത്. സമ്പിളനത്തിൻെറ ലക്ഷ്യം, സാമ്പിളിനെ അടിസ്ഥാനമാക്കി സമഷ്ടിയെപ്പററി ചില നിഗമനങ്ങളിൽ എത്തുക എന്നതാണല്ലോ. ഈ നിഗമനങ്ങൾ പൂർണമായും ശരിയായിരിക്കുമെന്ന് നാം പ്രതീക്ഷിക്കുന്നില്ല. ചില പിശകുകൾ സംഭവിക്കുക തന്നെ ചെയ്യും. ഇങ്ങനെ സംഭവിക്കുന്ന പിശകുകൾക്ക് സാമ്പിളനപിശകുകൾ എന്നാണ് പേർ പറയുക. സാമ്പിൾപരിമാണം വർധിക്കുന്നതനുസരിച്ച് ഈ പിശകിൻെറ സംഭാവ്യതയും കുറഞ്ഞുവരും. പൂർണമായ ഗണനത്തിൽ, അതായത് സാമ്പിൾപരിമാണം സമഷ്ടിയുടെ പരിമാണത്തോട് തുല്യമാവുമ്പോൾ, ഈ പിശക് ഇല്ലാതാകുകയും ചെയ്യും. ഇങ്ങനെ സംഭവിക്കുന്ന പിശക് എത്രയെന്ന് മനസ്സിലാക്കാൻ സഹായിക്കുന്ന ഒന്നാണ് ആകലങ്ങളുടെ സാമ്പിളനപ്രസരണം.

സാമ്പിളനശാസ്ത്രത്തിന് പ്രധാനമായി കൈകാര്യം ചെയ്യാനുള്ളത് ഈ വിഭാഗത്തിൽപ്പെട്ട പിശകാണ്. പക്ഷേ മററു ചില പിശകുകൾ അന്വേഷണഫലത്തെ ഗൗരവമായി ബാധിച്ചു എന്നുവരാം. അവ സാമ്പിളനശാസ്ത്രത്തിൻെറ പരിധിയിൽപ്പെട്ടവയല്ല എന്നു പറഞ്ഞ് ഒഴിഞ്ഞു മാറാനാവില്ല. അന്വേഷണഫലത്തിൻെറ സൂക്ഷ്മത ഉറപ്പു നൽകാനുള്ള ബാധ്യത സാമ്പിളനശാസ്ത്രജ്ഞനാണല്ലോ.

സാമ്പിളനേതരപിശകുകൾ എന്നു വ്യവഹരിക്കപ്പെടുന്ന ഈ പിശകുകൾ അന്വേഷണത്തിൻെറ ഏതു ഘട്ടത്തിലും ഉണ്ടാവാം. സമഷ്ടിയെ നിർവചിക്കുന്നിടത്ത് വന്നു പോകുന്ന പിശകുകൾ അന്വേഷണലക്ഷ്യത്തെ പ്രതികൂലമായി ബാധിച്ചെന്നു വരാം. നിർവചനത്തിൻെറ അവ്യക്തത മൂലം സമഷ്ടിയിൽ ഉൾപ്പെടുത്തേണ്ട പല വ്യക്തികളും അവഗണിക്കപ്പെട്ടു എന്നു വരാം. ഉൾപ്പെടുത്തരുതാ

ത്ത വ്യക്തികൾ ഉൾപ്പെടുത്തപ്പെട്ട എന്നും വരാവുന്നതാണ്. സാമ്പിളിലേയ്ക്ക് തിരഞ്ഞെടുക്കപ്പെട്ട എല്ലാ വ്യക്തികളെയും കണ്ടെത്താൻ കഴിഞ്ഞെന്നു വരികയില്ല. പലരും വിവരങ്ങൾ നൽകാൻ വിസമ്മതിക്കുകയോ തെററായ വിവരങ്ങൾ നൽകുകയോ ചെയ്തേക്കാം. വിവരശേഖരണത്തിനായി നിയോഗിക്കപ്പെട്ട അന്വേഷകൻ, ശ്രദ്ധക്കുറവു കൊണ്ടോ പരിശീലനക്കുറവു കൊണ്ടോ സ്വഭാവത്തിൻെറ പ്രത്യേകതകൾ കൊണ്ടോ തെററായ വിവരങ്ങൾ രേഖപ്പെടുത്തുകയോ വിവരം നൽകേണ്ട ആളിനെ നീരസപ്പെടുത്തി ക്ലടിക്കാഴ്ച പരാജയമാക്കി തീർക്കുകയോ ചെയ്തേക്കാം. വിവരശേഖരണത്തിൽ എന്തെങ്കിലും അളന്നു നോക്കുകയോ എണ്ണിനോക്കുകയോ ചെയ്യേണ്ട ആവശ്യം നേരിട്ടാൽ അവിടെയും പിശകുവരാൻ സാധ്യതയുണ്ട്. അന്വേഷകരുടെ പ്രവർത്തനം പരിശോധിക്കുന്നതിലുള്ള നടപടിക്രമത്തിൻെറ പോരായ്മകളും പിശകുകൾക്കു് കാരണമാവാം. അപഗ്രഥനഘട്ടത്തിലും പല തരത്തിലുള്ള പിശകുകൾക്കു് സാധ്യതയുണ്ട്. പകർത്തി എഴുതുന്നതിൽ തെററുകൾ സംഭവിക്കാവുന്നതാണ്. പട്ടികകൾ തയ്യാറാക്കുന്നിടത്തും ഗണനക്രിയകൾ ചെയ്യുന്നിടത്തും തെററുകൾ സംഭാവ്യമത്രെ. ഇങ്ങനെ സാമ്പിളനേതരപ്പിശകുകൾ നിരവധിയാണ്. ഇവയ്ക്കുള്ള ഒരു പ്രത്യേകത സാമ്പിൾപരിമാണം വർധിക്കുന്നതനുസരിച്ചു് ഇവയും വർധിക്കും എന്നതത്രെ ഇവയിൽ പ്രധാനപ്പെട്ട ചിലതും അവയ്ക്കുള്ള പരിഹാരമാർഗങ്ങളുമാണ് ഈ അധ്യായത്തിൽ ചർച്ചചെയ്യുന്നതു്. അടുത്ത കാലത്തു് ഇമ്മാതിരി പിശകുകളെപ്പററി ഗണിതശാസ്ത്രപരമായ ഒരു പഠനം ആരംഭിച്ചിട്ടുണ്ട്. പക്ഷേ അതു പ്രകാരമുള്ള ഒരു ചർച്ചയ്ക്കു് ഇവിടെ മുതിരുന്നില്ല.

2. അപൂർണമായ വിവരശേഖരണം

സമഷ്ടിയെ പൂർണമായി പരിഗണിക്കാതെ സാമ്പിളെടുക്കുക, സാമ്പിളിലെ എല്ലാ അംഗങ്ങളെയും കണ്ടെത്താൻ സാധിക്കാതെ വരിക, കണ്ടെത്തിയവർ തന്നെ വിവരശേഖരണത്തിൽ സഹകരിക്കാതിരിക്കുക തുടങ്ങി പല തരത്തിൽ വിവരശേഖരണം അപൂർണമായിതീരാവുന്നതാണ്. ഈ പ്രശ്നം രണ്ടാമധ്യായത്തിൽ ഭാഗികമായി പരിഗണിക്കുകയുണ്ടായി. അവിടെ ചർച്ച ചെയ്യാത്ത ചില കാര്യങ്ങളാണ് പ്രധാനമായും ഇവിടെ അവതരിപ്പിക്കുന്നതു്. സാമ്പിളന ഫ്രേമിൻെറ പൂർണത സൂക്ഷ്മത തുടങ്ങിയവ ഉറപ്പു വരുത്തുക, സാമ്പിളിൽ ഉൾപ്പെടുത്തിയ എല്ലാ അംഗങ്ങളിൽ നിന്നും വിവരം ശേഖരിക്കാൻ അക്ഷീണമായി പരിശ്രമിക്കുക തുടങ്ങി എല്ലാ കരുതലുകളും എടുത്താൽ തന്നെയും വിവരശേഖരണം പൂർണമാക്കാൻ കഴിഞ്ഞെന്നു വരികയില്ല. പല പ്രാവശ്യം അന്വേഷിച്ചാലും ചിലരെ കണ്ടെത്താൻ കഴിഞ്ഞില്ലെന്നു വരാം. കണ്ടെത്തിയവരിൽ പലരും വിവരം നല്കാൻ വിസമ്മതിച്ചെന്നു വരും. വിസമ്മതിക്കുന്നവരുടെ അനുപാതം എപ്പോഴും ഒന്നു തന്നെ ആയിരിക്കണമെന്നില്ല. അന്വേഷണവിധേയമായ വിഷയവുമായി ഇതു വളരെയേറെ ബന്ധപ്പെട്ടിരിക്കുന്നു. ഉദാഹരണമായി ലൈംഗികവിഷയങ്ങളെ പററിയുള്ള അന്വേഷണങ്ങൾ സാധാരണഗതിയിൽ വിവരം നൽകാൻ വളരെയേറെപ്പേരെ വിസമ്മതമുള്ളവരാക്കുന്ന ഒന്നാണ്. അതു പോലെ തന്നെ പ്രധാനപ്പെട്ട ഒന്നാണ്

വിവരം നൽകേണ്ട ആളിൻെറ പ്രായം, സ്ത്രീപുരുഷഭേദം എന്നിവ. വൃദ്ധന്മാരും സ്ത്രീകളും പൊതുവെ വിവരം നൽകാൻ വൈമനസ്യമുള്ളവരായി കാണപ്പെടുന്നു. അവരുടെ ഇടയിൽ വിസമ്മതത്തിന്റെ അനുപാതം വളരെ കൂടുതലാണ്. സാമ്പത്തികമായി മേലെക്കിടയിലുള്ളവരിലും വിസമ്മതക്കാരുടെ അനുപാതം കൂടുതലത്രെ. നഗരവാസികൾ ഗ്രാമീണരെ അപേക്ഷിച്ച് വിവരം നൽകാൻ താൽപര്യമില്ലാത്തവരാണെന്നാണ് അനുഭവങ്ങൾ സൂചിപ്പിക്കുന്നത്.

ഇതിൽ നിന്നെല്ലാം വന്നുചേരുന്നത് മിക്കവാറും എല്ലാ വിവരശേഖരണവും അപൂർണമായിരിക്കുമെന്നാണല്ലോ. ഇത് ആകലങ്ങളുടെ വിശ്വാസ്യതയെ ബാധിക്കുമെന്നതിനു സംശയമില്ല. അതുകൊണ്ട് ഈ പ്രശ്നം ഗൗരവമായി പരിഗണിക്കേണ്ട ഒന്നാണ്. ആദ്യമായി അപൂർണതയുടെ അളവ് ആകലനം ചെയ്യുകയാണ് ചെയ്യേണ്ടത്. അതായത്, എത്ര ശതമാനം വ്യക്തികളിൽ നിന്ന് വിവരം കിട്ടാതുണ്ട് എന്ന് ആകലനം ചെയ്യുക. സാമ്പിൾപരിമാണം ചെറുതാണെങ്കിൽ ഇത് കൃത്യമായിത്തന്നെ കണ്ടുപിടിക്കാം. പക്ഷേ വലിയ സാമ്പിളുകളെ സംബന്ധിച്ചിടത്തോളം ആകലങ്ങളെ ആശ്രയിക്കുകയെ നിർവാഹമുള്ളു. പലപ്പോഴും സാമ്പിളനം ആസൂത്രണം ചെയ്യുന്ന ഘട്ടത്തിൽ തന്നെ ഇതിനെപ്പററി സാമാന്യമായെങ്കിലും മനസ്സിലാക്കാൻ കഴിഞ്ഞാൽ അത് അത്യന്തം പ്രയോജനപ്രദമായിരിക്കും. ഇതിനു് അവലംബിക്കാവുന്ന മാർഗം ഉപസാമ്പിളനമാണ്. സാമ്പിളിൽ ഉൾപ്പെടുത്തേണ്ട വ്യക്തികളെ തിരഞ്ഞെടുത്തതിനുശേഷം അതിൽ നിന്ന് ഒരു സാമ്പിളെടുക്കുക. ഈ ഉപസാമ്പിളിലെ വ്യക്തികളിൽ നിന്ന് ആദ്യം വിവരശേഖരണം നടത്തുക. വിവരശേഖരണത്തിൻെറ അടുത്ത ഘട്ടത്തിൽ ഈ വ്യക്തികളെ ഒഴിവാക്കാമെന്നതുകൊണ്ട് ഇത് ഒരു നഷ്ടമായി പരിഗണിക്കേണ്ടതില്ല. ഈ ഉപസാമ്പിളിനെ അടിസ്ഥാനമാക്കി കണ്ടെത്താൻ സാധിക്കാതെ വരാവുന്ന വ്യക്തികളുടെയും വിവരം നൽകാൻ വിസമ്മതിച്ചേക്കാവുന്ന വ്യക്തികളുടെയും അനുപാതം ആകലനം ചെയ്യുക. അതിൻെറ വിശ്വാസ്യതാന്തരാളങ്ങൾ നിർണയിക്കുക. ഈ അനുപാതങ്ങൾ വളരെ ചെറുതാണെന്നു കണ്ടാൽ അവയെ അവഗണിക്കാം. ലഭ്യമായ വിവരങ്ങൾ ഉപയോഗിച്ച് ആകലനം നിർവഹിക്കുകയും സമഷ്ടിയെപ്പററി നിഗമനങ്ങളിൽ എത്തുകയും ചെയ്താൽ വലിയ തകരാറൊന്നും സംഭവിക്കാൻ ഇടയില്ലെന്നു സാരം. പക്ഷേ ഈ അനുപാതങ്ങളുടെ പരിമാണം വർധിക്കുന്തോറും അതു് കുറയ്ക്കുവാനുള്ള നടപടികൾ ആസൂത്രണം ചെയ്യേണ്ടതിൻെറ ആവശ്യവും വർധിക്കുന്നു.

വിവരം നൽകാൻ വിസമ്മതിക്കുന്നവരുടെ അനുപാതം കുറയ്ക്കുവാനുള്ള ചില പരിപാടികൾ രണ്ടാമധ്യായത്തിൽ സൂചിപ്പിക്കുകയുണ്ടായി. അതിൽ ഏററവും പ്രധാനപ്പെട്ടതു് അന്വേഷകൻെറ പ്രായോഗികബുദ്ധിയും കഴിവുമാണ്. വിവരം നൽകുന്ന ആൾ ആരാണെന്നുള്ളതു് രഹസ്യമായി സൂക്ഷിക്കപ്പെടുമെന്നു് അയാൾക്ക് വിശ്വാസമുണ്ടാക്കാൻ കഴിഞ്ഞാൽ വിസമ്മതക്കാരുടെ എണ്ണം വളരെയധികം കുറയും. വിവരം നൽകാൻ വിസമ്മതിക്കുന്നതു് അയാളുടെ താൽപര്യ

ങ്ങൾക്കു തന്നെ ഹാനികരമാകുന്ന സാഹചര്യം സൃഷ്ടിക്കാൻ കഴിഞ്ഞാൽ അതും വിസമ്മതിക്കുന്നവരുടെ എണ്ണത്തിൽ കുറവു വരുത്തും. ഉദാഹരണമായി, ഗവണ്മെൻറ് തലത്തിൽ നടത്തുന്ന അന്വേഷണങ്ങളിൽ വിവരം നൽകാൻ വിസമ്മതിക്കുന്നതു് കുററകരമായി പ്രഖ്യാപിക്കാം. പൊതുസ്ഥാപനങ്ങളാണു് അന്വേഷണം സംഘടിപ്പിക്കുന്നതെങ്കിൽ അതിനോട്ട് സഹകരിക്കാത്തവർക്കു് ആ സ്ഥാപനത്തിൽ നിന്നു് ലഭിക്കാവുന്ന എല്ലാ ആനുകൂല്യങ്ങളും നിഷേധിക്കാം. പക്ഷേ ഈ മാതിരി നടപടികൾ സംഘടിതമായ എതിർപ്പിൽ എത്തിച്ചേരാതിരിക്കാൻ ശ്രദ്ധിക്കണമെന്നുമാത്രം. കണ്ടെത്താൻ കഴിയാതെ വരിക എന്ന പ്രശ്നം പരിഹരിക്കാനുള്ള ഒരു മാർഗം ആവർത്തിച്ചുള്ള സന്ദർശനമാണു്. ആദ്യസന്ദർശനത്തിൽ കാണാൻ കഴിയാത്ത ആളിനെ അടുത്ത സന്ദർശനത്തിൽ കണ്ടെത്തുവരാം. പക്ഷേ ഈ മാർഗം വ്യയഹേതുകമാണെന്നു പറയാതെ തരമില്ല. അനാവശ്യമായ സമയനഷ്ടത്തിനും ഇതു് ഇടയാക്കാം. അന്വേഷണം സംഘടിപ്പിക്കുന്നതു് ഭരണകൂടമോ അതുപോലെ അധികാരശക്തിയുള്ള മററു സ്ഥാപനങ്ങളൊ ആണെങ്കിൽ ആദ്യസന്ദർശനത്തിൽ കണ്ടെത്താൻ കഴിയാത്ത വ്യക്തികൾ അന്വേഷകനെ കണ്ടുപിടിച്ചു് വിവരം നൽകണമെന്നു് നിർബന്ധിക്കാൻ സാധിച്ചേക്കാം. പക്ഷേ ഇതിനു് പല പ്രായോഗികവൈഷമ്യങ്ങളും ഉണ്ടെന്നുള്ള വസ്തുത വിസ്മരിക്കാനാവില്ല.

പ്രതികരണരാഹിത്യത്തെ കൈകാര്യം ചെയ്യാൻ നിർദേശിക്കപ്പെട്ടിട്ടുള്ള ഏതാനും നടപടിക്രമങ്ങൾ താഴെ കൊടുക്കുന്നു.

(1) പ്രതികരണരാഹിത്യഫലങ്ങളുടെ ആകലനം

പ്രതികരണരാഹിത്യത്തിൻെറ അനുപാതം ആകലനം ചെയ്തു് അതു കൂടി സാമ്പിളന റിപ്പോർട്ടിൽ ഉൾപ്പെടുത്തുക എന്നതു് ഇപ്പോൾ സർവസാധാരണമായി അംഗീകരിച്ചുവരുന്ന ഒരു നടപടിക്രമമാണു്. റിപ്പോർട്ടു് ഉപയോഗിക്കുന്ന ആളുകൾക്കു് വിലപ്പെട്ട ഒരു വിവരമത്രെ ഇതു്. പ്രതികരണരാഹിത്യം മൂലം ആകലങ്ങളിൽ വന്നിരിക്കാവുന്ന അഭിനതി എത്ര എന്നു് ആകലനം ചെയ്യാനും പരിശ്രമങ്ങൾ നടന്നു വരുന്നുണ്ടു്. പക്ഷേ, ഈ രംഗത്തു് ഭാഗികമായ വിജയം മാത്രമേ സിദ്ധിച്ചിട്ടുള്ളു. സമഷ്ടിയുടെ സമാന്തരമാധ്യമാണു് ആകലനം ചെയ്യേണ്ട പ്രാചലമെന്നിരിക്കട്ടെ. സാമ്പിൾ സമാന്തരമാധ്യമാണു് ആകലമായി ഉപയോഗിക്കുന്നതെന്നു സങ്കൽപിക്കുക. n അംഗങ്ങളുള്ള ഒരു സാമ്പിളാണു് എടുത്തതെന്നും അതിൽ n' അംഗങ്ങളിൽ നിന്നേ വിവരം ശേഖരിക്കാൻ കഴിഞ്ഞുള്ളു എന്നും ഇരിക്കട്ടെ. $x_1, x_2, \cdots x_{n'}$ ആണു് അവയുടെ അഭിലക്ഷണമൂല്യങ്ങൾ എന്നു സങ്കൽപിക്കുക. ബാക്കി വ്യക്തികളുടെ അജ്ഞാതങ്ങളായ അഭിലക്ഷണമൂല്യങ്ങൾ $x_{n'+1}, x_{n'+2}, \cdots x_n$ എന്നിവയാണെങ്കിൽ

$$\bar{x} = \frac{x_1 + x_2 + \cdots + x_{n'} + x_{n'+1} + x_{n'+2} + \cdots + x_n}{n}$$

പക്ഷേ നാം സാമ്പിളിൽ നിന്നു് ഗണിച്ചെടുക്കുന്ന സമാന്തരമാധ്യം $\bar{x}'$ എന്നു സങ്കൽപിച്ചാൽ,

$$\bar{x}' = \frac{x_1+x_2+\cdots+x_n'}{n'} \quad \text{അത്രെ.}$$

വിവരം ലഭിക്കാതെ വന്ന വ്യക്തികളെ വീണ്ടും സമീപിച്ചു് ഒരു ശ്രമം കൂടി നടത്തിയാൽ ഒരുപക്ഷേ കുറെ വ്യക്തികളിൽ നിന്നു കൂടി വിവരശേഖരണം നടത്താൻ കഴിഞ്ഞെന്നു വരാം. അങ്ങനെ ആകെ n'' വ്യക്തികളിൽ നിന്നു് വിവരം ലഭിച്ചെന്നിരിക്കട്ടെ. ഇതിനെ അടിസ്ഥാനമാക്കി ഗണിക്കുന്ന സാമ്പിൾ സമാന്തരമാധ്യം $\bar{x}''$ എന്നിരിക്കട്ടെ,

$$\bar{x}'' = \frac{x_1+x_2+\quad+x_n'+x_n'+_1+\cdots x_n''}{n''}$$

ഈ പദ്ധതി ആവർത്തിച്ചാൽ $\bar{x}'''$, $\bar{x}''''$ തുടങ്ങി കൂടുതൽ കൂടുതൽ വ്യക്തികളിൽ നിന്നുള്ള വിവരങ്ങളെ അടിസ്ഥാനമാക്കിയുള്ള സമാന്തരമാധ്യങ്ങൾ ലഭിക്കും. ഇവയുടെ പ്രവണത ഉപയോഗപ്പെടുത്തി n അംഗങ്ങളിൽ നിന്നു് വിവരം ശേഖരിച്ചിരുന്നെങ്കിൽ കിട്ടാമായിരുന്ന $\bar{x}$ ൻെറ മൂല്യം ആകലനം ചെയ്യാവുന്നതാണു്. ഇതു് ലേഖ ഉപയോഗിച്ചോ, ഗണിതപ്രക്രിയകൾ ഉപയോഗപ്പെടുത്തിയോ നിർവഹിക്കാവുന്നതത്രെ. ഇതാണു് ഒരു സമീപനരീതി. ഗണിതമാതൃകകൾ ഉപയോഗിച്ചുള്ള മററു പല സമ്പ്രദായങ്ങളും ഗവേഷണവിധേയമാക്കപ്പെട്ടുവരുന്നു.

(2) പ്രതികരണരഹിതങ്ങളായ വ്യക്തികളുടെ പ്രതിസ്ഥാപനം

കണ്ടുപിടിക്കാൻ സാധിക്കാത്ത വ്യക്തികളെയും വിവരം നൽകാൻ വിസമ്മതിക്കുന്നവരേയും ഉപേക്ഷിച്ചിട്ടു് പകരം അത്രയും പുതിയ വ്യക്തികളെ സാമ്പിളിൽ ഉൾപ്പെടുത്തുകയാണു് ഈ സമ്പ്രദായം. പ്രഥമദൃഷ്ട്യാ വളരെ നല്ലതു് എന്നു തോന്നാവുന്ന ഈ പരിഹാരമാർഗം അത്ര കണ്ടു് അഭികാമ്യമായ ഒന്നല്ലെന്നു് അൽപം ആലോചിച്ചാൽ മനസ്സിലാകും. വിവരം നൽകാൻ വിസമ്മതിക്കുകയോ കണ്ടെത്താൻ കഴിയാതെ വരികയോ ചെയ്യുന്ന വ്യക്തികൾക്കു് അവരുടേതായ ചില പ്രത്യേകതകളുണ്ടായിരിക്കും. അവർക്കു പകരം വിവരം ലഭിക്കാൻ സാധ്യതയുള്ള വ്യക്തികളെ സാമ്പിളിൽ ഉൾപ്പെടുത്തിയതുകൊണ്ടു് സ്ഥിതിഗതികൾ മെച്ചപ്പെടാനിടയില്ല. കാരണം, പുതുതായി സാമ്പിളിൽ ഉൾപ്പെടുത്തിയ വ്യക്തികൾക്കു്, പ്രതികരണരഹിതങ്ങളായ വ്യക്തികളെക്കാൾ സാമ്യം വിവരശേഖരണത്തിനു് ബുദ്ധിമുട്ടുണ്ടാക്കാതിരുന്ന ബാക്കി വ്യക്തികളോടായിരിക്കും. അതുകൊണ്ടു് ഇതു് പ്രതികരണരാഹിത്യത്തിനു് ശരിയായ ഒരു പ്രതിവിധിയാവുകയില്ല.

ഇവിടെ സാമ്പിളിൽ ഉൾപ്പെടുത്തപ്പെട്ട വ്യക്തികളെ വിവിധ വിഭാഗങ്ങളായി വിഭജിക്കാം. പ്രതികരണരാഹിത്യത്തിൻെറ അളവു്, അഭിലക്ഷ്യമൂല്യങ്ങളുടെ പരിമാണം എന്നിവയാണു് ഈ വിഭജനത്തിനു് മാനദണ്ഡങ്ങളായി

സ്വീകരിക്കാവുന്ന അടിസ്ഥാനവസ്തുതകൾ. ഈ ഓരോ ഉപവിഭാഗത്തിൽ നിന്നും ശേഖരിക്കപ്പെടുന്ന അഭിലക്ഷണമൂല്യങ്ങളിൽ നിന്നു് പ്രതികരണക്രമങ്ങളുടെ വിപരീതാനുപാതത്തിലുള്ള ഭാരങ്ങൾ ഉപയോഗിച്ചു് ഭാരിതങ്ങളായ ആകലങ്ങൾ എടുക്കുകയാണെങ്കിൽ പ്രതിസ്ഥാപനം ചെയ്യുന്നതിൽ നിന്നു് ലഭിക്കാവുന്ന പ്രയോജനം ഏകദേശമായി നേടിയെടുക്കാവുന്നതാണ്. പക്ഷേ, ഇതു് ആകലനത്തിൻെറ ലാളിത്യം നശിപ്പിക്കുമെന്നു് മാത്രമല്ല, അഭിനതിയെ ക്ഷണിച്ചുവരുത്തുകയും ചെയ്യുന്ന വരാം.

3. സാമ്പിളനേതരപ്പിശകുകളുടെ അഭിജ്ഞാനവും ആകലനവും

വിവരശേഖരണ ഘട്ടത്തിൽ സംഭവിക്കാവുന്ന ചില സാമ്പിളനേതരപ്പിശകുകളെപ്പററിയാണ് നാമിവിടെ ചർച്ചചെയ്തതു്. 8.1 ൽ സൂചിപ്പിച്ചതുപോലെ പിശകുകൾ അപഗ്രഥനഘട്ടത്തിലും സംഭവിക്കാവുന്നതാണ്. വൻതോതിലുള്ള സാമ്പിളനത്തിൽ ഈ വിധത്തിലുള്ള പിശകുകൾ പൂർണമായി ഒഴിവാക്കാൻ കഴിഞ്ഞെന്നു വരികയില്ല. അവയുടെ സംഭാവ്യത കഴിവതും കുറയ്ക്കാനുള്ള നടപടികൾ സ്വീകരിക്കുകയും, സംഭവിച്ചേക്കാവുന്ന പിശകിൻെറ അളവു് ആകലനം ചെയ്യുകയുമാണു് നമുക്കു് ചെയ്യാനുള്ള കാര്യങ്ങൾ. ഈ കാര്യം സാമ്പിളന ആസൂത്രണം ചെയ്യുന്ന ഘട്ടത്തിൽ തന്നെ പരിഗണിക്കേണ്ടതാണ്. സാമ്പിളനവുമായി ബന്ധപ്പെട്ട എല്ലാ വിഭാഗത്തിൽപ്പെട്ട ജോലിക്കാർക്കും അതിനനുസൃതമായ നിർദേശങ്ങൾ നൽകുകയും വേണം.

ഇവിടെ ഒരു സംഗതി പ്രത്യേകം ഓർമയിരിക്കണം. താരതമ്യേന വലിയ സാമ്പിളുകളെടുത്തെങ്കിൽ മാത്രമേ സാമ്പിളനപ്പിശകു് കുറഞ്ഞിരിക്കുകയുള്ളു. സാമ്പിൾ പരിമാണം വർധിക്കുമ്പോൾ സാമ്പിളനേതരപിശകും വർധിക്കും. അതുകൊണ്ടു് സാമ്പിളനേതരപ്പിശകു് കുറയ്ക്കുവാനുള്ള ചെലവും പരിശ്രമവും സാമ്പിളനപ്പിശകു് കുറയുന്നതനുസരിച്ചു് വർധിച്ചുവരും. അതുകൊണ്ടു് സാമ്പിളനപ്പിശകും സാമ്പിളനേതരപ്പിശകും, ആകെ ചെലവാക്കാൻ സാധിക്കുന്ന പണവും കണക്കിലെടുത്തുവേണം സാമ്പിളനേതരപ്പിശകുകൾ കുറയ്ക്കുവാനുള്ള പരിപാടികൾക്കു് രൂപം നൽകാൻ. എന്താവശ്യത്തിനു വേണ്ടിയാണോ അന്വേഷണഫലങ്ങൾ ഉപയോഗിക്കുവാൻ ഭാവകുന്നതു് അതിനു് മതിയാവുന്നത്ര സൂക്ഷ്മത ലക്ഷ്യം വെച്ചു മതി പിശകുകളിൽ കുറവു വരുത്താൻ ശ്രമിക്കുന്നതു്. മറ്റൊരു തരത്തിൽ പറഞ്ഞാൽ, പിശകുകൾ കുറയ്ക്കുവാനുള്ള ചെലവും ബുദ്ധിമുട്ടും അതുകൊണ്ടുള്ള നേട്ടങ്ങളുമായി പൊരുത്തപ്പെട്ടു പോകണം. ഉദാഹരണമായി ഒരു സോപ്പു് നിർമാണ സ്ഥാപനത്തിനു് അവയുടെ ഉൽപന്നത്തിൻെറ പ്രതിമാസ വിൽപന എത്രയെന്നു് അറിയുവാൻ താൽപര്യമുണ്ടെന്നിരിക്കട്ടെ. ഓരോ മാസവും ശരാശരി എത്ര സോപ്പു വീതം ഉണ്ടാക്കണമെന്ന കാര്യത്തിൽ തീരുമാനമെടുക്കുകയാണു് അന്വേഷണലക്ഷ്യം. ഇതിനാവശ്യമായ വിവരങ്ങൾ വിവിധ വിപണനകേന്ദ്രങ്ങളിൽ നിന്നാണു് ശേഖരിക്കേണ്ടിയിരിക്കുന്നതു്. ഇവിടെ അതിരുകവിഞ്ഞ സൂക്ഷ്മതയൊന്നും ആവശ്യമില്ല. വലിയ പിശകില്ലാതെ വിവരം ലഭി

കുണമെന്നേയുള്ളൂ. ഇങ്ങനെയുള്ള ഒരു സാഹചര്യത്തിൽ പിശകുകൾ കുറയ്ക്കാൻ വ്യാമിശ്രമായ നടപടിക്രമങ്ങൾ അവലംബിച്ച് ഭീമമായ ചെലവും ബുദ്ധിമുട്ടും ഉളവാക്കുന്നതിനു് ന്യായീകരണമൊന്നുമില്ല.

സാമ്പിളനേതരപ്പിശകുകളുടെ സാന്നിധ്യം അറിയുന്നതിനും അതു് ആകലനം ചെയ്യുന്നതിനും അവലംബിക്കാവുന്ന ചില നടപടിക്രമങ്ങളാണ് താഴെ കൊടുക്കുന്നതു്.

(1) അവിരോധിതപരിശോധന

ഇതു് പ്രധാനമായും വിവരശേഖരണത്തിനായി നിയോഗിക്കപ്പെട്ടവർ അവരുടെ ജോലി തൃപ്തികരമായി നിർവഹിച്ചിട്ടുണ്ടോ എന്നു പരിശോധിക്കാൻ ഉപകരിക്കുന്ന ഒന്നാണ്. കേരളത്തിലെ ഒരു ഇടത്തരം കുടുംബത്തിന്റെ ശരാശരി ജീവിതച്ചെലവാണ് അന്വേഷണ വിധേയമായ വസ്തുത എന്നിരിക്കട്ടെ. കുടുംബങ്ങളാണ് ഇതിലെ പ്രാഥമികവ്യക്തികൾ. സാമ്പിളിൽ ഉൾപ്പെടുത്തപ്പെടുന്ന കുടുംബങ്ങൾ സന്ദർശിച്ച് ചോദ്യാവലി പൂരിപ്പിക്കുകയാണ് അന്വേഷകന്റെ ജോലി. എത്തിച്ചേരാൻ ബുദ്ധിമുട്ടുള്ള കുടുംബങ്ങളെപ്പററിയുള്ള വിവരങ്ങൾ സ്വന്തം വീട്ടിലിരുന്ന് സ്വയമേയോ മററാരോടെങ്കിലും ചോദിച്ചോ ചോദ്യാവലിയിൽ രേഖപ്പെടുത്തുവാനുള്ള പ്രവണത പല അന്വേഷകരും പ്രകടിപ്പിക്കാറുണ്ട്. ഇതു് പൂർണമായി തടയുക അത്ര എളുപ്പമല്ല. അന്വേഷകന്റെ പ്രവർത്തനങ്ങൾ പരിശോധിക്കാൻ നിയുക്തനായ ഉദ്യോഗസ്ഥനും എത്തിച്ചേരാൻ ബുദ്ധിമുട്ടുള്ള സ്ഥലങ്ങളിലേക്ക് പോയി എന്നു വരികയില്ല. ഇങ്ങനെയുള്ള സാഹചര്യങ്ങളിൽ പ്രയോജനകരമായി ഉപയോഗപ്പെടുത്താവുന്ന ഒരു നടപടിക്രമമാണ് അവിരോധിതപരിശോധന. സമഷ്ടിയെ സംബന്ധിക്കുന്ന ചില അനുപാതങ്ങൾ പലപ്പോഴും മററു മാർഗങ്ങളിൽ കൂടി അറിയാൻ കഴിയും. ഉദാഹരണമായി കേരളത്തിലെ ജനസംഖ്യയിൽ എത്ര ശതമാനം സ്ത്രീകളാണ് എന്നതു് സെൻസസ് കണക്കുകളിൽ നിന്നോ, സമീപകാലത്തു നടത്തപ്പെട്ട സാമ്പിളനങ്ങളിൽ നിന്നോ അറിയാം. മുൻ സൂചിപ്പിച്ച അന്വേഷണത്തിലെ ചോദ്യാവലിയിൽ 'ഓരോ കുടുംബത്തിലെയും അംഗസംഖ്യ, അതിൽ സ്ത്രീകളുടെ എണ്ണം' എന്ന ഒരു ചോദ്യം കൂടി ഉൾപ്പെടുത്തി എന്നിരിക്കട്ടെ. സാമ്പിൾ സാമാന്യം വലുതാണെങ്കിൽ അതിൽ നിന്നു് ലഭിക്കുന്ന, സ്ത്രീകളുടെ അനുപാതത്തിന്റെ ആകലം സമഷ്ടിയിലെ അനുപാതവുമായി വളരെയൊന്നും വ്യത്യസ്തമായിരിക്കുകയില്ല. പക്ഷേ ലഭിച്ച സാമ്പിളിൽ അതു് വളരെ വ്യത്യസ്തമായിട്ടാണ് കാണുന്നതെങ്കിൽ അതിന്റെ അർഥം അന്വേഷകൻ ശരിയായിട്ടല്ല വിവരം ശേഖരിച്ചതെന്നാണ്. സാമ്പിളിൽ നിന്നു ലഭിച്ച അനുപാതവും യഥാർഥ അനുപാതവുമായുള്ള വ്യത്യാസം സാമ്പിളനപ്പിശകിന്റെ പരിധിക്കപ്പുറമായി കണ്ടാൽ ഈ നിഗമനം സ്ഥിരീകരിക്കപ്പെടും. ഈ വ്യത്യാസത്തിന്റെ അളവ് സാമ്പിളനേതരപ്പിശകിന്റെ വലുപ്പത്തിന്റെ നിദർശകമത്രെ.

ചുരുക്കത്തിൽ, അവിരോധിതപരിശോധന എന്നതു കൊണ്ടു് നാം അർഥമാക്കുന്നതു് അറിവുള്ള ചില കാര്യങ്ങളെപ്പററിയുള്ള ചോദ്യങ്ങൾ കൂടി ചോദ്യാവ

ലിയിൽ ഉൾപ്പെടുത്തി, അതിൽ നിന്നു ലഭിക്കുന്ന വിവരങ്ങൾ യാഥാർത്ഥ്യവുമായി പൊരുത്തപ്പെടുന്നുണ്ടോ എന്നു പരിശോധിക്കുക എന്നതാണ്. വലിയ ചെലവും ബുദ്ധിമുട്ടും കൂടാതെ നിർവഹിക്കാവുന്ന ഒരു പരിശോധനയത്രെ ഇതു്. അന്വേഷകരായി നിയമിക്കപ്പെട്ടിട്ടുള്ളവരിൽ ആത്മാർത്ഥതയും സേവനസന്നദ്ധതയുമുള്ളവരെ തിരിച്ചറിയുവാനും ഈ പരിശോധന ഉപകരിക്കും.

(2) സാമ്പിളനപരീക്ഷ

സാമ്പിളനേതരപ്പിശകുകൾ കണ്ടുപിടിക്കുന്നതിനും ഭാഗികമായി അവ തിരുത്തുന്നതിനും ഉപകരിക്കുന്ന ഒരു സമ്പ്രദായമാണിതു്. സാമ്പിളിൽ ഉൾപ്പെടുത്തേണ്ട വ്യക്തികൾ ഏതെല്ലാമെന്നു് തീരുമാനിച്ചു്കഴിഞ്ഞാൽ അവരിൽ നിന്നു് വിവരശേഖരണം നിർവഹിക്കാനായി അന്വേഷകരെ നിയോഗിക്കുയാണല്ലൊ ചെയ്യുന്നതു്. ഭൂമിശാസ്ത്രപരമായും മററും സമഷ്ടിയെ വിവിധഭാഗങ്ങളായി ഭാഗിച്ചു് ഓരോ ഭാഗത്തുനിന്നും സാമ്പിളിൽ ഉൾപ്പെടുത്തിയിട്ടുള്ള വ്യക്തികളിൽ നിന്നുള്ള വിവരശേഖരം ഓരോ അന്വേഷകൻെറ ചുമതലയിലാക്കുകയാണു് സാധാരണ ചെയ്യുന്നതു്. വിവരശേഖരണഘട്ടത്തിൽ സംഭവിക്കാവുന്ന സാമ്പിളനേതരപ്പിശകുകൾ പ്രധാനമായും അന്വേഷകൻെറ കഴിവില്ലായ്മ, ശ്രദ്ധക്കുറവു്, ആത്മാർത്ഥതയില്ലായ്മ തുടങ്ങിയവ കൊണ്ടും ഓരോ അന്വേഷകനും നൽകപ്പെട്ടിരിക്കുന്ന വ്യക്തികളുടെയോ, ഭൂവിഭാഗത്തിൻെറയോ, പ്രതികരണരാഹിത്യം, അനഭിഗമ്യത മുതലായവ കൊണ്ടുമാണു് സംഭവിക്കുക. അതുകൊണ്ടു് പരിശോധനയെ സംബന്ധിച്ചിടത്തോളം ഓരോ അന്വേഷകനും ശേഖരിക്കുന്ന വിവരങ്ങൾ പ്രത്യേകം പരിഗണിക്കേണ്ടതു് ആവശ്യമത്രെ. ഓരോ അന്വേഷകനും വിവരശേഖരണം നടത്തിയ വ്യക്തികളിൽ നിന്നു് ഓരോ സാമ്പിളെടുത്തു്, കൂടുതൽ സമർത്ഥന്മാരും, വിശ്വസ്തന്മാരും, പരിചയസമ്പന്നരുമായ ഒരു കൂട്ടം അന്വേഷകരെക്കൊണ്ടു് ഈ പുതിയ സാമ്പിളുകളിലെ വ്യക്തികളിൽ നിന്നു് വിവരങ്ങൾ ശേഖരിപ്പിച്ചു്, നേരത്തെ ശേഖരിച്ചു വെച്ചിരിക്കുന്ന വിവരങ്ങളുമായി താരതമ്യപ്പെടുത്തുന്ന സമ്പ്രദായത്തിനു് സാമ്പിളനപരീക്ഷാരീതി എന്നു പേർ പറയുന്നു. ഏതെങ്കിലും അന്വേഷകൻെറ ചുമതലയിൽപ്പെട്ട ഭാഗത്തു നിന്നും രണ്ടു തരത്തിൽ കിട്ടിയ വിവരങ്ങളും ഒന്നു തന്നെയാണെങ്കിൽ അയാൾ ശേഖരിച്ച വിവരങ്ങളിൽ സാമ്പിളനേതരപ്പിശകു് കാര്യമായിട്ടല്ലെന്നു കരുതാം. നേരെമറിച്ചു്, ചെറിയ വ്യത്യാസങ്ങൾ കാണപ്പെട്ടാൽ നേരത്തെ ശേഖരിച്ചതിൽ തിരുത്തലുകൾ വരുത്തി ആ വിവരങ്ങൾ തന്നെ സ്വീകരിക്കാം. ഇവിടെ രണ്ടാമത്തെ സാമ്പിളിൽ ഉൾപ്പെട്ടുവന്ന വ്യക്തികളിൽ നിന്നു ലഭിച്ച വിവരങ്ങളിൽ മാത്രമേ തിരുത്തലുകൾ വരുത്തിയിട്ടുള്ളു എന്ന പോരായ്മയുണ്ടു്. പക്ഷേ സാമ്പിളനേതരപ്പിശകിൻെറ അളവു് അത്ര വലുതല്ലാത്തതു കൊണ്ടു് എല്ലാ വ്യക്തികളിൽ നിന്നും രണ്ടാമതു വിവരം ശേഖരിച്ചു് തെറ്റുകളെല്ലാം തിരുത്താൻ വേണ്ടിവരുന്ന ചെലവിനും അധ്വാനത്തിനും സമയനഷ്ടത്തിനും നീതീകരണമില്ല. പക്ഷേ ഏതെങ്കിലും അന്വേഷകൻെറ ചുമതലയിൽപ്പെട്ട വിഭാഗത്തിൽ നിന്നും രണ്ടാമതു ലഭിച്ച വിവരങ്ങൾ ആദ്യം ശേഖരിച്ച വിവരങ്ങളുമായി

ഗണ്യമായി വ്യത്യാസപ്പെട്ടിരുന്നാൽ ആ വിഭാഗത്തിലെ എല്ലാ വ്യക്തികളിൽ നിന്നും രണ്ടാമതു് വിവരശേഖരണം നടത്തിയെ തീരൂ.

അപഗ്രഥനഘട്ടത്തിലെ സാമ്പിളനേതരപ്പിശകുകൾ മനസ്സിലാക്കാനും പരിഹരിക്കാനും ഈ രീതി ഉപയോഗപ്പെടുത്താവുന്നതാണ്. അപഗ്രഥനത്തിനായി ഓരോരുത്തരേയും ഏൽപിച്ചിട്ടുള്ള വിവരങ്ങളിൽ നിന്നു് സാമ്പിളുകൾ എടുത്തു് അതിന്റെ അപഗ്രഥനത്തിൽ വന്നുപോയിട്ടുള്ള പിശകുകൾ കണ്ടുപിടിക്കുകയും, ആവശ്യമെന്നു കണ്ടാൽ മുഴുവൻ രണ്ടാമതു് അപഗ്രഥിക്കുകയും ചെയ്യണം. പിശകുകൾ അത്ര ഗൗരവതരമല്ലെങ്കിൽ അതിന്റെ അളവു് ആകലനം ചെയ്തു് അവസാനിപ്പിക്കാം. അനുവദിക്കാവുന്ന പിശകിനെപ്പററിയുള്ള സങ്കൽപമാണു് ഇവിടെയെല്ലാം നിർണായകമായ വസ്തുത.

വിവരശേഖരണം നടന്നുകൊണ്ടിരിക്കുമ്പോൾ തന്നെ സാമ്പിളനപരീക്ഷ ചെയ്യാവുന്നതാണ്. താമസം വന്നാൽ വിവരങ്ങളിൽ സമയമാററം കൊണ്ടുവരുന്ന വ്യത്യാസങ്ങൾ തെററിദ്ധാരണകൾക്കു് ഇട നൽകും. സാമ്പിളനപരീക്ഷ നടന്നുകൊണ്ടിരിക്കുന്നു എന്ന വസ്തുത അന്വേഷകർക്കു് കൂടുതൽ ശ്രദ്ധയോടെ വിവര ശേഖരണം നടത്താൻ പ്രേരകവുമത്രെ.

സാമ്പിളനപരീക്ഷ താരതമ്യേന ബുദ്ധിമുട്ടും പണച്ചെലവും ഉൾക്കൊള്ളുന്ന ഒരു നടപടിയാണ്. രണ്ടാം ഘട്ടത്തിൽ ശേഖരിച്ച വിവരങ്ങൾ, ആ വ്യക്തികളുടെ ആദ്യഘട്ടത്തിലെ വിവരങ്ങൾ തേടിപ്പിടിച്ചു് അവയുമായി താരതമ്യപ്പെടുത്തുന്നതും ശ്രമകരമായ ഒരു ജോലിയത്രെ. പക്ഷേ ഈ വൈഷമ്യങ്ങൾ തരണം ചെയ്താൻ, സാമ്പിളനേതരപ്പിശകിന്റെ സ്വഭാവവും, പരിമാണവും മനസ്സിലാക്കാൻ ഏററവും ഉപകരിക്കുന്ന ഒരു സമ്പ്രദായമാണു് ഇതു് എന്നു പറയാതെ തരമില്ല.

രണ്ടാം സാമ്പിളിൽ നിന്നു് ആകലങ്ങൾ ഗണിച്ചെടുത്തു്, അവയെ ആദ്യസാമ്പിളിലെ ആകലങ്ങളുമായി താരതമ്യപ്പെടുത്തുന്ന രീതിയും നിലവിലുണ്ടു്. ഇതു് സാമ്പിളനേതരപ്പിശകിന്റെ വലുപ്പം ആകലനം ചെയ്യാൻ ഉപകരിക്കുമെന്നല്ലാതെ, എവിടെയാണു പിശകു വന്നതെന്നു കണ്ടുപിടിക്കാൻ സഹായിക്കുകയില്ല.

മുൻപു് നടത്തിയിട്ടുള്ള സാമ്പിളനങ്ങളിൽ നിന്നു് ലഭിച്ച വിവരങ്ങളുമായി പുതിയ സാമ്പിളനത്തിലെ വിവരങ്ങൾ താരതമ്യപ്പെടുത്തുന്നതും സാമ്പിളനേതരപ്പിശകുകൾ മനസ്സിലാക്കുന്നതിനു് ഉതകും. അതുപോലെ തന്നെ ലഭ്യമായ റിക്കാർഡുകളിലെ വിവരങ്ങളും സാമ്പിളനത്തിൽ നിന്നു ലഭിച്ച വിവരങ്ങളും തട്ടിച്ചുനോക്കാം. ഉദാഹരണമായി കൈവശഭൂമികളെപ്പററിയുള്ള അന്വേഷണത്തിൽ റവന്യൂ റെക്കാർഡുകൾ താരതമ്യപഠനത്തിനു് ഉപകരിക്കും.

ചുരുക്കത്തിൽ സാമ്പിളനേതരപ്പിശകുകൾ മനസ്സിലാക്കുകയും പരിഹരിക്കുകയും ചെയ്യാൻ സന്ദർഭാനുസരണം പ്രസക്തമായ ഏതെങ്കിലും രീതി സ്വീകരി

ക്കുകയേ നിർവാഹമുള്ളു. സാമ്പിളനം കൈകാര്യം ചെയ്യുന്ന വ്യക്തികളുടെ പരിചയസമ്പത്തും, സാമാന്യബുദ്ധിയുമാണ് ഇതിൽ മാർഗദർശകങ്ങൾ.

4. സാമ്പിളനറിപ്പോർട്ട്

വിവരശേഖരണഘട്ടത്തിലും അപഗ്രഥനഘട്ടത്തിലും എന്നപോലെ റിപ്പോർട്ട തയ്യാറാക്കുന്നതിലും സാമ്പിളനേതരപ്പിശകുകൾക്ക് സാധ്യതയുണ്ട്. പ്രസക്തമായ എല്ലാ വിവരങ്ങളും ഉൾക്കൊള്ളാത്ത റിപ്പോർട്ടുകൾ വായനക്കാരിൽ തെറ്റിദ്ധാരണകൾ ഉളവാക്കിയേക്കാം. അതുപോലെ തന്നെ അശ്രദ്ധ കൊണ്ടും അനവധാനത കൊണ്ടും സംഭവിക്കാവുന്ന എല്ലാ പിശകുകളും ഒഴിവാക്കുകയും വേണം. റിപ്പോർട്ട തയ്യാറാക്കുമ്പോൾ പ്രത്യേകം ശ്രദ്ധ പതിപ്പിക്കേണ്ട ചില കാര്യങ്ങൾ താഴെ കൊടുക്കുന്നു.

1. എന്തു് ആവശ്യത്തിനു വേണ്ടിയാണ് അന്വേഷണം സംഘടിപ്പിച്ചതെന്ന് വ്യക്തമായി രേഖപ്പെടുത്തിയിരിക്കണം. അന്വേഷണഫലങ്ങൾ ഏതെല്ലാം തരത്തിലുള്ള ഉപയോഗത്തിനാണ് ഉദ്ദേശിക്കപ്പെട്ടിട്ടുള്ളതു് എന്നും വ്യക്തമാക്കിയിരിക്കണം.

2. പഠനവിധേയമായ സമഷ്ടിയെപ്പററി വ്യക്തമായ വിവരണം റിപ്പോർട്ടിലുണ്ടാവണം. സമഷ്ടിയിലെ വ്യക്തികൾ ഏതെല്ലാമെന്നു് അർഥശങ്കയ്ക്കു് ഇട നൽകാത്ത വിധം വിശദീകരിച്ചിരിക്കണമെന്നു സാരം.

3. ഏതെല്ലാം വിവരങ്ങളാണ് ശേഖരിച്ചതെന്ന് രേഖപ്പെടുത്തിയിരിക്കണം. ചോദ്യാവലിയും അന്വേഷകർക്കു നൽകിയിരുന്ന നിർദേശങ്ങളുടെ പകർപ്പും റിപ്പോർട്ടിൽ ഉൾക്കൊള്ളിക്കേണ്ടതു് ആവശ്യമത്രെ.

4. വിവരശേഖരണത്തിനു് അവലംബിച്ച മാർഗങ്ങൾ വ്യക്തമായി രേഖപ്പെടുത്തണം. അന്വേഷകർക്കു് അഭിമുഖീകരിക്കേണ്ടിവന്ന പ്രയാസങ്ങളും അവയെ തരണം ചെയ്യാൻ അവലംബിച്ച രീതികളും വിശദീകരിക്കുന്നതു കൊള്ളാം. പിന്നീട് അന്വേഷണം സംഘടിപ്പിക്കുന്നവർക്കു് അത്യന്തം വിലപ്പെട്ട ഒരു വിവരമാണ് ഇതു്.

5. സാമ്പിളനരീതി, ആകലനസമ്പ്രദായങ്ങൾ എന്നിവയെപ്പററിയുള്ള വിശദവിവരങ്ങൾ റിപ്പോർട്ടിലുണ്ടാവണം. സാമ്പിൾപരിമാണം, സാമ്പിളനഫ്രേമിൻെറ സ്വഭാവം തുടങ്ങിയവയും അതിലുണ്ടായിരിക്കണം.

6. ഏതു കാലയളവിനുള്ളിലാണ് അന്വേഷണം നടത്തപ്പെട്ടതെന്നുള്ളതു് രേഖപ്പെടുത്തേണ്ട മറെറാരു പ്രധാന വിവരമത്രെ.

7. അന്വേഷണഫലങ്ങൾ വ്യക്തമാക്കുന്ന തരത്തിലുള്ള പട്ടികകൾ, ശതമാനങ്ങൾ, അനുപാതങ്ങൾ തുടങ്ങിയവ തയ്യാറാക്കി റിപ്പോർട്ടിൽ ഉൾപ്പെടുത്തണം. ആവശ്യമെങ്കിൽ ചിത്രങ്ങളും ലേഖകളും ഉപയോഗിക്കാം. ആകലങ്ങളും വിശ്വാസ്യതാന്തരാളങ്ങളും ഗണിച്ചിരിക്കണം.

8. സൂക്ഷ്മതയെപ്പററി ആവശ്യമായ വിവരങ്ങൾ നൽകിയിരിക്കണം. സാമ്പിളനേതരപ്പിശകിൻെറ അളവും, അതു് കുറയ്ക്കുവാൻ അവലംബിച്ച നടപടികളും സൂചിപ്പിക്കാം. പ്രതികരണരാഹിത്യത്തെപ്പററിയുള്ള വിവരങ്ങൾ പ്രത്യേകിച്ചും റിപ്പോർട്ടിൽ ഉൾപ്പെടുത്തേണ്ടതാണു്.

9. അന്വേഷണത്തിനു് ആവശ്യമായി വന്ന ചെലവിൻെറ വിശദാംശങ്ങൾ കൂടി റിപ്പോർട്ടിൽ ചേർക്കണം. അന്വേഷണത്തിൻെറ വിവിധ ഘട്ടങ്ങൾ, ഓരോ ഘട്ടത്തിലും സഹകരിച്ച പ്രവർത്തകർ തുടങ്ങിയവയെല്ലാം റിപ്പോർട്ടിൽ ഉണ്ടാവണം.

10. സാമ്പിളനത്തിൻെറ ശാസ്ത്രീയപശ്ചാത്തലവും അതിനായി ഉപയോഗിച്ച ഗ്രന്ഥങ്ങളുടെയും മററും സൂചികയും റിപ്പോർട്ടിൻെറ അവസാനം ഉണ്ടായിരിക്കേണ്ടതത്രെ.

വായിക്കുന്നവർക്കു് അന്വേഷണത്തെയും അന്വേഷണഫലങ്ങളെയും പററി പൂർണ വിവരങ്ങൾ ലഭ്യമാകത്തക്കവണ്ണം വേണം റിപ്പോർട്ട തയ്യാറാക്കാൻ. സൂക്ഷ്മതയും, വ്യക്തതയും, പൂർണതയും അതിനു് ഉണ്ടായിരിക്കണം.

5. സംഗ്രഹം

1. സമഷ്ടിയുടെ നിർവചനത്തിൽ വരാവുന്ന അപൂർണ്ണത, വിവരശേഖരണത്തിനു് നിയോഗിക്കപ്പെട്ടവരുടെ കഴിവുകുറവു്, അശ്രദ്ധ മുതലായവകൊണ്ടുള്ള പിശകുകൾ, വിവരം നൽകേണ്ടവരെ കണ്ടെത്താൻ കഴിയാതെ വരികയോ കണ്ടെത്തിയവർ തന്നെ വിവരം നൽകാൻ വിസമ്മതിക്കുകയൊ ചെയ്യുക, അപഗ്രഥന ഘട്ടത്തിൽ ഗണനപരമായ തെററുകൾ സംഭവിക്കുക തുടങ്ങിയവാണു് സാമ്പിളനേതരപ്പിശകുകൾ.

2. എത്ര ശ്രമിച്ചാലും എല്ലാ വ്യക്തികളിൽ നിന്നും വിവരം ശേഖരിക്കാൻ കഴിഞ്ഞെന്നു വരികയില്ല. പലരേയും കണ്ടെത്താൻ കഴിഞ്ഞില്ലെന്നും പലരും വിവരം നൽകാൻ വിസമ്മതിച്ചു എന്നും വരാം. ഒരു ഉപസാമ്പിളെടുത്തു് പ്രതികരണരാഹിത്യത്തിൻെറ അനുപാതം ആകലനം ചെയ്യാം. അതു വലുതാണെങ്കിൽ സങ്കീർണ്ണങ്ങളായ നടപടികൾ സ്വീകരിച്ചു് അതു കുറയ്ക്കാൻ ശ്രമിക്കണം.

3. പ്രതികരണരാഹിത്യഫലങ്ങളുടെ ആകലനം, പ്രതികരണരഹിതങ്ങളായ വ്യക്തികളുടെ പ്രതിസ്ഥാപനം തുടങ്ങി പല നടപടികളും പ്രതികരണരാഹിത്യത്തിൻെറ ഫലങ്ങൾ ഒഴിവാക്കാനായി നിർദേശിക്കപ്പെട്ടിട്ടുണ്ടു്

4. സാമ്പിളനേതരപ്പിശകുകൾ മനസ്സിലാക്കുന്നതിനായി അവിരോധിതപരിശോധന, സാമ്പിളനപരീക്ഷ തുടങ്ങിയ മാർഗങ്ങൾ അവലംബിക്കാം.

5. സാമ്പിളനറിപ്പോർട്ടിൽ സാമ്പിളനത്തെപ്പററിയുള്ള എല്ലാ വിവരങ്ങളും അടങ്ങിയിരിക്കണം. വായിക്കുന്നവർക്കു് അന്വേഷണം സംഘടിപ്പിച്ച ഘട്ടം മുതൽ അന്തിമനിഗമനങ്ങൾ വരെയുള്ള എല്ലാ വിവരങ്ങളും വ്യക്തമായും പൂർണമായും മനസ്സിലാകത്തക്കവണ്ണമായിരിക്കണം അതു് തയ്യാറാക്കുക.

6. ചില സഹായകഗ്രന്ഥങ്ങളും ലേഖനങ്ങളും

1 Survey Sampling (Ch. 13), Leelie Kish, Wiley.

2 Sampling Theory and Methods. (Ch. 13) M. N. Murthy, Statistical Publishing Society, Calcutta.

3 Non-response and Call-backs in Surveys (1954), Durbin. J, Bull. Inter. Stat. Inst, 34 (2), 72-86

4 Non-sampling errors in Surveys (1952), Sukhatme P. V & Seth. G. R, J. Ind. Soc. Agr. Stat, 4, 5-41.

ചില മാതൃകാചോദ്യാവലികൾ

1. അന്വേഷണലക്ഷ്യം: കേരളത്തിലെ തൊഴിലില്ലായ്മയെപ്പററി പഠനം നടത്തുക.

ചോദ്യാവലി

A. ജില്ല താലൂക്ക്/നഗരം

വില്ലേജ്/വാർഡ്

B. വീട്ടനംബർ കുടുംബനാഥൻെറ പേർ

മതം-ജാതി.

C. വരുമാനം. വരുമാനമാർഗം

കുടുംബാംഗങ്ങൾ

ക്രമസംഖ്യ	1	2	3	4	5
1. കുടുംബനാഥനുമായുള്ള ബന്ധം					
2. സ്ത്രീയോ പുരുഷനോ ?					
3. വയസ്സ്					
4. വൈവാഹികസ്ഥിതി					
5. വിദ്യാഭ്യാസം:					
പൊതുവിദ്യാഭ്യാസം					
സാങ്കേതികവിദ്യാഭ്യാസം					
മററു യോഗ്യതകൾ.					
6. തൊഴിൽ ഉണ്ടോ ?					
7. സ്ഥിരം ജോലി— കാലയളവ്					
a) ഗവണ്മെൻറ് ഉദ്യോഗം—ശമ്പളം					
b) മററു് ഉദ്യോഗം—സ്വഭാവം—ശമ്പളം					
c) സ്വന്തം ജോലി—സ്വഭാവം—വരവ്					
d) പെൻഷൻ, വാടക					
8. സ്ഥിരമല്ലാത്ത ജോലി-കഴിഞ്ഞ വർഷം, എത്രകാലത്തേക്ക്—സ്വഭാവം—കൂലി					
9. ജോലി ഇല്ലെങ്കിൽ:					
a) എത്ര കാലമായി ഇല്ലാതായിട്ട്					
b) കാരണം					
c) കിട്ടാനാഗ്രഹിക്കുന്ന ജോലി					
d) ആഗ്രഹിക്കുന്ന ശമ്പളം/കൂലി					

തീയതി.................

അന്വേഷകൻെറ ഒപ്പ് ....

നിർദ്ദേശങ്ങൾ

1. ഒരേ അടുക്കളയിൽ നിന്നു് ആഹാരം കഴിക്കുകയും ഒന്നിച്ചു ജീവിക്കുകയും ചെയ്യുന്ന ആളുകളുടെ സമുച്ചയമാണു് കുടുംബമെന്നു നിർവചിച്ചിരിക്കുന്നതു്. ഒരു വീട്ടിൽ ഒന്നിലധികം കുടുംബങ്ങൾ താമസിക്കുന്നുണ്ടെങ്കിൽ ഓരോ കുടുംബത്തിനും പ്രത്യേകം പ്രത്യേകം ചോദ്യാവലികൾ ഉപയോഗിക്കണം. ആ വീട്ടു നമ്പർ തന്നെ എല്ലാ കുടുംബത്തിനും ഉപയോഗിക്കാം.

2. കുടുംബാംഗങ്ങൾ ആരാണു് അവരുടെ പ്രധാനി എന്നു് പറയുന്നോ അയാളാണു് കുടുംബനാഥൻ.

3. ഒന്നിലധികം വരുമാനമാർഗങ്ങളുണ്ടെങ്കിൽ അവയെല്ലാം രേഖപ്പെടുത്തണം. ആകെ വാർഷികവരുമാനമാണു് വരുമാനമായി കാണിക്കേണ്ടതു്. കുടുംബത്തിലെ എല്ലാ അംഗങ്ങൾക്കും കൂടിയുള്ള വരുമാനം വേണം കണക്കാക്കുവാൻ.

4. കുടുംബനാഥനുൾപ്പടെ അഞ്ചു് അംഗങ്ങളെപ്പററിയുള്ള വിവരങ്ങൾ എഴുതാനുള്ള സ്ഥലമാണു് ഓരോ ചോദ്യാവലിയിലും ഇട്ടിരിക്കുന്നതു്. കൂടുതൽ അംഗങ്ങളുണ്ടെങ്കിൽ മറെറാരു ചോദ്യാവലി കൂടി ഉപയോഗിക്കാം. ക്രമസംഖ്യയിൽ ആവശ്യമായ തിരുത്തലുകൾ വരുത്താൻ മറക്കരുതു്.

5. കുടുംബനാഥൻ മറ്റൊരിടത്തു് ജോലിയായി താമസിക്കുകയാണെങ്കിൽ, കുടുംബത്തിലെ അടുത്ത പ്രധാന വ്യക്തിയെ വേണം നാഥനായി കാണിക്കുവാൻ. വരുമാനത്തിൽ യഥാർഥ കുടുംബനാഥൻ അയച്ചുകൊടുക്കുന്ന പണവും ഉൾപ്പെടുത്താം.

6. കുടുംബാംഗങ്ങളെപ്പററിയുള്ള വിവരങ്ങൾ എഴുതാനുള്ള സ്ഥലത്തു് ആദ്യത്തെ കോളത്തിൽ കുടുംബനാഥനെപ്പററിയുള്ള വിവരങ്ങൾ എഴുതുക, 2, 3,... കോളങ്ങളിൽ പ്രാധാന്യമനുസരിച്ചു് ഓരോ അംഗത്തെപ്പററിയുള്ള വിവരങ്ങൾ രേഖപ്പെടുത്തുക.

7. കുടുംബാധിപനുമായുള്ള ബന്ധം എന്നിടത്തു് താഴെക്കൊടുത്തിരിക്കുന്ന പ്രതീകങ്ങൾ ഉപയോഗിക്കാം.

കുടുംബനാഥൻ	H
ഭാര്യ	W
അമ്മ	M
അച്ഛൻ	F
മകൾ	D
മകൻ	S
മരുമകൻ / മരുമകൾ	N
മകന്റെ മകൻ/മകന്റെ മകൾ മകളുടെ മകൻ/മകളുടെ മകൾ	G.S/G D
മററു ബന്ധങ്ങൾ	O
ബന്ധമില്ലാത്തവർ	U

8. സ്ത്രീയോ പുരുഷനൊ എന്നിടത്തു് 'സ്ത്രീ' അല്ലെങ്കിൽ 'പു' എന്നു് എഴുതുക.

9. വയസ്സു് എഴുതുമ്പോൾ ഏറ്റവുമടുത്ത ജന്മദിനത്തിൽ പൂർത്തിയായ അല്ലെങ്കിൽ പൂർത്തിയാകുന്ന വയസ്സു് എഴുതുക.

10. വൈവാഹികസ്ഥിതി എഴുതുമ്പോൾ അവിവാഹിതർക്കു് S എന്നും കുടുംബജീവിതം നയിക്കുന്നവർക്കു് M എന്നും ഭാര്യ/ഭർത്താവു് മരിച്ചവർക്കു് W എന്നും വിവാഹമോചനം നേടിയവർക്കു് D എന്നും എഴുതുക.

11. വിദ്യാഭ്യാസം– ഡിഗ്രി എടുത്ത വരെ സംബന്ധിച്ചു്, എടുത്തിട്ടുള്ള ഏറ്റവും ഉയർന്ന ഡിഗ്രി എഴുതുക. ലോവർ പ്രൈമറി വിദ്യാഭ്യാസം മാത്രമുള്ളവർക്കു് L.P എന്നും, അപ്പർ പ്രൈമറി വിദ്യാഭ്യാസം നേടിയവർക്കു് U.P എന്നും ഹൈസ്കൂൾ വിദ്യാഭ്യാസം ഉണ്ടെങ്കിൽ H.S എന്നും പ്രീഡിഗ്രി വരെ എത്തിയവർക്കു് P.D C എന്നും എഴുതുക.

12. സാങ്കേതികവിദ്യാഭ്യാസമുണ്ടെങ്കിൽ അതിൽ നേടിയിട്ടുള്ള പ്രധാന യോഗ്യതകൾ രേഖപ്പെടുത്തുക.

13. ശമ്പളം ഒരു വർഷത്തേക്കുള്ളതാണു് എഴുതേണ്ടതു്.

14. സ്ഥിരമല്ലാത്ത ജോലിയുടെ പൊതുസ്വഭാവവും അതിൽ നിന്നു് കഴിഞ്ഞ വർഷം ലഭിച്ച ആകെ വരുമാനവുമാണു് എഴുതേണ്ടതു്.

15. തൊഴിലില്ലാതായിട്ടു് എത്ര കാലമായി എന്നിടത്തു് ജോലിക്കു് യോഗ്യത നേടിയശേഷം കഴിഞ്ഞ വർഷാവസാനംവരെ എത്രകാലമായി എന്നെഴുതണം.

16. കിട്ടാനാഗ്രഹിക്കുന്ന ജോലിയും ശമ്പളവും, യോഗ്യതയുടെ അടിസ്ഥാനത്തിലുള്ളതായിരിക്കാൻ ശ്രദ്ധിക്കണം.

2. അന്വേഷണലക്ഷ്യം: ജനങ്ങളുടെ രാഷ്ട്രീയചിന്താഗതിയും അതിൽ സ്വാധീനം ചെലുത്തുന്ന ഘടകങ്ങളും

ചോദ്യാവലി

1. കഴിഞ്ഞ രണ്ടുമാസമായി കാണാതിരുന്ന ഒരു സുഹൃത്തിനെ കണ്ടാൽ നിങ്ങൾ അദ്ദേഹത്തോടു പറയുന്ന പ്രധാന കാര്യം എന്തായിരിക്കും. } ...

2. a) നിങ്ങൾ വർത്തമാനപ്പത്രങ്ങൾ വായിക്കാറുണ്ടോ? }
ദിവസവും
ചിലപ്പോഴെല്ലാം
ഒരിക്കലും ഇല്ല.

b) നിങ്ങൾ വായിക്കാറുള്ള വർത്തമാനപ്പത്രങ്ങൾ പൊതുജനങ്ങൾക്ക് എത്ര മാത്രം വിവരം നൽകുന്നുണ്ടെന്നാണ് അഭിപ്രായം. } നല്ലതു പോലെ / സാമാന്യം / മോശമായി

c) റേഡിയോവിലെ വാർത്താപ്രക്ഷേപണങ്ങൾ കേൾക്കാറുണ്ടോ ? } എന്നും / ചില ദിവസങ്ങളിൽ / വല്ലപ്പോഴും / ഒരിക്കലും ഇല്ല

d) വാർത്തകൾ അറിയാൻ മററെന്തു മാർഗങ്ങളാണ് നിങ്ങൾ ഉപയോഗപ്പെടുത്താറുള്ളത്. ? } ...

e) ലോകവാർത്തകൾ സാമാന്യം തൃപ്തികരമായി അറിയാൻ കഴിയുന്നുണ്ടോ ? } ...

3. a) രാഷ്ട്രീയപ്രശ്നങ്ങൾ ചർച്ച ചെയ്യാൻ ഇഷ്ടപ്പെടുന്നുണ്ടോ ? } വളരെയധികം / സാമാന്യം / തീരെയില്ല

b) നിങ്ങളുടെ സ്നേഹിതന്മാർക്കും നിങ്ങളുടെ രാഷ്ട്രീയാഭിപ്രായങ്ങൾ തന്നെ ഉണ്ടായിരുന്നാൽ കൊള്ളാമെന്ന് ആഗ്രഹിക്കുന്നുണ്ടോ ? } വളരെയധികം / അൽപമെങ്കിലും / ഒട്ടുമില്ല

c) നിങ്ങൾ രാഷ്ട്രീയസമ്മേളനങ്ങളിൽ പങ്കെടുക്കാറുണ്ടോ ? } മിക്കപ്പോഴും / ചിലപ്പോഴെല്ലാം / ഒരിക്കലുമില്ല

d) കഴിഞ്ഞ തിരഞ്ഞെടുപ്പിനെപ്പറ്റി ആരെങ്കിലുമായി ചർച്ച ചെയ്യാറുണ്ടോ ? ഉണ്ടെങ്കിൽ ആരുമായി ? } സഹപ്രവർത്തകരുമായി / സുഹൃത്തുക്കളുമായി / കുടുംബാംഗങ്ങളുമായി / അന്യരുമായി

4. a) കഴിഞ്ഞ തിരഞ്ഞെടുപ്പിൽ നിങ്ങൾ വോട്ടു ചെയ്തോ ? } ...

b) വോട്ടു ചെയ്തില്ലെങ്കിൽ എന്തുകൊണ്ട് ? } ...

5. a)	അടുത്ത തിരഞ്ഞെടുപ്പിൽ കഴിഞ്ഞ പ്രാവശ്യം വോട്ടു ചെയ്ത പാർടിക്കു തന്നെയാണോ വോട്ടു ചെയ്യുക.	... ...
b)	അല്ലെങ്കിൽ ആ മനസ്സുമാററത്തിൻെറ കാരണമെന്തു്?	......
6. a)	നിങ്ങൾ വോട്ടു ചെയ്യാൻ ഇഷ്ടപ്പെടുന്ന രാഷ്ട്രീയപാർടിയിൽ എന്തുമാത്രം വിശ്വാസമുണ്ടു്?	പൂർണ്ണവിശ്വാസം മറ്റുള്ളവയെക്കാൾ മെച്ചം വിശ്വാസമില്ല
b)	ഏതു കാര്യങ്ങൾക്കാണു് ഈ പാർടി കൂടുതൽ പ്രാധാന്യം നൽകുന്നതു്.?	......
c)	നിങ്ങളുടെ അഭിപ്രായത്തിൽ ഈ പാർടിയുടെ പരിപാടിയിൽ ഏററവും പ്രധാനപ്പെട്ട ഇനം ഏതാണു്?	... ...
d)	ഈ പാർടിയുടെ ഘടനയിൽ എന്തെങ്കിലും മാററങ്ങൾ വരുത്തണമെന്നു് ആഗ്രഹിക്കുന്നുണ്ടോ? ഉണ്ടെങ്കിൽ എന്തു്?	... ...
e)	ഈ പാർടിയിൽ ഇവയുടെ പ്രാധാന്യം ഏതു ക്രമത്തിലാണെന്നാണു് നിങ്ങളുടെ അഭിപ്രായം?.	തത്വം പരിപാടി നേതാക്കന്മാർ അച്ചടക്കം

7. a) താഴെ പറയുന്നവയിൽ ഏതെല്ലാം കാര്യങ്ങൾ പാർടിക്കുവേണ്ടി ചെയ്തിട്ടുണ്ടു്.?

1) പണം കൊടുക്കുക.

2) കൂടുതലാളുകളെ പാർടിയിൽ ചേർക്കുക.

3) പാർടിപ്രസിദ്ധീകരണങ്ങൾ വിൽക്കുക.

4) പാർടി സംഘടിപ്പിക്കുന്ന മീറ്റിംങ്ങുകളിൽ പ്രസംഗിക്കുക.

5) പാർടിഭരണത്തിൽ പങ്കാളിയാവുക.

6) മററു തരത്തിൽ പാർടിയെ വളർത്തുക.

b) നിങ്ങൾ വോട്ടുചെയ്ത പാർട്ടി നിങ്ങളിൽ നിന്നു് സഹായം ആവശ്യപ്പെടാറുണ്ടോ ? — മിക്കവാറും / ചുരുക്കമായി / ഒരിക്കലുമില്ല

8. a) നിങ്ങളുടെ പാർട്ടിക്കു് അടുത്ത തിരഞ്ഞെടുപ്പിൽ നിയമസഭയിൽ / പാർലമെൻറിൽ ഭൂരിപക്ഷം കിട്ടുമെന്നു് പ്രതീക്ഷയുണ്ടോ ? — നിയമസഭയിൽ / പാർലമെൻറിൽ / രണ്ടിടത്തുമില്ല

b) നിങ്ങളുടെ പാർട്ടി ഭരണത്തിൽ വരുന്നതോ പ്രതിപക്ഷമായി നിൽക്കുന്നതോ നല്ലതു്.

c) ഭൂരിപക്ഷമില്ലെങ്കിലും ഭരണം പിടിച്ചെടുക്കാൻ നിങ്ങളുടെ പാർട്ടിക്കു് ആഗ്രഹമുണ്ടോ ?

9. നിങ്ങളുടെ പാർട്ടിയുടെ ഔദ്യോഗിക ഭരണാധികാരിയിൽ വിശ്വാസമുണ്ടോ?

10. നിങ്ങൾ ഈ പാർട്ടിയുമായി ബന്ധപ്പെടാൻ താഴെ കൊടുത്തിരിക്കുന്ന കാരണങ്ങളിൽ ഏതെങ്കിലും പ്രേരകമായോ ?

1. നിങ്ങളുടെ അവകാശങ്ങൾ സംരക്ഷിക്കുക.
2. രാജ്യത്തിന്റെ അഭിവൃദ്ധിക്കായി പരിശ്രമിക്കുക.
3. പാർട്ടിയിലുള്ള സുഹൃത്തുക്കളുടെ പ്രേരണ
4. നിങ്ങളുടെ വർഗത്തോടുള്ള സ്നേഹം
5. രാഷ്ട്രീയകാര്യങ്ങളിലുള്ള താൽപര്യം
6. സമാധാനത്തിനു വേണ്ടി പ്രവർത്തിക്കാനുള്ള ആഗ്രഹം

11. a) ഒരു രാഷ്ട്രീയ പാർട്ടിക്കു് എന്തെല്ലാം ഗുണങ്ങൾ ഉണ്ടായിരിക്കണമെന്നാണു് നിങ്ങളുടെ ആഗ്രഹം ?

b) ഇന്ത്യയിലുള്ള ഏതെല്ലാം പാർട്ടികൾക്കു് ഈ ഗുണങ്ങളുണ്ടെന്നു് നിങ്ങൾ കരുതുന്നു.

c) ഏതെങ്കിലും പാർട്ടിയെ നിരോധിക്കുന്നതു് ശരിയാണോ ?

d) ഇന്ത്യയിൽ നിരോധിക്കപ്പെടേണ്ട പാർട്ടികൾ ഏതെല്ലാം ?

e) അടുത്തു തന്നെ തിരഞ്ഞെടുപ്പുണ്ടായാൽ നിങ്ങൾ താഴെപ്പറയുന്നതിൽ ഏതെല്ലാം കാര്യങ്ങൾ പരിഗണിച്ചിട്ടായിരിക്കും വോട്ടു ചെയ്യുക.

1. രാജ്യം അഭിവൃദ്ധിപ്പെട്ടുകൊണ്ടിരിക്കുന്നു.
2. വിലകൾ വർധിക്കുന്നു.
3. തൊഴിൽ കിട്ടാൻ പ്രയാസമാണു്.
4. സാമൂഹ്യ അസമത്വങ്ങൾ കുറഞ്ഞുകൊണ്ടിരിക്കുന്നു.
5. വിദേശനയം പരാജയമാണു്.
6. മററു കാരണങ്ങൾ.

12 a) ഇപ്പോഴത്തെ ഭരണാധികാരികളെപ്പറ്റി എന്താണു് അഭിപ്രായം ?
- ഭൂരിപക്ഷവും നല്ലവരാണു്.
- ഭൂരിപക്ഷവും ചീത്തയാളുകളാണു്.
- പ്രത്യേകം അഭിപ്രായമില്ല.

b) ഇപ്പോഴത്തെ ഭരണാധികാരികൾ നിങ്ങളുടെ താൽപര്യങ്ങൾക്കു് അനുകൂലമായാണോ പ്രതികൂലമായാണോ പ്രവർത്തിക്കുന്നതു്?

c) എങ്ങനെയുള്ളവരാണു് നല്ല ഭരണാധികാരികളായിരിക്കുക?
- അതിസമർത്ഥന്മാർ
- സാമാന്യം സമർത്ഥന്മാരും നീതിമാന്മാരും
- സാമർത്ഥ്യമില്ലാത്തവർ
- വഞ്ചകന്മാർ

13. പ്രശ്നങ്ങൾ പരിഹരിക്കാൻ ഏതാണു് നല്ല മാർഗം?
- പടിപടിയായുള്ള മാറ്റം.
- പെട്ടെന്നുള്ള മാറ്റം.

14. ഇപ്പോൾ നിലവിലുള്ള വ്യവസ്ഥിതിയെപ്പററി നിങ്ങളുടെ ഏററവും പ്രധാനപ്പെട്ട പരാതി എന്താണു്?

15. a) കമ്യൂണിസമാണോ പാർലമെന്ററി ജനാധിപത്യമാണോ നിങ്ങൾ ഇഷ്ടപ്പെടുന്നതു് ?

b) എന്തുകൊണ്ടു് ?

16. a) ഭരണഘടനയിൽ മാറ്റം വരുത്തണമോ?

b) എന്തു കൊണ്ടു് ?

17. a) ഇന്നത്തെ രീതിയിലുള്ള കോടതികളെ ഇഷ്ടപ്പെടുന്നോ ?

b) ഇല്ലെങ്കിൽ എന്തു കൊണ്ടു് ?

18. a) ഏകാധിപത്യത്തെ ഇഷ്ടപ്പെടുന്നോ ?

b) എന്തുകൊണ്ടു് ?

19. സ്വകാര്യസ്വത്തവകാശം ഒരു മൗലികാവകാശമല്ലാതാക്കുന്നതു് ഇഷ്ടപ്പെടുന്നോ ?

20. ഇന്ത്യയുടെ വിദേശനയത്തെപ്പററി എന്താണു് അഭിപ്രായം ?

21. കഴിഞ്ഞ തിരഞ്ഞെടുപ്പിൽ ഏതു പാർട്ടിക്കാണു് വോട്ടു ചെയ്തതു് ?

22. ഉടനെ ഒരു തിരഞ്ഞെടുപ്പുണ്ടായാൽ ഏതു പാർട്ടിക്കാണു് വോട്ടു ചെയ്യുക ?

23. മററു സമുദായക്കാരനായ ഒരു സ്ഥാനാർത്ഥിക്കു് വോട്ടു ചെയ്യാൻ വൈമനസ്യമുണ്ടോ ?

24. ഇന്ത്യയുടെ പ്രധാനമന്ത്രി ആരായിരിക്കണമെന്നാണു് നിങ്ങളുടെ അഭിപ്രായം ?

25. വിവരം നൽകുന്ന ആളിനെപ്പററിയുള്ള വിവരങ്ങൾ

1. സ്ത്രീപുരുഷഭേദം
2. വയസ്സു്
3. തൊഴിൽ

4. വാർഷികവരുമാനം
5. ജനനസ്ഥലം
6. മതം—ജാതി
7. വിദ്യാഭ്യാസനില

കുറിപ്പ്:-ഉത്തരങ്ങൾ നൽകിയിട്ടുള്ള ചോദ്യങ്ങളെ സംബന്ധിച്ച് സംഗതങ്ങളല്ലാത്ത ഉത്തരങ്ങൾ വെട്ടിക്കളയുക. മറ്റുള്ള ചോദ്യങ്ങൾക്കു ചുരുക്കത്തിലും വ്യക്തമായും ഉത്തരമെഴുതുക. വിവരം നൽകുന്ന ആളിന്റെ പേരോ മററു സൂചനകളോ നൽകേണ്ടതില്ല. നൽകുന്ന വിവരങ്ങൾ പരമരഹസ്യമായി സൂക്ഷിക്കുന്നതായിരിക്കും.

3 അന്വേഷണലക്ഷ്യം: കുടുംബച്ചെലവുകളെപ്പററി പഠനം നടത്തുക.

ചോദ്യാവലി

1. ജില്ല...... താലൂക്ക്...... വില്ലേജ്......

വീട്ടുനംബർ...... മതം... ..

പ്രധാനതൊഴിൽ ... മാസവരുമാനം ....

സന്ദർശനത്തീയതി

2. കുടുംബാംഗങ്ങളെപ്പററിയുള്ള വിവരങ്ങൾ

ക്രമസംഖ്യ	കുടുംബനാഥനുമായുള്ള ബന്ധം	സ്ത്രീയോ പുരുഷനോ	വൈവാഹികനില	വിദ്യാഭ്യാസനില	തൊഴിൽ		ആരോഗ്യനില	
					പ്രധാനതൊഴിൽ	ഉപതൊഴിൽ	സ്ഥിരം രോഗം	ഇടക്കിടെ ഉണ്ടാകുന്ന രോഗം
1	2	3	4	5	6	7	8	9
1								
2								
3								
4								
5								
6								
7								
8								
9								
10								

3. കുടുംബച്ചെലവുകൾ

A. ഭക്ഷണം(മാസച്ചെലവ്) ചെലവ് (പണമായി)

1. അരി
2. ഗോതമ്പ്
3. മററു ധാന്യങ്ങൾ
4. കപ്പ
5. പയറുവർഗങ്ങൾ
6. പച്ചക്കറികൾ
7. ഇറച്ചി
8. മീൻ
9. മുട്ട
10. എണ്ണ
11. പാലും ബന്ധപ്പെട്ട സാധനങ്ങളും
12. പഴങ്ങൾ
13. പലവ്യംജനങ്ങൾ
14. കാപ്പി, ചായ
15. വിറക്, ഇലക്ട്രിസിറ്റി
16. ഹോട്ടൽച്ചെലവ്

B. വീട്(വാർഷികച്ചെലവ്)

1. വീട്ടുവാടക
2. വീട്ടുപകരണങ്ങൾ
3. കരം
4. അററകുററപണി

C. തുണി(വാർഷികച്ചെലവ്)

1. പുരുഷന്മാർക്കുള്ള വസ്ത്രം
2. സ്ത്രീകൾക്കുള്ള വസ്ത്രം
3. കുട്ടികൾക്കുള്ള വസ്ത്രം
4. അലക്കുകൂലി
5. കിടക്ക, ഷീററ്

D. പലവക

a) മാസച്ചെലവ്

1. സോപ്പ്, പൗഡർ തുടങ്ങിയവ
2. ക്ഷൗരം തുടങ്ങിയവ
3. പോസ്റ്റേജ്
4. പത്രം, മാസിക
5. വിനോദം
6. ബീഡി, സിഗററ്റ്, മുറുക്കാൻ
7. കുടിസാധനങ്ങൾ

b) വാർഷികച്ചെലവ്

1. യാത്ര
2. ആഘോഷങ്ങൾ, സംഭാവന
3. ചികിൽസ
4. വിദ്യാഭ്യാസം

സ്ഥലം ....

തീയതി ....   അന്വേഷകന്റെ ഒപ്പ്

4. അന്വേഷണലക്ഷ്യം: സർവകലാശാലാ വിദ്യാർഥികളുടെ ജീവിതസാഹചര്യങ്ങളെപ്പററി പഠനം നടത്തുക.

ചോദ്യാവലി

1. വ്യക്തിപരമായ വിവരങ്ങൾ
   1. പേര്: ....
   2. പഠിക്കുന്ന കോളേജ്: ...
   3. a) ക്ലാസ് .. ... b) വിഷയം ......
   4. a) സ്ത്രീപുരുഷഭേദം ...... b) മാതൃഭാഷ ......
   5. വയസ്സ് ....
   6. a) ഉയരം .... b) തൂക്കം ...... c) നെഞ്ചളവ് ......

7. വൈവാഹികനില
   1. അവിവാഹിതൻ / അവിവാഹിത
   2. വിവാഹിതൻ / വിവാഹിത
   3. ഭാര്യ മരിച്ചവൻ / ഭർത്താവു മരിച്ചവൾ
   4. വേർപെട്ടവൻ / വേർപെട്ടവൾ

8. താമസസ്ഥലം
   1. ഹോസ്റ്റൽ 2. സ്വഗൃഹം
   3. ബന്ധുഗൃഹം 4. പ്രൈവറ്റ് ലോഡ്ജ്

9. ജനനസ്ഥലം
   നഗരം / ഗ്രാമം .. ...
   ജില്ല ......
   സംസ്ഥാനം ......

10. a) രക്ഷകർത്താവിൻെറ മാസവരുമാനം ......
    b) മററു വരുമാനങ്ങൾ .. ...

    ആകെ .. ...

11. രക്ഷകർത്താവിൻെറ തൊഴിൽ

12. വീട്ടിൽ നിന്നു് വിദ്യാലയത്തിലേയ്ക്കുള്ള ദൂരം

2. ദിനചര്യ

1. എപ്പോഴാണു് ഉറങ്ങാൻ പോവുകയും എഴുന്നേൽക്കുകയും ചെയ്യുന്നതു് ?
   (a) ചൂടുകാലത്തു് രാവിലെ വൈകുന്നേരം
   (b) തണുപ്പുകാലത്തു് രാവിലെ വൈകുന്നേരം

2. ക്ലാസ് സമയത്തിനു പുറമെ പ്രതിദിനം എത്ര മണിക്കൂർ നിങ്ങൾ പഠിക്കുന്നു ?
   (a) ഒന്നാം ടേമിൽ
   (b) രണ്ടാം ടേമിൽ
   (c) മൂന്നാം ടേമിൽ

3. (a) നിങ്ങൾ പതിവായി വ്യായാമം ചെയ്യാറുണ്ടോ ?
   ഉണ്ടെങ്കിൽ എവിടെവെച്ചു് ?
   എപ്പോൾ മുതൽ എപ്പോൾ വരെ ?
   b) ഏതു തരം വ്യായാമമാണു് ചെയ്യുന്നതു് ?
      (1) കസർത്തു് (2) ഓട്ടം
      (3) നീന്തൽ (4) മററുള്ളവ

4. a) എപ്പോഴെല്ലാമാണു് പ്രധാനഭക്ഷണം കഴിക്കുന്നതു് ?
   b) " ഉപഭക്ഷണം കഴിക്കുന്നതു് ?

5. a) എപ്പോഴാണു് കോളേജിൽ പോകുന്നതു് ?
   b) " തിരിച്ചെത്തുന്നതു് ?

6. a) ഒരു പീരിയേഡിൻെറ ദൈർഘ്യം എന്തു് ?

b) എത്ര ഫ്രീ പീരിയേഡുകൾ ഒരാഴ്ചയിൽ ഉണ്ട്

c) ആഴ്ചയിൽ എത്ര മണിക്കൂർ വീതം താഴെ പറയുന്നവയ്ക്കു ചെലവാക്കുന്നു ?

(1) ലെക്ചർ, സെമിനാർ തുടങ്ങിയവ.

(2) പ്രായോഗികപരിശീലനം

(3) കളികൾ

7. വിശ്രമസമയം എങ്ങനെ ചെലവാക്കുന്നു?

(താഴെ കൊടുത്തിരിക്കുന്നവയുടെ നേരെ പ്രാധാന്യമനുസരിച്ച് 1, 2, 3, സംഖ്യകൾ കൊടുക്കുക.)

1. സംഭാഷണം
2. ദിവാസ്വപ്നം
3. ഉറക്കം
4. ചീട്ടുകളി തുടങ്ങിയ വീട്ടിൽ വെച്ചുള്ള കളികൾ
5. കളിസ്ഥലത്തു പോയുള്ള കളികൾ
6. പുസ്തകപാരായണം
7. പാർട്ട് ടൈം ജോലികൾ
8. വീട്ടുജോലി
9. സിനിമ കാണൽ
10. പാട്ട്, ഡാൻസ്, നാടകം തുടങ്ങിയവ
11. സുഹൃൽ സന്ദർശനം
12. ലൈംഗികസംതൃപ്തി
13. സമ്മേളനങ്ങളിൽ സംബന്ധിക്കുക
14. മററുള്ളവ

3. കളികളിലുള്ള താൽപര്യം

1. a) നിങ്ങൾ ബഹിർഗൃഹ കായികമത്സരങ്ങളിൽ ഏർപ്പെടാറുണ്ടോ ?

b) ഇല്ലെങ്കിൽ എന്തുകൊണ്ട്?
1. താൽപര്യമില്ലാത്തതുകൊണ്ടും
2. സമയമില്ലാത്തതുകൊണ്ട്
3. സൗകര്യമില്ലാത്തതുകൊണ്ട്
4, മററു കാരണങ്ങൾ കൊണ്ട്

c) ഉണ്ടെങ്കിൽ ഏതിലെല്ലാം ?

(1) ഹോക്കി (2) ടെന്നീസ് (3) ക്രിക്കററ്
(4) ഫുട്ട്ബാൾ (5) ബാററ്മിൻറൺ (6) നീന്തൽ
(7) കസർത്ത് (8) മററുള്ളവ

2. a) കായികവിനോദങ്ങളിൽ സമ്മാനങ്ങൾ കിട്ടിയിട്ടുണ്ടോ ?
   b) ഉണ്ടെങ്കിൽ അതിൻെറ സ്വഭാവം
   c) കളികളിലുള്ള താൽപര്യം നിങ്ങളുടെ വിദ്യാഭ്യാസത്തെ ബാധിച്ചിട്ടുണ്ടോ ?

   (1) ഒട്ടുമില്ല (2) കുറച്ചെല്ലാം
   (3) വളരെയധികം (4) വേറെ ഏതെങ്കിലും തരത്തിൽ

3. കളികൾക്കു വേണ്ടി മാസം എന്തു ചെലവു ചെയ്യുന്നു ?

4. വിശ്രമ പരിപാടികൾ

1. അന്തർ ഗൃഹ മത്സരക്കളികളിൽ ഏർപ്പെടാറുണ്ടോ ?
   ഉണ്ടെങ്കിൽ ഏതിലെല്ലാം ?

   (1) ടേബിൾ ടെന്നിസ് (2) ചീട്ടുകളി
   (3) ചതുരംഗം (4) കാരംസ്സ് (5) മററുള്ളവ

2. ഒരു മാസത്തിൽ എത്ര സിനിമ കാണാറുണ്ട് ?

3. ഏതുതരം പടങ്ങളാണ് കാണാറുള്ളത് ?

   (1) ഇന്ത്യൻ/വിദേശി
   (2) സുഖപര്യവസായി/ദുരന്തങ്ങൾ
   (3) ചരിത്രപരം/മതപരം/സാമൂഹ്യം/മററുതരം

4. വിദേശ ചിത്രങ്ങൾ ഇഷ്ടപ്പെട്ടുന്നെങ്കിൽ താഴെ പറയുന്ന ഏതു കാരണം കൊണ്ട് ?

   (1) വിദ്യാഭ്യാസപരമായ പ്രാധാന്യം
   (2) സാങ്കേതിക മെച്ചം
   (3) കഥയുടെയും ആശയത്തിൻെറയും പുതുമ
   (4) വിദേശ സംസ്കാരത്തെപ്പററി മനസ്സിലാക്കാൻ
   (5) മററു കാരണങ്ങൾ

5. ക്ലബ്ബിലോ അതുപോലുള്ള മറേറതെങ്കിലും സംഘടനയിലോ അംഗമാണോ ?

6. നിങ്ങൾ വിശ്രമത്തിനായി താഴെ പറയുന്ന ഏതെങ്കിലും സ്വീകരിച്ചിട്ടുണ്ടോ ?

   (1) സംഗീതം (2) സാഹിത്യം
   (3) ചിത്രകല (4) പൂന്തോട്ടനിർമ്മാണം
   (5) സ്റ്റാമ്പ് ശേഖരണം (6) ഫോട്ടോഗ്രാഫി

(7) ദിവാസ്വപ്നം (8) യാത്ര
(9) മറേറതെങ്കിലും

7. വിശ്രമപരിപാടികൾക്കായി മാസം എന്തു ചെലവാക്കുന്നു ?

5. സാമൂഹ്യവും മതപരവുമായ പ്രവർത്തനങ്ങൾ

1. a) പാഠ്യേതര പ്രവർത്തനങ്ങളിൽ ഏർപ്പെടാറുണ്ടോ ?
b) ഉണ്ടെങ്കിൽ ഏതിലെല്ലാം ?

(1) വാദപ്രതിവാദയോഗങ്ങൾ, പ്രസംഗങ്ങൾ
(2) നാടകം
(3) സംഗീതക്കച്ചേരി
(4) സംഘടനാപ്രവർത്തനം
(5) സാമൂഹ്യസേവനപരിപാടി
(6) മററുള്ളവ

c) ഇല്ലെങ്കിൽ എന്തുകൊണ്ട് ?

(1) താൽപര്യമില്ല
(2) സൗകര്യമില്ല
(3) ഒരു സമയനഷ്ടമായി കരുതുന്നു
(5) മറേറതെങ്കിലും കാരണം

2. a) വിദ്യാർത്ഥികൾ നിർമാണ പ്രവർത്തനങ്ങളിൽ ഏർപ്പെടുന്നതു് നല്ലതാണോ ?

(1) അതെ (2) അല്ല (3) അറിഞ്ഞുകൂട
(4) താൽപര്യമില്ല

b) ഏതു തരത്തിലുള്ള നിർമാണ പ്രവർത്തനത്തിലാണ് ഏർപ്പെടേണ്ടതു് ?

1. a) സാഹിത്യപ്രവർത്തനം നല്ലതാണോ ?
b) ആണെങ്കിൽ ഏതു തരം ? (1) കഥ (2) ലേഖനം
(3) കവിത (4) മറ്റേതെങ്കിലും

c) നിങ്ങളുടെ ഏതെങ്കിലും സാഹിത്യം പ്രസിദ്ധീകരിച്ചിട്ടുണ്ടോ ?
d) ഉണ്ടെങ്കിൽ എവിടെ ?

1. കോളേജ് മാസിക
2. സാഹിത്യ പ്രസിദ്ധീകരണങ്ങൾ
3. പത്രങ്ങൾ
4. മററുള്ളവ

4. a) പാഠ്യേതര പ്രവർത്തനങ്ങൾക്കു് സമ്മാനമോ
ബഹുമതിയോ ലഭിച്ചിട്ടുണ്ടോ ?

b) ഉണ്ടെങ്കിൽ അതിൻെറ സ്വഭാവമെന്തു് ?

5. a) നിങ്ങൾക്കു് എത്ര പേരോടു് പ്രേമമോ അതിനോടു് സമാനമായ
അടുപ്പമോ ഉണ്ടു് ?

b) നിങ്ങൾ പ്രേമിക്കുന്നവർ എവിടെയാണു് ?

(1) പഠിക്കുന്ന കോളേജിൽ തന്നെ
(2) വീടിനടുത്തു്
(3) മററവിടെയെങ്കിലും

6. നിങ്ങളുടെ മതം ഏതാണു് ?

7. a) നിങ്ങൾ ദൈവത്തിൽ വിശ്വസിക്കുന്നുണ്ടോ ?

(1) ഉണ്ടു് (2) ഇല്ല (3) നിശ്ചയമില്ല

b) മതപരമായ അനുഷ്ഠാനങ്ങൾ നടത്താറുണ്ടോ ?

(1) പതിവായി (2) ഇടയ്ക്കിടയ്ക്കു്
(3) ചിലപ്പോൾ (4) ഒരിക്കലുമില്ല

6. വരുമാനമാർഗം

1. പ്രതിമാസം ആരു് എന്തുമാത്രം വിദ്യാഭ്യാസത്തിനു് നിങ്ങൾക്കു് ധന
സഹായം ചെയ്യുന്നു ?

(1)	പിതാവു്	രൂപ
(2)	അമ്മ	,,
(3)	തന്നത്താൻ	,,
(4)	മററു ബന്ധുക്കൾ	,,
(5)	മററു സ്ഥാപനങ്ങൾ	,,

2. a) നിങ്ങൾക്കു് സ്കോളർഷിപ്പു് കിട്ടുന്നുണ്ടോ ?

b) ആരാണു് സ്കോളർഷിപ്പു തരുന്നതു് ?

c) പ്രതിമാസം എന്തു തുക തരുന്നു ?

3. a) ഫീസിളവു് എന്തെങ്കിലും ഉണ്ടോ ?

b) എത്ര (1) പകുതി (2) മുഴുവൻ (3) ഭാഗികം
(എത്ര)

4. a) ഉള്ള വരുമാനം പഠനത്തിനു മതിയാകുന്നുണ്ടോ ?

b) (1) ഇല്ലെങ്കിൽ ബാക്കി വേണ്ട പണം എങ്ങനെ സമ്പാദിക്കുന്നു ?

(2) അതിനായി എത്ര മണിക്കൂർ ജോലി ചെയ്യുന്നു?

(3) മാസം എത്ര പണം സമ്പാദിക്കുന്നു ?

5. a) പഠനത്തിനായി പണം കടം വാങ്ങിക്കാറുണ്ടോ ?

b) ഉണ്ടെങ്കിൽ മാസം എത്ര രൂപ?

c) ആരിൽ നിന്നു് ? (1) സുഹൃത്തുക്കളിൽ നിന്നു്
(2) ബന്ധുക്കളിൽ നിന്നു്
(3) മററുള്ളവരിൽ നിന്നു്

d) ഇപ്പോൾ എന്തു കടബാധ്യതയുണ്ടു് ?

6. വീട്ടും വീട്ടുപകരണങ്ങളും (വീട്ടിൽ താമസിക്കുന്നവർക്കു്)

1. ഏതു തരം വീട്ടിലാണു് താമസിക്കുന്നതു് ?

(1) നല്ല വീട്ടു് (2) മൺ വീട്ടു്
(3) കുടിൽ (4) മററു തരം

2. വീടിൻെറ വിസ്തീർണം

a) മുറികളുടെ എണ്ണം
b) മുറികളുടെ വിസ്തീർണം
c) വരാന്തയുടെ വിസ്തീർണം .

3. വീട്ടിലെ ശുചീകരണസംവിധാനം

(1) തൃപ്തികരം (2) സാമാന്യം
(3) മോശം

4. വെൻറിലേഷൻ (1) തൃപ്തികരം (2) സാമാന്യം
(3) മോശം

5. വീട്ടിൽ ഇലക്ട്രിക്കു് ലൈററുണ്ടോ ?

6. നിങ്ങൾ ഫാൻ ഉപയോഗിക്കുന്നുണ്ടോ ?

7. എവിടെ നിന്നാണു് വെള്ളം കിട്ടുന്നതു് ?

(1) ടാപ്പു് (2) കിണറു് (3) മററു സ്ഥലം

8. a) നിങ്ങൾക്കു് പഠിക്കാൻ പ്രത്യേകം മുറിയുണ്ടോ ?

b) ഇല്ലെങ്കിൽ നിങ്ങളുടെ മുറി എത്ര പേർ ഉപയോഗിക്കുന്നു ?

(1) യുവാക്കൾ
(2) വിദ്യാർത്ഥികൾ ആണു്
പെണ്ണു്
(3) കുട്ടികൾ

9. നിങ്ങളുടെ വീട്ടിൽ (a) റേഡിയോയുണ്ടോ ?
(b) ടെലഫോണുണ്ടോ ?

10. a) കോളേജിൽ നിന്നു് നിങ്ങളുടെ വീട്ടിലേക്കു് എന്തു ദൂരമുണ്ടു് ?

b) കോളേജിലേക്കു പോകുന്ന രീതി

(1) നടന്നു്
(2) സൈക്കിളിൽ
(3) സൈക്കിൾ റിക്ഷാ
(4) ബസ്സിൽ
(5) കാറിൽ
(6) മറേറതെങ്കിലും

c) കോളേജിൽ എത്താൻ എടുക്കുന്ന സമയം

11. a) വീട്ടിൽ താമസിച്ചു കൊണ്ടു് തൃപ്തികരമായി പഠനം നടത്താൻ സാധിക്കുന്നുണ്ടോ ?

b) ഹോസ്റ്റലിൽ പ്രവേശനത്തിനു് അപേക്ഷിച്ചിട്ടുണ്ടോ ?

c) ഹോസ്റ്റലിൽ കൂടുതൽ പഠനസൗകര്യങ്ങളുണ്ടെന്നു് കരുതുന്നുണ്ടോ?

8. താമസ സൗകര്യം (ഹോസ്റ്റലിൽ താമസിക്കുന്നവർ)

1. എന്തുകൊണ്ടാണു് നിങ്ങൾ ഹോസ്റ്റലിൽ താമസിക്കുന്നതു് ?
2. എത്ര നാളായി ഹോസ്റ്റലിൽ താമസിക്കുന്നു ?
3. നിങ്ങളുടെ മുറിയുടെ വിസ്തീർണം ?
4. നിങ്ങളുടെ മുറിയിൽ മററു് എത്ര കുട്ടികൾ കൂടി താമസിക്കുന്നുണ്ടു് ?
5. a) ഹോസ്റ്റലിൽ ഒരു പൊതു മുറിയുണ്ടോ ?
b) പത്രമാസികകളും അന്തർഗൃഹ കളി സൗകര്യങ്ങളും ഉണ്ടോ ?
6. ഹോസ്റ്റലിൽ ഭക്ഷണസൗകര്യമുണ്ടോ ?
7. ഹോസ്റ്റലിലെ ശുചീകരണസംവിധാനം എങ്ങനെ ?

1) പൊതു ശുചീകരണം
(a) തൃപ്തികരം
(b) സാമാന്യം
(c) മോശം

2) മററു സൗകര്യങ്ങൾ

(a) കുളിമുറി ഇല്ല/പൊതുവായുണ്ട്/പ്രത്യേകം ഉണ്ട്
(b) കക്കൂസ് ഇല്ല/പൊതു/പ്രത്യേകം
(c) ഇലക്ട്രിസിറ്റി ഉണ്ട്/ഇല്ല
(d) റേഡിയോ ഉണ്ട്/ഇല്ല

8. നിങ്ങൾക്കു മാത്രമായി താഴെ പറയുന്നവയുണ്ടോ ?

(1) ഫാൻ (2) മേശവിളക്ക് (3) മേശ
(4) കസേര (5) ഷെൽഫ്

9. ഹോസ്റ്റലും കോളേജുമായുള്ള അകലം

10. a) ഹോസ്റ്റലിലെ അന്തരീക്ഷം പഠിത്തത്തിനു സൗകര്യപ്രദമാണോ ?
b) അല്ലെങ്കിൽ ഉപദ്രവമുണ്ടാക്കുന്ന ഘടകങ്ങൾ ഏതെല്ലാം ?

(1)
(2)
(3)

9. പഠനം, പരീക്ഷ, ജീവിതലക്ഷ്യം മുതലായവ

1. നിങ്ങൾ പാസായിട്ടുള്ള ഏററവും ഉയർന്ന പരീക്ഷ ഏതു് ?
2. നിങ്ങൾ ഏതെല്ലാം പരീക്ഷകൾ പാസായിട്ടുണ്ട് ?
അവ ഓരോന്നിലും കിട്ടിയ ക്ലാസ് ഏതാണ് ?
3. ഇപ്പോൾ പഠിക്കുന്ന കോഴ്സിനു ചേരാൻ കാരണമെന്തു് ?
4. പഠനം കഴിഞ്ഞു് ഏതു തൊഴിൽ സ്വീകരിക്കാനാണ്
ആഗ്രഹിക്കുന്നതു് ?

[മുൻഗണനക്രമം 1, 2, 3 എന്നിങ്ങനെ സംഖ്യകൾ എഴുതി സൂചിപ്പിക്കുക.]

(1) ഗവണ്മെൻറുദ്യോഗം
(2) സർവകലാശാലാ അധ്യാപകൻ
(3) സ്കൂൾ അധ്യാപകൻ
(4) പ്രൈവററ് സ്ഥാപനങ്ങളിലെ ജോലി
(5) ചികിൽസ
(6) സ്വന്തം ബിസിനസ്
(7) വക്കീൽ പണി
(8) രാഷ്ട്രീയം
(9) കുടുംബഭരണം
(10) മറേറതെങ്കിലും

5. a) ആവശ്യമുള്ള പാഠപുസ്തകങ്ങൾ നിങ്ങൾ വാങ്ങിയിട്ടുണ്ടോ ?

b) ആവശ്യത്തിനു് പുസ്തകങ്ങൾ ഇല്ലെങ്കിൽ എങ്ങനെയാണു് ആവശ്യങ്ങൾ സാധിക്കുന്നതു് ?

(1) പുസ്തകങ്ങൾ കടം വാങ്ങി (2) പുസ്തകങ്ങളില്ലാതെ
(3) ലൈബ്രറി പുസ്തകങ്ങൾ (4) മററു മാർഗങ്ങൾ

6. നിങ്ങൾക്കു് ലൈബ്രറിക്കാർഡ് ഉണ്ടോ ?

7. a) ലൈബ്രറിയിൽ നിന്നു് ഒരു സമയത്തു് എത്ര പുസ്തകങ്ങൾ എടുക്കാം ?

b) ഇപ്പോൾ എത്ര പുസ്തകങ്ങൾ എടുത്തിട്ടുണ്ടു് ?

8. നിങ്ങൾ പതിവായി വായിക്കാറുണ്ടോ ?

a) പത്രങ്ങൾ
b) മാസികകൾ
c) പുസ്തകങ്ങൾ

9. പുസ്തകങ്ങൾ വായിക്കാറുണ്ടെങ്കിൽ ഏതു തരത്തിൽ പെട്ടവ

(a) നോവൽ, കഥ (b) ഡിററക്ടീവ്
(c) രാഷ്ട്രീയം (d) ശാസ്ത്രീയം
(e) മതപരം (f) കവിത
(g) ജീവചരിത്രം (h) ചരിത്രം
(i) മററുള്ളവ

10. a) പാഠ സഹായികൾ പഠനത്തിനു് ഉപയോഗിക്കാറുണ്ടോ ?

b) ഉണ്ടെങ്കിൽ എന്തുകൊണ്ടു് ?

(1) തയ്യാറാക്കിയ ഉത്തരങ്ങൾ കിട്ടും
(2) മനസ്സിലാക്കാൻ എളുപ്പമുണ്ടു്
(3) പാഠപുസ്തകങ്ങൾ വായിച്ചു പഠിക്കാൻ സമയമില്ല
(4) മറെറന്തെങ്കിലും (എന്തെന്നു വ്യക്തമാക്കുക)

11. പ്രത്യേകം ട്യൂഷൻ ലഭിക്കുന്നുണ്ടോ ?

12. a) ഇപ്പോഴത്തെ പരീക്ഷാസമ്പ്രദായം മാററണമോ ?

(1) വേണം (2) വേണ്ട (3) അറിഞ്ഞുകൂട
(4) താല്പര്യമില്ല

b) മാററണമെങ്കിൽ എന്തുകൊണ്ട് ?

(1) ഒരാളുടെ കഴിവിനെ ശരിക്കും പരീക്ഷിക്കുന്നില്ല
(2) ക്ലാസിൽ വെച്ചു് ചെയ്യുന്ന ജോലി പരിഗണിക്കുന്നില്ല
(3) യാദൃച്ഛികസംഭവങ്ങൾക്ക് വളരെയേറെ സ്വാധീനമുണ്ട്
(4) ഓർമശക്തിയാണ് പ്രധാനമായി പരീക്ഷിക്കപ്പെടുന്നതു്
(5) മറേറതെങ്കിലും കാരണം

13. a) പതിവായി ക്ലാസിൽ പോവാറുണ്ടോ ?

b) ഇല്ലെങ്കിൽ എന്തുകൊണ്ട് ?

(1) മററു ജോലി (2) ക്ലാസുകൾ രസകരമല്ല
(3) പ്രയോജനപ്രദമല്ല (4) അച്ചടക്കമില്ല
(5) മററു കാരണം

c) ഇത്ര ശതമാനം ഹാജരുണ്ടായിരിക്കണമെന്ന നിർബന്ധം ഉപേക്ഷിച്ചാൽ നിങ്ങൾ ക്ലാസിൽ പോകുമോ ?

d) ഇല്ലെങ്കിൽ എന്തുകൊണ്ട് ?

(1) ക്ലാസുകൾ രസകരങ്ങളല്ല (2) പ്രയോജനമില്ല
(3) മറേറതെങ്കിലും

e) പോകുമെങ്കിൽ എന്തുകൊണ്ട് ?

(1) ക്ലാസുകൾ പ്രയോജനപ്രദമാണ്
(2) രസകരമാണ്
(3) ഉപയോഗപ്രദമാണ്
(4) മറേറതെങ്കിലും കാരണം

14. നിങ്ങളുടെ അഭിപ്രായത്തിൽ അധ്യയന മാധ്യമം എന്തായിരിക്കണം ?

(1) ഇംഗ്ലീഷ് (2) ഹിന്ദി
(3) മാതൃഭാഷ (4) മറേറതെങ്കിലും

10. സാമൂഹ്യപദവി

1. നിങ്ങളുടെ അച്ഛനും മുത്തച്ഛനും എവിടെയാണ് താമസിച്ചിരുന്നതു് ?

(a) അച്ഛൻ നഗരം/ഗ്രാമം
(b) മുത്തച്ഛൻ നഗരം/ഗ്രാമം

2. നിങ്ങളുടെ അച്ഛൻെറയും മുത്തച്ഛൻെറയും ജോലി എന്തായിരുന്നു ?

(a) അച്ഛൻ
(b) മുത്തച്ഛൻ

3. പിതാവിന്റെ തൊഴിൽ തന്നെ തുടരാനാണോ ആഗ്രഹിക്കുന്നതു്? അല്ലെങ്കിൽ എന്തുകൊണ്ട് ?

11. ഭക്ഷണം

1. നിങ്ങൾ സസ്യഭുക്കോ മാംസഭുക്കോ ?

2. താഴെ പറയുന്ന സാധനങ്ങൾ ഏകദേശം എത്ര തൂക്കം നിങ്ങൾ ഉപയോഗിക്കാറുണ്ട് ?

ഭക്ഷണസാധനം (ഒരുദിവസം) തൂക്കം (ഗ്രാമിൽ) വില

a. പാൽ
b. പച്ചക്കറികൾ
c. തൈരു്, നെയ്യ്
d. എണ്ണ
e. പഴങ്ങൾ
f. ഇറച്ചി
g. മീൻ
h. അരി, ഗോതമ്പ്
i. പയർവർഗങ്ങൾ
j മുട്ട
k. പലവ്യംജനങ്ങൾ

3. a) ഈ മാസത്തിൽ വീട്ടിനു പുറത്തു നിന്ന് ആഹാരം കഴിച്ചിട്ടുണ്ടോ?

b) ഉണ്ടെങ്കിൽ എണ്ണം തുക കാരണം

പ്രധാനഭക്ഷണം

ലഘുഭക്ഷണം

4. ഏതെങ്കിലും രൂപത്തിൽ പുകയില ഉപയോഗിക്കുമോ ?
ഉപയോഗിക്കുമെങ്കിൽ ഏതു രൂപത്തിൽ

a. സിഗററ്റ് b. ചുരുട്ട്
c. മുറുക്കാൻ d. മറ്റേതെങ്കിലുംരൂപം

5. ഒരു ദിവസം എത്ര സിഗററ്റ് ഉപയോഗിക്കാറുണ്ട് ?

a) സുഹൃത്തുക്കൾ തന്നാൽ
b) സ്വന്തം പണം കൊടുത്ത്

6. a) ലഹരിപാനീയങ്ങൾ ഉപയോഗിക്കാറുണ്ടോ ?

b) ഉണ്ടെങ്കിൽ പതിവായോ വല്ലപ്പോഴുമോ ?

c) വല്ലപ്പോഴുമെങ്കിൽ എത്ര ദിവസം കൂടുമ്പോൾ ?

12. ആരോഗ്യം

1. a) അംഗഭംഗങ്ങൾ ഏതെങ്കിലും ഉണ്ടോ ?

b) ഉണ്ടെങ്കിൽ എന്ത്?

2. a) കഴിഞ്ഞ അധ്യയന വർഷത്തിൽ ഗൗരവമായ ഏതെങ്കിലും രോഗമുണ്ടായോ ?

b) ഉണ്ടായെങ്കിൽ എന്ത്?

c) ചികിത്സ എന്തെങ്കിലും ലഭിച്ചോ?

d) ചികിത്സക്ക വേണ്ടിവന്ന ഏകദേശ ചെലവ്.

3. ആരോഗ്യക്കുറവു കൊണ്ട് കഴിഞ്ഞ വർഷം എത്ര ദിവസം കോളേജിൽ പോകാതിരുന്നു ?

4. കഴിഞ്ഞ വർഷം നിങ്ങളെ കോളേജ് ഡോക്ടർ പരിശോധിക്കയുണ്ടായോ ?

5. എത്ര ദിവസം കൂടുമ്പോൾ ഡോക്ടറെ കണ്ടു പരിശോധന നടത്തിക്കാറുണ്ട്?

6 a) നിങ്ങൾ കണ്ണട ഉപയോഗിക്കുന്നുണ്ടോ ?

b) ഉണ്ടെങ്കിൽ എന്തിന് ?

13. വസ്ത്രധാരണം

1. നിങ്ങളുടെ പതിവു വസ്ത്രം എന്താണ് ?

	വീട്ടിൽ	പുറത്ത്
തണുപ്പുകാലത്ത്		
വേനൽ കാലത്ത്		

2. എത്ര ജോഡി വസ്ത്രങ്ങൾ ഇപ്പോഴുണ്ട് ?

14. ആകെ ചെലവ്

ആകെ ചെലവ് (കഴിഞ്ഞ വർഷ)

1. പുസ്തകങ്ങൾ	......	രൂപ
2. വസ്ത്രം	......	,,
3. ചെരിപ്പ്, കുട, തുടങ്ങിയവ	......	,,
4. കോളേജ് ഫീസ്	......	,,
5. ഹോസ്റ്റൽ ഫീസ്	......	,,
6. യാത്ര	......	,,
7. മററുള്ളവ [പെൻസിൽ, പേന തുടങ്ങിയവ]	......	,,
		,,
ആകെ	......	

15. അഭിപ്രായഗതി

1. a) സർവകലാശാലാ വിദ്യാർഥികൾ രാഷ്ട്രീയകാര്യങ്ങളിൽ ഇടപെടുന്നതു് നല്ലതോ?
   1. അല്ല 2. അതേ
   3. അറിഞ്ഞുകൂട
   4. താല്പര്യമില്ല.

   b) ആണെങ്കിൽ എന്തുകൊണ്ട്?

   1. ഭാവിയിലേക്ക് പരിശീലനം ലഭിക്കാൻ
   2. വിദ്യാർഥികളുടെ മതിപ്പ് ലഭിക്കാൻ
   3. രാജ്യത്തിനു് അഭിവൃദ്ധിയുണ്ടാവാൻ
   4. പഠിക്കുന്ന തത്വങ്ങൾ പരീക്ഷിക്കാൻ
   5. മററു കാരണങ്ങൾ.

2. സഹവിദ്യാഭ്യാസത്തെപ്പററി എന്താണ് അഭിപ്രായം?
   1. അനുകൂലം 2. പ്രതികൂലം 3. താല്പര്യമില്ല.

3. വാഹനങ്ങളിൽ യാത്ര ചെയ്യുമ്പോഴുള്ള വിദ്യാർഥികളുടെ പെരുമാററത്തെപററി എന്താണ് അഭിപ്രായം?
   1. നല്ലതു് 2. ചീത്ത 3. ഒന്നും ഇല്ല

4. a) നിങ്ങളുടെ കോളേജിലെ ക്ലാസ്സ് മുറിയിലെ അച്ചടക്കം എങ്ങനെ?

   b) അച്ചടക്കക്കുറവുണ്ടെങ്കിൽ എന്തുകൊണ്ട്?

   1. അധ്യാപകന്റെ കഴിവില്ലായ്മ
   2. അധ്യാപകന്റെ കർക്കശത്വം

3. ക്ലാസ്സ് രസകരമല്ല
4. ചില കുട്ടികളുടെ ദു:സ്വഭാവം
5. മറ്റേതെങ്കിലും കാരണം.

5. സർവകലാശാലകളിൽ ഇന്നു കാണപ്പെടുന്ന അച്ചടക്കരാഹിത്യത്തിന്റെ കാരണമെന്ത്?

(പ്രാധാന്യമനുസരിച്ച് 1, 2, 3 സംഖ്യകൾ കൊടുക്കുക)

1. അധ്യാപകരും വിദ്യാർഥികളും തമ്മിലുള്ള അടുത്ത സമ്പർക്കത്തിന്റെ കുറവ്.
2. മിച്ചം സമയവും ഉന്മേഷവും ചെലവഴിക്കാൻ മാർഗമില്ലായ്മ
3. രാഷ്ട്രീയകക്ഷികളുടെ സ്വാധീനം
4. വിദ്യാഭ്യാസത്തിനുശേഷം തൊഴിൽസാധ്യതകൾ കുറവ്
5. വിദ്യാർഥികൾക്ക് അവർ എന്തും ചെയ്യാൻ അവകാശമുള്ള ഒരു വർഗമാണെന്നുള്ള തോന്നൽ
6. സമൂഹത്തിലെ പ്രശ്നങ്ങളുടെ സ്വാധീനം
7. അധികാരികളുടെ മനോഭാവം
8. പഠനസൗകര്യങ്ങളുടെ കുറവ്
9. അധ്യാപകരുടെ രാഷ്ട്രീയപ്രവർത്തനം
10. വിദ്യാർഥികളുടെ ഉത്തരവാദിത്വബോധമില്ലായ്മ
11. മറ്റു കാരണങ്ങൾ

അന്വേഷകന്റെ പേര്

കൂടിക്കാഴ്ച നടത്തിയ തീയതി

അനുബന്ധം-2

സാമ്പിളനത്തിന് ഒരു ലഘുമാതൃക

സാമ്പിളന ലക്ഷ്യം: കേരളത്തിലെ തൊഴിലില്ലായ്മയെപ്പററി പഠനം നടത്തുക.

സാമ്പിളൽ ആസൂത്രണം ചെയ്യുകയാണ് ആദ്യത്തെ പടി. ഏതെല്ലാം വിവരങ്ങളാണ് ആവശ്യമുള്ളതെന്ന് വ്യക്തമായി നിർവചിക്കണം. സ്വാഭാവികമായും എത്രമാത്രം തൊഴിലില്ലായ്മയുണ്ടെന്നും അതിന്റെ സ്വഭാവമെന്താണെന്നും അറിയേണ്ടിയിരിക്കുന്നു. അതുതന്നെ പട്ടണങ്ങളിൽ എത്രമാത്രം ഗ്രാമങ്ങളിൽ എത്ര മാത്രം എന്നു് പരിശോധിക്കണം. ഏതു പ്രായത്തിലുള്ളവരുടെ ഇടയിലാണ് തൊഴിലില്ലായ്മ കൂടുതലുള്ളതെന്നു മനസ്സിലാക്കുന്നതു് ഉപകാരപ്രദമായിരിക്കും. വിദ്യാഭ്യാസമുള്ളവരുടെ ഇടയിലും, ഇല്ലാത്തവരുടെ ഇടയിലുമുള്ള തൊഴിലില്ലായ്മയുടെ അളവ് കണ്ടുപിടിക്കണം. തൊഴിലില്ലാത്തവരുടെ വിദ്യാഭ്യാസയോഗ്യത അറിയണം. സാമ്പത്തികമായി ഏതു നിലയിലുള്ളവരുടെ ഇടയിലാണ് തൊഴിലില്ലായ്മയുള്ളതു് എന്നതും പ്രാധാന്യമർഹിക്കുന്ന ഒരു കാര്യമാണ്. തൊഴിലില്ലാത്തവരിൽ കുടുംബജീവിതം നയിക്കുന്നവരുടെയും അല്ലാത്തവരുടെയും സംഖ്യ മനസ്സിലാക്കേണ്ടതാണ്. പൂർണ്ണമായും തൊഴിലില്ലാത്തവരെയും ഭാഗികമായി തൊഴിലില്ലാത്തവരെയും തിരിച്ചറിയേണ്ടതും ആവശ്യമത്രെ. തൊഴിലുള്ളവരെ സംബന്ധിച്ചും മുൻപറഞ്ഞ വിവരങ്ങളെല്ലാം ലഭിക്കുന്നതു് താരതമ്യപഠനത്തിനു് അത്യന്തം ഉപകാരപ്രദമായിരിക്കും. തൊഴിലില്ലായ്മയുടെ കാരണം, ആഗ്രഹിക്കുന്ന തൊഴിൽ കിട്ടാത്തതാണോ അതോ ഒരു തൊഴിലും കിട്ടാത്തതാണോ എന്നും അറിയണം. സാമാന്യേന ഇതെല്ലാമാണ് സാമ്പിളലിലൂടെ ശേഖരിക്കാൻ ആഗ്രഹിക്കുന്ന വിവരങ്ങൾ.

അടുത്തപടി ചോദ്യാവലി തയ്യാറാക്കലാണ്. മുൻപറഞ്ഞ വിവരങ്ങളെല്ലാം ലഭിക്കത്തക്കവണ്ണം ശ്രദ്ധാപൂർവം രൂപം കൊടുക്കേണ്ട ഒന്നാണിതു്. അനുബന്ധം–1 ൽ ആദ്യത്തേതായി കൊടുത്തിരിക്കുന്ന ചോദ്യാവലി ഇതിനു് ഉപകരിക്കും. കുടുംബത്തെ സാമ്പിളൽ വ്യക്തിയായി കരുതിയാണ് ആ ചോദ്യാവലി തയ്യാറാക്കിയിരിക്കുന്നതു്. തൊഴിലില്ലായ്മയെപ്പററിയുള്ള പഠനത്തിൽ ഒററപ്പെട്ട വ്യക്തികളിൽ നിന്നു് വിവരം ശേഖരിക്കുന്നതിനെക്കാൾ കുടുംബങ്ങളിൽ നിന്നു ശേഖരിക്കുന്നതാണ് കൂടുതൽ സൗകര്യം എന്നതു് ഇതിനു പ്രേരകമായ ഒരു വസ്തുതയാണ്. അതുപോലെതന്നെ തൊഴിലില്ലായ്മയുടെ കെടുതി

കളെപ്പററി മനസ്സിലാക്കുവാൻ ഒരു കുടുംബത്തിൽ ശരാശരി എത്ര അംഗങ്ങൾക്കു് തൊഴിൽ ലഭിക്കുന്നുണ്ടു് എന്നും മററും അറിയേണ്ടതും ആവശ്യമാണല്ലൊ.

സാമ്പിളനരീതി നിർണ്ണയിക്കുകയാണു് ഇനി ചെയ്യേണ്ടതു്. കേരള ജനതയുടെ പ്രാതിനിധ്യമുണ്ടെന്നു കരുതാവുന്ന ഒരു സാമ്പിളാണു് ആവശ്യം. കേരളത്തിൻെറ എല്ലാ ഭാഗങ്ങളും തൊഴിലില്ലായ്മ എന്ന അഭിലക്ഷണത്തെ സംബന്ധിച്ചിടത്തോളം ഒരുപോലെയല്ല. അതുകൊണ്ടു് സ്തരണം ആവശ്യമാണു്. പ്രായോഗികവും ചരിത്രപരവുമായ പരിഗണനകൾ വെച്ചു നോക്കുമ്പോൾ ഓരോ ജില്ലയും ഓരോ സ്തരമായി പരിഗണിക്കാം. ഭൂമിശാസ്ത്രപരമായി നോക്കുമ്പോൾ, തീരപ്രദേശം, മധ്യപ്രദേശം, മലമ്പ്രദേശം എന്നു മൂന്നായി തിരിക്കാം. മിക്കവാറും എല്ലാ ജില്ലകളും (ആലപ്പുഴ ഒഴികെ) തടൽതീരം മുതൽ സഹ്യപർവത താഴ്‌വാരങ്ങൾ വരെ വ്യാപിച്ചു കിടക്കുന്നതുകൊണ്ടു് അങ്ങനെയുള്ള ഒരോ ജില്ലയേയും മൂന്നു സ്തരങ്ങളായി തിരിക്കാം. ഇങ്ങനെ ലഭിച്ച സ്തരങ്ങളെത്തന്നെ നഗരപ്രദേശം, ഗ്രാമപ്രദേശം എന്നു് രണ്ടായി വിഭജിക്കുന്നതു് പ്രസക്തമാണു്. ഇത്രയുമായാൽ ഓരോ സ്തരവും ഏതാണ്ടു് ഏകാത്മകങ്ങളാണെന്നു് കരുതാം. കഴിഞ്ഞ സെൻസസ്സിലെ വിവരങ്ങൾ ഉപയോഗപ്പെടുത്തി ഓരോ സ്തരത്തിൽ നിന്നു് എത്ര വീടുകൾ വീതം സാമ്പിളിൽ ഉൾപ്പെടുത്തണമെന്നു് തീരുമാനിക്കാം. നഗരങ്ങളിൽ ആളുകൾ ഇടതിങ്ങി താമസിക്കുകയും ജോലിസാധ്യതകൾ കൂടുതലായിരിക്കുകയും മററും ചെയ്യുന്നതുകൊണ്ടു് നഗരങ്ങളിൽ നിന്നു് കൂടുതൽ ശതമാനം വീടുകൾ സാമ്പിളിൽ ഉൾപ്പെടുത്തേണ്ടതു് ആവശ്യമാണു്. നഗരങ്ങളിൽ നിന്നു് 5% വീടുകളും ഗ്രാമങ്ങളിൽ നിന്നു് 3% വീടുകളും സാമ്പിളിൽ ഉൾപ്പെടുത്തുന്നതു് സാമാന്യം തൃപ്തികരമായ ഒരു വിഭജനമത്രെ. സംഘസാമ്പിളനരീതിയിൽ ഓരോ സ്തരത്തിൽ നിന്നും വീടുകൾ തിരഞ്ഞെടുക്കുന്നതാണു് നല്ലതു്. ഗ്രാമപ്രദേശങ്ങളിൽ വില്ലേജുകളും പട്ടണത്തിൽ വാർഡും സംഘമായി സ്വീകരിക്കാം. ഓരോ പട്ടണത്തിൽ നിന്നും പ്രതിസ്ഥാപനമില്ലാത്ത ലഘുയാദൃച്ഛികസാമ്പിളന രീതിയിൽ വാർഡുകൾ ഒന്നൊന്നായി തിരഞ്ഞെടുക്കുക. ആ നഗരത്തിലെ ഏതാണ്ടു് 5% വീടുകൾ ഉൾപ്പെടുന്ന വാർഡുകൾ ലഭിക്കുന്നതുവരെ അതു് തുടരുക. ഗ്രാമപ്രദേശങ്ങളിൽ ഇതുപോലെ വില്ലേജുകൾ തിരഞ്ഞെടുക്കുക. സാമ്പിളനത്തിനു് ആകെ ചെലവഴിക്കാവുന്ന പണത്തിൻെറ തോതനുസരിച്ചു് സാമ്പിളനരീതിയിൽ ചില മാററങ്ങൾ വരുത്തേണ്ടതായി വരാം. പരിമിതമായ ചെലവിലാണു് സാമ്പിളനം നിർവഹിക്കേണ്ടിവന്നതെങ്കിൽ ഓരോ ജില്ലയിൽ നിന്നും ഒരു നിശ്ചിത എണ്ണം നഗരങ്ങളും, ഗ്രാമപ്രദേശങ്ങളും പ്രതിസ്ഥാപനമില്ലാത്ത ലഘുയാദൃച്ഛികസാമ്പിളനരീതിയിൽ തിരഞ്ഞെടുത്തു് അവയിൽ നിന്നു മാത്രം യഥാക്രമം 5% വും 3% വും വീടുകൾ ഉൾക്കൊള്ളുന്ന വാർഡുകളും വില്ലേജുകളും തിരഞ്ഞെടുക്കാം. ചെലവുകൾ നിർദിഷ്ടതുകയിൽ കവിയാതിരിക്കത്തക്കവണ്ണം തിരഞ്ഞെടുക്കുന്ന നഗരങ്ങളുടെയും ഗ്രാമപ്രദേശങ്ങളുടെയും എണ്ണം നിർണയിക്കുകയും ചെയ്യാം.

തിരഞ്ഞെടുക്കുന്ന വാർഡുകളിലും വില്ലേജുകളിലുമുള്ള എല്ലാ വീടുകളിൽ

നിന്നും വിവരശേഖരണം നടത്തണം. വിവരശേഖരണഘട്ടത്തിൽ വരാവുന്ന ബുദ്ധിമുട്ടുകൾ വളരെയേറെ ലഘൂകരിക്കാൻ ഈ രീതി സഹായിക്കും.

അടുത്തതു് അന്വേഷകരെ തിരഞ്ഞെടുക്കുകയാണ്. ഓരോ ജില്ലയിലേക്കും നിയോഗിക്കുന്ന ആളുകൾ അതതു ജില്ലയിൽ പെട്ടവരായിരുന്നാൽ നന്നായിരിക്കും. ഗ്രാമപ്രദേശത്തുനിന്നുള്ളവരെ ഗ്രാമപ്രദേശത്തേക്കും നഗരങ്ങളിൽ നിന്നുള്ളവരെ നഗരങ്ങളിലേക്കും അയക്കാൻ ശ്രമിക്കണം. ഈ ലക്ഷ്യം മനസ്സിൽ വെച്ചുകൊണ്ടുവേണം അന്വേഷകരെ തിരഞ്ഞെടുക്കാൻ. തിരഞ്ഞെടുത്ത അന്വേഷകർക്കു് ശരിയായ പരിശീലനം നൽകണം. പരിശോധകന്മാരായി കുറെയാളുകളെയും തിരഞ്ഞെടുക്കണം. ഓരോ അന്വേഷകനും അടുത്തടുത്തുള്ള ഏതാനും വാർഡുകളോ വില്ലേജുകളോ വിവരശേഖരണത്തിനായി നൽകണം. 10 അന്വേഷകർക്കു് ഒരു പരിശോധകൻ എന്ന ക്രമത്തിൽ പരിശോധകരുണ്ടായിരിക്കുന്നതു കൊള്ളാം. ഓരോ ദിവസവും വിവരശേഖരണം തീരുമ്പോൾ പൂരിപ്പിച്ച ചോദ്യാവലികൾ പരിശോധകനെ ഏല്പിക്കണം. അതിൽ നിന്നു് യാദൃച്ഛികമായി തിരഞ്ഞെടുത്ത ഏതാനും വീടുകളിൽ നിന്നു് പരിശോധകൻ നേരിട്ടു് വിവരശേഖരണം നടത്തുകയും അതും അന്വേഷകൻ നൽകിയിരിക്കുന്ന വിവരങ്ങളുമായി താരതമ്യപ്പെടുത്തി അന്വേഷകനു് ആവശ്യമായ നിർദേശങ്ങൾ നൽകുകയും ചെയ്യണം. മാത്രമല്ല, ലഭിക്കുന്ന പൂരിപ്പിച്ച ചോദ്യാവലികൾ നിഷ്കൃഷ്ടമായി പരിശോധിച്ചു് പോരായ്മകൾ പരിഹരിക്കേണ്ടതും പരിശോധകൻെറ കർത്തവ്യമത്രെ.

ശേഖരിച്ച വിവരങ്ങൾ അപഗ്രഥിച്ചു് റിപ്പോർട്ടു തയ്യാറാക്കുകയാണു് അടുത്തതായി ചെയ്യാനുള്ളതു്. അപഗ്രഥനത്തിൻെറ ആദ്യഘട്ടം ആവശ്യമായ പട്ടികകൾ തയ്യാറാക്കലത്രെ. ഓരോ സ്തരത്തിൽ നിന്നും ലഭിച്ച വിവരങ്ങൾ ഓരോ തരത്തിലുള്ള പ്രത്യേകം പ്രത്യേകം പട്ടികകളിൽ സംഗ്രഹിക്കണം. അതിനുശേഷം അവ ഒന്നിച്ചുചേർത്തു് ഒററ പട്ടികയാക്കണം. ഏതെല്ലാം പട്ടികകളാണു് ആവശ്യം എന്നു തീരുമാനിക്കുന്നതു് ഏതെല്ലാം തരത്തിലുള്ള വിവരങ്ങളാണു് നമുക്കു് ആവശ്യം എന്നതിനെ ആശ്രയിച്ചാണു്. ഉദാഹരണമായി ഈ അന്വേഷണത്തിൽ ആവശ്യമായി വരാവുന്ന രണ്ടു പട്ടികകളുടെ മാതൃകകൾ താഴെ കൊടുക്കുന്നു.

പട്ടിക-1

വിവിധ വയോന്തരാളങ്ങളിലുള്ള തൊഴിലില്ലാത്ത സ്ത്രീകളുടെയും പുരുഷന്മാരുടെയും, നഗരങ്ങളിലും ഗ്രാമപ്രദേശങ്ങളിലുമായുള്ള വിതരണം.

	നഗരം		ഗ്രാമപ്രദേശം	
	സ്ത്രീ	പുരുഷൻ	സ്ത്രീ	പുരുഷൻ
വയസ്സ്				
15—24				
25—34				
35—44				
45—54				
55 ൽ കൂടുതൽ				

ഇവിടെ ഓരോ വയോന്തരാളത്തിന്റെയും നേരെ യഥാർത്ഥത്തിലുള്ള ആളുകളുടെ എണ്ണം കൊടുക്കണമെന്നില്ല. ആകെ 1000 എന്നു സങ്കൽപിച്ചാൽ ഓരോ വിഭാഗത്തിലും എത്ര വീതം വരുമെന്നു കാണിച്ചാൽ മതി. ഈ പട്ടികയുടെ ഉപയോഗം ഓരോ വയോന്തരാളത്തിലുമുള്ള തൊഴിലില്ലാത്ത എത്ര ശതമാനം വീതം പുരുഷന്മാരും സ്ത്രീകളും യഥാക്രമം നഗരങ്ങളിലും ഗ്രാമങ്ങളിലുമായി താമസിക്കുന്നു എന്നു മനസ്സിലാക്കാൻ കഴിയുമെന്നതാണ്. തൊഴിലില്ലായ്മയെ പററിയുള്ള പഠനത്തിൽ അത്യന്തം വിലപ്പെട്ട വിവരമായിരിക്കുമല്ലോ ഇത്.

പട്ടിക-2

വിദ്യാഭ്യാസനിലവാരം തൊഴിലില്ലാത്തവരുടെ ഇടയിൽ

(1000 വീതം സ്ത്രീകളെയും പുരുഷന്മാരെയും പരിഗണിക്കുന്നു എന്നു കരുതുക.)

	വിദ്യാഭ്യാസനിലവാരം	പുരുഷന്മാർ	സ്ത്രീകൾ
1.	എഴുതാനും വായിക്കാനും അറിഞ്ഞുകൂടാ		
2.	എഴുതാനും വായിക്കാനും മാത്രമറിയാം		
3.	പ്രൈമറിവിദ്യാഭ്യാസം		
4.	ഹൈസ്കൂൾ വിദ്യാഭ്യാസം		
5.	ഡിഗ്രി ക്ലാസ്സു വരെ		
6.	സാങ്കേതികവിദ്യാഭ്യാസം		
7.	ബിരുദാനന്തര പഠനം		
8.	കലാപരമായ അഭ്യാസം		

ഈ രീതിയിൽ ആവശ്യാനുസരണം പല തരത്തിലുള്ള പട്ടികകൾ തയ്യാറാക്കണം. ഇവിടെ തന്നെ, കുടുംബവരുമാനമനുസരിച്ചും വരുമാനമാർഗമനുസരിച്ചും സ്ത്രീപുരുഷന്മാരുടെ എണ്ണം നൽകുന്ന പട്ടികകൾ, ഓരോ തൊഴിലിലും ഏർപ്പെട്ടിട്ടുള്ളവരുടെ സ്ത്രീ പുരുഷഭേദവും വയോന്തരാളവും അനുസരിച്ചുള്ള പട്ടികകൾ, ഓരോ തൊഴിലിലും ഏർപ്പെട്ടിട്ടുള്ളവരുടെ വിദ്യാഭ്യാസനിലവാരം കാണിക്കുന്ന പട്ടികകൾ, തൊഴിലില്ലാത്ത സ്ത്രീപുരുഷന്മാരുടെ വയോന്തരാളവും വിദ്യാഭ്യാസയോഗ്യതയും മററും കൂട്ടിച്ചേർത്തുള്ള പട്ടിക, സാങ്കേതിക വിദ്യാഭ്യാസവും ഉന്നതവിദ്യാഭ്യാസവും സിദ്ധിച്ചിട്ടുള്ള തൊഴിൽ രഹിതരായ സ്ത്രീപുരുഷന്മാരുടെ എണ്ണം കണിക്കുന്ന പട്ടികകൾ തുടങ്ങി അന്വേഷണം സംഘടിപ്പിക്കുന്ന ആളിന് പ്രസക്തമെന്നു തോന്നുന്ന നിരവധി പട്ടികകൾ തയ്യാറാക്കേണ്ടി വരും.

ഇങ്ങനെ ലഭിക്കുന്ന പട്ടികകളെ അടിസ്ഥാനമാക്കി കേരളത്തിനു പൊതുവായുള്ള പട്ടികകൾ തയ്യാറാക്കണം. ഒരു സ്തരത്തിലെ ജനസംഖ്യ a യും അതിൽ നിന്നും സാമ്പിളിൽ ഉൾപ്പെടുത്തിയ അംഗങ്ങളുടെ എണ്ണം b യുമാണെങ്കിൽ

ആ സ്തരത്തെ അടിസ്ഥാനമാക്കി തയ്യാറാക്കിയ പട്ടികയിലെ സംഖ്യകളെ a/b കൊണ്ടു ഗുണിക്കണം. ഇങ്ങനെ ഓരോ സ്തരത്തിൽ നിന്നു ലഭിച്ച പട്ടികയിൽ നിന്നും പുതിയ പട്ടികകളുണ്ടാക്കി അവയിലെ സംഖ്യകളെ യഥാസ്ഥാനം ഒന്നിച്ചു കൂട്ടിയാൽ കേരളത്തിന്നു് പൊതുവെയുള്ള പട്ടികകൾ കിട്ടും. ഇവയെ അടിസ്ഥാനമാക്കി വേണം പൊതുവായ നിഗമനങ്ങളിൽ എത്താൻ.

തൊഴിലില്ലായ്മയെ പററി പഠനം നടത്തുമ്പോൾ 15 വയസ്സിൽ കുറഞ്ഞവരെ പ്രത്യേകമായി പരിഗണിക്കുന്നതു കൊള്ളാം. അതുപോലെ തന്നെ 60 വയസ്സിൽ കൂടിയവരേയും. 15 നും 60 നും മധ്യേ പ്രായമുള്ളവരിൽ എത്ര ശതമാനം സ്ത്രീകളുണ്ടു് എത്ര ശതമാനം പുരുഷന്മാരുണ്ടു്; അവരിൽ ഓരോ വിഭാഗത്തിലും എത്ര ശതമാനത്തിനു് തൊഴിലുണ്ടു്; ഓരോ വയോന്തരാളത്തിലുമുള്ള എത്ര ശതമാനം സ്ത്രീകൾക്കും പുരുഷന്മാർക്കും തൊഴിലുണ്ടു്; ഓരോ വിദ്യാഭ്യാസനിലവാരത്തിലുള്ളവരുടെ തൊഴിലില്ലായ്മ എത്ര ശതമാനം വീതമാണു് എന്നു തുടങ്ങി പ്രയോജനപ്രദമെന്നു തോന്നുന്ന എല്ലാ പ്രാചലങ്ങളും ആകലനം ചെയ്തു് റിപ്പോർട്ടു തയ്യാറാക്കണം. ഇതിനു് മുൻ തയ്യാറാക്കിയ പട്ടികകൾ ഉപയോഗപ്പെടുത്താം. റിപ്പോർട്ടു തയ്യാറാക്കുമ്പോൾ കേരളത്തിൽ നിലവിലുള്ള സാമൂഹ്യ സാമ്പത്തികസ്ഥിതികളെപ്പററി പ്രസക്തമായ കാര്യങ്ങൾ കൂടി കണക്കിലെടുക്കുന്നതു കൊള്ളാം. ഇതിലെല്ലാം അന്വേഷണം സംഘടിപ്പിക്കുന്ന ആളിന്റെ സാമാന്യബോധവും പരിചയസമ്പത്തുമാണു് ഏററവുമധികം പ്രയോജനപ്രദമാവുന്നതു്.

നോർമൽ സാരണി

$$F(x) = \int_{-\infty}^{x} \frac{1}{\sqrt{2\pi}} e^{-t^2/2}\, dt$$

x	.00	.01	.02	.03	.04	.05	.06	.07	.08	.09
.0	.5000	.5040	.5080	.5120	.5160	.5199	.5239	.5279	.5319	.5359
.1	.5398	.5438	.5478	.5517	.5557	.5596	.5636	.5675	.5714	.5753
.2	.5793	.5832	.5871	.5910	.5948	.5987	.6026	.6064	.6103	.6141
.3	.6179	.6217	.6255	.6293	.6331	.6368	.6406	.6443	.6480	.6517
.4	.6554	.6591	.6628	.6664	.6700	.6736	.6772	.6808	.6844	.6879
.5	.6915	.6950	.6985	.7019	.7054	.7088	.7123	.7157	.7190	.7224
.6	.7257	.7291	.7324	.7357	.7389	.7422	.7454	.7486	.7517	.7549
.7	.7580	.7611	.7642	.7673	.7704	.7734	.7764	.7794	.7823	.7852
.8	.7881	.7910	.7939	.7967	.7995	.8023	.8051	.8078	.8106	.8133
.9	.8159	.8186	.8212	.8238	.8264	.8289	.8315	.8340	.8365	.8389
1.0	.8413	.8438	.8461	.8485	.8508	.8531	.8554	.8577	.8599	.8621
1.1	.8643	.8665	.8686	.8708	.8729	.8749	.8770	.8790	.8810	.8830
1.2	.8849	.8869	.8888	.8907	.8925	.8944	.8962	.8980	.8997	.9015
1.3	.9032	.9049	.9066	.9082	.9099	.9115	.9131	.9147	.9162	.9177
1.4	.9192	.9207	.9222	.9236	.9251	.9265	.9279	.9292	.9306	.9319
1.5	.9332	.9345	.9357	.9370	.9382	.9394	.9406	.9418	.9429	.9441
1.6	.9452	.9463	.9474	.9484	.9495	.9505	.9515	.9525	.9535	.9545
1.7	.9554	.9564	.9573	.9582	.9591	.9599	.9608	.9616	.9625	.9633
1.8	.9641	.9649	.9656	.9664	.9671	.9678	.9686	.9693	.9699	.9706
1.9	.9713	.9719	.9726	.9732	.9738	.9744	.9750	.9756	.9761	.9767
2.0	.9772	.9778	.9783	.9788	.9793	.9798	.9803	.9808	.9812	.9817
2.1	.9821	.9826	.9830	.9834	.9838	.9842	.9846	.9850	.9854	.9857
2.2	.9861	.9864	.9868	.9871	.9875	.9878	.9881	.9884	.9887	.9890
2.3	.9893	.9896	.9898	.9901	.9904	.9906	.9909	.9911	.9913	.9916
2.4	.9918	.9920	.9922	.9925	.9927	.9929	.9931	.9932	.9934	.9936
2.5	.9938	.9940	.9941	.9943	.9945	.9946	.9948	.9949	.9951	.9952
2.6	.9953	.9955	.9956	.9957	.9959	.9960	.9961	.9962	.9963	.9964
2.7	.9965	.9966	.9967	.9968	.9969	.9970	.9971	.9972	.9973	.9974
2.8	.9974	.9975	.9976	.9977	.9977	.9978	.9979	.9979	.9980	.9981
2.9	.9981	.9982	.9982	.9983	.9984	.9984	.9985	.9985	.9986	.9986
3.0	.9987	.9987	.9987	.9988	.9988	.9989	.9989	.9989	.9990	.9990
3.1	.9990	.9991	.9991	.9991	.9992	.9992	.9992	.9992	.9993	.9993
3.2	.9993	.9993	.9994	.9994	.9994	.9994	.9994	.9995	.9995	.9995
3.3	.9995	.9995	.9995	.9996	.9996	.9996	.9996	.9996	.9996	.9997
3.4	.9997	.9997	.9997	.9997	.9997	.9997	.9997	.9997	.9997	.9998

x	1.282	1.645	1.960	2.326	2.576	3.090	3.291	3.891	4.417
$F(x)$	.90	.95	.975	.99	.995	.999	.9995	.99995	.999995
$2[1 - F(x)]$	.20	.10	.05	.02	.01	.002	.001	.0001	.00001

അനുബന്ധം 4

സ്റ്റുഡൻറ് t സാരണി

$$F(t) = \int_{-\infty}^{t} \frac{\left(\frac{n-1}{2}\right)!}{\left(\frac{n-2}{2}\right)!\sqrt{\pi n}\left(1+\frac{x^2}{n}\right)^{(n+1)/2}}\,dx$$

n \ F	.75	.90	.95	.975	.99	.995	.9995
1	1.000	3.078	6.314	12.706	31.821	63.657	636 619
2	.816	1.886	2.920	4.303	6.965	9.925	31.598
3	.765	1.638	2.353	3.182	4.541	5.841	12.941
4	.741	1.533	2.132	2.776	3.747	4.604	8.610
5	.727	1.476	2 015	2.571	3.365	4.032	6.859
6	.718	1.440	1.943	2.447	3.143	3.707	5.959
7	.711	1.415	1.895	2.365	2.998	3.499	5.405
8	.706	1.397	1.860	2.306	2.896	3.355	5.041
9	.703	1.383	1.833	2.262	2.821	3.250	4.781
10	.700	1.372	1.812	2.228	2.764	3.169	4 587
11	.697	1.363	1.796	2.201	2.718	3.106	4.437
12	.695	1.356	1.782	2.179	2.681	3.055	4.318
13	.694	1.350	1.771	2.160	2.650	3.012	4.221
14	.692	1.345	1.761	2.145	2.624	2.977	4.140
15	.691	1.341	1.753	2.131	2.602	2.947	4.073
16	.690	1.337	1.746	2.120	2.583	2.921	4.015
17	.689	1.333	1.740	2.110	2.567	2.898	3.965
18	.688	1.330	1.734	2.101	2.552	2.878	3.922
19	.688	1.328	1.729	2.093	2.539	2.861	3.883
20	.687	1.325	1.725	2.086	2.528	2.845	3.850
21	.686	1.323	1.721	2.080	2 518	2.831	3.819
22	.686	1.321	1.717	2.074	2 508	2.819	3.792
23	.685	1.319	1.714	2.069	2.500	2.807	3.767
24	.685	1.318	1.711	2.064	2.492	2.797	3.745
25	.684	1.316	1.708	2.060	2.485	2.787	3.725
26	.684	1.315	1.706	2.056	2.479	2.779	3.707
27	.684	1.314	1.703	2.052	2.473	2.771	3.690
28	.683	1.313	1.701	2.048	2.467	2.763	3.674
29	.683	1.311	1.699	2.045	2.462	2.756	3.659
30	.683	1.310	1.697	2.042	2.457	2.750	3.646
40	.681	1.303	1.684	2.021	2.423	2.704	3.551
60	.679	1.296	1.671	2.000	2.390	2.660	3.460
120	.677	1.289	1.658	1.980	2.358	2.617	3.373
∞	.674	1.282	1.645	1.960	2.326	2.576	3.291

അനുബന്ധം 5

യാദൃച്ഛിക സംഖ്യാസാരണി

01	02	03	04	05	06	07	08	09	10	11	12	13	14	15
25	19	17	50	50	46	26	92	62	41	27	66	85	60	70
54	61	41	41	91	88	83	30	32	75	59	03	58	58	83
97	50	71	35	65	67	15	45	73	09	17	60	68	38	05
96	17	27	35	82	80	77	28	97	11	26	72	02	88	96
21	48	84	49	72	93	48	66	75	82	36	33	77	97	35
85	12	09	36	72	81	06	73	04	02	03	10	81	34	44
49	57	40	54	64	88	97	69	03	12	94	45	86	74	66
07	43	79	37	60	96	75	39	46	33	42	41	29	83	73
80	07	51	15	59	55	24	80	49	12	61	68	00	44	58
40	71	81	93	03	03	60	02	42	53	38	35	05	67	73
50	24	44	84	14	02	13	95	71	17	46	16	45	72	36
51	36	08	02	99	65	46	51	84	51	20	85	22	94	38
62	81	28	56	90	81	19	95	58	41	50	80	91	11	62
83	33	85	65	91	68	33	17	85	77	15	53	18	87	75
24	05	75	46	93	05	64	39	09	20	73	52	84	82	81
28	40	31	45	53	96	36	84	57	60	99	82	84	93	66
21	23	47	38	68	53	19	50	06	54	28	00	56	78	63
00	78	78	51	53	72	74	90	79	03	63	27	02	60	44
66	96	71	70	61	05	98	64	67	41	35	00	84	20	51
46	24	17	92	11	04	92	17	17	89	52	52	65	59	36
55	69	47	19	10	36	47	63	23	35	15	03	79	56	48
75	17	81	21	31	84	98	99	77	96	71	72	67	99	24
35	04	66	64	83	34	75	18	40	58	65	35	98	48	02
05	83	68	55	63	72	35	53	51	48	26	41	11	16	45
45	48	17	48	46	21	44	18	99	41	51	94	64	83	03
88	44	33	02	47	97	47	04	12	38	93	25	03	29	72
49	91	93	73	14	15	01	47	02	70	30	96	01	06	30
45	42	46	06	93	60	41	09	31	29	52	49	68	82	39
50	69	74	10	51	89	66	51	57	21	54	95	58	76	46
18	56	73	16	02	87	41	05	13	87	13	61	08	73	29
43	73	70	73	19	41	04	60	25	42	09	50	42	45	01
52	69	34	01	65	33	19	62	22	41	29	65	24	43	22
01	15	92	69	53	78	68	58	74	08	05	11	38	94	28
94	46	83	72	49	19	98	09	56	83	25	40	01	22	61
44	42	06	32	95	17	32	67	80	84	09	69	57	52	92

16	17	18	19	20	21	22	23	24	25	26	27	28	29	30
81	58	85	33	16	11	87	12	17	39	12	11	07	72	20
60	25	84	42	22	94	38	96	52	03	38	97	12	87	15
53	12	75	59	76	42	73	48	95	57	51	31	12	50	82
02	68	01	17	09	00	38	12	31	52	22	24	73	89	09
95	68	53	92	82	11	96	03	47	31	35	59	02	23	84
00	32	10	43	45	44	48	02	29	03	71	82	60	44	48
67	16	84	57	42	18	97	25	03	16	56	57	02	46	13
11	23	91	28	97	34	06	48	44	87	56	80	11	02	46
75	25	43	39	13	14	29	63	79	33	69	90	40	59	83
16	67	93	59	86	81	53	07	69	33	47	40	14	70	07
25	81	18	46	46	96	68	34	08	88	78	35	34	55	49
80	49	70	27	17	99	43	11	36	95	04	05	19	52	40
44	44	96	11	09	82	38	91	73	62	44	72	30	09	91
46	90	22	50	50	53	83	95	82	13	26	25	16	55	89
35	48	16	10	07	10	67	28	66	79	16	26	74	55	78
50	01	66	62	10	38	47	86	17	59	64	26	02	36	17
13	35	98	13	29	61	37	85	44	14	96	63	98	71	28
68	37	23	74	77	92	37	14	25	88	78	96	90	90	00
40	43	78	99	64	47	23	89	80	49	91	95	59	60	06
90	37	63	74	14	30	64	66	72	38	19	28	01	63	44
04	84	87	41	64	03	89	57	82	34	07	71	33	49	80
43	95	90	88	46	27	34	43	61	52	24	53	09	84	27
94	20	01	52	38	82	74	59	52	76	29	85	59	84	16
64	04	67	90	38	25	44	69	32	35	04	27	03	98	84
91	89	73	11	07	29	69	79	89	36	79	99	56	05	63
24	43	43	01	91	48	33	23	60	63	87	15	15	27	59
77	67	34	95	86	99	27	54	40	61	32	54	74	63	89
09	91	[illegible]5	96	96	59	13	33	76	69	65	15	88	82	08
36	59	12	33	44	28	85	77	72	84	23	05	57	14	43
67	03	48	83	77	15	39	38	60	87	93	20	89	37	55
87	07	87	94	15	70	33	87	92	20	44	52	85	28	63
70	83	47	08	44	92	03	01	69	36	54	02	85	92	92
35	61	24	35	08	63	55	43	88	72	23	80	06	83	24
33	90	47	53	07	64	57	02	75	91	23	41	95	06	18
10	86	00	20	21	25	38	66	72	50	88	21	00	24	82

ശബ്ദാവലിയും സൂചികയും

രാമകൃഷ്ണപിള്ള, കെ

സാമ്പിള സർവെ.