6 പ്രകീർണനം (Dispersion)

X8L1V6

ആമുഖം

വിവിധതരം കേന്ദ്രപ്രവണതാമാനങ്ങളാണ് കഴിഞ്ഞ അധ്യായത്തിൽ നാം പഠിച്ചത്. ഒരു കേന്ദ്രവില ഉപയോഗിച്ച് ഡാറ്റാവിതരണത്തിന്റെ സവിശേഷതകൾ എങ്ങനെ വിവരിക്കാം എന്നതാണവിടെ ചർച്ച ചെയ്തത്. എന്നാൽ ഈ അളവുകൾ മാത്രമുപയോഗിച്ച് ഒരു വിതരണത്തിന്റെ വ്യക്തമായ ധാരണ രൂപീകരിക്കാൻ നമുക്ക് കഴിയില്ല. താഴെ പറയുന്ന ഡാറ്റകൾ പരിശോധിക്കുക.

A	28	29	30	30	32
B	30	30	30	30	30
C	1	2	30	30	87

ഈ മൂന്ന് ഡാറ്റകളുടെയും മാധ്യം, മധ്യാങ്കം, മോഡ് എന്നിവ കണ്ടുപിടിച്ചാൽ ഇവ തുല്യമാണെന്നും വില 30 ആണെന്നും കാണാം. മൂന്ന് ഡാറ്റയിലും അംഗങ്ങളുടെ എണ്ണം തുല്യമാണ്. എങ്കിലും ഈ മൂന്ന് ഡാറ്റകളും വ്യത്യസ്തമാണ്. ഇവ തമ്മിലുള്ള പ്രധാന വ്യത്യാസം എന്താണ്?

മൂന്ന് ഡാറ്റകളുടെയും ശരാശരികൾ തുല്യമാണെങ്കിലും പ്രാപ്താങ്കങ്ങൾ തമ്മിലുള്ള വ്യതിയാനങ്ങൾ വ്യത്യാസപ്പെട്ടിരിക്കുന്നു. ഡാറ്റ C യിലെ വിലകൾ മറ്റ് രണ്ട് ഡാറ്റയിലെ വിലകളേക്കാൾ കൂടുതൽ ചിതറി കിടക്കുന്നു. ഈ മൂന്ന് ഡാറ്റകളും പരിശോധിക്കുമ്പോൾ 'C' യിലെ പ്രാപ്താങ്കങ്ങളാണ് അതിന്റെ മാധ്യത്തിൽ നിന്നും കൂടുതൽ അകലെയായി വിതരണം ചെയ്തിരിക്കുന്നത്. ശരാശരി വിലയിൽ നിന്നും വ്യതിചലിച്ചു വ്യാപിക്കാനുള്ള ഒരു സംഖ്യാഡാറ്റയുടെ പ്രവണതയുടെ അളവിനെയാണ് ആ ഡാറ്റയുടെ പ്രകീർണനം അല്ലെങ്കിൽ വ്യതിയാനം എന്നു വിളിക്കുന്നത്.

പ്രകീർണനം (Dispersion)

> ഒരു കേന്ദ്രവിലയിൽ നിന്നും ഒരു ചരത്തിന്റെ വിലകളുടെ വ്യാപനത്തിന്റെ അളവാണ് പ്രകീർണനം.

പ്രകീർണനത്തിന് വിവിധ അളവുകൾ ഉണ്ട് അവ താഴെ പറയുന്നു.

1. പരിധി (Range)
2. ചതുരംശവ്യതിയാനം (Quartile Deviation)
3. വ്യതിയാന മാധ്യം (Mean Deviation)
4. മാനക വ്യതിയാനം (Standard Deviation)

പ്രകീർണന അളവുകളുടെ സവിശേഷതകൾ

ഒരു നല്ല പ്രകീർണന അളവിന് താഴെ പറയുന്ന സവിശേഷതകൾ ഉണ്ടായിരിക്കണം.

- വ്യക്തമായി നിർവചിക്കപ്പെട്ടിരിക്കണം
- എല്ലാ പ്രാപ്താങ്കങ്ങളെയും അടിസ്ഥാനപ്പെടുത്തിയുള്ളതാവണം.
- മനസ്സിലാക്കാൻ ലളിതവും കണ്ടുപിടിക്കാൻ എളുപ്പവുമായിരിക്കണം.
- തുടർ ഗണിതപ്രക്രിയകൾക്ക് പറ്റുന്നതായിരിക്കണം.
- അഗ്ര വിലകളാൽ അമിതമായി സ്വാധീനിക്കപ്പെടുന്നതാവരുത്.

പ്രകീർണന അളവുകളുടെ പ്രാധാന്യം

- ഡാറ്റയുടെ വ്യതിയാനത്തെക്കുറിച്ച് ധാരണ നൽകുന്നു.
- രണ്ടോ അതിലധികമോ വിതരണങ്ങൾ താരതമ്യം ചെയ്യുന്നതിന് സഹായിക്കുന്നു.
- ഉയർന്ന തലത്തിലുള്ള സാംഖ്യക വിശകലനങ്ങളിൽ മുഖ്യ പങ്ക് വഹിക്കുന്നു.

6.1 പരിധി (Range)

ഒരു ഡാറ്റയിലെ ഏറ്റവും കൂടിയ വിലയും (H) ഏറ്റവും കുറഞ്ഞ വിലയും (L) തമ്മിലുള്ള വ്യത്യാസമാണ് പരിധി (R) .അതായത് $R = H - L$

- പരിധിയുടെ ഉയർന്ന വിലകൾ കൂടുതൽ വ്യതിയാനത്തെ സൂചിപ്പിക്കുന്നു.

മേന്മകൾ

- ഇത് ഏറ്റവും ലളിതമായ പ്രകീർണന അളവാണ്
- ഇത് ലളിതമായി മനസ്സിലാക്കുവാനും എളുപ്പത്തിൽ കണ്ടുപിടിക്കുവാനും കഴിയുന്നതാണ്.

ന്യൂനതകൾ

- അഗ്രവിലകൾ മാത്രം കണക്കിലെടുക്കുന്നതിനാൽ അവയാൽ അമിതമായി സ്വാധീനിക്കപ്പെടുന്നു.
- ഇത് പല സന്ദർഭങ്ങളിലും അവലംബിക്കാവുന്ന ഒരു പ്രകീർണന അളവ് അല്ല.

വിവരണം 6.1

2013 ആഗസ്റ്റിലെ ആദ്യ ആഴ്ചയിൽ ഒരു ഗ്രാം സ്വർണത്തിന്റെ വിലകൾ രൂപയിൽ താഴെ തന്നിരിക്കുന്നു.

തിങ്കൾ	ചൊവ്വ	ബുധൻ	വ്യാഴം	വെള്ളി	ശനി
2530	2500	2430	2560	2680	2600

ഈ ആഴ്ചയിലെ വിലകളുടെ പരിധി കണ്ടുപിടിക്കുക

പരിഹാരം

പരിധി = കൂടിയ വില - കുറഞ്ഞ വില

= 2680-2430

= 250 രൂപ

പ്രവർത്തനം

10, 30, 50, 100 എന്ന ശ്രേണി ശ്രദ്ധിക്കുക.

- 30 എന്ന വിലയ്ക്കു പകരം 20 ആയാൽ പരിധി കാണുക.
- 100 എന്ന വില ഒഴിവാക്കി പരിധി കാണുക.

6.2 ചതുരംശ വ്യതിയാനം (Quartile Deviation)

അഗ്ര വിലകളെ മാത്രം കണക്കിലെടുക്കുന്നു എന്നതും തുറന്ന ക്ലാസുള്ള വിതരണത്തിന് കണ്ടുപിടിക്കാൻ കഴിയില്ല എന്നതും പരിധിയുടെ പരിമിതികളാണ്. അഗ്ര വിലകൾക്ക് പകരം ചതുരംശങ്ങളെ ഉപയോഗിക്കുകയാണെങ്കിൽ ഈ പരിമിതിയെ മറികടക്കാൻ സാധിക്കും. മൂന്നാം ചതുരംശത്തിന്റെയും ഒന്നാം ചതുരംശത്തിന്റെയും വ്യത്യാസത്തെ ചതുരംശാന്തര പരിധി (Inter Quartile range) എന്നു പറയുന്നു. ഇതിന്റെ പകുതിയെ ചതുരംശവ്യതിയാനം (QD) എന്നും വിളിക്കുന്നു.

ചതുരംശവ്യതിയാനത്തെ അർധ ചതുരംശാന്തര പരിധി എന്നും വിളിക്കുന്നു.

ചതുരംശാന്തര പരിധി = $Q_3 - Q_1$

ചതുരംശ വ്യതിയാനം = $\frac{Q_3 - Q_1}{2}$

Q_3= മൂന്നാം ചതുരംശം, Q_1= ഒന്നാം ചതുരംശം

ഒരു അസംസ്കൃത ഡാറ്റയുടെ ചതുരംശ വ്യതിയാനം

n വിലകളുള്ള ഒരു അസംസ്കൃത ഡാറ്റയുടെ ചതുരംശ വ്യതിയാനം കാണുന്നതിന്, അവയെ ആരോഹണക്രമത്തിലാക്കി Q_1 , Q_3 എന്നിവ കണ്ടുപിടിച്ച് $\frac{Q_3 - Q_1}{2}$ കണ്ടാൽ മതിയാകും.

$$QD = \frac{Q_3 - Q_1}{2}$$

വിവരണം 6.2

11 സ്കൂൾ വിദ്യാർഥികൾ ഒരാഴ്ചയിൽ ടി.വി കാണാൻ ചെലവഴിക്കുന്ന സമയത്തിന്റെ ഡാറ്റ മണിക്കൂറിൽ തന്നിരിക്കുന്നു. 3, 8.5, 12, 9, 16.5, 9, 14, 20, 18, 19, 20 ചതുരംശവ്യതിയാനം കാണുക.

പരിഹാരം

തന്നിരിക്കുന്ന വിലകളെ ആരോഹണ ക്രമത്തിലാക്കിയാൽ

3, 8.5, 9, 9, 12, 14, 16.5, 18, 19, 20, 20

$$Q_1 = \left(\frac{n+1}{4}\right) \text{ -ാമത്തെ വില}$$

$$= \left(\frac{11+1}{4}\right) \text{ -ാമത്തെ വില}$$

$$= 3 \text{ -ാമത്തെ വില} = 9$$

$$Q_3 = 3\left(\frac{n+1}{4}\right) \text{ -ാമത്തെ വില}$$

$$= 3\left(\frac{11+1}{4}\right) \text{ -ാമത്തെ വില}$$

$$= 9 \text{ -ാമത്തെ വില} = 19$$

$$QD = \frac{Q_3 - Q_1}{2}$$

$$= \frac{19-9}{2} = 5$$

വേറിട്ട വിതരണത്തിന്റെ ചതുരംശവ്യതിയാനം

N വിലകളുള്ള ഒരു വേറിട്ട വിതരണത്തിന്റെ ചതുരംശ വ്യതിയാനം കാണുന്നതിന് അവയെ ആരോഹണക്രമത്തിലാക്കി Q_1, Q_3 എന്നിവ കണ്ടുപിടിച്ച് $QD = \frac{Q_3 - Q_1}{2}$ എന്ന സൂത്രവാക്യം ഉപയോഗിക്കാം.

വിവരണം 6.3

താഴെ തന്നിരിക്കുന്ന ഡാറ്റയുടെ QD കാണുക

സ്കോർ	5	10	15	20	25	30
കുട്ടികളുടെ എണ്ണം	4	7	15	8	7	2

പരിഹാരം

ആരോഹണ സഞ്ചിതാവൃത്തിപ്പട്ടിക തയാറാക്കിയാൽ

സ്കോർ (x)	കുട്ടികളുടെ എണ്ണം (f)	ആരോഹണസഞ്ചിതാവൃത്തി
5	4	4
10	7	11
15	15	26
20	8	34
25	7	41
30	2	43
	N = 43	

$Q_1 = \left(\frac{N+1}{4}\right)$ാമത്തെ വില $= \left(\frac{43+1}{4}\right)$ മത്തെ വില

= 11 ാമത്തെ വില

=10

$Q_3 = 3\left(\frac{N+1}{4}\right)$ ാമത്തെ വില $= 3\left(\frac{43+1}{4}\right)$ ാമത്തെ വില

= 33 ാമത്തെ വില

= 20

$$QD = \frac{Q_3 - Q_1}{2}$$

$$= \frac{20-10}{2}$$

= 5

തുടർ ആവൃത്തി വിതരണത്തിന്റെ ചതുരംശ വ്യതിയാനം

ഒരു തുടർ ആവൃത്തി വിതരണത്തിന്റെ Q_1, Q_3 എന്നിവ കണ്ടുപിടിക്കാൻ 5-ാം അധ്യായത്തിലെ 6-ാം ഭാഗത്ത് നാം പഠിച്ചിട്ടുണ്ട്.

$$QD = \frac{Q_3 - Q_1}{2}$$

ഉദാഹരണം 6.4

ഒരു കോളനിയിലെ ഗാർഹിക വൈദ്യുതി ഉപഭോഗസർവ്വേയുടെ വിവരങ്ങളാണ് താഴെ തന്നിരിക്കുന്നത്.

യൂണിറ്റ്	100ൽ താഴെ	100 - 200	200 - 300	300 - 400	400 - 500	500 - 600	600 - 700	700ന് മുകളിൽ
ഉപഭോക്താക്കളുടെ എണ്ണം	20	21	30	46	20	25	16	10

ചതുരംശവ്യതിയാനം കണ്ടുപിടിക്കുക.

പരിഹാരം

ആരോഹണ സഞ്ചിതാവൃത്തിപ്പട്ടിക തയാറാക്കുക.

യൂണിറ്റ്	ആവൃത്തി	സഞ്ചിതാവൃത്തി
100ൽ താഴെ	20	20
100 - 200	21	41
200 - 300	30	71
300 - 400	46	117
400 - 500	20	137
500 - 600	25	162
600 - 700	16	178
700ന് മുകളിൽ	10	188
	N=188	

$N = 188$

$\frac{N}{4} = 47$

$\frac{3N}{4} = 141$

$l_1 = 200$

$l_3 = 500$

$m_1 = 41$

$m_3 = 137$

$c_1 = 100$

$c_3 = 100$

$f_1 = 30$

$f_3 = 25$

$$Q_1 = \ell_1 + \left(\frac{\frac{N}{4} - m_1}{f_1}\right) c_1 = 200 + (47 - 41) \times \frac{100}{30} = 220$$

$$Q_3 = \ell_3 + \left(\frac{\frac{3N}{4} - m_3}{f_3} \right) c_3 = 500 + \frac{141 - 137}{25} \times 100 = 516$$

$$Q.D = \frac{516 - 220}{2} = 148$$

മേന്മകൾ

- തുറന്ന ക്ലാസുള്ള വിതരണത്തിനും ഇത് കണ്ടുപിടിക്കാം.
- അഗ്ര വിലകളാൽ അമിതമായി സ്വാധീനിക്കപ്പെടുന്നില്ല.

ന്യൂനതകൾ

- എല്ലാ പ്രാപ്താങ്കങ്ങളെയും കണക്കിലെടുക്കുന്നില്ല.
- തുടർ ബീജഗണിത പ്രക്രിയകൾക്ക് കഴിയാത്തവയാണ്.

നിങ്ങളുടെ പുരോഗതി അറിയുക

1. ഒരു വലിയ നഗരത്തിലെ 50 കച്ചവടക്കാരുടെ വാർഷിക ലാഭം (ലക്ഷം രൂപയിൽ) കാണിക്കുന്ന പട്ടികയാണ് താഴെ തന്നിരിക്കുന്നത് Q.D കണ്ടുപിടിക്കുക.

ലാഭം	20-29	30-39	40-49	50-59	60-69
കച്ചവടക്കാരുടെ എണ്ണം	8	12	20	7	3

2. താഴെ തന്നിരിക്കുന്നത് വിദ്യാർഥികളുടെ ഉയരങ്ങൾ (ഇഞ്ചിൽ) ആണ്. അവയുടെ Q.D കാണുക

   ഉയരം : 55, 54, 57, 67, 60, 61, 58, 63

3. ഒരു കമ്പനിയുടെ 40 ദിവസങ്ങളിലെ ഓഹരി വിലകളുടെ വിതരണം താഴെ തന്നിരിക്കുന്നു. Q.D കാണുക.

വില	150	200	190	210	230	180
ദിവസങ്ങളുടെ എണ്ണം	5	5	8	10	5	7

പ്രവർത്തനം

നിങ്ങളുടെ ക്ലാസിലെ കുട്ടികളുടെ ഉയരം ശേഖരിച്ച് ഝ.ഉ, റേഞ്ച് എന്നിവ കാണുക. ഉത്തരം വിശകലനം ചെയ്യുക.

6.3. വ്യതിയാനമാധ്യം (Mean Deviation)

പരിധിയേക്കാൾ മികച്ചതാണ് ചതുരംശവ്യതിയാന മാധ്യം എങ്കിലും ഇത് Q_1, Q_3 എന്നിവയെ മാത്രം ആശ്രയിക്കുന്നതാണ്. ഇതിനു പരിഹാരമായി എല്ലാ വിലകളേയും ആശ്രയിക്കുന്ന മറ്റൊരു അളവ് അവതരിപ്പിക്കാൻ ശ്രമിക്കുകയാണ് നാം. വ്യതിയാന മാധ്യം എന്നത് തിരഞ്ഞെടുത്ത ഒരു ശരാശരിയിൽ നിന്നും പ്രാപ്താങ്കങ്ങളുടെ കേവല വ്യതിയാന

ങ്ങളുടെ മാധ്യം ആണ്.

അസംസ്കൃത ഡാറ്റയുടെ വ്യതിയാന മാധ്യം

n പ്രാപ്താങ്കങ്ങളുള്ള ഒരു അസംസ്കൃത ഡാറ്റ പരിഗണിക്കുക. മീൻഡീവിയേഷൻ താഴെ പറയും പ്രകാരം കണ്ടുപിടിക്കാം.

വ്യതിയാന മാധ്യം $= \frac{\Sigma|x-A|}{n}$, ഇവിടെ A എന്നത് ഏതെങ്കിലും ഒരു ശരാശരിയാണ്

ഒരു ശരാശരി (A) കണ്ടുപിടിക്കുക. (മീൻ, മീഡിയൻ, മോഡ്) ഓരോ വിലയിൽ നിന്നും ഈ ശരാശരിയിലേക്കുള്ള കേവല വ്യതിയാനം കണ്ടുപിടിക്കുക. (ie|x A|)ഈ കേവല വ്യതിയാനങ്ങളുടെ മാധ്യം കണ്ടുപിടിക്കുക.

വിവരണം 6.5

11 വിദ്യാർഥികളെ തെരഞ്ഞെടുത്ത് അവർ ഓരോരുത്തരും സ്റ്റാറ്റിസ്റ്റിക്സ് പരീക്ഷയുടെ തലേ ദിവസം എത്ര മണിക്കൂർ വീതം പഠിച്ചെന്നു ചോദിക്കുന്നു. അവരുടെ ഉത്തരങ്ങൾ 8, 11, 5, 4, 5, 0, 2, 6, 9, 3, 2 ആണ്. മാധ്യത്തിൽ നിന്നുള്ള വ്യതിയാന മാധ്യം കണ്ടുപിടിക്കുക.

പരിഹാരം

മാധ്യം $= \frac{\sum x}{n} = \frac{55}{11} = 5$

മാധ്യത്തിൽനിന്നുള്ള വ്യതിയാനമാധ്യം $= \frac{\Sigma|x - \text{മാധ്യം}|}{n}$

പഠന സമയം	$\|x - \text{മാധ്യം}\|$
8	$\|8-5\| = 3$
11	$\|11-5\| = 6$
5	$\|5-5\| = 0$
4	$\|4-5\| = 1$
5	$\|5-5\| = 0$
0	$\|0-5\| = 5$
2	$\|2-5\| = 3$
6	$\|6-5\| = 1$
9	$\|9-5\| = 4$
3	$\|3-5\| = 2$
2	$\|2-5\| = 3$
ആകെ	$\sum\|x - \text{മാധ്യം}\| = 28$

മാധ്യത്തിൽ നിന്നുള്ള വ്യതിയാനമാധ്യം $= \frac{\Sigma|x - \text{മാധ്യം}|}{n} = \frac{28}{11} = 2.54$

വേറിട്ട ആവൃത്തി വിതരണത്തിന്റെ വ്യതിയാനമാധ്യം

N വിലകളുള്ള ഒരു വേറിട്ട ആവൃത്തി വിതരണത്തിന്റെ വ്യതിയാനമാധ്യം കാണാനുള്ള സമവാക്യം ചുവടെ നൽകുന്നു.

വ്യതിയാനമാധ്യം $= \frac{\Sigma f|x - A|}{N}$ ഇവിടെ A എന്നത് ഏതെങ്കിലും ഒരു ശരാശരി ആയിരിക്കും.

മീൻ ഡീവിയേഷൻ കണ്ടുപിടിക്കുന്ന രീതി ചുവടെ നൽകുന്നു.

ഘട്ടം 1

വ്യതിയാനമാധ്യം ഏത് ശരാശരിയെ അടിസ്ഥാനപ്പെടുത്തിയാണ് കണ്ടുപിടിക്കുന്നത് എന്ന് തീരുമാനിക്കുകയും ആ ശരാശരി കണ്ടുപിടിക്കുകയും ചെയ്യുക. ഇങ്ങനെ കിട്ടുന്ന ശരാശരിയുടെ വിലയെ A എന്ന അക്ഷരം കൊണ്ട് പ്രതിനിധീകരിക്കാം.

ഘട്ടം 2

ഈ ശരാശരിയിൽ നിന്നും ഓരോ വിലയിലേക്കുമുള്ള വ്യത്യാസത്തിന്റെ കേവല വില, $|x - A|$ കാണുന്നു.

ഘട്ടം 3

$|x - A|$ യെ അവയുടെ ആവൃത്തികൊണ്ട് ഗുണിക്കുന്നു. അതായത് $f|x - A|$ കാണുന്നു.

ഘട്ടം 4

$$MD = \frac{\sum f\,|X - A|}{N}$$

ഉദാഹരണം 6.6

ഒരു സാഹിത്യ ക്ലാസിലെ 28 വിദ്യാർഥികൾ ഒരു മാസത്തിൽ വായിച്ച പുസ്തകങ്ങളുടെ എണ്ണത്തിന്റെ ആവൃത്തിപ്പട്ടിക തന്നിരിക്കുന്നു.

പുസ്തകങ്ങളുടെ എണ്ണം	0	1	2	3	4
കുട്ടികളുടെ എണ്ണം	2	6	12	5	3

മോഡിൽ നിന്നുള്ള വ്യതിയാന മാധ്യം (മീൻ ഡീവിയേഷൻ) കാണുക

പരിഹാരം

$$\text{വ്യതിയാനമാധ്യം} = \frac{\Sigma f\left|x - \text{മോഡ്}\right|}{N}$$

ഇവിടെ മോഡ് = 2 (ഏറ്റവും കൂടുതൽ ആവൃത്തിയുള്ള വില)

പുസ്തകങ്ങളുടെ എണ്ണം	x - മോഡ്|	f	f|x - മോഡ്|
0	|0-2|= 2	2	4
1	|1-2|=1	6	6
2	|2-2|= 0	12	0
3	|3-2| = 1	5	5
4	|4-2| = 2	3	6
		N = 28	Σ f x - മോഡ്|=21

$$\text{വ്യതിയാനമാധ്യം} = \frac{21}{28}$$

$$= 0.75$$

തുടർ ആവ്യത്തി വിതരണത്തിന്റെ വ്യതിയാനമാധ്യം

തുടർ ആവൃത്തി വിതരണത്തിന്റെ വ്യതിയാന മാധ്യം കണ്ടുപിടിക്കുന്നതിന് വിതരണത്തിന്റെ ശരാശരി കണ്ടുപിടിച്ചതിനുശേഷം ഓരോ ക്ലാസിന്റെയും മധ്യവില കണ്ടുപിടിച്ച് തുടർ ആവൃത്തി വിതരണത്തെ വേറിട്ട ആവൃത്തി വിതരണമാക്കി മാറ്റുക. അതിനു ശേഷം വേറിട്ട ആവൃത്തി വിതരണരീതി പിന്തുടരുക.

വിവരണം 6.7

ചുവടെ കൊടുത്ത ഡാറ്റയുടെ മധ്യാങ്കത്തിൽ നിന്നുള്ള വ്യതിയാനമാധ്യം കണ്ടുപിടിക്കുക.

ഉയരം (സെ.മീ.യിൽ)	100 - 120	120 - 140	140 - 160	160 - 180	180 - 200
വിദ്യാർഥികളുടെ എണ്ണം	4	6	10	8	5

പരിഹാരം

$$\text{വ്യതിയാനമാധ്യം} = \frac{\Sigma f\left|x - \text{മധ്യാങ്കം}\right|}{N}$$

ക്ലാസ്	മധ്യവില (x)	ആവൃത്തി (f)	സഞ്ചിതാവൃത്തി	$\lvert x - \text{മധ്യാങ്കം} \rvert$	$f \lvert x - \text{മധ്യാങ്കം} \rvert$
100 - 120	110	4	4	$\lvert 110 - 153 \rvert = 43$	172
120 - 140	130	6	10	$\lvert 130 - 153 \rvert = 23$	138
140 - 160	150	10	20	$\lvert 150 - 153 \rvert = 3$	30
160 - 180	170	8	28	$\lvert 170 - 153 \rvert = 17$	136
180 - 200	190	5	33	$\lvert 190 - 153 \rvert = 37$	185
ആകെ		33			661

$$\frac{N}{2} = \frac{33}{2} = 16.5$$

16.5 സഞ്ചിതാവൃത്തിയുള്ള ക്ലാസ് 140 - 160 ആണ്

∴ മീഡിയൻ ക്ലാസ് = 140 - 160

$$\text{മധ്യാങ്കം} = \ell + \frac{\left(\frac{N}{2} - m\right)c}{f}$$

$$= 140 + \frac{(16.5 - 10) \times 20}{10} = 153$$

മധ്യാങ്കത്തിൽ നിന്നുള്ള വ്യതിയാനമാധ്യം

$$= \frac{\Sigma \lvert x - \text{മധ്യാങ്കം} \rvert}{N} = \frac{661}{33} = 20.03$$

വ്യതിയാന മാധ്യത്തിന്റെ മേന്മകളും ന്യൂനതകളും

മേൻമകൾ

- ഇത് എല്ലാ വിലകളെയും ആശ്രയിച്ചുള്ളതാണ്.
- ഏതു വിലയിൽ നിന്നു വേണമെങ്കിലും കണക്കാക്കാം.
- അഗ്രവിലകളുടെ സ്വാധീനം കാര്യമായി ഉണ്ടാകുന്നില്ല.

ന്യൂനതകൾ

- വ്യതിയാനങ്ങളുടെ ചിഹ്നങ്ങൾ അവഗണിക്കുന്നത് കൃത്രിമത്വം ഉണ്ടാക്കുന്നു.
- എളുപ്പത്തിൽ കണ്ടുപിടിക്കാൻ കഴിയില്ല.

നിങ്ങളുടെ പുരോഗതി അറിയുക

1) 30 കുട്ടികളുടെ ഗ്രേഡ് പോയിന്റുകളുടെ ആവൃത്തി വിതരണം ചുവടെ കൊടുക്കുന്നു

ഗ്രേഡ് പോയിന്റ്	8	6	4	2
കുട്ടികളുടെ എണ്ണം	4	20	5	1

മോഡിൽ നിന്നുള്ള വ്യതിയാനമാധ്യം കാണുക

2) നൂറ് കമ്പനികളുടെ വിറ്റുവരവിന്റെ ആവൃത്തിപ്പട്ടിക ചുവടെ നൽകുന്നു.

വിൽപന (ആയിരത്തിൽ)	40-50	50-60	60-70	70-80	80-90	90-100
ദിവസങ്ങളുടെ എണ്ണം	10	15	25	30	12	8

പ്രവർത്തനം

5, 8, 3, 2, 12 എന്നീ വിലകളുടെ

1) മാധ്യത്തിൽ നിന്നുള്ള വ്യതിയാനമാധ്യം

2) മധ്യാങ്കത്തിൽ നിന്നുള്ള വ്യതിയാനമാധ്യം

3) മോഡിൽ നിന്നുള്ള വ്യതിയാനമാധ്യം എന്നിവ കാണുക.

4) ഈ വ്യതിയാന മാധ്യങ്ങളിൽ ഏറ്റവും കുറഞ്ഞ വില ഏതെന്ന് കണ്ടെത്തുക.

6.4 മാനക വ്യതിയാനം (Standard Deviation)

ഒരു ഡാറ്റയുടെ കൂടുതൽ അർഥവത്തായ വ്യതിയാന അളവ് ലഭിക്കുന്നതിനായി നാം മാനക വ്യതിയാനം ഉപയോഗിക്കുന്നു. കാൾപിയേഴ്സനാണ് മാനക വ്യതിയാനം എന്ന ആശയം മുന്നോട്ടുവച്ചത്. ഒരു ഡാറ്റയുടെ മാനക വ്യതിയാനത്തിന്റെ വില കുറവാണ് എന്നത് ആ ഡാറ്റയിലെ വിലകൾ തമ്മിലുള്ള ഉയർന്ന ഏകാത്മകതയെ സൂചിപ്പിക്കുന്നു. എല്ലാ വിലകളും തുല്യമായാൽ മാനകവ്യതിയാനം പൂജ്യമായിരിക്കും.

മാധ്യത്തിൽ നിന്നുമുള്ള പ്രാപ്താങ്കങ്ങളുടെ വ്യതിയാനങ്ങളുടെ വർഗങ്ങളുടെ മാധ്യത്തിന്റെ പോസിറ്റീവ് വർഗമൂലമാണ് മാനകവ്യതിയാനം (SD). ഗ്രീക്ക് അക്ഷരമായ σ ഉപയോഗിച്ചാണ് ഇതു പ്രതിനിധീകരിക്കുന്നത്. ഇത് പ്രകീർണനത്തിന്റെ ഏറ്റവും മികച്ച അളവാണ്.

സവിശേഷതകൾ

1. മാനക വ്യതിയാനത്തിന്റെ ഏറ്റവും കുറഞ്ഞവില പൂജ്യമാകുന്നു.
2. വ്യതിയാനങ്ങളുടെ വർഗങ്ങളുടെ തുക ഏറ്റവും കുറവാകുന്നത് മാധ്യത്തിൽ നിന്നും വ്യതിയാനം കണക്കാക്കുമ്പോഴാണ്.

അസംസ്കൃത ഡാറ്റയുടെ മാനക വ്യതിയാനം

'n' വിലകളുടെ മാനക വ്യതിയാനം താഴെ പറയുന്ന രീതിയിൽ കണക്കാക്കാം.

$$\text{മാനക വ്യതിയാനം}, \sigma = \sqrt{\frac{\Sigma(x-\bar{x})^2}{n}}$$

ഇതിൽ നിന്നും ലഭിക്കുന്ന മറ്റൊരു സൂത്രവാക്യം താഴെ കൊടുക്കുന്നു.

മാനക വ്യതിയാനം , $\sigma = \sqrt{\frac{\Sigma x^2}{n} - \bar{x}^2}$

ഈ സൂത്രവാക്യം ഉപയോഗിച്ച് മാനകവ്യതിയാനം കണക്കാക്കുന്ന വിധം താഴെ വിവരിക്കുന്നു.

ഘട്ടം 1

തന്നിരിക്കുന്ന ഡാറ്റയുടെ മാധ്യം, $\bar{x}$ കണക്കാക്കുക.

ഘട്ടം 2

ഓരോ പ്രാപ്താങ്കങ്ങളുടെയും വർഗം (x^2) കണക്കാക്കുക.

ഘട്ടം 3

$\frac{\Sigma x^2}{n}$ കണക്കാക്കുക. (ഇവിടെ 'n' എന്നത് പ്രാപ്താങ്കങ്ങളുടെ എണ്ണമാണ്)

ഘട്ടം 4

$\sigma = \sqrt{\frac{\Sigma x^2}{n} - \bar{x}^2}$ കണക്കാക്കുക.

മാനകവ്യതിയാനം , $\sigma = \sqrt{\frac{\Sigma x^2}{n} - \bar{x}^2}$

വിവരണം 6.8

ഒരു പട്ടണത്തിൽ ഉണ്ടായ റോഡ് അപകടങ്ങളുടെ എണ്ണത്തെപ്പറ്റി തുടർച്ചയായ 11 ആഴ്ചകളിൽ ഒരു സർവെ നടത്തി. അതിന്റെ ഫലങ്ങളാണ് 8, 6, 3, 0, 5, 9, 2, 1, 3, 5, 2 എന്നിവ. റോഡ് അപകടങ്ങളുടെ എണ്ണത്തിന്റെ മാനകവ്യതിയാനം കണക്കാക്കുക.

പരിഹാരം

മാനകവ്യതിയാനം , $\sigma = \sqrt{\frac{\Sigma x^2}{n} - \bar{x}^2}$

x	8	6	3	0	5	9	2	1	3	5	2	$\Sigma x = 44$
x^2	64	36	9	0	25	81	4	1	9	25	4	$\Sigma x^2 = 258$

$$\bar{x} = \frac{44}{11} = 4$$

$$\sigma = \sqrt{\frac{258}{11} - 4^2}$$

$$= \sqrt{23.45 - 16} = 2.73$$

വേറിട്ട ആവൃത്തി വിതരണത്തിന്റെ മാനകവ്യതിയാനം

‘N’ പ്രാപ്താങ്കങ്ങളുള്ള ഒരു വേറിട്ട ആവൃത്തി വിതരണത്തിന്റെ മാനകവ്യതിയാനം കാണുന്നതിന് താഴെപ്പറയുന്ന രീതിയും സൂത്രവാക്യവുമാണ് ഉപയോഗിക്കുന്നത്.

ഘട്ടം 1 : $\bar{x} = \frac{\Sigma fx}{N}$ കാണുക

ഘട്ടം 2 : x^2, fx^2 എന്നിവ ഓരോ വിലകൾക്കും കണക്കാക്കുക.

ഘട്ടം 3 : Σfx^2 കാണുക

ഘട്ടം 4 : മാനകവ്യതിയാനം $= \sqrt{\frac{\Sigma fx^2}{N} - (\bar{x})^2}$ കണ്ടുപിടിക്കുക.

മാനകവ്യതിയാനം, $\sigma = \sqrt{\frac{\Sigma fx^2}{N} - (\bar{x})^2}$

ഉദാഹരണം 6.9

25 കുട്ടികൾക്ക് ഒരു അങ്കഗണിത പരീക്ഷ നടത്തി. അതു പൂർത്തീകരിക്കാനെടുത്ത സമയത്തിന്റെ (മിനിറ്റിൽ) വിതരണം ചുവടെ തന്നിരിക്കുന്നു. സമയത്തിന്റെ മാനകവ്യതിയാനം കണ്ടുപിടിക്കുക.

സമയം (മിനിറ്റിൽ)	1	2	3	4	5
വിദ്യാർഥികളുടെ എണ്ണം	4	3	10	5	3

പരിഹാരം

താഴെപ്പറയുന്ന രീതിയിൽ പട്ടിക തയാറാക്കുക

x	f	x^2	fx	fx^2
1	4	1	4	4
2	3	4	6	12
3	10	9	30	90
4	5	16	20	80
5	3	25	15	75
	N = 25		$\Sigma fx = 75$	$\Sigma fx^2 = 261$

$$\bar{x} = \frac{\Sigma fx}{N} = \frac{75}{25} = 3$$

$$\sigma = \sqrt{\frac{\Sigma fx^2}{N} - (\bar{x})^2} = \sqrt{\frac{261}{25} - 3^2}$$

$$= \sqrt{10.44 - 9} = \sqrt{1.44}$$

$$= 1.2$$

തുടർ ആവൃത്തി വിതരണത്തിന്റെ മാനകവ്യതിയാനം

തുടർ ആവൃത്തി വിതരണത്തിന്റെ ഓരോ ക്ലാസുകളുടേയും മധ്യവിലകൾ കണക്കാക്കി അവയെ വേറിട്ട ആവൃത്തി വിതരണമാക്കി മാറ്റുന്നു. സൂത്രവാക്യവും കണ്ടുപിടിക്കുന്ന രീതിയും വേറിട്ട ആവൃത്തി വിതരണത്തിലേതുപോലെ തന്നെയാണ്.

വിവരണം 6.10

100 എഞ്ചിനീയറിങ് കമ്പനികളെക്കുറിച്ചുള്ള പഠനത്തിൽ നിന്നും താഴെപ്പറയുന്ന വിവരങ്ങൾ ലഭിക്കുന്നു. മാനകവ്യതിയാനം കാണുക.

ലാഭം (കോടിയിൽ)	0-10	10-20	20-30	30-40	40-50	50-60
കമ്പനികളുടെ എണ്ണം	8	12	20	30	20	10

പരിഹാരം

താഴെ കൊടുത്തരീതിയിൽ പട്ടിക തയാറാക്കുക

x	f	x^2	fx	fx^2
5	8	25	40	200
15	12	225	180	2700
25	20	625	500	12500
35	30	1225	1050	36750
45	20	2025	900	40500
55	10	3025	550	30250
	N = 100		$\Sigma fx = 3220$	$\Sigma fx^2 = 122900$

$$\bar{x} = \frac{\Sigma fx}{N} = \frac{3220}{100} = 32.2$$

$$\sigma = \sqrt{\frac{\Sigma fx^2}{N} - (\bar{x})^2} = \sqrt{\frac{122900}{100} - (32.2)^2}$$

$$= \sqrt{192.16} = 13.86$$

മേന്മകൾ

1. എല്ലാ പ്രാപ്താങ്കങ്ങളെയും അടിസ്ഥാനമാക്കി വ്യക്തമായി നിർവചിക്കപ്പെട്ടിരിക്കുന്നു.
2. തുടർ ഗണിത പ്രവർത്തനങ്ങൾ നടത്തുവാൻ പര്യാപ്തമാണ്.
3. പ്രതിരൂപണത്തിലുണ്ടാകുന്ന മാറ്റം മറ്റുള്ള മാനകവ്യതിയാനങ്ങളെക്കാൾ കുറഞ്ഞ അളവിലാണ് ബാധിക്കുന്നത്.

ന്യൂനതകൾ

1. കണ്ടുപിടിക്കാൻ പ്രയാസമാണ്
2. അഗ്രം തുറന്ന ക്ലാസുകൾ‌ളുള്ള വിതരണങ്ങൾക്ക് ഇത് കണ്ടുപിടിക്കുവാൻ സാധ്യമല്ല.

നിങ്ങളുടെ പുരോഗതി അറിയുക

1. 10 വ്യക്തികളുടെ ദിവസ വേതനം താഴെ തന്നിരിക്കുന്നു. മാനകവ്യതിയാനം കണക്കാക്കുക.

   227, 235, 255, 269, 292, 299, 312, 321, 333, 348

2. ഒരു വില്ലേജിലെ 100 കുടുംബങ്ങളിലെ മാസച്ചെലവാണ് താഴെ പട്ടികയിൽ കൊടുത്തിരിക്കുന്നത്. മാനകവ്യതിയാനം കാണുക.

മാസച്ചെലവ്	0-1000	1000-2000	2000-3000	3000-4000	4000-5000	5000-6000
കുടുംബങ്ങളുടെ എണ്ണം	18	26	30	12	10	4

പ്രവർത്തനം

1. ഒരു വിദ്യാർഥി ഒരു പരീക്ഷയിലെ 6 വിഷയങ്ങൾക്ക് തുല്യ സ്കോർ നേടുന്നു. മാനകവ്യതിയാനം കണക്കാക്കി ഉത്തരം വ്യാഖ്യാനിക്കുക.
2. 40, 42, 38, 44, 46, 48, 50 എന്ന ഡാറ്റ പരിഗണിക്കുക.
   - ഇവയുടെ മാനകവ്യതിയാനം കാണുക
   - ഓരോ വിലയുടേയും കൂടെ 3 കൂട്ടിയാൽ കിട്ടുന്ന പുതിയ വിലകളുടെ മാനകവ്യതിയാനം കണക്കാക്കുക.
   - ഓരോ വിലയിൽ നിന്നും 3 കുറച്ചാൽ പുതിയ മാനകവ്യതിയാനം കണക്കാക്കുക.
   - ഓരോ വിലയേയും 3 കൊണ്ട് ഗുണിച്ചാൽ പുതിയ മാനകവ്യതിയാനം കണക്കാക്കുക.

- ഓരോ വിലയേയും 3 കൊണ്ട് ഹരിച്ചാൽ പുതിയ മാനകവ്യതിയാനം കണക്കാക്കുക.
- നിങ്ങളുടെ കണ്ടെത്തലുകൾ വ്യാഖ്യാനിക്കുക.

വ്യതിയാനം (Variance)

വ്യതിയാനം എന്ന ആശയം ആദ്യമായി അവതരിപ്പിച്ചത് ആർ.എ. ഫിഷറാണ്. മാനകവ്യതിയാനത്തിന്റെ വർഗമാണ് വ്യതിയാനം.

$$\text{വ്യതിയാനം} = \sigma^2$$

അല്ലെങ്കിൽ,

$$\text{മാനകവ്യതിയാനം}, \sigma = \sqrt{\text{വ്യതിയാനം}}$$

ഉയർന്നതലത്തിലുള്ള സാംഖ്യക വിശകലനത്തിന് ഇത് വളരെയധികം ഉപയോഗിക്കുന്നുണ്ട്.

6.5 ആപേക്ഷിക പ്രകീർണനമാനങ്ങൾ (Relative Measures of Dispersion)

നമ്മൾ ഇതുവരെ ചർച്ച ചെയ്തത് കേവല പ്രകീർണമാനങ്ങളെ (absolute measures of dispersion) കുറിച്ചാണ്. അതായത് പരിധി, ചതുരംശവ്യതിയാനം, വ്യതിയാന മാധ്യം, മാനക വ്യതിയാനം തുടങ്ങിയവ കേവല പ്രകീർണന മാനങ്ങളാണ്. കാരണം അവ ഒരു ഡാറ്റയിലുള്ള വിലകളുടെ വ്യതിയാനത്തെ അളക്കുകയും അവയുടെ യഥാർഥ യൂണിറ്റിൽ തന്നെ നൽകുകയും ചെയ്യുന്നു. വിവിധ ഡാറ്റകളിലെ യൂണിറ്റുകൾ വ്യത്യസ്തമാവുകയാണെങ്കിൽ ഡാറ്റകളെ തമ്മിൾ താരതമ്യം ചെയ്യുന്നതിന് കേവല പ്രകീർണനമാനങ്ങൾ ഒട്ടും അനുയോജ്യമാവുകയില്ല. മാത്രവുമല്ല അവ വ്യതിയാനത്തെക്കുറിച്ചുള്ള തെറ്റായ നിഗമനങ്ങളായിരിക്കും നമുക്ക് നൽകുന്നത്. ഇക്കാരണത്താൽ തന്നെ യൂണിറ്റുകൾക്കതീതമായ പ്രകീർണനമാനങ്ങൾ അത്യന്താപേക്ഷിതമാണ്. എങ്കിൽ മാത്രമേ വിവിധ യൂണിറ്റുകളിലായി ലഭ്യമാക്കിയിട്ടുള്ള ഡാറ്റകളുടെ വ്യതിയാനങ്ങളെ താരതമ്യ പഠനത്തിന് വിധേയമാക്കാൻ സാധിക്കുകയുള്ളൂ. അത്തരം പ്രകീർണനമാനങ്ങളെ ആപേക്ഷിക പ്രകീർണന മാനങ്ങൾ എന്നു പറയുന്നു. ഇത് പ്രകീർണനമാനത്തിന്റെയും അതിനുവേണ്ടി പരിഗണിച്ച ശരാശരിയുടെയും അംശബന്ധമായിരിക്കും.

ആപേക്ഷിക പ്രകീർണനമാനങ്ങളുടെ സവിശേഷതകൾ

- അത് ഒരു അംശബന്ധമായിരിക്കും
- അത് ഒരു സംഖ്യ മാത്രമായിരിക്കും
- അവ യൂണിറ്റുകൾക്കതീതമായിരിക്കും.
- രണ്ടോ അതിലധികമോ ഡാറ്റകളെ താരതമ്യം ചെയ്യാൻ വേണ്ടി ഉപയോഗിക്കുന്നു.

നമ്മൾ ഇവിടെ ചർച്ച ചെയ്യുന്ന പ്രധാനപ്പെട്ട രണ്ട് ആപേക്ഷിക പ്രകീർണനമാനങ്ങളാണ്,

1. വ്യതിയാന ഗുണാങ്കം (Coefficient of variation)
2. ചതുരംശ ഗുണാങ്കം (Coefficient of Quartile Deviation)

വ്യതിയാന ഗുണാങ്കം (Coefficient of variation)

രണ്ടോ അതിലധികമോ ഡാറ്റകളുടെ സ്ഥിരത അളക്കുന്നതിന് ഏറ്റവും കൂടുതലായി ഉപയോഗിക്കുന്ന ഗുണാങ്കമാണ് വ്യതിയാന ഗുണാങ്കം. മാനക വ്യതിചലന (σ) ത്തെ മാധ്യം $(\bar{x})$ കൊണ്ട് ഹരിച്ച് 100 കൊണ്ട് ഗുണിച്ചാൽ വ്യതിയാന ഗുണാങ്കം ലഭിക്കും.

$$\text{വ്യതിയാന ഗുണാങ്കം} \quad (CV) = \frac{\sigma}{\bar{x}} \times 100$$

രണ്ട് ഡാറ്റകളിൽ കൂടുതൽ സ്ഥിരത പുലർത്തുന്ന ഡാറ്റയെ കണ്ടെത്തണമെങ്കിൽ ആദ്യം നമ്മൾ രണ്ട് ഡാറ്റയുടേയും വ്യതിയാന ഗുണാങ്കങ്ങൾ (CV) കണ്ടുപിടിക്കണം. CV ഏറ്റവും കുറവുള്ള ഡാറ്റയെ നമുക്ക് ഏറ്റവും സ്ഥിരതയുള്ള ഡാറ്റയായി കണക്കാക്കാം.

വിവരണം 6.11

ഒരേ സ്ഥലത്ത് സ്ഥിതി ചെയ്യുന്ന A,B എന്നീ വ്യവസായശാലകളിലെ തൊഴിലാളികളുടെ ആഴ്ച വേതനങ്ങളുടെ ശരാശരികളും മാനക വ്യതിയാനങ്ങളും തന്നിരിക്കുന്നു

ഫാക്ടറി	ശരാശരി വേതനം ($\bar{x}$)	SD (σ)	തൊഴിലാളികളുടെ എണ്ണം
A	500	5	476
B	600	4	524

1. ഏതു വ്യവസായശാലയാണ് ആഴ്ച വേതനത്തിനായി കൂടുതൽ തുക ഉപയോഗിക്കുന്നത്?
2. ഏതു വ്യവസായശാലയിലാണ് വ്യക്തിഗത വേതനത്തിന്റെ സ്ഥിരത കൂടുതൽ ഉള്ളത്?

പരിഹാരം

ഇവിടെ $n_1 = 476$ $\quad n_2 = 524$

$\bar{x}_1 = 500$ $\quad \bar{x}_2 = 600$

$\sigma_1 = 5$ $\quad \sigma_2 = 4$

1. ഫാക്ടറി A യിലെ ആകെ ആഴ്ച വേതനം, $= 500 \times 476$

$= 238000$

2. ഫാക്ടറി B യിലെ ആകെ ആഴ്ച വേതനം, $= 600 \times 524$

$= 314400$

ഫാക്ടറി A യിലെ ആകെ ആഴ്ച വേതനം 238000 ഉം ഫാക്ടറി B യിലെ ആകെ ആഴ്ച വേതനം 314400 വും ആണ്. ഇതിൽ നിന്നും ഫാക്ടറി B ആണ് ആഴ്ച വേതനത്തിനായി കൂടുതൽ തുക ഉപയോഗിക്കുന്നതെന്ന് മനസ്സിലാക്കാം.

2) ഫാക്ടറി A യുടെ $CV = \frac{\sigma_1}{\bar{x}_1} \times 100$

$$= \frac{5}{500} \times 100 = 1$$

ഫാക്ടറി B യുടെ CV $= \frac{\sigma_2}{\bar{x}_2} \times 100$

$$= \frac{4}{600} \times 100 = 0.67$$

ഫാക്ടറി B യുടെ CV ഫാക്ടറി A യുടേതിനേക്കാൾ കുറവാണ്. അതിനാൽ ഫാക്ടറി B യാണ് വേതനത്തിന്റെ കാര്യത്തിൽ കൂടുതൽ സ്ഥിരത പുലർത്തുന്നത് എന്ന് പറയാൻ സാധിക്കും.

നിങ്ങളുടെ പുരോഗതി അറിയുക

കേരളത്തിലെയും തമിഴ്നാട്ടിലെയും രണ്ട് നഗരങ്ങളിൽ ഒരുൽപ്പന്നത്തിന്റെ വില തുടർച്ചയായ അഞ്ച് വർഷങ്ങളിൽ ചുവടെ കാണിച്ചിരിക്കുന്നത് പ്രകാരമാണ്.

തമിഴ്നാട് (രൂപ)	20	22	19	22	23
കേരളം (രൂപ)	18	12	10	20	15

ഏത് സംസ്ഥാനത്താണ് സ്ഥിരതയുള്ള വിലയുള്ളത്? നിങ്ങളുടെ ഉത്തരത്തെ സാധൂകരിക്കുക.

പ്രവർത്തനം

നിങ്ങളുടെ ക്ലാസിലെ കുട്ടികളെ രണ്ട് ഗ്രൂപ്പുകളാക്കുക. അവരുടെ സ്റ്റാറ്റിസ്റ്റിക്സ് പരീക്ഷയുടെ സ്കോറുകൾ ശേഖരിച്ച് കൂടുതൽ സ്ഥിരതയാർന്ന ഗ്രൂപ്പിനെ കണ്ടെത്തുക.

ചതുരംശ വ്യതിയാന ഗുണാങ്കം (Coefficient of Quartile Deviation)

ഒരു കേവല പ്രകീർണനമാനം ആണല്ലോ ചതുരംശ വ്യതിയാനം (QD). ചതുരംശ വ്യതിയാനത്തെ അടിസ്ഥാനപ്പെടുത്തി കണ്ടുപിടിക്കാവുന്ന ഒരു ആപേക്ഷിക പ്രകീർണനമാനമാണ് ചതുരംശ വ്യതിയാന ഗുണാങ്കം.

ചതുരംശ വ്യതിയാന ഗുണാങ്കം = $\frac{Q_3 - Q_1}{Q_3 + Q_1}$

6.6 സഹവ്യതിയാനം (Covariance)

രണ്ടു ചരങ്ങളുടെ രേഖീയ ബന്ധത്തിന്റെ ദൃഢതയുടെ അളവാണ് സഹവ്യതിയാനം ഒരു ദ്വിചര വിതരണത്തിലെ ചരങ്ങൾ തമ്മിലുള്ള ബന്ധം പോസിറ്റീവ് ആണോ നെഗറ്റീവ് ആണോ എന്നറിയാൻ സഹവ്യതിയാനം കണ്ടാൽ മതിയാകും.

$Cov(X,Y) = \frac{\Sigma(x-\bar{x})(y-\bar{y})}{n}$ ഇവിടെ n എന്നത് ഡാറ്റയിലെ വിലകളുടെ എണ്ണത്തെ സൂചിപ്പിക്കുന്നു.

കണക്ക് കൂട്ടാനുള്ള എളുപ്പത്തിന് വേണ്ടി നമുക്ക് മറ്റൊരു സൂത്രവാക്യം ഉപയോഗിക്കാം.

$$Cov(X,Y) = \frac{\Sigma xy}{n} - \bar{x} \times \bar{y}$$

Cov (X,Y) പോസിറ്റീവ് ആണെങ്കിൽ രണ്ട് ചരങ്ങളും ഒരേ ദിശയിലേക്ക് നീങ്ങുന്നതായി മനസ്സിലാക്കാം. Cov(X,Y) നെഗറ്റീവാകുന്നത് രണ്ടു ചരങ്ങൾ വിപരീതദിശകളിലേക്ക് ചലിക്കുമ്പോഴാണ്. രണ്ട് ചരങ്ങൾ തമ്മിൽ രേഖീയ ബന്ധമൊന്നുമില്ലെങ്കിൽ Cov(X,Y) പൂജ്യമായിരിക്കും.

വിവരണം 6.12

ഒരു തൊഴിൽശാലയിലെ എട്ട് തൊഴിലാളികളുടെ സേവന കാലവും വാർഷിക വരുമാനവും ചുവടെ കൊടുത്തിരിക്കുന്നു.

അവയുടെ സഹവ്യതിചലനം കണ്ടുപിടിക്കുക.

സേവന കാലം (x) (വർഷത്തിൽ)	6	8	9	10	11	13	15	16
വാർഷിക വരുമാനം (y) (ആയിരത്തിൽ)	14	17	15	18	16	22	25	25

പരിഹാരം

$$Cov(X,Y) = \frac{\Sigma xy}{n} - \bar{x} \times \bar{y}$$

x	y	xy
6	14	84
8	17	136
9	15	135
10	18	180
11	16	176
13	22	286
15	25	375
16	25	400
88	152	1772

$$\bar{x} = \frac{\Sigma x}{n} = \frac{88}{8} = 11$$

$$\bar{y} = \frac{\Sigma y}{n} = \frac{152}{8} = 19$$

$$Cov(x,y) = \frac{\Sigma xy}{n} - \bar{x} \times \bar{y}$$

$$= \frac{1772}{8} - 11 \times 19 = 12.5$$

നമുക്ക് സംഗ്രഹിക്കാം

ഈ അധ്യായത്തിൽ പ്രകീർണനം എന്ന ആശയവും കേവല പ്രകീർണന മാനകങ്ങളും ആപേക്ഷിക പ്രകീർണന മാനകങ്ങളും നമ്മൾ ചർച്ച ചെയ്തു. ഒരു കൂട്ടം വിലകളുടെ വ്യതിയാനത്തിന്റെ തോത് അളക്കാൻ കേവല പ്രകീർണനമാനങ്ങൾ ഉപയോഗിക്കുന്നു. ഇതിൽ പരിധി, ചതുരംശ വ്യതിയാനം എന്നിവ ഡാറ്റകൾക്ക് അവയിൽ നിന്ന് തന്നെയുണ്ടാകുന്ന വ്യതിയാനങ്ങൾ അളന്ന് തിട്ടപ്പെടുത്തുമ്പോൾ വ്യതിയാന മാധ്യം, മാനകവ്യതിയാനം എന്നിവ ഡാറ്റകൾക്ക് അവയുടെ ശരാശരിയിൽ നിന്നുള്ള വ്യതിയാനമാണ് അളന്നെടുക്കുന്നത്. ആപേക്ഷിക പ്രകീർണന മാനങ്ങൾ യൂണിറ്റുകൾക്ക് അതീതമാണ്. അവ ഉപയോഗിച്ച് രണ്ട് ഡാറ്റകളുടെ വ്യതിയാനങ്ങളെ താരതമ്യം ചെയ്യാൻ സാധിക്കും. രണ്ട് ചരങ്ങൾ തമ്മിലുള്ള ബന്ധത്തെ സഹവ്യതിചലനം ഉപയോഗിച്ച് മനസ്സിലാക്കാൻ സാധിക്കും.

പഠനനേട്ടങ്ങൾ

ഈ അധ്യായം പഠിക്കുന്നതിലൂടെ പഠിതാവ്:

- പ്രകീർണനമാനങ്ങളുടെ പ്രാധാന്യത്തെക്കുറിച്ച് തിരിച്ചറിയുന്നു.
- പരിധി, ചതുരംശവ്യതിയാനം, വ്യതിയാന മാധ്യം, മാനക വ്യതിയാനം എന്നീ പ്രകീർണനമാനങ്ങൾ എന്താണെന്ന് വിശദീകരിക്കാനും അവയുടെ വിലകൾ കണ്ടെത്താനും സാധിക്കുന്നു.
- കേവല പ്രകീർണന മാനങ്ങളും ആപേക്ഷിക പ്രകീർണന മാനങ്ങളും തമ്മിൽ വേർതിരിച്ചറിയുന്നു.

വിലയിരുത്തൽ ഇനങ്ങൾ

1. ഒരു ഡാറ്റയിലെ എല്ലാ വിലകളിൽ നിന്നും ഒരു നിശ്ചിത സംഖ്യ കുറച്ചാൽ വ്യതിചലനം ....

   (കൂടുന്നു, കുറയുന്നു, മാറുന്നില്ല)

2. ഒരു ഡാറ്റയിലെ മുഴുവൻ വിലകളെയും 10 കൊണ്ട് ഹരിച്ചാൽ മാനകവ്യതിയാനം ....

   (10 മടങ്ങാകുന്നു, പത്തിലൊന്നാകുന്നു, മാറ്റമില്ല)

3. ആദ്യത്തെ 25% വിലകൾ 20ഓ അതിൽ കുറവോ ആയതും അവസാനത്തെ 25% വിലകൾ 50 ഓ അതിൽ കൂടുതലോ ആയതുമായ ഒരു ഡാറ്റയുടെ ചതുരംശവ്യതിയാനം ....
4. ഒരു ഡാറ്റയുടെ ഏറ്റവും ചെറിയ വില 9 ഉം അതിന്റെ പരിധി 57 ഉം ആയാൽ ഡാറ്റയിലെ ഉയർന്ന വില എത്രയായിരിക്കും?
5. ഒരു ഡാറ്റയുടെ വ്യതിയാനഗുണാങ്കം 50 ഉം മാനകവ്യതിയാനം 20 ഉം ആയാൽ അതിന്റെ മാധ്യം....................
6. 5, 5, 5, 5, 5 എന്നീ വിലകളുടെ മാനക വ്യതിയാനം...................

7. മാധ്യത്തിൽ നിന്നുള്ള വ്യതിയാനങ്ങളുടെ വർഗങ്ങളുടെ മാധ്യം...................... ആണ്.

8. ചുവടെ കൊടുത്തിരിക്കുന്ന പ്രസ്താവനകൾ ശരിയോ തെറ്റോ എന്ന് പറയുക.

   (a) പരിധി ഏറ്റവും നല്ല പ്രകീർണനമാനമാണ്.

   (b) തുറന്ന ക്ലാസുകളുള്ള ആവൃത്തി വിതരണങ്ങളിൽ ചതുരംശവ്യതിയാനം ആണ് കൂടുതൽ അനുയോജ്യം.

   (c) താരതമ്യ പഠനത്തിന് കേവല പ്രകീർണനമാനങ്ങൾ ഉപയോഗിക്കാം.

   (d) മീഡിയനിൽ നിന്നുള്ള വ്യതിയാനങ്ങൾ പരിഗണിക്കുമ്പോഴാണ് വ്യതിയാന മാധ്യം ഏറ്റവും കുറവാകുന്നത്.

   (e) മാനക വ്യതിയാനം ഒരിക്കലും നെഗറ്റീവ് ആവുകയില്ല.

9. കേവല പ്രകീർണന മാനങ്ങളും ആപേക്ഷിക പ്രകീർണനമാനങ്ങളും തമ്മിലുള്ള വ്യത്യാസമെന്ത്?

10. താരതമ്യപഠനത്തിന് ആപേക്ഷിക പ്രകീർണന വ്യതിയാനമാണ് ഉപയോഗിക്കുന്നത് എന്തു കൊണ്ട്?

11. ഒരു കൂട്ടം വിലകളുടെ പ്രകീർണനം കൊണ്ട് ഉദ്ദേശിക്കുന്നത് എന്താണ്? എങ്ങിനെയാണ് പ്രകീർണനം അളക്കുന്നത്?

12. ഏതൊക്കെയാണ് വിവിധ കേവല പ്രകീർണനമാനങ്ങൾ?

13. എന്തുകൊണ്ടാണ് മാനകവ്യതിയാനത്തെ ഏറ്റവും മികച്ച പ്രകീർണനമാനമായി പരിഗണിക്കുന്നത്?

14. ഒരു ഡാറ്റയിലെ എല്ലാ വിലകളെയും 2 കൊണ്ട് ഗുണിച്ചാൽ ചുവടെ പറയുന്നവയ്ക്ക് എന്ത് സംഭവിക്കും.

   (a) മാധ്യം

   (b) വ്യതിചലനം

   (c) മാനകവ്യതിയാനം

15. ഒമ്പത് വിലകളുടെ തുക 360 ആണ്. അവയുടെ മാധ്യത്തിൽ നിന്നുള്ള വ്യതിയാനത്തിന്റെ വർഗങ്ങളുടെ തുക 288 ആണ്.

   (a) മാനകവ്യതിയാനം കാണുക.

   (b) വ്യതിയാന ഗുണാങ്കം കാണുക.

16. അശ്രദ്ധമായി വണ്ടിയോടിച്ചത് മൂലമുണ്ടായ അപകടങ്ങളുടെ എണ്ണങ്ങൾ ഒരാഴ്ചയിലെ ഏഴ് ദിവസങ്ങളിൽ തിട്ടപ്പെടുത്തിയിരിക്കുന്നു.

അപകടങ്ങളുടെ എണ്ണം	7	10	4	7	9	3	2

   മാധ്യത്തിൽ നിന്നുള്ള വ്യതിയാന മാധ്യം കണ്ടുപിടിക്കുക.

17. 52 വിദ്യാർഥികളിൽ നിന്ന് അവർ ദിവസവും ചെലവഴിക്കുന്ന തുകയുടെ വിവരങ്ങൾ ശേഖരിച്ചു. അവ ഇപ്രകാരമാണ്.

ചെലവഴിച്ച തുക	കുട്ടികളുടെ എണ്ണം
5	2
10	5
15	10
20	7
25	8
30	5
35	6
40	4
45	5

ഇവയുടെ ചതുരംശ വ്യതിയാനം കാണുക.

18. പുതിയതായി സൃഷ്ടിക്കപ്പെട്ട തസ്തികയിലേക്ക് അപേക്ഷകർക്കായി 5 ദിവസം മാനേജർ അഭിമുഖം നടത്തി. ഓരോ ദിവസവും ഉണ്ടായിരുന്ന ഉദ്യോഗാർഥികളുടെ എണ്ണം യഥാക്രമം 16, 19, 15, 15, 14 എന്നിങ്ങനെയാണ്. വ്യതിചലനം കാണുക.

19. ഒരു ഫാക്ടറി A,B എന്നീ രണ്ടുതരം വൈദ്യുത വിളക്കുകൾ നിർമിക്കുന്നു. അവയുടെ ആയുസ്സുമായി ബന്ധപ്പെട്ട ഒരു പരീക്ഷണത്തിൽ താഴെപ്പറയുന്ന വിവരങ്ങൾ ലഭ്യമായിരിക്കുന്നു.

ആയുർദൈർഘ്യം (മണിക്കൂറിൽ)	A വിളക്കുകളുടെ എണ്ണം	B വിളക്കുകളുടെ എണ്ണം
500 - 700	5	4
700 - 900	11	30
900 - 1100	26	12
1100 - 1300	10	8
1300 - 1500	8	6

a) ഏതു വിളക്കിനാണ് ശരാശരി ആയുസ് കൂടുതൽ?

b) ആയുർദൈർഘ്യത്തിൽ കൂടുതൽ സ്ഥിരതയുള്ളത് ഏതു വിളക്കിനാണ്?

20. ചില ജീവനക്കാർ അടച്ച വാർഷികനികുതി വിവരം താഴെ തന്നിരിക്കുന്നു.

നികുതി അടച്ചത് (ആയിരത്തിൽ)	5-10	10-15	15-20	20-25	25-30	30-35	35-40
ജീവനക്കാരുടെ എണ്ണം	18	30	46	28	20	12	6

അർധചതുരംശാന്തരപരിധി കാണുക

21. b mZrÝ rä anb n sXcsª Sq s¸ « 108 ഹയർസെക്കണ്ടറി വിദ്യാഥികളുടെ I.Q. വിതരണം ചുവടെ തന്നിരിക്കുന്നു.

I.Q.	90 - 98	99 - 107	108-116	117-125	126-134
വിദ്യാർഥികളുടെ എണ്ണം	6	10	25	15	4

I.Q. വിന്റെ മാനകവ്യതിയാനം കാണുക.

22. രണ്ടു നഗരങ്ങളിൽ കഴിഞ്ഞ അഞ്ച് വർഷമായുള്ള ഒരു പ്രത്യേക ഉൽപ്പന്നത്തിന്റെ വിലകൾ ചുവടെ തന്നിരിക്കുന്നു.

നഗരം A യിലെ വില	20	22	19	23	16
നഗരം B യിലെ വില	10	20	18	12	15

ചതുരംശ വ്യതിയാന ഗുണാങ്കങ്ങൾ ഉപയോഗിച്ച് ഏത് നഗരത്തിലാണ് വിലയിൽ സ്ഥിരതയുള്ളത് എന്നു കാണുക.

23. ഒരു ആശുപത്രിയിൽ റിപ്പോർട്ട് ചെയ്യപ്പെട്ട മോട്ടോർ സൈക്കിൾ അപകടങ്ങളുടെ എണ്ണം താഴെ തന്നിരിക്കുന്നു.

വയസ്സ്	0-10	10-20	20-30	30-40	40-50	50-60
അപകടങ്ങളുടെ എണ്ണം	15	49	37	20	6	1

മധ്യാങ്കം ആധാരമായ വ്യതിയാനമാധ്യം കാണുക.

24. 1000 ആളുകളുടെ വരുമാന വിതരണം ചുവടെ നൽകിയിരിക്കുന്നു.

വരുമാനം (കൂടുതൽ)	1000	900	800	700	600	500	400	300	200
ആളുകളുടെ എണ്ണം	0	50	110	200	400	650	825	950	1000

ചതുരംശ വ്യതിയാനം കാണുക

25. ഒരു കോളേജിലെ ബി.കോം. ക്ലാസിലെ 9 വിദ്യാർഥികൾക്ക് സ്റ്റാറ്റിസ്റ്റിക്സ് പരീക്ഷയിൽ ലഭ്യമായ മാർക്കുകൾ ചുവടെ തന്നിരിക്കുന്നു. മാനകവ്യതിയാനവും വ്യതിചലനവും കാണുക.

ക്രമനമ്പർ	1	2	3	4	5	6	7	8	9
സ്റ്റാറ്റിസ്റ്റിക്സിലെ മാർക്ക്	5	10	20	25	40	42	45	48	70

26. ഒരു ആശുപത്രിയിൽ ചികിത്സിച്ച ഡയബറ്റിക് രോഗികളുടെ എണ്ണവും അവരുടെ വയസ്സും താഴെ തന്നിരിക്കുന്നു.

വയസ്സ്	20	30	40	50	60	70	80
രോഗികളുടെ എണ്ണം	3	10	30	40	35	10	7

മധ്യാങ്കം അടിസ്ഥാനമായ വ്യതിയാന മാധ്യം കണ്ടുപിടിക്കുക.

27. ഒരു വീട്ടുപകരണ വിൽപ്പനശാലയിൽ ഒരാഴ്ച വിൽക്കപ്പെട്ട ടെലിവിഷൻ സെറ്റുകളുടെ എണ്ണം ചുവടെ തന്നിരിക്കുന്നു.

സെറ്റുകളുടെ എണ്ണം	5	6	7	8	4	3	1

മാനക വ്യതിയാനവും വ്യതിയാനവും കാണുക.

28. ഒരു ഗ്രാമത്തിലെ 75% വ്യക്തികളും 5,000 ത്തിന് മുകളിൽ സമ്പാദ്യമുള്ളവരാണ്. എന്നാൽ 25% വ്യക്തികൾക്ക് മാത്രമേ 10,000 ത്തിന് മുകളിൽ സമ്പാദ്യം ഉള്ളൂ. അനുയോജ്യമായ ആപേക്ഷിക വ്യതിയാനം , കേവല ആപേക്ഷിക വ്യതിയാനം എന്നിവ കണക്കാക്കുക.

29. ഒരു കമ്പനിയുടെ 10 വർഷക്കാലത്തെ പരസ്യച്ചെലവും (y) വിൽപ്പനയും (x) സംബന്ധിച്ച് ഒരു പട്ടിക ചുവടെ തന്നിരിക്കുന്നു. Cov (X,Y) കണക്കാക്കുക.

വിൽപ്പന (x)	50	5	50	40	30	20	20	15	10
പരസ്യച്ചെലവ് (y)	700	200	650	500	450	400	300	250	210

ഉത്തരങ്ങൾ

1) മാറ്റമില്ല
2) പത്തിലൊന്നാകും
3) 15
4) 66
5) 40
6) 0
7) മാനകവ്യതിയാനം
8) a) തെറ്റ് b) ശരി c) തെറ്റ് d) ശരി e) ശരി
14) a) 2 തവണ b) 4 തവണ c) 2 തവണ
15) $\sigma = 5.66$, CV= 6.28
16) മാധ്യം അടിസ്ഥാനമായ M.D = 2.57
17) Q.D = 10
18) $\sigma^2 = 2.96$
19) a) ശരാശരി ആയുസ് കൂടുതൽ ബൾബ് A യ്ക്ക്
    b) ബൾബ് A യായിരിക്കും കൂടുതൽ സ്ഥിരതയുള്ളത്
20) അർധ ചതുരംശാന്തര പരിധി = 5.48
21) $\sigma = 9.36$

22) A യുടെ ചതുരംശവ്യതിയാന ഗുണാങ്കം = 0.125, B യുടെ ചതുരംശവ്യതിയാന ഗുണാങ്കം = 0.267 അതുകൊണ്ട് നഗരം A യിലെ വില കൂടുതൽ സറിരതയുള്ളതാണ്.

23) മധ്യാങ്കം ആധാരമായ വ്യതിയാനമാധ്യം= 8.906

24) ചതുരംശ വ്യതിയാനം = 137.5

25) SD = 19.46, വേരിയൻസ് = 378.69

26) മധ്യാങ്കം ആധാരമായ വ്യതിയാനമാധ്യം = 10

27) SD = 2.231, വ്യതിയാനം = 4.979

28) ചതുരംശ വ്യതിയാനം = 2500, ചതുരംശവ്യതിയാന ഗുണാങ്കം = 0.33

29) Cov (X,Y) = 2694.44