த. பா. நி. (க.வெ.) வரிசை எண்—419

மாதிரி முறைகள்

(பட்டப்படிப்பிற்குரியது)

(திருத்தப்பட்ட பாடத்திட்டத்தின்கீழ் வெளியிடப்படுகிறது)

ஆசிரியர்

கி. பாலசுப்பிரமணியன், எம்.ஏ., எம்.ஸ்டேடிஸ்டிக்ஸ்,

பேராசிரியர், புள்ளியியல் துறை,

இலயோலா கல்லூரி

சென்னை-34.

தமிழ்நாட்டுப் பாடநூல் நிறுவனம்

First Edition — February, 1973

T.N.T.B.S. (C.P.) No. 419

© Tamil Nadu Text Book Society

SAMPLING TECHNIQUES

K. BALASUBRAMANIAN

Price Rs. 7-30

'Published by The Tamil Nadu Text Book Society under the Centrally Sponsored Scheme of Production of books and literature in regional languages at the University level, of the Government of India in the Ministry of Education and Social Welfare (Department of Culture), New Delhi.'

Printed by
Muthukumaran Press,
Madras-1.

அணிந்துரை

திரு. இரா. நெடுஞ்செழியன்
(தமிழகக் கல்வி-உள்ளாட்சித்துறை அமைச்சர்)

தமிழைக் கல்லூரிக் கல்வி மொழியாக ஆக்கிப் பன்னிரண்டாண்டுகள் ஆகிவிட்டன. குறிப்பிட்ட சில கல்லூரிகளில் பி.ஏ. வகுப்பு மாணவர்கள் தங்கள் பாடங்கள் அனைத்தையும் தமிழிலேயே கற்றுவந்தனர். 1968ஆம் ஆண்டின் தொடக்கத்தில் புகுமுக வகுப்பிலும் (P.U.C.), 1969ஆம் ஆண்டிலிருந்து பட்டப் படிப்பு வகுப்புகளிலும் அறிவியல் பாடங்களையும் தமிழிலேயே கற்பிக்க ஏற்பாடு செய்துள்ளோம். தமிழிலேயே கற்பிப்போம் என முன்வந்துள்ள கல்லூரி ஆசிரியர்களின் ஊக்கம், பிற பல துறைகளிலும் தொண்டு செய்வோர் இதற்கெனத் தந்த உழைப்பு, தங்கள் சிறப்புத் துறைகளில் நூல்கள் எழுதித் தர முன்வந்த நூலாசிரியர்கள் தொண்டுணர்ச்சி இவற்றின் காரணமாக இத் திட்டம் நம்மிடையே மகிழ்ச்சியும் மனநிறைவும் தரத்தக்க வகையில் நடைபெற்று வருகிறது. இவ்வகையில் கல்லூரிப் பேராசிரியர்கள் கலை, அறிவியல் பாடங்களை மாணவர்க்குத் தமிழிலேயே பயிற்றுவிப்பதற்குத் தேவையான பயிற்சியைப் பெறுவதற்கு மதுரைப் பல்கலைக்கழகம் ஆண்டுதோறும் எடுத்துவரும் பெரு முயற்சியைக் குறிப்பிட்டுச் சொல்லவேண்டும்.

பல துறைகளில் பணிபுரியும் பேராசிரியர்கள் எத்தனையோ நெருக்கடிகளுக்கிடையே குறுகிய காலத்தில் அரிய முறையில் நூல்கள் எழுதித் தந்துள்ளனர்.

வரலாறு, அரசியல், உளவியல், பொருளாதாரம், தத்துவம், புவியியல், புவியமைப்பியல், மனையியல், கணிதம், இயற்பியல், வேதியியல், உயிரியல், வானியல், புள்ளியியல், விலங்கியல், தாவரவியல், பொறியியல் ஆகிய எல்லாத் துறைகளிலும் தனி நூல்கள், மொழிபெயர்ப்பு நூல்கள் என்ற இருவகையிலும் தமிழ் நாட்டுப் பாடநூல் நிறுவனம் வெளியிட்டு வருகிறது.

இவற்றுள் ஒன்றான '**மாதிரி முறைகள்**' என்ற இந் நூல் தமிழ்நாட்டுப் பாடநூல் நிறுவனத்தின் **419**ஆவது வெளியீடாகும். கல்லூரித் தமிழ்க் குழுவின் சார்பில் வெளியான 35 நூல்களையும் சேர்த்து இதுவரை **454** நூல்கள் வெளிவந்துள்ளன. இந் நூல் மைய அரசு கல்வி, சமூக நல அமைச்சகத்தின் மாநில மொழியில் பல்கலைக்கழக நூல்கள் வெளியிடும் திட்டத்தின்கீழ் வெளியிடப்படுகிறது.

உழைப்பின் வாரா உறுதிகள் இல்லை ; ஆதலின், உழைத்து வெற்றி காண்போம். தமிழைப் பயிலும் மாணவர்கள் உலக மாணவர்களிடையே சிறந்த இடம் பெறவேண்டும். அதுவே தமிழன்னையின் குறிக்கோளுமாகும். தமிழ்நாட்டுப் பல்கலைக் கழகங்களின் பல்வகை உதவிகளுக்கும் ஒத்துழைப்புக்கும் நம் மனம் கலந்த நன்றி உரியதாகுக.

இரா. நெடுஞ்செழியன்

முன்னுரை

புள்ளியியலிலே நடைமுறையில் மிகவும் அதிக அளவில் கையாளப்படுவது மாதிரிமுறைதான் என்று கூறினால், அது மிகை யாகாது. புள்ளி விவரங்கள் என்று கூறும்போது எல்லோரும் அறிந்துகொள்ளத்தக்க விதத்தில் அமைக்கப்பட்ட விவரங்களைத் தாம் நாம் மனத்தில் கொள்கிறோம். இவ் விவரங்களை எப்படிச் சேகரிப்பது? எவ்விதத்தில் சேகரிக்கும்போது பிழை குறைவா யிருக்கும்? சேகரிக்கப்பட்ட விவரங்கள் எவ்வளவு தூரம் நம்பத் தகுந்தவை? ஒன்றுக்கு மேற்பட்ட மதிப்பீட்டு முறைகள் இருப்பின், அதில் எது மிகவும் சிறந்தது ? (திறன் வாய்ந்தது ?) மேற்கூறப் பட்ட கேள்விகளை நாம் நன்கு ஆராய்ந்து பார்த்தால், அவைகளி லுள்ள குறைபாடுகள் நமக்குப் புலப்படும்

(அ) சேகரிக்கும் விதத்திற்கும் சேகரிக்கப்படும் விவரங் களுக்கும் தொடர்பு ஏதாவது இருந்தாகவேண்டுமா, என்ன ? எவ்விதத்தில் விவரங்கள் சேகரிக்கப்பட்டன என்பதைக் குறிப் பிடாமல் விவரங்களைமட்டும் கொடுத்தால் போதாதா ?

(ஆ) விவரங்களிலுள்ள பிழைகள் என்று குறிப்பிட்டோம். எது பிழையென்று கூறமுடியுமா ? பிழையை அளக்கமுடியுமா? அவ்வளவெதற்கு ? பிழையென்றால் என்ன என்பதைக் கணித முறையில் ஐயத்திற்கிடமின்றி வரையறை செய்யமுடியுமா ?

(இ) விவரங்களை எவ்வளவு தூரம் நம்பமுடியும் என்று கேள்வி கேட்டோம். விவரங்களை நம்புவது என்றால் என்ன ? எவ்வளவு தூரம் என்று குறிப்பிடும்போது நம்முடைய நம்பிக் கையை அளப்பதற்கு ஏதோ அளவுகோல் இருப்பதுபோல நாம் உணர்வதில்லையா? நம்பிக்கையை எவ்வாறு அளப்பது ?

(ஈ) விவரங்களைச் சேகரிக்கப் பல வழிகள் இருக்கும்போது அவைகளில் எது சிறந்தது என்று எவ்வாறு கூறமுடியும்? இக் கேள்வியிலேயே, ஒவ்வொரு வழிமுறைக்கும் திறனென்று ஒன்று இருக்கும். இத் திறன்களை ஒப்பிடுவதன்மூலம் எம்முறை சிறந்தது என்பதைக் கண்டறியலாம் என்பது தொக்கிநிற்கவில்லையா? திறனென்பது என்ன? அதை எவ்வாறு அளப்பது ?

மேற்கூறப்பட்ட கேள்விகளுக்குத் தகுந்த முறையில் விடையிறுக்கவேண்டுமாயின் 'மாதிரி முறை' யென்னும் புள்ளியியலின் மிக முக்கியப் பகுதியை நாம் அமைக்க வேண்டி வரும்.

'ஒரு பானைச் சோற்றுக்கு ஒரு சோறு பதம்', என்பது பன்னெடுங்காலமாக வழங்கிவரும் பழமொழியாகும். இதிலிருந்து பண்டைக் காலந்தொட்டே மாதிரியை வைத்துக்கொண்டு முழுமைத்தொகுதியை ஆராய்ந்தறிய முடியும் என்னும் தொகுத்தறி முறையை மக்கள் உணர்ந்திருந்தனர் என்பது தெரிகிறது. ஆனால், இம் முறையை விஞ்ஞான பூர்வமாக எல்லோரும் ஒப்புக் கொள்ளத்தக்கவிதத்தில் வரையறுப்பதற்கு 'நிகழ்திறன்' (Probability) என்னும் இயல் உருவாக்கப்படவேண்டியிருந்தது. மாதிரிமுறை இருபதாம் நூற்றாண்டில்தான் அதிகமாகச் சீராக்கப் பட்டது. நிகழ்திறன் அல்லது ஊக அளவை என்பதும் இருபதாம் நூற்றாண்டில்தான் கணித வல்லுநர்கள் ஏற்கத்தக்க அளவில் உருவாக்கப்பட முடிந்தது.

நாம் ஒன்றை நினைத்து மிகவும் பெருமிதம் அடையலாம். அது புள்ளியியலில், குறிப்பாக மாதிரிமுறை இயலில், நம் இந்தியப் புள்ளியியலாளர்கள் மிகவும் வியக்கத்தக்க அளவில் ஆராய்ச்சிகள் செய்திருக்கின்றனர் என்பதே. உலகத்திலுள்ள தலைசிறந்த புள்ளியியலாளர்கள் பட்டியலில் நம் இந்தியர்கள் பெரும் பங்கு வகிக்கின்றனர்.

இந் நூல் என்னுடைய கன்னி முயற்சி. ஆகவே, ஆரம்ப வேகத்தில் சற்று மிகையாக எழுதியிருப்பேனோ என்று ஐயம் எனக்கு ஏற்படுகிறது. ஆனால், நான் இந் நூலில் எழுதியிருப்ப தெல்லாம் இதனுடன் ஒப்பிடத்தக்க பட்ட வகுப்பு மாணவர் களுக்கென்று எழுதப்பட்டிருக்கும் நூல்களிலுள்ளவையே. ஆனால் ஆரம்பத்தில், நிகழ்திறன் என்பதைத் தற்கால கணித இயலில் காணப்படும் கட்டுக்கோப்புக்கேற்றவாறு வரையறை செய்திருப்பது மாணவர்களுக்கு உபயோகமாக இருக்கும் என்று நம்புகிறேன். இது ஒன்றுதான் இந் நூலை மற்ற ஒப்பிடத்தக்க 'மாதிரி முறையியல்' நூல்களினின்று தனித்துக் காட்டலாம். ஆனால், இப் பகுதி மிகையென்று நினைப்பவர்கள் நேராக 13ஆம் பகுதியிலிருந்து தொடங்கலாம். இதனால் அவர்கள் அதிகம் இழக்கமாட்டார்களென்றே நான் நம்புகிறேன். முடிந்தவரை ஒவ்வொரு பகுதியிலும் பயிற்சிக்கென்று பல கணக்குகள் தந்திருக் கிறேன்; நிறையக் கணக்குகளை நானே விளக்கியிருக்கிறேன்.

எங்கெங்குப் புதிதான கணித முறையொன்று கையாளப்படுகிறதோ அங்கங்கு அதைப்பற்றிய சுருக்கமான விளக்கமும் தந்திருக்கிறேன். பலகட்ட மாதிரி முறை, விகிதமுறை மதிப்பீடு, தொடர்புமுறை மதிப்பீடு முதலியவைகளைப்பற்றி நான் அதிகம் விளக்கமாக எழுதவில்லையென்பது உண்மை. இதற்குக் காரணங்கள் என்று பல கூறமுடியும். முதலாவதாக, இவை மிகவும் அதிக அளவில் புழக்கத்திலில்லை. இரண்டாவதாக, இவைகளைப் புரிந்துகொள்வது சற்றுக் கடினம். கடைசியாக, என்னுடைய நூல் மாணவர்கள் புரிந்துகொள்ளத்தக்கவிதத்தில் எளியதாக இருக்கவேண்டுமென்ற என்னுடைய அவா. ஆனால், ஒன்றுமட்டும் உறுதியாக என்னால் கூற முடியும். இந்த நூலைப் படித்து முற்றிலும் புரிந்துகொண்ட மாணவன் தானாகவே இக் கடினமான பகுதிகளை மற்றத் தரமான நூல்களிலிருந்து படித்தறியலாம். இந் நூலைப் படிப்பதன்மூலம் மாதிரி இயலைப்பற்றி அதிகம் தெரிந்துகொள்ள மாணவர்கள் விழைவார்களேயானால், நான் இந் நூலை எழுதிய பயனை அடைந்தவனாவேன்.

நான் இந் நூலில் அளித்தவையெல்லாம் மாதிரி இயலின் கடைசி வார்த்தைகளல்ல. மாதிரி முறைகள் நாளுக்குநாள் அதிகரித்துக்கொண்டு வருகின்றன. ஆனால், இந் நூலில் மாதிரி முறைகளை மாணவர்களுக்கு அறிமுகம் மட்டும் செய்ய நான் எண்ணியிருப்பதால், இவ் வியலில் அறியத்தக்கவை இவ்வளவுதானா என்று மாணவர்கள் ஏமாற்றம் அடையக்கூடாது. அறிமுகத்தின் பின் அவர்களே முயன்று மாதிரி முறைகளைப்பற்றித் தரமான நூல்களிலிருந்து தெரிந்துகொள்ளவேண்டும். அதன் பொருட்டே, மாதிரி முறை நூல்களின் பட்டியலொன்று பின்னே தரப்பட்டிருக்கிறது.

ஒரு வார்த்தை. மாதிரி முறைகளைப்பற்றி நன்கு அறிய வேண்டுமாயின், புள்ளியியலில் பெரும் பங்கு வகிக்கும் மதிப்பீடு, செய்முறைத் திட்டம், எடுகோள்களைச் சோதித்தல் முதலிய தொடர்பு கொண்டவைகளையும் அறிந்துகொள்ளவேண்டும். ஏனெனில், இவைகளைத் தனித்தனியாகப் பிரிப்பது என்பது முடியாத காரியம். ஆகவே, இந் நூலைப் படிக்கும்போது மேற்கூறிய புள்ளியியல் பகுதிகளை நினைவுகூர்ந்தால், இவைகளிலுள்ள அடிப்படை ஒற்றுமை தெரியவரும். இவைகளுக்குள்ள ஒற்றுமையைச் சரிவரத் தெரிந்துகொண்டால், புள்ளியியலைப் பற்றி ஓரளவு தெரிந்துகொண்டிருக்கிறோம் என்று பெருமிதம் அடையலாம். இவை ஒன்றுக்கொன்று சம்பந்தப்படாமல் தனித்தனியே இயங்குகின்றன என்ற எண்ணம் மாணவர்களுக்கு ஏற்படுமாயின், என்னுடைய

தாழ்மையான வேண்டுகோள், மாணவர்கள் மேலும் நன்கு இவைகளைப் படிக்க வேண்டும் என்பதே.

இந் நூலில் வெவ்வேறு மாதிரி முறைகளைப்பற்றிய கணிதப் பகுதிகளை மட்டும் தந்திருக்கிறேன். ஆகவே, இந் நூல் நடைமுறையில் இம் முறைகளைக் கையாளப் போதுமானதாக இராது. ஏனெனில், நடைமுறையில் புள்ளியியல், கணிதம் முதலியவைகளைத் தவிர மனோதத்துவம், சமூகவியல், பொருளாதாரம் முதலியவைகளும் உபயோகப்படுகின்றன. என்னுடைய இந்தச் சிறிய நூலில் இவைகளைப்பற்றி ஒன்றும் தெரிவிக்கமுடியாது; அதற்கான முயற்சியையும் நான் செய்யவில்லை. ஆகவே, நடைமுறையில் இந் நூல் அதிகமான உதவியைஅளிக்கவில்லையானால், அதற்குக் காரணம் இந் நூல் 'மாதிரி முறைகள்' என்று பெயரிடப்பட்டிருக்கிறதே தவிர 'அளவெடுப்பு முறைகளும் அவைகளின் திட்டங்களும்,' என்றன்று.

கடைசியாக ஒரு வார்த்தை. இந் நூல் மாணவர்களுக்குப்புரியத் தக்க விதத்தில் நன்கு அமைந்திருக்கிறது என்று மாணவர்கள் கருதுவார்களேயானால், அப் புகழெல்லாம் நான் இந் நூலை எழுதுவதற்காகப் படித்த நூல்களின் ஆசிரியர்களையே சாரும். ஆனால், இந்நூலில் பிழைகளும், கடினமான விளக்கங்களும் மலிந்திருக்கின்றன என்று மாணவர்கள் எண்ணினால் அக் குற்றம் முழுவதும் என்னையே சாரும்.

கி. பாலசுப்பிரமணியன்.

பொருளடக்கம்

			பக்கம்
1.	அளவெடுப்பைத் திட்டமிடுதல்	...	1
2.	அளவெடுப்பின் படிகள்	...	5
3.	அளவெடுப்பில் நேரக்கூடிய பிழைகள்	...	11
4.	அளவெடுப்பில் பிழைகள் (தொடர்ச்சி)	...	15
5.	நிகழ்திறத்தின் அடிப்படை	...	19
6.	சார்புகள்	...	34
7.	நிகழ்திறம்	...	39
8.	சார்பிலா நிகழ்ச்சிகள்	...	45
9.	ராண்டம் மாறிகள்	...	46
10.	எளிய சார்புகளைத் தொகையிடல்	...	48
11.	ஊக்குவிக்கப்பட்ட நிகழ்திற வெளி	...	51
12.	சில பரவல்கள்	...	56
13.	மாதிரி, மாதிரி அளவை, முழுமைத் தொகுதியளவை	...	60
14.	நம்பிக்கை இடைவெளி	...	69
15.	மாதிரிகள்	...	70
16.	சாரதாரண ராண்டம் மாதிரிமுறை	...	73
	பயிற்சி 1	...	85
17.	ஈடு செய்யப்பட்ட சாதாரண ராண்டம் மாதிரிகளில் மதிப்பீடு	...	94
	பயிற்சி 2	...	105

			பக்கம்
18.	ஈடு செய்யப்படாத சாதாரண ராண்டம் மாதிரிகளில் மதிப்பீடுகள்	...	116
	பயிற்சி 3	...	125
19.	முறையுடை மாதிரி முறை	...	138
	பயிற்சி 4	...	156
20.	சமவாய்ப்பிலா மாதிரி முறை	...	163
	பயிற்சி 5	...	181
21.	படுகை மாதிரி முறை	...	184
	பயிற்சி 6	...	213
22.	திரள் மாதிரி முறை	...	229
	பயிற்சி 7	...	243
23.	விகித மதிப்பீடு	...	245
	பயிற்சி 8	...	253
24.	பலகட்ட மாதிரி முறை	...	264
	பயிற்சி 9	...	275
25.	தொடர்பு மதிப்பீடும் வேறுபாடு மதிப்பீடும்	...	279
	பயிற்சி 10	...	287
26.	சாதாரண ராண்டம் மாதிரி (தொடர்ச்சி)	...	290
	பயிற்சி 11	...	297
	பயிற்சி 12 பலவிதக் கணக்குகள்	...	307
27.	ராண்டம் எண்கள் அட்டவணை	...	339
28.	நூற்பட்டியல்	...	341
29.	கலைச்சொற்கள்	...	361

மாதிரி முறைகள்

1. அளவெடுப்பைத் திட்டமிடுதல்

(PLANNING A SURVEY)

1·1. அளவெடுப்பு என்பது நமக்குத் தேவையான விவரங்கள், செய்திகள், தகவல்கள் முதலியவற்றைச் சேகரிப்பதற்கான முறையாகும். இவ்வளவெடுப்பு என்பது பல விதங்களில் செயல்படுத்தப்படலாம். தகவல்களின் தன்மைக்கேற்றவாறே அளவெடுப்பும் அமையும். தகவல்கள் என்று குறிப்பிடும்போது செய்தித்தாள்களில் வெளியாகும் தகவல்கள் என்று மட்டும் இங்கு நாம் பொருள் கொள்ளலாகாது. அன்றுட அரசாங்கக் காரியங்களுக்காவும், ஆராய்ச்சிக் கூடங்களின் ஆராய்ச்சிகளுக்காகவும் அல்லது தனி நபர்களின் பொருளாதார மேம்பாட்டுக்காகவும் பல்வேறு வகையான விவரங்கள் தேவைப்படுகின்றன. இவ்விவரங்களைப் பொதுவாக இருவகையாகப் பிரிக்கலாம். அவையாவன:

(அ) பண்பின் விவரங்கள் ; (ஆ) அளக்கத்தகு விவரங்கள்.

1·2. பண்பின் விவரங்கள்

இவ்விவரங்கள் பொதுவாக எவ்வித அளவு கோல்களாலும், எண்களின் மூலமாகவும், அளக்க முடியாதனவாகும். உதாரணமாக ஒரு குறிப்பிட்ட மக்களினத்தின் முற்போக்குக் கொள்கைகள், விலங்குகளின் போக்கு, அரசாங்கத்தைப் பற்றிய மக்களின் கருத்துகள் போன்றவைகளைக் குறிப்பிடலாம். இவற்றைப் பற்றிய புள்ளி விவரங்கள் சேகரிப்பதில் புள்ளி விவர ஆய்வாளர்கள் மிகவும் கவனமாக இருத்தல் வேண்டும். ஏனெனில், புள்ளி விவரங்கள் ஆய்வாளர்களின் கொள்கைகளினுலும், முன்கூட்டியே மனதில் பதிந்த உறுதியான நம்பிக்கைகளினுலும் திரிக்கப்பட வாய்ப்புகள் உண்டு. நடுநிலை பிறழாது, தம் கடன் புள்ளி விவரங்கள் சேகரிப்பது ஒன்றே என்ற உறுதியான கட்டுப்பாட்டின் மூலமாகத்தான் குற்றங் குறையற்ற விவரங்களை ஆய்வாளர் தர முடியும்.

1·3. அளக்கத்தகு விவரங்கள்

இவ்விவரங்களை ஒரு குறிப்பிட்ட அலகு அல்லது அளவு கோல் மூலம் ஆய்வாளரால் அளக்க முடியும். உதாரணமாக, ஒரு வகுப்பிலுள்ள மாணவர்களின் உயரங்கள், ஒரு நாட்டின் ஏற்றுமதி இறக்குமதி அளவுகள், ஒரு மாநிலத்திலுள்ள கனி வளங்களின் மதிப்புகள் முதலியவற்றைக் குறிப்பிடலாம். இத்தகைய புள்ளி விவரங்கள் சேகரிப்பதில் ஆய்வாளர்கள் தவறுகள் செய்வதற்கு வாய்ப்புகள் குறைவு. பண்பின் விவரங்களைக் கூட ஒரு சில மாற்றங்களின் மூலமாக அளக்கத்தகு விவரங்களாக மாற்ற முடியும். உதாரணமாக, மக்களின் முற்போக்குக் கொள்கைகள் என்று குறிப்பிடப்படும்போது பண்பின் விவரங்களாகத் தோன்றுபவையே ஒரு குறிப்பிட்ட முற்போக்குக் கொள்கைகளைக் கொண்டவர் எத்தனைப்பேர் என்று எண்ணும் போது அளக்கத்தகு விவரமாகத் தோன்றும்.

1·4. அளவெடுப்பின் நோக்கங்கள்

அளவெடுப்புகளின் நோக்கங்கள் என்று கூறும்போது எதற்காக அளவெடுப்பு மேற்கொள்ளப்படுகிறது என்பதைப் பற்றிய ஆராய்ச்சியை நாம் இங்கு மனதில் கொள்ளவில்லை. ஒரு பண்பினையோ அல்லது அளவையையோ புள்ளி விவரங்களுக்குள் எப்படி அடக்குவது, நடைமுறையில் எவ்வாறு இதைச் செயலாற்றுவது என்பவையே நாம் மனதில் கொள்ளத்தக்கவையாகும்.

அளவிடுவது என்று வரும்பொழுது பிழைகளை முற்றிலும் தவிர்த்து விட முடியாது. ஆனால் எவ்வளவு பிழைபட்ட புள்ளி விவரங்களை நாம் பொறுத்து ஏற்றுக் கொள்ள முடியும், குறிப்பிட்ட பிழையளவைத் தாண்டாத புள்ளி விவரங்களை எவ்வகையில் குறைந்த செலவில் பெறமுடியும் என்பவற்றையே நாம் யோசிக்க வேண்டும். நம் அளவெடுப்பு முறைகளும் இவைகளையொட்டியே இருக்கும். ஓர் அளவெடுப்பு முறையைத் தேர்ந்தெடுத்த பின்னர் அதைச் சரிவர நடைமுறையில் எவ்வாறு செயலாற்றுவது; அளவெடுப்பின் ஒவ்வொரு கட்டத்திலும் எவ்வளவு பணம் செலவழிக்கப்படுகிறது; சேகரிக்கப்பட்ட விவரங்கள் உண்மையிலேயே நாம் எதிர்பார்க்கும் பிழையளவிற்குட் பட்டிருக்கின்றனவா என்பன போன்றவற்றை மிகுந்த கவனத்துடன் ஆராய்தல் வேண்டும்.

1·5. பொருளாதாரக் கட்டுப்பாடு

புள்ளியியல் ஆய்வாளரின் முக்கியமான நோக்கமே 'திறம் படச்' செயலாற்றுவதுதான். 'திறம்பட' என்று கூறும் பொழுது

பின்வரும் இரு முரண்பட்ட பிழைக் கூறுகளுக்கிடையே எப்பக்கத்திலும் அதிகம் ஒதுங்காது சீரிய துலாக்கோல் போல நடந்து கொள்வதையே நாம் இங்குக் குறிப்பிடுகிறோம்.

(அ) நாம் பொறுத்துக் கொள்ளக்கூடிய பிழையளவை விடக் குறைவான பிழை இருக்குமாறு அளவெடுப்பு முறையை அமைத்து அதனுல் அதிக நேரத்தையும் பொருளையும் விரயம் செய்வது. உதாரணமாக, விறகின் எடையைக் கண்டுபிடிப்பதற்குப் பொற்கொல்லர்கள் உபயோகிக்கும் துலாக்கோலை ஒருவன் உபயோகப் படுத்தினுல் அவனுடைய மதிநுட்பத்தை எவரும் பாராட்ட மாட்டார்கள்; மாருக அவனைக் கண்டு எள்ளி நகையாடுவர். இதற்குக் காரணம் யாதெனில் விறகின் எடை குன்றிமணி எடை கூட மேலோ, கீழோ போகாமல் தெரிவதனுல் நமக்கு யாதொரு பயனுமில்லை. அதாவது எடையில் குறிப்பிடும்படியான வேறு பாட்டைக்கூட நம்மால் பொறுத்துக்கொள்ள முடியும். ஆகவே தோராயமாகவும், சுலபமாகவும், குறைந்த கால அளவிலும் செய்து முடிக்கக்கூடிய முறைகளையே சில சமயங்களில் பின் பற்றுவது நல்லது.

(ஆ) நாம் பொறுத்துக் கொள்ளக்கூடிய பிழையளவைவிட அதிகமான பிழை உள்ளவாறு அளவெடுப்பு முறையை அமைத்து அதனுல் தரக்குறைவான, உபயோகத்திற்கு அருகதையற்ற விவரங்களைச் சேகரிப்பது. உதாரணமாக, ஒரு தங்க நகையை எடை போடுவதற்கு விறகுக் கடையில் உபயோகப்படுத்தப்படும் தராசை ஒருவன் பயன்படுத்தினுல் எப்படி பிழைபட்ட எடை கிடைக்கும் என்பதை எடுத்துக் கொள்ளலாம். தங்கத்தின் மதிப்பு மிகவும் அதிகம் என்பதால், இங்கு எடை குன்றிமணியளவுகூட மாருமல் நமக்குத் தெரிவது அவசியமாகிறது,

இவ்விரண்டு பிழைகளையும் மிகவும் குறைவாகச் செய்பவரையே சிறந்த புள்ளியியல்ஆய்வாளர் என நாம் கூறவேண்டும். இதை நாம் ஒரு வங்கிக்கு ஒப்பிடலாம். வங்கியின் செயலாளர் வங்கியில் கைவசமாக தினமும் எவ்வளவு பணம் வைத்துக் கொள்ள வேண்டும் என்பதை மிகவும் கவனத்துடன் தீர்மானிக்க வேண்டும். தேவைக்கு அதிகமாகக் கையிருப்பில் பணம் வைத்திருந்தால் வங்கியின் முக்கியக் கடமையான கடன் கொடுத்தல் மிகவும் பாதிக்கப்பட்டு அதனுல் வங்கிக்கு வரும் இலாபமும் குறைந்துவிடும். தேவைக்குக் குறைவாக கையிருப்பில் வைத்திருந்தால் அன்ருடத் தேவைகளுக்கு வங்கியில் கணக்கு வைத்துக் கொண்டிருப்பவருக்குப் பணம் கொடுக்க முடியாத நிலைமையுண்டாகி வங்கியே மூடப்படவேண்டி வரலாம். ஆகவே இவ்விரு பிழைகளும் அதிகம் நேராவண்ணம் பார்த்துக் கொள்ளும் வங்கிச்

செயலாளரையே நாம் சிறந்தவர் எனக் கூறமுடியும். இதைப் போலவே புள்ளியியல் ஆய்வாளரும் மேற் குறிப்பிட்ட இருவித பிழைகளும் நேராவண்ணம் பார்த்துக்கொள்ளக் கடமைப் பட்டவராவர்.

1·6. புள்ளியியல் பண்புகளுடன்கூறெடுக்கும் முறைகளின் தொடர்பு

கூறெடுக்கும் முறைகளைப் பற்றிய ஆராய்ச்சியென்பது புள்ளியியல் கலையின் ஒரு முக்கியமான பிரிவாகும். பரிசோதனை முறையமைப்பு (Experimental Design) கூறெடுக்கும் முறையை யொட்டியே அமையும். கூறெடுக்கும் முறைகளைப் பற்றிச் சிந்திக்கும் பொழுது பின்வருவனவற்றைக் கவனத்தில் கொள்ளல் வேண்டும்.

(அ) சேகரிக்கப்படும் தகவல்கள் முழுமையாகவும், உபயோகத்திற்கு உகந்தனவாகவும் இருத்தல் வேண்டும்.

(ஆ) எந்த நோக்கத்திற்காகத் தகவல்கள் சேகரிக்கப் படுகின்றனவோ அதைத் தகவல்கள் பிரதிபலிப்பனவாகவும், அதற்குத் தொடர்பு கொண்டனவாகவும் இருத்தல் வேண்டும். மேலும் தகவல்கள் நோக்கத்திற்கு இன்றியமையாதனவாகவும், பிழையளவிற்குட்பட்டும் இருத்தல் வேண்டும்.

(இ) தகவல்கள் எளிதில் புரியும்படியாக, நோக்கத்தை விளக்குவனவாகச் சித்தரிக்கப்பட்டிருத்தல் வேண்டும்.

(ஈ) தகவல்கள் குறிப்பிட்ட கால வரம்பிற்குள் வேகமாகச் சேகரிக்கப்படல் வேண்டும்.

(உ) தகவல்கள் குறைந்த செலவில் சேகரிக்கப்படல் வேண்டும்.

(ஊ) தகவல்கள் துல்லியமாகவும், சீரிய முறையிலும் விளக்கப்பட்டிருத்தல் வேண்டும்.

1·7. புள்ளியல் பணியின் முதற்கட்டம் குறிப்பிட்ட நோக்கத்திற்கு எத்தகைய தகவல்கள் தேவைப்படும் என்பதை நிர்ணயம் செய்வதுவேயாகும். அத்தகைய தகவல்களைச் சேகரிக்க முடியுமா, அப்படியே முடிந்தாலும் நியாயமான பொருள் வரம்பிற்குள் சேகரிக்க முடியுமா என்பனவற்றை ஆராய்வதுதான் இரண்டாம் கட்டமாகும். மூன்ருவது கட்டம் குறைந்த செலவில் தகவல்கள் சேகரிப்பதைச் செயல்படுத்துவதாகும். நான்காவதும் கடைசியுமான கட்டம் சேகரித்த தகவல்களைக் குறிப்பிட்ட நோக்கங்களுக்கு உதவும் வகையில் அட்டவணைப்படுத்துதலும், விளக்குதலும் ஆகும்.

2. அளவெடுப்பின் படிகள்

2·1. ஒரு குறிப்பிட்ட நோக்கத்திற்காக அளவெடுப்பொன்றை மேற்கொள்ளும் பொழுது பின் குறிப்பிட்டபடி, படிப்படியாகச் செயலாற்றல் வேண்டும்.

(அ) பிரச்சினையைப் புள்ளியியல் கலையின் மொழியில் விவரிப்பது முதற் படியாகும். ஒரு பிரச்சினையை அணுகுவதில் பலவித வழிகள் இருப்பினும் அவை யாவையும் புள்ளியியல் ஆய்வாளருக்கு உகந்தனவாக இருக்க முடியாது. புள்ளியியல் கலைக்கே உரிய சில வழிகளைப் பின்பற்றுவதன் மூலமாகத்தான் எத்தகைய தகவல்களைச் சேகரிக்க வேண்டும் என்பதைத்தெளிவாக உணர முடியும்.

(ஆ) தகவல்களை எந்த இனத்தொகுதியிலிருந்து பெறுவது என்பதைத் தெளிவாக அறிதல் வேண்டும். இனத்தொகுதி எது என்பது தெரியாமல் விவரங்களைச் சேகரிப்பதென்பது முடியாத காரியமாகும். உதாரணமாக, சென்னை மாநகரித்திலுள்ள மக்களின் சராசரி வருவாயைக் கண்டுபிடிக்க வேண்டும் என்றால், நகரத்திலுள்ள வருவாய் பெறும் மக்களடங்கிய பட்டியல்தான் இனத்தொகுதியாகக் கருதப்படல் வேண்டும். இப் பட்டியில் இடம் பெறுவோரை அணுகி, அவர்களது வருவாய் பற்றிய புள்ளி விவரங்களைச் சேகரிக்க வேண்டும். இதைப் போன்ற ஒவ்வோர் அளவெடுப்பிற்கும் ஓர் இனத் தொகுதியை நிர்ணயம் செய்து கொண்ட பிறகே தகவல்களைச் சேகரிக்க முடியும்.

(இ) பிரச்சினைக்குகந்த தகவல்களை முன்னமேயே வேறு யாராவது சேகரித்து வைத்திருக்கிறார்களா என்பதை ஆராய வேண்டும். இத்தகவல்கள் அச்சிடப்பட்டோ அல்லது அச்சிடப்படாமலோ எங்குக் கிடைக்கும் என்பதைப் பிறகு ஆராய

வேண்டும். விவரங்கள் இவ்விதமாக எளிதில் கிடைக்குமாயின், புதிதாக அளவெடுப்பொன்றை மேற்கொள்வதென்பது தேவையற்றதாகும். அரசாங்கத்தில் பல்வேறு துறைகளுக்கு இடையே அதிகத் தொடர்பில்லாக் காரணத்தால் சில சமயங்களில் ஒரே அளவெடுப்பு, பல்வேறு துறைகளால் மேற்கொள்ளப்பட்டு பொருள் விரயமாவதுண்டு.

(ஈ) எத்தகைய அளவெடுப்பு முறை தேவையான விவரங்களைக் குறைந்த செலவில் அளிக்கும் என்பதை ஆராய்ந்து முடிவு செய்தல் வேண்டும். இனத்தொகுதியின் முழுவதிலுமிருந்தும் அளவெடுக்க வேண்டுமா, அல்லது, ஒரு பகுதியை மட்டும் அளவெடுத்தால் போதுமா என்பதை உறுதி செய்து கொள்ளல் வேண்டும். ஆண்டில் எப்பருவத்தில் அளவெடுப்பை வைத்துக் கொள்ளவேண்டும், என்றைக்கு அளவெடுப்பைத் துவக்க வேண்டும் என்பனவற்றை முடிவு செய்துகொள்ளல் வேண்டும். அளவெடுப்பை ஒரு முறை மேற்கொண்டால் போதுமா அல்லது தொடர்ந்து பலமுறை மேற்கொள்ள வேண்டுமா என்பதையும் தீர்மானம் செய்து கொள்ளவேண்டும்.

(உ) இனத் தொகுதியிலுள்ள உறுப்புகளிடம் எத்தகைய முறையில் கேள்விகள் கேட்கப்பட்டால் தேவைப்படும் விவரங்கள் தெளிவாகக் கிடைக்கும் ; விவரங்கள் மறுக்கப்படமாட்டா என்பனவற்றைத் திட்டமிட்டுக்கொள்ளல்வேண்டும். தகவல்கள்சேகரிக்கும் புள்ளியியல் ஆய்வாளர்களையும், தகவல்களைத் தரும் இனத்தொகுதி உறுப்புகளையும் மனதில் கொண்டு புள்ளியியல் அளவெடுப்புப் பணியை விதம் விதமாக வகைப்படுத்திக் கொள்ளல் வேண்டும். கேள்விப் பட்டியலொன்றை நிறுவுவதைத் தொடங்கிவிட வேண்டும். புள்ளியியல் ஆய்வாளர்களையும், அவர்களுக்கான மேற்பார்வையாளர்களையும் பணிக்கு அமர்த்திக் கொண்டு விட வேண்டும். விவரங்கள் திரட்டப்படும்போது எத்தனை சதவீதம் விவரங்கள் மறுக்கப்படுவதை நாம் பொறுத்துக் கொள்ளலாம் என்பதை முடிவு செய்து கொள்ளவேண்டும். விவரங்கள் தர மறுப்பவர்களை எப்படித் தொடர்ந்து சென்று வேண்டுதலாலோ அல்லது வற்புறுத்தலாலோ சரிக்கட்டி விவரங்களைப் பெற வேண்டும் என்பதை ஆராய்ந்தறிதல் வேண்டும். நேரில் சென்று விவரங்களைப் பெறுவதா அல்லது அஞ்சல் மூலம் தொடர்பு கொண்டு விவரங்களைப் பெறுவதா என்பதை முடிவு செய்ய வேண்டும். அளவெடுப்பிற்காக எவ்வளவு பணமும் மற்ற உதவிகளும் கிடைக்கும் என்பதைச் சரியாகத் தெரிந்து கொள்ள வேண்டும்.

(ஊ) அளவெடுப்பு முறைகளில் ஒன்றுக்கு மேற்பட்ட முறைகள் இருப்பின் தோராயமாக ஒவ்வொரு முறையிலும் செலவு எவ்வளவு ஆகும், விவரங்கள் எவ்வளவு பிழையளவிற்குள் பெற முடியும் என்பனவற்றைக் கணக்கிட்டுக் கொள்ளவேண்டும். இக்கணக்கீட்டின்படி எந்த அளவீட்டுமுறை செயலகத்திலும், களத்திலும், கட்டுக்கோப்புடனும், சிக்கனமாகவும் செயல் புரிய அதிகம் உதவும் என்பதை முடிவு செய்து கொள்ளவேண்டும்.

எவ்வளவு பிழையளவை நாம் பொறுத்துக் கொள்ளலாம் என்பதைத் தீர்மானம் செய்துகொள்ளவேண்டும். (2 % ? ,5 % ?, 25 % ? , 50 % ?) இதற்குத் தகுந்தவாறு (அ), (ஆ) என்னும் படிகளை மாற்றியமைத்துக் கொள்ளவேண்டும்.

(எ) அளவெடுப்பு முறையில் பின்வருவனவற்றைக் கருத்திற் கொண்டு ஒரு கூறை (Sample) எவ்வாறு பெறலாம் என்பதைத் தெரிந்து கொள்ளவேண்டும். கூறிலிருந்து அளவெடுப்பின் வீச்சு (Coverage of Survey), தகவல்கள் மறுக்கப்படுவதன் மூலம் ஏற்படும் பிழைகளின் விளைவுகள், ஆய்வாளர்களிடையே உள்ள வேறுபாடுகளின் காரணமாக ஏற்படும் பிழைகளின் விளைவுகள், வெவ்வேறு அளவெடுப்பு முறைகளில் செலவு வேறுபாடு முதலியவை தோராயமாக அறியப்படுதல் வேண்டும்.

(ஏ) களத்தில் விவரங்கள் சேகரிக்கும் ஆய்வாளர்களுக்கான குறிப்புப் பட்டியல் தயாரிக்கவேண்டும். ஆய்வாளர்களுக்குக் களத்தில் ஏற்படும் எல்லா ஐயப்பாடுகளையும் போக்கும் வண்ணம் குறிப்புப் பட்டியல் அமைக்கப் படவேண்டும். ஆய்வாளர்களை எப்படித் தேர்ந்தெடுப்பது, தேர்ந்தெடுத்த பின்னர் அளவெடுப்பு முறைகளில் அவர்களுக்கு எவ்வாறு பயிற்சியை அளிப்பது, விவரங்கள் திரட்ட அவர்களுக்கு எவ்வளவு கால அவகாசமளிப்பது முதலியவற்றைச் செவ்வனே முடிவு செய்து கொள்ள வேண்டும்.

குறிப்பாக ஆய்வாளர்களுக்குப் பின்வருவனவற்றில் பயிற்சியளிக்க வேண்டும்.

(i) இனத்தொகுதிப் பட்டியலை எவ்வாறு நிறுவுவது?

(ii) அப் பட்டியலிலிருந்து ஒரு கூறை எவ்வாறு பிரிப்பது?

(iii) அக் கூறிலுள்ள உறுப்புகளிலிருந்து எவ்வாறு விவரங்களைக் கேட்டறிவது?

(iv) சேகரித்த விவரங்களடங்கிய பட்டியல் தாள்களை எவ்வாறு செயலகத்திற்கு அனுப்புவது?

(v) ஆய்வாளர்களின் மேற்பார்வையாளர்களுக்கு எப்படிப் பயிற்சி அளிப்பது?

(ஐ) சேகரித்த தகவல்களைப் பல பகுதிகளாக எப்படிப் பிரித்துக் கொள்வது, ஒவ்வொரு பகுதியின் அளவும் எவ்வளவு இருக்க வேண்டும்; அப்பகுதிகளுக்கு எப்படித் தலைப்புகளைக் கொடுக்க வேண்டும்; எல்லாவற்றிற்குமாகச் சேர்த்துப் பொதுவாக என்ன தலைப்புக் கொடுக்கவேண்டும் என்பனவற்றை முடிவு செய்து கொள்ளவேண்டும்.

(ஒ) தயாரித்த கேள்விப் பட்டியலை மாதிரிக்காகக் களத்தில் உபயோகப்படுத்தி வினாவுகளைப் பார்த்துக் கொள்ளவேண்டும். இம் முன்னோடி அளவெடுப்பு (Pilot Survey) மிகவும் இன்றியமையாததாகும். இதன் மூலமாகத்தான் அளவெடுப்பின் குறைபாடுகள், கேள்விப் பட்டியலிலுள்ள குறைபாடுகள் முதலியவை முதலிலேயே தெரியவரும். முக்கியமாக, எந்த அளவு விவரங்கள் மறுக்கப்படுகின்றன. தரப்படும் விவரங்களில் எவ்வளவு பிழைகள் இருக்கின்றன; என்பனவற்றை இம்முன்னோடி அளவெடுப்பின் மூலம் தெரிந்துகொள்ளமுடியும். இதன்மூலம் கிடைக்கும் அறிவானது அளவெடுப்பைத் தகுந்தபடி மாற்றியமைத்துக் கொள்ள உதவும்.

(ஓ) முன்னோடி அளவெடுப்பின் மூலம் கிடைத்த விவரங்களை வைத்துக் கொண்டு கேள்விப் பட்டியலையும், ஆய்வாளர்களுக்குக் கொடுக்கப்படும் குறிப்புப் பட்டியலையும் மாற்றியமைத்துக் கொள்ளவேண்டும்

(ஔ) இக்கட்டத்தில் எந்தக் கூறெடுக்கும் முறையை நாம் பின்பற்றவேண்டும் என்பதை உறுதி செய்து கொள்ளவேண்டும். ஆய்வாளர்களின் உதவிக்காக இனத்தொகுதியை விளக்கும் வரைப்படம், பட்டியல் முதலியவற்றைத் தயாரித்துக் கொள்ள வேண்டும். ஆய்வாளர்கள் இனத்தொகுதி உறுப்புகளை எப்படிப் பேட்டி காணவேண்டும்; மேற்பார்வையாளர்கள் ஆய்வாளர்களை எப்படிக் கண்காணிக்கவேண்டும் என்பனவற்றை உறுதியாகத் தீர்மானித்துக் கொள்ளவேண்டும்.

(ஃ) அளவெடுப்பை முன்கூட்டியே திட்டமிட்டபடி செயல் படுத்தவேண்டும். அளவெடுப்பின் மூலம் கிடைத்த விவரங்களைத்

தெளிவான முறையில் வகை வகையாகப் பிரித்துக்கொள்ள வேண்டும்.

அளவெடுப்பைச் செயல் படுத்துவதில் மிகவும் அதிகக் கவனம் செலுத்தவேண்டும். புள்ளியியலாளரின் கடமை அளவெடுப்பை எவ்வாறு செயல்படுத்தவேண்டும் என்று திட்டம் போடுவதுடன் முடிந்துவிடாது. அளவெடுப்பு செயல் படுத்தப்படும்பொழுது ஆய்வாளர்களுடனும், அவர்களின் மேற் பார்வையாளர்களுடனும் இடைவிடாத தொடர்பை ஏற்படுத்திக் கொண்டு மாறுதல்கள் தேவைப்படும்பொழுதெல்லாம் அவை களைச் சுட்டிக் காண்பிக்கவேண்டிய கடமையும் புள்ளியியலாள ருக்கு உண்டு. சரியான மேற்பார்வையில்லாத அளவெடுப்பு, பிழை மிகுந்தே இருக்கும் என்பதைப் புள்ளியியலாளர் எப்பொழுதும் கருத்தில் கொள்வது நல்லது. அதே சமயத்தில் புள்ளியிய லாளரின் மேற்பார்வையென்பது அளவுக்கதிகமாகி ஆய்வாளரின் எரிச்சலுக்கும், தன்மானத்திற்கும், ஊறு விளைவிக்கும் கெடுபிடிக்குக் காரணமாகிவிடக் கூடாது. இவ்விஷயத்தில் ஆய்வாளர், மேற் பார்வையாளரின் மனப்போக்குகளுக்கும் மதிப்புக் கொடுக்க வேண்டும்

(அ அ) ஆய்வாளர்கள் அனுப்பிய விவரங்களின் ஒரு சிறு பகுதியை எடுத்து அதன்மூலம் அளவெடுப்பில் ஏற்பட்ட பிழையளவுகளை மதிப்பிடல் வேண்டும். விவரங்களின் மொத்த அளவு சாதாரணமாக மிகவும் அதிகம் இருக்கும் என்பதால் ஒரு பகுதியை மட்டுமே பிழையளவை மதிப்பிடுவதற்கு உபயோகப் படுத்தவேண்டியது அவசியமாகிறது. பிழையளவை எப்படி மதிப்பிடுவது என்ற கேள்வி எழலாம். தேர்ந்தெடுத்த பகுதியைச் சார்ந்த உறுப்புகளிடமிருந்து மறுபடியும் புள்ளியியலாளரின் நேரிடை மேற்பார்வையில் தகவல்கள் சேகரிக்கப்படல் வேண்டும். இத்தகவல்களை முன்னமேயே சேகரித்த தகவல்களுடன் ஒப்பிட்டுப் பார்த்து பிழையளவை மதிப்பிடலாம்.

(அ ஆ) கிடைத்த விவரங்களைத் தகுந்த முறையில் வகைப் படுத்தி, அவைகளின் மூலம் தெரியவரும் உண்மைகளை விரிவாக அச்சிட வேண்டும்.

இந்த நிலையில்தான் கூறு முறைப் பிழை (Sampling error) எவ்வளவு ஏற்பட்டது என்பதை மதிப்பிட்டுத் தரவேண்டும். இவ்வளவெடுப்பின் முடிவுகளை இதற்கு முன்னர் மேற்கொள்ளப் பட்ட இதே வகையான அளவெடுப்புகளின் முடிவுகளோடு ஒப்பிடவேண்டும்.

ஆட்சியாளர்களின் துணைக்காகச் சிறிய, குறிப்பான உபயோகத்திற்குகந்த முடிவுகளோடு சிறிய அளவில் ஒரு வெளியீட்டைத் தயாரிக்கவேண்டும். இவ்வெளியீட்டில் புள்ளியியல் நுணுக்கங்களையோ, அல்லது, அவசியமற்ற இடைச் செருகல்களையோ சேர்க்கக்கூடாது. ஆட்சியாளர்களுக்குத் தேவையானவையெல்லாம் அவர்களது கொள்கைப்படி செயலாற்றுவதற்கு உபயோகப்படும் சரியான புள்ளிவிவரங்களேயன்றி, புள்ளியியலாளரின் கைத்திறன், அவர் அளவெடுப்பிற்காகப் பட்ட பாடு முதலியவைகளைப்பற்றிய விவரிப்பல்ல. இதை உணர்ந்து முடிவுகளைச் சுருக்கமாக, தேவைக்கேற்ப அளவில் வெளியிடுதல் நல்லது.

ஆராய்ச்சியாளர்களின் துணைக்காக விவரமான ஒரு வெளியீட்டையும் பிரசுரிக்கவேண்டும். இவ்வெளியீட்டில் புள்ளியியல் கலைநுணுக்கங்களையும், அளவெடுப்பு எம்முறையில் செயல்படுத்தப்பட்டது என்பதையும் தெளிவாகவும் விவரமாகவும் விளக்கவேண்டும். இவ்வளவெடுப்பின் தனித்தன்மை, செயல்படுத்துவதில் நேரிட்ட இன்னல்கள், எதிர்காலத்தில் இம்முறையில் செய்யத்தக்க மாறுதல்கள் முதலியவைகளை விளக்கமாகக் குறிப்பிட வேண்டும். மதிப்பீடுகள் சரிவரச் செய்யப்பட்டனவா என்பதைப் பகுத்தறிவதற்கு வேண்டிய விவரங்கள் குறிப்பிடப்படல் வேண்டும். அதாவது வேறொரு புள்ளியியலாளர் தாமாக இதேமுடிவுகளுக்கோ அல்லது வேறு முடிவுகளுக்கோ வருவதற்கான எல்லாவிவரங்களும் தெளிவாகத் தரப்படுதல் வேண்டும். புள்ளியியாலாளரின் கவனக் குறைவாலோ, அல்லது, குறையறிவாலோ, அல்லது ஆட்சியாளர்களைத் திருப்திப் படுத்துவதற்காகவென்றே உண்மைகளைத் திரித்துக் கூறுவதாலோ ஏற்படும் பிழைகள் இம்முறையில்தான் மற்ற புள்ளியியலாளர்களால் நீக்கப்படத்தக்கதாக இருக்கும். இதுதான் ஓர் உண்மையான புள்ளியியலாளர் கொள்ளத்தக்க பண்பாடாகும்.

3. அளவெடுப்பில் நேரக்கூடிய பிழைகள்

3·1. எல்லா அளவெடுப்புகளிலும் பிழைகள் இருக்கும் என்பது மறுக்கமுடியாத உண்மையாகும். பிழையிலா லட்சிய அளவெடுப்பு என்பது கற்பனையுலகைச்சேர்ந்தது; நடைமுறையில் கைக்கொள்ள முடியாதது. ஆயினும் பிழைகள் இருக்கின்றன என்பதற்காக அளவெடுப்பு உபயோகமற்றது என்று கூறுவது உண்மையை நேருக்கு நேர் சந்திக்க விருப்பமில்லாததையே குறிக்கும். சில பிழைகள் புள்ளியியலாளரின் திறமையையும் மீறி நேர்வதுண்டு. சில பிழைகளை, நேரும் என்று புள்ளியியலாளர் அறிந்திருந்தாலும், தவிர்க்கமுடியாத நிலையில் இருக்கக்கூடும். சில பிழைகளினால் எவ்வித இன்னலும் தோன்றாமலேயே போகக் கூடும். சில பிழைகள் ஒன்றுக்கொன்று நேரெதிர்ப்பட்டு அடிபட்டுப்போவதும் உண்டு.

சாதாரணமாக எல்லா விஞ்ஞான ஆராய்ச்சிகளிலும் பிழைகள் நேரிடா வண்ணம் தடுப்பு நடவடிக்கைகள் எடுக்கப்பட்டு, ஆராய்ச்சி முடிவில் பிழைகளின் அளவு புறக்கணிக்கப்படத்தக்க தாகத்தான் எடுத்துக்கொள்ளப்படுகிறது. ஆகவே பிழைகளின் அளவைக் கணிப்பதற்கு எவ்வித முயற்சியும் மேற்கொள்ளப் படுவதுமில்லை. ஆனால் புள்ளியியலாளர்கள் நடத்தும் பரிசோதனை களிலும், அளவெடுப்புகளிலும் பிழைகள் நேர்வது தவிர்க்க முடியாதது என்பது ஒப்புக்கொள்ளப்பட்டு அப்பிழைகளின் அளவு எவ்வளவு இருக்கும் என்பதை மதிப்பிடுவதற்கும் அளவெடுப்பிலேயே வழிவகைகள் செய்யப்படுகின்றன. ஆகவே தான் புள்ளியியலாளர் ஒரு மதிப்பீட்டைத் தரும்போது கூடவே அதில் இருக்கக்கூடிய பிழையளவையும் (1%, 2%, 5%, 10%) குறிப்பிடுவார். மேலும் பொறுத்துக் கொள்ளக்கூடிய பிழையளவிற் கேற்றவாறு அளவெடுப்பு முறையையும் மாற்றிக் கொள்வார்.

புள்ளியியலாளர் தன்னுடைய கலைநுணுக்கங்களின் மூலம் சிறந்த அளவெடுப்பு முறையைத் தேர்ந்தெடுத்து அதன்மூலம் பிழையளவை வெகுவாகக் குறைத்துவிட முடியும். ஆனால் எல்லா வகைப் பிழைகளையும் வெறும் கலைநுணுக்கத்தால் மட்டும் நீக்கிவிட முடியாது. உதாரணமாக, மக்களின் சராசரி வயதை மதிப்பிடுவதற்காக மேற்கொள்ளப்படும் ஓரளவெடுப்பில், மக்கள் வெவ்வேறு காரணங்களுக்காகப் பொய்யான வயதை ஆய்வாளர் களிடம் கூறுவார்களேயானால், புள்ளியியலாளர் தன்னுடைய புள்ளியியல் கலைத்திறன் முழுவதையும் உபயோகப்படுத்தினாலும் பிழையை நீக்கிவிட முடியாது. ஆனால் மக்களிடம் கேள்வி எவ்வகையில் கேட்கப்பட்டால் சரியான வயதைப் பற்றிய விவரம் கிடைக்கும் என்பதை ஆராய்ந்து அளவெடுப்பு முறையை மாற்றியமைத்தால் இப்பிழை வெகுவாகக் குறைந்துவிடும்.

ஆட்சியாளர்களுக்கு எவ்வாறு பிழைகள் ஏற்பட்டன என்பதைப் பற்றிய அக்கறை கிடையாது. பிழைகள் மிகவும் குறைவாக இருக்கவேண்டும் என்றுதான் அவர்களால் கூறமுடியும். புள்ளியியலாளர் பிழைகளுக்கான காரணத்தை ஆய்வாளர்கள், மேற்பார்வையாளர்கள், மக்கள் முதலியோர் மீது சுமத்துவது அழகல்ல. அப்பிழைகளை முன்கூட்டியே எதிர்பார்த்து அவற்றிற் காகத் தடுப்பு நடவடிக்கைகளை மேற்கொண்டிருத்தல் புள்ளியிய லாளரின் கடமை. அல்லது, முன்கூட்டியே இந்த, இந்த காரணங் களினால் இவ்வளவு பிழைகள் ஏற்படும் என்பதை எதிர்பார்த்துக் கூறிவிடவேண்டும். சில சமயங்களில் தெரிந்தே சில பிழைகளைச் செய்ய நேரிடும். இது அளவெடுப்பின் செலவைக் குறைப்பதற்காக இருக்கலாம் ; அல்லது, அப்பிழைகளால் ஆட்சியாளர்களின் காரியங்களில் குறிப்பிடத்தக்க அளவில் குந்தகம் விளையாது என்பதால் பிழைகள் புறக்கணிக்கப்படலாம்.

3·2. பிழைகளில் இருவகை

ஓர் அளவெடுப்பில் ஏற்படும் பிழைகளை இருவகையாகப் பிரிக்கலாம்.

(அ) மாதிரித் தேர்தற் பிழை (Sampling Error)

(ஆ) மாதிரித் தேர்தலாலல்லாத பிழை (Non-sampling Error)

(அ) மாதிரித் தேர்தற் பிழை

மாதிரித் தேர்தற் பிழையென்பது முழுமைத் தொகுதி (Population) முழுவதையும் ஆராயாது அதினின்று ஒரு கூறை யெடுத்து அதலுள்ள உறுப்புகளை மட்டும் ஆராய்ந்து முடிவுகளை

வெளியிடுவதால் ஏற்படும் பிழையாகும். முழுமைத் தொகுதியிலுள்ள உறுப்புகள் யாவும் சீராக ஒரே பண்புடன் இருந்தால் இப்பிழை நேராது. இதையே உறுப்புகள் யாவும் ஒரே படித்தானதாக (Homogeneous) இருந்தால் இம்மாதிரித் தேர்தற் பிழை நேராது எனக் கூறலாம். இப்பிழை உறுப்புகளின் பல படித்தானத் தன்மை (Heterogeneity) காரணமாகத்தான் நேரிடுகிறது.

'ஒரு பானைச் சோற்றுக்கு ஒரு சோறு பதம்' என்று கூறுவது பானையிலுள்ள எல்லாச் சோறும் ஒரு படித்தானதாக இருக்கும் எனும் நம்முடைய நம்பிக்கையைப் பிரதிபலிக்கும் கூற்றுகும். முழுமைத் தொகுதியிலுள்ள எல்லா உறுப்புகளும் பலபடித்தானதாக இருந்தால் மாதிரித் தேர்தற் பிழையை நாம் தவிர்க்க முடியாது. உதாரணமாக, ஒரு பானையில் அரிசி, பருப்பு, பட்டாணி, உருளைக்கிழங்கு சேர்ந்த கதம்பமொன்று வெந்து கொண்டிருந்தால் ஒரு சோற்றைப் பதம் பார்த்து எல்லாம் வெந்துவிட்டிருக்கும் என்று நம்மால் கூறமுடியாது. அப்படிக் கூறினுல் நிச்சயம் நாம் மாதிரித் தேர்தற் பிழையைச் செய்தோராவோம்.

(ஆ) மாதிரித் தேர்தலாலல்லாத பிழை

முழுமைத் தொகுதியிலுள்ள கூறு (Sample) ஒன்றை மட்டும் ஆராய்வதால் ஏற்படும் பிழைகளை மாதிரித் தேர்தற் பிழையென்று குறிப்பிட்டோம். ஆகவே இப்பிழையை முழுமைத் தொகுதியிலுள்ள எல்லா உறுப்புகளையும் ஆராய்வதன் மூலம் போக்கிவிட முடியும். இதையே முழுக்கணிப்பு முறை (Census method) மூலம் மாதிரித் தேர்தற் பிழையை முற்றிலும் நீக்கிவிடமுடியும் என்று கூறலாம். பொதுவாக மாதிரிக் கூறின் பருமன் (Sample size) அதிகரிக்கும்போது மாதிரித் தேர்தற் பிழை குறைந்துகொண்டே வரும். இப் பருமன் முழுமைத் தொகுதியின் பருமன் (Population size) அளவிற்கு வரும்பொழுது மாதிரித் தேர்தற் பிழை இல்லாதொழிந்துவிடுகிறது (vanishes).

ஆனுல் சில பிழைகள் இவ்விதம் நடந்து கொள்வதில்லை, முழுக்கணிப்பு முறையாலும் இப்பிழைகள் மறைந்தொழிவதில்லை. இப்பிழைகளுக்கே மாதிரித் தேர்தலாலல்லாப் பிழைகள் என்பது பெயர். பொதுவாக மாதிரிக் கூறின் பருமன் அதிகரிக்கும்போது மாதிரித் தேர்தலாலல்லாப் பிழைகளும் அதிகரிக்கும். அதாவது விரிந்த மாதிரிகளில் (large sample) மாதிரித் தேர்தலாலல்லாப்

பிழைகள் அதிகமாக இருக்கும். சிறு மாதிரிகளில் (small sample) மாதிரித் தேர்தற் பிழைகள் அதிகமாக இருக்கும்.

3·3. மாதிரிப் பருமன்

முன் கூறியவற்ருல் பொறுத்துக் கொள்ளக் கூடிய பிழையளவைப் பொறுத்து மாதிரிப் பருமனையும் நிர்ணயம் செய்து கொள்ளவேண்டும். மாதிரித் தேர்தற் பிழை, மாதிரிப் பருமன் அதிகரிக்கும் போது குறையும். ஆனல் மாதிரித் தேர்தலாலல்லாப் பிழை மாதிரிப் பருமன் அதிகரிக்கும்பொழுது அதிகரிக்கும். ஆகவே மாதிரிப் பிழையும் மாதிரித் தேர்தலாலல்லாப் பிழையும் மொத்தமாக எந்த மாதிரிப் பருமனுக்கு மிகக் குறைவாக இருக்கிறதோ அந்த மாதிரிப் பருமனையே எடுத்துக் கொள்ள வேண்டும். இப்படிப் பார்க்கும் போது சாதாரணமாக எல்லா அளவெடுப்பிலும் முழுக்கணிப்பு முறையை விட மாதிரித் தேர்தல் முறை.

ஓர் அளவெடுப்பில் மாதிரித்தேர்தலாலல்லாப் பிழைகள் மிகவும் அதிகமாக இருக்குமானல் சிறு மாதிரியை உபயோகிப்பதே தல்லது; மாதிரித் தேர்தல் பிழை மிகவும் அதிகமாக இருப்பின் விரிந்த மாதிரிகளை உபயோகிப்பது நல்லது.

4. அளவெடுப்பில் பிழைகள்

(தொடர்ச்சி)

4. 1. அளவெடுப்பு முறையைச் செயல்படுத்தும் பொழுது நேரும் பிழைகளைப் பொதுவாக இருவகையாகப் பிரித்தோம். இங்குப் பிழைகள் ஏற்படுவதற்கான காரணங்களை விரிவாகப் பார்க்கலாம்.

(அ) நோக்க மயக்கப் பிழைகள்

அளவெடுப்பு நடத்துவதன் நோக்கத்தைச் சரிவரப் புரிந்து கொள்ளாததாலும், அதனால் எத்தகைய தகவல்களைத் திரட்ட வேண்டும் என்பது தெரியாததாலும் நேரக்கூடிய பிழைகளை நோக்க மயக்கப் பிழைகள் எனக் கூறுவது பொருந்தும்.

(ஆ) கேள்விப் பட்டியல் குறைகள் (Errors in Schedules or Questionnaires)

கேள்விப் பட்டியலிலுள்ள கேள்விகள் தெளிவாக இல்லாத தாலும், கேள்விகள் தெளிவாக இருந்தாலும் விடைகள் சரிவரச் சொல்லமுடியாதபடி கேள்விகளின் வாசகங்கள் அமைந்திருப்ப தாலும் பதிலில் பிழைகள் உண்டாகின்றன. சில கேள்விகள் உணர்ச்சிகளைத் தூண்டிவிடும்படி அமைந்துவிட்டால் விருப்பு வெறுப்பின்றி, பதில் சொல்லாத நிலைமை ஏற்பட்டு அதன்மூலம் பிழைகள் நேரிடுகின்றன. அச்சுப் பிழைகள், கண்களுக்கு எரிச் சலையூட்டும்படி அமைந்த அச்சு, விடைகளுக்குச் சரியாக இடம் ஒதுக்காமை, உண்டு—இல்லையென்று குறிப்பிட வேண்டிய பதில்களை எப்படிக் குறிப்பது அல்லது அடிப்பது என்பதில் ஐயப் பாடு ஏற்படும்படியான நிலை முதலியவை கேள்விப் பட்டியலில் இடம் பெறும்போது பதிலில் பிழைகள் நேரிடுகின்றன.

(இ) ஒவ்வொரு கேள்வி முறைக்கும் தகுந்தவாறு கேள்விகள் அமையாமை

தகவல்கள் சேகரிப்பதிலும் பலவித முறைகள் உள்ளன. அவையாவன :

(i) நேரிடையாகப் பேட்டி காண்பது.

(ii) தொலைபேசிமூலம் தொடர்பு கொள்வது.

(iii) அஞ்சல்மூலம் தொடர்பு கொள்வது.

கேள்விகளும் இம்முறைகளுக்குத் தக்கவாறு அமைக்கப்பட வேண்டும். நேரே கேள்வி கேட்கும்போது உபயோகப்படுத்தும் கேள்விகளை மற்ற முறைகளில் சாதாரணமாகக் கேட்க முடியாது. நேரிடைப் பேட்டிகளிலும்கூட விரிவான பேட்டிக்கும், சுருக்கமான பேட்டிக்கும் கேள்விகள் வெவ்வேறு விதமாகக் கேட்கப் படவேண்டும். ஒரே வார்த்தையில் பதில் சொல்லவேண்டிய கேள்விகளையும் விரிவாகப் பதில் சொல்லவேண்டிய கேள்விகளையும் ஒரே மாதிரி அமைக்கக்கூடாது. இவைகளில் கவனக் குறைவாக இருந்தால் பிழைகள் நேரிடுகின்றன.

(ஈ) முழுமைத் தொகுதி மயக்கம்

அளவெடுப்பிற்கேற்ற முழுமைத் தொகுதி என்ன என்று சரிவரத் தெரியாததால் மாதிரி எடுப்பதிலும் அளவெடுப்பதிலும் பிழைகள் நேரிடுகின்றன.

(உ) ஆய்வாளர்கள் மயக்கம்

ஆய்வாளர்களுக்குக் குறிப்புகள் சரிவரக் கொடுக்காமையால்; அவர்களே கேள்விகளைச் சரியாகப் புரிந்து கொள்ளாமையால்; விடைகளை அவர்கள் தவறாக எழுதிக்கொள்வதால் பிழைகள் உண்டாகின்றன.

(ஊ) பதிலின்மைப் பிழை (Error of Non-response)

பேட்டி கொடுப்பவர், ஆய்வாளர் அணுகும்பொழுதெல்லாம் வீட்டில் இல்லாமல் இருக்கலாம். வீட்டிலேயே இருந்தாலும் சில கேள்விகளுக்கு பதில் சொல்ல மறுக்கலாம். அஞ்சல் முறையில், சோம்பலினால் பதிலையெழுதி அனுப்பாமலிருக்கலாம்; சில சமயங்களில் பதில் சொல்லத் தெரியாமல் போகலாம். இப்படிப் பட்ட காரணங்களால் பதிலின்மைப் பிழை உண்டாகிறது.

(சு) காலங்கடந்த பதில்

அஞ்சல் முறையில் காலங்கடந்து பதிலையனுப்புவதால் அவை அளவெடுப்பில் சேர்த்துக்கொள்ளப்படாமல் போகலாம். இதனால் பிழைகள் ஏற்படுகின்றன.

(ஏ) பேட்டி கொடுப்பவர் பிழைகள்

பேட்டி கொடுப்பவர் நேர்மையானவராக இருந்தாலும் கேள்விகளுக்குப் பதில் சொல்வதில் அவரையறியாமல் பிழைகள் நேரிடலாம். கையெழுத்துப் புரியாமலிருக்கலாம்; நினைவுக்குறைவால் தப்பான விடைகளைக் குறிக்கலாம்; நினைவு மயக்கத்தாலும் தவறான தகவல்களைக் கொடுக்கலாம். விடைகள் சரியாகத் தெரியாதபோது ஊகித்துப் பதில் கொடுப்பதில் பிழைகள் இருக்கலாம். தற்பெருமை, பதவிப் பெருமை போன்ற காரணங்களால் கல்வி, தொழில், வருமானம், செலவு முதலியவைகளில் கூட்டியும் வயது, நோய் போன்றவைகளில் குறைத்தும் பதில்கள் அளிக்கலாம்.

(ஐ) ஆய்வாளர்—பேட்டிகொடுப்பவர் தொடர்புப் பிழைகள்

ஆய்வாளர் பேட்டி காணும்போது அவரையறியாமல் கேள்விகளுக்குப் பதில்களைக் குறிப்பாலுணர்த்தலாம்; அல்லது விடைகளைத் திருத்தலாம். இக்காரணங்களால் பிழைகள் நேரிடுகின்றன. பேட்டி கொடுப்பவருக்குச் சில சமயங்களில் ஆய்வாளரைப் பிடிக்காமல் போய்விடலாம். ஆகவே, அவர் வேண்டுமென்றே ஆய்வாளரை மட்டம் தட்டுவதாக நினைத்துத் தவறான விடையளிக்கலாம்; அல்லது, பேட்டி கொடுப்பவருக்கு ஆய்வாளரை மிகவும் பிடித்துவிடலாம்; ஆகவே ஆய்வாளரை மகிழ்விப்பதாக எண்ணித் தவறான பதில்களைத் தரலாம்.

(ஒ) ஆய்வாளர் குறைகள்

ஆய்வாளர் முழுமைத் தொகுதியைச் சரிவரப் புரிந்து கொள்ளாமை; கேள்விகளைச் சரியாகப் புரிந்து கொள்ளாமை; அவருக்கு அளிக்கப்பட்ட குறிப்புகளைச் சரிவரப் புரிந்து கொள்ளாமை; தன்னுடைய களத்தை முழுவதும் சுற்றிப் பாராமை; தன்னுடைய களத்தையும் மீறி அதிகமாகப் பார்த்தல்; தான் யார் என்று பேட்டி கொடுப்பவருக்குக் கூறாமை; ஒரு வீட்டில் யாரைக் கேட்டால் தகுந்த பதில்கள் கிடைக்குமோ அவரை விட்டு விட்டு மற்றவர்களிடம் கேள்வி கேட்டல்; சில கேள்விகளை விட்டு விட்டுத் தானாகவே பதில்களை எழுதிக்கொள்ளல்; சோர்வினால் பதில்களைச் சரியாக எழுதிக்கொள்ளாமை.

மேற்கூறிய காரணங்களால் பிழைகள் உண்டாகின்றன.

(ஓ) திட்டக் குறைகள்

குறிப்புகள் தெளிவாக இல்லாமை ; ஆய்வாளர்கள், மேற்பார்வையாளர்கள் முதலியவர்களைச் சரிவரத் தகுதி பார்க்காமல் அமர்த்திக் கொள்ளல் ; அவர்களுக்குச் சரியாகப் பயற்சி அளிக்காமை ; முழுமைத் தொகுதியைக் கவனக் குறைவால் சரியாகக் குறிக்காதிருத்தல் ; அல்லது பலமுறை ஒரே உறுப்பைக் குறித்தல்.

மேற்கூறிய காரணங்களால் திட்டப் பிழைகள் ஏற்படுகின்றன.

(ஔ) மதிப்பீட்டுக் குறைகள்

விவரங்கள் கிடைத்த பின்னர் அவைகளைத் தரம் பார்த்துப் பிரிக்காமை ; தகுந்த தலைப்பின் கீழ் அவைகளைக் குறிக்காமை ; கூட்டல், கழித்தல், பெருக்கல், வகுத்தல் முதலியவற்றில் தவறு செய்தல்.

மேற்கூறிய காரணங்களால் மதிப்பீட்டுக் குறைகள் நேர்கின்றன.

(ஃ) பிரசுரிப்புக் குறைகள்

அச்சிடும்போது மாதிரித் தேர்தற் பிழையைத் தவருகக் குறித்தல், அளவெடுப்பில் இன்னல்களை மதியாமை ; அளவெடுப்பின் பின்னும் பிரசுரிப்பதற்கு முன்னும் உள்ள இடைக் காலத்தில் ஏற்பட்ட மாறுதல்களைக் கவனியாமை ; ஆகவே தவருன கருத்தை வெளியிடுதல் ; தன் கருத்தையும் சேர்த்துக் கொள்ளல் ; தவருன படங்களைச் சேர்த்தல் ; தகவல்களுக்கு ஒவ்வாத கருத்தினைக் கூறல்.

மேற்கூறிய காரணங்களால் பிரசுரிப்புக் குறைகள் ஏற்படுகின்றன.

5. நிகழ்திறத்தின் அடிப்படை

(Basis of Probability)

5.1. நிகழ்திறம் அல்லது ஊக அளவை (Probability) பல விதங்களில் வரையறுக்கப்படுகிறது. ஆனால் எல்லா விதங்களிலும் ஏதாவதொரு குறையிருக்கத்தான் செய்கிறது. இக் குறைகளையெல்லாம் நீக்கும் பொருட்டு சில எடுகோள்களை (Hypotheses or Axioms) மட்டும் உண்மையென்று எடுத்துக் கொண்டு அதன் மேல் ஊக அளவை எனும் ஓர் இயலே நிறுவப் பட்டது. இந்த எடுகோள்களின் உண்மை நம் மனதிற்கு வெளிப்படையாகத் தெரியலாம், அல்லது அனுபவத்தில் அறிந்த 'உண்மைகளை' எடுகோள்களாக எடுத்துக் கொள்ளலாம். காரணம் யாதாயினும் எந்த இயலில் எடுகோள்கள் முதலேயே எடுத்துக் கொள்ளப்படுகின்றனவோ அந்த இயலில் இவ்வெடு கோள்கள் நிரூபிக்கப்படுவதில்லை. ஆனால் அவ்வெடுகோள்கள் உண்மையென்று எடுத்துக் கொள்ளப்பட்டு, பகுத்தறி முறையால் (Deduction) பல தேற்றங்கள் நிரூபிக்கப்பட்டு, அத்தேற்றங் களும் நடைமுறைக்கு உபயோகமாக இருக்கும்பொழுது அவ் வெடுகோள்களின் அவசியம் நமக்குப் புலனாகிறது. எடுகோள் கள் மாற்றியமைக்கப்படும்பொழுது புதிய இயலொன்று உருவாகி றது. தற்காலத்தில் கணிதத்தின் பெரும் பகுதி இவ்விதமாகத் தான் உருவாக்கப்பட்டது. ஊக அளவை இயலிலும் இம்முறை கையாளப்பட்டதால் இவ்வியல் கணித விற்பன்னர்கள் எதிர் பார்க்கும் தர்க்கரீதியான இசைவு (Logical consistency) கொண்டதாக அமைந்திருக்கிறது.

5.2. மாதிரி வெளி (Sample Space)

ஒரு பரிசோதனையில் பலவித நிகழ்ச்சிகள் நிகழக்கூடும் என எடுத்துக் கொள்வோம். இந்த நிகழ்ச்சிகளின் மொத்தப் பட்டியலை மாதிரி வெளி (Sample space) எனக் குறிப்பிடலாம். மாதிரி வெளி சாதாரணமாக Ω என்னும் எழுத்தால் குறிக்கப்படும்.

மாதிரி வெளியில் முடிவுள்ள எண்ணிக்கையுடைய உறுப்புகள் (Finite number of elements) இருந்தால், அதற்கு முடிவுள்ள மாதிரி வெளி (Finite sample space) என்று பெயரிடலாம்.

மாதிரி வெளியில் 1, 2, 3, . . ∞ என்று எண்ணிடத்தக்க உறுப்புகள் (Countably infinite elements) இருந்தால் அதற்கு எண்ணிடத் தக்க மாதிரி வெளி (Countable sample space) என்று பெயரிடலாம்.

மாதிரி வெளியில் 1, 2, . ∞ என்று கணக்கிட முடியாத உறுப்புகள் (Non-countably infinite elements) இருந்தால் அதற்கு எண்ணிடத் தகா மாதிரி வெளி (Non-countably infinite sample space) எனப் பெயரிடலாம்.

5·3. நிகழ்ச்சிகள் (Events)

சாதாரணமாக மாதிரி வெளியில் இருக்கும் உறுப்புகளில் சிலவற்றை எடுத்து அமைக்கும் ஒவ்வொரு கணத்திற்கும் (Set) நிகழ்ச்சி (Event) எனப் பெயரிடலாம். மாதிரி வெளி முடிவுள்ளதாகவோ அல்லது எண்ணிடத் தக்கதாகவோ இருந்தால் மேற்கூறியதையே நிகழ்ச்சிக்கு வரை விலக்கணமாகக் (Definition) கொள்ளலாம். ஆனால் மாதிரி வெளி எண்ணிடத் தகாததாக இருப்பின் அதனின்று எடுக்கக்கூடிய எல்லாக் கணங்களையும் நிகழ்ச்சிகள் என்று எடுத்துக்கொண்டால் நிகழ்திறம் வரையறை செய்வதில் தர்க்கரீதியான குறைபாடுகள் ஏற்படுகின்றன. எண்ணிடத் தகா மாதிரி வெளியில் நிகழ்ச்சியென்பது என்ன என்பதைப் பிறகு வரையறை செய்வோம்.

5·3·1. எடுத்துக்காட்டு 1.

ஒரு நாணயத்தை ஒரு முறை சுண்டிவிட்டால் உண்டாகும் மாதிரி வெளி

$\Omega = [$ பூ, தலை $]$.

இம் மாதிரி வெளியில் இரண்டே உறுப்புகள் இருப்பதால் இது முடிவுள்ள மாதிரி வெளியாகும். இவ்வெளியில் பின்வரும் நிகழ்ச்சிகள் தாம் இருக்கும்.

(அ) $A = [$பூ$]$ (ஆ) $B = [$தலை$]$ (இ) $C = [$பூ, தலை$]$ (ஈ) $D = [\quad]$.

A, B, C, D என்பவை Ω-விருந்து எடுக்கக்கூடிய கணங்களின் பட்டியலாகும். இவைகளில் $C = \Omega$ என்பது தெளிவாகும்.

D என்பதற்குச் சற்று விளக்கம் தேவை. D என்பது Ω-விருந்து ஓர் உறுப்பு கூட எடுக்காமல் நிறுவிய கணமாகும். ஆகவே நிகழ்ச்சியென்பதன் வரைவிலக்கணப்படி D கூட ஒரு நிகழ்ச்சி தான்.

5·3·2. Ω-விருந்து ஓர் உறுப்புகூட எடுக்காமல் நிறுவப்படும் கணத்திற்கு நடவா நிகழ்ச்சி (Impossible event) என்பது பெயர். இது ϕ என்னும் எழுத்தால் குறிக்கப்படும்.

5·3·3. எடுத்துக்காட்டு 2.

ஆறு முகங்களையுடைய பகடையொன்றை உருட்டுவதால் உண்டாகும் மாதிரி வெளி.

$\Omega = [1, 2, 3, 4, 5, 6]$.

இம்மாதிரி வெளியிலிருந்து 64 நிகழ்ச்சிகளை நிறுவலாம்.

ஒரே ஓர் உறுப்பு கொண்ட நிகழ்ச்சிகள் $\binom{6}{1} = 6$

இரு உறுப்புகள் ,, ,, $\binom{6}{2} = 15$

[$\binom{n}{r}$ என்பது $nC\,r$ஐக் குறிக்கும்]

மூவுறுப்புகள் ,, ,, $\binom{6}{3} = 20$

நான்குறுப்புகள் ,, ,, $\binom{6}{4} = 15$

ஐந்துறுப்புகள் ,, ,, $\binom{6}{5} = 6$

ஆறுறுப்புகள் ,, ,, $\binom{6}{6} = 1$

ϕ : 1

64

5·3·4. துணை முடிவு

பொதுவாக 'n' உறுப்புகள் கொண்ட மாதிரி வெளியில் வரையறை செயத்தக்க நிகழ்ச்சிகளின் எண்ணிக்கை $= 1+\binom{n}{1}$

$$+\binom{n}{2}+\cdots+\binom{n}{n-1}+\binom{n}{n}=(1+1)^n=2^n.$$

5·3·5. எடுத்துக்காட்டு 3

ஒரு நாணயத்தைத் 'தலை' விழும் வரை சுண்டிக்கொண்டே இருந்தால் உண்டாகும் மாதிரி வெளி $\Omega =$ [(தலை); (பூ,தலை); (பூ, பூ, தலை); (பூ, பூ, பூ, தலை); ...]. இம் மாதிரி வெளி எண்ணிடத்தக்க மாதிரி வெளியாகும். இதில் ∞ உறுப்புகள் உள்ளன. இவ்வெளியில் எண்ணிடத்தகாத அளவிற்கு நிகழ்ச்சிகள் உள்ளன. [$2^{\infty} =$ 0-விருந்து 1 வரையுள்ள எல்லா மெய்யெண்களின் (real numbers) எண்ணிக்கையாகும். இது எண்ணிடத்தகாதது என்பதைக் கணிதப் பகுவியல் (Mathematical analysis) நூல்களில் காணலாம்]

5·4. நிகழ்ச்சிகளின் இயற் கணிதம் (Algebra of Events)

Ω என்பது ஒரு மாதிரி வெளி. $A, B, C, \ldots$ முதலியவை Ω-லுள்ள நிகழ்ச்சிகள். இவைகளை Ω-ன் பகுதிகளாகவும் கொள்ளலாம். Ω-லுள்ள உறுப்புகளைப் பொதுவாக ω எனும் எழுத்தால் குறிப்பிடலாம்.

ω எனும் உறுப்பு Ω-ல் உள்ளது என்பதை இயற்கணித முறையில் $\omega \in \Omega$ எனக் குறிப்பிடலாம் [ω, Ω-லுள்ளது எனப் படிக்கவும்].

$\omega \in A$ என்பது ω, A என்னும் நிகழ்ச்சியிலுள்ள உறுப்பு என்பதைக் குறிக்கும்.

5·4·1. அடக்கம்

A-லுள்ள எல்லா உறுப்புகளும் B-ல் இருந்தால் A, B-ல் அடக்கம் எனக் கூறலாம். இதையே இயற்கணித முறைப்படி

$A \subset B$ எனக் குறிப்பிடலாம் [A, B-ல் அடக்கம் எனப் படிக்கவும்].

இதையே $B \supset A$ எனவும் எழுதலாம் [B, A ஐ அடக்கும் எனப் படிக்கவும்].

கணவியல் (Set theory) முறையின்படி $A \subset B$ என்பதை A, B-ன் உபகணம் (Subset) எனக் குறிப்பிடலாம்.

5·4·2. a எனும் உண்மையிலிருந்து b எனும் உண்மையை வருவிக்க (dedue) முடிந்தால் $a \Longrightarrow b$ என இயற்கணித முறைப்படி குறிக்கலாம்.

[a, b ஐ வருவிக்கிறது எனப் படிக்கவும்].

எடுத்துக்காட்டுகள்

(அ) n எனும் எண் 4 ஆல் வகுபடும் $\Longrightarrow$ n இரண்டால் வகுபடும்.

(ஆ) n எனும் முழு எண் ஒற்றைப்படை எண் $\Longrightarrow n \neq 2$.

(இ) நான் சென்னையைச் சேர்ந்தவன் $\Longrightarrow$ நான் கல்கத்தாவைச் சேர்ந்தவனல்ல.

5·4·3. $a \Longrightarrow b$, $b \Longrightarrow a$ ஆயின் $a \Longleftrightarrow b$ எனக் குறிப்பிடலாம்.

எடுத்துக்காட்டுகள்

(அ) n எனும் எண் 12 ஆல் வகுபடும் $\Longleftrightarrow$ n, 3 ஆலும் 4 ஆலும் வகுபடும்.

(ஆ) நாணயத்தைச் சுண்டிவிட்டதில் 'தலை' விழுந்தது $\Longleftrightarrow$ 'பூ' விழவில்லை.

(இ) $a \Longrightarrow b, b \Longrightarrow a \Longleftrightarrow a \Longleftrightarrow b$

(ஈ) $A \subset B \Longleftrightarrow \omega \in A \Longrightarrow \omega \in B$

5·4·4. கூட்டு (Union)

A, B என்பவை Ω-லுள்ள கணங்களாயின், A-ல் மட்டுமுள்ள, B-ல் மட்டுமுள்ள A, B இரண்டிலுமுள்ள உறுப்புகளால் மட்டும் அமைக்கப்படும் கணத்திற்கு $A \cup B$ என்பது பெயர். இதை A கூட்டு B எனப் படிக்கவும்.

துணை முடிவுகள்

(அ) $x \in A$ அல்லது $x \in B \Longrightarrow x \in A \cup B$

(ஆ) $A \cup B = B \cup A$

(இ) $A \cup [B \cup C] = [A \cup B] \cup C$

(ஈ) $A \subset A \cup B$

(உ) $A \cup \phi = A$

(ஊ) $A \cup \Omega = \Omega$

(எ) $A \cup A = A$

5·4·5. வெட்டு (Intersection)

A, B என்பவை Ω-லுள்ள கணங்களாயின், A, B-லுள்ள பொது உறுப்புகளை மட்டும் கொண்டு அமைக்கப்படும் கணத்தை $A \cap B$ என்று குறிப்பிட வேண்டும். இதை A வெட்டு B எனப் படிக்கவும்.

துணை முடிவுகள்

(அ) $x \in A, x \in B \Longrightarrow x \in A \cap B$

(ஆ) $A \cap B = B \cap A$

(இ) $A \cap [B \cap C] = [A \cap B] \cap C$

(ஈ) $A \cap B \subset A$

(உ) $A \cap \phi = \phi$

(ஊ) $A \cap \Omega = A$

(எ) $A \cap A = A$

5·4·6. நிரப்பி (Complement)

A என்பது Ω-லுள்ள கணமானல் A-ல் இல்லாத Ω-லுள்ள உறுப்புகளை மட்டும் கொண்டு அமைக்கப்படும் கணத்தை A^c எனக் குறிப்பிட வேண்டும். இதை A-ன் நிரப்பி யெனப் படிக்கவும்.

துணை முடிவுகள்

(அ) $\phi^c = \Omega$; $\Omega^c = \phi$

(ஆ) $[A^c]^c = A$

(இ) டிமார்கன் விதிகள் (De Morgan Laws)

(i) $[A \cup B]^c = A^c \cap B^c$

(ii) $[A \cap B]^c = A^c \cup B^c$

(ஈ) $A \cup A^c = \Omega$; $A \cap A^c = \phi$

5·4·7. குறியிடும் கணம் (Index Set)

A_1, A_2, A_3, A_4 என்று நான்கு கணங்கள் இருந்தால், இவைகளை A_i; $i = 1, 2, 3, 4$ எனக் குறிக்கலாம். இதில் [1, 2, 3, 4] என்பது கணங்களைக் குறிக்கும் குறியீட்டை விளக்குவதால் இதைக் குறியிடும் கணம் எனக் கூறலாம். பொதுவாகக் குறியிடும் கணத்தை I என்று எழுதலாம்.

எடுத்துக்காட்டு

(அ) $[A_1, A_2, A_3, \ldots, A_n] = [A_i]$, $i \in [1, 2, 3, \ldots n]$
இங்கு I என்பது $[1, 2, \ldots, n]$ ஐக் குறிக்கும்,

(ஆ) $[A_1, A_2, \ldots, A_\infty] = [A_i]$, $i \in [1, 2, \ldots, \infty]$
இங்கு I என்பது $[1, 2, 3, \ldots \infty]$

(இ) $[\quad] = [A_i] \quad i \in \phi$

5·4·8. (அ) $[A_1 \cup A_2 \cup A_3 \ldots \cup A_n]$ என்பதை $\bigcup_{i=1}^{n} A_i$ என்று குறிக்கலாம்.

(ஆ) $[A_1 \cup A_2 \cup A_3 \ldots \cup A_\infty]$ என்பதை $\bigcup_{i=1}^{\infty} A_i$ என்று குறிக்கலாம்.

(இ) $\bigcup_{i \in I} A_i$ என்பது A_i-களின் கூட்டைக் குறிக்கும்; I என்பது i-ன் குறியிடும் கணமாகும்.

5·4·9. டிமார்கன் விதிளிகள் நீட்சி (Extension of De Morgan Laws)

(அ) $$\left[\bigcup_{i=1}^{\infty} A_i\right]^c = \bigcap_{i=1}^{\infty} A_i^c; \quad \left[\bigcup_{i=1}^{n} A_i^c\right]^c = \bigcap_{i=1}^{n} A_j^c$$

(ஆ) $$\left[\bigcap_{i=1}^{\infty} A_i\right]^c = \bigcup_{i=1}^{\infty} A_i^c; \quad \left[\bigcap_{i=1}^{n} A_i\right]^c = \bigcup_{i=1}^{n} A_i^c$$

5·4·10. பிரிந்த கணங்கள் (Disjoint Sets)

A, B என்ற இரண்டிற்கும் பொதுவான உறுப்பு ஒன்றுகூட இல்லையென்றால் இவைகளைப் பிரிந்த கணங்கள் எனக் கூறலாம். நிகழ்திறத்திற்குரிய மொழியில் இதையே ஒன்றையொன்று விலக்கும் கணங்கள் (நிகழ்ச்சிகள்) என்று கூறலாம்.

துணை முடிவு

(அ) A, B பிரிந்த கணங்கள் $\Longleftrightarrow A \cap B = \phi$

(ஆ) $A_1, A_2, \ldots A_n$ என்பவை பிரிந்த கணங்கள் என்பதை $A_i \cap A_j = \phi$; $i \neq j$; $i, j = 1, 2, 3, \ldots n$ என்று குறிப்பிடலாம்.

(இ) A, A^c பிரிந்த கணங்கள்.

(ஈ) $A - B$ என்பது A-ல் உள்ள B-ல் இல்லாத உறுப்புகளைக் கொண்டு அமைக்கப்படும் கணமாகும். இதிலிருந்து $B, A - B$ என்பவை பிரிந்த கணங்கள் என்பது தெரிகிறது.

5·4·11. எடுத்துக்காட்டுகள்

(அ) Ω என்பது $[1, 2, 3, 4, 5, 6]$

A ,, $[1, 3, 4, 6]$

B ,, $[2, 3, 4, 5]$

C ,, $[3, 4]$ என்றால் பின்வருபவை தெளிவாகும்.

(i) $A \cup B = [1, 2, 3, 4, 5, 6] = \Omega$

(ii) $A \cup C = [1, 3, 4, 6] = A$; $B \cup C = B$

(iii) $A \cap B = [3, 4] = C$; $A \cap C = B \cap C = C$

(iv) $A^c = [2, 5]$; $B^c = [1, 6]$; $C^c = [1, 2, 5, 6]$

(ஆ) Ω என்பது சென்னையிலுள்ள மக்கள்,

A : சென்னையிலுள்ள ஆண்கள்,

B : சென்னையில் பள்ளியில் படித்துக்கொண்டிருப்பவர்,

C : சென்னையிலுள்ள கிறித்தவர் என்றால்,

(i) A^c : சென்னையிலுள்ள பெண்கள். B^c = சென்னையில் பள்ளியில் படிக்காதவர்.

C^c : சென்னையிலுள்ள கிறித்தவரல்லாதார்.

(ii) $A \cap B$: சென்னையில் பள்ளியில் படித்துக் கொண்டிருக்கும் ஆண்கள்.

$A \cap C$: சென்னையிலுள்ள ஆண் கிறித்தவர்.

$B \cap C$: சென்னையில் பள்ளியில் படித்துக் கொண்டிருக்கும் கிறித்தவர்.

$A \cap B \cap C$: சென்னையில் பள்ளியில் படித்துக் கொண்டிருக்கும் கிறித்துவ ஆண்கள்.

$A^c \cap B \cap C$: சென்னையிள் பள்ளியில் படித்துக் கொண்டிருக்கும் பெண்கள்.

5·5. கணக்களம் (Field of Sets)

5·5·1. Ω என்பது ஒரு வெளியைக் குறிப்பிட்டால் இவ்வெளியிலிருந்து பல கணங்களை நிறுவலாம். இக்கணங்களைக் கொண்டு அமைக்கப்படும் தொகுதிகளுக்குக் கணத் தொகுதிகள் (class of sets) என்று பெயர்.

எடுத்துக்காட்டாக Ω என்பது [1, 2, 3, 4, 5, 6] என்பதைக் குறித்தால்,

$\{ [1], [1, 2], [3, 4, 5], [1, 3, 6] \}$ என்பது ஒரு கணத் தொகுதியாகும்.

$\{ \phi, [2, 3], [2, 3, 5, 6], \Omega \}$ என்பதும் ஒரு கணத் தொகுதியாகும்.

கணத்தொகுதியின் ஒவ்வோர் உறுப்பும் ஒருகணம் என்பதை நினைவில் நிறுத்தவேண்டும். ஆனால் கணத்திலுள்ள ஒவ்வோர் உறுப்புக்கும் உறுப்பு என்றுதான் பெயர்.

5·5·2. அடுக்குக் கணம் (Power set)

Ω என்பது ஒரு வெளியானால் அதிலிருந்து எடுக்கக்கூடிய எல்லாக் கணங்களையும் கொண்ட கணத் தொகுதிக்கு அடுக்குக் கணம் என்பது பெயர். இதை Ω^p என்று குறிப்பிடவேண்டும்.

எடுத்துக்காட்டு

Ω [1, 2, 3] என்றால்,

$\Omega^p = [\phi, [1], [2], [3], [1, 2], [1, 3], [2, 3], \Omega]$ ஆகும்.

பொதுவாக Ω-ல் n உறுப்புகள் இருந்தால் Ω^p-ல் 2^n கணங்கள் இருக்கும்.

5·5·3. கணக்களம்

Ω என்பது ஒரு வெளியாக இருக்கட்டும். F என்னும் ஒரு காலியில்லாத கணத் தொகுதியைப் பின்வரும் விதிகளை மதிக்குமாறு அமைத்தால் அதற்குக் கணக் களம் என்பது பெயர்.

விதி 1

A, B எனும் கணங்கள் F-ல் இருந்தால் $A \cup B$-ம் இருக்க வேண்டும்.

$$\text{அதாவது } A \in F,\ B \in F \Longrightarrow A \cup B \in F$$

விதி 2

A எனும் கணம் F-ல் இருந்தால் A^c-ம் F-ல் இருக்க வேண்டும்.

5·5·4. எடுத்துக்காட்டுகள்

(அ) $\Omega = [1, 2, 3, 4]$

$A = [1, 2]$, $B = [3, 4]$, $C = \phi$, $D = \Omega$ என்றால்

$F = [A, B, C, D]$ என்பது ஒரு கணக்களமாகும்.

நிறுவல்

$$A \cup B = D;\ A \cup C = A;\ A \cup D = D;$$
$$B \cup C = B;\ B \cup D = D;\ C \cup D = D;$$
$$A^c = B;\ B^c = A;\ C^c = D;\ D^c = C$$

என்பவைகளிலிருந்து F என்பது ஒரு கணக்களம் என்பது தெளிவாகிறது.

(ஆ) $\Omega = [1, 2, 3, 4]$

$$A = [1, 2];\quad B = [2, 3, 4];\quad C = \phi;\quad D = \Omega$$

என்றால்,

$F = [A, B, C, D]$ என்பது ஒரு கணக்களமல்ல.

ஏனெனில் $B^c = [1]$ என்பது இத்தொகுதியிலில்லை.

(இ) $F = \Omega^p$ என்பது ஒரு கணக் களமாகும். ஏனெனில்

$$A, B \in \Omega^p \Longrightarrow A \cup B \in \Omega^p$$

$$A \in \Omega^p \Longrightarrow A^c \in \Omega^p$$

இவையிரண்டும் Ω^p-ல் எல்லாக் கணங்களும் உள்ளன என்பதி லிருந்து புலனாகிறது.

(ஈ) $F = [\phi, \Omega]$ என்பது ஒரு கணக் களமாகும்.

ஏனெனில் $\phi^c = \Omega$, $\Omega^c = \phi$, $\phi \cup \Omega = \Omega$

தேற்றம் 5·5·5.

F என்பது ஒரு கணக்களமானால்

(அ) $\phi \in F$, $\Omega \in F$.

(ஆ) $A \in F$, $B \in F \Longrightarrow A \cap B \in F$.

(இ) $A \in F$, $B \in F \Longrightarrow A - B \in F$.

(ஈ) $A_1, A_2, \ldots, A_n \in F \Longrightarrow \bigcup_{i=1}^{n} A_i \in F$; n என்பது முடிவுள்ள எண்ணானால்

(உ) $A_1, A_2, \ldots, A_n \in F \Longrightarrow \bigcap_{i=1}^{n} A_i \in F$ „ „ „ „

நிறுவல்

(அ) F என்பது காலியில்லாத கணத் தொகுதியாதலால், அதில் A எனும் ஒரு கணம் இருக்கும்.

$$A \in F \Longrightarrow A^c \in F$$

$$\therefore \quad A \cup A^c \in F \Longrightarrow \Omega \in F$$

$$\Omega \in F \Longrightarrow \Omega^c = \phi \in F$$

(ஆ) $A, B \in F \Longrightarrow A^c, B^c \in F \Longrightarrow A^c \cup B^c \in F$

$$\Longrightarrow [A^c \cup B^c]^c \in F \Longrightarrow A \cap B \in F$$

[டிமார்கன் விதியால்]

(இ) $A, B \in F \Longrightarrow A, B^c \in F \Longrightarrow A \cap B^c \in F$

$\Longrightarrow A - B \in F$ [$A - B = A \cap B^c$ என்பது தெளிவாகும்.]

(ஈ) $A_1, A_2, \ldots, A_n \in F \Longrightarrow A_1 \cup A_2 \in F \Longrightarrow [A_1 \cup A_2] \cup A_3 \in F$

$$\Longrightarrow [A_1 \cup A_2 \cup A_3] \in F \Longrightarrow [A_1 \cup A_2 \cup A_3] \cup A_4 \in F \Longrightarrow \ldots$$

$$\ldots \Longrightarrow [A_1 \cup A_2 \cup A_3 \ldots \cup A_n] \in F \Longrightarrow \bigcup_{i=1}^{n} A_i \in F.$$

(உ) $A_1, A_2, \ldots A_n \in F \Longrightarrow A_1^c, A_2^c, \ldots, A_n^c \in F$

$$\Longrightarrow \bigcup_{i=1}^{n} A_i^c \in F \quad \text{[(ஈ)-ன்படி]}$$

$$\Longrightarrow \left[\bigcup_{i=1}^{n} A_i^c\right]^c \in F$$

$$\Longrightarrow \bigcap_{i=1}^{n} A_i \in F \quad \text{[டீமார்கன் விதியால்]}$$

5·5·6. σ-கணக் களம் (σ-field)

Ω எனும் வெளியில் F என்பது ஒரு கணக் களமாக இருக்கட்டும். பின்வரும் விதியையும் F பின்பற்றினுல் F ஐ ஸிக்மா கணக் களம் என்று கூறவேண்டும்.

σ- விதி

$A_1, A_2, \ldots, A_\infty$ என்பவை F-ல் இருந்தால்,

$\bigcup_{i=1}^{\infty} A_i$-ம் F-ல் இருக்கவேண்டும்.

அதாவது $A_i \in F, i = 1, 2, \ldots \infty \Longrightarrow \bigcup_{i=1}^{\infty} A_i \in F$

தேற்றம்

F என்பது ஒரு σ-கணக் களமானுல், $A_1, A_2, \ldots A_\infty$ என்பவை F-ல் உள்ள கணங்களானுல் $\bigcap_{i=1}^{\infty} A_i$-ம் F-ல் இருக்கும்.

அதாவது $A_1, A_2, \ldots, A_\infty \in F \Longrightarrow \bigcap_{i=1}^{\infty} A_i \in F$.

நிறுவல்

$$A_1, A_2, \ldots A_\infty \in F \Longrightarrow A_1^c, A_2^c, \ldots, A^c_\infty \in F$$

$$\Longrightarrow \bigcup_{i=1}^{\infty} A_i^c \in F \quad [\sigma\text{-விதியின்படி}]$$

$$\Longrightarrow \left[\overline{\bigcup_{i=1}^{\infty} A_i^c}\right]^c \in F$$

$$\Longrightarrow \bigcap_{i=1}^{\infty} A_i \in F \quad [\text{டிமார்கன் விதியின்படி}]$$

5.6 போரல் கணக்களம் (Borel Field)

5.6.1 $-\infty, +\infty$ அல்லாத மற்ற எண்கள் மெய்யெண்கள் (real numbers) எனக் கூறப்படுகின்றன.

அதாவது x ஒரு மெய்யெண் $\Longrightarrow -\infty < x < +\infty$.

a, b என்பவை மெய்யெண்களானால் $[a<b]$

(i) (a, b) என்பது a, b நீங்கலாக a-லிருந்து b வரையுள்ள எல்லா மெய்யெண்களையும் கொண்ட கணமாகும். இதைத் திறந்த இடைவெளி எனலாம் (open interval).

(ii) $[a, b]$ என்பது a, b உள்பட a-லிருந்து b வரையுள்ள எல்லா மெய்யெண்களையும் கொண்ட கணமாகும். இதை மூடிய இடைவெளி (closed interval) எனலாம்.

(iii) $[a, b)$ என்பது a உள்பட b நீங்கலாக a-லிருந்து b வரையுள்ள எல்லா மெய்யெண்களையும் கொண்ட கணமாகும். இதைப் பாதி திறந்த இடைவெளி (Semi open interval) எனலாம். இது வலப் புறம் திறந்த இடைவெளியாகும் (Right open interval).

(iv) $(a, b]$ என்பது a நீங்கலாக b உள்பட a-லிருந்து b வரையுள்ள எல்லா மெய்யெண்களையும் கொண்ட கணமாகும். இதுவும் பாதி திறந்த இடைவெளியாகும். இது இடப் புறம் திறந்த இடைவெளியாகும்.

5·6·2. இடைவெளிக் கணத் தொகுதி (Class of interval)

$\Omega = (-\infty, \infty)$ எனக் கொள்வோம். (a,b), $[a,b]$, $[a, b)$, $(a,b]$, என்னும் எல்லாவிதமான இடைவெளிகளையும் (எல்லா a,b எனும் மெய்யெண்களுக்கும், $a \leq b$) கொண்ட கணத் தொகுதிக்கு இடைவெளிக் கணத் தொகுதி என்பது பெயர். இதை I என்று குறிப்பிடலாம். I-ன் ஒவ்வோர் உறுப்பும் ஓர் இடைவெளியாகும்.

I என்பது ஒரு கணக்களம் ஆகாது. ஏனென்றால் $[0,1]$ $[10, 12]$ என்பவை I-ல் இருந்தாலும் $[0, 1] \cup [10,12]$ என்பது ஓர் இடைவெளியாக இல்லாததால் இது I-ல் இருக்க முடியாது. இது கணக்களத்தின் முதல் விதியை மீறுவதால் இது கணக்களம் ஆகாது.

5·6·3. போரல் கணக்களம்

I என்பது கணக்களம் இல்லையாயினும் I உடன் இன்னும் பல கணங்களைச் சேர்த்து ஒரு σ கணக்களத்தையே உருவாக்க முடியும்.

B_r எனும் σ-கணக்களத்தைப் பின்குறிப்பிட்ட கணங்களைக் கொண்டு உருவாக்க வேண்டும்.

(i) I-ல் உள்ள எல்லா இடைவெளிகள்.

(ii) I-ல் உள்ள எல்லா எண்ணிடத்தக்க இடைவெளி களின் கூட்டுகள்.

அதாவது $I_1, I_2, \ldots I_\infty$ என்பவை I-ல் இருந்தால்

$\bigcup_{i=1}^{\infty} I_i$ ஐயும் B_R-ல் சேர்க்கவேண்டும்,

B_R ஒரு σ-கணக்களம் என்பதும் (i), (ii) என்பவைகளைக் கொண்டு நிறுவலாம். இதுவே I ஐ அடக்கும் மிகச்சிறிய σ கணக்களமுமாகும்.

போரல் கணக்களம் நிகழ் திறத்தை வரையறுப்பதற்கும் லெபெக் தொகையிடலுக்கும் (Lebesgue Integration) இன்றி யமையாததாகும்.

5·7 நிகழ்ச்சியை வரையறுத்தல்

Ω என்பது ஒரு மாதிரி வெளியாக இருக்கட்டும் இம்மாதிரி வெளியில் F என்பது ஒரு σ-கணக்களமாக இருக்கட்டும். F-ல் உள்ள ஒவ்வொரு கணத்தையும் நிகழ்ச்சியென்று கூறலாம்.

மேற்கூறியபடி இயற்கணித முறையில் நிகழ்ச்சியை வரை யறுத்தலில் பலவித அனுகூலங்கள் இருக்கின்றன.

கவனிக்கவும் :

Ω என்பது முடிவுள்ள உறுப்புகளைக் கொண்டதாகவோ, அல்லது **எண்ணத்தக்க** உறுப்புகளைக் கொண்டதாகவோ இருந்தால் F என்பது எப்பொழுதும் Ω^p என்பதையே குறிக்கும். அதாவது நிகழ்ச்சிகளென்பன Ω-ன் அடுக்குக் கணத்திலுள்ள கணங்களைக் குறிக்கும். அதாவது Ω-லுள்ள எல்லாக் கணங் களையுமே நிகழ்ச்சிகளாகக் கருதவேண்டும்.

துணை முடிவு

(அ) A, B என்பவை நிகழ்ச்சிகளானல்

$A \cup B$, $A \cap B$, $A - B$, A^c முதலியவை யாவும் நிகழ்ச்சி களாகும்.

நிறுவல்

நிகழ்ச்சிகளின் தொகுதி ஒரு கணக்களம் என்பதிலிருந்து இவைகள் தொடர்கின்றன.

(ஆ) $A_1, A_2, \ldots\ldots\ldots A_n$ என்பவை நிகழ்ச்சிகளானல்

$$\bigcup_{i=1}^{n} A_i, \qquad \bigcap_{i=1}^{n} A_i$$ என்பவை நிகழ்ச்சிகளாகும்.

நிறுவல்

நிகழ்ச்சிகளின் தொகுதி ஒரு கணக்களம் என்பதிலிருந்து இவைகள் தொடர்கின்றன.

(இ) $A_1, A_2, \ldots\ldots A_\infty$ என்பவை நிகழ்ச்சிகளானல்

$$\bigcup_{i=1}^{\infty} A_i, \qquad \bigcap_{i=1}^{\infty} A_i$$ என்பவை நிகழ்ச்சிகளாகும்.

நிறுவல்

நிகழ்ச்சிகளின் தொகுதி ஒரு σ-கணக்களம் என்பதிலிருந்து இவைகள் தொடர்கின்றன.

6. சார்புகள்

(FUNCTIONS)

6·1 : D, R என்பவை இரு கணங்களாக இருக்கட்டும். D-லுள்ள ஒவ்வோர் உறுப்புக்கும் R-ல் உள்ள **ஓர்** உறுப்பைத் தொடர்புப்படுத்துவதாக வைத்துக் கொள்வோம். இத்தொடர்பையே சார்பு எனக் கூறுகிறோம்.

எடுத்துக்காட்டுகள்

(அ) $D: [a, b, c, d]$.

$R: [p, q, r, s, t]$.

அம்புக்குறி மூலம் தொடர்பை விளக்குவோம்.

$a \longrightarrow q$; $b \longrightarrow r$; $c \longrightarrow q$; $d \longrightarrow t$.

இத்தொடர்பையே சார்பு எனலாம். அம்புக் குறிக்குப் பதிலாக f என்ற எழுத்தையும் தொடர்பை விளக்குவதற்கு உபயோகப் படுத்தலாம். அதாவது மேலே கூறப்பட்ட சார்பு $f(a) = q$; $f(b) = r$; $f(c) = q$; $f(d) = t$ எனக் குறிப்பிடப்படலாம்.

(ஆ) $D: [0, 1]$

$R: [1, 3]$

$x \in D$ என்றால், $f(x) = x^2 + 1$ என எடுத்துக்கொண்டால், $f(x) \in R$ என்பது தெளிவு.

(இ) $D: [-1, 1]$

$R: [0, 1]$

$x \in D, f(x) = x^2 \in R$.

(ஈ) $D: [0, 1]$.

$R: [0, 1]$.

$x \in D, f(x) = x^2 \in R$.

6·2. D என்பதற்குச் சார்பின் அரங்கம் (Domian of function) என்பது பெயர்.

R என்பதற்குச் சார்பின் வீச்சு (Range of function) என்பது பெயர்.

ஆகவே சார்பானது தன்னுடைய அரங்கத்திலுள்ள ஒவ்வோர் உறுப்புடனும் தன்னுடைய வீச்சிலுள்ள ஓர் உறுப்பைத் தொடர்புப் படுத்துகிறது.

x என்பது அரங்கத்தில் இருந்தால் x-ன் சார்பான $f(x)$ வீச்சிலிருக்கும். சார்பு என்பதை உருமாற்றம் (Transformati n), படவாக்கம் (Mapping) முதலிய பெயர்களாலும் அழைக்கலாம். $f(x)$ஐ x-ன் நிழல் (Image) என்றும் கூறலாம்.

6·3. X என்பது D-லுள்ள கணமானல், X-லுள்ள ஒவ்வோர் உறுப்பினுடைய நிழலையும் உறுப்பாகக் கொண்ட R-லுள்ள கணத்தை X-ன் நிழல் எனலாம். இதை X-ன் நிழல் எனக் குறிப்பிடலாம். இதையே $f(x)$ என்று இயற்கணித முறையில் எழுதலாம்.

X என்பது ஒரு கணமானல் $f(X)$ என்பதும் ஒரு கணம் என்பதை உணர்க.

(அ) $X \subset D \Longrightarrow f(X) \subset R$.

(ஆ) $X = D$ எனக் கொண்டால் $f(D) \subset R$ என்று தெளிவாகிறது.

6·3·1. தேற்றம்

$X, Y \in D;\ f(D) \subset R$.

(அ) $f(X \cup Y) \subset f(X) \cup f(Y)$.

(ஆ) $f(X \cap Y) \subset f(X) \cap f(Y)$.

(இ) $X \subset Y \Longrightarrow f(X) \subset f(Y)$.

நிறுவல்

(அ) $x \in X \cup Y \Longrightarrow x \in X$ அல்லது $x \in Y$ அல்லது $x \in X \cap Y \Longrightarrow f(x) \in f(X)$ அல்லது $f(x) \in f(Y)$ அல்லது $f(x) \in f(X) \cap f(Y)$.

$\Longrightarrow f(x) \in f(X) \cup f(Y)$.

(ஆ), (இ)-இவைகளையும் இதைப்போலவே நிறுவலாம்.

6·3·2. ஒன்றுக்கொன்றான சார்பு

f என்பது D ஐ அரங்கமாகவும் R ஐ வீச்சாகவும் கொண்ட ஒரு சார்பாக இருக்கட்டும். வெவ்வேறு D-ன் உறுப்புகளுக்கு நிழல்களும், வெவ்வேருக இருப்பின் f ஐ ஒன்றுக்கொன்றான சார்பு (one-to-one function) எனக் கூறலாம்.

6·3·3. மேற்படவாக்கம் (Onto Mapping)

வீச்சிலுள்ள ஒவ்வோர் உறுப்பும் அரங்கத்திலுள்ள ஓர் உறுப்பின் நிழலாக இருப்பின் அச்சார்புக்கு மேற்படவாக்கம் என்பது பெயர்.

வீச்சிலுள்ள ஓர் உறுப்புக்காவது மேற்கூறப்பட்டது உண்மையில்லையாயின் சார்பிற்கு உள்படவாக்கம் (into mapping) என்பது பெயர்.

f என்பது ஒன்றுக்கொன்று மேற்படவாக்கமாக இருப்பின் D-க்கும் R-க்கும் ஒன்றுக்கொன்றான ஒத்தியைபு (one to one correspondence) இருக்கிறது என்று சொல்லப்படும்.

6·3·4. f என்பது ஒன்றுக்கொன்றான சார்பாக இருக்கட்டும். y என்பது x-ன் நிழலானால் x ஐ y-ன் மூலம் என்று குறிப்பிடலாம். f என்பது ஒன்றுக்கொன்றான சார்பாதலால் ஒவ்வொரு y-க்கும் ஒன்றுக்கு மேற்பட்ட மூலம் இருக்கமுடியாது.

இயற்கணித முறைப்படி இதையே,

$f(x) = y \Longrightarrow x = f^{-1}(y)$ என்று குறிப்பிடலாம்.

f^{-1} என்னும் சார்புக்கு R அரங்கமாகவும் D என்பது வீச்சாகவும் இருக்கும் என்பது தெளிவாகத் தெரிகிறது. f^{-1}ஐ f-ன் நேர்மாறான சார்பு (Inverse of f) என்று கூறலாம்.

6·3·5. தேற்றம்

$X, Y \in R$

(அ) $f^{-1}(X \cup Y) = f^{-1}(X) \cup f^{-1}(Y)$

(ஆ) $f^{-1}(X \cap Y) = f^{-1}(X) \cap f^{-1}(Y)$

(இ) $f^{-1}\left[\bigcup_{i \in I} X_i\right] = \bigcup_{i \in I} f^{-1}(X_i)$

(ஈ) $f^{-1}\left[\bigcap_{i \in I} X_i\right] = \bigcap_{i \in I} f^{-1}(X_i)$

(உ) $X \subset Y \Longrightarrow f^{-1}(X) \subset f^{-1}(Y)$

(ஊ) $f^{-1}(X - Y) = f^{-1}(X) - f^{-1}(Y)$

நிறுவல்

f^{-1} என்னும் சார்பின் வரைவிலக்கணத்திலிருந்து இவை தெளிவாகத் தெரிகின்றன.

6·3·6. ஒரு கணத்தின் குறியீட்டுச் சார்பு (Index Function or Characteristic Function of a Set).

A என்பது Ω-ன் ஒரு கணமாக இருக்கட்டும். X_A என்னும் சார்பு பின்வருமாறு வரையறை செய்யப்படுகிறது.

$$X_A(x) = 1, x \in A$$

$$X_A(x) = 0, x \in A^c$$

அதாவது A-ன் குறியீட்டுச் சார்பு A-லுள்ள உறுப்புகளின் மேல் ஒன்றையும் A-லில்லாத உறுப்புகளின் மேல் 0 ஐயும் எடுத்துக் கொள்கிறது. X_A-ன் அரங்கம் Ω. வீச்சு $[0, 1]$.

6·3·7. வீச்சு $(-\infty, \infty)$ என்னும் மெய்யெண் இடைவெளியானால் f ஐ மெய்ச்சார்பு (Real function) எனக் கூறலாம்.

X_A என்பது ஒரு மெய்ச்சார்பு.

6·3·8. எளிய சார்பு (Simple Function)

F என்பது ஒரு ஸிக்மா கணக்களமாக இருக்கட்டும். $A_1, A_2, \ldots$ என்பவை F-ன் உறுப்புகளாக இருக்கட்டும்.

$$f = \sum_{i=1}^{n} c_i \, X_{A_i}$$ என்பது ஒரு எளிய சார்பாகவும்

[$c_i \in (-\infty, \infty)$; n-முடிவுள்ளது].

துணை முடிவு

F என்பது ஒரு நிகழ்ச்சிகளின் σ-கணக்களமானல் இதில் நிறுவப்படும் எளிய சார்புகளுக்கு எளிய ராண்டம் மாறி என்பது பெயர் (Simple Random Variable).

6·3·9. $f_1, f_2, \ldots$ என்பது ஒரு எளிய சார்புகளாலான ஒருங்கும் முடிவில் தொடர் (Converging infinite sequence of simple functions) எனக் கொள்வோம். f என்பது இம் முடிவில் தொடரின் எல்லை (limit) யாக இருக்கட்டும். இத்தகைய f ஐ அளக்கத்தகு சார்பு (Measurable function) எனக் கூறலாம்.

6·3·10. f என்பது அளக்கத்தகு சார்பாக இருப்பதற்கு $(-\infty, a)$ ன் மூலங்கள் [$a \in (-\infty, \infty)$]. F-ல் இருப்பது வேண்டியதும் போதுமானதுமான நிபந்தனையாகும்.

அதாவது, f அளக்கத்தகு சார்பு $\Longleftrightarrow f^{-1}(-\infty, a) \in F \; \forall a \in (-\infty, \infty)$ [$\forall a$ என்பதை ஒவ்வொரு a-க்கும் எனப் படிக்கவும்.]

துணைமுடிவு

$f_1, f_2, \ldots f_\infty$ என்பவை அளக்கத்தகு சார்பானல் $\sum_{i=1}^{n} f_i$, $\sum_{i=1}^{n} \alpha_i f_i$ [$\alpha_i \in (-\infty, \infty)$], $\lim_{n \to \infty} f_n$ [எல்லையிருப்பின்], $G(f_1, f_2, \ldots, f_n)$ [G என்பது தொடர் சார்பாக இருப்பின்] முதலியவைகளும் அளக்கத் தகுந்தவை.

7. நிகழ்திறம்

(PROBABILITY)

Ω என்பது ஒரு மாதிரிவெளி; F என்பது Ω-லுள்ள ஒரு σ-நிகழ்ச்சிக் கணக்களம். P என்னும் மெய்ச்சார்பு F-லுள்ள கணங்களின்மேல் வரையறுக்கப்படுகிறது. அதாவது F-லுள்ள ஒவ்வொரு கணத்திற்கும் ஒரு மெய்யெண்ணை நிழலாகக் கொண்ட சார்பே P ஆகும். P என்னும் கணச்சார்பு (Set function) நிகழ்திறம் என்று கூறப்படுவதற்குப் பின்வரும் நிபந்தனைகள் உள்ளன.

7.1. (அ) $P(A) \geq 0 \quad A \in F$

(ஆ) $$P\left[\bigcup_{i=1}^{\infty} A_i\right] = \sum_{i=1}^{\infty} P(A_i) \quad A_i \in F \; i=1, 2, \dots \infty$$
$$A_i \cap A_j = \phi \; [i \neq j]$$

[அதாவது A_i-கள் ஒன்றையொன்று விலக்கும்போது]

(இ) $P(\Omega) = 1$.

இத்தகைய நிபந்தனைகளை ஏற்கும் P எனும் கணச்சார்பை நிகழ்திறம் என்று கூறலாம்.

7.2. நிகழ்திற வெளி (Prbability Space)

Ω என்பது மாதிரி வெளியாகவும், F என்பது அதிலுள்ள ஒரு σ-நிகழ்ச்சிக் கணக்களமாகவும், P என்னும் கணச்சார்பு (அ), (ஆ), (இ)-இவற்றைத் திருப்திப்படுத்தும் நிகழ்திறமாகவும் ஆனால் (Ω, F, P) என்னும் மும்மை (Triplet) நிகழ்திற வெளி எனப்படும்.

தேற்றம் 7·2·1.

$$A \in F \Longrightarrow P(A) + P(A^c) = 1.$$

நிறுவல்

$$P[A \cup A^c] = P(A) + P(A^c) = P(\Omega) \quad [\text{(இ)-யினால்}]$$

துணைமுடிவு

$P(\Omega) = 1 \Longrightarrow P(\phi) = 0$ $\because P(\Omega \cup \phi) = P(\Omega) = P(\Omega) + P(\phi) \Longrightarrow P(\phi) = P(\Omega) - P(\Omega) = 1-1=0.$

தேற்றம் 7·2·2.

$$A, B \in F, A \subset B \Longrightarrow P(A) \leq P(B)$$

நிறுவல்

$$A \subset B \Longrightarrow B = A \cup (B - A) \Longrightarrow P(B) = P(A) + P(B - A)$$

$$[\because A \cap (B - A) = \phi] \Longrightarrow P(B) \geq P(A)$$

தேற்றம் 7·2·3.

$$A, B \in F \Longrightarrow P(A \cup B) = P(A) + P(B) - P(A \cap B)$$

நிறுவல்

$$A \cup B = [A - A \cap B] \cup [B - A \cap B] \cup [A \cap B]$$

$$\therefore P[A \cup B] = P[A - A \cap B] + P[B - A \cap B] + P[A \cap B]$$

[$\because$ $[A - A \cap B], [B - A \cap B], [A \cap B]$ என்பவை ஒன்றையொன்று விலக்கும் நிகழ்ச்சிகள்].

$$\therefore P[A \cup B] = [P(A) - P(A \cap B)] + [P[B] - P(A \cap B)] + P[A \cap B] \quad [\because 7{\cdot}2{\cdot}2]$$

$$= P(A) + P(B) - P(A \cap B)$$

தேற்றம் 7·2·4.

$$P\left[\bigcup_{i=1}^{n} A_i\right] = \sum_{i=1}^{n} P(A_i) - \sum_{i>j=1}^{n} P(A_i \cap A_j)$$

$$+ \sum_{i>j>k=1}^{n} P(A_i \cap A_j \cap A_k) - \cdots\cdots + (-1)^{n-1} P\left[\bigcap_{i=1}^{n} A_i\right]$$

நிறுவல்

இத்தேற்றம் $n = 2$-க்கு முன்னமேயே நிறுவப்பட்டுவிட்டது. இத்தேற்றம் $n = 1, 2, 3, \ldots m$ என்னும் எண்வரைச் சரியெனக் கொண்டால்

$$P\left[\bigcup_{i=1}^{m} A_i\right] = \sum_{i=1}^{m} P(A_i) - \sum_{i>j=1}^{m} P\left[A_i \cap A_j\right]$$

$$+ \cdots\cdots\cdots + (-1)^{m-1} P\left[\bigcap_{i=1}^{m} A_i\right]$$

A_{m+1} என்ற புதிய நிகழ்ச்சியை $\bigcup_{i=1}^{n} A_i$ உடன் சேர்த்துக் கொண்டால்,

$$P\left[\bigcap_{i=1}^{m+1} A_i\right] = P\left\{A_{m+1} \cup \left[\bigcup_{i=1}^{m} A_i\right]\right\}$$

$$= P(A_{m+1}) + P\left[\bigcup_{i=1}^{m} A_i\right] - P\left[A_{m+1} \cap \bigcup_{i=1}^{m} A_i\right]$$

$$= P(A_{m+1}) + P\left[\bigcup_{i=1}^{m} A_i\right] - P\left[\bigcup_{i=1}^{m} (A_i \cap A_{m+1})\right]$$

$$= P(A_{m+1}) + \sum_{i=1}^{m} P(A_i) - \sum_{i>j=1}^{m} P(A_i \cap A_j) + \ldots$$

$$\cdots + (-1)^{m-1} P\left[\bigcap_{i=1}^{m} A_i\right] - \left[\sum_{i=1}^{m} P[A_i \cap A_{m+1}]\right.$$

$$- \sum_{i>j=1}^{n} P[A_i \cap A_j \cap A_{m+1}) + \cdots + (-1)^{m-1} P \left.\bigcap_{i=1}^{1+m} A_i\right]$$

$$= \sum_{i=1}^{m+1} P(A_i) - \sum_{i>j=1}^{m+1} P[A_i \cap A_j] + \cdots$$

$$+ (-1)^m P\left[\bigcap_{i=1}^{m+1} A_i\right]$$

இதிலிருந்து தேற்றம் $n = m + 1$-க்கும் சரியென்று தெரிகிறது. தொகுத்தறி முறை (Induction) மூலமாக இத்தேற்றம் எல்லா n-க்கும் சரியென்பது தெரிகிறது.

தேற்றம் 7·2·5.

$A_1, A_2, A_3, \ldots$ என்பது F-விலுள்ள ஒரேமுறை ஏறும் அல்லது இயங்கும் கணத்தொடரானல்

$$\lim_{n\to\infty} P(A_n) = P(\lim_{n\to\infty} A_n)$$

நிறுவல்

முதலில் $A_1, A_2, \ldots$ என்பவை ஏறுவரிசையிலுள்ளன எனக் கொள்வோம். அதாவது

$$A_1 \subset A_2 \subset A_3 \subset A_4 \cdots \subset A_n \subset A_{n+1} \cdots \cdots$$

இதிலிருந்து $A_n = A_1 \cup (A_2 - A_1) \cup (A_3 - A_2)$ $\cdots\cdots \cup (A_n - A_{n-1})$ என்பது தெரிகிறது.

$$\therefore \lim_{n\to\infty} A_n = A_1 \cup \left[\bigcup_{n=2}^{\infty} (A_n - A_{n-1})\right]$$

$$P\left[\lim_{n\to\infty} A_n\right] = P(A_1) + \sum_{n=2}^{\infty} P(A_n - A_{n-1})$$

[7·1 (ஆ) வினல்]

$$= \lim_{n\to\infty} \left\{ P(A_1) + \sum_{n=2}^{n} P(A_n - A_{n-1}) \right\}$$

$$= \lim_{n\to\infty} \Big\{ P(A_1) + P(A_2) - P(A_1)$$

$$+ P(A_3) - P(A_2) + \cdots + P(A_n) - P(A_{n-1}) \Big\}$$

$$\therefore \; P\left[\lim_{n\to\infty} A_n \right] = \lim_{n\to\infty} \left[P(A_n) \right]$$

$A_1, A_2, A_3, \ldots$ என்பவை இறங்குவரிசையிலிருப்பின்,

$A_1{}^c, A_2{}^c, \ldots$ என்பவை ஏறுவரிசையிலிருக்கும்.

$$\therefore \; P\left[\lim_{n\to\infty} A_n{}^c \right] = \lim_{n\to\infty} P(A^c{}_n).$$

$$\therefore \; P\left[\lim_{n\to\infty} (\Omega - A_n) \right] = \lim_{n\to\infty} P(\Omega - A_n).$$

$$= 1 - \lim_{n\to\infty} P(A_n)$$

$$\therefore \; P\left[\Omega - \lim_{n\to\infty} A_n \right] = 1 - \lim_{n\to\infty} P(A_n)$$

$$\therefore \; 1 - P\left[\lim_{n\to\infty} A_n \right] = 1 - \lim_{n\to\infty} P(A_n).$$

$$\therefore \; P\left[\lim_{n\to\infty} A_n \right] = \lim_{n\to\infty} P(A_n)$$

தேற்றம் 7·2·6.

$A, A_1, A_2, A_3, \ldots\ldots \in F$ and $A \subset \bigcup_{i=1}^{\infty} A_i$

$$\Longrightarrow P\,(A) \leq \sum_{i=1}^{\infty} P\,(A_i)$$

நிறுவல்

$$A \subset \bigcup_{i=1}^{\infty} A_i \Longrightarrow A \cap \left[\bigcup_{i=1}^{\infty} A_i\right] = A.$$

$$\therefore \quad P\,(A) = P\left\{A \cap \left[\bigcup_{i=1}^{\infty} A_i\right]\right\} = P\left\{\bigcup_{i=1}^{\infty}\left[A_i \cap A\right]\right\}$$

$$= P\,\{\,(A_1 \cup A) \cup (A_2 - A_1 \cup A) \cup [A_3 - (A_1 \cup A_2 \cap A)] \cdots\,\}$$

$$= P\,(A_1 \cap A) + P\,(A_2 - A_1 \cap A) + \cdots$$

$$\leq P\,(A_1) + P\,(A_2) + \cdots = \sum_{i=1}^{\infty} P\,(A_i)$$

8. சார்பிலா நிகழ்ச்சிகள்

(INDEPENDENT EVENTS)

8·1. பெருக்கல் கணங்கள் (Product Sets)

A, B என்று இரு கணங்கள் இருந்தால் (x, y) $[x \in A, y \in B]$ என்னும் இருமைகள் கொண்ட கணத்திற்குப் பெருக்கல் கணம் என்பது பெயர். இதை $A \times B$ என்று குறிப்பிடலாம். பொதுவாக $(x_1, x_2, \ldots x_n)$ என்னும் n உறுப்புகளால் $[x_i \in A_i]$ அமைக்கப் பட்ட கணத்தை $A_1 \times A_2 \times \ldots \times A_n$ அல்லது $\prod_{i=1}^{n} A_i$ என்று குறிப் பிடலாம்.

8.1·2 Ω_1, Ω_2 என்பவை இரு மாதிரி வெளிகளாக இருக் கட்டும். Ω_1-ல் F_1 என்ற σ-கணக்களத்தை நிகழ்ச்சிக் களமாகவும், Ω_2-ல் F_2 என்ற σ-கணக்களத்தை நிகழ்ச்சிக் களமாகவும் எடுத்துக் கொண்டால் $\Omega = \Omega_1 \times \Omega_2$-ல் $F = F_1 \times F_2$ என்பது ஒரு σ-கணக்களமாக இருக்கும். இதை Ω-ன் நிகழ்ச்சிக் களமாகக் கொள்ளலாம்.

F-லுள்ள உறுப்புகள் $(A \times B)$ $A \in F_1$, $B \in F_2$.

F_1-ல் P_1 என்றும் F_2-ல் P_2 என்றும் நிகழ் திறங்களை எடுத்துக் கொண்டால் F-ல் பெருக்கல் நிகழ் திறத்தைப் பின் வருமாறு வரையறுக்கலாம்.

$P(A \times B) = P_1(A) \cdot P_2(B)$. இப்படி வரையறுத்த பின்னர் $P = P_1 \times P_2$ என்று குறிப்பிடலாம். மேலும் F_1, F_2-லுள்ள நிகழ்ச்சிகளைச் சார்பிலா நிகழ்ச்சிகள் என்றும் குறிப்பிடலாம்.

(Ω_1, F_1, P_1), (Ω_2, F_2, P_2) இவைகளை (Ω, F, P)-ன் அங்கங்களாகக் (Components) கொள்ளலாம்.

9. ராண்டம் மாறிகள்

9·1·1. (Ω, F, P) என்பது ஒரு நிகழ்திற வெளியாக இருக்கட்டும். இவ்வெளியில் f என்பது ஓர் அளக்கத்தகு சார்பாக இருந்தால் அதை ராண்டம் மாறி (Random Variable) எனலாம்.

முன்பே குறிப்பிட்டது போல் $f^{-1} [(-\infty, a)] \in F \quad \forall a \in (-\infty, \infty)$ அளக்கத்தகு சார்புகளின் பண்புகளிலிருந்து பின்வரும் தேற்றம் தெளிவாகிறது.

9·1·2. தேற்றம்

$f_1, f_2, \ldots$ என்பவை ராண்டம் மாறிகளானல்

(அ) $\sum_{i=1}^{n} \alpha_i f_i \quad [\alpha_i \in (-\infty, \infty), \quad i = 1, 2, \ldots n]$

(ஆ) $G(f_1, f_2, \ldots f_n)$ [G என்பது தொடர் சார்பானல்]

(இ) $\lim_{n \to \infty} f_n$ [$\{ f_n \}$ என்னும் தொடர் ஒருங்கும் எனக் கொண்டால்]

முதலியவைகளும் ராண்டம் மாறிகளாகும்.

9.1·3. ராண்டம் மாறிகளின் பரவல் சார்பு (Distribution Function of Random Variables)

$[\Omega, F, P]$ என்பது ஒரு நிகழ்திறவெளியென்றும் $f(x)$ என்பது ஒரு ராண்டம் மாறி எனவும் கொள்வோம்.

முன்பே கூறியபடி $f^{-1} (-\infty, x)$ என்பது F-ல் உள்ள ஒரு கணமாகும். ஆகவே அதற்கு ஒரு நிகழ்திறம் இருக்கும். இந்த நிகழ்திறத்தை $F(x)$ எனக் குறிப்பிடுவோம்.

அதாவது $P[f^{-1}(\infty, x]] = F(x)$;

$F(x)$-க்கு f-ன் பரவல் சார்பு என்பது பெயர்.

9·1·4. பரவல் சார்பின் பண்புகள்

(அ) $F(-\infty) = 0, \ F(+\infty) = 1$

(ஆ) $0 \le F(x) \le 1$

(இ) $x \le y \Longrightarrow F(x) \le F(y)$.

நிறுவல்

(அ) $F(-\infty) = P[f^{-1}(-\infty \ \infty]] = P[\phi] = 0$

$F(+\infty) = P[f^{-1}(-\infty, \infty]] = P[\Omega] = 1.$

(ஆ) $F(x)$ என்பது நிகழ்திறம் ஆதலால் $0 \le F(x) \le 1$

(இ) $x \le y \Longrightarrow f^{-1}(-\infty, x] \subset f^{-1}(-\infty, y] \Longrightarrow$
$P[f^{-1}(-\infty \ x]] \le P[f^{-1}(-\infty, y]] \Longrightarrow F(x) \le F(y)$.

10. எளிய சார்புகளைத் தொகையிடல்

(INTEGRATION OF SIMPLE FUNCTIONS)

10·1·1.

$f(x) = \sum_{i=1}^{n} C_i X_{A_i}$ என்பது ஒரு எளிய சார்பாக இருக்கட்டும்.

$A_i \in F$. (Ω, F, P) என்பது நிகழ்திற வெளி.

Ω-ல் $f(x)$-ன் தொகையைப் (Integral) பின்வருமாறு நிறுவலாம்.

$$\int_{\Omega} f(x)\, dP = \sum_{i=1}^{n} C_i\, P(A_i).$$

இதிலிருந்து பின்வரும் தேற்றம் தெளிவாகிறது.

தேற்றம் 10·1·2.

f, g என்பவை எளிய சார்புகளானால்

(அ) $\int_{\Omega} (\alpha f + \beta g)\, dP = \alpha \int_{\Omega} f\, dP + \beta \int_{\Omega} g\, dP$

$(\alpha, \beta \in (-\infty, \infty))$

(ஆ) $f \leq g \Longrightarrow \int_{\Omega} f\, dP \leq \int_{\Omega} g\, dP.$

10·1·3.

A எனும் F-லுள்ள கணத்தில் $f(x)$ என்ற எளிய சார்பின் தொகையைப் பின்வருமாறு வரையறுக்கலாம்.

$$\int_A f(x)\, dP = \int_\Omega f\, X^A\, dP.$$

இதிலிருந்து பின்வரும் தேற்றம் தெளிவாகிறது.

10·1·4.

(அ) $\int_{A \cup B} f\, dP = \int_A f\, dP + \int_B f\, dP$ [$A \cap B = \phi$ என்றால்]

(ஆ) $\int_A (\alpha f + \beta g)\, dP = \alpha \int_A f\, dP + \beta \int_A g\, dP.$

(இ) $f \leq g \Longrightarrow \int_A f\, dP \leq \int_A g\, dP.$

10·1·5.

$f_1, f_2, \ldots$ என்பது எளிய சார்புகளாலான ஒரு தொடரெனவும், $f_n \leq K$ எனவும், இத்தொடர் f எனும் அளக்கத்தகு சார்புக்கு ஒருங்குகிறது எனவும் கொண்டால் f-ன் தொகையைப் பின்வருமாறு வரையறுக்கலாம்.

$$\int_\Omega f\, dP = \lim_{n \to \infty} \int f_n\, dP.$$

இதன் மூலம் எல்லா அளக்கத்தகு சார்புக்கும் தொகையை வரையறுக்கலாம். ஆனால் சில அளக்கத்தகு சார்புகளுக்குத் தொகை முடிவில்லாததாகப் போய்விடலாம். முடிவில்லாத தொகையையுடைய அளக்கத்தகு சார்புகளுக்குத் தொகையிட முடியாத சார்புகள் (Non-integrable functions) என்பது பெயர். தொகை முடிவுள்ளதாக இருப்பின் அச்சார்புக்குத் தொகையிட முடியும் சார்பு (Integrable functions) என்பது பெயர். இதிலிருந்து பின்வரும் தேற்றம் தெளிவாகிறது.

தேற்றம் 10·1·6.

f, g என்பவை தொகையிட முடியும் சார்புகளானுல்

(அ) $|f|$, $|g|$ என்பவை தொகையிட முடியும் சார்புகள்.

(ஆ) $\alpha f + b g$ தொகையிட முடியும் சார்பு

(இ) $\int f \, dP \le \int |f| \, dP.$

Ω

(ஈ) $\int_{\Omega} (\alpha f + bg) dP = \alpha \int_{\Omega} f \, dP + \beta \int_{\Omega} g \, dP.$

11. ஊக்குவிக்கப்பட்ட நிகழ்திற வெளி

(INDUCED PROBABILITY SPACE)

11·1·1.

$[\Omega, F, P]$ என்பது ஒரு நிகழ்திற வெளியாக இருக்கட்டும், இவ்வெளியில் f என்பது ஒரு ராண்டம் மாறியாக இருக்கட்டும். B என்பது மெய்யெண் இடைவெளியிலுள்ள போரல் களமானுல், B-லுள்ள ஒவ்வொரு கணத்தின் மூலமும் [f ஐப் பொருத்த வரை] F-ல் இருக்கவேண்டும்.

அதாவது $A \in B \Longrightarrow f^{-1}(A) \in F$.

ஆகவே B-லுள்ள ஒவ்வொரு கணத்திற்கும் ஒரு நிகழ்திறத்தைப் பின்வருமாறு வரையறுக்கலாம்.

307391

$P_1(A) = P\,[\overset{-1}{F}(A)]$. இப்படி நிறுவுவதால் ஒரு புதிய நிகழ்திற வெளி ஊக்குவிக்கப்படுகிறது. இது யாதெனில்

$$[\Omega_1, B, P_1]. \quad \Omega_1 = (-\infty, \infty), = R.$$

அதாவது Ω எனும் வெளியில் வரையறுக்கப்பட்ட நிகழ்திறம் மெய்யெண் இடைவெளியான $(-\infty, \infty) = R$-ல் ஒரு நிகழ்திறத்தை ஊக்குவிக்கிறது. R-ல் உள்ள σ கணக்களம் போரல் கணக்களமான B ஆகும்.

519.52

BAL;1

தேற்றம் 11·1·2.

(R, B, P_1) என்பது (Ω, F, P)-ல் f ஆல் ஊக்குவிக்கப்பட்ட நிகழ்திற வெளியென்றும், $G(f)$ என்பது (Ω, F, P)-ல் உள்ள ஒரு தொகையிடக்கூடிய அளத்தகு சார்பு என்றும், எடுத்துக் கொண்டால்,

$$\int_{\Omega} G(f)\, dP = \int_{-\infty}^{\infty} G(f)\, dP_1$$

F என்பது f-ன் பரவல் சார்பாகுல் முன்பே பார்த்தபடி $P_1 = F$.

$$\text{ஆகவே} \int_{\Omega} G(f)\, dP = \int_{-\infty}^{\infty} G(f)\, dF$$

மேற்கூறிய தேற்றத்திற்கு நிறுவல் வேண்டுவோர் Probability Theory—Michel Loeve என்ற நூலைப் பார்க்கவும். இத்தேற்றத்தின் மூலம் Ω-லுள்ள ஒவ்வொரு தொகையையும் $R[-\infty, \infty]$ லுள்ள தொகையிடலாகக் கொள்ளலாம்.

R-ல் கூறப்பட்ட தொகையிடலுக்கு லெபெக்-ஸ்டீல்ட்யெஸ் (Lobesque-Stieltjes) தொகையிடலென்பது பெயர்.

11·1·3.

$F(x)$ என்பது ஒரு வகையிடக்கூடிய சார்பு (differentiable function) எனவும், $f(x)$ என்பது $F(x)$-ன் வகையீட்டுக் கெழு (differential coefficient) எனவும் எடுத்துக்கொண்டால்

$$\int_{-\infty}^{\infty} g(x)\, dF(x) = \int_{-\infty}^{\infty} g(x) f(x)\, dx \quad \text{என எடுத்துக்}$$

கொள்ளலாம்.

வலப்புறமுள்ள தொகையிடல் சாதாரணமாக உபயோகப்படுத்தப்படும் தொகையிடலாகும்.

துணை முடிவு

(அ) $$\frac{dF(x)}{dx} = f(x), \quad F(\infty) = 1 \implies \int_{-\infty}^{\infty} f(x)dx = 1$$

(ஆ) $$F(b) - F(a) = \int_{a}^{b} f(x)\, dx.$$

(இ) $$f(x) = \underset{h \to 0}{Lt} \frac{F(x+h) - F(x)}{h} \implies f(x) \geq 0$$

[$\because$ $F(x+h) - F(x) > 0,\ h > 0$ எனக் கொண்டால்].

11·1·4.

$f(x)$ என்பது X எனும் ராண்டம் மாறியின் ஊக அளவு அடர்த்தி (Probability Density) என்பது பெயர். [X-ன் பரவல் சார்பு $F(x)$ என எடுத்துக்கொண்டால்.]

1·1·5· ராண்டம் மாறியின் எதிர்பார்க்கும் மதிப்பு (Expected value of a random variable)

X எனும் ராண்டம் மாறியின் ஊக அளவு அடர்த்தி $f(x)$ என்றால் $\int_{-\infty}^{\infty} x f(x)\, dx$ என்பதற்கு X-ன் எதிர்பார்க்கும் மதிப்பு என்பது பெயர். இதை $E(X)$ என்று குறிப்பிடலாம்.

[$\int_{-\infty}^{\infty} |x| f(x)\, dx$ தொகையிட முடியாதென்றால் $E(x)$ இல்லையென்று கூறவேண்டும்.]

X என்பது முடிவுள்ள அல்லது எண்ணிடத் தக்க மதிப்புகளையே கொண்டிருந்தால் $F(x)$ என்பது படிச்சார்பு (step function) எனப்படும். அதாவது $F(x)$-ம் முடிவுள்ள அல்லது எண்ணிடத்தக்க மதிப்புகளையே கொள்ளும். [$x_1, x_2, \ldots$] என்பவை X-ன் வெவ்வேறு மதிப்புகள் [$x_{i+1} > x_i$, $i = 1, 2 \ldots$] எனக்கொண்டால் $F(x_1)$, $F(x_2)$, .. என்பவை படிப்படியாக ஏறும் (increase monotonically). மேலும் $F(x)$ எனும் சார்பு ($x_1, x_2, \ldots$) எனும் புள்ளிகளில் தொடர்பற்று (discontinuous) இருக்கும். அதாவது x_i எனும் புள்ளியை $F(x)$ கடக்கும்போது அதன் மதிப்பு திடீரென்று அதிகரிக்கும். x_i என்ற புள்ளியில் $F(x)$-ன் அதிகரிப்பையே $f(x_i)$ எனக் கொள்ளவேண்டும்.

இப்படிப்பட்ட ராண்டம் மாறியின் எதிர்பார்க்கும் மதிப்பு

$$E(X) = \int_{-\infty}^{\infty} x\,dF(x) = \sum_{i=1}^{n} x_i f(x_i)$$

[n முடிவுள்ளது அல்லது ∞].

மேற்கூறப்பட்டது லெபெக் ஸ்டீல்ட்யெஸ் தொகையிடலின் சிறப்பாகும். அதாவது சாதாரணக்கூட்டல் தொகையைக்கூட, தொகையிடல் (Integration) மூலமாகக் குறிப்பிடலாம்.

11·1·6. X என்பது ஒரு ராண்டம் மாறியானல் $g(X)$ என்னும் ஓர் அளக்கத்தகு சார்பும் ராண்டம் மாறியாகும். $g(X)$-ன் எதிர் பார்க்கும் மதிப்பு

$$E[g(X)] = \int_{-\infty}^{\infty} g(x)\, d F(x)$$

[$\int_{-\infty}^{\infty} | g(x) |\, dF(x)$ தொகையிட முடியவில்லையென்ருல் $g(X)$-க்கு எதிர்பார்க்கும் மதிப்பு கிடையாது என்று குறிப்பிடப் படவேண்டும்.]

தேற்றம் 11·1·7.

X, Y என்பன இரண்டு ராண்டம் மாறிகளானல்

$E(X+Y) = E(X+Y)$.

துணை முடிவு

$X_1, X_2, \ldots \ldots, X_n$ [n முடிவுள்ளது] என்பவை ராண்டம் மாறிகளானல்,

$$E\left[\sum_{i=1}^{n} X_i\right] = \sum_{i=1}^{n} E(X_i)$$

தேற்றம் 11.1.8.

X, Y என்பன சார்பற்ற ராண்டம் மாறிகள் (Independent random variables) என்ருல் $E(XY) = E(X)\,E(Y)$.

துணை முடிவு

$X_1, X_2, \ldots X_n$ என்பவை சார்பற்ற ராண்டம் மாறிகளானல்

$$E(X_1\, X_2 \ldots X_n) = \pi_{i=1}^{n}\; E(X_i).$$

11·1.9. ராண்டம் மாறியின் திருப்புத் திறன்கள் (Moments of random variable)

(அ) X எனும் ராண்டம் மாறியின் a ஐச் சுற்றிய r ஆவது திருப்புத்திறன் என்பது $E(X-a)^r$ ஆகும். இதை μ'_r என்று குறிப்பிடலாம்.

(ஆ) $E(X)$-க்கு X-ன் சராசரியென்றும் பெயர். இதை $\bar{X}$ என்று குறிப்பிடலாம்.

(இ) X-ன் r ஆவது மையத்திருப்புத் திறன் (r th central moment) என்பது $E(X-\bar{X})^r$ ஆகும். இதை μ_r என்று குறிப்பிடலாம்.

(ஈ) $\mu_2 = E\,(X-\bar{X})^2$ என்பது X-ன் பரவற் படி (variance of X) என்று குறிப்பிடப்படும். இதை $V(X)$ என்றும் குறிப்பிடலாம். $\sigma_x{}^2$ என்றும் இது குறிப்பிடப்படுவதுண்டு.

(உ) $\sqrt{\mu_2} = \sqrt{E(X-\bar{X})^2} = \sqrt{V(X)} = \sigma_x$ என்பது X-ன் திட்டவிலக்கம் (Standard deviation) என்று குறிப்பிடப்படும்.

(ஊ) X, Y என்பவை இரு ராண்டம் மாறிகளானல் $E(X-\bar{X})\;(Y-\bar{Y})$ என்பது (X, Y)-ன் உடன்மாற்றம் [covariance of (X, Y)] என்று குறிப்பிடப்படும். இது cov (X, Y) அல்லது $\sigma_{x,\,y}$ என்றும் குறிப்பிடப்படும்.

(எ) $X; Y$ என்பவை சார்பற்ற ராண்டம் மாறிகளானல் cov $(X, Y) = 0$.

(ஏ) X, Y என்பவை இரு ராண்டம் மாறிகளானல்

$$\rho_{x.\,Y} = \frac{\text{cov}\,(X, Y)}{\sqrt{V(X)\,V(Y)}}$$ என்பது (X, Y)-ன் உடன் தொடர்புக் கெழு (correlation coefficient) எனப்படும்.

(ஐ) $\rho_{x,\,Y} = 0$ என்ருல் X-ம் Y-ம் உடன்தொடர் பற்றவை (uncorrelated) எனக் கூறப்படும்.

12. சில பரவல்கள்

(SOME DISTRIBUTION)

12·1. ஈருறுப்புப் பரவல் (Binomial Distribution)

ஒரு பரிசோதனையில் வெற்றி, தோல்வியெனும் இரு விளைவுகளே நிகழக்கூடியவையென்றும், வெற்றியின் நிகழ்திறம் P எனவும், தோல்வியின் நிகழ்திறம் $Q = 1 - P$ எனவும் எடுத்துக் கொள்வோம். இப் பரிசோதனையை n முறை தனித்தனியாகச் சார்பற்றவாறு நடத்தினால் இவைகளில் கிடைக்கும் வெற்றியின் எண்ணிக்கை 'X' என்பது ஒரு ராண்டம் மாறியாகும். இம் மாறியின் பரவல் பின் வருமாறு :

$$P[X = r] = \binom{n}{r} p^r q^{n-r} \quad r = 0, 1, 2, \ldots n.$$ இதையே

பின்வருமாறு எழுதலாம்.

ராண்டம் மாறி	$X = 0$	1	2	$3 \ldots$	r	$\ldots n$
நிகழ்திறம்	$P = q^n$	$n\,pq^{n-1}$	$\binom{n}{2} p^2 q^{n-2}$		$\binom{n}{r} p^r q^{n-r}$	$\ldots p^n$

இப் பரவலுக்கு ஈருறுப்புப் பரவல் என்பது பெயர். இதை $B(n, p)$ எனக் குறிப்பிடலாம்.

பின் வருவனவற்றை எளிதாக நிறுவலாம்.

12·2. (அ) $$E(X) = \sum_{r=0}^{n} r \binom{n}{r} p^r q^{n-r} = np.$$

ஆகவே ஈருறுப்புப் பரவல் $B(n, p)$-ன் சராசரி np ஆகும்.

$$(\text{ஆ}) \quad V(X) = \sum_{r=0}^{n} (r - np)^2 \binom{n}{r} p^r q^{n-r} = npq$$

ஆகவே ஈருறுப்புப் பரவல் $B(n, p)$-ன் பரவற் படி npq ஆகும்.

12·3. பாய்ஸான் பரவல் (Poisson Distribution)

ஈருறுப்புப் பரவலில் n மிக அதிகமாகவும், p மிகவும் சிறியதாகவும், np என்பது λ எனும் எண்ணை எல்லையாக உடையதாகவும் இருப்பின் இப்பரவல் பின்வரும் பரவலை எல்லையாகக் கொண்டிருக்கும்.

$$X = \quad 0 \qquad 1 \qquad 2 \quad \ldots \quad r \quad \infty$$

$$P = e^{-\lambda},\ e^{-\lambda}\lambda,\ e^{-\lambda}\frac{\lambda^2}{2},\ \ldots,\ e^{-\lambda}\frac{\lambda^r}{\angle r},$$

அதாவது $P[X = r] = e^{-\lambda}\dfrac{\lambda^r}{\angle r}$; $r = 0, 1, 2 \ldots \infty$.

பின்வருவனவற்றை எளிதாக நிறுவலாம்.

$$12·4. \quad (\text{அ}) \quad E(X) = \sum_{r=0}^{\infty} r\, e^{-\lambda} \frac{\lambda^r}{\angle r} = \lambda$$

ஆகவே பாய்ஸான் பரவலின் சராசரி λ ஆகும்.

$$(\text{ஆ}) \quad V(X) = \sum_{r=0}^{\infty} (r - \lambda)^2\, e^{-\lambda} \frac{\lambda^r}{\angle r} = \lambda$$

ஆகவே பாய்ஸான் பரவலின் பரவற் படி λ ஆகும்.

12·5. இயல்நிலைப் பரவல் (Normal distribution)

புள்ளியியலில் மிகவும் முக்கியமான இடத்தை வகிக்கும் இப்பரவல் பின்வரும் ஊக அளவு அடர்த்திச் சார்பின் மூலம் வரையறுக்கப்படுகிறது.

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

$$(\text{அ}) \quad E(X) = \int_{-\infty}^{\infty} x \frac{1}{\sqrt{2\pi}\,\sigma} e^{\frac{-(x-\mu)^2}{2\sigma^2}} dx = \mu$$

$$(\text{ஆ}) \quad V(X) = \int_{-\infty}^{\infty} (x-\mu)^2 \frac{1}{\sqrt{2\pi}\,\sigma} e^{\frac{-(x-\mu)^2}{2\sigma^2}} dx = \sigma^2$$

இப்பரவல் $N(\mu, \sigma^2)$ என்று குறிப்பிடலாம்.

கவனிக்கவும்

X என்னும் ராண்டம் மாறிக்கு $B(n, p)$ என்னும் பரவலிருந்தால் $X \cap B(n, p)$ என்று குறிப்பிடலாம்.

$$(\text{இ}) \quad X \cap N(\mu_1 \sigma^8) \Longrightarrow \frac{X-\mu}{\sigma} \cap N(0, 1).$$

$\dfrac{X-\mu}{\sigma}$ என்பது தரப்படுத்தப்பட்ட ராண்டம் மாறி (Standardised random variable) எனப்படும்.

$N(0, 1)$ என்பது தரமான இயல்நிலைப் பரவல் (Standard normal distribution) எனப்படும். இயல்நிலைப் பரவலின் முக்கியத்துவம் பின்வரும் தேற்றத்திலிருந்து புலனாகிறது.

தேற்றம் 12·5·1.

$X_1, X_2, \dots X_n$ என்பவை ஒரே பரவலையுடைய சார்பற்ற ராண்டம் மாறிகள் என்வும், இவைகளின் சராசரி μ, பரவற் படி σ^2 எனவும் கொண்டால் $\dfrac{\frac{1}{n}\sum_{i=1}^{n} X_i - \mu}{\frac{\sigma}{\sqrt{n}}}$ எனும் ராண்டம் மாறி

யின் பரவல் n, ∞ ஐ எட்டும்பொழுது $N(0, 1)$ பரவலை எல்லை யாகக் கொண்டிருக்கும்.

இத் தேற்றத்தின்படி எப்பரவலைக் கொண்ட ராண்டம் மாறி யாக இருந்தாலும், அவை சராசரி μ, பரவற் படி σ^2 கொண்ட ஒரு பரவலைக் கொண்டிருந்தால், அவை சார்பற்றிருக்கும்போது,

$$\frac{1}{n} \sum_{i=1}^{n} X_i$$

-ன் பரவல் n பெரியதாக இருக்கும்போது தோராய மாக $N\left(\mu, \frac{\sigma^2}{n}\right)$ என்று இருக்கும். இப் பண்புதான் பெருங் கூறு பரிசோதனையின் (Large sample test) உயிர்நாடியாகும்.

13. மாதிரி, மாதிரி அளவை, முழுமைத் தொகுதியளவை

(SAMPLE, STATISTIC, AND, PARAMETER)

13·1·1. X என்பது ஒரு ராண்டம் மாறியாக இருக்கட்டும்.

$X_i, i \in I$ என்பவை X-ன் மதிப்புகளாக இருக்கட்டும். I என்பது குறியீட்டுக் கணம். முன்பே குறிப்பிட்டபடி I-ல் முடிவுள்ள உறுப்புகளே இருந்தால் X — 2-க்கு தனித்த ராண்டம் மாறி (Discrete random variable) என்பது பெயர். X_i'கள் அடங்கிய கணத்திற்கு முழுமைத் தொகுதி (population) என்பது பெயர். X_i'கள் ஒரு மெய்யெண் வெளி (a, b)-லுள்ள எல்லா மதிப்புகளையும் ஏற்றுக்கொண்டால் அதற்குத் தொடர்ச்சியான ராண்டம் மாறி (continuous random variable) என்பது பெயர்.

உதாரணங்கள்

(அ) ஈருறுப்புப் பரவல், பாய்ஸான் பரவல் முதலியன தனித்த ராண்டம் மாறிகளின் பரவல்கள் [அல்லது தனித்தப் பரவல்கள்].

(ஆ) இயல்நிலைப் பரவல் ஒரு தொடர்ச்சியான பரவல்.

13·1·2. X எனும் ராண்டம் மாறியின் முழுமைத்தொகுதியிலிருந்து $(x_1, x_2, \dots x_n)$ என்று n உறுப்புகள் எடுக்கப்படுகின்றன என்று எடுத்துக் கொள்வோம். $(x_1, x_2, \dots, x_n)$-க்கு n பருமனையுடைய ஒரு மாதிரியென்பது பெயர். இம்மாதிரியின் பண்பு, இம் மாதிரியை எடுக்கும் முறையைப் பொருத்திருக்கிறது.

முழுமைத் தொகுதியிலுள்ள X_i'-களின் சார்புகளுக்கு முழுமைத் தொகுதியளவை (Parameter) என்பது பெயர்.

உதாரணத்திற்கு முழுமைத் தொகுதியில் N உறுப்புகள் $X_1, X_2, \ldots, X_N$ உள்ளன என்று எடுத்துக் கொண்டால்

$$X_1, \quad \frac{X_1+X_2}{3}, \; \frac{1}{N}\sum_{i=1}^{N} X_i, \; \frac{1}{N}\sum X_i^2, \; \sin X_3,$$

$e^{\frac{X_1}{X_2}}$ முதலியவை முழுமைத் தொகுதியளவைகளாகும்.

மாதிரியிலுள்ள $(x_1, x_2, \ldots x_n)$-களின் சார்புகளுக்கு மாதிரியின் அளவை (Statistic) என்பது பெயர். உதாரணமாக

$$x_1, (x_1 - x_2)^2, \bar{x} = \frac{1}{n}\sum_{i=1}^{n} x_i, \; \frac{1}{n-1}\sum_{i=1}^{n} (x_i - \bar{x})^2, \; \cos x_3,$$

$e^{x_3 + x_4}$ முதலியவைகளுக்கு மாதிரியளவைகள் என்பது பெயர்.

கவனிக்கவும்

முழுமைத் தொகுதியளவைகள் மாறிலிகள் (constants) ஆகும். ஆனால் மாதிரியளவைகள் ராண்டம் மாறிகளாகும்.

13·1·3. சில முக்கிய முழுமைத் தொகுதியளவைகள், மாதிரியளவைகள்

$(X_1, X_2, \ldots X_N)$ என்பவை முழுமைத் தொகுதி உறுப்புகளானல்

(அ) $\bar{X} = \frac{1}{N}\sum_{i=1}^{N} X_i$ என்பது முழுமைத் தொகுதிச் சராசரி (Population mean) யாகும்.

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

என்பது முழுமைத் தொகுதி பரவற் படி (Population variance) யாகும்.

(இ) $\sum_{i=1}^{N} X_i$ என்பது முழுமைத் தொகுதி மொத்தம் (Population total) ஆகும்.

$(x_1, x_2, \ldots \ldots x_n)$ என்பவை மாதிரி உறுப்புகளானால்

(அ) $\bar{x} = \frac{1}{n} \sum_{i=1}^{n} x_i$ என்பது மாதிரிச் சராசரி (sample mean) ஆகும்.

(உ) $s^2 = \frac{1}{(n-1)} \sum_{i=1}^{n} (x_i - \bar{x})^2$ என்பது மாதிரி பரவற் படி (sample variance) ஆகும்.

13·1·4. மாதிரியளவைப் பரவல்கள் (Sampling distributions of Statistic)

$t = (x_1, x_2, \ldots x_n)$ என்பது ஒரு மாதிரியானால் $T = g(t)$ எனும் மாதிரி உறுப்புகளின் சார்பு ஒரு ராண்டம் மாறியாகும். முன்பே குறிப்பிட்டபடி T என்பது ஒரு மாதிரியளவை.

T என்பது ராண்டம் மாறியாதலால் இதற்கு ஒரு பரவல் இருக்கும். இப்பரவலுக்கு மாதிரியளவைப் பரவல் என்பது பெயர். உதாரணத்திற்கு $\bar{x}$, s^2 முதலியவைகளுக்குப் பரவல்கள் இருக்கும்.

$\bar{x}$-ன் பரவலுக்கு மாதிரிச் சராசரியின் பரவல் என்பது பெயர்.

s^2-ன் பரவலுக்கு மாதிரிப் பரவற் படியின் பரவல் என்பது பெயர்.

13·2. பிறழ்ச்சியற்ற மதிப்பீட்டி (uubiased estimation)

13·2·1. T என்பது ஒரு மாதிரியளவையாக இருக்கட்டும். θ என்பது ஒரு முழுமைத் தொகுதியளவையாக இருக்கட்டும்.

T-ன் எதிர்பார்க்கும் மதிப்பு θ ஆனால் T ஐ, θ-ன் பிறழ்ச்சியற்ற மதிப்பீட்டி எனலாம்.

அதாவது $E(T) = \theta \Longleftrightarrow T$ என்பது θ-ன் பிறழ்ச்சியற்ற மதிப்பீட்டி.

13·2·2. மதிப்பீடு

மதிப்பீட்டி என்பது ஒரு ராண்டம் மாறியாகும். இதன் மதிப்புகளுக்கு மதிப்பீடு என்பது பெயர். உதாரணமாக

(1, 2, 3, 4) என்பது முழுமைத் தொகுதியாகவும் (1, 2), (1, 3), (1, 4), (2, 3), (2, 4), (3, 4) என்பவை மாதிரியாகவும் எடுத்துக் கொண்டால் $\frac{1}{2}(x_1 + x_2)$ என்பதை $\frac{1+2+3+4}{4}$ -ன் மதிப்பீட்டியாக எடுத்துக் கொள்ளலாம். [$(x_1 \ x_2)$ என்பது மாதிரி] இம்மாதிரியளவையின் மதிப்புகள் முறையே,

$$\frac{1+2}{2}, \frac{1+3}{2}, \frac{1+4}{2}, \frac{2+3}{2}; \frac{2+4}{2}, \frac{3+4}{2}$$ ஆகும்.

இவை ஒவ்வொன்றிற்கும் மதிப்பீடு என்பது பெயர்.

ஆகவே மதிப்பீட்டி என்பது ஒரு ராண்டம் மாறி. ஆனால் மதிப்பீடு என்பது மதிப்பீட்டியின் குறிப்பிட்ட மதிப்பாதலால் அது ஒரு மாறிலி (constant) ஆகும்.

13·2·3. வரைவிலக்கணம்

$E(T) = \theta$ என்றால், T என்பது θ-ன் பிறழ்ச்சியற்ற மதிப்பீட்டி என்று குறிப்பிடலாம். இதையே

$T \sim E^{-1}(\theta)$ என்றும் குறிப்பிடலாம்.

அதாவது $E^{-1}(\theta)$ என்பது θ-ன் பிறழ்ச்சியற்ற மதிப்பீட்டியாகும்.

$E(T) = EE^{-1}(\theta) = \theta$ என்பதால் EE^{-1} என்பது இயற்கணித முறைப்படி 1-க்குச் சமமாகும். அதாவது $EE^{-1} \equiv 1$.

$E^{-1}(\theta) = E^{-1}E(T) \sim T$ என்பதால் $E^{-1}E \sim 1$ என்பது தெளிவாகிறது.

தேற்றம் 13·2·4.

$$E^{-1}[\alpha\,\theta_1 + \beta\theta_2] \sim \alpha\,E^{-1}(\theta_1) + \beta\,E^{-1}(\theta_2)$$

[α, β மாறிலிகள்].

நிறுவல்

$E(T_1) = \theta_1$, $E(T_2) = \theta_2$ எனக் கொண்டால்

$$E(\alpha\,T_1 + \beta\,T_2) = \alpha\,E(T_1) + \beta\,E(T_2)$$
$$= \alpha\,\theta_1 + \beta\,\theta_2$$
$$\therefore\ \alpha\,T_1 + \beta\,T_2 \sim E^{-1}(\alpha\,\theta_1 + \beta\,\theta_2)$$

$T_1 \sim E^{-1}(\theta_1)$, $T_2 \sim E^{-1}(\theta_2)$ என்பதால்

$$\alpha\,E^{-1}(\theta_1) + \beta\,E^{-1}(\theta_2) \sim E^{-1}(\alpha\,\theta_1 + \beta\,\theta_2)$$

கவனிக்கவும்

$T_1 \sim T_2$ என்றால் $E(T_1) = E(T_2)$ என்பது பொருள். ஆனால் இதிலிருந்து $T_1 = T_2$ என்று சொல்லமுடியாது.

தேற்றம் 13·2·5.

(அ) $T_1 \sim T_2 \Longrightarrow T_2 \sim T_1$

(ஆ) $T_1 \sim T_2, T_2 \sim T_3 \Longrightarrow T_1 \sim T_3$.

(இ) $T_1 \sim T_1$.

நிறுவல்

(அ) $$T_1 \sim T_2 \Longrightarrow E(T_1) = E(T_2) \Longrightarrow E(T_2) = E(T_1) \Longrightarrow T_2 \sim T_1.$$

(ஆ) $$T_1 \sim T_2, T_2 \sim T_3 \Longrightarrow E(T_1) = E(T_2); E(T_2) = E(T_3) \Longrightarrow E(T_1) = E(T_3) \Longrightarrow T_1 \sim T_3.$$

(இ) $$E(T_1) = E(T_1) \Longrightarrow T_1 \sim T_1.$$

துணை முடிவு

$\sim$ என்னும் தொடர்புக் குறியானது (அ), (ஆ), (இ) என்னும் விதிகளையொட்டி நடப்பதால் இக்குறியைச் சமத்தொடர்பு (Equivalence relation) எனலாம். அதாவது "ஒரே எதிர்பார்க்கும் மதிப்பை உடையன" என்பது ஒரு சமத்தொடர்பு.

13·2·6. பிறழ்ச்சி

T என்பது ஒரு மாதிரியளவையாகவும், θ என்பது முழுமைத் தொகுதியளவையாகவும் இருக்கின்றன எனக் கொள்வோம்.

$E(T)$, θ-க்குச் சமமாக இல்லையென்றால் T என்பது θ-ன் பிறழ்ச்சியுடைய மதிப்பீட்டி எனலாம்.

$B_\theta(T) = E(T) - \theta$ என்பது T-ன் பிறழ்ச்சியென்று கூறப்படும்.

துணை முடிவு

T, θ-ன் பிறழ்ச்சியற்ற மதிப்பீட்டி $\Longleftrightarrow B_\theta(T) = 0$.

அதாவது $E^{-1}(\theta) \sim T \Longleftrightarrow B_\theta(T) = 0$.

13·2·7. சாலச் சிறந்த மதிப்பீட்டி (Best Estimator)

θ என்பது ஒரு முழுமைத் தொகுதியளவையானால் அதற்குச் சாதாரணமாகப் பல மதிப்பீட்டிகள் இருக்கும். ஆகவே இவைகளில் சாலச் சிறந்தது என்று எதைத் தேர்ந்தெடுப்பது எனும் கேள்வியெழலாம்.

இதற்குப் பிழைச்சார்பு என்று ஒன்றை வரையறுப்பதன் மூலம் தீர்வு காணலாம். T என்பது ஒரு மதிப்பீட்டியானால் $L_\theta(T)$ எனும் சார்பை, T-ன் நஷ்டச் சார்பு (Loss function) ஒன்றை, தகுந்த காரணங்களைக் கொண்டு வரையறுக்கலாம். $L_\theta(T)$ என்பது ராண்டம் மாறி என்பதால் இதற்கு எதிர்பார்க்கும் மதிப்பு ஒன்று இருக்கும். இம்மதிப்பைப் பிழைச்சார்பு எனலாம். இதை $R_\theta(T)$ என்று குறிப்பிடலாம்.

அதாவது $E[L_\theta(T)] = R_\theta(T)$.

T_1, T_2 என்று θ-ன் மதிப்பீட்டிகள் இரண்டை எடுத்துக் கொள்வோம். எல்லா θ-க்கும் $R_\theta(T) \leq R_\theta(T_2)$ என்பது உண்மையானால் T_2 ஐ விட T_1 ஆனது சீராக மேலானது (Uniformly better) எனலாம். $[T_i]\ i \in I$ என்பது θ மதிப்பீட்டிகளின் கணமெனக் கொள்வோம்.

$R_\theta(T_o) \leq R_\theta(T_i)\ \forall\, i \in I, \forall\, \theta$ என்பது உண்மையானால், T_0 என்பது சாலச்சிறந்த மதிப்பீட்டி எனலாம். ஆனால் பொதுவாக இத்தகைய T_0 ஆனது இல்லாமலேயே போய்விடக் கூடும் அதாவது சீரான சிறந்த மதிப்பீட்டி இல்லாமலேயே இருக்கக்கூடும்.

13·2·8. சில முக்கியப் பிழைச் சார்புகள்

(அ) T என்பது θ-ன் மதிப்பீட்டியானால் $E(T-\theta)^2$ என்பதை T-ன் பிழையாகக் கொள்ளலாம். இது θ ஐச் சுற்றிய வர்க்கச் சராசரி விலக்கம் ஆகும். (Mean square deviation about θ).

(ஆ) T என்பது θ-ன் பிறழ்ச்சியற்ற மதிப்பீட்டியானால் $\theta = E(T)$. ஆகவே $E(T-\theta)^2 = E[T - E(T)]^2 = V(T)$

ஆகவே பிறழ்ச்சியற்ற மதிப்பீட்டியின் பிழை அதனுடைய பரவற் படியேயாகும்.

$\sqrt{V(T)}$ என்பது T-ன் தரப்பிழை (Standard Error) என்று குறிப்பிடப்படும்.

13·2·9. பிறழ்ச்சியற்ற மதிப்பீட்டியின் பண்புகள்

θ எனும் முழுமைத் தொகுதியளவை மதிப்பிடுவதற்கு T எனும் மாதிரியளவையை உபயோகப்படுத்துவதாக வைத்துக் கொள்வோம். T-ன் மதிப்புகள் θ-க்கு மிகவும் நெருக்கமாக அமைந்திருக்கவேண்டும் என்பது மதிப்பீட்டிக்கு இருக்க வேண்டியதொரு முக்கியமான பண்பாகும். T-ன் மதிப்பையே நாம் θ ஆக எடுத்துக் கொள்வதால் இப்பண்பு மிகவும் முக்கிய மானதாகிறது.

T-ன் மதிப்புகள் θ-க்கு நெருக்கமாக இருக்கவேண்டு மெனக் குறிப்பிட்டோம். ஆனால் நெருக்கத்தை எவ்வாறு அளப்பது? நெருக்கத்தை அளப்பதற்குத்தான் $E(T-\theta)^2$ என்னும் பிழைச்சார்பைப் பயன்படுத்துகிறோம். இப்பிழை குறை வாக இருந்தால் T, θ-க்கு நெருங்கியுள்ளது எனக் கொள்ளலாம். பொதுவாக T என்பது θ-ன் பிறழ்ச்சியற்ற மதிப்பீட்டியாக இருந்தால், மற்ற பிறழ்ச்சியுள்ள மதிப்பீட்டிகளை விட T சிறந்தது எனக் கூறலாம். ஆனால் சிற்சில சமயங்களில் பிறழ்ச்சியுள்ள மதிப்பீட்டி, பிறழ்ச்சியற்ற மதிப்பீட்டியைவிட θ-உடன் அதிக நெருக்கமாக இருப்பதுண்டு. இப்படிப்பட்ட வேளைகளில் எம் மதிப்பீட்டிக்குப் பிழை குறைவாக உள்ளதோ அதையே சிறந்த தாகக் கொள்ளவேண்டும். ஆனால் மாதிரிமுறைகளில் பொதுவாகப் பிறழ்ச்சியற்ற மதிப்பீட்டிகளையே நாம் உபயோகிக்கிறோம். பிறழ்ச்சியற்ற மதிப்பீட்டிகள் ஒரு குறிப்பிட்ட θ-க்கு இல்லாமற் போவதும் உண்டு. அச்சமயங்களில் பிறழ்ச்சியுள்ள மதிப்பீட்டி களையே உபயோகப்படுத்தவேண்டிய நிர்ப்பந்தம் ஏற்படுகிறது.

தேற்றம் 13·3·1

T_1, T_2 என்பவை θ-ன் பிறழ்ச்சியற்ற மதிப்பீட்டிகளானால் $\lambda T_1 + (1-\lambda) T_2$ என்பது θ-ன் பிறழ்ச்சியற்ற மதிப் பீட்டியாகும்.

நிறுவல

$$E[\lambda (T_1) + (1-\lambda) T_2] = \lambda E(T_1) + (1-\lambda) E(T_2)$$
$$= \lambda\theta + (1-\lambda)\theta = \theta.$$

பயிற்சி 1

$V(T_1) = \sigma_1^2$, $V(T_2) = \sigma_2^2$, cov $(T_1, T_2) = \sigma_{12}$ என்றால் λ-ன் எந்த மதிப்புக்கு $V[\lambda T_1 + (1-\lambda) T_2]$ மிகவும் குறைவாக இருக்கும் என்பதைக் கண்டுபிடி.

$$S(\lambda) = V[\lambda T_1 + (1-\lambda) T_2] = \lambda^2 \sigma_1^2 + (1-\lambda)^2 \sigma_2^2 + 2\lambda(1-\lambda)\sigma_{12}$$

இது மிகவும் குறைவாக இருக்கவேண்டுமானால் $\frac{d}{d\lambda} S(\lambda) = 0$

$$\therefore \quad 2\lambda\sigma_1^2 - 2(1-\lambda)\sigma_2^2 + (2-4\lambda)\sigma_{12} = 0.$$

$$2\lambda[\sigma_1^2 + \sigma_2^2 - 2\sigma_{12}] = 2\sigma_2^2 - 2\sigma_{12}.$$

$$\therefore \quad \lambda = \frac{\sigma_2^2 - \sigma_{12}}{\sigma_1^2 + \sigma_2^2 - 2\sigma_{12}}$$

துணை முடிவு

$$\lambda T_1 + (1-\lambda) T_2 = \frac{\sigma_2^2 - \sigma_{12}}{\sigma_1^2 + \sigma_2^2 - 2\sigma_{12}} \cdot T_1 + \frac{\sigma_1^2 - \sigma_{12}}{\sigma_1^2 + \sigma_2^2 - 2\sigma_{12}} T_2.$$

$$V[\lambda T_1 + (1-\lambda) T_2] = \left[\frac{\sigma_2^2 - \sigma_{12}}{\sigma_1^2 + \sigma_2^2 - 2\sigma_{12}}\right]^2 \sigma_1^2 + \left[\frac{\sigma_1^2 - \sigma_{12}}{\sigma_1^2 + \sigma_2^2 - 2\sigma_{12}}\right]^2 \sigma_2^2 + \frac{2(\sigma_2^2 - \sigma_{12})(\sigma_1^2 - \sigma_{12})}{(\sigma_1^2 + \sigma_2^2 - 2\sigma_{12})^2}\sigma_{12}$$

$$= \frac{1}{(\sigma_1^2 + \sigma_2^2 - 2\sigma_{12})^2} [(\sigma_2^2 - \sigma_{12})^2 \sigma_1^2 + (\sigma_1^2 - \sigma_{12})^2 \sigma_2^2 + 2(\sigma_2^2 - \sigma_{12})(\sigma_1^2 - \sigma_{12})\sigma_{12}]$$

$$= \frac{1}{(\sigma_1^2 + \sigma_2^2 - 2\sigma_{12})^2} [\{\sigma_1(\sigma_2^2 - \sigma_{12}) + \sigma_2(\sigma_1^2 - \sigma_{12})\}^2 + 2(\sigma_2^2 - \sigma_{12})(\sigma_1^2 - \sigma_{12})(\sigma_{12} - \sigma_1\sigma_2)]$$

14. நம்பிக்கை இடைவெளி

(CONFIDENCE INTERVAL)

14·1·1. θ என்பது முழுமைத் தொகுதியளவை எனக் கொள்வோம். $(x_1, x_2, \dots, x_n)$ என்னும் ஒரு மாதிரியைத் துணையாகக் கொண்டு பல்வேறு மாதிரியளவைகள் நிறுவப்படுவதாகக் கொள்வோம். $100\alpha\%$ நம்பிக்கைக்கெழு (confidence coefficient) உள்ள நம்பிக்கை இடைவெளியென்பது (T_1, T_2) என்னும் இடைவெளியாக இருக்கவேண்டுமானுல்,

$P[T_1 \leq \theta \leq T_2] = \alpha$ என்றிருக்கவேண்டும். T_1, T_2 என்பவை மாதிரியளவைகள்.

(அ) உதாரணமாக இயல்நிலைப் பரவலிலிருந்து எடுக்கப்பட்ட $(x_1, x_2, \dots, x_n)$ எனும் சார்பற்ற உறுப்புகளின் மாதிரிச் சராசரி $\bar{x}$, மாதிரிப் பரவற் படி s^2 ஆனுல்,

$$P\left[\left|\frac{\bar{x}-\mu}{\sigma}\right| \leq 1{\cdot}96\right] = 0{\cdot}95$$

(μ, σ என்பவை இயல்நிலைப் பரவலின் சராசரி, பரவற் படியாகும்.)

$$\therefore \quad P[\bar{x} - 1{\cdot}96\,\sigma \leq \mu \leq \bar{x} + 1{\cdot}96\,\sigma] = 0{\cdot}95.$$

$\therefore$ μ -ன் 95% நம்பிக்கை இடைவெளியானது $[\bar{x} - 1{\cdot}96\,\sigma, \bar{x} + 1{\cdot}96\,\sigma]$.

(ஆ) σ என்பது தெரியாமலிருந்தால்,

$$P\left[\frac{|\bar{x}-\mu|}{\frac{s}{\sqrt{n}}} \leq t_{n-1}\left(\frac{\alpha}{2}\right)\right] = \alpha$$

$t_{n-1}\left(\frac{\alpha}{2}\right)$ என்பது $(n-1)$ வரையற்ற பாகை (Degrees of Freedom) யுடைய t-பரவலின் மேல் $\frac{\alpha}{2}$ புள்ளி (upper $\frac{\alpha}{2}$ point of t-distribution with $n-1$ degrees of freedom)

$$\therefore \quad P\left[\bar{x} - \frac{s}{\sqrt{n}}\, t_{n-1}\left(\frac{\alpha}{2}\right) \le \mu \le \bar{x} + \frac{s}{\sqrt{n}}\, t_{n-1}\left(\frac{\alpha}{2}\right)\right] = \alpha.$$

$\therefore$ μ-ன் 100 α % நம்பிக்கை இடைவெளி

$$\left[\bar{x} - \frac{s}{\sqrt{n}}\, t_{n-1}\left(\frac{\alpha}{2}\right),\ \bar{x} + \frac{s}{\sqrt{n}}\, t_{n-1}\left(\frac{\alpha}{2}\right)\right]$$ ஆகும்.

(இ) பொதுவாக மாதிரியளவைகளுக்கு இயல்நிலைப் பரவல் உள்ளதாகத்தான் நாம் கொள்ளவேண்டி வரும். ஆகவே உதாரணம் (அ)-ல் உள்ளதைப் போன்று தான் நம்பிக்கை இடைவெளிகள் அமையும். ஆனால் நம்பிக்கை இடைவெளிகள் யாவும் பொதுவாகச் சிறப்புகாண் சோதனை (Tests of significance) மூலமாகத்தான் நிறுவப்படுகின்றன.

15. மாதிரிகள்

(SAMPLES)

15·1·1. முழுமைத் தொகுதியிலுள்ள உறுப்புகள் $U_1, U_2, \ldots, U_N$ எனக் கொள்வோம். இவ்வுறுப்புகளுடன் $Y_1, Y_2, \ldots Y_N$ எனும் அளவுகளைத் தொடர்புப் படுத்துவோம். உதாரணமாக $U_1, U_2, \ldots, U_N$ என்பவை ஒரு வகுப்பிலுள்ள மாணவர்களைக் குறித்தால் $Y_1, Y_2, \ldots Y_N$ என்பவை முறையே இம் மாணவர்களின் வயதுகளாகக்கொள்ளலாம். இம் முழுமைத் தொகுதியிலிருந்து 'n' உறுப்புகள் கொண்ட மாதிரியொன்று தேவைப்பட்டால் இம் மாதிரியைப் பல விதங்களில் எடுக்கலாம்.

(அ) $(U_1, U_2, \ldots, U_N)$ என்னும் உறுப்புகளைப் பரிசோதித்து நம் மனதிற்குச் சிறந்ததாகப் படும் 'n' உறுப்புகளைத் தேர்ந்தெடுத்து இவைகளையே மாதிரியாகக் கொள்ளலாம். இதற்கு உத்தேச மாதிரி (Judgement Sample) என்பது பெயர். ஆனால் இத்தகைய மாதிரிகளை எடுப்பதில் பலவிதத் தொல்லைகள் உள்ளன. உத்தேச மாதிரியை எடுத்தால் அதில் நம்முடைய மனப் பிறழ்ச்சி (Bias) அதிகமாகக் காணப்படும். மனோதத்துவ முறைப்படி, எவராலும் இப் பிறழ்ச்சியை இல்லாதவாறு பார்த்துக்கொள்வது இயலாதது என்பது தெளிவாக நிரூபிக்கப்பட்டிருக்கிறது. ஒவ்வொருவருக்கும் தனித்தனியாக நிறம், சுவை, முதலியவற்றுள் வேண்டுவன, வேண்டாதன என்பவை இருப்பதால் இப்பிறழ்ச்சி யுண்டாகிறது. மேலும் இத்தகைய மாதிரிகளைப் பற்றி நிகழ்திறன் கணக்குகளைப் போடமுடியாது. ஏனெனில் நம் மனத்தின் விருப்பு வெறுப்புகளைத் துல்லியமாகக் கணக்கிட்டுக் கூற எவ்வாறு முடியும்? ஆகவே இத்தகைய மாதிரிகளைக் கொண்டு மேற்கொள்ளப்படும் முடிவுகளிலுள்ள பிழைகளைப் புள்ளியியல் முறையில் கணக்கிட்டுக் கூற இயலாது. ஆகவே புள்ளியியலாளர் இம் முறையை அறவே ஒதுக்கிவிடுவர்.

(ஆ) நிகழ்திறன் மாதிரிகள் (Probability Samples)

இம் மாதிரிகளில் முழுமைத் தொகுதியிலுள்ள ஒவ்வோர் உறுப்பும் ஒரு குறிப்பிட்ட (தெரிந்து) நிகழ்திறனுடன் எடுக்கப் படும். இப்படி எடுப்பதற்குப் பொதுவாகத் தானியங்கும் கருவி களையோ அல்லது அக்கருவிகளின் மூலம் நிறுவப்பட்ட அட்டவணைகளையோ நாம் உபயோகப்படுத்தவேண்டி வரும். உத்தேச முறையில் இதை எடுக்க இயலாது. மாதிரியை உபயோகித்துப் புள்ளியியல் கணக்குகளைச் செய்யும்பொழுது முழுமைத் தொகுதி உறுப்புகளின் நிகழ்திறன்களை உபயோகிக்க வேண்டியிருக்கும். பொதுவாக இதைப் பின்வரும் அட்டவணை மூலம் தெளிவாக்கலாம்.

Y_1	Y_2	Y_3	... ...	Y_N
P_1	P_2	P_3	... ...	P_N

Y_i, $i = 1, 2, \ldots N$ என்பவை முழுமைத் தொகுதி உறுப்புகள்.

P_i, $i = 1, 2, \; N$ என்பவை Y_i எனும் உறுப்பு மாதிரியில் தேர்ந்தெடுக்கப்படுவதற்கான நிகழ்திறன்கள்.

$$\sum_{i=1}^{N} P_i = 1$$ என்பது தெளிவாகத் தெரிகிறது.

முழுமைத் தொகுதியிலுள்ள ஒவ்வோர் உறுப்புக்கும் மாதிரி யில் வருவதற்குக் குறுப்பிட்ட நிகழ்திறன் இருந்தால் அம் மாதிரிக்கு நிகழ்திறன் மாதிரி அல்லது ராண்டம் மாதிரி (Random Sample) என்பது பெயர்.

மற்ற மாதிரிகளுக்கு ராண்டமற்ற மாதிரிகள் (Non-Random Samples) என்பது பெயர். உத்தேச மாதிரியென்பது இவ் வகையைச் சேர்ந்தது.

இப்புத்தகத்தில் பார்க்கப் போவது ராண்டம் மாதிரி முறைகளை மட்டும் தான்.

15·1·2. மாதிரிப் பின்னம் (Sampling Fraction)

முழுமைத் தொகுதியில் N உறுப்புகளும் மாதிரியில் n உறுப்புகளும் இருந்தால் $\frac{n}{N}$ என்பது மாதிரிப் பின்னம் ஆகும்.

15.1.3. முழுமைத் தொகுதிப் பட்டியல் (Sampling Frame)

முழுமைத் தொகுதியிலுள்ள உறுப்புகளின் பட்டியல் மாதிரிகளை எடுப்பதற்கு இன்றியமையாததாகும், இப்பட்டியலை முழுக் கணிப்பு (census) முறையில்தான் அமைக்க முடியும். இப் பட்டியலை $(U_1, U_2, ..., U_N)$ என்று குறிப்பிடலாம். பொதுவாக இப் பட்டியலில் உறுப்புகளைப் பற்றிய தேவையான விவரங்கள், வரைப் படம் (map) முதலியவை இடம் பெறும்.

15·1·4. சார்புப் பயனளவு (Relative Efficiency)

T_1, T_2 என்பவை θ-ன் பிறழ்ச்சியற்ற மதிப்பீட்டிகளானுல் $\frac{V(T_1)}{V(T_2)}$ என்பது T_2-ன் T_1ஐச் சார்ந்த சார்புப் பயனளவு எனப்படும். இது ஒன்றைவிட அதிகமாக இருந்தால் T_1 ஐவிட T_2 சிறந்ததாகக் கொள்ளப்பட வேண்டும். இது ஒன்றைவிடக் குறைவாக இருந்தால் T_2ஐ விட T_1 சிறந்தாகக் கொள்ளப்பட வேண்டும்.

16. சாதாரண ராண்டம் மாதிரிமுறை

(SIMPLE RANDOM SAMPLING)

16·1· இம் மாதிரி முறை ராண்டம் மாதிரி முறைகளிலேயே மிகவும் எளிதானதாகும்· இம்முறையில் உறுப்புகள் ஒவ்வொன்றாக எடுக்கப்படும்· ஒவ்வோர் உறுப்பு எடுக்கப்படும் போதும் எல்லா முழுமைத் தொகுதி உறுப்புகளுக்கும் சரிசம வாய்ப்பு, அதாவது சம நிகழ்திறன் தரப்படவேண்டும்· ஆகவே இம்முறையைச் சரிசம வாய்ப்புக் கூறு முறையென்றும் கூறலாம். இம்முறையையும் இருவகையாகப் பிரிக்கலாம். அவையாவன·

(அ) ஈடு செய்யப்படும் சாதாரண ராண்டம் மாதிரி முறை (Simple Random Sampling with replacement) [ஈ·செ·சா·ரா·மா.]

(ஆ) ஈடு செய்யப்படாத சாதாரண ராண்டம் மாதிரி முறை (Simple Random Sampling without replacement) [ஈ·செயா·கா·ரா.மா.]

16·2· ஈ·செ.சா·ரா·மா.

இம்முறையில் ஓர் உறுப்பு சரிசம வாய்ப்புடன் தேர்ந்தெடுக்கப்பட்ட பின்னர் திரும்பவும் முழுமைத் தொகுதியில் சேர்க்கப்பட்டு விடும். ஆகவே இரண்டாம் முறை உறுப்பொன்று தேர்ந்தெடுக்கப் படும்போது அது முதல் முறை எடுக்கப்பட்ட உறுப்பாகவே கூட இருக்கலாம்· அதாவது ஓர் உறுப்பே பல முறை மாதிரியில் வருவதற்கு வாய்ப்புகள் உள்ளன· ஆகவே எடுக்கப்படும் உறுப்புகள் யாவும் ஒன்றுக்கொன்று சார்பிலாமல் (independent) இருக்கும்· முழுமைத் தொகுதியையே x எனும் ராண்டம் மாறியின் பரவல் என எடுத்துக் கொள்வோமாயின் இம்முறையில் எடுக்கப்படும் உறுப்புகளை x-ன் சார்பிலாப் பிரதிகளாகக் (Independent copies) கொள்ளலாம்.

முழுமைத் தொகுதியானது தொடர்ச்சியாக (continuous) இருந்தால், ஓர் உறுப்பே ஒன்றுக்கு மேற்பட்ட முறைகள் வருதற்கு வாய்ப்பே இராதாகையால் சாதாரண ராண்டம் மாதிரி முறையை (அ), (ஆ) என்று இருவகையாகப் பிரிப்பதில் அர்த்தமே இல்லை. ஆகவே தொடர்ச்சியான முழுமைத் தொகுதியைப் பொறுத்த வரையில் சாதாரண ராண்டம் மாதிரி என்று மட்டும் சொன்னால் போதும். (அ), (ஆ) என்பதில் ஐயப்பாட்டிற்கிடமே யில்லை. பொதுவாகத் தொடர்ச்சியான முழுமைத் தொகுதியில் சாதாரண ராண்டம் மாதிரியை எடுக்கும்போது அம்மாதிரியை ஈ·செ·சா·ரா·மா என்று கொள்வதே நல்லது, புள்ளியியல் கணக்குகளைப் பொறுத்த மட்டில் (ஆ) ஐ விட (அ) தான் எளியது.

முழுமைத் தொகுதியில் N உறுப்புகள் இருந்தால் N^n, n, உறுப்புகள் கொண்ட ஈ·செ·சா·ரா·மா-க்களை எடுக்கலாம். எல்லா மாதிரிக்கும் சரிசமவாய்ப்பு இருப்பதால் ஒவ்வொரு மாதிரிக்கும் நிகழ்திறன் $\frac{1}{N^n}$ என்பது தெளிவாகத் தெரிகிறது. ஆகவே ஈ.செ.சா.ரா.மா. முறையில் வெக்டர் ராண்டம் மாறியைப் பின் வருமாறு விவரிக்கலாம்,

$\underset{\sim}{x}$ என்பது ஒரு மாதிரியானால் $\underset{\sim}{x} = (x_1\ x_2, \ldots, x_n)$. x_i என்பது i ஆவது முறை மாதிரியில் எடுக்கப்படும் உறுப்பின் பண்பாகும். முழுமைத் தொகுதியிலுள்ள உறுப்புகளை $(X_1, X_2, \ldots, X_N)$ என எடுத்துக் கொண்டால்

$$P[\underset{\sim}{x} = (X_{i1}, X_{i2}, X_{i3} \ldots, X_{in})] = \frac{1}{N^n}$$

$(i_1, i_2, \ldots, i_n)$ என்பன $(1, 2, \ldots, N)$ எனும் எண்களிலிருந்து எடுக்கப்பட்ட n எண்கள் (இவைகளில் சில சமமாகவும் கூட இருக்கலாம்.)

முழுமைத் தொகுதி, தொடர்ச்சியானது எனவும் $f(x)$ என்பது முழுமைத் தொகுதிப் பரவலின் ஊக அளவு அடர்த்தி (Probability density) எனவும் கொண்டால்,

$$P[\underset{\sim}{x} = (a_1, a_2, \ldots, a_n)] = f(a_1) f(a_2) \ldots f(a_n)$$

$$= \prod_{i=1}^{n} f(a_i).$$

a_i என்பவை மெய்யெண்களாகும் (real numbers). $P[\underset{\sim}{x} = (a_1, \ldots, a_n)]$ என்பது $\underset{\sim}{x}$-ன் நிகழ்திறம் அல்ல. இது $\underset{\sim}{x}$-ன் ஊக அளவு அடர்த்தியையே குறிக்கும். உண்மையில் பின்வருமாறு எழுதுவதே பொருத்தமானது.

$$P[a_1 \leq x_1 \leq a_1 + \triangle a_1, a_2 \leq x_2 \leq a_2 + \triangle a_2, \ldots, \ldots, a_n \leq x_n \leq a_n + \triangle a_n]$$
$$= \prod_{i=1}^{n} f(a_i) \triangle a_i$$

துணைமுடிவு

அதாவது $\underset{\sim}{x}$-ன் ஊக அளவு அடர்த்தி $\prod_{i=1}^{n} f(x_i)$ ஆகும்.

முழுமைத் தொகுதியின் பரவல் சார்பு $F(x)$ ஆனல் $\underset{\sim}{x}$-ன் பரவல் சார்பு $F(\underset{\sim}{x}) = \prod_{i=1}^{n} F(x_i)$ $\left[F(x) = \int_{-\infty}^{x} f(x)\,dx\right]$.

16·3. உதாரணங்கள்

(அ) முழுமைத் தொகுதியில் (2, 7, 9, 12) என்னும் நான்கு உறுப்புகளே இருக்கின்றன. என்று எடுத்துக்கொள்வோமாயின் 2 உறுப்புகள் கொண்டஈ. செ. சா. ரா. மா-களின் முழுப்பட்டியல் பின்வருமாறு.

(2, 2), (2, 7), (2, 9), (2, 12)
(7, 2), (7, 7), (7, 9), (7, 12)
(9, 2), (9, 7), (9, 9), (9, 12)
(12, 2), (12, 7), (12, 9), (12, 12)

இவை ஒவ்வொன்றிக்கும் நிகழ்திறம் $\frac{1}{16}$ ஆகும்.

(ஆ) $[a, b]$ எனும் மெய்யெண் வெளியில் வரையறுக்கப் படும் செவ்வகப் பரவல் (Rectangular distribution) என்பது பின்வரும் ஊக அளவு அடர்த்தி மூலம் விளக்கப்படுகிறது,

$$f(x) = \frac{1}{p-a}, \quad a \leq x \leq b$$
$$= 0 \qquad x \notin [a, b].$$

இச்செவ்வகப் பரவலிலிருந்து மூன்று உறுப்புகள் சா. ரா, மா. முறையில் எடுத்தால் (x_1, x_2, x_3) என்னும் மாதிரியின் ஊக அளவு அடர்த்தி

$$f(x_1, x_2, x_3) = \frac{1}{(b-a)^3} \quad x_i \in [a, b] \quad i = 1, 2, 3$$
$$= 0 \qquad x_i \notin [a, b].$$

(x_1, x_2, x_3)-ன் பரவல் சார்பு

$$F(x_1, x_2, x_3) = \left(\frac{x_1 - a}{b - a}\right)\left(\frac{x_2 - a}{b - a}\right)\left(\frac{x_3 - a}{b - a}\right)$$

$$= \frac{1}{(b-a)^3} \prod_{i=1}^{3} (x_i - a)$$

(இ) கோஷி பரவல் (Cauchy's Distribution) பின்வருமாறு வரையறுக்கப்படுகிறது.

$$f(x) = \frac{1}{\pi} \frac{1}{1+(x-\mu)^2}, \qquad -\infty \leq x \leq +\infty$$

இப்பரவலின் பரவல் சார்பு $F(x) = \int_{-\infty}^{x} \frac{1}{\pi} \frac{1}{1+(x-\mu)^2} dx$

$$= \frac{1}{\pi}\Big[\tan^{-1}(x-\mu)\Big]_{-\infty}^{x}$$

$$= \frac{1}{\pi}\left[\tan^{-1}(x-\mu) - \frac{\pi}{2}\right]$$

$$= \frac{1}{\pi}\tan^{-1}(x-\mu) - \frac{1}{2}$$

இப்பரவலிலிருந்து (x_1, x_2) எனும் இரண்டு உறுப்புகளை சா. ரா. மா. முறையில் எடுத்தால்

$$f(x_1, x_2) = \frac{1}{\pi^2} \frac{1}{1+(x_1-\mu)^2} \frac{1}{1+(x_2-\mu)^2}$$

$$F(x_1, x_2) = \left[\frac{1}{\pi}\tan^{-1}(x_1-\mu)-\frac{1}{2}\right]\left[\frac{1}{\pi}\tan^{-1}(x_2-\mu)-\frac{1}{2}\right]$$

16·4. சாதாரண ராண்டம் மாதிரியை எடுக்கும் முறைகள்

(அ) குலுக்கும் முறை (Lottery Method)

முழுமைத் தொகுதியில் N உறுப்புகள் இருக்கின்றன என்றும் மாதிரியில் n உறுப்புகள் இருக்கவேண்டும் என்றும் கொள்வோம். முழுமைத் தொகுதியிலுள்ள உறுப்பு ஒவ்வொன்றிற்கும் ஒரு சீட்டு எழுதி ஒரு குடத்திலோ அல்லது குலுக்குவதற்கென்றே அமைத்த ஓர் இயந்திரத்திலோ போடவேண்டும். சீட்டுகளை நன்கு குலுக்கிய பின்னர் ஒரு சீட்டை எடுக்க வேண்டும். அச்சீட்டிற்குரிய உறுப்பை ஒரு காகிதத்தில் குறித்து வைத்துக் கொண்ட பின்னர் அவ்வுறுப்பை மறுபடியும் குடத்தில் (இயந்திரத்தில்) போட்டுவிட்டு மறுபடியும் குலுக்கிய பின்னர் இரண்டாவது சீட்டை எடுக்கவேண்டும். இப்படித் தொடர்ந்து n உறுப்புகள் எடுத்தால் n உறுப்புகள் கொண்ட ஈ. செ. சா. ரா. மா. ஒன்று கிடைக்கும்.

ஈ. செயா. சா. ரா. மா. முறையில் n உறுப்புகள் வேண்டுமாயின் முதல் உறுப்பைத் தேர்ந்தெடுத்த பின்னர் அதைத் திருப்பி, குடத்தில் (இயந்திரத்தில்) வைத்துவிடத் தேவையில்லை. ஆகவே ஓர் உறுப்பைத் தேர்ந்தெடுத்த பின்னர் முழுமைத் தொகுதியில் (குடத்தில்) $(N-1)$ உறுப்புகளே இருக்கும். இரண்டாம் உறுப்பைத் தேர்ந்தெடுத்த பின்னர் குடத்தில் $(N-2)$ உறுப்புகளே இருக்கும்.

(ஆ) ராண்டம் எண் பட்டியல் முறை (Method of Random Number Tables)

சார்பிலா முறையில் எண்கள் எடுக்கப்பட்டு அவை பட்டியலாகத் தரப்பட்டிருக்கின்றன. இவ்வெண்கள் எடுப்பதற்கு இயந்திரங்களையே பெரிதும் உபயோகித்திருக்கின்றனர். புள்ளியியல் அட்டவணைகளில் (Statistical Tables) ராண்டம் எண் பட்டியலும் கொடுக்கப்பட்டிருக்கும். பின்வரும் பட்டியல் ராண்டம் எண் பட்டியல்களை யார் யார் தொகுத்திருக்கின்றனர் என்பதை விவரிக்கிறது.

(i) Tippett's (1927) random number tables cnrsisting of 41,600 random digits grouped into 10,400 sets of four-digited random numbers.

1927-ல் டிப்பெட் என்பவர் தயாரித்த, நான்கு நான்காக வகைப்படுத்தப்பட்ட 10,400 கணங்கள் கொண்ட 41,600 ராண்டாம் எண்கள்.

(ii) Fisher & Yates (1938) table of random numbers with 15,000 random digits arranged into 1,500 sets of ten-digited, random numbers.

1938-ல் ஃபிஷரும், யேட்ஸும் தயாரித்த பத்து பத்தாக 1,500 கணங்கள் கொண்ட 15,000 ராண்டம் எண்கள்.

(ii) Kendall & B. B. Smith (1936) table of random numbers haing 1,00,000 random digits grouped into 2,50,02 sets of foue-digited randoms numbers.

1939-ல் கெண்டாலும், பி. பி. ஸ்மித்தும் தயாரித்த நான்கு நான்காக 25,000 கணங்கள் கொண்ட 1,0),000 ராண்டம் எண்கள்.

(iv) Rand corporation (1955) table of random numbes consisting of 10,00,000 random digits grouped into 2,00,000 sets of five-digited random numbers.

ராண்ட் கார்ப்பரேஷன் 1955-ல் தயாரித்த ஐந்தைந்தாக 2,00,000 கணங்கள் கொண்ட 1,00,0000 ராண்டம் எண்கள்.

(v) C. R. Rao; Mitra and Matthai (1966) table of random numbers inth 20,000 random digits arranged into 5,000 sets of four-digited random numbers.

1966-ல் சி. ஆர். ராவ், மித்ரா, மத்தாய் முதலியோர் தயாரித்த நான்கு நான்காக 5 000 கணங்கள், கொண்ட 20,000 ராண்டம் எண்கள்.

மேற்கூறப்பட்ட ராண்டம் எண்கள் பட்டியல் சீட்டுகளைக் குடத்திலிட்டுக் குலுக்கி மாதிரிகளை எடுப்பதைத் தவிர்க்கிறது பொதுவாக மாதிரியின் பருமன் (sample size) அதிகமாக இருந்தால் குலுக்கும் முறையை உபயோகிப்பது நடைமுறையில்

சாத்தியமாகாது. இதனுல் பொருள் விரயமும் கால விரயமும் ஏற்படும். ஆனுல் ராண்டம் எண்களிலிருந்து எண்களை எடுத்தல் குலுக்குவதற்கொப்பாகும். நடைமுறையிலும் இம்முறை மிகவும் எளியது. இனி இம்முறையில் மாதிரியை எவ்வாறு எடுப்பது என்பதைக் கவனிப்போம்.

முழுமைத் தொகுதியில் 1000 உறுப்புகள் இருப்பதாகக் கொள்வோம். இவ்வுறுப்புகளுக்கு 000, 001; ......, 999 என்று வரிசையாக எண்களைக் கொடுக்கலாம். பிறகு ராண்டம் எண்கள் பட்டியலிலிருந்து மூன்றிலக்க ராண்டம் எண்களை ஏதாவதொரு இடத்திலிருந்து வரிசையாக, நமக்கு மாதிரியில் எத்தனை உறுப்புகள் வேண்டுமோ அத்தனை, எண்களைக் குறித்துக் கொள்ளவேண்டும். இவ்வெண்களின் தொடர்பு கொண்ட உறுப்புகளை மாதிரியில் சேர்த்துக்கொள்ளவேண்டும், இம் முறையில் ஈ. செ. சா. ரா. மா. ஒன்று கிடைக்கும்.

ஈ. செயா. சா. ரா. மா. வேண்டுமாயின் ராண்டம் எண்கள் பட்டியலிலிருந்து வரிசையாக வெவ்வேறு எண்களை நம் தேவைக் கேற்ப எடுத்துக்கொள்ளவேண்டும். அதாவது ஒரே எண் மறுமுறை பட்டியலில் வந்தால் அதை ஒதுக்கிவிட்டு அடுத்து, வேறு எண்ணுக்குச் சென்றுவிடவேண்டும்.

மேற்கூறிய முறையில் உள்ள இன்னல்களைப் பின்வரும் உதாரணங்களில் கூறப்படுவது போலத் தவிர்க்கலாம்.

உதாரணம் 1

முழுமைத் தொகுதியில் 250 உறுப்புகள் உள்ளன. 10 உறுப்புகள் கொண்ட ஈ. செ. சா. ரா. மா. ஒன்று தேவைப்படுகிறது.

முழுமைத் தொகுதியில் 250 உறுப்புகள் இருப்பதால் மூன்றிலக்க ராண்டம் எண்களையே நாம் எடுக்கவேண்டும். ஆனுல் மூன்றிலக்க ராண்டம் எண்களில் கிட்டத்தட்ட முக்கால் பங்கு 250-க்கு மேல் இருக்கும். இவைகளையெல்லாம் நிராகரிப்பது எனத் தீர்மானித்தால் 10 எண்கள் எடுப்பதற்கு சுமார் 4 ராண்டம் எண்களை வரிசையாக எழுதிக்கொள்ள நேரிடும் (கிட்டத்தட்ட 30 எண்கள் 250 ஐவிட மேலாக இருக்கும் என்பதால்) ஆகவே 250-க்கு மேற்பட்ட எண்களை ஒதுக்குவது சிக்கனமாகாது.

ஆகவே பின்வருமாறு 10 உறுப்புகள் எடுக்கலாம்.

முழுமைத் தொகுதியிலுள்ள உறுப்புகளுக்கு 000,001, ... ,249 என வரிசையாக எண்கள் இடப்பட்டிருந்தால் 10 எண்களை மூன்றிலக்க ராண்டம் நம்பர் பட்டியலிலிருந்து எடுக்கவும். வரும் எண் 249 ஐவிட அதிகமாக இருந்தால் அவ்வெண்ணை 250 ஆல் வகுத்து மீதியை மட்டும் எடுத்துக்கொள்ளவும். இவ்விதம் 250 ஐ விடக் குறைவான எண்கள் 10 கிடைக்கும்.

உதாரணம் 2

முழுமைத் தொகுதியில் 320 உறுப்புகள் உள்ளன. 10 உறுப்புகள் கொண்ட ஈ. செ. சா ரா. மா. ஒன்று தேவைப் படுகிறது.

முழுமைத் தொகுதியிலுள்ள உறுப்புகளுக்கு 000,001, ... ,319 என வரிசையாக எண்கள் இடவும். 959-க்கு மேலேயுள்ள மூன்றிலக்க எண்கள் ராண்டம் எண் பட்டியலில் வந்தால் அவைகளை ஒதுக்கிவிடவும். இவ்வாறு 960-க்குக் குறைவான 10 எண்களை ராண்டம் எண் பட்டியலிருந்து எடுக்கவும். வந்த எண் 319-க்கு மேல் இருந்தால் அதை 320 ஆல் வகுத்து மீதியை மட்டும் எடுத்துக்கொள்ளவும். இவ்விதம் 320-க்கும் குறைவான 10 எண்களை எடுக்கவும்.

இம்முறையில் சுமார் ·04 நிகழ்திறனே ஒதுக்கப்படும் எண் களுக்கு இருப்பதால் மிகக்குறைவான எண்களே ஒதுக்கப்பட வேண்டியிருக்கும்.

16·5 ஈ. செயா. சா. ரா. மா. : ஈடு செய்யப்படாத சாதாரண ராண்டம் மாதிரிமுறையில் ஓர் உறுப்பு எடுக்கப்பட்ட பின்னர் அடுத்த உறுப்பு எடுப்பதற்கு முன்னர் முதல் உறுப்பை முழுமைத் தொகுதியில் திருப்பிவைக்கவேண்டிய அவசியமில்லை. இத் தகைய மாதிரியை எவ்வாறு எடுப்பது என்பதை முன்னமேயே பார்த்தோம்.

இம்மாதிரியில் ஒருறுப்பு இருமுறை வர முடியாது. மேலும் மாதிரியில் உள்ள எந்த இரண்டு உறுப்பை எடுத்துக் கொண்டாலும் அவை சார்பிலாமல் இருக்க முடியாது. ஆகவே n உறுப்புகள் கொண்ட ஈ. செயா. சா. ரா. மா ஆனது முழுமைத் தொகுதிப் பரவல் ராண்டம் மாறியின் படிகள் தாமென்றாலும் சார்புள்ள படிகளாகும். ஆகவே $f(x_1, x_3, x_n) = \prod_{i=1}^{n} f(x_i)$ என்று எழுத முடியாது.

உதாரணம்

முழுமைத் தொகுதியில் (2, 7, 9, 12) என்று நான்கு உறுப்புகள் இருப்பதாகக் கொள்வோம். ஈருறுப்புள்ள மாதிரி ஈ. செயா. சா. ரா மா. முறையில் எடுத்தால் பின்வருவன அம்மாதிரிகளின் முழுப் பட்டியலாகும்.

(2, 7), (2, 9), (2, 12)

(7, 2), (7, 9), (7, 12)

(9, 2), (9, 7), (9, 12)

(12, 2), (12, 7), (12, 9)

இவை ஒவ்வொன்றிற்கும் நிகழ்திறம் $\frac{1}{12}$ ஆகும். (2, 2) (7, 7), (9, 9), (12, 12) ஆகியவை இப்பட்டியலில் இடம் பெறவில்லையென்பது குறிப்பிடத்தக்கது.

பொதுவாக முழுமைத் தொகுதியில் N உறுப்புகள் இருந்தால் n உறுப்புகள் கொண்ட மாதிரிகளின் மொத்த எண்ணிக்கையானது

$NP_n = N(N-1) \ldots\ldots (N-n+1)$ ஆகும்.

ஒவ்வொரு மாதிரியும் $(X_{i_1}, X_{i_2}, \ldots\ldots, X_{i_n})$ என்றிருக்கும் $i_1, i_2, \ldots\ldots, i_n$ என்பன, n வெவ்வேறு $(1, 2, \ldots\ldots, N)$-க்குள் அடங்கிய எண்களாகும்.

$$P[\underset{\sim}{x} = (Xi_1, Xi_2, \ldots\ldots Xi_n)] = \frac{1}{NP_n}$$

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

கவனிக்கவும்

$$P[x_1 = X_1, x_2 = X_2] = \frac{1}{N(N-1)}$$

ஆனல் $P(x_1 = X_1) = \frac{1}{N}$, $P(x_2 = X_2) = \frac{1}{N}$

$\therefore \quad P[x_1 = X_1, x_2 = X_2] \neq P[x_1 = X_1, x_2 = X_2]$,

$\therefore \quad x_1, x_2$, என்பவை சார்புள்ளனவை.

16·6. தொடர்ச்சியான பரவலிலிருந்து சா. ரா. மா. எடுத்தல்

X எனும் ராண்டம் மாறியின் பரவல் சார்பு $F(x)$ என்று எடுத்துக்கொள்வோமேயானால் n உறுப்புகள் கொண்ட சா.ரா.மா. ஒன்றைப் பின்வரும் முறையில் எடுக்கலாம். ராண்டம் எண்கள் பட்டியலிலிருந்து n எண்கள் எடுக்கவும். அவ்வெண்களுக்கு முன் ஒரு தசமப்புள்ளியை வைத்து அவைகளை 0-லிருந்து 1 வரையுள்ள ஓர் எண்ணாக மாற்றவும். உதாரணமாக ராண்டம் எண்கள் 759, 328, 020, 001, 901 என்றிருந்தால் அவைகளை 0·759, 0·328, 0·020, 0·001, 0·901 என்று மாற்றவும். r என்பது இத்தகைய எண்ணானால்,

$F(x) = r$ என்னும் சமன்பாட்டைத் தீர்ப்பதன் மூலம் x ஐக் கண்டு பிடிக்கலாம். இதையே $x = F^{-1}(r)$ எனக் குறிப்பிடலாம். ஆனால் இவ்வாறு குறிப்பிடுவதற்கு $F(x)$ எனும் சார்பு கட்டாயமாக ஒரே சீராக ஏறும் (strictly monotonic increasing) சார்பாக இருக்கவேண்டும். $r_1, r_2, \dots, r_n$ என்பவை தசமப்புள்ளி வைக்கப்பட்ட எண்களானால் சாதாரண ராண்டம் மாதிரி $F^{-1}(r_1), F^{-1}(r_2), \dots, F^{-1}(r_n)$ ஆகும்.

நிறுவல்

$f(r) = 1$. ஏனென்றால் r என்பது 0-லிருந்து 1 வரையுள்ள செவ்வகப் பரவலாகும்.

$$\therefore \quad f(x) = \frac{1}{b-a} = \frac{1}{1-0} = 1.$$

$$\therefore \quad F_1(r) = \int_0^r 1 \, dr = r$$

$F_1(r)$ என்பது r-ன் பரவல் சார்பாகும்.

$$P[X \leq x] = P[X \leq F^{-1}(r)] = F F^{-1}(r) = r.$$

ஆனால் $P[X \leq x] = F(x)$. $\quad \therefore \quad F(x) = r$.

$\therefore \quad F(x) = r$ என்னும் சமன்பாட்டைத் தீர்ப்பதன் மூலம் கிடைக்கும் x-ன் பரவல் படி $F(x)$ என்பது தெரிகிறது.

துணை முடிவு

$F(x)$ எனனும ராணடம மாறியின பரவல $[0\ 1]$ககுள அமைநத செவவகப பரவலாகும

ஏனெனில $P[F(x) \leq x] = P[X \leq F^{-1}(x)]$
$$= F F^{-1}(x) = x$$

$F(X)$ எனும ராணடம மாறியின பரவல சாாபு x ஆகும மேலும $0 \leq F(X) \leq 1$ $F(X)$ ன பரவல $[0\ 1]$ ல் அமைநதுளள செவவகப பரவலாகும

உணமையின ராணடம எண்களையெலலாம செவவகப பரவல ஒனறின மதிபபுகள் எனறே கொளளலாம ஆகவே அவைகளுககு ஒரு தசமப புளளியை வைபபதனுல அவை $[0\ 1]$ ல அமைநதுளள செவவகப பரவல ஆகிவிடுகினறன இப பணபையே நாம தொடாசசியான பரவலகளிலிருநது சா ரா மா ஐ எடுபபதறகு உபயோகபபடுததுகிறோம

16 7 ஓா அலைவுப பரவலின மூலம சா ரா மா எடுததல்

ஓா அலைவுப பரவல பினவருமாறு இருககடடும

X	x_1	x_2	x_3	..	x_n	மொததம
f	f_1	f_2	f_3		f	N

இபபரவலிலிருநது சா ரா மா எடுககவேண்டுமானுல முதறபடியாக குவிவுப பரவல சாாபை (Cumulative distribution function) பினவருமாறு அமைததுக கொளளவும

$$F(x_1) = \frac{f_1}{N} \qquad F(x_2) = \frac{f_1 + f_2}{N} \qquad F(x_3) = \frac{f_1 + f_2 + f_3}{N}$$

$$F(x\ ) = \frac{f_1 + f_2 + \cdots + f_n}{N} = 1$$

X	x_1	x_2	x_3		x_n
f	f_1	f_2	f_3		f_n
F	$F(x_1)$	$F(x_2)$	$F(x_3)$		1

ராண்டம் எண்கள் பட்டியலிலிருந்து எண்ணொன்றை எடுத்து அதற்கு முன் ஒரு தசமப் புள்ளியை வைத்துக் கொள்ளவும். இவ்வெண் r என்ருல் இதற்குரிய மாதிரியைப் பின்வரும் சமனின்மை (Inequality) மூலம் கண்டுபிடிக்கலாம்.

$F(x_i) < r \leq F(x_{i+1})$ என்ருல் x_{i+1}— ஐ மாதிரியாகக் கொள்ள வேண்டும். $0 \leq r \leq F(x_1)$ என்றிருந்தால் x_1 ஐ மாதிரியாகக் கொள்ளவேண்டும்.

உதாரணம்

$n = 8$, $p = \frac{1}{2}$ உள்ள ஈருறுப்புப் பரவலிலிருந்து 5 உறுப்புள்ள சா. ரா. மா. ஒன்றையெடு.

$$P[x = r] = \binom{n}{r} \left(\frac{1}{2}\right)^r \left(\frac{1}{2}\right)^{n-r} = \binom{8}{r} \frac{1}{2^8}$$

$$= \binom{8}{r} \frac{1}{256}$$

r	0	1	2	3	4	5	6	7	8
P	$\frac{1}{256}$	$\frac{8}{256}$	$\frac{28}{256}$	$\frac{56}{256}$	$\frac{70}{256}$	$\frac{56}{256}$	$\frac{28}{256}$	$\frac{8}{256}$	$\frac{1}{256}$
$F(r)$	$\frac{1}{256}$	$\frac{9}{256}$	$\frac{37}{256}$	$\frac{93}{256}$	$\frac{163}{256}$	$\frac{219}{256}$	$\frac{247}{256}$	$\frac{255}{256}$	1

ராண்டம் எண்கள் பட்டியலிலிருந்து எடுக்கப்பட்ட எண்கள் பின்வருவன. 0702, 7177, 5280, 4552, 6257

இவைகளுக்குத் தசமப்புள்ளியை வைத்தால் 0·0702, 0·7177, 0·528, 0·4552, 0·6257 ஆகின்றன.

$\frac{9}{256} < 0{\cdot}0702 < \frac{37}{256}$ $\therefore$ 2 ஒரு மாதிரி உறுப்பு

$\frac{163}{256} < 0{\cdot}7177 < \frac{219}{256}$ $\therefore$ 5 ஒரு மாதிரி உறுப்பு

$\frac{93}{256} < 0{\cdot}5280 < \frac{163}{256}$ $\therefore$ 4 ஒரு மாதிரி உறுப்பு

$\frac{93}{256} < 0{\cdot}4552 < \frac{169}{256}$ $\therefore$ 4 ஒரு மாதிரி உறுப்பு

$\frac{93}{256} < 0{\cdot}6257 < \frac{163}{256}$ $\therefore$ 4 ஒரு மாதிரி உறுப்பு

ஆகவே (2, 5, 4, 4, 4) என்பது மேலே தரப்பட்டுள்ள ஈருறுப்புப் பரவலிலிருந்து எடுக்கப்பட்ட 5 உறுப்புகள் கொண்ட மாதிரியாகும்.

பயிற்சி 1

1. x என்னும் ராண்டம் மாறி 1, 0 என்னும் எண்களை முறையே $p, q\ (= 1 - p)$ என்னும் ஊக அளவைகளுடன் ஏற்கிறது. $E(x) = p$, $V(x) = pq$ என்பனவற்றை நிரூபி. y என்னும் ராண்டம் மாறி 1, 0 என்னும் எண்களை முறையே p', q $(= 1 - p')$ என்னும் ஊக அளவைகளுடன் ஏற்றால் பின் வரும் ராண்டம் மாறிகளின் எதிர்பார்க்கும் மதிப்பையும், பரவற்படியையும் கண்டுபிடி (i) $x + y$, (ii) $x - y$, (iii) xy.

விடை

(அ) x : 0 1
p : q p

$\therefore\ E(x) = 0{\cdot}q + 1{\cdot}p = p.$

x^2	0	1
p	q	p

$$\therefore \quad E(x^2) = 0\cdot q + 1\cdot p = p.$$

$$\therefore \quad V(x) = E(x^2) - [E(x)]^2 = p - p^2 = p(1-p)$$
$$= pq.$$

(ஆ) $E(x+y) = E(x) + E(y) = p + p'$

$E(x-y) = E(x) - E(y) = p - p'$

$E(x-y) = E(x)\ E(y) = pp'$
[x, y சார்பற்றவை எனக் கொண்டால்]

$V(x+y) = V(x) + V(y) = pq + p'q'$ [,, ,, ,,]

$V(x-y) = V(x) + V(y) = pq + p'q'$ [,, ,, ,,]

$$V(xy) = E(x^2 y^2) - [E(xy)]^2$$
$$= E(x^2)\,E(y^2) - [E(x)\,E(y)]^2$$
$$= pp' - (pp')^2 = pp'(1-pp')$$ [,, ,, ,,]

2. $\theta_1, \theta_2, \ldots \theta_k$ என்னும் முழுமைத் தொகுதியின் அளவைகளை முறையே $t_1, t_2, \ldots t_k$, என்னும் மாதிரியளவைகள் பிறழ்ச்சியின்றி மதிப்பிட்டால் $\sum_{i=1}^{k}$ [a_i-மாறிலிகள்] என்னும் மாதிரியளவையின் எதிர்பார்க்கும் மதிப்பையும் பரவற் படியையும் கண்டுபிடி.

விடை

$$E\left[\sum_{i=1}^{k} a_i\ t_i\right] = \sum_{i=1}^{k} a_i\,E(t_i) = \sum_{i=1}^{k} a_i\,\theta_i$$

$$V\left[\sum_{i=1}^{k} a_i\, t_i\right] = E\left[\sum_{i=1}^{k} a_i\, t_i - \sum_{i=1}^{k} a_i\, \theta_i\right]^2$$

$$= E\left[\sum_{i=1}^{k} a_i\,(t_i - \theta_i)\right]^2$$

$$= \sum_{i=1}^{k} a_i^2\, E\,(t_i - \theta_i)^2$$

$$+ 2 \sum\sum_{i<j} a_i\, a_j\, E\,(t_i - \theta_i)\,(t_j - \theta_j)$$

$$= \sum_{i=1}^{k} a_i^2\, \sigma_i^2 + 2 \sum\sum_{i<j} a_i\, a_j\, \sigma_{ij}$$

$[\ \sigma_i = V(t_i) + \sigma_{ij} = \text{cov}\,(t_i,\, t_j)\ ]$

3. $t_1, t_2, \ldots\ldots t_k$ என்பவை θ என்னும் முழுமைத் தொகுதி யளவையின் சார்பற்ற பிறழ்ச்சியற்ற மாதிரியளவைகளானல்

$$t = \sum_{i=1}^{k} \omega_i\, t_i \quad \left[\sum_{i=1}^{k} \omega_i = 1\right]$$ என்பதும் θ ஐப் பிறழ்ச்சியின்றி

மதிப்பிடும் என்று நிரூபி. $V\,(t)$ சிறுமமாக (மிகவும் குறைவாக) இருக்கவேண்டுமானல் ω_i-களின் மதிப்புகளைக் கண்டுபிடி.

விடை

$$E\left[\sum_{i=1}^{k}\omega_i t_i\right] = \sum_{i=1}^{k}\omega_i E(t_i) = \sum_{i=1}^{k}\omega_i \theta$$

$$= \theta\sum_{i=1}^{k}\omega_i = \theta$$

$\therefore \sum_{i=1}^{k}\omega_i t_i$ என்பது θ-ன் பிறழ்ச்சியற்ற மதிப்பீட்டி.

$$V(t) = V\left[\sum_{i=1}^{k}\omega_i t_i\right] = \sum_{i=1}^{k}\omega_i^2 \sigma_i^2 \qquad V(t_i) = \sigma_i^2$$

$V(t)$ ஐச் சிறுமமாக்குவதற்கு லெக்ராஞ்சியின் முறையைப் பின்பற்ற வேண்டும். [சுருக்கமாக லெக்ராஞ்சியின் முறையென்பது பின்வருமாறு: $F(x_1, x_2, \ldots, x_n)$ என்னும் சார்பை சிறுமமாகவோ, பெருமமாகவோ $9_1(x_1, \ldots\ldots x_n) = 0$, $9_2(x_1\ x_2 \ldots \ldots x_n) \ldots\ldots 9_k(x_1, x_2, \ldots\ldots x_n) = 0$, $k < n$ என்னும் நிபந்தனைகளுக்குட்பட்டுச் செய்யவேண்டுமாயின் $S = F + \lambda_1\ 9_1 + \lambda_2\ 9_2 + \ldots + \lambda_k 9_k$ என்று எடுத்துக்கொள்ளவும். $\lambda_1, \lambda_2, \ldots \lambda_k$ என்பவை லெக்ராஞ்சியின் பெருக்குமெண்கள் (Lagrangean multipliers) என அழைக்கப்படும் மாறிலிகளாகும்.

$$\frac{\partial s}{\partial x_1} = 0,\ \frac{\partial s}{\partial x_2} = 0, \ldots, \frac{\partial s}{\partial x_n} = 0,\ 9_1 = 0,\ 9_2 = 0, \ldots 9_k = 0$$

என்னும் $n + k$ சமன்பாடுகளைத் தீர்த்து $x_1, x_2, \ldots x_n$, $\lambda_1, \lambda_2, \ldots, \lambda_k$ என்பவைகளைக் கண்டுபிடிக்கலாம்.]

$$S = V(t) + \lambda \left[\sum_{i=1}^{k} \omega_i - 1 \right] = \sum_{i=1}^{k} \omega_i^2 \sigma_i^2$$

$$+ \lambda \left[\sum_{i=1}^{k} \omega_i - 1 \right]$$

$$\frac{\partial s}{\partial \omega_i} = 2\omega_i \sigma_i^2 + \lambda = 0 \quad \therefore \ -\frac{\lambda}{2} = \omega_i \sigma_i^2$$

$$\therefore \ \omega_1 \sigma_1^2 = \omega_2 \sigma_2^2 = \omega_3 \sigma_3^2 = \dots = \omega_k \sigma_k^2 = -\frac{\lambda}{2}$$

$$\therefore \ 1 = \sum_{i=1}^{k} \omega_i = \sum_{i=1}^{k} -\frac{\lambda}{2} \frac{1}{\sigma_i^2} \quad \therefore \ -\frac{\lambda}{2} = \frac{1}{\sum_{i=1}^{k} \frac{1}{\sigma_i^2}}$$

$$\therefore \ \omega_1 = -\frac{\lambda}{2} \frac{1}{\sigma_1^2} = \frac{1}{\sigma_1^2} \div \sum_{i=1}^{k} \frac{1}{\sigma_i^2},$$

$$\omega_2 = \frac{1}{\sigma_2^2} \div \sum_{i=1}^{k} \frac{1}{\sigma_i^2} \dots \dots,$$

$$\omega_k = \frac{1}{\sigma_k^2} \div \sum_{i=1}^{k} \frac{1}{\sigma_i^2}$$

4. பின்வரும் அளவெடுப்புகளில் முழுமைத் தொகுதி, உறுப்பு, உறுப்பினின்று எடுக்கப்படவேண்டிய விவரம் முதலிய வைகளைக் குறிப்பிடுக.

(அ) நகரங்களின் கட்டட வேலைகளைப் பற்றிய அளவெடுப்பு.

(ஆ) ஒரு குறிப்பிட்ட மாவட்டத்தில் மாம்பழ உற்பத்தியைப் பற்றிய அளவெடுப்பு.

(இ) சென்னை நகரத்தில் மக்கள் உபயோகிக்கும் பண்டங்களின் சில்லறை விலைகளைப் பற்றிய அளவெடுப்பு.

(ஈ) அரசுடைமைப் போக்குவரத்தின் மொத்த அளவைப் பற்றிய அளவெடுப்பு.

(உ) சென்னையிலிருந்து மாமல்லபுரம் வரையுள்ள கடற்கரைப் பகுதியில் பிடிக்கப்படும் மீன்களின் அளவைப் பற்றிய அளவெடுப்பு.

(ஊ) புதிதாகச் சென்னை வானொலியில் ஒலி பரப்பப்படும் ஒரு நிகழ்ச்சியைப் பற்றிய நேயர்களின் கருத்தைப் பற்றிய அளவெடுப்பு.

(எ) செங்கற்பட்டு மாவட்டத்தில் தொழுநோயின் கடுமையைப் பற்றிய அளவெடுப்பு.

(ஏ) சென்னை நகர மக்களின் வருமானப் பரவலைப் பற்றிய அளவெடுப்பு.

5. பின்வரும் பரவலிலிருந்து 10 உறுப்புகள் கொண்ட சாதாரண ராண்டம் மாதிரியை எடுக்கவும்.

x	2	4	6	8	10	12	14	16	18	20	22	24	26	28	30
f	23	34	39	47	62	79	92	99	83	64	60	41	37	29	21

6. பின்வரும் எண்களிலிருந்து (அ) ஈ. செ. சா. ரா. மா. (ஆ) ஈ. செயா. சா. ரா. மா. முறைகளில் 10 உறுப்புகளைக் கொண்ட மாதிரிகளை எடுக்கவும்.

106	92	84	90	107	86	99	107	76	70
113	81	82	126	204	131	109	68	111	75
107	130	141	84	115	129	136	104	93	139
123	110	187	90	80	95	115	115	118	123
128	98	82	125	100	110	111	186	78	119

7. பின்வரும் பரவல்களிலிருந்து 5 உறுப்புகள் கொண்ட சா. ரா. மா. எடுக்கவும்.

(அ) $f(x) = \frac{1}{(1+x)^2} \qquad 0 \leq x \leq \infty$

(ஆ) பாரெடோ பரவல் (Pareto's Distribution)

$$f(x) = \frac{\alpha}{\beta}\left(\frac{\beta}{x}\right)^{\alpha+1} \qquad \beta \leq x \leq \infty$$

$$(இ)\ f(x) = \begin{cases} \frac{1}{16}(3+x)^2 & -3 \leq x \leq -1 \\ \frac{1}{16}(6-2x^2) & -1 \leq x \leq 1 \\ \frac{1}{16}(3-x)^2 & 1 \leq x \leq 3. \end{cases}$$

பிறழ்ச்சியற்ற நாணயம் ஒன்றை நூறு முறை சுண்டிவிட்டதில் பின்கண்ட முடிவுகள் குறிக்கப்பட்டன. 'த' என்பது தலையைக் குறிக்கிறது.

பூ பூ த த பூ	த த பூ த த	த பூ பூ த த	த பூ பூ த பூ
பூ பூ பூ த பூ	பூ பூ த பூ த	பூ பூ த த த	த பூ பூ த த
த த பூ த பூ	பூ த பூ பூ பூ	பூ த த பூ பூ	பூ த த த பூ
த த த த த	பூ த பூ பூ பூ	பூ த பூ பூ த	பூ த த பூ த
த பூ பூ த பூ	த த பூ த பூ	பூ பூ த பூ பூ	த பூ பூ த த

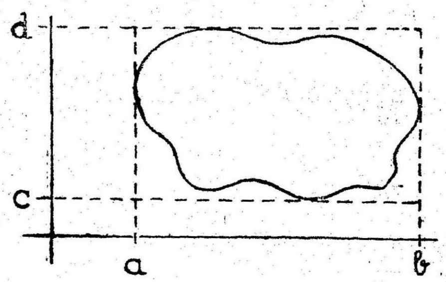
மேலே கொடுக்கப்பட்டவைகளை உபயோகித்து, 23 கிராமங்களிலிருந்து 4 கிராமங்களை எவ்வாறு ஈ. செயா. சா. ரா. மா. முறையில் எடுக்க முடியும் என்பதை விளக்குக.

9. 4 அங்குல ஆரமுள்ள ஒரு வட்டத்தினுள் ஒரு புள்ளியை எப்படி ராண்டம் முறையில் எடுப்பாய்?

விடை

பொதுவாக ஒரு வரைப்படத்தினுள்ளிருந்து ஒரு புள்ளியை ராண்டம் முறையில் எடுப்பதற்குப் பின்வரும் முறையைப் பின்பற்றலாம்.

வரைப்படத்தில் x அச்சுத் தூரங்கள் a-லிருந்து b வரை யென்றும், y அச்சுத் தூரங்கள் c-லிருந்து d வரையென்றும் எடுத்துக் கொண்டால்



x_1 எனும் எண்ணை $[a, b]$-லிருந்தும் y_1 எனும் எண்ணை $[c, d]$-லிருந்தும் ராண்டம் முறையில் எடுக்க வேண்டும். (x_1, y_1) எனும் புள்ளி வரைப்படத்தினுள்ளே இருந்தால் இதையே நம் புள்ளியாகக் கொள்ளலாம். இல்லையேல் மறுபடியும் இதே முறையில் (x_2, y_2) எனும் புள்ளியைத் தேர்தெடுத்து வரைப் படத்தினுள் இருக்கிறதாவெனப் பார்க்க வேண்டும். ஒரு புள்ளி வரைப்படத்தினுள்ளே அமையும்படி கிடைக்கும்வரை இம் முறையைத் தொடர்ந்து உபயோகிக்கவேண்டும்.

மேற்கண்ட முறையை வட்டத்திற்கு உபயோகப்படுத்தவும்.

10. ஈ. செயா. சா. ரா. மா. முறையில் ஓர் அகராதியிலிருந்து n வார்த்தைகள் எடுப்பதற்குப் பின்வரும் முறை உபயோகிக்கப் பட்டது.

முதலில் ஒரு பக்கத்தைச் சம வாய்ப்புடன் தேர்ந்தெடுத்தனர். k என்பது ஒரு பக்கத்திலுள்ள வார்த்தைகளின் பெருமம் (maximum number of words in a page) என்றால் ஒன்றிலிருந்து k வரையுள்ள எண்களிலிருந்து ஓர் எண் ராண்டம் முறையில் தேர்ந் தெடுக்கப்பட்டது. இது r_1 எனக் கொண்டால் தேர்ந்தெடுத்த பக்கத்திலிருந்து r_1 ஆவது வார்த்தை மாதிரியாகத் தேர்ந்தெடுக்

கப்பட்டது. அப்பக்கத்தில் r_1 வார்த்தைகள் இல்லையென்றால் மறுபடியும் ஒன்றிலிருந்து k வரையுள்ள எண்களிலிருந்து r_2 எனும், எண் ராண்டம் முறையில் தேர்ந்தெடுக்கப்பட்டு, r_2 ஆவது வார்த்தை தேர்ந்தெடுக்கப்பட்டது. r_2 வார்த்தைகள் அப் பக்கத்தில் இல்லையெனில் மறுபடியும் ஓர் எண் r_3 என்பது தேர்ந் தெடுக்கப்பட்டது. இம் முறையில் ஒரு வார்த்தை தேர்ந்தெடுக் கப்பட்ட பக்கத்தில் கிடைக்கும் வரை ராண்டம் எண்கள் எடுக்கப் பட்டன.

மேற்கூறிய முறையில் n உறுப்புகள் எடுக்கப்பட்டன. இம் முறையில் எல்லா வார்த்தைகளுக்கும் சம வாய்ப்புக் கிடைக்கும் என்று நிரூபி.

17. ஈடு செய்யப்பட்ட சாதாரண ராண்டம் மாதிரிகளில் மதிப்பீடு

(ESTIMATION IN SIMPLE RANDOM SAMPLES WITH REPLACEMENT)

17·1. $X_1, X_2, \ldots, X_N$ என்பவை முழுமைத்தொகுதி உறுப்புகள் எனக் கொள்வோம்.

(அ) $\bar{X} = \frac{1}{N}\sum_{i=1}^{N} X_i$ என்பது முழுமைத் தொகுதியின் சராசரி.

(ஆ) $\sigma^2 = \frac{1}{N}\sum_{i=1}^{N} (X_i - \bar{X})^2 = \frac{1}{N}\sum_{i=1}^{N} X_i^2 - \bar{X}^2$ என்பது முழுமைத் தொகுதியின் பரவற் படி.

(இ) $X = \sum_{i=1}^{N} X_i$ என்பது முழுமைத் தொகுதியின் மொத்தம்.

$(x_1, x_2, \ldots, x_n)$ என்பவை n உறுப்புகள் கொண்ட ஈ. செ. சா. ரா. மா. எனக் கொள்வோம்.

(ஈ) $\bar{x} = \frac{1}{n}\sum_{i=1}^{n} x_i$ என்பது மாதிரிச் சராசரி.

(உ) $s^2 = \frac{1}{n-1}\sum_{i=1}^{n}(x_i - \bar{x}^2)$ என்பது மாதிரிப் பரவற் படி.

(ஊ) $x = \sum_{i=1}^{n} x_i$ என்பது மாதிரி மொத்தம்.

தேற்றம் 17·2.

ஈடு செய்யப்பட்ட சாதாரண ராண்டம் மாதிரியின் சராசரி முழுமைத் தொகுதியின் சராசரியின் பிறழ்ச்சியற்ற மதிப்பீட்டியாகும். அதாவது $E(\bar{x}) = \bar{X}$ அல்லது $\bar{x} \sim E^{-1}(\bar{X})$.

நிறுவல்

x_i-ன் பரவல் பின்வருமாறு.

x_i	X_1	X_2	X_3	....	X_N
P	$\frac{1}{N}$	$\frac{1}{N}$	$\frac{1}{N}$		$\frac{1}{N}$

எந்தக் கட்டத்திலும் முழுமைத் தொகுதியிலுள்ள எல்லா உறுப்புகளுக்கும் சம வாய்ப்பு அளிக்கப்படுகிறது என்பதிலிருந்து மேற்கூறியது விளங்கும்.

$$\therefore \quad E(x_i) = \sum_{i=1}^{N} X_i \cdot \frac{1}{N} = \frac{1}{N}\sum_{i=1}^{N} X_i = \bar{X}$$

$$\therefore \quad E(\bar{x}) = E\left[\frac{1}{n}\sum_{i=1}^{n} x_i\right] = \frac{1}{n}\sum_{i=1}^{n} E(x_i) = \frac{1}{n}\sum_{i=1}^{n}\bar{X}$$

$$= \frac{1}{n}\, n\,\bar{X} = \bar{X}$$

$\therefore \quad E(\bar{x}) = \bar{X}$ (நி-து.) [(நி-து.) என்பது நிறுவப்பட்டது என்பதைக் குறிக்கும்.]

17·3. தேற்றம்

ஈடு செய்யப்பட்ட சாதாரண ராண்டம் மாதிரியின் சராசரியின் பரவற் படி $\frac{\sigma^2}{n}$ ஆகும். அதாவது $V(\bar{x}) = \frac{\sigma^2}{n}$

நிறுவல்

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

[$x_1, x_2, \ldots x_n$ ஆகியவை சார்பற்றவை]

x_i^2-ன் பரவல்

x_i^2	X_1^2	X_2^2	X_3^2		X_N^2
P	$\frac{1}{N}$	$\frac{1}{N}$	$\frac{1}{N}$		$\frac{1}{N}$

$$\therefore \quad E(x_i^2) = \sum_{j=1}^{n} x_j^2 \frac{1}{N} = \frac{1}{N}\sum_{j=1}^{n} x_j^2$$

$$\therefore \quad V(x_i) = E(x_i^2) - [E(x_i)]^2 = \frac{1}{N}\sum_{j=1}^{n} X_j^2 - \bar{X}^2 = \sigma^2.$$

$$\therefore \quad V(\bar{x}) = \frac{1}{N^2}\sum_{i=1}^{n} V(x_i) = \frac{1}{n^2}\sum_{i=1}^{n} \sigma^2 = \frac{1}{n^2} n\sigma^2 = \frac{\sigma^2}{n}$$

(நி-து).

17·4. தேற்றம்

ஈடு செய்யப்பட்ட சாதாரண ராண்டம் மாதிரியின் பரவற் படியின் எதிர்பார்க்கும் மதிப்பு σ^2 ஆகும்.

அதாவது $E(s^2) = \sigma^2$

றுவல்

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

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

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

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

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

$$\left. -2 \sum_{i=1}^{n} (\bar{x} - \bar{X})^2 \right]$$

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

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

$$= \frac{1}{(n-1)} \left[n\sigma^2 - n\frac{\sigma^2}{n} \right] = \sigma^2 \quad [\text{நி-து.}]$$

அல்லது $s^2 \sim E^{-1}(\sigma^2)$

தேற்றம் 17·5.

ஈடு செய்யப்பட்ட சாதாரண ராண்டம் மாதிரி முறையில்

(அ) σ^2-ன் பிறழ்ச்சியற்ற மதிப்பீட்டியொன்று s^2 ஆகும்.

(ஆ) $V(\bar{x})$-ன் பிறழ்ச்சியற்ற மதிப்பீட்டியொன்று $\frac{s^2}{n}$ ஆகும்.

நிறுவல்

(அ) தேற்றம் 17·4-ல் $E(s^2) = \sigma^2$ என்று நிரூபித்தோம்.

$\therefore \quad s^2 \sim E^{-1}(\sigma^2)$ [நி-து.]

(ஆ) $s^2 \sim E^{-1}(\sigma^2)$

$$\frac{s^2}{n} \sim E^{-1}\left(\frac{\sigma^2}{n}\right) = E^{-1}[V(\bar{x})] \; [\text{நி-து.}]$$

அதாவது $E^{-1}(\sigma^2) \sim s^2$ & $E^{-1}[V(\bar{x})] \sim \frac{s^2}{n}$

உதாரணம்

முழுமைத் தொகுதியிலுள்ள உறுப்புகள் (2, 7, 9, 12) 2 உறுப்புகள் கொண்ட ஈ. செ. சா. ரா. மா.கள் பின்வருமாறு:

(2, 2) (2, 7) (2, 9) (2, 12)

(7, 2) (7, 7) (7, 9) (7, 12)

(9, 2) (9, 7) (9, 9) (9, 12)

(12, 2) (12, 7) (12, 9) (12, 12)

இவைகளின் சரரசரி பின்வருமாறு: $\bar{x}$:

2	$\frac{9}{2}$	$\frac{11}{2}$	7
$\frac{9}{2}$	7	8	$\frac{19}{2}$
$\frac{11}{2}$	8	9	$\frac{21}{2}$
7	$\frac{19}{2}$	$\frac{21}{2}$	2

இவைகளின் பரவற் படி பின்வருமாறு: s^2 :

0	$\frac{25}{2}$	$\frac{49}{2}$	50
$\frac{25}{2}$	0	2	$\frac{25}{2}$
$\frac{49}{2}$	2	0	$\frac{9}{2}$
50	$\frac{25}{2}$	$\frac{9}{2}$	0

முழுமைத் தொகுதியின் சராசரி $\bar{X} = \frac{1}{4} [2 + 7 + 9 + 12]$

$= 7{\cdot}5$

முழுமைத் தொகுதியின் பரவற் படி $\sigma^2 = \frac{1}{4} [4 + 49 + 81 + 144] - \frac{225}{4}$

$= \frac{53}{4} = 13{\cdot}25.$

மாதிரிச சராசரிகளின எதிரபாாககும மதிபபு $E(\bar{x}) =$

$$\frac{1}{16}\left[2+\frac{9}{2}+\frac{11}{2}+7+\frac{9}{2}+7+8+\frac{19}{2}+\frac{11}{2}+8+9+\frac{21}{2}+7+\frac{19}{2}+\frac{21}{2}+12\right]=\frac{120}{16}=7\ 5$$

$$E(\bar{x}) = \bar{X} = 7\ 5$$

மாதிரிப பரவறபடியின எதிரபாாககும மதிபபு $E(s^2) =$

$$\frac{1}{16}\left[0+\frac{25}{2}+\frac{49}{2}+50+\frac{25}{2}+0+2+\frac{25}{2}+\frac{49}{2}+2+0+\frac{9}{2}+50+\frac{25}{2}+\frac{9}{2}+0\right]=\frac{212}{16}=13\ 25$$

$$E(s^2) = \sigma^2 = 13\ 25$$

17 6 முழுமைத தொகுதியின சராசரிககு மாதிரிச சராசரியை விடத திறன அதிகமுளள பிறழசசியறற மதிபபீட்டி.

ஈ செ சா ரா மா முறையில வரும உறுபபுகள யாவையும வெவவேறானவை எனறு கூறமுடியாது ஏனெனில இமமுறையில ஒருறுபபு பலமுறை வருவதறகு வாயபபு உளளது n உறுபபுகள் கொண்ட மாதிரியொனறை எடுககுமபொழுது அவைகளில m $[m \leq n]$ உறுபபுகளதாம வெவவேறானவை எனறு கொள்வோம.

இவை x_1 x_2 x_3 .. x_m எனக கொணடால $\bar{x} = \frac{1}{m}\sum_{i=1}^{m} x$ எனபது x ஐ விட அதிகததிறன வாயநத பிறழசசியறற மதிபபீட்டியாகும

அதாவது $E(\bar{x}) = \bar{X}$

$$\infty\ V(x') \leq V(x)$$

$$E^{-1}[V(\bar{x})] = \left\{\left(\frac{1}{m}-\frac{1}{N}\right)+\frac{N-1}{N}\ \frac{}{N}\right\}$$

$$\frac{1}{m-1}\sum_{i=1}^{m}(x-x)^2$$

மேற்கூறியவறறைப பறறி அதிகம தெரிநதுகொளள வேண்டு வோர

Basu 1958) Desraj and Khamis (1958) and Pathak (196?) எனபவாகளின ஆராயசசிகளைப படிககவும

17 7 கிளைததேறறம

ஈ செ சா ரா மா முறையில

(அ) $E^{-1}(x) \sim N\bar{x}$

(ஆ) $V[E^{-1}(x)] = N^2 \frac{\sigma^2}{n}$

(இ) $E^{-1}\{V[E^{-1}(x)]\} = \frac{N^2 s^2}{n}$

நிறுவல

(அ) $E(\bar{x}) = \bar{X}$

$$= \frac{X}{N}$$

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

$$E^{-1}(x) \sim N\bar{x} \quad \text{(நி து)}$$

(ஆ) $V[E^{-1}(x)] = V(N\bar{x}) = N^2 V(\bar{x}) = \frac{N^2 \sigma^2}{n}$ (நி து)

(இ) $E^{-1}\{V[E^{-1}(X)]\} = E^{-1}\left(\frac{N^2 \sigma}{n}\right)$

$$= \frac{N^2}{n} E^{-1}(\sigma^2)$$

$$\sim \frac{N S^2}{n} \quad \text{(நி து)}$$

17 8.

இனத தொகுதியிலுளள (முழுமைத தொகுதியிலுள்ள) உறுப்புகளை ஒரு பண்பின வழி ஒரு பாகமாகப பிரிக்க முடியும எனக் கொள்வோம உதாரணமாகச் சென்னை நகரமக்களை இனத தொகுதியெனக் கொண்டால இததொகுதியைத் தமிழ பேசுவோர் தமிழ பேசாதார என இரு பாகமாகப பிரிககலாம இனததொகுதி

யில் N உறுப்புகள் இருந்தால் ஒரு குறிப்பிட்ட பண்புடைய உறுப்புகள் N_1 எனவும் மற்றவை N_2 எனவும் கொள்வோம். $N_1 + N_2 = N$ என்பது தெளிவு.

$\frac{N_1}{N}$ என்பது சார்கின்ற நிகழ்வெண் (relative frequency) என்று அக்குறிப்பிட்ட பண்பைப் பொருத்தவரை கூறலாம். அதாவது ஒரு குறிப்பிட்ட பண்பின் சார்கின்ற நிகழ்வெண் என்பது மொத்தத்தில் எவ்வளவு பாகம் இப்பண்புள்ளவை என்பதைக் குறிக்கிறது. இதைச் சுருக்கமாகப் பண்பின் பாகம் (proportion) எனக் குறிப்பிடுவோம்.

ஓர் இனத்தொகுதியில் ஒரு பண்பின் பாகம் $P\left[= \frac{N_1}{N}\right]$ எனக் கொள்வோம். இதை ஒரு மாதிரியின் வழி எவ்வாறு மதிப்பிடுவது? இதற்கு இவ்வினத்தொகுதியிலிருந்து n உறுப்புகள் கொண்ட ஈ. செ. சா. ரா. மா. ஒன்றை எடுத்து அம்மாதிரியில் எத்தனை உறுப்புகள் அப்பண்பினைக் கொண்டவை என்பதைக் கணக்கிடுவோம். n_1 உறுப்புகள் அப்பண்பினைக் கொண்டிருந்தால் $\frac{n_1}{n} = p$ என்பது மாதிரி பாகத்தைக் குறிக்கும். $\frac{N_1}{N} = P$ என்பதை இனத்தொகுதிப் பாகம் எனக் குறிப்பிடலாம்.

இனத் தொகுதியிலுள்ள உறுப்புகள் $U_1, U_2, \ldots U_N$ எனக் கொள்வோம், இவ்வுறுப்புகளுடன் $X_1, X_2, \ \ , X_N$ என்னும் எண்களைப் பின் வருமாறு தொடர்புப் படுத்துவோம்.

U_1 என்னும் உறுப்பு குறிப்பிட்ட பண்பினைக் கொண்டிருந்தால் $X_1 = 1$
இல்லையேல் $X_1 = 0$

U_2 என்னும் உறுப்பு குறிப்பிட்ட பண்பினைக் கொண்டிருந்தால் $X_2 = 1$
இல்லையேல் $X_2 = 0$

... ...

... ...

U_N என்னும் உறுப்பு குறிப்பிட்ட பண்பினைக் கொண்டிருந்தால் $X_N = 1$
இல்லையேல் $X_N = 0$

மேற்கூறியவற்றிலிருந்து இனத்தொகுதியில் பண்பின் பாகம்

$$P = \frac{N_1}{N} = \frac{\sum_{i=1} X_i}{N} = \bar{X}$$ என்பது தெரிகிறது.

இதைப் போலவே மாதிரியிலும் $x_1, x_2, ..., x_n$ என்பன வற்றை வரையறுக்கலாம்.

அதாவது $x_i = 1$, i ஆவது மாதிரியுறுப்பு குறிப்பிட்ட பண்புடையதாக இருந்தால்.

இல்லையேல் $x_i = 0$

இதிலிருந்து $P = \frac{n_1}{n} = \frac{\sum_{1=i}^{n} x_i}{n} = \bar{x}$ என்பது தெரிகிறது.

இவைகளிலிருந்து பின்வரும் தேற்றம் தெளிவு.

தேற்றம் 17·9.

ஈ. செ. சா. ரா. மா. முறையில் இனத்தொகுதிப் பண்பின் பாகம் P-க்கு மாதிரிப் பண்பின் பாகம் P பிறழ்ச்சியற்ற மதிப்பீட்டியாகும். P-ன் பரவற்படி $\frac{PQ}{n}$. $\frac{PQ}{n}$-ன் பிறழ்ச்சியற்ற மதிப்பீட்டியொன்று $\frac{pq}{(n-1)}$ ஆகும்.

$Q = 1 - P$; $q = 1 - p$

நிறுவல்: 17·8-ல் குறிப்பிட்டபடி $\bar{X} = P$; $\bar{x} = p$

(அ) $E(\bar{x}) = \bar{X} \therefore E(p) = P$ (நி-து.)

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

$[\because X_i^2 = X_i]$

$$= P - P^2 = P(1 - P) = PQ$$

$V(\bar{x}) = \frac{\sigma^2}{n}$. $\therefore V(p) = \frac{PQ}{n}$ (நி-து.)

(இ) $E^{-1}[V(\bar{x})] \sim \frac{s^2}{n}$

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

$[\because \quad x_i^2 = x_i]$

$$= \frac{1}{n-1}[np - np^2] = \frac{n}{n-1}pq$$

$$\therefore \quad \frac{s^2}{n} = \frac{pq}{n-1}$$

$$\therefore \quad E^{-1}\left(\frac{PQ}{n}\right) = E^{-1}[V(\bar{x})] \sim \frac{s^2}{n} = \frac{pq}{n-1}$$

அதாவது $E^{-1}\left(\frac{PQ}{n}\right) \sim \frac{pq}{n-1}$ (நி-து.)

உதாரணம்

சென்னை நகரின் மக்கட் தொகை 25 லட்சம். இங்கிருந்து 20 மக்கள் ஈ. செ. சா. ரா. மா. முறையில் தேர்ந்தெடுக்கப்பட்டனர். இம் மாதிரியில் 45 பேர்கள் புகை பிடிப்பவர். இதிலிருந்து சென்னை நகரில் புகை பிடிப்பவர் எவ்வளவு பாகம் என்பதை மதிப்பிடு. இம் மதிப்பீட்டின் பரவற் படியென்ன? இப் பரவற் படியின் மதிப்பீட்டியென்ன?

விடை

இங்கு $n = 200$; $n_1 = 45$

$$\therefore \quad P = \frac{45}{200} = 0{\cdot}225$$

$\therefore$ சென்னை நகரில் புகை பிடிப்பவர் 1000-ல் 225 பேர்.

$V(p) = \frac{PQ}{n} = \frac{PQ}{200}$ [P தெரியாததால் இதைக் கண்டு பிடிக்க இயலாது].

$V\ (p)$-ன் மதிப்பீட்டி $\dfrac{pq}{n-1}$ ஆகும்.

$$\therefore \quad E^{-1}\,[V\,(p)] \sim \frac{pq}{199} = \frac{0{\cdot}225 \times 0{\cdot}775}{199} = 0{\cdot}0008667$$

தேற்றம் 17·10.

இனத் தொகுதியில் ஒரு குறிப்பிட்ட பண்பையுடைய உறுப்புகளின் மொத்தத்திற்குப் பிறழ்ச்சியற்ற மதிப்பீடொன்று NP ஆகும். இதன் பரவற் படி $N^2\,\dfrac{PQ}{n}$. இப்பரவற் படியின் பிறழ்ச்சியற்ற மதிப்பீடொன்று $N^2\,\dfrac{pq}{n-1}$.

நிறுவல்

$$\frac{N_1}{N} = P \;\therefore\; N_1 = NP.$$

$$E\,(p) = P \text{ என்பதால் } E\,(Np) = NP = N_1. \qquad \text{(நி-து.)}$$

$$V\,(Np) = N^2\,V\,(p) = N^2\,\frac{PQ}{n}.\; E^{-1}\left(\frac{N^2\,PQ}{n}\right)$$

$$= N^2\,E^{-1}\left(\frac{PQ}{n}\right) \sim N^2\,\frac{pq}{n-1} \qquad \text{(நி-து.)}$$

பயிற்சி 2

1. ஒரு முழுமைத் தொகுதியில் பின்வரும் எண்கள் உள்ளன.

7 9 12 14 8

ஈ. செ. சா. ரா. மா. முறையில் 2 உறுப்புகள் கொண்ட 25 மாதிரிகளை எழுது. ஒவ்வொன்றின் மாதிரிச் சராசரியையும் கண்டு பிடி. இம்மாதிரிச் சராசரிகளின் எதிர்பார்க்கும் மதிப்பு (சராசரி) முழுமைத் தொகுதியின் சராசரிக்குச் சமம் என்பதைக் கண்டறி.

2. இரண்டிலக்க ராண்டம் எண்கள் பட்டியலிலிருந்து (x, y) எனும் இரண்டு எண்கள் எடுக்கப்படுகின்றன. பின்வருவன வற்றைக் கண்டுபிடி.

(அ) $E(x)$, $E(y)$, $V(x)$, $V(y)$

(ஆ) $E(x-y)^2$

விடை

(அ) x-ன் பரவல் பின்வருமாறு.

x	00	01	02	…………………………	99
P	$\frac{1}{100}$	$\frac{1}{100}$	$\frac{1}{100}$	…………………………	$\frac{1}{100}$

$$\therefore \quad E(x) = \sum_{r=0}^{99} \frac{1}{100} r = \frac{1}{100} \sum_{r=0}^{99} r = \frac{1}{100} \frac{99 \times 100}{2}$$

$$= 49 \cdot 5$$

$$E(x^2) = \sum_{r=0}^{99} \frac{1}{100} r^2 = \frac{1}{100} \sum_{r=0}^{99} r^2 = \frac{1}{100} \frac{99 \times 100 \times 199}{6}$$

$$= 3283 \cdot 5$$

$$\therefore \quad V(x) = E(x^2) - [E(x)]^2 = 3283 \cdot 5 - (49 \cdot 5)^2$$

$$= 833 \cdot 25$$

x-ம் y-ம் ஒரே பரவலையுடையவையாதலால்

$E(y) = E(x) = 49 \cdot 5 \quad V(y) = V(x) = 833 \cdot 25$

(ஆ) $E(x-y)^2 = E[(x - 49 \cdot 5) - (y - 49 \cdot 5)]^2$

$= V(x-y) = V(x) + V(y)$ [x, y சார்பற்றவை]

$= 2\,V(x) = 1666 \cdot 50$

3. இரண்டு உறுப்புகளிலிருந்து ஓருறுப்பைச் சம வாய்ப்புடன் எடுக்கவேண்டும். இதற்காக ஒரு நாணயம் கொடுக்கப்பட்டிருக்கிறது. ஆனால் துரதிருஷ்டவசமாக நாணயம் பிறழ்ச்சியுடையது (Biased). இந்த நாணயத்தை உபயோகித்து எப்படி ஓருறுப்பைச் சமவாய்ப்புடன் தேர்ந்தெடுப்பாய்? இதே முறையை 8 உறுப்புகளிலிருந்து ஓருறுப்பைச் சம வாய்ப்புடன் எடுக்க எப்படி உபயோகிப்பாய்?

விடை

(அ) நாணயத்தை இரு முறை சுண்டிவிடு. (பூ, தலை) என்று வந்தால் முதல் உறுப்பைத் தேர்ந்தெடு. (தலை, பூ) என்று வந்தால் இரண்டாம் உறுப்பைத் தேர்ந்தெடு. (பூ, பூ) அல்லது (தலை, தலை) என்று வந்தால் மறுபடியும் நாணயத்தை இருமுறை சுண்டிவிடு. இதே முறையை (பூ, தலை) அல்லது (தலை, பூ) வரும் வரை உபயோகி.

P (பூ, தலை) $=$ P (தலை, பூ) என்பதால் இரு உறுப்புகளுக்கும் சமவாய்ப்புக் கிடைக்கிறது.

(ஆ) 8 உறுப்புகளுக்கும் பின்கண்ட எண்களை அளி.

000, 001, 010, 011, 100, 101, 110, 111.

(பூ, தலை) என்பதை 1 எனவும், (தலை, பூ) என்பதை 0 எனவும் கொள். ஒரு நாணயத்தை இருமுறை சுண்டிவிடு. (பூ, தலை) வந்தால் 1 எனவும், (தலை, பூ) வந்தால் 0 எனவும் கொள். (அ)-ல் குறிப்பிட்டபடி 1 அல்லது 0 வரும் வரை நாணயத்தை இருமுறை சுண்டிவிட்டுக் கொண்டேயிரு. இவ்விதத்தில் 1 அல்லது 0 என்ற எண் கிடைக்கும். இன்னமும் இருமுறை இம்முறையை உபயோகித்து, மேலும் இரண்டு எண்களைக் கண்டுபிடி. இவ்விதம் மூன்று எண்கள் வரிசையாகக் கிடைக்கும். (000) வந்தால் முதல் உறுப்பு, (001) வந்தால் இரண்டால் உறுப்பு,......... (111) வந்தால் 8ஆவது உறுப்பு என்று தேர்ந்தெடு.

4. N உறுப்புகள் கொண்ட முழுமைத் தொகுதியில் $\alpha\%$ உறுப்புகளின் மதிப்பு 0 ஆகும். மற்ற உறுப்புகளின் மதிப்பு 0 அல்லாதது. முழுமைத் தொகுதியிலுள்ள எண்களின் மொத்தம் மதிப்பிடப்படவேண்டியிருக்கிறது. இதற்குப் பின்வரும் இரண்டு முறைகள் கையாளப்பட்டிருக்கின்றன.

(அ) முழுமைத் தொகுதியிலிருந்து n உறுப்புகள் ஈ. செ. சா. ரா. மா. முறையில் எடுக்கப்பட்டன. $\bar{x}_1$ என்பது இம் மாதிரியின் சராசரி. $N\bar{x}_1$ என்ற மதிப்பீட்டி முழுமைத் தொகுதியின் மொத்தத்தை மதிப்பிடுவதற்கு உபயோகப்படுத்தப்பட்டது.

(ஆ) முழுமைத் தொகுதியிலுள்ள குறிப்பிட்ட $\alpha\%$ உறுப்புகளை நீக்கிவிட்டு மிச்சமுள்ள $\left(N - \frac{\alpha N}{100}\right)$ உறுப்புகளிலிருந்து n உறுப்புகள் கொண்ட ஈ, செ. சா. ரா. மா. ஒன்று எடுக்கப்பட்டது. இதன் மாதிரிச் சராசரி $\bar{x}_2$. $\left(N - \frac{\alpha N}{100}\right)\bar{x}_2$ என்பது முழுமைத்தொகுதியின் மொத்தத்தை மதிப்பிடுவதற்கு உபயோகப் படுத்தப்பட்டது.

மேற்கூறிய இரு முறைகளில் எது அதிகத் திறன் வாய்ந்தது? அதன் திறன் விகிதம் (சார்புப் பயனளவு) என்ன?

விடை

$\frac{N\alpha}{100} = N_1$ எனவும், $N - N_1 = N_2$ எனவும் கொள்க.

N_1 உறுப்புகளின் சராசரி $\bar{X}_1 = 0$, இவற்றின் பரவற் படி $\sigma_1^2 = 0$

N_2 உறுப்புகளின் சராசரி $\bar{X}_2$ இவற்றின் சராசரி σ_2^2

N உறுப்புகளின் கூட்டுச் சராசரி (Combined Mean)

$$= \frac{N_1\bar{X}_1 + N_2\bar{X}_2}{N_1 + N_2} = \frac{N_2\bar{X}_2}{N} = \bar{X} \text{ எனக் கொள்க.}$$

N உறுப்புகளின் பரவற் படி (Combined Vairance)

$$= \frac{N_1\sigma_1^2 + N_2\sigma_2^2}{N_1 + N_2} + \frac{N_1(\bar{X} - \bar{X}_1)^2 + N_2(\bar{X} - \bar{X}_2)^2}{N_1 + N_2}$$

$$= \frac{N_2}{N}\sigma_2^2 + \frac{N_1\bar{X}^2 + N_2\left(\bar{X} - \frac{N}{N_2}\bar{X}\right)^2}{N}$$

$$= \left[N_2\sigma_2^2 + \left(\frac{N_1^2}{N_2} + N_1\right)\bar{X}^2\right] \div N = \sigma^2 \text{ எனக்}$$

கொள்க.

$$V(N\bar{x}_1) = N^2 \frac{\sigma^2}{n}$$

$$V(N_2 \bar{x}_2) = N_2^2 \frac{\sigma_2^2}{n}$$

ஆனல் $\dfrac{N_2^2 \sigma_2^2}{n} < \dfrac{NN_2 \sigma_2^2}{n} < \dfrac{NN_2 \sigma_2^2}{n}$

$$+ \frac{N_1 N^2}{N_2 n} \bar{X}^2 = N^2 \frac{\sigma^2}{n}$$

அதாவது $V(N_2 \bar{x}_2) < V(N \bar{x}_1)$

∴ இரண்டாம் மதிப்பீட்டியே திறன் வாய்ந்தது.
இரண்டாம் மதிப்பீட்டியின் சார்புப் பயனளவு(Relative Efficiency)

$$= \frac{V(N\bar{x}_1)}{V(N_2\bar{x}_2)} = \left[\frac{NN_2}{n}\sigma_2^2 + \frac{N_1 N^2}{N_2 n}\bar{X}^2\right] \div \frac{N_2^2 \sigma_2^2}{n}$$

$$= 1 + \frac{N_1}{N_2} + \frac{N_1 N^2}{N_2^3} \cdot \frac{\bar{X}^2}{\sigma^2} > 1.$$

5. இயல்நிலைப் பரவலிலிருந்து $2n$ உறுப்புகள் சா. ரா. மா. முறையில் எடுக்கப்பட்டன. இவற்றுள் முதல் n உறுப்புகளின் சராசரி $\bar{y}_1$, கடைசி n உறுப்புகளின் சராசரி $\bar{y}_2$, $2n$ உறுப்புகளின் சராசரி $\bar{y}$. s_1^2 என்பது முதல் n உறுப்புகளின் மாதிரிப் பரவற் படி. s_2^2 என்பது கடைசி n உறுப்புகளின் மாதிரிப் பரவற் படி. s^2 என்பது $2n$ உறுப்புகளின் மாதிரி பரவற் படி.

பின்வரும் ராண்டம் மாறிகளின் பரவற் படியைக் கண்டுபிடி.

(அ) $v_1 = \dfrac{s^2}{2n}$ (ஆ) $v_2 = \dfrac{(\bar{y}_1 - \bar{y}_2)^2}{4}$

(இ) $v_3 = \dfrac{s_1^2 + s_2^2}{4n}$

விடை

இயல்நிலைப்பரவலின் பரவற்படி σ^2 எனக் கொண்டால் பின்வரும் முடிவுகள் தெளிவானவை.

(அ) $\dfrac{(n-1)\, s_1^2}{\sigma^2} \frown X^2_{n-1}, \dfrac{(n-1)\, s_2^2}{\sigma^2} \frown X^2_{n-1}$

(ஆ) $\frac{(2n-1)\, s^2}{\sigma^2} \frown X^2_{2n-1}$

(இ) $V[\, X_m^2 \,] = 2m$

(அ) $V(v_1) = V\left(\frac{s^2}{2n}\right) = \frac{1}{4n^2}\, V(s^2)$

$$= \frac{1}{4n^2}\, V\left[\frac{\sigma^2}{2n-1} \cdot \frac{2n-1}{\sigma^2}\, s^2\right]$$

$$= \frac{1}{4n^2}\, \frac{\sigma^4}{(2n-1)^2}\, V\left(\frac{2n-1}{\sigma^2}\, s^2\right)$$

$$= \frac{\sigma^4}{4n^2\,(2n-1)^2}\, V(X^2_{2n-n})$$

$$= \frac{\sigma^4}{4n^2\,(2n-1)^2}\, 2\,(2n-1) = \frac{\sigma^4}{2n^2\,(2n-1)}$$

(ஆ) $V(v_2) = V\left[\frac{(\bar{y}_1 - \bar{y}_2)^2}{4}\right]$

$$= \frac{1}{16}\, V(\bar{y}_1 - \bar{y}_2)^2$$

ஆனல் $\bar{y}_1 - \bar{y}_2 \frown N\left(0, \frac{2\sigma^2}{n}\right)$

$\therefore \quad \frac{\bar{y}_1 - \bar{y}_2}{\sqrt{\frac{2\sigma^2}{n}}} \frown N(0, 1)$

$\therefore \quad \frac{(\bar{y}_1 - \bar{y}_2)^2}{\frac{2\sigma^2}{n}} \frown X_2^2$

$\therefore \quad V\left[\frac{(\bar{v}_1 - \bar{p}_2)^2}{\frac{2\upsilon^2}{n}}\right] = V(X_1^2) = 2$

$$\therefore \quad V[(\bar{y}_1 - \bar{y}_2)^2] = 2 \cdot \frac{4\sigma^4}{n^2}$$

$$\therefore \quad V(v_2) = \frac{1}{16} \cdot 2 \cdot \frac{4\sigma^4}{n^2} = \frac{\sigma^4}{2n^2}$$

(இ) $v_8 = \frac{s_1^2 + s_2^2}{4n}$

ஆனல் $(n-1)\frac{s_1^2}{\sigma^2} + (n-1)\frac{s_2^2}{\sigma^2} \cap \chi^2_{2n-2}$

$$\therefore \quad 4n\,\frac{(n-1)}{\sigma^2} v_8 \cap \chi^2_{2n-2}$$

$$\therefore \quad V\left[\frac{4n(n-1)}{\sigma^2} v_8\right] = V(\chi^2_{2n-2}) = 2(2n-2)$$

$$\therefore \quad V(v_8) = \frac{\sigma^4}{16n^2(n-1)^2} 2(2n-2) = \frac{\sigma^4}{4n^2(n-1)}$$

6. n உறுப்புகள் கொண்ட ஈ. செ. சா. ரா. மா. ஒன்று ஒரு பரவலிலிருந்து எடுக்கப்பட்டது. இப்பரவலின் மாறுபாட்டுக் கெழு (coefficient of variation) $= C = \frac{\sigma}{\bar{Y}}$ தெரிந்திருக்கிறது. C ஐ உபயோகித்து மாதிரிச் சராசரியை விடக் குறைவான சராசரி வர்க்கப் பிழையுடைய மதிப்பீட்டியொன்றைக் கண்டுபிடி. இப் புதிய மதிப்பீட்டியின் சராசரி வர்க்கப் பிழையை $V(\bar{y})$ உடன் ஒப்பிட்டுத் திறன் விகிதம் (Relative efficiency) கண்டுபிடி. (Goodman, L.A., Ann. Math. Stat. 1953)

விடை

$$V(\bar{y}) = \frac{\sigma^2}{n}$$

புதிய மதிப்பீட்டியை $k\,\bar{y}$ எனக் கொள்வோம்.

$k\bar{y}$-ன் சராசரி வர்க்கப் பிழை $MSE(k\bar{y}) = E(k\bar{y} - \bar{Y})^2$

$$= E[(k(\bar{y} - \bar{Y}) + (k-1)\bar{X}]^2$$

$$= k^2\,\frac{\sigma^2}{n} + (k-1)^2\,\bar{Y}^2 \quad [E(\bar{y} - \bar{Y}) = 0]$$

$MSE\,(k\bar{y})$ சிறுமமாக (minimum) இருக்கவேண்டுமாயின்

$$\frac{d}{dk}\left[\,MSE\,(k\bar{y})\,\right]=0$$

$$\therefore\quad \frac{d}{dk}\left[k^2\frac{\sigma^2}{n}+(k-1)^2\,\bar{Y}^2\right]=2k\frac{\sigma^2}{n}+2(k-1)\,\bar{Y}^2=0.$$

$$\therefore\quad k=\frac{\bar{Y}^2}{\bar{Y}^2+\frac{\sigma^2}{n}}=\frac{1}{1+\frac{c^2}{n}}$$

$$\therefore\quad \text{புதிய மதிப்பீட்டி}\ \frac{1}{\left(1+\frac{c^2}{n}\right)}\,\bar{y}$$

$$\text{இதன் திறன் விகிதம்}=\frac{MSE\,(\bar{y})}{MSE\,(k\bar{y})}=\frac{V\,(\bar{y})}{MSE\,(k\bar{y})}$$

$$=\frac{\frac{\sigma^2}{n}}{k^2\frac{\sigma^2}{n}+(k-1)^2\,\bar{Y}^2}$$

$$=\frac{1}{k^2+(k-1)^2\,\frac{n}{c^2}}=\frac{1}{\frac{n^2}{(n+c^2)^2}+\left(\frac{c^2}{n+c^2}\right)^2\frac{n}{c^2}}$$

$$=\frac{(n^2+c^2)^2}{n^2+n^2c^2}$$

$$=1+\frac{c^2}{n}>1.$$

7. ஈ. செ. சா. ரா. மா முறையில் 3 உறுப்புகள் எடுக்கப்படுகின்றன. இம்மாதிரியில் பின்வரும் நிகழ்ச்சிகளின் நிகழ்திறனைக் கண்டுபிடி.

(அ) மூன்று உறுப்புகளும் வெவ்வேறானவை.

(ஆ) இரண்டு மட்டும் வெவ்வேறானவை.

(இ) மூன்றும் ஒரே உறுப்பு.

இம்மாதிரியில் வெவ்வேருன உறுப்புகளைக் கொண்டு கண்டு பிடிக்கப்படும் சராசரி, முழுமைத் தொகுதியின் சராசரியைப் பிறழ்ச்சியில்லாமல் மதிப்பிடுகிறது என்பதை நிரூபி. இம் மதிப் பீட்டின் பரவற் படி, அப்பரவற் படியின் பிறழ்ச்சியற்றதொரு மதிப்பீட்டி முதலியவைகளைக் கண்டுபிடி.

8. ஒரு முழுமைத் தொகுதியில் அரிதான பண்பு வாய்ந்தவை $100\,P$ சதவீதமாகும். P ஐ மதிப்பிடுவதற்கு உறுப்புகள் ஒவ்வொன்ருக ஈ. செ. சா. ரா. மா. முறையில் எடுக்கப்படுகின்றன. 'm' முறை அரிதான பண்பு வாய்ந்த உறுப்பு வந்ததும் எடுப்பது நிறுத்தப்படுகிறது. இதற்காக 'n' முறை உறுப்புகளை எடுக்க நேர்ந்தது. P-க்கான பிறழ்ச்சியற்ற மதிப்பீட்டியொன்றினைக் கண்டுபிடி. அதன் தோராயமான பரவற் படியென்ன? அப் பரவற் படியின் தோராயமான மதிப்பீட்டியென்ன? (J.B.S. Haldane, Biometrika 1945)

விடை

n உறுப்புகளில் கடைசியானது அரிதான பண்பு வாய்ந்த தாகையால் அதன் நிகழ்திறம் P முதலிலுள்ள $(n-1)$ உறுப்பு களில் $(m-1)$ அரிதான பண்பு வாய்ந்தவையாதலால் அவைகளின் நிகழ்திறன் $\binom{n-1}{m-1} P^{m-1}(1-P)^{n-m}$

$$\therefore \quad \text{மாதிரியின் நிகழ்திறன் } P_{n,m} = \binom{n-1}{m-1} P^m (1-P)^{n-m}$$

$$E\left(\frac{1}{n-1}\right) = \sum_{n=m}^{\infty} \frac{1}{n-1} \cdot P_{n,m}$$

$$= \sum_{n=m}^{\infty} \frac{1}{n-1} \cdot \frac{n-2}{m-2} \binom{n-1}{m-1} P^m (1-P)^{n-m}$$

$$= \frac{1}{(m-1)} P \cdot \sum_{n=m}^{\infty} \binom{n-1}{m-1} P^{m-1} \cdot (1-P)^{(n-1)-(m-1)}$$

$$= \frac{P}{m-1} \cdot \sum_{n=m}^{\infty} P_{n-1, m-1} = \frac{P}{m-1} \sum_{n-1=m-1}^{\infty} P_{n-1, m-1}$$

$$= \frac{P}{m-1} \left[\because \sum_{n=m}^{\infty} P_{n,m} = 1 \right.$$

$$\left. \Longrightarrow \sum_{n-1=m-1}^{\infty} P_{n-1, m-1} = 1 \right]$$

$$\therefore \quad E\left(\frac{m-1}{n-1}\right) = P$$

$$\therefore \quad E^{-1}(P) \sim \frac{m-1}{n-1}$$

அதாவது P-ன் பிறழ்ச்சியற்ற மதிப்பீடொன்று $\frac{m-1}{n-1}$ ஆகும்.

$$V\left(\frac{m-1}{n-1}\right) = E\left(\frac{m-1}{n-1}\right)^2 - P^2$$

$$= (m-1)^2 E\left[\frac{1}{(n-1)(n-2)}\right] - P^2$$

[n பெரியதாக இருந்தால் $(n-1)^2 = (n-1)(n-2)$ தோராயமாக].

$$E \frac{1}{(n-1)(n-2)} = \sum_{n=m}^{\infty} \frac{1}{(n-1)(n-2)} \binom{n-1}{m-1} \cdot P^m (1-P)^{n-m}$$

$$= \frac{P^2}{(m-1)(m-2)} \sum_{n=m}^{\infty} \binom{n-3}{m-3} P^{n-2} (1-P)^{(n-2)-(m-2)}$$

$$= \frac{P^2}{(m-1)(m-2)}$$

$$\therefore V\left(\frac{m-1}{n-1}\right) = (m-1)^2 \frac{P^2}{(m-1)(m-2)} - P^2 = \frac{m-1}{m-2}P^2 - P^2$$

$$= P^2\left[\frac{m-1}{m-2} - 1\right] = \frac{P^2}{m-2} \quad [m > 2]$$

9. $B(n, p)$ என்னும் ஈருறுப்புப் பரவலிலிருந்து r_1, r_2 என்னும் இரண்டு உறுப்புகள் ஈ. செ. சா. ரா. மா. முறையில் எடுக்கப் பட்டன. n என்பது தெரிந்த எண், ஆனால் p தெரியாதது. பின் வரும் மாதிரியளவைகள் p ஐப் பிறழ்ச்சியின்றி மதிப்பிடுகின்றன என்பதை நிரூபி. இவற்றுள் எது சிறந்தது ?

(அ) r_1 (ஆ) r_2 (இ) $\frac{2r_1+5r_2}{7}$ (ஈ) $\frac{r_1+r_2}{2}$

10. பாய்ஸான் பரவல் ஒன்றிலிருந்து (x_1, x_2, x_3) என்னும் 3 உறுப்பு கொண்ட ஈ. செ. சா. ரா. மா. எடுக்கப்பட்டது. பின் வரும் மாதிரியளவைகள் பாய்ஸான் பரவலின் சராசரியைப் பிறழ்ச்சி யின்றி மதிப்பிடுகின்றன என்பதை நிரூபி. இவற்றுள் எது சிறந்தது ?

(அ) x_1 (ஆ) $\frac{x_1+x_2}{2}$ (இ) $\frac{(x_2-x_3)^2}{2}$

18. ஈடு செய்யப்படாத சாதாரண ராண்டம் மாதிரிகளில் மதிப்பீடுகள்

(ESTIMATION IN SIMPLE RANDOM SAMPLING WITHOUT REPLACEMENT)

18·1 ஈ. செயா. சா. ரா. மா. முறையில் $(x_1, x_2, ..., x_n)$ என்பது n உறுப்புகள் கொண்ட மாதிரியானுல் எல்லா உறுப்புகளுக்கும் ஒரே பரவல்தான் இருக்கும். அதாவது $(x_1, x_2, ... x_n)$ என்னும் n ராண்டம் மாதிரிகள் முழுமைத் தொகுதியைப் பரவலாகக் கொண்ட n பிரதிகளாகும். ஆனுல் இவை சார்புள்ளவை. அதாவது $P(x_1 = a, x_2 = b) \neq P(x_1 = a)\, P(x_2 = b) \;\forall\; a, b$. ஆகவே இது ஈ. செ. சா. ரா. மா. ஐ விடச் சற்றுக் கடினமானதாகும்.

$\bar{x}, s^2, n$ என்பவை முறையே மாதிரிச் சராசரி, மாதிரிப் பரவற்படி, மாதிரிப் பருமன் (Sample size) எனக் கொள்வோம்.

$\bar{X}, \sigma^2, N$ என்பவை முறையே முழுமைத் தொகுதியின் சராசரி, பரவற் படி, பருமன் எனக் கொள்வோம்.

18·2. தேற்றம்

மாதிரிச் சராசரி ஈ. செயா. சா. ரா. மா. முறையில் முழுமைத் தொகுதிச் சராசரியைப் பிறழ்ச்சியின்றி மதிப்பிடுகிறது.

அதாவது $\overset{-1}{E}(\bar{X}) \sim \bar{x}$

நிறுவல்

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

x_i-ன் பரவல் பின்வருமாறு:

x_i	X_1	X_2	X_3	....	X_N
P	$\frac{1}{N}$	$\frac{1}{N}$	$\frac{1}{N}$	....	$\frac{1}{N}$

$$\therefore \quad E(x_i) = \sum_{i=1}^{N} \frac{1}{N} X_i = \frac{1}{N} \sum_{i=1}^{N} X_i = \overline{X}$$

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

$$\therefore \quad E^{-1}(\overline{X}) \sim \overline{X}$$

18·3. தேற்றம்

ஈ. செயா. சா. ரா. மா. முறையில் மாதிரிச் சராசரியின் பரவற்படி $\frac{N-n}{N-1}\,\frac{\sigma^2}{n}$ ஆகும்.

நிறுவல்

x_i^2-ன் பரவல்

x_i^2	X_1^2	X_2^2	X_3^2	....	X_N^2
P	$\frac{1}{N}$	$\frac{1}{N}$	$\frac{1}{N}$	....	$\frac{1}{N}$

$$\therefore \quad E(x_i^2) = \frac{1}{N} \sum_{j=1}^{N} X_j^2$$

$x_i\, x_j$-ன் பரவல்

$x_i\, x_j$	$X_1\, X_2$	$X_2\, X_1$	....	$X_{N-1}\, X_N$	$X_N\, X_{N-1}$
P	$\frac{1}{N\,(N-1)}$	$\frac{1}{N\,(N-1)}$	....	$\frac{1}{N\,(N-1)}$	$\frac{1}{N\,(N-1)}$

$$\therefore \quad E\,(x_i\, x_j) = \frac{1}{N\,(N-1)} \sum_{i \neq j} \sum X_i\, X_j$$

$$= \frac{2}{N\,(N-1)} \sum_{i<j} \sum X_i\, X_j$$

$$V\,(\bar{x}) = E\,(\bar{x}^2) - [\,E\,(\bar{x})\,]^2 = E\,(\bar{x}^2) - \bar{X}^2 \qquad \ldots\ldots I$$

$$E\,(\bar{x}^2) = E\left[\frac{x_1 + x_2 + \cdots + x_n}{n}\right]^2 = \frac{1}{n^2} \sum_{i=1}^{n} E\,(x_i^2)$$

$$+ \frac{2}{n^2} \sum_{i<j} \sum E\,(x_i\, x_j)$$

$$= \frac{1}{n^2} \cdot n \cdot \frac{1}{N} \sum_{j=1}^{N} X_j^2$$

$$+ \frac{2}{n^2}\, \frac{n\,(n-1)}{2} \cdot \frac{2}{N\,(N-1)} \sum_{i<j} \sum X_i\, X_j$$

$$= \frac{1}{nN} \sum_{j=1}^{N} X_j^2$$

$$+ \frac{n-1}{nN\,(N-1)} \left[\left(\sum_{j=1}^{N} X_j \right)^2 - \sum_{j=1}^{N} X_j^2 \right]$$

$$= \frac{1}{nN}\sum_{j=1}^{N} X_j^2 + \frac{n-1}{nN(N-1)}\left[N^2\bar{X}^2 - \sum_{j=1}^{N} X_j^2 \right]$$

$$\therefore \quad I\text{-லிருந்து } V(\bar{x}) = \frac{1}{nN}\sum_{j=1}^{N} X_j^2$$

$$+ \frac{n-1}{nN(N-1)} N^2 \bar{X}^2 - \frac{n-1}{nN(N-1)}\sum_{j=1}^{N} X_j^2 - \bar{X}^2$$

$$\therefore \quad V(\bar{x}) = \frac{1}{nN}\sum_{j=1}^{N} X_j^2 \left[1 - \frac{n-1}{N-1}\right]$$

$$- \bar{X}^2 \left[1 - \frac{N(n-1)}{n(N-1)}\right]$$

$$= \frac{N-n}{nN(N-1)}\sum_{j=1}^{N} X_j^2 - \frac{N-n}{n(N-1)}\bar{X}^2$$

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

$$= \frac{N-n}{n(N-1)}\sigma^2$$ (நி—து.)

18·4. தேற்றம்.

ஈ. செயா. சா. ரா. மா. முறையில் $E(s^2) = \frac{N\sigma^2}{N-1}$

நிறுவல்

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

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

$$= \frac{1}{n-1}\cdot n\cdot\frac{1}{N}\sum_{j=1}^{N} X_j^2 - \frac{n}{n-1}[V(\bar{x}) + \bar{X}^2]$$

$$[\because\ V(\bar{x}) = E(\bar{x}^2) - \bar{X}^2\ ]$$

$$= \frac{n}{(n-1)N}\sum_{j=1}^{n} X_j^2 - \frac{n}{n-1}\bar{X}^2$$

$$- \frac{n}{n-1}\,\frac{N-n}{n(N-1)}\,\sigma^2$$

$$= \frac{n}{n-1}\sigma^2 - \frac{N-n}{(n-1)(N-1)}\sigma^2$$

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

$$= \frac{N}{N-1}\sigma^2 \qquad \text{(நி-து.)}$$

18·5. தேற்றம்.

(அ) σ^2-ன் பிறழ்ச்சியற்ற மதிப்பீட்டியொன்று $\frac{N-1}{N}\,s^2$ ஆகும்.

அதாவது $E^{-1}(\sigma^2) \sim \frac{N-1}{N}\,s^2$

(ஆ) $V(\bar{x})$-ன் பிறழ்ச்சியற்ற மதிப்பீட்டியொன்று $\frac{N-n}{N} \frac{s^2}{n}$

(இ) ஈ. செயா. சா. ரா. மா-ன் சராசரி ஈ. செ. சா. ரா. மா. -ன் சராசரியை விட முழுமைத் தொகுதிச் சராசரியை மதிப்பிடுவதில் திறன் அதிகம் வாய்ந்தது.

நிறுவல்

(அ) $E(s^2) = \frac{N}{N-1} \sigma^2$ என்று 18·4-ல் நிறுவப்பட்டது.

$$\therefore \quad E\left(\frac{N-1}{N} s^2\right) = \sigma^2$$

$$\therefore \quad \overset{-1}{E}(\sigma^2) \sim \frac{N-1}{N} s^2 \qquad \text{(நி-து.)}$$

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

$$\therefore \quad \overset{-1}{E}[V(\bar{x})] = \frac{N-n}{(N-1)\,n} \overset{-1}{E}(\sigma^2)$$

$$\sim \frac{N-n}{(N-1)\,n} \frac{N-n}{N} s^2 \qquad [\because \text{ (அ)}]$$

$$= \frac{N-1}{nN} s^2 \qquad \text{(நி-து.)}$$

(இ) ஈ. செ. சா. ரா. மா. முறையில் $V(\bar{x}) = V_1 = \frac{\sigma^2}{n}$

ஈ. செயா. சா. ரா. மா. முறையில் $V(\bar{x}) = V_2 = \frac{N-n}{N-1} \frac{\sigma^2}{n}$

$$V_2 = \frac{N-n}{N-1} \frac{\sigma^2}{n} = \frac{(N-1)-(n-1)}{N-1} \frac{\sigma^2}{n}$$

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

$n \leq N$ என்பதால் $0 \leq \frac{n-1}{N-1} \leq 1.$

$$\therefore \quad V_2 \leq \frac{\sigma^2}{n} = V_1 \quad ie\ V_2 \leq V_1 \qquad \text{(நி-து.)}$$

$$\text{திறன் விகிதம் (Relative efficiancy)} = \frac{V_1}{V_2} = \frac{\frac{\sigma^2}{n}}{\frac{N-n}{N-1}\frac{\sigma^2}{n}}$$

$$= \frac{N-1}{N-n} \geq 1$$

18·6. ஈ. செ. சா. ரா. மா & ஈ. செயா. சா. ரா. மா. ஒப்பீடு

மாதிரிப் பருமன் n ஆனது முழுமைத் தொகுதிப் பருமன் N உடன் ஒப்பிடும் போது மிகவும் சிறியதாக இருந்தால் மாதிரியில் ஓர் உறுப்பு ஒன்றுக்கு மேற்பட்ட முறை வருவதற்கு நிகழ் திறன் மிகவும் குறைவாகவே இருக்கும். ஆகவே மேற் கூறப்பட்ட இரு முறைகளுக்கும் அதிக வேறுபாடு இராது.

$$V_2 = \frac{N-n}{N-1}\frac{\sigma^2}{n} \doteq \frac{\sigma^2}{n} = V_1$$

அதாவது இரண்டு மாதிரி முறைகளும் N பெரியதாகவும், n சிறியதாகவும் இருக்கும்போது கிட்டத்தட்ட ஒரே பயனளவு (Efficiancy) கொண்டவை.

ஆனல் N சிறியதாக இருக்கும்போது ஈ. செயா. சா. ரா. மா. முறை மற்ற முறையை விடத் திறன் அதிகம் வாய்ந்ததாகவே இருக்கும்.

18·7. முழுமைத் தொகுதிப் பண்புப் பாகம் P ஐ ஈ. செயா. சா. ரா. மா. முறையில் மதிப்பிடுதல்

ஈ. செ. சா. ரா. மா. முறையில் வரையறுத்தது போலவே

$$\begin{cases} X_i = 1 \text{ [i ஆவது முழுமைத் தொகுதி உறுப்புக்குக் குறிப்பிட்ட பண்பு இருந்தால்]} \\ \quad = 0 \text{ [,, ,, ,, ,, ,, ,, இல்லையேல்]} \end{cases}$$

$$\begin{cases} x_i = 1 \text{ [}i\text{ ஆவது மாதிரி உறுப்புக்குக் குறிப்பிட்ட பண்பு இருந்தால்]} \\ \quad = 0 \text{ [,, ,, ,, ,, ,, ,, இல்லையேல்]} \end{cases}$$

என்று இங்கும் வரையறுப்போம்.

முன்பே கண்டபடி $\bar{X} = P$ $\sigma^2 = PQ$

$$\bar{x} = P \qquad s^2 = \frac{n}{n-1} pq$$

18·8. தேற்றம்

ஈ. செ. சா. ரா. மா. முறையில்,

(அ) முழுமைத் தொகுதிப் பண்பு பாகம் P-ன் பிறழ்ச்சியற்ற மதிப்பீடொன்று மாதிரிப் பண்பு பாகம் P ஆகும்.

அதாவது $\overset{-1}{E} (P) \sim P$

(ஆ) $V(P) = \dfrac{N-n}{N-1} \dfrac{PQ}{n}$

(இ) $V(p)$-ன் பிறழ்ச்சியற்ற மதிப்பீடொன்று $\dfrac{N-n}{N(n-1)} pq$ ஆகும்.

அதாவது $\overset{-1}{E} [V(p)] \sim \dfrac{N-n}{N(n-1)} pq$

நிறுவல்

(அ) $E(\bar{x}) = \bar{X}$

$\bar{x} = p$ $\bar{X} = P$ என்பதிலிருந்து $E(p) = P$ என்பது தெரிகிறது.

$\therefore \quad E^{-1}(P) \sim p$ [நி-து.]

(ஆ) $V(\bar{x}) = \dfrac{N-n}{n(N-1)} \sigma^2$

$\bar{x} = p \quad \sigma^2 = PQ$ என்பதிலிருந்து

$V(p) = \dfrac{N-n}{n(N-1)} PQ$ என்பது தெரிகிறது. [நி-து.]

(இ) $\overset{-1}{E} [V(\bar{x})] \sim \dfrac{N-n}{N} \dfrac{s^2}{n}$

$\bar{x} = p, s^2 = \dfrac{n}{n-1} pq$ என்பதிலிருந்து.

$\overset{-1}{E} [V(p)] \sim \dfrac{N-n}{Nn} \dfrac{n}{n-1} pq = \dfrac{N-n}{N(n-1)} pq$

என்பது விளங்குகிறது. [நி-து.]

18·9. தேற்றம்

(அ) ஈ. செயா. சா. ரா. மா. முறையில் முழுமைத் தொகுதி மொத்தமான X-ன் பிறழ்ச்சியற்ற மதிப்பீடொன்று $N\bar{x}$ ஆகும்.

அதாவது $E^{-1}(X) \sim N\bar{x}$.

(ஆ) $V(N\bar{x}) = N^2 \dfrac{N-n}{N-1} \cdot \dfrac{\sigma^2}{n}$

(இ) $\overset{-1}{E} [V(N\bar{x}] \sim \dfrac{N(N-n)}{n} s^2.$

நிறுவல்

(அ) $E(N\bar{x}) = NE(\bar{x}) = N\bar{X} = X$

$\therefore \quad \overset{-1}{E}(X) \sim N\bar{x}$ (நி-து.)

(ஆ) $V(N\bar{x}) = N^2 V(\bar{x}) = N^2 \dfrac{N-n}{N-1} \dfrac{\sigma^2}{n}$ (நி-து.)

(இ) $\overset{-1}{E} [V(N\bar{x})] = \overset{-1}{E} \left[N^2 \dfrac{N-n}{N-1} \dfrac{\sigma^2}{n} \right]$

$$= N^2 \frac{N-n}{(N-1)} \frac{1}{n} \overset{-1}{E} (\sigma^2)$$

$$\sim \frac{N^2}{n} \frac{N-n}{N-1} \frac{N-1}{N} s^2 = \frac{N(N-n)}{n} s^2$$

அதாவது $E^{-1}[V(N\bar{x})] \sim \frac{N(N-n)}{n} s^2$ (நி-து.)

பயிற்சி 3

1. ஒரு முழுமைத் தொகுதியிலிருந்து சார்பில்லாமல் தனித் தனியாக $n_1, n_2, \ldots, n_k$ உறுப்புகள், (அ) ஈ.செ.சா.ரா.மா. முறையில் (ஆ) ஈ. செயா. சா. ரா. மா. முறையில் எடுக்கப்பட்டன. இம் மாதிரிகளின் சராசரி முறையே $\bar{x}_1, \bar{x}_2, \ldots \ldots \bar{x}_k$; பரவற் படிகள் $s_1^2, s_2^2, \ldots, s_k^2$.

முழுமைத் தொகுதியின் சராசரிக்குப் பிறழ்ச்சியற்றதொரு மதிப்பீட்டியைக் கண்டுபிடி. அதன் பரவற் படியென்ன ? இப் பரவற் படியின் பிறழ்ச்சியற்றதொரு மதிப்பீட்டியென்ன ?

விடை

(அ) $\bar{x} = \frac{1}{n} \sum_{i=1}^{K} n_i \bar{x}_i \qquad n = \sum_{i=1}^{K} n_i$ எனக் கொள்க.

$$E(\bar{x}) = \frac{1}{n} \sum_{i=1}^{k} E(n_i \bar{x}_i) = \frac{1}{n} \sum_{i=1}^{k} n_i \bar{X} = \bar{X}$$

$$\therefore \quad E^{-1}(\bar{X}) \sim \bar{x}$$

$$V\bar{x} = V\left[\frac{1}{n} \sum_{i=1}^{K} n_i \bar{x}_i\right] = \frac{1}{n^2} \sum_{i=1}^{K} n_i^2 V(\bar{x}_i)$$

[$\bar{x}_1, \bar{x}_2, \ldots \bar{x}_k$ என்பவை சார்பற்றவை]

$$= \frac{1}{n^2} \sum_{i=1}^{k} n_i^2 \frac{\sigma^2}{n_i} = \frac{\sigma^2}{n}$$

$$E(s_i^2) = \sigma^2 \qquad E\left[\frac{1}{n}\sum_{i=1}^{k} n_i\, s_i^2\right] = \frac{1}{n}\sum_{i=1}^{k} n_i\, \sigma^2 = \sigma^2.$$

$$E^{-1}[V(\bar{x})] = E^{-1}\left(\frac{\sigma^2}{n}\right) \sim \frac{1}{n^2}\sum_{i=1}^{k} n_i\, s_i^2$$

(ஆ) $$\bar{x} = \frac{1}{n}\sum_{i=1}^{k} n_i\, \bar{x}_i$$

$$E(\bar{x}) = \frac{1}{n}\sum_{i=1}^{k} n_i\, \bar{X} = \bar{X}$$

$$E^{-1}(\bar{X}) \sim \bar{x}$$

$$V(\bar{x}) = \frac{1}{n^2}\sum_{i=1}^{k} n_i^2\, \frac{N-n_i}{N-1}\, \frac{\sigma^2}{n_i} = \frac{1}{n^2}\sum_{i=1}^{k} n_i\, \frac{N-n_i}{N-1}\sigma^2.$$

$$E(s_i^2) = \frac{N}{N-1}\sigma^2 \qquad \therefore\ E\left(\frac{s_i^2}{N}\right) = \frac{\sigma^2}{N-1}$$

$$E\left[\frac{1}{n^2}\sum_{i=1}^{k} n_i\, \frac{N-n_i}{N}\, s_i^2\right] = V(\bar{x})$$

$$E^{-1}[V(\bar{x})] \sim \frac{1}{n^2\bar{X}}\sum_{i=1}^{k} n_i (N-n_i)\, s_i^2.$$

2 ஈ செயா சா ரா மா முறையில n_1 n_2, n_3 உறுபபுகளை யுடைய மாதிரிகளின வாயிலாக நெல கருமபு, வோககடலை முதலியவைகளின் மாதிரிச சராசரி விளைசசல் $\bar{x}$, $\bar{y}$, $\bar{z}$, [கிலோ

கிராமில்] இவை N_1, N_2, N_3 உறுப்புகளையுடைய முழுமைத் தொகுதிகளினின்று சார்பில்லா வகையில் எடுக்கப்பட்டன

நெல, கருமபு, வேர்க்கடலை முதலியவைகளின் மொத்த விளைச்சலை (கிலோ கிராமில்) பிறழ்ச்சியின்றி மதிப்பிட எந்த மதிப்பீட்டியைக் கையாள்வாய ? அம்மதிப்பீட்டியின் பரவற் படி யென்ன ?

3 N உறுப்புகள் கொண்ட (x, y) என்னும இரு மாறிகளின பரவலின் சராசரி $(\bar{X}, \bar{Y})$, பரவல் படிகள் (σ_x^2, σ_y^2), உடன் தொடர்புக் கெழு (Correlation coefficient) P

$(\bar{x}, \bar{y})$ என்பது n உறுப்புகள் கொண்ட ஈ செய்யா சா ரா மா முறையில் எடுக்கப்பட்ட மாதிரியின் சராசரி $(\bar{x}, \bar{y})$-க்கிடையே உள்ள உடன் தொடர்புக் கெழு யாது ? $(\bar{x} - \bar{y}, \bar{x} + \bar{y})$-க் கிடையே உள்ள உடன் தொடர்புக் கெழு யாது ?

விடை ρ, $\dfrac{\sigma_x^2 - \sigma_y^2}{\sqrt{(\sigma_x^2 + \sigma_y^2)^2 - 4\rho^2 \sigma_x^2 \sigma_y^2}}$

4. ஒரு பெட்டியில் 1 முதல் N வரை வரிசையாக எண்ணிடப் பட்ட N பொருள்கள் உள்ளன N என்ன என்பது தெரியாததால் அதை மதிப்பிட n உறுப்புகள் கொண்ட மாதிரி ஈ செய்யா சா ரா மா முறையில் எடுக்கப்பட்டன இம மாதிரி எண்கள் $y_1, y_2, \cdots y_n$ N-க்குப பிறழ்ச்சியற்ற மதிப்பீட்டியொன்றினைக் கண்டுபிடி. இம மதிப்பீட்டியின் பரவற் படி யாது? இப்பரவற் படிக்குப பிறழ்ச்சியற்ற மதிப்பீட்டி யாது?

விடை

$$\bar{Y} = \frac{1 + 2 + 3 + \cdots + N}{N} = \frac{N(N+1)}{2N} = \frac{N+1}{2}$$

$$\bar{y} = \frac{1}{n} \sum_{i=1}^{n} y_i$$ என்று எடுத்துக்கொண்டால்

$$E(\bar{y}) = \bar{Y} = \frac{N+1}{2}$$

$$E(2\bar{y} - 1) = N \qquad E^{-1}(N) \sim (2 - 1\bar{y})$$

$$V(2\bar{y}-1)=4\,V(\bar{y})=4\,\frac{N-n}{N-1}\;\frac{\sigma^2}{n}.$$

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

$$=\frac{1}{N}\;\frac{N(N+1)(2N+1)}{6}-\frac{(N+1)^2}{4}$$

$$=\frac{N+1}{12}\,[4N+2-3N-3]=\frac{N^2-1}{12}.$$

$$\therefore\quad V(2\bar{y}-1)=\frac{4}{n}\,\frac{N-n}{N-1}.\;\frac{N^2-1}{12}=\frac{(N-n)(N+1)}{3n}$$

$$E^{-1}[V(2\bar{y}-1)]=4\,E^{-1}[(V(\bar{y}))]\sim 4\,\frac{N-n}{N}\,\frac{s^2}{n}$$

$$s^2=\frac{1}{n-1}\sum_{i=1}^{n}(y_i-\bar{y})^2.$$

5. ஈ. செயா. சா. ரா. மா. முறையில் எடுக்கப்பட்ட மாதிரியில் iஆவது உறுப்பு y_i. y_i-ன் பரவற் படியைக் கண்டுபிடி. y_i, $y_{i'}$ $(i \neq i')$-இவைகளின் உடன் மாற்றத்தைக் (Covariance) கண்டுபிடி. இவைகளின் வழியாக மாதிரிச் சராசரி $\bar{y}$-ன் பரவற் படியைக் கண்டுபிடி.

6. ஒரு முழுமைத் தொகுதியில் 4, 6, 8, 12, 10, 14 என்னும் எண்கள் மட்டுமே உள்ளன. இத்தொகுதியிலிருந்து ஈ. செயா. சா. ரா. மா. முறையில் 3 உறுப்புகள் கொண்ட 20 மாதிரிகளையும் எழுது. இம்மாதிரிகளின் சராசரிகளைக் கண்டுபிடி. இம்மாதிரிச் சராசரிகளின் சராசரி முழுமைத் தொகுதியின் சராசரியான 9 என்பதைக் கண்டறி.

7. ஒரு குளத்திலிருந்து r மீன்கள் பிடிக்கப்படுகின்றன. அம் மீன்கள் அடையாளமிடப்பட்டு மறுபடியும் குளத்திலேயே விடப்படுகின்றன. சில நாட்கள் கழித்து அடையாளமிடப்பட்ட

மீன்கள் 'm' வரும்வரை ஒவ்வொன்ருக ஈ: செயா. சா. ரா. மா. முறையில் பிடிக்கப்படுகின்றன. இவ்விதமாக மொத்தம் n மீன்கள் எடுக்கவேண்டியதாயிற்று. மேற்கூறப்பட்ட சில நாட்களில் மீன்களின் எண்ணிக்கை மாறவில்லையென்று எடுத்துக்கொண்டு குளத்திலுள்ள மொத்த மீன்களின் எண்ணிக்கைக்குப் பிறழ்ச்சியற்ற மதிப்பீடொன்றைக் கண்டுபிடி. இம் மதிப்பீட்டியின் பரவற் படிக்குப் பிறழ்ச்சியற்றதொரு மதிப்பீட்டியைக் கண்டுபிடி.

[Bailey, N. T. J., Biometrika, 38, (1951)].

விடை

முதல் $(n-1)$ முயற்சிகளில் $(m-1)$ அடையாளமிடப்பட்ட மீன்கள் வந்திருக்க வேண்டும்.

இதன் நிகழ் திறம் $$\frac{\binom{r}{m-1}\binom{N-r}{n-m}}{\binom{N}{n-1}}$$

n ஆவது முயற்சியில் அடையாளமிடப்பட்ட மீன் வருவதற்கு

நிகழ் திறம் $\dfrac{r-m+1}{N-n+1}$ [N = குளத்திலுள்ள மொத்த மீன்கள்]

ஆகவே m அடையாளமிட்ட மீன்கள் சரியாக n முயற்சிகளில் வருவதற்கான நிகழ் திறம்

$$P_n = \frac{r-m+1}{N-n+1}\;\frac{\binom{r}{m-1}\binom{N-r}{n-m}}{\binom{N}{n-1}} = \frac{\frac{m}{n}\binom{r}{m}\binom{N-r}{n-m}}{\binom{N}{n}}.$$

$\sum_{n\geq m} P_n = 1$ என்பது தெளிவானது.

$$\therefore \sum_{n\geq m} \frac{\frac{m}{n}\binom{r}{m}\binom{N-r}{n-m}}{\binom{N}{n}} = 1.$$

$$\therefore \sum_{n \geq m} \frac{\binom{N-r}{n-m}}{n\binom{N}{n}} = \frac{1}{m\binom{r}{m}} \qquad \ldots\ldots \text{I}$$

$$E(n-m) = \frac{\sum_{n \geq m} (n-m)\frac{m}{n}\binom{r}{m}\binom{N-r}{n-m}}{\binom{N}{n}}$$

$$= m\binom{r}{m}(N-r)\sum_{n \geq m} \frac{\binom{N-r-1}{n-m-1}}{n\binom{N}{n}}$$

$$= m\binom{r}{m}(N-r)\frac{1}{(m+1)\binom{r+1}{m+1}}$$

[I-ல் r, m என்பவைகளை $r+1$, $m+1$ என்று மாற்றுவதால்]

$$= \frac{m\binom{r}{m}(N-r)}{(m+1)\frac{r+1}{m+1}\binom{r}{m}} = \frac{m(N-r)}{r+1}$$

$$\therefore E\left[\frac{(n-m)(r+1)}{m}\right] = N-r$$

$$\therefore E\left[\frac{(n-m)(r+1)}{m} + r\right] = N$$

$$\therefore E\left[\frac{n(r+1)}{m} - 1\right] = N$$

$\therefore$ N-ன் பிறழ்ச்சியற்ற மதிப்பீடொன்று $\frac{n(r+1)}{m} - 1$ ஆகும்.

இம்மதிப்பீட்டியின் பரவற் படி : $V\left[\frac{n(r+1)}{m}-1\right]$

$$= \left(\frac{r+1}{m}\right)^2 V(n)$$

$$= \left(\frac{r+1}{m}\right)^2 V(n-m)$$

$$= \left(\frac{r+1}{m}\right)^2 \left\{E[(n-m)^2] - \frac{m^2(N-r)^2}{(r+1)^2}\right\}$$

$$= \left(\frac{r+1}{m}\right)^2 \left\{E(n-m)(n-m-1) + E(n-m) - \frac{m^2(N-r)^2}{(r+1)^2}\right\} \quad \ldots II$$

$E(n-m)(n-m-1)$

$$= \sum_{n \geq m} (n-m)(n-m-1)\, \frac{\frac{m}{n}\binom{r}{m}\binom{N-r}{n-m}}{\binom{N}{n}}$$

$$= \sum_{n \geq m} \frac{\frac{m}{n}\binom{r}{m}(N-r)(N-r-1)\binom{N-r-2}{n-m-2}}{\binom{N}{n}}$$

$$= m\binom{r}{m}(N-r)(N-r-1) \sum_{n \geq m} \frac{\binom{N-r-2}{n-m-2}}{n\binom{N}{n}}$$

$$= m\binom{r}{m}(N-r)(N-r-1)\, \frac{1}{(m+2)\binom{r+2}{m+2}}$$

[I-லிருந்து]

$$= \frac{m(N-r)(N-r-1)\binom{r}{m}}{\frac{(r+2)(r+1)}{m+1}\binom{r}{m}}$$

$$= \frac{m(m+1)(N-r)(N-r-1)}{(r+1)(r+2)}$$

$$V\left[n\frac{(r+1)}{m}-1\right]$$

$$= \frac{(r+1)^2}{m^2}\left[\frac{m(m+1)(N-r)(N-r-1)}{(r+1)(r+2)} + \frac{m(N-r)}{r+1} - \frac{m^2(N-r)^2}{(r+1)^2}\right]$$

$$= \frac{(r+1)}{m}\frac{(m+1)(N-r)(N-r-1)}{(r+2)} + \frac{r+1}{m}(N-r) - (N-r)^2$$

$$= \frac{(N-r)}{m(r+2)}[(r+1)(m+1)(N-r-1) + (r+1)(r+2) - m(r+2)(N-r)]$$

$$= \frac{(N-r)}{m(r+2)}[(m+1)(r+1)(N+1-r-2) + (r+1)(r+2) - m(r+2)(N+1-r-1)]$$

$$= \frac{(N-r)}{m(r+2)}[(N+1)\{(m+1)(r+1) - m(r+2)\} - (m+1)(r+1)(r+2) + m(r+1)(r+2) + (r+1)(r+2)]$$

$$= \frac{N-r}{m(r+2)}[(N+1)(r-m+1)]$$

∴ N-ன் மதிப்பீட்டியின் பரவற் படி

$$\frac{(N-r)(N+1)(r-m+1)}{m(r+2)}$$

இப் பரவற் படியை மதிப்பிடுவதற்கு $\frac{n(n-m)(r+1)}{m^2(m+1)}$ $\cdot(r-m+1)$ என்னும் மதிப்பீட்டியைக் கையாளவேண்டும் என்பதை எளிதாக நிரூபிக்கலாம். [II.கையாண்டு]

8. இரண்டு பெயர்ப் பட்டியல்கள் உள்ளன. முதல்பட்டியலில் M பெயர்களும், இரண்டாம் பட்டியலில் N பெயர்களும் உள்ளன. இப் பட்டியல்களில் D பெயர்கள் (D தெரியாது) பொதுவாக உள்ளன. D ஐ மதிப்பிடுவதற்கு முதல் பட்டியலிலிருந்து m உறுப்புகளும் இரண்டாம் பட்டியலிலிருந்து n உறுப்புகளும் ஈ. செயா. சா. ரா. மா. முறையில் எடுக்கப்பட்டு, இவைகளுக்குப் பொதுவாக d பெயர்கள் உள்ளன என்று கண்டுபிடிக்கப்படுகிறது. d-னுடைய பரவல் பின்வருமாறு :

$$P(d) = \binom{D}{d} f^d (1-f)^{D-d} \quad \left[f = \frac{mn}{MN}\right]$$

M, N முடிவிலியை (∞) எட்டுகின்றன. $D, \frac{m}{M}, \frac{n}{N}$ முதலியவை மாறிலிகளாக இருக்கின்றன. இந்த நிலையில் D ஐப் பிறழ்ச்சியின்றி எவ்வாறு மதிப்பிடுவாய்? அம்மதிப்பீட்டியின் பரவற் படி யாது? அப்பரவற் படிக்குப் பிறழ்ச்சியற்ற மதிப்பீட்டியொன்று யாது?

[Deming, W. E. and Glasser, G. J., J. Amer. Stat. Assn. 54, (1959)]

விடை

$$E^{-1}(D) \sim \left(\frac{MN}{mn}\right)d \;, \quad V\left(\frac{MN}{mn}d\right) = \left(\frac{MN}{mn} - 1\right)D$$

$$E^{-1}\left[V\left(\frac{MN}{mn}d\right)\right] = \frac{MN}{mn}\left(\frac{MN}{mn} - 1\right)d$$

9. ஒரு நாட்டில் N தொழிற்சாலைகள் இருக்கின்றன. இவைகளின் மொத்த உற்பத்தியை மதிப்பிட வேண்டியிருக்கிறது. மதிப்பீடு உண்மையான மதிப்பிலிருந்து 10%-க்கு மேல் வேறுபடக்கூடாதென்பதற்கு நிகழ்திறம் 95% என்றால் பின்வரும் நிலைமைகளில் மாதிரிப் பருமன் 'n' எவ்வளவாக இருக்க வேண்டும்?

(அ) ஈ. செயா. சா. ரா. மா. முறையை உபயோகப்படுத்தும் பொழுது

(i) $N=500$ (ii) 1000 (iii) 2500 (iv) 5000 (v) 10000

(ஆ) ஈ. செ. சா. ரா. மா. முறையை உபயோகப்படுத்தும் பொழுது

(i) $N=500$ (ii) 1000 (iii) 2500 (iv) 5000 (v) 10000

முழுமைத் தொகுதியின் மாறுபாட்டுக் கெழு (coefficient of variation) 60% என்றும், மதிப்பீடுகள் இயல்நிலைப் பரவலைக் கொண்டவை எனவும் எடுத்துக் கொள்ளவும்.

விடை

(அ) உண்மையான மொத்த உற்பத்தி Y எனக்கொண்டால் அதன் மதிப்பீட்டி $N\bar{y}$ ஆகும்.

$$P\left[\,|N\bar{y}-Y| \leq \frac{Y}{10}\right] = 0\cdot95$$

$$\therefore \quad P\left[\,|\bar{y}-\bar{Y}| \leq \frac{\bar{Y}}{10}\right] = 0\cdot95$$

$$P\left[\frac{|\bar{y}-\bar{Y}|}{\sqrt{V(\bar{y})}} \leq \frac{\bar{Y}}{10\sqrt{V(\bar{y})}}\right] = 0\cdot95$$

ஆனால் $V(\bar{y}) = \dfrac{N-n}{N-1}\,\dfrac{\sigma^2}{n}$

$$\therefore \quad \frac{\bar{Y}}{\sqrt{V(\bar{y})}} = \frac{\bar{Y}}{\sqrt{\dfrac{N-n}{n(N-1)}}\,\sigma} = \frac{5}{3\sqrt{\dfrac{N-n}{n(N-1)}}}$$

$$\left[\because \frac{\sigma}{\bar{Y}} = \cdot6 = \frac{3}{5}\right]$$

$$\therefore P\left[\,|Z| \le \frac{1}{6\sqrt{\frac{N-n}{n(N-1)}}}\right] = 0{\cdot}95$$

$$[Z \frown N(0,1)]$$

$$\therefore \frac{1}{6\sqrt{\frac{N-n}{n(N-1)}}} = 1{\cdot}96$$ [இயல் நிலைப் பரவல் அட்டவணையிலிருந்து]

$$\frac{N-n}{n(N-1)} = \frac{1}{36\times(1{\cdot}96)^2}$$

$$\therefore n\,[\,(N-1) + 36\,(1{\cdot}96)^2\,] = N\times 36\,(1{\cdot}96)^2$$

$$\therefore n = \frac{N\times 36\,(1{\cdot}96)^2}{N-1+36(1{\cdot}96)^2}$$

இதில் $N = 500$, 1000, 2500, 5000, 10000 புகுத்தி n-ன் மதிப்பை 109, 123, 132, 136, 137 எனக் கண்டுபிடிக்கலாம்.

(ஆ) $V(\bar{y}) = \dfrac{\sigma^2}{n}$

$$\therefore P\left[\,|Z| \le \frac{\bar{Y}}{10}\frac{\sqrt{n}}{\sigma}\right] = 0{\cdot}95$$

$$\therefore P\left[\,|Z| \le \frac{\sqrt{n}}{6}\right] = 0{\cdot}95 \quad \therefore \frac{\sqrt{n}}{6} = 1{\cdot}96$$

$$\therefore n = 36\times(1{\cdot}96)^2 = 139.$$

இதில் ஆச்சரியப்படத்தக்க விவரம் என்னவெனில் 'n', N ஐச் சாராமலேயே இருக்கிறது. அதாவது N=500, 1000, 2500, 5000, 10000 முதலிய எல்லா மதிப்புகளுக்கும் n=139 தான். இவ் விதத்தில் இது (அ) -விருந்து மாறுபடுகிறது.

10. 2000 கல்லூரிகளிலிருந்து 200 கல்லூரிகள் ஈ. செயா. சா. ரா. மா. முறையில் எடுக்கப்படுகின்றன. இவைகளில் 140 கல்லூரிகள் புதிய தேர்வு முறையை ஆதரிக்கின்றன.

(அ) முழுமைத் தொகுதியில் புதிய முறையை ஆதரிக்கும் கல்லூரிகளின் எண்ணிக்கைக்கு 95% நம்பிக்கை இடை வெளியைக் கண்டுபிடி.

(ஆ) முழுமைத் தொகுதியில் புதிய முறையை 50% கல்லூரிகள்தாம் ஆதரிக்கின்றன என்னும் எடுகோளை (Hypothesis) நிராகரிப்பதற்குக் கொடுக்கப்பட்ட விவரங்கள் போதுமானவையா?

(இ) மாதிரிப் பருமன் n எவ்வளவாக இருந்தால் (அ) -ல் நிறுவப்பட்ட நம்பிக்கை இடைவெளியின் நீளம் பாதியாகக் குறையும்?

விடை :

(அ) $p = \dfrac{140}{200} = 0{\cdot}7.$

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

$$= \frac{2000-200}{200 \times 1999} 0{\cdot}7 \times 0{\cdot}3$$

$$= \frac{1800 \times 0{\cdot}7 \times 0{\cdot}3}{200 \times 1999}$$

$$\therefore \quad \sqrt{V(p)} = 0{\cdot}03075$$

P-க்கு இயல்நிலைப் பரவல் இருப்பதாகக் கொண்டால் P-ன் 95% நம்பிக்கை இடைவெளி $[P \pm 1{\cdot}96 \times {\cdot}03075]$

$$= [0{\cdot}7 \pm {\cdot}0603] = [0{\cdot}6397, 0{\cdot}7603]$$

$\therefore$ 2000 p-க்கு 95% நம்பிக்கை இடைவெளி
$[0{\cdot}6397 \times 2000, 0{\cdot}7603 \times 2000] = [1279, 1521]$

(ஆ) $H_0 ; \ P = 0{\cdot}5$
$H_1 : \ P \neq 0{\cdot}5$

$$\frac{|p-P|}{\sqrt{V(P)}} = \frac{|0{\cdot}7-0{\cdot}5|}{\sqrt{\dfrac{1800}{200 \times 1999} 0{\cdot}5 \times 0{\cdot}5}}$$

$$= \frac{0{\cdot}2}{0{\cdot}5 \times 3} \sqrt{1999} > 1{\cdot}96$$

ஆகவே H_0 எனும் எடுகோளைக் கண்டிப்பாய் நிராகரிக்கலாம்.

அதாவது 95% அளவில் 50% கல்லூரிகள் புதிய முறையை வரவேற்கின்றன என்னும் எடுகோளை நிராகரித்து விடலாம்.

(இ) நம்பிக்கை இடைவெளியின் நீளம் பாதியாக வேண்டுமாயின் $\sqrt{V(p)} = \frac{\cdot 03075}{2}$ ஆக இருக்கவேண்டும்.

$$\text{அதாவது} \quad \frac{2000 - n}{n \times 1999} \, 0{\cdot}7 \times 0{\cdot}3 = \left(\frac{\cdot 03075}{2}\right)^2$$

$$\therefore \quad n\left[1999 \times \left(\frac{\cdot 03075}{2}\right)^2 + 0{\cdot}21\right] = 2000 \times 0{\cdot}21.$$

$$\therefore \quad n = \frac{2000 \times 0{\cdot}21}{1999 \times \left(\frac{\cdot 03075}{2}\right)^2 + 0{\cdot}21} = \frac{420}{0{\cdot}4725 + 0{\cdot}21}$$

$$= \frac{420}{0{\cdot}6825} = 615.$$

19. முறையுடை மாதிரி முறை

(SYSTEMATIC SAMPLING)

19·1. சாதாரண ராண்டம் மாதிரி முறையில் எல்லா உறுப்புகளுக்கும் சம வாய்ப்பு அளிக்கப்படுகிறது. மேலும் ஈ. செ. சா. ரா. மா. முறையில் மொத்தம் வரக்கூடிய N^n மாதிரிகளுக்கும் சம வாய்ப்பு அளிக்கப்படுகிறது. பொதுவாக n உறுப்புகளுள்ள மாதிரிகளை எடுத்துக்கொண்டால் இம்முறையில்தான் மிக அதிகமான மாதிரிகள் கிடைக்கின்றன. ஈ. செயா. சா. ரா. மா. முறையில் NP_n மாதிரிகள் உள்ளன. இவை யாவைக்கும் சமவாய்ப்பு அளிக்கப்படுகிறது.

பொதுவாகச் சா.ரா.மா. முறைகள் நடைமுறையில் பலவிதத் தொந்தரவுகளை அளிக்கின்றன. மாதிரிகள் எடுப்பது சாதாரணமாகக் கடினம்தான். இதற்கு முக்கியக் காரணமே மிக அதிகமான அளவில் மாதிரிகள் இருப்பதுதான். ஆகவே மாதிரிகள் எடுப்பது எளிதாக இருக்கவேண்டுமாயின் மாதிரிகளின் மொத்த எண்ணிக்கை குறைவாக இருக்கவேண்டும் என்பது தெரிகிறது. உதாரணமாக 10 உறுப்புகளுள்ள முழுமைத் தொகுதியிலிருந்து 4 உறுப்புகள் கொண்ட ஒரு மாதிரி எடுக்கவேண்டுமாயின் ஈ. செ. சா. ரா. மா. முறையில் மொத்தம் 10000 மாதிரிகள் உள்ளன. முழுமைத் தொகுதியின் பருமனை ஒப்பிடும்போது இது மிகவும் அதிகம். ஈ. செயா. சா. ரா. மா. முறையில் மொத்தம் 5040 மாதிரிகள் உள்ளன. இதுவும் மிக அதிகமே. சிறிய முழுமைத் தொகுதிக்கே இந்த நிலையெனில் பெரிய முழுமைத் தொகுதிகளைப் பற்றிக் கேட்கவே வேண்டியதில்லை.

மேற்கூறிய இன்னலை நீக்குவதற்கு முறையுடை மாதிரி முறைகள் கையாளப்படுகின்றன. இம்முறைகளில் மொத்த மாதிரிகளின் எண்ணிக்கையே சிறியதாகத்தான் இருக்கும். ஆகவே நடைமுறையில் இதைக் கையாள்வது மிகவும் எளிது.

19·2. நேர்க்கோட்டு முறையுடை மாதிரி முறை (Linear Systematic Sampling) [நே. மு. மா. முறை].

முழுமைத் தொகுதியில் U_i, $i = 1, 2, \ldots 12$ என்று 12 உறுப்புகள் இருக்கின்றன எனக் கொள்வோம். 4-உறுப்புகள் கொண்ட மாதிரியொன்று தேவைப்படுகிறதென்றும் கொள்வோம். இதைப் பின்வருமாறு எடுக்கலாம். 1-லிருந்து 3-க்குள் ஓர் எண்ணை ராண்டம் முறையில் எடுக்கவேண்டும். இந்த ராண்டம் எண் r என்றால் மாதிரியானது

$$U_r \quad U_{r+3} \quad U_{r+6} \quad U_{r+9}.$$

இம்முறையில் இந்த உதாரணத்தில் மொத்த மாதிரிகளின் எண்ணிக்கை 3 தான். அவையாவன.

1. $U_1 \quad U_4 \quad U_7 \quad U_{10}$
2. $U_2 \quad U_5 \quad U_8 \quad U_{11}$
3. $U_3 \quad U_6 \quad U_9 \quad U_{12}$

முழுமைத் தொகுதியில் 13 உறுப்புகள் இருந்தால் பின்வருவன மாதிரிகளாகும்.

1. $U_1 \quad U_4 \quad U_7 \quad U_{10} \quad U_{13}$
2. $U_2 \quad U_5 \quad U_8 \quad U_{11}$
3. $U_3 \quad U_6 \quad U_9 \quad U_{12}$

இந்த உதாரணத்தில் முதல் மாதிரியில் 4 உறுப்புகளுக்குப் பதிலாக 5 உறுப்புகள் இருக்கின்றன. பொதுவாக, முழுமைத் தொகுதியில் N உறுப்புகள் இருந்து மாதிரியில் n உறுப்புகள் வேண்டுமாயின் பின்வரும் முறையைக் கையாளவேண்டும்.

$\dfrac{N}{n}$ என்னும் பின்னத்தின் முழு எண் பகுதி k ஆக இருக்கட்டும்.

உதாரணமாக $N = 25$, $n = 6$ என்றால், $k = 4$, $N = 30$, $n = 6$ என்றால் $k = 5$. 1-லிருந்து k வரையுள்ள எண்களுள் ஒன்றை ராண்டம் முறையில் தேர்ந்தெடு. இது r என்றால், U_r என்னும் ராண்டம் துவக்கத்திலிருந்து துவங்கி U_{r+k}, U_{r+2k}, ... என்று $(k-1)$ உறுப்புகளை விட்டுவிட்டு வரும் எல்லா உறுப்புகளையும் எடுக்கவும். இதுதான் நேர்க்கோட்டு முறையுடை மாதிரி ஒன்று.

எடுக்கும் முறையிலிருந்தே மொத்தம் k மாதிரிகள்தாம் இருக்கும் என்பது தெரிகிறது. அவையாவன:

1 U_1, U_{k+1}, U_{2k+1}, ...

2. U_2, U_{k+2}, U_{2k+2}, ...

3. U_3, U_{k+3}, U_{2k+3}, ...

⋮

k U_k, U_{2k}, U_{3k}, ...

கவனிக்கவும்

$\frac{N}{n}$ என்பது முழு எண்ணுக இல்லாமல் கலப்பு எண்ணுக இருந்தால் எல்லா மாதிரிகளிலும் சரியாக n உறுப்புகள் இரா. சில மாதிரிகளில் சில உறுப்புகள் அதிகமாகவே இருக்கும். உதாரணமாக $N = 199$, $n = 80$ எனக் கொண்டால் $k = 2$.

∴ முதல் மாதிரியில் 100 உறுப்புகளும் இரண்டாவது மாதிரியில் 99 உறுப்புகளும் இருக்கும்.

மேற்கூறிய பண்பு நேர்க்கோட்டு முறையுடை முறையின் விரும்பத்தகாத ஒன்ருகும்.

19·3. தேற்றம்

நேர்க்கோட்டு முறையுடை மாதிரி முறையில், $\frac{N}{n}$ என்பது முழு எண்ணுக இருந்தால், முழுமைத் தொகையின் சராசரிக்கு மாதிரிச் சராசரி பிறழ்ச்சியற்ற மதிப்பீட்டியாகும்.

நிறுவல்

$\frac{N}{n} = k =$ ஒரு முழு எண் என எடுத்துக்கொண்டால் மொத்தம் k மாதிரிகள் இருக்கும். ஒவ்வொரு மாதிரியிலும் சரியாக n உறுப்புகள் இருக்கும்.

முழுமைத் தொகுதியிலுள்ள உறுப்புகள் $X_1, X_2, \ldots X_N$ என எடுத்துக் கொள்வோம். $\bar{x}$ என்பது மாதிரிச் சராசரியானல் அதன் பரவல் பின்வருமாறு:

$\bar{x}$	$\frac{1}{n}\sum_{r=0}^{n-1} X_{1+rk}$	$\frac{1}{N}\sum_{r=0}^{n-1} X_{s+rk}$	......	$\frac{1}{n}\sum_{r=0}^{n-1} X_{k+rk}$
P	$\frac{1}{k}$	$\frac{1}{k}$		$\frac{1}{k}$

$$\therefore \quad E(\bar{x}) = \sum_{s=1}^{k} \frac{1}{k}\frac{1}{n}\sum_{r=0}^{n-1} X_{s+rK}$$

$$= \frac{1}{kn}\sum_{s=1}^{k}\sum_{r=0}^{n-1} X_{s+rK}$$

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

$\therefore \quad E^{-1}(\bar{X}) \sim \bar{x}$ (நி-து.)

19·4. $\frac{N}{n}$ முழு எண்ணாக இல்லாமல் கலப்பு எண்ணாக இருந்தால் மாதிரிச் சராசரி முழுமைத் தொகுதியின் சராசரியின் பிறழ்ச்சியற்ற மதிப்பீட்டியாக இருக்கவேண்டிய கட்டாயமில்லையென்பது பின் வரும் உதாரணத்தின் மூலம் தெரியவரும்.

இனத்தொகுதி 2, 4, 6, 8, 10, 16, 18

$n = 3$

$\therefore$ முதல் மாதிரி 2, 6, 10, 18. இதன் சராசரி 9

இரண்டாம் மாதிரி 4, 8, 16. இதன் சராசரி $\frac{28}{3}$

$$E(\bar{x}) = \frac{1}{2}\cdot 9 + \frac{1}{2}\cdot\frac{28}{3} = \frac{9}{2} + \frac{14}{3} = \frac{55}{6}$$

$$\bar{X} = \frac{1}{7}[2+4+6+8+10+16+18] = \frac{1}{7} \times 64 = \frac{64}{7}$$

இதிலிருந்து $E(\bar{x}) \neq \bar{X}$ என்பது தெரிகிறது.

ஆகவே $\frac{N}{n}$ கலப்பு எண்ணாக இருந்தால் மாதிரிச் சராசரி முழுமைத் தொகுதியின் சராசரிக்குப் பிறழ்ச்சியுள்ள மதிப்பீட்டியாகும்.

19·5. தேற்றம்

x என்பது மாதிரியின் மொத்தமானால் $\frac{k}{N}$ x என்பது முழுமைத் தொகுதியின் சராசரியின் பிறழ்ச்சியற்ற மதிப்பீட்டியாகும்.

நிறுவல்

ஒவ்வொரு மாதிரிக்கும் நிகழ்திறம் $\frac{1}{k}$. $x^{(1)}, x^{(2)}$ ....$x^{(k)}$ என்பவை k மாதிரிகளின் மொத்தங்களானால் $\frac{k}{N}$ x-ன் பரவல் பின்வருமாறு:

$\frac{kx}{N}$	$\frac{kx^{(1)}}{N}$	$\frac{kx^{(2)}}{N}$	$\frac{kx^{(3)}}{N}$	....	$\frac{kx^{(K)}}{N}$
P	$\frac{1}{k}$	$\frac{1}{k}$	$\frac{1}{k}$	.	$\frac{1}{k}$

$$\therefore E\left(\frac{k\,x}{N}\right) = \sum_{i=1}^{k} \frac{1}{k} \cdot \frac{k}{N} x^{(i)} = \frac{1}{N} \sum_{i=1}^{k} x^{(i)} = \frac{1}{N}\sum_{i=1}^{N} X_i \; \bar{X}$$

$$\therefore E^{-1}(\bar{X}) \sim \frac{kx}{N} \qquad \text{(நி — து.)}$$

கவனிக்கவும்

தேற்றம் 19·3, தேற்றம் 19·5-லிருந்து தொடர்கிறது. ஏனெனின் $\frac{N}{n}$ முழு எண்ணாக இருந்தால் $\frac{N}{n} = k. \therefore \frac{k}{N} = \frac{1}{n}$

$$\therefore \frac{kx}{N} = \frac{1}{n} x = \bar{x}.$$

$$\therefore E(x) = E\left(\frac{kx}{N}\right) = \bar{X}$$

உதாரணம்

முழுமைத் தொகுதி 2, 4, 6, 8, 10, 16, 18

$n=3 \qquad \therefore k=2$

முதல் மாதிரி 2, 6, 10, 18. இதன் மொத்தம் 36.

$$\therefore \frac{k}{N} x = \frac{72}{7}$$

இரண்டாம் மாதிரி 4, 8, 16. இதன் மொத்தம் 28.

$$\therefore \frac{k}{N} x = \frac{56}{7}$$

$$\therefore E(\bar{x}) = \frac{1}{2} \cdot \frac{72}{7} + \frac{1}{2} \cdot \frac{56}{7} = \frac{1}{14} [72+56] = \frac{128}{14}$$

$$= \frac{64}{7} = \bar{X}$$

19·6. நேர்க்கோட்டு முறையுடை மாதிரி முறையில், $\frac{N}{n}$ முழு எண்ணாக இருக்கும் பொழுது மாதிரிச் சராசரியின் பரவற் படி.

$\bar{x}^{(1)}, \bar{x}^{(2)}, \ldots; \bar{x}^{(K)}$ என்பவை மாதிரிச் சராசரிகளாக இருக்கட்டும்.

$$\bar{x}^{(i)} = \frac{1}{n} \sum_{r=0}^{n-1} X_{i+rk}$$

$$\therefore \quad V(\bar{x}) \; E(\bar{x}-\bar{X}) = \frac{1}{k}\sum_{i=1}^{k} (\bar{x}^{(i)} - \bar{X})^2$$

$$= \frac{1}{k}\sum_{i=1}^{k}\left[\frac{1}{n}\sum_{r=0}^{n-1} X_{i+rk}-\bar{X}\right]^2$$

$$= \frac{1}{kn^2}\sum_{i=1}^{k}\left[\sum_{r=0}^{n-1} (X_{i+rk}-\bar{X})\right]^2$$

$$= \frac{1}{kn^2}\sum_{i=1}^{k}\sum_{r=0}^{n-1}(X_{i+rk}-\bar{X})^2 + \frac{2}{kn^2}\sum_{i=1}^{k}\sum_{r<r'}(X_{i+rk}-\bar{X})(\bar{X}_{i+r'K}-\bar{X})$$

$$= \frac{1}{n}\sigma^2 + \frac{2}{Nn}\sum_{i=1}^{k}\sum_{r<r'}(X_{i+rk}-\bar{X})(X_{i+r'K}-\bar{X})$$

$$\therefore \quad V(\bar{x}) = \frac{\sigma^2}{n}\left[1 + \frac{2}{N\sigma^2}\sum_{i=1}^{k}\sum_{r<r'}(X_{i+rk}-\bar{X})(X_{i+r'k}-\bar{X})\right]$$

K மாதிரிகளை வகுப்புகளாக எடுத்துக்கொண்டால் இவைகளுக்கு வகுப்புள் ஒட்டுறவுக் கெழு (Intra-class correlation coefficient) ρ_c என்பதைப் பின்வருமாறு வரையறுக்கலாம்.

$$\rho_c = \frac{2}{kn(n-1)\sigma^2}\sum_{i=1}^{k}\sum_{r<r'}(X_{i+rk}-\bar{X})(X_{i+r'k}-\bar{X})$$

$$\therefore \quad V(\bar{x}) = \frac{\sigma^2}{n}[1+(n-1)\rho_c]$$

19·7. கிளைத்தேற்றம்

நேர்க்கோட்டு முறையுடை மாதிரிச் சராசரி, ஈ. செ. சா. ரா. மா. முறைச் சராசரியை விட முழுமைத் தொகுதிச் சராசரியை மதிப்பிடுவதில் திறனதிகம் உள்ளதாக இருக்கவேண்டுமானால் $\rho_c < 0$ என்பது வேண்டியதும் போதுமானதுமான நிபந்தனை யாகும் (Necessary and sufficient condition).

நிறுவல்

$$\frac{\sigma^2}{n}\left[1+(n-1)\rho_c\right] < \frac{\sigma^2}{n} \Longleftrightarrow 1+(n-1)\rho_c < 1$$

$$\Longleftrightarrow (n-1)\rho_c < 0 \Longleftrightarrow \rho_c < 0 \qquad \text{(நி—து.)}$$

இத்தேற்றத்திலிருந்து $\rho_c < 0$ ஆக இருக்கும்படி முழுமைத் தொகுதியுறுப்புகள் வரிசைப்படுத்தப் பட்டிருந்தால் நேர்க்கோட்டு முறையுடை மாதிரி முறை ஈ. செ. சா. ரா, மா. முறையை விடத் திறனதிகமுள்ளதாக இருக்குமென்பது தெரிகிறது. மேலும் $V(\bar{x})$ முழுமைத் தொகுதியுறுப்புகள் எவ்வாறு வரிசைப்படுத்தப் பட்டுள்ளன என்பதைப் பொறுத்திருக்கிறது என்பதும் தெளிவாகிறது. ஆகவே நே. மு. மா. முறையில் முழுமைத் தொகுதியுறுப்புகள் தகுந்த முறையில் வரிசைப்படுத்தப் பட வேண்டியது அவசியமாகிறது.

19·8. மாதிரிச் சராசரியின் பரவற் படிக்கு மற்றொரு வடிவம்

$$\sigma^2 = \frac{1}{N}\sum_{i=1}^{k}\sum_{r=0}^{n-1}\left(X_{i+rk}-\bar{X}\right)^2$$

$$= \frac{1}{N}\sum_{i=1}^{k}\sum_{r=0}^{n-1}\left(X_{i+rk}-\bar{x}^{(i)}+\bar{x}^{(i)}-\bar{X}\right)^2$$

$$= \frac{1}{N}\sum_{i=1}^{k}\sum_{r=0}^{n-1}\left[X_{i+rk}-\bar{x}^{(i)}\right]^2$$

$$+\frac{1}{N}\sum_{i=1}^{k}\sum_{r=0}^{n-1}\left(\bar{x}^{(i)}-\bar{X}\right)^2$$

ஏனென்றால் $\sum_{i=1}^{k} \sum_{r=0}^{n-1} [X_{i+rk} - \bar{x}^{(i)}] [\bar{x}^{(i)} - \bar{X}]$

$$= \sum_{i=1}^{k} (\bar{x}^{(i)} - \bar{X}) \sum_{r=0}^{n-1} [X_{i+rk} - \bar{x}^{(i)}] = 0$$

$$\therefore \quad \sigma^2 = \frac{1}{N} \sum_{i=1}^{k} \sum_{r=0}^{n-1} (X_{i+rk} - \bar{x}^{(i)})^2$$

$$+ \frac{1}{k} \sum_{i=1}^{k} (\bar{x}^{(i)} - \bar{X})^2$$

$$= \frac{1}{N} \sum_{i=1}^{k} \sum_{r=0}^{n-1} (X_{i+rk} - \bar{x}^{(i)})^2 + V(\bar{x})$$

$\frac{1}{N} \sum_{i=1}^{k} \sum_{r=0}^{n-1} (X_{i+rk} - \bar{x}^{(i)})^2$ என்பது மாதிரியுள் பரவற் படி (Within-sample variance) எனப்படும். இதை σ_w^2 எனக் குறிப்பிடுவோம்.

$$\therefore \quad \sigma^2 = \sigma_w^2 + V(\bar{x})$$

$$\therefore \quad V(\bar{x}) = \sigma^2 - \sigma_w^2$$

19·9. கிளைத்தேற்றம்

முழுமைத் தொகுதியைப் பொறுத்தவரை σ^2 என்பது மாறிலியாகும். ஆனால் σ^2_w என்பது உறுப்புகள் மாதிரிகளில் அமைக்கப்பட்டிருப்பதைப் பொறுத்திருக்கிறது. ஆகவே நே.மு.மா. முறை அதிகப் பயனை அளிக்கவேண்டுமாயின் σ_w^2 மிகவும் அதிகமாக இருக்கவேண்டும். அதாவது ஒவ்வொரு மாதிரிக்குள்ளும் உறுப்புகள் மிகவும் வேறுபட்டிருக்கவேண்டும்.

19·10. கிளைத்தேற்றம்

முன்பே $V(\bar{x}) = \frac{\sigma^2}{n}[1+(n-1)\rho_c]$ என்று நிறுவப்பட்டது. ஆனால் $V(\bar{x}) = \sigma^2 - \sigma_w^2$

$$\therefore \quad \sigma^2 - \sigma_w^2 = \frac{\sigma^2}{n}[1 + (n-1)\rho_c]$$

$$\therefore \quad n - \frac{n\,\sigma_w^2}{\sigma^2} = 1+(n-1)\rho_c$$

$$\therefore \quad \rho_c = \frac{1}{n-1}\left[n - \frac{n\,\sigma_w^2}{\sigma^2} - 1\right]$$

$$= 1 - \frac{n}{n-1}\,\frac{\sigma\omega^2}{\sigma^2}$$

இதிலிருந்து $\rho_c \leq 1$ என்பது தெரிகிறது.

மேலும் $V(\bar{x}) = \sigma^2 - \sigma_w^2 \geq 0$

$$\therefore \quad \sigma^2_w \leq 1$$

$$\therefore \quad \rho_c \geq 1 - \frac{n}{n-1}\cdot 1 = \frac{n-1-n}{n-1} = -\frac{1}{n-1}$$

ஆகவே $\frac{-1}{n-1} \leq \rho_c \leq 1$.

19·11. கிளைத்தேற்றம்

நே. மு. மா. முறையில் $V(\bar{x}) = \frac{\sigma^2}{n}[1 + \rho_c(n-1] = V_1$ எனக் கொள்க.

ஈ. செயா. சா. ரா. மா. முறையில் $V(\bar{x}) = \frac{N-n}{N-1}\,\frac{\sigma^2}{n} = V_2$ எனக் கொள்க.

நே. மு. மா. முறை . ஈ. செயா. சா. ரா. மா. முறையை விடத் திறனதிகம் உடையதாக இருந்தால் $V_1 < V_2$.

$$\therefore \quad \frac{\sigma^2}{n}[1+(n-1)\rho_c] < \frac{N-n}{N-1}\,\frac{\sigma^2}{N}$$

$$\therefore \quad 1+(n-1)\rho_c < \frac{N-n}{N-1} = 1-\frac{n-1}{N-1}$$

$$\therefore \quad (n-1)\rho_c < -\frac{n-1}{N-1}.$$

$$\therefore \quad \rho_c < -\frac{1}{N-1}.$$

இதிலிருந்து நே. மு மா. முறை சில சமயங்களில் ஈ. செயா. சா. ரா. மா. முறையை விடச் சிறந்ததாக (திறனதிகம் வாய்ந்ததாக) இருக்க முடியுமென்பது விளங்குகிறது.

19·12. நே. மு. மா. முறையில் $V(\bar{x})$ ஐ மதிப்பிடல்

இம்முறையில் ஒரே ஒரு மாதிரியைக் கொண்டு $V(\bar{x})$ ஐப் பிறழ்ச்சியின்றி மதிப்பிட முடியாது. இதைப் பிறழ்ச்சியற்ற மதிப்பீட்டிகளின் பண்புகளின் வாயிலாகத் தெரிந்து கொள்ளலாம்.

$$E^{-1}(\bar{X}) \sim \bar{x}$$

$$\therefore \quad E^{-1}(X) \sim N\bar{x}.$$

$$E^{-1}(\sigma^2) \sim E^{-1}\left[\frac{1}{N}\sum_{i=1}^{N} X_i^2 - \bar{X}^2\right]$$

$$\sim \frac{1}{N}E^{-1}\left[\sum_{i=1}^{N} X_i^2\right] - \frac{1}{N^2}E^{-1}\left[\sum_{i=1}^{N} X_i^2\right]$$

$$-\frac{2}{N^2}E^{-1}\left[\sum_{i<j} X_i X_j\right]$$

$$\therefore \quad \frac{2}{N^2} E^{-1}\left[\sum_{i<j} X_i X_j\right] \sim E^{-1}(\sigma^2) - \frac{N-1}{N^2} E^{-1}\left[\sum_{i=1}^{N} X_i^2\right]$$

இதிலிருந்து σ^2 ஐயும் $\sum_{i=1}^{N} X_i^2$ ஐயும் பிறழ்ச்சியின்றி மதிப்பிட முடிந்தால் $\sum_{i<j} X_i X_j$ என்பதையும் பிறழ்ச்சியின்றி மதிப்பிட முடியும்.

ஆகவே σ^2 ஐப் பிறழ்ச்சியின்றி மதிப்பிட முடியும் என்று எடுத்துக்கொள்வோம். $\sum_{i=1}^{N} X_i^2$ ஐ $k \sum_{i=1}^{n} x_i^2$ என்பது பிறழ்ச்சியின்றி மதிப்பிடும். ஆகவே $\sum_{i<j} X_i X_j$ என்பது மதிப்பிட முடியும் என்பது தெரிகிறது, இந்தக் கூட்டுத் தொகையில் மொத்தம் $\binom{N}{2}$ உறுப்புகள் உள்ளன. ஆனால் நே. மு. மாதிரி ஒவ்வொன்றிலும் $\binom{n}{2}$ உறுப்புகள் ($X_i X_j$ என்பன போன்ற)தாம் இருக்கும். ஆகவே எல்லா மாதிரிகளையும் எடுத்துக்கொண்டால் கூட மொத்தம் $k\binom{n}{2}$ இரட்டைகள் (Pairs) தானிருக்கும்.

$$k\binom{n}{2} = \frac{kn(n-1)}{2} = \frac{N(n-1)}{2} < \frac{N(N-1)}{2} =$$

$\binom{N}{2}$. ஆகவே இவைகளைக் கொண்டு $\sum_{i<j} X_i X_j$ பிறழ்ச்சி

யின்றி மதிப்பிடமுடியாது. என்பது விளங்குகிறது. ஆகவே σ^2 பிறழ்ச்சியின்றி மதிப்பிடப்படமுடியாது.

ஆனால் ஒன்றுக்கு மேற்பட்ட நே. மு. மாதிரிகள் எடுக்கப் பட்டால் அவைகளிலிருந்து $V(\bar{x})$ பிறழ்ச்சி யின்றி மதிப்பிடப்பட முடியும்.

19·13. தேற்றம்

$\bar{x}_1, \bar{x}_2, \ldots \bar{x}_m$ என்பவை m நே. மு. மாதிரிச் சராசரிகளானல் (ஒவ்வொன்றிலும் சார்பின்றித் தனித்தனியாக ஈ. செ. சா. ரா. மா. முறையில் ராண்டம் துவக்கம் தேர்ந்தெடுக்கப்பட்டது எனக் கொள்க.)

$$\bar{x} = \frac{1}{m}\sum_{i=1}^{m} \bar{x}_i$$ என்பது $\bar{X}$-ன் பிறழ்ச்சியற்ற மதிப்பீட்டி யாகும். $\bar{x}$-ன் பரவற் படிக்கு பிறழ்ச்சியற்றதொரு மதிப்பீட்டி

$$\frac{1}{m(m-1)}\sum_{i=1}^{m} (\bar{x}_i - \bar{x})^2$$ ஆகும்.

நிறுவல்

$$E(\bar{x}) = \frac{1}{m}\sum_{i=1}^{m} E(\bar{x}_i) = \frac{1}{m} m\bar{X} = \bar{X}.$$

$\bar{x}_i$-ன் பரவல் :

$\bar{x}_i$	$\bar{x}^{(1)}$	$\bar{x}^{(2)}$	$\bar{x}^{(3)}$	....	$\bar{x}^{(k)}$
P	$\frac{1}{k}$	$\frac{1}{k}$	$\frac{1}{k}$		$\frac{1}{k}$

$$V(\bar{x}) = V\left[\frac{1}{m}\sum_{i=1}^{m} \bar{x}_i\right] = \frac{1}{m^2}\sum_{i=1}^{m} V(\bar{x}_i)$$

$$= \frac{1}{m^2}\, m.\frac{1}{k} \sum_{i=1}^{k} (x^{(i)} - \bar{X})^2$$

$$= \frac{1}{mk}\sum_{i=1}^{k} (\bar{x}^{(i)} - \bar{X})^2$$

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

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

$$= \frac{1}{(m-1)}\left[\frac{1}{m}\ \frac{m}{k}\sum_{i=1}^{k} (\bar{x}^{(i)})^2 - V(\bar{x})\ \ \bar{X}^2\right]$$

$$= \frac{1}{(m-1)}\left[\frac{1}{k}\sum_{i=1}^{k} (\bar{x}^{(i)})^2 - \bar{X}^2 - V(\bar{x})\right]$$

$$= \frac{1}{(m-1)}\left[\frac{1}{k}\sum_{i=1}^{k} (\bar{x}^{(i)} - \bar{X})^2 - V(\bar{x})\right]$$

$$= \frac{1}{(m-1)}[m\,V(\bar{x}) - V(\bar{x})] = V(\bar{x})$$

$$\therefore\ E^{-1}[V(\bar{x})] \sim \frac{1}{m(m-1)}\sum_{i=1}^{m} (\bar{x}_i - \bar{x})^2$$

19·14. வட்ட முறையுடை மாதிரி முறை (Circular systematic sampling)

நேர்க்கோட்டு முறையுடை மாதிரி முறையில் நமக்குத் தேவையான மாதிரிப் பருமன் சில சமயங்களில் $\left(\frac{N}{n}\right.$ முழு எண்ணாக இல்லாத போது $\left.\right)$ கிடைப்பதில்லையென்பதை நாம் பார்த்தோம். இது ஒரு முக்கிய பலவீனமாகும். இதைத் தவிர்ப்பதற்கென்றே வட்ட முறையுடை மாதிரி முறை அமைக்கப்பட்டது. இம்முறை பின்வருவாறு.

முழுமைத் தொகுதியில் $U_1, U_2, \ldots, U_N$ என்ற N உறுப்புகள் இருப்பதாகக் கொள்ளவோம். இவைகளை ஒரு வட்டத்தின் விளிம்பில் வரிசையாக வைப்பதாகக் கொள்வோம். k என்பது $\left(\frac{N}{n}\right)$ என்னும் எண்ணின் முழு எண் பகுதியாக இருக்கட்டும். 1-லிருந்து N-வரையுள்ள எண்களிலிருந்து ஓர் எண்ணை ராண்டம் முறையில் எடுக்கவும். இது r என்று எடுத்துக் கொண்டால் U_r என்னும் ராண்டம் துவக்கத்திலிருந்து $(k-1)$ உறுப்புகளை விட்டு விட்டு வரிசையாக n உறுப்புகள் எடுக்கவேண்டும். அதாவது $U_r, U_{r+k}, U_{r+2k}, \ldots, U_{r+(n-1)k}$ [$r + sk$ என்பது N ஐவிட அதிகமாக இருந்தால் $r + sk - N$ என்று அதை மாற்றிவிட வேண்டும்.]

உதாரணமாக $N = 13$, $n = 4$ எனக் கொள்வோம். k என்பது 3 என்பது விளங்கும். r என்னும் ராண்டம் துவக்கம் 7 எனக் கொண்டால் மாதிரியிலுள்ள உறுப்புகள்

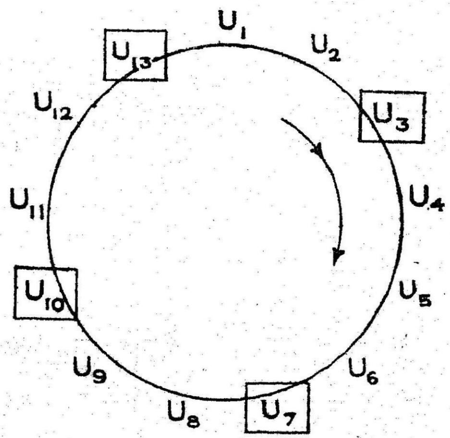
U_7 U_{10} U_{13} U_{16} 16 என்பது 13 ஐ விட அதிகமாகையால் அதை $16 - 13 = 3$ என்று மாற்ற வேண்டும். இம்மாற்றத்திற்குப் பிறகு மாதிரி U_7 U_{10} U_{13} U_3 ஆகும்.

இந்த உதாரணத்தில் மொத்தம் 13 மாதிரிகள் இருக்கும். அவையாவன.

(U_1, U_4, U_7, U_{10})
$(U_4, U_7, U_{10}, U_{13})$
$(U_7, U_{10}, U_{13}, U_3)$
$(U_{10}, U_{13}; U_3, U_6)$
(U_{13}, U_3, U_6, U_9)

(U_3, U_6, U_9, U_{12})
(U_6, U_9, U_{12}, U_2)
(U_9, U_{12}, U_2, U_5)
(U_{12}, U_2, U_5, U_8)

(U_2, U_5, U_8, U_{11})
(U_5, U_8, U_{11}, U_1)
(U_8, U_{11}, U_1, U_4)
(U_{11}, U_1, U_4, U_7)



மேலேயுள்ள படம் $r = 7$ என்னும் ராண்டம் துவக்கத்திலிருந்து எடுக்கும் வ.மு.மா. ஐக் குறிக்கிறது.

பொதுவாக இம் முறையில் மொத்தம் N மாதிரிகள் இருக்கும். அதாவது முழுமைத் தொகுதியில் எத்தனை உறுப்புகள் உள்ளனவோ அத்தனை வ.மு. மாதிரிகள் இருக்கும்.

19·15. தேற்றம்

வ.மு.மா. சராசரி முழுமைத் தொகுதியின் சராசரிக்கும் பிறழ்ச்சியற்றதொரு மதிப்பீட்டியாகும்.

நிறுவல்

r என்னும் ராண்டம் துவக்கத்திலிருந்து எடுக்கப்படும் வ.மு.மா. $Y_{r1}, Y_{r2}, Y_{r3}, \ldots, Y_{rn}$ எனக் கொள்வோம், இதன் சராசரியை $\bar{x}_r$ எனக் குறிப்பிடுவோம், $\bar{x}$ என்பது ஒரு வ. மு.மா-ன் சராசரியானல், அதன் பரவல் பின்வருமாறு :

$\bar{x}$	$\bar{x}_1$	$\bar{x}_2$	$\bar{x}_3$	....	$\bar{x}_N$
P	$\frac{1}{N}$	$\frac{1}{N}$	$\frac{1}{N}$		$\frac{1}{N}$

$$\therefore \quad E(\bar{x}) = \frac{1}{N}\sum_{i=1}^{N} \bar{x}_i = \frac{1}{N}\sum_{i=1}^{N}\frac{1}{n}\sum_{m=1}^{n} Y_{im}$$

$$= \frac{1}{Nn}\sum_{i=1}^{N}\sum_{m=1}^{n} Y_{im}$$

இத்தக் கூட்டுத்தொகையில் மொத்தம் Nn உறுப்புகள் உள்ளன. இவை $(X_1, X_2, \ldots, X_N)$ என்பதாலும், ஒழுங்குமுறை (Symmetry) காரணமாக ஒவ்வொரு X_i-ம் n முறை என்பதாலும் இதைப் பின்வருமாறு எழுதலாம்.

$$E(\bar{x}) = \frac{1}{Nn}\; n \sum_{i=1}^{n} X_i = \frac{1}{N}\sum_{i=1}^{N} X_i = \bar{X} \quad \text{(நி-து.)}$$

$$\therefore \quad E^{-1}(\bar{X}) \sim \bar{x}.$$

19·16. $V(\bar{x})$, $E^{-1} V(\bar{x})$ முதலியவை நே. மு. மா. முறையிலுள்ளவை போன்றே சற்று மாறுதலுடன் இருக்கும். இவைகளை மாணவர்களின் முயற்சிக்கே விட்டுவிடுகிறேன்.

நே. மு. மா-ஐ விட வ. மு. மா. பலவிதத்திலும் அழகானது. இதிலுள்ள ஒழுங்கு (Symmetry) நம் மனதைக் கவரக்கூடியது. மற்றபடி நே. மு. மா.க்கு என்னென்ன சிறப்புகள் உண்டோ அவை யாவையும் வ. மு மா. -க்கும் உள்ளன. இக் காரணத்தால் வட்ட முறையுடை மாதிரி முறையை நாம் நே. மு. மா. முறையை விட அதிகமாகக் கையாளலாம்.

19.17. முறையுடை மாதிரியும் சாதாரண ராண்டம் மாதிரியும்

முறையுடை மாதிரிச் சராசரியின் பரவற் படி முழுமைத் தொகுதியில் உறுப்புகள் வரிசைப்படுத்தும் தன்மைக் கேற்ப மாறும் என்பதைப் பார்த்தோம். முறையுடை மாதிரியை எடுக்கும் முன்னர், முழுமைத் தொகுதி உறுப்புகளை ராண்டம் வரிசையில் அமைத்தோமானால் வரும் மாதிரி ஈ. செயா. சா. ரா. மாதிரியாகத் தான் இருக்க வேண்டுமென்பது தெளிவாகத் தெரிகிறது. இது எப்படியெனில் முழுமைத் தொகுதி உறுப்புகள் ராண்டம் வரிசை யில் அமைந்திருப்பதால் N உறுப்புகளில் எந்த n உறுப்பும் வருவதற்கு வாய்ப்பு $1 \div NP_n$. ஆகவே இது ஈ. செயா சா. ரா. மா. முறையாகத்தான் இருக்க வேண்டும்.

உதாரணமாக, சீட்டுக்கட்டை ராண்டம் வரிசையில் அமைத்து விட்டு 4 பேர்களுக்கு வரிசையாகப் போட்டோமானால் ஒவ்வொருவருக்கும் வருவது ஒரு முறையுடை மாதிரியே. இருந்தாலும் ஒவ்வொருவருக்கும் எந்த 13 சீட்டுகளும் வருவதற்குச் சம வாய்ப்பே இருக்கிறது.

இப்படி ராண்டம் வரிசையில் அமைக்கப்பட்ட முழுமைத் தொகுதியிலிருந்து n உறுப்புகள் கொண்ட மாதிரி எடுத்தால் அதன் சராசரி $\bar{x}$ முழுமைத் தொகுதிச் சராசரியைப் பிறழ்ச்சியின்றி மதிப்பிடும்.

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

$$E^{-1}\,[V(\bar{x})] \sim \frac{N-n}{Nn}\,s^2$$

இந்தக் காரணத்தால் முழுமைத் தொகுதி உறுப்புகள் ராண்டம் வரிசையில் அமைக்கப்படாவிட்டாலும் கூட $\frac{N-n}{Nn}\,s^2$ என்பதை $V(\bar{x})$-ன் மதிப்பீட்டியாக எடுத்துக்கொள்ளலாம்.

ஆனால் இது பிறழ்ச்சியற்றதாக இராது. ஆனால் முழுமைத் தொகுதி உறுப்புகள் சாதாரணமாக ஓரளவு ராண்டம் வரிசையிலேயே அமைந்திருக்கும் என எடுத்துக்கொள்வதில் தவறு அதிகம் இருக்க முடியாது. நாமே ஒரு குறிப்பிட்ட வரிசையில் அமைத்தால்தான் அவ்வரிசை ராண்டம் வரிசையாக இராது. நடைமுறையில் $\frac{N-n}{Nn}s^2$ என்பதையே $V(\bar{x})$-ன் மதிப்பீட்டியாக எடுத்துக்கொள்ளலாம்.

பயிற்சி 4

1. N உறுப்புடைய முழுமைத் தொகுதியிலிருந்து n உறுப்பு கொண்ட வ. மு மா. ஒன்று எடுக்கப்படுகிறது.

(அ) K என்பது $\left(\frac{N}{n}\right)$ என்ற எண்ணின் மிகவும் அருகிலுள்ள முழு எண் என எடுத்துக் கொண்டால் எந்த நிபந்தனையின் பேரில் ராண்டம் துவக்கத்தை மறுமுறை சந்திக்காமலேயே அல்லது தாண்டிப் போகாமலேயே n உறுப்புகளை எடுக்க முடியும்?

(ஆ) தகுந்த காரணங்களுடன், k என்பதை, $\frac{N}{n}$ என்ற எண்ணின் முழு எண் பாகமாகவோ அல்லது $\frac{N+n}{n}$ என்ற எண்ணின் முழு எண் பாகமாகவோ தேர்ந்தெடுப்பாய் என்பதைக் கூறு.

2. பின்வரும் 40 எண்கள் ஒரு முழுமைத் தொகுதியாகும்.

கணம் I :	0	1	2	1	4	5	7	7	9	8
கணம் II :	10	11	13	12	12	15	14	17	20	23
கணம் III :	22	25	29	30	32	35	33	38	40	41
கணம் IV :	39	43	46	50	53	52	57	59	63	62.

(அ) முழுமைத் தொகுதியின் சராசரியை மதிப்பிடுவதில் நே. மு. மா. முறையின் ஈ, செயா. சா. ரா. மா. முறையைச் சார்ந்த திறன் விகிதத்தைக் (relative efficiency) கண்டுபிடி. மாதிரிப் பருமனை 4 எனக் கொள்க.

(ஆ) இரண்டாவது, நான்காவது கணங்களில் உள்ள எண்களை மட்டும் எதிர் வரிசையில் (revesre way) எழுதுவதால் முன் கூறிய திறன் விகிதத்தில் என்ன மாற்றம் ஏற்படுகிறது? இம் மாறுதலை எப்படி விளக்குவாய் ?

விடை

(அ) 3·32 (ஆ) 143·9.

(3) ஒரு பட்டியலிலிருந்து 10% முறையுடை மாதிரியொன்று எடுக்கப்படுகிறது. அப் பட்டியலில் குடும்பங்கள் மதங்களின் வாரியாக வரிசைப்படுத்தப்பட்டன. அக் குடும்பங்களுள் ஆட்கள் வயது வாரியாக வரிசைப்படுத்தப்பட்டனர். பின் கண்ட பண்புகளுக்கு இம் மாதிரி ஈ. செயா. சா. ரா. மா. ஐ விடத் திறனதிகம் கொண்டதாக இருக்குமா?

(அ) முழுமைத் தொகுதியில் வெவ்வேறு மதத்தவரின் எண்ணிக்கை,

(ஆ) ஆண்களின் சராசரி வயது.

(இ) பள்ளிக்குச் செல்லும் குழந்தைகளின் சதவீதம்.

4. முறையுடை மாதிரியின் பரவற் படியை வரிசை ஒட்டுறவுக் கெழு (Serial correlation coefficient) ρ_{α} மூலமாக வெளிப்படுத்து. முழுமைத் தொகுதிப் பருமன் N எனவும் மாதிரிப் பருமன் n எனவும் கொள்க.

$$\rho_{\alpha} = \frac{1}{kn(n-1)\sigma^2} \sum_{i=1}^{k} \sum_{i'=1}^{n-\alpha} (Y_{ii'} - \bar{Y})(Y_{i(i'+\alpha)} - \bar{Y})$$

$\alpha = 1, 2, \ldots (n-1)$, σ^2 = முழுமைத் தொகுதிப் பரவற் படி. [Madow, W.G. and Madow, L.H., Ann. Math. Stat., 15, (1944)]

விடை

$$V(\bar{y}) = \frac{\sigma^2}{n} \left[1 + \frac{2}{n} \sum_{\alpha=1}^{n-1} (n-\alpha)\rho_{\alpha} \right]$$

5. ஓர் இரு பரிமாண முழுமைத் தொகுதி (Two-dimensional population) Y_{ij} என்பது பின்வருமாறு :

$$Y_{ij} = i + j \qquad i = 1, 2, \dots N$$
$$j = 1, 2, \dots M$$

Y_{ij}-க்கள் ஒரு செவ்வக அணியாக அமைக்கப்பட்டிருக்கின்றன (in Matix Form). அதாவது Y_{ij} என்னும் உறுப்பு இவ்வணியின் i ஆவது நிரை j ஆவது நிரலில் (i^{th} row, j^{th} column) உள்ளது.

இம்முழுமைத் தொகுதியிலிருந்து nm பருமனுடைய இரு பரிமாண முறையுடை மாதிரி பின்வருமாறு எடுக்கப்பட்டது.

$k = \frac{N}{n}$, $k' = \frac{N}{m}$. k, k' முழு எண்கள்.

i' என்னும் ராண்டம் துவக்கம் 1-லிருந்து k வரையுள்ள எண்களிலிருந்து எடுக்கப்பட்டது.

,, ,, ,, 1-லிருந்து k' ,,

பின்வரும் அணி உறுப்புகள் மாதிரியில் இருக்கின்றன.

$(i'+\alpha k, j'+\beta k')$; $\alpha=0,1,2,\dots(n-1)$, $\beta=0,1,2,\dots(m-1)$

(அ) முழுமைத் தொகுதியின் சராசரிக்கு மாதிரிச் சராசரி பிறழ்ச்சியற்ற மதிப்பீட்டியென நிரூபி.

(ஆ) மாதிரிச் சராசரியின் பரவற் படியைக் கண்டுபிடி. இதை NM உறுப்புகளிலிருந்து nm பருமனுடைய ஈ. செயா. சா. ரா. மா.-ன் பரவற் படியுடன் ஒப்பிடு.

விடை

$(i', \alpha k, j'+\beta k')$ என்னும் அணியுறுப்பு $(i'+j'+\alpha k+\beta k')$ ஆகும். (i', j') என்னும் ராண்டம் துவக்கத்துடன் எடுக்கப்படும் மாதிரியின் மொத்தம்

$$S_{i',j'} = \sum_{\beta=0}^{m-1} \sum_{\alpha=0}^{n-1} (i' + j' + \alpha k + \beta k') = nm\,(i' + j')$$
$$+ mk \sum_{\alpha=0}^{n-1} \alpha + nk' \sum_{\beta=0}^{m-1} \beta$$

$$= nm\,(i' + j') + \frac{mn(n-1)}{2}\,k + \frac{nm(m-1)}{2}\,k'$$

ஆகவே இம்மாதிரியின் சராசரி $\frac{S_{i',j'}}{nm}$

$$= \left[i'+j'+\frac{(n-1)k}{2} + \frac{(m-1)k'}{2} \right]$$

$$\therefore\ E\left(\frac{S_{i'j'}}{nm}\right) = \sum_{j'=1}^{k} \sum_{i'=1}^{k} \frac{1}{kk'}$$

$$\left[i'+j'+\frac{(n-1)k}{2} + \frac{(m-1)k'}{2} \right]$$

$$= \frac{1}{kk'}\left[kk' \left(\frac{\overline{n-1}\,k}{2} + \frac{\overline{m-1}\,k'}{2}\right)\right.$$

$$\left. + k' \sum_{i'=1}^{k} i' + k \sum_{j'=1}^{k} j' \right]$$

$$= \frac{(n-1)k}{2} + \frac{(m-1)k'}{2} + \frac{k+1}{2} + \frac{k'+1}{2} = \frac{nk}{2} + \frac{mk'}{2} + 1$$

ஆனல் $k = \frac{N}{n},\ k' = \frac{M}{m}$

$$\therefore\ E\left(\frac{S_{i',j'}}{nm}\right) = \frac{N}{2} + \frac{M}{2} + 1 = \frac{M+N+2}{2}$$

முழுமைத் தொகுதியின் சராசரி $\frac{1}{NM} \sum_{j=1}^{M} \sum_{i=1}^{N} (i + j)$

$$= \frac{1}{NM}\left[M \sum_{i=1}^{N} i + N \sum_{j=1}^{M} j \right]$$

$$= \frac{1}{NM}\left[\frac{MN(N+1)}{2} + \frac{NM(M+1)}{2}\right]$$

$$= \frac{N+1}{2} + \frac{M+1}{2} = \frac{N+M+2}{2}$$

இதிலிருந்து முழுமைத் தொகுதியின் சராசரிக்கு மாதிரிச் சராசரி பிறழ்ச்சியற்ற மதிப்பீட்டி என்பது தெரிகிறது.

மாதிரிச் சராசரியின் பரவற் படி $V = V\left[\frac{S_{i'j'}}{nm}\right]$

$$= V\left[i' + j' + \frac{(n-1)k}{2} + \frac{(m-1)k'}{2}\right]$$

$$= V[i' + j']\left[\because \frac{(n-1)k}{2} + \frac{(m-1)k'}{2}\right.$$ என்பது மாறிலி.]

$= V(i') + V(j')$ [$\because$ i', j' சார்பற்றவை.]

$V(i')$ ஐ எடுத்துக்கொள்வோம்.

i'-ன் பரவல்

i'	1	2	3	4	....	k
P	$\frac{1}{k}$	$\frac{1}{k}$	$\frac{1}{k}$	$\frac{1}{k}$	....	$\frac{1}{k}$

$\therefore$ $V(i') = E(i'^2) - [E(i')]^2$

$$= \frac{1}{k}\sum_{r=1}^{k} r^2 - \left[\frac{1}{k}\sum_{r=1}^{k} r\right]^2$$

$$= \frac{k(k+1)(2k+1)}{6k} - \left(\frac{k(k+1)}{2k}\right)^2$$

$$= \frac{k^2-1}{12}$$

இதைப்போலவே $V(j') = \frac{k'^2-1}{12}$

$$\therefore \quad V(i'+j') = \frac{k^2-1}{12} + \frac{k'^2-1}{12} = \frac{k^2+k'^2-2}{12}$$

இந்த இரு பரிமாண முழுமைத் தொகுதியிலிருந்து ஈ. செயா. சா. ரா. மா. முறையில் nm உறுப்புகள் எடுத்தால் மாதிரிச் சராசரியின் பரவற் படி

$$V_2 = \frac{NM - mn}{NM - 1} \quad \frac{\sigma^2}{nm}, \text{ இதில் } \sigma^2 = V(i+j)$$

$i = 1, 2, \ldots, N, j = 1, 2, \ldots M$ என்பதால்

$$\sigma^2 = V(i+j) = V(i) + V(j) = \frac{N^2-1}{12} + \frac{M^2-1}{12}$$

$$= \frac{N^2+M^2-2}{12}$$

$$\therefore \quad V_2 = \frac{(NM-nm)}{(NM-1)} \cdot \frac{(N^2+M^2-2)}{12\,nm}$$

6. N உறுப்புகள் உள்ள முழுமைத் தொகுதி $Y_1, Y_2, \ldots Y_N$. இதை ஒரு பாகமாகக் கொண்டுள்ள பெரும் முழுமைத் தொகுதி (Super population) ஒன்றிலிருந்து எடுக்கப்பட்டிருக்கிறது. இப்பெரும் முழுமைத் தொகுதியில்

(அ) $E(Y_i) = \mu$; $V(Y_i) = \sigma_i^2$;

$$\text{cov}(Y_i, Y_{j'}) = 0_{i'}; j' \neq i'$$

(ஆ) $E(Y_i) = \alpha + \beta i$; $V(Y_i) = \sigma^2$;

$$\text{cov}(Y_i, Y_{i'}) = 0, i \neq i'$$

மேற்கூறப்பட்ட (அ), (ஆ) என்ற இரண்டிலிருந்தும் தனித் தனியாக n பருமன் கொண்ட மாதிரி (i) நே. மு. மா. முறையில்

(ii) ஈ. செயா. சா. ரா. மா. முறையில் எடுக்கப்பட்டிருந்தால் அம்மாதிரியின் சராசரியின் பரவற் படியின் எதிர்பார்க்கும் மதிப்பு என்ன என்பதைக் கண்டுபிடி.

[Cochran, W.G., Sampling Techniques, ch. 8 (1963)]

விடை

(அ) (i) $\frac{k-1}{k} \frac{\sigma^2}{n}$ (ii) $\frac{k-1}{k} \frac{\sigma^2}{n}$

$$\sigma^2 = \frac{1}{N} \sum_{i=1}^{N} \sigma_i^2$$

(ஆ) (i) $\frac{k-1}{k} \frac{\sigma^2}{n} + \beta^2 \frac{k^2-1}{12}$

(ii) $\frac{k-1}{k} \frac{\sigma^2}{n} + \beta^2 \frac{k^2-1}{12} + \frac{\beta^2}{12} k(k-1)(n-1)$

7. $Y_1, Y_2, \ldots Y_N$ என்னும் முழுமைத் தொகுதி உறுப்புகள் பின்கண்ட தற்றொடர்பாக்கமுள்ள (Autocorrelated) பெரும் முழுமைத் தொகுதி (Super population) ஒன்றிலிருந்து எடுக்கப்பட்டிருக்கிறது.

$E(Y_i) = \mu$, $V(Y_i) = \sigma^2$, $\text{cov}(Y_i, Y_{i+u}) = \rho_u$

எல்லா i-க்கும், $u \geq 1$

$u < v$ ஆக இருந்தால் $\rho_u \geq \rho_v \geq 0$

$\delta_i^2 = p_{i+1} - p_{i-1} + 2p_i \geq 0, \quad i = 2, 3 \ldots\ldots (N-2)$

மேற்கூறப்பட்ட நிபந்தனைகளின் பேரில் n உறுப்புகள் கொண்ட நே. மு. மா.-ன் சராசரியின் பரவற் படியின் எதிர்பார்க்கும் மதிப்பு, n உறுப்புகள் கொண்ட ஈ. செயா. சா, ரா. மா. -ன் சராசரியின் பரவற் படியின் எதிர்பார்க்கும் மதிப்பை விடக் குறைவானது என நிரூபி. $\left[\frac{N}{n}\right.$ ஒரு முழு எண் எனக் கொள்க.$\left.\right]$

[Cochran, W. G., Ann. Math. Stat., 17, (1946)]

8. மேற்கூறிய கணக்கில் முதலிரண்டு நிபந்தனைகளை

$E(Y_i) = \mu_i$, $V(Y_i) = \sigma_i^2$ என்று தளர்த்தினாலும் முடிவு மாறாது என்பதை நிரூபி.

[Quenouille, M.H., Ann Math. Stat., 20, (1949)]

20. சமவாய்ப்பிலா மாதிரி முறை

(VARYING PROBABILITY SAMPLING)

20·1. இதுவரை நாம் பார்த்த மாதிரி முறைகளில் எல்லா உறுப்புகளுக்கும் மாதிரியில் வருவதற்குச் சம வாய்ப்பு அளிக்கப்பட்டிருந்தது. முழுமைத் தொகுதியிலுள்ள உறுப்புகளின் தன்மை தெரியாதபோது எல்லா உறுப்புகளுக்கும் சம வாய்ப்பு அளிக்க வேண்டியது நியாயமும் கூட. ஆனுல் சில சமயங்களில் சில உறுப்புகளுக்கு அதிக வாய்ப்பும் சில உறுப்புகளுக்குக் குறைவான வாய்ப்பும் கொடுக்க வேண்டிய அவசியம் ஏற்படுகிறது. சம வாய்ப்பு கொடுக்கப்படாமல் எடுக்கப்பட்ட மாதிரிகளிலிருந்து மிகச் சிறந்த மதிப்பீடுகளைப் பெறமுடியும் என்பதைப் பிறகு பார்ப்போம்.

முழுமைத் தொகுதிச் சராசரியை மதிப்பிடும்பொழுது பெரிய உறுப்புகளுக்கு அதிக வாய்ப்பும், சிறிய உறுப்புகளுக்குக் குறைவான வாய்ப்பும் கொடுப்பது அவசியம். ஏனெனில் பெரிய உறுப்பு சராசரியை அதிகம் பாதிக்கும், சிறிய உறுப்பு சராசரியை அதிகம் பாதிக்காது. ஒவ்வோர் உறுப்புக்கும் எவ்வளவு வாய்ப்பளிக்கப்படவேண்டும் என்பதைப் பார்ப்போம். முழுமைத் தொகுதியிலுள்ள உறுப்புகளின் பருமன் (size) முன் கூட்டியே தெரிந்திருந்தால் அப்பருமன்களின் விகிதங்களிலேயே நிகழ்திறன்களையும் அமைத்துக் கொள்ளலாம். ஆனுல் உறுப்புகளின் பருமன் முன்கூட்டியே தெரிந்திருந்தால் அதிலிருந்து மாதிரி எடுக்கவேண்டிய அவசியமே இராது. முழுமைத் தொகுதி உறுப்புகளின் பருமன்கள் தெரியாது என்று எண்ணித்தான் நாம் அவைகளுக்கு நிகழ்திறன்களையும் அளிக்கவேண்டும்.

[கவனிக்கவும் : U_i என்னும் உறுப்பைச் சார்ந்த X_i என்னும் பண்பைப் பற்றி நாம் ஆராயும்போது X_i என்பதைத்தான் நாம் U_i-ன் பருமன் என்று எடுத்துக்கொள்கிருேம்.]

ஆகவே $U_1, U_2, \ldots U_N$ என்னும் உறுப்புகளில் $X_1, X_2, \ldots X_N$ என்னும் பண்புகளைப்பற்றி நாம் கணக்கிடும்போது இந்த X_i-களுடன் சார்ந்த S_i என்னும் மற்றொரு பண்பினை நாம் அளக்க வேண்டிய அவசியமேற்படுகிறது. $S_1, S_2, \ldots S_N$ என்பவை முற்றிலும் தெரியும் என எடுத்துக்கொண்டால் $U_1, U_2, \ldots U_N$ என்னும் உறுப்புகளுக்கு முறையே $\frac{S_1}{S}, \frac{S_2}{S}, \ldots, \frac{S_N}{S}$

$\left[S = \sum_{i=1}^{N} S_i \right]$ என்னும் நிகழ்திறன்களை மாதிரியில் வருவதற்கு அளிக்கலாம். இப்படி நிகழ்திறன்களை அளிப்பது பருமனுக்கு விகித சமமுள்ள (Proportional) நிகழ்திறன் அளிப்பது என்று கூறப்படும். சுருக்கமாக இதை ப.வி.நி. என்று எழுதுவோம். சரி வாய்ப்பில்லாத மாதிரி முறையை ச. இலா. மா. முறை என்று எழுதுவோம். $X_1, X_2, \ldots X_N$ என்னும் உறுப்புகளுக்கு $P_1, P_2, \ldots P_N$ என்னும் நிகழ்திறன்களை மாதிரியில் வருவதற்கு அளிப்பதாகக் கொள்வோம். மாதிரியில் $x_1, x_2, \ldots x_n$ என்னும் உறுப்புகளைச் சார்ந்த நிகழ்திறன்கள் $p_1, p_2, \ldots p_n$ எனவும் கொள்வோம்.

20·2. தேற்றம்

ச. இலா. மா. முறையில் ஈடு செய்யப்படாமல் n உறுப்புகள் எடுக்கப்பட்டால் X என்னும் முழுமைத் தொகுதியின் மொத்தத்திற்கு $\frac{1}{n} \sum_{i=1}^{n} \frac{x_i}{p_i}$ என்பது பிறழ்ச்சியற்ற மதிப்பீட்டியாகும்.

இதன் பரவற் படியானது $\frac{1}{n} \left[\sum \frac{X_i^2}{p_i} - X^2 \right]$ ஆகும்.

நிறுவல்

x_i-ன் பரவல் :

x_i	$X_1 \quad X_2 \quad X_3 \ldots X_N$
P	$P_1 \quad P_2 \quad P_3 \ldots P_N$

ஆகையால் $\left(\frac{x_i}{p_i}\right)$-ன் பரவல்

$\frac{x_i}{p_i}$	$\frac{X_1}{P_1}$	$\frac{X_2}{P_2}$	$\frac{X_3}{P_3}$	......	$\frac{X_N}{P_N}$
P	P_1	P_2	P_3	......	P_N

மேலும் ஈடுசெய்யப்பட்ட முறையில் உறுப்புகள் எடுக்கப்படுவதால் $\frac{x_1}{p_1}, \frac{x_2}{p_2}, \dots, \frac{x_n}{p_n}$ முதலியவை சார்பற்றவை (Independent).

$$\therefore \quad E\left(\frac{x_i}{P_i}\right) = \sum_{i=1}^{N} \frac{X_i}{P_i} P_i = \sum_{i=1}^{N} X_i = X.$$

$$\therefore \quad E\left[\frac{1}{n}\sum_{i=1}^{n}\frac{x_i}{p_i}\right] = \frac{1}{n}\sum_{i=1}^{n} X = \frac{1}{n}\cdot nX = X.$$

$$\therefore \quad E^{-1}(X) \sim \frac{1}{n}\sum_{i=1}^{n}\frac{x_i}{p_i}. \qquad \text{(நி-து)}$$

$$V\left(\frac{x_i}{p_i}\right) = E\left(\frac{x_i^2}{p_i^2}\right) - \left[E\left(\frac{x_i}{p_i}\right)\right]^2$$

$$= \sum_{i=1}^{N}\frac{X_i^2}{P_i^2}\cdot P_i - X^2 = \sum_{i=1}^{N}\frac{X_i^2}{P_i} - X^2$$

$$\therefore \quad V\left[\frac{1}{n}\sum_{i=1}^{n}\frac{x_i}{p_i}\right] = \frac{1}{n^2}\sum_{i=1}^{n}\left(\frac{x_i}{p_i}\right)$$

$\left[\frac{x_1}{p_1}, \dots, \frac{x_n}{p_n}\right.$ முதலியவை சார்பற்றவை $\left.\right]$

$$= \frac{1}{n^2}\sum_{i=1}^{n}\left[\sum_{i=1}^{N}\frac{X_i^2}{P_i} - X^2\right]$$

$$= \frac{1}{n^2}\cdot n\left[\sum_{i=1}^{N}\frac{X_i^2}{P_i} - X^2\right]$$

$$= \frac{1}{n}\left[\sum_{i=1}^{N}\frac{X_i^2}{P_i} - X^2\right] \qquad \text{(நி-து)}$$

20·3. தேற்றம்

ஈடு செய்யப்பட்ட சம வாய்ப்பிலா மாதிரி முறையில்

$V\left[\frac{1}{n}\sum_{i=1}^{n}\frac{x_i}{p_i}\right]$-ன் பிறழ்ச்சியற்ற மதிப்பீடொன்று

$$\frac{1}{n(n-1)}\left[\sum_{i=1}^{n}\frac{x_i^2}{p_i^2} - n\left(\frac{1}{n}\sum_{i=1}^{n}\frac{x_i}{p_i}\right)^2\right] \text{ ஆகும்.}$$

$$V\left[\frac{1}{n}\sum_{i=1}^{n}\frac{x_i}{p_i}\right] = \frac{1}{n}\left[\sum_{i=1}^{N}\frac{X_i^2}{P_i} - X^2\right] \text{ என்று}$$

முன்பே பார்த்தோம்.

$$E\left(\frac{x_i^2}{p_i^2}\right) = \sum_{i=1}^{N}\frac{X_i^2}{P_i^2}P_i = \sum_{i=1}^{N}\frac{X_i^2}{P_i}$$

$$\therefore\quad E\left[\frac{1}{n}\sum_{i=1}^{n}\frac{x_i^2}{p_i^2}\right] = \frac{1}{n}\cdot n\sum_{i=1}^{N}\frac{X_i^2}{P_i} = \sum_{i=1}^{N}\frac{X_i^2}{P_i}$$

$$\therefore \quad E^{-1}\left[\frac{1}{n}\sum_{i=1}^{N}\frac{X_i^2}{P_i}\right] \sim \frac{1}{n^3}\sum_{i=1}^{n}\frac{x_i^2}{p_i^2} \qquad \ldots\ldots \; I$$

$i \neq j$ என்றால்

$$E\left(\frac{x_i\, x_j}{p_i\, p_j}\right) = \sum_{k=1}^{N}\sum_{l=1}^{N}\frac{X_k\, X_l}{P_k\, P_l}P_k\, P_l \quad [x_i,\, x_j \text{ சார்பற்றவை}]$$

$$= \sum_{k=1}^{N}\sum_{l=1}^{N} X_k\, X_l = \left[\sum_{i=1}^{N} X_i\right]^2 = X^2$$

$$\therefore \quad E\left[\sum_{i \neq j}^{n}\sum^{n}\frac{x_i\, x_j}{p_i\, p_j}\right] = n\,(n-1)\, X^2$$

$$\therefore \quad E^{-1}\left(\frac{X^2}{n}\right) \sim \frac{1}{n^2\,(n-1)}\sum_{i \neq j}^{n}\sum^{n}\frac{x_i\, x_j}{P_i\, P_j} \qquad \ldots\ldots \; II$$

$$\therefore \quad E^{-1}\left\{V\left[\frac{1}{n}\sum_{i=1}^{n}\frac{x_i}{P_i}\right]\right\}$$

$$\sim E^{-1}\left[\frac{1}{n}\sum_{i=1}^{n}\frac{X_i^2}{P_i}\right] - E^{-1}\left[\frac{1}{n}X^2\right]$$

$$\sim \frac{1}{n^2}\sum_{i=1}^{n}\frac{x_i^2}{p_i^2} - \frac{1}{n^2\,(n-1)}\sum_{i \neq j}\sum\frac{x_i\, x_j}{p_i\, p_j}$$

$$\sim \frac{1}{n^2}\sum_{i=1}^{n}\frac{x_i^2}{p_i^2} - \frac{1}{n^2(n-1)}\left[\left\{\sum_{i=1}^{n}\frac{x_i}{p_i}\right\}^2 - \sum_{i=1}^{n}\frac{x_i^2}{p_i^2}\right]$$

$$\sim \left[\frac{1}{n^2} + \frac{1}{n^2(n-1)}\right]\sum_{i=1}^{n}\frac{x_i^2}{p_i^2} - \frac{1}{n^2(n-1)}\left\{\sum_{i=1}^{n}\frac{x_i}{p_i}\right\}^2$$

$$\sim \frac{1}{n(n-1)}\left[\sum_{i=1}^{n}\frac{x_i^2}{p_i^2} - n\left\{\frac{1}{n}\sum_{i=1}^{n}\frac{x_i}{p_i}\right\}^2\right] \text{ (நி-து.)}$$

20·4. ஈடு செய்யப்பட்ட சம வாய்ப்பிலா மாதிரி முறை ஈ. செ. சா. ரா. மா. முறையை விட எப்பொழுது திறன் அதிகம் வாய்ந்ததாயிருக்கும் ?

ஈடு செய்யப்பட்ட சம வாய்ப்பிலா மாதிரி முறையில் முழுமைத் தொகுதி மொத்தத்தை மதிப்பிடும் மதிப்பீட்டியின் பரவற் படி

$$V_1 = \frac{1}{n}\left[\sum_{i=1}^{N}\frac{X_i^2}{P_i^2} - X^2\right].$$ ஈ. செ. சா. ரா. மா. முறையில் முழுமைத் தொகுதி மொத்தத்தை மதிப்பிடும் மதிப்பீட்டியின் பரவற் படி $V_2 = N^2\frac{\sigma^2}{n} = \frac{N^2}{n}\left[\frac{1}{N}\sum X_i^2 - \bar{X}^2\right]$

$$= \frac{N}{n}\sum_{i=1}^{N}X_i^2 - \frac{X^2}{n}.$$

$$\therefore \quad V_2 - V_1 = \frac{N}{n}\sum_{i=1}^{N} X_i^2 - \frac{X^2}{n} - \frac{1}{n}\sum \frac{X_i^2}{P_i} + \frac{X^2}{n}$$

$$= \frac{1}{n}\sum_{i=1}^{N}\left(N - \frac{1}{P_i}\right) X_i^2$$

ஆகவே $\sum_{i=1}^{N}\left(N - \frac{1}{P_i}\right) X_i^2 > 0$ ஆக இருக்கும்போது ஈடு செய்யப்பட்ட சம வாய்ப்பிலா மாதிரி முறை திறன் அதிகம் வாய்ந்ததாயிருக்கும்.

கிளைத்தேற்றம் 20·5.

P_i-கள் X_i-களின் விகிதத்திலேயே இருந்தால்; அதாவது P_i-களும் X_i-களும் விகித சமமாக (Proportional) இருந்தால் ஈடு செய்யப்பட்ட சம வாய்ப்பிலா மாதிரி முறை ஈ. செ. சா. ரா. மா. முறையை விடத் திறனதிகமுள்ளதாக இருக்கும்.

நிறுவல்

$n(V_2 - V_1) = \sum_{i=1}^{N}\left(N - \frac{1}{P_i}\right) X_i^2$ என்று முன்பே பார்த்தோம்.

$\frac{P_1}{X_1} = \frac{P_2}{X_2} = \frac{P_n}{X_n} = k$ எனக்கொண்டால்,

$P_1 = KX_1 \,;\, P_2 = kX_2, \ldots, P_n = KX_n$

$\sum_{i=1}^{N} P_i = 1$ என்பதால் $k \sum_{i=1}^{N} X_i = 1 \quad \therefore\ k = \frac{1}{\sum_{i=1}^{N} X_i} = \frac{1}{X}$

$$\therefore \quad P_1 = \frac{X_1}{X},\ P_2 = \frac{X_2}{X},\ \ldots\ldots,\ P_n = \frac{X_n}{X}$$

$$\therefore \quad n(V_2 - V_1) = \sum_{i=1}^{N} \left(N - \frac{X}{X_i} \right) X_i^2$$

$$= N \sum_{i=1}^{N} X_i^2 - X \sum_{i=1}^{N} X_i$$

$$= N \sum_{i=1}^{N} X_i^2 - X^2$$

$$= N^2 \left[\frac{1}{N} \sum X_i^2 - \bar{X}^2 \right]$$

$$= N^2 \sigma^2 \geq 0 \qquad [\text{நி-து.}]$$

கவனிக்கவும்

P_i-களும், X_i-களும் விகித சமமாக இருந்தால் ஈடு செய்யப் பட்ட சம வாய்ப்பிலா மாதிரி முறை மற்ற எல்லா மாதிரி முறைகளை விடத்திறனதிகம் வாய்ந்ததாயிருக்கும். ஏனென்றால் $n(V_2 - V_1) = N^2 \sigma^2 = n V_2$.

$\therefore \quad V_1 = 0$ அதாவது இம் முறையில் மதிப்பீட்டியின் பரவற் படி 0 ஆகும். அதாவது இம்மதிப்பீட்டி X ஐப் பிழையின்றி மதிப்பிடும் எந்த முறையிலும் பிழை சிறிதாவது இருக்குமென்று கொண்டால் மேற்கூறப்பட்டது விளங்கும்.

20·6. கிளைத்தேற்றம்

ஈடு செய்யப்பட்ட மாதிரியை ஈ. செ. சா. ரா. மா. -க்குப் பதிலாகக் கையாள்வதால் ஏற்படும் இலாபம்,

$$= V_2 - V_1 = \frac{1}{n} \sum_{i=1}^{N} \left(N - \frac{1}{P_i} \right) X_i^2.$$

இந்த இலாபத்தைப் பிறழ்ச்சியின்றி மதிப்பிடமுடியும்.

$$E\left(\frac{x_i^2}{P_i}\right) = \sum_{i=1}^{N} \frac{X_i^2}{P_i} P_i = \sum_{i=1}^{N} X_i^2$$

$$\therefore \quad E\left[\frac{1}{n}\sum_{i=1}^{n} \frac{x_i^2}{p_i}\right] = \frac{1}{n} \cdot n \sum_{i=1}^{N} X_i^2 = \sum_{i=1}^{N} X_i^2$$

$$\therefore \quad E^{-1}\left[N \sum_{i=1}^{N} X_i^2 \right] \sim \frac{N}{n} \sum_{i=1}^{N} \frac{x_i^2}{p_i}$$

$$E\left(\frac{x_i^2}{p_i^2}\right) = \sum_{i=1}^{N} \frac{X_i^2}{P_i^2} P_i = \sum_{i=1}^{N} \frac{X_i^2}{P_i}$$

$$\therefore \quad E\left[\frac{1}{n}\sum_{i=1}^{N} \frac{x_i^2}{p_i^2}\right] = \frac{1}{n} \cdot n \sum_{i=1}^{N} \frac{X_i^2}{P_i} = \sum_{i=1}^{N} \frac{X_i^2}{P_i}$$

$$\therefore \quad E^{-1}\left[\sum_{i=1}^{N} \frac{X_i^2}{P_i}\right] \sim \frac{1}{n} \sum_{i=1}^{n} \frac{x_i^2}{p_i^2}$$

$$\therefore \quad E^{-1} (V_2 - V_1)$$

$$\sim \frac{1}{n} E^{-1}\left[N \sum_{i=1}^{N} X_i^2 \right] - \frac{1}{n} E^{-1}\left[\sum_{i=1}^{N} \frac{X_i^2}{P_i}\right]$$

$$\sim \frac{N}{n^2} \sum_{i=1}^{n} \frac{x_i^2}{p_i^2} - \frac{1}{n^2} \sum \frac{x_i^2}{p_i^2}$$

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

இம்மதிப்பீட்டின் மூலம் இலாபத்தின் அளவு நமக்குத் தெரியும். இதைக்கொண்டு, ஓரளவு, மேலே குறிப்பிட்ட முறைகளில் எது திறனதிகம் வாய்ந்தது என்பதை அனுபவ பூர்வமாகத் தெரிந்து கொள்ளலாம்.

20·7. ஈடு செய்யப்படாத சம வாய்ப்பிலா மாதிரி முறை

ஈடு செய்யப்பட்ட சம வாய்ப்பிலா மாதிரி முறையில் $X_1, X_2, \ldots, X_N$ என்னும் முழுமைத் தொகுதி உறுப்புகளுக்கு முறையே $P_1, P_2, \ldots, P_N$ என்னும் நிகழ்திறன்களை அளித்தோம். ஒரு முறை எடுக்கப்பட்ட உறுப்பு மறுபடியும் முழுமைத் தொகுதியில் திருப்பி வைக்கப்பட்டுவிடுவதால் மறுபடி ஓர்உறுப்பு எடுக்கப்படும்போது இதே நிகழ்திறன்கள் மறுபடியும் செயல்படும். ஆனல் எடுக்கப்பட்ட உறுப்பு மறுபடியும் முழுமைத் தொகுதியில் திருப்பி வைக்கப்படாவிட்டால் நிலைமை மாறிவிடுகிறது. பழைய நிகழ்திறன்களை மாற்றவேண்டிய நிலைமை ஏற்படும்.

உதாரணமாக, முழுமைத் தொகுதியில் 4 உறுப்புகள் இருப்பதாகக் கொள்வோம். இவைகளை $\frac{1}{2}, \frac{1}{4}, \frac{1}{6}, \frac{1}{12}$ என்னும் நிகழ்திறன்களுடன் எடுப்பதாக வைத்துக்கொள்வோம். முதலில் இரண்டாவது உறுப்பு எடுக்கப்பட்டுவிடுவதாக வைத்துக் கொள்வோம். இரண்டாவது தேர்வுக்கு முதல், மூன்றும், நான்காம் உறுப்புகளுக்கு முறையே

$$\frac{\frac{1}{2}}{1-\frac{1}{4}}, \frac{\frac{1}{6}}{1-\frac{1}{4}}, \frac{\frac{1}{12}}{1-\frac{1}{4}},$$

அதாவது $\frac{2}{3}, \frac{2}{9}, \frac{1}{9}$ என்பவை நிகழதிறன்களாக இருக்கும். இரண்டாவது தேர்வில் மூன்றும் உறுப்பு தேர்ந்தெடுக்கப்பட்டு விடுவதாக வைத்துக்கொண்டால் முதல் உறுப்புக்கும், நான்காம்

உறுப்புக்கும் முறையே

$$\frac{\frac{2}{3}}{1-\frac{2}{9}}, \frac{\frac{1}{9}}{1-\frac{2}{9}}$$

அதாவது $\frac{6}{7}, \frac{1}{7}$ என்பவை நிகழ்திறன்களாக இருக்கும். இதிலிருந்து இம்முறையிலுள்ள சிக்கல் தெளிவாக விளங்கும்.

பொதுவாக, முதல் தேர்வில் X_{i_1} என்னும் உறுப்பு தேர்ந்தெடுக்கப்பட்டால் அடுத்த தேர்வுக்கு X_j $[j \neq i]$ என்னும் உறுப்பு வருவதற்கு நிகழ்திறன் $\frac{P_j}{1-P_{i_1}}$ ஆகும். இரண்டாவது தேர்வில் X_{i_2} என்னும் உறுப்பு தேர்ந்தெடுக்கப்பட்டுவிட்டால் X_j $[j \neq i_1\ j \neq i_2]$ என்னும் உறுப்பு மூன்றுவது தேர்வில் வருவதற்கு நிகழ்திறன் $\frac{P_j}{1-P_{i_1}-P_{i_2}}$ ஆகும். பொதுவாக முதல் r தேர்வுகளில் $X_{i_1}; X_{i_2}, \dots X_{i_r}$ என்னும் உறுப்புகள் தேர்ந்தெடுக்கப்பட்டு விட்டால் $(r+1)$ ஆவது தேர்வில் X_j என்னும் உறுப்பு வருவதற்கு நிகழ்திறன்

$$\frac{P_j}{1_i - P_{i_1} - P_{i_2} \dots P_{ir}}$$

$[j \neq i\alpha, \alpha = 1, 2, \dots\dots r]$,

20·8. முதல் தேர்வுக்கு $X_1, X_2, \dots X_N$ என்னும் முழுமைத் தொகுதி உறுப்புகளுக்கு $P_1, P_2, \dots P_N$ என்பவை நிகழ்திறன்களாக இருக்கட்டும். n உறுப்பு கொண்ட ஈடு செய்யப்படாத மாதிரி முறையில் X_i என்னும் உறுப்பு வருவதற்கு நிகழ்திறன் என்ன என்பதை இங்குப் பார்ப்போம்.

X_i முதல் தேர்விலேயே வருவதற்கு நிகழ்திறன் P_i

X_i இரண்டாம் தேர்வில் வருவதற்கு ,,

$$\sum_{j \neq i}^{n} P_j \frac{P_j}{1-P}$$

X_i மூன்றும் ,, ,, ,,

$$\sum_{j_1 \neq j_i}^{n} \sum_{j_2 \neq i}^{n} P_{j1} \frac{P_{j2}}{1-P_{j1}} \quad \frac{P}{1-P_{i1}-P_{i2}}$$

∴ n உறுப்புகள் கொண்ட மாதிரியின் X_i என்னும் உறுப்பு வருவதற்கான நிகழ்திறன்.

$$\pi_i = P_i \left[1 + \sum_{j \neq i}^{n} \frac{P_i}{1-P_j} + \sum_{j_1 \neq j_2 \neq i}^{n} \frac{P_{j1}\, P_{j2}}{(1-P_{j1})(1-P_1-P_{-i2})} \right.$$

$$+ \dots\dots +$$

$$\left. \sum_{j_1 \neq j_2 \neq \dots \neq j_{z-1} \neq i}^{n} \frac{P_{j1}\, P_{j2} \dots P_{n-1}}{(1-P_{j1})(1-Pj_1-Pj_2)\dots(1-Pj_1-Pj_2 \dots\dots P_{jy-4})} \right]$$

உதாரணம்

ஒரு முழுமைத் தொகுதியிலுள்ள உறுப்புகள் 10, 20, 30, 40, 50, 60, 70 ஆகும். இதிலிருந்து உறுப்புகளின் பருமனுக்குச் சம விகிதமுள்ளபடி சம வாய்ப்பிலா மாதிரி எடுக்கப்படுகிறது. மாதிரியின் பருமன் 2 என்றால் ஒவ்வோர் உறுப்பும் மாதிரியில் வருவதற்கான நிகழ்திறன்கள் $\pi_1, \pi_2, \dots \pi_7$ என்பவைகளின் மதிப்பைக் கண்டுபிடி. $\sum_{i=1}^{7} \pi_i = 2$ என்பதை நிரூபி.

விடை

இங்கு $P_1, P_2\ P_3, P_5, P_5, P_6, P_7$ என்பவை முறையே $\frac{1}{28}, \frac{2}{28}, \frac{3}{28}, \frac{4}{28}, \frac{5}{28}, \frac{6}{28}, \frac{7}{28}$ ஆகும்.

$$\pi_1 = P_1 + P_1 \sum_{j \neq 1} \frac{P_j}{1-P_j} = P_1 \left[1 + k - \frac{P_1}{1-P_1} \right]$$

$$k = \sum_{j=1}^{7} \frac{P_j}{1-P_j}$$

$$P_1 = \left[2 + k - \frac{1}{1-P_1} \right]$$

$$K = \sum_{j=1}^{7} \frac{P_j}{1-P_j} = -7 + \sum_{j=1}^{7} \frac{1}{1-P_j}$$

$$= -7 + \frac{28}{27} + \frac{28}{26} + \frac{28}{25} + \frac{28}{24} + \frac{28}{23} + \frac{28}{22} + \frac{28}{21}$$

$$= \frac{1}{27} + \frac{2}{26} + \frac{3}{25} + \frac{4}{24} + \frac{5}{23} + \frac{6}{22} + \frac{7}{21}$$

$$= 0{\cdot}0370 + 0{\cdot}0769 + 0{\cdot}1200 + 0{\cdot}1667 + 0{\cdot}2174 + 0{\cdot}2727 + 0{\cdot}3333$$

$$= 1{\cdot}2240.$$

$$\therefore \pi_1 = \frac{1}{28}[3{\cdot}2240 - 1{\cdot}0370] = \frac{2{\cdot}1870}{28} = 0{\cdot}0781$$

$$\pi_2 = P_2\left[2+k-\frac{1}{1-P_2}\right] = \frac{2}{28}[3{\cdot}2240 - 1{\cdot}0769]$$

$$= \frac{2{\cdot}1471 \times 2}{28}$$

$$= 0{\cdot}1534$$

$$\pi_3 = P_3\left[2+k-\frac{1}{1-P_3}\right] = \frac{3}{28}[3{\cdot}2240 - 1{\cdot}1200] = \frac{2{\cdot}1040 \times 3}{28}$$

$$= 0{\cdot}2254$$

$$\pi_4 = P_4\left[2+k-\frac{1}{1-P_4}\right] = \frac{4}{28}[3{\cdot}2240 - 1{\cdot}1667] = \frac{2{\cdot}0573}{7}$$

$$= 0{\cdot}2939$$

$$\pi_5 = P_5\left[2+k-\frac{1}{1-P_5}\right] = \frac{5}{28}[3{\cdot}2240 - 1{\cdot}2174] = \frac{2{\cdot}0066 \times 5}{28}$$

$$= 0{\cdot}3583$$

$$\pi_6 = P_6\left[2+k-\frac{1}{1-P_6}\right]=\frac{6}{28}[3\cdot2240-1\cdot2727]=\frac{1\cdot9513\times6}{28}$$

$$= 0\cdot4181$$

$$\pi_7 = P_7\left[2+k-\frac{1}{1-P_7}\right]$$

$$=\frac{7}{28}[3\cdot2240-1\cdot3333] = \frac{1\cdot8907}{4} = 0\cdot4727$$

$\sum_{i=1}^{7}\pi_i = 1\cdot9999$ என்று வருகிறது. இது தோராயமாக π_i பைகளைக்கண்டுபிடித்ததால் வருகிறது. உண்மையில் இது $2\cdot0000$ என்றிருக்கவேண்டும்.

உண்மையில் $$\sum_{i=1}^{7}\pi_i = \sum_{i=1}^{7}P_i\left[2+K-\frac{1}{1-P_i}\right]$$

$$= (2+k)\sum_{i=1}^{7}P_i - \sum_{i=1}^{7}\frac{P_i}{1-P_i}$$

$$= (2+k)\,1 - k$$

$$= 2. \qquad \text{(நி—து.)}$$

20·9. தேற்றம்

$$E^{-1}(X) \sim \sum_{i=1}^{n}\frac{x_i}{\pi_i},\ \text{[Horvitz \& Thompson]}$$

π_i' என்பது x_i என்ற உறுப்பு மாதிரியில் வருவதற்கான நிகழ் திறன்.

நிறுவல்

முழுமைத் தொகுதியிலிருந்து NP_n மாதிரிகள் எடுக்க முடியும் இவற்றுள் α ஆவது மாதிரி வருவதற்கான நிகழ்திறன் $P(\alpha)$ எனக் கொள்வோம்.

$$E \sum_{i=1}^{n} \left(\frac{x_i}{\pi_i'}\right) = \sum_{\alpha=1}^{NP_n} \left\{ \sum_{i=1}^{n} \frac{x_i}{\pi_i'} \right\}_{\alpha} P(\alpha)$$

$$= \sum_{i=1}^{N} \frac{X_i}{\pi_i'} \sum_{2\subset\alpha} P(\alpha) = \sum_{i=1}^{N} \frac{X_i}{\pi_i} \pi_i = \sum_{i=1}^{N} X_i$$

$= X$ [$i\subset\alpha$ என்பது X_i என்னும் உறுப்பு α ஆவது மாதிரியில் இருக்கிறது என்ற பொருள்படும்.]

$$\therefore \overset{-1}{E}(X) = \sum_{i=1}^{n} \frac{x_i}{\pi_i'} \qquad \text{(நி—து.)}$$

20·10. தேற்றம்

$$V\{\overset{-1}{E}(X)\} = \sum_{i=1}^{N} \sum_{j>i}^{N} (\pi_i \pi_j - \pi_{ij}) \left(\frac{X_i}{\pi_i} - \frac{X_j}{\pi_j}\right)^2$$

π_{ij} என்பது (x_i, X_j) என்னும் இரு உறுப்புகளும் மாதிரியில் வருவதற்கான நிகழ்திறன்.

நிறுவல்

$$V[\overset{-1}{E}(X)] = E[\overset{-1}{E}(X)]^2 - [E\overset{-1}{E}(X)]^2$$

$$= E[\overset{-1}{E}(X)]^2 - X^2$$

$$= E\left[\sum_{i=1}^{n} \frac{x_i^2}{\pi_i'^2} + 2 \sum_{j>i} \sum \frac{x_i\, x_j}{\pi_i'\, \pi_j'}\right] - X^2$$

$$= \sum_{\alpha=1}^{NP_n} \left[\sum_{i=1}^{n} \frac{x^2}{\pi_i'} + 2 \sum_{j>i}\sum \frac{x_i\, x_i}{\pi_i'\pi_j'} \right] P(\alpha) - X^2$$

$$= \sum_{i=1}^{N} \frac{X^2}{\pi i_2} \sum_{i<\alpha} P(\alpha) + \sum_{i=1}^{N} \sum_{j\neq i}^{N} \frac{X_i\, X_j}{\pi_i\, \pi_j}$$

$$\sum_{(i,j)\subset \alpha} P(\alpha) - X^2$$

[$(i, j) \subset \alpha$ என்பது X_i, X_j என்னும் உறுப்புகள் α ஆவது மாதிரியில் இருக்கும் எனப் பொருள்படும்.]

$$= \sum_{i=1}^{N} \frac{X_i^2}{\pi_i^2} \pi_i + \sum_{i=1}^{N} \sum_{j\neq i}^{N} \frac{X_i\, X_j}{\pi_i\, \pi_i} \pi_{ij} - X^2.$$

ஆனால் $X^2 = \sum_{i=1}^{N} X_i^2 + \sum_{j\neq 1}\sum X_i\, X_j.$

$$\therefore \quad V[E^{-1}(X)] = \sum_{i=1}^{N} \left(\frac{X_i^2}{\pi_i} - X_i^2 \right)$$

$$+ \sum_{i=1}^{N} \sum_{j\neq i}^{N} \left(\frac{X_i\, X_j\, \pi_{ij}}{\pi_i\, \pi_j} - X_i\, X_j \right)$$

$$= \sum_{i=1}^{N} \frac{1-\pi_i}{\pi_i} X_i^2$$

$$+ \sum_{i=1}^{N} \sum_{j\neq 1}^{N} \frac{\pi_{ij} - \pi_i\, \pi_i}{\pi_i\, \pi_i} X_i\, X_j.$$

மேலும் $\sum_{i=1}^{U} \pi_i = n, \sum_{j\neq 1}^{N} \pi_{ij} = (n-1)\pi_i$

[இதை நிருபி !].

$$\therefore \quad V[\overset{-1}{E}(X)] = \sum_{i=1}^{N} \pi_i (1-\pi_i) \frac{\pi_i^2}{\pi_i^2}$$

$$+ \sum_{i=1}^{N} \sum_{j\neq i}^{N} (\pi_{ij} - \pi_i \pi_j) \frac{X_i}{\pi_j} \frac{X_j}{\pi_j}.$$

$$= \sum_{i=1}^{N} \left[\sum_{j\neq i} \pi_i \pi_j - \pi_{ij} \right] \frac{X_i^2}{\pi_i^2}$$

$$+ \sum_{i=1}^{N} \sum_{j\neq i}^{N} (\pi_{ij} - \pi_i \pi_j) \frac{X_i X_j}{\pi_i \pi_j}$$

$$= \sum_{i=1}^{N} \sum_{j>i}^{N} (\pi \pi_{ij} - \pi_{ij}) \left(\frac{X_i}{\pi_i} - \frac{X_j}{\pi_j} \right)^2$$

(நி-து.)

20·11. தேற்றம்

$V[\overset{-1}{E}(X)]$ என்பதன் பிறழ்ச்சியற்ற மதிப்பீட்டிகள்

$$(அ) \sum_{i=1}^{n} (1-\pi_i') \frac{x_i^2}{\pi_i'^2} + \sum_{i=1}^{n} \sum_{j\neq 1}^{n} \frac{\pi_{ij}' - \pi_i' \pi_j'}{\pi'_{ij}}$$

$$\cdot \frac{x_i x_j}{\pi_i' \pi_j'} \quad \text{(Horvitz, Thompson)}$$

$$(ஆ)\ \sum_{i=1}^{n} \sum_{j>i} \frac{\pi_i' \pi_j' - \pi_{ij}'}{\pi_{ij}'} \left(\frac{x_i}{\pi_i'} - \frac{x_j}{\pi_j'} \right)^2$$

(Yates, Grundy)

மேற்கூறியவை இரண்டும் பின்வரும் தேற்றத்திலிருந்து விளங்கும்.

$X_1, X_2, \ldots X_N$ என்னும் உறுப்புகளோடு $A_1, A_2, \ldots A_N$ என்பதைச் சார்ந்திருக்கின்றன எனக் கொள்வோம். இதிலிருந்து சம வாய்ப்பிலா மாதிரி மாதிரி எடுத்தால் அவை $x_1, x_2, \ldots x_n$ எனவும், அவற்றைச் சார்ந்தவை $a_1, a_2, \ldots a_n$ எனவும் கொண்டால்

$$E\left[\sum_{i=1}^{n} a_i f(x_i)\right] = \sum_{\alpha}\left[\sum_{i=1}^{n} a_i f(x_i)\right]_{\alpha} P(\alpha)$$

$$= \sum_{i=1}^{N} A_i f(X_i) \sum_{i \subset \alpha} P(\alpha) = \sum_{i=1}^{N} A_i f(X_i)\, \pi_i$$

20·12. $X_1, X_2, \ldots, X_N$ என்னும் உறுப்புகளை $P_1, P_2, \ldots P_N$ என்னும் நிகழ்திறன்களுடன் எடுப்பதற்குப் பின்வரும் முறையைப் பின்பற்றலாம்.

ஒரு ராண்டாம் எண்ணை எடுத்து அதற்கு ஒரு தசமப் புள்ளியை முன்னே வைக்கவும். இது r எனக் கொண்டால் $0 \leq r \leq 1$.

$Q_0 = 0, Q_1 = P_1,\ Q_2 = P_1 + P_2,\ Q_3 = P_1 + P_2 + P_3, \ldots\ldots, Q_N = P_1 + P_2 + \ldots\ldots + P_N = 1$ எனக் கொள்ளவும்.

$Q_m < r \leq Q_{m+1}$ என்றிருந்தால் X_{m+1} என்னும் உறுப்பைத் தேர்ந்தெடுக்கவும் $[m = 1, 2, \ldots (N-1)]$.

$0 \leq r \leq Q_1$ என்றால் X_1 என்னும் உறுப்பைத் தேர் தெடுக்கவும். இம்முறையில் தேர்ந்தெடுக்கப்பட்ட உறுப்பு x ஆனது X_i ஆக இருப்பதற்கு நிகழ்திறன் $= P\,[Q_{i-1} < r \leq Q_i]$

r என்னும் ராண்டம் மாறிக்கு $[0, 1]$ என்னும் இடைவெளியில் செவ்வகப் பரவல்.

$$\therefore \quad P[r \leq y] = y.$$

$$\therefore \quad P[Q_{i-1} < r \leq Q_i] = P[r \leq Q_i] - P[r \leq Q_{i-1}]$$
$$= Q_i - Q_{i-1} = P_i.$$

பயிற்சி 5

1, $X_1, X_2, \ldots X_N$ என்னும் உறுப்புகளை முதல் தேர்வில் எடுப்பதற்கு நிகழ்திறன்கள் $P_1, P_2, \ldots, P_N$. ஈடு செய்யப்படாத சமவாய்ப்பிலா n உறுப்பு கொண்ட மாதிரியில் X_i வருவதற்கான நிகழ்திறன் π_i; (X_i, X_j) $j \neq i$ என்னும் உறுப்புகள் வருவதற்கான நிகழ்திறன் π_{ij}. π_i, π_{ij} முதலியவைகளை $P_1, P_2, \ldots$ P_N மூலமாக எழுது. $\sum_{i=1}^{N} \pi_i = n, \quad \sum_{i \neq j}^{n} \pi_{ij} = \pi_j \, (n-1).$

என்று நிரூபி.

விடை

π_i என்னவென்பதை முதலிலேயே பார்த்தோம். π_{ij} என்பதைப் பின்வருமாறு கண்டு பிடிக்கலாம்.

பின்வரும் மாதிரியின் நிகழ்திறனை முதலில் கணக்கிடலாம்.

$$\left[X_{j_1}, X_{j_2}, \ldots X_{j_{\alpha-1}}, X_i, X_{j_{\alpha+1}}, \ldots X_{j_{\beta-1}}, X_j, \ldots, X_{j_n}\right]$$
$$j_k \neq j_l, \; k \neq l$$

இது $P_{j_1} \cdot \dfrac{P_{j_2}}{1 - P_{j_1}} \cdot \dfrac{P_{j_3}}{1 - P_{j_1} - P_{j_2}} \cdots\cdots \dfrac{P_{j_{\alpha-1}}}{1 - P_{j_1} \ldots P_{j_{\alpha-2}}}$.

$$\frac{P_i}{1 - P_{j_1} \cdot P_{j_{\alpha-1}}} \cdot \frac{P_{j_{\alpha+1}}}{1 - P_{j_1} \ldots P_{j_{\alpha-1}} - P_i} \cdot \frac{P_{j_n}}{1 - P_{j_1} \ldots P_{j_{\alpha+1}}}$$

$j_1, j_2, \ldots j_n$-க்குச் சமமில்லாத 1-லிருந்து N வரையுள்ள எண்களை (i, j தவிர்த்து) எடுத்துக்கொண்டு மேற்கூறியவற்றின் கூட்டுத்தொகையைக் கண்டுபிடி. அது $P_{\alpha, \beta}$ என இருக்கட்டும்

$$\pi_{ij} = \sum_{\alpha<}^{n}\sum_{\beta}^{n} P_{\alpha\beta} + \sum_{\beta<}^{n}\sum_{\alpha}^{n} P_{\alpha\beta}.$$

இப்பொழுது $\sum_{i=1}^{N} \pi_i = n$ என்பதை நிரூபிப்போம்.

$$E\left[\sum_{i=1}^{n} a_i f(x_i)\right] = \sum_{i=1}^{N} A_i f(X_i)\pi_i$$

என்று முன்பே பார்த்தோம். இதில் $a_i = 1, f(x_i) = 1$ என எடுத்துக்கொண்டால் $A_i=1, f(X_i) = 1$ என்பது விளங்குகிறது.

$$\therefore \quad E\left(\sum_{i=1}^{n} 1\cdot 1\right) = \sum_{i=1}^{N} 1\cdot 1\cdot \pi_i$$

$$\therefore \quad n = \sum_{i=j}^{N} \pi_i$$

$\sum_{i\neq j}^{N} \pi_{ij} = \pi_{j(n-1)}$ என்பதை நிரூபிப்போம்.

X_j என்னும் உறுப்பு n உறுப்பு கொண்ட மாதிரியில் வருவதற்கு நிகழ்திறன் π_j ஆகும். S என்பது ஒரு மாதிரியானல்

$$P[X_i \in S/X_j \in S] = \frac{P[X_i \in S, X_j \in S]}{P[X_j \in S]} = \frac{\pi_{ij}}{\pi_j}.$$

ஆகவே X_j என்னும் உறுப்பு இருக்கும் மாதிரிகளை மட்டும் எடுத்துக் கொள்வோமேயானல் முன் தேற்றத்தின்படி

$$\sum_{i\neq j}^{N} \frac{\pi_{ij}}{\pi_j} = (n-1)$$

$$\therefore \quad \sum_{i \neq j}^{n} \pi_{ij} = (n-1)\,\pi_j$$

2. ஈடு செய்யப்பட்ட மூன்று உறுப்புள்ள சம வாய்ப்பிலா மாதிரியில் இரண்டு உறுப்புகளே சமமல்லாதவை. பின்வரும் மதிப்பீட்டிகள் முழுமைத் தொகுதியின் மொத்தமான X ஐப் பிறழ்ச்சியின்றி மதிப்பிடுகின்றன என்பதை நிரூபி.

(அ) $\frac{1}{3}\left[\frac{x_1}{p_1} + \frac{x_2}{p_2} + \frac{x_1 + x_2}{p_1 + p_2}\right]$

(ஆ) $\frac{x_1}{1-(1-p_1)^3} + \frac{x_2}{1-(1-p_2)^3}$

P_i-கள் X_i-களின் விகிதத்திலிருந்தால் மேற்கூறியவற்றுள் எது திறன் அதிகம் வாய்ந்தது ?

விடை : (அ).

சமவாய்ப்பில் ஈடு செய்யப்படாத n உறுப்புகள் கொண்ட மாதிரியிலிருந்து $\frac{t}{\lambda\ P(S)}$ எப்பொழுது முழுமைத் தொகுதியின் மொத்தமான Y-க்குப் பிறழ்ச்சியற்ற மதிப்பீட்டியாகும்?

[$t = \sum_{i=1}^{n} y_i$, $\lambda =$ மாறிலி, $P(S) =$ வரும் மாதிரியின் நிகழ்திறன்.]

அதிலிருந்து $\frac{(y_1 + y_2)(1-p_1)(1-p_2)}{(N-1)\,p_1 p_2,\,(2-p_1-p_2)}$ என்னும் 2 உறுப்புகள் கொண்ட மாதிரியின் மதிப்பீட்டி Y-க்குப் பிறழ்ச்சியற்றது என்பதை நிரூபி.

மேற்கூறப்பட்ட மதிப்பீட்டி $\frac{(y_1 + y_2)(1-p_2)}{2(N-1)\,p_1\,p_2}$ ஐ விடத் திறன் அதிகம் வாய்ந்தது என்பதை நிரூபி.

Midzuno, H , Ann. Inst Stat. Math , 3 (1952) ; Murthy, M.N., Sankhya 18, (1957).

விடை : $\lambda = \binom{N-1}{n-1}$.

21. படுகை மாதிரி முறை

(STRATIFIED SAMPLING)

21·1. இதுவரை நாம் பார்த்த மாதிரி முறைகளில் உறுப்புகள் நேராக முழுமைத் தொகுதியிலிருந்தே எடுக்கப்பட்டன. ஆனால் சில காரணங்களுக்காக முழுமைத் தொகுதி பல பகுதிகளாகப் பிரிக்கப்பட்டிருப்பதுண்டு. இப்படி, பகுதிகளாகப் பிரிக்கப் பட்டிருக்கும் முழுமைத் தொகுதியில் ஒவ்வொரு பகுதியிலிருந்தும், தனித்தனியாக, சார்பில்லாமல், ஒவ்வொரு பகுதி மாதிரிகளைத் தேர்ந்தெடுத்து, இப்பகுதி மாதிரிகளை ஒருங்கிணைத்து ஒரு மாதிரியை உண்டாக்கினால் அம்முறைக்குப் படுகை மாதிரி முறை யென்பது பெயர்.

இம்முறையைப் பின்வரும் காரணங்களுக்காகக் கையாளலாம்.

(அ) முழுமைத் தொகுதி முதலிலேயே பல பகுதிகளாகப் பிரிக்கப்பட்டிருக்கலாம். உதாரணமாக இந்தியா முழுவதையும் முழுமைத் தொகுதியெனக் கொண்டோமானால் இந்தியாவிலுள்ள மாநிலங்களைப் பகுதிகளாகக் கொள்ளலாம். அளவெடுப்பைப் பொறுத்தவரை மாநிலங்களையே பொறுப்புள்ளவையாக ஆக்கலாம். ஒவ்வொரு மாநிலமும் அதன் வசதிக்கேற்ப ஒரு மாதிரி முறையைக் கையாளலாம். இக்காரணத்தால் மாதிரியைப் பொறுத்த வரையில் ஒருவிதச் சுதந்திரம் கிடைக்கிறது. அதாவது குறிப்பிட்ட முறையொன்றையே இந்தியா முழுவதிலும் பின்பற்ற வேண்டும் என்ற கட்டுப்பாடு கிடையாது. அளவெடுப்பைப் பொறுத்த வரையில் இது மிகவும் சௌகரியமானதும் கூட.

(ஆ) முழுமைத் தொகுதிப் பட்டியல் (Sampling Frame) பல பகுதிகளாகப் பிரிக்கப்பட்டிருப்பதுண்டு. உதாரணமாக வயது வாரியாக, மத வாரியாக, கல்வி நிலை வாரியாக, மொழி வாரியாக, ஆண்-பெண் என்று தனித்தனியாக முழுமைத் தொகுதிப் பட்டியல்

பிரிக்கப்பட்டிருக்கலாம். இப்படிப் பிரிக்கப்பட்டுள்ள முழுமைத் தொகுதி ஒன்று கூட்டாமல் ஒவ்வொரு பகுதியிலிருந்தும் தனித் தனியாக மாதிரி எடுப்பது புள்ளியியலாளர்களுக்குச் சற்று எளிதானது.

(இ) முழுமைத் தொகுதியிலுள்ள உறுப்புகள் ஒன்றே போல் இல்லாமல் மிகவும் அதிகம் வேறுபட்டிருந்தால் (Heterogeneous) சாதாரண மாதிரி முறை முழுமைத் தொகுதியை முற்றிலும் பிரதிபலிக்கும் மாதிரி ஒன்றைத் தருமா என்பதில் ஐயமிருக்கலாம். ஒருவேளை மிகப் பெரிய உறுப்புகள் மட்டுமோ அல்லது மிகச்சிறிய உறுப்புகள் மட்டுமோ மாதிரியில் வந்து விடக்கூடும் என்ற அச்சமிருக்கலாம். இந்த நிலையில் முழுமைத் தொகுதியைப் பல, வேறுபாடு அதிகமிலா, பாகங்களாகப் பிரிப்பது மிகவும் இன்றியமையாததாகும். இப்படிப் பிரிக்கப்பட்ட பகுதிகள் ஒவ்வொன்றிலிருந்தும் தனித்தனியே மாதிரிகள் எடுப்பதனுல் எல்லா வித உறுப்புகளுக்கும் பிரதிநிதித்துவம் கொடுக்க முடிகிறது.

(ஈ) படுகை மாதிரி முறையில், குறிப்பிட்ட பிழையளவுக்கு அளவெடுப்பின் செலவைக் குறைக்கும் மாதிரியை எடுக்க வசதியுள்ளது. அல்லது குறிப்பிட்ட செலவுக்குப் பிழையளவைக் குறைக்கவும் இம்முறையில் முடியும். ஆகவே குறிப்பிடப்பட்ட கட்டுட்பாட்டுக்குட்பட்ட மிகச் சிறந்த மாதிரியை எடுக்க இம் முறையில் முடியும்.

(உ) முழுமைத் தொகுதியிலுள்ள வெவ்வேறு பகுதிகளில் வெவ்வேறு விவரங்கள் தெரிய வந்தால் இவ் விவரங்களைக் கொண்டு ஒவ்வொரு பகுதியிலும் சிறந்த மாதிரியை எடுக்க இம் முறை வசதியளிக்கிறது,

உதாரணம்

முழுமைத் தொகுதியில் 1, 2, 100, 101 என்று நான்கே உறுப்புகள் இருப்பதாகக் கொள்வோம். 2 உறுப்பு கொண்ட மாதிரிகளின் சராசரி பின்வருமாறு:

மாதிரி :	(1,	2)	மாதிரிச் சராசரி	1·5
மாதிரி :	(1,	100)	,,	50·5
,, :	(1,	101)	,,	51·0
,, :	(2,	100)	,,	51·0
,, :	(2,	101)	,,	51·5
,, :	(100,	101)	,,	100·5.

அதாவது மாதிரிச் சராசரிகள் 1·5-லிருந்து 100·5 வரை பரவியுள்ளன. இம்முழுமைத் தொகுதியை இரண்டு பகுதிகளாக அதாவது, (1, 2), (100, 101) என்று பிரித்து ஒவ்வொரு பகுதியி லிருந்தும் ஓர் உறுப்பைத் தேர்ந்தெடுப்பதாக வைத்துக் கொண்டோமானால் 2 உறுப்பு கொண்ட மாதிரிச் சராசரிகள் பின்வருமாறு:

மாதிரி	(1,	100)	மாதிரிச் சராசரி	50·5
,,	(1,	101)	,,	51·0
,,	(2,	100)	,,	51·0
,,	(2,	101)	,,	51·5.

அதாவது மாதிரிச் சராசரிகள் 50·5-லிருந்து 51·5 வரையே பரவியுள்ளன. இதிலிருந்து பிந்தைய முறையில் மாதிரிச் சராசரி யின் பரவற் படி மிகக் குறைவாக இருக்கும் என்பது தெரிகிறது.

21·2. வரை இலக்கணம்

முழுமைத் தொகுதி K பகுதிகளாகப் பிரிக்கப்பட்டிருக்கிறது எனக் கொள்வோம். இவைகளில் $N_1, N_2, \ldots N_k$ உறுப்புகள் இருப்பதாகக் கொள்வோம். ஆகவே முழுமைத் தொகுதியிலுள்ள மொத்த உறுப்புகள் $N = N_1 + N_2 + \ldots + N_k$ ஆகும்

r ஆவது பகுதியிலுள்ள உறுப்புகள்

$X_{r1}, Y_{r2}, X_{r3}, \ldots, X_{rN_r}$ எனக் கொள்வோம்.

r ஆவது பகுதியின் சராசரி $\bar{X}_r = \dfrac{1}{N_r}\displaystyle\sum_{i=1}^{N_r} X_{ri}$

r ஆவது பகுதியின் பரவற் படி $\sigma_r^2 = \dfrac{1}{N}\displaystyle\sum_{i=1}^{N_r} X_{ri}^2 - \bar{X}_r^2$.

முழுமைத் தொகுதியின் சராசரி $X = \dfrac{1}{N}\displaystyle\sum_{r=1}^{k}\sum_{i=1}^{N_r} X_{ri}$

$$= \frac{1}{N} \sum_{r=1}^{k} - N_r \frac{1}{N_r} \sum_{i=1}^{N_r} X_{ri}$$

$$= \frac{1}{N} \sum_{r=1}^{k} N_r \bar{X}_r$$

முழுமைத் தொகுதியின் பரவற்படி

$$\sigma^2 = \frac{1}{N} \sum_{r=1}^{k} \sum_{i=1}^{N_r} (X_{ri} - \bar{X})^2$$

$$\therefore \quad \sigma^2 = \frac{1}{N} \sum_{r=1}^{k} \left[\sum_{i=1}^{N_r} X_{ri} - \bar{X}_r + \bar{X}_r - \bar{X})^2 \right]$$

$$= \frac{1}{N} \sum_{r=1}^{k} N_r \cdot \frac{1}{N_r} \sum_{i=1}^{N_r} (X_{ri} - \bar{X}_r)^2$$

$$+ \frac{1}{N} \sum_{r=1}^{k} N_r (\bar{X}_r - \bar{X})^2$$

$$= \frac{1}{N} \sum_{r=1}^{k} N_r \sigma_r^2 + \frac{1}{N} \sum_{r=1}^{k} N_r (\bar{X}_r - \bar{X})^2$$

$$= \sigma^2_{\text{உள்}} + \sigma^2_{\text{இடை}}$$

$\sigma^2_{\text{உள்}}$ என்பது பகுதிகளுக்குள் உறுப்புகளின் பரவற் படி.

$\sigma^2_{\text{இடை}}$ என்பது பகுதிகளுக்கிடையே உள்ள பரவற் படி.

$\frac{N_r}{N}$ என்பதை W_r என்று குறிப்பிடுவோம்.

21·3. முழுமைத் தொகுதியின் சராசரியை மதிப்பிடவேண்டும் எனக் கொள்வோம். $T_1, T_2, \dots T_k$ என்பவை $\bar{X}_1, \bar{X}_2, \dots, \bar{X}_k$ என்பவைகளின் பிறழ்ச்சியற்ற மதிப்பீட்டிகளாக இருக்கட்டும். T_r என்னும் மதிப்பீட்டி rஆவது பகுதியிலிருந்து எடுக்கப்பட்ட மாதிரியைக் கையாண்டு அமைக்கப்பட்டது எனக் கொள்வோம். $V(T_r)$ என்பது S_r^2 எனக் கொள்வோம். $V(T_r)$-ன் மதிப்பீட்டி $\overset{-1}{E}(S_r^2)$ எனக் கொள்வோம்.

21·4. தேற்றம்

(அ) $$\overset{-1}{E}(\bar{X}) \sim \frac{1}{N}\sum_{r=1}^{k} N_r T_r = \sum_{r=1}^{k} W_r T_r$$

(ஆ) $$V[\overset{-1}{E}(X)] = \frac{1}{N^2}\sum_{r=1}^{k} N_r^2 S_r^2 = \sum_{r=1}^{k} W_r^2 S_r^2$$

(இ) $$\overset{-1}{E}\{V[\overset{-1}{E}(X)]\} \sim \frac{1}{N^2}\sum_{r=1}^{k} N_r^2 \overset{-1}{E}(S_r^2)$$

$$= \sum_{r=1}^{k} W_r^2 E^{-1}(S_r^2)$$

நிறுவல்

(அ) $$E\left[\frac{1}{N}\sum_{r=1}^{k} N_r T_r\right] = E\left[\sum_{r=1}^{k} W_r T_r\right]$$

$$= \sum_{r=1}^{k} W_r E(T_r) = \sum W_r \bar{X}_r = \bar{X}$$

$$\therefore \quad E^{-1}(\bar{X}) \sim \sum_{r=1}^{k} W_r T_r \qquad \text{(நி-து.)}$$

(ஆ) $$V[E^{-1}(\bar{X})] = V\left[\sum_{r=1}^{k} W_r T_r\right]$$

$$= \sum_{r=1}^{k} W_r^2 V(T_r)$$

[$\because$ $T_r, Ts; r \neq s$ என்பவை சார்பற்றவை].

$$= \sum_{r=1}^{k} W_r^2 S_r^2 \qquad \text{(நி-து.)}$$

(இ) $$E^{-1}\{V[E^{-1}(\bar{X}]\} = E^{-1}\left[\sum_{r=1}^{k} W_r^2 S_r^2\right]$$

$$\sim \sum_{r=1}^{k} W_r^2 E^{-1}(S_r^2) \qquad \text{(நி-து.)}$$

மேற்கூறியவைகளிலிருந்து பகுதி மதிப்பீட்டிகளின் மூலம் முழுமைத் தொகுதியளவை எவ்வாறு மதிப்பிடுவது, அம்மதிப்பீட்டியின் பரவற் படியென்ன, அப்பரவற் படியின் பிறழ்ச்சியற்ற மதிப்பீட்டியென்ன என்பவை விளங்கும்.

21·5. கிளைத்தேற்றம்

$$V[E^{-1}(\bar{X})] = \sum_{r=1}^{k} W_r^2 V(T_r)$$

ஆகவே $V[\overline{E}^{1}(\overline{X})]$ என்பது பகுதிகளுக்குள் உள்ள உறுப்புகளின் பரவற் படிகளையே பொறுத்திருக்கிறது. பகுதிகளுக்கிடையே உள்ள பரவற் படி இதில் வரவில்லை. ஆகவே $V[\overline{E}^{1}(X)]$ ஐக் குறைக்கவேண்டுமென்றால் பகுதிகளுக்குள் உள்ள பரவற் படிகளைக் குறைக்கவேண்டும்.

$$\sigma^2 = \sigma^2_{\text{உள்}} + \sigma^2_{\text{இடை}}.$$

$$\therefore \quad \sigma^2_{\text{உள்}} = \sigma^2 - \sigma^2_{\text{இடை}}.$$

$\therefore$ $\sigma^2_{\text{உள்}}$-ஐக் குறைப்பதென்பது $\sigma^2_{\text{இடை}}$ ஐப் பெருமமாக்குவதற்குச் சமம். ஆகவே படுகை முறை மாதிரி அதிகப் பயனை அளிக்க வேண்டுமென்றால் பின்வருமாறு பகுதிகளை அமைப்பது சாலச் சிறந்தது.

பகுதிகளுக்குள் உறுப்புகள் அதிக வேறுபாடின்றி இருக்க வேண்டும். பகுதிகளுக்கிடையே வேறுபாடு அதிகமாக இருக்க வேண்டும்.

21·6. பங்கீட்டு முறை (Allocation)

k பிரிவுகளாகப் பிரிக்கப்பட்ட முழுமைத் தொகுதியிலிருந்து n உறுப்பு கொண்ட மாதிரியொன்று தேவைப்படுகிறதெனக் கொள்வோம். k பிரிவுகளிலிருந்தும் முறையே $n_1, n_2, \ldots, n_k$, $[n_1 + n_2 + \ldots + n_k = n]$ உறுப்புகள் கொண்ட மாதிரிகளைத் தனித்தனியாக எடுப்பதன் மூலம் இதை நிறைவேற்றலாம். இதில் ஒரே ஒரு நிபந்தனை

$n_1 + n_2 + \ldots + n_k = n$ என்பதாகும்.

மேற்கூறப்பட்ட நிபந்தனையை ஏற்குமாறு பலவித $(n_1, n_2, \ldots, n_k)$-க்களை நாம் அமைக்க முடியும். இப்படியமைக்கப்படும் ஒவ்வொரு பகிர்வுக்கும் பங்கீடு என்பது பெயர். அதாவது n எனும் எண்ணை எல்லாப் பகுதிகளுக்கும் பங்கிடுவதையே இது குறிக்கும்.

இப்பங்கீட்டினைப் பலமுறைகளில் செய்யலாம்.

21·7. (அ) விகித சமப் பங்கீடு (Proportional Allocation)

இம்முறையில் ஒவ்வொரு பகுதியிலிருக்கும் உறுப்புகளின் எண்ணிக்கைக்குச் சமமான விகிதத்தில் மாதிரிப் பருமன்களை

அமைக்க வேண்டும். அதாவது பெரிய பகுதியிலிருந்து அதிக மான உறுப்புகளும் சிறிய பகுதியிலிருந்து குறைவான உறுப்புகளும் வருவதை இப்பங்கீட்டு முறை நிறைவேற்றும்.

$$\frac{n_1}{N_1} = \frac{n_2}{N_2} = \cdots = \frac{n_k}{N_k} = \frac{n_1 + n_2 + \cdots + n_k}{N_1 + N_2 + \cdots + N_k} = \frac{n}{N}$$

என்பதால் $n_1 = \frac{N_1}{N} \cdot n$, $n_2 = \frac{N_2}{N} \cdot n$, $\ldots$, $n_k = \frac{N_k}{N} \cdot n$ என்பது விளங்கும்.

இப் பங்கீட்டு முறைக்குப் பௌலி பங்கீட்டு முறை (Bowley's Allocations) என்றும் பெயர்.

உதாரணம்

40, 50, 60, 100 உறுப்புகளையுடைய நான்கு பகுதிகளையுடைய முழுமைத் தொகுதியிலிருந்து 50 உறுப்புகளைக் கொண்ட விகித சமப் பங்கீடு யாது?

இங்கு $N_1 = 40$, $N_2 = 50$, $N_3 = 60$, $N_4 = 100$

$\therefore \quad N = 40 + 50 + 60 + 100 = 250.$

$$n = 50. \qquad \therefore \quad \frac{n}{N} = \frac{50}{250} = \frac{1}{5}.$$

$$\therefore \quad n_1 = N_1 \cdot \frac{n}{N} = \frac{40}{5} = 8; \quad n_2 = N_2 \cdot \frac{n}{N} = 50 \times \frac{1}{5} = 10;$$

$$n_3 = N_3 \cdot \frac{n}{N} = 60 \times \frac{1}{5} = 12;$$

$$n_4 = N_4 \cdot \frac{n}{N} = 100 \times \frac{1}{5} = 20.$$

$\therefore$ பங்கீடு : (8, 10, 12, 20)

(ஆ) உத்தமப் பங்கீட்டு முறை (Optimum Allocation)

இம்முறையை மூன்று பிரிவுகளாகப் பிரிக்கலாம்.

(i) ஒரு குறிப்பிட்ட மாதிரிப் பருமனுக்கு மதிப்பீட்டியின் பரவற் படியைச் சிறுமமாக்கும் உத்தமப் பங்கீடு.

(ii) ஒரு குறிப்பிட்ட செலவுக்கு மதிப்பீட்டியின் பரவற் படியைச் சிறுமமாக்கும் உத்தமப் பங்கீடு.

(iii) ஒரு குறிப்பிட்ட மதிப்பீட்டியின் பரவற் படிக்குச் செலவைச் சிறுமமாக்கும் உத்தமப் பங்கீடு.

(i) ஒரு குறிப்பிட்ட மாதிரிப் பருமனுக்கு மதிப்பீட்டியின் பரவற் படியைச் சிறுமமாக்கும் உத்தமப் பங்கீடு

இப்பங்கீட்டு முறையில் $n = n_1 + r_2 + \dots + n_k$ என்பது கொடுக்கப்பட்டிருக்கிறது. இந்த நிபந்தனைக்குட்பட்டு $V[E^{-1}(\overline{X})]$ஐச் சிறுமமாக்கும் $n_1, n_2, \dots n_k$ என்பவைகளை நாம் கண்டுபிடிக்க வேண்டும்.

நிபந்தனைக்குட்பட்ட சிறுமம், பெருமம் மதிப்பீட்டிற்கு நாம் லெக்ராஞ்ச் பெருக்கி முறை (Method of Lagrangean Multipliers) ஒன்றைத்தான் கையாளவேண்டி வருகிறது. (இம்முறை சுருக்கமாகப் பயிற்சி 1, கணக்கு 3-ல் விளக்கப்பட்டிருக்கிறது.)

$$V[E^{-1}(X)] = \sum_{r=1}^{k} W_r^2 S_r^2$$

நிபந்தனை $n_1 + n_2 + \dots + n_k = n$.

$$F = \sum_{r=1}^{k} W_r^2 S_r^2 + \lambda [n_1 + n_2 + \dots + n_k - n]$$ எனக் கொள்க.

இதில் n_i என்பது $W_i^2 S_i^2$ என்பதிலும் λn_i என்பதிலும் தான் வரும் என்பது தெளிவு.

$$\frac{\partial F}{\partial n_i} = \frac{\partial}{\partial n_i}(W_i^2 S_i^2) + \lambda = 0. \quad i = 1, 2, \dots k.$$

இந்த k சமன்பாடுகளுடன் $n_1 + n_2 + \dots + n_k = n$ என்னும் சமன்பாட்டையும் உபயோகித்து $n_1, n_2, \dots, n_k, \lambda$ முதலியவைகளை $W_1, \dots, W_k, S_1^2, \dots S_k^2$-களின் சார்பாகக் கண்டுபிடிக்கலாம்.

(ii) ஒரு குறிப்பிட்ட செலவுக்கு மதிப்பீட்டியின் பரவற் படியைச் சிறுமமாக்கும் உத்தமப் பங்கீடு

$(n_1, n_2, ..., n_k)$ என்னும் பங்கீட்டிற்கு ஆகும் செலவு $C(n_1, n_2, ... n_k)$ எனக் கொள்வோம். சாதாரணமாக $C(n_1, ..., n_k)$ என்பதை $C_0 + C_1 n_1 + C_2 n_2 + ... + C_k n_k$ என்று எடுத்துக் கொள்வது வழக்கம்.

மொத்தமாக ஒதுக்கப்படும் செலல C' எனக் கொண்டால் நிபந்தனை

$C(n_1, n_2, ..., n_k) = C'$ என்பதாகும்,

இந்த நிபந்தனைக்குட்பட்ட $V[\overset{-1}{E}(\bar{X})]$ ஐச் சிறுமமாக்கும் $(n_1, n_2, ... n_k)$-க்களையே நாம் இம்முறையில் கணக்கிட வேண்டும்.

$$F = \sum_{r=1}^{k} W_r S_r^2 + \lambda [C(n_1, ..., n_k) - C']$$

$$\frac{\partial F}{\partial n_i} = \frac{\partial}{\partial n_i}(W_i S_i^2) + \lambda \frac{\partial C}{\partial n_i} = 0 \qquad i = 1, 2, ... k$$

இந்த k சமன்பாடுகளுடன் $C(n_1, ..., n_k) = C'$ என்னும் சமன்பாட்டையும் சேர்த்து $n_1, n_2, ..., n_k, \lambda$ முதலியவைகளைக் கண்டு பிடிக்கலாம்.

(iii) ஒரு குறிப்பிட்ட மதிப்பீட்டியின் பரவற் படிக்குச் செலவைச் சிறுமமாக்கும் உத்தமப் பங்கீடு

இங்கு $V[\overset{-1}{E}(\bar{X})]$ என்பது தரப்பட்டிருக்கிறது. இதை V எனக் கொள்வோம். இது $\sum_{r=1}^{k} W_r^2 S_r^2 = V'$ என்னும் சமன்பாட்டிற்குச் சமம். இதுதான் இங்கு நிபந்தனையாகும். இந்த நிபந்தனைக்குட்பட்ட $C(n_1, ..., n_k)$ என்னும் செலவுச் சார்பை (Cost function) சிறுமமாக்கும் $(n_1, ..., n_k)$ என்பவையே இப் பங்கீடாகும்.

$$F = C(n_1, ..., n_k) + \lambda \left[\sum_{r=1}^{k} W_r^2 S_r^2 - V' \right]$$

$$\frac{\partial F}{\partial n_i} = \frac{\partial C}{\partial n_i} + \lambda \frac{\partial}{\partial n_i} (W_r^2 S_r^2) = 0 \qquad i = 1, 2, \ldots, k$$

என்னும் k சமன்பாடுகளுடன் நிபந்தனையான $\sum_{r=1}^{k} W_r^2 S_r^2 = V'$ என்னும் சமன்பாட்டையும் சேர்த்து n_1, n_2, n_k, λ முதலியவைகளைக் கண்டுபிடிக்கலாம்.

கவனிக்கவும்

(ii)-ல் $C(n_1, \ldots, n_k) = n_1 + n_2 + \ldots + n_k$ என்றும், $C' = n$ என்றும் எடுத்துக் கொண்டால் (i) கிடைக்கிறது. ஆகவே (i) என்பது (ii)-ன் குறிப்பிட்ட அம்சமாகும் (Particular case).

21·8. படுகை மாதிரி முறையில் ஈ. செ. சா. ரா. மா. முறை

படுகை மாதிரி முறையில் ஒவ்வொரு பகுதியிலும், தனித் தனியாக, சார்பற்ற முறையில் ஈ. செ. சா. ரா. மா. முறை கையாளப் படுகிறதெனக் கொள்வோம்.

இங்கு $\bar{X}_r$ என்னும் r ஆவது பகுதிச் சராசரிக்கு $\bar{x}_r$ என்னும் r ஆவது பகுதியிலிருந்து எடுக்கப்பட்ட n_r உறுப்புகள் கொண்ட ஈ. செ. சா. ரா. மா. சராசரி பிறழ்ச்சியற்ற மதிப்பீட்டியாகும்.

அதாவது $T_r = \bar{x}_r$.

$$S_r^2 = V(T_r) = V(\bar{x}_r) = \frac{\sigma_r^2}{n_r}.$$

$\overset{-1}{E} [V(T_r)] \sim \frac{s_r^2}{n_r}$. s_r^2 என்பது r ஆவது பகுதியிலிருந்து எடுக்கப்பட்ட மாதிரியின் மாதிரிப் பரவற்படி. இவைகளிலிருந்து பின்வரும் தேற்றம் தெளிவாகிறது.

தேற்றம் 21·9.

ஒவ்வொரு பகுதியிலும் ஈ. செ. சா. ரா. மா. முறை உபயோகப் படுத்தப்பட்ட படுகை மாதிரி முறையில், $\bar{x}_r, s_r^2, n_r$ என்பவை r ஆவது பகுதியிலிருந்து எடுக்கப்பட்ட மாதிரியின் சராசரி, பரவற் படி, பருமன் எனக்கொண்டால்,

$$(\text{அ}) \quad E^{-1}\,(\overline{X}) \sim \sum_{r=1}^{k} W_r\,\bar{x}_r \qquad W_r = \frac{N_r}{N}.$$

$$(\text{ஆ}) \quad V\,[E^{-1}\,(X)] = \sum_{r=1}^{k} W_r^2\,\frac{\sigma_r^2}{n_r}$$

$$(\text{இ}) \quad E^{-1}\,\{V\,[E^{-1}\,(X)]\} = \sum_{r=1}^{k} W_r^2\,\frac{s_r^2}{n_r}$$

நிறுவல்

$$(\text{அ}) \quad E\left[\sum_{r=1}^{k} W_r\,\bar{x}_r\right] = \sum_{r=1}^{k} W_r\,\overline{X}_r = \overline{X}$$

$$\therefore \quad E^{-1}\,(\overline{X}) \sim \sum_{r=1}^{k} W_r\,\bar{x}_r \qquad (\text{நி-து.})$$

$$(\text{ஆ}) \quad V\,[E^{-1}\,(\overline{X})] = V\left[\sum_{r=1}^{k} W_r\,\bar{x}_r\right] = \sum_{r=1}^{k} W_r^2\,\frac{\sigma_r^2}{n_r}$$

(நி-து.)

$$(\text{இ}) \quad E^{-1}\,\{V\,[E^{-1}\,(\overline{X})\,]\} \sim E^{-1}\left[\sum_{r=1}^{k} W_r^2\,\frac{\sigma_r^2}{n_r}\right]$$

$$\sim \sum_{r=1}^{k} \frac{W_r^2}{n_r}\,E^{-1}\,(\sigma_r^2)$$

$$\sim \sum_{r=1}^{k} W_r^2 \frac{s_r^2}{n_r} \qquad \text{(நி-து.)}$$

21·10. படுகை மாதிரி முறையில் ஒவ்வொரு பகுதியிலும் ஈ. செ. சா. ரா. மா. முறை கையாளப்படும்பொழுது சமவிகிதப் பங்கீடு

சம விகிதப் பங்கீட்டில் $n_r = \frac{N_r}{N} n = W_r n$

$$V[\overset{-1}{E}(X)] = \sum_{r=1}^{k'} W_r^2 \frac{\sigma_r^2}{n_r} = \sum_{r=1}^{k} W_r^2 \frac{\sigma_r^2}{W_r n}$$

$$= \frac{1}{n} \sum_{r=1}^{k} W_r \sigma_r^2$$

இதை $V_{\text{ச.வி.ப.}}$ என்று குறிப்பிடுவோம். [ச.வி.ப: சமவிகிதப் பங்கீடு.]

அதாவது $V_{\text{ச. வி. ப.}} = \frac{1}{n} \sum_{r=1}^{k} W_r \sigma_r^2$

21·11. படுகை மாதிரி முறையில் ஒவ்வொரு பகுதியிலும் ஈ. செ. சா. ரா. மா. முறை கையாளப்படும்பொழுது உத்தமப் பங்கீடு

(i) குறிப்பிட்ட மாதிரிப் பருமனுக்கு மதிப்பீட்டின் பரவற் படியைச் சிறுமமாக்கும் உத்தமப் பங்கீடு

$$V[\overset{-1}{E}(\bar{X})] = \sum_{r=1}^{k} W_r^2 \frac{\sigma_r^2}{n_r}$$

நிபந்தனை $n_1 + n_2 + \dots\dots + n_k = n$

$$F = \sum_{r=1}^{K} W_r^2 \frac{\sigma_r^2}{n_r} + \lambda \, [n_1 + n_2 + \ldots\ldots + n_k - n]$$

$$\frac{\partial F}{\partial n_r} = - \frac{W_r^2 \, \sigma_r^2}{n_r^2} + \lambda = 0. \quad r = 1, 2, \ldots, k$$

$$n_r = \frac{W_r \, \sigma_r}{\sqrt{\lambda}}$$

$$n = \sum_{r=1}^{k} n_r = \frac{\sum_{r=1}^{k} W_r \, \sigma_r}{\sqrt{\lambda}}$$

$$\therefore \quad \frac{1}{\sqrt{\lambda}} = \frac{n}{\sum_{r=1}^{k} W_r \, \sigma_r}$$

$$\therefore \quad n_r = \frac{W_r \, \sigma^2}{\sum_{r=1}^{k} W_r \, \sigma_r} \cdot n$$

இதிலிருந்து இந்த உத்தம முறையில் n_r-கள் $W_r \, \sigma_r$-களின் விகிதத்தில் இருக்கின்றன என்பது புலப்படும்.

இந்தப் பங்கீட்டில் $V \, [E^{-1} (\overline{X})] = V_{\text{உத். பங்.}}$ எனக் கொள்வோம்.

$$V_{\text{உத். பங்.}} = \sum_{r=1}^{k} W_r^2 \frac{\sigma_r^2}{n_r} = \sum_{r=1}^{k} W_r^2 \frac{\sigma_r^2}{\frac{W_r \, \sigma_r}{\sum_{r=1}^{k} W_r \, \sigma_r} \cdot n}$$

$$= \frac{1}{n}\left[\sum_{r=1}^{k} W_r \, \sigma_r\right]^2$$

கிளைத்தேற்றம்

$V_{\text{சா. ரா. மா.}}$ என்பது படுகை முறையல்லாத, சாதாரண, ஈ. செ. சா. ரா. மா-ன் சராசரியின் பரவற் படியென்றால்

$$V_{\text{சா. ரா. மா.}} = \frac{\sigma^2}{n} = \frac{1}{n}\left[\frac{1}{N}\sum_{r=1}^{k} N_r \, \sigma_r^2\right.$$

$$\left. + \frac{1}{N}\sum_{r=1}^{k} N_r \, (\overline{X}_r - \overline{X})^2\right]$$

$$= \frac{1}{n}\sum_{r=1}^{k} W_r \, \sigma_r^2 + \frac{1}{n}\sum_{r=1}^{k} W_r \, (\overline{X}_r - \overline{X})^2$$

$$= V_{\text{ச. வி. ப.}} + \frac{1}{n}\sum_{r=1}^{k} W_r \, (X_r - \overline{X})^2.$$

$$\therefore \quad V_{\text{சா. ரா. மா.}} - V_{\text{ச. வி. ப.}} = \frac{1}{n}\sum_{r=1}^{k} W_r \, (X_r - X)^2 \geq 0$$

$$\therefore \quad V_{\text{சா. ரா. மா.}} \geq V_{\text{ச. வி. ப.}}$$

$$V_{\text{ச. வி. ப.}} = \frac{1}{n} \sum_{r=1}^{k} W_r \, \sigma_r^2$$

$$V_{\text{உத் . பங்.}} = \frac{1}{n} \left[\sum_{r=1}^{k} W_r \, \sigma_r \right]^2$$

$$V_{\text{ச·வி·ப.}} - V_{\text{உத்.பங்.}} = \frac{1}{n} \left[\sum_{r=1}^{k} W_r \, \sigma_r^2 - \left[\sum_{r=1}^{k} W_r \, \sigma_r \right]^2 \right]$$

$$\left[\sum_{r=1}^{k} a_r \, b_r \right]^2 \leq \sum_{r=1}^{k} a_r^2 \sum_{r=1}^{k} b_r^2$$ என்பது

கோஷி-ஷ்வார்ஸ் (Cauchy-Schwartz) சமனிலியாகும்,

இதில் $a_r = \sqrt{W_r} \, \sigma_r$ $b_r = \sqrt{W_r}$ எனக் கொண்டால்,

$$\left[\sum_{r=1}^{k} W_r \sigma_r \right]^2 \leq \sum_{r=1}^{k} W_r \, \sigma_r \cdot \sum_{r=1}^{k} W_r$$

$$W_r = \frac{N_r}{N} \text{ என்பதால் } \sum_{r=1}^{k} W_r = \frac{\sum_{r=1}^{k} N_r}{N} = 1$$

$$\therefore \left[\sum_{r=1}^{k} W_r \sigma_r \right]^2 \leq \sum_{r=1}^{k} W_r \sigma_r^2$$

$$\therefore V_{\text{ச.வி.ப.}} - V_{\text{உத்.பங்.}} \geq 0$$

$$\therefore V_{\text{ச.வி.ப}} \geq V_{\text{உத்.பங்.}} \qquad \ldots \ldots \text{II}$$

I, II—இரண்டையும் இணைத்து

$$V_{\text{சா.ரா.மா.}} \geq V_{\text{ச.வி.ப.}} \geq V_{\text{உத்.பங்.}}$$

என்று எழுதலாம்.

இதிலிருந்து இம்மூன்று முறைகளில் உத்தமப் பங்கீட்டு முறை திறன் அதிகம் வாய்ந்தது என்பது தெளிவாகத் தெரிகிறது.

11·11. குறிப்பிட்ட செலவுக்கு மதிப்பீட்டியின் பரவற் படியைச் சிறுமமாக்கும் உத்தமப் பங்கீட்டு முறை

இங்குச் செலவுச் சார்பு (Cost function) $C = C_0 + C_1 n_1 + C_2 n_2 + \dots + C_K n_K$ எனக் கொள்வோம். இதில் C_0 என்பது மாதிரியாலல்லாது மற்ற விதங்களில் ஏற்படும் செலவு. C_r என்பது r ஆவது பகுதியில் ஓர் உறுப்பினிடமிருந்து விவரங்களைப் பெருவதற்கான செலவு. C' என்பது குறிப்பிட்ட செலவு எனக் கொண்டால் நிபந்தனை $C_0 + C_1 n_1 + C_2 n_2 + \dots + C_k n_k = C'$ ஆகும்.

$$V[E^{-1}(\overline{X})] = \sum_{r=1}^{k} W_r^2 \frac{\sigma_r^2}{n_r}$$

$$\therefore \quad F = \sum_{r=1}^{k} W_r^2 \; \frac{\sigma_r^2}{n_r} +$$

$$\lambda [C_0 + C_1 n_1 + C_2 n_2 + \ldots + C_k n_k - C']$$

$$\frac{\partial F}{\partial n_r} = - \frac{W_r^2 \sigma_r^2}{n_r^2} + \lambda C_r = 0 ; \qquad r = 1, 2, \ldots, k$$

$$\therefore \quad n_r = \frac{W_r \sigma_r}{\sqrt{\lambda c_r}}$$

$C_0 + C_1 n_1 + C_2 n_2 + \ldots + C_K n_K = C'$ ஆகையால்

$$C_0 + \sum_{r=1}^{k} \frac{W_r \sigma_r}{\sqrt{\lambda C_r}} C_r = C'$$

$$\therefore \quad \frac{1}{\sqrt{\lambda}} \sum_{r=1}^{k} \frac{W_r \sigma_r}{\sqrt{C_r}} C_r = C' - C_0$$

$$\therefore \quad \frac{1}{\sqrt{\lambda}} = \frac{C' - C_0}{\sum_{r=1}^{k} \frac{W_r \sigma_r}{\sqrt{c_r}} C_r} = \frac{C' - C_0}{\sum_{r=1}^{k} W_r \sigma_r \sqrt{C_r}}$$

$$\therefore \quad n_r = \frac{W_r \sigma_r}{\sqrt{C_r}} \; \frac{C' - C_0}{\sum_{r=1}^{k} W_r \sigma_r \sqrt{C_r}}$$

$$\therefore \quad V[\overset{-1}{E}(\bar{X})] = \sum_{r=1}^{k} W_r^2 \frac{\sigma_r^2}{n_r}$$

$$= \sum_{r=1}^{k} W_r^2 \sigma_r^2 \frac{\sum_{r=1}^{k} W_r \sigma_r \sqrt{C_r}}{(C'-C_0)\frac{W_r \sigma_r}{\sqrt{C_r}}}$$

$$= \frac{1}{(C'-C_0)} \left[\sum_{r=1}^{k} W_r \sigma_r \sqrt{C_r} \right]^2$$

(iii) ஒரு குறிப்பிட்ட மதிப்பீட்டியின் பரவற் படிக்குச் செலவைச் சிறுமமாக்கும் உத்தமப் பங்கீடு

இங்கு $\sum_{r=1}^{k} W_r^2 \frac{\sigma_r^2}{n_r} = V'$ என்பது நிபந்தனை.

$C = C_0 + C_1 n_1 + C_2 n_2 + \dots + C_k n_k$ என்னும் செலவுச் சார்பைச் சிறுமமாக்கவேண்டும்.

$$F = C_0 + C_1 n_1 + C_2 n_2 + \dots + C_k n_k$$

$$+ \lambda \left[\sum_{r=1}^{k} W_r^2 \frac{\sigma_r^2}{n_r} - V' \right]$$

$$\frac{\partial F}{\partial n_r} = C_r + \lambda \left(- \frac{W_r^2 \sigma_r^2}{n_r^2} \right) = 0 \quad r = 1, 2, \dots, k$$

$$\therefore \quad n_r = W_r \sigma_r \sqrt{\frac{\lambda}{C_r}} .$$

$$V' = \sum_{r=1}^{k} W_r^2 \frac{\sigma_r^2}{n_r} = \sum_{r=1}^{k} W_r^2 \sigma_r^2 \frac{1}{W_r \sigma_r} \sqrt{\frac{C_r}{\lambda}}$$

$$\therefore \quad \sqrt{\lambda} = \frac{1}{V'} \sum_{r=1}^{k} W_r \sigma_r \sqrt{C_r}$$

$$\therefore \quad n_r = \frac{W_r \sigma_r}{\sqrt{C_r}} \frac{1}{V'} \sum_{r=1}^{k} W_r \sigma_r \sqrt{C_r}$$

$\therefore$ சிறுமச் செலவு (Minimum cost)

$$= C_0 + \sum_{r=1}^{k} C_r n_r = C_0 + \sum_{r=1}^{k} C_r \frac{W_r \sigma_r}{\sqrt{C_r}} \frac{1}{V'} \cdot \sum_{r=1}^{k} W_r \sigma_r \sqrt{C_r}$$

$$= C_0 + \frac{1}{V'} \left[\sum_{r=1}^{k} W_r \sigma_r \sqrt{C_r} \right]^2$$

21·12. ஒவ்வொரு பகுதியிலும் ஈ. செயா. சா. ரா. மா. முறை கையாளப்படும் படுகை மாதிரி முறை

தேற்றம்

ஒவ்வொரு பகுதியிலும் ஈ. செயா. சா. ரா. மா. முறை கையாளப்படும் படுகை மாதிரி முறையில்.

$$(\text{அ}) \quad E^{-1}(\bar{X}) \sim \sum_{r=1}^{k} W_r \bar{x}_r$$

(ஆ) $$V[E^{-1}(\bar{X})] = \sum_{r=1}^{k} W_r^2 \frac{N_r - n_r}{N_r - 1} \frac{\sigma_r^2}{n_r}$$

(இ) $$E^{-1}\{V[E^{-1}(\bar{X}))\} \sim \sum_{r=1}^{k} W_r^2 \frac{N_r - n_r}{N_r - n_r} s_r^2.$$

நிறுவல்

(அ) $E(\bar{x}_r) = \bar{X}_r$

$$\therefore \quad E\left[\sum_{r=1}^{k} W_r \bar{x}_r\right] = \sum_{r=1}^{k} W_r \bar{X}_r = \bar{X}.$$

$$\therefore \quad E^{-1}(\bar{X}) \sim \sum_{r=1}^{k} W_r \bar{x}_r$$ (நி-து.)

(ஆ) $$V(\bar{x}_r) = \frac{N_r - n_r}{N_r - 1} \frac{\sigma_r^2}{n_r}.$$

$$\therefore \quad V\left[\sum_{r=1}^{k} \bar{x}_r W_r\right] = \sum_{r=1}^{k} W_r^2 V(\bar{x}_r)$$

$$= \sum_{r=1}^{k} W_r^2 \frac{N_r - n_r}{N_r - 1} \frac{\sigma_r^2}{n_r}$$ (நி-து.)

(இ) $$E^{-1}\left(\frac{N_r - n_r}{N_r - 1} \frac{\sigma_r^2}{n_r}\right) \sim \frac{N_r - n_r}{N_r n_r} s_r^2$$

$$\therefore \quad E^{-1}\{V[E^{-1}(\bar{X})]\} \sim \sum_{r=1}^{k} \frac{N_r - n_r}{N_r n_r} W_r^2 s_r^2$$ (நி-து.)

21·13. ஒவ்வொரு பகுதியிலும் ஈ. செயா. சா. ரா. மா. முறை கையாளப்படும் படுகை மாதிரி முறையில் சமவிகிதப் பங்கீடு

சம விகிதப் பங்கீட்டில் $n_r = \frac{N_r}{N} \cdot n$

$$\therefore \; V[E^{-1}(\bar{X}) = \sum_{r=1}^{k} W_r^2 \frac{N_r - \frac{N_r}{N}n}{N_r - 1} \frac{\sigma_r^2}{\frac{N_r}{N} \cdot n}$$

$$= \sum_{r=1}^{k} W_r^2 \frac{N-n}{N_r - 1} \frac{\sigma_r^2}{n}$$

$$= \frac{N-n}{n} \sum_{r=1}^{k} \frac{W_r^2 \sigma_r^2}{N_r - 1}$$

இதை $V_{\text{ச. வி. ப.}}$ எனக் குறிப்பிடுவோம்.

$$\therefore \; V_{\text{ச. வி. ப.}} = \frac{N-n}{n} \sum_{r=1}^{k} \frac{W_r^2 \sigma_r^2}{N_r - 1}$$

21·14. முழுமைத் தொகுதியைப் பகுதிகளாகப் பிரிக்காமல் நேராக n உறுப்புகள் கொண்ட ஈ. செயா. சா. ரா. மா. ஒன்றையெடுத்தால் மாதிரிச் சராசரியின் பரவற் படி.

$$V_{\text{சா. ரா. மா.}} = \frac{N-n}{N-1} \frac{\sigma^2}{n} = \frac{N-n}{(N-1)n} \left[\frac{1}{N} \sum_{r=1}^{k} N_r \sigma_r^2 + \frac{1}{N} \sum_{r=1}^{k} N_r (\bar{X}_r - \bar{X})^2 \right]$$

$$= \frac{N-n}{(N-1)\,n}\left[\sum_{r=1}^{k} W_r\,\sigma_r^2 + \sum_{r=1}^{k} W_r\,(\bar{X}_r - \bar{X})^2\right]$$

21·15. ஒவ்வொரு பகுதியிலும் ஈ. செயா. சா. ரா. மா. முறை கையாளப்படும் படுகை மாதிரி முறையில் உத்தமப் பங்கீடு

(i) குறிப்பிட்ட மாதிரிப் பருமன் n-க்கு மதிப்பீட்டியின் பரவற்படியைச் சிறுமமாக்கும் உத்தமப் பங்கீடு

$$V\,[\,E^{-1}(\bar{X})\,] = \sum_{r=1}^{k} W_r^2\,\frac{N_r - n_r}{N_r - 1}\,\frac{\sigma_r^2}{n_r}$$

$$= \sum_{r=1}^{k}\left(A_r + \frac{B_r}{n_r}\right)$$

$$A_r = -\,\frac{W_r^2}{N_r - 1}\,\sigma_r^2, \qquad B_r = \frac{W_r^2\,N_r}{N_r - 1}\,\sigma_r^2$$

நிபந்தனை $n_1 + n_2 + \cdots\cdots + n_k = n.$

$$F = \sum_{r=1}^{k}\left(A_r + \frac{B_r}{n_r}\right) + \lambda\,[\,n_1 + n_2 + \cdots\cdots + n_k - n\,].$$

$$\frac{\partial F}{\partial n_r} = -\,\frac{B_r}{n_r^2} + \lambda = 0; \quad r = 1, 2, \ldots, k.$$

$$\therefore\ n_r = \sqrt{\frac{B_r}{\lambda}}$$

$$\therefore\ n = \sum_{r=1}^{k} n_r = \sum_{r=1}^{k}\sqrt{\frac{B_r}{\lambda}} = \frac{1}{\sqrt{\lambda}}\sum_{r=1}^{k}\sqrt{B_r}$$

$$\therefore \quad \frac{1}{\sqrt{\lambda}} = \frac{n}{\sum_{r=1}^{k} \sqrt{B_r}}$$

$$\therefore \quad n_r = \sqrt{B_r} \cdot \frac{n}{\sum_{r=1}^{k} \sqrt{B_r}}$$

அதாவது $n_r = \dfrac{W_r \, \sigma_r \sqrt{\dfrac{N_r}{N_r - 1}}}{\sum_{r=1}^{k} W_r \, \sigma_r \sqrt{\dfrac{N_r}{N_r - 1}}} \cdot n$

அதாவது n_r-கள் $W_r \, \sigma_r \sqrt{\dfrac{N_r}{N_r - 1}}$ -களின் விகிதத்தில் இருக்கும்.

இங்கு $V [E^{-1} (\overline{X})]$ என்பதை $V_{\text{உ.த்.பங்.}}$ எனக் குறிப்பிட்டால்,

$$V_{\text{உ.த்.பங்.}} = \sum_{r=1}^{k} \left(A_r + \frac{B_r}{n_r} \right)$$

$$= \sum_{r=1}^{k} \left[A_r + \frac{B_r}{\dfrac{\sqrt{B_r} \cdot n}{\sum_{r=1}^{k} \sqrt{B_r}}} \right]$$

$$= \sum_{r=1}^{K} \left[A_r + \frac{\sqrt{B_r} \sum_{r=1}^{k} \sqrt{B_r}}{n} \right]$$

$$V_{\text{உத் . பங்}} = \sum_{r=1}^{k} A_r + \frac{1}{n} \left[\sum_{r=1}^{k} \sqrt{B_r} \right]^2$$

$$= - \sum_{r=1}^{k} \frac{W_r^2 \sigma_r^2}{N_r - 1}$$

$$+ \frac{1}{n} \left[\sum_{r=1}^{k} W_r \sigma_r \sqrt{\frac{N_r}{N_r - 1}} \right]^2$$

$$= \frac{1}{n} \left[\sum_{r=1}^{k} W_r \sigma_r \sqrt{\frac{N_r}{N_r - 1}} \right]^2$$

$$- \sum_{r=1}^{k} \frac{W_r^2 \sigma_r^2}{N_r - 1}$$

(ii) குறிப்பிட்ட செலவுக்கு மதிப்பீட்டியின் பரவற் படியைச் சிறுமமாக்கும் உத்தமப் பங்கீடு

செலவுச் சார்பு $C + C_1 n_1 + C_2 n_2 + \dots + C_k n_k$ எனவும்,

நிபந்தனையை $C_0 + C_1 n_1 + C_2 n_2 + \dots + C_k n_k = C'$ எனவும் கொள்வோம்.

$$\bar{V}[E^{-1}(\bar{X})] = \sum_{r=1}^{k} \left[A_r + \frac{B_r}{n_r} \right]$$

$$\therefore \quad F = \sum_{r=1}^{k} \left[A_r + \frac{B_r}{n_r} \right]$$

$$+ \lambda [C_0 + C_1 n_1 + \dots + C_k n_k - C']$$

$$\frac{\partial F}{\partial n_r} = - \frac{B_r}{n_r^2} + \lambda C_r = 0 ; \quad r = 1, 2, \dots, k.$$

$$\therefore \quad n_r = \sqrt{\frac{B_r}{\lambda C_r}}$$

$$C' = C_0 + \sum_{r=1}^{k} C_r n_r = C_0 + \sum_{r=1}^{k} C_r \frac{\sqrt{B_r}}{\sqrt{\lambda C_r}}$$

$$\therefore \quad \frac{1}{\sqrt{\lambda}} \sum_{r=1}^{k} C_r \frac{\sqrt{B_r}}{\sqrt{C_r}} = C' - C_0$$

$$\therefore \quad \frac{1}{\sqrt{\lambda}} = \frac{C' - C_0}{\sum_{r=1}^{k} C_r \frac{\sqrt{B_r}}{\sqrt{C_r}}} = \frac{C' - C_0}{\sum_{r=1}^{k} \sqrt{C_r B_r}}$$

$$\therefore \quad n_r = \frac{\sqrt{\frac{B_r}{C_r}} \; \frac{C' - C_0}{k}}{\sum_{r=1} \sqrt{C_r B_r}}$$

$$= \frac{C' - C_0}{\sqrt{C_r}} \; \frac{W_r \sigma_r \sqrt{\frac{N_r}{N_r - 1}}}{\sum_{r=1}^{k} \sqrt{C_r} W_r \sigma_r \sqrt{\frac{N_r}{N_r - 1}}}$$

$$\therefore \quad V[\overset{-1}{E} (\bar{X})] = \sum_{r=1}^{k} \left(A_r + \frac{B_r}{n_r} \right)$$

$$= - \sum_{r=1}^{k} \frac{W_r^2 \sigma_r^2}{N_r - 1} + \sum_{r=1}^{k} W_r^2 \sigma_r^2 \frac{N_r}{N_r - 1}$$

$$\cdot \frac{\sqrt{C_r} \sum_{r=1}^{k} \sqrt{C_r} W_r \sigma_r \sqrt{\frac{N_r}{N_r - 1}}}{(C' - C_0) \; W_r \sigma_r \sqrt{\frac{N_r}{N_r - 1}}}$$

$$= \frac{1}{(C' - C_0)} \left[\sum_{r=1}^{k} W_r \sigma \sqrt{C_r} \sqrt{\frac{N_r}{N_r - 1}} \right]$$

$$- \sum_{r=1}^{k} \frac{W_r^2 \sigma_r^2}{N_r - 1}$$

இந்தப் பங்கீட்டில் மாதிரிப் பருமன் n

$$n = \sum_{r=1}^{k} n_r = \frac{\sum_{r=1}^{k} \frac{(C'-C_0)}{\sqrt{C_r}} W_r \sigma_r \sqrt{\frac{N_r}{N_r-1}}}{\sum_{r=1}^{k} W_r \sigma_r \sqrt{C_r} \sqrt{\frac{N_r}{N_r-1}}}$$

$$= \frac{(C'-C_0) \sum_{r=1}^{k} W_r \sigma_r \frac{1}{\sqrt{C_r}} \sqrt{\frac{N_r}{N_r-1}}}{\sum_{r=1}^{k} W_r \sigma_r \sqrt{C_r} \sqrt{\frac{N_r}{N_r-1}}}$$

21·16. படுகை முறை மாதிரியில் முழுமைத் தொகுதிப் பாகத்தை (Population Proportion) மதிப்பிடுதல்

முழுமைத் தொகுதிப் பாகம் P எனவும் r ஆவது பகுதியின் பாகம் P_r எனவும் கொள்வோம். r ஆவது பகுதியிலிருந்து எடுக்கப்பட்ட மாதிரியின் பாகம் p_r எனவும் கொள்வோம்.

$$P = \sum_{r=1}^{k} W_r P_r$$ என்பது தெளிவு.

$$E^{-1}(P) \sim \sum_{r=1}^{k} W_r P_r$$ என்பதும் தெளிவு.

[பகுதிகளில் ஈ. செ. சா. ரா. மா. முறை அல்லது ஈ. செயா. சா. ரா. மா. முறை உபயோகப்படுத்தினால்]

(*i*) ஒவ்வொரு பகுதியிலும் ஈ. செ. சா. ரா. மா. முறையை உபயோகித்தால்

(அ) $V[\overset{-1}{E}(P)] = \sum_{r=1}^{k} W_r^2 \frac{P_r(1-P_r)}{n_r}$

(ஆ) இதன் பிறழ்ச்சியற்ற மதிப்பீட்டி $\overset{-1}{E}\{V[\overset{-1}{E}(P)]\}$

$$\sim \sum_{r=1}^{k} W_r^2 \frac{p_r(1-p_r)}{n_r-1}$$

(இ) குறிப்பிட்ட மாதிரிப் பருமனுக்குப் பரவற் படியைச் சிறுமமாக்கும் உத்தமப் பங்கீட்டில் $n_r = \frac{W_r \sqrt{P_r(1-P_r)}}{\sum_{r=1}^{k} W_r \sqrt{P_r(1-P_r)}}\, n.$

(ஈ) மேற் கூறப்பட்ட உத்தமப் பங்கீட்டில்,

$$V_{\text{உத்.பங்.}} = V[\overset{-1}{E}(P)] = \frac{1}{n}\left[\sum_{r=1}^{k} W_r \sqrt{P_r(1-P_r)}\right]^2$$

(உ) விகித சமப் பங்கீடு கையாளும்பொழுது,

$$V_{\text{வி. ச. ப.}} = V[\overset{-1}{E}(P)] = \frac{1}{n}\sum_{r=1}^{k} W_r P_r (1-P_r).$$

(ii) ஒவ்வொரு பகுதியிலும் ஈ. செயா. சா. ரா. மா. முறையைக் கையாளும் பொழுது,

(அ) $V[\overset{-1}{E}(P) = \sum_{r=1}^{k} W_r^2 \frac{N_r - n_r}{N_r - 1} \frac{P_r(1-P_r)}{n_r}.$

(ஆ) $\overset{-1}{E}\{V[\overset{-1}{E}(P)]\} \sim \frac{1}{N}\sum_{r=1}^{k} W_r \frac{N-n}{n_r - 1} p_r(1-p_r).$

மேற் கூறப்பட்ட ஏழு பண்புகளும் மிக எளிதாக நிறுவப்படக் கூடியவை. இவைகளை மாணவர்கள் நிறுவுவது ஒரு நல்ல பயிற்சியாகும். இதற்குப் படுகை மாதிரி முறையைத் துவக்கும்பொழுது கூறப்பட்டவைகளை நினைவு கூறவேண்டும்.

பயிற்சி 6

1. ஓர் அளவெடுப்பில் ஒரு மாநிலத்திலுள்ள கிராமங்களனைத்தும் கொண்ட முழுமைத் தொகுதி 4 பகுதிகளாகப் பிரிக்கப்பட்டது. ஒவ்வொரு பகுதியிலிருந்தும் 10 கிராமங்கள் ஈ.செ.சா.ரா. மா. முறையில் எடுக்கப்பட்டன. எடுக்கப்பட்ட மாதிரிக் கிராமங்களில் வீடுகளின் எண்ணிக்கை பின்வருமாறு:

பகுதி எண்	மொத்தக் கிராமங்கள்	மாதிரிக் கிராமங்களில் வீடுகளின் எண்ணிக்கை									
		1	2	3	4	5	6	7	8	9	10
1	1411	43	84	98	0	10	44	0	124	13	0
2	4705	50	147	62	87	84	158	170	104	56	160
3	2558	228	262	110	232	139	178	334	0	63	220
4	14997	17	34	25	34	36	0	25	7	15	31

(அ) மாநிலத்திலுள்ள மொத்த வீடுகளின் எண்ணிக்கையைப் பிறழ்ச்சியின்றி மதிப்பிடு. உன் மதிப்பீட்டியின் மாறுபாட்டுக் கெழுவைக் (coefficient of variation) கண்டுபிடி

(ஆ) பகுதியாகப் பிரித்ததனால் இலாபம் ஏற்பட்டிருக்கிறதா என்பதை ஈ.செ.சா.ரா மா. முறை (பகுதியாகப் பிரிக்கப்படாத) மாதிரியின் பரவற் படியைக் கண்டுபிடிப்பதன் மூலம் காண்.

(இ) மேலே கொடுக்கப்பட்ட பங்கீட்டை 40 உறுப்புகள் கொண்ட உத்தமப் பங்கீட்டுடன் ஒப்பிட்டு இதன் திறன் விகிதத்தைக் (Relative efficiency) கண்டுபிடி.

விடை

(அ) $E^{-1}(X) = 1353572$, இதன் மாறுபாட்டுக் கெழுவின் மதிப்பீடு = 9·12%

(ஆ) 407% (இ) 86·38%

2. ஒரு முன்னோடி அளவெடுப்பு (Pilot Survey) மூலம் மூன்று வகுப்பினராகப் பிரிக்கப்பட்ட ஒரு நகரத்தினரின் கல்வியறிவுச் சதவீதம் (Percentage literacy) பின்வருமாறு மதிப்பிடப்பட்டது.

வகுப்பு	மக்கட்தொகை	கல்வியறிவுடையோர் சதவீதம்
1	60000	40
2	10000	80
3	30000	60

(i) வகுப்புகளைப் பகுதிகளாகக் கருதி ஒவ்வொரு பகுதியிலும் ஈ.செ.சா.ரா.மா. முறை கையாளப்படுகிறது என்று எடுத்துக் கொண்டு, 2,000 பருமனுடைய மாதிரியின் உத்தமப் பங்கீட்டைக் கண்டுபிடி. [மதிப்பிடப்படவேண்டிய பண்பு முழுமைத் தொகுதியில் கல்வியறிவுடையோர் சதவீதம்.]

(ii) உத்தமப் பங்கீட்டின் திறனைப் பகுதியாகப் பிரிக்காத ஈ.செ.சா.ரா.மா. முறையின் திறனுடன் ஒப்பிடு.

விடை

(i) 1222, 167, 611 (ii) திறன் விகிதம் 108·1%

3. ஒரு நகரத்தின் வட்டத்தை இரண்டு பகுதிகளாகப் பிரிக்க நேர்ந்தது. இப்பகுதிகளில் 1000, 1500 பேர் வசிக்கின்றனர். இப்பகுதிகளில் முறையே 30%, 70% பேர் (தோராயமாக) இயந்திர உற்பத்தியில் ஈடுபட்டிருக்கின்றனர் என்பதை மதிப்பிடவேண்டும். மதிப்பீடு உண்மையான அளவிலிருந்து 5%-க்கு மேல்

வேறுபட்டிக்கக்கூடாது என்பதற்கு 95% நிகழ்திறன் இருக்க வேண்டுமானுல் மாதிரிப்பருமன் என்னவாக இருக்கவேண்டும் ? இரு பகுதிகளிலும் ஈ.செ.சா.ரா.மா.முறை கையாளப்படும் உத்தமப் பங்கீடு என்று எடுத்துக்கொள்ளவும்.

விடை

n என்பது மாதிரிப் பருமனுக இருக்கட்டும். (n_1, n_2) என்பது பங்கீடாக இருக்கட்டும். இங்கு $P_1 = 0{\cdot}3$, $P_2 = 0{\cdot}7$, $N_1 = 1000$, $N_2 = 1500$, $N = 2500$.

$$\text{மதிப்பீட்டி } x = W_1 p_1 + W_2 p_2$$
$$= \frac{1000}{2500} p_1 + \frac{1500}{2500} p_2 = \frac{2}{5} p_1 + \frac{3}{5} p_2$$

p_1, p_2 என்பவை மாதிரிப் பாகங்கள் (Proportions).

$$V(x) = \frac{4}{25} \frac{P_1(1 - P_1)}{n_1} + \frac{9}{25} \frac{P_2(1 - P_2)}{n_2}$$

$$= \frac{4}{25} \frac{0{\cdot}3 \times 0{\cdot}7}{n_1} + \frac{9}{25} \frac{0{\cdot}7 \times 0{\cdot}3}{n_2}$$

$$= \frac{0{\cdot}84}{25n_1} + \frac{1{\cdot}89}{25n_2} = \frac{1}{25}\left(\frac{0{\cdot}84}{n_1} + \frac{1{\cdot}89}{n_2}\right).$$

P என்பது உண்மையான முழுமைத் தொகுதிப் பாகமாயின் கணக்கின்படி

$$P\left[\ |x - P| \leq \frac{1}{20} P\ \right] = 0{\cdot}95.$$

$$\therefore\ P\left[\frac{|x - P|}{\sqrt{\frac{1}{25}\left(\frac{0{\cdot}84}{n_1} + \frac{1{\cdot}89}{n_2}\right)}} \leq \frac{\frac{1}{20} P}{\sqrt{\frac{1}{25}\left(\frac{0{\cdot}84}{n_1} + \frac{1{\cdot}89}{n_2}\right)}}\right] = 0{\cdot}95.$$

x-க்குத் தோராயமாக இயல் நிலைப் பரவலிருக்கிறது எனக் கொண்டால்,

$$\frac{P}{20} \quad \frac{1}{\sqrt{\frac{1}{25}\left(\frac{0\cdot84}{n_1}+\frac{1\cdot89}{n_2}\right)}} = 1\cdot96 \text{ என்பது விளங்கும்.}$$

அதாவது $$\frac{0\cdot84}{n_1}+\frac{1\cdot89}{n_2}=\frac{P^2}{16\times(1\cdot96)^2} \qquad \ldots I$$

உத்தமப் பங்கீடாகையால் $V(x)$ஐச் சிறுமமாக்கவேண்டும்.

$$F=\frac{1}{25}\left(\frac{0\cdot84}{n_1}+\frac{1\cdot89}{n_2}\right)+\lambda\,(n_1+n_2-n).$$

$$\frac{\partial F}{\partial n_1}=\frac{-0\cdot84}{25\,n_1^{\,2}}+\lambda=0. \qquad \therefore \quad n_1=\sqrt{\frac{0\cdot84}{25\,\lambda}}$$

$$\frac{\partial F}{\partial n_2}=\frac{-1\cdot89}{25\,n_2^{\,2}}+\lambda=0. \qquad \therefore \quad n_2=\sqrt{\frac{1\cdot89}{25\,\lambda}}$$

$$\therefore \quad n_1 : n_2 :: \sqrt{0\cdot84} : \sqrt{1\cdot89} = 2 : 3.$$

$$\therefore \quad n_2=\frac{3}{2}n_1$$

$$\therefore \quad I\text{-லிருந்து } \frac{0\cdot84}{n_1}+\frac{1\cdot89}{\frac{3}{2}n_1}=\frac{P^2}{16\times(1\cdot96)^2}$$

P என்பது தோராயமாக $$\frac{N_1P_1+N_2P_2}{N}=\frac{300+1050}{2500}$$

$$=\frac{1350}{2500}=0\cdot54$$

$$\therefore \quad \frac{1}{n_1}\,[0\cdot84+1\cdot26]=\frac{(0\cdot54)^2}{16\times(1\cdot96)^2}$$

$$\therefore \quad n_1 = \frac{2\cdot1 \times 16 \times (1\cdot96)^2}{(0\cdot54)^2} \doteq 443$$

$$n_2 = \frac{3}{2}\, n_1 \doteq 664$$

$$\therefore \quad n_{.} + n_2 = 444 + 664 = 1107$$

4. ஒவ்வொரு பகுதியிலுமுள்ள முழுமைத் தொகுதி உறுப்புகள் சம எண்ணிக்கைக் கொண்டவை. ஒவ்வொரு பகுதியிலும் ஈ. செ. சா. ரா. மா. முறை கையாளப்படும் படுகை மாதிரி முறையில் முழுமைத் தொகுதியின் சராசரி மதிப்பிடப்பட வேண்டும். ஒவ்வொரு பகுதியிலும் மாதிரிப் பருமன் சமமானவை எனக் கொண்டால் மதிப்பீட்டியின் பரவற் படி V_2 ஐப் பின் வருமாறு எழுத முடியும் என்பதை நிறுவு.

$$V_2 = V_1 - \frac{1}{k_n} \sum_{r=1}^{k} (\bar{X}_r - \bar{X})^2$$

V_1 என்பது பகுதியாகப் பிரிக்கப்படாத ஈ.செ சா.ரா மா-ன் சராசரியின் பரவற் படி.

நிறுவல்

$N_1 = N_2 = \dots = N_k$ என்பதால்

$$W_1 = W_2 = \dots W_k = \frac{1}{k} \,.$$

$n_1 = n_2 = \dots = n_k$ என்பதால் ஒவ்வொன்றும் $\frac{n}{k}$-க்குச் சமமாகும்.

$$V_1 = \frac{1}{n} \left[\sum_{r=1}^{k} W_r \,\sigma_r^2 + \sum_{r=1}^{k} W_r \,(\bar{X}_r - \bar{X})^2 \right]$$

$$= \frac{1}{kn}\left[\sum_{r=1}^{k} \sigma_r^2 + \sum_{r=1}^{k} (\bar{X}_r - \bar{X})^2\right]$$

$$V_2 = \sum_{r=1}^{k} W_r^2 \frac{\sigma_r^2}{n_r} = \frac{1}{k^2 . \frac{n}{k}} \sum_{r=1}^{k} \sigma_r^2 = \frac{1}{kn}\sum_{r=1}^{k} \sigma_r^2 .$$

$$\therefore \quad V_1 - V_2 = \frac{1}{kn}\sum_{r=1}^{k} (\bar{X}_r - \bar{X})^2$$

$$\therefore \quad V_2 = V_1 - \frac{1}{kn}\sum_{r=1}^{k} (\bar{X}_r - \bar{X})^2 \qquad \text{(நி-து.)}$$

5. இரு கிராமங்களில் ஒரு குறிப்பிட்ட நோயினால் பாதிக்கப் பட்டவர்களின் எண்ணிக்கையின் வித்தியாசம் தேவைப்படுகிறது. முதல் கிராமத்தில் N_1 பேர்கள். அவர்களுள் P_1 பாகம் நோயினால் பாதிக்கப்பட்டவர்கள். இரண்டாம் கிராமத்தில் N_2 பேர்கள். அவர் களுள் P_2 பாகம் நோயினால் பாதிக்கப்பட்டவர்கள். ஒருவரை வைத்தியப் பரிசோதனை செய்வதற்கு முதல் கிராமத்தில் C_1 ரூபாய்களும், இரண்டாம் கிராமத்தில் C_2 ரூபாய்களும் செலவாகிறது. மொத்தச் செலவு C' ரூபாய்கள் தான் ஆக வேண்டும் என்ற நிபந்தனையின் பேரில் உத்தமப் பங்கீட்டினைக் கண்டுபிடி. ஒவ்வொரு கிராமத்திலும் ஈ.செ சா.ரா மா. முறை கையாளப்படுகிறது.

விடை

மொத்தச் செலவு $C' = C_1 n_1 + C_2 n_2$.

$$V[\overset{-1}{E}(P)] = W_1^2 \frac{P_1(1-P_1)}{n_1} + W_2^2 \frac{P_2(1-P_2)}{n_2}$$

[P என்பது பாகங்களின் வித்தியாசம்].

W_1, W_2 என்பவை $\frac{N_1}{N}$, $\frac{N_2}{N}$ அல்ல.

$$F = \frac{W_1^2 P_1' (1 - P_1)}{n_1} + \frac{W_2^2 P_2 (1 - P_2)}{n_2}$$

$$+ \lambda [C_1 n_1 + C_2 n_2 - C'].$$

$$\frac{\partial F}{\partial n_1} = \frac{- W_1^2 P_1 (1 - P_1)}{n_1^2} + \lambda C_1 = 0.$$

$$\therefore \quad n_1 = | W_1 | \sqrt{\frac{P_1 (1 - P_1)}{\lambda C_1}}$$

இதைப்போலவே $n_2 = | W_1 | \sqrt{\frac{P_2 (1 - P_2)}{\lambda C_2}}$

$$C' = C_1 n_1 + C_2 n_2 = | W_1 | \sqrt{\frac{P_1 (1 - P_1) C_1}{\lambda}}$$

$$+ | W_2 | \sqrt{\frac{P_2 (1 - P_2) C_2}{\lambda}}$$

$$\therefore \quad \frac{1}{\sqrt{\lambda}} = \frac{C'}{| W_1 | \sqrt{C_1 P_1 (1 - P_1)} + | W_2 | \sqrt{C_2 P_2 (1 - P_2)}}$$

$$\therefore \quad n_1 = | W_1 | \sqrt{\frac{P_1 (1 - P_1)}{C_1}}$$

$$\cdot \frac{C'}{| W_1 | \sqrt{C_1 P_1 (1 - P_1)} + | W_2 | \sqrt{C_2 P_2 (1 - P_2)}}$$

$$n_2 = | W_2 | \sqrt{\frac{P_2 (1 - P_2)}{C_2}}$$

$$\cdot \frac{C'}{| W_1 | \sqrt{C_1 P_1 (1 - P_1)} + | W_2 | \sqrt{C_2 P_2 (1 - P_2}}$$

இங்கு $W_1 = 1$, $W_2 = -1$ என்று எடுத்துக்கொள்ள வேண்டும்.

ஏனெனில் $E^{-1}(P) = p_1 - p_2 = W_1 p_1 + W_2 p_2$.

$$\therefore \quad n_1 = \sqrt{\frac{P_1(1-P_1)}{C_1}} \cdot \frac{C'}{\sqrt{C_1 P_1 (1-P_1)} + \sqrt{C_2 P_2 (1-P_2)}}$$

$$n_2 = \sqrt{\frac{P_2(1-P_2)}{C_2}} \cdot \frac{C'}{\sqrt{C_1 P_1 (1-P_1)} + \sqrt{C P_2 (1-P_2)}}$$

6. t_{s_1}, t_{s_2} என்பவை s ஆவது பகுதியின் மொத்தத்திற்குப் பிறழ்ச்சியற்ற சார்பற்ற மதிப்பீட்டிகள். பின்வரும் சார்புகள் $V\left[\sum_{s=1}^{k}\left(\frac{t_{s_1}+t_{s_2}}{1}\right)\right]$ என்பதன் பிறழ்ச்சியற்ற மதிப்பீட்டிகள் என்பதை நிறுவு. அவைகளின் பரவற் படியை ஒப்பிடு.

$$\text{(i)} \quad v_1 = \sum_{s=1}^{k} \frac{(t_{s_1} - t_{s_2})^2}{4}$$

$$\text{(ii)} \quad v_2 = \frac{\left[\sum_{s=1}^{k} t_{s_1} - \sum_{s=1}^{k} t_{s_2}\right]^2}{4}$$

[Murthy, M.N., Sankhya, 24, (*B*), (1962)].

விடை

$$E(t_{s_1}) = E(t_{s_2}) = \overline{X}_s$$

$$V\left[\sum_{s=1}^{k}\left(\frac{{}^{t}s_1 + {}^{t}s_2}{1}\right)\right] = \frac{1}{4}\sum_{s=1}^{k}[V({}^{t}s_1) + V({}^{t}s_2)]$$

[$\therefore$ ${}^{t}s_1, {}^{t}s_2$ சார்பற்றவை.]

$$E(\nu_1) = \frac{1}{4}\sum_{s=1}^{k} E({}^{t}s_1 - \bar{X}_s + \bar{X}_s - {}^{t}s_2)^2$$

$$= \frac{1}{4}\sum_{s=1}^{k}[E({}^{t}s_1 - \bar{X}_s)^2 + E({}^{t}s_2 - \bar{X}_s)^2]$$

$$= \frac{1}{4}\sum_{s=1}^{k}[V({}^{t}s_1) + V({}^{t}s_2)] \qquad \text{(நி-து)}$$

$$E(\nu_2) = \frac{1}{4} E\left[\sum_{s=1}^{k}({}^{t}s_1 - \bar{X}_s) - \sum_{s=1}^{k}({}^{t}s_2 - \bar{X}_s)^2\right]$$

$$= \frac{1}{4}\sum_{s=1}^{k}[E({}^{t}s_1 - \bar{X}_s)^2 + E({}^{t}s_2 - \bar{X}_s)^2]$$

$$= \frac{1}{4}\sum_{s=1}^{k}[V({}^{t}s_1) + V({}^{t}s_2)] \qquad \text{(நி-து.)}$$

7. ஒரு முழுமைத் தொகுதி N_1, N_2 உறுப்புகளுள்ள இரு பகுதிகளாகப் பிரிக்கப்பட்டது. n_1, n_2 உறுப்புகள் உள்ள மாதிரிகள், இவைகளிலிருந்து ஈ. செ. சா. ரா. மா. முறையில் எடுக்கப்பட்டன. முழுமைத் தொகுதியின் சராசரியை மதிப்பிடுவதில் இம் மதிப்பீட்டியின் திறன் விகிதம் $(n_1 + n_2)$ உறுப்புகளுள்

உத்தமப் பங்கீட்டு முறை கையாளப்பட்ட மதிப்பீட்டியைப் பொறுத்தவரை $\frac{4t}{1+t^2}$-க்குக் குறையாது என நிறுவு.

$t = \frac{n_1 n_2'}{n_2 n_1'}$ ...,n_1', n_2' என்பவை உத்தமப் பங்கீட்டு மாதிரிப் பருமன்களைக் குறிக்கும்.

[Cochran, W.G., Sampling Technique, (1953)].

$$V_{(n_1, n_2)} = \frac{W_1^2 \sigma_1^2}{n_1} + \frac{W_2^2 \sigma_2^2}{n_2}$$

$$V_{(n_1', n_2')} = \frac{1}{n_1 + n_2}(W_1 \sigma_1 + W_2 \sigma_2)^2.$$

$$\frac{V_{(n_1', n_2')}}{V_{(n_1, n_2)}} = \frac{\frac{1}{n_1 + n_2}(W_1 \sigma_1 + W_2 \sigma_2)^2}{\frac{W_1^2 \sigma_1^2}{n_1} + \frac{W_2^2 \sigma_2^2}{n^2}}$$

$$n_1' = (n_1 + n_2)\frac{W_1 \sigma_1}{W_1 \sigma_1 + W_2 \sigma_2}; \; n_2' = \frac{(n_1 + n_2) W_2 \sigma_2}{W_1 \sigma_1 + W_2 \sigma_2}$$

$$\therefore \; \frac{n_2'}{n_1'} = \frac{W_2 \sigma_2}{W_1 \sigma_1}.$$

$\therefore$ நிறுவ வேண்டியது $\frac{1}{n_1 + n_2} \quad \frac{(W_1 \sigma_1 + W_2 \sigma_2)^2}{\frac{W_1^2 \sigma_1^2}{n_1} + \frac{W_2 \sigma_2^2}{n^2}}$

$$\geq \frac{4 \frac{n_1}{n_2} \frac{W_2 \sigma_2}{W_1 \sigma_1}}{\left(\frac{n_1}{2} \frac{W_2 \sigma_2}{W_1 \sigma_1} \right)^2}$$

அல்லது $$\frac{1}{n_1+n_2}\;\frac{(n_1+n_2)^2}{\frac{n_1'^2}{n_1}+\frac{n_2'^2}{n^2}} \geq \frac{4\frac{n_2\,n_2'}{n_1\,n_1'}}{\left(\frac{n_2}{n_1}+\frac{n_2'}{n_1'}\right)^2}$$

அதாவது $$\frac{(n_1+n_2)}{\frac{n_1^{\,2}}{n_1}+\frac{r_2'^2}{n_2}} \geq \frac{4\frac{n_2 n_2'}{n_1 n_1'}}{\left(\frac{n_2}{n_1}+\frac{n_2'}{n_1'}\right)^2}$$

$\frac{n_2}{n_1} = x$ எனவும், $\frac{n_2'}{n_1'} = y$ எனவும் எடுத்துக் கொண்டால் இதை எளிதாக நிறுவலாம்.

8. ஈ.செயா.சா.ரா.மா. முறை கையாளப்படும் படுகை மாதிரி முறையில் முழுமைத் தொகுதியின் சராசரி மதிப்பிடப்பட வேண்டியிருக்கிறது. மதிப்பீட்டியின் பரவற் படிகளைப் பின்கண்ட நிலைமைகளில் கண்டுபிடி.

(i) விகிதசமப் பங்கீடு (ii) n உறுப்புகள் கொண்ட உத்தமப் பங்கீடு. இங்கு $\frac{N_r}{N_r-1} \doteq 1$ எனக் கொள்ளவும்.

இவைகளுடன் பகுதிகளாகப் பிரிக்கப்படாத n உறுப்புகள் கொண்ட ஈ.செயா.சா.ரா.மா.-ன் சராசரியின் பரவற் படியை ஒப்பிடு.

விடை

$$\text{(i)}\quad V\,[\,\overset{-1}{E}\,(\bar{X})\,] = \sum_{r=1}^{k} W_r^2\,\frac{\sigma_r^{\,2}}{n_r}\cdot\frac{N_r-n_r}{N_r-1}$$

இங்கு விகிதசமப் பங்கீடாதலால் இது $\frac{N-n}{n}\sum_{r=1}^{k}\frac{W_r^2\,\sigma_r^2}{N_r-1}$

என்றாகும்.

$N_r - 1 \doteq N_r$ எனக் கொண்டால் இது $\frac{N-n}{n} \sum_{r=1}^{k} W_r^2 \sigma_r^2$

என்றாகும்.

$$\therefore \quad V_{\text{ச.வி.ப}} = \frac{N-n}{Nn} \sum_{r=1}^{k} W_r^2 \sigma_r^2$$

(ii) உத்தமப் பங்கீட்டில்

$$V_{\text{உத். பங்}} \doteq \frac{1}{n} \left[\sum_{r=1}^{k} W_r \sigma_r \right]^2 - \frac{1}{N} \sum_{r=1}^{k} W_r \sigma_r^2$$

$$V_{\text{சா.ரா.மா.}} = \frac{N-n}{N-1} \frac{1}{n} \left[\sum_{r=1}^{k} W_r \sigma_r^2 + \sum_{r=1}^{k} W_r (\overline{X}_r - \overline{X})^2 \right]$$

9. ஒரு முழுமைத் தொகுதி உத்தம முறையில் இரு பகுதி களாகப்பிரிக்கப்பட்டிருக்கிறது. அதாவது மதிப்பீட்டியின் பரவற் படி சிறுமமாக இருக்குமாறு ஒரு குறிப்பிட்ட இடத்தில் (y) ஏறு வரிசையில் அமைக்கப் பெற்ற முழுமைத் தொகுதி பிரிக்கப்பட்டி ருக்கிறது. இரண்டாவது பகுதியில் N_2 பெரிய உறுப்புகளும், முதல் பகுதியில் N_1 சிறிய உறுப்புகளும் உள்ளன. மாதிரியில் இரண்டாம் பகுதி உறுப்புகள் யாவையும் எடுக்கப்பட்டன. முதல் பகுதியிலிருந்து $n_1 = n - N_2$ உறுப்புகள் எடுக்கப்பட்டன. (ஈ. செயா. சா. ரா. மா. முறையில்)

$$y = Y_1 + \sigma_1 \sqrt{\frac{N_1}{n_1}}$$ என நிறுவு.

$\overline{Y}_1, \sigma_1^2$ என்பவை முதல் பகுதியின் சராசரி, பரவற் படியாகும்.

(Dalenius, T., Skand Akt., 35, (1952)]

10. y என்னும் ராண்டம் மாறியின் ஊக அளவு அடர்த்தி (Probability density) $f(y) = e^{-y} \qquad y > 0$

இது $(0, y_0]$, (y_0, ∞) என்று இரண்டு பகுதிகளாகப் பிரிக்கப்பட்டது. ஒவ்வொரு பகுதியிலும் ஈ. செ. சா. ரா. மா. முறை கையாளப்படும் சமவிகிதப் பங்கீட்டில் முழுமைத் தொகுதியின் சராசரியின் மதிப்பீட்டியின் பரவற்படியைக் கண்டுபிடி. இப் பரவற்படி சிறுமமாக இருக்கவேண்டுமானால் y_0 என்ன மதிப்புடையதாக இருக்கும்? இச்சிறும பரவற் படியைக் கணக்கிடு.

விடை

$$N_1 = N \int_0^{y_0} e^{-y}\, dy = N(1-e^{-y_0})$$

$$N_2 = N - N_1 = N[1 - 1 + e^{-y_0}] = N e^{-y_0}$$

சம விகிதப் பங்கீட்டில் $n_1 = \frac{N_1}{N} n = n(1 - e^{-y_0})$

$$n_2 = \frac{N_2}{N} n = n e^{-y_0}$$

$$\overline{Y} = \int_0^{\infty} y e^{-y}\, dy = \Gamma(2) = 1$$

$$\sigma^2 = V(Y) = \int_0^{\infty} y^2 e^{-y}\, dy - \overline{Y}^2 = \Gamma(3) - 1 = 2 - 1 = 1$$

$$\overline{Y}_2 = \int_{y_0}^{\infty} y \frac{e^{-y}}{e^{-y_0}} dy = \int_{y_0}^{\infty} y\, e^{-(y-y_0)} dy = \int_0^{\infty} (y_0+t)\, e^{-t} dt$$

$$t = y - y_0$$

$$= y_0 + 1$$

$$\sigma_2 = \int_{y_0}^{\infty} y^2 \frac{e^{-y}}{e^{-y_0}} dy - (1+y_0)^2 = \int_0^{\infty} (t + y_0)^2 e^{-t} dt - (1 + y_0)^2$$

$$= \Gamma(3) + 2 y_0\, \Gamma(2) + y_0^2\, \Gamma(1) - (1 + y_0)^2$$

$$= 2 + 2 y_0 + y_0^2 - (1 + y_0)^2$$

$$= 1.$$

ஆனல் $\overline{Y} = \dfrac{N_1 \overline{Y}_1 + N_2 \overline{Y}_2}{N}$ $\quad \therefore \overline{Y}_1 = \dfrac{N\overline{Y} - N_2 \overline{Y}_2}{N_1}$

$$\therefore \overline{Y}_1 = \frac{N - N_2 (1+y_0)}{N_1} = \frac{N_1 - N_2 y_0}{N_1} = 1 - \frac{N_2}{N_1} y_0$$

$$\therefore \overline{Y}_1 = 1 - \frac{e^{-y_0}}{1-e^{-y_0}} y_0$$

$$\sigma^2 = 1 = \frac{N_1 \sigma_1^2 + N_2 \sigma_2^2}{N} + \frac{N_1 (\overline{Y}_1 - \overline{Y})^2 + N_2 (\overline{Y}_2 - \overline{Y})^2}{N}$$

$$\therefore \sigma_1^2 = \frac{N - N_2 \sigma_2^2 - N_1 (\overline{Y}_1 - \overline{Y}) - N_2 (\overline{Y}_2 - \overline{Y})^2}{N_1}$$

$$= \frac{N_1 - N_1 \left[\dfrac{e^{-y_0}}{1-e^{-y_0}} y_0 \right]^2 - N_2 y_0^2}{N_1}$$

$$= 1 - \frac{y_0^2 e^{-2y_0}}{(1-e^{-y_0})^2} - \frac{e^{-y_0}}{1-e^{-y_0}} y_0^2$$

$$V_{\text{ச·வி·ப}} = \frac{W_1^2 \sigma_1^2}{n_1} + \frac{W_2^2 \sigma_2^2}{n_2} = \frac{W_1^2 \sigma_1^2}{W_1 n} + \frac{W_2^2 \sigma_2^2}{W_2 n}$$

$$= \frac{1}{n} [W_1 \sigma_1^2 + W_2 \sigma_2^2]$$

$$= \frac{1}{n} [(1-e^{-y_0}) \sigma_1^2 + e^{-y_0} \sigma_2^2]$$

$$= \frac{1}{n} \left[(1-e^{-y_0}) - \frac{y_0^2 (e^{-2y_0})}{1-e^{-y_0}} - e^{-y_0} y_0^2 + e^{-y_0} \right]$$

$$\frac{1}{n(1-e^{-y_0})} [1 + e^{-2y_0} - 2e^{-y_0} - e^{-2y_0} y_0^2 - e^{-y_0} y_0^2$$
$$+ e^{-2y_0} y_0^2 + e^{-y_0} - e^{-2y_0}]$$

$$= \frac{1}{n(1-e^{-y_0})} [1 - e^{-y_0}(1+y_0^2)] = \frac{1}{n} - \frac{e^{-y_0} y_0^2}{n(1-e^{-y_0})}$$

(நி—து.)

ச. வி. ப. சிறுமமாக இருக்கவேண்டுமாயின்

$$\frac{d}{dy_0} [V_{\text{ச. வி. ப.}}] = 0. \quad \therefore \frac{d}{dy_0} \left[\frac{1}{n} - \frac{1}{n} \frac{e^{-y_0} y_0^2}{1-e^{-y_0}} \right] = 0.$$

$$\therefore \frac{d}{dy_0} \left[\frac{e^{-y_0} y_0^2}{1-e^{-y_0}} \right] = 0.$$

$$\therefore (1-e^{-y_0}) [2y_0 e^{-y_0} - y_0^2 e^{-y_0}] - e^{-y_0} y_0^2 \cdot e^{-y_0} = 0.$$

$$\therefore (1-e^{-y_0}) (2y_0 - y_0^2) - y_0^2 e^{-y_0} = 0.$$

$$\therefore (1-e^{-y_0}) (2-y_0) - y_0 e^{-y_0} = 0.$$

$$\therefore \quad (e^{y_0} - 1)(2 - y_0) - y_0 = 0.$$

$$\therefore \quad e^{y_0} = 1 + \frac{y_0}{2 - y_0} = \frac{2}{2 - y_0}.$$

$$\therefore \quad e^{y_0} = \frac{2}{2 - y_0}.$$

இந்தச் சமன்பாடு தோராயமாகத்தான் தீர்க்கப்பட முடியும்.

$e^{y_0} > 0$ என்பதால் $y_0 < 2$ என்பது தெரிகிறது.

$9(y_0) = e^{y_0} - \dfrac{2}{2 - y_0}$ என்று எடுத்துக்கொண்டால்

$9(1) = e - 2 > 0.$

$9(2) = -\infty.$

$\therefore \quad 1 < y_0 < 2$ என்பது தெரிகிறது.

$y_0 = 1 + x$ என எடுத்துக்கொண்டால் $0 < x < 1.$

$$\therefore \quad e^{(1+x)} = \frac{2}{1 - x} = 2[1 + x + x^2 + \cdots].$$

$$\therefore \quad e\left[1 + x + \frac{x^2}{2} + \frac{x^3}{6} + \cdots\right] = 2(1 + x + x^2 \cdots).$$

x^3 என்பது மிகவும் சிறியது என எடுத்துக்கொண்டால்

$$e\left[1 + x + \frac{x^2}{2}\right] = 2 + 2x + 2x^2.$$

$$\therefore \quad x^2\left(\frac{e}{2} - 2\right) + x(e - 2) + (e - 2) = 0,$$

இதிலிருந்து $x \doteq 0{\cdot}6$ என்று வரும்.

$$\therefore \quad y_0 \doteq 1{\cdot}6.$$

22. திரள் மாதிரி முறை

(CLUSTER SAMPLING)

இம்முறையில் முழுமைத் தொகுதி பல பகுதிகளாகப் (திரள்களாகப்) பிரிக்கப்பட்டு ஒரு சில பகுதிகள் முற்றிலும் அளவெடுக்கப்படுகின்றன. இது படுகை மாதிரி முறையினின்றும் அதிகம் வேறுபட்டது. படுகை மாதிரி முறையிலும் முழுமைத் தொகுதி பல பகுதிகளாகப் பிரிக்கப்படுகிறது. ஆனால் இம்முறையில் ஒவ்வொரு பகுதியிலிருந்தும் ஒரு மாதிரி எடுக்கப்படுகிறது. எப்பகுதியும் முற்றிலும் ஒதுக்கப்படுவதில்லை. மேலும் பொதுவாக ஒரு பகுதியிலுள்ள எல்லா உறுப்புகளும் மாதிரியில் சேர்த்துக் கொள்ளப்படுவதில்லை. ஆனால் திரள் மாதிரி முறையில் ஒரு சில பகுதிகளே மாதிரியில் சேர்த்துக்கொள்ளப்படுகின்றன. எடுக்கப்பட்ட பகுதியிலுள்ள எல்லா உறுப்புகளும் மாதிரியில் சேர்த்துக் கொள்ளப்படுகின்றன.

முழுமைத் தொகுதியிலுள்ள எல்லாப் பகுதிகளிலிருந்தும் தி. மா. முறையில் உறுப்புகள் எடுத்துக்கொள்ளப் படுவதில்லையாதலால், எடுக்கப்பட்ட மாதிரி உண்மையில் முழுமைத் தொகுதியைப் பிரதிபலிப்பதாக இராதோ என்ற ஐயம் எழலாம். இது ஓரளவு நியாயமான ஐயமும் கூட. ஆனால் இது முழுமைத் தொகுதி எவ்விதத்தில் திரள் படுத்தப்பட்டிருக்கிறது என்பதைப் பொறுத்தே இருக்கிறது. முழுமைத் தொகுதியிலுள்ள எல்லாத் திரள்களும் எல்லா விதத்திலும் சமமாக இருந்தால் தி. மா. முறை சிறந்த பயனை அளிக்கும் என்பது எளிதில் விளங்கும். ஆனால் திரள்களுக்கிடையே அதிக வேறுபாடு இருப்பின் இம்முறையில் எடுக்கப்படும் மாதிரி முழுமைத் தொகுதியைச் சரியாகப் பிரதிபலிக்காது. ஆகவே மதிப்பீட்டியும் அதிகப் பிழையுள்ளதாக இருக்கும்.

ஆனால் திரள் மாதிரி முறை நடைமுறையில் மிகவும் சௌகரியமானது. பொதுவாக அருகருகே உள்ள உறுப்புகள் ஒரே திரளில் வருமாகையால் அளவெடுப்பு எளிதாகிறது. ஏனெனில் உறுப்புகள் பல்வேறு இடங்களில் பரவலாக இருந்தால் அளவெடுப்பது அசௌகரியமானது. பொதுவாகத் திரள் மாதிரி முறை சாதாரண ராண்டம் மாதிரி முறையை விடத் திறன் குறைந்த தாகவே இருக்கும். ஆனால் நடைமுறை சௌகரியத்திற்காகவும், செலவு குறைவாகவே ஆகும் என்ற காரணத்திற்காகவும் திறன் குறைவு என்பதைப் பொருட்படுத்தாது இம்முறையைக் கையாளுகின்றனர் புள்ளியியலாளர்கள்.

சாதாரணமாகத் திரள்கள் ஒன்றையொன்று விலக்கும் (mutually exclusive) தன்மையுடையனவாகவே கருதப்படுகின்றன அதாவது எந்த இரண்டு திரள்களை எடுத்துக்கொண்டாலும் அவைகளுக்குப் பொதுவான உறுப்பு ஒன்று கூட இராது. ஆனால் திரள்கள் ஒன்றையொன்று விலக்காமலும் இருக்கலாம் என்று கருதினால் எல்லா மாதிரி முறைகளும் திரள் மாதிரி முறையில் அடங்கிவிடும். ஏனெனில் ஒரு குறிப்பிட்ட மாதிரி முறையில் வரக்கூடிய எல்லா மாதிரிகளையும் திரள்களாகக் கருதுவோமேயானால் இம் மாதிரிமுறை திரள் முறையாகிவிடும், எடுத்துக்காட்டாக N உறுப்புகள் உள்ள முழுமைத் தொகுதியில் n வெவ்வேறு உறுப்புகளைக் கொண்ட $\binom{N}{n}$ மாதிரிகளையும் திரள் களாகக் கொண்டு ஒரு திரளை $\frac{1}{\binom{N}{n}}$ என்னும் நிகழ்திறனுடன் எடுத்தால் இம்முறை ஈ. செயா. சா. ரா. மா. முறையாகிவிடும். முறையுடை மாதிரி முறையையும் இவ்வாறே திரள் மாதிரி முறையாகக் கருதலாம்.

ஆனால் பொதுவாக அருகருகே உள்ள உறுப்புகளை இணைத்துத் திரளாக்கி, திரள்களும் ஒன்றையொன்று விலக்குவனவாக இருந்தால் மட்டுமே மாதிரி முறை திரள் மாதிரி முறையென்று நாம் கொள்வோம். இவ்வாறு கருதும் பட்சத்தில் எல்லா மாதிரி முறைகளும் திரள் மாதிரி முறைதான் என்று கூறுவது பொருத்தமுடையதாகாது.

திரள் ஒவ்வொன்றையும் ஒருறுப்பு எனக் கருதி நாம் சில திரள்களை எடுக்க (அ) ஈ. செ. சா. ரா. மா முறை (ஆ) ஈ. செயா.

சா. ரா. மா. முறை (இ) முறையுடை மாதிரி முறை (ஈ) படுகை முறை மாதிரிமுறை (உ) சம வாய்ப்பிலா மாதிரி முறை முதலிய எல்லா முறைகளையும் கையாளலாம். இப்படிச்செய்வது இரு கட்ட மாதிரி முறை (Two stage Sampling) போன்று இருக்கும். முதல் கட்ட உறுப்புகள் திரள்கள். இரண்டாம் கட்ட உறுப்புகள் திரள்களிலுள்ள உறுப்புகள் எனக் கொள்ளலாம். பல கட்ட மாதிரி முறையை நாம் பிறகு பார்க்கலாம்.

22·1. சம பருமன் உள்ள திரள்கள்

திரள் மாதிரி முறையில் எல்லாத் திரள்களும் சம பருமன் உள்ளவையாக இருக்கவேண்டிய கட்டாயமில்லை. ஆனால் சம பருமன் உடையவையாக இருந்தால் கணக்கீடுகள் எளியனவாக இருக்கும். ஆகவே முதலில் சம பருமன் உள்ள திரள்களையே நாம் கவனிப்போம்.

முழுமைத் தொகுதியில் k திரள்கள் உள்ளன என்றும் ஒவ்வொரு திரளிலும் M உறுப்புகள் உள்ளன என்றும் கொள்வோம். முழுமைத் தொகுதியில் மொத்தம் $N = kM$ உறுப்புகள் இருக்கும். r ஆவது திரளில் உள்ள உறுப்புகள் $X_{r1}, X_{r2}, \ldots, X_{rM}$ எனக் கொள்வோம். r ஆவது திரளின் சராசரி

$$\overline{X}_r = \frac{1}{M}\sum_{i=1}^{M} X_{ri}\text{, பரவற் படி } \sigma_r^2 = \frac{1}{M}\sum_{i=1}^{M}(X_{ri} - \overline{X}_r)^2$$

$$\text{முழுமைத் தொகுதியின் சராசரி } \overline{X} = \frac{1}{k}\sum_{r=1}^{k}\overline{X}_r$$

$$\text{முழுமைத் தொகுதியின் பரவற் படி } \sigma^2 = \frac{1}{k}\sum_{r=1}^{k}\sigma_r^2$$

$$+\frac{1}{k}\sum_{r=1}^{k}(\overline{X}_r - \overline{X})^2$$

முழுமைத் தொகுதியிலிருந்து சா. ரா. மா. முறையில் ஒரே ஒரு திரளை எடுத்தால் அதன் சராசரியை $\bar{x}$ என்றும் மாதிரிப் பரவற்படியை s^2 என்றும் எழுதலாம்.

22·2. தேற்றம்

திரளின் சராசரி முழுமைத் தொகுதியின் சராசரிக்குப் பிறழ்ச்சியற்ற மதிப்பீட்டியாகும். [திரள் சராசரியின் பரவற் படி

$$\frac{1}{k}\sum_{r=1}^{k}(\bar{X}_r-\bar{X})^2$$ ஆகும்.

நிறுவல்

திரள் சராசரி $\bar{x}$-ன் பரவல் பின்வருமாறு :

$\bar{x}$	$\bar{X}_1$	$\bar{X}_2$	$\bar{X}$	... ...	$\bar{X}_k$
P	$\frac{1}{k}$	$\frac{1}{\kappa}$	$\frac{1}{k}$	... ...	$\frac{1}{k}$

$$\therefore\quad E(\bar{x})=\sum_{r=1}^{k}\bar{X}_r\cdot\frac{1}{k}=\frac{1}{k}\sum_{r=1}^{k}\bar{X}_r=\bar{X}$$

$$\therefore\quad \overset{-1}{E}\,(\bar{X})\sim\bar{x}$$ (நி-து.)

$$E(\bar{x}^2)=\frac{1}{k}\sum_{r=1}^{k}\bar{X}_r^2$$

$$\therefore\quad V(\bar{x})=E(\bar{x}^2)-[E(\bar{x})]^2=\frac{1}{k}\sum_{r=1}^{k}\bar{X}_r^2-\bar{X}^2$$

$$= \frac{1}{k} \sum_{r=1}^{k} (\bar{X}_r - \bar{X})^2 \qquad \text{(நி-து.)}$$

$\frac{1}{k} \sum_{r=1}^{k} (\bar{X}_r - \bar{X})^2$ என்பதை $\sigma^2_{\text{தி. இடை}}$ எனக் குறிப்பிடலாம்.

இது திரள்களுக்கு இடையிலுள்ள பரவற் படியாகும்.

கிளைத் தேற்றம்

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

s^2-ன் பரவல் பின்வருமாறு :

$$S_r^2 = \frac{1}{M-1} \sum_{i=1}^{M} (X_{ri} - \bar{X}_r)^2 \text{ எனக் கொள்க.}$$

s^2	S_1^2	S_2^2	S_3^2	....	S_k^2
P	$\frac{1}{k}$	$\frac{1}{k}$	$\frac{1}{k}$	....	$\frac{1}{k}$

$$\therefore \quad E(s^2) = \sum_{r=1}^{k} \frac{1}{k} S_r^2 = \frac{M}{k(M-1)} \sum_{r=1}^{k} \sigma_r^2$$

$$\left[\because \quad \frac{M-1}{M} S_r^2 = \sigma_r^2 \right]$$

$\frac{1}{k}\sum_{r=1}^{k}\sigma_r^2$ என்பதை σ^2 தி.உள். எனக் குறிப்பிடலாம்.

σ^2 தி.உள். என்பது திரளுக்குள் உள்ள உறுப்புகளின் பரவற் படி.

$$\therefore \quad E(s^2) = \frac{M}{M-1}\,\sigma^2_{\text{தி. உள்.}}$$

முழுமைத் தொகுதியிலிருந்து ஈ.செ.சா.ரா.மா. முறையில் M உறுப்புகள் எடுத்தால் மாதிரிச் சராசரியின் பரவற் படி

$$V_{\text{சா.ரா.மா.}} = \frac{\sigma^2}{M} = \frac{1}{Mk}\sum_{r=1}^{k}\sigma_r^2 + \frac{1}{Mk}\sum_{r=1}^{k}(\bar{X}_r - \bar{X})^2$$

$$= \frac{1}{M}\left[\sigma^2_{\text{தி. உள்.}} + \sigma^2_{\text{தி. இடை}}\right]$$

$$V_{\text{தி. மா.}} = \sigma^2_{\text{தி.இடை.}}$$

$$\therefore \quad \frac{V_{\text{சா.ரா.மா.}}}{V_{\text{தி.மா.}}} = \frac{1}{M}\left[\frac{\sigma^2_{\text{தி.உள்}} + \sigma^2_{\text{தி.இடை}}}{\sigma^2_{\text{தி.இடை}}}\right]$$

$$= \frac{1}{M}\left[1 + \frac{\sigma^2_{\text{தி.இடை}}}{\sigma^2_{\text{தி.உள்.}}}\right]$$

இதுதான் ஒரே ஒரு திரள் எடுக்கப்படும்போது தி.மா. முறையின் சா.ரா.மா. முறையுடன் ஒப்பிடப்பட்ட திறன் விகிதம். இது ஒன்றை விட அதிகமாக இருக்கவேண்டுமாயின்,

$$\frac{\sigma^2_{\text{தி உள்.}}}{\sigma^2_{\text{தி.இடை.}}} > (M-1)$$

$$\text{அதாவது } \sigma^2_{\text{தி.உள்}} > (M-1)\,\sigma^2_{\text{தி.இடை.}}$$

ஆனுல் நடைமுறையில் $\sigma^2_{\text{தி.இடை}} \geq \sigma^2_{\text{தி.உள்}}$ என்றிருக்குமாகையால் தி.மா முறை ஈ.செ.சா.ரா.மா. முறையைவிடத் திறன் மிகவும் குறைந்ததாகவே இருக்கும்.

முறையுடை மாதிரியில் கூறியபடி $\sigma^2_{\text{தி.இடை}}$ என்பதை வகுப்புள் ஒட்டுறவுக் கெழு (Intra-class correlation) ρ_c என்பதன் மூலம் எழுதலாம்.

$$\sigma^2_{\text{தி.இடை.}} = \frac{\sigma^2}{M}\left[1 + (M-1)\,\rho_c\right]$$

$$\rho_c = \frac{\sum_{r=1}^{k}\ \sum_{i=1}^{M}\ \sum_{j\neq i}^{M}(X_{ri}-\bar{X})\ (X_{rj}-\bar{X})}{kM\,(M-1)\,\sigma^2}$$

22·3. n திரள்களை ஈ.செ.சா.ரா.மா. முறையில் எடுத்தல்

k திரள்களிலிருந்து n திரள்கள் ஈ.செ.சா.ரா.மா. முறையில் எடுக்கப்படுவதாக வைத்துக்கொள்வோம். இவைகளின் மாதிரிச் சராசரிகள் $\bar{x}_1, \bar{x}_2, \ldots, \bar{x}_n$ எனவும், மாதிரிப் பரவற் படிகள் $s_1^2; s_2^2, \ldots, s_n^2$ எனவும் கொள்வோம்.

22·4. தேற்றம்

(அ) $\bar{x} = \frac{1}{n}\sum_{r=1}^{n}\bar{x}_r$ என்பது முழுமைத் தொகுதிச் சராசரிக்குப் பிறழ்ச்சியற்ற மதிப்பீட்டியாகும்.

(ஆ) $V(\bar{x}) = \frac{1}{n} \sigma^2_{\text{தி.இடை.}}$

(இ) $\overset{-1}{E}\,[V(\bar{x})] \sim \frac{1}{n(n-1)} \sum_{r=1}^{n} (\bar{x}_r - \bar{x})^2$

நிறுவல்

$E(\bar{x}_r) = \bar{X}, \quad V(\bar{x}_r) = \sigma^2_{\text{தி.இடை.}}$

$E(s^2) = \frac{M}{(M-1)} \sigma^2_{\text{தி.உள்.}}$ என்பனவற்றை அண்மையில் பார்த்தோம். மேலும் இங்கு $\bar{x}_1, \bar{x}_2, \ldots, \bar{x}_n$ என்பன சார்பற்றவை.

(அ) $E(\bar{x}) = \frac{1}{n} \sum_{r=1}^{n} E(\bar{x}_r) = \frac{1}{n} n \bar{X} = \bar{X}$

$\therefore \overset{-1}{E}(\bar{X}) \sim \bar{x}$ (நி-து.)

(ஆ) $V(\bar{x}) = \frac{1}{n^2} \sum_{r=1}^{n} V(\bar{x}_r) = \frac{1}{n^2} \cdot n \cdot \sigma^2_{\text{தி.இடை}}$

$= \frac{1}{n} \sigma^2_{\text{தி.இடை}}$ (நி-து.)

(இ) $E\left[\frac{1}{(n-1)} \sum_{r=1}^{n} (\bar{x}_r - \bar{x})^2\right]$

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

$$= \frac{1}{(n-1)} \left\{ \sum_{r=1}^{n} [V(\bar{x}_r) + \bar{X}^2] - n [V(\bar{x}) + \bar{X}^2] \right\}$$

$$= \frac{1}{(n-1)} \left[n\sigma^2_{\text{தி.இடை.}} + n\bar{X}^2 - n \cdot \frac{1}{n}\sigma^2_{\text{தி.இடை}} - n\bar{X}^2 \right]$$

$$= \sigma^2_{\text{தி.இடை.}}$$

$$\therefore \quad E\left[\frac{1}{n(n-1)} \sum_{r=1}^{n} (\bar{x}_r - \bar{x})^2 \right] = \frac{1}{n}\sigma^2_{\text{தி. இடை.}}$$

$$= V(\bar{x}).$$

$$\therefore \quad \overset{-1}{E}[V(\bar{x})] \sim \frac{1}{n(n-1)} \sum_{r=1}^{n} (\bar{x}_r - \bar{x})^2. \quad (\text{நி-து.}).$$

கிளைத்தேற்றம்

n திரள்களை ஈ. செயா. சா. ரா. மா. முறையில் எடுத்தால்

$\bar{x} = \frac{1}{n}\sum_{r=1}^{n} \bar{x}_r$ என்பது $\bar{X}$-ன் பிறழ்ச்சியற்ற மதிப்பீட்டியாகும்.

$$V(\bar{x}) = \frac{k-n}{k-1} \frac{\sigma^2_{\text{தி. இடை.}}}{n}$$

$$= \frac{k-n}{k-1} \frac{\sigma^2}{nM} \{1 + (M-1)\rho_c\}.$$

$$E^{-1}\{V(\bar{x})\} \sim \frac{k-n}{k} \frac{1}{n(n-1)} \sum_{r=1}^{n} (\bar{x}_r - \bar{x})^2$$

மேற்கூறியவற்றை எளிதாக நிறுவலாம். மாணவர்களுக்கு இது நல்லதொரு பயிற்சி.

கிளைத்தேற்றம்

முழுமைத் தொகுதியின் பாகம் P ஐ மதிப்பிட, மேற்கூறியவைகளில் சிறிய மாற்றங்களையே செய்யவேண்டி வரும். r ஆவது மாதிரிப் பகுதி p_r எனக் கொண்டால், மேற்கூறியவைகளில்

$$\bar{x}_r = p_r,\ \bar{X} = P,\ \sigma^2_{\text{தி. இடை.}} = \frac{1}{k} \sum_{i=1}^{k} (P_i - P)^2,\ \bar{X}_i = P_i,$$

$$\sigma^2 = PQ,\ \sigma^2_{\text{தி உள்.}} = \frac{1}{k} \sum_{i=1}^{k} P_i\ (1 - P_i)$$ முதலிய மாற்றங்களைச் செய்வதன் மூலம் எல்லாவற்றையும் நிறுவலாம். ஈ. செ. சா ரா. மா. முறை, ஈ. செயா. சா. ரா. மா. முறை ஆகிய இரண்டிற்கும் மேற்கூறப்பட்டது பொருந்தும். இதையும் மாணவர்கள் பயிற்சியாகக் கொள்ளவேண்டும்.

22·4. சம பருமனிலாத் திரள்கள்

இங்கு எல்லாத் திரள்களிலும் சமமான உறுப்புகள் இல்லையெனக் கொள்ளலாம். மொத்தம் k திரள்கள் உள்ளதாகக் கொள்வோம். r ஆவது திரளில் M_r உறுப்புகள் இருப்பதாகக் கொள்வோம். r ஆவது திரளில் உள்ள உறுப்புகள் $X_{r1}, X_{r2}, \ldots,$

X_{rM_r} எனக் கொண்டால் இத்திரளின் சராசரி $\bar{X}_r = \frac{1}{M_r} \sum_{i=1}^{M_r} X_{ri},$

பரவற்படி $\sigma_r^2 = \frac{1}{M_r} \sum_{i=1}^{M_r} (X_{ri} - \bar{X}_r)^2$

முழுமைத் தொகுதியின் சராசரி

$$\bar{X} = \frac{1}{N}\sum_{r=1}^{k} M_r \bar{X}_r \quad , \quad N = \sum_{r=1}^{k} M_r$$ என்பது முழுமைத் தொகுதி மொத்த உறுப்புகள்.

$$\sigma^2 = \frac{\sum_{r=1}^{k} M_r \sigma_r^2 + \sum_{r=1}^{k} M_r (\bar{X}_r - X)^2}{N}$$

சாதாரண ராண்டம் மாதிரி முறையில் n திரள்கள் எடுக்கப்பட்டால், $\bar{x}_1, \bar{x}_2, \dots \bar{x}_n$ என்பவை மாதிரித் திரள்களின் சராசரியானுல், $m_1, m_2, \dots m_n$ என்பவை முறையே திரள்களின் பருமனுனல் $\bar{x}_r$-ன் பரவல் பின்வருமாறு:

$\bar{x}_r$	$\bar{X}_1$	$\bar{X}_2$	....	$\bar{X}_k$
P	$\frac{1}{k}$	$\frac{1}{k}$	....	$\frac{1}{k}$

$$\therefore \quad E(\bar{x}_r) = \frac{1}{k}\sum_{i=1}^{k} \bar{X}_i \neq X.$$

ஆகவே இங்கு $\bar{x}_r$ என்பது முழுமைத் தொகுதியைப் பிறழ்ச்சியின்றி மதிப்பிடாது.

22·5. தேற்றம்

$$\frac{k}{nN}\sum_{r=1}^{n} m_r \bar{x}_r$$ என்பது $\bar{X}$-ன் பிறழ்ச்சியற்ற மதிப்பீடாகும்.

நிறுவல்

$$E(m_r \bar{x}_r) = \frac{1}{k}\sum_{r=1}^{k} M_r \bar{X}_r = \frac{N}{k}\bar{X}$$

$$\therefore \quad E\left[\frac{k}{nN}\sum_{r=1}^{n} M_r \bar{x}_r\right] = \frac{k}{nN} n . \frac{N}{k}\bar{X} = \bar{X}$$

$$\therefore \quad \overset{-1}{E}(\bar{X}) \sim \frac{k}{nN}\sum_{r=1}^{n} m_r \bar{x}_r$$

கிளைத்தேற்றம்

ஈ. செ. சா. ரா. மா. முறை கையாளப்பட்டால்

$$V[\overset{-1}{E}(\bar{X})] = \frac{k^2}{n^2 N^2}\sum_{r=1}^{n}\left[\frac{1}{k}\sum_{r=1}^{k} M_r^2 \bar{X}_r^2 - \frac{N^2}{k^2} X^2\right]$$

$$= \frac{k^2}{nN^2}\left[\frac{1}{k}\sum_{r=1}^{k} M_r^2 \bar{X}_r^2 - \frac{N^2}{k^2}\bar{X}^2\right]$$

$$= \frac{k}{nN^2}\sum_{r=1}^{k} M_r^2 \bar{X}_r^2 - \frac{1}{n}\bar{X}^2$$

22·6. தேற்றம்

$$\overset{-1}{E}\{V[\overset{-1}{E})\bar{X})]\} \sim \frac{1}{n(n-1)}\sum_{r=1}^{n}\left[\frac{k\, m_r}{N}\bar{x}_r - \frac{k}{nN}\sum_{r=1}^{n} m_r \bar{x}_r\right]^2$$

நிறுவல்

$$E\left\{\frac{1}{n(n-1)}\sum_{r=1}^{n}\left[\frac{k\,m_r}{N}\bar{x}_r-\frac{k}{nN}\sum_{r=1}^{n}m_r\bar{x}_r\right]^2\right\}$$

$$=\frac{1}{n(n-1)}E\left\{\sum_{r=1}^{n}\frac{k^2\,m_r^2}{N^2}\bar{x}_r^2-\right.$$

$$\left.n\left(\frac{k}{nN}\sum_{r=1}^{n}m_r\bar{x}_r\right)^2\right\}$$

$$=\frac{1}{n(n-1)}\left\{\frac{k^2}{N^2}\,\frac{n\sum_{r=1}^{k}M_r^2\bar{X}_r^2}{k}\right.$$

$$\left.-n\left(V[\overset{-1}{E}(\bar{X})]+\bar{X}^2\right)\right\}$$

$$=\frac{1}{n(n-1)}\left\{\frac{kn}{N^2}\sum_{r=1}^{n}M_r^2\bar{X}_r^2-n\bar{X}^2-nV[\overset{-1}{E}(\bar{X})]\right\}$$

$$=\frac{1}{(n-1)}\{n\,V\,[\overset{-1}{E}(\bar{X})]-V[\overset{-1}{E}(\bar{X})]\}$$

$$=V(\overset{-1}{E}\,(\bar{X})] \qquad \text{(நி-து.)}$$

கவனிக்கவும்

$\overset{-1}{E}(\bar{X})=\dfrac{k}{nN}\sum_{r=1}^{k}m_r\bar{x}_r$ என்பதில் m_r, $\bar{x}_r$, n, k முதலியவை மாதிரி எடுக்கையிலேயே தெரிந்திருக்கும். ஆனால் N என்பது

மாதிரியிலிருந்தே கண்டுபிடிக்க முடியாது. ஆகவே மேற்கூறிய மதிப்பீட்டியில் N தெரியும் என்று நாம் கொள்ள வேண்டியிருக்கிறது.

$$N \text{ தெரியாது என்றால் } E^{-1}(\bar{X}) \doteq \frac{\sum_{r=1}^{n} m_r \bar{x}_r}{\sum_{r=1}^{n} m_r} \text{ என்று}$$

$$\text{கொள்ளலாம்.} \quad E\left[\sum_{r=1}^{k} m_r \bar{x}_r\right] = \frac{nN}{k}\bar{X}$$

$$E\left[\sum_{r=1}^{k} m_r\right] = \sum_{r=1}^{k} \frac{1}{k}\sum_{r=1}^{n} m_r = \frac{nN}{k}$$

$$\therefore \quad \frac{E\left[\sum_{r=1}^{n} m_r \bar{x}_r\right]}{E\left[\sum_{r=1}^{n} m_r\right]} = \frac{\frac{nN}{k}\bar{X}}{\frac{nN}{k}} = \bar{X}$$

$$\text{ஆனால் } E\left[\frac{\sum_{r=1}^{n} m_r \bar{x}_r}{\sum_{r=1}^{n} m_r}\right] \neq \frac{E\left[\sum_{r=1}^{n} m_r \bar{x}_r\right]}{E\left[\sum_{r=1}^{n} m_r\right]} \text{ என்னும்}$$

காரணத்தால் $\dfrac{\sum_{r=1}^{n} m_r \bar{x}_r}{\sum_{r=1}^{n} m_r}$ என்னும் மதிப்பீட்டி $\bar{X}$ ஐப் பிறழ்ச்சி யின்றி மதிப்பிடாது. ஆகவே, இம் மதிப்பீட்டி தோராயமாகத்தான் $\bar{X}$ ஐ மதிப்பிடும்.

22·7. தேற்றம்

ஈ. செயா. சா. ரா. மா. முறை திரள் மாதிரி முறையில் உபயோகப்படுத்தப்பட்டால் $\bar{x} = \dfrac{k}{nN} \sum m_r \bar{x}_r$ என்னும் $\bar{X}$-ன் மதிப்பீட்டியின் பரவற் படி,

$$V(\bar{x}) = \frac{k-n}{k-1}\,\frac{1}{nk} \sum_{r=1}^{k} \left[\frac{k\,M_r}{N}\,\bar{X}_r - \bar{X}\right]^2$$

$$E^{-1}\,[V(\bar{x})] \sim \frac{k-n}{kn}\,\frac{1}{(n-1)} \sum_{r=1}^{n} \left[\frac{k\,m_r}{N}\,\bar{x}_r - E^{-1}\,(\bar{X})\right]^2$$

இவையிரண்டும் மாணவர்களுக்குப் பயிற்சியாக ஒதுக்கப்படு கின்றன.

பயிற்சி 7

1. முழுமைத் தொகுதியிலுள்ள NM உறுப்புகள் ராண்டம் முறையில் M உறுப்புகள் கொண்ட N திரள்களாகப் பிரிக்கப்படு கின்றன. n திரள்கள் ஈ. செயா. சா. ரா. மா. முறையில் எடுக்கப் படுகின்றன. இதன் திறன் முழுமைத் தொகுதியிலிருந்து nM உறுப்புகள் கொண்ட ஈ. செயா. சா. ரா. மா. முறையின் திறனுக்குச் சமம் என நிறுவு.

2. M_0 உறுப்புகள் கொண்ட முழுமைத் தொகுதி N திரள்களாகப் பிரிக்கப்படுகிறது. i ஆவது திரளில் M_i உறுப்புகள் உள்ளன. இம்முழுமைத் தொகுதியிலிருந்து m உறுப்புகள் ஈ.செயா. சா. ரா. மா. முறையில் எடுக்கப்பட்டு அவைகள் திரள்வாரியாகப் பிரிக்கப்படுகின்றன. N திரள்களிலிருந்து n திரள்கள் ஈ. செயா. சா. ரா. மா. முறையில் எடுக்கப்படுகின்றன. இத்திரள்களிலுள்ள முன்னமேயே எடுக்கப்பட்ட உறுப்புகள் (m உறுப்புகளிலிருந்து) மட்டும் ஒரு மாதிரியாக எடுத்துக் கொள்ளப்படுகின்றன. இம் மாதிரியின் மொத்தம் y என்றால் $\frac{Ny}{mn}$ என்பது முழுமைத் தொகுதி மொத்தம் Y-ன் பிறழ்ச்சியற்ற மதிப்பீட்டியென நிறுவு. இதன் பரவற் படியென்ன?

[Ghosh, S P., Ann. Math. Stat. 34 (1963)]

விடை $(M_0-1)mn\, V\left(\frac{Ny}{mn}\right) = (M_0-m)\left\{ N\sigma^2 + (N-n)\bar{Y}^2 \right\}$

$$+ N^2 (N-n)(m-1) \left\{\frac{V_c}{M_0(N-1)}\right\}$$

$$\sigma^2 = \frac{1}{M_0} \sum_{i=1}^{N} \sum_{j=1}^{M_i} Y_{ij}^2 - \bar{Y}^2, \quad V_c = \frac{1}{N} \sum_{i=1}^{N} \left(M_i \bar{Y}_i - \frac{Y}{N} \right)$$

3. ஒரு குறிப்பிட்ட வட்டத்தில் ஆண்களின் விகிதம் $\left(\frac{\text{ஆண்களின் எண்ணிக்கை}}{\text{மொத்த ஜனத் தொகை}}\right)$ தேவைப்படுகிறது. இதற்கு நேராக நபர்களை அளவெடுப்பதற்குப் பதிலாக, வீடுகள் அளவெடுக்கப் பட்டால் திறன் விகிதம் என்ன என்பது கணக்கிடப்படவேண்டும். இதற்குப் பின் கண்ட நிபந்தனைகள் சரியெனக் கொள்ளப்பட்டன.

(அ) ஒவ்வொரு வீட்டிலும் கணவன், மனைவி, இரு குழந்தைகள், ஆக மொத்தம் சரியாக நான்கு பேர்களே உள்ளனர்.

(ஆ) குழந்தை ஆணாக இருப்பதற்கு ஈருறுப்புப் பரவல் (Binomial distribution) வகுப்புள் ஒட்டுறவுக் கெழு (Intra class rrelation coefficient) $= -\frac{1}{6}$ என்றும், திறன் விகிதம் 200% ன்றும் நிறுவுக.

[Sukhatme, P. V., Sampling Theory of Surveys with Applications (1953)]

23. விகித மதிப்பீடு

(RATIO ESTIMATION)

23·1. இதுவரை நாம் பார்த்த மாதிரி முறைகளில் முழுமைத் தொகுதி மொத்தம், முழுமைத் தொகுதிச் சராசரி, முழுமைத் தொகுதி விகித சமம் (Proportion) முதலியவைகளையே மதிப்பிட்டோம். முழுமைத்தொகுதி உறுப்புகள் ஒவ்வொன்றிற்கும் (X_i, Y_i) என்னும் இரு அளவுகளை நாம் நிறுவ முடியுமென்றால்

$$\frac{Y}{X} = \frac{\sum_{i=1}^{n} Y_i}{\sum_{i=1}^{n} X_i}$$ என்பது ஒரு விகிதமாகும். இவ்

விகிதத்தை எவ்வாறு மதிப்பிடுவது? இத்தகைய விகிதங்களை மதிப்பிடுவது நடைமுறையில் மிகவும் அவசியமாகிறது. உதாரணமாக ஒரு நாட்டின் சராசரி வருமானம் என்பது மொத்த வருமானம், ஜனத்தொகை ஆகியவற்றின் விகிதமாகும். ஒரு நகரத்தில் ஆண் பெண் விகிதம் என்பது மொத்த ஆண்களின் எண்ணிக்கை, மொத்தப் பெண்களின் எண்ணிக்கை ஆகியவற்றின் விகிதம். இதைப் போலவே பிறப்பு விகிதம், இறப்பு விகிதம் ; வேலையில்லாதோர்—வேலையிலுள்ளோர் விகிதம் முதலியவைகளும் இதே பாகுபாட்டில் வரும்.

பொதுவாக $R = \frac{Y}{X}$ என்பதை மதிப்பிடுவதற்கு, Y-ன் மதிப்பீட்டி : X-ன் மதிப்பீட்டி என்பதையே மதிப்பீட்டியாகக்

கொள்வது வழக்கம். X, Y, R முதலியவைகளின் மதிப்பீட்டிகளை முறையே $\hat{X}, \hat{Y}, \hat{R}$ என்று எடுத்துக் கொண்டால்,

$\hat{R} = \frac{\hat{Y}}{\hat{X}}$ ஆகும். இங்கு $\hat{Y}, \hat{Y}$ என்பவை, X, Y-ன் பிறழ்ச்சியற்ற மதிப்பீட்டிகளாக இருந்தால்கூட $\hat{R}$ என்பது R-ன் பிறழ்ச்சியற்ற மதிப்பீட்டி யென்று கூறமுடியாது.

$$\text{ஏனெனில் } E\left(\hat{R}\right) \neq \frac{E\left(\hat{Y}\right)}{E\left(\hat{X}\right)} = \frac{Y}{X}$$

ஆகவே $\hat{R}$ என்பது பொதுவாக R-ன் பிறழ்ச்சியுடைய மதிப்பீட்டியாகவே இருக்கும்.

X என்பது தெரிந்திருந்தால் $\hat{R} = \frac{\hat{Y}}{X}$ என்று எடுத்துக் கொள்ளலாமேயெனும் கேள்வியெழலாம். ஆனால் $\frac{Y_i}{X_i}$ என்பது தோராயமாக மாறிலியெனக் கொண்டால்,

$$\hat{R}_1 = \frac{\hat{Y}}{X} ; \hat{R}_2 = \frac{\hat{Y}}{\hat{X}} \text{ என்பனவற்றுள்}$$

$\hat{R}_1$ஐ விட $\hat{R}_2$ திறன் அதிகம் வாய்ந்தது என்பதை நிறுவ முடியும்.

ஆகவே Y என்பதை மதிப்பிடுவதற்குக்கூட $\hat{Y} = \hat{R} \cdot \hat{X}$ என்று R, X ஆகியவற்றின் மதிப்பீட்டிகளின் பெருக்குத் தொகையை Y-ன் மதிப்பீட்டியாக எடுத்துக்கொள்வது நல்ல முறையாக இருக்கக்கூடும்.

பொதுவாக விகித மதிப்பீடுகளில் மதிப்பீட்டிகள் பிறழ்ச்சியுடையனவாக இருக்குமாதலால் அவைகளின் பிழையளவுகளைப் பரவற் படியாகக் கொள்வது சிறந்ததாகாது. வர்க்கவிலக்கப் பிழையே (Mean Square Error) பிழையளவாகக் கொள்ளப்படும்.

அதாவது R என்பது மதிப்பிடப்படுவதற்கு $\hat{R}$ என்னும் மதிப்பீட்டி கையாளப்பட்டால் அதன் பிழையளவு $= MSE(\hat{R}) = E(\hat{R}-R)^2$.
ஆகவே $\hat{R}_1, \hat{R}_2$ என்பவை R-ன் மதிப்பீட்டிகளானால்

$$\frac{MSE(\hat{R}_2)}{MSE(\hat{R}_1)} = \frac{E(\hat{R}_2 - R)^2}{E(\hat{R}_1 - R)^2}$$ என்பதுதான் $\hat{R}_2$ ஐப்

பொறுத்த $\hat{R}_1$-ன் திறன் விகிதமாகக் கருதப்படும்.

23·2 விகித மதிப்பீட்டில் பிறழ்ச்சி

$\hat{R}$ என்பது R-ன் மதிப்பீட்டியானால் $B(\hat{R}) = E(\hat{R}) - R$ $\hat{R}$-ன் பிறழ்ச்சியாகும்.

$\hat{R} = \dfrac{\hat{Y}}{\hat{X}}$ எனக் கொள்வோம்.

$y = \dfrac{\hat{Y} - Y}{Y}$, $x = \dfrac{\hat{X} - X}{X}$ எனவும், $\hat{X}$, $\hat{Y}$ என்பவை

X, Y-ன் பிறழ்ச்சியற்ற மதிப்பீட்டிகள் எனவும் கொள்வோம்,

$$E(y) = \frac{E(\hat{Y} - Y)}{Y} = 0 \qquad E(x) = \frac{E(\hat{X}) - X}{X} = 0$$

$$V(y) = \frac{V(\hat{Y})}{Y^2} \qquad V(x) = \frac{V(\hat{X})}{X^2}$$

$$\frac{\text{cov}(\hat{X}, \hat{Y})}{XY} = \text{cov}(x, y).$$

ஆகவே $$R = \frac{\hat{Y}}{\hat{X}} = \frac{\hat{Y} - Y + Y}{\hat{X} - X + X} = \frac{Y}{X}\left[\frac{\dfrac{\hat{Y} - Y}{Y} + 1}{\dfrac{\hat{X} - X}{X} + 1}\right]$$

$$\therefore \quad \hat{R} = R\left(\frac{y+1}{x+1}\right)$$

இங்கு $|x| < 1$ எனக்கொண்டால் $\frac{1}{1+x}$ என்பதை ஈருறுப்புத் தேற்றத்தின் மூலம் $(1 - x + x^2 \ldots\ldots)$ என்று எழுதலாம். $x^3, x^4, \ldots$ என்பவை மிகச் சிறிய எண்கள் எனக் கருதி அவைகளை ஒதுக்கிவிட்டால்

$$\hat{R} = R\left(\frac{y+1}{x+1}\right) \doteq R(1+y)(1-x+x^2)$$

$$\doteq R\,[1 + (y-x) + (x^2 - xy)]$$

{ x^2y ஐயும் ஒதுக்கினால் }

$$\therefore \quad E(\hat{R}) \doteq R + R\,E(y-x) + E(x^2 - xy)$$

$$\doteq R + R\cdot 0 + R\cdot\frac{V(\hat{X})}{X^2} - R\cdot\frac{\text{cov}(X, Y)}{XY}$$

$$\doteq R + \frac{R}{X^2}\left[V(\hat{X}) - \frac{X}{Y}\text{cov}(\hat{X}, \hat{Y})\right]$$

$$\therefore \quad B(\hat{R}) \doteq \frac{1}{X^2}\,[\,R\,V(\hat{X}) - \text{cov}(\hat{X}, \hat{Y})\,]$$

இதிலிருந்து விகித மதிப்பீட்டில் பிறழ்ச்சியில்லாமல் (தோராயமாக) இருக்கவேண்டுமாயின்

$$R = \frac{\text{cov}(\hat{X}, \hat{Y})}{V(\hat{X})}$$ என்றிருக்கவேண்டும்.

$\hat{Y}, \hat{X}$ என்னும் மாறித்தொடர்பு (Regression) ஆதியின் (origin) வழியாகச் செல்லும் நேர்க்கோடாக இருப்பின் மேற்கூறிய

நிபந்தனை நிறைவேறும். ஆகவே இந்த நிலைமையில் விகித மதிப்பீடு பிறழ்ச்சி அதிகமின்றியிருக்கும்.

23·3. சராசரி வர்க்கவிலக்கப் பிழை (Mean Square Error)

$$MSE(\hat{R}) = E(\hat{R} - R)^2$$

$$= E\left[R\frac{y+1}{x+1} - R\right]^2 = R^2 E\left[\frac{y-x}{x+1}\right]^2$$

$$\doteq R^2 E[(y-x)^2(1-2x+3x^2)]$$

$$\doteq R^2 E[(y-x)^2]$$

[x^2, xy, y^2 முதலியவைகளை ஒதுக்கினால்]

$$\therefore \quad MSE(\hat{R}) \doteq R^2 [E(y^2) - 2E(xy) + E(x^2)]$$

$$\doteq R^2\left[\frac{V(\hat{Y})}{Y^2} - \frac{2\operatorname{cov}(\hat{X}, \hat{Y})}{XY} + \frac{V(\hat{X})}{X^2}\right]$$

$$\doteq \frac{1}{X^2}\left[V(\hat{Y}) - 2R\operatorname{cov}(\hat{X},\hat{Y}) + R^2 V(\hat{X})\right]$$

இதை மதிப்பிடுவதற்கு

$$\frac{1}{\hat{X}^2}\left[\hat{V}(\hat{Y}) - 2\hat{R}\,\widehat{\operatorname{cov}}\,(\hat{X}, \hat{Y}) + \hat{R}^2\,\hat{V}(\hat{X})\right]$$

என்பதைக் கையாளலாம். இங்கு $\hat{V}(\hat{Y})$, $\hat{V}(\hat{X})$, $\widehat{\operatorname{cov}}(\hat{X},\hat{Y})$ என்பன முறையே $V(\hat{Y}), V(\hat{X})$, $\operatorname{cov}(X, Y)$ என்பனவற்றின் மதிப்பீட்டிகளாகும். முன்பே கூறியபடி இதுவும் பிறழ்ச்சியுள்ள மதிப்பீட்டியே ஆகும்.

23·4. விகித முறையில் Y ஐ மதிப்பிடல்

முன்பே கூறியபடி $\hat{Y}_{\text{வி.மு.}} = \dfrac{\hat{Y}}{\hat{X}} X$ என்று எடுத்துக் கொள்ளலாம்.

அதாவது $\hat{Y}_{\text{வி.மு.}} = \hat{R}\, X.$

$$\therefore \quad B(\hat{Y}_{\text{வி.மு.}}) = B(\hat{R}\, X) = X\, B(\hat{R})$$

$$= \frac{1}{x}\, [R\, V(\hat{X}) - \operatorname{cov}(\hat{X}, \hat{Y})].$$

$$V(\hat{Y}_{\text{வி.மு.}}) \doteq X^2\, V(\hat{R}) \doteq V(\hat{Y}) - 2R \operatorname{cov}(\hat{X}, \hat{Y}) + R^2\, V(\hat{X})$$

[$V(\hat{Y}_{\text{வி.மு.}}) \doteq MSE(Y_{\text{வி.மு.}})$ என்று தோராயமாக எடுத்துக் கொண்டுள்ளோம்.]

ஆகவே இம்முறை சாதாரண முறையை விடத் திறன் அதிகம் உடையதாக இருக்கவேண்டுமாயின்,

$V(\hat{Y}_{\text{வி.மு.}}) < V(\hat{Y})$ என்றிருக்கவேண்டும்.

ஆதாவது $V(\hat{Y}) - 2R \operatorname{cov}(\hat{X}, \hat{Y}) + R^2\, V(\hat{X}) < V(\hat{Y})$

$$\therefore \quad 2R \operatorname{cov}(\hat{X}, \hat{Y}) > R^2\, V(\hat{X})$$

R என்பது > 0 என்று எடுத்துக் கொண்டால்

$$\operatorname{cov}(\hat{X}, \hat{Y}) > \frac{R\, V(\hat{X})}{2}$$

$(\hat{X}, \hat{Y})$-ன் ஒட்டுறவுக் கெழு $\rho_{\hat{X}, \hat{Y}} = \dfrac{\operatorname{cov}(\hat{X}, \hat{Y})}{\sqrt{V(\hat{X})\, V(\hat{Y})}}$

$$\therefore \quad \operatorname{cov}(\hat{X}, \hat{Y}) = \rho_{x, \hat{Y}} \sqrt{V(\hat{X})\, V(\hat{Y})}$$

ஆகவே $\operatorname{cov}(\hat{X}, \hat{Y}) > \dfrac{R\, V(\hat{X})}{2}$ என்பது

$$\rho_{\hat{X},\hat{Y}} > \frac{Y}{X}\sqrt{\frac{V(\hat{X})}{V(\hat{Y})}}\ \frac{1}{2}$$ என்ருகும்.

$\frac{V(\hat{X})}{X}$, $\frac{V(\hat{Y})}{Y}$ என்பவை $\hat{X}$, $\hat{Y}$-ன் மாறுபாட்டுக் கெழுக்கள் ஆகும். இவைகளை $C(\hat{X})$, $C(\hat{Y})$ எனக் குறிப்பிட்டால்

$$\rho_{\hat{X},\hat{Y}} > \frac{1}{2}\ \frac{C(\hat{X})}{C(\hat{Y})}$$ என்பது நிபந்தனை.

ஆகவே $\hat{X}$ என்பது மிகவும் கவனத்துடன் மேற்கூறிய நிபந்தனைக்கேற்ப தேர்ந்தெடுக்கப்பட்டாலன்றி விகித முறை மதிப்பீடு அதிக பயனையளிக்காது.

$R < 0$ என்று எடுத்துக் கொண்டால் மேற்கூறிய நிபந்தனை

$$\rho_{\hat{X},\hat{Y}} < -\frac{1}{2}\ \frac{C(\hat{X})}{C(\hat{Y})}$$ என்ருகும்.

23·5. விகிதமுறை மதிப்பீட்டில் ஈ. செ. சா. ரா. மா. முறை

n இரு பரிமாண உறுப்புகள் (x_1, y_1), (x_2, y_2), ..., (x_n, y_n) என்பனவற்றை முழுமைத் தொகுதியிலிருந்து ஈ. செ. சா. ரா. மா. முறையில் எடுப்பதாக வைத்துக்கொள்வோம்.

$$\bar{x} = \frac{1}{n}\sum_{i=1}^{n} x_i$$ என்பது $\bar{X}$-ன் பிறழ்ச்சியற்ற மதிப்பீட்டியாகும்.

$$\bar{y} = \frac{1}{n}\sum_{i=1}^{n} y_i$$ என்பது $\bar{Y}$-ன் பிறழ்ச்சியற்ற மதிப்பீட்டியாகும்.

ஆகவே $\frac{Y}{X} = \frac{\bar{Y}}{\bar{X}}$ என்பதன் மதிப்பீட்டியாக $\frac{\bar{y}}{\bar{x}}$ என்பதை எடுத்துக் கொள்ளலாம்.

$$V(\hat{\bar{X}}) = \frac{\sigma_x^2}{n}, V(\hat{\bar{Y}}) = \frac{\sigma_y^2}{n}, \text{cov}(\hat{\bar{X}}, \hat{\bar{Y}}) = \rho \frac{\sigma_x \sigma_y}{n}$$

σ_x^2, σ_y^2, ρ என்பவை முழுமைத் தொகுதியின் பரவற் படிகள், உடன் தொடர்புக் கெழு (Correlation coefficient) ஆகும்.

$$\therefore \quad B(\hat{R}) = \frac{1}{n\bar{X}^2}[R\sigma_x^2 - \rho\sigma_x\sigma_y] = \frac{R}{n}[C^2(x) - \rho C(x) C(y)];$$

$C(x)$, $C(y)$ என்பவை x, y-ன் மாறுபாட்டுக் கெழுக்களாகும்.

$$V(\hat{R}) = \frac{1}{n\bar{X}^2}[\sigma_y^2 - 2R\rho\sigma_x\sigma_y + R^2\sigma_x^2] = \frac{R^2}{n}[C^2(y) - 2\rho C(x) C(y) + C^2(x)]$$

இவைகளிலிருந்து,

$$\hat{B}(\hat{R}) = \frac{1}{n\bar{x}^2}[\hat{R}s_x^2 - s_{x,y}]$$

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

$$\hat{V}(\hat{R}) = \frac{1}{n\bar{x}^2}[s_y^2 - 2\hat{R}s_{x,y} + \hat{R}^2 s_x^2]$$

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

என்பன விளங்கும்.

$$s_x^2 = \frac{1}{(n-1)}\sum_{i=1}^{n}(x_i - \bar{x})^2, \quad s_y^2 = \frac{1}{(n-1)}\sum_{i=1}^{n}(y_i - \bar{y})^2$$

$$s_{x,y} = \frac{1}{(n-1)}\sum_{i=1}^{n}(x_i - \bar{x})(y_i - \bar{y}^2)$$

கவனிக்கவும்

மேற்கூறியபடியே ஈ. செயா. சா. ரா. மா. முறையிலும் எல்லாவற்றையும் நிறுவலாம். இவை மாணவர்களுக்குப் பயிற்சியாக விடப்படுகின்றன.

பயிற்சி 8

1. கீழே கொடுக்கப்பட்டுள்ள அட்டவணையின் மூலம், Y-ன் 2 உறுப்புகள் கொண்ட விகித முறை மதிப்பீட்டி, ஈ. செயா. சா. ரா. மா. மதிப்பீட்டிகளின் திறன்களை ஒப்பிடுக. 2 உறுப்புகளுள்ள எல்லா மாதிரிகளையும் எடுப்பதன் மூலம் திறன் விகிதத்தைக் கண்டுபிடி.

மாறி	U_1	U_2	U_3	U_4	U_5	U_6
x	0	1	3	5	8	10
y	1	3	11	18	29	46

விடை :

	மாதிரி	$\hat{R} = \frac{\bar{y}}{\bar{x}}$	$\hat{Y} = \hat{R} \times X$	$\hat{Y} = 6\bar{y}$
$X = 27$	0, 1 1, 3	4	108	12
	0, 3 1, 11	4	108	36

0 | 5 1 | 18	$\frac{19}{5} = 3.8$	102.6	57
0 | 8 1 | 29	$\frac{30}{8} = 3.75$	101.25	90
0 | 10 1 | 46	$\frac{47}{10} = 4.7$	126.9	141
1 | 3 3 | 11	$\frac{14}{4} = 3.5$	94.5	42
1 | 5 3 | 18	$\frac{21}{6} = 3.5$	94.5	63
1 | 8 3 | 29	$\frac{32}{9} = 3.\dot{5}$	96	96
1 | 10 3 | 46	$\frac{49}{11} = 4.\dot{5}$	121.5	147

3	5
11	18

$\frac{29}{15} \doteq 1{\cdot}9\dot{3}$ 52·2 87

3	8
11	29

$\frac{40}{11} \doteq 3{\cdot}6$ 97·2 120

3	10
11	46

$\frac{57}{13} \doteq 4{\cdot}4$ 118·8 171

5	8
18	29

$\frac{47}{13} \doteq 3{\cdot}6$ 97·2 141

5	10
18	46

$\frac{64}{15} \doteq 4{\cdot}2\dot{6}$ 115·2 192

8	10
29	46

$\frac{75}{18} = 4{\cdot}1\dot{6}$ 112·5 225

$R = \frac{Y}{X} = \frac{108}{27} = 4.$ $Y = 108$

விகிதமுறை மதிப்பீட்டி

$$MSE\left[\hat{Y}\right] = \frac{1}{15} \sum \left(\hat{Y} - 108\right)^2$$

$$= \frac{1}{15}[0 + 0 + 5{\cdot}4^2 + 6{\cdot}75^2 + 18{\cdot}9^2 + 13{\cdot}5^2 + 13{\cdot}5^2$$

$$+ 12^2 + 13{\cdot}5^2 + 55{\cdot}8^2 + 10{\cdot}8^2 + 10{\cdot}8^2 + 10{\cdot}8^2 + 7{\cdot}2^2 + 4{\cdot}5^2]$$

$$\doteq \frac{1}{15}[1578] = 105{\cdot}2$$

ஈ. செயா. சா. ரா. மா மதிப்பீட்டி

$$MSE\left[\hat{Y}\right] = V\left(\hat{Y}\right) \frac{1}{15}[96^2 + 72^2 + 51^2 + 18^2 + 32^2$$

$$+ 66^2 + 45^2 + 12^2 + 39^2 + 21^2 + 12^2 + 63^2 + 33^2$$

$$+ 84^2 + 117^2] \doteq 3523.$$

ஆகவே விகித முறை மதிப்பீட்டியின் ஈ. செயா. சா. ரா. மா முறை மதிப்பீட்டியைச் சார்ந்த திறன் விகிதம் $\frac{3523}{105{\cdot}2} = 3349\%$

ஆக, இந்த உதாரணத்தில் விகிதமுறை மதிப்பீட்டி மிகமிகச் சிறந்தது என்பது விளங்குகிறது.

2. ஒரு பெரிய நெல் விளையும் நிலத்தில் வைக்கோலுடன் சேர்த்து நெல்லின் எடை (x), நெல்லின் எடை (y) முதலியவை ராண்டம் முறையில் குறிப்பிட்ட பரப்பளவிலிருந்து எடுக்கப் பட்டன. இவைகளிலிருந்து $C^2(x)$, $C^2(y)$, $C(x, y)$ என்பன வற்றின் மதிப்பீடுகள் முறையே 1·20, 1·24, 0·81 ; $C(x)$, $C(y)$ என்பன x, y-ன் மாறுபாட்டுக் கெழுக்கள். $C(x, y) = \rho\ C(x)\ C(y)$, ρ என்பது (x, y)-ன் உடன் தொடர்புக் கெழு. x-ன் மொத்தம் தெரியும் என எடுத்துக் கொண்டால் y-ன் மொத்தத்திற்கு விகித முறை மதிப்பீட்டியின், ஈ. செயா. சா. ரா. மா. முறை மதிப் பீட்டியைச் சார்ந்த திறன் விகிதத்தைக் கண்டுபிடி.

விடை 151%

3. x, y என்பவை X, Yஐப் பிறழ்ச்சியின்றி மதிப்பிடுகின்றன. $\left(\frac{y}{x}\right)X$ என்னும் விகித முறை மதிப்பீட்டியின் பிறழ்ச்சிக்கும்,

தரப்பிழைக்கும் (Standard error) உள்ள விகிதம் C_xஐ விட அதிகமானதல்ல என்பதை நிறுவு. $\frac{y}{x}$ -ன் மாறுபாட்டுக்கெழு C_xஐ விடக் குறைவாக இருந்தால் $\frac{y}{x}X$-ன் சார்புப் பிறழ்ச்சி C_x^2ஐ விடக் குறைவாக இருக்கும் என்பதையும் நிறுவு.

விடை

$$B\left(\frac{y}{x}X\right) = X\left[E\left(\frac{y}{x}\right) - \frac{Y}{X}\right]$$

$$V\left[\frac{y}{x}X\right] = X^2\, V\left(\frac{y}{x}\right)$$

$$\therefore \frac{B\left(\frac{y}{x}X\right)}{\sqrt{V\left(\frac{y}{x}X\right)}} = \frac{XE\left(\frac{y}{x}\right) - Y}{X\sqrt{V\left(\frac{y}{x}\right)}}$$

$$= \frac{E(x)\,E\left(\frac{y}{x}\right) - E(y)}{X\sqrt{V\left(\frac{y}{x}\right)}}$$

$$= \frac{-\operatorname{cov}\left(x, \frac{y}{x}\right)}{X\sqrt{V\left(\frac{y}{x}\right)}} = -\rho_{x,\frac{y}{x}}\sqrt{\frac{V(x)}{|X|}}$$

$$\therefore \left|\frac{B\left(\frac{y}{x}\right)X}{\sqrt{V\left(\frac{y}{x}X\right)}}\right| \leq \left|\rho_{x,\,y/x}\right| \frac{\sqrt{V(x)}}{|X|} \leq \frac{\sqrt{V(x)}}{|X|}$$

$[\because |\rho| \leq 1]$.

$$\therefore \left| \frac{B\left(\frac{y}{x}X\right)}{\sqrt{V\left(\frac{y}{x}X\right)}} \right| \leq |C_x| \qquad \text{(நி-து).}$$

$\left[\frac{y}{x}X\right]$-ன் சார்புப் பிறழ்ச்சி யென்பது $\dfrac{B\left(\frac{y}{x}X\right)}{E\left(\frac{y}{x}X\right)}$ ஆகும்.

முன்னமேயே $\left| \dfrac{B\left(\frac{y}{x}X\right)}{\sqrt{V\left(\frac{y}{x}X\right)}} \right| \leq |C_x|$ என்பதை நிறுவினோம்.

$$\therefore \frac{\left|B\left(\frac{y}{x}X\right)\right|}{\left|E\left(\frac{y}{x}X\right)\right|} \leq \frac{|C_x| \sqrt{V\left(\frac{y}{x}X\right)}}{\left|E\left(\frac{y}{x}X\right)\right|} = |C_x| \left| C^{y}/_{x} \right|$$

$\left| C^{y}/_{x} \right| \leq C_x$ என்று கொண்டால்,

$$\left| \frac{B\left(\frac{y}{x}X\right)}{E\left(\frac{y}{x}X\right)} \right| \leq |C_x| \; |C_x| = C_x^2. \qquad \text{(நி-து.)}$$

4. y, x என்பவை Y, X-ன் பிறழ்ச்சியற்ற மதிப்பீட்டிகள். $\frac{y}{x}$ -ன் சார்புப் பரவற் படி, $\rho_x{}^{y}/_{x}$ மிகவும் குறைவென எடுத்துக் கொண்டால், தோராயமாக $C_y{}^2 - C_x{}^2$ என நிறுவு.

[M.H. Hansen, W.N. Hurvitz and W.G. Madow, Sample Survey Methods and Theory Vol II]

சார்புப் பரவற்படி (Relative Variance,) என்பது C^2 ஆகும்.

ஆகவே $C^{2}_{y/x} = C_y^2 - C_x^2$ என்பதை நாம் நிறுவவேண்டும். இதைச் சென்ற கணக்கைப் போலவே நிறுவலாம்.

5. ஈ. செ. சா. ரா. மா முறையில் n உறுப்புகளைக் கொண்ட $(\bar{y}, \bar{x})$ என்பவை $(\bar{Y}, \bar{X})$ -களின் பிறழ்ச்சியற்ற மதிப்பீட்டிகளாகும். $\frac{1}{n^2}$, $\frac{1}{n^3}$ முதலியவைகளை ஒதுக்கிக் காணும் $\frac{\bar{y}}{\bar{x}}$ -ன் சாரசரி வர்க்கவிலக்கம் M_1. $\frac{1}{n^3}$, $\frac{1}{n^4}$ முதலியவைகளை ஒதுக்கிக் காணும் $\frac{\bar{y}}{\bar{x}}$-ன் சராசரி வர்க்க விலக்கம் M_2. $(\bar{x}, \bar{y})$ என்பதற்கு இரு பரிமாண இயல் நிலைப் பரவல் எனக் கொண்டு பின் வருவனவற்றை நிறுவு.

(i) $\frac{M_2 - M_1}{M_1} \leq \frac{g}{n} C_x^2$

(ii) $C_x = C_y$ எனக் கொண்டால் $\frac{M_2 - M_1}{M_1} \doteq \frac{3}{n}(2 - \rho) C_x^2$ என நிறுவு.

$\rho = \rho_{x, y}$.

[P.V. Sukhatme., Sampling theory of Surveys with applications; W.G. Cochran., Sampling Techinques]

$$E\left(\frac{\bar{y}}{\bar{x}} - \frac{\bar{Y}}{\bar{X}}\right)^2 = E\left(\frac{\eta + \bar{Y}}{\xi + X} - \frac{\bar{Y}}{\bar{X}}\right]^2 \quad \left.\begin{matrix} \xi = \bar{x} - \bar{X} \\ \eta = \bar{y} - \bar{Y} \end{matrix}\right\}$$

$$= \frac{1}{\bar{X}^4} E \left\{[\eta \bar{X} - \xi \bar{Y}]^2 \left[1 + \frac{\xi}{\bar{X}}\right]^{-2}\right\}$$

$$= \frac{1}{\bar{X}^4} E\left[\xi^2 \bar{Y}^2 + \eta^2 \bar{X}^2 - 2\bar{X}\bar{Y}\xi\eta\right.$$

$$- \frac{2}{\bar{X}} (\xi^3 \bar{Y}^2 + \xi \eta^2 \bar{X}^2 - 2\xi^2 \eta \bar{X} \bar{Y}) + \frac{3}{\bar{X}^2} (\xi^4 \bar{Y}^2$$

$$+ \xi^2 \eta^2 \bar{X}^2 - 2\bar{X}\bar{Y} \xi^3 \eta) + \dots + \dots \Big]$$

$$= \frac{1}{\bar{X}^4} \Big[\bar{Y}^2 \frac{\sigma_x^2}{n} + \frac{\bar{X}^2 \sigma_y^2}{n} - \frac{2\bar{X}\bar{Y}}{n} \sigma_{x,y}$$

$$- \frac{2}{\bar{X}} (\bar{Y}^2 \cdot 0 + \bar{X}^2 \cdot 0 - 2\bar{X}\bar{Y} \cdot 0) + \frac{3}{\bar{X}^2} \Big(\frac{3\sigma_x^4}{n^2} \bar{Y}^2$$

$$+ \frac{(1 + 2\rho^2)}{n^2} \sigma_x^2 \sigma_y^2 \bar{X}^2 - \frac{3\rho \sigma_x^3 \sigma_y}{n^2} \cdot 2 \bar{X}\bar{Y} \Big) + \dots \Big]$$

[மேற்கூறப்பட்டது இரு பரிமாண இயல்நிலைப் பரவலின் மோமெண்டுகளை உபயோகிப்பதனால் வரும்]

$$\therefore \quad M_1 = \frac{1}{n \bar{X}^4} (\bar{Y}^2 \sigma_x^2 + \bar{X}^2 \sigma_y^2 - 2\bar{X}\bar{Y} \sigma_{x,y}]$$

$$M_2 = M_1 + \frac{3}{n^2 \bar{X}^6} [3 \bar{Y}^2 \sigma_x^4 + (1+2\rho^2) \sigma_x^2 \sigma_y^2 \bar{X}^2$$

$$- 6 \bar{X} \bar{Y} \rho \sigma_x^3 \sigma_y \Big].$$

$$\therefore \quad M_2 = M_1 + \frac{3}{n^2} [3R^2 C_x^4 + (1 + 2\rho^2) R^2 C_x^2 C_y^2$$

$$- 6\rho C_x^3 C_y R^2]; \quad R = \frac{\bar{Y}}{\bar{X}}$$

$$\therefore \quad \frac{M_2 - M_1}{M_1} = \frac{\frac{3}{n} R^2 [3C_x^4 + (1+2\rho^2) C_x^2 C_y^2 - 6\rho C_x^3 C_y]}{R^2 [C_x^2 + C_y^2 - 2\rho C_x C_y]}$$

$$= \frac{9\,C_x^2}{n} \left[\frac{C_x^2 - 2\rho\, C_x C_y + \frac{1}{3}(1 + 2\rho^2)\ C_y^2}{C_x^2 - 2\rho\, C_x C_y + C_y^2} \right]$$

$$\leq \frac{9\,C_x^2}{n}$$

$$\because \frac{1 + 2\rho^2}{3} \leq 1.$$

மேலும் $C_x = C_y$ என்று கொண்டால்

$$\frac{M_2 - M_1}{M_1} \doteq \frac{\frac{3C_x^2}{n}\, 3 \left[1 - 2\rho + \frac{1 + 2\rho^2}{3} \right]}{2 - 2\rho}$$

$$= \frac{9\,C_x^2}{3n} \left[\frac{2 - 3\rho + \rho^2}{1 - \rho} \right]$$

$$= \frac{9\,C_x^2}{3n}\ \frac{(1 - \rho)(2 - \rho)}{1 - \rho}$$

$$= \frac{3\,C_x^2}{n}\,(2 - \rho)$$

6. (அ) (x, y) என்னும் மாறிகளின் தொடர்புப் போக்கு நேர் கோடானது, அதாவது $Y_i = \alpha + \beta\, X_i,\ i = 1, 2, \ldots N.$ $\bar{Y}$ ஐ மதிப்பிடும் விகித முறை மதிப்பீட்டி ஈ. செயா. சா. ரா. மா. முறையை விடத் திறன் அதிகம் இருப்பதற்கான நிபந்தனையைக் கண்டுபிடி. $\alpha = 0$ என்றால் இந்த நிபந்தனை என்னாகும்?

(ஆ) (x, y) என்னும் மாறிகளுக்கு நேர்க்கோட்டுத் தொடர்புப் போக்கு கிடையாது. ஆனால் Y_1, Y_2, Y_n என்னும் உறுப்புகள் பின்வரும் பெரும் முழுமைத் தொகுதியிலிருந்து எடுக்கப்படுகின்றன. [Desraj, J. Ind. Soc. Agr. Stat. 6, (1954)].

$$E\left(\frac{y}{x}\right) = \alpha + \beta\, x \qquad V\left(\frac{y}{x}\right) = \sigma_y^2\,(1 - \rho^2)$$

ρ என்பது (x, y)-களின் உடன்தொடர்புக் கெழு. இந்த நிலையில் (அ)-லுள்ள கேள்விகளுக்குப் பதிலென்ன ?

விடை

(அ) $V(\hat{Y}_R) = V(\hat{Y}) - 2R \operatorname{cov}(\hat{X}, \hat{Y}) + R^2 V(\hat{X})$

$$= \frac{N-n}{N-1} \cdot \frac{1}{n} [\sigma_y^2 - 2R\,\sigma_{x,y} + R^2 \sigma_x^2]$$

ஆனால் $\sigma_y^2 = \beta^2 \sigma_x^2 \quad \sigma_{x,y} = \beta\,\sigma_x^2$.

$$\therefore \quad V(\hat{Y}_R) = \frac{N-n}{N-1} \cdot \frac{1}{n} \sigma_x^2 [\beta^2 - 2\beta R + R^2]$$

$$V(Y_{\text{சா. ரா. மா.}}) = \frac{N-n}{N-1} \frac{1}{n} \sigma_y^2 = \frac{N-n}{(N-1)} \frac{\beta^2 \sigma_x^2}{n}$$

$$\therefore \quad \text{திறன் விகிதம்} = 1 \div \left[\frac{\beta^2 - 2\beta R + R^2}{\beta^2}\right]$$

இது ஒன்றைவிட அதிகமாக இருக்கவேண்டுமாயின் $(\beta - R)^2 \leq \beta^2$. அதாவது $R \geq 0$ · etc.

7. r_1, r_2 என்பவை, இரு வெவ்வேறு சமயங்களில் எடுத்த, முழுமைத் தொகுதி விகிதத்தின், ஒரே முறையில் அமைந்த மதிப்பீட்டிகள்.

$$\frac{B(r_1 - r_2)}{\sqrt{V(r_1 - r_2)}} \leq \sqrt{2}\, Cx_2 ; \quad \rho(r_1, r_2) \leq 0 \text{ ஆக இருந்}$$

தால் $\dfrac{B(r_1 - r_2)}{\sqrt{V(r_1 - r_2)}} \leq \sqrt{2}\, Cx_2 \sqrt{(\sigma_{r_1}^2 + \sigma_{r_2}^2) \div (\sigma_{r_1} - \sigma_{r_2})^2}$

$\rho(r_1, r_2) > 0$, என்று நிறுவு. Cx_2 என்பது $MSE(\bar{x})$ [இரண்டாவது முறை எடுத்த மாதிரியில்]

[J.C.Koop., Metrika, 5, (1962)]

8. m உறுப்புகளுள்ள k மாதிரிகள் சா. ரா. மா. முறையில் சார்பில்லாமல் எடுக்கப்படுகின்றன. இவைகளை இணைத்துப் பெற்ற $mk = n$ உறுப்புகள் மாதிரியாகக் கருதப்படுகிறது.

$$B(\hat{R}) = \frac{C}{n} + O\left(\frac{1}{n^2}\right) \text{ என்றால், } \hat{R}_Q = k\hat{R} -$$

$$\frac{k-1}{k} \sum_{i=1}^{k} \hat{R}_i \text{ என்பதன் பிறழ்ச்சி அதிக பட்சம் } O\left(\frac{1}{n^2}\right)$$

என்பதை நிறுவு. $\hat{R}_i$ என்பது i ஆவது சிறு மாதிரி (k மாதிரிகளிலிருந்து) நீக்கப்பட்ட பெரும் மாதிரியிலிருந்து நிறுவப்படும் மதிப்பீட்டி.

[Quenouille, M.A., Biometrika, 43, (1956)].

24. பலகட்ட மாதிரி முறை

(MULTI-STAGE SAMPLING)

24·1. இதுவரை நாம் பார்த்த மாதிரி முறைகளில் மாதிரி உறுப்புகள் ஒரே கட்டத்தில் எடுக்கப்பட்டன. ஆனால் இது எப்பொழுதும் எளிதானதல்ல. உதாரணமாக, சென்னையிலிருந்து 100 உறுப்பினர் கொண்ட மாதிரியொன்று தேவைப்படுகிறதென்று கொள்வோம். இதற்குச் சென்னை மக்களின் முழுப் பட்டியலொன்று தேவைப்படும். ஒரே கட்டத்தில் எடுக்கப்படவேண்டுமென்றால், படுகை முறையைக் கையாண்டால் கூட, ஒவ்வொரு பகுதியிலுமுள்ள மக்களின் முழுப் பட்டியல் தேவைப்படுமாதலால் இம்முறையும் மற்ற முறைகளைப்போலத்தான். திரள் மாதிரி முறையைக் கையாளலாம். ஆனால் இது திறன் மிகக் குறைந்தது. ஆகவே இதற்குப் பின்னே கொடுக்கப்பட்டுள்ள முறையைக் கையாளலாம்.

சென்னையை 200 பகுதிகளாகப் பிரித்துக்கொள்ளவும். ஒவ்வொரு பகுதியையும் தெருக்களாகப் பிரித்துக் கொள்ளவும். ஒவ்வொரு தெருவையும் வீடுகளாகப் பிரித்துக் கொள்ளவும். ஒவ்வொரு வீட்டையும் மக்களாகப் பிரித்துக் கொள்ளவும். 200 பகுதிகளிலிருந்து 5 பகுதிகளை ஏதாவதொரு மாதிரி முறையில் தேர்ந்தெடுக்கவும். இங்குப் பகுதிகளென்பன **முதல் கட்ட உறுப்புகள்**. தேர்ந்தெடுத்த பகுதிகளினின்று 5 தெருக்களை ஏதாவதொரு மாதிரி முறையில் எடுக்கவும். தெருக்கள் **இரண்டாம் கட்ட உறுப்புகள்**. தேர்ந்தெடுக்கப்பட்ட தெருக்கள் ஒவ்வொன்றிலிருந்தும் 4 வீடுகளை ஏதாவதொரு மாதிரி முறையில் எடுக்கவும். வீடுகள் **மூன்றும் கட்ட உறுப்புகள்**. தேர்ந்தெடுக்கப்பட்ட ஒவ்வொரு வீட்டிலிருந்தும் ஒருவரை ஏதாவதொரு மாதிரி முறையில் எடுக்கவும். மக்கள் **நான்காவது கட்ட உறுப்பினர்**. நான்காவது கட்டம் இங்குக் கடைசியானதால் மக்கள் **கடைசி**

கட்ட உறுப்பினராகக் கருதப்படுவர். இம்முறையில் மொத்தம் $5 \times 5 \times 4 \times 1 = 100$ உறுப்பினர் கொண்ட மாதிரி கிடைக்கும். இது நான்கு கட்ட மாதிரி முறையாகும்.

மேற்கூறப்பட்ட முறையில் முதலில் 200 **பகுதிகளின் பெயர்கள் மட்டும்** தெரிந்திருந்தால் போதும். இவைகளிலிருந்து தேர்ந்தெடுக்கப்பட்ட 5 பகுதிகளிலுள்ள **தெருக்களின் பெயர்கள் மட்டும்** தெரிந்தால் போதும். தேர்ந்தெடுக்கப்பட்ட **தெருக்களிலுள்ள வீடுகளின் எண்கள்** மட்டும் தெரிந்தால் போதும். இவைகளிலிருந்து எடுக்கப்பட்ட **வீடுகளில் உள்ளோர் பெயர்கள்** மட்டும் தெரிந்தால் போதும். ஆகவே சென்னை நகர மக்களின் மொத்தப் பட்டியல் இம்முறையில் தேவையேயில்லை. இதிலிருந்து இம்முறை எவ்வளவு எளிதானது என்பது புலப்படும்.

பொதுவாக n கட்ட மாதிரி முறையில் முதல் கட்டம், இரண்டாம் கட்டம், ..., $(n-1)$ ஆவது கட்டம் வரை திரள் மாதிரி முறை போன்றதொரு முறையையே நாம் கையாள்கிறோம். ஆனால் திரள் மாதிரி முறையில் எல்லா உறுப்புகளையும் தேர்ந்தெடுக்கப்பட்ட திரள்களிலிருந்து எடுக்கவேண்டி வரும். இம்முறையில் அப்படியல்ல. ஆகவே இது திரள் மாதிரி முறைக்கும் சாதாரண (ஒரு கட்ட மாதிரி) முறைக்கும் இடையிலுள்ள ஒரு முறை.

நடைமுறையில் சில சமயங்களில் பல கட்ட முறையொன்று தான் சாத்தியமானதாக இருக்கும். மேலும் ஒவ்வொரு கட்டத்திலும் எடுக்கப்படும் கட்ட உறுப்புகளைப்பற்றி விவரங்கள் தெரிந்திருந்தால் இவ்விவரங்களை உபயோகப்படுத்தி அடுத்த கட்ட உறுப்புகளை மிகச் சிறந்த முறையில் எடுக்கலாம். ஆகவே ஒவ்வொரு கட்டத்திலும் உத்தம முறையில் கட்ட உறுப்புகள் எடுக்கப்பட்டால் இம்முறை சாதாரண ஒரு கட்ட முறையை விட அதிகத் திறன் வாய்ந்ததாக இருக்கமுடியும். மேலும் ஒவ்வொரு கட்டத்திலும் மிகவும் அனுகூலமான மாதிரி முறையை நாம் கையாளலாம். இவ்விஷயத்தில் இது படுகை முறையைப் பின் பற்றுகிறது.

ஆனால் இம்முறையில் சங்கடங்கள் இல்லாமலும் இல்லை. ஒவ்வொரு கட்டத்திலும் கட்ட உறுப்புகளுக்கு ஒரு நிகழ்திறன் அளிக்கப்படுவதால் கடைசி கட்ட உறுப்புகளிலிருந்து அமைக்கப் படும் மதிப்பீட்டிகளின் எதிர்பார்க்கும், பரவற் படி முதலியவைகளைக் கணக்கிடுவது பொதுவாக மிகவும் கடினம்தான். இதற்கு

நிபந்தனைக்கு உட்பட்ட நிகழ்திறன்களைக் (Conditional Probability) கையாளவேண்டும். பொதுவாகக் கட்டங்கள் அதிகமாக, ஆக, கணக்கீடுகளும் மேலும் மேலும் கடினமாகும்.

24·2. நிபந்தனைக்குட்பட்ட எதிர்பார்க்கும் மதிப்பு, பரவற் படி (Conditional Expectation and Variance)

இரு கட்ட மாதிரி முறையில் இரண்டாம் கட்ட உறுப்புகளிலிருந்து நிறுவப்பெறும் மதிப்பீட்டி T-ன் எதிர்பார்க்கும் மதிப்பைப் பின்வருமாறு கணக்கிடலாம். E_2 என்பது எடுக்கப்பட்ட முதற்கட்ட உறுப்புகள் மாறிலிகள் என எண்ணி, கண்டு பிடிக்கப்படும் எதிர்பார்க்கும் மதிப்பை அளக்க உதவும் குறியாகும். E_1 என்பது முதற்கட்ட உறுப்புகளின் சார்புக்கு எதிர்பார்க்கும் மதிப்பை அளக்க உதவும் குறியாகும். E_{12} என்பது நேராக இரண்டாம் கட்ட உறுப்புகளின் சார்புகளுக்கு நிபந்தனையேதுமின்றி எதிர்பார்க்கும் மதிப்பை அளக்க உதவும் குறியாகும்.

$$E_{12}(T) = E_1 E_2 (T)$$

என்பது தேற்றம். இதை எளிதில் நிறுவலாம்.

பொதுவாக n-கட்ட மாதிரி முறையில் கடைசி கட்ட உறுப்புகள் சார்பின் எதிர்பார்க்கும் மதிப்பை

$$E_{12\cdots n}(T) = E_1 E_2 E_3 \ldots E_n (T)$$ என்று எழுதலாம்.

இதைப்போலவே இரு கட்ட மாதிரி முறையில்,

$$V_{12}(T) = V_1 E_2 (T) + E_1 V_2 (T).$$

n கட்ட மாதிரி முறையில்,

$$V_{123\cdots n}(T) = \sum_{r=1}^{n} E_1 E_2 \ldots E_{r-1} V_r E_{r+1} \ldots E_n (T)$$

மேற்கூறப்பட்டவை இருவழி பாகுபாடு (Two way classifications), பல வழி பாகுபாடு (multi way classifications) முதலியவற்றில் அதிகமாகக் கையாளப்படுகின்றன.

பரவற் படி n எண்களின் கூட்டுத் தொகையாக வருவதால், இவைகளை, ஒவ்வொரு கட்டமும் பரவற் படிக்கு எவ்வளவு பங்கைத் தருகிறது என்பதை அறியக் கையாளலாம்.

உதாரணம் :

(1, 2, 3, 4, ... , 12) என்னும் எண்கள் (1, 2, 3, 4), (5,6,7,8), (9, 10, 11, 12) என்று மூன்று பிரிவுகளாகப் பிரிக்கப் படுகின்றன. இப்பிரிவுகளிலிருந்து ஒரு பிரிவு சா. ரா. மா. முறையில் எடுக்கப் படுகிறது. எடுக்கப்பட்ட பிரிவிலிருந்து ஓர் உறுப்பு சா. ரா. மா. முறையில் எடுக்கப்படுகிறது. இவ்வெண்ணின் எதிர்பார்க்கும் மதிப்பு, பரவற் படி முதலியவைகளைக் கண்டுபிடி.

விடை

எடுக்கப்பட்ட எண் x எனக் கொள்க.

இது முதல் பிரிவிலிருந்தால் $E_2(x) = \frac{1}{4}(1 + 2 + 3 + 4)$

$= 2{\cdot}5$

,, இரண்டாம் ,, $E_2(x) = \frac{1}{4}(5 + 6 + 7 + 8)$

$= 6{\cdot}5$

,, மூன்றும் ,, $E_2(x) = \frac{1}{4}(9 + 10 + 11 + 12)$

$= 10{\cdot}5$

$$\therefore \quad E_{12}(x) = E_1 E_2(x) = \frac{1}{3}[2{\cdot}5 + 6{\cdot}5 + 10{\cdot}5]$$

$$= \frac{19{\cdot}5}{3} = 6{\cdot}5$$

x முதல் பிரிவிலிருந்தால் $V_2(x) = \frac{1}{4}[1^2 + 2^2 + 3^2 + 4^2]$

$- (2{\cdot}5)^2 = 1{\cdot}25$

x இரண்டாம் ,, $V_2(x) = \frac{1}{4}[5^2 + 6^2 + 7^2 + 8^2]$

$- (6{\cdot}5)^2 = 1{\cdot}25$

x மூன்றும் பிரிவிலிருந்தால் $V_3(x) = \frac{1}{4}[9^2 + 10^2+11^2+12^2]$

$$- (10{\cdot}5)^2 = 1{\cdot}25$$

$$\therefore \quad V_{12}(x) = V_1\, E_2(x) + E_1\, V_2(x).$$

$$E_1\, V_2(x) = \frac{1}{3}\,[1{\cdot}25 + 1{\cdot}25 + 1{\cdot}25] = 1{\cdot}25.$$

$$V_1\, E_2(x) = \frac{1}{3}\,[2{\cdot}5^2 + 6{\cdot}5^2 + 10{\cdot}5^2] - 6{\cdot}5^2 = 10\frac{2}{3}\,.$$

$$\therefore \quad V_{12}(x) = 10\frac{2}{3} + 1{\cdot}25 = 11\,\frac{11}{12}$$

மூன்று பிரிவாகப் பிரிக்காமல் ஒரே ஓர் உறுப்பை நேராக எடுத்தால்

$$E(x) = \frac{1}{11}\,(1 + 2 + \dots + 12) = 6{\cdot}5$$

$V(x) = \frac{1}{12}(1^2+2^2+ \dots + 12^2) - 6{\cdot}5^2 = 11\,\frac{11}{12}$. ஆகவே இந்த உதாரணத்தில் இரு முறைகளும் ஒரே பரவற் படியை அளிக்கின்றன. ஆனல் எப்பொழுதும் இது உண்மையென்று கூற முடியாது.

24·3. இருகட்ட மாதிரி முறை (Two-Stage Sampling)

முழுமைத்தொகுதி N முதற் கட்ட உறுப்புகளாகப் பிரிக்கப் பட்டிருக்கிறதென்றும், i ஆவது முதற் கட்ட உறுப்பில் $X_{i1}, X_{i2}, \dots\dots X_{i\,i}$ என்று M_i உறுப்புகள் உள்ளன என்றும் கொள்வோம்.

i ஆவது முதற் கட்ட உறுப்பிலுள்ள இரண்டாம் கட்ட உறுப்புகளின் மொத்தம் $X_i = \sum_{j=1}^{M_i} X_{ij}$

இவைகளின் சராசரி $\bar{X}_i = \frac{1}{M_i} \sum_{j=1}^{M_i} X_{ij}$

இவைகளின் பரவற்படி $\sigma_i^2 = \frac{1}{M_i} \sum_{j=1}^{M_i} (X_{ij} - \bar{X}_i)^2$

ஆகவே முழுமைத் தொகுதி மொத்தம்

$$X = \sum_{i=1}^{N} X_i = \sum_{i=1}^{N} \sum_{j=1}^{M_i} X_{ij}$$

முதற்கட்டத்தில் n உறுப்புகளும், இவைகளில் r ஆவது உறுப்பிலிருந்து $x_{r1}, x_{r2}, \ldots, x_{rm_r}$ என்று m_r உறுப்புகள் எடுக்கப்படுகின்றன என்றும் கொள்வோம். r ஆவது மாதிரி முதற் கட்ட உறுப்பின் மொத்தம் X_r' எனவும், பரவற்படி $\sigma_r'^2$ என்றும் கொள்வோம். M_r' என்பது அதிலுள்ள இரண்டாம் கட்ட உறுப்புகள்.

r ஆவது மாதிரி முதற் கட்டத்தில் மாதிரி மொத்தம்

$$x_r = \sum_{s=1}^{m_r} x_{rs} \text{ எனவும், சராசரி } \bar{x}_r = \frac{1}{m_r} \sum_{s=1}^{m_r} x_{rs} \quad \text{எனவும்,}$$

மாதிரிப் பரவற் படி $s_r^2 = \frac{1}{(m_r - 1)} \sum_{s=1}^{m_r} (x_{rs} - \bar{x}_r)^2$ எனவும்

கொள்வோம்.

24·4 தேற்றம்

இரு கட்ட மாதிரிமுறையில் இரண்டு கட்டங்களிலும் சாதாரண ராண்டம் மாதிரி முறை கையாளப்பட்டால் முழுமைத் தொகுதியின் மொத்தம் X-க்குப் பிறழ்ச்சியற்றதொரு மதிப்பீட்டி

$$\hat{X} = \frac{N}{n} \sum_{r=1}^{n} \frac{M_r'}{m_r'} \sum_{s=1}^{m_r} x_{rs} \text{ ஆகும்.}$$

நிறுவல்

$$E[\hat{X}] = E_{12}[\hat{X}] = E_1 E_2 [\hat{X}]$$

$$= E_1 \left\{ \frac{N}{n} \sum_{r=1}^{n} \frac{M_r'}{m_r'} E\left[\sum_{s=1}^{m_r} x_{rs} \right] \right\}$$

$$= E_1 \left\{ \frac{N}{n} \sum_{r=1}^{n} \frac{M_r'}{m_r'} m_r' \bar{X}_r' \right\}$$

$$= E_1 \left\{ \frac{N}{n} \sum_{r=1}^{n} X_r' \right\} = \frac{N}{n} n \frac{X}{N} = X.$$

கவனிக்கவும்: மேலே கொடுக்கப்பட்டது ஈ.செ.சா.ரா.மா. முறை, ஈ.செயா.சா.ரா.மா முறை ஆகிய இரண்டிற்கும் பொருந்தும்.

24.5. இருகட்டங்களிலும் ஈ.செயா.சா.ரா.மா முறை

$\hat{\hat{X}} = \frac{N}{n} \sum_{r=1}^{n} \frac{M_r'}{m_r} \sum_{s=1}^{m_r} x_{rs}$ என்று முன்பே பார்த்தோம்.

$$V(\hat{X}) = V_{12}(\hat{X}) = V_1 \; E_2(\hat{X}) + E_1 V_2(\hat{X})$$

$$= V_1 \left[\frac{N}{n} \sum_{r=1}^{n} X_r' \right]$$

$$+ E_1 \left[\frac{N^2}{n^2} \sum_{r=1}^{n} M_r'^2 \frac{M_r' - m_r}{M_r' - 1} \frac{\sigma_r'^2}{m_r} \right]$$

ஆனால் X_r' என்பது $(X_1, X_2, \ldots, X_N)$ என்னும் மதிப்புகளை $\left(\frac{1}{N}, \frac{1}{N}, \ldots \frac{1}{N}\right)$ என்னும் நிகழ்திறன்களுடன் ஏற்கும்.

$$\therefore \quad V_1(X_r') = \frac{1}{N}\sum_{i=1}^{N}\left(X_i - \frac{X}{N}\right)^2$$

$$= S^2_{\text{இடை}} \text{ எனக் குறிப்பிடுவோம்.}$$

$$\text{மேலும் } V_1\left[\frac{N}{n}\sum_{r=1}^{n} X_r'\right] = N^2\,\frac{N-n}{N-1}\,\frac{1}{n}\,S^2_{\text{இடை}}.$$

$$E_1\left[M_r'^2\,\frac{M_r' - M_r}{M_r' - 1}\,\frac{\sigma_r'^2}{m_r}\right] = \sum_{i=1}^{N} M_i^2\,\frac{M_i - m_r}{M_i - 1}\,\frac{\sigma_i^2}{m_r}$$

$$\therefore \quad V(\hat{X}) = \frac{(N-n)\,N^2\,S^2_{\text{இடை}}}{(N-1)\,n}$$

$$+\frac{N^2}{n^2}\sum_{r=1}^{n}\sum_{i=1}^{N} M_i^2\,\frac{M_i - m_r}{M_i - 1}\,\frac{\sigma_i^2}{m_r}$$

$$\frac{(N-n)N^2S^2_{\text{இடை}}}{(N-1)\,n} + \frac{N^2}{n^2}\sum_{i=1}^{n}\frac{M_i^2\sigma_i^2}{M_i - 1}\sum_{r=1}^{n}\left(\frac{M_i}{m_r} - 1\right)$$

$$\frac{(N-n)\,N^2\,S^2_{\text{இடை}}}{(N-1)\cdot n} + \frac{N^2}{n^2}\sum_{i=1}^{n}\frac{M_i^2}{M_i - 1}\,\sigma_i^2\,(M_i\,h\,n - n)$$

h என்பது m_r-களின் இனச் சராசரி

$$\therefore \quad V(\hat{X}) = \frac{(N-n)}{(N-1)} N^2 \frac{S^2 \text{ இடை}}{n}$$

$$+ \frac{N^2}{n} \sum_{i=1}^{N} \frac{M_i^2 (M_i h - 1)}{M_i - 1} \sigma_i^2.$$

$V(\hat{X})$-ன் பிறழ்ச்சியற்ற மதிப்பீட்டி

$$E^{-1} [V(\hat{X})] \sim N^2 M'^2 \frac{N-n}{N} \frac{s_b^2}{n}$$

$$+ \frac{N}{n} \sum_{r=1}^{n} M_r'^2 \frac{(M_r' - m_r)}{M_r'} \frac{1}{m_r} s_{wr}^2$$

இதில் $M' = \frac{1}{N} \sum_{i=1}^{N} M_i$

$$s_b^2 = \frac{1}{(n-1)} \sum_{r=1}^{n} \left(\frac{M_r'}{M'} \bar{x}_i - \frac{\hat{X}}{NM'} \right)^2$$

$$s_{wr}^2 = \frac{1}{m_r - 1} \sum_{s=1}^{m_r} (x_{rs} - \bar{x}_r)^2$$

மேற்கூறியதை வலப் புற உறுப்பின் எதிர் பார்க்கும் மதிப்பைக் கணக்கிடுவதன் மூலம் நிறுவலாம். மாணவர் இதைப் பயிற்சியாக எடுத்துக் கொள்ளலாம்.

24·6. முதற்கட்டத்தில் ஈ. செ. சா. ரா. மா. முறை, இரண்டாம் கட்டத்தில் ஈ. செயா. சா. ரா. மா. முறை

இங்கு $$V(\hat{X}) = \frac{N^2 M'^2}{n} \cdot \frac{1}{N} \sum_{i=1}^{N} \left(\frac{M_i}{M'} \overline{X}_i - \overline{X} \right)^2$$

$$+ \frac{N}{n} \sum_{i=1}^{N} \frac{M_i^2 (M_i - m_i)}{(M_i - 1) M_i m_i} \sum_{j=1}^{M_i} (X_{ij} - X_i)^2$$

$$= \frac{N M'^2}{n} \sum_{i=1}^{N} \left(\frac{M_i}{M'} \overline{X}_i - X \right)^2$$

$$+ \frac{N}{n} \sum_{i=1}^{N} \frac{(M_i - m_i)}{(M_i - 1)} \frac{M_i}{m_i} \sum_{j=1}^{M_i} (X_{ij} - X_i)^2$$

$$E^{-1} [V(\hat{X})] \sim \frac{1}{n(n-1)} \sum_{r=1}^{n} (NM_r' \overline{x}_r - \hat{X})^2$$

24·7. இரண்டு கட்டங்களிலும் ஈ. செ. சா. ரா. மா. முறை

இங்கு $$V(\hat{X}) = N^2 M'^2 \left[\frac{1}{n} \frac{1}{N} \sum_{i=1}^{N} \left(\frac{M_i}{M'} \overline{X}_i - \overline{X} \right)^2 + \frac{1}{nmN} \sum_{i=1}^{N} \frac{M_i}{M'} \sigma_{wi}^2 \right]$$

$$= \frac{NM'^2}{m}\left[m \sum_{i=1}^{N} \left(\frac{M_i}{M'} \overline{X}_i - \overline{X} \right)^2 + \sum_{i=1}^{N} \frac{M_i}{M'} \sigma_{wi}^2 \right]$$

$$\sigma_{wi}^2 = \frac{1}{M_i} \sum_{j=1}^{M} (X_{ij} - \overline{X}_i)^2$$

$$E^{-1}[V(\hat{X})] \sim \frac{1}{n(n-1)} \sum_{r=1}^{n} (NM_r' \overline{x}_r - \hat{X})^2.$$

24·8. பல்தோற்ற மாதிரி முறை (Multi-phase Sampling)

இம் முறையிலும் பல கட்டங்கள் உள்ளன. ஆனாலும் இது பல கட்ட மாதிரி முறை ஆகாது. இம்முறையில் முழுமைத் தொகுதியிலிருந்து முதற் கட்டமாகப்பெரிய மாதிரி ஒன்று எடுக்கப் படும். இம் மாதிரியிலுள்ள உறுப்புகளைப் பற்றிய துணை விவர மொன்று இக் கட்டத்தில் அளவிடப்படும். இவ் விவரத்தை வைத்துக் கொண்டு இம் முதற் கட்ட மாதிரியிலுள்ள உறுப்புகள் தகுந்த விதத்தில் வரிசைப்படுத்தப்படும் அல்லது பல படுகை களாகவோ, திறள்களாகவோ அமைக்கப்படும். பிறகு இரண்டாம் கட்டமாக முதல் மாதிரியிலிருந்து இரண்டாம் மாதிரி எடுக்கப்படும். மறுபடியும் இம் மாதிரியிலுள்ள உறுப்புகளின் மற்றொரு துணை விவரம் அளவிடப்படும். இவ் விவரத்தையும் முதல் துணை விவரத்தையும் வைத்துக் கொண்டு மறுபடியும் மூன்றும் கட்ட மாதிரியொன்று இரண்டாம் கட்ட மாதிரியிலிருந்து எடுக்கப்படும். இம் முறை கடைசி கட்டம் வரும் வரை கையாளப் படும்.

இது பல கட்ட முறையினின்றும் மாறுபட்டது. ஒவ்வொரு கட்டத்திலும் கடைசி கட்ட உறுப்பைப் பற்றிய விவரங்களே சேகரிக்கப்படுகின்றன. ஆனால் பல கட்ட முறையில் ஒவ்வொரு கட்டத்திலும் வெவ்வேறு உறுப்புகள் எடுக்கப்படுகின்றன. பல் தோற்ற மாதிரி முறையில் பெரிய மாதிரியிலிருந்து படிப்படியாகக்

கீழிறங்கி வந்து சிறிய மாதிரியொன்றை எடுக்கவேண்டும். ஒவ்வொரு கட்டத்திலும் உறுப்புகள் மட்டும் கடைசி கட்ட உறுப்புகளே. இதில் துணை விவரங்களை அதிகச் செலவின்றி எடுக்க முடிந்தால் அளவெடுப்பே சிக்கனமாக முடிந்துவிடும்.

ஆனால் மதிப்பீடுகளின் எதிர்பார்க்கும் மதிப்பு, பரவற் படி ஆகியவைகளைப் பொறுத்த மட்டில் பலகட்ட மாதிரி முறையும் பல் தோற்ற மாதிரி முறையும் ஒருவழிப்பட்டனவாகும்.

பயிற்சி 9

1. ஓர் ஆங்கில அகராதியிலுள்ள வார்த்தைகளை மதிப்பிடுவதற்கு, 26 எழுத்துகளிலிருந்து 10 எழுத்துகள் அவை எத்தனை பக்கங்களில் வருகின்றனவோ அவைகளுக்குச் சம விகிதத்தில் ஈ.செ சா.ரா.மா. முறையில் எடுக்கப்பட்டன. எடுக்கப் பட்ட எழுத்து ஒவ்வொன்றுக்கும் 2 பக்கங்கள் ஈ.செயா.சா ரா.மா. முறையில் எடுக்கப்பட்டன. இவைகளிலுள்ள வார்த்தைகளின் எண்ணிக்கை பின்வருமாறு :

மாதிரி எழுத்து	பக்கங்கள்	மாதிரிப் பக்கத்தில் வார்த்தைகளின் எண்ணிக்கை	
		முதல் பக்கம்	இரண்டாம் பக்கம்
S	131	34	27
C	97	27	26
N	21	44	38
S	131	24	29
F	43	25	32
J	7	42	48
U	18	24	21
P	85	53	24
A	49	47	55
D	54	38	57

[அகராதியிலுள்ள மொத்தப் பக்கங்கள் 980].

(அ) அகராதியிலுள்ள மொத்த வார்த்தைகளின் மதிப்பீடென்ன? அம்மதிப்பீட்டின் சார்புத் தரவிலக்கம் (Coefficient of variation) என்ன என்று மதிப்பிடு.

(ஆ) மேற்கூறிய முறையின் திறனை 20 பக்கம் ஈ.செ.சா. ரா.மா. முறையில் மதிப்பிடப்படும் மதிப்பீட்டியின் திறனுடன் ஒப்பிடு.

விடை : (அ) $\hat{X} = 35035$, $C_{\hat{X}} = 8{\cdot}93\%$ (ஆ) $68{\cdot}02\%$

2. ஒரு முழுமைத் தொகுதியில் N திரள்கள் உள்ளன. இவை ஒவ்வொன்றிலும் M உறுப்புகள் உள்ளன. முதல் கட்டத்தில் n திரள்களும், இரண்டாம் கட்டத்தில் தேர்ந்தெடுக்கப் பட்ட திரள்களிலிருந்து m உறுப்புகளும் ஈ.செ.சா.ரா. மா. முறையில் எடுக்கப்பட்டன. இவைகளின் மூலம் $\overline{X}$ மதிப்பிடப் பட்டது.

(அ) செலவுச் சார்பு $C = C_0 + C_1 n + C_2 nm$ ஆனால் $C = C'$ என்ற நிபந்தனையின் பேரில் n, m ஆகியவைகளின் உத்தம அளவைக் கண்டுபிடி.

(ஆ) $C' = 1000$, $C_0 = 300$, $C_1 = 9$, $C_2 = 1$ என்றால் பின் வரும் பரவற்படி ஆய்வு அட்டவணை (Analysis of variance table) மூலம் n, m ஆகியவைகளின் மதிப்புகளைக் கண்டுபிடி,

தோற்றுவாய் (Source)	வரையயற்ற பாகைகள் (Degrees of freedom)	வர்க்கங்களின் கூட்டுத் தொகை (Sum of squares)	சாரசரி வர்க்கம் (Mean square)
திரள்களுக் கிடையில்	89	$20 \sum_{i=1}^{90} (\overline{Y}_i - \overline{Y})^2$	180·9
திரள்களுக்குள்	1710	$\sum_{i=1}^{90} \sum_{j=1}^{20} (Y_{ij} - \overline{Y}_i)^2$	49·5
மொத்தம்	1799	$\sum_{i=1}^{90} \sum_{j=1}^{20} (Y_{ij} - \overline{Y})^2$	56·0

மொத்தத் திரள்கள் 90 ; ஒவ்வொரு திரளிலும் உள்ள உறுப்புகள் 20.

விடை :

(அ) $m = \dfrac{\sigma_w \sqrt{C_1}}{\sigma_b \sqrt{C_2}}, \ n = \dfrac{C' - C_0}{C_1 + mC_2}$

(ஆ) $m = 7 , \quad n = 44.$

3. MP கறுப்புப் பந்துகள், MQ வெள்ளைப் பந்துகள் உள்ள முழுமைத் தொகுதியிலிருந்து N உறுப்புகள் சா.ரா.மா. முறையில் எடுக்கப்பட்டன. $[P+Q = 1]$. இந்த N பந்துகளிலிருந்து n பந்துகள் சா.ரா.மா. முறையில் எடுக்கப்பட்டன. இவைகளில் r பந்துகள் கறுப்பு. r-ன் எதிர்பார்க்கும் மதிப்பு, பரவற் படி ஆகியவை பின்வரும் நிலைமைகளில் என்ன என்பதைக் கண்டு பிடி.

(அ) n பந்துகள் ஈ.செ.சா.ரா.மா. முறையில் எடுக்கப் பட்டன.

(ஆ) n பந்துகள் ஈ.செயா.சா.ரா.மா. முறையில் எடுக்கப் பட்டன.

மேற்கூறிய மதிப்புகளுக்கு மதிப்பீட்டிகளைக் கண்டுபிடி.
M என்பது மிக மிகப் பெரியது எனக் கொள்ளவும்.

விடை :

(அ) $E(r) = nP,\ V(r) = n(N + n - 1)\dfrac{PQ}{N},$

$$E^{-1}\ V(r) \sim \frac{(N + n - 1)\, r\,(n - r)}{(n - 1)(N - 1)};$$

(ஆ) $E(r) = nP,\ V(r) = nPQ,\ E^{-1}\ V(r) \sim \dfrac{r\,(n - r)}{n - 1}.$

25. தொடர்பு மதிப்பீடும் வேறுபாடு மதிப்பீடும்

(Regression Estimation and Difference Estimation)

25·1. வேறுபாடு மதிப்பீடு

சில சமயங்களில் ஒரு குறிப்பிட்ட பண்பளவு இரு வெவ்வேறு காலங்களில் மதிப்பிடப்படும்பொழுது வெவ்வேறு மதிப்பீடுகளைத் தரலாம். உதாரணமாக இந்தியாவின் மக்கட்தொகை 1971, 1972-ல் மதிப்பிடப்பட்டால் வெவ்வேறு மதிப்பீடுகள் கிடைக்கும். இவ்விரு அளவைகளின் வேறுபாட்டினை எவ்வாறு மதிப்பிடுவது ? காலத்தால் மட்டுமன்றி வேறு பல காரணங்களாலும் வேறுபாடுகள் ஏற்படுவதுண்டு. இந்த வேறுபாட்டினை மதிப்பிடும் முறைக்குத் தான் வேறுபாட்டு மதிப்பீட்டு முறை என்பது பெயர்.

பொதுவாக $X_2 - X_1$ என்பதை மதிப்பிட வேண்டுமெனவும், X_1, X_2 ஆகியவைகளின் மதிப்பீட்டிகள் $\hat{X}_1$, $\hat{X}_2$ எனவும் எடுத்துக் கொண்டால் $X_2 - X_1 = \triangle$-ன் மதிப்பீட்டி $\hat{\triangle} = \hat{X}_2 - \hat{X}_1 = \delta$ எனக் கொள்ளலாம்.

$$V(\delta) = V(\hat{X}_2 - \hat{X}_1) = V(\hat{X}_2) - 2\,\mathrm{cov}\,(\hat{X}_2, \hat{X}_1) + V(\hat{X}_1)$$

இதிலிருந்து $\mathrm{cov}\,(\hat{X}_2, \hat{X}_1)$ என்பது எவ்வளவுக் கெவ்வளவு அதிகமாக இருக்கிறதோ அவ்வளவு δ என்னும் மதிப்பீட்டியும் திறன் அதிகம் உள்ளதாயிருக்கும்.

25·2. தொடர்பு மதிப்பீட்டி முறை

Y என்னும் முழுமைத் தொகுதி அளவை மதிப்பிடப்பட வேண்டுமென எடுத்துக் கொள்வோம். X என்னும் மற்றோர்

அளவையை மதிப்பிடுவதன் மூலம் Y-ன் மதிப்பீட்டியைச் சிறந்ததாக ஆக்க முடியும். இதற்கு ஒரு வழி விகித முறை மதிப்பீடு. அதாவது $\hat{Y} = \left[\frac{\hat{Y}}{\hat{X}}\right] X$ என்பது விகித முறை மதிப்பீடு. இங்கு மற்றொரு முறையான தொடர்பு மதிப்பீட்டு முறையை ஆராய்வோம். X, Y ஆகியவற்றின் பிறழ்ச்சியற்ற மதிப்பீட்டிகள் $\hat{X}, \hat{Y}$ எனக் கொள்வோம். $\hat{Y}_{\text{தொ.}} = \hat{Y} + k(\hat{X} - X)$ என்பதும் Y-ன் பிறழ்ச்சியற்ற மதிப்பீடென்பது தெளிவாகத் தெரிகிறது, k என்பது மாறிலி எனக் கொண்டால் kஐ உத்தம முறையில் தேர்ந்தெடுப்பதன் மூலம் $\hat{Y}_{\text{தொ.}}$ என்பதைப் பிழை மிகக் குறைந்த மதிப்பீட்டியாக மாற்ற முடியும்.

$$V(\hat{Y}_{\text{தொ.}}) = V(\hat{Y}) + K^2 V(\hat{X}) + 2k \operatorname{cov}(\hat{X}, \hat{Y}).$$

$V(\hat{Y}_{\text{தொ.}})$ சிறுமமாக இருக்க வேண்டுமாயின்,

$$\frac{d}{dk}[V(\hat{Y}_{\text{தொ.}})] = 2k\, V(\hat{X}) + 2 \operatorname{cov}(\hat{X}, \hat{Y}) = 0$$

$$\therefore \quad k = -\frac{\operatorname{cov}(\hat{X}, \hat{Y})}{V(\hat{X})}$$

$y = a + bx$ என்னும் நேர்க்கோட்டுத் தொடர்பைக் கணக்கிடும் போது $b = \frac{\operatorname{cov}(x, y)}{V(x)}$ என்பதும், b என்பது y-ன் x மேலுள்ள தொடர்புக் கெழு (Regression coefficient of y on x) என்பதும் தெரிந்ததே. ஆகவே $[-k]$ என்பது $\hat{Y}$-க்கு $\hat{X}$-ன் மேலுள்ள தொடர்புக் கெழு என்பது தெரிகிறது. இதை β என்று குறித்தால் $k = -\beta$

ஆகவே $V(\hat{Y}_{\text{தொ.}})$-ன் சிறுமம் $= V(\hat{Y})$

$$+ \left[\frac{\operatorname{cov}(\hat{X}, \hat{Y})}{V(\hat{X})}\right]^2 V(\hat{X}) - 2\frac{[\operatorname{cov}(\hat{X}, \hat{Y})]^2}{V(\hat{X})}$$

$$= V(\hat{Y}) - \frac{[\text{cov}(\hat{X}, \hat{Y})]^2}{V(\hat{X})}$$

$$= V(\hat{Y}) \left\{ 1 - \frac{[\text{cov}(\hat{X}, \hat{Y})]^2}{V(\hat{X})\, V(\hat{Y})} \right\}$$

$$= V(\hat{Y}) \left\{ 1 - \rho^2_{\hat{X}, \hat{Y}} \right\}$$

$\hat{Y}_{\text{தொ.}}$ என்பதற்கு k-ன் உத்தம அளவின் போது தொடர்பு மதிப்பீட்டியென்பது பெயர். இது $\hat{Y}$ ஐ விடச் சிறந்தது என்பது

$$\frac{V(\hat{Y})}{V(\hat{Y}_{\text{தொ.}})} = \left[1 - \rho^2_{\hat{X}, \hat{Y}} \right]^{-1} \geq 1$$ என்பதிலிருந்து

தெரிகிறது. இதில் $\rho^2_{\hat{X}, \hat{Y}} \cdot V(\hat{Y})$ என்பது தொடர்பு மதிப்பீட்டு முறையைக் கையாள்வதால் ஏற்படும் இலாபத்தைக் குறிக்கிறது.

மேற்கூறியவற்றில் உள்ள ஒரே ஒரு சிக்கல் $\beta = -k$ என்பது என்ன என்பது நமக்குத் தெரியாது என்பதுதான். ஆகவே β ஐயும் நாம் மதிப்பிடவேண்டும். β-ன் மதிப்பீட்டி $\hat{\beta}$ என்று எடுத்துக்கொண்டால்

$$\hat{Y}_{\text{தொ.}} = \hat{Y} - \hat{\beta}\,(\hat{X} - X)$$

மேலும் $V(\hat{Y}_{\text{தொ.}}) = V(\hat{Y})\ [1 - \rho^2_{\hat{X}, \hat{Y}}]$ என்பதால் $\rho^2_{\hat{X}, \hat{Y}}$ எவ்வளவுக்கெவ்வளவு அதிகமாக இருக்கிறதோ அவ்வளவு தொடர்முறை மதிப்பீட்டியும் திறன் அதிகம் உள்ளதாக இருக்கும் எனத் தெரிகிறது. இதிலிருந்து Yஐ மதிப்பிடுவதற்கு அதனுடன் அதிகம் தொடர்புடைய X எனும் பண்பளவை யையும் எடுத்துக்கொள்ளவேண்டுமென்பதும் புரிகிறது.

மேலும் $Y_{\text{வி.மு.}} = \left[\frac{\hat{Y}}{\hat{X}}\right] X$ என்னும் மதிப்பீட்டியைவிட $\hat{Y}_{\text{தொ.}}$ திறன் அதிகம் வாய்ந்ததாக இருக்குமென நிறுவ முடியும்.

ஏனெனில்

$$V(\hat{Y}_{\text{வி.மு.}}) = V(\hat{Y}) - 2R\,\text{cov}(\hat{X}, \hat{Y}) + V(\hat{X}) \cdot R^2$$

$$\therefore \quad V(\hat{Y}_{\text{வி.மு.}}) - V(\hat{Y}_{\text{தொ.}})$$

$$= \rho^2_{\hat{X}, \hat{Y}} \cdot V(\hat{Y}) - 2R\,\text{cov}(\hat{X}, \hat{Y}) + R^2 V(\hat{X})$$

$$= \rho^2_{\hat{X}, \hat{Y}} V(\hat{Y}) - 2R\rho_{\hat{X}, \hat{Y}} \sqrt{V(\hat{X})V(\hat{Y})} + R^2 V(\hat{X})$$

$$= [\rho\sqrt{V(\hat{Y})} - R\sqrt{V(\hat{X})}]^2 \geq 0$$

$$\therefore \quad V(\hat{Y}_{\text{வி.மு.}}) \geq V(\hat{Y}_{\text{தொ.}})$$

ஆகவே தொடர்பு மதிப்பீட்டி, விகிதமுறை மதிப்பீட்டியை விடச் சிறந்தது. இவ்விரண்டு முறைகளும் சமத்திறன் வாய்ந்தவைகளாக இருக்க வேண்டுமாயின்

$$\rho\sqrt{V(\hat{Y})} - R\sqrt{V(\hat{X})} = 0$$

$$\therefore \quad Y = \rho \frac{\sqrt{V(\hat{Y})}}{\sqrt{V(\hat{X})}} X$$

அதாவது Y-க்கு X-ன் மேலுள்ள தொடர்பு ஆதி (Origin) வழிச் செல்லவேண்டும்.

25·3. தொடர்பு முறை மதிப்பீட்டில் பிறழ்ச்சியும் பரவற்படியும்

$\hat{Y}_{\text{தொ.}} = \hat{Y} - \hat{\beta}(\hat{X} - X)$ என்பதில்

$E[\hat{\beta}(\hat{X} - X)] \neq E(\hat{\beta})E(\hat{X} - X) = 0$ என்பதால் $\hat{Y}_{\text{தொ.}}$ என்பது Y-ன் பிறழ்ச்சியுடைய மதிப்பீட்டியாகத்தான் இருக்கும்.

$\dfrac{\hat{Y} - Y}{Y} = y$ என்றும், $\dfrac{\hat{X} - X}{X} = x$ என்றும் எடுத்துக் கொண்டால், $\dfrac{\hat{\beta} - \beta}{\beta} = b$ என்று எடுத்துக்கொள்ளலாம்.

$$\therefore \quad \hat{Y}_{\text{தொ.}} = Y(1+y) - \beta(1+b)[Xx]$$

$$= Y + (yY - x\beta X) - bx\beta X$$

$$\hat{Y}_{\text{தொ.}} - Y = (yY - x\beta X) - bx\beta X$$

$$\therefore \quad B(\hat{Y}_{\text{தொ.}}) = E(\hat{Y}_{\text{தொ.}} - Y)$$

$$= Y E(y) - \beta X E(x) - \beta X E(bx)$$

$E(x) = E(y) = 0.\quad E(bx) = \dfrac{\text{cov}(\hat{X}, \hat{\beta})}{X\beta}$ என்பதால்

$$B(\hat{Y}_{\text{தொ.}}) = -\text{cov}(\hat{X}, \hat{\beta})$$

மாதிரிப் பருமன் அதிகமானல் பொதுவாக $\text{cov}(\hat{X}, \hat{\beta})$ என்பது ஒதுக்கத்தக்கதாகவே இருக்குமாதலால் $\hat{Y}_{\text{தொ.}}$-ன் பிறழ்ச்சி மாதிரிப் பருமன் அதிகமாக ஆகக் குறையும்.

25·6. முழுமைத் தொகுதியிலிருந்து n சார்பற்ற மாதிரிகளை எடுப்பதன் மூலம் Y, β முதலியவைகளைப் பின்வரும் சார்புகள் மூலம் மதிப்பிடலாம். $\hat{\hat{Y}}_1, \hat{\hat{Y}}_2, \ldots, \hat{Y}_n;\ \hat{X}_1, \hat{X}_2, \ldots \hat{X}_n;\ \hat{\beta}_1, \hat{\beta}_2, \ldots, \hat{\beta}_n$ என்பவை இந்த n மாதிரிகளிலிருந்து அமைக்கப்பட்ட மதிப்பீட்டிகளானல்,

$$\hat{Y}_{\text{தொ.}} = \hat{\hat{Y}} - \hat{\hat{\beta}}(\hat{\hat{X}} - X)$$

$$\hat{B}[\hat{Y}_{\text{தொ.}}] = -\frac{1}{n(n-1)}\sum_{i=1}^{n}(\hat{X}_i - \hat{\hat{X}})(\hat{\beta}_i - \hat{\hat{\beta}}).$$

$$\hat{\hat{Y}} = \frac{1}{n}\sum_{i=1}^{n}\hat{Y}_i, \quad \hat{\hat{X}} = \frac{1}{n}\sum_{i=1}^{n}\hat{X}_i, \quad \hat{\hat{\beta}} = \frac{1}{n}\sum_{i=1}^{n}\hat{\beta}_i$$

நிறுவல்

$\hat{Y}_{\text{தொ.}} = \hat{\bar{Y}} - \hat{\bar{\beta}} \; (\hat{\bar{X}} - X)$ என்பதற்கு விளக்கம் தேவையில்லை.

$$B\,(\hat{Y}_{\text{தொ.}}) = E\,[\hat{Y}_{\text{தொ.}} - Y]$$

$$= E\,(\hat{\bar{Y}} - Y) - E\,[\hat{\bar{\beta}}\,(\hat{\bar{X}} - X)]$$

$$= E\left[\frac{1}{n}\sum_{i=1}^{n} \hat{Y}_i\right] - Y - E\,[\hat{\bar{\beta}}\;(\hat{\bar{X}} - X]]$$

$$= \frac{1}{n}\,n\,Y - Y - E\;[(\hat{\bar{\beta}} - \beta)\,(\hat{\bar{X}} - X) + \beta\,(\hat{\bar{X}} - X)]$$

$$= -E\,[(\hat{\bar{\beta}} - \beta)\,(\hat{\bar{X}} - X)]$$

$$= -E\left[\frac{1}{n}\sum_{i=1}^{n} (\hat{\beta}_i - \beta)\cdot\frac{1}{n}\sum_{i=1}^{n}(\hat{X}_i - X)\right]$$

$$= -E\left[\frac{1}{n^2}\sum_{i=1}^{n} (\hat{\beta}_i - \beta)\,(\hat{X}_i - X) + \frac{1}{n^2}\sum\sum_{i \neq j} (\hat{\beta}_i - \beta)\,(\hat{X}_j - X)\right]$$

$$= -\frac{1}{n^2}\sum_{i=1}^{n} \operatorname{cov}\,(\hat{\beta}_i, \hat{X}_i) \quad [\because \; \hat{\beta}_i, \hat{X}_j \text{ சார்பற்றவை.}]$$

$$T = \sum_{i=1}^{n} (\hat{X}_i - \hat{\bar{X}})(\hat{\beta}_i - \hat{\bar{\beta}})$$ என்று எடுத்துக்

கொள்ளவும்.

$$B(\hat{Y}_{\text{தொ.}}) = -\frac{1}{n^2} E \sum_{i=1}^{n} (\hat{\beta}_i - \beta)(\hat{X}_i - X)$$

$$= -\frac{1}{n^2} E \sum_{i=1}^{n} (\hat{\beta}_i - \hat{\bar{\beta}} + \hat{\bar{\beta}} - \beta)$$

$$(\hat{X}_i - \hat{\bar{X}} + \hat{\bar{X}} - X).$$

$$= -\frac{1}{n^2} E \sum_{i=1}^{n} (\hat{\beta}_i - \hat{\bar{\beta}})(\hat{X}_i - \bar{X})$$

$$-\frac{n}{n^2} E (\hat{\bar{\beta}} - \beta)(\hat{\bar{X}} - X)$$

[மற்ற இரண்டு உறுப்புகளும் 0].

$$= -\frac{1}{n^2} E \left[\sum_{i=1}^{n} (\hat{\beta}_i - \hat{\bar{\beta}})(\hat{X}_i - \hat{\bar{X}}) \right.$$

$$\left. + n \frac{1}{n} \sum_{i=1}^{n} (\hat{\beta}_i - \beta) . \frac{1}{n} \sum_{i=1}^{n} (\hat{X}_i - X) \right]$$

$$= -\frac{1}{n^2} E \left[T + \frac{1}{n} \sum_{i=1}^{n} (\hat{\beta}_i - \beta)(\hat{X}_i - X) \right]$$

[மற்றவை 0. ஏனென்றால் $\hat{\beta}_i$, $\hat{X}_j$ ஆகியவை சார்பற்றவை.]

$$\therefore \quad B\,(\hat{Y}_{\text{தொ.}}) = -\frac{1}{n^2}E\,(T) + \frac{1}{n}\,B\,(\hat{Y}_{\text{தொ.}}).$$

$$\therefore \quad \left(1 - \frac{1}{n}\right) B\,(\hat{Y}_{\text{தொ.}}) = -\frac{1}{n^2}\,E\,(T).$$

$$\therefore \quad B\,(\hat{Y}_{\text{தொ.}}) = -\frac{1}{n\,(n-1)}E\,(T)$$

$$= -\frac{1}{n\,(n-1)}\sum_{i=1}^{n} (\hat{X}_i - \hat{X})\,(\hat{\beta}_i - \hat{\beta})$$

25·7. சாதாரண ராண்டம் மாதிரி

தொடர்பு முறை மதிப்பீட்டியைக் கண்டுபிடிக்க. ஈ.செயா. சா.ரா.மா. முறை உபயோகப்படுத்தப்பட்டால்

$$\hat{Y} = N\bar{y}$$

$$\hat{X} = N\bar{x}$$

$$\hat{\beta} = \frac{\sum_{i=1}^{n} (x_i - \bar{x})\,(y_i - \bar{y})}{\sum_{i=1}^{n} (x_i - \bar{x})^2}$$

$(x_1, y_1), (x_2, y_2), \ldots, (x_n, y_n)$ என்பது மாதிரி. $\bar{x} = \frac{1}{n}\sum_{i=1}^{n} x_i$, $\bar{y} = \frac{1}{n}\sum_{i=1}^{n} y_i$. N என்பது முழுமைத் தொகுதியிலுள்ள மொத்த உறுப்புகள்.

$$\therefore \quad \hat{Y}_{\text{தொ.}} = N\left[\bar{y} - \hat{\beta}\,(\bar{x} - \bar{X})\right]$$

$$\therefore \quad B\,(\hat{Y}_{\text{தொ.}}) = -\operatorname{cov}\,(\hat{\beta}, \bar{x})$$

$V\,(\hat{Y}_{\text{தொ.}})$ ஐத் தோராயமாகக் கண்டுபிடிக்கலாம்.

$$V\,(\bar{y}) = \frac{N-n}{N-1}\,\frac{\sigma_y^2}{n}$$

$$\therefore \quad V\,(\hat{Y}_{\text{தொ.}}) \doteq N^2\,\frac{N-n}{N-1}\,\frac{\sigma_y^2}{n}\,(1-\rho^2) \qquad \rho = \rho_{x,y}$$

இதனுடைய மதிப்பீட்டி

$$\hat{V}\,(\hat{Y}_{\text{தொ.}}) = N^2\,\frac{(N-n)}{N\,n\,(n-1)}\sum_{i=1}^{n}\{(y_i-\bar{y}) - \hat{\beta}(x_i-\bar{x})\}^2$$

$$= \frac{N\,(N-n)}{n\,(n-1)}\sum_{i=1}^{n}[(y_i-\bar{y}) - \hat{\beta}\,(x_i-\bar{x})]^2$$

பயிற்சி 10

1. ஒரு மாவட்டத்திலுள்ள ஆடுமாடுகளின் எண்ணிக்கையை மதிப்பிட வேண்டியிருக்கிறது. அதற்காக அம்மாவட்டத்திலுள்ள 1,238 கிராமங்களிலிருந்து 24 கிராமங்கள் ஈ.செ.சா.ரா.மா.முறையில் எடுக்கப்பட்டன. ஒவ்வொரு மாதிரி கிராமத்தில் இருக்கும் ஆடுமாடுகளின் எண்ணிக்கையும், அதே கிராமத்தில் முன் வருடத்திலிருந்த ஆடுமாடுகளின் எண்ணிக்கையும் பின்வரும் அட்டவணையில் கொடுக்கப்பட்டிருக்கின்றன. இதைக்கொண்டு தொடர்பு முறை மதிப்பீட்டியின், விகித முறை மதிப்பீட்டியைச் சார்ந்த திறன் விகிதத்தை மதிப்பிடு.

விடை : 105%

கிராமம்	ஆடுமாடுகள் எண்ணிக்கை	
	தற்போது	முன்பு
1	654	623
2	696	690
3	530	534
4	315	293
5	78	69
6	640	842
7	692	475
8	292	371
9	210	161
10	555	298
11	2110	2045
12	592	1069
13	707	706
14	1890	1795
15	1123	1406
16	115	118
17	375	330
18	212	218
19	147	160
20	297	210
21	401	262
22	252	204
23	199	185
24	564	574

2. தொடர்பு முறை மதிப்பீட்டியின் பரவற் படியின் அதிகரிப்பு விகிதம், β என்பதற்குப் பதில் λ என்பதைக் கையாளும் பொழுது, αஐ விடக் குறைவாகக் இருக்கவேண்டுமாயின்

$$\left|\frac{\lambda-\beta}{\beta}\right| < \sqrt{\frac{\alpha(1-\rho^2)}{\rho^2}}$$ என்று நிறுவு.

[W. G. Cochran, Sampling Techniques (1963)]

விடை

λ கையாளப்படும்போது பரவற் படி $V\lambda = V(\hat{Y}) - 2\lambda\,\mathrm{cov}(\hat{X},\hat{Y}) + \lambda^2 V(\hat{X})$

β கையாளப்படும்போது பரவற் படி $V\beta = V(\hat{Y})[1-\rho^2]$

$$\therefore \frac{V_\lambda - V_\beta}{V_\beta} = \frac{\rho^2 V(\hat{Y}) - 2\lambda\,\mathrm{cov}(\hat{X},\hat{Y}) + \lambda^2 V(\hat{X})}{V(\hat{Y})[1-\rho^2]} < \alpha$$

என்றால்,

$$\rho^2 V(\hat{Y}) - 2\lambda\,\mathrm{cov}(\hat{X},\hat{Y}) + \lambda^2 V(\hat{X}) - \alpha V(\hat{Y})(1-\rho^2) < 0$$

ஆனால் $$\beta = \frac{\mathrm{cov}(\hat{X},\hat{Y})}{V(\hat{X})} = \rho\sqrt{\frac{V(\hat{Y})}{V(\hat{X})}}$$

$$\therefore \rho^2\frac{V(\hat{Y})}{V(\hat{X})} - 2\lambda\frac{\mathrm{cov}(\hat{X},\hat{Y})}{V(\hat{X})} + \lambda^2 - \frac{\alpha(1-\rho^2)V(\hat{Y})}{V(\hat{X})} < 0$$

$$\therefore \beta^2 - 2\lambda\beta + \lambda^2 - \alpha(1-\rho^2)\frac{\beta^2}{\rho^2} < 0$$

$$\therefore \left[\frac{\beta-\lambda}{\beta}\right]^2 < \frac{\alpha(1-\rho^2)}{\rho^2}$$

$$\therefore \left|\frac{\beta-\lambda}{\beta}\right| < \sqrt{\frac{\alpha(1-\rho^2)}{\rho^2}}$$ (நி-து.)

26. சாதாரண ராண்டம் மாதிரி [தொடர்ச்சி]

ஈ.செ.சா.ரா.மா. முறையில் சில முக்கியமான பரவல்களிலிருந்து மாதிரிகள் எடுத்தால் மாதிரிமொத்தத்தின் பரவல் கண்டுபிடிப்பது சுவையானதொரு பயிற்சியாகும்.

26·1. தேற்றம்

ஈருறுப்புப் பரவல் $B(n, p)$-லிருந்து $(x_1, x_2, \ldots\ldots, x_k)$ என்றொரு மாதிரியை ஈ.செ.சா.ரா.மா. முறையில் எடுத்தால் $x = x_1 + x_2 + \ldots\ldots + x_k$ என்னும் ராண்டம் மாறியின் பரவல் $B(nk, p)$.

நிறுவல்

இதைத் தொகுத்தறி முறையில் (Induction) நிறுவலாம். தேற்றம் 26·1. என்பது $k = m$ என்னும் எண்ணுக்கு உண்மையானது எனக் கொள்வோம்.

$x_1 + x_2 + \ldots\ldots + x_m = y$ என்றும் $x_1 + x_2 + \ldots + x_{m+1} = z$ என்றும் கொள்வோம். $z = y + x_{m+1}$ என்பது தெளிவு. தொகுத்தறி எடுகோள் (Induction hypothesis) மூலமாக y-ன் பரவல் $B(nm, p)$. அதாவது,

$$y \frown B(nm, p).$$

$$x_{m+1} \frown B(n, p).$$

$$P[z = r] = \sum_{s=0}^{r} P[y = s, x_{m+1} = r - s].$$

$$= \sum_{s=0}^{r} P(y = s)\, P(x_{m+1} = r - s)$$

[∵ y, x_{m+1} சார்பற்றவை].

$$= \sum_{s=0}^{r} \binom{nm}{s} p^s q^{nm-s} . \binom{n}{r-s} p^{r-s} q^{n-r+s}$$

$$= p^r q^{nm+n-r} \sum_{s=0}^{r} \binom{nm}{s} \binom{n}{r-s}$$

$\sum_{s=0}^{r} \binom{nm}{s} \binom{n}{r-s}$ என்பது $(1+x)^{nm}(1+x)^{n}$-ல் x^r-ன் குணகம் (கெழு.)

ஆகவே இது $\binom{nm+n}{r}$

$$\therefore \quad P(z = r) = \binom{nm+n}{r} p^r q^{nm+n-r}$$

$r = 0, 1, 2, \ldots nm+n$

$\therefore \quad Z \frown B\, [n(m+1), p]$

∴ தேற்றம் $k = m + 1$-க்கும் உண்மையானது. $k = 1$ என்னும் போது தேற்றம் எளிதில் விளங்கும். ஆகவே தொகுத்தறி முறைப்படி தேற்றம் எல்லா k-க்கும் உண்மையானது.

26·2. தேற்றம்

$(x_1, x_2, \ldots\ldots, x_n)$ என்பது பாய்ஸான் பரவல் $P(\lambda)$-லிருந்து எடுக்கப்பட்ட ஈ.செ.சா.ரா.மா. ஆனல் $x_1 + x_2 + \ldots\ldots + x_n = x \frown P(n\lambda)$.

நிறுவல்

இதையும் தொகுத்தறி முறையில் நிறுவுவோம். மேற்கூறிய தேற்றம் $n = m$ என்னும்போது உண்மையானது எனக் கொள்வோம்.

$y = x_1 + x_2 + \ldots\ldots + x_m$, $z = x_1 + x_2 + \ldots\ldots + x_{m+1}$ எனக் கொள்க.

$z = y + x_{m+1}$ $\qquad \therefore \quad y \cap P(m\lambda),\ x_{m+1} \cap P(\lambda)$

$$P[z = r] = \sum_{s=0}^{r} P[y = s,\ x_{m+1} = r - s]$$

$$= \sum_{s=0}^{r} P[y = s]\ P[x_{m+1} = r - s]$$

[$\because$ y, x_{m+1} சார்பற்றவை]

$$= \sum_{s=0}^{r} e^{-m\lambda} \frac{(m\lambda)^s}{\lfloor s} \cdot e^{-\lambda} \frac{\lambda^{r-s}}{\lfloor r-s}$$

$$= e^{-(m+1)\lambda} \lambda^r \sum_{s=0}^{r} \frac{m^s}{\lfloor s \ \lfloor r-s}$$

$$= e^{-(m+1)\lambda} \frac{\lambda^r}{\lfloor r} \sum_{s=0}^{r} \binom{r}{s} m^s$$

$$= e^{-(m+1)\lambda} \frac{\lambda^r}{\lfloor r} (1 + m)^r$$

$$= e^{-(m+1)\lambda} \frac{[\lambda (1+m)]^r}{\underline{|r}}, \; r = 0, 1, 2, \ldots$$

$\therefore \quad Z \frown P\,[(m+1)\,\lambda\,]$.

ஆகவே தேற்றம், $n = m + 1$-க்கும் உண்மையானது.

ஆனால் தேற்றம் $n = 1$ என்னும்பொழுது உண்மையானது என்பது தெளிவு. ஆகவே தொகுத்தறி முறையில் $x_1 + x_2 \ldots\ldots + x_n \frown P\,(n\,\lambda)$.

26·3. தேற்றம்

$(x_1, x_2, \ldots, x_n)$ என்பது இயல் நிலைப் பரவல் $N\,(\mu, \sigma^2)$ என்பதிலிருந்து எடுக்கப்பட்ட சாதாரண ராண்டம் மாதிரியானால்

$x_1 + x_2 + \ldots + x_n = x \frown N\,(n\mu, n\sigma^2)$.

நிறுவல்

$$\frac{x_i - \mu}{\sigma} \frown N\,(0, 1).$$

$\therefore \quad y_i = \dfrac{x_i - \mu}{\sigma}$ எனக் கொண்டால்,

$y_1 + y_2 + \ldots + y_n \frown N\,(0, n)$ என்று நிறுவினால் போதுமானது. இதைத் தொகுத்தறி முறையில் நிறுவுவோம். இத்தேற்றம் $n = m$ என்னும்பொழுது உண்மையானது எனக் கொள்வோம்.

$y_1 + y_2 + \ldots + y_m = y$ என்றும், $y_1 + y_2 + \ldots + y_{m+1} = z$ என்றும் கொள்க. $z = y + y_{m+1}$.

$y \frown N\,(0, m)$.

$y_{m+1} \frown N\,(0, 1)$.

$$\therefore \quad f(z) = \frac{1}{2\pi\sqrt{m}} \int_{-\infty}^{\infty} e^{-\frac{x^2}{2m}} \cdot e^{-\frac{1}{2}(z-x)^2}\, dx.$$

$$= \frac{1}{2\pi\sqrt{m}} \cdot \int_{-\infty}^{\infty} e^{-\frac{1}{2m}[x^2 + mx^2 + mz^2 - 2mxz]} dx$$

$$= \frac{1}{2\pi\sqrt{m}} \cdot \int_{-\infty}^{\infty} e^{-\frac{m+1}{2m}\left[x^2 - 2x\frac{mz}{m+1} + \frac{m}{m+1}z^2\right]} dx$$

$$= \frac{1}{2\pi\sqrt{m}} e^{-\frac{m+1}{2m}\left[\frac{mz^2}{m+1} - \frac{m^2z^2}{(m+1)^2}\right]} \int_{-\infty}^{\infty} e^{-\frac{\left(x - \frac{mz}{m+1}\right)^2}{2\frac{m}{m+1}}} dx.$$

$$= \frac{1}{2\pi\sqrt{m}} e^{-\frac{1}{2(m+1)}[z^2]} \sqrt{2\pi}\sqrt{\frac{m}{m+1}}$$

$$= \frac{1}{\sqrt{2\pi}\ \sqrt{m+1}} e^{-\frac{z^2}{2(m+1)}}.$$

$\therefore \quad z \frown N(0, m+1).$

இதிலிருந்து தொகுத்தறி முறையில் தேற்றம் உண்மையென்றுகிறது.

26·4. தேற்றம்

$(x_1, x_2, \ldots, x_n)$ என்பது $G(\alpha, p)$ என்னும் காமா பரவல் (Gamma distribution) ஒன்றிலிருந்து எடுக்கப்பட்ட சாதாரண ராண்டம் மாதிரியானல்

$x_1 + x_2 + \ldots + x_n \frown G(\alpha, np).$

நிறுவல்

$G(\alpha, p)$-ன் ஊக அளவு அடர்த்தி $\dfrac{\alpha^p}{\Gamma(p)} e^{-\alpha x} x^{p-1}$;

$$0 \leq x \leq \infty$$

இத் தேற்றத்தைத் தொகுத்தறி முறையில் நிறுவுவோம். $n = m$ என்னும்பொழுது இத்தேற்றம் உண்மையென்று கொள்வோம்.

$x_1 + x_2 + \dots + x_m = x, \qquad x_1 + x_2 + \dots + x_{m+1} = y$

$y = x + x_{m+1}$,

$x \cap G(\alpha, mp) \qquad x_{m+1} \cap G(\alpha, p)$.

$$\therefore \; f(y) = \int_0^y \frac{\alpha^{mp}}{\Gamma(mp)} \cdot e^{-\alpha x} x^{mp-1} \frac{\alpha^p}{\Gamma(p)} \cdot e^{-\alpha(y-x)} (y-x)^{p-1} dx$$

$$= \frac{\alpha^{(m+1)p} e^{-\alpha y}}{\Gamma(mp)\,\Gamma(p)} \int_0^y x^{mp-1} (y-x)^{p-1} dx$$

$x = yt$ என்று மாற்றவும். $dx = ydt$.

$$\therefore \; f(y) = \frac{\alpha^{(m+1)p} e^{-\alpha y}}{\Gamma(mp)\,\Gamma(p)} y^{(m+1)p-1} \cdot \int_0^1 t^{mp-1} (1-t)^{p-1} dt.$$

$$= \frac{\alpha^{(m+1)p} e^{-\alpha y}}{\Gamma(mp)\,\Gamma(p)} y^{(m+1)p-1} \cdot \beta(mp, p)$$

$$= \frac{\alpha^{(m+1)p}}{\Gamma[(m+1)p]} e^{-\alpha y} y^{(m+1)p-1}$$

$$\left[\therefore \quad \beta(mp, p) = \frac{\Gamma(mp)\,\Gamma(p)}{\Gamma(m+1)p}\right]$$

$$\therefore \quad y \frown G(\alpha, (m+1)p)$$

இதிலிருந்தும் தொகுத்தறி முறையிலிருந்தும் தேற்றம் தெளிவாகிறது.

26·5. வரிசை அளவைகள் (Order Statistics)

$(x_1, x_2, \dots x_n)$ என்பது ஊக அளவு அடர்த்தி $f(x)$ என்றுள்ள பரவல் ஒன்றிலிருந்து எடுத்த சாதாரண ராண்டம் மாதிரியாக இருக்கட்டும். இம்மாதிரி உறுப்புகளை ஏறு வரிசையில் வரிசைப் படுத்தினால் $(x_{[1]}, x_{[2]}, \dots, x_{[n]})$ என்னும் வரிசைப்படுத்திய மாதிரி கிடைக்கிறது. $x_{[1]}$ என்பது முதல் வரிசை அளவை (First order Statistic) எனப்படும். பொதுவாக $x_{[r]}$ என்பது r ஆவது வரிசை அளவை (r^{th} order statistic) எனப்படும்.

$x_{[n]} - x_{[1]}$ என்பது மாதிரி வீச்சு (Sample range) எனப் படும்.

$x_{[r]}$ -ன் பரவல்

$P[x < x_{[r]} \le x + dx] = P[(r-1)$ மாதிரி உறுப்புகள் $\le x$; ஓர் உறுப்பு $(x, x+dx)$ நடுவில்; $(n-r)$ உறுப்புகள் $x+dx)$-க்கு மேலே.]

$$= \frac{\lfloor n}{\lfloor r-1 \; \lfloor n-r} [F(x)]^{r-1} f(x)\, dx\, [1-F(x)]^{n-r}$$

$$F(x) = P[X \le x] = \int_{\infty}^{x} f(x)\, dx$$ ஆகும்.

ஆகவே $f_{x_{[r]}}(x) = r\binom{n}{r}[F(x)]^{r-1}[1-F(x)]^{n-r}f(x)$

குறிப்பாக $x_{(1)}$-ன் பரவல் : $f_{x_{[1]}}(x) = n[1-F(x)]^{n-1}f(x)$

,, $x_{(n)}$-ன் பரவல் : $f_{x_{[n]}}(x) = n[F(x)]^{n-1}f(x)$.

$x_{[r]}$, $x_{[s]}$ $(r<s)$ ஆகியவையின் இணைந்த பரவல்

$$P\ [x < x_{[r]} \leq x + dx\ ;\ y < x_{[s]} \leq y + dy]$$

$= P$ [$(r-1)$ மாதிரி உறுப்புகள் $< x$; ஓர் உறுப்பு $(x, x+dx)$-க்கு நடுவில்; $(s-r-1)$ உறுப்புகள் $(x+dx, y)$-க்கு நடுவில்; ஓர் உறுப்பு $(y, y+dy)$-க்கு நடுவில்; $(n-s)$ உறுப்புகள் $y+dy$-க்கு மேல்.]

$$= \frac{\lfloor n}{\lfloor r-1\ \lfloor s-r-1\ \lfloor n-s} [F(x)]^{r-1}f(x)dx.$$

$$[F(y)-F(x)]^{s-r-1}f(y)\,dy\,[1-F(y)]^{n-s}$$

குறிப்பாக $[x_{(1)}, x_{(n)}]$-ன் இணைந்த பரவல்

$\frac{\lfloor n}{\lfloor n-2}[F(y)-F(x)]^{n-2}f(x)f(y)$ ஆகும்.

அதாவது $n(n-1)[F(y)-F(x)]^{n-2}f(x)f(y);\ y \geq x$

மேற்கண்டவைகளைத் தனித்த ராண்டம் மாறிகளுக்கும் கையாளலாம்.

பயிற்சி 11

1. $[x_{(1)}, x_{(2)}, \ldots x_{(n)}]$ என்பன $F(x)$ என்னும் பரவல் சார்புடைய தொடர்ச்சியான பரவலிலிருந்து எடுக்கப்பட்ட வரிசை

அளவைகள். மாதிரி வீச்சு $R=x_{(n)}-x_{(1)}$-ன் எதிர்பார்க்கும்

$$\text{மதிப்பு } E(R) = \int_{-\infty}^{\infty} [1-F^n(x)-(1-F(x))^n]dx \text{ என்று நிறுவு.}$$

[Tippet, (1925)]

$$2.\quad R\text{-ன் பரவல் சார்பு} = n\int_{-\infty}^{\infty}[F(x+h)-F(x)]^{n-1}dF(x)$$

என்று நிறுவு.

3. $(x_1, x_2, \ldots, x_n)$ என்பது செவ்வகப் பரவல் $R[0, 1]$-லிருந்து எடுக்கப்பட்ட சாதாரண ராண்டம் மாதிரியானால்

$$y = (x_1\, x_2 \ldots x_n)^{\frac{1}{n}} \text{-ன் பரவல் } f(y) = \frac{n^n\, y^{n-1}}{\lfloor n-1}(-\log y)^{n-1}$$

$0 \leq y \leq 1$ என்று நிறுவு.

4. $[x_{(1)}, x_{(2)}, \ldots, x_{(n)}]$ என்பவை பின்வரும் பரவலிலிருந்து எடுக்கப்பட்ட சாதாரண ராண்டம் மாதிரியின் வரிசை அளவைகள்.

$$f(x) = \frac{1}{\sigma}\, e^{-\left[\frac{x-\mu}{\sigma}\right]} \qquad \mu \leq x \leq \infty.$$

$$W = \frac{1}{(n-1)}\sum_{\xi=2}^{n}[x_{(\xi)}-x_{(1)}]$$

$$\text{இதிலிருந்து } t = \frac{x_{(1)}-\mu}{W} \text{-ன் அடர்த்தி } g(t) \text{ என்பது}$$

$$g(t) = \frac{n}{\left[1+\frac{nt}{n-1}\right]^n} \quad t>0 \text{ என நிறுவு.}$$

[Guttman (1960)]

26·6. பரவற் படியை ஸ்திரப்படுத்தும் மாறி மாற்றம் [Variance Stabilising Transformations)

X என்னும் ராண்டம் மாறியின் சராசரி θ என்றும், பரவற் படி σ_θ^2 என்றும் இருக்கட்டும். பலவித காரணங்களுக்குப் பரவற் படியில் θ என்னும் முழுமைத் தொகுதி அளவை வராமலிருந்தால் நலம். முக்கியமாகப் பெரிய மாதிரியிலிருந்து எடுக்கப்பட்ட அளவை X எனக் கொண்டால் $\dfrac{X - E(x)}{\sqrt{V(x)}}$ என்பது இயல்நிலைப் பரவலைக் கொண்டதாக இருக்கலாம்.

ஆனால், இப்பரவலின் பரவற் படியில் θ இருக்குமாதலால் இதன் போக்கு ஸ்திரமாக இராது. X-க்குப் பதிலாக $f(X)$ என்னும் மாறியைக் கையாண்டால் $V[f(X)]$-ல் θ வராமற் போகக்கூடும். இத்தகைய மாறி மாற்றத்திற்குப் பரவற் படியை ஸ்திரப்படுத்தும் மாறி மாற்றம் என்பது பெயர்.

டெய்லர் தேற்றத்தின்படி,

$$f(X) = f(\theta) + (X - \theta) f'(\theta) + \frac{(X - \theta)^2}{2} f''(\theta) + \cdots$$

இவைகளுள் $(X - \theta)^2, (X - \theta)^3$ முதலியவை மிகவும் சிறியவையென்று ஒதுக்க முடிந்தால்,

$$f(X) \doteq f(\theta) + (X - \theta) f'(\theta)$$

$$\therefore \quad E[f(X)] \doteq f(\theta) \quad [\because E(X - \theta) = 0].$$

$$V[f(X) = [f'(\theta)]^2 V(X) = [f'(\theta)]^2 \sigma_\theta^2.$$

$V[f(X)]$-ல் θ இல்லாமலிருக்க வேண்டுமாயின்

$V[f(X)] = k^2$ எனக் கொள்ளலாம் [k, θ-ன் சார்பற்றது.]

$$\therefore \quad [f'(\theta)]^2 \sigma_\theta^2 = k^2.$$

$$f'(\theta) = \frac{k}{\sigma_\theta}.$$

$\therefore \quad f(\theta) = \displaystyle\int \frac{k}{\sigma_\theta} d\theta$ என எடுத்துக் கொள்ளலாம்.

இதிலிருந்து f என்னும் சார்பையறியலாம்.

உதாரணம் 1

r என்பது ஈருறுப்புப் பரவல் மாறியானால்

$$E(r) = np = \theta.$$

$$V(r) = npq = \theta\left(1 - \frac{\theta}{n}\right)$$

$$\therefore \quad \sigma_\theta = \sqrt{\theta\left(1 - \frac{\theta}{n}\right)}.$$

$$\therefore \quad f(\theta) = \int \frac{k\, d\theta}{\sqrt{\theta\left(1 - \frac{\theta}{n}\right)}} \qquad \theta = n \sin^2 \alpha\,;$$

$$= k \int \frac{2\, n \sin \alpha \cos \alpha \, d\alpha}{\sqrt{n} \sin \alpha \cos \alpha}.$$

$$= 2\, k \sqrt{n} \int d\alpha = 2\, k \sqrt{n}\, \alpha.$$

$$= 2\, k \sqrt{n} \sin^{-1} \sqrt{\frac{\theta}{n}}$$

$$= 2\, k \sqrt{n} \sin^{-1} \sqrt{p}.$$

$$\therefore \quad f(r) = 2\, k \sqrt{n} \sin^{-1} \sqrt{\frac{r}{n}}.$$

$k = \dfrac{1}{2\sqrt{n}}$ எனக் கொண்டால் $f(r) = \sin^{-1} \sqrt{\dfrac{r}{n}}.$

$$\therefore \quad E\left[\sin^{-1} \sqrt{\frac{r}{n}}.\right] \doteq \sin^{-1} \sqrt{P}$$

$$V\left[\sin^{-1} \sqrt{\frac{r}{n}}\right] \doteq k^2 = \frac{1}{4n}.$$

உதாரணம் 2

x என்பது பாய்ஸான் மாறி.

$E(x) = \lambda = \theta$

$V(x) = \lambda = \sigma_{\theta}^{2}$.

$$f(\theta) = \int \frac{k\, d\theta}{\sqrt{\theta}} = 2k\sqrt{\theta}$$

$\therefore \quad k = \frac{1}{2}$ என்று எடுத்துக்கொண்டால்

$$f(\theta) = \sqrt{\theta} \quad \therefore \quad f(x) = \sqrt{x}$$

$$\therefore \quad E(\sqrt{x}) \doteq \sqrt{\lambda} \quad V(\sqrt{x}) = \frac{1}{4}.$$

உதாரணம் 3

ஒட்டுறவுக்கெழு (Correlation coefficient) ρ உள்ள ஓர் இரு பரிமாண இயல்நிலைப் பரவல் தொகுதியிலிருந்து எடுக்கப்பட்ட சாதாரண ராண்டம் மாதிரியின் மாதிரி ஒட்டுறவுக் கெழு r என்பது n என்பது அதிகமென்று கொண்டால்

$E(r) = \rho, \quad V(r) = \frac{(1-\rho^2)^2}{n}$ உடையதாக இருக்கும்.

$$\therefore \quad \theta = \rho, \quad \sigma_{\theta}^{2} = \frac{(1-\rho^2)^2}{n}, \quad \sigma_{\theta} = \frac{1-\rho^2}{\sqrt{n}}$$

$$\therefore \quad f(\rho) = \sqrt{n}\int \frac{k\,d\rho}{\sqrt{1-\rho^2}} = \frac{k\sqrt{n}}{2}\log\left[\frac{1+\rho}{1-\rho}\right]$$

$k = \frac{1}{\sqrt{n}}$ எனக் கொண்டால் $f(\rho) = \frac{1}{2}\log\left[\frac{1+\rho}{1-\rho}\right]$

$$= \tan h^{-1}\rho.$$

$$\therefore \quad f(r) = \tan h^{-1} r.$$

$$\therefore \quad E[\tan h^{-1} r] \doteq \tan h^{-1} \rho.$$

$$V[\tan h^{-1} r] \doteq k^2 = \frac{1}{n}.$$

26·7. 26·6-ல் உள்ளது பல பரிமாண ராண்டம் மாறிகளுக்கும் பொருந்தும்.

$E[x_1, x_2, \ldots, x_n] = (\theta_1, \theta_2, \ldots, \theta_n)$ என்றும்,

$V(x_i) = \sigma_i^2$ என்றும், $\text{cov}(x_i, x_j) = \sigma_{i,j}$ $i \neq j$ என்றும் கொண்டால் $f(x_1, \ldots, x_n)$ என்பதை டெய்லர் தேற்றத்தின்படி

$$f(x_1, \ldots, x_n) \doteq f(\theta_1, \theta_2, \ldots, \theta_n) + \sum_{i=1}^{n} (x_i - \theta_i) \frac{\partial f}{\partial \theta_i}$$

என்று கொள்ளலாம்.

$$\therefore \quad E[f(x_1, \ldots, x_n)] = f(\theta_1, \theta_2, \ldots, \theta_n)$$

$$V[f(x_1, \ldots, x_n)] = \sum_{i=1}^{n} \left[\frac{\partial f}{\partial \theta_i}\right]^2 \sigma_i^2 + 2 \sum_{i \neq j}\sum \left[\frac{\partial f}{\partial \theta_i} \frac{\partial f}{\partial \theta_i}\right] \sigma_{i,j}$$

26·8. மாதிரிவிலக்கப் பெருக்குத்தொகை (Sample Moments)

ஒரு பரவலின் r ஆவது பெருக்குத்தொகை μ_r' எனவும், r ஆவது சராசரியைச் சுற்றிய பெருக்குத்தொகை μ_r எனவும் கொள்வோம். இப்பரவலிலிருந்து $(x_1, x_2, \ldots x_r)$ என்னும் சாதாரண ராண்டம் மாதிரிக்குப் பெருக்குத் தொகைகளைப் பின் வருமாறு வரையறுக்கலாம்.

$$m_r' = \frac{1}{n} \sum_{i=1}^{n} x_i^r \qquad m_1' = \frac{1}{n} \sum_{i=1}^{n} x_i = \bar{x}$$

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

இதிலிருந்து $E(m_r') = \frac{1}{n} \sum_{i=1}^{n} E(x_i^r) = \frac{1}{n} n \mu_r' = \mu_r'$

$$V(m_r') = E(m_r'^2) - [\mu_r']^2$$

$$= \frac{1}{n^2} \left[\sum_{i=1}^{n} E(x_i^{2r}) + \sum_{i \neq j} \sum E(x_i^r x_j^r) \right] - [\mu_r']^r$$

$$= \frac{1}{n^2} n \mu_{2r}' + \frac{1}{n^2} n(n-1) \mu_r'^2 - \mu_r'^2$$

$$= \frac{1}{n} [\mu_{2r}' - \mu_r'^2]$$

$E(m_r)$, $V(m_r)$:

$$m_r = \frac{1}{n} \sum_{r=1}^{n} (x_i - \bar{x})^r = \frac{1}{n} \sum_{r=1}^{n} [(x_i - \mu_1') - (\bar{x} - \mu_1')]^r$$

$$= \frac{1}{n} \sum_{r=1}^{n} (x_i - \mu_1')^r - \frac{1}{n} \binom{r}{1} \sum (x_i - \mu_1')^{r-1} (\bar{x} - \mu_1') + \dots$$

$$\therefore \quad E(m_r) = \frac{1}{n} n \, \mu_r -$$

$$\binom{r}{1} \frac{1}{n} \sum_{i=1}^{n} E\,[\,(x_i - \mu_1')^{r-1} \;(\bar{x} - \mu_1')\,] + \cdots$$

$$= \mu_r - \binom{r}{1} \frac{1}{n^2} \sum_{i=1}^{n} E\,(x_i - \mu_1')^{r-1}$$

$$\sum_{i=1}^{n} (x_i - \mu_1') + \cdots$$

$$= \mu_r - \frac{1}{n} \mu_r + 0\left(\frac{1}{n^2}\right)$$

$$= \mu_r + 0\left(\frac{1}{n}\right)$$

$$\therefore \quad V(m_r) \doteq E\,(m_r - \mu_r)^2$$

$$\doteq E\left[\frac{1}{n} \sum_{i=1}^{n} (x_i - \mu_1')^r - \mu_r - \right.$$

$$\left. \binom{r}{1} \frac{(\bar{x} - \mu_1')}{n} \sum_{i=1}^{n} (x_i - \mu_1')^{r-1} \right]^2$$

$$= E\left[\frac{1}{n} \sum_{i=1}^{n} (x_i - \mu_1')^r - \mu_r - \right.$$

$$\left. \frac{r}{n^2} \left\{ \sum (x_i - \mu_1')^r + \sum_{i \neq j}\sum (x_i - \mu_1')^{r-1} \; (x_j - \mu_1') \right\} \right]^2$$

$(x_i - \mu_1')^r$ என்பதை a_{ri} என்றெழுதினால்,

$$V(m_r) = E\left[\frac{1}{n}\sum_{i=1}^{n} a_{ri} - \mu_r - \frac{r}{n^2}\left\{\sum_{i=1}^{n} a_{ri} + \sum\sum_{i \neq j} a_{(r-1)i}\, a_{1j}\right\}\right]^2$$

$$= E\left[\frac{n-r}{n^2}\sum_{i=1}^{n} a_{ri} - \mu_r - \frac{r}{n^2}\sum\sum_{i \neq j} a_{(r-1)i}\, a_{1j}\right]^2$$

$$= E\left[\left(\frac{n-r}{n^2}\right)^2\left\{\sum a_{ri}^2 + \sum\sum_{i \neq j} a_{ri}\, a_{rj}\right\} + \mu_r^2 + \frac{r^2}{n^4}\sum\sum\sum a^2_{(r-1)i}\, a_{ij}^2 + \frac{r^2}{n^2}\sum\sum\sum a_{(r-1)i}\, a_{ij}^2 + \cdots\cdots\right]$$

$$= \frac{1}{n}\left[\mu_{2r} - 2r\,\mu_{r-1}\,\mu_{r+1} - \mu_r^2 + r^2\,\mu_2\,\mu_{r-1}^2\right] + 0\left(\frac{1}{n^2}\right).$$

இதைப்போலவே

$$\text{cov}(m_r, m_s) = \frac{1}{n}\left[\mu_{r+s} - r\,\mu_{r-1}\,\mu_{s+1} - s\,\mu_{r+1}\,\mu_{s-1} - \mu_r\,\mu_s + rs\,\mu_{r-1}\,\mu_{s-1}\right] + 0\left(\frac{1}{n^2}\right).$$

குறிப்பாக

$$V(m_2) = \frac{1}{n}\left[\mu_4 - 4\,\mu_1\,\mu_3 - \mu_2^2 + 4\,\mu_2\,\mu_1^2\right] + 0\left(\frac{1}{n^2}\right).$$

$$= \frac{1}{n}\left[\mu_4 - \mu_2{}^2 \right] + 0\left(\frac{1}{n^2}\right).$$

$$\text{cov}(m_1, m_2) = \frac{1}{n}\left[\mu_3 - \mu_0\mu_3 - 2\mu_2\mu_1 - \mu_1\mu_2 + 2\mu_0\mu_1 \right] + 0\left(\frac{1}{n^2}\right).$$

$$= 0\left(\frac{1}{n^2}\right).$$

மாதிரியின் மாறுபாடுக்கெழு (Coefficient of variations)

$$C_x = \frac{\sqrt{m_2}}{m_1'} = f_0(m_1', m_2), = f(m_1', m_2')$$

$$= \frac{\sqrt{m_2' - m_1'^2}}{m_1'}$$

$$E(m_1') = \mu_1' \qquad E(m_2') = \mu_2'$$

$$V(m_1') = \frac{\mu_2' - \mu_1'^2}{n} \qquad V(m_2') = \frac{\mu_4' - \mu_2'}{n}$$

$$f(m_1', m_2') = f(\mu_1', \mu_2') + \left[(m_1', \mu_1') \frac{\partial f}{\partial m_1'} + (m_2' - \mu_2') \frac{\partial f}{\partial m_2'} \right]$$

$$\therefore \quad V[C_x] = V[f(m_1', m_2')] = V(m_1')\left[\frac{\partial f}{\partial \mu_1'}\right]^2$$

$$+ V(m_2')\left[\frac{\partial f}{\partial \mu_2'}\right]^2 + 2\,\text{cov}(m_1', m_2')\left[\frac{\partial f}{\partial \mu_1'}\;\frac{\partial f}{\partial \mu_2'}\right]$$

$$= \frac{\mu_1'^2(\mu_4 - \mu_2{}^2) - 4\mu_1'\mu_2\mu_3 + 4\mu_2{}^3}{4\mu_1'^4\mu_2 n}$$

[$\mu_1' \neq 0$ எனக் கொண்டால்].

பயிற்சி 11

பின் வருவனவற்றை நிறுவு.

1. $g_1 = \dfrac{m_3}{m_2^{\frac{3}{2}}}$; $g_2 = \dfrac{m_4}{m_2^2} - 3$ என்றால்,

$$E(g_1) \doteq r_1,\ E(g_2) = r^2$$

$$V(g_1) \doteq \frac{4\mu_2^2\mu_6 - 12\mu_2\mu_3\mu_5 - 24\mu_2^2\mu_4 + 9\mu_3^2\mu_4 + 35\mu_2^2\mu_3^2 + 36\mu_2^5}{4n\mu_2^5}$$

$$V(g_2) = \frac{\mu_2^2\mu_8 - 4\mu_2\mu_4\mu_6 - 8\mu_2^2\mu_3\mu_5 + 4\mu_4^3 - \mu_2^2\mu_4^2 + 16\mu_2\mu_3^2\mu_4 + 16\mu_2^3\mu_3^2}{n\mu_2^6}$$

பரவல் இயல்நிலைப்பரவலாயின்

$E(g_1) = E(g_2) = 0$ என்றும், $V(g_1) = \dfrac{6}{n}$, $V(g_2) = \dfrac{24}{n}$

என்றும் நிறுவு.

2. இரு பரிமாணப் பரவலிலிருந்து எடுக்கப்பட்ட சாதாரண ராண்டம் மாதிரியின் மாதிரி உடன் தொடர்புக் கெழு (Correlation Coefficient) r ஆயின் பின்வருவனவற்றை நிறுவு.

$$E(r) = \rho$$

$$V(r) = \frac{\rho^2}{4n}\left[\frac{\mu_{40}}{\mu_{20}^2} + \frac{\mu_{04}}{\mu_{02}^2} + \frac{2\mu_{22}}{\mu_{20}\mu_{02}} + \frac{4\mu_{22}}{\mu_{11}^2} - \frac{4\mu_{31}}{\mu_{11}\mu_{20}} - \frac{4\mu_{13}}{\mu_{11}\mu_{02}}\right]$$

பயிற்சி 12

[பலவிதக் கணக்குகள்]

1. 1-லிருந்து 72 வரையுள்ள எண்களிலிருந்து ஓர் எண்ணைச் சமவாய்ப்புடன் எடுக்கவேண்டும். ராண்டம் எண்கள் பட்டியலை இதற்கு எவ்வாறு பயன்படுத்துவாய் ?

2. 759 சீட்டுகளிலிருந்து 100 சீட்டுகளைச் சமவாய்ப்புகளுடன் தேர்ந்தெடுக்கவேண்டும். ராண்டம் எண்கள் பட்டியலை எவ்வாறு இதற்குப் பயன்படுத்துவாய் ?

3. 1-லிருந்து 100 வரையுள்ள எண்களிலிருந்து 10 எண்களைச் சம வாய்ப்புகளுடன் எடுக்கவேண்டும். இரண்டு இலக்க ராண்டம் எண்கள் பட்டியலை இதற்கு எவ்வாறு பயன்படுத்த முடியும் ?

4. கணித முறையைப் பயன்படுத்தாமல் தர்க்க ரீதியாக முறையுடை மாதிரி மதிப்பீட்டின் பரவற் படி முழுமைத் தொகுதியின் வரிசை முறையைப் பொறுத்திருக்கிறது என்பதை எவ்வாறு நிறுவுவாய் ?

5. ஒரு பட்டியலில் ஒரு வட்டத்திலுள்ள மக்களின் பெயர்களுடன் ஆணு, பெண்ணு? வயது, குடும்பத் தலைவருடன் உறவு முறை, பள்ளிக்குப் போதல், தொழில் முதலிய விவரங்கள் தரப்பட்டுள்ளன. ஆனல் சில பெயர்கள் பலவித காரணங்களுக்காகத் தராமல் விடப்பட்டிருக்கின்றன. இப்பட்டியலிலிருந்து 5 சதவீத மாதிரியொன்று எடுக்கப்படவிருக்கிறது. இதற்காக 5 சதவீதக் குடும்பங்கள் தேர்ந்தெடுக்கப்பட்டு அக் குடும்பத்திலுள்ள யாவரும் சேர்த்துக் கொள்ளப்படுகின்றனர்.

பின்வரும் 8 முறைகளில் எது சிறந்தது என்பதைக் காரணங்களுடன் விளக்குக.

(அ) ராண்டம் துவக்கத்துடன் ஒவ்வொரு 20 ஆவது குடும்பம்.

(ஆ) ,, ,, ,, ,, ,, ,, ,, பெயர்.

(இ) ,, ,, ,, ,, ,, ,, ,, குடும்பத் தலைவர்.

(ஈ) ,, ,, ,, ,, ,, , ,, பக்கத்திலுள்ள எல்லா மக்கள்

(உ) ஒவ்வொரு 20 குடும்பங்களிலிருந்தும் ஒரு குடும்பம் ராண்டம் முறையில்

(ஊ) ,, ,, பெயர்களிலிருந்தும் ஒரு பெயர் ராண்டம் முறையில்.

(எ) ,, ,, குடும்பத் தலைவர்களிலிருந்தும் ஒரு குடும்பத் தலைவர் ராண்டம் முறையில்.

(ஏ) ஒவ்வொரு 20 பக்கங்களிலிருந்தும் ஒரு பக்கம் ராண்டம் முறையில். அப் பக்கத்திலுள்ள எல்லாக் குடும்பத் தலைவர்களும்.

6. ஒரு சீட்டுத் தொகுதியில் ஒவ்வொரு சீட்டிலும் ஒரு விலாசம் தரப்பட்டிருக்கிறது. சில விலாசங்களில் ஒரு குடும்பமும், சில விலாசங்களில் இரு குடும்பங்களும், சில விலாசங்களில் 3 குடும்பங்களும், சிலவற்றில் 10 குடும்பங்களும் உள்ளன. இவ் விலாசங்களிலுள்ளோர்களின் மொத்த வருமானம் A மதிப்பிடப்பட வேண்டியிருக்கிறது.

பின்கண்ட முறைகளில் A மதிப்பிடப்படுகிறது.

(அ) மொத்தமுள்ள M சீட்டுகளிலிருந்து m சீட்டுகள் சாதாரண ராண்டம் மாதிரி முறையில் எடுக்கப்படுகின்றன. i ஆவது சீட்டிலுள்ளோரின் மொத்த வருமானம் x_i ஆனால்

$$\hat{X} = \frac{M}{m} \sum_{i=1}^{m} x_i$$

$\hat{X}$, A-ன் பிறழ்ச்சியற்ற மதிப்பீடென்று நிறுவு.

(ஆ) மொத்தமுள்ள M சீட்டுகளிலிருந்து m சீட்டுகள் சாதாரண ராண்டம் மாதிரி முறையில் எடுக்கப்படுகின்றன. எடுக்கப்பட்ட சீட்டுகள் ஒவ்வொன்றிலிருந்தும் ஒரு குடும்பம் சாதாரண ராண்டம் மாதிரி முறையில் தேர்ந்தெடுக்கப்படுகிறது. y_i என்பது இக்குடும்பத்தின் வருமானமாயின்,

$$\hat{X} = \frac{N}{n} \sum_{i=1}^{m} x_i \qquad [N = \text{மொத்தக் குடும்பங்கள்.}]$$

என்பது A-ன் மதிப்பீட்டியாக எடுத்துக் கொள்ளப்படு கிறது.

$\hat{X}$, A-ன் பிறழ்ச்சியுடைய மதிப்பீட்டியென நிறுவு. ஆனால் இப் பிறழ்ச்சி அவ்வளவு அதிகமாக இராதென்றும் நிறுவு.

(இ) மொத்தமுள்ள M சீட்டுகளிலிருந்து m சீட்டுகள் சாதாரண ராண்டம் மாதிரி முறையில் எடுக்கப்படுகின்றன. இவைகளில் i ஆவது சீட்டில் N_i குடும்பங்கள் உள்ளன. இந்த

N_i குடும்பங்களிலிருந்து ஒரு குடும்பம் சாதாரண ராண்டம் மாதிரி முறையில் எடுக்கப்படுகிறது. இக் குடும்பத்தின் வருமானம் y_i

$$\hat{X} = \frac{M}{m} \sum_{i=1}^{m} N_i y_i$$

என்பது A-ன் பிறழ்ச்சியற்ற மதிப்பீடென்று நிறுவு.

$$\hat{X} = \frac{M}{m^2} \left[\sum_{i=1}^{m} N_i \right] \left[\sum_{i=1}^{m} y_i \right]$$

என்பது A-ன் பிறழ்ச்சியுடைய மதிப்பீடென்று நிறுவு.

7. (அ) மேற்கூறப்பட்ட வினாவில் M சீட்டுகளுக்கு முறையே $P_1, P_2, \ldots, P_m$ என்னும் நிகழ்திறன்களையளித்து m சீட்டுகள் எடுக்கப்படுகின்றன. A_i என்பது i ஆவது சீட்டிலுள்ளோரின் மொத்த வருமானம். $A = \sum_{i=1}^{M} A_i$ என்பது மொத்த வருமானம்.

$$\hat{X} = \frac{1}{m} \sum_{i=1}^{m} \frac{x_i}{P_i}$$ என்பது A-ன் பிறழ்ச்சியற்ற மதிப்பீட்டியென நிறுவு. x_i என்பது i ஆவது மாதிரிச் சீட்டிலுள்ளோரின் மொத்த வருமானம்.

(ஆ) (அ)-ல் $m = 1$ எனக் கொண்டால் P_i-கள் A_i-களின் விகித சமமாக இருந்தால்தான் $V(\hat{X})$ என்பது சிறுமமாக இருக்குமென நிறுவு.

8. வங்கியில் கணக்கு வைத்துக் கொண்டிருப்பவர்களுக்காக அட்டவணையொன்று இருக்கிறது. சிலருக்கு இவ்வட்டவணையில் ஒரே ஒரு பக்கம் மட்டும் ஒதுக்கியிருக்கப்பட்டிருக்கிறது. மற்றவர்களுக்கு அவர்கள் கொடுக்கல் வாங்கல்களுக்கேற்ப இரண்டு, மூன்று, நான்கு என்று பக்கங்கள் ஒதுக்கப்பட்டிருக்கின்றன.

சில வங்கி உறுப்பினர்களைத் தேர்ந்தெடுக்க வேண்டியிருக்கிறது. எல்லாப் பக்கங்களுக்கும் சமவாய்ப்பு அளிக்கப்பட்டால் உறுப்பினர்களுக்குச் சமவாய்ப்பு அளிக்க முடியாது ; ஏனெனில் அதிகப் பக்கம் கொண்ட உறுப்பினர் மாதிரியில் வருவதற்கு நிகழ்திறன் அதிகமாக இருக்கும். 2 சதவீத மாதிரி தேவைப் பட்டால் பிற்கண்ட முறைகளை ஒப்பிடு.

(அ) ராண்டம் துவக்கத்திலிருந்து ஒவ்வொரு 50 ஆவது உறுப்பினர் தேர்ந்தெடுக்கப்படுகிறார்.

(ஆ) ராண்டம் துவக்கத்திலிருத்து ஒவ்வொரு 50 ஆவது பக்கம் தேர்ந்தெடுக்கப்பட்டு அப்பக்கம் ஓர் உறுப்பினரின் முதல் பக்கமாக இருந்தால் அவ்வுறுப்பினர் தேர்ந்தெடுக்கப்படுகிறார்.

9. ஒரு மின்சார நிறுவனத்தினரிடம் 9,800 பக்கங்கள் கொண்ட மின்சார விளக்குக் கம்பங்களைப் பற்றிய விவரங்கள் அடங்கிய அட்டவணையொன்றுள்ளது. சில பக்கங்களில் ஒரே ஒரு விளக்குக் கம்பமும், சில பக்கங்களில் இரண்டு விளக்குக் கம்பங்களும், மற்றவைகளில் 25 கம்பங்கள் வரையும் உள்ளன. விளக்குக் கம்பங்களைப் பரிசோதிப்பதற்கு ஒரு மாதிரி தேவைப்படு கிறது. மொத்தம் 1,05,000 விளக்குக் கம்பங்கள் உள்ளன. மாதிரியில் 1,000 கம்பங்கள் தோராயமாக இருக்கவேண்டுமென்று தீர்மானிக்கப்பட்டது. ஒவ்வொரு கம்பத்திற்கும் 'சதவீத நிலைமை' என்று ஒன்று அளக்க முடியும். இது 0-விருந்து 100 வரை இருக்க முடியும். மொத்தக் கம்பங்களின் சராசரி சதவீத நிலைமை மதிப்பிடப்பட வேண்டியிருக்கிறது. இதற்காக இரு பார்வை யாளர்கள் மாதிரிக் கம்பங்களைப் பார்வையிட அனுப்பப்பட்டனர். 9,800 பக்கங்களில் 1;05,000 கம்பங்கள் இருப்பதால் ஒரு பக்கத் திற்கு சராசரி 10·7 கம்பங்கள் உள்ளதாகக் கொள்ளலாம். ஆகவே பத்துப் பக்கத்திற்கு ஒரு கம்பமென்று தேர்ந்தெடுத்தால் போதுமானது.

(அ) 1-லிருந்து 10 வரையுள்ள எண்களிலிருந்து ராண்டம் துவக்கம் ஒன்று எடுத்து அதிலிருந்து ஒவ்வொரு பத்தாவது பக்கத்தையும் எடுத்து அப் பக்கத்திலிருந்து ஒரு கம்பம் ராண்டம் முறையில் எடுக்கப்படுகிறது. இம் முறை எல்லாக் கம்பங்களுக்கும் சம வாய்ப்பளிக்காது என நிறுவு.

(ஆ) m என்பது மாதிரியிலுள்ள பக்கங்கள். x_i என்பது i ஆவது மாதிரிக் கம்பத்தின் 'சதவீத நிலைமை'.

$$\bar{x}_1 = \frac{1}{m}\sum_{i=1}^{m} x_i$$ என்பது பிறழ்ச்சியுடைய மதிப்பீட்டியென

நிறுவு. ஆனால் இப் பிறழ்ச்சி குறைவாகத்தானிருக்குமெனவும் நிறுவு.

(இ) 1-லிருந்து 105 வரையுள்ள எண்களிலிருந்து ஓர் எண்ணை ராண்டம் துவக்கமாக எடுத்து அதிலிருந்து ஒவ்வொரு 105 ஆவது பக்கத்தையும் எடுத்து அதிலுள்ள எல்லாக் கம்பங்களையும் எடுத்தால்,

$$\bar{x}_2 = \frac{1}{n}\sum_{i=1}^{n} x_i$$ [n என்பது மாதிரியிலுள்ள மொத்தக்

கம்பங்கள்: x_i என்பது i ஆவது மாதிரிக் கம்பத்தின் சதவீத நிலைமை.]

இது சிறிதளவே பிறழ்ச்சியுடைய மதிப்பீடென நிறுவு. ஆனால் இதன் பரவற் படி, அடுத்தடுத்துள்ள கம்பங்களின் நிலைமை அதிகத் தொடர்புடையதாக இருக்கும் என்ற காரணத்தால், அதிகமாகவே இருக்குமென நிறுவு.

(ஈ) ஒன்றிலிருந்து 105 வரையுள்ள எண்களிலிருந்து ஒரு ராண்டம் துவக்கம் எடுத்து அத் துவக்கத்திலிருந்து ஒவ்வொரு 105 ஆவது கம்பமும் மாதிரியில் எடுத்துக் கொள்ளப்படுகிறது. x_i என்பது i ஆவது மாதிரிக் கம்பத்தின் சதவீத நிலைமையாயின்,

$$\bar{x}_3 = \frac{1}{1000}\sum_{i=1}^{1000} x_i$$ என்பது பிறழ்ச்சியில்லாத மதிப்பீட்டி

யென நிறுவு. இது (இ)-லுள்ள மதிப்பீட்டியை விடப் பரவற் படியில் குறைந்ததாக இருக்கும் என்பதை நிறுவு.

(உ) முதலில் 1-லிருந்து 5 வரையுள்ள எண்களிலிருந்து ஓர் எண்ணை ராண்டம் துவக்கமாக எடுத்து ஒவ்வொரு 5 ஆவது பக்கத்திலுள்ள எல்லாக் கம்பங்களும் வரிசையாக எடுத்துக் கொள்ளப்படுகின்றன. இது முதல் கட்டம். இம் முதல் கட்டக் கம்பங்களிலிருந்து, ஒன்றிலிருந்து 21 வரையுள்ள எண்களிலிருந் து

ஓர் எண் ராண்டம் துவக்கமாக எடுத்து, ஒவ்வொரு 21ஆவது கம்பமும் எடுத்துக் கொள்ளப்படுகிறது. இம் முறையிலும் வேண்டிய மாதிரிப் பருமன் கிடைக்குமென்பதை நிறுவு.

(ஊ) ஒவ்வொரு 10 ஆவது பக்கத்திலிருந்தும் ஒரு கம்பம் ராண்டம் முறையில் தேர்ந்தெடுக்கப்படுகிறது. N_i என்பது i ஆவது கம்பம் தேர்ந்தெடுக்கப்பட்ட பக்கத்திலுள்ள மொத்தக் கம்பங்களென்றும், x_i என்பது i ஆவது தேர்ந்தெடுக்கப்பட்ட கம்பத்தின் சதவீத நிலைமை என்றும் ஆனுல்,

$$\bar{x} = \frac{\sum N_i x_i}{\sum N_i}$$ என்னும் விகித மதிப்பீடு சற்று பிறழ்ச்சியுடைய மதிப்பீடென நிறுவு.

10. 9 ஆம் கணக்கில் (ஆ) பிரிவிலுள்ள மதிப்பீட்டியின் பிறழ்ச்சியை

$$\frac{1}{M\overline{N}} \sum_{1}^{M} (\overline{N} - N_i)\, \overline{a}_i$$ என நிறுவு

$\overline{N} = \frac{N}{M}$, N என்பது மொத்தக் கம்பங்கள். $\overline{a}_k$ என்பது முழுமைத் தொகுதியில் k ஆவது பக்கத்திலுள்ள N_k கம்பங்களின் சராசரிச் சதவீத நிலைமை.

11. அமெரிக்காவிலுள்ள நகரங்களில் ஒவ்வொரு வீட்டிலுமுள்ள சராசரி நபர் எண்ணிக்கையின் மாறுபாட்டுக் கெழு (Coefficient of Variation) 0·7-லிருந்து 0·9 வரை உள்ளது. இதைத் தோராயமாக ஒன்று என்று எடுத்துக்கொள்வோம். 40,000 வீடுகளையுடைய நகரம் ஒன்றின் மக்கட்தொகை 1% மாறுபாட்டுக் கெழுவுடைய மதிப்பீட்டியினுல் மதிப்பிடப்படவேண்டுமாயின் எத்தனை வீடுகளை மாதிரியில் எடுக்கவேண்டுமென்பதைக் கண்டுபிடி. ஈ. செயா. சா. ரா. மா. முறையில் வீடுகள் எடுக்கப்படுகின்றன என்று கொள்.

விடை : $n = 8000$.

12. ஓர் இயந்திரத்தின் திறன் 80%-லிருந்து 100% வரை இருக்க முடியும்.

(அ) திறனின் பரவல் செவ்வகப் பரவல் $R[80, 100]$ எனக் கொண்டு, எல்லா இயந்திரங்களின் மொத்தத் திறனை, 3σ-எல்லை உண்மை மதிப்பின் $1\frac{1}{2}$ சதவீத வேறுபாட்டிலேயே இருக்கும் படியான மதிப்பீட்டியைச் சாதாரண ராண்டம் மாதிரி முறையில் அமைக்கவேண்டுமாயின் மாதிரிப் பருமன் 133 தேவைப் படுமென நிறுவு.

(ஆ) திறனின் பரவல் பின் கூறப்பட்ட முக்கோணப் பரவலாயின் மாதிரிப் பருமன் 89 போதுமென நிறுவு.

x என்பது திறனானல்

$$f(x) = \frac{x-80}{100}; \qquad 80 \leq x \leq 90$$

$$= \frac{100-x}{100}; \qquad 90 \leq x \leq 100.$$

(இ) திறனின் பரவல் பின்கூறப்பட்ட இயல்நிலைப் பரவலானல் மாதிரிப் பருமன் 45 போதுமென நிறுவு.

x என்பது திறனானல் $x \frown N\left(90, \frac{100}{9}\right)$.

13. புகையிலைக் கட்டுகளின் எடைகளுக்கு மாறுபாட்டுக் கெழு 5% என்பது அனுபவத்திலிருந்து தெரிய வருகிறது. 1,267 கட்டு களின் மொத்த எடை மதிப்பிடப்படவேண்டியிருக்கிறது. இம் மதிப்பீட்டி அதனுடைய 3σ எல்லைக்குள் இருப்பது உண்மையான மதிப்பினின்று $1\frac{1}{2}$ சதவீத மாறுபாட்டுடன் இருப்பதற்குச் சமமாக இருக்கவேண்டுமாயின் மாதிரிப் பருமன் $n = 93$ என்று நிறுவு. ஈ. செயா. சா. ரா. மா. முறை கையாளப்படுகிறது என்று கொள்ளவும்.

14. ஒரு முழுமைத் தொகுதியின். விகிதசமம் P (Population proportion) மதிப்பிடுவதற்கு ஈ. செ. சா. ரா. மா. முறையில் $\frac{r}{n} = p$ என்னும் மாதிரிப் பாகம் உபயோகப்படுத்தப்படுகிறது. $q = 1 - p$ என்றால்

$$\frac{p\text{-ன் மாறுபாட்டுக்கெழு}}{q\text{-ன் மாறுபாட்டுக்கெழு}} = \frac{Q}{P}$$ என நிறுவு. $[Q = 1 - P]$.

15. ஹைபர் ஜியாமெட்ரிக் பரவல் (Hyper geometric distribution) என்பது பின்வருமாறு வரையறுக்கப்படுகிறது.

$$P[X = r] = \frac{\binom{Np}{r}\binom{Nq}{n-r}}{\binom{N}{n}} \quad r = 0, 1, 2, \ldots n.$$

$[p + q = 1]$.

இப்பரவலின் சராசரி np என்றும் பரவற் படி $\frac{N-n}{N-1}\, npq$ என்றும் நிறுவு.

16. மேற்கூறப்பட்டப் பரவலில் $N \to \infty$ என்றால், பரவல் ஈருறுப்புப் பரவலை எல்லையாகக் கொண்டிருக்கும் என நிறுவு.

[M என்னும் எண் பெரியதாயிருந்தால் $\underline{|M} \doteq \sqrt{2\pi M}\, M^{M} C^{-M}$ எனக் கொள்]. [இது ஸ்டர்லிங் தோராயம் எனப்படும்].

17. $\bar{X} = 0, \bar{Y} = 0$ என்றிருக்கும் இரு பரிமாணப் பரவலில் உடன்தொடர்புக் கெழு $\rho = \frac{E[XY]}{\sqrt{V(X)\, V(Y)}}$ ஆகும். இப்பரவலில் N இருமைகள் (Pairs) உள்ளன எனக் கொள்ளவும். இப் பரவலிலிருந்து n உறுப்புகள் ஈ. செயா. சா. ரா. மா. முறையில் எடுக்கப்படுகின்றன. $(\bar{X}, \bar{Y})$ என்பன மாதிரிச் சராசரிகளானால்

$$\frac{E(\bar{x}\,\bar{y})}{\sqrt{V(\bar{x})\;V(\bar{y})}} = \rho$$ என நிறுவு.

விடை

(X_1, Y_1), (X_2, Y_2), ..., (X_N, Y_N) என்பது முழுமைத் தொகுதி உறுப்புகளானால்

$$\bar{x} = \frac{1}{n} \sum_{i=1}^{N} \delta_i X_i \qquad \bar{y} = \frac{1}{n} \sum_{i=1}^{N} \delta_i Y_i$$

δ_i என்பது பின்வரும் ராண்டம் மாறி.

$\delta_i = 1$, (X_i, Y_i) என்னும் உறுப்பு மாதிரியில் இருந்தால்.
$= 0$, ,, ,, ,, ,, இல்லையேல்.

இதிலிருந்து $E(\delta_1)\, E(\delta_s{}^2) = \dfrac{n}{N}$ என்பதும்

$$E(\delta_i\, \delta_j) = \frac{n}{N}\left(\frac{n-1}{N-1}\right) \qquad (i \neq j)$$

என்பதும் தெளிவு.

$$\therefore \quad E(\bar{x}\,\bar{y}) = \frac{1}{n^2}\, E\left[\sum_{i=1}^{N} \delta_i X_i\right] \left[\sum_{i=1}^{N} \delta_i Y_i\right]$$

$$= \frac{1}{n^2}\, E\left[\sum_{i=1}^{N} \delta_i^2 X_i Y_i + \sum_{i \neq j}\sum \delta_i\, \delta_j X_i Y_j\right]$$

$$= \frac{1}{n^2}\left[\frac{n}{N} \sum_{i=1}^{N} X_i Y_i + \frac{n(n-1)}{N(N-1)} \sum_{i \neq j}\sum X_i Y_j\right]$$

$$= \frac{1}{nN}\left[\sum_{i=1}^{N} X_i Y_i + \frac{n-1}{N-1}\right.$$

$$\left.\left\{\sum_{i=1}^{N} X_i \sum_{i=1}^{N} Y_i - \sum_{i=1}^{N} X_i Y_i\right\}\right]$$

$$\left[\because \sum X_i = \sum Y_i = 0\right]$$

$$= \frac{1}{nN}\left[1 - \frac{n-1}{N-1}\right] \sum_{i=1}^{N} X_i Y_i$$

$$= \frac{1}{nN}\,\frac{N-n}{N-1} \sum_{i=1}^{N} X_i Y_i = \frac{N-n}{N-1}\,\frac{1}{n}\,E(XY)$$

ஆனல் $V(\bar{x}) = \dfrac{N-n}{N-1}\,\dfrac{1}{n}\,\sigma_x^2 \quad V(\bar{y}) = \dfrac{N-n}{N-1}\,\dfrac{1}{n}\,\sigma_y^2$

$$\therefore \quad \frac{E(\bar{x}\,\bar{y})}{\sqrt{V(\bar{x})\,V(\bar{y})}} = \frac{\dfrac{N-n}{N-1}\,\dfrac{1}{n}\,E(X\,Y)}{\dfrac{N-n}{N-1}\,\dfrac{1}{n}\,\sigma_x\,\sigma_y} = \frac{E(XY)}{\sigma_x\,\sigma_y} = \rho$$

(நி-து.)

18. N_1, N_2 வீடுகளுள்ள இரு பகுதிகள் உள்ளன. முதலில் முதல் பகுதி மட்டும் முழுமைத் தொகுதியின் பாகத்திற்கு மதிப்பீட்டியின் மாறுபாட்டுக் கெழு C உள்ளதாக எடுத்துக் கொள்ளப்பட்டது. இதன் பிறகு இரு பகுதிகளையும் இணைத்து அதே C ஐ மாறுபாட்டுக் கெழுவாகக் கொண்ட மதிப்பீட்டியை உருவாக்க இரண்டாம் பகுதியில் எத்தனை உறுப்புகள் எடுக்கப் படவேண்டும்? ஈ. செயா. சா. ரா. மா. முறை கையாளப்படு வதாகக் கொள்ளவும்.

விடை

$$C^2 = \frac{N_1 - n_1}{N-1} \frac{\sigma_1{}^2}{n_1} \cdot \frac{1}{\mu_1{}^2};$$ μ_1, $\sigma_1{}^2$ என்பவை முதற் பகுதியின் சராசரி, பரவற் படி.

$$C^2 = \frac{N_1{}^2 \frac{N_1 - n_1}{N_1 - 1} \frac{\sigma_1{}^2}{n_1} + N_2{}^2 \frac{N_2 - n_2}{N_2 - 1} \frac{\sigma_2{}^2}{n_2}}{(N_1 \mu_1 + N_2 \mu_2)^2}$$

இவைகளிலிருந்து

$$\frac{N_2}{n_2} = 1 + \left(\frac{\sigma_1}{\sigma_2} \frac{\mu_2}{\mu_1}\right)^2 \left(\frac{2\mu_1}{\mu_2} + \frac{N}{N_1}\right) \left(\frac{N_1}{n_1} - 1\right)$$

18. $a_1, a_2, \ldots, a_n$ என்று N உறுப்புகளுள்ள முழுமைத் தொகுதியிலிருந்து $x_1, x_2, \ldots, x_n$ என்னும் மாதிரி சா. ரா. மா. முறையில் எடுக்கப்படுகிறது. $x = \sum_{i=1}^{n} x_i$ என்ருல்

$$E(x) = n\bar{a} \quad \bar{a} = \frac{1}{N} \sum_{i=1}^{N} a_i$$

$$V(x) = n\sigma^2 [1 + (n-1) \rho] \quad \sigma^2 = \frac{1}{N} \sum a_i^2 - \bar{a}^2$$

$$\rho = \rho x_i, \; x_j \qquad i \neq j$$

என்பவற்றை நிறுவு.

விடை

$$E(x) = E \sum_{i=1}^{n} x_i = n\, E(x_i)$$

$$= n \frac{1}{N} \sum_{i=1}^{N} a_i = n\bar{a}$$

$$V(x) = V\left[\sum_{i=1}^{n} x_i\right] = \sum_{i=1}^{n} V(x_i) + \sum\sum_{i \neq j} \text{cov}(x_i, x_j)$$

$$= n\sigma^2 + \sum\sum_{i \neq j} \rho\, \sigma_{xi}\, \sigma_{xj} \qquad \sigma_{xi} = \sigma_{xj} = \sigma$$

என்பதால் $V(x) = n\sigma^2 + n(n-1)\rho\sigma^2$

$$= n\sigma^2 [1 + (n-1)\rho].$$

கவனிக்கவும்

ஈ. செ. சா. ரா. மா. முறையில் $\rho = 0$. ஆகவே $V(x) = n\sigma^2$

ஈ. செயா. சா. ரா. மா. முறையில்

$$\text{cov}(x_i, x_j) = \frac{1}{N(N-1)} \sum_{i=1}^{N} \sum_{\neq j \neq 1}^{N} (\bar{a}_i - \bar{a})(a_j - \bar{a}),$$

$$= \frac{1}{N(N-1)}\left[\left(\sum_{i=1}^{N} a_i - \bar{a}\right)^2 - \sum_{i=1}^{N} (a_i - \bar{a})^2\right]$$

$$= \frac{1}{N(N-1)} [0 - N\sigma^2] = -\frac{\sigma^2}{N-1}$$

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

$$= n\,\frac{N-n}{N-n}\,\sigma^2.$$

19. ஒரு நகரம் இரு பகுதிகளாகப்பிரிக்கப்பட்டிருக்கிறது. முதல் பகுதியில் படித்தவர்கள் 10%, இரண்டாம் பகுதியில் படித்தவர் 9% நகரத்தில் படித்தவர்களின் சதவீதத்தை மதிப்பிட, படுகைமாதிரி முறையில்

(அ) சமவிகிதப் பங்கீடு

(ஆ) உத்தமப்பங்கீடு

ஆகிய இருமுறைகளும் ஒரே பங்கீட்டையளிக்கின்றன என்பதை நிறுவு.

விடை

சதவிகிதப் பங்கீட்டில் $n_1 = \frac{N_1}{N} n$, $n_2 = \frac{N_2}{N} n.$

உத்தமப் பங்கீட்டில் $n_1 = \frac{N_1 \sigma_1}{N_1 \sigma_1 + N_2 \sigma_2} \cdot n$,

$$n_2 = \frac{N_2 \sigma_2}{N_1 \sigma_1 + N_2 \sigma_2} n.$$

ஆனால் $\sigma_1^2 = P_1 Q_1 = 0{\cdot}1 \times {\cdot}9 = {\cdot}09.$

$\sigma_2^2 = P_2 Q_2 = 0{\cdot}9 \times 0{\cdot}1 = {\cdot}09.$

$\therefore \sigma_1 = \sigma_2.$

$\therefore$ உத்தமப் பங்கீட்டில் $n_1 = \frac{N_1}{N_1 + N_2} n = \frac{N_1}{N} n.$

$$n_2 = \frac{N_2}{N_1 + N_2} n = \frac{N_2}{N} n. \text{ (நி-து.)}$$

20. சென்ற கணக்கில் $N_1 = N_2$ என்றால், படுகை முறையில் $n_1 = n_2$ என்னும் பங்கீடு, படுகை முறை கையாளாமல் சாதாரணராண்டம் மாதிரி முறையைக் கையாள்வதைக் காட்டிலும் 64% அதிகத் திறன் வாய்ந்தது என்பதை நிறுவு.

விடை

$n_1 = n_2$ என்னும் பங்கீட்டில்

$$V(\bar{x}) = \frac{1}{N^2} \left[N_1^2 \frac{\sigma_1^2}{n_1} + N_2 \frac{\sigma_2^2}{n_2} \right] = \frac{N_1^2}{N^2 n_1} [\sigma_1^2 + \sigma_2^2]$$

$$= \frac{1}{4n_1} 2\sigma_1^2 = \frac{\sigma_1^2}{2n_1} = \frac{\sigma_1^2}{n}$$

சாதாரண ராண்டம் மாதிரி முறையில் $V(\bar{x}) = \frac{\sigma^2}{n}$.

$$\sigma^2 = \frac{N_1\sigma_1^2 + N_2\sigma_2^2}{N} + \frac{N_1(\bar{X}_1 - \bar{X})^2 + N_2(\bar{X}_2 - \bar{X})^2}{N}$$

$$= \frac{2N_1\sigma_1^2}{N} + \frac{N_1}{N}[(P_1 - P)^2 + (P_2 - P)^2]$$

$$P = \frac{N_1P_1 + N_2P_2}{N}$$

$$= \frac{N_1}{N} = \frac{1}{2}$$

$$= \sigma_1^2 + \frac{1}{2}[(0{\cdot}1 - 0{\cdot}5)^2 + (0{\cdot}9 - 0{\cdot}5)^2].$$

$$\therefore \quad \sigma^2 = P_1(1 - P_1) + \frac{1}{2} \times 2 \times 0{\cdot}16$$

$$= 0{\cdot}09 + 0{\cdot}16 = 0{\cdot}25$$

$\therefore$ $n_1 = n_2$ என்னும் பங்கீட்டின் திறன் விகிதம் $= \dfrac{\frac{\sigma^2}{n}}{\frac{\sigma_1^2}{n}}$

$$= \frac{\sigma^2}{\sigma_1^2} = \frac{0{\cdot}25}{0{\cdot}09}$$

திறன் வேறுபாடு $= \dfrac{\sigma^2}{n} - \dfrac{\sigma_1^2}{n} = \dfrac{1}{n}[0{\cdot}25 - 0{\cdot}09] = \dfrac{0{\cdot}16}{n}$.

திறன் வேறுபாட்டின் சதவீதம் $= \dfrac{\frac{0{\cdot}16}{n}}{\frac{0{\cdot}25}{n}} \times 100$

$$= \frac{16}{25} \times 100 = 64\%.$$

21. (அ) $P_1 = 0{\cdot}25$, $P_2 = 0{\cdot}75$ என்றால் திறன் வேறுபாட்டின் சதவீதம் 25% என்பதை நிறுவு. (ஆ) $P_1 = 0{\cdot}6$, $P_2 = 0{\cdot}4$ என்றால் திறன் வேறுபாட்டின் சதவீதம் 4% என்பதை நிறுவு.

விடை : திறன் வேறுபாடு விகிதம் $= \dfrac{\dfrac{\sigma^2}{n} - \dfrac{\sigma_1^2}{n}}{\dfrac{\sigma^2}{n}} \times 100$

$$= \left(1 - \frac{\sigma_1^2}{\sigma^2}\right) 100\ \%$$

$$\sigma_1^2 = P_1 Q_1 \qquad \sigma^2 = \frac{2 N_1 \sigma_1^2}{N} + \frac{N_1}{N} [P_1 - 0{\cdot}5)^2 + (P_2 - 0{\cdot}05)^2]$$

$$\therefore \quad \sigma^2 = \sigma_1^2 + (P_1 - 0{\cdot}5)^2.$$

$$\therefore \quad \sigma^2 - \sigma_1^2 = (P_1 - 0{\cdot}5)^2$$

$$\therefore \quad \left(1 - \frac{\sigma_1^2}{\sigma^2}\right) 100 = 100 \cdot \frac{(P_1 - 0{\cdot}5)^2}{\sigma_1^2 + (P_1 - 0{\cdot}5)^2}$$

$$= 100 \ \frac{(P_1 - 0{\cdot}5)^2}{P_1 Q_1 + (P_1 - 0{\cdot}5)^2}$$

(அ) $P_1 = 0{\cdot}25$ என்றால் திறன் வேறுபாடு விகிதம்

$$= 100 \frac{(\cdot 25)^2}{\cdot 25 \times \cdot 75 + (\cdot 25)^2} = \frac{25}{\cdot 75 + \cdot 25} = 25\%.$$

(ஆ) $P_1 = 0{\cdot}6$ என்றால் திறன் வேறுபாடு விகிதம்

$$= 100 \frac{\cdot 1 \times \cdot 1}{0{\cdot}6 \times 0{\cdot}4 + \cdot 1 \times \cdot 1} = \frac{1}{0{\cdot}24 + \cdot 01} = 4\%$$

22. இரு படுகைகளின் சராசரி μ_1, μ_2; பரவற் படி σ_1^2, σ_2^2; உறுப்புகளின் எண்ணிக்கை N_1, N_2. n பருமன் உள்ள படுகை மாதிரி இரு படுகைகளிலும் ஈ. செயா. சா. ரா. மா. முறையில் எடுக்கப்படுகின்றன. $\mu_2 - \mu_1$-ன் மதிப்பீட்டியின் $[\bar{x}_2 - \bar{x}_1]$ பரவற்படி சிறுமமாக இருக்கும் உத்தமப் பங்கீட்டினைக் கணக்கிடு.

விடை

$$V(\bar{x}_2 - \bar{x}_1) = V(\bar{x}_2) + V(\bar{x}_1) = \frac{N_1 - n_1}{N_1 - 1}\,\frac{\sigma_1^2}{n_1} + \frac{N_2 - n_2}{N_2 - 1}\,\frac{\sigma_2^2}{n_2}.$$

$n_1 + n_2 = n$ என்பது நிபந்தனை.

$$S = \frac{N_1 - n_1}{N_1 - 1}\,\frac{\sigma_1^2}{n_1} + \frac{N_2 - n_2}{N_2 - 1}\,\frac{\sigma_2^2}{n_2} + \lambda\,[n_1 + n_2 - n].$$

$$\frac{\partial s}{\partial n_1} = -\frac{N_1}{N_1 - 1}\,\frac{\sigma_1^2}{n_1^2} + \lambda = 0.$$

$$\frac{\partial s}{\partial n_2} = -\frac{N_2}{N_2 - 1}\,\frac{\sigma_2^2}{n_2^2} + \lambda = 0.$$

$$\therefore \quad \frac{N_1}{N_1 - 1}\,\frac{\sigma_1^2}{n_1^2} = \frac{N_2}{N_2 - 1}\,\frac{\sigma_2^2}{n_2^2} = \lambda.$$

$$\therefore \quad n_1 = \sqrt{\frac{N_1}{N_1 - 1}}\,\frac{\sigma_1}{\sqrt{\lambda}}\ ;\ n = \sqrt{\frac{N_2}{N_2 - 1}}\,\frac{\sigma_2}{\sqrt{\lambda}}$$

$$n = n_1 + n_2 = \sqrt{\frac{N_1}{N_1 - 1}}\,\frac{\sigma_1}{\sqrt{\lambda}} + \sqrt{\frac{N_2}{N_2 - 1}}\,\frac{\sigma_2}{\sqrt{\lambda}}$$

$$\therefore \quad \frac{1}{\sqrt{\lambda}} = \frac{n}{\sqrt{\frac{N_1}{N_1 - 1}}\cdot\sigma_1 + \sqrt{\frac{N_2}{N_2 - 1}}\cdot\sigma_2}$$

$$\therefore \quad n_1 = n\cdot\frac{\sqrt{\frac{N_1}{N_1 - 1}}\cdot\sigma_1}{\sqrt{\frac{N_1}{N_1 - 1}}\cdot\sigma_1 + \sqrt{\frac{N_2}{N_2 - 1}}\cdot\sigma_2}$$

$$n_2 = n\,\frac{\sqrt{\frac{N_2}{N_2 - 1}}\,\sigma_2}{\sqrt{\frac{N_1}{N_1 - 1}}\cdot\sigma_1 + \sqrt{\frac{N_2}{N_2 - 1}}\cdot\sigma_2}.$$

23. ஒரு கார் கம்பெனியில் கார்களை இரு படுகைகளாகப் பிரித்திருக்கின்றனர். முதல் படுகையிலுள்ள கார்களின் சராசரி மதிப்பு 5,000 ரூபாய். இரண்டாம் படுகையிலுள்ள கார்களின் சராசரி மதிப்பு 2,000 ரூபாய். இவைகளின் தரவிலக்கங்கள் (Standard deviations) சராசரிகளின் விகிதத்திலிருக்கின்றன. 100 பருமனுள்ள மாதிரியைப் பரவற் படியைச் சிறுமமாக்கும் உத்தமப் பங்கீடு உள்ளதாகப் பிரித்தால் பின் வருவனவற்றை நிறுவு.

(அ) $\frac{N_1}{n} : \frac{N_2}{n_2} = 2 : 5$ (ஆ) $N_1 = 93, N_2 = 796$

என்றால் $n_1 = 23, n_2 = 77$.

(அ) உத்தமப் பங்கீட்டில் $n_1 : n_2 = N_1 \sigma_1 : N_2 \sigma_2$ என்பது நமக்குத் தெரியும்.

$$\therefore \frac{N_1}{n_1} : \frac{N_2}{n_2} = \frac{1}{\sigma_1} : \frac{1}{\sigma_2} = \frac{1}{5000} : \frac{1}{2000} = 2 : 5$$

(ஆ) $N_1 = 93, N_2 = 796$ என்றால்

$$\frac{93}{n_1} : \frac{796}{n_2} = 2 : 5.$$

$$\therefore \frac{\frac{93}{2}}{n_1} = \frac{\frac{796}{5}}{n_2} = \frac{\frac{93}{2} + \frac{796}{5}}{n_1 + n_2} = \frac{\frac{93}{2} + \frac{796}{5}}{100} = \frac{2057}{1000}$$

$$\therefore n_1 = \frac{93}{2} \times \frac{1000}{2057} \doteq 23. \qquad \therefore n_2 \doteq 100 - n_1 = 77.$$

24. படுகை மாதிரி முறையில் முழுமைத் தொகுதியின் சராசரியை மதிப்பிடுவதற்கு,

$$\hat{X} = \frac{\sum_{i=1}^{k} n_i \bar{x}_i}{\sum_{i=1}^{k} n_i}$$

என்னும் மதிப்பீட்டியின் பிறழ்ச்சியைக் கண்டுபிடி. விகிதசமப் பங்கீட்டில் பிறழ்ச்சி மறைந்து விடும் என்பதை அதிலிருந்து நிறுவு.

விடை

$$E[\hat{\bar{X}}] = \frac{\sum_{i=1}^{k} n_i \bar{X}_i}{\sum_{i=1}^{k} n_i}$$

$$\therefore \quad B[\hat{\bar{X}}] = \frac{\sum_{i=1}^{k} n_i \bar{X}_i}{\sum_{i=1}^{k} n_i} - \bar{X} = \frac{\sum_{i=1}^{k} n_i (\bar{X}_i - \bar{X})}{\sum_{i=1}^{k} n_i}$$

விகிதசமப் பங்கீட்டில் $n_i = \frac{N_i}{N} n.$

$$\therefore \quad B[\hat{\bar{X}}] = \frac{\sum_{i=1}^{k} \frac{N_i}{N} n (\bar{X}_i - \bar{X})}{n}$$

$$= \frac{1}{N} \sum_{i=1}^{k} (N_i \bar{X}_i - N_i \bar{X})$$

$$= \bar{X} - \bar{X} = 0$$

25. படுகை மாதிரி முறையில் முழுமைத் தொகுதியின் சராசரியை மதிப்பிட

$\hat{\bar{X}} = \frac{1}{k} \sum_{i=1}^{k} \bar{x}_i$ என்னும் மதிப்பீட்டியின் பிறழ்ச்சியைக் கண்டுபிடி. $N_1 = N_2 = \dots = N_k$ என்றால் பிறழ்ச்சி மறைந்து விடுமென நிறுவு.

விடை

$$B(\hat{X}) = E(\hat{X}) - \bar{X} = \frac{1}{k}\sum_{i=1}^{k}\bar{X}_i - \bar{X} = \frac{1}{k}\sum_{i=1}^{k}(\bar{X}_i - \bar{X})$$

$N_1, N_2, \ldots\ldots, N_k$ ஆகியவை சமமாக இருக்கும்பொழுது

$$\bar{X} = \frac{1}{N}\sum_{i=1}^{k} N_i \bar{X}_i = \frac{1}{k}\sum \bar{X}_i.$$

$$\therefore \frac{1}{k}\sum_{i=1}^{k}(\bar{X}_i - \bar{X}) = 0$$

$$\therefore B(\hat{X}) = 0$$

26. 25 ஆம் கணக்கில் மாதிரிப் பருமனை அதிகரிப்பதால் பிறழ்ச்சியைக் குறைக்க முடியாது என்பதை நிறுவு.

விடை

$B(\hat{X}) = \frac{1}{k}\sum_{i=1}^{k}(\bar{X}_i - \bar{X})$. இதில் n_i-கள் வராததால் இது ஒரு மாறிலியாகும். ஆகவே இதை n_i-களை அதிகரிப்பதன் மூலம் குறைக்க முடியாது.

27. முழுமைத் தொகுதி சராசரியைக் கணக்கிட $\hat{\bar{X}} = \sum_{i=1}^{k} a_i \bar{x}_i$ என்னும் மதிப்பீட்டி கையாளப்பட்டால் அதன் பிறழ்ச்சியென்ன ? $a_i = \frac{N_i}{N}$ ஆக இருக்கும்பொழுது பிறழ்ச்சி மறைந்துவிடும் என நிறுவு. $\left[\sum_{i=1}^{k} a_i = 1\right.$ எனக் கொள்க. $\left.\right]$

விடை

$$E(\hat{\bar{X}}) = \sum_{i=1}^{k} a_i \bar{X}_i$$

$$\therefore \quad B(\hat{\bar{X}}) = \sum_{i=1}^{k} a_i \bar{X}_i - \bar{X}$$

$$= \sum_{i=1}^{k} a_i \bar{X}_i - \sum_{i=1}^{k} a_i \bar{X}$$

$$= \sum_{i=1}^{k} a_i (\bar{X}_i - \bar{X})$$

$a_i = \frac{N_i}{N}$ என்றால் $B(\hat{\bar{X}}) = \sum_{i=1}^{N} \frac{N_i}{N}(\bar{X}_i - \bar{X})$

$$= \bar{X} - \bar{X} = 0$$

28. (அ) 27ஆம் கணக்கில் பிறழ்ச்சி $\sum_{i=1}^{k} a_i (\bar{X}_i - \bar{X})$

$= \sum \left(a_i - \frac{N_i}{N}\right)\bar{X}_i$ என்று நிறுவு.

(ஆ) $m = MSE(\hat{\bar{X}}) = \sum_{i=1}^{k} a_i^2 \bar{X} \sigma_{\bar{x}_i}^2$

$+ \left[\sum_{i=1}^{k} a_i \bar{X}_i - \bar{X}\right]^2$ என்று நிறுவு.

(இ) $V(\hat{\bar{X}})$ சிறுமமாக இருக்கவேண்டுமாயின்

$a_i = \dfrac{\dfrac{1}{\sigma_{\bar{x}_i}^2}}{\sum_{i=1}^{k} \dfrac{1}{\sigma_{\bar{x}_i}^2}}$ என நிறுவு.

விடை

(அ) $$B(\hat{\bar{X}}) = \sum_{i=1}^{k} a_i (\bar{X}_i - \bar{X})$$

$$= \sum_{i=1}^{k} a_i(\bar{X}_i) - \bar{X} = \sum_{i=1}^{k} a_i \bar{X}_i - \sum_{i=1}^{k} \frac{N_i}{N} \bar{X}_i$$

$$= \sum_{i=1}^{k} \left(a_i - \frac{N_i}{N}\right) \bar{X}_i$$

(ஆ) $$MSE(\hat{\bar{X}}) = E[\hat{\bar{X}} - \bar{X}]^2$$

$$= E\left[\sum_{i=1}^{k} a_i [\bar{x}_i - \hat{\bar{X}}]\right]^2$$

$$= E\left[\sum_{i=1}^{k} a_i (\bar{x}_i - \bar{X}_i) + \sum_{i=1}^{k} a_i (\bar{X}_i - \bar{X})\right]^2$$

$$= \sum_{i=1}^{k} a_i^2 \sigma_{\bar{x}_i}^2 + \left[\sum_{i=1}^{k} a_i (\bar{X}_i - \bar{X})\right]^2$$

(இ) $V(\hat{\bar{X}}) = \sum_{i=1}^{k} a_i^2 \sigma_{\bar{x}_i}^2$. இது சிறுமமாக இருக்க வேண்டும்.

$$\sum_{i=1}^{k} a_i = 1$$ என்பது நிபந்தனை.

ஆகவே $S = \sum_{i=1}^{k} a_i^2 \sigma_{\bar{x}_i}^2 + \lambda \left[\sum_{i=1}^{k} a_i - 1 \right]$

$$\frac{\partial S}{\partial a_i} = 2a_i \sigma_{\bar{x}_i}^2 + \lambda = 0.$$

$$\therefore \frac{a_1}{\frac{1}{\sigma_{\bar{x}_1}^2}} = \frac{a_2}{\frac{1}{\sigma_{\bar{x}_2}^2}} = \cdots = \frac{a_k}{\frac{1}{\sigma_{\bar{x}_k}^2}} = \frac{1}{\sum_{i=1}^{k} \frac{1}{\sigma_{\bar{x}_i}^2}}$$

$$\therefore a_i = \frac{\frac{1}{\sigma_{\bar{x}_i}^2}}{\sum_{i=1}^{k} \frac{1}{\sigma_{\bar{x}_i}^2}}$$

29. 28 ஆம் கணக்கில் $MSE(\hat{\bar{X}})$ சிறுமமாக்கப்படும் உத்தமப் பங்கீட்டினைக் கண்டுபிடி. a_i-களையும் கண்டுபிடி.

விடை

$$MSE(\hat{\bar{X}}) = \sum_{i=1}^{k} a_i^2 \sigma_{\bar{x}_i}^2 + \left[\sum_{i=1}^{k} a_i \bar{X}_i - \bar{X} \right]^2$$

$$\therefore S = MSE(\hat{\bar{X}}) + \lambda \left[\sum_{i=1}^{k} a_i - 1 \right] + \mu \left[\sum_{i=1}^{k} n_i - n \right]$$

$$\frac{\partial S}{\partial n_i} = a_i^2 \frac{\partial}{\partial n_i}\left[\sigma_{\bar{x}_i^2} \right] + \mu = 0.$$

$$\frac{\partial S}{\partial a_i} = 2a_i \sigma_{\bar{x}_i}{}^2 + \lambda = 0.$$

$\sigma_{\bar{x}_i}{}^2 = \frac{N_i - n_i}{N_i - 1} \frac{\sigma_i^2}{n_i}$ எனக் கொண்டால்,

$$- a_i^2 \frac{N_i}{N_i - 1} \frac{\sigma_i^2}{n_i^2} + \mu = 0.$$

$$\therefore \frac{n_i}{a_i} = \sqrt{\frac{N_i}{N_i -} \frac{\sigma_i^2}{\mu}}$$

$\dfrac{a_i}{\dfrac{1}{\sigma_{\bar{x}_i}{}^2}} = -\dfrac{\lambda}{2}$ $\therefore$ முன்னைப் போலவே

$$a_i = \frac{\dfrac{1}{\sigma_{\bar{x}_i}{}^2}}{\displaystyle\sum_{i=1}^{k} \frac{1}{\sigma_{\bar{x}_i}{}^2}} = \frac{\dfrac{(N_i - 1)}{N_i - n_i} \dfrac{n_i}{\sigma_i^2}}{\displaystyle\sum_{i=1}^{k} \frac{(N_i - 1)\, n_i}{(N_i - n_i)\, \sigma_i^2}}$$

$$\therefore n_i = \sqrt{\frac{N_i}{N_i - 1} \frac{\sigma_i^2}{\mu}}\, a_i$$

30. M_p கறுப்புப் பந்துகளும் M_q வெள்ளைப் பந்துகளும் உள்ள M பருமன் கொண்ட முழுமைத் தொகுதியிலிருந்து N பந்துகள் சா. ரா. மா. முறையில் எடுக்கப்படுகின்றன. இந்த N பந்துகளிலிருந்து n பந்துகள் சா. ரா. மா. முறையில் எடுக்கப் படுகின்றன. இந்த n பந்துகளில் r பந்துகள் கறுப்பு என்றால் பின் வரும் 4 நிலைமைகளில் $E\left(\frac{r}{n}\right) = p$ என்பதையும், $V\left(\frac{r}{n}\right)$ என்பது பின்வருமாறு என்பதையும் நிறுவு.

N பந்துகள் எடுத்த முறை	n பந்துகள் எடுத்த முறை	
	ஈ. செ. சா. ரா. மா.	ஈ.செயா.சா ரா.மா.
ஈ. செ. சா. ரா. மா.	(அ) $V\left(\frac{r}{n}\right)=\frac{pq}{n}\left[1+\frac{n-1}{N}\right]$	(ஆ) $V\left(\frac{r}{n}\right)=\frac{pq}{n}$
ஈ.செயா.சா.ரா.மா.	(இ) $V\left(\frac{r}{n}\right)=\frac{pq}{n}\left[1+\frac{n-1}{N}\ \frac{M-N}{M-1}\right]$	(ஈ) $V\left(\frac{r}{n}\right)=\frac{pq}{n^2}\ \frac{M-n}{M-1}$

விடை

N பந்துகளில் x பந்துகள் கறுப்பு எனக் கொள்வோம்.

நிகழ்திறன்	(அ)	(ஆ)	இ	ஈ
$P(x)$	$\binom{N}{x}p^{x}q^{N-x}$	$\binom{N}{x}p^{x}q^{N-x}$	$\dfrac{\binom{Mp}{x}\binom{Mq}{N-x}}{\binom{M}{N}}$	$\dfrac{\binom{MP}{x}\binom{Mq}{N-x}}{\binom{M}{N}}$
$P\left(\frac{r}{x}\right)$	$\binom{n}{r}\left(1-\frac{x}{N}\right)^{n-r}\left(\frac{x}{N}\right)^{r}$	$\dfrac{\binom{x}{r}\binom{N-x}{n-r}}{\binom{N}{n}}$	(அ)	(ஆ)

$$E(r) = \sum_{r=0}^{n} \sum_{x=0}^{N} r\, P\left(\frac{r}{x}\right) P(x);$$

$$V(r) = \sum_{r=0}^{n} \sum_{x=0}^{N} r^2\, P\left(\frac{r}{x}\right) P(x) - [E(r)]^2$$

(அ) $$E(r) = \sum_{r=0}^{n} \sum_{x=0}^{N} r \binom{n}{r} \left(1 - \frac{x}{N}\right)^{n-r} \left(\frac{x}{N}\right)^{r} \binom{N}{x} p^x q^{N-x}$$

$$= \sum_{x=0}^{N} \binom{N}{x} p^x q^{N-x} \frac{nx}{N} \sum_{r=1}^{n} \binom{n-1}{r-1} \left(\frac{x}{N}\right)^{r-1} \left(1 - \frac{x}{N}\right)^{(n-1)-(r-1)}$$

$$= \frac{n}{N} \sum_{x=0}^{N} x \binom{N}{x} p^x q^{N-x}$$

$$= \frac{n}{N} Np \sum_{x=1}^{N} \binom{N-1}{x-1} p^{x-1} q^{(N-1)-(x-1)}$$

$$= np.$$

$$V(r) = \sum_{r=0}^{n} \sum_{x=0}^{N} r^2 \binom{n}{r} \left(1 - \frac{x}{N}\right)^{n-r} \left(\frac{x}{N}\right)^{r} \binom{N}{x} p^x q^{N-x} - (np)^2$$

$$= \sum_{r=0}^{n} \sum_{x=0}^{N} [r(r-1)+r] \binom{n}{r} \left(1-\frac{x}{N}\right)^{n-r} \left(\frac{x}{N}\right)^{r} \binom{N}{x} p^{x} q^{N-x} - (np)^{2}$$

$$= n(n-1) \sum_{x=0}^{N} \binom{N}{x} p^{x} q^{N-x} \frac{x^{2}}{N^{2}} \sum_{r=2}^{n} \binom{n-2}{r-2} \left(\frac{x}{N}\right)^{r-2} \left(1-\frac{x}{N}\right)^{\overline{n-2}-\overline{r-2}} + np - n^{2} p^{2}$$

$$= \frac{n(n-1)}{N^{2}} \sum_{x=0}^{N} x^{2} \binom{N}{x} p^{x} q^{N-x} + np - n^{2} p^{2}$$

$$= \frac{n(n-1)}{N^{2}} [Npq + N^{2} p^{2}] + np - n^{2} p^{2}$$

$$= \frac{n(n-1)}{N} pq + n(n-1) p^{2} + np - n^{2} p^{2}$$

$$= npq \left[1 + \frac{n-1}{N}\right]$$

(ஆ) $$E(r) = \sum_{r=0}^{n} \sum_{x=0}^{N} r \frac{\binom{x}{r}\binom{N-x}{n-r}}{\binom{N}{n}} \binom{N}{x} p^{x} q^{N-x}$$

$$= \sum_{x=0}^{N} \binom{N}{x} p^{x} q^{N-x} \sum_{r=0}^{n} r \frac{\binom{x}{r}\binom{N-x}{n-r}}{\binom{N}{n}}$$

$\left[\because \text{ ஹைபர் ஜியோமெட்ரிக் பரவலின் சராசரி } n\frac{x}{N} \right]$

$$= \sum_{x=0}^{N} \frac{nx}{N} \binom{N}{x} p^{x} q^{N-x} = \frac{n}{N} Np = np.$$

$$V(r) = E(r^2) - n^2 p^2 = \sum_{x=0}^{N} \binom{N}{x} p^{x} q^{N-x}$$

$$\cdot \sum_{r=0}^{n} \frac{r^2 \binom{x}{r} \binom{N-x}{n-r}}{\binom{N}{n}} - n^2 p^2$$

$$= \sum_{x=0}^{N} \left[n \frac{x}{N} \left(1 - \frac{x}{N}\right) \frac{N-n}{N-1} + n^2 \frac{x^2}{N^2} \right]$$

$$\cdot \binom{N}{x} p^{x} q^{N-x} - n^2 p^2$$

[∵ ஹைபர் ஜியோமெட்ரிக் பரவலின் பரவற் படி

$$= n \frac{x}{N} \left(1 - \frac{x}{N}\right) \cdot \left(\frac{N-n}{N-1}\right)$$]

$$\therefore \quad V(r) = \sum_{x=0}^{N} \left[\frac{N-n}{N-1} \frac{n}{N} x + \frac{x^2}{N^2} \left(n^2 - n \frac{N-n}{N-1} \right) \right]$$

$$\cdot \binom{N}{x} p^{x} q^{N-x} - n^2 p^2$$

$$= \frac{N-n}{N-1} \frac{n}{N} Np + \frac{n}{N^2} \left[n - \frac{N-n}{N-1} \right]$$

$$[Npq + N^2 p^2] - n^2 p^2$$

$$= npq$$

இ, ஈ. முதலியவற்றை இதைப்போலவே நிறுவலாம்.

31. ஒரு மிகப்பெரிய முழுமைத் தொகுதியில் குறிப்பிட்ட பண்புள்ளவை P பாகம். இம்முழுமைத் தொகுதியிலிருந்து சா. ரா. மா. முறையில் n உறுப்புகள் எடுத்ததில் r உறுப்புகள் குறிப்பிட்ட பண்புள்ளவையாயிருந்தன. P-க்கு உச்சநிகழ் வாய்ப்பு (Maximum likelihood) மதிப்பீட்டியைக் கண்டுபிடி.

விடை

$$P(r) = P^r (1 - P)^{n-r}$$

இது பெருமமாக இருக்கவேண்டுமாயின் $\frac{\partial}{\partial P} P(r) = 0$.

அல்லது $\frac{\partial}{\partial P} \log P(r) = 0 = \frac{r}{P} - \frac{n-r}{1-P}$.

$$\therefore \quad r(1-P) - P(n-r) = 0.$$

$$\therefore \quad -P[r + n - r] + r = 0.$$

$$\therefore \quad P = \frac{r}{n}.$$

$\therefore$ உச்ச நிகழ் வாய்ப்பு மதிப்பீட்டி $\frac{r}{n}$ ஆகும்.

32. ஒரு முழுமைத் தொகுதியில் மொத்தம் N பந்துகள் உள்ளன. அவைகளில் x பந்துகள் கறுப்பு; மற்றவை வெள்ளை. இத்தொகுதியிருந்து n பந்துகள் ஈ. செயா. சா. ரா. மா. முறையில் எடுக்கப்பட்டபொழுது அவைகளில் r பந்துகள் கறுப்பு. x-க்கு உச்ச நிகழ் வாய்ப்பு மதிப்பீட்டியைக் கண்டுபிடி.

விடை

$$P_x(r) = \binom{x}{r}\binom{N-x}{n-r} \div \binom{N}{n}.$$

$$P_x(r) - P_{x-1}(r) = \frac{1}{\binom{N}{n}}\left[\binom{x}{r}\binom{N-x}{n-r} - \binom{x-1}{r}\binom{N-x+1}{n-r}\right]$$

$$= \frac{1}{\binom{N}{n}} \left[\frac{\underline{|x}}{\underline{|r}\ \underline{|x-r}} \; \frac{\underline{|N-x}}{\underline{|n-r}\ \underline{|N-x-n+r}} - \frac{\underline{|x-1}}{\underline{|r}\ \underline{|x-1-r}} \; \frac{\underline{|N-x+1}}{\underline{|n-r}\ \underline{|N-x+1-n+r}} \right]$$

$$= \frac{1}{\binom{N}{n}} \; \frac{\underline{|x-1}}{\underline{|r}} \; \frac{\underline{|N-x}}{\underline{|x-r-1}\ \underline{|n-r}} \; \frac{1}{\underline{|N-x-n+r}}$$

$$\left[\frac{x}{(x-r)} - \frac{N-x+1}{N-x+1-n+r} \right]$$

$$\therefore \quad P_x(r) - P_{x-1}(r) \gtreqless 0 \iff x(N-x+1-n+r)$$

$$-(N-x+1)(x-r) \gtreqless 0$$

$$\iff x \gtreqless (N+1)\frac{r}{n}.$$

$\therefore$ x ன் மதிப்பீட்டி $(N+1)\frac{r}{n}$ ஐ விட அதிகமில்லாத மிகப் பெரிய முழு எண்ணாகும். $(N+1)\frac{r}{n}$ முழு எண்ணானல் $\frac{(N+1)r}{n-1}$ என்பது கூட ஒரு மதிப்பீட்டியே.

ஈ. செ. சா. ரா. மா. முறையில் ஒரு முழுமைத் தொகுதியிலிருந்து $x_1, x_2, \ldots, x_n$ என்னும் உறுப்புகள் எடுக்கப்பட்டன.

$$s^2 = \frac{1}{n}\sum_{i=1}^{n} x_i^2 - \bar{x}^2 \text{ என்றால்,}$$

$V(s^2)$, $C{\cdot}V{\cdot}(s^2)$ ஆகியவைகளைக் கண்டுபிடி.

விடை

μ என்பது முழுமைத் தொகுதிச் சராசரியானல் $y_i = x_i - \mu$ எனக் கொள்க. $s^2 = \frac{1}{n}\sum_{i=1}^{n} x_i^2 - \bar{x}^2 = \frac{1}{n}\sum_{i=1}^{n} y_i^2 - \bar{y}^2$ என்பது தெளிவு.

மேலும் $E(y_i^r) = \mu_r$ ஆகும்.

$$E(s^2) = \frac{1}{n}\sum_{i=1}^{n} E(y_i^2) - E(\bar{y}^2)$$

$$= \mu_2 - [V(\bar{y}) + 0^2]$$

$$= \mu_2 - \left[\frac{\mu_2}{n} + 0^2\right] = \frac{n-1}{n}\mu_2$$

$$s^4 = \frac{1}{n^2}\left(\sum y_i^2\right)^2 - \frac{2}{n^3}\sum_{i=1}^{n} y_i^2\left(\sum_{i=1}^{n} y_i\right)^2 + \frac{\left(\sum y_i\right)^4}{n^4}$$

$$= \frac{1}{n^2}\left[\sum_{i=1}^{n} y_i^4 + \sum\sum_{i \neq j} y_i^2 y_j^2\right]$$

$$- \frac{2}{n}\left[\sum_{i=1}^{n} y_i^4 + \sum\sum_{i \neq 1} y_i^2 y_j^2\right]$$

$$+ \frac{1}{n^4}\left[\sum_{i=1}^{n} y_i^4 + 3\sum\sum_{i \neq j} y_i^2 y_j^2\right] + 0$$

$$\therefore \quad E(s^4) = \frac{\mu_4}{n} + \frac{n(n-1)}{n^2}\mu_2^2 - \frac{2\mu_4}{n^2} - \frac{2n(n-1)}{n^3}\mu_2^2 + \frac{\mu_4}{n^3} + \frac{3n(n-1)}{n^4}\mu_2^2$$

$$= \mu_2^2 + \frac{\mu_4 - 3\mu_2^2}{n} - \frac{2\mu_4 - 5\mu_2^2}{n^2} + \frac{\mu_4 - 3\mu_2^2}{n^3}$$

$$\therefore \quad V(s^2) = E(s^4) - [E(s^2)]^2$$

$$= \frac{\mu_4 - \mu_2^2}{n} - 2\,\frac{\mu_4 - 2\mu_2^2}{n^2} + \frac{\mu_4 - 3\mu_2^2}{n^3}$$

$$[C.V.(s^2)]^2 = \frac{V(s^2)}{[E(s^2)]^2} = \frac{\beta_2 - 1}{n-1} - \frac{\beta_2 - 3}{(n-1)^2} \cdot \frac{n+1}{n}$$

27. ராண்டம் எண்கள் அட்டவணை

(RANDOM NUMBER TABLE)

02946	96520	81881	56247	17623	47441	27821
85697	62000	87957	07258	45054	58410	92081
26734	68426	52067	23123	73700	58730	06111
47829	32353	95941	72169	58374	03905	06865
76603	99339	40571	41186	04981	17531	97372
47526	26522	11045	83565	45639	02485	43905
70100	85732	19741	92951	98832	38188	24080
86819	50200	50889	06493	66638	03619	90906
41614	30070	23403	03656	77580	87772	86877
17930	26194	53836	53692	67125	98175	00912
24649	31845	25736	75231	83808	98997	71829
79899	34061	54308	59358	56462	58166	97302
76801	49594	81002	30397	52728	15101	72070
62567	08480	61873	63162	44873	35302	04511
49729	15275	09399	11211	67352	41526	23497
42658	70183	89417	57676	35370	14915	16569
65080	35569	79392	14937	06081	74957	87787
02906	38119	72407	71427	58478	99297	43519
75153	86376	63852	60557	21211	77299	74967
14192	49525	78844	13664	98964	64425	33536
32059	11548	86264	74406	81496	23996	56872
81716	80301	96704	57214	71361	41989	92589
43315	50483	02950	09611	36341	20326	37489
27510	10769	09921	46721	34183	22856	18724
81782	04769	36716	82519	98272	13969	12429

19975	48346	91029	78902	75689	70722	88553
98356	76855	18769	52843	64204	95212	31320
29708	17814	31556	68610	16574	42305	56300
88014	27583	78167	25057	93552	74363	30951
94491	19238	17396	10592	48907	79840	34607
56957	05072	53948	07850	42569	82391	20435
50915	31924	80621	17495	81618	15125	48087
49631	93771	80200	84622	31413	33756	15218
99683	58162	45516	39671	77600	15175	67415
86017	20264	94618	85979	42009	78616	45210
77339	64605	82585	85011	02955	84348	46436
61714	57933	37342	26000	93611	93346	71212
15232	48027	15832	62924	11509	95853	02747

"மாதிரிமுறை" நூல் பட்டியல்

(BIBLIOGRAPHY)

1. *R. L. Ackoff* (1953) : Design of Social Research.
2. *C. H. Backstrom & G. D. Hursh* (1963) : Survey Research.
3. *H. Cantril* (1945) : Gauging Public Opinion.
4. *W. G. Cochran* (1963) : Sampling Techniques.
5. *T. Da'enius* (1957) : Sampling in Sweden.
6. *W. E. Deming* : (1950) : Some Theory of Sampling.
7. *W. E. Deming* (1960) : Sample Design in Business Research.
8. *G. Gallup* (1948) : A Guide to Public Opinion Polls.
9. *M. H. Hansen & W. N. Hurwitz & W. G. Madow* (1953) : Sample Survey Method and Theory.
10. *W. A. Hendricks* (1956) : The Mathematical Theory of Sampling.
11. *H. H. Hyman* (1955) : Survey Design and Analysis.
12. *M. V. Jambunathan* (1953) : Some Aspects of Sampling.
13. *C. D. Jones* (1949) : Social Surveys.
14. *M. G. Kendall and A. Stuart* (1966) : The Advanced Theory of Statistics Vol. 3—Design, Analysis and Time Series.
15. *L. Kish* (1965) : Survey Sampling.

16. *J. B. Lansing, G. P. Ginsberg and K. Braaten* (1961) : An Investigation of Response Error.

17. *A. E. Mace* (1964) : Sample-size Determination.

18. *P. C. Mahalanobis* (1961) : Experiments in Statistical Sampling in the Indian Statistical Institute.

19. *P. C. Mahalanobis* (1967) : Sample Census of Area under Jute in Bengal in 1940.

20. *M. N. Murthy* (1967) : Sampling Theory and Methods

21. *V. G. Panse* (1958) : Estimation of Crop Yields.

22. *M. B. Parten* (1950) : Surveys, Polls & Samples.

23. *J. G. Peatman* (1947) : Descriptive and Sampling Statistics.

24. *M. R. Sampford* (1962): Introduction to Sampling Theory with Application to Agriculture.

25. *F. H. Sanderson* (1954) : Methods of Crop Forecasting.

26. *M J. Slonim* (1960) : Sampling in a Nutshell.

27. *R. K. Som* (1967) : Recall Lapse in Demographic Enquiries.

28. *F. F. Stephan and P. J. Mccarthy* (1958) : Sampling Opinions.

29. *A. Stuart* (1962) : Basic Ideas of Scientific Sampling.

30. *P. V. Sukhatme* (1953) : Sampling Theory of Surveys with Applications.

31. *R. M. Trueblood and R. M. Cyert* (1957) : Sampling Techniques in Accountancy.

32. *United Nations Statistical Office* (1960) : A Short Manual on Sampling; Volume I.

33. *L. L. Vance and J. Neter* (1950) : Statistical Sampling for Auditors and Accountants.

34. *F. Yates* (1949) : Sampling Methods for Censuses and Surveys.

35. *S. S. Zarkovich* (1961) : Sampling Methods and Census ; Volume I.

36. *S. S. Zarkovich (Edited by)* (1965) : Estimation of Areas in Agricultural Statistics.

புள்ளியியல் பத்திரிகைகள் (சுருக்கங்களுடன்)

(JOURNALS OF STATISTICS)

1. Agricultural Economic Research ; AER
2. American Journal of Public Health ; AJPH
3. American Sociological Review ; ASR
4. American Statistician ; Amer. Stat.
5. Annals of Applied Biology ; AAB
6. Annals of Institute of Statistical Mathematics ; AISM
7. Annals of Mathematical Statistics ; AMS
8. Applied Statistics ; Appl. Stat.
9. Australian Journal of Statistics ; AJS
10. Bulletin of Calcutta Statistical Association ; BCSA
11. Bulletin of International Statistical Institute ; BISI
12. Bulletin of Mathematical Statistics ; BMS
13. Incorporated Statistician ; Inc. Stat.
14. International Journal of Opinion and Attitude Research ; IJOAR
15. Journal of Agricultural Research ; JAR
16. Journal of Agricultural Science ; JAS
17. Journal of American Statistical Association ; JASA
18. Journal of Farm Economics ; JFE
19. Journal of Indian Society of Agricuitural Statistics ; JISAS

20. Journal of Indian Statistical Association ; JISA

21. Journal of Royal Statistical Society ; JRSS

22. Monthly Bulletin of Agricultural and Economic Statistics ; MBAES

23. Public Opinion Quarterly ; POQ

24. Report on Statistics of the Applied Research : Union of Japanese Engineers and Scintists ; JUSE

25. Review of International Statistical Institute ; RISI

26. Skandinavisk Aktuaritidskrift ; Skand Akt.

சில முக்கிய வெளியீடுகள் (Some Important Published Papers)

(அ) சாதாரண ராண்டம் மாதிரி முறை (Simple Random Sampling)

1. *L. A. Aroian* (1944) : Some methods for evaluation of a sum; JASA 39

2. *M. S. Avadhani and B. V. Sukhatme* (1965) : Controlled simple random sampling; JISAS, 18

3. *E. F. Baker* (1949) : The variance of the proportions of samples falling within a fixed interval for a normal population ; AMS, 20

4. *E F. Brayer* (1957) : Calculating the standard error of a proportion ; Appl. Stat, 6

5. J. *Cornfield* (1944) : On sampling from finite populations ; JASA, 39

6. *Desraj and H. S. Khamis* (1958) : Some remarks on sampling with replacements ; AMS, 29

7. *W. D. Evans* (1951) : On the variance of estimates of standard deviation and variance ; JASA, 46

8. *G. J. Glasser* (1962) : On the complete coverage of large units in a statistical survey ; RISI, 30

9. *H. Hart* (1926) : Reliability of a percentage; JASA, 21

10. *W. A. Hendricks* (1964) : Estimation of the probability that an observation will fall into a specified class ; JASA, 59

11. *L. Katz* (1953): Confidence interval for the number showing a certain characteristic in the population when sampling is without replacement; JASA, 48

12. *M. G. Kendall* (1952): Moment statistics in sampling from a finite population; Biometrika, 39

13. *J. Kozniewska* (1957): Comparison of the efficiency of drawing samples with and without replacement when the variance of the general population is unknown; Colloquium Mathematicum, 4

14. *K. W. Morris* (1963): A note on direct and inverse binomial sampling; Biometrika, 50

15. *P. K. Pathak* (1962): On simple random sampling with replacement; Sankhya, 24

16. *P. K. Pathak* (1961): On the evaluation of moments of district units in sample; Sankhya, 23

17. *P. J. Sandiford* (1960): A new Binomial approximation for use in sampling from finite populations; JASA, 55

18. *D. T. Searls* (1964): The utilization of a known coefficient of variation in the estimation procedure; JASA, 59

19. *D. T. Searls* (1966): An estimator for population mean which reduces the effect of large true observations; JASA, 61

20. *G. R. Seth and J. N. K. Rao* (1964): On the comparison between simple random sampling with and without replacement; Sankhya, 26

21. *J. G. Skellam* (1949): The distributions of the moment statistics of samples drawn without replacement from a finite population; JRSS (B), 11

22. *M. Weibull* (1959): Moments of the difference between means in two samples from a finite population; Skand. Akt., 42

(ஆ) முறையுடை மாதிரிமுறை (Systematic Sampling)

23. *H. Aoyama* (1951): On Practical systematic sampling; AISM, 3

24. *W. G. Cochran* (1946): Relative accuracy of systematic and stratified random samples for a certain class of populations ; AMS, 17

25. *A. C. Das* (1949): Two dimensional systematic sampling ; Science and Culture, 15

26. *A. C. Das* (1950): Two dimensional systematic sampling and the associated stratified and random sampling; Sankhya, 10

27. *A C. Das* (1951) : Systematic sampling; BISI, 33

28. *D. J. Finney* (1948): Random and systematic sampling in timber surveys ; Forestry, 22

29. *W. Gautschi* (1957) : Some remarks on systematic sampling; AMS, 28

30. *E.J.Hannan*(1962): Systematic sampling; Biometrika, 49

31. *H. H. Jowett* (1952): The accuracy of systematic sampling from covneyor belts; Appl Stat., 1

32. *W.G. Madow and L. H. Madow* (1944) : On the theory of systematic sampling; AMS, 15

33. *L. H. Madow* (1946): Systematic sampling and its relation to other sampling designs ; JASA, 41

34. *W. G. Madow* (1949) : On the theory of systematic sampling-II; AMS, 20

35. *W. G. M dow* (1953) : On the theory of systematic sampling-III; AMS, 24

36. *A. Milne* (1959) : The centric systematic area sample treated as a random sample; Biometrics, 15

37. *V. K. Mokashi* (1954) : Efficiency of sampling methods in forest surveys ; JISAS, 6

38. *A. W. Nordskog and S. L. Crump* (1948): Systematic and random sampling for estimating egg production in poultry ; Biometrics, 4

39 *J. G. Osborne* (1942) : Sampling errors of systematic and random surveys of cover type areas ; JASA, 37

40. *V. K. Sethi* (1965) : On optimum pairing of units; Sankhya, 27

41. *P. V. Sukhatme and V. G Panse and K V. R. Sastry*: Sampling techniques for estimating the catch of sea fish in India; Biometrics, 14

42. *R. M Sundrum* (1953) : A method of systematic sampling based on order properties; Biometrika, 40

43. *R. M. Williams* (1956) : Variance of the mean of systematic samples; Biometrika, 43

44 *F. Yates* (1948) : Systematic sampling; Philosophical Transactions of Royal Society (A), 241

45. *A. Zinger* (1964) : Systematic sampling in forestry; Biometrics, 209

(இ) சமவாய்ப்பிலா மாதிரி முறை : (Varying Probability Sampling)

46. *K. R. W. Brewer and G. C. Undy* (1962) : Samples of two units drawn with unequal probabilities without replacement ; AJS, 4.

47. *K. R. W. Brewer* (1963) : A model of systematic sampling with unequal probabilities, AJS, 5.

48. *E. Cansado* (1957) : Sampling with replacement from finite populations ; Trab. Est , 8.

49. *A. C. Das* (1951) : On two-phase sampling and sampling with varying probabilities ; BISI, 33.

50. *A. C. Das* (1962) : On MVU estimates of parameters of a finite population based on varying probability samples ; BCSA, 11.

51. *Desraj* (1954) : Ratio estimation in sampling with equal and unequal probabilities ; JISAS, 6.

52. *Desraj* (1956) : On the method of overlapping maps in sample surveys ; Sankhya, 17.

53. *Desraj* (1956) : A note on the determination of optimum probabilities in sampling without replacement; Sankhya, 17.

54. *Desraj* (1956) : Some estimators in sampling with varying probabilities without replacement , JASA, 51.

55. *Desraj* (1958) : On the relative accuracy of some sampling techniques : JASA, 53.

56. *Desraj* (1964) : The use of systematic sampling with probability proportional to size in a large-scale survey ; JASA, 59.

57. *Desraj* (1964) : On sampling with probabilities proportionate to aggregate size ; JISAS, 16

58. *Desr j* (1965) : Variance estimation in randomized systematic sampling wit hprobability porportionate to size; JASA, 60.

59. *J Durbin* (1953) : Some results in sampling theory when the units are selected with unequal probabilities ; JRSS, (B) 15.

60. *P. M. Grundy* (1964) : A method of sampling with probability exactly proportional to size, JRSS (B), 16.

61. *J. Hajek* (1964) : Asymptotive theory of rejective sampling with varying probabilities from a finite population ; AMS, 35.

62. *A. Haldar* (1961) : A note on the amount of rejection in Lahiri's method of pps sampling; Sankhya, 23 (B)

63. *M. H. Hansen and W. N. Hurwitz* (1943) : On the theory of sampling from finite population; AMS, 14

64. *M. H. Hansen and W. N. Hurwitz* (1949) : On the determination of optimum probabilities in sampling ; AMS, 20.

65. *T. V. Hanurav* (1962) : On Horvitz and Thompson estimator ; Sankhya, 24, (A).

66. *H. O. Hartley and J. N K. Rao* (1962): Sampling with unequal probabilities and without replacement ; AMS, 33.

67. *H.O. Hartley* (1966): Systematic sampling with unequal probabilities and without replacement ; JASA, 61

68. *D. G. Horvitz and D.J. Thompson* (1952): A generalization of sampling without replacement from a finite universe ; JASA, 47,

69. *R. J. Jessen* (1955) : Determining the fruit count on a tree by randomised branch sampling ; Biometrics, 11.

70. *N Keyfitz* (1951): Sampling with probability proportional to size ; adjustment for changes in probabilities ; JASA, 46.

71. *D. B. Lahiri* (1951) : A method of sample selection providing unbiased ratio estimates ; BISI, 33.

72. *D. B. Lahiri* (1954) : On the question of bias of systematic sampling ; Proceedings of World population Conference, 6.

73. *N C. Meier and C. F. Haner* (1951) : The adoptability of area probability sampling to public optimum measurements ; POQ, 15.

74. *H. Midzuno* (1952) : On the sampling-system with probability proportionals to sum of sizes ; AISM, 3.

75. *M. N. Murthy* (1957): Ordered and unordered estimators in sampling without replacement ; Sankhya, 18.

76. *R. D. Narain* (1951) : On sampling without replacement with varying probabilities ; JISAS, 3.

77. *P. K. Pathak* (1962) : On sampling units with unequal probabilities ; Sankhya, 24.

78. *P. K. Pathak* (1966) : An estimator in pps sampling for multiple characteristics ; Sankhya, 28.

79. *P. K. Pathak anl Shukla, N.D.* (1966) : Non-negativity of variance estimator ; Sankhya, 28

80 *S. C. Pearce and D. A. Holland* (1957) : Randomised branch sampling for estimating fruit mumber Biometrics, 13.

81. *S. G. Prabhu Ajgaonkar* (1965): On a class of linear estimators in sampling with varying probabilities without replacement ; JASA, 13.

82. *J. N. K. Rao* (1961) : On sampling with varying probabilities with replacement in sub-sampling designs; JISAS, 13.

83. *J.N.K. Rao* (1961) : On the estimate of variance in unequal probability sampling; AISM, 13.

84. *J N.K Rao, H.O. Hartley and W.G. Cochran* (1962) ; On a sample procedure of unequal probability sampling without replacement; JRSS, 24.

85. *J N.K. Rao* (1963) : On the three procedures of unequal probability sampling without replacement; JASA, 58.

86. *J.N.K. Rao* (1963) : On two systems of unequal probability sampling without replacement; AISM, 15.

87. *D.K. Roy Choudhury* (1956) : Integration of several pps surveys; Science and Culture, 22.

88. *M.R.Sampford* (1962): Methods of cluster sampling with and without replacement for clusters of unequal sizes; Biometrika, 49.

89. *A.R. Sen* (1953) : Recent advances in sampling with varying probabilities; BCSA, 5.

90. *A.R. Sen* (1955) : A simple design in sampling with varying probabilities ; JISAS, 7.

91. *G.R. Seth* (1966) : On estimates of variance of estimate of population total in varying probabilities; JISAS, 18.

92. *V.K. Sethi* (1962) : Some consequences of an interpretation of varying probability sampling;Sankhya, 24.

93. *D. Singh* (1954) : On efficiency of sampling with varying probabilities without replacement; JISAS, 6.

94. W.L. *Stevens* (1958) : Sampling without replacement with probability proportional to size; JRSS, (B), 20.

95. *A. Stuart* (1963) : Some remarks on sampling with unequal probabilities; BISI, 40.

96. *A. Stuart* (1964) : Multi-stage sampling with preliminary random stratification of first stage units; RISI, 32.

97. *P.V Sukhatme* (1946) : Bias in the use of small size plots in sample surveys for yield; Current Science, 15.

98. *S.Yamamoto* (1955) : On the theory of sampling with probability proportional to given values; AISM, 7.

99. *S.S. Zarcovich* (1960) : On the efficiency of sampling with varying probabilities and the selection of units with replacement; Metrika, 3.

(ச) படுகை மாதிரி முறை (Stratified Sampling)

100. *P.H. Anderson* (1942): Distributions in stratified sampling; AMS, 13.

101. *H. Aoyama* (1954) : A study of stratified random sampling ; AISM, 6.

102. *H. Aoyama* (1963) : Stratified random sampling with optimum allocation for multi-variate population; AISM, 14.

103. *P. Armitage* (1947) : A comparison of stratified with unrestricted random sampling from a finite population; Biometrika, 34.

104. *E C. Bryant, H.O. Hartley and R.J. Jessen* (1960): Design and estimation in two way stratification.

105. *W.G. Cochran* (1961) : Comparision of methods for determining stratum boundaries; BISI, 38.

106. *F. G. Cornell* (1947) : A stratified random sample of a small finite population ; JASA, 42.

107. *T. Dalenius* (1950) : The problem of optimum stratification I ; Skand. Akt., 33.

108. *T. Dalenius and M. Gurney* (1951) : The problem of optimum stratification II, Skand. Akt., 34.

109. *T. Dalenius* (1952) : The problem of optimum stratification in a special type of design ; Skand Akt. 35.

110. *T. Dalenius* (1953) : The economics of one-stage stratified sampling ; Sankhya, 12.

111. *T. Dalenius and J. L. Hodges (Jr)* (1957) : The choice of stratification points ; Skand. Akt., 40.

112. *T. Dalenius and J. L. Hodges (Jr)* (1959) : Minimum variance stratification ; JASA, 54.

113. *W. E. Deming* (1963) : Some stratified sampling plan in replicated designs ; Estadistica, 21.

114. *Desraj* (1957) : On estimating parametric functions in stratified sampling designs ; Sankhya, 17.

115. *Desraj* (1964) : On forming strata of equal aggregate size ; JASA, 59

116. *W. A. Ericson* (1965) : Optimum stratified sampling using prior information ; JASA, 60.

117. *W. D. Fisher* (1958) : On grouping for maximum homogeneity ; JASA, 53.

118. *B. Ghosh* (1947) : Bias due to change is stratification ; BCSA, 1.

119. *S.P. Ghosh* (1963) : Optimum stratification with two characters ; AMS, 34.

120. *S. P. Ghosh* (1963) : Optimum allocation in stratified sampling with replacement ; Metrika, 9.

121. *R. Jagannathan* (1965) : The programming approach in multiple character studies ; Econometrica, 33.

122. *A. E. Jones* (1948) : Systematic sampling of continuous parameter populations; Biometrika, 35.

123. *L. Kish* (1961) : Efficient allocation of multi-purpose sample; Econometrica, 29.

124 *E. Maruyama, C. Hayashi and M. D. Isida* (1950): On some criteria for stratification, AISM, 2.

125. *V. K. Mokashl* (1954): Efficiency of stratification in subsampling designs for the ratio method of estimation; JISAS, 6.

126 *S. Nordbotten* (1956): Allocation in stratified sampling by means of linear programming; Skand., Akt. 39.

127. *A. Ross* (1961): Variance estimates in optimum sampling designs, JASA, 56.

128. *G. R. Seth* (1966): On collapsing strata; JISAS, 18.

129. *V. K. Sethi* (1963): A note on optimum stratification of populations for estimating the population means; AJS, 5.

130. *B.D. Singh and D. Singh* (1965): Some remarks on double sampling for stratification; Biometrika, 52.

131. *F. F. Stephan* (1941): Stratification in representative sampling; J. Marketing, 6.

132. *W.L. Seevens* (1952): Samples with same number in each stratum; Biometrika, 39.

133. *J. S. Stock and L. R. Frankel* (1939): The allocation of sampling among serval strata; AMS, 10.

134. *P. V. Sukhatme* (1935): Contributions to the theory of representative method; JRSS (Supplement), 2.

135. *Y. Taga* (1935): On optimum balancing between sample side and number of strata in sub-sampling; AISM, 4.

136. *M. Weibull* (1950): The distribution of *t* and *z* variables in the case of stratified sample with individuals taken from normal parent populations with varying means; Skand. Akt., 33.

137. *W. H. Williams* (1964): Sample selection and choice of estimators in two-way stratified population; JASA, 59.

(உ) திரள் மாதிரி முறை (Cluster Sampling)

138. *B. Ghosh* (1956): Optimum structure of rectangular sample units; BCSA, 6.

139. *M. H. Hansen and W. N. Hurwitz* (1942): Relative efficiencies of various sampling units in population inquiries; JASA, 37.

140. *W. A. Hendricks* (1944): The relative efficiencies of groups of farms as sampling units; JASA, 39.

141. *S. H. Justesen* (1932): Influence of size and shape of plots on the precision of field experiments with potatoes; JAS, 22.

142. *L. Kish* (1957): Confidence intervals for clustered samples, ASR, 22.

143. P. C. *Mahalanobis* (1946): Use of small size plots in sample survey for crop yields; Nature, 158.

144. *V. G. Panse* (1946): Plot size in yield surveys on cotton; Current Science, 15.

145. *J. M. Sengupta* (1964): On perimeter bias in sample cuts of small size; Sankhya, 26.

146. *P. V. Sukhatme* (1946): Size of sampling unit in yield surveys; Nature, 158.

(ஊ) பலகட்ட மாதிரிமுறை (Multi-stage Sampling)

147. *P. K. Banerjee* (1952): Error of one-stage and two-stage sampling; BCSA, 4.

148. *M. S. Bartlett* (1937) : Sub sampling for attributes : JRSS, Supplement, 4.

149. *D. Basu* (1954) : On the optimum character of some estimators used in multi-stage sampling problems ; Sankhya, 13.

150. Z. *W. Birnbaum and W. C. Healy* (1960) : Estimates with prescribed variance based on two-stage sampling; AMS, 31.

151. *S. H. Brooks* (1955) : The estimation of an optimum sub-sampling number ; JASA, 50.

152. *E. Cansado* (1952) : Expectations and variances in multi stage sampling ; Trab. Est., 3.

153. *I. M. Chakravarty* (1952) : Use of analysis of covariance in two-stage sampling ; BCSA, 4

154. *T. Corltet* (1963) : Rapid methods of estimating standard errors of stratified multi-stage samples-a preliminary investigation ; The statistician, 13.

155. *J.P.Ecimovic* (1956) : Three stage sampling with varying probabilities of selection ; JISAS, 8.

156. *M. Ganguly* (1941) : A note on nested sampling ; Sankhya, 5.

157. *B. Ghosh* (1949) : A multi-stage stochastic model for natural fields , BCSA, 2.

158. *S. P. Ghosh* (1963) : Estimating the mean by two-stage sampling with replacement ; BCSA, 12.

159. *V. P. Godambe* (1951) : On two stage sampling ; JRSS, (B) 13.

160. *E.H. Jebe* (1952) : Estimation for sub sampling designs employing the country as a primary sampling unit ; JASA, 47.

161. *L. Kish* (1949) : A procedure for objective respondent selection within the household ; JASA, 44.

162. *L. H. Madow* (1950) : On the use of the country as a primary sampling unit for state estimates ; JASA, 45.

163. *R. Rangarajan* (1957) : A note on two-stage sampling ; Sankhya, 17.

164. *J. N. K. Rao* (1964) : Unbiased ratio and regression estimators in multi-stage sampling, JISAS, 16.

165. *J. Roy* (1957) : A note on estimations of variance components in multi-stage sampling with varying probabilities ; Sankhya, 17.

166. *J. Sedransk* (1965) : Analytical surveys with cluster sampling ; JRSS, (B), 27.

167. *A. R. Sen, R. L. Anderson and A. L. Finkner* (1954) : A comparison of stratified two-stage sampling systems ; JASA, 49.

168. *D. Singh* (1958) : Estimates of variance components in finite population, JISAS, 10.

169. *H. E. Smith* (1947) : Standard errors of means in sampling surveys with two stage sampling; JRSS, (A), 110.

170. *P. V. Sukhatme and R. D. Narain* (1952) : Sampling with replacement ; JISAS, 4.

171. *S. S. Wilks* (1960) : A two-stage scheme for sampling without replacement, BISI, 37.

விகிதமுறை மதிப்பீடு (Ratio Estimation)

172. *M.S. Avadhani and B.V. Sukhatme* (1966) : A note on the ratio and regression methods of estimation in controlled simple random sampling ; JISAS, 18.

173. *K.R.W. Brewer* (1963) : Ratio Estimation and finite populations ; some results deducible from the assumption of an underlying stochastic process : AJS, 5.

174. *Desraj* (1964) : A note on the variance of the ratio estimate ; JASA, 59.

175. *J. Durbin* (1959) : A note on the application of Quenouille's method of bias reduction to the estimation of ratios ; Biometrika, 46.

176. *J.M Elkin* (1953) : Estimating the ratio between the proportions of two classes when one is a sub class of the other; JASA, 48.

177. *A.F. Fung* (1961) : Interviewer differences among automobile purchasers ; Appl. Stat. 10.

178. *L.A. Goodman and H.O. Hartley* (1958) : The precision of unbiased ratio type estimators ; JASA, 53.

179. J. *Hajek* (1960) : On the theory of ratio estimates ; BISI, 37.

180. *H.O. Hartley and A. Ross* (1954) : Unbiased ratio estimators ; Nature, 174.

181. *L. Kish and I. Hess* (1959) : On variances of ratios and their differences in multi stage samples; JASA, 54 ;

182. *A.R. Kokan* (1963) : A note on the stability of the estimates of standard errors of the ordinary estimates and ratio estimate, BCSA, 12.

183. *J C. Koop* (1951) : A note on the bias of the ratio estimate ; BISI, 33.

184. J.C. *Koop* (1964) : On an identity for the variance of a ratio of two random variables ; JRSS, (B), 26.

185. *V.K. Mokashi* (1953) : Investigations on sampling for estimations of crop acreages I, JISAS, 5.

186. *I. Olkin* (1958) : Multi variate ratio estimation for finite population ; Biometrika, 45.

187. *J.N. Pascual* (1961) : Unbiased ratio estimators in stratified sampling ; JASA, 56

188. *P.K. Pathak* (1964) : On sampling schemes providing unbiased ratio estimates AMS, 35.

189. *M.H. Quenouille* (1956) : Notes on bias in estimation; Biometrika, 43.

190. *M.J. Rao* (1966) : On certain unbiased ratio Estimators, AISM, 18.

191. *T.J. Rao* (1966) : On the variance of the ratio estimator for Midzuno-Sen sampling scheme ; Metrika, 10.

192. *D.S. Robson and Vithayasai* (1961) : Unbiased component wise ratio estimation, JASA, 56.

193. *K.V.R. Sastry* (1965) : Unbiased ratio estimators ; JISAS, 17.

194. *M.P. Singh* (1965) : On the estimation of ratio and product of population parameters ; Sankhya, 27.

195. *B.V. Sukhatme* (1962) : Generalized Hartley-Ross unbiased ratio type estimator ; Nature, 199

196. *B.V. Sukhatme* (1962) : Some ratio type estimators in two-phase sampling ; JASA, 57.

197. *A.K P.C. Swain* (1964) : The use of systematic sampling in ratio estimate ; JISA, 2.

198. *M. Tin* (1965) : Comparison of some ratio type estimators ; JASA, 60.

(ஏ) பல்தோற்ற மாதிரி முறையும் தொடர்புப் போக்கு மதிப்பீடும்
(Multi-phase Sampling and Regression Estimation)

199. *K.S. Bonerjee* (1955) : A note on successive sampling ; BCSA, 6.

200. *C. Bose* (1943) : Note on the sampling error in the method of double sampling ; Sankhya, 6.

201. *C Bose* (1951): Some further results on errors in double sampling technique; Sankhya 11.

202. *D. R. Cox* (1952) Estimations by double-sampling ; Biometrika, 34.

203. *Desraj* (1965) : On a method of using multi-auxiliary information in sample surveys ; JASA, 60.

204. *Desraj* (1965) : On sampling over two occasions with probability proportional to size ; AMS, 36.

205. *A. R. Eckler* (1955) : Rotation sampling ; AMS, 26.

206. *I. P. Fellegi* (1963) : Sampling with varying probabilities without replacement—rotating and non-rotating samples ; JASA, 58.

207. *B. Ghosh* (1947) : Double sampling with many auxiliary variates ; BCSA, 1.

208. *G. Koulldorf* (1963) : Some problems of optimum allocation for sampling on two occasions ; RISI, 31.

209. *E. S. Marks and W. P. Maudlin* (1950) : Response errors in census research ; JASA.

210. *R. D. Narain* (1954) : The general theory of sampling on successive occasions ; BISI, 34.

211. *H. D. Patterson* (1950) : Sampling on successive occasions with partial replacement of units ; JRSS, (B), 12.

212. *K. C. Seal* (1951) : On errors of estimates in various double-sampling procedure ; Sankhya, 11.

213. *K. C. Seal* (1953) : On certain extended cases of double sampling ; Sankhya, 12.

214. *J. Sedransk* (1965) : A double-sampling scheme for analytical surveys ; JASA, 60.

215. J. *D. Shaffer* (1954) : A plan for sampling a changing population over time ; JFE, 36.

216. J. *Silber* (1948) : Multiple sampling for variables ; AMS, 19.

217. *B. V. Sukhatme and R. S. Koshal* (1959) : A contribution to double sampling ; JISAS, 9.

218. *B. D. Tikkiwal* (1953) : Optimum allocation in successive sampling; JISAS, 9

219. *A. D. Tikkiwal* (1960) : On the theory of classical regression and double-sampling estimation ; JRSS, (B), 22,

220. *B. D Tikkiwal* (1964) : A note on two-stage sampling on successive occasions ; Sankhya, 26.

221. *W. H. Williams* (163) : The precision of some unbiased regression estimators ; Biometrika, 19.

கலைச்சொற்கள்

A

Algebra — இயற்கணிதம்
Allocation — பங்கீடு
Attribute — பண்பு
Associate — தொடர்புப்படுத்து
Auto correlation — தற்றொடர்பாக்கம்
Average — சராசரி
Axiom — எடுகோள்

B

Best — சாலச் சிறந்த
Biased — பிறழ்ச்சியுடைய
Binomial distribution — ஈருறுப்புப் பரவல்
Borel field — போரல் கணக்களம்

C

Cauchy-Schwartz inequality — கோஷி ஷ்வார்ட்ஸ் சமனிலி.
Characteristic function of a set — கணத்தின் குறியீட்டுச் சார்பு
Circular Systematic Sampling — வட்டமுறையுடை மாதிரிமுறை
Central moments — மையத் திருப்புத் திறன்கள்.
Class of sets — கணத்தொகுதி
Class of intervels — இடைவெளித்தொகுதி
Closed intervel — மூடிய இடைவெளி
Cluster sampling — திரள்மாதிரிமுறை
Components — அங்கங்கள்
Coefficient of variation — மாறுபாட்டுக் கெழு

Conditional Probability — நிபந்தனைக்குட்பட்ட நிகழ்திறன்
Confidence Interval — நம்பிக்கை இடைவெளி
Constant — மாறிலி
Contained — அடக்கம்
Continuous — தொடர்ச்சியான
Continuous function — தொடர் சார்பு
Continuous random variable — தொடர்ச்சியான ராண்டம் மாறி
Copy — பிரதி
Correlation Coefficient — உடன் தொடர்புக்கெழு
Convergent infinite sequence — ஒருங்கும் முடிவில் தொடர்
Countable — எண்ணிடத்தக்க
Cost function — செலவுச்சார்பு
Covariance — உடன் மாற்றம்
Cumulative distribution function — குவிவுப் பரவல் சார்பு

D

Deduction — பகுத்தறிதல், வருவித்தல்
Definition — வரைவிலக்கணம்
De Morgan Laws — டிமார்கன் விதிகள்
Difference Estimation — வேறுபாடு மதிப்பீடு
Differentiable — வகையிடக்கூடிய
Differential coefficient — வகையீட்டுக்கெழு
Discrete random variable — தனித்த ராண்டம் மாறி
Disjoint sets — பிரிந்த கணங்கள்
Distribution function — பரவல் சார்பு
Domain of a function — சார்பின் அரங்கம்

E

Efficiency — திறன், பயனளவு
Equivalence Relation — சமத்தொடர்பு
Error function — பிழைச் சார்பு
Errors in Schedules and questionnaires — கேள்விப்பட்டியல் குறைகள்
Errors of non-response — பதிலின்மைப் பிழை
Estimate — மதிப்பிடு, மதிப்பீடு
Estimator — மதிப்பீட்டி

Events — நிகழ்ச்சிகள்
Expected value — எதிர்பார்க்கும் மதிப்பு
Experimental design — பரிசோதனை முறையமைப்பு
Extension — நீட்சி

F

Finite — முடிவுள்ள
Field of sets — கணக்களம்
Functions — சார்புகள்

H

Heterogeneous — பல படித்தான
Homogeneous — ஒரே படித்தான
Hypothesis — எடுகோள்

I

Image — நிழல்
Impossible Event — நடவா நிகழ்ச்சி
Independent — சார்பிலா
Index set — குறியிடும் கணம்
Independent Events — சார்பிலா நிகழ்ச்சிகள்
Increasing — ஏறு வரிசையில்
Induced Probability space — ஊக்குவிக்கப்பட்ட நிகழ்திற வெளி
Inequality — சமனிலி, சமனின்மை
Induction — தொகுத்தறிதல்
Integral — தொகை
Integrable — தொகையிட முடியும்
Integration — தொகையிடல்
Into transformation — உள்படவாக்கம்
Intersection — வெட்டு
Intra class correlation co-efficient — வகுப்புள் ஒட்டுறவுக் கெழு
Inverse element — மூலம்
Inverse function — நேர்மாறான சார்பு

J

Judgement sample — உத்தேச மாதிரி

L

Lebesgue Integration — லெபெக் தொகையிடல்
Lebesgue Stieltjes Integration — லெபெக் ஸ்டீல்ட்யென் தொகையிடல்
Lagrangean multiplier method — லெக்ராஞ்சின் பெருக்கி முறை
Left open interval — இடப்புறம் திறந்த இடைவெளி
Linear Systematic Sampling — நேர்க்கோட்டு முறையுடை மாதிரி முறை
Limit — எல்லை
Logical consistency — தர்க்க ரீதியான இசைவு
Loss function — நஷ்டச் சார்பு
Lottery method — குலுக்கும் முறை

M

Mapping — படவாக்கம்
Mathematical Analysis — கணிதப் பகுவியல்
Mean — சராசரி
Mean square deviation — சராசரி வர்க்கவிலக்கம்
Mean Square Error — சராசரி வர்க்கப் பிழை
Measurable function — அளக்கத்தகு சார்பு
Minimum cost — சிறுமச்செலவு
Moments — திருப்புத்திறன்கள்
Minimum — சிறுமம்
Maximum — பெருமம்
Multi phase sampling — பல் தோற்ற மாதிரி முறை
Multi stage sampling — பலகட்ட மாதிரி முறை
Mutually Exclusive — ஒன்றையொன்று விலக்கும்

N

Necessary and Sufficient — வேண்டியதும் போதுமானதுமான
Non-countable — எண்ணிடத்தகா
Non-empty — காலியில்லாத
Non-integrable — தொகையிடமுடியாத
Non-random Sample — ராண்டமற்ற மாதிரி
Non-sampling Errors — மாதிரித் தேர்தலாலல்லாத பிழைகள்

Normal distribution	—	இயல் நிலைப்பரவல்

O

One to one correspondence	—	ஒத்தியைபு
One to one transformation	—	ஒன்றுக்கொன்று சார்பு
On to transformation	—	மேற்படவாக்கம்
Open interval	—	திறந்த இடைவெளி
Optimum	—	உத்தம
Optimum allocation	—	உத்தமப் பங்கீடு
Order Statistics	—	வரிசையளவுகள்
Origin	—	மூலம், ஆதி

P

Parameter	—	முழுமைத்தொகுதியளவை
Pareto distribution	—	பாரெடோ பரவல்
Pilot Survey	—	முன்னோடி அளவெடுப்பு
Population	—	முழுமைத் தொகுதி, ஜனத்தொகை
Population Proportion	—	முழுமைத் தொகுதிபாகம்
Population Size	—	முழுமைத் தொகுதி பருமன்
Poisson distribution	—	பாய்ஸன் பரவல்
Power set	—	அடுக்குக் கணம்
Probability	—	நிகழ்திறம். நிகழ்திறன், நிகழ் தகவு ஊக அளவை, வாய்ப்பு
Probability density	—	ஊக அளவை அடர்த்தி
Probabity space	—	நிகழ்திறவெளி
Product sets	—	பெருக்கல் கணங்கள்
Proportional	—	விகித சமமுள்ள
Proportional allocation	—	விகித சமப்பங்கீடு

Q

Quality	—	பண்பு

S

Sample	—	மாதிரி, கூறு
Sample Space	—	மாதிரி வெளி

Sample size — மாதிரிப் பருமன்
Sampling distribution — மாதிரிப் பரவல்
Sampling Error — கூறுமுறைப்பிழை, மாதிரித் தேர்தற் பிழை
Set — கணம்
Set theory — கணவியல்
Semi open interval — பாதி திறந்த இடைவெளி
Sequence of sets — கணத்தொடர்
Simple function — எளிய சார்பு
Size — பருமன்
Serial correlation Coefficient — வரிசை ஒட்டுறவுக் கெழு
Simple random Sample — சாதாரண ராண்டம் மாதிரி
Small samples — சிறு மாதிரிகள்
Standard deviation — திட்டவிலக்கம்
Standard Error — திட்டப்பிழை
Standard normal distribution — தரமான இயல்நிலைப் பரவல்
Standardised random variable — தரப்படுத்தப்பட்ட ராண்டாம் மாதிரி
Stabilizing — ஸ்திரப்படுத்தும்
Statistic — மாதிரியளவை
Statistics — புள்ளியியல்
Statistical tables — புள்ளியியல் அட்டவணை
Step function — படிச்சார்பு
Stratified Sampling — படுகை மாதிரி முறை
Subset — உபகணம்
Super population — பெரும் முழுமைத் தொகுதி
Survey — அளவெடுப்பு
Symmetry — ஒழுங்கு, ஒழுங்கு முறை

T

Test of significance — சிறப்பு காண் சோதனை
Transformation — உருமாற்றம், படவாக்கம்
Triplet — மும்மை
Two-dimensional — இருபரிமாண
Two-stage sampling — இருகட்ட மாதிரி முறை

U

Unbiased	—	பிறழ்ச்சியற்ற
Uncorrelated	—	உடன்தொடர்பற்ற
Uniformly better	—	சீராக மேலானது
Union	—	கூட்டு, பிணைப்பு
Unit	—	உறுப்பு

V

Vanishing	—	இல்லாதொழிதல், மறைதல்
Variance	—	பரவற்படி
Varying Probability Sampling	—	சமவாய்ப்பிலா மாதிரிமுறை

W

With replacement	—	ஈடுசெய்யப்பட்ட
Without replacement	—	ஈடுசெய்யப்படாத