நிகழ்தகவு
(probability)

நிகழ்தகவு - விளக்கம் :

இன்று மழை பெய்யலாம்; இன்று பேராசிரியர் வருகை தராமல் போகலாம், விளையாட்டுப் போட்டியில் நமது அணி வெற்றி வாய்ப்பு உண்டு. இக்கூற்றுகள் எல்லாம் ஒரு சம்பவம் நடைபெறுமா என்ற சந்தேகத்தின் அடிப்படையில் எழுந்தவையாகும்.

ஆனால் புள்ளியியலில் நாம் இத்தகைய கூற்றை நிகழ்தகவு எனும் கருத்தின் அடிப்படையில் துல்லியமாக எண் சார்ந்த கூறுகளாக மாற்றுகின்றோம்.

நிகழ்தகவு கருத்தின் பயன்பாடு :-

i) முன்னர் சூதாட்டங்களில் உள்ள சிக்கல்களைத் தீர்க்கவும் உலகத்தை அளவிடவும் இது பயன்பட்டது.

ii) தற்போது புள்ளியியலில் முடிவுகளை ஆய்வதற்கு இது பயன்படுகிறது.

iii) நடைபெறுமா என்ற சந்தேகத்திற்குரிய நிகழ்ச்சிகளைப் பற்றித் தீர்மானம் செய்ய நிகழ்தகவு பயன்படும்.

iv) பொருளியல் அளவை மாதிரிகளை அமைக்க உதவு-கிறது.

v) நிர்வாகம், திட்டமிடல் ஆகியவற்றில் தீர்வுகள் காண இது பயன்படுகிறது.

vi) அறிவியல் மேலாண்மை போன்ற பலதுறைகளிலும் பயன்படுகிறது.

நிகழ்தகவு – இலக்கணம் :

ஒரு நிகழ்ச்சி நடைபெறுவதற்கான வாய்ப்பை எண் வடிவில் கூறுவதை நிகழ்தகவு என்கிறோம். நடைபெற எவ்வித வாய்ப்பும் இல்லை எனில் '0' எனவும்; கட்டாயம் நடைபெறும் எனில் '1' எனவும் கொள்ளலாம்.

1) தொன்மை அணுகுமுறை :

இதுவும் பழமையான எளிமையான கருத்தாகும். இக்கருத்து சூதாட்டங்களில் பலவகைகளான நாணயம் சுண்டுதல், தாயக்கட்டை உருட்டுதல், சீட்டுக்கட்டில் சீட்டு உருவுதல் போன்றவற்றில் பயன்படுத்தப்பட்டது.

$$\text{நிகழ்தகவு}, P = \frac{\text{சாதகமான நிகழ்ச்சிகளின் எண்ணிக்கை}}{\text{சமவாய்ப்புடைய மொத்த நிகழ்ச்சிகளின் எண்ணிக்கை}}$$

குறிப்பு :

ஒரு நாணயத்தை சுண்டினால்.

தலைவிழும் நிகழ்தகவு $(p) = 1/2$

பூவிழும் நிகழ்தகவு $(q) = 1/2$

$$p+q = 1/2 + 1/2 = 1$$
$$p = 1-q$$
$$q = 1-p$$

மாதிரி 1 :

ஒரு பெட்டியில் இரண்டு வெள்ளை பந்துகளும் எட்டு கருப்பு பந்துகளும் உள்ளன. வெள்ளைப்பந்து எடுப்பதற்கான நிகழ்தகவு என்ன? கருப்புபந்து எடுப்பதற்கான நிகழ்தகவு என்ன? வெள்ளை அல்லது கருப்பு பந்து எடுப்பதற்கான நிகழ்தகவு என்ன?

தீர்வு :

வெள்ளைப்பந்து எடுக்கும் நி.த $= \frac{2}{2+8} = \frac{2}{10} = \frac{1}{5}$

கருப்பு பந்து எடுக்கும் நி.த $= \frac{8}{2+8} = \frac{8}{10} = \frac{4}{10}$

வெள்ளை அல்லது கறுப்பு பந்து எடுக்கும் நி.த $= \frac{1}{5} + \frac{4}{5}$

2) சார்பு அலைவெண் கோட்பாடு :

ஒரு நாணயத்தை 10 தடவை சுண்டினால் 6 முறை தலையும் 4 முறை பூவும் விழலாம். சம அளவில் விழாது. ஆனால், சுண்டுதலை மிக அதிக தடவை நிகழ்த்தினால் பூவும் தலையும் சம அளவில் விழுவதற்கான வாய்ப்பு கூடுகின்றது. எனவே (n) மொத்த நிகழ்ச்சிகளின் எண்ணிக்கை கூடும்போது விழுவதற்கான வாய்ப்பு 0.5 ஐ நெருங்கும்.

$$\left(P(A) = \lim_{n \to \infty} \frac{a}{n} \right)$$

$$P(A) = \frac{a}{n}$$

அதாவது ஒரு நிகழ்ச்சியின் நிகழ்தகவு மிகவும் அதிக அளவிலான சோதனைகளில் நிகழும் அலைவு எண்களின் விகிதமாகும்.

3) தனிநபர் கண்ணோட்டம் :

இது ஒரு நிகழ்ச்சி நடைபெறுவதற்கான வாய்ப்பை நாம் கருதுவதாக வெளியிடும் முறையாகும். இது தனி நபரது நம்பிக்கையை சார்ந்துள்ளது.

4) கொள்கைசார் அணுகு முறை :

இதனை 1933-ல் கோலமொகரவ் (Kolmogorov) அறிமுகம் செய்தார். நிகழ்தகவுகளை தீர்மானிக்கும் மூன்று கொள்கைகளை வலியுறுத்தினார்.

i) ஒரு நிகழ்ச்சி நடைபெறும் நிகழ்தகவு 0 முதல் 1 வரை இருக்கும்.

ii) மொத்த கூறுபரப்பின் நிகழ்தகவு 1 அதாவது $P(S) = 1$

iii) Aயும் Bயும் ஒன்றை ஒன்று புறக்கணிக்கும் நிகழ்ச்சிகள் எனில் A அல்லது B நிகழும் நிகழ்தகவு

$P(A \cup B) = P(A) + P(B)$ ஆகும்.

நிகழ்தகவு - விளக்கச் சொற்கள் :

(Terms in probablity)

i) **நிகழ்ச்சி** : சோதனையின் முடிவு நிகழ்ச்சியாகும். ஒரு நாணயத்தை சுண்டுதல் சோதனை. அதில் தலை (அ) பூ விழுதல் நிகழ்ச்சி ஆகும்.

ii) **எளிய நிகழ்ச்சிகள்** :
ஒரு நிகழ்ச்சி நடைபெறுவதையோ நடைபெறாததையோ காண்பது எளிய நிகழ்ச்சியாகும்.

iii) **கூட்டு நிகழ்ச்சிகள்** :
இரண்டு அல்லது அதற்கு மேலான நிகழ்ச்சிகள் ஒரே சமயத்தில் நடைபெறுவது கூட்டு நிகழ்ச்சி.

iv) **முழுமையான நிகழ்ச்சிகள்** :
ஒரு சோதனையில் கிடைக்கின்ற எல்லா முடிவுகளும் முழுமை-யான நிகழ்ச்சிகளாகும். (எ.கா) ஒரு தாயக்கட்டையை உருட்டுக கிடைக்கின்ற முடிவுகள் 1, 2, 3, 4, 5, 6 ஆகும்.

v) **இணை நிகழ்ச்சிகள்** :
A, B ஆகிய நிகழ்ச்சி ஒன்றையொன்று புறக்கணிப்பதாகவும், அதே சமயம் முழுமையாகவும் இருந்தால் அவை ஒன்றுக்கு ஒன்று இணையான நிகழ்ச்சிகளாகும்.
(எ.கா) ஒரு தாயக்கட்டை உருட்டப்பட்டால் ஒற்றைப்படை எண்கள் (1, 3, 5) இரட்டைப்படை எண்கள் (2, 4, 6) கிடைப்பதை கூறலாம்.

vi) **புறக்கணிக்கும் நிகழ்ச்சிகள்** :
ஒரு சோதனையில் இரு நிகழ்ச்சிகள் ஒரு சேர நடக்காவிடில் அவை ஒன்றை ஒன்று புறக்கணிக்கும் நிகழ்ச்சி ஆகும். தாயக்கட்டையை உருட்டினால் 6 எண்களில் ஏதேனும் ஒன்று தான் விழும்.

vii) **தனித்த நிகழ்ச்சி** :
ஏதேனும் ஒரு நிகழ்ச்சி நடைபெறுவது மற்ற நிகழ்ச்சியை பாதிக்காது. அதே சமயம் மற்றது நிகழ்வதால் இந்நிகழ்ச்சியும் பாதிக்காது. அப்படியானால் அவை தனித்த நிகழ்ச்சியாகும்.

viii) சார்புள்ள நிகழ்ச்சிகள் :

இரு நிகழ்ச்சிகளில் ஒன்றது மற்றதைப் பாதித்தாலோ, அல்லது மற்றதால் பாதிக்கப்பட்டாலோ அவை இரண்டும் சார்புள்ள நிகழ்ச்சிகளாகும்.

ix) சமவாய்ப்பு நிகழ்ச்சி :

ஒரு சோதனையில் நிகழக்கூடிய எல்லா நிகழ்ச்சிகளும் சமவாய்ப்புள்ள நிகழ்ச்சிகளாகும்.

நிகழ்தகவுத் தேற்றங்கள் :

i) கூட்டல் தேற்றம் :

A, B என்ற நிகழ்ச்சிகள் ஒன்றை ஒன்று புறக்கணிக்கும் நிகழ்ச்சிகள் எனில், A அல்லது B நிகழ்வதற்கான நிகழ்தகவு அவற்றின் தனித்தனி நிகழ்தகவுகளின் கூடல் ஆகும்.

$$p(A \text{ அல்லது } B) = p(A) + p(B)$$

(எ.கா) ஒரு ~~சீட்டுக்~~ சீட்டுக்கட்டில் இருந்து ஒரு ராஜா (அ) ராணி சீட்டு எடுக்கும் நிகழ்தகவைக் காண்க :

தீர்வு:

ராஜா சீட்டு எடுப்பதன் நிகழ்தகவு = $4/52$

ராணி சீட்டு எடுப்பதன் நிகழ்தகவு = $4/52$

ராஜா (அ) ராணி சீட்டு எடுப்பதன் நி.த = $4/52 + 4/52$

$= 8/52 = 2/13$

குறிப்பு :

A, B என்ற இரு நிகழ்ச்சிகளும், ஒன்றை ஒன்று புறக்கணிக்காத நிகழ்ச்சிகளானால்

$$P(A \text{ (அ) } B) = p(A) + (B) - p(A \& B)$$

இயல்நிலைப் பரவலின் இயல்புகள்

(properties of Normal Distribution)

1) இயல்நிலைப் பரவல் மணி வடிவமானது.

2) இது ஒரு கோட்டமற்ற பரவலாகும்.

3) இயல் நிலைப் பரவலின் கூட்டுச் சராசரி, இடைநிலை, முகடு ஆகியவை ஒரே மதிப்பாக அமையும்.

4) இது ஒரு சமச்சீரான பரவலாகும். கூட்டுச்சராசரியிலிருந்து இப்பரவலை இரு பாதிகளாகப் பகுக்கின்ற போது இரு பாதிகளும் ஒன்றுபோல இருக்கும்.

5) இயல்நிலை பரவல் கூட்டுசராசரியின் மதிப்பிற்கு நேராக மிக உயர்ந்து காணப்படும்.

6) இது ஒரே முகடைக் கொண்டிருக்கும்.

7) இப்பரவலின் வளைவுமாற்றப் புள்ளி $\bar{x} \pm \sigma$ ஆகிய புள்ளிகளில் அமையும்.

8) இது ஒரு தொடர்ச்சியான பரவல்.

9) முதலாம் (ம) மூன்றாம் கால்மானங்கள் இயல் நிலையிலிருந்து சமமான தூரத்தில் இருக்கும்.

10) இயல்நிலைப் பரவலின் பரப்பு கீழ்கண்டவாறு அமையும்.

அ) $\bar{x} \pm \sigma$ 68.27 சதவிதப்பரப்பினைக் கொண்டிருக்கும்

ஆ) $\bar{x} \pm 2\sigma$ 95.45 சதவிதப்பரப்பினைக் கொண்டிருக்கும்.

இ) $\bar{x} \pm 3\sigma$ 99.73 சதவிதப் பரப்பினைக் கொண்டிருக்கும்.

ஈ) $\bar{x} \pm 1.96\sigma$ 95 சதவிதப்பரப்பினைக் கொண்டிருக்கும்.

உ) $\bar{x} \pm 2.57\sigma$ 99 சதவிதப் பரப்பினைக் கொண்டிருக்கும்.

Scanned by TapScanner

unit - II

புள்ளியியல் பரவல்கள் (statistical Distributions)

புள்ளியியல் பரவல்களை இரண்டு வகைகளாகப் பிரிக்கலாம். சேகரிக்கப்பட்ட விவரங்களிலிருந்தும், கோட்பாட்டின் அடிப்படையிலும் பரவல்களை உண்டாக்கலாம்.

நிகழ்தகவுக் கொள்கையின் அடிப்படையிலும் பரவல்களை உருவாக்கலாம். அதனைப் நிகழ்தகவு பரவல் (அ) கோட்பாட்டு அலைவெண் பரவல் எனலாம்.

1) ஈருறுப்பு பரவல்:

இது ஒரு தொடர்ச்சியற்ற நிகழ்தகவுப் பரவலாகும். இதனை ஜேகப் பெர்னௌலி (1654 - 1705) என்னும் சுவிடன் கணித அறிஞர் உருவாக்கினார். ஆதலால் பெர்னௌலி பரவல் என்றும் கூறுவர்.

ஈருறுப்புப் பரவலை பொருத்துதல்:

சேகரிக்கப்பெற்ற விவரங்களுக்கு எவ்வாறு ஈருறுப்பு பரவலைப் பொருத்த முடியும் என்பதை காணலாம்.

1) முதல் நிகழ்ச்சிகளின் வெற்றி (p), தோல்வி (q) ஆகியவை காரணமான நிகழ்தகவினை கணக்கிட வேண்டும்.

2) $(q+p)^n$ ஐ விரிவாக்கம் செய்ய வேண்டும்.

3) $(q+p)^n$ இதன் ஒவ்வொரு உறுப்பினையும் மொத்த அலைவெண்ணால் (n) பெருக்கினால் ஒவ்வொரு பிரிவிற்குமான எதிர்பார்க்கின்ற அலைவெண் கிடைக்கும்.

2) இயல்நிலைப் பரவல்:

இது ஒரு தொடர்ச்சியான பரவல் ஆகும். இது ஈருறுப்பு பரவலின் அடிப்படையில் அமைந்தது. ஒர் ஈருறுப்பு பரவலின் p யும் q வும் சமமாக இருப்பினும், இல்லாவிடினும், சோதனைகளின் எண்ணிக்கை n மிகவும் அதிகமாக இருக்கின்ற போது ஈருறுப்பு பரவல் ஒர் இயல்நிலைப் பரவலாக மாறும்.

இயல்நிலைப் பரவலின் இயல்புகள் :

இது ஒரு கோட்பாட்டுப் பரவலாகும்.
இயல்நிலை பரவல் மணி வடிவமானது.
இது ஒரு சமச்சீரான பரவல்.

நேரான இயல்நிலை பரவல்:

X என்ற இயல்நிலை மாறியின் நிகழ்தகவு $-\frac{1}{2}\left(\frac{x-\bar{x}}{\sigma}\right)$
அடர்த்திச் சார்புகள் $p(x) = \frac{1}{\sqrt{2\pi}\sigma}$ ஆகும்.
இந்த இயல்நிலைப் பரவலின் கூட்டு சராசரி $\bar{x}$, திட்ட விலக்கம் σ ஆகும். இதனை சுருக்கமாக $X \sim N(\bar{x}, \sigma)$ என்று குறிப்பிடலாம்.

உண்மைப்பிரிவு இடைவெளி	கீழ் எல்லை	அலைவெண்	$z = \frac{x-\bar{x}}{\sigma}$	பரப்பு 0 – z	ஒவ்வொரு பிரிவிலும் பரப்பு	எதிர் அலைவெண்
49.5 – 52.5	49.5	5	−2.63	0.4957	0.0483	4.825
52.5 – 55.5	52.5	20	−1.62	0.4474	0.2183	21.8 = 22
55.5 – 58.5	55.5	45	−0.61	0.2291	0.3845	38.8 = 39
58.5 – 61.5	58.5	20	0.40	0.1554	0.2647	26.5 = 27
61.5 – 64.5	61.5	10	1.41	0.4201	0.0721	7.5 = 7
	64.5		2.42	0.2022		

z மதிப்புகள் பரப்பளவை அடுத்தடுத்து கழிக்க ஒவ்வொரு பிரிவுக்குமுள்ள பரப்பு கிடைக்கும். ஆனால் z அடையாளம் மாறுகின்ற போது மட்டும் கூட்ட வேண்டும்.

−0.61 க்கும் 0.40 க்கும் இடையேயுள்ள பரப்பு மட்டும் கூட்டப்பட்டுள்ளதைக் கவனிக்கவும் ஒவ்வொரு பிரிவுக்குமுள்ள பரப்பை மொத்த அலைவெண்ணால் பெருக்க எதிர்பார்க்கின்ற அலைவெண்கள் கிடைக்கும்.

unit - III

Testing of hypothesis

எடுகோள் சோதனை :

ஒரு முழுமையான தொகுதிகளின் பண்பளவைப் பற்றி கூறின் பண்பளவைகளைக் கொண்டு முடிவெடுக்கும் சோதனை எடுகோள் சோதனை ஆகும்.

எடுகோள்கள் :

முழுமை தொகுதியின் பண்பளவைகள் பற்றி அனுமானிக்கப் -படும் கூற்று எடுகோள் எனப்படும்.

i) **முழுமை தொகுதியின் பண்பளவைகள்** : பற்றி (values)

ஒவ்வொரு புள்ளியியல் பண்பு அளவைகளையான சராசரி, திட்டவிலக்கம், இடைநிலை, மற்றும் பரவற்படி, ஒட்டுறவு கெழு ஆகியவை ஒரு முழுமை தொகுதியை குறிப்பிடுவதாக இருப்பின் அவை முழுமை தொகுதியின் பண்பளவைகள் எனப்படும்.

ii) **கூறுப் பண்பளவைகள்** : (Sample)

கூறின் உறுப்புகளைப் பயன்படுத்திக் கணக்கிடப்படும் சராசரி, திட்டவிலக்கம், பரவற்படி, ஒட்டுறவு கெழு போன்றவை கூறின் பண்பளவைகள் எனப்படும்.

எடுகோளின் வகைகள் :

எடுகோள் 2 வகைப்படும் :

i) பூஜ்ஜிய எடுகோள்கள் (சூன்ய எடுகோள்கள்)
ii) எதிர் எடுகோள்கள் (எதிர்மறை எடுகோள்கள்)

i) **பூஜ்ஜிய எடுகோள்** :

முழுமை தொகுதியின் பண்பளவைகளும் கூறின் பண்பளவைக்கும் எவ்வித வேறுபாடும் இல்லை அல்லது கொடுக்கப்பட்ட கூறு முழுமை தொகுதியிலிருந்து எடுக்கப்பட்டது என்பதை கூறும் எடுகோள்கள் பூஜ்ஜிய எடுகோள் எனப்படும். இதனை H_0 என்றும் கூறலாம்.

(ex) $\mu \neq \bar{x}$; $H_0 = \mu > 0$ $\mu = \bar{x}$ $\mu = \leq 0$

$H_0 : \mu_1 = \mu_2$

i) $P(\mu - \sigma < x < \mu + \sigma) = 68.26$

ii) $P(\mu - 2\sigma < x < \mu + 2\sigma) = 95.44$

iii) $P(\mu - 3\sigma < x \; \mu + 3\sigma) = 99.73$

ii) **எதிர் எடுகோள்கள் :**

பூஜ்ஜிய எடுகோளுக்கு எதிர் மறையான கூற்று எதிர்கோள் எனப்படும். இதனை H_1 என்று குறிப்பர்.

(ex) $H_1 \neq (\mu_1 \neq \mu_2, \mu \neq \bar{x})$ (or) $\mu \neq \mu_0$

i) $P(\mu - \sigma < x < \mu + \sigma) = 6826$

ii) $P(\mu - 2\sigma - < x < \mu + 2\sigma) = 9544$

iii) $P(\mu - 3\sigma - < x + 3\sigma) = 9973$

=> **இரண்டு வகைப் பிழைகள்: (Types of Errors)**

எடுகோள் சோதனையில் எடுக்கப்படும் பொதுவாக தவறான முடிவுகள் எனப்படும். எடுகோள் சோதனையில் இரண்டு வகையான பிழைகள் ஏற்படுகின்றன. அவை,

i) முதல் வகை பிழைகள்

ii) இரண்டாம் வகை பிழைகள்

$P(\mu - \sigma < x < \mu + \sigma) = 6826$

$P(\mu \, 2\sigma < x < \mu + 2\sigma) = 9544$

$P(\mu \, 3\sigma < x < \mu + 3\sigma) = 9973$

i) **முதல் வகை பிழைகள் :**

H_0 உண்மையான கூற்றாக இருக்கும் போது அதை நிராகரித்த P முதல் வகை பிழைகள் செய்வதற்கான நிகழ்தகவு α எனப்படும்.

ii) இரண்டாம் வகை பிழைகள் :

H_0 பொய்யானதாக இருக்கும் போது அதை ஏற்றுக் கொள்ளுதல் இரண்டாம் வகை பிழை எனப்படும். இரண்டாம் வகை பிழை செய்வதற்கான நிகழ்தகவு β எனப்படும்.

திட்ட பிழை : (Standard Error)

கூறுப் பரவலின் திட்ட விலக்கம் திட்ட பிழை எனப்படும்.

திட்ட பிழையின் பயன்கள்

⇒ எடுகோள் சோதனைகளின் முடிவுகளை எடுப்பதற்கு திட்டப் பிழை இன்றியமையாததாக உள்ளது.

⇒ முழுமை தொகுதியின் பண்பளவைகளின் எல்லையை கணக்கிட உதவுகிறது.

⇒ கூறின் நம்பகத்தன்மையை அறிய உதவுகிறது.