CURVE FITTING

Distributions

1. Binominal
2. Poisson
3. Exponential
4. Geometric
5. Hypergeometric
6. Normal
7. Negative Binomial

$\boxed{\text{STATISTICS – SEM-IV}}$

By
Biswajit Mukherjee.

Fittings

1. Lets assume. $X \sim b(4, P)$ and $E(X) = 4P$

BINOMIAL

X	Frequency
0	7
1	24
2	34
3	29
4	6
	100

$\sum frequency = 100$

$f(x) = {}^nC_x p^x q^{n-x}$

Ans:-

Step 1: Calculate $E(X)$ from the data table

$$E(x) = \frac{0\times7 + 1\times24 + 2\times34 + 3\times29 + 4\times6}{100} = 2.03$$

Step 2: Calculate probability of success and trials

Given,

$4P = E(x)$

we know,

$E(x) = nP$

Comparing these we get

$\boxed{n = 4}$

Now,

$4\hat{P} = 2.03 \quad \therefore \hat{P} = \frac{2.03}{4} \simeq 0.51$

$\therefore \hat{q} = (1 - 0.51) = 0.49$

Step 3:- Find the probabilities distribution at each trial of the table

$f(0) = {}^4C_0 (0.51)^0 (0.49)^{4-0} = 0.057 \simeq \underline{0.6}$

$f(1) = {}^4C_1 (0.51)^1 (0.49)^3 = 0.2496 \simeq \underline{0.25}$

$f(2) = {}^4C_2 (0.51)^2 (0.49)^2 \simeq \underline{0.39}$

$f(3) = {}^4C_3 (0.51)^3 (0.49)^1 \simeq \underline{0.27}$

$f(4) = {}^4C_4 (0.51)^4 (0.49)^0 \simeq \underline{0.09}$

Step 3 : Calculating Expected Frequency

X	f(x)	$E_{f(x)}$
0	54	$90 \times 0.58 = 52$
1	26	$90 \times 0.31 \simeq 28$
2	8	$90 \times 0.08 = 7$
3	2	$90 \times 0.02 = 2$

Σ expected Frequency = 89

Let, $X \sim \exp(\lambda)$

Life time (hrs)	No. of bulbs (f)
< 100	121
100 - 200	78
200 - 300	43
300 & more	58

$\Sigma f = 300$

$f(x) = \lambda e^{-\lambda x}$

EXPONENTIAL

Step 1 : converting the table

Life time (X_{m})	No. of bulbs
50	121
150	78
250	43
350	58

Step 2: Calculate mean

$$E(X) = \frac{1}{\lambda}$$

$$\bar{X} = \frac{(50 \times 121) + (150 \times 78) + (43 \times 250) + (350 \times 58)}{300}$$

$$\cong 163$$

$$\therefore E(x) = \frac{1}{163} \cong 0.066.$$

Step 4: Calculating Probabiliti frequency

$$f(x) = 0.006\, e^{-0.006x}$$

Now, $P[a \leq x < b] = F(b) - F(a) = (1 - e^{-\lambda b}) - (1 - e^{-\lambda a})$

$$= e^{-0.006a} - e^{-0.006b}$$

a and b are upper limit and lower limit

X	f(x)	Probability	E f(x)
0 - 100	121	$e^{0} - e^{-0.006 \times 100} = 0.45$	135
100 - 200	78	$e^{-0.6} - e^{-1.2} = 0.25$	75
200 - 300	43	$e^{-1.2} - e^{-1.8} = 0.14$	42
300 - more	58	$e^{-1.8} - 0 = 0.16$	48

Σ Expected Frequency = 300.

Let, $X \sim Geom(P)$

X (trials until success)	frequency
1	30
2	20
3	10
4	6
5	5

$\boxed{\text{GEOMETRIC}}$

$\Sigma f = 70$

$f(x) = (1-P)^{x-1} P$
[as $x = 1, 2, 3 \dots$]

$f(x) = (1-P)^x P$
[as $x = 0, 1, 2, 3 \dots$]

Step 1: Calculate mean

$$\bar{X} = \frac{(30\times 1) + (20\times 2) + (3\times 10) + (4\times 6) + (5\times 5)}{70} \cong \frac{114}{70} \cong 2.05$$

Step 2: Find probability of success [value of P]

we know,

$$\bar{x} = \frac{1}{P} \quad \text{or,} \quad P = \frac{1}{\bar{x}} = \frac{1}{2.05} \simeq 0.48$$

$\therefore P = 0.48$

$q = (1 - 0.48) = 0.52$

Step 3: Find Probability distribution at each trial

$f(1) = (0.52)^0 \times 0.48 = 0.48$

$f(2) = (0.52)^1 \times 0.48 = 0.25$

$f(3) = (0.52)^2 \times 0.48 = 0.13$

$f(4) = (0.52)^3 \times 0.48 = 0.07$

$f(5) = (0.52)^4 \times 0.48 = 0.04$

Step 4: Calculating Expected Frequency. $[N \cdot f(x)]$

X	f	$Ef(x) = N \cdot f(x)$	
1	30	$70 \times 0.48 =$	34
2	20	$70 \times 0.25 =$	18
3	10	$70 \times 0.13 =$	9
4	6	$70 \times 0.07 =$	5
5	4	$70 \times 0.04 =$	3

$\sum$ Expected Frequency = 69.

Lets assume, $X \sim Hgeo(N, K, n)$

HYPERGEOMERIC

Total balls = 20
White = 12
red = 8

5 balls are drawn.

Trial = 100

$$P(x=x) = \frac{\binom{K}{x}\binom{N-K}{n-x}}{\binom{N}{n}}$$

X (Red balls)	Observed frequency
0	5
1	20
2	40
3	25
4	8
5	2

$\sum$ frequency = 100

Step 1:- Calculate Expected probabilities

X	f(x)	Expected probability $E(x)$	$Ef(x)$
0	5	0.511	5
1	20	0.223	22
2	40	0.342	34
3	25	0.263	26
4	8	0.097	10
5	2	0.022	2
	$\sum f(x) = 100$		$\sum Ef(x) = 99$

$$P(x=x) = \frac{\binom{8}{x}\binom{12}{5-x}}{\binom{20}{5}}$$

$$P(x=0) = \frac{\binom{8}{0}\binom{12}{5}}{\binom{20}{5}}$$

$$P(x=1) = \frac{\binom{8}{1}\binom{12}{4}}{\binom{20}{5}}$$

$\vdots$

$$P(x=5) = \frac{\binom{8}{5}\binom{12}{0}}{\binom{20}{5}}$$

Marks of 100 Students

NORMAL

$X \sim N(\mu, \sigma)$

Marks (x)	f
40 - 50	5
50 - 60	18
60 - 70	42
70 - 80	27
80 - 90	8

Step 1 Find Midpoint and S.D an mean

$\overline{X}$	Midpoint(x_m)	f	$f(x_m)$	$(x - x_m)^2$	$f \cdot (x - x_m)^2$
40 - 50	45	5	225	400	2000
50 - 60	55	18	990	100	1800
60 - 70	65	42	2730	0	0
70 - 80	75	27	2025	100	2700
80 - 90	85	8	680	400	3200
		$\Sigma f = 100$	$\Sigma f x_m = 6650$		$\Sigma = 9700$

$\therefore$ Mean $= \dfrac{6650}{100} = 66.5$

$\sigma = \sqrt{\dfrac{9700}{100}} = \sqrt{97} \approx 9.85$

Step-2 :- Convert into Z-Scores.

$Z = \dfrac{X - \mu}{\sigma}$

For, Range 40-50

L. Boundary = 39.5

U. Boundary = 50.5

$Z_1 = \dfrac{39.5 - 66.5}{9.85}$

≈ -2.74

$Z_2 = \dfrac{49.5 - 66.5}{9.85}$

≈ -1.73

Expected Frequency = Area between Z_1 & Z_2 $\times \sum f$

Step 3 : Find Area between z_1 and z_2

$$P(z_1 < Z < z_2) = \Phi(z_2) - \Phi(z_1)$$

$$= \left[\Phi(-2.74) - \Phi(-1.73)\right] = \cancel{0.000} \; 0.0418 - 0.0031$$

$$= 0.0387$$

Step 4 : Final table [Expected Frequency]

Find the Z-scores for all values of the table and find Area between z_2 on z_1

X	Z-range	Area	Expected Frequency
40-50	-2.74 to -1.73	0.0387	~~3.87~~ 4
50-60	-1.73 to -0.71	0.1325	13
60-70	-0.71 to 0.30	0.3602	36
70-80	0.30 to 1.32	0.3021	30
80-90	1.32 to 2.34	0.1585	15
			$\Sigma = 98.$

NEGATIVE BINOMIAL

$$X \sim NB(\mu, \sigma)$$

X (failure of the 3 suc)	f
0	8
1	18
2	24
3	20
4	10
5	5
	$\Sigma f = 85$

$$f(x) = \binom{x+r-1}{r-1}(1-p)^x p^r$$

$$\boxed{r=3}$$

Step 1: Compute Mean

$$\bar{x}_\mu = \frac{\sum x \cdot f}{N} = \frac{0\times8 + 1\times18 + 2\times24 + 3\times20 + 4\times10 + 5\times5}{85}$$

$$= \frac{217}{85} \approx 2.55$$

for negative binomial.

$$\mu = \frac{r(1-p)}{p} \Rightarrow 2.55 = \frac{3(1-p)}{p}$$

$$p = \frac{3}{5.55} \approx 0.5405$$

Step 2: Calculate Expected Frequency

$$f(x) = \binom{x+3-1}{3-1}(1-0.5405)^x (0.5405)^3$$

$f(0) = 0.1581$

$f(1) = 0.2178$

$f(2) = 0.2001$

$f(3) = 0.1536$

$f(4) = 0.1061$

$f(5) = 0.0684$

X	f	$f(x) \times N$
0	8	$0.1581 \times 85 = 13$
1	18	$0.2178 \times 85 = 19$
2	24	$0.2001 \times 85 = 17$
3	20	$0.1536 \times 85 = 13$
4	10	$0.1061 \times 85 = 9$
5	5	$0.0684 \times 85 = 6$
	$\Sigma f = 85$	$\Sigma Nf(x) = 77$

ance = $np(1-p)$

$$\boxed{\text{Variance} < \text{Mean}}$$

Negative Binomial

Mean = $\dfrac{r(1-p)}{p}$; Variance = $\dfrac{r(1-p)}{p^2}$

$$\boxed{\text{Variance} > \text{Mean}}$$

Poisson Distribution

Mean = λ Variance = λ

$$\boxed{\text{Mean} \approx \text{Variance}}$$

Exponential Distribution

It's continious distribution.

It will always have a range of X

Lifetime data / waiting data

Normal Distribution

Mean $\approx$ Median $\approx$ Mode.

Continous ~~data~~ distribution

grouped data.