Manash Dev Bhatta
221224

2024, Spring

Date
Page

1a) The factories produces two type of car batteries i.e. battery A & battery B. An experiment shows the life of batteries in days which were recorded as follows.

| Life in days: | 500-700 | 700-900 | 900-1100 | 1100-1300 | 1300-1500 |
|---|---|---|---|---|---|
| Battery A | 5 | 11 | 26 | 10 | 8 |
| Battery B | 4 | 30 | 12 | 8 | 6 |

compare the variability of two type of batteries using coefficient of variation.

↳ Solution:

Case-I:

| Life(in days) | Mid (x) | Battery A (f) | $x^2$ | $fx$ | $fx^2$ |
|---|---|---|---|---|---|
| 500-700 | 600 | 5 | $3.6 \times 10^5$ | 3000 | $1.8 \times 10^6$ |
| 700-900 | 800 | 11 | $6.4 \times 10^5$ | 8800 | $7.04 \times 10^6$ |
| 900-1100 | 1000 | 26 | $1 \times 10^6$ | 26000 | $2.6 \times 10^7$ |
| 1100-1300 | 1200 | 10 | $1.44 \times 10^6$ | 12000 | $1.44 \times 10^7$ |
| 1300-1500 | 1400 | 8 | $1.96 \times 10^6$ | 11200 | $1.568 \times 10^7$ |
| | | $N = 60$ | | $\Sigma fx = 61000$ | $\Sigma fx^2 = 64.92 \times 10^6$ |

Now,

$$S.D = \sqrt{\left(\frac{\Sigma fx^2}{N}\right) - \left(\frac{\Sigma fx}{N}\right)^2}$$

$$\bar{x} = \Sigma fx / N = 1016.67$$

$$= \sqrt{\left(\frac{64.92 \times 10^6}{60}\right) - \left(\frac{61000}{60}\right)^2}$$

$$= \sqrt{1.082 \times 10^6 - 1.033 \times 10^6}$$

$$= \sqrt{49000}$$

$$= 221.35$$

CamScanner

coefficient of variation

$$C.V = \frac{\sigma}{\bar{x}} * 100\%$$

$$= \frac{221.35}{1016.67} * 100\%$$

$$= 21.77\%$$

Case II:

| Life (in days) | Mid (x) | Battery (B) (f) | $x^2$ | $fx$ | $fx^2$ |
|---|---|---|---|---|---|
| 500 - 700 | 600 | 4 | $3.6 * 10^5$ | 2400 | $1.44 * 10^6$ |
| 700 - 900 | 800 | 30 | $6.4 * 10^5$ | 24000 | $1.92 * 10^7$ |
| 900 - 1100 | 1000 | 12 | $1 * 10^6$ | 12000 | $1.2 * 10^7$ |
| 1100 - 1300 | 1200 | 8 | $1.44 * 10^6$ | 9600 | $1.152 * 10^7$ |
| 1300 - 1500 | 1400 | 6 | $1.96 * 10^6$ | 8400 | $1.176 * 10^7$ |
| | | $N = 60$ | | $\Sigma fx = 56400$ | $\Sigma fx^2 = 55.92 * 10^6$ |

$$\bar{x} = \Sigma fx / N = 56400 / 60 = 940$$

$$\sigma = \sqrt{\frac{\Sigma fx^2}{N} - \left(\frac{\Sigma fx}{N}\right)^2}$$

$$= \sqrt{\frac{55.92 * 10^6}{60} - (940)^2}$$

$$= \sqrt{932000 - 883600}$$

$$= \sqrt{48400}$$

$$= 220$$

Coefficient of variation:

$$C.V. = \frac{\sigma}{\bar{x}} * 100\%$$

$$= 23.40\ \%$$

$\therefore$ As ~~coff~~ coefficient of variation of Battery B is greater than that of Battery A. so variability of Battery A is more than that of Battery B.

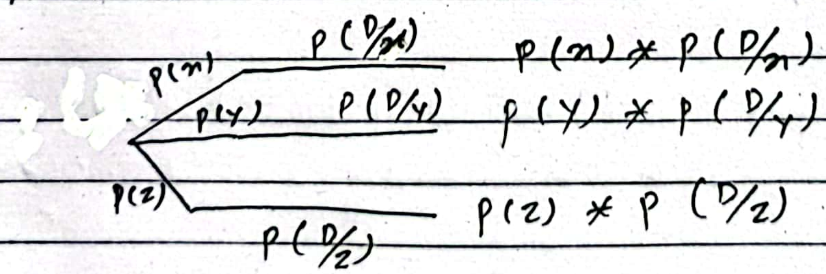
1b) If $E_1, E_2, E_3, \ldots E_n$ be mutually disjoint events of sample space S with $P(E_i) \neq 0$, $i = 1,2,3,\ldots n$ then for any arbitary events A which is subset of S such that $P > 0$,

We have,

$$P(E_i/A) = \frac{P(E_i) \cdot P(A/E_i)}{\sum_{i=1}^{n} P(E_i) * P(A/E_i)}$$

Now,

| | Machine X | Machine Y | Machine Z |
|---|---|---|---|
| production: | 25% | 35% | 40% |
| defective | 5% | 4% | 2% |

i) 

i) The probability that it is defective.

$$P(D) = P(D/X) \cdot P(X) + P(D/Y) \cdot P(Y) + P(D/Z) \cdot P(Z)$$

$$= 0.05 * 0.25 + 0.04 * 0.35 + 0.02 * 0.40$$

$$= 0.034$$

$$= 3.4\%$$

ii) $$P(Y/D) = \frac{P(Y) \cdot P(D/Y)}{P(D)}$$

$$= \frac{0.35 * 0.04}{0.034}$$

$$= 0.41$$

$$= 41\%$$

CamScanner

$E(x) \rightarrow$ Mean
$V(x) \rightarrow$ variance

Date :
Page :

2a.

| X | P(x) | x·p(x) | $x^2$ |
|---|---|---|---|
| -2 | 0.2 | -0.4 | 4 |
| -1 | 0.1 | -0.1 | 1 |
| 0 | 0.3 | 0 | 0 |
| 1 | 0.3 | 0.3 | 1 |
| 2 | 0.1 | 0.2 | 4 |
| Total | 1 | 0 | |

i) $E(x) = \sum x \cdot p(x)$
$= 0$

ii) $E(2x-3) = 2.E(x) - 3$
$= 2 * 0 - 3$
$= -3$

iii) $V(x) = E(x^2) - (E(x))^2$

$E(x^2) = 4(0.2) + 1(0.1) + 0(0.3) + 1(0.3) + 4(0.1)$
$= 1.6$

$\therefore V(x) = 1.6 - 0$
$= 1.6$

iv) $V(2x-3) = 2^2 \, v(x)$
$= 4 * 1.6$
$= 6.4$

$p(x=x) = (e^{-\lambda} \lambda^{x})/x!$ , $\lambda = np$ & $\bar{x} = \lambda$ = variance.



# Poisson Distribution

26→ $\bar{x}$ = 3 calls per minute on an average.
Receiving of calls follows a poisson distribution

i) No call in one minute interval.

$$p(X=0) = \frac{e^{-3}(+3)^0}{0!}$$

$= 0.049$

ii) At least 3 calls in one minute interval.

$$p(X \geq 3) = 1 - p(X < 3)$$
$$= 1 - \{p(X=0) + p(X=1) + p(X=2)\}$$
$$= 1 - \{0.049 + (+0.149) + 0.224\}$$
$$= 1 - \{0.422\}$$
$$= 0.578$$

iii) At most 2 calls in 5 minute interval.

$p(Y \leq 2)$

Where Y follows a poisson distribution with mean $(\lambda) = 15$

i.e. $5 \times 3$ (for 5 minute interval)

Then,

$$p(Y \leq 2) = p(Y=0) + p(Y=1) + P(Y=2)$$
$$= 0.30 \times 10^{-6} + 4.58 \times 10^{-6} + 3.44 \times 10^{-5}$$
$$= 3.92 \times 10^{-5}$$

$p(x=x) = {}^nC_x p^x q^{n-x}$, mean $= np$, variance $= npq$, ${}^nC_x = n!/(n-x)!x!$

Binomial Distribution'

Date
Page

26- OR

Here,

$n=5$ (Five independent trials)

$p(x=1) = 0.4096$

$p(x=2) = 0.2048$

$p=?$

$p(x=1) = {}^nC_x p^x q^{n-x}$

$0.4096 = {}^5C_1 p q^4$

$0.4096 = 5pq^4$ --- (i)

Mean $(\bar{x}) = np$
$= 5*p$
$= 5p$

$p(x=2) = {}^5C_2 p^2 q^3$

$0.2048 = 10p^2q^3$ -- (ii)

From (i)

$p = \frac{0.4096}{5q^4}$

$= \frac{0.4096}{5(1-p)^4}$

From (ii)

$p = \frac{0.2048}{10pq^3}$

$= \frac{0.2048}{10p(1-p)^3}$

comparing both'

$\frac{0.4096}{5(1-p)^4} = \frac{0.2048}{10p(1-p)^3}$

$0.4096(10p(1-p)^3) = 0.2048(5(1-p)^4)$

Dividing (ii) by (i)

or, $\frac{0.2048}{0.4096} = \frac{10p^2q^3}{5pq^4}$

or, $\frac{1}{2} = \frac{2p}{q}$

or, $(1-p) = 4p$

or, $1 = 5p$

or $p = \frac{1}{5} = 0.2$

parameter p of the distribution is

0.2

3a→ Rectangular/uniform distribution:-

A continuous random variable x is said to have uniform distribution over an interval [a,b], if its probability density density function is given by

$$f(x) = \begin{cases} \frac{1}{b-a} & \text{if } a \le x \le b \\ 0 & \text{otherwise.} \end{cases}$$

Mean $(E(x)) = (b+a)/2$

Variance $(V(x)) = \frac{(b-a)^2}{12}$

Date ____
Page ____

Derivation of Mean (E(n))

$$E(x) = \int_a^b x f(x)\, dx$$

$$= \int_a^b x \cdot 1/_{b-a} \, dx$$

$$= 1/_{b-a} \int_a^b x \, dx$$

$$= 1/_{b-a} \left( x^2/_2 \right)_a^b$$

$$= 1/_{2(b-a)} \cdot (b^2 - a^2)$$

$$= \frac{(b-a)(b+a)}{2(b-a)}$$

$$= (b+a)/_2$$

Derivation of variance (V(n))

$$V(x) = E(x^2) - (E(x))^2$$

$$E(x^2) = \int_a^b x^2 f(x) \, dx$$

$$= 1/_{3(b-a)} \quad (b - a)\,(b^2 + ab + a^2)$$

$$= (b^2 + ab + a^2)/_3$$

Then,

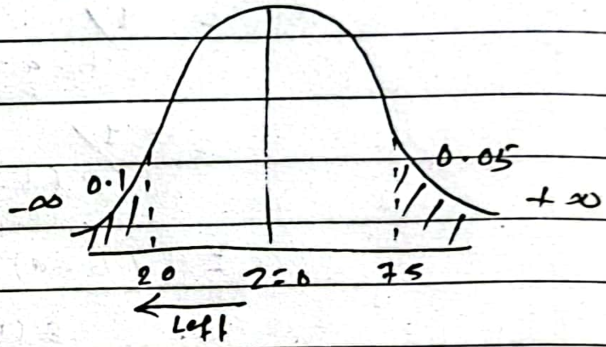
$$V(x) = \frac{(b^2 + ab + a^2)}{3} - \frac{(b+a)^2}{4}$$

$$= \frac{4b^2 + 4ab + 4a^2 - b^2 - 2ab - b^2}{}$$

$$= \frac{b^2 - 2ab + b^2}{12}$$

$$= \frac{(b-a)^2}{12}$$

36→ 10% of the students get less than 20 marks.
5% " " " " over 75 marks.
Normal Distribution
Mean & SD = ?




Here,

$p(x < 20) = 0.1$

$p(x > 75) = 0.05$

Now,

$p(x<20) = 0.1$

$= 0.5 - 0.1$

$= 0.4$

$Z \Rightarrow 1.29$

As it is left of $z=0$, $p(X<20) = -1.29$

$$Z = \frac{x-\mu}{\sigma}$$

$$-1.29 = \frac{20-\mu}{\sigma} \quad \text{--- (1)}$$

Again

$p(X>75) = 0.05$

$= 0.5 - 0.05$

$= 0.45$

$Z \Rightarrow 1.64$

$$Z = \frac{x - \mu}{\sigma}$$

$$1.64 = \frac{75 - \mu}{\sigma} \quad \text{--- (ii)}$$

solving (i) & (ii)

$\sigma = 18.77$

$\mu = 44.21$

4a. Joint probability function of random variable $X$ & $Y$

$$f(x,y) = \begin{cases} k(2x+y) & \text{for } 0 \le x \le 2,\ 0 \le y \le 3 \\ 0 & \text{otherwise.} \end{cases}$$

i) Marginal Density of $X$ & $Y$

$$f(x) = f_x(x) = \int_{-\infty}^{\infty} f(x,y)\,dy \ ; \quad -\infty < y < \infty$$

$$= \int_0^3 k(2x+y)\,dy \quad \text{--- (i)}$$

$$f(y) = f_y(y) = \int_{-\infty}^{\infty} f(x,y)\,dx \ ; \quad -\infty < x < \infty$$

$$= \int_0^2 k(2x+y)\,dx \quad \text{--- (ii)}$$

For value of $k$

Date
Page

For continuous random variable, joint probability densitiy function

$$\int_{-\infty}^{\infty}\int_{-\infty}^{\infty} f(x,y)\,dx\,dy = 1$$

or, $$\int_0^3\int_0^2 k(2x+y)\,dx\,dy = 1$$

or, $$k\int_0^3 \left(2x^2/2 + yx\right)_0^2 dy = 1$$

or, $$k\int_0^3 (2^2+2y)\,dy = 1$$

or, $$k\left(2y^2/2 + 4y\right)_0^3 = 1$$

or, $$18/2 + 12 = 1/k$$

or $$9+12 = 1/k$$

or, $$k = 1/21$$

From ② $$\int_0^3 1/21\,(2x+y)\,dy$$

or, $$f_x(x) = 1/21\left(2xy + y^2/2\right)_0^3$$

$$= 1/21\left(6x + 9/2\right)$$

CamScanner

From ④

$$\int_0^2 \frac{1}{21}(2x+y)\,dx$$

$$f_Y(y) = \frac{1}{21}\left(\frac{2x^2}{2} + xy\right)_0^2$$

$$= \frac{1}{21}(4 + 2y)$$

ii) Are X & Y independent ?

To test the independent.

$$f(x,y) = f(x) \times f(y)$$

$$\frac{1}{21}(2x+y) = \frac{1}{21}(6x + 9/2) \times \frac{1}{21}(4+2y)$$

$$\frac{1}{21}(2x+y) \neq \frac{1}{21 \times 21}(24x + 12xy + 18 + 18y)$$

As $f(x,y) \neq f(x) \times f(y)$, They aren't independent.

$n < 30 \rightarrow t$- test

Date
Page

4.6-) Here: size
Sample (n) = 20
Mean ($\bar{x}$) = 990
S.D ($\sigma$) = 22 $\Leftarrow S$
Find 95 % confidence interval for mean.

solution:

~~① Null Hypothesis ($H_0$): $\bar{x}$ = 990 hours~~
~~(Mean life of the sample bulb is 990)~~
~~② Alternative Hypothesis ($H_1$): $\bar{x} \neq 990$ hrs~~
~~(Mean life of the sample buld is not equal to 990 hours)~~

~~③ Test statistics under $H_0$~~

$$\cancel{t_{cal} = \frac{\bar{x} - \mu}{S/\sqrt{n}}}$$

confidence interval for single mean

$$C.I = [\bar{x} \pm t_{\alpha/2,(n-1)} * S/\sqrt{n}]$$

$t_{\alpha/2}(n-1) = t_{\alpha/2}(19)$ at 95 % confidence interval

$$df = n-1$$
$$= 19$$

$$t_{\alpha/2}(19) = 2.093$$

CS CamScanner

Date
Page

$$CI = \left[990 \pm 2.093 \times \frac{22}{\sqrt{20}}\right]$$

$$= [990 \pm 10.296]$$

$$= [979.7, 1000.29]$$

So, the 95 % confidence interval for the mean life of the bulbs is 979.7 to 1000.29

5a. → Criterias for a good estimator:

(i) Unbiasedness:
Any estimator $\hat{\theta}$ of $\theta$ is said to be unbiased if it measure what it expected to measure.

i.e. $E(\hat{\theta}) = \theta$

(ii) Consistent estimator:
Any estimator $\hat{\theta}$ is said to be consistent estimator of population parameter of $\theta$ if as the sample size increase, the value of sample statistic becomes very close to value of population parameters.

i.e. $\lim_{n \to \infty} E(\hat{\theta}) \longrightarrow \theta$

(iii) Efficiency:
An estimator $\hat{\theta}$ is said to be efficient estimator of parameter $\theta$ if the variance of estimator $\hat{\theta}$ is least with comparison to variance of all other estimator of $\theta$. Thus, if $\hat{\theta}_1$ & $\hat{\theta}_2$ are consistent estimator of $\theta$. i.e. $var(\hat{\theta}_1) < var(\hat{\theta}_2)$ for all $n$

$$\mathcal{E} = \frac{var(\hat{\theta_1})}{var(\hat{\theta_2})} < 1$$

(iv) sufficient:

A statistics is said to be sufficient estimator of parameter if it contains all the information in the comple about parameters.

56.

| (prior) X | (After) Y | $d = y - x$ | $d^2$ |
|---|---|---|---|
| 84 | 90 | 6 | 36 |
| 48 | 58 | 10 | 100 |
| 36 | 56 | 20 | 400 |
| 37 | 49 | 12 | 144 |
| 54 | 62 | 8 | 64 |
| 69 | 81 | 12 | 144 |
| 83 | 84 | 1 | 1 |
| 96 | 86 | -10 | 100 |
| 90 | 84 | -6 | 36 |
| 65 | 75 | +10 | 100 |
| | | $\Sigma d = 63$ | $\Sigma d^2 = 1125$ |

Is training effective? $\alpha = 5\%$.

$\bar{d} = \Sigma d / n$ $\quad n = 10$

$= 63/10 = 6.3$

$$s_d^2 = 1/_{n-1} \left( \Sigma d^2 - \frac{(\Sigma d)^2}{n} \right)$$

$$= 1/_9 \left( 1125 - \frac{(63)^2}{10} \right)$$

$= 80.9$

i.e. $s_d = \sqrt{80.9}$
$= 8.99$

① Null Hypothesis : ($H_0$) $\mu_x = \mu_y$
( Training isn't effective)
② Alternative Hypothesis ($H_1$) $\mu_x < \mu_y$
( Training is effective)

③ Test statistics : Under $H_0$

$$t_{cal} = \bar{d} \Big/ \frac{s_d}{\sqrt{n}} = 6.3 \Big/ \frac{8.99}{\sqrt{10}} = 2.216$$

④ Level of significance ($\alpha$) = 5 %.
⑤. Degree of freedom (df) = n-1 = 9

$t_{tab} = t_{\alpha\,(n-1)} = 1.833$ at df = 9

⑥ Decision,
$t_{cal} > t_{tab}$ ($H_0$ rejected)

⑦ conclusion:
Training isn't effective.

Date
Page

5b. OR $1^{st}$ case

Sample size $(n_1) = 500$

$x_1 = 16$ (imperfect articles)

$2^{nd}$ case

Sample size $(n_2) = 100$

$x_2 = 3$ (imperfect articles)

Has machine improved? $\alpha = 5\%$.

$P_1 = x_1/n_1 = 16/500 = 0.032$

$P_2 = x_2/n_2 = 3/100 = 0.03$

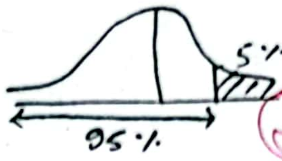
$Q_1 = 1 - P_1 = 0.968$

$Q_2 = 1 - P_2 = 0.97$

1. Null Hypothesis $(H_0)$: $P_1 = P_2$
   Machine hasn't improved
2. Alternative Hypothesis $(H_1)$: $P_1 > P_2$ (One tail)
   Machine has improved
3. Test statistics under $H_0$

$$Z_{cal} = \frac{(P_1 - P_2)}{\sqrt{\frac{P_1 Q_1}{n_1} + \frac{P_2 Q_2}{n_2}}}$$

$$= 0.1064$$




Date ______
Page ______

4. Level of significance ($\alpha$) = 5%.

$Z_{tab} = p(z = (0.95 - 0.5))$ ← From 50% z-table
$= p(z = 0.45)$
$= 1.645$

5. Decision

$Z_{cal} < Z_{tab}$ $H_0$ (accepted)

6. conclusion :
Machine hasn't improved

6a.

| Gender / smoking | Men | Women | Total |
|---|---|---|---|
| Yes | 54 | 32 | 86 |
| NO | 46 | 68 | 114 |
| Total | 100 | 100 | 200 ← N |

Association between smoking habits & gender?
$\alpha = 5\%$
Using chi-squared test.

| Observed (O) | Expected (E) $= \frac{R.T \times C.T}{N}$ | $\frac{(O-E)^2}{E}$ |
|---|---|---|
| 54 | 43 | 2.81 |
| 32 | 43 | 2.81 |
| 46 | 57 | 2.12 |
| 68 | 57 | 2.12 |

$\Sigma \frac{(O-E)^2}{E} = 9.86$

Date ______
Page ______

(i) Null Hypothesis ($H_0$):
There is no association between smoking habit & gender.

(ii) Alternative Hypothesis ($H_1$):
There is association between smoking habit & gender.

(iii) Test statistics under $H_0$

$$\chi^2 = \sum \frac{(O-E)^2}{N} = 9.86$$

(iv) Level of significance ($\alpha$) = 5%.

(v) Degree of freedom

$$df = \underset{\text{No. of columns}}{(c-1)} * \underset{\text{No. of rows}}{(r-1)} = (2-1) * (2-1)$$
$$= 1$$

(vi) At $\alpha = 0.05$ with $df = 1$

$$\chi^2 = 3.841$$

(vii) Decision

$$\chi^2_{cal} > \chi^2_{tab} \quad (H_0 \text{ rejected})$$

(viii) There is association between smoking habit & gender.

obtain

$$r = \frac{n\Sigma XY - \Sigma X \Sigma Y}{\sqrt{n\Sigma x^2 - (\Sigma x)^2}\sqrt{n\Sigma y^2 - (\Sigma y)^2}}$$

Coefficient of determination $= r^2$

Date
Page

66. Age (x) 56 42 36 47 49 42 60 72
Weight (Y) 147 125 118 128 145 140 155 160

i) correlation coefficient between x & y & coefficient of determination.

| X | Y | $X^2$ | $Y^2$ | XY |
|---|---|---|---|---|
| 56 | 147 | 3136 | 21609 | 8232 |
| 42 | 125 | 1764 | 15625 | 5250 |
| 36 | 118 | 1296 | 13924 | 4248 |
| 47 | 128 | 2209 | 16384 | 6016 |
| 49 | 145 | 2401 | 21025 | 7105 |
| 42 | 140 | 1764 | 19600 | 5880 |
| 60 | 155 | 3600 | 24025 | 9300 |
| 72 | 160 | 5184 | 25600 | 11520 |

n = 8

$\Sigma X = 404$ $\Sigma Y = 1118$ $\Sigma X^2 = 21354$ $\Sigma Y^2 = 157792$ $\Sigma XY = 57551$

Now,

correlation coefficient (r) $= \frac{n\Sigma XY - \Sigma X \Sigma Y}{\sqrt{n\Sigma X^2 - (\Sigma X)^2}\sqrt{n\Sigma Y^2 - (\Sigma Y)^2}}$

$= \frac{8736}{87.26 * 111.40}$

$= 0.89$

∴ There is high degree of correlation between Age(x) & Weight (Y)

coefficient of determination $(r^2) = (0.89)^2$

$= 0.80$

ii) Regression line of y on x

$$y = a + bx \quad \text{--- (I)}$$

The value of constants 'a' & 'b' can be estimated by fitting two normal equation by least square estimation.

$$\Sigma y = an + b\Sigma x \quad \text{--- (II)}$$

$$\Sigma xy = a\Sigma x + b\Sigma x^2 \quad \text{--- (III)}$$

From above table' equation (II) & (III) becomes:

$$1118 = 8a + 404b \quad \text{--- (IV)}$$

$$57551 = 404a + 21854b \quad \text{--- (V)}$$

Solving (IV) & (V)

$a = 81.82$

$b = 1.14$

From (I)

Regression line of y on x

$\Rightarrow y = 81.82 + 1.14x$

iii) Weight of women whose age is 45 yrs.

i.e. $y = 81.82 + 1.14 \times 45$

$= 133.12$

$\therefore$ Weight of women whose age is 45 $\Rightarrow$ 133.12 kg.

7a. Hypergeometric distribution:

→ The distribution of discrete random variable 'x' is said to be hypergeometric distribution with parameter $n, m, N$, if it assume only non-negative value & its probability mass function is given by

$$p(X=x) = h(x; n, m, N)$$

$$= \frac{\binom{m}{x}\binom{N-m}{n-x}}{\binom{N}{n}}$$

Where: $n$ = sample size

$N$ = population size

$m$ = No. of success in a lot.

Its Mean $E(x) = n \cdot m/N = np$

Its variance $V(x) = \left(\frac{N-n}{N-1}\right) npq$

c. Properties of correlation

i. A measure of strength & direction of linear relationship between two variables is correlation.

mathematically:

$$\text{correlation}(r) = \frac{\text{cov}(x,y)}{\sqrt{\text{var}(x)}\sqrt{\text{var}(y)}}$$

$$= \frac{n\Sigma XY - \Sigma X \Sigma Y}{\sqrt{n\Sigma X^2 - (\Sigma X)^2}\sqrt{n\Sigma Y^2 - (\Sigma Y)^2}}$$

It's range lies between -1 to 1

* Properties:

(i) If $r=0$, there is no correlation coefficient.
(ii) If $r=1$; there is perfectly positive correlation.
(iii) If $r=-1$; there is perfectly negative correlation.
(iv) If $r = 0-0.5$; Low degree of correlation.
(v) If $r = 0.5-0.75$; Moderate correlation.
(vi) If $r = 0.75-0.99$; High degree of correlation.

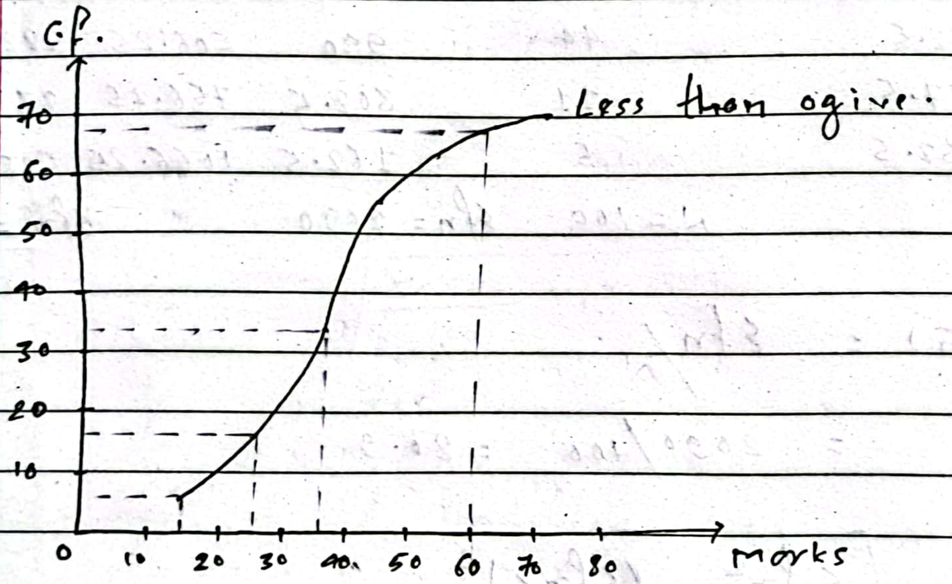
2024, ~~Spring~~ Fall

Date
Page

1a.

| obtained marks | 10-20 | 20-30 | 30-40 | 40-50 | 50-60 | 60-70 | 70-8 |
|---|---|---|---|---|---|---|---|
| No. of students | 7 | 10 | 17 | 20 | 8 | 6 | 2 |

| Marks | No. of students | c.f. |
|---|---|---|
| 10-20 | 7 | 7 |
| 20-30 | 10 | 17 |
| 30-40 | 17 | 34 |
| 40-50 | 20 | 54 |
| 50-60 | 8 | 62 |
| 60-70 | 6 | 68 |
| 70-80 | 2 | 70 |




i) No. of students securing less than 35 marks is 34

ii) For between 25 & 60 marks

No. of students securing less than 25 is 17

No. of " " " " 60 is 67

Between 25 & 60 marks is 67-17

i.e. 50

iii) More than 15 marks
No. of students securing marks less than 15 is 7

More than 15 marks is 70-7
i.e. 63

1b → Case-I

| Bursting pressure | Mid value $x$ | No. of polythene bags A $(f)$ | $fx$ | $x^2$ | $fx^2$ |
|---|---|---|---|---|---|
| 5-10 | 7.5 | 2 | 15 | 56.25 | 112.5 |
| 10-15 | 12.5 | 9 | 112.5 | 156.25 | 1406.25 |
| 15-20 | 17.5 | 29 | 507.5 | 306.25 | 8881.25 |
| 20-25 | 22.5 | 44 | 990 | 506.25 | 22275 |
| 25-30 | 27.5 | 11 | 302.5 | 756.25 | 8318.75 |
| 30-35 | 32.5 | 5 | 162.5 | 1056.25 | 5281.25 |
| | | $N = 100$ | $\Sigma fx = 2090$ | | $\Sigma fx^2 = 46275$ |

$$\text{Mean } (\bar{x}) = \Sigma fx / N$$

$$= 2090/100 = 20.9$$

$$S.D\ (\sigma) = \sqrt{\frac{\Sigma fx^2}{N} - \left(\frac{\Sigma fx}{N}\right)^2}$$

$$= \sqrt{\frac{46275}{100} - (20.9)^2}$$

$$= 5.09$$

Coefficient of variation (C.V)

$$C.V = \sigma/\bar{x} \times 100\%$$

$$= \frac{5.09}{20.9} \times 100\% = 24.36\%$$

Case - II

| Bursting pressure | Mid value $x$ | No. of polythene bags B ($f$) | $fx$ | $x^2$ | $fx^2$ |
|---|---|---|---|---|---|
| 5-10 | 7.5 | 9 | 67.5 | 56.25 | 506.25 |
| 10-15 | 12.5 | 11 | 137.5 | 156.25 | 1718.75 |
| 15-20 | 17.5 | 18 | 315 | 306.25 | 5512.5 |
| 20-25 | 22.5 | 32 | 720 | 506.25 | 16200 |
| 25-30 | 27.5 | 17 | 467.5 | 756.25 | 12856.25 |
| 30-35 | 32.5 | 13 | 422.5 | 1056.25 | 13731.25 |

$N = 100$ $\quad \Sigma fx = 2130$ $\quad \Sigma fx^2 = 50525$

$$\text{Mean } (\bar{x}) = \Sigma fx / N$$

$$= 2130/100 = 21.3$$

$$SD\ (\sigma) = \sqrt{\frac{\Sigma fx^2}{N} - \left(\frac{\Sigma fx}{N}\right)^2}$$

$$= \sqrt{\frac{50525}{100} - (21.3)^2}$$

$$= 7.18$$

Coefficient of variation (C.V)

$$C.V = \sigma/\bar{x} * 100\%$$

$$= 7.18/21.3 * 100\%$$

$$= 33.71\%$$

i) Avg. life of Model A & Model B is 20.9 & 21.3 respectively

ii) Model A has greater uniformity because C.V. of A is less than that of B.

$$24.36\% < 33.71\%$$

2a. A husband & wife appear in an interview for two vacancies in the same post.

$p(H) = 1/7$

$p(W) = 1/5$

i) Both of them will be selected.

$p(\text{Both selected}) = p(H) * p(W)$

$= 1/7 * 1/5 = 1/35$

ii) only one of them will be selected.

$p(\text{only one selected}) = p(H \text{ and } W') + p(W \text{ and } H')$

$= p(H) * p(W') + p(W) * p(H')$

$= 1/7 * 4/5 + 1/5 * 6/7$

$= 4/35 + 6/35$

$= 10/35$

$= 2/7$

iii) None of them will be selected

$p(\text{None selected}) = 4/5 * 6/7 = 24/35$

CamScanner

Date
Page

iv) Husband will be selected but not wife.

$p(\text{Husband only selected}) = \frac{1}{7} * \frac{4}{5} = \frac{4}{35}$

v) Wife will be selected but not husband.

$p(\text{Wife only selected}) = \frac{1}{5} * \frac{6}{7}$

$= \frac{6}{35}$

26. Here:

$p(C) = 0.45$ $\quad p(T/C) = 0.05$

$p(E) = 0.30$ $\quad p(T/E) = 0.04$

$p(O) = 0.25$ $\quad p(T/O) = 0.02$

$p(E/T) = ?$

$$p(E/T) = \frac{p(E) * p(T/E)}{p(T)} \quad \text{--- (i)}$$

$p(T) = p(T/C) * p(C) + p(T/E) * p(E) + p(T/O) * p(O)$

$= 0.05 * 0.45 + 0.04 * 0.30 + 0.02 * 0.25$

$= 0.0395$

(i) is $p(E/T) = \frac{0.30 * 0.04}{0.0395} = 0.30$

30.

| X | 2 | 3 | 4 | 5 | 6 | 7 |
|---|---|---|---|---|---|---|
| F(x) | 0.1 | k | 0.2 | 2k | 0.3 | k |

For value of k

$\sum f(x) = 1$

$4k + 0.6 = 1$

$k = 0.4/4$

$= 0.1$

| X | F(x) | X F(x) | $x^2$ | $x^2 F(x)$ |
|---|---|---|---|---|
| 2 | 0.1 | 0.2 | 4 | 0.4 |
| 3 | 0.1 | 0.3 | 9 | 0.9 |
| 4 | 0.2 | 0.8 | 16 | 3.2 |
| 5 | 0.2 | 1 | 25 | 5.0 |
| 6 | 0.3 | 1.8 | 36 | 10.8 |
| 7 | 0.1 | 0.7 | 49 | 4.9 |
| | | $\sum x F(x) = 4.8$ | | $\sum x^2 F(x) = 25.2$ |

Mean $(E(x)) = \sum x F(x)$

$= 4.8$

$S.D = \sqrt{E(x^2) - (E(x))^2}$

$= \sqrt{25.2 - (4.8)^2}$

$= \sqrt{2.16}$

$= 1.46$

$E(2x-9) = 2 \times E(x) - 9$

$= 2 \times 4.8 - 9$

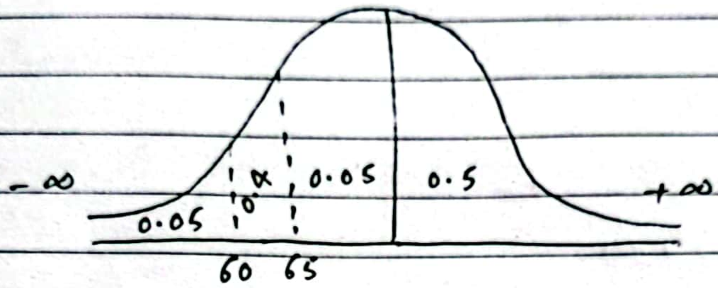
$= 0.6$

$v(5x+2) = 5^2 v(x)$

$= 25 * 2.16$

$= 54$

4a.




$p(x<60) = 0.5 - 0.05$

$= 0.45$

$= 1.64$ (from normal table)

$z = (x - \mu)/\sigma$

$-1.64 = \dfrac{60 - \mu}{\sigma}$ --- (i)

$p(x>65) = 0.5 - 0.45$

$= 0.05$

$= 0.13$ (from normal table)

$z = (x - \mu)/\sigma$

$-0.13 = (65 - \mu)/\sigma$ --- (ii)

Solving (i) & (ii)

$\sigma = 3.31$ ← standard deviation

$\mu = 65.43$ ← mean

46. Joint probability distribution of X & Y

$$f(x,y) = 4xy\,e^{-(x^2+y^2)} \;;\; x>0,\; y>0$$

i) Independence:

$$f(x) = \int_{-\infty}^{\infty} f(x,y)\,dy \qquad -\infty < y < \infty$$

$$= \int_{1}^{\infty} 4xy\,e^{-(x^2+y^2)}\,dy$$

$$= 4x\int_{1}^{\infty} y\,e^{-x^2} * e^{-y^2}\,dy$$

$$= 4x \cdot e^{-x^2}\int_{1}^{\infty} y e^{-y^2}\,dy$$

$$= 4x \cdot e^{-x^2}\int_{1}^{\infty} e^{-u}\,\frac{du}{2}$$

Let $u = y^2$

$\frac{du}{dy} = 2y$

$\frac{du}{2} = y\,dy$

Limits

$y \to 1,\; u \to 1$

$y \to \infty,\; u \to \infty$

$$= \frac{4x \cdot e^{-x^2}}{2}\int_{1}^{\infty} e^{-u}\,du$$

$$= 2x e^{-x^2}\left[\frac{e^{-u}}{-1}\right]_1^{\infty}$$

$$= -2x e^{-x^2}\left(e^{-\infty} - e^{-1}\right)$$

$$= -2x \cdot e^{-x^2}(-e^{-1})$$

$$= 2x\,e^{-(x^2+1)}$$

$$f(y) = \int_1^\infty f(x)\,dx$$

$$= \int_1^\infty 4xy\,e^{-(x^2+y^2)}\,dx$$

$$= 4y\,e^{-y^2}\int_1^\infty x e^{-x^2}\,dx$$

$$= 2y\,e^{-(y^2+1)}$$

For independence:

$$f(x,y) = f(x) \times f(y)$$

$$4xy\,e^{-(x^2+y^2)} \neq 2x\,e^{-(x^2+1)} \times 2y\,e^{-(y^2+1)}$$

$$\neq 4xy\,e^{-x^2-1+(-y^2-1)}$$

$$\neq 4xy\,e^{-(x^2+y^2)-2}$$

X & Y are not independent.

ii) conditional density X given Y = y

i.e. $$f(x/y) = \frac{f(x,y)}{f(y)}$$

$$= \frac{4xy\,e^{-(x^2+y^2)}}{2y\,e^{-(y^2+1)}}$$

$$= 2x\,e^{-x^2-y^2+y^2+1}$$

$$= 2x\,e^{-x^2+1}$$

Date
Page

5a. Solution:

Sample size$(n) = 25$

$S.D = 100$

$\bar{x} = 350$ (sample mean)

Standard error of the mean
(When $\sigma$ is known & population $\infty$)

$$S.E(\bar{x}) = \sigma/\sqrt{n}$$

$$= 100/\sqrt{25}$$

$$= 100/5 = 20$$

iii) 95% CI

$$CI = \bar{x} \pm t_{\alpha/2\,(n-1)} \cdot s/\sqrt{n}$$

$$= 350 \pm t_{\alpha/2\,(24)} * 20$$

$$= 350 \pm 2.093 * 20$$

$$= 350 \pm 41.86$$

$$= [308.14, 391.86]$$

Degree of freedom

$t_{\alpha/2}$ 24 at 5% level of significance

$\Rightarrow$ 2.093

As $\sigma$ is known

For 95% confidence, $Z_{\alpha/2} = 1.96$

$CI = \bar{x} \pm Z_{\alpha/2} * SE$

$= 350 \pm 39.2$

$= [310.8, 389.2]$