17.12.24 (TUESDAY)

Prof Mustafiz Sir

## Mean and variance of a discrete random variable :

The means or expected value of the discrete random variable X, denoted by $\mu$ or $E(x)$ is :

$$\mu = E(x) = \sum_x x f(x) .$$

The variance of X, denoted as $\sigma^2$ or $V(x)$ is $\sigma^2 = V(x) = E(x-\mu)^2$

$$= \sum_x (x-\mu)^2 f(x) = \left(\sum_x x^2 f(x)\right) - \mu^2$$

* Probabilities justified কিনা তা experiment এর সাথে সম্পর্ক 1.

**Ex-1:** Digital channel there is a chance that a bit transmitted through a digital transmission channel is recieved in error. ~~Lex~~ let X equal the number of bit in error is the next four bits transmitted. The possible values of X are {0, 1, 2, 3 and 4} Suppose that the probabilities are $P(X=0) = 0.6516$, $P(X=1) = 0.2916$, $P(X=2) = 0.0486$, $P(X=3) = 0.0036$, $P(X=4) = 0.0001$. calculate mean and variance of X.

Capital X = set value
x → member

Soln: Mean $\mu = E(x) = \sum_x x P(X)$

~~= 0.X6~~

~~= 0X0.6516 + 1X0.2916 + 1X0.6516 +~~

$= 0X\,0.6516 + 1X0.2916 + 2X0.0486 + 0.0036X3 + 4X0.0001 = 0.3982$

$\approx 0.4$

Variance

| $x$ | $x-\mu$ <br> $x-0.4$ | $(x-\mu)^2$ <br> $(x-0.4)^2$ | $P(x)$ | $(x-0.4)^2 P(x)$ |
|---|---|---|---|---|
| 0 | -0.4 | 0.16 | 0.6561 | ~~0.104976~~ 0.105072 |
| 1 | 0.6 | 0.36 | 0.2916 | 0.104976 |
| 2 | 1.6 | 2.56 | 0.0486 | 0.124416 |
| 3 | 2.6 | 6.76 | 0.0036 | 0.024336 |
| 4 | 3.6 | 12.96 | 0.0001 | $1.296 \times 10^{-3}$ |

$$V(x) = \sigma^2 = \sum_x (x-\mu)^2 P(x)$$

$$= 0.360096$$

Ex-2: Marketing two new ~~prod~~ product designs are to be compared on the basis of revenue potential. ~~Marken~~ Marketing believes that the revenue from design A can be predicted quite accurately to be \$3 million. The revenue potential of design B is more difficult to access. Marketing concludes that there is a probability of 0.3 that the revenue from design B will be \$7 million but there is a 0.7 probability that the revenue will be only \$2 million. Which design do you prefer?

Sol$^n$: Let X denote the revenue from design A.

$E(x) = \$3$ million.

Let Y denote the revenue from design B,

$E(Y) = 2 \times 0.3 + 2 \times 0.7 = \$3.5$

Ex-3: The no of e-mail messages recieved per hour has the following distribution:

| x=Msg | 10 | 11 | 12 | 13 | 14 | 15 |
|---|---|---|---|---|---|---|
| f(x) | 0.08 | 0.15 | 0.3 | 0.2 | 0.2 | 0.07 |

Determine the mean and SD of the number of msg recieved per hour.

$\mu = E(x) = 10 \times 0.08 + 11 \times 0.15 + 12 \times 0.3 + 13 \times 0.2 + 14 \times 0.2 + 15 \times 0.07$

$= 12.5$

$V(x) = \sum_x x^2 f(x) - \mu^2$

$= 10^2 \times 0.08 + 11^2 \times 0.15 + 12^2 \times 0.3 + 13^2 \times 0.2 + 14^2 \times 0.2 + 15^2 \times 0.07 - 12.5^2$

$= 1.85$

$SD = \sqrt{V(x)} = 1.36$

* Total 100

| x = Passenger | 10 | 20 | 30 | 15 | 25 |
|---|---|---|---|---|---|
| f(x) | 0.1 | 0.2 | 0.3 | 0.15 | 0.25 |

## Discrete probability distributions:

1. Binomial distribution
2. Poisson distribution

**1) Binomial distribution:** The binomial distribution must possess the following properties –

i) The expt consits of repeated trials

ii) Each trial results in an outcome that may be classified as a success or failure.

iii) The probability of success, denoted by P remains constant from trial to trial

iv) The repeated trials are independent.

The no of x of succeesses in n Bernoulli's trials is called a ~~bio~~ binomial random variable.

The probability distribution of this discrete random variable is called the ~~bio~~ binomial distribut

and its value will be denoted by $b(x;n,p)$ and define by

$$b(x;n,p) = P(x)$$

$$= {}^{n}C_{x}\, p^{x}(1-p)^{n-x}, \quad x = 0, 1, 2, 3, \ldots$$

$x$ = random variable

$n$ = trial no

$p$ = probability of success

~~$\mu = E(x) = x_i f(x)$~~

$$b(5; 15, 0.9) = {}^{15}C_{5}\, 0.9^{5}(1-0.9)^{15-5}$$

$$= 0.18$$

# 24.12.24 (Tuesday)

## Prof Mustafiz Sir

The probability distribution of X is

$b(x, n, p) = {}^{n}C_x p^x (1-p)^{n-x}$ $x = 0, 1, 2, \ldots, n$

[Important]

Ex1: Organic ~~pollut~~ pollution each sample of water has a 10% chance of containing a particular organic pollutant. Assume that the samples are independent with regard to the presence of the pollutant. Determine

1) The probability of that in the next 18 samples, exactly 2 contain pollutant

2) The probability that at ~~the~~ least four samples contain the pollutant

3) The probality that $3 \le X \le 7$

Soln:

① Let X be the no of samples that contains the pollutant in the next 18 samples. Then X is a binomial random variable with $P = 0.1$ and $n = 18$

Therefore $P(X=2) = {}^{18}C_2 (0.1)^2 (1-0.1)^{18-2}$

$= 0.284$

② The requested probality that at least four samples contain the pollutant is $P(X \ge 4)$

$= 1 - P(X \le 4)$

$= 1 - \sum_{x=0}^{3} \binom{18}{x} (0.1)^x (0.9)^{18-x}$

$= 1 - [0.15 + 0.3 + 0.284 + 0.168]$

$= 0.098$

③ The requested probability that $3 \le X \le 7$ is

$P(3 \le X \le 7) = \sum_{x=3}^{6} \binom{18}{x} (0.1)^x \times (0.9)^{18-x}$

$= 0.168 + 0.07 + 0.0221 + 0.005$

$= 0.265$

Example – 5.10 : Let X represent the number (P-10)

i) $P(10) = {}^{10}C_{10} \left(\frac{1}{5}\right)^{10} \left(\frac{4}{5}\right)^{10-10} = \frac{1}{5^{10}} = 1.024 \times 10^{-7}$

ii) $1 - P(\text{all are defective}) = 1 - \frac{1}{5^{10}} = 0.999$

iii) $1 - P(0) = 0.893$

iv) 0.859

Exp: 5.11 :

a) $P(X \geq 10)$

$= \sum_{x=10}^{15} b(x, 15, 0.4)$

$= {}^{15}C_{10} (0.4)^{10} (0.6)^{5} + {}^{15}C_{11} (0.4)^{11} (0.6)^{15-11} + {}^{15}C_{12} (0.4)^{12} (0.6)^{15-12}$

$+ {}^{15}C_{13} (0.4)^{13} (0.6)^{15-13} +$

${}^{15}C_{14} (0.4)^{14} (0.6)^{15-14} +$

${}^{15}C_{15} (0.4)^{15} (0.6)^{15-15}$

$= 0.0338$

c) $P(X=5) = b(5, 15, 0.4)$

$= 0.1859$

Exam 5.12 : $P = 0.1$, $n = 15$

a) $1 - P(X<5)$

$= 1 - {}^{15}C_{0}($

Exp 5.14 : Important

$P(X \geq 3)$

$= \sum_{x=3}^{5} b(x, 5, 0.29)$

$= 0.1035$

Exp$-5.15 (V.V. Important)

$P = 0.2 \quad q = 0.8$

i) $P(X=1) = \binom{4}{1} (0.2)^1 (0.8)^3$

$= 0.4096$

Box = $2000 \times 0.4096$

$= 819$

ii) $P(X=2) = \binom{4}{2} (0.2)^2 (0.8)^2$

$= 0.1536$

Box = $2000 \times 0.1536$

$= 307$

iii) $\sum_{x=0}^{2} = 0.972$

Box = 999

Define binonmial distribution with practical interpretion.

State the areas of application of binomial distribution

Properties

~~Binomi~~

Binomial ⟶ 1 set ques
Theory & Math

# class ... syllabus

Exp-5.13:

(a) $b(3, 10, 0.3) = 0.2668$

31.12.24 (Tuesday)

Prof Mustafiz sir

## Random exp

### Sample space:

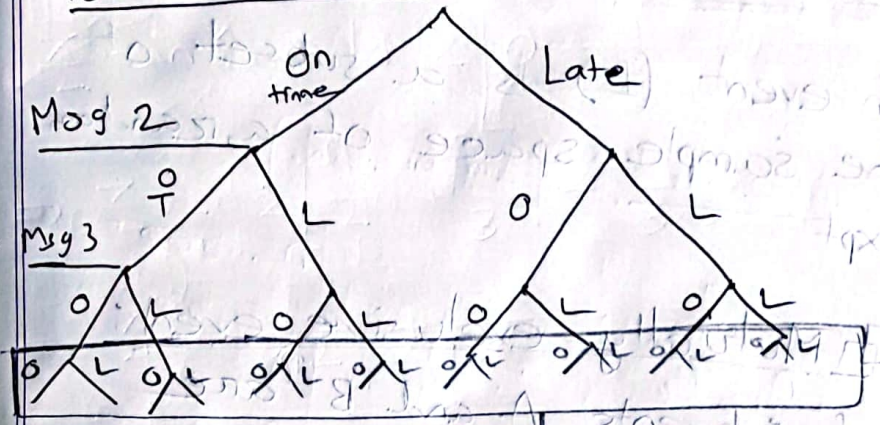
Example-1.1: Randomly select a a camera and record the recyle time of a flash

$S = R = \{x \mid x > 0\}$, the +ve real number

Example-1.2: [Important]

$S = \{000, 001, 010, 011, 100, 101, 110, 111\}$

Msg 1




→ For 4 message

Example-1.3:

$\therefore S = \{0000, 000L, 00LO, 00LL, 0L00, 0L0L, 0LL0, 0LLL, L000, L00L, L0L0, L0LL, LL00, LL0L, LLL0, LLLL\}$

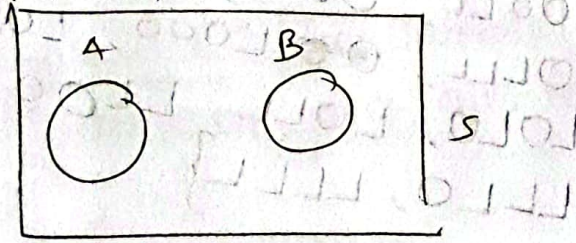
## Events are sets of outcomes

An event (E) is a subset of the sample space of a random expt

## Mutually exclusive event:

Defn: Events A and B are mutually exclusive if they share no common outcomes.

$A \cap B = \emptyset$




## Counting techniques:

${}^{n}P_{r}$ ${}^{n}C_{r}$ → Combination

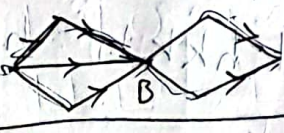
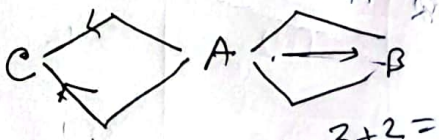
↓ Permutation

$${}^{n}P_{r} = \frac{n!}{(n-r)!}$$

$${}^{n}C_{r} = \frac{n!}{r!\,(n-r)!}$$

$${}^{15}C_{7} = \frac{15!}{7!\,(15-7)!} = 6435$$

$${}^{15}P_{2} = \frac{15!}{(15-2)!} = 32432400$$

| | |
|---|---|
| #   $3 \times 2 = 6$ | Dependent → Multiplication |
| #   $3+2=5$ | Independent → Addition |

Example-1.5:

$4 \times 3 \times 3 = 36$

Example-1.7: [Important]

$$\frac{n!}{n_1!\, n_2!} = \frac{5!}{3!\, 2!} = 10$$

$n = 3+2 = 5$

$n_1 = 3$

$n_2 = 2$

K, h
3, 2

kkkhh kkhhk khhkk hhkkk

~~kkhkh khkkh khhkk~~

kkhkh kkkhk khkkh hkkkh

hkhkk hkkhk

Probability: ~~Defn~~ It is the likelihood or chance that a p

* $^{n}P_n = \frac{n!}{(n-n)!} = \frac{n!}{0!} = n!$

$^{n}C_n = \frac{n!}{n!\,(n-n)!} = 1$

Example-1.9 [Important]

$A = \{a, b\}$ ~~$B = \{c, d\}$~~ $B = \{b, c, d\}$

$C = \{d\}$

$P(A) = 0.1 + 0.3 = 0.4$

$P(B) = 0.3 + 0.5 + 0.1 = 0.9$

$P(C) = 0.1$

$P(A') = 1 - 0.4 = 0.6$

$P(B') = 0.1$

$P(C') = 0.9$

$A \cap B = \{b\}$ $P(A \cap B) = 0.3$

$A \cup B = \{a, b, c, d\}$ $P(A \cup B) = 1$

$A \cap C = \{\text{null}\}$ $P(A \cap C) = 0$

## Axioms of probability: [Important]

**Example-1.10:** A semiconductor wafer is randomly selected from a batch that is classified by contamination and location as shown.

| Contamination | Location of Foil | | Total |
|---|---|---|---|
| | Center | Edge | |
| Low | 514 | 68 | 582 |
| High | 112 | 296 | 358 |
| Total | 626 | 314 | 940 |

## Probability of union.

## Conditional probability Rule:

Defn ~~of~~ ⟶ Important

$P(B/A) = P(A \cap B) / P(A)$ for $P(A) > 0$

**Example-1.11:** [Important]

Parts classified:

| Defective | Surface flaws | | Total |
|---|---|---|---|
| | Yes (F) | No (F) | |
| Yes (D) | 10 | 18 | 28 |
| No (D) | 30 | 342 | 372 |
| Total | 40 | 360 | 400 |

$P(F) = 40/400 = 0.1$

$P(D) = 0.07$

$P(D/F) = P(D \cap F)/P(F) = 0.25$

$P(D'/F) = P(D' \cap F)/P(F) = 0.75$

$P(D/F') = P(D \cap F')/P(F') = 0.05$

$P(D'/F') = P(D' \cap F')/P(F) = 0.95$

~~Random~~

Random sample [Important]

7.1.25 (Tuesday)

Prof Mustafiz Sir

## Total Probability Rule:

For any 2 events A & B,

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

$= P(B|A) \cdot P(A) + P(B|A') \cdot P(A')$

### Example-1.14

$P(F|H) = 0.1 \quad P(H) = 0.2$

$P(F \cap H) = 0.02$

$P(F|H') = 0.005$

$P(H') = 0.8$

$P(F \cap H') = 0.004$

$P(F) = P(F \cap H) + P(F \cap H')$

$= 0.02 + 0.004 = 0.024$

* $P(A/B) = \frac{P(A \cap B)}{P(B)}$

$\Rightarrow P(A \cap B) = P(A/B) \cdot P(B)$

## Total Probality Rule (Multiple events):

[Important]

$P(B) = P(B \cap E_1) + P(B \cap E_2) + - - - + P(B \cap E_k)$

$= P(B/E_1) \cdot P(E_1) + P(B/E_2) P(E_2) + - - - + P(B/E_k) P(E_k)$

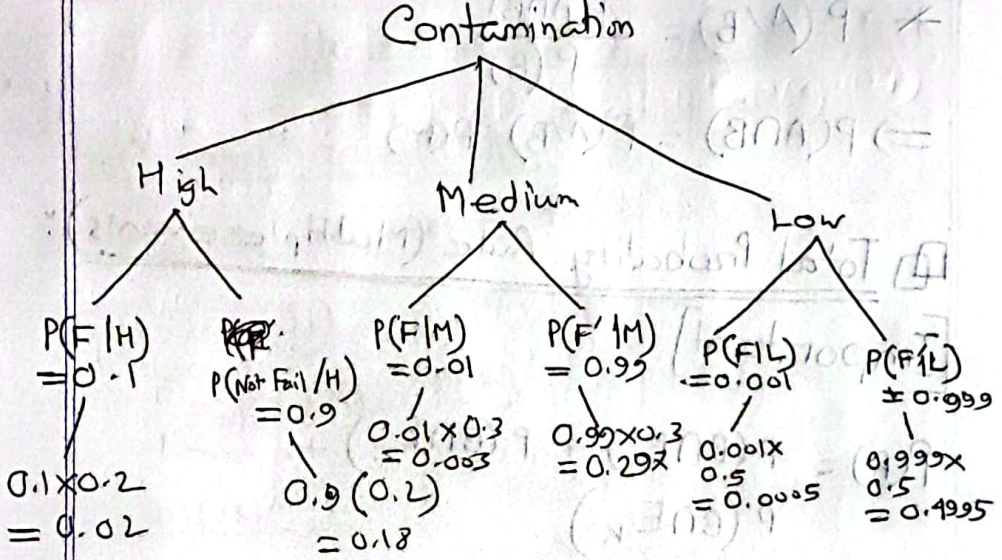
### Example: 1.15 [Important]

$P(F) = P(F \cap H) + P(F \cap M) + P(F \cap L)$

$= P(F|H) P(H) + P(F|M) P(M) + P(F|L) P(L)$

$= 0.1 \times 0.2 + 0.01 \times 0.3 + 0.001 \times 0.5$

$= 0.0235$




$P(F) = 0.02 + 0.003 + 0.0005 = 0.0235$

Event independence.

Lec-1 → Iset ques

**Exp-20: [Important]**

a)

$P[A \cap B \cap (C \cup D)] = P(A)\,P(B)\,P(C \cup D)$

$= P(A)\,P(B)\,[1 - P(C' \cap D')]$

$= P(A)\,P(B)\,[1 - P(C')\,P(D')]$

$= 0.9 \times 0.9\,(1 - (1-0.8)(1-0.8)] = 0.7776$

b)

$P = P(\text{the system works but } C \text{ does not}) \,/$
$P(\text{the system works})$

$= P(A \cap B \cap C^c \cap D) \,/\, P(\text{the system works})$

$= [0.9 \times 0.9 \times (1-0.8)(0.8)] / 0.7776$

$= 0.1667$

* C doesn't work

$P(A \cap B \cap C \cap D') / P(\text{the system works})$

$= 0.166\%$

Exp

**Exp-1.23: [Important]**

a) $\frac{8C_3}{20C_3} = 0.049$

b) $\frac{3C_3}{20C_3} = 0.0008\%$

c) $\frac{8C_2\ 3C_1}{20C_3} = 0.07\%$

d) 1 is white & 2 are others.

$\frac{3C_1\ 17C_2}{20C_3} = 0.35$

2 are white, 1 is other

$\frac{3C_2\ 17C_1}{20C_3} = 0.099$

a All are white,

$\frac{3C_3}{20C_3} = 0.0008\%$

$\therefore 0.35 + 0.049 + 0.0008\% = 0.399\%$

**Exp-1.24:**

$P(H) = 0.4 \quad P(W) = 0.5 \quad P(H/W) = 0.7$

1) $P(\text{Married Couple}) = P(W \cap H) = P(W) \cdot P(H/W)$

$= 0.5 \times 0.7 = 0.35$

2) $P(W/H) = \frac{P(W \cap H)}{P(H)} = \frac{0.35}{0.4} = 0.875$

3) $P(\text{At least one}) = P(H \cup W) = P(W) + P(H) - P(W \cap H)$

$= 0.5 + 0.4 - 0.35 = 0.55$

## Exp-1.25: [Important]

i) P(both will be selected) = $P(W \cap H)$

$= \frac{1}{6} \frac{2}{5} = \frac{1}{15} = 0.067$

ii) P(Only one of them) =

$P(H \cap W') + P(H' \cap W)$

$= \frac{1}{6}\left(1 - \frac{2}{5}\right) + \left(1 - \frac{1}{6}\right)\frac{2}{5}$

$= \frac{1}{10} + \frac{1}{3} = \frac{13}{30} = 0.43$

iii) None of them selected

$P(H' \cap W') = \left(1 - P(H)\right) \times \left(1 - P(W)\right)$

$= \left(1 - \frac{1}{6}\right)\left(1 - \frac{2}{5}\right)$

$= \frac{1}{2} = 0.5$

iv) At least one selected

$1 - P(H' \cap W')$

$= 1 - \frac{1}{2} = \frac{1}{2}$

$P(E \cup F) = P(E) + P(F) - P(E \cap F)$

$= \frac{1}{6} + \frac{2}{5} - \frac{1}{15}$

$= \frac{1}{2} = 0.5$

## Baye's Theorem: [Important]

$P(A|B) = \frac{P(B|A) \cdot P(A)}{P(B)}$ for $P(B) > 0$

## Exp-1.28:

| | | | | |
|---|---|---|---|---|
| 85 | 105 | 100 | 100 | 390 |
| | | | | 210 |
| | | | | 160 |
| | | | | 20 |

1) $P(E_1) = \frac{210}{390} = 0.5385$

2) $P(E_2) = \frac{60}{390} = 0.154$

3) $P(E_3) = \frac{2/13}{X/13} = 0.286$

4) $P(E_4) = \frac{5/39}{X/13} = 0.238$

Exp-1.29:

Lec-1 → সব Note Assignment

# 14.01.25 (Tuesday)

## Prof Mustafiz Sir

### Poisson distribution

1) Properties.

2) Example → [IMPORTANT]

3) $\mu = \lambda t$

4) $P(x, \lambda t) = \frac{e^{-\lambda t}(\lambda t)^x}{x!}$, $x = 0, 1, 2 \ldots$

where $\lambda$ is the avg number of outcomes per unit time, distance, area, or volume and $e = 2.71828$

~~Me The poisson di~~

$x$ = poisson random variable

Mean = $\lambda = p$

Standard deviation = $\sqrt{\lambda}$

Skewness, $\beta_1 = \frac{1}{\lambda}$

Kurtosis, $\beta_2 = 3 + \frac{1}{\lambda}$

Variance = $\lambda$

**Exp-5.17:**

$$P(6,4) = \frac{e^{-4}4^6}{6!}$$

$$= 0.1042$$

**Exp-5.18: [Important]**

~~$\lambda = 10$~~ $\lambda t = 10$

$$P(x>15) = 1 - P(x \le 15)$$

$$= 1 - \sum_{x=0}^{15} P(x;10)$$

$$= 1 - \{[P(0;10) + P(1;10) + P(2;10) + P(3;10) + P(4;10) + P(5;10) + P(6;10) + \text{——} + P(15;10)$$

$$= 1 - \frac{e^{-10}(10)^0}{0!} + \frac{e^{-10}(10)^1}{1!} + \frac{e^{-10}(10)^2}{2!} + \frac{e^{-10}(10)^3}{3!} + \text{——} + \frac{e^{-10}(10)^{15}}{15!}$$

$= 1 - 0.9513 = 0.0487$ ~~$= 0.048$~~

**Exp-5.19: [Important]**

$\lambda t = 6$

| 3 → 2 |
|---|
| 9 → $\frac{2\times 9}{3} = 6$ |

$$f(x) = \frac{e^{-6}6^x}{x!} \quad x = 0, 1, 2, \ldots$$

$$\lambda = \frac{2\times 9}{3} = 6$$

$$f(x) = \frac{e^{-6}6^x}{x!}$$

$$P(x \ge 5) = 1 - P(x \le 4)$$

$$= 1 - \sum_{x=0}^{4} \frac{e^{-6}6^x}{x!}$$

$$= 1 - \left(\frac{e^{-6}6^0}{0!} + \frac{e^{-6}6^1}{1!} + \frac{e^{-6}6^2}{2!} + \frac{e^{-6}6^3}{3!} + \frac{e^{-6}6^4}{4!}\right)$$

$= 1 - 0.285 = 0.715$

Math: Telephone calls enter a switchboard on the avg of 3 every 2 min interval. What is the probability of 6 or more call arriving in a 8 min?

Ans: Let X denote the number of calls in a 8-min period. We see the at the avg number of calls in a 8 min interval is $\lambda = \frac{3 \times 8}{2} = 12$

$$f(x) = \frac{e^{-12} 12^x}{x!}$$

$$P(x \geq 6) = 1 - P(x \leq 5)$$

$$= 1 - \sum_{x=0}^{5} \frac{e^{-12} (12)^x}{x!}$$

$$= 1 - 0.02 = 0.98$$

## Continuous Random variable:

### Probability density funcn: [Important]

For a continuous random variable X, a probability density funcn is a funcn such that

1) $f(x) \geq 0$

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

3) $P(a \leq x \leq b) = \int_a^b f(x)\, dx$ = area under $f(x)$ from a to b for any a and b

### Cumulative distribution funcn: [Important]

$$F(x) = P(X \leq x) = \int_{-\infty}^{x} f(u)\, du \quad \text{for } -\infty < x < \infty$$

## Mean and variance of a continuous random variable:

[Important]

$$\mu = E(x) = \int_{-\infty}^{\infty} x f(x)\,dx$$

$$\sigma^2 = V(x) = \int_{-\infty}^{\infty} (x-\mu)^2 f(x)\,dx$$

$$= \int_{-\infty}^{\infty} x^2 f(x)\,dx - \mu^2$$

## Normal distribution

$$n(x;\mu,\sigma) = \frac{1}{\sigma\sqrt{2\pi}} e^{-\frac{(x-\mu)^2}{2\sigma^2}}, \quad -\infty < x < \infty$$

$\mu$ → mean, $\sigma$ → deviation

## Properties:

1) The mode, which is the point on the horizontal axis where the curve is a maximum occurs at $x = \mu$

2) The curve is symmetric about a vertical axis through the mean $\mu$

3) The curve has its points of inflection at $x = \mu \pm \sigma$; it is concave downward if $\mu - \sigma < x < \mu + \sigma$ and upward otherwise.

4) The normal curve approaches the horizontal axis asymptotically

asymptotically as we proceed in either direction away from the mean.

5) The total area under the curve and above the horizontal axis is 1

$P(x_1 < X < x_2) = \int_{x_1}^{x_2} n($

$P(x_1 < X < x_2) = \int_{x_1}^{x_2} n($

(Lec-3 Page-9)

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

There are many types of statistical software that can be used in calculating areas under the normal curve. The difficulty encountered in solving integra

# 4.2.25 (Tuesday)

## Prof. Mustafiz Sir

~~Fig-3.9~~ Lec-3

Fig-3.9 [P-9] Graph

$P(x_1 < x < x_2) = \int_{x_1}^{x_2}$

P-10

Exp-3.8 :- ~~PC~~ P-10 Graph

$$z = \frac{2.3 - 3}{0.5} = -1.4$$

$P(x < 2.3) = P(Z < -1.4) = 0.0808$

## Exp-3.9; (P-11) Graph

$$Z_1 = \frac{778-800}{40} = -0.55$$

$$Z_2 = \frac{839-800}{40} = 0.85$$

$P(778 < x < 834)$

## Exp-3.10 (P-11) Graph

$$Z_1 = \frac{2.99-3}{0.005} = -2$$

$$Z_2 = \frac{3.01-3}{0.005} = 2.0$$

$P(2.99 < x < 3.01) = P(-2.0 < z < 2)$

$P(z < -2$

## Exp-3.11 (P-12) Graph

## Exp-3.12 (P-13) [Important] Graph

$$Z_1 = \frac{2.5-2.1}{0.6} = 0.67$$

$$Z_2 = \frac{3.5-2.1}{0.6} = 2.33$$

$P(2.5 < x < 3.5) = P(0.67 < z < 2.33)$

$= P(z < 2.33) - P(z < 0.67)$

$= 0.9901 - 0.7748 = 0.2921$

Hence 24.21% or approximately $0.2421 \times 80 = 20$ students of the 80 are expected to make a score between 2.5 and 3.5

এই পর্যন্ত Syllabus

Exp-313 (P-13)

$$Z_1 = \frac{4.75 - 5.25}{0.6} = -0.83$$

$$Z_2 = \frac{5.69 - 5.25}{0.6} = 0.73$$

$$P(4.75 < x < 5.69) =$$

$$P(-0.83 < Z < 0.73)$$

$$= P(Z < 0.73) - P(Z < -0.83)$$

$$= P(Z < 0.73) - [1 - P(Z < 0.83)]$$

$$= 0.7673 - (1 - 0.7967)$$

$$= 0.564 = 56.4\%$$

Prof Mustafiz Sir [Instead Lec Jakaria sir]

**Regression**: The term 'regression' literally means stepping back towards the average.

Regression analysis is a mathematical measure of the average relationship between 2 or more variables in terms of the original units of the data.

If the variable in a bivariate distribution are related, then the points in the scatter diagram (which is a diagrammatic representation of bivariate data) will cluster round some curve called of regression.

If the curve is a straight line, it is called the line of regression and said to be linear regression between the variables, otherwise regression is said to be curvelinear.

## Eqn of lines of Regression

$$y - \bar{y} = \rho \frac{\sigma_y}{\sigma_x}(x - \bar{x})$$

$$b_{yx} = \rho \frac{\sigma_y}{\sigma_x}$$

$$x - \bar{x} = \rho \frac{\sigma_x}{\sigma_y}(y - \bar{y})$$

$$b_{xy} = \rho \frac{\sigma_x}{\sigma_y}$$

## Properties [Important] (5.5.3)

* ~~$b_{yx}$~~

$$b_{yx} = \rho \frac{\sigma_y}{\sigma_x} = \frac{n\Sigma xy - (\Sigma x)(\Sigma y)}{n\Sigma x^2 - (\Sigma x)^2}$$

$$= \frac{\Sigma (x - \bar{x})(y - \bar{y})}{\Sigma (x - \bar{x})^2}$$

$$b_{xy} = \rho \frac{\sigma_x}{\sigma_y} = \frac{n\Sigma xy - (\Sigma x)(\Sigma y)}{n\Sigma y^2 - (\Sigma y)^2}$$

$$= \frac{\Sigma (x - \bar{x})(y - \bar{y})}{\Sigma (y - \bar{y})^2}$$

## Angle [5.5.4]

## Exp-5.69 [P-2]

| $x$ | $y$ | $xy$ | $x^2$ | $y^2$ |
|---|---|---|---|---|
| 65 | 67 | 4355 | 4225 | 4489 |
| 66 | 68 | 4488 | 4356 | 4624 |
| 67 | 65 | 4355 | 4489 | 4225 |
| 67 | 68 | 4556 | 4489 | 4624 |
| 68 | 72 | 4896 | 4624 | 5184 |
| 69 | 72 | 4968 | 4761 | 5184 |
| 70 | 69 | 4830 | 4900 | 4761 |
| 72 | 71 | 5112 | 5184 | 5041 |
| $\Sigma=549$ | $\Sigma=552$ | $\Sigma=37560$ | $\Sigma=37028$ | $\Sigma=38132$ |

$$\bar{X} = \frac{549}{8} \approx 69 \qquad \bar{Y} = \frac{552}{8} = 69$$

$$\bar{X}\bar{Y} = 68 \times 69 \qquad \bar{X}\,\bar{Y} = 69 \times 69 = 4692$$

$$\sigma_x = \sqrt{\frac{1}{n}\Sigma x^2 - \bar{X}^2} = \sqrt{\frac{37028}{8} - 4624} = 2.1$$

$$b_{yx} = \frac{8 \times 37560 - 549 \times 552}{8 \times 37028 - (549)^2}$$

$$= 0.67$$

$$b_{xy} = \frac{8 \times 37560 - 549 \times 552}{8 \times 38132 - (552)^2}$$

$$= 0.595$$

$$r(x,y) = \frac{\frac{1}{n}\Sigma XY - \bar{X}\cdot\bar{Y}}{\sigma_x \sigma_y}$$

$$r(x,y) = \sqrt{b_{xy}\, b_{yx}}$$

$$= 0.629$$

# 11.2.25 (Tuesday)

## Prof Mustafiz Sir

Exp - 5.20 [IMPORTANT]

| $x$ | $y$ | $X = x - 58$ | $Y = y - 40$ | $XY$ | $X^2$ | $Y^2$ |
|---|---|---|---|---|---|---|
| 55 | 35 | −3 | −5 | 15 | 9 | 25 |
| 56 | 38 | −2 | −2 | 4 | 4 | 4 |
| 58 | 32 | 0 | −3 | 0 | 0 | 9 |
| 59 | 39 | 1 | −1 | −1 | 1 | 1 |
| 60 | 44 | 2 | 4 | 8 | 4 | 16 |
| 60 | 43 | 2 | 3 | 6 | 4 | 9 |
| 62 | 44 | 4 | 4 | 16 | 16 | 16 |
| | | $\Sigma X = 9$ | $\Sigma Y = 0$ | $\Sigma XY = 48$ | $\Sigma X^2 = 38$ | $\Sigma Y^2 = 80$ |

$$\bar{X} = \frac{\Sigma x}{n} = \frac{4}{7} = 0.5714$$

$$\bar{Y} = \frac{\Sigma y}{n} = \frac{0}{7} = 0$$

$$Cov(X,Y) = \frac{\Sigma xy}{n} = \bar{X}\bar{Y} = \frac{48}{7} - 0$$

$$= 6.857$$

$$\sigma_x = \sqrt{\frac{\Sigma x^2}{n} - \bar{X}^2} = \sqrt{\frac{38}{7} - (0.5714)^2}$$

$$= 2.2588$$

$$\sigma_Y = \sqrt{\frac{\Sigma Y^2}{n} - \bar{X}\bar{Y}} = \sqrt{\frac{80}{7} - 0}$$

$$= 3.38$$

$$r(X,Y) = \frac{cov(X,Y)}{\sigma_x \sigma_y} = 0.898$$

~~$y - \bar{y} = \rho \frac{\sigma_y}{\sigma_x}(x - \bar{x})$~~

$$Y - \bar{Y} = \rho \frac{\sigma_Y}{\sigma_x}(x - \bar{x})$$

$$b_{xy} = \rho \frac{\sigma_y}{\sigma_x}$$

$$X - \bar{X} = \rho \frac{\sigma_x}{\sigma_y}(Y - \bar{Y})$$

$$Y - 0 = \frac{3.38}{2.2588}(X + 0.5714)$$

$$\Rightarrow Y = 1.496(X - 0.5714)$$

$$X - 0.5714 = \frac{2.2588}{3.38}(Y - 0)$$

$$\Rightarrow X - 0.5714 = 0.66Y$$

$$(X,Y) = (0.5714, 0)$$

$$\tan\theta = \frac{1-\rho^2}{\rho} =$$

Iset Ques – Exp – 5.69, 5.20

$$Y - \bar{Y} = \rho \frac{\sigma_y}{\sigma_x}(x - \bar{x})$$

$$\Rightarrow Y - 0 = 0.898 \frac{3.38}{2.2588}(x - 0.5714)$$

$$\Rightarrow Y = 1.34 (x - 0.5714)$$

$$x - 0.5714 = 0.898 \frac{2.2588}{3.38}(Y - 0)$$

$$\Rightarrow x - 0.5714 = 0.06\, Y.$$

$$\therefore (x, Y) = (0.5714, 0)$$

* $U_i = \frac{x_i - A}{h}$

1) Find the regression line of y on x and x on y

2) Find the correlation co-efficient

3) Angle between the lines.

Exp – 5.72

| x | y | X = x − 70 | Y = y − 165 | XY | $X^2$ | $Y^2$ |
|---|---|---|---|---|---|---|
| 62 | 126 | −8 | −39 | 312 | 64 | 1521 |
| 64 | 125 | −6 | −40 | 240 | 36 | 1600 |
| 65 | 139 | −5 | −26 | 130 | 25 | 676 |
| 69 | 145 | −1 | −20 | 20 | 1 | 400 |
| 70 | 165 | 0 | 0 | 0 | 0 | 0 |
| 71 | 152 | 1 | −13 | −13 | 1 | 169 |
| 72 | 180 | 2 | 15 | 30 | 4 | 225 |
| 74 | 208 | 4 | 43 | 172 | 16 | 1849 |
| | | $\Sigma X = -13$ | $\Sigma Y = -80$ | $\Sigma XY = 891$ | $\Sigma X^2 = 142$ | $\Sigma Y^2 = 6440$ |

$$\bar{X} = \frac{\Sigma X}{n} = \frac{-13}{8} = -1.625$$

$$\bar{Y} = \frac{\Sigma Y}{n} = -\frac{80}{8} = -10$$

$$\sigma_X^2 = \frac{1}{n}\Sigma X^2 - \bar{X}^2 = \frac{197}{8} - \left(\frac{-13}{8}\right)^2$$

$$= 15.73$$

$$\sigma_Y^2 = \frac{\Sigma Y^2}{n} - \bar{Y}^2 = \frac{6440}{8} - (-10)^2$$

$$= 705$$

$$\rho_{XY} = \frac{\frac{1}{n}\Sigma XY - \bar{X}\bar{Y}}{\sigma_X \sigma_Y}$$

$$= \frac{\frac{891}{8} - \left(\frac{-13}{8}\right)(-10)}{3.96 \times 26.55} = 0.9$$

$$Y - \bar{Y} = \rho \frac{\sigma_Y}{\sigma_X}(X - \bar{X})$$

$$\Rightarrow Y + 10 = 0.9 \frac{26.55}{3.96}(X + 1.625)$$

$$\Rightarrow Y + 10 = 6.0$$

$$\bar{x} = \bar{X} + 70 = \frac{-13}{8} + 70 = 68.375$$

$$\bar{y} = \bar{Y} + 165 = -10 + 165 = 155$$

$$x - 68.375 = 0.9 \times \frac{3.96}{26.55}(y - 155)$$

$$\Rightarrow x = 0.134y + 47.5$$

$$y - 68.375 = 0.9 \frac{26.55}{3.96}(x - 68.375)$$

* Define regression and correlation with illustration

* Properties of regression and correlation co-efficient.

5.70, 5.72, 5.74, 5.75

Ques হল এই form

# Prof Mustafiz Sir

Lec-8 & 9 important

↓

Hypothesis testing

- Confidence interval
  - Hypothesis testing
  - Statistical hypothesis
  - Null hypothesis
  - Alternative 11
  - Type I error
  - 11 2 11
  - p-values
  - Critical-value approve
  - General procedure for hypothesis test

[Important p-9]

## Summary of tests on the mean

variance known

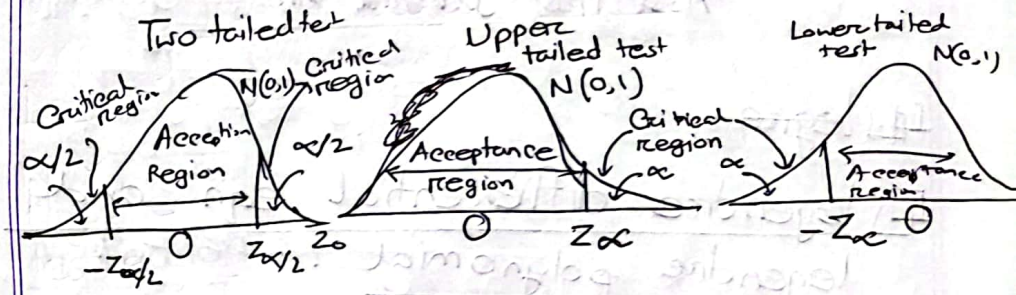
Ans: Null

Alternative

Test statistic formula

| Alternative hypotheses | Critical value |
|---|---|
| | |

Fig - 8.2 Drawing.




Example - 8.1

Reject $H_0$ → ~~on~~ Critical value ~~from~~

Example - 8.2 [IMPORTANT]

# 25.2.25 (Tuesday)

## Prof Mustafiz Sir

### Example – 8.3

A manufacturer of sports equipment has developed a new synthetic fishing line that the company claims has a mec

~~Ans: Parameter of interest~~

Ans:

Parameter of interest $\mu$

Null hypothesis : $H_0 : \mu = 20$ years

Alternative " : $H_1 : \mu > 20$

Test statistic : $z_0 = \frac{\bar{x} - \mu_0}{\sigma/\sqrt{n}}$

Reject $H_0$ : If the p-value is less that $\alpha = 0.05$

Computations:

$\bar{x} = 21.8$

$\sigma = 8.9$

$z_0 = \frac{\bar{x} - \mu_0}{\sigma/\sqrt{n}} = \frac{21.8 - 20}{8.9/\sqrt{100}} = 2.02$

Conclusion: The tabular value of the critical region : $z_{table} > 1.645$. Here the value of test statistic is $z_0 = 2.02 > 1.645$ we reject $H_0$ at the 0.05 level of significance and conclude that the mean

| Alternative hypotheses | Rejection Criterion for fixed level test |
|---|---|
| $H_1 : \mu \neq \mu_0$ | $t_0 > t_{\alpha/2, n-1}$ or $t_0 < -t_{\alpha/2, n-1}$ |
| $H_1 : \mu > \mu_0$ | $t_0 > t_{\alpha, n-1}$ |
| $H_1 : \mu < \mu_0$ | $t_0 < -t_{\alpha, n-1}$ |

**Example-8.4** [Important]

The edison electric institute has published figures on the number of ~~kw hour~~ kwh used annually by various home appliances. It is claimed that a vaccum cleaner uses an avg of 46kwh per year. If a random sample of 12 homes included in a planned study indicates that vaccum cleaners use an avg of 42 kwh per year with a std deviation of 11.9 kwh, does this suggest at the 0.05 level of significance that vaccum cleaners use, on avg, less than 46 kwh anually?

sample $= 12$

$\bar{x} = 42$

$s = 11.9$

$\alpha = 0.05$

Ques ⇒ std deviation
⇒ from assume

Parameter of interest: $\mu$

Null hypothesis: $H_0: \mu = 96$

Alternative " : $H_1: \mu < 96$

Test statistic: $t_0 = \frac{\bar{x} - \mu_0}{s/\sqrt{n}}$

Reject $H_0$: If the value of test statistic is in the critical region

Computation:

$$t_0 = \frac{92 - 96}{11.9/\sqrt{12}} = -1.16 \text{ with } 11^\circ$$

of freedom

Conclusion: The critical region at $\alpha = 0.05$ with $\nu = 11$ degrees of freedom is $t < -1.796$. Here the value of test statistic is $t = -1.16 > -1.796$. Hence don't reject $H_0$ and conclude that the avg number of kwh used annually by home vaccum cleaners isn't significantly less than 96.

* Table এ যদি value +ve হলে $t_0$ এর value টি -ve হবে তবে হবে।

$\nu = n - 1$ → Degree of freedom

* +ve হলে ছোট হলে → Not reject

-ve " বড় " → " "

Example – 8.5: An exp was performed in which 15 driverers produced by a particular club maker were selected at random and their co-efficient of refund measured. In the expt, the golf balls were fired from an air canon so that the incoming velocity and spin rate of the ball could be precisely

controlled. It is of interest to determine whether there is evidence (with $\alpha = 0.05$) to support ——

Parameter of interest : $\mu$

Null hypothesis

Standard deviation $s = \sqrt{\frac{\Sigma(x_i - \bar{x})^2}{N}}$

$\bar{x} = \frac{\Sigma x}{15} = 0.83725$

$= 0.02456$

$H_0 : \mu = 0.82$

$H_1 : \mu > 0.82$

$t_0 = \frac{\bar{x} - \mu_0}{s/\sqrt{n}} = \frac{0.83725 - 0.82}{0.02456/\sqrt{15}} = 2.72$

0.83725

| $x_i$ | $\bar{x} - x_i$ | $(\bar{x} - x_i)^2$ |
|---|---|---|
| 0.8411 | $3.85 \times 10^{-3}$ | $1.48225 \times 10^{-5}$ |
| 0.8580 | 0.02025 | $4.305625 \times 10^{-4}$ |
| 0.8042 | $-0.03305$ | $+1.092302 5 \times 10^{-3}$ |
| 0.8191 | $-0.01815$ | $3.29 \times 10^{-4}$ |
| 0.8532 | $-0.01595$ | $2.599 \times 10^{-4}$ |
| 0.823 | $0.035 \times 5$ | $1.22 \times 10^{-3}$ |
| 0.8182 | $-0.01905$ | $3.62 \times 10^{-4}$ |
| 0.8483 | $0.01105$ | $1.22 \times 10^{-4}$ |
| 0.8282 | $-9.05 \times 10^{-3}$ | $8.19 \times 10^{-5}$ |
| 0.8125 | $-0.02425$ | $5.88 \times 10^{-4}$ |
| 0.8276 | $-9.65 \times 10^{-3}$ | $6.12 \times 10^{-5}$ |
| 0.8359 | $-1.35 \times 10^{-3}$ | $9.31 \times 10^{-6}$ |
| 0.8750 | $0.03725$ | $1.8225 \times 10^{-3}$ |
| 0.7983 | $-0.03895$ | $1.92 \times 10^{-3}$ |
| 0.864 | $0.02875$ | $1.51 \times 10^{-4}$ |
| | | $8.26 \times 10^{-4}$ |

*Std deviation table কে দেখাতে হবে।

$\bar{x} = .83725$

$s = 0.02456$

① Random variable

② How can you generate a random variable from an experiment?

③ How can you distinguish discrete and continuous random variable with example?

④ Binomial distribution
Real life example:
Properties

⑤ Poission distribution
Real life exp
Properties

⑥ Normal distribution
Real life exp
Properties

⑦ Define and illustrate with example random exp, sample space, event, Assignment

## Prof Mustafiz Sir

The statistical method for testing the null hypothesis that the means of several population are equal is accomplished by F-test. The F-test is based on the technique of analysis of variance or abbriviated ANOVA. In the present context, we discuss about one way analysis of variance, in which only one type of treatment or factor is involved in the experiment.

Assume that we have k independent samples from k normal populations with population mean $\mu_1, \mu_2, \mu_3, \ldots, \mu_k$ respectively with common ~~variance~~ variance $\sigma^2$. An appropriate for testing the hypotheses is ~~$H_0 \neq$~~ $H_0 : \mu_1 = \mu_2 = \mu_3 = \ldots = \mu_k$

$H_1$ : At least ~~two~~ of the mean are not equal.

Assignment → Summary table 3 → With formula

## Model for One-Way Anova

| Sample 1 | Sample 2 | — | Sample k |
|---|---|---|---|
| $y_{11}$ | $y_{21}$ | | $y_{k1}$ |
| $y_{12}$ | $y_{22}$ | | $y_{k2}$ |
| — | — | — | — |
| — | — | — | — |
| $y_{1n_1}$ | $y_{2n_2}$ | | $y_{kn_k}$ |
| Total: $T_1$ | $T_2$ | | $T_k$ |
| Mean: $\bar{y}_1 = T_1/n_1$ | $\bar{y}_2 = T_2/n_2$ | | $\bar{y}_k = T_k/n_k$ |

Grand Total $= \sum_{j=1}^{k} \sum_{j=1}^{n_i} y_{ij}$

$= \sum_{i=1}^{k} T_i$

Grand mean $= \dfrac{\sum_{j=1}^{k} \sum_{j=1}^{n_i} y_{ij}}{N}$

$= \dfrac{G}{N}$

$N = n_1 + n_2 + n_3 + \cdots + n_k$

$SST = SSB + SSW$

$= \sum_{j=1}^{k} \sum_{j=1}^{n_i} (y_{ij} - \bar{y})^2 = \sum_{i=1}^{k} n_i (\bar{y}_i - \bar{y})^2 +$

$\sum_{j=1}^{k} \sum_{j=1}^{n_i} (y_{ij} - \bar{y}_i)^2$

Total variance = Variance between samples + Variance within samples.

| Variations | Formula | Degree of freedom |
|---|---|---|
| SSB | $\sum_{i=1}^{k} \frac{T_i^2}{n_i} - \frac{G^2}{N}$ | $k-1$ |
| SSW | $SST - SSB$ | $N-k$ |
| SST | $\sum_{i=1}^{k} \sum_{j=1}^{n_i} y_{ij}^2 - \frac{G^2}{N}$ | $N-1$ |

| Source of Variance | Sum of squares | Degree of freedom | Mean Squares | Computed F |
|---|---|---|---|---|
| Between samples | SSB | $k-1$ | $MSB = \frac{SSB}{k-1}$ | $F = \frac{MSB}{MSW}$ |
| Within samples | SSW | $N-k$ | $MSW = \frac{SSW}{N-k}$ | |
| Total | SST | $N-1$ | | |

Two-way analysis

11.3.25 (Tuesday)

Math-203

Prof Mustafiz Sir

Page - 9 math

| Brand | | | |
|---|---|---|---|
| A | B | C | D |
| 20 | 25 | 29 | 23 |
| 19 | 23 | 20 | 20 |
| 21 | 21 | 22 | 20 |

| $x_1$ | $x_1^2$ | $x_2$ | $x_2^2$ | $x_3$ | $x_3^2$ | $x_4$ | $x_4^2$ |
|---|---|---|---|---|---|---|---|
| 20 | 400 | 25 | 625 | 29 | 576 | 23 | 529 |
| 19 | 361 | 23 | 529 | 20 | 400 | 20 | 400 |
| 21 | 441 | 21 | 441 | 22 | 484 | 20 | 400 |
| $\Sigma x_1 = 60$ | $\Sigma x_1^2 = 1202$ | $\Sigma x_2 = 69$ | $\Sigma x_2^2 = 1595$ | $\Sigma x_3 = 66$ | $\Sigma x_3^2 = 1460$ | $\Sigma x_4 = 63$ | $\Sigma x_4^2 = 1329$ |

$G = \Sigma T_i = 60 + 69 + 66 + 63 = 258$

$\frac{G^2}{N} = \frac{258^2}{N} = 5542$

$$SST = \sum\sum x_{ij}^2 - \frac{G^2}{N}$$

$$= 1202 + 1595 + 1980 + 13229 - 5542$$

$$= 39$$

$$SSB = \sum \frac{T_i^2}{n_i} - \frac{G^2}{N}$$

$$= 15$$

| Source of variation | Sum of sqr | Deg of freedom | Mean Sqr | Computed F |
|---|---|---|---|---|
| Between samples | SSB = 15 | K − 1 = 3 − 1 = 2 | MSB = SSB / (k−1) = 15/3 = 5 | $F = \frac{MSB}{MSW}$ |
| Within | SSW = 39 − 15 = 24 | N − K = 12 − 3 = 8 | MSW = SSW/(N−K) = 24/8 = 3 | = 5/3 = 1.16 |
| Total | SST = 39 | N − 1 = 12 − 1 = 11 | | |

3 → 4.07 Table value

1.67 < 4.07

Null Hypothesis → acceptable