Discrete Distribution –

Probability distribution having its random variable discrete is called as 'discrete distribution'.

e.g.: PDF of X
where X: No. of heads in just one toss

X	0	1
$f(x)$	$1/2$	$1/2$

Bernoulli Distribution :

Let X be a random variable taking only 2 values '0' & '1' and corresponding probabilities as follows –

$$f(x_i) = P(X = x_i) = \begin{cases} p & \text{if } X = 1 \\ 1-p & \text{if } X = 0 \end{cases}$$

e.g.: X defined in above example is bernoulli random variable.

- Note I: Bernoulli distribution is used when there are only 2 possible outcomes. Outcome of Event A of interest is usually coded as $x=1$ with the probability 'p'

Notation – A : Event of interest ("Success")
& A^c : Event which is not of interest. ("failure")

Scanned with OKEN Scanner

Page No. | Date

e.g.: Tossing a coin 4 times & let,
X: no. of heads find PDF of X.

→

X	0	1	2	3	4
$P(X=x)$ = $f(x)$	$\frac{1}{16}$	$\frac{4}{16}$	$\frac{6}{16}$	$\frac{4}{16}$	$\frac{1}{16}$
	$\Downarrow$ $\frac{^4C_0}{2^4}$	$\Downarrow$ $\frac{^4C_1}{2^4}$	$\Downarrow$ $\frac{^4C_2}{2^4}$	$\Downarrow$ $\frac{^4C_3}{2^4}$	$\Downarrow$ $\frac{^4C_4}{2^4}$

Note: Max^m value is $\frac{6}{16}$
$\Rightarrow$ Mode of X is 2

Let A: Getting Head ("Success")

→ This not an example of Bernoulli distb^n
~~It~~ is of binomial distrib^n.

• Note:

$$P(X=0) = \frac{^4C_0}{2^4} = {}^4C_0 \left(\frac{1}{2}\right)^0 \left(\frac{1}{2}\right)^4$$

$$P(X=1) = \frac{^4C_1}{2^4} = {}^4C_1 \left(\frac{1}{2}\right)^1 \left(\frac{1}{2}\right)^3$$

$$P(X=2) = \frac{^4C_2}{2^4} = {}^4C_2 \left(\frac{1}{2}\right)^2 \left(\frac{1}{2}\right)^2$$

$$P(X=3) = \frac{^4C_3}{2^4} = {}^4C_3 \left(\frac{1}{2}\right)^3 \left(\frac{1}{2}\right)^1$$

$$P(X=4) = \frac{^4C_4}{2^4} = {}^4C_4 \left(\frac{1}{2}\right)^4 \left(\frac{1}{2}\right)^0$$

$$\Rightarrow P(X=x) = {}^4C_x \, (p)^x (q)^{4-x}$$

Scanned with OKEN Scanner

Page No.
Date

BINOMIAL PROBABILITY FORMULA

* n: no. of trials

then $$\boxed{P(X=x) = {}^{n}C_x \, p^x q^{n-x}}$$

where $q = 1-p$; p : Prob (success)
q : Prob (failure)

and $x = 0, 1, 2, \ldots, n$

BINOMIAL DISTRIBUTION –

- Note – I: Consider 'n' independent trials or repetitions of a bernoulli experiment.

- Note – II: In each trial, we may get Event 'A' or 'A^c' where 'A' is the event of interest and 'A^c' is the event which is not of interest and we count it as a failure. Event 'A' of interest is counted as a success.

- Note – III: For simplicity, we denote success by number '1' & failure is denoted by number '0'.

- Note – IV: Probability of A is denoted by 'p' & probability of A^c is $q = 1-p$.

Q.e) A company organizes a raffle at the end of year function. There are 300 Lottery tickets in total. Out of which 50 are marked as winning tickets. The event 'A' of interest is 'Ticket wins'. Find the probability 'p' of having a winning ticket.

Scanned with OKEN Scanner

* Binomial experiments are performed with replacement.

Page No.
Date

$\therefore p = \frac{50}{300} = \boxed{\frac{1}{6}}$

Que.) Find the probability that we are having 3 winning tickets out of 10 randomly chosen tickets.

$$\therefore p(X=3) = {}^{10}C_3 \left(\frac{1}{6}\right)^3 \left(\frac{5}{6}\right)^{10-3}$$

$$= {}^{10}C_3 \left(\frac{1}{6^3} \times \frac{5^7}{6^7}\right)$$

$$= \frac{10 \times 9 \times 8}{3 \times 2} \times \frac{5^7}{6^{10}}$$

$$= \frac{120 \times 5^7}{6^{10}}$$

$$= \boxed{0.155}$$ Ans

* The important thing here to note is the binomial experiments are performed with repetition allowed, so that the probability of success remains constant from trial to trial

Que.) What is the probability of having no winning ticket in 10 trials?

$$P(X=0) = {}^{10}C_0 \left(\frac{1}{6}\right)^0 \left(\frac{5}{6}\right)^{10-0}$$

$$= \frac{5^{10}}{6^{10}}$$

$$= \boxed{0.162}$$ Ans.

Scanned with OKEN Scanner

Page No. | Date

Que. 2) Consider an experiment where 3 items are selected at random from a manu company which are classified as defective D and non-defective (N). Total items manufactured are 100 out of which 25 are defective. Find the probability that we are having 2 defective items out of 3. ⇒ (Repition is allowed)

→

$$\therefore P(X=2) = {}^{3}C_{2}\left(\frac{1}{4}\right)^{2}\left(\frac{3}{4}\right)^{3-2}$$

$$= 3 \times \frac{1}{16} \times \frac{3}{4}$$

$$= \boxed{\frac{9}{64}}$$

Que.) Answer the same que where the replacement of the item back is not allowed.

→

$$\therefore \cancel{P = \frac{25}{100} \times \frac{24}{99} \times 23}$$

wrong

$$P = \frac{{}^{25}C_2 \times {}^{75}C_1}{{}^{100}C_3} = \frac{\frac{25 \times 24}{2} \times 75}{\frac{100 \times 99 \times 98}{3 \times 2}}$$

$$= \frac{\cancel{25} \times 24 \times 75 \times \cancel{3} \times \cancel{2}}{\cancel{100}_{4} \times \cancel{99}_{33} \times 98 \times \cancel{2}}$$

$$= \boxed{0.14}$$

Scanned with OKEN Scanner

Page No.
Date

Que) Find the expected value & the variance of Bernoulli random variable.

→

X	0	1
$f(x)$	~~$\frac{1}{2}$~~ $1-p$	~~E~~ p

$\Rightarrow$ ~~$E(X) = 0 \times p + 1 \times (1-p)$~~

~~$= 1-p$~~

$$\Rightarrow E(X) = 0 \times (1-p) + 1 \times p$$
$$= \boxed{p}$$

$$V(X) = (0-p)^2(1-p) + (1-p)^2 p$$
$$= p^2 - p^3 + (1 + p^2 - 2p)\,p$$
$$= p^2 - \cancel{p^3} + p + \cancel{p^3} - 2p^2$$
$$= p - p^2$$
$$= \boxed{p(1-p)} = \boxed{pq}$$

- THEOREM — Let X be a binomial random variable. Then Expected value of X i.e. $E(X) = np$ and variance of X i.e. $Var(X) = npq$ where all the parameters n, p & q have their usual meaning.

→ Missed a class on Poisson's Ratio

Scanned with OKEN Scanner

Page No.
Date

Generalization of binomial distribution when $n \to \infty$ & $p \to 0$

- **Poisson Distribution** –

↳ formula →
$$P(X=x) = \frac{e^{-\lambda}\lambda^{x}}{x!}$$

where $x = 0, 1, 2, \ldots\ldots, \infty$

$$\boxed{E(X) = var(X) = \lambda}$$

where $\lambda \Rightarrow$ Avg. no. of successes per unit time per unit area or volume.

Tut
Que.) What is prob. of atleast 10 will survive.

$n = 15$

$p = 0.4$

$q = 1-p = 0.6$

$$P(X \geq 10) = P(X=10) + P(X=11) + P(X=12) + P(X=13) + P(X=14) + P(X=15)$$

$= 0.0338$ (Using formula of binomial distⁿ)

- Another way using statistical tables –

$$\because P(X \geq 10) = 1 - P(X<10)$$
$$= 1 - P(X \leq 9)$$
$$= 1 - \sum_{x=0}^{9} b(x; n, p)$$
$$= 1 - \sum_{x=0}^{9} b(x; 15, 0.4)$$
$$= 1 - 0.9662$$
$$= \boxed{0.0338}$$ Ans.

Page No. Date

• What is the probability that 3 to 8 will survive?

$\longrightarrow$

$$P(3 \le X \le 8) = P(X \le 8) - P(X < 3)$$
$$= P(X \le 8) - P(X \le 2)$$
$$= \sum_{x=0}^{8} b(x; 15, 0.4) - \sum_{x=0}^{2} b(x; 15, 0.4)$$
$$= 0.9050 - 0.0271$$
$$= \boxed{0.8779} \text{ Ans.}$$

• Find probability that exactly 5 survives.

$\longrightarrow$

$$P(X=5) = {}^{15}C_5 (0.4)^5 (0.6)^{15-5} = \cancel{0.3099}$$
$$= \boxed{0.1859}$$

OR

$$= P(X \le 5) - P(X \le 4)$$

R Command

$\boxed{\text{dbinom}(x, n, p)}$

dbinom(5, 15, 0.4)

$$= \sum_{x=0}^{5} b(x; 15, 0.4) - \sum_{x=0}^{4} b(x; 15, 0.4)$$
$$= 0.4032 - 0.2173$$
$$= \boxed{0.1859}$$

* R commands -

1) To find probability at single point ⇒ { dbinom(x, n, p) }

2) $P(X \le x)$ ⇒ { pbinom(x, n, p) }

3) $P(a \le X \le b)$ ⇒ pbinom(b, n, p) - pbinom(a, n, p)

Scanned with OKEN Scanner

for $P(X \leq x) \rightarrow \{1 - ppois(x, \lambda)\}$

Que.) The average number of oil tankers arriving each day in a certain city is 10. The facilities at a port can handle atmost 15 tankers per day. What is the probability that on the given day, tankers will have to sent away.

→

Prob(that tankers have to be sent away)

$= P(X > 15)$

$= 1 - ppois(15, 10)$

$= 1 - \sum_{x=0}^{15} p(x, \lambda)$

$\lambda = \mu = 10$

$$\begin{array}{r} 1.0000 \\ -\,0.9513 \\ \hline 0.0487 \end{array}$$

$= 1 - 0.9513$

$= \boxed{0.0487}$

Que.) A ~~basketball~~ cricket player's batting average is 0.25. What is the probability that he gets exactly one hit in his next 4 times at a bat?

→ $\Rightarrow \lambda = 0.25 = E(X) = np \Rightarrow p = \frac{0.25}{4}$

$n = 4$

~~Success~~

$\therefore \boxed{p = 0.0625}$

Scanned with OKEN Scanner

Normal Distribution : (Guassian / Continuous Distribution) -

A r.v. X is said to follow a normal distrib^n with parameters μ & σ^2 if its PDF is given by -

$$f(x) = \frac{1}{\sigma\sqrt{2\pi}} e^{-\frac{(x-\mu)^2}{2\sigma^2}} \quad ; \quad x \in R, \ \mu \in R, \ \sigma^2 > 0$$

→ Property -

1) Graph of this $f(x)$ is symmetric to line $x = \mu$.

2) Max. value of $f(x)$ is attained at $x = \mu$.

Proof ⇒ $f' = 0 \Rightarrow \frac{1}{\sigma\sqrt{2\pi}} e^{-\frac{(x-\mu)^2}{2\sigma^2}} \times \frac{(-2)\cdot(x-\mu)}{2\sigma^2} = 0$

$\Rightarrow x - \mu = 0 \Rightarrow x = \mu$ is a critical point.

3) Now, $f'' = 0 \Rightarrow \frac{d}{dx}\left(\frac{-1}{\sigma^2\sqrt{2\pi}} e^{-\frac{(x-\mu)^2}{2\sigma^2}} \cdot (x-\mu)\right) = 0$

$$\Rightarrow \frac{-1}{\sigma^2\sqrt{2\pi}}\left[e^{-\frac{(x-\mu)^2}{2\sigma^2}} \times 1 + (x-\mu)e^{-\frac{(x-\mu)^2}{2\sigma^2}} \cdot \frac{(-2)(x-\mu)}{2\sigma^2}\right] = 0$$

$$\Rightarrow e^{-\frac{(x-\mu)^2}{2\sigma^2}} = \frac{(x-\mu)^2}{\sigma^2} e^{-\frac{(x-\mu)^2}{2\sigma^2}}$$

$\Rightarrow (x-\mu)^2 = \sigma^2$

~~$\Rightarrow \boxed{x = \mu + \sigma^2}$~~

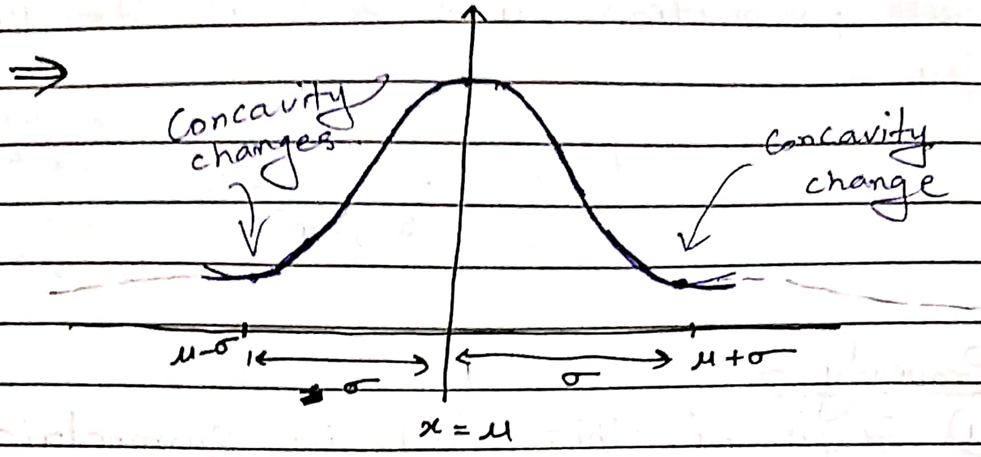
$\Rightarrow \boxed{x = \mu + \sigma}$ and $\boxed{x = +\mu - \sigma}$

Scanned with OKEN Scanner

Date

$\therefore$ There are two inflexion points $x = \mu + \sigma$ & $x - \sigma$.

$\Rightarrow$



4) $\lim_{x \to \infty} f(x) = 0$ and $\lim_{x \to -\infty} f(x) = 0$

$\therefore$ Graph of f tends to zero as $x \to \pm\infty$.

- **Note:**

1) Expected value of $X = \mu$ and the variance of X is σ^2.

$$\Rightarrow E(X) = \mu \quad \& \quad Var(X) = \sigma^2.$$

where $X \to$ normal r.v.

Proof: $\because$ $E(X) = \int_{-\infty}^{\infty} x f(x)\, dx$

$$= \int_{-\infty}^{\infty} x \left(\frac{1}{\sqrt{2\pi}\,\sigma} e^{-\frac{(x-\mu)^2}{2\sigma^2}} \right) dx$$

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

use by-parts

On evaluation,

$$\boxed{E(X) = \mu}$$

Scanned with OKEN Scanner

Page No.
Date

2) If $\mu = 0$ & $\sigma^2 = 1$, we call X as 'Standard Normal Random Variable'.
In particular, when $\mu = 0$ & $\sigma = 1$, its PDF is given by :

$$\left\{ \phi(x) = \frac{1}{\sqrt{2\pi}} e^{\frac{-x^2}{2}} \right\}$$ PDF of standard normal r.v.

We denote standard normal r.v. by 'Z'.

3) We write (1) $X \sim N(\mu, \sigma^2)$
$\Rightarrow$ X is equivalent to ——

(2) $Z \sim N(0, 1)$.

4) Cumulative probability distribution function CDF ~~and~~ of Z is denoted by -

$$\left\{ \Phi(x) = \int_{-\infty}^{x} \phi(t)\, dt \right\}$$

5) Lower value of σ indicates high concn around $x = \mu$. ($\because$ σ is Standard deviation).

6) Higher value of σ indicates flatter curve.

Scanned with OKEN Scanner

* **THEOREM** –

If X is a normal r.v. with mean μ and variance σ^2 then $Z = \frac{X-\mu}{\sigma}$ is a standard normal r.v.

• **Proof:**

i.e. To prove that $E(Z) = 0$ and $var(Z) = 1$

Now, $E(Z) = E\left(\frac{X-\mu}{\sigma}\right)$

$= \frac{1}{\sigma} E(X-\mu)$

$= \frac{1}{\sigma}[E(X) - E(\mu)]$

$= \frac{1}{\sigma}[E(X) - \mu]$

$= \frac{1}{\sigma}[\mu - \mu]$ --- [$\because$ given that, mean $= \mu$]

$\therefore \boxed{E(Z) = 0}$

Also, $var(Z) = var\left(\frac{X-\mu}{\sigma}\right)$

$= \frac{1}{\sigma^2}\left(var(X-\mu)\right)$

$= \frac{1}{\sigma^2}\left(var(X) - \cancel{var(\mu)}\right)$ --- [$\because var(\mu) = 0$]

$= \frac{1}{\sigma^2}(\sigma^2)$ --- [given]

$\therefore \boxed{var(Z) = 1}$

* Calculation Rules :

Let X : normal r.v. $\quad a, b \in \mathbb{R}$

$\leq$ i) $P(X \leq b) = P\left(\frac{X-\mu}{\sigma} \leq \frac{b-\mu}{\sigma}\right)$

$$= P\left(Z \leq \frac{b-\mu}{\sigma}\right).$$

e.g.: If x is a normal r.v. with mean $\mu = 100$ & S.D $\sigma = 5$. Then find i) $P(X < 50)$.

$\rightarrow$

$$= P(X < 50)$$

$$= P\left(\frac{x-100}{5} \leq \frac{50-100}{5}\right)$$

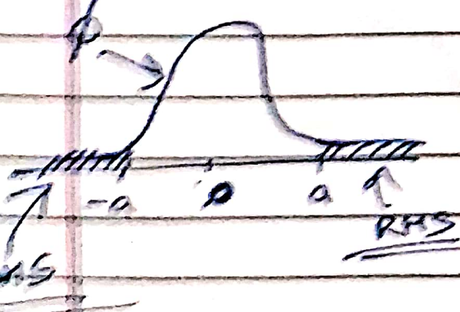
$$= P(Z \leq -10) = 0 \qquad \because -10 < -3.4$$

$\leq$ ii) $P(a \leq X \leq b) = P\left(\frac{a-\mu}{\sigma} \leq Z \leq \frac{b-\mu}{\sigma}\right)$

$$= P\left(Z \leq \frac{b-\mu}{\sigma}\right) - P\left(Z \leq \frac{a-\mu}{\sigma}\right)$$

$\leq$ iii) $\Phi(-a) = 1 - \Phi(a)$

Proof: LHS: $\Phi(-a) = \int_{-\infty}^{-a} \phi(t)\,dt$ = Area under the curve of ϕ from $-\infty$ to $-a$



RHS : $1 - \Phi(a) = 1 - \int_{-\infty}^{a} \phi(t)\,dt$

= Total Area under curve − Area under ϕ from $-\infty$ to a

= Area on right hand side of a.

$\therefore$ By symmetry of ϕ ; LHS = RHS

4) $\{P(-a < z < a) = 2\Phi(a) - 1\}$

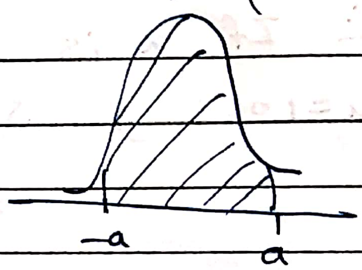
Proof:

LHS : $P(-a < z < a)$

$= P(z < a) - P(z < -a)$

$= \left(\text{Area to the left of } z = a\right) - \left(\text{Area to the left of } z = -a\right)$

=



$= \Phi(a) - \Phi(-a)$

$= \Phi(a) - (1 - \Phi(a))$

$= 2\Phi(a) - 1$

$= \underline{\text{RHS}}$

Que.) A scientist is giving rodents the diet ~~an~~ average of 40 months when their diets are sharpely restricted & enriched with vitamins & proteins. Assuming that the lifetime of such rodents are normally distributed with a S.D. of 6.3 months find the probability that the given rodent will survive for more than 32 months.

→ Given: $\mu = 40$

$\sigma = 6.3$

X : lifetime of rodents.

To find: $P(X > 32)$.

Scanned with OKEN Scanner

Page No.
Date

Sol^n: $\because P(x>32) = 1 - P(x \leq 32)$

$$= 1 - P\left(z \leq \frac{32-40}{6.3}\right)$$

$$= 1 - P\left(z \leq \frac{-8}{6.3}\right)$$

$$= 1 - P(z \leq -1.27)$$

$$= 1 - 0.1020$$

$$= \boxed{0.8980} \text{ Ans.}$$

(missed 1 lec on χ^2 statistics)

Scanned with OKEN Scanner

* RECALL :

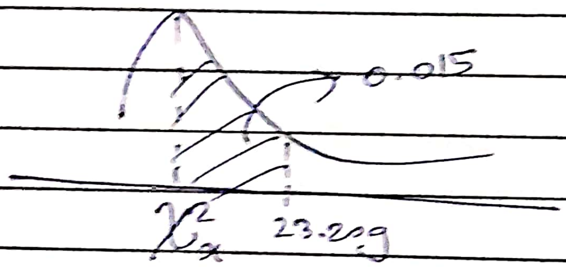
1) $\chi^2 = \sum_{i=1}^{n} Z_i^2$

2) We have shown,

$$\frac{(n-1)s^2}{\sigma^2} \sim \chi^2_{(n-1)} \leftarrow \text{dof.}$$

Que.) Using statistical tables, find χ^2_α if $P(\chi^2_\alpha < \chi^2 < 23.209) = 0.015$ for dof = 10.

$\rightarrow$



$$P(\chi^2_\alpha < \chi^2 < 23.209) = 0.015$$

$\Rightarrow$ (Total area to the right of χ^2_α) − (Total area to the right of 23.209) = 0.015

$$\Rightarrow P(\chi^2 > \chi^2_\alpha) - P(\chi^2 > 23.209) = 0.015$$

$$\therefore P(\chi^2 > \chi^2_\alpha) = 0.015 + \underline{0.01} \quad \leftarrow \text{from table}$$

$$= \boxed{0.025}$$

$$\Rightarrow \boxed{\chi^2_\alpha = 20.483} \leftarrow \text{from table.}$$

Ans.

population parameters $\mu \longleftrightarrow \bar{X}$, $\sigma \longleftrightarrow S$ Analogous sample parameters

Page No.
Date

T-DISTRIBUTION (Student's t-distribution):

- **t-statistics**: Let X & Y be two random variables such that, $X \sim N(0,1)$ and $Y \sim \chi^2_{n-1}$ dof then the ratio:

$$\frac{X}{Y} = \frac{N(0,1)}{\sqrt{\frac{\chi^2}{n-1}}}$$ is called as t-statistics.

- **Use of t-statistics**:

This statistics is used when σ is unknown.

Now consider the expression:

$$\frac{\bar{X}-\mu}{\left(\frac{\sigma}{\sqrt{n}}\right)} = \frac{\bar{X}-\mu}{\left(\frac{S}{\sqrt{n}}\right)} = \frac{\sigma(\bar{X}-\mu)}{\sigma\left(\frac{S}{\sqrt{n}}\right)}$$

$$= \frac{\frac{\bar{X}-\mu}{\sigma/\sqrt{n}}}{S/\sigma} = \frac{Z}{\sqrt{S^2/\sigma^2}}$$

$$= \frac{Z}{\sqrt{\frac{\chi^2}{(n-1)}}}$$

$\therefore$ We conclude that, $\left\{\frac{\bar{X}-\mu}{\left(S/\sqrt{n}\right)} = t_{n-1}\right\}$

being a ratio of normal & χ^2-r.v.

Scanned with OKEN Scanner

* Application of t-distribution -

Q.) A chemical engi— claims that the population mean of a certain batch process is 500 g/mL. To check this claim, he samples 25 batches & if the computed value of t-statistics lies in betⁿ $-t_{0.05}$ & $t_{0.05}$ then he is satisfied with his claim. What conclusion can be drawn from a sample whose mean is 518 g/mL (Note that: Sample S.D. is 40grams).

<u>Given:</u> $\mu = 500$ gm/mL

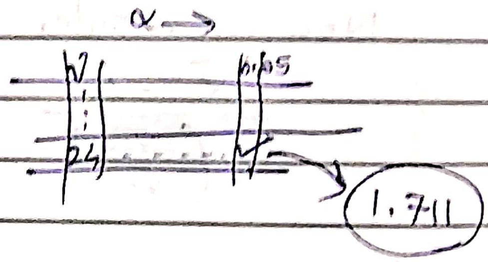
$n = 25$

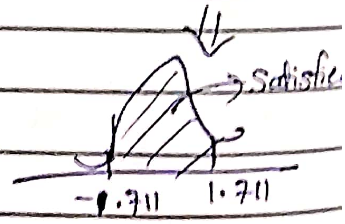
he is satisfied with the claim if $t \in (-t_{0.05}, \cancel{t_0}\ t_{0.05})$.

$S = 40$ gm

$\bar{X} = 518$ gm/mL

$\alpha = 0.05$





⇒ Note:

$$t = \frac{\bar{X} - \mu}{S/\sqrt{n}}$$

$$= \frac{518 - 500}{40/\sqrt{25}}$$

$$= \frac{18}{8} = \boxed{2.25}$$

f- STATISTICS :

Ratio of two χ^2 r.v.'s with m & n dofs, resp.

$$\therefore \quad f = \frac{\chi^2_m}{\chi^2_n}$$

- **Note :** f-statistics comes with two degrees of freedom.

* Reciprocal Rule :

$$\left\{ f_\alpha(\underbrace{v_1, v_2}_{\text{dofs}}) = \frac{1}{f_{1-\alpha}(v_2, v_1)} \right\}$$

Que.) Find $f_{0.95}(10, 6)$

$$\therefore f_{0.95}(10, 6) = \frac{1}{f_{0.05}(6, 10)} = \frac{1}{3.22}$$

$$= \sim \text{Ans.}$$

Scanned with OKEN Scanner