18th JUNE 2023

# Probability & Statistics

| | | | | | | |
|---|---|---|---|---|---|---|
| Q 1 | Five Transistors | Probability | Easy | 30 | 28 % |
| Q 2 | 3 Ants | Probability | Easy | 30 | 46 % |
| Q 3 | 2 Cabinets | Probability | Medium | 40 | 37 % |
| Q 4 | Macbook preference | Statistics | Medium | 40 | 65 % |
| Q 5 | Who is that Smoker | Probability | Hard | 50 | 63 % |
| Q 6 | Big Family | Probability | Medium | 40 | 46 % |
| Q 7 | Chemical Plant Impurity | Statistics | Hard | 50 | 41 % |
| Q 8 | Testing efficacy of improving ... | Fundamentals | Medium | 40 | 39 % |
| Q 9 | Poverty and Dietary Calcium | Statistics | Medium | 40 | 33 % |
| Q 10 | Significance Level | Statistics | Very Easy | 20 | 39 % |

Q 1 Five Transistors Probability Easy 30 28 %

A set of five transistors are to be tested, one at a time in a random order, to see which of them are defective.

Suppose that three of the five transistors are defective, and let N1 denote the number of tests made until the first defective is spotted,

and let N2 denote the number of additional tests until the second defective is spotted.

Find the probability P(N1 = 1, N2 = 2)

$T_1 \quad T_2 \quad T_3 \quad T_4 \quad T_5$

$3 \to D$

$2 \to ND$

$P(N_1 = 1, N_2 = 2) \to$ D, ND, D $\to$ Prob. of this sequence.

$2 \to D$

$1 \to ND$

$$= \left(\frac{\cancel{3}}{5} \times \frac{\cancel{2}}{\cancel{4}} \times \frac{2}{\cancel{3}}\right) = 1/5$$

D , ND , D

$N_1 = 1$

1 2

$N = 2$

2

5

3 def

2 not def

$\Rightarrow P(D) = \frac{3}{5}$ [3D, 2ND]

$\Rightarrow P(ND) = \frac{2}{4}$ [2D, 2ND]

$\Rightarrow P(D) = \frac{2}{3}$ [2D, 1ND]

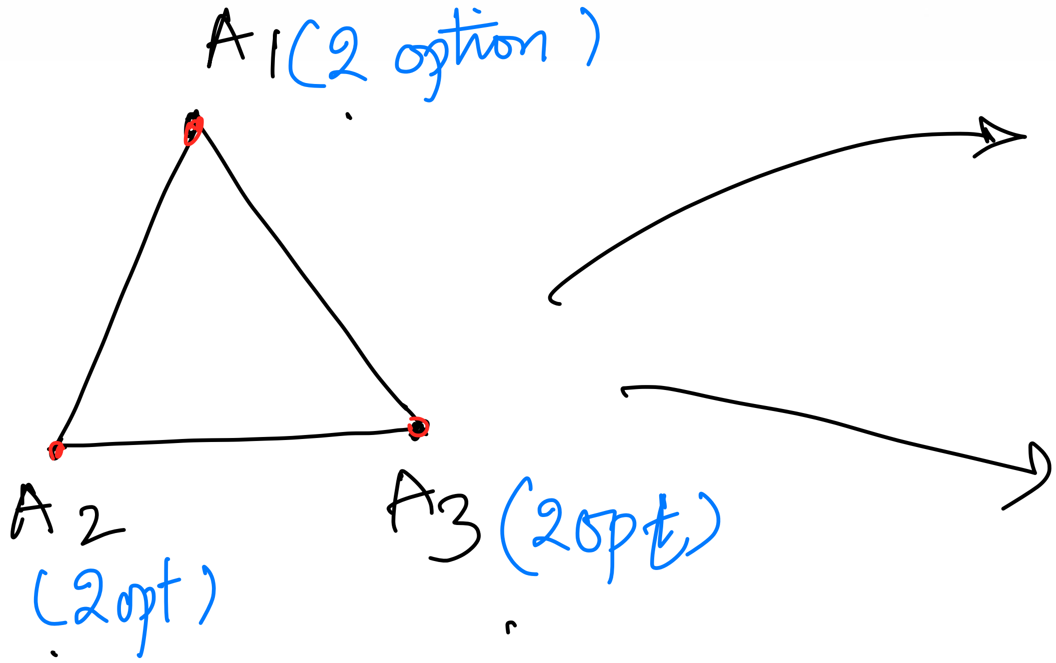
$$\frac{3}{5} \times \frac{2}{4} \times \frac{2}{3}$$

$$= 1/5$$

*

Three ants are sitting on the corners of a triangular path. An ant have equal probability to move along any of the two edges.

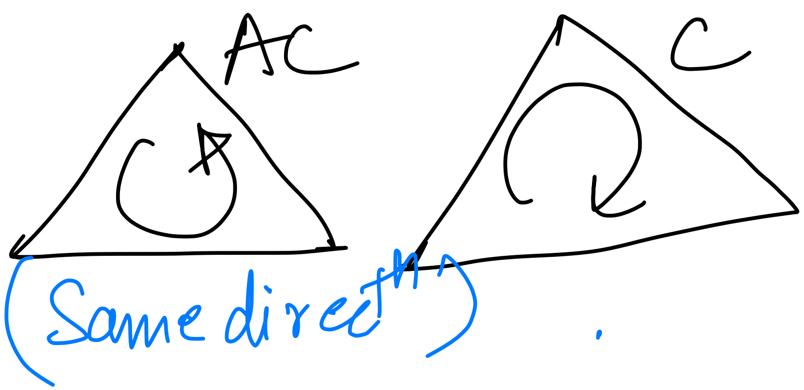
If all three ants start at the same time what is the probability that there will be no collision among them?




Tossing a coin 3 times (H/T)

| | |
|---|---|
| $\boxed{HHH}$ | $\boxed{TTT}$ |
| HHT | TTH |
| HTH | THT |
| THH | HTT |

No collision Cases.




$$= \frac{2}{8} = P(\text{no collision}) = \frac{1}{4}$$

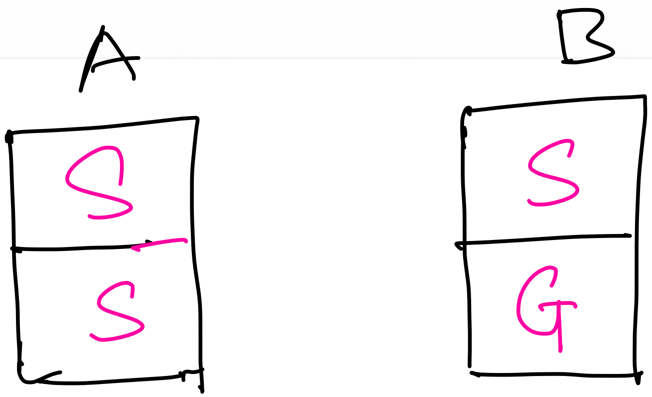
Q 3 2 Cabinets Probability Medium 40 37 %

Each of 2 cabinets identical in appearance has 2 drawers.

Cabinet A contains a silver coin in each drawer, and cabinet B contains a silver coin in one of its drawers and a gold coin in the other.

A cabinet is randomly selected, one of its drawers is opened, and a silver coin is found.

What is the probability that there is a silver coin in the other drawer?




$$P(A) = P(B) = 1/2$$

$$P(2^{nd} S \mid 1^{st} S) = ?$$

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

$$P(2^{nd} S \mid 1^{st} S) = \frac{P(2^{nd} S \cap 1^{st} S)}{P(1^{st} S)}$$

$$= \frac{P(A)}{P(1^{st} S)} = \frac{1/2}{P(1^{st} S)}$$

$$P(1^{st}S) = \left(\frac{1}{2} \times 1\right) + \left(\frac{1}{2} \times \frac{1}{2}\right)$$

$$= \frac{1}{2} + \frac{1}{4} = \frac{3}{4}$$

$$P(2^{nd}S \mid 1^{st}S) = \frac{1/2}{3/4} = \frac{4}{2 \times 3} = \frac{2}{3}$$

Q 4 Macbook preference Statistics Medium 40 65 %

A company has 10% software developers. 70% of software developers prefer Macbook. 40% of non-software developers prefer Macbook. What fraction of people prefer macbook?

Soft. Dev.

$$100 \rightarrow 10 \rightarrow \begin{cases} 7 & \text{(prefer MAC)} \\ 3 & \text{(don't prefer MAC)} \end{cases}$$

non. S.D

$$100 \rightarrow 90 \rightarrow \begin{cases} 36 & \text{(prefer MAC)} \\ 54 & \text{(don't prefer MAC)} \end{cases}$$

$\Rightarrow$ # of People Prefer MAC $= 7+36=43$

fraction $= 43/100 = 43\%$

0.43

Q 5 Who is that Smoker Probability Hard 50 63 %

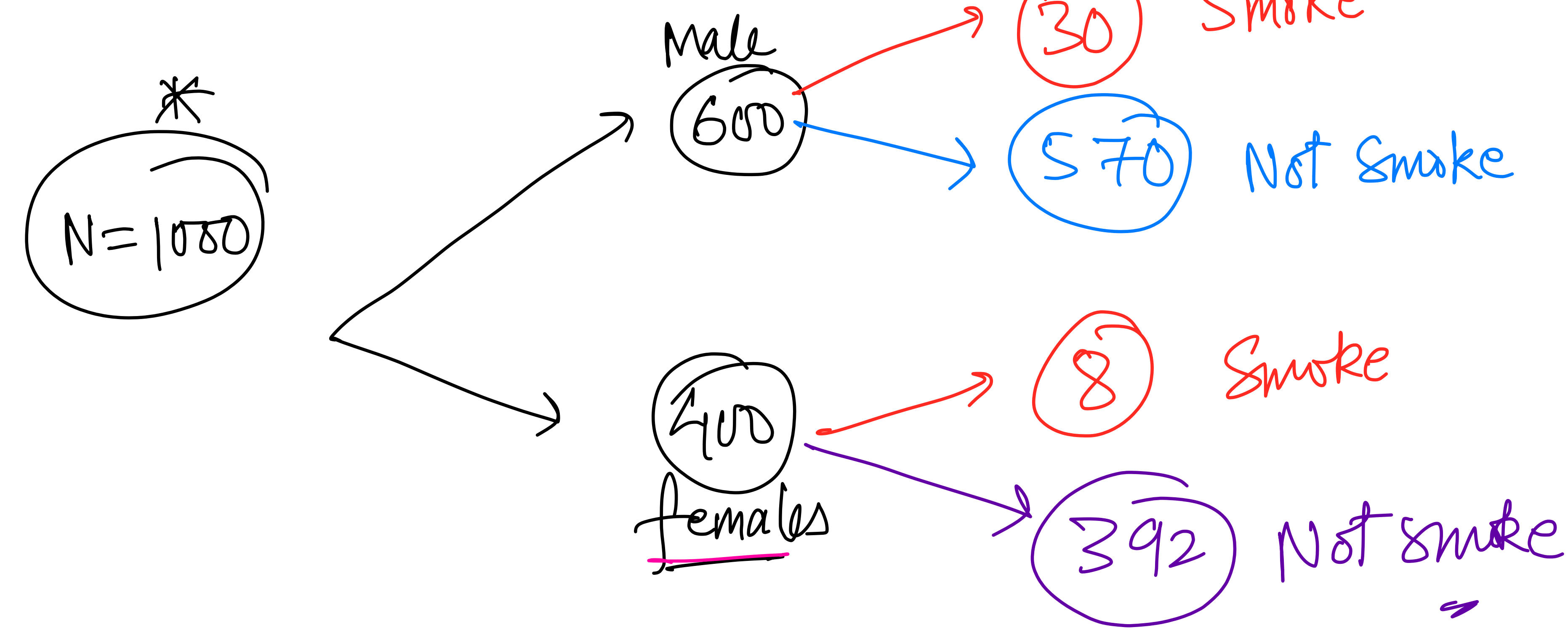
In a certain population:

Male are 60% and rest are Female.

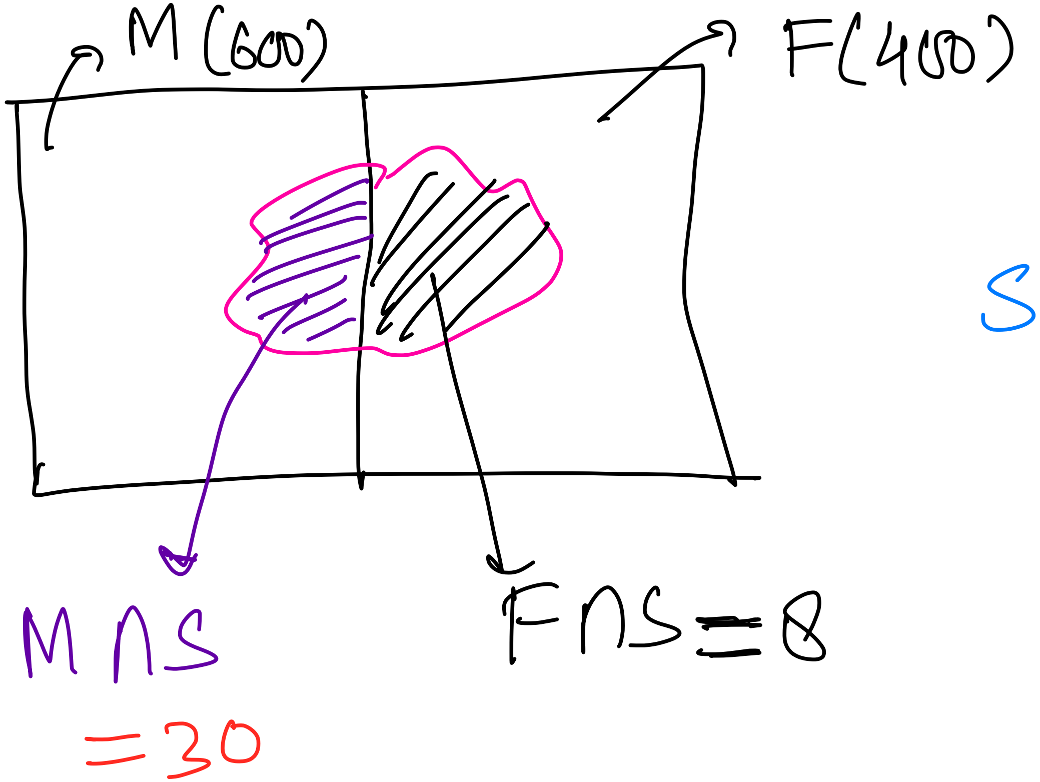
5% of Male and 2% Female Smokes.

a) What is the probability that a randomly selected person is a Smoker ?

b) A randomly selected person found to be a Smoker, what is the probability that this person is MALE ?




(a) $P(S) = \frac{30+8}{1000} = \frac{38}{1000} = 0.038 \rightarrow 3.8\%$




$$S = (M \cap S) + (F \cap S)$$
$$= 30 + 8$$
$$= \underline{38} \rightarrow \underline{\text{Total smoker}}$$

(b) $$P(M \mid S) = \left(\frac{30}{38}\right) \simeq 0.79$$
$$\simeq \underline{79\%}$$

Q 6 Big Family Probability Medium 40 46 %

A family has 8 children, both girl and boy are equally likely.

What is the probability that this family has less than 3 boys ?

$P(G) = 1/2$

$P(B) = 1/2$

P( less than ~~3~~ 3 boys).

$\text{Child} \to B$ . $\to P(B) = 0.5$

$\text{Child} \to G$ . $\to P(G) = 0.5$

$X \to$ # of boys

$0, 1, 2, 3, - - - - , 8 \to$ Random Variable

$P(X < 3) = ?$

$$\left.\begin{aligned} &P(X=0)\\ &P(X=1)\\ &P(X=2)\\ &\vdots\\ &P(X=8)\end{aligned}\right\} \text{Binomial Dist}^n$$

$$\left.\begin{aligned} &P(X=0)\\ &+P(X=1)\\ &+P(X=2)\end{aligned}\right\}$$

Binomial ( $n \to$ trials.

$k \to$ success

$p \to$ prob. of success)

$$\Rightarrow \; {}^{n}C_{k} \cdot (p)^{k} \cdot (1-p)^{n-k}$$

$\{ n = 8$, $p = 0.5$

$$P(X < 3)$$

$$= P(X=0) + P(X=1) + P(X=2)$$

$$\to {}^{8}C_{0}(0.5)^{0}(1-0.5)^{8} + {}^{8}C_{1}(0.5)^{1}(1-0.5)^{7} + {}^{8}C_{2}(0.5)^{2}(1-0.5)^{6}$$

In just single line

Python

binom.cdf ( n = 8,
k = 2,
p = 0.5 ) → 0, 1, 2

+

= 0.144

---

C C C C C C C C

2 2 2 _ _ _ 2

= 256

x = 0 ⇒ 1 = $^8C_0$

x = 1 ⇒ $^8C_1$ = 8

x = 2 ⇒ $^8C_2 = \frac{8 \cdot 7}{2} = 28$

$$P(X<3)=\left(\frac{1+8+28}{256}\right)=\underline{\underline{0.144}}\ \#$$

Q 7 Chemical Plant Impurity Statistics Hard 50 41 %

The amount of impurity in a batch of a chemical product is a random variable with mean value 4.0 g and standard deviation 1.5 g.

If 50 batches are independently prepared,

what is the probability that the average amount of impurity in these 50 batches is between 3.5 and 3.8 g?

3.5 , 3.2 , 4.8 , 4.1 , 3.2 , 4.2 , . . .

mean = 4g = 4.

$\sigma = 1.5$

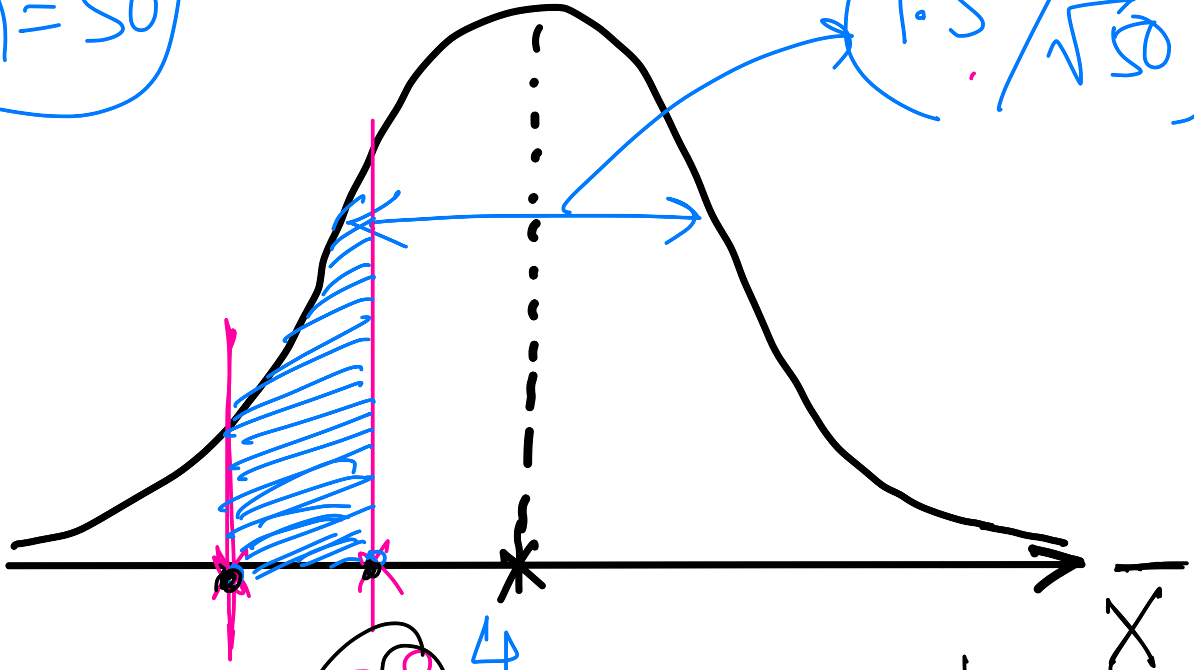
$$\left(\frac{X_1 + X_2 + - - - - + X_{50}}{50}\right) = \bar{X}$$

Also a Random Variable

$$\bar{X} \sim N\left(4g,\ \text{std dev} = \sigma/\sqrt{n} = \frac{1.5}{\sqrt{50}}\right)$$

$n = 50$

$(1.5/\sqrt{50})$



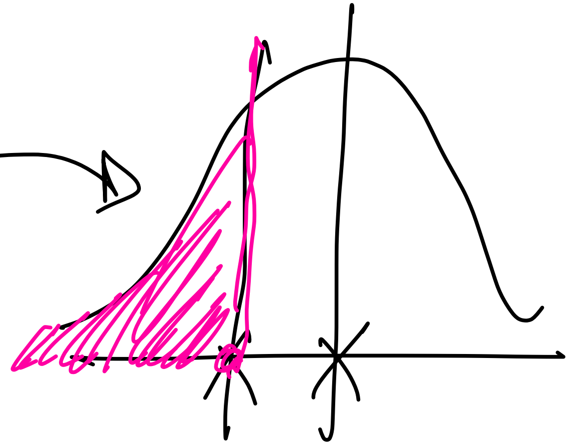

$$Z = \left(\frac{x-\mu}{\text{std. dev.}}\right)$$

$Z_{3.5}$

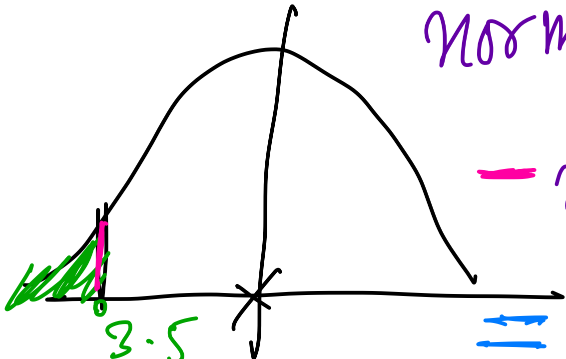
$$= \frac{3.5-4}{(1.5/\sqrt{50})}$$

$Z_{3.8}$

$$\frac{3.8-4}{(1.5/\sqrt{50})}$$



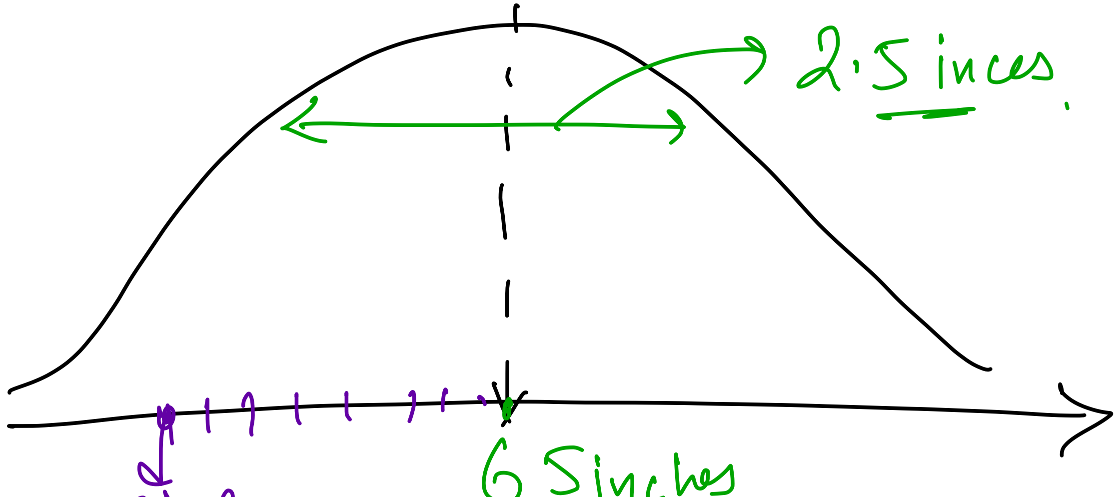

norm.cdf$(Z_{3.8})$




$-$ norm.cdf$(Z_{3.5})$

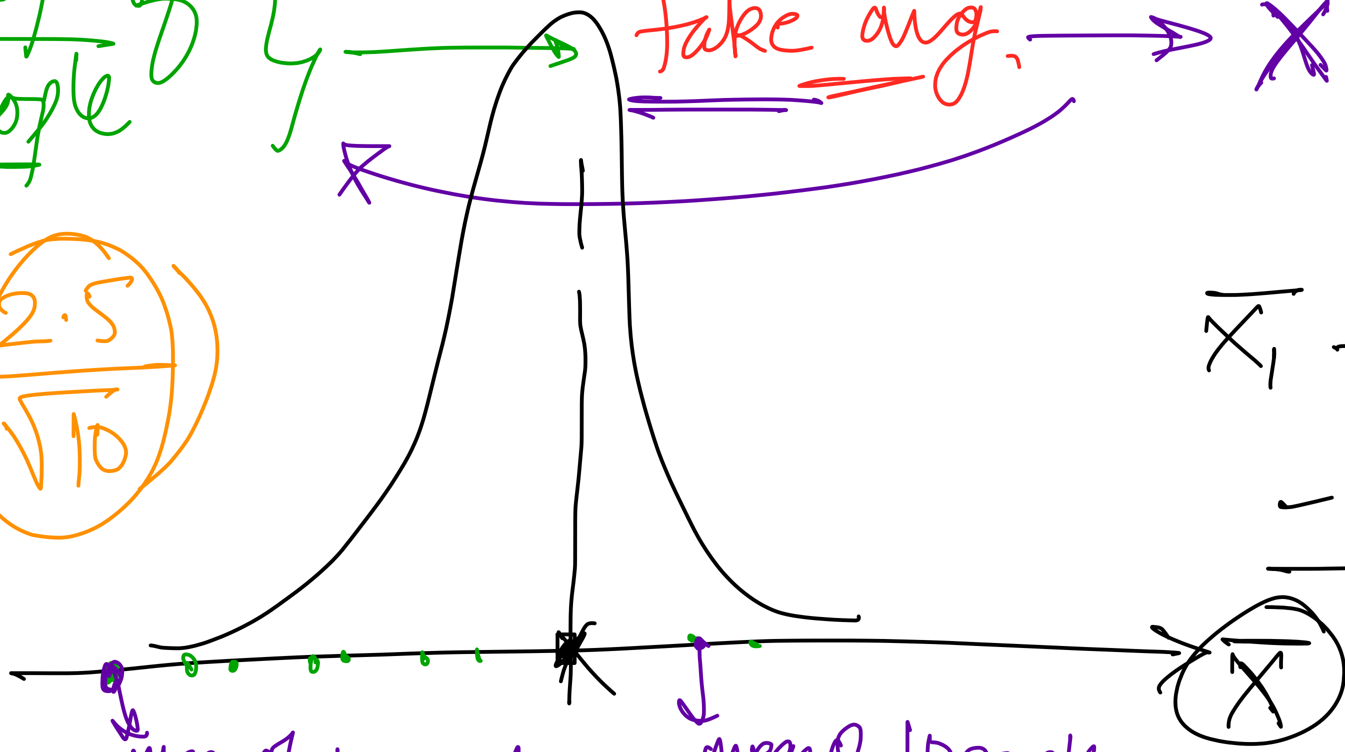
$= 0.16$

$X \rightarrow$ height




$X \sim N(65, 2.5)$

random sample of 10 people $\}$ → take avg. → $\bar{X}$




$\bar{X} \sim N\left(\mu, \frac{2.5}{\sqrt{10}}\right)$

$$\frac{\bar{X}_1 + \bar{X}_2 + \bar{X}_3 + \text{- - -} + \bar{X}_n}{n}$$

The verbal reasoning in GRE has an average score of 150, and a standard deviation of 8.5. A coaching center claims that their students are better. An average of 10 people showed that the students from this coaching center have an average of 155. At a 5% significance level (or 95% confidence level), can we conclude that students from the coaching center are better? Use the Z-test, and compute the p-value.

$\mu = 150$

$\boxed{\sigma = 8.5}$ → population std. devⁿ

$n = 10$ → $\bar{x} = 155$

Z test $= \dfrac{(x - \mu)}{\sigma}$

⇒ $H_0$: not better than avg. $(\bar{x} = \mu = 150)$

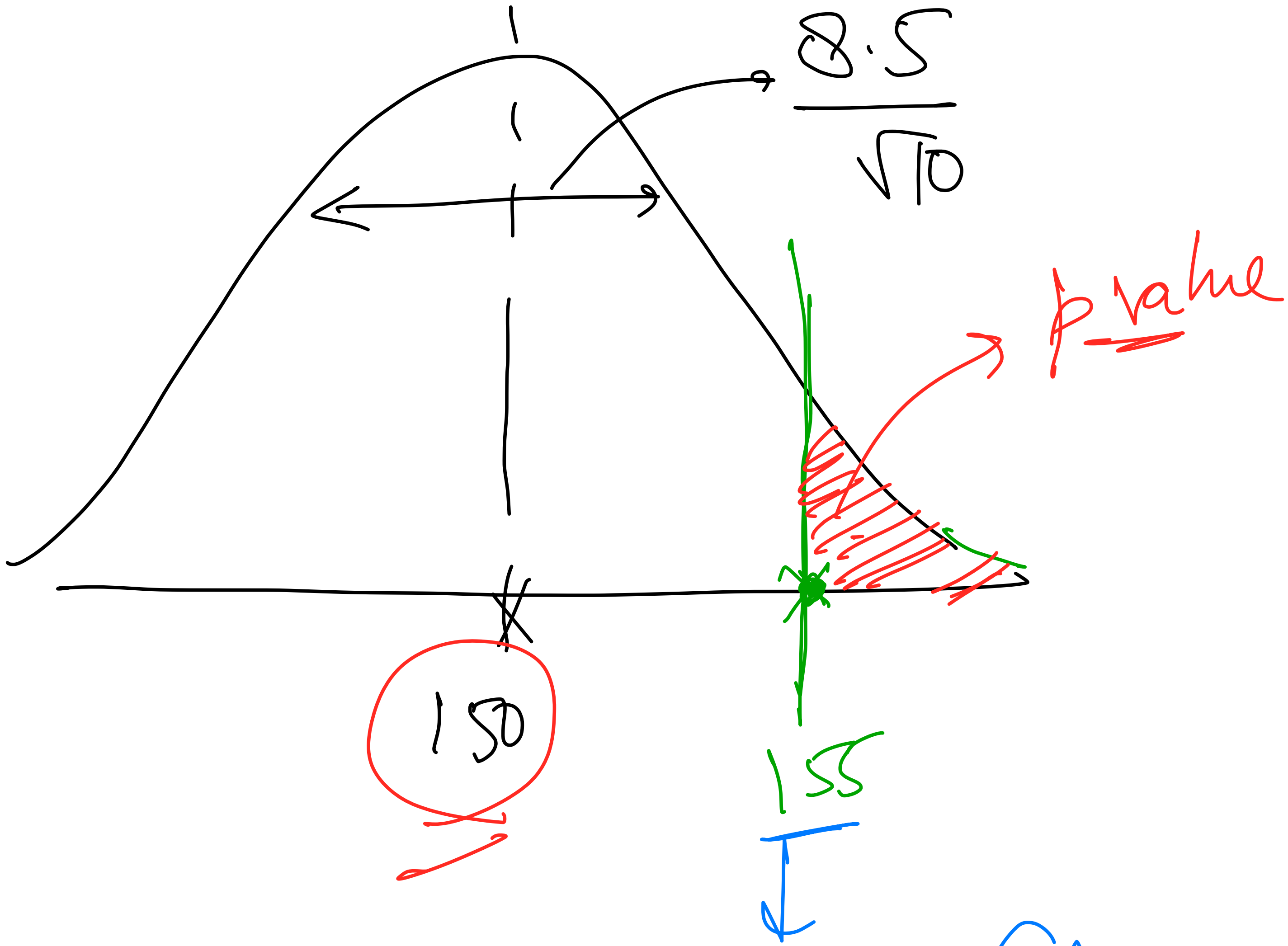
$H_A$: better $(\bar{x} > \mu)$

Right

⇒ $\bar{X} \sim \text{Normal}\left(150, \dfrac{8.5}{\sqrt{10}}\right)$.

$n = 10$

Acc. to CLT




Right

$\alpha = 0.05$

$$\text{Pvalue} = 1 - \text{norm.cdf}\left(\left(\frac{155-150}{\sigma/\sqrt{n}}\right)\right)$$

$$\text{pvalue} = 0.03 \Rightarrow p\text{-value} < \alpha\,(0.05)$$

⇒ Reject $H_0$, ∴ Student of this coaching center are better.

Pop. std. dev$^n$

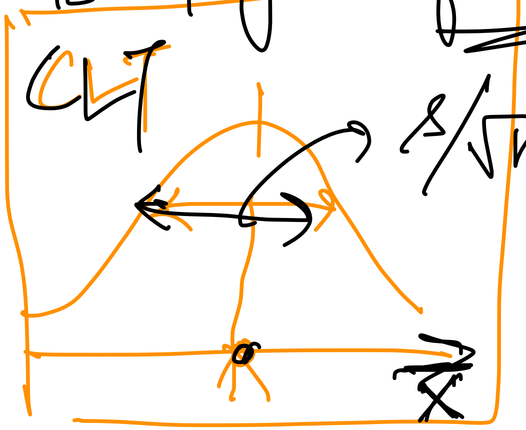
$\sigma$

- known → z test *
- Unknown
  - $n < 30$ → t-test *
  - $n > 30$ → z-test, t test *

Unknown → sample std. dev

[ , , , , ]. std( )

Sampling dist$^n$ of means.



A simple random sample of 18 adults with incomes below the poverty level gives the daily calcium intakes given below. (BPL)

**886, 633, 943, 847, 934, 841, 1193, 820, 774, 834, 1050, 1058, 1192, 975, 1313, 872, 1079, 809** → $\bar{x}$ (mean)

At the 5% significance level, do the data provide sufficient evidence to conclude that the mean calcium intake of all adults with incomes below the poverty level is less than the Recommended Adequate Intake of 1000 mg? Assuming the population standard deviation σ = 188 mg.

Choose the appropiate test and compute the p-value.

Select the right option below regarding condidering right test, p-value and conclusion.

$\left.\begin{array}{l}\mu = 1000 \\ \sigma = 188\end{array}\right\}$ $n < 30$

Z test    t test

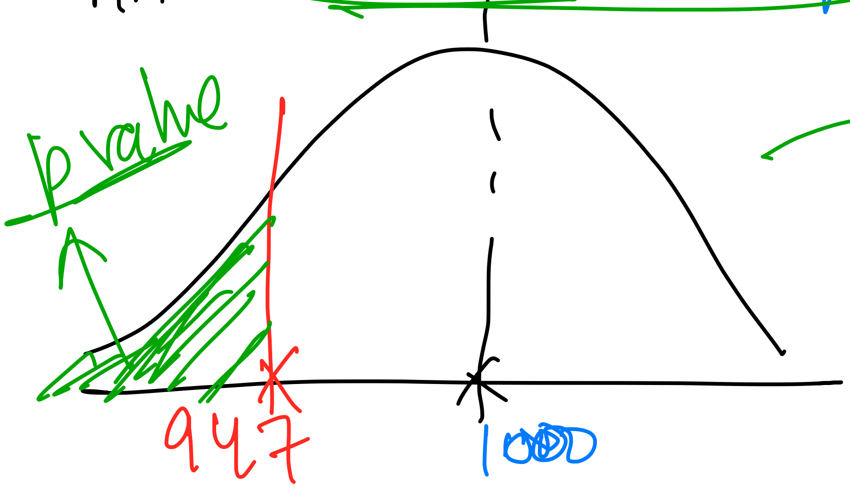
$n = 18$ ($n < 30$)

Recommended intake = 1000 mg

$H_0$: mean intake by poor is same as 1000 mg ($\bar{x} = 1000$ mg)

$H_A$: $\bar{x} < 1000$ mg.

Left




$$p\text{-value} = \text{norm.cdf}\left(\frac{947 - 1000}{188/\sqrt{18}}\right)$$

$$= 0.117$$

$$\boxed{p\text{-value} \not< \alpha\ (0.05)}$$

failed to reject $H_0$

Q 10 Significance Level Statistics Very Easy 20 39 %

What does Significance Level (alpha) represents in a Hypothesis Test ?

$\alpha = 0.05 \rightarrow$ confidence 95%.

$\alpha = 0.10 \rightarrow$ confidence 90%.

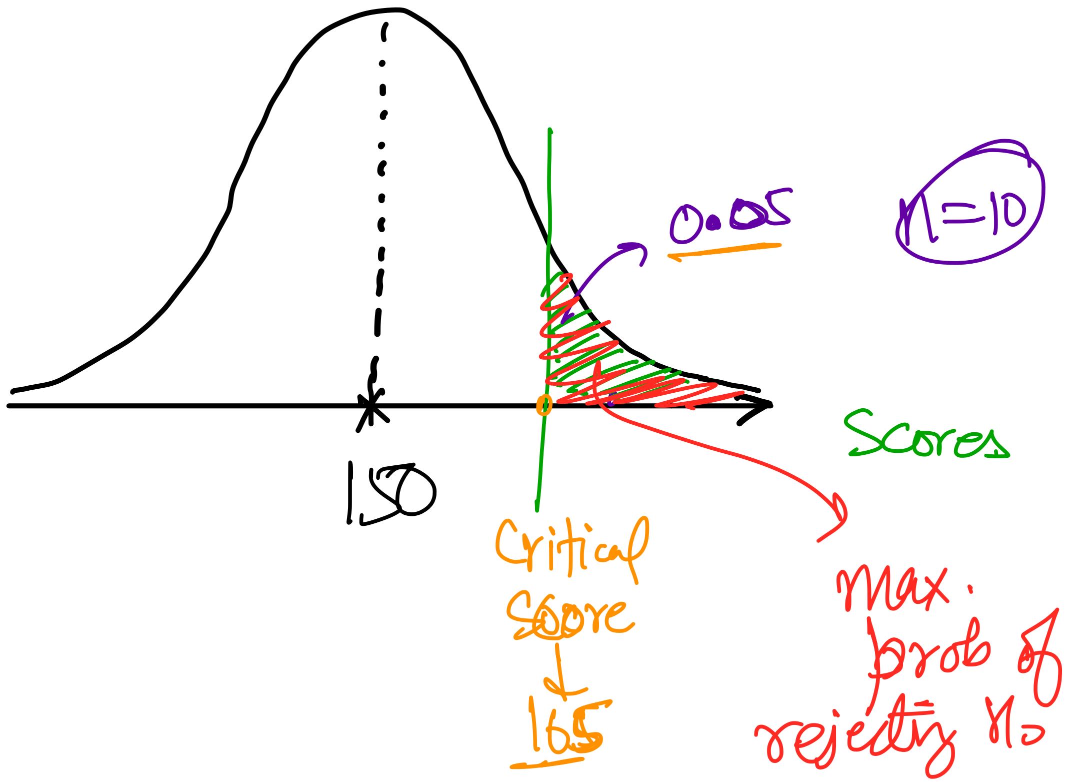
Maximum Probability of rejecting FALSE Null hypothesis

Minimum Probability of rejecting FALSE Null hypothesis

Maximum Probability of rejecting TRUE Null hypothesis

Minimum Probability of rejecting TRUE Null hypothesis

$\alpha = 0.05$



$\Rightarrow$ max. prob. of rejecting $\boxed{H_0}$ (Null hypothesis) $= \alpha$.

$H_0$ $\rightarrow$ True $\rightarrow$ Error

$H_0$ $\rightarrow$ False $\rightarrow$ Correct decision

Type I Error.

$\alpha$ : Max. prob. of $\underbrace{\text{Rejecting True } H_0}$