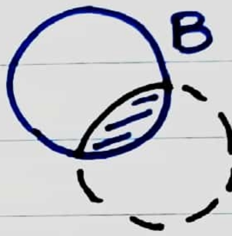
Probability Refresher

Conditional Probability

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



← acts as our reduced sample space

Probability of A given B

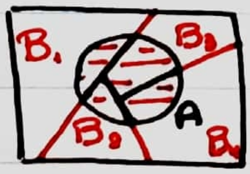
Joint Probability of A, B (Probability of both occuring simult.)

- If B is guaranteed to occur what's the Probability of A?
- $n(A)$ is the no. of ways in which A can occur.

• Total Probability

→ Given a set of events $B_1, B_2, \ldots, B_K$ that disjointly partition the sample space the Probability of any event A of the same Probability space is

$$P(A) = \sum_{i=1}^{K} P(A \cap B_i) = \sum_{i=1}^{K} P(A|B_i)\,P(B_i)$$



← A is constructed through its intersections with B_i.

- Check PS

Bayes Rule

$$P(B_i|A) = \frac{P(A|B_i)\,P(B_i)}{P(A)}$$

$P(B_i|A)$ → Posterior Probability (After learning about A)

$P(B_i)$ → Prior Probability (Before learning about A)

$P(A)$ → Marginal Probability

$$\sum_{j=1}^{K} P(A|B_j)\,P(B_j)$$

<bonus>

- In Bayesian inference, B can be a parameter and we're often interested in finding its distribution given the observed data A.
- In this case $P(A|B)$ is a function in the parameter B and is regarded as the likelihood of the parameter B given the data A. (the data is known, the parameter isn't).

</bonus>

Statistical Independence

- if $P(A \cap B) = P(A)P(B)$ then A and B are independent events

$$P(A|B) = P(A) \quad \text{and} \quad P(B|A) = P(B)$$ • implicitly

→ whether A has happened or not does not affect B and vice versa.

Marginal Distribution

$$P_x(x) = \int_{-\infty}^{\infty} P(x,y)\,dy$$

helps find the probability of $X = x$ irrespective to any outcome of Y.

• Sum for discrete (Σ)

Two random variables are independent if

$$P(x,y) = P_X(x)\,P_Y(y)$$ (for all x, y in their range)

- joint density
- Marginal densities

* Expected Value

$$\mu_x = E(X) = \int_{-\infty}^{\infty} x\,P_x(x)\,dx$$

- More generally

$$\mu_{g(x)} = E(g(X,Y)) = \int_{-\infty}^{\infty}\int_{-\infty}^{\infty} g(x,y)\,P(x,y)\,dy\,dx$$

* Variance

$$Var(g(X)) = E(g(X) - E(g(X)))^2 = E((g(x))^2) - (E[g(x)])^2$$
$$= E[(g(x))^2] - (E[g(x)])^2$$

- also applies for $g(X,Y)$

So $Var(X) = E(X^2) - (E(X))^2$

* Covariance

$$COV(X,Y) = E((X - E(X))(Y - E(Y)))$$
$$= E(XY) - E(X)E(Y)$$

* Correlation

$$Corr(X,Y) = \frac{COV(X,Y)}{\sigma_X\,\sigma_Y}$$

$-1 \leqslant Corr(X,Y) \leqslant 1$

- Perfectly anti-correlated ↗↙
- Perfectly correlated ↗↗

NN Sheet 1

Q1)

	Red Box	Blue Box	Green Box
Apples	3	5	3
Oranges	4	5	3
Limes	3	0	4

a) $P(\text{Fruit} = \text{Orange} \mid \text{Box} = \text{red}) = \dfrac{n(F = \text{Orange} \cap B = \text{Red})}{n(B = \text{red})}$

• n is the no. of fruits

$$= \frac{4}{3+4+3} = 0.4$$

b) $P(F = \text{Orange}) = P(F = \text{Orange} \mid B = \text{red}) . P(B = \text{red})$

• By the law of total Probability.

$+ P(F = \text{Orange} \mid B = \text{blue}) . P(B = \text{blue})$

$+ P(F = \text{Orange} \mid B = \text{green}) . P(B = \text{green})$

$$= \frac{4}{3+4+3} \cdot \frac{1}{5} + \frac{5}{5+5+0} \cdot \frac{1}{5} + \frac{3}{3+3+4} \cdot \frac{3}{5}$$

$= 0.36$

c) $P(B=\text{Red} \mid F=\text{Orange}) = \dfrac{P(F=\text{Orange} \mid B=\text{red}) . P(B=\text{red})}{P(F=\text{Orange})}$

• By Bayes Rule

$$= \frac{0.4 \times 0.2}{0.36} = 0.\dot{2}4$$

» Alternatively,

$P(B=\text{Red} \mid F=\text{Orange}) = \dfrac{n(B=\text{red} \cap F=\text{Orange})}{n(F=\text{Orange})}$

• By Conditional Probability.

$$= \frac{4}{4+5+3\times3} = 0.\dot{2}$$

d) $P(B=\text{green} \mid F=\text{Apple}) = \dfrac{n(B=\text{green} \cap F=\text{Apple})}{n(\text{Apple})}$

$$= \frac{3\times3}{3+5+3\times3} = \frac{9}{17}$$

» Alternatively,

$$P(F=\text{Apple}) = P(F=\text{Apple} \mid B=\text{Red}) . P(B=\text{Red}) + P(F=\text{Apple} \mid B=\text{blue}) . P(B=\text{blue}) + P(F=\text{Apple} \mid B=\text{green}) . P(B=\text{green})$$

$$= \frac{3}{3+4+3} \cdot \frac{1}{5} + \frac{5}{5+5+10} \cdot \frac{1}{5} + \frac{3}{3+3+4} \cdot \frac{3}{5} = \frac{17}{50}$$

$$P(B=\text{green} \mid F=\text{Apple}) = \left(\frac{3}{3+3+4} \cdot \frac{3}{5}\right) / \left(\frac{17}{50}\right) = \frac{9}{17}$$

Q2) Let x_1 be the result from the 1st flip and let x_2 be the result from the 2nd flip

a) $P(x_1=h \wedge x_2=h \mid x_1=h) = \dfrac{P(x_1=h \wedge x_2=h)}{P(x_1=h)}$

(h,h) only

$= \dfrac{1/4}{2/4} = \dfrac{1}{2}$

(h,t), (h,h)

Can also do $\dfrac{n(x_1=h \wedge x_2=h)}{n(x_1=h)} = \dfrac{1}{2}$

b) $P(x_1=h \wedge x_2=h \mid x_1=h \vee x_2=h)$

(h,h)

$= \dfrac{P(x_1=h \wedge x_2=h)}{P(x_1=h \vee x_2=h)} = \dfrac{1/4}{3/4} = \dfrac{1}{3}$

(h,h), (h,T), (T,h)

Note: $(A \vee B)(A \wedge B) = A \wedge B$

• Normally, the Sample Space shouldn't be given.

Q3)

• The two events E, F are independent iff
$P(E \cap F) = P(E)P(F)$

→ The Sample Space is $\{(h,h), (h,t), (t,h), (t,t)\}$

$P(E) = \dfrac{2}{4}$ ← (h,h), (h,t)

1st Coin lands on head

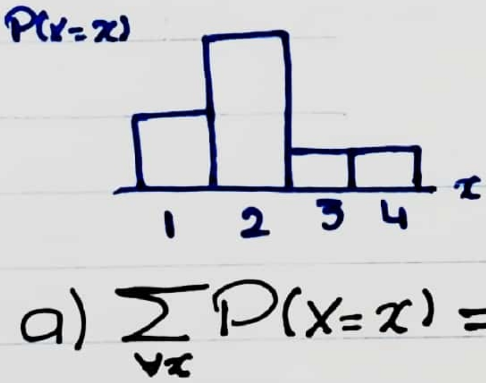
$P(F) = \dfrac{2}{4}$ ← (h,t), (t,t)

2nd Coin lands on tail

$P(E \cap F) = \dfrac{1}{4}$ ← (h,t)

• Clearly, $P(E \cap F) = P(E)P(F)$ and the two events are thereby independent.

Q4)



a) $\sum_{\forall x} P(X=x) = 0.25 + 0.5 + 0.125 + 0.125 = 1$

$P(X=x) \geq 0 \quad \forall x$

b) $E(X) = \sum_{x} x P(X=x) = 1\times0.25 + 2\times0.5$

$+ 3\times0.125 + 4\times0.125$

$= 2.125$

c) $E(X^2) = \sum_{x} x^2 P(X=x)$

$= 1^2\times0.25 + 2^2\times0.5 + 3^2\times0.125 + 4^2\times0.125$

$= 5.375$

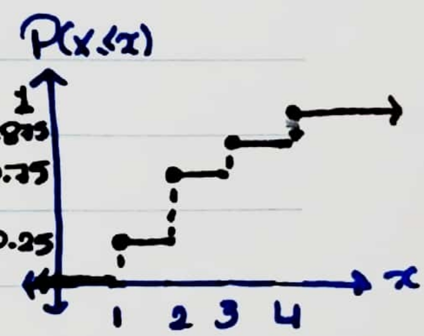
$Var(X) = E(X^2) - (E(X))^2$ (σ^2)

$= 5.375 - 2.125^2 = 0.859375$

$STD(X) = \sqrt{Var(X)} = 0.927$ (σ)

d) $P(X \leq x) = \sum_{t \leq x} P(t)$

x	1	2	3	4
$P(X=x)$	0.25	0.5	0.125	0.125
$P(X\leq x)$	0.25	0.75	0.875	1



e)

x	1	2	3	4
$y=\ln(x)$	0	0.693	1.0986	1.386
$P(Y=y)$	0.25	0.5	0.125	0.125

• under the Same Steps

$E(Y) = 0.657$

$Var(Y) = 0.1994$

$STD(Y) = 0.4465$

Q5) $E(X) = 1$ and $Var(X) = 5$

. $Var(X) = 5 \rightarrow E(X^2) = Var(X) + (E(X))^2$
$= 6$

a) $E((X+2)^2) = E(X^2 + 4X + 4)$:

$= E(X^2) + \underbrace{4E(X)}_{E(4X)} + \underbrace{4}_{E(4)}$

$= 6 + 4 \times 1 + 4 = 14$

b) $Var(4 - 3X) = (-3)^2 \, Var(X) = 9 \times 5 = 45$

Q6)

- Box Contains 5 green and 5 blue balls.
- Two balls are drawn at random

The no. of ways to

Draw two balls $^{10}C_2$

Draw two different balls $^{5}C_1 \times {}^{5}C_1$ • For every chosen red there are 5 choices of green

Draw two similar balls $^{5}C_2 + {}^{5}C_2$ • can choose 2 of the red or 2 of the green

Scanned with CamScanner

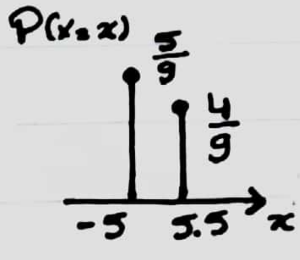
- Note that the no. of ways to draw two similar balls is also the no. of ways to not draw two different balls. $(2 \times {}^5C_2 = {}^{10}C_2 - ({}^5C_1)^2)$

- let X be a random variable representing the no. of points earned.

  → $X \in \{-5, +5.5\}$ depending on whether the two balls are similar or not.

$$P(X=-5) = \frac{{}^5C_1 \times {}^5C_1}{{}^{10}C_2} = \frac{5}{9}$$

$$P(X=5.5) = \frac{{}^5C_2 + {}^5C_2}{{}^{10}C_2} = \frac{4}{9}$$

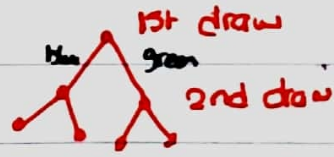


a) $E(X) = -5 \times \frac{5}{9} + 5.5 \times \frac{4}{9} = -\frac{1}{3}$

b) $E(X^2) = 25 \times \frac{5}{9} + 5.5^2 \times \frac{4}{9}$
$= \frac{82}{3}$

$$Var(X) = \frac{82}{3} - \left(\frac{1}{3}\right)^2 = \frac{245}{9}$$

- Can also compute the two probabilities using a probability tree



Q7) • Let X be a Continuous random Variable representing the amount of time that a Computer runs before breaking down.

→ The Probability density function is

$$f(x) = \begin{cases} \lambda e^{-x/100} & x \geq 0 \\ 0 & x < 0 \end{cases}$$

$$\int_{-\infty}^{\infty} f(x)\,dx = 1 \rightarrow \Big|^{\infty} \lambda e^{-x/100}(-100) = 1$$

$$\therefore -\lambda(e^{0}.(-100)) = 1 \quad (\lambda = \frac{1}{100})$$

a) $P(50 < X < 150) = \int_{50}^{150} \frac{1}{100} e^{-x/100}\,dx = 0.3834$

b) $P(X < 100) = \int_{-\infty}^{100} f(x)\,dx = \int_{0}^{100} \frac{1}{100} e^{-x/100}\,dx = 0.632$

Q8) • Since its a PDF

$$\int_{-\infty}^{\infty} f(x)\,dx = \int_{0}^{1} (a+bx^2)\,dx = {}_0\Big|^{1} ax + \frac{bx^3}{3}$$

$$= (a + \frac{b}{3}) = 1$$

Scanned with CamScanner

• Since $E(x) = 35$

$$\int_{-\infty}^{\infty} x f(x)\,dx = \int_0^1 (ax + bx^3)\,dx = \left.\frac{ax^2}{2} + \frac{bx^4}{4}\right|_0^1$$

$$= \frac{a}{2} + \frac{b}{4} = 35$$

Solving the System

$$a + \frac{1}{3}b = 1$$

$$\frac{1}{2}a + \frac{1}{4}b = 35$$

yields

$$a = -137 \qquad b = 414$$

$$f(x) = \begin{cases} 414x^2 - 137 & 0 \leq x \leq 1 \\ 0 & \text{otherwise} \end{cases}$$

Q9)

a) $E(aX+b) = \int_{-\infty}^{\infty} (ax+b) P(x)\, dx$

$= a\int_{-\infty}^{\infty} x P(x)\, dx + b\int_{-\infty}^{\infty} P(x)\, dx$

$= a.\ E(X) \quad + b\ .\ 1$

$= aE(X) + b$

• Also holds for Σ if X was discrete.

b) $Var(aX+b) = E((aX+b)^2) - (E(aX+b))^2$

$= E(a^2X^2 + 2abX + b^2) - (\underbrace{E(aX+b)}_{aE(X)+b})^2$

$= a^2E(X^2) + 2abE(X) + b^2$

$- \quad a^2(E(X))^2 + 2abE(X) + b^2$

$= a^2(E(X^2) - (E(X))^2) = a^2 Var(X)$

Scanned with CamScanner

c) $Var(X+Y) = E((X+Y)^2) - (E(X+Y))^2$

$$= E(X^2+2XY+Y^2) - (E(X)+E(Y))^2$$

$$= E(X^2) + 2E(XY) + E(Y^2) - \left((E(X))^2 + 2E(X)E(Y) + (E(Y))^2 \right)$$

$$= \left(E(X^2) - (E(X))^2\right) + \left(E(Y^2) - (E(Y))^2\right) + 2\left(E(XY) - E(X)E(Y)\right)$$

$$= Var(X) + Var(Y) + 2Cov(X,Y)$$

d) $Cov(X,Y) = E((X-\mu_x)(Y-\mu_y))$

• writing μ_x, μ_y for $E(X), E(Y)$ for easier notation.

$$= E(XY - X\mu_y - Y\mu_x + \mu_x\mu_y)$$

$$= E(XY) - \mu_y E(X) - \mu_x E(Y) + \mu_x\mu_y$$

$$= E(XY) - \mu_y\mu_x - \mu_x\mu_y + \mu_x\mu_y$$

$$= E(XY) - \mu_x\mu_y = E(XY) - E(X)E(Y)$$

• Notice $E(\underbrace{E(X)}_{\text{Treated as a const.}}) = E(X)$ as $\int_{-\infty}^{\infty} (\underbrace{\int_{-\infty}^{\infty} x P(x)\,dx}_{\text{not a function in } x \text{ (const.)}}) P(x)\,dx$

Q10)

a) $E(XY) = \int_{-\infty}^{\infty}\int_{-\infty}^{\infty} xy \, P(x,y) \, dx \, dy$

• Due to independence

$= \int_{-\infty}^{\infty}\int_{-\infty}^{\infty} xy \, P_X(x) P_Y(y) \, dx \, dy$

$= \int_{-\infty}^{\infty} P_X(x) \, dx \int_{-\infty}^{\infty} P_Y(y) \, dy = E(X) E(Y)$

$\neq 0$

→ Won't be computed to zero unless either $E(x)$ or $E(Y)$ is zero

c) $Cov(X,Y) = E(XY) - E(X)E(Y) = 0$

d) $Corr(X,Y) = \dfrac{Cov(X,Y)}{Var(X)Var(Y)} = 0$ • They are uncorelated

b) $Var(X+Y) = Var(X) + Var(Y) + 2Cov(X,Y)$

$= Var(X) + Var(Y)$

→ Only zero if both $Var(X)$, $Var(Y)$ are zero

e) $E(XY) - E(X+Y) = E(X)E(Y) - (E(X) + E(Y))$

→ Not zero unless $E(X)E(Y) = E(X) + E(Y)$

f) • Check C

Scanned with CamScanner

g) $\mathrm{Var}(XY) = E(X^2Y^2) - (E(XY))^2$

$$E(X^2Y^2) = \int_{-\infty}^{\infty}\int_{-\infty}^{\infty} x^2 y^2 P(x,y)\,dx\,dy$$

$$= \int_{-\infty}^{\infty} x^2 P_x(x)\,dx \int y^2 P_Y(y)\,dy$$

$$= E(X^2)\,E(Y^2)$$

$$\mathrm{Var}(XY) = E(X^2)E(Y^2) - (E(X))^2 (E(Y))^2$$

• which isn't necessarily zero

h) $\underbrace{\mathrm{Var}(X+Y)}_{\mathrm{Var}(X)+\mathrm{Var}(Y)} - (\mathrm{Var}(X) + \mathrm{Var}(Y)) = 0$

(as in b)

Q11)

$$P(x) = \frac{1}{\sqrt{2\pi\sigma^2}} \cdot e^{-\frac{(x-\mu)^2}{2\sigma^2}}$$

a) $E(X) = E(X-\mu) + \mu$

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

→ let $z = \frac{x-\mu}{\sigma}$, $dz = \frac{dx}{\sigma}$, $x \to -\infty \rightarrow z \to -\infty$

$x \to \infty \rightarrow z \to \infty$

$$E(X-\mu) = \frac{\sigma}{\sqrt{2\pi}} \int_{-\infty}^{\infty} z e^{-z^2/2} dz$$

$$= \frac{\sigma}{\sqrt{2\pi}} \lim_{a\to\infty} \underbrace{\int_{-a}^{a} z e^{-z^2/2} dz}_{\text{zero}}$$

• as $f(z) = z e^{-z^2/2}$ is odd ($f(-z) = -f(z)$)

b)

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

• let $z = \frac{x-\mu}{\sigma}$, $dx = \sigma dz$, $\begin{matrix} x\to\infty \\ x\to-\infty \end{matrix} \to \begin{matrix} z\to\infty \\ z\to-\infty \end{matrix}$

$$= \frac{\sigma^2}{\sqrt{2\pi}} \int_{-\infty}^{\infty} z^2 e^{-z^2/2} dz$$

• Integrating by Parts:

u	dv
z	$z e^{-z^2/2}$
1	$-e^{-z^2/2}$

$$= \frac{\sigma^2}{\sqrt{2\pi}} \left(\overset{①}{\Big|_{-\infty}^{\infty} -z e^{-z^2/2}} + \overset{②}{\int_{-\infty}^{\infty} e^{-z^2/2} dz} \right)$$

① $\lim_{z\to\pm\infty} \frac{z}{e^{z^2/2}}$

$= \lim_{z\to\pm\infty} \frac{1}{z e^{z^2/2}} = \frac{1/z}{e^{z^2/2}} = 0$

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

• let $z = \frac{(x-\mu)}{\sigma}$

$$\int_{-\infty}^{\infty} \frac{e^{-z^2/2}}{\sqrt{2\pi}} dz = 1$$

• Thus, ② $= \sqrt{2\pi}$

Thus, we're left with $Var(x) = \frac{\sigma^2}{\sqrt{2\pi}}(0+\sqrt{2\pi})$

$= \sigma^2$

• let $X \sim N(\mu, \sigma^2)$

$\rightarrow E(x) = \mu$ and $Var(x) = \sigma^2$

» For $Z = \frac{X-\mu}{\sigma}$ (using the results proven in Q9)

• $E(Z) = \frac{E(X)-\mu}{\sigma} = \frac{\mu-\mu}{\sigma} = 0$

• $Var(Z) = Var\left(\frac{X}{\sigma} - \frac{\mu}{\sigma}\right) = \frac{Var(x)}{\sigma^2} = 1$

Q12)

$$E(e^{-\alpha x}) = \int_{-\infty}^{\infty} e^{-\alpha x} \cdot \frac{1}{\sqrt{2\pi}} e^{-x^2/2} dx$$

$$= \int_{-\infty}^{\infty} \frac{1}{\sqrt{2\pi}} \cdot e^{-\left(\frac{x^2}{2}+\alpha x\right)} dx$$

• Not easy to integrate.

Observe

$$-\left(\frac{x^2}{2} + \alpha x\right) = -\frac{1}{2}(x^2 + 2\alpha x)$$

$$= -\frac{1}{2}((x+\alpha)^2 - \alpha^2) \quad \text{• Complete the square}$$

$$= \frac{-1}{2}(x+\alpha)^2 + \frac{\alpha^2}{2}$$

$$E(e^{-\alpha x}) = \int_{-\infty}^{\infty} \frac{1}{\sqrt{2\pi}} \cdot e^{-\frac{1}{2}(x-(-\alpha))^2} \cdot e^{\alpha^2/2} \, dx$$

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

$$= e^{\alpha^2/2}$$

↳ A Normal distribution with $\mu = -\alpha$ and $\sigma = 1$ (must integrate to 1)

Q13)

$$E(e^{-x} + e^{-2x} + x^2) = \underset{\alpha=1}{E(e^{-x})} + \underset{\alpha=2}{E(e^{-2x})} + E(X^2)$$

$$= e^{0.5} + e^{2} + 1$$

$X \sim N(\mu = 0, \sigma^2 = 1)$

• $\sigma^2 = E(X^2) - \mu^2$

→ So $E(X^2) = 1$

Scanned with CamScanner

Q14)

Consider the numbers

X
-2
3
0.5
0.3
4.2
9
-3.4

1. Compute the mean

2. Compute the variance

3. Compute the standard deviation

4. What happens to the standard deviation (without computing it) if we add the point 10,000.

1. $\bar{X} = \frac{\sum_1^n X_i}{n} = 1.6571$

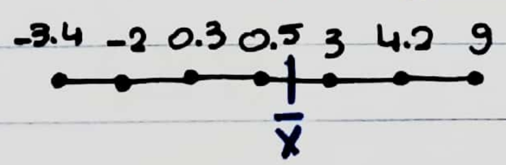
2. $S^2 = \frac{1}{n-1} \sum_1^n (X_i - \bar{X})^2 = \frac{1}{n-1}\left(\sum_1^n X_i^2 - n\bar{X}^2\right)$

$= \frac{1}{7-1}(123.54 - 7 \times 1.657^2)$

$= 17.3867$

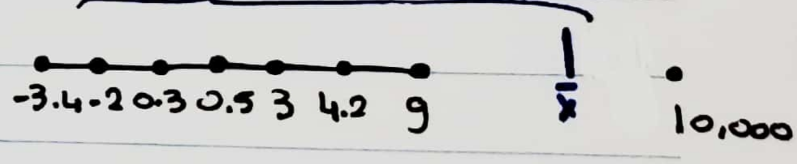
3. $S = 4.16967$

4. It will greatly increase



$\Rightarrow$

squared distances to $\bar{X}$ are much larger



Scanned with CamScanner