# Assignment #1

ABHINAV KUMAR
180050003

BHASKAR GUPTA
180050022

# Assignment #1

Date:

Q.1 $|x_i - \mu| \leqslant \sqrt{\sum_{j=1}^{N} (x_j - \mu)^2} \leqslant \sqrt{\sigma^2 (N-1)}$

$= \sigma \sqrt{N-1}$

$\Rightarrow |x_i - \mu| \leqslant \sigma \sqrt{N-1}$

For large $n$,
Probability of $S_K = \{x_i : |x_i - \mu| \geqslant \sigma \sqrt{n-1}\}$
is $\frac{|S_K|}{n} \leqslant \frac{1}{n-1}$ according to Chebychev inequality

For $n \to \infty$
$\frac{|S_K|}{n} \to 0$

While, from above inequality, $\frac{|S_K|}{n} = 0$ since $\forall$
$i \in [1, n]$, $|x_i - \mu| \leq \sigma \sqrt{n-1}$

Q.2 Chebychev-Cantelli's inequality - NOTE for $\sigma=0 \Rightarrow \tau=0, \mu=0$ equality holds

$S_K = \{x_i \mid x_i - \bar{x} \geq K\sigma\}$

$\frac{|S_K|}{N} < \frac{1}{1+K^2}$ for $K=1$,

$\frac{|S_1|}{N} < \frac{1}{2}$

But $S_U = \{x_i \mid x_i \geq \tau\}$

$\frac{|S_U|}{N} \geq \frac{1}{2}$ $\boxed{\text{Hence } \tau \text{ cannot lie in } S_K}$

$\Rightarrow \tau \leq \bar{x} + \sigma$

Similarly $S_{K'} = \{x_i \mid x_i - \bar{x} \leq -K\sigma\}$

$\frac{|S_{K'}|}{N} < \frac{1}{1+K^2}$ for $K=1$, $\frac{|S_{1'}|}{N} < \frac{1}{2}$

$S_L = \{x_i \mid x_i \leq \tau\}$ $\frac{|S_U|}{N} \geq \frac{1}{2}$ [Median def.]

$\boxed{\text{Hence } \tau \text{ can't lie in } S_{K'}}$

$\Rightarrow \tau \geq \bar{x} - \sigma$

$\Rightarrow \bar{x} - \sigma \leq \tau \leq \bar{x} + \sigma \Rightarrow \boxed{|\bar{x} - \tau| \leq \sigma}$

Q.3

$$P(C_R|R) = \frac{P(R|C_R) \times P(C_R)}{P(R|C_R)P(C_R) + P(R|C_B)P(C_B)}$$

where $P(C_R)$ is probabillity of auto being Red

$P(C_B)$ " " " Blue

~~P(R) is probabillity of XYZ seeing red object as re~~

P(B) " "

$$P(C_R|R) = \frac{\frac{99}{100} \times \frac{1}{100}}{\frac{99}{100} \times \frac{1}{100} + \frac{2}{100} \times \frac{99}{100}} = \frac{1}{3} = 33\%$$

Main Argument of lawyer will be the probabillity of XYZ seeing a red auto really when he observed & reported it to be red is only 33%. so, he is not a reliable.

4.

(a) $P(C_i | Z_1)$ is $1/3$ for all $i \in (1,3)$.
as $C_i$ and $Z_1$ are independent events.

(b) $P(H_3 | (C_1 \cap Z_1)) = 1/2$ as it was equally probable that he would open door 2 or 3 as neither contains the car and can't open door 1 as contestant chose it and $P(C_1|Z_1) = P(C_1)$

$P(H_3 | (C_2 \cap Z_1)) = 1$ as he can't chose door 1 as the contestant chose it and can't chose door 2 as car is in it and $P(C_1|Z_1) = P(C_1)$

$P(H_3 | C_3 \cap Z_1) = 0$ as he can't open the door with car present behind it.

(c) $\dfrac{P(H_3 | (C_2 \cap Z_1))}{P(H_3 | (C_1 \cap Z_1))} = \dfrac{1}{1/2} = 2.$

$\therefore$ Host opening door 3 given contestant chose door 1 implies car behind door 2 is twice as probable behind door 1.

Let $p$ = Prob behind door 1

$\therefore 2p$ = Prob behind door 2

$p + 2p = 1$. (As prob behind door 3 is zero)

$\therefore p = 1/3 \Rightarrow$ Behind door 1.

$2p = 2/3 \Rightarrow$ Behind door 2.

c) $2/3$ d) $1/3$.

(e) It is beneficial to swap as probability doubles.

(f) $p_1 = P(H_3 | (C_1 \cap Z_1)) = 1/2$ as it was equally probable he would open door 2 or door 3 as he is whimsical.

$p_2 = P(H_3 | (C_2 \cap Z_1)) = 1/2$ as it was equally probable he would open door 2 or door 3

$\therefore$ as $\frac{P_1}{P_2} = 1$.

and $p_1 + p_2 = 1$. (As only 2 choices are left).

$\therefore p_1 = 1/2 = 1/2$.

So no need to change

5. Quartile (first quartile) method produces the best relative mean squared error as mean is easily affected by outliers and lower quartile is even less affected than median for increased outliers because for lower quartile to be affected 75% of values have to be increased while for median to be affected just 50% of the values in the range need to be affected (As the values are being increased). ~~But~~ 60% disturbance increases the values for all three values than 30%.

Q.6

a) Let newMean be $\mu'$

Then $\mu'(n+1) =$ Sum of all elements of new array

$= \mu(n) + X$

$$\Rightarrow \boxed{\mu' = \frac{\mu(n)+X}{(n+1)}} \qquad \mu' \longrightarrow \text{newMean}$$

b) NewMedian

Case-I $n$ is even

newMedian is going to be $\left(\frac{n+2}{2}\right)^{th}$ element in the array

if newValue < $A(n/2)$

$\Rightarrow A'\left(\frac{n+2}{2}\right) = A(n/2)$ newMedian = $A(n/2)$

if newValue > $A((n+2)/2)$

$\Rightarrow A'\left(\frac{n+2}{2}\right) = A\left(\frac{n+2}{2}\right)$ newMedian = $A\left(\frac{n+2}{2}\right)$

else

newMedian = newValue

Case-II $n$ is odd.

newMedian is going to be $\left(A'\left(\frac{n+1}{2}\right) + A'\left(\frac{n+3}{2}\right)\right)/2$

where $A'$ is the new array.

if newValue <= $A(n-1/2)$

$A'\left(\frac{n+1}{2}\right) = A\left(\frac{n-1}{2}\right)$, $A'\left(\frac{n+3}{2}\right) = A\left(\frac{n+1}{2}\right)$

if newValue >= $A((n+3)/2)$

$A'\left(\frac{n+1}{2}\right) = A\left(\frac{n+1}{2}\right)$, $A'\left(\frac{n+3}{2}\right) = A\left(\frac{n+3}{2}\right)$

else

newMedian = $\frac{\left(\text{newValue} + A\left(\frac{n+1}{2}\right)\right)}{2}$

Scanned by CamScanner

Date:

Q.6c

$$\sigma_{old}^2 = \frac{\sum_{i=1}^{n} (x_i - \bar{x}_{old})^2}{n-1}$$

$$\sigma_{new}^2 \times n = \sum_{i=1}^{n+1} (x_i - \bar{x}_{new})^2$$

$$= \sum_{i=1}^{n} (x_i - \bar{x}_{new})^2 + (A_{new} - \bar{x}_{new})^2$$

$$= \sum_{i=1}^{n} (x_i - \bar{x}_{old} + \bar{x}_{old} - \bar{x}_{new})^2 + (A_{new} - \bar{x}_{new})^2$$

$$= \sum_{i=1}^{n} (x_i - \bar{x}_{old})^2 + (n)(\bar{x}_{old} - \bar{x}_{new})^2$$

$$+ \quad 2(\bar{x}_{old} - \bar{x}_{new}) \sum_{i=1}^{n} (x_i - \bar{x}_{old}) + (A_{new} - \bar{x}_{new})^2$$

$$= \left[\sigma_{old}^2 \times (n-1)\right] + \left[A_{new} - \bar{x}_{new}\right]^2 + (n)(\bar{x}_{old} - \bar{x}_{new})^2$$

$$\Rightarrow \sigma_{new}^2 = \frac{\sigma_{old}^2 \times (n-1) + \left[A_{new} - \bar{x}_{new}\right]^2 + n(\bar{x}_{old} - \bar{x}_{new})^2}{n}$$

Q.E.D.

Q.6 To update the histogram, we will simply choose the bin in which newDataValue lies and increase the height of the bin by 1 unit if on y-axis number of elements in the bin is present.

Q.7 $$1 - \frac{(365)(365-1)\cdots(366-n)}{(365)^n} \geq p$$

Since first person has 365 choices, second person has 364 choices of day & so on.