ASSIGNMENT-1

Middepogu Manoj:160050075

Pavan Bhargav Tangirala:160050076

11th AUG, 2017

<u>PROBLEM 1</u>):- Given n values $\{x_i\}_{i=0}^{n}$ having mean μ, and median ν and standard deviation σ and n is even.

<u>R.T.P</u>:-$\langle\mu - \nu\rangle \leq \sigma$. Let $|\mu - \nu|$ be P.

$$\langle \mu - \nu \rangle = \left\langle \frac{\sum_{i=1}^{n} x_i}{n} - \nu \right\rangle$$

$$=> \langle \mu - \nu \rangle = \left\langle \frac{\sum_{i=1}^{n} (x_i - \nu)}{n} \right\rangle$$

(Here, <> denotes modulus here.)

As we know that $|\sum_{i=1}^{n}(x_i - \nu)| \leq \sum_{i=1}^{n}|x_i - \nu|$.So,nP is less than $\sum_{i=1}^{n}|x_i - \nu|$ and the equation $\sum_{i=1}^{n}|x_i - x|$ has it's minimum at x=ν.

So we can say that $nP \leq \sum_{i=1}^{n}|x_i - \mu|$. We Know that R.M.S $\geq$ A.M for any 2 set of Positive No.'s.

$$\text{So,} \sqrt{\frac{\sum_{i=1}^{n}(x_i-\mu)^2}{n}} >= \left(\frac{\sum_{i=1}^{n}|x_i-\mu|}{n}\right)$$

$$(i) \sqrt{\frac{\sum_{i=1}^{n}(x_i-\mu)^2}{n-1}} >= \sqrt{\frac{\sum_{i=1}^{n}(x_i-\mu)^2}{n}} >= \left(\frac{\sum_{i=1}^{n}|x_i-\mu|}{n}\right)$$

$$(ii) \left(\frac{\sum_{i=1}^{n}|x_i-\mu|}{n}\right) >= \left(\frac{\sum_{i=1}^{n}|x_i-\nu|}{n}\right)$$

So,We can prove that $\sigma >= |\mu - \nu|$, and the equality happens when all values are equal(from (i)).

PROBLEM 2):- Given 4 sets of n values $\{x_i\}_{i=1}^{n}$; $\{y_i\}_{i=1}^{n}$; $\{z_i\}_{i=1}^{n}$; $\{w_i\}_{i=1}^{n}$ where $z_i = ax_i + b; w_i = cy_i + d$.
Proof:-
Let μ_x be the mean of $\{x_i\}_{i=1}^{n}$ then $a\mu_x + b$ is the mean of $\{z_i\}_{i=1}^{n}$.
Similarly $c\mu_y + d$ is the mean of $\{w_i\}_{i=1}^{n}$ where μ_y is the mean of$\{y_i\}_{i=1}^{n}$.

$$r(z,w) = \frac{\sum_{i=1}^{n}(z_i-\mu_z)(w_i-\mu_w)}{\sqrt{\sum_{i=1}^{n}(z_i-\mu_z)^2 * \sum_{i=1}^{n}(w_i-\mu_w)^2}}$$

$$r(z,w) = \frac{\sum_{i=1}^{n}(ax_i+b-(a\mu_x+b))(cy_i+d-(c\mu_y+d))}{\sqrt{\sum_{i=1}^{n}(ax_i+b-(a\mu_x+b))^2 * \sum_{i=1}^{n}(cy_i+d-(c\mu_y+d))^2}}$$

$$r(z,w) = \frac{\sum_{i=1}^{n}(ax_i-a\mu_x)(cy_i-c\mu_y)}{\sqrt{\sum_{i=1}^{n}(ax_i-a\mu_x)^2 * \sum_{i=1}^{n}(cy_i-c\mu_y)^2}}$$

$$r(z,w) = \frac{(ac) * \sum_{i=1}^{n}(x_i-\mu_x)(y_i-\mu_y)}{\sqrt{(ac)^2 \sum_{i=1}^{n}(x_i-\mu_x)^2 * \sum_{i=1}^{n}(y_i-\mu_y)^2}}$$

$$r(z,w) = (ac/|ac|) * \frac{\sum_{i=1}^{n}(x_i-\mu_x)(y_i-\mu_y)}{\sqrt{\sum_{i=1}^{n}(x_i-\mu_x)^2 * \sum_{i=1}^{n}(y_i-\mu_y)^2}}$$

By definition of correlation coefficient r for all x,y:-

$$r(x,y) = \frac{\sum_{i=1}^{n}(x_i-\mu_x)(y_i-\mu_y)}{\sqrt{\sum_{i=1}^{n}(x_i-\mu_x)^2 * \sum_{i=1}^{n}(y_i-\mu_y)^2}}$$

$(ac/|ac|)$ is 1 if $ab > 0$ and -1 if $ab \ < \ 0$.
So,$r(z,w) = +r(x,y) \ if \ ab \ > \ 0$, i.e., a and b are of same sign.
else $r(z,w) = -r(x,y) \ if \ ab \ < 0$, i.e., a and b are of different sign.

<u>PROBLEM 3</u>):- Given set of n values $\{x_i\}_{i=1}^{n}$ and μ is the mean and σ is the standard deviation.

R.T.P:-$|x_i - \mu| \leq \sigma\sqrt{n-1}$

$$\sigma = \sqrt{\frac{\sum_{i=1}^{n}(x_i-\mu)^2}{n-1}}$$

$$\text{So,}\sigma\sqrt{n-1} = \sqrt{\sum_{i=1}^{n}(x_i-\mu)^2}$$

$$\sigma\sqrt{n-1} = \sqrt{\sum_{j!=i}^{n}(x_j-\mu)^2 + (x_i-\mu)^2}$$

$$(\sigma\sqrt{n-1})^2 - (x_i-\mu)^2 = \sum_{j!=i}^{n}(x_j-\mu)^2$$

$$\text{As,}\sum_{j!=i}^{n}(x_j-\mu)^2 >= 0$$

$$(\sigma\sqrt{n-1})^2 - (x_i-\mu)^2 >= 0$$

$$\text{So, } \sigma\sqrt{n-1} >= |x_i-\mu|$$

PROBLEM 4):- Given:-C_1, C_2, C_3 are events that the Car is behind doors 1,2,3 respectively.
$P(C - i) = \frac{1}{3}, i\epsilon\{1, 2, 3\}$.

(i)Z_i be the event that contestant chose door i.

$$P(C_i|Z_i) = \frac{P(C_i.Z_i)}{P(Z_i)}$$

As, C_i *and* Z_i are independent

$$P(C_i|Z_i) = \frac{P(C_i)P(Z_i)}{P(Z_i)}$$

$$P(C_i|Z_i) = \frac{\frac{1}{3}*\frac{1}{3}}{\frac{1}{3}}$$

$$\text{So, } P(C_i|Z_i) = \frac{1}{3}.$$

Same for $z = 1$ so,$P(C_i|Z_1) = \frac{1}{3}$ for all $i\epsilon\{1, 2, 3\}$.

(ii)

$$P(H_3|C1, Z1) = \frac{1}{2}.$$

$$\frac{C}{Z_1} - -$$

As the contestant chose door 1 and the car is in door 1 the host has 2 choices door 2 and door 3 as he doesn't choose the door in which there is a car or the door the contestant chose.

$$\overline{Z_1}\ \underline{C}\ -$$

$$P(H_3|C2, Z1) = 1$$

As the contestant chose door 1 and the car is in door 2 host has only one choice i.e., to go with 3.

$$\overline{Z_1}\ -\ \underline{C}$$

$$P(H_3|C3, Z1) = 0$$

As the contestant chose door 1 and the car is in door 3 host has only one chance i.e., to go with 2(not 3).So,probability that he chooses door 3 when car is in door 3 and contestant chose door 1 is Zero.

(iii)

$$P(C_2|H_3,Z_1) = \frac{P(H3|C_2,Z_1)*P(C_2Z_1)}{P(H_3Z_1)}$$

We have calculated the value of $P(H3|C_2,Z_1)$ is 1 in (ii) and $P(C_2|Z_1)$ is $\frac{1}{9}$ as C_2 *and* Z_1 are 2 independent events so $P(C_2|Z_1)$ *is equal to* $P(C_2) * P(Z_1)$ i.e., $\frac{1}{9}$.

$$\underline{\ \ }_{Z_1}\ \underline{H}\ \underline{H}\ \ |\ \ \underline{H}\ \underline{\ \ }_{Z_1}\ \underline{H}\ \ |\ \ \underline{H}\ \underline{H}\ \underline{\ \ }_{Z_1}.$$

These are the 3 cases possible for contestant and host for selecting doors respectively .So totally there are 6 cases and probability that host chooses door3 and contestant choosing door 1 is $\frac{1}{6}$ so $P(H_3\ Z_1)$ *is* $\frac{1}{6}$

So, the value of $\frac{P(H3|C_2,Z_1)*P(C_2Z_1)}{P(H_3Z_1)}$ is $\frac{1*1/9}{1/6} = \frac{2}{3}$.

(iv)

$$P(C_1|H_3,Z_1) = \frac{P(H3|C_1,Z_1)*P(C_1Z_1)}{P(H_3Z_1)}$$

$$P(C_1|H_3,Z_1) = \frac{1}{2} \text{ (by ii). } P(C_1\ Z_1) = \frac{1}{9}$$

As,C_1 *and* Z_1 are independent $P(C_1Z_1) = P(C_1) * P(Z_1) = \frac{1}{9}$ and $P(H_3,Z_1) = \frac{1}{6}$ as we have calculate in (iii).

$$\text{So, } \frac{P(H3|C_1,Z_1)*P(C_1Z_1)}{P(H_3Z_1)} = \frac{\frac{1}{2}*\frac{1}{9}}{6} = \frac{1}{3}.$$

(v)

As,$P(C_2|H_3,Z_1) > P(C_1|H_3,Z_1)$ we can say that Shifting is beneficial to win the Car. where $P(C_2|H_3,Z_1)$ and $P(C_1|H_3,Z_1)$ meanings remain as mentioned in the question.

PROBLEM 5):- Given: There is an island where people speak one or more of n different languages.The proportion of people speaking those languages is $p_1, p_2, ..., p_n$ where for all i $0 \leq p_i \leq 1$.

The proportion of people who can speak exactly one language, assuming the ability to speak different languages are all independent is

$$\sum_{i=0}^{n} p_i(1 - p_{i+1}) ....$$

$$= \prod_{i=1}^{n}(1 - p_i)(\tfrac{p_1}{1-p_1} + \tfrac{p_2}{1-p_2} + ...)$$

$$= \prod_{i=1}^{n}(1 - p_i)(\textstyle\sum_{i=1}^{n} \tfrac{p_i}{1-p_i})$$

Note:- Because all languages are independent we can say portion that are speaking only language p_i is $p_i * (1 - p_{i+1}) * ....$ i.e., in terms of probability people speaking p_i for every i and not speaking any other language for all j except i.So,we get the above summation.