Statistics

Question solved

Suman Pradhan

2019

1.

(a) What is variable? Define with example different types of variable?

* A variable is a quantity that may change within the context of a mathmatical problem or experiment. Typically, we use a single letter to represent a variable. The letters x, Y, Z are common generic symbols used for variables.

* The two main categories of variable are categorical and numeric. A variable is characteristic that can be measured and that can assume different values. Heigh, age, income, province are all examples of variables.

1(b)

(b) What do you mean by dispersion? Discuss different measures of dispersion?

(*) dispersion: measures the fluctuations (variability) around the characteristics of central tendency.

* Measures of Dispersion:

(*) Mean Deviation.

(*) Range.

(*) Variance.

(*) Standard Deviation.

(*) Coefficient of Variation.

1

1 (c) Calculate the standard deviation from the following data:

40, 60, 65, 65, 65, 68, 68, 70, 70, 70, 70, 70, 70, 74, 75, 75, 90, 95.

We know:

$$SD. = \sqrt{\frac{\Sigma f_i x_i^2}{N} - \left(\frac{\Sigma f_i x_i}{N}\right)^2}$$

x_i	f_i	$f_i x_i$	$f_i x_i^2$
40	1	40	40
60	1	60	60
65	3	65×3	65×9
68	2	68×2	68×4
70	6	70×6	70×36
74	1	74	74
75	2	75×2	75×4
90	1	90	90
95	1	95	95
$\Sigma f_i = 18$		$\Sigma f_i x_i = 1260$	$\Sigma f_i x_i^2 = 4036$

$$= \sqrt{\frac{4036}{18} - \left(\frac{1260}{18}\right)^2}$$

$$= \sqrt{224.22 - 4900}$$

$$= \sqrt{\frac{90374}{18} - \left(\frac{1260}{18}\right)^2}$$

$$= \sqrt{5020.77 - 4900}$$

$$= 10.98$$

$$\sum \frac{x_i^2}{n} - \left(\frac{\sum x_i}{n}\right)^2$$

$$\left(\frac{\Sigma x_i - \bar{x}}{N}\right)^2$$

2(a)

(a) Define outcome, sample space and event of an experiment?

Outcome: A result of a random expirenment

Sample space: The set of all possible outcomes.

Event: A subset of sample space.

Union and intersection: If A and B are events, then $A \cup B$ and $A \cap B$ are also events.

(b) State the law of total probability?

The total probability Rule is a fundamental rule in statistics. relating to conditional and marginal probabilities. The rules states that if the probabilites of an event is unknown, it can be calculated using the known probabilities of several distinct events.

$$p(A) = p(A|B)\,p(B) + p(A|B')\,p(B')$$

law of total probability

$$p(A) = \Sigma\, p(A \cap B)$$
$$= \Sigma\, p(A|B)\,p(B_i)$$

2@:

What is conditional probability?

* $p(A|B)$ - the conditional probability. the probability of event A occurring given that B has already occured.

If $p(A) = 0.4$, $p(B) = 0.5$ and $p(A \cap B) = 0.3$ then find $p(A|B)$ and $p(B|A)$?

* $p(A|B) = \dfrac{p(A \cap B)}{p(B)} = \dfrac{0.3}{0.5} = 0.6$

* $p(B|A) = \dfrac{p(A \cap B)}{p(A)} = \dfrac{0.3}{0.4} = 0.75$

3

(a) What dog you mean by random variable?

A random variable is a numerical description of the outcome of a statistical experinment.

A random variable that may assume only a finite number or an Infinite sequence of values is said to be discrete

(c) Given that X is a random variable whose mean = 3, find the variance of 4x+2

$E(x) = 3$

$V(x) = 4^2 V(x) -$

$= 16 \times 3$

2. (c) Show that variance is independent of origin but not of scale

$$\sigma^2 = \frac{\Sigma(x-\bar{x})^2}{n}$$

Now; $u = \frac{x-A}{h}$

$\Rightarrow uh = x - A$

$\Rightarrow x - A = uh$

$\Rightarrow x = uh + A$

$5 \times 5 \cdot 5 = 20$
$4 \times 5 = 20$

$\Rightarrow \Sigma x = \Sigma A + h\Sigma u$

$\Rightarrow \Sigma x = nA + h\Sigma u$

$\Rightarrow \frac{\Sigma x}{n} = A + h\frac{\Sigma u}{n}$ A, h constant

$\Rightarrow \boxed{\bar{x} = A + h\bar{u}}$

$x - \bar{x} = uh + \cancel{A} - \cancel{A} - h\bar{u}$

$= h(u - \bar{u})$

$\sigma_x^2 = \frac{\Sigma(x-\bar{x})^2}{n}$

$= \frac{\Sigma[h(u-\bar{u})]^2}{n}$

$\sigma_x^2 = h^2\frac{\Sigma(u-\bar{u})^2}{n} = h^2\sigma_u^2$

$\sigma_x = h\sigma_u$

$\Rightarrow \sigma_u = \frac{\sigma_x}{h}$

(c) Show that variance is independent of origin but not of scale

$$\sigma_x^2 = \frac{\sum (x-\bar{x})^2}{n}$$

Proof: $u = \frac{x-A}{h}$

$\Rightarrow hu = x - A$

$\Rightarrow x - A = hu$

$\Rightarrow x = hu + A$

$\Rightarrow \sum x = \sum hu + \sum A$

$\Rightarrow \sum x = h\sum u + nA$

$\Rightarrow \frac{\sum x}{n} = h\frac{\sum u}{n} + A$

$\Rightarrow \bar{x} = h\bar{u} + A$

$\Rightarrow x - \bar{x} = hu + \cancel{A} - h\bar{u} - \cancel{A}$

$= h(u - \bar{u})$

$\sigma_x^2 = \frac{\sum (x-\bar{x})^2}{n}$

$= \frac{\sum [h(u-\bar{u})]^2}{n}$

$\sigma_x^2 = \frac{h^2 \sum (u-\bar{u})^2}{n} = h^2\sigma_u^2$

$\sigma_x = h\sigma_u$

$\Rightarrow \sigma_u = \frac{\sigma_x}{h}$

2018

1 (c)

Profits (in lac)	50-60	60-70	70-80	80-90	90-100
No of Shops	18	32	24	16	10

* Mean $= a + \frac{\Sigma f_i u_i}{N} \times h$

* median $= L + \frac{n/2 - f_c}{f_m} \times h$

* mode $= L + \frac{f_1}{f_1 + f_2} \times h$

mode $= 60 + \frac{32-18}{(32-18)+(32-24)} \times 10$

$= 60 + \frac{14}{14+8} \times 10$

$= 66.36$

Class interval	frequency f_i	midvalue x_i	$u_i = \frac{x_i - a}{h}$	$f_i u_i$	Comulated frequency
50-60	18	55	-1	-18	18
60-70	32	65	0	0	50
70-80	24	75	1	24	74
80-90	16	85	2	32	90
90-100	10	95	3	30	100

Mean $\bar{x} = a + \frac{\Sigma f_i u_i}{N} \times h$

$= 65 + \frac{68}{100} \times 10$

$= 71.8$

median $= L + \frac{n/2 - f_c}{f_m} \times h$

$= 60 + \frac{100/2 - 18}{32} \times 10$

$= 70.$

1.

(a) What is central tendency? Write down the measures of central tendency?

* In statistics, a central tendency is a central or typical value for a probability distribution.

There are three type of central tendency.

1. Mean
2. Median
3. Mode.

CT

(b)

* For non-zero positive observations show that

$\boxed{AM \geq GM \geq HM}$ (Notations are usual)

$$AM = \frac{(a_1 + a_2 + a_3 \cdots a_n)}{n}$$

$$GM = (a_1 a_2 a_3 \cdots a_n)^{1/n}$$

$$HM = \frac{n}{\frac{1}{a_1} + \frac{1}{a_2} + \frac{1}{a_3} \cdots \frac{1}{a_n}}$$

$$AM = \frac{a+b}{2}$$

$$GM = \sqrt{ab}$$

$$HM = \frac{2ab}{a+b}$$

$$AM \times HM = \frac{a+b}{2} \times \frac{2ab}{a+b}$$

$$= ab$$

$$= (GM)^2 = \sqrt{ab}$$

$$= ab$$

$$AM \times HM = (GM)^2$$

$AM \geq GM$

$\Rightarrow$ if $a = 2$, $b = 3$ non-zero positive

AM	GM	HM
$\frac{a+b}{2}$	$\sqrt{a.b}$	$\frac{2}{\frac{1}{2}+\frac{1}{3}}$
$= \frac{5}{2}$	$= \sqrt{2 \times 3}$	
$= 2.5 \geq$	$= 2.49 \geq$	$= 2.4$

$\boxed{1.c}$

C.I	Midvalue x_i	Cumulated frequency f_i	Cumulated frequency
50-60	55	18	18
60-70	65	32	50
70-80	75	24	74
80-90	85	16	90
90-100	95	10	100

$$\bar{x} = a + \frac{\Sigma x_i f_i}{N} \times h$$

3

(a) Define probability, probability function and probability density function?

probability: Probability is the chance that some event will haven.

pobability function:

P function is a discreate random variable that gives the probability that the outcome associated with that variable will occur.

3 (b)

(b) State and prove Bay's theorm:

* The probability of two events A and B happing, $p(A \cap B)$. Is the probability A, $p(A)$. Times the probability of B given that A has occured, $p(B/A)$

$$p(A \cap B) = p(A) \cdot p(B/A)$$

On the other hand, the probability of A and B is also equal to the probability of B times the probability of A given B.

$$p(A \cap B) = p(B) \cdot p(A/B)$$

$$\therefore \quad p(A) \cdot p(B/A) = p(B) \cdot p(A/B)$$

thus: $$p(A \backslash B) = \frac{p(A) \times p(B \backslash A)}{p(B)}$$

This equation konwn as Bayes theorm is the basis of statistical inference

(c) Two unbiased dice are to tossed simultaneously. What is the probability of getting a total of point 6 or even numbers from both the dice?

	1	2	3	4	5	6
1	11	12	13	14	15	16
2	21	22	23	24	25	26
3	31	32	33	34	35	36
4	41	42	43	44	45	46
5	51	52	53	54	55	56
6	61	62	63	64	65	66

Total element $= 6 \times 6 = 36$

total of point 6 = 11 times.

total of even number = 18

$\cancel{\frac{6}{36} + \frac{18}{36}} = \cancel{\frac{6+18}{36}} = \cancel{\frac{24}{36}} = \cancel{\frac{2}{3}}$

$= \frac{2}{3}$

(2,3)
(3,5)

P(A) Total 9 = 22, 42, 62, 64, 44, 24, 26, 46, 66

P(B) total 6 = (1,5), (4,2), (3,3), (5,1), (2,4)

point 6 or even number

$= P(A) + P(B) - P(A \cap B)$

$= \frac{9}{36} + \frac{5}{36} - \frac{2}{36} = \frac{1}{3}$

2017

1 (c) Calculate the mean and variance of the first n natural numbers.

$\bar{x} = \left(\frac{n+1}{2}\right)$

Variance: $\sigma^2 = \sum_{i=1}^{n} \frac{(x_i - \bar{x})^2}{N}$

$$= \sum_{i=1}^{n} \frac{\left(x_i - \left(\frac{n+1}{2}\right)\right)^2}{N}$$

$$= \frac{1}{N}\sum_{i=1}^{n} x^2 - \frac{1}{n}\sum_{i=1}^{n} 2.x\left(\frac{n+1}{n}\right) + \sum_{i=1}^{n} \frac{(n+1)^2}{4n}$$

$$= \frac{1}{N}\,\frac{n(n+1)(2n+1)}{6} - \frac{(n+1)}{n}\,\frac{n(n+1)}{2} + \frac{(n+1)^2}{4n}\times n$$

$$= \frac{(n+1)(2n+1)}{6} - \frac{(n+1)^2}{4}$$

$$= (n+1)\left(\frac{4n+2-3n-3}{12}\right)$$

$$= \frac{(n+1)(n-1)}{12}$$

$$= \frac{n^2-1}{12}$$

$$\frac{\sum x_i^2}{n} - (\bar{x})^2$$

$$= \frac{n(n+1)(2n+1)}{6n} - \frac{(n+1)^2}{4}$$

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

$$= \frac{(n+1)}{2}\,\frac{4n+2-3n-3}{6}$$

$$= \frac{n^2-1}{12}$$

$$\text{Mean} = \frac{\text{sum of all observations}}{\text{Number of observations}}$$

$$= \frac{\frac{n(n+1)}{2}}{n}$$

$$= \frac{(n+1)}{2}$$

2

(b) <u>State and prove the additive law of probability for two events.</u>

If A and B are any two elements of a random experiment and P is a probability function then the probability of happening of at least one of the events is defined as

$$P(A \cup B) = P(A) + P(B) - P(A \cap B).$$

Now we have to prove the addition theorem of probability.

<u>Proof:</u> From the set theory we know that

$$n(A \cup B) = n(A) + n(B) - n(A \cap B).$$

Suppose $n(S)$ denote the total number of the possible events of random experiment. and then dividing both left ~~and~~ hand side and right hand side of the above equation we get

$$\frac{n(A\cup B)}{n(S)} = \frac{n(A)}{n(S)} + \frac{n(B)}{n(S)} - \frac{n(A\cap B)}{n(S)}$$

$P(X) = \frac{n(X)}{n(S)}$. By applying this we can write.

$$P(A\cup B) = P(A) + P(B) - P(A\cap B)$$

<u>Note:</u> if two events A and B are mutually exclusive, then $A\cap B = 0$, then we mathmatically write $P(A\cup B) = P(A) + P(B)$.

2 (c) three events A_1, A_2, A_3 are mutually exclusive and their union is the sample space S.

If $P(A_1) = \frac{3}{2} P(A_2)$, then $P(A_1)$, $P(A_2)$ and $P(A_3)$

$$\frac{P(A_1)}{P(A_2)} = \frac{3}{2} \qquad \frac{P(A_2)}{P(A_3)} = 2$$

$$\frac{P(A_1)}{P(A_2)} \times \frac{P(A_2)}{P(A_3)} = \frac{3}{2} \times 2$$

$$\Rightarrow 20 \text{ Ans}$$

(b) $H_0 : \mu = 65$ $\qquad n = 10$

$H_1 : \mu \neq 65$

$$\bar{x} = \frac{66 + 65 + 69 + 70 + 69 + 68 + 71 + 63 + 64 + 6}{10}$$

$$= 67.3$$

$$\sigma = \sqrt{\frac{\Sigma(x_i - \bar{x})^2}{n-1}}$$

$$= \sqrt{\frac{1.69 + 5.29 + 2.89 + 7.29 + 2.89 + 0.49 + 13.69 + 18.49 + 10.89 + 0.49}{9}}$$

$$= 2.668$$

$$Z = \frac{\bar{x} - \mu}{\sigma/\sqrt{n}}$$

$$= \frac{67.3 - 65}{2.668/\sqrt{10}}$$

$$= 2.72 > 2.26$$

Rejection Region: Reject H_0 if $Z > 2.26$

Comment: $Z = 2.72$ falls in the rejection region and H_0 is rejected at 5% level of significance.

(b) If x is distributed as binomial, show that <u>mean $\geq$ variance</u>

(*) Let x is distributed

parameters n and p.

Then mean $= np$

Variance $= npq$

Mean − Variance $= np - npq$

$= np(1-q) = np^2$

Mean − Variance > 0.

Mean $>$ Variance

---

19 | Men sample size $= 50$ n_1

" mean $= 756$ $\bar{x}_1$

" standard deviation $= 35$ $\sigma_1 =$

Women: sample size $= 50$ n_2

" mean $= 762$ $\bar{y}_2$

" standard deviation $= 30$ σ_2

$H_0 : \bar{y}_1 = \bar{y}_2$

$H_1 : \bar{y}_1 \neq \bar{y}_2$

$$z = \frac{(\bar{x}_1 - \bar{x}_2) - (\bar{y}_1 - \bar{y}_2)}{\sqrt{\frac{\sigma_1^2}{n_1} + \frac{\sigma_2^2}{n_2}}}$$

$$= \frac{(756 - 762)}{\sqrt{\frac{(35)^2}{50} + \frac{(30)^2}{50}}} = \frac{-6}{\sqrt{\frac{1225}{50} + \frac{900}{50}}} = -0.920 > -1.96$$ <u>accepted</u> region at 5% level of significance.

3 (a) Define binomial distribution?

The binomial distribution is the discrete probability distribution that gives only two possible results in an experiment, either success or failure.

3 (c) Let three unbaised coins are tossed at a time. What is the probability that (i) no head (ii) at least two heads will apper.

* total no of probabilites $= 2^3 = 8$

S = (HHH, TTT, HHT, THH, THT, HTH, TTH, HTT)

(i) no head: probability $= \frac{1}{8}$

(ii) at least two heads will apper: $= \frac{3}{8}$

Part II B

2020

ⓐ Define Null hypothesis and alternative hypothesis?

Null hypothesis: predicts there's no relation between the 2 variables.

Alternative hypothesis: predicts there's relation between the 2 variables.

2017 distinguish between correlation and regression?

Correlation	Regression
Relationship between two are more variables	① Average Relationship between two are more variables
X and Y are random Variable	② X is a the random variable and Y is the fixed variable
It gives limited information after verifying the Relationship between variables	③ It is used for the prediction of one value relation to the other given value
The range of relationship lies between and +1	④ Regression value is an absolute figure

(5) It studies the linear relationship between the variable	(5) It studies the linear and non-linear relationship between the variables.
(6) Relationship	(6) One affects the other
(7) Variable move together	(7) Cause and effect
(8) x and y can be interchanged	(8) x and y cannot be interchanged.
(9) Data represented in single point	(9) Data represented by line.

(b) Show that correlation coefficient is independent of change of origin and scale of the variable?

$r_{xy} = r_{uv}$

for $\frac{cov(xy)}{\sqrt{V(x)\,V(y)}}$

$$r_{xy} = \frac{E[(x-\bar{x})(y-\bar{y})]}{\sqrt{E(x-\bar{x})^2}\sqrt{E(y-\bar{y})^2}}$$

Suppose: $u = \frac{x-a}{h}$ $v = \frac{y-b}{k}$

Where a, b, h and K are constant

$x = uh + a \qquad y = vk + a$

$\bar{x} = \bar{u}h + a \qquad \bar{y} = \bar{v}k + a$

$$r_{xy} = \frac{E(x-\bar{x})(y-\bar{y})}{\sqrt{E(x-\bar{x})^2}\sqrt{E(y-\bar{y})^2}}$$

$$= \frac{E(uh+a-\bar{u}h-a)(vk+a-\bar{v}k-a)}{\sqrt{E(uh+a-\bar{u}h-a)^2}\sqrt{(vk+a-\bar{v}k-a)^2}}$$

$$= \frac{E(uh-\bar{u}h)\,E(vk-\bar{v}k)}{\sqrt{E(uh-\bar{u}h)^2}\,E\sqrt{(vk-\bar{v}k)^2}}$$

$$= \frac{hk\,E(u-\bar{u})\,E(v-\bar{v})}{hk\sqrt{(u-\bar{u})^2}\sqrt{(v-\bar{v})^2}}$$

$$= \frac{E[(u-\bar{u})\,E(v-\bar{v})]}{\sqrt{(u-\bar{u})^2}\sqrt{(v-\bar{v})^2}}$$ (Show that)

$= r_{uv}$

(6) Calculate the rank correlation coefficient from the following data on hourly sales (x) and expenses (y) of 10 stores

x	50	50	55	60	65	65	65	60	60	60
y	11	13	14	16	16	15	15	14	13	13

solution:

We know Rank Correlation

$$\rho = 1 - \frac{6\Sigma d^2}{n(n^2-1)}$$

But when repeated value present:

$$\rho = 1 - \frac{6(\Sigma d^2 + F)}{n(n^2-1)}$$

$$F = \frac{m(m^2-1)}{12}$$

x	y	Rx	Ry	dx-Ry	d^2
50	11	9.5	10	-0.5	0.25
50	13	9.5	8	1.5	2.25
55	14	8	5.5	2.5	6.25
60	16	5.5	1.5	4	16
65	16	2	1.5	0.5	0.25
65	15	2	3.5	-1.5	2.25
65	15	2	3.5	-1.5	2.25
60	14	5.5	5.5	0	0
60	13	5.5	8	-2.5	6.25
60	13	5.5	8	-2.5	6.25

$\Sigma d^2 = 42$

$$F = F_{50} + F_{60} + F_{65} + F_{16} + F_{13} + F_{15}$$

$$= \frac{2(2^2-1)}{12} + \frac{4(4^2-1)}{12} + \frac{3(3^2-1)}{12} + \frac{2(2^2-1)}{12} + \frac{3(3^2-1)}{12} + \frac{2(2^2-1)}{12}$$

$$= 0.5 + 5 + 2 + 0.5 + 2 + 0.5$$

$$= 10.5$$

$$\rho = 1 - \frac{6 \times (42 + 0.5)}{10(10^2 - 1)}$$

$$= 0.68$$ (Ans:)

2016

5(a) What is correlation? Write down few of its properties-

Correlation: A measure of the strength of a linear relationship between two quantitative variables (heights weight)

7 properties of correlation:

(I) Mean / Median / Mode

(II) Independent / Dependent Variable

(III) Deviation

(IV) Correlation

(V) Sampling.

(VI) Distributions.

(VI) Probability.

5 (b) Derive the correlation co-efficient and interpret it?

$$r = \frac{S_{xy}}{S_x \cdot S_y} = \frac{S_{xy}}{\sqrt{S_x^2 \cdot S_y^2}}$$

$$S_{xy} = \frac{\Sigma(x-\bar{x})(y-\bar{y})}{n}$$

$$S_x^2 = \frac{\Sigma(x-\bar{x})^2}{n}$$

$$S_y^2 = \frac{\Sigma(y-\bar{y})^2}{n}$$

$$\bar{x} = \frac{\Sigma x}{n}$$

$$\bar{y} = \frac{\Sigma y}{n}$$

$$r = \frac{\frac{\Sigma(x-\bar{x})(y-\bar{y})}{n}}{\sqrt{\frac{\Sigma(x-\bar{x})^2}{n}}\sqrt{\frac{\Sigma(y-\bar{y})^2}{n}}}$$

$$= \frac{\Sigma(x-\bar{x})(y-\bar{y})}{\sqrt{\Sigma(x-\bar{x})^2 \Sigma(y-\bar{y})^2}}$$

$$= \frac{\Sigma xy - n\bar{x}\bar{y}}{\sqrt{\Sigma(x^2 - n\bar{x}^2)\Sigma(\Sigma y^2 - n\bar{y}^2)}}$$

$$= \frac{\Sigma xy - n\left(\frac{\Sigma x}{n}\right)\left(\frac{\Sigma y}{n}\right)}{\sqrt{\Sigma x^2 - n\left(\frac{\Sigma x}{n}\right)^2 \left(\Sigma y^2 - n\left(\frac{\Sigma y}{n}\right)^2\right)}}$$

$$= \frac{\Sigma xy - \frac{(\Sigma x)(\Sigma y)}{n}}{\sqrt{\left(\Sigma x^2 - \frac{\Sigma x^2}{n}\right)\left(\Sigma y^2 - \frac{\Sigma y^2}{n}\right)}}$$

$$= \frac{\frac{n\Sigma xy - \Sigma x \Sigma y}{\cancel{n}}}{\sqrt{\frac{n\{\Sigma x^2 - (\Sigma x)^2\}}{\cancel{n}} \cdot \frac{\{\Sigma y^2 - (\Sigma y)^2\}}{\cancel{n}}}}$$

$$= \frac{n\Sigma xy - \Sigma x \cdot \Sigma y}{\sqrt{\left(n\{\Sigma x^2 - (\Sigma x)^2\}\right)\left(n\{\Sigma y^2 - (\Sigma y)^2\}\right)}}$$

A correlation coefficient is a numerical measure of some type of correlation, meaning a statistical relationship between two variables. The variables may be two columns of a given data set of observations.

(c) If x and y are independent then show that they are uncorrelated?

* x & y two independent variable:

$$cov(x,y) = E(x-\bar{x})(y-\bar{y})$$

$$= E\{xy - x\bar{y} - \bar{x}y + \bar{x}\bar{y})$$

$$= E(xy) - \bar{y}E(x) - \bar{x}E(y) + \bar{x}\bar{y}$$

$$= \bar{x}\cdot\bar{y} - \bar{y}\cdot\bar{x} - \bar{x}\cdot\bar{y} + \bar{x}\bar{y}$$

$$= 0$$

$$\rho_{xy} = \frac{cov(x,y)}{\sigma_x \sigma_y} = \frac{0}{\sigma_x \sigma_y} = 0$$ (Show that)

6

ⓐ **parameter:**

parameter is a number describing a whole population, while a statistic is a number describing a sample (sample mean).

ⓘⓘ **level of significance:**

The level of significance, also denoted as alpha, α, is the probability of rejecting the null hypothesis when it is true.

ⓑ Prices of shares in (Tk) of a company on the different days in a month were found to be 66, 65, 69, 70, 69, 68, 71, 63, 64 and 68. Test whether the mean price of shares in the month is 65 or not (critical value at 5% level of significance is 2.26)

$$\bar{x} = \frac{(66+65+69+70+69+68+71+63+64+68)}{10} = 67.3$$

$H_0 : \mu_0 = 2.65$

$H_1 : \mu \neq 65$

$n = 10$

$n \ngeq 30$

So it is T-test case

$$S = \sqrt{\frac{\Sigma(x-\bar{x})^2}{n-1}} = \sqrt{\frac{64.1}{9}} = 2.65$$

$$T = \frac{\bar{x} - \mu_0}{s/\sqrt{n}}$$

$$= \frac{67.3 - 65}{2.66/\sqrt{10}}$$

$$= 2.72$$

2017

4(a) correlation coefficient: measures the strength of the relationship between x and y.

$$r = \frac{S_{xy}}{\sqrt{S_{xx} \cdot \sqrt{S_{yy}}}}$$

$$S_{xx} = \Sigma x^2 - \frac{(\Sigma x)^2}{N}$$

$$S_{yy} = \Sigma y^2 - \frac{(\Sigma y)^2}{N}$$

$$S_{xy} = \Sigma xy - \frac{\Sigma x \cdot \Sigma y}{n}$$

(b) partial coefficient: measures the strength of a relationship between two variables, while controlling for the effect of one or more other variables.

$$r_{AB} = \frac{N\Sigma XAB - \Sigma XA \cdot XB}{\sqrt{N\Sigma XA^2 - (\Sigma XA)^2 - N\Sigma XB^2 - (\Sigma XB)^2}}$$

$$r_p = r_{AB.C} = \frac{r_{AB} - r_{AC} \cdot r_{BC}}{\sqrt{(1 - r_{AC}^2)(1 - r_{BC}^2)}}$$

Multiple correlation :

The multiple correlation coefficient denoting a correlation of one variable with multiple other variables.

$$R_{A.BC} = \sqrt{\frac{r_{AB}^2 + r_{AC}^2 - 2r_{AB}.r_{AC}.r_{BC}}{1 - r_{BC}^2}}$$

(c) Find the value of the correlation coefficient from the following table :

(x) Age	Glucose level (Y)	xy	x^2	y^2
43	99	4257	1849	9801
21	65	1365	441	4225
25	79	1975	625	6241
42	75	3150	1764	5625
57	87	4959	3249	7569
59	81	4779	3481	6561
$\Sigma x = 247$	$\Sigma y = 486$	$\Sigma xy = 20485$	11409	40022

$$S_{xy} = \frac{\Sigma xy - \Sigma x . \Sigma y}{n} = 20485 - \frac{247 \times 486}{6} = 478$$

$$S_{xx} = \Sigma x^2 - \frac{(\Sigma x)^2}{N} = 11409 - \frac{(247)^2}{6} = 1240.83$$

$$S_{yy} = \Sigma y^2 - \frac{(\Sigma y)^2}{N} = 40022 - \frac{(486)^2}{6} = 656$$

$$r = \frac{S_{xy}}{\sqrt{S_{xx} . S_{yy}}} = \frac{478}{\sqrt{1240.83 \times 656}} = 0.5298$$

2019

(c) $r_{xy} = \frac{Cov(x,y)}{\sigma_x \cdot \sigma_y}$

$Cov(x,y) = \Sigma xy - \frac{\Sigma x \cdot \Sigma y}{n}$

$\sigma_x = \sqrt{\Sigma x^2 - \frac{(\Sigma x)^2}{N}}$

$\sigma_y = \sqrt{\Sigma y^2 - \frac{(\Sigma y)^2}{N}}$

x	y	xy	x^2	y^2
48	99	4752	9801	9801
21	65	1365	441	4225
25	79	1975	625	6241
42	75	3150	1764	5625
57	87	4959	3249	7569
59	81	4779	3481	6561
$\Sigma x = 247$	$\Sigma y = 486$	$\Sigma xy = 20485$	11409	40022

$Cov(x,y) = 20485 - \frac{247 \times 486}{6} = 478$

$\sigma_x = \sqrt{11409 - \frac{(247)^2}{6}} = 35.225$

$\sigma_y = \sqrt{40022 - \frac{(486)^2}{6}} = 25.61$

$r_{xy} = \frac{478}{35.225 \times 25.61}$

$r_{xy} = 0.5298$

CT-1

1. Establish the relationship among AM, GM, HM for the case of two non-zero positive numbers and verify it for the following two positive integers 2 and 4

$$AM = \frac{a+b}{2}$$

$$GM = \sqrt{a+b}$$

$$HM = \frac{2}{\frac{1}{a}+\frac{1}{b}}$$

$$= \frac{2ab}{a+b}$$

$$AM - GM = \frac{a+b}{2} - \sqrt{a+b}$$

$$= \frac{a+b-2\sqrt{a+b}}{2}$$

$$= \frac{(\sqrt{a})+(\sqrt{b})^2 - 2\sqrt{a}\sqrt{b}}{2}$$

$$AM - GM = \frac{(\sqrt{a}-\sqrt{b})^2}{2} \geq 0$$

$$\Rightarrow AM - GM \geq 0$$

$$\Rightarrow AM \geq GM$$

$$GM - HM = \sqrt{ab} - \frac{2ab}{a+b}$$

$$= \sqrt{ab}\left(1 - \frac{2\sqrt{ab}}{a+b}\right)$$

$$= \sqrt{ab}\left(\frac{a+b-2\sqrt{ab}}{a+b}\right)$$

$$= \frac{\sqrt{ab}}{a+b}(\sqrt{a}-\sqrt{b})^2 \geq 0$$

$$GM - HM \geq 0 \Rightarrow GM \geq HM$$

$AM \geq GM \geq HM$

alphabetical order

Again $(AM \times HM) - (GM)^2$

$$= \left(\frac{a+b}{2} \times \frac{2ab}{a+b}\right) - \left(\sqrt{ab}\right)^2$$

$$= ab - ab$$

$$= 0$$

$AM \times HM \geq GM$ (Proved)

Verify $a = 2$, $b = 4$

$AM = \frac{a+b}{2} = \frac{2+4}{2} = 3$

$GM = \sqrt{ab} = \sqrt{2 \times 4} = 2.82$

$HM = \frac{2ab}{a+b} = \frac{2 \times 2 \times 4}{2+4} = 2.67$

$AM > GM > HM$

b) Show that sum of deviations of a set of observations ~~of a set of~~ from their respective arithmetic mean is zero

$$\Sigma x = x_1 + x_2 + x_3 \cdots x_n$$

$$\bar{x} = \frac{\Sigma x}{n}$$

sum of deviations mean:

$$MD(\bar{x}) = \frac{\Sigma(x_i - \bar{x})}{n} = \frac{1}{n}\left[\Sigma x_i - n\bar{x}\right]$$

$$= \frac{\Sigma x_i}{n} - \bar{x}$$

$$= \bar{x} - \bar{x}$$

$$= 0$$

$$MD(\bar{x}) = \frac{\Sigma|x_i - \bar{x}|}{n}$$

$$= \frac{1}{n}\Sigma x_i - n.\bar{x}$$

$$= \frac{\Sigma x_i}{n} - \frac{n.\bar{x}}{n}$$

$$= \bar{x} - \bar{x}$$

$$= 0$$

(c) Frequency distribution:

A representation, either in a graphical or tabular format that displays the number of observations within a given interval. (f_i)

Heighest value = 100

Lowest value = 51

Range = (100 − 51) = 49

Now $h = 10$,

Number of row in table $\frac{49}{10} = 4.9 \approx 5$

class Interval	Mid value x_i	Freq f_i	Cumulate freq	$f_i x_i$
50 – 60	55	8	8	440
60 - 70	65	8	16	520
70 – 80	75	2	18	150
80 – 90	85	1	19	85
90 – 100	95	1	20	95

Mean

$$\frac{\Sigma x_i f_i}{N}$$

$$= \frac{1290}{20}$$

$$= 64.5$$

Mode:

$$= L + \frac{f_1}{f_1 + f_2} \times h$$

$$= 60 + \frac{(8-8)}{(8-8)+(8-2)} \times 10$$

$$= 60 \text{ (Ans)}$$

$$=$$

Median:

$$= L + \frac{(N/2 - f_c)}{f_m} \times h$$

$$= 60 + (10-8) \times 10/8$$

$$= 62.5 \text{ (Ans:)}$$

Arithmetic mean is not an ideal measure for the data set because 400 is the out layer and for this value the mean is effected.

(a) Random variable:

A variable whose value is unknown or a function that assigns values to each of an experiment's outcomes.

How does it differ from variable? Give some examples:

A variable is a symbol that represents some quantity. x is the variable whose value is unknown and we are trying to find its value. After evaluating (মূল্যায়ন). A random variable is different from the variable in algebra as it has whole set of values and it can take any of those randomly. or A random variable is a value that follows some - probability distribution.

2

For a discrete random variable X, show that

(i) $E(aX+b) = aE(X)+b$ (ii) $V(aX+b) = a^2V(X)$

We know:

$E(aX+b)$

$= \sum_x (ax+b)\, p(x)$

$= \sum_x ax \cdot p(x) + \sum_x b \cdot p(x)$

$= a \sum_x x p(x) + b \sum_x p(x)$

$= aE(X) + b \cdot 1$

$= aE(X)+b$ (Ans)

(ii) We know:

$V(aX+b)$

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

$= E(a^2X^2+b^2+2abX) - (aE(X)+b)(aE(X)+b)$

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

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

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

$= a^2V(X)$. (Ans)

(e) The pdf of a random variable X is given below:

x	-2	-1	0	1	2
$p(x)$	$1/8$	$2/8$	$3/8$	$1/8$	$1/8$

compute $E(x) = ?$

$V(x) = ?$

$$E(x) = \Sigma x \cdot p(x) = x_1 p(x_1) + x_2 p(x_2) + x_3 p(x_3) + x_4 p(x_4) + x_5 p(x_5) = -0.125$$

$$(E(x^2)) = \Sigma x^2 p(x)$$

$$= x_1^2 p(x) + x_2^2 p(x_2) + x_3^2 p(x_3) + x_4^2 p(x_4) + x_5^2 p(x_5) = 1.375$$

$$V(x) = E(x^2) - (E(x))^2$$

$$= \cancel{(-0.125) - 1.375}$$

$$= 1.375 - (-0.125)^2$$

$$= 1.35$$

is discreate prob-bility distribution that gives my two possible results in an experiment either success or failure

ⓐ Define binomial distribution? with some application

Binomial distribution is a probability distribution used in statistics that summarize the likelihood that a value will take one of two independent values under a given set of parameters or assumptions. (It is used when there are exactly two mutually exclusive outcomes of a trial.)

for example:

When tossing a coin, the probability of obtaining a head is 0.5. If there are 50 trials, the expected value of the number of heads is 25 (50×0.5). The binomial distribution is used in statistics as a building block for dichotomous variables such as the likelihood that either candidate A or B will emerge in position 1 in the midterm exams.

(b) Compute mean and variance of binomial distribution. Show that mean is always greater than variance.

The binomial distribution is:

$$p(X=x) = {}^nC_x\, p^x \cdot q^{n-x} \qquad \boxed{x = 0, 1, 2, \ldots n}$$

Mean of $x = \mu =$ $E(x) = \sum x \cdot p(x)$

$$= \sum_0^n x \cdot {}^nC_x\, p^x q^{n-x}$$

$a^x b^y$

$$= \sum_1^n x \frac{n!}{x!(n-x)!} p^x \cdot q^{n-x}$$

$= (ab)^{x+y}$

$$= \sum_{x=1}^n \frac{x \cdot n(n-1)!}{x(x-1)!(n-x)!} p^x \cdot q^{n-x}$$

$$= n \sum_{x=1}^n \frac{(n-1)!}{(x-1)!(n-x)!} p^x \cdot q^{n-x}$$

$$= np \sum_{x=1}^n \frac{(n-1)!}{(x-1)!(n-x)!} p^{x-1} q^{n-x}$$

$n \cdot p^{x-1}$

$$= np\,(p+q)^{n-1}$$

$p+q=1$ By Binomial expansion

$$\boxed{\text{Mean} = np}$$

Variance $\sigma^2 = E(x^2) - (E(x))^2$

$= E(x(x-1)+x) - (np)^2$

$= E(x(x-1)) + E(x) - n^2p^2$

$= \sum_{x=0}^{n} x(x-1)\left(\frac{n!}{x!(n-x)!} p^x q^{n-x}\right) + np - n^2p^2$

$= \sum_{x=2}^{n} \frac{n!}{(x-2)!(n-x)!} p^x \cdot q^{n-x} + np - n^2p^2$

$= n(n-1) \sum_{x=2}^{n} \frac{(n-2)!}{(x-2)!(n-x)!} p^x q^{n-x} + np - n^2p^2$

$= n(n-1)p^2 \sum_{x=2}^{n} \frac{(n-2)!}{(x-2)!(n-x)!} p^{x-2} + np - n^2p^2$

$= n(n-1)p^2 (p+q)^{n-2} + np - n^2p^2$

$p+q=1$

$\Rightarrow 1-p=q$

$= \cancel{n^2p^2} - np^2 + np - \cancel{n^2p^2}$

$= np(1-p)$

$= npq$

Mean $= npq$

Variance $= npq$

$\frac{\text{Mean}}{\text{Variance}} = \frac{\cancel{np}}{\cancel{np}q} = \frac{1}{q} > 1$

Thus $\boxed{\text{Mean} > \text{Variance}}$ always.

(c) In an Independent 7 trials of a binomial distribution with mean and variance are obtained 3 and 2 respectively. Find the probability that the variate the values (i) less than or equal to 2 (ii) greather than or equal to 5

mean $= np = 3$

variance $= npq = 2$

$\Rightarrow q = 2/3$

$p = 1 - 2/3 = 1/3$

(i) $p(x \leq 2) = \sum_{0}^{2} {}^{n}c_{x}\, p^{x} q^{n-x}$

$= {}^{7}c_{0}\, p^{0} q^{7} + {}^{7}c_{1}\, p^{1} q^{6} + {}^{7}c_{2}\, p^{2} q^{5}$

$= {}^{7}c_{0} (1/3)^{0} \cdot (2/3)^{7} + {}^{7}c_{1} (1/3) \cdot (2/3)^{6} + {}^{7}c_{2} (1/3)^{2} (2/3)^{5}$

$= 128 \cdot \frac{128}{2187} + \frac{448}{2187} + \frac{224}{9375} = 0.2873$ (Ans)

(iii) $p(X \geq 5)$

$$= \sum_{x=5}^{7} {}^{n}c_{x}\, p^{x} q^{n-x}$$

$$= {}^{7}c_{5}\, p^{5} q^{2} + {}^{7}c_{6}\, p^{6} q^{1} + {}^{7}c_{7}\, p^{7} q$$

$$= {}^{7}c_{5} (1/3)^{5} (2/3)^{2} + {}^{7}c_{6} (1/3)^{6} (2/3)^{1} + {}^{7}c_{7} (1/3)^{7} \cdot (2/3)^{0}$$

$$= \frac{28}{729} + \frac{14}{2187} + \frac{1}{2187}$$

$$= 0.045$$

বাকী প্রশ্ন:

2020

(b) Describe the test procedure for comparing two means?

when variance are different

$$Z = \frac{(\bar{x}_1 - \bar{x}_2) - (\mu_1 - \mu_2)}{\sqrt{\frac{\sigma_1^2}{n} + \frac{\sigma_2^2}{n}}}$$

$\sigma_1^2 = \sigma_2^2 = \sigma^2$

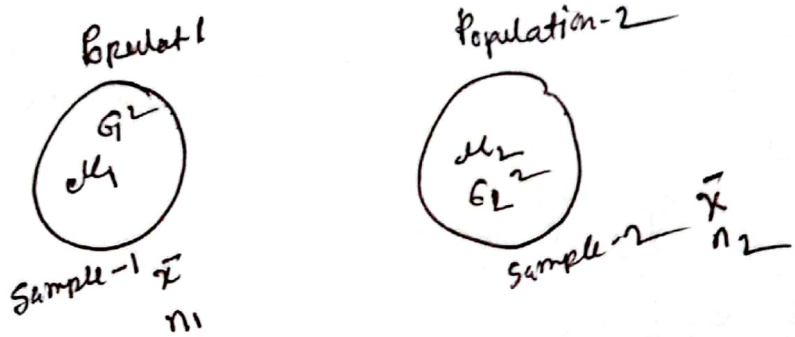
when variance are equal

$$t = \frac{(\bar{x}_1 - \bar{x}_2) - (\mu_1 - \mu_2)}{\sqrt{s^2\left(\frac{1}{n_1} + \frac{1}{n_2}\right)}}$$

$$s^2 = \frac{(n_1 - 1)s_1^2 + (n_2 - 1)s_2^2}{n_1 + n_2 - 2}$$

$H_0 : \mu_1 = \mu_2$

$H_1 : \mu_1 \neq \mu_2$



2015 (6)

(a) Write down the goal and different steps of hypothesis testing?

goal: The purpose/goal of hypothesis testing is to test whether the null hypothesis can be rejected or accepted.

different steps of hypothesis testing:

(i) Identify Population and Sample

(ii) State the hypothesis

(iii) Check conditions

(iv) Calculate the test statistic.

(5) Calculate the p-value

(6) State the conclusion

* 2018 3(b)

Prove and state Bay's theorm:

Bay's theorm:

Let $E_1, E_2, E_3 \cdots E_n$ be a set of events associated with a sample sapce S. where all the events $E_1, E_2 \cdots E_n$ have nonzero probability of occurrence and they form a partition of S.

Let A be any event associated with S. then according to Bayes theorm,

$$P(E_i/A) = \frac{P(E_i) \cdot P(A/E_i)}{P(E_1) \cdot P(A|E_1) + P(E_2) \cdot P(A|E_2) \cdots P(E_n) \cdot P(A/E_n)}$$

$i = 1, 2, 3 \cdots n$

$k = 1, 2, 3 \cdots n$

$$= \frac{P(E_i)\, P(A|E_i)}{\sum_{k=1}^{n} P(E_k)\, P(A|E_k)}$$

Proof:

According to the conditional probability formula:

$$P(E_i|A) = \frac{P(E \cap A)}{P(A)} \quad \text{—— (I)}$$

Using the multiplication rule of probability,

$$P(E \cap A) = P(E_i) \cdot P(A|E_i) \quad \text{—— (II)}$$

Using total probability theorem

$$P(A) = \sum_{k=1}^{n} P(E_k)\, P(A/E_k) \quad \text{--- (III)}$$

putting the values from equations ② and ③ in equation 1, we get

$$P(E/A) = \frac{P(E) \cdot P(A/E)}{\sum_{k=1}^{n} P(E_k) \cdot P(A/E_k)}$$

ⓑ For two observations, show that standard deviation is the half of the range?

Let $x_1 = a$, $n = 2$
$x_2 = b$

$$\bar{x} = \frac{x_1 + x_2}{n} = \frac{a+b}{2}$$

$$S.D = \sqrt{\sum_{i=1}^{n} \frac{(x_i - \bar{x})^2}{n}}$$

$$= \sqrt{\sum_{i=1}^{2} \frac{(x_i - \bar{x})^2}{2}}$$

$$= \sqrt{\frac{(x_1 - \bar{x})^2 + (x_2 - \bar{x})^2}{2}}$$

$$= \sqrt{\frac{\frac{R^2}{4} + \frac{R^2}{4}}{2}}$$

$$= \sqrt{\frac{R^2}{4}} = \frac{R}{2}$$

$$= \frac{\text{Range}}{2}$$

(Proved)

$$(x_1 - \bar{x})^2$$
$$= \left(a - \frac{a+b}{2}\right)^2$$
$$= \left(\frac{2a - a - b}{2}\right)^2 = \frac{(a-b)^2}{2}$$
$$= \frac{R^2}{4}$$

$$(x_2 - \bar{x})^2$$
$$= \left(b - \frac{a+b}{2}\right)^2$$
$$= \left(\frac{2b - a - b}{2}\right)^2$$
$$= \frac{(b-a)^2}{4}$$
$$= \frac{R^2}{4}$$

Probability distributions: provides the possible values of random variable and their corresponding probabilities. A probability distribution can be in the form of a table, graph or mathmatical formula.

Random variable: numerical measure of the outcome from a probability experiment. so its value is determined by chance. Random variable are denoted X, Y, Z etc.

Random Variable

Discreate

↓

a random variable either infinite number of possible values or a countable number of possible values.

Continous.

is a random variable that has an infinite number of possible values that is not countable

Poisson

$$p(x) = \sum_{x=0}^{n} \frac{e^{-\mu}\mu^{x}}{x!}$$

$$= e^{-\mu} \sum \frac{\mu^{x}}{x!}$$

$$= e^{-\mu}\left(\frac{\mu^0}{0!} + \frac{\mu^1}{1!} + \frac{\mu^2}{2!} + \frac{\mu^3}{3!} + \frac{\mu^4}{4!} + \cdots\right)$$

$$= e^{-\mu} \cdot e^{\mu}$$

$$= 1$$

Normal distribution

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

Let $z = \frac{x-\mu}{\sigma}$

$\sigma dz = dx$

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

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

$$= \frac{1}{\sqrt{2\pi}} \sqrt{2\pi}$$

$$= 1$$

Ans:

Physics department 2019

3 (b) if $\bar{x}$ and s denote the mean and standard deviation of a set of n non-negative quantities then show that $\boxed{\bar{x}\sqrt{n-1} \geq s}$

Proof:

let $x_1, x_2, x_3 \cdots x_n$ be non-negative observations

$$\bar{x} = \frac{\sum_{i=1}^{n} x_i}{n} \Rightarrow \sum_{i=1}^{n} x_i = n\bar{x} \rightarrow (i)$$

$$\therefore s = \sqrt{\frac{\sum_{i=1}^{n} x_i^2}{n} - \left(\frac{\sum x_i}{n}\right)^2}$$

$$\Rightarrow s^2 = \frac{1}{n}\left[\sum x_i^2 - \frac{\left(\sum x_i\right)^2}{n}\right]$$

$$\Rightarrow ns^2 = \sum x_i^2 - \frac{\left(\sum x_i\right)^2}{n} \quad \text{--- (ii)}$$

Now $\left(\sum_{i=1}^{n} x_i\right)^2 = \sum_{i=1}^{n} x_i^2 + \sum_{i \neq j}^{n} x_i x_j$

$$\Rightarrow \left(\sum_{i=1}^{n} x_i\right)^2 \geq \sum_{i=1}^{n} x_i^2$$

$$\Rightarrow \left(\sum_{i=1}^{n} x_i\right)^2 - \frac{\left(\sum_{i=1}^{n} x_i\right)^2}{n} \geq \sum_{i=1}^{n} x_i^2 - \frac{\left(\sum_{i=1}^{n} x_i\right)^2}{n}$$

$$\Rightarrow \left(\sum_{i=1}^{n} x_i\right)^2 \left\{1 - \frac{1}{n}\right\} \geq ns^2 \quad [\text{From (ii)}]$$

$$\Rightarrow \frac{n-1}{n}(n\bar{x})^2 \geq ns^2$$

$$\Rightarrow n(n-1)(\bar{x})^2 \geq ns^2$$

$$\Rightarrow \bar{x}\sqrt{n-1} \geq s$$

Imp 2017 ~~Normal distribution — Mean/Variance~~

① Three events A_1, A_2, A_3 are mutually exclusive and their union is the sample space S. If $P(A_1) = \frac{3}{2}P(A_2)$, $P(A_2) = 2P(A_3)$ then find $P(A_1)$, $P(A_2)$ and $P(A_3)$

$P(A_1) = x$

$P(A_1) = \frac{3}{2}P(A_2)$

$\Rightarrow P(A_2) = \frac{2}{3}x$

$P(A_2) = 2P(A_3)$

$\Rightarrow \frac{2}{3}x = 2P(A_3)$

$\Rightarrow \frac{x}{3} = P(A_3)$

$\Rightarrow$

Note: Since, A, B, C are mutually exclusive and exhaustive events associated with a random experiment

$A_1 \cup A_2 \cup A_3 = S$

$\Rightarrow P(A_1 \cup A_2 \cup A_3) = P(S)$

$\Rightarrow P(A_1) + P(A_2) + P(A_3) = 1$

$\Rightarrow x + \frac{2x}{3} + \frac{x}{3} = 1$

$\Rightarrow \frac{3x + 2x + x}{3} = 1$

$\Rightarrow 6x = 3$

$\Rightarrow 2x = 1$

$\Rightarrow x = 1/2$

$P(A_1) = \frac{1}{2}$

$P(A_2) = \frac{2}{3} \times \frac{1}{2} = \frac{1}{3}$

$P(A_3) = \frac{1}{3} \times \frac{1}{2} = \frac{1}{6}$

3 (c) three unbiased coins are tossed

(i) no head —

(ii) at least two heads —

$p(S) = (HHT, HTT, THT, HTH, TTT, HHH, THH, \cancel{HHH}, TTH)$

(i) no head = TTT

probability of no head = $\frac{1}{8}$

(ii) at least two heads $p(x \geq 2)$

$= \frac{3}{8} + \frac{1}{8} = \frac{1}{2}$

1.

(a) Define coefficient of correlation. Define the situation when coefficient of correlation $r = 0.95$ and $r = -0.95$

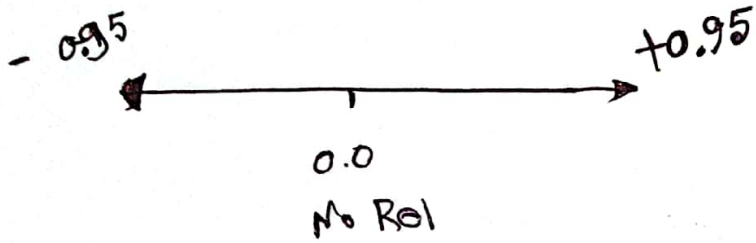
* correlation coefficient measures the strength of the relationship between x and y.

$$r_{xy} = \frac{cov(x,y)}{\sigma_x \cdot \sigma_y}$$

$$\sigma_x^2 = \Sigma x^2 - \frac{(\Sigma x)^2}{n}$$

$$\sigma_y = \Sigma y^2 - \frac{(\Sigma y)^2}{n}$$

$$cov(x,y) = \Sigma xy - \frac{(\Sigma x)(\Sigma y)}{n}$$



$\boxed{r \geq 0.95}$ a correlation of $r = 0.95$ suggests a strong positive association between two variables.

whereas a correlation of $r = -0.95$ suggests a weak negative association,

CT Question-sched-2

2 (c) A machine is designed to produce insulting washers for electrical devices of average thickness of 0.025 cm. A random sample of 10 washers was found to have an average thickness of 0.029 cm with standard deviation of 0.002 cm. Test whether the thickness equal to 0.025 at 5% level where the tabulated value of "t" is 2.262.

$\bar{x} = 0.029$ cm

$n = 10$

$s = 0.002$ cm

$H_0: \mu = 0.025$ cm

$H_1: \mu \neq 0.025$ cm

$$t = \frac{\bar{x} - \mu}{s/\sqrt{n}} = \frac{0.029 - 0.025}{0.002/\sqrt{10}}$$

$$= \frac{4 \times 10^{-3}}{6.324 \times 10^{-4}} = 6.32$$

Rejection Region:

Rejected H_0 if $z > 2.262$

Comment: $z = 6.32$ falls in rejection region and H_0 is rejected at the 5% level of significance.

1 (c) In a trivariate case, find the partial correlation coefficient $r_{23.1}$ when $r_{13} = 0.86$, $r_{32} = 0.48$ and $r_{12} = 0.75$ and multiple correlation coefficient $R_{1.23}$.

partial correlation:

$$r_p = r_{AB.C} = \frac{r_{AB} - r_{AC} \cdot r_{BC}}{\sqrt{(1 - r_{AC}^2)(1 - r_{BC}^2)}}$$

$$r_{23.1} = \frac{r_{23} - r_{31} \cdot r_{21}}{\sqrt{(1 - r_{31}^2)(1 - r_{21}^2)}}$$

$$= \frac{0.48 - 0.86 \times 0.75}{\sqrt{(1 - 0.86^2) \cdot (1 - 0.75^2)}}$$

multiple correlation:

$$R_{A.BC} = \sqrt{\frac{r_{AB}^2 + r_{AC}^2 - 2 r_{AB} r_{AC} r_{BC}}{1 - r_{BC}^2}}$$

$$r_{1.23} = \sqrt{\frac{r_{12}^2 + r_{13}^2 - 2 \cdot r_{12} \cdot r_{13} \cdot r_{23}}{1 - (r_{23})^2}}$$

$$= \sqrt{\frac{(0.75)^2 + (0.86)^2 - 2 \times 0.86 \times 0.48 \times 0.75}{1 - (0.48)^2}}$$

(b) $r_{xy} = 0.80$

$\Sigma xy = 60$

$\sigma_x = 2.5$

$\Sigma y^2 = 90$

$\bar{x} = 0$

$\bar{y} = 0$

$n = ?$

$$r_{xy} = \frac{cov(x,y)}{\sigma_x \sigma_y}$$

$$= \frac{\Sigma xy - \frac{\Sigma x \Sigma y}{n}}{\sqrt{\frac{\Sigma x^2}{n} - \frac{(\Sigma x)^2}{n}} \cdot \sigma_x}$$

$$\sigma_x = \frac{(\Sigma x - \bar{x})^2}{n}$$

$$\Rightarrow 0.80 = \frac{60 - 0}{\sqrt{\frac{(\Sigma x - \bar{x})^2}{n}} \cdot \sigma_x}$$

$$\Rightarrow 0.80 = \frac{60}{\sqrt{\frac{\Sigma y^2}{n}} \times 2.5}$$

$$\Rightarrow 0.80 = \frac{60}{\sqrt{\frac{90}{n}} \times 2.5}$$

$$\Rightarrow 0.80 \times \sqrt{\frac{90}{n}} \times 2.5 = 60$$

$$\Rightarrow \sqrt{\frac{90}{n}} = 30$$

$$\Rightarrow \frac{90}{n} = (30)^2$$

$$\Rightarrow n = \frac{900}{90}$$

$$\Rightarrow n = 10$$

$$\Rightarrow (0.80 \times 2.5)^2 = \frac{60^2}{\frac{\Sigma y^2}{n}}$$

$$\Rightarrow n = \frac{60^2}{(.80 \times 2.5)^2 \times 90}$$

$$= 10$$

2 (a) Define level of significance and critical Region?

The level of significance is the measurment of the statistical significance. It was denoted as alpha α, is the probability of rejecting the null hypothesis when it is true.

A critical region, also known as the rejection region, is a set of values for the test statistic for which the null hypothesis is rejected

* Divide the distribution into rejection and non-rejection region.

- Show that correlation coefficient lies between -1 and $+1$

Proof: let x and y be two variables

σ_x, σ_y

Consider. $\Sigma\left[\frac{x-\bar{x}}{\sigma_x} \pm \frac{y-\bar{y}}{\sigma_y}\right]^2 \geq 0$

$$\Rightarrow \Sigma\left[\left(\frac{x-\bar{x}}{\sigma_x}\right)^2 + \left(\frac{y-\bar{y}}{\sigma_y}\right)^2 \pm 2\left(\frac{x-\bar{x}}{\sigma_x}\right)\cdot\left(\frac{y-\bar{y}}{\sigma_y}\right)\right] \geq 0$$

$$\Rightarrow \frac{1}{\sigma_x^2}\Sigma(x-\bar{x})^2 + \frac{1}{\sigma_y^2}\Sigma(y-\bar{y})^2 \pm 2\frac{(x-\bar{x})(y-\bar{y})}{\sigma_x\cdot\sigma_y} \geq 0$$

Dividing both side by n:

$$\Rightarrow \frac{1}{\sigma_x^2}\frac{\Sigma(x-\bar{x})^2}{n} + \frac{1}{\sigma_y^2}\frac{\Sigma(y-\bar{y})^2}{n} \pm \frac{2}{\sigma_x\cdot\sigma_y}\underline{Cov(x,y)} \geq 0$$

$$\Rightarrow 1 + 1 \pm 2r \geq 0$$

$$\Rightarrow 2 \pm 2r \geq 0$$

$$\Rightarrow 1 \pm r \geq 0$$

Either $1 + r \geq 0$ or $1 - r \geq 0$

$\Rightarrow r \geq -1$ or $1 \geq r$

$\boxed{-1 \leq r \leq 1}$ (প্রমাণিত)