Aparajita Mondal
21/1/19

FOR GOD AND COUNTRY

ST. LAWRENCE HIGH SCHOOL

A JESUIT CHRISTIAN MINORITY INSTITUTION

Pre- Annual Examination– 2019

Sub: Statistics Model Answer **Class: XI** **F. M. 70**

Duration: **Date: 18/01/2019**

Group-A

1. Select the correct alternative. 1x10=10
   a) The relation between two variables x and y is 2x-5y=10 and the quartile deviation of x is 5. Quartile deviation of y is iii) 2
   b) If a variable assumes the values 1,2,3,....10 with frequencies 1,2,3,.....10 respectively, then arithmetic mean is iii) 7
   c) Numerically the measure of skewness interms of quartiles can not exceed i)1

   d) The probability of getting 9 dots with two unbiased dice is iii) 1/9

   e) If A and B are two independent events. P(A)=0.5, P(B)=0.7 then P (A-B)= equals i) 0.15

   f)For any frequency distribution ii) $b_2 > b_1$

   g) Which one of the following is correct? ii) $15 \equiv 27 \pmod 3$

   h)Grades obtained in an examination is i) attribute

   i)For any distribution first central moment is ii) 0

   j)Number of single digit Fermat number is ii) 2

Group-B

2.Answer the following questions. 1x8=8

a) Write down a situation where C.V is the appropriate measures of dispersion.

Ans. If we want to compare the dispersion of height of a group with the dispersion of weights.

b) If the number of educated persons is 65% of the population , then in Pie diagram 234º is needed to mark the sector.

Or

Frequency density is plotted along Y axis in Histogram.

c)Write down Paasche's Price Index formula.

Ans. $I_{01} = \frac{\sum_{i=1}^{n} p1i\ q1i}{\sum_{i=1}^{n} p0i\ q1i}$

d) Write down the sample space of the event when a coin is tossed repeatedly until a tail appears.

Ans.{T, HT,HHT,HHHT,……..}

Or

Define exhaustive events.

Ans. several events are said to be exhaustive if atleast one of them will necessarily occurs whenever the random experiment is performed.

e)Define vital Index.

Ans. ratio of birth to death of a given population during a given period.

Or

Define crude Birth rate.

Ans. simplest measure of fertility is Crude Birth Rate

CBR=B/P x 1000

B= total number of live births occurring in the given region during given period

P=total number of population in a given region during given period.

f) If P(AUB)= 2/3 , then find the value of $P(A^c \cap B^c)$.

Ans. $P(A^c \cap B^c) = P(AUB)^c = 1 - P(AUB) = 1/3$

g) Define real wage.

Ans. Real wage= actual wage/ cost of living index number x 100

Or

What are the tests proposed by Fisher for checking the goodness of an Index number.

Ans. Factor reversal test and Time reversal test.

h)Define secondary data.

Ans. The statistical data which have already been collected by some agency and are compiled from that source by the enquirer for his use are called secondary data.

Or

Define continuous data.

Ans. A quantitative character that can take any value with in its range of variation is termed as continuous variable . Data on this variable is called continuous data.

Group-C

3. Answer the following questions. 2x4=8

a) If A.M = 10 and C.V=50% . Find var(5-2x)

Ans. C.V= S.D/Mean x 100 or, 50= S.D/ 10 x 100 or, S. D=5

Var (5-2x)= $(-2)^2$Var x= 4x 25=100

b) Distinguish between qualitative data and quatitative data.

Ans. A qualitative character that can not be expressed numerically is known as attribute. Data on attribute is called as qualitative data.

On the other hand a quantitave character that can be expressed numerically is known as variable. Data on variable is called quantitative data.

Or

DefineTime series and Cross sectional data.

Ans.Values of a variable recorded for different point of time or interval of time for an individual or for a population are called time series data. In Time series data we may have cross sectional data when each member of a group is classified on two or more characters.

c) Find the Standard deviation 1,3,5,...(2n-1)

Ans. Variane= 1/n $[1^2+3^2+....+$upto n terms] – 1/n [1+3+5+.....+up to n terms]

$= 1/n \frac{n(2n-1)(2n+1)}{3}$ - $[1/n\{n/2(2a+(n-1)d)\}]^2 = 1/3(4n^2-1) - n^2$

$= 1/3(n^2-1)$ therefore SD=$\sqrt{1/3(n2-1)}$

d) Write down the Theorem of total probability

Ans. If $A_1, A_2, A_3....A_k$ are mutually exclusive events then $p(A_1 U A_2 U......U A_k)=$ $P(A_1) + P(A_2) + ...+ P(A_k)$

Or

If P(A) =a and P(B) =b show that P(A/B)≤a/b

Ans. $n(A) \geq n(A \cap B)$ or, $P(A) \geq P(A \cap B)$ or, $P(A)/P(B) \geq P(A \cap B) / P(B)$

or, a/ b ≥ P(A/B)

Group – D

4. Answer the following questions. 3x8=24

a) Describe different parts of a table.

Ans.Table number---Title----stub----caption---body----Foot note---source

b) What are the advantages of Interview method.

Ans.i)It can be used even if the informants are illiterate

ii)scope of cross examining is there

iii)the chance of non response is totally diminished.

Or

What are the disadvantages of mail-questionnaire Method.

Ans.i) There is scope for non-response due to unwillingness

ii) This method is applicable for educated people

iii) There is chance to get back the questionnaire that arew not fully filled in.

c) The variance of 1,2,3..........,n is 24 . Find n.

Ans.Variance = 1/n[1+2+3+...+n] – [1/n(1+2+3+..+n)]^2= 1/6 (n+1) (2n+1) – ¼ (n+1)^2= 1/12 [2(n+1)(2n+1) -3(n+1)^2]= 1/12 (n^2-1)

According to the problem -1/12 (n^2-1)= 24 or, n^2=289 or, n=17

Or

Show that if s^2 be the variance of n given values $x_1, x_2, \ldots\ldots x_n$ of a variable x, then $s^2 = \frac{1}{2n^2} \sum_{i=1}^{n} \sum_{j=1}^{n} (x_i - x_j)^2$

Ans.

$$\sum_{i=1}^{n}\sum_{j=1}^{n}(x_i-x_j)^2 = \sum_{i=1}^{n}\sum_{j=1}^{n}\{(x_i-\bar{x})-(x_j-\bar{x})\}^2$$

$$= \sum_{i=1}^{n}\sum_{j=1}^{n}[(x_i-\bar{x})^2 - 2(x_i-\bar{x})(x_j-\bar{x})+(x_j-\bar{x})^2]$$

$$= \sum_{i=1}^{n}\sum_{j=1}^{n}(x_i-\bar{x})^2 - 2\sum_{i=1}^{n}\sum_{j=1}^{n}(x_i-\bar{x})(x_j-\bar{x}) + \sum_{i=1}^{n}\sum_{j=1}^{n}(x_j-\bar{x})^2$$

$$= n\sum_{i=1}^{n}(x_i-\bar{x})^2 + n\sum_{j=1}^{n}(x_j-\bar{x})^2$$

$$= n^2s^2 + n^2s^2$$

$$= 2n^2s^2$$

$$s^2 = \frac{1}{2n^2}\sum_{i=1}^{n}\sum_{j=1}^{n}(x_i-x_j)^2$$

d)Prove that for any frequency distribution $b_2 \geq 1$

Ans. Suppose a variable x takes n values

$x_1, x_2, \dots, x_n$ with mean $\bar{x} = \sum_{i=1}^{n} x_i/n$

In Cauchy-Schwarz inequality

$$\left(\sum_{i=1}^{n} a_i^2\right)\left(\sum_{i=1}^{n} b_i^2\right) \geq \left(\sum_{i=1}^{n} a_i b_i\right)^2,$$

if we put $a_i = (x_i - \bar{x})^2$ and $b_i = 1$, for each i, then we get

$$\left\{\sum_{i=1}^{n} (x_i - \bar{x})^4\right\}\left(\sum_{i=1}^{n} 1\right) \geq \left\{\sum_{i=1}^{n} (x_i - \bar{x})^2\right\}^2$$

or $n\sum_{i=1}^{n} (x_i - \bar{x})^4 \geq \left\{\sum_{i=1}^{n} (x_i - \bar{x})^2\right\}^2$ or $\frac{1}{n}\sum_{i=1}^{n} (x_i - \bar{x})^4 \geq \left\{\frac{1}{n}\sum_{i=1}^{n} (x_i - \bar{x})^2\right\}^2$

or $m_4 \geq m_2^2$ or $\frac{m_4}{m_2^2} \geq 1$, [assuming $m_2 = s^2 \neq 0$]

or $b_2 \geq 1$.

e)write down different measures of Skewness.

Ans. $g_1 = m_3/s^3$ provided $s \neq 0$, sk= 1/s (mean- mode) , sk = 1/s[3(mean-median)]

$Sk = \frac{1}{Q3-Q1}(Q_3-2Q_2+Q_1)$

Or

The lower and the upper quartiles of a distribution are 14.6 and 25.2 respectively and the coefficient of skewness is 0.5. Find the median of the distribution.

Ans. Q_1 =14.6 , Q_3=25.2 ,SK= 0.5 now $SK = \frac{1}{Q3-Q1}(Q_3-2Q_2+Q_1)$

Or, $0.5 = \frac{1}{25.2-14.6}(25.2-2Q_2+14.6)$ or, 10.6 x0.5= 39.8 – $2Q_2$ or, 39.8 – $2Q_2$ = 5.3

Or, Q_2= ½ (39.8-5.3)= 17.25

f)State and prove Bonferroni's inequality.

Ans. (II) Since probability of an event cannot exceed 1,

$P(A_1 \cup A_2) \leq 1$

or $P(A_1) + P(A_2) - P(A_1 \cap A_2) \leq 1$

$\therefore$ $P(A_1 \cap A_2) \geq P(A_1) + P(A_2) - 1$

Using this result, we have

$$\begin{aligned} P(A_1 \cap A_2 \cap A_3) &= P[(A_1 \cap A_2) \cap A_3] \\ &\geq P(A_1 \cap A_2) + P(A_3) - 1 \\ &\geq P(A_1) + P(A_2) - 1 + P(A_3) - 1 \\ &= P(A_1) + P(A_2) + P(A_3) - 2. \end{aligned}$$

Or

If the events A,B and C are independent such that P(A) =1/2 , P(B)=1/3 and P(C)= ¼ then find the probability that at least one of the three events occurs.

Ans. $P(A \cup B \cup C) = P(A^c \cap B^c \cap C^c) = 1 - P(A^c).P(B^c).P(C^c) = 1 - \frac{1}{2} \times 2/3 \times 3/4 = 3/4$

g) Why is Fisher's index number called 'ideal'?

proof :

$$P_{01} = \sqrt{\frac{\sum p_1 q_0}{\sum p_0 q_0} \times \frac{\sum p_1 q_1}{\sum p_0 q_1}}$$

Change p to q and q to p

$$q_{01} = \sqrt{\frac{\sum q_1 p_0}{\sum q_0 p_0} \times \frac{\sum q_1 p_1}{\sum q_0 p_1}}$$

$$P_{01} \times q_{01} = \sqrt{\frac{\sum p_1 q_0}{\sum p_0 q_0} \times \frac{\sum p_1 q_1}{\sum p_0 q_1} \times \frac{\sum q_1 p_0}{\sum q_0 p_0} \times \frac{\sum q_1 p_1}{\sum q_0 p_1}}$$

$$= \frac{\sum p_1 q_1}{\sum p_0 q_0}$$

∴ The factor reversal test is satisfied by the Fisher's Ideal Index.

Again, $I_{01} = \sqrt{\frac{\sum p_1 q_0}{\sum p_0 q_0} \times \frac{\sum p_1 q_1}{\sum p_0 q_1}}$ and $I_{10} = \sqrt{\frac{\sum p_0 q_1}{\sum p_1 q_1} \times \frac{\sum p_0 q_0}{\sum p_1 q_0}}$

∴ $I_{01} I_{10} = 1$

∴ The Time Reversal test is satisfied by the Fisher's Ideal Index.

Ans.

h) what are the merits and demerits of Crude Death Rate?

Ans. Merits-i) easy nto interpret ii) easy to compute

Demerits-i) Here composition of the population by sex, age, race occupation, locality is not considered ii) because of the first defect CDR can not be used for comparing the mortility over regions.

Group-E

5. Answer the following questions. **5x4= 20**

a)Derive Lagrange's interpolation formula.

Ans.

Let $y = f(x)$ represent a function which assumes the values $y_0, y_1, y_2, \ldots, y_n$ orresponding to $(n+1)$ values $x_0, x_1, x_2, \ldots, x_n$ of the argument x. Here the nterpolation formula may be considered as a polynomial of degree n.

Suppose the polynomial is taken in the form

$$\phi(x) = c_0(x-x_1)(x-x_2)\ldots(x-x_n) + c_1(x-x_0)(x-x_2)\ldots(x-x_n) + c_2(x-x_0)(x-x_1)(x-x_3)\ldots(x-x_n) + \ldots + c_r(x-x_0)(x-x_1)\ldots(x-x_{r-1})(x-x_{r+1})\ldots(x-x_n) + \ldots + c_n(x-x_0)(x-x_1)\ldots(x-x_{n-1}) \quad \ldots (i)$$

The constants $c_0, c_1, c_2, \ldots, c_n$ are obtained such that

$$\phi(x_0) = y_0, \phi(x_1) = y_1, \phi(x_2) = y_2, \ldots, \phi(x_n) = y_n.$$

Putting $x = x_0$ in (i), $\phi(x_0) = c_0(x_0-x_1)(x_0-x_2)\ldots(x_0-x_n)$ and hence

$c_0(x_0-x_1)(x_0-x_2)\ldots(x_0-x_n) = y_0$

or, $c_0 = \dfrac{y_0}{(x_0-x_1)(x_0-x_2)\cdots(x_0-x_n)}$

Let $y = f(x)$ represent a function which assumes the values $y_0, y_1, y_2, \ldots, y_n$ orresponding to $(n+1)$ values $x_0, x_1, x_2, \ldots, x_n$ of the argument x. Here the nterpolation formula may be considered as a polynomial of degree n.

Suppose the polynomial is taken in the form

$$\phi(x) = c_0(x-x_1)(x-x_2)\ldots(x-x_n) + c_1(x-x_0)(x-x_2)\ldots(x-x_n) + c_2(x-x_0)(x-x_1)(x-x_3)\ldots(x-x_n) + \ldots + c_r(x-x_0)(x-x_1)\ldots(x-x_{r-1})(x-x_{r+1})\ldots(x-x_n) + \ldots + c_n(x-x_0)(x-x_1)\ldots(x-x_{n-1}) \quad \ldots (i)$$

The constants $c_0, c_1, c_2, \ldots, c_n$ are obtained such that

$$\phi(x_0) = y_0, \phi(x_1) = y_1, \phi(x_2) = y_2, \ldots, \phi(x_n) = y_n.$$

Putting $x = x_0$ in (i), $\phi(x_0) = c_0(x_0-x_1)(x_0-x_2)\ldots(x_0-x_n)$ and hence

$c_0(x_0-x_1)(x_0-x_2)\ldots(x_0-x_n) = y_0$

or, $c_0 = \dfrac{y_0}{(x_0-x_1)(x_0-x_2)\cdots(x_0-x_n)}$

b) Prove that Standard Deviation is the least root-mean-square deviation.

Ans. Square deviation of x about an arbitrary value c is $\sqrt{1/n\sum(x_i - c)^2}$

$$\sum(x_i - c)^2 = \sum[(x_i - mean\ of\ x) - (mean\ of\ x - c)]^2$$

$= \sum(x_i - mean\ of\ x)^2 + 2(mean\ of\ x - c)\sum(x_i - mean\ of x)$ +n(mean ofx-c)^2

$= \sum(x_i - mean\ of\ x)^2$ + n(mean ofx-c)^2

$therefore\ \sum(x_i - c)^2 \leq \sum(x_i - mean\ of\ x)^2$

Therefore $\sqrt{1/n\sum(x_i - mean\ of\ x)^2} \leq \sqrt{1/n\sum(x_i - c)^2}$

Or

A variable assumes the values 1,2,.....,n with corresponding frequencies 1,2,....n,. Calculate variance of the variable.

Ans. Mean of x= $1/n\sum x_i f_i = \frac{1^1+2^2 \ldots +n^2}{1+2+3+..+n} = \frac{n(n+1)(2n+1)}{6} X \frac{2}{n(n+1)} = \frac{2n+1}{3}$

$1/n\sum x_i^2 f_i = \frac{1^3+2^3+3^3+\cdots..n^3}{1+2+3+\cdots+n} = (\frac{n(n+1)}{2})^2 X \frac{2}{n(n+1)}$=n(n+1)/2

Variance= $1/n\sum x_i^2 f_i$ - (mean of x)^2=n(n+1)/2 – (2n+1)^2/9= $\frac{9n(n+1)-2(4n^2+4n+1)}{18}$

=(n+2)(n-1)/18

c)Describe different parts of a life table.

Ans.

x	l_x	d_x	q_x	L_x	T_x	e_x^o
1	2	3	4	5	6	7

1) x- integral value of age in years
2) l_x- the number of persons out of an assumed number of births l_o who attain exact age x.
3) d_x–the number of persons out of l_x persons attaining precise age x ,who die before reaching age x+1 . $d_x = l_x - l_{x+1}$

4) q_x- the probability that a person of exact age x will die before attaining age x+1 . Hence $q_x = dx / lx$

5) L_x- Total numbers of years lived by the cohort of l_o persons between age x and x+1. Thus we have $L_x = \int_0^1 l_{x+t}\, dt$

Now out of l_x persons at age x , lx+1 persons live one complete year in the age interval;l x –x+1 and the remaining d_x persons who die in that age interval live for varying fractions of a year .Denoting by a_x the average of these fraction we get $L_x = l_{x+1} + a_x d_x = (l_x - d_x) + a_x d_x = l_x - (1-a_x)\, d_x$

If we assume $a_x = 1/2$ then we have $L_x = l_x - 1/2 d_x = ½ (l_x + l_{x+1})$

6) T_x –The total number of years by the cohort after age x or the total life time of l_x persons who reach age x. thus $T_x = \int_0^\infty l_{x+t}\, dt$

7) e_x^o - The average number of years lived by each of the l_x persons after age x. it is called expectation of life at age x.

$$e_x^o = T_x / l_x$$

d) The first of two urns contains 3 white and 2 black balls, and the second contains 3 white, 1 black and 2 red balls. One ball is taken at random from the first urn and is placed in the second urn. Then if one ball is taken at random from the second urn, find the probabilities that i) it is black , ii) it is either red or white.

Ans.Let $A_1$and A_2 be the events that first and second urn are selected respectively and B be the event that the ball taken at random is black .

i) P(B)= P(A_1) P(B/A_1) + P(A_2) P(B/A_2)= 3/5 x 1/7 + 2/5 x 2/7 =1/5

ii) P(R or W)= P(R) + [P(A_1) P(W/A_1) +P(A_2) P(W/A_2)= 2/7 + (3/5x 4/7+ 2/5 x 3/7)= 1/35 (10+12 +6)=4/5

Or

Sate and prove Bayes' Theorem.

ANs. suppose the events A_1, A_2,A_3 ....A_n are exhaustive and mutually exclusive events , none of them has zero probability . Further , let B be an event which too has non zero probability then

$P(A_i/B) = P(A_i)\, P(B/A_i) / \sum P(Aj)P(\frac{B}{Aj})$ this is known as Bayes' theorem.

Since $P(A_i) > 0$ for i=1,2,3,...n and $P(B) > 0$ we have for each i

$P(Ai \cap B) = P(A_i)\, P(B/A_i)$

and also, $P(A_i \cap B) = P(B)\, P(A_i \mid B)$.

Hence, $P(B)\, P(A_i \mid B) = P(A_i)\, P(B \mid A_i)$ or, $P(A_i \mid B) = \dfrac{P(A_i)P(B|A_i)}{P(B)}$

But, as the events $A_1, A_2, \ldots, A_n$ are exhaustive and mutually exclusive,

$$P(B) = \sum_{j=1}^{n} P(A_j \cap B) = \sum_{j=1}^{n} P(A_j) \cdot P(B|A_j)$$

So, $$P(A_i|B) = \frac{P(A_i)P(B|A_i)}{\sum_{j=1}^{n} P(A_j)P(B|A_j)}, \text{ for } i = 1, 2, \ldots, n.$$

It may be said that Bayes' theorem gives the posterior probability of A_i in terms of the prior probabilities $P(A_i)$, $i = 1, 2, \ldots, n$ and the conditional probabilities of B.