Aparajita Mondal
16/01/2020

FOR GOD AND COUNTRY

ST. LAWRENCE HIGH SCHOOL

A JESUIT CHRISTIAN MINORITY INSTITUTION

PRE-ANNUAL EXAMINATION

Subject: STATISTICS **Class: XI** **F. M. 70**

Duration:3 HRS 15 MINS **Date: 15/01/20**

Group-A

1. Write down the correct option. **1x10=10**

i) The first moment about the value 4 of a variable is 2,then its mean is
ans. d) none of these
ii) If p,q ,r are the roots of $x^3+mx-n=0$ then p+q+r is
ans. c)0
iii) If P(A) =$\frac{1}{3}$, P(B)=$\frac{1}{5}$,then the highest possible value of P(A∪B) is
ans. b) $\frac{8}{15}$
iv)Which of the following is correct?
Ans. b) 15≡27 (mod 3)
v) The harmonic mean of the reciprocals of first seven positive integers is
ans. b) $\frac{1}{4}$
vi) Indicate the type of data : 'Blood group of any person'
ans. attribute
vii) Variance of first n natural numbers is
ans. c) $\frac{n^2-1}{12}$
viii) If 5v +2y=17 is the relation between variables y and v , and mean deviation of y about its mean 6 is 5, then mean deviation of v about its mean is
ans. c) 2
ix) If P(A)=0.2 , P(B)=0.4 , P(A∩B)=0.08 then $P(B/A^c)$ equals
ans. 0.4
x) The absolute value of measure of skewness based on quartiles can not exceed
ans. b) 1

Group -B

2. Answer the following questions . **1x8=8**

i) State De Morgan's law for two arbitrary events A and B .
ans $(A\cap B)^c=A^c\cup B^c$
ii) Under what condition the weighted average becomes identical to the simple average?
Ans. when each value has same frequency.

Or

If the mean of the five consecutive positive integers is 18 , find the mean of the highest and least values.

Ans, least =16, highest=20

iii) In a moderately skewed distribution A.M =24 .6 and the median= 26.1 . Find the value of mode.

Ans.mode=29.1

iv) If a coin is tossed repeatedly until a head appears then form the sample space.

Ans.{H,TH,TTH,TTTH......}

v) What is a compound event?

Ans.when an event can not be decomposed into any other elementary events.

Or

State one limitation of the classical definition of probability.

Ans. it uses concept of probability in the definition of probability.

vi) If $P(A/B) = \frac{1}{4}$, then what is the value of $P(A^c/B)$?

Ans. $P(A^c/B)=3/4$

Or

If $P(A_1)=0.2$, $P(A_2)=0.1$ and $P(A_3)=0.3$ and those events are mutually independent ,then what is the value of $P(A_1 \cap A_2 \cap A_3)$?

Ans. $P(A_1 \cap A_2 \cap A_3)=0.006$

vii) If the cost of living index number is 120 and the salary of a worker is Rs 800, what is the real wage?

Ans. Rs666 (approx)

Or

State one important use of price index.

Ans. It is used to measure the change in the retail prices of a group of commodities and services representing consumption level of a group of people.

viii) Define crude rate of natural increase.

Ans. crude rate of natural increase= crude birth rate- crude death rate

Group-C

3 . Answer the following . 2x4=8

i) Show that the sum of deviations of a set of observations about their mean is zero.

Ans. $\sum$ (x-mean of x)=(n x mean of x) -(n x mean of x)=0

ii)Find the standard deviation of 1,3,5,.......(2n-1)

ans. Variane= $1/n\,[1^2+3^2+....+\text{upto n terms}] - 1/n\,[1+3+5+.....+\text{up to n terms}]$

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

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

or

If the relationship between the two variables x and y is 2x+3y=7 and first quartile of x is 9 then find the third quartile.

Ans. $Q_3(y)=7/3 - 2/3 \times 9= -11/3$

iii)Find the geometric mean of n observations a, ar, ar^2,...........ar^{n-1}

Ans. $GM= (a.ar.ar^2..................ar^{n-1})^{1/n}= (a^n)^{1/n}(r\,.\,r^2.r^3.........r^{n-1})^{1/n}= a\,[r^{(1+2+3...+n)}]^{1/n}$

$= ar^{n-1/2}$

iv)Distinguish between discrete variable and continuous variable .

Ans. the variable which can take only isolated values with in the range of variation is known as discrete variable. Data on discrete variable is known as discrete data. On the other hand the variable which can take any value with in the range of variation is known as continuous variable. Data on continuous variable is known as continuous data.

or

How will you draw a pie diagram.
Ans. First find out the % values of the data --------then multiply 3.6 with each value -------then draw the angles inside a circle.

Group –D

4. Answer the following. 3x8=24

i) If a, b, x ,y are all positive then show that (ab+xy)(ax+by)≥4abxy
ans. using the concept AM≥GM we can write $\frac{1}{2}(ab+xy) \geq \sqrt{abxy}$ and $1/2(ax+by) \geq \sqrt{abxy}$
multiply above two inequations we get (ab+xy)(ax+by)≥4abxy

ii) For any distribution prove that $b_2>b_1$

or

express the central moments m_3 and m_4 interms of ordinary moments about an arbitrary origin A.
ans. $m_3 = m'_3(A) - 3\, m'_2(A)\, m'_1(A) + 2\, m'^3_3(A)$ and $m_4 = m'_4(A) - 4\, m'_3(A)\, m'_1(A) + 6\, m'_2(A)\, m'^2_1(A) - 3\, m'^4_1(A)$

iii) Examine whether fisher's index satisfies both the 'Time reversal' and 'Factor reversal test'.

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.

Or

Write down the steps of construction of cost of living index number.
Ans. selection of base period------selection of commodities by judgement sampling------ selection of items-----determination of price index for each group---- weight selection------- then weighted average of the group index has to be found out.

iv) Define Ordinal data and Nominal data with the help of example.

Ans. when data on attribute can be ordered is known as ordinal data. Example- primary education---secondary education---- higher secondary education----graduation
When data on attribute can not be ordered the data is known as nominal data. Example – religion of different people.

v)For three events A_1 ,A_2 and A_3 state and prove Boole's inequality.

Ans. $P(A_1 U A_2)=P(A_1)+P(A_2)- P(A_1 \cap A_2) \leq P(A_1)+P(A_2)$
$P(A_1 U A_2 U A_3)= P\{P(A_1 U\ A_2) U A_3\} \leq P(A_1)+P(A_2) +P(A_3)$

Or

If an integer x is randomly selected from the first 50 positive integers then find the value of $P(x+\frac{96}{x} >50)$

Ans. there are 3 such numbers. therefore $P(x+\frac{96}{x} >50)=3/50$

Vi) The probabilities of solving a problem by three students A,B,C are $\frac{3}{7}$,$\frac{3}{8}$ and $\frac{1}{3}$. If all of them try independently, find the probability that the problem could be solved by one person only.

Ans. Probability that the problem can be solved by one person only=$P[(A_1 \cap A^c_2 \cap A^{C}_3) U (A^{C}_1 \cap A_2 \cap A^{C}_3) U (A^{C}_1 \cap A^c_2 \cap A_3)]$=3/7x5/8 x 2/3 + 4/7x3/8 x 2/3 + 4/7 x 5/8 x 1/3 = 37/84

vii) A variable assumes the values a and b and (n-2) other values all equal to $\frac{a+b}{2}$. Find the standard deviation.

Ans. variance= 1/n [{a-(a+b)/2 }²+ {b -(a+b)/2}²+{(a+b)/2 -(a+b)/2}²(n-2)]= (a-b)²/2n
Sd=$\sqrt{(a-b)^2/2n}$

Or

Show that if S^2 be the variance of n given values $x_1, x_2, x_3 .....x_n$ 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.

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

viii) Write down merits and demerits of mail-questionnaire method.
Ans. merits- less time consuming---cost effective---informant will not feel hesitant to answer. Demerit- possibility of getting half filled questionnaire----proper education is needed

Group-E

5. Answer the following. 5x4=20

i) Derive Newton's forward interpolation formula.

Ans. let y=f(x) which assumes the values $y_0, y_1, y_2,.... y_n$
Corresponding to the values $x_0, x_1, x_2,.... x_n$ having same diff erence h

$\Phi(x)= a_0+a_1(x-x_0)+a_2(x-x_0)(x-x_1)+ a_3(x-x_0)(x-x_1)(x-x_2)+\ldots\ldots+a_n(x-x_0)\ldots(x-x_{n-1})$

Putting $x=x_0, x=x_1, x=x_2$ **we got**
$a_0= y_0, a_1=\Delta y_0/h,\ a_2 =\Delta^2 y_0/2!h$

$\Phi(x)= y_0 +u\Delta y_0 +u(u-1)\Delta^2 y_0/2!+\ldots\ldots+u(u-1)\ldots(u-n+1)\Delta^n y_0/n!$ **where** $u=(x-x_0)/h$

ii) State and Bayes' theorem.

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

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,

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

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.

Or,

A man is known to speak truth 2 out of 3 times. He throws a die and reports that the number obtained is a four. Find the probability that the number obtained is actually a four.

Ans. A be the event that the man reports that the number is four.

E be the event that four occurs and F is the complementary event.

$P(E/A)= P(E)P(A/E)/ P(E)\,P(A/E) + P(F)\,P(A/F)=(1/6\text{x}2/3)\div[(1/6\text{x}2/3) +(5/6\text{x}1/3)]= 2/7$

iii) Describe the structure 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 $L_x= l_{x+1}+a_xd_x= (l_x-d_x) + a_xd_x= l_x- (1-a_x)\, d_x$

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

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.

$$e_x^o=T_x/l_x$$

iv) Suppose two groups of values of a variable x are given below:

$\overline{x_1}$= the mean of first group

$\overline{x_2}$=The mean of 2nd group

S_1^2= the variance of 1st group

S_2^2= the variance 2nd group

n_1= the number of observations in the first group

n_2= the number of observations in the second group

then show that combined variance of ($n_1 + n_2$) observations is

$$S^2=\frac{n_1s_1^2+n_2s_2^2}{n_1+n_2}+\frac{n_1n_2(\overline{x_1}-\overline{x_2})^2}{(n_1+n_2)^2}$$

or

For a set of n values $x_1, x_2, \ldots\ldots x_n$, prove that $\frac{R^2}{2n} \leq S^2 \leq \frac{R^2}{4}$, where S^2 and R are, respectively, variance and range of the values.

$\sum_{i=1}^{n}(x_i - c)^2$ is least when $c = \bar{x}$.

Hence, $\sum_{i=1}^{n}(x_i - \bar{x})^2 \le \sum\left(x_i - \frac{a+b}{2}\right)^2 = \sum_1\left(x_i - \frac{a+b}{2}\right)^2 + \sum_2\left(x_i - \frac{a+b}{2}\right)^2$,

where Σ_1 and Σ_2 include respectively those values of x which are less than or equal to $\frac{a+b}{2}$ and greater than $\frac{a+b}{2}$,

or $\sum_{i=1}^{n}(x_i - \bar{x})^2 \le \sum_1\left(a - \frac{a+b}{2}\right)^2 + \sum_2\left(b - \frac{a+b}{2}\right)^2 = \sum_1\frac{R^2}{4} + \sum_2\frac{R^2}{4} = n\frac{R^2}{4}$

Hence, $\frac{1}{n}\sum_{i=1}^{n}(x_i - \bar{x})^2 \le \frac{R^2}{4}$ i.e., $s^2 \le \frac{R^2}{4}$.

Here the equality sign holds either when all the values are equal or when the variable takes only two distinct values with same frequency.

Again, we have

$ns^2 = \sum_{i=1}^{n}(x_i - \bar{x})^2 = (a - \bar{x})^2 + (b - \bar{x})^2 + \sum_1(x_i - \bar{x})^2$,

where Σ_1 includes all values of x except the minimum and maximum values.

So, $ns^2 \ge (a - \bar{x})^2 + (b - \bar{x})^2$

$= \frac{1}{2}[2(a - \bar{x})^2 + 2(b - \bar{x})^2] = \frac{1}{2}[(a + b - 2\bar{x})^2 + (a - b)^2]$,

$\because 2(p^2 + q^2) = (p+q)^2 + (p-q)^2$.

$\ge \frac{1}{2}(a - b)^2 = \frac{R^2}{2}$.

Hence, $s^2 \ge \frac{R^2}{2n}$.

The equality sign holds either if all the values of the variable are equal or if all the values except the maximum and minimum values are equal to $\frac{a+b}{2}$.

Hence, we have,

$$\frac{R^2}{2n} \le s^2 \le \frac{R^2}{4}.$$

Ans.