Cultures

ST. LAWRENCE HIGH SCHOOL

FOR GOD AND COUNTRY

Pre-annual Examination - 2018

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

Duration: 3 hrs **Date: 17/02/2018**

Group A

1. **Choose the correct alternative.** **1x10=10**

(i) Monthly income of works of a factory is

a) Attribute b) Discrete variable

c) Continuous Variable d) None of these

(ii) The vertical axis in case of an ogive shows

a) Cumulative frequencies b) Absolute frequencies

c) Frequency Densities d) Class boundaries

(iii)Monthly income of works of a factory is

a) 1 b) 2

c) 3 d) None of these

(iv) A train ran at x km per hr from A to B and returned from B to A at y km per hr. The average speed (in km per hour) is

a) $\frac{x+y}{2}$ b) $\sqrt{xy}$

c) $\frac{2xy}{x+y}$ d) none of these

(v) If A.M. and Co-efficient of variation of x are 6 and 50% respectively, then variance of x is

a) 3 b) 6

c) 9 d) None of these

(vi) If A and B independent events and P(A) = 0.5, P(B) = 0.7 then P(A-B) equals

a) 0.15 b) 0.35

c) 0.15 d) None of these

(vii) The probability of having at most one tail in 3 tosses of a fair coin is

a) $\frac{3}{8}$ b) $\frac{1}{2}$

c) $\frac{1}{8}$ d) None of these

(viii) If b_2= 3.25 then the distribution is

a) Platykurtic b) Leptokurtic

c) Mesokurtic d) None of these

ix) The harmonic mean of the reciprocals of first seven positive integers is

a) $\frac{1}{6}$ b) $\frac{1}{4}$

c) 3 d) None of these

x) If the relation between two variables x and y is 2y-6x=6 and mode of x is 21, then mode of y is

a) 19 b) 23

c) 63 d) 66

Group B

1x8=8

2. i) Differentiate between primary and secondary data.

ii) For a symmetrical distribution Q_1=28, Q_3=46. Find the median.

iii) Define frequency density.

OR

Write down different parts of a table.

iv) If $P(A)=\frac{1}{2}$, $P(B)=\frac{1}{4}$, $P(A/B)=\frac{1}{3}$, Find $P(A \cap B)$.

OR

Let A,B,C be three events, Give expression for the event 'None of the events occurs' using set theory notations.

v) State Bonferroni's inequality for two events A and B.

vi) Write down Paasche's price Index Formula.

OR

Write down one demerit of crude birth rate.

vii) State three conditions necessary for the events A,B and C to be pairwise independent.

OR

Define composite event.

viii) State the relation between d_x, l_x and q_x in a life table.

Group C

3. **2x4=8**

i) If A,B and C are three inpedendent events such that $P(A) = \frac{1}{4}$, $P(B) = \frac{1}{3}$, $P(C) = \frac{2}{5}$. Find P(AUBUC).

ii) Distinguish between discrete variable and continuous variable.

OR

How will you draw a pie-diagram?

iii) Find the mean of first n odd positive integers.

iv) Find the GM of n observations a, ar, ar^2,.... ,ar^{n-1}.

OR

If the relation between two variables x and y be 5x+7y=28 and median of y be 3, then find out median of x.

Group D

4. **3X8=24**

i) Prove that standard deviation is the least root-mean-square deviation.

ii) State and prove Theorem of Total Probability.

OR

If the events $E_1, E_2, \ldots\ldots, E_n$ are independent and such that $P(E_i^c) = \frac{i}{1+i}$, i=1,2,3....n. Then find the probability that at least one of the n events occurs.

iii) Write down merits and demerits of crude death rate.

iv) If $P(A) = \frac{1}{2}$, $P(B) = \frac{1}{3}$, and $P(A^c \cup B^c) = \frac{5}{12}$ Find P(A/B)

OR

If $P(AUB)=\frac{7}{12}$, $P(A) = \frac{1}{3}$ for two independent events A and B, then Find the value of P(B).

v) If the standard deviation of the set of numbers 1,2,3,...,k is 2 , then find the value of k.

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

OR

Write any three measures of skewness of a distribution.

vii) Three boxes of the same appearance have the following proportions of black and white balls. Box I – 5 black and 3 White; Box II – 6 Black and 2 White; Box III – 3 black and 5 white. One of the boxes is selected at random and one ball is drawn randomly from it what is the probability that the ball is black?

viii) Write down demerits of personal observation method.

OR

Write down merits of mail questionnaire method.

Group E

5. **5X4=20**

(i) Suppose S and R are respectively the standard deviation and range of n values of a variable x. Then prove that $\frac{R^2}{2n} \leq S^2 \leq \frac{R^2}{4}$.

(ii) State and prove Bayes' Theorem.

OR

The probabilities of X,Y and Z becoming the principal of a certain college are respectively 0.3, 0.5 and 0.2. The probabilities that student aid fund will be introduced in the courage if X,Y,Z become principal, are 0.4, 0.6 and 0.1 respectively. Given that 'Student-aid-fund' has been introduced, find the probability that Y has been appointed as the principal.

(iii) Why is Fisher's index number called 'ideal'?

OR

Write down the steps of constructing cost of Living Index Number.

(iv) (a) Define Class Width. 2

(b) What do you mean by ordinal and nominal data? Explian with example. 3

Aparajita Mondal.
19/2/18.

FOR GOD AND COUNTRY

ST. LAWRENCE HIGH SCHOOL

Pre-annual Examination – 2018

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

Duration: 3 hrs **MODEL ANSWER** **Date-17/02/2018**

Group A

1. **Choose the correct alternative.** **1x10=10**

(i) Monthly income of works of a factory is

c) Continuous Variable

(ii) The vertical axis in case of an ogive shows

a) Cumulative frequencies

(iii)Numerically the measure of skewness in terms of quartiles can not exceed

a) 1

(iv) A train ran at x km per hr from A to B and returned from B to A at y km per hr. The average speed (in km per hour) is

c)$\frac{2xy}{x+y}$

(v) If A.M. and Co-efficient of variation of x are 6 and 50% respectively, then variance of x is

c) 9

(vi) If A and B independent events and P(A) = 0.5, P(B) = 0.7 then P(A-B) Equals

c) 0.15

(vii) The probability of having at most one tail in 3 tosses of a fair coin is

b) $\frac{1}{2}$

(viii) If b_2= 3.25 then the distribution is

b) Leptokurtic

ix) The harmonic mean of the reciprocals of first seven positive integers is

b) $\frac{1}{4}$

x) If the relation between two variables x and y is 2y-6x=6 and mode of x is 21, then mode of y is

d) 66

Group B **1x8=8**

2. i) Differentiate between primary and secondary data.

Ans. Primary data are directly collected from the field and secondary data are the data which have already been collected by some agency and used by others for different purpose.

ii) For a symmetrical distribution Q_1=28, Q_3=46. Find the median.

Ans.28+46/2= 74/2=37

iii) Define frequency density.

Ans.Freqency of the class/ class width

OR

Write down different parts of a table.

Ans.table number,title,stub,caption,body,footnote,source

iv) If $P(A)=\frac{1}{2}$, $P(B)=\frac{1}{4}$, $P(A/B)=\frac{1}{3}$, Find $P(A\cap B)$.

ans. $P(A\cap B) = P(B)\,P(A/B) = \frac{1}{4} \times 1/3 = 1/12$

OR

Let A,B,C be three events, Give expression for the event 'None of the events occurs' using set theory notations.

Ans. $A^c \cap B^c \cap C^c$

v) State Bonferroni's inequality for two events A and B.

ans. $P(A\cap B) \geq p(A) + p(B) - 1$

vi) Write down Paasche's price Index Formula.

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

OR

Write down one demerit of crude birth rate.

Ans. CBR does not take into consideration the age –composition and the sex-composition of the population. Though these are the most important factors affecting fertility.

vii) State three conditions necessary for the events A,B and C to be pairwise independent.

Ans. $P(A \cap B) = p(A).P(B)$, $P(B \cap c) = P(B).P(c)$, $P(C \cap A) = P(C).P(A)$

OR

Define composite event.

Ans. Events which can be further subdivided into smaller events are called composite events.

viii) State the relation between d_x, l_x and q_x in a life table.

Ans. $Q_x = d_x / l_x$

Group C

3. **2x4=8**

i) If A,B and C are three inpedendent events such that $P(A) = \frac{1}{4}$, $P(B) = \frac{1}{3}$, $P(C) = \frac{2}{5}$.

Find P(AUBUC).

Ans. $P(AUBUC) = P(A^c \cap B^c \cap C^c)^c = 1 - P(A^c \cap B^c \cap C^c) = 1 - P(A^c).P(B^c).P(C^c) = 1 - 3/4 \times 2/3 \times 3/5 = 7/10$

ii) Distinguish between discrete variable and continuous variable.

Ans. A quantitative character that can take certain isolated values only in its range of variation is called discrete variable. Where as a quantitative character that can take any value with in its range of variation.

OR

How will you draw a pie-diagram?

Ans.Convert data into percentage then convert it into degrees by multiplying with 3.6 and then according to the measure in degree mark inside the circle.

iii) Find the mean of first n odd positive integers.

Ans:d=2,a=1 S_n=n/2 [2x1 + (n-1) 2]=n^2 therefore mean n^2/n=n

iv) Find the GM of n observations a, ar, ar^2,.... ,ar^{n-1}.

Ans:Gm= $(a.ar.ar^2ar^3.......ar^{n-1})^{1/n}$

=(a.a.a.a.....nterms. $r^{1+2+3...n-1})^{1/n}$

$=[a^n r^{n(n-1)/2}]^{1/n}= a\ r^{n-1/2}$

OR

If the relation between two variables x and y be 5x+7y=28 and median of y be 3, then find out median of x.

Ans: Me (x)=28/5 – 7/5 x 3 = 7/5 =1.4

Group D

4. **3X8=24**

i) Prove that standard deviation is the least root-mean-square deviation.

$$\text{Now } \sum_{i=1}^{n}(x_i-c)^2 = \sum_{i=1}^{n}\{(x_i-\bar{x})+(\bar{x}-c)\}^2$$

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

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

$$\therefore \sum_{i=1}^{n}(x_i-c)^2 \geq \sum_{i=1}^{n}(x_i-\bar{x})^2, \quad \text{since } n(\bar{x}-c)^2 \geq 0,$$

Ans:

$$\text{or } \boxed{\sqrt{\frac{1}{n}\sum_{i=1}^{n}(x_i-\bar{x})^2} \leq \sqrt{\frac{1}{n}\sum_{i=1}^{n}(x_i-c)^2}}$$

ii) State and prove Theorem of Total Probability.

Ans.

If $A_1, A_2, \ldots, A_k$ are mutually exclusive events, then,

$$P(A_1 \cup A_2 \cup \ldots \cup A_k) = P(A_1) + P(A_2) + \ldots + P(A_k).$$

This is known as ***theorem of total probability***.

Proof **:** Let the total number of elementary events of the random experiment be n, where n is finite and the elementary events are equally likely. Also let $n(A_i)$ of them be favourable to the event A_i, $i = 1(1)k$. Since the events are mutually exclusive, the elementary events that are favourable to any of the events, are entirely different from those favourable to others. So the number of elementary events that are favourable to A_1 or A_2 or ... or A_k is $n(A_1) + n(A_2) + \ldots + n(A_k)$.

Hence,

$$P(A_1 \cup A_2 \cup \ldots \cup A_k) = \frac{n(A_1)+n(A_2)+\ldots+n(A_k)}{n}$$

$$= \frac{n(A_1)}{n}+\frac{n(A_2)}{n}+\ldots+\frac{n(A_k)}{n}$$

$$= P(A_1) + P(A_2) + \ldots + P(A_k).$$

:

OR

If the events $E_1, E_2, \ldots, E_n$ are independent and such that $P(E_i^c) = \frac{i}{1+i}$, i=1,2,3....n.

Then find the probability that at least one of the n events occurs.

iii) Write down merits and demerits of crude death rate.

Ans.

Merits

(i) It is easy to interpret; it gives the number of deaths, on the average, per thousand people.

(ii) It is simple to compute, as it requires only total number of deaths and total population size.

(iii) If all the persons are taken to be equally exposed to the risk of dying during an interval, from this cause or that, CDR is a true probability rate.

Demerits

(i) CDR does not take into account the fact that mortality of a region depends on the composition of its population by age, sex, race, occupation, locality of dwelling, etc.

(ii) Because of the first defect, CDR cannot be used for comparing the mortality over regions.

iv) If $P(A) = \frac{1}{2}$, $P(B) = \frac{2}{3}$, and $P(A^c \cup B^c) = \frac{5}{12}$ Find P(A/B)

ans :$P(A/B)=\frac{7}{12} \div \frac{2}{3} = 7/8$

OR

If $P(A\cup B)=\frac{7}{12}$, $P(A) = \frac{1}{3}$ for two independent events A and B, then Find the value of P(B).

Ans

$$P(A\cup B) = P(A^c\cap B^c)^c$$
$$= 1 - P(A^c \cap B^c)$$
$$= 1 - P(A^c)\cdot P(B^c)$$

or. $\frac{7}{12} = 1 - \frac{2}{3}P(B^c)$

or $1 - \frac{7}{12} = \frac{2}{3}P(B^c)$.

or $\frac{2}{3}P(B^c) = \frac{5}{12}$

or $P(B^c) = \frac{5}{\cancel{12}_4} \times \frac{\cancel{3}}{2} = \frac{5}{8}$

$\therefore P(B) = \frac{3}{8}$

v) If the standard deviation of the set of numbers 1,2,3,...,k is 2 , then find the value of k.

Ans.

$$SD^2 = \frac{1^2+2^2+3^2+ k^2}{k} - \left[\frac{1+2+3+\dots+k}{k}\right]^2$$
$$= \frac{\cancel{k}(k+1)(2k+1)}{6\cancel{k}} - \left[\frac{\cancel{k}(k+1)}{2\cancel{k}}\right]^2$$
$$= (k+1)\left[\frac{4k+2-3k-3}{12}\right]$$
$$= \frac{(k+1)(k-1)}{12} = \frac{k^2-1}{12}.$$

$\therefore \frac{k^2-1}{12} = 4$

or $k^2-1 = 48$

$\therefore k^2 = 49$ $\qquad \therefore k = \pm 7$

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

Ans.

(A) For any frequency distribution, $b_2 \geq 1$.

Suppose a variable x takes n values

$x_1, x_2, \ldots, x_n$ 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 $b_2 \geq 1$.

The equality sign holds when the variable takes only two distinct values with same frequency.

OR

Write any three measures of skewness of a distribution.

Ans: Pearson's co-efficient of skewness:

Sk= mean – mode / s.d.

$g_1 = m_3/s^3$

sk = 3(mean – median)/s.d.

vii) Three boxes of the same appearance have the following proportions of black and white balls. Box I – 5 black and 3 White; Box II – 6 Black and 2 White; Box III – 3 black and 5 white. One of the boxes is selected at random and one ball is drawn randomly from it what is the probability that the ball is black?

Ans.

Let A_1, A_2 and A_3 denote the selection of box I, box II and box III respectively, and B denote the event of drawing a black ball.

Here the events A_1, A_2 and A_3 are exhaustive and mutually exclusive. As the box is selected at random,

$$P(A_1) = P(A_2) = P(A_3) = \frac{1}{3}.$$

Also $P(B \mid A_1) = \frac{5}{8}$, $P(B \mid A_2) = \frac{6}{8}$ and $P(B \mid A_3) = \frac{3}{8}$.

(i) The required probability is

$$P(B) = \sum_{i=1}^{3} P(A_i).P(B|A_i) = \frac{1}{3}\times\frac{5}{8}+\frac{1}{3}\times\frac{6}{8}+\frac{1}{3}\times\frac{3}{8} = \frac{1}{3}\times\frac{14}{8}=\frac{7}{12}.$$

viii) Write down demerits of personal observation method.

Ans. Demerits:

1. It is an expensive method.
2. The enumerators must be efficient and loyal to their task, otherwise this method may fail to yield correct information.
3. This procedure is not appropriate for a large area.

OR

Write down merits of mail questionnaire method.

Merits:

1. It is not a costly method.
2. It does not need long period.
3. It is free from interviewer bias.

Group E

5. **5X4=20**

(i) Suppose S and R are respectively the standard deviation and range of n values of a variable x. Then prove that $\frac{R^2}{2n} \le S^2 \le \frac{R^2}{4}$.

Ans.

$\sum_{i=1}^{n}(x_i - c)^2$ 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

$ns^2 = \sum_{i=1}^{n}(x_i-\bar{x})^2 = (a-\bar{x})^2 + (b-\bar{x})^2 + \Sigma_1(x_i-\bar{x})^2$,

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,

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

16

(ii) State and prove Bayes' Theorem.

Ans

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

The probabilities of X,Y and Z becoming the principal of a certain college are respectively 0.3, 0.5 and 0.2. The probabilities that student aid fund will be introduced in the courage if X,Y,Z become principal, are 0.4, 0.6 and 0.1 respectively. Given that 'Student-aid-fund' has been introduced, find the probability that Y has been appointed as the principal.

Ans

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.

(iii) Why is Fisher's index number called 'ideal'?

Proof: $P_{01} = \sqrt{\frac{\Sigma P_1 q_0}{\Sigma P_0 q_0} \times \frac{\Sigma P_1 q_1}{\Sigma P_0 q_1}}$

Change p to q and q to p

$$q_{01} = \sqrt{\frac{\Sigma q_1 P_0}{\Sigma q_0 P_0} \times \frac{\Sigma q_1 P_1}{\Sigma q_0 P_1}}$$

$$P_{01} \times q_{01} = \sqrt{\frac{\Sigma P_1 q_0}{\Sigma P_0 q_0} \times \frac{\Sigma P_1 q_1}{\Sigma P_0 q_1} \times \frac{\Sigma q_1 P_0}{\Sigma q_0 P_0} \times \frac{\Sigma q_1 P_1}{\Sigma q_0 P_1}}$$

$$= \frac{\Sigma P_1 q_1}{\Sigma P_0 q_0}$$

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

Again, $I_{01} = \sqrt{\frac{\Sigma P_1 q_0}{\Sigma P_0 q_0} \times \frac{\Sigma P_1 q_1}{\Sigma P_0 q_1}}$ and $I_{10} = \sqrt{\frac{\Sigma P_0 q_1}{\Sigma P_1 q_1} \times \frac{\Sigma P_0 q_0}{\Sigma P_1 q_0}}$

∴ $I_{01} I_{10} = 1$

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

OR

Write down the steps of constructing cost of Living Index Number.

Ans: 1) Commodity Selection

2) Mostly used commodity's price quotation

3) Weight Selection

(iv) (a) Define Class Width.

Ans. The difference between the upper and lower class boundaries of a class interval is called the class width. 2

(b) What do you mean by ordinal and nominal data? Explain with example. 3

Ans.When attribute can be ordered. They are called ordinal data. Example- economic status

When there is no such ordered maintained in attribute it is called nominal data. Example – Religion of people.