MAHARASHTRA AGRICULTURAL UNIVERSITIES EXAMINATION BOARD, PUNE

SEMESTER END EXAMINATION

Model Answer Paper

B.B.M. (Agri)

Semester : IV (New) Academic Year : 2013-14
Course No. : STAT 241 Title : Business Statistics
Credits : 3(2+1)
Day & Date :
Time : 14.00 to 17.00 Total Marks: 80

Note: 1. Solve ANY Eight questions from SECTION "A".
2. All questions from SECTION "B" are compulsory.
3. All questions carry equal marks.
4. Draw neat diagrams wherever necessary.

SECTION "A"

Q. 1 Define Statistics. State the limitations and uses of statistics.

Ans. Statistics may be defined as the science of collection, presentation analysis and interpretation of numerical data from the logical analysis.

Limitation of Statistics:

1. It does not study qualitative phenomena.
2. Statistical studies are true only on an average.
3. It does not study individuals.
4. It does not reveal the entire story.
5. It is liable to be misused.

Uses of statistics

In the early part of the growth period, statistics use was restricted to the states only. But today its use has spread to the study of problems which may be social, religious, economic, political, administrative, commercial, agricultural, financial, medical, relating to business management planning research education,. Psychology, forecasting and to so many other spheres.

This widely use of statistics is on account of the fact that the statistical principals have a very wide scope of application and their knowledge is very essential for any research work in any branch of study.

Statistics has pervaded almost all spheres of human activities. Statistics is useful in the administration of various states, Industry, business, economics, research workers, banking, insurance companies etc.

Q. 2 Define Correlation. Write down the properties of correlation coefficient. Give the formula for Karl Pearson's correlation coefficient with specification.

Ans. Correlation is an analysis of the co-variation between two or more variables.

Properties of Correlation:

1. Correlation coefficient lies between -1 and $+1$ (i.e) $-1 \leq r \leq +1$.
2. ' r' is independent of change of origin and scale.
3. It is a pure number independent of units of measurement.
4. Independent variables are uncorrelated but the converse is not true.
5. Correlation coefficient is the geometric mean of two regression coefficients.
6. The correlation coefficient of x and y is symmetric. $r_{xy} = r_{yx}$.

Karl pearson' s coefficient of correlation:

The formula for calculating ' r' is

$$r = \frac{Cov(xy)}{\sigma_x \sigma_y} = \frac{Cov(xy)}{S.D(x) \times S.D(y)}$$

Where, $\sigma_x \sigma_y$ are SD of X and Y respectively.

Cov (xy) is covariance between X & Y.

Q. 3. Enlist the different methods of dispersion. Give the formulae for standard error of mean and coefficient of variation. Write properties of a good measure of dispersion.

Ans. The various absolute and relative measures of dispersion are listed below.

Absolute measure	Relative measure
1. Range	1.Co-efficient of Range
2.Quartile deviation	2.Co-efficient of Quartile deviation
3.Mean deviation	3. Co-efficient of Mean deviation
4.Standard deviation	4.Co-efficient of variation

Properties of a good measure of dispersion:

An ideal measure of dispersion is expected to possess the following properties:-

1.It should be rigidly defined.
2. It should be based on all the items.
3. It should not be unduly affected by extreme items.
4. It should be capable of algebraic manipulation.
5. It should be simple to understand and easy to calculate.

Formula for CV and SE

Coefficient of variation (C.V) = $\frac{\sigma}{\overline{X}} \times 100$

Where, σ = Standard deviation and $\overline{X}$ = Arithmetic mean.

Standard Error (S.E.) of the mean = $\frac{\sigma}{\sqrt{n}}$

Where, σ = Standard deviation and n = Number of observations.

Q. 4. Discuss exhaustive events and mutually exclusive events. State law of addition and multiplication.

Ans. **Exhaustive Events:**

The total number of outcomes in any trial is known as exhaustive events.

Ex. In tossing of a coin there are two exhaustive cases i.e. head and trial.

1. **Mutually Exclusive Events:**

Events are said to be mutually exclusive if the happening of any one of them precludes the happening of all the others i.e. no two or more of them can happen simultaneously in the same trial. Ex. In tossing a coin the events head and tail are mutually exclusive.

Law of addition:

If A and B are two arbitrary (possible) events in the sample space S, the probability of the union of A and B is governed by the law,

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

Law of Multiplication:

If two events A and B are independent, the probability of their product of their individual probabilities.

$$P(AB) = P(A \cap B) = P(A)\,P(B).$$

Q. 5 Define Average. State the properties of good average. Give the Empirical relationship between Mean, Median and Mode.

Ans. Average is defined as the sum of the observations divided by the number of observations.

Properties for good average

The following properties should possess for an ideal average.

1. It should be rigidly defined.
2. It should be easy to understand and compute.
3. It should be based on all items in the data.
4. Its definition shall be in the form of a mathematical formula.
5. It should be capable of further algebraic treatment.
6. It should have sampling stability.

Empirical relationship between mean, median and mode

In a symmetrical distribution the three simple averages mean = median = mode. For a moderately asymmetrical distribution, the relationship between them are brought by Prof. Karl Pearson as

mode = 3median - 2mean.

Q. 6 Define test of significance. State the assumptions and applications of t-test.

Ans. **Test of significance:**

Student's t is the deviation of estimated mean from its population mean expressed in terms of standard error. The test criteria is; reject H_0 if $t \geq t_{\alpha/2,n-1}$ or $-t \leq -t_{\alpha/2,n-1}$, otherwise accept H_0. Let Xi (i = 1,2,...n) be a random sample of size n from a normal population with mean μ and variance σ^2. Then t –test is defined as

$$t = \sqrt{n}\,(\bar{x} - \mu_0)/s, \quad -\infty \leq t \leq \infty$$

Where, $$s = \sqrt{\Sigma(x-\bar{x})^2/(n-1)}$$

Assumption for t-test

The assumptions under which t-test is applied are given below:

1. Parent population is normal.
2. Sample has been selected randomly.
3. Sample size is small.
4. Population standard deviation is not known.

Applications of t-test

In small samples, students t-test is applied in the following cases.:

1. Comparison of a sample mean with the population mean (or a hypothetical mean).
2. Testing the significance of a mean difference (paired observations).
3. Comparison of two means from two independent samples.

Q. 7 Write short notes of following. (Any two)
i) 'O' give curve. ii) Normal Distribution
iii) Poisson Distribution

Ans. **i) Ogives:**

For a set of observations, we know how to construct a frequency distribution. In some cases we may require the number of observations less than a given value or more than a given value. This is obtained by an accumulating (adding) the frequencies upto (or above) the give value. This accumulated frequency is called cumulative frequency. These cumulative frequencies are then listed in a table is called cumulative frequency table. The curve table is obtained by plotting cumulative frequencies is called a cumulative frequency curve or an ogive. There are two methods of constructing ogive namely:

1. The 'less than ogive' method
2. The 'more than ogive' method.

In less than ogive method we start with the upper limits of the classes and go adding the frequencies. When these frequencies are plotted, we get a rising curve. In more than ogive method, we start with the lower limits of the classes and from the total frequencies we subtract the frequency of each class. When these frequencies are plotted we get a declining curve.

ii) Normal Distribution:

A r. v. X is said to have a normal distribution with parameters μ (called mean) and σ^2 (called variance) if its density function is given by the prob. Law:

$$f(x) = \frac{1}{\sigma\sqrt{2\pi}} e^{-\frac{1}{2}\left(\frac{x-\mu}{\sigma}\right)^2}; \quad -\infty < x < \infty, -\infty < \mu < \infty, \sigma > 0.$$

Properties of Normal Distribution.

1. The curve is bell shape and symmetrical about the line $x = \mu$.
2. Mean, Median and mode are coincide i.e. mean = median =mode.
3. Quartiles are equidistant from median.
4. Linear combination of independent normal variates is also a normal variate.
5. The mode of the normal curve lies at the point $x = \mu$.
6. The area under the normal curve within its range $-\infty$ to ∞ is always unity.
7. The normal curve is unimodal.
8. All odd order moments of the normal distribution are zero.

iii) Poisson Distribution:

A r. v. X is said to follow a Poisson distribution if it assumes on ly non – negative values; and its probability mass function is given by

$$P(X = x) = \begin{cases} \frac{e^{-\lambda}\lambda^x}{x!}; x = 0,1,2,\ldots,n \\ 0; otherwise \end{cases}$$

Where λ is known as parameter of the distribution. A r. v. X is follows P.D. with parameter λ is denote as $X \sim P(\lambda)$.

Properties of Poisson distribution:

1. Poisson distribution has only one parameter.
2. Mean of Poisson distribution (variate) is λ.
3. Variance of Poisson distribution is also λ. This is the only distribution of which mean and variance are equal.
4. Moment generating function of Poisson distribution is $e^{\mu(e^{it}-1)}$.
5. If X and Y are two Poisson variates with mean μ and λ respectively, the conditional distribution of X given (X+Y) is binomial.

Q. 8 Define Geometric Mean and Harmonic Mean. Write down merits and demerits of Geometric Mean. Give the formula for Harmonic Mean of grouped data.

Ans. **Geometric Mean:** The geometric mean of a series containing n observations is the n_{th} root of the product of the values.

Harmonic mean (H.M) : Harmonic mean(H.M.) of a set of observations is defined as the reciprocal of the arithmetic average of the reciprocal of the given values.

Merits of Geometric mean :

1. It is rigidly defined
2. It is based on all items
3. It is very suitable for averaging ratios, rates and percentages
4. It is capable of further mathematical treatment.
5. Unlike AM, it is not affected much by the presence of extreme values

Demerits of Geometric mean:

1. It cannot be used when the values are negative or if any of the observations is zero
2. It is difficult to calculate particularly when the items are very large or when there is a frequency distribution.
3. It brings out the property of the ratio of the change and not the absolute difference of change as the case in arithmetic mean.
4. The GM may not be the actual value of the series.

Harmonic Mean is given as

$$H.M. = \frac{N}{\sum_{i=1}^{n} \frac{f}{x}}$$

Q. 9 Define Skewness and Kurtosis. Explain in brief the different measures of skewness.

Ans. **Skewness:** Lack of symmetry is said to be Skewness.

Kurtosis: The convexity of curve is said to be Kurtosis.

Different measures of Skewness:

i. $Sk = M - Mo$
ii. $Sk = M - Md$
iii. $Sk = (Q_3 - Md) - (Md - Q_1)$

These are the absolute measures of Skewness.

iv. **Karl Pearson's Coefficient of Skewness**

$$Skp = \frac{\text{Mean - Mode}}{\sigma}$$

Where, Skp is karl pearsons coefficient of skewness.

σ is standard deviation.

Bowleys coefficient of Skewness:

$$Skb = \frac{Q_3 + Q_1 - 2Md}{Q_3 - Q_1},$$

Where, Q_1 = First Quartile

Q_3 = Third Quartile

Q. 10 Explain the terms degrees of freedom, type I & type II errors, Null hypothesis, level of significance and critical region.

Ans. **Degrees of Freedom:** The number of independent observations in a set is called degrees of freedom.

Type I error: Reject H_0 when it is true. It is also known as rejection error and it is denoted by α.

Type II error: Accept H_0 when H_1 is true. It is also known as acceptance error and it is denoted by β.

Null hypothesis: The null hypothesis is defined as a particular hypothesis which is tested for rejection under the assumption that it is true, as distinct from the alternative hypotheses which are under consideration.

Level of significance: the maximum probability of rejecting the hypothesis when it is true is known as level of significance.

Critical Region: The hypothesis is rejected when an observed difference is equal to or greater thanδ, all the values, which are beyond from a region known as Critical region.

SECTION "B"

Q. 11 Fill up the blanks.

1. Standard deviation of the series 5,5,5,5,5 and 5 is **Zero.**
2. The sum of the deviations taken from the arithmetic mean is equal to **Zero.**
3. For a mesokurtic curve, β_2 equal to **3.**
4. The number of seeds per pod is **Discrete** variable.
5. The middle value of an ordered series is called **Second** quartile.
6. Homogeneity of two sample variances is tested by **F test.**
7. Large sample test is applied when sample size is greater than **50.**
8. In dependence of attributes is tested by **Chi-square** test.

Q. 12. Match the pairs.

'A'	'B'
1. t-test	a) The real number lies between 0 & 1
2. f-test	b) random experiment
3. χ^2 test	c) The value lies between -1 & +1
4. Normal Distribution	d) Discrete distribution
5. Poisson Distribution	e) Continuous distribution
6. Binomial Distribution	f) Mean & Variance are same
7. Probability	g) Testing equality of Variances
8. Correlation	h) Testing goodness of fit
	i) Used for testing equality of two small sample mean

Ans. [1] i [2] g [3] h [4] e
[5] f [6] d [7] a [8] c