Model Answers

Integrated UG/PG Biotechnology, V Semester Examination, 2014-15
Subject Code: LBTC 504
Subject Name: Biostatistics and Computer Applications

Section 1: Answer all the questions. Each question carries 1.0 mark. Choose one correct answer for the following questions. **(10 x 1 = 10)**

i. One use of a regression line is
~~(c) to determine if a change in x causes a change in y.~~
(a) to estimate the change in y for a specified change in x

ii. The weight of seeds are 5, 7, 8, 6 and 4 mg. The geometric mean of weight of seeds is
(b) 5.9 mg

iii. From a well-shuffled deck of 52 cards, a card is drawn at random. Find the probability of getting an ace.
(a) 1/13

iv. An important application of the chi-square distribution is
~~(b) Testing for goodness of fit~~
(b) all of the above

v. The mean of a binomial variate with parameter n and p is:
(a) np

vi. Which technology is used to produce microprocessor chips in fifth generation computer?
(a) ULSI

vii. Supercomputers use ---------- technology to solve the complex problems faster.
(c) both multiprocessing and parallel processing

viii. Decimal equivalent of binary number 11001 is ----------
(b) 25

ix. Hexadecimal equivalent of decimal number 428 is ---------
(a) 1AC

x. Which one of the following is use as a numerical data analysis tool that allows us to create a computerized ledger?
(b) Spreadsheet

*Section 2: Each question carries 5 marks. Attempt only **four** questions out of the following **eight** questions.* **(4 x 5 = 20)**

2. Derive the mathematical expression for the calculation of standard deviation for both the discrete and grouped data applying algebra method. Calculate the standard deviation of the following data set.

No. of pods/plant	15-17	18-20	21-23	24-26	27-29	30-32	33-35	36-38	39-41
Frequency	5	6	8	12	22	18	15	9	5

Answer:

Algebraic Derivation of Standard Deviation

Let, d=deviation,
n = number of items, and
$\bar{x}$ = sample mean

Thus $d = x - \bar{x}$ from the definition

Therefore, $d^2 = (x - \bar{x})^2 = x^2 - 2x\bar{x} + \bar{x}^2$

Since the deviations are summed, the equation becomes,

$\Sigma d^2 = \Sigma x^2 - 2\Sigma x\bar{x} + \Sigma\bar{x}^2$

The mean is a constant for all observations, thus the equation can be rewritten by transferring the x outside the summation. $\Sigma\bar{x}^2$ is the $\bar{x}^2$ added by n times which can be written as $n\bar{x}^2$. By substituting all these figures in the equation we get :

$$\Sigma d^2 = \Sigma x^2 - 2\Sigma x\bar{x} + n\bar{x}^2$$

Since the mean, $\bar{x}$ is $\frac{\Sigma x}{n}$ by definition, we can replace $\bar{x}$ by $\frac{\Sigma x}{n}$. Thus the equation will be

$$\Sigma d^2 = \Sigma x^2 - 2\Sigma x \cdot \frac{\Sigma x}{n} + n\frac{(\Sigma x)^2}{n^2}$$

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

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

In case of grouped data (with frequency values) as is done in most of the biological studies, the d^2 may be substituted by fd^2. The equation will be

$$\Sigma fd^2 = \Sigma fx^2 - \frac{(\Sigma fx)^2}{n} \quad \text{or} \quad \Sigma f_x{}^2 - \frac{(\Sigma fx)^2}{\Sigma f}$$

Whatever be the way of calculation the result will be the same. The process of calculation using the formulae

$$\Sigma d^2 = \Sigma x^2 - \frac{(\Sigma x)^2}{n} \quad \text{or} \quad \Sigma fd^2 = \Sigma fx^2 - \frac{(\Sigma fx)^2}{\Sigma f}$$

becomes easier with the help of a calculating machine. The value calculated in the way outlined above will be equal to the variance calculated taking squares of deviations.

No. of Pods/Plant (class of 3)	*Midpoint m*	*No. of Plants (frequency) f*	fm	m^2	fm^2
15–17	16	5	80	256	1280
18–20	19	6	114	361	2166
21–23	22	8	176	484	3872
24–26	25	12	300	625	7500
27–29	28	22	616	784	17248
30–32	31	18	558	961	17298
33–35	34	15	510	1156	17340
36–38	37	9	333	1369	12321
39–41	40	5	200	1600	8000
		100	2887		87025

$$\text{Mean} = \frac{2887}{100} = 28{\cdot}87$$

$$\text{Variance, } s^2 = \frac{\Sigma fm^2 - \frac{(\Sigma fm)^2}{\Sigma f}}{\Sigma f}$$

$$= \frac{87025 - \frac{(2887)^2}{100}}{100}$$

$$= \frac{87025 - 83347{\cdot}69}{100}$$

$$= \frac{3677{\cdot}31}{100} = 36{\cdot}77$$

$$\text{Standard deviation, } s = \sqrt{s^2} = \sqrt{36{\cdot}77} = 6{\cdot}064$$

3. A die is tossed once. What is the probability of getting:
(a) the number 4? (b) an even number? (c) a number less than 5?
(d) a number greater than 4? (e) a number less than 8?

Answer:

The sample space associated with the random experiment of rolling a die is given by S = {1,2,3,4,5,6}. Total number of elementary events = 6.

(a) The probability of getting a number 4 is 1/6.

(b) An even number is obtained, if we obtain any one of 2,4,6 as an outcome. So favourable number of elementary events = 3.

Hence required probability = 3/6 = ½ = 0.5

(c) A number less than 5 is obtained, if we get any one of 1,2,3,4 as an outcome. So favourable number of elementary events = 4.

Hence, required probability = 4/6 = 2/3.

(d) A number greater than 4 is obtained, if we get any one of 5 and 6 as an outcome. So favourable number of elementary events = 2.

Hence, required probability = 2/6 = 1/3.

(e) Since every face of a die is marked with a number less than 8, so favourable number of elementary events = 6.

Hence, required probability = 6/6 = 1.

4. If the probability that an individual suffers a bad reaction from an injection of a given serum is 0.001, determine the probability that out of 2000 individuals (i) exactly 3 individuals and (ii) more than 2 individuals suffer from bad reaction. (Given $e^{-2} = 0.13534$).

SOLUTION. Let p be the probability that an individual suffers a bad reaction from an injection of a given serum and n be the total number of individuals who took the injection.

It is given that $n = 2000$ and $p = 0.001$

Since p is very small and n is large, therefore we shall use Poisson's distribution.

We have, $m = np \Rightarrow m = 2000 \times 0.001 = 2$.

Let X denote the number of individuals who suffer from bad reaction. Then, X is a Poisson variate such that

$$P(X = r) = \frac{m^r e^{-m}}{r!}, r = 0, 1, 2, \ldots, \infty$$

$$\Rightarrow \quad P(X = r) = \frac{2^r e^{-2}}{r!}, r = 0, 1, 2, \ldots, \infty$$

(i) Required probability $= P(X = 3)$

$$= \frac{2^3 e^{-2}}{3!}$$

$$= \frac{4}{3} e^{-2} = \frac{4}{3} \times 0.13534 \qquad [\because e^{-2} = 0.13534]$$

$$= 0.18045$$

(ii) Required probability $= P(X > 2)$

$$= 1 - P(X \leq 2)$$

$$= 1 - [P(X = 0) + P(X = 1) + P(X = 2)]$$

$$= 1 - \left[e^{-2} + 2 e^{-2} + \frac{2^2 e^{-2}}{2!}\right]$$

$$= 1 - e^{-2}\left[1 + 2 + \frac{4}{2}\right]$$

$$= 1 - \frac{5}{e^2} = 1 - 5 \times 0.1353$$

$$= 1 - 0.67670 = 0.323$$

5. In a grassland the earthworm population was sampled from ten randomly located plots of 1 m^2 area. The following table gives the number of earthworms obtained. Examine the distribution pattern of earthworms.

Quadrates	1	2	3	4	5	6	7	8	9	10
Earthworm No./m^2	25	32	17	23	15	39	27	19	22	26

Answer:

Null hypothesis (Ho): The earthworm population is distributed equally among quadrats.

Alternative hypothesis (H1): The earthworm population is not distributed equally among the quadrats.

The test statistic that will be used to test the hypothesis is chi-square test.

The expected number of earthworms is determined taking into consideration that the population is equally distributed in all quadrates. Thus the expected number of earthworm in each quadrat is the mean number of earthworms.

Observed	25	32	17	23	15	39	27	19	22	26
Expected	24.5	24.5	24.5	24.5	24.5	24.5	24.5	24.5	24.5	24.5
Difference	0.5	7.5	-7.5	-1.5	-9.5	14.5	2.5	-5.5	-2.5	1.5

$$\chi^2 = \frac{(0.5)^2}{24.5} + \frac{(7.5)^2}{24.5} + \frac{(-7.5)^2}{24.5} + \frac{(-1.5)^2}{24.5} + \frac{(-9.5)^2}{24.5} + \frac{(14.5)^2}{24.5} + \frac{(2.5)^2}{24.5} + \frac{(-5.5)^2}{24.5} + \frac{(-2.5)^2}{24.5} + \frac{(1.5)^2}{24.5}$$

$$= 0.01 + 2.3 + 2.3 + \cdots + 0.09$$

$$= 18.80$$

Chi-square – 18.80

The tabulated value of chi-square with d.f. 9 at p = 0.05 is 16.92. Since the tabulated value is less than the calculated value, the null hypothesis is rejected, i.e. the earthworm population is not distributed equally.

6. What is "generation" in computer terminology? List various computer generations along with key characteristics of computers of each generation.

Answer: Generation in computer talk provides a framework for the growth of computer industry based on key technologies developed. Originally it was used to distinguish between hardware technologies but was later extended to include both hardware and software technologies.

Brief description of each generation computer along with key features is needed.

7. Find the equivalent number of the following numbers given below as mentioned.

(i) $101110_2 = ?_8$ (ii) $1101010_2 = ?_8$ (iii) $562_8 = ?_2$

(iv) $ABC_{16} = ?_2$ (v) $101101011100_2 = ?_{16}$

Answer:

(i) $101110_2 = ?_8$

Step 1 Convert 101110_2 to base 10

$101110_2 = 1\times2^5 + 0\times2^4 + 1\times2^3 + 1\times2^2 + 1\times2^1 + 0\times2^0$

$= 32 + 0 + 8 + 4 + 2 + 0 = 46_{10}$

Step 2 Convert 46_{10} to base 8

8	46	Remainders
	5	6
	0	5

$46_{10} = 56_8$

Hence $101110_2 = (56)_8$

~~(ii) $1101010_2 = ?_8$~~ 0 4

$101110_2 = ?_8$

Step 1: Divide the binary digits into groups of 3 starting from right

101 110

$(101)_2 = 1\times2^2 + 0\times2^1 + 1\times2^0 = 4+0+1 = 5_8$ Hence $101110_2 = 56_8$

$(110)_2 = 1\times2^2 + 1\times2^1 + 0\times2^0 = 4+2+0 = 6_8$

(ii) $1101010_2 = ?_8$

$1101010_2 = 001\ 101\ 010 = 152_8$ (convert each group to octal digit)

(iii) $562_8 = ?_2$

Step 1: Convert each octal digit to 3 binary digits.

Step 2: Combine the binary digits

$(562)_8$: 101 110 010

(iv) $ABC_{16} = ?_2$

$(ABC)_{16} = \underset{A}{1010} \quad \underset{B}{1011} \quad \underset{C}{1100}$

$= (101010111100)_2$

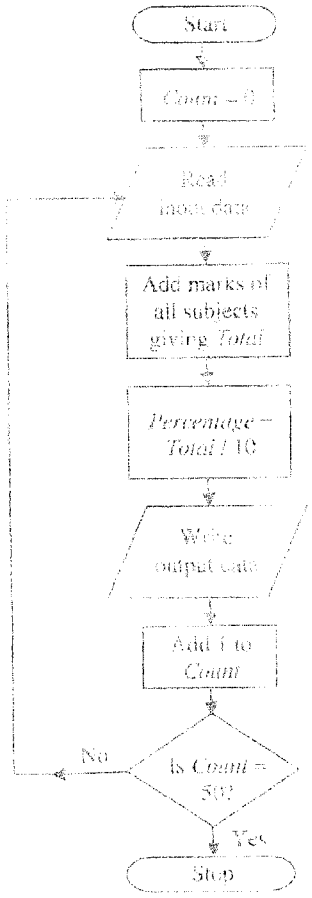
(v) $101010111100_2 = ?_{16}$

= 0101 1010 1100 (Group 4 digits from right)

= 5AC (convert each group to hexadecimal digit)

Hence $(101010111100)_2 = (5AC)_{16}$

8. Hundred students of a class appear in the examination consisting of 10 subjects with each subject having maximum marks of 100. Write an algorithm and draw a flow chart to calculate the percentage marks obtained by each student.



9. What is a word-processing package? Write key features supported by modern word-processing packages.

Answer: The term word-processing describes use of hardware and software to create, edit, view, format, store, retrieve and print documents. A word-processing package enables us to do all these on a computer system.

Discuss the different features of word processing package with following headings

i. Entering text and Editing text
ii. Formatting page style
iii. Formatting text (selection of font type, font size, font style, justification, creating number or bulleted, indenting text)
iv. Entering mathematical symbols
v. Displaying document
vi. Saving, retrieving and deleting document
vii. Printing documents, Importing text, graphics and images
viii. Checking spelling, Checking grammar and style