Statistics for Decision Making - I

End Semester Examination

Full Marks: 40 Time : 2 hr 30 minutes

Answer 1 and 2 and any two from the rest.

 a) In a group of 500 persons, 220 are literate and employed, 20 are literate and unemployed and 180 are illiterate and unemployed. Is there any association between literacy and employment? (5)

b) Data from a case-control study of 200 esophageal cancer cases and 775 community-based controls are shown below. The data tries to explore relationship between alcohol consumption (dichotomized at 80 grams per day) and esophageal cancer. Use odds ratio and give your comments.
 (5)

Alcohol	Esophageal cancer			
g/day	+	-		
+	96	109		
-	104	666		

2. Following are the marks obtained by a group of 43 students in a science test:

47	26	45	19	7	30	27	23	12	
47	35	28	26	15	36	23	26	29	
46	37	39	28	29	37	8	30	36	
28	32	29	23	28	21	13	24		
37	38	22	27	32	24	20	13		

Draw a SRSWR and SRSWOR both of size 10. In each case, estimate the average marks and also estimate the corresponding standard error. (10)

b) Define positive predictive value and negative predictive value. (3).

a) 'In multiple linear regression, if we go on increasing the number of variables, the predicted values will be more closely related to the observed values'
Prove or disprove. (6)

b) Show that, mean deviation about mean can not exceed standard deviation. When are the two equal? (4)

5. a) Let X_1 and X_2 be a random sample of size 2 from a distribution having pdf

 $f(x) = \begin{cases} e^{-x}, 0 < x < \infty \\ 0, otherwise \end{cases}$

Define $Y_1 = X_1 + X_2$ and $Y_2 = \frac{X_1}{X_1 + X_2}$. Show that, Y_1 and Y_2 are stochastically independent. (6)

b) In case of simple random sampling, show that, $P(y_r = Y_i) = \frac{1}{N}$, irrespective of whether the samples are drawn with replacement or without. (4)