Answer **any one** essay question out of 2. **(1×10=10)**

32. The accompanying table presents data on yields relating to resistance to stain for three materials. (M1, M2 and M3) treated with four chemicals in a randomized block design. (A low value indicates good stain resistance)

	Material			
Chemical	**M1**	**M2**	**M3**	**Total**
A	5	9	7	21
B	3	8	4	15
C	8	13	9	30
D	4	6	8	18
Total	20	36	28	84

Is there evidence of difference in mean resistance among the four chemicals ? Give bounds for the p-value. What would you conclude at the $\alpha = .05$ level of significance ?

33. A response Y is a function of three independent variables x_1, x_2 and x_3 that are related as follows :

$Y = \beta_0 + \beta_1 x_1 + \beta_2 x_2 + \beta_3 x_3 + \varepsilon$,

a) Fit this model to the n = 7 data points shown in the accompanying table :

y	x_1	x_2	x_3
1	– 3	5	– 1
0	– 2	0	1
0	– 1	– 3	1
1	0	– 4	0
2	1	– 3	– 1
3	2	0	– 1
3	3	5	1

b) Predict Y when $x_1 = 1$, $x_2 = -3$, $x_3 = -1$. Compare with the observed response in the original data. Why are these two not equal ?

c) Do the data present sufficient evidence to indicate that x_3 contributes information for the prediction of Y ? (Test the hypothesis $H_0 : \beta_3 = 0$, using $\alpha = .05$)

d) Find a 95% confidence interval for the expected value of Y, given $x_1 = 1$, $x_2 = -3$ and $x_3 = -1$.

e) Find a 95% prediction interval for Y, given $x_1 = 1$, $x_2 = -3$ and $x_3 = -1$.

Reg. No. : .....

Name : .....

IV Semester B.Sc. Hon's (Mathematics) Degree (Supple. / Improv.) Examination, May 2018
(2013-15 Admissions)
BHM 405 : ADVANCED STATISTICS – II

Time : 3 Hours Max. Marks : 80

Answer **all** the **10** questions. **(10×1=10)**

1. Give an unbiased estimator for σ^2.
2. Give the test statistic for testing the significance of correlation coefficient.
3. What is the standard error of $\bar{Y}_1 - \bar{Y}_2$ in estimating $\mu_1 - \mu_2$ based on independent samples ?
4. What is a treatment in experimental design ?
5. Give the statistical model for two-way layout.
6. Give any two applications of chi-square distribution.
7. Give the 100 $(1 - \alpha)$% confidence interval for β_1.
8. What is coefficient of determination ?
9. What is the variance of the least square estimate of β_1 in a simple linear regression model ?
10. What is the degrees of freedom corresponding to a contingency table with 4 rows and 3 columns ?

Answer **any 10** short answer questions. **(10×3=30)**

11. What are the assumptions underlying a randomised block design ?
12. If n observations are to be used to estimate $\mu_1 - \mu_2$ on independent samples from two populations of interest, find n_1 and n_2 so that $V(\bar{Y}_1 - \bar{Y}_2)$ is minimised.
13. S. T. the prediction interval for an actual value of Y when $x = x^*$ is longer than confidence interval for E(Y) at $x = x^*$.

14. Prove that in one-way ANOVA :
    Total SS=SST+SSE

15. Describe a completely randomised design.

16. Describe with suitable example matched pair experiment.

17. Describe method of least squares for estimating the parameters of a line.

18. What are the properties of the least square estimators in a simple linear regression ?

19. Using the information given below estimate σ^2.
    $n = 5, \Sigma y_i = 5, S_{xy} = 7, \beta_1\hat{} = 0.7, \Sigma y_i^2 = 11$

20. Find the moment estimator of ρ.

21. Give the test procedure to test β_i in the simple linear regression model.

22. We have fit a model with k independent variables and wish to test $H_0 : \beta_1 = \beta_2 = ... = \beta_k = 0$. Give an appropriate test statistic.

Answer **any 6** short essay questions out of 9 : **(6×5=30)**

23. Describe chi-square goodness of fit test.

24. Derive the 100 $(1 - \alpha)$% confidence interval for $\theta = a_0\beta_0 + a_1\beta_1$.

25. Fit a line to the given data. Give the estimates of β_0 and β_1.

Y :	3	2	1	1	0.5
X :	−2	−1	0	1	2

26. A weight reduction programme was administered to 10 persons. The results are given below.

Candidates	1	2	3	4	5	6	7	8	9	10
Weight before (kg)	86	92	100	93	88	80	88	92	95	106
Weight after (kg)	77	84	92	87	80	74	80	85	95	96

Test whether there is any change in weight after training programme (Use 5% level of significance).

27. The correlation coefficient for the heights and weights of 10 football players was determined to be r = 0.8261
    1) What percentage of the variation in weight was explained by the heights of the players ?
    2) Is there significance evidence at $\alpha = 0.01$ level to claim that heights and weights are positively correlated ?

28. Suppose $Y_1, Y_2, ... Y_n$ are independent normal variates with $E(Y_i) = \beta_0 + \beta_1 x_i$ and $V(Y_i) = \sigma^2$ for i = 1, 2, ... n. S.T. mle of β_0 and β_1 are the same as the least square estimates.

29. In a breeding experiment, the ratio of off-springs in four classes was expected to be 1 : 3 : 3 : 9. The experiment yielded the data as follows :

Classes :	AA	Aa	aA	aa
No. of off-springs :	8	29	37	102

Test whether the given data is in agreement with the hypothetical ratio.

30. The model $Y = \beta_0 + \beta_1 x_1 + \beta_2 x_2 + \beta_3 x_1 x_2 + \beta_4 x_2^2 + \varepsilon$, where Y = annual salary (in thousands of rupees), $x_1 = 1$ if female and $x_1 = 0$ if male. X_2 = amount of experience (in years). When this model was fit to data obtained from the records of 200 faculty members, SSE = 783.90. The reduced model $Y = \beta_0 + \beta_1 x_2 + \beta_2 x_2^2 + \varepsilon$ was also fit and produced a value SSE = 795.23. Do the data provide sufficient evidence to support the claim that the mean salary depends on the gender of the faculty members ($\alpha = 0.05$).

31. A plant manager in deciding whether to purchase a machine of design A or design B, checks the times for completing a certain task on each machine. Eight technicians were used in the experiment with each technician using both machine A and machine B in a randomised manner. The times (in seconds) required to complete the task are given below :

Technician :	1	2	3	4	5	6	7	8
A :	32	40	42	26	35	29	45	22
B :	30	39	42	23	36	27	41	21

Test to see if there is significance difference between mean completion times at 5% level of significance.