EASTERN UNIVERSITY, SRI LANKA

FACULTY OF COMMERCE AND MANAGEMENT

THIRD YEAR FIRST SEMESTER EXAMINATION IN BBA/ B.COM

(GENERAL & SPECIALIZATION) – 2007/ 2008 (PROPER)

DAF 3134 BUSINESS STATISTICS

Answer all questions **Time : 03 Hours**

---

01. (A) (i) Distinguish between a pair of terms 'continuous random variable' and 'discrete random variable' with suitable examples.

(ii) A student majoring in accounting has been told by a placement counselor that she can expect to receive a job offer from 80% of the firms to which she applies. The student applies to only five firms. Let X be the random variable represent that the number of offers she receives.

(a) What are the possible values of X?

(b) Is the random variable X is discrete or continuous? Explain.

(c) Construct the probability distribution of X.

(d) Express the probability distribution of X graphically.

(e) What assumptions did you make in constructing the probability distribution of X?

(f) Find the probability that the student receives the following :

a) No offers

b) At most two offers

c) Between 2 and 4 offers (Inclusive)

d) Exact 5 offers.

(g) What is the expected number of offers she receives?

(h) What is the variance of offers she receives?

(B) A normally distributed population has a mean of 40 and a standard deviation of 12.

(i) What does the central limit theorem say about the sampling distribution of the mean if samples of size 100 are drawn from this population?

(ii) Suppose that population is not normally distributed. Does this change your answer? Explain.

02. (A) (i) Using suitable examples explain how does a binomial random variable differ from a Poisson random variable.

(ii) An average of 3 accidents per week occur on an assembly line. Let X be the number of accidents per week.

(a) State the appropriate probability distribution of X clearly.

(b) What is the probability that a particular week will be accident free?

(c) What is the probability that there will be at least 3 accidents in a 2 weeks period?

(B) (i) (a) What are the parameters of a normal distribution?

(b) Describe the effect of changing the value of variance of a normal distribution with the same mean.

(ii) A contractor is planning the construction of a new office building, and has identified, three major critical activities involved in the project. He estimates that excavating and foundation work will take an expected time of 2 weeks with a variance of 0.10, framing will take an expected 3 weeks with a variance of 0.15, and finishing will take an expected 5 weeks with a variance of 0.90. Assume that the completion times for these three activities are statistically independent, and that the total time to complete the project is normally distributed.

(a) Compute the values of the parameters of the probability distribution of total time to complete the project.

(b) Find the probability that the project will take at most 12 weeks to complete.

(c) The contractor will receive a bonus if the project is completed in less than 9 weeks. Find the probability that a bonus will be paid.

03. (A) (i) Define the term consistency in the context of estimation.

(ii) Draw diagrams representing what happens to the sampling distribution of a consistent estimator when the sample size increases.

(B) A study was conducted to investigate the use of humour in television advertisements in the United States and in the United Kingdom. A random sample of 400 television advertisements in the United Kingdom revealed that 142 of the advertisements used humour, while a random sample of 500 television advertisements in the United States revealed that 122 of the advertisements used humour. It is wanted to test whether the proportion of television advertisements using humour in the United Kingdom is the same as the proportion of advertisements using humour in the United States at the 5% level of significance.

(i) State the null and alternative hypotheses for the test.

(ii) Show that the large sample conditions are satisfied in this case to approximate the sampling distribution of $\hat{P}_1 - \hat{P}_2$ by the normal distribution.

(iii) Calculate the value of the test statistics?

(iv) Sketch the rejection region of the test.

(v) What is the decision rule for the test?

(vi) What is the decision of the test?

(vii) Explain how you reached your decision in (vi).

(viii) Interpret the result of this test.

(C) In a time study in the banking industry, 30 randomly selected managers spent a mean of 24 hours each day on paper work with a standard deviation of 13 hours.

(i) Construct a 95% confidence interval for the mean paper work time of all the managers.

(ii) Interpret the confidence interval in the context of the question.

(iii) For the results in part (i) to be valid, must the population distribution of hours spent by bank managers on paper work be normal? Explain why or why not.

(iv) What sample size would be required to estimate the mean number of hours spent each day on paper work by bank managers to within ½ hour with 99% confidence?

04. (A) State the assumptions of a simple linear regression model.

(B) The owner of Maumee Motors wants to study the relationship between the age of a car and its selling price. Listed below is a random sample of 12 used cars sold at Maumee Motors during the last year.

X Age (years)	9	7	11	12	8	7	8	11	10	12	6	6
Y Selling price (Rs.1000)	8.1	6.0	3.6	4.0	5.0	10.8	7.6	8.0	8.0	6.0	9.4	8.6

(i) Draw a scatter diagram of Y against X.

(ii) Examine the scatter plot and decide whether a straight line model is a reasonable model.

Suppose that the estimated least squares line is $\hat{Y}$ = 5.531 – 2.327.

(iii) Interpret the slope of the regression line in the context of the problem.

(iv) Test to see whether there is a significant linear relationship between age and selling price.

(v) Use the estimated regression line to predict the selling price of a car that is 8 years old.

(vi) If the correlation coefficient is 0.59, compute the amount of the variation in selling price explained by the linear regression model based on age.

(vii) Calculate the predicted values of Y and the residuals.

(viii) Using the values calculated in part (vii), draw the appropriate diagrams to identify whether ;

a) there is heteroscedasticity problem,

b) the errors are normally distributed,

and comment the diagrams.

05. (A) Describe the components of time series with suitable examples.

(B) The revenues (in million of rupees) of a chain of ice cream stores are listed for each quarter during the years 2003 – 2007 in the accompanying table.

Quarter	Year				
	2003	2004	2005	2006	2007
1	16	14	17	18	21
2	25	27	31	29	30
3	31	32	40	45	52
4	24	23	27	24	32

(i) Determine the trend line using regression analysis.

(ii) Compute the seasonal indices using the trend line.

(iii) Interpret the seasonal pattern.

(iv) Deseasonalize the revenues.

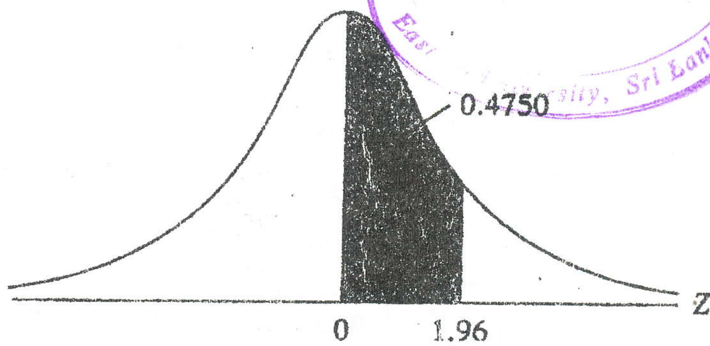
(v) Forecast the revenue for the four quarters of 2008.

Areas under the standardized normal distribution

Example

$\Pr(0 \leq Z \leq 1.96) = 0.4750$

$\Pr(Z \geq 1.96) = 0.5 - 0.4750 = 0.025$



Z	.00	.01	.02	.03	.04	.05	.06	.07	.08	.09
0.0	.0000	.0040	.0080	.0120	.0160	.0199	.0239	.0279	.0319	.0359
0.1	.0398	.0438	.0478	.0517	.0557	.0596	.0636	.0675	.0714	.0753
0.2	.0793	.0832	.0871	.0910	.0948	.0987	.1026	.1064	.1103	.1141
0.3	.1179	.1217	.1255	.1293	.1331	.1368	.1406	.1443	.1480	.1517
0.4	.1554	.1591	.1628	.1664	.1700	.1736	.1772	.1808	.1844	.1879
0.5	.1915	.1950	.1985	.2019	.2054	.2088	.2123	.2157	.2190	.2224
0.6	.2257	.2291	.2324	.2357	.2389	.2422	.2454	.2486	.2517	.2549
0.7	.2580	.2611	.2642	.2673	.2704	.2734	.2764	.2794	.2823	.2852
0.8	.2881	.2910	.2939	.2967	.2995	.3023	.3051	.3078	.3106	.3133
0.9	.3159	.3186	.3212	.3238	.3264	.3289	.3315	.3340	.3365	.3389
1.0	.3413	.3438	.3461	.3485	.3508	.3531	.3554	.3577	.3599	.3621
1.1	.3643	.3665	.3686	.3708	.3729	.3749	.3770	.3790	.3810	.3830
1.2	.3849	.3869	.3888	.3907	.3925	.3944	.3962	.3980	.3997	.4015
1.3	.4032	.4049	.4066	.4082	.4099	.4115	.4131	.4147	.4162	.4177
1.4	.4192	.4207	.4222	.4236	.4251	.4265	.4279	.4292	.4306	.4319
1.5	.4332	.4345	.4357	.4370	.4382	.4394	.4406	.4418	.4429	.4441
1.6	.4452	.4463	.4474	.4484	.4495	.4505	.4515	.4525	.4535	.4545
1.7	.4454	.4564	.4573	.4582	.4591	.4599	.4608	.4616	.4625	.4633
1.8	.4641	.4649	.4656	.4664	.4671	.4678	.4686	.4693	.4699	.4706
1.9	.4713	.4719	.4726	.4732	.4738	.4744	.4750	.4756	.4761	.4767
2.0	.4772	.4778	.4783	.4788	.4793	.4798	.4803	.4808	.4812	.4817
2.1	.4821	.4826	.4830	.4834	.4838	.4842	.4846	.4850	.4854	.4857
2.2	.4861	.4864	.4868	.4871	.4875	.4878	.4881	.4884	.4887	.4890
2.3	.4893	.4896	.4898	.4901	.4904	.4906	.4909	.4911	.4913	.4916
2.4	.4918	.4920	.4922	.4925	.4927	.4929	.4931	.4932	.4934	.4936
2.5	.4938	.4940	.4941	.4943	.4945	.4946	.4948	.4949	.4951	.4952
2.6	.4953	.4955	.4956	.4957	.4959	.4960	.4961	.4962	.4963	.4964
2.7	.4965	.4966	.4967	.4968	.4969	.4970	.4971	.4972	.4973	.4974
2.8	.4974	.4975	.4976	.4977	.4977	.4978	.4979	.4979	.4980	.4981
2.9	.4981	.4982	.4982	.4983	.4984	.4984	.4985	.4985	.4986	.4986
3.0	.4987	.4987	.4987	.4988	.4988	.4989	.4989	.4989	.4990	.4990

Note: This table gives the area in the right-hand tail of the distribution (i.e., $Z \geq 0$). But since the normal distribution is symmetrical about $Z = 0$, the area in the left-hand tail is the same as the area in the corresponding right-hand tail. For example. $P(-1.96 \leq Z \leq 0) = 0.4750$. Therefore, $P(-1.96 \leq Z \leq 1.96) = 2(0.4750) = 0.95$.

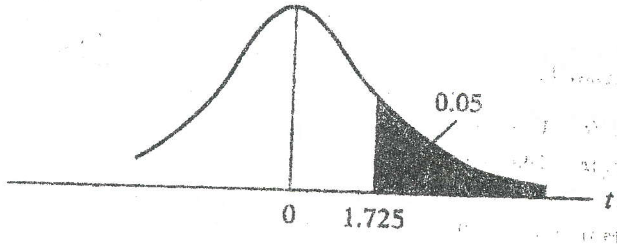
Percentage points of the *t* distribution

Example

$Pr(t > 2.086) = 0.025$

$Pr(t > 1.725) = 0.05$ for df = 20

$Pr(|t| > 1.725) = 0.10$



df \ Pr	0.25 0.50	0.10 0.20	0.05 0.10	0.025 0.05	0.01 0.02	0.005 0.010	0.001 0.002
1	1.000	3.078	6.314	12.706	31.821	63.657	318.31
2	0.816	1.886	2.920	4.303	6.965	9.925	22.327
3	0.765	1.638	2.353	3.182	4.541	5.841	10.214
4	0.741	1.533	2.132	2.776	3.747	4.604	7.173
5	0.727	1.476	2.015	2.571	3.365	4.032	5.893
6	0.718	1.440	1.943	2.447	3.143	3.707	5.208
7	0.711	1.415	1.895	2.365	2.998	3.499	4.785
8	0.706	1.397	1.860	2.306	2.896	3.355	4.501
9	0.703	1.383	1.833	2.262	2.821	3.250	4.297
10	0.700	1.372	1.812	2.228	2.764	3.169	4.144
11	0.697	1.363	1.796	2.201	2.718	3.106	4.025
12	0.695	1.356	1.782	2.179	2.681	3.055	3.930
13	0.694	1.350	1.771	2.160	2.650	3.012	3.852
14	0.692	1.345	1.761	2.145	2.624	2.977	3.787
15	0.691	1.341	1.753	2.131	2.602	2.947	3.733
16	0.690	1.337	1.746	2.120	2.583	2.921	3.686
17	0.689	1.333	1.740	2.110	2.567	2.898	3.646
18	0.688	1.330	1.734	2.101	2.552	2.878	3.610
19	0.688	1.328	1.729	2.093	2.539	2.861	3.579
20	0.687	1.325	1.725	2.086	2.528	2.845	3.552
21	0.686	1.323	1.721	2.080	2.518	2.831	3.527
22	0.686	1.321	1.717	2.074	2.508	2.819	3.505
23	0.685	1.319	1.714	2.069	2.500	2.807	3.485
24	0.685	1.318	1.711	2.064	2.492	2.797	3.467
25	0.684	1.316	1.708	2.060	2.485	2.787	3.450
26	0.684	1.315	1.706	2.056	2.479	2.779	3.435
27	0.684	1.314	1.703	2.052	2.473	2.771	3.421
28	0.683	1.313	1.701	2.048	2.467	2.763	3.408
29	0.683	1.311	1.699	2.045	2.462	2.756	3.396
30	0.683	1.310	1.697	2.042	2.457	2.750	3.385
40	0.681	1.303	1.684	2.021	2.423	2.704	3.307
60	0.679	1.296	1.671	2.000	2.390	2.660	3.232
120	0.677	1.289	1.658	1.980	2.358	2.617	3.160
∞	0.674	1.282	1.645	1.960	2.326	2.576	3.090

Note: The smaller probability shown at the head of each column is the area in one tail; the larger probability is the area in both tails.

Source: From E. S. Pearson and H. O. Hartley, eds., *Biometrika Tables for Statisticians*, vol. 1, 3d ed., table 12, Cambridge University Press, New York, 1966. Reproduced by permission of the editors and trustees of *Biometrika*.