EASTERN UNIVERSITY, SRI LANKA

FACULTY OF COMMERCE AND MANAGEMENT

THIRD YEAR FIRST SEMESTER EXAMINATION IN BBA/ B.COM – 2005/ 2006

(AUGUST 2007) PROPER/ REPEAT

DAF 3134 BUSINESS STATISTICS

Answer all questions Time : 03 Hours

01. (A) (i) Define the following terms.

(a) Random variable (b) Probability distribution

(ii) What are the conditions that must be satisfied by the probabilities in a probability distribution?

(iii) Leakage from underground petrol tanks at service stations can damage the environment. It is estimated that 25% of these tanks leak. 10 tanks are chosen at random, independently of each other, and examined.

(a) Explain briefly whether or not this experiment satisfies the conditions for a binomial distribution.

(b) Compute the following probabilities.

(i) None of the tanks has leak

(ii) Fewer than three leaks

(iii) At least nine leaks

(c) What is the expected number and standard deviation of leaking tanks?

(20 Marks)

(B) (i) Describe how to compute the mean and the standard deviation of a discrete random variable, and interpret what these quantities tell about the observed values of the random variable.

(ii) A company is trying to decide which of two product lines to select, both of which require the same rupees investment. The probabilities of market acceptance and the corresponding profits are shown in the following table :

Market acceptance	Probability	Line 1 (Rs)	Line 2 (Rs)
Poor	0.05	3000	2250
Below average	0.15	7500	3700
Average	0.60	24000	25500
Above average	0.15	40500	44250
Excellent	0.05	45000	48750

(a) Compute the expected value and the variance of the profits for each line.

(b) Which line has higher expected profit?

(c) Which line has the higher risk?

(d) Which line would you select? Explain, why?

02. (A) (i) What is meant by the standard normal distribution?

(ii) The amount of petrol, which an estate agent used in driving prospective buyers around the city to inspect home units, was recorded each week for 200 weeks. The amounts were found to follow an approximate normal distribution with a mean of 75 liters and standard deviation of 12 liters.

(a) Find the probability that the fuel consumption was more than 70 liters.

(b) Find the probability that the fuel consumption was less than 60 liters.

(c) Estimate the number of weeks the fuel consumption was less than 6 liters.

(d) What the fuel consumption was exceeded in only 20 out of the 200 weeks

(B) (i) State the sampling distribution clearly for the statistics given below.

(a) Sample mean

(b) Sample proportion

(c) Difference between the two sample means.

(ii) (a) The population distribution of gripping strengths of industrial workers is known to have a mean of 110 and standard deviation of 10. A sample of 75 workers is selected randomly.

(i) What is the distribution of the sample mean grapping strength?

(ii) Compute the probability that the sample mean gripping strength will be between 109 and 112.

(b) The Laurier company's brand has a market share of 30%. In a survey 1000 consumers were asked which brand they prefer. What is the probability that more than 32% of all the respondents say prefer the Laurier brand?

(20 Marks)

03. (A) (i) Explain the difference between a population parameter and its point estimate.

(ii) Draw a diagram representing the sampling distribution of two unbiased estimators, one of which is relatively efficient.

(iii) Explain why it is important to calculate a confidence interval in addition to calculating a point estimate of a population parameter.

(B) A financial controller is interested in the number of defective items produced each hour by a machine in the factory. A random sample of 25 hours produced the following number of defectives per hours.

10	6	5	7	8	4	5	5	6	8
4	3	7	8	4	10	5	3	2	0
5	8	9	3	7					

(i) Find a point estimate of the population mean number of defectives that the machines produces per hour.

(ii) Find the standard error of the estimate in (i).

(iii) Construct a 95% confidence interval estimate for the mean number of defectives that the machine produces per hour? What is the assumption you made?

(iv) Interpret what the interval estimate in (iii) tells you.

(v) If you were told that the population standard deviation of the number of defectives that the machine produced per hour was 1.80, would it change your answer to part (iii)?

If so, find the new answer. If not explain why?

(20 Marks)

04. (A) (i) Define the following terms in the context of correlation and regression.

(a) Product moment correlation coefficient

(b) Partial correlation coefficient

(c) Independent variable

(d) Dependent variable

(B) (i) What are the assumptions of a simple linear regression model?

(ii) What is the purpose of the error term in regression?

(iii) The following data represent yearly sales volume (y) and the advertising expenditures (x) of a carpet manufacturing firm :

Year	Advertising Expenditure (Rs. millions)	Sales volume (Rs. millions)
1981	1.8	26
1982	2.3	31
1983	2.6	28
1984	2.4	30
1985	2.8	34
1986	3.0	38
1987	3.4	41
1988	3.2	44
1989	3.6	40
1990	3.8	43

(a) Draw a scatter diagram of y against x.

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

(c) Calculate the correlation coefficient and comment on it.

(d) Obtain the linear least squares regression line of y on x for the data (work to two decimal places) and sketch it on your diagram for (a).

(e) Estimate the sales volume for advertising expenditure 2.5 million.

(20 Marks)

05. (A) Define the following terms in the context of hypothesis testing.

(i) Hypothesis testing

(ii) Rejection region

(iii) Level of significance

(iv) p – value

(B) The researchers surveyed a total of 1165 business owners, of which 115 are women. Of the 115 women who asked for a loan, 14 were turned down. A total of 98 men who asked for a loan were rejected.

(i) Conduct a test at 5% level of significance to determine whether the proportion of loans denied is greater for women than for men.

(ii) The table given below provides statistics regarding the rates of interest for all loans that were granted.

Statistics	Interest rate	
	Women	Men
Mean	1.55	1.28
Standard deviation	0.64	0.67
Sample size	101	952

(a) Test whether there is any difference for the interest rates of loans that were granted, between men and women.

(b) What are the assumptions you made to conduct the test in part (a)?

(iii) Using the information given below;

Correlation of rates and sales	$= -0.27$
p – value	= 0.0000

(a) Interpret the relationship between annual gross sales and interest rates;

(b) Test whether banks grant lower interest rates to business with higher sales.

(iv) Using the results obtained in part (a) and (b), discuss whether banks discriminate against women business owners.

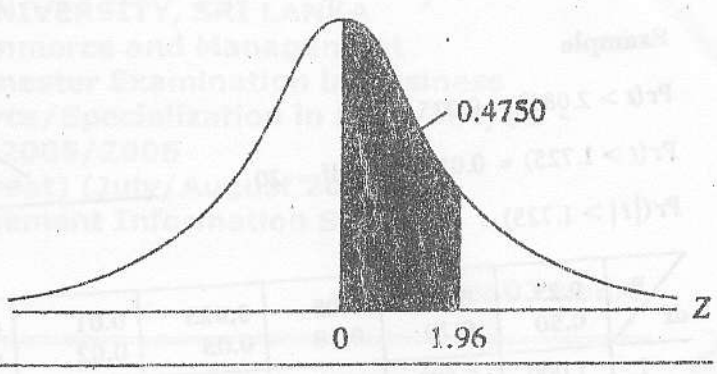
(20 Marks)

eas un er the standardized normal distribution

ample

$(0 \le Z \le 1.96) = 0.4750$

$(Z \ge 1.96) = 0.5 - 0.4750 = 0.025$



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

e: This table gives the area in the right-hand tail of the distribution (i.e., $Z \ge 0$). 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 t-hand tail. For example. $P(-1.96 \le Z \le 0) = 0.4750$. Therefore, $P(-1.96 \le Z \le 1.96) = 2(0.4750) = 0.95$.

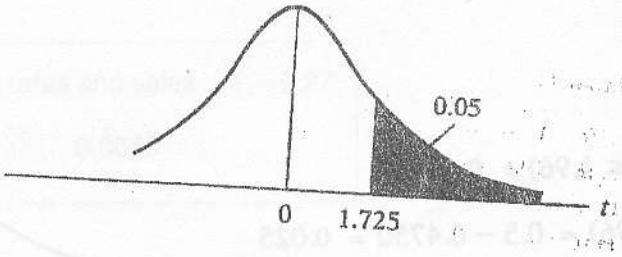
Percentage points of the t distribution

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*.