

2 FEB 2018



EASTERN UNIVERSITY, SRI LANKA
FACULTY OF COMMERCE AND MANAGEMENT
SECOND YEAR SECOND SEMESTER EXAMINATION IN
DEPARTMENT OF BUSINESS ADMINISTRATION/ BACHELOR OF COMMERCE - 2014/2015 (July 2017)
(PROPER/ REPEAT/ RE-REPEAT)
COM 2053 BUSINESS STATISTICS

All Questions.
Answers permitted.

Time: 03 Hours

Choose the letter of the best choice for each question in the given answer script:

A sample of 400 households in the given region is selected and several variables are recorded. Which of the following statements is correct?

- a) Total household income (in rupees) is interval level data.
- b) Socioeconomic status recorded as "low income", "middle income", or "high income" is nominal level data.
- c) The number of people living in a household is a discrete variable.
- d) The primary language spoken in the household is ordinal level data.
- e) None of the above.

Which of the following exhibits the correct relationship between variables and values?

- a) Red; price of an automobile
- b) Color of an automobile; price of an automobile
- c) Red; Rs.23,000
- d) Color of an automobile; Rs.23,000
- e) Rs. 23000; Color of an automobile

Variables in which measurement is always approximate because they permit an unlimited number of intermediate values are

- a) Nominal
- b) Discrete
- c) Ordinal
- d) Continuous
- e) Interval

Which of the following would be most suitable for displaying the proportions of a city's budget spent on different items?

- a) Pie chart
- b) Bar chart
- c) Line graph
- d) Histogram
- e) Scatter plot

When constructing a frequency distribution, which of the following rules must be followed?

- a) The midpoint of each class must be an integer.
- b) The width of each class is equal to the lowest value in the data set.
- c) Adjacent classes cannot overlap.
- d) The number of classes must be an even number.
- e) The lower limit of the first class interval is equal to the lowest value in the data set

Questions from 6 to 8 are based on the following grouped frequency table of the income, X , of 30 employees at a local small business (in Rs.1000s).

Income	$26 < X \leq 28$	$28 < X \leq 30$	$30 < X \leq 32$	$32 < X \leq 34$	$34 < X \leq 36$
Frequency	2	11	8	5	4

- 6) The relative cumulative frequency of the $28 < X \leq 30$ class is
- a) 11 b) 0.43 c) 0.06 d) 13 e) 0.7
- 7) The class that contains the 80th percentile is
- a) $26 < X \leq 28$ b) $28 < X \leq 30$ c) $30 < X \leq 32$ d) $32 < X \leq 34$ e) $34 < X \leq 36$
- 8) If the boss' income (the "31st employee") is Rs. 250,000, the mean income for all 31 workers is approximately equal to
- a) Rs. 8,000 b) Rs. 30,000 c) Rs. 38,000 d) Rs. 140,000 e) Rs. 220,000
- 9) Which of the following is not a condition of a binomial probability distribution?
- a) Only two possible outcomes b) Constant probability of success
 c) Must have at least three trials d) Independent trials
 e) None of the above
- 10) If $P(-2 < Z < k) = .6$, where Z is a standard normal random variable, then k is
- a) 0.5773 b) 0.195 c) 0.73 d) 0.55 e) -0.40
- 11) The mean and the variance are equal in
- a) all probability distributions b) the binomial distribution
 c) the Poisson distribution d) the uniform distribution
 e) the student -t distribution
- 12) A probability function is a rule of correspondence or equation that
- a) finds the mean value of the random variable
 b) assigns values of a random variable to the events of a probability experiment
 c) assigns probabilities to the various values of a random variable
 d) defines the variability in the experiment
 e) none of the above is correct
- 13) The normal approximation to the binomial is used when
- a) the sample size is at least 30 b) both np and $np(1-p)$ are at least 5
 c) the mean and the variance are the same d) the z value is greater than 0
 e) none of the above
- 14) Which of the following statements about hypothesis tests is **incorrect**?
- a) Null hypothesis must always include equality
 b) In a one-tailed test, alternative hypothesis involves either " $>$ " or " $<$ ", but not " \neq ."
 c) If the test statistic lies in the rejection region, we reject the claim.
 d) The claim can be either null hypothesis or alternative hypothesis
 e) The critical value indicates the start of the rejection region

A sample of twenty-five observations is taken from a normal population with variance 9. 90% confidence limits corresponding to a sample mean of 30 are best represented by

- a) 30 ± 9.00 b) $30 \pm .79$ c) 30 ± 1.03 d) $30 \pm .47$ e) $30 \pm .99$

Which of the following is a true statement regarding the comparison of t-distributions to the standard normal distribution?

- a) The normal distribution is symmetrical whereas the t-distribution is slightly skewed.
- b) The proportion of area beyond a specific value of t is less than the proportion of area beyond the corresponding value of z.
- c) The greater the df, the more the t-distributions resemble the standard normal distribution
- d) The mean of the standard normal is zero whereas the mean of the t-distribution is greater than zero
- e) All of the above.

The size of a confidence interval for a mean is affected by changes in which of the following?

- a) The size of the sample
- b) The confidence coefficient
- c) The variance of the sample
- d) b and c only
- e) a, b, and c

The distribution of means of all possible samples of the same size n drawn from a population will approximate the normal curve if

- a) the n is large enough.
- b) the population is large.
- c) the population is symmetrical.
- d) the mean of each sample equals the mean of the population.
- e) none of the above is correct.

sampling distribution

- a) is a distribution of all the various sample statistics that can be found for one sample.
- b) of the mean is a distribution of the means taken from all possible samples of a given size n that could be taken from the population.
- c) of any statistic has an approximately normal distribution.
- d) is a histogram showing the distribution of the sample.
- e) all of the above are correct.

What test statistic is used for a global test of significance in the regression analysis?

- a) Z test b) t-test c) Chi-square test d) F test e) p test

Spearman's co-efficient of rank correlation is equal to one, then

- a) the rankings of the two variables totally agree
- b) all the total variation is explained by the regression line
- c) the rankings of the two variables is totally different
- d) the rankings of the two variables partially agree
- e) none of the above is correct

- 22) What does the slope represent in the simple linear regression analysis?
- The predicted value of Y when $X = 0$
 - The estimated change in average Y per unit change in X
 - The predicted value of Y
 - The variation around the line of regression
 - All the above are correct
- 23) Finding the centered four-quarter moving average helps us to identify the
- trend component
 - cyclical component
 - seasonal component
 - irregular component
 - none of the above
- 24) In time series seasonal variations can occur within a period of
- One year
 - Four years
 - Nine years
 - Three years
 - Five years
- 25) Variations due to floods, droughts, strikes, fires and political disturbances are:
- Trend
 - Seasonal
 - Cyclical
 - Irregular
 - All of the above

02. I) Thousands of customers have accounts at a large department store. An accountant claims average unpaid balance for these accounts is Rs. 7500, a figure obtained by computing average of the unpaid balances for 50 of the accounts.

- Identify the population and its parameter.
- What is the sample?
- Is the figure of Rs. 7500 a parameter or a statistic? Explain it.

II) Thirty adults were asked which of the following conveniences they would find most difficult without: television (T), refrigerator (R), air conditioning (A), public transportation (P), or (M). Their responses are listed below.

R	A	R	P	P	T	R	M	P	A
A	R	R	T	P	P	T	R	A	A
R	P	A	T	R	P	R	A	P	R

- Prepare a frequency distribution table.
- Calculate the relative frequencies and percentages for all categories.
- What percentage of these adults named refrigerator or air conditioning find most difficult without?
- Would find most difficult to do without?
- Draw a bar graph for the relative frequency distribution

III) A department store manager is interested in the number of complaints received by the customer department about the quality of electrical products sold by the store. Records over a 10-week yield the data shown in the table.

Week	1	2	3	4	5	6	7	8	9	10
No. of Complaints	13	15	08	16	08	04	21	11	03	15

- Find the mean, median and mode of complaints.
- Find the variance, standard deviation, range and interquartile range.
- Based on the measures calculated in the above two parts, what would you conclude about the complaints made on electrical products sold by the store?

(20 Marks)

An economist wishes to estimate the total cost of a project to offer a proper price for it. She values its job on a fixed quantity of Rs. 12000 and a variable quantity of Rs. 300 per day of work. It is known that the job will take between 7 and 11 days according to the following probability function for $X =$ "number of days that the job will take"

x	7	8	9	10	11
$P(X = x)$	0.10	0.20	0.30	0.30	0.10

- Compute the probability that the project takes 9 or 10 days.
- Compute the mean number of days the project will take.
- Find the standard deviation for the number of days the project will take and interpret this value.
- Determine the expected cost of the project and its standard deviation.

A popular soft drink is sold in 2-liter (2000- milliliters) bottles. Because of variation in the filling process, bottles have a mean of 2000 milliliters and a standard deviation of 20 milliliters, normally distributed.

- If the process fills the bottle by more than 60 milliliters, the overflow will cause a machine malfunction. What is the probability of this occurring?
- If the manufacturer samples 100 bottles, what is the probability that the mean is less than 1950 milliliters?

The numbers of accidents in a production facility has a Poisson distribution with mean 2.6 per month.

- For a given month, what is the probability there will be fewer than two accidents?
- What is the probability that there will be exactly 10 accidents next year?

Every day 5 units are selected in an industrial selection process. The quality control department stops the production process if two or more of the samples of five units have defects. If the percentage of defective units is $p = 0.05$,

- Compute the probability that the process is stopped.
- Find the expected value and the standard deviation of defective units

(18 Marks)

A researcher works for a business regulatory agency, finds that in a sample of 1762 people in a particular country 1004 of them believe that government regulation of business does more harm than good.

- Construct a 95% confidence interval for the proportion of people who believe that government regulation of business does more harm than good and interpret it.

- b) Based on your answer in part (a), what would you conclude about the claim that 60% of people believe that government regulation of business does more harm than good at 5% significance?

II) An insurance company wants to know if the average speed at which men drive cars is greater than that of women drivers. The company took a random sample of 27 cars driven by men on a highway and found the mean speed to be 72 km per hour with a standard deviation of 2.2 km per hour. A random sample of 18 cars driven by women on the same highway gave a mean speed of 68 km per hour with a standard deviation of 2.5 km per hour.

- What is the point estimate of the difference between the mean speeds of cars driven by all men and all women on this highway and interpret it.
- Construct a 95% confidence interval for the difference between the mean speeds of cars driven by all men and all women on this highway and interpret it.
- Testing at the 1% significance level, can you conclude that the mean speed of cars driven by all men drivers on this highway is greater than that of cars driven by all women drivers?
- State the assumptions you made to solve this problem.

05. I) A random sample of eight drivers insured with a company and having similar minimum requirements for insurance policies was selected. The following table lists their driving experiences (in years) and their monthly auto insurance premiums (in Rs. 100s).

Driving Experience (years)	Monthly Auto Insurance Premium (in Rs. 100s)
5	64
2	87
12	50
9	71
15	44
6	56
25	42
16	60

- Does the insurance premium depend on the driving experience, or does the driving experience depend on the insurance premium? Do you expect a positive or a negative relationship between these two variables?
- Compute the correlation coefficient and interpret it.
- Calculate the coefficient of determination and interpret its value based on the given problem.
- Estimate the least squares regression equation in an attempt to predict the monthly auto insurance premium by the years of driving.
- Predict the monthly auto insurance premium for a driver with 10 years of driving.

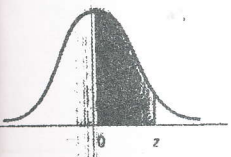
II) Galaxy International manufactures and sells toys all around the world. Management accountants of the company wish to determine seasonal indexes for the quarterly data on revenue. The following table shows the quarterly sales for Galaxy International for the years 2014 through 2017. The sales are reported in millions of rupees.

Year	Sales			
	Q 1	Q 2	Q 3	Q 4
2014	30	50	70	25
2015	40	58	88	42
2016	62	75	105	45

- Calculate the centered four-point moving average for the given time series
Determine the seasonal index for each of the four quarters using the ratio to moving average method.
Forecast the demand for the four quarters of 2017 using trend forecasts of 65, 80, 95 and 52.

(20 Marks)

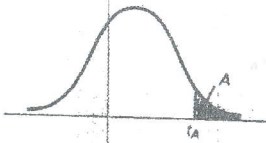
Normal Probabilities



	.00	.01	.02	.03	.04	.05	.06	.07	.08	.09
0	.0000	.0040	.0080	.0120	.0160	.0199	.0239	.0279	.0319	.0359
1	.0398	.0438	.0478	.0517	.0557	.0596	.0636	.0675	.0714	.0753
2	.0793	.0832	.0871	.0910	.0948	.0987	.1026	.1064	.1103	.1141
3	.1179	.1217	.1255	.1293	.1331	.1368	.1406	.1443	.1480	.1517
4	.1554	.1591	.1628	.1664	.1700	.1736	.1772	.1808	.1844	.1879
5	.1915	.1950	.1985	.2019	.2054	.2088	.2123	.2157	.2190	.2224
6	.2257	.2291	.2324	.2357	.2389	.2422	.2454	.2486	.2517	.2549
7	.2580	.2611	.2642	.2673	.2704	.2734	.2764	.2794	.2823	.2852
8	.2881	.2910	.2939	.2967	.2995	.3023	.3051	.3078	.3106	.3133
9	.3159	.3186	.3212	.3238	.3264	.3289	.3315	.3340	.3365	.3389
0	.3413	.3438	.3461	.3485	.3508	.3531	.3554	.3577	.3599	.3621
1	.3643	.3665	.3686	.3708	.3729	.3749	.3770	.3790	.3810	.3830
2	.3849	.3869	.3888	.3907	.3925	.3944	.3962	.3980	.3997	.4015
3	.4032	.4049	.4066	.4082	.4099	.4115	.4131	.4147	.4162	.4177
4	.4192	.4207	.4222	.4236	.4251	.4265	.4279	.4292	.4306	.4319
5	.4332	.4345	.4357	.4370	.4382	.4394	.4406	.4418	.4429	.4441
6	.4452	.4463	.4474	.4484	.4495	.4505	.4515	.4525	.4535	.4545
7	.4554	.4564	.4573	.4582	.4591	.4599	.4608	.4616	.4625	.4633
8	.4641	.4649	.4656	.4664	.4671	.4678	.4686	.4693	.4699	.4706
9	.4713	.4719	.4726	.4732	.4738	.4744	.4750	.4756	.4761	.4767
0	.4772	.4778	.4783	.4788	.4793	.4798	.4803	.4808	.4812	.4817
1	.4821	.4826	.4830	.4834	.4838	.4842	.4846	.4850	.4854	.4857
2	.4861	.4864	.4868	.4871	.4875	.4878	.4881	.4884	.4887	.4890
3	.4893	.4896	.4898	.4901	.4904	.4906	.4909	.4911	.4913	.4916
4	.4918	.4920	.4922	.4925	.4927	.4929	.4931	.4932	.4934	.4936
5	.4938	.4940	.4941	.4943	.4945	.4946	.4948	.4949	.4951	.4952
6	.4953	.4955	.4956	.4957	.4959	.4960	.4961	.4962	.4963	.4964
7	.4965	.4966	.4967	.4968	.4969	.4970	.4971	.4972	.4973	.4974
8	.4974	.4975	.4976	.4977	.4977	.4978	.4979	.4979	.4980	.4981
9	.4981	.4982	.4982	.4983	.4984	.4984	.4985	.4985	.4986	.4986
0	.4987	.4987	.4987	.4988	.4988	.4989	.4989	.4989	.4990	.4990

SOURCE: Abridged from Table 1 of A. Hald, *Statistical Tables and Formulas* (New York: Wiley & Sons, Inc.), 1952. Reproduced by permission of A. Hald and the publisher, John Wiley & Sons, Inc.

Critical Values of t



DEGREES OF FREEDOM						DEGREES OF FREEDOM			
	$t_{.100}$	$t_{.050}$	$t_{.025}$	$t_{.010}$	$t_{.005}$		$t_{.100}$	$t_{.050}$	$t_{.025}$
1	3.078	6.314	12.706	31.821	63.657	24	1.318	1.711	2.064
2	1.886	2.920	4.303	6.965	9.925	25	1.316	1.708	2.060
3	1.638	2.353	3.182	4.541	5.841	26	1.315	1.706	2.056
4	1.533	2.132	2.776	3.747	4.604	27	1.314	1.703	2.052
5	1.476	2.015	2.571	3.365	4.032	28	1.313	1.701	2.048
6	1.440	1.943	2.447	3.143	3.707	29	1.311	1.699	2.045
7	1.415	1.895	2.365	2.998	3.499	30	1.310	1.697	2.042
8	1.397	1.860	2.306	2.896	3.355	35	1.306	1.690	2.030
9	1.383	1.833	2.262	2.821	3.250	40	1.303	1.684	2.021
10	1.372	1.812	2.228	2.764	3.169	45	1.301	1.679	2.014
11	1.363	1.796	2.201	2.718	3.106	50	1.299	1.676	2.009
12	1.356	1.782	2.179	2.681	3.055	60	1.296	1.671	2.000
13	1.350	1.771	2.160	2.650	3.012	70	1.294	1.667	1.994
14	1.345	1.761	2.145	2.624	2.977	80	1.292	1.664	1.990
15	1.341	1.753	2.131	2.602	2.947	90	1.291	1.662	1.987
16	1.337	1.746	2.120	2.583	2.921	100	1.290	1.660	1.984
17	1.333	1.740	2.110	2.567	2.898	120	1.289	1.658	1.980
18	1.330	1.734	2.101	2.552	2.878	140	1.288	1.656	1.977
19	1.328	1.729	2.093	2.539	2.861	160	1.287	1.654	1.975
20	1.325	1.725	2.086	2.528	2.845	180	1.286	1.653	1.973
21	1.323	1.721	2.080	2.518	2.831	200	1.286	1.653	1.972
22	1.321	1.717	2.074	2.508	2.819	∞	1.282	1.645	1.960
23	1.319	1.714	2.069	2.500	2.807				

SOURCE: From M. Merrington, "Table of Percentage Points of the t -Distribution," *Biometrika* 32 (1941): 300. Reproduced by the Biometrika Trustees.