



PRE-Ph.D. COURSE WORK
Subject: Quantitative Techniques(QT)
Subject Code: MSP1003
March 2022
Theory (External): 70 Marks
Time: 03 Hours

Instructions to the Students

1. This Question paper consists of two Sections. All sections are compulsory.
2. Section A comprises 10 questions of objective type in nature. All questions are compulsory. Each question carries 2 marks.
3. Section B comprises 8 essay type questions out of which students need to do any 5. Each question carries 10 marks.
4. Read the questions carefully and write the answers in the answer sheets provided.
5. Do not write anything on the question paper.
6. Wherever necessary, the diagram drawn should be neat and properly labelled

Roll Number											

SECTION –A (SHORT/OBJECTIVE TYPE QUESTIONS)
(10x2=20 Marks)

- A. A survey of 430 business travelers found 155 business travelers used a travel agent to make the travel arrangements (USA Today, November 20, 2003). [2]
- a. Develop a descriptive statistic that can be used to estimate the percentage of all business travelers who use a travel agent to make travel arrangements.
 - b. The survey reported that the most frequent way business travelers make travel arrangements is by using an online travel site. If 44% of business travelers surveyed made their arrangements this way, how many of the 430, business travelers used an online travel site?
- B. Mention at least two types of graphical methods used for each of qualitative and quantitative data in descriptive statistics. [2]
- C. Consider a sample with data values of 27, 25, 20, 15, 30, 34, 28, and 25. Compute the 25th, 50th, 75th and 90th percentiles. [2]
- D. A simple random sample of size 100 is selected from a population with $p=.40$. [2]
- a. What is the expected value of population proportion?
 - b. What is the standard error of population proportion?
- E. In an effort to estimate the mean amount spent per customer for dinner at a major Atlanta restaurant, data were collected for a sample of 49 customers. Assume a population standard deviation of \$5. At 95% confidence, what is the margin of error? [2]
- F. A sample of 20 items provides a sample standard deviation of 5. Compute the 90% confidence interval estimate of the population variance. [2]
- G. An automotive part must be machined to close tolerances to be acceptable to customers. Production specifications call for a maximum variance in the lengths of the parts of .0004. Suppose the sample variance



for 30 parts turns out to be $s_2 = .0005$. Use $\alpha = .05$ to test whether the population variance specification is being violated. [2]

H. A 95% confidence interval for a population mean was reported to be 152 to 160. If $\sigma = 15$, what sample size was used in this study? [2]

I. Suppose we have a multinomial population with four categories: A, B, C, and D. The null hypothesis is that the proportion of items is the same in every category. The null hypothesis is [2]

$$H_0: p_A = p_B = p_C = p_D = 0.25$$

A sample of size 300 yielded the following results

A: 85 B: 95 C: 50 D: 70

Use $\alpha = .05$ to determine whether H_0 should be rejected. What is the p-value?

J. The estimated regression equation for a model involving two independent variables and 10 observations follows. [2]

$$\hat{y} = 29.1270 + .5906x_1 + 0.4980x_2$$

a. Interpret b_1 and b_2 in this estimated regression equation.

b. Estimate y when $x_1 = 180$ and $x_2 = 310$.

SECTION – B (ESSAY TYPE QUESTIONS)

(10 × 5 = 50)

Q. 1

(i) A doctor's office staff studied the waiting times for patients who arrive at the office with a request for emergency service. The following data with waiting times in minutes were collected over a one-month period. [5]

2	5	10	12	4	4	5	17	11	8	9
	8	12	21	6	8	7	13	18	3	

Use classes of 0–4, 5–9, and so on in the following:

a. Show the frequency distribution.

b. Show the relative frequency distribution.

c. Show the cumulative frequency distribution.

d. Show the cumulative relative frequency distribution.

e. What proportion of patients needing emergency service wait 9 minutes or less?

- (ii) Consider the following two frequency distributions. The first frequency distribution provides an approximation of the annual adjusted gross income in the United States (Internal Revenue Service, March 2003). The second frequency distribution shows exam scores for students in a college statistics course.

Income (\$1000s)	Frequency (millions)	Exam score	Frequency
0-24	60	20-29	2
25-49	33	30-39	5
50-74	20	40-49	6
75-99	6	50-59	13
100-124	4	60-69	32
125-149	2	70-79	78
150-174	1	80-89	43
175-199	1	90-99	21
Total	127		200

- Develop a histogram for the annual income data. What evidence of skewness does it show? Does this skewness make sense? Explain.
- Develop a histogram for the exam score data. What evidence of skewness does it show? Explain. [5]

Q. 2

Ebby Halliday Realtors provide advertisements for distinctive properties and estates located throughout the United States. The prices listed for 22 distinctive properties and estates are shown here (The Wall Street Journal, January 16, 2004). Prices are in thousands. [10]

1500 700 2995 895 619 880 719 725 3100
 619 739 1699 625 799 1120 4450 2495 1250
 2200 1395 912 1280

- Provide a five-number summary.
- Compute the lower and upper limits.
- The highest priced property, \$4,450,000, is listed as an estate overlooking White Rock Lake in Dallas, Texas. Should this property be considered an outlier? Explain.
- Should the second highest priced property, listed for \$3,100,000, be considered an outlier? Explain.
- Show a box plot.

Q. 3

(i) The population proportion is .30. What is the probability that a sample proportion will be within $\pm .04$ of the population proportion for each of the following sample sizes? [5]

a. $n = 100$

b. $n = 200$

c. $n = 500$

d. $n = 1000$

e. What is the advantage of a larger sample size?

(ii) Consider a sample with a mean of 30 and a standard deviation of 5. Use Chebyshev's theorem to determine the percentage of the data within each of the following ranges. [5]

a. 20 to 40

b. 5 to 45

c. 22 to 38

d. 18 to 42

e. 12 to 48

Q. 4 Playbill magazine reported that the mean annual household income of its readers is \$119,155 (Playbill, January 2006). Assume this estimate of the mean annual household income is based on a sample of 80 households, and based on past studies, the population standard deviation is known to be $\sigma = \$30,000$. [10]

a. Calculate the standard error?

b. Develop a 90% confidence interval estimate of the population mean.

c. Develop a 95% confidence interval estimate of the population mean.

d. Develop a 99% confidence interval estimate of the population mean.

e. Discuss what happens to the width of the confidence interval as the confidence level is increased. Does this result seem reasonable? Explain.

Q. 5

i. Consider the following hypothesis test: [5]

$$H_0 : \mu = 15$$

$$H_a : \mu \neq 15$$

A sample of 50 provided a sample mean of 14.15. The population standard deviation is 3.

b. What are the implications of your statistical conclusions in terms of driving safety recommendations?

ii. The following table reports prices and usage quantities for two items in 2004 and 2006. [5]

Item	Quantity		Unit Price (\$)	
	2004	2006	2004	2006
A	1500	1800	7.50	7.75
B	2	1	630.00	1500.00

- Compute price relatives for each item in 2006 using 2004 as the base period.
- Compute an unweighted aggregate price index for the two items in 2006 using 2004 as the base period.
- Compute a weighted aggregate price index for the two items using the Laspeyres method.
- Compute a weighted aggregate price index for the two items using the Paasche method.

Q. 7

i. The enrolment data (in thousands) for a state college over the past six years are shown. [5]

Year	1	2	3	4	5	6
Enrolment	20.5	20.2	19.5	19.0	19.1	18.8

Develop the equation for the linear trend component of this time series. Comment on what is happening to enrolment at this institution?

ii For the data given below: [5]

x_i :	1	2	3	4	5
y_i :	3	7	5	11	14

The estimated regression equation for these data is $\hat{y} = .20 + 2.60x$

- Compute SSE, SST, and SSR.
- Compute the coefficient of determination r^2 . Comment on the goodness of fit.
- Compute the sample correlation coefficient.

- a. Compute the value of the test statistic.
- b. What is the p-value?
- c. At $\alpha = .05$, what is your conclusion?
- d. At $\alpha = .01$, what is your conclusion?
- e. What is the rejection rule using the critical value? What is your conclusion

ii. Joan's Nursery specializes in custom-designed landscaping for residential areas. The estimated labor cost associated with a particular landscaping proposal is based on the number of plantings of trees, shrubs, and so on to be used for the project. For cost-estimating purposes, managers use two hours of labor time for the planting of a medium-sized tree. [5]

Actual times from a sample of 10 plantings during the past month follow (times in hours).

1.7 1.5 2.6 2.2 2.4 2.3 2.6 3.0 1.4 2.3

With a .05 level of significance, test to see whether the mean tree-planting time differs from two hours.

- a. State the null and alternative hypotheses.
- b. Compute the sample mean.
- c. Compute the sample standard deviation.
- d. What is the p-value?
- e. What is your conclusion?

Q. 6

i. A research hypothesis is that the variance of stopping distances of automobiles on wet pavement is substantially greater than the variance of stopping distances of automobiles on dry pavement. In the research study, 16 automobiles travelling at the same speeds are tested for stopping distances on wet pavement and then tested for stopping distances on dry pavement. On wet pavement, the standard deviation of stopping distances is 32 feet. On dry pavement, the standard deviation is 16 feet. [5]

- a. At a .05 level of significance, do the sample data justify the conclusion that the variance in stopping distances on wet pavement is greater than the variance in stopping distances on dry pavement? What is the p-value?

Q. 8 The following data is given about the factors in a system.

[10]

Factor Number	Eigen value	% of Variance
1	1.067	13.336
2	3.057	38.206
3	0.736	9.205
4	0.958	11.980
5	0.622	7.770
6	0.571	7.135
7	0.543	6.788
8	0.446	5.580

- Which four factors should be discarded if only four factors are to be retained? Give reason in support of your answer.
- Which factors should be retained to explain 70% of variance?
- Which factors should be retained to explain 90% of variance?
- Draw the Scree plot of the given component matrix.
- Which factors are to be discarded in Scree plot?

END OF PAPER