

THE HONG KONG POLYTECHNIC UNIVERSITY

Department of Applied Mathematics

Examination

Subject Code: AMA1501/AMA1502 Subject Title: Introduction to Statistics for
Business / Introduction to
Statistics

Session: Semester 1, 2017/2018

Date: 8 December 2017 Time: 8:45 a.m. – 11:45 a.m.

Time Allowed: 3 Hours

This question paper has 13 pages (attachments included).

Instructions to Candidates: This question paper has SIX questions.
Attempt FIVE questions.
All questions carry equal marks.

Attachments: Standard Normal Distribution Table, *t*-distribution Table,
 χ^2 -distribution Table, F-distribution Table, Formula Sheets

Subject Examiner: Dr. Chi Kin Chan
Dr. Benedict Wong

DO NOT TURN OVER THE PAGE UNTIL YOU ARE TOLD TO DO SO

1. The sales manager of a publisher investigates the customers' monthly expenses in purchasing supplementary exercises for primary school students through an online shop. The amounts of expenses in the last month for a random sample of customers are listed below:

Amount of expenses (\$)	Number of customers
50 to less than 150	4
150 to less than 200	7
200 to less than 250	15
250 to less than 300	24
300 to less than 350	17
350 to less than 400	11
400 to less than 450	6

- (a) Calculate the mean, mode and standard deviation of monthly expenses. (7 marks)
- (b) Estimate, from the frequency distribution table, the inter-quartile range of monthly expenses. (5 marks)
- (c) Estimate, from the frequency distribution table, the proportion of customers having monthly expenses between \$240 and \$380. (4 marks)
- (d) Construct a 98% confidence interval for the proportion of all customers who spent between \$240 and \$380 in the last month. (4 marks)
2. (a) A manager has 2 jobs of category A, 4 jobs of category B and 6 jobs of category C. He would like to assign them to 2 assistants at random in a way that they will work on the same number of jobs in each of those categories. How many possible ways are there? (4 marks)
- (b) From the sales records, 65% of customers purchased clothes, 38% of customers purchased small electricity appliances. Among customers who purchased clothes, 42% of them purchased small electricity appliances.
- (i) Calculate the probability that a randomly selected customer purchased clothes or small electricity appliances. (3 marks)
- (ii) Calculate the probability that a purchaser of small electricity appliances did not purchase clothes. (3 marks)
- (iii) It is known that a randomly selected customer did not purchase clothes, what is the probability that he/she did not purchase small electricity appliances? (4 marks)

(Question 2 Cont'd)

- (c) A manufacturing company has 3, 4 and 5 production lines in City A, B and C, respectively. From past records, the defective rates among products produced by production lines in City A, B and C are 5%, 4.5% and 4%, respectively. A production line is selected at random, in which twenty products are selected. If there is one defective product, calculate the probability that this production line is located in City B. (6 marks)
3. (a) The percentage of revenue spent on research and development among the companies of innovative technology in City A follows the normal distribution approximately with a mean of 60% and a standard deviation of 12%.
- (i) Calculate the probability that a randomly selected company of innovative technology in City A spends at least 45% of revenue on research and development. (3 marks)
- (ii) A random sample of one hundred companies of innovative technology in City A is selected. Calculate the probability that at least half of them spend at most 65.4% of revenue on research and development. (6 marks)
- (iii) Five companies of innovative technology in City A are selected at random. Calculate the probability that the minimum percentage of revenue spent on research and development, out of the five, is less than 42%. (4 marks)
- (b) The number of click-throughs received by a website in a 5-minute period follows the Poisson distribution with a mean of 8.
- (i) Calculate the probability that in a randomly selected 5-minute period, there are at least 5 click-throughs received. (4 marks)
- (ii) It is known that there are at least 5 click-throughs received in a randomly selected 5-minute period, find the probability that the number of click-throughs received is more than 8. (3 marks)
4. (a) The IQ scores of students of schools A and B follow normal distribution approximately. Mean scores for school A and school B students are 120 and 132, respectively; and standard deviations are 12 and 13, respectively. Random sample of size 10 is selected from each of the schools, calculate the probability that the difference in the mean IQ scores of the two samples is within 10. (4 marks)
- (b) An analyst studies the pollutant concentration of NO_2 ($\mu\text{g}/\text{m}^3$) in a certain location from 9:00am to 8:00pm in a day. The hourly observations in a randomly selected day are listed below:
- | | | | | | |
|------|------|------|------|------|------|
| 57.2 | 63.0 | 54.5 | 63.3 | 62.6 | 60.9 |
| 67.1 | 71.0 | 73.6 | 84.7 | 83.7 | 71.2 |
- Construct a 95% confidence interval for the mean concentration of NO_2 . Assume that the concentrations of NO_2 are normally distributed. (5 marks)
- (c) The director of human resources of a company would like to produce a point estimate for the population proportion of salespersons who favour a certain bonus system. Determine the sample size required if the director wishes to be at least 95% confident that the error of his estimate will not be more than 0.08. (5 marks)

(Question 4 Cont'd)

- (d) In comparing the job satisfaction level among employees who worked for the company at most 3 years and those worked for the company more than 15 years, job satisfaction indices are evaluated for the selected employees with summaries provided below:

Year of services	Sample size	Mean	Standard deviation
At most 3 years	12	82	8
More than 15 years	14	75	7

Can we conclude that the mean job satisfaction index of employees who worked for the company at most 3 years is higher than those worked for the company more than 15 years at 1% level of significance? State your assumptions. (6 marks)

5. (a) In order to assess the growth of social responsibility of students after the completion of a social services programme, level of social responsibility of students are assessed before joining the programme and after completion of programme. Eight students are selected at random and their assessment scores of social responsibility are listed below.

Student	1	2	3	4	5	6	7	8
Before	65	58	70	52	48	60	53	62
After	79	82	89	78	76	70	72	80

Test whether the mean assessment score of all students has increased after joining the programme at 5% level of significance. (7 marks)

- (b) The marketing officer of a company investigates the preference of customers on the design of bed linen sets. From past records, 50% of customers prefer classic design, 20% of customers prefer floral pattern, 15% prefer animal pattern and the rest prefer cartoon figures. A random sample of customers is selected and their most preferable design are tabulated below:

Design	Floral pattern	Cartoon figures	Animal pattern	Classic design
Number of customers	28	20	24	48

Are there any changes in preference of customers on the design of bed linen sets at 1% level of significance? (6 marks)

- (c) In order to study the association of the views of employees on the new promotion policy and their ranking, a random sample of employees is selected and the following cross-tabulation table shows the result.

Ranking of employee	Views on new promotion policy		
	Disagree	Neutral	Agree
Junior	58	32	17
Middle	45	63	57
Senior	27	45	56

Test whether the views of employees on the new promotion policy and their ranking are independent at 5% level of significance. (7 marks)

6. (a) The production manager would like to model the length of time (in hours) per month that a particular type of machine of a factory will be shut down for repairs on the age of the machine (in year). The summary of machine down time (y) and age of machine (x) of 10 randomly selected machines are shown below:

$$\sum x = 23.6, \sum x^2 = 71.02, \sum y = 228, \sum y^2 = 6376 \text{ and } \sum xy = 668.8$$

- (i) Find the least squares linear regression equation for predicting machine down time on age of machine. (4 marks)
- (ii) Calculate the sample correlation coefficient of machine down time and age of machine. (2 marks)
- (iii) Interpret the coefficient estimate of the independent variable. (2 marks)
- (b) The manager of a chain café is interested in predicting the weekly sales of the café on the weekly pedestrian flow on the street where the café is located and another three independent variables (namely x_2, x_3 and x_4). A random sample of 24 cafés is selected and a regression model is fitted using the observed data with some results provided below:

	<i>Coefficients</i>	<i>p-value (Sig.)</i>
Intercept	1.0834	0.0032
Pedestrian flow	0.1037	0.0001
x_2	-1.2158	0.0001
x_3	-0.5308	0.0779
x_4	-1.0765	0.0001

- (i) Write down the fitted multiple linear regression equation. (2 marks)
- (ii) Test whether the overall model is significant at 5% level of significance given that the corresponding test statistic is 222.173. (3 marks)
- (iii) Test whether weekly pedestrian flow is a significant explanatory variable at 1% level of significance. (3 marks)
- (iv) Based on the test statistic specified in (ii) and $SST=119.1496$, calculate and interpret the coefficient of determination. (4 marks)

*** END ***