

**THE UNIVERSITY OF HONG KONG**  
**School of Economics & Finance**  
**2003-2004 1st Semester Examination**

Economics: ECON1003  
Analysis of Economic Data  
Dr K F Wong

Candidates may use any self-contained, silent, battery-operated and pocket-sized calculator. The calculator should have numeral-display facilities only and should be used only for the purposes of calculation. It is the candidate's responsibility to ensure that his calculator operates satisfactorily.

Candidates must record the name and type of their calculators on the front page of their examination scripts.

The test paper consists of **twelve** questions. Please answer all of them, a total of **110** points. You are advised to allocate your time based on the relative weights of credits given in each question: approximately 5 minutes on a 5-point question, for example.

This is a close book exam. You are not allowed to refer to any notes or books during the exam.

Selected statistical formula and tables are provided.

- 
1. (10 points) The Rivertown city council is attempting to choose one of two sites (A or B) as the location for its new emergency facility. After the new emergency facility becomes available for service, the current emergency facility will be shut down. The project manager has estimated the following response times in minutes from each of the proposed sites to the four areas that must be served by the emergency facility.

Proposed Site	Area Served			
	1	2	3	4
A	5.2	4.4	3.6	6.5
B	6.0	7.4	3.4	4.0

The number of emergency runs from the current emergency facility to each of the four areas over the past year is as follows:

Area	1	2	3	4
Number of runs	150	65	175	92

Compute the weighted mean response time from both proposed locations and determine which proposed site should be selected for the new emergency facility.

2. (6 points) In one of our in-class experiments, we estimated the percentage of earth surface covered by water. We did that by throwing an inflated globe and checking whether a mark on our thumb landed on water of the globe. After 50 random throws, we estimated that 30 percent of earth surface is covered by water.
  - a. (3 points) Construct a 95% confidence interval for the percentage of earth covered by water.
  - b. (3 points) What is the probability of getting an estimated of 30 percent of earth surface covered by water if actually 70 percent of earth surface is covered by water?
  
3. (6 points) Four employees who work as drive-through attendees at a local fast food restaurant are being evaluated. As a part of quality improvement initiative and employee evaluation these workers were observed over three days. One of the statistics collected is the proportion of time employee forgets to include a napkin in the bag. Related information is given in the table below.

Worker	% of Dinners Packed	% Forgot Napkin
Joe	25%	6%
Jan	20%	2%
Cheryl	20%	10%
Clay	35%	4%

- a. What is the probability that Cheryl prepared your dinner and forgot to include a napkin?
  - b. What is the probability that there is not a napkin included for a given order?
  - c. You just purchased a dinner and found that there is no napkin in your bag, what is the probability that Cheryl has prepared your order?
  - d. You just purchased a dinner and found that there is no napkin in your bag, what is the probability that Jan has prepared your order.
  
4. (10 points) In Hong Kong, about 50 out of every 1000 children are gifted and talented. Children are classified as gifted and talented based on some IQ test or public examinations. Currently in undergraduate admission in Hong Kong universities, students with six or more distinctions (“A” grade) in Certificate Examination are treated as gifted and talented and can enter university a year earlier than their peers. This test (whether the student has six or more distinctions in Certificate Examination) is not perfect. One out of every 1000 tests given to students who have are actually gifted and talented comes out negative, that is,  $P(\text{less than six distinctions in Certificate Examination} \mid \text{gifted and talented}) = 1/1000$ . Out of every

1000 people who are not gifted and talented, 50 of them test positive, that is,  $P(\text{six distinctions in Certificate Examination} \mid \text{not gifted and talented}) = 5/100$ .

- a. (4 points) Draw a tree diagram describing probability of these events.
- b. (3 points) Using your figures from your tree diagram find  $P(\text{gifted and talented} \mid \text{six distinctions in Certificate Examination})$ , i.e., the probability of a student is gifted and talented when he/she gets six or more distinctions in Certificate Examination.
- c. (3 points) Using your figures from your tree diagram find  $P(\text{not gifted and talented} \mid \text{less than six distinctions in Certificate Examination})$ , i.e., the probability of a student is not gifted and talented when he/she gets less than six distinctions in Certificate Examination.

5. During off hours, cars arrive at a tollbooth on the East-West toll road at an average rate of 0.5 cars per minute. The arrivals are distributed according to the Poisson distribution.
  - a. What is the probability that during the next minute three cars will arrive?
  - b. What is the probability that during the next five minutes three cars will arrive?
6. (7 points) A pharmaceutical company has determined that if a new cholesterol-reducing drug is manufactured (introduced to the market), the following probability distribution will describe this drug's contribution to the company's profits during the next six months.

<u>Profit Contribution</u>	<u>Probability of profit contribution</u>
-\$ 30,000	.20
\$ 50,000	.50
\$200,000	.30

The company management has decided to market this product if the expected contribution to profit for the next six months is more than \$90,000. Based on the information given above, should the company begin manufacturing the new drug?

7. State and explain briefly
  - a. the Central Limit Theorem and
  - b. the Law of large numbers.
8. In a survey of 1,000 people, 420 are opposed to the tax increase.
  - a. Construct a 95 percent confidence interval for the proportion of those people opposed to the tax increase.
  - b. What is the required sample size in a follow-up survey if we want the maximum error to be less than xxx.

9. (10 points) An insurance company estimates that 40 percent of its claims have errors. The insurance company wants to estimate with 90 percent confidence the proportion of claims with errors. What sample size is needed if they wish to be within 5 percent of the actual?
10. (10 points) A car insurance company would like to determine the proportion of accident claims covered by the company. According to a preliminary estimate 60% of the claims are covered. How large a sample should be taken to estimate the proportion of accident claims covered by the company if we want to be 98% confident that the sample percentage is within  $\pm 3\%$  of the actual percentage of the accidents covered by the insurance company?
11. A recent study conducted by the state government attempts to determine whether the voting public supports further increase in cigarette taxes. The opinion poll recently sampled 1500 voting age citizens. 1020 of the sampled citizens were in favor of an increase in cigarette taxes. The state government would like to decide if there is enough evidence to establish whether the proportion of citizens supporting an increase in cigarette taxes is significantly greater than .66.
- Write the null hypothesis for this problem.
  - Write the alternative hypothesis.
  - What is the rejection point (given in terms of the value of the test statistic) at  $\alpha = .10$ ?
  - What is the rejection point (given in terms of the value of the test statistic) at  $\alpha = .01$ ?
  - What is the sample value of the test statistic?
  - What is the decision at  $\alpha = .10$ ?
  - What is the decision at  $\alpha = .05$ ?
  - What is the p-value for this test?
12. (5 points) A marketing research company surveyed grocery shoppers in the east and west coasts to see the percentage of the customers who prefer chicken to other meat. The data are given below.

	Sample size	Number of customers in the sample that prefer chicken
East Coast	492	156
West Coast	386	172

- Is the proportion of customers who prefer chicken the same for the two regions? Test at  $\alpha = .10$ .
- Is the proportion of customers who prefer chicken higher at the West Coast? Test at  $\alpha = .05$ .

- c. Determine the 95% confidence interval for the difference between the proportion of customers who prefer chicken at the West Coast and the proportion of customers who prefer chicken at the East Coast.
- d. The 95% confidence interval for the difference between the proportion of customers who prefer chicken at the West Coast and the proportion of customers who prefer chicken at the East Coast is .0642 to .193. Provide a one-sentence interpretation of this interval.

13. (10 points) A researcher has used a one-way analysis of variance model to test whether the average starting salaries differ among the recent graduates from nursing, engineering, business and education disciplines. She has randomly selected four graduates from each of the four areas.

- a. If  $MSE = 4$ , and  $SSTO = 120$  complete the following ANOVA table and determine the value of the F statistic.

Source	SS	DF	MS	F
Treatment				
Error				
Total				

- b. Determine degrees of freedom treatment, degrees of freedom error and degrees of freedom total and state the critical value of the F statistic at  $\alpha = .05$
- c. If  $MSE = 4$ , and  $SSTO = 120$  complete the following ANOVA table. Is there a significant difference in the starting salaries among the four disciplines?

Source	SS	DF	MS	F
Treatment				
Error				
Total				

14. Complete the following partial ANOVA table from a simple linear regression analysis with a sample size of 16 observations. Use the F test to test the significance of the model at  $\alpha = .05$ .

Source	SS	DF	MS	F
Regression			309.9	
Error				
Total	995.95			

15. A local tire dealer wants to predict the number of tires sold each month. He believes that the number of tires sold is a linear function of the amount of money invested in advertising. He randomly selects 6 months of data consisting of monthly tire sales (in thousands of tires) and monthly advertising expenditures (in thousands of dollars). The simple linear regression equation is  $\hat{y} = 3 + 1X$  and sample correlation coefficient  $(r^2) = .6364$ . Test to determine if there is a significant correlation between the monthly tire sales and monthly advertising expenditures. Use  $H_0: \rho = 0$  vs.  $H_A:$

$\rho \neq 0$  at  $\alpha = .05$ .

16. The manufacturer of a light fixture believes that the dollars spent on advertising, the price of the fixture and the number of retail stores selling the fixture in a particular month influence the light fixture sales. The manufacturer randomly selects 10 months and collects the following data:

	Sales	Advertising	Price	# of stores
	41	20	40	1
	42	40	60	3
	59	40	20	4
	60	50	80	5
	81	50	10	6
	80	60	40	6
	100	70	20	7
	82	70	60	8
	101	80	30	9
	110	90	40	10

The sales are in thousands of units per month, the advertising is given in hundreds of dollars per month, the price is the unit retail price for the particular month. Using MINITAB the following computer output is obtained.

The regression equation is

$$\text{Sales} = 31.0 + 0.820 \text{ Advertising} - 0.325 \text{ Price} + 1.84 \text{ Stores}$$

Predictor	Coef	StDev	T	P
Constant	30.992	7.728	4.01	0.007
Advertising	0.8202	0.5023	1.63	0.154
Price	-0.32502	0.08935	-3.64	0.011
Stores	1.841	3.855	0.48	0.650

S = 5.465

R-Sq = 96.7% R-Sq(adj) = 95.0%

Analysis of Variance

Source	DF	SS	MS	F	P
Regression	3	5179.2	1726.4	57.81	0.000
Residual Error	6	179.2	29.9		

Based on the multiple regression model given above, the point estimate of the monthly light fixture sales corresponding to second sample data is 49.82 or 49,820 units. This point estimate is calculated based on the assumption that the company spends \$4000 on advertising, the price of the fixture is \$60 and the fixture is being sold at 3 retail stores. Additional information related to this point estimate is given below.

Obs.	Advert	Sales	Fit	StDev Fit	Residual	St Resid.
2	40.0	42.00	49.82	3.53	-7.82	-1.87

- Determine the 95% confidence interval for this point estimate and interpret its meaning.
- Calculate the 95% prediction interval for this point estimate.
- The 95% prediction interval for the point estimate given above is from 33.9 to 65.74. Interpret the meaning of this interval.
- Determine the 95% interval for  $\beta_1$  (beta coefficient for the advertising variable).
- The 95% confidence interval for  $\beta_1$  is from -0.4089 to 2.0493. Interpret the meaning of this interval.
- Test the usefulness of variable “price” in the model using the null hypothesis  $H_0: \beta_2 \leq 0$ , at  $\alpha = 0.05$ , and state your conclusions.