國立政治大學
108 學年度

統計學系碩士班甄試考試命題紙

考試科目	統計學	所(組)別	統計學系	考試時間	107 年 11 月 4 日 星期日 10：00－11：40

請注意，每一大題的作答分別寫在一份答案本上。答案紙上請清楚標示各大題及小題的題號，無法辨識者，以零分計算。每一大題皆為 25 分，各小題配分以 points 或 pts 表示之。

第一大題

1. (5 points) Which ones of the following variables are interval-scaled? (multiple choice)
   (a) Hourly income of restaurant waiters
   (b) IQ ratings
   (c) Movie ratings (great, very good, average, poor, awful)
   (d) Ranking of students in a class
   (e) Temperature
2. (5 points) Below are summary statistics of hourly income earned by a sample of 100 waiters at Orange city. Which ones of the following statements are correct? (multiple choice)

min	Q1	mean	Q2	Q3	max
75	114	121	134	141	145

   (a) The minimum value 75 is an outlier.
   (b) The maximum value 145 is an outlier.
   (c) The average of minimum and maximum, (max+min)/2 = 110, is a statistic.
   (d) The distribution of hourly income is negatively skewed.
   (e) The distribution of hourly income is right skewed.
3. (10 points) The return on investment earned by Alice for three successive years was: 20 percent, 5 percent, and -10 percent. Suppose that the original investment and the total return each year is reinvested for the next year. Find the annual rate of return on investment.

第二大題

1. (8 pts) Suppose that in a university, campus shuttles leave from the library in campus to the dormitory area every 30 minutes. Suppose that students arrive at the library stop randomly and the average waiting time that a student waits for the shuttle at that stop is 15 minutes. Let T denote the waiting time (in minutes) for the next student arriving at the library stop. For each of the following distribution, determine whether it is reasonable to use it as the distribution of T. Justify your answer.
   (a) the normal distribution with mean 30 and standard deviation 10
   (b) the Poisson distribution with mean 15
   (c) the uniform distribution on $(0, 30)$
   (d) the binomial distribution with number of trials $n = 60$ and success probability $p = 0.25$

考試科目	統計學	所(組)別	統計學系	考試時間	107 年 11 月 4 日 星期日 10：00－11：40

2. (6 pts) Suppose that lifetimes (in hours) for a brand of light bulbs follow the exponential distribution with mean 25000. Write down the probability that a new light bulb of this brand will still work after being used for 20000 hours. If your answer involves e^x for some x, you do not have to evaluate it, where e is the base for the natural exponential function.

3. (6 pts) State and explain the central limit theorem.

第三大題

Past experience indicates that the monthly long-distance telephone call bill is normally distributed with a mean of \$17.85 and a standard deviation of \$3.87. After an advertising campaign aimed at increasing long-distance telephone usage, a random sample of 25 household bills was taken. The data are as follows:

19.61	20.14	19.57	19.26	14.03	19.24	15.98	24.85	26.00
19.46	18.29	16.91	26.15	19.64	16.75	20.52	25.47	18.19
12.56	28.47	14.13	19.72	17.05	13.92	12.38		

a. (6 pts) Do the data allow us to infer at the 10% significant level that the campaign was successful?
b. (2 pts) What assumption(s) must you make to answer Part (a)?
c. (6 pts) Determine the probability of a Type II error when the mean μ is 17.95.
d. (2 pts) Assume that the mean μ is unknown, construct a 98% confidence interval for μ.
e. (4 pts) Repeat Part (a) when the standard deviation is unknown.

You may use the following quantiles for this problem.
$z_{.1}=1.28$, $z_{.05}=1.645$, $z_{.025}=1.96$, $z_{.01}=2.328$, $z_{.005}=2.576$
$t_{.1,24}=1.32$, $t_{.05,24}=1.71$, $t_{.025,24}=2.06$, $t_{.01,24}=2.49$, $t_{.005,24}=2.80$

第四大題

觀察十五位使用自動櫃員機（ATM）的顧客，記錄其身高（*Height*，公分）、性別（*Gender*，1=男性，0 = 女性）及操作 ATM 完成交易所需時間（*Time*，秒）。若以完成交易所需時間為反應變數，進行迴歸分析；下表為不同解釋變數的迴歸模型之估計結果：

考試科目	統計學	所(組)別	統計學系	考試時間	107 年 11 月 4 日 星期日 10：00－11：40

	模型 A		模型 B		模型 C	
	Estimate	Std. Error	Estimate	Std. Error	Estimate	Std. Error
Intercept	210.778	3.537	36.973	74.407	228.232	90.303
Height			1.085	0.439	-0.106	0.550
Gender	24.222	5.592			25.558	9.024
Residual standard error	10.61		13.68		11.03	
F-statistic	18.760		6.095		8.705	

回答以下問題：

（1） (12 pts) 建構模型 C 的變異數分析（ANOVA）表，並解釋其意義。

（2） (8 pts) 討論模型 A、B 和 C 中之最佳模型為何？詳細說明模型選擇的依據(至少兩種準則或判斷方式)。

第五大題

An analyst at a local bank wonders if the age distribution of customers coming for service at his branch in town is the same as at the branch located near the mall. He selects 100 transactions at random from each branch and researches the age information for the associated customers. Here are the data:

	Age			
	<30	30-55	>55	Total
In-Town Branch	20	40	40	100
Mall Branch	30	50	20	100
Total	50	90	60	200

(a) (2 pts) What is the null hypothesis?
(b) (3 pts) What type of test is this, goodness-of-fit, homogeneity, or independence? Explain.
(c) (6 pts) Carry the test, and find the P-value.
(d) (2 pts) State your conclusion. ($\alpha = 0.05$)
(e) (4 pts) If the analyst had selected 50 transactions at random from each branch and counted exactly half as many in each category, i.e., 10, 20, 20 for in-town branch, and 15, 25, 10 for mall branch, what conclusion would he have reached?
(f) (3 pts) Is there a discrepancy between the two conclusions? If yes, why is there a discrepancy between the two conclusions? Explain.