國立政治大學
114學年度

統計學系碩士班甄試考試試題

考試科目	統計學	所(組)別	統計學系	考試時間	113年11月10日 星期日 10：00－11：40

注意：本科目有四份試題，第一份試題答案寫在第一卷答案本上，依此類推。每份試題配分為25分，合計共100分。

註：請自行利用查表本找到臨界值

第一份試題(共25分，答案請寫在第一卷答案本上)

1. 政治大學統計系某位老師調查大學生使用手機的習慣，以隨機抽樣訪問 100 位受訪者，發現每天使用手機時間的95%信賴區間為(5.4 小時, 6.6 小時)：
   (a) 請提供這個信賴區間的統計詮釋（4%）；
   (b) 另一位老師在同時間以相同隨機抽樣方法執行類似研究，得出每天手機使用時間的95%信賴區間為(5.5 小時, 6.3 小時)，說明哪些因素造成兩位老師的估計值不同，以及兩者是否會獲得相同結論（6%）？

2. 報章雜誌經常提到問卷調查的結果，像是臺灣或美國某位總統候選人支持度，出現以下描述「某候選人的支持度為 48%，根據 1068 位受訪者的調查結果，在百分之 95 的信心水準下，抽樣誤差在正負 3 個百分點以內」。請說明上述文字敘述必須在哪些條件下才會成立（5%）。

3. 政治大學學生每天午餐花費平均 100 元新台幣，午餐費用的標準差為 60 元新台幣。如果以簡單隨機抽樣蒐集36位學生的樣本：
   (a) 計算這36位學生午餐的平均費用，請問這個平均數服從什麼分配，期望值、標準差又是多少(4%)。
   (b) 計算 36 位學生平均午餐費用至少 80 元新台幣的機率（3%）。
   (c) 計算 36 位學生平均午餐費用介於 90～120 元新台幣的機率（3%）

備考：試題隨卷繳交

統計學系碩士班甄試考試試題

考試科目	統計學	所(組)別	統計學系	考試時間	113 年 11 月 10 日 星期日 10：00－11：40

第二份試題(共 25 分，答案請寫在第二卷答案本上)

1. A company wants to show that they are not gender biased. They take a sample of 20 woman in the company and get their salaries. They then take 20 men who have the identical jobs to each of the women, and get their salaries. Suppose that the normality assumption is satisfied. What tests would you use for this problem? (Multiple Choice; 3%)

   A. z test based on a single sample for μ

   B. t test based on a single sample for μ

   C. t test based on two independent samples for $\mu_1 - \mu_2$

   D. t test based on matched samples for μ_D

   E. z test based on two independent samples for $p_1 - p_2$

2. The following contingency table shows the results of a random sample of adults classified by their gender and color preference. The sample proportions are displayed in cells.

	Red	Yellow	Blue
Men	0.13	0.19	0.28
Women	0.07	0.11	0.22

   (a) Write down the relevant hypotheses if we want to determine whether there is no difference in the color preferences of men and women by a chi-square analysis. (2%)

   (b) Suppose the sample consists 100 adults. Find the *p*-value for the test. What is your conclusion at 10% significance level? (6%)

3. A magazine claims that 50% of college students have student loans. A college dean feels that this is not true at his college. He decides to conduct a test at the 10% significance level and randomly selects a sample of 8 students. Let X be the number of students who have student loans in the sample.

   (a) Find the rejection region for the test in terms of X. (6%)

   (b) Calculate the actual significance level of this test. (4%)

   (c) If 4 out of the 8 students have student loans, should the null hypothesis be rejected using the rejection region in part (a)? (4%)

統計學系碩士班甄試考試試題

考試科目	統計學	所(組)別	統計學系	考試時間	113 年 11 月 10 日 星期日 10：00－11：40

第三份試題(共 25 分，答案請寫在第三卷答案本上)

1. (25 pts) Suppose that Mr. Chao always goes to a restaurent for lunch, and the restaurent provides three dinning choices: Meal 1, Meal 2, and Meal 3. For $i, j \in \{1, 2, 3\}$, let $p_{i,j}$ denote the conditional probability that Mr. Chao chooses Meal j for the next lunch given that his current lunch is Meal i. Suppose that

$$\begin{pmatrix} p_{1,1} & p_{1,2} & p_{1,3} \\ p_{2,1} & p_{2,2} & p_{2,3} \\ p_{3,1} & p_{3,2} & p_{3,3} \end{pmatrix} = \begin{pmatrix} 0 & 0.6 & 0.4 \\ 0.5 & 0 & 0.5 \\ 0.4 & 0.6 & 0 \end{pmatrix}$$

Suppose that the probability that Mr. Chao chooses Meal j for his lunch tomorrow is $1/3$ for $j \in \{1, 2, 3\}$.

(a) (9 pts) Find the probability that Mr. Chao chooses Meal j for lunch the day after tomorrow for $j \in \{1, 2, 3\}$.

(b) (8 pts) Find the conditional probability that Mr. Chao chooses Meal 2 for lunch on the second day after tomorrow given that he chooses Meal 2 for lunch tomorrow.

(c) (8 pts) Find the conditional probability that Mr. Chao chooses Meal 1 for lunch tomorrow given that he chooses Meal 2 for lunch the day after tomorrow.

統計學系碩士班甄試考試試題

考試科目	統計學	所(組)別	統計學系	考試時間	113 年 11 月 10 日 星期日 10：00 – 11：40

第四份試題(共 25 分，答案請寫在第四卷答案本上)

1. Consider a simple regression model

$$Y = \beta_0 + \beta_1 X + \epsilon,$$

where $\epsilon \sim N(0, \sigma^2)$. Suppose that the dataset is given by $\{(X_i, Y_i) : i = 1, \cdots, n\}$.

   (a) (5%) Determine $\widehat{\boldsymbol{\beta}} \triangleq (\widehat{\beta}_0, \widehat{\beta}_1)^\top$ and show that $\widehat{\boldsymbol{\beta}}$ is the unbiased estimator of $\boldsymbol{\beta} \triangleq (\beta_0, \beta_1)^\top$.

   (b) (5%) Find the estimator of σ^2 and argue that whether your derived estimator is unbiased.

   (c) (5%) Determine the covariance $\text{cov}(\widehat{\beta}_0, \widehat{\beta}_1)$.

2. Continue on the question #1, suppose that we have a dataset with sample size $n = 20$. It is also known from the scatter plot that two variables Y and X have positive correlation. Based on the data, $\sum_{i=1}^{n}(Y_i - \overline{Y})^2 = 300$ and $\sum_{i=1}^{n}(X_i - \overline{X})^2 = 100$. Moreover, the coefficient of determination is $R^2 = 0.9$.

   (a) (2%) Find the estimator of β_1.

   (b) (2%) Complete the following ANOVA table:

Source of variation	Sum of Squares	Degrees of Freedom	Mean Square	F
Regression				
Error				
Total				

   (c) (2%) Find 90% confidence interval of β_1.

   (d) (2%) Test $H_0 : \rho = 0$ vs $H_1 : \rho \neq 0$, where ρ is the population correlation coefficient of Y and X.

   (e) (2%) Suppose $\overline{X} = 0$ and $\overline{Y} = 0$. Please find a 95% confidence interval for the mean response when a new observation is $X_0 = 2$.