**100學年度上學期 金門大學 資工系三年級 機率統計 期末考 出題者 : 陳鍾誠**

**學號 : 姓名 : 分數 :**

**注意：請盡可能包含說明與過程，答案對但沒寫過程與說明只能得到很少分數。**

|  |  |
| --- | --- |
| 1. **請簡要說明以下概念。** 2. **何謂機率？何謂統計？兩者有何不同？(5%)**   機率：注重從母體到樣本的過程，探討各種機率分布的特性，以及如何從母體產生樣本。  統計：注重從樣本到母體的過程，探討如何描述樣本集合的特性，以及如何從樣本推論母體的特性。  兩者的不同點在於觀察的角度與方法，機率是從母體產生樣本，而統計是從樣本推論母體。   1. **何謂敘述統計？何謂推論統計？兩者有何不同？(5%)**   敘述統計：探討如何以數字或圖形描述樣本集合特性的統計學問。  推論統計：探討如何用樣本推論母體特性的統計學問。  兩者的不同點在於敘述統計只描述樣本。而推論統計則會推論母體的分布與參數。   1. **何謂點估計？何謂區間估計？兩者有何不同？(5%)**   點估計：是用樣本統計值推論出母體最可能參數值的估計方式，例如用樣本平均值估計母體平均數，通常採用不偏估計式。  區間估計：是用樣本集推論母體某參數之信賴區間的學問，例如樣本平均值估計母體平均數的 95% 信賴區間。  兩者的不同點在於點估計只找出最可能參數點，但區間估計則是找出一個較可能的範圍。   1. **何謂假設檢定？何謂顯著性檢定？兩者有何不同？(5%)**   假設檢定：先確定信賴水準，然後再計算信賴區間，以便檢定某個假設是否可信賴的檢定法。  顯著性檢定：直接檢驗某個假設的顯著性，以判斷可能性的檢定法。  假設檢定找出的是一個範圍，落在範圍外的就會被否決。而顯著性檢定則找出一個 P 值，讓使用者根據 P 值進行判斷。 | 1. **請詳細說明以下概念。** 2. **何謂中央極限定理？(10%) (必須寫出公式並解釋其意義，越詳細越好)**   中央極限定理：樣本平均數的分布會趨近於常態分布，其數學公式如下所示。  其條件是 X1, X2,…,Xn必須是相互獨立且來自同一個母體的樣本 (但母體的分布並沒有限制於常態分布，母體可以是任意的分布，像是均等分布，二項分布、poisson 等母體，其樣本平均數都會趨向於以 母體平均數 與變異數 為參數的常態分布。   1. **何謂虛無假設 H0？何謂研究假設H1？中央極限定理與平均值 的檢定之間有何關係呢？ (10%)**   虛無假設：在檢定時先入為主的一個假設，用來判斷是否要被否決的假設。通常用 H0 表示，例如 H0: 就是用來檢定母體平均數的虛無假設。  研究假設：虛無假設的相反假設(否定版)，又稱為對立假設，也就是不符合虛無假設，就必定符合研究假設。通常用 H1 表示，例如 H1: 就是上述 H0 的對立假設。  推論統計當中的大部份推論，都是以中央極限定理為核心的，我們之所以可以用 推論母體的 ，並對母體進行信賴區間檢定，這完全是因為中央極限定理的緣故。  因為 會趨近於 ，所以 會趨近於標準常態分布，所以我們才能用常態分布檢定 。(但是在母體變異數 未知的情況之下，只好改用樣本變異數 S2 估計，此時會引入另一層誤差，因此必需改用 T 分布進行檢定) (T 分布與常態分布很像，而且當自由度越大時，T 分布會越趨近於常態分布。 |

**學號 : 姓名 :**

|  |  |
| --- | --- |
| 1. **以下是母體變異數未知情況下用 T 分布估計母體平均數 信賴區間的公式，以及對應的 R 程式** 2. **請分別說明 的意義。(5%)**   ：樣本平均數  T：學生T 分布  ：母體平均數  ：犯下型 I 錯誤的機率   1. **以下是對應上述公式的 R 程式，請分別說明 分別對應到哪個變數。(5%)**   **mean.range2 = function(x, alpha=0.05) {**  **n = length(x)**  **mx = mean(x)**  **S = sqrt(var(x))**  **r1 = qt(alpha/2, df=n-1)**  **r2 = qt(1-alpha/2, df=n-1)**  **L1 = mx+r1\*S/sqrt(n)**  **L2 = mx+r2\*S/sqrt(n)**  **range = c(L1, mx, L2)**  **}**  ：對應程式中的mx  T：qt 為 T的累積分布函數， r1 = **，**r2 =  ：被估計的母體平均數，在程式中不存在，但其分布的 95% 信賴區間為 L1<<L2。  ：對應程式中的 alpha | 1. **有一樣本序列X=( 8, 9, 1, 3, 9, 1, 2, 4 ,8, 4, 5, 1, 3, 9, 6, 2, 4, 6, 8, 5)，請回答下列問題** 2. **請計算 X 的樣本平均數 (5%)** 3. **請計算 X 的樣本變異數 (5%)** 4. **假如已知X 的母體變異數 ，請估計 X 母體平均值 的 95% 信賴區間。(10%)**   下限   上限  所以平均值 的 95% 信賴區間為 (3.66, 6.13)   1. **假如X 的母體變異數未知，請估計 X 母體平均值 的 95% 信賴區間。(10%)**   下限   上限  所以平均值 的 95% 信賴區間為 (3.56, 6.23)  **已知：Z 與 T 都對稱於 0 點，且其累積分布符合下列算式** |

**學號 : 姓名 :**

|  |  |
| --- | --- |
| 1. **請說明下列 R 操作過程的意義 (必需特別説明信賴區間的範圍，顯著水準的值與意義) (10%)。**   **> x = runif(25, 1, 5) # 取 (1,5) 之間的均等分布樣本 25 個**  **> x**  **[1] 1.914426 1.249618 4.786053 4.364882 2.592136 4.185278 3.522799 3.020045**  **[9] 3.068202 1.499622 3.445676 3.998780 4.306535 4.031798 3.514701 4.501148**  **[17] 4.894386 3.031645 1.722967 3.323432 2.454091 3.169133 1.064701 4.726883**  **[25] 4.651454**  **> t.test(x, mu=3) # 對x進行 H0: 的T檢定**  **One Sample t-test**  **data: x**  **t = 1.384 , df = 24, p-value = 0.1791**  **# t 值為 1.384, 自由度為 24 # 顯著水準P值為 0.1791**  **alternative hypothesis:true mean is not equal to 3**  **# 對立假設為 H1:**  **95 percent confidence interval:**  **2.841988 3.801243**  **# 95% 信賴區間為**  **sample estimates:**  **mean of x # 平均值**  **3.321616**  **# 結果在 95%信賴區間下無法否決 H0:，其P值為 0.1791 > 0.025。** | 1. **請說明下列 R 操作過程的意義 (必需特別説明信賴區間的範圍，顯著水準的值與意義) (10%)。**   **> x = rnorm(20, mean=5, sd=2) # 取均值為5,標準差為 2 的常態分布樣本 20 個**  **> y = rnorm(20, mean=5.1, sd=1.9) # 取均值為5.1,標準差為 1.9 的常態分布樣本 20 個**  **> t.test(x,y,alternative="two.sided") # 對 (x,y) 進行兩組數據的T檢定，檢驗**  **Welch Two Sample t-test**  **data: x and y**  **t = 1.6057, df = 37.094, p-value = 0.1168**  **# t 值為1.6057, 自由度為 37.094 # 顯著水準P值為 0.1168**  **alternative hypothesis: true difference in means is not equal to 0 # 對立假設為**  **95 percent confidence interval:**  **-0.3376277 2.9175453**  **# 95% 信賴區間為**  **sample estimates:**  **mean of x mean of y**  **5.977734 4.687776**  **# 平均值**  **>**  **# 結果在 95%信賴區間下無法否決 ，其P值為 0.1168 > 0.025。** |