

注意：

1. 本試題共七大題，合計 100 分。
2. 請依題號依序作答，否則不予計分。
3. 請詳述理由或計算推導過程，否則不予計分。

一、(5 分) 假設 (X, Y) 為抽自均等分配 (uniform distribution) $U[-1, 1]$ 的隨機樣本。
試求條件機率

$$P(X < Y | Y < 0).$$

二、令 X 為間斷隨機變數，其機率函數為：

$$f_X(x) = P(X = x) = \begin{cases} p, & x = -1, \\ q, & x = 0, \\ r, & x = 1. \end{cases}$$

試求以下期望值：

1. (5 分) $E(X^2 | X)$
2. (5 分) $E(X | X^2)$
3. (5 分) $E(E(X | X^2))$

三、 $\{X_i\}_{i=1}^n$ 為一抽自均等分配 (uniform distribution) $U[0, 1]$ 的隨機樣本。 $\{Y_i\}_{i=1}^n$ 為抽自一個未知分配的隨機樣本，其中 $E(Y_1) = 0$, $Var(Y_1) = 1$ 。令

$$Z_n = \frac{X_1 + X_2 + \cdots + X_n}{X_1^2 + X_2^2 + \cdots + X_n^2},$$

$$W_n = \frac{\sqrt{n}(Y_1 + Y_2 + \cdots + Y_n)}{X_1^2 + X_2^2 + \cdots + X_n^2}.$$

1. (5 分) 若 Z_n 機率收斂 (convergence in probability) 到常數 c ，試求算 c 值。
2. (10 分) 試找出 W_n 的極限分配 (asymptotic distribution)。

接背面

四、若 $(X_1, \dots, X_n)$ 是從常態母體 $X \sim N(\mu, \sigma^2)$ ，以簡單隨機抽樣法(SRS)抽出的隨

機樣本，且令 $s^2 = \frac{1}{n-1} \sum_{i=1}^n (X_i - \bar{X})^2$ ， $\hat{\sigma}^2 = \frac{1}{n} \sum_{i=1}^n (X_i - \mu)^2$ ，請依序回答下列各問題：

1. (5分) s^2 與 $\hat{\sigma}^2$ 都是 σ^2 (母體變異數) 的不偏估計式(Unbiased Estimator)嗎？
(請證明或說明，否則不予計分。)
2. (5分) s^2 與 $\hat{\sigma}^2$ 都是 σ^2 (母體變異數) 的一致性估計式(Consistent Estimator)嗎？
為什麼？又相對於 $\hat{\sigma}^2$ 而言， s^2 具有有效性(Efficiency)嗎？為什麼？
3. (5分) 設若母體平均數(μ)值為已知，試問 σ^2 的區間估計式(信任係數為 $1-\alpha$)
為何？為什麼？

五、台大的某教授為檢定台灣的大學生「每週讀書時數的平均數(μ_x)是否低於(含) μ_0 」，
因此，他選擇樣本平均數($\bar{X}$)做為檢定統計量，並從大學生當中，以簡單隨機抽樣法
(SRS)，抽出 100 名大學生，發現他們每週讀書時數的平均值($\bar{X}_0$)是 10.328 小時，
且計算得到 P-Value = 0.05，請依序回答下列各問題：

1. (9分) 何謂 α 風險 (α Risk)？設若該教授決定採「拒絕「大學生每週讀書時數
的平均數(μ_x)低於(含) μ_0 」的決策」，試問 α Risk 多大？又若知台灣
的全體大學生每週讀書時數的標準差是 2 小時，試問 $\mu_0 = ?$
【Note: $Z_{0.0015} = 3, Z_{0.005} = 2.58, Z_{0.023} = 2, Z_{0.025} = 1.96, Z_{0.05} = 1.64, Z_{0.1585} = 1$;
 Z 是標準常態隨機變數。】

2. (6分) 今若該位教授也想要探討「台灣的大學生從高中進入大學後，每週讀書時
數的平均數有何變化(增或減多少小時)。」試問此教授應採用哪一個統
計方法？如何抽樣？如何探討該問題？

六、若想探討依變數(Y)與自變數(X)間的關係，設立下列的模型：

$$(1): Y = \alpha_1 + \beta_1 X + \varepsilon_1$$

1. (6分) 設立模型後，你發現影響Y的重要因素尚有Z變數，因此將模型設為(2):

$$(2): Y = \alpha_2 + \beta_2 X + \gamma_2 Z + \varepsilon_2$$

若你利用 OLS (最小平方法) 去估計模型(1)與(2)，那麼 $\hat{\beta}_1$ 和 $\hat{\beta}_2$ 的關係為何？請證明之。

2. (4分) 若模型設為(3):

$$(3): \frac{Y}{P} = \alpha_3 + \beta_3 \frac{X}{P} + \gamma_3 \frac{Z}{P} + \varepsilon_3$$

其中P為隨機變數

$\hat{\beta}_3$ 為 OLSE (最小平方法估計式)，那麼 $\hat{\beta}_2$ 和 $\hat{\beta}_3$ 會相同嗎？理由。

3. (5分) 若模型設為(4):

$$(4): \ln \frac{Y}{P} = \alpha_4 + \beta_4 \ln \frac{X}{P} + \gamma_4 \ln \frac{Z}{P} + \varepsilon_4$$

你是否可用最大概似法 (ML) 去估計嗎？若可以，請簡單說明估計的方法；若不可以，請說明理由。

4. (5分) 若模型設為(5):

$$(5): \ln Y = \alpha_5 + \beta_5 \ln X + \gamma_5 \ln Z + \delta_5 \ln P + \varepsilon_5$$

請問利用 OLS 去估計模型(4)與(5)時， $\hat{\beta}_4$ 和 $\hat{\beta}_5$ 會相同嗎？請證明。

七、樂透彩券的收益是政府的財政收入來源之一，但是許多人批評樂透彩券是對低所得者及低教育程度者的租稅。假設你是經濟學家，你想利用統計方法檢驗上述的批評，因此抽取 100 個成人的隨機樣本，調查購買樂透支出佔所得的百分比 (Y) 及家戶所得 (I)、教育程度 (分三級，高中以下、大學、大學以上) 與一些社會及經濟變數，如年齡 (Age)、小孩數 (Child)、是否就業 (EMP) 等。此外，已知年齡為二次式的影響，且是否就業與年齡有交叉影響。

1. (6分) 請建立一個適當的複迴歸模型 (Multiple regression model)。(模型設立不完整，不予計分。)

2. (3分) 若教育程度以教育年數來衡量，是否較佳？理由。

3. (6分) 若要檢定樂透彩券是對低所得者及低教育程度者的租稅的批評，應如何進行聯合檢定，請列出兩個假設並說明檢定的統計量。