1. (15分) (Random Variable and Expectation) Let X be a normally distributed random variable

$$X = \ln P \sim \mathcal{N}(\mu, \sigma^2)$$

with mean μ and variance σ^2. The probability density function for X is then given by

$$f_X(x) = \frac{1}{\sqrt{2\pi\sigma^2}} \exp\left\{-\frac{(x-\mu)^2}{2\sigma^2}\right\}, \quad x \in \mathbb{R}.$$

(a) Derive the probability density function for P.

(b) Given $\gamma \in \mathbb{R}$, compute $\mathbb{E}\left[e^{-\gamma X}\right]$ and $Var\left(e^{-\gamma X}\right)$.

(c) Observe the fact that $\mathbb{E}\left[P \cdot \mathbb{I}_{\{P>K\}}\right] = \mathbb{E}[P]\Phi\left(\frac{\mu - \ln K}{\sigma} + \sigma\right)$, where $\mathbb{I}_{\{\cdot\}} = 1$ if $P > K$, $\mathbb{I}_{\{\cdot\}} = 0$ if $P \leq K$, $K > 0$, and $\Phi(\cdot)$ denotes the standard Normal cumulative distribution function. Now, compute $\mathbb{E}\left[\max\{P-K, 0\}\right]$.

2. (35分) (Maximum Likelihood Estimation and The Best Critical Region) Consider a random sample $(X_1, X_2, \ldots, X_n)$ and each random variable X has density

$$f_X(x) = (1+\theta)x^\theta,$$

where θ is the parameter, $\theta > 1$, and $0 < x < 1$.

(a) Write down the log likelihood function in terms of θ based on n observations $x_1, \ldots, x_n$.

(b) Calculate $\mu := \mathbb{E}[X]$ and obtain the variance of $\hat{\mu} := \left(\sum_{i=1}^n x_i\right)/n$ in terms of μ only. (*Hint:* There is a relation between θ and μ).

(c) Obtain the maximum likelihood estimator of μ based on the log likelihood function constructed in part (a). Denote the resulting ML estimator as $\hat{\mu}_{ML}$.

(d) Derive the asymptotic variance of $\hat{\mu}_{ML}$, and then compare it with $Var(\hat{\mu})$ which you have obtained in part (b).

(e) Suppose that an economist is interested in testing $H_0 : \mu = \mu_0$ against $H_A : \mu = \mu_A$, you are asked to describe the best critical region of size α. To this end, you should use the following fact which is known as the Neyman-Pearson Lemma: in testing $H_0 : \mu = \mu_0$ against $H_A : \mu = \mu_A$, the best critical region of size α is given by

$$CR = \left\{(x_1, \ldots, x_n) \,\middle|\, \frac{L(x_1, \ldots, x_n|\mu_A)}{L(x_1, \ldots, x_n|\mu_0)} > c\right\},$$

where L is the likelihood function and c (the critical value) is determined so as to satisfy $\mathbb{P}(CR|\mu_0) = \alpha$.

見背面

3. (10分) 某經濟學家想利用實驗設計去評估三種獎勵方法 (A_1, A_2, A_3) 對生產水準的影響，但因員工經驗不同 (可以以年資表示)，亦會影響生產水準。

(1) 請問你如何進行實驗設計，利用何種變異數分析法去分析獎勵方法是否會影響生產水準?

(2) 除了變異數分析外，請寫出另一種可以評估三種獎勵方法是否有差異的統計方法及其步驟。

4. (25分) 下列是以最小平方法將睡眠時間對工作時間，教育年數，年齡及性別迴歸的結果:

$$\widehat{Sleep} = \underset{}{3840.83} - \underset{(0.018)}{0.163\,work} - \underset{(5.86)}{11.71\,edu} - \underset{(11.21)}{8.70\,age} + \underset{(0.042)}{0.128\,age^2} + \underset{(34.33)}{87.75\,male}$$

$n = 706,\ R^2 = 0.123,\ \bar{R}^2 = 0.117$

Sleep: 一個禮拜晚上睡眠時間 (分鐘)，*work*: 一個禮拜工作總時間 (分鐘)，*edu*: 教育年數，*age*: 年齡 (年)，*male*: 性別虛擬變數 ($male = 1$ 男性，$male = 0$ 女性)。

試回答下列小題:

(1) 請檢定迴歸模型是否可被接受? ($\alpha = 0.05$) (請自行與適當的臨界值比較，說明是否可接受該模型。)

(2) 請問工作與睡眠時間是否有顯著的替代 (trade off) 關係? 如果有，請問在其他變數固定下，多工作 1 小時，睡眠時間會如何?

(3) 你如何比較迴歸模型中各變數的影響力大小，請說明之。

(4) 在其他變數固定下，你如何檢定年齡是否會影響睡眠時間 (請寫出其方法及檢定統計量)，跟據模型結果，幾歲時睡眠時間最少?

(5) 你如果認爲工作時間，教育年齡會因性別不同而不同，請問如何將此影響考慮在模型中，請寫出應估計的迴歸模型及其預期的影響。

5. (15分) 令 Y 的母體迴歸模型如下:

$$Y = \beta_0 + \beta_1 X_1 + \beta_2 X_2 + \beta_3 X_3 + \epsilon.$$

(1) $\hat{\beta}_0, \hat{\beta}_1, \hat{\beta}_2, \hat{\beta}_3$ 爲 OLS 的迴歸結果，設 C_0, C_1, C_2, C_3, 爲非零常數，若將 $C_0 Y$ 對 $C_1 X_1$, $C_2 X_2$, $C_3 X_3$ 迴歸，其 OLS 的結果表爲 $\tilde{\beta}_0, \tilde{\beta}_i$, $i = 1, 2, 3$，請證明 $\tilde{\beta}_0 = C_0 \hat{\beta}_0$, $\tilde{\beta}_i = \frac{C_0}{C_i} \hat{\beta}_i$, $i = 1, 2, 3$。

接次頁

(2) 設迴歸模型滿足 Gauss Markov 定理的假設，若我們估計模型時，忽略 X_3，令 $\check{\beta}_0, \check{\beta}_1, \check{\beta}_2$ 爲 Y 對 X_1 與 X_2 的 OLS 估計式，請證明

$$\mathbb{E}[\check{\beta}_1] = \beta_1 + \beta_3 \frac{\sum \widehat{U} X_3}{\sum \widehat{U}^2}$$

其中 $\widehat{U}$ 爲 X_1 對 X_2 迴歸的殘差項。

(*Hint:* $\check{\beta}_1$ 可表爲 $\frac{\sum \widehat{U} Y}{\sum \widehat{U}^2}$，你可由此式來幫助你證明。)

(3) 若利用上述模型擬預測當 $X_1 = X_{10}, X_2 = X_{20}, X_3 = X_{30}$ 時之 $\mathbb{E}[Y_0]$ 及其信賴區間，某研究生指出若將 Y 對 $X_1 - X_{10}, X_2 - X_{20}, X_3 - X_{30}$ 進行迴歸，則其迴歸結果的截距項可估計 $\mathbb{E}[Y_0]$，截距項的信賴區間即爲 $\mathbb{E}[Y_0]$ 之信賴區間，請證明此方法的正確性。(*Note:* $\mathbb{E}[Y_0] = \beta_0 + \beta_1 X_{10} + \beta_2 X_{20} + \beta_3 X_{30}$)