> **Instructions:**
> - Write in Chinese or English only.
> - Make sure all your answers are legible, unquestionably labeled, and clearly explained (with equations if possible).
> - A standard normal probability table is attached on the last page.

1. (16 points) A retail chain has set each of its stores a target of earning \$2 million in revenue during a month. The 100 stores are in locations with similar demographics. The head office believes that each store has a probability of 0.4 of reaching the revenue target, and that different stores' performances are independent of one another. The managers of a store that reaches the revenue target receive \$80,000 in bonuses.

   (a) (8 points) Calculate the expected value and standard deviation of the number of stores that will reach the revenue target.

   > **解：** 設 X 為達成營收目標的商店數量。由題意知，每家商店達成目標的機率為 $p = 0.4$，共有 $n = 100$ 家商店，且各商店表現獨立。因此 $X \sim \text{Binomial}(n = 100, p = 0.4)$。
   > **期望值：**
   > $$\mathbb{E}[X] = np = 100 \times 0.4 = \boxed{40}$$
   > **變異數與標準差：**
   > $$\text{Var}(X) = np(1-p) = 100 \times 0.4 \times 0.6 = 24$$
   > $$標準差 = \sqrt{\text{Var}(X)} = \sqrt{24} = 2\sqrt{6} \approx \boxed{4.899}$$

   (b) (8 points) Calculate the probability that the chain pays out more than \$3 million in bonuses.

   > **解：** 每家達成目標的商店經理可獲得 \$80,000 獎金。若總獎金超過 \$3,000,000，則達成目標的商店數必須滿足：
   > $$80,000 \times X > 3,000,000 \implies X > 37.5$$
   > 由於 X 為整數，故需要 $X \geq 38$。
   > 我們需要計算 $P(X \geq 38)$。
   > 由於 $n = 100$ 夠大，且 $np = 40 > 5$、$n(1-p) = 60 > 5$，可使用常態近似：
   > $$X \overset{\text{approx}}{\sim} N(\mu = 40, \sigma^2 = 24)$$
   > 使用連續性校正 (continuity correction)：
   > $$P(X \geq 38) = P(X > 37.5) \approx P\left(Z > \frac{37.5 - 40}{\sqrt{24}}\right) = P\left(Z > \frac{-2.5}{4.899}\right) = P(Z > -0.51)$$
   > 由標準常態分佈的對稱性：
   > $$P(Z > -0.51) = 1 - P(Z \leq -0.51) = 1 - 0.3050 = \boxed{0.6950}$$
   > （查表：$P(Z \leq -0.5) = 0.3085$，$P(Z \leq -0.51) \approx 0.3050$）
   > **備註：** 若不使用連續性校正：
   > $$P(X \geq 38) \approx P\left(Z \geq \frac{38 - 40}{\sqrt{24}}\right) = P(Z \geq -0.408) = 1 - P(Z \leq -0.41) \approx 1 - 0.3409 = 0.6591$$

2. (18 points) The admissions office at a university is determining how much weight to put on an applicant's high school GPA in making admissions decisions. It has collected the high school GPAs ($GPAHS$) and the college first-year GPAs ($GPAFY$) of a random sample of 75 students who just completed their first year at the university. Table 1 gives the summary statistics.

Assume that the sample satisfies the assumptions of the random sampling regression model. The estimation results for the linear regression $\widehat{GPAFY}_i = \beta_0 + \beta_1 GPAHS_i$ are summarized in Table 2.

Table 1: Summary Statistics

Variable	Obs	Mean	S.D.	Min	Max
$GPAFY$	75	2.97	0.51	1.77	4.00
$GPAHS$	75	3.44	0.32	2.62	4.00

Table 2: Regression Estimates

Variable	Estimate	Std. Error
$GPAHS$	0.3109	0.1491
Constant	1.9046	0.5106

$n = 75$. The dependent variable is $GPAFY$.
Heteroskedasticity-robust standard errors are presented.

(a) (6 points) Do the data provide convincing evidence that a student's high-school GPA is associated with his or her college first-year GPA? State the hypotheses and conduct the test at a 5% significance level.

> **解： 假設檢定：**
>
> $$H_0 : \beta_1 = 0 \quad \text{（高中 GPA 與大學 GPA 無關聯）}$$
> $$H_1 : \beta_1 \neq 0 \quad \text{（高中 GPA 與大學 GPA 有關聯）}$$
>
> **檢定統計量：**
>
> $$t = \frac{\hat{\beta}_1 - 0}{SE(\hat{\beta}_1)} = \frac{0.3109}{0.1491} \approx 2.085$$
>
> **自由度：** $df = n - 2 = 75 - 2 = 73$
> **臨界值：** 在顯著水準 $\alpha = 0.05$ 下，雙尾檢定的臨界值約為 $t_{0.025,73} \approx 1.993$（當 df 大時接近 1.96）。
> **決策：** 由於 $|t| = 2.085 > 1.993$，我們**拒絕虛無假設**。
> **結論：** 在 5% 顯著水準下，有充分證據顯示學生的高中 GPA 與其大學第一年 GPA 有顯著關聯。

(b) (6 points) A student improves her high-school GPA by 1.2. Construct a 90% confidence interval for change in her college first-year GPA.

> **解：** 當高中 GPA 增加 $\Delta GPAHS = 1.2$ 時，大學 GPA 的預期變化為：
>
> $$\Delta GPAFY = \hat{\beta}_1 \times \Delta GPAHS = 0.3109 \times 1.2 = 0.37308$$
>
> 變化量的標準誤為：
>
> $$SE(\Delta \widehat{GPAFY}) = |1.2| \times SE(\hat{\beta}_1) = 1.2 \times 0.1491 = 0.17892$$
>
> 在 90% 信賴水準下，$df = 73$ 時，$t_{0.05,73} \approx 1.666$（或使用常態近似 $z_{0.05} = 1.645$）。
> **90% 信賴區間：**
>
> $$\Delta GPAFY \pm t_{0.05,73} \times SE = 0.37308 \pm 1.666 \times 0.17892$$

$$= 0.37308 \pm 0.2981$$
$$= \boxed{(0.075, 0.671)}$$

解釋：我們有 90% 的信心認為，當高中 GPA 增加 1.2 時，大學第一年 GPA 的增加量介於 0.075 至 0.671 之間。

(c) (6 points) The office estimates a different model $\widehat{GPAHS}_i = \gamma_0 + \gamma_1 GPAFY_i$ using the data. Calculate the least-squares estimate of the slope coefficient $\hat{\gamma}_1$.

解：在原迴歸模型 $GPAFY_i = \beta_0 + \beta_1 GPAHS_i + \epsilon_i$ 中：

$$\hat{\beta}_1 = \frac{\sum_{i=1}^n (GPAHS_i - \overline{GPAHS})(GPAFY_i - \overline{GPAFY})}{\sum_{i=1}^n (GPAHS_i - \overline{GPAHS})^2} = \frac{S_{XY}}{S_{XX}}$$

在新模型 $GPAHS_i = \gamma_0 + \gamma_1 GPAFY_i + u_i$ 中：

$$\hat{\gamma}_1 = \frac{\sum_{i=1}^n (GPAFY_i - \overline{GPAFY})(GPAHS_i - \overline{GPAHS})}{\sum_{i=1}^n (GPAFY_i - \overline{GPAFY})^2} = \frac{S_{XY}}{S_{YY}}$$

因此：

$$\hat{\beta}_1 \times \hat{\gamma}_1 = \frac{S_{XY}}{S_{XX}} \times \frac{S_{XY}}{S_{YY}} = \frac{S_{XY}^2}{S_{XX} \cdot S_{YY}} = r^2$$

其中 r 是 $GPAHS$ 和 $GPAFY$ 之間的樣本相關係數。
另一方面，我們知道：

$$\hat{\beta}_1 = r \times \frac{S_Y}{S_X} = r \times \frac{0.51}{0.32}$$

所以：

$$r = \hat{\beta}_1 \times \frac{S_X}{S_Y} = 0.3109 \times \frac{0.32}{0.51} = 0.3109 \times 0.6275 = 0.1951$$

因此：

$$\hat{\gamma}_1 = r \times \frac{S_X}{S_Y} = 0.1951 \times \frac{0.32}{0.51} = 0.1224$$

或者直接使用關係式：

$$\hat{\gamma}_1 = \hat{\beta}_1 \times \frac{(S_X)^2}{(S_Y)^2} = 0.3109 \times \frac{(0.32)^2}{(0.51)^2} = 0.3109 \times \frac{0.1024}{0.2601} = 0.3109 \times 0.3937 = \boxed{0.1224}$$

3. (16 points) A chain of convenience stores wants the stores' stocks of bottled water to average 50 boxes at any given time, and plans to change its inventory policy if there is strong evidence that such an average is not being maintained.

(a) (4 points) State appropriate null and alternative hypotheses in symbols and in words for this scenario.

> **解： 以符號表示：**
>
> $$H_0 : \mu = 50$$
> $$H_1 : \mu \neq 50$$
>
> **以文字敘述：**
>
> - H_0（虛無假設）：商店瓶裝水庫存的平均數等於 50 箱。
>
> - H_1（對立假設）：商店瓶裝水庫存的平均數不等於 50 箱。
>
> 這是一個**雙尾檢定**，因為我們關心平均庫存是否偏離目標值（無論偏高或偏低）。

(b) (6 points) Suppose that the chain obtains the current bottled water inventories from 82 randomly selected stores. The sample mean is 52.5 boxes, and sample standard deviation is 9.8 boxes. Find the p-value and evaluate whether to reject the null hypothesis using a significance level of 0.05.

> **解： 給定資訊：**
>
> - 樣本數 $n = 82$
>
> - 樣本平均數 $\bar{x} = 52.5$
>
> - 樣本標準差 $s = 9.8$
>
> - 顯著水準 $\alpha = 0.05$
>
> **檢定統計量：**
>
> $$t = \frac{\bar{x} - \mu_0}{s/\sqrt{n}} = \frac{52.5 - 50}{9.8/\sqrt{82}} = \frac{2.5}{9.8/9.055} = \frac{2.5}{1.082} \approx 2.31$$
>
> **自由度：** $df = n - 1 = 81$
> **p-value 計算：** 由於這是雙尾檢定：
>
> $$\text{p-value} = 2 \times P(T > 2.31) \quad \text{where } T \sim t_{81}$$
>
> 當自由度較大時，t 分布接近標準常態分布。查表或計算：
>
> $$P(Z > 2.31) \approx 0.0104$$
>
> $$\text{p-value} \approx 2 \times 0.0104 = \boxed{0.0208}$$
>
> **決策：** 由於 p-value $= 0.0208 < 0.05 = \alpha$，我們**拒絕虛無假設**。
> **結論：** 在 5% 顯著水準下，有充分證據顯示商店瓶裝水的平均庫存不等於 50 箱。連鎖店應該考慮改變其庫存政策。

(c) (6 points) If the true mean inventory is 54 boxes, what is the probability that the test from part

(b) results in a Type II error?

解： **Type II 錯誤**：當虛無假設為假時，未能拒絕虛無假設的機率。
設定：真實平均數 $\mu = 54$，我們在 $\alpha = 0.05$ 下進行雙尾檢定。
拒絕域：在虛無假設 $H_0 : \mu = 50$ 下，當 $|t| > t_{0.025,81} \approx 1.99$ 時拒絕 H_0。
轉換為 $\bar{x}$ 的範圍：

$$\bar{x} < 50 - 1.99 \times \frac{9.8}{\sqrt{82}} = 50 - 1.99 \times 1.082 = 50 - 2.15 = 47.85$$

或

$$\bar{x} > 50 + 2.15 = 52.15$$

當 $\mu = 54$ 時的 Type II 錯誤機率：

$$\beta = P(47.85 \leq \bar{x} \leq 52.15 \mid \mu = 54)$$

將 $\bar{x}$ 標準化（在 $\mu = 54$ 下）：

$$Z = \frac{\bar{x} - 54}{9.8/\sqrt{82}} = \frac{\bar{x} - 54}{1.082}$$

$$\beta = P\left(\frac{47.85 - 54}{1.082} \leq Z \leq \frac{52.15 - 54}{1.082}\right) = P(-5.68 \leq Z \leq -1.71)$$

$$\beta = \Phi(-1.71) - \Phi(-5.68) \approx 0.0436 - 0 = \boxed{0.0436}$$

（查表：$P(Z \leq -1.71) \approx 0.0436$，$P(Z \leq -1.7) = 0.0446$）
結論：當真實平均庫存為 54 箱時，犯 Type II 錯誤的機率約為 4.36%。這意味著檢定力（power）約為 $1 - 0.0436 = 0.9564$，即 95.64%。

4. (30 points) Please answer the following questions with either "True" or "False," then briefly justify your answers. Answers without justifications will receive *no* points.

$\{y_i\}_{i=1}^n$ are random variables such that for $i = 1, 2, \ldots, n$, $\mathbb{E}(y_i) = b_0 + b_1 x_{i1} + b_2 x_{i2} + b_3 x_{i3}$ for some $\mathbf{b} = [b_0, b_1, b_2, b_3]^\top$, where $\{\mathbf{x}_i\}_{i=1}^n = \{[1, x_{i1}, x_{i2}, x_{i3}]^\top\}_{i=1}^n$ are non-random and *not* multicollinear, $\mathrm{Var}(y_i) = \sigma_0^2$, and $\mathrm{corr}(y_i, y_j) = \rho$ for $j \neq i$.

(a) (6 points) Suppose that $\rho \neq 0$. Then $\hat{\mathbf{b}}_{OLS} = [\hat{b}_0, \hat{b}_1, \hat{b}_2, \hat{b}_3]^\top$, the ordinary least squares (OLS) estimator, is a *biased* estimator for $\mathbf{b}$.

> **解：False（錯誤）**
> **理由：**OLS 估計量的不偏性（unbiasedness）只依賴於以下條件：
>
> 1. $\mathbb{E}(y_i|\mathbf{x}_i) = \mathbf{x}_i^\top \mathbf{b}$（正確設定的條件期望值）
>
> 2. $\mathbf{x}_i$ 為非隨機（fixed）或與誤差項獨立
>
> OLS 估計量為：
> $$\hat{\mathbf{b}}_{OLS} = (\mathbf{X}^\top \mathbf{X})^{-1} \mathbf{X}^\top \mathbf{y}$$
>
> 其期望值：
> $$\mathbb{E}[\hat{\mathbf{b}}_{OLS}] = \mathbb{E}[(\mathbf{X}^\top \mathbf{X})^{-1} \mathbf{X}^\top \mathbf{y}] = (\mathbf{X}^\top \mathbf{X})^{-1} \mathbf{X}^\top \mathbb{E}[\mathbf{y}] = (\mathbf{X}^\top \mathbf{X})^{-1} \mathbf{X}^\top \mathbf{X} \mathbf{b} = \mathbf{b}$$
>
> 誤差項之間的相關性（$\rho \neq 0$）會影響 OLS 估計量的**效率**（efficiency）和**變異數估計**，但**不會**影響其不偏性。

(b) (6 points) If we have a sample size $n = 500$, then $\hat{\mathbf{b}}_{OLS} = [\hat{b}_0, \hat{b}_1, \hat{b}_2, \hat{b}_3]^\top$, the OLS estimator for $\mathbf{b}$, is *normally-distributed* in general.

> **解：False（錯誤）**
> **理由：**OLS 估計量是否服從常態分布取決於誤差項的分布：
>
> 1. 若 y_i（或等價地，誤差項 ϵ_i）服從常態分布，則 $\hat{\mathbf{b}}_{OLS}$ 精確地服從常態分布（對任何樣本大小）。
>
> 2. 若 y_i 不服從常態分布，則只有當 $n \to \infty$ 時，根據中央極限定理，$\hat{\mathbf{b}}_{OLS}$ 才**漸近地**服從常態分布。
>
> 題目說的是「in general」（一般而言），且沒有假設 y_i 服從常態分布。雖然 $n = 500$ 是一個相當大的樣本，但我們不能說 $\hat{\mathbf{b}}_{OLS}$ 在有限樣本下**精確**服從常態分布。
> 更精確地說，當 $n = 500$ 時，$\hat{\mathbf{b}}_{OLS}$ **近似**服從常態分布（by CLT），但不是**精確**常態分布。

(c) (6 points) In the beginning we said that $\{\mathbf{x}_i\}_{i=1}^n = \{[1, x_{i1}, x_{i2}, x_{i3}]^\top\}_{i=1}^n$ are *not* multicollinear. It is important because weird things happen when the independent variables are multicollinear. For example, if $x_{i1} = 2x_{i2}$, then $\hat{b}_1 = 0.5\hat{b}_2$.

> **解：False（錯誤）**
> **理由：**若存在完美多重共線性（perfect multicollinearity），例如 $x_{i1} = 2x_{i2}$，則矩陣 $\mathbf{X}^\top \mathbf{X}$ 是**奇異的**（singular），不可逆。
> 在這種情況下，OLS 估計量 $\hat{\mathbf{b}}_{OLS} = (\mathbf{X}^\top \mathbf{X})^{-1} \mathbf{X}^\top \mathbf{y}$ **無法計算**，因為 $(\mathbf{X}^\top \mathbf{X})^{-1}$ 不存在。
> 因此，$\hat{b}_1$ 和 $\hat{b}_2$ 都無法被唯一確定，說 $\hat{b}_1 = 0.5\hat{b}_2$ 是沒有意義的。

多重共線性的問題在於無法分離各個自變量的個別效果，而不是它們之間存在某種固定的數學關係。

(d) (6 points) Consider the *Eicker-White* variance-covariance matrix estimator

$$\hat{D} = \left(\frac{1}{n}\sum_{i=1}^n \mathbf{x}_i\mathbf{x}_i^\top\right)^{-1} \left(\frac{1}{n}\sum_{i=1}^n \hat{\epsilon}_i^2 \mathbf{x}_i\mathbf{x}_i^\top\right) \left(\frac{1}{n}\sum_{i=1}^n \mathbf{x}_i\mathbf{x}_i^\top\right)^{-1},$$

where $\hat{\epsilon}_i = y_i - \mathbf{x}_i^\top \hat{\mathbf{b}}_{OLS}$ is the residual. Suppose that $\rho = 0$, i.e., $\{y_i\}_{i=1}^n$ are homoskedastic. Then $\hat{D}$ is an *inconsistent* estimator for the variance-covariance matrix of $\mathbf{b}_{OLS}$.

解： False（錯誤）

理由：Eicker-White（或稱 heteroskedasticity-robust）變異數-共變異數矩陣估計量在以下情況下都是一致的：

1. 存在異質變異（heteroskedasticity）時

2. 同質變異（homoskedasticity）時

當 $\rho = 0$（誤差項不相關）且 $\mathrm{Var}(y_i) = \sigma_0^2$（同質變異）時：

$$\mathrm{Var}(\hat{\mathbf{b}}_{OLS}) = \sigma_0^2 (\mathbf{X}^\top\mathbf{X})^{-1}$$

在這種情況下，Eicker-White 估計量仍然是一致的，因為：

$$\frac{1}{n}\sum_{i=1}^n \hat{\epsilon}_i^2 \mathbf{x}_i\mathbf{x}_i^\top \xrightarrow{p} \sigma_0^2 \cdot \mathbb{E}[\mathbf{x}_i\mathbf{x}_i^\top]$$

因此 $\hat{D}$ 會收斂到真實的變異數-共變異數矩陣。

然而，在同質變異下，Eicker-White 估計量雖然一致，但**效率較低**（less efficient）。傳統的 OLS 變異數估計量 $\hat{\sigma}^2(\mathbf{X}^\top\mathbf{X})^{-1}$ 在這種情況下更有效率。

(e) (6 points) Suppose that $\{x_{i1}, x_{i2}, x_{i3}\}_{i=1}^n$ are random variables. Consider two models.

$$y_i = c_0 + c_1 x_{i1} + u_i, \qquad\qquad\qquad \text{(Model 1)}$$
$$y_i = d_0 + d_1 x_{i1} + d_2 x_{i2} + v_i. \qquad\qquad \text{(Model 2)}$$

We estimate both models by the OLS estimator. Suppose $\hat{c}_1$, the OLS estimate for c_1 in Model 1, is numerically equivalent to $\hat{d}_1$, the OLS estimate for d_1 in Model 2, then we can conclude that $\{x_{i1}\}_{i=1}^n$ and $\{x_{i2}\}_{i=1}^n$ are independent.

解： False（錯誤）

理由：當 $\hat{c}_1 = \hat{d}_1$ 時，這只意味著 x_{i1} 和 x_{i2} 在樣本中是**不相關的**（uncorrelated），而非獨立（independent）。

具體來說，在 OLS 中，加入一個新的解釋變量 x_{i2} 後，$\hat{d}_1$ 與 $\hat{c}_1$ 相等的條件是：

$$\sum_{i=1}^n (x_{i1} - \bar{x}_1)(x_{i2} - \bar{x}_2) = 0$$

即樣本相關係數 $r_{x_1, x_2} = 0$。

然而，「不相關」（uncorrelated）不等於「獨立」（independent）：

- 不相關：$\text{Cov}(x_{i1}, x_{i2}) = 0$（或樣本中 $r = 0$）

- 獨立：對所有 Borel 集合 A, B，$P(x_{i1} \in A, x_{i2} \in B) = P(x_{i1} \in A) \cdot P(x_{i2} \in B)$

獨立一定導致不相關，但不相關**不一定**導致獨立。例如，若 $x_{i1} \sim N(0, 1)$，$x_{i2} = x_{i1}^2$，則 $\text{Cov}(x_{i1}, x_{i2}) = 0$，但 x_{i1} 和 x_{i2} 明顯不獨立。

5. (8 points) Consider the following dynamic panel model. For $i = 1, 2, \ldots, n$ and $t = 1, 2, \ldots, T$,

$$y_{i,t} = \alpha_i + \beta y_{i,t-1} + \epsilon_{i,t},$$

where $|\beta| < 1$, and $\{\epsilon_{i,t}\}$ are independent and identically distributed (i.i.d.) with a finite second moment. To eliminate the heterogeneous intercepts $\{\alpha_i\}$, we take the first difference:

$$\Delta y_{i,t} = \beta \Delta y_{i,t-1} + \Delta \epsilon_{i,t},$$

where $\Delta y_{i,t} = y_{i,t} - y_{i,t-1}$. Let

$$\hat{\beta}_{FD} = \frac{\sum_{i=1}^{n} \sum_{t=2}^{T} (\Delta y_{i,t-1})(\Delta y_{i,t})}{\sum_{i=1}^{n} \sum_{t=2}^{T} (\Delta y_{i,t-1})^2}.$$

Suppose that $n \to \infty$ while T is fixed. Is $\hat{\beta}_{FD}$ a consistent estimator for β? Justify your answer. Answers without justifications will receive *no* points.

解： False（$\hat{\beta}_{FD}$ 不是 β 的一致估計量）

證明：

首先，將 $\hat{\beta}_{FD}$ 改寫為：

$$\hat{\beta}_{FD} = \frac{\sum_{i=1}^{n} \sum_{t=2}^{T} (\Delta y_{i,t-1})(\Delta y_{i,t})}{\sum_{i=1}^{n} \sum_{t=2}^{T} (\Delta y_{i,t-1})^2}$$

將 $\Delta y_{i,t} = \beta \Delta y_{i,t-1} + \Delta \epsilon_{i,t}$ 代入分子：

$$\hat{\beta}_{FD} = \frac{\sum_{i=1}^{n} \sum_{t=2}^{T} (\Delta y_{i,t-1})(\beta \Delta y_{i,t-1} + \Delta \epsilon_{i,t})}{\sum_{i=1}^{n} \sum_{t=2}^{T} (\Delta y_{i,t-1})^2}$$

$$= \beta + \frac{\sum_{i=1}^{n} \sum_{t=2}^{T} (\Delta y_{i,t-1})(\Delta \epsilon_{i,t})}{\sum_{i=1}^{n} \sum_{t=2}^{T} (\Delta y_{i,t-1})^2}$$

現在分析 $\Delta y_{i,t-1}$ 與 $\Delta \epsilon_{i,t}$ 的相關性：

$$\Delta y_{i,t-1} = y_{i,t-1} - y_{i,t-2}$$
$$\Delta \epsilon_{i,t} = \epsilon_{i,t} - \epsilon_{i,t-1}$$

關鍵觀察：$y_{i,t-1}$ 包含 $\epsilon_{i,t-1}$（因為 $y_{i,t-1} = \alpha_i + \beta y_{i,t-2} + \epsilon_{i,t-1}$）。
因此：

$$\mathbb{E}[(\Delta y_{i,t-1})(\Delta \epsilon_{i,t})] = \mathbb{E}[(y_{i,t-1} - y_{i,t-2})(\epsilon_{i,t} - \epsilon_{i,t-1})]$$

$$= -\mathbb{E}[y_{i,t-1} \cdot \epsilon_{i,t-1}] = -\mathbb{E}[\epsilon_{i,t-1}^2] = -\sigma^2 \neq 0$$

由於 $\Delta y_{i,t-1}$ 與 $\Delta \epsilon_{i,t}$ 相關（correlation $\neq 0$），即使當 $n \to \infty$：

$$\text{plim}_{n \to \infty} \hat{\beta}_{FD} = \beta + \frac{\mathbb{E}[(\Delta y_{i,t-1})(\Delta \epsilon_{i,t})]}{\mathbb{E}[(\Delta y_{i,t-1})^2]} \neq \beta$$

這是著名的 **Nickell 偏誤**（Nickell bias），在動態面板模型中使用一階差分法時會出現。當 T 固定且 $n \to \infty$ 時，偏誤不會消失。

結論： $\hat{\beta}_{FD}$ **不是** β 的一致估計量，因為解釋變量 $\Delta y_{i,t-1}$ 與誤差項 $\Delta \epsilon_{i,t}$ 相關。

---

6. (12 points) Suppose that $\{y_i\}_{i=1}^n$ are generated according to

$$y_i = \beta_0 x_i + \epsilon_i,$$

where $\{[x_i, \epsilon_i]^\top\}_{i=1}^n$ are independent and identically distributed (i.i.d.) random vectors such that $\mathbb{E}(x_i) = \mu_x$, $\mathbb{E}(\epsilon_i) = 0$, $\mathrm{Var}(x_i) = \sigma_x^2$, $\mathrm{Var}(\epsilon_i) = \sigma_\epsilon^2$, and $\mathbb{E}(x_i\epsilon_i) = 0$.

Unfortunately, $\{x_i\}_{i=1}^n$ cannot be observed directly. Instead, we observe $\{w_i\}_{i=1}^n$ and $\{z_i\}_{i=1}^n$ with

$$w_i = x_i + u_i, \quad \text{and} \quad z_i = x_i + v_i,$$

where $\{u_i\}_{i=1}^n$ and $\{v_i\}_{i=1}^n$ are i.i.d. random variables such that $\mathbb{E}(u_i) = \mathbb{E}(u_ix_i) = \mathbb{E}(u_i\epsilon_i) = 0$, $\mathrm{Var}(u_i) = \sigma_u^2$, and $\mathbb{E}(v_i) = \mathbb{E}(v_ix_i) = \mathbb{E}(v_i\epsilon_i) = 0$, $\mathrm{Var}(v_i) = \sigma_v^2$, and $\mathbb{E}(u_iv_i) = 0$.

(a) (6 points) Regress y_i on w_i without intercept and obtain the ordinary least squares (OLS) estimate $\tilde{\beta}$. Determine the probability limit of $\tilde{\beta}$.

解：無截距項的 OLS 估計量為：

$$\tilde{\beta} = \frac{\sum_{i=1}^n w_iy_i}{\sum_{i=1}^n w_i^2}$$

將 $y_i = \beta_0 x_i + \epsilon_i$ 和 $w_i = x_i + u_i$ 代入：

$$\tilde{\beta} = \frac{\sum_{i=1}^n (x_i + u_i)(\beta_0 x_i + \epsilon_i)}{\sum_{i=1}^n (x_i + u_i)^2}$$

展開分子：

$$\sum_{i=1}^n w_iy_i = \sum_{i=1}^n (\beta_0 x_i^2 + x_i\epsilon_i + \beta_0 u_ix_i + u_i\epsilon_i)$$

展開分母：

$$\sum_{i=1}^n w_i^2 = \sum_{i=1}^n (x_i^2 + 2x_iu_i + u_i^2)$$

由大數法則，當 $n \to \infty$：

$$\frac{1}{n}\sum_{i=1}^n x_i^2 \xrightarrow{p} \mathbb{E}[x_i^2] = \mu_x^2 + \sigma_x^2$$

$$\frac{1}{n}\sum_{i=1}^n x_i\epsilon_i \xrightarrow{p} \mathbb{E}[x_i\epsilon_i] = 0$$

$$\frac{1}{n}\sum_{i=1}^n u_ix_i \xrightarrow{p} \mathbb{E}[u_ix_i] = 0$$

$$\frac{1}{n}\sum_{i=1}^n u_i\epsilon_i \xrightarrow{p} \mathbb{E}[u_i\epsilon_i] = 0$$

$$\frac{1}{n}\sum_{i=1}^n x_iu_i \xrightarrow{p} 0$$

$$\frac{1}{n}\sum_{i=1}^n u_i^2 \xrightarrow{p} \sigma_u^2$$

因此：

$$\text{plim}_{n \to \infty} \tilde{\beta} = \frac{\beta_0(\mu_x^2 + \sigma_x^2) + 0 + 0 + 0}{(\mu_x^2 + \sigma_x^2) + 0 + \sigma_u^2} = \frac{\beta_0(\mu_x^2 + \sigma_x^2)}{\mu_x^2 + \sigma_x^2 + \sigma_u^2}$$

由於 $\mathbb{E}[x_i^2] = \text{Var}(x_i) + (\mathbb{E}[x_i])^2 = \sigma_x^2 + \mu_x^2$，可簡化為：

$$\boxed{\text{plim}\tilde{\beta} = \beta_0 \cdot \frac{\sigma_x^2 + \mu_x^2}{\sigma_x^2 + \mu_x^2 + \sigma_u^2}}$$

這顯示了經典的**衰減偏誤**(attenuation bias)：由於測量誤差，估計量向零偏移($|\text{plim}\tilde{\beta}| < |\beta_0|$，假設 $\sigma_u^2 > 0$)。

(b) (6 points) How would you estimate β_0 consistently? Write down your estimator and briefly justify your answer. Answers without justifications will receive *no* points.

解：工具變量法 (Instrumental Variables, IV)

使用 z_i 作為 w_i 的工具變量，IV 估計量（無截距項）為：

$$\hat{\beta}_{IV} = \frac{\sum_{i=1}^n z_i y_i}{\sum_{i=1}^n z_i w_i}$$

驗證一致性：

將 $y_i = \beta_0 x_i + \epsilon_i$、$w_i = x_i + u_i$、$z_i = x_i + v_i$ 代入：

分子：

$$\sum_{i=1}^n z_i y_i = \sum_{i=1}^n (x_i + v_i)(\beta_0 x_i + \epsilon_i) = \sum_{i=1}^n (\beta_0 x_i^2 + x_i \epsilon_i + \beta_0 v_i x_i + v_i \epsilon_i)$$

分母：

$$\sum_{i=1}^n z_i w_i = \sum_{i=1}^n (x_i + v_i)(x_i + u_i) = \sum_{i=1}^n (x_i^2 + x_i u_i + v_i x_i + v_i u_i)$$

由大數法則：

$$\text{plim}\frac{1}{n}\sum z_i y_i = \beta_0(\sigma_x^2 + \mu_x^2) + 0 + 0 + 0 = \beta_0(\sigma_x^2 + \mu_x^2)$$

$$\text{plim}\frac{1}{n}\sum z_i w_i = (\sigma_x^2 + \mu_x^2) + 0 + 0 + 0 = \sigma_x^2 + \mu_x^2$$

因此：

$$\text{plim}\hat{\beta}_{IV} = \frac{\beta_0(\sigma_x^2 + \mu_x^2)}{\sigma_x^2 + \mu_x^2} = \boxed{\beta_0}$$

IV 方法有效的原因：

1. **相關性** (Relevance)：z_i 與 w_i 相關，因為它們都包含 x_i：

$$\text{Cov}(z_i, w_i) = \text{Cov}(x_i + v_i, x_i + u_i) = \text{Var}(x_i) = \sigma_x^2 \neq 0$$

2. **外生性** (Exogeneity)：z_i 與迴歸誤差 $(\epsilon_i - \beta_0 u_i)$ 不相關：

$$\text{Cov}(z_i, \epsilon_i) = \text{Cov}(x_i + v_i, \epsilon_i) = 0$$
$$\text{Cov}(z_i, u_i) = \text{Cov}(x_i + v_i, u_i) = 0$$

z_i 是 w_i 的有效工具變量，因為 z_i 與內生的測量誤差 u_i 不相關，但與真實變量 x_i 相關。

Standard normal probability table

(標準常態分佈曲線，陰影區域表示 $P(Z \leq z)$)

此表給出負值 Z 的累積機率 $P(Z \leq z)$。對於正值 Z，使用對稱性：$P(Z \leq z) = 1 - P(Z \leq -z)$。

Z	0.00	0.01	0.02	0.03	0.04	0.05	0.06	0.07	0.08	0.09
−3.4	0.0003	0.0003	0.0003	0.0003	0.0003	0.0003	0.0003	0.0003	0.0003	0.0002
−3.3	0.0005	0.0005	0.0005	0.0004	0.0004	0.0004	0.0004	0.0004	0.0004	0.0003
−3.2	0.0007	0.0007	0.0006	0.0006	0.0006	0.0006	0.0006	0.0005	0.0005	0.0005
−3.1	0.0010	0.0009	0.0009	0.0009	0.0008	0.0008	0.0008	0.0008	0.0007	0.0007
−3.0	0.0013	0.0013	0.0013	0.0012	0.0012	0.0011	0.0011	0.0011	0.0010	0.0010
−2.9	0.0019	0.0018	0.0018	0.0017	0.0016	0.0016	0.0015	0.0015	0.0014	0.0014
−2.8	0.0026	0.0025	0.0024	0.0023	0.0023	0.0022	0.0021	0.0021	0.0020	0.0019
−2.7	0.0035	0.0034	0.0033	0.0032	0.0031	0.0030	0.0029	0.0028	0.0027	0.0026
−2.6	0.0047	0.0045	0.0044	0.0043	0.0041	0.0040	0.0039	0.0038	0.0037	0.0036
−2.5	0.0062	0.0060	0.0059	0.0057	0.0055	0.0054	0.0052	0.0051	0.0049	0.0048
−2.4	0.0082	0.0080	0.0078	0.0075	0.0073	0.0071	0.0069	0.0068	0.0066	0.0064
−2.3	0.0107	0.0104	0.0102	0.0099	0.0096	0.0094	0.0091	0.0089	0.0087	0.0084
−2.2	0.0139	0.0136	0.0132	0.0129	0.0125	0.0122	0.0119	0.0116	0.0113	0.0110
−2.1	0.0179	0.0174	0.0170	0.0166	0.0162	0.0158	0.0154	0.0150	0.0146	0.0143
−2.0	0.0228	0.0222	0.0217	0.0212	0.0207	0.0202	0.0197	0.0192	0.0188	0.0183
−1.9	0.0287	0.0281	0.0274	0.0268	0.0262	0.0256	0.0250	0.0244	0.0239	0.0233
−1.8	0.0359	0.0351	0.0344	0.0336	0.0329	0.0322	0.0314	0.0307	0.0301	0.0294
−1.7	0.0446	0.0436	0.0427	0.0418	0.0409	0.0401	0.0392	0.0384	0.0375	0.0367
−1.6	0.0548	0.0537	0.0526	0.0516	0.0505	0.0495	0.0485	0.0475	0.0465	0.0455
−1.5	0.0668	0.0655	0.0643	0.0630	0.0618	0.0606	0.0594	0.0582	0.0571	0.0559
−1.4	0.0808	0.0793	0.0778	0.0764	0.0749	0.0735	0.0721	0.0708	0.0694	0.0681
−1.3	0.0968	0.0951	0.0934	0.0918	0.0901	0.0885	0.0869	0.0853	0.0838	0.0823
−1.2	0.1151	0.1131	0.1112	0.1093	0.1075	0.1056	0.1038	0.1020	0.1003	0.0985
−1.1	0.1357	0.1335	0.1314	0.1292	0.1271	0.1251	0.1230	0.1210	0.1190	0.1170
−1.0	0.1587	0.1562	0.1539	0.1515	0.1492	0.1469	0.1446	0.1423	0.1401	0.1379
−0.9	0.1841	0.1814	0.1788	0.1762	0.1736	0.1711	0.1685	0.1660	0.1635	0.1611
−0.8	0.2119	0.2090	0.2061	0.2033	0.2005	0.1977	0.1949	0.1922	0.1894	0.1867
−0.7	0.2420	0.2389	0.2358	0.2327	0.2296	0.2266	0.2236	0.2206	0.2177	0.2148
−0.6	0.2743	0.2709	0.2676	0.2643	0.2611	0.2578	0.2546	0.2514	0.2483	0.2451
−0.5	0.3085	0.3050	0.3015	0.2981	0.2946	0.2912	0.2877	0.2843	0.2810	0.2776
−0.4	0.3446	0.3409	0.3372	0.3336	0.3300	0.3264	0.3228	0.3192	0.3156	0.3121
−0.3	0.3821	0.3783	0.3745	0.3707	0.3669	0.3632	0.3594	0.3557	0.3520	0.3483
−0.2	0.4207	0.4168	0.4129	0.4090	0.4052	0.4013	0.3974	0.3936	0.3897	0.3859
−0.1	0.4602	0.4562	0.4522	0.4483	0.4443	0.4404	0.4364	0.4325	0.4286	0.4247
−0.0	0.5000	0.4960	0.4920	0.4880	0.4840	0.4801	0.4761	0.4721	0.4681	0.4641

*For $Z \leq -3.50$, the probability is less than or equal to 0.0002.