Multiple Choice Questions. Notes:

(1) Please choose only one of the answer choices (a)-(e).

(2) Write down your answers on the scantron answer sheet.

(3) Each question is worth 5 points.

1. Let $\{(e_i, X_i')\}$ be a sequence of independent and $N(0, I_{k+1})$-distributed random vectors for some $k := \dim(X_i) > 1$, and X_{ij} be the jth element of X_i for $j = 1, \ldots, k$. Assume that Y_i is a random variable defined in the following way:

$$Y_i = \sum_{j=1}^{k} X_{ij} + \left(\sum_{j=1}^{k} X_{ij}^2\right)^2 + e_i.$$

Let $g(X_i)$ be an arbitrary transformation of X_i such that $\mathbb{E}[(Y_i - g(X_i))^2]$ is defined, and $g^*(X_i)$ be the optimal choice of $g(X_i)$ which minimizes $\mathbb{E}[(Y_i - g(X_i))^2]$. Which of the following is right when $k = 10$?

(a) $\mathbb{E}[g^*(X_i)] = 60$

(b) $\mathbb{E}[g^*(X_i)] = 80$

(c) $\mathbb{E}[g^*(X_i)] = 160$

(d) $\mathbb{E}[g^*(X_i)] = 180$

(e) None of the above choices (a)-(d).

2. Let $\{(X_{1i}, X_{2i})\}_{i=1}^{n}$ be a sequence of IID random vectors with the distribution:

$$\begin{bmatrix} X_{1i} \\ X_{2i} \end{bmatrix} \sim N\left(\begin{bmatrix} 99 \\ 89 \end{bmatrix}, \begin{bmatrix} 2 & 0.5 \\ 0.5 & 4 \end{bmatrix}\right).$$

Denote $\bar{X}_1 = n^{-1}\sum_{i=1}^{n} X_{1i}$ and $\bar{X}_2 = n^{-1}\sum_{i=1}^{n} X_{2i}$. Which of the following is right when $n = 100$?

(a) $\text{var}[\bar{X}_1 + \bar{X}_2] = 0.160$

(b) $\text{var}[\bar{X}_1 + \bar{X}_2] = 0.155$

(c) $\text{var}[\bar{X}_1 + \bar{X}_2] = 0.150$

(d) $\text{var}[\bar{X}_1 + \bar{X}_2] = 0.070$

(e) None of the above choices (a)-(d).

3. Let $\{Y_i\}_{i=1}^n$ be a sequence of independent and $t(4)$-distributed random variables. Consider a linear regression:

$$Y_i = \alpha + e_i,$$

where α is a parameter, and e_i is the error term. Let $\hat{\alpha}$ be the least squares estimator of α. Which of the following is right?

(a) $\mathbb{E}[\hat{\alpha}] = 0$ and $\text{var}[\hat{\alpha}] = \frac{1}{4n}$

(b) $\mathbb{E}[\hat{\alpha}] = 0$ and $\text{var}[\hat{\alpha}] = \frac{1}{n}$

(c) $\mathbb{E}[\hat{\alpha}] = 0$ and $\text{var}[\hat{\alpha}] = \frac{1}{2n}$

(d) $\mathbb{E}[\hat{\alpha}] = 0$ and $\text{var}[\hat{\alpha}] = \frac{2}{n}$

(e) None of the above choices (a)-(d).

4. Let $\{X_i\}_{i=1}^n$ be a sequence of IID random variables with the probability density function:

$$f(x) = \exp(-x), \quad \text{for} \quad x \in \mathbb{R}.$$

Denote $\sigma^2 := \text{var}[X_i]$, $\bar{X} := n^{-1}\sum_{i=1}^n X_i$ and $\hat{\sigma}^2 := n^{-1}\sum_{i=1}^n (X_i - \bar{X})^2$. Using a suitable large-sample method, we obtain that $\sqrt{n}(\hat{\sigma} - \sigma)$ has the limiting distribution $N(0, \kappa)$, as $n \to \infty$, for some $v > 0$. Which of the following is right?

(a) $\kappa = 1$

(b) $\kappa = 2$

(c) $\kappa = 4$

(d) $\kappa = 8$

(e) None of the above choices (a)-(d).

接次頁

5. Assume that $\{Y_i\}_{i=1}^n$ and $\{Z_i\}_{i=1}^n$ are two independent sequences of IID random variables with finite fourth moments. Consider the following two regressions:

$$Y_i = \alpha_Y + e_i$$

and

$$Z_i = \alpha_Z + u_i,$$

where α_Y and α_Z are parameters, and

$$\begin{bmatrix} e_i \\ u_i \end{bmatrix} \sim N\left(\begin{bmatrix} 0 \\ 0 \end{bmatrix}, \begin{bmatrix} 2 & 0.5 \\ 0.5 & 1 \end{bmatrix}\right).$$

Denote $\bar{Y} := \frac{1}{n}\sum_{i=1}^n Y_i$ and $\bar{Z} := \frac{1}{n}\sum_{i=1}^n Z_i$. Let $\Phi^{-1}(\cdot)$ be the quantile function of $N(0,1)$, and set $\Phi^{-1}(0.9) = 1.282$, $\Phi^{-1}(0.95) = 1.645$, $\Phi^{-1}(0.975) = 1.96$ and $\Phi^{-1}(0.99) = 2.326$. Suppose that we estimate these two regressions using the least squares method. Which one of the following is the 95% confidence interval of $\alpha_Y + \alpha_Z$ implied by a suitable large-sample method when $n = 100$, $\bar{X} = 0.4$ and $\bar{Y} = 0.6$?

(a) $(0.822, 1.626)$

(b) $(0.716, 1.426)$

(c) $(0.608, 1.392)$

(d) $(0.416, 2.266)$

(e) None of the above choices (a)-(d).

6. Let $\{X_i\}_{i=1}^n$ be a sequence of independent and $\chi^2(1)$-distributed random variables. Denote $\bar{X} := n^{-1}\sum_{i=1}^n X_i$. By Chebyshev's inequality, we have the result:

$$P(\bar{X} \le 2) \ge \beta,$$

for some $\beta \in (0,1)$. Which of the following is right when $n = 50$?

(a) $\beta = 0.95$

(b) $\beta = 0.96$

(c) $\beta = 0.97$

(d) $\beta = 0.98$

見背面

(e) None of the above choices (a)-(d).

7. Let $\{(e_i, X_i')\}_{i=1}^n$ be a sequence of IID random vectors with the distribution:

$$\begin{bmatrix} e_i \\ X_i \end{bmatrix} \sim N\left(\begin{bmatrix} 0 \\ \mu \end{bmatrix}, \begin{bmatrix} 1 & 0_{1\times k} \\ 0_{k\times 1} & \Sigma \end{bmatrix}\right),$$

for some $k := \dim(X_i) > 1$. Consider the following regression:

$$Y_i = X_i'\gamma + e_i,$$

where γ is a vector of regression coefficients, and e_i is an error term. Let R^2 be the "uncentered R^2" of this regression, based on the least squares method, and R_*^2 be the probability limit of R^2 as $n \to \infty$. Which of the following is right?

(a) $R_*^2 = \frac{\gamma'\gamma}{1+2\gamma'\gamma}$

(b) $R_*^2 = \frac{\gamma'\Sigma\gamma}{1+\gamma'\Sigma\gamma}$

(c) $R_*^2 = \frac{\gamma'(\mu\mu'+\Sigma)\gamma}{1+\gamma'\Sigma\gamma}$

(d) $R_*^2 = \frac{\gamma'\Sigma\gamma}{1+\gamma'(\mu\mu'+\Sigma)\gamma}$

(e) None of the above choices (a)-(d).

8. Let $\{(Y_i, Z_i)\}$ be a sequence of IID random vector, in which Y_i is a Bernoulli random variable, and Z_i is a random vector with the distribution $U(\alpha, \beta)$ for some $0 < \alpha < \beta < 1$. Consider the following regression:

$$Y_i = \gamma Z_i + e_i,$$

for some $0 < \gamma < \frac{1}{\beta}$, and e_i is an error term with the property $\mathbb{E}[e_i|Z_i] = 0$. Also, let S be the skewness coefficient of Y_i. Which of the following is right?

(a) $S = \frac{1-3\left(\frac{\gamma(\alpha+\beta)}{2}\right)+2\left(\frac{\gamma(\alpha+\beta)}{2}\right)^2}{\left(1-\frac{\gamma(\alpha+\beta)}{2}\right)\left(\frac{\gamma(\alpha+\beta)}{2}\left(1-\frac{\gamma(\alpha+\beta)}{2}\right)\right)^{1/2}}$

(b) $S = \frac{1-3\left(\frac{\gamma(\alpha+\beta)}{2}\right)+2\left(\frac{\gamma(\alpha+\beta)}{2}\right)^2}{\left(1-\frac{\gamma(\alpha+\beta)}{2}\right)^{1/2}\left(\frac{\gamma(\alpha+\beta)}{2}\left(1-\frac{\gamma(\alpha+\beta)}{2}\right)\right)^{1/2}}$

(c) $S = \frac{1-3\left(\frac{\gamma(\alpha+\beta)}{2}\right)+2\left(\frac{\gamma(\alpha+\beta)}{2}\right)^2}{\left(1+\frac{\gamma(\alpha+\beta)}{2}\right)\left(\frac{\gamma(\alpha+\beta)}{2}\left(1-\frac{\gamma(\alpha+\beta)}{2}\right)\right)^{1/2}}$

接次頁

(d) $S = \frac{1-3\left(\frac{\gamma(\alpha+\beta)}{2}\right)+2\left(\frac{\gamma(\alpha+\beta)}{2}\right)^2}{\left(1+\frac{\gamma(\alpha+\beta)}{2}\right)^{1/2}\left(\frac{\gamma(\alpha+\beta)}{2}\left(1-\frac{\gamma(\alpha+\beta)}{2}\right)\right)^{1/2}}$

(e) None of the above choices (a)-(d).

9. Let $\{X_i\}_{i=1}^n$ be a sequence of independent and $N(0,1)$-distributed random variables. Denote $Y_n := \frac{1}{n_1}\sum_{i=1}^{n_1} X_i - \frac{1}{n-n_1}\sum_{i=n_1+1}^{n} X_i$, for some $n_1 \leq n$. Let $K_n(t)$ be the cumulant generating function of the statistic, with t denoting a real number. Which of the following is right?

(a) $\ln K_n(t) = \ln 4 + 2\ln t + \ln\left(\frac{n}{2n_1(n-n_1)}\right)$

(b) $\ln K_n(t) = 2\ln t + \ln\left(\frac{n}{n_1(n-n_1)}\right)$

(c) $\ln K_n(t) = \ln\frac{1}{2} + 2\ln t + \ln\left(\frac{n}{n_1(n-n_1)}\right)$

(d) $\ln K_n(t) = \ln\frac{1}{4} + \frac{1}{2}\ln t + \ln\left(\frac{n}{2n_1(n-n_1)}\right)$

(e) None of the above choices (a)-(d).

10. Let $\{X_i\}_{i=1}^n$ be a sequence of independent and $N(0,1)$-distributed random variables. Consider the following variable:

$$Y_i := \begin{cases} X_1, & i = 1, \\ X_i + \beta X_{i-1}, & i > 1, \end{cases}$$

for some $\beta \in (0,1)$, and denote the statistic $W_n := \frac{1}{n}\sum_{i=1}^n Y_i$. Which of the following is right?

(a) $\mathrm{var}[W_n] = \frac{1}{n^2}\left(1+(1+\beta)(n-1)\right)^2$

(b) $\mathrm{var}[W_n] = \frac{1}{n^2}\left(1+(1+\beta)^2(n-1)\right)$

(c) $\mathrm{var}[W_n] = \frac{1}{n^2}\left(1+(1+\beta)^2(n-1)+2\beta(n-1)^2\right)$

(d) $\mathrm{var}[W_n] = \frac{1}{n^2}\left(1+(1+\beta)(n-1)+(1+\beta)^2(n-1)^2\right)$

(e) None of the above choices (a)-(d).

見背面

Piske and Usagi made a big fortune by selling Line stickers. They are planning to invest the money in the stock market. They are now studying the property of a certain stock X. They have collected data of monthly returns of X in the past 8 months (r_x) and the corresponding market returns (r_m) as well as the risk-free rates (r_f).

Month$_t$	r_x (%)	r_m (%)	r_f (%)
1	3	5	2
2	5	5	2
3	3	5	2
4	0	-5	2
5	-2	-5	1
6	-6	-5	1
7	10	10	1
8	15	10	1

11. Piske is interested in the systematic risk of stock X and would like to estimate the following regression:

$$r_{xt} - r_{ft} = b_0 + b_1(r_{mt} - r_{ft}) + e_t \tag{1}$$

Assume that $e_t \sim IID(0, \sigma^2)$ and that $E[e_t | r_{mt}, r_{ft}] = 0$. Please help Piske estimate the ordinary least square (OLS) estimates $\widehat{b_0^{OLS}}$ and $\widehat{b_1^{OLS}}$.

a. $\widehat{b_0^{OLS}} = 0.011$, $\widehat{b_1^{OLS}} = 0.927$
b. $\widehat{b_0^{OLS}} = 0.025$, $\widehat{b_1^{OLS}} = 0.927$
c. $\widehat{b_0^{OLS}} = 0.011$, $\widehat{b_1^{OLS}} = 0.955$
d. $\widehat{b_0^{OLS}} = 0.025$, $\widehat{b_1^{OLS}} = 0.955$
e. None of the above choices (a)-(d).

12. Usagi learned from his investment class that there should be no intercept in a Capital Asset Pricing Model. Thus, he plans to perform the following regression using the Ordinary Least Square (OLS) method:

接次頁

$$r_{xt} - r_{ft} = \beta(r_{mt} - r_{ft}) + \epsilon_t \tag{2}$$

Which of the following statement(s) is (are) correct?

I. $E[\widehat{\beta^{OLS}}] = b_1$.

II. $\sum_t \widehat{\epsilon_t} = 0$.

III. $\sum_t (r_{mt} - r_{ft})\widehat{\epsilon_t} = 0$.

IV. Model (2) should yield a lower centered-R^2 than model (1).

a. I and IV.

b. II and III.

c. III and IV.

d. II and IV.

e. II, III, and IV.

13. Piske and Usagi are interested in assessing the goodness-of-fit of the Ordinary Least Squares (OLS) estimations for the two models above. Let R_1^2 represent the centered-R^2 for the first model, and R_2^2 represent the centered-R^2 for the second model. In addition, let $(\widehat{r_{xt} - r_{ft}})_1$ denote the predicted value of the first model, and $(\widehat{r_{xt} - r_{ft}})_2$ denote the predicted value of the second model. Lastly, let ρ_1 denote the sample correlation coefficient between $(\widehat{r_{xt} - r_{ft}})_1$ and $r_{xt} - r_{ft}$, and ρ_2 denote the sample correlation coefficient between$(\widehat{r_{xt} - r_{ft}})_2$ and $r_{xt} - r_{ft}$. Which of the following statements is correct?

a. $\rho_1^2 = \rho_2^2 = R_1^2$

b. $\rho_1^2 = \rho_2^2 = R_2^2$

c. $\rho_1^2 = R_1^2 = R_2^2$

d. $\rho_2^2 = R_1^2 = R_2^2$

e. None of the above choices (a)-(d).

見背面

14. Let's come back to model (1). Assume that $E[e_t|r_{mt}, r_{ft}] = 0$, but the error term (e_t) follows the distribution below:

$$e_t = \rho e_{t-1} + \varepsilon_t, |\rho| < 1$$
$$\varepsilon_t \sim IID(0, \sigma^2)$$

Which of the following statement is correct?

a. $\widehat{b_1^{OLS}}$ is biased.
b. $var(e_t) = \frac{\sigma^2}{1-\rho^2}$
c. $cov(e_t, e_{t-1}) = cov(e_t, e_{t-2}) = 0$
d. $\widehat{b_0^{OLS}}$ and $\widehat{b_1^{OLS}}$ will no longer be BUE but just $BLUE$.
e. None of the above choices (a)-(d).

15. T is a random variable with the following probability density function:

$$f(t) = \beta e^{-\beta t}, \text{for } t > 0$$

where $\beta > 0$. The expected value of T is $E(T) = \frac{1}{\beta}$.
A sample consisting of n independent realizations of T ($\{t_1, t_2, \cdots, t_n\}$) was collected. Which of the following statement is correct?

a. $\widehat{\beta_{MLE}} = \frac{\sum_{i=1}^{n} t_i}{n}$.
b. $\widehat{\beta_{MLE}}$ is an unbiased estimator.
c. $\widehat{\beta_{MLE}}$ is a consistent estimator.
d. $\widehat{\beta_{MLE}}$ is neither unbiased nor consistent.
e. None of the above choices (a)-(d).

16. Lisa and Gaspard are interested in the NTU Master program in Finance and collect the salary information of their recent alumni. To investigate potential gender discrimination in the finance industry in Taiwan, they group the observations by gender and estimate the statistics of their annual salary (in thousand NTD). Denote

接次頁

the population mean salary of men as μ_{men} and women as μ_{women}. Assume that the samples of men and women are distributed independently. Please help Lisa and Gaspard to conduct hypotheses testing using the sample information below.

	sample average salary ($\bar{Y}$)	sample standard deviation (S_Y)	n
Men	\$1,080	\$270	30
Women	\$960	\$360	20

Which of the following statement is correct?

a. To examine potential gender discrimination, we may establish a null hypothesis of $H_A : \mu_{men} \neq \mu_{women}$.

b. The t-statistic for $H_0 : \mu_{men} = \mu_{women}$ is 1.271.

c. The t-statistic for $H_0 : \mu_{men} = \$1,000$ is 2.623.

d. The t-statistic for $H_0 : \mu_{women} = \$1,000$ is -1.497.

e. We can reject the null hypothesis that there is no gender discrimination at 5% significance level.

17. There are machines A, B, and C in a factory all producing screws. Of their production, machines A, B, and C produce 1%, 2% and 4% defective screws respectively. Of the total production in the factory, machine A produces 50%, machine B produces 25%, and machine C produces 25%. If one screw is selected at random from the entire screws produced in a day, which of the following is true?

a. The probability that it is defective is 3%.

b. If the selected product is defective, the conditional probability that it was produced by machine A is $\frac{1}{4}$.

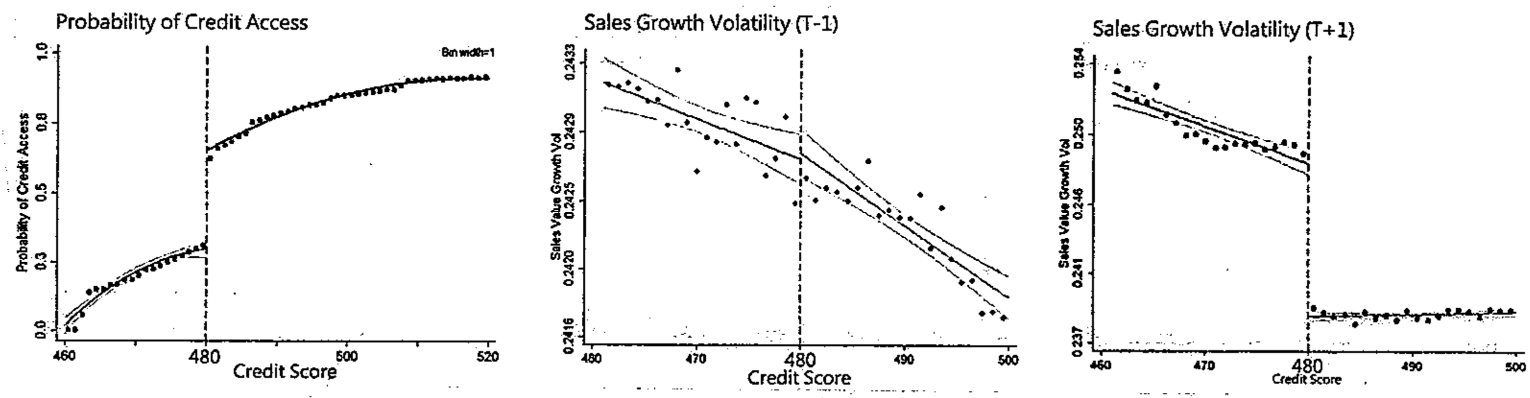
c. If the selected product is defective, the conditional probability that it was produced by machine B is $\frac{2}{4}$.

d. If the selected product is defective, the conditional probability that it was produced by machine C is $\frac{3}{4}$.

e. None of the above choices (a)-(d).

見背面

18. Chen, Huang, Lin, and Shen (2022,*MS*) are interested in the causal relationship between credit supply and growth stability of small enterprises in China. They use the data from Alibaba and provide the following figures. The horizontal axis shows the credit scores of the entrepreneurs. The vertical axes are (from left to right) the probability that the entrepreneurs are granted with credit, the sales growth volatility one year *before* the event year, and the sales growth volatility one year *after* the event year. Which of the following statement is **incorrect**?



a. There is a sudden jump in credit access when one's credit score surpasses 480.

b. There is a sudden drop in the sales growth volatility one year *before* the borrowing when one's credit score surpasses 480.

c. There is a sudden drop in the sales growth volatility one year *after* the borrowing when one's credit score surpasses 480.

d. The findings suggest that the relationship between credit access and future growth stability may be causal.

e. If there is a discontinuity in the sales growth volatility one year *before* the credit access, we cannot make a causal inference between credit access and future growth stability.

19. Consider the following model:

$$Y_t = \gamma Y_{t-1} + u_t$$
$$u_t = \phi u_{t-1} + \epsilon_t$$
$$\epsilon_t \sim IID(0, \sigma^2)$$

接次頁

It is known that $|\gamma| < 1$ and $|\phi| < 1$. We also have a reasonably large T number of observations. Now, you are planning to run an Ordinary Least Square (OLS) regression of $\{Y_t\}$ on its lag term $\{Y_{t-1}\}$:

$$Y_t = a + bY_{t-1} + e_t$$

Which of the following statements is correct?

a. $\widehat{b^{OLS}}$ is an unbiased estimator for γ.

b. $\widehat{b^{OLS}}$ is biased since there is no intercept in the data generating process for Y_t.

c. $\widehat{b^{OLS}}$ is biased because u_t is not stationary.

d. $cov(Y_{t-1}, u_t) = \frac{\phi\sigma^2}{(1-\gamma\phi)(1-\phi^2)}$

e. None of the above choices (a)-(d).

20. Suppose that a data generating process is as the following:

$$Y_i = a + bX_i^* + \epsilon_i$$

However, X_i^* cannot be directly observed. There are two observable proxies for X_i^*:

$$X_{i1} = X_i^* + u_{i1}$$
$$X_{i2} = X_i^* + u_{i2}$$

We know that $\epsilon_i \sim N(0, \sigma_\epsilon^2)$, $X_i^* \sim N(\mu_X, 6)$, $u_{i1} \sim N(0, 1)$, and $u_{i2} \sim N(0, 3)$. u_{i1} and u_{i2} are also uncorrelated to X_i^*, Y_i, ϵ_i, and each other.
We now construct two additional regressors:

$$X_{i3} = \frac{2}{3}X_{i1} + \frac{1}{3}X_{i2}$$
$$X_{i4} = \frac{1}{5}X_{i1} + \frac{4}{5}X_{i2}$$

For $n = 1, 2, 3,$ or 4, denote $\widehat{\beta_n}$ as the OLS coefficient when we regress Y on X_{in}. For a positive b, which of the following has the largest value?

a. $\frac{1}{2}b$;

b. $\widehat{\beta_1}$

c. $\widehat{\beta_2}$

d. $\widehat{\beta_3}$

e. $\widehat{\beta_4}$.

試題隨卷繳回