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.

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	10	10	2
2	0	10	2
3	10	10	3
4	0	10	3
5	-10	-4	3
6	-20	-4	3
7	0	6	2
8	6	6	2

1. 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. $\widehat{b_0^{OLS}} = -0.058$, $\widehat{b_1^{OLS}} = 0.944$

b. $\widehat{b_0^{OLS}} = -0.073$, $\widehat{b_1^{OLS}} = 0.944$

c. $\widehat{b_0^{OLS}} = -0.058$, $\widehat{b_1^{OLS}} = 1.437$

d. $\widehat{b_0^{OLS}} = -0.073$, $\widehat{b_1^{OLS}} = 1.437$

e. $\widehat{b_0^{OLS}} = -0.084$, $\widehat{b_1^{OLS}} = 1.443$

2. 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:

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

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

I. $E[\widehat{\beta}] \neq E[\widehat{b_1}]$.

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

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

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

a. I

b. I and II.

c. I and III.

d. I, II, and III.

e. I, II, and IV.

3. 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?

接次頁

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

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

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

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

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

4. 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. $var(e_t) = \frac{\sigma^2}{1-\rho}$

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).

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

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

where $\lambda > 0$. The expected value of T is $E(T) = \frac{1}{\lambda}$.
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. $\widehat{\lambda_{MLE}} = \frac{\sum_{i=1}^n t_i}{n}$.

b. $\widehat{\lambda_{MLE}}$ is an unbiased estimator.

c. $\widehat{\lambda_{MLE}}$ is a consistent estimator.

d. $\widehat{\lambda_{MLE}}$ is unbiased and consistent.

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

6. 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 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	\$1,200	\$250	20
Women	\$900	\$300	15

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 2.861.

c. The t-statistic for $H_0 : \mu_{men} = \$1,000$ is 3.578.

d. The t-statistic for $H_0 : \mu_{women} = \$1,000$ is -1.789.

e. We cannot reject the null hypothesis that there is no gender discrimination at 5% significance level.

7. There are machines I, II, and III in a certain factory all producing the same product. Of their production, machines I, II, and III produce 1%, 2% and 3% defective products respectively. Of the total production in the factory, machine I produces 30%, machine II produces 30%, and machine III produces 40%. If one product is selected at random from the total product produced in a day, which of the following is true?

a. The probability that it is defective is 2%

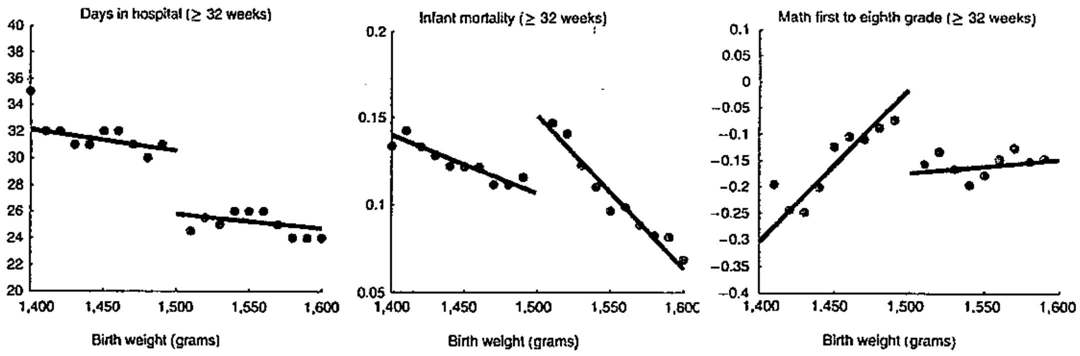
b. If the selected product is defective, the conditional probability that it was produced by machine I is $\frac{1}{7}$.

c. If the selected product is defective, the conditional probability that it was produced by machine II is $\frac{3}{7}$.

d. If the selected product is defective, the conditional probability that it was produced by machine III is $\frac{5}{7}$.

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

8. Bharadwaj, Løken, and Neilson (2013,*AER*) are interested in the effects of early life health interventions and provide the following figures. It is widely known that the birth weight of an infant is negatively associated with the mortality rate and positively associated with future health condition. The researchers found a natural experiment in both Chile and Norway that the governments require hospitals to provide additional care for infants born less than 1,500 grams. Below are the days stayed in the hospital, mortality rates, and the math performance of the infants when they turn to first to eighth grade across their birth weights. Which of the following statement is **incorrect**?



a. Infants born right below 1,500 grams stay longer in the hospital than infants born right above 1,500 grams.

b. The negative slopes of infant mortality rates and birth weights below and above the 1,500-gram birth weight are what the researchers are interested in.

c. Infants born right below 1,500 grams tend to perform better in math than infants born right above 1,500 grams when they grow up.

d. The findings suggest that the relationship between early life health care and future intellectual development may be causal.

e. Except for the discontinuity at exactly 1,500 grams, there seems to be a positive relationship between the birth weights and future math performance.

9. Consider the following model:

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

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. $cov(Y_{t-1}, u_t) = \frac{\phi\sigma^2}{(1-\gamma\phi)(1-\phi)}$
e. None of the above choices (a)-(d).

10. 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 $\epsilon_i \sim N(0, \sigma_\epsilon^2)$, $X_i^* \sim N(\mu_X, 4)$, $u_{i1} \sim N(0, 1)$, and $u_{i2} \sim N(0, 2)$. u_{i1} and u_{i2} are also uncorrelated to X_i^*, Y_i, ϵ_i, and each other.

We now construct two additional regressors:

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

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

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

b. $\widehat{\beta_1}$

c. $\widehat{\beta_2}$

d. $\widehat{\beta_3}$

e. $\widehat{\beta_4}$.

11. Let $\{Y_i\}_{i=1}^n$ be a sequence of independent and $N(0,1)$-distributed random variables. Denote $\hat{\mu}_2 := n^{-1}\sum_{i=1}^n Y_i^2$. By Chebyshev's inequality, we obtain the result:

$$P(0.5 \leq \hat{\mu}_2 \leq 1.5) \geq \alpha,$$

for some $\alpha \in (0,1)$. Which of the following is right when $n = 100$?

(a) $\alpha = 0.88$

(b) $\alpha = 0.90$

(c) $\alpha = 0.92$

(d) $\alpha = 0.94$

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

12. Let $\{Y_i\}_{i=1}^n$ be a sequence of independent and $\chi^2(1)$-distributed random variables. Consider a linear regression:

$$Y_i = \beta + e_i,$$

where β is a parameter, and e_i is the error term with $\mathbb{E}[e_i] = 0$. Let $\hat{\beta}$ be the least squares estimator of β. Which of the following is right?

(a) $\mathbb{E}[\hat{\beta}] = 1$ and $\text{var}[\hat{\beta}] = \frac{2}{n}$

(b) $\mathbb{E}[\hat{\beta}] = \frac{1}{n}$ and $\text{var}[\hat{\beta}] = \frac{3}{n}$

(c) $\mathbb{E}[\hat{\beta}] = 1$ and $\text{var}[\hat{\beta}] = \frac{3}{n}$

(d) $\mathbb{E}[\hat{\beta}] = \frac{1}{n}$ and $\text{var}[\hat{\beta}] = \frac{4}{n}$

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

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

$$\begin{bmatrix} Y_{1i} \\ Y_{2i} \end{bmatrix} \sim N\left(\begin{bmatrix} 72 \\ 28 \end{bmatrix}, \begin{bmatrix} 1 & 0.25 \\ 0.25 & 0.5 \end{bmatrix}\right).$$

Denote $\bar{Y}_1 = n^{-1}\sum_{i=1}^{n} Y_{1i}$ and $\bar{Y}_2 = n^{-1}\sum_{i=1}^{n} Y_{2i}$. Which of the following is right when $n = 50$?

(a) $\text{var}[\bar{Y}_1 - \bar{Y}_2] = 0.015$

(b) $\text{var}[\bar{Y}_1 - \bar{Y}_2] = 0.020$

(c) $\text{var}[\bar{Y}_1 - \bar{Y}_2] = 0.025$

(d) $\text{var}[\bar{Y}_1 - \bar{Y}_2] = 0.030$

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

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

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

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

(c) $\ln M_n(t) = t^2\left(\frac{n}{n_1(n-n_1)}\right)$

(d) $\ln M_n(t) = 2t^2\left(\frac{n}{n_1(n-n_1)}\right)$

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

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

$$X_i := \begin{cases} Y_1, & i = 1, \\ Y_i - \alpha Y_{i-1}, & i > 1, \end{cases}$$

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

(a) $\mathbb{E}[Z_n^2] = \frac{1}{n^2}\left(1 + (1-\alpha)^2(n-1)\right)$

(b) $\mathbb{E}[Z_n^2] = \frac{1}{n^2}\left(1 + 2\alpha^2 + (1-\alpha)^2(n-1)\right)$

(c) $\mathbb{E}[Z_n^2] = \frac{1}{n^2}\left(1 + (1-\alpha)^2(n-1) + 2\alpha(n-1)^2\right)$

(d) $\mathbb{E}[Z_n^2] = \frac{1}{n^2}\left(1 + 2(1-\alpha)(n-1) + (1-\alpha)^2(n-1)^2\right)$

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

16. Let $\{Y_i\}_{i=1}^{n}$ be a sequence of independent and $U(0,1)$-distributed random variables with $\sigma^2 := \text{var}[Y_i]$, $\bar{Y} := n^{-1}\sum_{i=1}^{n} Y_i$ and $\hat{\sigma}^2 := n^{-1}\sum_{i=1}^{n}(Y_i - \bar{Y})^2$. By a suitable large-sample method, it can be shown that $\sqrt{n}(\hat{\sigma}^2 - \sigma^2)$ has the limiting distribution $N(0,v)$, as $n \to \infty$, for some $v > 0$. Which of the following is right?

(a) $v = \frac{1}{360}$

(b) $v = \frac{1}{280}$

(c) $v = \frac{1}{240}$

(d) $v = \frac{1}{180}$

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

17. Let $\{(e_i, X_i')\}$ be a sequence of independent and $N(0, I_{k+1})$-distributed random vectors for some $k := \dim(X_i) > 1$. Assume that Y_i is a random variable generated by the formula:

$$Y_i = X_i'X_i + e_i.$$

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

(a) $\text{var}[h^*(X_i)] = k$

(b) $\text{var}[h^*(X_i)] = 3k$

(c) $\text{var}[h^*(X_i)] = 6k$

(d) $\text{var}[h^*(X_i)] = 9k$

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

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

$$Y_i = \beta X_i + e_i,$$

for some $0 < \beta < \frac{1}{b}$, and e_i is an error term with the property $\mathbb{E}[e_i|X_i] = 0$. Which of the following is right?

(a) $\text{var}[Y_i] = \frac{\beta(a+b)}{2}\left(1 + \frac{\beta(a+b)}{2}\right)$
(b) $\text{var}[Y_i] = \frac{\beta(a+b)}{2}\left(1 - \frac{\beta(a+b)}{2}\right)$
(c) $\text{var}[Y_i] = \frac{\beta(a+b)}{2}\left(1 + \frac{\beta(a+b)}{2}\right)^2$
(d) $\text{var}[Y_i] = \frac{\beta(a+b)}{2}\left(1 - \frac{\beta(a+b)}{2}\right)^2$
(e) None of the above choices (a)-(d).

19. Let $\{(e_i, X_i')\}_{i=1}^n$ be a sequence of independent and $N(0, I_{k+1})$-distributed random vectors for some $k := \dim(X_i) > 1$. Consider the following regression:

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

where β is a vector of regression coefficients, and e_i is an error term. Let R^2 be the "centered 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) $R_*^2 = \frac{\beta'\beta}{1+2\beta'\beta}$
(b) $R_*^2 = \frac{\beta'\beta}{1+(\beta'\beta)^2}$
(c) $R_*^2 = \frac{2\beta'\beta}{1+2\beta'\beta}$
(d) $R_*^2 = \frac{\beta'\beta}{1+2(\beta'\beta)^2}$
(e) None of the above choices (a)-(d).

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

$$X_i = \alpha_X + e_i$$

接次頁

and

$$Y_i = \alpha_Y + u_i,$$

where α_X and α_Y are parameters, and

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

Denote $\bar{X} := \frac{1}{n}\sum_{i=1}^{n} X_i$ and $\bar{Y} := \frac{1}{n}\sum_{i=1}^{n} Y_i$. 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 $\alpha_X - \alpha_Y$ implied by a suitable large-sample method when $n = 100$, $\bar{X} = 0.56$ and $\bar{Y} = 0.53$?

(a) $(-0.166, 0.226)$

(b) $(-0.188, 0.244)$

(c) $(-0.206, 0.262)$

(d) $(-0.264, 0.288)$

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

試題隨卷繳回