※ 注意：請用 2B 鉛筆作答於答案卡，並先詳閱答案卡上之「畫記說明」。

Part I. 統計與計量 (50%, 選擇題，答案可能不只一個選項，每題 5 分)

1. Given the function:
$$h(x) := \exp\left(-\frac{x^2}{2}\right),$$
for $x \in \mathbb{R}$, we define
$$\kappa_1 := \int_{-\infty}^{\infty} (x + x^2 + x^3 + x^4) h(x) \mathrm{d}x$$
and
$$\kappa_2 := \int_{-\infty}^{\infty} (\sin(x) + \cos(x) + \exp(2x)) h(x) \mathrm{d}x.$$
Which of the following is right?
   a. $\kappa_1 = 6\sqrt{2\pi}$.
   b. $\kappa_1 = 4\sqrt{2\pi}$.
   c. $\kappa_2 = (\exp(2) + \exp(-1/2))\sqrt{2\pi}$.
   d. $\kappa_2 = 2\exp(2)\sqrt{2\pi}$.
   e. None of the above choices (a)-(d).

2. Let $\{Y_i\}_{i=1}^n$ be a sequence of independently and $N(0,1)$-distributed random variables. Suppose that we utilize Chebyshev's inequality to establish the following result:
$$P(B) \geq \gamma,$$
where
$$B := \left\{ \left| \frac{1}{n} \sum_{i=1}^n Y_i^3 \right| \leq \sqrt{3} \right\}.$$
Which of the following is right?
   a. $\gamma \to \frac{1}{2}$, as $n \to \infty$.
   b. $\gamma = \frac{1}{2} - \frac{2}{n}$.
   c. $\gamma = \frac{1}{2} - \frac{3}{n}$.
   d. $\gamma = 1 - \frac{5}{n}$.
   e. None of the above choices (a)-(d).

3. Let X and Y be two continuous random variables with finite variances. Which of the following is right?
   a. $cov(X, Y) = E[XE[Y|X]] - E[E[X|Y]]E[E[Y|X]]$.
   b. $var[E[Y|X]] \leq var[Y]$.
   c. $E[|X|Y]^2 \leq var(|X|)$ if Y is $N(0,1)$-distributed.

見背面

d. Suppose that X and Y has the joint density function:

$$f_{X,Y}(x,y) = \frac{1}{2\pi\sqrt{1-\rho^2}} \exp\left(-\frac{x^2 - 2\rho xy + y^2}{2(1-\rho^2)}\right),$$

for $(x,y) \in \mathbb{R}^2$ and for some $\rho \in [-1,1]$. Then, $E[X^2Y^2] = 1$ and $E[X^4Y^4] = 9$ when $\rho = 0$.

e. All of the above choices (a)-(d).

4. Assume that $Y = X + X^2 + W + W^2$, where X and W are two independent $U(0,1)$-distributed random variables. Which of the following is right?

   a. $E[Y] = \frac{4}{3}$.
   b. $E[(X+X^2)^2] = \frac{31}{30}$.
   c. $E[(X+X^2)(W+W^2)] = \frac{25}{36}$.
   d. $E[Y^2] = \frac{310}{89}$.
   e. None of the above choices (a)-(d).

5. Suppose that $\{Y_i\}_{i=1}^n$ is a sequence of independently and identically distributed random variables with the probability density function:

$$f(y|\nu) = \frac{\Gamma(\frac{\nu+1}{2})}{\Gamma(\frac{\nu}{2})\sqrt{\nu\pi}} \left(1 + \frac{y^2}{\nu}\right)^{-(\frac{1+\nu}{2})},$$

for $y \in \mathbb{R}$ and for some parameter $\nu > 2$. Also, let $\hat{\nu}$ be the maximum likelihood estimator for ν which is solved by maximizing the log-likelihood function:

$$L_n(\nu) := \frac{1}{n}\sum_{i=1}^{n} \log f(Y_i|\nu).$$

Denote $\psi(x) := \frac{d}{dx}\log\Gamma(x)$. Which of the following is right?

   a. $\frac{1}{n}\sum_{i=1}^n \log(1+\frac{Y_i^2}{\hat{\nu}}) - (\frac{1+\hat{\nu}}{\hat{\nu}})\frac{1}{n}\sum_{i=1}^n(\frac{Y_i^2/\hat{\nu}}{1+Y_i^2/\hat{\nu}}) = \psi(\frac{\hat{\nu}+1}{2}) - \psi(\frac{\hat{\nu}}{2}) - \frac{1}{\hat{\nu}}$.
   b. The estimator $\hat{\nu}$ is consistent for ν. In addition, the estimator $\tilde{\nu}$, which is solved from the restriction: $\frac{\tilde{\nu}}{\tilde{\nu}-2} = \frac{1}{n}\sum_{i=1}^n(Y_i - \bar{Y})^2$ with $\bar{Y} := \frac{1}{n}\sum_{i=1}^n Y_i$, is also consistent for ν.
   c. The estimator $\hat{\nu}$ is consistent for ν. However, the estimator $\tilde{\nu}$, which is solved from the restriction: $\frac{\tilde{\nu}}{\tilde{\nu}-2} = \frac{1}{n}\sum_{i=1}^n(Y_i - \bar{Y})^2$ with $\bar{Y} := \frac{1}{n}\sum_{i=1}^n Y_i$, is inconsistent for ν.
   d. $\psi(\frac{\nu+1}{2}) - \psi(\frac{\nu}{2}) - \frac{1}{\nu} = E\left[\log(1+\frac{Y_i^2}{\nu})\right] - (\frac{1+\nu}{\nu})E\left[\frac{Y_i^2/\nu}{1+Y_i^2/\nu}\right]$
   e. None of the above choices (a)-(d).

接次頁

6. Let $\{Y_i\}_{i=1}^n$ be a sequence of independently and $N(\mu, \sigma^2)$-distributed random variables. Denote $\bar{Y} := \frac{1}{n}\sum_{i=1}^n Y_i$ and $\hat{\sigma}^2 := \frac{1}{n}\sum_{i=1}^n (Y_i - \bar{Y})^2$. Also, let $\Phi(\cdot)$ be the distribution function of $N(0,1)$, and $\Phi^{-1}(\cdot)$ be the quantile function of $N(0,1)$. Suppose that we examine the null hypothesis: $\mu = 0$ using the t-statistic:

$$T := n^{1/2}\bar{Y}/\hat{\sigma}.$$

Let c be a positive constant. Which of the following is right?

a. $\lim_{n\to\infty} P(|T| > c) = 2 - 2\Phi(c)$ under the null hypothesis.
b. $\Phi(T)$ is $U(0,1)$-distributed, as $n \to \infty$, under the null hypothesis.
c. If $\lim_{T\to\infty} P(T > c) = \alpha$ holds under the null hypothesis for some $\alpha \in (0,1)$, then $c = \Phi^{-1}(1-\alpha)$.
d. $\lim_{T\to\infty} P(T > c) = 1 - \Phi\left(c - \frac{\delta}{\sigma}\right)$ holds under the hypothesis: $\mu = \delta/\sqrt{n}$ for some fixed $\delta \neq 0$.
e. None of the above choices (a)-(d).

7. Let Z be an $n \times 4$ matrix of Gaussian random variables with $n > 4$, and I_n be an $n \times n$ identity matrix. Assume that $Z'Z$ is positive definite. Define the matrix:

$$W := I_n - Z(Z'Z)^{-1}Z'.$$

Which of the following is right?

a. W is singular.
b. $W = WWW$.
c. $\text{rank}(WW) = 4$.
d. $\text{trace}(WWW) = n - 4$.
e. None of the above choices (a)-(d).

8. Let $\{(Y_i, X_{1,i}, \ldots, X_{K,i})\}_{i=1}^n$ be a sequence of independently and identically distributed random vectors with a finite covariance matrix. Consider the following linear regression:

$$Y_i = \sum_{j=1}^K \beta_j X_{j,i} + e_i,$$

where β_j is the regression coefficient of $X_{j,i}$, for $j = 1, 2, \ldots, K$, and e_i is the error term. Let $\hat{\beta} := (\hat{\beta}_1, \ldots, \hat{\beta}_K)'$ be the least squares estimator for $\beta := (\beta_1, \ldots, \beta_K)'$, and $\hat{e} := Y_i - \sum_{j=1}^K \hat{\beta}_j X_{j,i}$ be the least sequares residual. Which of the following is right?

a. $\frac{1}{n}\sum_{i=1}^n \hat{e}_i = 0$.

b. Define $R^2 := 1 - \frac{ESS}{TSS}$, where $ESS := \sum_{i=1}^{n} \hat{e}_i^2$ and $TSS := \sum_{i=1}^{n}(Y_i^2 - \frac{1}{n}\sum_{i=1}^{n} Y_i)^2$. Then, R^2 must be positive and not greater than one.

c. $\sum_{i=1}^{n} X_{j,i}\hat{e}_i = \sum_{i=1}^{n} X_{\ell,i}\hat{e}_i$ for all $j, \ell = 1, 2, \ldots, K$.

d. $\left(\sum_{i=1}^{n} X_{j,i}\hat{e}_i\right)^2 = 1 - \exp\left(\sum_{i=1}^{n} X_{\ell,i}\hat{e}_i\right)$ for all $j, \ell = 1, 2, \ldots, K$.

e. All of the above choices (a)-(d).

9. Let $\{(Y_i, X_i, Z_i)\}_{i=1}^{n}$ be a sequence of independently and identically distributed random vectors, where X_i has the distribution $\chi^2(k_x)$, Z_i is a $\chi^2(k_z)$-distributed random variable which is independent of X_i, and $Y_i = X_i + Z_i$ for all i's. Consider a simple regression:

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

where β is the regression coefficient, and e_i is the regression error. Let $\hat{\beta}$ be the least squares estimator for β. Which of the following is right?

a. $\hat{\beta}$ converges in probability to 1, as $n \to \infty$.

b. $\hat{\beta}$ converges in probability to $1 + \frac{k_z}{3+k_x}$, as $n \to \infty$.

c. $\hat{\beta}$ converges in probability to $1 + \frac{k_z}{2+k_x}$, as $n \to \infty$.

d. $\hat{\beta}$ converges in probability to $1 + \frac{k_z}{3+2k_x}$, as $n \to \infty$.

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

10. Let $\{Y_i\}_{i=1}^{n}$ be a sequence of independently and $N(0, 1)$-distributed random variables. Denote $\bar{Y} := \frac{1}{n}\sum_{i=1}^{n} Y_i$ and $\hat{\sigma}^2 := \frac{1}{n}\sum_{i=1}^{n}(Y_i - \bar{Y})^2$. Which of the following is right?

a. The limiting distribution of $\sqrt{n}(\hat{\sigma}^2 - 1)$ is $N(0, 1)$, as $n \to \infty$.

b. The limiting distribution of $\sqrt{n}(\hat{\sigma}^2 - 1)$ is $N(0, 2)$, as $n \to \infty$.

c. The limiting distribution of $\sqrt{n}(\hat{\sigma} - 1)$ is $N(0, 1)$, as $n \to \infty$.

d. The limiting distribution of $\sqrt{n}(\hat{\sigma} - 1)$ is $N\left(0, \frac{1}{2}\right)$, as $n \to \infty$.

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

接次頁

Part II. 統計個案分析 (50%, 選擇題，答案可能不只一個選項，每題 5 分)

Ash Ketchum, the owner of Pikachu, surveyed attack powers of all 23 Pokémon monsters after he won the 2022 Pokémon World Championships. He studied all monsters for five years and regress attack powers of Pokémon on their weights. A simple linear regression analysis was specified as

$y_{it}=a_0+ a_1 x_{it} + e_{it},$

where i and t indicate subscripts for monster and year, respectively. y is the attack power, and x is the weight. Regression was performed by ordinary least squares (OLS) method with results below.

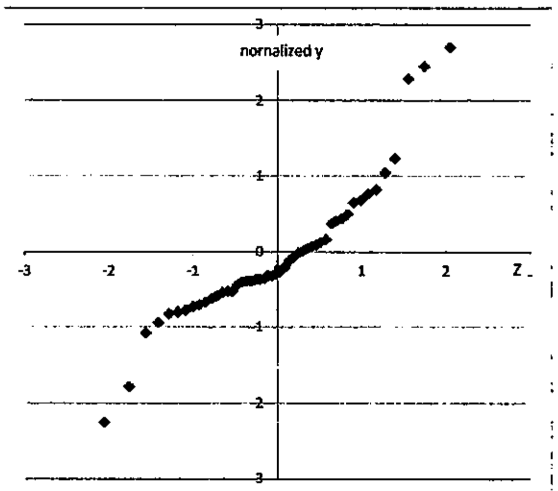
Summary	
R-sq	0.111
Adj. R-sq	[ ]
SE	56.375
N	115.000

ANOVA	DF	SS	MS	F	P-value
Regression	1.000	44768.233	44768.233	14.086	0.000
Residual	113.000	359132.495	3178.164		
Total	114.000	403900.728			

Regression estimates	Coefficient	SE	t-value	P-value
Intercept	83.305	17.278	4.821	0.000
Weight	1.994	0.531	3.753	0.000

11. The data form of Ash's study could be known as
    a. Cross-sectional
    b. Time-series
    c. Panel
    d. Longitudinal
    e. Categorical

12. Ash wants to know the distribution of the y variable, i.e., the attack power, by QQ-plot. To plot the QQ-plot, he obtained a figure in the right. Which one(s) of following statement should be true?
    a. Data do not follow a normality
    b. Data are right skewed
    c. Data are left skewed
    d. Data are fat-tailed
    e. Data are generally a normal distribution



13. Based on information from ANOVA, the unconditional variance of y, which is equal to $\Sigma(y_{it} - \bar{y})^2/(n-1)$, is about
    a. 56.375
    b. 59.523
    c. 3178.140
    d. 3542.988
    e. 4156.213

見背面

14. Ash does not interpret the volatility of y by standard error but by variance. Which of following one(s) is the problem of the standard error (SE) under assumptions of *i.i.d.*?
   a. SE is a non-linear function of y
   b. SE is a biased estimator
   c. SE is an inconsistent estimator
   d. SE has the same unit of y
   e. SE is more volatile than the variance

15. In the ANOVA table, SS indicates
   a. Sum of Squares
   b. Single Source
   c. Speed Symbol
   d. Standard Security
   e. Sum of Submissions

16. The adj. R-sq is not presented in the tables. The adj. R-sq is equal to.
   a. 0.103
   b. 0.111
   c. 0.897
   d. 0.889
   e. 0.998

17. What is SS of the independent variable x?
   a. 106.112
   b. 9987.121
   c. 11261.963
   d. 44768.233
   e. 403900.728

18. Given the data form (preferring to the answer in question 1 of this case study),which of following statement(s) is(are) true?
   a. The time series variations in the data generally violate the independence assumption
   b. The cross-sectional variations in the data generally violate the normality assumption in a large sample
   c. We can further a model autoregressive regression model for the residual e to fit the independent assumption.
   d. The cross-sectional variations in the data generally can be modeled by a moving average model for the residual e.
   e. Regression model using panel form of data can be estimated by maximum likelihood method.

19. Ash considered an alternative regression model that include a new variable and year fixed effect:

接次頁

$y_{it}=a_0+ a_1 x_{it} +a_2 w_t +b_1 v2018_t +b_2 v2019_t +b_3 v2020_t +b_4 v2021_t +b_5 v2022_t + e_{it}$.

v_t indicates year fixed effect. New variable, w_t, is the time trend, which is equal to 1, 2, 3, 4 and 5 for year 2018, 2019, 2020, 2021 and 2022, respectively. *v2018, v2019,...* and *v2022* are dummies for year fixed effects. Which one(s) of following statements is(are) true?

a. Coefficient a_2 is generally positive.
b. Intercept a_0 will become smaller.
c. Results cannot be estimated
d. A linear combination of w_t, *v2018, v2019,...* and *v2022* could be a constant.
e. There is a perfect collinearity issue.

20. Ash also plans to estimate following regression model:

$x_{it}=d_0+ d_1 y_{it} + \varepsilon_{it}$,

where i and t indicate subscripts for monster and year, respectively. y is the attack power, and x is the weight. Regression was performed by ordinary least squares (OLS) method. Which one(s) of following statements is(are) true?

a. $d_1 < 0.01$
b. $d_1 = 0.056$
c. $d_1 = 0.502$
d. $d_1 > 1$
e. R-sq of regressing x on y is 0.111

試題隨卷繳回