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. The number of red chips and white chips in an urn is not known, but it is known that the proportion, p, of reds is either 1/5, 1/3, 1/2, or 3/4. A sample of size 5, drawn with replacement, yields the sequence red, white, red, white, white. The MLE for p is:

   a. 1/5;
   b. 2/5;
   c. 1/3;
   d. 1/2;
   e. 3/4.

2. Let $\{X_i\}_{i=1}^{10}$ be an IID sequence of random variables drawn from a normal distribution $N(\mu, \sigma^2)$. Define $\bar{X} := \frac{1}{9}\sum_{i=1}^{9} X_i$. Suppose that Z is the standard normal distribution ($Z \sim N(0,1)$). What is the probability that $X_{10} \in [\bar{X} - 2\sigma, \bar{X} + 2\sigma]$?

   a. $Pr(X_{10} \in [\bar{X} - 2\sigma, \bar{X} + 2\sigma]) = Pr(Z \in [-\sqrt{3}, \sqrt{3}])$;
   b. $Pr(X_{10} \in [\bar{X} - 2\sigma, \bar{X} + 2\sigma]) = Pr(Z \in [\frac{-3\sqrt{10}}{5}, \frac{3\sqrt{10}}{5}])$;
   c. $Pr(X_{10} \in [\bar{X} - 2\sigma, \bar{X} + 2\sigma]) = Pr(Z \in [-2, 2])$;
   d. $Pr(X_{10} \in [\bar{X} - 2\sigma, \bar{X} + 2\sigma]) = Pr(Z \in [-1.96, 1.96])$;
   e. None of the above choices (a)-(d).

3. Suppose we want to examine a policy effect on manufacturing output growth in 1960s. There is a policy that affected some cities starting 1965 in the US. Assume that the city-year level manufacturing output growth, Y_{it}, follows the model below:

   $Y_{it} = \lambda_i + \psi_t + \delta D_{it} + \epsilon_{it}$

   where λ_i is the city fixed effect; ψ_t is the year fixed effect, $t \in \{1960, 1961, ..., 1969, 1970\}$;

   $D_{it} = \begin{cases} 1 \ , & if\ t \geq 1965\ \&\ if\ city\ i\ is\ hit\ by\ the\ policy. \\ 0 \ , & otherwise. \end{cases}$; $\epsilon_{it} \sim N(0, \sigma_\epsilon^2)$.

   Which of the following is true?

   a. The expectation on the output growth of a city j that is not affected by the policy in year $\tau > 1965$ is $\lambda_j + \psi_\tau + \delta$;
   b. δ captures the average difference in the absolute value of output growth in the entire sample period between cities affected by the policy and cities not affected by the policy;

見背面

c. Some cities may intrinsically enjoy different output growth. This difference will be reflected by ψ_t;

d. Before 1965, cities with and without the policy should have the same expectation on the output growth: $\lambda_i + \psi_t$;

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

4. A random variable X is normally distributed, $X \sim N(\mu, \sigma^2)$. A sample of size $n = 9$ is drawn and yields: $\sum_{i=1}^{9} X_i = 90$, and $\sum_{i=1}^{9}(X_i - \bar{X})^2 = 72$, where $\bar{X} = \frac{1}{9}\sum_{i=1}^{9} X_i$. We would like to test the following hypothesis: $H_0 : \mu = 4$ and $H_1 : \mu \neq 4$. We are given that $t_{8,0.99} = 2.896$, $t_{8,0.975} = 2.306$, $t_{8,0.95} = 1.860$, and $t_{8,0.90} = 1.397$, where $t_{d,\alpha}$ is the t-statistic with degree of freedom d and cumulative probability of α. Which of the following is true?

a. We cannot reject the null at 90% confidence level;

b. We can reject the null at 90% confidence level but not at 95% confidence level;

c. We can reject the null at 95% confidence level but not at 97.5% confidence level;

d. We can reject the null at 97.5% confidence level but not at 99% confidence level;

e. We can reject the null at 99% confidence level.

5. Consider a normal distribution of the form N$(\mu, 6)$. You want to test the hypothesis that $H_0 : \mu = 3$ against the alternative $H_1 : \mu > 3$. A random sample $X_1, X_2, ..., X_n$ was obtained.

Suppose you would like to test the above one-sided hypothesis using a 5% level of significance. How would you test it?

a. Set up a Z test where $z = \frac{\bar{X}-3}{\sqrt{6n}}$ and reject the null if $z > 1.96$;

b. Set up a Z test where $z = \frac{\bar{X}-3}{\sqrt{6n}}$ and reject the null if $z > 1.645$;

c. Set up a Z test where $z = \frac{\bar{X}-3}{\sqrt{6/n}}$ and reject the null if $z > 1.96$;

d. Set up a Z test where $z = \frac{\bar{X}-3}{\sqrt{6/n}}$ and reject the null if $z > 1.645$;

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

6. Follow the previous question, denote $\Phi(.)$ as the CDF of normal distribution. Provide the power of this test at θ_1 where $\theta_1 > 3$.

a. $1 - \Phi(1.645 + \frac{\theta_1-3}{\sqrt{6/n}})$;

接次頁

b. $1-\Phi(1.645+\frac{3-\theta_1}{\sqrt{6/n}})$;

c. $1-\Phi(1.96+\frac{\theta_1-3}{\sqrt{6/n}})$;

d. $\Phi(1.96+\frac{\theta_1-3}{\sqrt{6n}})$;

e. $\Phi(1.645+\frac{3-\theta_1}{\sqrt{6/n}})$.

7. Bob is testing the following empirical model using a sample of 1,500 observations:
$Y_i = \alpha + \beta_1 X_{1,i} + \beta_2 X_{2,i} + \beta_3 X_{3,i} + \beta_4 X_{4,i} + \epsilon_i$.
He found that the F-statistic of the joint test for the entire model is 300. He now considers expressing the model's strength using (unadjusted) R-square. Please calculate the R-square value for him.

a. 0.376;

b. 0.445;

c. 0.501;

d. 0.668;

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

8. In the two-variable model:
$Y_i = \beta_1 X_{1i} + \beta_2 X_{2i} + \epsilon_i, \quad i = 1, 2, 3, ..., 11$
Suppose that $X_1'X_1 = 2$, $X_2'X_2 = 2$, $X_1'X_2 = 1$, $X_1'Y = 1$, $X_2'Y = 1$, and $Y'Y = 4/3$, where X_1, X_2, and Y are the column vectors with typical elements X_{1i}, X_{2i}, and Y_i respectively. Furthermore, X_1', X_2', and Y' are the transpose of X_1, X_2, and Y respectively. Assume $\epsilon_i \sim IID\ N(0, \sigma_\epsilon^2)$.
You are considering of testing the following hypotheses: $\beta_1 = 0$. Please calculate the test statistic: $\frac{\hat{\beta}_1 - 0}{SE(\hat{\beta}_1)}$.
Note: To calculate the standard errors, please estimate the variance-covariance matrix of $\hat{\beta}$ as $s^2(X'X)^{-1}$ where s^2 is the residual sum of squares divided by the degree of freedom and $X = \begin{bmatrix} X_1 & X_2 \end{bmatrix}$.

a. 1.0;

b. 1.5;

c. 2.0;

d. 2.5;

e. 3.0.

9. Consider 4 events, A, B, C, D. $Pr(A) = Pr(B) = Pr(C) = Pr(D) = 0.4$. $Pr(C \cap D) = Pr(C \cap A) = Pr(B \cap D) = 0$. $Pr(A \cap B) = 0.1$. Which of the following is possible?

見背面

a. $Pr(B \cap C) = 0.2$;
b. A and D are independent;
c. $Pr((B \cap C) \cup (A \cap D)) = 0.45$;
d. $Pr(D \cup A^c) = 0.55$;
e. None of the above choices (a)-(d).

10. X and ϵ are independent random variables. $X \sim N(\mu_X, \sigma_X^2)$ and $\epsilon \sim N(\mu_\epsilon, \sigma_\epsilon^2)$. a and b are real numbers. Y is given by: $Y = a + bX + \epsilon$. Which of the follow is the correlation coefficient ρ between variable X and Y?

a. $\frac{b}{\sqrt{b^2+\sigma_X^2\sigma_\epsilon^2}}$;
b. $\frac{b}{\sqrt{b^2\sigma_X^2+\sigma_\epsilon^2}}$;
c. $\frac{b}{\sqrt{(b^2+(\sigma_X^2/\sigma_\epsilon^2)}}$;
d. $\frac{b}{\sqrt{(b^2+(\sigma_\epsilon^2/\sigma_X^2)}}$;
e. None of the above choices (a)-(d).

11. Let X be a $N(0,1)$-distributed random variable, and Y be a $\chi^2(n)$-distributed random variable for some positive integer n. Suppose that X and Y are independent. Please calculate $E(X^2Y)$ and $E(X^4Y^2)$.

a. $E(X^2Y) = n$ and $E(X^4Y^2) = 2n + 7n^2$
b. $E(X^2Y) = n$ and $E(X^4Y^2) = 4n + 5n^2$
c. $E(X^2Y) = n$ and $E(X^4Y^2) = 6n + 3n^2$
d. $E(X^2Y) = n$ and $E(X^4Y^2) = 8n + n^2$
e. None of the above choices (a)-(d).

12. Let $\{(Y_i, X_i')\}_{i=1}^n$ be a sequence of IID random vectors, where $X_i := (X_{i,1}, X_{i,2}, \ldots, X_{i,k})'$. Suppose that we have a multiple linear regression:

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

for $i = 1, 2, \ldots, n$. Denote $e := (e_1, e_2, \ldots, e_n)'$. Assume that $e|(X_1, \ldots, X_n) \sim N(0, \sigma^2 I_n)$, where I_n is the $n \times n$ identity matrix. Let $f(\cdot|X_i, \beta, \sigma^2)$ be the conditional density function of $Y_i|X_i$ implied by this regression, with $\beta := (\beta_1, \beta_2, \ldots, \beta_k)'$, and $\hat{\beta}_{ML}$ and $\hat{\sigma}_{ML}^2$ be the maximum likelihood estimators for β and σ^2, respectively, that are defined by maximizing the objective function:

$$\frac{1}{n} \sum_{i=1}^{n} \ln f(Y_i|X_i, \beta, \sigma^2).$$

接次頁

Which of the followings is right?

a. $E(\hat{\beta}_{ML}) = \beta$ and $E(\hat{\sigma}^2_{ML}) = \sigma^2$
b. $E(\hat{\beta}_{ML}) = \beta$ and $E(\hat{\sigma}^2_{ML}) = \frac{n-k}{n}\sigma^2$
c. $E(\hat{\beta}_{ML}) = \beta$ and $E(\hat{\sigma}^2_{ML}) = \frac{n+k}{n}\sigma^2$
d. $E(\hat{\beta}_{ML}) = \beta$ and $E(\hat{\sigma}^2_{ML}) = \frac{n}{n+k}\sigma^2$
e. None of the above choices (a)-(d).

13. Let Y and X be two random variables with finite variances. Suppose that we have a misspecified regression:

$$Y = \beta_0 + \beta_1 X + e,$$

with $E[e|X] = h(X)$ for some $h(\cdot) \neq 0$. Let $g(X)$ be a prediction of Y generated from X, and $E[(Y - g(X))^2]$ be the mean squared error (MSE) of this prediction. We also let $m(X)$ be the optimal choice of $g(X)$ which minimizes the MSE. Which of the following is right?

a. $m(X) = \beta_0 + \beta_1 X + h(X)$
b. $m(X) = \beta_1 X$
c. $m(X) = h(X)$
d. $m(X) = \beta_0 + \beta_1 X$
e. None of the above choices (a)-(d).

14. Let $\{X_i\}_{i=1}^n$ be an IID sequence of normal random variables with $\mu = E(X_i)$ and $\sigma^2 = var(X_i)$. Denote $\bar{X} := n^{-1}\sum_{i=1}^n X_i$, $\hat{e}_i = X_i - \bar{X}$ and $e_i := X_i - \mu$. Which of the following is right?

a. $n^{-1/2}\sum_{i=1}^n \hat{e}_i - n^{-1/2}\sum_{i=1}^n e_i$ degenerates to zero, as $n \to \infty$
b. $n(\bar{X} - \mu)^2$ degenerates to zero, as $n \to \infty$
c. $n^{-1/2}\sum_{i=1}^n \hat{e}_i^2 - n^{-1/2}\sum_{i=1}^n e_i^2$ degenerates to zero, as $n \to \infty$
d. $\left(n^{-1}\sum_{i=1}^n e_i^2\right) n^{1/2}(\bar{X} - \mu)$ degenerates to zero, as $n \to \infty$
e. None of the above choices (a)-(d).

15. Let X be a random variable that has the probability density function:

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

where $x \in \mathbb{R}$ and $\kappa > 0$ is the parameter. In addition, we define $Z_1 := \frac{\mathrm{d}}{\mathrm{d}\kappa}\ln f(X, \kappa)$ and $Z_2 := Z_1^2 + \frac{\mathrm{d}^2}{\mathrm{d}\kappa^2}\ln f(X, \kappa)$. Which of the following is right?

見背面

a. $E(Z_1) = \kappa\pi$ and $E(Z_2) = \kappa^2\Gamma''\left(\frac{\kappa+1}{2}\right)$
b. $E(Z_1) = \kappa\Gamma'\left(\frac{\kappa+1}{2}\right)$ and $E(Z_2) = \frac{\kappa+1}{2}\pi^2\Gamma''\left(\frac{\kappa+1}{2}\right)/\Gamma'\left(\frac{\kappa+1}{2}\right)$
c. $E(Z_1) = \kappa\pi\Gamma'\left(\frac{\kappa+1}{2}\right)$ and $E(Z_2) = \kappa^2\pi^2\Gamma''\left(\frac{\kappa+1}{2}\right)$
d. $E(Z_1) = 0$ and $E(Z_2) = 0$
e. None of the above choices (a)-(d).

16. Let I_n be the $n \times n$ identity matrix, and X be a $n \times 2$ matrix with $n > 2$:

$$X := \begin{bmatrix} x_{11} & x_{12} \\ x_{21} & x_{22} \\ \vdots & \vdots \\ x_{n1} & x_{n2} \end{bmatrix},$$

in which the x_{ij}'s are independent and $N(0,1)$-distributed random variables. Note that X' is the transpose of X. Suppose that $X'X$ is positive definite. Denote $Y := X(X'X)^{-1}X'$, $Z := I_n - Y$, $M_1 := E[\text{trace}(YZ)]$ and $M_2 := E[\text{trace}(ZZ)]$. Please calculate M_1 and M_2.

a. $M_1 = 0$ and $M_2 = (n-2)$
b. $M_1 = 0$ and $M_2 = (n-2)^2$
c. $M_1 = 0$ and $M_2 = (n-4)$
d. $M_1 = 0$ and $M_2 = (n-4)^2$
e. None of the above choices (a)-(d).

17. Let $\{X_i\}_{i=1}^n$ be an IID sequence of $N(0,1)$ random variables, and $\{Z_i\}_{i=1}^n$ be another IID sequence of $N(0,1)$ random variables. In addition, $\{X_i\}_{i=1}^n$ is independent of $\{Z_i\}_{i=1}^n$, and we define $Y_i := 3X_i^2 + 4Z_i^3$. Consider a simple linear regression:

$$Y_i = \beta_0 + \beta_1 X_i + e_i,$$

where β_0 and β_1 are regression coefficients, and e_i is a zero-mean error. Let $\hat{\beta}_0$ and $\hat{\beta}_1$ be the ordinary least squares estimators for β_0 and β_1, respectively. Suppose that $\hat{\beta}_0$ and $\hat{\beta}_1$ are, respectively, consistent for β_0^* and β_1^*, as $n \to \infty$. Please calculate β_0^* and β_1^*.

a. $\beta_0^* = 6$ and $\beta_1^* = 3$
b. $\beta_0^* = 5$ and $\beta_1^* = 2$
c. $\beta_0^* = 4$ and $\beta_1^* = 1$
d. $\beta_0^* = 3$ and $\beta_1^* = 0$
e. None of the above choices (a)-(d).

接次頁

18. Let $\{(Y_i, X_{1i}, X_{2i})\}_{i=1}^n$ be an IID sequence of trivariate normal random variables. The mean vector and the covariance matrix of $(Y_i, X_{1i}, X_{2i})'$ are unknown. Consider the following linear regression:

$$Y_i = \beta_0 + \beta_1 X_{1i} + \beta_2 X_{2i} + e_i,$$

where e_i is a zero-mean error. Let P be the p-value of the t statistic for the following hypotheses:

$$H_0 : \beta_2 = 0;$$
$$H_1 : \beta_2 > 0.$$

Please calculate the third and fourth moments of P under H_0.

a. $E(P^3) = 0.25$ and $E(P^4) = 0.20$
b. $E(P^3) = 0.45$ and $E(P^4) = 0.30$
c. $E(P^3) = 0.15$ and $E(P^4) = 0.10$
d. $E(P^3) = 0.35$ and $E(P^4) = 0.15$
e. None of the above choices (a)-(d).

19. Let $\{X_i\}_{i=1}^n$ be an IID sequence of random variables with the zero mean and the variance $\sigma^2 < \infty$. Define $\bar{X} := n^{-1}\sum_{i=1}^n X_i$. Let $M_n(\cdot)$ be the moment generating function of the statistic $n^{1/2}\bar{X}/\sigma$, and $M(\cdot)$ be the moment generating function of the random variable X_i. We also let t be an arbitrary real number. Which of the following is right?

a. $nM_n(t) = \frac{1}{\sigma^2\sqrt{n}} M(\frac{t}{\sigma\sqrt{n}})^{n-1}$
b. $M_n(t) = \frac{1}{\sigma^2\sqrt{n}} M(\frac{t}{\sigma\sqrt{n}})^{n-2}$
c. $M_n(t) = \frac{1}{\sigma\sqrt{n}} M(\frac{t}{\sigma\sqrt{n}})^{n}$
d. $M_n(t) = M(\frac{t}{\sigma\sqrt{n}})^{n}$
e. None of the above choices (a)-(d).

20. Let $\{Y_i\}_{i=1}^n$ and $\{X_i\}_{i=1}^n$ be two sequence of IID normal random variables. Consider a simple linear regression:

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

where e_i is a zero-mean error. Let $\hat{\beta}$ be the ordinary least squares estimator for β. Denote the fitted value $\hat{Y}_i := \hat{\beta} X_i$, the residual $\hat{e}_i := Y_i - \hat{Y}_i$ and the sample averages: $\bar{Y} := n^{-1}\sum_{i=1}^n Y_i$ and $\bar{\hat{Y}} := n^{-1}\sum_{i=1}^n \hat{Y}_i$. We also define

$$R^2 := \frac{\sum_{i=1}^n (Y_i - \bar{\hat{Y}})^2}{\sum_{i=1}^n (Y_i - \bar{Y})^2},$$

見背面

and let $\hat{\rho}$ be the sample correlation coefficient of $\{(Y_i, \hat{Y}_i)\}_{i=1}^n$. Which of the following is right.

a. $n^{-1}\sum_{i=1}^n \hat{e}_i = 0$;
b. $0 \le R^2 \le 1$;
c. $n^{-1}\sum_{i=1}^n X_i\hat{e}_i = 0$;
d. $R^2 = \hat{\rho}^2$;
e. All of the above choices (a)-(d).

試題隨卷繳回