國立中山大學104學年度碩士暨碩士專班招生考試試題

科目名稱：商用統計學丙【企管系企管甲班碩士班丙組選考】　　題號：441004

※本科目依簡章規定「不可以」使用計算機(混合題)　　

Notation and some information:
iid: identically independently distributed.
$N(\mu, \sigma^2)$: normal distribution with mean μ and variance σ^2.
$exp(\lambda)$ is the exponential distribution with mean $\frac{1}{\lambda}$.
cdf: cumulative distribution function. pdf: probability density function.
$\Phi(t)$: cdf of standard normal distribution.
$\Phi(0.90)$=1.28 and $\Phi(0.95)$=1.645 and $\Phi(0.975)$=1.96 and $\Phi(0.99)$=2.32.
MGF:moment generating function. $M_X(t)$: moment generating function for the random variable X.
$Cov(X, Y)$ is the covariance of the random variables X and Y. $Var(X)$ is the variance of the random variable X. $E(X)$ is the expectation of the random variable X.

1. Consider the following linear regression model:

$$Y_i = bX_i + \epsilon_i, i = 1, \ldots, n,$$

where Y_i is the ith response variable, X_i is the ith predictor variable, b is the regression coefficient, and $\epsilon_i, i = 1, \ldots, n$ are the iid normal errors with mean 0 and variance σ^2. Please answer the following questions:
(10%)(a) Please calculate the least squared estimator of b.
(5%)(b) Denote the least squared estimator of b as $\hat{b}$. Give a reasonable unbiased estimator for σ^2 expressed by $Y_i, i = 1, \ldots, n$, $X_i, i = 1, \ldots, n$ and $\hat{b}$.
Note: The intercept in this linear regression model is 0.

(15%)2. $X_1, \ldots, X_m$ are iid random variables with distribution $N(\mu_1, \sigma^2)$ and $Y_1, \ldots, Y_n$ are iid $N(\mu_2, \sigma^2)$. $X_i, i = 1, \ldots, m$ and $Y_j, j = 1 \ldots, n$ are independent. Please conduct a two-sample t test (two-sided test) to test the null hypothesis $\mu_1 = \mu_2$ with significance level α.
Note: You need to specify the null hypothesis, alternative hypothesis, the test statistic, and the rejection region.

3. Assume $X_1, \ldots, X_n$ are identically independently distributed with probability density function

$$f(x, \theta) = \frac{1}{2\theta} exp(-\frac{|x|}{\theta})$$

(10%)(a) Find the method of moment estimate of θ.
(10%)(b) Find the maximum likelihood estimate of θ.

Multiple Choice Questions (Note: only one choice is correct). 5 points for each question.

1. Assume $X_1, \ldots, X_n$ are iid random variables with distribution $N(\mu, \sigma^2)$. We construct a confidence interval for μ, then the length of the confidence interval for μ would decrease if

(a) standard error for μ increases.

(b) σ^2 increases.

(c) n increases.

(d) sample mean $\bar{X}$ increases.

2. A survey of the viewing habits of a group reveals the following information:
(a) 30% watch gymnastics
(b) 30% watch basketball
(c) 20% watch soccer
(d) 15% watch gymnastics and basketball
(e) 10% watch basketball and soccer
(f) 10% watch gymnastics and soccer
(g) 4% watch all three sports
What is the percentage of the group that watch none of the three sports?

(a) 0.49.

(b) 0.51.

(c) 0.53.

(d) 0.55.

3. Assume the $U \sim Unif[0,1]$. Which of the following statements is **wrong**?

(a) E(U)=1/2.

(b) Let $X \sim exp(1)$ and F(t) be its cdf, then the distribution of F(X) is $Unif[0,1]$.

(c) $1-U \sim Unif[0,1]$.

(d) $Cov(U, 1-U) \leq -\frac{1}{6}$.

Note: If $X \sim Unif[0,1]$, then the pdf (of X) $f(x) = 1$ if $x \in [0,1]$; $f(x) = 0$ otherwise.

4. Assume $X_1, \ldots, X_n$ are iid normal random variables with mean μ and variance σ^2. We conduct a test to test the null hypothesis $\mu = \mu_0$, the power of the test would increase if

(a) n decreases.

(b) type II error increases.

(c) the significance level α increases.

(d) standard error for μ increases.

5. $X_1, \ldots, X_n$ are iid random variables with probability density function $f(x, \theta)$. Which of the following statements about maximum likelihood estimator (MLE) for θ is **wrong**?

(a) MLE is usually asymptotically normally distributed.

(b) MLE is not always unbiased.

(c) MLE uses the information of likelihood.

(d) MLE uses only the first and the second moments.

6. Toss a fair coin 10000 times. Let A be the event that the number of heads is between 4900 and 5100. Which number is closest to the probability of A?

(a) 0.87

(b) 0.91

(c) 0.95

(d) 0.99.

7. X and Y are two normal random variables. $X \sim N(0,4)$ and $Y \sim N(0,4)$. Which of the following statements is correct?

(a) (X,Y) are bivariate normal.

(b) $Cov(X,Y) \leq 4$.

(c) $Var(X+Y) = 8$.

(d) $E(X+Y) \geq 2$.

8. X and Y are two independent random variables with MGF $M_X(t)$ and $M_Y(t)$, respectively and a and b are two positive constants. Which of the following statements is wrong?

(a) $M_{aX}(t) = M_X(at)$

(b) $M_{X+Y}(t) = M_X(t)M_Y(t)$

(c) $M''_X(0) \geq Var(X)$. where $M''_X(0)$ is the second derivative of $M_X(t)$ at 0.

(d) $M_{(X+b)}(t) = e^b M_X(t)$

9. Which of the following statements about distribution is **wrong**?

(a) The expectation and variance of Poisson distribution are the same.

(b) Sum of two independent Bernoulli random variables with same success rates p are Binomial random variables.

(c) If $X \sim exp(1)$, then X^2 has gamma distribution.

(d) Assume X and Y are two independent random variables. $X \sim exp(2)$ and $Y \sim exp(2)$. $X + Y$ has gamma distribution.

10. A researcher wishes to estimate the average height of all students at National Sun Yat-sen University (NSYSU). He would **randomly sample** 30 students and measures their heights. He considers the following four ways of random sampling. Which way of random sampling do you think is most appropriate?

(a) Go to the gym of the university and randomly select 30 students in the gym.

(b) Stay at the main entrance of the university and randomly select 30 students who pass the entrance.

(c) Go to one of the classes of the university and randomly select 30 students in that class.

(d) Go to the registration office of the university and take the list of all students and randomly select 30 students from the list.