國立高雄大學 115 學年度研究所碩士班招生考試試題

科目：統計學　　系所：財務金融學系　　是否使用計算機：是
考試時間：100 分鐘　　本科原始成績：100 分

A. Multiple Choices (70%)

1. There is a certain clinic with unknown appointment status. The waiting time for each patient (X) follows an exponential distribution. The clinic's service rate is 0.05 patient per minute. Furthermore, the opportunity cost in dollar (C) for each patient is the function of the waiting time. Now, it is known that the cost function is $\mathrm{C} = 100 + 5\mathrm{X} + 0.5X^2$.What is the expected value of the opportunity cost in dollar (C)?

(A) 400　(B) 500　(C) 600　(D) 700　(E) 800

2. Suppose that your waiting time for a bus in the morning is uniformly distributed on [0, 5]. The waiting time in the evening is uniformly distributed [0, 10], which is independent of morning waiting time. What are the expected value and variance of the difference between total morning waiting time and total evening waiting time for a particular week? (Note: There are seven days in a week.)

(A) expected value = -17.5, variance = 72.94
(B) expected value = -2.5, variance = 10.42
(C) expected value = 17.5, variance = 10.42
(D) expected value = -7.5, variance = 72.94
(E) expected value = 7.5, variance = 52.5

3. Based on historical performance, a statistics instructor established the following facts. Students who adhere to the recommended intensive study plan (two hours per class hour) pass the course (Grade C or better) with a 90% success rate. Students who do not commit to a regular study routine pass the course with only a 20% success rate. Across all enrolled students, the proportion that commits to the intensive study plan is 0.7. If a student is randomly selected from those who successfully passed the course (achieved a C or better), what is the likelihood that this student was one of those who followed the intensive study plan?

(A) 0.8524　(B) 0.5798　(C) 0.9130　(D) 0.9347　(E) 0.8431

4. Let $X_1, X_2, \ldots\ldots, X_n \sim N(0, \sigma^2)$. $\mathrm{Y} = \frac{\sum X^2}{\sigma^2}$. Please identify the distribution of Y.

(A) $Y \sim \chi^2_{(1)}$　(B) $Y \sim \chi^2_{(\mathrm{n})}$　(C) $Y \sim \chi^2_{(\mathrm{n}-1)}$　(D) $Y \sim t_{(n)}$　(E) $Y \sim t_{(n-1)}$

5. Given the following constraints. Please find the 95% prediction interval of Y when $X_0 = 1$.
$\sum X = 90,\ \sum Y = 150,\ \sum XY = 800,\ \sum X^2 = 450,\ \sum Y^2 = 1600,$

n = 20 (Sample Size), $t_{(1-\frac{\alpha}{2},18)} = 2.101,\ t_{(1-\frac{\alpha}{2},19)} = 2.093$

(A) -8.135 $\leq Y_{X_0=1} \leq$ 4.001　(B) -8.659 $\leq Y_{X_0=1} \leq$ 4.215　(C) -8.135 $\leq Y_{X_0=1} \leq$ 4.857

(D) -8.059 $\leq Y_{X_0=1} \leq$ 4.001　(E) -8.088 $\leq Y_{X_0=1} \leq$ 4.983

6. Given X~N(μ, 25), we would like to perform a *one-tailed* test for $H_0: \mu = 100$ against $H_1: \mu = 102$. How large a sample should be taken if the probability of accepting H_0 when H_1 is true is 0.025 and rejecting H_0 when H_0 is true is 0.05? $Z_{0.05} = 1.645, Z_{0.025} = 1.96$.

(A) 78　(B) 79　(C) 80　(D) 81　(E) 82

7. If the linear regression equation of $\hat{Y} = a + bX$, and we are given the following: The regression line of X on Y is $\hat{X} = 3 + 0.5Y$. The correlation coefficient is r = 0.75 and $\bar{X} = 10$. What are the values of a and b?

(A) a = 2.75, b = 1.125

(B) a = 2.15, b = 1.250

(C) a = 3.75, b = 1.250

(D) a = 3.25, b = 1.500

(E) a = 2.50, b = 1.750

8. A beer company randomly samples n bottles from a production line in the quality control department and measures their weights $X_1, X_2, \ldots\ldots, X_n$. Assume the bottle weights X~$N(\mu_X, \sigma_X^2)$. The quality control department wants to estimate the true mean weight of the beers, μ_X. Given the following five estimators, $E_1, E_2, E_3, E_4,$ and E_5, which one is the *worst* estimator? (n ≥ 2)

$$E_1 = \frac{1}{n}\sum_{1}^{n} X$$

$$E_2 = X_1$$

$$E_3 = \frac{1}{2}(X_1 + X_2)$$

$$E_4 = 0.9X_1 + \frac{0.1}{n-1}\sum_{2}^{n} X_n$$

$$E_5 = 0.5X_1 + \frac{0.5}{n-1}\sum_{2}^{n} X_n$$

(A) E_1 (B) E_2 (C) E_3 (D) E_4 (E) E_5

9. Suppose the joint probability density function of random variables X and Y is given by

$$f(x,y)=\begin{cases}4y(x-y)e^{-(x+y)}, & 0<x<\infty, 0\le y\le x\\ 0, & otherwise\end{cases}$$

Please compute E[X|Y=y].

(A) y-2 (B) y-1 (C) y (D) y+1 (E) y+2

10. Fifteen people were given two types of cereal, Brand X and Brand Y. Two people preferred Brand X and thirteen people preferred Brand Y. We want to determine whether or not customers prefer one brand over the other. The *p*-value for this test is (Note: $(0.5)^{15}=0.0000305$)

(A) 0.0014 (B) 0.0028 (C) 0.0037 (D) 0.0074 (E) 0.0098

B. Calculation (30%)

1. (10%) Consider the simple linear regression model:

$$Y_i=\alpha+\beta X_i+\epsilon_i$$

where the error terms ϵ_i satisfy the following classical linear regression model (CLRM) assumptions:

$$E(\epsilon_i)=0,\quad E(\epsilon_i^2)=\sigma^2,\quad E(\epsilon_i\epsilon_j)=0\quad (i\neq j)$$

X_i is a nonstochastic variable. We have 20 observations, ordered by X values: $(X_1,Y_1),(X_2,Y_2),\ldots,(X_{20},Y_{20})$. Let the estimator for the slope coefficient β be b,

$$b=\frac{\overline{Y_2}-\overline{Y_1}}{\overline{X_2}-\overline{X_1}}$$

where

$$\overline{X_1}=\frac{1}{10}\sum_{i=1}^{10}X_i,\quad \overline{Y_1}=\frac{1}{10}\sum_{i=1}^{10}Y_i$$

$$\overline{X_2}=\frac{1}{10}\sum_{i=11}^{20}X_i,\quad \overline{Y_2}=\frac{1}{10}\sum_{i=11}^{20}Y_i$$

(1) Is *b* the unbiased estimator of β? Explain and prove your assertion.

(2) State the content of the Gauss-Markov Theorem. Based on the theorem, is *b* the Best Linear Unbiased Estimator (BLUE) of β? Explain your conclusion.

科目：統計學　　系所：財務金融學系　　是否使用計算機：是
考試時間：100 分鐘　　本科原始成績：100 分

2. (12%) Two independent populations, $P_x \sim N(\mu_x, \sigma_x^2)$ and $P_y \sim N(\mu_y, \sigma_y^2)$ are considered. A random sample of size *n*, denoted by $X_1, X_2, \ldots\ldots, X_n$, is drawn from P_x, yielding sample statistics $\overline{\mathrm{X}}$ and S_x^2. Separately, a random sample of size *m*, $Y_1, Y_2, \ldots\ldots, Y_m$, is drawn from P_y, yielding $\overline{\mathrm{Y}}$ and S_y^2.

(1) Prove that the difference between the sample means, $\overline{\mathrm{X}} - \overline{\mathrm{Y}}$, is an unbiased estimator for the difference between the population means, $\mu_x - \mu_y$.

(2) Identify the probability distribution of the statistic $\overline{\mathrm{X}} - \overline{\mathrm{Y}}$, and state its corresponding mean and variance.

(3) Assuming $\sigma_x^2 = \sigma_y^2 = \sigma^2$, determine the distribution of the pooled sample variance term, $\frac{(n-1)S_x^2+(m-1)S_y^2}{\sigma^2}$, and state its sole parameter.

3. (8%) Let X and Y have the joint probability density function (p.d.f.) given by:

$$\mathrm{f(X,Y)} = X^2 + \frac{XY}{3}, 0 < X < 1, 0 < Y < 2$$

Please find (1) P(X>$\frac{1}{2}$), (2) P(Y<X)