國立彰化師範大學114學年度碩士班招生考試試題

系所： 統計資訊研究所(選考乙)　　　　科目： 統計學

☆☆請在答案紙上作答☆☆

1. (10%) Let $f(x) = \frac{c}{(x+1)(x+2)}$, $x = 0, 1, 2, 3, \ldots$ Determine the constant c so that $f(x)$ satisfies the conditions of being a probability mass function for a random variable X.

2. A random variable X has a binomial distribution $b(n, p)$ with the probability mass function

$$f(x) = \binom{n}{x} p^x (1-p)^{n-x}, \quad x = 0,1,2,\ldots,n.$$

(1) (20%) Find the expected value $E\left(\frac{1}{X+1}\right)$.

(2) (10%) Suppose that $n = 3$ and $p = 0.5$, find the expected value $E\left(\frac{1}{X+1}\right)$.

3. (10%) Let $M_X(t)$ be the moment-generating function of a random variable X, and define $S(t) = \log(M_X(t))$. Show that $\frac{d}{dt}S(t)|_{t=0} = E(X)$ and $\frac{d^2}{dt^2}S(t)|_{t=0} = Var(X)$.

4. A study develops two multiple linear regression models to predict a continuous dependent variable Y, based on a sample of 25 observations. The total sum of square is 100. The models are specified as follows:

Model 1: $Y=\beta_0+\beta_1X_1+\beta_2X_2+\varepsilon$ (regression sum of squares=60)

Model 2: $Y=\beta_0+\beta_1X_1+\beta_2X_2+\beta_3X_3+\varepsilon$ (regression sum of squares=61)

Here ε represents the random error assumed to follow the $N(0,\sigma^2)$.

(1) (10%) Calculate the coefficient of multiple determination (R^2) and the adjusted coefficient of multiple determination ($adj\text{-}R^2$) for both models. Explain the meaning of these values and evaluate how well each model explains the variation in the dependent variable. Based on your analysis, discuss which model is more suitable and justify your choice.

(2) (10%) For Model 2, the estimated coefficient for β_3 is 0.297, with a standard error of 0.15. Perform a t-test for β_3 to test the hypothesis: H_0: $\beta_3=0$ vs. H_1: $\beta_3\neq0$ at a significance level of 0.05. Based on the results, determine whether to reject H_0 and explain the implications for including X_3 in Model 2. Additionally, compare the findings from part (1) and (2). Discuss which model you would recommend and justify your reasoning.
(Reference t-values: $t_{0.025,\text{ df=1}}=12.706$, $t_{0.05,\text{ df=1}}=6.314$,
$t_{0.025,\text{ df=21}}=2.080$, $t_{0.05,\text{ df=21}}=1.721$, $t_{0.025,\text{ df=22}}=2.074$, $t_{0.05,\text{ df=22}}=1.717$,
$t_{0.025,\text{ df=23}}=2.069$, $t_{0.05,\text{ df=23}}=1.714$, $t_{0.025,\text{ df=24}}=2.064$, $t_{0.05,\text{ df=24}}=1.711$)

5. (10%) Let $X_1, X_2, \ldots, X_n$ be a random sample from a normal distribution $N(\mu,\sigma^2)$, where μ and σ^2 are unknown parameters. Define a new parameter $\tau = \mu^2+\sigma^2$. Please derive the maximum likelihood estimator (MLE) of τ.

6. Given a random sample $X_1, X_2, \ldots, X_n$ from a Bernoulli distribution with an unknown parameter p, answer the following:
   (1) (5%) Identify the sufficient statistic for the parameter p and provide a justification.

   (2) (5%) Find the uniformly minimum variance unbiased estimator (UMVUE) for p.

   (3) (7%) For the hypotheses H_0: $p \le 0.5$ vs. H_1: $p > 0.5$ with n=10 and significance level α=0.1, derive the uniformly most powerful (UMP) test and compute the critical value.

   (4) (3%) If the observed sample gives $X_1+X_2+\ldots+X_{10}$ =7, determine whether H_0 is rejected based on the test derived in part (3).