臺灣大學數學系

八十六學年度第一學期碩博士班資格考試試題

統計與機率

1. (10 points) Show that if $X_1, \ldots, X_{2n}$; $Y_1, \ldots, Y_n$ are independent and identically distributed random variables with $EX_1 = EY_1 = \lambda = Var(X_1) = Var(X_2)$. Find the limiting distribution of the sequence of random variables

$$F_n = \sqrt{n}\frac{\sum_{i=1}^{2n}(X_i - \lambda)}{\sum_{j=1}^{n} Y_j}.$$

2. (10 points) Let $\{X_n\}$ be a sequence of random variables satisfying

$$X_1 > X_2 > \cdots > 0 \quad \text{almost surely.}$$

Then show that X_n converges to 0 in probability implies that X_n converges to 0 almost surely.

3. (10 points) Consider an urn which contains $b \geq 1$ black and $w \geq 1$ white balls which are well mixed. Repeated drawings are made from the urn, and after each drawing the ball drawn is replaced, along with c balls of the same color. Here $c \geq 1$ is an integer. Let $X_0 = b/(b+w)$, and X_n be the proportion of black balls in the urn after the nth draw. Find the probability that the nth ball drawn is black.

4. (20 points) Let $X_1, \ldots, X_n$ be a sample from uniform (θ_1, θ_2) where θ_1, θ_2 are unknown.
    1. Show that $T(\mathbf{X}) = (\min(X_1, \ldots, X_n), \max(X_1, \ldots, X_n))$ is sufficient. (10 points)
    2. Assuming that $T(\mathbf{X})$ is complete find a U.M.V.U. estimate of $(\theta_1 + \theta_2)/2$. (10 points)

5. (15 points) Suppose $(X_1, \ldots, X_n)$ is a sample from a population with density

$$f(x, \theta) = \frac{9}{10\sigma}\phi\left(\frac{x-\mu}{\sigma}\right) + \frac{1}{10}\phi(x-\mu)$$

where ϕ is the standard normal density and

$\theta = (\mu, \sigma^2) \in \Theta = \{(\mu, \sigma^2) : -\infty < \mu < \infty, \sigma^2 > 0\}$.

1. Show that maximum likelihood estimates do not exist, but that $\sup_\sigma p(\mathbf{x}, \hat{\mu}, \sigma^2) = \sup_{\mu,\sigma} p(\mathbf{x}, \hat{\mu}, \sigma^2)$ if, and only if, $\hat{\mu}$ equals one of the numbers $x_1, \dots, x_n$. Here $\mathbf{x} = (x_1, \dots, x_n)$. (10 points)
2. Give a brief explanation why (a) holds. (5 points)
   Hint: continuous pdf versus discrete pdf

6. (15 points) The normally distributed random variables $X_1, \dots, X_n$ are said to be *serially correlated* if we can write

$$X_i = \theta X_{i-1} + \epsilon_i, \quad i = 1, \dots, n,$$

where $X_0 = 0$ and $\epsilon_1, \dots, \epsilon_n$ are independent $N(0, \sigma^2)$ random variables. Derive the likelihood ratio test of $H : \theta = 0$ (independence) versus $K : \theta \neq 0$ (serial correlation).

7. (20 points) Assume that $Y_i = \beta X_i + \epsilon_i$ for $1 \leq i \leq n$, where $X_1, \dots, X_n$ are independent $N(\mu, \tau^2)$ random variables ($\mu \neq 0$), and $\epsilon_1, \dots, \epsilon_n$ are iid $N(0, \sigma^2)$, and the Xs and ϵs are independent. Statistician A proposes to estimate β by $\sum_{i=1}^n Y_i / \sum_{i=1}^n X_i$, denoted it by $\hat{\beta}_{n1}$, since Y_i is normally distributed with mean $\beta\mu$. Statistician B proposes to estimate β by $\sum_{i=1}^n X_i Y_i / \sum_{i=1}^n X_i^2$, denoted it by $\hat{\beta}_{n2}$, since (X_i, Y_i) follows a linear regression model if Xs are viewed as constants. For these two estimates $\hat{\beta}_{n1}$ and $\hat{\beta}_{n2}$, which one will you use? Please give reasons to support your conclusion.