PhD entrance exam 90 <u>MathStat</u> Dept of Statistics

1. (30 points) Let $X_1, ..., X_n$ be a random sample from $N(\mu, \sigma^2)$ population, where μ and σ are unknown parameters. Define

$$\bar{X}_n = \frac{1}{n}\sum_{i=1}^{n} X_i$$

and

$$S^2 = \frac{1}{\sigma^2}\sum_{i=1}^{n}(X_i - \bar{X}_n)^2.$$

Show that
(a) $\bar{X}_n$ and S^2 are independent.
(b) S^2 is distributed as χ^2_{n-1}.

2. (10 points) Let $X_1, ..., X_n$ be a random sample from $N(0,1)$ population. Define

$$T_n = \frac{(1/k)\sum_{i=1}^{k} X_i^2}{[1/(n-k)]\sum_{j=k+1}^{n} X_j^2}.$$

Suppose that k is fixed. Find the limiting distribution of T_n as $n \to \infty$.

3. (30 points) Let $X_1, ..., X_n$ be normally distributed random variables satisfying the following model

$$X_i = \theta X_{i-1} + e_i, \quad i = 1, ..., n,$$

where $X_0 = 0$ and $e_1, ..., e_n$ are independent $N(0, \sigma^2)$ random variables.
(a) Find the joint density of $X_1, ..., X_n$.
(b) Derive the likelihood ratio statistic of $H_0 : \theta = 0$ v.s. $H_1 : \theta \neq 0$.

4. (30 points) Let $X_{(1)} < ... < X_{(n)}$ be the order statistics of a sample of size n from an *Exponential*(1) population. Show that $nX_{(1)}$, $(n-1)(X_{(2)} - X_{(1)})$, $(n-2)(X_{(3)} - X_{(2)})$,..., $(X_{(n)} - X_{(n-1)})$ are i.i.d. random variables according to *Exponential*(1).

1. **(30%)** Consider the model $y_{ij} = \mu_i + \gamma_1 z_{ij} + \gamma_2 w_{ij} + \epsilon_{ij}$, where $i = 1, 2, \cdots, I$; $j = 1, 2, \cdots, J$; and $\epsilon_{ij} \overset{iid}{\sim} N(0, \sigma^2)$. Let $\boldsymbol{\gamma} = (\gamma_1 \;\; \gamma_2)^T$.

(1) Please derive the least squares estimate (LSE) of $\boldsymbol{\gamma}$ and also show that it is unbiased. (10%)

(2) Find the variance-covariance matrix of the LSE $\hat{\boldsymbol{\gamma}}$ of $\boldsymbol{\gamma}$. (5%)

(3) Under what conditions are $\hat{\gamma}_1$ and $\hat{\gamma}_2$ statistically independent? (5%)

(4) Obtain a test statistic for testing the hypothesis $H_0 : \gamma_i = \gamma_0$, where $i = 1, 2$. (10%)

2. **(20%)** Given the model $y_i = \beta x_i + \epsilon_i \quad (i = 1, \cdots, n)$. Let us suppose that $Var(Y) = diag(1/\omega_1, 1/\omega_2, \cdots, 1/\omega_n)\sigma^2$, where $\omega_i > 0$ for all i.

(1) Please find the estimate of β, $\hat{\beta}$, and $Var(\hat{\beta})$. (10%)

(2) Show how to predict y_* for a given value x_* of x and also construct the confidence interval for it. (5%)

(3) What would the estimate of β be if $Var(y_i) = kx_i$? (5%)

3. **(30%)** Suppose $y_t = \beta + \epsilon_t$, where $\epsilon_t = \rho\epsilon_{t-1} + a_t$, $t = 1, 2, \cdots, T$, $0 \le \rho \le 1$, $a_0 = 0$, and $a_t \overset{iid}{\sim} N(0, \sigma^2)$.

(1) Show that the sample mean of $\{y_t;\ t = 1, \cdots, T\}$, $\bar{Y}$, is still unbiased for β. (10%)

(2) Let

$$\hat{\beta} = \frac{y_1 + (1-\rho)\sum_{t=2}^{T}(y_t - \rho y_{t-1})}{(T-1)(1-\rho)^2 + 1}.$$

Show that $\hat{\beta}$ is unbiased. (10%)

(3) Show that $Var(\bar{Y}) \ge Var(\hat{\beta})$ with strict inequality unless $\rho = 0$. (10%)

4. **(20%)** Consider the model $\boldsymbol{Y} = \boldsymbol{X}\boldsymbol{\beta} + \boldsymbol{\epsilon}$ with $\boldsymbol{X}$ and $\boldsymbol{\beta}$ partitioned as

$$\boldsymbol{X} = [\boldsymbol{X}_1 \mid \boldsymbol{X}_2] \text{ and } \boldsymbol{\beta} = \begin{bmatrix} \boldsymbol{\gamma}_1 \\ -- \\ \boldsymbol{\gamma}_2 \end{bmatrix}$$

where $\boldsymbol{X}_1$ is an $n \times r$ matrix of rank r, $\boldsymbol{X}_2$ is an $n \times (p-r)$ matrix of rank $p - r$, $\boldsymbol{\gamma}_1$ is an $r \times 1$ vector, and $\boldsymbol{\gamma}_2$ is a $(p-r) \times 1$ vector. Suppose $\boldsymbol{\epsilon} \sim N(\boldsymbol{0}, \sigma^2 I)$. Please obtain a test statistic for testing $H_0 : \boldsymbol{\gamma}_2 = \boldsymbol{0}$.