Ph.D. Prelim: Exam A

Distribution Theory & Regression Analysis

Jan 2018

1. Let $X_1, X_2, \ldots, X_n$ be independent identically distributed random variables each having a normal distribution with mean μ and variance σ^2. Define the random variables $\bar{X} = \frac{\sum_{i=1}^n X_i}{n}$, and $S^2 = \frac{\sum_{i=1}^n (X_i - \bar{X})^2}{n-1}$. Show that $\bar{X}$ and S^2 are independent and $\frac{(n-1)S^2}{\sigma^2}$ has a chi-square, with $n-1$ degrees of freedom, distribution.

2. Let $X_1, X_2, \ldots, X_n$ denote a random sample from a distribution of the continuous type having a probability density function $f_X(x)$ and cumulative distribution function $F_X(x)$ that has support $S = (a, b)$, where $\infty \leq a < b \leq \infty$. Let Y_1 be the smallest of these X_i, Y_2 the next X_i in order of magnitude, . . ., and Y_n the largest of X_i. That is, $Y_1 < Y_2 < \ldots < Y_n$ represent $X_1, X_2, \ldots, X_n$ when the latter are arranged in ascending order of magnitude. We call $Y_i, i = 1, 2, \ldots, n$, the i th order statistic of the random sample $X_1, X_2, \ldots, X_n, n > 10$. Derive the joint probability density function of $(Y_i, Y_j, Y_k, Y_l, Y_m)$, where $1 < i < j < k < l < m < n$.

3. Let $\{a_1, \ldots, a_n\}$ be a set of positive real numbers. Use methods in probability distributions to show that $\frac{1}{\frac{1}{n}\sum_{i=1}^n \frac{1}{a_i}} \leq (a_1 \cdots a_n)^{1/n} \leq \frac{1}{n}\sum_{i=1}^n a_i$.

4. (a) Prove the following Theorem: If $y \sim N(0, \sigma^2 I)$ and M is a symmetric idempotent matrix of rank m ($m \leq n =$ rank of I), then

$$\frac{y'My}{\sigma^2} \sim \chi^2(tr(M)).$$

   (b) Use above Theorem to show that in a normal error simple linear regression $Y_i = \beta_0 + X_i\beta_1 + \epsilon_i$ for $i \in \{1, 2, \ldots, n\}$ with $\epsilon = (\epsilon_1, \epsilon_2, \ldots, \epsilon_n)' \sim N(0, \sigma^2 I)$, the residual sum of squares

$$ESS/\sigma^2 = \sum_{i=1}^{n}(y_i - \hat{y}_i)^2/\sigma^2 \sim \chi^2_{n-2}$$

under $H_0 : \beta_1 = 0$.

(c) Use above Theorem to show that in a normal error simple linear regression $Y_i = \beta_0 + X_i\beta_1 + \epsilon_i$ for $i \in \{1, 2, \ldots, n\}$ with $\epsilon = (\epsilon_1, \epsilon_2, \ldots, \epsilon_n)' \sim N(0, \sigma^2 I)$, the regression sum of squares

$$RSS/\sigma^2 = \sum_{i=1}^{n}(\bar{y} - \hat{y}_i)^2/\sigma^2 \sim \chi^2_1$$

under $H_0 : \beta_1 = 0$.

5. (a) Prove the following Theorem: Let the $n \times 1$ vector $y \sim N(0, I)$, let A be an $n \times n$ idempotent matrix of rank $m \leq n$, let B be an $n \times n$ idempotent matrix of rank $s \leq n$, and suppose $BA = 0$. Then $y'Ay$ and $y'By$ are independently distributed χ^2 variables.

(b) Show that in a normal error simple linear regression $Y_i = \beta_0 + X_i\beta_1 + \epsilon_i$ for $i \in \{1, 2, \ldots, n\}$ with $\epsilon = (\epsilon_1, \epsilon_2, \ldots, \epsilon_n)' \sim N(0, \sigma^2 I)$, $F = RMS/EMS$ follows $F_{1,n-2}$ under $H_0 : \beta_1 = 0$, where RMS is defined as the regression mean squares and EMS is defined as error mean squares.

6. (a) Consider a simple linear regression that goes through the origin:

$$Y_i = \beta_1 X_i + \epsilon_i$$

where $\epsilon_i \sim N(0, \sigma^2)$ are independent, for $i = 1, 2, ..., n$. Obtain $\hat{\beta}_1$, the least square estimator of β_1, and construct a 95% confidence interval for β_1.

(b) Consider an extended model:

$$Y_i = \beta_1 X_i + \beta_2 Z_i + \epsilon_i$$

where $\epsilon_i \sim N(0, \sigma^2)$ $(i = 1, 2, \ldots, n)$ are independent, and Z is a second predictor variable. Specify a condition under which the least square estimator of β_1 in part (b) is identical to the the least square estimator of β_1 obtained in part (a). Prove this fact mathematically.