

The Hong Kong University of Science & Technology
MATH3423 - Statistical Inference
Midterm Examination - Fall 2013/2014

Answer ALL Questions

Date: 17 October 2013

Full marks: 30 + 3 marks for Bonus

Time Allowed: 75 minutes

- DO NOT open the exam paper until instructed to do so.
- It is a closed-book examination.
- Only the calculator approved by H.K.E.A. is allowed in the final examination.
- Three questions are included in the paper.
- Give all your answers in 4 decimal points.

Name : _____

Student Number : _____

Signature : _____

1. (8 marks, 1 mark each) Let X_1, \dots, X_n be a random sample from $N(\mu, \sigma^2)$. Define

$$\bar{X}_k = \frac{1}{k} \sum_{i=1}^k X_i$$

$$\bar{X}_{n-k} = \frac{1}{n-k} \sum_{i=1+k}^n X_i$$

and

$$S_k^2 = \frac{1}{k-1} \sum_{i=1}^k (X_i - \bar{X}_k)^2,$$

$$S_{n-k}^2 = \frac{1}{n-k-1} \sum_{i=k+1}^n (X_i - \bar{X}_{n-k})^2,$$

Answer the following:

- What is the distribution of $((k-1)S_k^2 + (n-k-1)S_{n-k}^2)/\sigma^2$?
- What is the distribution of S_k^2/S_{n-k}^2 ?
- What is the distribution of $(\bar{X}_k + \bar{X}_{n-k})/2$?
- What is the distribution of $(\bar{X}_n - \mu)/(S_n/\sqrt{n})$?

If $\mu = 0$ $\sigma = 1$,

- What is the distribution of $k\bar{X}_k^2 + (n-k)\bar{X}_{n-k}^2$?
- What is the distribution of X_1^2/X_2^2 ?
- What is the distribution of X_1/X_2 ?
- What is the distribution of $(X_1 + X_2)^2/(X_1 - X_2)^2$?

Solution:

$$\begin{aligned} \text{(a) Since } \frac{(k-1)S_k^2}{\sigma^2} &\sim \chi_{k-1}^2, \frac{(n-k-1)S_{n-k}^2}{\sigma^2} \sim \chi_{n-k-1}^2 \text{ and } S_k^2 \text{ and } S_{n-k}^2 \text{ are independent} \\ \Rightarrow \frac{(k-1)S_k^2}{\sigma^2} + \frac{(n-k-1)S_{n-k}^2}{\sigma^2} &\sim \chi_{n-2}^2 \end{aligned}$$

$$\begin{aligned} \text{(b) Since } \frac{(k-1)S_k^2}{\sigma^2} &\sim \chi_{k-1}^2, \frac{(n-k-1)S_{n-k}^2}{\sigma^2} \sim \chi_{n-k-1}^2 \text{ and } S_k^2 \text{ and } S_{n-k}^2 \text{ are independent} \\ \Rightarrow \frac{(k-1)S_k^2/(k-1)\sigma^2}{(n-k-1)S_{n-k}^2/(n-k-1)\sigma^2} &= \frac{S_k^2}{S_{n-k}^2} \sim F_{k-1, n-k-1} \end{aligned}$$

$$\begin{aligned} \text{(c) Since } \bar{X}_k &\sim N(\mu, \frac{\sigma^2}{k}), \bar{X}_{n-k} \sim N(\mu, \frac{\sigma^2}{n-k}) \text{ and } \bar{X}_k \text{ and } \bar{X}_{n-k} \text{ are independent} \\ \Rightarrow \frac{1}{2}(\bar{X}_k + \bar{X}_{n-k}) &\sim N(\mu, \frac{n\sigma^2}{4(n-k)k}) \end{aligned}$$

$$\begin{aligned} \text{(d) Since } \frac{\bar{X}_n - \mu}{\sigma/\sqrt{n}} &\sim N(0, 1), \frac{(n-1)S_n^2}{\sigma^2} \sim \chi_{n-1}^2 \text{ and they are independent} \\ \Rightarrow \frac{\bar{X}_n - \mu}{\sigma/\sqrt{n}} / \sqrt{\frac{(n-1)S_n^2}{(n-1)\sigma^2}} &= \frac{\bar{X}_n - \mu}{S_n/\sqrt{n}} \sim t_{n-1} \end{aligned}$$

If $\mu = 0$ $\sigma = 1$,

$$\begin{aligned} \text{(e) } \bar{X}_k &\sim N(0, \frac{1}{k}), \bar{X}_{n-k} \sim N(0, \frac{1}{n-k}) \text{ and they are independent} \\ \Rightarrow k\bar{X}_k^2 &\sim \chi_1^2, (n-k)\bar{X}_{n-k}^2 \sim \chi_1^2 \Rightarrow k\bar{X}_k^2 + (n-k)\bar{X}_{n-k}^2 \sim \chi_2^2 \end{aligned}$$

$$\text{(f) } X_1^2 \sim \chi_1^2, X_2^2 \sim \chi_1^2 \text{ and they are independent} \Rightarrow \frac{X_1^2}{X_2^2} \sim F_{1,1}$$

- (g) Notice $\frac{X_1}{\sqrt{X_2^2/1}} \sim t$, and $X_2 \sim N(0, 1)$ is symmetrically distributed $\Rightarrow \frac{X_1}{\sqrt{X_2^2/1}} = \frac{X_1}{X_2} \sim t_1$
- (h) Notice $\begin{pmatrix} 1 & 1 \\ 1 & -1 \end{pmatrix} \begin{pmatrix} X_1 \\ X_2 \end{pmatrix} = \begin{pmatrix} X_1 + X_2 \\ X_1 - X_2 \end{pmatrix} \Rightarrow X_1 + X_2$ and $X_1 - X_2$ are independent
- Since $X_1 + X_2 \sim N(0, 2)$ and $X_1 - X_2 \sim N(0, 2) \Rightarrow \frac{1}{2}(X_1 + X_2)^2 \sim \chi_1^2, \quad \frac{1}{2}(X_1 - X_2)^2 \sim \chi_1^2$
- $\Rightarrow \frac{\frac{1}{2}(X_1 + X_2)^2}{\frac{1}{2}(X_1 - X_2)^2} = \frac{(X_1 + X_2)^2}{(X_1 - X_2)^2} \sim F_{1,1}$

2. (14 marks) Let U_1, \dots, U_n be a random sample from the $U(0, \theta)$, where θ is the unknown parameter.

- (a) (2 marks) Find the moment estimator of θ . Is it unbiased? Hence or otherwise, find an unbiased estimator for θ .
- (b) (3 marks) Find the maximum likelihood estimator of θ . Is it unbiased? Hence or otherwise, find an unbiased estimator for θ .
- (c) (4 marks) Find the variances of unbiased estimators from (a) and (b). Which unbiased estimator for θ is more efficient?
- (d) (2 marks) Suppose a random sample with sample size six is drawn. The values are as follow.

0.3, 1.2, 1.8, 2.4, 4.1, 5.5

Calculate estimates from methods of moment and maximum likelihood. Hence or otherwise, state one problem of moment estimator **other than efficiency**

- (e) (2 marks) Given that $\theta > 1$, find the maximum likelihood estimator of θ .
- (f) (1 mark) Find the MLE for θ^3 .
- (g) (Bonus: 3 marks) Calculate the variance of MLE for θ^3 . Hence or otherwise, find its Mean Squared Error (MSE).

Solution:

- (a) Since $E(U_i) = \frac{\theta}{2}$, the moment estimator of θ is $\tilde{\theta} = 2\bar{U}_n$.
 $E(\tilde{\theta}) = 2\bar{U}_n = \theta \Rightarrow \tilde{\theta}$ is an unbiased estimator.
- (b) Likelihood function $L(\theta) = \frac{1}{\theta^n} \prod_{i=1}^n I(0 < U_i < \theta)$. The maximum of $L(\theta)$ is achieved when $\hat{\theta} = U_{(n)}$.
 Let $Y = U_{(n)}$, then $f(y) = \frac{ny^{n-1}}{\theta^n}$.
 $E(\hat{\theta}) = E(U_{(n)}) = \frac{n\theta}{n+1} \Rightarrow$ an unbiased estimator should be $\hat{\theta}_{\text{unbiased}} = \frac{n+1}{n}U_{(n)}$
- (c) $Var(\tilde{\theta}) = \frac{4}{n}Var(U_i) = \frac{\theta^2}{3n}$
 $Var(\hat{\theta}_{\text{unbiased}}) = \frac{(n+1)^2}{n^2}Var(U_{(n)}) = \frac{(n+1)^2}{n^2} \cdot \frac{n\theta^2}{(n+1)^2(n+2)} = \frac{\theta^2}{n(n+2)}$
 Therefore, when $n+2 = 3$, i.e., $n = 1$, $\tilde{\theta}$ and $\hat{\theta}_{\text{unbiased}}$ are equally efficient; when $n > 1$, $\hat{\theta}_{\text{unbiased}}$ is more efficient than $\tilde{\theta}$.
- (d) moment estimator: $\tilde{\theta} = 2\bar{U}_6 = 2 \times 2.55 = 5.1$
 maximum likelihood estimator: $\hat{\theta} = U_{(n)} = 5.5$
 Some data points may have values greater than the moment estimator, i.e., it is poor estimator.

- (e) Likelihood function $L(\theta) = \frac{1}{\theta^n} \prod_{i=1}^n I(0 < U_i < \theta)I(\theta > 1)$. The maximum of $L(\theta)$ is achieved when $\tilde{\theta}_a = \max\{U_{(n)}, 1\}$.
- (f) By invariance property, m.l.e. of θ^3 is $U_{(n)}^3$.
- (g) By the density function obtained in (b)

$$\begin{aligned} E(U_{(n)}^3) &= \frac{n}{\theta^n} \int_0^\theta y^3 \cdot y^{n-1} dy = \frac{n}{(n+3)\theta^n} \int_0^\theta dy^{n+3} = \frac{n}{n+3} \theta^3 \\ E(U_{(n)}^6) &= \frac{n}{\theta^n} \int_0^\theta y^6 \cdot y^{n-1} dy = \frac{n}{(n+6)\theta^n} \int_0^\theta dy^{n+6} = \frac{n}{n+6} \theta^6 \end{aligned}$$

$$\begin{aligned} \Rightarrow \text{Var}(U_{(n)}^3) &= E(U_{(n)}^6) - [E(U_{(n)}^3)]^2 = \left[\frac{n}{n+6} - \frac{n^2}{(n+3)^2} \right] \theta^6 = \frac{9n\theta^6}{(n+6)(n+3)^2} \\ \Rightarrow \text{MSE}(U_{(n)}^3) &= \text{Var}(U_{(n)}^3) + [E(U_{(n)}^3) - \theta^3]^2 = \left[\frac{n}{n+6} - \frac{n^2-9}{(n+3)^2} \right] \theta^6 = \frac{18\theta^6}{(n+6)(n+3)} \end{aligned}$$

3. (8 marks) Let U_1, \dots, U_n be a random sample from the $U(0,1)$.

(a) (2 marks) Let $X = -\log(U)$. Find the distribution of X .

(b) (6 marks) Let $Y = \frac{1}{\prod_{i=1}^n U_i^{\frac{1}{n}}}$, where U_1, \dots, U_n be a random sample from the $U(0,1)$ and n is very large. Using Central Limit Theorem and Delta method to find the approximate distribution of Y .

Solution:

(a) Let $X \sim F_X(x)$ where $F_X(x)$ is the CDF of X .

$$\begin{aligned} F_X(x) &= P(X \leq x) = P(-\log(U) \leq x) \\ &= P(\log(U) \geq -x) = P(U \geq \exp(-x)) = 1 - P(U \leq \exp(-x)) \\ &= 1 - \exp(-x) \end{aligned}$$

Therefore, $X \sim \text{Exp}(1)$

(b) By (a), $\log(Y) = -\frac{1}{n} \sum_{i=1}^n \log(U_i)$, where $-\log(U_i) \sim_{iid} \text{Exp}(1)$.

Since $E(-\log(U_i)) = 1$ and $\text{Var}(-\log(U_i)) = 1$, by central limit theorem,

$$\sqrt{n} \left[-\frac{1}{n} \sum_{i=1}^n \log(U_i) - 1 \right] \rightarrow_d N(0, 1)$$

Since $Y = \exp \left\{ -\frac{1}{n} \sum_{i=1}^n \log(U_i) \right\}$, by delta method,

$$\sqrt{n}(Y - \exp(1)) \rightarrow_d N(0, \exp(2))$$

Therefore, for large enough n , $Y \rightarrow N(e, \frac{e^2}{n})$

***** END *****