MATH 243 Wong 50
final 98/99 Solutions for post final 6 DEC 2004 a

1. (a) Two hundred randomly selected electronic devices of a certain type were tested and the following frequency distribution of their 'lift-times', measured in months, was compiled:

Lift time (months)	Observed freq.
$0 \leq x < 3$	53
$3 \leq x < 6$	42
$6 \leq x < 9$	35
$9 \leq x$	70

Using a test at significance level 0.01, examine the null hypothesis H_0 that the life time distribution is an exponential distribution with mean 12.

H_0 : the life time distribution is an exponential distribution with mean 12

H_1 : otherwise

$[a_i, b_i)$	$[0, 3)$	$[3, 6)$	$[6, 9)$	$[9, \infty)$
$P_r(a_i \leq X < b_i)$	0.2212	0.1723	0.1342	0.4723
exp. freq.	44.24	34.46	26.84	94.46

4

the Pearson's statistic $= \frac{(53-44.24)^2}{44.24} + \frac{(42-34.46)^2}{34.46} + \frac{(35-26.84)^2}{26.84} + \frac{(70-94.46)^2}{94.46}$

$= 1.735 + 1.650 + 2.481 + 6.334$

$= 12.20$ 4

$\chi^2_{(4-1-0,\, 0.01)} = \chi^2_{(3,\, 0.01)} = 11.34$ 2

$\Rightarrow$ Reject H_0 at $\alpha = 0.01$

(b) The data of the following table were obtained from a random sample of 300 car owners, each of whom was classified both according to age and to the number of accidents he or she had been involved in during the past two years. Using a test at the 5% level of significance, examine the null hypothesis that there is no dependence of accident rate on age in the sample population.

		Number of Accidents			
		0	1	2 or more	
Age	Under 22 year	10	21	14	45
	Between 22 and 32 years	22	43	10	75
	Over 32 years	81	80	19	180
		113	144	43	300

the test statistic = 19.34 just as (a).

$\chi^2_{(4, 0.05)} = 9.49$

$\Rightarrow$ Reject Ho at $\alpha = 0.05$.

8

2

2. (a) If X is the number of successes in n binomial trials, $\hat{\theta}_1 = X/n$, and $\hat{\theta}_2 = (X+1)/(n+2)$, for what values of θ is $E\left[\left(\hat{\theta}_2 - \theta\right)^2\right]$ less than $E\left[\left(\hat{\theta}_1 - \theta\right)^2\right]$? Do you consider $\hat{\theta}_1$ as a better estimator than $\hat{\theta}_2$?

$$E(\hat{\theta}_1) = E\left(\frac{X}{n}\right)$$

$$= \theta \quad (\text{unbiased})$$ 1

$$E(\hat{\theta}_2) = E\left(\frac{X+1}{n+2}\right)$$

$$= \frac{n\theta+1}{n+2} \quad (\text{biased})$$ 1

$$E(\hat{\theta}_1-\theta)^2 = Var(\hat{\theta}_1)$$

$$= \frac{\theta(1-\theta)}{n}$$ 1

$$E(\hat{\theta}_2-\theta)^2 = Var(\hat{\theta}_2) + (E(\hat{\theta}_2)-\theta)^2$$

$$= \frac{n\theta(1-\theta)}{(n+2)^2} + \left(\frac{1-2\theta}{n+2}\right)^2$$

$$= \frac{n\theta(1-\theta)+(1-2\theta)^2}{(n+2)^2}$$ 2

$$E(\hat{\theta}_1-\theta)^2 > E(\hat{\theta}_2-\theta)^2$$

$$\Rightarrow \frac{\theta(1-\theta)}{n} > \frac{n\theta(1-\theta)+(1-2\theta)^2}{(n+2)^2}$$

$$\Rightarrow 4(2n+1)\theta^2 - 4(2n+1)\theta + n < 0$$

$$\Rightarrow \frac{1}{2} - \frac{1}{2}\sqrt{\frac{n+1}{2n+1}} < \theta < \frac{1}{2} + \frac{1}{2}\sqrt{\frac{n+1}{2n+1}}$$ 3

2

(b) Let $Y_1 < Y_2 < Y_3$ be the order statistics of a random sample of size 3 from the uniform distribution having p.d.f. $f(x;\theta) = 1/\theta,\ 0 < x < \theta,\ 0 < \theta < \infty$, zero elsewhere. Show that $4Y_1$ and $\frac{4}{3}Y_3$ are unbiased statistics for θ. Find the variance of each of these unbiased statistics.

$$f_{Y_1}(y_1) = 3[1 - F_X(y_1)]^2 f_X(y_1)$$
$$= 3\left(1 - \frac{y_1}{\theta}\right)^2 \left(\frac{1}{\theta}\right)$$ 2

$$E(Y_1) = \int_0^\theta y_1 f_{Y_1}(y_1)\,dy_1$$
$$= \int_0^\theta y_1 \frac{3}{\theta}\left(1 - \frac{y_1}{\theta}\right)^2 dy_1$$
$$= \frac{3}{\theta^3}\int_0^\theta y_1(\theta - y_1)^2\,dy_1$$
$$= \frac{3}{\theta^3}\int_0^\theta (y_1^3 - 2\theta y_1^2 + \theta^2 y_1)\,dy_1$$
$$= \frac{3}{\theta^3}\left[\frac{y_1^4}{4} - \frac{2}{3}\theta y_1^3 + \frac{\theta^2 y_1^2}{2}\right]_0^\theta$$
$$= \frac{\theta}{4}$$

$\Rightarrow 4Y_1$ is unbiased for θ

$$E(Y_1^2) = \int_0^\theta y_1^2 f_{Y_1}(y_1)\,dy_1$$
$$= \int_0^\theta y_1^2 \frac{3}{\theta}\left(1 - \frac{y_1}{\theta}\right)^2 dy_1$$
$$= \frac{3}{\theta^3}\int_0^\theta y_1^2(\theta - y_1)^2\,dy_1$$
$$= \frac{\theta^2}{10}$$

$$Var(Y_1) = \frac{3}{80}\theta^2$$
$$Var(4Y_1) = \frac{3}{5}\theta^2$$ 3

$$f_{Y_3}(y_3) = 3[F_X(y_3)]^2 f_X(y_3)$$
$$= 3\left(\frac{y_3}{\theta}\right)^2\left(\frac{1}{\theta}\right)$$ 2

$$E(Y_3) = \int_0^\theta y_3 \frac{3}{\theta^3} y_3^2\,dy_3$$
$$= \frac{3}{\theta^3}\int_0^\theta y_3^3\,dy_3$$
$$= \frac{3}{4}\theta$$

$\Rightarrow \frac{4}{3}Y_3$ is unbiased for θ

$$E(Y_3^2) = \int_0^\theta y_3^2 \frac{3}{\theta^3} y_3^2\,dy_3$$
$$= \frac{3}{\theta^3}\int_0^\theta y_3^4\,dy_3$$
$$= \frac{3}{5}\theta^2$$

$$Var(Y_3) = \frac{3}{80}\theta^2$$
$$Var\left(\frac{4}{3}Y_3\right) = \frac{\theta^2}{15}$$ 3

You also can see the solutions in Q22 of Ex 2.

2. Final 98/99 Q3

Type I error is the error of rejecting H_0 when it is in fact true.
Type II error is the error of not rejecting H_0 when it is in fact false.
H_0 is rejected when $\underset{\sim}{X} \in C_1$ where C_1 is the critical region.
The power of a test for testing $H_0: \theta \in \Theta_0$ v.s. $H_1: \theta \in \Theta_1$ is $Q(\theta) = P(\underset{\sim}{X} \in C_1 \mid \theta)$ $\forall \theta \in \Theta_1$.

Let $(X_1, X_2, \ldots, X_n)$ and $(Y_1, Y_2, \ldots, Y_m)$ be random samples from the independent normal distributions $N(\theta_1, \theta_3)$ and $N(\theta_2, \theta_4)$, respectively.

(a) $\begin{cases} H_0: \theta_1 = \theta_2, \ \theta_3 = \theta_4 \\ H_1: \text{otherwise} \end{cases}$

$L(\theta_1, \theta_2, \theta_3, \theta_4, \underset{\sim}{x}, \underset{\sim}{y}) = f(\underset{\sim}{x}, \underset{\sim}{y}; \theta_1, \theta_2, \theta_3, \theta_4)$

$= (2\pi\theta_3)^{-\frac{n}{2}} \exp\{-\frac{1}{2\theta_3}\sum_{i=1}^{n}(x_i - \theta_1)^2\} (2\pi\theta_4)^{-\frac{m}{2}} \exp\{-\frac{1}{2\theta_4}\sum_{j=1}^{m}(y_j - \theta_2)^2\}$

The likelihood ratio $\lambda(\underset{\sim}{x}, \underset{\sim}{y}) = \dfrac{\sup\{L(\theta_1, \theta_2, \theta_3, \theta_4, \underset{\sim}{x}, \underset{\sim}{y}) : (\theta_1, \theta_2, \theta_3, \theta_4) \in \Theta_0\}}{\sup\{L(\theta_1, \theta_2, \theta_3, \theta_4, \underset{\sim}{x}, \underset{\sim}{y}) : (\theta_1, \theta_2, \theta_3, \theta_4) \in \Theta\}}$

where $\Theta_0 = \{(\theta_1, \theta_2, \theta_3, \theta_4) : \theta_1 = \theta_2 = \mu, \ \theta_3 = \theta_4 = \sigma^2, \mu \in \mathbb{R}, \sigma^2 \in \mathbb{R}^+\}$
and $\Theta = \{(\theta_1, \theta_2, \theta_3, \theta_4) : \theta_1 \in \mathbb{R}, \theta_2 \in \mathbb{R}, \theta_3 \in \mathbb{R}^+, \theta_4 \in \mathbb{R}^+\}$.

Denominator:

$\log L = -\frac{n}{2}\log 2\pi - \frac{n}{2}\log\theta_3 - \frac{1}{2\theta_3}\sum_{i=1}^{n}(x_i - \theta_1)^2 - \frac{m}{2}\log 2\pi - \frac{n}{2}\log\theta_4 - \frac{1}{2\theta_4}\sum_{j=1}^{m}(y_j - \theta_2)^2$

$$\begin{cases} \frac{\partial \log L}{\partial \theta_1} = -\frac{1}{\theta_3}\sum_{i=1}^{n}(x_i - \theta_1)(-1) = \frac{1}{\theta_3}\sum_{i=1}^{n}(x_i - \theta_1) \\ \frac{\partial \log L}{\partial \theta_2} = -\frac{1}{\theta_4}\sum_{j=1}^{m}(y_j - \theta_2)(-1) = \frac{1}{\theta_4}\sum_{j=1}^{m}(y_j - \theta_2) \\ \frac{\partial \log L}{\partial \theta_3} = -\frac{n}{2\theta_3} + \frac{1}{2\theta_3^2}\sum_{i=1}^{n}(x_i - \theta_1)^2 \\ \frac{\partial \log L}{\partial \theta_4} = -\frac{n}{2\theta_4} + \frac{1}{2\theta_4^2}\sum_{j=1}^{m}(y_j - \theta_2)^2 \end{cases}$$

$$\begin{cases} \frac{\partial \log L}{\partial \theta_1} = 0 \\ \frac{\partial \log L}{\partial \theta_2} = 0 \\ \frac{\partial \log L}{\partial \theta_3} = 0 \\ \frac{\partial \log L}{\partial \theta_4} = 0 \end{cases} \Rightarrow \begin{cases} \hat{\theta}_1 = \frac{1}{n}\sum_{i=1}^{n} X_i = \bar{X} \\ \hat{\theta}_2 = \frac{1}{m}\sum_{j=1}^{m} Y_j = \bar{Y} \\ \hat{\theta}_3 = \frac{1}{n}\sum_{i=1}^{n}(X_i - \hat{\theta}_1)^2 = \frac{1}{n}\sum_{i=1}^{n}(X_i - \bar{X})^2 \\ \hat{\theta}_4 = \frac{1}{m}\sum_{j=1}^{m}(Y_j - \hat{\theta}_2)^2 = \frac{1}{m}\sum_{j=1}^{m}(Y_j - \bar{Y})^2 \end{cases}$$

$\therefore$ the denominator $= [(2\pi)\frac{1}{n}\sum_{i=1}^{n}(x_i - \bar{x})^2]^{-\frac{n}{2}} \exp\{-\frac{n}{2}\} [(2\pi)\frac{1}{m}\sum_{j=1}^{m}(y_j - \bar{y})^2]^{-\frac{m}{2}} \exp\{-\frac{m}{2}\}$

Numerator: $\theta_1 = \theta_2 = \mu$, $\theta_3 = \theta_4 = \sigma^2$

$L_0 = (2\pi\sigma^2)^{-\frac{n}{2}} \exp\{-\frac{1}{2\sigma^2}\sum_{i=1}^{n}(x_i - \mu)^2\} (2\pi\sigma^2)^{-\frac{m}{2}} \exp\{-\frac{1}{2\sigma^2}\sum_{j=1}^{m}(y_j - \mu)^2\}$

$= (2\pi\sigma^2)^{-\frac{1}{2}(m+n)} \exp\{-\frac{1}{2\sigma^2}[\sum_{i=1}^{n}(x_i - \mu)^2 + \sum_{j=1}^{m}(y_j - \mu)^2]\}$

$\log L_0 = -\frac{1}{2}(m+n)\log 2\pi - \frac{1}{2}(m+n)\log\sigma^2 - \frac{1}{2\sigma^2}[\sum_{i=1}^{n}(x_i - \mu)^2 + \sum_{j=1}^{m}(y_j - \mu)^2]$

2 (a) (cont.)

$$\begin{cases} \frac{\partial \log L_0}{\partial \mu} = -\frac{1}{\sigma^2}\left[\sum_{i=1}^{n}(x_i-\mu)(-1) + \sum_{j=1}^{m}(y_j-\mu)(-1)\right] \\ \frac{\partial \log L_0}{\partial \sigma^2} = -\frac{1}{2\sigma^2}(m+n) + \frac{1}{2\sigma^4}\left[\sum_{i=1}^{n}(x_i-\mu)^2 + \sum_{j=1}^{m}(y_j-\mu)^2\right] \end{cases}$$

$$\begin{cases} \frac{\partial \log L_0}{\partial \mu} = 0 \\ \frac{\partial \log L_0}{\partial \sigma^2} = 0 \end{cases} \Rightarrow \begin{cases} \sum_{i=1}^{n} x_i - n\mu + \sum_{j=1}^{m} y_j - m\mu = 0 \\ \hat{\sigma}^2 = \frac{1}{m+n}\left[\sum_{i=1}^{n}(X_i-\hat{\mu})^2 + \sum_{j=1}^{m}(Y_j-\hat{\mu})^2\right] \end{cases}$$

$$\Rightarrow \begin{cases} \hat{\mu} = \frac{1}{m+n}\left(\sum_{i=1}^{n} X_i + \sum_{j=1}^{m} Y_j\right) = \frac{1}{m+n}(n\bar{X} + m\bar{Y}) = U \\ \hat{\sigma}^2 = \frac{1}{m+n}\left[\sum_{i=1}^{n}(X_i-U)^2 + \sum_{j=1}^{m}(Y_j-U)^2\right] \end{cases}$$

where $U = (n\bar{X} + m\bar{Y})/(n+m)$.

$\therefore$ the numerator $= \left\{(2\pi)\left(\frac{1}{m+n}\right)\left[\sum_{i=1}^{n}(x_i-u)^2 + \sum_{j=1}^{m}(y_j-u)^2\right]\right\}^{-\frac{1}{2}(m+n)} \exp\left\{-\frac{m+n}{2}\right\}$

$$\therefore \lambda(\underset{\sim}{x}, \underset{\sim}{y}) = \frac{\left\{(2\pi)\left(\frac{1}{m+n}\right)\left[\sum_{i=1}^{n}(x_i-u)^2 + \sum_{j=1}^{m}(y_j-u)^2\right]\right\}^{-\frac{1}{2}(m+n)} \exp\left\{-\frac{m+n}{2}\right\}}{\left[(2\pi)\left(\frac{1}{n}\right)\sum_{i=1}^{n}(x_i-\bar{x})^2\right]^{-\frac{n}{2}} \exp\left\{-\frac{n}{2}\right\}\left[(2\pi)\frac{1}{m}\sum_{j=1}^{m}(y_j-\bar{y})^2\right]^{-\frac{m}{2}} \exp\left\{-\frac{m}{2}\right\}}$$

$$= \frac{\left[\sum_{i=1}^{n}(x_i-\bar{x})^2/n\right]^{n/2}\left[\sum_{j=1}^{m}(y_j-\bar{y})^2/m\right]^{m/2}}{\left\{\left[\sum_{i=1}^{n}(x_i-u)^2 + \sum_{j=1}^{m}(y_j-u)^2\right]/(m+n)\right\}^{(m+n)/2}}$$

where $u = (n\bar{x} + m\bar{y})/(n+m)$

(b) $\begin{cases} H_0: \theta_3 = \theta_4, \ \theta_1 \text{ and } \theta_2 \text{ unspecified} \\ H_1: \theta_3 \neq \theta_4, \ \theta_1 \text{ and } \theta_2 \text{ unspecified} \end{cases}$

$L(\theta_1, \theta_2, \theta_3, \theta_4, \underset{\sim}{x}, \underset{\sim}{y}) = f(\underset{\sim}{x}, \underset{\sim}{y}; \theta_1, \theta_2, \theta_3, \theta_4)$

$= (2\pi\theta_3)^{-\frac{n}{2}} \exp\left\{-\frac{1}{2\theta_3}\sum_{i=1}^{n}(x_i-\theta_1)^2\right\}(2\pi\theta_4)^{-\frac{m}{2}} \exp\left\{-\frac{1}{2\theta_4}\sum_{j=1}^{m}(y_j-\theta_2)^2\right\}$

~~The likelihood ratio~~ $\lambda(\underset{\sim}{x}, \underset{\sim}{y}) = \dfrac{\sup\{L(\theta_1, \theta_2, \theta_3, \theta_4, \underset{\sim}{x}, \underset{\sim}{y}) : (\theta_1, \theta_2, \theta_3, \theta_4) \in \Theta_0\}}{\sup\{L(\theta_1, \theta_2, \theta_3, \theta_4, \underset{\sim}{x}, \underset{\sim}{y}) : (\theta_1, \theta_2, \theta_3, \theta_4) \in \Theta\}}$

where $\Theta_0 = \{(\theta_1, \theta_2, \theta_3, \theta_4) : \theta_3 = \theta_4 = \sigma^2, \ \theta_1 \in \mathbb{R}, \ \theta_2 \in \mathbb{R}, \ \sigma^2 \in \mathbb{R}^+\}$

and $\Theta = \{(\theta_1, \theta_2, \theta_3, \theta_4) : \theta_1 \in \mathbb{R}, \ \theta_2 \in \mathbb{R}, \ \theta_3 \in \mathbb{R}^+, \ \theta_4 \in \mathbb{R}^+\}$

The denominator is same as the denominator of part (a).

Numerator: $\theta_3 = \theta_4 = \sigma^2$

$L_0 = (2\pi\sigma^2)^{-\frac{n}{2}} \exp\left\{-\frac{1}{2\sigma^2}\sum_{i=1}^{n}(x_i-\theta_1)^2\right\}(2\pi\sigma^2)^{-\frac{m}{2}} \exp\left\{-\frac{1}{2\sigma^2}\sum_{j=1}^{m}(y_j-\theta_2)^2\right\}$

$\log L_0 = -\frac{m+n}{2}\log(2\pi) - \frac{m+n}{2}\log\sigma^2 - \frac{1}{2\sigma^2}\left[\sum_{i=1}^{n}(x_i-\theta_1)^2 + \sum_{j=1}^{m}(y_j-\theta_2)^2\right]$

$$\begin{cases} \frac{\partial \log L_0}{\partial \theta_1} = -\frac{1}{\sigma^2}\sum_{i=1}^{n}(x_i-\theta_1)(-1) \\ \frac{\partial \log L_0}{\partial \theta_2} = -\frac{1}{\sigma^2}\sum_{j=1}^{m}(y_j-\theta_2)(-1) \\ \frac{\partial \log L_0}{\partial \sigma^2} = -\frac{m+n}{2\sigma^2} + \frac{1}{2\sigma^4}\left[\sum_{i=1}^{n}(x_i-\theta_1)^2 + \sum_{j=1}^{m}(y_j-\theta_2)^2\right] \end{cases}$$

$$\begin{cases} \frac{\partial \log L_0}{\partial \theta_1} = 0 \\ \frac{\partial \log L_0}{\partial \theta_2} = 0 \\ \frac{\partial \log L_0}{\partial \sigma^2} = 0 \end{cases} \Rightarrow \begin{cases} \hat{\theta}_1 = \frac{1}{n}\sum_{i=1}^{n} X_i = \bar{X} \\ \hat{\theta}_2 = \frac{1}{m}\sum_{j=1}^{m} Y_j = \bar{Y} \\ \hat{\sigma}^2 = \frac{1}{m+n}\left[\sum_{i=1}^{n}(X_i-\bar{X})^2 + \sum_{j=1}^{m}(Y_j-\bar{Y})^2\right] \end{cases}$$

2 (b) (Cont.)

the numerator $= \left\{(2\pi)\left(\frac{1}{m+n}\right)\left[\sum_{i=1}^{n}(x_i-\bar{x})^2+\sum_{j=1}^{m}(y_j-\bar{y})^2\right]\right\}^{-\frac{m+n}{2}} \exp\left\{-\frac{m+n}{2}\right\}$

$$\therefore \lambda(\underset{\sim}{x},\underset{\sim}{y}) = \frac{\left\{(2\pi)\left(\frac{1}{m+n}\right)\left[\sum_{i=1}^{n}(x_i-\bar{x})^2+\sum_{j=1}^{m}(y_j-\bar{y})^2\right]\right\}^{-\frac{m+n}{2}} \exp\left\{-\frac{m+n}{2}\right\}}{\left[2\pi\left(\frac{1}{n}\right)\sum_{i=1}^{n}(x_i-\bar{x})^2\right]^{-\frac{n}{2}}\exp\left\{-\frac{n}{2}\right\}\left[(2\pi)\left(\frac{1}{m}\right)\sum_{j=1}^{m}(y_j-\bar{y})^2\right]^{-\frac{m}{2}}\exp\left\{-\frac{m}{2}\right\}}$$

$$= \frac{\left\{\frac{1}{m+n}\left[\sum_{i=1}^{n}(x_i-\bar{x})^2+\sum_{j=1}^{m}(y_j-\bar{y})^2\right]\right\}^{-\frac{m+n}{2}}}{\left[\frac{1}{n}\sum_{i=1}^{n}(x_i-\bar{x})^2\right]^{-\frac{n}{2}}\left[\frac{1}{m}\sum_{j=1}^{m}(y_j-\bar{y})^2\right]^{-\frac{m}{2}}}$$

H_0 is rejected when $\lambda(\underset{\sim}{x},\underset{\sim}{y}) \le k$

i.e. $$\frac{\left[\frac{1}{n}\sum_{i=1}^{n}(x_i-\bar{x})^2\right]^{\frac{n}{2}}\left[\frac{1}{m}\sum_{j=1}^{m}(y_j-\bar{y})^2\right]^{\frac{m}{2}}}{\left\{\frac{1}{m+n}\left[\sum_{i=1}^{n}(x_i-\bar{x})^2+\sum_{j=1}^{m}(y_j-\bar{y})^2\right]\right\}^{\frac{m+n}{2}}} \le k$$

iff $$\frac{\left[\sum_{i=1}^{n}(x_i-\bar{x})^2\right]^{\frac{n}{2}}\left[\sum_{j=1}^{m}(y_j-\bar{y})^2\right]^{\frac{m}{2}}}{\left[\sum_{i=1}^{n}(x_i-\bar{x})^2+\sum_{j=1}^{m}(y_j-\bar{y})^2\right]^{\frac{m+n}{2}}} \le k'$$

iff $$\frac{\left[\sum_{i=1}^{n}(x_i-\bar{x})^2 / \sum_{j=1}^{m}(y_j-\bar{y})^2\right]^{\frac{n}{2}}}{\left[1+\sum_{i=1}^{n}(x_i-\bar{x})^2 / \sum_{j=1}^{m}(y_j-\bar{y})^2\right]^{\frac{m+n}{2}}} \le k' \qquad (*)$$

iff $\sum_{i=1}^{n}(x_i-\bar{x})^2 / \sum_{j=1}^{m}(y_j-\bar{y})^2 \le k_1$ or $\sum_{i=1}^{n}(x_i-\bar{x})^2 / \sum_{j=1}^{m}(y_j-\bar{y})^2 \ge k_2$

iff $F \le K_1$ or $F \ge K_2$ where $F = \dfrac{\sum_{i=1}^{n}(x_i-\bar{x})^2/(n-1)}{\sum_{j=1}^{m}(y_j-\bar{y})^2/(m-1)}$

(A closer look on $(*)$:

consider $g(z) = \dfrac{z^{\frac{n}{2}}}{(1+z)^{\frac{m+n}{2}}}$ for $z>0$

$$= \frac{1}{(1+z)^{\frac{m}{2}}\left(1+\frac{1}{z}\right)^{\frac{n}{2}}}$$

$\therefore$ when z is small, $\left(1+\frac{1}{z}\right)^{\frac{n}{2}}$ is large and $(1+z)^{\frac{m}{2}}$ close to 1.

but when z is large, $(1+z)^{\frac{m}{2}}$ is large and $(1+z)^{\frac{n}{2}}$ close to 1.

Thus, in both cases, $g(z)$ will be large.)

Note that $F \sim F$-distribution with d.f. $(n-1)$ and $(m-1)$.

The critical region at significance level α is:

$C_1 = \{(\underset{\sim}{x},\underset{\sim}{y}) : F \le F_{1-\frac{\alpha}{2}}(n-1, m-1)$ or $F \ge F_{\frac{\alpha}{2}}(n-1, m-1)\}$

4. (a) Let a random sample of size n be taken from a distribution of the discrete type with p.d.f. $f(x;\theta) = 1/\theta,\ x = 1, 2, ..., \theta$, zero elsewhere, where θ is an unknown positive integer.

(i) Show that the largest item, say Y, of the sample is a complete sufficient statistic for θ.

(ii) Prove that

$$[Y^{n+1} - (Y-1)^{n+1}]/[Y^n - (Y-1)^n]$$

is the best statistic for θ.

(i) $P_r(Y \le y) = P_r(\max\{X_i\} \le y)$

$= P_r(X_1 \le y, \cdots, X_n \le y)$

$= \left(\frac{y}{\theta}\right)^n$

$\therefore f_Y(y) = P_r(Y = y)$

$= P_r(Y \le y) - P_r(Y \le y-1)$

$= \left(\frac{1}{\theta}\right)^n (y^n - (y-1)^n)$

$f_{X_1,\cdots,X_n|Y}(x_1,\cdots,x_n|y) = \dfrac{f_{X_1,\cdots,X_n,Y}(x_1,\cdots,x_n,y)}{f_Y(y)}$

$= \dfrac{f_{X_1,\cdots,X_n}(x_1,\cdots,x_n)}{f_Y(y)}$

$= \dfrac{\prod_{i=1}^{n}\left(\frac{1}{\theta}\right)}{\left(\frac{1}{\theta}\right)^n (y^n - (y-1)^n)}$

$= \dfrac{1}{y^n - (y-1)^n}$ which does not depend on θ

$\therefore Y = \max(X_1, \cdots, X_n)$ is suff. for θ 3

$E(Z(Y)) = \sum_{y=1}^{\theta} Z(y) \left(\frac{1}{\theta}\right)^n (y^n - (y-1)^n) = 0$

$\Rightarrow \sum_{y=1}^{\theta} Z(y) \left(\frac{1}{\theta}\right)^n (y^n - (y-1)^n) = 0$

$\Rightarrow \sum_{y=1}^{\theta} Z(y) (y^n - (y-1)^n) = 0$

$\Rightarrow Z(y) = 0$

$\therefore Y$ is complete 2

(ii) $E\left(\dfrac{Y^{n+1} - (Y-1)^{n+1}}{Y^n - (Y-1)^n}\right)$

$= \sum_{y=1}^{\theta} \left(\dfrac{y^{n+1} - (y-1)^{n+1}}{y^n - (y-1)^n}\right) P_r(Y = y)$

$= \sum_{y=1}^{\theta} \dfrac{y^{n+1} - (y-1)^{n+1}}{y^n - (y-1)^n} \left(\frac{1}{\theta}\right)^n (y^n - (y-1)^n)$

$= \left(\frac{1}{\theta}\right)^n \sum_{y=1}^{\theta} (y^{n+1} - (y-1)^{n+1})$

$= \left(\frac{1}{\theta}\right)^n \theta^{n+1} = \theta$

$\Rightarrow$ it is unbiased for θ 5

$\Rightarrow$ it is the best statistic for θ

MATH 243 Final Examination – Fall 98/99

5. (a) $\log f(x;\lambda) = \log\lambda - \lambda x$

$\frac{\partial}{\partial\lambda}\log f(x;\lambda) = \frac{1}{\lambda} - x$

$\frac{\partial^2}{\partial\lambda^2}\log f(x;\lambda) = -\frac{1}{\lambda^2}$

$E\left(\frac{\partial^2}{\partial\lambda^2}\log f(x;\lambda)\right) = -\frac{1}{\lambda^2}$ $\Rightarrow$ C R. Lower bound $= \frac{\lambda^2}{n}$

(b) $E(e^{t\bar{X}}) = \left(\frac{1}{1-\frac{t}{\lambda n}}\right)^n$

which is m.g.f. of Gamma dist with p.d.f. $f_{\bar{X}}(y) = \frac{1}{\Gamma(n)}(\lambda n)^n y^{n-1} e^{-\lambda n y}$

$\Rightarrow E\left(\frac{1}{\bar{X}}\right) = \int_0^\infty \frac{1}{y} f_{\bar{X}}(y)\,dy$

$= \int_0^\infty \frac{1}{y}\frac{1}{\Gamma(n)}(\lambda n)^n y^{n-1} e^{-\lambda n y}\,dy$

$= (\lambda n)\left(\frac{1}{\Gamma(n)}\right)\int_0^\infty (\lambda n)^{n-1} y^{n-2} e^{-\lambda n y}\,dy$

$= (\lambda n)\left(\frac{1}{\Gamma(n)}\right)\Gamma(n-1) = \frac{\lambda n}{n-1} \neq \lambda$

But $E(T) = E\left(\frac{n-1}{\sum_{i=1}^n X_i}\right)$

$= \frac{n-1}{n} E\left(\frac{1}{\bar{X}}\right)$

$= \lambda$

$\therefore T = \frac{n-1}{\sum_{i=1}^n X_i}$ is an unbiased estimator for λ

(c) $E\left(\frac{1}{\bar{X}^2}\right) = \int_0^\infty \frac{1}{y^2} f_{\bar{X}}(y)\,dy$

$= \int_0^\infty \frac{1}{y^2}\frac{1}{\Gamma(n)}(\lambda n)^n y^{n-1} e^{-\lambda n y}\,dy$

$= \frac{\lambda^2 n^2}{(n-1)(n-2)}$

$E(T^2) = \frac{(n-1)\lambda^2}{n-2}$

$Var(T) = E(T^2) - \lambda^2 = \frac{\lambda^2}{n-2} >$ C R. lower bound