THE UNIVERSITY OF HONG KONG

DEPARTMENT OF STATISTICS AND ACTUARIAL SCIENCE

STAT1302/STAT2602 PROBABILITY AND STATISTICS II

May 14, 2015 **Time: 2:30 p.m. - 4.30 p.m.**

Only approved calculators as announced by the Examinations Secretary can be used in this examination. It is candidates' responsibility to ensure that their calculator operates satisfactorily, and candidates must record the name and type of the calculator used on the front page of the examination script.

Answer ALL SEVEN questions. Marks are shown in square brackets.

1. The quality control engineer at a bottling plant keeps track of the number of bottles that are either overfilled or under-filled. She believes that the percentage of such bottles is less than 4%. To examine the issue she randomly selects 400 bottles and counts the number of overfilled and under-filled bottles. Assume that the number of under-filled bottles is 3 and that of overfilled bottles is 9. Can the engineer conclude at the 5% significance level that she is right? Check whether your method used is suitable.

[Total: 9 marks]

2. A survey was conducted to investigate interest of middle-aged adults in physical fitness programs in Rhode Island, Colorado, California and Florida. The data were recorded in the following table.

Participation	Rhode Island	Colorado	California	Florida
Yes	46	63	108	121
No	149	178	192	179

Do the data indicate differences among the rates of adult participation in physical fitness programmes from one state to another? Test at $\alpha = 0.01$.

[Total: 12 marks]

3. Suppose a population is uniformly distributed over the interval $[\mu - \theta, \mu + \theta]$, and X_1, $X_2, \ldots, X_n$ constitute a random sample of size n from the population.

(a) Find estimators of μ and θ by the method of moments. **[7 marks]**

(b) Find estimators of μ and θ by the method of maximum likelihood. **[7 marks]**

[Total: 14 marks]

4. Suppose a population has a probability function $f(x; \theta) = \theta(1-\theta)^{x-1}$ for $x = 1, 2, \ldots$, where θ is a constant in $(0, 1)$. Suppose $X_1, X_2, \ldots, X_n$ constitute a random sample of size n from the population. Find the generalised likelihood ratio for testing

$$H_0: \theta = \theta_0 \quad \text{against} \quad H_1: \theta \neq \theta_0$$

where θ_0 is a constant in $(0, 1)$.

[Total: 11 marks]

5. A professor took two samples, one of 12 males and another of 18 females, from students who were enrolled in a statistics course. The mean score of male students in a mid-term test was 76.2 with a standard deviation of 7.3, and the mean score of female students was 78.5 with a standard deviation of 6.7. Assume that the scores of all male and all female students are normally distributed with equal standard deviations. Using the 5% significance level, can we conclude that the mean score in statistics for all male students is lower than that for all female students?

[Total: 9 marks]

6. A manufacturer of light bulbs advertises that, on average, its long-life bulb will last more than 5,000 hours. Assume that the lifetime of a randomly selected bulb of this type has a standard deviation of 400 hours. A statistician took a random sample of 100 bulbs and measured, for each bulb, the amount of time until the bulb burned out. The sample mean is 5,065 hours.

   (a) Using p-value, conclude at the 5% level of significance whether the manufacturer's claim is true. **[7 marks]**

   (b) Find a 95% confidence interval (correct to integers) for the expected number of hours a randomly selected bulb of this type will last. **[4 marks]**

[Total: 11 marks]

7. Suppose a population has a density function $\frac{\delta}{2}|x|^{\delta-1}\,1_{[-1,\,1]}(x)$, where δ is a positive constant. Suppose $X_1, X_2, \ldots, X_n$ constitute a random sample of size n from the population.

   (a) Find an estimator of δ by the method of moments. **[6 marks]**

   (b) Find an estimator of δ by the method of maximum likelihood. **[7 marks]**

   (c) Find a sufficient statistic for δ. **[4 marks]**

   (d) Find the Cramér-Rao lower bound for estimating δ. **[7 marks]**

   (e) Are $\frac{1}{n}\sum_{i=1}^{n}|X_i|$ and $\prod_{i=1}^{n}|X_i|$ asymptotically unbiased estimators of δ? **[10 marks]**

[Total: 34 marks]

Formulae

$\mathrm{M}_X(t) = \mathrm{E}(e^{tX})$, $\left.\left(\dfrac{\mathrm{d}^r}{\mathrm{d}t^r}\mathrm{M}_X(t)\right)\right|_{t=0} = \mu'_r$.

Normal population $\Rightarrow$ $\dfrac{(n-1)S^2}{\sigma^2} \sim \chi^2(n-1) = \mathrm{G}\left(\dfrac{n-1}{2}, \dfrac{1}{2}\right)$, $\dfrac{\overline{X}-\mu}{S/\sqrt{n}} \sim \mathrm{t}(n-1)$.

$\mathrm{b}(\hat{\Theta}) = \mathrm{E}(\hat{\Theta}) - \theta$. $\mathrm{E}([\hat{\Theta} - \theta]^2) = \mathrm{Var}(\hat{\Theta}) + [\mathrm{b}(\hat{\Theta})]^2$.

$\mathrm{I}(\theta) = \mathrm{E}\left(\left[\dfrac{\partial \ln \mathrm{f}(X;\theta)}{\partial\theta}\right]^2\right) = \mathrm{E}\left(-\dfrac{\partial^2 \ln \mathrm{f}(X;\theta)}{\partial\theta^2}\right)$. $\mathrm{Var}(\hat{\Theta}) \geq \dfrac{1}{n\mathrm{I}(\theta)}$.

$\mathrm{P}(|X-\mu| \geq c) \leq \dfrac{\sigma^2}{c^2}$. Consistency: $\lim_{n\to\infty} \mathrm{P}(|\hat{\Theta} - \theta| \geq c) = 0$.

$\dfrac{\overline{X}-\mu_0}{\sigma/\sqrt{n}}$; $\dfrac{\overline{X}-\mu_0}{S/\sqrt{n}}$; $\dfrac{(n-1)S^2}{\sigma_0^2}$; $\dfrac{\hat{p}-p_0}{\sqrt{p_0(1-p_0)/n}}$; $\dfrac{\overline{X}-\overline{Y}-\delta}{\sqrt{\dfrac{\sigma_1^2}{n_1}+\dfrac{\sigma_2^2}{n_2}}}$; $\dfrac{\overline{X}-\overline{Y}-\delta}{S_p\sqrt{\dfrac{1}{n_1}+\dfrac{1}{n_2}}}$, $S_p^2 =$

$\dfrac{(n_1-1)S_1^2+(n_2-1)S_2^2}{n_1+n_2-2}$; $\dfrac{\overline{X}-\overline{Y}-\delta}{\sqrt{\dfrac{S_1^2}{n_1}+\dfrac{S_2^2}{n_2}}}$, $df \approx \left(\dfrac{s_1^2}{n_1}+\dfrac{s_2^2}{n_2}\right)^2 \Big/ \left[\dfrac{(s_1^2/n_1)^2}{n_1-1}+\dfrac{(s_2^2/n_2)^2}{n_2-1}\right]$; $\dfrac{S_1^2}{S_2^2}$, $f_{1-\alpha,m,n} =$

$\dfrac{1}{f_{\alpha,n,m}}$; $\dfrac{\hat{p}_1-\hat{p}_2}{\sqrt{\hat{p}(1-\hat{p})\left(\dfrac{1}{n_1}+\dfrac{1}{n_2}\right)}}$, $\hat{p} = \dfrac{n_1\hat{p}_1+n_2\hat{p}_2}{n_1+n_2}$; $-2\ln\Lambda \approx \sum_i \dfrac{(O_i-E_i)^2}{E_i} = \sum_i \dfrac{O_i^2}{E_i} - n$.

$\bar{x} \pm z_{\alpha/2}\dfrac{\sigma}{\sqrt{n}}$; $\bar{x} \pm t_{\alpha/2,n-1}\dfrac{s}{\sqrt{n}}$; $\left(\dfrac{(n-1)s^2}{\chi^2_{\alpha/2,n-1}}, \dfrac{(n-1)s^2}{\chi^2_{1-\alpha/2,n-1}}\right)$; $\hat{p} \pm z_{\alpha/2}\sqrt{\dfrac{\hat{p}(1-\hat{p})}{n}}$;

$\bar{x}_1 - \bar{x}_2 \pm z_{\alpha/2}\sqrt{\dfrac{\sigma_1^2}{n_1}+\dfrac{\sigma_2^2}{n_2}}$; $\bar{x}_1 - \bar{x}_2 \pm t_{\alpha/2,n_1+n_2-2}\, s_p\sqrt{\dfrac{1}{n_1}+\dfrac{1}{n_2}}$; $\bar{x}_1 - \bar{x}_2 \pm t_{\alpha/2,df}\sqrt{\dfrac{s_1^2}{n_1}+\dfrac{s_2^2}{n_2}}$;

$\left(\dfrac{s_1^2}{s_2^2}\cdot\dfrac{1}{f_{\alpha/2,n_1-1,n_2-1}}, \dfrac{s_1^2}{s_2^2}\cdot f_{\alpha/2,n_2-1,n_1-1}\right)$; $\hat{p}_1 - \hat{p}_2 \pm z_{\alpha/2}\sqrt{\dfrac{\hat{p}_1(1-\hat{p}_1)}{n_1}+\dfrac{\hat{p}_2(1-\hat{p}_2)}{n_2}}$.

$U \sim \chi^2(m)$ independent of $V \sim \chi^2(n)$ $\Rightarrow$ $\dfrac{U/m}{V/n} \sim \mathrm{F}(m, n)$.

************ **END OF PAPER** ************