THE UNIVERSITY OF HONG KONG

DEPARTMENT OF STATISTICS AND ACTUARIAL SCIENCE

STAT2602 PROBABILITY AND STATISTICS II

May 21, 2020 **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 FIVE questions. Marks are shown in square brackets.

1. Suppose that a random variable X has a (probability) density function given by

$$f(x) = \begin{cases} 2e^{-2x}, & \text{for } x > 0; \\ 0, & \text{otherwise,} \end{cases}$$

(i) Calculate the moment generating function of X. **[5 marks]**

(ii) Calculate $\mathrm{E}(X)$ and $\mathrm{E}(X^2)$. **[5 marks]**

(iii) Calculate $\mathrm{E}(e^{X/2})$, $\mathrm{E}(e^{X})$ and $\mathrm{E}(e^{3X})$, if they exist. **[5 marks]**

(iv) Based on an independent random sample $\mathbf{X} = \{X_1, X_2, ..., X_n\}$ from the distribution of X, provide a consistent estimator for $\theta = \mathrm{E}(e^{\sin(X)})$, where $\sin(\cdot)$ is the sine function. **[5 marks]**

[Total: 20 marks]

2. Let $X_1, X_2, \cdots, X_n$ be an independent random sample from $N\left(\frac{n}{n+1}\mu, \sigma^2\right)$.

(i) Show that $\overline{X} = \frac{1}{n}\sum_{i=1}^{n} X_i$ is an asymptotically unbiased estimator of μ.

[5 marks]

(ii) Show that

$$\mathrm{Var}(\overline{X}) = \frac{\sigma^2}{n}.$$

Hence show that $\overline{X}$ is a consistent estimator of μ. **[5 marks]**

(iii) If $\frac{n}{n+1}\mu$ is μ_0 (a constant), show that $S^2 = \frac{1}{n}\sum_{i=1}^{n}(X_i - \overline{X})^2$ is a consistent estimator of σ^2 by using the weak law of large numbers. **[5 marks]**

(iv) If $\mathrm{E}(X^4) = s^4 + 6s^2\sigma^2 + 3\sigma^2$ when $X \sim N(s, \sigma^2)$, show that $S^2 = \frac{1}{n}\sum_{i=1}^{n}(X_i - \overline{X})^2$ is a consistent estimator of σ^2. **[5 marks]**

[Total: 20 marks]

3. Let $X_1, X_2, \cdots, X_n$ be an independent random sample from a Laplace(θ) distribution, which has the density given by

$$f(x;\theta) = \frac{1}{2\theta}\exp\left(-\frac{|x|}{\theta}\right) \quad \text{for} \quad -\infty < x < \infty.$$

Here, $\theta > 0$ is a parameter. Note that if $X \sim$ Laplace(θ), $\mathrm{E}X = 0$, $\mathrm{E}|X| = \theta$ and $\mathrm{E}(X^2) = 2\theta^2$.

(i) Find the MLE of θ on $\Omega_1 = \{\theta : \theta > 0\}$. **[5 marks]**

(ii) Calculate the CRLB with respect to θ. **[5 marks]**

(iii) Is the MLE in (i) the UMVUE of θ? **[5 marks]**

(iv) Find the rejection region of the most powerful test for hypotheses:

$$H_0 : \theta = 1 \quad \text{versus} \quad H_1 : \theta = 2.$$

[5 marks]

[Total: 20 marks]

4. Let $X_1, X_2, \cdots, X_n$ be an independent random sample from $N(\mu_X, \sigma_X^2)$, and $Y_1, Y_2, \cdots, Y_m$ be an independent random sample from $N(\mu_Y, \sigma_Y^2)$. Assume that these two random samples are independent.

(i) Show that

$$U = \frac{(\overline{X} - \overline{Y}) - (\mu_X - \mu_Y)}{\sqrt{\frac{\sigma_X^2}{n} + \frac{\sigma_Y^2}{m}}} \sim N(0, 1),$$

by using the moment generating function. **[5 marks]**

(ii) If $V_1 = \dfrac{nS_X^2}{\sigma_X^2} \sim \chi^2_{n-1}$, $V_2 = \dfrac{mS_Y^2}{\sigma_Y^2} \sim \chi^2_{m-1}$, and V_1 and V_2 are independent, show that

$$V = V_1 + V_2 \sim \chi^2_{n+m-2},$$

by using the moment generating function. **[5 marks]**

(iii) If U and V are independent, $\sigma_X^2/\sigma_Y^2 = d$, and d is a known constant, construct a random variable W which has a t distribution. **[3 marks]**

(iv) Using W, construct a $1 - \alpha$ confidence interval for $\mu_X - \mu_Y$. **[5 marks]**

(v) If $d = 2$, $n = m = 6$, and the observed values of $\overline{X} = 2$, $\overline{Y} = 1$, $S_X^2 = 1/3$, and $S_Y^2 = 2/3$, determine whether the 95% confidence interval estimate contains 0. **[2 marks]**

[Total: 20 marks]

5. Let $X_1, X_2, \cdots, X_n$ be an independent random sample from a Bernoulli distribution with parameter p.

   (i) Show that the most powerful test for hypotheses:

   $$H_0 : p = \frac{1}{2} \quad \text{versus} \quad H_1 : p = \frac{1}{3}$$

   has the rejection region $\left\{ \sum_{i=1}^{n} X_i \leq c \right\}$ for some constant c. **[5 marks]**

   (ii) Use the central limit theorem to find n and c so that this test approximately has the size 0.1 under H_0 and the power 0.8 under H_1. **[15 marks]**

[Total: 20 marks]

A LIST OF STATISTICAL FORMULAE

1. $M_X(t) = \mathrm{E}(e^{tX})$. $\quad \mathrm{E}(X^r) = \left(\dfrac{\mathrm{d}^r}{\mathrm{d}t^r} M_X(t)\right)\Big|_{t=0}$.

2. $\overline{X} = \dfrac{1}{n}\sum_{i=1}^{n} X_i$. $\quad S^2 = \dfrac{1}{n}\sum_{i=1}^{n}\left(X_i - \overline{X}\right)^2$.

3. CLT: $\dfrac{\overline{X} - \mu}{\sigma/\sqrt{n}} \approx N(0,1)$ for large n.

4. Normal population $\Longrightarrow \dfrac{nS^2}{\sigma^2} \sim \chi^2_{n-1}$, $\dfrac{\overline{X} - \mu}{S/\sqrt{n-1}} \sim t_{n-1}$.

5. $\mathrm{Bias}(\hat{\theta}) = \mathrm{E}(\hat{\theta}) - \theta$. $\quad \mathrm{E}\left[(\hat{\theta} - \theta)^2\right] = \mathrm{Var}(\hat{\theta}) + \left[\mathrm{Bias}(\hat{\theta})\right]^2$.

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

7. $\overline{x} \pm z_{\alpha/2}\dfrac{\sigma}{\sqrt{n}}$. $\quad \overline{x} \pm t_{\alpha/2,n-1}\dfrac{s}{\sqrt{n-1}}$. $\quad \left(\dfrac{ns^2}{\chi^2_{\alpha/2,n-1}}, \dfrac{ns^2}{\chi^2_{1-\alpha/2,n-1}}\right)$.

8. $\overline{x} - \overline{y} \pm z_{\alpha/2}\sqrt{\dfrac{\sigma_x^2}{n_1} + \dfrac{\sigma_y^2}{n_2}}$. $\quad \overline{x} - \overline{y} \pm t_{\alpha/2,n_x+n_y-2}\sqrt{\dfrac{n_x s_x^2 + n_y s_y^2}{n_x + n_y - 2}\left(\dfrac{1}{n_x} + \dfrac{1}{n_y}\right)}$.
$\left(\dfrac{n_x(n_y-1)s_x^2}{n_y(n_x-1)s_y^2} \cdot \dfrac{1}{F_{\alpha/2,n_x-1,n_y-1}}, \dfrac{n_x(n_y-1)s_x^2}{n_y(n_x-1)s_y^2} \cdot F_{\alpha/2,n_y-1,n_x-1}\right)$.

9. $z_{0.9} = -1.28, z_{0.2} = 0.85, t_{0.025,10} = 2.228$

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