WPI Mathematical Sciences Ph.D. General Comprehensive Exam

MA 541 Probability and Mathematical Statistics II

January, 2020

Note: The problems are arbitrarily ordered, not necessarily according to difficulty. Please show a clear logic of your solution. If you cannot solve a problem perfectly, still show your idea on solving the problem.

1. (20 points) Suppose that $X_1, X_2, \cdots, X_n$ is a sample of i.i.d. observations drawn from a distribution function F. The empirical distribution function is defined as

$$\hat{F}_n(t) = \frac{1}{n}\sum_{i=1}^{n} I(X_i \le t), \quad \forall t \in (-\infty, \infty)$$

where $I(\cdot)$ is the indicator function.

   (a) (6 points) Show that $\hat{F}_n(t)$ is an unbiased estimator of $F(t)$.

   (b) (6 points) Specify the distribution of $\hat{F}_n(t)$.

   (c) (8 points) For any fixed t, show that

$$\sqrt{n}\{\hat{F}_n(t) - F(t)\} \xrightarrow{d} \mathrm{N}(0, \nu(t))$$

   as $n \to \infty$. Determine the value of asymptotic variance $\nu(t)$. Here $\xrightarrow{d}$ represents convergence in distribution.

2. (20 points) Let $X_1, \ldots, X_n \mid \theta \overset{ind}{\sim} \text{Normal}(\theta, \theta^2), \theta > 0$. Let $T = \omega_1 \bar{X} + \omega_2 a^{-1} S$, where $\bar{X}$ and S^2 are respectively the sample mean and the sample variance. Find ω_1 and ω_2 so that T is the minimum variance unbiased estimator of θ. [Hint: $\mathrm{E}(S) = a\theta$, where $a = \frac{\Gamma(n/2)\sqrt{2}}{\Gamma((n-1)/2)\sqrt{n-1}}$, $n \ge 2, a > 1$.]

3. (20 points) Let $X_1, \ldots, X_n$ are iid from Bernoulli(p) where $n \ge 2$ and $0 < p < 1$ is the known parameter.

   (a) (10 points) Derive the uniformly minimum-variance unbiased estimator (UMVUE) of $\tau(p)$, where $\tau(p) = e^2(p(1-p))$.

   (b) (10 points) Find the Cramér–Rao lower bound for estimating $\tau(p) = e^2(p(1-p))$.

4. (20 points) Suppose that $X_1, \ldots, X_n$ are iid samples from a common discrete distribution

$$P(X_k = x) = \frac{\exp(\theta x)}{\exp(-\theta) + 1 + \exp(\theta)}, \quad x = -1, 0, 1,$$

   for $k = 1, \ldots, n$, where θ is a real-valued unknown parameter.

   (a) (10 points) Find a minimal sufficient statistic for θ based on $(X_1, \ldots, X_n)$.

   (b) (10 points) Find the maximum likelihood estimator of θ based on $(X_1, \ldots, X_n)$.

5. (20 points) Let $X_{(1)}, \ldots, X_{(n)}$ denote the order statistics of a sample $X_1, \ldots, X_n \mid \theta \overset{iid}{\sim}$ Uniform$(0, \theta)$. Let $R = X_{(n)} - X_{(1)}$ denote the sample range. In the text book by Casella and Berger, the pdf of R is

$$f(r) = \frac{n(n-1)r^{n-2}(\theta - r)}{\theta^n}, 0 < r < \theta.$$

Show that, using the pivotal method, a $100(1-\alpha)\%$ equal-tailed confidence interval for θ is $(b^{-1}R, a^{-1}R)$, where

$$[n - (n-1)a]a^{n-1} = \alpha/2 \quad \text{and} \quad [n - (n-1)b]b^{n-1} = 1 - \alpha/2.$$

6. (20 points) Let $X_1, X_2, \cdots, X_n$ be a sample of size n with common probability density function $f(x|\theta)$. Denote $\hat{\theta}$ the maximum likelihood estimator of θ. Let $g(\cdot)$ be a continuous function. It is well known that if $f(x|\theta)$ is "well-behaved" (e.g., under some regularity conditions on $f(x|\theta)$), $g(\hat{\theta})$ is a consistent and asymptotically efficient estimator of $g(\theta)$.

   (a) (6 points) Use statistical notations to express the statement that "$g(\hat{\theta})$ is a consistent and asymptotically efficient estimator of $g(\theta)$".

   (b) (14 points) Give an outline of proof for the above statement. You can focus on the key ideas and steps, ignoring secondary details for the regularity conditions on $f(x|\theta)$. [Hint: You can consider applying Taylor expansion to the first derivative of the log likelihood function.]