STAT5005 Qualifying Exam 2019/20

Question 1: Let $X_1, X_2, \ldots$ be a sequence of independent, identically distributed random variables. They may not have finite expectation. Let $S_n = X_1 + \cdots + X_n$. Fix a constant $0 < p < 1$. Prove that $E(|X_1|^p) < \infty$ if and only if as $n \to \infty$,

$$\frac{S_n}{n^{1/p}} \to 0 \quad a.s.$$

Question 2:

(a) If $X_1, X_2, \ldots$ are independent random variables with $\frac{1}{2} = P(X_n = a_n) = 1 - P(X_n = -a_n)$, characterize the sequences $\{a_n, n \geqslant 1\}$ for which $\sum_{i=1}^{\infty} X_i$ converges almost surely.

(b) Suppose $\{X, Y, X_n, Y_n, n \geqslant 1\}$ are random variables defined on the same probability space. Suppose further that $X_n \xrightarrow{d} X$ and $Y_n \xrightarrow{d} Y$ as $n \to \infty$. Is it true that $X_n + Y_n \xrightarrow{d} X + Y$? What if we assume that $\{X, X_1, X_2, \ldots\}$ are independent of $\{Y, Y_1, Y_2, \ldots\}$? Justify your answer.

Question 3: Pólya's urn. A bag contains red and blue balls, with initially r red and b blue where $rb > 0$. A ball is drawn from the bag at random, its colour noted, and then returned to the bag together with a new ball of the same colour. Let R_n be the number of red balls after n such operations.

(a) Show that $\{Y_n = R_n/(n + r + b), n \geqslant 0\}$ is a martingale.

(b) Show that Y_n converges almost surely.

(c) Let T be the number of balls drawn until the first blue ball appears, and suppose that $r = b = 1$. Compute $E[T/(T+2)]$.

(d) Suppose $r = b = 1$. Show that $P(Y_n \geqslant \frac{3}{4}$ for some $n) \leqslant \frac{2}{3}$.

[Note: You may use the following version of Doob's Optional Stopping Theorem: If the sequence $\{Y_n, n \geqslant 0\}$ is a bounded martingale and T is a stopping time, then the expected value of Y_T is Y_0.]

Department of Statistics, The Chinese University of Hong Kong
Qualifying Examination in June 2020

STAT5010 Advanced Statistical Inference

Instruction: All questions are compulsory. Throughout, the abbreviations 'i.i.d', 'pdf/pmf' and 'MLE' stand for 'independent and identically distributed', 'probability density/mass function' and 'maximum likelihood estimator', respectively. A normal distribution in $\mathbb{R}^d$ with mean vector $\boldsymbol{\mu}$ and covariance matrix Σ is denoted by $N_d(\boldsymbol{\mu}, \Sigma)$ and $N(\boldsymbol{\mu}, \Sigma)$ corresponds to the univariate case $d = 1$. For observations X arising from a parametric model $\{f(\cdot, \theta) : \theta \in \Theta\}$, $\Theta \subseteq \mathbb{R}$, the quadratic risk of a decision rule $\delta(X)$ is defined to be $R(\delta, \theta) = E_\theta(\delta(X) - \theta)^2$.

1. Consider an i.i.d. sample $X_1, \ldots, X_n$ arising from the model

$$\{f(\cdot, \theta) : \theta \in \mathbb{R}\}, \quad f(x, \theta) = \frac{1}{2} e^{-|x - \theta|}, \quad x \in \mathbb{R},$$

of *Laplace distributions*. Assuming n to be odd for simplicity, find the MLE. Discuss what happens when n is even. Calculate also the Fisher information.

2. Given $X_1, \ldots, X_n$ i.i.d. random variables such that $E(X_1) = 0$, $E(X^2) \in (0, \infty)$, the *Student t-statistic* is given by

$$t_n = \frac{n^{1/2} \bar{X}_n}{S_n}, \quad \bar{X}_n = \frac{1}{n} \sum_{i=1}^n X_i, \quad S_n^2 = \frac{1}{(n-1)^2} \sum_{i=1}^n (X_i - \bar{X}_n)^2.$$

Show that $t_n \xrightarrow{d} N(0, 1)$ as $n \to \infty$. Assuming now $E(X_1) = \mu \in \mathbb{R}$, deduce an asymptotic level $1 - \alpha$ confidence interval for μ.

3. (a) Consider estimation of $\theta \in \Theta = [0, 1]$ in a $Bin(n, \theta)$ model under quadratic risk.

   i. Find the unique minimax estimator $\tilde{\theta}$ of θ and deduce that the maximum likelihood estimator $\hat{\theta}$ of θ is *not* minimax for fixed sample size $n \in \mathbb{N}$.

   ii. Show, however, that the maximum likelihood estimator dominates $\tilde{\theta}_n$ in the large sample limit by proving that, as $n \to \infty$,

$$\lim_{n \to \infty} \frac{\sup_\theta R(\hat{\theta}_n, \theta)}{\sup_\theta R(\tilde{\theta}_n, \theta)}$$

   and that

$$\lim_{n \to \infty} \frac{R(\hat{\theta}_n, \theta)}{R(\tilde{\theta}_n, \theta)} < 1 \quad \text{for all } \theta \in [0, 1], \theta \neq \frac{1}{2}.$$

   ∞

   (b) Consider $X_1, \ldots, X_n$ i.i.d. from a $N(\theta, 1)$-model where $\theta \in \Theta = [0, \theta)$. Show that the sample mean $\bar{X}_n$ is inadmissible for quadratic risk, but that it is still minimax. What happens if $\Theta = [a, b]$ for some $0 < a < b < \infty$.

4. Let $(X, X_n : n \in \mathbb{N})$ be random vectors in $\mathbb{R}^k$.

   (a) Suppose $E\|X_n - X\| \to 0$ as $n \to \infty$ where $\|\cdot\|$ is the Euclidean norm on $\mathbb{R}^k$. Deduce that $X_n \xrightarrow{P} X$ as $n \to \infty$.

   (b) Show that the converse in (b) is false.

5. Suppose that $X_1, \ldots, X_n$ are independent random variables, and $X_i \sim N(\theta_i, \sigma^2)$ for $i = 1, \ldots, n$, where σ^2 is a known constant. Find a size α UMP test for

$$H_0 : \theta_1 = \ldots = \theta_n = 0 \quad \text{versus} \quad H_1 : \theta_i = \theta_{i0}, \quad i = 1, \ldots, n,$$

where $\theta_{10}, \ldots, \theta_{n0}$ are given constants. You are required to identify the test statistic and the rejection region.