Department of Statistics, The Chinese University of Hong Kong
STAT5010 Advanced Statistical Inference (Term 1, 2022–23)

Assignment 1 · due on 3 October 2022

Please submit your answers in `.pdf` *format via Blackboard.*

1. Show that if X is a continuous random variable, then
$$\min_a E|X-a| = E|X-m|,$$
where m is the median of X.

2. Let X and Y be independent standard normal random variables.
   (a) Show that $\frac{X}{X+Y}$ has a Cauchy distribution.
   (b) Find the distribution of $X/|Y|$.

3. (a) Show that the Axiom of Countable Additivity implies Finite Additivity.
   (b) Let $A_1 \supset A_2 \cdots \supset A_n \supset \cdots$ be an infinite sequence of nested sets whose limit is the empty set, which we denote by $A_n \downarrow \emptyset$. Axiom of Continuity means if $A_n \downarrow \emptyset$, then $P(A_n) \to 0$. Prove that the Axiom of Continuity and the Axiom of Finite Additivity together imply Countable Additivity.

4. Prove
$$\lim_{n\to\infty} \frac{n!}{n^{(n+1/2)}e^{-n}} = C,$$
where C is a positive constant. You are not allowed to use Stirling's Formula to prove the claim.

5. Random variables $X_1, X_2, \ldots$ are called "m-dependent" if X_i and X_j are independent whenever $|i-j| \geq m$. Suppose $X_1, X_2, \ldots$ are m-dependent, with $E(X_j) = \mu$ and $Var(X_j) = \sigma^2 < \infty$ for $j \geq 1$. Show that $\bar{X}_n \xrightarrow{p} \mu$ as $n \to \infty$, where $\bar{X}_n = n^{-1}\sum_{i=1}^n X_i$.

6. Prove, by definition, that if $A_n \xrightarrow{p} 1$ and $Y_n \xrightarrow{d} Y$, then $A_nY_n \xrightarrow{d} Y$.

7. Let $X_1, X_2, \ldots$ be iid from the uniform distribution $U(1,2)$ and let H_n denote the harmonic average of the first n variables:
$$H_n = \frac{n}{X_1^{-1} + \ldots + X_n^{-1}}.$$
Show that $H_n \xrightarrow{p} c$ as $n \to \infty$. Identify the constant c.

8. Let $X_1, \cdots, X_n$ be independently distributed with exponential density

$$\frac{1}{2\theta} \exp(-x/2\theta) I(x \geq 0) \ ,$$

and let the ordered X's be denoted by $Y_1 \leq Y_2 \leq \cdots \leq Y_n$. It is assumed that Y_1 becomes available first, then Y_2, and so on, and that observation is continued until Y_r has been observed. This might arise, for example, in life testing where each X measures the length of life of, say, an electron tube, and n tubes are being tested simultaneously. Another application is to the disintegration of radioactive material, where n is the number of atoms, and observation is continued until r α-particles have been emitted.

   (a) Show that the joint distribution of $Y_1 \leq Y_2 \leq \cdots \leq Y_r$ has density

$$\frac{1}{(2\theta)^r} \frac{n!}{(n-r)!} \exp\left\{ -\frac{\sum_{i=1}^r y_i + (n-r) y_r}{2\theta} \right\} \ , \qquad 0 \leq y_1 \leq \cdots \leq y_r \ .$$

   (b) Argue that the distribution of $\{\sum_{i=1}^r y_i + (n-r) y_r\}/\theta$ is χ^2 with $2r$ degrees of freedom.

   (c) Let $Y_1, Y_2, \ldots$ denote the time required until the first, second, $\ldots$ event occurs in a Poisson process with parameter $1/2\theta'$. Prove that $Z_1 = Y_1/\theta'$, $Z_2 = (Y_2 - Y_1)/\theta'$, $Z_3 = (Y_3 - Y_2)/\theta', \cdots$ are independently distributed as χ^2 with 2 degrees of freedom, and the joint density of $Y_1, \ldots, Y_r$ has the density

$$\frac{1}{(2\theta')^r} \exp\left(-\frac{y_r}{2\theta'} \right) \ , \qquad 0 \leq y_1 \leq \cdots \leq y_r \ .$$

   The distribution of Y_r/θ' is again χ^2 with $2r$ degrees of freedom.

9. In statistics, a simple random sample is a subset of individuals chosen (one by one) from a population. Each individual is chosen randomly such that each individual has the same probability of being chosen at any stage during the sampling process, and each subset of k individuals has the same probability of being chosen for the sample as any other subset of k individuals.

   From a population of size N with finite variance, a simple random sample of size n is drawn without replacement, and a real-valued characteristic X measured to yield observations $X_j (j = 1, \ldots, n)$.

   (a) Show that the sample mean $\bar{X}_n$ is an unbiased estimator of the population mean m.

   (b) Show that the expected squared error of $\bar{X}$ as an estimator of m is smaller than that of the mean of a simple random sample of the same size n drawn with replacement.

   (c) Show that as $n, N \to +\infty$ and $r = n/N \to 0$ and the population variance is always less than M for all N, the difference between the expected squared errors of the two estimators is $O(N^{-1})$.

- End -