WPI Mathematical Sciences Ph.D. General Comprehensive Exam
MA 540 Probability and Mathematical Statistics - I
January, 2019

Note: Please make sure to write down your thinking process in bullet points even if you cannot solve the problems perfectly. Also, the problems are arbitrarily ordered, not necessarily according to difficulty.

1. (20 points) Let $P_1, ..., P_n$ be n i.i.d. Uniform(0,1) random variables, and $P_{(1)} \leq P_{(2)} \leq ... \leq P_{(n)}$ are their order statistics. The Simes statistic is defined as

$$T_n = \min_{1 \leq i \leq n} \left\{ P_{(i)} \cdot \frac{n}{i} \right\}.$$

Through the following steps show that

$$T_n \sim U(0, 1).$$

(a) (7 points) Let $f(x)$ be the density function of $P_{(n)}$. Find the formula of $f(x)$.

(b) (7 points) Show that, for any given $\alpha \in [0, 1]$,

$$P(T_n \leq \alpha) = \int_0^\alpha f(x)dx + \int_\alpha^1 P(T_n \leq \alpha | P_{(n)} = x) f(x) dx.$$

(c) (6 points) Apply induction to show that $T_n \sim U(0, 1)$. (Hint: Conditional on $P_{(n)} = x$ for any given $x \in [0, 1]$, the other P values are i.i.d. Uniform(0,x) random variables.)

2. (20 points) Suppose (X, Y) are uniformly distributed within the area of a circle centered at $(0, 0)$ with radius 1. Are the two random variables X and Y correlated? Why? Are X and Y dependent? Why?

3. (20 points) Let $f(y) = \frac{n!}{y!(n-y)!} \frac{B(y+\alpha, n-y+\beta)}{B(\alpha,\beta)}, y = 0, \ldots, n$, where $B(\cdot, \cdot)$ is the beta function. Find E(Y) and Var(Y).

4. (20 points) Let $X \sim$ Normal$(0, \sigma^2)$. Find a closed form expression (as a double integral of a fully specified bivariate normal pdf) for $E\{(\Phi(X))^2\}$, where $\Phi(\cdot)$ is the cdf of the standard normal density.

5. (20 points) Let $(X_1, \ldots, X_n)$ be a random sample from a Poisson distribution with mean θ . We are interested in estimating $g(\theta) = P(X_1 = 0) = e^{-\theta}$. Consider the following two estimators:

$$T_{n,1} = e^{-\bar{X}_n} \qquad \text{and} \qquad T_{n,2} = \frac{1}{n} \sum_{i=1}^n I(X_i = 0),$$

where $\bar{X}_n = \frac{1}{n} \sum_{i=1}^n X_i$, and I is an indicator function.

(a) (7 points) Find the asymptotic distribution of $T_{n,1}$.

(b) (7 points) Find the asymptotic distribution of $T_{n,2}$.

(c) (6 points) Which estimator is more efficient in estimating $g(\theta)$ when a large sample size is available? Show your argument.

6. (20 points) Suppose $X \sim N(0,1)$ and let $Y = f(X)$ for some function $f: \boldsymbol{R} \to \boldsymbol{R}$. It is well known that if f is a linear function (i.e. $f(x) = a + bx$), then Y has a normal distribution. Are there any nonlinear functions f for which Y has a normal distribution? Either prove there are not or supply an example of one and prove it works.