Ma 6251

Problems I

1. Let X_1, X_2, X_3 be a random sample from a distribution of the continuous type having probability density function $f(x) = 2x\, I_{(0,1)}(x)$. Compute the probability that the smallest of these X_i exceeds the median of the distribution.

2. Let $Y_1 < \dots < Y_n$ be the order statistics from a random sample of size n from an exponential distribution with mean 1.

(a) Show that

$$Z_1 = n Y_1, Z_2 = (n-1)(Y_2 - Y_1), \dots, Z_n = Y_n - Y_{n-1}$$

are stochastically independent and that each has an exponential distribution.

(b) Prove that all linear functions $\sum_{i=1}^{n} a_i Y_i$ of $Y_1, \dots, Y_n$ can be expressed as linear combinations of stochastically independent random variables.

3. Let F be a continuous distribution and $X_{(1)} < \dots < X_{(n)}$ be the order statistics from a random sample of size n. Set

$$C_1 = F(X_{(1)}), C_2 = F(X_{(2)}) - F(X_{(1)}), \dots, C_{n+1} = 1 - F(X_{(n)}).$$

Show that the joint probability density function of

$$\mathbf{C} = (C_1, C_3, C_4, \dots, C_{n+1})$$

is

$$h(c_1, c_3, c_4, \dots, c_{n+1}) = \begin{cases} n! & \text{if } \mathbf{c} \quad D \\ 0 & \text{otherwise} \end{cases},$$

where

$$D = \{ \mathbf{c} = (c_1, c_3, c_4, \dots, c_{n+1}) : c_i \geq 0 \text{ for } i = 1,3,4,5,\dots n+1 \text{ and } \sum_{i=1\, i\neq 2}^{n+1} c_i \leq 1 \}.$$

4. Prove the following. (Hint: You'll need Chebyshev's inequality, the weak law of large numbers, and the fact that if

$$U_n \xrightarrow{L} U \text{ and } V_n \xrightarrow{p} 0, \text{ then } U_n + V_n \xrightarrow{L} U.)$$

(a) $\dfrac{S_n}{\sigma} \xrightarrow{p} 1.$

(b) $\sqrt{\dfrac{n-1}{2}}\left(\dfrac{S_n^2}{\sigma^2} - 1\right) \xrightarrow{L} N(0, 1 + \dfrac{\gamma_2}{2}).$

5. Let X have the Beta distribution with parameters $\alpha > 0$ and $\beta > 0$, written $X \sim Be(\alpha, \beta)$. The probability density of X is

$$f(x) = \frac{\Gamma(\alpha+\beta)}{\Gamma(\alpha)\Gamma(\beta)}\, x^{\alpha-1}(1-x)^{\beta-1} I_{(0,1)}(x).$$

(a) Find the mean and variance of X.

(b) Find $E(X^k)$ for $k \in \{1,2, \ldots\}$ if $X \sim Be(m,n)$ and m and n are positive integers.

6. Find an approximate 90% distribution-free confidence interval for $\pi_{.4}$ for the continuous distribution from which the following data constitute a random sample.

100.89 98.76 109.75 97.94 110.04 100.95 119.27 98.97

100.74 100.51 99.11 101.46 106.02 101.02 98.35 102.97

7. Find an approximate 95% upper 75% tolerance region for the continuous distribution from which the following data constitute a random sample.

71.86 102.79 68.55 131.42 95.68 80.50 89.27

86.95 88.89 101.89 78.79 83.23 60.26 89.39 116.93

8. Let $y_1, \ldots, y_N$ denote N arbitrary numbers. Let Y denote the random variable which is the result of sampling one of the y-values at random (each y_i has probability $1/N$ of being sampled.)

a) Find E(Y).

b) Find Var(Y).

c) Specialize these to the case $y_i = i$, $i = 1, 2, \ldots, N$ and show that

$$E(Y) = \frac{N+1}{2}$$

and

$$\mathrm{Var}(Y) = \frac{N^2 - 1}{12}.$$

(Hint: Define the indicator random variables

$$Z_i = \begin{cases} 1 & \text{if } y_i \text{ is sampled} \\ 0 & \text{otherwise} \end{cases}$$

and write $Y = \sum_{i=1}^{N} Z_i \, y_i$.)