EASTERN UNIVERSITY, SRI LANKA

FIRST EXAMINATION IN SCIENCE 2002/2003

and 2002/2003(A)

SECOND SEMESTER

ST-104 - DISTRIBUTION THEORY

Answer all questions

Time : Three hours

---

1. (a) An environmental engineer measures the amount(by weight) of a particular pollution in air samples (of a certain volume) collected over the smoke stack of a coal - operated power plant. Let X_1 denote the amount of pollutant per sample when a certain cleaning device on the stack is not operating, and X_2 the amount of pollutant per sample when the cleaning device is operating, under similar environmental conditions. It is observed that X_1 is always greater than $2X_2$, and the relative frequency behavior of (X_1, X_2) can be modeled by

$$f(x_1, x_2) = \begin{cases} k & ; 0 \leq x_1 \leq 2, 0 \leq x_2 \leq 1, 2x_2 \leq x_1, \\ 0 & ; \text{elsewhere.} \end{cases}$$

i. Find the value of k that makes this a probability density function.

ii. Find $P(X_1 \geq 3X_2)$.

iii. Find the marginal density functions for X_1 and X_2.

iv. Are X_1 and X_2 independent?

v. Find $P(X_2 \leq 0.4)$.

(b) If X is a random variable with mean μ and variance σ^2, then for any positive number k, prove that
$P\{|X-\mu| \geq k\sigma\} \leq \frac{1}{k^2}$.

[70+30]

2. (a) Suppose that X and Y are continuous independent random variables with the following probability functions:

$$f_X(x) = \frac{\theta^\alpha x^{\alpha-1} e^{-\theta x}}{\Gamma(\alpha)},\ x > 0 \qquad f_Y(y) = \frac{\theta^\beta y^{\beta-1} e^{-\theta y}}{\Gamma(\beta)},\ y > 0$$

where α, β and θ are all greater than 0 and Γ denotes the gamma function. Define new random variables U and V as follows:

$$U = \frac{X}{X+Y}, \qquad V = X+Y$$

Find the distributions of U and V and show that they are independent.

(b) A particular job consists of two consecutive tasks. Duration times of these tasks are independent and identically distributed exponential random variables. The ratio of the time spent on the first task to total time for job is denoted by W. Using the results in (a). Find the distribution of W.

[75+25]

3. (a) Suppose that the length of time Y that it takes a worker to com-

plete a certain task has the probability density function

$$f(y) = \begin{cases} e^{-(y-\theta)} & y > \theta, \\ 0 & \text{otherwise.} \end{cases}$$

where θ is a positive constant that represents the minimum time to task completion. Let $Y_1, Y_2, \cdots, Y_n$ denote a random sample of completion times from this distribution.

i. Find the density function for $Y_{(1)} = min(Y_1, \cdots, Y_n)$.

ii. Find $E(Y_{(1)})$.

(b) Let $Y_1, Y_2, \cdots, Y_n$ be a random sample of size n from a normal distribution with a mean of μ and a variance of σ^2 and we define $Z_i = \dfrac{(Y_i - \mu)}{\sigma}$. Then show that $\sum_{i=1}^{n} Z_i^2 = \sum_{i=1}^{n} \left[\dfrac{Y_i - \mu}{\sigma}\right]^2$ is a χ^2 distribution with n degrees of freedom.

(c) Let Y have the probability density function given by

$$f_Y(y) = \begin{cases} \dfrac{y+1}{2} & ; -1 \le y \le 1, \\ 0 & ; \text{elsewhere.} \end{cases}$$

Find the density function for $U = Y^2$.

[40+30+30]

4. (a) The variables X and Y with zero means and standard deviations σ_1 and σ_2 are normally distributed and have correlation coefficient ρ. Also assume that $(X, Y) \sim BVN(0, 0, \sigma_1^2, \sigma_2^2, \rho)$.Show that $U = \dfrac{X}{\sigma_1} + \dfrac{Y}{\sigma_2}$ and $V = \dfrac{X}{\sigma_1} - \dfrac{Y}{\sigma_2}$ are independent normal variate with variance $2(1+\rho)$ and $2(1-\rho)$ respectively.

If $\sigma_1 = \sigma_2 = 1$ then show that

i. Regression of Y on X is linear.

ii. $X+Y$ and $X-Y$ are independent.

iii. $Q = \dfrac{(X^2 - 2\rho XY + Y^2)}{1-\rho^2}$ is a chi-square random variable.

(b) Suppose that $X_1, \cdots, X_m$ and $Y_1, \cdots, Y_n$ are two independent random samples, with the $X_i's$ normally distributed with mean μ_1 and variance σ_1^2 and the $Y_i's$ normally distributed with mean μ_2 and variance σ_2^2.

i. Find $E(\bar{X} - \bar{Y})$.

ii. Find $V(\bar{X} - \bar{Y})$.

iii. Suppose σ_1^2=2, σ_2^2=2.5, and $m = n$. Find the sample sizes so that $(\bar{X} - \bar{Y})$ will be within one unit of $(\mu_1 - \mu_2)$ with probability 0.95.

[50+50]

5. (a) The number of vehicles arriving at the campus junction in a period of 5 minutes is a Poisson random variable with expected value μ. A proportion $\theta(0 < \theta < 1)$ of all the vehicles which arrive at this junction intend to turn left. Derive the probability distribution of the number of vehicles which arrive at this junction intending to turn left in a period of 5 minutes.

(b) Let X and Y be independent discrete random variables which follow Poisson distribution with expected value μ and λ respectively. Find the distribution of the random variable $X+Y$.

(c) Electronic components of a certain type have a length of life Y, with a probability density given by

$$f(y) = \begin{cases} (1/100)e^{\frac{-y}{100}}, & y > 0, \\ 0 & \text{otherwise.} \end{cases}$$

(length of life measured in hours). Suppose the two such components operate independently and in series in a certain system (That is, the system fails when either component fails). Find the density function for X, the length of life of the system.

[30+30+40]

6. A particular fast-food outlet is interested in the joint behavior of the random variables Y_1, defined as the total time between a customer's arrival at the store and leaving the service window, and Y_2, the time that a customer waits in line before reaching the service window. Because Y_1 contains the time a customer waits in line, we must have $Y_1 \geq Y_2$, The relative frequency distribution of observed values of Y_1 and Y_2 can be modeled by the probability density function

$$f(y_1, y_2) = \begin{cases} e^{-y_1} \quad ; & 0 \leq y_2 \leq y_1 < \infty, \\ 0 & \text{otherwise.} \end{cases}$$

(a) Find $P(Y_1 < 2, Y_2 > 1)$

(b) find $P(Y_1 \geq 2Y_2)$

(c) If 2 minutes elapse between a customer's arrival at the store and departure from the service window, find the probability that he waited in line less than 1 minute to reach the window.

(d) Are Y_1 and Y_2 independent?

(e) The random variable $Y_1 - Y_2$ represents the time spent at the service window. Find $E(Y_1 - Y_2)$ and $V(Y_1 - Y_2)$. Is it highly likely that a customer would spend more than 2 minutes at the service window?

[$5 \times 20 = 100$]