Indian Institute of Information Technology Allahabad
Probability and Statistics
Marking Scheme

---

1. Consider a random experiment with the sample space $S = \{x_1, x_2, x_3, x_4, x_5\}$ and the event space $\Sigma = \{\emptyset, \{x_1\}, \{x_5\}, \{x_1, x_5\}, \{x_2, x_3, x_4\}, \{x_1, x_2, x_3, x_4\}, \{x_2, x_3, x_4, x_5\}, S\}$. Then, check whether the function $X : S \to \mathbb{R}$ defined by $X(x_1) = 3$, $X(x_2) = X(x_3) = X(x_4) = 0$ and $X(x_5) = 2$ is a random variable. [4]

   Solution: A function $X : S \longrightarrow \mathbb{R}$ is a random variable if and only if $X^{-1}\big((-\infty, x]\big) \in \sum \quad \forall x \in \mathbb{R}$

$$X^{-1}\big((-\infty, x]\big) = \begin{cases} \phi & x < 0 & [1] \\ \{x_2, x_3, x_4\} & 0 \le x < 2 & [1] \\ \{x_2, x_3, x_4, x_5\} & 2 \le x < 3 & [1] \\ S & x \ge 3 & [1] \end{cases}$$

   Thus X is a random variable.

2. Let $E_1, E_2, \cdots, E_n$ be a collection of exhaustive and mutually exclusive events. Let E and F be two events such that they are conditionally independent given E_i for every $1 \le i \le n$, i.e., $P(E \cap F | E_i) = P(E|E_i)P(F|E_i)$ for every $1 \le i \le n$. Then, show that E and F are independent if F and E_i are independent for every $1 \le i \le n$. [4]

   Solution:

$$P(E \cap F) = \sum_{i=1}^{n} P(E \cap F|E_i)P(E_i) \quad [1]$$

$$= \sum_{i=1}^{n} P(E|E_i)P(F|E_i)P(E_i) \quad [1]$$

$$= \sum_{i=1}^{n} P(E|E_i)P(F)P(E_i), \text{ (Since } F \text{ and } E_i \text{ are independent)} [1]$$

$$= P(F) \sum_{i=1}^{n} P(E|E_i)P(E_i)$$

$$= P(F)P(E) \quad [1]$$

3. Let X be a random variable with probability density function [9]

$$f_X(x) = \begin{cases} c(x+4), & -1 < x < 1 \\ 0, & otherwise \end{cases},$$

where c is a real constant. Let $Y = X^3$. Find

(a) the value of constant c.

(b) the distribution function of Y and hence find its PDF (don't use any direct formula).

(c) the expectation and variance of Y.

Solution: (a) Since

$$\int_{-\infty}^{\infty} f_X(x)dx = 1 \qquad [1]$$

$$\implies \int_{-1}^{1} c(x+4)dx = 1$$

$$\implies c\left(\frac{x^2}{2} + 4x\right)\Bigg|_{-1}^{1} = 1$$

$$\implies 8c = 1$$

$$\implies c = \frac{1}{8}. \qquad [1]$$

(b) The distribution function of Y

$$F_Y(y) = P(Y \le y) = P(X^3 \le y) = P(X \le y^{\frac{1}{3}})$$

$$F_Y(y) = \begin{cases} 0, & y < -1 \\ \frac{1}{16}(y^{\frac{2}{3}} + 8y^{\frac{1}{3}} + 7), & -1 \le y < 1 \\ 1, & y \ge 1 \end{cases} \qquad [1/2+2+1/2]$$

Now, PDF is given by

$$f_Y(y) = \begin{cases} \frac{1}{24}\left(y^{\frac{-1}{3}} + 4y^{\frac{-2}{3}}\right), & -1 \le y \le 1 \\ 0, & \text{otherwise} \end{cases} \qquad [1]$$

(c)

$$\begin{aligned}
E(Y) = E(X^3) &= \int_{-\infty}^{\infty} x^3 f_X(x)dx \\
&= \frac{1}{8}\int_{-1}^{1} x^3(x+4)dx \\
&= \frac{1}{8}\left(\frac{x^5}{5} + x^4\right)_{-1}^{1} \\
&= \frac{1}{20}. \qquad [1]
\end{aligned}$$

$$\begin{aligned}
Var(Y) = E[(Y-\mu)^2], \mu = E(Y) \\
&= E\left[\left(X^3 - \frac{1}{20}\right)^2\right] \\
&= \int_{-\infty}^{\infty} (x^3 - \frac{1}{20})^2 f_X(x)dx \\
&= \frac{1}{8}\int_{-1}^{1} (x^3 - \frac{1}{20})^2 (x+4)dx \\
&= \frac{393}{2800}. \qquad [1+1]
\end{aligned}$$

4. Let X be a random variable with p.m.f $F = f_X(x) = \begin{cases} \frac{2}{3^{x+1}}, & \text{if } x \in \{0,1,2,\cdots\} \\ 0, & \text{otherwise} \end{cases}$.

   Using Markov's Inequality, find a lower bound on $P(-5000 < X < 1011)$.

   Solution: Markov inequality $P(X \geq a) \leq \frac{E(X)}{a}$
   Also, $P(-5000 < X < 1011) = P(X < 1011)$. [1]
   Now, $E(X) = \sum_{x=0}^{\infty} \frac{2x}{3^{x+1}} = \frac{2}{3}\sum_{x=0}^{\infty} \frac{x}{3^x}$
   Series is convergent using Ratio test which implies $E(X)$ exists. [1]
   $S_n = \frac{2}{3}\left(\frac{1}{3} + \frac{2}{3^2} + \frac{3}{3^3} \cdots \frac{n}{3^n}\right) \qquad \cdots (1)$
   $\frac{1}{3}S_n = \frac{2}{3}\left(\frac{1}{3^2} + \cdots + \frac{n}{3^{n+1}}\right) \qquad \cdots (2)$
   Subtracting (2) from (1), we get
   $\frac{2}{3}S_n = \frac{2}{3}\left(\frac{1}{3} + \frac{1}{3^2} \cdots + \frac{1}{3^n} - \frac{n}{3^{n+1}}\right)$
   $\lim_{n\to\infty} S_n = \frac{1}{3}\left(\frac{1}{1-\frac{1}{3}}\right) = \frac{1}{2}$. [1]

Hence, $E(X) = \frac{1}{2}$.

Using Markov inequality, we get $P(X \geq 1011) \leq \frac{1}{2022}$ [1]

$1 - P(X < 1011) \leq \frac{1}{2022}$ and hence, $P(X < 1011) \geq \frac{2021}{2022}$. [1]

5. A student has taken a MCQ (multiple choice) examination. Every question has 5 options out of which one is correct. He continues to answer questions until he gets five correct answers. What is the probability that he gets them on twenty fifth question if he guesses at each answer? [3]

Solution: P(success) $= \frac{1}{5}$, P(failure) $= \frac{4}{5}$

He needs $5th$ success on $25th$ question

$X \sim NB(5, \frac{1}{5})$

$P(X = x) = \binom{r+x-1}{r-1} p^r q^x$

$P(X = 20) = \binom{5+20-1}{5-1} (\frac{1}{5})^5 (\frac{4}{5})^{20}$

$P(X = 20) = \binom{24}{4} (\frac{1}{5})^5 (\frac{4}{5})^{20}$. [3]