

# Probability and Statistical Techniques BCNS 1206C

## Unit 2: Discrete Random Variables

### 1 Univariate Discrete Random Variables

#### 1.1 Probability Mass Function (PMF)

**Definition 1.** A probability mass function for a discrete random variable  $X$  is a function  $p_X : \mathbb{R} \rightarrow [0, 1]$  defined by:

$$p_X(x) = \mathbb{P}(X = x),$$

with the condition that:

$$\sum_{x \in \mathbb{R}} p_X(x) = 1.$$

#### 1.2 Frequently Used Discrete Distributions

##### 1.2.1 Bernoulli Distribution

A random variable  $X$  follows a Bernoulli distribution with success probability  $p$  if it only takes values in  $\{0, 1\}$ . This is commonly used to model a single trial with two outcomes (e.g., coin toss). The probabilities are:

$$\mathbb{P}(X = 1) = p, \quad \mathbb{P}(X = 0) = 1 - p.$$

##### 1.2.2 Binomial Distribution

If  $X$  represents the number of successes in  $n$  independent Bernoulli trials with success probability  $p$ , then  $X$  follows a binomial distribution. The probability of observing exactly  $x$  successes is:

$$\mathbb{P}(X = x) = \binom{n}{x} p^x (1 - p)^{n-x}, \quad x = 0, 1, \dots, n.$$

##### 1.2.3 Poisson Distribution

A random variable  $X$  has a Poisson distribution with rate parameter  $\lambda > 0$  if it models the number of events occurring in a fixed interval. The probability mass function is:

$$\mathbb{P}(X = x) = \frac{\lambda^x e^{-\lambda}}{x!}, \quad x = 0, 1, 2, \dots$$

##### 1.2.4 Geometric Distribution

A variable  $X$  follows a geometric distribution with parameter  $p \in (0, 1)$  if it counts the number of trials until the first success. The PMF is:

$$\mathbb{P}(X = x) = p(1 - p)^{x-1}, \quad x = 1, 2, 3, \dots$$

**Exercise 1.** Verify that for each distribution above, the total probability sums to 1:  $\sum_x p_X(x) = 1$ .

#### 1.3 Transformations of Random Variables

Let  $X$  be a discrete random variable defined on a probability space  $(\Omega, \mathcal{F}, \mathbb{P})$ , and let  $g : \mathbb{R} \rightarrow \mathbb{R}$  be a function. Then  $Y = g(X)$  is also a discrete random variable. For example:

- If  $g(x) = ax + b$ , then  $Y = aX + b$ .
- If  $g(x) = cx^2$ , then  $Y = cX^2$ .

### 1.3.1 Expected Value and Variance

**Definition 2.** The expected value of a discrete random variable  $X$  is:

$$\mathbb{E}(X) = \sum_x x \cdot \mathbb{P}(X = x).$$

**Theorem 1** (Law of the Unconscious Statistician). If  $g : \mathbb{R} \rightarrow \mathbb{R}$  and  $X$  is a discrete random variable, then:

$$\mathbb{E}(g(X)) = \sum_x g(x) \cdot \mathbb{P}(X = x).$$

**Exercise 2.** Prove that  $\mathbb{E}(aX + b) = a \cdot \mathbb{E}(X) + b$  for constants  $a, b \in \mathbb{R}$ .

**Exercise 3.** Let  $X$  follow a Poisson distribution with parameter  $\lambda$ . Compute  $\mathbb{E}(e^X)$ .

**Exercise 4.** Let  $X$  follow a geometric distribution with parameter  $p$ . Find  $\mathbb{E}(e^{tX})$  for  $t \in \mathbb{R}$ .

**Definition 3.** The variance of a discrete random variable  $X$  is:

$$\text{Var}(X) = \mathbb{E}[(X - \mu)^2] = \sum_x (x - \mu)^2 \cdot \mathbb{P}(X = x),$$

where  $\mu = \mathbb{E}(X)$ .

**Exercise 5.** Show that  $\text{Var}(X) = \mathbb{E}(X^2) - [\mathbb{E}(X)]^2$ .

**Exercise 6.** Prove that  $\text{Var}(aX + b) = a^2 \cdot \text{Var}(X)$  for constants  $a, b \in \mathbb{R}$ .

## 2 Bivariate Distributions

**Definition 4.** Let  $X$  and  $Y$  be discrete random variables defined on a probability space  $(\Omega, \mathcal{F}, \mathbb{P})$ . The joint probability mass function (pmf) of  $X$  and  $Y$  is a function  $p_{X,Y} : \mathbb{R}^2 \rightarrow [0, 1]$  such that:

$$p_{X,Y}(x, y) = \mathbb{P}(X = x, Y = y),$$

with the condition that:

$$\sum_x \sum_y p_{X,Y}(x, y) = 1.$$

### 2.0.1 Marginal Distributions

The individual distributions of  $X$  and  $Y$  can be obtained by summing over the joint pmf:

$$p_X(x) = \sum_y p_{X,Y}(x, y), \quad p_Y(y) = \sum_x p_{X,Y}(x, y).$$

### 2.0.2 Expected Value in Two Dimensions

Given the joint pmf  $p_{X,Y}$ , the expected value of the product  $XY$  is:

$$\mathbb{E}(XY) = \sum_x \sum_y xy \cdot p_{X,Y}(x, y).$$

More generally, for any function  $g(x, y)$ :

$$\mathbb{E}[g(X, Y)] = \sum_x \sum_y g(x, y) \cdot p_{X,Y}(x, y).$$

**Exercise 7.** Prove that for constants  $a, b \in \mathbb{R}$ , we have:

$$\mathbb{E}(aX + bY) = a \cdot \mathbb{E}(X) + b \cdot \mathbb{E}(Y).$$

**Exercise 8.** Let  $(X, Y)$  have joint pmf:

$$\mathbb{P}(X = i, Y = j) = \theta^{i+j+1}, \quad \text{for } i, j \in \{0, 1, 2\}.$$

Compute  $\mathbb{E}(X)$  and  $\mathbb{E}(XY)$  in terms of  $\theta$ .

**Theorem 2.** If  $X$  and  $Y$  are independent discrete random variables, then:

$$\mathbb{E}(XY) = \mathbb{E}(X) \cdot \mathbb{E}(Y).$$

**Exercise 9.** Prove the above theorem using the definition of independence.

### 2.0.3 Covariance

**Definition 5.** The covariance between  $X$  and  $Y$  is defined as:

$$\text{Cov}(X, Y) = \mathbb{E}[(X - \mathbb{E}(X))(Y - \mathbb{E}(Y))] = \mathbb{E}(XY) - \mathbb{E}(X)\mathbb{E}(Y).$$

If  $X$  and  $Y$  are independent, then  $\text{Cov}(X, Y) = 0$ .

### 2.0.4 Correlation Coefficient

**Definition 6.** The correlation coefficient  $\rho(X, Y)$  is given by:

$$\rho(X, Y) = \frac{\text{Cov}(X, Y)}{\sqrt{\text{Var}(X) \cdot \text{Var}(Y)}},$$

provided that both variances are non-zero.

## 2.1 Practice Problems

**Problem 1.** Let  $X$  follow a geometric distribution with success probability  $p$ . Show that:

$$\mathbb{P}(X > k) = (1 - p)^k.$$

**Problem 2.** Consider a sequence of independent Bernoulli trials with success probability  $p$ . Let  $X$  be the trial number of the first success.

(a) Derive the pmf of  $X$  and verify that it sums to 1.

(b) Suppose the chance of a drone breaching a security zone is 0.01.

(i) What is the probability that the 80th drone is the first to succeed?

(ii) What is the probability that more than 80 drones are needed for a breach?

**Problem 3.** A yarn manufacturer sources from three suppliers:  $A$ ,  $B$ , and  $C$ , in the ratio 1:2:3. The number of flaws per unit length from each supplier follows a Poisson distribution with means 3, 2, and 1 respectively. If a randomly selected unit has 2 flaws, determine whether it is more likely to have come from supplier  $B$  or  $C$ .

**Problem 4.** Let  $X$  be uniformly distributed over  $\{1, 2, 3, 4\}$ . Given  $X = k$ , define  $Y$  as a random integer chosen uniformly from  $\{1, 2, \dots, k\}$ .

(a) Find the joint pmf of  $(X, Y)$ .

(b) Compute the marginal distributions of  $X$  and  $Y$ , and calculate  $\mathbb{E}(X)$ ,  $\mathbb{E}(Y)$ ,  $\text{Var}(X)$ , and  $\text{Var}(Y)$ .

(c) Determine  $\text{Cov}(X, Y)$  and the correlation coefficient  $\rho(X, Y)$ .

**Problem 5.** In a dice game, a player bets \$1 on a number from 1 to 6. Three fair dice are rolled. If the chosen number appears  $k$  times ( $k = 1, 2, 3$ ), the player wins \$ $k$  in addition to their stake. Otherwise, they lose their stake.

(a) Calculate the expected gain or loss per game.

(b) Compute the variance of the player's winnings.

**Problem 6.** Let  $X$  be a discrete random variable with pmf:

$$\mathbb{P}(X = i) = \frac{c\lambda^i}{i!}, \quad i = 0, 1, 2, \dots$$

Find the value of  $c$  that makes this a valid pmf, and compute  $\mathbb{P}(X > 2)$ .

**Problem 7.** Suppose  $X$  and  $Y$  are discrete random variables with joint pmf:

$$\mathbb{P}(X = x, Y = y) = c \cdot |x + y|, \quad x \in \{-2, 0, 2\}, \quad y \in \{-1, 0, 1\}.$$

Determine the value of  $c$  and calculate the correlation coefficient  $\rho(X, Y)$ .

**Problem 8.** Let  $X$  and  $Y$  be discrete random variables with joint pmf:

$$\mathbb{P}(X = x, Y = y) = \frac{1}{12}, \quad \text{for } (x, y) \in \{(1, 1), (1, 2), (2, 1), (2, 2), (3, 1), (3, 2)\}.$$

Find the marginal distributions of  $X$  and  $Y$ . Are  $X$  and  $Y$  independent?

**Problem 9.** Suppose  $X$  and  $Y$  are discrete random variables with joint pmf given by:

$$\mathbb{P}(X = x, Y = y) = \frac{x + y}{36}, \quad \text{for } x, y \in \{1, 2, 3\}.$$

(a) Verify that this is a valid joint probability distribution.

(b) Compute  $\mathbb{E}(X)$ ,  $\mathbb{E}(Y)$ , and  $\mathbb{E}(XY)$ .

**Problem 10.** Let  $X$  and  $Y$  be the number of heads in the first and second tosses of two fair coins, respectively. Define  $Z = X + Y$ . Find the joint pmf of  $X$  and  $Y$ . Determine the distribution of  $Z$ . Compute  $\text{Cov}(X, Y)$ .

**Problem 11.** A discrete random vector  $(X, Y)$  has the following joint distribution:

$X \backslash Y$	0	1	2
0	0.1	0.2	0.1
1	0.1	0.2	0.3

(a) Find the marginal distributions of  $X$  and  $Y$ .

(b) Calculate  $\mathbb{E}(X)$ ,  $\mathbb{E}(Y)$ , and  $\text{Cov}(X, Y)$ .

(c) Are  $X$  and  $Y$  positively or negatively correlated?

**Problem 12.** Let  $X$  and  $Y$  be discrete random variables such that:

$$\mathbb{P}(X = x, Y = y) = \begin{cases} \frac{1}{8}, & \text{if } (x, y) \in \{(0, 0), (0, 1), (1, 0), (1, 1)\}, \\ \frac{1}{4}, & \text{if } (x, y) = (2, 2), \\ 0, & \text{otherwise.} \end{cases}$$

Compute the correlation coefficient  $\rho(X, Y)$ . Interpret the result in terms of the relationship between  $X$  and  $Y$ .