Chapter IV

DISCRETE RANDOM VARIABLES - Part II

1 Functions of Random Variables

Let X be a discrete r.v., with PMF given by

$$P_X(x) = \begin{cases} p_i, \ x = x_i \\ \\ 0, \ otherwise \end{cases}, \quad 0 < p_i < 1, p_1 + p_2 + ... + p_n = 1,$$

D_X the range of P, and $g : D_X \to \mathbb{R}$ a real, bijective function.

Then, the PMF of the r.v. $Y = g(X)$ is:

$$P(Y = y) = P(g(X) = y) = P(X = g^{-1}(\{y\})) = P(X = x) = p,$$

therefore the PMF of r.v. Y is:

$$P_Y(y) = \begin{cases} p_i, \ y = g(x_i) \\ \\ 0, \ otherwise \end{cases}, \quad 0 < p_i < 1, p_1 + p_2 + ... + p_n = 1.$$

Example The amplitude V of a sinusoidal signal is a r.v. with PMF

$$P_V(x) = \begin{cases} \frac{1}{7}, \ x = -3, -2, ..., 2, 3 \\ \\ 0, \ otherwise. \end{cases}$$

Let $Y = \frac{V^2}{2}$ be the average power of the transmitted signal. Find the PMF of V.

The possible values of Y are $0,\ 0.5,\ 2,\ 4.5$, so the PMF of Y is:

$$P_Y(x) = \begin{cases} \frac{1}{7},\ x = 0 \\ \\ \frac{2}{7},\ x = 0.5,\ 2,\ 4.5 \\ \\ 0,\ otherwise. \end{cases}$$

Consider now two discrete r.v. X and Y, D_X the range of X, D_Y the range of Y. Then

$$h : D_X \times D_Y \to \mathbb{R},\ h(x_i, y_j) \overset{not}{=} q_{ij}$$

with

$$P(h(X,Y) = q) = P((X,Y) \in h^{-1}(q)).$$

2 Averages. Variance and Standard Deviation

A mode of a discrete r.v. X is a number x_{mod} satisfying

$$P_X(x_{mod}) \geq P_X(x),$$

for each value of x.
A median of a discrete r.v. X is a number x_{med} that satisfies

$$P(X < x_{med}) = P(X > x_{med}).$$

The expected value of X is

$$E(X) = \sum_{x \in S_X} x \cdot P_X(x).$$

Remark 2.1. Neither the mode nor the median of r.v. X need to be unique. A r.v. can have several modes or medians.

Example 1. For one quiz, ten students have the following grades:

$$9,\ 5,\ 10,\ 8,\ 4,\ 7,\ 5,\ 5,\ 8,\ 7.$$

Find the mean, the median and the mode.

The mean (*media aritmetică*) is:

$$m = \frac{9 + 5 \cdot 3 + 10 + 2 \cdot 8 + 4 + 3 \cdot 5 + 2 \cdot 7}{10} = 6.8$$

The median is 7 since are four scores below 7 and four score above 7.
The mode is 5 since it occurs more often than any other.

2. Determine the expected value of r.v. X whose PMF is

$$P_X(x) = \begin{cases} \frac{1}{8}, \ x = 0 \\ \frac{3}{8}, \ x = 1 \\ \frac{3}{8}, \ x = 2 \\ \frac{1}{8}, \ x = 3 \\ 0, \ otherwise. \end{cases}$$

The expected value is:

$$E(X) = 0 \cdot \frac{1}{8} + 1 \cdot \frac{3}{8} + 2 \cdot \frac{3}{8} + 3 \cdot \frac{1}{8} = \frac{3}{2}.$$

Proposition 2.1. *For any r.v. X, the following properties hold:*
1. $E(a \cdot X + b) = a \cdot E(X) + b, \ \forall a, b \in \mathbb{R}.$
2. $E(X - E(X)) = 0;$
3. $E(X + Y) = E(X) + E(Y).$

Proposition 2.2. *1. The Bernoulli (p) r.v. X has the expected value $E(X) = p$.*
2. The geometric (p) r.v. X has the expected value $E(X) = \dfrac{1}{p}$.

3. The Poisson (α) r.v. X has the expected value $E(X) = \alpha$.

4. The binomial (n, p) r.v. X has the expected value $E(X) = n \cdot p$.

5. The Pascal (k,p) r.v. X has the expected value $E(X) = \dfrac{k}{p}$.

6. The discrete (k,l) r.v. X has the expected value $E(X) = \frac{k+l}{2}$.

The variance of r.v. X is the number given by

$$var(X) = E[(X - E(X))^2].$$

The standard deviation of r.v. X is

$$\sigma_X = \sqrt{var(X)}.$$

Remark 2.2. We think of sample values within σ_X of the expected values $x \in [E(X) - \sigma_X,\ E(X) + \sigma_X]$ as "typical" values of X and other values as "unusual."

Proposition 2.3. *1. For any r.v. X, the variance can be computed using the formula:*

$$var(X) = E(X^2) - E(X)^2.$$

2. The variance of any r.v X is a positive number:

$$var(X) > 0.$$

3.

$$var(a \cdot X) = a^2 \cdot var(X), \ \ \forall a \in \mathbb{R}.$$

Proposition 2.4. *1. The Bernoulli (p) r.v. X has the variance $var(X) = p(1-p)$.*

2. The geometric (p) r.v. X has the variance $var(X) = \dfrac{1-p}{p^2}$.

3. The Poisson (α) r.v. X has the variance $var(X) = \alpha$.

4. The binomial (n,p) r.v. X has the variance $var(X) = np(1-p)$.

5. The Pascal (k,p) r.v. X has the variance $var(X) = \dfrac{k(1-p)}{p^2}$.

6. The discrete (k,l) r.v. X has the variance $var(X) = \dfrac{(l-k)(l-k+2)}{12}$.

The n-th moment for r.v. X is $E(X^n)$.

The n-th central moment for r.v. X is $E[(X - E(X))^n]$.

Example In an experiment to monitor two calls, the PMF of $X-$ the number of voice calls, is

$$P_X(x) = \begin{cases} \frac{1}{10}, & x = 0 \\ \frac{2}{5}, & x = 1 \\ \frac{1}{2}, & x = 2 \\ 0, & otherwise. \end{cases}$$

Find:
a) The expected value of X;
b) The expected value of X^2;
c) The variance of X.
d) The standard deviation for X.

Solution: a) $E(X) = 0 \cdot \frac{1}{10} + 1 \cdot \frac{2}{5} + 2 \cdot \frac{1}{2} = \frac{14}{10}$.

b) $E(X^2) = 0^2 \cdot \frac{1}{10} + 1^2 \cdot \frac{2}{5} + 2^2 \cdot \frac{1}{2} = \frac{24}{10}$.

c) $var(X) = E(X^2) - E(X)^2 = \frac{24}{10} - (\frac{14}{10})^2 = 0.44$

d) $\sigma(X) = \sqrt{var(X)} = \sqrt{0.44}$.

3 Conditional Probability Mass Function

Given an event $B, P(B) > 0$, **the conditional PMF of** X is:

$$P_{X|B}(x) = P(X = x \mid B).$$

Proposition 3.1. *If $B_1, B_2, ..., B_m$ is an event space, then:*

$$P_X(x) = \sum_{i=1}^{m} P_{X|B_i} \cdot P(B_i).$$

Proposition 3.2. *When a conditioning event B is included in the range of X, then the conditional PMF of X given B is:*

$$P_{X|B}(x) = \frac{P(X = x \mid B)}{P(B)} = \begin{cases} \dfrac{P_X(x)}{P(B)}, & x \in B \\ \\ 0, & otherwise \end{cases}$$

Example The length of a fax is a discrete r.v. which has PMF

$$P_X(x) = \begin{cases} 0.15, & x = 1, 2, 3, 4 \\ \\ 0.1, & x = 5, 6, 7, 80, \quad otherwise \end{cases}$$

Suppose the company has two fax machines, one for shorter than 5 pages and the other for faxes that have more pages.
a) What is the PMF of fax length in the second machine?
b) Find the expected value, the variance and the standard deviation of $X \mid B$.

Solution a) The condition is $B = \{5, 6, 7, 8\}$, so the probability of B is $P(B) = 4 \cdot 0.1 = 0.4$. The conditional PMF of X given B is:

$$P_{X|B}(x) = \begin{cases} \dfrac{P_X(x)}{P(B)}, & x \in B \\ \\ 0, & otherwise \end{cases} = \begin{cases} \dfrac{0.1}{0.4}, & x \in B \\ \\ 0, & otherwise \end{cases} = \begin{cases} \dfrac{1}{4}, & x \in \{5, 6, 7, 8\} \\ \\ 0, & otherwise. \end{cases}$$

b) $E(X \mid B) = 5 \cdot \dfrac{0.1}{0.4} + 6 \cdot \dfrac{0.1}{0.4} + 7 \cdot \dfrac{0.1}{0.4} + 8 \cdot \dfrac{0.1}{0.4} = 26 \cdot 0.25 = 6.5$ pages.

Now,

$$var(X \mid B) = E(X^2 \mid B) - [E(X \mid B)]^2.$$

First, we compute

$E(X^2 \mid B) = 25 \cdot 0.25 + 36 \cdot 0.25 + 49 \cdot 0.25 + 64 \cdot 0.25 = 174 \cdot 0.25 = 43.5$ pages, so

$var(X \mid B) = 43.5 - 6.5^2 = 1.25$ pages.

4 Solved Problems

1. A student takes two courses. In each course, the student will earn a B with probability 0.6, or a C with probability 0.4, independent to the other course.

   To calculate a grade point average, a B worth 3 points and a C is worth 2 points. The student's average is the sum of the points for each course divided by 2.

   Make a table of a simple space of the experiment and the corresponding values of the student's average, X. Write the corresponding cumulative distribution function (CDF).

   Solution

   $$\begin{array}{ccccc} X: & BB & BC & CB & CC \\ & 3 & \frac{5}{2} & \frac{5}{2} & 2 \end{array}$$

   The corresponding probabilities are: $P(X = 3) = 0.6 \cdot 0.6 = 0.36$, $P(X = \frac{5}{2}) = 2 \cdot 0.6 \cdot 0.4 = 0.48, P(X = 2) = 0.4 \cdot 0.4 = 0.16$ therefore, the PMF for X given B is:

   $$P_X(x) = \begin{cases} 0.36, & x = 3 \\ 0.48, & x = \frac{5}{2} \\ 0.16, & x = 2 \\ 0, & otherwise. \end{cases}$$

   The corresponding cumulative distribution function is given by:

   $$F_X(x) = \begin{cases} 0, & x < 30.36, \quad x \in [3, \frac{5}{2}) \\ 0.84, & x \in [\frac{5}{2}, 2) \\ 1, & x \in [2, \infty). \end{cases}$$