# Chapter 3: Discrete Random Variables and Probability Distributions

Defn: A random variable (r.v.) is a function that takes a sample point in S and assigns to it a real number.

- That is, the r.v. $Y$ maps the sample space (its domain) to the real number line (its range), i.e., $Y: S \longrightarrow \mathbb{R}$.

Simple example: In an experiment, we toss two coins. Let the r.v. $Y$ = the number of "heads".

$S =$

The r.v. $Y$ as a function:

Defn: A discrete random variable is a r.v. that takes on a finite or countably infinite number of different values.

- Since each sample point has an associated probability, we can similarly determine the probability for each value of $Y$.
- This collection of $Y$-values and their probabilities is called the probability distribution of the discrete r.v.

Notation: We use a capital letter (like $Y$) to denote a random variable and a lowercase letter (like $y$) to denote a particular value the r.v. may take.

Previous example: $Y$ represents the number of "heads" in the experiment. If we toss the coins and both are heads, then $y=2$ for that trial.

- Hence $Y$ is random (not certain), but an observed value $y$ is fixed, not random.

- We denote the probability distribution of a discrete r.v. with the notation $P(Y=y)$, or simply $p(y)$.
- For some number $y$, consider the set of sample points that yield $Y=y$. We define $P(Y=y)$ to be the sum of the probabilities for the sample points in this set.

<u>Example</u> <u>above</u>: (Suppose the two coins are fair)

- Each sample point has probability

- We may represent a discrete probability distribution using a table, graph, or formula.

Example:

Example 2: Consider an urn of 4 balls labeled 1, 2, 3, and 4. We draw 2 balls at random (without replacement). Let $Y$ = the sum of the numbers on the selected balls. Find the probability distribution for $Y$:

- The sample space is

Theorem: For any discrete probability distribution:

(i)

(ii)

- These facts follow readily from the basic laws of probability.

Note: Probability distributions are used as models (approximations) of the behavior of data in the real world. The usefulness of the probabilities we calculate depends on how well the model fits the real data's pattern.

## 3.3 Expected Value

- A parameter is a numerical characteristic of a probability distribution (or of the population of data that it models).

Defn. Let $Y$ be a discrete r.v. with probability distribution $p(y)$. Then the expected value of $Y$, denoted $E(Y)$, is:

Note: $E(Y)$ exists if this sum is absolutely convergent (if $\sum_y |y| p(y) < \infty$). This condition will typically hold for the examples we cover.

Note: If $p(y)$ characterizes some population distribution, then $E(Y)$ is often called the population mean $\mu$.

<u>Theorem</u>: If $Y$ is a discrete r.v. with probability function $p(y)$ and $g(Y)$ is a real-valued function of $Y$, then:

$$E[g(Y)] = \sum_{\text{all } y} g(y)\,p(y).$$

<u>Proof</u> (assuming $Y$ has a finite range):

- The expected value (or mean) $E(Y) = \mu$ describes the <u>center</u> of a population.
- The spread, or variability, of a population is described by the variance $\sigma^2 = E\left[(Y-\mu)^2\right]$.
- The <u>variance</u> $V(Y) = E\left[(Y-\mu)^2\right]$ is the expected squared difference between a random variable and its mean.
- The <u>standard deviation</u> of $Y$ is defined as:

<u>Theorem:</u> Let $Y$ have discrete probability function $p(y)$. Let $g(Y)$ be a function of $Y$ and $c$ be a constant. Then

$$E[c\, g(Y)] = c\, E[g(Y)].$$

<u>Proof:</u>

**Corollary**: For any constant c, $E(c) = c$.

**Proof**:

**Theorem**: Let $g_1(Y), g_2(Y), \ldots, g_k(Y)$ be k functions of Y. Then

$$E[g_1(Y) + g_2(Y) + \cdots + g_k(Y)]$$
$$= E[g_1(Y)] + E[g_2(Y)] + \cdots + E[g_k(Y)].$$

**Proof** ($k=2$ case):

Note: Since $\mu$ is a constant and $E(Y) = \mu$, we can write the variance of $Y$ as:

Example 1: Suppose $Y$ has probability distribution:

| $y$ | 0 | 1 | 2 | 3 |
|---|---|---|---|---|
| $p(y)$ | $\frac{1}{125}$ | $\frac{12}{125}$ | $\frac{48}{125}$ | $\frac{64}{125}$ |

Find $E(Y)$, $E(Y^2)$, $V(Y)$, and $E[(3Y+2)^2]$.

Example 2 (Petersburg Paradox): Consider a series of flips of a fair coin. A player will receive \$2 if the first head occurs on flip 1, \$4 if it occurs on flip 2, \$8 if on flip 3, etc. What is the player's expected earnings?