# CP Statistics – Chapter 6: Discrete Random Variables

## 6.1: Discrete and Continuous Random Variables

**RANDOM VARIABLE**
A **random variable** is a variable whose value is a numerical outcome of a random phenomenon.

**DISCRETE RANDOM VARIABLE**
A **discrete random variable** *X* has a *countable* number of possible values. Generally, these values are limited to integers (whole numbers). The probability distribution of X lists the values and their probabilities.

| **Value of X** | $x_1$ | $x_2$ | $x_3$ | … | $x_k$ |
|---|---|---|---|---|---|
| **Probability** | $p_1$ | $p_2$ | $p_3$ | … | $p_k$ |

**The probabilities $p_i$ must satisfy two requirements:**

1. Every probability $p_i$ is a number between 0 and 1.
2. $p_1 + p_2 + \ldots + p_k = 1$

Find the probability of any event by adding the probabilities $p_i$ of the particular values $x_i$ that make up the event.

**CONTINUOUS RANDOM VARIABLE**
A **continuous random variable** *X* takes all values in an interval of numbers and is *measurable*.

> **For example:** a person's height in inches can take on any value and is not limited to a whole number of inches.

**MEAN OF A DISCRETE RANDOM VARIABLE**
Suppose that X is a discrete random variable whose distribution is

| **Value of X** | $x_1$ | $x_2$ | $x_3$ | … | $x_k$ |
|---|---|---|---|---|---|
| **Probability** | $p_1$ | $p_2$ | $p_3$ | … | $p_k$ |

To find the **mean** of X, multiply each possible value by its probability, then add all the products:

$$\boldsymbol{\mu_x = E(x) = \sum x_i \cdot p_i = x_1 \cdot p_i + x_2 \cdot p_2 + \cdots + x_k \cdot p_k}$$

**For example, for the distribution below:**

| X | 1 | 2 | 3 | 4 | 5 |
|---|---|---|---|---|---|
| *P*(X) | 0.1 | 0.2 | 0.3 | 0.3 | .1 |

**The mean would be computed as**

$$\mu_X = 1(.1) + 2(.2) + 3(.3) + 4(.3) + 5(.1) = 3.1$$

**LAW OF LARGE NUMBERS**
Draw independent observations at random from any population with finite mean μ. As the number of observations drawn increases, the mean of the observed values eventually approaches the mean μ.

For example, if a coin has probability .5 of being flipped tails, the proportion of tails will be very close to .5 *in the long run*. We cannot say what will happen in the short term, but after *many, many* flips, this proportion will be approximately .5.

# 6.3 – Binomial Random Variables

A **binomial probability distribution** occurs when the following requirements are met.

**B** – Each outcome is ***binary***, having only two possibilities – call them "success" or "failure."
**I** – The observations must be ***independent*** – result of one does not affect another.
**N** – The procedure has a fixed ***number*** of trials – we call this value *n*.
**S** – The probability of ***success*** – call it *p* - remains the same for each observation.

**Notation for binomial probability distribution**

***n*** = the number of fixed trials

***k*** = the number of successes in the *n* trials

***p*** = the probability of success

***1 – p*** = the probability of failure

**Binomial Probability Formula**

$$P(X=k)=\frac{n!}{k!(n-k)!}(p)^{k}(1-p)^{n-k}$$

**Mean (expected value) of a Binomial Random Variable**

Formula: $\mu = np$ Meaning: Expected number of successes in *n* trials (think *average*)

Example: *Suppose you are a 80% free throw shooter. You are going to shoot 4 free throws.*

For *n* = 4, *p* = .8, $\mu=(4)(.8)=3.2$, which means we expect 3.2 makes out of 4 shots, on average

---

# 6.3 – Geometric Random Variables

A **geometric probability distribution** occurs when the following requirements are met.

**B** – Each outcome is ***binary***, having only two possibilities – call them "success" or "failure."
**I** – The observations must be ***independent*** – result of one does not affect another.
**T** – The variable of interest is the number of ***trials*** required to obtain the first success.*
**S** – The probability of ***success*** – call it *p* - remains the same for each observation.
* As such, the geometric setting is also called a "waiting-time" distribution

**Notation for geometric probability distribution**

***n*** = the number of trials required to obtain the first success

***p*** = the probability of success

***1 – p*** = the probability of failure

**Geometric Probability Formula**

$$P(X=n)=(1-p)^{n-1}(p)$$

**Mean (expected value) of a Geometric Random Variable**

Formula: $\mu=\frac{1}{p}$ Meaning: Expected number of *n* trials to achieve first success (*average*)

Example: *Suppose you are a 80% free throw shooter. You are going to shoot until you make.*

For *p* = .8, $\mu=\frac{1}{.8}=1.25$, which means we expect to take 1.25 shots, on average, to make first