4.1 – Probability Distribution Function (PDF) for a Discrete Random Variable

In this chapter we will construct *discrete probability distribution functions*, by combining the descriptive statistics that we learned from chapters 1 and 2 and the probability from chapter 3.

A random variable, often denoted by $x$, is a variable whose value is determined by the outcome of a random experiment.

A random variable that assumes countable values is called a **discrete random variable**.

A random variable that can assume any value contained in one or more intervals is called a **continuous random variable**.

A **probability distribution of a discrete random variable** lists all the possible outcomes of the random variable and their corresponding probabilities.

**Requirements for Probability Distributions:**

1. $0 \leq P(x) \leq 1$ for each value of $x$
2. $\sum P(x) = 1$

Below is an example of a probability distribution, presented as a table on the left and also as a bar graph on the right.
1. Using the probability distribution for number of vehicles owned (on previous page), determine
   
   (a) \( P( x = 4 ) \)
   
   (b) \( P( x < 3 ) \)
   
   (c) \( P( 0 < x \leq 3 ) \)
   
   (d) the probability that a family owns more than 2 cars

2. Suppose we toss a coin three times. Let \( x \) denote the total number of heads. Write the
   probability distribution of \( x \).

3. In a group of 12 persons, 3 are left-handed. Suppose that 2 persons are randomly selected from
   this group. Let \( x \) denote the number of left-handed persons in this sample. Write the probability
   distribution of \( x \).
4.2 – Mean or Expected Value and Standard Deviation

The **mean** of a discrete random variable, also called the **expected value**, and sometimes also referred to as the “long-term average”, is  

\[ E(x) = \mu_x = \sum x \cdot P(x) \]

The **variance** of a discrete random variable is  

\[ \sigma_x^2 = \sum [(x - \mu_x)^2 \cdot P(x)] \]

The **standard deviation** of \( x \) is  

\[ \sigma_x = \sqrt{\sigma_x^2} \]

How to find the mean and standard deviation by using TI83/84:

4. Find the mean and standard deviation in problem #3 on the previous page.

**Interpretation of the Mean of a Random Variable**

The mean, also called the expected value, can be used to predict what value of a random variable one would "expect" to find if one could repeat the random variable process an infinite number of times and take the average of those values.

5. Interpret the mean that you found in problem 4.

The expected value can also be used to predict how much someone is likely to gain or lose when taking certain actions.

6. An insurance company sells a one-year term life insurance policy to an 80-year-old woman. The woman pays a premium of $1000. If she dies within one year, the company will pay $20,000 to her beneficiary. According to the U.S. Center of Disease Control and Prevention, the probability that an 80-year-old woman will be alive one year later is 0.9516. Let \( X \) be the profit made by the insurance company. Find the probability distribution and the expected value of the profit. Interpret the expected value.
4.3 – The Binomial Probability Distribution

For a binomial experiment, the probability of exactly \( x \) successes in \( n \) trials is

\[
P(x) = \binom{n}{x} \cdot p^x \cdot q^{n-x}
\]

where
- \( n \) = total number of trials
- \( p \) = probability of success
- \( q \) = probability of failure
- \( x \) = number of successes in \( n \) trials
- \( (n-x) \) = number of failures in \( n \) trials

Conditions of a Binomial Experiment:

1. There are in identical trials (repetitions).
2. Each trial has exactly two possible outcomes. We call them “success” vs. “failure”.
3. The probability of each success, denoted \( p \), remains the same in all trials, as do the probability for each failure, \( 1-p \).
4. The trials are independent.
5. The random variable \( x \) represents the number of successes that occur.

The two outcomes are commonly referred to as success and failure; although a success is not necessarily representing something good or positive.

The mean and standard deviation of the binomial distribution are

\[
\mu = np \quad \text{and} \quad \sigma = \sqrt{npq}
\]

We can use our calculators to find the probabilities for a binomial random variable. This is how to do it on a TI-83/84:

Press \[ \text{PRGM} \rightarrow \text{select BINOML83} \rightarrow \text{ENTER} \rightarrow \text{ENTER} \rightarrow \] choose the option that fits your specific problem.

If you are asked to find the probability distribution or want to find many single probabilities, option 3 might be a good choice, which will give you the probabilities for all possible \( x \)'s.

OR

You can use the \text{binompdf} or \text{binomcdf} functions, which you find by pressing \[ \text{2ND}, \text{VARS} \] to access the distribution menu, from which you would choose:

- \text{binompdf} if you want to find \( P(x) \), and you would enter \text{binompdf}(n, p, x) or \text{binomcdf} if you want to find \( P(\text{less than or equal to } x) \), and you would enter \text{binomcdf}(n, p, x).
$X \sim B(n,p)$ reads “$X$ is a random variable with a binomial distribution and the parameters $n$ and $p.$”

**ex.** Suppose we roll a die 20 times. Find

(a) $P($getting 5 sixes$)$

(b) $P($getting 6 sixes$)$

(c) $P($getting 5 or 6 sixes$)$

(d) $P($getting less than 5 sixes$)$

(e) $P($getting at least 6 sixes$)$

**ex.** An airline has a policy of booking as many as 15 persons on an airplane that can seat only 14. Past studies have revealed that only 85% of the booked passengers actually arrive for the flight. Find the probability that if the airline books 15 persons, not enough seats will be available. Is the probability low enough so that overbooking is not a real concern for passengers?

**ex.** After being rejected for employment, Kim Kelly learns that the Bellevue Credit Company has hired only two women among the last 20 new employees. She also learns that the pool of applicants is very large, with an approximately equal number of qualified men and women. Help her address the charge of gender discrimination by finding the probability of getting two or fewer women when 20 people are hired, assuming that there is no discrimination based on gender. Does the resulting probability really support such a charge?