Chapter 4: Discrete Random Variables

Introduction

Random Variable Notation

- **Random variable** - a variable whose outcome is determined by chance and defines a numerical quantity to each outcome in the sample space.

  Upper case letters such as X or Y denote a random variable. Lower case letters like x or y denote the value of a random variable. If X is a random variable, then X is written in words, and x is given as a number.

  For example, let $X =$ the number of heads you get when you toss three fair coins. The sample space for the toss of three fair coins is

  $$TTT; THH; HTH; HHT; HTT; THT; TTH; HHH.$$

  Then, $x = 0, 1, 2, 3$. X is in words and x is a number. Notice that for this example, the x values are countable outcomes. Because you can count the possible values that X can take on and the outcomes are random (the x values $0, 1, 2, 3$), X is a discrete random variable.

4.1: Probability Mass Function (PMF) for a Discrete Random Variable

- **Probability distribution** - a description that gives the probability for each value of a random variable.

- **Probability mass function (PMF)** - a probability distribution for a discrete random variable
  - Notated as $f(x)$ where $f(x) = P(X = x)$
  - Properties
    - $0 \leq f(x) \leq 1$ (each probability is between 0 and 1, inclusive)
    - $f(x) = 0$ for every x not in the sample space
    - The sum of all probabilities must be one

4.2: Mean or Expected Value and Standard Deviation

- **Expected value** - the mean value of a random variable averaged over the sample space
- Mean of a discrete random variable - defined by, $\mu = \sum x f(x)$
- Standard deviation of a discrete random variable - defined by

$$\sigma = \sqrt{\sum_S (x-\mu)^2 f(x)} = \sqrt{\sum_S [x^2 f(x)] - \mu^2}$$

Example 1. Consider rolling a die two times and recording the sum of the dice. The random variable X is the sum of the two dice. X can take on the values from two through twelve. Find the PMF for the sum of the two rolled dice.

X	$f(x)$
2	1/36
3	2/36
4	3/36
5	4/36
6	5/36
7	6/36
8	5/36
9	4/36
10	3/36
11	2/36
12	1/36

(a) Find the expected value of the random variable X.

Example 2. Consider a baseball team that plays a four game series with another team. Let the random variable X represent the number of games that team wins. X can take the values of 0, 1, 2, 3, or 4. The PMF for the number of games this team wins in a four game series is given as

X	$f(x)$
0	0.0256
1	0.1536
2	0.3456
3	0.3456
4	0.1296

(a) Find the expected value of the random variable X.

(b) Find the standard deviation.

4.3: Binomial Distribution

- A Bernoulli trial is defined as a random experiment where only one of two mutually exclusive outcomes can occur.
- A binomial distribution is a sequence of Bernoulli trials such that we observe a fixed number of independent trials and the probability of success remains constant from trial to trial.
- If the random variable X is a binomial random variable, we denote it as $X \sim B(n, p)$ where n is the possible number of trials and p is the probability of a success in a single trial.
- If the sample size is less than 5% of the population, it is acceptable to treat without replacement sampling as with replacement sampling.

- Examples: Consider a standard deck of 52 playing cards.
  - Picking three cards at random and recording their value is not a binomial trial because we do not have two mutually exclusive possible outcomes.
  - Picking three cards at random with replacement and recording whether they are face cards or not is a binomial trial, and we can say the random variable $X \sim B(3, \frac{3}{13})$. Picking three cards at random without replacement and recording whether they are face cards or not is not a binomial trial because the trials are dependent.
  - Picking cards at random and recording whether they are a face card until we find three face cards is not a binomial trial because we do not have a fixed number of trials.

- Properties
  - The PMF of the binomial distribution is $f(x) = \binom{n}{x} p^x q^{n-x}$, $x = 0, 1, \ldots, n$, where $\binom{n}{x} = \dfrac{n!}{x!(n-x)!}$ and $q = 1 - p$.
  - The mean of a binomial distribution is np.
  - The variance of a binomial distribution is npq.

Example 3. The earlier baseball example actually follows a binomial distribution where $n = 4$ and $p = 0.6$.

(a) This explains why $\mu = 2.4$ and $\sigma = \sqrt{4(0.6)(0.4)} = 0.98$.

(b) However, we can find the probability of this team winning two of the four games by using the binomial PMF by:

(c) We can find the probability of this team winning at least one game by using the binomial PDF by

(d) We can find the probability of the four game series not being a sweep by the PDF by

Example 4. A business advertises by cold calling potential customers at their homes. From experience, it is known that only 1 out of 10 people called will buy their product. One day, the company plans to call 50 people.

(a) Find the probability that any individual sales call is successful.

(b) Find the expected number of successful sales calls.

(c) Find the variance of this distribution.

(d) Find the standard deviation of this distribution.

(e) Find the probability that exactly two sales calls are successful.

(f) Find the probability that at least three sales calls are successful.

