4.6: Key Terms

Bernoulli Trials

an experiment with the following characteristics:

1. There are only two possible outcomes called "success" and "failure" for each trial.
2. The probability p of a success is the same for any trial (so the probability $q = 1 - p$ of a failure is the same for any trial).

Binomial Experiment

a statistical experiment that satisfies the following three conditions:

1. There are a fixed number of trials, n.
2. There are only two possible outcomes, called "success" and, "failure," for each trial. The letter p denotes the probability of a success on one trial, and q denotes the probability of a failure on one trial.
3. The n trials are independent and are repeated using identical conditions.

Binomial Probability Distribution

a discrete random variable (RV) that arises from Bernoulli trials; there are a fixed number, n, of independent trials. "Independent" means that the result of any trial (for example, trial one) does not affect the results of the following trials, and all trials are conducted under the same conditions. Under these circumstances the binomial RV X is defined as the number of successes in n trials. The mean is $\mu = np$ and the standard deviation is $\sigma = \sqrt{npq}$. The probability of exactly x successes in n trials is

$P(X = x) = \binom{n}{x} p^{\mathrm{x}} q^{\mathrm{n-x}}$

Geometric Distribution

a discrete random variable (RV) that arises from the Bernoulli trials; the trials are repeated until the first success. The geometric variable X is defined as the number of trials until the first success. The mean is $\mu = \dfrac{1}{p}$ and the standard deviation is $\sigma = \sqrt{\dfrac{1}{p}\left(\dfrac{1}{p} - 1\right)}$. The probability of exactly x failures before the first success is given by the formula:
$P(X = x) = p(1 - p)^{x-1}$ where one wants to know the probability for the number of trials until the first success: the xth trial is the first success.
An alternative formulation of the geometric distribution asks the question: what is the probability of x failures until the first success? In this formulation the trial that resulted in the first success is not counted. The formula for this presentation of the geometric is: $P(X = x) = p(1 - p)^x$

The expected value in this form of the geometric distribution is $\mu = \dfrac{1-p}{p}$.

The easiest way to keep these two forms of the geometric distribution straight is to remember that p is the probability of success and (1−p) is the probability of failure. In the formula the exponents simply count the number of successes and number of failures of the desired outcome of the experiment. Of course the sum of these two numbers must add to the number of trials in the experiment.

Geometric Experiment

a statistical experiment with the following properties:

1. There are one or more Bernoulli trials with all failures except the last one, which is a success.
2. In theory, the number of trials could go on forever. There must be at least one trial.
3. The probability, p, of a success and the probability, q, of a failure do not change from trial to trial.

Hypergeometric Experiment

a statistical experiment with the following properties:

1. You take samples from two groups.
2. You are concerned with a group of interest, called the first group.
3. You sample without replacement from the combined groups.

4. Each pick is not independent, since sampling is without replacement.

Hypergeometric Probability

a discrete random variable (RV) that is characterized by:

1. A fixed number of trials.
2. The probability of success is not the same from trial to trial.

We sample from two groups of items when we are interested in only one group. *X* is defined as the number of successes out of the total number of items chosen.

Poisson Probability Distribution

a discrete random variable (RV) that counts the number of times a certain event will occur in a specific interval; characteristics of the variable:

- The probability that the event occurs in a given interval is the same for all intervals.
- The events occur with a known mean and independently of the time since the last event.

The distribution is defined by the mean μ of the event in the interval. The mean is $\mu = np$. The standard deviation is $\sigma = \sqrt{\mu}$. The probability of having exactly x successes in r trials is $P(x) = \dfrac{\mu^x e^{-\mu}}{x!}$. The Poisson distribution is often used to approximate the binomial distribution, when n is "large" and p is "small" (a general rule is that np.

Probability Distribution Function (PDF)

a mathematical description of a discrete random variable (*RV*), given either in the form of an equation (formula) or in the form of a table listing all the possible outcomes of an experiment and the probability associated with each outcome.

Random Variable (RV)

a characteristic of interest in a population being studied; common notation for variables are upper case Latin letters *X*, *Y*, *Z*,...; common notation for a specific value from the domain (set of all possible values of a variable) are lower case Latin letters *x*, *y*, and *z*. For example, if *X* is the number of children in a family, then *x* represents a specific integer 0, 1, 2, 3,.... Variables in statistics differ from variables in intermediate algebra in the two following ways.

- The domain of the random variable (RV) is not necessarily a numerical set; the domain may be expressed in words; for example, if *X* = hair color then the domain is {black, blond, gray, green, orange}.
- We can tell what specific value *x* the random variable *X* takes only after performing the experiment.