Chapter 6

Binomial coefficient The number of ways of arranging k successes among n observations is given by the binomial coefficient $\binom{n}{k} = \frac{n!}{k!(n-k)!}$ for k = 0, 1, 2, … n where $n! = n(n-1)(n-2)\cdot\ldots\cdot 3\cdot 2\cdot 1$ and 0! = 1.

Binomial distribution In a binomial setting, suppose we let X = the number of successes. The probability distribution of X is a binomial distribution with parameters n and p, where n is the number of trials of the chance process and p is the probability of a success on any one trial. The possible values of X are the whole numbers from 0 to n.

Binomial probability If X has the binomial distribution with n trials and probability p of success on each trial, the possible values of X are 0, 1, 2, …, n. If k is any one of these values, $P(X = k) = \binom{n}{k} p^k (1-p)^{n-k}$.

Binomial random variable The count X of successes in a binomial setting.

Binomial setting Arises when we perform several independent trials of the same chance process and record the number of times that a particular outcome occurs. The four conditions for a binomial setting are:
- **B**inary? The possible outcomes of each trial can be classified as "success" or "failure."
- **I**ndependent? Trials must be independent; that is, knowing the result of one trial must not have any effect on the result of any other trial.
- **N**umber? The number of trials n of the chance process must be fixed in advance.
- **S**uccess? On each trial, the probability p of success must be the same.

Continuous random variable Takes all values in an interval of numbers. The probability distribution of a continuous random variable is described by a density curve. The probability of any event is the area under the density curve and above the values of the variable that make up the event.

Discrete random variable Takes a fixed set of possible values with gaps between. The probability distribution of a discrete random variable gives its possible values and their probabilities. The probability of any event is the sum of the probabilities for the values of the variable that make up the event.

Factorial For any positive whole number n, its factorial $n!$ is

$$n! = n \cdot (n-1) \cdot (n-2) \cdot \ldots \cdot 3 \cdot 2 \cdot 1$$

In addition, we define 0! = 1.

Geometric distribution In a geometric setting, suppose we let Y = the number of trials required to get the first success. The **probability distribution** of Y is a geometric distribution with parameter p, the probability of a success on any trial. The possible values of Y are 1, 2, 3, ….

Geometric random variable The number of trials Y that it takes to get a success in a geometric setting.

Geometric setting A geometric setting arises when we perform independent trials of the same chance process and record the number of trials until a particular outcome occurs. The four conditions for a geometric setting are:
• **B**inary? The possible outcomes of each trial can be classified as "success" or "failure."
• **I**ndependent? Trials must be independent; that is, knowing the result of one trial must not have any effect on the result of any other trial.
• **T**rials? The goal is to count the number of trials until the first success occurs.
• **S**uccess? On each trial, the probability p of success must be the same.

Geometric probability If Y has the geometric distribution with probability p of success on each trial, the possible values of Y are 1, 2, 3, …. If k is any one of these values, $P(Y=k)=(1-p)^{k-1}\cdot p$.

Independent random variables If knowing whether any event involving X alone has occurred tells us nothing about the occurrence of any event involving Y alone, and vice versa, then X and Y are independent random variables. That is, there is no association between the values of one variable and the values of the other.

Linear transformation A linear transformation of a random variable involves adding a constant a, multiplying by a constant b, or both. We can write a linear transformation of the random variable X in the form $Y = a + bX$. The shape, center, and spread of the probability distribution of Y are as follows:
Shape: Same as the probability distribution of X.
Center: $\mu_Y = a + b\mu_X$
Spread: $\sigma_Y = |b|\sigma_X$

Mean (expected value) of a random variable The mean of a random variable X, denoted by μ_X, is the balance point of the probability distribution histogram or density curve. Since the mean is the long-run average value of the variable after many repetitions of the chance process, it is also known as the expected value $E(X)$ of the random variable.

Mean (expected value) of a discrete random variable *X* To find the mean (expected value) of X, multiply each possible value by its probability, then add all the products:
$$\mu_X = E(X) = x_1p_1 + x_2p_2 + x_3p_3 + \ldots$$
$$= \sum x_i p_i$$

Mean of the sum (difference) of random variables For any two random variables *X* and *Y*, if $T = X + Y$ then the mean of *T* is $\mu_T = \mu_X + \mu_Y$ If $D = X - Y$, then the mean of *D* is $\mu_D = \mu_X - \mu_Y$. In general, the mean of the sum (difference) of several random variables is the sum (difference) of their means.

Mean (expected value) of a geometric random variable If *Y* is a geometric random variable with probability of success *p* on each trial, then its mean (expected value) is $\mu_Y = E(Y) = \frac{1}{p}$. That is, the expected number of trials required to get the first success is $1/p$.

Mean and standard deviation of a binomial random variable If a count *X* has the binomial distribution with number of trials *n* and probability of success *p*, the mean and standard deviation of *X* are

$$\mu_X = np$$

$$\sigma_X = \sqrt{np(1-p)}$$

Normal approximation for binomial distributions If *X* is a count having the binomial distribution with parameters *n* and *p*, then when *n* is large, *X* is approximately Normally distributed with mean *np* and standard deviation $\sqrt{np(1-p)}$. We will use this approximation when $np \geq 10$ and $n(1-p) \geq 10$.

Probability distribution The probability distribution of a random variable gives its possible values and their probabilities.

Random Variable Takes numerical values that describe the outcomes of some chance process.

Standard deviation of a random variable The square root of the variance of a random variable σ_X^2. The standard deviation measures the variability of the distribution about the mean.

Variance of a random variable σ_X^2 The average squared deviation of the values of the variable from their mean.

Variance of the sum (difference) of independent random variables For any two *independent* random variables *X* and *Y*, if $T = X + Y$, then the variance of *T* is $\sigma_T^2 = \sigma_X^2 + \sigma_Y^2$. If $D = X - Y$, then the variance of *D* is $\sigma_D^2 = \sigma_X^2 + \sigma_Y^2$