Chapter 5: Random Variables

November 27, 2019

Section 5.1

Random Variable

A random variable is a quantitative variable whose value depends on chance.

Discrete Random Variable

A discrete random variable is a random variable with only finite number of possible values, it usually involves a count of something.

Probability Distribution and Probability Histogram

The probability distribution and probability histogram of a discrete variable show its possible values and their likelihood.

- Probability distribution: A listing of the possible values and corresponding probabilities of discrete random variable, or a formula for the probabilities.
- Probability histogram: A graph of the probability distribution that displays the possible values of a discrete random variable on the horizontal axis and the probability of those values on the vertical axis. The probability of each value is represented by a vertical bar whose height equals the probability.

Sum of the Probability of a Discrete Random Variable

For any discrete random variable X, we have $\sum P(X = x) = 1$. The sum of the probabilities of the possible values of a discrete random variable equals 1.

Interpretation of a Probability Distribution

In a large number of independent observations of a random variable X, the proportion of times each possible value occurs will approximate the probability distribution of X; or, equivalently, the proportion histogram will approximate the probability histogram for X.

Section 5.2

Mean of a Discrete Random Variable (Expectation)

The mean of a discrete random variable X is denoted μ_x or, when no confusion will arise, simply μ. It is defined by

$$\mu = \sum xP(X = x)$$

The terms expected value and expectation are commonly used in place of the term mean.

Interpretation of the Mean of a Random Variable (Expectation)

In a large number of independent observations of a random variable X, the average value of those observations will approximately equal the mean, μ, of X. The larger the number of observations, the closer the average tends to be to μ.

Standard Deviation of a Discrete Random Variable

The standard deviation of a discrete random variable X is denoted σ_X or, when no confusion will arise, simply σ. It is defined as

$$\sigma = \sqrt{\sum (x - \mu)^2 P(X = x)}$$

The standard deviation of a discrete random variable can also be obtained from the computing formula

$$\sigma = \sqrt{\sum x^2 P(X = x) - \mu^2}$$

Roughly speaking, the standard deviation of a random variable X indicates how far, on average, an observed value of X is from its mean. In particular, the smaller the standard deviation of X, the more likely it is that an observed value of X will be close to its mean.

Variance of X

The square of the standard deviation, σ^2 is called the variance of X. It is defined as

$$\sigma^2 = \sum (x-\mu)^2 P(X=x)$$

The computing formula for the variance of X is defined as

$$\sigma^2 = \sum x^2 P(X=x) - \mu^2$$

Section 5.3

Binomial Coefficients

If n is a positive integer and x is a nonnegative integer less than or equal to n, then the binomial coefficient $\binom{n}{k}$ is defined as

$$\binom{n}{k} = \frac{n!}{k!(n-k)!}$$

The binomial coefficient is equivalent to ${}_nC_k$ and is often termed "n choose k".

Bernoulli Trials

Bernoulli trials are identical and independent repetitions of an experiment with two possible outcomes. Repeated trials of an experiment are called Bernoulli trials if the following three conditions are satisfied:

1. The experiment (each trial) has two possible outcomes, denoted generically s, for success, and f, for failure.

2. The trials are independent, meaning that the outcome of one trial does not affect the outcome of other trials.

3. The probability of a success, called the success probability and denoted p, remains the same from trial to trial.

Binomial Distribution

The binomial distribution is the probability distribution for the number of successes in a sequence of Bernoulli trials.

Number of Outcomes Containing a Specified Number of Successes

In n Bernoulli trials, the number of outcomes that contain exactly x successes equals the binomial coefficient $\binom{n}{k}$. In other words, there are $\binom{n}{k}$ ways of getting exactly x successes in n Bernoulli trials.

Binomial Probability Formula

Let X denote the total number of successes in n Bernoulli trials with success probability p. The the probability distribution of the random variable X is given by

$$P(X = x) = \binom{n}{x} p^x (1-p)^{n-x}, x \in \mathbb{N}$$

Procedure to find a Binomial Probability Formula.

Assumptions:

1. n trials are to be performed.
2. Two outcomes, success or failure, are possible for each trial.
3. The trials are independent.
4. The success probability, p, remains the same from trial to trial.

Step 1 Identify a success.

Step 2 Determine p, the success probability.

Step 3 Determine n, the number of trials.

Step 4 The binomial probability formula for the number of successes, X, is

$$P(X = x) = \binom{n}{x} p^x (1-p)^{n-x}.$$

Mean and Standard Deviation of a Binomial Random Variable

The mean and standard deviation of a binomial random variable with parameters n and p are

$$\mu = npand\sigma = \sqrt{np(1-p)},$$

respectively.

Sampling and the Binomial Distribution

Suppose that a simple random sample of size n is taken from a finite population in which the proportion of members that have a specified attribute

- has exactly a binomial distribution with parameters n and p if the sampling is done with replacement and
- has approximately a binomial distribution with parameters n and p if the sampling is done without replacement and the sample size does not exceed 5% of the population size.

When a simple random sample is taken from a finite population, you can use a binomial distribution for the number of members obtained having a specified attribute, regardless of whether the sampling is with or without replacement, provided that, in the latter case, the sample size is small relative to population size.

Section 5.4

Poisson Probability Formula

Probabilities for a random variable X that has a Poisson distribution are given by the formula

$$P(X = x) = e^{-\lambda}\frac{\lambda^x}{x!}, x \in \mathbb{N}$$

where λ is a positive real number and e is approximately 2.718. The random variable X is called a Poisson random variable and is said to have the Poisson distribution with parameter λ. A Poisson random variable can be any integer value. Consequently, we cannot display all the probabilities for a Poisson random variable in a probability distribution table.

Mean and Standard Deviation of a Poisson Random Variable

The mean and standard deviation of a Poisson random variable with parameter λ are

$$\mu = \lambda and \sigma = \sqrt{\lambda}$$

respectively.

To Approximate Binomial Probabilities by Using a Poisson Probability Formula

Step 1 Find n, the number of trials, and p, the success probability.

Step 2 Continue only if $n \geq 100$ and $np \leq 10$.

Step 3 approximate the required binomial probabilities by using the Poisson probability formula

$$P(X = x) = e^{-np}\frac{(np)^x}{x!}.$$