Lecture 11

Bernoulli and Binomial Distributions

Text: Chapter 4

STAT 8010 Statistical Methods I
September 13, 2019

Whitney Huang
Clemson University

Agenda

1. Bernoulli Trials

2. Bernoulli and Binomial Distributions

Bernoulli Trials

Many problems in probability and its applications involve independently repeating a random experiment and observing at each repetition whether a specified event occurs. We label the occurrence of the specified event a success and the nonoccurrence of the specified event a failure.

Example:
Tossing a coin several times

Bernoulli Trials Cont'd

Bernoulli trials:

- Each repetition of the random experiment is called a trial
- We use p to denote the probability of a success on a single trial

Properties of Bernoulli trials:

- Exactly two possible outcomes success and failure
- The outcomes of trials are independent of one another
- The success probability, p, and therefore the failure probability, $(1 - p)$, remains the same from trial to trial

Bernoulli Random Variable

Characteristics of the Bernoulli random variable:
Let X be a Bernoulli r.v.

Bernoulli Random Variable

Characteristics of the Bernoulli random variable:
Let X be a Bernoulli r.v.

- The definition of X: The number of successes in a single trial of a random experiment

Bernoulli Random Variable

Characteristics of the Bernoulli random variable:
Let X be a Bernoulli r.v.

- The definition of X: The number of successes in a single trial of a random experiment
- The support (possible values for X): 0: “failure" or 1: “success"

Bernoulli Random Variable

Characteristics of the Bernoulli random variable:
Let X be a Bernoulli r.v.

- The definition of X: The number of successes in a single trial of a random experiment
- The support (possible values for X): 0: “failure" or 1: “success"
- Its parameter and definition: p: the probability of success on 1 trial

Bernoulli Random Variable

Characteristics of the Bernoulli random variable:
Let X be a Bernoulli r.v.

- The definition of X: The number of successes in a single trial of a random experiment
- The support (possible values for X): 0: “failure" or 1: “success"
- Its parameter and definition: p: the probability of success on 1 trial
- The probability mass function (pmf):

$$p_X(x) = p^x(1-p)^{1-x}, \qquad x = 0, 1$$

Bernoulli Random Variable

Characteristics of the Bernoulli random variable:
Let X be a Bernoulli r.v.

- The definition of X: The number of successes in a single trial of a random experiment
- The support (possible values for X): 0: "failure" or 1: "success"
- Its parameter and definition: p: the probability of success on 1 trial
- The probability mass function (pmf):

$$p_X(x) = p^x(1-p)^{1-x}, \qquad x = 0, 1$$

- The expected value:

$$\mathbb{E}[X] = 0 \times (1-p) + 1 \times p = p$$

Bernoulli Random Variable

Characteristics of the Bernoulli random variable:
Let X be a Bernoulli r.v.

- The definition of X: The number of successes in a single trial of a random experiment
- The support (possible values for X): 0: "failure" or 1: "success"
- Its parameter and definition: p: the probability of success on 1 trial
- The probability mass function (pmf):

$$p_X(x) = p^x(1-p)^{1-x}, \qquad x = 0, 1$$

- The expected value:

$$\mathbb{E}[X] = 0 \times (1-p) + 1 \times p = p$$

- The variance:

$$Var(X) = \mathbb{E}[X^2] - (\mathbb{E}[X])^2 = p - (p)^2 = p(1-p)$$

Binomial Random Variable

We can define the Binomial r.v. as the number of successes in n Bernoulli trials, where the probability of success in one trial is p. Let X be a Binomial r.v.

- The definition of X: The number of successes in n trials of a random experiment, where sampling is done with replacement (or trials are independent)

Binomial Random Variable

We can define the Binomial r.v. as the number of successes in n Bernoulli trials, where the probability of success in one trial is p. Let X be a Binomial r.v.

- The definition of X: The number of successes in n trials of a random experiment, where sampling is done with replacement (or trials are independent)
- The support: $0, 1, \cdots, n$

Binomial Random Variable

We can define the Binomial r.v. as the number of successes in n Bernoulli trials, where the probability of success in one trial is p. Let X be a Binomial r.v.

- The definition of X: The number of successes in n trials of a random experiment, where sampling is done with replacement (or trials are independent)
- The support: $0, 1, \cdots, n$
- Its parameter(s) and definition(s): p: the probability of success on 1 trial and n is the sample size

Binomial Random Variable

We can define the Binomial r.v. as the number of successes in n Bernoulli trials, where the probability of success in one trial is p. Let X be a Binomial r.v.

- The definition of X: The number of successes in n trials of a random experiment, where sampling is done with replacement (or trials are independent)
- The support: $0, 1, \cdots, n$
- Its parameter(s) and definition(s): p: the probability of success on 1 trial and n is the sample size
- The probability mass function (pmf):

$$p_X(x) = \binom{n}{x} p^x (1-p)^{n-x}, \; x = 0, 1, \cdots, n$$

Binomial Random Variable

We can define the Binomial r.v. as the number of successes in n Bernoulli trials, where the probability of success in one trial is p. Let X be a Binomial r.v.

- The definition of X: The number of successes in n trials of a random experiment, where sampling is done with replacement (or trials are independent)
- The support: $0, 1, \cdots, n$
- Its parameter(s) and definition(s): p: the probability of success on 1 trial and n is the sample size
- The probability mass function (pmf):

$$p_X(x) = \binom{n}{x} p^x (1-p)^{n-x}, \; x = 0, 1, \cdots, n$$

- The expected value:

$$\mathbb{E}[X] = np$$

Binomial Random Variable

We can define the Binomial r.v. as the number of successes in n Bernoulli trials, where the probability of success in one trial is p. Let X be a Binomial r.v.

- The definition of X: The number of successes in n trials of a random experiment, where sampling is done with replacement (or trials are independent)
- The support: $0, 1, \cdots, n$
- Its parameter(s) and definition(s): p: the probability of success on 1 trial and n is the sample size
- The probability mass function (pmf):

$$p_X(x) = \binom{n}{x} p^x (1-p)^{n-x}, \; x = 0, 1, \cdots, n$$

- The expected value:

$$\mathbb{E}[X] = np$$

- The variance:

$$Var(X) = np(1-p)$$

Example

To test for Extrasensory perception (ESP), we have 4 cards. They will be shuffled and one randomly selected each time, and you are to guess which card is selected. This is repeated 10 times. Suppose you do not have ESP. Let R be the number of times you guess a card correctly. What are the distribution and parameter(s) of R? What is the expected value of R? Furthermore, suppose that you get certified as having ESP if you score at least an 8 on the test. What is the probability that you get certified as having ESP?

Example

In Whitney's Stat 8020 class, 80% of the students passed (got a C or better) on Exam 1. If you were to pick a student at random and asked them whether or not they passed. Let X represent the number of student who passed.

1. What type of random variable is this? How do you know? Additionally, write down the pmf, the expected value, and the variance for X
2. What about if you picked 10 students with replacement and let Y be the number of student(s) who passed. What type of random variable is this? Write down the pmf, the expected value, and the variance for Y

Example

Suppose that 95% of consumers can recognize Coke in a blind taste test. Assume consumers are independent of one another. The company randomly selects 4 consumers for a taste test. Let X be the number of consumers who recognize Coke.

1. Write out the pmf table for X
2. What is the probability that X is at least 1?
3. What is the probability that X is at most 3?

Binomial and Hypergeometric Distributions

The binomial distribution describes the probability of k successes in n trials with replacement.

We want a distribution to describe the probability of k successes in n trials without replacement from a finite population of size N containing exactly K successes.

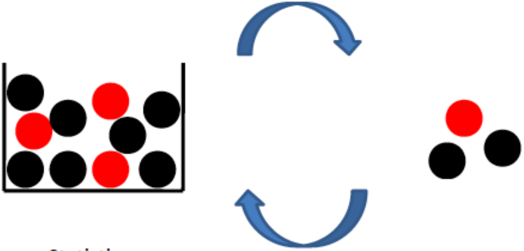
$\Rightarrow$ Hypergeometric Distribution

Important applications are **quality control** and statistical **estimation of population proportions**. The hypergeometric r.v. is the equivalent of a Binomial r.v. except that sampling is done without replacement.

An Example of Hypergeometric r.v.

<u>Probability</u>:

What is the probability to get 1 red and 2 black balls?



<u>Statistics:</u>

What percentage of balls in the box are red?

Hypergeometric Distributions

Let X be a hypergeometric r.v.

- The definition of X: It is the number of successes in n trials of a random experiment, where sampling is done without replacement (or trials are dependent)

Hypergeometric Distributions

Let X be a hypergeometric r.v.

- The definition of X: It is the number of successes in n trials of a random experiment, where sampling is done without replacement (or trials are dependent)
- The support: $k \in \{\max(0, n+K-N), \cdots, \min(n, K)\}$

Hypergeometric Distributions

Let X be a hypergeometric r.v.

- The definition of X: It is the number of successes in n trials of a random experiment, where sampling is done without replacement (or trials are dependent)
- The support: $k \in \{\max(0, n + K - N), \cdots, \min(n, K)\}$
- Its parameter(s) and definition(s): N: the population size, n: the sample size, and K: number of success in the population

Hypergeometric Distributions

Let X be a hypergeometric r.v.

- The definition of X: It is the number of successes in n trials of a random experiment, where sampling is done without replacement (or trials are dependent)
- The support: $k \in \{\max(0, n+K-N), \cdots, \min(n, K)\}$
- Its parameter(s) and definition(s): N: the population size, n: the sample size, and K: number of success in the population
- The probability mass function (pmf): $p_X(x) = \frac{\binom{K}{k} \times \binom{N-K}{n-k}}{\binom{N}{n}}$

Hypergeometric Distributions

Let X be a hypergeometric r.v.

- The definition of X: It is the number of successes in n trials of a random experiment, where sampling is done without replacement (or trials are dependent)
- The support: $k \in \{\max(0, n+K-N), \cdots, \min(n, K)\}$
- Its parameter(s) and definition(s): N: the population size, n: the sample size, and K: number of success in the population
- The probability mass function (pmf): $p_X(x) = \frac{\binom{K}{k} \times \binom{N-K}{n-k}}{\binom{N}{n}}$
- The expected value: $\mathbb{E}[X] = n\frac{K}{N}$

Hypergeometric Distributions

Let X be a hypergeometric r.v.

- The definition of X: It is the number of successes in n trials of a random experiment, where sampling is done without replacement (or trials are dependent)
- The support: $k \in \{\max(0, n+K-N), \cdots, \min(n, K)\}$
- Its parameter(s) and definition(s): N: the population size, n: the sample size, and K: number of success in the population
- The probability mass function (pmf): $p_X(x) = \frac{\binom{K}{k} \times \binom{N-K}{n-k}}{\binom{N}{n}}$
- The expected value: $\mathbb{E}[X] = n\frac{K}{N}$
- The variance: $Var(X) = n\frac{K}{N}\frac{N-K}{N}\frac{N-n}{N-1}$

Example

There are 100 identical looking 52" TVs at Best Buy in Anderson, SC. Let 10 of them be defective. Suppose we want to buy 8 of the aforementioned TVs (at random). What is the probability that we don't get any defective TVs?

Solution.

Example

There are 100 identical looking 52" TVs at Best Buy in Anderson, SC. Let 10 of them be defective. Suppose we want to buy 8 of the aforementioned TVs (at random). What is the probability that we don't get any defective TVs?

Solution.

Let D be the number of defective TVs in the sample.

$D \sim Hyp(N = 100, n = 8, K = 10)$

$\mathbb{P}(D = 0) = \frac{\binom{10}{0}\binom{90}{8}}{\binom{100}{8}} = 0.4166$