Lecture 10

Mean and Variance of Discrete Random Variables

Text: Chapter 4

STAT 8010 Statistical Methods I

September 11, 2019

Whitney Huang
Clemson University

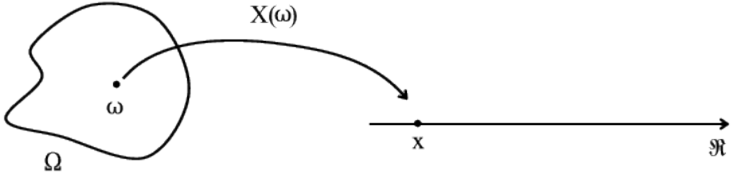
Agenda

Random Variables (r.v.)

A random variable is a real–valued function whose domain is the sample space of a random experiment. In other words, a random variable is a function

$$X : \Omega \mapsto \mathbb{R}$$

where Ω is the sample space of the random experiment under consideration and $\mathbb{R}$ represents the set of all real numbers.



A random variable X that can take on at most a countable number of possible values is said to be a discrete random variable $\Rightarrow$ use the probability mass function to describe X.

Probability Mass Function

Let X be a discrete random variable. Then the probability mass function (pmf) of X is the real–valued function defined on $\mathbb{R}$ by

$$p_X(x) = \mathbb{P}(X = x)$$

The capital letter, X, is used to denote random variable. Lowercase letter, x, is used to denote possible values of the random variable.

Example

Flip a fair coin 3 times. Let X denote the number of heads tossed in the 3 flips. Create a pmf for X

Solution.

The random variable X maps any outcome to an integer (e.g. $X((\text{T, T, T})) = 0$)

x	0	1	2	3
$p_X(x)$	$\frac{1}{8}$	$\frac{3}{8}$	$\frac{3}{8}$	$\frac{1}{8}$

Properties of a PMF

- $0 \leq p_X(x) \leq 1, \ \forall x \in R$

Properties of a PMF

- $0 \leq p_X(x) \leq 1, \ \forall x \in R$

Properties of a PMF

- $0 \leq p_X(x) \leq 1, \ \forall x \in R$
- $\{x \in \mathbb{R} : p_X(x) \neq 0\}$ is countable

Properties of a PMF

- $0 \leq p_X(x) \leq 1, \ \forall x \in R$
- $\{x \in \mathbb{R} : p_X(x) \neq 0\}$ is countable

Properties of a PMF

- $0 \leq p_X(x) \leq 1, \ \forall x \in R$
- $\{x \in \mathbb{R} : p_X(x) \neq 0\}$ is countable
- $\sum_x p_X(x) = 1$

Example

Let X be a random variable with pmf defined as follows:

$$p_X(x) = \begin{cases} k(5-x) & \text{if } \quad x = 0,1,2,3,4 \\ 0 & \text{otherwise} \end{cases}$$

1. Find the value of k that makes $p_X(x)$ a legitimate pmf.
2. What is the probability that X is between 1 and 3 inclusive?
3. If X is not 0, what is the probability that X is less than 3?

Mean of Discrete Random Variables

The mean of a discrete r.v. X, denoted by $\mathbb{E}[X]$, is defined by

$$\mathbb{E}[X] = \sum_x x \times p_X(x)$$

Remark:

The mean of a discrete r.v. is a weighted average of its possible values, and the weight used is its probability. Sometimes we refer to the expected value as the expectation (expected value), or the first moment.

For any function, say $g(X)$, we can also find an expectation of that function. It is

$$\mathbb{E}[g(X)] = \sum_x g(x) \times p_X(x)$$

Example

$$\mathbb{E}[X^2] = \sum_x x^2 \times p_X(x)$$

Properties of Mean

Let X and Y be discrete r.v.s defined on the same sample space and having finite expectation (i.e. $\mathbb{E}[X], \mathbb{E}[Y] < \infty$). Let a and b be constants. Then the following hold:

- $\mathbb{E}[X + Y] = \mathbb{E}[X] + \mathbb{E}[Y]$
- $\mathbb{E}[aX + b] = a \times \mathbb{E}[X] + b$

Number of Siblings Example Revisited

Siblings (X)	Frequency	Relative Frequency
0	8	.200
1	17	.425
2	11	.275
3	3	.075
4	1	.025
Total	40	1

Find the expected value of the number of siblings

Solution.

Number of Siblings Example Revisited

Siblings (X)	Frequency	Relative Frequency
0	8	.200
1	17	.425
2	11	.275
3	3	.075
4	1	.025
Total	40	1

Find the expected value of the number of siblings

Solution.

$\mathbb{E}[X] = \sum_x x p_X(x) =$
$0 \times .200 + 1 \times .425 + 2 \times .275 + 3 \times .075 + 4 \times .025 = 1.3$

Variance/standard deviation of Discrete r.v.'s

The **variance** of a (discrete) r.v., denoted by $Var(X)$, is a measure of the spread, or variability, in the r.v. $Var(X)$ is defined by

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

or

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

The **standard deviation**, denoted by $Sd(X)$, is the square root of its variance

Properties of variance

Let c be a constant. Then the following hold:

- $Var(cX) = c^2 \times Var(X)$
- $Var(X + c) = Var(X)$

Example

Suppose X and Y are random variables with $\mathbb{E}[X] = 3$, $\mathbb{E}[Y] = 4$ and $Var(X) = 4$. Find:

1. $\mathbb{E}[2X + 1]$
2. $\mathbb{E}[X - Y]$
3. $\mathbb{E}[X^2]$
4. $\mathbb{E}[X^2 - 4]$
5. $\mathbb{E}[(X - 4)^2]$
6. $Var(2X - 4)$

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?