Probability and Statistics

Nels Grevstad

Metropolitan State University of Denver

ngrevsta@msudenver.edu

February 10, 2019

Topics

Objectives

Objectives:

- Distinguish discrete from continuous random variables
- Use discrete probability distributions to find probabilities
- For discrete random variables, compute and interpret:
  - The expected value
  - The expected value of a function of the random variable
  - The variance and standard deviation
  - The variance and standard deviation of a linear function of the random variable

Random Variables (3.1)

- Numerical values that are determined by **chance** are modeled as **random variables**.

Random Variables (3.1)

- Numerical values that are determined by **chance** are modeled as **random variables**.
- Random variables are denoted by capital letters X, Y, Z, etc.

Example

If we toss a coin, the **sample space** is

$$\mathcal{S} = \{H, T\}.$$

Let

$$X = \begin{cases} 1 & \text{if the outcome is } H \\ 0 & \text{if the outcome is } T \end{cases}$$

Then X is a **random variable.**

Example

Randomly select a person from a population. Then the **sample space** consists of the individuals in the population:

$$
\begin{aligned}
S &= \{\text{Stephanie Lawson}, \\
&\quad \text{Jeffrey Miller}, \\
&\quad \text{Angela DuPont}, \\
&\quad \vdots \\
&\quad \text{Karl Stevenson}\}
\end{aligned}
$$

Example

Randomly select a person from a population. Then the **sample space** consists of the individuals in the population:

$$S = \{\text{Stephanie Lawson}, \text{Jeffrey Miller}, \text{Angela DuPont}, \vdots, \text{Karl Stevenson}\}$$

Now let

$$X = \text{The selected person's age}$$

Then X is a **random variable**.

Example

Randomly select a person from a population. Then the **sample space** consists of the individuals in the population:

$$\begin{aligned} S &= \{\text{Stephanie Lawson} & 48, \\ & \quad \text{Jeffrey Miller} & 28, \\ & \quad \text{Angela DuPont} & 27, \\ & \quad \vdots & \vdots \\ & \quad \text{Karl Stevenson} & 34\} \end{aligned}$$

Now let

$$X = \text{The selected person's age}$$

Then X is a **random variable**.

Example

Consider rolling two dice, one red one and the other green. The **sample space** consists of the 36 outcomes:

		Number on Green Die					
		1	2	3	4	5	6
	1	(1,1)	(1,2)	(1,3)	(1,4)	(1,5)	(1,6)
	2	(2,1)	(2,2)	(2,3)	(2,4)	(2,5)	(2,6)
Number on Red Die	3	(3,1)	(3,2)	(3,3)	(3,4)	(3,5)	(3,6)
	4	(4,1)	(4,2)	(4,3)	(4,4)	(4,5)	(4,6)
	5	(5,1)	(5,2)	(5,3)	(5,4)	(5,5)	(5,6)
	6	(6,1)	(6,2)	(6,3)	(6,4)	(6,5)	(6,6)

Let

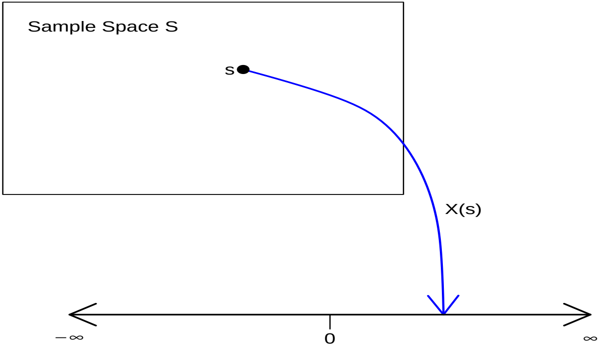
$$X = \text{The sum of the two numbers on the dice.}$$

Then X is a **random variable.**

- A **random variable** arises when a numerical value is associated with each outcome in the sample space $\mathcal{S}$.

- A **random variable** arises when a numerical value is associated with each outcome in the sample space $\mathcal{S}$.
- Formally, a ***random variable*** is a real-valued function whose domain is $\mathcal{S}$.

- A **random variable** arises when a numerical value is associated with each outcome in the sample space $\mathcal{S}$.
- Formally, a ***random variable*** is a real-valued function whose domain is $\mathcal{S}$.



- A random variable is ***discrete*** if its set of possible values is finite or countably infinite.

- A random variable is ***discrete*** if its set of possible values is finite or countably infinite.

  It's ***continuous*** if its set of possible values is a continuous interval.

- A random variable is ***discrete*** if its set of possible values is finite or countably infinite.

  It's ***continuous*** if its set of possible values is a continuous interval.

- The ***probability distribution*** of a random variable specifies:
  1. The set of possible values for the variable.
  2. The probabilities of those values.

Discrete Probability Distributions (3.2)

Example

The table below shows the vehicle occupancy rates in Miami-Dade County, Florida.

Number of Occupants	Percentage of Vehicles
1	82 %
2	12 %
3	4 %
4	2 %

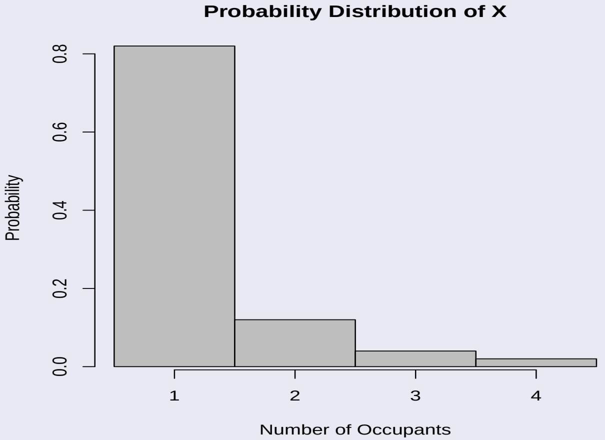
Consider a randomly selected vehicle, and let

$$X = \text{The number of occupants in the vehicle}$$

X is a **discrete** random variable whose **probability distribution** is below.

x	1	2	3	4
$p(x)$	0.82	0.12	0.04	0.02

Here's a graph of the **probability distribution** of X.



- In general, the probability distribution of a **discrete** random variable is represented by a ***probability mass function*** (or ***pmf***), denoted $p(x)$ and defined as

$$p(x) \;=\; P(X \;=\; x)$$

- In general, the probability distribution of a **discrete** random variable is represented by a ***probability mass function*** (or ***pmf***), denoted $p(x)$ and defined as

Particular value $\boldsymbol{x}$

$$p(x) \ = \ P(X \ = \ x)$$

- In general, the probability distribution of a **discrete** random variable is represented by a ***probability mass function*** (or ***pmf***), denoted $p(x)$ and defined as

Particular value $\boldsymbol{x}$

$$p(x) \;=\; P(X \;=\; x)$$

Random variable $\boldsymbol{X}$

Example (Cont'd)

Here's the **pmf** for the random variable

$$X = \text{The number of occupants in a vehicle}$$

x	1	2	3	4
$p(x)$	0.82	0.12	0.04	0.02

For example, that the probability of a randomly selected vehicle having only one occupant is

$$p(1) = P(X = 1) = 0.82.$$

- In order for a **pmf** to be legitimate, it must satisfy the following conditions:

  1. $p(x) \geq 0$ for all x.
  2. $\sum p(x) = 1$.

     where the summation is over all possible values x of X.

Expected Values (3.3)

- The ***expected value*** of a **discrete** random variable X, also called the ***mean*** of its distribution, is denoted $\boldsymbol{E(X)}$ or $\boldsymbol{\mu_X}$ and defined as:

Expected Value:

$$E(X) = \mu_X = \sum x\, p(x)$$

where the summation is over all possible values x of X.

- $E(X)$ is a **weighted average** of the possible values x of X.

- The **expected value** (or **mean**) has a few interpretations:
  - It's the **long-run average** value of X.

- The **expected value** (or **mean**) has a few interpretations:
  - It's the **long-run average** value of X.
  - It's the **center** ("balancing point") of the probability distribution.

- The **expected value** (or **mean**) has a few interpretations:
  - It's the **long-run average** value of X.
  - It's the **center** ("balancing point") of the probability distribution.
- Later, we'll use probability distributions to represent **populations.** The expected value will be the **population mean.**

Example

When a roulette wheel is spun, the ball is equally likely to land in any of 38 slots, 18 of which are red, 18 black, and 2 green.

A bet of \$1.00 on red pays a dollar if the ball lands in a red slot. Otherwise you lose your dollar. Let

$$X = \text{Your winnings after a \$1.00 bet on red}$$

The probability distribution of X is:

x	-\$1.00	\$1.00
$p(x)$	$\frac{20}{38}$	$\frac{18}{38}$

The expected value of X is

$$E(X) = -1.00\left(\frac{20}{38}\right) + 1.00\left(\frac{18}{38}\right) = \mathbf{-0.053}.$$

Example

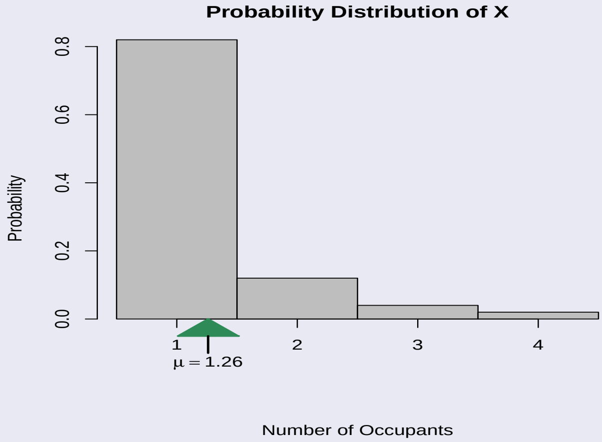
Recall that the probability distribution of the number of occupants X in a randomly selected vehicle is:

x	1	2	3	4
$p(x)$	0.82	0.12	0.04	0.02

The **expected value** is

$$E(X) = 1(0.82) + 2(0.12) + 3(0.04) + 4(0.02) = \mathbf{1.26}.$$

This is the **center** ("balancing point") of the distribution.



It's also the **population mean.**

- To see why μ is the **population mean**, suppose there are $N = 100,000$ vehicles in the population. Then using

$$\mu = \frac{1}{N}\sum_{i=1}^{N} x_i$$

gives

$$\begin{aligned}
\mu &= \frac{1}{100,000}(1 + 1 + \cdots + 1 && \text{(82,000 ones)}\\
&\qquad +2 + 2 + \cdots + 2 && \text{(12,000 twos)}\\
&\qquad +3 + 3 + \cdots + 3 && \text{(4,000 threes)}\\
&\qquad +4 + 4 + \cdots + 4) && \text{(2,000 fours)}\\
&= \frac{1(82,000) + 2(12,000) + 3(4,000) + 4(2,000)}{100,000}\\
&= 1(0.82) + 2(0.12) + 3(0.04) + 4(0.02) = \mathbf{1.26}.
\end{aligned}$$

- If X is a random variable, then any **function** $h(X)$ is also a random variable.

Proposition

If X is a discrete random variable with pmf $p(x)$, then the expected value of any function $h(X)$, denoted $\boldsymbol{E\left(h(X)\right)}$ or $\boldsymbol{\mu_{h(X)}}$, is computed by

$$E\left(h(X)\right) \;=\; \mu_{h(X)} \;=\; \sum h(x)p(x),$$

where the summation is over all possible values x of X.

Example

Suppose a random variable X has pmf given by

x	-2	-1	0	1
$p(x)$	0.4	0.3	0.2	0.1

and suppose we want the **expected value** of X^2.

Letting $h(X) = X^2$, we have

$h(x)$	$(-2)^2$	$(-1)^2$	0^2	1^2
$p(x)$	0.4	0.3	0.2	0.1

and

$$\begin{aligned} E(X^2) &= E(h(X)) \\ &= (-2)^2(0.4) + (-1)^2(0.3) + 0^2(0.2) + 1^2(0.1) \\ &= \mathbf{2.0}. \end{aligned}$$

- The next proposition can be derived from the previous one by setting $h(X) = aX + b$.

Proposition

If X is any random variable, then for any constants a and b,

$$E\left(aX + b\right) \;=\; aE(X) + b$$

(or, using alternative notation, $\mu_{aX\,+\,b} \;=\; a\mu_X + b$).

- Two special cases (for which $b = 0$ and $a = 1$):
  1. $E(aX) \;=\; aE(X)$.

- The next proposition can be derived from the previous one by setting $h(X) = aX + b$.

Proposition

If X is any random variable, then for any constants a and b,

$$E\left(aX + b\right) \;=\; aE(X) + b$$

(or, using alternative notation, $\mu_{aX\,+\,b} \;=\; a\mu_X + b$).

- Two special cases (for which $b = 0$ and $a = 1$):
  1. $E(aX) \;=\; aE(X)$.
  2. $E(X + b) \;=\; E(X) + b$.

- The ***variance*** and ***standard deviation*** of a **discrete** random variable X, denoted $\boldsymbol{V(X)}$ or $\boldsymbol{\sigma_X^2}$ and $\boldsymbol{SD(X)}$ or $\boldsymbol{\sigma_X}$, are defined as follows.

Variance and Standard Deviation:

$$\begin{aligned} V(X) &= \sigma_X^2 = E\left((X-\mu)^2\right) \\ &= \sum (x-\mu)^2\, p(x), \end{aligned}$$

where $\mu = E(X)$ and the summation is over all possible values x of X, and

$$SD(X) = \sigma_X = \sqrt{V(X)}.$$

- The **variance** is a **weighted average** of the **squared deviations** of X away from μ.

- The **variance** is a **weighted average** of the **squared deviations** of X away from μ.
- The **standard deviation** is interpreted as a **typical deviation** of X away from μ.

- The **variance** is a **weighted average** of the **squared deviations** of X away from μ.
- The **standard deviation** is interpreted as a **typical deviation** of X away from μ.
- Both are measures of the **variation** in X, that is, of the **spread** of the probability distribution of X.

- The **variance** is a **weighted average** of the **squared deviations** of X away from μ.
- The **standard deviation** is interpreted as a **typical deviation** of X away from μ.
- Both are measures of the **variation** in X, that is, of the **spread** of the probability distribution of X.
- They're the **population variance** and **population standard deviation** when the probability distribution represents a population.

Example

Consider a randomly selected rented housing unit in the U.S., and let

$$X = \text{The number of rooms in the unit}$$

The U.S. Census Bureau gives the **probability distribution** of X:

x	1	2	3	4	5	6	7	8
$p(x)$	0.01	0.03	0.25	0.35	0.20	0.10	0.04	0.02

The **mean** of this distribution is

$$\boldsymbol{\mu = 4.26}.$$

The **variance** of the distribution is

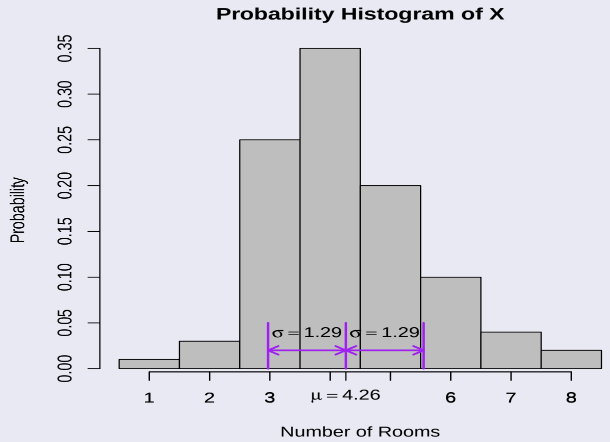
$$\begin{aligned}\boldsymbol{\sigma^2} &= (1-4.26)^2(0.01) + (2-4.26)^2(0.03) + (3-4.26)^2(0.25) + \\ &\quad \cdots + (8-4.26)^2(0.02) \\ &= \mathbf{1.67}\end{aligned}$$

so the **standard deviation** is

$$\boldsymbol{\sigma} = \sqrt{\mathbf{1.67}} = \mathbf{1.29}.$$

A typical rented housing unit has **4.26** rooms, on average, plus or minus about **1.29** rooms.

Also, μ and σ are the **population mean** and **population standard deviation** in the population of rented housing units.


Probability Histogram of X
Probability
0.00
0.05
0.10
0.15
0.20
0.25
0.30
0.35
σ = 1.29
σ = 1.29
1
2
3
μ = 4.26
6
7
8
Number of Rooms

- By expanding the square in the definition

$$V(X) = \sum (x-\mu)^2 p(x)$$

of a variance, we can derive the following.

Proposition

$$V(X) = E(X^2) - \mu^2$$

where $\mu = E(X)$.

- The variance of a **function** $h(X)$ is

$$V\,(h(X)) \;=\; E\left((h(X) - \mu_{h(X)})^2\right)$$

Setting $h(X) = aX + b$, we can derive the following.

Proposition

$$V(aX + b) \;=\; \sigma^2_{aX\,+\,b} \;=\; a^2\,\sigma^2_X$$

and so

$$SD(aX + b) \;=\; \sigma_{aX\,+\,b} \;=\; |a|\,\sigma_X.$$

- Two special cases of the previous proposition (for which $b = 0$ and $a = 1$):

1. $V(aX) \;=\; \sigma^2_{aX} \;=\; a^2\,\sigma^2_X$

   and

   $SD(aX) \;=\; \sigma_{aX} \;=\; |a|\,\sigma_X$.

- Two special cases of the previous proposition (for which $b = 0$ and $a = 1$):

1. $V(aX) \;=\; \sigma^2_{aX} \;=\; a^2\,\sigma^2_X$

   and

   $SD(aX) \;=\; \sigma_{aX} \;=\; |a|\,\sigma_X$.

2. $V(X+b) \;=\; \sigma^2_{X\,+\,b} \;=\; \sigma^2_X$

   and

   $SD(X+b) \;=\; \sigma_{X\,+\,b} \;=\; \sigma_X$.