# Discrete Random Variables
# Chapter 3

Kevin Lutz
PhD Candidate of Statistics
Wednesday, September 6, 2022

Department of Mathematical Sciences
The University of Texas at Dallas

# Table of Contents

# What is a Random Variable?

**Definition: Random Variable**

A random variable (RV) is a numerical outcome of a probability experiment where the outcome is determined randomly and by chance.

- We denote a RV using capitol letters: $X, Y, Z, \ldots$
- We denote a particular value of a RV using lower case letters: $x, y, z, \ldots$
  - The expression $\boxed{X = x}$ tells us the random variable $(X)$ and a particular outcome $(x)$.
- Mathematically, the RV is a rule or mapping from the sample space to a numerical outcome.

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

# Example

A coin is flipped 10 times. The results are:

$$\{TTHTHTTHHH\}$$

- The value of a RV can be thought of as the "answer to a question".
  1. Let $X =$ the number heads.

     $$X = 5$$

  2. Let $Y =$ the number of tosses until the first $H$ occurs.

     $$Y = 3$$

# Table of Contents

# Definition

## Discrete Random Variable

A discrete r.v. is a variable whose value is obtained by counting. The value of the random variable $X$ can only take on values from $\{0, 1, 2, 3, 4, \ldots\}$ and so $X$ is countably finite.

- $\mathbb{N}$ is the set of natural numbers: $\{0, 1, 2, 3, 4, \ldots\}$ so we say $X \in \mathbb{N}$ for any discrete r.v.
- Discrete can NOT be negative, decimals/fractions, $\pi$, etc.
- If a r.v. can take on *any* real number, then it is continuous (i.e., $X \in \mathbb{R}$).

# Examples

Discrete or Continuous?

| Random Variable | Answer |
|---|---|
| 1. Number of students in this class. | Discrete |
| 2. Number of people attending the World Series. | Discrete |
| 3. Today's temperature at UTD. | Continuous |
| 4. Amount of time spent working on HW1. | Continuous |
| 5. Number of heads when flipping a coin 20 times. | Discrete |
| 6. The distance or time it takes to get to school. | Continuous |

# Probability Distribution for a Discrete R.V.

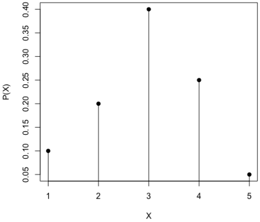
- Probability Distribution
  1. A table/list of the probabilities associated with each value of $X$.
  2. Called the "probability mass function" or p.m.f.
  3. Properties: The 3 Axioms!
  4. The probability that $X = x$ is notated as

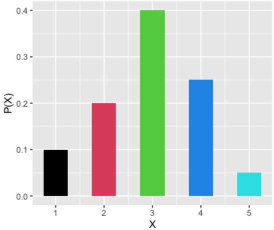
$$P(X = x) \text{ or } f_X(x)$$

- Example

| $X$ | 1 | 2 | 3 | 4 | 5 |
|---|---|---|---|---|---|
| $P(X)$ | 0.1 | 0.2 | 0.4 | 0.25 | 0.05 |

- Plots of the p.m.f.



# Cumulative Probability Distribution for a Discrete R.V.

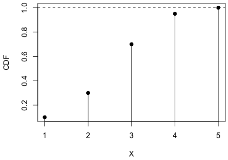
- Cumulative Probability Distribution
  1. A function giving the probability that the r.v. X is less than or equal to x.
  2. Called the "cumulative distribution function" or c.d.f.
  3. The probability that $X \leq x$ is notated as

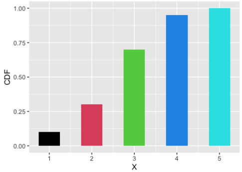
$$P(X \leq x) \text{ or } F_X(x)$$

- Example

| $X$ | 1 | 2 | 3 | 4 | 5 |
|---|---|---|---|---|---|
| $P(X)$ | 0.1 | 0.2 | 0.4 | 0.25 | 0.05 |
| $P(X \leq x)$ | 0.1 | 0.3 | 0.7 | 0.95 | 1 |

- Plots of the c.d.f.



## Your Turn

The cumulative probability distribution for the r.v. $X$ is given:

| $X$ | $F_X(x)$ |
|---|---|
| 0 | 0.08 |
| 1 | 0.30 |
| 2 | 0.43 |
| 3 | 0.71 |
| 4 | 1 |

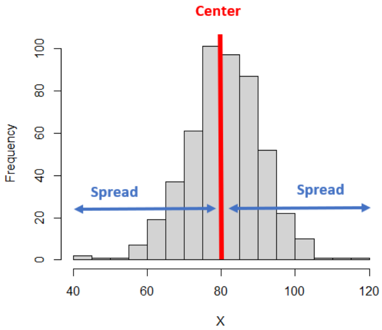
Compute and interpret:

1. $F_X(2)$
2. $f_x(2)$

# Table of Contents

# Statistical Summaries of a R.V.

- Statistical summaries are numbers that describe data or a r.v.
- Measures of center and spread (width) are common summaries.




1. Measure of Center
   - Expected Value (which is the mean)
2. Measure of Spread
   - Variance
   - Standard Deviation

# Calculating the Mean

## Formula: The Mean of a Discrete R.V.

The mean or expected value of a discrete r.v. $X$ is calculated as

$$E(X) = \sum_{i=1}^{n} x_i \cdot P(x_i)$$

where the expected value of $X$ is denoted as $E(X)$.

- The symbol $\mu$ is sometimes used in place of $E(X)$.
- Example: Mean of a Probability Distribution

| $X$ | 1 | 2 | 3 | 4 | 5 |
|---|---|---|---|---|---|
| $P(X)$ | 0.1 | 0.2 | 0.4 | 0.25 | 0.05 |

$$\begin{aligned} \mu &= E(X) \\ &= (1)(0.1) + (2)(0.2) + (3)(0.4) \\ &\quad +(4)(0.25) + (5)(0.05) \\ &= \boxed{2.95} \end{aligned}$$

# Additional Examples

(a) Suppose the following are your overall scores for this class. What is your final grade based on my grading criteria?

| Category | HW | Quizzes | Exam 1 | Exam 2 |
|---|---|---|---|---|
| Score | 90 | 80 | 75 | 88 |
| Weight | 30% | 10% | 30% | 30% |

(b) You play a game that costs \$5. One winner receives a prize of $1000$. If the probability of winning is $1/10000$, what is your expected return?

See handwritten solutions.

# Variance/Standard Deviation

| Measure | Abbreviation | Notation | Units |
|---|---|---|---|
| 1. Variance | $Var(X)$ | $\sigma^2$ | squared units |
| 2. Standard Deviation | $StD(X)$<br>$SD(X)$ | $\sigma$ | Un-squared units<br>Same as data |

- Variance measures the expected squared distance (or "deviation") of all the data from the mean.
  - "On average, the data are ___ $units^2$ from the mean."

**Formula: Variance and Standard Deviation of a Discrete R.V.**

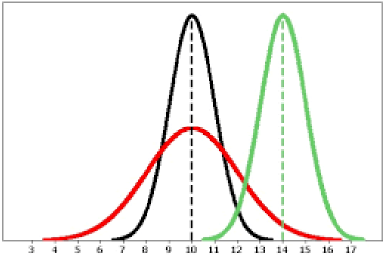
The variance of a discrete r.v. is computed as

$$\sigma^2 = \sum_{i=1}^{n} (x_i - \mu)^2 \cdot P(x_i)$$

and the standard deviation is simply $\sigma = \sqrt{variance}$.

# Variance/Standard Deviation

- Variance (and standard deviation) indicates the wideness of a distribution.
- Larger means wider!
- Example:




- Red is wider than black and green.
  - ∴ Red's variance is the greatest.
- Black and green look equally as wide.
  - ∴ The variances of black and green should be about the same, but less than red's variance.

# Example

Calculate the variance and standard deviation of $X$ given by:

| $X$ | 1 | 2 | 3 | 4 | 5 |
|---|---|---|---|---|---|
| $P(X)$ | 0.1 | 0.2 | 0.4 | 0.25 | 0.05 |

$\Longrightarrow$ We already found the mean $\mu = E(X) = 2.95$. Then,

$$\begin{aligned}
\Longrightarrow \sigma^2 &= \sum_{i=1}^{n} (x_i - \mu)^2 \cdot P(x_i) \\
&= (1 - 2.95)^2 \cdot (0.1) + (2 - 2.95)^2 \cdot (0.2) + (3 - 2.95)^2 \cdot (0.4) \\
&\quad + (4 - 2.95)^2 \cdot (0.25) + (5 - 2.95)^2 \cdot (0.05) \\
&= 1.0475
\end{aligned}$$

Variance: $\sigma^2 = 1.0475$ units$^2$

Standard Deviation: $\sigma = \sqrt{1.0475} \approx 1.023474$ units

# Alternative Formula for Variance

## Alternative Formula for Variance

The variance of the r.v. $X$ can also be found using

$$\sigma^2 = E(X^2) - [E(X)]^2$$

Example: Find the variance of the r.v. $X$ :

| $X$ | $P(X)$ |
|---|---|
| 1 | 0.1 |
| 2 | 0.2 |
| 3 | 0.4 |
| 4 | 0.25 |
| 5 | 0.05 |

Step 1: We can use the fact that $E(X) = 2.95$ which gives so far

$$Var(X) = E(X^2) - [2.95]^2$$

# Alternative Formula for Variance (Continued)

## Alternative Formula for Variance

The variance of the r.v. $X$ can also be found using

$$\sigma^2 = E(X^2) - [E(X)]^2$$

Step 2: Calculate $E(X^2)$: Square each $X$ value and use the fact that

| $X$ | $X^2$ | $P(X)$ |
|---|---|---|
| 1 | 1 | 0.1 |
| 2 | 4 | 0.2 |
| 3 | 9 | 0.4 |
| 4 | 16 | 0.25 |
| 5 | 25 | 0.05 |

$$\begin{aligned} E(X^2) &= \sum_{i=1}^{n} x^2 \cdot P(x_i) \\ &= 1 \cdot (0.1) + 4 \cdot (0.2) + 9 \cdot (0.4) \\ &\quad +16 \cdot (0.25) + 25 \cdot (0.05) \\ &= 9.75 \end{aligned}$$

# Alternative Formula for Variance (Continued)

## Alternative Formula for Variance

The variance of the r.v. $X$ can also be found using

$$\sigma^2 = E(X^2) - [E(X)]^2$$

| $X$ | $X^2$ | $P(X)$ |
|---|---|---|
| 1 | 1 | 0.1 |
| 2 | 4 | 0.2 |
| 3 | 9 | 0.4 |
| 4 | 16 | 0.25 |
| 5 | 25 | 0.05 |

Step 3: Put it all together:

$$\begin{aligned}\sigma^2 &= E(X^2) - [E(X)]^2 \\ &= 9.75 - [2.95]^2 \\ &= 1.0475\checkmark\end{aligned}$$

Same answer as before!

# Your Turn

Calculate the mean and variance of $X$:

| $X$ | $P(X)$ |
|---|---|
| 0 | 0.01 |
| 1 | 0.10 |
| 2 | 0.38 |
| 3 | 0.51 |