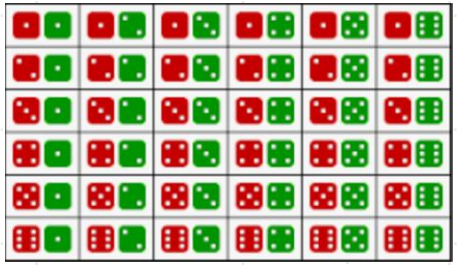
# Random Variables and Probability Distributions

From Outcomes to Real Numbers



# Revisit Dice Rolling Example

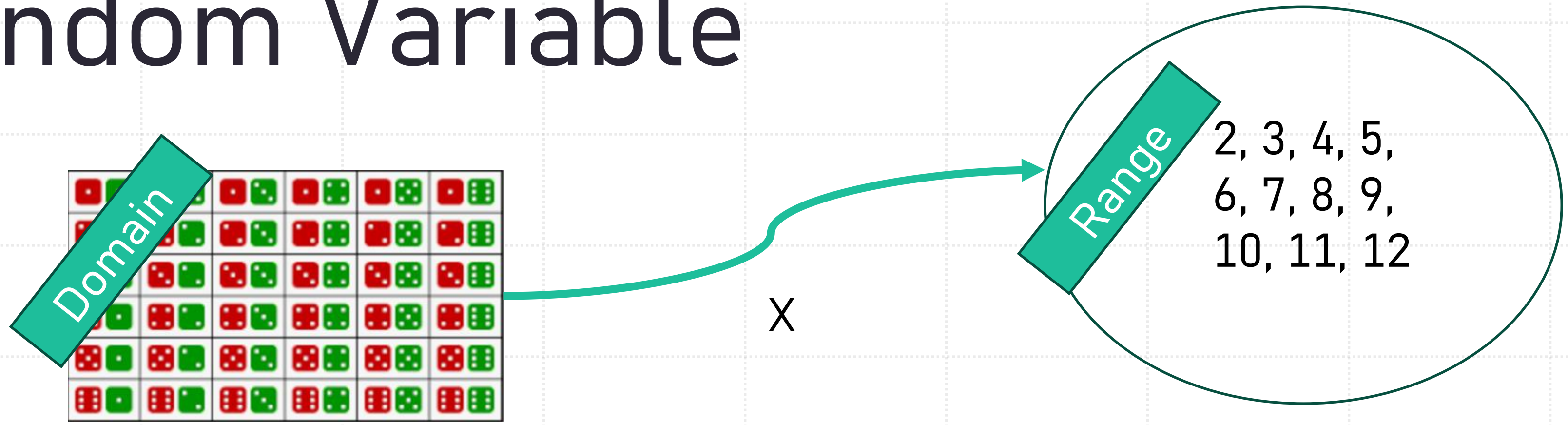
| | 2 | 3 | 4 | 5 | 6 | 7 | 8 | 9 | 10 | 11 | 12 |
|---|---|---|---|---|---|---|---|---|---|---|---|
| prob | 1/36 | 2/36 | 3/36 | 4/36 | 5/36 | 6/36 | 5/36 | 4/36 | 3/36 | 2/36 | 1/36 |

Define a new variable: The sum of the dots facing upward when rolling two dice.

Let’s call this variable X.

Thus, X= the sum of dots facing upward when rolling two dice.

# Random Variable




Thus, X= the sum of dots facing upward when rolling two dice.

X is called a Random Variable.

X can take values from the set {2, 3, 4, 5, 6, 7, 8, 9, 10, 11, 12}.

Suppose that X takes 7, then we say $X = x, where\ x = 7$ or in short $X = 7$.

7 is one of the realizations of X, meaning 7 is an observed value of the measurement X.

What is the definition of a random variable?

A **random variable** is a function that assigns a numerical value to each outcome in a sample space of a random experiment.

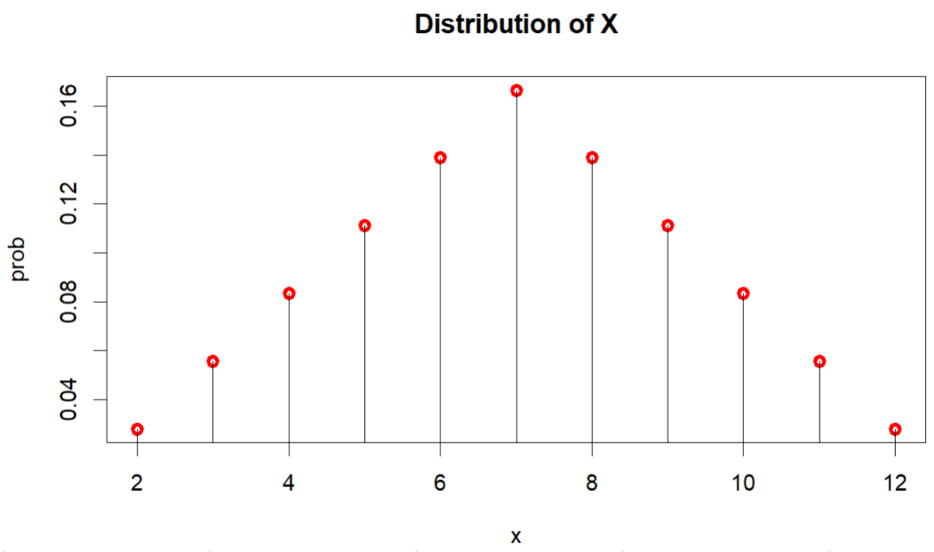
## Formal Definition:

A **random variable** is a function $X : \Omega \to \mathbb{R}$, where:

- $\Omega$ is the **sample space** (the set of all possible outcomes of a random process),
- $\mathbb{R}$ is the set of **real numbers**, and
- For each outcome $\omega \in \Omega$, $X(\omega)$ is a real number associated with that outcome.

# The Distribution of a Random Variable

| $x$ | 2 | 3 | 4 | 5 | 6 | 7 | 8 | 9 | 10 | 11 | 12 |
|---|---|---|---|---|---|---|---|---|---|---|---|
| prob | 1/36 | 2/36 | 3/36 | 4/36 | 5/36 | 6/36 | 5/36 | 4/36 | 3/36 | 2/36 | 1/36 |




$$prob = p(x) = \begin{cases} \frac{1}{36}, & x = 2,12 \\ \frac{2}{36}, & x = 3,11 \\ \frac{3}{36}, & x = 4,10 \\ \frac{4}{36}, & x = 5,9 \\ \frac{5}{36}, & x = 6,8 \\ \frac{6}{36}, & x = 7 \\ 0, & Otherwise \end{cases}$$

# Probability Mass Function

$$prob = p(x) = \begin{cases} \frac{1}{36}, & x = 2,12 \\ \frac{2}{36}, & x = 3,11 \\ \frac{3}{36}, & x = 4,10 \\ \frac{4}{36}, & x = 5,9 \\ \frac{5}{36}, & x = 6,8 \\ \frac{6}{36}, & x = 7 \\ 0, & Otherwise \end{cases}$$

What is the definition of a probability mass function?

A **probability mass function (PMF)** is a function that gives the probability that a **discrete random variable** is exactly equal to a specific value.

## Formal Definition:

Let $X$ be a discrete random variable. The **probability mass function** $p(x)$ is defined as:

$$\boxed{p(x) = P(X = x)}$$

for all values $x$ in the range of $X$.

# Example:

- X is a discrete random variable.
- The distribution of X is discrete.
- The p.m.f. of X is given as p(x), where

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

$X$ = number of declared majors before BYU graduation (a student who never changes his or her major has 1 major)

| $x$ | $p(x)$ |
|---|---|
| 1 | 0.15 |
| 2 | 0.30 |
| 3 | 0.40 |
| 4 | 0.10 |
| 5 | 0.05 |

or $$p(x) = \begin{cases} 0.15 & \text{if } x = 1 \\ 0.30 & \text{if } x = 2 \\ 0.40 & \text{if } x = 3 \\ 0.10 & \text{if } x = 4 \\ 0.05 & \text{if } x = 5 \\ 0 & \text{otherwise} \end{cases}$$

# Remarks

- $p(x)$ is a function.
- $P(X = x)$ is a probability of the event $\{X = x\}$.
- The function $p(x)$ takes value of the probability of $\{X = x\}$.

# Example: Rolling Dice

| $x$ | 2 | 3 | 4 | 5 | 6 | 7 | 8 | 9 | 10 | 11 | 12 |
|---|---|---|---|---|---|---|---|---|---|---|---|
| prob | 1/36 | 2/36 | 3/36 | 4/36 | 5/36 | 6/36 | 5/36 | 4/36 | 3/36 | 2/36 | 1/36 |

$$P(X \geq 7) = p(7) + p(8) + \cdots + p(12) = \sum_{x=7}^{12} p(x)$$

Note that $\{X \geq 7\} = \{X = 7\} \cup \{X = 8\} \cup \cdots \cup \{X = 12\} = \cup_{x=7}^{12}\{X = x\}$. The event is the union of a series of mutually exclusive events. Thus, we can use probability Axiom 3 to find $P(X \geq 7)$.

# Example:

$X$ = number of declared majors before BYU graduation (a student who never changes his or her major has 1 major)

| $x$ | $p(x)$ |
|---|---|
| 1 | 0.15 |
| 2 | 0.30 |
| 3 | 0.40 |
| 4 | 0.10 |
| 5 | 0.05 |

or

$$p(x) = \begin{cases} 0.15 & \text{if } x = 1 \\ 0.30 & \text{if } x = 2 \\ 0.40 & \text{if } x = 3 \\ 0.10 & \text{if } x = 4 \\ 0.05 & \text{if } x = 5 \\ 0 & \text{otherwise} \end{cases}$$

$$P(X = 4) = p(4) = 0.1$$

$$P(X > 1) = p(2) + p(3) + P(4) + p(5) = 0.85$$

Or

$$P(X > 1) = 1 - \mathrm{p}(1) = 1 - 0.15 = 0.85$$

## Properties of a PMF:

1. **Non-negativity:** $p(x) \geq 0$ for all values of $x$
2. **Normalization:** $\sum_{x \in \mathcal{X}} p(x) = 1$, where $\mathcal{X}$ is the set of all possible values that $X$ can take
3. **Support:** $p(x) > 0$ only for values $x \in \mathcal{X}$; for all other $x$, $p(x) = 0$

# Notes

## Random Variable


Thus, X= the sum of dots facing upward when rolling two dice.

X can take values from the set {2, 3, 4, 5, 6, 7, 8, 9, 10, 11, 12}.

Suppose that X takes 7, then we say $X = x, where\ x = 7$ or in short $X = 7$.

7 is one of the realizations of X, meaning 7 is an observed value of the measurement X.

## Probability Axioms

1. $P(\mathcal{S}) = 1,$ *where* $\mathcal{S}$ *is the sample space*
2. $P(A) \geq 0,$ where $A \subseteq \mathcal{S}$
3. $P(\bigcup_{i=1}^{\infty} A_i) = \sum_{i=1}^{\infty} P(A_i)$, *where* $A_i'$s are disjoint events

## Properties of a PMF:

1. **Non-negativity:** $p(x) \geq 0$ for all values of $x$
2. **Normalization:** $\sum_{x \in \mathcal{X}} p(x) = 1$, where $\mathcal{X}$ is the set of all possible values that $X$ can take
3. **Support:** $p(x) > 0$ only for values $x \in \mathcal{X}$; for all other $x$, $p(x) = 0$

# Verifying PMF Properties

(a)

| $x$ | 2 | 3 | 4 | 5 | 6 | 7 | 8 | 9 | 10 | 11 | 12 |
|---|---|---|---|---|---|---|---|---|---|---|---|
| prob | 1/36 | 2/36 | 3/36 | 4/36 | 5/36 | 6/36 | 5/36 | 4/36 | 3/36 | 2/36 | 1/36 |

(b)

| $x$ | $p(x)$ |
|---|---|
| 1 | 0.15 |
| 2 | 0.30 |
| 3 | 0.40 |
| 4 | 0.10 |
| 5 | 0.05 |

(c) Which of the following functions is a p.m.f.?

$$p(x) = \begin{cases} 0.5, & x = 1 \\ 0.1, & x = 2 \\ 0.2, & x = 3 \\ 0, & Otherwise \end{cases}$$

$$\mathrm{f(x)} = \begin{cases} 0.5, & x = 1 \\ 0.3, & x = 2 \\ 0.2, & x = 3 \\ 0, & Otherwise \end{cases}$$

# Example: Binary Outcomes

- Suppose we play a game with only two possible outcomes:
  - Success – The probability of succeeding in the game is 0.7
  - Failure – The probability of failing in the game is 0.3
- We introduce a random variable X, where

$$X = \begin{cases} 1, & if\ it\ is\ a\ success \\ 0, & if\ it\ is\ a\ failure \end{cases}$$

- The p.m.f. is then given by

$$p(x) = \begin{cases} 0.7, & for\ x = 1 \\ 1 - 0.7, & for\ x = 0 \end{cases}$$

# Remarks

- The random variable in the Binary Outcome example is called a Bernoulli random variable.
- The p.m.f. can also be written as

$$p(x) = 0.7^x(1 - 0.7)^{1-x}, \text{for x} = 0 \text{ or } 1$$

- If we denote the probability of a success by a generical notation $p$, then the above function is written as

$$p(x) = p^x(1 - p)^{1-x}, \text{for x} = 0 \text{ or } 1$$

# Bernoulli Distribution

- The random variable in the Binary Outcome example is called a Bernoulli random variable.