A Random Variable ($x$): assigns a numerical value, assigned by chance, to each outcome of a procedure.

† Note: the random variable is more properly denoted as capital $X$, until it is assigned a numerical value (that we call a realization), then is referred to as little $x$. However, your book does not distinguish the two, so we will generally go with the less proper notation $x$ to symbolize a random variable.

** Example: ** rolling a die-the random variable (RV), $x$ is the result of the roll

$$x \in \{1, 2, 3, 4, 5, 6\}$$
The **probability distribution** \( P(x) \) gives the probability that a RV takes each of the possible values (may be displayed graphically, in a table, or a formula).

**Example:** rolling a die

<table>
<thead>
<tr>
<th>Roll: ( x )</th>
<th>Probability: ( P(x) )</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>1/6</td>
</tr>
<tr>
<td>2</td>
<td>1/6</td>
</tr>
<tr>
<td>3</td>
<td>1/6</td>
</tr>
<tr>
<td>4</td>
<td>1/6</td>
</tr>
<tr>
<td>5</td>
<td>1/6</td>
</tr>
<tr>
<td>6</td>
<td>1/6</td>
</tr>
</tbody>
</table>

The probability histogram for coin flips is not so interesting...

**Example:** (slightly more interesting) is the number of times heads comes up in three coin flips:

HHH, HTT, THT, TTH, THH, HTH, HHT, TTT

<table>
<thead>
<tr>
<th># heads: ( x )</th>
<th>Probability: ( P(x) )</th>
</tr>
</thead>
<tbody>
<tr>
<td>0</td>
<td>1/8</td>
</tr>
<tr>
<td>1</td>
<td>3/8</td>
</tr>
<tr>
<td>2</td>
<td>3/8</td>
</tr>
<tr>
<td>3</td>
<td>1/8</td>
</tr>
</tbody>
</table>

The probability histogram for coin flips is not so interesting...
Random Variables may be **discrete** or **continuous**.

→ The previous examples are **discrete**, and thus the corresponding probability distribution, \( P(x) \), is a histogram. (e.g. binomial and Poisson distributions)

→ **Continuous** random variables take on values with no gaps in between, thus the corresponding probability distribution, \( P(x) \), is a smooth curve: (e.g. normal distribution)

![Normal Distribution](image)

### Requirements for a Probability Distribution

1. \( \sum_{i=1}^{N} P(x_i) = 1 \) The probability of all events must add to 1 (and these events are mutually exclusive)

2. \( 0 \leq P(x) \leq 1 \)

**Example:** # of heads in 3 coin tosses.

<table>
<thead>
<tr>
<th># heads: ( x )</th>
<th>Probability: ( P(x) )</th>
</tr>
</thead>
<tbody>
<tr>
<td>0</td>
<td>1/8</td>
</tr>
<tr>
<td>1</td>
<td>3/8</td>
</tr>
<tr>
<td>2</td>
<td>3/8</td>
</tr>
<tr>
<td>3</td>
<td>1/8</td>
</tr>
</tbody>
</table>

\[ \sum_{i=1}^{4} P(x_i) = (1/8) + (3/8) + (3/8) + (1/8) = 8/8 = 1 \]
A probability distribution has a **theoretical** mean and sd, like a population mean and sd. The theoretical mean is also called the **Expected Value**:

- \( \mu = E[x] = \sum_{i=1}^{N} [x_i P(x_i)] \)
- \( \sigma^2 = Var(x) = \sum_{i=1}^{N} [(x_i - \mu)^2 P(x_i)] \)
- \( \sigma = \sqrt{\sum_{i=1}^{N} [(x_i - \mu)^2 P(x_i)]} \)

**Example**: # of heads in 3 coin tosses.

\[
\mu = E[x] = 0 \left( \frac{1}{8} \right) + 1 \left( \frac{3}{8} \right) + 2 \left( \frac{3}{8} \right) + 3 \left( \frac{1}{8} \right) = \left( \frac{12}{8} \right) = 1.5
\]

\[
\sigma = \sqrt{\left[(0 - 1.5)^2 \frac{1}{8} + (1 - 1.5)^2 \frac{3}{8} + (2 - 1.5)^2 \frac{3}{8} + (3 - 1.5)^2 \frac{1}{8}\right]} = 0.8660
\]

**Binomial Probability Distribution**: If the following are met

1. Fixed number of trials, \( n \)
2. Trials are independent
3. Each trial has only two possible outcomes (“success”, “failure”)
4. The probability of success, \( p \), is the same for each trial

\( \Rightarrow \) Then, the # of successes is \( X \sim Bin(n, p) \)

**Are the following scenarios binomial or not?**

→ Flip a coin twice, count the # of heads.
→ # of guesses right on a multiple choice exam.
→ In a random sample of 50 votes, who voted for Schwarzenegger.
→ Stand at an intersection for 10 minutes and count the # of cars that turn left.
→ Someone comes to this class with the flu, the # of students present that day that caught it.
There are two ways to get the probability distribution of the Binomial:
→ The table in A-1
→ OR the binomial formula:

\[ P(x) = \frac{n!}{(n-x)!x!} p^x (1-p)^{n-x} \]

where the \(!\) symbolizes the factorial:

\[ n! = n(n-1)(n-2) \cdots (2)(1) \]
\[ 1! = 1 \]
\[ 0! = 1 \text{ (by definition)} \]
\[ 4! = 4 \cdot 3 \cdot 2 \cdot 1 = 24 \]

The term \( \frac{n!}{(n-x)!x!} \) is also known as the choose function, verbally expressed as “n choose x”.

\[ nC_x = \binom{n}{x} = \frac{n!}{(n-x)!x!} \]

**Example**: Suppose you are taking a pop quiz with 5 multiple choice questions, each with five choices. You guess randomly. Let \( X \) be the \# you guess correctly.

\[ X \sim Bin(n = 5, p = 1/5 = 0.2) \]

○ What is the probability that you answer three questions correctly? \( P(X = 3) =? \)

From the formula:
\[ P(X = 3) = \frac{5!}{2!3!} (0.2)^3 (0.8)^2 = 0.0512 \]
From the table?

○ What is the probability that you answer all questions correctly? \( P(X = 5) =? \)

From the formula:
\[ P(X = 5) = \frac{5!}{0!5!} (0.2)^5 (0.8)^0 = (0.2)^5 = 0.0003 \]
From the table?
What is the probability that you get at least two correct?

\[ P(X \geq 2) = ? \]

\[
P(X \geq 2) = P(2) + P(3) + P(4) + P(5) = 1 - P(0) - P(1)
\]

\[ = 1 - 0.328 - 0.410 \]

\[ = 0.262 \]

What is the probability that you get an odd number of questions correct?

\[ P(X = 1 \text{ or } X = 3 \text{ or } X = 5) = P(1) + P(3) + P(5) \]

\[ = 0.410 + 0.051 + 0 \]

\[ = 0.461 \]

Recall that probability distributions have a theoretical mean (expected value) and sd. For the binomial (and ONLY binomial) distribution there is a shortcut formula to find the mean and sd:

- \( \mu = E[X] = np \)
- \( \sigma^2 = Var(X) = np(1 - p) \)
- \( \sigma = \sqrt{np(1 - p)} \)

**Example**: The multiple choice pop quiz -

\( X \sim Bin(n = 5, p = 1/5 = 0.2) \)

- \( \mu = E[X] = 5(0.2) = 1 \)
- \( \sigma = \sqrt{5(0.2)(0.8)} = 0.89 \)
When is an event considered to be “unusual”? Recall that when the distribution looked bell-shaped, we said that values more than 2 sd’s away from the mean were unusual.

- This corresponded to the observation having a probability < 0.05 of occurring (95% of the observations were expected to fall within 2 sd’s of the mean).
- We can use the same rule of thumb, i.e. the Empirical formula, for the binomial distribution:

**Example:** If a student gets 4 question right on that 5-question pop quiz, do we think it is reasonable that he guessed randomly?

\[
\begin{align*}
\mu &= 1 \\
\sigma &= 0.89 \\
\Rightarrow: \mu + 2\sigma &= 1 + 2(0.89) = 2.78
\end{align*}
\]
Since 4 > 2.78, then it is unlikely that this student was guessing randomly.

**Poisson Probability Distribution:** If the following are met

1. the RV is the number of occurrences of an event over some interval
2. Occurrences must be at random
3. Occurrences must be independent of each other

- the data are counts with no upper limit
  † typos in a paper
  † earthquakes in a year
  † chips in a cookie
  † pepperonis on a pizza slice
• Differences between the binomial and Poisson distribution:

<table>
<thead>
<tr>
<th>binomial</th>
<th>Poisson</th>
</tr>
</thead>
<tbody>
<tr>
<td>effected by $n$ and $p$</td>
<td>effected by $\mu$ only</td>
</tr>
<tr>
<td>$x \in {0, 1, 2, \ldots, n}$</td>
<td>$x \in {0, 1, 2, \ldots}$ (no upper limit)</td>
</tr>
</tbody>
</table>

• As with the Binomial distribution, you can get the probability distribution via a table (which your book does NOT provide) or the Poisson formula:

$$P(x) = \frac{\mu^x e^{-\mu}}{x!}$$

where $e \approx 2.71828$

☆ The mean and the sd for the Poisson are easy cheesy!!!!!!
  • $E[x] = \mu$ (typically just given, sometimes minor algebra needed)
  • $Var(x) = \mu$
  • $\sigma = \sqrt{\mu}$
** Example: Suppose you typically make 4 typos per page, and you type a 9-page paper. Would it be unusual for you to make only 20 typos in the paper?

→ For a single page, the mean would be 4 typos
  - But we are interested in 9 pages, so

\[
\mu = 9(4) = 36 \quad \text{and} \quad \sigma = \sqrt{\mu} = \sqrt{36} = 6
\]

⇒

\[
\mu - 2\sigma = 36 - 2(6) = 24
\]

Since 20 < 24, then, yes, we would say that 20 is an unusually small number of typos.

---

Key Concepts!!!!

★ Probability Distribution
★ Binomial Distribution
★ Expected Value
★ Poisson Distribution