Random Variable.
- assigns a number to each mutually exclusive event.

Eg. Roll a die.
R.V. is the result of the roll
\[ X \in \{1, 2, 3, 4, 5, 6\} \]

Flip two coins
\[ Y \] be # heads
\[ Y \in \{0, 1, 2\} \]

Probability Distribution.
gives the probability that a Random Variable takes each of the possible values.
Table X:

<table>
<thead>
<tr>
<th>Roll</th>
<th>Prob.</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>$\frac{1}{6}$</td>
</tr>
<tr>
<td>2</td>
<td>$\frac{1}{6}$</td>
</tr>
<tr>
<td>3</td>
<td>$\frac{1}{6}$</td>
</tr>
<tr>
<td>4</td>
<td>$\frac{1}{6}$</td>
</tr>
<tr>
<td>5</td>
<td>$\frac{1}{6}$</td>
</tr>
<tr>
<td>6</td>
<td>$\frac{1}{6}$</td>
</tr>
</tbody>
</table>

Graph of $P(x) = \frac{1}{6}$

Table Y:

<table>
<thead>
<tr>
<th># heads</th>
<th>Prob.</th>
</tr>
</thead>
<tbody>
<tr>
<td>0</td>
<td>$\frac{1}{4}$</td>
</tr>
<tr>
<td>1</td>
<td>$\frac{1}{2}$</td>
</tr>
<tr>
<td>2</td>
<td>$\frac{1}{4}$</td>
</tr>
</tbody>
</table>

Mutually exclusive events

Probabilities sum to 1

Probability distributions can be defined over

Discrete or continuous random variables.

E.g.,

Binomial

Poisson

E.g.

Normal
Binomial Distribution.

A bag containing one red ball
nine green balls

Make 5 draws with replacement.

What is \( p(\text{exactly 2 reds in the 5 draws}) \)?

\[
\begin{align*}
\text{RRGGG} & \quad \frac{1}{10} \times \frac{1}{10} \times \frac{9}{10} \times \frac{9}{10} \times \frac{9}{10} = \left(\frac{1}{10}\right)^2 \times \left(\frac{9}{10}\right)^3 \\
\text{RRGGR} & \quad \left(\frac{1}{10}\right)^2 \times \left(\frac{9}{10}\right)^3 \\
\text{GRGGR} & \quad \left(\frac{1}{10}\right)^2 \times \left(\frac{9}{10}\right)^3 \\
\text{GGRGR} & \quad \left(\frac{1}{10}\right)^2 \times \left(\frac{9}{10}\right)^3 \\
\text{GGRGR} & \quad \left(\frac{1}{10}\right)^2 \times \left(\frac{9}{10}\right)^3 \\
\text{GGGRR} & \quad \left(\frac{1}{10}\right)^2 \times \left(\frac{9}{10}\right)^3 \\
\text{GGGRR} & \quad \left(\frac{1}{10}\right)^2 \times \left(\frac{9}{10}\right)^3 \\
\text{GGGRR} & \quad \left(\frac{1}{10}\right)^2 \times \left(\frac{9}{10}\right)^3 \\
\text{GGGRR} & \quad \left(\frac{1}{10}\right)^2 \times \left(\frac{9}{10}\right)^3 \\
\end{align*}
\]

\[ p(\text{exactly 2 reds in 5 draws}) = 10 \times \left(\frac{1}{10}\right)^2 \times \left(\frac{9}{10}\right)^3 \]

\[ \text{Is there a formula that tells us how many possibilities there are?} \]
Binomial Coefficient

- \( n \) - # trials
- \( k \) - # successes

\[
\binom{n}{k} = \frac{n!}{(n-k)! \cdot k!}
\]

- \( n! = n \) factorial.
- \( 4! = 4 \times 3 \times 2 \times 1 \)

# ways of arranging \( n \) objects.

0! = 1

\[
n! = n \times (n-1) \times (n-2) \times \ldots \times 2 \times 1
\]

\[
\binom{n}{k} = \# \text{ of ways to pick } k \text{ out of } n \text{ if order doesn't matter.}
\]

Binomial Distribution

\[
p(k) = \frac{n!}{k! \cdot (n-k)!} \cdot p^k \cdot (1-p)^{n-k}
\]

- \( p \) - probability of success on any trial
Example.

A family has 4 children.

What's the probability that they have more girls than boys?

\[ P(3 \text{ girls out of 4 kids}) + P(4 \text{ girls out of 4}) = \frac{4!}{3!(4-3)!} \left(\frac{1}{2}\right)^3 \cdot \frac{1}{2} \left(1-\frac{1}{2}\right)^{4-3} + \frac{4!}{4!(4-4)!} \left(\frac{1}{2}\right)^4 \cdot \left(1-\frac{1}{2}\right)^0 \]

\[ = \frac{24}{6 \times 1} \times \frac{1}{8} \times \frac{1}{2} + \frac{24}{24 \times 1} \times \frac{1}{16} \]

\[ = \frac{1}{4} + \frac{1}{16} \]

\[ = \frac{5}{16} = 31\% \]

In class 14 families with 4 children, 7 with more girls than boys.

\[ \frac{7}{14} = \frac{1}{2} \]
Conditions for the Binomial Distribution to apply.

1. Fixed number of trials, $n$

2. Trials are independent

3. Each trial has only two possible outcomes
   ("success" or "failure")

4. The probability of success, $p$, is the same for each trial.

Let $k \sim \text{Bin}(n, p)$

$\uparrow$

has the distribution.

Flip coin $n$ times, count # heads.

# correct guesses on a multiple choice test.

Random sample of 100 students, count by gender
Multiple choice test with 5 questions.
Each with 5 options, one of which is correct.
You guess randomly.

Let $Z = 4$ guesses correctly.

$$p(Z) = \frac{5!}{(5-Z)! Z!} \cdot \left( \frac{1}{5} \right)^Z \left( \frac{4}{5} \right)^{5-Z}$$

$$p(5) = \frac{5!}{0! 5!} \cdot \left( \frac{1}{5} \right)^5 \left( \frac{4}{5} \right)^0 = 0.0003$$

$$p(3) = \frac{5!}{2! 3!} \cdot \left( \frac{1}{5} \right)^3 \left( \frac{4}{5} \right)^2 = 0.0512$$

<table>
<thead>
<tr>
<th>$Z$</th>
<th>$p(Z)$</th>
<th>$p$ (at least 2 right)</th>
</tr>
</thead>
<tbody>
<tr>
<td>0</td>
<td>0.328</td>
<td></td>
</tr>
<tr>
<td>1</td>
<td>0.0410</td>
<td>$= p(2 \text{ or } 3 \text{ or } 4 \text{ or } 5)$</td>
</tr>
<tr>
<td>2</td>
<td>0.205</td>
<td>$= 0.205 + 0.051 + 0.006 + 0$</td>
</tr>
<tr>
<td>3</td>
<td>0.051</td>
<td>$= 0.262.$</td>
</tr>
<tr>
<td>4</td>
<td>0.006</td>
<td></td>
</tr>
<tr>
<td>5</td>
<td>0</td>
<td></td>
</tr>
</tbody>
</table>

$\text{EV Probability Distribution}$
A probability distribution has a theoretical mean \( \mu \) and standard deviation.

The mean of a probability distribution is called the expected value \( \mu \) or

\[
\mu = E[x] = \sum [x \cdot p(x)]
\]

"the expectation of \( x \)"

\[ E[z] = 0 \cdot 0.328 + 1 \cdot 0.41 + 2 \cdot 0.205 + 3 \cdot 0.051 + 4 \cdot 0.006 + 5 \cdot 0 = \frac{1}{\mu} \]
\[ Y = \pm H \text{ or } 2 \cos. \]

\[ E[Y] = 0 \times \frac{1}{4} + 1 \times \frac{1}{2} + 2 \times \frac{1}{4} = 1 \]

\[ X = \text{outcome of rolling a single die}. \]

\[ E[X] = 1 \times \frac{1}{6} + 2 \times \frac{1}{6} + 3 \times \frac{1}{6} + 4 \times \frac{1}{6} + 5 \times \frac{1}{6} + 6 \times \frac{1}{6} = 3. \frac{1}{2}. \]

\[ \sigma^2 = \text{Var}[X] = \sum [ (x - \mu)^2 \cdot p(x) ] \]

\[ = E \left[ (X-\mu)^2 \right] \]

\[ \sigma = \sqrt{\sum [ (x-\mu)^2 \cdot p(x) ]} \]

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

\[ = E[X^2] - \mu^2 \]

\[ = \sum x^2 \cdot p(x) - \mu^2 \text{ is often easier to compute}. \]
For a Binomial:

\[\mu = np\]
\[\sigma = \sqrt{np(1-p)}\]

\[Y: \# \text{H or 2 coins.}\]
\[\sigma_Y = \sqrt{2 \times \frac{1}{2} \times (1-\frac{1}{2})} = \frac{1}{\sqrt{2}}\]

\[Z: \# \text{correct guesses.}\]
\[\sigma^2_Z = \sqrt{5 \times \frac{1}{5} \times \frac{1}{5}} = \frac{2}{\sqrt{5}} = 0.89.\]

Recall: Typically 95% of data lie in the range \(\mu \pm 2\sigma\).

\[Z: \mu = 1 \quad \mu + 2\sigma = 2.79.\]
\[\sigma = 0.89\]

If a student gets 3 correct, it starts to suggest that something other than guessing is going on.
Poisson Distribution.

When events occur randomly at a certain rate, the # of events that occur in an interval of time (or space) has a Poisson distribution.

Examples.

# injuries
# typos in a paper
# earthquakes in a year
# chips in a cookie

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

\( \mu \) - mean

Variance \( \sigma^2 = \mu \)

Std. deviation \( \sigma = \sqrt{\mu} \)
Example.

Suppose you typically make 4 typos/page.

Type a 9 page paper.

Would it be unusual to make only 20 typos?

1 page, \( \mu_x = 4 \)

9 pages, \( \mu_y = 4 \times 9 = 36 \)

\( \sigma_y = \sqrt{36} = 6 \).

Empirical rule:

consider \( \mu - 2\sigma \)

\( 36 - 2 \times 6 \)

\( = 24 \)

\( 20 < 24 \) so this would be considered

unusual.
<table>
<thead>
<tr>
<th>Year</th>
<th># Lost Time Injuries</th>
</tr>
</thead>
<tbody>
<tr>
<td>2005</td>
<td>3</td>
</tr>
<tr>
<td>2006</td>
<td>4</td>
</tr>
<tr>
<td>2007</td>
<td>3</td>
</tr>
<tr>
<td>2008</td>
<td>2</td>
</tr>
<tr>
<td>2009</td>
<td>3</td>
</tr>
<tr>
<td>2010</td>
<td>0</td>
</tr>
<tr>
<td>2011</td>
<td>1</td>
</tr>
</tbody>
</table>

The chances are there in a clump of buses after the long intervals. Otherwise there would not have been a long interval for you to arrive in.