Outline

1. Probability Model
   - Experiments and events
   - Rules for probability calculations
   - Independence

2. Random Variables
   - Definition and examples
   - Probability distribution of a RV
   - Expectation
   - Variance and Standard deviation
   - Independence
Notion of chance

What is a probability?

- **Long-run relative frequency** such as past flood records used to predict the chance of flood tomorrow.

- **Subjective notion** such as expert opinions about the reliability of a new GPS unit.

- **Classical notion** based on assumed symmetry in situations. Example: a bag of 50 M&Ms contains 5 blue ones (45 are brown). If 1 piece is picked at “random”, then the probability of it being blue is

Probability theory models/quantifies/describes any of these notions of chance. Our focus: classical notion.
Elements in a probability model

- **Experiment**: action/process that generates data. Ex: recording the weight of a bag of soil, counting the # of cows, rolling a die, tossing a coin, etc. Usually has more than one possible outcomes, is theoretically repeatable.

- Individual possible outcomes are **elementary outcomes**.

- **Sample space**: the entire group of elementary outcomes, noted $S$. Ex: roll a die once and record the result: $S = \{1, 2, 3, 4, 5, 6\}$

- **Event**: a collection of elementary outcomes, i.e. a subset of $S$. Ex: pick a nest of super fairy wrens at random and count the number of eggs. Possible events are:
  
  - $A = \text{“the result is even”} = \{0,2,4,6,\ldots\}$
  - $B = \text{“the result is not 4 nor 5”} = \{0,1,2,3,6,\ldots\}$
  - $C = \text{“the result is 3”} = \{3\}$
  - $D = \text{“there is at least one egg”}$
  - $S = \{0, 1, 2, 3, 4, 5, 6, \ldots\}$
  - $\emptyset = \{\}$ is an empty set.
Operations on events

- The event “$U$ or $V$” consists of the elementary outcomes in $U$, in $V$, or in both. Written as $U \cup V$, reads as ”$U$ union $V$”. $A$ or $C = \{0,2,4,6,\ldots\}$ or $\{3\} =$ $A$ or $B =$
- The event “$U$ and $V$” consists of the elementary outcomes in both $U$ and $V$. Written as $U \cap V$, reads ”$U$ intersect $V$”. $A$ and $B = \{0,2,4,6,\ldots\}$ and $\{0,1,2,3,6,\ldots\} =$ $B$ and $C =$ “even number of eggs” and “odd number” =
- “not $U$” consists of all elementary outcomes in $S$ that are not in $U$. Written as $\bar{U}$, reads “the complement of $U$”. not $D =$ not “at least one egg”= not “at least 2 eggs”=
- Two events are **mutually exclusive** if they do not have any elementary outcomes in common. Examples:
A probability model consists of a probability assignment to each of the events in \( S \).

### Basic rules an assignment must follow

(i) For any event \( U \), \( 0 \leq \mathbb{P}(U) \leq 1 \)
(ii) \( \mathbb{P}(S) = 1 \)
(iii) Addition rule: If \( U \) and \( V \) are mutually exclusive, then

\[
\mathbb{P}(U \text{ or } V) = \mathbb{P}(U) + \mathbb{P}(V)
\]

These rules ensure that the assignment is consistent with our intuitive notion of chance.
Example: fair die rolled once

\( S = \{1, 2, 3, 4, 5, 6\} \). If the die is fair, we choose the probability model that says

\[
\mathbb{P}(1) = \mathbb{P}(2) = \mathbb{P}(3) = \mathbb{P}(4) = \mathbb{P}(5) = \mathbb{P}(6) = \frac{1}{6}.
\]

Following the three basic rules we get:

\[
\mathbb{P}(\{2, 3, 4, 6\}) = \mathbb{P}(2 \text{ or } 3 \text{ or } 4 \text{ or } 6) = \mathbb{P}(2) + \mathbb{P}(3) + \mathbb{P}(4) + \mathbb{P}(6) = 2/3
\]

\[
\mathbb{P}(\text{“3 or less” or “3 or more”}) = \mathbb{P}(\text{“anything”}) = 1
\]

\[
\neq \mathbb{P}(\text{“3 or less”}) + \mathbb{P}(\text{“3 or more”}) =
\]
Derived rules

What if we cannot apply (iii) to compute \( \mathbb{P}(A \text{ or } B) \) because \( A \) and \( B \) are not mutually exclusive? From (i)–(iii):

(iv) For any two events \( U, V \),

\[
\mathbb{P}(U \text{ or } V) = \mathbb{P}(U) + \mathbb{P}(V) - \mathbb{P}(U \text{ and } V)
\]

(iv) is consistent with (iii). Now

\[
\begin{align*}
\mathbb{P}(\text{"3 or less" or "3 or more"}) &= \mathbb{P}(\text{"3 or less"}) + \mathbb{P}(\text{"3 or more"}) - \mathbb{P}(\text{" } ) \\
&= 3/6 + 4/6 - 1/6 = 1. \text{ Pfew!}
\end{align*}
\]

Can be generalized: For any three events \( U, V, W \):

\[
\begin{align*}
\mathbb{P}(U \text{ or } V \text{ or } W) &= \mathbb{P}(U) + \mathbb{P}(V) + \mathbb{P}(W) \\
&- \mathbb{P}(U \text{ and } V) - \mathbb{P}(U \text{ and } W) - \mathbb{P}(U \text{ and } W) + \mathbb{P}(U \text{ and } V \text{ and } W)
\end{align*}
\]
Derived rules

$U$ and “not $U$” are always mutually exclusive. By (iii),

$$\mathbb{P}(U) + \mathbb{P}(\text{not } U) = \mathbb{P}(S) = 1$$

Thus the derived rule:

(v) For any event $U$, $\mathbb{P}(\text{not } U) = 1 - \mathbb{P}(U)$

Ex: $\mathbb{P}(\text{ not } \{6\}) = \mathbb{P}(\{1, 2, 3, 4, 5\}) =$

or simply by (v): $\mathbb{P}(\text{not } \{6\}) = 1 - \mathbb{P}(6) =$
Conditional probability

- Again roll a fair die. I am told the result is an even number. What is the probability that it is a 2? Not 1/6 any longer!

\[ \mathbb{P}(2 \text{ given it is an even number}) = \]

- Additional information can alter the probability of an event.
- The **conditional probability** of an event \( U \) given \( V \), \( \mathbb{P}(U \mid V) \), is the probability of \( U \) given (or knowing) that \( V \) has occurred. Exact definition:

\[
\mathbb{P}(U \mid V) = \frac{\mathbb{P}(U \text{ and } V)}{\mathbb{P}(V)},
\]

provided that \( \mathbb{P}(V) \neq 0 \).

- Fair die:

\[
\mathbb{P}(2 \mid \text{even}) = \frac{\mathbb{P}(2 \text{ and even})}{\mathbb{P}(\text{even})} = \frac{\mathbb{P}(2)}{\mathbb{P}(\{2, 4, 6\})} = \frac{1/6}{3/6} = 1/3.
\]
Independence

- Probability model: fairywren eggs & helper presence

<table>
<thead>
<tr>
<th>egg #</th>
<th>2</th>
<th>2</th>
<th>3</th>
<th>3</th>
<th>4</th>
<th>4</th>
<th>5</th>
<th>5</th>
</tr>
</thead>
<tbody>
<tr>
<td>helper</td>
<td>N</td>
<td>Y</td>
<td>N</td>
<td>Y</td>
<td>N</td>
<td>Y</td>
<td>N</td>
<td>Y</td>
</tr>
<tr>
<td>probability</td>
<td>.05</td>
<td>.05</td>
<td>.15</td>
<td>.15</td>
<td>.18</td>
<td>.18</td>
<td>.12</td>
<td>.12</td>
</tr>
</tbody>
</table>

Under this model, what is ...
\( \mathbb{P}(2 \text{ eggs}) = \) \\
\( \mathbb{P}(\text{helper present}) = \) \\
\( \mathbb{P}(2 \text{ eggs} \mid \text{helper}) = \)

- Does the extra knowledge of a helper’s presence alter the probability on the number of eggs?

This model corresponds to a hypothesis that egg # is independent of helper presence.

Two events \( U, V, U \) and \( V \) are independent if

\[ \mathbb{P}(U \mid V) = \mathbb{P}(U) \]
Independence

If $U$ and $V$ are independent: $\mathbb{P}(U) = \mathbb{P}(U|V) = \frac{\mathbb{P}(U \text{ and } V)}{\mathbb{P}(V)}$

Multiplication rule

$U$ and $V$ are independent if

$$\mathbb{P}(U \text{ and } V) = \mathbb{P}(U) \times \mathbb{P}(V)$$

Are “4 eggs or more” and “helper present” independent, in the model below?

<table>
<thead>
<tr>
<th>egg #</th>
<th>3</th>
<th>3</th>
<th>4</th>
<th>4</th>
<th>5</th>
<th>5</th>
</tr>
</thead>
<tbody>
<tr>
<td>helper</td>
<td>N</td>
<td>Y</td>
<td>N</td>
<td>Y</td>
<td>N</td>
<td>Y</td>
</tr>
<tr>
<td>probability</td>
<td>.12</td>
<td>.18</td>
<td>.20</td>
<td>.30</td>
<td>.08</td>
<td>.12</td>
</tr>
</tbody>
</table>

And with the model below?

<table>
<thead>
<tr>
<th>egg #</th>
<th>3</th>
<th>3</th>
<th>4</th>
<th>4</th>
<th>5</th>
<th>5</th>
</tr>
</thead>
<tbody>
<tr>
<td>helper</td>
<td>N</td>
<td>Y</td>
<td>N</td>
<td>Y</td>
<td>N</td>
<td>Y</td>
</tr>
<tr>
<td>probability</td>
<td>.12</td>
<td>.18</td>
<td>.18</td>
<td>.32</td>
<td>.04</td>
<td>.16</td>
</tr>
</tbody>
</table>
Independence

The multiplication rule can be used to find \( \mathbb{P}(U \text{ and } V) \), if we know \( U \) and \( V \) are independent.

Example: we know egg # and helper presence are independent

<table>
<thead>
<tr>
<th>helper</th>
<th>Y</th>
<th>N</th>
</tr>
</thead>
<tbody>
<tr>
<td>probability</td>
<td>.40</td>
<td>.60</td>
</tr>
</tbody>
</table>

| egg # | 3 | 4 | 5 |
| probability | .30 | .50 | .20 |

\[ \mathbb{P}(3 \text{ eggs and no helper}) = \]
Independence

If we know that egg # and helper presence are not independent, then we need to know conditional probabilities to calculate probabilities. General form of the multiplication rule:

\[ P(U \text{ and } V) = P(U) \times P(V|U) = P(V) \times P(U|V) \]

Example: we assume this probability model

<table>
<thead>
<tr>
<th>helper</th>
<th>Y</th>
<th>N</th>
</tr>
</thead>
<tbody>
<tr>
<td>prob.</td>
<td>.40</td>
<td>.60</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>egg #</th>
<th>3</th>
<th>4</th>
<th>5</th>
</tr>
</thead>
<tbody>
<tr>
<td>prob.</td>
<td>.40</td>
<td>.50</td>
<td>.10</td>
</tr>
<tr>
<td></td>
<td>.20</td>
<td>.55</td>
<td>.25</td>
</tr>
</tbody>
</table>

if no helper

if helper

\[ P(3 \text{ eggs and no helper}) = \]
Outline

1. Probability Model
   - Experiments and events
   - Rules for probability calculations
   - Independence

2. Random Variables
   - Definition and examples
   - Probability distribution of a RV
   - Expectation
   - Variance and Standard deviation
   - Independence
Random Variables - Definition

- A random variable (RV) is a variable that depends on the outcome of a chance situation.
- A RV is often denoted by capital letters (e.g. $Y$).
- More rigorously, given a sample space $S$ of an experiment, a RV is a function (or rule) that assigns a number to each elementary outcome in the sample space $S$.

Elementary outcome $\rightarrow$ number
Example: Number of nests destructed by predators

Pick 3 nests at random independently and record the number of times nests are destroyed by a predator.

\[ Y = \# \text{ of nests destroyed.} \]

Sample space had 8 elementary outcomes:

<table>
<thead>
<tr>
<th>Nest #</th>
<th>1</th>
<th>2</th>
<th>3</th>
<th>Y</th>
</tr>
</thead>
<tbody>
<tr>
<td>destroyed?</td>
<td>Yes</td>
<td>Yes</td>
<td>Yes</td>
<td>3</td>
</tr>
<tr>
<td></td>
<td>Yes</td>
<td>Yes</td>
<td>No</td>
<td>2</td>
</tr>
<tr>
<td></td>
<td>Yes</td>
<td>No</td>
<td>Yes</td>
<td>2</td>
</tr>
<tr>
<td></td>
<td>Yes</td>
<td>No</td>
<td>No</td>
<td>1</td>
</tr>
<tr>
<td></td>
<td>No</td>
<td>Yes</td>
<td>Yes</td>
<td>2</td>
</tr>
<tr>
<td></td>
<td>No</td>
<td>Yes</td>
<td>No</td>
<td>1</td>
</tr>
<tr>
<td></td>
<td>No</td>
<td>No</td>
<td>Yes</td>
<td>1</td>
</tr>
<tr>
<td></td>
<td>No</td>
<td>No</td>
<td>No</td>
<td>0</td>
</tr>
</tbody>
</table>
Example: Twinning in cows

Sample 100 (independent) cows give birth. Record the number of cows giving birth to twins. Let $Y = \#$ of cows with twins.

<table>
<thead>
<tr>
<th>Cow #</th>
<th>1</th>
<th>2</th>
<th>3</th>
<th>...</th>
<th>100</th>
<th>Y</th>
</tr>
</thead>
<tbody>
<tr>
<td>S=single calf</td>
<td>S</td>
<td>S</td>
<td>S</td>
<td>...</td>
<td>S</td>
<td>0</td>
</tr>
<tr>
<td>T=twins</td>
<td>T</td>
<td>S</td>
<td>S</td>
<td>...</td>
<td>S</td>
<td>1</td>
</tr>
<tr>
<td></td>
<td>S</td>
<td>T</td>
<td>S</td>
<td>...</td>
<td>S</td>
<td>1</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td>...</td>
</tr>
<tr>
<td></td>
<td>S</td>
<td>S</td>
<td>S</td>
<td>...</td>
<td>T</td>
<td>1</td>
</tr>
<tr>
<td></td>
<td>T</td>
<td>T</td>
<td>S</td>
<td>...</td>
<td>S</td>
<td>2</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td>...</td>
</tr>
<tr>
<td></td>
<td>T</td>
<td>T</td>
<td>T</td>
<td>...</td>
<td>T</td>
<td>100</td>
</tr>
</tbody>
</table>
Discrete RV and probability distribution

- Like for data, RV’s can be discrete or continuous, but RV cannot be categorical.
- A RV is discrete if it takes either a finite number of possible values (e.g. \( y_1, \ldots, y_n \)) or at most there is one for every integer (e.g. \( y_1, y_2, \ldots \)).
- The probability distribution of a discrete RV is described by the probability of each possible value of the RV.
Example: Number of nests destructed by predators

Suppose there is a 60% predation rate.
Use $\times = \text{destroyed}, \ o = \text{not destroyed}$

\[
P\{Y = 3\} = P(\times \times \times) = P(\times) \times P(\times) \times P(\times) \\
= .216
\]

\[
P\{Y = 2\} = P\{\times \times o \text{ or } \times o \times \text{ or } o \times \times\} \\
= P\{\times \times o\} + P\{\times o \times\} + P\{o \times \times\} \\
= P(\times)P(\times)P(o) + P(\times)P(o)P(\times) + P(o)P(\times)P(\times) \\
= \ 3 \times 0.144 = 0.432
\]

Similarly $P\{Y = 1\} =$ \\
$= 3 \times 0.096 = 0.288$ \\
and \\
$P\{Y = 0\} =$ \\
$= 0.064.$
We define $p(y) = \mathbb{P}\{Y = y\}$, e.g. $p(2) = \mathbb{P}\{Y = 2\}$.

Frequency table for $Y$:

<table>
<thead>
<tr>
<th>$y$</th>
<th>0</th>
<th>1</th>
<th>2</th>
<th>3</th>
</tr>
</thead>
<tbody>
<tr>
<td>$p(y)$</td>
<td>.064</td>
<td>.288</td>
<td>.432</td>
<td>.216</td>
</tr>
</tbody>
</table>

A line graph shows the probability distribution of $Y$: 
Summary measures

- The probability distribution of a discrete RV $Y$ gives complete information about $Y$ and hence complete information about the population.
- Helpful to have some numerical summaries such as the center/location or spread/variability of the population (as with sample data).

<table>
<thead>
<tr>
<th></th>
<th>population (RV)</th>
<th>sample (observed data)</th>
</tr>
</thead>
<tbody>
<tr>
<td>mean</td>
<td>$\mu_Y$</td>
<td>$\bar{y}$</td>
</tr>
<tr>
<td>variance</td>
<td>$\sigma^2_Y$</td>
<td>$s^2$</td>
</tr>
<tr>
<td>standard deviation</td>
<td>$\sigma_Y$</td>
<td>$s$</td>
</tr>
</tbody>
</table>
Example: Number of nests destructed by predators

Experiment: pick 3 nests at random, independently. Repeat this 1000 times.

With a predation rate of 60% we got

<table>
<thead>
<tr>
<th>y</th>
<th>0</th>
<th>1</th>
<th>2</th>
<th>3</th>
</tr>
</thead>
<tbody>
<tr>
<td>p(y)</td>
<td>.064</td>
<td>.288</td>
<td>.432</td>
<td>.216</td>
</tr>
</tbody>
</table>

Roughly, we will find 0 destroyed nests 64 times, ...

So we expect to find a total # of destroyed nests of:

\[ 0 \times 64 + \text{ } \times \text{ } = 1800 \]

and the average # of destroyed nests per experiment is: \( Y \). This is the expected value of \( Y \).
Expectation of a RV

The **Expectation** of a RV $Y$ is the population mean of the probability distribution of $Y$. Denoted as $\mathbb{E}(Y)$ or $\mu_Y$. Can be thought of as a typical value.

For a discrete RV $Y$, it is

\[
\mathbb{E}(Y) = \sum y \times P\{Y = y\}
\]

summing over all possible values $y$ of the RV $Y$.

In the nest predation problem:

\[
\mathbb{E}(Y) = 0 \times 0.064 + 1 \times 0.288 + 2 \times 0.432 + \ldots = 1.8 \text{ nests}
\]
Properties of expectation

For any RV $Y$,

- $\mathbb{E}(Y + c) = \mathbb{E}(Y) + c$ if $c$ is fixed (non-random) number.
- $\mathbb{E}(2Y) = 2\mathbb{E}(Y)$. More generally, $\mathbb{E}(kY) = k\mathbb{E}(Y)$ if $k$ is fixed (non-random) number.

Ex: If there are exactly 3 eggs per nests, then $\mathbb{E}($eggs destroyed by predators$) =$

If $X =$ # of nests not destroyed by predators then $\mathbb{E}X =$
Variance and Standard deviation of a RV

The variance of a RV $Y$ – noted $\text{var}(Y)$ or $\sigma^2_Y$ – measures the population spread/variability of the distribution of $Y$. Can be thought of as the amount $Y$ typically deviates from $\mu_Y$.

For a discrete RV $Y$, it is

Variance

$$\text{var}(Y) = \mathbb{E}(Y - \mu_Y)^2 = \sum (y - \mu_Y)^2 \times \mathbb{P}\{Y = y\}$$

summing over all possible values $y$ of the RV $Y$.

The standard deviation of $Y$ is $\sigma_Y = \sigma = \sqrt{\text{var}(Y)}$.

Nest destruction by predators:

$$\text{var}(Y) =$$

$$= 0.72$$

and $\sigma_Y = \sqrt{0.72} = 0.848$ nests.
Properties of the variance

For any RV $Y$,

- $\text{var}(Y) \geq 0$. If $\text{var}(Y) = 0$, then $Y$ is a constant, i.e. non-random (no variation).
- $\text{var}(Y + c) = \text{var}(Y)$ if $c$ is a fixed (non-random) number.
- $\text{var}(2Y) = 4\text{var}(Y)$ so $\sigma_{2Y} = 2\sigma_Y$.
- $\text{var}(kY) = k^2\text{var}(Y)$ if $k$ is a fixed (non-random) number.

Ex: If there are exactly 3 eggs per nests, then
$\text{var}(\# \text{ eggs destroyed by predators}) = 6.48$ and
SD is $= 2.55$

If $X = \# \text{ of nests not destroyed by predators}$ then
$\text{var}(X) =$ and $\sigma_X =$
Independent Random Variables

Two RV’s are independent if knowledge of the value of one RV has no effect on the probability about the other RV.

For discrete RV’s, $X$ and $Y$ are independent if

$$\mathbb{P}\{X = x \text{ and } Y = y\} = \mathbb{P}\{X = x\} \mathbb{P}\{Y = y\}$$

for any $x$ and $y$.

Ex: the size and the weight of a cow are probably not independent.

For super fairy wrens, the # of eggs and the destruction by a snake might be independent.
More on two Rvs

For any Rvs $X$ and $Y$,

- $\mathbb{E}(X + Y) = \mathbb{E}(X) + \mathbb{E}(Y)$
- $\mathbb{E}(X - Y) = \mathbb{E}(X) - \mathbb{E}(Y)$
- $\text{var}(X + Y) = \text{var}(X) + \text{var}(Y)$ if $X$ and $Y$ are independent.
- $\text{var}(X - Y) = \text{var}(X) + \text{var}(Y)$ if $X$ and $Y$ are independent.

Ex: $X =$ milk yield from cow #1 and $Y =$ milk yield from cow #2.
How about $X =$ milk yield from cow #1, day #1 and $Y =$ milk yield from the same cow, day #2?