Last lecture

- **Probability models** describe real, complex systems

- Probability models have the following components:
  1. Description of a phenomenon
  2. List of possible outcomes ("sample space", denoted by S)
  3. Probability associated with each outcome

- An **event** is a subset of outcomes from the model

Last lecture

Example of a probability model:

1. We flip the same coin 3 times and count the number of heads
2. The possible outcomes: $S = \{0, 1, 2, 3\}$
3. Probabilities of each event (in order above): 0.125, 0.375, 0.375, 0.125. ($\rightarrow$ extremely close to real-life proportions)

Example events:

1. $A = \{0\}$ = "no heads"
2. $B = \{2, 3\}$ = "more than 1 head"
3. $C = \{0, 3\}$ = "some heads but not all heads"

Last lecture

Event terminology:

1. Events A and B are *disjoint* if they do not share any outcomes
2. Events A and B are *complements* if event A contains every outcome in S that B does *not* contain

Basic rules of probability calculations:

1. For any event A, $0 \leq P(A) \leq 1$
2. Probability of the whole sample space is 1 ($P(S) = 1$)
3. If events A and B are disjoint, $P(A \text{ or } B) = P(A) + P(B)$
4. If B is the complement of A, $P(A) = 1 - P(B)$.

Last lecture

Example events (from 3-flips example):

1. $A = \{0\} =$ "no heads"
2. $B = \{2, 3\} =$ "more than 1 head"
3. $C = \{0, 3\} =$ "some heads but not all heads"

Probability calculation examples:

1. $P(B) = P(2) + P(3) = 0.375 + 0.125 = 0.5$
2. $P(B \text{ or } C) = P(0, 2, \text{ or } 3) = P(0) + P(2) + P(3) =$
   $0.5 + 0.125 = 0.625.$
3. Number 2 using complement rule:
   $P(B \text{ or } C) = P(0, 2, \text{ or } 3) = 1 - P(1) = 1 - 0.375 = 0.625.$

Last lecture

Independence:

- Events A and B are *independent* if knowing whether A occurs does not change the probability of B
- Formally: events A and B are independent **if and only if** the equation
  $$P(A \text{ and } B) = P(A)P(B)$$
  is satisfied.

Last lecture

Examples of independence:

1. Flip a coin twice. $A = \{\text{first flip is heads}\}$, $B = \{\text{second flip is heads}\}$. Second flip heads probability is $\frac{1}{2}$ regardless of first flip. Thus A and B are independent, and $P(\text{both flips are heads}) = P(A)P(B) = \frac{1}{2} \times \frac{1}{2} = \frac{1}{4}$.

2. See boat/plane/criminal example from last lecture. We checked that

$$P(A \text{ and } B) = P(A)P(B)$$

and therefore A and B are independent. The point: whether we decide to (say) catch a boat first or catch any vehicle at random does not change our chances of catching a criminal.

Random Variables (Section 4.3)

New concept: Random Variable

- A **random variable** is a variable whose value is a numerical outcome of a probability model.

Simple example:

- Random phenomenon: Flip a coin 3 times.
- **Number of heads observed** is a random variable.
- **Number of tails observed** is also a random variable.

Random Variables (Section 4.3)

New concept: Random Variable

- A **random variable** is a variable whose value is a numerical outcome of a probability model.

Simple example:

- Random phenomenon: Flip a coin 3 times.
- **Number of heads observed** is a random variable.
- **Number of tails observed** is also a random variable.

Random variables are written as capital letters. So, "let $X =$ number of heads observed" is how to define a r.v. Could also be Y, Z, etc.

Random Variables (Section 4.3)

More complex example: A fair 6-sided die is rolled, and then a card is drawn from a 52-card deck with the aces and face cards removed (so only numbers left).

What are some random variables associated with the situation?

Random Variables (Section 4.3)

More complex example: A fair 6-sided die is rolled, and then a card is drawn from a 52-card deck with the aces and face cards removed (so only numbers left).

What are some random variables associated with the situation?

- X = number on the die
- Y = number of the card
- Z = number of the card plus number on the die $= X + Y$
- $X_2 = X/3$, number on the die divided by 3.

Random Variables (Section 4.3)

- Random variables have properties given by the probability distribution of the phenomenon.

For random variables taking discrete values (die roll, card draw):

- Every possible value a random variable can take has a probability of occuring. (could even be zero or 1)

- The probabilities associated with the possible values of a random variable is called a **probability distribution**.

Random Variables (Section 4.3)

The **probability distribution** of a (discrete) random variable X is a list of the possible values of X and their probabilities.

⋆ If we know the probabilities for outcomes in a random process, we can *derive* the distribution of a random variable.

Random Variables (Section 4.3)

The **probability distribution** of a (discrete) random variable X is a list of the possible values of X and their probabilities.

$\star$ If we know the probabilities for outcomes in a random process, we can *derive* the distribution of a random variable.

Example: Pick a person at random from this class. Let X be *number* associated with their year in school (i.e. Freshman = 1). Suppose we know the class has:

Freshman	Sophomore	Junior	Senior
65%	35%	8%	2%

Then $P(X = 1) = 0.65, P(X = 2) = 0.35, P(X = 3) = 0.08$, and $P(X = 4) = 0.02$.

Random Variables (Section 4.3)

Note: A random variable taking a value (like $X = 2$) is an event.

Sometimes the probability distribution is not direct or obvious.

More complicated example: A bag has 2 marbles: 1 red, 1 green. Three separate times, we draw a marble and place it back in the bag. Let Y be the number of times we draw the red marble.

$\star$ To get the probability distribution, we need to

- List the possible values of Y
- Find the probability of each value

Random Variables (Section 4.3)

A bag has 2 marbles: 1 red, 1 green. Three separate times, we draw a marble and place it back in the bag. Let Y be the number of times we draw the red marble.

1. What are the possible values of Y?

Random Variables (Section 4.3)

A bag has 2 marbles: 1 red, 1 green. Three separate times, we draw a marble and place it back in the bag. Let Y be the number of times we draw the red marble.

1. What are the possible values of Y? Answer: $0, 1, 2, 3$

Random Variables (Section 4.3)

A bag has 2 marbles: 1 red, 1 green. Three separate times, we draw a marble and place it back in the bag. Let Y be the number of times we draw the red marble.

1. What are the possible values of Y? Answer: $0, 1, 2, 3$
2. What are the probabilities of each value?

Let's solve #2 in multiple ways.

Random Variables (Section 4.3)

A bag has 2 marbles: 1 red, 1 green. Three separate times, we draw a marble and place it back in the bag. Let Y be the number of times we draw the red marble.

Option 1: Brute-force counting. The possible outcomes of this experiment are: green-green-green, green-green-red, green-red-green, red-green-green, red-red-green, red-green-red, green-red-red, red-red-red. Therefore,

$$
\begin{aligned}
P(Y = 0) &= 1/8 = 0.125 \\
P(Y = 1) &= 3/8 = 0.375 \\
P(Y = 2) &= 3/8 = 0.375 \\
P(Y = 3) &= 1/8 = 0.125
\end{aligned}
$$

Random Variables (Section 4.3)

A bag has 2 marbles: 1 red, 1 green. Three separate times, we draw a marble and place it back in the bag. Let Y be the number of times we draw the red marble.

Option 2: Probability rules. The draws are independent, since we place the marble back each time. Therefore

$$\begin{aligned} P(Y=0) =& P(\textit{greengreengreen}) = P(\textit{green})P(\textit{green})P(\textit{green}) \\ =& \frac{1}{2}\cdot\frac{1}{2}\cdot\frac{1}{2} = \frac{1}{8} = 0.125, \text{ and} \end{aligned}$$

$$\begin{aligned} P(Y=1) =& P(\textit{redgreengreen} \text{ or } \textit{greenredgreen} \text{ or } \textit{greengreenred}) \\ =& P(\textit{redgreengreen}) + P(\textit{greenredgreen}) + (\textit{greengreenred}) \\ =& \frac{1}{8}+\frac{1}{8}+\frac{1}{8} = \frac{3}{8} = 0.375 \end{aligned}$$

Random Variables (Section 4.3)

A bag has 2 marbles: 1 red, 1 green. Three separate times, we draw a marble and place it back in the bag. Let Y be the number of times we draw the red marble.

- For this example, brute force counting and using the rules seemed equal amounts of work.
- But for most examples, using probability rules is **much** easier.

Random Variables (Section 4.3)

Example 2: Lucy, Darla, Stefon, Henrique are freshman, Asha and Liz are sophomores, and Sally is a senior. For each of three separate (and independent) class questions, a professor calls on one student randomly. Let $F =$ the number of total freshman called upon.

What are the possible values for F?

Example 2: Lucy, Darla, Stefon, Henrique are freshman, Asha and Liz are sophomores, and Sally is a senior. For each of three separate (and independent) class questions, a professor calls on one student randomly. Let $F =$ the number of total freshman called upon.

What are the possible values for F? $0, 1, 2, 3$
$\star$ Note: same as in marble example, but different probabilities!

Let's find the probability distribution for F.

Example 2: Lucy, Darla, Stefon, Henrique are freshman, Asha and Liz are sophomores, and Sally is a senior. For each of three separate (and independent) class questions, a professor calls on one student randomly. Let F = the number of total freshman called upon.

Option 1: Brute-force counting. It would take very long to list all the possible outcomes and count them.

Instead, we will find F's distribution using probability rules

Random Variables (Section 4.3)

Example 2: Lucy, Darla, Stefon, Henrique are freshman, Asha and Liz are sophomores, and Sally is a senior. For each of three separate (and independent) class questions, a professor calls on one student randomly. Let F = the number of total freshman called upon.

Option 2: Probability rules. Let "f" denote the outcome of a freshman for a single draw, and "u" and upperclass-student (non-freshman). $P(f) = 4/7$ and by the complement rule, $P(u) = 1 - P(f) = 3/7$. Therefore

$$P(F = 0) = P(uuu) = P(u)P(u)P(u) = \frac{3}{7} \cdot \frac{3}{7} \cdot \frac{3}{7} \approx 0.08$$

Random Variables (Section 4.3)

Example 2: Lucy, Darla, Stefon, Henrique are freshman, Asha and Liz are sophomores, and Sally is a senior. For each of three separate (and independent) class questions, a professor calls on one student randomly. Let F = the number of total freshman called upon.

Option 2: Probability rules. Let "f" = freshman for a single draw, and "u" = non-freshman. Then $P(f) = 4/7$ and by the complement rule, $P(u) = 1 - P(f) = 3/7$.

$$
\begin{aligned}
P(F=1) =& P(fuu \text{ or } ufu \text{ or } uuf) = P(fuu) + P(ufu) + (uuf) \\
=& P(f)P(u)P(u) + P(u)P(f)P(u) + P(u)P(u)P(f) \\
=& \frac{4}{7} \cdot \frac{3}{7} \cdot \frac{3}{7} + \frac{3}{7} \cdot \frac{4}{7} \cdot \frac{3}{7} + \frac{3}{7} \cdot \frac{3}{7} \cdot \frac{4}{7} \\
=& 3 \cdot \frac{4}{7} \cdot \frac{3}{7} \cdot \frac{3}{7} \approx 0.31
\end{aligned}
$$

Example 2: Lucy, Darla, Stefon, Henrique are freshman, Asha and Liz are sophomores, and Sally is a senior. For each of three separate (and independent) class questions, a professor calls on one student randomly. Let F = the number of total freshman called upon.

Option 2: Probability rules. Let "f" = freshman for a single draw, and "u" = non-freshman. Then $P(f) = 4/7$ and by the complement rule, $P(u) = 1 - P(f) = 3/7$. Full probability distribution for F:

Y	0	1	2	3
Probability	0.08	0.31	0.42	0.19

Random Variables (Section 4.3)

Example 2: Lucy, Darla, Stefon, Henrique are freshman, Asha and Liz are sophomores, and Sally is a senior. For each of three separate (and independent) class questions, a professor calls on one student randomly. Let $F =$ the number of total freshman called upon. Full probability distribution for F:

Y	0	1	2	3
Probability	0.08	0.31	0.42	0.19

Note: many different kinds of events with r.v.'s. For example:

- $P(F > 1) =$

Random Variables (Section 4.3)

Example 2: Lucy, Darla, Stefon, Henrique are freshman, Asha and Liz are sophomores, and Sally is a senior. For each of three separate (and independent) class questions, a professor calls on one student randomly. Let F = the number of total freshman called upon. Full probability distribution for F:

Y	0	1	2	3
Probability	0.08	0.31	0.42	0.19

Note: many different kinds of events with r.v.'s. For example:

- $P(F > 1) = P(F = 2 \text{ or } F = 3) = P(F = 2) + P(F = 3) = 0.42 + 0.19 = 0.61$
- $P(F < 3) =$

Random Variables (Section 4.3)

Example 2: Lucy, Darla, Stefon, Henrique are freshman, Asha and Liz are sophomores, and Sally is a senior. For each of three separate (and independent) class questions, a professor calls on one student randomly. Let $F =$ the number of total freshman called upon. Full probability distribution for F:

Y	0	1	2	3
Probability	0.08	0.31	0.42	0.19

Note: many different kinds of events with r.v.'s. For example:

- $P(F > 1) = P(F = 2 \text{ or } F = 3) = P(F = 2) + P(F = 3) = 0.42 + 0.19 = 0.61$
- $P(F < 3) = 1 - P(F = 3) = 1 - 0.19 = 0.81$

Random Variables (Section 4.3)

Recall: data can be **discrete** (taking countable number of values) or **continuous** (taking any real numbers in a *range* of values).

Discrete random variable: a random variable that can be a countable (list-able) number of values. Each value has probability in $[0, 1]$ and *at least one* value has positive probability.
Example: $F =$ number of freshman in 3 independent draws.

Random Variables (Section 4.3)

Recall: data can be **discrete** (taking countable number of values) or **continuous** (taking any real numbers in a *range* of values).

Discrete random variable: a random variable that can be a countable (list-able) number of values. Each value has probability in $[0, 1]$ and *at least one* value has positive probability.
Example: $F =$ number of freshman in 3 independent draws.

Continuous random variable: a random variable that can be any real number in a range of values. Each *range* of values has probability in $[0, 1]$ and *at least one* range has positive probability.
Example: $H =$ height of a randomly drawn student from a large population.

Random Variables (Section 4.3)

⋆ For **discrete** random variables, can write out probs.

Example: F = number of freshman in 3 independent draws:

Y	0	1	2	3
Probability	0.08	0.31	0.42	0.19

Random Variables (Section 4.3)

⋆ For **discrete** random variables, can write out probs.

Example: $F =$ number of freshman in 3 independent draws:

Y	0	1	2	3
Probability	0.08	0.31	0.42	0.19

⋆ For **continuous** random variables, **density** gives probabilities.

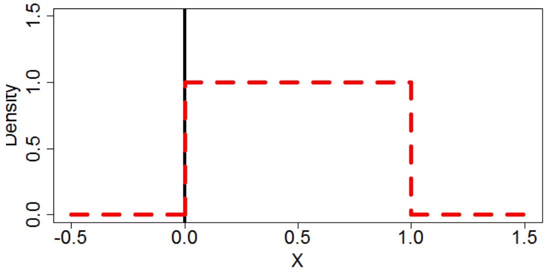
Example: If student heights are $\mathcal{N}(\mu, \sigma)$, then probabilities about $H =$ height of a randomly drawn student follow Normal probabilities (more on this to follow).

Continuous Random Variables (Section 4.3)

- Let X be a continuous random variable
- Events about X are: "X in a *range* $[a, b]$" $= \{a \leq X \leq b\}$
- Interested in $P(a \leq X \leq b)$
- $P(a \leq X \leq b) =$ area under X's density between a and b

Continuous Random Variables (Section 4.3)

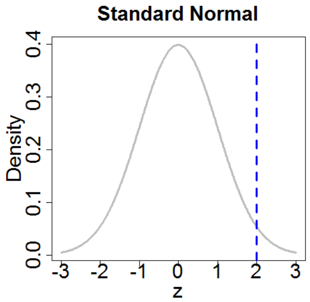
Example 1: X has **uniform** distribution if its probabilities follow uniform density:



- $P(0.5 \leq X \leq 0.75) = 1 \times (0.75 - 0.5) = 0.25$
- $P(X \leq 0.2) = 1 \times (0.2 - 0) = 0.2$

Continuous Random Variables (Section 4.3)

Example 2: Z has **Standard Normal** distribution if its probabilities follow Standard Normal density:



- $P(Z \leq 0.75) =$ area under curve below $z = 0.75$
- $P(-0.2 \geq Z) =$ area under curve above $z = -0.2$

Continuous Random Variables (Section 4.3)

Technicality of continuous random variables:

- Area under any *point* is 0.
- So $P(X = a) = 0$ for any a

Continuous Random Variables (Section 4.3)

Technicality of continuous random variables:

- Area under any *point* is 0.
- So $P(X = a) = 0$ for any a
- ... and $\{a \leq X \leq b\}$ and $\{a < X < b\}$ are the same
- ... and $P(a \leq X \leq b) = P(a < X < b)$

Continuous Random Variables (Section 4.3)

Technicality of continuous random variables:

- Area under any *point* is 0.
- So $P(X = a) = 0$ for any a
- ... and $\{a \leq X \leq b\}$ and $\{a < X < b\}$ are the same
- ... and $P(a \leq X \leq b) = P(a < X < b)$
- Above not necessarily true if X is discrete

Continuous Random Variables (Section 4.3)

Probabilities for general Normal random variables:

- Let X have distribution $\mathcal{N}(10, 2)$
- Consider the event $A = \{X < 8\}$. What is $P(A)$?

Continuous Random Variables (Section 4.3)

Probabilities for general Normal random variables:

- Let X have distribution $\mathcal{N}(10, 2)$
- Consider the event $A = \{X < 8\}$. What is $P(A)$?

Option 1: use $68 - 95 - 99.7$ rule:

- 68% of area between $[10 - 2, 10 + 2] = [8, 12]$
- Density is symmetric, so $32\%/2 = 16\%$ to left of $[8, 12]$
- Thus $P(X < 8) = 0.16$

Continuous Random Variables (Section 4.3)

What is $P(X < 8)$? Option 2: transform the variable.

Fact: If X is $\mathcal{N}(\mu, \sigma)$, the random variable $Z = \frac{X-\mu}{\sigma}$ is $\mathcal{N}(0,1)$.

- The event $\{X < 8\}$ is the same as $\left\{\frac{X-10}{2} < \frac{8-10}{2}\right\}$
- Let $Z = \frac{X-10}{2}$. Then Z is $\mathcal{N}(0,1)$
- So the event $\left\{\frac{X-10}{2} < \frac{8-10}{2}\right\}$ is the same as $\{Z < -1\}$
- Therefore $P(X < 8) = P(Z < -1) = 0.16$.

Continuous Random Variables (Section 4.3)

Example: Combining densities and probability rules.
Trains on a certail railway are approximately $\mathcal{N}(3000\text{ft}, 600\text{ft})$. Assume trains that pass by are independent. What is the probability of observing two trains in a row that are over 4000ft long?

Continuous Random Variables (Section 4.3)

Example: Combining densities and probability rules.
Trains on a certail railway are approximately $\mathcal{N}(3000\text{ft}, 600\text{ft})$. Assume trains that pass by are independent. What is the probability of observing two trains in a row that are over 4000ft long?

- Let $A = \{\text{observe 1 train that is over 4000ft}\}$.
- $P(A) =$? Standardize (previous slide) and use Table A
- Then $P(\text{2 independent trains over 4000ft}) = P(A)^2$