 Updates

- HW2 is online on the course website and on webcafe (see announcement). It will be due **Wed, Jun. 9**.
- Lectures are now available online before the beginning of class (if you would like to print out the slides before hand). Try not to peak ahead!
- I plan on grading HW1 over the weekend.
Example 1

- This is known as the *points problem*. It is somewhat similar to the previous example with sequences of successes and failures.

- **Question**: Independent trials resulting in a success with probability $p$ and a failure with probability $1 - p$ are performed. What is the probability that $n$ successes occur before $m$ failures?
Example 1

This is known as the *points problem*. It is somewhat similar to the previous example with sequences of successes and failures.

**Question**: Independent trials resulting in a success with probability $p$ and a failure with probability $1 - p$ are performed. What is the probability that $n$ successes occur before $m$ failures?

Another way of thinking of it is: $A$ and $B$ are playing a game such that $A$ gains 1 point when a success occurs and $B$ gains 1 point when a failure occurs. Then, the desired probability is the probability that $A$ would win if the game were to be continued where $A$ needed to score $n$ points and $B$ needed to score $m$ more points to win.
Example 1 (cont.)

- Denote $P_{n,m}$ as the probability that $n$ successes occur before $m$ failures. There are two ways of solving this.

- **Solution 1:** It is necessary and sufficient that there be at least $n$ successes in the first $m + n - 1$ trials. What is that implying?
Example 1 (cont.)

- Denote $P_{n,m}$ as the probability that $n$ successes occur before $m$ failures. There are two ways of solving this.

- **Solution 1**: It is necessary and sufficient that there be at least $n$ successes in the first $m + n - 1$ trials. What is that implying?

- There were **at least** $n$ successes and **at most** $m - 1$, which would mean that $n$ successes most certainly happened before $m$ failures.
Example 1 (cont.)

- Denote $P_{n,m}$ as the probability that $n$ successes occur before $m$ failures. There are two ways of solving this.

- **Solution 1**: It is necessary and sufficient that there be at least $n$ successes in the first $m + n - 1$ trials. What is that implying?

- There were **at least** $n$ successes and **at most** $m - 1$, which would mean that $n$ successes most certainly happened before $m$ failures.

- On the other hand, if in the first $m + n - 1$ trials there are fewer than $n$ successes, then there would have to be at least $m$ failures in the same number of trails and, thus, $n$ successes would not occur before $m$ failures.
Example 1 (cont.)

- Denote $P_{n,m}$ as the probability that $n$ successes occur before $m$ failures. There are two ways of solving this.

- **Solution 1**: It is necessary and sufficient that there be at least $n$ successes in the first $m + n - 1$ trials. What is that implying?

  There were **at least** $n$ successes and **at most** $m - 1$, which would mean that $n$ successes most certainly happened before $m$ failures.

  On the other hand, if in the first $m + n - 1$ trials there are fewer than $n$ successes, then there would have to be at least $m$ failures in the same number of trails and, thus, $n$ successes would not occur before $m$ failures.

- How does this relate to that previous example?
Example 1 (cont.)

- Denote $P_{n,m}$ as the probability that $n$ successes occur before $m$ failures. There are two ways of solving this.

- **Solution 1:** It is necessary and sufficient that there be at least $n$ successes in the first $m + n - 1$ trials. What is that implying?

- There were **at least** $n$ successes and **at most** $m - 1$, which would mean that $n$ successes most certainly happened before $m$ failures.

- On the other hand, if in the first $m + n - 1$ trials there are fewer than $n$ successes, then there would have to be at least $m$ failures in the same number of trials and, thus, $n$ successes would not occur before $m$ failures.

- How does this relate to that previous example?
  - **Hint:** probability relating to $n$ successes in $m + n - 1$ trials.
Example 1 (cont.)

The probability of exactly $k$ successes in $m + n - 1$ trials (we showed) is

$$\binom{m + n - 1}{k} p^k (1 - p)^{(m+n-1)-k}$$

So, it follows that the *desired* probability of $n$ successes before $m$ failures (i.e. $P_{n,m}$) is . . .
Example 1 (cont.)

- The probability of exactly $k$ successes in $m+n-1$ trials (we showed) is
  \[
  \binom{m+n-1}{k} p^k (1-p)^{m+n-1-k}
  \]

- So, it follows that the desired probability of $n$ successes before $m$ failures (i.e. $P_{n,m}$) is . . .
  \[
P_{n,m} = \sum_{k=n}^{m+n-1} \binom{m+n-1}{k} p^k (1-p)^{m+n-1-k}
  \]
Example 1 (cont.)

- **Solution 2**: Think about breaking down the \( P_{n,m} \) to previous "events":

\[
P_{n,m} = pP_{n-1,m} + (1 - p)P_{n,m-1} \quad n \geq 1, \quad m \geq 1
\]

Why can we do this?
Example 1 (cont.)

**Solution 2:** Think about breaking down the $P_{n,m}$ to previous "events":

$$P_{n,m} = pP_{n-1,m} + (1 - p)P_{n,m-1} \quad n \geq 1, m \geq 1$$

Why can we do this?

- The last trial is a success $p$ percent of the time and a failure $1 - p$ percent of the time.
- When it is a success, the "leftover" sequence is $n - 1$ successes in $n + m - 1$ trials, so we can consider that probability. If it is a failure, then the "leftover" sequence is $n$ successes in $n + m - 1$ trials, so we consider that probability related to it.
Example 1 (cont.)

- **Solution 2**: Think about breaking down the $P_{n,m}$ to previous "events":

  $$P_{n,m} = pP_{n-1,m} + (1 - p)P_{n,m-1} \quad n \geq 1, m \geq 1$$

Why can we do this?

- The last trial is a success $p$ percent of the time and a failure $1 - p$ percent of the time.
- When it is a success, the "leftover" sequence is $n - 1$ successes in $n + m - 1$ trials, so we can consider that probability. If it is a failure, then the "leftover" sequence is $n$ successes in $n + m - 1$ trials, so we consider that probability related to it.
- All that is left is stating that $P_{n,0} = 0$ (i.e. the probability that $n$ successes occur before 0 failures is 0) and $P_{0,m} = 1$ (i.e the probability that 0 successes occur before $m$ failures is 1) and we can use a form of recursion to get the answer.
Example 2

- Another famous problem known as *Polya’s urn*.
- Suppose there is an urn which initially contains one white chip and one yellow chip. At each stage, we select one chip from all that are in the urn *uniformly* at random. Then we replace that chip and add another chip of the same color as the chip just selected.
  - By *uniformly*, we mean each chip has an equal chance of being drawn.

Example: 1 White chip drawn \(\Rightarrow\) urn has two white chips, one yellow chip.  
2 White chip drawn \(\Rightarrow\) urn has three white chips, one yellow chip.  
3 Yellow chip drawn \(\Rightarrow\) urn has three white chips, two yellow chips.  

Question: What is the probability that the first four chips selected are yellow and the next two are white?
Example 2

- Another famous problem known as Polyá’s urn.
- Suppose there is an urn which initially contains one white chip and one yellow chip. At each stage, we select one chip from all that are in the urn uniformly at random. Then we replace that chip and add another chip of the same color as the chip just selected.
  - By uniformly, we mean each chip has an equal chance of being drawn.
- **Ex:**
  1. White chip drawn ⇒ urn has two white chips, one yellow chip.
  2. White chip drawn ⇒ urn has three white chips, one yellow chip.
  3. Yellow chip drawn ⇒ urn has three white chips, two yellow chips.
Example 2

- Another famous problem known as *Polya’s urn*.
- Suppose there is an urn which initially contains one white chip and one yellow chip. At each stage, we select one chip from all that are in the urn *uniformly* at random. Then we replace that chip and add another chip of the same color as the chip just selected.
  - By *uniformly*, we mean each chip has an equal chance of being drawn.
- **Ex:**
  1. White chip drawn $\Rightarrow$ urn has two white chips, one yellow chip.
  2. White chip drawn $\Rightarrow$ urn has three white chips, one yellow chip.
  3. Yellow chip drawn $\Rightarrow$ urn has three white chips, two yellow chips.
- **Question:** What is the probability that the first four chips selected are yellow and the next two are white?
Example 2 (cont.)

**Solution**: Let $E_i$ be the event that the $i$th chip selected is yellow. Then the probability we are looking for is . . .
Example 2 (cont.)

Solution: Let $E_i$ be the event that the $i$th chip selected is yellow. Then the probability we are looking for is . . .

$$P(E_1 E_2 E_3 E_4 E_5^c E_6^c) = P(E_1)P(E_2 | E_1)P(E_3 | E_1 E_2)P(E_4 | E_1 E_2 E_3$$

$$P(E_5^c | E_1 E_2 E_3 E_4)P(E_6^c | E_1 E_2 E_3 E_4 E_5^c)$$

We have to use the multiplication rule because what chips that are left in the urn depends on the previous events. What are some of these values?

The first probability is . . .
Example 2 (cont.)

Solution: Let $E_i$ be the event that the $i$th chip selected is yellow. Then the probability we are looking for is . . .

$$P(E_1 E_2 E_3 E_4 E_5^c E_6^c) = P(E_1) P(E_2 | E_1) P(E_3 | E_1 E_2) P(E_4 | E_1 E_2 E_3$$

$$P(E_5^c | E_1 E_2 E_3 E_4) P(E_6^c | E_1 E_2 E_3 E_4 E_5^c)$$

We have to use the multiplication rule because what chips that are left in the urn depends on the previous events. What are some of these values?

The first probability is . . .

$$P(E_1) = \frac{1}{2}$$

If $E_1$ occurs, then there are two yellow chips and one white chip left in the urn, so . . .
Example 2 (cont.)

- **Solution:** Let $E_i$ be the event that the $i$th chip selected is yellow. Then the probability we are looking for is ...

$$P(E_1 E_2 E_3 E_4 E_5^c E_6^c) = P(E_1) P(E_2|E_1) P(E_3|E_1 E_2) P(E_4|E_1 E_2 E_3^c) P(E_5^c|E_1 E_2 E_3 E_4) P(E_6^c|E_1 E_2 E_3 E_4 E_5^c)$$

We have to use the **multiplication rule** because what chips that are left in the urn depends on the previous events. What are some of these values?

- The first probability is ...

$$P(E_1) = \frac{1}{2}$$

- If $E_1$ occurs, then there are two yellow chips and one white chip left in the urn, so ...

$$P(E_2|E_1) = \frac{2}{3}$$
Example 2 (cont.)

- The other probabilities follow like this:

\[
P(E_3 | E_1 E_2) = \frac{3}{4} \\
P(E_4 | E_1 E_2 E_3) = \frac{4}{5} \\
P(E_5 | E_1 E_2 E_3 E_4) = \frac{1}{6} \\
P(E_6 | E_1 E_2 E_3 E_4 E_5) = \frac{2}{7} \\
\]

Thus, the final probability is:

\[
P(E_1 E_2 E_3 E_4 E_5 E_6) = \frac{1}{2} \cdot \frac{2}{3} \cdot \frac{3}{4} \cdot \frac{4}{5} \cdot \frac{1}{6} \cdot \frac{2}{7} = \frac{4}{2 \cdot 7} 
\]
Example 2 (cont.)

The other probabilities follow like this:

\[ P(E_3|E_1E_2) = \frac{3}{4} \]
Example 2 (cont.)

- The other probabilities follow like this:

\[ P(E_3|E_1 E_2) = \frac{3}{4} \]
\[ P(E_4|E_1 E_2 E_3) = \frac{4}{5} \]
Example 2 (cont.)

- The other probabilities follow like this:

\[
P(E_3 | E_1 E_2) = \frac{3}{4}
\]

\[
P(E_4 | E_1 E_2 E_3) = \frac{4}{5}
\]

\[
P(E_5^c | E_1 E_2 E_3 E_4) = \frac{1}{6}
\]
Example 2 (cont.)

- The other probabilities follow like this:

\[
P(E_3 | E_1 E_2) = \frac{3}{4}
\]

\[
P(E_4 | E_1 E_2 E_3) = \frac{4}{5}
\]

\[
P(E_5^c | E_1 E_2 E_3 E_4) = \frac{1}{6}
\]

\[
p(E_6^c | E_1 E_2 E_3 E_4 E_5^c) = \frac{2}{7}
\]
Example 2 (cont.)

- The other probabilities follow like this:

\[
P(E_3|E_1E_2) = \frac{3}{4}
\]
\[
P(E_4|E_1E_2E_3) = \frac{4}{5}
\]
\[
P(E_5^c|E_1E_2E_3E_4) = \frac{1}{6}
\]
\[
p(E_6^c|E_1E_2E_3E_4E_5^c) = \frac{2}{7}
\]

- Thus, the final probability is:

\[
P(E_1E_2E_3E_4E_5^cE_6^c) = \frac{1}{2} \cdot \frac{2}{3} \cdot \frac{3}{4} \cdot \frac{4}{5} \cdot \frac{1}{6} \cdot \frac{2}{7} = \frac{4!2!}{7!}
\]
Example 2 (cont.)

**Question**: What is the probability that the first two chips selected are white and the next four are yellow?
Example 2 (cont.)

- **Question:** What is the probability that the first two chips selected are white and the next four are yellow?

- **Solution:** The probability we are looking for here is . . .

\[
P(E_1^c E_2^c E_3 E_4 E_5 E_6) = P(E_1^c) P(E_2^c | E_1^c) P(E_3 | E_1^c E_2^c) P(E_4 | E_1^c E_2^c E_3^c E_4 E_5^c) \\
\quad P(E_5^c | E_1^c E_2^c E_3 E_4) P(E_6 | E_1^c E_2^c E_3 E_4 E_5^c)
\]

- Is that the same probability as before? Let’s find out . . .
Example 2 (cont.)

- Again, we use the **multiplication rule** getting the following values:
Example 2 (cont.)

Again, we use the **multiplication rule** getting the following values:

\[ P(E_1^c) = \frac{1}{2} \]
Example 2 (cont.)

- Again, we use the **multiplication rule** getting the following values:

\[
P(E_1^c) = \frac{1}{2}
\]

\[
P(E_2^c|E_1^c) = \frac{2}{3}
\]
Example 2 (cont.)

Again, we use the **multiplication rule** getting the following values:

\[
P(E_1^c) = \frac{1}{2}
\]

\[
P(E_2^c | E_1^c) = \frac{2}{3}
\]

\[
P(E_3 | E_1^c E_2^c) = \frac{1}{4}
\]
Example 2 (cont.)

- Again, we use the **multiplication rule** getting the following values:

\[
\begin{align*}
P(E_1^c) &= \frac{1}{2} \\
P(E_2^c|E_1^c) &= \frac{2}{3} \\
P(E_3|E_1^c E_2^c) &= \frac{1}{4} \\
P(E_4|E_1^c E_2^c E_3) &= \frac{2}{5}
\end{align*}
\]
Example 2 (cont.)

Again, we use the **multiplication rule** getting the following values:

\[
\begin{align*}
P(E_1^c) &= \frac{1}{2} \\
P(E_2^c|E_1^c) &= \frac{2}{3} \\
P(E_3|E_1^c E_2^c) &= \frac{1}{4} \\
P(E_4|E_1^c E_2^c E_3) &= \frac{2}{5} \\
P(E_5|E_1^c E_2^c E_3 E_4) &= \frac{3}{6}
\end{align*}
\]
Example 2 (cont.)

Again, we use the **multiplication rule** getting the following values:

\[
\begin{align*}
P(E_1^c) &= \frac{1}{2} \\
P(E_2^c | E_1^c) &= \frac{2}{3} \\
P(E_3 | E_1^c E_2^c) &= \frac{1}{4} \\
P(E_4 | E_1^c E_2^c E_3) &= \frac{2}{5} \\
P(E_5 | E_1^c E_2^c E_3 E_4) &= \frac{3}{6} \\
p(E_6 | E_1^c E_2^c E_3 E_4 E_5) &= \frac{4}{7}
\end{align*}
\]
Example 2 (cont.)

Thus, the final probability is:

$$P(E_1^c E_2^c E_3 E_4 E_5 E_6) = \frac{1}{2} \cdot \frac{2}{3} \cdot \frac{1}{4} \cdot \frac{2}{5} \cdot \frac{3}{6} \cdot \frac{4}{7} = \frac{4!2!}{7!}$$

So it is the same.
Example 2 (cont.)

Thus, the final probability is:

\[
P(E_1^c E_2^c E_3 E_4 E_5 E_6) = \frac{1}{2} \cdot \frac{2}{3} \cdot \frac{1}{4} \cdot \frac{2}{5} \cdot \frac{3}{6} \cdot \frac{4}{7} = \frac{4!2!}{7!}
\]

So it is the same.

**Question:** What is the probability that the first 6 selected chips include 4 yellow and 2 white?
Example 2 (cont.)

**Solution**: We showed that each sequence is, in fact, equally likely. How many such sequences are there?
Example 2 (cont.)

- **Solution**: We showed that each sequence is, in fact, equally likely. How many such sequences are there? \( \binom{6}{4} \). Thus, the probability can be expressed as:
Example 2 (cont.)

- Solution: We showed that each sequence is, in fact, equally likely. How many such sequences are there?
- \( \binom{6}{4} \). Thus, the probability can be expressed as:

\[
P(4 \text{ of first 6 chips are yellow}) = \binom{6}{4} \frac{4!2!}{7!}
\]
Example 2 (cont.)

- **Solution**: We showed that each sequence is, in fact, equally likely. How many such sequences are there?

\[
\binom{6}{4}. \text{ Thus, the probability can be expressed as:}
\]

\[
P(4 \text{ of first 6 chips are yellow}) = \binom{6}{4} \frac{4!2!}{7!} = \frac{6!}{4!2!7!} \frac{4!2!}{7!} = \frac{1}{7}
\]
Example 2 (cont.)

**Solution:** We showed that each sequence is, in fact, equally likely. How many such sequences are there? \( \binom{6}{4} \). Thus, the probability can be expressed as:

\[
P(4 \text{ of first 6 chips are yellow}) = \binom{6}{4} \frac{4!2!}{7!}
\]

\[
= \frac{6!}{4!2!7!} = \frac{1}{7}
\]

We can generalize this to:

\[
P(m \text{ of first } n \text{ chips are yellow}) = \frac{n!}{m!(n-m)!} \frac{4!2!}{7!} = \frac{1}{n+1}
\]

Note that the answer doesn’t depend on \( m \), the number of yellow chips selected.
Formalization

- **Def**: Events $E_1$ and $E_2$ are *conditionally independent* given $F$ if given that $F$ occurs, the conditional probability that $E_1$ occurs is unchanged by information as to whether or not $E_2$ occurs.

- More formally, *conditional independence* holds if

$$P(E_1|E_2F) = P(E_1|F)$$

or, equivalently,

$$P(E_1E_2|F) = P(E_1|F)P(E_2|F)$$

- Note that the problems related are very similar to previous examples. The only difference is that it all depends (i.e. conditional) on some event occurring.
  - For examples, see section 3.5.
Motivation

Until this point, much of what has been learned is to prepare for Chapter 4 and onwards. Much of what we have done might not make sense in the grand scheme, but *random variables* will serve to do that. And, as you’ll see, they are *vital* to our understanding.
Motivation

- Until this point, much of what has been learned is to prepare for Chapter 4 and onwards. Much of what we have done might not make sense in the grand scheme, but *random variables* will serve to do that. And, as you’ll see, they are *vital* to our understanding.

- In most problems, we aren’t interested in the actual experiment, but the *outcomes* of those experiments.

- **Ex:** When we toss 10 coins, we may be interested in the total number of heads, instead of the outcome of each coin.
Motivation

- Until this point, much of what has been learned is to prepare for Chapter 4 and onwards. Much of what we have done might not make sense in the grand scheme, but random variables will serve to do that. And, as you’ll see, they are vital to our understanding.

- In most problems, we aren’t interested in the actual experiment, but the outcomes of those experiments.

  - **Ex:** When we toss 10 coins, we may be interested in the total number of heads, instead of the outcome of each coin.

  - In other cases, we could be interested in the coarsest portion of the experiment.

  - **Ex:** Everyone in the class takes the midterm and we want each person’s score.
Formalization

- **Def**: Let $S$ be the sample space with some relating probability $P$. A *random variable* $X$ is a function with domain $S$.

- Essentially, a *random variable* reflects the aspect of a random experiment that is of interest to us.
Formalization

**Def:** Let $S$ be the sample space with some relating probability $P$. A *random variable* $X$ is a function with domain $S$.

Essentially, a *random variable* reflects the aspect of a random experiment that is of interest to us.

**Ex:** $X =$ the total number of heads out of 10 flips. $X : S \rightarrow \{1, 2, \ldots, 10\}$. This is a *discrete random variable* because it takes on at most a countable number of outcomes (note: countable refers to the analysis term).
Formalization

- **Def:** Let \( S \) be the sample space with some relating probability \( P \). A *random variable* \( X \) is a function with domain \( S \).

- Essentially, a *random variable* reflects the aspect of a random experiment that is of interest to us.

  - **Ex:** \( X = \) the total number of heads out of 10 flips. 
    \( X : S \rightarrow \{1, 2, \ldots, 10\} \). This is a *discrete random variable* because it takes on at most a countable number of outcomes (note: countable refers to the analysis term).

  - **Ex:** \( Y = \) the score of the midterm by James. 
    \( Y : S \rightarrow \{y : 0 \leq y \leq 100\} \). This is a *continuous random variable* because it could denote an uncountable number of outcomes.
Example 3

Suppose our experiment consists of tossing three coins. If we let $Y$ denote the number of heads that appear, then $Y$ is a random variable taking on one of four values: 0, 1, 2 and 3.

**Question:** What is the probability relating to each of the four outcomes?
Example 3

Suppose our experiment consists of tossing three coins. If we let $Y$ denote the number of heads that appear, then $Y$ is a random variable taking on one of four values: 0, 1, 2 and 3.

**Question:** What is the probability relating to each of the four outcomes?

**Solution:**

$$P(Y = 0) = P(\{TTT\}) = \frac{1}{8}$$
Example 3

- Suppose our experiment consists of tossing three coins. If we let $Y$ denote the number of heads that appear, then $Y$ is a random variable taking on one of four values: 0, 1, 2 and 3.

- **Question**: What is the probability relating to each of the four outcomes?

- **Solution**:

  $$P(Y = 0) = P\left(\{TTT\}\right) = \frac{1}{8}$$

  $$P(Y = 1) = P\left(\{TTH\}, \{THT\}, \{HTT\}\right) = \frac{3}{8}$$
Example 3

Suppose our experiment consists of tossing three coins. If we let $Y$ denote the number of heads that appear, then $Y$ is a random variable taking on one of four values: 0, 1, 2 and 3.

**Question:** What is the probability relating to each of the four outcomes?

**Solution:**

\[
P(Y = 0) = P(\{TTT\}) = \frac{1}{8}
\]
\[
P(Y = 1) = P(\{TTH\}, \{THT\}, \{HTT\}) = \frac{3}{8}
\]
\[
P(Y = 2) = P(\{THH\}, \{HTH\}, \{HHT\}) = \frac{3}{8}
\]
Example 3

- Suppose our experiment consists of tossing three coins. If we let $Y$ denote the number of heads that appear, then $Y$ is a random variable taking on one of four values: 0, 1, 2 and 3.
- **Question**: What is the probability relating to each of the four outcomes?
- **Solution**:

  \[
  P(Y = 0) = P(\{TTT\}) = \frac{1}{8}
  \]

  \[
  P(Y = 1) = P(\{TTH\}, \{THT\}, \{HTT\}) = \frac{3}{8}
  \]

  \[
  P(Y = 2) = P(\{THH\}, \{HTH\}, \{HHT\}) = \frac{3}{8}
  \]

  \[
  P(Y = 3) = P(\{HHH\}) = \frac{1}{8}
  \]
Example 4

- Suppose three balls are to be randomly selected \textit{without} replacement from an urn containing 20 balls numbered 1 through 20.

- \textbf{Question}: If we bet that at least one of the balls that are drawn has a number as large as or larger than 17, what is the probability that we win the bet?
Example 4

Suppose three balls are to be randomly selected \textit{without} replacement from an urn containing 20 balls numbered 1 through 20.

\textbf{Question:} If we bet that at least one of the balls that are drawn has a number as large as or larger than 17, what is the probability that we win the bet?

\textbf{Solution:} Let $X$ be the largest number selected. So, $X$ is a \textit{random variable} taking on one of the values: 3, 4, \ldots, 20.
Example 4 (cont.)

First off, how many possible selections are there, since these events are equally likely?

\[ \text{P}(X = i) = \frac{(i - 1)}{\binom{20}{3}} \]

because you need to choose 2 of the \(i - 1\) smaller numbers.
Example 4 (cont.)

- First off, how many possible selections are there, since these events are equally likely?

\[
\binom{20}{3}
\]

- So, with that in mind, what is the probability that the largest number selected is \(i\)?
Example 4 (cont.)

- First off, how many possible selections are there, since these events are equally likely?
- \( \binom{20}{3} \). So, with that in mind, what is the probability that the largest number selected is \( i \)?

\[
P(X = i) = \frac{\binom{i-1}{2}}{\binom{20}{3}}
\]

because you need to choose 2 of the \( i - 1 \) smaller numbers.
What probability are we ultimately looking for and how do we relate it to $X$?

\[
P(X \geq 17) = P(X = 17) + P(X = 18) + P(X = 19) + P(X = 20) \\ 
= (\frac{16}{20})^2 \cdot (\frac{3}{20}) + (\frac{17}{20})^2 \cdot (\frac{3}{20}) + (\frac{18}{20})^2 \cdot (\frac{3}{20}) + (\frac{19}{20})^2 \cdot (\frac{3}{20}) \\ 
\approx 0.508
\]
Example 4 (cont.)

What probability are we ultimately looking for and how do we relate it to $X$?

$$P(X \geq 17) = P(X = 17) + P(X = 18) + P(X = 19) + P(X = 20) \approx 0.508$$
Example 4 (cont.)

- What probability are we ultimately looking for and how do we relate it to $X$?

$$P(X \geq 17) = P(X = 17) + P(X = 18) + P(X = 19) + P(X = 20)$$

$$= \frac{16}{20} \cdot \frac{2}{3} + \frac{17}{20} \cdot \frac{2}{3} + \frac{18}{20} \cdot \frac{2}{3} + \frac{19}{20} \cdot \frac{2}{3}$$

$$\approx 0.508$$

- We can break up this probability because
Example 4 (cont.)

- What probability are we ultimately looking for and how do we relate it to $X$?

\[
P(X \geq 17) = P(X = 17) + P(X = 18) + P(X = 19) + P(X = 20)
\]
\[
= \frac{16}{20} + \frac{17}{20} + \frac{18}{20} + \frac{19}{20}
\]
\[
\approx 0.508
\]

- We can break up this probability because

(a) It is a *discrete random variable*. 
Example 4 (cont.)

- What probability are we ultimately looking for and how do we relate it to $X$?

$$P(X \geq 17) = P(X = 17) + P(X = 18) + P(X = 19) + P(X = 20)$$

$$= \frac{\binom{16}{2}}{\binom{20}{3}} + \frac{\binom{17}{2}}{\binom{20}{3}} + \frac{\binom{18}{2}}{\binom{20}{3}} + \frac{\binom{19}{2}}{\binom{20}{3}}$$

$$\approx 0.508$$

- We can break up this probability because
  
  (a) It is a discrete random variable.
  
  (b) The event $\{X \geq 17\}$ is the union of the disjoint events described above.
Def: If $X$ is a *discrete random variable*, then the function $p(x)$,

$$p(x) = P(X = x) = P(\text{all } s \in S : X(s) = x)$$

for each $x$ within the range of $X$, is called the *probability mass function* of $X$. 
**Cumulative Distribution Function**

- **Def**: The *cumulative distribution function*, $F(x)$, of a discrete random variable, $X$, with *probability mass function* (pmf), $p(x)$, is given by

  $$F(x) = P(X \leq x) = \sum_{y \leq x} p(y)$$

- Informally, you can think about it like this: for any $x$, $F(x)$ is the probability that the observed value of $X$ (r.v.!) will be at most $x$ (some "outcome").
Example 3 (cont.)

- Back to Example 3.
- We can represent each probability using \( pmf \) as:

\[
\begin{align*}
p(0) &= P(Y = 0) & p(1) &= P(Y = 1) \\
p(2) &= P(Y = 2) & p(3) &= P(Y = 3)
\end{align*}
\]
Example 3 (cont.)

- Back to Example 3.
- We can represent each probability using *pmf* as:
  
  \[
  p(0) = P(Y = 0) \quad p(1) = P(Y = 1) \\
  p(2) = P(Y = 2) \quad p(3) = P(Y = 3) 
  \]

- And the *cumulative distribution function* (*c.d.f.*) of \( Y \) is:
Example 3 (cont.)

- Back to Example 3.
- We can represent each probability using *pmf* as:

  \[
  p(0) = P(Y = 0) \quad p(1) = P(Y = 1) \\
  p(2) = P(Y = 2) \quad p(3) = P(Y = 3)
  \]

- And the *cumulative distribution function (c.d.f.)* of \( Y \) is:

  \[
  F(0) = P(Y \leq 0) = \frac{1}{8}
  \]
Example 3 (cont.)

- Back to Example 3.
- We can represent each probability using pmf as:

  \[ p(0) = P(Y = 0) \quad p(1) = P(Y = 1) \]
  \[ p(2) = P(Y = 2) \quad p(3) = P(Y = 3) \]

- And the cumulative distribution function (c.d.f.) of \( Y \) is:

  \[ F(0) = P(Y \leq 0) = \frac{1}{8} \]
  \[ F(1) = P(Y \leq 1) = \frac{4}{8} \]
Example 3 (cont.)

- Back to Example 3.
- We can represent each probability using *pmf* as:
  \[
  p(0) = P(Y = 0) \quad p(1) = P(Y = 1) \\
  p(2) = P(Y = 2) \quad p(3) = P(Y = 3)
  \]

- And the *cumulative distribution function* (c.d.f.) of \( Y \) is:
  \[
  F(0) = P(Y \leq 0) = \frac{1}{8} \\
  F(1) = P(Y \leq 1) = \frac{4}{8} \\
  F(2) = P(Y \leq 2) = \frac{7}{8}
  \]
Example 3 (cont.)

- Back to Example 3.
- We can represent each probability using \( pmf \) as:
  \[
  p(0) = P(Y = 0) \quad p(1) = P(Y = 1) \\
  p(2) = P(Y = 2) \quad p(3) = P(Y = 3)
  \]

- And the \textit{cumulative distribution function} (c.d.f.) of \( Y \) is:
  \[
  F(0) = P(Y \leq 0) = \frac{1}{8} \\
  F(1) = P(Y \leq 1) = \frac{4}{8} \\
  F(2) = P(Y \leq 2) = \frac{7}{8} \\
  F(3) = P(Y \leq 3) = 1
  \]
Def: If $X$ is a discrete random variable having a pmf $p(x)$, then the expectation, or the expected value, of $X$, is defined by (and denoted as)

$$E(X) = \sum_{x:p(x) > 0} xp(x)$$

$E(X)$ is the weighted average of all of the possible values that $X$ can take on.

The weight of each value is the probability that $X$ takes on that value (similar to Solution 2 in Example 2 with Polya’s urn).
Example 3 (cont.)

The *expectation* of the *r.v.* $Y$ referring to tossing three coins is:
Example 3 (cont.)

The *expectation* of the *r.v.* $Y$ referring to tossing three coins is:

$$E(Y) = 0 \cdot p(0) + 1 \cdot p(1) + 2 \cdot p(2) + 3 \cdot p(3) = \frac{3}{2}$$
Properties

- We often say that $E$, the *expected value* operator, is **linear** in the sense that
  - $E(X + c) = E(X) + c$, where $c$ is a constant and $X$ is a r.v.
  - $E(X + Y) = E(X) + E(Y)$, where $X$ and $Y$ are r.v.'s.
    - Note that this is true regardless of whether $X$ is independent of $Y$ (more on that later).
  - $E(aX) = aE(X)$, where $a$ is some constant and $X$ is a r.v.

- By combining these three facts, we can produce the general form:

$$E(aX + bY) = aE(X) + bE(Y)$$
Example 5

A roulette wheel consists of 38 spaces (0, 00, 1, 2, …, 36). One bet is called topline and it pays out $6 for every dollar bet when the balls lands on 0, 00, 1, 2, or 3. If it doesn’t land on those slots, then the better receives nothing. Another bet to make is to bet on red. If the ball lands on a red number, then the better is paid $1 ontop of their $1 cost. There are 18 red numbers out of the 38 total.

Let’s define our random variables relating to each bet, assuming we bet $1. For the topline bet, we will give it $X$: 

\[ X = \begin{cases} 7 & \text{if } s \text{ is } 0, 00, 1, 2, \text{ or } 3, \\ 0 & \text{else.} \end{cases} \]

Note that we define \( X \) in terms of total money left over after the bet. We could think of the outcomes as 6 and -1, too.
Example 5

- A roulette wheel consists of 38 spaces (0, 00, 1, 2, ..., 36). One bet is called topline and it pays out $6 for every dollar bet when the balls land on 0, 00, 1, 2, or 3. If it doesn’t land on those slots, then the better receives nothing. Another bet to make is to bet on red. If the ball lands on a red number, then the better is paid $1 on top of their $1 cost. There are 18 red numbers out of the 38 total.

- Let’s define our *random variables* relating to each bet, assuming we bet $1. For the topline bet, we will give it $X$:

$$X = \begin{cases} 
7 & \text{if } s \text{ is 0, 00, 1, 2, or 3.} \\
0 & \text{else.} 
\end{cases}$$

- Note that we define $X$ in terms of total money left over after the bet. We could think of the outcomes as 6 and -1, too.
Example 5 (cont.)

- Now, for the red bet, we will denote it $Y$: 

\[
Y = \begin{cases} 
2 & \text{if } s \text{ is red.} \\
0 & \text{else.} 
\end{cases}
\]
Example 5 (cont.)

Now, for the red bet, we will denote it $Y$:

$$Y = \begin{cases} 2 & \text{if s is red.} \\ 0 & \text{else.} \end{cases}$$

**Question:** What is $E(X)$ and $E(Y)$?
Example 5 (cont.)

Solution: The *expected* total money left after a $1 topline bet is
Example 5 (cont.)

Solution: The expected total money left after a $1 topline bet is

\[ E(X) = 0 \cdot \frac{33}{38} + 7 \cdot \frac{5}{38} = \frac{35}{38} \]

and the expected total money left after a $1 bet on red is
Example 5 (cont.)

- **Solution:** The *expected* total money left after a $1 topline bet is

  \[ E(X) = 0 \cdot \frac{33}{38} + 7 \cdot \frac{5}{38} = \frac{35}{38} \]

  and the *expected* total money left after a $1 bet on red is

  \[ E(Y) = 0 \cdot \frac{20}{38} + 2 \cdot \frac{18}{38} = \frac{36}{38} \]
Example 5 (cont.)

- **Solution**: The *expected* total money left after a $1 topline bet is

\[ E(X) = 0 \cdot \frac{33}{38} + 7 \cdot \frac{5}{38} = \frac{35}{38} \]

and the *expected* total money left after a $1 bet on red is

\[ E(Y) = 0 \cdot \frac{20}{38} + 2 \cdot \frac{18}{38} = \frac{36}{38} \]

- **Question**: What is the *expected* total money left after a $5 topline bet and a $3 bet on red?
Example 5 (cont.)

**Solution**: Use the fact that *expectation* is *linear*:

\[
E(5X + 3Y) = 5E(X) + 3E(Y)
\]

\[
= 5 \cdot \frac{35}{38} + 3 \cdot \frac{36}{38} 
\approx 7.45
\]
Example 5 (cont.)

- **Solution**: Use the fact that *expectation* is linear:

\[
E(5X + 3Y) = 5E(X) + 3E(Y)
\]

\[
= 5 \cdot \frac{35}{38} + 3 \cdot \frac{36}{38} \approx 7.45
\]
Finish reading 3.4 - 3.5 and try to get through 4.1 - 4.3.
Next lecture *starts* reviewing 4.3 and begin more with 4.4.
Start working on HW2.