TESTING FOR A DISEASE
TESTING FOR A DISEASE

• Suppose 1% of the population has a disease
TESTING FOR A DISEASE

• Suppose 1% of the population has a disease

• there is a diagnostic test, that finds it 80% of the time (assuming the subject has it)
TESTING FOR A DISEASE

• Suppose 1% of the population has a disease

• there is a diagnostic test, that finds it 80% of the time (assuming the subject has it)

• the test also produces false positives, at a rate of 9.6% (you're fine, but the test says you're not)
TESTING FOR A DISEASE

• Suppose 1% of the population has a disease
• there is a diagnostic test, that finds it 80% of the time (assuming the subject has it)
• the test also produces false positives, at a rate of 9.6% (you're fine, but the test says you're not)

If someone tests positive, what are the odds that they have the disease?
<table>
<thead>
<tr>
<th></th>
<th>Have disease</th>
<th>Don't have disease</th>
</tr>
</thead>
<tbody>
<tr>
<td>Test α</td>
<td>80%</td>
<td>9.6%</td>
</tr>
<tr>
<td>Test β</td>
<td>20%</td>
<td>90.4%</td>
</tr>
<tr>
<td></td>
<td>1% Have disease</td>
<td>99% Don't have disease</td>
</tr>
<tr>
<td>----------------</td>
<td>------------------</td>
<td>------------------------</td>
</tr>
<tr>
<td>Test &quot;ø&quot;</td>
<td>80%</td>
<td>9.6%</td>
</tr>
<tr>
<td>Test &quot;Ω&quot;</td>
<td>20%</td>
<td>90.4%</td>
</tr>
<tr>
<td></td>
<td>1% Have Disease</td>
<td>99% Don't Have Disease</td>
</tr>
<tr>
<td>----------------</td>
<td>-----------------</td>
<td>------------------------</td>
</tr>
<tr>
<td>Test ✗</td>
<td>80%</td>
<td>9.6%</td>
</tr>
<tr>
<td>Test ⚪</td>
<td>20%</td>
<td>90.4%</td>
</tr>
</tbody>
</table>

4 events

\[
P(\text{disease} | \text{test ✗}) = \ ?
\]
### 4 events

<table>
<thead>
<tr>
<th></th>
<th>1% Have disease</th>
<th>99% Don't have disease</th>
</tr>
</thead>
<tbody>
<tr>
<td>Test ☐</td>
<td>80%</td>
<td>9.6%</td>
</tr>
<tr>
<td>Test ☐</td>
<td>20%</td>
<td>90.4%</td>
</tr>
</tbody>
</table>

\[
P(\text{disease} | \text{test } \checkmark) = \frac{P(\text{disease } \cap \text{ test } \checkmark)}{P(\text{test } \checkmark)}
\]
There are 4 events:

<table>
<thead>
<tr>
<th>Have disease</th>
<th>Don't have disease</th>
</tr>
</thead>
<tbody>
<tr>
<td>Test ☐️</td>
<td>1% 80%</td>
</tr>
<tr>
<td>Test ☑️</td>
<td>99% 9.6%</td>
</tr>
</tbody>
</table>

\[
P(\text{disease|test ☐}) = \frac{P(\text{disease } \cap \text{ test ☐})}{P(\text{test ☐})} = \frac{P(\text{test ☐}|\text{disease}) \cdot P(\text{disease})}{P(\text{test ☐})}
\]
<table>
<thead>
<tr>
<th>Event</th>
<th>Disease</th>
<th>No Disease</th>
</tr>
</thead>
<tbody>
<tr>
<td>Test 0</td>
<td>1%</td>
<td>99%</td>
</tr>
<tr>
<td>4 events</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Test 1</td>
<td>80%</td>
<td>9.6%</td>
</tr>
<tr>
<td>Test 0</td>
<td>20%</td>
<td>90.4%</td>
</tr>
</tbody>
</table>

\[
P(\text{disease} | \text{test 1}) = \frac{P(\text{disease} \land \text{test 1})}{P(\text{test 1})} = \frac{P(\text{test 1} | \text{disease}) \cdot P(\text{disease})}{P(\text{test 1})}
\]

\[
P(\text{test 1} | \text{disease}) \cdot P(\text{disease}) = ?
\]
4 events

<table>
<thead>
<tr>
<th></th>
<th>1% Have disease</th>
<th>99% Don't have disease</th>
</tr>
</thead>
<tbody>
<tr>
<td>Test ✗</td>
<td>80%</td>
<td>9.6%</td>
</tr>
<tr>
<td>Test ✚</td>
<td>20%</td>
<td>90.4%</td>
</tr>
</tbody>
</table>

\[
P(\text{disease} | \text{test ✗}) = \frac{P(\text{disease} \cap \text{test ✗})}{P(\text{test ✗})} = \frac{P(\text{test ✗ | disease}) \cdot P(\text{disease})}{P(\text{test ✗})} \\
P(\text{test ✗ | disease}) \cdot P(\text{disease}) = 0.8 \cdot 0.01 = 0.008 \text{ (0.8%)}
\]

\[P(\text{test ✗}) = ?\]
<table>
<thead>
<tr>
<th>Event</th>
<th>Disease (%)</th>
<th>No Disease (%)</th>
</tr>
</thead>
<tbody>
<tr>
<td>Test (\boxed{\text{a}})</td>
<td>80%</td>
<td>9.6%</td>
</tr>
<tr>
<td>Test (\boxed{\text{b}})</td>
<td>20%</td>
<td>90.4%</td>
</tr>
</tbody>
</table>

\[
P(\text{disease}|\text{test } \boxed{\text{a}}) = \frac{P(\text{disease} \cap \text{test } \boxed{\text{a}})}{P(\text{test } \boxed{\text{a}})} = \frac{P(\text{test } \boxed{\text{a}}|\text{disease}) \cdot P(\text{disease})}{P(\text{test } \boxed{\text{a}})}
\]

\[
P(\text{test } \boxed{\text{b}}) = \left\{ \begin{array}{l}
P(\text{test } \boxed{\text{b}}|\text{disease}) \cdot P(\text{disease}) = 0.8 \cdot 0.01 = 0.008 (0.8\%)
\end{array} \right.
\]
<table>
<thead>
<tr>
<th></th>
<th>1% Have disease</th>
<th>99% Don't have disease</th>
</tr>
</thead>
<tbody>
<tr>
<td>Test ☐</td>
<td>80%</td>
<td>9.6%</td>
</tr>
<tr>
<td>Test ☐</td>
<td>20%</td>
<td>90.4%</td>
</tr>
</tbody>
</table>

\[
P(\text{disease}|\text{test ☐}) = \frac{P(\text{disease} \cap \text{test ☐})}{P(\text{test ☐})} = \frac{P(\text{test ☐}|\text{disease}) \cdot P(\text{disease})}{P(\text{test ☐})}
\]

\[
P(\text{test ☐}) = \left\{ \begin{array}{l}
P(\text{test ☐}|\text{disease}) \cdot P(\text{disease}) = 0.8 \cdot 0.01 = 0.008 (0.8\%)
\end{array} \right.
\]

\[
+ P(\text{test ☐}|\text{no disease}) \cdot P(\text{no disease})
\]
4 events

<table>
<thead>
<tr>
<th></th>
<th>1% Have disease</th>
<th>99% Don't have disease</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>Test ✗</strong></td>
<td>80%</td>
<td>9.6%</td>
</tr>
<tr>
<td><strong>Test ☑</strong></td>
<td>20%</td>
<td>90.4%</td>
</tr>
</tbody>
</table>

\[
P(\text{disease} | \text{test ✗}) = \frac{P(\text{disease} \cap \text{test ✗})}{P(\text{test ✗})} = \frac{P(\text{test ✗} | \text{disease}) \cdot P(\text{disease})}{P(\text{test ✗})}
\]

\[
P(\text{test ✗}) = P(\text{test ✗} | \text{disease}) \cdot P(\text{disease}) + P(\text{test ✗} | \text{no disease}) \cdot P(\text{no disease}) = 0.8 \cdot 0.01 = 0.008 (0.8\%) + 0.096 \cdot 0.99 \approx 0.095 (9.5\%)
\]
\[
P(\text{disease} | \text{test } \checkmark) = \frac{P(\text{disease} \land \text{test } \checkmark)}{P(\text{test } \checkmark)} = \frac{P(\text{test } \checkmark | \text{disease}) \cdot P(\text{disease})}{P(\text{test } \checkmark)}
\]

\[
P(\text{test } \checkmark) = \left\{ \begin{array}{l}
P(\text{test } \checkmark | \text{disease}) \cdot P(\text{disease}) = 0.8 \cdot 0.01 = 0.008 \ (0.8\%)
\end{array} \right. + \left\{ \begin{array}{l}
P(\text{test } \checkmark | \text{no disease}) \cdot P(\text{no disease}) = 0.096 \cdot 0.99 = 0.095 \ (9.5\%)
\end{array} \right.
\]

\[
P(\text{disease} | \text{test } \checkmark) = \frac{0.008}{0.008 + 0.095} \approx 7.8\%
\]
Bayes theorem
\[ P(A \cap B) = P(A|B) \cdot P(B) = P(B|A) \cdot P(A) \]
\[ P(A \cap B) = P(A|B) \cdot P(B) = P(B|A) \cdot P(A) \]

\[ P(A|B) = P(B|A) \cdot \frac{P(A)}{P(B)} \]
Bayes' theorem

\[ P(A \cap B) = P(A|B) \cdot P(B) = P(B|A) \cdot P(A) \]

\[ P(A|B) = \frac{P(B|A) \cdot P(A)}{P(B)} \]

\[ P(A|B) = \frac{P(B|A) \cdot P(A)}{P(B|A) \cdot P(A) + P(B|\bar{A}) \cdot P(\bar{A})} \]
\[ P(A|B) = \frac{P(B|A) \cdot P(A)}{P(B|A) \cdot P(A) + P(B|\overline{A}) \cdot P(\overline{A})} \]
\[ P(A|B) = \frac{P(B|A) \cdot P(A)}{P(B|A) \cdot P(A) + P(B|\overline{A}) \cdot P(\overline{A})} \]

A: have disease
B: test ☑
\[ P(A|B) = \frac{P(B|A) \cdot P(A)}{P(B|A) \cdot P(A) + P(B|\bar{A}) \cdot P(\bar{A})} \]

A: have disease
B: test

<table>
<thead>
<tr>
<th></th>
<th>( A )</th>
<th>( \bar{A} )</th>
</tr>
</thead>
<tbody>
<tr>
<td>( B ) \mid ( A )</td>
<td>80%</td>
<td>9.6%</td>
</tr>
<tr>
<td>( B ) \mid ( \bar{A} )</td>
<td>1%</td>
<td>99%</td>
</tr>
</tbody>
</table>
\[ P(A|B) = \frac{P(B|A) \cdot P(A)}{P(B|A) \cdot P(A) + P(B|\bar{A}) \cdot P(\bar{A})} \]

- **A**: have disease
- **B**: test

<table>
<thead>
<tr>
<th></th>
<th>1% A</th>
<th>99% ( \bar{A} )</th>
</tr>
</thead>
<tbody>
<tr>
<td>( B</td>
<td>A )</td>
<td>80%</td>
</tr>
</tbody>
</table>

- \( P(A) = 0.01 \)
- \( P(\bar{A}) = 0.99 \)
- \( P(B|A) = 0.8 \)
- \( P(B|\bar{A}) = 0.096 \)
\[ P(A|B) = \frac{P(B|A) \cdot P(A)}{P(B|A) \cdot P(A) + P(B|\bar{A}) \cdot P(\bar{A})} \]

\( A: \) have disease
\( B: \) test

\begin{tabular}{c|c|c}
1\%\( A \) & 99\%\( \bar{A} \) \\
B\( |A\) = 80\% & B\( |\bar{A}\) = 9.6\% \\
\end{tabular}

\[
P(A) = 0.01 \\
P(\bar{A}) = 0.99 \\
P(B|A) = 0.8 \\
P(B|\bar{A}) = 0.096
\]

\[ P(A|B) = \frac{0.8 \cdot 0.01}{0.8 \cdot 0.01 + 0.096 \cdot 0.99} \]
RANDOM VARIABLES
RANDOM VARIABLES

A quote from your textbook:

“A random variable is neither random nor variable”
We have been using random variables, implicitly.

ex: roll 2 dice, examine probability that sum = k, or = even.
We have been using random variables, implicitly.

ex: roll 2 dice, examine probability that \( \text{sum} = k \), or \( k \) even.

\[ \text{Define random variable } X : \text{sum of two dice rolls.} \]

So, \( X[(1,2)] = 3 \)
\[ X[(5,5)] = 10 \]
We have been using random variables, implicitly.

ex: roll 2 dice, examine probability that \( \text{sum} = k \), or \( \text{even} \).

\[ \rightarrow \text{define random variable } X : \text{sum of two dice rolls}. \]

So, \[ X[(1,2)] = 3 \]
\[ X[(5,5)] = 10 \]

\[ \rightarrow \text{define random variable } Y : \text{parity of two dice rolls}. \]
We have been using random variables, implicitly.

ex: roll 2 dice, examine probability that sum = k, or = even.

\[ \text{define random variable } X : \text{ sum of two dice rolls.} \]

\[ X[(1,2)] = 3 \]
\[ X[(5,5)] = 10 \]

\[ \text{define random variable } Y : \text{ parity of two dice rolls.} \]

\[ Y[(1,2)] = 1 \]
\[ Y[(5,5)] = 0 \]
We have been using random variables, implicitly.

ex: roll 2 dice, examine probability that sum = k, or = even.

→ define random variable $X$ : sum of two dice rolls.

So, $X[(1,2)] = 3$
$X[(5,5)] = 10$

→ define random variable $Y$ : parity of two dice rolls.

So, $Y[(1,2)] = 1$
$Y[(5,5)] = 0$ or $13 = 2$ (arbitrary)
Think of a r.v. as a function, mapping sample space to whatever you like (usually a number).
Think of a r.v. as a function, mapping sample space to whatever you like (usually a number).

Then we can express questions neatly:

\[ P(X < 3) = \frac{1}{36} \]
\[ P(Y = 1) = \frac{1}{2} \]
Think of a r.v. as a function, mapping sample space to whatever you like, usually a number.

Then we can express questions neatly: 

\[ P(X < 3) = \frac{1}{36} \]
\[ P(Y = 1) = \frac{1}{2} \]

We can also eliminate absurd events, e.g., \( P(X = 13) = 0 \).
Think of a r.v. as a function, mapping sample space to whatever you like, usually a number.

Then we can express questions neatly:

\[
P(X < 3) = \frac{1}{36}
\]

\[
P(Y = 1) = \frac{1}{2}
\]

We can also eliminate absurd events, e.g., \( P(X = 13) = 0 \).

Section 33 discusses in more detail; not required for our class.
EXPECTATION : the very basics
EXPECTATION: the very basics

As mentioned, a r.v. $X$ can have several values. It is based on outcomes that result from a random process.
EXPECTATION: the very basics

As mentioned, a r.v. $X$ can have several values. It is based on outcomes that result from a random process. So we don't know what value it will have. But we can expect it to have some value.
EXPECTATION: the very basics

As mentioned, a r.v. $X$ can have several values. It is based on outcomes that result from a random process. So we don't know what value it will have. But we can expect it to have some value

\[
\text{Expected value} = \text{weighted average}
\]

(e.g. average parity for unfair coin = $\frac{1}{2}$; weighted average $\neq \frac{1}{2}$)
Expected value = weighted average

\[ E(X) = \sum y \cdot P(X = y) \]

*over all possible values y, compatible with X.*
Expected value = weighted average

\[ E(X) = \sum y \cdot P(X=y) \]

*over all possible values of \( y \), compatible with \( X \).

(however we only care about finitely many)
Expected value = weighted average

\[ E(X) = \sum y \cdot P(X = y) \]

* over all possible values \( y \), compatible with \( X \).

(However we only care about finitely many)

\[ E(X) = \sum_{s \in S} \left[ X(s) \cdot P(s) \right] \]

over all samples that define \( X \).
Expected value = weighted average

$$E(X) = \sum y \cdot P(X = y)$$

* over all possible values $y$, compatible with $X$.
  (however we only care about finitely many)

$$E(X) = \sum_{s \in S} [X(s) \cdot P(s)]$$

\[\{\text{over all samples that define } X.}\] \(\{\text{a finite number}\)
Expected value = weighted average

\[ E(X) = \sum y \cdot P(X=y) \]

*over all possible values y, compatible with X. (however we only care about finitely many)

\[ E(X) = \sum_{s \in S} [X(s) \cdot P(s)] \]

over all samples that define X. \{a finite number\}

\[ E(X) = \sum_{i=1}^{6} X(i) \cdot P(i) \]

e.g. roll 1 die. X = number observed.
Expected value = weighted average

\[ E(X) = \sum_y y \cdot P(X=y) \]

*over all possible values \( y \), compatible with \( X \). (however we only care about finitely many)

\[ E(X) = \sum_{s \in S} [X(s) \cdot P(s)] \]

over all samples that define \( X \).

\{ \text{a finite number} \}

e.g. roll 1 die. \( X = \) number observed.

\[ E(X) = \sum_{i=1}^{6} X(i) \cdot P(i) \]

\[ = 1 \cdot \frac{1}{6} + 2 \cdot \frac{1}{6} + 3 \cdot \frac{1}{6} + 4 \cdot \frac{1}{6} + 5 \cdot \frac{1}{6} + 6 \cdot \frac{1}{6} \]
Expected value = weighted average

\[ E(X) = \sum y \cdot P(X = y) \]

* over all possible values y, compatible with X.
* (however we only care about finitely many)

\[ E(X) = \sum_{s \in S} [X(s) \cdot P(s)] \]

over all samples that define X.

\[ \sum_{i=1}^{6} X(i) \cdot P(i) \]

e.g. roll 1 die. X = number observed.

\[ = 1 \cdot \frac{1}{6} + 2 \cdot \frac{1}{6} + 3 \cdot \frac{1}{6} + 4 \cdot \frac{1}{6} + 5 \cdot \frac{1}{6} + 6 \cdot \frac{1}{6} \]

\[ = \frac{1}{6} \cdot (1+2+3+4+5+6) = 3.5 \]
Expected value = weighted average

\[ E(X) = \sum y \cdot P(X=y) \]

*over all possible values of \( y \), compatible with \( X \).

(However we only care about \textbf{finitely many})

\[ E(X) = \sum_{s \in S} [X(s) \cdot P(s)] \]

\textit{over all samples that define \( X \).} \quad \textit{a finite number}

\[ E(X) = \sum_{i=1}^{6} X(i) \cdot P(i) \]

\[ = 1 \cdot \frac{1}{6} + 2 \cdot \frac{1}{6} + 3 \cdot \frac{1}{6} + 4 \cdot \frac{1}{6} + 5 \cdot \frac{1}{6} + 6 \cdot \frac{1}{6} \]

\[ = \frac{1}{6} \cdot (1+2+3+4+5+6) = 3.5 \]

\textit{e.g. roll 1 die. \( X = \) number observed.}
$E(X) = \sum y \cdot P(X = y)$

Typically use this when you have a good way of aggregating outcomes.
\[ E(X) = \sum y \cdot P(X = y) \]

typically use this when you have a good way of aggregating outcomes

example: roll 2 dice. \( X = |\text{difference between the 2}| \)
$E(X) = \sum y \cdot P(X=y)$

Typically use this when you have a good way of aggregating outcomes.

Example: roll 2 dice. $X = |\text{difference between the 2}|$

Possible values of $X \rightarrow 0 \ 1 \ 2 \ 3 \ 4 \ 5$
$E(X) = \sum y \cdot P(X=y)$

typically use this when you have a good way of aggregating outcomes

eexample: roll 2 dice. $X = |\text{difference between the 2}|$

possible values of $X \rightarrow 0 \ 1 \ 2 \ 3 \ 4 \ 5$

# outcomes supporting value $\rightarrow 6 \ ?$
\[ E(X) = \sum y \cdot P(X=y) \] typically use this when you have a good way of aggregating outcomes.

e.g. roll 2 dice. \( X = |\text{difference between the 2}| \)

possible values of \( X \) \( \rightarrow 0 \quad 1 \quad 2 \quad 3 \quad 4 \quad 5 \)

\# outcomes supporting value \( \rightarrow 6 \quad 5 \cdot 2 \quad ? \)
\[ E(X) = \sum y \cdot P(X = y) \] typically use this when you have a good way of aggregating outcomes

example: roll 2 dice. \( X = |\text{difference between the 2}| \)

possible values of \( X \) → 0 1 2 3 4 5

# outcomes supporting value → 6 5 2 4 2 ?
\[ E(X) = \sum y \cdot P(X=y) \] typically use this when you have a good way of aggregating outcomes

eexample: roll 2 dice. \( X = |\text{difference between the 2}| \)

possible values of \( X \) \( \rightarrow \) 0 1 2 3 4 5

# outcomes supporting value \( \rightarrow \) 6 5 2 4 2 3 2 2 2 1 2
\[ E(X) = \sum y \cdot P(X = y) \] typically use this when you have a good way of aggregating outcomes

example: roll 2 dice. \( X = | \text{difference between the 2} | \)

possible values of \( X \) \( \rightarrow 0 \ 1 \ 2 \ 3 \ 4 \ 5 \)

# outcomes supporting value \( \rightarrow 6 \ \ 5 \ .2 \ \ 4 \ .2 \ \ 3 \ .2 \ \ 2 \ .2 \ \ 1 \ .2 \)

(for probability, divide by 36)
\[ E(X) = \sum y \cdot P(X=y) \]

Typically use this when you have a good way of aggregating outcomes.

Example: roll 2 dice. \( X = | \text{difference between the 2} | \)

Possible values of \( X \) → 0 1 2 3 4 5

# outcomes supporting value → 6 5 2 4 2 3 2 2 2 1 2

(for probability, divide by 36)

\[ E(x) = \ ? \]
\[ E(X) = \sum y \cdot P(X=y) \]  

\[ \text{typically use this when you have a good way of aggregating outcomes} \]

\text{example: roll 2 dice. } \quad X = |\text{difference between the 2}| 

possible values of \( X \) → 0 1 2 3 4 5  

# outcomes supporting value → 6 5 2 4 2 3 2 2 2 1 2  

(for probability, divide by 36)  

\[ E(x) = \frac{0 + 10 + 16 + 18 + 16 + 10}{36} \]  

\( \approx 1.944 \)
EXPECTATION : PROPERTIES
**Expectation: Properties**

\[ E(X + Y) = E(X) + E(Y) \]
$c_1, c_2 \in \mathbb{R}$

$$E(c_1X + c_2Y) = c_1 \cdot E(X) + c_2 \cdot E(Y)$$
LINEARITY OF EXPECTATION (important)

\[ c_1, c_2 \in \mathbb{R} \quad E(c_1X + c_2Y) = c_1 \cdot E(X) + c_2 \cdot E(Y) \]
LINEARITY OF EXPECTATION (important)

\[ c_1, c_2 \in \mathbb{R} \quad E(c_1X + c_2Y) = c_1 \cdot E(X) + c_2 \cdot E(Y) \]

Generally,

\[ E(c_1X_1 + c_2X_2 + \cdots + c_nX_n) = c_1E(X_1) + c_2E(X_2) + \cdots + c_nE(X_n) \]
LINEARITY OF EXPECTATION (important)

\[ c_1, c_2 \in \mathbb{R} \quad E(c_1X + c_2Y) = c_1 \cdot E(X) + c_2 \cdot E(Y) \]

Generally, \[ E(c_1X_1 + c_2X_2 + \cdots + c_nX_n) = c_1E(X_1) + c_2E(X_2) + \cdots + c_nE(X_n) \]

\[ E(\sum c_iX_i) = \sum c_iE(X_i) \quad \rightarrow \text{Does NOT assume independence} \]
**EXPECTATION : PROPERTIES**

**LINEARITY OF EXPECTATION**  (**important**)  

\[ c_1, c_2 \in \mathbb{R} \quad E(c_1 X + c_2 Y) = c_1 \cdot E(X) + c_2 \cdot E(Y) \]

Generally, \( E(c_1 X_1 + c_2 X_2 + \cdots + c_n X_n) = c_1 E(X_1) + c_2 E(X_2) + \cdots + c_n E(X_n) \)

\[ E(\sum c_i X_i) = \sum c_i E(X_i) \quad \text{→ Does NOT assume independence} \]

Independence: \( P(X=a \ & \ Y=b) = P(X=a) \cdot P(Y=b) \)

for all \( a, b \) ...

(Def. 3.3.6)
LINEARITY OF EXPECTATION (important)

\[ E(c_1X + c_2Y) = c_1 \cdot E(X) + c_2 \cdot E(Y) \]

Generally, 

\[ E(c_1X_1 + c_2X_2 + \cdots + c_nX_n) = c_1E(X_1) + c_2E(X_2) + \cdots + c_nE(X_n) \]

\[ E(\sum c_iX_i) = \sum c_iE(X_i) \quad \rightarrow \text{Does NOT assume independence} \]

Independence: 

\[ P(X = a \& Y = b) = P(X = a) \cdot P(Y = b) \]

for all \( a, b \) ...

(Def. 33.6)

2 dice, A, B. \( X = \) result of A. \( Y = \) result of B. \( Z = X + Y \)

\[ E(Z) = E(X + Y) = E(X) + E(Y) = 2 \cdot 3 \cdot 5 = 7 \]
**EXPECTATION : PROPERTIES**

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

Linearity of expectation doesn't assume independence
**Expectation: Properties**

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

Linearity of expectation doesn't assume independence.

But \( E(X \cdot Y) \neq E(X) \cdot E(Y) \) in general.
**EXPECTATION : PROPERTIES**

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

Linearity of expectation doesn't assume independence

but \[ E(X \cdot Y) \neq E(X) \cdot E(Y) \] in general.

If \( X \) & \( Y \) are independent, then \( E(X \cdot Y) = E(X) \cdot E(Y) \)
EXPECTATION: PROPERTIES

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

Linearity of expectation doesn't assume independence

but \[ E(X \cdot Y) \neq E(X) \cdot E(Y) \] in general.

If \( X \) & \( Y \) are independent, then \( E(X \cdot Y) = E(X) \cdot E(Y) \)

However, \( E(X \cdot Y) = E(X) \cdot E(Y) \) does NOT imply \( X \) & \( Y \) are independent.

(see example 34.15)
— We are skipping the proofs of most statements in this section.

— You are not required to study these, but it would probably be beneficial.

— We are also skipping variance, which is an important concept to learn independently.
INDICATOR RANDOM VARIABLES

(taking value 0 or 1)
INDICATOR RANDOM VARIABLES

(taking value 0 or 1)

We already saw this: \( Y \): parity of rolling one die.
INDICATOR RANDOM VARIABLES
(taking value 0 or 1)

We already saw this: $Y$: parity of rolling one die.

Another example: flip a coin 10 times.

$X = \# \text{times we see pattern HT}$
INDICATOR RANDOM VARIABLES
(taking value 0 or 1)

We already saw this: \( Y \) : parity of rolling one die.

Another example: flip a coin 10 times.
\[ X = \text{#times we see pattern HT} \]

\[ E(X) = ? \]

Try solving this "directly"
INDICATOR RANDOM VARIABLES

Flip a coin 10 times. \( X = \) #times we see pattern HT
INDICATOR RANDOM VARIABLES

flip a coin 10 times. $X$ = # times we see pattern HT

HT could appear at flips 1 & 2, or 2 & 3, ..., or 9 & 10
flip a coin 10 times. \( X = \# \text{times we see pattern HT} \)

HT could appear at flips 1\&2, or 2\&3, \ldots, or 9 \& 10.

Define r.v. \( X_i = ? \)
flip a coin 10 times. \( X = \# \text{times we see pattern HT} \)

HT could appear at flips 1 & 2, or 2 & 3, ..., or 9 & 10

Define r.v. \( X_i = \begin{cases} 1 & \text{if flips } i \text{ & } i+1 \text{ produce HT} \\ 0 & \text{otherwise} \end{cases} \)
flip a coin 10 times. \[ X = \# \text{times we see pattern HT} \]

HT could appear at flips 1&2, or 2&3, \ldots, or 9&10

Define r.v. \[ X_i = \begin{cases}1 & \text{if flips } i \text{ & } i+1 \text{ produce HT} \\0 & \text{otherwise} \end{cases} \]

\[ X = X_1 + X_2 + \cdots + X_9 \]
flip a coin 10 times. \( X = \# \text{times we see pattern HT} \)

HT could appear at flips 1 & 2, or 2 & 3, \ldots, or 9 & 10

Define r.v. \( X_i = \begin{cases} 1 & \text{if flips } i & \& i+1 \text{ produce HT} \\ 0 & \text{otherwise} \end{cases} \)

\[ X = X_1 + X_2 + \cdots + X_9 \]

Notice \( X_1 & X_2 \) are not independent. \( P(X_i = 1) = \frac{1}{4} \)

\( P(X_1 \land X_2) = 0 \)
flip a coin 10 times. \( X = \# \) times we see pattern HT

HT could appear at flips 1 & 2, or 2 & 3, \ldots , or 9 & 10

Define r.v. \( X_i = \begin{cases} 1 & \text{if flips } i \text{ & } i+1 \text{ produce HT} \\ 0 & \text{otherwise} \end{cases} \)

\[
X = X_1 + X_2 + \cdots + X_9
\]

\[
E(X) = E(X_1 + X_2 + \cdots + X_9)
\]

Notice \( X_1 \) & \( X_2 \) are not independent. \( P(X_i=1) = \frac{1}{4} \)

\( P(X_1 \land X_2) = 0 \)
flip a coin 10 times. \( X \) = \# times we see pattern HT

HT could appear at flips 1\&2, or 2\&3, \ldots, or 9 \& 10

Define r.v. \( X_i = \begin{cases} 1 & \text{if flips } i \& i+1 \text{ produce HT} \\ 0 & \text{otherwise} \end{cases} \)

\[
X = X_1 + X_2 + \cdots + X_9
\]

\[
E(X) = E(X_1 + X_2 + \cdots + X_9)
\]

Notice \( X_1 \& X_2 \) are not independent.

\[
P(X_i=1) = \frac{1}{4}
\]

\[
P(X_1 \& X_2) = 0
\]

\text{linearity of expectation}
INDICATOR RANDOM VARIABLES

Flip a coin 10 times. \( X = \) \#times we see pattern HT

HT could appear at flips 1&2, or 2&3, \ldots, or 9&10

Define r.v. \( X_i = \begin{cases} 
1 & \text{if flips } i \text{ & } i+1 \text{ produce } HT \\
0 & \text{otherwise}
\end{cases} \)

\( X = X_1 + X_2 + \cdots + X_9 \)

\[
E(X) = E(X_1 + X_2 + \cdots + X_9) \\
= E(X_1) + E(X_2) + \cdots + E(X_9)
\]

Notice \( X_1 \) & \( X_2 \) are not independent. \( P(X_1=1) = \frac{1}{4} \)

\( P(X_1 \land X_2) = 0 \)

linearity of expectation
flip a coin 10 times. \( X = \# \) times we see pattern HT
HT could appear at flips 1 & 2, or 2 & 3,..., or 9 & 10

Define r.v. \( X_i = \begin{cases} 1 & \text{if flips } i \& i+1 \text{ produce HT} \\ 0 & \text{otherwise} \end{cases} \)

\[ X = X_1 + X_2 + \cdots + X_9 \]

\[ E(X) = E(X_1 + X_2 + \cdots + X_9) \]
\[ = E(X_1) + E(X_2) + \cdots E(X_9) \]

Notice \( X_1 \& X_2 \) are not independent. \( P(X_1 = 1) = \frac{1}{4} \)
\( P(X_1 \& X_2) = 0 \)

linearity of expectation

\[ E(X_i) = ? \]
Flip a coin 10 times. \( X = \# \text{times we see pattern HT} \)

HT could appear at flips 1&2, or 2&3, \ldots, or 9 & 10

Define r.v. \( X_i = \begin{cases} 1 & \text{if flips } i \& i+1 \text{ produce HT} \\ 0 & \text{otherwise} \end{cases} \)

\[
X = X_1 + X_2 + \cdots + X_9
\]

\[
E(X) = E(X_1 + X_2 + \cdots + X_9) = E(X_1) + E(X_2) + \cdots + E(X_9)
\]

Notice \( X_1 \) & \( X_2 \) are not independent. \( P(X_i=1) = \frac{1}{4} \)

\[ P(X_1 \land X_2) = 0 \]

Linearity of expectation

\[
E(X_i) = 0 \cdot P(X_i=0) + 1 \cdot P(X_i=1)
\]
Indicators of Random Variables

Flip a coin 10 times. \( X = \) \# times we see pattern HT

HT could appear at flips 1 & 2, or 2 & 3, \ldots, or 9 & 10

Define r.v. \( X_i = \begin{cases} 1 & \text{if flips } i \& i+1 \text{ produce HT} \\ 0 & \text{otherwise} \end{cases} \)

\[
X = X_1 + X_2 + \cdots + X_9
\]

\[
E(X) = E(X_1 + X_2 + \cdots + X_9) = E(X_1) + E(X_2) + \cdots E(X_9)
\]

Notice \( X_1 \& X_2 \) are not independent. \( P(X_i = 1) = \frac{1}{4} \)

\[
P(X_1 \& X_2) = 0
\]

Linearity of expectation

\[
E(X_i) = 0 \cdot P(X_i = 0) + 1 \cdot P(X_i = 1)
\]
INDICATOR RANDOM VARIABLES

Flip a coin 10 times. \( X = \text{# times we see pattern HT} \)

HT could appear at flips 1 & 2, or 2 & 3, …, or 9 & 10

Define r.v. \( X_i = \begin{cases} 1 \text{ if flips } i \text{ & } i+1 \text{ produce HT} \\ 0 \text{ otherwise} \end{cases} \)

\[
X = X_1 + X_2 + \cdots + X_9
\]

\[
E(X) = E(X_1 + X_2 + \cdots + X_9)
= E(X_1) + E(X_2) + \cdots + E(X_9)
\]

\[
E(X_i) = 0 \cdot P(X_i = 0) + 1 \cdot P(X_i = 1) = \frac{1}{4}
\]

Notice \( X_1 \) & \( X_2 \) are not independent. \( P(X_1 = 1) = \frac{1}{4} \)

\[P(X_1 \land X_2) = 0\]

(linearity of expectation)

\[E(X) = 9 \cdot \frac{1}{4}\]

(p. 240)