2.18 Independent Random Variables

Intro / Definition
Consequences of Independence
Covariance and Correlation
Anti-University of Georgia Example
Theorems Involving Covariance
Random Samples
Intro / Definition

Recall that two events are independent if \( \Pr(A \cap B) = \Pr(A) \Pr(B) \).

Then

\[
\Pr(A|B) = \frac{\Pr(A \cap B)}{\Pr(B)} = \frac{\Pr(A) \Pr(B)}{\Pr(B)} = \Pr(A).
\]

And similarly, \( \Pr(B|A) = \Pr(B) \).

Now want to define independence for RV’s, i.e., the outcome of \( X \) doesn’t influence the outcome of \( Y \).
Definition: $X$ and $Y$ are independent RV’s if, for all $x$ and $y$,

$$f(x, y) = f_X(x)f_Y(y).$$

Equivalent definitions:

$$F(x, y) = F_X(x)F_Y(y), \ \forall x, y$$

or

$$\Pr(X \leq x, Y \leq y) = \Pr(X \leq x)\Pr(Y \leq y), \ \forall x, y$$

If $X$ and $Y$ aren’t indep, then they’re dependent.
Theorem: If $X$ and $Y$ are indep, then $f(y|x) = f_Y(y)$.

Proof:

$$f(y|x) = \frac{f(x,y)}{f_X(x)} = \frac{f_X(x)f_Y(y)}{f_X(x)} = f_Y(y).$$

Similarly, $X$ and $Y$ indep implies $f(x|y) = f_X(x)$. 
Example (discrete): \( f(x, y) = \Pr(X = x, Y = y) \).

\[
\begin{array}{c|cc|c}
Y = 2 & X = 1 & X = 2 & f_Y(y) \\
0.12 & 0.28 & 0.4 \\
Y = 3 & 0.18 & 0.42 & 0.6 \\
f_X(x) & 0.3 & 0.7 & 1
\end{array}
\]

\( X \) and \( Y \) are indep since \( f(x, y) = f_X(x)f_Y(y), \forall x, y. \)
Example (cts): Suppose $f(x, y) = 6xy^2$, $0 \leq x \leq 1$, $0 \leq y \leq 1$.

After some work (which can be avoided by the next theorem), we can derive

$$f_X(x) = 2x, \text{ if } 0 \leq x \leq 1, \text{ and}$$

$$f_Y(y) = 3y^2, \text{ if } 0 \leq y \leq 1.$$ 

$X$ and $Y$ are indep since $f(x, y) = f_X(x)f_Y(y)$, $\forall x, y$. 
Easy way to tell if $X$ and $Y$ are indep... 

Theorem: $X$ and $Y$ are indep iff $f(x, y) = a(x)b(y)$, $\forall x, y$, for some functions $a(x)$ and $b(y)$ (not necessarily pdf’s).

So if $f(x, y)$ factors into separate functions of $x$ and $y$, then $X$ and $Y$ are indep.
Example: \( f(x, y) = 6xy^2, \, 0 \leq x \leq 1, \, 0 \leq y \leq 1. \) Take 
\[ a(x) = 6x, \, 0 \leq x \leq 1, \text{ and } b(y) = y^2, \, 0 \leq y \leq 1. \]
Thus, \( X \) and \( Y \) are indep (as above).

Example: \( f(x, y) = \frac{21}{4}x^2y, \, x^2 \leq y \leq 1. \) “Funny” (non-rectangular) limits make factoring into marginals impossible. Thus, \( X \) and \( Y \) are not indep.
Example: \( f(x, y) = \frac{c}{x+y}, \ 1 \leq x \leq 2, \ 1 \leq y \leq 3. \)

Can’t factor \( f(x, y) \) into fn’s of \( x \) and \( y \) separately. Thus, \( X \) and \( Y \) are not indep.

Now that we can figure out if \( X \) and \( Y \) are indep, what can we do with that knowledge?
2.18 Independent RV's

Consequences of Independence

Definition/Theorem (another Unconscious Statistician): Let \( h(X, Y) \) be a fn of the RV's \( X \) and \( Y \). Then

\[
\mathbb{E}[h(X, Y)] = \begin{cases} 
\sum_x \sum_y h(x, y) f(x, y) & \text{discrete} \\
\int_{\mathbb{R}} \int_{\mathbb{R}} h(x, y) f(x, y) \, dx \, dy & \text{continuous}
\end{cases}
\]

Theorem: Whether or not \( X \) and \( Y \) are indep,

\[
\mathbb{E}[X + Y] = \mathbb{E}[X] + \mathbb{E}[Y].
\]
Proof (cts case):

\[ E[X + Y] = \int_{\mathbb{R}} \int_{\mathbb{R}} (x + y) f(x, y) \, dx \, dy \]

\[ = \int_{\mathbb{R}} \int_{\mathbb{R}} x f(x, y) \, dx \, dy + \int_{\mathbb{R}} \int_{\mathbb{R}} y f(x, y) \, dx \, dy \]

\[ = \int_{\mathbb{R}} x \int_{\mathbb{R}} f(x, y) \, dy \, dx + \int_{\mathbb{R}} y \int_{\mathbb{R}} f(x, y) \, dx \, dy \]

\[ = \int_{\mathbb{R}} x f_X(x) \, dx + \int_{\mathbb{R}} y f_Y(y) \, dy \]

\[ = E[X] + E[Y]. \]
2.18 Independent RV's

Can generalize this result to more than two RV's.

Theorem: If $X_1, X_2, \ldots, X_n$ are RV's, then

$$E\left[\sum_{i=1}^{n} X_i\right] = \sum_{i=1}^{n} E[X_i].$$

Proof: Induction.
Theorem: If $X$ and $Y$ are *indep*, then $E[XY] = E[X]E[Y]$.

Proof (cts case):

\[
E[XY] = \int_{\mathbb{R}} \int_{\mathbb{R}} xyf(x, y) \, dx \, dy
\]

\[
= \int_{\mathbb{R}} \int_{\mathbb{R}} xyf_X(x)f_Y(y) \, dx \, dy \quad (X \text{ and } Y \text{ are indep})
\]

\[
= \left( \int_{\mathbb{R}} xf_X(x) \, dx \right) \left( \int_{\mathbb{R}} yf_Y(y) \, dy \right)
\]

\[
= E[X]E[Y].
\]
Remark: The above theorem is *not* necessarily true if $X$ and $Y$ are *dependent*. See the upcoming discussion on covariance.

**Theorem:** If $X$ and $Y$ are *indep*, then

$$\text{Var}(X + Y) = \text{Var}(X) + \text{Var}(Y).$$

Remark: The assumption of independence really is important here.
2.18 Independent RV's

Proof:

\[
\begin{align*}
\text{Var}(X + Y) &= \mathbb{E}[(X + Y)^2] - (\mathbb{E}[X + Y])^2 \\
&= \mathbb{E}[X^2 + 2XY + Y^2] - (\mathbb{E}[X] + \mathbb{E}[Y])^2 \\
&= \mathbb{E}[X^2] + 2\mathbb{E}[XY] + \mathbb{E}[Y^2] \\
&\quad - (\mathbb{E}[X])^2 - 2\mathbb{E}[X]\mathbb{E}[Y] - (\mathbb{E}[Y])^2 \\
&= \mathbb{E}[X^2] + 2\mathbb{E}[X]\mathbb{E}[Y] + \mathbb{E}[Y^2] \\
&\quad - (\mathbb{E}[X])^2 - 2\mathbb{E}[X]\mathbb{E}[Y] - (\mathbb{E}[Y])^2 \\
&= \mathbb{E}[X^2] - (\mathbb{E}[X])^2 + \mathbb{E}[Y^2] - (\mathbb{E}[Y])^2 \\
&= \text{Var}(X) + \text{Var}(Y).
\end{align*}
\]
Covariance and Correlation

These are measures used to define the degree of association between $X$ and $Y$ if they don’t happen to be indep.

Definition: The covariance between $X$ and $Y$ is

$$\text{Cov}(X, Y) \equiv \sigma_{XY} \equiv \mathbb{E}[(X - \mathbb{E}[X])(Y - \mathbb{E}[Y])].$$

Remark: $\text{Cov}(X, X) = \mathbb{E}[(X - \mathbb{E}[X])^2] = \text{Var}(X)$. 
If $X$ and $Y$ have positive covariance, then $X$ and $Y$ move “in the same direction.” Think height and weight.

If $X$ and $Y$ have negative covariance, then $X$ and $Y$ move “in opposite directions.” Think snowfall and temperature.
Theorem (easier way to calculate Cov):

\[ \text{Cov}(X, Y) = E[XY] - E[X]E[Y]. \]

Proof:

\[
\text{Cov}(X, Y) = E[(X - E[X])(Y - E[Y])]
\]

\[
\]

\[
\]

\[
= E[XY] - E[X]E[Y].
\]
Theorem: $X$ and $Y$ indep implies $\text{Cov}(X, Y) = 0$.

Proof:

\[ \text{Cov}(X, Y) = \mathbb{E}[XY] - \mathbb{E}[X]\mathbb{E}[Y] \]

\[ = \mathbb{E}[X]\mathbb{E}[Y] - \mathbb{E}[X]\mathbb{E}[Y] \quad (X, Y \text{ indep}) \]

\[ = 0. \]
Danger Will Robinson: Cov(X, Y) = 0 does not imply X and Y are indep!!

Example: Suppose X ∼ U(−1, 1) and Y = X^2 (so X and Y are clearly dependent).

But

\[ E[X] = \int_{-1}^{1} x \cdot \frac{1}{2} dx = 0 \] and

\[ E[XY] = E[X^3] = \int_{-1}^{1} x^3 \cdot \frac{1}{2} dx = 0, \]

so Cov(X, Y) = E[XY] − E[X]E[Y] = 0.
Definition: The **correlation** between $X$ and $Y$ is

$$
\rho = \text{Corr}(X, Y) \equiv \frac{\text{Cov}(X, Y)}{\sqrt{\text{Var}(X)\text{Var}(Y)}} = \frac{\sigma_{XY}}{\sigma_X\sigma_Y}.
$$

Remark: Cov has “square” units; corr is unitless.

Corollary: $X, Y$ indep implies $\rho = 0$. 
Theorem: It can be shown that $-1 \leq \rho \leq 1$.

$\rho \approx 1$ is “high” corr

$\rho \approx 0$ is “low” corr

$\rho \approx -1$ is “high” negative corr.

Example: Height is highly correlated with weight.
Temperature on Mars has low corr with IBM stock price.
2.18 Independent RV's

Anti-UGA Example: Suppose $X$ is the avg yards/carry that a UGA fullback gains, and $Y$ is his grade on an astrophysics test. Here's the joint pmf $f(x, y)$.

<table>
<thead>
<tr>
<th></th>
<th>$X = 2$</th>
<th>$X = 3$</th>
<th>$X = 4$</th>
<th>$f_Y(y)$</th>
</tr>
</thead>
<tbody>
<tr>
<td>$Y = 40$</td>
<td>.0</td>
<td>.2</td>
<td>.1</td>
<td>.3</td>
</tr>
<tr>
<td>$Y = 50$</td>
<td>.15</td>
<td>.1</td>
<td>.05</td>
<td>.3</td>
</tr>
<tr>
<td>$Y = 60$</td>
<td>.3</td>
<td>.0</td>
<td>.1</td>
<td>.4</td>
</tr>
</tbody>
</table>

$f_X(x)$ | .45 | .3 | .25 | 1
2.18 Independent RV's

\[ E[X] = \sum_x x f_X(x) = 2.8 \]
\[ E[X^2] = \sum_x x^2 f_X(x) = 8.5 \]
\[ \text{Var}(X) = E[X^2] - (E[X])^2 = 0.66 \]

Similarly, \( E[Y] = 51 \), \( E[Y^2] = 2670 \), and \( \text{Var}(Y) = 60 \).

\[ E[XY] = \sum_x \sum_y xy f(x, y) \]
\[ = 2(40)(.0) + \cdots + 4(60)(.1) = 140 \]
\[ \text{Cov}(X, Y) = E[XY] - E[X]E[Y] = -2.8 \]
\[ \rho = \frac{\text{Cov}(X, Y)}{\sqrt{\text{Var}(X)\text{Var}(Y)}} = -0.415. \]
Cts Example: Suppose $f(x, y) = 10x^2y$, $0 \leq y \leq x \leq 1$.

$$f_X(x) = \int_0^x 10x^2y \, dy = 5x^4, \quad 0 \leq x \leq 1$$

$$E[X] = \int_0^1 5x^5 \, dx = 5/6$$

$$E[X^2] = \int_0^1 5x^6 \, dx = 5/7$$

$$\text{Var}(X) = E[X^2] - (E[X])^2 = 0.01984$$
Similarly,

\[ f_Y(y) = \int_y^1 10x^2 y \, dx = \frac{10}{3} y(1 - y^3), \quad 0 \leq y \leq 1 \]

\[ E[Y] = \frac{5}{9}, \quad \text{Var}(Y) = 0.04850 \]

\[ E[XY] = \int_0^1 \int_0^x 10x^3 y^2 \, dy \, dx = \frac{10}{21} \]

\[ \text{Cov}(X, Y) = E[XY] - E[X]E[Y] = 0.1323 \]

\[ \rho = \frac{\text{Cov}(X, Y)}{\sqrt{\text{Var}(X)\text{Var}(Y)}} = 0.4265 \]
Theorems Involving Covariance

Theorem: \( \text{Var}(X+Y) = \text{Var}(X) + \text{Var}(Y) + 2\text{Cov}(X, Y) \), whether or not \( X \) and \( Y \) are indep.

Remark: If \( X, Y \) are indep, the Cov term goes away.

Proof: By the work we did on a previous proof,

\[
\text{Var}(X + Y) = \mathbb{E}[X^2] - (\mathbb{E}[X])^2 + \mathbb{E}[Y^2] - (\mathbb{E}[Y])^2 + 2(\mathbb{E}[XY] - \mathbb{E}[X]\mathbb{E}[Y])
\]

\[
= \text{Var}(X) + \text{Var}(Y) + 2\text{Cov}(X, Y).
\]
2.18 Independent RV’s

Theorem:

\[
\text{Var}(\sum_{i=1}^{n} X_i) = \sum_{i=1}^{n} \text{Var}(X_i) + 2\sum\sum_{i<j} \text{Cov}(X_i, X_j).
\]

Proof: Induction.

Remark: If all \(X_i\)’s are \textit{indep}, then

\[
\text{Var}(\sum_{i=1}^{n} X_i) = \sum_{i=1}^{n} \text{Var}(X_i).
\]
Theorem: $\text{Cov}(aX, bY) = ab\text{Cov}(X, Y)$.

Proof:

\[
\text{Cov}(aX, bY) = E[aX \cdot bY] - E[aX]E[bY]
\]

\[
= abE[XY] - abE[X]E[Y]
\]

\[
= ab\text{Cov}(X, Y).
\]
2.18 Independent RV's

Theorem:

\[ \text{Var}( \sum_{i=1}^{n} a_i X_i ) = \sum_{i=1}^{n} a_i^2 \text{Var}(X_i) + 2 \sum \sum_{i<j} a_i a_j \text{Cov}(X_i, X_j). \]

Proof: Put above two results together.
Example: \( \text{Var}(X - Y) = \text{Var}(X) + \text{Var}(Y) - 2 \text{Cov}(X, Y) \).

Example:

\[
\text{Var}(X - 2Y + 3Z) = \text{Var}(X) + 4 \text{Var}(Y) + 9 \text{Var}(Z) - 4 \text{Cov}(X, Y) + 6 \text{Cov}(X, Z) - 12 \text{Cov}(Y, Z).
\]
Random Samples

Definition: $X_1, X_2, \ldots, X_n$ form a random sample if

- $X_i$’s are all independent.
- Each $X_i$ has the same pmf/pdf $f(x)$.

Notation: $X_1, \ldots, X_n \overset{iid}{\sim} f(x)$ ("indep and identically distributed")
Example/Theorem: Suppose $X_1, \ldots, X_n \sim f(x)$ with $E[X_i] = \mu$ and $\text{Var}(X_i) = \sigma^2$. Define the **sample mean** as

$$ \bar{X} \equiv \frac{1}{n} \sum_{i=1}^{n} X_i. $$

Then

$$ E[\bar{X}] = E\left[\frac{1}{n} \sum_{i=1}^{n} X_i \right] = \frac{1}{n} \sum_{i=1}^{n} E[X_i] = \frac{1}{n} \sum_{i=1}^{n} \mu = \mu. $$

So the mean of $\bar{X}$ is the same as the mean of $X_i$. 
Meanwhile, \ldots

\[ \text{Var}(\bar{X}) = \text{Var}\left(\frac{1}{n} \sum_{i=1}^{n} X_i\right) \]

\[ = \frac{1}{n^2} \text{Var}\left(\sum_{i=1}^{n} X_i\right) \]

\[ = \frac{1}{n^2} \sum_{i=1}^{n} \text{Var}(X_i) \quad (X_i \text{'s indep}) \]

\[ = \frac{1}{n^2} \sum_{i=1}^{n} \sigma^2 = \frac{\sigma^2}{n}. \]

So the mean of \( \bar{X} \) is the same as the mean of \( X_i \), but the variance decreases!