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

$$\text{Pr}(X \le x, Y \le y) = \text{Pr}(X \le x)\text{Pr}(Y \le y), \ \ \forall x,y$$

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): $f(x,y) = \Pr(X = x, Y = y)$.

	$X = 1$	$X = 2$	$f_Y(y)$
$Y = 2$	.12	.28	.4
$Y = 3$	.18	.42	.6
$f_X(x)$	.3	.7	1

X and Y are indep since $f(x,y) = f_X(x)f_Y(y)$, $\forall x, y$.

Example (cts): Suppose $f(x, y) = 6xy^2$, $0 \le x \le 1$, $0 \le y \le 1$.

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

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

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

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: $f(x,y) = 6xy^2$, $0 \le x \le 1$, $0 \le y \le 1$. Take

$$a(x) \;=\; 6x,\;\; 0 \le x \le 1,\; \text{and } b(y) \;=\; y^2,\;\; 0 \le y \le 1.$$

Thus, X and Y are indep (as above).

Example: $f(x,y) = \frac{21}{4}x^2 y$, $x^2 \le y \le 1$. "Funny" (non-rectangular) limits make factoring into marginals impossible. Thus, X and Y are *not* indep.

Example: $f(x,y) = \frac{c}{x+y}$, $1 \le x \le 2$, $1 \le y \le 3$.

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?

Consequences of Independence

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

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

Theorem: *Whether or not* X and Y are indep,

$$\mathsf{E}[X+Y] = \mathsf{E}[X] + \mathsf{E}[Y].$$

Proof (cts case):

$$
\begin{aligned}
\mathsf{E}[X+Y] &= \int_{\Re}\int_{\Re}(x+y)f(x,y)\,dx\,dy \\
&= \int_{\Re}\int_{\Re} xf(x,y)\,dx\,dy + \int_{\Re}\int_{\Re} yf(x,y)\,dx\,dy \\
&= \int_{\Re} x\int_{\Re} f(x,y)\,dy\,dx + \int_{\Re} y\int_{\Re} f(x,y)\,dx\,dy \\
&= \int_{\Re} xf_X(x)\,dx + \int_{\Re} yf_Y(y)\,dy \\
&= \mathsf{E}[X] + \mathsf{E}[Y].
\end{aligned}
$$

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

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

$$\mathsf{E}\Big[\sum_{i=1}^{n} X_i\Big] = \sum_{i=1}^{n} \mathsf{E}[X_i].$$

Proof: Induction.

Theorem: If X and Y are *indep*, then $\mathsf{E}[XY] = \mathsf{E}[X]\mathsf{E}[Y]$.

Proof (cts case):

$$
\begin{aligned}
\mathsf{E}[XY] &= \int_{\Re}\int_{\Re} xyf(x,y)\,dx\,dy \\
&= \int_{\Re}\int_{\Re} xyf_X(x)f_Y(y)\,dx\,dy \quad (X \text{ and } Y \text{ are indep}) \\
&= \left(\int_{\Re} xf_X(x)\,dx\right)\left(\int_{\Re} yf_Y(y)\,dy\right) \\
&= \mathsf{E}[X]\mathsf{E}[Y].
\end{aligned}
$$

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.

Proof:

$$
\begin{aligned}
&\mathsf{Var}(X+Y) \\
&= \mathsf{E}[(X+Y)^2] - (\mathsf{E}[X+Y])^2 \\
&= \mathsf{E}[X^2 + 2XY + Y^2] - (\mathsf{E}[X] + \mathsf{E}[Y])^2 \\
&= \mathsf{E}[X^2] + 2\mathsf{E}[XY] + \mathsf{E}[Y^2] \\
&\qquad\qquad -(\mathsf{E}[X])^2 - 2\mathsf{E}[X]\mathsf{E}[Y] - (\mathsf{E}[Y])^2 \\
&= \mathsf{E}[X^2] + 2\mathsf{E}[X]\mathsf{E}[Y] + \mathsf{E}[Y^2] \\
&\qquad\qquad -(\mathsf{E}[X])^2 - 2\mathsf{E}[X]\mathsf{E}[Y] - (\mathsf{E}[Y])^2 \\
&= \mathsf{E}[X^2] - (\mathsf{E}[X])^2 + \mathsf{E}[Y^2] - (\mathsf{E}[Y])^2 \\
&= \mathsf{Var}(X) + \mathsf{Var}(Y).
\end{aligned}
$$

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

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

Remark: $\mathsf{Cov}(X,X) = \mathsf{E}[(X-\mathsf{E}[X])^2] = \mathsf{Var}(X)$.

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):

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

Proof:

$$\begin{aligned}
\mathrm{Cov}(X,Y) &= \mathrm{E}[(X-\mathrm{E}[X])(Y-\mathrm{E}[Y])] \\
&= \mathrm{E}\Big[XY - X\mathrm{E}[Y] - Y\mathrm{E}[X] + \mathrm{E}[X]\mathrm{E}[Y]\Big] \\
&= \mathrm{E}[XY] - \mathrm{E}[X]\mathrm{E}[Y] - \mathrm{E}[Y]\mathrm{E}[X] + \mathrm{E}[X]\mathrm{E}[Y] \\
&= \mathrm{E}[XY] - \mathrm{E}[X]\mathrm{E}[Y].
\end{aligned}$$

Theorem: X and Y indep implies $\mathsf{Cov}(X, Y) = 0$.

Proof:

$$
\begin{aligned}
\mathsf{Cov}(X, Y) &= \mathsf{E}[XY] - \mathsf{E}[X]\mathsf{E}[Y] \\
&= \mathsf{E}[X]\mathsf{E}[Y] - \mathsf{E}[X]\mathsf{E}[Y] \quad (X,\ Y \text{ indep}) \\
&= 0.
\end{aligned}
$$

Danger Will Robinson: $\text{Cov}(X,Y) = 0$ *does not imply* X and Y are indep!!

Example: Suppose $X \sim \text{U}(-1,1)$ and $Y = X^2$ (so X and Y are clearly *dependent*).

But

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

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

so $\text{Cov}(X,Y) = \text{E}[XY] - \text{E}[X]\text{E}[Y] = 0$.

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 $-1 \le \rho \le 1$.

$\rho \approx 1$ 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.

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 $f(x,y)$.

	$X=2$	$X=3$	$X=4$	$f_Y(y)$
$Y=40$	.0	.2	.1	.3
$Y=50$	.15	.1	.05	.3
$Y=60$	.3	.0	.1	.4
$f_X(x)$	.45	.3	.25	1

$$
\begin{aligned}
\mathsf{E}[X] &= \sum_x x f_X(x) = 2.8 \\
\mathsf{E}[X^2] &= \sum_x x^2 f_X(x) = 8.5 \\
\mathsf{Var}(X) &= \mathsf{E}[X^2] - (\mathsf{E}[X])^2 = 0.66
\end{aligned}
$$

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

$$
\begin{aligned}
\mathsf{E}[XY] &= \sum_x \sum_y xyf(x,y) \\
&= 2(40)(.0) + \cdots + 4(60)(.1) = 140 \\
\mathsf{Cov}(X,Y) &= \mathsf{E}[XY] - \mathsf{E}[X]\mathsf{E}[Y] = -2.8 \\
\rho &= \frac{\mathsf{Cov}(X,Y)}{\sqrt{\mathsf{Var}(X)\mathsf{Var}(Y)}} = -0.415.
\end{aligned}
$$

Cts Example: Suppose $f(x,y) = 10x^2y$, $0 \le y \le x \le 1$.

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

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

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

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

Similarly,

$$f_Y(y) \;=\; \int_y^1 10x^2 y\,dx \;=\; \frac{10}{3}y(1-y^3),\;\; 0 \le y \le 1$$

$$\mathsf{E}[Y] \;=\; 5/9, \quad \mathsf{Var}(Y) \;=\; 0.04850$$

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

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

$$\rho \;=\; \frac{\mathsf{Cov}(X,Y)}{\sqrt{\mathsf{Var}(X)\mathsf{Var}(Y)}} \;=\; 0.4265$$

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,

$$\begin{aligned} \text{Var}(X+Y) &= \text{E}[X^2] - (\text{E}[X])^2 + \text{E}[Y^2] - (\text{E}[Y])^2 \\ &\qquad +2(\text{E}[XY] - \text{E}[X]\text{E}[Y]) \\ &= \text{Var}(X) + \text{Var}(Y) + 2\text{Cov}(X,Y). \end{aligned}$$

Theorem:

$$\mathsf{Var}\Big(\sum_{i=1}^{n} X_i\Big) = \sum_{i=1}^{n} \mathsf{Var}(X_i) + 2\sum\sum_{i<j}\mathsf{Cov}(X_i, X_j).$$

Proof: Induction.

Remark: If all X_i's are *indep*, then

$$\mathsf{Var}\Big(\sum_{i=1}^{n} X_i\Big) = \sum_{i=1}^{n} \mathsf{Var}(X_i).$$

Theorem: $\mathsf{Cov}(aX, bY) = ab\mathsf{Cov}(X, Y)$.

Proof:

$$\begin{aligned}
\mathsf{Cov}(aX, bY) &= \mathsf{E}[aX \cdot bY] - \mathsf{E}[aX]\mathsf{E}[bY] \\
&= ab\mathsf{E}[XY] - ab\mathsf{E}[X]\mathsf{E}[Y] \\
&= ab\mathsf{Cov}(X, Y).
\end{aligned}$$

Theorem:

$$\mathsf{Var}\Big(\sum_{i=1}^{n} a_i X_i\Big)$$

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

Proof: Put above two results together.

Example: $\text{Var}(X-Y) = \text{Var}(X)+\text{Var}(Y)-2\text{Cov}(X,Y)$.

Example:

$$\begin{aligned} &\text{Var}(X - 2Y + 3Z) \\ &= \text{Var}(X) + 4\text{Var}(Y) + 9\text{Var}(Z) \\ &\quad -4\text{Cov}(X,Y) + 6\text{Cov}(X,Z) - 12\text{Cov}(Y,Z). \end{aligned}$$

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: $X_1, \ldots, X_n \overset{\text{iid}}{\sim} f(x)$ ("indep and identically distributed")

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

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

Then

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

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

Meanwhile,. . .

$$\begin{aligned}
\mathrm{Var}(\bar{X}) &= \mathrm{Var}\left(\frac{1}{n}\sum_{i=1}^{n} X_i\right) \\
&= \frac{1}{n^2}\mathrm{Var}\left(\sum_{i=1}^{n} X_i\right) \\
&= \frac{1}{n^2}\sum_{i=1}^{n} \mathrm{Var}(X_i) \quad (X_i\text{'s indep}) \\
&= \frac{1}{n^2}\sum_{i=1}^{n} \sigma^2 = \sigma^2/n.
\end{aligned}$$

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