Covariance, and a small taste of bivariate probability

[Not in the textbook!]

Q.: Suppose that X and Y are random variables and we look at $Z = X+Y$. Then $E(Z) = E(X+Y) = E(X)+E(Y)$, but if X and Y are not independent the $V(Z)$ is harder to calculate [$V(X+Y) = V(X)+V(Y)$ if they are independent].

If they are not independent we fall on the definition of $V(Z) = V(X+Y)$

$$= E\left((X+Y)^2\right) - \left[E(X)+E(Y)\right]^2$$
$$= E\left(X^2+2XY+Y^2\right) - \left[\left(E(X)\right)^2 + 2E(X)E(Y) + \left(E(Y)\right)^2\right]$$
$$= E(X^2) + 2E(XY) + E(Y^2)$$
$$- [E(X)]^2 - 2E(X)E(Y) - [E(Y)]^2$$
$$= V(X) + V(Y) + 2\left[\underbrace{E(XY) - E(X)E(Y)}_{Cov(X,Y)}\right]$$

is the covariance of X and Y.

We will only play with this when X and Y are discrete random variables with only finitely many possible values each.

⌈ It's too complicated at this level to handle anything more complicated.
e.g. the continuous case requires multivariate calculus. ⌋

Example:

$Y \backslash X$	-1	1
0	0.1	0.2
1	0.3	0.4

The sum of all the combined probabilities must be 1

$= 0.1 + 0.2 + 0.3 + 0.4$

so the "joint probability function" of X & Y

is
$m(0,-1) = 0.1$
$m(0,1) = 0.2$
$m(1,-1) = 0.3$
$m(1,1) = 0.4$

also
$m_X(-1) = 0.1 + 0.3 = 0.4$
$m_X(1) = 0.2 + 0.4 = 0.4$

$m_Y(0) = 0.1 + 0.2 = 0.3$
$m_Y(1) = 0.3 + 0.4 = 0.7$

Suppose $Z = X+Y$. We want to know $E(Z)$ and $V(Z)$.

$$E(X) = (-1)\cdot(0.1+0.3) \quad +1\cdot(0.2+0.4) = -0.4+0.6=0.2$$

$$E(Y) = 0\cdot(0.1+0.2) + 1\cdot(0.3+0.4) = 0+0.7 = 0.7$$

so $$E(Z) = E(X+Y) = E(X)+E(Y) = 0.2+0.7=0.9$$

$$E(X^2) = (-1)^2\cdot(0.1+0.3)+1^2(0.2+0.4) = 0.4+0.6 = 1$$

$$E(Y^2) = 0^2\cdot(0.1+0.2)+1^2(0.3+0.4) = 0+0.7 = 0.7$$

$$E(XY) = \overset{x}{(-1)}\cdot\overset{y}{0}\cdot 0.1 + (-1)\cdot 1\cdot 0.3 \quad = 0 - 0.3$$
$$+1\cdot 0\cdot 0.2 + 1\cdot 1\cdot 0.4 \qquad +0+0.4 = 0.1$$

$$V(X) = E(X^2) - [E(X)]^2 = 1 - [0.2]^2 = 1-0.04$$
$$= 0.96$$

$$V(Y) = E(Y^2) - [E(Y)]^2 = 0.7 - [0.7]^2 = 0.7-0.49$$
$$= 0.21$$

$$Cov(X,Y) = E(XY) - E(X)\cdot E(Y)$$
$$= 0.1 - 0.2\cdot 0.7 = 0.1 - 0.14 = -0.04$$

$$V(X+Y) = V(X)+V(Y) - 2Cov(X,Y)$$
$$= 0.96+0.21-2\cdot(-0.04)$$
$$= 1.17+0.08 = 1.25$$

If X & Y are independent, then
$$E(XY) = E(X)E(Y),$$
so $Cov(X,Y) = 0$. However,

if $Cov(X,Y) = 0$, the X and Y might <u>or might not be</u> independent.

<u>Example:</u>

$Y \backslash X$	0	1	2
-1	0.2	0	0.2
0	0	0.2	0
1	0.2	0	0.2

X & Y are <u>not</u> independent since
$P(X=1 \mid Y=0)$
$= 1 \neq P(X=1) = 0.2$.

$E(X) = 0\cdot 0.4 + 1\cdot 0.2 + 2\cdot 0.4 = 1$

$E(Y) = -1\cdot 0.4 + 0\cdot 0.2 + 1\cdot 0.4 = 0$

$E(X^2) = 0^2\cdot 0.4 + 1^2\cdot 0.2 + 2^2\cdot 0.4 = 1.8$

$E(Y^2) = (-1)^2\cdot 0.4 + 0^2\cdot 0.2 + 1^2\cdot 0.4 = 0.8$

$E(X+Y) = 1+0 = 1$

$V(X) = 1.8 - 1^2 = 0.8$

$V(Y) = 0.8 - 0^2 = 0.8$

$$E(XY) = \cancel{0\cdot(-1)\cdot 0.2} + \cancel{0\cdot 0\cdot 0} + \cancel{0\cdot 1\cdot 0.2} + \cancel{1\cdot(-1)\cdot 0} + \cancel{1\cdot 0\cdot 0.2} + \cancel{1\cdot 1\cdot 0} + 2\cdot(-1)\cdot 0.2 + \cancel{2\cdot 0\cdot 0} + 2\cdot 1\cdot 0.2 = -0.4 + 0.4 = 0$$

so $Cov(X,Y) = E(XY) - E(X)E(Y) = 0 - 1\cdot 0 = 0$

So X and Y are "uncorrelated" but not independent

Basic Properties of Covariance

(0) If X and Y are independent, then

$$Cov(X,Y) = 0.$$

(1) $V(X) = Cov(X,X)$

(2)
$$\begin{aligned} Cov(X+Z,Y) &= Cov(X,Y) + Cov(Z,Y) \\ Cov(X,Y+Z) &= Cov(X,Y) + Cov(Y,Z) \\ Cov(cX,Y) &= c\,Cov(X,Y) \\ Cov(X,cY) &= c\,Cov(X,Y) \end{aligned}$$

(3) $Cov(X,Y) = Cov(Y,X)$

(4) $V(X+Y) = V(X) + V(Y) - 2Cov(X,Y)$