Activity Solution: Variance and Covariance of Random Variables

The **mean** (**expectation**, or **expected value**) of a random variable X is denoted $E(X)$. Formally, it is defined as

$$E(X) = \sum_x xP(X = x)$$

if X is discrete, where the sum is over all possible values of X, or as

$$E(X) = \int xf(x)\,dx$$

if X is continuous, where $f(x)$ is the density function of X and again the integral is over all possible values. The **variance** of a random variable X is defined by

$$\mathrm{Var}(X) = E\left((X - E(X))^2\right).$$

It is a measure of how "spread out" the distribution of X is, in relation to its mean.

1. *Clickers question:* Expand the square to give an alternative form for Var(X). Describe the result in words.
   So

$$\begin{aligned}
\mathrm{Var}(X) &= E\left(X^2 - 2XE(X) + E(X)^2\right) \\
&= E\left(X^2\right) - 2E(X)^2 + E(X)^2 \\
&= E\left(X^2\right) - E(X)^2.
\end{aligned}$$

   the variance is also described in words as "the mean of the square minus the square of the mean."

2. *Clickers question:* If a is a constant and X any random variable, then what is Var(aX) in terms of Var(X) and a?

$$\begin{aligned}
\mathrm{Var}(aX) &= E\left((aX - E(aX))^2\right) \\
&= E\left(a^2(X - E(X))^2\right) \\
&= a^2E\left((X - E(X))^2\right) \\
&= a^2\mathrm{Var}(X).
\end{aligned}$$

3. *Clickers question:* If X and Y are two random variables then find $\mathrm{Var}(X+Y)$ in terms of expectations of X and Y and any functions of X and Y required.
We find in fact that

$$\mathrm{Var}\,(X+Y) = \mathrm{Var}\,(X) + \mathrm{Var}\,(Y) + 2\mathrm{Cov}\,(X,Y)$$

where $Cov(X,Y)$ is called the ***covariance*** *between X and Y, and is defined as*

$$\mathrm{Cov}\,(X,Y) := E\,((X - E\,(X))(Y - E\,(Y)))\,.$$

This follows since

$$\begin{aligned}\mathrm{Var}\,(X+Y) &= E([(X+Y) - E\,(X+Y)]^2) \\ &= E([(X - E\,(X)) + (Y - E\,(Y))]^2) \\ &= E((X - E\,(X))^2 + (Y - E\,(Y))^2 \\ &\quad +2\,(X - E\,(X))(Y - E\,(Y))) \\ &= \mathrm{Var}\,(X) + \mathrm{Var}\,(Y) + 2\mathrm{Cov}\,(X,Y)\end{aligned}$$

as required.

4. With $\mathrm{Cov}(X,Y)$ the covariance between X and Y, show that

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

5. Show that $\mathrm{Cov}(X,Y) = \mathrm{Cov}(Y,X)$.
This is obvious from the definition, which is symmetric between X and Y.

6. What is $\mathrm{Cov}(X,Y)$ if X and Y are independent?
With X and Y independent we have

$$E\,(XY) = E\,(X)\,E\,(Y)$$

(Why?), and so $Cov(X,Y) = 0$.

7. *Clickers question:* Find $\mathrm{Cov}(X+Z,Y)$ in terms of $\mathrm{Cov}(X,Y)$ and $\mathrm{Cov}(Z,Y)$, for any random variables X, Y, and Z.

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

8. *Clickers question:* If a and b are constants, and X and Y random variables, simplify $\mathrm{Cov}(aX, bY)$.

$$\begin{aligned}
\mathrm{Cov}\left(aX, bY\right) &= E\left(abXY\right) - E\left(aX\right)E\left(bY\right) \\
&= abE\left(XY\right) - abE\left(X\right)E\left(Y\right) \\
&= ab\mathrm{Cov}\left(X, Y\right).
\end{aligned}$$