A Simple Proof about the Correlation of Two Random Variables

David Darmon

November 1, 2011

The following proof is due to Dr. Roger Coleman, my Probability professor from Ursinus College. I expect he did not originate it but instead borrowed it from someone else.

Theorem 1. *Let X and Y be random variables with finite variances. Then*

$$-1 \leq \mathit{Corr}(X, Y) \leq 1.$$

Proof. We will consider the linear combination of X and Y given by $Z = X + tY$ where t is some scalar. Note that Z itself is a random variable. Thus, we may takes its variance. By a property of variance,

$$\mathrm{Var}(X + tY) \geq 0, \text{ for all values of } t. \tag{1}$$

Moreover, using an identity for the variance of linear combinations of variables,

$$\begin{aligned} \mathrm{Var}(X + tY) &= \mathrm{Var}(X) + t^2 \mathrm{Var}(Y) + 2t\mathrm{Cov}(X, Y) & (2) \\ &= t^2 \mathrm{Var}(Y) + 2t\mathrm{Cov}(X, Y) + \mathrm{Var}(X). & (3) \end{aligned}$$

Thus, combining our two results so far, we see that

$$t^2 \mathrm{Var}(Y) + 2t\mathrm{Cov}(X, Y) + \mathrm{Var}(X) \geq 0. \tag{4}$$

But note that (4) is just a quadratic inequality in t. That is, we have (by another name)

$$at^2 + bt + c \geq 0 \tag{5}$$

where $a = \mathrm{Var}(Y)$, $b = 2\mathrm{Cov}(X, Y)$, and $c = \mathrm{Var}(X)$. This quadratic is greater than or equal to 0 for all values of t. This tells us something about the roots of this quadratic. Namely, we know that it must have either a single real root or two imaginary roots (i.e. no real roots). Draw a picture of the two scenarios where $at^2 + bt + c \geq 0$ for all t to convince yourself of this. We can make statements about the roots of a quadratic using the quadratic formula. Recall

$$x = \frac{-b \pm \sqrt{b^2 - 4ac}}{2a}. \tag{6}$$

The quantity $b^2 - 4ac$ under the square root above is called the discriminant because it helps us discriminate the types of roots a quadratic will have. Because we know that (5) has either one real root or two imaginary roots, it follows that the discriminant must be negative or zero (if it were positive, we would have two real roots). Thus, we have that

$$b^2 - 4ac \leq 0. \tag{7}$$

Substituting for a, b, and c and performing a few algebraic manipulations, we arrive at

$$\begin{aligned}
[2\text{Cov}(X,Y)]^2 - 4\text{Var}(X)\text{Var}(Y) &\leq 0 \\
[2\text{Cov}(X,Y)]^2 &\leq 4\text{Var}(X)\text{Var}(Y) \\
\sqrt{[2\text{Cov}(X,Y)]^2} &\leq 2\sqrt{\text{Var}(X)\text{Var}(Y)} \\
|2\text{Cov}(X,Y)| &\leq 2\sqrt{\text{Var}(X)\text{Var}(Y)} \\
-2\sqrt{\text{Var}(X)\text{Var}(Y)} \leq 2\text{Cov}(X,Y) &\leq 2\sqrt{\text{Var}(X)\text{Var}(Y)} \\
-1 \leq \frac{\text{Cov}(X,Y)}{\sqrt{\text{Var}(X)\text{Var}(Y)}} &\leq 1 \\
-1 \leq \text{Corr}(X,Y) &\leq 1,
\end{aligned}$$

as was to be shown. □

A simpler proof uses an important inequality from analysis.

Theorem 2. *(The Cauchy-Schwarz Inequality) Let ξ and η be random variables with finite variances. Then*

$$[\mathbb{E}(\xi\eta)]^2 \leq \mathbb{E}(\xi^2)\mathbb{E}(\eta^2).$$

Take $\xi = X - \mathbb{E}(X)$ and $\eta = Y - \mathbb{E}(\eta)$ and Theorem 1 proves itself! Of course, we need to prove the Cauchy-Schwarz inequality before using it. Fortunately (and perhaps unsurprisingly), we may generalize the method used in Theorem 1 to prove Cauchy-Schwarz.

We had to sneak up on the proof of Theorem 1. This occurs frequently in mathematics. Cleverness pays off. But don't be disheartened. Remember one of my favorite quotes about mathematics:

> It is important to remember that differential equations have been studied by a great many people over the last 300 years, and during most of that time there were no TVs, no cell phones, and no Internet. Given 300 years with nothing else to do, you might get quite adept at changing variables. - from *Differential Equations* by Paul Blanchard, Robert Devaney, and Glen Hall

The same certainly applies to probability!