

**Exercise Number: 4.5.11**

**Proposition.** Let  $X$  and  $Y$  be independent, each with standard normal distribution, let  $a, b \in \mathbb{R}$  and let  $Y = aX + bZ$ . Let  $\rho$  denote the correlation coefficient, with its input being two random variables. Then

- $\rho(X, Y) = \frac{a}{\sqrt{a^2 + b^2}}$ .
- $|\rho(X, Y)| \leq 1$ .
- $\{b = 0, a \in \mathbb{R}^+\} \iff \rho(X, Y) = 1$ .
- $\{b = 0, a \in \mathbb{R}^-\} \iff \rho(X, Y) = -1$ .

**Proof.** We prove the proposition in the order presented.

- This proof consists solely of computations guided by the definitions, and linearity of expectation. First we compute the covariance of the variables.

$$\begin{aligned} \text{Cov}(X, Y) &= \mathbb{E}[(X - \mu_X)(Y - \mu_Y)] = \mathbb{E}[(X - \mu_X)(aX + bY - \mathbb{E}[aX + bY])] = \mathbb{E}[(X - \mu_X)(a(X - \mu_X) + b(Y - \mu_Y))] \\ &= \mathbb{E}[a(X - \mu_X)^2] \mathbb{E}[b(X - \mu_X)(Z - \mu_Z)] = a\text{Var}(X) + b\text{Cov}(X, Z) = a\text{Var}(X) \end{aligned}$$

where the last step follows by page 47 of the text: the covariance of independent variables is 0. Now by definition of correlation,

$$\rho(X, Y) = \frac{\text{Cov}(X, Y)}{\sqrt{\text{Var}(X)\text{Var}(Y)}} = \frac{a\text{Var}(X)}{\sqrt{\text{Var}(X)\text{Var}(Y)}} = a\sqrt{\frac{\text{Var}(X)}{\text{Var}(aX + bZ)}} = \sqrt{\frac{a^2\text{Var}(X)}{a^2\text{Var}(X) + b^2\text{Var}(Z)}}$$

where the last step follows since the covariance of the sum of independent variables is merely the sum of the variances. Since  $X$  and  $Z$  have standard normal distribution (variance 1), we have as our final result

$$\rho(X, Y) = \frac{a}{\sqrt{a^2 + b^2}}.$$

- Since  $\sqrt{a^2 + b^2} \geq |a| \in \mathbb{R}$  this follows immediately.
- Follows directly by elementary algebra.
- Follows directly by elementary algebra.

□

**Discussion.**

While this problem deals with, in particular, standard normal distributions, it is revealing concerning properties of random variables in general with respect to correlation.

If two variables are independent, their covariance is necessarily 0. Intuitively, covariance is a generalization of variance: on two r.v.'s  $X, Y$ , it is the expectation that the product of their joint differences from their means. So if the covariance of two random variables is negative, this implies a tendency for them to diverge in separate direction on similar inputs. Likewise, if it is positive, they tend to move together relative to their means.

The problem is that, in and of itself covariance is not a “relative” measure: that is to say, the magnitude is somewhat meaningless. Algebraically, covariance arises as the middle term of taking the variance of the sum of two random variables, and so cannot be expected to have much meaning in and of itself.

Fortunately, as this exercise somewhat touches upon, we can “give covariance meaning” by normalizing by square root of the product of the variances. This new measure, correlation, is bounded inclusively by  $-1$  and  $1$ , so now magnitude may come into play: if two variables have close to negative one correlation, they almost maximally move in opposite directions on a shared event. If two variables are positive, they tend to move in tandem almost as much as possible. Finally, if the correlation is close to 0, this implies almost no relation: i.e. they are close to independent.

In the above problem, we saw that when two r.v.'s that are independent are added together, a certain level of preservation of their independence emerges when compared to one of the original variables: that is to say, what determines the new correlation are the scalars in the linear combination. If  $a$  is large compared to  $b$ , then  $X$  and  $Y$  are either highly correlated or highly inversely correlated. If  $b$  is large wrt  $a$ , then there will be little correlation.