Proof Practice

Emanuel Mejía

Week 3

1.Maximizing Correlation

Correlation is a measure of linear dependence. Then, what are the possible values for correlation when one random variable is a linear function of another? To fix terms, suppose that X and Y are random variables, and a is a constant where $a \neq 0$ and b is any constant in $\mathbb{R}$. Furthermore, suppose that Y is a function of X that takes the following form:

$$Y = aX + b$$

What are the possible values for $\rho(X, Y)$, the correlation between X and Y?

$$\begin{aligned}
\rho(x,y) &= \frac{Cov\left[X,Y\right]}{\sigma\left[X\right]\sigma\left[Y\right]} = \frac{Cov\left[X,Y\right]}{\sqrt{V\left[X\right]}\sqrt{V\left[Y\right]}} \\
&= \frac{Cov\left[X,aX+b\right]}{\sqrt{V\left[X\right]}\sqrt{V\left[aX+b\right]}} \\
&= \frac{Cov\left[X,aX+b\right]}{\sqrt{V\left[X\right]}\sqrt{a^2V\left[X\right]}} \\
&= \frac{Cov\left[X,aX+b\right]}{|a|\left(\sqrt{V\left[X\right]}\right)^2} \\
&= \frac{Cov\left[X,aX+b\right]}{|a|\,V\left[X\right]} \\
&= \frac{E\left[\left(X-E\left[X\right]\right)\left(aX+b-E\left[aX+b\right]\right)\right]}{|a|\,V\left[X\right]} \\
&= \frac{E\left[\left(X-E\left[X\right]\right)\left(aX\cancel{+b}-aE\left[X\right]\cancel{-b}\right)\right]}{|a|\,V\left[X\right]} \\
&= \frac{E\left[\left(X-E\left[X\right]\right)\left(a\left(X-E\left[X\right]\right)\right)\right]}{|a|\,V\left[X\right]} \\
&= \frac{E\left[a\left(X-E\left[X\right]\right)^2\right]}{|a|\,V\left[X\right]} \\
&= \frac{a\cancel{V\left[X\right]}}{|a|\,\cancel{V\left[X\right]}} = \frac{a}{|a|}
\end{aligned}$$

And since we already know that $a \neq 0$:

$$= \begin{cases} 1 & a > 0 \\ -1 & a < 0 \end{cases}$$