Math Stat 465

Covariance and Correlation

Suppose that X and Y have a joint distribution.

Def. The **covariance** of X and Y is $\mathrm{Cov}(X, Y) = \mathrm{E}[(X - \mu_X)(Y - \mu_Y)]$.
If $\mathrm{Cov}(X, Y) = 0$, we say that X and Y are **uncorrelated**.

Here are some properties:

C1. $\mathrm{Cov}(X, Y) = \mathrm{E}(XY) - \mu_X\mu_Y$. (Computational formula)

C2. $-\sigma_X\sigma_Y \leq \mathrm{Cov}(X, Y) \leq \sigma_X\sigma_Y$, with equality iff $Y = aX + b$ or $X = c$.

(More precisely: If $Y = aX + b$ then one of the $\leq$–signs is $=$, depending on whether a is positive or negative. If X is a constant then all three members on the left are zero.)

C3. X and Y independent $\Rightarrow$ X and Y uncorrelated.

In general, the converse ($\Leftarrow$) is false. It is true, however, if (X, Y) is bivariate normal.

C4. $\mathrm{Cov}(X, Y) = \mathrm{Cov}(Y, X)$

C5. $\mathrm{Cov}(X, X) = \mathrm{V}(X)$.

C6. $\mathrm{Cov}(aX + bY, cZ + dW) = ac\mathrm{Cov}(X, Z) + ad\mathrm{Cov}(X, W) + bc\mathrm{Cov}(Y, Z) + bd\mathrm{Cov}(Y, W)$ (like FOIL).

Example. Use C4, C5, and C6 to calculate

$\mathrm{V}(X + Y)$

$\mathrm{V}(X + Y + Z)$

How would this change if X, Y and Z were independent?

Def.: $\rho = \dfrac{\mathrm{Cov}(X, Y)}{\sigma_X\sigma_Y}$ is the **correlation coefficient** between X and Y. Thus X and Y are uncorrelated if $\rho = 0$.

R1. $-1 \leq \rho \leq 1$, with equality as in (C2) above.

R2. Independence $\Rightarrow$ uncorrelatedness.