5.3.1 Covariance and Correlation

Consider two random variables X and Y. Here, we define the **covariance** between X and Y, written $\textrm{Cov}(X,Y)$. The covariance gives some information about how X and Y are statistically related. Let us provide the definition, then discuss the properties and applications of covariance.

> The **covariance** between X and Y is defined as
>
> $$\textrm{Cov}(X,Y)=E\big[(X-EX)(Y-EY)\big]=E[XY]-(EX)(EY).$$

Note that

$$\begin{aligned}
E\big[(X-EX)(Y-EY)\big] &= E\big[XY-X(EY)-(EX)Y+(EX)(EY)\big]\\
&= E[XY]-(EX)(EY)-(EX)(EY)+(EX)(EY)\\
&= E[XY]-(EX)(EY).
\end{aligned}$$

Intuitively, the covariance between X and Y indicates how the values of X and Y move relative to each other. If large values of X tend to happen with large values of Y, then $(X-EX)(Y-EY)$ is positive on average. In this case, the covariance is positive and we say X and Y are positively correlated. On the other hand, if X tends to be small when Y is large, then $(X-EX)(Y-EY)$ is negative on average. In this case, the covariance is negative and we say X and Y are negatively correlated.

Example 5.32
Suppose $X \sim Uniform(1,2)$, and given $X=x$, Y is exponential with parameter $\lambda=x$. Find $\textrm{Cov}(X,Y)$.

Solution

We can use $\textrm{Cov}(X,Y)=EXY-EXEY$. We have $EX=\frac{3}{2}$ and

$$
\begin{aligned}
EY &= E[E[Y|X]] && \big(\text{law of iterated expectations (Equation 5.17)}\big) \\
&= E\left[\frac{1}{X}\right] && \big(\text{since } Y|X \sim Exponential(X)\big) \\
&= \int_1^2 \frac{1}{x} dx \\
&= \ln 2.
\end{aligned}
$$

We also have

$$
\begin{aligned}
EXY &= E[E[XY|X]] && \big(\text{law of iterated expectations}\big) \\
EXY &= E[XE[Y|X]] && \big(\text{since} E[X|X=x]=x\big) \\
&= E\left[X\frac{1}{X}\right] && \big(\text{since } Y|X \sim Exponential(X)\big) \\
&= 1.
\end{aligned}
$$

Thus,

$$
\mathrm{Cov}(X,Y) = E[XY] - (EX)(EY) = 1 - \frac{3}{2}\ln 2.
$$

Now we discuss the properties of covariance.

Lemma 5.3

The covariance has the following properties:

1. $\mathrm{Cov}(X,X) = \mathrm{Var}(X)$;
2. if X and Y are independent then $\mathrm{Cov}(X,Y) = 0$;
3. $\mathrm{Cov}(X,Y) = \mathrm{Cov}(Y,X)$;
4. $\mathrm{Cov}(aX,Y) = a\mathrm{Cov}(X,Y)$;
5. $\mathrm{Cov}(X+c,Y) = \mathrm{Cov}(X,Y)$;
6. $\mathrm{Cov}(X+Y,Z) = \mathrm{Cov}(X,Z) + \mathrm{Cov}(Y,Z)$;
7. more generally,

$$
\mathrm{Cov}\left(\sum_{i=1}^{m} a_i X_i, \sum_{j=1}^{n} b_j Y_j\right) = \sum_{i=1}^{m}\sum_{j=1}^{n} a_i b_j \mathrm{Cov}(X_i, Y_j).
$$