Lecture 9: Correlation & Characteristic Function

26 March, 2018

Sunil Kumar Gauttam

Department of Mathematics, LNMIIT

Covariance of X and Y is a measure of a linear relationship of X and Y in the following sense: $\mathrm{cov}(X, Y)$ will be positive when $X - EX$ and $Y - EY$ tend to have the same sign with high probability, and $\mathrm{cov}(X, Y)$ will be negative when $X - EX$ and $Y - EY$ tend to have opposite signs with high probability. Thus the sign of $\mathrm{cov}(X, Y)$ gives information about the linear relationship of X and Y; however, its actual magnitude does not have much meaning since it depends on the variability of X and Y. Therefore $\mathrm{cov}(X, Y)$ the number itself does not give information about the strength of the relationship between X and Y.

The correlation coefficient removes, in a sense, the individual variability of each X and Y by dividing the covariance by the product of the standard deviations, and thus the correlation coefficient is a better measure of the linear relationship of X and Y than is the covariance. Also, the correlation coefficient is unitless.

Definition 9.1 *The correlation coefficient of two random variables X and Y, denoted by $\rho(X, Y)$ is defined as*

$$\rho(X, Y) = \frac{cov(X, Y)}{\sqrt{var(X)var(Y)}},$$

provided $var(X) > 0$ and $var(Y) > 0$.

Example 9.2 *A standard normal random variable X satisfies: $EX = 0, EX^2 = 1, EX^3 = 0, EX^4 = 3$. Let $Y = a + bX + cX^2$. Find the correlation coefficient $\rho(X, Y)$.*

Solution:

$$\begin{aligned}
\mathrm{cov}(X, Y) &= E[XY] - E[X]E[Y] = E\left[aX + bX^2 + cX^3\right] - 0 \times E[Y] \\
&= aEX + bEX^2 + cEX^3 = b \\
\mathrm{var}(X) &= EX^2 - (EX)^2 = 1 \\
\mathrm{var}(Y) &= EY^2 - (EY)^2 = E(a^2 + b^2X^2 + 2abX + c^2X^4 + 2c(a + bX)X^2) - (a + c)^2 \\
&= a^2 + b^2 + 3c^2 + 2ac - a^2 - c^2 - 2ac = b^2 + 2c^2
\end{aligned}$$

Therefore

$$\rho(X, Y) = \frac{\mathrm{cov}(X, Y)}{\sqrt{\mathrm{var}(X)\mathrm{var}(Y)}} = \frac{b}{\sqrt{b^2 + 2c^2}}.$$

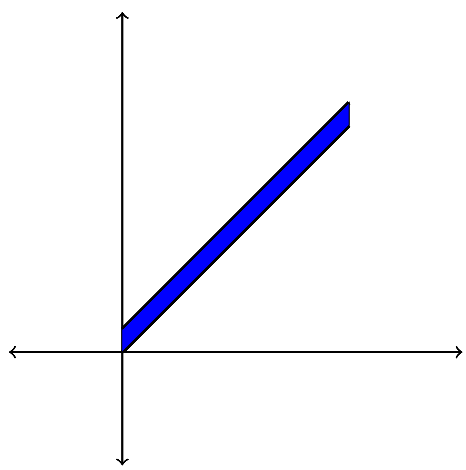
∎

The nature of the linear relationship measured by the covariance and correlation is somewhat explained in the following theorem.

Proposition 9.3 *The correlation coefficient between two random variables X and Y satisfies the following properties.*

1. $|\rho(X,Y)| \le 1$.
2. $|\rho(X,Y)| = 1$ *if and only if there exits real numbers a, b with $a \neq 0$ such that $Y = aX + b$. If $\rho(X,Y) = 1$ then $a > 0$ and if $\rho(X,Y) = -1$, then $a < 0$.*

Remark 9.4 *Intuitively, if there is a line $y = ax + b$, with $a \neq 0$, such that values of (X,Y) have high probability being near to this line, then the correlation between X and Y will be near 1 or -1. But if no such line exists, the correlation will be near zero.*



In the figure above the blue region represents the set of point where the following joint density is positive.

$$f(x,y) = \begin{cases} 10 & ; \quad 0 < x < 1, x < y < x + \frac{1}{10} \\ 0 & ; \quad \textit{otherwise} \end{cases}$$

With the above joint pdf, one can show that $\rho(X,Y) = \sqrt{\dfrac{100}{101}}$, *which is close to* 1.

Remark 9.5 *Covariance and correlation measure only a particular kind of linear relationship. But it may happen that X and Y have a strong relationship but their covariance and correlation are small or even zero, because the relationship is not linear. In fact in Example 9.2, we see that $\rho(X,Y) \le \frac{b}{\sqrt{2c}}$. Therefore if b is small and c is large then correlation is small. If $b = 0$, then $\operatorname{cov}(X,Y) = 0$ and $\rho(X,Y) = 0$ but $Y = a + cX^2$.*

Example 9.6 *Let X and Y be two random variables. Suppose that $var(X) = 4$, and $var(Y) = 9$. If we know that the two random variables $Z = 2X - Y$ and $W = X + Y$ are independent, find $\rho(X, Y)$.*

Solution: Since independent random variables are uncorrelated, therefore we have

$$
\begin{aligned}
0 = \mathrm{cov}(Z, W) &= \mathrm{cov}(2X - Y, X + Y) = \mathrm{cov}(2X - Y, X) + \mathrm{cov}(2X - Y, Y) \\
&= \mathrm{cov}(2X, X) + \mathrm{cov}(-Y, X) + \mathrm{cov}(2X, Y) + \mathrm{cov}(-Y, Y) \\
&= 2\mathrm{cov}(X, X) - \mathrm{cov}(Y, X) + 2\mathrm{cov}(X, Y) - \mathrm{cov}(Y, Y) = 2var(X) + cov(X, Y) - var(Y) \\
&= 8 + cov(X, Y) - 9 \implies cov(X, Y) = 1 \\
\rho(X, Y) &= \frac{1}{\sqrt{4 \times 9}} = \frac{1}{6}
\end{aligned}
$$

■

Complex-valued Random Variables

A complex-valued random variable $Z : \Omega \to \mathbb{C}$ can be written in the form $Z = X + iY$, where X and Y are real-valued random variables. Its expectation EZ is defined as $EZ = E(X + iY) = EX + iEY$ whenever EX and EY are well defined and finite. The formula $E(a_1 Z_1 + a_2 Z_2) = aEZ_1 + a_2 EZ_2$ is valid whenever a_1 and a_2 are complex constants and Z_1 and Z_2 are complex-valued random variables having finite expectation.

Characteristic Function

We introduce the notion of characteristic function of a random variable and study its properties. Characteristic function serves as an important tool for analyzing random phenomenon.

Definition 9.7 *The characteristic function of a random variable X is defined*

$$\phi_X(t) = E\left[e^{itX}\right], \quad t \in \mathbb{R}$$

So basically $\phi_X : \mathbb{R} \to \mathbb{C}$.

The advantage of the characteristic function is that it is defined for all real-valued random variables. Because for any real-valued random variable X and for any real number t, the random variables $\cos tX, \sin tX$ are bounded by 1. Therefore, both have finite expectation bounded by 1, hence $\phi_X(t)$ is defined for all t and for all X.

Example 9.8 *Let $X \sim$ Bernoulli (p). Find its characteristic function.*

Solution:

$$\begin{aligned}
\phi_X(t) &:= E\left[e^{itX}\right] \\
&= e^{it}P(X=1) + e^{0}P(X=0) \\
&= e^{it}p + (1-p)
\end{aligned}$$

■