COVARIANCE AND CORRELATION LECTURE NOTES

PROFESSOR HOHN

Contents

1. Covariance

Definition 1.1. Let X and Y be jointly distributed random variable. The *covariance* of X and Y is defined by

$$\mathrm{Cov}(X, Y) = \mathbb{E}[XY] - \mathbb{E}[X]\mathbb{E}[Y]$$

This is equivalent to

$$\mathrm{Cov}(X, Y) = \mathbb{E}[(X - \mu_X)(Y - \mu_Y)]$$

where $\mu_X = \mathbb{E}[X]$ and $\mu_Y = \mathbb{E}[Y]$.

Properties 1.2. Let X, Y, and Z be jointly distributed random variables. From the definition of covariance, we derive the following:

(1) The covariance generalizes variance: $\mathrm{Cov}(X, X) = \mathrm{Var}(X)$.
(2) The covariance is symmetric: $\mathrm{Cov}(X, Y) = \mathrm{Cov}(Y, X)$.
(3) For any fixed scalars $a, b \in \mathbb{R}$, $\mathrm{Cov}(aX + b, Y) = a\,\mathrm{Cov}(X, Y)$.
(4) The covariance is bilinear: $\mathrm{Cov}(X + aY, Z) = \mathrm{Cov}(X, Z) + a\,\mathrm{Cov}(Y, Z)$ for any fixed $a \in \mathbb{R}$.
(5) If X and Y are independent, then $\mathrm{Cov}(X, Y) = 0$.

Exercise 1.3. *Show that each of the listed properties of the covariance is true.*

Solution. Click here. ■

Remark 1.4. *It is true that if X and Y are independent, then* $\operatorname{Cov}(X,Y)=0$*. However, if* $\operatorname{Cov}(X,Y)=0$ *then we can* **not** *immediately conclude that X and Y are independent. Put graphically:*

$$X\&Y \textit{ independent } \implies \operatorname{Cov}(X,Y)=0$$

$$\operatorname{Cov}(X,Y)=0 \not\Longrightarrow X\&Y \textit{ independent.}$$

Example 1.5. *Let $Z \stackrel{d}{=} N(0,1)$ and define $X=Z^2$. It is clear that X and Z are not independent (clearly X depends on Z). Show that* $\operatorname{Cov}(X,Z)=0$ *even though X and Z are dependent.*

Proof. With $X=Z^2$ we have

$$\operatorname{Cov}(X,Z)=\mathbb{E}[XZ]-\mathbb{E}[X]\mathbb{E}[Z]=\mathbb{E}[Z^3]-\mathbb{E}[Z^2]\mathbb{E}[Z].$$

Now, we know that $\mathbb{E}[Z]=0$ (since $Z \stackrel{d}{=} N(0,1)$, so the mean is 0). Also,

$$\mathbb{E}[Z^3]=\frac{1}{\sqrt{2\pi}}\int_{-\infty}^{\infty} x^3 e^{-x^2/2}\,dx.$$

We can perform an integration by parts with $u=x^2$ and $dv=xe^{-x^2/2}$. Then $du=2xdx$ and $v=e^{-x^2/2}$. So,

$$\begin{aligned}\int_{-\infty}^{\infty} x^3 e^{-x^2/2}\,dx &= x^2e^{-x^2/2}\Big|_{-\infty}^{\infty}-2\int_{-\infty}^{\infty} xe^{-x^2/2}\,dx\\ &= 0-2\int_{-\infty}^{\infty} xe^{-x^2/2}\,dx\\ &= -2\int_{-\infty}^{\infty} xe^{-x^2/2}\,dx\end{aligned}$$

Recognize that this last integral is (up to a scalar constant) the same integral we calculate to find $\mathbb{E}[Z]$, which is 0. Therefore $\mathbb{E}[Z^3]=0$. Hence $\operatorname{Cov}(X,Z)=0-0=0$.

Another way to come upon $\mathbb{E}[Z^3]=0$ is by noticing that x^3 is an odd function, and $e^{-x^2/2}$ is an even function. Hence, $x^3e^{-x^2/2}$ is an odd function. We know that the integral of an odd function from $-\infty$ to ∞ is zero. Thus,

$$\mathbb{E}[Z^3]=\frac{1}{\sqrt{2\pi}}\int_{-\infty}^{\infty} x^3 e^{-x^2/2}\,dx=0$$

■

Exercise 1.6. *Let X be a random selection of number $\{2,3,4,5\}$, equally likely to be any of the found numbers. Once X is drawn, Y is randomly drawn from the numbers $\{1,2,...,X\}$ with equal probability of getting each number. For example, given that $X=3$, then Y is drawn from the numbers $\{1,2,3\}$ with $P(Y=1 \mid X=3)=1/3$, $P(Y=2 \mid X=3)=1/3$, and $P(Y=3 \mid X=3)=1/3$. Find the joint probability mass function of X and Y and find* $\mathrm{Cov}(X,Y)$.

Solution. Click here ■

Exercise 1.7. *Let X and Y be jointly continuous random variables with joint density*

$$f_{X,Y}(s,t)=\begin{cases} c\,(s^2e^{-2t}+e^{-t}) & 0<s<1, 0<t<\infty \\ 0 & \textit{otherwise} \end{cases}$$

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

Solution. Let's first find c. We have

$$c\int_0^\infty\int_0^1 (s^2e^{-2t}+e^{-t})\,dsdt = c\int_0^\infty \left(\frac{1}{3}e^{-2t}+e^{-t}\right)dt = c\left(\frac{1}{6}+1\right)=\frac{7c}{6}.$$

Therefore, $c=\frac{6}{7}$. For $\mathrm{Cov}(X,Y)=\mathbb{E}[XY]-\mathbb{E}[X]\mathbb{E}[Y]$, we will need the following:

$$\begin{aligned}\mathbb{E}[XY] &= \frac{6}{7}\int_0^\infty\int_0^1 st(s^2e^{-2t}+e^{-t})\,dsdt \\ &= \frac{6}{7}\int_0^\infty\int_0^1 (s^3te^{-2t}+ste^{-t})\,dsdt \\ &= \frac{6}{7}\int_0^\infty \left(\frac{t}{4}\,e^{-2t}+\frac{t}{2}\,e^{-t}\right)dt \\ &= \frac{6}{7}\left(\frac{1}{4}\cdot\frac{1}{4}+\frac{1}{2}\cdot 1\right) \\ &= \frac{6}{7}\cdot\frac{9}{16}.\end{aligned}$$

$$\begin{aligned}\mathbb{E}[X] &= \frac{6}{7}\int_0^\infty \int_0^1 s(s^2e^{-2t} + e^{-t})\, dsdt \\ &= \frac{6}{7}\int_0^\infty \int_0^1 (s^3e^{-2t} + se^{-t})\, dsdt \\ &= \frac{6}{7}\int_0^\infty \left(\frac{1}{4}e^{-2t} + \frac{1}{2}e^{-t}\right) dt \\ &= \frac{6}{7}\left(\frac{1}{4}\cdot\frac{1}{2} + \frac{1}{2}\cdot 1\right) \\ &= \frac{6}{7}\cdot\frac{5}{8}\end{aligned}$$

and

$$\begin{aligned}\mathbb{E}[Y] &= \frac{6}{7}\int_0^\infty \int_0^1 t(s^2e^{-2t} + e^{-t})\, dsdt \\ &= \frac{6}{7}\int_0^\infty \int_0^1 (s^2te^{-2t} + te^{-t})\, dsdt \\ &= \frac{6}{7}\int_0^\infty \left(\frac{t}{3}e^{-2t} + te^{-t}\right) dt \\ &= \frac{6}{7}\left(\frac{1}{3}\cdot\frac{1}{4} + 1\right) \\ &= \frac{6}{7}\cdot\frac{13}{12}\end{aligned}$$

Therefore,

$$\mathrm{Cov}(X,Y) = \mathbb{E}[XY] - \mathbb{E}[X]\mathbb{E}[Y] = \frac{6}{7}\Big(\frac{9}{16} - \frac{5}{8}\cdot\frac{13}{12}\cdot\frac{6}{7}\Big) = -\frac{3}{196}.$$

■

2. Correlation

Definition 2.1. Suppose that X and Y are jointly distributed random variables with non-zero variances. Then the *correlation* of X and Y is

$$\mathrm{Corr}(X,Y) = \frac{\mathrm{Cov}(X,Y)}{\sqrt{\mathrm{Var}(X)\,\mathrm{Var}(Y)}}.$$

Intuition 2.2. Remember back in multivariable calculus that if you take vectors $\mathbf{v}, \mathbf{w} \in \mathbb{R}^n$ we could define the dot-product $\mathbf{v}\cdot\mathbf{w}$ of the two vectors. Also remember that if you wanted to find the length of a vector $\mathbf{v}$, you could do this by $\|\mathbf{v}\| = \sqrt{\mathbf{v}\cdot\mathbf{v}}$. Now, speaking abstractly, the covariance of two random variables acts like a generalized "dot-product" between the two variables. That is, we can think of $\mathrm{Cov}(X,Y)$ roughly as a dot product of X and Y. In this analogy then, the "length" of a random variable is $\|X\| = \sqrt{\mathrm{Cov}(X,X)} = \sqrt{\mathrm{Var}(X)}$. Let's also remember that with two vectors $\mathbf{v}$ and $\mathbf{w}$, we had an interpretation of the dot-product as $\mathbf{v}\cdot\mathbf{w} = \|\mathbf{v}\|\,\|\mathbf{w}\|\cos(\theta)$ where θ is the angle between the vectors. Solving for $\cos(\theta)$ we have $\cos(\theta) = \frac{\mathbf{v}\cdot\mathbf{w}}{\|v\|\,\|w\|}$. Hence with respect to the interpretation of $\mathrm{Cov}(X,Y)$ as the dot product of X and Y, $\|X\| = \sqrt{\mathrm{Var}(X)}$, and $\|Y\| = \sqrt{\mathrm{Var}(Y)}$, the formula for correlation gives $\mathrm{Corr}(X,Y) = \frac{\mathrm{Cov}(X,Y)}{\|X\|\|Y\|}$. With this interpretation then, you can invision $\mathrm{Corr}(X,Y) = \cos(\theta)$ where θ is a (very roughly) the "angle between" the random variables X and Y.

Proposition 2.3. *For any random variables X and Y, it holds that $-1 \le \mathrm{Corr}(X,Y) \le 1$.*

Intuitive Proof. With the interpretation that $\mathrm{Corr}(X,Y) = \cos(\theta)$ where θ is some generalized notion of the angle between X and Y, since cosine is always bounded between -1 and 1, so must be the correlation. ∎

Definition 2.4. If $\mathrm{Corr}(X,Y) = 0$, we say that X and Y are *uncorrelated*.

Remark 2.5. *Notice that $\mathrm{Corr}(X,Y) = 0$ if and only if $\mathrm{Cov}(X,Y) = 0$. So, all the previous properties discussing when the covariance is zero still hold for the correlation.*

Example 2.6. *Show that $\mathrm{Corr}(aX+b,Y) = \mathrm{Corr}(X,Y)$ for any fixed scalars $a > 0$ and $b \in \mathbb{R}$.*

Solution. For random variables X and Y, and scalars $a, b \in \mathbb{R}$, we have shown that $\mathrm{Cov}(aX+b,Y) = a\,\mathrm{Cov}(X,Y)$, and we have seen that $\mathrm{Var}(aX+b) = a^2\,\mathrm{Var}(X)$. Also, note that since $a > 0$, $\sqrt{a^2} = |a| = a$. Therefore,

$$\begin{aligned}\mathrm{Corr}(aX+b,Y) &= \frac{\mathrm{Cov}(aX+b,Y)}{\sqrt{\mathrm{Var}(aX+b)}\sqrt{\mathrm{Var}(Y)}} = \frac{a\,\mathrm{Cov}(X,Y)}{\sqrt{a^2\,\mathrm{Var}(X)}\sqrt{\mathrm{Var}(Y)}} \\ &= \frac{\mathrm{Cov}(X,Y)}{\sqrt{\mathrm{Var}(X)}\sqrt{\mathrm{Var}(Y)}} = \mathrm{Corr}(X,Y).\end{aligned}$$

Let's make a note that if $a < 0$, then $\sqrt{a^2} = |a| = -a$, so our previous calculation would have left us with $\mathrm{Corr}(aX+b,Y) = -\mathrm{Corr}(X,Y)$, but this makes sense with our "dot product" intuition, since if $a < 0$, then aX "switches the direction" of X. ∎

Example 2.7. *Let X be a random selection of number $\{2,3,4,5\}$, equally likely to be any of the found numbers. Once X is drawn, Y is randomly drawn from the numbers $\{1,2,...,X\}$*

with equal probability of getting each number. For example, given that $X = 3$*, then* Y *is drawn from the numbers* $\{1,2,3\}$ *with* $P(Y=1 \mid X=3) = 1/3$*,* $P(Y=2 \mid X=3) = 1/3$*, and* $P(Y=3 \mid X=3) = 1/3$*. Find* $\mathrm{Corr}(X,Y)$*.*

Solution. Much of the work we've already done in Exercise 1.6. We found $\mathrm{Cov}(X,Y) = \frac{5}{8}$, $\mathbb{E}[X] = \frac{9}{4}$ and $\mathbb{E}[Y] = \frac{7}{2}$. We have left to find $\mathbb{E}[X^2]$ and $\mathbb{E}[Y^2]$. To this end,

$$\begin{aligned}\mathbb{E}[X^2] = \sum_s \sum_t s^2 \, p_{X,Y}(s,t) &= \sum_{s=2}^{5} \sum_{t=1}^{s} s^2 \, \frac{1}{4s} = \sum_{s=2}^{5} \sum_{t=1}^{s} \frac{s}{4} \\ &= \sum_{s=2}^{5} \frac{s^2}{4} = \frac{4}{4} + \frac{9}{4} + \frac{16}{4} + \frac{25}{4} = \frac{27}{2}\end{aligned}$$

and

$$\begin{aligned}\mathbb{E}[Y^2] &= \sum_s \sum_t t^2 \, p_{X,Y}(s,t) = \sum_{s=2}^{5} \sum_{t=1}^{s} t^2 \, \frac{1}{4s} = \sum_{s=2}^{5} \frac{1}{4s} \sum_{t=1}^{s} t^2 \\ &= \frac{1}{4\cdot 2}(1+4) + \frac{1}{4\cdot 3}(1+4+9) + \frac{1}{4\cdot 4}(1+4+9+16) + \frac{1}{4\cdot 5}(1+4+9+16+25) \\ &= \frac{77}{12}\end{aligned}$$

From before, we now have

$$\mathrm{Var}(X) = \frac{27}{2} - \left(\frac{7}{2}\right)^2 = \frac{5}{4}$$

and

$$\mathrm{Var}(Y) = \frac{77}{12} - \left(\frac{9}{4}\right)^2 = \frac{65}{48}$$

Therefore,

$$\mathrm{Corr}(X,Y) = \frac{\mathrm{Cov}(X,Y)}{\sqrt{\mathrm{Var}(X)}\sqrt{\mathrm{Var}(Y)}} = \frac{\frac{5}{8}}{\sqrt{(65/48)}\sqrt{(5/4)}} \approx .4804$$

■

Exercise 2.8. *Let* X *and* Y *be jointly continuous random variables with joint density*

$$f_{X,Y}(s,t) = \begin{cases} c(s^2 e^{-2t} + e^{-t}) & 0 < s < 1, 0 < t < \infty \\ 0 & \textit{otherwise} \end{cases}$$

Find $\mathrm{Corr}(X,Y)$*.*

Solution. Click here. ■

3. Solutions to Exercises

Solution to 1.3. Let X, Y, and Z be random variables and $a, b \in \mathbb{R}$ be scalars. Then,

(1) For a random variable X,

$$\operatorname{Cov}(X, X) = \mathbb{E}[X \cdot X] - \mathbb{E}[X]\mathbb{E}[X] = \mathbb{E}[X^2] - \big(\mathbb{E}[X]\big)^2 = \operatorname{Var}(X).$$

(2) For random variables X and Y,

$$\operatorname{Cov}(X, Y) = \mathbb{E}[XY] - \mathbb{E}[X]\mathbb{E}[Y] = \mathbb{E}[YX] - \mathbb{E}[Y]\mathbb{E}[X] = \operatorname{Cov}(Y, X).$$

(3) For random variables X and Y, and scalars $a, b \in \mathbb{R}$,

$$\begin{aligned}\operatorname{Cov}(aX + b, Y) &= \mathbb{E}[(aX + b)Y] - \mathbb{E}[aX + b]\mathbb{E}[Y] \\ &= a\mathbb{E}[XY] + b\mathbb{E}[Y] - \big(a\mathbb{E}[X]\mathbb{E}[Y] + b\mathbb{E}[Y]\big) \\ &= a\big(\mathbb{E}[XY] - \mathbb{E}[X]\mathbb{E}[Y]\big) + b\big(\mathbb{E}[Y] - \mathbb{E}[Y]\big) \\ &= a \operatorname{Cov}(X, Y).\end{aligned}$$

(4) For random variables X, Y, and Z, and scalar $a \in \mathbb{R}$,

$$\begin{aligned}\operatorname{Cov}(X + aY, Z) &= \mathbb{E}[(X + aY)Z] - \mathbb{E}[X + aY]\mathbb{E}[Z] \\ &= \mathbb{E}[XZ] + a\mathbb{E}[YZ] - \big(\mathbb{E}[X]\mathbb{E}[Z] + a\mathbb{E}[Y]\mathbb{E}[Z]\big) \\ &= \big(\mathbb{E}[XZ] - \mathbb{E}[X]\mathbb{E}[Z]\big) + a\big(\mathbb{E}[YZ] - \mathbb{E}[Y]\mathbb{E}[Z]\big) \\ &= \operatorname{Cov}(X, Z) + a \operatorname{Cov}(Y, Z).\end{aligned}$$

(5) If X and Y are independent random variables, then $\mathbb{E}[XY] = \mathbb{E}[X]\mathbb{E}[Y]$, so

$$\operatorname{Cov}(X, Y) = \mathbb{E}[XY] - \mathbb{E}[X]\mathbb{E}[Y] = \mathbb{E}[X]\mathbb{E}[Y] - \mathbb{E}[X]\mathbb{E}[Y] = 0.$$

■

Solution to 1.6. For the joint probability mass function, we have

$$p_{X,Y}(s, t) = P(X = s, Y = t) = P(Y = t \mid X = s)\, P(X = s).$$

Further, by the assumptions of the problem, $P(X = s) = \frac{1}{4}$ for any choice of $s \in \{2, 3, 4, 5\}$, and

$$P(Y = t \mid X = s) = \begin{cases} \frac{1}{s} & t \leq s \\ 0 & \text{otherwise} \end{cases}$$

So, we have

$$p_{X,Y}(s,t) = P(Y = t \mid X = s)\, P(X = s) = \begin{cases} \frac{1}{4s} & t \leq s \\ 0 & \text{otherwise} \end{cases}.$$

As a table, this is

X \ Y	1	2	3	4	5	$X = k$
2	$\frac{1}{8}$	$\frac{1}{8}$	0	0	0	$\frac{1}{4}$
3	$\frac{1}{12}$	$\frac{1}{12}$	$\frac{1}{12}$	0	0	$\frac{1}{4}$
4	$\frac{1}{16}$	$\frac{1}{16}$	$\frac{1}{16}$	$\frac{1}{16}$	0	$\frac{1}{4}$
5	$\frac{1}{20}$	$\frac{1}{20}$	$\frac{1}{20}$	$\frac{1}{20}$	$\frac{1}{20}$	$\frac{1}{4}$
$Y = k$	$\frac{77}{240}$	$\frac{77}{240}$	$\frac{47}{240}$	$\frac{9}{80}$	$\frac{1}{20}$	

To find $\mathrm{Cov}(X,Y) = \mathbb{E}[XY] - \mathbb{E}[X]\mathbb{E}[Y]$ we have

$$\begin{aligned} \mathbb{E}[XY] &= \sum_s \sum_t st\, p_{X,Y}(s,t) \\ &= \sum_{s=2}^{5} \sum_{t=1}^{s} st\, \frac{1}{4s} \\ &= \sum_{s=2}^{5} \sum_{t=1}^{s} \frac{t}{4} \\ &= \left(\frac{1}{4} + \frac{2}{4}\right) + \left(\frac{1}{4} + \frac{2}{4} + \frac{3}{4}\right) + \left(\frac{1}{4} + \frac{2}{4} + \frac{3}{4} + \frac{4}{4}\right) + \left(\frac{1}{4} + \frac{2}{4} + \frac{3}{4} + \frac{4}{4} + \frac{5}{4}\right) \\ &= \frac{17}{2} \end{aligned}$$

$$\mathbb{E}[X] = 2\left(\frac{1}{4}\right) + 3\left(\frac{1}{4}\right) + 4\left(\frac{1}{4}\right) + 5\left(\frac{1}{4}\right) = \frac{7}{2}$$

and

$$\mathbb{E}[Y] = 1\left(\frac{77}{240}\right) + 2\left(\frac{77}{240}\right) + 3\left(\frac{47}{240}\right) + 4\left(\frac{9}{80}\right) + 5\left(\frac{1}{20}\right) = \frac{9}{4}$$

Therefore $\mathrm{Cov}(X,Y) = \frac{17}{2} - \frac{7}{2}\cdot\frac{9}{4} = \frac{5}{8}$.

■

Solution to 2.8. Much of the work has already been done in Example 1.7. We found $\mathrm{Cov}(X,Y) = -\frac{3}{196}$, $\mathbb{E}[X] = \frac{6}{7}\cdot\frac{5}{8} = \frac{15}{28}$, and $\mathbb{E}[Y] = \frac{6}{7}\cdot\frac{13}{12} = \frac{13}{14}$. We have left to find $\mathbb{E}[X^2]$ and $\mathbb{E}[Y^2]$. To this end,

$$\mathbb{E}[X^2] = \frac{6}{7}\int_0^\infty\int_0^1 s^2(s^2e^{-2t} + e^{-t})\,dsdt = \frac{6}{7}\int_0^\infty\int_0^1 (s^4e^{-2t} + s^2e^{-t})\,dsdt$$
$$= \frac{6}{7}\int_0^\infty \left(\frac{1}{5}e^{-2t} + \frac{1}{3}e^{-t}\right)dt = \frac{6}{7}\left(\frac{1}{10} + \frac{1}{3}\right) = \frac{13}{35}.$$

and

$$\mathbb{E}[Y^2] = \frac{6}{7}\int_0^\infty\int_0^1 t^2(s^2e^{-2t} + e^{-t})\,dsdt = \frac{6}{7}\int_0^\infty\int_0^1 \left(s^2t^2e^{-2t} + t^2e^{-t}\right)dsdt$$
$$= \frac{6}{7}\int_0^\infty \left(\frac{1}{3}t^2e^{-2t} + t^2e^{-t}\right)dt = \frac{6}{7}\left(\frac{1}{12} + 2\right) = \frac{25}{14}.$$

We then find

$$\mathrm{Var}(X) = \frac{13}{35} - \left(\frac{15}{28}\right)^2 = \frac{331}{3920}$$

and

$$\mathrm{Var}(Y) = \frac{25}{14} - \left(\frac{13}{14}\right)^2 = \frac{181}{196}$$

Therefore,

$$\mathrm{Corr}(X,Y) = \frac{\mathrm{Cov}(X,Y)}{\sqrt{\mathrm{Var}(X)}\sqrt{\mathrm{Var}(Y)}} = \frac{-3/196}{\sqrt{(331/3920)}\sqrt{(181/196)}} \approx -0.0548$$

■