CHAPTER 12

Expectations

12.1. Expectation and covariance for multivariate distributions

As we discussed earlier for two random variables X and Y and a function $g(x, y)$ we can consider $g(X, Y)$ as a random variable, and therefore

$$\mathbb{E}g(X,Y) = \sum\sum_{x,y} g(x,y)p(x,y)$$

in the discrete case and

$$\mathbb{E}g(X,Y) = \iint g(x,y)f(x,y)dx\,dy$$

in the continuous case.

In particular for $g(x,y) = x + y$ we have

$$\begin{aligned}\mathbb{E}(X+Y) &= \iint (x+y)f(x,y)dx\,dy \\ &= \iint xf(x,y)dx\,dy + \iint yf(x,y)dx\,dy.\end{aligned}$$

If we now set $g(x,y) = x$, we see the first integral on the right is $\mathbb{E}X$, and similarly the second is $\mathbb{E}Y$. Therefore

$$\mathbb{E}(X+Y) = \mathbb{E}X + \mathbb{E}Y.$$

Proposition 12.1

If X and Y are two jointly continuous *independent* random variables, then for functions h and k we have

$$\mathbb{E}[h(X)k(Y)] = \mathbb{E}h(X) \cdot \mathbb{E}k(Y).$$

In particular, $\mathbb{E}(XY) = (\mathbb{E}X)(\mathbb{E}Y)$.

PROOF. By the above with $g(x, y) = h(x)k(y)$, and recalling that the joint density function factors by independence of X and Y as we saw in (11.1), we have

$$\begin{aligned}
\mathbb{E}[h(X)k(Y)] &= \iint h(x)k(y)f_{XY}(x, y)dx\,dy \\
&= \iint h(x)k(y)f_X(x)f_Y(y)dx\,dy \\
&= \int h(x)f_X(x)\int k(y)f_Y(y)dy\,dx \\
&= \int h(x)f_X(x)(\mathbb{E}k(Y))dx \\
&= \mathbb{E}h(X)\cdot\mathbb{E}k(Y).
\end{aligned}$$

□

Note that we can easily extend Proposition 12.1 to any number of *independent* random variables.

Definition 12.1

The *covariance* of two random variables X and Y is defined by

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

As with the variance, $\operatorname{Cov}(X, Y) = \mathbb{E}(XY) - (\mathbb{E}X)(\mathbb{E}Y)$. It follows that if X and Y are independent, then $\mathbb{E}(XY) = (\mathbb{E}X)(\mathbb{E}Y)$, and then $\operatorname{Cov}(X, Y) = 0$.

Proposition 12.2

Suppose X, Y and Z are random variables and a and c are constants. Then
(1) $\operatorname{Cov}(X, X) = \operatorname{Var}(X)$.
(2) if X and Y are independent, then $\operatorname{Cov}(X, Y) = 0$.
(3) $\operatorname{Cov}(X, Y) = \operatorname{Cov}(Y, X)$.
(4) $\operatorname{Cov}(aX, Y) = a\operatorname{Cov}(X, Y)$.
(5) $\operatorname{Cov}(X + c, Y) = \operatorname{Cov}(X, Y)$.
(6) $\operatorname{Cov}(X + Y, Z) = \operatorname{Cov}(X, Z) + \operatorname{Cov}(Y, Z)$.
More generally,

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

Note

$$\begin{aligned}
&\operatorname{Var}(aX+bY)\\
&\quad=\mathbb{E}[((aX+bY)-\mathbb{E}(aX+bY))^2]\\
&\quad=\mathbb{E}[(a(X-\mathbb{E}X)+b(Y-\mathbb{E}Y))^2]\\
&\quad=\mathbb{E}[a^2(X-\mathbb{E}X)^2+2ab(X-\mathbb{E}X)(Y-\mathbb{E}Y)+b^2(Y-\mathbb{E}Y)^2]\\
&\quad=a^2\operatorname{Var}X+2ab\operatorname{Cov}(X,Y)+b^2\operatorname{Var}Y.
\end{aligned}$$

We have the following corollary.

Proposition 12.3

If X and Y are independent, then

$$\operatorname{Var}(X+Y)=\operatorname{Var}X+\operatorname{Var}Y.$$

Proof. We have

$$\operatorname{Var}(X+Y)=\operatorname{Var}X+\operatorname{Var}Y+2\operatorname{Cov}(X,Y)=\operatorname{Var}X+\operatorname{Var}Y.$$

□

Example 12.1. Recall that a binomial random variable is the sum of n independent Bernoulli random variables with parameter p. Consider the *sample mean*

$$\overline{X}:=\sum_{i=1}^{n}\frac{X_i}{n},$$

where all $\{X_i\}_{i=1}^{\infty}$ are independent and have the same distribution, then $\mathbb{E}\overline{X}=\mathbb{E}X_1=p$ and $\operatorname{Var}\overline{X}=\operatorname{Var}X_1/n=p\,(1-p)$.

12.2. Conditional expectation

Recall also that in Section 11.3 we considered conditional random variables $X \mid Y=y$. We can define its expectation as follows.

Definition (Conditional expectation)

The *conditional expectation* of X given Y is defined by

$$\mathbb{E}[X\mid Y=y]=\sum_x x f_{X\mid Y=y}(x)$$

for discrete random variables X and Y, and by

$$\mathbb{E}[X\mid Y=y]=\int x f_{X\mid Y=y}(x)dx$$

for continuous random variables X and Y

Here the conditional density is defined by Equation (11.3) in Section 11.3. We can think of $\mathbb{E}[X\mid Y=y]$ is the mean value of X, when Y is fixed at y. Note that unlike the expectation of a random variable which is a number, the conditional expectation, $\mathbb{E}[X \mid Y = y]$, is a *random variable* with randomness inherited from Y, not X.

12.3. Further examples and applications

12.3.1. Expectation and variance.

Example 12.2. Suppose the joint PMF of X and Y is given by

$X \diagdown Y$	0	1
0	0.2	0.7
1	0	0.1

Find $\mathbb{E}[XY]$.

Solution: Using the formula we have

$$\begin{aligned}\mathbb{E}[XY] &= \sum_{i,j} x_i y_j p(x_i, y_j) \\ &= 0 \cdot 0 \cdot p(0,0) + 1 \cdot 0 \cdot p(1,0) + 0 \cdot 1 \cdot p(0,1) + 1 \cdot 1 \cdot p(1,1) \\ &= 0.1\end{aligned}$$

Example 12.3. Suppose X, Y are independent exponential random variables with parameter $\lambda = 1$. Set up a double integral that represents

$$\mathbb{E}\left[X^2 Y\right].$$

Solution:

Since X, Y are independent then $f_{X,Y}$ factorizes

$$f_{X,Y}(x,y) = e^{-1 \cdot x} e^{-1 \cdot y} = e^{-(x+y)}, 0 < x, y < \infty.$$

Thus

$$\mathbb{E}\left[X^2 Y\right] = \int_0^\infty \int_0^\infty x^2 y e^{-(x+y)} dy dx.$$

Example 12.4. Suppose the joint PDF of X, Y is

$$f(x,y) = \begin{cases} 10xy^2 & 0 < x < y, 0 < y < 1 \\ 0 & \text{otherwise} \end{cases}.$$

Find $\mathbb{E}XY$ and $\operatorname{Var}(Y)$.

Solution: We first draw the region (try it!) and then set up the integral

$$\begin{aligned}\mathbb{E}XY &= \int_0^1 \int_0^y xy\left(10xy^2\right)dxdy = 10\int_0^1\int_0^y x^2y^3dxdy \\ &= \frac{10}{3}\int_0^1 y^3y^3dy = \frac{10}{3}\frac{1}{7} = \frac{10}{21}.\end{aligned}$$

First note that $\operatorname{Var}(Y) = \mathbb{E}Y^2 - (\mathbb{E}Y)^2$. Then

$$\begin{aligned}\mathbb{E}Y^2 &= \int_0^1 \int_0^y y^2\left(10xy^2\right)dxdy = 10\int_0^1\int_0^y y^4xdxdy \\ &= 5\int_0^1 y^4y^2dy = \frac{5}{7}.\end{aligned}$$

and

$$\begin{aligned}\mathbb{E}Y &= \int_0^1 \int_0^y y\left(10xy^2\right)dxdy = 10\int_0^1\int_0^y y^3xdxdy \\ &= 5\int_0^1 y^3y^2dy = \frac{5}{6}.\end{aligned}$$

So that $\operatorname{Var}(Y) = \frac{5}{7} - \left(\frac{5}{6}\right)^2 = \frac{5}{252}$.

12.3.2. Correlation.

We start with the following definition.

Definition

The *correlation coefficient* of X and Y is defined by

$$\rho(X,Y) = \frac{\operatorname{Cov}(X,Y)}{\sqrt{\operatorname{Var}(X)\operatorname{Var}(Y)}}$$

In addition, we say that X and Y are

positively correlated, if $\rho(X,Y) > 0$,
negatively correlated, if $\rho(X,Y) < 0$,
uncorrelated, if $\rho(X,Y) = 0$.

Proposition (Properties of the correlation coefficient)

Suppose X and Y are random variables. Then
(1) $\rho(X,Y) = \rho(Y,X)$.
(2) If X and Y are independent, then $\rho(X,Y) = 0$. The converse is not true in general.
(3) $-1 \leqslant \rho(X,Y) \leqslant 1$.
(4) If $\rho(X,Y) = 1$, then $Y = aX + b$ for some $a > 0$.
(5) If $\rho(X,Y) = -1$, then $Y = aX + b$ for some $a < 0$.
(6) If $\rho(aX+b, cY+d) = \rho(X,Y)$ if $a, c > 0$.

Example 12.5. Suppose X, Y are random variables whose joint PDF is given by

$$f(x,y) = \begin{cases} \frac{1}{y} & 0 < y < 1, 0 < x < y \\ 0 & \text{otherwise} \end{cases}.$$

(a) Find the covariance of X and Y.
(b) Find $\text{Var}(X)$ and $\text{Var}(Y)$.
(c) Find $\rho(X,Y)$.

Solution:

(a) Recall that $\text{Cov}\,(X,Y) = \mathbb{E}XY - \mathbb{E}X\mathbb{E}Y$. So

$$\begin{aligned} \mathbb{E}XY &= \int_0^1 \int_0^y xy\frac{1}{y}dxdy = \int_0^1 \frac{y^2}{2}dy = \frac{1}{6}, \\ \mathbb{E}X &= \int_0^1 \int_0^y x\frac{1}{y}dxdy = \int_0^1 \frac{y}{2}dy = \frac{1}{4}, \\ \mathbb{E}Y &= \int_0^1 \int_0^y y\frac{1}{y}dxdy = \int_0^1 ydy = \frac{1}{2}. \end{aligned}$$

Thus

$$\begin{aligned} \text{Cov}\,(X,Y) &= \mathbb{E}XY - \mathbb{E}X\mathbb{E}Y \\ &= \frac{1}{6} - \frac{1}{4}\frac{1}{2} \\ &= \frac{1}{24}. \end{aligned}$$

(b) We have that

$$\begin{aligned} \mathbb{E}X^2 &= \int_0^1 \int_0^y x^2\frac{1}{y}dxdy = \int_0^1 \frac{y^2}{3}dy = \frac{1}{9}, \\ \mathbb{E}Y^2 &= \int_0^1 \int_0^y y^2\frac{1}{y}dxdy = \int_0^1 y^2dy = \frac{1}{3}. \end{aligned}$$

Thus recall that

$$\begin{aligned} \text{Var}\,(X) &= \mathbb{E}X^2 - (\mathbb{E}X)^2 \\ &= \frac{1}{9} - \left(\frac{1}{4}\right)^2 = \frac{7}{144}. \end{aligned}$$

Also

$$\begin{aligned} \text{Var}\,(Y) &= \mathbb{E}Y^2 - (\mathbb{E}Y)^2 \\ &= \frac{1}{3} - \left(\frac{1}{2}\right)^2 = \frac{1}{12}. \end{aligned}$$

(c)

$$\rho\,(X,Y) = \frac{\text{Cov}\,(X,Y)}{\sqrt{\text{Var}\,(X)\,\text{Var}\,(Y)}} = \frac{\frac{1}{24}}{\sqrt{\left(\frac{7}{144}\right)\left(\frac{1}{12}\right)}} \approx 0.6547.$$

12.3.3. Conditional expectation: examples. We start with properties of conditional expectations which we introduced in Definition 12.2 for discrete and continuous random variables. We skip the proof of this as it is very similar to the case of (unconditional) expectation.

Proposition 12.4

For the conditional expectation of X given $Y = y$ it holds that

(i) for any $a, b \in \mathbb{R}$, $\mathbb{E}[aX + b \mid Y = y] = a\mathbb{E}[X \mid Y = y] + b$.

(ii) $\operatorname{Var}(X \mid Y = y) = \mathbb{E}[X^2 \mid Y = y] - (\mathbb{E}[X \mid Y = y])^2$.

Example 12.6. Let X and Y be random variables with the joint PDF

$$f_{XY}(x, y) = \begin{cases} \frac{1}{18} e^{-\frac{x+y}{6}} & \text{if } 0 < y < x, \\ 0 & \text{otherwise.} \end{cases}$$

In order to find $\operatorname{Var}(X \mid Y = 2)$, we need to compute the conditional PDF of X given $Y = 2$, i.e.

$$f_{X|Y=2}(x \mid 2) = \frac{f_{XY}(x, 2)}{f_Y(2)}.$$

To this purpose, we compute first the marginal of Y.

$$f_Y(y) = \int_y^\infty \frac{1}{18} e^{-\frac{x+y}{6}}\, dx = \frac{1}{3} e^{-\frac{y}{6}} \left. -e^{-\frac{y}{6}} \right|_y^\infty = \frac{1}{3} e^{-\frac{y}{3}} \qquad \text{for } y \geqslant 0.$$

Then we have

$$f_{X|Y=2}(x \mid 2) = \begin{cases} \frac{1}{6} e^{\frac{2-x}{6}} & \text{if } x > 2, \\ 0 & \text{otherwise.} \end{cases}$$

Now it only remains to find $\mathbb{E}[X^2 \mid Y = 2]$ and $\mathbb{E}[X \mid Y = 2]$. Applying integration by parts twice we have

$$\mathbb{E}[X^2 \mid Y = 2] = \int_2^\infty \frac{x^2}{6} e^{\frac{2-x}{6}}\, dx = -\left. x^2 e^{\frac{2-x}{6}} \right|_2^\infty - \left. 12 x e^{\frac{2-x}{6}} \right|_2^\infty - \left. 12\, 6 e^{\frac{2-x}{6}} \right|_2^\infty = 4 + 24 + 72 = 100.$$

On the other hand, again applying integration by parts we get

$$\mathbb{E}[X \mid Y = 2] = \int_2^\infty \frac{x}{6} e^{\frac{2-x}{6}}\, dx = -\left. x e^{-\frac{x-2}{6}} \right|_2^\infty - \left. 6 e^{-\frac{x-2}{6}} \right|_2^\infty = 2 + 6 = 8.$$

Finally, we obtain $\operatorname{Var}(X \mid Y = 2) = 100 - 8^2 = 36$.

12.4. Exercises

Exercise 12.1. Suppose the joint distribution for X and Y is given by the joint probability mass function shown below:

$Y \backslash X$	0	1
0	0	0.3
1	0.5	0.2

(a) Find the covariance of X and Y.
(b) Find $\operatorname{Var}(X)$ and $\operatorname{Var}(Y)$.
(c) Find $\rho(X, Y)$.

Exercise 12.2. Let X and Y be random variables whose joint probability density function is given by

$$f(x, y) = \begin{cases} x + y & 0 < x < 1, 0 < y < 1 \\ 0 & \text{otherwise.} \end{cases}$$

(a) Find the covariance of X and Y.
(b) Find $\operatorname{Var}(X)$ and $\operatorname{Var}(Y)$.
(c) Find $\rho(X, Y)$.

Exercise 12.3. Let X be normally distributed with mean 1 and variance 9. Let Y be exponentially distributed with $\lambda = 2$. Suppose X and Y are independent. Find $\mathbb{E}\left[(X-1)^2 Y\right]$. (Hint: Use properties of expectations.)

Exercise* 12.1. Prove Proposition 12.2.

Exercise* 12.2. Show that if random variables X and Y are uncorrelated, then $\operatorname{Var}(X+Y) = \operatorname{Var}(X) + \operatorname{Var}(Y)$. Note that this is a more general statement than Proposition 12.3 since independent variables are uncorrelated.

12.5. Selected solutions

Solution to Exercise 12.1(A): First let us find the marginal distributions.

$Y\backslash X$	0	1	
0	0	0.3	0.3
1	0.5	0.2	0.7
	0.5	0.5	

Then

$$\begin{aligned}\mathbb{E}XY &= (0\cdot 0)\cdot 0 + (0\cdot 1)\cdot 0.5 + (1\cdot 0)\cdot 0.3 + (1\cdot 1)\cdot 0.2 = 0.2\\ \mathbb{E}X &= 0\cdot 0.5 + 1\cdot 0.5 = 0.5\\ \mathbb{E}Y &= 0\cdot 0.3 + 1\cdot 0.7 = 0.7.\end{aligned}$$

Solution to Exercise 12.1(B): First we need

$$\begin{aligned}\mathbb{E}X^2 &= 0^2 0.5 + 1^2 0.5 = 0.5\\ \mathbb{E}Y^2 &= 0^2 0.3 + 1^2 0.7 = 0.7\end{aligned}$$

Therefore

$$\begin{aligned}\operatorname{Var}(X) &= \mathbb{E}X^2 - (\mathbb{E}X)^2 = 0.5 - (0.5)^2 = 0.25,\\ \operatorname{Var}(Y) &= \mathbb{E}Y^2 - (\mathbb{E}Y)^2 = 0.7 - (0.7)^2 = 0.21,\end{aligned}$$

Solution to Exercise 12.1(C):

$$\rho(X,Y) = \frac{\operatorname{Cov}(X,Y)}{\sqrt{\operatorname{Var}(X)\cdot\operatorname{Var}(Y)}} \approx -0.6547.$$

Solution to Exercise 12.2(A): We need to find $\mathbb{E}[XY], \mathbb{E}X$, and $\mathbb{E}Y$.

$$\begin{aligned}\mathbb{E}[XY] &= \int_0^1\int_0^1 xy\,(x+y)\,dydx = \int_0^1\left(\frac{x^2}{2}+\frac{x}{3}\right)dx = \frac{1}{3},\\ \mathbb{E}X &= \int_0^1\int_0^1 x(x+y)dydx = \int_0^1 (x^2+\frac{x}{2})dx = \frac{7}{12}\\ \mathbb{E}Y &= \frac{7}{12},\ \text{by symmetry with the } \mathbb{E}X \text{ case.}\end{aligned}$$

Therefore

$$\operatorname{Cov}(X,Y) = \mathbb{E}[XY] - \mathbb{E}[X]\,\mathbb{E}[Y] = \frac{1}{3} - \left(\frac{7}{12}\right)^2 = -\frac{1}{144}.$$

Solution to Exercise 12.2(B): We need $\mathbb{E}X^2$ and $\mathbb{E}Y^2$,

$$\mathbb{E}X^2 = \int_0^1 \int_0^1 x^2(x+y)dydx = \int_0^1 (x^3 + \frac{x^2}{2})dx = \frac{5}{12}$$

so we know that $\mathbb{E}Y^2 = \frac{5}{12}$ by symmetry. Therefore

$$\operatorname{Var}(X) = \operatorname{Var}(Y) = \frac{5}{12} - \left(\frac{7}{12}\right)^2 = \frac{11}{144}.$$

Solution to Exercise 12.2(C):

$$\rho(X,Y) = \frac{\operatorname{Cov}(X,Y)}{\sqrt{\operatorname{Var}(X) \cdot \operatorname{Var}(Y)}} = -\frac{1}{11}.$$

Solution to Exercise 12.3: Since X, Y are independent then

$$\begin{aligned}\mathbb{E}\left[(X-1)^2 Y\right] &= \mathbb{E}\left[(X-1)^2\right]\mathbb{E}\left[Y\right] \\ &= \operatorname{Var}(X)\frac{1}{\lambda} = 9/2 = 4.5\end{aligned}$$