Chapter 4 - Part 2

Joint Probability

Joint Probability Distribution - Example

There is a bag with 2 blue and 3 red marbles. Two marbles are drawn without replacement.

$$S = \{RR, RB, BR, BB\}$$

Let

- X : number of blue marbles in both draws
- Y : number of red marbles in last draw

The assignment

$$S = \{ \underbrace{RR}_{x=0,y=1}, \underbrace{RB}_{x=1,y=0}, \underbrace{BR}_{x=1,y=1}, \underbrace{BB}_{x=2,y=0} \}$$

gives the joint distribution of X and Y. Mostly, we will show this as:

y \ x	0	1	2
0	no way	R B	B B
1	R R	B R	no way

Example (cont'd)

- 2 blue and 3 red marbles

$$P(RR) = \frac{3}{5} \cdot \frac{2}{4} = \frac{3}{10}$$
$$P(RB) = \frac{3}{5} \cdot \frac{2}{4} = \frac{3}{10}$$
$$P(BR) = \frac{2}{5} \cdot \frac{3}{4} = \frac{3}{10}$$
$$P(BB) = \frac{2}{5} \cdot \frac{1}{4} = \frac{1}{10}$$

y \ x	0	1	2
0	no way	R B	B B
1	R R	B R	no way

The joint probability distribution function of X and Y is given by

y \ x	0	1	2	$\sum$
0	0	0.3	0.1	0.4
1	0.3	0.3	0	0.6
$\sum$	0.3	0.6	0.1	1

Example (cont'd)

The joint probability distribution function of X and Y is given by

y \ x	0	1	2	$\sum$
0	0	0.3	0.1	0.4
1	0.3	0.3	0	0.6
$\sum$	0.3	0.6	0.1	1

$p(0,0) = 0, \quad p(0,1) = 0.3, \quad p(1,0) = 0.3, \quad p(1,1) = 0.3, \quad p(2,0) = 0.1, \quad p(2,1) = 0$

<u>Another notation:</u>

$$P(X = 2, Y = 0) = 0.1, \quad P(Y = 0) = 0.4,$$

$$P(X > 1, Y = 0) = 0.1, \quad P(X < 2, Y < 1) = 0.3$$

Joint Probability Distribution - Definition

Given two random variables X and Y that are defined on the same sample space, the joint probability distribution function of them is given by

$$p(a, b) = P(X = a, Y = b) = P(X = a) \cap P(Y = b)$$

Note: This can also be defined for $n > 2$ many random variables.

Covariance - Definition

Let X and Y be two discrete random variables on the same sample space. Let the values that X can take be

$$x_1, x_2, \ldots, x_n$$

and the values that Y can take be

$$y_1, y_2, \ldots, y_m.$$

We define the covariace of X and Y as

$$\sigma_{XY} = \mathrm{Cov}(X, Y) = \sum_{i=1}^{n} \sum_{j=1}^{m} p(x_i, y_j)(x_i - E(X))(y_j - E(Y))$$

Example

Consider again

y \ x	0	1	2	$\sum$
0	0	0.3	0.1	0.4
1	0.3	0.3	0	0.6
$\sum$	0.3	0.6	0.1	1

First,

$$E(X) = 0 \cdot 0.3 + 1 \cdot 0.6 + 2 \cdot 0.1 = 0.8, \qquad E(Y) = 0 \cdot 0.4 + 1 \cdot 0.6 = 0.6$$

Hence

$$\begin{aligned} \mathrm{Cov}(X, Y) = \; & \underbrace{p(0,0)}_{0}(0-0.8)(0-0.6) + \underbrace{p(0,1)}_{0.3}(0-0.8)(1-0.6) \\ & + \underbrace{p(1,0)}_{0.3}(1-0.8)(0-0.6) + \underbrace{p(1,1)}_{0.3}(1-0.8)(1-0.6) \\ & + \underbrace{p(2,0)}_{0.1}(2-0.8)(0-0.6) + \underbrace{p(2,1)}_{0}(2-0.8)(1-0.6) \end{aligned}$$

Example (cont'd)

$$\begin{aligned}
\mathrm{Cov}(X,Y) = \quad & \underbrace{p(0,0)}_{0}(0-0.8)(0-0.6) + \underbrace{p(0,1)}_{0.3}(0-0.8)(1-0.6) \\
& + \underbrace{p(1,0)}_{0.3}(1-0.8)(0-0.6) + \underbrace{p(1,1)}_{0.3}(1-0.8)(1-0.6) \\
& + \underbrace{p(2,0)}_{0.1}(2-0.8)(0-0.6) + \underbrace{p(2,1)}_{0}(2-0.8)(1-0.6) \\
= \quad & (0.3)\cdot(-0.8)\cdot(0.4) + (0.3)\cdot(0.2)\cdot(-0.6) \\
& + (0.3)\cdot(0.2)\cdot(0.4) + (0.1)\cdot(1.2)\cdot(-0.6) \\
= \quad & -0.18
\end{aligned}$$

Random Variable vs Dataset

The **covariance** of the joint distribution

y \ x	0	1	2	$\sum$
0	0	0.3	0.1	0.4
1	0.3	0.3	0	0.6
$\sum$	0.3	0.6	0.1	1

is the same as the **population covariance** of a dataset of the form

$$\{\underbrace{(0,1),(0,1),\ldots,(0,1)}_{30\%},\underbrace{(1,0),(1,0),\ldots,(1,0)}_{30\%},\underbrace{(1,1),(1,1),\ldots,(1,1)}_{30\%},\underbrace{(2,0),\ldots,(2,0)}_{10\%}\}$$

An example of such dataset:

$$\{(0,1),(0,1),(0,1),(1,0),(1,0),(1,0),(1,1),(1,1),(1,1),(2,0)\}$$

Multivariable Function of Random Variables - Example

y \ x	0	1	2	$\sum$
0	0	0.3	0.1	0.4
1	0.3	0.3	0	0.6
$\sum$	0.3	0.6	0.1	1

The function $t = f(x, y) = 2x + 4y + 3$ gives a random variable with outcomes

(x,y)	(0,0)	(0,1)	(1,0)	(1,1)	(2,0)	(2,1)
t=f(x,y)	3	7	5	9	7	11

$$P(T = 3) = P(X = 0, Y = 0) = 0,$$

$$P(T = 5) = P(X = 1, Y = 0) = 0.3,$$

$$P(T = 7) = P(X = 0, Y = 1) + P(X = 2, Y = 0) = 0.3 + 0.1 = 0.4,$$

$$P(T = 9) = P(X = 1, Y = 1) = 0.3,$$

$$P(T = 11) = P(X = 2, Y = 1) = 0$$

Example (cont'd)

y \ x	0	1	2	$\sum$
0	0	0.3	0.1	0.4
1	0.3	0.3	0	0.6
$\sum$	0.3	0.6	0.1	1

$$T = 2X + 4Y + 3$$

t	5	7	9
p(t)	0.3	0.4	0.3

$$E(T) = 5 \cdot 0.3 + 7 \cdot 0.4 + 9 \cdot 0.3 = 7$$

Also notice that

$$2E(X) + 4E(Y) + 3 = 2 \cdot 0.8 + 4 \cdot 0.6 + 3 = 7$$

Proposition

Let $a, b, c \in \mathbb{R}$ be constants and X, Y be random variables. Then,

$$E(aX + bY + c) = aE(X) + bE(Y) + c$$

Example

Roll a pair of dice (one black one white). Let X be the random variable representing sum of two dice. Let's find $E(X)$.
Call X_b and X_w for the random variables representing the outcomes of black and white die, respectively. Then

$$X = X_b + X_w$$

We already seen that $E(X_b) = E(X_w) = 3.5$. Therefore

$$E(X) = E(X_b + X_w) = E(X_b) + E(X_w) = 3.5 + 3.5 = 7$$

Proposition

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

Proof.

It follows from the definition of covariance:

$$\mathrm{Cov}(X, Y) = \sum_{i=1}^{n}\sum_{j=1}^{m} p(x_i, y_j)(x_i - E(X))(y_j - E(Y)) = E\left((X - E(X))(Y - E(Y))\right).$$

$\square$

Proposition

$$\mathrm{Cov}(X, X) = \mathrm{Var}(X)$$

Proof.

$$\mathrm{Cov}(X, X) = E[(X - E(X))(X - E(X))] = E\left[(X - E(X))^2\right] = \mathrm{Var}(X)$$

□

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

Proof.

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

□

Proposition

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

Proof.

$$\begin{aligned}
\mathrm{Cov}(X, Y) &= E\left[(X - E(X))(Y - E(Y))\right] \\
&= E\left[XY - XE(Y) - YE(X) + E(X)E(Y)\right] \\
&= E(XY) - E(XE(Y)) - E(YE(X)) + E(E(X)E(Y)) \\
&= E(XY) - E(X)E(Y) - E(X)E(Y) + E(X)E(Y) \\
&= E(XY) - E(X)E(Y)
\end{aligned}$$

$\square$

Example

Let's calculate again the covariance of the following

y \ x	0	1	2	$\sum$
0	0	0.3	0.1	0.4
1	0.3	0.3	0	0.6
$\sum$	0.3	0.6	0.1	1

using $\text{Cov}(X, Y) = E(XY) - E(X)E(Y)$.
Call $T = XY$. Then possible outcomes of T are

(x,y)	(0,1)	(1,0)	(1,1)	(2,0)
t=xy	0	0	1	0

$$P(T = 0) = 0.3 + 0.3 + 0.1 = 0.7, \quad P(T = 1) = 0.3$$

Example (cont'd)

y \ x	0	1	2	$\sum$
0	0	0.3	0.1	0.4
1	0.3	0.3	0	0.6
$\sum$	0.3	0.6	0.1	1

$$T = XY$$

t	0	1
p(t)	0.7	0.3

$$E(XY) = E(T) = 0.3$$

Hence

$$\mathrm{Cov}(X, Y) = E(XY) - E(X)E(Y) = 0.3 - 0.8 \cdot 0.6 = -0.18$$

Independence

Two random variables X and Y are called independent if

$$P(X = a, Y = b) = P(X = a) \cdot P(Y = b)$$

for all possible outcome a of X and b of Y.

Otherwise, they are called dependent.

Example

y \ x	0	1	2	$\sum$
0	0	0.3	0.1	0.4
1	0.3	0.3	0	0.6
$\sum$	0.3	0.6	0.1	1

X and Y are not independent since for example

$$P(X = 0, Y = 1) = 0.3 \neq 0.3 \cdot 0.6 = P(X = 0) \cdot P(Y = 1)$$

Example

y \ x	1	2	$\sum$
1	0.14	0.56	0.7
2	0.06	0.24	0.3
$\sum$	0.2	0.8	1

X and Y are independent since

$$P(X = 1, Y = 1) = 0.14 = 0.2 \cdot 0.7 = P(X = 1) \cdot P(Y = 1)$$
$$P(X = 1, Y = 2) = 0.06 = 0.2 \cdot 0.3 = P(X = 1) \cdot P(Y = 2)$$
$$P(X = 2, Y = 1) = 0.56 = 0.8 \cdot 0.7 = P(X = 2) \cdot P(Y = 1)$$
$$P(X = 2, Y = 2) = 0.24 = 0.8 \cdot 0.3 = P(X = 2) \cdot P(Y = 2)$$

Example

Roll a pair of dice (one black one white). Call X_b and X_w for the random variables representing the outcomes of black and white die, respectively.

X_b and X_w are independent since

$$P(X_b = i, X_w = j) = \frac{1}{36} = \frac{1}{6} \cdot \frac{1}{6} = P(X_b = i) \cdot P(X_w = j)$$

for all $i, j \in \{1, 2, 3, 4, 5, 6\}$.

Covariance of Independent Random Variables - Example

Roll a pair of dice (one black one white). Call X_b and X_w for the random variables representing the outcomes of black and white die, respectively.

$$E(X_bX_w) = \sum_{j=1}^{6}\sum_{i=1}^{6} i\cdot j\cdot \underbrace{P(x_b = i, x_w = j)}_{\frac{1}{36}} = \frac{1}{36}\sum_{j=1}^{6}\sum_{i=1}^{6} i\cdot j$$

$$= \frac{1}{36}(1\cdot 1 + 1\cdot 2 + \ldots + 1\cdot 6 + 2\cdot 1 + 2\cdot 2 + \ldots + 2\cdot 6 + 3\cdot 1 + \ldots + \ldots + 6\cdot 6)$$

$$= \frac{1}{36}(1\cdot(1+2+\ldots+6) + 2\cdot(1+2+\ldots+6) + \ldots + 6\cdot(1+2+\ldots+6))$$

$$= \frac{1}{36}\cdot\underbrace{(1+2+\ldots+6)}_{21}\cdot\underbrace{(1+2+\ldots+6)}_{21} = \left(\frac{21}{6}\right)^2$$

$$\implies \mathrm{Cov}(X_b, X_w) = E(X_bX_w) - E(X_b)E(X_w) = \left(\frac{21}{6}\right)^2 - \frac{21}{6}\cdot\frac{21}{6} = 0.$$

Covariance of Independent Random Variables - Theorem

If X and Y are independent, then

$$\mathrm{Cov}(X, Y) = 0.$$

Proof.

If X and Y are independent,

$$\begin{aligned} E(XY) &= \sum_j \sum_i x_i \cdot y_j \cdot p(x_i, y_j) \\ &= \sum_j \sum_i x_i \cdot y_j \cdot p(x_i)p(y_j) \\ &= \left(\sum_i x_i \cdot p(x_i) \right) \left(\sum_j y_j \cdot p(y_j) \right) \\ &= E(X)E(Y) \end{aligned}$$

$$\implies \mathrm{Cov}(X, Y) = E(XY) - E(X)E(Y) = 0$$

Covariance of Independent Random Variables - Theorem

If X and Y are independent, then

$$\mathrm{Cov}(X, Y) = 0.$$

Remark: Converse of this theorem is **not true!** $\mathrm{Cov}(X, Y) = 0$ does not imply X and Y being independent.

Example

y \ x	0	3	5	$\sum$
0	0.1	0.1	0.1	0.3
1	0.1	0.1	0.2	0.4
2	0.1	0.1	0.1	0.3
$\sum$	0.3	0.3	0.4	1

$$E(X) = 0 \cdot 0.3 + 3 \cdot 0.3 + 5 \cdot 0.4 = 2.9$$
$$E(Y) = 0 \cdot 0.3 + 1 \cdot 0.4 + 2 \cdot 0.3 = 1$$
$$E(XY) = 3 \cdot 0.1 + 5 \cdot 0.2 + 6 \cdot 0.1 + 10 \cdot 0.1 = 2.9$$
$$\implies \operatorname{Cov}(X, Y) = E(XY) - E(X)E(Y) = 0$$

But X and Y are not independent since for example

$$P(X = 0, Y = 0) = 0.1 \neq 0.3 \cdot 0.3 = P(X = 0) \cdot P(Y = 0)$$

Independent vs Uncorrelated

Two random variables X and Y are called uncorrelated if

$$\mathrm{Cov}(X, Y) = 0.$$

- independent $\implies$ uncorrelated
- uncorrelated $\not\Longrightarrow$ independent

Proposition

Let $a, b, c \in \mathbb{R}$ be constants and X, Y be random variables. Then,

$$\mathrm{Var}(aX + bY + c) = a^2\mathrm{Var}(X) + b^2\mathrm{Var}(Y) + 2ab\mathrm{Cov}(X, Y).$$

In particular, if $\mathrm{Cov}(X, Y) = 0$ (for example if X and Y are independent),

$$\mathrm{Var}(aX + bY + c) = a^2\mathrm{Var}(X) + b^2\mathrm{Var}(Y).$$

Taking $a = b = 1, c = 0$:

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

Proposition

More generally, if $X_1, X_2, \ldots, X_n$ are random variables such that

$$\mathrm{Cov}(X_i, X_j) = 0$$

for all i, j, then

$$\mathrm{Var}(X_1 + X_2 + \ldots + X_n) = \mathrm{Var}(X_1) + \mathrm{Var}(X_2) + \ldots + \mathrm{Var}(X_n).$$

Threfore, if $X_1, X_2, \ldots, X_n$ are pairwise independent random variables

$$\mathrm{Var}(X_1 + X_2 + \ldots + X_n) = \mathrm{Var}(X_1) + \mathrm{Var}(X_2) + \ldots + \mathrm{Var}(X_n).$$