ST221 Introduction to Statistics

Dr Rafael Moral

Dept of Mathematics and Statistics
Maynooth University

rafael.deandrademoral@mu.ie

4 Bivariate Discrete Random Variables

Many times we may wish to study the joint behaviour of two or more random variables

For this seciton, we will look at the joint distribution of two discrete random variables

4.1 Joint Distributions

A family has 3 children. Assume that the probability of a boy and of a girl is 0.5. We have seen the sample space for this experiment before.

Let X be the random variable defined as the "total number of boys" and Y equal to 1 if the firstborn is female and 0 otherwise.

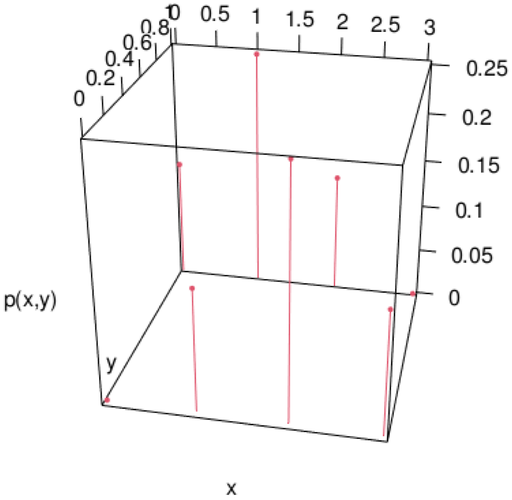
We have

Outcome	Probability	X	Y
bbb	1/8	3	0
bbg	1/8	2	0
bgb	1/8	2	0
gbb	1/8	2	1
ggb	1/8	1	1
gbg	1/8	1	1
bgg	1/8	1	0
ggg	1/8	0	1

We can write the joint probability mass function of X and Y $\mathrm{P}(X = x, Y = y)$:

Y	X			
	0	1	2	3
0	0	1/8	2/8	1/8
1	1/8	2/8	1/8	0

We can graph it using a 3-d plot.



4.2 Marginal and Conditional Distributions

The marginal distributions of X and Y can be obtained from the joint distribution:

$$P(X = x) = \sum_y P(X = x, Y = y)$$

$$P(Y = y) = \sum_x P(X = x, Y = y)$$

e.g.

$$\begin{aligned} P(X = 0) &= \sum_x P(X = 0, Y = y) \\ &= P(X = 0, Y = 0) + P(X = 0, Y = 1) \\ &= 0 + 1/8 \\ &= 1/8 \end{aligned}$$

Y	X = 0	1	2	3	$P(Y = y)$
0	0	1/8	2/8	1/8	1/2
1	1/8	2/8	1/8	0	1/2
$P(X = x)$	1/8	3/8	3/8	1/8	

The expected values for X and Y, as well as their variances are computed the same way we have learned before, using their marginal distributions.

The conditional distribution of $X|Y = y$ is

$$P(X = x|Y = y) = \frac{P(X = x, Y = y)}{P(Y = y)}$$

e.g.

$$\begin{aligned}
P(X = 0|Y = 0) &= \frac{P(X = 0, Y = 0)}{P(Y = 0)} = \frac{0}{1/2} = 0 \\
P(X = 1|Y = 0) &= \frac{P(X = 1, Y = 0)}{P(Y = 0)} = \frac{1/8}{1/2} = 1/4 \\
P(X = 2|Y = 0) &= \frac{P(X = 2, Y = 0)}{P(Y = 0)} = \frac{2/8}{1/2} = 1/2 \\
P(X = 3|Y = 0) &= \frac{P(X = 3, Y = 0)}{P(Y = 0)} = \frac{1/8}{1/2} = 1/4
\end{aligned}$$

1. Find the conditional distribution of $X|Y = 1$.

2. Find the conditional distribution of $Y|X = 2$.

3. Find the conditional distribution of $Y|X = 3$.

4.3 Covariance between two discrete random variables

The covariance between random variables X and Y is defined as

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

The covariance can be any real value.

In our example,

$$
\begin{aligned}
\mathrm{E}[XY] &= \sum_x \sum_y xy\mathrm{P}(X = x, Y = y) \\
&= 0 \times 0 \times \mathrm{P}(X = 0, Y = 0) + 0 \times 1 \times \mathrm{P}(X = 0, Y = 1) + \\
&\quad 1 \times 0 \times \mathrm{P}(X = 1, Y = 0) + 1 \times 1 \times \mathrm{P}(X = 1, Y = 1) + \\
&\quad 2 \times 0 \times \mathrm{P}(X = 2, Y = 0) + 2 \times 1 \times \mathrm{P}(X = 2, Y = 1) + \\
&\quad 3 \times 0 \times \mathrm{P}(X = 3, Y = 0) + 3 \times 1 \times \mathrm{P}(X = 3, Y = 1) \\
&= 0 + 0 + 0 + 2/8 + 0 + 2/8 + 0 + 0 = 1/2
\end{aligned}
$$

So we have

$$\text{Cov}(X, Y) = \frac{1}{2} - \frac{3}{2} \times \frac{1}{2} = -\frac{1}{4}$$

Exercise: Take the joint p.m.f. of X and Y given by

Y	X		
	0	1	2
1	3/20	3/20	2/20
2	1/20	1/20	2/20
3	4/20	1/20	3/20

Compute $\text{Cov}(X, Y)$.

4.4 Correlation between two discrete random variables

The correlation coefficient ρ_{XY} is given by

$$\rho_{XY} = \frac{\text{Cov}(X, Y)}{\sqrt{\text{Var}(X)\text{Var}(Y)}}$$

We can show that $-1 \leq \rho_{XY} \leq 1$.

In our example:

$$\rho_{XY} = \frac{-1/4}{\sqrt{3/4 \times 1/4}} = -\frac{\sqrt{3}}{3} = -0.5774$$

If X and Y are independent, then $\text{Cov}(X, Y) = 0$ (and hence $\rho_{XY} = 0$ as well), however a covariance of zero **does not** imply independence.

4.5 Independence

X and Y are independent if

$$P(X = x, Y = y) = P(X = x)P(Y = y)$$

Are X and Y independent in the children example?

Exercise: Check if X and Y are independent if their joint p.m.f. is given by:

Y	X = 0	X = 1
0	1/16	3/16
1	2/16	3/8
2	1/16	3/16

Exercise: Check if X and Y are independent if their joint p.m.f. is given by:

Y	X: -1	0	1
-1	$1/8$	$1/8$	$1/8$
0	$1/8$	0	$1/8$
1	$1/8$	$1/8$	$1/8$