STAT 2857A – Lecture 21 Examples and Exercises

Solutions

Example 21.1

a) By definition

$$\begin{aligned} E(XY) &= \sum_x \sum_y xyp(x,y) \\ &= 0(0)p_{00} + 0(1)p_{01} + 1(0)p_{10} + 1(1)p_{11} \\ &= p_{11} \end{aligned}$$

b) First note that the marginal pmf of X is given by

$$\begin{aligned} P(X=0) &= p_{00} + p_{10} \\ P(X=1) &= p_{10} + p_{11} \end{aligned}$$

which implies that

$$X \sim \text{Bernoulli}(p_{10} + p_{11})$$

or

$$X \sim \text{Binomial}(1, p_{10} + p_{11}).$$

Then

$$\begin{aligned} E(X) &= p_{10} + p_{11} = \mu_x \\ V(X) &= (p_{10} + p_{11})(1 - p_{10} - p_{11}) = \mu_x(1 - \mu_x). \end{aligned}$$

Similarly

$$\begin{aligned} E(Y) &= p_{01} + p_{11} = \mu_y \\ V(Y) &= (p_{01} + p_{11})(1 - p_{01} - p_{11}) = \mu_y(1 - \mu_y). \end{aligned}$$

Then

$$
\begin{aligned}
\operatorname{Cov}(X,Y) &= E[(X-\mu_x)(Y-\mu_y)] \\
&= \sum_{x=0}^{1}\sum_{y=0}^{1}(x-\mu_x)(y-\mu_y) \\
&= (-\mu_x)(-\mu_y)p_{00} + (1-\mu_x)(-\mu_y)p_{10} + (-\mu_x)(1-\mu_y)p_{01} + (1-\mu_x)(1-\mu_y)p_{11} \\
&= \ldots \\
&= p_{11} - (p_{10}+p_{11})(p_{01}+p_{11}).
\end{aligned}
$$

Alternatively, we can apply the shortcut formula:

$$
\begin{aligned}
\operatorname{Cov}(X,Y) &= E(XY) - E(X)E(Y) \\
&= p_{11} - (p_{10}+p_{11})(p_{01}+p_{11}).
\end{aligned}
$$

To compute the correlation, we divide the covariance by the standard deviation of both X and Y

$$
\begin{aligned}
\operatorname{Corr}(X,Y) &= \frac{\operatorname{Cov}(X,Y)}{\sqrt{V(X)V(Y)}} \\
&= \frac{p_{11} - (p_{10}+p_{11})(p_{01}+p_{11})}{\sqrt{(p_{10}+p_{11})(1-p_{10}-p_{11})\cdot(p_{01}+p_{11})(1-p_{01}-p_{11})}} \\
&= \frac{p_{11} - (p_{10}+p_{11})(p_{01}+p_{11})}{\sqrt{(p_{10}+p_{11})(p_{00}+p_{01})\cdot(p_{01}+p_{11})(p_{00}+p_{10})}}.
\end{aligned}
$$

Note that if $p_{01} = p_{10} = 0$ then $p_{00} = 1 - p_{11}$ and

$$
\begin{aligned}
\operatorname{Corr}(X,Y) &= \frac{p_{11} - p_{11}^2}{\sqrt{p_{11}(1-p_{11})\cdot p_{11}(1-p_{11})}} \\
&= \frac{p_{11}(1-p_{11})}{p_{11}(1-p_{11})} \\
&= 1.
\end{aligned}
$$

Similarly if $p_{00} = p_{11} = 0$ then $\operatorname{Corr}(X,Y) = -1$.

c) The property for sums of random variables tells us that the expected value is

$$
\begin{aligned}
E(Z) &= E(2X+4Y) \\
&= 2E(X) + 4E(Y) \\
&= 2(p_{10}+p_{11}) + 4(p_{01}+_{11}) \\
&= 2p_{10} + 4p_{01} + 6p_{11}.
\end{aligned}
$$

The variance is

$$\begin{aligned}V(Z) =& 4V(X) + 8\text{Cov}(X,Y) + 16V(Y)\\ =& 4[(p_{10} + p_{11})(p_{00} + p_{01})] +\\ & 8[p_{11} - (p_{10} + p_{11})(p_{01} + p_{11})] +\\ & 16[(p_{01} + p_{11})(p_{00} + p_{10})]\end{aligned}$$

d) The random variables X and Y are independent if $p(x,y) = p(x)p(y)$ for all X and Y. Let $p_x = P(X = 1)$ and $p_y = P(Y = 1)$. Then X and Y are independent if

$$\begin{aligned}p(0,0) &= p_{00} = (1 - p_x)(1 - p_y)\\ p(1,0) &= p_{10} = p_x(1 - p_y)\\ p(0,1) &= p_{01} = p_y(1 - p_x)\\ p(1,1) &= p_{11} = p_x p_y.\end{aligned}$$

In this case,

$$\begin{aligned}\text{Cov}(X,Y) &= p_{11} - (p_{10} + p_{11})(p_{01} + p_{11})\\ &= p_x p_y - [p_x(1 - p_y) + p_x p_y][p_y(1 - p_x) + p_x p_y]\\ &= p_x p_y - [p_x - p_x p_y + p_x p_y][p_y - p_x p_y + p_x p_y]\\ &= p_x p_y - p_x p_y\\ &= 0.\end{aligned}$$

Hence, if X and Y are independent then $\text{Cov}(X,Y) = 0$ and $E(XY) = E(X)E(Y)$.

This is a special case of the property of the expectation of products of independent random variables.

Note that the opposite is not generally true. It is not generally the case that if $\text{Cov}(X,Y) = 0$ then X and Y are independent.

Example 21.2

The order is

- D: $\text{Corr}(X,Y) = .00$
- C: $\text{Corr}(X,Y) = .29$
- B: $\text{Corr}(X,Y) = .90$
- A: $\text{Corr}(X,Y) = 1.00$

Exercise 21.1

a) The possible values of S are 2, 3, 4, 5, and 6, and the possible values of D are 0, 1, and 2. The pmf is

	D		
S	0	1	2
2	1/9	0	0
3	0	2/9	0
4	1/9	0	2/9
5	0	2/9	0
6	1/9	0	0

b) The marginal pmf of S has values

$$\begin{aligned}
P(S = 2) &= 1/9 \\
P(S = 3) &= 2/9 \\
P(S = 4) &= 3/9 = 1/3 \\
P(S = 5) &= 2/9 \\
P(S = 6) &= 1/9.
\end{aligned}$$

The marginal pmf of D has values

$$\begin{aligned}
P(D = 0) &= 3/9 = 1/3 \\
P(D = 1) &= 4/9 \\
P(D = 2) &= 2/9
\end{aligned}$$

c) By direct computation

$$\begin{aligned}
E(S) &= 2(1/9) + 3(2/9) + 4(3/9) + 5(2/9) + 6(1/9) \\
&= 36/9 \\
&= 4,
\end{aligned}$$

$$\begin{aligned}
E(S^2) &= 2^2(1/9) + 3^2(2/9) + 4^2(3/9) + 5^2(2/9) + 6^2(1/9) \\
&= 156/9 \\
&= 52/3.
\end{aligned}$$

Then

$$V(S) = E(S^2) - E(S)^2 = \frac{52}{3} - 4 = \frac{4}{3} = 1.333.$$

Similarly

$$\begin{aligned}
E(D) &= 0(1/3) + 1(4/9) + 2(2/9) \\
&= 8/9
\end{aligned}$$

$$
\begin{aligned}
E(D^2) &= 0(1/3) + 1^2(4/9) + 2^2(2/9) \\
&= 16/9
\end{aligned}
$$

Then

$$
V(D) = E(D^2) - E(D)^2 = \frac{16}{9} - \left(\frac{8}{9}\right)^2 = \frac{80}{81} = .98765.
$$

d) Applying the shortcut formula,

$$
\begin{aligned}
E(SD) &= (1/9)(2)(0) + (2/9)(3)(1) + (1/9)(4)(0) + (2/9)(4)(2) + (2/9)(5)(1) + (1/9)(6)(0) \\
&= 0 + 6/9 + 0 + 16/9 + 10/9 + 0 \\
&= 32/9
\end{aligned}
$$

so

$$
\begin{aligned}
\mathrm{Cov}(S, D) &= E(SD) - E(S)E(D) \\
&= 32/9 - 4(8/9) \\
&= 0.
\end{aligned}
$$

Then

$$
\mathrm{Corr}(S, D) = \frac{\mathrm{Cov}(S, D)}{\sqrt{V(S)V(D)}} = 0.
$$

e) No, S and D are not independent – they are dependent. An immediate reason is that the support of one variable depends on the other. E.g., D can only take the value 0 is $S = 2$, but can only take the value 1 is $S = 3$. Alternatively, consider that

$$
P(S = 2, D = 1) = 0 \neq P(S = 2)P(D = 1) = \frac{4}{81}.
$$

This shows that there exists s and d such that

$$
P(S = s, D = d) \neq P(S = s)P(D = d).
$$

This is an example in which the $\mathrm{Cov}(S, D)$ is 0, but the random variables are not independent.