Recitation 7: Multiple R.V.'s, Covariance, Correlation

1 Review

1.1 Discrete R.V.s

- Two discrete random variables, X and Y, can be characterized by their *joint PMF*:

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

defined for every x, y pair where either $P(X = x) \geq 0$ or $P(Y = y) \geq 0$.

- Given a joint PMF, the *marginal PMFs* of X and Y:

$$f_X(x) = \sum_{y \in S(X,Y)} f_{X,Y}(x,\, y)$$
$$f_Y(y) = \sum_{x \in S(X,Y)} f_{X,Y}(x,\, y)$$

- The *expectation* of a function $g(X,\, Y)$ of two random variables

$$E[g(X,\, Y)] = \sum\sum_{(x,y) \in S(X,Y)} g(x,\, y)\, f_{X,Y}(x,\, y)$$

1.2 Continuous R.V.'s

- Two continuous random variables, X and Y, can be characterized by their *joint PDF* $f_{X,Y}(x,\, y)$.
- Marginal distributions and expectations can be computed as above (replace sums with integrals).

1.3 For both discrete and continuous R.V.'s

- Two random variables X and Y are *independent* if for all real numbers x and y

$$f_{X,Y}(x,\, y) = f_X(x)\, f_Y(y).$$

Equivalently, they are independent if for all sets A and B,

$$P(X \in A, Y \in B) = P(X \in A)P(Y \in B).$$

- The *covariance* between two random variables X and Y is

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

- If X and Y are independent, then $\mathrm{Cov}(X,\, Y) = 0$. However, if $\mathrm{Cov}(X,\, Y) = 0$, then X and Y may or may not be independent.

- Properties of covariance

$$\mathrm{Cov}(aX+b,\, Y) = a\mathrm{Cov}(X,\, Y)$$
$$\mathrm{Cov}(X,\, X) = \mathrm{Var}(X)$$

- The *correlation* between two random variables X and Y is

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

  which always falls between -1 and 1.

2 Exercises

1. True or false.

   (a) If X and Y are independent, then $\rho(X,Y)=0$.
   True **False**

   (b) If $\rho(X,Y)=0$, then X and Y are independent.
   True **False**

   (c) If $\mathrm{Var}(X+Y) > \mathrm{Var}(X)+\mathrm{Var}(Y)$, then X and Y are dependent.
   True **False**

   (d) If X and Y are dependent, then $\mathrm{Var}(X+Y) > \mathrm{Var}(X)+\mathrm{Var}(Y)$.
   True **False**

   (e) If X and Y are independent, then for any functions g and h, $E\big(g(X)h(Y)\big) = E\big(g(X)\big)E\big(h(Y)\big)$.
   True **False**

   (f) We always have $E\big(g_1(X,Y)+g_2(X,Y)\big) = E\big(g_1(X,Y)\big)+E\big(g_2(X,Y)\big)$, regardless of what the joint distribution of X and Y is.
   True **False**

2. A furniture store sells desks and armchairs. Let X and Y represent the number of desks and armchairs sold on a given day, respectively, and assume that their distribution is given by the following joint PMF:

			y	
$f_{X,Y}(x,y)$		0	1	2
	0	0.10	0.04	0.02
x	1	0.08	0.20	0.06
	2	0.06	0.14	0.30

   (a) Find the marginal distributions $f_X(x)$ and $f_Y(y)$.

   (b) Compute $E[X]$ and $E[Y]$.

   (c) Suppose desks sell for \$200 each, and armchairs for \$100 each. Let R be the revenue that the furniture store makes on a given day. Compute $E[R]$.

   (d) Compute $\mathrm{Var}(R)$.

(e) Are X and Y independent? Why or why not?

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

(g) What is $\rho(X, Y)$, the correlation between X and Y?

3. Each tire on a particular motorcycle should be filled to a pressure of 15 psi. Suppose that the actual air pressure in each tire is a random variable. Let X and Y be the pressure in the front and rear tires, respectively, and assume they have joint PDF

$$f_{X,Y}(x,y) = \begin{cases} kxy & 0 \leq x \leq 30,\ 0 \leq y \leq 20 \\ 0 & \text{otherwise} \end{cases}.$$

(a) Find k.

(b) What is the probability that both tires are underfilled?

(c) Compute the marginal densities $f_X(x)$ and $f_Y(y)$.

(d) Compute $E[X - Y]$. (**Hint**: Linearity of expectations.)

(e) Are X and Y independent? Why are why not?

4. Suppose X and Y are independent Uniform(0, 1) random variables. Compute the probabilities $P(\min\{X, Y\} \leq 0.5)$ and $P(\max\{X, Y\} \leq 0.5)$. **Hint**: You could integrate over the joint density, but there is a simpler way using the fact that if X and Y are independent, then

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

regardless of whether X and Y are discrete or continuous.

5. (This problem is relatively hard and optional.) A bakery sells 3 varieties of cupcakes. You visit the bakery every morning, and ask the person at the counter to randomly select one cupcake for your breakfast. Assume that each variety is equally likely to be chosen, and that cupcakes are selected independently of one another. Let X be the number of days needed to sample each type of cupcake at least once. Compute $E[X]$ and $\mathrm{Var}(X)$.

(**Hint**: For $i = 1, 2, 3$, let Y_i denote the number of days between the first day you try $i - 1$ types of cupcakes at least once, and the first day you try i types of cupcakes at least once. Note that

$$X = Y_1 + Y_2 + Y_3$$

What kind of random variable is Y_1 i.e. the number of days it takes to sample the first cupcake? What about Y_2 i.e. the number of days, after sampling the first cupcake, that you sample a second type of cupcake?)