STA 111 (Summer Session I)

Lecture Seven – Joint Distributions, Covariance and Correlation
D.S. Sections 3.4, 3.5, 3.5 and 4.6

Instructor: Olanrewaju (Michael) Akande

Department of Statistical Science, Duke University

May 26, 2016

Outline

- Questions from Last Lecture.
- Joint Distributions
- Covariance and Correlation
- Recap

Introduction

– We have already talked about random variables, their distributions, density (or mass) functions and expectations.

– Today we will extend those concepts to two or more variables jointly. We will cover joint distributions for bivariate distributions as well as conditional densities, independence, covariance and correlation.

Joint Distributions

In general, a joint density (or mass) function gives the probability distribution for more than one random variable (i.e., a random vector). These joint distributions may be discrete, continuous, or mixed.

We will focus on the continuous case with two random variables in defining the different concepts but all the definitions extend to the discrete case.
First,

$$\mathbb{P}[a \leq X \leq b \text{ and } c \leq Y \leq d] = \int_c^d \int_a^b f(x,y)\, dxdy$$

where $f(x,y)$ is the joint density function.
So that the joint cumulative distribution function (cdf) is:

$$F(x,y) = \mathbb{P}[X \leq x \text{ and } Y \leq y] = \int_{-\infty}^{x} \int_{-\infty}^{y} f(x,y)\, dxdy$$

In order to be a joint density function, $f(x,y)$ must be non-negative, and it must integrate to 1.

Example

Example 1: You compare the share price of two stocks, say Apple, Inc. and Orange Julius Co. Apple share values (X) cannot exceed \$6.00, and Orange Julius shares (Y) cannot exceed \$9.00.

Random market forces cause the values to vary. Assume that on a given day, the share prices in X and Y have joint probability density function

$$f(x,y) = \begin{cases} \frac{1}{4374}xy^2 & \text{for } 0 \le x \le 6, 0 \le y \le 9 \\ 0 & \text{otherwise.} \end{cases}$$

Qualitatively, this joint density function puts less probability on extremely low values, which seems reasonable. It also suggests that the probability of large values for Orange Julus increases faster than the probability of large values for Apple, which may also be plausible.

Marginal Densities

Is the function in our stock value example a joint density? Yes, since:

$$\begin{aligned}\int_0^9 \int_0^6 \frac{1}{4374} xy^2 \, dxdy &= \frac{1}{4374} \int_0^9 18y^2 \, dy \\ &= \frac{1}{4374}(18 * 243) = 1.\end{aligned}$$

The marginal densities of X and Y are, respectively,

$$\begin{aligned}f_1(x) &= \int_{-\infty}^{\infty} f(x,y)\, dy \text{ for } -\infty < x < \infty \\ f_2(y) &= \int_{-\infty}^{\infty} f(x,y)\, dx \text{ for } -\infty < y < \infty\end{aligned}$$

The marginals give us the density for just the X random variable or just the Y random variable, ignoring the other.

Example

Example 1 Cont'd: If I just wanted to know the marginal density function for the value of an Apple share, it would be

$$\begin{aligned} f_1(x) &= \int_{\infty}^{\infty} f(x,y)\,dy \text{ for } -\infty < x < \infty \\ &= \frac{1}{4374}\int_0^9 xy^2\,dy \text{ for } 0 \leq x \leq 6 \\ &= \frac{1}{18}x \text{ for } 0 \leq x \leq 6. \end{aligned}$$

Similarly, if I wanted to know just the probability density function for the Orange Julius share, a similar integration would give

$$\begin{aligned} f_2(y) &= \int_{-\infty}^{\infty} f(x,y)\,dx \text{ for } -\infty < x < \infty \\ &= \frac{1}{243}y^2 \text{ for } 0 \leq y \leq 9. \end{aligned}$$

Independence and Conditional Densities

Two random variables are independent if and only if

$$f(x, y) = f_1(x) * f_2(y).$$

In our example, are the the two share prices independent?

Yes, since $(1/243)y^2 * (1/18)x = (1/4374)xy^2$.

The conditional density of X given that $Y = y$ is

$$g_1(x|y) = \frac{f(x, y)}{f_2(y)}.$$

The conditional density $g_2(y|x)$ for Y given $X = x$ is similarly defined.

Example

Example 1 Cont'd: For our example, what is the conditional density for the Apple stock value if an Orange Julius share costs \$5.00?

$$
\begin{aligned}
g_1(x|y=5) &= \frac{f(x,5)}{f_2(5)} \\
&= \frac{\left(\frac{1}{4374}x5^2\right)}{\left(\frac{1}{243}5^2\right)} \\
&= \frac{1}{18}x \text{ for } 0 \leq x \leq 6.
\end{aligned}
$$

Since the random variables are independent, knowing the value of Orange Julius does not change our probability for the value of Apple.

Note: As usual, we use X and Y to denote random variables, and x and y to denote values they may take. Above, we observed that the outcome for Y was \$5.00 and sought the corresponding density for X.

Covariance

The expectation of a function $h(X, Y)$ is

$$\mathbb{E}[h(X, Y)] = \int_{-\infty}^{\infty}\int_{-\infty}^{\infty} h(x, y)f(x, y)\, dxdy.$$

For example, the expected value of the product of X and Y is

$$\mathbb{E}[XY] = \int_{-\infty}^{\infty}\int_{-\infty}^{\infty} xyf(x, y)\, dxdy.$$

A particularly useful function is $h(X, Y) = (X - \mu_X)(Y - \mu_Y)$. Its expectation is called the covariance.
One can show that the covariance between X and Y can be calculated more simply than the definitional formula suggests. It reduces to

$$\text{Cov}(X, Y) = \mathbb{E}[XY] - \mu_X * \mu_Y.$$

Correlation

When there are more than two random variables, e.g., there are X, Y, and Z, then one can calculate the covariance matrix whose entries are all possible pairs of covariances.

The covariance is important because it allows us to calculate the correlation between X and Y:

$$\text{Corr}(X, Y) = \frac{\text{Cov}(X, Y)}{\sigma_X * \sigma_Y}.$$

Recall that σ_X is the standard deviation of X, or the square-root of $\mathbb{E}[X^2] - \mu_X^2$.

Correlation is a measure of strength of association/relationship between two random variables. See graph on the board.

Correlation Cont'd

Some useful facts about the correlation and covariance:

- If X and Y are independent, then $\text{Corr}(X, Y) = 0$.
- The converse fails: if $\text{Corr}(X, Y) = 0$, then X and Y may be dependent.
- $-1 \leq \text{Corr}(X, Y) \leq 1$.
- $\text{Cov}(aX, b) = 0$.
- $\text{Cov}(aX + b, cY + d) = ac\text{Cov}(X, Y)$.
- $\text{Corr}(aX + b, cY + d) = \text{Corr}(X, Y)$.
- The $\text{Corr}(X, Y) = \pm 1$ if and only if $Y = aX + b$ for some $a \neq 0$.

The book covers this material in Chapter 4, and treats both continuous and discrete distributions. As we have seen before, that simply means that summations replace integrals in the formulæ shown here.

A Discrete Example

Example 2 (To be done in class): Consider the joint probability mass function for bivariate discrete random variables defined by the following table:

	x=1	x=2	x=3
y=0	0	0.09	0.15
y=4	0.15	0.04	0.25
y=5	0.07	0.05	c

1. What is the value of c?
2. What is the expected value of the product XY?
3. What is $\mathbb{P}(X = 1|y = 4)$?
4. Are X and Y independent?
5. What is the expected value of X? What is the expected value of Y?
6. What is the covariance between X and Y?
7. What is the correlation between X and Y?

Another Example

Example 3 (To be done in class): Suppose $f(x, y) = 6x$ for $x + y \leq 1$ with both x and y restricted to be between 0 and 1.

Hint: When working with bivariate densities, it is always a good idea to draw the support.

1. What is the marginal density of X?
2. What is the expected value of X?
3. What is the marginal density of Y?
4. What is the expected value of Y?
5. Are X and Y independent?

Recap

We discussed the following:

- Joint Distributions
- Covariance
- Correlation