

# Tutorial 9: Further Topics on Random Variables 2

Weiwen LIU  
wwliu@cuhk.edu.hk  
March 27, 2017

# Covariance and Correlation

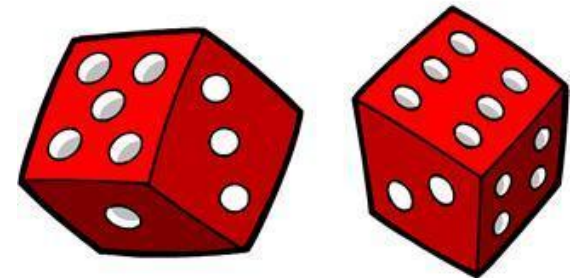
- Covariance and correlation describe the degree to which two random variables or sets of random variables tend to deviate from their expected values in similar ways.
- Independent random variables are uncorrelated, but NOT vice versa.

# Example 1

- Let  $X$  and  $Y$  be continuous random variables with joint pdf

$$f_{X,Y}(x,y) = \begin{cases} 3x, & 0 \leq y \leq x \leq 1, \\ 0, & \text{otherwise.} \end{cases}$$

- Compute  $\text{cov}(X, Y)$ ;
- Are  $X$  and  $Y$  independent?



# Example 1

- The marginal pdfs and expectations of X and Y are

$$f_X(X) = \int_0^x 3x dy = 3x^2, 0 \leq x \leq 1,$$

$$E[X] = \int_0^1 x \cdot 3x^2 dx = \left[ \frac{3}{4} x^4 \right]_0^1 = \frac{3}{4},$$

$$f_Y(y) = \int_y^1 3x dx = \left[ \frac{3}{2} x^2 \right]_y^1 = \frac{3}{2} (1 - y^2), 0 \leq y \leq 1,$$

$$E[Y] = \int_0^1 y \cdot \frac{3}{2} (1 - y^2) dy = \left[ \frac{3}{2} \left( \frac{y^2}{2} - \frac{y^4}{4} \right) \right]_0^1 = \frac{3}{8},$$

$$E[XY] = \int_0^1 \int_0^x xy \cdot 3x dy dx = \frac{3}{2} \left[ \frac{x^5}{5} \right]_0^1 = \frac{3}{10}$$

# Example 1

- Then the covariance is

$$\text{cov}(X, Y) = E[XY] - E[X]E[Y] = \frac{3}{10} - \frac{3}{4} \times \frac{3}{8} = \frac{3}{160}.$$

- $X$  and  $Y$  are not independent as it is **not true** that

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

# Variance of Summations

- Let  $X_1, X_2, \dots, X_n$  be random variables with finite variance, then we have

$$\text{var}(\sum_i X_i) = \sum_i \text{var}(X_i) + \sum_{i \neq j} \text{cov}(X_i, X_j)$$

## Example 2

- If the variance of verbal GRE were 64, the variance of quantitative SAT were 81 and the correlation between these two tests were 0.50, then what is the variance of total SAT (verbal + quantitative)?



## Example 2

- Denote  $X_1, X_2$  as the score for verbal and quantitative respectively, then

$$\begin{aligned}\text{var}(X_1 + X_2) &= \text{var}(X_1) + \text{var}(X_2) + 2\text{cov}(X_1, X_2) \\ &= 64 + 81 + 2 \times 0.5 \times \sqrt{64} \times \sqrt{81} \\ &= 217\end{aligned}$$



# Conditional Expectation Revisit

- The conditional expectation  $E[X|Y]$  of a random variable  $X$  given another random variable  $Y$ , is a new random variable determined by  $Y$ .
- It's distribution is determined by the distribution of  $Y$ .

# Expectation of CE

- The conditional expectation  $E[X|Y]$  is a random variable. It has expectation (over  $Y$ )

$$E[E[X|Y]] = \sum_y E[X|Y = y] p_Y(y),$$

- or for continuous  $Y$

$$E[E[X|Y]] = \int_R E[X|Y = y] f_Y(y) dy.$$

# Expectation of CE - Properties

- If  $X$  has a finite expectation, we have, concluded straightforwardly from the total expectation theorem of the Law of iterated expectations

$$E[E[X|Y]] = E[X].$$

- For any function  $g$ , we have

$$E[Xg(Y)] = g(Y)E[X|Y].$$

## Example 3

- A class has  $n$  students with score  $x_1, \dots, x_n$ . Let  $m = \frac{1}{n} \sum_i x_i$ . Divide them into  $k$  disjoint subsets  $A_1, \dots, A_n$ . Let  $n_s = |A_s|$ , and  $m_s = \frac{1}{n_s} \sum_{i \in A_s} x_i$  be the averaged score of the  $s$ -th section. We have

$$\sum_s \frac{n_s}{n} m_s = m.$$

## Example 3

- The result  $\sum_s \frac{n_s}{n} m_s = m$  can be viewed from a CE perspective. Let  $X$  be the score of a random student, and  $Y$  be the section of a random student, we have

$$\begin{aligned}\sum_s \frac{n_s}{n} m_s &= \sum_s E[X|Y = s]P(Y = s) \\ &= E[E[X|Y]] \\ &= E[X] = m\end{aligned}$$

## Example 4

- Let  $X$  be the sales of the next year, and  $Y$  be the sales of the first quarter of next year. Suppose we have a forecast system giving the joint distribution of  $X$  and  $Y$ . We view  $Z = E[X|Y] - E[X]$  as the forecast revision

$$E[Z] = E[E[X|Y]] - E[E[X]] = 0$$

- Intuitively, if  $Z$  is positive, the forecast system underestimates  $E[X|Y]$ .

# CE as Estimator

- If we could observe  $Y$  who provides information about  $X$ , it's natural we estimate  $Y$  using  $X$ , as

$$\hat{X} = E[X|Y].$$

- The estimation error

$$\tilde{X} = \hat{X} - X$$

- is a random variable satisfying

$$E[\tilde{X}|Y] = E[\hat{X}|Y] - E[X|Y] = 0$$

- And hence  $E[\tilde{X}] = 0$

# CE as Estimator

- An important property is that the estimation  $\hat{X}$  is uncorrelated with the estimator error  $\tilde{X}$ . In fact,

$$E[\hat{X}\tilde{X}] = E[E[\hat{X}\tilde{X}|Y]] = E[\hat{X}E[\tilde{X}|Y]] = 0.$$

- Hence

$$\text{cov}(\hat{X}, \tilde{X}) = E[\hat{X}\tilde{X}] - E[\hat{X}]E[\tilde{X}] = 0.$$

- As a result,

$$\text{var}(X) = \text{var}(\tilde{X}) + \text{var}(\hat{X})$$



# Conditional Variance

- The conditional variance is defined as

$$\text{var}(X|Y) = E[\tilde{X}^2|Y] = E[(X - E[X|Y])^2|Y].$$

- As usual,

$$\text{var}(X|Y = y) = E[\tilde{X}^2|Y = y].$$

- Also,

$$\text{var}(\tilde{X}) = E[\tilde{X}^2] = E[E[\tilde{X}^2|Y]] = E[\text{var}(X|Y)].$$

# Law of Total Variance

- Law of Total Variance

$$\text{var}(X) = E[\text{var}(X|Y)] + \text{var}(E[X|Y]).$$

- It's especially useful to calculate variances of random variables.
- See the following examples

## Example 5

- Consider we tossing a coin, of whom the probability of heads is a uniform random variable  $Y$ , the number of heads  $X$  satisfies  $E[X|Y] = nY$  and

$$\text{var}(E[X|Y]) = \text{var}(nY) = \frac{n^2}{12}$$

## Example 5

- Meanwhile we have  $\text{var}(X|Y) = nY(1 - Y)$ , and

$$E[\text{var}(X|Y)] = \frac{n}{6}$$

- Then using law of total variance,

$$\begin{aligned}\text{var}(X) &= E[\text{var}(X|Y)] + \text{var}(E[X|Y]). \\ &= \frac{n}{6} + \frac{n^2}{12}\end{aligned}$$

# Transforms

- The transform provides us with an **alternative representation** of its probability law (PMF or PDF).
- It is not particularly intuitive, but it is often **convenient** for certain types of mathematical manipulations.
- The transform (also referred to as the **associated moment generating function**) is defined as:

$$M_X(s) = E[e^{sX}],$$

which is a function of a scalar parameter  $s$ .

# Inversion Property

- Suppose that  $M_X(s)$  is finite for all  $s$  in an interval of the form  $[-a, a]$ , where  $a$  is a positive number. Then,  $M_X$  **determines uniquely** the CDF of the random variable  $X$ .
- If  $M_X(s) = M_Y(s) < \infty$ , for all  $s \in [-a, a]$ , where  $a$  is a positive number, then the random variables  $X$  and  $Y$  have the same CDF.

# Sums of Independent Variables

- Generally,  $X_1, \dots, X_n$  is a collection of independent random variables, and

$$Z = X_1 + \dots + X_n$$

- Then,

$$M_Z(s) = M_{X_1}(s) \cdots M_{X_n}(s)$$

# Sums of A Random Number of Independent Random Variables

- Let  $Y = X_1 + \cdots + X_N$ , where  $N$  is a **random variable** that takes nonnegative integer values.
- Then,

$$\begin{aligned}\mathbf{E}[Y] &= \mathbf{E}[\mathbf{E}[Y|N]] = \mathbf{E}[N]\mathbf{E}[X], \\ \text{var}\mathbf{E}(Y) &= \mathbf{E}[N]\text{var}(X) + \mathbf{E}[X]^2\text{var}(N), \\ M_Y(s) &= \sum_{n=0}^{\infty} M_X(s)^n p_N(n).\end{aligned}$$



## Example 6

- Let  $N$  be geometrically distributed with parameter  $p$ , and let each random variable  $X_i$  be geometrically distributed with parameter  $q$ . We assume that all of these random variables are independent. Let  $Y = X_1 + \cdots + X_N$ . We have

$$M_N(s) = \frac{pe^s}{(1 - (1 - p)e^s)},$$
$$M_X(s) = \frac{qe^s}{1 - (1 - q)e^s}.$$

What is the distribution of  $Y$ ?

## Example 6

- To determine  $M_Y(s)$ , we start with the formula for  $M_N(s)$  and replace each occurrence of  $e^s$  with  $M_X(s)$ . This yields

$$M_Y(s) = \frac{pM_X(s)}{1 - (1 - p)M_X(s)}$$

and, after some algebra

$$M_Y(s) = \frac{pqe^s}{1 - (1 - pq)e^s}.$$

We conclude that  $Y$  is geometrically distributed, with parameter  $pq$ .