EE 381 Probability & Statistic with Applications to Computing (Fall 2020)

Lecture 6
Discrete Random Variables (p3)

Duc H. Tran, PhD

Discrete Random Variable’s topics

- Discrete probability distributions
- Joint distributions
- Mathematical expectation
- Variance, Standardized random variables
- **Moments and Moment generating function**
- **Covariance & correlation**
- Special Probability Distributions

Moments and Moment Generating Function

- The *rth moment of a random variable X about the mean* μ, also called the *rth central moment*, is defined as:

$$\mu_r = E[(X-\mu)^r]$$

Where *r=0, 1, 2…* It follows that $\mu_0 = 1$, $\mu_1 = 0$, and $\mu_2 = \sigma^2$,i.e., the second moment about the mean is the variance.

- The *rth moment of X about the origin*, also called the *rth raw moment*, is defined as:

$$\mu'_r = E(X^r)$$

Moments and Moment Generating Function

- For many probability mass functions, computing expected values is not easy since the sum may be difficult to compute → we can use *moment generating function*.
- The moment generating function of X is defined by:

$$\mathscr{M}_X(u) = E\left[e^{uX}\right] = \sum_k e^{uk} p(k)$$

- Expanding the moment generating function by applying the Maclaurin series to the exponential term:

$$e^{uX} = 1 + uX + \frac{u^2 X^2}{2!} + \frac{u^3 X^3}{3!} + \cdots$$

$$\boxed{\begin{aligned} \mathscr{M}(u) = E\left[e^{uX}\right] &= E[1] + E[uX] + E\left[\frac{u^2 X^2}{2!}\right] + E\left[\frac{u^3 X^3}{3!}\right] + \cdots \\ &= 1 + uE[X] + \frac{u^2}{2!}E[X^2] + \frac{u^3}{3!}E[X^3] + \cdots \end{aligned}}$$

Moments and Moment Generating Function

- Now take a derivative with respect to *u*:

$$\frac{d}{du}\mathcal{M}(u) = 0 + E[\mathbf{X}] + uE[\mathbf{X}^2] + \frac{u^2}{2!}E[\mathbf{X}^3] + \cdots$$

- Finally set *u=0*: $$\frac{d}{du}\mathcal{M}(u)\Big|_{u=0} = 0 + E[\mathbf{X}] + 0 + 0 + \cdots = E[\mathbf{X}]$$

Notice that the only term that "survive" both steps is *E[X]*.

- The second moment is found by taking two derivatives, then setting *u=0*:

$$E[\mathbf{X}^2] = \frac{d^2}{du^2}\mathcal{M}(u)\Big|_{u=0}$$

- In general, the *kth* moment can be found as:

$$E[\mathbf{X}^k] = \frac{d^k}{du^k}\mathcal{M}(u)\Big|_{u=0}$$

Moment Generating Functions

Example: The random variable X can assume the values 1 and -1 with probability ½ each. Find:

a) The moment generating function

b) The first four moments about the origin

Solution:

a)
$$E(e^{tX}) = e^{t(1)}\left(\frac{1}{2}\right) + e^{t(-1)}\left(\frac{1}{2}\right) = \frac{1}{2}(e^t + e^{-t})$$

b) We have:
$$e^t = 1 + t + \frac{t^2}{2!} + \frac{t^3}{3!} + \frac{t^4}{4!} + \cdots$$
$$e^{-t} = 1 - t + \frac{t^2}{2!} - \frac{t^3}{3!} + \frac{t^4}{4!} - \cdots$$

Then
$$\frac{1}{2}(e^t + e^{-t}) = 1 + \frac{t^2}{2!} + \frac{t^4}{4!} + \cdots \qquad (*)$$

But,
$$\mathcal{M}_X(t) = 1 + tE[X] + \frac{t^2}{2!}E[X^2] + \frac{t^3}{3!}E[X^3] + \frac{t^4}{4!}E[X^4] + \cdots \qquad (**)$$

Comparing (*) and (**), we have: $E[X] = 0$, $E[X^2] = 1$, $E[X^3] = 0$, $E[X^4] = 1$

The odd moments are all zero, and the even moments are all one.

In-class Exercise

Let X is a discrete random variable, and X is described as follow:

$$X = \begin{cases} 1/2 & \text{prob. } 1/2 \\ -1/2 & \text{prob. } 1/2 \end{cases}$$

a) Find the moment generating function of X.

b) The first four moments about the origin.

Some Theorems on Moment Generating Functions

Theorem 3-8: If $M_X(t)$ is the moment generating function of the random variable *X*, and *a* and *b* are constants, then the moment generating function of *(X+a)/b* is

$$M_{(X+a)/b}(t) = e^{at/b} M_X\left(\frac{t}{b}\right)$$

Theorem 3-9: If X and Y are independent random variables having moment generating function $M_X(t)$ and $M_Y(t)$, respectively, then

$$M_{X+Y}(t) = M_X(t) M_Y(t)$$

Theorem 3-10 (**Uniqueness Theorem**): Suppose that X and Y are random variables having moment generating functions $M_X(t)$ and $M_Y(t)$, respectively. Then X and Y have the same probability distribution if and only if $M_X(t) = M_Y\ (t)$ identically.

Variance for Joint Distribution: Covariance

Covariance shows the relationship between the random variables. Specifically, it is a measure of joint variability of two random variables.

If X and Y are two discrete random variables having joint density function f(x, y), the means, or expectations, of X and Y are:

$$\mu_X = \sum_x \sum_y x f(x,y) \qquad \mu_Y = \sum_x \sum_y y f(x,y)$$

Their variances are:

$$\sigma_X^2 = E[(X-\mu_X)^2] \qquad \sigma_Y^2 = E[(Y-\mu_Y)^2]$$

The covariance of two variables X and Y is defined by:

$$\sigma_{XY} = Cov(X,Y) = E[(X-\mu_X)(Y-\mu_Y)]$$
$$= \sum_x \sum_y (x-\mu_X)(y-\mu_Y) f(x,y)$$

Variance for Joint Distribution: Covariance

Important theorems on covariance:

Theorem 3-14 $\sigma_{XY} = E(XY) - E(X)E(Y) = E(XY) - \mu_X\mu_Y$

Theorem 3-15 If X and Y are independent random variables, then

$$\sigma_{XY} = \mathrm{Cov}(X, Y) = 0$$

Theorem 3-16 $\mathrm{Var}(X \pm Y) = \mathrm{Var}(X) + \mathrm{Var}(Y) \pm 2\,\mathrm{Cov}(X, Y)$

or $\sigma^2_{X \pm Y} = \sigma^2_X + \sigma^2_Y \pm 2\sigma_{XY}$

Theorem 3-17 $|\sigma_{XY}| \leq \sigma_X \sigma_Y$

Variance for Joint Distribution: Covariance

Example: The joint probability function of 2 discrete random variable X and Y is: $f(x, y) = c(2x+y)$, where x and y are integers such that $0 \le x \le 2$, $0 \le y \le 3$, and $f(x, y) = 0$ otherwise.

Since the grand total, 42c, must equal 1, then **c = 1/42.**

$X \backslash Y$	0	1	2	3	Totals ↓
0	0	c	$2c$	$3c$	$6c$
1	$2c$	$3c$	$4c$	$5c$	$14c$
2	$4c$	$5c$	$6c$	$7c$	$22c$
Totals →	$6c$	$9c$	$12c$	$15c$	$42c$

Find $E(X)$, $E(Y)$, $E(XY)$, $E(X^2)$, $Var(X)$, $Cov(X, Y)$?

Variance for Joint Distribution: Covariance

Example (ctn.):

a)

$$E(X) = \sum_x \sum_y x f(x, y) = \sum_x x \left[\sum_y f(x, y) \right]$$

$$= (0)(6c) + (1)(14c) + (2)(22c) = 58c = \frac{58}{42} = \frac{29}{21}$$

b)

$$E(Y) = \sum_x \sum_y y f(x, y) = \sum_y y \left[\sum_x f(x, y) \right]$$

$$= (0)(6c) + (1)(9c) + (2)(12c) + (3)(15c) = 78c = \frac{78}{42} = \frac{13}{7}$$

c)

$$E(XY) = \sum_x \sum_y xy f(x, y)$$

$$= (0)(0)(0) + (0)(1)(c) + (0)(2)(2c) + (0)(3)(3c)$$

$$+ (1)(0)(2c) + (1)(1)(3c) + (1)(2)(4c) + (1)(3)(5c)$$

$$+ (2)(0)(4c) + (2)(1)(5c) + (2)(2)(6c) + (2)(3)(7c)$$

$$= 102c = \frac{102}{42} = \frac{17}{7}$$

Variance for Joint Distribution: Covariance

Example (ctn.):

d)

$$E(X^2) = \sum_x \sum_y x^2 f(x, y) = \sum_x x^2 \left[\sum_y f(x, y) \right]$$

$$= (0)^2(6c) + (1)^2(14c) + (2)^2(22c) = 102c = \frac{102}{42} = \frac{17}{7}$$

e)

$$\sigma_X^2 = \mathrm{Var}(X) = E(X^2) - [E(X)]^2 = \frac{17}{7} - \left(\frac{29}{21}\right)^2 = \frac{230}{441}$$

f)

$$\sigma_{XY} = \mathrm{Cov}(X, Y) = E(XY) - E(X)E(Y) = \frac{17}{7} - \left(\frac{29}{21}\right)\left(\frac{13}{7}\right) = -\frac{20}{147}$$

Different formulas for Variance & Covariance

In case the probability function is not given, use the following formulas to calculate variance and covariance:

Variance

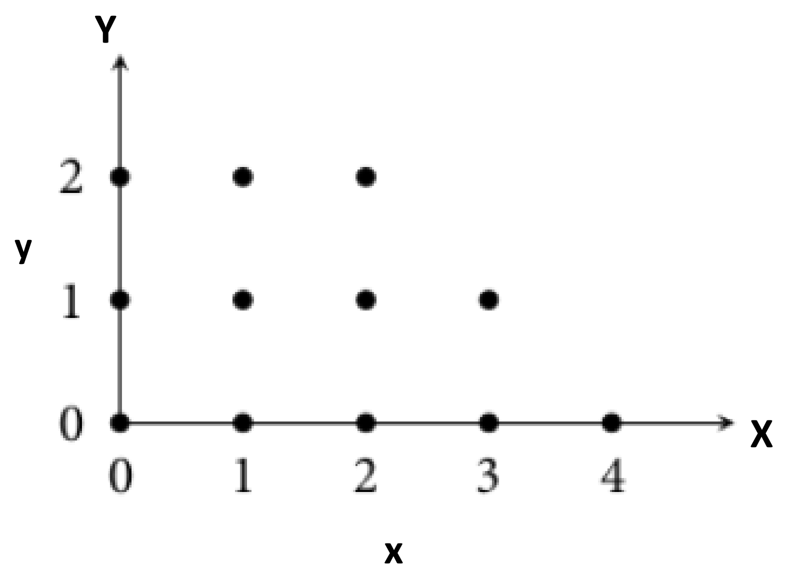
- For Population: $Var(X) = \frac{1}{N}\sum_{i=1}^{N}(x_i - E(X))^2$
- For Sample: $Var(X) = \frac{1}{N-1}\sum_{i=1}^{N}(x_i - E(X))^2$

Covariance

- For Population: $Cov(X,Y) = \frac{1}{N}\sum_{i=1}^{N}(x_i - E(X))(y_i - E(Y))$
- For Sample: $Cov(X,Y) = \frac{1}{N-1}\sum_{i=1}^{N}(x_i - E(X))(y_i - E(Y))$

In-class Exercise

Let X and Y be discrete random variables which are restricted to 12 points shown in the graph below. Assuming that each point is equally likely. Find the covariance of X and Y, σ_{XY}.



Correlation Coefficient

If X and Y are independent, then $\sigma_{XY} = \mathrm{Cov}(X, Y) = 0$

If X and Y are completely dependent ($X = Y$), then $\mathrm{Cov}(X, Y) = \sigma_{XY} = \sigma_X \sigma_Y$.

A *measure of the dependence* of the variables X and Y are given by:

$$\rho = \frac{\sigma_{XY}}{\sigma_X \sigma_Y}$$

ρ: the correlation coefficient, or coefficient of correlation.

Based on theorem 3-17, we know that $-1 \leq \rho \leq 1$.

If $\rho = 0$ (i.e., the covariance is zero), X and Y are uncorrelated.

Correlation Coefficient

Example: Calculate the correlation coefficient of random variables X and Y.

$X \backslash Y$	0	1	2	3	Totals ↓
0	0	c	$2c$	$3c$	$6c$
1	$2c$	$3c$	$4c$	$5c$	$14c$
2	$4c$	$5c$	$6c$	$7c$	$22c$
Totals →	$6c$	$9c$	$12c$	$15c$	$42c$

Previously calculated: $\sigma_X^2 = Var\,(X) = \frac{230}{441}$

$$Cov\,(X,Y) = -\frac{20}{147}$$

Correlation Coefficient

Example (ctn.):

$$E(Y^2) = \sum_x \sum_y y^2 f(x, y) = \sum_y y^2 \left[\sum_x f(x, y) \right]$$

$$\sigma_Y^2 = \mathrm{Var}(Y) = E(Y^2) - [E(Y)]^2 = \frac{32}{7} - \left(\frac{13}{7}\right)^2 = \frac{55}{49}$$

$$\boxed{\rho = \frac{\sigma_{XY}}{\sigma_X \sigma_Y} = \frac{-20/147}{\sqrt{230/441}\sqrt{55/49}} = \frac{-20}{\sqrt{230}\sqrt{55}} = -0.2103}$$

Covariance vs Correlation

- In simple words, both the terms measure the relationship and the dependency between two random variables.
- Covariance is not standardized, and its value can range from $-\infty$ to $+\infty$
- Correlation is a normalized version of covariance. We want to normalize covariance to make it unitless. This provide better information for the relationship between two random variables. Its value range is [-1,1].

Covariance vs Correlation

Example: We would like to check the relationship of the daily return for two stocks using the closing price.

Day	Stock X	Stock Y
1	1.10%	3%
2	1.70%	4.20%
3	2.10%	4.90%
4	1.40%	4.10%
5	0.20%	2.50%

$$\mu_X = \frac{1.1 + 1.7 + 2.1 + 1.4 + 0.2}{5} = 1.3$$

$$\mu_Y = \frac{3 + 4.2 + 4.9 + 4.1 + 2.5}{5} = 3.74$$

$$Cov\ (X, Y) = \frac{\sum(Return_X - \mu_X) * (Return_Y - \mu_Y)}{Sample\ size\ - 1}$$

$$= \frac{[(1.1 - 1.3) * (3 - 3.74)] + [(1.7 - 1.3) * (4.2 - 3.74) + \cdots}{5 - 1} = 0.665$$

Covariance vs Correlation

Example (ctn.):

The covariance between the two stock returns is 0.665. This is positive → the stocks move in the same direction. In other words, when stock X had a high return, stock Y also had a high return.

However, it only show how the stocks move together. To determine the strength of the relationship (i.e., the degree to which both stocks move together), we need to look at their correlation.

$$\rho = \frac{\sigma_{XY}}{\sigma_X \sigma_Y} = \frac{0.665}{\sqrt{0.515 * 0.943}} = 0.9543$$

Investors usually select stocks that move in opposite directions because the risk will be lower.

Other Measures of Central Tendency

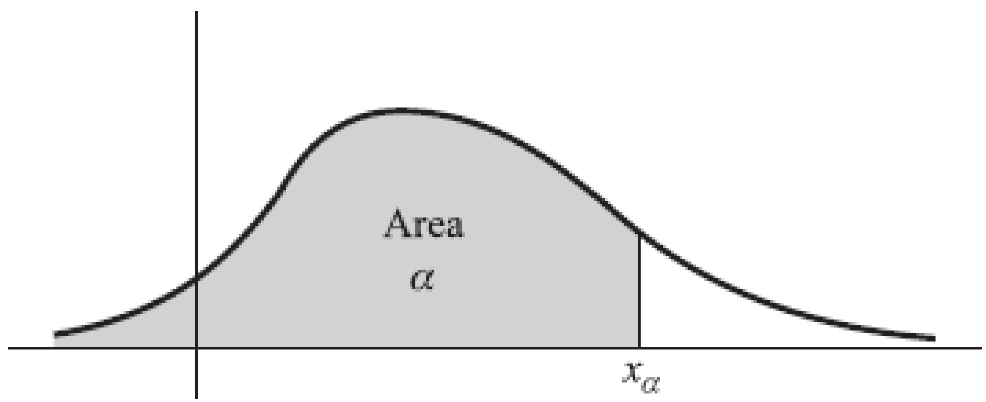
- The mean, or expectation, of a random variable X is a measure of central tendency for the values of a distribution.
- There are also other measures of central tendency: *mode* and *median*.

Mode: The mode of a discrete random variable is that value which occurs most often or has the greatest probability of occurring. Sometimes we have two (bimodal), three (trimodal), or more values (multimodal) that have large probability of occurrence.

Median: The median is that value x for which $P(X < x) \leq 1/2$ and $P(X > x) \geq 1/2$.

Percentiles

- Subdivide the area under a density curve by use of ordinates so that the area to the left of the ordinate is some percentage of the total unit area.
- The values corresponding to such areas are called *percentiles*.
- Percentiles are commonly used to report scores in tests (SAT, GRE, LSAT,…)



Reference

Notes, equations, and figures in the lecture are based on or taken from materials in the course textbook:

“Probability and Statistics”, by Spiegel, Schiller and Srinivasan, ISBN 987-007-179557-9 (McGraw-Hill/Schaun’s)