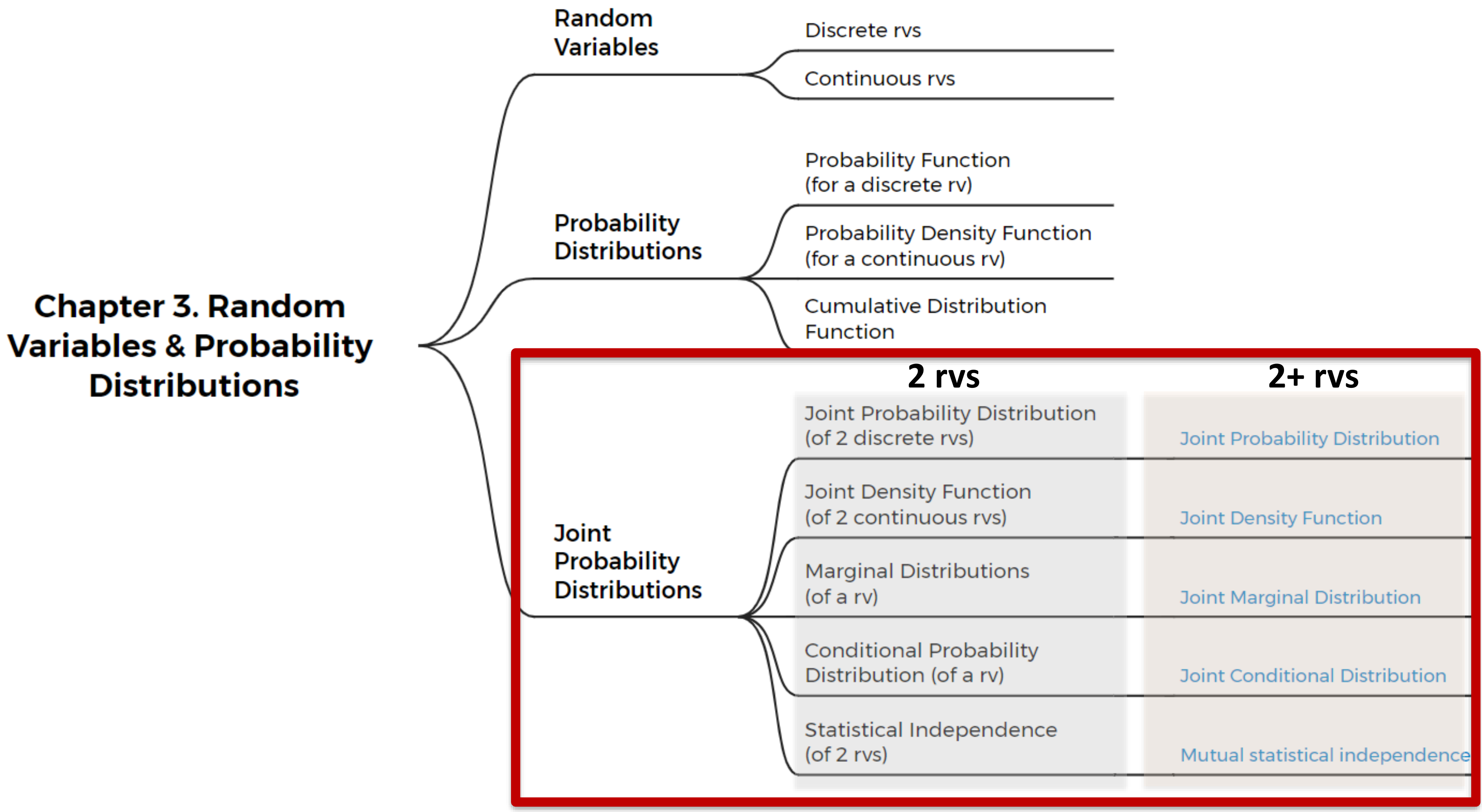
STAT 1151
Introduction to Probability

Lecture 5 Mathematical Expectation

Xiaomei Tan

Last Lecture



Outline

- Chapter 4 Mathematical Expectation
  - Mean of A Random Variable
  - Mean of A Function
  - Mean of Joint Random Variables
  - Means of Linear Combinations of Random Variables

Mean of A Random Variable

Quiz scores of five students: 10, 10, 9, 8, 8

- Arithmetic mean: $\dfrac{10+10+9+8+8}{5} = 9$
- Geometric mean: $\sqrt[5]{10 \times 10 \times 9 \times 8 \times 8} = 8.96$
- Harmonic mean: $\dfrac{5}{\frac{1}{10}+\frac{1}{10}+\frac{1}{9}+\frac{1}{8}+\frac{1}{8}} = 8.91$

These measures are used to show the **central tendency** of a set of observations.

Mean of A Random Variable

Quiz scores of five students: 10, 10, 9, 8, 8

- Mean: $\dfrac{10+10+9+8+8}{5} = \dfrac{(10)(2)+9+(8)(2)}{5} = 9$

Quiz scores of five students: 10, 10, 9, 9, 8

- Mean: $\dfrac{10+10+9+9+8}{5} = \dfrac{(10)(2)+(9)(2)+8}{5} = 9.2$

The mean, or average, is not necessarily a possible outcome for the experiment.

Mean of A Random Variable

Quiz scores of five students: 10, 10, 9, 8, 8

- Mean: $\frac{10+10+9+8+8}{5} = \frac{(10)(2)+9+(8)(2)}{5} = 9$

$$= (10)\left(\frac{2}{5}\right) + (9)\left(\frac{1}{5}\right) + (8)\left(\frac{2}{5}\right)$$

Fractions of total observations resulting in each outcome

We can calculate the **mean, or average,** of a set of data by knowing the <u>distinct values</u> that occur and their <u>relative frequencies</u>, without any knowledge of the total number of observations.

Mean of A Random Variable

This method of relative frequencies can be used to calculate the **mean of a random variable X,** or the **mean of the probability distribution X**, denoted by μ_x or simply μ.

It is also common to refer to this mean as the **mathematical expectation**, or the **expected value** of the random variable X, denoted by E(X).

Example:

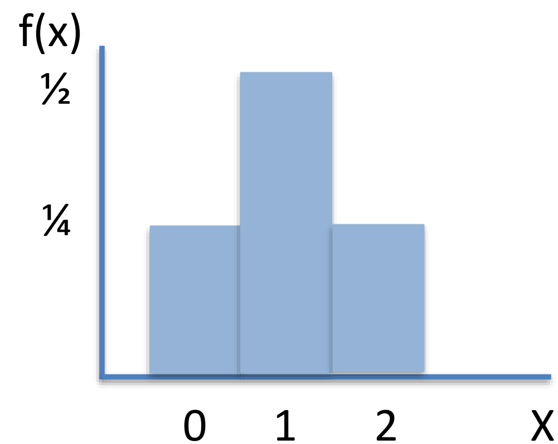
Assuming that two fair coins (red and green) were tossed, the sample space for the experiment is S = {HH, HT, TH, TT}. Define X as a random variable for the number of heads.

P(X=0) = P(TT) = ¼

P(X=2) = P(HH) = ¼

P(X=1) = P(HT)+P(TH) = ½

$$\mu_x = E(X) = (0)\left(\frac{1}{4}\right) + (1)\left(\frac{1}{2}\right) + (2)\left(\frac{1}{4}\right) = 1$$



Mean of A Random Variable

Definition:
Let X be a random variable with probability distribution f(x).
The **mean**, or **expected value**, of X is

$$\mu = E(X) = \sum_{x} x f(x) \qquad \text{(if X is discrete)}$$

$$\mu = E(X) = \int_{-\infty}^{\infty} x f(x) dx \qquad \text{(if X is continuous)}$$

The mean, or expected value, of a rv describes where the probability distribution is centered.

Mean of A Discrete RV

Example:
Suppose we are buying some new microphones from a local store. They have 4 good and 3 defective microphones. We need to randomly buy 3 from the store. What is the expected value of the number of good microphones that we buy from this store?

Let X represent the number of good components in the sample.

- Possible values of X:
- The probability distribution of X is:
- The expected value of X is:

Mean of A Discrete RV

Example:
A company produces some simple pens which are composed by the cap of the pen and the body of the pen. The probability of being defective for the cap and the body are 0.1 and 0.2, respectively. Rework a defective cap or pen body costs \$1 and \$2, respectively. Find the expected cost to rework a random pen.

Let X be the cost to rework a random pen.

- Possible values of X:
- The probability distribution of X:
- The expected value of X is:

Mean of A Continuous RV

Example:

An engineer is interested in the *mean life* of a certain type of electronic device. It is an important parameter for its evaluation. Let X be the random variable that denotes the life in hours of a certain electronic device, with the probability density function:

$$f(x) = \begin{cases} \dfrac{20{,}000}{x^3}, & x > 100 \\ 0, & elsewhere \end{cases}$$

Find the expected life of this type of device.

Expected Value of A Function

Consider a new random variable g(X), which depends on X.

- Ex. g(x) = X^2
- Whenever X assumes the value 2, g(X) assumes the value g(2)=4.
- For a discreate rv X with probability distribution f(x), for x = -1, 0, 1, 2
- Possible values of g(X):

x	0	-1	1	2
$g(x) = x^2$	0	1	1	4

- Probability distribution of g(X):

$$P[g(X) = 0] = P(X = 0) = f(0)$$
$$P[g(X) = 1] = P(X = -1) + P(X = 1) = f(-1) + f(1)$$
$$P[g(X) = 4] = P(X = 2) = f(2)$$

- Expected value of g(X):

$$\mu_{g(X)} = E[g(x)] = (0)f(0) + (1)[f(-1) + f(1)] + (4)f(2) = \sum_x g(x)f(x)$$

Expected Value of A Function

Theorem:
Let X be a random variable with probability distribution *f*(x).
The **expected value** of a random variable *g*(X) is

$$\mu_{g(X)} = E[g(X)] = \sum_x g(x)f(x) \qquad \text{(if X is discrete)}$$

$$\mu_{g(X)} = E[g(X)] = \int_{-\infty}^{\infty} g(x)f(x)dx \qquad \text{(if X is continuous)}$$

Expected Value of A Function

Example:

Suppose that the number of cars X that pass through a car wash between 4 PM and 5 PM on any sunny Friday has the following probability distribution:

x	4	5	6	7	8	9
$P(X=x)$	$\frac{1}{12}$	$\frac{1}{12}$	$\frac{1}{4}$	$\frac{1}{4}$	$\frac{1}{6}$	$\frac{1}{6}$

Let g(X) = 10(X+5) represent the amount of money, in RMB, paid to the attendant by the manager. Find the attendant's expected earnings for this particular time period.

Expected Value of A Function

Example:
Let X be a random variable with density function:

$$f(x) = \begin{cases} \dfrac{x^2}{3}, & -1 < x < 2 \\ 0, & elsewhere \end{cases}$$

Find the expected value of g(X) = 4X+3.

Expected Value of Joint RVs

Definition:
Let X and Y be random variables with joint probability distribution *f*(x, y). The **mean**, or **expected value**, of the random variable *j*(X, Y) is

$$\mu_{j(X,Y)} = E[j(X,Y)] = \sum_{x}\sum_{y} j(x,y)f(x,y) \qquad \text{(if X and Y are discrete)}$$

$$\mu_{j(X,Y)} = E[j(X,Y)] = \int_{-\infty}^{\infty}\int_{-\infty}^{\infty} j(x,y)f(x,y)\,dx\,dy \qquad \text{(if X and Y are continuous)}$$

**The definition can be generalized to calculate the mathematical expectations of functions of 2+ random variables.*

Expected Value of Joint Discrete RVs

Example:
Let X and Y be the random variables with joint probability distribution indicated in the table below. Find the expected value of j(X,Y) = XY.

	$f(x, y)$	$x = 0$	$x = 1$	$x = 2$	Row Totals
y	0	$\frac{3}{28}$	$\frac{9}{28}$	$\frac{3}{28}$	$\frac{15}{28}$
	1	$\frac{3}{14}$	$\frac{3}{14}$	0	$\frac{3}{7}$
	2	$\frac{1}{28}$	0	0	$\frac{1}{28}$
Column Totals		$\frac{5}{14}$	$\frac{15}{28}$	$\frac{3}{28}$	1

Expected Value of Joint Continuous RVs

Example:
Find E(Y/X) for the density function

$$f(x,y) = \begin{cases} \dfrac{x(1+3y^2)}{4}, & 0 < x < 2, 0 < y < 1 \\ 0, & elsewhere \end{cases}$$

Expected Value of Joint RVs

Special case:
If j(X, Y) = X, we have

$$\mu_{j(X,Y)} = E(X) = \sum_x \sum_y x f(x, y) = \sum_x x\, g(x)$$

$$g(x) = \sum_y f(x, y)$$

(if X and Y are discrete)

$$\mu_{j(X,Y)} = E(X) = \int_{-\infty}^{\infty} \int_{-\infty}^{\infty} x f(x, y) dy dx = \int_{-\infty}^{\infty} x g(x) dx$$

(if X and Y are continuous)

$$g(x) = \int_{-\infty}^{\infty} f(x, y) dy$$

where g(x) is the marginal distribution of X.

Expected Value of Joint RVs

Special case:
Similarly, if j(X,Y) = Y, we have

$$\mu_{j(X,Y)} = E(Y) = \sum_x \sum_y y f(x,y) = \sum_y y\, h(y)$$

$$h(y) = \sum_x f(x,y)$$

(if X and Y are discrete)

$$\mu_{j(X,Y)} = E(Y) = \int_{-\infty}^{\infty}\int_{-\infty}^{\infty} y f(x,y)\,dx\,dy = \int_{-\infty}^{\infty} y h(y)\,dy$$

(if X and Y are continuous)

$$h(y) = \int_{-\infty}^{\infty} f(x,y)\,dx$$

where h(y) is the marginal distribution of Y.

**To calculate E(Y) over a two-dimensional space, we can use either the joint probability distribution of X and Y or the marginal distribution of Y.*

Expected Value of Joint RVs

Revisit example:
Let X and Y be the random variables with joint probability distribution indicated in the table below. Find E(X).

	$f(x,y)$	x = 0	1	2	Row Totals
y	0	$\frac{3}{28}$	$\frac{9}{28}$	$\frac{3}{28}$	$\frac{15}{28}$
	1	$\frac{3}{14}$	$\frac{3}{14}$	0	$\frac{3}{7}$
	2	$\frac{1}{28}$	0	0	$\frac{1}{28}$
Column Totals		$\frac{5}{14}$	$\frac{15}{28}$	$\frac{3}{28}$	1

Means of Linear Combinations of RVs

Useful properties to simplify calculations (valid for both discrete and continuous rvs):

- **Theorem:**
  If a and b are constants, then $E(aX + b) = aE(X) + b$

  Proof (for the continuous case):

$$E(aX + b) = \int_{-\infty}^{\infty} (ax + b) f(x) dx = a \underbrace{\boxed{\int_{-\infty}^{\infty} x f(x) dx}}_{E(X)} + b \underbrace{\boxed{\int_{-\infty}^{\infty} f(x) dx}}_{1}$$

- **Corollaries:**
  Setting a = 0, then $E(b) = b$

  Setting b = 0, then $E(aX) = aE(X)$

Means of Linear Combinations of RVs

Revisit example:
Suppose that the number of cars X that pass through a car wash between 4 PM and 5 PM on any sunny Friday has the following probability distribution:

x	4	5	6	7	8	9
$P(X=x)$	$\frac{1}{12}$	$\frac{1}{12}$	$\frac{1}{4}$	$\frac{1}{4}$	$\frac{1}{6}$	$\frac{1}{6}$

Let $g(X) = 10(X+5)$ represent the amount of money, in RMB, paid to the attendant by the manager. Find the attendant's expected earnings for this particular time period.

Means of Linear Combinations of RVs

Revisit example:

Let X be a random variable with density function:

$$f(x) = \begin{cases} \dfrac{x^2}{3}, & -1 < x < 2 \\ 0, & elsewhere \end{cases}$$

Find the expected value of g(X) = 4X+3.

Means of Linear Combinations of RVs

- **Theorem:**

  The expected value of the sum or difference of two or more functions of a random variable X is the sum or difference of the expected values of the functions. That is,

$$E(r_1(X) \pm r_2(X)) = E[r_1(X)] \pm E[r_2(X)]$$

Proof (for the continuous case):

$$E[r_1(X) \pm r_2(X)] = \int_{-\infty}^{\infty} [r_1(x) \pm r_2(x)] f(x) dx$$

$$= \underbrace{\boxed{\int_{-\infty}^{\infty} r_1(x) f(x) dx}}_{E[r_1(X)]} \pm \underbrace{\boxed{\int_{-\infty}^{\infty} r_2(x) f(x) dx}}_{E[r_2(X)]}$$

Means of Linear Combinations of RVs

Example:
Let X be a random variable with probability distribution as follows:

x	0	1	2	3
$f(x)$	$\frac{1}{3}$	$\frac{1}{2}$	0	$\frac{1}{6}$

Find the expected value of $Y = (X - 1)^2$.

Means of Linear Combinations of RVs

Example:

The weekly demand for a certain drink, in thousands of liters, at a chain of convenience stores is a continuous random variable $g(X) = X^2 + X - 2$, where X has the density function:

$$f(x) = \begin{cases} 2(x-1), & 1 < x < 2 \\ 0, & elsewhere \end{cases}$$

Find the expected value of the weekly demand for the drink.

Means of Linear Combinations of RVs

- **Theorem:**
  The expected value of the sum or difference of two or more functions of the random variables X and Y is the sum or difference of the expected values of the functions. That is,

$$E[r_1(X,Y) \pm r_2(X,Y)] = E[r_1(X,Y)] \pm E[r_2(X,Y)]$$

- **Corollaries:**
  Setting $r_1(X,Y) = g(X)$ and $r_2(X,Y) = h(Y)$, we see that,

$$E[g(X) \pm h(Y)] = E[g(X)] \pm E[h(Y)]$$

  Setting $r_1(X,Y) = X$ and $r_2(X,Y) = Y$, we see that,

$$E(X \pm Y) = E(X) \pm E(Y)$$

Means of Linear Combinations of RVs

- **Theorem:**

  Let X and Y be two independent random variables. Then,

$$E(XY) = E(X)E(Y)$$

Proof (for the continuous case):

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

X and Y are independent

$$\boxed{f(x,y) = g(x)h(y)}$$

$$= \int_{-\infty}^{\infty}\int_{-\infty}^{\infty} xyg(x)h(y)dxdy$$

$$= \underbrace{\boxed{\int_{-\infty}^{\infty} xg(x)dx}}_{E(X)}\ \underbrace{\boxed{\int_{-\infty}^{\infty} yh(y)dy}}_{E(Y)}$$

Means of Linear Combinations of RVs

Example:
Let X and Y be two independent random variables with the joint density function:

$$f(x) = \begin{cases} \dfrac{x(1+3y^2)}{4}, & 0 < x < 2, 0 < y < 1 \\ 0, & elsewhere \end{cases}$$

Show that E(XY) = E(X)E(Y).