MATH 3339
Statistics for the Sciences
Sec 4.1-4.3

Wendy Wang, Ph.D.
wwang60@central.uh.edu

Lecture 6 - 3339

UNIVERSITY of HOUSTON
DEPARTMENT OF MATHEMATICS

Outline

1. Discrete Random Variables

2. Expected Values of Discrete Variables

Chapter 4

Discrete Distributions

UNIVERSITY of HOUSTON
DEPARTMENT OF MATHEMATICS

Random Variables

Suppose an experiment is conducted. A **random variable** is a function that assigns values to the outcomes of the experiment.
Example: A coin is flipped resulting in either heads or tails and the random variable X is defined by

$$X(heads) = 1 \qquad X(tails) = 0$$

We say that the random variable X indicates heads.

Example: A fair six sided die is tossed, let the random variable X indicate the event that an even number is rolled. What is $X(2)$? $X(3)$?

✓ $X=1$ if even number

$X(2) = 1$ $\quad$ $X(3) = 0$

$X(5) = 0$

UNIVERSITY of HOUSTON
DEPARTMENT OF MATHEMATICS

Example

Suppose we have 2 gas stations having six pumps each. Define a random variable X = total number of pumps in use at the two stations What are the possible values of X?

X can be $0, 1, 2, 3, \ldots\ldots, 11, 12$

Define a random variable Y = total number of pumps not currently being used at the two stations What are the possible values of Y? How are X and Y related?

Y can be $0, 1, 2, \ldots\ldots, 12$

$$X + Y = 12$$

Example

Suppose pump 1 at gas station number one is observed as the next customer begins to use it. Let T = the length of time the customer is at the pump. What are the possible values of T?

$T \geq 0$

continuous distribution (chapter 5)

UNIVERSITY of HOUSTON
DEPARTMENT OF MATHEMATICS

Random Variables

There are two types of random variables:

- **Discrete Random Variables:** A random variable whose possible values are either finite or may be listed.
- **Continuous Random Variables:** A random variable whose possible values consists of an interval of values or several intervals of values and for which the probability of any one value is zero.

For continuous random variables X, they take on all values in an interval of numbers. In fact, the probability of X equaling an individual number is **0**! The probability distribution of X is described by a density curve. The probability of any event is the area under the curve and above the values of X that make up the event.

Which of the above variables were discrete? Which were continuous?

Discrete Random Variables

Def: The **probability mass function** (pmf) of a discrete random variable (r.v.) is defined for every number x_i by $f(x_i) = P(X = x_i)$.

Example: Six lots of components are ready to be shipped by a certain supplier. The number of defective components in each lot is as follows:

Lot	1	2	3	4	5	6
No. Defective	0	2	0	1	2	0

$X =$

One of these lots is to be randomly selected for shipment to a particular customer. Let X = number of defectives in the selected lot. What are the possible values of X? Determine the values of the pmf.

X	0	1	2
$P(X=x_i)$	$P(X=0) = \frac{3}{6} = \frac{1}{2}$	$P(X=1) = \frac{1}{6}$	$P(X=2) = \frac{2}{6} = \frac{1}{3}$

Discrete Random Variables

Properties of f:

- $f(x) \geq 0$ for all $x \in \mathbb{R}$
- $\sum_i f(x_i) = 1$
- $P(X \in A) = \sum_{x \in A} f(x)$, where $A \subset \mathbb{R}$ is a discrete set.

Suppose we toss a fair coin 10 times. Let X = number of heads in the 10 tosses.What are the possible values of X?

X can be 0, 1, 2, ... 10

How many heads do we expect to get in the 10 tosses?

5

What is $f(0)$? $f(1)$?

$f(0) = P(X=0) = 0.5 * 0.5 * 0.5 \cdots\cdots * 0.5 = 0.5^{10}$

$f(3) \cdots \cdots ?$

$f(10) = P(X=10) = 0.5 * 0.5 * \cdots = 0.5^{10}$

Discrete Random Variables

pmf: $f(4) = P(X=4)$

cdf: $F(4) = P(X \le 4)$

Example: Suppose you are given the following distribution table:

X	1	2	3	4	5	6	7
P(X)	0.15	0.05	0.10	0.30	0.10	0.15	0.15

Find the following:

$\sum P(X=x) = 1$

$P(X = 4)$ $= 1 - .15 - 0.05 - \cdots - 0.15 = \boxed{0.3}$

$P(X < 2)$ $= P(X \text{ can be } 1) = P(X=1) = 0.15$

$P(2 < X \le 5)$ $= P(X \text{ can be } 3, 4, 5) = P(X=3) + P(X=4) + P(X=5)$
$= 0.10 + 0.30 + 0.10$
$= 0.50$

$P(X > 3)$
$= P(X \text{ can be } 4, 5, 6, 7)$
$= 0.30 + 0.10 + 0.15 + 0.15 = 0.7$

$P(X > 3) = 1 - P(X \le 3) = 1 - P(X = 1, 2, 3) = \boxed{0.7}$

UNIVERSITY of HOUSTON
DEPARTMENT OF MATHEMATICS

The Cumulative Distribution Function: Discrete R.V.

Def: The **cumulative distribution function** (cdf) $F(x)$ of a discrete rv X with pmf $f(x)$ is defined for every number x by $F(x) = P(X \leq x)$.

For any number x, $F(x)$ is the probability that the observed value of X will be at most x.

pmf: probability mass function

$$f(x) = P(X = x)$$

cdf: cumulative distribution function

$$F(x) = P(X \leq x)$$

UNIVERSITY of HOUSTON
DEPARTMENT OF MATHEMATICS

$F(1) = P(X \le 1) = 0.05$
$F(2) = P(X \le 2) = 0.15$
$F(4) = P(X \le 4) = 0.5$

Example: A store carries flash drives with either 1 GB, 2 GB, 4 GB, 8 GB, or 16 GB of memory. The table below gives the distribution of the rv X = the amount of memory in a purchased drive.

$P(17.9) = 1$

$F(1.6) = F(1)$

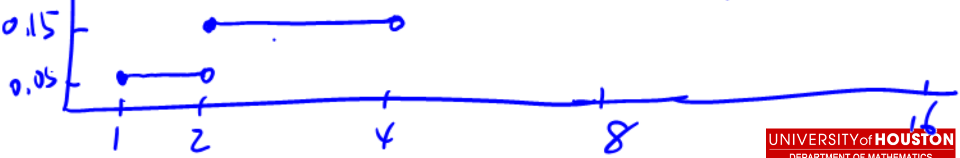
X	1	2	4	8	16
P(X)	0.05	0.10	0.35	0.40	0.10

Determine the values of F for the possible values of X.

What is the value of $F(2.5)$?

$F(2.5) = P(X \le 2.5)$
$= P(X \text{ can be } 1, 2)$
$= P(X=1) + P(X=2) = \boxed{0.15}$

Sketch a graph of F.



Discrete Random Variables

Example: Suppose the random variable X takes on possible values $x = 0, 1, 2, 3$ and has pmf given by $f(x) = \frac{x+1}{k}$, determine the value of k.

X	0	1	2	3
$f(x) = \frac{x+1}{k}$	$\frac{0+1}{k} = \frac{1}{k}$	$\frac{1+1}{k} = \frac{2}{k}$	$\frac{2+1}{k} = \frac{3}{k}$	$\frac{3+1}{k} = \frac{4}{k}$

$$\frac{1}{k} + \frac{2}{k} + \frac{3}{k} + \frac{4}{k} = 1$$

$$\frac{1+2+3+4}{k} = \frac{10}{k} = 1 \Rightarrow k = 10$$

UNIVERSITY of HOUSTON
DEPARTMENT OF MATHEMATICS

Expected Values of Discrete Variables

UNIVERSITY of HOUSTON
DEPARTMENT OF MATHEMATICS

Expected Values

The **expected value** or mean of the distribution of a random variable X is given by:

$$E[X] = \mu = \sum_x x \cdot f(x) = \sum_{i=1}^{n} x_i \cdot p_i$$

Example: From a previous example where we had number of defectives per lot, find the expected number of defective items.

Lot	1	2	3	4	5	6
No. Defective	0	2	0	1	2	0

UNIVERSITY of HOUSTON
DEPARTMENT OF MATHEMATICS

Expected Values

Another Example: Consider the table below which gives the number of years required to obtain a Bachelor's degree for graduates of high school A, and the number of students who needed each:

Years	3	4	5	6
No. of Students	17	23	38	19

How would compute the average number of years required by graduates of high school A?

UNIVERSITY of HOUSTON
DEPARTMENT OF MATHEMATICS

Properties of Expected values

- $E[c] = c$ for any constant $c \in \mathbb{R}$
- $E[aX + b] = aE[X] + b$
- $E[aX + bY] = aE[X] + bE[Y]$
- $E[h(x)] = \sum_x h(x) \cdot f(x)$

Variance

The variance of a rv X is

$$\sigma^2 = Var[X] = E[(X-\mu)^2] = E[X^2] - E[X]^2$$

UNIVERSITY of HOUSTON
DEPARTMENT OF MATHEMATICS

Expected Values and Variance

Properties of Expected values and Variance

- $E[c] = c$ for any constant $c \in \mathbb{R}$
- $E[aX + b] = aE[X] + b$
- $E[aX + bY] = aE[X] + bE[Y]$
- $E[h(x)] = \sum_x h(x) \cdot f(x)$
- $Var[aX + b] = a^2 Var[X]$
- $Var[aX + bY] = a^2 Var[X] + b^2 Var[Y]$

UNIVERSITY of HOUSTON
DEPARTMENT OF MATHEMATICS

Expected Values and Variance $Var(X) = E(X^2) - [E(x)]^2$

Example: Let X have pmf. given by

x	1	2	3	4
f(x)	0.4	0.2	0.3	0.1

Determine $E[X]$, $E[X^2]$, $Var[X]$ and the standard deviation of X.

$$E(x) = \mu = \sum x \cdot f(x)$$
$$= 1 * 0.4 + 2 * 0.2 + 3 * 0.3 + 4 * 0.1$$
$$= \boxed{2.1}$$

$$E(X^2) = \sum x^2 \cdot f(x)$$
$$= 1^2 * 0.4 + 2^2 * 0.2 + 3^2 * 0.3 + 4^2 * 0.1$$
$$= 5.5$$

UNIVERSITY of HOUSTON
DEPARTMENT OF MATHEMATICS

$$\mathrm{Var}(X) = \underbrace{E(X^2)} - (E(X))^2$$
$$= 5.5 - (2.1)^2$$
$$= \boxed{1.09}$$

$$\mathrm{sd}(X) = \sqrt{\mathrm{Var}(X)}$$
$$= \sqrt{1.09} = \boxed{1.044}$$

Expected Values and Variance

Example: Determine the expected value and variance of the rv Y defined by $Y = 5X - 1$, where X is given in the previous problem.

$$E(Y) = E(5X-1)$$
$$= 5 * E(X) - 1$$
$$= 5 * 2.1 - 1 = \boxed{9.5}$$

$$Var(Y) = Var(5X-1)$$
$$= 5^2 Var(X) = 25*(1.09)$$
$$= \boxed{27.25}$$

UNIVERSITY of HOUSTON
DEPARTMENT OF MATHEMATICS

Expected Values and Variance

Example: A store carries flash drives with either 1 GB, 2 GB, 4 GB, 8 GB, or 16 GB of memory. The table below gives the distribution of the rv X = the amount of memory in a purchased drive.

X	1	2	4	8	16
P(X)	0.05	0.10	0.35	0.40	0.10

Determine $E[X]$ and $Var[X]$ using R.

Finding Expected Value and Standard Deviation in R

```
> X=c(1,2,4,8,16)
> PX=c(.05,.10,.35,.40,.10)
> EX=sum(X*PX)
> EX
[1] 6.45
> VarX=sum(X^2*PX)-(EX)^2
> VarX
[1] 15.6475
```

UNIVERSITY of HOUSTON
DEPARTMENT OF MATHEMATICS