Discrete Probability Distributions

Section 3.1

Cathy Poliak, Ph.D.
cathy@math.uh.edu
Office hours: T Th 2:30 - 5:15 pm 620 PGH

Department of Mathematics
University of Houston

February 9, 2016

Outline

UNIVERSITY of HOUSTON
DEPARTMENT OF MATHEMATICS

Popper Set Up

- Fill in all of the proper bubbles.
- Use a #2 pencil.
- This is popper number 03.

Rules of Probability

1. The probability $P(E)$ of any event E satisfies $0 \leq P(E) \leq 1$.
2. If S is the sample space in a probability model, then $P(S) = 1$.
3. **Complement rule**: For any event E, $P(E^C) = 1 - P(E)$.
4. **General rule for addition**: For any two events E and F, $P(E \cup F) = P(E) + P(F) - P(E \cap F)$.
5. **General rule for multiplication**: For any two events E and F, $P(E \cap F) = P(E) \times P(F|E)$ or $P(E \cap F) = P(F) \times P(E|F)$.
6. **Conditional probability**: For any two events E and F, P(F, given E) = $P(F|E) = \frac{P(F \cap E)}{P(E)}$.

UNIVERSITY of HOUSTON
DEPARTMENT OF MATHEMATICS

Example of Rules

The table below gives the results of a survey of the diet and exercise habits of 1200 adults:

	Diet	Don't diet	Total
Exercise	315	165	480
Don't exercise	585	135	720
Total	900	300	1200

1. What is the probability that someone in this group exercises?

2. What is the probability that a dieter is also an exerciser?

UNIVERSITY of HOUSTON
DEPARTMENT OF MATHEMATICS

Random Variables

- A **random variable** is a variable whose value is a numerical outcome of a random phenomenon.
- Examples
  - X = the sum of two dice, $X = 2, 3, 4, \ldots, 12$.
  - X = the number of customers that order a muffin in a coffee shop between 7:00 am and 9:00 am, $X = 0, 1, \ldots$.
  - X = weight of a box of Lucky Charms, $X \geq 0$.
- Random variables can either be **discrete** or **continuous**.

Discrete random variables

- **Discrete random variables** has either a finite number of values or a countable number of values, where countable refers to the fact that there might be infinitely many values, but they result from a counting process.
- Example of discrete random variable:
  - The sum of two dice.
  - The number of customers who order a muffin in a coffee shop between 7:00 am and 9:00 am.
- The possible values for a discrete random variable has "gaps" between each value.

UNIVERSITY of HOUSTON
DEPARTMENT OF MATHEMATICS

Continuous random variables

- **Continuous random variables** are random variables that can assume values corresponding to any of the points contained in one or more intervals.
- Example of continuous random variable:
  - The weight of a box of Lucky Charms.
- We want to know how both of these types of random variables are **distributed**.

UNIVERSITY of HOUSTON
DEPARTMENT OF MATHEMATICS

Probability Distribution

The **probability distribution** of a random variable X tells us what values X can take and how to assign probabilities to those values. This is the "ideal" distribution for a random variable. Requirements for a probability distribution:

1. The sum of all the probabilities equal 1.
2. The probabilities are between 0 and 1, including 0 and 1.

We will determine the three characteristics of a probability distribution: shape, center, and spread.

UNIVERSITY of HOUSTON
DEPARTMENT OF MATHEMATICS

Popper 03 Question #1

Many stores run ”secret sales“: Shoppers receive cards that determine how large a discount they get, but the percentage is revealed by scratching off that black stuff only after the purchase has been totaled at the cash register. the store is required to reveal the probability distribution of discount available. Which one of these is a correct probability distribution?

	10% off	20% off	30% off	50% off
a)	0.2	0.2	0.2	0.2
b)	0.5	0.3	0.2	0.1
c)	0.8	0.1	0.05	0.05
d)	0.75	0.25	0.25	-0.25
e)	1	0	0	0

a. only a) b. only b) c. only c) d. both c) and e)

UNIVERSITY of HOUSTON
DEPARTMENT OF MATHEMATICS

Discrete Probability Distribution

- The probability distribution of a discrete random variable X lists the possible values of X and their probabilities.
- A **probability distribution table of** X consists of all possible values of a discrete random variable with thier corresponding probabilities.

UNIVERSITY of HOUSTON
DEPARTMENT OF MATHEMATICS

Example of Discrete Probability Distribution

The following is a probability distribution of the number of repairs needed for a car.

X	0	1	2	3
$P(X)$	0.72	0.17	0.07	0.04

- $P(X \geq 2) = P(X = 2) + P(X = 3) = 0.11$
- $P(X > 2) = P(X = 3) = 0.04$
- $P(X \geq 1) = 1 - P(X = 0) = 1 - 0.72 = 0.28$

UNIVERSITY of HOUSTON
DEPARTMENT OF MATHEMATICS

Popper 03 Questions

Suppose you are given the following probability distribution table. Determine the probabilities below.

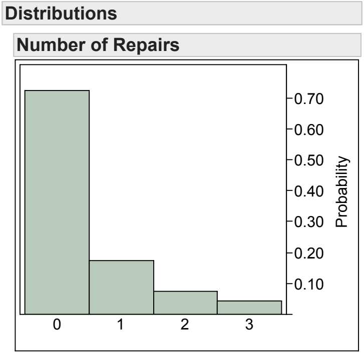
X	1	2	3	4	5	6	7
$P(X)$	0.15	0.05	0.10	?	0.10	0.15	0.15

2. $P(X = 4)$ a. 0 b. 1 c. 0.7 d. 0.3
3. $P(X < 2)$ a. 0.15 b. 0.20 c. 0.85 d. 0.80
4. $P(2 < X \leq 5)$ a. 0.05 b. 0.55 c. 0.5 d. 0.2
5. $P(X > 3)$ a. 0.1 b. 0.7 c. 0.8 d. 0.2

UNIVERSITY of HOUSTON
DEPARTMENT OF MATHEMATICS

Probability Histogram

One way to see the probability distribution of a discrete random variable is using a probability histogram. The following is a histogram for X = the number of repairs needed for a car. What is the "shape" of this distribution.



UNIVERSITY of HOUSTON
DEPARTMENT OF MATHEMATICS

Determining Center of a Discrete Random Variable

Suppose that X is a discrete random variable whose distribution is

Values of X	x_1	x_2	x_3	$\cdots$	x_k
Probability	p_1	p_2	p_3	$\cdots$	p_k

To find the **mean** of the random variable X, multiply each possible value by its probability, then add all the products:

$$\mu_X = E[X] = x_1p_1 + x_2p_2 + x_3p_3 + \cdots + x_kp_k.$$

This is also called the **expected value** $E[X]$.
Note: The list here is not a list of observations but a list of all possible outcomes. So we are finding μ, the population mean not $\bar{x}$, the sample mean.

UNIVERSITY of HOUSTON
DEPARTMENT OF MATHEMATICS

Example 2 for Expected Value

For the example of the number of repairs for a car.

X	$P(X)$	$X \times P(X)$
0	0.72	
1	0.17	
2	0.07	
3	0.04	
		Sum =

UNIVERSITY of HOUSTON
DEPARTMENT OF MATHEMATICS

Popper 03 Question

Let X=The number of traffic accidents daily in a small city. The following table is the probability distribution for X.

X	Probability
0	0.10
1	0.20
2	0.45
3	0.15
4	0.05
5	0.05

6. Compute the expected number of accidents per day, $E[X]$.

a. 2.5 b. 2 c. 3 d. 0.1667

UNIVERSITY of HOUSTON
DEPARTMENT OF MATHEMATICS

Determining the spread of a Discrete Random Variable

The **variance** of a discrete random variable X is

$$\begin{aligned}\sigma_X^2 &= (x_1-\mu_X)^2 p_1 + (x_2-\mu_X)^2 p_2 + \cdots + (x_k-\mu_X)^2 p_k \\ &= \sum_{i=1}^{k} (x_i-\mu_X)^2 p_i\end{aligned}$$

The standard deviation of X is the square root of the variance $\sigma_X = \sqrt{\sigma_X^2}$.

UNIVERSITY of HOUSTON
DEPARTMENT OF MATHEMATICS

Easier Calculation for Variance and Standard Deviation from a Discrete Probability Distribution

$$\text{VAR(X)} = \sigma_x^2 = E(X^2) - [E(X)]^2$$

Where $E(X^2) = x_1^2 P(X_1) + x_2^2 P(X_2) + x_3^2 P(X_2) + \ldots + x_n^2 P(X_n)$.

UNIVERSITY of HOUSTON
DEPARTMENT OF MATHEMATICS

Example of standard deviation

Determine the standard deviation for the number of repairs needed for a car using the following probability distribution.

X	$P(X)$
0	0.72
1	0.17
2	0.07
3	0.04

Example of standard deviation

Step 1: Determine the mean (expected value) of the probability distribution, μ_X. $\mu_X = E(X) = 0.43$

X	$P(X)$
0	0.72
1	0.17
2	0.07
3	0.04

UNIVERSITY of HOUSTON
DEPARTMENT OF MATHEMATICS

Example of standard deviation

Step 2: Determine the expected value of X^2.

X	$P(X)$	X^2	$X^2 \times P(X)$
0	0.72	0	
1	0.17	1	
2	0.07	4	
3	0.04	9	
			$E(x^2) =$

UNIVERSITY of HOUSTON
DEPARTMENT OF MATHEMATICS

Example of standard deviation

Step 3: Find the variance by taking the difference of the expected value of X^2 and the square of the mean (expected value of X).

$$VAR(X) = \sigma_X^2 = E[X^2] - (E[X])^2 = 0.81 - 0.43^2 = 0.6251$$

Step 4: Take the square root of the variance. This is the standard deviation.

$$SD(X) = \sigma_X = \sqrt{0.6251} = 0.70963$$

UNIVERSITY of HOUSTON
DEPARTMENT OF MATHEMATICS

Meanings of σ_X and μ_X

- The mean of the random variable X = number of repairs for a vehicle is $\mu_X = 0.43$ and the standard deviation of X is $\sigma_X = 0.79$.
- This implies that the expected number of repairs for a vehicle is 0.43 repairs, give or take 0.79 repairs or so.

UNIVERSITY of HOUSTON
DEPARTMENT OF MATHEMATICS

Popper 03 Question

7. Determine the standard deviation of the number of accidents on a given day. The expected value is $E[X] = 2$.

X	0	1	2	3	4	5
$P(X)$	0.10	0.20	0.45	0.15	0.05	0.05

a. 19
b. 1.4
c. 0.2333
d. 1.1832

UNIVERSITY of HOUSTON
DEPARTMENT OF MATHEMATICS

Example

Suppose that in the small city on a given day there was rain. So there would be at least one accident, the probability distribution of the number of accidents will be:

Number of accidents	1	2	3	4	5	6
Probability	0.10	0.20	0.45	0.15	0.05	0.05

- Compute the mean number of accidents.

- Compute the variance of the number of accidents.

UNIVERSITY of HOUSTON
DEPARTMENT OF MATHEMATICS

Rule 1a for Means and Variances

- If X is any random variable and a is a fixed numbers that is added to all of the values of the random variable then
- the mean increases by that number:

$$E[a + X] = a + E[X].$$

- the variance remains the same:

$$\sigma^2_{a+X} = \sigma^2_X.$$

- In the example of a rainy day, we added 1 accident to each value, $1 + X$ and the probabilities remained the same.

$$\begin{aligned} E[1 + X] &= 1 + E[X] = 1 + 2 = 3 \\ \sigma^2_{1+X} &= \sigma^2_X = 1.4 \end{aligned}$$

Rule 1b for Means and Variances

- If X is any random variable and b is a fixed numbers that is multiplied to all of the values of the random variable then
- the mean is changed by that multiplier:

$$E[bX] = bE[X].$$

- the variance is also changed by the square of the multiplier:

$$\sigma^2_{bX} = b^2\sigma^2_X.$$

UNIVERSITY of HOUSTON
DEPARTMENT OF MATHEMATICS

Playing Roulette

Suppose you play the Roulette wheel in Las Vegas and bet \$10 on red. Let X = the amount won, the probability distribution is as follows:

Winnings	P(X)
10	0.4737
-10	0.5263

The expected winnings is $\mu_X = -\$0.53$ with a standard deviation of $\sigma_X = \$10$. Suppose we double the bet to \$20. That is $2X$, using rule 1b determine the following.

1. Calculate the expected winnings, $E[2X]$.
2. Calculate the variance of the winnings, $\sigma^2_{(2X)}$.
3. Calculate the standard deviation of the winnings, $\sigma_{(2X)}$.

UNIVERSITY of HOUSTON
DEPARTMENT OF MATHEMATICS

Rule 1 for Means and Variances

Suppose X is a random variable with mean $E[X]$ and variance σ_X^2, and we define W as a new variable such that $W = a + bX$, where a and b are real numbers. We can find the mean and variance of W by:

$$\begin{aligned} E[W] &= E[a+bX] = a + bE[X] \\ \sigma_W^2 &= Var[W] = Var[a+bX] = b^2 Var[X] \\ \sigma_W &= SD[W] = \sqrt{Var[W]} = \sqrt{b^2 Var[X]} = |b|(SD[X]) \end{aligned}$$

UNIVERSITY of HOUSTON
DEPARTMENT OF MATHEMATICS

#14 from text

Suppose you have a distribution X, with mean = 22 and standard deviation = 3. Define a new random variable $Y = 3X + 1$.

1. Find the variance of X.

2. Find the mean of Y.

3. Find the variance of Y.

UNIVERSITY of HOUSTON
DEPARTMENT OF MATHEMATICS

Rule 2 for Means

- If X and Y are two different random variables, then the mean of the sums of the pairs of the random variable is the same as the sum of their means:

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

This is called the addition rule for means.

- The mean of the difference of the pairs of the random variable is the same as the difference of their means:

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

UNIVERSITY of HOUSTON
DEPARTMENT OF MATHEMATICS

Rule 2 for Variances

If X and Y are independent random variables

$$\sigma^2_{X+Y} = Var[X+Y] = Var[X] + Var[Y].$$

and

$$\sigma^2_{X-Y} = Var[X-Y] = Var[X] + Var[Y].$$

UNIVERSITY of HOUSTON
DEPARTMENT OF MATHEMATICS

Example of Rule 2

Tamara and Derek are sales associates in a large electronics and appliance store. The following table shows their mean and standard deviation of daily sales. Assume that daily sales among the sales associates are independent.

Sales associate	Mean	Standard deviation
Tamara X	$E[X] = \$1100$	$\sigma_X = \$100$
Derek Y	$E[Y] = \$1000$	$\sigma_Y = \$80$

1. Determine the mean of the total of Tamara's and Derek's daily sales, $E[X + Y]$.
2. Determine the variance of the total of Tamara's and Derek's daily sales, $\sigma^2_{(X+Y)}$.
3. Determine the standard deviation of the total of Tamara's and Derek's daily sales, $\sigma_{(X+Y)}$.

UNIVERSITY of HOUSTON
DEPARTMENT OF MATHEMATICS

Example