CHAPTER 4 : DISCRETE PROBABILITY DISTRIBUTIONS

Probability distributions can be represented by tables or by formulas.
The simplest type of probability distribution can be displayed in a table.

Discrete Probability Distributions using PDF Tables

EXAMPLE D1:   Students who live in the dormitories at a certain four year college must buy a meal plan.
They must select from four available meal plans: 10 meals, 14 meals, 18 meals, or 21 meals per week.
The Food and Housing Office has determined that the 15% of students purchase 10 meal plan,
45% purchase the 14 meal plan of students, 30% purchase the 18 meal plan ,10% purchase the 21 meal plan.

a. What is the random variable? X = _____

_____

Notation: P(Event) = probability value
  P(X = 10) is the probability that a student purchases a meal plan with 10 meals per week
  P(X >14)   is the probability that a student purchases a meal plan with more than 14 meals per week

b. Make a table that shows the probability distribution

This table is called the PDF   *We can create an extra column next to*
  Probability Distribution Function   *the PDF table to help calculate the mean*

x =Number of Meals	Probability P(x)
10	
14	
18	
21	

xP(x)

c. Find the probability that a student purchases more than 14 meals:

d. Find the probability that a student does not purchase 21 meals.

e. On average, how many meals does a student purchase per week in their meal plan?

  Calculate the mean.   Mean = Expected Value: μ = ΣxP(x) μ = _____

f. Write a sentence that interprets the mean in the context of the problem.

NOTE that it is acceptable that the mean is not whole number; it can have a fraction or a decimal.

CHAPTER 4 : DISCRETE PROBABILITY DISTRIBUTIONS
Discrete Probability Distributions using PDF Tables

- **PDF: Probability Distribution Function**

   All probabilities are between 0 and 1, inclusive   AND   All probabilities must sum to 1.

- **Mean = Expected Value = μ = $\Sigma xP(x)$**

   Interpreted as a long term average over many observations
   Formula is a "weighted" average where each value is "weighted" according to how likely is its to occur

- **Standard Deviation = σ = $\sqrt{\sum (x-\mu)^2 P(x)}$**

   Measures variation in the probability distribution
   Formula is a "weighted" average of the squared distances between each data value and the mean

- **Variance = (standard deviation)^2 σ^2 = $\sum (x-\mu)^2 P(x)$**

   Measures variation in the probability distribution
   Before widespread technology, variance was easier to calculate than standard deviation
   Variance is used in some types of "statistical tests" instead of standard deviation

---

EXAMPLE D2: The Highway Commissioner wants to know how many people are in cars that use the carpool lanes on Ocean Expressway. The best way to estimate occupancy of a car accurately is with human observers; electronic methods often are not accurate.
A team of "observers" is sent to watch the highway from overpasses and count the number of occupants in a sample of cars passing below the overpass. Based on the data, the number of occupants in cars in the carpool lane on Ocean Expressway follows the probability distribution below.

X	P(X)				
1	0.1				
2	0.5				
3	0.3				
4	0.1				

a.   Draw a relative frequency histogram
 of this probability distribution

b. Find the expected value and
write a sentence that interprets its meaning
in the context of the problem


```
0     1     2     3     4 people
```
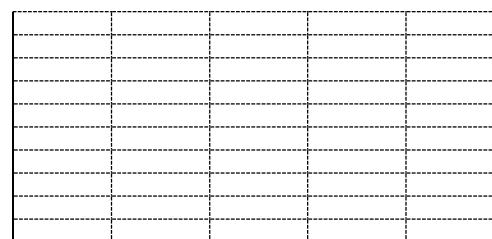
c. Find the variance $\sigma^2 = \sum (x-\mu)^2 P(x)$ and the standard deviation: $\sigma = \sqrt{\sum (x-\mu)^2 P(x)}$

   by doing the calculation step by step using the table.
   NOTE: We'll do this in class together once. After that we'll always use our calculator to do this for us.

d. Use your calculator 1VarStats to find the mean, standard deviation , and variance.

CHAPTER 4: DISCRETE PROBABILITY DISTRIBUTIONS USING PDF TABLES

EXAMPLE D3:   *At the county fair, a booth has a coin flipping game.*
You pay $1 to flip **three fair coins**.
If the result contains three heads, you win $4.   If the result is two heads, you win $1. Otherwise there is no prize.
We are interested in the net amount of money gained or lost in one game.
a. Define the random variable and the values it can have.
b. Write the PDF for the amount gained or lost in one game.
c. Find the expected value for this game (Expected NET GAIN OR LOSS)
d. Find the expected total net gain or loss if you play this game 50 times.

EXAMPLE D4:   Suppose you play a different game. In this game, you flip a **biased coin** twice.
A biased or unfair coin has different probabilities for landing on heads and tails.
Suppose that for this coin, P(HEAD) = 2/3 and P(TAIL) = 1/3.
In this game you do not pay in order to play.
You toss the coin twice, and then win or lose according to the following:
        win $3 if you toss two tails
        win $1 if you toss two heads
        pay (lose) $2 if you toss one head and one tail.
We are interested in the net amount of money gained or lost in one game.
a. Define the random variable and the values it can have.
b. Write the PDF for the amount gained or lost in one game.
c. Find the expected value for this game (Expected NET GAIN OR LOSS)

CHAPTER 4 : DISCRETE PROBABILITY DISTRIBUTIONS USING PDF TABLES

EXAMPLE D5: In this game we roll ONE fair EIGHT SIDED DIE once.(The eight sides of the die are
numbered 1, 2, 3, 4, 5, 6, 7, 8 and the die has an equal chance of landing on each side.)
Suppose that you win $6 if you roll an 8, win $2.50 if you roll a 2,
lose $2 if you roll an odd number, and
if you roll a 4 or 6 you neither win anything nor lose anything.
We are interested in the monetary outcome for one game.
a. Define the random variable and the values it can have.
b. Write the PDF for the amount gained or lost in one game.
c. Find the expected value for this game (Expected NET GAIN OR LOSS)
d. Find the expected total net gain or loss if you play this game 100 times.

EXAMPLE D6 A real estate developer is presenting plans to the Planning Commissioner for a development of
houses and apartments he proposes to build. He needs to estimate the impact on the local schools so he must
estimate the number of children expected to attend the schools. He hires a statistician who studies the
demographics of the neighborhood and of similar housing developments; she provides the estimates below.
Let X = the number of school age children per household.
a. Find the expected number of children per household
b. Find the expected number of school age children in the new housing development if 120 housing units are built.

X	P(X)
0	0.30
1	0.18
2	
3	0.15
4	0.08
5	0.05
6 ore more	0

CHAPTER 4 : BINOMIAL PROBABILITY DISTRIBUTION

The Binomial Distribution is a special discrete probability distribution that arises often in problems.
A BINOMIAL probability experiment has
- a fixed number **n** of repeated trials
- each trial has outcomes that we can classify as "success or "failure"
- outcome of trials are independent (*Outcome of a trial does not influence outcome of future trials*)
The probability of success on a single trial, **p**, is constant (the same) for all trials

We are interested in the number of successes, x, in n trials   Notation: X~B(n,p)

EXAMPLE B1: A college claims that 70% of students receive financial aid. Suppose that 4 students at the college are randomly selected. We are interested in the number of students in the sample who receive financial aid.

X = _____

p = the probability that a student receives financial aid: p = _____   q = 1−p = _____
X ~ B(4, 0.7) : Binomial with n = 4 and p = 0.7

X	P(x)
0	
1	
2	
3	
4	

Ways to get x successes in n trials

n = 4	n = 4	n = 4
x = 1	**x = 2**	**x = 3**
Abcd	**AB**cd	a**BCD**
a**B**cd	**A**b**C**d	**A**b**CD**
ab**C**d	**A**bc**D**	**AB**c**D**
abc**D**	a**BC**d	**ABC**d
	a**B**c**D**	
	ab**CD**	

a. Find the probability that AT MOST 2 of the students in the sample receive financial aid:

b. Find the probability that AT LEAST 3 of the students in the sample receive financial aid:

c. Find the **mean and the standard deviation** using the **shortcut formulas for the binomial distribution**:

$$\mu = np \quad ; \quad \sigma = \sqrt{npq} \quad \text{where } q = 1 - p \quad \textbf{only for Binomial distribution.}$$

These shortcut formulas for μ and σ give the same results as the definitions $\mu=\Sigma xP(x)$, $\sigma = \sqrt{\sum (x-\mu)^2 P(x)}$

Formulas for Binomial Distribution: **X~B(n,p)** $P(X = x) = {_nC_x}\, p^x\, (1-p)^{n-x}$

$P(X = x)$ is the probability of obtaining x successes in n independent trial

$\mu = np \quad ; \quad \sigma = \sqrt{npq}$ where $q = 1 - p$ **only for binomial distribution.**

${_nC_x}$ represents the number of ways (patterns) in which it is possible to get x successes in n trials

$${_nC_x} = \binom{n}{x} = \frac{n!}{x!(n-x)!} \qquad \textit{Where n! = n(n−1)(n−2)(n−3)...(3)(2)(1) for integers n > 0}$$

Example 4! = (4)(3)(2)(1) = 24   3! = (3)(2)(1) = 6   2! = (2)(1) = 6   also 0! = 1 by definition

$${_4C_2} = \frac{4!}{2!(4-2)!} = \frac{4!}{(2!)(2!)} = \frac{(4)(3)(2)(1)}{(2)(1)(2)(1)} = 6$$

${_nC_x}$ using calculator MATH PROB nCr :
Example: 4 MATH PROB nCr 2 ENTER

CHAPTER 4: BINOMIAL PROBABILITY DISTRIBUTION

Binomial Distribution TI 83, 84 Calculator

Use binompdf or binomcdf found at   2ⁿᵈ Distr
pdf : P(X = value) probability distribution function
cdf : P(X ≤ value) cumulative distribution function

P(X= x)	**binompdf (n,p,x)**
P(X ≤ x)	**binomcdf (n,p,x)**
P(X < x)	**binomcdf (n,p,x − 1)**
P(X > x)	**1 − binomcdf (n,p,x)**
P(X ≥ x)	**1 − binomcdf (n,p,x − 1)**

EXAMPLE B2:

http://www.pewresearch.org/fact-tank/2016/01/05/pew-research-center-will-call-75-cellphones-for-surveys-in-2016/?utm_source=Pew+Research+Center&utm_campaign=4a62041804-Methods_Newsletter_for_June6_24_2015
http://www.pewresearch.org/fact-tank/2016/01/05/pew-research-center-will-call-75-cellphones-for-surveys-in-2016/?utm_source=Pew+Research+Center&utm_campaign=4a62041804-Methods_Newsletter_for_June6_24_2015

Many of the major survey organizations that conduct "public opinion polls" gather their data through telephone surveys. These types of polls include political polls, polls about current events, and other subjects about demographics, lifestyle, economic issues, and more. In recent years, these survey organizations have had to change their data gathering techniques, as sampling from landline phones only will no longer provide a sample that is representative of the population.

Kyley McGeeney, a research methodologist at Pew Research Center, wrote that "All major survey organizations that conduct telephone surveys include cellphones in their samples. They have to, because the kinds of people who rely only on a cellphone are different from those reachable on a landline, even though being cellphone-only is becoming more mainstream. Cellphone-only individuals are considerably younger than people with a landline. They tend to have less education and lower incomes than people with a landline. They are also more likely to be Hispanic and to live in urban areas. For this reason, excluding cellphones from a poll – or not including enough of them – would provide a sample that is not representative of all U.S. adults."

Drew DeSilver at Pew Research Center cites that 65.7% of 25- to 29-year-olds live in wireless-only households and do not have landlines.

Suppose we took a sample consisting of 100 people age 25-29 and we are interested in counting the number of people in the sample who have only cell phone service.

X = (*description*)_____

p = (*description*)_____

p = (*value*) _____   X ~ _____

   a. Find the probability that 60 have only cell phone service

   b. Find the probability that at most (≤)60 have only cell phone service

   c. Find the probability that less than 60 have only cell phone service

   d. Find the probability that the number who have only cell phone service exceeds (is more than) 60

   e. Find the probability that at least (≥) 60 have only cell phone service

   f. Find the probability that exactly half of the people in the sample have only cell phone service

CHAPTER 4: BINOMIAL PROBABILITY DISTRIBUTION

Binomial Distribution TI 83, 84 Calculator

Use binompdf or binomcdf found at   2nd Distr
pdf : P(X = value) **probability distribution function**
cdf : P(X ≤ value) **cumulative distribution function**

P(X= x)	binompdf (n,p,x)
P(X ≤ x)	binomcdf (n,p,x)
P(X < x)	binomcdf (n,p,x − 1)
P(X > x)	1 − binomcdf (n,p,x)
P(X ≥ x)	1 − binomcdf (n,p,x − 1)

EXAMPLE B3: Make sure that the probability of success matches the definition of a success

A recent study showed that about 60% of California voters voted by mail. Suppose we are selecting a random sample of 75 voters and we are interested in the number of people in the sample who vote in person at the polls.

X = (*description*)_____

p = (*description*)_____

p = (*value*) _____   X ~ _____

    a. Find the probability that more than 25 of the voters in the sample voted in person at the polls.

    b. How many people in the sample would you expect to vote in person at the polls.

EXAMPLE B4: PRACTICE *Try-It 4.13 Introductory Statistics from Openstax download for free at www.openstax.org*

About 32% of students participate in a community volunteer program outside of school.
Suppose that 30 students are selected at random.
X = the number of students who participate in a community volunteer program outside of school **X ~ _____**

    a. Find the probability that at most 14 participate in a community volunteer program outside of school

    b. Find the probability that at least 15 participate in a community volunteer program outside of school

    c. Find the probability that more than 20 participate in a community volunteer program outside of school

    d. For many samples of 30 students, on average, what is the expected number per sample who participate in a community volunteer program outside of school?

    e. Find the standard deviation

NOTE: Recognizing Scientific Notation on your Calculator:

Sometimes probabilities are very small numbers.
If the number is very close to 0, your calculator automatically uses scientific notation:
      $0.000068 = 6.8 \times 10^{-5}$ appears on the calculator as 6.80000000E-5
      4.26710E-6 = 4.26713×10^{-6} = 0.0000426713