Chapter 6
Discrete Probability Distributions

Section 6.1 Discrete Random Variables

Random Variable (RV): A random variable is a numerical measure of the outcome of a probability experiment, so its value is determined by chance. Random variables are typically denoted using capital letters such as X.

Discrete: A variable is discrete if its value results from counting. (Number of kids in a family, number of TV's in a house, number of m&m's in a bag)
Continues: The variable is continuous if its value is measured. (Time, money, height, weight)

Example 1: Is it discrete or continuous random variable (rv):

(a) # of defective vending machines at VSU; Discrete
(b) The amount of time required to complete your homework each day for your math2620 class; Continues
(c) Heights of male students at VSU; Continues
(d) The number of students in your math class at VSU; Discrete
(e) Distance required to stop a car traveling at 70 mph; Continues
(f) # of chess games you will have to play before winning a game of chess. Discrete

Probability Distribution: A listing of all possible values and corresponding probabilities of a discrete random variable.

Example 2: Toss two coins. Let X = "the number of head of heads." List the values of X and the associated probabilities.
The experimental outcomes are: S = {HH, HT, TH, TT}
The values of the random variable X corresponding to the experimental outcomes are: S = {0, 1, 2}, see probability table below.

X	f(X)=P(X)	Outcome(s)
0	¼	TT
1	2/4	HT, TH
2	¼	HH

Note: **f(x)** above is called probability density function(pdf) for X. For a discrete random variable f(X)=P(X). we can use either one to represent a probability.

Discrete Probability Distributions. Let X be a discrete random variable. For the function f(x) to be a pdf for X, the following two conditions must be satisfied:

$$1.\quad \sum P(x) = 1 \quad and \quad 2.\quad 0 \le P(X) \le 1$$

Example 3: Given the probability distribution of a discrete random variable X, find (a) P(X= 2) (b) P($1 < X \le 4$) (c) P(X < 5).

X	0	1	2	3	4	5
f(X)=P(X=x)	0.1	0.2	?	0.3	0.25	0.1

(a) Since this is a probability distribution all the probabilities must sum to 1. The value that is missing is the value that will make the sum equal to 1. Hence, P(X= 2) = 1-(0.1+0.2+0.3+0.25+0.1)=1-0.95=0.05; i.e. P(X= 2) = 0.05

(b) The discrete values in the range of $1 < X \le 4$ are the values strictly greater than 1 but less than or equal to 4. From the probability table we can find only three values in this range, 2, 3, and 4. Hence, P($1 < X \le 4$) = P(X=2)+P(X=3)+P(X=4)=0.05+0.3+0.25 = 0.6 .

(c) P(X < 5) = P($X \le 4$) =P(X=0)+P(X=1)+P(X=2)+P(X=3)+P(X=4) = 0.1+0.2++0.05+0.3+0.25=0.9. A better approach is to use the compliment. If X< 5, its complement is x=5. Find the probability of X=5 and then subtract it from 1; i.e. P(X<5) =1-P(X=5) = 1-0.1 = 0.9

Identifying Discrete Probability Distributions

Problem Which of the following is a discrete probability distribution?

(a)

x	$P(x)$
1	0.20
2	0.35
3	0.12
4	0.40
5	−0.07

(b)

x	$P(x)$
1	0.20
2	0.25
3	0.10
4	0.14
5	0.49

(c)

x	$P(x)$
1	0.20
2	0.25
3	0.10
4	0.14
5	0.31

Approach In a discrete probability distribution, the sum of the probabilities must equal 1, and all probabilities must be between 0 and 1, inclusive.

Solution

(a) This is not a discrete probability distribution because $P(5) = -0.07$, which is less than 0.

(b) This is not a discrete probability distribution because

$$\sum P(x) = 0.20 + 0.25 + 0.10 + 0.14 + 0.49 = 1.18 \neq 1$$

(c) This is a discrete probability distribution because the sum of the probabilities equals 1, and each probability is between 0 and 1, inclusive. •

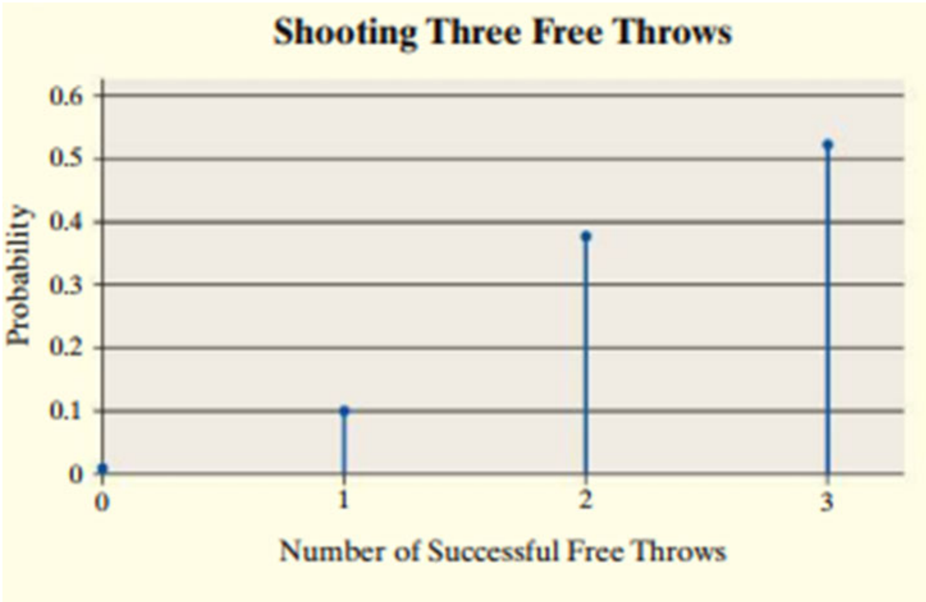
Example 3: Suppose we ask a basketball player to shoot three free throws. Let the random variable X represent the number of shots made, so x=0, 1, 2, or 3. Table 1 shows a probability distribution for the random variable X. We denote probabilities using the notation P(x), where x is a specific value of the random variable. We read P(x) as "the probability that the random variable X equals x." For example, P(3) =0.51 is read "the probability that the random variable X equals 3 is 0.51."

Table 1

x	$P(x)$
0	0.01
1	0.10
2	0.38
3	0.51

Graphing a Discrete Probability Distribution

Graph the discrete probability distribution given in Table 1 from Example 3. In the graph of a discrete probability distribution, the horizontal axis represents the values of the discrete random variable and the vertical axis represents the corresponding probability of the discrete random variable. Draw the graph using vertical lines above each value of the random variable to a height that is the probability of the random variable. Below is the graph of the distribution in Table 1.



Graphs of discrete probability distributions help determine the shape of the distribution. Recall that we describe distributions as skewed left, skewed right, or symmetric. The graph above is skewed left.

Mean and Variance for a Discrete RV

The Mean or Expected Value of a Discrete Random Variable

Expected Value of X (Mean): $E(X) = \mu_X = \sum x P(x)$

The mean or expected value of a discrete random variable is given by the formula $E(X) = \mu_X = \sum x P(x)$ where x is the value of the random variable and P(x) is the probability of observing the value x.

Computing the Mean of a Discrete Random Variable

Compute the mean of the discrete random variable given in Table 1 from Example 3. Find the mean of a discrete random variable by multiplying each value of the random variable by its probability and adding these products. Refer to Table 3. The first two columns represent the discrete probability distribution. The third column represents $x P(x)$. Substitute into $E(X) = \mu_X = \sum x P(x)$ to find the mean number of free throws made. $E(X) = \mu_X = \sum x P(x)$=0(0.01) +1(0.10) +2(0.38) +3(0.51) =0+0.1+0.76+1.53= 2.39 ≈ 2.4.

Table 3

x	$P(x)$	$x \cdot P(x)$
0	0.01	$0 \cdot 0.01 = 0$
1	0.10	$1 \cdot 0.1 = 0.1$
2	0.38	$2 \cdot 0.38 = 0.76$
3	0.51	$3 \cdot 0.51 = 1.53$

How to Interpret the Mean of a Discrete Random Variable?
The mean of a discrete random variable can be thought of as the mean outcome of the probability experiment if we repeated the experiment many times. If we repeated the experiment in of shooting three free throws many times, we would expect the mean number of free throws made to be around 2.4.

Interpretation of the Mean of a Discrete Random Variable
Suppose an experiment is repeated n independent times and the value of the random variable X is recorded. As the number of repetitions of the experiment increases, the mean value of the n trials will approach μ_X, the mean of the random variable X. In other words, let x_1 be the value of the random variable X after the first experiment, x_2 be the value of the random variable X after the second experiment, and so on. Then $\bar{X} = \frac{x_1+x_2+...+x_n}{n}$. The difference between $\bar{X}$ and μ_X gets closer to 0 as n increases.

Interpretation of the Mean of a Discrete Random Variable
The basketball player from Example 2 is asked to shoot three free throws 100 times. Compute the mean number of free throws made. The player shoots three free throws and the number made is recorded. We repeat this experiment 99 more times and then compute the mean number of free throws made (see table 4 below).

Solution Table 4 shows the results.

Table 4

First experiment →	3	2	3	3	3	3	1	2	3	2
Second experiment →	2	3	3	1	2	2	2	2	2	3
Third experiment →	3	3	2	2	3	2	3	2	2	2
	3	3	2	3	2	3	3	2	3	1
	3	2	2	2	2	0	2	3	1	2
	3	3	2	3	2	3	2	1	3	2
	2	3	3	3	1	3	3	1	3	3
	3	2	2	1	3	2	2	2	3	2
	3	2	2	2	3	3	2	2	3	3
← Hundredth experiment	2	3	2	1	2	3	3	2	3	3

In the first experiment, the player made all three free throws. In the second experiment, the player made two out of three free throws. In the hundredth experiment, the player made three free throws. The mean number of free throws made was

$$\bar{x} = \frac{3 + 2 + 3 + \cdots + 3}{100} = 2.35$$

This is close to the theoretical mean of 2.4 (from Example 5). As the number of repetitions of the experiment increases, we expect $\bar{x}$ to get even closer to 2.4. •

Find the Expected Value for a Real World Problem

Problem A term life insurance policy will pay a beneficiary a certain sum of money upon the death of the policyholder. These policies have premiums that must be paid annually. Suppose an 18-year-old male buys a \$250,000 1-year term life insurance policy for \$350. According to the *National Vital Statistics Report*, Vol. 58, No. 21, the probability that the male will survive the year is 0.998937. Compute the expected value of this policy to the insurance company.

Approach The experiment has two possible outcomes: survival or death. Let the random variable X represent the *payout* (money lost or gained), depending on survival or death of the insured. Assign probabilities to each payout and substitute these values into $E(X) = \mu_X = \sum xP(x)$

Solution

Step 1 Because $P(\text{survives}) = 0.998937$, $P(\text{dies}) = 0.001063$. If the client survives the year, the insurance company makes \$350, or $x = \$350$. If the client dies during the year, the insurance company must pay \$250,000 to the client's beneficiary, but still keeps the \$350 premium, so $x = \$350 - \$250{,}000 = -\$249{,}650$. The value is negative because it is money paid by the insurance company. The probability distribution is listed in Table 5.

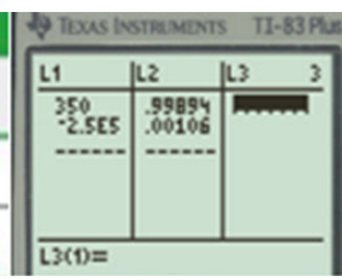
Step 2 The expected value of the policy (from the point of view of the insurance company) is

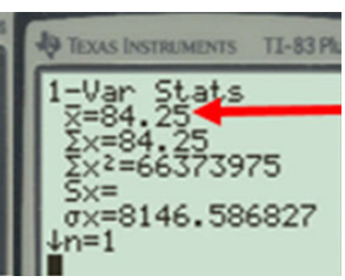
$$E(X) = \mu_X = \sum[x \cdot P(x)] = \$350(0.998937) + (-\$249{,}650)(0.001063) = \$84.25$$

Interpretation The company expects to make \$84.25 for each 18-year-old male client it insures. The \$84.25 profit of the insurance company is a long-term result. It does not make \$84.25 on each 18-year-old male it insures, but rather the average profit per 18-year-old male insured is \$84.25. Because this is a long-term result, the insurance "idea" will not work with only a few insured.

Table 5

x	$P(x)$
\$350 (survives)	0.998937
−\$249,650 (dies)	0.001063





How to do it on the calculator please see instructions below using the TI 83/84 to find the Mean and Standard Deviation.

Standard Deviation of a Discrete Random Variable

The standard deviation of a discrete random variable X is given by

$\sigma_x = \sqrt{\sum (x-\mu_X)^2 P(x)}$ (formula **2a**) or the computational formula

$\sigma_x = \sqrt{\sum x^2 P(x) - (\mu_X)^2}$ (formula **2b)** where x is the value of the random variable, μ_X is the mean of the random variable, and P(x) is the probability of observing a value of the random variable.

EXAMPLE: Computing the Standard Deviation of a Discrete Random Variable Find the standard deviation of the discrete random variable given in Table 1 from Example 3.

Approach We will use Formula (2a) with the unrounded mean $\mu_X = 2.39$.

Solution Refer to Table 6. Columns 1 and 2 represent the discrete probability distribution. Column 3 represents $(x-\mu_X)^2 \cdot P(x)$. Find the sum of the entries in Column 3.

Table 6

x	$P(x)$	$(x-\mu_X)^2 \cdot P(x)$
0	0.01	$(0-2.39)^2 \cdot 0.01 = 0.057121$
1	0.10	$(1-2.39)^2 \cdot 0.10 = 0.19321$
2	0.38	$(2-2.39)^2 \cdot 0.38 = 0.057798$
3	0.51	$(3-2.39)^2 \cdot 0.51 = 0.189771$
		$\sum[(x-\mu_X)^2 \cdot P(x)] = 0.4979$

The standard deviation of the discrete random variable X is

$$\sigma_X = \sqrt{\sum[(x-\mu_X)^2 \cdot P(x)]} = \sqrt{0.4979} \approx 0.7$$

Approach We will use Formula (2b) with the unrounded mean $\mu_X = 2.39$.

Solution Refer to Table 7. Columns 1 and 2 represent the discrete probability distribution. Column 3 represents $x^2 \cdot P(x)$. Find the sum of the entries in Column 3.

Table 7

x	$P(x)$	$x^2 \cdot P(x)$
0	0.01	$0^2 \cdot 0.01 = 0$
1	0.10	$1^2 \cdot 0.10 = 0.10$
2	0.38	$2^2 \cdot 0.38 = 1.52$
3	0.51	$3^2 \cdot 0.51 = 4.59$
		$\sum[x^2 \cdot P(x)] = 6.21$

The standard deviation of the discrete random variable X is

$$\sigma_X = \sqrt{\sum[x^2 \cdot P(x)] - \mu_X^2} = \sqrt{6.21 - 2.39^2}$$
$$= \sqrt{0.4979} \approx 0.7$$

The variance, σ_X^2, of the discrete random variable is the value under the square root in the computation of the standard deviation. The variance of the discrete random variable in Example above is $\sigma_X^2 = 0.4979 \approx 0.5$.

Using the TI 83/84 to find the Mean and Standard Deviation for a Discrete Random Variable

Problem Use statistical software or a graphing calculator to find the mean and the standard deviation of the random variable whose distribution is given in Table 1.

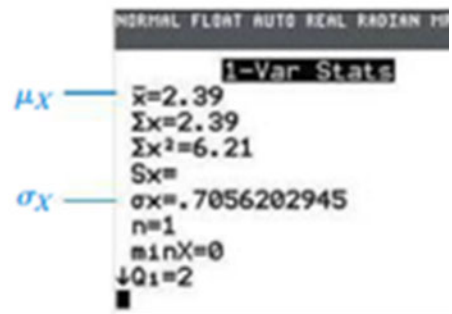
Approach We will use a TI-84 Plus C graphing calculator to obtain the mean and standard deviation. The steps for determining the mean and standard deviation using a TI-83/84 Plus graphing calculator and given in the Technology Step-by-Step below.

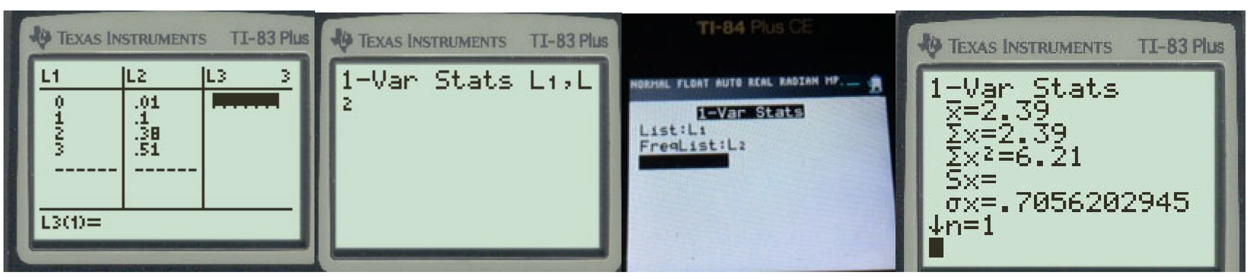
Solution Figure 4 shows the results from a TI-84 Plus C graphing calculator.

Note: The TI-84 does not find s_X when the sum of L_2 is 1. •

TI-83/84 Plus

1. Enter the values of the random variable in L1 and their corresponding probabilities in L2.
2. Press STAT, highlight CALC, select `1: 1-VAR Stats`, and press ENTER.
3. Select L1 for List :; Select L2 for FreqList:. Highlight Calculate and press ENTER.





Homework-Section 6.1 Online - MyStatLab

Section 6.2 The Binomial Distribution

Properties of a Binomial Experiment

1. The experiment consists of a sequence of n identical and independent trials, that is, repeat a process n times.
2. There are two outcomes possible on each trial -- success or failure.
3. p = P[success] remains constant from trial to trial.

The general notation for a Binomial r.v. is **X~B(n,p).**

The probability density function(pdf) or probability distribution function (pdf) of the r.v. X is given by $f(x) = \binom{n}{x} p^{x}(1\text{-}p)^{n-x}$; x=0,1, 2, …, n . Probabilities for the binomial distribution will be computed using the **TI 83/84** calculator.

Example 1: Flip a coin twice. Let the r.v. X be the number of heads. Does the r.v. X follow a binomial distribution? (Yes) **X~ B(2, 0.5)**

The pdf is: $f(x) = \binom{2}{x}(0.5)^{x}(1-0.5)^{2-x}$; where x = 0, 1, 2

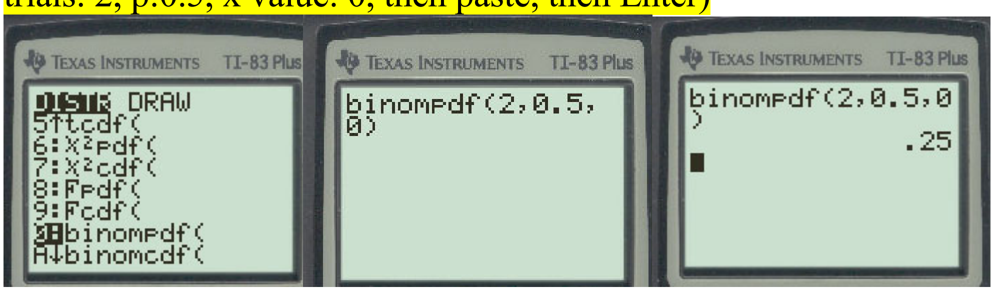
The probability distribution for this experiment is:
Let us verify the pdf using the TI-83/84.

X	0	1	2
P(X=x)	.25	.5	.25

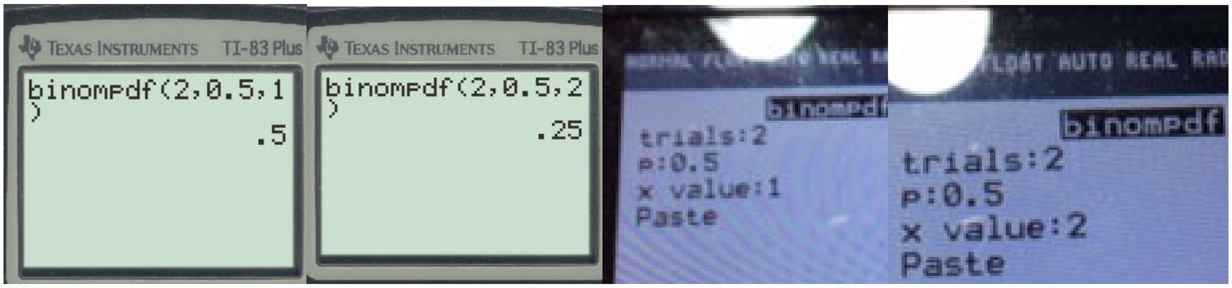
The TI-83/84 Commands are: $P(X = x) = binompdf(n, p, x)$.
The binopdf is used to find the probability for a single value of x.

P(X=0)=binopdf(2, 0.5, 0)=0.25

Using the TI 83/84
2nd and Vars to access the DISTR, scroll down to 0: binopdf(, hit Enter, Type 2, .5, 0), and Hit Enter(On newer calculators: Scroll down to A and Enter, trials: 2, p:0.5, x value: 0, then paste, then Enter)



Similarly: P(X=1)=binopdf(2, 0.5, 1)=0.5 and P(X=2)=binopdf(2, 0.5, 2)=0.25



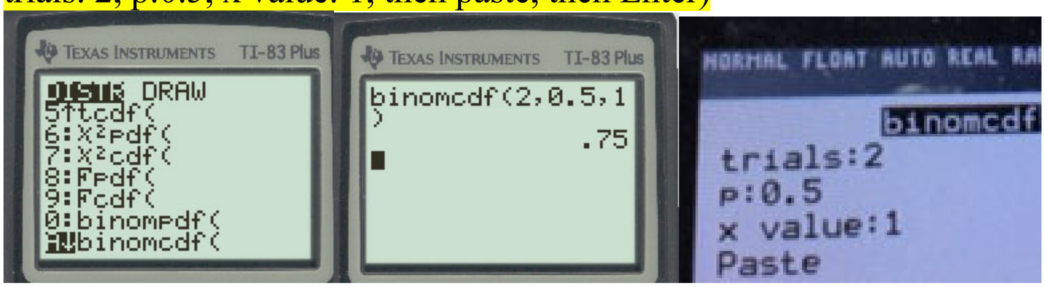
$P(X \le x) = binomcdf\,(n,\ p,\ x)$.

The binocdf is used to find the sum of probabilities less than or equal to x.

$P(X \le 1)$ = binocdf(2, 0.5, 1) = 0.75

Using the TI 83/84

2nd and Vars to access the DISTR, scroll down to A: binocdf(, hit Enter, Type 2, 0.5, 1), and Hit Enter(On newer calculators: Scroll down to B and Enter, trials: 2, p:0.5, x value: 1, then paste, then Enter)



Example 2: A baseball player with a batting average 30% comes to bat four times in a game. What is the probability he will hit the ball. (a) P(X=0), (b) P(X=1), (c) $P(1 < X \le 2)$, (d) $P(X \le 3)$, (e) $P(X \ge 3)$.

X~B(n=4, p=0.3)

(a) P(X=0) = binompdf(4 , .3, 0) = 0.2401 (Using TI-83)

(b) P(X=1) = binompdf(4 , .3, 1) = 0.4116 (Using TI-83)

(c) $P(1<X \le 2)$ = P(X=2) = binompdf(4 , .3, 2) = 0.2646 (Using TI-83)

(d) $P(X \le 3)$ = binomcdf(4 , .3, 3) = 0.9919 (Using TI-83)

(e) $P(X \ge 3)$ = 1 - $P(X \le 2)$=1- binomcdf(4 , .3, 2) = 0.0837 or You can also use binopdf to answer this question. $P(X \ge 3)$ = P(X=3) + P(X=4)= binopdf(4, 0.3, 3) + binopdf(4, 0.3, 4) (Using TI-83)

Binomial Mean and Variance:
If the r.v. X follows a binomial distribution with parameters **n** and **p**, **X ~ B(n,p)**, then the expected value and S.D. of X are give by
$\mu_X = E(X) = n.p$, $\sigma_X^2 = np(1-p)$, and $\sigma_X = \sqrt{np(1-p)}$

Example 3. In example 1, X ~ B(n=2, p=0.5).

$$\mu_X = E(X) = n.p = (2)(0.5) = 1$$
$$\sigma_X^2 = np(1-p) = 2(0.5)(1-0.5) = 2(0.5)(0.5) = 0.5$$
$$\sigma_X = \sqrt{.05} = 0.7071$$

Example 4. According to CTIA, 41% of all U.S. households are wireless-only households (no landline). In a random sample of 20 households, what is the probability that; X~**B(n=20, p=0.41)**

(a) exactly 5 are wireless-only? P(X=5) = binopdf(20, 0.41, 5) = 0.06563

(b) fewer than 3 are wireless-only? P(X<3) = $P(X \le 2)$= binocdf(20, 0.41, 2)=0.0028

(c) at least 3 are wireless-only? $P(X \ge 3) = 1 - P(X \le 2) = 1-$ binomcdf(20 , 0.41, 2) = 1-0.0028=0.9972

(d) the number of households that are wireless-only is between 5 and 7, inclusive? $P(5 \le X \ge 7) = P(X \le 7) - P(X \le 4) =$ binomcdf(20 , 0.41, 7) - binomcdf(20 , 0.41, 4) = 0.3804-0.0423=0.3381 or You can also use binopdf to answer this question. $P(5 \le X \ge 7)$= P(X=5) + P(X=6)+P(X=7)= binopdf(20, 0.41, 5) + binopdf(20, 0.41, 6) + binopdf(20, 0.41, 7) (Using TI-83)

(e) What is the expected number and standard deviation that are wireless?

$$\mu_X = E(X) = n.p = (20)(0.41) = 8.2$$
$$\sigma_X = \sqrt{(20)(0.41)(1-0.41)} = 2.1995$$

Homework-Section 6.2 Online - MyStatLab