AP Statistics
Solutions to Packet 8

<u>The Binomial and Geometric Distributions</u>
The Binomial Distributions
The Geometric Distributions

HW #1   1 – 5, 7, 8

8.1 **BINOMIAL SETTING?** In each situation below, is it reasonable to use a binomial distribution for the random variable *X*? Give reasons for your answer in each case.

(a) An auto manufacturer chooses one car from each hour's production for a detailed quality inspection. One variable recorded is the count *X* of finish defects (dimples, ripples, etc.) in the car's paint. No: There is no fixed *n* (i.e., there is no definite upper limit on the number of defects).

(b) The pool of potential jurors for a murder case contains 100 persons chosen at random from the adult residents of a large city. Each person in the pool is asked whether he or she opposes the death penalty; *X* is the number who say "Yes." Yes

B: Only two choices, yes or no
I: It is reasonable to believe that all responses are independent (ignoring any "peer pressure")
N: $n = 100$
S: All have the same probability of saying "yes" since they are randomly chosen from the population

(c) Joe buys a ticket in his state's "Pick 3" lottery game every week; *X* is the number of times in a year that he wins a prize. Yes

B: Only two choices, win or lose
I: All responses are independent
N: $n = 52$
S: In a "Pick 3" game, Joe's chance of winning the lottery is the same every week

8.2 **BINOMIAL SETTING?** In each of the following cases, decided whether or not a binomial distribution is an appropriate model, and give your reasons.

(a) Fifty students are taught about binomial distributions by a television program. After completing their study, all students take the same examination. The number of students who pass is counted. YES

B: Only two choices, pass or fail
I: It is reasonable to assume that the results for the 50 students are independent
N: $n = 50$
S: Each student has the same chance of passing

(b) A student studies binomial distributions using computer-assisted instruction. After the initial instruction is completed, the computer presents 10 problems. The student solves each problem and enters the answer; the computer gives additional instruction between problems if the student's answer is wrong. The number of problems that the students solves correctly is counted.

No: Since the student receives instruction after incorrect answers, her probability of success is likely to increase.

(c) A chemist repeats a solubility test 10 times on the same substance. Each test is conducted at a temperature 10° higher than the previous test. She counts the number of times that the substance dissolves completely. No: Temperature may affect the outcome of the test.

8.3 **INHERITING BLOOD TYPE** Each child born to a particular set of parents has probability 0.25 of having blood type O. Suppose these parents have 5 children. Let X = number of children who have type O blood. Then X is $B(5, 0.25)$.

(a) What is the probability that exactly 2 children have type O blood? 0.2637 binompdf(5, .25, 2)

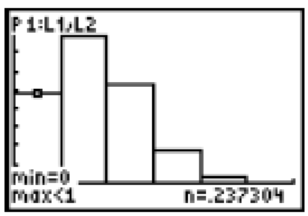
(b) Make a table for the pdf of the random variable X. Then use the calculator to find the probabilities of all possible values of X, and complete the table.

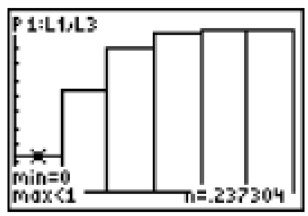
x_i:	0	1	2	3	4	5	
probability:	0.2373	0.3955	0.2637	0.0879	0.0146	0.0010	[binompdf(5, .25)]
cum prob:	0.2373	0.6328	0.8965	0.9844	0.9990	1	[binomcdf(5, .25)]

(c) Verify that the sum of the probabilities is 1.

(d) Construct a histogram of the pdf. See below

(e) Use the calculator to find the cumulative probabilities, and add these values to your pdf table. Then construct a cumulative distribution histogram.





8.4 **GUESSING ON A TRUE-FALSE QUIZ** Suppose that James guesses on each question of a 50-item true-false quiz. Find the probability that James passes if

Let X = the number of correct answers. X is binomial with $n = 50$, $p = 0.5$.

(a) a score of 25 or more correct is needed to pass.

$P(X \geq 25) = 1 - P(X \leq 24) = 1 - \text{binomcdf}(50, .5, 24) = 1 - .444 = .556.$

(b) a score of 30 or more correct is needed to pass.

$P(X \geq 30) = 1 - P(X \leq 29) = 1 - \text{binomcdf}(50, .5, 29) = 1 - .899 = .101.$

(c) a score of 32 or more correct is needed to pass.

$P(X \geq 32) = 1 - P(X \leq 31) = 1 - \text{binomcdf}(50, .5, 31) = 1 - .968 = .032.$

8.5 **GUESSING ON A MULTIPLE-CHOICE QUIZ** Suppose that Erin guesses on each question of a multiple-choice quiz with four different choices.

Let X = the number of correct answers. X is binomial with n = 10, p = 0.25.

(a) If each question has four different choices, find the probability that Erin gets one or more correct answers on a 10-item quiz.

The probability of at least one correct answer is $P(X \geq 1) = 1 - P(\text{no correct answers}) =$ $1 - P(X = 0) = 1 - \text{binompdf}(10, .25, 0) = 1 - 0.056 = 0.944$.

(b) If the quiz consists of three questions, question 1 has 3 possible answers, question 2 has 4 possible answers, and question 3 has 5 possible answers, find the probability that Erin gets one or more correct answers.

Let X = the number of correct answers. We can write $X = X1 + X2 + X3$, where Xi = the number of correct answers on question i. (Note that the only possible values of Xi are 0 and 1, with 0 representing an incorrect answer and 1 a correct answer.) The probability of at least one correct answer is $P(X \geq 1) = 1 - P(X = 0) = 1 - [P(X1 = 0)\,(X2 = 0)\,P(X3 = 0)]$ (since the Xi are independent) $= 1 - \left(\frac{2}{3}\right)\left(\frac{3}{4}\right)\left(\frac{4}{5}\right) = 1 - \frac{24}{60} = 0.6$.

8.7 **DO OUR ATHLETES GRADUATE?** A university claims that 80% of its basketball players get degrees. An investigation examines the fate of all 20 players who entered the program over a period of several years that ended six years ago. Of these players, 11 graduated and the remaining 9 are no longer in school. If the university's claim is true, the number of players among the 20 who graduate would have the binomial distribution with n = 20 and p = 0.8. What is the probability that exactly 11 out of 20 players graduate?

Let X = the number of players out of 20 who graduate.

$P(X = 11) = \text{binompdf}(20, .8, 11) = 0.0074$.

8.8 **MARITAL STATUS** Among employed women, 25% have never been married. Select 10 employed women at random.

(a) The number in your sample who have never been married has a binomial distribution. What are n and p? n = 10 and p = 0.25

(b) What is the probability that exactly 2 of the 10 women in your sample have never been married?

$$P(X = 2) = \binom{10}{2}(0.25)^2(0.75)^8 = 0.28157 \qquad [\text{binomialpdf}(10, .25, 2)]$$

(c) What is the probability that 2 or fewer have never been married?

$$P(X \leq 2) = \binom{10}{0}(0.25)^0(0.75)^{10} + \binom{10}{1}(0.25)^1(0.75)^9 + \binom{10}{2}(0.25)^2(0.75)^8 = 0.52559$$

HW #2 9, 11 – 13, 15 – 17, 19

In each of the following exercises, you are to use the binomial probability formula to answer the question. Do not use the binomial pdf command on your calculator. Begin with the formula, and show substitution into the formula.

8.9 **BLOOD TYPES** The count X of children with type O blood among 5 children whose parents carry genes for both the O and A types is $B(5, 0.25)$. Use the binomial probability formula to find $P(X = 3)$.

$$P(X=3)=\binom{5}{3}(0.25)^3(0.75)^2=10\,(0.25)^3(0.75)^2=0.088 \quad [\text{binomialpdf}(5,.25,3)]$$

8.11 **MORE ON BLOOD TYPES** Use the binomial probability formula to find the probability that at least one of the children in the preceding exercise has blood type O. (*Hint:* Do not calculate more than one binomial formula.) Let X = the number of children with blood type O. *X is B*(5, .25).

$$P(X\geq 1)=1-P(X=0)=1-\binom{5}{0}(0.25)^0(0.75)^5=1-(.75)^5=0.763 \quad [1-\text{binomialpdf}(5,.25,0)]$$

8.12 **GRADUATION RATES FOR ATHLETES** See Exercise 8.7*(preceding page)*. The number of athletes who graduate is $B(20, 0.8)$. Use the binomial probability formula to find the probability that all 20 graduate. What is the probability that they do not all graduate?

Probability that all 20 graduate:

$$P(X=20)=\binom{20}{20}(0.8)^{20}(0.2)^0=(0.8)^{20}=0.0115 \quad [\text{binomialpdf}(20,.8,20)]$$

Probability that not all 20 graduate:

$$P(X<20)=1-P(X=20)=0.9885 \quad [1-\text{binomialpdf}(20,.8,20)]$$

8.13 **HISPANIC REPRESENTATION** A factory employs several thousand workers, of whom 30% are Hispanic. If the 15 members of the union executive committee were chosen from the workers at random, the number of Hispanics on the committee would have the binomial distribution with $n = 15$ and $p = 0.3$.

(a) What is the probability that exactly 3 members of the committee are Hispanic?

$$P(X=3)=\binom{15}{3}(0.3)^3(0.7)^{12}=0.17004 \quad [\text{binomialpdf}(15,.3,3)]$$

(b) What is the probability that none of the committee members are Hispanic?

$$P(X=0)=\binom{15}{0}(0.3)^0(0.7)^{15}=0.00475 \quad [\text{bonomialpdf}(15,.3,0)]$$

8.15 **ATTITUDES ON SHOPPING** Are attitudes toward shopping changing? Sample surveys show that fewer people enjoy shopping than in the past. A recent survey asked a nationwide random sample of 2500 adults if they agreed or disagreed that "I like buying new clothes, but shopping is often frustrating and time-consuming." The population that the poll wants to draw conclusions about is all U.S. residents aged 18 and over. Suppose that in fact 60% of all adults U.S. residents would say "Agree" if asked the same question. What is the probability that 1520 or more of the sample agree?

(a) Verify that the rule of thumb conditions are satisfied for using the normal approximation to the binomial distribution.

$np = 2500(0.6) = 1500 \geq 10$ $\quad$ $n(1-p) = 2500(0.4) = 1000 \geq 10$

(b) Use your calculator and the cumulative binomial function to verify the exact answer for the probability that at least 1520 people in the sample find shopping frustrating is 0.2131. What is the probability correct to 6 decimal places?

Let X = the number of people in the sample who find shopping frustrating. $X \sim B(2500, .6)$.
Then $P(X \geq 1520) = 1 - P(X \leq 1519) = 1 -$ binomcdf (2500, .6, 1519) =
$1 - 0.7868609113 = 0.2131390887$, which rounds to 0.213139 (correct to 6 decimal places)

(c) What is the probability that at most 1468 people in the sample would agree with the statement that shopping is frustrating?

$P(X \leq 1468)$ = binomcdf (2500, .6, 1468) = 0.0994.
Using the normal approximation to the binomial, ncdf($-\infty$, 1468, 1500, 24.5) = 0.0957,
(a difference of 0.0037.)

8.16 **HISPANIC COMMITTEE MEMBERS**

(a) A factory employs several thousand workers, of whom 30% are Hispanic. If the 15 members of the union executive committee were chosen from the workers at random, the number of Hispanics on the committee would have the binomial distribution with $n = 15$ and $p = 0.3$. What is the mean number of Hispanics on randomly chosen committees of 15 workers?

$m = np = 15(0.3) = 4.5$

(b) What is the standard deviation σ of the count X of Hispanic members?

$s = \sqrt{np(1-p)} = \sqrt{15(0.3)(0.7)} = 1.7748$

(c) Suppose that 10% of the factory workers were Hispanic. Then $p = 0.1$. What is σ in this case? What is σ if $p = 0.01$? What does your work show about the behavior of the standard deviation of a binomial distribution as the probability of a success gets closer to 0?

If $p = 0.1,\ s = \sqrt{15(0.1)(0.9)} = 1.1619$

If $p = 0.01,\ s = \sqrt{15(0.01)(0.99)} = 0.3854$

As p gets closer to 0, s *gets closer to* 0.

8.17 **DO OUR ATHLETES GRADUATE?**

(a) Find the mean number of graduates out of 20 players in the setting of Exercise 8.12 (packet p. 5).

$\mu = np = (20)(.8) = 16.$

(b) Find the standard deviation σ of the count X.

$$s = \sqrt{npq} = \sqrt{20(0.8)(0.2)} = 1.7888$$

(c) Suppose that the 20 players came from a population of which $p = 0.9$ graduated. What is the standard deviation σ of the count of graduates? If $p = 0.99$, what is σ? What does your work show about the behavior of the standard deviation of a binomial distribution as the probability p of success gets closer to 1?

$$s = \sqrt{20(0.9)(0.1)} = 1.3416$$

$$s = \sqrt{20(0.99)(0.01)} = 0.44497$$

As p gets closer to 1, *s gets closer to* 0.

8.19 **POLLING** Many local polls of public opinion use samples of size 400 to 800. Consider a poll of 400 adults in Richmond that asks the question "Do you approve of President George W. Bush's response to the World Trade Center terrorists attacks in September 2001?" Suppose we know that President Bush's approval rating on this issue nationally is 92% a week after the incident.

(a) What is the random variable X? Is X binomial? Explain.

X = the number of people in the sample of 400 adult Richmonders who approve of the President's reaction.

B: Only two choices, approve or not

I: Because the sample size is small compared to the population size (all adult Richmonders), it is reasonable to consider the individual responses independent

N: $n = 400$

S: All have the same probability of success (approval) = 0.92

(b) Calculate the binomial probability that at most 358 of the 400 adults in the Richmond poll answer "Yes" to this question.

$$P(X \le 358) = \texttt{binomcdf}(400, 0.92, 358) = .0441$$

(c) Find the expected number of people in the sample who indicate approval. Find the standard deviation of X. $m = np = 400(0.92) = 368$, $s = \sqrt{npq} = \sqrt{400(0.92)(0.08)} = 5.426$

(d) Perform a normal approximation to answer the question in (b), and compare the results of the binomial calculation and the normal approximation. Is the normal approximation satisfactory?

$$P(X \le 358) = P\left(Z \le \frac{358-368}{5.426}\right) = P(Z \le -1.843) = 0.0327$$

The approximation is not very accurate (note that p is close to 1).

HW #3 23, 26, 27, 29, 30, 32, 34

8.23 **SIMULATING COMMITTEE SELECTION** Refer to Exercise 8.13 (packet p. 5). Construct a simulation to estimate the probability that in a committee of 15 members, 3 or fewer members are Hispanic. Describe the design of your experiment, including the correspondence between digits and outcomes in the experiment, and report the relative frequency for 30 repetitions.

There are $n = 15$ people on the committee, and the probability that a randomly selected person is Hispanic is $p = 0.3$.

Let 0, 1, 2 ⇔Hispanic and let 3–9 ⇔non-Hispanic.

Using the calculator, repeat the command 30 times:

randBin (1, .3, 15) →L1: sum (L1) →L2 (1)

where 0 = non-Hispanic, and 1 = Hispanic.

Our frequencies were:

x_i	0	1	2	3	4	5	6	7	8
freq	0	1	2	4	5	11	3	4	0

For this simulation, the relative frequency of 3 or fewer Hispanics was $= \frac{7}{30} =$ 0.233.

Compare this with the theoretical result: $P(X \le 3) = 0.29687$, where X = number of Hispanics on the committee.

8.26 **DRAWING POKER CHIPS** There are 50 poker chips in a container, 25 of which are red, 15 white, and 10 blue. You draw a chip without looking 25 times, each time returning the chip to the container.

(a) What is the expected number of white chips you will draw in 25 draws?

The probability of drawing a white chip is 15/50 = 0.3. The number of white chips in 25 draws is $B(25, .3)$. Therefore, the expected number of white chips is $np = (25)(0.3) = 7.5$.

(b) What is the standard deviation of the number of blue chips that you will draw?

The probability of drawing a blue chip is 10/50 = 0.2. The number of blue chips in 25 draws is $B(25, .2)$. Therefore, the std dev of the number of blue chips is $\sqrt{npq} = \sqrt{25(0.2)(0.8)} = 2$.

(c) Simulate 25 draws by hand or by calculator. Repeat the process as many times as you think necessary.

Let the digits 0, 1, 2, 3, 4 ⇔red chip, 5, 6, 7 ⇔white chip, and 8, 9 ⇔blue chip.

Draw 25 random digits from Table B and record the number of times that you get chips of various colors.

Using the TI-84, you can draw 25 random digits using the command randInt(0, 9, 25)→L1. You can then sortA(L1) to make it easier to count the data. Repeat this process 30 times (or however many times you like) to simulate multiple draws of 25 chips.

A sample simulation of a single 25-chip draw using the TI-84 yielded the following result:

Digit	0	1	2	3	4	5	6	7	8	9
Frequency	4	3	4	2	1	2	0	2	1	6

This corresponds to drawing 14 red chips, 4 white chips, and 7 blue chips.

(d) Based on your answers to parts (a) and (c), is it likely or unlikely that you will draw 9 or fewer blue chips? The expected number of blue chips is (25)(0.2) = 5, and the standard deviation = 2 by part (b). It seems extremely likely that you will draw at most 9 blue chips.

The actual probability is binomcdf (25, .2, 9) = 0.9827.

(e) Is it likely or unlikely that you will draw 15 or fewer blue chips? It seems virtually certain that you will draw 15 or fewer blue chips; the probability is even larger than in part (d).

The actual probability is binomcdf (25, .2, 15) = 0.999998.

8.27 **RANDOM DIGITS** Each entry in a table of random digits like Table B has probability 0.1 of being a 0, and digits are independent of each other.

The count of 0s among *n* random digits has a binomial distribution with *p* = 0.1.

(a) What is the probability that a group of five digits from the table will contain at least one 0?

P(at least one 0) = 1 − P(no 0) = 1 − $(0.9)^5$ = 0.40951.

(b) What is the mean number of 0s in lines 40 digits long?

$\mu = np = 40(0.1) = 4$

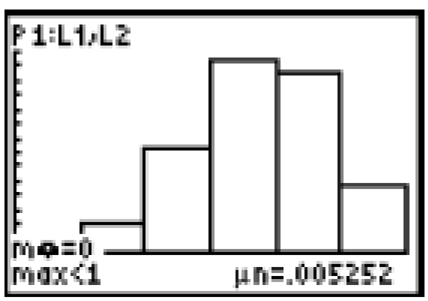
8.29 **RANDOM STOCK PRICES** A believer in the "random walk" theory of stock markets thinks that an index of stock prices has probability 0.65 of increasing in any year. Moreover, the change in the index in any given year is not influenced by whether it rose or fell in earlier years. Let *X* be the number of years among the next 5 years in which the index rises.

(a) *X* has a binomial distribution. What are *n* and *p*? *n* = 5 and *p* = 0.65.

(b) What are the possible values that *X* can take? *X* can take any value between 0 and 5.

(c) Find the probability of each value *X*. Draw a probability histogram for the distribution of *X*.

X_i:	0	1	2	3	4	5
probability	0.0053	0.0488	0.1812	0.3364	0.3124	0.1160



(d) What are the mean and standard deviation of this distribution? Mark the location of the mean on the histogram. $m = np = 5(0.65) = 3.25$, $s = \sqrt{npq} = \sqrt{5(0.65)(0.35)} = 1.067$

8.30 **LIE DETECTORS** A federal report finds that lie detector tests given to truthful persons have probability about 0.2 of suggesting that the person is deceptive.

(a) A company asks 12 job applicants about thefts from previous employers, using a lie detector to assess their truthfulness. Suppose that all 12 answer truthfully. What is the probability that the lie detector says all 12 are truthful? What is the probability that the lie detector says at least 1 is deceptive?

The probability that all are assessed as truthful is $\binom{12}{0}(0.2)^0\ (0.8)^{12} = 0.06872$

The probability that at least one is reported to be a liar is $1 - 0.06872 = 0.93128$.

(b) What is the mean number among 12 truthful persons who will be classified as deceptive? What is the standard deviation of this number?

$m = np = 12(0.2) = 2.4, \qquad s = \sqrt{npq} = \sqrt{12(0.2)(0.8)} = \sqrt{1.92} = 1.38564$

(c) What is the probability that the number classified as deceptive is less than the mean?

P(X < the mean) = P($X < 2.4$) = P($X \le 2$) = binomcdf(12, .2, 2) = 0.5583.

8.32 **PLANNING A SURVEY** You are planning a sample survey of small businesses in your area. You will choose an SRS of businesses listed in the telephone book's Yellow Pages. Experience shows that only about half the businesses you contact will respond.

(a) If you contact 150 businesses, it is reasonable to use the binomial distribution with $n = 150$ and $p = 0.5$ for the number X who respond. Explain why.

B: Only two choices, respond or not
I: It is reasonable to believe that all responses are independent
N: $n = 150$
S: All have the same probability of success (response) = 0.5

(b) What is the expected number (the mean) who will respond?

$m = 150(0.5) = 75 \textit{ responses}$

(c) What is the probability that 70 or fewer will respond? (Use the normal approximation.)

$s = \sqrt{npq} = \sqrt{150\,(0.5)(0.5)} = 6.1237$

$$P(X \le 70) = P\left(Z \le \frac{70-75}{6.1237}\right) = P\left(Z \le -0.82\right) = 0.2071$$

(d) How large a sample must you take to increase the mean number of respondents to 100?

You must increase the sample to 200, since (200)(0.5) = 100.

8.34 **AIDS TEST** A test for the presence of antibodies to the AIDS virus in blood has probability 0.99 of detecting the antibodies when they are present. Suppose that during a year 20 units of blood with AIDS antibodies pass through a blood bank.

(a) Take X to be the number of these 20 units that the test detects. What is the distribution of X?

X has a binomial distribution with $n = 20$ and $p = 0.99$.

(b) What is the probability that the test detects all 20 contaminated units? What is the probability that at least 1 unit is not detected?

$$P(X = 20) = \binom{20}{20}(0.99)^{20}(0.01)^{0} = 0.81791;$$ [binompdf(20, .99, 20)]

$$P(X < 20) = 1 - P(X = 20) = 1 - 0.81791 = 0.18209.$$

(c) What is the mean number of units among the 20 that will be detected? What is the standard deviation of the number detected?

$$m_X = np = 20(.99) = 19.8$$

$$s_X = \sqrt{npq} = \sqrt{20(.99)(.01)} = .445$$

HW #4   37, 39, 41, 43, 45, 46 (skip h)

8.37 **GEOMETRIC SETTING** For each of the following, determine if the experiment describes a geometric distribution. If it does, describe the two events of interest (success and failure), what constitutes a trial, and the probability of success on one trial. If the random variable is not geometric, identify a condition of the geometric setting that is not satisfied.

(a) Flip a coin until you observe a tail. Yes, geometric.

B: Only two choices, head or tail
I: Each flip is independent
T: Flip until you observe a tail
S: Probability of success, $p = 0.5$

(b) Record the number of times a player makes both shots in a one-and-one foul-shooting situation. (In this situation, you get to attempt a second shot only if you make your first shot.)

Not a geometric setting. You are not counting the number of trials before the first success is obtained.

(c) Draw a card from a deck, observe the card, and replace the card within the deck. Count the number of times you draw a card in this manner until you observe a jack. Yes, geometric.

B: Only two choices, success = jack, failure = any other card
I: Each draw is independent because the card is replaced each time
T: Draw until you observe a jack
S: Probability of success, $p = 4/52 = 1/13$

(d) Buy a "Match 6" lottery ticket every day until you win the lottery. (In a "Match 6" lottery, a player chooses 6 different numbers from the set {1, 2, 3, . . ., 44}. A lottery representative draws 6 different numbers from this set. To win, the player must match all 6 numbers, in any order.)

Yes, geometric.

B: Only two choices, success = match all 6 numbers, failure = do not match all 6 numbers.
I: Trials are independent because the setting of a drawing is always the same and the results of different drawings do not influence each other.
T: Keep drawing until you match all 6
S: Probability of success, $p = \frac{6!}{\binom{44}{6}} = 0.000102$

(e) There are 10 red marbles and 5 blue marbles in a jar. You reach in and, without looking, select a marble. You want to know how many marbles you will have to draw (without replacement), on average, in order to be sure that you have 3 red marbles.

Not a geometric setting. The trials (draws) are not independent because you are drawing without replacement. Also, you are interested in getting 3 successes, rather than just the first success.

8.39 **TESTING HARD DRIVES** Suppose we have data that suggest that 3% of a company's hard disk drives are defective. You have been asked to determine the probability that the first defective hard drive is the fifth unit tested.

(a) Verify that this is a geometric setting. Identify the random variable; that is, write *X* = number of _______________ and fill in the blank. What constitutes a success in this situation?

X = number of drives tested in order to find the first defective.
Success = defective drive.

B: Only two choices, defective or not
I: Tests on successive drives are independent
T: keep testing until you find a defective drive
S: Probability of success, $p = 0.03$

(b) Answer the original question: What is the probability that the first defective hard drive is the fifth unit tested?

$$P(X=5) = (1-0.03)^{5-1}(0.03)$$
$$= (0.97)^4(0.03)$$
$$= 0.0266$$

(c) Find the first four entries in the table of the pdf for the random variable *X*.

X	1	2	3	4
P(X)	0.03	0.0291	0.0282	0.0274

8.41 **FLIP A COIN** Consider the following experiment: flip a coin until a head appears.

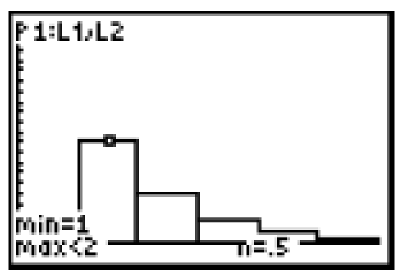
(a) Identify the random variable *X*.

X = number of flips required in order to get the first head. *X* is a geometric random variable with $p = 0.5$.

(b) Construct the pdf table for *X*. Then plot the probability histogram.

$$P(X=x) = (0.5)^{x-1}(0.5) = (0.5)^x \quad \textit{for} \;\; x = 1, 2, 3, 4\ldots$$

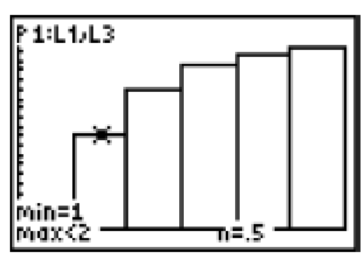
x	1	2	3	4	5	. . .
$P(X \le x)$	0.5	0.25	0.125	0.0625	0.03125	. . .



(c) Compute the cdf and plot its histogram.

$$P(X \le x) = (0.5)^1 + \ldots + (0.5)^x \quad \textit{for} \;\; x = 1, 2, 3, 4\ldots$$

x	1	2	3	4	5	. . .
$P(X \le x)$	0.5	0.75	0.875	0.9375	0.96875	. . .

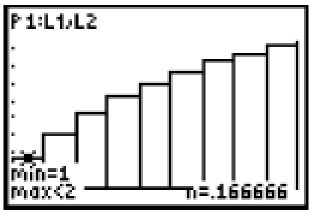


8.43 **ROLL A DIE**

An experiment consists of rolling a single die. The event of interest is rolling a 3; this event is called a success. The random variable is defined as *X* = the number of trials until a 3 occurs. To verify that this is a geometric setting, note that rolling a 3 will represent a success, and rolling any other number will represent a failure. The observations are independent. We roll the die until a 3 appears. The probability of rolling a 3 on each roll is the same: 1/6.

(a) Plot the cumulative distribution histogram (out to X = 10) for the die-rolling experiment described above.

The cumulative distribution histogram (out to $X = 10$) for rolling a die is shown below. Note that the cumulative function value for $X = 10$ is only 0.8385. Many more bars are needed for it to reach a height of 1.



(b) Find the probability that it takes more than 10 rolls to observe a 3.

$P(X > 10) = (1-1/6)^{10} = (5/6)^{10} = 0.1615.$

(c) Find the smallest positive integer *k* for which $P(X \le k) > 0.99$.

The smallest positive integer k for which $P(X \le k) > 0.99$ is k = 26 (see second calculator screen above).

8.45 **SHOOTING FREE THROWS** A basketball player makes 80% of her free throws. We put her on the free-throw line and ask her to shoot free throws until she misses one. Let *X* = the number of free throws the player takes until she misses.

(a) What assumption do you need to make in order for the geometric model to apply? With this assumption, verify that *X* has a geometric distribution. What action constitutes "success" in this context?

X = number of free throws shot until a miss
Success = a miss

B: Only two choices, hit or miss
I: The shots are independent
T: keep shooting until you have a miss
S: Probability of success (a miss) is the same for each shot, $p = 0.2$

(b) What is the probability that the player will make 5 shots before she misses?

The first "success" (miss) is the sixth shot, so $X = 6$ and

$P(X = 6) = (1-p)^{n-1} p = (0.8)^5 (0.2) = 0.0655.$

(c) What is the probability that the player will make at most 5 shots before she misses?

$P(X \le 6) = 1 - P(X > 6) = 1 - (1-p)^6 = 1 - (0.8)^6 = 0.738$

$or\ P(X \le 6) = geometcdf\ (0.2, 6) = 0.738$

8.46 **GAME OF CHANCE** Three friends each toss a coin. The odd man wins; that is, if one coin comes up different from the other two, that person wins that round. If the coins all match, then no one wins and they toss again. We're interested in the number of times the players will have to toss the coins until someone wins.

(a) What is the probability that no one will win on a given coin toss?

Out of 8 possible outcomes, HHH and TTT do not produce winners. So P(no winner) = 0.25.

(b) Define a success as "someone wins on a given toss." What is the probability of a success?

P(winner) = 1 − 0.25 = 0.75.

(c) Define the random variable of interest: X = number of coin tosses until someone wins. Is X binomial? Geometric? Justify your answer. X is geometric

B: Only two choices, win or lose

I: The tosses are independent

T: keep tossing until you have a win

S: Probability of success is the same for each toss, $p = 0.75$

(d) Construct a probability distribution table for X. Then extend your table by the addition of cumulative probabilities in a third row.

X	1	2	3	4	5	. . .
P(X)	0.75	0.1875	0.04688	0.01172	0.00293	
cdf	0.75	0.9375	0.9844	0.9961	0.9990	

(e) What is the probability that it takes no more than 2 rounds for someone to win?

$P(X \leq 2) = 0.9375$ from the table.

(f) What is the probability that it takes more than 4 rounds for someone to win?

$P(X > 4) = (0.25)^4 = 0.0039$.

(g) What is the expected number of tosses needed for someone to win?

$$m = \frac{1}{p} = \frac{1}{0.75} = 1.33$$

HW #5 49, 55, 57, 59

8.49 **MULTIPLE-CHOICE** Carla makes random guesses on a multiple-choice test that has five choices for each question. We want to know how many questions Carla answers until she gets one correct.

(a) Define a success in this context, and define the random variable *X* of interest. What is the probability of success?

Success = getting a correct answer. *X* = number of questions Carla must answer in order to get the first correct answer. p = 1/5 = 0.2. (all 5 choices equally likely to be selected).

(b) What is the probability that Carla's first correct answer occurs on problem 5?

$P(X=5)=(1-1/5)^{5-1}(1/5)=(4/5)^4(1/5)=0.082$

(c) What is the probability that it takes more than 4 questions before Carla answers one correctly?

$P(X>4)=(1-1/5)^4=(4/5)^4=0.4096$

(d) Construct a probability distribution table for *X*.

X	1	2	3	4	5	. . .
P(X)	0.2	0.16	0.128	0.1024	0.082	. . .

(e) If Carla took a test like this test many times and randomly guessed at each question, what would be the average number of questions she would have to answer before she answered one correctly?

$$m=\frac{1}{p}=\frac{1}{1/5}=5.$$

8.55 **BINOMIAL SETTING?** In each of the following cases, decide whether or not a binomial distribution is an appropriate model, and give your reasons.

(a) You want to know what percent of married people believe that mothers of young children should not be employed outside the home. You plan to interview 50 people, and for the sake of convenience you decide to interview both the husband and wife in 25 married couples. The random variable *X* is the number among the 50 persons interviewed who think mothers should not be employed.

No. It is not reasonable to assume that the opinions of a husband and wife (especially on such an issue as mothers working outside the home) are independent.

(b) You are interested in attitudes toward drinking among the 75 members of a fraternity. You choose 25 members at random to interview. One question is "Have you had five or more drinks at one time during the last week?" Suppose that in fact 20% of the 75 members would say "Yes." Explain why you cannot safely use the *B*(25, 0.2) distribution for the count *X* in your sample who say "Yes."

No. The sample size (25) is so large compared to the population size (75) that the probability of success ("Yes") will substantially change as we move from person to person within the sample. The population should be at least 10 times larger than the sample in order for the binomial setting to be valid.

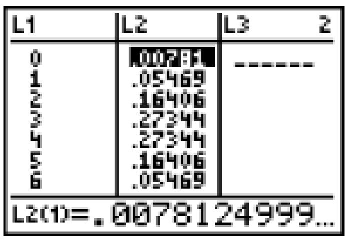
8.57 **SEVEN BROTHERS!** There's a movie classic entitled *Seven Brides for Seven Brothers.* Even if these brothers had a few sisters, this many brothers in unusual. We will assume then that there are no sisters.

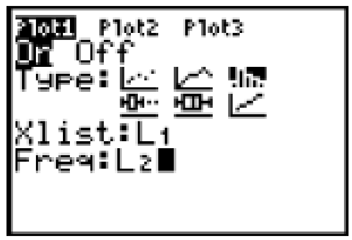
(a) Let X = number of boys in a family of 7 children. Assume that sons and daughters are equally likely outcomes. Do you think the distribution of X will be skewed left, symmetric, or skewed right? The answer to this question depends on what fact?

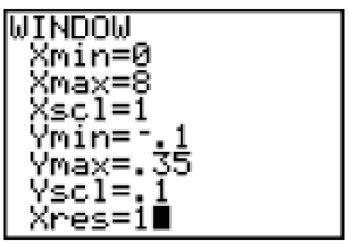
The distribution of X = B(7, 0.5) is symmetric; the shape depends on the value of the probability of success. Since 0.5 is halfway between 0 and 1, the histogram is symmetric.

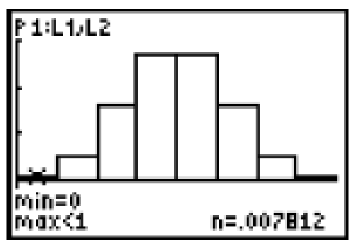
(b) Use the binompdf command to construct a pdf table for X. Then construct a probability distribution histogram and a cumulative distribution histogram for X. Keep a written record of your numerical results as they are produced by your calculator, as well as sketches of the histograms.

With the values of X = 0, 1, . . ., 7 in L1, define L2 to be binompdf (7, 0.5).
Then the probability table for B(7, 0.5) is installed in L1 and L2. Here is a histogram of the pdf.:

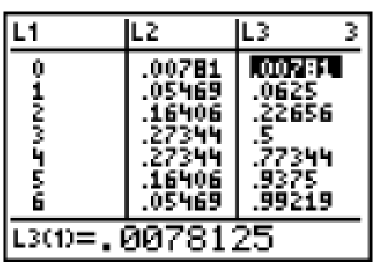


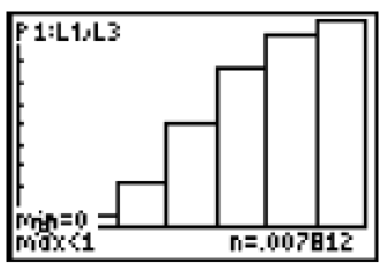






Now, define L3 to be binomcdf (7, 0.5, L1). Then the cdf for the B(7, 0.5) distribution is installed in L3. Here is the histogram of the cdf:





(c) What is the probability that all of the 7 children are boys?

P(X = 7) = binompdf (7, 0.5, 7)= 0.0078125.

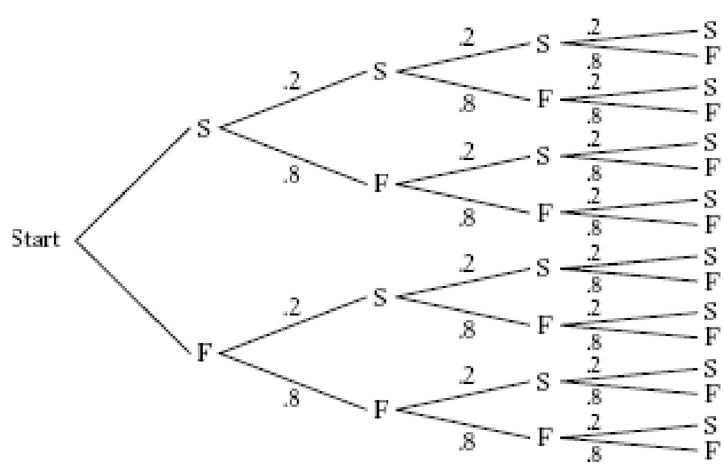
8.59 **ARMED AND DANGEROUS** According to a 1997 Centers for Disease Control study of risky behavior, roughly one in five teenagers carries a weapon. Hispanics were most likely to arm themselves, with 23% carrying a gun, knife or club, compared with 22% of blacks and 17% of whites. Suppose that 4 teenagers are selected at random and subjected to a search. Suppose that success is defined as "the teen is carrying a gun, knife, or club."

(a) What is the probability, *p*, of success? *p* = one in five = 0.2.

(b) Make a list of all the possible results of the search of 4 teenagers. Use S to represent success, and F for failure. For each of these responses, write a product of four factors for that combination of successes and failures. For example, SSFS is one such response, and the probability of that outcome is (0.2)(0.2)(0.8)(0.2) = 0.0064. Display the probabilities to four decimal places.

Result	Probability
SSSS	(.2)(.2)(.2)(.2) = .0016
SSSF	(.2)(.2)(.2)(.8) = .0064
SSFS	(.2)(.2)(.8)(.2) = .0064
SFSS	(.2)(.8)(.2)(.2) = .0064
FSSS	(.8)(.2)(.2)(.2) = .0064
SSFF	(.2)(.2)(.8)(.8) = .0256
SFSF	(.2)(.8)(.2)(.8) = .0256
SFFS	(.2)(.8)(.8)(.2) = .0256
FSFS	(.8)(.2)(.8)(.2) = .0256
FSSF	(.8)(.2)(.2)(.8) = .0256
FFSS	(.8)(.8)(.2)(.2) = .0256
SFFF	(.2)(.8)(.8)(.8) = .1024
FSFF	(.8)(.2)(.8)(.8) = .1024
FFSF	(.8)(.8)(.2)(.8) = .1024
FFFS	(.8)(.8)(.8)(.2) = .1024
FFFF	(.8)(.8)(.8)(.8) = .4096

(c) Draw a tree diagram to show the possible outcomes.



(d) List the outcomes in which exactly 2 of the 4 students are found to carry a gun, knife, or club.

SSFF, SFSF, SFFS, FSFS, FSSF, FFSS.

(e) What are the probabilities of the outcomes in part (d)? Briefly explain why all of these probabilities are the same.

Each outcome has probability 0.0256. The probabilities all involve the same number of each factor; they are simply multiplied in different orders, which will not affect the product.