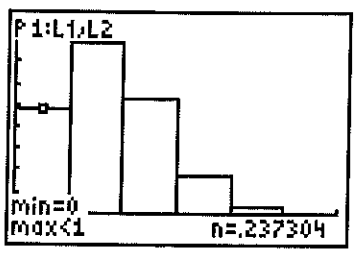
8

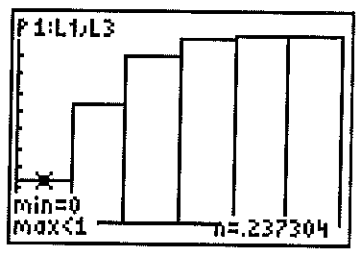
The Binomial and Geometric Distributions

8.1 (a) No: There is no fixed n (i.e., there is no definite upper limit on the number of defects). (b) Yes: It is reasonable to believe that all responses are independent (ignoring any "peer pressure"), and all have the same probability of saying "yes" since they are randomly chosen from the population. Also, a "large city" will have a population over 1000 (10 times as big as the sample). (c) Yes: In a "Pick 3" game, Joe's chance of winning the lottery is the same every week, so assuming that a year consists of 52 weeks (observations), this would be binomial.

8.2 (a) Yes: It is reasonable to assume that the results for the 50 students are independent, and each has the same chance of passing. (b) No: Since the student receives instruction after incorrect answers, her probability of success is likely to increase. (e) No: Temperature may affect the outcome of the test.

8.3





(a) .2637. (b) The binomial probabilities for $x = 0, \ldots, 5$ are: .2373, .3955, .2637, .0879, .0146, .0010. (e) The cumulative probabilities for $x = 0, \ldots, 5$ are: .2373, .6328, .8965, .9844, .9990, 1. Compared with Corinne's cdf histogram, the bars in this histogram get taller, sooner. Both peak at 1 on the extreme right.

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

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

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

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

8.5 (a) Let X = the number of correct answers. X is binomial with $n = 10$, $p = 0.25$. The probability of at least one correct answer is $P(X \geq 1) = 1 - P(X = 0) = 1 -$ `binompdf (10, .25, 0)` $= 1 - .056 = .944$.

(b) Let X = the number of correct answers. We can write $X = X_1 + X_2 + X_3$, where X_i = the number of correct answers on question i. (Note that the only possible values of X_i are 0 and 1, with 0 representing an incorrect answer and 1 a correct answer.) The probability of at least one

correct answer is $P(X \geq 1) = 1 - P(X = 0) = 1 - [P(X_1 = 0)P(X_2 = 0)\, P(X_3 = 0)]$ (since the X_i are independent) $= 1 - \left(\frac{2}{3}\right)\left(\frac{3}{4}\right)\left(\frac{4}{5}\right) = 1 - \frac{24}{60} = 0.6$.

8.6 (a) Yes, if the 100 children are randomly selected, it is extremely likely that the result for one child will not be influenced by the result for any other child (e.g., the children are siblings). "Success" in this context means having an incarcerated parent. $n = 100$, since 100 children are selected, and $p = 0.02$.

(b) $P(X = 0)$ = the probability of none of the 100 selected children having an incarcerated parent. $P(X = 0) =$ `binompdf (100, .02, 0)` $= .133$. $P(X = 1) =$ `binompdf (100, .02, 1)` $= .271$.

(c) $P(X \geq 2) = 1 - P(X \leq 1) = 1 -$ `binomcdf (100, .02, 1)` $= 1 - .403 = .597$. Alternatively, by the addition rule for mutually exclusive events, $P(X \geq 2) = 1 - (P(X = 0) + P(X = 1)) = 1 - (.133 + .271) = 1 - .404 = .596$. (The difference between answers is due to roundoff error.)

8.7 Let X = the number of players out of 20 who graduate. $P(X = 11) =$ `binompdf (20, .8, 11)` $= .0074$.

8.8 (a) $n = 10$ and $p = 0.25$. (b) $\binom{10}{2}(0.25)^2(0.75)^8 = 0.28157$. (c) $P(X \leq 2) = \binom{10}{0}(0.25)^0(0.75)^{10} + \ldots + \binom{10}{2}(0.25)^2(0.75)^8 = 0.52559$.

8.9 $P(X = 3) = \binom{5}{3}(.25)^3(.75)^2 = (10)(.25)^3(.75)^2 = .088$.

8.10 Let X = the number of broccoli plants that you lose. X is $B(10, .05)$.

$$P(X \leq 1) = P(X = 0) + P(X = 1) = \binom{10}{0}(.05)^0(.95)^{10} + \binom{10}{1}(.05)^1(.95)^9 = (.95)^{10} + (10)(.05)(.95)^9 = .914.$$

8.11 Let X = the number of children with blood type O. X is $B(5, .25)$.

$$P(X \geq 1) = 1 - P(X = 0) = 1 - \binom{5}{0}(.25)^0(.75)^5 = 1 - (.75)^5 = .763.$$

8.12 Probability that all 20 graduate: $P(X = 20) = \binom{20}{20}(.8)^{20}(.2)^0 = (.8)^{20} = .0115$.

Probability that not all 20 graduate: $P(X < 20) = 1 - P(X = 20) = .9885$.

8.13 (a) $\binom{15}{3}(0.3)^3(0.7)^{12} = 0.17004$. (b) $\binom{15}{0}(0.3)^0(0.7)^{15} = .00475$.

8.14 Let X = the number of free throws that Corinne makes. X is $B(12, .7)$.

$$P(X = 7) = \binom{12}{7}(.75)^7(.25)^5 = (792)(.75)^7(.25)^5 = .1032.$$

8.15 (a) $np = 1500$, $n(1 - p) = 1000$; both values are greater than or equal to 10. (b) Let X = the number of people in the sample who find shopping frustrating. X is $B(2500, .6)$. Then $P(X \geq 1520) = 1 - P(X \leq 1519) = 1 -$ `binomcdf (2500, .6, 1519)` $= 1 - .7868609113 = .2131390887$, which rounds to .2131. The probability correct to six decimal places is .213139. (c) $P(X \leq 1468) =$ `binomcdf (2500, .6, 1468)` $= .0994$. Using the normal approximation to the binomial yields .0957, a difference of .0037.

8.16 (a) $\mu = 4.5$. (b) $\sigma = \sqrt{3.15} = 1.77482$. (c) If $p = 0.1$, then $\sigma = \sqrt{1.35} = 1.16190$. If $p = 0.01$, then $\sigma = \sqrt{0.1485} = 0.38536$. As p gets close to 0, σ gets closer to 0.

8.17 (a) $\mu = (20)(.8) = 16$. (b) $\sigma = \sqrt{(20)(.8)(.2)} = \sqrt{3.2} = 1.789$. (c) For $p = 0.99$, $\sigma = \sqrt{(20)(.99)(.01)} = \sqrt{0.198} = .445$. As the probability of success gets closer to 1, the standard deviation decreases. (Note that as p approaches 1, the probability histogram of the binomial distribution becomes increasingly skewed, and thus there is less and less chance of seeing an observation of the binomial at an appreciable distance from the mean.)

8.18 $\mu = 2.5, \sigma = \sqrt{1.875} = 1.36931$.

8.19 (a) X = the number of people in the sample of 400 adult Richmonders who approve of the President's reaction. X is approximately binomial because the sample size is small compared to the population size (all adult Richmonders), and as a result, the individual responses may be considered independent and the probability of success (approval) remains essentially the same from trial to trial. $n = 400$ and $p = .92$.

(b) $P(X \le 358) =$ `binomcdf (400, .92, 358)` $= .0441$.

(c) $\mu = (400)(.92) = 368, \sigma = \sqrt{(400)(.92)(.08)} = \sqrt{29.44} = 5.426$.

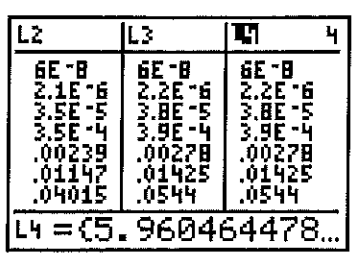
(d) $P(X \le 358) \approx P(Z \le \frac{358 - 368}{\sqrt{29.44}}) = P(Z \le -1.843) \approx .0327$. The approximation is not very accurate (note that p is close to 1).

8.20 (a) Yes. Assuming that the sample size is small compared to the population size—a reasonable assumption if your area is heavily populated—this study satisfies the requirements of a binomial setting. $n = 200$ because 200 households are surveyed, and $p = 0.4$ because the probability of success (being committed to eating nutritious food away from home) is 40% and remains essentially the same from trial to trial.

(b) $\mu = (200)(.4) = 80$. $\sigma = \sqrt{200(.4)(.6)} = \sqrt{48} = 6.923$.

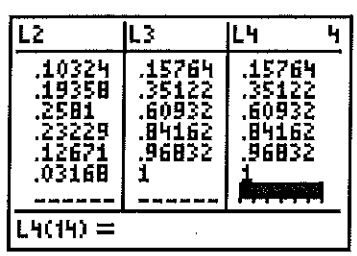
(c) $np = 80$ and $n(1 - p) = 120$, so the rule of thumb is satisfied here. $P(75 \le X \le 85) \approx P(\frac{75 - 80}{\sqrt{48}} \le Z \le \frac{85 - 80}{\sqrt{48}}) = P(-0.72 \le Z \le 0.72) = 0.5285$.

8.21 The command `cumSum (L2) → L3` calculates and stores the values of $P(X \le x)$ for $x = 0, 1, 2, \ldots, 12$. The entries in L_3 and the entries in L_4 defined by `binomcdf (12, .75, L1) → L4` are identical (see below).



L2	L3	L4 4
6E-8	6E-8	6E-8
2.1E-6	2.2E-6	2.2E-6
3.5E-5	3.8E-5	3.8E-5
3.5E-4	3.9E-4	3.9E-4
.00239	.00278	.00278
.01147	.01425	.01425
.04015	.0544	.0544

L4 =5.960464478...



L2	L3	L4 4
.10324	.15764	.15764
.19358	.35122	.35122
.2581	.60932	.60932
.23229	.84162	.84162
.12671	.96832	.96832
.03168	1	1
------	------	[redacted]

L4(14) =

The first screen holds the probabilities for X = 0, 1, 2, 3, 4, 5, 6. The second holds the probabilities for X = 7, 8, 9, 10, 11, 12.

8.22 (a) We simulate 10 observations of X = number of defective switches (for which $n = 10$, $p = .1$) by using the command `randBin (1, .1, 10) → L1: sum (L1) → L2 (1)`. (Press ENTER 10 times.) The observations for one sample simulation are: 0, 0, 4, 0, 1, 0, 1, 0, 0, 1. For these data, $\bar{x} = .7$. To generate 25/50 observations of x, replace 10 with 25/50 in the `randBin` command above. As the number of observations increases, the resulting $\bar{x}$ should approximate the known mean $\mu = 1$ more closely, by the law of large numbers.

(b) Simulate 10 observations of X = number of free throws Corinne makes (where the number of trials is $n = 12$ and the probability of a basket on a given trial is $p = .75$) by using the command `randBin (1, .75, 12)` $\rightarrow L_1$: `sum (`L_1`)` $\rightarrow L_2$ `(1)`. (Press ENTER 10 times.) The observations for one sample simulation are: 9, 8, 10, 10, 9, 8, 10, 9, 8, 8. For these data, $\bar{x} = 8.9$. Compare this with the known mean $\mu = np = (12)(.75) = 9$. To generate 25/50 observations of $\bar{x}$, replace 10 with 25/50 in the `randBin` command above. As the number of observations increases, the resulting $\bar{x}$ should approximate the known mean $\mu = 1$ more closely, by the law of large numbers.

8.23 There are $n = 15$ people on the committee, and the probability that a randomly selected person is Hispanic is $p = .3$. Let 0, 1, 2 $\Leftrightarrow$ Hispanic and let 3–9 $\Leftrightarrow$ non-Hispanic. Use the random digit table. Or, using the calculator, repeat the command 30 times: `randBin (1, .3, 15)` $\rightarrow L_1$: `sum (`L_1`)` $\rightarrow L_2$ `(1)` where 0 = non-Hispanic, and 1 = Hispanic. Our frequencies were:

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

For this simulation, the relative frequency of 3 or fewer Hispanics was 7/30 = .233. Compare this with the theoretical result: $P(X \leq 3) = 0.29687$, where X = number of Hispanics on the committee.

8.24 The probability of a success (having four or more credit cards) in this case is $p = 0.33$. Let the two-digit groups 00, 01, 02, ..., 32 represent students who have four or more credit cards, and let 33, 34, 35, ..., 99 represent students who do not. Choose 30 two-digit groups from the table of random digits (Table B). Starting at Line 141 in the table, for example, we get

96	76	73	59	64	23	82	29	60	12
94	59	16	51	94	50	84	25	33	72
72	82	95	02	32	97	89	26	34	08

The total number of "students with four or more credit cards" here is 9. Repeat this process a total of 30 times to obtain 30 observations of the binomial variable X (# of students with four or more cards) and estimate $P(X > 12)$ by the relative frequency of observations of X that exceed 12.

If the TI-83 is used, we can generate 30 observations of X by the command `randBin (1, .33, 30)` $\rightarrow L_1$: `sum (`L_1`)` $\rightarrow L_2$ `(1)`, with ENTER being pressed 30 times to produce 30 simulated observations. A sample simulation yielded the following results:

10	13	11	11	6	13	9	9	11	11
10	6	8	6	8	7	13	10	13	10
6	13	3	14	9	9	11	8	14	12

Here, $P(X > 12) = 7/30 = 0.233$. The actual value of $P(X > 12)$ is `1 − binomcdf (30, .33, 12)` = 0.1563.

8.25 The sample size is $n = 10$, and the probability that a randomly selected employed woman has never been married is $p = 0.25$. Let 0 $\Leftrightarrow$ never married, let 1, 2, 3, $\Leftrightarrow$ married, and use Table B. Or, using the calculator, repeat the command `randBin (1, .25, 10)` $\rightarrow L_1$: `sum (`L_1`)`. Our results for 30 repetitions were:

2	3	14	5	4	2
0	1	2	3	4	5

Our relative frequency of "2 or fewer never married" was $19/30 = .63$. The actual value of $P(X \leq 2) =$ `binomcdf (10, .25, 2)` $= .525$.

8.26 (a) The probability of drawing a white chip is $\frac{15}{50} = \frac{3}{10} = .3$. The number of white chips in 25 draws is $B(25, .3)$. Therefore, the expected number of white chips is $(25)(.3) = 7.5$.

(b) The probability of drawing a blue chip is $\frac{10}{50} = \frac{2}{10} = .2$. The number of blue chips in 25 draws is $B(25, .2)$. Therefore, the standard deviation of the number of blue chips is $\sqrt{(25)(.2)(.8)} = 2$.

(c) Let the digits 0, 1, 2, 3, 4 $\Leftrightarrow$ red chip, 5, 6, 7 $\Leftrightarrow$ white chip, and 8, 9 $\Leftrightarrow$ blue chip. Draw 25 random digits from Table B and record the number of times that you get chips of various colors. If you are using the TI-83, you can draw 25 random digits using the command `randInt (0, 9, 25)` $\rightarrow$ L_1. 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-83 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.

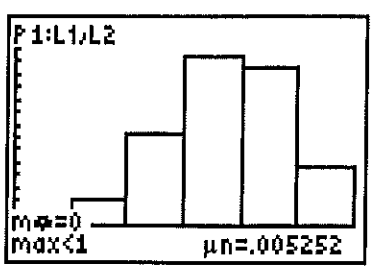
(d) The expected number of blue chips is $(25)(.2) = 5$, and the standard deviation $= 2$ by part (b). It seems extremely likely that you will draw at most 9 blue chips. The actual probability is `binomcdf (25, .2, 9)` $= .9827$.

(e) 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)` $= .999998$.

8.27 The count of 0s among n random digits has a binomial distribution with $p = 0.1$. (a) $P(\text{at least one } 0) = 1 - P(\text{no } 0) = 1 - (0.9)^5 = 0.40951$. (b) $\mu = (40)(0.1) = 4$.

8.28 (a) $n = 20$ and $p = 0.25$. (b) $\mu = 5$. (c) $\binom{20}{5}(0.25)^5(0.75)^{15} = 0.20233$.

8.29 (a) $n = 5$ and $p = 0.65$. (b) X takes values from 0 to 5. (c) $P(X = 0) = .0053$, $P(X = 1) = .0488$, $P(X = 2) = .1815$, $P(X = 3) = .3364$, $P(X = 4) = .3124$, $P(X = 5) = .11603$. Histogram below. (d) $\mu = (5)(.65) = 3.25$, $\sigma = \sqrt{(5)(.65)(.35)} = \sqrt{1.1375} = 1.067$.



8.30 (a) 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) $\mu = 2.4$, $\sigma = \sqrt{1.92} = 1.38564$.

(c) $P(X < 2.4) = P(X \leq 2) =$ `binomcdf(12, .2, 2)` $= .5583$.

8.31 In this case, $\mu = (200)(.4) = 80$ and $\sigma = \sqrt{(200)(.4)(.6)} = 6.9282$. Using the normal approximation to the binomial distribution, $P(X \geq 100) \approx P(Z \geq 2.89) = .0019$. Using the exact binomial distribution, $P(X \geq 100) = 1 - P(X \leq 99) = 1 -$ `binomcdf (200, .4, 99)` $= .0026$. Regardless of how we compute the probability, this is strong evidence that the local percentage of households concerned about nutrition is higher than 40%.

8.32 (a) There are 150 independent observations, each with probability of success (response) = .5. (b) $\mu = (150)(.5) = 75$ responses. (c) $P(X \leq 70) \approx P(Z \leq -0.82) = .2061$; using unrounded values and software yields .2071. (d) Use 200, since $(200)(.5) = 100$.

8.33 (a) Let X = the number of orders shipped on time. X is approximately $B(100, .9)$, since there are 100 independent observations, each with probability .9 (the population size is more than 10 times the sample size). $\mu = (100)(.9) = 90$ and $\sigma = \sqrt{(100)(.9)(.1)} = 3$. The normal approximation can be used, since the second rule of thumb is *just* satisfied; $n(1 - p) = 10$. $P(X \leq 86) \approx P(Z \leq -1.33) = .0918$; the software value is .0912.

(b) Even when the claim is correct, there will be some variation in the sample proportions. In particular, in about 9% of all samples, we can expect to see 86 or fewer orders shipped on time.

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

(b) $P(X = 20) = 0.81791$; $P(X < 20) = 0.18209$.

(c) $\mu = 19.8$, $\sigma = \sqrt{0.198} = 0.44497$.

8.35 Identical to Exercise 8.7, except that in this case we calculate $P(X \leq 11)$ for the 30 simulated observations. The actual value of $P(X \leq 11)$ is `binomcdf (20, .8, 11)` $= .00998$ or approximately .01.

8.36 (a) $P(X = 3) = .0613$. (b) $P(X \leq 2) = .9257$. (c) $P(X < 2) = P(X \leq 1) = .7206$. (d) $P(3 \leq X \leq 5) = P(X \leq 5) - P(X \leq 2) = .9999 - .9257 = .0742$. (e) $P(X < 2 \text{ or } X > 5) = P(X < 2) + P(X > 5) = P(X \leq 1) + (1 - P(X \leq 5)) = .7206 + (1 - .9999) = .7207$.

8.37 (a) Geometric setting; success = tail, failure = head; trial = flip of coin; $p = 1/2$.

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

(c) Geometric setting; success = jack, failure = any other card; trial = drawing of a card; $p = 4/52 = 1/13$. (Trials are independent because the card is replaced each time.)

(d) Geometric setting, success = match all 6 numbers, failure = do not match all 6 numbers; trial = drawing on a particular day; the probability of success is the same for each trial; $p = \frac{6!}{\binom{44}{6}} = .000102$; and trials are independent because the setting of a drawing is always the same and the results on different drawings do not influence each other.

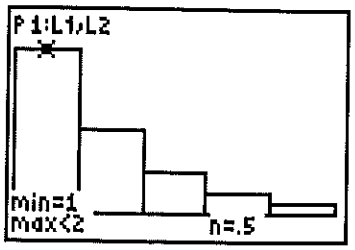
(e) Not a geometric setting. The trials (draws) are not independent because you are drawing without replacement. Also, you are interested in getting 3 successes, rather than just the first success.

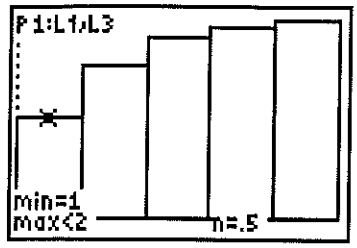
8.38 (a) The four conditions of a geometric setting hold, with probability of success 5 1/2.

(b) and (d)

X	1	2	3	4	5	...
P(X)	.5	.25	.125	.0625	.03125	
c.d.f.	.5	.75	.875	.9375	.96875	

(c) and (d)





(e) Sum $= \dfrac{a}{1-r} = \dfrac{.5}{1-r} = 1.$

8.39 (a) X = number of drives tested in order to find the first defective. Success = defective drive. This is a geometric setting because the trials (tests) on successive drives are independent, $p = .03$ on each trial, and X is counting the number of trials required to achieve the first success.

(b) $P(X = 5) = (1 - .03)^{5-1}(.03)$
$= (0.97)^4(.03)$
$= .0266.$

(c)

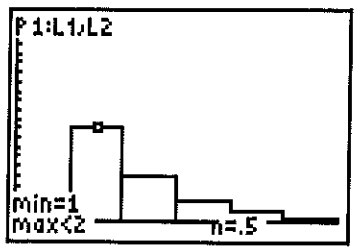
X	1	2	3	4
$P(X)$	.03	.0291	.0282	.0274

8.40 Parts (a), (c), and (d) of Exercise 8.37 constituted geometric settings. (a) $P(X = 4) = (.5)^3(.5) = (.5)^4 = .0625$. (c) $P(X = 4) = (\frac{12}{13})^3(\frac{1}{13}) = .0605$. (d) $P(X = 4) = (.999898)^3(.000102) = .0001$.

8.41 (a) X = number of flips required in order to get the first head. X is a geometric random variable with $p = .5$.

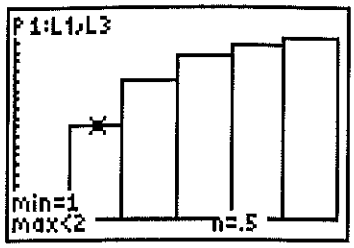
(b) $P(X = x) = (.5)^{x-1}(.5) = (.5)^x$ for $x = 1, 2, 3, 4, \ldots$

x	1	2	3	4	5	...
$P(X = x)$	.5	.25	.125	.0625	.03125	...



(c) $P(X \le x) = (.5)^1 + \ldots + (.5)^x$ for $x = 1, 2, 3, 4, \ldots$

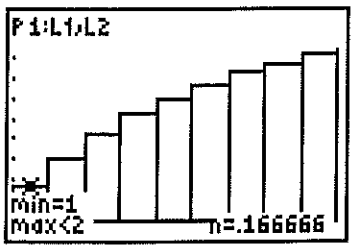
x	1	2	3	4	5	...
$P(X \le x)$	.5	.75	.875	.9375	.96875	...

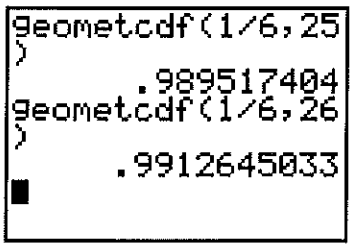


8.42 (a) $P(X > 10) = (1 - \frac{1}{12})^{10} = (\frac{11}{12})^{10} = .419$.

(b) $P(X > 10) = 1 - P(X \leq 10) = 1 -$ `geometcdf(1/12, 10)` $= .419$.

8.43 (a) 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 .8385. Many more bars are needed for it to reach a height of 1.





(b) $P(X > 10) = (1 - 1/6)^{10} = (5/6)^{10} = 1615$. (c) The smallest positive integer k for which $P(X \leq k) > .99$ is $k = 26$ (see second calculator screen above).

8.44 Let X = the number of applicants who need to be interviewed in order to find one who is fluent in Farsi. X is geometric with $p = 4\% = .04$. (a) $\mu = \frac{1}{p} = \frac{1}{.04} = 25$. (b) $P(X > 25) = (1 - .04)^{25} = (.96)^{25} = .3604$; $P(X > 40) = (.96)^{40} = .1954$.

8.45 (a) Assumptions needed for the geometric model to apply are that the shots are independent, and that the probability of success is the same for each shot. A "success" is a missed shot, so the probability of success is $p = 0.2$. The four conditions for a geometric setting are satisfied.

(b) The first "success" (miss) is the sixth shot, so $X = 6$ and $P(X = 6) = (1 - p)^{n-1}p = (.8)^5(.2) = .0655$.

(c) $P(X \leq 6) = 1 - P(X > 6) = 1 - (1 - p)^6 = 1-(.8)^6=.738$ or $P(X \leq 6) =$ `geometcdf (.2, 6)` $= .738$.

8.46 (a) Out of 8 possible outcomes, HHH and TTT do not produce winners. So P(no winner) $= 2/8 = .25$.

(b) P(winner) $= 1 - .25 = .75$.

(c) Let X = number of coin tosses until someone wins. Then X is geometric because all four conditions for a geometric setting are satisfied.

(d)

X	1	2	3	4	5	...
P(X)	.75	.1875	.04688	.01172	.00293	
c.d.f.	.75	.9375	.9844	.9961	.9990	

(e) $P(X \le 2) = .9375$ from the table. (f) $P(X > 4) = (.25)^4 = .0039$. (g) $\mu = 1/p = 1/.75 = 1.33$. (h) Let 1 ⇔ heads and 0 ⇔ tails, and enter the command `randInt (0, 1, 3)` and press ENTER 25 times. In our simulation, we recorded the following frequencies:

X	1	2	3
Freq.	21	3	1
Rel. freq.	.84	.12	.04

These compare with the calculated probabilities of .75, .1875, and .04688, respectively. A larger number of trials should result in somewhat better agreement (Law of Large Numbers).

8.47 (a) Geometric setting; X = number of marbles you must draw to find the first red marble. We choose geometric in this case because the number of trials (draws) is the variable quantity.

(b) $P = 20/35 = 4/7$ in this case, so

$$P(X = 2) = (1 - 4/7)^{2-1}(4/7) = (3/7)(4/7) = 12/49 = .2449$$

$$P(X \le 2) = 4/7 + (3/7)(4/7) = 4/7 + 12/49 = 40/49 = .8163$$

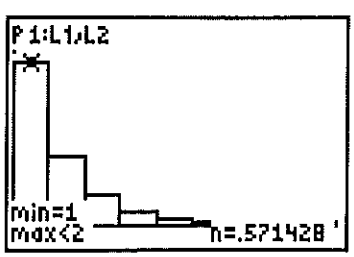
$$P(X > 2) = (1 - 4/7)^2 = (3/7)^2 = 9/49 = .1837$$

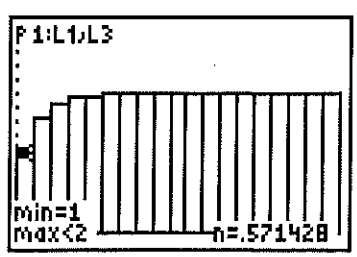
(c) Use the TI-83 commands `seq (X, X, 1, 20)` → L_1, `geometpdf (4/7, L1)` → L_2, `cumSum (L2)` → L_3 (or `geometcdf (4/7, L1)` → L_3).

X	1	2	3	4	5	6	7	8	9	10
P(X)	.571	.245	.105	.045	.019	.008	.004	.002	.001	.000
F(X) c.d.f.	.571	.816	.921	.966	.986	.994	.997	.999	.99951	.9998

X	11	12	13	14	15	16	17	18	19	20
P(X)	.000	.000	.000	.000	.000	.000	.000	.000	.000	.000
F(X)	.9999	.9999	.9999	1	1	1	1	1	1	1

(d) The probability distribution histogram is below left; the cumulative distribution is below right.





8.48 (a) No. Since the marbles are being drawn *without* replacement and the population (the set of all marbles in the jar) is so small, the results of any draw will clearly be dependent upon the results of previous draws. Also, the geometric variable measures the number of trials

required to get the *first* success; here, we are looking for the number of trials required to get the *second* success.

(b) No. Even though the results of the draws are now independent, the variable being measured is still not the geometric variable.

(c) The probability of getting a red marble on any draw is $\frac{20}{35} = \frac{4}{7}$. Let the digits 0, 1, 2, 3 $\Leftrightarrow$ a red marble is drawn, 4, 5, 6 $\Leftrightarrow$ some other color marble is drawn, and 7, 8, 9 $\Leftrightarrow$ digit is disregarded. Start choosing random digits from Table B, or use the TI-83 command `randInt(0, 9, 1)` repeatedly. After two digits in the set 0, 1, 2, 3 have been chosen, stop the process and count the number of digits in the set 0, 1, 2, 3, 4, 5, 6 that have been chosen up to that point; this represents an observation of X. Repeat the process until the desired number of observations of X have been obtained. Here are some sample simulations using the TI-83 (with R = red marble, O = other color marble, D = disregard):

7	0	4	3		X = 3	
D	R	O	R			
9	0	8	6	2	X = 3	
D	R	D	O	R		
9	7	3	2		X = 2	etc.
D	D	R	R			

8.49 (a) Success = getting a correct answer. X = number of questions Carla must answer in order to get the first correct answer. $p = 1/5 = .2$ (all 5 choices equally likely to be selected).

(b) $P(X = 5) = (1 - 1/5)^{5-1}(1/5) = (4/5)^4(1/5) = .082$.

(c) $P(X > 4) = (1 - 1/5)^4 = (4/5)^4 = .4096$.

(d)

X	1	2	3	4	5	....
$P(X)$	.2	.16	.128	.1024	.082	....

(e) $\mu = 1/(1/5) = 5$.

8.50 (a) If "success" = son and p (success) = .5, then the average number of children per family is $\mu = 1/p = 1/.5 = 2$.

(b) If the average size of the family is 2, and the last child is a boy, then the average number of girls per family is 1.

(c) Let even digit = boy, and odd digit = girl. Read random digits until an even digit occurs. Count number of digits read. Repeat many times, and average the counts. Beginning on line 101 in the random digit table and simulating 50 trials, the average number of children per family is 1.96, and the average number of girls is .96. These are very close to the expected values.

8.51 (a) Letting G = girl and B = boy, the outcomes are: {G, BG, BBG, BBBG, BBBB}. Success = having a girl.

(b) X = number of boys can take values of 0, 1, 2, 3, or 4. The probabilities are calculated by using the multiplication rule for independent events:

$P(X = 0) = 1/2$
$P(X = 1) = (1/2)(1/2) = 1/4$
$P(X = 2) = (1/2)(1/2)(1/2) = 1/8$
$P(X = 3) = (1/2)(1/2)(1/2)(1/2) = 1/16$
$P(X = 4) = (1/2)(1/2)(1/2)(1/2) = 1/16$

X	0	1	2	3	4
P(X)	1/2	1/4	1/8	1/16	1/16

Note that $\Sigma P(X) = 1$.

(c) Let Y = number of children produced until first girl is seen. Then Y is a geometric variable for Y = 1 up to Y = 4, but then "stops" because the couple plans to stop at 4 children if it does not see a girl by that time. By the multiplication rule,

$$P(Y = 1) = 1/2$$
$$P(Y = 2) = 1/4$$
$$P(Y = 3) = 1/8$$
$$P(Y = 4) = 1/16$$

Note that the event $\{Y = 4\}$ can only include the outcome BBBG. BBBB must be discarded. The probability distribution table would begin

Y	1	2	3	4
P(Y)	1/2	1/4	1/8	1/16

But note that this table is incomplete and this is not a valid probability model since $\Sigma P(Y) < 1$. The difficulty lies in the way Y was defined. It does not include the possible outcome BBBB.

(d) Let Z = number of children per family. Then

Z	1	2	3	4
P(Z)	1/2	1/4	1/8	1/16

and $\mu_z = \Sigma(Z \times P(Z)) = (1)(1/2) + (2)(1/4) + (3)(1/8) + (4)(1/8)$
$= 1/2 + 1/2 + 3/8 + 1/2$
$= 1.875.$

(e) $P(Z > 1.875) = P(2) + P(3) + P(4) = .5.$

(f) The only way in which a girl cannot be obtained is BBBB, which has probability 1/16. Thus the probability of having a girl, by the complement rule, is $1 - 1/16 = 15/16 = .938$.

8.52 Let $0 - 4 \Leftrightarrow$ girl and $5 - 9 \Leftrightarrow$ boy. Beginning with line 130 in the random digit table:

690	51	64	81	7871	74	0	951	784
BBG	BG	BG	BG	BBBG	BG	G	BBG	BBG
3	2	2	2	4	2	1	3	3

53	4	0	64	89872	0	1	972	4	50	50
BG	G	G	BG	BBBBG	G	G	BBG	G	BG	BG
2	1	1	2	5	1	1	3	1	2	2

0	71	663	2	81
B	BG	BBG	G	BG
1	2	3	1	2

The average number of children is 52/25 = 2.08. This compares with the expected value of 1.875.

8.53 We will approximate the expected number of children, μ, by making the mean $\bar{x}$ of 25 randomly generated observations of X. We create a suitable string of random digits (say of length 100) by using the command `randInt (0, 9, 100)` $\rightarrow L_1$. Now we scroll down the list L_1. Let the digits 0 to 4 represent a boy and 5 to 9 represent a girl. We read digits in the string until we get a "5 to 9" (girl) or until four "0 to 4"s (boys) are read, whichever comes first. In each case, we record X = the number of digits in the string = the number of children. We continue until 25 X-values have been recorded. Our sample string L_1 yielded the following values of X:

245 /	8 /	06 /	37 /	9 /	6 /	6 /	6 /	2443 /	9 /	9 /	16 /	45 /
(3)	(1)	(2)	(2)	(1)	(1)	(1)	(1)	(4)	(1)	(1)	(2)	(2)

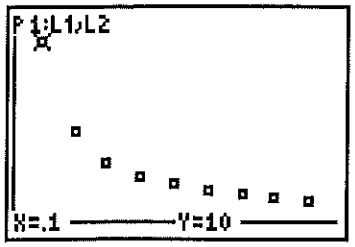
8 /	336 /	15 /	9 /	1331 /	38 /	8 /	48 /	37 /	119 /	5 /	8 /
(1)	(3)	(2)	(1)	(4)	(2)	(1)	(2)	(2)	(3)	(1)	(1)

This yields $\bar{x} = 45/25 = 1.8$, compared with the known mean $\mu = 1.875$.

8.54 (a) Recall that if p = probability of success, then $\mu = 1/p$. Then the table is as follows:

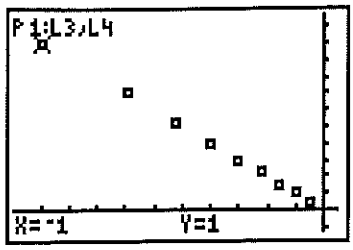
X	.1	.2	.3	.4	.5	.6	.7	.8	.9
Y	10	5	3.33	2.5	2	1.67	1.43	1.25	1.1

(b)

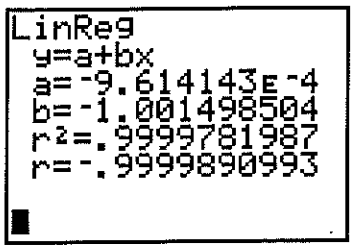


(c) Assuming the power function model $y = ab^x$, we transform the data by taking logs of both sides: $\log y = a + b \log x$. We thus compute and store the logs of the x's and y's.

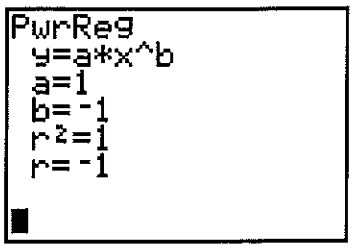
(d) Here is the plot of the transformed data log y vs. log x:

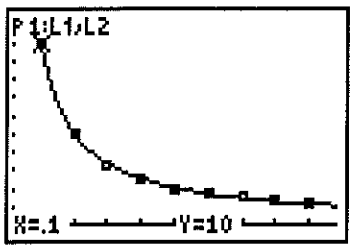


(e) The correlation is $r = -.999989$, or approximately -1.



(f) The equation of the power function is $\hat{y} = 1x^{-1} = 1/x$.





(g) The power function illustrates the fact that the mean of a geometric random variable is the *reciprocal* of the probability p of success: $\mu = 1/p$.

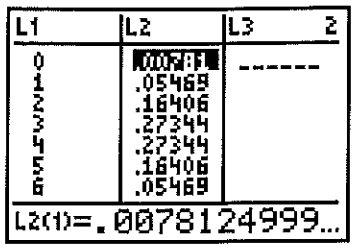
8.55 (a) No. It is not reasonable to assume that the opinions of a husband and wife (especially on such an issue as mothers working outside the home) are independent.

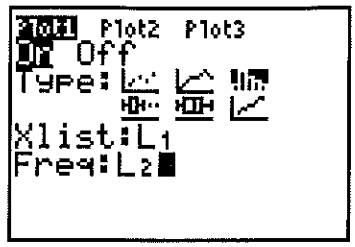
(b) No. The sample size (25) is so small compared to the population size (75) that the probability of success ("Yes") will substantially change as we move from person to person within the sample. The population should be at least 10 times larger than the sample in order for the binomial setting to be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

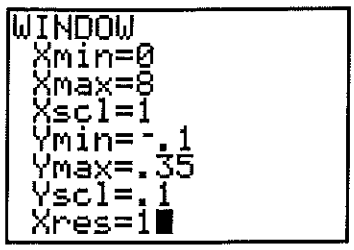
8.56 P(alcohol related fatality) $= 346/869 = .398$. If X = number of alcohol related fatalities, then X is $B(25, .398)$. $\mu = np = 25(.398) = 9.95$. $\sigma = \sqrt{((9.95)(.602))} = 2.45$. $P(X \leq 5) =$ `binomcdf(25, .398, 5)` $= .0307$.

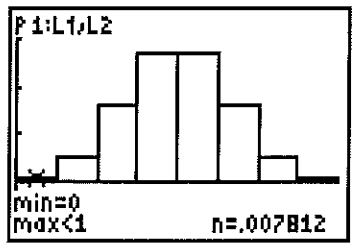
8.57 (a) The distribution of $X = B(7, .5)$ is symmetric; the shape depends on the value of the probability of success. Since .5 is halfway between 0 and 1, the histogram is symmetric.

(b) With the values of $X = 0, 1, \ldots, 7$ in L_1, define L_2 to be `binompdf (7, .5)`. Then the probability table for B(7, .5) is installed in L_1 and L_2. Here is a histogram of the p.d.f.:

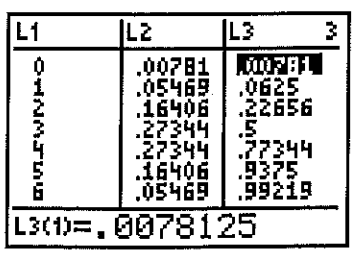


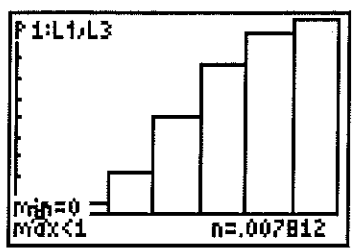






Now, define L_3 to be `binomcdf (7, .5, L1)`. Then the cdf for the B(7, .5) distribution is installed in L_3. Here is the histogram of the cdf:





(c) $P(X = 7) =$ `binompdf (7, .5, 7)` $= .0078125$.

8.58 (a) In our simulation, we obtained the following results:

Outcome	1	2	3	4	5	6	7	8
Frequency	28	17	2	2				1
Rel. freq.	.56	.34	.04	.04				.02

(b) We observed heads on the first toss 56% of the time. Our estimate of the probability of heads on the first toss is 0.5.

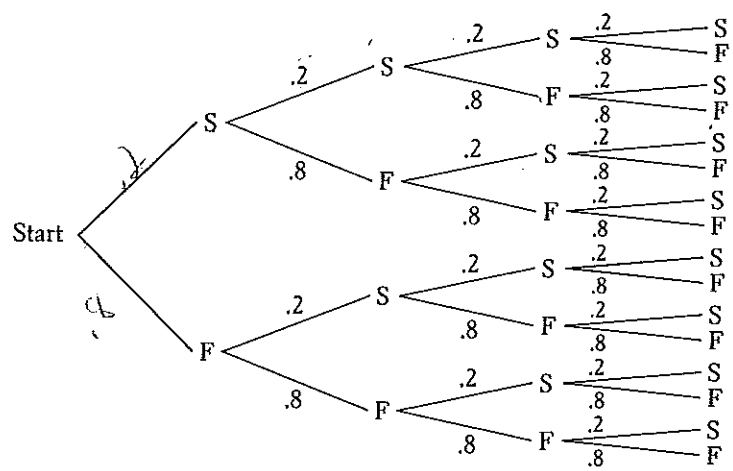
(c) An estimate of the probability that the first head appears on an odd-numbered toss is 2/3.

8.59 (a) $p = .2$.

(b)

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)



(d) SSFF, SFSF, SFFS, FSFS, FSSF, FFSS.

(e) Each outcome has probability .0256. The probabilities all involve the same number of each factor; they are simply multiplied in different orders, which will not affect the product.

8.60 Let X = the number of schools out of 20 that say they have a soft drink contract. X is binomial with $n = 20$ and $p = .62$.

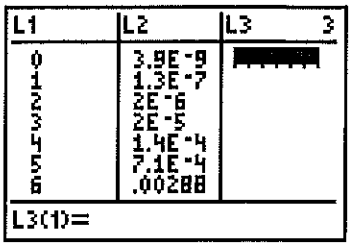
(a) $P(X = 8)$ = `binompdf (20, .62, 8)` = .0249.

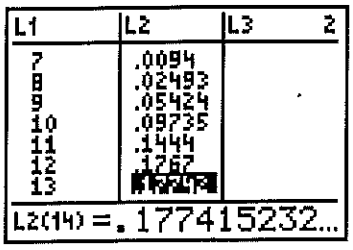
(b) $P(X \le 8)$ = `binomcdf (20, .62, 8)` = .0381.

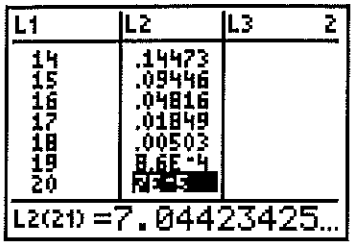
(c) $P(X \ge 4) = 1 - P(X \le 3) = 1 -$ `binomcdf (20, .62, 3)` = .99998.

(d) $P(4 \le X \le 12) = P(X \le 12) - P(X \le 3) =$ `binomcdf (20, .62, 12)` − `binomcdf (20, .62, 3)` = .5108 − .0381 = .4727.

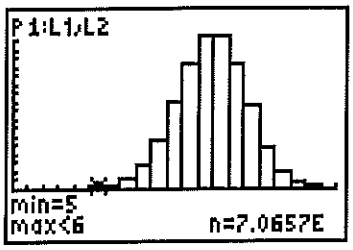
(e) X = the number of schools out of 20 that say they have a soft drink contract. Create the probability distribution table for X using the TI-83 by entering the values 0, 1, 2, ..., 20 into column L_1 and then entering the command `binompdf (20, .62, L1)` → L_2 to store the probabilities in L_2. The pdf table is displayed in the three screens below.







(f) The TI-83 version of the histogram is given below. (Note that some of the bars are too short to appear in the histogram.)



8.61 Let X = the number of southerners out of 20 that believe they have been healed by prayer. X is binomial with $n = 20$ and $p = .46$.

(a) $P(X = 10)$ = `binompdf (20, .46, 10)` = .1652.

(b) $P(10 < X < 15) = P(11 \le X \le 14) = P(X \le 14) - P(X \le 10) =$ `binomcdf (20, .46, 14)` − `binomcdf (20, .46, 10)` = .9917 − .7209 = .2708.

(c) $P(X > 15) = 1 - P(X \le 15) = 1 -$ `binomcdf (20, .46, 15)` = 1 − .9980 = .0020.

(d) $P(X < 8) = P(X \le 7) =$ `binomcdf (20, .46, 7)` = .2241.

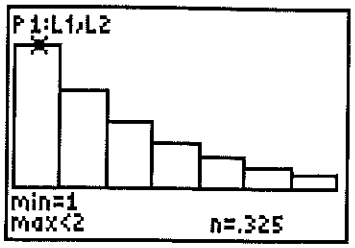
8.62 (a) $P(X = 2) = .2753$.

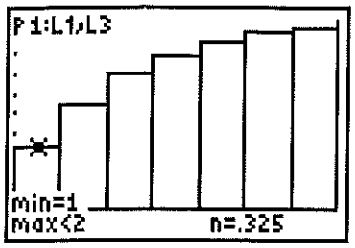
(b) $P(X \ge 2) = P(X = 2) + \ldots + P(X = 7) = .4234$.

(c) $P(X < 2) = P(X = 0) + P(X = 1) = .5767$.

(d) $P(2 \le X \le 5) = P(X = 2) + \ldots + P(X = 5) = .4230$.

8.63 X is geometric with $p = .325$. $1 - p = .675$. (a) $P(X = 1) = .325$. (b) $P(X \le 3) = .325 + (.675)(.325) + (.675)^2(.325) = .69245$. (c) $P(X > 4) = (.675)^4 = .208$. (d) The expected number of at-bats until Roberto gets his first hit is $\mu = 1/p = 1/.325 = 3.08$. (e) To do this, use the commands `seq (x, x, 1, 10)` → L_1, `geometpdf (.325, L1)` → L_2, and `geometcdf (.325, L1)` → L_3. (f) See the top of the facing page for the plot of the histogram.





8.64 (a) By the 68–95–99.7 rule, the probability of any one observation falling within the interval $\mu - \sigma$ to $\mu + \sigma$ is .68. Let X = the number of observations out of 5 that fall within this interval. Assuming that the observations are independent, X is $B(5, .68)$. Then, $P(X = 4) =$ `binompdf (5, .68, 4)` = .3421

(b) By the 68–95–99.7 rule, 95% of all observations fall within the interval $\mu - 2\sigma$ to $\mu + 2\sigma$. Thus, 2.5% (half of 5%) of all observations will fall above $\mu + 2\sigma$. Let X = the number of observations that must be taken before we observe one falling above $\mu + 2\sigma$. Then X is geometric with $p = .025$. $P(X = 4) = (1 - .025)^3(.025) = (.975)^3(.025) = .0232$.

8.65 $P(X \geq 1) = 1 - P(X = 0) = 1 - \binom{n}{0} p^0 (1 - p)^{n-0} = 1 - (1)(1)(1 - p)^n = 1 - (1 - p)^n$.