Chapter 14 – Binomial Distributions

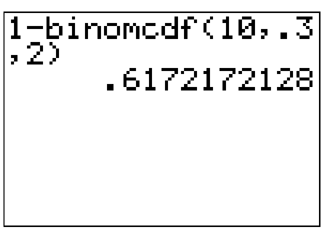
14.1 Binomial. (1) We have a fixed number of observations (n = 15). (2) It is reasonable to believe that each call is independent of the others. (3) Success means reaching a working cell phone number; failure is any other outcome. (4) Each randomly dialed number has chance p = 0.55 of reaching a working cell phone number.

14.2 Not binomial. We do not have a fixed number of observations.

14.3 Not binomial. The trials aren't independent. If one tile in a box is cracked, there are likely more tiles cracked (probably due to rough handling of the box).

14.4 The sample size n = 4000 is smaller than 5% of all Canadians aged 15 and over. The condition is satisfied for the count to approximately have the binomial distribution with sample size n = 4000 and success probability p = 0.137.

14.5 **(a)** C, the number caught, is binomial with n = 10 and p = 0.7. M, the number missed, is binomial with n = 10 and p = 0.3. **(b)** We find $P(M = 3) = \binom{10}{3}(0.3)^3(0.7)^7$ = 120(0.027)(0.08235) = 0.2668. With software, we find $P(M \geq 3)$ = 0.6172. Here, we used the fact that the event 3 or more is the complement of the event 2 or fewer with graphing calculators and most software. Output from a graphic calculator is provided.



14.6 (Let N be the number of working cell phone numbers contacted among the 15 calls observed. Then, N has the binomial distribution with n = 15 and p = 0.55.) **(a)** $P(N = 3) = \binom{15}{3}(0.55)^3(0.45)^{12} = 0.0052$. **(b)** $P(N \leq 3) = P(N = 0) + \cdots + P(N = 3) = 0.0063$. **(c)** $P(N \geq 3) = 1 - P(N \leq 2) = 1 - (P(N = 0) + \cdots + P(N = 2)) = 0.9989$. **(d)** $P(N < 3) = 1 - P(N \geq 3) = 1 - 0.9989 = 0.0011$. **(e)** $P(N > 3) = P(N \geq 3) - P(N = 3) = 0.9989 - 0.0052 = 0.9937$.

The figure shown illustrates the use of the TI-83/84 calculator's `binompdf` and `binomcdf` functions (found near the bottom of the `DISTR` menu) to compute the first two probabilities. The first of these finds individual binomial probabilities, and the second finds cumulative probabilities (that is, it sums the probability from 0 up to and including a given number). Excel offers similar features with its BINOM.DIST function. Calculators that do not have binomial probabilities may have a built-in function to compute the factorials involved in the binomial coefficient, for example,

$\binom{15}{3}$, which can then be multiplied by the appropriate probabilities.

```
binompdf(15,0.55,3)
     0.005219749
binomcdf(15,0.55,3)
     0.006326773
```

14.7 (a) `5 choose 2` returns 10. **(b)** `500 choose 2` returns 124,750, and `500 choose 100` returns 2.041694 × 10^{107}. **(c)** `(10 choose 1)*0.11*0.89^9` returns 0.38539204407.

14.8 (a) With n = 15 and p = 0.55, we have $\mu = np = (15)(0.55) = 8.25$ calls. **(b)** $\sigma = \sqrt{np(1-p)} = \sqrt{15(0.55)(0.45)} = 1.927$ calls. **(c)** When $p = 0.7$, $\sigma = 1.775$ calls; with $p = 0.8$, $\sigma = 1.549$ calls. As p approaches 1, the standard deviation decreases (that is, it approaches 0).

14.9 (a) X is binomial with n = 10 and p = 0.3; Y is binomial with n = 10 and p = 0.7. **(b)** The mean of Y is (10)(0.7) = 7 errors caught, and for X the mean is (10)(0.3) = 3 errors missed. **(c)** The standard deviation of Y (or X) is $\sigma = \sqrt{10(0.7)(0.3)}$ = 1.4491 errors.

14.10 Let X be the number of 1s and 2s; then X has a binomial distribution with n = 90 and p = 0.477 (in the absence of fraud). This should have mean $\mu = np = 42.93$ and standard deviation $\sigma = \sqrt{90(0.477)(1-0.477)}$ = 4.7384. Since np and $n(1-p)$ are at least 10, we can use the Normal approximation. Therefore, $P(X \leq 29)$ = $P(Z \leq 29 - 42.93 \;/\; 4.7384) = P(Z \leq -2.94)$ = 0.0016. (Using software, shown below, we find that the exact binomial probability is 0.0021.) Either way, this probability is quite small, so we have reason to be suspicious.

```
binomcdf(90,.477
,29)
      .0020818796
■
```

14.11 (a) μ = (1175)(0.37) = 434.75 and $\sigma = \sqrt{1175(0.37)(1-0.37)} = \sqrt{273.8925} = 16.550$ students. **(b)** We observe that $np = (1175)(0.37) = 434.75 \geq 10$ and $n(1-p) = (1175)(0.63) = 740.25 \geq 10$, so n is large enough for the Normal approximation to be reasonable. The college wants 450 students, so $P(X \geq 451)$ = $P(Z \geq 451 - 434.75 \;/\; 16.550) = P(Z \geq 0.98)$ = 0.1635. **(c)** The exact binomial probability is 0.1706 (obtained from software), so the Normal approximation is 0.0071 (not quite three-fourths of 1%) too low. For a better approximation, consider using the continuity correction described in Exercise 14.43. **(d)** To decrease the chance of more students than they want, the college needs to decrease the number admitted (this will decrease both μ and σ). Using technology, if $n = 1150$, the probability of more than 450 students is 0.0638. Similarly, with $n = 1145$, we have 0.0506; with $n = 1144$, we have 0.0482. They should admit at most 1144 students to have no more than a 5% chance of too many. (Answers will vary slightly if the Normal approximation is used to compute the probabilities.)

14.12 (a) μ = (1500)(0.22) = 330 and $\sigma = \sqrt{1500(0.22)(1 - 0.22)} = 16.044$ first-generation Canadians. **(b)** To check whether the Normal approximation can be applied, note that np = 330 and $n(1 - p)$ = 1170 are both at least 10. We compute $P(340 \leq X \leq 390) = P(340 - 330/16.044 \leq Z \leq 390 - 330/16.044) = P(0.62 \leq Z \leq 3.74) = 0.2675$.

14.13 (b) binomial with n = 3 and p = 1/4. Larry has three independent eggs, each with probability 1/4 of containing salmonella.

14.14 (b) 0.58. $P(S \geq 1) = P(S > 0) = 1 - P(S = 0) = 1 - 0.4219 = 0.5781$.

14.15 (c) not binomial. The selections are not independent; once we choose one student, it changes the probability that the next student is a business major.

14.16 (c) $\binom{4}{2} = 6$. We must choose two of the four guesses to be correct; $\binom{4}{2} = 6$. [Note that option (b) is only wrong for its computation.]

14.17 (c) 0.035. This probability is $(0.25)(0.75)^2(0.25) = 0.0352$.

14.18 (b) 0.211. Guessing correctly twice means missing twice, so this probability is $\binom{4}{2}(0.25)^2(0.75)^2 = 0.2109$.

14.19 (b) 0.2. The numbers 8 and 9 are two of the ten possible digits, so the probability is 0.20.

14.20 (a) binomial with n = 80 and p = 0.2. Two lines of the table have 2(40) = 80 independent digits. The number of successes (8 or 9) is binomial with n = 80 and p = 0.20.

14.21 (a) 16. The mean is np = (80)(0.20) = 16.

14.22 (a) No. There are more than two possible outcomes (the number of defects could be 0, 1, 2, etc.). **(b)** Yes. A binomial distribution is reasonable here; a large city will have a population much larger than 100 (the sample size), and each randomly selected juror has the same (unknown) probability p of opposing the death penalty (supposing they are asked so that one person's opinion will not influence another). However, if the jurors are asked the question as a whole, for example, one person's answer could influence another and independence would be lost. **(c)** Yes. In a Pick 3 game, Joe's chance of winning the lottery is the same every week and each week's drawing is independent of the last, so assuming that a year consists of 52 weeks (observations), this would be binomial.

14.23 (a) A binomial distribution is not an appropriate choice for field goals made because, given the different situations the kicker faces (wind, distance, etc.), the probability of success is likely to change from one attempt to another. **(b)** It would be reasonable to use a binomial distribution for free throws made because each is from the same position with respect to the basket with no interference allowed for the shot and, presumably, each is independent of any others.

14.24 (a) The distribution is binomial with $n = 8$ and $p = 0.30$. **(b)** $\mu = (8)(0.3) = 2.4$ and $\sigma = \sqrt{8(0.3)(0.7)} = 1.2961$. **(c)** $P(X = 1) = \binom{8}{1}(0.3)^1(0.7)^7 = 0.1977$ and $P(X \geq 1) = 1 - P(X = 0) = 1 - \binom{8}{0}(0.3)^0(0.7)^8 = 0.9424$.

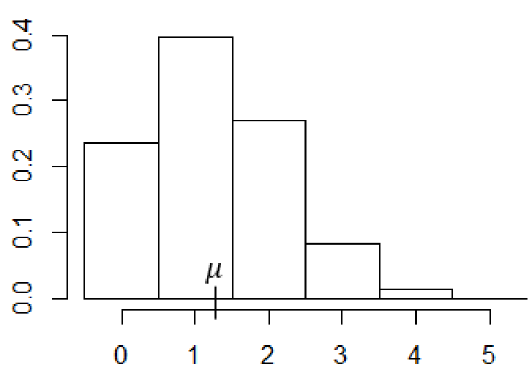
14.25 $n = 5$ and $p = 0.25$. **(b)** The possible values of X are the integers 0, 1, 2, 3, 4, and 5. **(c)** The probabilities are computed:

$P(X = 0) = \binom{5}{0}(0.25)^0(0.75)^5 = 0.2373$ $\quad$ $P(X = 1) = \binom{5}{1}(0.25)^1(0.75)^4 = 0.3955$

$P(X = 2) = \binom{5}{2}(0.25)^2(0.75)^3 = 0.2637$ $\quad$ $P(X = 3) = \binom{5}{3}(0.25)^3(0.75)^2 = 0.0879$

$P(X = 4) = \binom{5}{4}(0.25)^4(0.75)^1 = 0.0146$ $\quad$ $P(X = 5) = \binom{5}{5}(0.25)^5(0.75)^0 = 0.00098$

(d) $\mu = np = (5)(0.25) = 1.25$ years and $\sigma = \sqrt{5(0.25)(0.75)} = 0.9682$ year. The mean μ is indicated on the probability histogram.



14.26 (a) This probability is 18/38 = 0.47368. **(b)** X has the binomial ($n = 4$, $p = 0.47368$) distribution. **(c)** $P(\text{break even}) = P(X = 2) = \binom{4}{2}(0.47368)^2(1 - 0.47368)^2 = 0.37292$. **(d)** $P(\text{lose money}) = P(X < 2) = P(X = 0) + P(X = 1) = \binom{4}{0}(0.47368)^0(1 - 0.47368)^4 + \binom{4}{1}(0.47368)^1(1 - 0.47368)^3 = 0.07674 + 0.27625 = 0.35299$.

14.27 (a) Provided none of the children are related, all children should be independent in terms of immune response. There is a fixed number of children to be observed, and (we assume) each has the same probability of having an immune response. **(b)** Let X be the number who developed an immune response. Observe that at least one child not developing an immune response is equivalent to no more than 19 developing an immune response, or $X \leq 19$. For the aerosolized vaccine, $P(X \leq 19) = 1 - P(X = 20) = 1 - \binom{20}{20}(0.85)^{20}(0.15)^0 = 0.9612$. For the subcutaneous injection, the probability is $1 - \binom{20}{20}(0.95)^{20}(0.05)^0 = 0.6415$.

14.28 X, the number of wins betting on red 200 times, is binomial with n = 200 and p = 0.47368 (using the information from Exercise 14.26). The Normal approximation is quite safe: np = 94.736 ≥ 10 and $n(1 - p)$ = 105.264 ≥ 10. The mean is $\mu = np$ = 94.736 and the standard deviation is $\sigma = \sqrt{200(0.47368)(1 - 0.47368)}$= 7.06126, so $P(X < 100) = P(X \leq 99) = P(Z \leq 99 - 94.736\ /\ 7.06126)$= $P(Z \leq 0.60)$ = 0.7257. The exact binomial probability is 0.7502. As the number of plays (n) increases, the probability of losing money will increase. For example, if n = 400, $P(X < 200)$ = standardize to $P(Z \leq 0.95)$ = 0.8289. The exact binomial probability is 0.8424.

14.29 (a) $n = 100$ and $p = 0.85$. Observe that $np = 85 \geq 10$ and $n(1 - p) = 15 \geq 10$, so the Normal approximation can be applied. Also, $\mu = 100(0.85) = 85$ and $\sigma = \sqrt{100(0.85)(0.15)} = 3.5707$ reactions. $P(X \geq 90) = P(Z \geq 90 - 85/3.5707)P(Z \geq 1.40) = 0.0808$. Using software, we get the exact probability to be 0.0994, which is 0.0186 larger than the approximate probability. **(b)** If the subcutaneous injection is used, then $p = 0.95$ and $n(1 - p) = 100(0.05) = 5$, which is not at least 10.

14.30 (a) Let X denote the number of heads. Then X is binomial with $n = 10$ and $p = 0.5$. The distribution of X is given in the table below. Collecting $160 or more means at least 8 heads were flipped. This event has probability 0.0439 + 0.0098 + 0.00098 = 0.05468.

Heads	0	1	2	3	4	5
Probability	0.00098	0.0098	0.0439	0.1172	0.2051	0.2461
Heads	6	7	8	9	10	
Probability	0.2051	0.1172	0.0439	0.0098	0.00098	

(b) From the distribution in part (a), $P(X \geq 6) = 0.2051 + 0.1172 + \cdots + 0.00098 = 0.377$. In the control group, the proportion who tossed 6 or more heads is $\frac{14 + 6 + 2 + 1 + 2}{67} = 0.3731$. In the treatment group, the proportion was $\frac{15 + 7 + 6 + 0 + 5}{61} = 0.541$. **(c)** The proportion of the treatment group that reported tossing 6 or more heads is much higher than it would be if they were telling the truth; whereas the proportion for the control group was very similar to what it would be if they were telling the truth. This implies cheating may have occurred in the treatment group.

14.31 (a) If R is the number of red-blossomed plants out of a sample of 4, then $P(R =$

3) = $\binom{4}{3}(0.75)^3(0.25)^1$ = 0.4219, using a binomial distribution with n = 4 and p = 0.75. **(b)** With n = 60, the mean number of red-blossomed plants is np = (60)(0.75) = 45. **(c)** If R is the number of red-blossomed plants out of a sample of 60, then $P(R \geq 45) = P(Z \geq 0)$ = 0.5000. (Software gives 0.5688 using the binomial distribution.)

14.32 (a) X, the number of positive tests, is binomial with n = 1000 and p = 0.004, assuming the individuals being tested are random (and not sexual partners, for example). **(b)** $\mu = np$ = (1000)(0.004) = 4 positive tests. **(c)** To use the Normal approximation, we need np and $n(1 - p)$ to both be at least 10; as we saw in part (b), np = 4.

14.33 (a) Of the 761,710 total vehicles sold, Elantras accounted for a proportion of 241706/761710 = 0.3173. **(b)** If E is the number of Elantra buyers in the 1000 surveyed buyers, then E has the binomial distribution with n = 1000, and p = 0.3173. $\mu = np$ = (1000)(0.3173) = 317.3 and $\sigma = \sqrt{100(0.3173)(1-0.3173)} = 14.718$ Elantra buyers. **(c)** $P(E < 300) = P(E \leq 299) = P(Z \leq 299 - 317.3/14.718) = P(Z \leq -1.24) = 0.1075$.

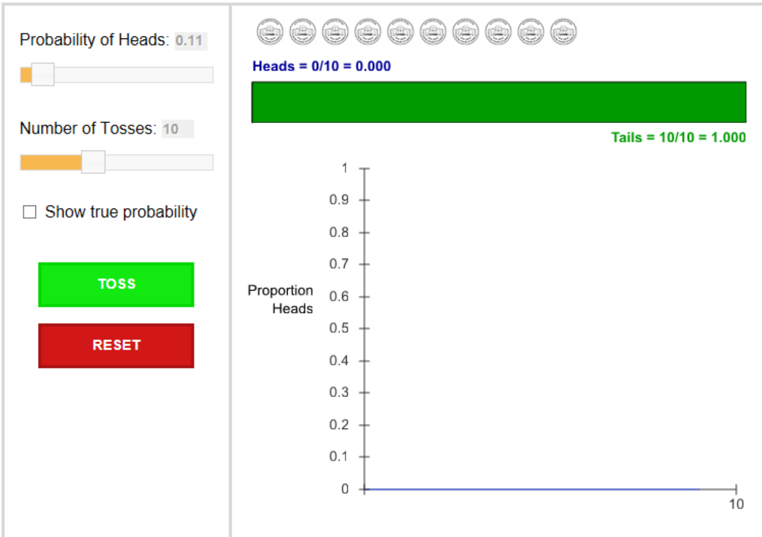
14.34 (a) If each subject selects the socks at random, the probability they would choose them from the center position is 1/5 or 0.2. Let X be the number who choose their socks from the center position. Then X has the binomial distribution with $n = 100$ and $p = \frac{1}{5} = 0.2$. **(b)** $\mu = 100(0.2) = 20$ and $\sigma = \sqrt{100(0.2)(0.8)} = 4$ subjects. **(c)** Observe that $np = 100(0.2) = 20 \geq 10$ and $n(1-p) = 80 \geq 10$. $P(X \geq 34) = P(Z \geq 34 - 20/4) = P(Z \geq 3.5) = 0.0002$. Using software, the exact probability is 0.0007, which is 0.0005 larger than the approximate probability. **(d)** The experiment does support the center stage effect. If participants were truly picking the socks at random, it would be highly unlikely for 34 or more to choose the center pair.

14.35 (a) With n = 100, the mean and standard deviation are μ = 75 and σ = 4.3301 questions, so $P(70 \leq X \leq 80) = P(-1.15 \leq Z \leq 1.15)$ = 0.7498 (software gives 0.7518). **(b)** With n = 250, we have μ = 187.5 and σ = 6.8465 questions, and a score between 70% and 80% means 175 to 200 correct answers, so $P(175 \leq X \leq 200) = P(-1.83 \leq Z \leq 1.83)$ = 0.9328 (software gives 0.9428).

Note: *If one used the more mathematical idea that "between" does not include the endpoints of the interval, we would have* $P(70 < X < 80) = P(71 \leq X \leq 79) = 0.6424$ *for the 100-question test and* $P(175 < X < 200) = 0.9070$ *for the 250-question test.*

14.36 We have μ = 5000 and $\sigma = \sqrt{10{,}000(0.5)(1-0.5)} = 50$ heads, so using the Normal approximation, we compute $P(X \geq 5067$ or $X \leq 4933) = 2P(Z \geq 1.34)$ = 0.1802. If Kerrich's coin were fair, we would see results at least as far from 5000 as what he observed in about 18% of all repetitions of the experiment of flipping the coin 10,000 times. This is not unreasonable behavior for a fair coin.

14.37 (a) Answers will vary, but over 99.8% of samples should have 0 to 4 bad tomatoes. The result of one such sample by the applet is shown.



(b) Each time we choose a sample of size 10, the probability that we have exactly 1 bad tomato is 0.3854; therefore, out of 20 samples, the number of times that we have exactly 1 bad tomato has a binomial distribution with parameters n = 20 and p = 0.3854. This means that most students—99.8% of them—will find that between 2 and 14 of their 20 samples have exactly 1 bad tomato, giving a proportion between 0.10 and 0.70. (If anyone has an answer outside of that range, it would be significant evidence that he or she did the exercise incorrectly.)

14.38 (The number N of new infections is binomial with n = 20 and p = 0.80 [for unvaccinated children] or 0.05 [for vaccinated children].) **(a)** For vaccinated children, the mean is (20)(0.05) = 1 new infection, and $P(N \leq 2) = 0.9245$. **(b)** For unvaccinated children, the mean is (20)(0.80) = 16 new infections, and $P(N \geq 18)$ = 0.2061.

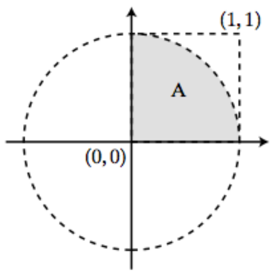
14.39 The number N of infections among untreated BJU students (assuming independence) is binomial with n = 1400 and p = 0.80, so the mean is 1120 and the standard deviation is 14.9666 students. In addition, 75% of that group is 1050, and the Normal approximation is safe because (1400)(0.80) = 1120 and (1400)(0.20) = 280 are both at least 10. $P(N \geq 1050) = P(Z \geq 1050 - 1120/14.9666) = P(Z \geq -4.68)$, which is very near to 1. (Exact binomial computation gives 0.999998.)

14.40 (Let V and U be, respectively, the number of new infections among the vaccinated and unvaccinated children.) **(a)** V is binomial with n = 17 and p = 0.05, with mean 0.85 infection. **(b)** U is binomial with n = 3 and p = 0.80, with mean 2.4 infections. **(c)** The overall mean is 3.25 infections (2.4 + 0.85).

14.41 (Define V and U as in Exercise 14.40.) **(a)** $P(V = 1) = 0.3741$ and $P(U = 1) = 0.0960$. Because these events are (assumed) independent, $P(V = 1$ and $U =$

1) = $P(V = 1)P(U = 1)$ = (0.3741)(0.0960) = 0.0359. **(b)** Considering all the possible ways to have a total of two infections, we have P(2 infections) = $P(V = 0$ and $U = 2)$ + $P(V = 1$ and $U = 1)$ + $P(V = 2$ and $U = 0)$ = $P(V = 0)P(U = 2) + P(V = 1)P(U = 1) + P(V = 2)P(U = 0)$ = (0.4181)(0.3840) + (0.3741)(0.0960) + (0.1575)(0.0080) = 0.1977.

14.42 (a) and **(b)** The unit square and circle are shown; the intersection A is shaded.



(c) The circle has area π, and A is a quarter of the circle, so the area of A is $\pi/4$. This is the probability that a randomly selected point (X, Y) falls in A, so T is binomial with n = 2000 and $p = \pi/4$ = 0.7854. **(d)** The mean of T is $np = 2000(\pi/4) = 500\pi$ = 1570.7963, and the standard deviation is $\sqrt{np(1-p)} = \sqrt{2000\frac{\pi}{4}(1-\pi/4)}$ = 18.3602. **(e)** Because the mean of T is 500π, $T/500$ is an estimate of π.

14.43 (a) $np = 40(0.3) = 12 \geq 10$ and $n(1-p) = 40(0.7) = 28 \geq 10$. **(b)** Observe that $\mu = 12$ and $\sigma = \sqrt{40(0.3)(0.7)} = 2.898$ cases. Using the Normal approximation, $P(X \geq 15) = P(Z \geq 15 - 12/2.898) = P(Z \geq 1.04) = 0.1492$. **(c)** Using the continuity correction, $P(X \geq 15) = P(X \geq 14.5) = P(Z \geq 14.5 - 12/2.898) = P(Z \geq 0.86) = 0.1949$.