Section 6.1: Summarizing Possible Outcomes and Their Probabilities

6.1 Rolling dice

a) Uniform distribution; P(1) = P(2) = P(3) = P(4) = P(5) = P(6) = 1/6

b) The probabilities below correspond to the stems on the graph in this exercise. Each probability is calculated by counting how many rolls of the dice add up to a particular number. For example, there are three rolls that add up to four (1,3; 2,2; 3,1); thus, the probability of four is 3/36 = 0.083.

x	P(x)	x	P(x)
2	1/36	8	5/36
3	2/36	9	4/36
4	3/36	10	3/36
5	4/36	11	2/36
6	5/36	12	1/36
7	6/36		

c) The probabilities in (b) satisfy the two conditions for a probability distribution: $0 \le P(x) \le 1$ for each x and $\sum P(x) = 1$.

6.2 Dental Insurance

a) The value of the payout is determined by a random phenomenon, depending on whether you need major, minor, or no dental repair over the next 5 years.

b) X is discrete, with possible values \$0, \$100, or \$1000.

c) P(\$0) = 0.35, P(\$100) = 0.60, and P(\$1000) = 0.05

6.3 San Francisco Giants hitting

a) The probabilities give a legitimate probability distribution because each one is between 0 and 1 and the sum of all of them is 1.

b) $\mu = 0P(0) + 1P(1) + 2P(2) + 3P(3) + 4P(4) = 0(0.7429) + 1(0.1704) + 2(0.0517) + 3(0.0055) + 4(0.0295) = 0.4083$; The expected number of bases for a random time at bat for a San Francisco Giants player is 0.4083.

c) The mean is the expected value of X; that is, what we expect for the average in a long run of observations. Since the mean is a long term average of values, it doesn't have to be one of the possible values for the random variable.

6.4 Basketball shots

Let A (respectively B) denotes the event that competitor A (respectively B) has a successful attempt on a three-points shot in a round.

a) The series ends within the second round if we get (AB at the first round and $A\bar{B}$ at the second round) or (AB at the first round and $\bar{A}B$ at the second round) or ($\bar{A}\bar{B}$ at the first round and $A\bar{B}$ at the second round) or ($\bar{A}\bar{B}$ at the first round and $\bar{A}B$ at the second round). These events occur with probabilities 0.0144, 0.0084, 0.1344 and 0.0784, respectively. The probability that the series ends within the second round is 0.0144 + 0.0084 + 0.1344 + 0.0784 = 0.2356.

b) X = 1 or 2 or 3, with P(1) = 0.3 × 0.8 + 0.7 × 0.2 = 0.38, P(2) = 0.2356 and P(3) = 1 – 0.38 – 0.2356 = 0.3844.

c) μ = 1 × 0.38 + 2 × 0.2356 + 3 × 0.3844 = 2.0044.

6.5 WhatsApp reviews

a)

Rate	Probability
5	0.71
4	0.15
3	0.06
2	0.03
1	0.05

b) μ = 1(0.05) + 2(0.03) + 3(0.06) + 4(0.15) + 5(0.71) = 4.43. The average rating is 4.43 stars.

6.6 Selling houses

a) The probabilities are not the same for each outcome; that is, each *x* value (0, 1, 2, 3 and 4) does not carry the same weight.

b) $\mu = 0(0.68) + 1(0.19) + 2(0.09) + 3(0.03) + 4(0.01) = 0.5$; We expect the agent to sell 0.5 houses per month.

6.7 Playing the lottery

a) Your 3-digit number only matches one of the 1000 options, so P($500) = 1/1000 = 0.001.

b)

x	P(*x*)
0	0.999
500	0.001

c) $\mu = 0(0.999) + 500(0.001) = 0.5$; Your expected return on a $1.00 bet is $0.50.

d) For Pick 4, your expected return is $\mu = 0(0.9999) + 5000(0.0001) = 0.5$ or $0.50. Since both games have the same expected return, we are indifferent as to which we play.

6.8 Roulette

a) Answers will vary.

b) Expected profit for wager (1): $\mu = 350\left(\frac{1}{38}\right) + (-10)\left(\frac{37}{38}\right) = -\$0.53$

Expected profit for wager (2): $\mu = 10\left(\frac{18}{38}\right) + (-10)\left(\frac{20}{38}\right) = -\$0.53$

In this sense, the two wagers are the same and both have a negative expected return.

6.9 More Roulette

The wager on 23 will have the higher standard deviation since the winnings are, on average, further from the mean for this bet.

6.10 Ideal number of children

a) The mean for females is $\mu = 0(0.01) + 1(0.03) + 2(0.55) + 3(0.31) + 4(0.11) = 2.50$. The mean for males is $\mu = 0(0.02) + 1(0.03) + 2(0.60) + 3(0.28) + 4(0.08) = 2.39$. The means for the two distributions are quite similar.

b) Although the means are similar, the responses for males tend to be slightly closer to the mean than the responses for females. Thus, males seem to hold slightly more consistent views than females about ideal family size.

6.11 Profit and the weather

a)

x	P(*x*)
$80,000	0.70
$50,000	0.20
$20,000	0.10

b) $\text{P}(X \geq \$50{,}000) = \text{P}(X = \$50{,}000 \text{ or } X = \$20{,}000) = 0.20 + 0.10 = 0.30$

c) $\mu = 80{,}000(0.70) + 50{,}000(0.20) + 20{,}000(0.10) = 68{,}000$; The wheat farmer's expected profit is $68,000.

d)

x	P(*x*)
$77,000	0.70
$47,000	0.20
$37,000	0.10

6.11 (continued)

$\mu = 77,000(0.70) + 47,000(0.20) + 37,000(0.10) = 67,000$; With the insurance policy, the farmer's expected profit is \$67,000. The insurance policy actually results in a lower expected profit for the farmer.

6.12 Buying on eBay

a) and b)

Sample Space	Probability
WW	(0.1)(0.2) = 0.02
WL	(0.1)(0.8) = 0.08
LW	(0.9)(0.2) = 0.18
LL	(0.9)(0.8) = 0.72

c)

x	P(*x*)
\$50	0.02
\$30	0.08
\$20	0.18
\$0	0.72

d) $\mu = 50(0.02) + 30(0.08) + 20(0.18) + 0(0.72) = 7$, or \$7

6.13 Selling at the right price

a) μ = 150(0.45) + 250(0.15) + 350(0.40) = 245. The expected selling price for the sale of a car insurance plan is \$245.

x	P(*x*)
\$150	0.45
\$250	0.15
\$350	0.40

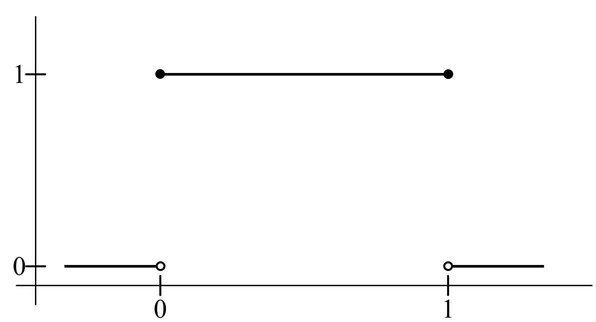
b) μ = 170(0.15) + 250(0.40) + 310(0.45) = 265. The expected selling price for the sale of a car insurance plan under the new pricing strategy is \$265.

x	P(*x*)
\$170	0.15
\$250	0.40
\$310	0.45

c) By comparing the means in part (a) and (b) the company can prospect higher profits with second pricing set in the long run.

6.14 Uniform distribution

a)



b) The mean of this distribution is 0.50.

c) There is a 0.40 probability that this random variable falls between 0.35 and 0.75.

d) There is a 0.80 probability that this random variable falls below 0.80.

6.15 TV watching

a) TV watching is, in theory, a continuous random variable because someone could watch exactly one hour of TV or 1.835 hours or 2.07 hours of TV.
b) Histograms were used because TV watching was measured to the nearest integer. The histograms would display the frequencies of the rounded (to the nearest integer) values obtained in the sample.
c) The two smooth curves would represent the approximate (based on the histograms) shape of the probability distribution for TV watching in the population if we could measure it in a continuous manner. Then, the area under the curve above an interval would represent the probability of an observation falling in that interval.

Section 6.2: Probabilities for Bell-Shaped Distributions

6.16 Probabilities in tails

a) The probability that an observation is at least one standard deviation above the mean is 0.159.
b) The probability that an observation is at least one standard deviation below the mean is 0.159.
c) The probability that an observation is within one standard deviation of the mean is $1 - 2(0.159) = 0.682$.

6.17 Probability in graph

a) The observation would fall 0.67 standard deviations above the mean, and thus, would have a *z*-score of 0.67. Looking up this *z*-score in Table A, we see that this corresponds to a cumulative probability of 0.749. The probability that an observation falls above this point (in the shaded region) is $1 - 0.749 = 0.251$.
b) The observation would fall between 0.50 standard deviations below the mean and 0.50 standard deviations above the mean. Looking up the *z*-score of 0.50 in Table A, we see that this corresponds to a cumulative probability of 0.6915. The probability that an observation falls above this point is $1 - 0.6915 = 0.3085$. Because of the symmetry of the normal distribution, the probability of falling below the opposite of this *z*-score is also 0.3085. This means that the probability of falling more than 0.5 standard deviations away from the mean is $2(0.3085) = 0.6170$. Therefore, the probability of falling within 0.5 standard deviations of the mean is $1 - 0.6170 = 0.383$.

6.18 Empirical rule

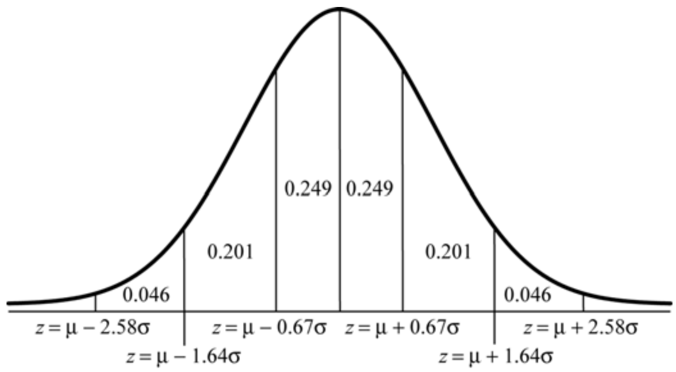
a) The cumulative probability for one standard deviation above the mean (*z*-score of 1) is 0.8413, and for one standard deviation below the mean (*z*-score of –1) is 0.1587. The difference, $0.8413 - 0.1587 = 0.6826$ (rounds to 0.68), the probability of a normally distributed random variable falling within 1 standard deviation of the mean on either side.
b) Similarly, we can look up a *z*-score of 2 to find that its cumulative probability is 0.9772. The cumulative probability for –2 is 0.0228. $0.9772 - 0.0228 = 0.9544$ (rounds to 0.95).
c) Finally, we can do the same for *z*-scores of 3 and –3, to get cumulative probabilities of 0.9987 and 0.0013, respectively. $0.9987 - 0.0013 = 0.9974$, or roughly 1.00.

6.19 Central probabilities

a) If we look up 1.64 on Table A, we see that the cumulative probability is 0.9495. The cumulative probability is 0.0505 for –1.64. $0.9495 - 0.0505 = 0.899$, which rounds to 0.90.
b) Using similar logic, we see 0.9951 and 0.0049 on the table for 2.58 and –2.58, respectively. $0.9951 - 0.0049 = 0.9902$, which rounds to 0.99.
c) Finally, we can use the same logic for *z*-scores of 0.67 and –0.67. We find cumulative probabilities of 0.7486 and 0.2514. The difference between these is $0.7486 - 0.2514 = 0.4972$, which rounds to 0.50.

6.19 (continued)

d)



6.20 *z*-score for given probability in tails

a) First we calculate the cumulative probability for the total probability of 0.04 in the tails. We divide 0.04 by two to determine the probability in each tail, 0.02. Then subtract the probability in the top tail from 1 to get 0.98. We can look up the closest cumulative probabilities to 0.02 and 0.98 in Table A to get *z*-scores of –2.054 and 2.054.

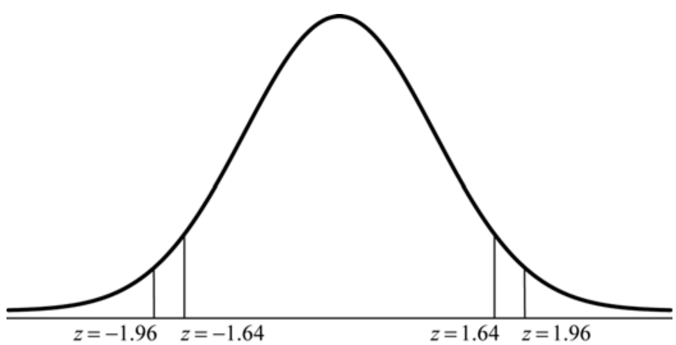
b) The probability more than 2.054 standard deviations below the mean equals 0.02 both because it is half of 0.04 and the probability that it would fall beyond this *z*-score is the same in either direction.

c) 2.054 standard deviations below the mean is the 2nd percentile because only 2% (0.02) of the population falls 2.054 standard deviations below the mean.

6.21 Probability in tails for given *z*-score

a) We have to divide this probability by two to find the amount in each tail, 0.01/2 = 0.005. We then subtract this from 1.0 to determine the cumulative probability associated with this *z*-score, 0.995. We can look up this probability on Table A to find the *z*-score of 2.58.

b) For both (a) and (b), we divide the probability in half, subtract from one, and look it up on Table A
(a) 1.96
(b) 1.645



6.22 *z*-score for right-tail probability

a) 0.20 is associated with a cumulative probability of 1 – 0.20 = 0.80. When we look up 0.80 in Table A, we find a *z*-score of 0.84.

b) (i) 0.05 is associated with a cumulative probability of 1 – 0.05 = 0.95. When we look up 0.95 in Table A, we find a *z*-score of 1.645.

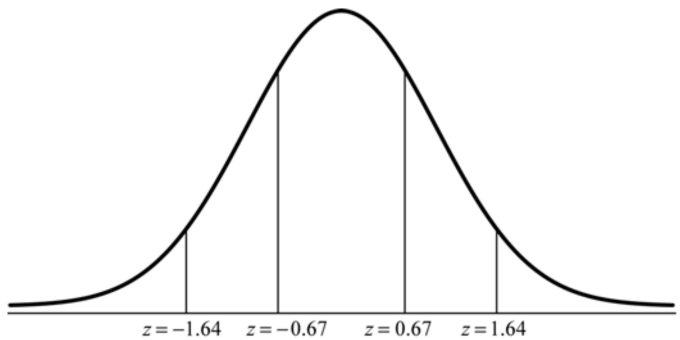
(ii) 0.05 is associated with a cumulative probability of 1 – 0.05 = 0.995. When we look up 0.995 in Table A, we find a *z*-score of 2.58.

6.23 *z*-score and central probability

a) The middle 50% means that 25% of the scores fall between the mean and each z-score (negative and positive). This corresponds to a cumulative probability of 0.75. When looked up in Table A, we find a z-score of 0.67.

b) Using the same logic as in (a), we find a cumulative probability of 0.95, and a z-score of 1.645.

c)



6.24 U.S. Air Force

For the lower bound, we have $z = \frac{x-\mu}{\sigma} = \frac{18-38}{22.67} = -0.88$, which corresponds to a cumulative probability of 0.19.

For the upper bound, we have $z = \frac{x-\mu}{\sigma} = \frac{34-38}{22.67} = -0.18$, which corresponds to a cumulative probability of 0.43.

Thus, 0.43 – 0.19 = 0.24 is the proportion of U.S. citizens eligible to be an officer in the U.S. Air Force with respect to age and 1 – 0.24 = 0.76 is the proportion of U.S. citizens who are not.

6.25 Blood pressure

a) $z = \frac{x-\mu}{\sigma} = \frac{140-121}{16} = 1.19$

b) A z-score of 1.19 corresponds, from Table A, to a cumulative probability of 0.8830. The amount above this z-score would be 1 – 0.88 = 0.12.

c) $z = \frac{x-\mu}{\sigma} = \frac{100-121}{16} = -1.31$; which corresponds to a cumulative probability of 0.0951. If we subtract 0.0951 from 0.8830 (the cumulative probability for 140), we get 0.79 as the probability between 100 and 140.

d) The 90th percentile is associated with a cumulative probability of 0.90. When we look up 0.90 in Table A, we find a z-score of 1.28. Using $x = \mu + z\sigma$, the 90th percentile is 121 + 1.28(16) = 141.5.

6.26 Coffee Machine

a) $z = \frac{x-\mu}{\sigma} = \frac{12-13}{0.6} = -1.67$; which corresponds to a cumulative probability of 0.0475. Therefore, 4.75% of cups will be filled with less than 12 ounces of coffee.

b) $z = \frac{x-\mu}{\sigma} = \frac{12.5-13}{0.6} = -0.83$; which corresponds to a cumulative probability of 0.2033. Therefore, 1 – 0.2033 = 0.7967, or about 80% of cups will be filled with more than 12.5 ounces of coffee.

c) Since half the cups will be filled with 13 or less ounces of coffee, the percentage of cups with between 12 and 13 ounces of coffee would be 50% – 4.8% = 45.2%.

6.27 Lifespan of phone batteries

a) According to the assumption of normality and the usual range of the number of charge and discharge cycles, the suggested mean is $\frac{300+500}{2} = 400$ and the standard deviation is approximately $\frac{500-400}{3} = 33.33$.

b) The 95th percentile corresponds to a z-score of 1.645. Therefore, the 95th percentile is $\mu + 1.645 \times \sigma =$ $400 + 1.645 \times 33.33 = 454.83$.

6.28 Birth weight for boys

a) $z = \frac{x-\mu}{\sigma} = \frac{2.5-3.41}{0.55} = -1.65$; which corresponds to a cumulative probability of 0.049. Therefore, the proportion of baby boys born with low birth weight is 0.49.

b) $z = \frac{x-\mu}{\sigma} = \frac{1.5-3.41}{0.55} = -3.47$

c) $z = \frac{x-\mu}{\sigma} = \frac{4.0-3.41}{0.55} = 1.07$; which corresponds to a cumulative probability of 0.858. Thus, the probability that a baby boy is born with a weight between 2.5 kg and 4.0 kg is 0.858 – 0.049 = 0.809, or about 80.9%.

d) $z = \frac{x-\mu}{\sigma} = \frac{3.6-3.41}{0.55} = 0.345$; which corresponds to a cumulative probability of 0.635. Therefore, Matteo falls at the 63.5th percentile.

e) The 96th percentile is associated with a cumulative probability of 0.96, which corresponds to a *z*-score of 1.75. Using $x = \mu + z\sigma$, Max weighs 3.41 + 1.75(0.55) = 4.37 kilograms.

6.29 MDI

a) (i) $z = \frac{x-\mu}{\sigma} = \frac{120-100}{16} = 1.25$; which corresponds to a cumulative probability of 0.894. This indicates that the proportion of children with MDI of at least 120 is 1 – 0.894 = 0.106.

(ii) $z = \frac{x-\mu}{\sigma} = \frac{80-100}{16} = -1.25$; which corresponds to a cumulative probability of 0.106. This indicates that the proportion of children with MDI of at least 80 is 1 – 0.106 = 0.894.

b) The 99th percentile for MDI is 137.3. The *z*-score corresponding to the 99th percentile is 2.33. To find the value of *x*, we calculate $x = \mu + z\sigma = 100 + (2.33)(16) = 137.3$.

c) The 1st percentile for MDI is 62.7. The *z*-score corresponding to the 1st percentile is –2.33. To find the value of *x*, we calculate $x = \mu + z\sigma = 100 + (-2.33)(16) = 62.7$.

6.30 Quartiles and outliers

a) The *z*-score corresponding to the lower quartile (25%) and the upper quartile of the standard normal distribution are –0.6745 (rounded to –0.67) and 0.6745 (rounded to 0.67), respectively.

b) The lower quartile of *X* is the 25th percentile Q1 = $\mu - 0.67\sigma$ = 200 + (–0.67)(36) = 175.88. The upper quartile of *X* is the 75th percentile Q3 = $\mu + 0.67\sigma$ = 200 + (0.67)(36) = 224.12. Thus, 175.88 is the value of *X* that marks the lowest 25% of values, and 224.12 is the value that marks the highest 25% of i values.

c) IQR = Q3 – Q1 = 224.12 – 175.88 = 48.24

d) Q1 – 1.5 × IQR = 175.88 – 1.5 × 48.24 = 103.52 and Q3 + 1.5 × IQR = 224.12 + 1.5 × 48.24 = 296.48. Therefore, the values of *X* that are less than 103.52 or more than 296.48 are considered as outliers.

6.31 April precipitation

a) $z = \frac{x-\mu}{\sigma} = \frac{8.4-3.6}{1.6} = 3$; If the distribution were roughly normal, this would be unusually high since 99.9% of observations fall below that level.

b) $z = \frac{x-\mu}{\sigma} = \frac{4.5-3.6}{1.6} = 0.563$; which corresponds to a cumulative probability of 0.713. Therefore, 4.5 inches of precipitation falls at the 71.3rd percentile.

c) The given percentages are similar to the percentages within 1, 2, or 3 standard deviations for the normal distribution, so it appears the distribution of April precipitation is approximately normal.

6.32 Automatic filling machine

a) $z = \frac{x-\mu}{\sigma} = \frac{500-502}{3} = -0.67$, which corresponds to a cumulative probability of 25% approximately. Thus, approximately 75% of nut packets are considered to be conformant.

b) $z = \frac{x-\mu}{\sigma} = \frac{510-502}{3} = 2.67$, which corresponds to a cumulative probability of 0.9962. Thus, approximately 99.62% of nut packets are less than 510 grams net weighted.

c) 0.25 + (1 – 0.9962) = 0.2538 of nut packets are either less than 500 grams net or more than 510 grams net weighted.

d) The median of the net weights of nuts in the packets corresponds to a z-score of 0. Therefore, the median is exactly the mean of the normal distribution of the net weights of nuts in the packets ($Q2 = \mu = 502$ grams)

6.33 SAT versus ACT

Joe's SAT score: $z = \frac{x-\mu}{\sigma} = \frac{600-500}{100} = 1.0$; Kate's ACT score: $z = \frac{x-\mu}{\sigma} = \frac{25-21}{4.7} = 0.85$

The SAT of 600 is relatively higher than the ACT of 25 because it is further from the mean in terms of the number of standard deviations, so Joe did relatively better.

6.34 Relative grades

Relative grades in Statistics: $z = \frac{x-\mu}{\sigma} = \frac{85-80}{4.5} = 1.11$

Relative grades in English: $z = \frac{x-\mu}{\sigma} = \frac{90-85}{4.0} = 1.25$

The English grade is relatively better than the Statistics grade score because it is further from the mean in terms of standard deviations. This is because Statistics grades are more spread out than English grades.

Section 6.3: Probabilities When Each Observation Has Two Possible Outcomes

6.35 Kidney transplants

a)

Sample Space	Probability
SSS	$(0.1)^3$
SSF	$(0.1)^2(0.9)$
SFS	$(0.1)^2(0.9)$
FSS	$(0.1)^2(0.9)$
SFF	$(0.1)(0.9)^2$
FSF	$(0.1)(0.9)^2$
FFS	$(0.1)(0.9)^2$
FFF	$(0.9)^3$

Compatible Donors	Probability
3	$(0.1)^3 = 0.001$
2	$3(0.1)^2(0.9) = 0.027$
1	$3(0.1)(0.9)^2 = 0.243$
0	$(0.9)^3 = 0.729$

6.35 (continued)

b) The formula for the binomial distribution is $P(x) = \frac{n!}{x!(n-x)!} p^x (1-p)^{n-x}$.

$$P(0) = \frac{3!}{0!(3-0)!} 0.1^0 (1-0.1)^{3-0} = 0.729, \quad P(1) = \frac{3!}{1!(3-1)!} 0.1^1 (1-0.1)^{3-1} = 0.243,$$

$$P(2) = \frac{3!}{2!(3-2)!} 0.1^2 (1-0.1)^{3-1} = 0.027, \quad P(3) = \frac{3!}{3!(3-3)!} 0.1^3 (1-0.1)^{3-3} = 0.001$$

6.36 Compatible donors

a) A trial consists of selecting a donor from the registry and observing whether the donor is compatible (two possible outcomes, i.e., binary). There are 3 trials.

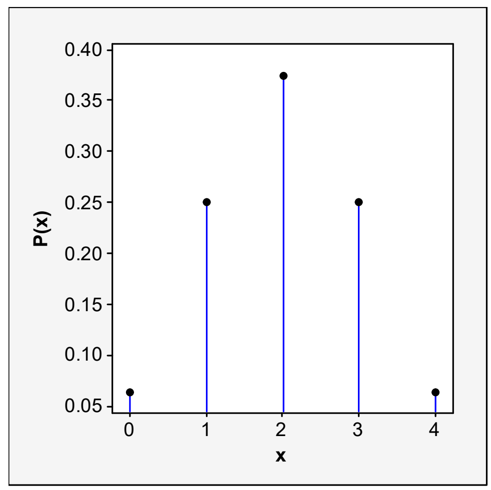
b) Yes, for each trial there is a 10% chance that the selected donor is compatible.

c) Whether one donor is compatible does not depend on some other donor being compatible, so the trials are independent.

6.37 Symmetric binomial

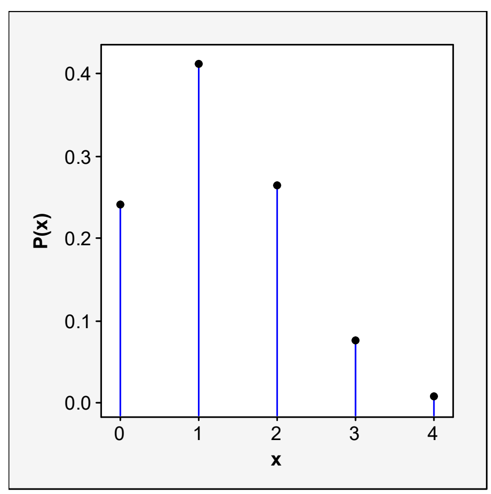
a)

x	P(x)
0	0.0625
1	0.2500
2	0.3750
3	0.2500
4	0.0625



b)

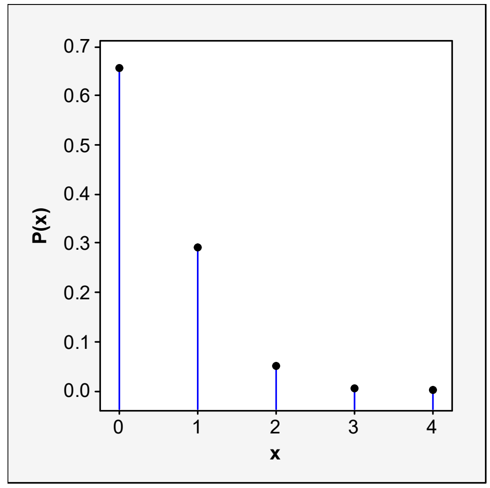
x	P(x)
0	0.2401
1	0.4116
2	0.2646
3	0.0756
4	0.0081



6.37 (continued)

c)

x	P(x)
0	0.6561
1	0.2916
2	0.0486
3	0.0036
4	0.0001



d) Only the graph in (a) is symmetric. The case $n = 20$ and $p = 0.5$ would be symmetric.

e) The graph in (c) is the most heavily skewed. The case $n = 4$ and $p = 0.01$ would exhibit more skewness than the graph in (c) since even more of the mass would be centered at 0.

6.38 Unfair wealth distribution sentiment

a) The data are binary (fair, unfair), there is the same probability of success for each trial (0.63), and the trials are independent (an opinion of an American has no bearing on the opinion of another American).

b) For this distribution, n is 10, and p is 0.63.

c) The probability that 2 Americans in the sample said wealth distribution is unfair is:

$$P(2) = \frac{10!}{2!(10-2)!}(0.63)^2(1-0.63)^{10-2} = (45)\left[(0.3969)(0.000351)\right] = 0.0063.$$

6.39 Bidding on eBay

a) Each bid can be thought of as a trial which can result in two outcomes (win or lose). The bids are independent, and for each bid there is a constant probability (25%) of winning.

b) The probability of winning exactly two bids is 0.211.

$$P(2) = \frac{4!}{2!(4-2)!}(0.25)^2(1-0.25)^{4-2} = 0.211$$

c) The probability of winning at most two bids is 0.949.

$$P(X \le 2) = P(0)+P(1)+P(2) = \frac{4!}{0!(4-0)!}(0.25)^0(1-0.25)^{4-0}$$
$$+\frac{4!}{1!(4-1)!}(0.25)^1(1-0.25)^{4-1}$$
$$+\frac{4!}{2!(4-2)!}(0.25)^2(1-0.25)^{4-2}$$
$$= 0.949$$

d) The probability of winning at more than two bids is 0.051.
Using the result from (c): $P(X > 2) = 1 - P(X \le 2) = 1 - 0.949 = 0.051.$

6.40 More eBay bidding

a) This is not a binomial distribution, the probability of success is not constant.

b) This is not a binomial distribution, X does not count number of successes.

6.41 Test generator

The distribution of the number of HARD questions among the 5 randomly selected one is Binomial with $n = 5$ and $p = 0.3$.

a) $P(5) = \binom{5}{5}(0.3)^5(0.7)^0 = 0.00243.$

b) $P(0) = \binom{5}{0}(0.3)^0(0.7)^5 = 0.1681.$

c) $P(\text{less than half of the questions are HARD}) = P(0) + P(1) + P(2)$

$$= 0.1681 + \binom{5}{1}(0.3)^1(0.7)^4 + \binom{5}{2}(0.3)^2(0.7)^3$$

$$= 0.1681 + 0.3087 + 0.3602 = 0.837.$$

6.42 NBA shooting

a) We must assume that the data are binary (which they are – free throw made or missed), that there is the same probability of success for each trial (free throw), and that the trials are independent.

b) $n = 10$; $p = 0.90$

c) (i) P(Makes all 10 free throws) = 0.349

$$P(10) = \frac{10!}{10!(10-10)!}(0.9)^{10}(1-0.9)^{10-10} = 0.3487$$

(ii) P(Makes 9 free throws) = 0.387

$$P(9) = \frac{10!}{9!(10-9)!}(0.9)^{9}(1-0.9)^{10-9} = 0.3874$$

(iii) P(Makes more than seven free throws) = P(8) + P(9) + P(10) = 0.3487 + 0.3874 + 0.1937 = 0.9298

$$P(8) = \frac{10!}{8!(10-8)!}(0.9)^{8}(1-0.9)^{10-8} = 0.1937$$

6.43 Season performance

a) $\mu = np = (400)(0.90) = 360,\ \ \sigma = \sqrt{np(1-p)} = \sqrt{400(0.9)(0.1)} = 6$

b) By the Empirical Rule, we would expect the number to fall almost certainly in the range from 360 – 3(6) = 342 to 360 + 3(6) = 378. This is the case because the probability within 3 standard deviations of the mean is close to 1.0.

c) We can calculate the proportion indicated by each end of the range. 342/400 = 0.855, and 378/400 = 0.945

6.44 Is the die balanced?

a) $n = 60$; $p = 1/6 = 0.1667$ (Which rounds to 0.167.)

b) $\mu = np = (60)\left(\frac{1}{6}\right) = 10,\ \ \sigma = \sqrt{np(1-p)} = \sqrt{60(1/6)(5/6)} = 2.887$ (Which rounds to 2.89.)

This indicates that if we sample 60 rolls from a population in which 1/6 are 6s, we would expect that about 10 in the sample are 6s. We also would expect a spread for this sample of about 2.887.

c) We would be skeptical because 0 is well over three standard deviations from the mean.

d) $\mathrm{P}(0) = \frac{60!}{0!(60-0)!}(0.1667)^{0}(1-0.1667)^{60-0} = 0.0000177$

6.45 Exit poll

a) This scenario satisfies the three conditions needed to use the binomial distribution because 1) the data are binary (voted for proposition or not), 2) there is the same probability of success for each trial (i.e., 0.50), and the trials are independent (the vote of one individual likely does not affect the vote of the next individual; n < 10% of the population size). $n = 3000$; $p = 0.50$.

6.45 (continued)

b) $\mu = np = (3000)(0.50) = 1500$, $\sigma = \sqrt{np(1-p)} = \sqrt{3000(0.50)(1-0.50)} = 27.4$

c) We would expect *X* to fall almost certainly between 1500 – 3(27.4) = 1418 and 1500 + 3(27.4) = 1582.

d) If the exit poll had $x = 1706$, this would suggest that the actual value of *p* is higher than 0.50.

6.46 Jury duty

a) One can assume that *X* has a binomial distribution because 1) the data are binary (Hispanic or not), 2) there is the same probability of success for each trial (i.e., 0.40), and 3) each trial is independent of the other trials (whom you pick for the first juror is not likely to affect whom you pick for the other jurors, and $n < 10\%$ of the population size). $n = 12$; $p = 0.40$

b) The probability that no Hispanic is selected is 0.002.

$$P(0) = \frac{12!}{0!(12-0)!}(0.40)^0(1-0.40)^{12-0} = 0.002$$

c) If no Hispanic is selected out of a sample of size 12, this does cast doubt on whether the sampling was truly random. There is only a 0.2 % chance that this would occur if the selection were done randomly.

6.47 Poor, poor, Pirates

a) $\mu = np = (162)(0.42) = 68.04$

b) Using technology: $P(X \geq 81) = 0.0242$.

c) It is unlikely that the probability of winning remains constant from one trial to the next since the team may use different pitchers/players for different games. It is also possible that the trials are not independent. After losing several games in a row, the players may become discouraged and not play as well in future games.

6.48 Checking guidelines

a) The sample is less than 10% of the size of the population size; thus, the guideline about the relative sizes of the population and the sample was satisfied.

b) The binomial distribution is not likely to have a bell shape because neither the expected number of success (*np*) nor the expected number of failures [$n(1-p)$] is at least 15. $np = 3.2$; $n(1-p) = 4.8$.

6.49 Movies sample

X would not have a binomial distribution because the trials are not independent. *n* needs to be less than 10% of the population size, but it is 25% of the population size.

6.50 Binomial needs fixed *n*

a) The formula for the probabilities for each possible outcome in a binomial distribution relies on a given number of trials, *n*.

b) The binomial applies for *X*. We only have *n* for *X*, not *Y*. $n = 3$, $p = 0.50$

6.51 Binomial assumptions

a) Not binomial. The disease is contagious; thus, it is very likely that a given family member get the disease to be dependent on others. The independence assumption is not plausible.

b) Binomial. The data are binary (married, unmarried), there is the same probability of success for each trial (0.4), and the trials are independent (the marital status of a woman has no bearing with the marital status of another woman in the population). Therefore, *X* is binomial with $n = 100$ and $p = 0.4$.

c) Not binomial. Five is not less than 10% of the population size of 20. We cannot assume independence.

d) Not binomial. The probability of dining out a day is the same in each day. Thus, the probability of success is not the same at each trial.

Chapter Problems: Practicing the Basics

6.52 Grandparents

a) This refers to a discrete random variable because there can only be whole numbers of grandparents. One can’t have 1.78 grandparents.

b) The probabilities satisfy the two conditions for a probability distribution because they each fall between 0 and 1, and the sum of the probabilities of all possible values is 1.

c) The mean of this probability distribution is $\mu = 0(0.71) + 1(0.15) + 2(0.09) + 3(0.03) + 4(0.02) = 0.50$.

6.53 Investment choices

Your expected earnings under Plan (ii) are 1000 × 2/3 + 1100 × 1/3 = \$1033.33. The expected earnings under Plan (iii) are 10,000 × 0.05 + 0 × 0.95 = \$500. Note that both Plans (ii) and (iii) will proceed with expected earnings less than the secured amount that will be obtained under Plan (i). Your expected profit under Plan (iii) is negative which is the worst scenario!

6.54 Auctioning paintings

a) Letting W = wins the bid, and L = loses the bid, the sample space is WW, WL, LW, LL.

b) No, the probability of winning the second bid changes with the outcome of the first bid, so the events are dependent.

c)

Outcome	Probability
WW	(0.3)(0.8) = 0.24
WL	(0.3)(0.2) = 0.06
LW	(0.7)(0.1) = 0.07
LL	(0.7)(0.9) = 0.63

d)

x	P(x)
\$5000	0.24
\$3000	0.06
\$2000	0.07
\$0	0.63

e) $\mu = 0(0.63) + 2000(0.07) + 3000(0.06) + 5000(0.24) = 1520$, or \$1520

6.55 Paths to school

a) The possible combinations of paths are (A, A), (A, B), (B, A) and (B, B), where the first path in each combination is that to get to school and the second one is that to get from school.

b) Your choices of path combinations to get from and to school are independent of each other, thus, P[(A,A)] = 3/4 × 1/2 = 3/8; P[(A,B)] = 3/4 × (1 – 1/2) = 3/8;
P[(B,A)] = (1 – 3/4) × 1/2 = 1/8 and P[(B,B)] = (1 – 3/4) × (1 – 1/2) = 1/8;

c) The lengths of the possible combinations of paths found in part (a) are 3 km, 2.8 km, 2.8 km and 2.6 km, respectively. Therefore, the average distance to get to and from school in both directions in a given day is 3 × 3/8 + 2.8 × 3/8 + 2.8 × 1/8 + 2.6 × 1/8 = 2.85 km

6.56 Are you risk averse?

a) The expected outcome for Program 1 is 200.
The expected outcome for Program 2 is 0.6667(0) + 0.3333(600) = 200.
Thus, the expected outcomes for the two programs are the same.

b) Expected deaths with Program 3 = 400
Expected deaths with Program 4 = 0.3333(0) + 0.6667(600) = 400
Thus, the expected deaths are the same for both programs.

c) Programs 1 and 3 are similar because the outcomes are known; these are risk-averse strategies.
Programs 2 and 4 are similar because they involve risk-taking.

6.57 Flyers' insurance

a)

Money	Probability
\$0	0.999999
\$100,000	0.000001

b) $\mu = 0(0.999999) + 100{,}000(0.000001) = 0.10$, or \$0.10

c) The company is very likely to make money in the long run because the return on each \$1 spent per flyer averages to \$0.10.

6.58 Normal probabilities

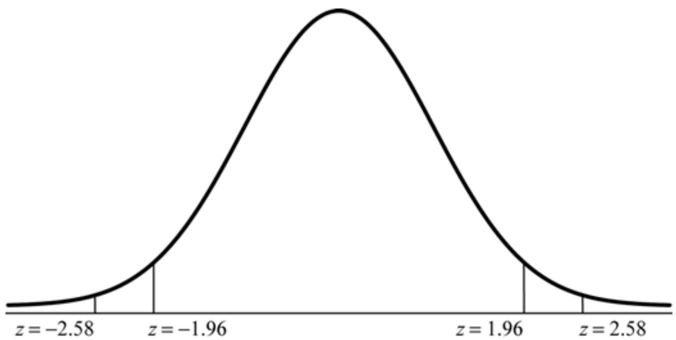
a) From Table A, we see that 0.9750 falls below a *z*-score of 1.96 and 0.0250 falls below –1.96. Thus, 0.9750 – 0.0250 = 0.95 would fall within 1.96 standard deviations of the mean.

b) From Table A, we see that 0.9901 falls below a *z*-score of 2.33 and 0.0099 falls below –2.33. Since 0.9901 – 0.0099 = 0.9802 would fall within 2.33 standard deviations of the mean, 1 – 0.9802 = 0.02 would fall more than 2.33 standard deviations from the mean.

6.59 *z*-scores

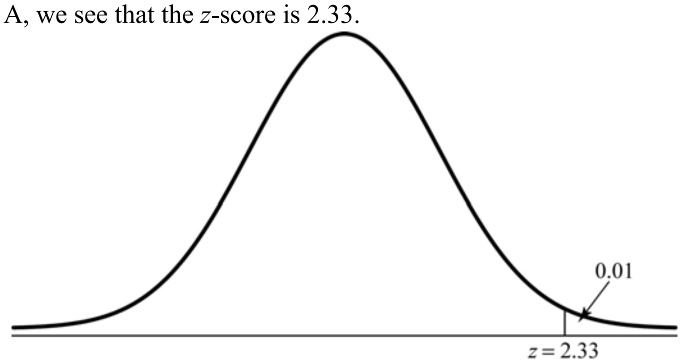
a) First, divide 0.95 in half to determine the amount between the mean and the *z*-score. This amount is 0.475, which means that 0.975 falls below the *z*-score in which we're interested. According to Table A, this corresponds to a *z*-score of 1.96.

b) First, divide 0.99 in half to determine the amount between the mean and the *z*-score. This amount is 0.495, which means that 0.995 falls below the *z*-score in which we're interested. According to Table A, this corresponds to a *z*-score of 2.58.



6.60 *z*-score and tail probability

a) The *z*-score that is less than only 1% of the values would be greater than 99% of the values. If we look up 0.99 on Table A, we see that the *z*-score is 2.33.



b) (i) The *z*-score that is above 0.90 is 1.28.
(ii) The *z*-score that is above 0.99 is 2.33.

6.61 Quartiles

a) If the interval contains 50% of a normal distribution, then there is 25% between the mean and the positive *z*-score. Added to 50% below the mean, 75% of the normal distribution is below the positive *z*-score. If we look this up in Table A, we find a *z*-score of 0.67.

b) The first quartile is at the 25th percentile, and the third is at the 75th percentile. We know that the 75th percentile has a *z*-score of 0.67. Because the normal distribution is symmetric, the *z*-score at the 25th percentile must be –0.67.

c) $\text{Q1} = \mu - 0.67\sigma$ and $\text{Q3} = \mu + 0.67\sigma$

The interquartile range is $\text{Q3} - \text{Q1} = (\mu + 0.67\sigma) - (\mu - 0.67\sigma) = 2(0.67)\sigma$.

6.62 Boys and girls birth weight

a) For boys, $z = \dfrac{x-\mu}{\sigma} = \dfrac{2.5-3.41}{0.55} = -1.645$; which corresponds to a cumulative probability of 0.049, so 2.5 kg falls at the 4.9th percentile for boys.

For girls, $z = \dfrac{x-\mu}{\sigma} = \dfrac{2.5-3.29}{0.52} = -1.519$; which corresponds to a cumulative probability of 0.064, so 2.5 kg falls at the 6.4th percentile for girls.

b) Since 2.5 kg falls at a lower percentile for boys, it is a more extreme weight for boys. Also, a z-score of –1.645 is more extreme than a z-score of –1.519, so 2.5 kg is a more extreme weight for boys.

6.63 Normal heart rate

a) For the lower bound $z = \dfrac{x-\mu}{\sigma} = \dfrac{60-80}{12} = -1.67$. This z-score corresponds with a proportion of 0.04746.

For the upper bound $z = \dfrac{x-\mu}{\sigma} = \dfrac{100-80}{12} = 1.67$. This z-score corresponds with a proportion of 0.95254 to the left of 1.67. Therefore, the proportion of individuals in the sample whose heart rate in the normal range is 0.95254 – 0.04746 = 0.906.

b) For Tachycardia, the z-score is $z = \dfrac{x-\mu}{\sigma} = \dfrac{100-80}{12} = 1.67$. Thus, the estimated proportion of individuals who have Tachycardia is 1 – 0.95254 = 0.0475 which corresponds to a number of about 19 in a sample of size 400.

6.64 Female heights

a) $z = \dfrac{x-\mu}{\sigma} = \dfrac{60-65}{3.5} = -1.43$; According to Table A, 0.076 of women are below this z-score; so 0.076 of women are below five feet in height.

b) $z = \dfrac{x-\mu}{\sigma} = \dfrac{72-65}{3.5} = 2.0$; 0.977 of women are below this z-score, which indicates that 1 – 0.977 = 0.023 of women are beyond this z-score; that is, over 6 feet in height.

c) 60 and 70 are equidistant from the mean of 65. If 0.076 of women are below 60 inches, that indicates that 0.076 women are above 70 inches, leaving 0.848 of women between 60 and 70 inches.

d) For North American males:

  a) $z = \dfrac{x-\mu}{\sigma} = \dfrac{60-70}{4.0} = -2.5$; According to Table A, 0.006 of men are below this z-score; therefore, 0.006 of men are below five feet in height.

  b) $z = \dfrac{x-\mu}{\sigma} = \dfrac{72-70}{4.0} = 0.5$; 0.691 of men are below this z-score, which indicates that 1 – 0.691 = 0.309 of men are beyond this z-score; that is, over 6 feet.

  c) 70 is the mean, and 60 has a cumulative probability of 0.006 (from part a). Thus, 0.50 – 0.006 = 0.494 of men are between 60 and 70 inches.

6.65 Cloning butterflies

a) $z = \dfrac{x-\mu}{\sigma} = \dfrac{8-9}{0.75} = -1.33$; According to Table A, 0.092 of the butterflies have a wingspan less than 8 cm.

b) $z = \dfrac{x-\mu}{\sigma} = \dfrac{10-9}{0.75} = 1.33$; 0.092 of the butterflies have a wingspan wider than 10 cm.

c) From (a) and (b), 1 – 2(0.092) = 0.816 of the butterflies have wingspans between 8 and 10 cm.

d) From Table A, the 90th percentile is 1.28. Using $x = \mu + z\sigma$, 10% of the butterflies have wingspan wider than 9 + 1.28(0.75) = 9.96 cm.

6.66 Gestation times

$z = \dfrac{x-\mu}{\sigma} = \dfrac{258-281.9}{11.4} = -2.10$; According to Table A, 0.018 fall below this z-score; thus, 0.018 of babies would be classified as premature.

6.67 Used car prices

a) $z = \dfrac{x-\mu}{\sigma} = \dfrac{25,000-23,800}{4380} = 0.27$; which corresponds to a cumulative probability of 0.608. Thus, 1 – 0.608 = 0.392, about 39.2% of used Audi A4s cost more than $25,000.

b) $z = \dfrac{x-\mu}{\sigma} = \dfrac{22,000-23,800}{4380} = -0.41$; which corresponds to a cumulative probability of 0.341 and $z = \dfrac{18,000-23,800}{4380} = -1.32$; which corresponds to a cumulative probability of 0.093. Thus, 0.341 – 0.093 = 0.248, about 24.8% of used Audi A4s cost between $18,000 and $25,000.

c) The z-score that corresponds to the lowest 10% is –1.2816. Using $x = \mu + z\sigma$, 10% of used Audi 4As cost less than 23,800 – 1.2816(4380) = 18,187, or about $18,187.

6.68 Used car deals

a) According to Table A, a z-score of 1.5 indicates that 0.93 of the curve falls below this point. Therefore, 1 – 0.93 = 0.067, or 6.7% of used Audi 4As will be highlighted.
b) 6.7% of used Honda Civics will be highlighted for the same reasons as in (a). Since the z-score is given, the mean and standard deviation are not needed.
c) The percentage will most likely be larger because more cars will be priced at the lower end if distribution is right-skewed.

6.69 Global warming

The third quartile indicates the point that 75% of observations fall below, and corresponds to a z-score of 0.67. We can solve for μ algebraically in the equation below; the mean weekly gas usage, therefore, would have to be reduced to 16.0.

$$0.67 = (20-\mu)/6 \Rightarrow 20-\mu = 4.02 \Rightarrow \mu = 16$$

6.70 Fast food profits

a) $z = \dfrac{x-\mu}{\sigma} = \dfrac{0-140}{80} = -1.75$; Table A indicates that 0.04 of observations fall below this z-score. The probability that the restaurant loses money on a given day is 0.04.
b) Probability = $(1 - 0.04)^7 = (0.96)^7 = 0.75$. For this calculation to be valid, each day's take must be independent of every other day's take.

6.71 Metric height

a) The distribution would still be normal because it is the same distribution in another metric; just as converting to z-scores would give us the same distribution, converting to centimeters would give us the same distribution.
b) In centimeters, the mean would be (72)(2.54) = 182.88 cm, and the standard deviation would be (4)(2.54) = 10.16 cm.
c) $z = \dfrac{x-\mu}{\sigma} = \dfrac{200-182.88}{10.16} = 1.69$; On Table A, this z-score corresponds to a proportion of 0.95. Thus, 0.95 fall below this height, and 1 –0.95 = 0.05 fall above it.

6.72 Manufacturing tennis balls

a) 56.7 grams: $z = \dfrac{x-\mu}{\sigma} = \dfrac{56.7-57.6}{0.3} = -3.0$; 58.5 grams: $z = \dfrac{x-\mu}{\sigma} = \dfrac{58.5-57.6}{0.3} = 3.0$

The proportion of observations below a z-score of 3.0 is 0.9987 and below –3.0 is 0.0013.
0.9987 – 0.0013 = 0.997 (essentially 1.0) is the probability that a ball manufactured with this machine satisfies the rules.

6.72 (continued)

b) 56.7 grams: $z = \dfrac{x-\mu}{\sigma} = \dfrac{56.7-57.6}{0.6} = -1.5$; 58.5 grams: $z = \dfrac{x-\mu}{\sigma} = \dfrac{558.5-57.6}{0.6} = 1.5$

The proportion of observations below a z-score of 1.5 is 0.933 and below –1.5 is 0.067. 0.933 – 0.067 = 0.866 (rounds to 0.87), which is the probability that a ball manufactured with this machine satisfies the rules.

6.73 Bride's choice of surname

If each marriage is independent of the others (as they would be in a random sample), the probability would be (0.9)(0.9)(0.9)(0.9) = 0.6561.

6.74 ESP

a) Sample space: (SSS, SSF, SFS, SFF, FSS, FSF, FFS, FFF), where S = guessing correctly.
The probability of each is (0.5)(0.5)(0.5) = 0.125. For each number of successes, the probability equals the number of outcomes with that number of successes multiplied by 0.125. For example, there is one outcome with zero successes, giving a probability of 0.125.

x	**P(*x*)**
0	0.125
1	0.375
2	0.375
3	0.125

b) The probability distribution will be the same as in (a).

$$\mathrm{P}(0) = \frac{3!}{0!(3-0)!}(0.5)^0(1-0.5)^{3-0} = 0.125,\ \ \mathrm{P}(1) = \frac{3!}{1!(3-1)!}(0.5)^1(1-0.5)^{3-1} = 0.375,$$

$$\mathrm{P}(2) = \frac{3!}{2!(3-2)!}(0.5)^2(1-0.5)^{3-2} = 0.375,\ \mathrm{P}(3) = \frac{3!}{3!(3-3)!}(0.5)^3(1-0.5)^{3-3} = 0.125$$

6.75 More ESP

a) For the analogy with coin flipping, heads would represent a correct guess and tails would represent an incorrect guess on any trial.

b) It is sensible to assume the same probability for each trial because with random guessing, she has a constant probability of 0.2 of guessing the right number, for each of the three trials.

c) It is sensible to assume independent trials because trials are not affected by the outcomes of previous trials.

6.76 Yale babies

Yes, since the probability that 14 or more infants would choose the helpful figure randomly is so small (0.002), this can be considered evidence that the infants are exhibiting a preference for the helpful object.

$$P(14)+P(15)+P(16)$$
$$= \frac{14!}{14!(16-14)!}(0.5)^{14}(1-0.5)^{16-14} + \frac{15!}{15!1!}(0.5)^{15}(0.5)^1 + \frac{16!}{16!0!}(0.5)^{16}(0.5)^0$$
$$= 0.002$$

6.77 Weather

a) If R = rain and D = dry, the possibilities for the weekend's weather are RR, RD, DR and DD, all equally likely because P(rain) = 0.5. Because it rains on at least one of the days for 3 out of the 4 possibilities, the probability is 0.75.

b) $1-\mathrm{P}(0) = 1 - \dfrac{2!}{0!(2-0)!}(0.5)^0(1-0.5)^{2-0} = 1-0.25 = 0.75$

6.78 Dating success

a) There are three conditions to use the binomial distribution.
1) The outcomes must be binary (e.g., yes or no as the only two options).
2) He must have the same probability of success for each call.
3) Each trial (phone call) must be independent of the others.

b) If he calls the same girl five times, her responses are not likely to be independent of her other responses!

c) $n = 5$; $p = 0.60$; $\mu = np = 5(0.60) = 3$

6.79 Total loss

a) The distribution of X is binomial with $n = n$ and $p = 0.0015$. Therefore, the mean of X is $\mu = np = n(0.0015)$.

b) For the mean to be 10, the number of policies in the portfolio should be $n = \mu/0.0015 = 10/0.0015 = 6666.67$ (rounded to 6667).

c) The expected number of total loss claims in a portfolio of 6667 is 10. The cost of these claims is 10 × 15,000 = \$150,000. However, the amount of premiums collected is 6667 × 500 = \$3333,500 which makes an expected return of 3333,500 – 150,000 = \$3183,500 of this company's portfolio.

6.80 Likes on Facebook

a) The data are binary (like message, not like message), there is the same probability of success (liking the message) for each trial, and the trials are independent (one user's response will not affect another's). $n = 15{,}000{,}000$, $p = 0.00001$

b) $\mu = np = 15{,}000{,}000(0.00001) = 150,\ \ \sigma = \sqrt{np(1-p)} = \sqrt{15{,}000{,}000(0.00001)(1-0.00001)} = 12.25$

c) Using $x = \mu + z\sigma$, 150 – 3(12.25) = 113.25 and 150 + 3(12.25) = 186.75, so the interval is [113.25, 186.75].

d) Users may not have the same probability of liking the message and responses may not be independent. (If you see your friend liking the message, you may be more inclined to like it, too, if you received it.)

6.81 Survival

The distribution of X = number of deaths to occur between ages 1 and 4 is binomial with $n = 98{,}500$ and $p = 0.0002$.

a) $E(X) = np = 98{,}500 \times 0.0002 = 19.7$

b) $np = 19.7$ and $np(1 - p) = 19.7 \times 0.9998 = 19.696$. Both are greater than 15. Therefore, we can reasonably approximate the distribution of X by a normal distribution with mean $\mu = 19.7$ and standard deviation $\sigma = \sqrt{19.696} = 4.44$. Almost all values of X will fall within 3 standard deviations of the mean, i.e., between $\mu - 3\sigma = 6.38$ (rounded to 6) and $\mu + 3\sigma = 33.02$ (rounded to 33).

c) The z-score of 10 is $\dfrac{10-19.7}{4.44} = -2.1847$, which corresponds to a cumulative probability of about 0.0145. Thus, the probability that the number of deaths to occur between ages 1 and 4 does not exceed 10 is about 1.45%.

6.82 Which distribution for sales?

a) Since (i) each trial has two outcomes, (ii) the probability of a successful phone call is the same for each call (2%), and (iii) the trials are independent (the outcome of one call does not affect the outcome of another call) the distribution is binomial.

b) $\mu = np = 200(0.02) = 4.0,\ \ \sigma = \sqrt{np(1-p)} = \sqrt{200(0.02)(0.98)} = 2.0$; The expected number of successful calls out of 200 is 4.

c) $\mathrm{P}(0) = \binom{200}{0}(0.98)^{200}(0.02)^{0} = 0.018$

Chapter Problems: Concepts and Investigations

6.83 Best of five

a) Let A = team A wins (P(A) = 0.5) and B = Team B wins (P(B) = 0.5).

Outcome	x	Probability	Outcome	x	Probability
AAA	3	$(0.5)^3 = 0.125$	BAAA	4	$(0.5)^4 = 0.0625$
AABA	4	$(0.5)^4 = 0.0625$	BAABA	5	$(0.5)^5 = 0.03125$
AABBA	5	$(0.5)^5 = 0.03125$	BAABB	5	$(0.5)^5 = 0.03125$
AABBB	5	$(0.5)^5 = 0.03125$	BABAA	5	$(0.5)^5 = 0.03125$
ABAA	4	$(0.5)^4 = 0.0625$	BABAB	5	$(0.5)^5 = 0.03125$
ABABA	5	$(0.5)^5 = 0.03125$	BAAB	4	$(0.5)^4 = 0.0625$
ABABB	5	$(0.5)^5 = 0.03125$	BBAAA	5	$(0.5)^5 = 0.03125$
ABBAA	5	$(0.5)^5 = 0.03125$	BBAAB	5	$(0.5)^5 = 0.03125$
ABBAB	5	$(0.5)^5 = 0.03125$	BBAB	4	$(0.5)^4 = 0.0625$
ABBB	4	$(0.5)^4 = 0.0625$	BBB	3	$(0.5)^3 = 0.125$

b) See table in (a)

c) See table in (a)

d) P(3) = 2(0.125) = 0.25, P(4) = 6(0.0625) = 0.375, P(5) = 12(0.031250) = 0.375

6.84 More best of five

a) $\mu = 3(0.25) + 4(0.375) + 5(0.375) = 4.125$

b) Let A = team A wins (P(A) = 0.8) and B = Team B wins (P(B) = 0.2).

Outcome	x	Probability	Outcome	x	Probability
AAA	3	$(0.8)^3 = 0.512$	BAAA	4	$(0.8)^3(0.2) = 0.1024$
AABA	4	$(0.8)^3(0.2) = 0.1024$	BAABA	5	$(0.8)^3(0.2)^2 = 0.02048$
AABBA	5	$(0.8)^3(0.2)^2 = 0.02048$	BAABB	5	$(0.8)^2(0.2)^3 = 0.00512$
AABBB	5	$(0.8)^2(0.2)^3 = 0.00512$	BABAA	5	$(0.8)^3(0.2)^2 = 0.02048$
ABAA	4	$(0.8)^3(0.2) = 0.1024$	BABAB	5	$(0.8)^2(0.2)^3 = 0.00512$
ABABA	5	$(0.8)^3(0.2)^2 = 0.02048$	BAAB	4	$(0.8)(0.2)^3 = 0.0064$
ABABB	5	$(0.8)^2(0.2)^3 = 0.00512$	BBAAA	5	$(0.8)^3(0.2)^2 = 0.02048$
ABBAA	5	$(0.8)^3(0.2)^2 = 0.02048$	BBAAB	5	$(0.8)^2(0.2)^3 = 0.00512$
ABBAB	5	$(0.8)^2(0.2)^3 = 0.00512$	BBAB	4	$(0.8)(0.2)^3 = 0.0064$
ABBB	4	$(0.8)(0.2)^3 = 0.0064$	BBB	3	$(0.2)^3 = 0.008$

P(3) = 0.512 + 0.008 = 0.52, P(4) = 3(0.1024) + 3(0.0064) = 0.3264, P(5) = 6(0.02048) + 6(0.00512) = 0.1536; $\mu = 3(0.52) + 4(0.3264) + 5(0.1536) = 3.63$

6.85 Family size in Gaza

The median would be the score that falls at 50%. If we add the probabilities from either end, we find that 4 falls at 0.50. The median is 4.

6.86 Longest streak made

a) The mean increases by one for each doubling of number of shots. The chance of a streak becomes longer with more trials.

b) (i) For 400, we would expect 8 because 400 is double 200.
   (ii) For 3200, we would expect 11 because we would have to double 400 to 800 to 1600 to 3200.

c) 95% of a bell-shaped curve falls within about two standard deviations of the mean, and 4 is a bit more than two standard deviations.

6.87 Stock market randomness

A streak of 7 is likely to occur just by chance given enough trials.

6.88 Airline overbooking

a) Using the binomial distribution with $n = 190$ and $p = 0.8$, $\mu = np = 190(0.8) = 152$ and $\sigma = \sqrt{npq}$ $= \sqrt{190(0.8)(0.2)} = 5.5136$. Since both $np = 152$ and $n(1 - p) = 38$ are greater than 15, the normal distribution can be used to approximate the binomial distribution. Thus, we expect nearly all counts to fall within 3 standard deviations of the mean, i.e., approximately between 152 – 3(5.5) = 135 and 152 + 3(5.5) = 169 seats. Since 170 falls outside of this range, the airline is justified in selling more tickets for the flight.

b) There might be situations where large groups buy tickets and travel together. This would violate the assumption of independent trials.

6.89 Babies in China

Based on the original birth rate and a sample size of 2000, we'd have $\mu = np = 2000(0.49) = 980$ and $\sigma = \sqrt{np(1-p)} = \sqrt{2000(0.49)(1-0.49)} = 22.4$; we'd expect most samples to have the number of female babies fall within three standard deviations of the mean for a range of 980 – 3(22.4) = 913 to 980 + 3(22.4) = 1047. Because only 800 females were born in this particular year, we can conclude that the current probability of a female birth in this town does seem to be less than it used to be. This relies on an assumption of independence which might not hold true in a given town.

6.90 True or false? IQR for normal distribution

False, as show in Exercise 6.30, Q1 and Q3 represent z-scores of –0.67 and 0.67, so the IQR would cover the interval from $\mu \pm 0.67\sigma$, which is smaller that the interval of $\mu \pm \sigma$ given in this exercise.

6.91 Multiple choice: Guess answers

The best answer is (d).

6.92 Multiple choice: Terrorist coincidence?

The best answer is (a).

♦♦6.93 SAT and ethnic groups

a) Group A: $z = \dfrac{1200-1500}{300} = -1$; which corresponds to a cumulative probability of 0.16. The proportion not admitted for ethnic group A is 0.16.

Group B: $z = \dfrac{1200-1350}{200} = -0.75$; which corresponds to a cumulative probability of 0.23.

b) (0.16+0.23) = 0.39 so 0.23/0.39 = 0.59 of those not admitted are from ethnic group B.

c) Group A: $z = \dfrac{600-1500}{300} = -3$; which corresponds to a cumulative probability of 0.0013.

Group B: $z = \dfrac{600-1350}{200} = -3.75$; which corresponds to a cumulative probability of 0.00009.

Now, (0.00009)/(0.0013+0.00009) = 0.065 are from group B.

♦♦6.94 College acceptance

a) Since the scores are close to continuous and likely to follow a bell-shaped distribution with most of the scores falling within 3 standard deviations of the mean, the appropriate distribution is the normal.

b) The top 20% corresponds to a cumulative probability of 80%, which yields a z-score of 0.842. Using $x = \mu + z\sigma$, the corresponding ACT score is 21.1 + 0.842(5.3) = 25.56.

c) $P(0)=\binom{5}{0}(0.2)^0\,(0.8)^5 = 0.328$

♦♦6.95 Standard deviation of a discrete probability distribution

a) $\sigma = \sqrt{(4-5.8215)^2(0.1250)+\cdots+(7-5.8215)^2(0.3215)} = \sqrt{1.02734} = 1.01$

b) The standard deviation will be smaller since the observations are centered at 4 games, with very few series lasting longer than that. For the 99% chance scenario, $\sigma = 0.194$.

6.96 Mean and standard deviation for a binary random variable

a) $\mu = 0(1-p)+1(p) = p$

b) $\sigma = \sqrt{\sigma^2} = \sqrt{(0-p)^2(1-p)+(1-p)^2(p)} = \sqrt{p^2-p^3+p^3-2p^2+p} = \sqrt{p-p^2} = \sqrt{p(1-p)}$

♦♦6.97 Linear transformations: Taxes and fees

a) The distribution of the new prices will still be normal with $\mu = 1.06(23{,}800) = \$25{,}228$ and $\sigma = 1.06(4380) = \$4643$.

b) The distribution of the new prices will still be normal with $\mu = 23{,}800+199 = \$23{,}999$ and $\sigma = \$4380$.

♦♦6.98 Binomial probabilities

If we want to see the probability of multiple independent events occurring, we have to multiply the probabilities of each of those events. If the events have the same probabilities, we can use p^x rather than multiply *p* by itself *x* times. We also want to know the probability of the event NOT occurring multiple times. The probability of the event not occurring is $1-p$, and the number of times we're interested in is all the times that the event doesn't occur, or $n-x$. The logic for using $(1-p)^{n-x}$ is the same as the logic for using p^x.

♦♦6.99 Waiting time for doubles

a) The probability of rolling doubles is 1/6. Whatever you get on the second die has a 1/6 chance of matching the first. To have doubles occur first on the second roll, you'd have to have no match on the first roll; there's a 5/6 chance of that. You'd then have to have a match on the second roll, and there's a 1/6 chance of that. To get the probability of both occurring is (5/6)(1/6). For no doubles until the third roll, both the first and the second rolls would not match, a 5/6 chance for each, followed by doubles on the third, a 1/6 chance. The probability of all of three events is $(5/6)^2(1/6)$.

b) By the logic in (a), P(4) = (5/6)(5/6)(5/6)(1/6) or $(5/6)^3(1/6)$. By extension, we could calculate P(*x*) for any *x* by $(5/6)^{(x-1)}(1/6)$.

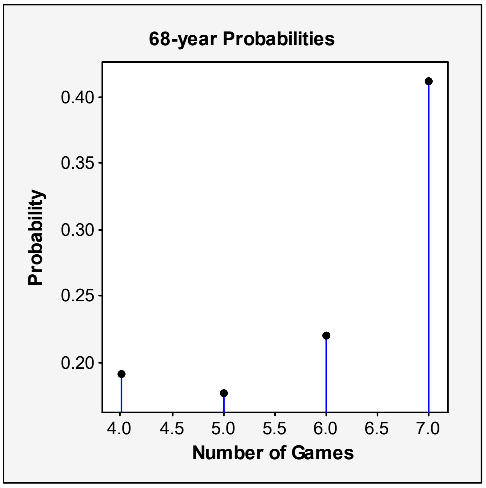
6.100 Geometric mean

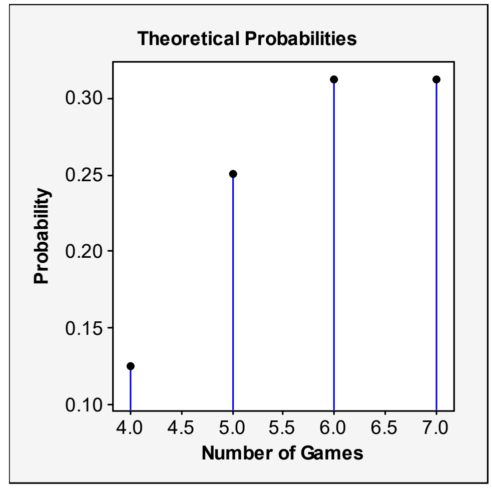
We would expect another 1/0.024 = 41.7 losing seasons, so their next win is expected in 2042 (given the assumptions hold, in particular, the constant probability of a win being 0.42.

6.101 World series in baseball

a) $\mu = 4(13/68) + 5(12/68) + 6(15/68) + 7(28/68) = 5.853$

b)





The 68-year distribution is much more left-skewed, with a mode at 7 games. Among the reasons for the difference between the 68-yeardistribution and the theoretical distribution is the likely fact that win probabilities are not constant (home-field advantage).

c) You would expect $68(0.3125) = 21.25$ series to be decided in seven games.

Chapter Problems: Student Activities

6.102 Best of seven games

The results will be different each time this exercise is conducted.