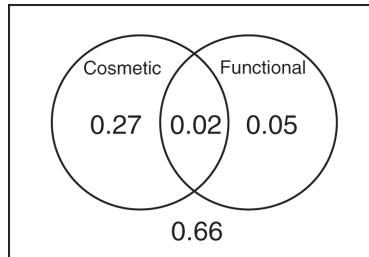


Review of Part IV – Randomness and Probability

R4.1. Quality Control.

a) $P(\text{defect}) = P(\text{cosmetic defect}) + P(\text{functional defect}) - P(\text{cosmetic and functional defects})$
 $= 0.29 + 0.07 - 0.02$
 $= 0.34$

Or, from the Venn: $0.27 + 0.02 + 0.05 = 0.34$



b) $P(\text{cosmetic defect and no functional defect}) = P(\text{cosmetic defect}) - P(\text{cosmetic and functional defects})$
 $= 0.29 - 0.02$
 $= 0.27$

Or, from the Venn: 0.27 (region inside Cosmetic circle, outside Functional circle)

c) $P(\text{func.} \mid \text{cosm.}) = \frac{P(\text{func. and cosm.})}{P(\text{cosm.})} = \frac{0.02}{0.29} \approx 0.069$

From the Venn, consider only the region inside the Cosmetic circle. The probability that the car has a functional defect is 0.02 out of a total of 0.29 (the entire Cosmetic circle).

- d) The two kinds of defects are not disjoint events, since 2% of cars have both kinds.
e) Approximately 6.9% of cars with cosmetic defects also have functional defects. Overall, the probability that a car has a cosmetic defect is 7%. The probabilities are estimates, so these are probably close enough to say that the two types of defects are independent.

R4.2. Workers.

Organize the counts in a two-way table.

| Job Type | Male | Female | Total |
|-------------|------|--------|-------|
| Management | 7 | 6 | 13 |
| Supervision | 8 | 12 | 20 |
| Production | 45 | 72 | 117 |
| Total | 60 | 90 | 150 |

a) i) $P(\text{female}) = \frac{90}{150} = 0.6$

ii) $P(\text{female or production}) = P(\text{female}) + P(\text{production}) - P(\text{female and production})$
 $= \frac{90}{150} + \frac{117}{150} - \frac{72}{150}$
 $= 0.9$

R4.2. (continued)

iii) Consider only the production row of the table. There are 72 women out of 117 production workers, and $72/117 \approx 0.615$. Or, use the formula.

$$P(\text{female} \mid \text{production}) = \frac{P(\text{female and production})}{P(\text{production})} = \frac{\cancel{72}/\cancel{150}}{\cancel{117}/\cancel{150}} \approx 0.615$$

iv) Consider only the female column. There are 72 production workers out of a total of 90 women, and $72/90 = 0.8$. Or, use the formula.

$$P(\text{production} \mid \text{female}) = \frac{P(\text{production and female})}{P(\text{female})} = \frac{\cancel{72}/\cancel{150}}{\cancel{90}/\cancel{150}} = 0.8$$

- b)** No. Fewer than half of the managers are female, but 60% of all workers are female.

R4.3. Airfares.

- a)** Let C = the price of a ticket to China
Let F = the price of a ticket to France

$$\text{Total price of airfare} = 3C + 5F$$

b) $\mu = E(3C + 5F) = 3E(C) + 5E(F) = 3(1000) + 5(500) = \5500

$$\sigma = SD(3C + 5F) = \sqrt{3^2(Var(C)) + 5^2(Var(F))} = \sqrt{3^2(150^2) + 5^2(100^2)} \approx \$672.68$$

c) $\mu = E(C - F) = E(C) - E(F) = 1000 - 500 = \500

$$\sigma = SD(C - F) = \sqrt{Var(C) + Var(F)} = \sqrt{150^2 + 100^2} \approx \$180.28$$

- d)** No assumptions are necessary when calculating means. When calculating standard deviations, we must assume that ticket prices are independent of each other for different countries but all tickets to the same country are at the same price.

R4.4. Bipolar.

Let X = the number of people with bipolar disorder in a city of $n = 10,000$ residents.

These may be considered Bernoulli trials. There are only two possible outcomes, having bipolar disorder or not having bipolar disorder. Psychiatrists estimate that the probability that a person has bipolar is about 1 in 100, so $p = 0.01$. We will assume that the cases of bipolar disorder are distributed randomly throughout the populations. The trials are not independent, since the population is finite, but 10,000 people represent fewer than 10% of all people. Therefore, the number of people with bipolar disorder in a city of 10,000 may be modeled by $\text{Binom}(10000, 0.01)$.

Since $np = 100$ and $nq = 9900$ are both greater than 10, $\text{Binom}(10000, 0.01)$ may be approximated by the Normal model, $N(100, 9.95)$.

$$E(X) = np = 10,000(0.01) = 100 \text{ residents}; SD(X) = \sqrt{npq} = \sqrt{10,000(0.01)(0.99)} \approx 9.95 \text{ residents}$$

We expect 100 city residents to have bipolar disorder. According to the Normal model, 200 cases would be over 10 standard deviations above this mean. The probability of this occurring is essentially zero.

Technology can compute the probability according to the Binomial model. Again, the probability that 200 cases of bipolar disorder exist in the city is essentially zero. We use the Normal model in this case, since it gives us a more intuitive idea of just how unlikely this event is.

R4.5. A game.

- a) Let X = net amount won

| | | | |
|--------|------|------|------|
| X | \$0 | \$2 | -\$2 |
| $P(X)$ | 0.10 | 0.40 | 0.50 |

$$\mu = E(X) = 0(0.10) + 2(0.40) - 2(0.50) = -\$0.20$$

$$\sigma^2 = \text{Var}(X) = (0 - (-0.20))^2(0.10) + (2 - (-0.20))^2(0.40) + (-2 - (-0.20))^2(0.50) = 3.56$$

$$\sigma = SD(X) = \sqrt{\text{Var}(X)} = \sqrt{3.56} \approx \$1.89$$

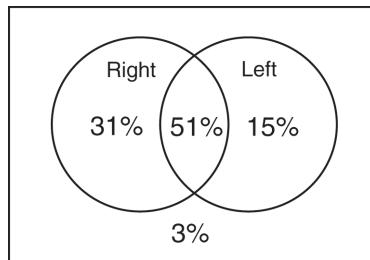
- b) $X + X$ = the total winnings for two plays.

$$\mu = E(X + X) = E(X) + E(X) = (-0.20) + (-0.20) = -\$0.40$$

$$\sigma = SD(X + X) = \sqrt{\text{Var}(X) + \text{Var}(X)} = \sqrt{3.56 + 3.56} \approx \$2.67$$

R4.6. Emergency switch.

Construct a Venn diagram of the disjoint outcomes.



- a) From the Venn diagram, 3% of the workers were unable to operate the switch with either hand.
- b) $P(\text{left} \mid \text{right}) = \frac{P(\text{left and right})}{P(\text{right})} = \frac{0.51}{0.82} \approx 0.622$; About 62% of the workers who could operate the switch with their right hands could also operate it with left hands. Overall, the probability that a worker could operate the switch with his right hand was 66%. Workers who could operate the switch with their right hands were less likely to be able to operate the switch with their left hand, so success is not independent of hand.
- c) Success with right and left hands are not disjoint events. Fifty-one percent of the workers had success with both hands.

R4.7. Twins.

The selection of these women can be considered Bernoulli trials. There are two possible outcomes, twins or no twins. As long as the women selected are representative of the population of all pregnant women, then $p = 1/90$. (If the women selected are representative of the population of women taking Clomid, then $p = 1/10$.) The trials are not independent since the population of all women is finite, but 10 women are fewer than 10% of the population of women.

Let X = the number of twin births from $n = 10$ pregnant women.

Let Y = the number of twin births from $n = 10$ pregnant women taking Clomid.

288 Part IV Randomness and Probability

R4.7. (continued)

- a) Use $\text{Binom}(10, 1/90)$.

$$\begin{aligned} P(\text{at least one has twins}) &= 1 - P(\text{none have twins}) \\ &= 1 - P(X = 0) \\ &= 1 - {}_{10}C_0 \left(\frac{1}{90} \right)^0 \left(\frac{89}{90} \right)^{10} \\ &\approx 0.106 \end{aligned}$$

- b) Use $\text{Binom}(10, 1/10)$.

$$\begin{aligned} P(\text{at least one has twins}) &= 1 - P(\text{none have twins}) \\ &= 1 - P(Y = 0) \\ &= 1 - {}_{10}C_0 \left(\frac{1}{10} \right)^0 \left(\frac{9}{10} \right)^{10} \\ &\approx 0.651 \end{aligned}$$

- c) Use $\text{Binom}(5, 1/90)$ and $\text{Binom}(5, 1/10)$.

$$\begin{aligned} P(\text{at least one has twins}) &= 1 - P(\text{no twins without Clomid})P(\text{no twins with Clomid}) \\ &= 1 - \left({}_5C_0 \left(\frac{1}{90} \right)^0 \left(\frac{89}{90} \right)^5 \right) \left({}_5C_0 \left(\frac{1}{10} \right)^0 \left(\frac{9}{10} \right)^5 \right) \\ &\approx 0.442 \end{aligned}$$

R4.8. Deductible.

$$\mu = E(\text{cost}) = 500(0.005) = \$2.50$$

$$\sigma^2 = \text{Var}(\text{cost}) = (2.50 - 500)^2(0.005) + (2.50 - 0)^2(0.995) = 1243.75$$

$$\sigma = \text{SD}(\text{cost}) = \sqrt{\text{Var}(\text{cost})} = \sqrt{1243.75} \approx \$35.27$$

Expected (extra) cost of the cheaper policy with the deductible is \$2.50, much less than the \$12 surcharge for the policy with no deductible, so on average she will save money by going with the deductible. The standard deviation, at \$35.27, is quite high compared to the \$12 surcharge, indicating a high amount of variability. The value of the car shouldn't influence the decision.

R4.9. More twins.

In Exercise 7, it was determined that these were Bernoulli trials. Use $\text{Binom}(5, 0.10)$.

Let X = the number of twin births from $n = 5$ pregnant women taking Clomid.

- a) $P(\text{none have twins}) = P(X = 0)$

$$\begin{aligned} &= {}_5C_0(0.1)^0(0.9)^5 \\ &\approx 0.590 \end{aligned}$$

- b) $P(\text{one has twins}) = P(X = 1)$

$$\begin{aligned} &= {}_5C_1(0.1)^1(0.9)^4 \\ &\approx 0.328 \end{aligned}$$

- c) $P(\text{at least three will have twins}) = P(X = 3) + P(X = 4) + P(X = 5)$

$$\begin{aligned} &= {}_5C_3(0.1)^3(0.9)^2 + {}_5C_4(0.1)^4(0.9)^1 + {}_5C_5(0.1)^5(0.9)^0 \\ &= 0.00856 \end{aligned}$$

R4.10. At fault.

If we assume that these drivers are representative of all drivers insured by the company, then these insurance policies can be considered Bernoulli trials. There are only two possible outcomes, accident or no accident. The probability of having an accident is constant, $p = 0.005$. The trials are not independent, since the populations of all drivers is finite, but 1355 drivers represent fewer than 10% of all drivers. Use $\text{Binom}(1355, 0.005)$.

- a) Let X = the number of drivers who have an at-fault accident out of $n = 1355$.

$$E(X) = np = 1355(0.005) = 6.775 \text{ drivers; } SD(X) = \sqrt{npq} = \sqrt{1355(0.005)(0.995)} \approx 2.60 \text{ drivers}$$

- b) Since $np = 6.775 < 10$, the Normal model cannot be used to model the number of drivers who are expected to have accidents. The Success/Failure condition is not satisfied.

R4.11. Twins, part III.

In Exercise 7, it was determined that these were Bernoulli trials. Use $\text{Binom}(152, 0.10)$.

Let X = the number of twin births from $n = 152$ pregnant women taking Clomid.

- a) $E(X) = np = 152(0.10) = 15.2$ births; $SD(X) = \sqrt{npq} = \sqrt{152(0.10)(0.90)} \approx 3.70$ births.

- b) Since $np = 15.2$ and $nq = 136.8$ are both greater than 10, the Success/Failure condition is satisfied and $\text{Binom}(152, 0.10)$ may be approximated by $N(15.2, 3.70)$.

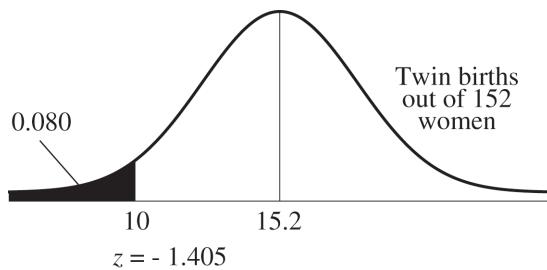
- c) **Using Binom(152, 0.10):**

$$\begin{aligned} P(\text{no more than } 10) &= P(X \leq 10) \\ &= P(X = 0) + \dots + P(X = 10) \\ &= {}_{152}C_0(0.10)^0(0.90)^{152} + \dots + {}_{152}C_{10}(0.10)^{10}(0.90)^{142} \\ &\approx 0.097 \end{aligned}$$

According to the Binomial model, the probability that no more than 10 women would have twins is approximately 0.097.

Using $N(15.2, 3.70)$:

$$z = \frac{x - \mu}{\sigma} = \frac{10 - 15.2}{3.70} \approx -1.405$$



$P(X \leq 10) \approx P(z < -1.405) \approx 0.080$; According to the Normal model, the probability that no more than 10 women would have twins is approximately 0.080.

R4.12. Child's play.

- a) Let X = the number indicated on the spinner

| | | | |
|--------|-----|------|------|
| X | 5 | 10 | 20 |
| $P(X)$ | 0.5 | 0.25 | 0.25 |

b) $\mu = E(X) = 5(0.5) + 10(0.25) + 20(0.25) = 10$

$$\sigma^2 = Var(X) = (5 - 10)^2(0.5) + (10 - 10)^2(0.25) + (20 - 10)^2(0.25) = 37.5$$

$$\sigma = SD(X) = \sqrt{Var(X)} = \sqrt{37.5} \approx 6.12$$

- c) Let Y = the number indicated on the die

| | | | | | |
|--------|---------------|---------------|---------------|---------------|---------------|
| Y | 0 | 1 | 2 | 3 | 4 |
| $P(Y)$ | $\frac{1}{3}$ | $\frac{1}{6}$ | $\frac{1}{6}$ | $\frac{1}{6}$ | $\frac{1}{6}$ |

d) $\mu = E(Y) = 0\left(\frac{1}{3}\right) + 1\left(\frac{1}{6}\right) + 2\left(\frac{1}{6}\right) + 3\left(\frac{1}{6}\right) + 4\left(\frac{1}{6}\right) = \frac{10}{6} \approx 1.67$

$$Var(Y) = \left(0 - \frac{10}{6}\right)^2\left(\frac{1}{3}\right) + \left(1 - \frac{10}{6}\right)^2\left(\frac{1}{6}\right) + \left(2 - \frac{10}{6}\right)^2\left(\frac{1}{6}\right) + \left(3 - \frac{10}{6}\right)^2\left(\frac{1}{6}\right) + \left(4 - \frac{10}{6}\right)^2\left(\frac{1}{6}\right) \approx 2.22$$

$$\sigma = SD(Y) = \sqrt{Var(Y)} = \sqrt{2.22} \approx 1.49$$

e) $\mu = E(X + Y) = E(X) + E(Y) \approx 10 + 1.67 \approx 11.67$ spaces

$$\sigma = SD(X + Y) = \sqrt{Var(X) + Var(Y)} \approx \sqrt{37.5 + 2.22} \approx 6.30$$
 spaces

R4.13. Language.

Assuming that the freshman composition class consists of 25 randomly selected people, these may be considered Bernoulli trials. There are only two possible outcomes, having the specified language center or not having the specified language center. The probabilities of the specified language centers are constant at 80%, 10%, or 10%, for right, left, and two-sided language center, respectively. The trials are not independent, since the population of people is finite, but we will select fewer than 10% of all people.

- a) Let L = the number of people with left-brain language control from $n = 25$ people.

Use $\text{Binom}(25, 0.80)$.

$$\begin{aligned} P(\text{no more than } 15) &= P(L \leq 15) \\ &= P(L = 0) + \dots + P(L = 15) \\ &= {}_{25}C_0(0.80)^0(0.20)^{25} + \dots + {}_{25}C_{15}(0.80)^{15}(0.20)^{10} \\ &\approx 0.0173 \end{aligned}$$

According to the Binomial model, the probability that no more than 15 students in a class of 25 will have left-brain language centers is approximately 0.0173.

- b) Let T = the number of people with two-sided language control from $n = 5$ people.

Use $\text{Binom}(5, 0.10)$.

$$\begin{aligned} P(\text{none have two-sided language control}) &= P(T = 0) \\ &= {}_5C_0(0.10)^0(0.90)^5 \\ &\approx 0.590 \end{aligned}$$

R4.13. (continued)

- c) Use Binomial models:

$$E(\text{left}) = np_L = 1200(0.80) = 960 \text{ people}$$

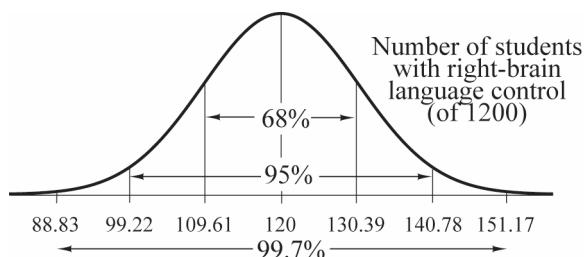
$$E(\text{right}) = np_R = 1200(0.10) = 120 \text{ people}$$

$$E(\text{two-sided}) = np_T = 1200(0.10) = 120 \text{ people}$$

- d) Let R = the number of people with right-brain language control.

$$E(R) = np_R = 1200(0.10) = 120 \text{ people}; SD(R) = \sqrt{np_R q_R} = \sqrt{1200(0.10)(0.90)} \approx 10.39 \text{ people.}$$

- e) Since $np_R = 120$ and $nq_R = 1080$ are both greater than 10, the Normal model, $N(120, 10.39)$, may be used to approximate $\text{Binom}(1200, 0.10)$. According to the Normal model, about 68% of randomly selected groups of 1200 people could be expected to have between 109.61 and 130.39 people with right-brain language control. About 95% of randomly selected groups of 1200 people could be expected to have between 99.22 and 140.78 people with right-brain language control. About 99.7% of randomly selected groups of 1200 people could be expected to have between 88.83 and 151.17 people with right-brain language control.

**R4.14. Play, again.**

$$\mu = E(X - Y) = E(X) - E(Y) \approx 10 - 1.67 \approx 8.33 \text{ spaces}$$

$$\sigma = SD(X - Y) = \sqrt{Var(X) + Var(Y)} \approx \sqrt{37.5 + 2.22} \approx 6.30 \text{ spaces}$$

R4.15. Beanstalks.

- a) The greater standard deviation for men's heights indicates that men's heights are more variable than women's heights.
- b) Admission to a Beanstalk Club is based upon extraordinary height for both men and women, but men are slightly more likely to qualify. The qualifying height for women is about 2.4 SDs above the mean height of women, while the qualifying height for men is about 1.75 SDs above the mean height for men.
- c) Let M = the height of a randomly selected man from $N(69.1, 2.8)$.

Let W = the height of a randomly selected woman from $N(64.0, 2.5)$.

$M - W$ = the difference in height of a randomly selected man and woman.

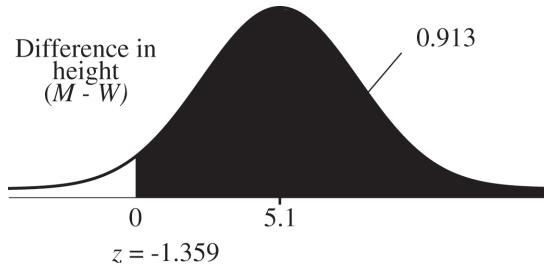
d) $E(M - W) = E(M) - E(W) = 69.1 - 64.0 = 5.1 \text{ inches}$

e) $SD(M - W) = \sqrt{Var(M) + Var(W)} = \sqrt{2.8^2 + 2.5^2} \approx 3.75 \text{ inches}$

R4.15. (continued)

- f) Since each distribution is described by a Normal model, the distribution of the difference in height between a randomly selected man and woman is $N(5.1, 3.75)$.

$$z = \frac{y - \mu}{\sigma} = \frac{0 - 5.1}{\sqrt{2.8^2 + 2.5^2}} \approx -1.359$$



According to the Normal model, the probability that a randomly selected man is taller than a randomly selected woman is approximately 0.913.

- g) If people chose spouses independent of height, we would expect 91.3% of married couples to consist of a taller husband and shorter wife. The 92% that was seen in the survey is close to 91.3%, and the difference may be due to natural sampling variability. Unless this survey is very large, there is not sufficient evidence of association between height and choice of partner.

R4.16. Stocks.

- a) $P(\text{market will rise for 3 consecutive years}) = (0.73)^3 \approx 0.389$
- b) Use $\text{Binom}(5, 0.73)$. $P(\text{market will rise in 3 out of 5 years}) = {}_5C_3(0.73)^3(0.27)^2 \approx 0.284$
- c) $P(\text{fall in at least 1 of next 5 years}) = 1 - P(\text{no fall in 5 years}) = 1 - (0.73)^5 \approx 0.793$
- d) Let $X = \text{the number of years in which the market rises}$. Use $\text{Binom}(10, 0.73)$.
 $P(\text{rises in the majority of years in a decade}) = P(X \geq 6)$

$$\begin{aligned} &= P(X = 6) + \dots + P(X = 10) \\ &= {}_{10}C_6(0.73)^6(0.27)^4 + \dots + {}_{10}C_{10}(0.73)^{10}(0.27)^0 \\ &\approx 0.896 \end{aligned}$$

R4.17. Multiple choice.

Guessing at questions can be considered Bernoulli trials. There are only two possible outcomes, correct or incorrect. If you are guessing, the probability of success is $p = 0.25$, and the questions are independent. Use $\text{Binom}(50, 0.25)$ to model the number of correct guesses on the test.

- a) Let $X = \text{the number of correct guesses}$.
 $P(\text{at least 30 of 50 correct}) = P(X \geq 30)$
- $$\begin{aligned} &= P(X = 30) + \dots + P(X = 50) \\ &= {}_{50}C_{30}(0.25)^{30}(0.75)^{20} + \dots + {}_{50}C_{50}(0.25)^{50}(0.75)^0 \\ &\approx 0.00000016 \end{aligned}$$

You are **very** unlikely to pass by guessing on every question.

R4.17. (continued)

- b)** Use Binom(50, 0.70).

$$\begin{aligned} P(\text{at least 30 of 50 correct}) &= P(X \geq 30) \\ &= P(X = 30) + \dots + P(X = 50) \\ &= {}_{50}C_{30}(0.70)^{30}(0.30)^{20} + \dots + {}_{50}C_{50}(0.70)^{50}(0.30)^0 \\ &\approx 0.952 \end{aligned}$$

According to the Binomial model, your chances of passing are about 95.2%.

- c)** Use Geom(0.70).

$$P(\text{first correct on third question}) = (0.30)^2(0.70) = 0.063$$

R4.18. Stock strategy.

- a)** This does not confirm the advice. Stocks have risen 75% of the time after a 2-year fall, but there have only been eight occurrences of the 2-year fall. The sample size is very small, and therefore highly variable.
- b)** Stocks have actually risen in 73% of years. This is not much different from the strategy of the advisors, which yielded a rise in 75% of years (from a very small sample of years.)

R4.19. Insurance.

The company is expected to pay \$100,000 only 2.6% of the time, while always gaining \$520 from every policy sold. When they pay, they actually only pay \$99,480.

$$E(\text{profit}) = \$520(0.974) - \$99,480(0.026) = -\$2,080.$$

The expected profit is actually a **loss** of \$2,080 per policy. The company had better raise its premiums if it hopes to stay in business.

R4.20. Teen smoking.

Randomly selecting high school students can be considered Bernoulli trials. There are only two possible outcomes, smoker or nonsmoker. The probability that a student is a smoker is $p = 0.18$. The trials are not independent, since the population is finite, but we are not sampling more than 10% of all high school students.

- a)** $P(\text{none of the first 4 are smokers}) = (0.82)^4 = 0.452$

- b)** Use Geom(0.18).

$$P(\text{first smoker is the sixth person}) = (0.82)^5(0.18) \approx 0.067$$

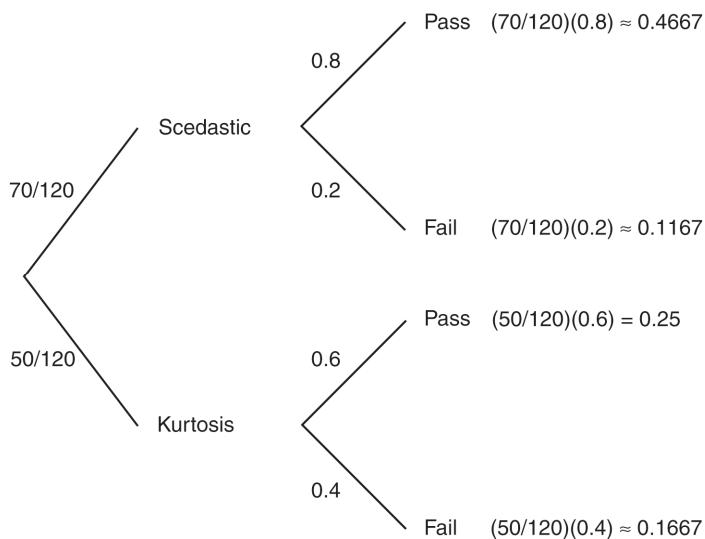
- c)** Use Binom(10, 0.18). Let $X =$ the number of smokers among $n = 10$ students.

$$P(\text{no more than 2 smokers of 10}) = P(X \leq 2)$$

$$\begin{aligned} &= P(X = 0) + P(X = 1) + P(X = 2) \\ &= {}_{10}C_0(0.18)^0(0.82)^{10} + {}_{10}C_1(0.18)^1(0.82)^9 + {}_{10}C_2(0.18)^2(0.82)^8 \\ &\approx 0.737 \end{aligned}$$

R4.21. Passing stats.

Organize the information in a tree diagram.



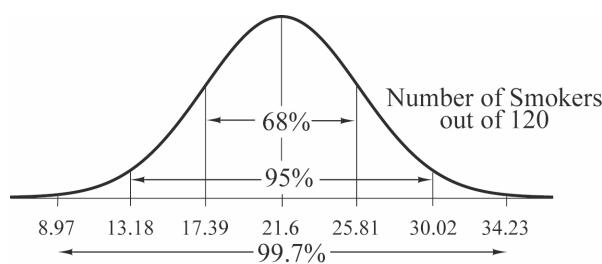
$$\begin{aligned}
 \text{a)} \quad P(\text{Passing Statistics}) &= P(\text{Scedastic and Pass}) + P(\text{Kurtosis and Pass}) \\
 &\approx 0.4667 + 0.25 \\
 &\approx 0.717
 \end{aligned}$$

$$\text{b)} \quad P(\text{Kurtosis} \mid \text{Fail}) = \frac{P(\text{Kurtosis and Fail})}{P(\text{Fail})} \approx \frac{0.1667}{0.1167 + 0.1667} \approx 0.588$$

R4.22. Teen smoking II.

In Exercise 20, it was determined that the selection of students could be considered to be Bernoulli trials.

- a) Use $\text{Binom}(120, 0.18)$ to model the number of smokers out of $n = 120$ students.
 $E(\text{number of smokers}) = np = 120(0.18) = 21.6$ smokers.
- b) $SD(\text{number of smokers}) = \sqrt{npq} = \sqrt{120(0.18)(0.82)} \approx 4.21$ smokers.
- c) Since $np = 21.6$ and $nq = 98.4$ are both greater than 10, the Success/Failure condition is satisfied and $\text{Binom}(120, 0.18)$ may be approximated by $N(21.6, 4.21)$.
- d) According to the Normal model, approximately 68% of samples of size $n = 120$ are expected to have between 17.39 and 25.81 smokers, approximately 95% of the samples are expected to have between 13.18 and 30.02 smokers, and approximately 99.7% of the samples are expected to have between 8.97 and 34.23 smokers.



R4.23. Random variables.

- a) $\mu = E(X + 50) = E(X) + 50 = 50 + 50 = 100$
 $\sigma = SD(X + 50) = SD(X) = 8$
- b) $\mu = E(10Y) = 10E(Y) = 10(100) = 1000$
 $\sigma = SD(10Y) = 10SD(Y) = 60$
- c) $\mu = E(X + 0.5Y) = E(X) + 0.5E(Y) = 50 + 0.5(100) = 100$
 $\sigma = SD(X + 0.5Y) = \sqrt{Var(X) + 0.5^2 Var(Y)} = \sqrt{8^2 + 0.5^2(6^2)} \approx 8.54$
- d) $\mu = E(X - Y) = E(X) - E(Y) = 50 - 100 = -50$
 $\sigma = SD(X - Y) = \sqrt{Var(X) + Var(Y)} = \sqrt{8^2 + 6^2} = 10$
- e) $\mu = E(X_1 + X_2) = E(X) + E(X) = 50 + 50 = 100$
 $\sigma = SD(X_1 + X_2) = \sqrt{Var(X) + Var(X)} = \sqrt{8^2 + 8^2} \approx 11.31$

R4.24. Merger.

Small companies may run into trouble in the insurance business. Even if the expected profit from each policy is large, the profit is highly variable. There is a small chance that a company would have to make several huge payouts, resulting in an overall loss, not a profit. By combining two small companies together, the company takes in profit from more policies, making the larger company more resistant to the possibility of a large payout. This is because the total profit is increasing by the expected profit from each additional policy, but the standard deviation is increasing by the square root of the sum of the variances. The larger a company gets, the more the expected profit outpaces the variability associated with that profit.

R4.25. Youth survey.

- a) Many boys play computer games and use email, so the probabilities can total more than 100%. There is no evidence that there is a mistake in the report.
- b) Playing computer games and using email are not mutually exclusive. If they were, the probabilities would total 100% or less.
- c) Emailing friends and being a boy or girl are not independent. Seventy-six percent of girls emailed friends in the last week, but only 65% of boys emailed. If emailing were independent of being a boy or girl, the probabilities would be the same.
- d) Let X = the number of students chosen until the first student is found who does not use the Internet. Use Geom(0.07). $P(X = 5) = (0.93)^4(0.07) \approx 0.0524$.

R4.26. Meals.

Let X = the amount the student spends daily.

- a) $\mu = E(X + X) = E(X) + E(X) = 13.50 + 13.50 = \27.00
 $\sigma = SD(X + X) = \sqrt{Var(X) + Var(X)} = \sqrt{7^2 + 7^2} \approx \9.90
- b) In order to calculate the standard deviation, we must assume that spending on different days is independent. This is probably not valid, since the student might tend to spend less on a day after he has spent a lot. He might not even have money left to spend!

296 Part IV Randomness and Probability

R4.26. (continued)

c) $\mu = E(X + X + X + X + X + X + X) = 7E(X) = 7(\$13.50) = \$94.50$

$$\begin{aligned}\sigma &= SD(X + X + X + X + X + X + X) \\ &= \sqrt{Var(X) + Var(X) + Var(X) + Var(X) + Var(X) + Var(X) + Var(X)} \\ &= \sqrt{7(7^2)} \approx \$18.52\end{aligned}$$

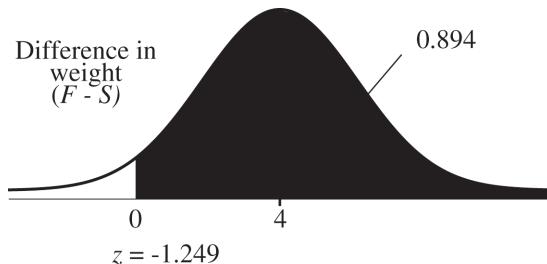
- d) Assuming once again that spending on different days is independent, it is unlikely that the student will spend less than \$50. This level of spending is about 2.4 standard deviations below the weekly mean. Don't try to approximate the probability! We don't know the shape of this distribution.

R4.27. Travel to Kyrgyzstan.

- a) Spending an average of 4237 soms per day, you can stay about $\frac{90,000}{4237} \approx 21$ days.
- b) Assuming that your daily spending is independent, the standard deviation is the square root of the sum of the variances for 21 days.
- $$\sigma = \sqrt{21(360)^2} \approx 1649.73 \text{ som}$$
- c) The standard deviation in your total expenditures is about 1650 som, so if you don't think you will exceed your expectation by more than 2 standard deviations, bring an extra 3300 som. This gives you a cushion of about 157 soms for each of the 21 days.

R4.28. Picking melons.

- a) $\mu = E(\text{First} - \text{Second}) = E(\text{First}) - E(\text{Second}) = 22 - 18 = 4 \text{ lbs.}$
- b) $\sigma = SD(\text{First} - \text{Second}) = \sqrt{Var(\text{First}) + Var(\text{Second})} = \sqrt{2.5^2 + 2^2} \approx 3.20 \text{ lbs.}$
- c) $z = \frac{y - \mu}{\sigma} = \frac{0 - 4}{\sqrt{2.5^2 + 2^2}} \approx -1.249$



According to the Normal model, the probability that a melon from the first store weighs more than a melon from the second store is approximately 0.894.

R4.29. Home sweet home.

Since the homes are randomly selected, these can be considered Bernoulli trials. There are only two possible outcomes, owning the home or not owning the home. The probability of any randomly selected resident home being owned by the current resident is 0.66. The trials are not independent, since the population is finite, but as long as the city has more than 8200 homes, we are not sampling more than 10% of the population. The Binomial model, $\text{Binom}(820, 0.66)$, can be used to model the number of homeowners among the 820 homes surveyed. Let H = the number of homeowners found in $n = 820$ homes.

$$E(H) = np = 820(0.66) = 541.2 \text{ homes}; \quad SD(H) = \sqrt{npq} = \sqrt{820(0.66)(0.34)} \approx 13.56 \text{ homes}$$

R4.29. (continued)

The 523 homeowners found in the candidate's survey represent a number of homeowners that is only about 1.34 standard deviations below the expected number of homeowners. It is not particularly unusual to be 1.34 standard deviations below the mean. There is little support for the candidate's claim of a low level of home ownership.

R4.30. Buying melons.

The mean price of a watermelon at the first store is $22(0.32) = \$7.04$.

At the second store the mean price is $18(0.25) = \$4.50$.

The difference in the price is expected to be $\$7.04 - \$4.50 = \$2.54$.

The standard deviation in price at the first store is $2.5(0.32) = \$0.80$.

At the second store, the standard deviation in price is $2(0.25) = \$0.50$.

The standard deviation of the difference is $\sqrt{0.80^2 + 0.50^2} \approx \0.94 .

R4.31. Who's the boss?

a) $P(\text{first three owned by women}) = (0.30)^3 \approx 0.027$

b) $P(\text{none of the first four are owned by women}) = (0.70)^4 \approx 0.240$

c) $P(\text{sixth firm called is owned by women} \mid \text{none of the first five were}) = 0.30$

Since the firms are chosen randomly, the fact that the first five firms were owned by men has no bearing on the ownership of the sixth firm.

R4.32. Jerseys.

a) $P(\text{all four kids get the same color}) = 4\left(\frac{1}{4}\right)^4 \approx 0.0156$

(There are four different ways for this to happen, one for each color.)

b) $P(\text{all four kids get white}) = \left(\frac{1}{4}\right)^4 \approx 0.0039$

c) $P(\text{all four kids get white}) = \left(\frac{1}{6}\right)\left(\frac{1}{4}\right)^3 \approx 0.0026$

R4.33. When to stop?

- a) Since there are only two outcomes, 6 or not 6, the probability of getting a 6 is $1/6$, and the trials are independent, these are Bernoulli trials. Use $\text{Geom}(1/6)$.

$$\mu = \frac{1}{p} = \frac{1}{1/6} = 6 \text{ rolls}$$

- b) If 6's are not allowed, the mean of each die roll is $\frac{1+2+3+4+5}{5} = 3$. You would expect to get 15 if you rolled 5 times.

c) $P(5 \text{ rolls without a } 6) = \left(\frac{5}{6}\right)^5 \approx 0.402$

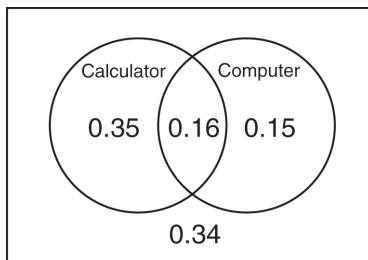
298 Part IV Randomness and Probability

R4.34. Plan B.

- a) If 6's are not allowed, the mean of each die roll is $\frac{1+2+3+4+5}{5} = 3$.
- b) Let X = your current score. You expect to lose it all $\frac{1}{6}$ of the time, so your expected loss per roll is $\frac{1}{6}X$.
- c) Expected gain equals expected loss when $\frac{1}{6}X = 3$. So, $X = 18$.
- d) Roll until you get 18 points, then stop.

R4.35. Technology on campus.

a)



$$\begin{aligned}
 P(\text{neither tech.}) &= 1 - P(\text{either tech.}) \\
 &= 1 - [P(\text{calculator}) + P(\text{computer}) - P(\text{both})] \\
 &= 1 - [0.51 + 0.31 - 0.16] \\
 &= 0.34
 \end{aligned}$$

Or, from the Venn: 0.34 (the region outside both circles) This is MUCH easier.

Thirty-four percent of students use neither type of technology.

- b) $P(\text{calc. and no comp.}) = P(\text{calc.}) - P(\text{calc. and comp.}) = 0.51 - 0.16 = 0.35$

Or, from the Venn: 0.35 (region inside the Calculator circle, outside the Computer circle)

Thirty-five percent of students use calculators, but not computers.

- c) $P(\text{computer} | \text{calculator}) = \frac{P(\text{comp. and calc.})}{P(\text{calc.})} = \frac{0.16}{0.51} \approx 0.314$

About 31.4% of calculator users have computer assignments.

- d) Thirty-one point four percent of classes that used calculators used computer assignments, while in classes that didn't use calculators, 30.6% used computer assignments. These are close enough to think the choice is probably independent.

R4.36. Dogs.

Since the outcomes are disjoint, probabilities may be added and subtracted as needed.

- a) $P(\text{no dogs}) = (0.77)(0.77) = 0.5929$
- b) $P(\text{some dogs}) = 1 - P(\text{no dogs}) = 1 - (0.77)(0.77) = 0.407$
- c) $P(\text{both dogs}) = (0.23)(0.23) = 0.053$
- d) $P(\text{more than one dog in each}) = (0.05)(0.05) \approx 0.0025$

R4.37. O-rings.

- a) A Poisson model would be used.
- b) If the probability of failure for one O-ring is 0.01, then the mean number of failures for 10 O-rings is $E(X) = \lambda = np = 10(0.01) = 0.1$ O-ring. We are able to calculate this because the Poisson model scales according to sample size.
- c) $P(\text{one failed O-ring}) = \frac{e^{-0.1}(0.1)^1}{1!} \approx 0.090$
- d) $P(\text{at least one failed O-ring}) = 1 - P(\text{no failures})$
 $= 1 - \frac{e^{-0.1}(0.1)^0}{0!} \approx 0.095$

R4.38. Volcanoes.

- a) A Poisson model would be used.
- b) Since the mean number of volcanic episodes is 2.4 per year, then the mean number of volcanic episodes for 2 years is 4.8.
- c) $P(\text{no volcanic episodes}) = \frac{e^{-4.8}(4.8)^0}{0!} \approx 0.008$
- d) $P(\text{more than three volcanic episodes}) = 1 - [P(X = 0) + P(X = 1) + P(X = 2) + P(X = 3)]$
 $= 1 - \left[\frac{e^{-4.8}(4.8)^0}{0!} + \frac{e^{-4.8}(4.8)^1}{1!} + \frac{e^{-4.8}(4.8)^2}{2!} + \frac{e^{-4.8}(4.8)^3}{3!} \right]$
 ≈ 0.706

(or, use technology. For example: $1 - \text{poissoncdf}(4.8,3)$ on the TI-83.)

R4.39. Socks.

Since we are sampling without replacement, use conditional probabilities throughout.

- a) $P(\text{2 blue}) = \left(\frac{4}{12}\right)\left(\frac{3}{11}\right) = \frac{12}{132} = \frac{1}{11}$
- b) $P(\text{no gray}) = \left(\frac{7}{12}\right)\left(\frac{6}{11}\right) = \frac{42}{132} = \frac{7}{22}$
- c) $P(\text{at least one black}) = 1 - P(\text{no black}) = 1 - \left(\frac{9}{12}\right)\left(\frac{8}{11}\right) = \frac{60}{132} = \frac{5}{11}$
- d) $P(\text{green}) = 0$ (There aren't any green socks in the drawer.)
- e) $P(\text{match}) = P(\text{2 blue}) + P(\text{2 gray}) + P(\text{2 black})$
 $= \left(\frac{4}{12}\right)\left(\frac{3}{11}\right) + \left(\frac{5}{12}\right)\left(\frac{4}{11}\right) + \left(\frac{3}{12}\right)\left(\frac{2}{11}\right) = \frac{19}{66}$

R4.40. Coins.

Coin flips are Bernoulli trials. There are only two possible outcomes, the probability of each outcome is constant, and the trials are independent.

- a) Use $\text{Binom}(36, 0.5)$. Let H = the number of heads in $n = 36$ flips.

$$\mu = E(H) = np = 36(0.5) = 18 \text{ heads}; \quad \sigma = SD(H) = \sqrt{npq} = \sqrt{36(0.5)(0.5)} = 3 \text{ heads}$$

- b) Two standard deviations above the mean corresponds to 6 “extra” heads observed.
- c) The standard deviation of the number of heads when 100 coins are flipped is $\sigma = \sqrt{npq} = \sqrt{100(0.5)(0.5)} = 5$ heads. Getting 6 “extra” heads is not unusual.
- d) Following the “two standard deviations” measurement, 10 or more “extra” heads would be unusual.
- e) What appears surprising in the short run becomes expected in a larger number of flips. The “Law of Averages” is refuted, because the coin does not compensate in the long run. A coin that is flipped many times is actually less likely to show exactly half heads than a coin flipped only a few times. The Law of Large Numbers is confirmed, because the percentage of heads observed gets closer to the percentage expected due to probability.

R4.41. The Drake equation.

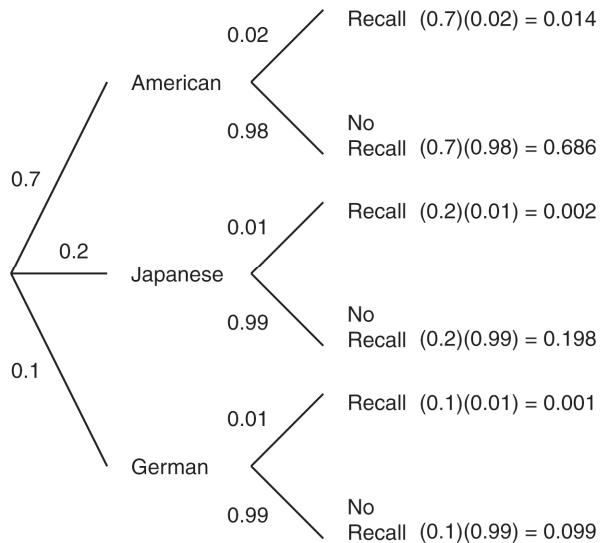
- a) $N \cdot f_p$ represents the number of stars in the Milky Way Galaxy expected to have planets.
- b) $N \cdot f_p \cdot n_e \cdot f_i$ represents the number of planets in the Milky Way Galaxy expected to have intelligent life.
- c) $f_l \cdot f_i$ is the probability that a planet has a suitable environment and has intelligent life.
- d) $f_l = P(\text{life} \mid \text{suitable environment})$; This is the probability that life develops, if a planet has a suitable environment.

$f_i = P(\text{intelligence} \mid \text{life})$; This is the probability that the life develops intelligence, if a planet already has life.

$f_c = P(\text{communication} \mid \text{intelligence})$; This is the probability that radio communication develops, if a planet already has intelligent life.

R4.42. Recalls.

Organize the information in a tree diagram.

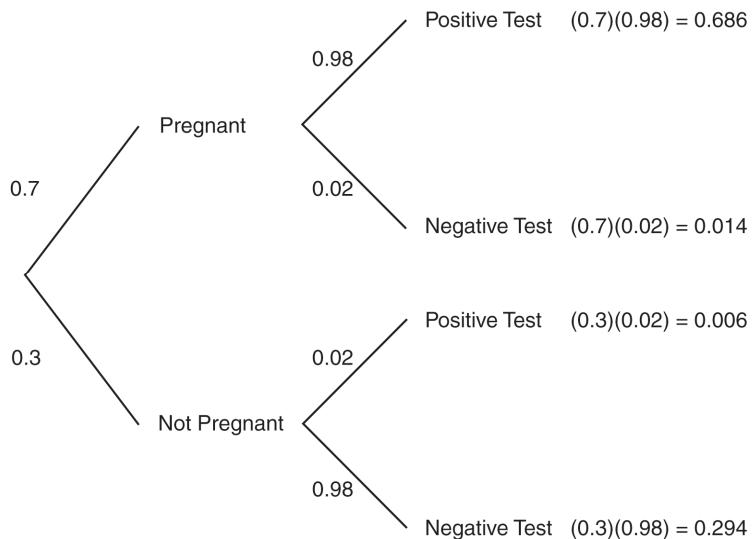


$$\begin{aligned}
 \text{a)} \quad P(\text{recall}) &= P(\text{American recall}) + P(\text{Japanese recall}) + P(\text{German recall}) \\
 &= 0.014 + 0.002 + 0.001 \\
 &= 0.017
 \end{aligned}$$

$$\text{b)} \quad P(\text{American} \mid \text{recall}) = \frac{P(\text{American and recall})}{P(\text{recall})} = \frac{0.014}{0.014 + 0.002 + 0.001} \approx 0.824$$

R4.43. Pregnant?

Organize the information in a tree diagram.



$$P(\text{pregnant} \mid + \text{test}) = \frac{P(\text{preg. and} + \text{test})}{P(+ \text{test})} = \frac{0.686}{0.686 + 0.006} \approx 0.991$$

R4.44. Door prize.

- a) The probability that the first person in line wins is 1 out of 100, or 0.01.
- b) If you are third in line, the two people ahead of you must not win in order for you to win. The probability is $(0.99)(0.99)(0.01) = 0.009801$.
- c) There must be 100 losers in a row. The probability is $(0.99)^{100} \approx 0.366$.
- d) The first person in line has the greatest chance of winning at $p = 0.01$. The probability of winning decreases from there, since winning is dependent upon everyone else in front of you in line losing.
- e) Position is irrelevant now. Everyone has the same chance of winning, $p = 0.01$. One way to visualize this is to imagine that one ball is handed out to each person. Only one person out of the 100 people has the red ball. It might be you!

If you insist that the probabilities are still conditional, since you are sampling without replacement, look at it this way:

$$\text{Consider } P(\text{sixth person wins}) = \left(\frac{99}{100}\right)\left(\frac{98}{99}\right)\left(\frac{97}{98}\right)\left(\frac{96}{97}\right)\left(\frac{95}{96}\right)\left(\frac{1}{95}\right) = \frac{1}{100}$$