

1 Chapter 3

1. Since $P(A_1 \cap A_2) = 0.64 = (0.8)(0.8) = P(A_1)P(A_2)$, we can conclude that A_1 and A_2 are independent; hence, the answer is (c).
2. (a) $P(A \cap B) = P(B \cap A) = P(B|A) = (0.95)(0.05) = 0.0475$.
 (b) $P(B) = P(B \cap A) + P(B \cap A^c) = 0.0475 + (0.03)(1 - 0.05) = 0.076$.
 (c) $P(A|B) = P(A \cap B) / P(B) = 0.0475 / 0.076 = 0.625$.
3. Let A be the event that an adult gets the flu and let B be the event that an adult gets the flu shot.
 (a) $P(A \cap B) = P(A|B)P(B) = (0.1)(0.42) = 0.042$.
 (b) $P(A) = P(A \cap B) + P(A \cap B^c) = 0.042 + (0.7)(1 - 0.42) = 0.448$.
4. Let X denote the number of people who have asthma. Then $X \sim \text{Binomial}(n = 50, p = 0.2)$. (Think about why!)
 (a) $P(X = 19) = \binom{50}{19}(0.2)^{19}(0.8)^{50-19} \approx 0.001579$.
 (b) The standard deviation is $\sigma = \sqrt{np(1-p)} = \sqrt{(50)(0.2)(0.8)} = 2.828427$ and the mean/expected value is $\mu = np = (50)(0.2) = 10$. Hence, the z -score is $(19 - 10) / 2.828427 \approx 3.181981$ which implies that 19 is a little over three standard deviations above the mean.
 (c) Using the empirical rule, we have that $P(X \geq 19) \approx P(X \geq \mu + 3\sigma) \approx (1 - 0.997) / 2 = 0.003 / 2 = 0.0015$. (Draw a picture!). The exact answer is


```
1 - pbinom(18, size = 50, prob = 0.2)
## [1] 0.002511203
```
5. $E[X] = \sum_{r=1}^3 rP(X=r) = (1)(1/3) + (2)(1/3) + (3)(1/3) = 2$ envelopes.
6. (a) $E[X] = \mu = \sum_{r=0}^4 rP(X=r) = (0)(0.2) + (1)(0.3) + (2)(0.3) + (3)(0.1) + (4)(0.1) = 1.6$ egg masses.
 (b) $Var[X] = \sum_{r=0}^4 (r - \mu)^2 P(X=r) = (0 - 1.6)^2(0.2) + (1 - 1.6)^2(0.3) + (2 - 1.6)^2(0.3) + (3 - 1.6)^2(0.1) + (4 - 1.6)^2(0.1) = 1.44$. Hence, the standard deviation is $\sqrt{1.44} = 1.2$ egg masses.
7. Let A be the event that a subject is taking the drug (then A^c is the event that the subject is taking the placebo) and let B be the event that a subject improves.
 (a) $P(B \cap A) = P(B|A)P(A) = (0.6)(0.5) = 0.3$.
 (b) $P(B) = P(B \cap A) + P(B \cap A^c) = 0.3 + (0.35)(0.5) = 0.475$.

8. Let Y be a random variable denoting the number of chickens out of 20 with the bird flu. It then follows that $Y \sim \text{Binomial}(n = 20, p = 0.1)$.
- (a) $P(Y = 5) = \binom{20}{5}(0.1)^5(0.9)^{15} = (15504)(10^{-5})(0.2058911) \approx 0.031921$.
 - (b) $E[Y] = np = (20)(0.1) = 2$ chickens.
 - (c) $\sqrt{\text{Var}[Y]} = \sqrt{np(1-p)} = \sqrt{(20)(0.1)(0.9)} \approx 1.3416$ chickens.
9. Let X be a random variable denoting the number of frog eggs that hatch out of 100. Then, $X \sim \text{Binomial}(n = 100, p = 0.87)$. (Since the frog eggs hatch independently of each other!)
- (a) $P(X = 80) = \binom{100}{80}(0.87)^{80}(0.13)^{20} \approx 0.01477606$. You should be able to do this with a calculator, but in R we would just use

```
dbinom(80, size = 100, prob = 0.87)
## [1] 0.01477606
```

- (b) For a binomial random variable, $E[X] = np = (100)(0.87) = 87$ eggs.
- (c) To use the empirical rule, we must first calculate the standard deviation of X . The standard deviation of a binomial random variable is given by $\sqrt{\text{Var}[X]} = \sqrt{np(1-p)} = \sqrt{3.31} = 3.363034$. Computing the z -score yields

$$Z = \frac{77 - 87}{3.363034} = -2.973505 \approx -3.$$

In other words, 77 eggs is about three standard deviations below the mean. So, $P(X \leq 77) \approx$ the probability of being three or more standard deviations below the mean $\approx 0.003/2 = 0.0015$.

2 Chapter 4

1. No (it looks bimodal).
2. The best answer is (d); bimodal.
3. The population might consist of both males and females, and each of these subpopulations probably has its own mean.
4. The best answer is (a); 31.
5. The best answer is (c); 0.22.
6. The best answer is (e); 0.94.
7. It will remain the same. Go back to the properties about correlation and linear transformations!

8. The best answer is (d); 0.27.
9. The best answer is (d); 0.58.
10. Use the fact that $X \sim N(\mu = 5.28, \sigma = 0.4)$.

(a)

$$\begin{aligned}
 P(X > 5.4) &= 1 - P(X \leq 5.4) \\
 &= 1 - P\left(\frac{X - 5.28}{0.4} < \frac{5.4 - 5.28}{0.4}\right) \\
 &= 1 - P\left(Z < \frac{5.4 - 5.28}{0.4}\right) \\
 &= 1 - P(Z < 0.3) \\
 &= 1 - \Phi(0.3)
 \end{aligned}$$

In R, we get

```
1 - pnorm(0.3)
## [1] 0.3820886
```

(b)

$$\begin{aligned}
 P(5 < X < 6) &= P(X < 6) - P(X < 5) \\
 &= P\left(\frac{X - 5.28}{0.4} < \frac{6 - 5.28}{0.4}\right) - P\left(\frac{X - 5.28}{0.4} < \frac{5 - 5.28}{0.4}\right) \\
 &= P\left(Z < \frac{6 - 5.28}{0.4}\right) - P\left(Z < \frac{5 - 5.28}{0.4}\right) \\
 &= P(Z < 1.8) - P(Z < -0.7) \\
 &= \Phi(1.8) - \Phi(-0.7)
 \end{aligned}$$

In R, we get

```
pnorm(1.8) - pnorm(-0.7)
## [1] 0.722106
```

- (c) The general formula for the p -th percentile, denoted x_p , of a normal distribution with mean μ and standard deviation σ is

$$x_p = \mu + \sigma z_p,$$

where z_p is the p -th percentile of a standard normal distribution (which we can obtain using `qnorm(p)` in R). Hence, the 95-th percentile is $x_{0.95} = 5.28 + 0.4z_{0.95}$. Using `qnorm(0.95)` in R, we obtain $x_{0.95} = 5.28 + 0.4(1.644854) = 5.937941$.

- (d) Since the data are a random sample from a normal distribution, we know that the sample mean also has a normal distribution; in particular, $\bar{X} \sim N(\mu = 5.28, \sigma = 0.4/\sqrt{50})$. Hence,

$$\begin{aligned} P(\bar{X} > 5.4) &= 1 - P(\bar{X} \leq 5.4) \\ &= 1 - P\left(\frac{\bar{X} - 5.28}{0.4/\sqrt{50}} < \frac{5.4 - 5.28}{0.4/\sqrt{50}}\right) \\ &= 1 - P\left(Z < \frac{5.4 - 5.28}{0.4/\sqrt{50}}\right) \\ &= 1 - P(Z < 2.12132) \\ &= 1 - \Phi(2.12132) \end{aligned}$$

In R, we get

```
1 - pnorm(2.12132)
## [1] 0.01694744
```

11. Use the fact that $X \sim N(\mu = 170, \sigma = 20)$.

(a)

$$\begin{aligned} P(X > 200) &= 1 - P(X \leq 200) \\ &= 1 - P\left(\frac{X - 170}{20} < \frac{200 - 170}{20}\right) \\ &= 1 - P\left(Z < \frac{200 - 170}{20}\right) \\ &= 1 - P(Z < 1.5) \\ &= 1 - \Phi(1.5) \end{aligned}$$

In R, we get

```
1 - pnorm(1.5)
## [1] 0.0668072
```

- (b) Using the fact that $\bar{X} \sim N(\mu = 170, \sigma = 20/\sqrt{20})$, we get

$$\begin{aligned} P(\bar{X} > 200) &= 1 - P(\bar{X} \leq 200) \\ &= 1 - P\left(\frac{\bar{X} - 170}{20/\sqrt{20}} < \frac{200 - 170}{20/\sqrt{20}}\right) \\ &= 1 - P\left(Z < \frac{200 - 170}{20/\sqrt{20}}\right) \\ &= 1 - P(Z < 6.708204) \\ &= 1 - \Phi(6.708204) \end{aligned}$$

In R, we get

```
1 - pnorm(6.708204)
## [1] 9.851675e-12
```

(c) $x_{0.95} = \mu + \sigma z_{0.95} = 170 + 20 (1.644854) = 202.8971$ (mg/dL).