Chapter 13
Estimation with Confidence

13.1 **a.** A 95% confidence interval for μ is of the form $\bar{x} \pm z_{.975}\dfrac{\sigma}{\sqrt{n}} = 155.4 \pm 1.96 \cdot \dfrac{8}{\sqrt{9}}$
($z_{.975}$ can be found in the last row of Table 4).
So an approximate 95% confidence interval for μ is $150.2 \le \mu \le 160.6$.

b. A 99% confidence interval for μ is of the form $\bar{x} \pm z_{.995}\dfrac{\sigma}{\sqrt{n}} = 155.4 \pm 2.5758 \cdot \dfrac{8}{\sqrt{9}}$
($z_{.995}$ can be found in the last row of Table 4).
So an approximate 99% confidence interval for μ is $148.5 \le \mu \le 162.3$.

13.3 **a.** An interval that includes 95% of the population is of the form $\mu \pm z_{.025}\sigma$. If we estimate the population mean by the sample mean (replace μ by 200), then an interval that includes approximately 95% of the population is 121.60 through 278.40.

A second approach, which avoids estimating μ, is to imagine taking one more random sample X_{new} from this population. Then, using results from Section 10.4, we can compute the mean and standard deviation of $X_{new} - \overline{X}$ to be 0 and $\sigma\sqrt{1+\frac{1}{n}}$, respectively. Since $X_{new} - \overline{X}$ is a linear combination of independent normal random variables, it is normally distributed. Thus there is a 95% chance that $z_{.025} \le \dfrac{X_{new} - \overline{X}}{\sigma\sqrt{1+\frac{1}{n}}} \le z_{.975}$. So there is a 95% chance that $\overline{X} + z_{.025}\sigma\sqrt{1+\frac{1}{n}} \le X_{new} \le \overline{X} + z_{.975}\sigma\sqrt{1+\frac{1}{n}}$, and this results in the slightly longer interval 121.21 through 278.79.

b. A 95% confidence interval for μ is of the form $\bar{x} \pm z_{.975}\sigma_{\overline{X}}$. Substituting the known values into this formula gives $200 \pm 1.96 \cdot \dfrac{40}{\sqrt{100}}$. So a 95% confidence interval for μ is $192.2 \le \mu \le 207.8$.

c. Assuming the 50 represent a random sample from the population, their total weight has mean 50μ and standard deviation $\sqrt{50}\sigma$. The probability that the plane will be overloaded is $Pr\left(\sum_{i=1}^{50} X_i \ge 10900\right) = 1 - \Phi\left(\dfrac{10900 - 50\cdot\mu}{\sqrt{50}\cdot\sigma}\right)$ where Φ denotes the standard normal cumulative distribution function (as tabulated in Table 3). Since we are 95% confident that μ is between 192.16 and 207.84 (computed in part b), we are also 95% confident that the probability the 50 passengers will overload the plane is between $1 - \Phi\left(\dfrac{10900 - 50\cdot\mu_2}{\sqrt{50}\cdot\sigma}\right)$ and $1 - \Phi\left(\dfrac{10900 - 50\cdot\mu_1}{\sqrt{50}\cdot\sigma}\right)$, where μ_1 and μ_2 are the end points of the confidence interval for μ .

Recall σ = 40, so the interval is $1-\Phi(4.568)$ and $1-\Phi(1.796)$, or 0.000002 to 0.036243.

A point estimate for the chance of over load can be found by replacing μ by 200 in the formula, giving $Pr\left(\sum_{i=1}^{50} X_i \geq 10900\right) = 1-\Phi\left(\frac{10900-50\cdot\mu}{\sqrt{50}\cdot\sigma}\right) \approx 1-\Phi(3.182) \approx 0.000731$.

13.5 **a.** Given this data, $n=16, \bar{x}=18.3, \sigma=1.5,$ and $z_{.95}=1.6449$, conclude that 95% of the time the sample mean would be no more than $z_{.95}=1.6449$ standard errors above the population mean (you may use the last row of Table 4 to verify that $z_{.95}=1.6449$). So a 95%CI: $\mu \geq \bar{x}-z_{.95}\frac{\sigma}{\sqrt{n}}$ and upon substitution find $\mu \geq 18.3-1.6449\frac{1.5}{\sqrt{16}}$, so $\mu \geq 17.68$.

b. A one-sided 99% confidence interval for μ is of the form $\mu \geq \bar{x}-z_{.99}\frac{\sigma}{\sqrt{n}}$ (you may use the last row of Table 4 to verify that $z_{.99}=2.3263$). Upon substitution find

$\mu \geq 18.3-2.3263\frac{1.5}{\sqrt{16}}$, so $\mu \geq 17.43$.

13.7 **a.** Given this data, $n=100, \bar{x}=28.2, s=9.1,$ and $z_{.975}=1.96$, conclude that an approximate 95% confidence interval for mean age is of the form $\bar{x}\pm z_{.975}\frac{\sigma}{\sqrt{n}}$. Since the sample size is large, we can approximate σ by the sample standard deviation s. Upon substitution find $28.2\pm 1.96\cdot\frac{9.1}{\sqrt{100}}$, or 28.2 ± 1.78.

So an approximate 95% confidence interval for μ is 95%CI: 26.42 years $\leq \mu \leq 29.98$ years.

b. An approximate one-sided 90% confidence interval for mean age is of the form $\mu \leq \bar{x}-z_{.05}\frac{\sigma}{\sqrt{n}}$ so $\mu \leq 28.2+1.6449\frac{9.1}{\sqrt{100}}$. So an approximate one-sided 90% confidence interval for μ is 99%CI: $\mu \leq 29.7$ years.

13.9 **a.** This exercise is similar to Exercise 13.1 a. Here $n=70,\ \bar{x}=87,\ s=0.5,$ and $z_{.975}=1.96$ (or $t_{69,.975}=1.96$). The limits for a 95% confidence interval for μ are $\bar{x}\pm z_{.975}\frac{\sigma}{\sqrt{n}}$. Approximating σ by s gives $87\pm 1.96\cdot\frac{0.5}{\sqrt{70}}$, so an approximate 95% confidence interval for μ is $86.88 \leq \mu \leq 87.12$. In Section 13.4 you will learn that when the population roughly follows a normal curve, and the sample size is moderate, then a slightly better approximation is to replace $z_{.975}$ by $t_{69,.975}=1.9949$.

b. This exercise is similar to Exercise 13.8a:. Here $n=70,\ \bar{x}=87,\ s=0.5,$ and $z_{.99}=2.3264$ (or $t_{69,.99}=2.3816$). The limits for a 98%

confidence interval for μ are $\bar{x} \pm z_{.99}\dfrac{\sigma}{\sqrt{n}}$. Estimating σ by s gives $87 \pm 2.3264 \cdot \dfrac{0.5}{\sqrt{70}}$, so an approximate 98% confidence interval for μ is $86.86 \le \mu \le 87.14$. In Section 13.4 you will learn that when the population roughly follows a normal curve, and the sample size is moderate, then a slightly better approximation is to replace $z_{.975}$ by $t_{69,.975} = 1.9949$.

c. It is not too important that we do not know the exact shape of the population because the sample size, n = 70, is not small (specifically $n \ge 30$), so $T = \sqrt{n}\,\dfrac{\bar{X} - \mu}{s}$ will be approximately normally distributed, unless the parent distribution is badly skewed or has extremely thick tails. (It is a good idea to plot the data whenever it is available.) We are also assuming that we are sampling at random with replacement (or that the population size is large relative to the sample size).

13.11 **a.** Here $n = 4, \bar{x} = 85, s = 7.7,$ and $t_{3,.975} = 3.1824$. The limits for a 95% confidence interval for μ are $\bar{x} \pm t_{3,.975}\dfrac{s}{\sqrt{n}}$ = $85 \pm 3.1824 \cdot \dfrac{7.7}{\sqrt{4}}$. A 95% confidence interval for μ is $72.75 \le \mu \le 97.25$.

b. The limits for a 99% confidence interval for μ are $\bar{x} \pm t_{3,.995}\dfrac{s}{\sqrt{n}}$ and approximating σ by s gives $85 \pm 5.8408 \cdot \dfrac{7.7}{\sqrt{4}}$.

An approximate 99% confidence interval for μ is $62.52 \le \mu \le 107.49$.

13.13 Here $n = 12$, $\bar{x} = 150$ mg/dL, $s = 25$ mg/dL, and $t_{11,.95} = 1.7959$. Assume the population is approximately normally distributed. The limits for a 90% confidence interval for μ are $\bar{x} \pm t_{11,.95}\dfrac{s}{\sqrt{n}}$ gives $150 \pm 1.7959 \cdot \dfrac{25}{\sqrt{12}}$, so a 90% confidence interval for the mean serum-cholesterol level μ is $137 \text{ mg/dL} \le \mu \le 163 \text{ mg/dL}$.

13.15 Assume the 10 burn times are a random sample.

a. Here $n = 10$, $\bar{x} = 924.8$ hrs, $s = 136.6$ hrs, and $t_{9,.95} = 1.8331$.

b. The limit for a one-sided 90% confidence interval for μ are $\mu \le \bar{x} + t_{9,.95}\dfrac{s}{\sqrt{n}}$ = $924.8 + 1.8331 \cdot \dfrac{136.6}{\sqrt{10}}$, so a 90% confidence interval for the mean burn time μ is $\mu \le 1003.98$ hrs.

c. Since 1000 is in the 95% confidence interval, we cannot reject the manufacturer's claim.

13.19 One of the common rules, based on the normal curve, asserts that about 95% of the population is within 2 standard deviations of its mean. The length of this interval is 4 standard deviations. That is why the divisor 4 is used. A 99% confidence interval for the population mean is of the form $\bar{x} \pm z_{0.995} \dfrac{\sigma}{\sqrt{n}}$. Substituting in the values for $z_{0.995}$ and σ then solving for n shows n must be at least $(2.5758 \cdot 12.5)^2 = 1037$.

13.21 This exercise is similar to the previous exercise. Using the values $z_{0.995} = 2.5758$ and $\sigma = 9$, and then solving for the sample size n, gives $n \geq \left(\dfrac{2.5758 \cdot 15}{2}\right)^2 \approx 374$.

13.23 **a.** A conservative-method, one-sided 95% confidence interval for the population proportion is of the form $\pi \geq p - \dfrac{z_{0.95}}{2\sqrt{n}}$. Here p = 66/150, n = 150, and $z_{0.95} = 1.6448$ (you can find $z_{0.95}$ in the last row of Table 4). So $\pi \geq 0.3728$.

b. A bootstrap-method, one-sided 95% confidence interval for the population proportion is of the form $\pi \geq p - z_{0.95}\sqrt{\dfrac{p(1-p)}{n}}$. Again n = 150, p = 66/150, and $z_{0.95} = 1.6448$. So $\pi \geq 0.3733$.

13.25 The end points of a bootstrap-method, two-sided 99% confidence interval for the long-run percentage of heads are of the form $100p \pm 100 z_{0.995}\sqrt{\dfrac{p(1-p)}{n}}$. Here n = 24000, p = 12012/24000 = 0.5005, and $z_{0.995} = 2.5758$. So the end points are 50.05% ± 0.83% and the interval is from 49.22% to 50.88%.

13.31 This is similar to Exercise 13.30. Here n = 100 and k = 7. Suppose K has a binomial distribution with parameters n = 100 and π. To find a 99% confidence interval for π, solve the equation $\Pr(K \leq 7) = 0.005$ for π (find $\pi = 0.163$). Then solve the equation $\Pr(K \geq 7) = 0.005$ for π (find $\pi = 0.021$). So a 99% confidence interval for the percentage of women who are shorter than 5 feet is 2.1% to 16.3%.

13.33 **a.** This is similar to Exercise 13.30. Here n = 40, p = 0.10, and k = 4. Suppose K has a binomial distribution with parameters n = 40 and π. To find a 95% confidence interval for π, solve the equation $\Pr(K \leq 4) = 0.025$ for the upper bound π (find $\pi = 0.2366$). Then solve the equation $\Pr(K \geq 4) = 0.025$ for the lower bound π (find $\pi = 0.0279$). So a 95% confidence interval for the percentage of pacemaker users with some cell-phone interference is 2.8% to 23.7%.

b. This is similar to Exercise 13.32. The end points of a bootstrap-method, two-sided 95% confidence interval for the long-run percentage of interference are of the form

$100p \pm 100z_{0.975}\sqrt{\frac{p(1-p)}{n}}$. Here n = 40, p = 0.10, and k = 4, and $z_{0.975}$ = 1.9600. So the bootstrap, normal theory, 95% confidence interval for the percentage of users who suffer some interference is 0.7% to 19.3%. It is symmetric about the observed 10% interference rate.

13.35 This is similar to Exercise 13.30. Here n = 34, p = 14/34, and k = 14. Suppose K has a binomial distribution with parameters n = 34 and π. To find a 95% confidence interval for π, solve the equation $\Pr(K \le 14) = 0.025$ for the upper bound π (find π = 0.5930). Then solve the equation $\Pr(K \ge 14) = 0.025$ for the lower bound π (find π = 0.2465). So a 95% confidence interval for the proportion of all such subjects who would show improvement of symptoms under the treatment versus the control is 0.246 to 0.593.

13.39 **a.** Each reading was about <u>12.37</u>g/l give or take <u>0.21</u>g/l or so.

b. The true concentration of copper is estimated as <u>12.37</u>g/l; this estimate is likely to be off by <u>0.03</u>g/l. (Because the standard error for the sample average as an estimate of the true mean is $SE_{\bar{x}} = \frac{\sigma}{\sqrt{n}} \approx \frac{0.21}{\sqrt{49}} = 0.03$)

c. If another reading were taken it would be about <u>12.37</u>g/l give or take <u>0.21</u>g/l.or so.

d. The 95% confidence interval for the true concentration is <u>12.37</u>g/l ± <u>0.06</u>g/l (using either $\pm z_{0.975}$ standard errors or $\pm t_{48,0.975}$ standard errors)

13.41 An approximate 98% two-sided confidence interval for the mean length of catchable bass at Lake Wobegon is of the form $\bar{x} \pm z_{0.99}\frac{s}{\sqrt{n}}$. Here n = 100, $\bar{x} = 35.5cm$, and s = 5.0*cm,* and $z_{0.99}$ = 2.32635. So $34.3cm \le \mu \le 36.7cm.$

13.43 **a.** The standard deviation of the sample mean is $\sigma_{\bar{X}} = \frac{\sigma}{\sqrt{n}} = \frac{50}{\sqrt{9}}$ = 16.6667 micrograms.

b. The 68% confidence interval for the true weight is of the form $\bar{x} \pm z_{0.84}\frac{\sigma}{\sqrt{n}}$. Here $\bar{x} = 0.100$ grams, σ = 50 micrograms, and $z_{0.84}$ = 0.9945. So the interval is 0.100000 grams ±16.57 micrograms or 0.099983g ≤ μ ≤ 1.00017g.

c. The 95% confidence interval for the true weight is of the form $\bar{x} \pm t_{8,0.975}\frac{s}{\sqrt{n}}$. Here $\bar{x} = 0.100$ grams, s = 100 micrograms, and $t_{8,0.975}$ = 2.3060. So the interval is 0.100000 grams ±76.87 micrograms or 0.099923g ≤ μ ≤ 1.00077g.

d. Note: X – μ has a normal distribution with mean 0 and standard deviation 50 micrograms. So $\Pr(|x-\mu| \le 64) = \Pr\left((|Z| \le \frac{64}{50})\right) = \Pr(|Z| \le 1.28) = 0.79945$.

e. The probability all three result in success is $\pi^3 = (0.79945)^3 = 0.51095$.

f. Yes, K is a binomial random variable with parameters n = 10 and π. If the null hypothesis is true, then K has mean $\mu = n\pi = 9$ and variance $n\pi(1-\pi) = 0.9$.

g. Type 1 error is when a true null hypothesis is rejected. In this case the probability of a Type 1 error is $\Pr(K < 7 \mid n = 10, \pi = 0.9) = 1 - \sum_{k=7}^{10} \binom{10}{k} 0.9^k 0.1^{10-k} = 0.0128$ and the Type II error probability is $\Pr(K \geq 7 \mid n = 10, \pi = 0.8) = \sum_{k=7}^{10} \binom{10}{k} 0.8^k 0.2^{10-k} = 0.8791$.

13.47 For this data, n = 34, $\bar{x} = 0.817$, s = 1.126 and $t_{33,0.95} = 1.6924$. The 90% confidence interval for the mean change in BMI is of the form $\bar{x} \pm t_{33,0.975} \frac{s}{\sqrt{n}}$ or 0.49 to 1.14 BMI units. The population size is large compared to the sample size, so the finite population correction of the standard error is not needed here. However, a box-plot revealed that the data is skewed to the right with 2 outliers.

13.49 The bootstrap-method, one-sided 99% confidence interval for the percentage of Hispanic boys in the ECLS study whose BMI is at or above the 95th percentile is of the form $\pi \geq p + z_{0.01}\sqrt{\frac{p(1-p)}{n}}$. Here n = 59, p = 11/59, and k = 11, and $z_{0.01} = -2.3263$. So this 99% confidence is $\pi \geq 6.8\%$. The small-sample 99% one-sided confidence interval is found by solving the equation $\Pr(K \geq 11 \mid n = 59, \pi) = 0.01$ for π. This gives $\pi \geq 8.5\%$.

13.51 For this data, n = 2500, $\bar{x} = 35$ minutes, s = 25 minutes, and $z_{0.975} = 1.9600$. The 95% confidence interval for the mean head-of-household commute time is of the form $\bar{x} \pm z_{0.975} \frac{s}{\sqrt{n}}$ or 34.0 to 36.0 minutes.

13.53 **a.** The true acceleration due to gravity is estimated as <u>978.721</u> cm/sec^2; this estimate is likely to be off by <u>0.003</u> cm/sec^2 or so.
b. False. The mean of the 225 readings *is* 978.721 cm/sec^2.
c. True.
d. False.
e. False.
f. False. However, for about 95% of samples of size 225, intervals of the form $\bar{x} \pm 2\frac{s}{\sqrt{225}}$ will include μ.