CHAPTER 8

Section 8.1

1.

a. $z_{\alpha/2} = 2.81$ implies that $\alpha/2 = 1-\Phi(2.81) = .0025$, so $\alpha = .005$ and the confidence level is $100(1-\alpha)\% = 99.5\%$.

b. $z_{\alpha/2} = 1.44$ for $\alpha = 2\left[1-\Phi(1.44)\right] = .15$, and $100(1-\alpha)\% = 85\%$.

c. 99.7% implies that $\alpha = .003$, $\alpha/2 = .0015$, and $z_{.0015} = 2.96$. (Look for cumulative area .9985 in the main body of table A.3, the Z table.)

d. 75% implies $\alpha = .25$, $\alpha/2 = .125$, and $z_{.125} = 1.15$.

3.

a. A 90% confidence interval will be narrower. The *z* critical value for a 90% confidence level is 1.645, smaller than the *z* of 1.96 for the 95% confidence level, thus producing a narrower interval.

b. Not a correct statement. Once and interval has been created from a sample, the mean μ is either enclosed by it, or not. None of μ, 7.8, or 9.4 is random. Instead, we have 95% confidence in the general procedure, under repeated and independent sampling.

c. Not a correct statement. The interval is an estimate for the population *mean*, not a boundary for individual population values.

d. Not a correct statement, but close. In theory, if the process were repeated an infinite number of times, 95% of the intervals would contain the population mean μ. We *expect* 95 out of 100 intervals will contain μ, but it won't necessarily be the case that exactly 95 out of 100 intervals will include μ.

5.

a. $4.85 \pm \dfrac{(1.96)(.75)}{\sqrt{20}} = 4.85 \pm .33 =$ (4.52, 5.18).

b. $z_{\alpha/2} = z_{.02/2} = z_{.01} = 2.33$, so the interval is $4.56 \pm \dfrac{(2.33)(.75)}{\sqrt{16}} =$ (4.12, 5.00).

c. $n = \left[\dfrac{(1.96)(.75)}{.40/2}\right]^2 = 54.02$, so n = 55.

d. $w/2 = .2 \rightarrow n = \left[\dfrac{(2.58)(.75)}{.2}\right]^2 = 93.61$, so n = 94.

7. CI width = $L = 2z_{\alpha/2}\dfrac{\sigma}{\sqrt{n}}$. Dividing this by two (halving the width) requires quadrupling the sample size: $L \div 2 = 2z_{\alpha/2}\dfrac{\sigma}{\sqrt{n}} \div 2 = 2z_{\alpha/2}\dfrac{\sigma}{2\sqrt{n}} = 2z_{\alpha/2}\dfrac{\sigma}{\sqrt{4n}}$. If $n' = 25n$, then CI width becomes $2z_{\alpha/2}\dfrac{\sigma}{\sqrt{25n}} = 2z_{\alpha/2}\dfrac{\sigma}{5\sqrt{n}} = 2z_{\alpha/2}\dfrac{\sigma}{\sqrt{n}} \div 5 = L \div 5$, so the width is decreased by a factor of 5.

9.

a. Imitating Example 8.7, we may write $P\left(\dfrac{\bar{X}-\mu}{\sigma/\sqrt{n}} < z_\alpha\right) = 1-\alpha$. Solve the inequality inside parentheses for μ:
$\dfrac{\bar{X}-\mu}{\sigma/\sqrt{n}} < z_\alpha \Rightarrow \bar{X}-\mu < z_\alpha\sigma/\sqrt{n} \Rightarrow \mu-\bar{X} > -z_\alpha\sigma/\sqrt{n} \Rightarrow \mu > \bar{X}-z_\alpha\sigma/\sqrt{n}$. Thus, a lower confidence bound for μ with confidence level 100(1 – α)% is $\bar{x}-z_\alpha\sigma/\sqrt{n}$.

b. $\bar{x}-z_{.005}\sigma/\sqrt{n} = 4.85-2.576(.75)/\sqrt{20} = 4.85-.432 = 4.418$.

c. $\bar{x}+z_\alpha\dfrac{\sigma}{\sqrt{n}}$; from 4**a**, $\bar{x}=58.3$, $\sigma=3.0$, $n=25$ $\Rightarrow 58.3+2.33\dfrac{3}{\sqrt{25}} = 59.70$.

11. Y is a binomial rv with n = 1000 and p = .95. Using software, $P(940 \le Y \le 960)$ = .8731. Or, using the normal approximation with $E(Y) = np = 950$, the expected number of intervals that capture μ, and $V(Y) = npq = 47.5$, $P(940 \le Y \le 960) = P(939.5 \le Y \le 960.5) \approx P(-1.52 \le Z \le 1.52)$ = .9357 – .0643 = .8714.

Section 8.2

13.

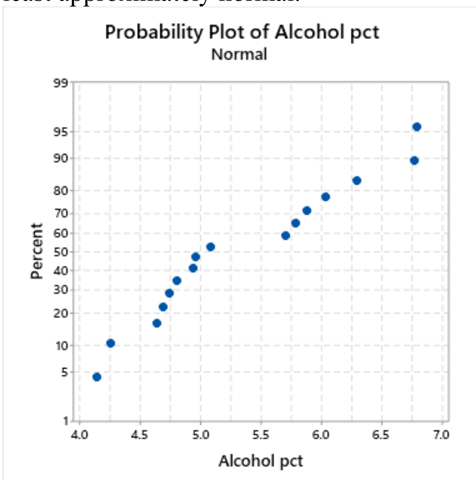
a. 1.341
b. 1.753
c. 1.708
d. 1.684
e. 2.704

15.

a. $t_{.025,10} = 2.228$
b. $t_{.025,15} = 2.131$
c. $t_{.005,15} = 2.947$
d. $t_{.005,4} = 4.604$
e. $t_{.01,24} = 2.492$
f. $t_{.005,37} \approx 2.712$

17.

a. The probability plot below is sufficiently straight to assume the alcohol percentage distribution is at least approximately normal.



b. From the sample, n = 16, $\bar{x}$ = 5.34, and s = .8483, so the 95% CI is

$$\bar{x} \pm t_{.025,15}\frac{s}{\sqrt{n}} = 5.34 \pm (2.131)\frac{.8483}{\sqrt{16}} = 5.34 \pm .45 = (4.89, 5.79).$$

c. Multiply the above percents (in decimal form) by 12 fl oz: .0489(12) = .5868 fl oz to .0579(12) = .6948 fl oz.

19.

a. From the sample data, n = 77, $\bar{x}$ = 64.887, and s = 7.803, from which the 95% CI is

$$\bar{x} \pm t_{.025,77-1}\frac{s}{\sqrt{n}} = 64.887 \pm (1.992)\frac{7.803}{\sqrt{77}} = (63.12, 66.66).$$

b. No. A normal probability plot of the noise level data shows a stark deviation from normality. While the *confidence* interval in **a** is still valid for non-normal data with n = 77, the *prediction* interval is only legitimate for "normal data" no matter the sample size.

21.

a. $\bar{x} \pm t_{.025,44-1}\frac{s}{\sqrt{n}} = 35.02 \pm (2.017)\frac{18.94}{\sqrt{44}} = 35.02 \pm 5.76 = (29.26, 40.78)$. At the 95% confidence level, the true average time to change jobs for this population is between 29.26 and 40.78 months.

b. $\bar{x} \pm t_{.025,44-1}s\sqrt{1+\frac{1}{n}} = 35.02 \pm (2.017)(18.94)\sqrt{1+\frac{1}{44}} = 35.02 \pm 38.63 = (-3.61, 73.65)$. This interval assumes the population distribution of job-change times is normal, which is not realistic here. First, the interval implies negative times are possible, which is not true. Second, the fact that the mean is less than twice the sd implies that the data is likely not consistent with a normal distribution — we'd expect data from a normal distribution to extend 3 sd's in either direction from the mean.

23. From the sample, $n = 10$, $\bar{x} = 21.90$, $s = 4.134$, and $t_{.025,10-1} = 2.262$.

a. $21.90 \pm 2.262\dfrac{4.134}{\sqrt{10}} = 21.90 \pm 2.96 = (18.94, 24.86)$.

b. All other things being equal, a 90% CI would be narrower than a 95% CI.

c. All other things being equal, a larger sample size results in a narrower CI.

d. $21.90 \pm 2.262(4.134)\sqrt{1+\dfrac{1}{10}} = 21.90 \pm 9.81 = (12.09, 31.71)$.

25. $n = 14$, $\bar{x} = 8.48$, $s = .79$; $t_{.05,13} = 1.771$

a. A 95% lower confidence bound: $8.48 - 1.771\left(\dfrac{.79}{\sqrt{14}}\right) = 8.48 - .37 = 8.11$. With 95% confidence, the value of the true mean proportional limit stress of all such joints is greater than 8.11 MPa. We must assume that the sample observations were taken from a normally distributed population.

b. A 95% lower prediction bound: $8.48 - 1.771(.79)\sqrt{1+\dfrac{1}{14}} = 8.48 - 1.45 = 7.03$. If this bound is calculated for sample after sample, in the long run 95% of these bounds will provide a lower bound for the future value of the proportional limit stress of a single joint of this type.

27. $n = 26$, $\bar{x} = 370.69$, $s = 24.36$; $t_{.05,25} = 1.708$.

a. A 95% upper confidence bound: $370.69 + (1.708)\left(\dfrac{24.36}{\sqrt{26}}\right) = 370.69 + 8.16 = 378.85$.

b. A 95% upper prediction bound: $370.69 + 1.708(24.36)\sqrt{1+\dfrac{1}{26}} = 370.69 + 42.45 = 413.14$.

c. Following a similar argument as in the text, we need to find the variance of $\bar{X} - \bar{X}_{new}$:

$$\begin{aligned} V\left(\bar{X} - \bar{X}_{new}\right) &= V\left(\bar{X}\right) + V\left(\bar{X}_{new}\right) = V\left(\bar{X}\right) + V\left(\tfrac{1}{2}\left(X_{27} + X_{28}\right)\right) \\ &= V\left(\bar{X}\right) + V\left(\tfrac{1}{2}X_{27}\right) + V\left(\tfrac{1}{2}X_{28}\right) = V\left(\bar{X}\right) + \tfrac{1}{4}V\left(X_{27}\right) + \tfrac{1}{4}V\left(X_{28}\right) \\ &= \frac{\sigma^2}{n} + \frac{1}{4}\sigma^2 + \frac{1}{4}\sigma^2 = \sigma^2\left(\frac{1}{2} + \frac{1}{n}\right) \end{aligned}$$

We eventually arrive at $T = \dfrac{\bar{X} - \bar{X}_{new}}{S\sqrt{\frac{1}{2}+\frac{1}{n}}} \sim t_{n-1}$ distribution, so the new prediction interval is $\bar{x} \pm t_{\alpha/2,n-1} \cdot s\sqrt{\frac{1}{2}+\frac{1}{n}}$. For this situation, we have

$$370.69 \pm 1.708(24.36)\sqrt{\frac{1}{2}+\frac{1}{26}} = 370.69 \pm 30.53 = (340.16, 401.22).$$

29. Using the values from Example 8.8, a PI is $53,907.2 \pm 2.131(28,287.2)\sqrt{1+\frac{1}{16}} = 53,907.2 \pm 62135.2 =$ (–8228.0, 116,042.4). The value 19,815 does fall in the PI and thus is consistent with the prediction; however, it's worth noting that the interval includes (impossible) negative values for the prediction, which casts some doubt on its reliability.

31. For this data, $\bar{x}$ = 174.37, s = 19.89, n = 63.

a. $174.37 \pm t_{.025,62}(19.89)/\sqrt{63}$ = (169.36,179.37).

b. $174.37 \pm t_{.025,62}(19.89)\sqrt{1+\frac{1}{63}}$ = (134.30, 214.43), which includes 152.

c. The second interval is much wider, because it allows for the variability of a single observation.

d. A normal probability plot gives no reason to doubt normality. This is especially important for part (b), but the large sample size implies that normality is not so critical for (a).

33. $n = 35$, $\bar{x} = 19.257$, $s = 1.832$

a. $\bar{x} \pm t_{.005,35-1}\dfrac{s}{\sqrt{n}} = 19.257 \pm 2.728\dfrac{1.832}{\sqrt{35}}$ = (18.413, 20.102). At the 99% confidence level, the true mean distilled alcohol content for port wines is between 18.413% and 20.102%.

b. $\bar{x} - t_{.10,35-1}\dfrac{s}{\sqrt{n}} = 19.257 - 1.307\dfrac{1.832}{\sqrt{35}}$ = 18.852. At the 90% confidence level, the true mean distilled alcohol content for port wines is greater than 18.852%.

c. A normal probability plot of the data suggests the population distribution is *not* normal, so the PI formula is not valid.

35. $\bar{x} \pm t_{.10/2,539-1}\dfrac{s}{\sqrt{n}} = \bar{x} \pm z_{.05}\dfrac{s}{\sqrt{n}} = 19 \pm 1.645\dfrac{10}{\sqrt{539}}$ = 19 ± .7 = (18.3, 19.7) days.

37. A 95% lower confidence bound for μ is $\bar{x} - z_{.05}\dfrac{s}{\sqrt{n}} = 2.08 - 1.645\dfrac{7.88}{\sqrt{41}}$ = 0.056. We're 95% confident that $\mu > 0.056$; this suggests that the mean population stress change is positive, though potentially by a very small amount. (Notice that if we increased the confidence level to 99%, the lower bound would be negative and we could no longer say with confidence that μ is positive.)

39. Using $\sigma \approx 15$, $z_{.025} = 1.96$, and $w/2 = 3$, $n = \left(\dfrac{(1.96)(15)}{3}\right)^2$ = 96.04. Round up to 97.

41.

a. $z_\alpha = .84$, and $\Phi(.84) = .7995 \approx .80$, so the confidence level is 80%.

b. $z_\alpha = 2.05$, and $\Phi(2.05) = .9798 \approx .98$, so the confidence level is 98%.

c. $z_\alpha = .67$, and $\Phi(.67) = .7486 \approx .75$, so the confidence level is 75%.

Section 8.3

43. $n = 1044$, $\hat{p} = 859/1044 = .8228$, $z = 1.96$: $\dfrac{\hat{p} + z^2/(2n) \pm z\sqrt{\hat{p}(1-\hat{p})/n + z^2/(4n^2)}}{1 + z^2/n} =$

$$\frac{.8228 + 1.96^2/(2(1044)) \pm 1.96\sqrt{.8228(1-.8228)/1044 + 1.96^2/(4(1044)^2)}}{1 + 1.96^2/1044} = (.798, 845).$$

45.

a. $n = 1018$, $\hat{p} = .53$, $z = z_{.05} = 1.645$: $\dfrac{\hat{p} + z^2/(2n) - z\sqrt{\hat{p}(1-\hat{p})/n + z^2/(4n^2)}}{1 + z^2/n} =$

$\dfrac{.53 + 1.645^2/(2(1018)) - 1.645\sqrt{.53(1-.53)/1018 + 1.645^2/(4(1018)^2)}}{1 + 1.645^2/1018} = .504$. At the 95% confidence level, the proportion of all U.S. adults that support the government sponsoring a manned mission to Mars is greater than .504.

b. Yes: since the entire interval exceeds .5, we may infer that $p > .5$, i.e., more than half (a majority) of all U.S. adults feel this way.

47. $n = 2020$, $\hat{p} = .57$, $z = z_{.10} = 1.28$: $\dfrac{\hat{p} + z^2/(2n) + z\sqrt{\hat{p}(1-\hat{p})/n + z^2/(4n^2)}}{1 + z^2/n} =$

$\dfrac{.57 + 1.28^2/(2(2020)) + 1.28\sqrt{.57(1-.57)/2020 + 1.28^2/(4(2020)^2)}}{1 + 1.28^2/2020} = .584$. At the 90% confidence level, the proportion of all U.S. adults that shop online at work is less than .584.

49. $n = 365$, $\hat{p} = 201/356 = .5646$, $z = 1.96$: $\dfrac{\hat{p} + z^2/(2n) \pm z\sqrt{\hat{p}(1-\hat{p})/n + z^2/(4n^2)}}{1 + z^2/n} =$

$$\frac{.5646 + 1.96^2/(2(356)) \pm 1.96\sqrt{.5646(1-.5646)/356 + 1.96^2/(4(356)^2)}}{1 + 1.96^2/356} = (.513, .615).$$

51. $n = 1515$, $\hat{p} = .42$, $z = z_{.05} = 1.645$: $\dfrac{\hat{p}+z^2/(2n)+z\sqrt{\hat{p}(1-\hat{p})/n+z^2/(4n^2)}}{1+z^2/n} =$

$\dfrac{.42+1.645^2/(2(1515))+1.645\sqrt{.42(1-.42)/1515+1.645^2/(4(1515)^2)}}{1+1.645^2/1515} = .441$. The bound indicates that less than 44.1% of all U.S. adults favor arming teachers. So, yes, we can be confident that a majority of the population does not favor this policy.

53.

a. $n = \dfrac{(1.96)^2\left[2(.5)(.5)-(.10)^2+\sqrt{(2(.5)(.5))^2+(1-4(.5)(.5))(.10)^2}\right]}{(.10)^2} \approx 381$.

b. $n = \dfrac{(1.96)^2\left[2(2/3)(1/3)-(.10)^2+\sqrt{(2(2/3)(1/3))^2+(1-4(2/3)(1/3))(.10)^2}\right]}{(.10)^2} \approx 339$.

55. $n = 57$, $\hat{p} = 4/57 = .070$, $z = 1.96$: $\dfrac{\hat{p}+z^2/(2n)\pm z\sqrt{\hat{p}(1-\hat{p})/n+z^2/(4n^2)}}{1+z^2/n} =$

$\dfrac{.07+1.96^2/(2(57))\pm 1.96\sqrt{.07(1-.07)/57+1.96^2/(4(57)^2)}}{1+1.96^2/57} = (.028, .167)$.

Section 8.4

57.

a. $\chi^2_{.1,15} = 22.307$ (.1 column, 15 df row)

b. $\chi^2_{.1,25} = 34.381$

c. $\chi^2_{.01,25} = 44.313$

d. $\chi^2_{.005,25} = 46.925$

e. $\chi^2_{.99,25} = 11.523$ (from .99 column, 25 df row)

f. $\chi^2_{.995,25} = 10.519$

59. A 90% CI for σ^2 is $\left(\dfrac{n-1}{\chi^2_{.05,n-1}}s^2, \dfrac{n-1}{\chi^2_{.95,n-1}}s^2\right) = \left(\dfrac{16-1}{24.996}(.8483)^2, \dfrac{16-1}{7.261}(.8483)^2\right) = (0.4318, 1.4866)$. Then, a 90% CI for σ is $(\sqrt{0.4318}, \sqrt{1.4866}) = (.657, 1.219)$.

61. For these data, $\bar{x}$ = 69.4, s = 3.3, n = 12.

a. A normal probability plot verifies that the data is plausibly normal.

b. At 12 – 1 = 11df, the chi-square critical values are 3.816 and 21.920. Hence, a 95% CI for σ is $\left(\sqrt{\frac{11}{21.920}}(3.3), \sqrt{\frac{11}{3.816}}(3.3)\right)$ = (2.34, 5.60). We are 95% confident the standard deviation for performance times of Beethoven's 9th for all conductors is between 1.65 minutes and 5.60 minutes.

c. False, though we didn't need the CI to confirm this. There's obviously variability in the performance times of Beethoven's 9th.

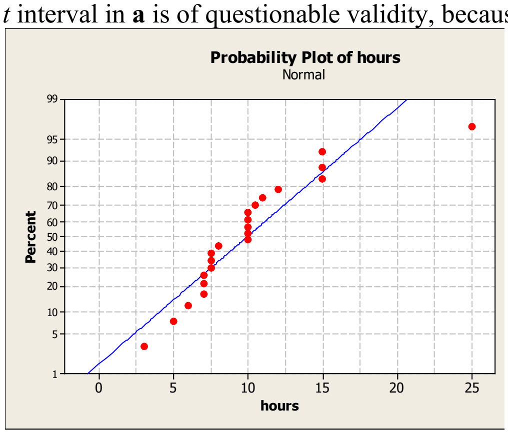
Section 8.5

Note: Most of the answers in this section will differ slightly from students' answers due to simulation variability.

63. For this data, $\bar{x}$ = 9.955, s = 4.603, n = 22.

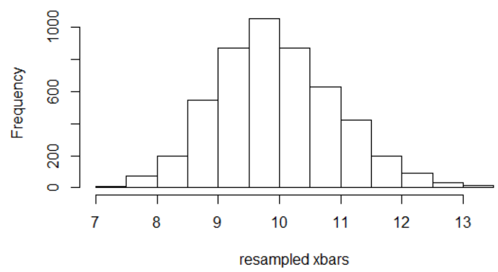
a. $t_{.025,21}$ = 2.080, so a 95% CI for μ is 9.955 ± 2.080(4.603)/ $\sqrt{22}$ = (7.91, 12.00).

b. The probability plot below shows gross deviation from a straight line, suggesting the data is highly non-normal. Our use of a t interval in **a** is of questionable validity, because n = 22 is not large.



c. The accompanying R code assumes the data has been read in as the vector x.

```
N=5000
xbar=rep(0,N)
for (i in 1:N){
  resample = sample(x,length(x),replace = T)
  xbar[i]=mean(resample)
}
```



d. From R, the standard deviation of the 5000 $\bar{x}^*$ values is s_{boot} = .9854. Thus, a 95% bootstrap t CI for μ is $\bar{x} \pm t_{.025,21} s_{\text{boot}}$ = 9.955 ± 2.080(.9854) = (7.905, 12.005). This interval is likely more reliable than the one in part **a**; however, the bootstrap distribution looks somewhat skewed, rather than symmetric and bell-shaped, which calls into question the validity of any t interval.

e. The R command `quantile(xbar,probs=c(.025,.975))` returns the 2.5th and 97.5th percentile of the bootstrap distribution from part **c**. For this particular simulation, the returned values were (8.204, 12.091), and that's our 95% confidence bootstrap percentile interval for μ.

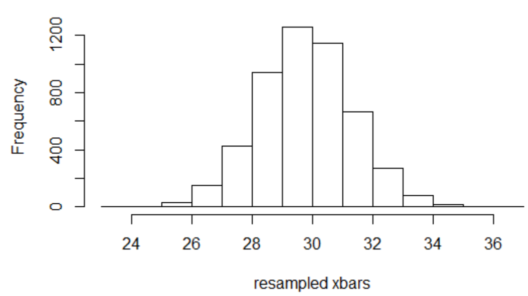
f. Given our concerns about the non-normality of both the population and the sampling distribution of $\bar{X}$ (the latter seen through the bootstrap distribution), the interval in **e** is arguably the only trustworthy interval.

65. For this data, $\bar{x}$ = 29.78, s = 13.07, n = 68.

a. $t_{.025,67}$ = 1.996, so a 95% CI for μ is 29.78 ± 1.996(13.07)/$\sqrt{68}$ = (26.61, 32.94).

b. Because of outliers, the weight gains do not seem normally distributed. However, with n = 68, the effects of the CLT might be enough to validate use of t procedures anyway.

c. Use the R code displayed in the solution to Exercise 63**c**.



d. From the preceding simulation, the sd of the 5000 $\bar{x}^*$ values is s_{boot} = 1.561. Thus a 95% bootstrap t CI for μ is $\bar{x} \pm t_{.025,67} s_{\text{boot}}$ = 29.78 ± 1.996(1.561) = (26.66, 32.90). Because the histogram in **c** is symmetric and bell-shaped, this bootstrap t interval is valid.

e. The 2.5th and 97.5th percentiles of the bootstrap distribution were 26.69 and 32.88, so a 95% confidence bootstrap percentile interval for μ is (26.69, 32.88).

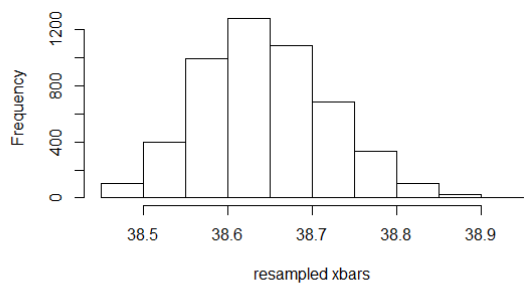
f. All three intervals are quite close. So, in particular, the traditional one-sample t interval in part **a** should be considered valid. (Again, this comports with using one-sample t whenever $n > 40$.)

67. For this data, $\bar{x} = 38.65$, $s = 0.233$, $n = 8$.

a. $t_{.025,7} = 2.365$, so a 95% CI for μ is $38.65 \pm 2.365(0.233)/\sqrt{8} = (38.46, 38.84)$.

b. Although a normal probability plot is not perfectly straight, there is not enough deviation to reject normality.

c. Use the R code in the solution to Exercise 63**c**.



d. The previous simulation gave $s_{\text{boot}} = .07667$. Hence a bootstrap t CI for μ is $\bar{x} \pm t_{.025,7} s_{\text{boot}} = 38.65 \pm 2.365(.07667) = (38.47, 38.83)$. The bootstrap distribution is positively skewed, rather than symmetric and bell-shaped, suggesting that the bootstrap t interval is of questionable validity.

e. The 2.5th and 97.5th percentiles of the bootstrap distribution created in **c** were 38.51 and 38.81. So, a 95% confidence bootstrap percentile interval for μ is (38.51, 38.81).

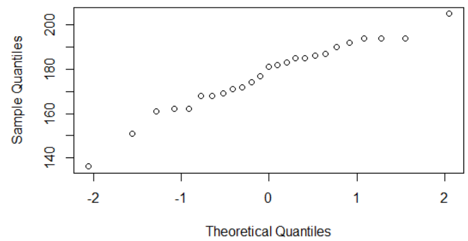
f. All three intervals are surprisingly similar, despite the dubious validity of the bootstrap t interval. In particular, this suggests the one-sample t interval in **a** was fine.

g. Normal body temperature is ~37 degrees Celsius, well below our CIs. This suggests the extreme conditions suffered by these Australian soldiers indeed influenced their body temperatures.

69. For this random sample, $n = 25$, $\bar{x} = 177.16$, and $s = 15.53$.

a. The t critical value is $t_{.025,25-1} = 2.064$, so the CI is $177.16 \pm 2.064(15.53)/\sqrt{25} = (170.75, 183.57)$.

b. The accompanying normal probability plot is reasonably linear, so the one-sample t CI in part **a** is legitimate.



c. Use the R code from the solution to Exercise 63**c**.

d. The bootstrap standard deviation from part **c** was $s_{\text{boot}} = 3.05$. The resulting 95% bootstrap t CI for μ is $177.16 \pm 2.064(3.05) = (170.86, 183.46)$. This is very similar to the one-sample t CI.

e. The 2.5th and 97.5th percentiles of the bootstrap distribution were (171.04, 183.00).

f. All three intervals should be legitimate, so take the shortest one: (171.04, 183.00), the bootstrap percentile interval. Interestingly, *none* of the three intervals capture the true value of μ! It is not too surprising here that all three intervals behave the same, since they're based on the same sample. This particular sample of 25 game times is among the 5 percent whose resulting 95% CI misses μ.

Supplementary Exercises

71. The specified condition is $w/2 = .1$, so $n = \left[\dfrac{(1.96)(.8)}{.1}\right]^2 = 245.86$, and $n = 246$ should be used.

73.

a. With $n = 300$, $\hat{p} = 198/300 = .66$, and $z = z_{.05} = 1.645$, $\dfrac{\hat{p} + z^2/(2n) - z\sqrt{\hat{p}(1-\hat{p})/n + z^2/(4n^2)}}{1 + z^2/n} =$

$\dfrac{.66 + 1.645^2/(2(300)) - 1.645\sqrt{.66(1-.66)/300 + 1.645^2/(4(300)^2)}}{1 + 1.645^2/300} = .614$. With 95% confidence, at least 61.4% of all households in this town would patronize the sandwich franchise (assuming people are honest in their responses).

b. Based on part **a**, a 95% LCB for the total number of households that will eat at the new store is .614(7700) = 4727.8. Since this is less than the required 5000, the company *cannot* be confident they will have enough customers.

c. Note that .66(7700) = 5082, so naively using the point estimate for p would lead the company to the wrong conclusion!

75.

a. For this data, $\bar{x}$ = 0.16835 and s = 0.01938. Assuming normality, $t_{.025,47} \approx 2.01$, so a 95% CI for μ is $0.16835 \pm 2.01\dfrac{0.01938}{\sqrt{48}}$ = (.163, .174).

b. Nine of the 48 reaction times are below .15, so $\hat{p}$ = 9/48 = .1875. Using the formula from this chapter, a 95% CI for the true p is (.089, .326).

77. $\hat{p} = \dfrac{11}{55} = .2 \Rightarrow$ a 90% CI for p is $\dfrac{.2 + \dfrac{1.645^2}{2(55)} \pm 1.645\sqrt{\dfrac{(.2)(.8)}{55} + \dfrac{1.645^2}{4(55)^2}}}{1 + \dfrac{1.645^2}{55}} = \dfrac{.2246 \pm .0887}{1.0492} = (.1295, .2986)$.

79.

a. At the 95% confidence level, the average TV viewing time for the population of all 0-11 months old children is between 0.8 and 1.0 hours per day. The corresponding intervals are 1.4 to 1.8 hours per day for 12-23 months old children and 2.1 to 2.5 hours per day for 24-35 months old children.

b. Samples with larger sample standard deviations will naturally result in wider intervals. Also, samples with smaller sample sizes will result in wider intervals (both because of $\sqrt{n}$ in the denominator of the formula and because of the df for the t critical value).

c. Yes. Since none of the intervals overlap, we have clear evidence that $\mu_{0-11} < \mu_{12-23} < \mu_{24-35}$. That is, there is a direct association between children's ages and mean TV viewing time.

81.

a. The likelihood function is proportional to $\exp([2\sigma^2]^{-1}\sum (y_i - \beta x_i)^2)$. This is maximized exactly when $\sum (y_i - \beta x_i)^2$ is maximized, so differentiate this squared sum with respect to β and solve: $\sum 2(y_i - \beta x_i)(-x_i) = 0 \Rightarrow -2[\sum y_i x_i - \beta \sum x_i^2] = 0 \Rightarrow \beta = \sum y_i x_i / \sum x_i^2$. That is the mle of β is $\hat{\beta} = \Sigma x_i Y_i / \Sigma x_i^2$.

b. $E[\hat{\beta}] = E[\Sigma x_i Y_i / \Sigma x_i^2] = \Sigma x_i E[Y_i] / \Sigma x_i^2 = \Sigma x_i \beta x_i / \Sigma x_i^2 = \beta \Sigma x_i^2 / \Sigma x_i^2 = \beta$.

c. $V[\hat{\beta}] = V[\Sigma x_i Y_i / \Sigma x_i^2] = \Sigma x_i^2 V[Y_i] / (\Sigma x_i^2)^2 = \Sigma x_i^2 \sigma^2 / (\Sigma x_i^2)^2 = \sigma^2 / \Sigma x_i^2$. Thus, the standard deviation of $\hat{\beta}$ is $\sigma / \sqrt{\Sigma x_i^2}$.

d. The variance of $\hat{\beta}$ is inversely proportional to the sum of squares of the x values. Hence, to estimate $\hat{\beta}$ precisely — i.e., to reduce its variance — we should spread out the x values as far from zero as possible.

e. Solving the given expression for β gives the CI $\hat{\beta} \pm t_{.025,n-1} s / \sqrt{\Sigma x_i^2}$. By direct computation on these n = 10 values, we get a 95% CI of (29.93, 30.15).

83. The length of the interval is $\left(z_\gamma + z_{\alpha-\gamma}\right)\dfrac{s}{\sqrt{n}}$, which is minimized when $z_\gamma + z_{\alpha-\gamma}$ is minimized, i.e. when $\Phi^{-1}(1-\gamma)+\Phi^{-1}(1-\alpha+\gamma)$ is minimized. Taking $\dfrac{d}{d\gamma}$ and equating to 0 yields $\dfrac{1}{\phi\left(\Phi^{-1}(1-\gamma)\right)}(-1)+\dfrac{1}{\phi\left(\Phi^{-1}(1-\alpha+\gamma)\right)}=0$ or $\dfrac{1}{\phi\left(z_\gamma\right)}(-1)+\dfrac{1}{\phi\left(z_{\alpha-\gamma}\right)}=0$, where $\phi(z)$ is the standard normal pdf. Simplifying, we have $\phi\left(z_\gamma\right)=\phi\left(z_{\alpha-\gamma}\right)$, which from the symmetry of the standard normal pdf can only be true if $z_\gamma = z_{\alpha-\gamma}$ (both *z* values are positive, so we can discard the – option). Finally, by the uniqueness of quantiles, it follows that $\gamma=\alpha-\gamma$, whence $\gamma=\dfrac{\alpha}{2}$.

85.

a. Since $2\lambda\Sigma X_i$ has a chi-squared distribution with 2*n* df and the area under this chi-squared curve to the *right* of $\chi^2_{.95,2n}$ is .95, $P\left(2\lambda\Sigma X_i > \chi^2_{.95,2n}\right)=.95$. Solving for *λ*, this implies that $\dfrac{\chi^2_{.95,2n}}{2\Sigma X_i}$ is a 95% lower confidence bound for *λ*. The chi-squared critical value for *n* = 10 is $\chi^2_{.95,20}$ = 10.851, so the bound is $\dfrac{10.851}{2(550.87)}=.0098$. We can be 95% confident that *λ* exceeds .0098.

b. Arguing as in **a**, $P\left(2\lambda\Sigma X_i < \chi^2_{.05,2n}\right)=.95$. The following inequalities are equivalent to the one in parentheses:

$$\lambda < \frac{\chi^2_{.05,2n}}{2\Sigma X_i} \quad \Rightarrow -\lambda t > \frac{-t\chi^2_{.05,2n}}{2\Sigma X_i} \quad \Rightarrow e^{-\lambda t} > \exp\left[\frac{-t\chi^2_{.05,2n}}{2\Sigma X_i}\right].$$

Replacing the ΣX_i by Σx_i in the expression on the right-hand side of the last inequality gives a 95% lower confidence bound for $e^{-\lambda t}$. Substituting *t* = 100, $\chi^2_{.05,20}=31.410$ and $\Sigma x_i = 550.87$ gives .058 as the lower bound for the probability that time until breakdown exceeds 100 minutes.

87.

a. Let the 100*p*th percentile of a standard normal population be *z*, so the corresponding percentile for a general normal population is $\theta=\mu+\sigma z$, or $z=(\theta-\mu)/\sigma$. Substitute this into the given expression for *t*, and cancel: $t=\dfrac{\sigma}{s}\left[\dfrac{\bar{x}-\mu}{\sigma/\sqrt{n}}-\left(\dfrac{\theta-\mu}{\sigma}\right)\sqrt{n}\right]=\dfrac{\bar{x}-\theta}{s/\sqrt{n}}$, so this quantity has a non-central *t* distribution. Set this equal to the appropriate critical values (those provided), and a 95% CI for *θ* is $(\bar{x}-t_{.025,n-1,\delta}\cdot s/\sqrt{n}, \bar{x}-t_{.975,n-1,\delta}\cdot s/\sqrt{n})$.

b. The 5th percentile of the *Z* distribution is *z* = –1.645, so $\delta = -(-1.645)\sqrt{16}$ = 6.58. Hence, we can use the critical values provided. Using the data from Exercise 17 gives a 95% CI of (3.01, 4.46).

89.

a. Using independence, the probability is still just $(1/2)(1/2)\ldots(1/2) = (1/2)^n$.

b. There are n possible X_i's that could fall below the median, so the probability exactly one does is n times the answer from (a), or $n(1/2)^n$. Or, let Y = the number of X_i's less than the median. Then $Y \sim$ Bin(n, ½) and $P(Y=1) = \binom{n}{1}(1/2)^1(1-1/2)^{n-1} = n(1/2)^n$.

c. The event $Y_2 > \tilde{\mu}$ occurs if either all n observations lie above the median or exactly one of the observations lies below the median. Combining the previous exercise and (b), $P(Y_2 > \tilde{\mu}) = (1/2)^n + n(1/2)^n = (n+1)(1/2)^n$.

d. Using part **c** and the same rationale as the previous exercise, the CI (y_2, y_{n-1}) for $\tilde{\mu}$ has confidence coefficient $1 - 2[(n+1)(1/2)^n] = 1 - (n+1)(1/2)^{n-1}$.

e. For $n = 10$, this confidence level is 97.85%, and the CI from the data is $(y_2, y_{n-1}) = (29.9, 39.3)$.

91.

a. If A_1 and A_2 are independent, then $P(A_1 \cap A_2) = P(A_1)P(A_2) = (.95)^2 = 90.25\%$.

b. For any events A and B, $P(A \cup B) = P(A) + P(B) - P(A \cap B) \le P(A) + P(B)$. Apply that here: $P(A_1' \cup A_2') \le P(A_1') + P(A_2') = (1-.95) + (1-.95) = .10$, so that $P(A_1 \cap A_2) = 1 - P(A_1' \cup A_2') \ge 1 - .10 = .90$.

c. Replace .05 with α above, and you find $P(A_1 \cap A_2) \ge 100(1 - 2\alpha)\%$. In general, the simultaneous confidence level for k separate CIs is *at least* $100(1 - k\alpha)\%$.