5.51

Individual times are given by $X \sim N(10, 2)$. For day 1, $n = 5$, and so

$$P(\bar{X} \le 11) = P\left(Z \le \frac{11-10}{2/\sqrt{5}}\right) = P(Z \le 1.12) = .8686.$$

For day 2, $n = 6$, and so

$$P(\bar{X} \le 11) = P(\bar{X} \le 11) = P\left(Z \le \frac{11-10}{2/\sqrt{6}}\right) = P(Z \le 1.22) = .8888 .$$

Finally, assuming the results of the two days are independent (which seems reasonable), the probability the sample average is at most 11 min on both days is (.8686)(.8888) = .7720.

5.52

We have $X \sim N(10,1)$, $n = 4$, $\mu_{T_o} = n\mu = (4)(10) = 40$ and $\sigma_{T_o} = \sigma\sqrt{n} = 2$. Hence, $T_o \sim N(40, 2)$. We desire the 95$^{\text{th}}$ percentile of T_o: 40 + (1.645)(2) = 43.29 hours.

5.53

a. With the values provided,

$$P(\bar{X} \ge 51) = P\left(Z \ge \frac{51-50}{1.2/\sqrt{9}}\right) = P(Z \ge 2.5) = 1 - .9938 = .0062 .$$

b. Replace $n = 9$ by $n = 40$, and

$$P(\bar{X} \ge 51) = P\left(Z \ge \frac{51-50}{1.2/\sqrt{40}}\right) = P(Z \ge 5.27) \approx 0 .$$

5.60

Y is normally distributed with $\mu_Y = \frac{1}{2}(\mu_1 + \mu_2) - \frac{1}{3}(\mu_3 + \mu_4 + \mu_5) = -1$, and

$$\sigma_Y^2 = \frac{1}{4}\sigma_1^2 + \frac{1}{4}\sigma_2^2 + \frac{1}{9}\sigma_3^2 + \frac{1}{9}\sigma_4^2 + \frac{1}{9}\sigma_5^2 = 3.167 \Rightarrow \sigma_Y = 1.7795 .$$

Thus, $P(0 \le Y) = P\left(\frac{0-(-1)}{1.7795} \le Z\right) = P(.56 \le Z) = .2877$ and

$$P(-1 \le Y \le 1) = P\left(0 \le Z \le \frac{2}{1.7795}\right) = P(0 \le Z \le 1.12) = .3686.$$

6.3

a. We use the sample mean, $\bar{x} = 1.3481$.

b. Because we assume normality, the mean = median, so we also use the sample mean $\bar{x} = 1.3481$. We could also easily use the sample median.

c. We use the 90th percentile of the sample:
$\hat{\mu} + (1.28)\hat{\sigma} = \bar{x} + 1.28s = 1.3481 + (1.28)(.3385) = 1.7814$.

d. Since we can assume normality,
$P(X < 1.5) \approx P\left(Z < \frac{1.5 - \bar{x}}{s}\right) = P\left(Z < \frac{1.5 - 1.3481}{.3385}\right) = P(Z < .45) = .6736.$

e. The estimated standard error of $\bar{x} = \frac{\hat{\sigma}}{\sqrt{n}} = \frac{s}{\sqrt{n}} = \frac{.3385}{\sqrt{16}} = .0846.$

6.5

Let θ = the total audited value. Three potential estimators of θ are $\hat{\theta}_1 = N\bar{X}$, $\hat{\theta}_2 = T - N\bar{D}$, and $\hat{\theta}_3 = T \cdot \frac{\bar{X}}{\bar{Y}}$. From the data, $\bar{y}$ = 374.6, $\bar{x}$ = 340.6, and $\bar{d}$ = 34.0. Knowing N = 5,000 and T = 1,761,300, the three corresponding estimates are $\hat{\theta}_1 = (5{,}000)(340.6) = 1{,}703{,}000$, $\hat{\theta}_2 = 1{,}761{,}300 - (5{,}000)(34.0) = 1{,}591{,}300$, and $\hat{\theta}_3 = 1{,}761{,}300\left(\frac{340.6}{374.6}\right) = 1{,}601{,}438.281$.

6.15

a. $E(X^2) = 2\theta$ implies that $E\left(\frac{X^2}{2}\right) = \theta$. Consider $\hat{\theta} = \frac{\sum X_i^2}{2n}$. Then
$E(\hat{\theta}) = E\left(\frac{\sum X_i^2}{2n}\right) = \frac{\sum E(X_i^2)}{2n} = \frac{\sum 2\theta}{2n} = \frac{2n\theta}{2n} = \theta$, implying that $\hat{\theta}$ is an unbiased estimator for θ.

b. $\sum x_i^2 = 1490.1058$, so $\hat{\theta} = \frac{1490.1058}{20} = 74.505.$

6.28

a. $\left(\frac{x_1}{\theta}\exp\left[-x_1^2/2\theta\right]\right)\ldots\left(\frac{x_n}{\theta}\exp\left[-x_n^2/2\theta\right]\right) = (x_1\ldots x_n)\frac{\exp\left[-\Sigma x_i^2/2\theta\right]}{\theta^n}$. The natural log of the likelihood function is $\ln(x_i\ldots x_n) - n\ln(\theta) - \frac{\Sigma x_i^2}{2\theta}$. Taking the derivative with respect to θ and equating to 0 gives $-\frac{n}{\theta} + \frac{\Sigma x_i^2}{2\theta^2} = 0$, so $n\theta = \frac{\Sigma x_i^2}{2}$ and $\theta = \frac{\Sigma x_i^2}{2n}$. The mle is therefore $\hat{\theta} = \frac{\Sigma X_i^2}{2n}$, which is identical to the unbiased estimator suggested in Exercise 15.

b. For $x > 0$ the cdf of X is $F(x;\, \theta) = P(X \le x) = 1 - \exp\left[\frac{-x^2}{2\theta}\right]$. Equating this to .5 and solving for x gives the median in terms of θ: $.5 = \exp\left[\frac{-x^2}{2\theta}\right] \Rightarrow x = \tilde{\mu} = \sqrt{-2\theta \ln(.5)} = \sqrt{1.3863\theta}$.

The mle of $\tilde{\mu}$ is therefore $\sqrt{1.3863\hat{\theta}}$.

7.3

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. 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 population values.

d. Not a correct statement. 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 we don't know this to be true.

7.6

a. $8439 \pm \frac{(1.645)(100)}{\sqrt{25}} = 8439 \pm 32.9 = (8406.1, 8471.9)$.

b. $1 - \alpha = .92 \Rightarrow \alpha = .08 \Rightarrow \alpha/2 = .04$ so $z_{\alpha/2} = z_{.04} = 1.75$.

7.13

a. $\bar{x} \pm z_{.025}\frac{s}{\sqrt{n}} = 654.16 \pm 1.96\frac{164.43}{\sqrt{50}} = (608.58, 699.74)$. We are 95% confident that the true average CO_2 level in this population of homes with gas cooking appliances is between 608.58ppm and 699.74ppm

b. $w = 50 = \frac{2(1.96)(175)}{\sqrt{n}} \Rightarrow \sqrt{n} = \frac{2(1.96)(175)}{50} = 13.72 \Rightarrow n = (13.72)^2 = 188.24$, which rounds up to 189.

7.25

a. $$n = \frac{2(1.96)^2(.25) - (1.96)^2(.01) \pm \sqrt{4(1.96)^4(.25)(.25 - .01) + .01(1.96)^4}}{.01} \approx 381$$

b. $$n = \frac{2(1.96)^2\left(\frac{1}{3}\cdot\frac{2}{3}\right) - (1.96)^2(.01) \pm \sqrt{4(1.96)^4\left(\frac{1}{3}\cdot\frac{2}{3}\right)\left(\frac{1}{3}\cdot\frac{2}{3} - .01\right) + .01(1.96)^4}}{.01} \approx 339$$