## Homework 7 Solutions

## Assignment

Chapter 23: 13, 16, 36, 44

## Chapter 23

### 23.13] Normal temperature.

a) Randomization condition: The adults were randomly selected.
$\mathbf{1 0 \%}$ condition: 52 adults are less than $10 \%$ of all adults.
Nearly Normal condition: The sample of 52 adults is large, and the histogram shows no serious skewness, outliers, or multiple modes.
b) We have $n-1=51$ degrees of freedom. The $98 \%$ confidence interval is:

$$
\bar{y} \pm t_{n-1}^{*} \frac{s}{\sqrt{n}}=98.285 \pm 2.403 \frac{0.6824}{\sqrt{52}}=98.285 \pm 0.2274=(98.06,98.51)
$$

I had to use row 50 of the $t$-table, because there is no row 51 .
c) We are $98 \%$ confident that the interval $98.06^{\circ} \mathrm{F}$ to $98.51^{\circ} \mathrm{F}$ contains the true mean body temperature for adults.
d) It means that $98 \%$ of all random samples of size 52 will produce intervals that contain the true mean body temperature of adults.
e) Since the interval is completely below the body temperature of $98.6^{\circ} \mathrm{F}$, there is strong evidence that the true mean body temperature of adults is lower than $98.6^{\circ} \mathrm{F}$.

### 23.16] Parking II.

a) The $95 \%$ confidence interval would be wider than the $90 \%$ confidence interval. We can be more confident that our interval contains the mean parking revenue when we are less precise. This would be better for the city because the $95 \%$ confidence interval is more likely to contain the true mean parking revenue.
b) The $95 \%$ confidence interval is wider than the $90 \%$ confidence interval, and therefore less precise. It would be difficult for budget planners to use this wider interval, since they need precise figures for the budget.
c) By collecting a larger sample of parking revenue on weekdays, they could create a more precise interval without sacrificing confidence.
d) All sample size computations are approximate in some way. We can look at the margin of error from a confidence interval for help:

$$
M E=t_{n-1}^{*} \frac{s}{\sqrt{n}}
$$

Ideally, we could substitute some values and solve for $n$. There are some problems though. We need a value for $s$, the standard deviation. We also need a $t^{*}$ value, but this depends on $n$ !

We'll use $s=15$ as our guess, from problem 14. To get the critical value, we could try two tactics. First, we could use the $Z^{*}$ value instead.

$$
\begin{gathered}
M E=t_{n-1}^{*} \frac{s}{\sqrt{n}} \\
3=1.96 \frac{15}{\sqrt{n}} \\
n=\frac{1.96^{2}\left(15^{2}\right)}{3^{2}}=96.04
\end{gathered}
$$

Another approach would be to use information from problem 14 and treat it as a pilot study. There, $n=44$. Then use $t_{43}^{*}=2.014$ [Row 45 of the problem]. This would give us

$$
\begin{gathered}
M E=t_{n-1}^{*} \frac{s}{\sqrt{n}} \\
3=2.014 \frac{15}{\sqrt{n}} \\
n=\frac{2.014^{2}\left(15^{2}\right)}{3^{2}}=101.4
\end{gathered}
$$

Using the first method, we'd collect data from 97 trees. We'd need data from 102 days using the other method.

### 23.36] Ski wax.

a) Let $\mu$ be the mean race time. We want to test

$$
\begin{aligned}
& H_{0}: \mu=55 \\
& H_{A}: \mu<55
\end{aligned}
$$

Bjork decided not to buy the wax, when in reality it would have helped. He failed to reject the null hypothesis when the alternative was true. This is an example of a Type II error.
b) The hypotheses have already been stated in part (a). We should check the assumptions: Independence assumption: Since the times are not randomly selected, we will assume that the times are independent, and representative of all times.

Nearly Normal condition: I used RCmdr to plot both a histogram of the data and a normal Q-Q plot. These are given below. The histogram of the times is unimodal and roughly symmetric. The normal Q-Q plot shows no great deviations from normality.


The needed assumptions seem to be satisfied.
The times in the sample had a mean of 53.1 seconds and a standard deviation of 7.029 seconds. We will perform a one-sample $t$-test.

The test statistic is $t=\frac{\bar{y}-\mu_{0}}{s / \sqrt{n}}=\frac{53.1-55}{7.029 / \sqrt{8}}=\frac{-1.9}{2.4851}=-0.7645$.
I used a $t$-table to look up an approximate $p$-value. The $t=-0.7645$ is off the scale on the left side of row 7. Therefore the $p$-value must be greater than 0.10 . With such a high $p$ value, we fail to reject the null hypothesis. There is no evidence to suggest the mean time is less than 55 seconds. Bjork should not buy the new ski wax.

### 23.44] Wind power.

a) Let $\mu$ be the mean wind speed at the site. We wish to test

$$
\begin{aligned}
& H_{0}: \mu=8 \mathrm{mph} \\
& H_{A}: \mu>8 \mathrm{mph}
\end{aligned}
$$

Now, we check the conditions needed. They seem to be satisfied.
Independence assumption: Data taken in a time series such as here can be a problem, as nearby measurements tend to be similar to each other. However, the timeplot shows no pattern, so it seems reasonable that the measurements are independent.
Randomization condition: This is not a random sample, but an entire year is measured. These wind speeds should be representative of all wind speeds at this location. $\mathbf{1 0 \%}$ condition: These wind speeds certainly represent fewer than $10 \%$ of all wind speeds.
Nearly Normal condition: The Normal probability plot is reasonably straight, and the histogram of the wind speeds is unimodal and reasonably symmetric.
b) We will perform a one-sample $t$-test.

The test statistic is $t=\frac{\bar{y}-\mu_{0}}{s / \sqrt{n}}=\frac{8.019-8}{3.813 / \sqrt{1114}}=\frac{0.019}{0.1142}=0.1663$.
I used a $t$-table to look up an approximate $p$-value. We have $\mathrm{n}-1=1113$ degrees of freedom. The $t=-0.7645$ is off the scale on the left side of row 1000 . Therefore the $p$ value must be greater than 0.10 . With such a high $p$-value, we fail to reject the null hypothesis. There is no evidence to suggest the mean wind speed is over 8 mph . I wouldn't recommend building the turbine at this site.

