

## 1 Chapter 21 review exercises

### # 21.4

$172 + 207 = 379$  households in the sample than had one or more cars. Thus  $\frac{379}{500} * 100 = 75.8\%$  of households in the sample had one or more cars. This is our estimated percentage for the population. The standard error for the number is  $\sqrt{500} \times \sqrt{.758 \times .242} = 9.58$ . Using this we can find the standard error for the percentage by dividing by the number of draws,  $\frac{9.58}{500} = .02$ , or 2%.

### # 21.6

This is false. It is very likely that the sample percentage will be close to the population percentage plus or minus chance error. It is not very likely that the sample percentage will be exactly equal to the population percentage.

### # 21.7

No, the standard error has not been calculated correctly. In fact, we cannot calculate any standard errors here. To see this, recall that the data consists of observations taken on 252 consecutive days. That is certainly not a random sample and therefore no box model applies, indicating that it makes no sense to calculate a standard error.

### # 21.11

These are the same. The average is the sum of the draws divided by the number of draws, so having a sum of 100 draws greater than 710 is equivalent to the average of the draws being greater than 7.1.

## 2 Chapter 23 review exercises

### # 23.3

(a) Part of the answer to this question is given directly in the text: the average commute distance is estimated as 8.7 miles. However, in the second gap we have to fill in the value of the SE for the average, which is not given directly, but rather has to be calculated. We can find the SE for the average as follows:

$$SE_{average} = \frac{\sqrt{\# \text{ of draws}} \times SD}{\# \text{ of draws}} = \frac{\sqrt{1000} \times 9 \text{ miles}}{1000} = 0.2846 \text{ miles}$$

The complete sentence should thus read, “The average commute distance of all 50,000 heads of households in the town is estimated as 8.7 miles, and this estimate is likely to be off by 0.2846 miles or so.”

(b) To obtain a 95% confidence interval, we can use the values from (a). The 95% confidence interval can be calculated via  $8.7 \text{ miles} \pm (2 \times 0.2846 \text{ miles})$ , or  $[8.1308 \text{ miles}, 9.2692 \text{ miles}]$ .

### # 23.4

Given that the sample at hand is large, we can use the sample SD to estimate the SD of the “box” (underlying population). As always, finding a 95% confidence interval requires us to have information on the SE, in this case on the SE for the average. This information can be found via:

$$SE_{average} = \frac{\sqrt{\# \text{ of draws}} \times SD}{\# \text{ of draws}} = \frac{\sqrt{2500} \times 10.2 \text{ miles}}{2500} = 0.204 \text{ miles}$$

Using this value for the  $SE_{average}$ , we can compute the 95% confidence interval for the average commute distance for all people age 16 and over in the town as 7.1 miles  $\pm$  ( $2 \times 0.204$  miles), or [6.692 miles, 7.508 miles].

### # 23.7

(iii) is the best option, because we cannot be sure that we are dealing with a random sample here and thus cannot make statements about the sample’s standard errors. In all likelihood, we are confronted with selection bias, meaning that predominantly those women that are concerned about stress filled out the survey forms. The simple fact that the questionnaires were distributed widely does not mean that we are dealing with a random sample.

### # 23.8

(a) True. Given the information in the question, we can assert that a 68% confidence interval ( $\pm 1$  S.E. around the mean) will run from 3,375 to 4,025.

(b) False. The confidence interval will either include the true population parameter or not, i.e., there is no such thing as a 68% chance that this will be the case. Either the interval includes it or not, read: the chance that the true population parameter is included in the confidence interval is either 0 or 1.

(c) False. The only thing about which we can assume normality is sampling distribution of average enrollment, not enrollment itself. In fact, one can imagine situations where a confidence interval for the sampling distribution of average enrollment contains regions over which no schools actually have enrollments at that level. This is not the quantity to which the confidence level refers.

(d) False. Recall that the 68% for the confidence interval refer to the estimation procedure, not the parameter in question. For a “68% confidence interval,” we say that in repeated sampling, 68% of the confidence intervals will include the true population parameter, while 32% will not.

(e) False. Given the large sample size, we can assume a normal distribution of the data used in the study. Recall that “[w]hen drawing from a box, the probability histogram for the average of the draws will follow the normal curve, even if the contents of the box do not.” (FPP, p. 412)

### 3 Chapter 26 review exercises

1. (a) True. Observed significance level is the definition of the  $P$ -value (see FPP, 4th ed., p. 479).
- (b) False. The alternative hypothesis is that there is a real difference in the population, so any observed difference is *not* due to chance.
4. No, the TA cannot present a strong case. In order to answer this question, we have to compute the  $SE_{\text{average}}$ , which will then be used to find the  $z$ -score and determine whether an average of 55 is statistically different from 63 given the data at hand. In addition, it is helpful to write out the null and the alternative hypothesis:

$$H_0 : \mu = 63$$

$$H_1 : \mu < 63$$

The standard error of the average is

$$SE_{\text{average}} = \frac{\sqrt{n} \times SD}{n} = \frac{\sqrt{30} \times 20}{30} \approx 3.65$$

Next, the test statistic (here a one-sample  $z$ -score) can be calculated:

$$\text{test statistic} = \frac{\text{observed} - \text{expected}}{SE_{\text{average}}} = \frac{55 - 63}{3.65} = \frac{-8}{3.65} = -2.19$$

Finally, we can look up the area that will be in the lower tail for a  $z$ -score of  $-2.19$ ; it is approximately 1.39%.

This low  $P$ -value means that, contrary to the TA's claims, it would be highly unlikely to have a test average of 55 or less in a random sample of 30 tests from a class with an average of 63. In other words, we reject the null hypothesis, and conclude that there is some systematic difference between this TA's section and the rest of the class.

6. (a) When 50% of the population is female, we expect 50% of the jurors to be female ( $350/2 = 175$  jurors), which gives us the null hypothesis. The question asks how likely it is to get "102 women *or fewer*," so we will use a one-sided alternative.

$$H_0 : \mu = 50\%$$

$$H_1 : \mu < 50\%$$

Under the null hypothesis, the population SD is  $\sqrt{0.5 \times 0.5} = 0.5$ , so the SE of the percentage for a sample of 350 is

$$SE_{\text{percentage}} = \frac{\sqrt{350} \times 0.5}{350} \times 100\% = 2.67\%.$$

The expected percentage under the null hypothesis is 50%, and the observed percentage is  $102/350 \times 100\% = 29\%$ . The  $z$ -score is

$$\frac{29 - 50}{2.67} \approx -8$$

The  $P$ -value corresponding to this test statistic is less than 0.02% (the lowest value on the provided  $Z$ -table). Therefore, there is a less than 0.02% chance of randomly drawing 102 women in a sample of 350 if the true population proportion of women is 50%. If the population proportion were greater than 50%, the chance would be even lower.

- (b) The population proportion of women in the “venire” of 350 people is  $102/350 \approx 29\%$ . We want to make inferences about the rule used to select the final panel of 100 potential jurors. Our null hypothesis is that the selection rule is fair, so we expect 29% women in the final panel. This part also asks about the chance of getting 9 women or fewer, so we again use a one-sided alternative.

$$H_0 : \mu = 29\%$$

$$H_1 : \mu < 29\%$$

Under the null hypothesis, the population SD is  $\sqrt{0.29 \times 0.71} \approx 0.452$ , so the SE for the percentage in a sample of 100 is

$$SE_{\text{percentage}} = \frac{\sqrt{100} \times 0.452}{100} \times 100\% = 4.52\%.$$

The observed percentage is 9%, so our  $z$ -score is

$$\frac{9 - 29}{4.52} = -4.42.$$

As in part (a), the corresponding  $P$ -value is less than 0.02%. Therefore, there is a less than 0.02% chance of randomly drawing a sample with 9 women or fewer.

- (c) Based upon the results from (a) and (b), we conclude that the selection process for the jury was heavily biased against women. In both stages, there was a virtual zero chance that as few women would be selected as were actually selected had the process been “fair,” i.e., undertaken without regard to gender. Dr. Spock’s right to a fair trial was clearly undermined even before it started.
8. The null hypothesis is that the county is no different from the nation as a whole in its educational level. The alternative in this case is two-sided, as the question does not ask about “greater” or “lesser.”

$$H_0 : \mu = 13$$

$$H_1 : \mu \neq 13$$

To conduct a significance test, we need to compute the standard error of the average for a sample of 1,000. Since the SD of the “box” under the null hypothesis (education level of the nation) is known, we should use this—and not the sample SD—to compute the standard error of the average. (See section 26.5 of FPP.)

$$SE_{\text{average}} = \frac{\sqrt{n} \times SD}{n} = \frac{\sqrt{1000} \times 3}{1000} \approx 0.1.$$

The test statistic for our significance test is the z-score,

$$\frac{\text{observed} - \text{expected}}{\text{SE}_{\text{average}}} = \frac{14 - 13}{0.1} = 10.$$

Based upon this very large z-score (which corresponds to an infinitesimal  $P$ -value), we reject the null hypothesis that the educational level in the county is the same as in the nation as a whole.

A possible explanation for the pattern found is that the county is one where most jobs require high levels of education, such as in Silicon Valley or near a large NASA center.

#### 4 Chapter 27 review exercises

3. (a) In this case, we should use a two-sample z-test. We have two independent samples from different years and want to compare them to each other.
- (b) The null hypothesis is that there is no difference between 2000 and 2005 in terms of the percentage of people who rate clergymen's honesty and ethics as high or very high. To formulate this in terms of a box model, we need two boxes, a 2000 box and a 2005 box, since there are two different samples. Each box contains millions of tickets, one for each adult in the country being sampled at the time. There are two ticket values: 1 for each person who the clergy's honesty as "high or very high," and 0 for those who do not. There are 1,000 draws from each box. In terms of the box model, the null hypothesis is that the percentage of 1s is the same between the two boxes, whereas the alternative hypothesis is that this percentage differs between them.
- (c) To perform a significance test for the difference, we first need to compute the SE of the percentage for each sample. We have

$$\begin{aligned} \text{SE}_{2000} &= \frac{\sqrt{n} \times \text{SD}_{2000}}{n} \times 100\% = \frac{\sqrt{1000} \times \sqrt{0.6 \times 0.4}}{1000} \times 100\% = 1.54\% \\ \text{SE}_{2005} &= \frac{\sqrt{n} \times \text{SD}_{2005}}{n} \times 100\% = \frac{\sqrt{1000} \times \sqrt{0.54 \times 0.46}}{1000} \times 100\% = 1.58\%. \end{aligned}$$

Therefore, the SE of the difference is

$$\text{SE}_{\text{difference}} = \sqrt{(\text{SE}_{2000})^2 + (\text{SE}_{2005})^2} = \sqrt{1.54^2 + 1.58^2} = 2.21\%.$$

The observed difference between the samples is  $60\% - 54\% = 6\%$ , whereas the expected value under the null hypothesis of no difference is  $0\%$ . We compute the test statistic (a z-score) in the familiar way:

$$\text{test statistic} = \frac{\text{observed} - \text{expected}}{\text{SE}} = \frac{6\% - 0\%}{2.21\%} = 2.71.$$

This corresponds to a  $P$ -value of about  $100\% - 99.31\% = 0.69\%$ . Therefore, we reject the null hypothesis that the difference between the samples can be

explained by chance variation. However, this does not necessarily mean the scandals were responsible. Other intervening factors between 2000 and 2005 (e.g., declining religiosity in the population or an overall decline in public trust) could be the cause.

4. In this question, we have the same set of respondents answering two different questions. This means the samples are not independent, since an individual's trust in druggists may be correlated with his or her trust in doctors. With non-independent samples, we need the individual data; we cannot simply use the overall averages to perform hypothesis tests, because the square root law for the SE of the difference does not apply.
7. (a) The first step here is to set up the null and alternative hypothesis. The null hypothesis is that income support has no effect on recidivism, so the re-arrest rate is the same for supported and unsupported prisoners. The question asks about whether support *reduces* recidivism, so we have a one-sided alternative hypothesis. The population parameter of interest is  $\mu = (\% \text{ recidivism for supported} - \% \text{ recidivism for unsupported})$ .

$$H_0 : \mu = 0\%$$

$$H_1 : \mu < 0\%$$

To compute the SE of the difference, we need the SE of the percentage for each group. Let  $SE_{\text{percentage}}^t$  be the SE for the treatment group, and  $SE_{\text{percentage}}^c$  for the control group. The SDs are  $SD^t = \sqrt{0.483 \times 0.517} \approx 0.5$  for the treated group and  $SD^c = \sqrt{0.494 \times 0.506} \approx 0.5$  for the control group.

$$SE_{\text{percentage}}^t = \frac{\sqrt{592} \times 0.5}{592} \times 100\% = 2.06\%$$

$$SE_{\text{percentage}}^c = \frac{\sqrt{154} \times 0.5}{154} \times 100\% = 4.03\%$$

Using the square root rule, the standard error of the difference is

$$SE_{\text{difference}} = \sqrt{(SE_{\text{percentage}}^t)^2 + (SE_{\text{percentage}}^c)^2} = \sqrt{2.06^2 + 4.03^2} = 4.53\%.$$

The expected difference under the null hypothesis is 0%. The observed difference is  $48.3\% - 49.4\% = -1.1\%$ . Our test statistic (a z-score) is therefore

$$\frac{-1.1 - 0}{4.53} = -0.243.$$

This corresponds to a  $P$ -value of roughly 40.52%, which is much larger than the conventional significance level of 5%, so we fail to reject the null hypothesis. There is not strong enough evidence that income support reduces recidivism.

- (b) This question requires the same approach as in (a), but now we are dealing with averages rather than percentages. The SE of the average in each group is

$$\begin{aligned} \text{SE}_{\text{average}}^t &= \frac{\sqrt{592} \times 15.9}{592} = 0.653 \\ \text{SE}_{\text{average}}^c &= \frac{\sqrt{154} \times 17.3}{154} = 1.394. \end{aligned}$$

Using the square root rule, the SE of the difference is

$$\text{SE}_{\text{difference}} = \sqrt{(\text{SE}_{\text{average}}^t)^2 + (\text{SE}_{\text{average}}^c)^2} = \sqrt{0.653^2 + 1.394^2} = 1.539.$$

The expected difference under the null hypothesis of no effect is 0. The question asks about whether income support *reduces* hours worked, so the alternative hypothesis is one-sided. The observed difference in average hours worked is  $16.8 - 24.3 = -7.5$ . Our test statistic (a z-score) is

$$\frac{-7.5 - 0}{1.539} = -4.87.$$

This corresponds to an infinitesimal  $P$ -value, so we would reject the null hypothesis. It would be highly unlikely to observe such a large difference in sample averages if income support truly has no effect.