Chapter 10

Answering Questions About Population Characteristics

Statistical inference is the process of using the characteristics of a sample to make statements about the population from which it is drawn. There are many forms of statistical inference, including interval estimation and hypothesis testing. These procedures can be used to:

- compute the range of plausible values for scores on an IQ test;
- determine whether mammals get, on average, 9 hours of sleep per day;
- test that the proportion of minority firefighter applicants is 26%, the same as the general population;
- determine whether reading scores for students exposed to a particular instruction program change over time.

10.1 AN INTERVAL OF PLAUSIBLE VALUES FOR A MEAN

The sample mean is a point estimate of the population mean, but it is usually not exactly equal to the population mean. The degree of accuracy can be indicated by reporting the estimate of the population mean as a range of values, that is, a confidence interval. The standard error, a measure of precision of the point estimate, is incorporated into the confidence interval.

In instances in which the population standard deviation (σ) is known, it can be used directly to compute the standard error and obtain the confidence interval. In most cases, however, the value of σ is not known and the sample standard deviation (s) must be substituted to give the estimate of the standard error of the mean, $s_{\bar{x}}$. SPSS for Windows bases its computations on the sample standard deviation, and we will illustrate only this situation.

We will compute the confidence interval for the mean number of calories per hot dog, using the data file "hotdog.sav."

1. Click on **Analyze** from the menu bar.
2. Click on **Descriptive Statistics** from the pull-down menu.
3. Click on **Explore**.
4. Click on the "calories" variable and move it into the Dependent List box with the **top right arrow button**.
5. Click on **Statistics** to open the Explore: Statistics dialog box.
6. Note that **Descriptives: Confidence Interval for Mean** is checked and 95% is the default. You may change this if you require a different level of confidence by moving your cursor to this box and typing in the desired level. For this exercise, we will maintain the default 95% level of confidence.
7. Click on **Continue**.
8. Click on **Statistics** in the Display box. This is an optional step, but recommended here because it suppresses tables that are unnecessary to address the question at hand.
9. Click on **OK**.

Your output will appear as shown in Figure 10.1. Note that the sample mean is 145.44 calories. The confidence interval is labeled "95% Confidence Interval for Mean" with a lower bound of 137.42 and an upper bound of 153.46. This indicates that we are 95% confident that the population mean is in the range 137.42 to 153.46 calories.

10.2 TESTING A HYPOTHESIS ABOUT A MEAN

The procedures for testing hypotheses about a mean differ slightly depending upon whether the standard deviation of the population (σ) is known or unknown. In this book, we illustrate the more common case, when the population standard deviation is unknown.

Case Processing Summary

	Cases					
	Valid		Missing		Total	
	N	Percent	N	Percent	N	Percent
CALORIES	54	100.0%	0	.0%	54	100.0%

Descriptives

			Statistic	Std. Error
CALORIES	Mean		145.4444	3.99857
	95% Confidence Interval for Mean	Lower Bound	137.4243	
		Upper Bound	153.4646	
	5% Trimmed Mean		146.0123	
	Median		145.0000	
	Variance		863.384	
	Std. Deviation		29.38339	
	Minimum		86.00	
	Maximum		195.00	
	Range		109.00	
	Interquartile Range		41.7500	
	Skewness		-.167	.325
	Kurtosis		-.691	.639

Figure 10.1 Explore Output with 95% Confidence Interval

The first step in hypothesis testing is generating two competing hypotheses for the question you are asking — the null hypothesis, denoted H_0, and the alternative hypothesis, denoted H_1. These hypotheses are about the value of the population mean, and are determined by the specific research question being asked. Hypothesis testing involves drawing a random sample from the population, computing the sample mean, and converting it to a *test statistic* that indicates how far the sample mean is from the hypothesized population mean. The test statistic is compared to percentage points of the appropriate probability distribution to decide if H_0 is maintained or rejected.

Validity Conditions

It is good practice to check the data for normality prior to conducting the test of significance. A visual way to inspect for normality is to plot a histogram of the variable. There is an option that directs SPSS to impose a normal curve on the graph, which makes it easier to evaluate the normality of the data. We will illustrate this using the calories variable of the "hotdog.sav" data file. You can obtain this graph using the Frequencies procedure as follows:

1. Click on **Graphs** from the menu.
2. Click on **Histogram** from the pull-down menu.
3. Click on and move the "calories" variable to the Variable box of the Histogram dialog box by clicking on the **right arrow button**.
4. Click on the **Display normal curve** option.
5. Click on **OK**.

Your output should look like that shown in Figure 10.2.

Although the distribution is not precisely normal, it is not highly skewed either. The histogram fits fairly well under the normal curve superimposed on the graph. Because the test is fairly "robust" with respect to normality, we conclude that it is an appropriate method for hypothesis testing in this instance.

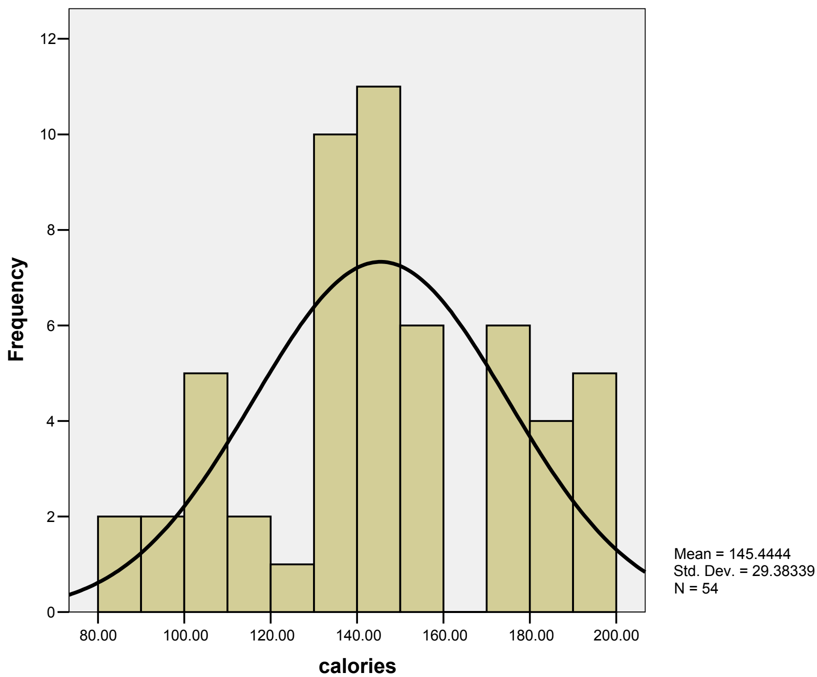

Figure 10.2 Sample Histogram with Normal Curve Superimposed

Conducting the Hypothesis Test

When the standard deviation of the population is not known, you will need to estimate σ in order to compute the test statistic. SPSS computes the sample standard deviation, s, and uses it to calculate the appropriate test statistic (t). The procedure for using SPSS is the same regardless of whether you are making a one-tailed or a two-tailed test.

We will illustrate again using the calories variable from "hotdog.sav." Treating this sample as representative of the population of hotdogs, we will test the hypothesis that the population mean (represented as μ) number of calories is 150. The null and alternative hypotheses are H_0: $\mu = 150$ and H_1: $\mu \neq 150$.

Before computing the test statistic, as the data analyst you must select your error level, α. This represents the probability of committing a Type I error; that is, the probability of rejecting the null hypothesis when it is true. In this example, we select an α level of .05. That is, we have chosen to accept a 5% chance of incorrectly rejecting the null hypothesis (incorrectly stating that hot dogs do not have, on average, 150 calories).

We can now use SPSS to determine the test statistic for this sample and use it to conduct the hypothesis test. Once the data file is open:

1. Click on **Analyze** from the menu bar.
2. Click on **Compare Means** from the pull-down menu.
3. Click on **One-Sample T Test** from the pull-down menu to open the One-Sample T Test dialog box (see Fig. 10.3).
4. Highlight the "calories" variable and move it to the Test Variable(s) box by clicking on the **right arrow button.**

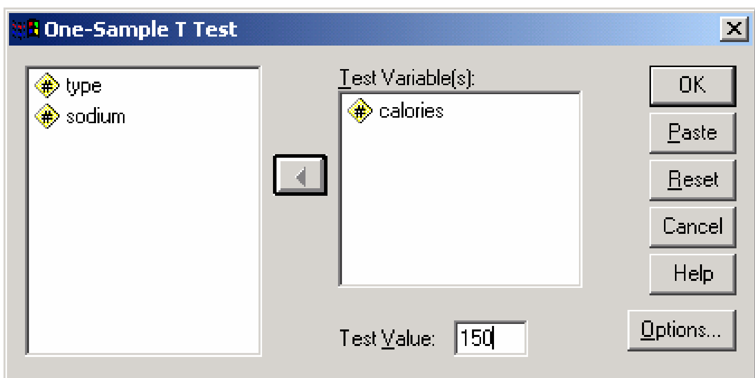

Figure 10.3 One-Sample T Test Dialog Box

One-Sample Statistics

	N	Mean	Std. Deviation	Std. Error Mean
CALORIES	54	145.4444	29.38339	3.99857

One-Sample Test

	Test Value = 150					
					95% Confidence Interval of the Difference	
	t	df	Sig. (2-tailed)	Mean Difference	Lower	Upper
CALORIES	-1.139	53	.260	-4.5556	-12.5757	3.4646

Figure 10.4. Sample Output for One-Sample T-Test

5. In the Test Value box, enter the number 150.
6. Click on **OK** to run the procedure.

The output should look like that in Figure 10.4.

From this listing we see that the mean of the 54 hotdog brands is 145.44 calories and the standard deviation is 29.38 calories (the same as was displayed in Figure 10.1). The *t*-statistic for the test is $t = -1.139$. This is found by computing the following:

$$t = \frac{\bar{x} - \mu_0}{\frac{s}{\sqrt{n}}} = \frac{145.44 - 150}{\frac{29.38}{\sqrt{54}}} = -1.139$$

Thus, the test statistic represents the number of standard errors the sample mean (here, 145.44 calories) is greater or less than the test mean (here, 150 calories).

In order to determine whether to reject the null hypothesis based on this test statistic, we must determine the probability of obtaining a value more extreme than our test statistic when the null hypothesis is true. This probability is called the *P* value. If the *P* value is less than our chosen α level, we reject the null hypothesis; if the *P* value is greater than our chosen α level, we do not reject the null hypothesis.

In Figure 10.4, the column labeled "Sig. (2-tailed)" is the *P* value for this test. Notice that the *P* value is listed as .260. So, if the null hypothesis were true (hot dogs have, on average, 150 calories), then the probability of obtaining a test statistic with an absolute value at least 1.139 is less than .260. This is greater than our chosen α level of .05, so we cannot reject the null hypothesis. We accept H_0 and conclude that the average number of calories in hot dogs does not differ from 150.

SPSS reports *P* values for two-tailed tests. If we were performing a one-tailed test, we would be concerned only with the upper (or lower) tail of the *t*-

distribution. In order to obtain the correct achieved significance level, the *P* value produced by SPSS must be divided by 2. To reject the null hypothesis when conducting a one-tailed test, (a) $P/2$ must be less than α *and* (b) the sample mean must be in the direction specified by the alternative hypothesis (H_1).

Refer again to the hotdog example (Fig. 10.4). Suppose we want to test the null hypothesis H_0: $\mu \geq 150$ against the alternative H_1: $\mu < 150$, using $\alpha = .05$. The *P* value reported by SPSS is $P < .260$. We divide *P* by 2 and find that the achieved significance level is less than .130. This is not less than our α of .05. Thus, we do not reject H_0; we conclude that, on average, hot dogs do not have less than 150 calories.

Relationship Between Two-Tailed Tests and Confidence Intervals

There is a specific relationship between two-tailed tests and confidence intervals. If μ_0 lies within the $(1 - \alpha)$ confidence interval for the mean of a variable, the null hypothesis would not be rejected at the α level of significance. Conversely, if μ_0 is not within the interval, then the null hypothesis would be rejected. Return to Figure 10.1, and note that the 95% confidence interval contains the values of μ_0, 150 calories.

10.3 TESTING HYPOTHESES ABOUT A PROPORTION

The steps in testing hypotheses about a proportion are similar to testing hypotheses about a mean demonstrated in the previous section of this chapter. The null and alternative hypotheses are stated in terms of a probability, p, and a hypothesized value, p_0. The hypothesis is tested by computing the test statistic:

$$z = \frac{(\hat{p} - p_0)}{\sigma_{\hat{p}}}$$

where $\hat{p}$ is the sample proportion, and the standard error is

$$\sigma_{\hat{p}} = \sqrt{\frac{p_0(1-p_0)}{n}}$$

We can test hypotheses about a proportion using the Binomial procedure in SPSS. Using the “fire.sav” data file, let us test the hypothesis that the proportion of white applicants is less than the national population percentage, which is approximately 74%. The hypotheses are H_0: $p \geq .74$ and H_1: $p < .74$. After you open the data file:

1. Click on **Analyze** from the menu bar.
2. Click on **Nonparametric Tests** from the pull-down menu.
3. Click on **Binomial** from the pull-down menu to open the Binomial Test dialog box (Fig. 10.5).
4. Click on and move the "race" variable to the Test Variable List box using the **right arrow button**.
5. Type .74 in the Test Proportion box. The .74 represents p_0.
6. Click on **OK** to run the procedure.

The output listing is displayed in Figure 10.6. The Test Proportion is the p_0 you entered in Step 5. The Obs. Prop. is the proportion of cases with value of 1 in the data file ($\frac{17}{28}$ in the example). (Note that SPSS always calculates the proportion using the value of the variable with the larger number of cases. Therefore, you must be careful to enter the appropriate test proportion.)

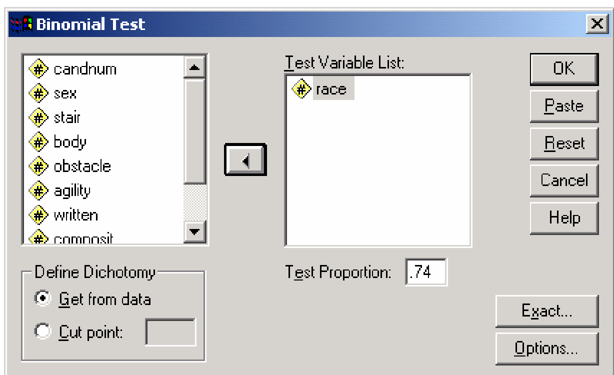

Figure 10.5 Binomial Test Dialog Box

Binomial Test

		Category	N	Observed Prop.	Test Prop.	Asymp. Sig. (1-tailed)
race	Group 1	white	17	.61	.74	.086[a,b]
	Group 2	minority	11	.39		
	Total		28	1.00		

a. Alternative hypothesis states that the proportion of cases in the first group < .74.

b. Based on Z Approximation.

Figure 10.6 Sample Output for Binomial Test

The printout also contains the P value using the normal approximation to the binomial distribution. SPSS employs a continuity correction for $\hat{p}$, adding $\frac{1}{2n}$ to the sample proportion. The "corrected" sample proportion is $\frac{17}{28} + \frac{1}{56} = .625$, and the test statistic is

$$z = \frac{(.625 - .74)}{\sqrt{\frac{.74 \times .26}{28}}} = -1.39$$

The corresponding P value from the approximation to the standard normal distribution is $P = .086$. Note that the footnotes indicate this is a one-tailed P value. If we were using an α level of .05, we would accept the null hypothesis that the proportion of white applicants is greater than or equal to .74.

10.4 PAIRED MEASUREMENTS

There are two main types of paired measurements that occur in statistics. One involves two measurements being made on one unit of observation at two times, such as body weight before and after a diet, or the scores made on a college admissions test before and after a preparatory class. Matched samples are also used to choose individuals with similar characteristics but assigned to different experimental conditions. We shall demonstrate how to use SPSS to test hypotheses about the mean of a population of differences (often referred to as difference scores), and hypotheses about the equality of proportions from paired measurements.

Testing Hypotheses About the Mean of a Population of Differences

SPSS conducts a test of whether the mean of a population of difference scores is equal to 0. We illustrate with the "reading.sav" data file. This file contains reading scores for 30 students obtained on the same test administered before and after second grade. We want to determine whether reading skill increases, on average, throughout second grade. Suppose we choose an α level of .05. After opening the data file:

1. Click on **Analyze** from the menu bar.
2. Click on **Compare Means** from the pull-down menu.
3. Click on **Paired-Samples T Test** from the pull-down menu. This opens the Paired-Samples T Test dialog box (see Fig. 10.7).
4. Click on the "before" variable. It will appear in the Current Selections box as Variable 1.
5. Click on the "after" variable. It will appear in the Current Selections box as Variable 2.
6. Move the paired variables into the Paired Variables box by clicking on the **right arrow button**.
7. Click on **OK** to run the procedure.

Figure 10.8 displays the output from this procedure. The top portion lists the means, standard deviations, and standard errors of reading scores before and after second grade. The scores increased, on average, from 1.52 to 2.03 points.

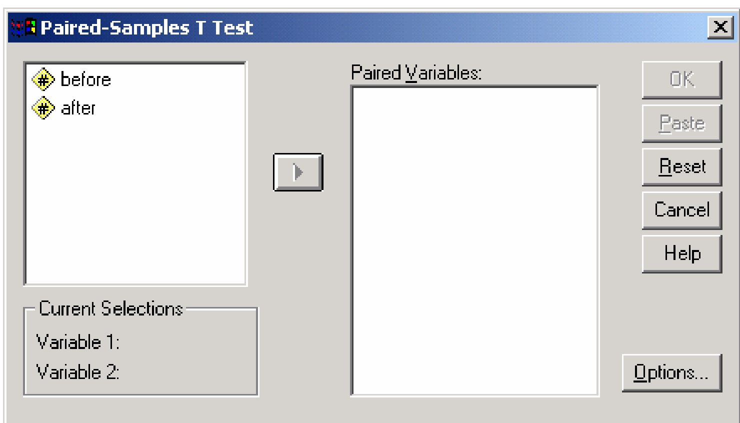

Figure 10.7 Paired-Samples T Test Dialog Box

The middle portion includes the sample correlation coefficient between the pre-test and post-test (.573) and a test of significance of the correlation ($P <$ .001). This is not a component of the hypothesis test we are conducting in this example.

The lowest portion of the output contains information regarding the test of the hypothesis that the mean difference is equal to 0. The mean difference, –.5100, is equivalent to 1.5233 – 2.0333. The table also displays the standard deviation and the standard error of the difference. The t statistic is

$$t = \frac{-.5100 - 0}{\frac{.492}{\sqrt{30}}} = -5.683 \text{ .}$$

The two-tailed P value is less than .0005 (and is rounded to .000 in SPSS). However, we are conducting a one-tailed test because we began by speculating that reading scores would increase. Therefore, we must compare the $P/2$ value to α and verify that the sample post-test mean is higher than the sample pretest mean. In this case, $P/2 < .00025$, which is less than our selected significance level, .05. In addition, mean score after second grade is higher than that before second grade. Consequently, we reject the null hypothesis and conclude that reading skills increase during grade 2.

Paired Samples Statistics

		Mean	N	Std. Deviation	Std. Error Mean
Pair 1	Reading Score Before 2nd Grade	1.5233	30	.27628	.05044
	Reading Score After 2nd Grade	2.0333	30	.59442	.10853

Paired Samples Correlations

		N	Correlation	Sig.
Pair 1	Reading Score Before 2nd Grade & Reading Score After 2nd Grade	30	.573	.001

Paired Samples Test

		Paired Differences					t	df	Sig. (2-tailed)
		Mean	Std. Deviation	Std. Error Mean	95% Confidence Interval of the Difference Lower	95% Confidence Interval of the Difference Upper			
Pair 1	Reading Score Before 2nd Grade - Reading Score After 2nd Grade	-.5100	.49155	.08974	-.6935	-.3265	-5.683	29	.000

Figure 10.8 Sample Output for *t* Test for Paired Samples

Testing the Hypothesis of Equal Proportions

You may also examine changes in time for dichotomous variables by testing the equality of two proportions obtained from a single sample. Some textbooks refer to this a "turnover table," and SPSS labels it the McNemar test for correlated proportions. Using SPSS, the procedure produces a chi-square test statistic.

We will illustrate this using the "war.sav" file. This file contains the data from a study examining changes in attitudes regarding the likelihood of war. In both June and October of 1948, subjects were asked to indicate whether or not they expected a war in the next ten years. The "war.sav" data file is coded so that a 2 represents "Expects War" and a 1 represents "Does Not Expect War." Let us test this at the .05 level of significance.

To test the hypothesis of equal proportions, open the data file and:

1. Click on **Analyze** from the menu bar.
2. Click on **Nonparametric Tests** from the pull-down menu.
3. Click on **2 Related Samples** from the pull-down menu. This opens the Two-Related-Samples Tests dialog box (Fig. 10.9).
4. Click on the variable name "June." It will appear as Variable 1 in the Current Selections box.
5. Click on the variable name "October." It will appear as Variable 2 in the Current Selections box.
6. Move the paired variables to the Test Pair(s) List box by clicking on the **right arrow button**.
7. In the Test Type box, click off the **Wilcoxon** box and click on the **McNemar** box.
8. Click on **OK**.

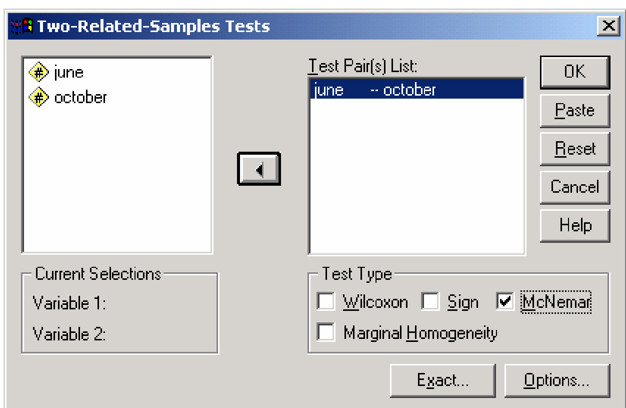

Figure 10.9 Two-Related-Samples Test Dialog Box

june & october

june	october 1	october 2
1	194	45
2	147	211

Test Statistics[b]

	june & october
N	597
Chi-Square[a]	53.130
Asymp. Sig.	.000

a. Continuity Corrected

b. McNemar Test

Figure 10.10 McNemar Test for Likelihood of War

The output (Fig. 10.10) displays the data in the two-way frequency table. We see that 45 people who did not think there would be war when questioned in June changed their minds for the October polling. Similarly, there were 147 individuals who in June thought there would be war, but in October did not believe this to be the case. The listing also reports the chi-square test statistic (53.130) and the corresponding P value ($P < .0005$). Therefore, we would reject the null hypothesis of equal proportions at any α level greater than .0005.

Note that the McNemar test is a two-tailed test. To perform a one-tailed test, compare $P/2$ to α and check that the sample proportions are in the direction indicated by the alternative hypothesis.

Chapter Exercises

10.1 Use SPSS and the "noise.sav" data file to:

- **a.** Test the null hypothesis that the average speed of automobiles is equal to 35 mph. Use $\alpha = .05$, and state the test statistic and conclusion.
- **b.** Would your decision change if you used $\alpha = .10$?

10.2 Use SPSS and the "football.sav" data file to complete the following:

a. Find the 90% confidence interval for the number of points by which football games are won. State the sample mean as well.

b. Would you reject the null hypothesis H_0: $\mu = 10$ points using $\alpha = .10$? (Hint: refer to your conclusions in part (a).)

c. What are your conclusions for the hypothesis H_0: $\mu \leq 10$ points, using $\alpha = .10$?

d. What is the *P* value for the hypothesis in part (c)?

10.3 Use SPSS and the "bodytemp.sav" data file, which contains body temperature (in degrees Fahrenheit) for 130 adults, to complete the following:

a. What is the *P* value for the test of the null hypothesis that the average body temperature is 98.6° Fahrenheit?

b. Would you reject the hypothesis at $\alpha = .05$? $\alpha = .01$?

10.4 Use the "popular.sav" data file to complete the following:

a. Estimate the proportion of students who stated that their goal was to "make good grades."

b. Using $\alpha = .10$, test the hypothesis that the proportion of students wanting to make good grades is greater than 50%. (Hint: You my need to recode the "goals" variable, or use the cut point option.)

10.5 Using the "conform.sav" data file, answer the following questions:

a. Are wives more comformist, on average, than their husbands (use $\alpha = .05$)?

b. What is the minimum α for which you would reject the null hypothesis indicated in part (a)?