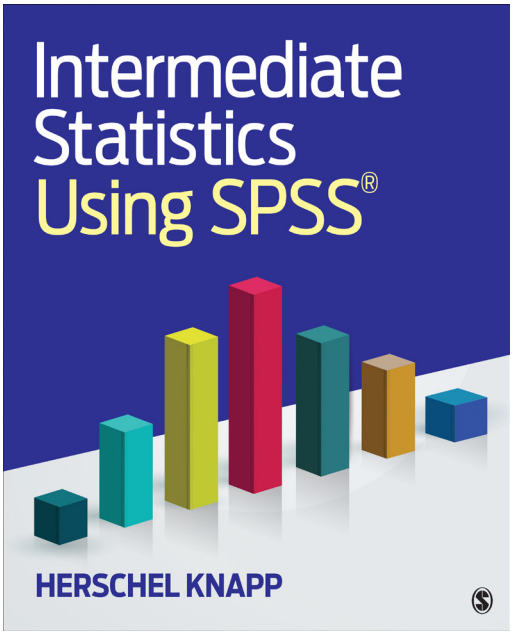


C H A P T E R 8

Paired *t* Test and Wilcoxon Test

Solutions to Odd-Numbered Exercises

EXERCISE 8.1, DATA SET A

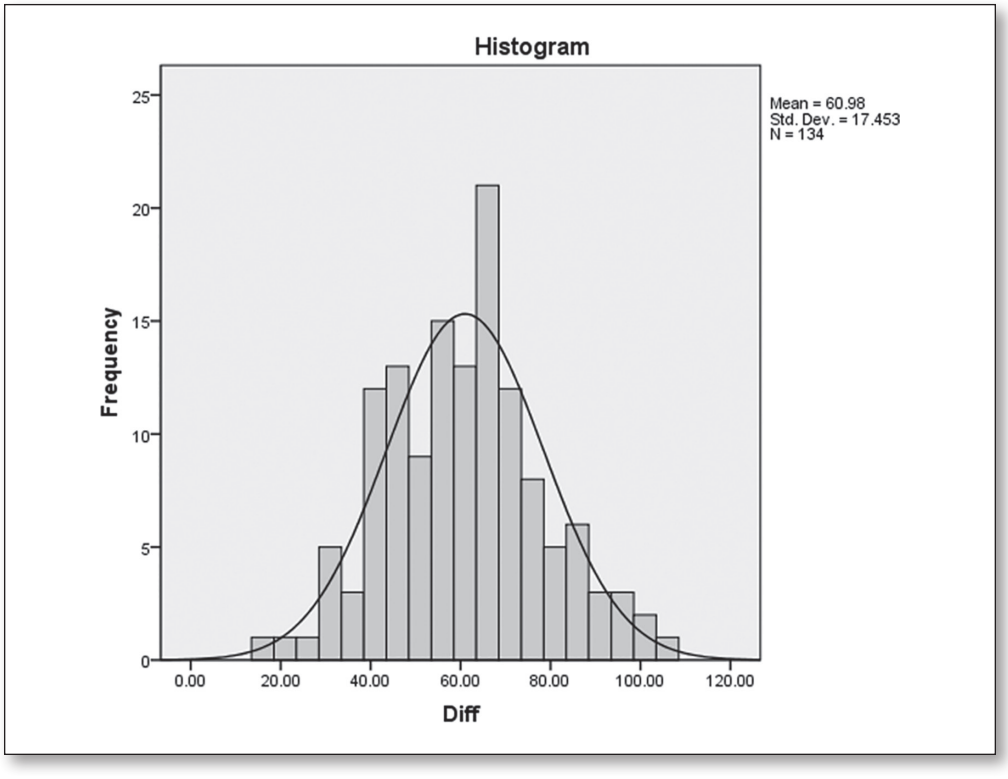
(a)

H_0: Data entry time will not change.

H_1: Data entry time will change.

(b)

After computing the difference between the pretest score and posttest score (*Diff* = *Week1* – *Week2*), a histogram with normal curve was plotted for this difference (*Diff*). The graph below presents a symmetrical (bell-shaped) normal curve for *Diff*, thus satisfying this criterion.



(c)

The paired *t* test revealed the following:

Paired Samples Statistics

		Mean	N	Std. Deviation	Std. Error Mean
Pair 1	Week1	301.89	134	14.550	1.257
	Week2	240.91	134	12.636	1.092

Paired Samples Test

		Paired Differences					t	df	Sig. (2-tailed)
		Mean	Std. Deviation	Std. Error Mean	95% Confidence Interval of the Difference Lower	Upper			
Pair 1	Week1 - Week2	60.978	17.453	1.508	57.995	63.960	40.443	133	.000

During week 1, it required a mean of 302 seconds for employees to make their entries on the website; by week 2, that was down to a mean of 241 seconds. This 61 second reduction is statistically significant since the *p* value of < .001 is less than the specified α level of .05. Based on these findings, we would reject H_0, and not reject H_1.

(d)

In order to assess the efficiency of a new website for employees to enter their hours, we gathered data on 134 employees tracking how long it took them to make the necessary entries correctly. In the first week, mean completion time was 302 seconds; by week 2, that was down to 241 seconds. This 61 second reduction is statistically significant ($p < .001$) suggesting a relatively brief learning curve for this website. Based on these findings, we reject H_0 and accept H_1.

EXERCISE 8.1, DATA SET B

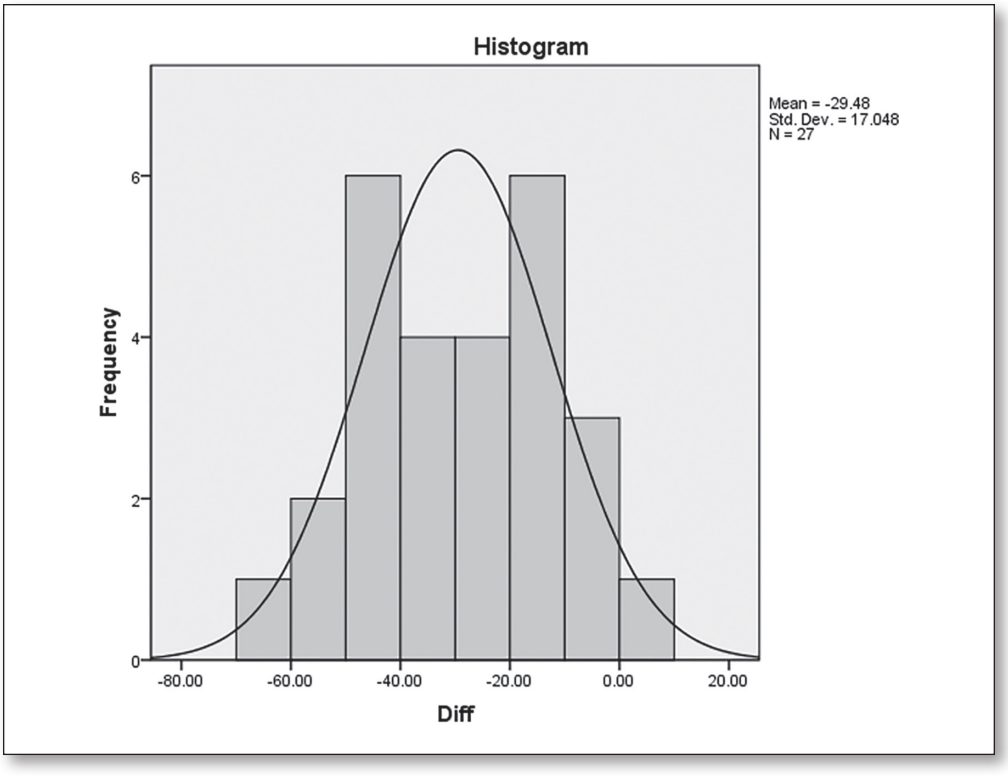
(a)

H_0: Data entry time will not change.

H_1: Data entry time will change.

(b)

After computing the difference between the pretest score and posttest score (*Diff* = *Week1* – *Week2*), a histogram with normal curve was plotted for this difference (*Diff*). The graph below presents a symmetrical (bell-shaped) normal curve for *Diff*, thus satisfying this criterion.



(c)

The paired *t* test revealed the following:

Paired Samples Statistics

		Mean	N	Std. Deviation	Std. Error Mean
Pair 1	Week1	301.78	27	10.282	1.979
	Week2	331.26	27	13.008	2.503

Paired Samples Test

		Paired Differences					t	df	Sig. (2-tailed)
		Mean	Std. Deviation	Std. Error Mean	95% Confidence Interval of the Difference Lower	Upper			
Pair 1	Week1 - Week2	-29.481	17.048	3.281	-36.226	-22.737	-8.986	26	.000

During week 1, it required a mean of 302 seconds for employees to make their entries on the website; by week 2, that was up to a mean of 331 seconds. This 30 second increase is statistically significant since the p value of $< .001$ is less than the specified α level of .05. Based on these findings, we would reject H_0, and not reject H_1.

(d)

In order to assess the efficiency of a new website for employees to enter their hours, we gathered data on 27 employees tracking how long it took them to make the necessary entries correctly. In the first week, mean completion time was 302 seconds; by week 2, that was up to 331 seconds. This statistically significant 9.6% increase in data entry time was unexpected ($p < .001$). We will be reviewing the functionality of this website and our training protocol to better comprehend and improve this outcome. Based on these findings, we reject H_0 and accept H_1.

EXERCISE 8.3, DATA SET A

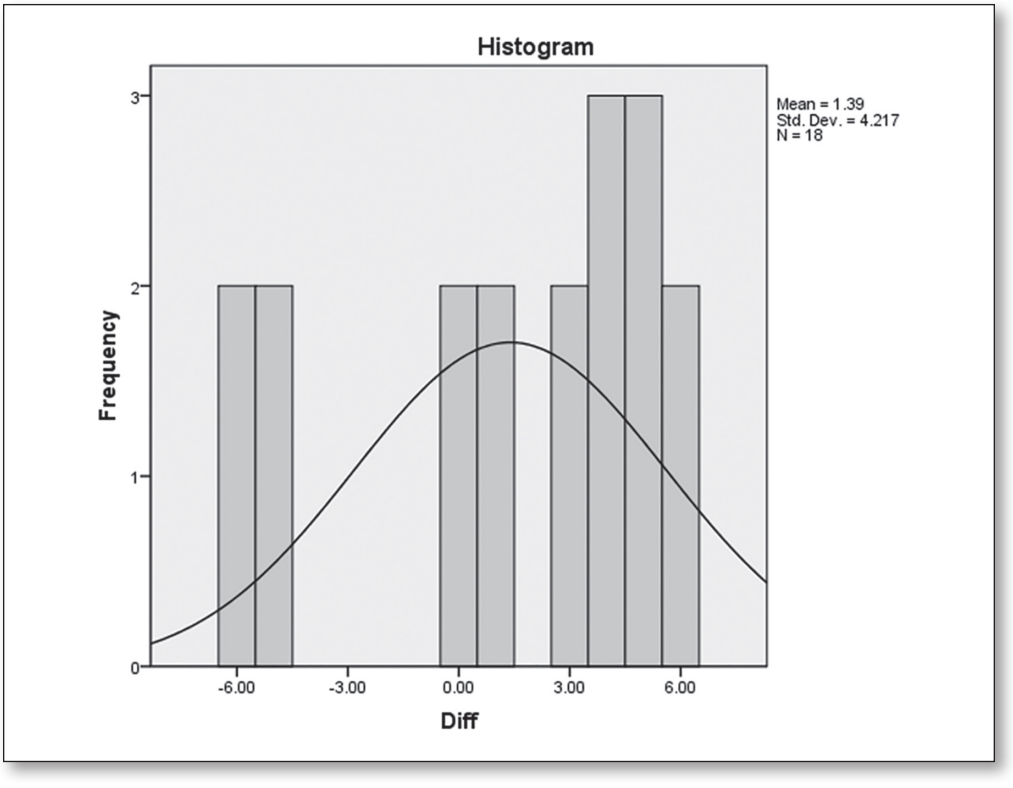
(a)

H_0: Short-term therapy is effective in reducing depression.

H_1: Short-term therapy is ineffective in reducing depression.

(b)

After computing the difference between the pretest score and posttest score (*Diff* = *Baseline* – *Week5*), a histogram with normal curve was plotted for this difference (*Diff*). The graph below presents a symmetrical (bell-shaped) normal curve for *Diff*, thus satisfying this criterion.



(c)

The paired *t* test revealed the following:

Paired Samples Statistics

		Mean	N	Std. Deviation	Std. Error Mean
Pair 1	Baseline	53.83	18	8.893	2.096
	Week05	52.44	18	10.782	2.541

Paired Samples Test

		Paired Differences					t	df	Sig. (2-tailed)
		Mean	Std. Deviation	Std. Error Mean	95% Confidence Interval of the Difference Lower	Upper			
Pair 1	Baseline - Week05	1.389	4.217	.994	-.708	3.486	1.397	17	.180

The mean baseline score on the ADI was 53.83; after 5 weeks of therapy, it was 52.44. This 1.39-point decrease is statistically insignificant ($p = .180$, $\alpha = .05$). Based on these findings, we would not reject H_0, and reject H_1.

(d)

To determine the effectiveness of short-term therapy to reduce depression, we gathered baseline depression data from 18 participants using the Acme Depression Inventory (ADI), which renders a score from 0 to 75 (0 = low depression . . . 75 = high depression). We readministered the ADI after 5 weeks of the therapy. We detected a statistically insignificant 2.6% drop in mean depression score from 53.83 to 52.44 ($p = .180$, $\alpha = .05$). Based on these findings, we would not reject H_0, and reject H_1. Considering the depression is moving in the right direction, we will continue to research this topic and refine our methods in the hope of enhancing the strength of this treatment protocol.

EXERCISE 8.3, DATA SET B

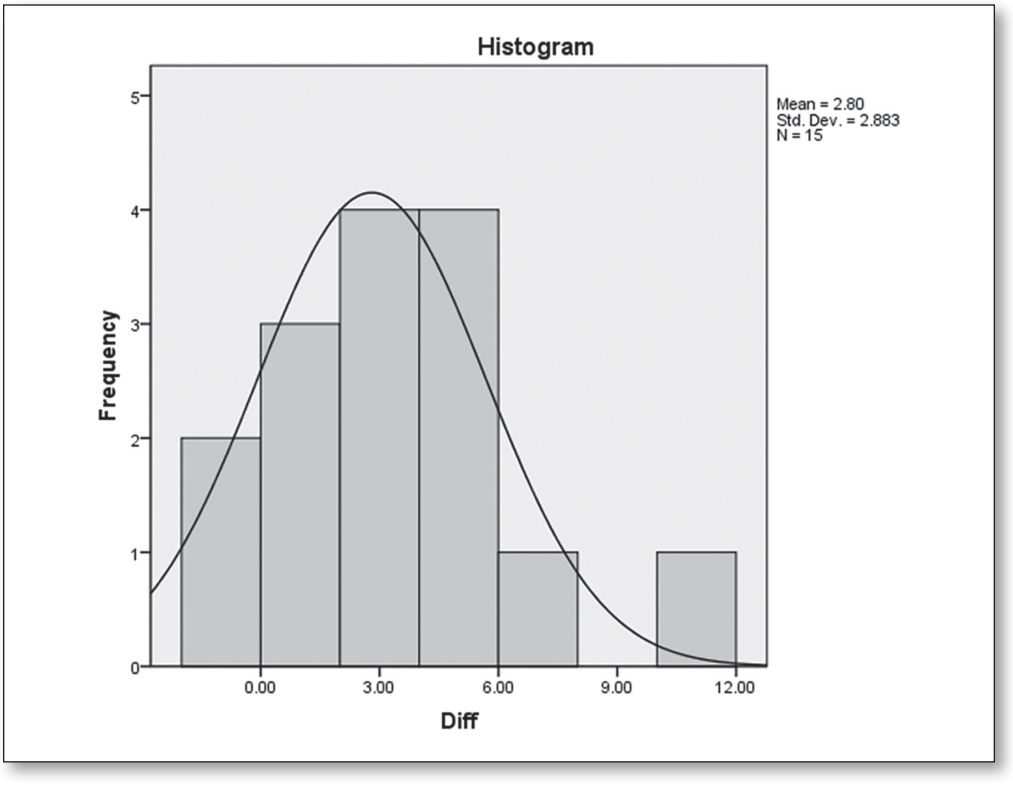
(a)

H_0: Short-term therapy is effective in reducing depression.

H_1: Short-term therapy is ineffective in reducing depression.

(b)

After computing the difference between the pretest score and posttest score (*Diff = Baseline – Week5*), a histogram with normal curve was plotted for this difference (*Diff*). The graph below presents a symmetrical (bell-shaped) normal curve for *Diff*, hence this criterion is not satisfied. The Wilcoxon test will be used here.



(c)

The paired *t* test revealed the following:

Test Statistics[b]

	Week05 - Baseline
Z	-2.995[a]
Asymp. Sig. (2-tailed)	.003

a. Based on positive ranks.
b. Wilcoxon Signed Ranks Test

NOTE: The Wilcoxon test computed the p value of .003, but the output does not include the means for the variables. To facilitate the documentation process, run descriptive statistics for Baseline and Week05:

Statistics

		Baseline	Week05
N	Valid	15	15
	Missing	0	0
Mean		69.80	67.00
Median		70.00	66.00
Mode		70	65[a]
Std. Deviation		2.007	3.485
Variance		4.029	12.143
Range		6	13
Minimum		67	61
Maximum		73	74

a. Multiple modes exist. The smallest value is shown

The mean baseline score on the ADI was 69.8; after 5 weeks of therapy, it was 67.0. We found this 2.8-point decrease in ADI to be statistically significant ($p = .003$, $\alpha = .05$). Based on these findings, we would reject H_0, and not reject H_1.

(d)

To determine the effectiveness of short-term therapy to reduce depression, we gathered baseline depression data from 15 participants using the Acme Depression Inventory (ADI), which renders a score from 0 to 75 (0 = low depression . . . 75 = high depression). We readministered the ADI after 5 weeks of the therapy. We detected a statistically significant 4% drop in mean depression score from 69.8 to 67.0 ($p = .003$, $\alpha = .05$). Based on these findings, we would reject H_0, and not reject H_1. We will continue to research this topic and refine our methods in the hope of enhancing the strength of this treatment protocol.

EXERCISE 8.5, DATA SET A

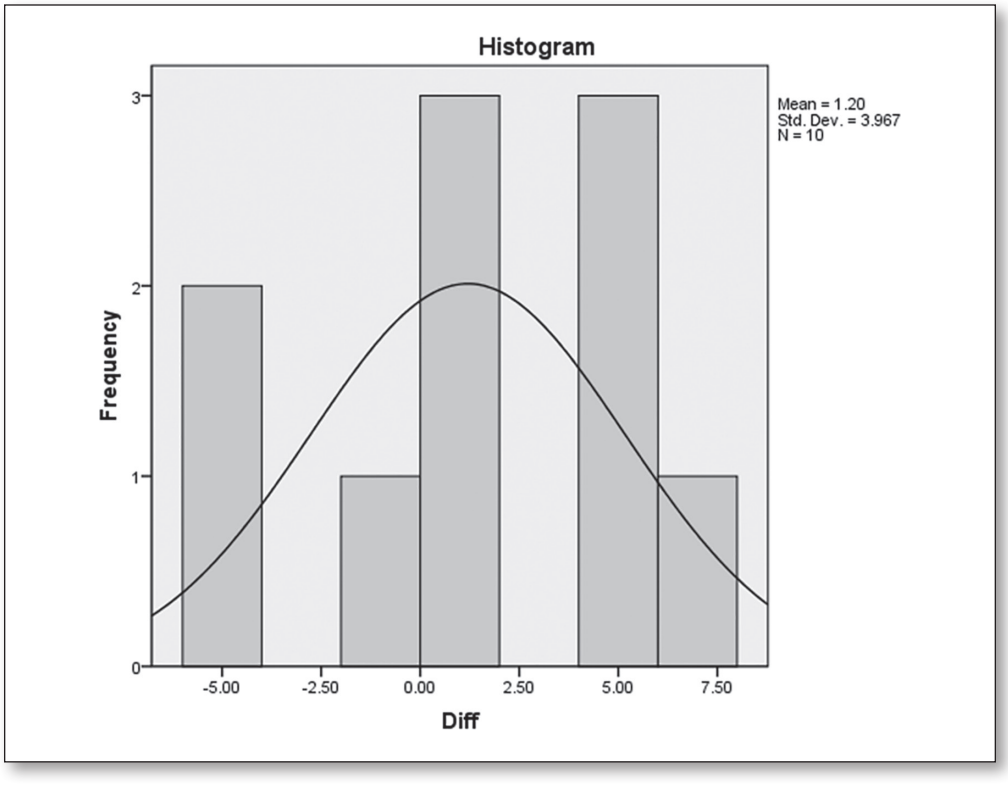
(a)

H_0: Providing free coffee has no effect on productivity.

H_1: Providing free coffee has an effect on productivity.

(b)

After computing the difference between the pretest score and posttest score (*Diff = Productivity_1 - Productivity_2*), a histogram with normal curve was plotted for this difference (*Diff*). The graph below presents a symmetrical (bell-shaped) normal curve for *Diff*, thus satisfying this criterion.



(c)

The paired *t* test revealed the following:

Paired Samples Statistics

		Mean	N	Std. Deviation	Std. Error Mean
Pair 1	Productivity_1	80.00	10	2.357	.745
	Productivity_2	78.80	10	2.974	.940

Paired Samples Test

		Paired Differences					t	df	Sig. (2-tailed)
		Mean	Std. Deviation	Std. Error Mean	95% Confidence Interval of the Difference Lower	Upper			
Pair 1	Productivity_1 - Productivity_2	1.200	3.967	1.254	-1.637	4.037	.957	9	.364

Before introducing the coffee, mean weekly employee productivity was at 80.0, compared to 78.80 after the coffee. This 1.20 drop in mean productivity is notable, but not statistically significant ($p = .364$). Based on these findings, we would accept H_0, and reject H_1.

(d)

In an effort to boost weekly productivity, a manger introduced free gourmet coffee to all (10) employees. Mean productivity went from 80.0 in the week before the coffee to 78.0 in the week after introducing the coffee but it is statistically insignificant ($p = .364$, $\alpha = .05$). These findings suggest that either the coffee was distracting or perhaps factors other than coffee have more influence over productivity than expected. Based on this, I would accept H_0 and reject H_1.

EXERCISE 8.5, DATA SET B

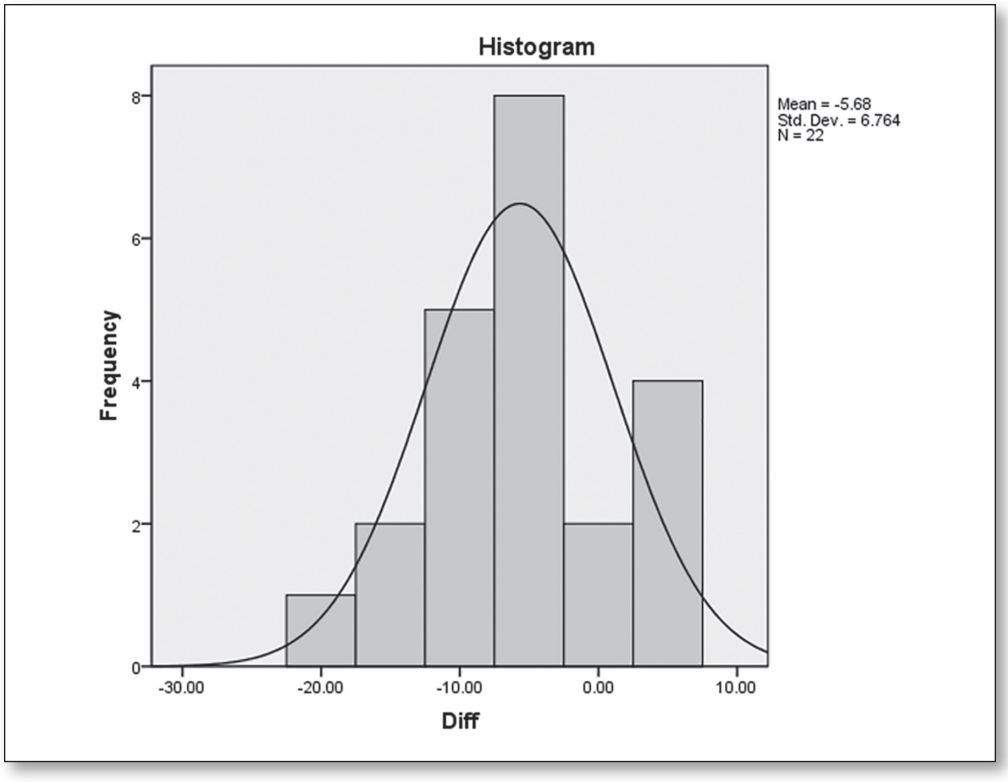
(a)

H_0: Providing free coffee has no effect on productivity.

H_1: Providing free coffee has an effect on productivity.

(b)

After computing the difference between the pretest score and posttest score (*Diff = Productivity_1 - Productivity_2*), a histogram with normal curve was plotted for this difference (*Diff*). The graph below presents a symmetrical (bell-shaped) normal curve for *Diff*, thus satisfying this criterion.



(c)

The paired *t* test revealed the following:

Paired Samples Statistics

		Mean	N	Std. Deviation	Std. Error Mean
Pair 1	Productivity_1	181.00	22	5.674	1.210
	Productivity_2	186.68	22	6.806	1.451

Paired Samples Test

		Paired Differences					t	df	Sig. (2-tailed)
		Mean	Std. Deviation	Std. Error Mean	95% Confidence Interval of the Difference Lower	Upper			
Pair 1	Productivity_1 - Productivity_2	-5.682	6.764	1.442	-8.681	-2.683	-3.940	21	.001

Before introducing the coffee, mean weekly employee productivity was at 181.00; compared to 186.68 after the coffee. This 5.68 increase in productivity is statistically significant ($p = .001$). Based on these findings, we would reject H_0, and accept H_1.

(d)

In an effort to boost weekly productivity, a manger introduced free gourmet coffee to all (22) employees. Mean productivity went from 181.00 in the week before the coffee to 186.68 in the week after introducing the coffee. This 3.13% increase is statistically significant ($p = .001$, $\alpha = .05$). Based on this, I would reject H_0 and accept H_1. Our goal is to continue offering the free coffee and tracking weekly productivity.

EXERCISE 8.7, DATA SET A

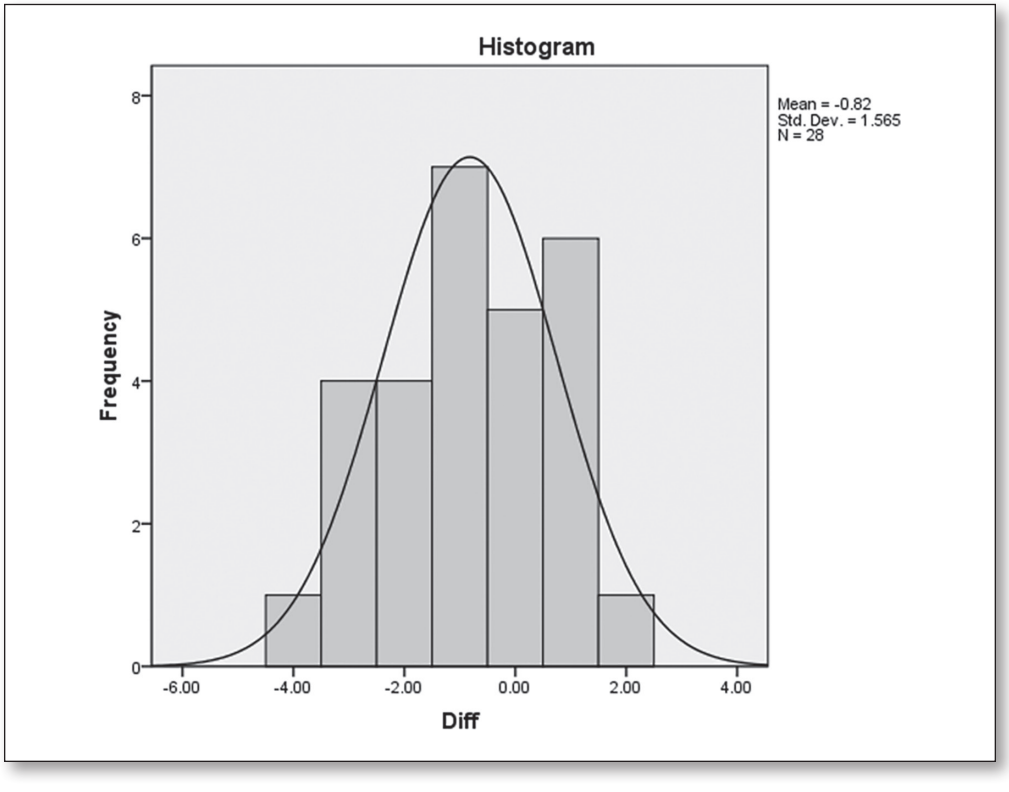
(a)

H_0: Media influences voter opinion.

H_1: Media does not influence voter opinion.

(b)

After computing the difference between the pretest score and posttest score (*Diff* = *Opinion1_baseline* - *Opinion2_audio*), a histogram with normal curve was plotted for this difference (*Diff*). The graph below presents a symmetrical (bell-shaped) normal curve for *Diff*, thus satisfying this criterion.



(c)

The paired *t* test revealed the following:

Paired Samples Statistics

		Mean	N	Std. Deviation	Std. Error Mean
Pair 1	Opinion1_Baseline	2.64	28	1.311	.248
	Opinion2_Audio	3.46	28	1.232	.233

Paired Samples Test

		Paired Differences					t	df	Sig. (2-tailed)
		Mean	Std. Deviation	Std. Error Mean	95% Confidence Interval of the Difference Lower	Upper			
Pair 1	Opinion1_Baseline - Opinion2_Audio	-.821	1.565	.296	-1.428	-.215	-2.778	27	.010

The baseline mean of the focus group is 4.02; after hearing the audio advertisement, their mean went up to 4.18; since the p value of .010 is less than .05, this is a statistically significant difference. As such, we would reject H_0, and not reject H_1.

(d)

A political consultant convened a focus group consisting of 28 registered voters to evaluate the effectiveness of an audio (radio) commercial promoting a candidate. Prior to running any media, the participants were asked one question: *Do you intend to vote for Jones in the upcoming election?* The participants responded using a 7-point scale (1 = absolutely will not vote for Jones . . . 7 = absolutely will vote for Jones). The baseline mean of the focus group is 2.64; after hearing the audio advertisement, their mean significantly increased by 31% to 3.46 ($p = .010$, $\alpha = .05$). As such, we would reject H_0, and not reject H_1.

EXERCISE 8.7, DATA SET B

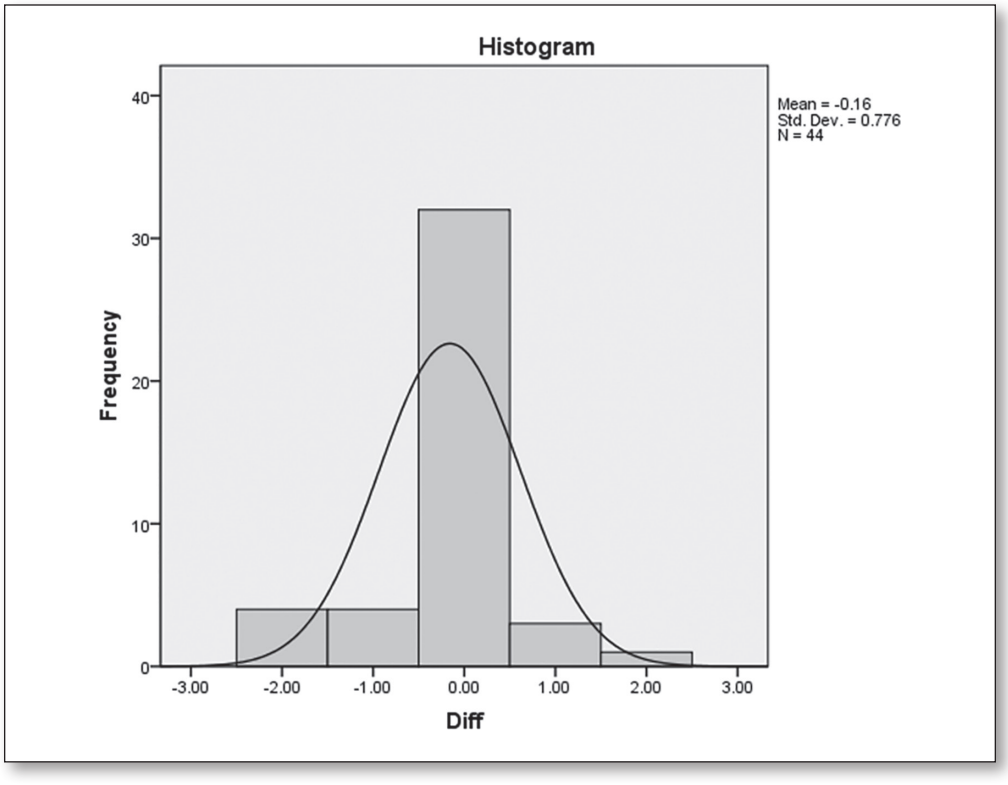
(a)

H_0: Media influences voter opinion.

H_1: Media does not influence voter opinion.

(b)

After computing the difference between the pretest score and posttest score (*Diff* = *Opinion1_baseline* - *Opinion2_audio*), a histogram with normal curve was plotted for this difference (*Diff*). The graph below presents a symmetrical (bell-shaped) normal curve for *Diff*, thus satisfying this criterion.



(c)

The paired *t* test revealed the following:

Paired Samples Statistics

		Mean	N	Std. Deviation	Std. Error Mean
Pair 1	Opinion1_Baseline	4.02	44	1.607	.242
	Opinion2_Audio	4.18	44	1.483	.224

Paired Samples Test

		Paired Differences					t	df	Sig. (2-tailed)
		Mean	Std. Deviation	Std. Error Mean	95% Confidence Interval of the Difference Lower	Upper			
Pair 1	Opinion1_Baseline - Opinion2_Audio	-.159	.776	.117	-.395	.077	-1.360	43	.181

The baseline mean of the focus group is 4.02; after hearing the audio advertisement, their mean went up to 4.18; since the p value of .181 is greater than .05, this is not a statistically significant difference. As such, we would not reject H_0, and reject H_1.

(d)

A political consultant convened a focus group consisting of 44 registered voters to evaluate the effectiveness of an audio (radio) commercial promoting a candidate. Prior to running any media, the participants were asked one question: *Do you intend to vote for Jones in the upcoming election?* The participants responded using a 7-point scale (1 = absolutely will not vote for Jones . . . 7 = absolutely will vote for Jones). The baseline mean of the focus group is 4.02; after hearing the audio advertisement, their mean increased by 3.98% to 4.18 ($p = .181$, $\alpha = .05$). As such, we would not reject H_0, and reject H_1.

EXERCISE 8.9, DATA SET A

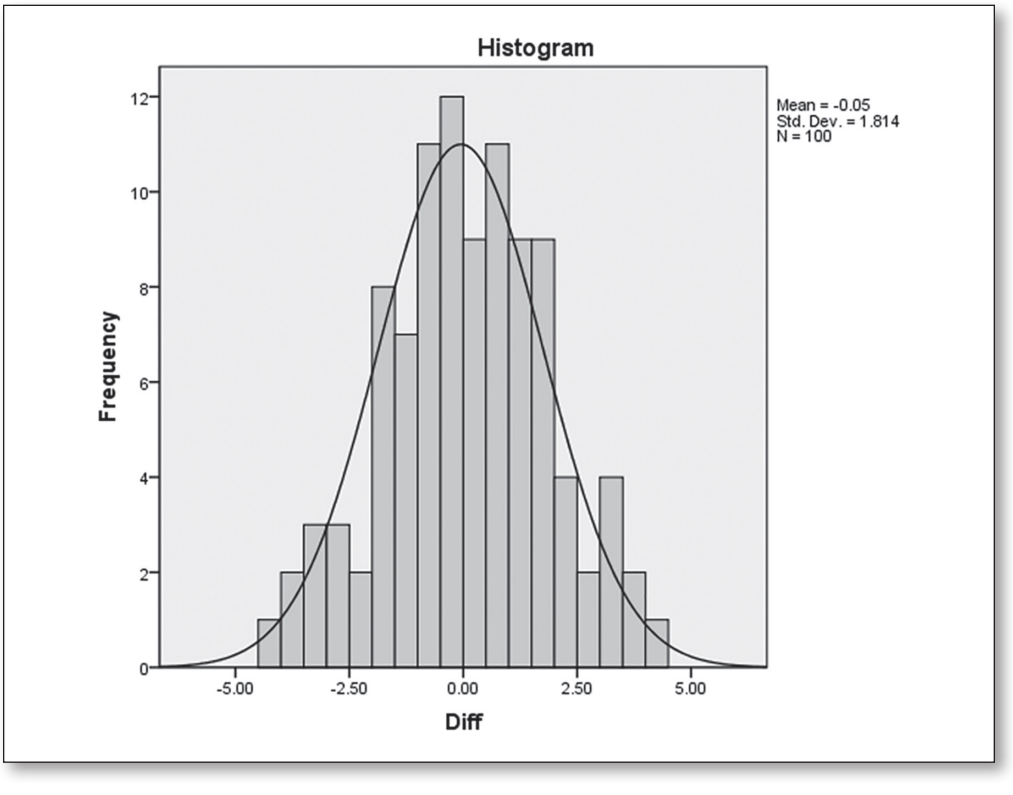
(a)

H_0: *Zzzleep Zzzound* has no effect on sleep hours.

H_1: *Zzzleep Zzzound* has an effect on sleep hours.

(b)

After computing the difference between the pretest score and posttest score (*Diff = Sleep1 - Sleep2*), a histogram with normal curve was plotted for this difference (*Diff*). The graph below presents a symmetrical (bell-shaped) normal curve for *Diff*, thus satisfying this criterion.



(c)

The paired *t* test revealed the following:

Paired Samples Statistics

		Mean	N	Std. Deviation	Std. Error Mean
Pair 1	Sleep1	7.3050	100	1.21053	.12105
	Sleep2	7.3600	100	1.24839	.12484

Paired Samples Test

		Paired Differences					t	df	Sig. (2-tailed)
		Mean	Std. Deviation	Std. Error Mean	95% Confidence Interval of the Difference Lower	Upper			
Pair 1	Sleep1 - Sleep2	-.05500	1.81380	.18138	-.41490	.30490	-.303	99	.762

The mean sleep time during the baseline week was 7.305 hours per night; during the first week of use, that mean went up to 7.360. Since the p value of .762 is greater than the α level of .05, this is not a statistically significant increase; hence, we would not reject H_0, and reject H_1.

(d)

The *Zzzleep Zzzound* app provides audio selections (e.g., gentle rain, ocean waves, soothing music) to help induce peaceful sleep. During the night, the app uses the camera and motion sensor to gather sleep data. If the user wakes during the night, the app senses it and plays the selected sound for 10 minutes. For the first week, the app runs without any audio to gather baseline data. Once a week, the software transmits the mean sleep time per night for that week to the sleep researcher's database. Data gathered on 100 subscribers revealed a mean of 7.305 hours of sleep during the baseline week, followed by 7.360 hours in the next week. This .75% increase is statistically insignificant ($p = .792$, $\alpha = .05$). Per these findings, we will review our research on sleep science, which will inform our next release of the app.

EXERCISE 8.9, DATA SET B

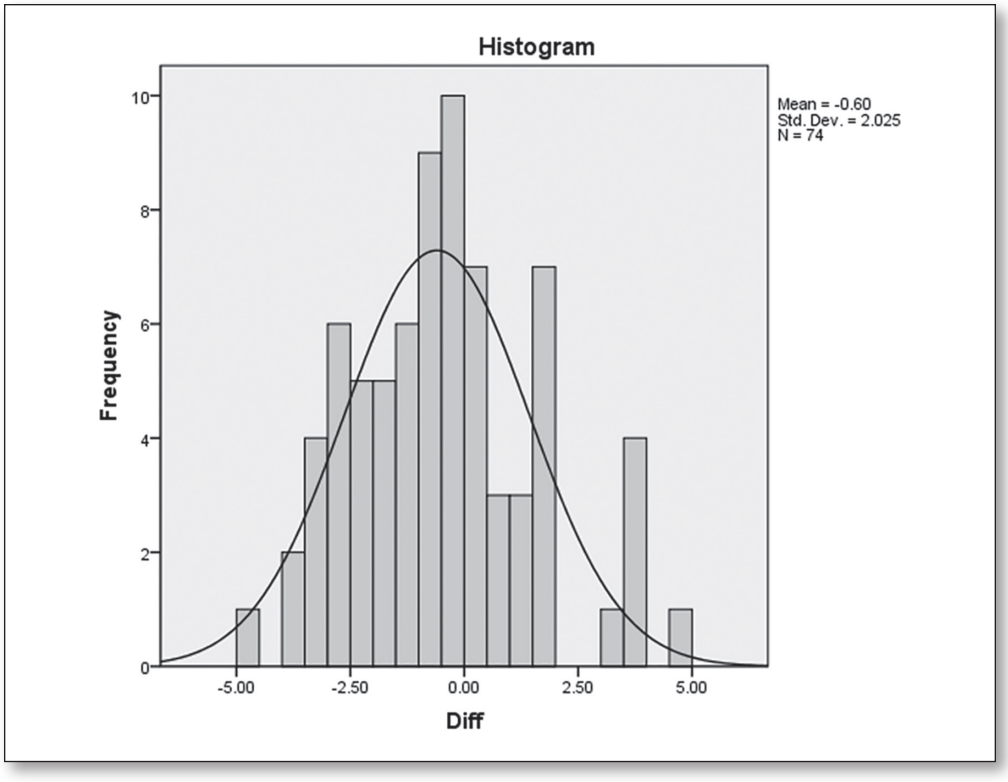
(a)

H_0: *Zzzleep Zzzound* has no effect on sleep hours.

H_1: *Zzzleep Zzzound* has an effect on sleep hours.

(b)

After computing the difference between the pretest score and posttest score (*Diff = Sleep1 - Sleep2*), a histogram with normal curve was plotted for this difference (*Diff*). The graph below presents a symmetrical (bell-shaped) normal curve for *Diff*, thus satisfying this criterion.



(c)

The paired *t* test revealed the following:

Paired Samples Statistics

		Mean	N	Std. Deviation	Std. Error Mean
Pair 1	Sleep1	6.9324	74	1.42825	.16603
	Sleep2	7.5338	74	1.39736	.16244

Paired Samples Test

		Paired Differences					t	df	Sig. (2-tailed)
		Mean	Std. Deviation	Std. Error Mean	95% Confidence Interval of the Difference Lower	Upper			
Pair 1	Sleep1 - Sleep2	-.60135	2.02507	.23541	-1.07052	-.13218	-2.554	73	.013

The mean sleep time during the baseline week was 6.9324 hours per night; during the first week of use, that mean went up to 7.5338. Since the p value of .013 is less than the α level of .05, this is a statistically significant increase; hence, we would reject H_0, and not reject H_1.

(d)

The *Zzzleep Zzzound* app provides audio selections (e.g., gentle rain, ocean waves, soothing music) to help induce peaceful sleep. During the night, the app uses the camera and motion sensor to gather sleep data. If the user wakes during the night, the app senses it and plays the selected sound for 10 minutes. For the first week, the app runs without any audio to gather baseline data. Once a week, the software transmits the mean sleep time per night for that week to the sleep researcher's database. Data gathered on 74 subscribers revealed a mean of 6.9324 hours of sleep during the baseline week, followed by 7.5338 hours in the next week. This 8.7% increase is statistically significant ($p = .013$, $\alpha = .05$).