Confidence Intervals and Hypothesis Tests: Two Samples

9.1 Z-Interval to Compare Two Population Means: Independent Samples

To complete this section of homework watch Chapter Nine, Lecture Examples: 131 and 132.

1. A study comparing hotel rates in Rome and Paris was done using 50 rooms randomly selected from Rome and 50 rooms randomly selected from Paris. The average price for a room in Rome was 100 Euros per night with a standard deviation of 6.25 Euros, and the average price for a room in Paris was 150 Euros per night with a standard deviation of 9.37 Euros. Find the 98% confidence interval for the true mean difference between hotel room rates in Rome and Paris. Is there a significant difference between the rates? VS

2. Is there a difference between the problem solving skills of math majors and business majors? A test was given to both groups the results are summarized below. Find a 95% confidence interval for the true mean difference between the scores of math majors and business majors. Is there a significant difference? If so who does better? VS

Math majors	Business majors
$n = 38$	$n = 42$
$\bar{x} = 84.6$	$\bar{x} = 64.6$
$s = 4.4$	$s = 6.3$

3. A researcher theorized that average heights of boys and girls of the same age prior to undergoing puberty should be the same. Fifty-eight randomly selected boys of age 8 had an average height of 123.5 cm with a variance of 98cm. Fifty-five randomly selected girls had an average height of 126.1 cm with a variance of 119cm. Find the 99% confidence interval for the true mean difference between boys and girls heights at age 8. Is there a significant difference? VS

4. Which batteries last longer? A randomly selected sample of 92 Duralife AA batteries had an average lifespan of 46.2 hrs with a standard deviation of 2.67 hrs. A randomly selected sample of 91 Energy AA batteries had an average lifespan of 42.9hrs with a standard deviation of 3.17 hrs. Find the 90% confidence interval for the true mean difference between the lifespans. Is there a significant difference?

: indicates the exercise has a video devoted to it in the corresponding section of STATSprofessor.com

5. Who gets better grades: males or females? A randomly selected sample of 61 female students had an average college gpa of 3.12 points with a standard deviation of 0.32 points. A randomly selected sample of 64 male students had an average college gpa of 2.9 points with a standard deviation of 0.34 points. Find the 95% confidence interval for the true mean difference between female and male college gpas. Is there a significant difference?

6. A large sample confidence interval for the true average difference between IQ's of male students and female students was created. The result was as follows: $-11 < \mu_m - \mu_f < 9$. Is there a significant difference? Which group had the larger sample mean? VS

7. A large sample confidence interval for the true average difference between the maximum bench press of eighteen year old males and 25 year old males (in the US army) was created. The result was as follows: $-21 < \mu_{18} - \mu_{25} < -19$. Is there a significant difference? Which group had the larger sample mean? VS

8. A large sample confidence interval for the true average difference between vertical leap of male track athletes and the vertical leap female track athletes was created. The result was as follows: $3 < \mu_m - \mu_f < 5$. Is there a significant difference? Which group had the larger sample mean? VS

9.1 Answers

1. $[-53.71, -46.29]$; Yes, Paris is significantly more expensive.
2. $[17.64, 22.36]$; Yes, Math majors do significantly better.
3. $[-7.7, 2.5]$; Since zero is inside the interval we cannot conclude a significant difference exists because the mean difference could be zero.
4. $[2.6, 4.0]$; Yes, Duralife batteries last significantly longer.
5. $[0.10, 0.34]$; Yes, females score significantly higher.
6. No significant difference, but females had the higher sample IQ because the subtraction went male – female since there is a bigger number on the negative side (absolute value) than the positive side it means the mean for females must have been larger.
7. Yes, the 25 year olds are significantly stronger.
8. Yes, males jump significantly higher.

Need more exercises?

: indicates the exercise has a video devoted to it in the corresponding section of STATSprofessor.com

9.2 Z-Test to Compare Two Population Means: Independent Samples

To complete this section of homework watch Chapter Nine, Lecture Examples 133 and 134.

9. A study comparing hotel rates in Rome and Paris was done using 50 rooms randomly selected from Rome and 50 rooms randomly selected from Paris. The average price for a room in Rome was 100 Euros per night with a standard deviation of 6.25 Euros, and the average price for a room in Paris was 150 Euros per night with a standard deviation of 9.37 Euros. Use a 1% significance level to test the claim that Parisian rooms are more expensive on average than Roman rooms.

10. Is there a difference between the problem solving skills of math majors and business majors? A test was given to both groups the results are summarized below. Use a 2.5% significance level to test the claim that math majors are better problem solvers on average than business majors.

Math majors	Business majors
$n = 38$	$n = 42$
$\bar{x} = 84.6$	$\bar{x} = 64.6$
$s = 4.4$	$s = 6.3$

11. A researcher theorized that average heights of boys and girls of the same age prior to undergoing puberty should be the same. Fifty-eight randomly selected boys of age 8 had an average height of 123.5 cm with a variance of 98cm. Fifty-five randomly selected girls had an average height of 126.1 cm with a variance of 119cm. Use a 1% significance level and the p-value method to test the claim that there is no difference between the average heights of boys and girls at age eight. VS

12. Which batteries last longer? A randomly selected sample of 92 Duralife AA batteries had an average lifespan of 46.2hrs with a standard deviation of 2.67 hrs. A randomly selected sample of 91 Energy AA batteries had an average lifespan of 42.9hrs with a standard deviation of 3.17 hrs. Use a 5% significance level to test the claim that Duralife batteries last longer on average. VS

13. Who gets better grades: males or females? A randomly selected sample of 61 female students had an average college gpa of 3.08 points with a standard deviation of 0.32 points. A randomly selected sample of 64 male students had an average college gpa of 2.9 points with a standard deviation of 0.34 points. Use a 5% significance level and the p-value method to test the claim that there is a difference between the average gpas of male and female students. VS

: indicates the exercise has a video devoted to it in the corresponding section of STATSprofessor.com

9.2 Answers

9. $Claim: \mu_R < \mu_p$, $\begin{matrix} H_o: \mu_R \geq \mu_p \\ H_a: \mu_R < \mu_p \end{matrix}$, $TestStat: -31.39$, $CriticalValue: -2.326$

Initial Conclusion: Reject the Null, support the alternative

Final Conclusion: The sample data support the claim…

10. $Claim: \mu_M > \mu_B$, $\begin{matrix} H_o: \mu_M \leq \mu_B \\ H_a: \mu_M > \mu_B \end{matrix}$, $TestStat: 16.58$, $CriticalValue: 1.96$

Initial Conclusion: Reject the Null, support the alternative

Final Conclusion: The sample data support the claim…

11. $Claim: \mu_B = \mu_G$, $\begin{matrix} H_o: \mu_B = \mu_G \\ H_a: \mu_B \neq \mu_G \end{matrix}$, $TestStat: -1.32$, $P-Value: 0.1868$

Initial Conclusion: Do not reject the Null, do not support the alternative

Final Conclusion: The sample data does not reject the claim…

12. $Claim: \mu_D > \mu_E$, $\begin{matrix} H_o: \mu_D \leq \mu_E \\ H_a: \mu_D > \mu_E \end{matrix}$, $TestStat: 7.61$, $CriticalValue: 1.645$

Initial Conclusion: Reject the Null, support the alternative

Final Conclusion: The sample data support the claim…

13. $Claim: \mu_F \neq \mu_M$, $\begin{matrix} H_o: \mu_F = \mu_M \\ H_a: \mu_F \neq \mu_M \end{matrix}$, $TestStat: 3.05$, $P-Value: 0.0022$

Initial Conclusion: Reject the Null, support the alternative

Final Conclusion: The sample data supports the claim…

[Need more exercises?]

9.3 t-Interval to Compare Two Population Means: Independent Samples (Equal Variances)

To complete this section of homework watch Chapter Nine, Lecture Example 135.

14. A researcher wants to know if statisticians in the private sector are paid better than statisticians in the public sector. She selects random samples from both areas; the results are summarized below. Form a 95% confidence interval for the average difference between the salaries of government statisticians and private sector statisticians (Assume equal variances). VS

Government	Private
$n = 26$	$n = 28$
$\bar{x} = 35.50$ dollars per hour	$\bar{x} = 54.6$ dollars per hour
$s = 4.16$	$s = 4.4$

15. A parent theorized that his daughter spent more time on her cell phone than his son. He randomly selected 27 calls made by his son and 25 calls made by his daughter from his past phone records. His son's calls had an average length of 23.5 minutes with a variance of 28 minutes. His daughter's calls had an average length of 36.1 minutes with a variance of 35 minutes. Find the 99% confidence interval for the true mean difference between the length of calls made by his son and his daughter (Assume equal variances). Is there a significant difference? VS

16. Who gives a higher real estate assessment? In an experiment, a single home was assessed by both real estate appraisers and local government tax assessors. A randomly selected sample of 22 appraisals done by the real estate appraisers had an average value (in thousands) of $212.1 with a standard deviation of $8.48. A randomly selected sample of 23 appraisals done by tax assessors had an average of $225.3 with a standard deviation of $9.01. Find the 90% confidence interval for the true mean difference between the two types of appraisals (Assume equal variances). Is there a significant difference? VS

17. Who recovers from childbirth faster? A randomly selected sample of 16 insured women spent an average of 2.4 days in the hospital for a routine childbirth with a standard deviation of 0.62 days. Another randomly selected sample of 16 uninsured women spent an average of 1.8 days in the hospital for a routine childbirth with a standard deviation of 0.53 days. Find the 98% confidence interval for the true mean difference between the amount of time spent in the hospital after childbirth for insured women versus uninsured women (Assume equal variances). Is there a significant difference? VS

9.3 Answers

14. $[-21.44 < \mu_G - \mu_P < -16.76]$, $S_p^2 = 18.372308$, $df = 52$; Yes, the private pay is significantly better.

15. $[-16.76 < \mu_S - \mu_D < -8.44]$, $S_p^2 = 31.36$, $df = 50$; Yes, the daughter talks significantly longer.

16. $[-17.59 < \mu_R - \mu_T < -8.81]$, $S_p^2 = 76.65303727$, $df = 43$; Yes, the tax men say the house is significantly more valuable.

17. $[0.099 < \mu_I - \mu_{UI} < 1.101]$, $S_p^2 = 0.3326500005$, $df = 30$; Yes, the insured stay in the hospital a sig amount longer.

Need more exercises?

9.4 t-Test to Compare Two Population Means: Independent Samples (Equal Variances)

To complete this section of homework watch Chapter Nine, Lecture Example 136.

18. Does marijuana use make you slow? A random selection of 28 heavy marijuana users spent an average of 38.28 minutes to complete a set of logic puzzles. Their standard deviation was 4.51 minutes. A random selection of 29 non users spent an average of 25.4 minutes to complete the same set of logic puzzles. Their standard deviation was 3.98 minutes. Use a 0.01 significance level to test the claim that the population of heavy marijuana users takes longer on average to complete the set of problems than non users (assume equal variances). VS
19. Does alcohol impair your visual/motor skills? Two randomly selected groups of 20 people were given either alcohol to drink or placebo, and they then had their visual and motor skills tested. Those who drank alcohol made an average of 4.3 errors with a standard deviation of 2.5 errors. Those who drank the placebo made an average of 1.7 errors with a standard deviation of 0.7 errors. Use a 0.02 significance level to test the claim that there is a difference between the two groups (assume equal variances). Does alcohol lead to more mistakes? VS

20. A study looked at the relationship between low birth weight and IQ. A random selection of 17 low birth weight children had an average IQ of 95.5 points with a standard deviation of 16. A random selection of 19 normal birth weight children had an average IQ of 104.9 points with a standard deviation of 14.2. Use a 0.10 significance level to test the claim that the two groups of children have the same average IQ (assume equal variances). VS

21. Do men and women consume the same average number of vegetable servings per day? Twenty-five men and twenty-five women were randomly selected to keep a food journal. The men ate an average of 2.9 servings of vegetables per day with a standard deviation of 0.6 servings. The women ate an average of 4.3 servings of vegetables per day with a standard deviation of 0.7 servings. Use a 0.05 significance level to test the claim that women eat more servings of vegetables per day than men on average (assume equal variances). VS

9.4 Answers

18. $Claim: \mu_M > \mu_N$, $\begin{matrix} H_o: \mu_M \le \mu_N \\ H_a: \mu_M > \mu_N \end{matrix}$, $TestStat: 11.44$, $CriticalValue: 2.396$, $df = 55$, $S_p^2 =$ 18.04934363

Initial Conclusion: Reject the Null, support the alternative; Final Conclusion: The sample data support the claim…

19. $Claim: \mu_A \ne \mu_P$, $\begin{matrix} H_o: \mu_A = \mu_P \\ H_a: \mu_A \ne \mu_P \end{matrix}$, $TestStat: 4.48$, $CriticalValues: \pm 2.429$, $df = 38$, $S_p^2 =$ 3.370000018

Initial Conclusion: Reject the Null, support the alternative

Final Conclusion: The sample data supports the claim…

20. $Claim: \mu_{LW} = \mu_{NW}$, $\begin{matrix} H_o: \mu_{LW} = \mu_{NW} \\ H_a: \mu_{LW} \ne \mu_{NW} \end{matrix}$, $TestStat: -1.87$, $CriticalValues: \pm 1.691$, $df = 34$, $S_p^2 = 227.2211769$

Initial Conclusion: Reject the Null, support the alternative
Final Conclusion: The sample data allows rejection of the claim…

21. $Claim: \mu_M < \mu_W$, $\begin{matrix} H_o: \mu_M \ge \mu_W \\ H_a: \mu_M < \mu_W \end{matrix}$, $TestStat: -7.59$, $CriticalValue: -1.678$, $df = 48$, $S_p^2 =$ 0.4249999472 (note: I used interpolation to derive the critical value. If you use the table, you will get -1.679.)

Initial Conclusion: Reject the Null, support the alternative

Final Conclusion: The sample data support the claim…

Need more exercises?

: *indicates the exercise has a video devoted to it in the corresponding section of STATSprofessor.com*

9.5 t-Interval to Compare Two Population Means: Independent Samples (Unequal Variances)

To complete this section of homework watch Chapter Nine, Lecture Examples 137 and 137.5.

22. A researcher wants to know if statisticians in the private sector are paid better than statisticians in the public sector. She selects random samples from both areas. The results are summarized below. Form a 95% confidence interval for the average difference between the salaries of government statisticians and private sector statisticians (Do not assume equal variances). VS

Government	Private
$n = 26$	$n = 28$
$\bar{x} = 35.50$ dollars per hour	$\bar{x} = 54.6$ dollars per hour
$s = 4.16$	$s = 4.4$

23. A parent theorized that his daughter spent more time on her cell phone than his son. He randomly selected 27 calls made by his son and 25 calls made by his daughter from his past phone records. His son's calls had an average length of 23.5 minutes with a variance of 28 minutes. While his daughter's calls had an average length of 36.1 minutes with a variance of 35 minutes. Find the 99% confidence interval for the true mean difference between the length of calls made by his son and his daughter (Do not assume equal variances). Is there a significant difference? VS

24. Who gives a higher real estate assessment? In an experiment, a single home was assessed by both real estate appraisers and local government tax assessors. A randomly selected sample of 22 appraisals done by the real estate appraisers had an average value (in thousands) of \$212.1 with a standard deviation of \$8.48. A randomly selected sample of 23 appraisals done by tax assessors had an average of \$225.3 with a standard deviation of \$9.01. Find the 90% confidence interval for the true mean difference between the two types of appraisals (Do not assume equal variances). Is there a significant difference? VS

25. Who recovers from childbirth faster? A randomly selected sample of 16 insured women spent an average of 2.4 days in the hospital for a routine childbirth with a standard deviation of 0.62 days. Another randomly selected sample of 16 uninsured women spent an average of 1.8 days in the hospital for a routine childbirth with a standard deviation of 0.53 days. Find the 98% confidence interval for the true mean difference between the amount of time spent in the hospital after childbirth for insured women versus uninsured women (Do not assume equal variances). Is there a significant difference? VS

9.5 Answers

22. $[-21.44 < \mu_G - \mu_P < -16.76]$, $df = 51.9803 = 51$ (remember we truncate here we do not round); Yes, the private pay is significantly better.

23. $[-16.79 < \mu_S - \mu_D < -8.41]$, $df = 48.27 = 48$; Yes, the daughter talks significantly longer.

24. $[-17.58 < \mu_R - \mu_T < -8.82]$, $df = 42.99 = 42$; Yes, the tax men say the house is significantly more valuable.

25. $[0.098 < \mu_I - \mu_{UI} < 1.102]$, $df = 29.29 = 29$; Yes, the insured stay in the hospital a sig amount longer. Note: here since the sample sizes are the same it is acceptable to use the simpler degrees of freedom: $n_1 + n_2 - 2 = 30$

Need more exercises?

9.6 t-Test to Compare Two Population Means: Independent Samples (Unequal Variances)

To complete this section of homework watch Chapter Nine, Lecture Example 138.

26. Does marijuana use make you slow? A random selection of 28 heavy marijuana users spent an average of 38.28 minutes to complete a set of logic puzzles. Their standard deviation was 4.51 minutes. A random selection of 29 non users spent an average of 25.4 minutes to complete the same set of logic puzzles. Their standard deviation was 3.98 minutes. Use a 0.01 significance level to test the claim that the population of heavy marijuana users takes longer on average to complete the set of problems than non users (Do not assume equal variances). VS

27. Does alcohol impair your visual/motor skills? Two randomly selected groups of 20 people were given either alcohol to drink or placebo, and they then had their visual and motor skills tested. Those who drank alcohol made an average of 4.3 errors with a standard deviation of 2.5 errors. Those who drank the placebo made an average of 1.7 errors with a standard deviation of 0.7 errors. Use a 0.02 significance level to test the claim that there is a difference between the two groups (Do not assume equal variances). Does alcohol lead to more mistakes? VS

28. A study looked at the relationship between low birth weight and IQ. A random selection of 17 low birth weight children had an average IQ of 95.5 points with a standard deviation of 16. A random selection of 19 normal birth weight children had an average IQ of 104.9 points with a standard deviation of 14.2. Use a 0.10 significance level to test the claim that the two groups of children have the same average IQ (Do not assume equal variances). VS

29. Do men and women consume the same average number of vegetable servings per day? Twenty-five men and twenty-five women were randomly selected to keep a food journal. The men ate an average of 2.9 servings of vegetables per day with a standard deviation of 0.6 servings. The women ate an average of 4.3 servings of vegetables per day with a standard deviation of 0.7 servings. Use a 0.05 significance level to test the claim that women eat more servings of vegetables per day than men on average (Do not assume equal variances).

9.6 Answers

26. $Claim: \mu_M > \mu_N$, $\begin{matrix} H_o: \mu_M \le \mu_N \\ H_a: \mu_M > \mu_N \end{matrix}$, $TestStat: 11.42$, $CriticalValue: 2.403$, df =53.63=53, S =4.24845191 [note: I used the more conservative (larger) of the two values on the table—this is an alternative to using interpolation]

Initial Conclusion: Reject the Null, support the alternative

Final Conclusion: The sample data support the claim…

27. $Claim: \mu_A \ne \mu_P$, $\begin{matrix} H_o: \mu_A = \mu_P \\ H_a: \mu_A \ne \mu_P \end{matrix}$, $TestStat: 4.48$, $CriticalValues: \pm 2.429$, df =20+20-2=38,

Initial Conclusion: Reject the Null, support the alternative

Final Conclusion: The sample data supports the claim…

28. $Claim: \mu_{LW} = \mu_{NW}$, $\begin{matrix} H_o: \mu_{LW} = \mu_{NW} \\ H_a: \mu_{LW} \ne \mu_{NW} \end{matrix}$, $TestStat: -1.86$, $CriticalValues: \pm 1.694$, df = 32.257=32,

Initial Conclusion: Reject the Null, support the alternative

Final Conclusion: The sample data allows rejection of the claim…

29. $Claim: \mu_M < \mu_W$, $\begin{matrix} H_o: \mu_M \ge \mu_W \\ H_a: \mu_M < \mu_W \end{matrix}$, $TestStat: -7.59$, $CriticalValues: -1.679$, df =25+25-2 = 48 (note: since both have the same sample size we don't need to use the complicated formula for the degrees of freedom), Initial Conclusion: Reject the Null, support the alternative

Final Conclusion: The sample data support the claim…

Need more exercises?

: indicates the exercise has a video devoted to it in the corresponding section of STATSprofessor.com

9.7 Hypothesis Test to Compare Two Population Means: Matched-Pair Experiments

To complete this section of homework watch Chapter Nine, Lecture Example 139 and 140.

30. A researcher suspects that men, if given the chance, would lie about their heights. She suspects that they would try to say they are taller than they really are. To test this, the researcher first asks the male subjects to report their heights, and then she actually measures them. The subjects do not know they will be measured. Use the data below and a 5% significance level to test the claim that men would report a taller height than what they are in reality. These subjects were between 12 and 16 years old. Do you think this could have affected the results? VS

Reported height	68	71	63	70	71	60	65	64	54	63	66	72
Measured height	67.9	69.9	64.9	68.3	70.3	60.6	64.5	67	55.6	74.2	65	70.8

* S = 3.5196

31. A strength training program is designed to improve core strength. To test its effectiveness, 12 patients are timed in seconds while holding a position called the plank before and after a 3 week strength program. Use the results below and a 1% significance level to test the claim that the program increases core strength.

Pre-Program	38	47	63	50	41	30	35	34	44	43	46	52
Post-program	67	92	120	75	69	60	68	65	131	122	120	135

* S = 24.4

32. An English teacher wants to test if her grammar instruction is effective. Ten students are pre and post tested by counting the number of errors missed by the students reading an essay. Use the results below and a 2% significance level to test the claim that the program produces some change in the student's ability to spot errors in written work. VS

Pre-test	12	14	5	21	17	18	4	7	13	10	
Post-test	8	11	0	12	7	10	1	2	4	7	

* S = 2.807

33. A fitness researcher has decided to test the weight loss effects of a sprinting program versus a traditional jogging plan. Each participant engaged in the running programs for 3 months, but some did the jogging first and others did the sprinting first. Use the results below and a 5% significance level to test the claim that the sprinting program produces more weight loss than the jogging. VS

Sprinting weight loss	10	6	5	12	15	12	4	8	9	
Jogging weight loss	8	4	0	12	8	10	1	2	11	

* S = 2.8626

: indicates the exercise has a video devoted to it in the corresponding section of STATSprofessor.com

34. A lot of test prep programs claim that they will improve student scores, but a retake may improve test scores without the expensive test prep. Eight students took the LSAT twice to see if there was an improvement on the second attempt. Use the results below and a 1% significance level to test the claim that there is a difference between the two attempts. What do these results say about the test prep industry? VS

1st score	161	143	142	152	145	147	143	155
2nd score	165	148	150	154	145	152	150	159

* S= 2.560

9.7 Answers

30. $Claim: \mu_M < \mu_R \rightarrow \mu_{M-R} = \mu_d < 0$, $\begin{matrix} H_o: \mu_d \geq 0 \\ H_a: \mu_d < 0 \end{matrix}$, $TestStat: 0.98$, $CriticalValue(s): -1.796$,

Initial Conclusion: Do not reject the Null, do not support the alternative

Final Conclusion: The sample data does not support the claim…

The fact that the participants were so young would have a major effect on the study. It is possible that they are too young to feel the need to lie about their height; especially since they are probably too young to be worried about competing for mates.

31. $Claim: \mu_{post} > \mu_{pre} \rightarrow \mu_{post-pre} = \mu_d > 0$, $\begin{matrix} H_o: \mu_d \leq 0 \\ H_a: \mu_d > 0 \end{matrix}$, $TestStat: 7.11$,

$CriticalValue(s): 2.718$, (Note: if you write post before pre in your claim you will perform the subtraction post - pre. If you do the subtraction consistently with how you expressed your claim, the overall outcome of the test will always be correct. Do not worry if you should do post - pre or pre - post, it won't matter if you are consistent throughout.)

Initial Conclusion: Reject the Null, support the alternative

Final Conclusion: The sample data supports the claim…

32. $Claim: \mu_{pre} \neq \mu_{post} \rightarrow \mu_{pre-post} = \mu_d \neq 0$, $\begin{matrix} H_o: \mu_d = 0 \\ H_a: \mu_d \neq 0 \end{matrix}$, $TestStat: 6.65$,

$CriticalValues: \pm 2.821$,

Initial Conclusion: Reject the Null, support the alternative

Final Conclusion: The sample data supports the claim…

33. $Claim: \mu_S > \mu_J \rightarrow \mu_{S-J} = \mu_d > 0$, $\begin{matrix} H_o: \mu_d \le 0 \\ H_a: \mu_d > 0 \end{matrix}$, $TestStat: 2.91$, $CriticalValue(s): 1.860$,

Initial Conclusion: Reject the Null, support the alternative

Final Conclusion: The sample data supports the claim…

34. $Claim: \mu_{1st} \ne \mu_{2nd} \rightarrow \mu_{1st-2nd} = \mu_d \ne 0$, $\begin{matrix} H_o: \mu_d = 0 \\ H_a: \mu_d \ne 0 \end{matrix}$, $TestStat: -4.834$,

$CriticalValues: \pm 3.499$,

Initial Conclusion: Reject the Null, support the alternative

Final Conclusion: The sample data supports the claim…

There seems to be improvement just by retaking the exam, so maybe those prep courses are a waste of money.

Need more exercises?

9.8 Confidence Interval to Compare Two Population Means: Matched-Pair Experiments

To complete this section of homework watch Chapter Nine, Lecture Example 141 and 142.

35. A researcher suspects that men, if given the chance, would lie about their heights. She suspects that they would try to say they are taller than they really are. To test this, the researcher first asks the male subjects to report their heights, and then she actually measures them. The subjects do not know they will be measured. Use the data below to construct a 95% confidence interval for the true μ_d between measured heights and reported heights. Is zero in the interval? If so, what does this tell us? VS

Reported height	68	71	63	70	71	60	65	64	54	63	66	72
Measured height	67.9	69.9	64.9	68.3	70.3	60.6	64.5	67	55.6	74.2	65	70.8

* S = 3.5196

36. A lot of test prep programs claim that they will improve student scores, but a retake may improve test scores without the expensive test prep. Eight students took the LSAT twice to see if there was an improvement on the second attempt. Use the results below and a 99% confidence level to create a confidence interval for the true mean difference between the scores.

1st score	161	143	142	152	145	147	143	155
2nd score	165	148	150	154	145	152	150	159

* S= 2.560

37. One of the skills that is necessary in order for you to be able to use the statistical techniques learned in this class is the ability to recognize when a particular test is appropriate. For example, what makes the scenario in the question above a matched pairs t-test as opposed to an independent t-test?

9.8 Answers

35. [-1.24, 3.24] We are 95% confident that the true mean difference lies within this interval. Since zero is inside the interval, it is possible that the true mean difference is zero. In other words, it's possible that there is no difference between reported and measured heights.

36. [1.21, 7.54] We are 99% confident the true mean difference between the second score and the first is between 1.21 to 7.54 points. This implies that there is an improvement when retaking the exam—no test prep required.

37. The two sets of exam scores are not unrelated or independent. They are connected by the test takers. Since the first exam in every column was taken by the same person as the second exam in the column, we have a matched pair scenario. The exam scores will be affected by the students' individual abilities. We are not interested if the students are different we know that they are, we are only concerned with the question of the before and after effect. We need to therefore block out the individual differences among the students, so that we can detect the probably smaller differences between the two exam attempts.

Need more exercises?

9.9 Inference Procedures to Compare Two Population Proportions: Independent Sampling

To complete this section of homework watch Chapter Nine, Lecture Example 143, 144, 145 and 146.

38. In a study of the opinions of 60 men and 58 women, fifty-two of the men said that they believe they work longer hours today than they did 15 years ago. Fifty-six of the women said they believe they work longer hours today than they did 15 years ago. Construct a 90% confidence interval for the difference of the proportions of men and women who think they work more hours than they did in the past. Can we say that the proportion of women who feel they work longer hours now is higher than the proportion of men who feel the same?

39. Alternate day fasting (ADF) is an alternative to the traditional carbohydrate reduction (CR) weight loss strategy. ADF requires patients to eat 75% less calories every other day for a prescribed period (usually several weeks). Research has been done to determine which method is more sustainable. One hundred patients were placed on the ADF plan for six months; at the end of the study 79 of the patients had remained on the plan. One hundred ten patients were given a CR plan to follow for six months, and 81 patients remained on the plan after six months. Using a 5% significance level, test the claim that the proportion of people that can remain on the ADF plan is the same as the proportion of those that can remain on the CR plan for six months. VS

40. Which sex exercises more? A recent poll of 1000 men revealed that 390 of them exercise at least three days per week. A similar poll of 800 women revealed that 256 of them exercised at least three days per week. Use a 2.5% significance level to test the claim that the proportion of men exercise three or more days per week is greater than the proportion of women. VS

41. In 2007, researchers looked at 15,024 deaths of US citizens overseas. In the study, it was found that 13% of those deaths were injury related. In the same year 121,599 deaths in the US mainland were due to injuries out of a total of 2,423,995 deaths. At the 2.5% significance level test the claim that the proportion of deaths of US citizens living in the mainland due to injury is less than the proportion of injury related deaths of US citizens abroad. VS

9.9 Answers

38. $-0.18 < p_{men} - p_{women} < -0.02$; Yes, because both limits are negative.

39. $Claim: p_{ADF} = p_{CR}$, $\begin{matrix} H_0: p_{ADF} = p_{CR} \\ H_a: p_{ADF} \neq p_{CR} \end{matrix}$, $TestStat: Z = 0.91$, $CriticalValue(s): \pm 1.960$,

$InitialConclusion:$ Do not reject the null, do not support the alternative.

$FinalConclusion:$ The sample data does not allow us to reject the claim…

40. $Claim: p_{men} > p_{women}$, $\begin{matrix} H_0: p_{men} \leq p_{women} \\ H_a: p_{men} > p_{women} \end{matrix}$, $TestStat: Z = 3.08$, $CriticalValue(s): 1.960$,

$InitialConclusion:$ Reject the null, support the alternative.

$FinalConclusion:$ The sample data supports the claim…

41. $Claim: p_{overseas} > p_{usa}$, $\begin{matrix} H_0: p_{overseas} \leq p_{usa} \\ H_a: p_{overseas} > p_{usa} \end{matrix}$, $TestStat: Z = 44.48$, $CriticalValue(s): 1.960$,

$InitialConclusion:$ Reject the null, support the alternative.

$FinalConclusion:$ The sample data supports the claim…

Need more exercises?

9.10 Using the f-table to Find Critical Values

To complete this section of homework watch Chapter Nine, Lecture Example 147 and 148.

Use the f-table to find the critical value for each of the given scenarios below:

42. Sample 1: $s_1^2 = 128, n_1 = 23$ Sample2: $s_2^2 = 162, n_2 = 16$ Claim: $\sigma_1^2 \neq \sigma_2^2$
    Significance Level: $\alpha = 0.02$

43. Sample 1: $s_1^2 = 37, n_1 = 14$ Sample2: $s_2^2 = 89, n_2 = 25$ Claim: $\frac{\sigma_2^2}{\sigma_1^2} > 1$
    Significance Level: $\alpha = 0.01$ VS

44. Sample 1: $s_1^2 = 232, n_1 = 30$ Sample2: $s_2^2 = 387, n_2 = 46$ Claim: $\frac{\sigma_2^2}{\sigma_1^2} \neq 1$
    Significance Level: $\alpha = 0.05$ VS

45. Sample 1: $s_1^2 = 164, n_1 = 21$ Sample2: $s_2^2 = 53, n_2 = 17$ Claim: $\sigma_1^2 \neq \sigma_2^2$
    Significance Level: $\alpha = 0.10$

46. Sample 1: $s_1^2 = 92.8, n_1 = 11$ Sample2: $s_2^2 = 43.6, n_2 = 11$ Claim: $\frac{\sigma_1^2}{\sigma_2^2} > 1$
    Significance Level: $\alpha = 0.05$

9.10 Answers

42. Critical Value: 2.98
43. Critical Value: 3.59
44. Critical Value: 2.03
45. Critical Value: 2.28
46. Critical Value: 2.98

Need more exercises?

9.11 Hypothesis Test to Compare Two Population Variances: Independent Sampling

To complete this section of homework watch Chapter Nine, Lecture Example 149 and 150.

: indicates the exercise has a video devoted to it in the corresponding section of STATSprofessor.com

47. Is there a difference between the variances of the number of weeks on the best seller lists for nonfiction and fiction books? Fifteen *New York Times* bestselling fiction books had a standard deviation of 6.17 weeks on the list. Sixteen *New York Times* bestselling nonfiction books had a standard deviation of 13.12 weeks. At the 10% significance level, can we conclude there is a difference in the variances? VS

48. Variation and Quality: A sign of quality is low variation in the dimensions of key manufactured components for a product. For example, a study looked at 25 shelves that were designed for a bookcase sold at the retail chain Wal-mart. This bookcase needed to be put together by the customer. They found that the variance for the shelf lengths was 9.23 mm^2 (the shelves were supposed to be 670 mm long in order to fit into the case properly). Another 27 shelves produced by Bassett Furniture were examined, and it was found that the variance for those shelf lengths was 0.25 mm^2. At the 5% significance level, test the claim that the shelves sold at Wal-Mart are manufactured with more variation than the shelves produced by Bassett. VS

49. Arrival time variation: Variation in the on-time arrival percentage was analyzed for 21 time blocks accounting for every hour of the day at Miami International Airport and Fort Lauderdale International Airport. The standard deviation for MIA was 14.2 (with a mean on-time arrival rate of 66.7%). The standard deviation for FLL was 12.5 (with a mean on-time arrival rate of 78.0%). At the 5% significance level, test the claim that the two airports have the same variation in on-time rates. VS

9.11 Answers

47. $Claim: \sigma^2_{fic} \neq \sigma^2_{nonfict}$, $\begin{matrix} H_0: \sigma^2_{fic} = \sigma^2_{nonfict} \\ H_a: \sigma^2_{fic} \neq \sigma^2_{nonfict} \end{matrix}$, $TestStat: F = 4.52$, $CriticalValue(s): 2.46$,

$InitialConclusion$: Reject the null, support the alternative.

$FinalConclusion$: The sample data supports the claim…

48. $Claim: \sigma^2_W > \sigma^2_B$, $\begin{matrix} H_0: \sigma^2_W \leq \sigma^2_B \\ H_a: \sigma^2_W > \sigma^2_B \end{matrix}$, $TestStat: F = 36.92$, $CriticalValue(s): 1.95$,

$InitialConclusion$: Reject the null, support the alternative.

$FinalConclusion$: The sample data supports the claim…

49. $Claim: \sigma^2_{MIA} = \sigma^2_{FLL}$, $\begin{matrix} H_0: \sigma^2_{MIA} = \sigma^2_{FLL} \\ H_a: \sigma^2_{MIA} \neq \sigma^2_{FLL} \end{matrix}$, $TestStat: F = 1.290$, $CriticalValue(s): 2.46$,

$InitialConclusion$: Do not reject the null, do not support the alternative.

$FinalConclusion$: The sample data does not allow rejection of the claim…

Need more exercises? Take a sample exam for chapter 9

: indicates the exercise has a video devoted to it in the corresponding section of STATSprofessor.com

Chapter 9 Mixed Review

50. In a random sample of 1500 people in their freshman year of college, 68% believe they will have children by the time they are 30 years old. In a random sample of 987 people in their senior year of college, 54% believe they will have children by the time they turn 30. The 95% confidence interval for the difference between the population proportions is:

$$0.101 < \rho_{freshman} - \rho_{senior} < 0.179$$

What does this imply?

51. A researcher believes the variation in resting heart rates for those people who exercise is less than the variation for those who do not exercise. Use the sample data below and a 5% significance level to test the claim that people who exercise have less variation in their resting heart rates.

Do Exercise	Don't Exercise
n = 28	n = 25
$\bar{x}$ = 69.9	$\bar{x}$ = 73.5
s = 8.2	s = 10.9

52. A doctor wants to test the claim, at the 2.5% significance level, that open surgery patients spend more time in the hospital recovering than patients who had less invasive procedures. Independent samples were collected. The data appears below (do not assume equal variances):

Open	Less Invasive
n = 25	n = 28
$\bar{x}$ = 48.6	$\bar{x}$ = 22.1
s = 11.2	s = 7.4

Claim: $\left(\mu_{lessInvasive} < \mu_{open}\right)$

Use the Welch-Satterthwaite method to find the degrees of freedom for the test and determine the critical value(s).

53. What kind of statistical procedure should be used to analyze data from the following study: An engineer measures the energy generated from two different types of solar cells over a random selection of sunny days in Phoenix. The engineer would like to run a hypothesis test to determine which type of solar cell produces more energy on average. He plans to test 22 solar cells made by ACME and 24 solar cells made by MPE. Past sample data indicates that the variances for the two cell types are about the same.

54. Who spends more time studying? A recent survey of the study habits of STA 3123 students at Florida International University produced the following interval: $-40.1 < \mu_{psy} - \mu_{bio} < -25.9$. The 95% confidence interval compares the average time spent studying per day in minutes for psychology majors (μ_{psy}) and biology majors (μ_{bio}). Interpret the interval.

55. The following data was used to produce the 95% confidence interval below to compare the endurance of athletes while taking placebo and while taking creatine ($\mu_{placebo} - \mu_{creatine}$). Participants sprinted on a road bike until they could no longer maintain a speed of 20 miles per hour. Each of the nine athletes were tested twice, once while taking creatine and once while taking a placebo. The amount of time (in minutes) that each athlete maintained the pace is listed below. The resulting interval was: $-1.40 < \mu_d < -0.27$. Interpret the results of the research.

Subject	1	2	3	4	5	6	7	8	9
Placebo	3.5	4.1	2.0	3.9	5.6	2.2	3.6	8.1	2.5
Creatine	4.6	4.2	3.2	4.1	6.6	2.1	4.7	10.4	3.1

56. A study looked at the relationship between height and human growth hormone levels at the end of puberty. A random selection of 22 low growth hormone, post-pubescent males had an average height of 66.5 inches with a standard deviation of 2.6 inches. A random selection of 19 normal growth hormone, post-pubescent males had an average height of 68.9 inches with a standard deviation of 2.7 inches. Use a 0.10 significance level to test the claim that the two groups of males have the same average height (assume equal variances).

57. A strength training program is designed to improve barbell squat weights for women. To test its effectiveness, 12 women are tested for their two-rep maximum squat before and after a four-week strength program. Use the results below and a 5% significance level to test the claim that the program increases squat maximums in women.

Pre-Program	125	145	175	110	95	225	200	100	110	125	130	155
Post-program	140	160	185	130	125	235	215	140	135	135	145	165
Differences	-15	-15	-10	-20	-30	-10	-15	-40	-25	-10	-15	-10

$\bar{x}_1 = 141.25, \bar{x}_2 = 159.17, \bar{x}_d = -17.92, s_1 = 40.74, s_2 = 35.28, s_p = 38.11, s_d = 9.40$

58. A bank manager wants to reduce the average wait time at her branch. She runs a test of two different line types: a single line and multiple lines. Using the computer output below and a 5% level of significance, test the claim that the two line types produce different average wait times.

	Single Line	Multiple Lines
Mean	211.9	235.6
Standard Dev.	57.3	76.5
Hypothesized difference	0	
Test Statistic	-1.19	
P-value (one-tailed)	0.118	
P-value (two-tailed)	0.235	

Chapter 9 Mixed Review Answers:

50. There is a significant difference between the proportion of freshmen and seniors who believe they will have children by the age of thirty. A larger proportion of freshmen believe they will have children by thirty. We are 95% confident the difference between the two proportions is between 10.1% and 17.9%.

51. $Claim: \sigma^2_{Don't} > \sigma^2_{Do}$, $\begin{matrix} H_0: \sigma^2_{Don't} \le \sigma^2_{Do} \\ H_a: \sigma^2_{Don't} > \sigma^2_{Do} \end{matrix}$, $TestStat: F = 1.767$, $CriticalValue(s): 1.9299$,

$InitialConclusion:$ Do not reject the null, do not support the alternative.

$FinalConclusion:$ The sample data do not support the claim…

52. Degrees of freedom: 40.84... which is truncated to **40**. The critical value is: **-2.021**.

53. This data should be analyzed using an independent t-test with the assumption of equal variances (i.e., pool the sample variances).

54. Biology majors spend more time studying on average. We are 95% confident that the average difference in study time is somewhere 25.9 and 40.1 minutes per day.

55. The placebo group had (statistically) significantly less endurance. We are 95% confident that the riders on creatine will maintain a speed of 20 mph between 0.27 and 1.4 minutes longer than riders on placebo.

56. $Claim: \mu_L = \mu_N$, $\begin{matrix} H_o: \mu_L = \mu_N \\ H_a: \mu_L \ne \mu_N \end{matrix}$, $TestStat: -2.895$, $CriticalValues: \pm 1.686$, $df = 39$, $S^2_p =$ 7.00462, Initial Conclusion: Reject the Null, support the alternative; Final Conclusion: The sample data allow us to reject the claim…

57. $Claim: \mu_{pre} < \mu_{post} \rightarrow \mu_{pre-post} = \mu_d < 0$, $\begin{matrix} H_o: \mu_d \ge 0 \\ H_a: \mu_d < 0 \end{matrix}$, $TestStat: -6.599$,

$CriticalValue: -1.796$ Initial Conclusion: Reject the null, support the alternative; Final Conclusion: The data support the claim that the program increases squat strength.

58. Since the test is two-tailed and the p-value for the two-tailed test is 0.235, we do not support the claim of different average wait times. The lines do not have significantly different average wait times.