Note that the 95% lower bound is just barely consistent with the alternate hypothesis. This indicates that the P-value is just barely less than 0.05 (although it is given by 0.050 to two significant digits).

Summary

Let $(X_1, Y_1), \ldots, (X_n, Y_n)$ be a sample of ordered pairs whose differences $D_1, \ldots, D_n$ are a sample from a *normal* population with mean μ_D. Let s_D be the sample standard deviation of $D_1, \ldots, D_n$.

To test a null hypothesis of the form $H_0: \mu_D \le \mu_0$, $H_0: \mu_D \ge \mu_0$, or $H_0: \mu_D = \mu_0$:

- Compute the test statistic $t = \dfrac{\overline{D} - \mu_0}{s_D/\sqrt{n}}$.
- Compute the P-value. The P-value is an area under the Student's t curve with $n - 1$ degrees of freedom, which depends on the alternate hypothesis as follows:

Alternate Hypothesis	P-value
$H_1: \mu_D > \mu_0$	Area to the right of t
$H_1: \mu_D < \mu_0$	Area to the left of t
$H_1: \mu_D \ne \mu_0$	Sum of the areas in the tails cut off by t and $-t$

- If the sample is large, the D_i need not be normally distributed, the test statistic is $z = \dfrac{\overline{D} - \mu_0}{s_D/\sqrt{n}}$, and a z test should be performed.

Exercises for Section 6.8

1. The article "Improved Bioequivalence Assessment of Topical Dermatological Drug Products Using Dermatopharmacokinetics" (B. N'Dri-Stempfer, W. Navidi, R. Guy, and A. Bunge, *Pharmaceutical Research*, 2009:316–328) described a study comparing the amounts of econozole nitrate absorbed into human skin for several formulations of antifungal ointment. Both a brand name and generic drug were applied to the arms of 14 subjects, and the amounts absorbed, in μg/cm^2, were measured. Following are the results. Can you conclude that the mean amount absorbed differs between the brand name and the generic drug?

Brand Name	Generic	Difference
2.23	1.42	0.81
1.68	1.95	−0.27
1.96	2.58	−0.62
2.81	2.25	0.56
1.14	1.21	−0.07
3.20	3.01	0.19
2.33	2.76	−0.43
4.06	3.65	0.41
2.92	2.89	0.03
2.92	2.85	0.07
2.83	2.44	0.39
3.45	3.11	0.34
2.72	2.64	0.08
3.74	2.82	0.92

2. The article "Estimation of Mean Arterial Pressure from the Oscillometric Cuff Pressure: Comparison of Different Techniques" (D. Zheng, J. Amoore, et al., *Med Biol Eng Comput*, 2011:33–39) describes a study comparing two methods of measuring mean arterial blood pressure. The auscultatory method is based on listening to sounds in a stethoscope, while the oscillatory method is based on oscillations in blood flow. Following are measurements on six subjects in mmHg, consistent with means and standard deviations presented in the article.

Auscultatory	Oscillatory	Difference
92.9	86.3	6.6
101.5	97.3	4.2
74.3	79.8	−5.5
95.0	98.1	−3.1
91.4	82.1	9.3
80.6	84.5	−3.9

Can you conclude that the mean reading is greater for the auscultatory method?

3. A dry etch process is used to etch silicon dioxide (SiO_2) off of silicon wafers. An engineer wishes to study the uniformity of the etching across the surface of the wafer. A total of 10 wafers are sampled after etching, and the etch rates (in Å/min) are measured at two different sites, one near the center of the wafer, and one near the edge. The results are presented in the following table.

Wafer	Center	Edge
1	586	582
2	568	569
3	587	587
4	550	543
5	543	540
6	552	548
7	562	563
8	577	572
9	558	559
10	571	566

Can you conclude that the etch rates differ between the center and the edge?

4. In an experiment to determine the effect of ambient temperature on the emissions of oxides of nitrogen (NO_x) of diesel trucks, 10 trucks were run at temperatures of 40°F and 80°F. The emissions, in ppm, are presented in the following table.

Truck	40°F	80°F
1	0.8347	0.8152
2	0.7532	0.7652
3	0.8557	0.8426
4	0.9012	0.7971
5	0.7854	0.7643
6	0.8629	0.8195
7	0.8827	0.7836
8	0.7403	0.6945
9	0.7480	0.7729
10	0.8486	0.7947

Can you conclude that the mean emissions differ between the two temperatures?

5. Two formulations of a certain coating, designed to inhibit corrosion, are being tested. For each of eight pipes, half the pipe is coated with formulation A and the other half is coated with formulation B. Each pipe is exposed to a salt environment for 500 hours. Afterward, the corrosion loss (in μm) is measured for each formulation on each pipe.

Pipe	A	B
1	197	204
2	161	182
3	144	140
4	162	178
5	185	183
6	154	163
7	136	156
8	130	143

Can you conclude that the mean amount of corrosion differs between the two formulations?

6. Two microprocessors are compared on a sample of six benchmark codes to determine whether there is a difference in speed. The times (in seconds) used by each processor on each code are given in the following table.

	Code					
	1	2	3	4	5	6
Processor A	27.2	18.1	27.2	19.7	24.5	22.1
Processor B	24.1	19.3	26.8	20.1	27.6	29.8

Can you conclude that the mean speeds of the two processors differ?

7. The compressive strength, in kilopascals, was measured for concrete blocks from five different batches of concrete, both three and six days after pouring. The data are presented in the following table.

	Batch				
	1	**2**	**3**	**4**	**5**
After 3 days	1341	1316	1352	1355	1327
After 6 days	1376	1373	1366	1384	1358

Can you conclude that the mean strength after six days is greater than the mean strength after three days?

8. The article "Effect of Granular Subbase Thickness on Airfield Pavement Structural Response" (K. Gopalakrishnan and M. Thompson, *Journal of Materials in Civil Engineering*, 2008:331–342) presents a study of the effect of the subbase thickness (in mm) on the amount of surface deflection caused by aircraft landing on an airport runway. Two landing gears, one simulating a Boeing 747 aircraft, and the other a Boeing 777 aircraft, were trafficked across four test sections of runway. The results are presented in the following table.

	Section			
	1	**2**	**3**	**4**
Boeing 747	4.01	3.87	3.72	3.76
Boeing 777	4.57	4.48	4.36	4.43

Can you conclude that the mean deflection is greater for the Boeing 777?

9. A crossover trial is a type of experiment used to compare two drugs. Subjects take one drug for a period of time, then switch to the other. The responses of the subjects are then compared using matched pair methods. In an experiment to compare two pain relievers, seven subjects took one pain reliever for two weeks, then switched to the other. They rated their pain level from 1 to 10, with larger numbers representing higher levels of pain. The results were

	Subject						
	1	**2**	**3**	**4**	**5**	**6**	**7**
Drug A	6	3	4	5	7	1	4
Drug B	5	1	5	5	5	2	2

Can you conclude that the mean response differs between the two drugs?

10. A group of eight individuals with high cholesterol levels were given a new drug that was designed to lower cholesterol levels. Cholesterol levels, in mg/dL, were measured before and after treatment for each individual, with the following results:

Subject	Before	After
1	283	215
2	299	206
3	274	187
4	284	212
5	248	178
6	275	212
7	293	192
8	277	196

a. Can you conclude that the mean cholesterol level after treatment is less than the mean before treatment?

b. Can you conclude that the reduction in mean cholesterol level after treatment is greater than 75 mg/dL?

11. The management of a taxi cab company is trying to decide if they should switch from bias tires to radial tires to improve fuel economy. Each of 10 taxis was equipped with one of the two tire types and driven on a test course. Without changing drivers, tires were then switched to the other tire type and the test course was repeated. The fuel economy (in mpg) for the 10

cars is as follows:

Car	Radial	Bias
1	32.1	27.1
2	36.1	31.5
3	32.3	30.4
4	29.5	26.9
5	34.3	29.9
6	31.9	28.7
7	33.4	30.2
8	34.6	31.8
9	35.2	33.6
10	32.7	29.9

a. Because switching tires on the taxi fleet is expensive, management does not want to switch unless a hypothesis test provides strong evidence that the mileage will be improved. State the appropriate null and alternate hypotheses, and find the P-value.

b. A cost-benefit analysis shows that it will be profitable to switch to radial tires if the mean mileage improvement is greater than 2 mpg. State the appropriate null and alternate hypotheses, and find the P-value, for a hypothesis test that is designed to form the basis for the decision whether to switch.

12. The following MINITAB output presents the results of a hypothesis test for the difference $\mu_X - \mu_Y$ between two population means.

```
Paired T for X - Y
               N       Mean      StDev    SE Mean
X             12    134.233     68.376     19.739
Y             12    100.601     94.583     27.304
Difference    12    33.6316    59.5113    17.1794

95% lower bound for mean difference: 2.7793
T-Test of mean difference = 0 (vs > 0):   T-Value = 1.96 P-Value = 0.038
```

a. Is this a one-tailed or two-tailed test?
b. What is the null hypothesis?
c. Can H_0 be rejected at the 1% level? How can you tell?
d. Use the output and an appropriate table to compute a 98% confidence interval for $\mu_X - \mu_Y$.

13. The following MINITAB output presents the results of a hypothesis test for the difference $\mu_X - \mu_Y$ between two population means. Some of the numbers are missing. Fill in the numbers for (a) through (d).

```
Paired T for X - Y
               N       Mean      StDev    SE Mean
X              7    12.4141     2.9235      (a)
Y              7     8.3476       (b)      1.0764
Difference     7       (c)     3.16758    1.19723

95% lower bound for mean difference: 1.74006
T-Test of mean difference = 0 (vs > 0):   T-Value = (d) P-Value = 0.007
```