```
The SAS System
The TTEST Procedure

                              Statistics           Lower CL       Upper CL
                       Lower CL     Upper CL         Std      Std      Std      Std
Variable  group        N    Mean   Mean    Mean      Dev      Dev      Dev      Err
-----------------------------------------------------------------------------------
pressure  C           10  110.49  126.1  141.71  15.008    21.82  39.834      6.9
pressure  SCI         10  110.08  133.1  156.12  22.133   32.178  58.745   10.176
pressure  Diff (1-2)      −32.83     −7   18.83  20.773   27.491  40.655   12.294

                                 T-Tests
-----------------------------------------------------------------------------------
Variable     Method           Variances     DF       t Value     Pr > |t|

pressure     Pooled           Equal         18       −0.57       0.5761
pressure     Satterthwaite    Unequal     15.8       −0.57       0.5771

                           Equality of Variances
-----------------------------------------------------------------------------------
Variable     Method           Num DF       Den DF   F Value     Pr > F
pressure     Folded F              9            9      2.17     0.2626
```

FIGURE 7.3.3 SAS® output for Example 7.3.2 (data in Table 7.3.1).

EXERCISES

In each of the following exercises, complete the ten-step hypothesis testing procedure. State the assumptions that are necessary for your procedure to be valid. For each exercise, as appropriate, explain why you chose a one-sided test or a two-sided test. Discuss how you think researchers or clinicians might use the results of your hypothesis test. What clinical or research decisions or actions do you think would be appropriate in light of the results of your test?

7.3.1 Subjects in a study by Dabonneville et al. (A-9) included a sample of 40 men who claimed to engage in a variety of sports activities (multisport). The mean body mass index (BMI) for these men was 22.41 with a standard deviation of 1.27. A sample of 24 male rugby players had a mean BMI of 27.75 with a standard deviation of 2.64. Is there sufficient evidence for one to claim that, in general, rugby players have a higher BMI than the multisport men? Let $\alpha = .01$.

7.3.2 The purpose of a study by Ingle and Eastell (A-10) was to examine the bone mineral density (BMD) and ultrasound properties of women with ankle fractures. The investigators recruited 31 postmenopausal women with ankle fractures and 31 healthy postmenopausal women to serve as controls. One of the baseline measurements was the stiffness index of the lunar Achilles. The mean stiffness index for the ankle fracture group was 76.9 with a standard deviation of 12.6. In the control group, the mean was 90.9 with a standard deviation of 12.5. Do these data provide sufficient evidence to allow you to conclude that, in general, the mean stiffness index is higher in healthy postmenopausal women than in postmenopausal women with ankle fractures? Let $\alpha = .05$.

7.3.3 Hoekema et al. (A-11) studied the craniofacial morphology of 26 male patients with obstructive sleep apnea syndrome (OSAS) and 37 healthy male subjects (non–OSAS). One of the variables of interest was the length from the most superoanterior point of the body of the hyoid bone to the Frankfort horizontal (measured in millimeters).

Length (mm) Non–OSAS				Length (mm) OSAS		
96.80	97.00	101.00	88.95	105.95	114.90	113.70
100.70	97.70	88.25	101.05	114.90	114.35	116.30
94.55	97.00	92.60	92.60	110.35	112.25	108.75
99.65	94.55	98.25	97.00	123.10	106.15	113.30
109.15	106.45	90.85	91.95	119.30	102.60	106.00
102.75	94.55	95.25	88.95	110.00	102.40	101.75
97.70	94.05	88.80	95.75	98.95	105.05	
92.10	89.45	101.40		114.20	112.65	
91.90	89.85	90.55		108.95	128.95	
89.50	98.20	109.80		105.05	117.70	

Source: Data provided courtesy of A. Hoekema, D.D.S.

Do these data provide sufficient evidence to allow us to conclude that the two sampled populations differ with respect to length from the hyoid bone to the Frankfort horizontal? Let $\alpha = .01$.

7.3.4 Can we conclude that patients with primary hypertension (PH), on the average, have higher total cholesterol levels than normotensive (NT) patients? This was one of the inquiries of interest for Rossi et al. (A-12). In the following table are total cholesterol measurements (mg/dl) for 133 PH patients and 41 NT patients. Can we conclude that PH patients have, on average, higher total cholesterol levels than NT patients? Let $\alpha = .05$.

Total Cholesterol (mg/dl)						
Primary Hypertensive Patients					Normotensive Patients	
207	221	212	220	190	286	189
172	223	260	214	245	226	196
191	181	210	215	171	187	142
221	217	265	206	261	204	179
203	208	206	247	182	203	212
241	202	198	221	162	206	163
208	218	210	199	182	196	196
199	216	211	196	225	168	189
185	168	274	239	203	229	142
235	168	223	199	195	184	168
214	214	175	244	178	186	121
134	203	203	214	240	281	
226	280	168	236	222	203	

(*Continued*)

Subject	Baseline FACT-G	Follow-up FACT-G	Subject	Baseline FACT-G	Follow-up FACT-G
23	19	7	56	27	10
24	23	20	57	22	16
25	19	19	58	16	14
26	21	24	59	26	24
27	24	23	60	17	19
28	21	15	61	23	22
29	28	27	62	23	23
30	18	26	63	13	3
31	25	26	64	24	22
32	25	26	65	17	21
33	28	28	66	22	21

Source: Data provided courtesy of Johnny Beney, Ph.D. and E. Beth Devine, Pharm.D., M.B.A. et al.

7.4.3 The purpose of an investigation by Morley et al. (A-17) was to evaluate the analgesic effectiveness of a daily dose of oral methadone in patients with chronic neuropathic pain syndromes. The researchers used a visual analogue scale (0–100 mm, higher number indicates higher pain) ratings for maximum pain intensity over the course of the day. Each subject took either 20 mg of methadone or a placebo each day for 5 days. Subjects did not know which treatment they were taking. The following table gives the mean maximum pain intensity scores for the 5 days on methadone and the 5 days on placebo. Do these data provide sufficient evidence, at the .05 level of significance, to indicate that in general the maximum pain intensity is lower on days when methadone is taken?

Subject	Methadone	Placebo
1	29.8	57.2
2	73.0	69.8
3	98.6	98.2
4	58.8	62.4
5	60.6	67.2
6	57.2	70.6
7	57.2	67.8
8	89.2	95.6
9	97.0	98.4
10	49.8	63.2
11	37.0	63.6

Source: John S. Morley, John Bridson, Tim P. Nash, John B. Miles, Sarah White, and Matthew K. Makin, "Low-Dose Methadone Has an Analgesic Effect in Neuropathic Pain: A Double-Blind Randomized Controlled Crossover Trial," *Palliative Medicine, 17* (2003), 576–587.

7.4.4 Woo and McKenna (A-18) investigated the effect of broadband ultraviolet B (UVB) therapy and topical calcipotriol cream used together on areas of psoriasis. One of the outcome variables is the Psoriasis Area and Severity Index (PASI). The following table gives the PASI scores for 20 subjects measured at baseline and after eight treatments. Do these data provide sufficient evidence, at the .01 level of significance, to indicate that the combination therapy reduces PASI scores?

EXERCISES

For each of the following exercises, carry out the ten-step hypothesis testing procedure at the designated level of significance. For each exercise, as appropriate, explain why you chose a one-sided test or a two-sided test. Discuss how you think researchers or clinicians might use the results of your hypothesis test. What clinical or research decisions or actions do you think would be appropriate in light of the results of your test?

7.5.1 Jacquemyn et al. (A-21) conducted a survey among gynecologists-obstetricians in the Flanders region and obtained 295 responses. Of those responding, 90 indicated that they had performed at least one cesarean section on demand every year. Does this study provide sufficient evidence for us to conclude that less than 35 percent of the gynecologists-obstetricians in the Flanders region perform at least one cesarean section on demand each year? Let $\alpha = .05$.

7.5.2 In an article in the journal *Health and Place,* Hui and Bell (A-22) found that among 2428 boys ages 7 to 12 years, 461 were overweight or obese. On the basis of this study, can we conclude that more than 15 percent of the boys ages 7 to 12 in the sampled population are obese or overweight? Let $\alpha = .05$.

7.5.3 Becker et al. (A-23) conducted a study using a sample of 50 ethnic Fijian women. The women completed a self-report questionnaire on dieting and attitudes toward body shape and change. The researchers found that five of the respondents reported at least weekly episodes of binge eating during the previous 6 months. Is this sufficient evidence to conclude that less than 20 percent of the population of Fijian women engage in at least weekly episodes of binge eating? Let $\alpha = .05$.

7.5.4 The following questionnaire was completed by a simple random sample of 250 gynecologists. The number checking each response is shown in the appropriate box.

1. When you have a choice, which procedure do you prefer for obtaining samples of endometrium?
   (a) Dilation and curettage $\boxed{175}$
   (b) Vobra aspiration $\boxed{75}$

2. Have you seen one or more pregnant women during the past year whom you knew to have elevated blood lead levels?
   (a) Yes $\boxed{25}$
   (b) No $\boxed{225}$

3. Do you routinely acquaint your pregnant patients who smoke with the suspected hazards of smoking to the fetus?
   (a) Yes $\boxed{238}$
   (b) No $\boxed{12}$

Can we conclude from these data that in the sampled population more than 60 percent prefer dilation and curettage for obtaining samples of endometrium? Let $\alpha = .01$.

7.5.5 Refer to Exercise 7.5.4. Can we conclude from these data that in the sampled population fewer than 15 percent have seen (during the past year) one or more pregnant women with elevated blood lead levels? Let $\alpha = .05$.

7.5.6 Refer to Exercise 7.5.4. Can we conclude from these data that more than 90 percent acquaint their pregnant patients who smoke with the suspected hazards of smoking to the fetus? Let $\alpha = .05$.

MINITAB Output

Test and CI for Two Proportions

```
Sample    X    N   Sample p
1        24   44  0.545455
2        11   29  0.379310


Difference = p (1) − p (2)
Estimate for difference: 0.166144
95% lower bound for difference: −0.0267550
Test for difference = 0 (vs > 0):  Z = 1.39  P-Value = 0.082
```

NCSS Output

Test Name	Test Statistic's Distribution	Test Statistic Value	Prob Level	Conclude H1 at 5% Significance?
Z-Test	Normal	1.390	0.0822	No

FIGURE 7.6.1 MINITAB and partial NCSS output for the data in Example 7.6.1.

EXERCISES

In each of the following exercises use the ten-step hypothesis testing procedure. For each exercise, as appropriate, explain why you chose a one-sided test or a two-sided test. Discuss how you think researchers or clinicians might use the results of your hypothesis test. What clinical or research decisions or actions do you think would be appropriate in light of the results of your test?

7.6.1 Ho et al. (A-25) used telephone interviews of randomly selected respondents in Hong Kong to obtain information regarding individuals' perceptions of health and smoking history. Among 1222 current male smokers, 72 reported that they had "poor" or "very poor" health, while 30 among 282 former male smokers reported that they had "poor" or "very poor" health. Is this sufficient evidence to allow one to conclude that among Hong Kong men there is a difference between current and former smokers with respect to the proportion who perceive themselves as having "poor" and "very poor" health? Let $\alpha = .01$.

7.6.2 Landolt et al. (A-26) examined rates of posttraumatic stress disorder (PTSD) in mothers and fathers. Parents were interviewed 5 to 6 weeks after an accident or a new diagnosis of cancer or diabetes mellitus type I for their child. Twenty-eight of the 175 fathers interviewed and 43 of the 180 mothers interviewed met the criteria for current PTSD. Is there sufficient evidence for us to conclude that fathers are less likely to develop PTSD than mothers when a child is traumatized by an accident, cancer diagnosis, or diabetes diagnosis? Let $\alpha = .05$.

7.6.3 In a *Kidney International* article, Avram et al. (A-27) reported on a study involving 529 hemodialysis patients and 326 peritoneal dialysis patients. They found that at baseline 249 subjects in the hemodialysis treatment group were diabetic, while at baseline 134 of the subjects in the peritoneal dialysis group were diabetic. Is there a significant difference in diabetes prevalence at baseline between the two groups of this study? Let $\alpha = .05$. What does your finding regarding sample significance imply about the populations of subjects?

7.6.4 In a study of obesity the following results were obtained from samples of males and females between the ages of 20 and 75:

	n	Number Overweight
Males	150	21
Females	200	48

Can we conclude from these data that in the sampled populations there is a difference in the proportions who are overweight? Let $\alpha = .05$.

7.7 HYPOTHESIS TESTING: A SINGLE POPULATION VARIANCE

In Section 6.9 we examined how it is possible to construct a confidence interval for the variance of a normally distributed population. The general principles presented in that section may be employed to test a hypothesis about a population variance. When the data available for analysis consist of a simple random sample drawn from a normally distributed population, the test statistic for testing hypotheses about a population variance is

$$\chi^2 = (n-1)s^2/\sigma^2 \tag{7.7.1}$$

which, when H_0 is true, is distributed as χ^2 with $n-1$ degrees of freedom.

EXAMPLE 7.7.1

The purpose of a study by Wilkins et al. (A-28) was to measure the effectiveness of recombinant human growth hormone (rhGH) on children with total body surface area burns > 40 percent. In this study, 16 subjects received daily injections at home of rhGH. At baseline, the researchers wanted to know the current levels of insulin-like growth factor (IGF-I) prior to administration of rhGH. The sample variance of IGF-I levels (in ng/ml) was 670.81. We wish to know if we may conclude from these data that the population variance is not 600.