3. In a between-subjects design there are two independent groups of subjects. From each group is obtained the mean of that group and an estimate of the population variance. The one best estimate of the population variance is then obtained by computing a weighted average of the two estimates from the two groups. The weighting is in terms of degrees of freedom; and the final estimate is based on $n_1 - 1 + n_2 - 1 = n_1 + n_2 - 2$ degrees of freedom, where n_1 and n_2 are the sample sizes of the two groups. In the case where $n_1 = n_2$, the weighting is equal, and therefore the one best estimate of σ^2 is simply the mean of the two estimates from the two groups.

   Finally the hypothesis-testing steps involve testing the difference between the two sample means against zero using the formula

$$t(n_1 + n_2 - 2) = \frac{(M_1 - M_2) - 0}{\text{est } \sigma_{M_1 - M_2}}$$

4. The t-test as well as the z-test and the analysis of variance are based on certain assumptions. Violation of these assumptions is generally not a cause for alarm; but when it is, one may do nonparametric tests, which are the topic of a later chapter.

PROBLEMS

1. In the three questions below, the mean M of a random sample from an approximately normal population is given, and the value of the sample standard deviation S computed from that sample is also given. In each problem test the specified hypothesis about the population mean μ. Use (i) a one-tailed t-test at the 0.05 significance level; (ii) a two-tailed t-test at the 0.05 significance level; (iii) a one-tailed t-test at the 0.01 significance level; and (iv) a two-tailed t-test at the 0.01 significance level. Also use the given data to find (v) a 95% confidence interval for the mean of the population and (vi) a 99% confidence interval for the mean of the population.
   a. The hypothesis is that a sample of 20 clinical interview tapes have $M = 52$ minutes and $S = 19$ minutes comes from a population having $\mu = 50$ minutes.
   b. The hypothesis is that a sample of 12 weights of males having $M = 182$ pounds and $S = 16$ comes from a population having $\mu = 170$ pounds.
   c. The hypothesis is that a sample of 64 heights having $M = 5$ feet 9 inches and $S = 3$ inches comes from a population having $\mu = 5$ feet 6 inches.

2. Gail Burns reads in the newspaper that Americans see an average (mean) of 10 homicides a night on TV. Wondering whether her sorority sisters differ from this national average, she asks each of the four other members of her sorority how many homicides each of them saw on TV the previous night. Their answers are as follows:

lent groups of subjects.
and an estimate of the
ulation variance is then
estimates from the two
n; and the final estimate
ees of freedom, where
le case where $n_1 = n_2$,
stimate of σ^2 is simply

the difference between

ce are based on certain
not a cause for alarm,
are the topic of a later

dom sample from an
value of the sample
also given. In each
ulation mean μ. Use
a two-tailed t-test at
the 0.01 significance
ce level. Also use the
iean of the population
population.
interview tapes have
a population having

les having $M = 182$
g $\mu = 170$ pounds.
$M = 5$ feet 9 inches
$\mu = 5$ feet 6 inches.

an average (mean) of
sorority sisters differ
her members of her
the previous night.

Sister	Number of Homicides
1	9
2	8
3	10
4	9

a. What is the mean number of homicides and the 95% confidence interval?
b. Perform the relevant t-test on these data.

3. Consider Chapter 5, problems 3, 4, 5, and 8.
   a. Reevaluate the data from these problems performing a t-test rather than a sign test. (*Hint:* Each of these t-tests should be a within-something t-test.)
   b. Compute 95% confidence intervals around the sample mean difference score in each case.

4. In questions a–c of this problem the means M_1 and M_2 of random samples of sizes N_1 and N_2, respectively, and the corresponding sample standard deviations, S_1 and S_2, are given. In each problem test the null hypothesis that the two samples came from the same normal population. Use (i) a one-tailed t-test at the 0.05 significance level; (ii) a two-tailed t-test at the 0.05 significance level; (iii) a one-tailed t-test at the 0.01 significance level; and (iv) a two-tailed t-test at the 0.01 significance level. Also find the 95% confidence interval around each of the two means.
   a. For students' graduating averages: $M_1 = 78$, $n_1 = 16$, $S_1 = 6$; $M_2 = 84$, $n_2 = 16$, $S_2 = 6$.
   b. For female heights: $M_1 = 5$ ft 9 inches, $n_1 = 81$, $S_1 = 3$ feet; $M_2 = 5$ ft 4 inches, $n_2 = 11$, $S_2 = 4.5$ inches.
   c. For weights of U.S. males: $M_1 = 170$ pounds, $n_1 = 16$, $S_1 = 15$; $M_2 = 185$ pounds, $n_2 = 10$, $S_2 = 12$.

5. A developmental psychologist is interested in whether dating behavior is different for first-born versus second-born adolescent males. He gathers two samples of six first-born and six second-born teenagers and measures how many dates each has in a month. He gets the following data:

Number of Dates in a Month	
First Born	**Second Born**
6	2
4	5
5	4
7	2
3	1
5	4

   a. Plot the two means, including 95% confidence intervals around the means.
   b. Perform a t-test on the data.

6. Consider the following data (assume that 10 *S*'s each participate in both conditions):

Subject	Condition 1	Condition 2
1	9	7
2	2	2
3	7	4
4	12	13
5	14	13
6	10	7
7	6	4
8	7	6
9	12	8
10	10	9

   a. Get the set of difference scores (condition 1 minus condition 2 for each *S*).
   b. Compute the mean difference score and the 95% confidence interval around the mean difference score.
   c. Use a *t*-test to test the hypothesis that condition 1 is better than condition 2.
   d. Do the same thing with a sign test.

7. An experiment is done on the effect of marijuana on memory. Two groups of six subjects per group are given a free-recall test on a 15-word list. The experimental group has been given two marijuana cigarettes to smoke, whereas the control group has been given two standard cigarettes. The results (number of correct responses out of 15) are as follows:

Number of Words Remembered

Experimental (marijuana)	Control (cigarette)
6	10
5	5
8	8
7	8
5	9
7	8

   a. Plot the two sample means along with 95% confidence intervals around the means.
   b. Test the hypothesis that the marijuana causes a decrease in memory performance.

8. Photi's Fish Shop claims that it sells heavier salmon than does Ivan's across the street. Samples of $n_1 = 4$ salmon and $n_2 = 3$ salmon are obtained from Photi's and Ivan's, respectively. The weights (in pounds) are as follows:

Group 1 (Photi's)	Group 2 (Ivan's)
6	3
8	5
7	4
7	

a. Plot the means for the two groups along with the 95% confidence intervals.
b. Test the hypothesis that Photi's salmon weigh more, in general, than do Ivan's.

9. Plot means and 95% confidence intervals for the data from Chapter 8, problems 7, 8, 9, 10, 11, 12, and 13. Assume for each that σ^2, the population variance, is *unknown*.

10. An experiment is done to test whether vitamin C is effective in preventing colds using two groups of four subjects per group. Each subject in group 1 takes 1000 milligrams of vitamin C per day for a year, whereas each subject in group 2 takes a 1000-milligram sugar tablet every day. The number of colds during the year is recorded for each subject. Unfortunately, one of the group 2 subjects catches pneumonia and has to leave the experiment. The data are as follows:

Group 1 (vitamin C)	Group 2 (sugar tablet)
$x_{11} = 1$	$x_{12} = 3$
$x_{21} = 0$	$x_{22} = 3$
$x_{31} = 1$	$x_{32} = 3$
$x_{41} = 2$	

a. Plot the two sample means along with 95% confidence intervals.
b. Can you conclude from these data that vitamin C reduces the number of colds? Carry out the appropriate test.

11. A developmental psychologist believes that if a child is forced to eat a certain food, he or she will grow to like that food. An experiment is designed to test that notion. Two groups of 10 five-year-old children per group are picked. In the "liver group" each child is made by his or her parents to take two bites of liver a week for a year. In the control group liver is offered to the child, but he or she is not forced to eat it. Five years later, a liver-preference test is given to the two groups. Each child is asked to rate how much he or she likes liver on a scale ranging from 1 = very much to 5 = not at all. The families of two of the children who were in the liver group have moved away from town. Four of the families of children who were in the control group have moved away. The data are as follows:

Rating for How Much Liver Is Liked

Liver Group	Control Group
4	3
5	3
3	2
3	1
4	2
5	5
2	
4	

a. Plot the two sample means along with 95% confidence intervals.
b. Test the hypothesis that the liver group likes liver *more* than the control group.

12. An experiment is done to test the effect of communicator "peer similarity" on attitudes. Subjects (college students) are randomly assigned to two groups: a "similar peer" (SP) and "dissimilar peer" (DP) group, with eight students per group. In both groups a 19-year-old communicator gives a lecture on why the presidential incumbent should be reelected. In the SP condition the communicator is long-haired and dressed in sandals, jeans, and a t-shirt. In the DP condition the communicator has a crew-cut and is wearing glasses. At the end of the talk the students are asked to rate on a 1–7 scale how strongly they think the incumbent should be reelected. Unfortunately, in the DP condition two subjects ask to be excused in the middle of the lecture due to nausea; while in the SP condition one subject leaves to go to the bathroom and never returns. This leaves six subjects in the DP condition and seven subjects in the SP condition.

The results are as follows:

Rating that incumbent should be reelected (1 means definitely should be reelected; 7 means definitely should not be reelected):

SP Group	DP Group
4	7
3	5
3	7
2	6
5	6
4	7
6	

a. Plot the two means along with 95% confidence intervals.
b. Test the hypothesis that a similar peer produces *different* attitudes from a dissimilar peer.

13. Marine e
sink. In
tankers te
looks at r
one sunk

Am

Li

a. Test t
than v
b. Comp

14. Dr. Joe
reduce fe
experime
subjects
sugar pil
control g
mistaken
eliminate
as follow

Ex

a. Can J
Perfo
b. Do yo
in this
c. Suppo
these
their
d. What
aspirin

13. Marine engineer Joe Smith is interested in spills that occur when oil tankers sink. In particular, Joe is interested in the question of whether Liberian tankers tend to spill more oil when they sink than do American tankers. He looks at records from six sunken Liberian tankers, but he is only able to find one sunken American tanker. The data are as follows.

Amount of Oil Spilled (Millions of Gallons)

Liberian Tankers	American Tanker
$x_{11} = 125$	$x_{12} = 112$
$x_{21} = 120$	
$x_{31} = 131$	
$x_{41} = 132$	
$x_{51} = 119$	
$x_{61} = 126$	

a. Test the hypothesis that more oil is spilled when a Liberian tanker sinks than when an American tanker sinks. Use the 0.05 α-level.
b. Compute a 98% confidence interval around the mean in each condition.

14. Dr. Joe Smith is doing an experiment to determine whether aspirin will reduce fever. Twelve feverish hospital patients are randomly divided into six experimental group and six control group subjects. The experimental group subjects are given aspirin and the control group subjects are given aspirinlike sugar pills. The change in fever is then noted. Unfortunately, five of the control group subjects and three of the experimental group subjects are mistakenly given cyanide instead of sugar pills or aspirin and have to be eliminated from the experiment. The data from the remaining subjects are as follows. (Negative sign indicates decrease in temperature.)

$n_1 = 3$ Experimental Group	$n_2 = 1$ Control Group
$x_{11} = -2°$	$x_{12} = -1°$
$x_{21} = 0°$	
$x_{31} = -4°$	

a. Can Joe conclude that aspirin reduces fever more than do sugar pills? Perform the appropriate test.
b. Do you think the homogeneity of variance assumption is being violated in this experiment?
c. Suppose it is known that the population variance of fever reduction under these conditions is $\sigma^2 = 1$. Test whether aspirin and sugar pills differ in their effect on fever reduction.
d. What is the power of the test against the alternative hypothesis that aspirin reduces fever 1° more than sugar pills?

15. An experiment is done to test the effects of activity on time perception. The experiment involves 18 subjects and consists of two phases:

Phase 1: The subjects are randomly divided into two groups: nine subjects to a high-activity group (group 1) and nine subjects to a low-activity group (group 2). To get baseline measures, each subject in the two groups is placed alone in a room and asked to estimate when 10 minutes has elapsed. The dependent variable is thus the subject's estimate of 10 minutes. Note that no manipulation has been performed between the two groups yet.

The data may be represented as follows:

Group 1 (High Activity)	Group 2 (Low Activity)
x_{11}	x_{12}
x_{21}	x_{22}
$\vdots$	$\vdots$
x_{91}	x_{92}

Phase 2: The subjects in group 1 are given a high-activity task to perform while subjects in group 2 are given a low-activity task. The subjects then go through the time estimation procedure again. The data may be represented as follows (here we will use the notation x to distinguish the data from the baseline data):

Group 1 (High Activity)	Group 2 (Low Activity)
x'_{11}	x'_{12}
x'_{21}	x'_{22}
$\vdots$	$\vdots$
x'_{91}	x'_{92}

a. How would you test the hypothesis that the time estimates would be different during high versus low activity.
   What you would use as your scores.
   What means you would calculate.
   What estimates of the population variance you would calculate and how you would combine them into one best estimate.
   How many degrees of freedom your test would be based on.
b. In similar fashion describe how you would test the hypothesis that baseline time estimates are different in the two conditions.
c. Suppose the experimenter wishes to compute an estimate of the population variance of the baseline scores (that is, the phase 1 data only). The experimenter claims that with the phase 1 data only, he could compute such an estimate based on 17 degrees of freedom. Is this possible? Why or why not?

16. Gail Burns believes that chewing gum during a statistics examination will improve exam scores. Hence, she recruits four of her sorority sisters, Betty, Connie, Carrie, and Rosie, to take part in an experiment. In the first hour

exam they do not chew gum, and in the second hour exam they do. Their scores are as follows:

Sister	Exam 1	Exam 2
Betty	91	92
Connie	73	80
Carrie	41	41
Rosie	80	90

a. Test the hypothesis that chewing gum improves examination scores.
b. Compute a confidence interval that will include the population mean difference score with a probability of 0.95.
c. What is wrong with this experimental design?

17. A college dean is interested in whether student ratings from an introductory chemistry class (of 50 students) differ from the ratings of an introductory psychology class (of 40 students). Each student rates the quality of the class on the following scale:

1 = excellent
2 = good
3 = average
4 = poor
5 = abominable

Thus, one rating is obtained from each student in each of the two classes. The data are as follows:

Chemistry ($N_1 = 50$)	Psychology ($N_2 = 40$)
$M_1 = \sum_{i=1}^{50} \frac{X_{i1}}{50} = 2.40$	$M_2 = \sum_{i=1}^{40} \frac{X_{i2}}{40} = 1.80$
$SS_1 = \sum_{i=1}^{50} (X_{i1} - M_1)^2 = 100$	$SS_2 = \sum_{i=1}^{40} (X_{i2} - M_2)^2 = 76$

a. Calculate est σ^2.
b. Calculate est $\sigma^2_{M_1}$ and est $\sigma^2_{M_2}$.
c. Test the hypothesis that the ratings for introductory chemistry *differ* from those for introductory psychology.
d. Test the hypothesis that the ratings for introductory chemistry *differ* from an "average" rating of 3.