Chapter 15

MISCELLANEOUS PROBLEMS

SECTION 15.1 Introduction

The purpose of this chapter is to give the student practice in choosing the appropriate statistical analysis for a particular dataset. Problems are selected from the material in all the chapters and are arranged in random order. Although this is not a substitute for experience in "real" data analysis, it should increase confidence in applying the knowledge obtained from this text in new situations. If you are unsure as to which method to use for a particular problem, consult the Flow Chart of Methods of Statistical Inference (in the text).

PROBLEMS

Hypertension

Suppose we wish to design an experiment to assess the effectiveness of a new drug for treating hypertensive patients. We decide to declare the drug "effective" if the mean diastolic blood pressure (DBP) of the 50 patients participating in the study drops by at least 2 mm Hg after using the drug daily for 1 month.

Let us set up this study as a hypothesis-testing problem and assume that the population standard deviation of the before-after blood-pressure differences is known to be 10 mm Hg.

15.1 If we identify declaring the drug effective with rejecting the null hypothesis, then what is the significance level of the test?

15.2 Suppose that the underlying population before–after DBP decrease from using the drug is 3 mm Hg. What is the power of the test in Problem 15.1?

Gynecology

In a 1985 study, 89 of 283 women with primary tubal infertility (cases) and 640 of 3833 control women reported ever having used an IUD [1].

15.3 What is the best point estimate of the rate of IUD use among case and control women, respectively?

15.4 Provide a 95% CI for the estimates in Problem 15.3.

Emergency Medicine

Mannitol and Decadron are drugs that are often administered to patients with severe head injury when they are admitted to the emergency room of a hospital. One hypothesis is that this type of treatment would be more beneficial if administered by paramedics to patients in the field before they are transported to the hospital. To plan such a study, a pilot study is performed, whereby 4 of 10 patients with field treatment and 6 of 10 patients with no field treatment die before discharge from the hospital.

15.5 If a clinical trial is planned based on a two-tailed test with $\alpha = .05$, assuming that the pilot study mortality rates are the true mortality rates in the 2 groups, how much power would such a study have if 20 patients are randomized to each of the two groups?

15.6 How many patients are needed in each group to achieve an 80% power with the preceding study design if the differences found in the pilot study are assumed to be valid and an equal number of patients are randomized to each group?

15.7 If the true mortality rate in the field-treated group were actually .50 rather than .40, how would these new data affect the power estimate in Problem 15.5 and the sample-size estimate in Problem 15.6 (Provide only a qualitative answer.)

Cardiovascular Disease
Much controversy has arisen concerning the possible association between myocardial infarction (MI) and coffee drinking. Suppose the information in Table 15.1 on coffee drinking and prior MI status is obtained from 200 60–64-year-old males in the general population.

Table 15.1 Coffee drinking and prior MI status

Coffee drinking (cups/day)	MI in last 5 years	Number of people
0	Yes	3
0	No	57
1	Yes	7
1	No	43
2	Yes	8
2	No	42
3 or more	Yes	12
3 or more	No	28
	Total yes	30
	Total no	170

15.8 Test for the association between history of MI and coffee-drinking status, which is categorized as follows: 0 cups, 1 or more cups.

15.9 Suppose coffee drinking is categorized as follows: 0 cups, 1 cup, 2 cups, 3 or more cups. Perform a test to investigate whether or not there is a "dose-response" relationship between these two variables (i.e., does the prevalence of prior MI increase or decrease as the number of cups of coffee per day increases?).

Cardiology
A study was conducted among a group of people who underwent coronary angiography at Baptist Memorial Hospital, Memphis, Tennessee, between January 1, 1972, and December 31, 1986 [2]. A group of 1493 people with coronary-artery disease were identified and were compared with a group of 707 people without coronary-artery disease (the controls). Both groups were age 35–49 years. Risk-factor information was collected on each group. Among cases, the mean serum cholesterol was 234.8 mg/dL with standard deviation = 47.3 mg/dL. Among controls, the mean serum cholesterol was 215.5 mg/dL with standard deviation = 47.3 mg/dL.

15.10 What test is appropriate to determine if the true mean serum cholesterol is different between the two groups?

15.11 Implement the test in Problem 15.10 and report a two-sided *p*-value.

15.12 What power did the study have to detect a significant difference using a two-sided test with $\alpha = .05$ if the true mean difference is 10 mg/dL between the two groups and the true standard deviations are the same as the sample standard deviations in the study?

Obstetrics
A new drug therapy is proposed for the prevention of low-birthweight deliveries. A pilot study undertaken, using the drug on 20 pregnant women, found that the mean birthweight in this group is 3500 g with a standard deviation of 500 g.

15.13 What is the standard error of the mean in this case?

15.14 What is the difference in interpretation between the standard deviation and standard error in this case (in words)?

15.15 Suppose $(\bar{x} - \mu_0)/(s/\sqrt{n}) = 2.73$ and a one-sample *t* test is performed based on 20 subjects. What is the two-tailed *p*-value?

Health-Services Administration
A comparison is made between demographic characteristics of patients using fee-for-service practices and prepaid group health plans. Suppose the data presented in Table 15.2 are found.

Table 15.2 Characteristics of patients using fee-for-service practices and prepaid group health plans

	Fee-for-service			Prepaid group health plans		
Characteristic	Mean	sd	*n*	Mean	sd	*n*
Age (years)	58.1	6.2	57	52.6	4.3	48
Education (years)	11.8	0.7	57	12.7	0.8	48

15.16 Test for a significant difference in the variance of age between the two groups.

15.17 What is the appropriate test to compare the mean ages of the two groups?

15.18 Perform the test in Problem 15.17 and report a *p*-value.

15.19 Compute a 95% CI for the mean age difference between the two groups.

15.20 Answer Problem 15.16 for the number of years of education.

15.21 Answer Problem 15.17 for the number of years of education.

15.22 Answer Problem 15.18 for the number of years of education.

15.23 Answer Problem 15.19 for the number of years of education.

Hypertension
An investigator wishes to determine if sitting upright in a chair versus lying down on a bed will affect a person's blood pressure. The investigator decides to use each of 10 patients as his or her own control and collects systolic blood-pressure (SBP) data in both the sitting and lying positions, as given in Table 15.3.

Table 15.3 Effect of position on SBP level (mm Hg)

Patient	Sitting upright	Lying down
1	142	154
2	100	106
3	112	110
4	92	100
5	104	112
6	100	100
7	108	120
8	94	90
9	104	104
10	98	114

15.24 What is the distinction between a one-sided and a two-sided hypothesis test in this problem?

15.25 Which hypothesis test is appropriate here? Why?

15.26 Using an α level of .05, test the hypothesis that position affects the level of blood pressure.

15.27 What is the *p*-value of the preceding test? Compute either the specific *p*-value or a range within which the *p*-value lies.

Hypertension
A frequently encountered phenomenon in Western society is the positive correlation of blood pressure and age. One finding of recent interest is the apparent absence of this correlation in many underdeveloped countries. Suppose that in a sample of 903 U.S. males, the observed correlation is .402, whereas in a sample of 444 Polynesian males, the observed correlation is .053.

15.28 Test for whether or not a significant difference exists between the two underlying correlations.

Cardiovascular Disease
Many studies have demonstrated that the level of HDL (high-density lipoprotein) cholesterol is positively related to alcohol consumption. This relationship has intrigued researchers because the level of HDL cholesterol is inversely correlated with the incidence of heart disease.

One possible mechanism was explored by Kuller et al. in an analysis of participants in the MRFIT study relating the level of HDL cholesterol to the level of SGOT, a parameter commonly used to assess liver function [3]. The data in Table 15.4 were presented.

Table 15.4 The relationship of HDL cholesterol to SGOT in the MRFIT population

SGOT	Mean HDL cholesterol (mg/dL)
≤ 9	40.0
10–12	41.2
13–14	42.3
15–16	42.8
17–18	43.8
19–20	43.6
≥ 21	46.5

Source: Reprinted with permission of the *American Journal qf Epidemiology*, **117**(4),406–418,1983.

15.29 Fit a regression line predicting mean HDL cholesterol as a function of SGOT. For this purpose assume that all levels of SGOT ≤ 9 and ≥ 21 are 9.5 and 20.5, respectively, and that levels of SGOT within all other groups occur at the midpoint of the group. Assume also that group assignments are made after rounding; for example, 10–12 represents an actual range from 9.5 to 12.5. Using this convention, if $x =$ SGOT and $y =$ mean HDL cholesterol, then

$$\sum_{i=1}^{7} x_i = 107,\ \sum_{i=1}^{7} x_i^2 = 1740.5,\ \sum_{i=1}^{7} y_i = 300.2,$$
$$\sum_{i=1}^{7} y_i^2 = 12{,}900.2,\ \sum_{i=1}^{7} x_i y_i = 4637.6$$

15.30 Test the line fitted in Problem 15.29 for statistical significance and report a *p*-value.

15.31 What is the best estimate of the HDL cholesterol for an average person with SGOT level of 11, and what is the standard error of this estimate?

15.32 Suppose that a mistake was made in Problem 15.29 and the SGOT groups were assigned by truncation rather than rounding (e.g., 10–12 would represent 10.0–12.99 rather than 9.5–12.5). It would follow that average levels of SGOT within each group would increase by 0.5 (assume this is also the case for the groups ≤ 9 and ≥ 21, respectively). What effect does the truncation-assignment rule have on the estimates of the regression parameters in Problem 15.29? (No further calculation should be necessary.)

As part of the same study, a multiple-regression analysis was performed relating the change in HDL cholesterol to baseline levels and changes in other risk factors. The results are presented in Table 15.5.

15.33 Interpret the coefficients for change in SGOT and drinks/week in the multiple-regression model.

Table 15.5 Multiple regression of change in HDL cholesterol from the second screening visit to the 48-month follow-up visit (follow-up minus baseline) among participants receiving standard medical care (n = 5,112)

Variable	Regression coefficient	Regression coefficient/ standard error
Age (years)	–0.0054	–0.3
Screen 2 HDL cholesterol (mg/dL)	–0.3109	–29.3
Screen 1 diastolic blood pressure (mm Hg)	–0.0322	–1.7
Screen 1 cigarettes/day	–0.0324	–4.6
Change in diastolic blood pressure (mm Hg)	0.1233	9.1
Change in thiocyanate[a] (mg/L)	–0.0106	–4.4
Change in body mass index (kg/ m^2)	–1.2702	–18.0
Change in drinks/week	0.1053	9.5
Change in SGOT (IU/L)	0.1160	8.4

[a] A biochemical marker used to measure the actual number of cigarettes smoked recently.

Nutrition

In a dietary reproducibility study, sucrose intake was assessed on the 1st day of each of 4 weeks separated in time by approximately 3 months. A one-way ANOVA was run with results given in Table 15.6.

Table 15.6 One-way ANOVA results for sucrose data, dietary reproducibility study (n = 173)

	SS	*df*	MS
Between	473,103.2	172	2750.6
Within	366,725.4	519	706.6

15.34 Is there a significant difference between mean sucrose intake for different individuals?

15.35 Estimate the between-person and within-person components of variation for sucrose intake.

15.36 What is the intraclass correlation for sucrose intake? What does it mean, in words?

Cardiology

Suppose a drug to relieve anginal pain is effective within 8 hours in 30% of 100 patients studied.

15.37 Derive an upper one-sided 95% CI for the percentage of patients who could get pain relief from the drug within 8 hours.

15.38 Suppose that if the patients are untreated, 15% will be free from pain within 8 hours. Assuming that the drug either benefits the patients or has no effect at all, what is your opinion on the effectiveness of the drug?

Suppose that in the same study 20% of the patients become pain-free within 4 hours of administration of the drug.

15.39 Derive an upper one-sided CI for the percentage of patients who could get pain relief from the drug within 4 hours.

15.40 Suppose 10% of untreated patients will be pain-free within 4 hours. Assuming that the drug either benefits the patients or has no effect at all, what is your opinion on the effectiveness of the drug?

Ophthalmology

A new drug is developed to relieve the ocular symptoms of hay fever. The drug is composed of two components, A and B: A is supposed to relieve itching of the eye and B is supposed to prevent redness; the combination is supposed to relieve both itching and redness. FDA regulations require that each component be proven effective both separately and in combination. Three experiments are performed on a group of 25 patients with hay fever. In the first experiment, drug A is administered to a randomly selected eye and a placebo is administered to the other eye, and the change from baseline is noted for each eye. The data are given in Table 15.7. A plus sign represents more improvement in the drug-treated eye; a minus sign represents more improvement in the placebo-treated eye; 0 represents equal improvement in both eyes.

Table 15.7 Comparison of drug A versus placebo for the relief of redness and itching in hayfever patients

Subject	Redness	Itching	Subject	Redness	Itching
1	0	+	13	+	+
2	0	+	14	0	0
3	–	–	15	0	0
4	+	–	16	–	0
5	0	+	17	–	+
6	+	+	18	–	0
7	+	+	19	0	0
8	–	0	20	–	+
9	–	+	21	–	0
10	0	–	22	–	+
11	+	+	23	0	0
12	0	+	24	–	0
			25	–	0

15.41 Why might a nonparametric statistical test be useful in comparing drug A with placebo for this experiment?

15.42 Why is it important to administer the placebo to the second eye of the same person rather than to a different group of people with hay fever?

15.43 What nonparametric statistical test would you use to compare drug A with placebo for redness or itching?

15.44 Compare redness between the drug-A vs placebo eyes and report a *p*-value.

15.45 Compare itching between the drug-A vs placebo eyes and report a *p*-value.

In the second experiment, drug B is administered to a randomly selected eye and a placebo is administered to the other eye. The data are given in Table 15.8.

Table 15.8 Comparison of drug B versus placebo for the relief of redness and itching in hayfever patients

Subject	Redness	Itching	Subject	Redness	Itching
1	+	–	13	+	0
2	+	–	14	+	0
3	+	+	15	+	0
4	+	+	16	0	0
5	+	+	17	+	–
6	0	–	18	+	–
7	+	0	19	+	+
8	+	+	20	+	+
9	+	0	21	+	0
10	0	0	22	+	0
11	+	–	23	+	0
12	+	+	24	+	0
			25	+	0

15.46 Compare redness between drug-B vs. placebo eyes and report a *p*-value.

15.47 Compare itching between the drug-B vs. placebo eyes and report a *p*-value.

In the third experiment, the combination of drugs A and B is administered to a randomly selected eye, and a placebo is administered to the other eye. The data are given in Table 15.9.

Table 15.9 Comparison of combination of drugs A and B versus placebo for the relief of redness and itching in hay-fever patients

Subject	Redness	Itching	Subject	Redness	Itching
1	+	+	13	+	+
2	+	0	14	+	0
3	+	+	15	+	0
4	+	+	16	+	+
5	+	+	17	+	+
6	+	–	18	+	+
7	+	+	19	0	+
8	+	0	20	0	+
9	+	+	21	+	0
10	+	+	22	+	0
11	+	–	23	+	0
12	+	–	24	+	0
			25	+	+

15.48 Compare redness between the combination vs. placebo eyes and report a *p*-value.

15.49 Compare itching between the combination vs. placebo eyes, and report a *p*-value.

15.50 Summarize the results of the three experiments as concisely as possible.

Cancer

Suppose a clinical trial is performed to assess the effect of a new treatment for cancer of the esophagus. No attempt is made to match the patients in any way because the number of cases is too small. It is found that of 100 patients who were given the standard treatment, 6 lived for 3 years and 5 lived for 5 years. Correspondingly, of 47 patients who were given the new treatment, 10 survived for 3 years, whereas 2 survived for 5 years.

15.51 Is there any evidence that the new treatment is helpful for the 3-year prognosis of the patient?

15.52 Is there any evidence that the new treatment is helpful for the 5-year prognosis?

15.53 Suppose a person has survived for 3 years. Is there any evidence for a treatment effect on the prognosis for the next 2 years?

15.54 Suppose we want to use the Wilcoxon signed-rank test and have a rank sum of 27 based on 7 untied pairs. Evaluate the significance of the results.

15.55 Answer Problem 15.54 for a rank sum of 65 based on 13 untied pairs.

15.56 Answer Problem 15.54 for a rank sum of 90 with 15 untied pairs.

Cancer

A study was conducted to look at the relationship between self-reports of mole count and prevalence of malignant melanoma [4]. 98 cases of malignant melanoma occurring between 1976 and 1982 and 190 age-matched controls among members of the Nurses' Health Study (age 30–54) were selected for analysis. The total number of moles in the left and right arms combined was compared between cases and controls. The data are given in Table 15.10.

Table 15.10 Number of moles in left and right arm (combined)

	0	1–5	6–15	16–30	31+	Total
Cases	16	20	20	12	30	98
Controls	57	40	45	20	28	190

Source: Reprinted with permission of the *American Journal of Epidemiology*, **127**(4),703–712,1988.

15.57 Test the hypothesis that cases, on average, have more moles (or less moles) than controls.

15.58 Estimate the odds ratio relating specific mole-count levels (1–5, 6–15, 16–30, 31+, respectively) versus having no moles to malignant melanoma. Provide a 95% CI for each odds-ratio estimate.

The prevalence in 1973 of a rare disease in two communities, A and B, was compiled and it was found that of 100,000 people in community A, 5 have the disease, whereas of 200,000 people in community B, only 1 has the disease.

15.59 Test whether the underlying prevalence rates are significantly different in the 2 communities.

Hypertension
Refer to Table 12.30 (Chapter 12, text).

15.60 Compute the intraclass correlation between replicate SBP measurements, and provide a 95% CI about this estimate.

Hypertension
Suppose we are interested in investigating the effect of race on level of blood pressure. A study was conducted in Evans County, Georgia, comparing the mean level of blood pressure between whites and blacks for different age-sex groups [5]. The mean and standard deviation of systolic blood pressure (SBP) among 25–34-year-old white males were reported as 128.6 mm Hg and 11.1 mm Hg, respectively, based on a large sample.

15.61 Suppose the mean SBP of 38 25–34-year-old black males is reported as 135.7 mm Hg. If the standard deviation of white males is assumed to hold for black males as well, then test if the underlying mean SBPs are the same in the two groups.

15.62 The actual reported standard deviation among 25–34-year-old black males is 12.5 mm Hg. Test the hypothesis that the underlying variance of blood pressure for white and black males is the same.

15.63 Suppose the actual underlying mean for black males is 135 mm Hg. What is the power of the test in Problem 15.61 in this case if we assume that the standard deviations for white males and black males are the same?

Suppose we do *not* assume that the standard deviation of 11.1 mm Hg is correct for 25–34-year-old black males.

15.64 Test the hypothesis that the underlying mean SBPs are the same in the two groups.

15.65 Derive a 95% CI for the mean SBP for 25–34-year-old black males under the assumptions in Problem 15.64.

15.66 Relate your answers to Problems 15.64 and 15.65.

15.67 Suppose we have two normally distributed samples of sizes 12 and 16 with sample standard deviations of 12.6 and 6.2, respectively. Test if the variances are significantly different in the two samples.

Endocrinology, Bone and Joint Disease
A study was conducted to relate serum estrogens to bone metabolism.

15.68 Suppose that a correlation of .52 is found between serum estradiol and the bone density of the lumbar region of the spine, as computed by CT scan among 23 post-menopausal women. Test for the statistical significance of these results and report a *p*-value.

15.69 Similarly, a correlation of .34 is found between serum estradiol and the bone density of the distal radius of the nondominant arm among the same 23 post-menopausal women. Test for the statistical significance of these results and report a *p*-value.

15.70 Finally, a correlation coefficient of .395 was computed between the two preceding measures of bone metabolism in Problems 15.68 and 15.69 based on the same group of women. Test for the statistical significance of these results and report a *p*-value.

15.71 Suppose in a previous large study, other investigators found a correlation of .5 between the two measures of bone metabolism in Problems 15.68 and 15.69. Test for whether or not the sample results in Problem 15.70 are compatible with these previous results. Report a *p*-value.

Cancer
A case-control study was performed to look at the effect of cigarette smoking on the risk of cervical cancer [6]. There were 230 cervical-cancer cases and 230 controls consisting of women ages 22–74 matched to the cases for age within 5-year intervals and admitted to the hospital for acute conditions other than malignancy, hormonal, or gynecologic disorders. The smoking habits of these women are given in Table 15.11. Suppose we wish to compare each smoking group versus nonsmokers.

15.72 Compute the odds ratio and associated 95% CI's comparing ex-smokers, current smokers of < 15 cigarettes per day, and current smokers of ≥ 15 cigarettes per day, respectively, versus nonsmokers.

15.73 What is your overall interpretation of the data?

Cancer
An investigation was conducted looking at the incidence of leukemia after treatment for Hodgkin's disease according to the type of treatment received [7]. The data in Table 15.12 were reported.

Table 15.11 Association between cigarette smoking and the risk of cervical cancer

			Current smokers	
	Nonsmokers	Ex-smokers	< 15 cigarettes/day	≥ 15 cigarettes/day
Cervical-cancer cases	155	13	33	29
Controls	169	21	22	18

Source: Reprinted with permission by the *American Journal of Epidemiology*, **123**(1), 22–29, 1986.

Table 15.12 Relationship of type of treatment for Hodgkin's disease and incidence of leukemia

Treatment procedure	Number of leukemias	Number of person-years of follow-up after treatment
(A) No intensive therapy	2	3322.8
(B) Either intensive radiotherapy or chemotherapy	4	1858.2

Source: Reprinted with permission of the *Journal of the National Cancer Institute*, **67**(4), 751–760, 1981.

Suppose we assume under the null hypothesis that the type of treatment has no effect on the incidence of leukemia.

15.74 What is the expected proportion of leukemias in the no-intensive-therapy group under the null hypothesis?

15.75 What test procedure can be used to test if there is a different incidence density of leukemia in group A than expected under the null hypothesis, if the total number of leukemias (= 6) are considered fixed?

15.76 Implement the test procedure in Problem 15.75 and report a *p*-value.

Infectious Disease, Pulmonary Disease, Hospital Epidemiology
A study was performed to compare reactivity to tuberculin among nurses working in different hospitals vs. the general population. In particular, suppose that according to statewide, age-specific rates, 31% of nurses in a hospital are expected to have positive tuberculin skin tests. In the hospital under study, 93 out of 221 nurses are tuberculin positive.

15.77 Provide a 95% CI for the proportion of positive tuberculin skin tests among nurses in the hospital.

15.78 How does the nurses' observed rate compare with the expected rate of 31%?

Occupational Health
To assess quantitatively the association between benzene exposure and cancer mortality, a study was conducted among 1165 men who worked in a rubber-hydrochloride plant with occupational exposure to benzene. A comparison was made between observed cancer mortality during the period 1950–1981 and expected cancer mortality as derived from U.S. white male mortality rates over the same time period. The data in Table 15.13 were obtained [8].

Table 15.13 Mortality data for workers exposed to benzene

	Number of deaths	
Cause of death	Observed	Expected
Total cancer	69	66.8
Multiple myeloma	4	1.0

15.79 What are the standardized mortality ratios (SMRs) for total cancer and multiple myeloma in this group?

15.80 Perform a test to comparing the observed and expected total cancer mortality in this cohort. Report a *p*-value.

15.81 Perform a test comparing the observed and expected multiple myeloma rates in this cohort. Report a *p*-value.

Infectious Disease
Refer to the lymph-node data in Table 8.18 (p. 319, text).

15.82 Use nonparametric methods to assess whether there are systematic differences between the assessments of Doctor A versus Doctor B regarding the number of palpable lymph nodes among sexual contacts of AIDS or ARC patients.

15.83 Are there differences between your conclusions in Problem 15.82 and your conclusions using parametric methods in Problem 8.69 (Chapter 8, text)?

15.84 Use nonparametric methods to estimate the between-observer reproducibility in the assessment of the number of palpable lymph nodes.

Cardiovascular Disease, Nutrition
A study was performed of the effects of walnut consumption on serum lipids and blood pressure [9]. The same subjects were administered a regular cholesterol-lowering diet (called the *reference diet*) for 4 weeks and an experimental diet substituting 3 servings of walnuts per day for portions of some foods in the reference diet for 4 weeks. The total calories, protein, carbohydrates, and fat were kept the same during each

period. The order in which the diets were administered was randomized for individual subjects. 18 subjects completed the protocol. Table 15.14 shows the values obtained for mean HDL cholesterol at the end of each 4-week dietary period.

15.85 What does it mean that the "order in which the diets were administered was randomized"?

15.86 From the information provided, estimate the standard error of the difference in mean HDL cholesterol between the two diets.

15.87 Perform a statistical test to assess if the mean HDL cholesterol was the same for the reference versus walnut diets. Report a *p*-value as precisely as possible using the tables in the text.

Table 15.14 Comparison of HDL cholesterol while on walnut diet versus reference diet

	Reference diet	Walnut diet	Mean difference (walnut diet-reference diet)	95% CI for mean difference
HDL cholesterol (mg/dL)	47 ± 11[a]	45 ± 10[a]	–2.3	(–3.9, –0.7)

[a] Mean ± sd.

15.88 Which of the following statements, if any, are inconsistent with the data in Table 15.14 and why? (no additional calculation is necessary). (a) *p*-value in Problem 15.87 $< .05$; (b) 90% CI for mean difference $= (-4.1, -0.5)$, (c) 98% CI for mean difference $= (-4.2, -0.4)$.

Pulmonary Disease

A study was performed comparing radiologic findings between bronchitic and normal individuals. Eighteen randomly selected 18–29-year-old bronchitic women and 17 randomly selected 18–29-year-old normal women were chosen for study. It was found that 7 out of 18 bronchitic women had abnormal chest x-ray findings, while 0 out of 17 normal women had abnormal findings.

15.89 What test procedure can be used to compare the percentage of abnormal findings between bronchitic and normal women?

15.90 Implement the test procedure and report a *p*-value (versus the one-sided alternative that bronchitic women have more abnormal radiologic findings than normal women).

15.91 Is the preceding design a paired- or independent-samples design?

Cardiovascular Disease

Some reports in the literature have noted that there is an inverse relationship between moderate alcohol consumption and cholesterol levels. To test this hypothesis, a questionnaire is administered to a group of workers at a particular company as to their alcohol consumption, and blood samples are taken to measure their cholesterol levels. The workers are subdivided into those who report no alcohol consumption, those who drink ≤ 2 oz of alcohol, and those who drink more than 2 oz of alcohol on an average day. The 23 nondrinkers had an average cholesterol level of 205.6 mg/dL with a standard deviation of 25.3 mg/dL. The 15 light drinkers had an average cholesterol level of 182.7 mg/dL with a standard deviation of 21.9 mg/dL. The 12 heavy drinkers had an average cholesterol level of 199.8 mg/dL with a standard deviation of 30.3 mg/dL.

15.92 Test the hypothesis that there is an overall difference in underlying mean cholesterol levels among the three groups.

15.93 Use the method of multiple comparisons to test for significant differences in mean cholesterol levels between each pair of groups.

Obstetrics

Assume that the distribution of birthweights in the general population is normal with mean 120 oz and standard deviation 20 oz. We wish to test a drug that, when administered to mothers in the prenatal period, will reduce the number of low-birthweight infants. We anticipate that the mean birthweight of the infants whose mothers are on the drug will be 1/2 lb heavier than that for the general newborn population.

15.94 How large a sample is needed to have an 80% probability of finding a significant difference if a one-sided test with significance level .05 is used?

15.95 Answer Problem 15.94 for a 90% probability.

SOLUTIONS

15.1 Under H_0, we have $\bar{d} \sim N(0,\ \sigma^2/n)$ where $\bar{d}$ = mean difference of before-after, $\sigma = 10$, $n = 50$. Therefore,

$$\alpha = \Pr\left[\bar{d} > 2.0 \middle| d \sim N\left(0,\ \frac{\sigma^2}{n}\right)\right]$$
$$= 1 - \Phi\left(\frac{2}{\frac{\sigma}{\sqrt{n}}}\right) = 1 - \Phi\left(\frac{2\sqrt{n}}{\sigma}\right)$$
$$= 1 - \Phi\left(\frac{2\sqrt{50}}{10}\right) = 1 - \Phi(1.414) = .079$$

15.2 We have

$$\text{Power} = \Pr\left[\bar{d} > 2.0 \middle| \bar{d} \sim N\left(3,\ \frac{\sigma^2}{n}\right)\right]$$
$$= 1 - \Phi\left[\frac{(2-3)\sqrt{n}}{\sigma}\right] = 1 - \Phi\left(\frac{-\sqrt{n}}{\sigma}\right)$$
$$= \Phi\left(\frac{\sqrt{n}}{\sigma}\right) = \Phi\left(\frac{\sqrt{50}}{10}\right) = \Phi(0.71) = .760$$

Thus, if the underlying reduction in blood pressure is 3 mm Hg, then there is a 76% chance of showing that the drug is effective, based on a study with 50 subjects.

15.3 Case women,

$$\frac{89}{283} = .314;$$

control women,

$$\frac{640}{3833} = .167.$$

15.4 A 95% CI for case women is given by

$$.314 \pm 1.96\sqrt{\frac{.314(.686)}{283}} = .314 \pm 1.96(.0276)$$
$$= .314 \pm .054 = (.260,\ .369)$$

A 95% CI for control women is given by

$$.167 \pm 1.96\sqrt{\frac{.167(.833)}{3833}} = .167 \pm 1.96(.0060)$$
$$= .167 \pm .012 = (.155,\ .179)$$

15.5 The power is given in Equation 10.15 (text, Chapter 10) by

$$\text{Power} = \Phi\left[\frac{\Delta}{\sqrt{\frac{p_1q_1 + p_2q_2}{n}}} - z_{1-\alpha/2}\frac{\sqrt{2\bar{p}\bar{q}}}{\sqrt{p_1q_1 + p_2q_2}}\right]$$

In this case, $p_1 = .6$, $p_2 = .4$, $\Delta = |.6 - .4| = .2$, $\alpha = .05$, $n = 20$, $\bar{p} = (.6 + .4)/2 = .5$. We have

$$\text{Power} = \Phi\left[\frac{0.2}{\sqrt{\frac{.6(.4)+.4(.6)}{20}}} - z_{.975}\frac{\sqrt{2(.5)(.5)}}{\sqrt{.6(.4)+.4(.6)}}\right]$$
$$= \Phi\left(\frac{.2}{.155} - 1.96 \times \frac{.707}{.693}\right) = \Phi(1.291 - 2.000)$$
$$= \Phi(-0.709) = .239$$

Thus, such a study would only have a 24% chance of detecting a significant difference.

15.6 We use the sample-size formula in Equation 10.14 (text, Chapter 10) as follows:

$$n = \frac{\left(\sqrt{2\bar{p}\,\bar{q}}\,z_{1-\alpha/2} + \sqrt{p_1q_1 + p_2q_2}\,z_{1-\beta}\right)^2}{\Delta^2}$$
$$= \frac{\left[\sqrt{2(.5)(.5)}\,z_{.975} + \sqrt{.6(.4)+.4(.6)}\,z_{.8}\right]^2}{.2^2}$$
$$= \frac{\left[0.707(1.96) + 0.693(0.84)\right]^2}{.04} = \frac{1.968^2}{.04} = 96.8$$

Thus, we need to enroll 97 patients in each group in order to achieve 80% power.

15.7 The estimated power in Problem 15.5 would decrease and the required sample size in Problem 15.6 would increase.

15.8 We form the following 2×2 table:

		Coffee drinking		
		0	1+	
MI status	Yes	3	27	30
	No	57	113	170
		60	140	200

We use the chi-square test for 2×2 tables since all expected values are ≥ 5

$$\left(\text{the smallest expected value} = \frac{60 \times 30}{200} = 9.0\right).$$

We have the following test statistic:

$$X^2 = \frac{n(|ad - bc| - \frac{n}{2})^2}{(a+b)(c+d)(a+c)(b+d)}$$
$$= \frac{200[|3(113) - 57(27)| - 100]^2}{30 \times 170 \times 60 \times 140}$$
$$= 5.65 \sim \chi_1^2 \text{ under } H_0$$

From the chi-square table (Table 6, Appendix, text), we see that $\chi^2_{1,\ .975} = 5.02$, $\chi^2_{1,\ .99} = 6.63$, and thus because

$$5.02 < 5.65 < 6.63$$

it follows that $.01 < p < .025$. Therefore, there is a significant association between prior MI status and coffee drinking, with coffee drinkers having a higher incidence of prior MI.

15.9 The results in Problem 15.8 would be more convincing if we were able to establish a "dose-response" relationship between coffee drinking and MI status with the risk of MI increasing as the number of cups per day of coffee consumption increases. For this purpose, we form the following 2×4 table:

		Current coffee consumption (cups per day)				
		0	1	2	3+	
MI status	Yes	3	7	8	12	30
	No	57	43	42	28	170
		60	50	50	40	200

We perform the chi-square test for trend in binomial proportions using the score statistic 1, 2, 3, 4 for the coffee-consumption groups 0, 1, 2, 3+ respectively. From Equation 10.24 (text, Chapter 10), we have the test statistic $X_1^2 = A^2/B$ where

$$A = \sum_{i=1}^{k} x_i S_i - x\overline{S} = 3(1) + 7(2) + 8(3) + 12(4)$$
$$-30 \times \left[\frac{60(1) + \ldots + 40(4)}{200}\right]$$
$$= 89 - \frac{30(470)}{200} = 89 - 70.5 = 18.5$$
$$B = \overline{pq}\left[\sum_{i=1}^{k} n_i S_i^2 - \frac{\left(\sum_{i=1}^{k} n_i S_i\right)^2}{N}\right]$$
$$= \frac{30}{200} \times \frac{170}{200} \times \left[60(1)^2 + \ldots + 40(4)^2 - \frac{470^2}{200}\right]$$
$$= .1275(1350 - 1104.50) = .1275(245.5) = 31.30$$

Therefore,

$$X_1^2 = \frac{18.5^2}{31.30} = 10.93 \sim \chi_1^2 \text{ under } H_0.$$

Since $\chi^2_{1,\ .999} = 10.83 < X_1^2$, it follows that $p < .001$. Thus, there is a significant linear trend, with the rate of prior MI increasing as the number of cups/day of coffee consumed increases.

15.10 Let x_i be the serum cholesterol for the *i*th case and y_j be the serum cholesterol for the *j*th control. We assume $x_i \sim N(\mu_1, \sigma_1^2)$, $y_j \sim N(\mu_2, \sigma_1^2)$. We wish to test the hypothesis $H_0: \mu_1 = \mu_2$ versus $H_1: \mu_1 \neq \mu_2$. Since $s_1 = s_2$, we will assume equal variances, i.e., $\sigma_1^2 = \sigma_2^2$. Thus, we use the two-sample *t* test for independent samples with equal variances.

15.11 We have the test statistic

$$t = \frac{\bar{x} - \bar{y}}{\sqrt{s^2\left(\frac{1}{n_1} + \frac{1}{n_2}\right)}} = \frac{234.8 - 215.5}{\sqrt{47.3^2\left(\frac{1}{1493} + \frac{1}{707}\right)}} = \frac{19.3}{2.159}$$
$$= 8.94 \sim t_{1493+707-2} = t_{2198}$$

Since $t > t_{120,\ .9995} = 3.373 > t_{2198,\ .9995}$, it follows that $p < 2 \times (1 - .9995)$ or $p < .001$.

15.12 We use the power formula

$$\text{Power} = \Phi\left(-z_{1-\alpha/2} + \frac{|\Delta|}{\sqrt{\sigma_1^2/n_1 + \sigma_2^2/n_2}}\right)$$

In this case $\alpha = .05$, $z_{1-\alpha/2} = z_{.975} = 1.96$, $\Delta = 10$, $n_1 = 1493$, $n_2 = 707$, $\sigma_1^2 = \sigma_2^2 = 47.3^2$. Thus, we have

$$\text{Power} = \Phi\left(-1.96 + \frac{10}{\sqrt{47.3^2/1493 + 47.3^2/707}}\right)$$
$$= \Phi(-1.96 + 4.63) = \Phi(2.671) = .996$$

Thus, there is a 99.6% chance of finding a significant difference.

15.13 $\text{sem} = 500/\sqrt{20} = 111.8$

15.14 The standard deviation is a measure of variability for the birthweight of *one* infant. The standard error of the mean is a measure of variability for the *mean* birthweight of a group of *n* infants (in this case $n = 20$). The standard error will always be smaller than the standard deviation because a mean of more than one birthweight will be less variable in repeated samples than an individual birthweight.

15.15 $p = 2 \times Pr(t_{19} > 2.73)$. We refer to Table 5 (Appendix, text) and note that $t_{19,\ .99} = 2.539$, $t_{19,\ .995} = 2.861$. Since $2.539 < 2.73 < 2.861$, it follows that

$$2 \times (1 - .995) < p < 2 \times (1 - .99)$$

or $.01 < p < .02$. The exact p-value obtained by computer is $p = .013$.

15.16 We assume that the distribution of age is normally distributed in each group. We have the test statistic $F = s_1^2/s_2^2 = 6.2^2/4.3^2 = 2.08 \sim F_{56,\ 47}$ under H_0. Since $F = 2.08 > F_{24,\ 40,\ .975} = 2.01 > F_{56,\ 47,\ .975}$, it follows that $p < .05$ and there are significant differences between the variances.

15.17 A two-sample t test with unequal variances.

15.18 We have the test statistic

$$t = \frac{\bar{x}_1 - \bar{x}_2}{\sqrt{\frac{s_1^2}{n_1} + \frac{s_2^2}{n_2}}} = \frac{58.1 - 52.6}{\sqrt{\frac{6.2^2}{57} + \frac{4.3^2}{48}}} = \frac{5.5}{1.029} = 5.34$$

We determine the effective df from Equation 8.21 (Chapter 8, text) as follows:

$$d' = \frac{\left(\frac{6.2^2}{57} + \frac{4.3^2}{48}\right)^2}{\frac{(6.2^2/57)^2}{56} + \frac{(4.3^2/48)^2}{47}} = 99.5$$

Therefore, $t = 5.34 \sim t_{99}$ under H_0. Since

$$t > t_{60,\ .9995} = 3.460 > t_{99,\ .9995},$$

it follows that $p < .001$ and there are significant differences in mean age between the two groups.

15.19 The 95% CI is given by

$$\bar{x}_1 - \bar{x}_2 \pm t_{99,\ .975}\sqrt{\frac{s_1^2}{n_1} + \frac{s_2^2}{n_2}} = 5.5 \pm t_{99,\ .975}(1.029)$$

We estimate $t_{99,\ 0.975}$ using the computer and obtain 1.984. Therefore, the 95% CI is given by

$$5.5 \pm 1.984(1.029) = 5.5 \pm 2.0 = (3.5,\ 7.5)$$

15.20 We assume that number of years of education is normally distributed in each group. We have the test statistic $F = s_2^2/s_1^2 = 0.8^2/0.7^2 = 1.31 \sim F_{47,\ 56}$ under H_0. Since $F < F_{\infty,\ 60,\ .975} = 1.48 < F_{47,\ 56,\ .975}$, it follows that $p > .05$ and there is no significant difference between the variances.

15.21 The two-sample t test with equal variances.

15.22 The pooled-variance estimate is given by

$$s^2 = \frac{0.7^2(56) + 0.8^2(47)}{103} = 0.558$$

Therefore, the test statistic is obtained from

$$t = \frac{\bar{x}_1 - \bar{x}_2}{\sqrt{s^2\left(\frac{1}{n_1} + \frac{1}{n_2}\right)}} = \frac{11.8 - 12.7}{\sqrt{0.558\left(\frac{1}{57} + \frac{1}{48}\right)}} = \frac{-0.9}{0.146}$$
$$= -6.1 \sim t_{103} \text{ under } H_0$$

Since $|t| > t_{60,\ .9995} = 3.460 > t_{103,\ .9995}$ it follows that $p < .001$. Therefore, there are highly significant differences in mean educational level between the two groups.

15.23 A 95% CI is given by $-0.9 \pm t_{103,\ .975}(0.146)$. We obtain $t_{103,\ .975}$ from the computer $= 1.983$. Therefore, the 95% CI is

$$-0.9 \pm 1.983(0.146) = -0.9 \pm 0.29 = (-1.19,\ -0.61).$$

Thus, in summary, patients in fee-for-service practices are significantly older and have significantly fewer years of education than patients in prepaid group health plans.

15.24 The distinction between a one-sided and two-sided test in this case is that for a one-sided test we would test the hypothesis $H_0: \mu_1 = \mu_2$ versus $H_1: \mu_1 > \mu_2$ or, alternatively, $H_0: \mu_1 = \mu_2$ versus $H_1: \mu_1 < \mu_2$, where μ_1 represents mean SBP (systolic blood pressure) sitting upright and μ_2 represents mean SBP lying down. For a two-sided test we would test the hypothesis $H_0: \mu_1 = \mu_2$ versus $H_1: \mu_1 \neq \mu_2$.

15.25 A two-sided test is appropriate here, since we have no preconceived notions as to the relative orderings of μ_1 and μ_2 and would be equally interested in the outcomes $\mu_1 < \mu_2$ and $\mu_1 > \mu_2$, (i.e., we don't know if SBP is higher while sitting upright or lying down in a bed).

15.26 Because each person is serving as his or her own control, we are dealing with highly dependent samples and must use the paired t test. We test the hypothesis $H_0: \mu_d = 0$ versus $H_1: \mu_d \neq 0$, where d_i = sitting SBP – lying SBP for the ith person and

$$d_i \sim N(\mu_d,\ \sigma_d^2).$$

We have the following set of within-pair differences: –12, –6, +2, –8, –8, 0, –12, +4, 0, –16. Compute the test statistic

$$t = \frac{\bar{d}}{\frac{s_d}{\sqrt{n}}} = \frac{-5.60}{\frac{6.786}{\sqrt{10}}} = \frac{-5.60}{2.146} = -2.61$$

under H_0, $t \sim t_9$ and we have from Table 5 (Appendix, text) that $t_{9,\ .975} = 2.262 < |t|$.

Therefore, H_0 would be rejected at the 5% level and the hypothesis that position affects level of SBP, with the sitting upright position having the lower blood pressure, would be accepted.

15.27 We have $t_{9,\ .975} = 2.262$, $t_{9,\ .99} = 2.821$. Since

$$t_{9,\ .99} = 2.821 > |t|,$$

$.02 < p < .05$. Using the computer, the exact p-value $= 2 \times Pr(t_9 > 2.61) = .028$.

15.28 We wish to test the hypothesis $H_0: \rho_1 = \rho_2$ versus $H_1: \rho_1 \neq \rho_2$. We use the test statistic

$$\lambda = \frac{z_1 - z_2}{\sqrt{\frac{1}{n_1 - 3} + \frac{1}{n_2 - 3}}} \sim N(0,\ 1) \text{ under } H_0$$

We have

$$z_1 = .5 \times \ln\left(\frac{1 + .402}{1 - .402}\right) = .5 \times \ln(2.344) = 0.426$$
$$z_2 = .5 \times \ln\left(\frac{1 + .053}{1 - .053}\right) = .5 \times \ln(1.112) = 0.053$$

Thus,

$$\lambda = \frac{0.426 - 0.053}{\sqrt{\frac{1}{900} + \frac{1}{441}}} = \frac{0.373}{0.058} = 6.42 \sim N(0,\ 1) \text{ under } H_0$$

Clearly, $p < .001$ and thus, there is a highly significant difference between the two correlations.

15.29 We fit the least-squares line with $b = L_{xy}/L_{xx}$ and $a = (\sum y_i - b\sum x_i)/n$. We have that

$$L_{xx} = 1740.5 - \frac{107^2}{7} = 104.93$$
$$L_{xy} = 4637.6 - \frac{107(300.2)}{7} = 48.83$$
$$L_{yy} = 12{,}900.2 - \frac{(300.2)^2}{7} = 25.93$$

It follows that $b = 48.83/104.93 = 0.465$,

$$a = \frac{300.2 - 0.465(107)}{7} = 35.77.$$

Thus, the least-squares line is $y = 35.77 + 0.465x$.

15.30 We use the regression F test. We have that

$$\text{Reg SS} = \frac{L_{xy}^2}{L_{xx}} = \frac{48.83^2}{104.93} = 22.72 = \text{Reg MS}$$
$$\text{Res SS} = \text{Total SS} - \text{Reg SS} = 25.93 - 22.72 = 3.21$$
$$\text{Res MS} = \frac{\text{Res SS}}{n - 2} = \frac{3.21}{5} = 0.641$$

Therefore, we have the test statistic

$$F = \frac{\text{Reg MS}}{\text{Res MS}} = \frac{22.72}{0.641} = 35.4 \sim F_{1,\ 5} \text{ under } H_0.$$

Since $F_{1,\ 5,\ .995} = 22.78$, $F_{1,\ 5,\ .999} = 47.18$ and

$$22.78 < F < 47.18,$$

it follows that $1 - .999 < p < 1 - .995$ or

$$.001 < p < .005.$$

Thus, there is a significant association between HDL cholesterol and SGOT.

15.31 We have that $\hat{y} = 35.77 + 0.465(11) = 40.89$ mg/dL. The standard error of this estimate is given by

$$se = \sqrt{\text{Res MS}\left[\frac{1}{n} + \frac{(x - \bar{x})^2}{L_{xx}}\right]}$$
$$= \sqrt{0.641\left[\frac{1}{7} + \frac{(11 - \frac{107}{7})^2}{104.93}\right]}$$
$$= \sqrt{0.641(0.143 + 0.175)} = 0.45 \text{ mg/dL}$$

15.32 If we represent the old and new x-values by x_{old} and x_{new}, then we are given that $x_{\text{new}} = x_{\text{old}} + .5$. It follows that $\bar{x}_{\text{new}} = \bar{x}_{\text{old}} + .5$. This should have no effect on L_{xy} or L_{xx} since the terms in x in these expressions are of the form $x_i - \bar{x}$, which should be unaffected by a change in origin. Thus

$$b_{\text{new}} = \frac{(L_{xy})_{\text{new}}}{(L_{xx})_{\text{new}}} = \frac{(L_{xy})_{\text{old}}}{(L_{xx})_{\text{old}}} = b_{\text{old}},$$

and the estimate of the slope is unaffected by the change in x. However, the estimate of the intercept will change in the following way:

$$a_{\text{new}} = \bar{y}_{\text{new}} - b_{\text{new}}\bar{x}_{\text{new}}$$
$$= \bar{y}_{\text{old}} - b_{\text{old}}(\bar{x}_{\text{old}} + .5)$$
$$= (\bar{y}_{\text{old}} - b_{\text{old}}\bar{x}_{\text{old}}) - .5b_{\text{old}}$$
$$= a_{\text{old}} - .5b_{\text{old}} < a_{\text{old}}$$

Thus, the estimate of the intercept is reduced by $0.5\ b_{\text{old}}$.

15.33 The results indicate that after simultaneously controlling for the effects of all other risk factors in the multiple-regression model, that increases in alcohol consumption and increases in SGOT levels, respectively, were associated with increases in HDL-cholesterol levels after controlling for the effects of all other risk factors. The results as regards SGOT levels supports the data presented in Table 15.4 relating the HDL-cholesterol levels to SGOT levels at one point in time.

15.34 We perform the F test for one-way ANOVA. We have the F statistic

$$F = \frac{2750.6}{706.6} = 3.89 \sim F_{172,\,519} \text{ under } H_0 .$$

Since $F > F_{24,\,120,\,.999} = 2.40 > F_{172,\,519,\,.999}$, it follows that $p < .001$. Thus, there is a significant difference between mean sucrose intake for different people.

15.35 The between-person variance $= \sigma_A^2$ is estimated by $(2750.6 - 706.6)/4 = 511.0$. The within-person variance $= \sigma^2$ is estimated by 706.6.

15.36 The intraclass correlation

$$\frac{511.0}{511.0 + 706.6} = \frac{511.0}{1217.6} = .42 .$$

It represents the proportion of the total variability that is attributable to between-person variation. Sucrose is a much less variable nutrient than dietary cholesterol, where the intraclass correlation was .13 (see Section 12.6, Chapter 12, Study Guide).

15.37 The upper one-sided 95% CI is obtained from the formula

$$p > \hat{p} - z_{.95}\sqrt{\frac{\hat{p}\hat{q}}{n}} = .30 - 1.645\sqrt{\frac{.3(.7)}{100}}$$
$$= .30 - 1.645 \times .046 = .30 - .075 = .225$$

15.38 Since 15% is not within the one-sided 95% CI computed in Problem 15.37, it follows that the true rate of pain relief is greater with the drug than if the patients were untreated. Therefore, the drug has proven benefit.

15.39 The upper one-sided 95% CI is given by

$$p > \hat{p} - z_{.95}\sqrt{\frac{\hat{p}\hat{q}}{n}} = .20 - 1.645\sqrt{\frac{.2(.8)}{100}}$$
$$= .20 - 1.645 \times .04 = .20 - .066 = .134$$

15.40 Since 10% is not within the one-sided 95% CI computed in Problem 15.39 we conclude that the drug is beneficial as regards pain relief within 4 hours.

15.41 A nonparametric statistical test would be useful because the change in redness or itching cannot be numerically quantified; only which eye has improved more can be assessed.

15.42 The degree of redness or itching is probably similar in two eyes of the same person at baseline. Thus, this procedure would eliminate the substantial between-person variability that can occur if the drug were administered to one group of people and the placebo to another group.

15.43 The sign test should be used here because we are comparing two paired samples consisting of alternate eyes from the same people and we only know which eye did better but not how much better.

15.44 There are 5 +'s and 11 –'s for redness. Since there are < 20 nonzero differences, the exact test must be used. In particular, the p-value is given by

$$p = 2 \times \sum_{k=0}^{5} \binom{16}{k} \left(\frac{1}{2}\right)^{16}$$

From the binomial tables (Table 1, Appendix, text) using $n = 16$, $p = .50$

$$Pr(0) = .0000 \qquad Pr(1) = .0002$$
$$Pr(2) = .0018 \qquad Pr(3) = .0085$$
$$Pr(4) = .0278 \qquad Pr(5) = .0667$$

Thus,

$$p = 2 \times (.0000 + .0002 + .0018 + .0085 + .0278 + .0667)$$
$$= .21.$$

Thus, there is no significant difference in redness between the drug A and placebo treated eyes.

15.45 There are 12 +'s and 3 –'s for itching. Again use the exact test, calculated from the binomial tables, with the p-value given by

$$p = 2 \times \sum_{k=0}^{3} \binom{15}{k} \left(\frac{1}{2}\right)^{15}$$
$$= 2 \times (.0000 + .0005 + .0032 + .0139)$$
$$= .035$$

Thus, there is a significant difference in itching between the drug-treated and placebo-treated eyes. This is what is expected, since drug A is only supposed to be effective for itching.

15.46 For redness there are 22 +'s and 3 0's. We can use the sign test using the normal approximation, since we have 22 untied pairs. The test statistic is given by

$$z = \frac{|C - \frac{n}{2}| - .5}{\sqrt{\frac{n}{4}}} = \frac{|22 - 11| - .5}{\sqrt{\frac{22}{4}}} = \frac{10.5}{2.345}$$
$$= 4.48 \sim N(0, 1) \text{ under } H_0$$

The p-value is

$$2 \times [1 - \Phi(z)] = 2 \times [1 - \Phi(4.48)] < .001$$

There is clearly overwhelming evidence that drug B is effective for redness.

15.47 For itching there are 7 +'s and 6 –'s. We use the exact binomial version of the sign test. The p-value is obtained from the binomial tables as follows:

$$p = 2 \times \sum_{k=0}^{6} \binom{13}{k} \left(\frac{1}{2}\right)^{13}$$
$$= 2 \times (.0001 + .0016 + .0095 + .0349$$
$$+ .0873 + .1571 + .2095)$$
$$= 1.0$$

There is clearly no significant difference in itching between drug B treated and placebo treated eyes. These results are again consistent with our expectations for drug B, which is only supposed to be effective for redness.

15.48 For redness there are 23 +'s and 2 0's. This result is even more extreme than that in Problem 15.46, and if we use the sign test with the normal approximation, we surely will find $p < .001$. Thus, the combination drug is very effective for redness.

15.49 For itching there are 14 +'s and 3 –'s. We use the exact binomial version of the sign test, as follows:

$$p = 2 \times \sum_{k=0}^{3} \binom{17}{k} \left(\frac{1}{2}\right)^{17}$$
$$= 2 \times (.0000 + .0001 + .0010 + .0052) = .013$$

Thus, the combination drug is significantly better than the placebo for itching as well, but not as strongly as for redness.

15.50 The results of the three experiments agree with our prior hypotheses. Indeed, drug A is significantly better than the placebo for itching ($p = .035$) but not for redness; drug B is significantly better than the placebo for redness ($p < .001$) but not for itching; the combination drug is significantly better than the placebo for both redness ($p < .001$) and itching ($p = .013$).

15.51 We have the following 2×2 table for the 3-year prognosis:

		3-Year Mortality Status: Alive	Dead	
Treatment	Standard	6	94	100
	New	10	37	47
		16	131	147

We can use the chi-square test here since the smallest expected value $= (16 \times 47)/147 = 5.12 \geq 5$. The chi-square statistic with Yates' correction is given as follows:

$$X^2 = \frac{n(|ad - bc| - \frac{n}{2})^2}{(a+b)(c+d)(a+c)(b+d)}$$
$$= \frac{147[|6(37) - 10(94)| - 73.5]^2}{16 \times 131 \times 100 \times 47}$$
$$= \frac{147(718 - 73.5)^2}{16 \times 131 \times 100 \times 47}$$
$$= \frac{147(644.5)^2}{16 \times 131 \times 100 \times 47}$$
$$= 6.20 \sim \chi_1^2 \text{ under } H_0$$

Since $\chi^2_{1,\,.975} = 5.02$, $\chi^2_{1,\,.99} = 6.63$ and

$$5.02 < 6.20 < 6.63,$$

we have $.01 < p < .025$. Therefore, the new treatment is significantly better than the standard treatment as regards the 3-year prognosis.

15.52 We have the following 2×2 table for the 5-year prognosis:

		5-Year Mortality Status: Alive	Dead	
Treatment	Standard	5	95	100
	New	2	45	47
		7	140	147

We use Fisher's exact test here because the smallest expected value $= (7 \times 47)/147 = 2.24 < 5$. We must enumerate all the possible tables with the same row and column margins as follows:

0	100	1	99	2	98	3	97	4	96
7	40	6	41	5	42	4	43	3	44

5	95	6	94	7	93
2	45	1	46	0	47

We then use the HYPGEOMDIST function of Excel and obtain:

$$Pr(0) = 2.470 \times 10^{-4},\ Pr(1) = .004,\ Pr(2) = .030,$$
$$Pr(3) = .113,\ Pr(4) = .250,$$
$$Pr(5) = .320,\ Pr(6) = .220,\ Pr(7) = .063.$$

Since our table is the "5" table, it follows that the two-sided p-value

$$2 \times \min(2.470 \times 10^{-4} + \ldots + .320,\ .320 + .220 + .063,\ .5)$$
$$= 2 \times \min(.717,\ .603,\ .5) = 2 \times .50 = 1.0,$$

which is clearly not statistically significant. Therefore, there is no significant difference between the two treatments as regards the 5-year prognosis.

15.53 We can look at the set of people who are alive after 3 years to answer this question. We have the following 2×2 table:

		Status after 5 years: Alive	Dead	
Treatment	Standard	5	1	6
	New	2	8	10
		7	9	16

We again use Fisher's exact test, because some of the expected values are < 5. We enumerate all the possible tables with the preceding row and column totals:

0	6	1	5	2	4	3	3	4	2
7	3	6	4	5	5	4	6	3	7

5	1	6	0
2	8	1	9

We use the HYPGEOMDIST function of Excel and obtain:

$$Pr(0) = .0105,\ Pr(1) = .110,\ Pr(2) = .330,$$
$$Pr(3) = .367,\ Pr(4) = .157,\ Pr(5) = .024,\ Pr(6) = .001.$$

Our table is the "5' table and the two-tailed p-value

$$2 \times \min(.0105 + \ldots + .024,\ .024 + .001,\ .5) = 2 \times .024$$
$$= .049$$

which is statistically significant. Thus, if a person has survived for 3 years he or she is significantly more likely to survive for the next 2 years with the standard treatment.

15.54 We refer to Table 11 (Appendix, text). Because the upper two-tailed critical value $= 26$ for $\alpha = .05$ and 28 for $\alpha = .02$, it follows that $.02 \le p < .05$.

15.55 We refer to Table 11 (Appendix, text) under $n = 13$, $\alpha = .10$. We note that the lower and upper critical values are 21 and 70. Because $21 < 65 < 70$, it follows that $p \ge .10$ and we accept H_0 at the 5% level.

15.56 The exact tables for the signed-rank test must be used, because the number of untied pairs is less than 16. Refer to Table 11 (Appendix, text) under $n = 15$ and note that the upper critical value for $\alpha = .10$ is 90, whereas the upper critical value for $\alpha = .05$ is 95. Since

$$R_1 = 90 \ge 90 \text{ and } 90 < 95,$$

we have $.05 \le p < .10$ using a two-sided test.

15.57 We will use the chi-square test for trend. We have the test statistic $X_1^2 = A^2/B$, where we use scores of 1, 2, 3, 4, and 5 for the 5 mole groups. We have

$$A = 1(16) + 2(20) + 3(20) + 4(12) + 5(30)$$
$$-\left(\frac{98}{288}\right)[1(73) + 2(60) + 3(65) + 4(32) + 5(58)]$$
$$= 314 - 274.26 = 39.74$$
$$B = \left[\frac{98(190)}{288^2}\right]\left\{1^2(73) + 2^2(60) + 3^2(65) + 4^2(32) + 5^2(58) - \frac{[1(73) + 2(60) + 3(65) + 4(32) + 5(58)]^2}{288}\right\}$$
$$= .2245\left(2860 - \frac{806^2}{288}\right) = 135.66$$

Thus,

$$X_1^2 = \frac{39.74^2}{135.66} = 11.64 \sim \chi_1^2 \text{ under } H_0.$$

Since $11.64 > 10.83 = \chi^2_{1,\ .999}$, it follows that $p < .001$. Therefore, we reject H_0. Since $A > 0$, we conclude that the mole count of cases is significantly higher than for controls.

15.58 The odds ratio comparing the 1-5 mole group to the 0 mole group $= 20(57)/[16(40)] = 1.78$. Using the Woolf approach,

$$se\left[\ln\left(\hat{\text{OR}}\right)\right] = \sqrt{\frac{1}{20} + \frac{1}{57} + \frac{1}{16} + \frac{1}{40}} = 0.3938.$$

Therefore, a 95% CI for $\ln(\text{OR})$ is

$$\ln(1.78) \pm 1.96(0.3938) = (-0.1944,\ 1.3491).$$

The corresponding 95% CI for

$$\text{OR} = [\exp(-0.1944),\ \exp(1.3491)] = (0.82,\ 3.85).$$

Similar methods are used for each of the other 3 mole-count groups versus the 0 mole-count group, with results summarized as follows:

Odds ratio and 95% CI for the 1–5, 6–15,16–30, and 31+ mole-count groups, respectively, versus the 0 mole-count group

Mole-count group	$\hat{OR}$	95% CI
1–5 moles	1.78	(0.82, 3.85)
6–15 moles	1.58	(0.74, 3.40)
16–30 moles	2.14	(0.86, 5.29)
31+ moles	3.82	(1.79, 8.14)

Interestingly, there is only a significant difference between the 31+ mole-count group versus the 0 mole-count group. However, the overall trend reveals an increasing odds ratio as the mole count increases. Therefore, the chi-square test for trend is the most appropriate method for analyzing these data.

15.59 The 2×2 contingency table is shown as follows:

	Disease status		
Community	+	–	Total
A	5	99,995	100,000
B	1	199,999	200,000
Total	6	299,994	300,000

The expected table is formed as follows:

2.0	99,998
4.0	199,996

Since two of the four cells have expected values < 5, Fisher's exact test rather than the χ^2 test must be used. Thus, the following seven tables with the same margins as the observed table are formed:

0	100,000
6	199,994

1	99,999
5	199,995

2	99,998
4	199,996

3	99,997
3	199,997

4	99,996
2	199,998

5	99,995
1	199,999

6	99,994
0	200,000

We use the HYPGEOMDIST function of Excel to compute the probability of each table as follows:

$Pr(0) = .0878$ $Pr(1) = .263$ $Pr(2) = .329$
$Pr(3) = .219$ $Pr(4) = .082$ $Pr(5) = .016$
$Pr(6) = .001$

The one-sided p-value associated with the "5' table is $.016 + .001 = .018$. The two-sided p-value is .036. Thus, there is a significant difference between the prevalence rates in the 2 communities.

15.60 We first perform a one-way ANOVA using MINITAB, as follows:

Analysis of variance on SBP					
Source	DF	SS	MS	F	P
Day	9	152.0	16.9	1.16	0.405
Error	10	145.0	14.5		
Total	19	297.0			

From Equation 12.33 (text, Chapter 12) we have $\hat{\sigma}^2 = 14.5$, $\hat{\sigma}_A^2 = (16.9 - 14.5)/2 = 1.2$. The estimated intraclass correlation

$$\hat{\rho}_I = \frac{1.2}{1.2 + 14.5} = \frac{1.2}{15.7} = .076.$$

From Equation 12.36 (text, Chapter 12) a 95% CI for ρ_I is given by (c_1, c_2), where

$$c_1 = \max\left(\frac{\frac{F}{F_{9,\,10,\,.975}-1}}{\frac{F}{F_{9,\,10,\,.975}+1}},\ 0\right);\ c_2 = \max\left(\frac{\frac{F}{F_{9,\,10,\,.025}-1}}{\frac{F}{F_{9,\,10,\,.025}+1}},\ 0\right)$$

We use MINITAB to obtain $F_{9,\,10,\,.975} = 3.78$ and $F_{9,\,10,\,.025} = 0.253$. Thus,

$$c_1 = \max\left(\frac{\frac{1.16}{3.78}-1}{\frac{1.16}{3.78}+1},\ 0\right) = \max\left(\frac{-0.692}{1.308},\ 0\right) = 0$$

$$c_2 = \max\left(\frac{\frac{1.16}{0.253}-1}{\frac{1.16}{0.253}+1},\ 0\right) = \max\left(\frac{3.604}{5.604},\ 0\right) = .64$$

Therefore, the 95% CI for $\rho_I = (0,\ .64)$.

15.61 We use a two-sided one-sample z test here, because the standard deviation is known. We have $H_0: \mu = 128.6$, $\sigma = 11.1$ versus $H_1: \mu \neq 128.6$, $\sigma = 11.1$. It is appropriate to use a two-sided test because it is not known in which direction the mean blood pressure will fall. We base our significance test on

$$z = \frac{\bar{x} - \mu_0}{\frac{\sigma}{\sqrt{n}}} = \frac{135.7 - 128.6}{\frac{11.1}{\sqrt{38}}} = 3.94$$

We have that $p = 2 \times \Pr[z \geq 3.94 | z \sim N(0,\ 1)] < .001$. Thus, this difference is very highly significant.

15.62 We use a one-sample χ^2 test for the variance of a normal distribution here to test the hypothesis $H_0: \sigma^2 = 11.1^2$ versus $H_1: \sigma^2 \neq 11.1^2$. We base our significance test on

$$X^2 = \frac{(n-1)s^2}{\sigma^2} = \frac{37(12.5)^2}{(11.1)^2} = 46.92$$

which under H_0 follows a χ^2_{37} distribution. We note from Table 6 (Appendix, text) that

$$\chi^2_{37,\ .975} > \chi^2_{30,\ .975} = 46.98 > X^2 .$$

Thus, it follows that $p > 2(.025) = .05$, and the variances of systolic blood pressure for white and black males are not significantly different.

15.63 The power of a test $= Pr$ (reject $H_0 | H_1$ true). If we use a two-sided 5% significance test, then the power is given by the formula

$$\Phi\left(z_{\alpha/2} + |\mu_1 - \mu_0| \frac{\sqrt{n}}{\sigma} \right)$$

We have $\mu_0 = 128.6$, $\mu_1 = 135$, $\sigma = 11.1$, $n = 38$, and $\alpha = .05$. It follows that

$$\text{Power} = \Phi\left[-1.96 + \frac{(135 - 128.6)\sqrt{38}}{11.1} \right] = \Phi(1.594) = .94$$

Thus, we have a 94% chance of finding a significant difference using a two-sided test with an α level $= .05$ and a sample size of 38.

15.64 We have the test statistic

$$t = \frac{\bar{x} - \mu_0}{\frac{s}{\sqrt{n}}} = \frac{135.7 - 128.6}{\frac{12.5}{\sqrt{38}}} = \frac{7.1}{2.028}$$
$$= 3.50 \sim t_{37} \text{ under } H_0$$

We see from Table 5 (Appendix, text) that

$$t_{30,\ .995} = 2.750,\ t_{30,\ .9995} = 3.646,$$
$$t_{40,\ .995} = 2.704,\ t_{40,\ .9995} = 3.551.$$

Thus, because $2.750 < t < 3.646$, and $2.704 < t < 3.551$, it follows that if we had either 30 or 40 *df*, then we would obtain $.001 < p < .01$. The exact *p*-value obtained by computer is

$$p = 2 \times Pr(t_{37} > 3.50) = .0012 .$$

Therefore, we again conclude that there is a significant difference between the mean SBPs of 25–34-year-old black and white males.

15.65 A 95% CI for μ is given by $\bar{x} \pm t_{37,\ .975}\, s/\sqrt{n}$. We compute $t_{37,\ .975}$ using MINITAB, and obtain 2.026. Thus, the 95% CI for μ is given by

$$135.7 \pm \frac{2.026(12.5)}{\sqrt{38}} = 135.7 \pm 4.1 = (131.6,\ 139.8) .$$

15.66 The 95% CI in Problem 15.65 does not contain the mean for white males $= 128.6$. This must be the case, because the *p*-value for the test in Problem 15.64 was $< .05$.

15.67 We test the hypothesis $H_0: \sigma_1^2 = \sigma_2^2$ versus $H_1: \sigma_1^2 \neq \sigma_2^2$. We compute the *F* statistic

$$\frac{s_1^2}{s_2^2} = \frac{12.6^2}{6.2^2} = 4.13 \sim F_{11,\ 15} \text{ under } H_0 .$$

Since $F > F_{8,\ 14,\ .975} = 3.29 > F_{11,\ 15,\ .975}$, it follows that $p < 2 \times (1 - .975)$ or $p < .05$, and the variances are significantly different.

15.68 We have the test statistic

$$t = \frac{r\sqrt{n-2}}{\sqrt{1-r^2}} = \frac{.52\sqrt{21}}{\sqrt{1-(.52)^2}} = \frac{2.383}{0.854}$$
$$= 2.79 \sim t_{21} \text{ under } H_0$$

Since $t_{21,\ .99} = 2.518$, $t_{21,\ .995} = 2.831$ and

$$2.518 < 2.79 < 2.831 ,$$

it follows that

$$2 \times (1 - .995) < p < 2 \times (1 - .99)$$

or $.01 < p < .02$.

15.69 We have the test statistic

$$t = \frac{.34\sqrt{21}}{\sqrt{1-(.34)^2}} = \frac{1.558}{0.940} = 1.66 \sim t_{21} \text{ under } H_0$$

Since $t_{21,\ .975} = 2.080 > t$, it follows that $p > .05$ and there is no significant association between serum estradiol and the bone density of the distal radius.

15.70 We have the test statistic

$$t = \frac{.395\sqrt{21}}{\sqrt{1-(.395)^2}} = \frac{1.810}{0.919} = 1.970 \sim t_{21} \text{ under } H_0$$

Since $t_{21,\ .975} = 2.080$, $t_{21,\ .95} = 1.721$ and

$$1.721 < 1.970 < 2.080 ,$$

it follows that $.05 < p < .10$ and there is a trend toward a significant association between the two measures of bone metabolism.

15.71 We use the one-sample z test for correlation to test the hypothesis $H_0: \rho = \rho_0$ versus $H_1: \rho \neq \rho_0$. We have the test statistic

$$\lambda = \sqrt{n-3}(z - z_0) \sim N(0,\ 1) \text{ under } H_0$$

where

$$z_0 = .5 \times \ln\left(\frac{1+.5}{1-.5}\right) = 0.549$$

from Table 13 (Appendix, text) and

$$z = .5 \times \ln\left(\frac{1+.395}{1-.395}\right) = .5 \times \ln(2.306) = 0.418$$

Thus,

$$\lambda = \sqrt{20}(0.418 - 0.549)$$
$$= -0.589 \sim N(0,\ 1) \text{ under } H_0$$

This yields a p-value of

$$2 \times \Phi(-0.589) = 2(.2781) = .56.$$

Thus, our results are compatible with the previous results, because no statistical significance was established.

15.72 The odds ratio for ex-smokers versus nonsmokers is

$$\hat{\text{OR}} = \frac{13(169)}{155(21)} = 0.67.$$

Using the Woolf approach,

$$se\left[\ln\left(\hat{\text{OR}}\right)\right] = \sqrt{\frac{1}{13} + \frac{1}{21} + \frac{1}{155} + \frac{1}{169}} = 0.3700.$$

Therefore, a 95% CI for $\ln(\text{OR})$ is

$$\ln(0.67) \pm 1.96(0.3700)$$
$$= -0.3931 \pm 0.7252 = (-1.1183,\ 0.3321)$$

The corresponding 95% CI for OR is

$$[\exp(-1.1183),\ \exp(0.3321)] = (0.33,\ 1.39).$$

The odds ratio for current smokers of < 15 cigarettes per day versus nonsmokers is

$$\frac{33(169)}{155(22)} = 1.64.$$

We have $se\left[\ln\left(\hat{\text{OR}}\right)\right] = \sqrt{\frac{1}{33} + \frac{1}{169} + \frac{1}{155} + \frac{1}{22}} = 0.2969$. A 95% CI for $\ln(\text{OR})$ is

$$\ln(1.64) \pm 1.96(0.2969)$$
$$= 0.4919 \pm 0.5818 = (-0.0899,\ 1.0738)$$

The corresponding 95% CI for OR is

$$[\exp(-0.0899),\ \exp(1.0738)] = (0.91,\ 2.93).$$

The odds ratio for current smokers of ≥ 15 cigarettes per day versus nonsmokers is

$$\frac{29(169)}{18(155)} = 1.76.$$

We have $se\left[\ln\left(\hat{\text{OR}}\right)\right] = \sqrt{\frac{1}{29} + \frac{1}{169} + \frac{1}{18} + \frac{1}{155}} = 0.3200$. A 95% CI for $\ln(\text{OR})$ is

$$\ln(1.76) \pm 1.96(0.3200)$$
$$= 0.5634 \pm 0.6272 = (-0.0638,\ 1.1906)$$

The corresponding 95% CI for OR is

$$[\exp(-0.0638),\ \exp(1.1906)] = (0.94,\ 3.29).$$

15.73 The overall impression is that there is an excess number of current smokers among cervical-cancer cases versus controls. Although the light and heavy current smokers, when considered separately do not differ significantly from the nonsmokers, if one combines them into one group there is a significant excess number of current smokers among cervical-cancer cases versus controls.

15.74 Under the null hypothesis, the type of treatment has no effect on the proportion of leukemias. Therefore, the proportion of leukemias in each group should be in direct proportion to the number of person-years of follow-up after treatment in each group. Thus,

$$p_0 = \frac{3322.8}{3322.8 + 1858.2} = \frac{3322.8}{5181.0} = .641.$$

15.75 We wish to compare incidence-density measures between groups A and B. Since

$$V_1 = 6(.641)(.359) = 1.38 < 5,$$

we must use the exact method in Equation 14.9 (text, Chapter 14).

15.76 Since $\hat{p} = 2/6 = .33 < p_0 = .641$, the p-value for the test in Problem 15.75 is given by

$$p = 2 \times \sum_{k=0}^{2} {}_6C_k p_0^k q_0^{6-k}$$
$$= 2 \times \left[(.359)^6 + 6(.641)(.359)^5 + {}_6C_2(.641)^2(.359)^4\right]$$
$$= 2 \times (.00213 + .02284 + .10209) = 2 \times .12706$$
$$= .254$$

Thus, there is no significant difference between the incidence density of leukemia in the no-intensive-therapy group versus the intensive-therapy group.

15.77 We have that $\hat{p} = 93/221 = .421$. Thus, a 95% CI for p is given by

$$\hat{p} \pm 1.96\sqrt{\frac{\hat{p}\hat{q}}{n}} = .421 \pm 1.96\sqrt{\frac{.421(.579)}{221}}$$
$$= .421 \pm .065 = (.356,\ .486)$$

15.78 Because 31% does not fall in this interval, we can conclude that the reactivity rate in this hospital is significantly higher than the statewide rate.

15.79 The SMR for total cancer

$$100\% \times \frac{69}{66.8} = 103\%.$$

The SMR for multiple myeloma

$$100\% \times \frac{4}{1} = 400\%.$$

15.80 Since the expected number of deaths $= 66.8 \geq 5$, we can use the large-sample test in Equation 7.49 (text, Chapter 7). We have the test statistic

$$X^2 = \frac{(O-E)^2}{E} = \frac{(69-66.8)^2}{66.8}$$
$$= 0.07 \sim \chi^2_1 \text{ under } H_0$$

Since $\chi^2_{1,\ .10} = 0.02$, $\chi^2_{1,\ .25} = 0.10$, and $.02 < .07 < .10$, it follows that $1 - .25 < p < 1 - .10$ or $.75 < p < .90$. Clearly, the results are not statistically significant; i.e., the cancer mortality rate in this cohort is not significantly different from the U.S. white male mortality rates.

15.81 Since the expected number of deaths $= 1.0 < 5.0$, we must use the small-sample test given in Equation 7.48 (text, Chapter 7). Since the observed number of deaths is greater than expected, we compute the p-value using the formula

$$p = 2 \times \sum_{k=4}^{\infty} \frac{e^{-1.0}1.0^k}{k!} = 2 \times \left(1 - \sum_{k=0}^{3} \frac{e^{-1.0}}{k!}\right)$$

Referring to the Poisson table (Table 2, Appendix, text) under $\mu = 1.0$, we obtain

$$p = 2 \times [1 - (.3679 + \ldots + .0613)]$$
$$= 2 \times (1 - .9810) = .038.$$

Thus, there is a significant excess of multiple-mycloma deaths among workers exposed to benzene.

15.82 We will use the Wilcoxon signed-rank test. The difference scores are ranked by absolute value as follows:

	Neg		Pos				
$\lvert d_i \rvert$	d_i	f_i	d_i	f_i	Total	Rank Range	Average Rank
10	–10	0	10	1	1	29	29.0
8	–8	0	8	1	1	28	28.0
6	–6	0	6	2	2	26–27	26.5
5	–5	0	5	3	3	23–25	24.0
4	–4	0	4	4	4	19–22	20.5
3	–3	1	3	10	11	8–18	13.0
2	–2	2	2	1	3	5–7	6.0
1	–1	1	1	3	4	1–4	2.5
		4		25	29		
0	0	3					

The rank sum of the positive differences is given by $R_1 = 29(1) + 28(1) + \ldots + 2.5(3) = 407.5$. Since there are 29 untied pairs, we can use the normal-theory test. The test statistic is

$$T = \frac{\left|407.5 - \frac{29(30)}{4}\right| - .5}{\sqrt{\frac{29(30)(59)}{24} - \frac{\left[\left(2^3-2\right)+\left(3^3-3\right)+\ldots+\left(4^3-4\right)\right]}{48}}}$$

$$= \frac{189.5}{\sqrt{2138.75 - 31.125}} = \frac{189.5}{\sqrt{2107.625}} = \frac{189.5}{45.91}$$

$$= 4.13 \sim N(0, 1)$$

The p-value $= 2 \times [1 - \Phi(4.13)] < .001$. Thus, there are systematic differences between the assessments of Doctor A and Doctor B, with Doctor A detecting a significantly greater number of palpable lymph nodes.

15.83 The results are same: Doctor A significantly detects more palpable lymph nodes than Doctor B.

15.84 We will compute a Spearman rank correlation between the lymph-node assessments of Doctors A and B. We first rank-order the assessments of Doctors A and B, separately.

To compute the Spearman rank-correlation coefficient, we must compute the ordinary Pearson correlation between ranks. The Spearman rank-correlation coefficient (r_s) is given by

$$r_s = \frac{2228.5}{\sqrt{2688(2674)}} = \frac{2228.5}{2681.0} = .831$$

This correlation is high. In general, the rank correlation and the signed-rank test provide complementary information. It is possible that the signed-rank test is not significant, whereby, on average, Doctors A and B are comparable, yet they may vary substantially for individual subjects, which would result in a low rank correlation. Conversely, the signed-rank test may be highly significant, which would indicate that on average Doctors A and B differ, but if this difference is consistent for each subject, then the rank correlation will be high. This seems to be the situation here.

15.85 This means that a random-number table (or a set of computer-generated random numbers) was used to select the order in which the diets were administered to individual subjects. Each subject had a probability of 1/2 of having the reference diet or the walnut diet first for 4 weeks, followed by the other diet for an additional 4 weeks. This does not guarantee that in a small study there will be exactly an equal number of subjects assigned to each diet first. Indeed, in the actual study, 10 subjects received the walnut diet first, while 8 subjects received the reference diet first.

15.86 The 95% CI for the true mean difference is

$$\bar{d} \pm t_{n-1,\ .975}(se) = -2.3 \pm t_{17,\ .975}(se).$$

The width of the 95% CI $= 2t_{17,\ .975}(se)$. Thus

$$2t_{17,\ .975}(se) = -0.7 - (-3.9) = 3.2$$

or

$$se = \frac{3.2}{2t_{17,\ .975}} = \frac{1.6}{2.110} = 0.758$$

15.87 Since we are using each person as his or her own control, we will use the paired t test. We wish to test the hypothesis $H_0: \mu_d = 0$ versus $H_1: \mu_d \neq 0$ where μ_d = true mean for HDL cholesterol while on the walnut diet − true mean for HDL cholesterol while on the reference diet. The test statistic is

$$t = \frac{\bar{d}}{\frac{s_d}{\sqrt{n}}} = \frac{\bar{d}}{se} = \frac{-2.3}{0.758} = -3.03 \sim t_{17} \text{ under } H_0$$

Therefore, the p-value $= 2 \times Pr(t_{17} < -3.03)$. From Table 5 (Appendix, text), we see that $t_{17,\ .995} = 2.898$, $t_{17,\ .9995} = 3.965$. Since $2.898 < |t| = 3.03 < 3.965$, it follows that $1 - .9995 < \frac{p}{2} < 1 - .995$ or

$$.0005 < \frac{p}{2} < .005$$

or $.001 < p < .01$. Thus, there is a significant difference between the mean HDL-cholesterol level while on the walnut diet versus the mean HDL- cholesterol while on the reference diet, with a significantly lower mean under the walnut diet.

15.88 (a) Since the two-sided 95% CI excludes zero, it follows that the p-value must be $< .05$. Thus, (a) is *true*. (b) The 90% CI must be narrower than the 95% CI. Thus, (b) must be *false* because the interval (–4.1, –0.5) is wider than the 95% CI $= (-3.9, -0.7)$. (c) The 98% CI must be wider than the 95% CI. Since the interval (–4.2, –0.4) is wider than (–3.9, –0.7), this is *not inconsistent* with the data in the table.

15.89 We first form the following 2×2 table:

		Chest X ray: Abnormal	Chest X ray: Normal	Total
Diagnosis	Normal	0	17	17
	Bronchitis	7	11	18
		7	28	35

The smallest expected value = 7(17)/35 = 3.40 < 5. Thus, we must use Fisher's exact test to analyze these data.

15.90 We wish to test the hypothesis $H_0: p_1 = p_2$ versus $H_1: p_1 < p_2$, where

p_1 = proportion of abnormal chest x - ray findings among normal women

p_2 = proportion of abnormal chest x - ray findings among bronchitic women

Thus, the p-value $= Pr(0)$ where $Pr(0)$ is the exact probability of the 0 table (i.e., the observed table) under H_0 given that the margins are fixed. This is obtained by using the HYPGEOMDIST function of Excel as follows: $Pr(0) = .005$. Thus, there is a significant association, with bronchitic women having significantly more abnormal chest x-ray findings than normal women.

15.91 This is an example of an independent-samples design because women in the two groups are *not matched* on an individual basis.

15.92 We use the fixed-effects one-way ANOVA. We have the test statistic

$$F = \frac{\text{Between MS}}{\text{Within MS}} \sim F_{k-1,\ n-k} \text{ under } H_0$$

We have

$$\text{Between MS} = \frac{\text{Between SS}}{2}$$
$$= \frac{(205.6)^2 23 + (182.7)^2 15 + (199.8)^2 12}{2} - \frac{\left[(205.6)23 + (182.7)15 + (199.8)12\right]^2}{50(2)}$$
$$= 975{,}985.56 - 973{,}557.16$$
$$= 2428.40$$

$$\text{Within MS} = s^2 = \frac{\left[22(25.3)^2 + 14(21.9)^2 + 11(30.3)^2\right]}{47}$$
$$= 657.35$$

Thus,

$$F = \frac{\text{Between MS}}{\text{Within MS}} = \frac{2428.40}{657.35}$$
$$= 3.69 \sim F_{2,\ 47} \text{ under } H_0$$

Since $F_{2,\ 47,\ .95} < F_{2,\ 40,\ .95} = 3.23 < F$, it follows that $p < .05$, and there is an overall significant difference among the groups.

15.93 We compute the test statistic, critical value, and level of significance using the Bonferroni approach as follows:

Group	Test statistic	Critical value	p-value
Non, light	$t = \frac{205.6 - 182.7}{\sqrt{657.35(\frac{1}{23} + \frac{1}{15})}} = \frac{22.9}{8.51} = 2.69$	$t_{47,\ .9917} = 2.48$	$p < .05$
Non, heavy	$t = \frac{205.6 - 199.8}{\sqrt{657.35(\frac{1}{23} + \frac{1}{12})}} = \frac{5.8}{9.13} = 0.64$	2.48	NS
Light, heavy	$t = \frac{182.7 - 199.8}{\sqrt{657.35(\frac{1}{15} + \frac{1}{12})}} = \frac{-17.1}{9.93} = -1.72$	2.48	NS

Thus, there is a significant difference between the mean cholesterol levels for nondrinkers versus light drinkers, but no significant difference between the mean cholesterol levels for either nondrinkers versus heavy drinkers or light drinkers versus heavy drinkers.

15.94 We use the sample-size formula given in Equation 7.26 (text, Chapter 7) as follows:

$$n = \frac{\sigma^2 (z_{1-\beta} + z_{1-\alpha})^2}{(\mu_1 - \mu_0)^2}$$

In this case,

$$\mu_0 = 120 \text{ oz},\ \mu_1 = 120 + 8 = 128 \text{ oz},\ \sigma = 20 \text{ oz},$$
$$\alpha = .05,\ \beta = 1 - .8 = .2.$$

Thus, we have

$$n = \frac{20^2 (z_{.8} + z_{.95})^2}{(120 - 128)^2} = \frac{400(0.84 + 1.645)^2}{64} = 38.6$$

Thus, we would need to administer the drug to 39 pregnant women to have an 80% chance of finding a significant difference using a one-sided test with $\alpha = .05$.

15.95 We use the same formula as in Problem 15.94, but with $\beta = 1 - .9 = .1$. We have

$$n = \frac{20^2 (z_{.9} + z_{.95})^2}{(120 - 128)^2} = \frac{400(1.28 + 1.645)^2}{64} = 53.5$$

Thus, we need 54 women in the study to have a 90% chance of finding a significant difference using a one-sided test with $\alpha = .05$.

REFERENCES

[1] Cramer, D. W., Schiff, I., Schoenbaum, S. C., Gibson, M., Belisle, S., Albrecht, B., Stillman, R. J., Berger, M. J., Wilson, W., Stadel, B. V., & Seibel, M. (1985). Tubal infertility and the intrauterine device. *New England Journal of Medicine*, **312**(15), 941–947.

[2] Applegate, W. B., Hughes, J. P., & Vander Zwaag, R. (1991). Case-control study of coronary heart disease risk factors in the elderly. *Journal of Clinical Epidemiology*, **44**(4), 409–415.

[3] Kuller, L. H., Hulley, S. B., Laporte, R. E., Neaton, J., & Dai, W. S. (1983). Environmental determinants, liver function, and high density lipoprotein cholesterol levels. *American Journal of Epidemiology*, **117**(4), 406–418.

[4] Bain, C., Colditz, G. A., Willett, W. C., Stampfer, M. J., Green, A., Bronstein, B. R., Mihm, M. C., Rosner, B., Hennekens, C. H., & Speizer, F. E. (1988). Self-reports of mole counts and cutaneous malignant melanoma in women: Methodological issues and risk of disease. *American Journal of Epidemiology*, **127**(4), 703–712.

[5] McDonough, J. R., Garrison, G. E., & Hames, C. G. (1967). Blood pressure and hypertensive disease among negroes and whites in Evans County, Georgia. In J. Stamler, R. Stamler, & T. N. Pullman (eds.), *The epidemiology of hypertension*, New York: Grune & Stratton.

[6] La Vecchia, C., Franceschi, S., Decarli, A., Fasoli, M., Gentile, A., & Tognoni, G. (1986). Cigarette smoking and the risk of cervical neoplasia. *American Journal of Epidemiology*, **123**(1), 22–29.

[7] Boivin, J. F., & Hutchison, G. B. (1981). Leukemia and other cancers after radiotherapy and chemotherapy for Hodgkin's disease. *Journal of the National Cancer Institute*, **67**(4), 751–760.

[8] Rinsky, R., Smith, A. B., Hornung, R., Filloon, T. G., Young, R. J., Okun, A. H., & Landrigan, P. J. (1987). Benzene and leukemia: An epidemiologic risk assessment. *New England Journal of Medicine*, **316**(17), 1044–1050.

[9] Sabate, J., Fraser, G. E., Burke, K., Knutsen, S. F., Bennett, H., & Lindsted, K. D. (1993). Effects of walnuts on serum lipid levels and blood pressure in normal men. *New England Journal of Medicine*, **328**(9), 603–607.