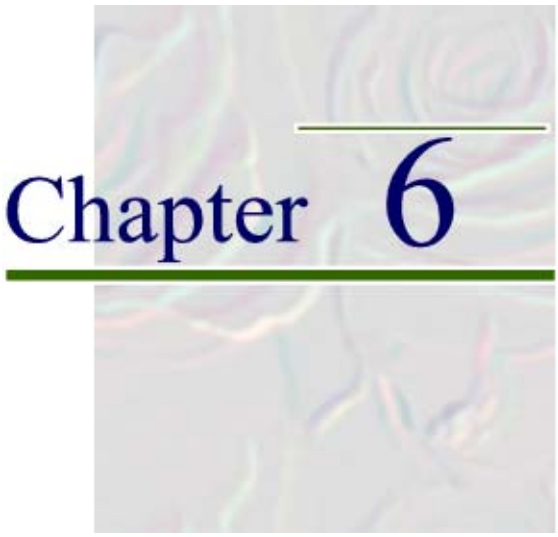


ESTIMATION

REVIEW OF KEY CONCEPTS

SECTION 6.1 Relationship of Population to Sample

Reference population Collection of all sampling units that potentially might be under study (or collection of all sampling units to whom we wish to generalize our findings).

Sample A finite collection of sampling units from the reference population.

Random sample A sample chosen such that each member of the reference population has an equal chance of being chosen. This is actually a simple random sample; a random sample can allow for unequal probabilities of selection. In this text, however, we will use the term *random sample* to mean a simple random sample.

SECTION 6.2 Random-Number Tables

Suppose we had a registry of 5000 children with a rare cancer. How can we select a random sample of 100 children from this registry?

Random digit Each digit 0, 1, … , 9 has a probability of 1/10 of being selected; successive digits are independent.

Table 4 (Appendix, text) contains a collection of 1000 random digits. To select a random sample of 100 children,

1. Start at row 10 of Table 4 and obtain the following groups of 5 consecutive random digits: *39446 01375 75264 51173 16638 04680 98617 90298*.
2. Use the first 4 digits from each group of 5 digits and obtain *3944 0137 7526 5117 1663 0468 9861 9029*.
3. Sort the original list of 5000 children according to alphabetical order, ID order, or some other prespecified order.

4. Select the children from the list according to the numbers obtained in step 2; e.g., the 1st child selected would be the 3944th on the list; the 2nd child selected would be the 137th on the list, ... , etc. If some of the numbers in step 2 are either zero, > 5000 or are duplicates of previously used numbers, then ignore them and go on to the next set of 5 numbers.

SECTION 6.3 Randomized Clinical Trials

An important use of random numbers is in randomized clinical trials. Patients are assigned to treatment groups at random, using random numbers obtained from either a table or generated using a computer. The primary rationale for randomization is that it ensures approximate comparability of treatment groups with respect to baseline characteristics.

6.3.1 Techniques of Study Design in Randomized Clinical Trials

1. **Blocking** An equal number of patients are assigned to each treatment group for every 2*n* patients. The purpose is to ensure that an equal number of patients will be assigned to each treatment group over the short run (i.e., every 2*n* patients), because study protocols sometimes change over time.
2. **Stratification** Separate random-number tables are used for each of a number of homogeneous strata (e.g., age groups). The purpose is to ensure balance between treatment groups within strata. The strata are usually constructed to have very different risk profiles for disease (e.g., different age groups).
3. **Blinding** The purpose of blinding is to avoid bias on the part of either the patient or the physician in the ascertainment of outcome. In **double-blind** clinical trials, neither the patient nor the physician is aware of the treatment assignment. In **single-blind** clinical trials, one of either the physician or the patient is aware of the treatment assignment, but not both. It is usually used when it is either impossible or impractical to conduct a double-blind study.

The gold standard of modern clinical research is the *randomized double-blind placebo-controlled* clinical trial. By placebo control, we mean that one treatment arm is an active treatment, while the other treatment arm is a *placebo*—a treatment that appears similar to active treatment to the patient, but actually has no biologic effect (e.g., sugar pills). In some instances, it is unethical to use a placebo control (e.g., when a standard treatment already exists for a life-threatening condition.)

SECTION 6.4 Sampling Distribution

A distribution of all possible values of a statistic of interest (e.g., $\overline{X}$) over all possible random samples that can be selected from the reference population.

SECTION 6.5 Estimation of the Mean of a Distribution

The sample mean is used to estimate the underlying or population mean of a distribution. Why is the sample mean a good estimator of the population mean?

1. It is unbiased (i.e., $E(\overline{X}) = \mu$).
2. It has minimum variance among all unbiased estimators.

By an *unbiased estimator*, we mean that the average value of the sample mean over all possible random samples of size *n* that can be selected from the reference population = the population mean.

SECTION 6.6 Standard Error of the Mean

The standard error of the mean (*se*) reflects the variability of the sample mean in repeated samples of size n drawn from the reference population. It differs from the standard deviation, which reflects the variation of individual sample points. It is given by $\sigma/\sqrt{n}$, where $\sigma =$ standard deviation and n = sample size. In situations where σ is unknown, the *se* is estimated by $s/\sqrt{n}$.

Example: Suppose we take blood pressure measurements on twenty 30–49-year-old males. The mean ± sd of diastolic blood pressure (DBP) is 78.5 ± 10.3 mm Hg. Our best estimate of μ is 78.5. The standard error (*se*) of this estimate is $10.3/\sqrt{20} = 2.3$. Notice that the standard error of the mean (2.3) is much smaller than the standard deviation (10.3).

The standard error is dependent on

1. $\sigma =$ the variability of individual sample points
2. n = sample size

The following table gives the *se* for mean DBP for different sample sizes (n):

n	*se*
1	10.3
20	2.3
100	1.0

SECTION 6.7 The Central-Limit Theorem

If the underlying distribution of the individual sample points is normal, then it can be shown that $\bar{X} \sim N\left(\mu,\ \sigma^2/n\right)$. What if our distribution is not normal? For large enough n, $\bar{X} \dot{\sim} N(\mu,\ \sigma^2/n)$ even if the distribution of the individual observations is not normal. This is crucial for the validity of many hypothesis-testing and confidence-interval methods that assume approximate normality for $\bar{X}$ in repeated samples of size n from the reference population. Based on the central-limit theorem, this will be true if n is large enough even if the original distribution is not normal.

SECTION 6.8 Interval Estimation for the Mean

Our best estimate of μ from the preceding data = 78.5 mm Hg. However, there is uncertainty in our estimate of μ as reflected by the standard error. A confidence interval for μ is a range of likely values for μ. A 95% confidence interval (CI) for μ based on the sample data is given by

$$\left(\bar{x} - \frac{t_{n-1,\ .975}s}{\sqrt{n}} < \mu < \bar{x} + \frac{t_{n-1,\ .975}s}{\sqrt{n}}\right)$$

where $t_{n-1,\ 0.975} = 97.5\text{th}$ percentile of a t distribution with $n-1$ degrees of freedom (df). The t distribution is a family of distributions indexed by the degrees of freedom (d). Percentiles of the t distribution are given in Table 5 (Appendix, text) and are also available in computer packages such as

Excel or MINITAB (see Appendix). The term 95% confidence interval means that 95% of the CI's constructed in this manner from repeated samples of size n from the reference population (i.e., 30–49-year-old males) will contain the true mean (μ).

Example: Compute a 95% CI for $\mu =$ true mean DBP for 30–49-year-old males. A 95% CI for μ is given by $78.5 \pm t_{19,\ .975}(10.3)/\sqrt{20} = 78.5 \pm 2.093(10.3)/\sqrt{20} = 78.5 \pm 4.8 = (73.7,\ 83.3)$.

6.8.1 Use of Confidence Intervals for Decision-Making Purposes

Suppose the community we are studying is near a major highway and we conjecture whether there is an increased BP level in this community relative to the national average. Suppose the national average SBP for age 65+-year-olds = 140.0 mm Hg based on previous national health surveys. We draw a sample of size 50 and find $\bar{x} = 150$ mmHg and $s = 20$ mmHg. Can we conclude that the average SBP for 65+-year-olds in our community is higher? We will construct a 95% CI for μ and use it to help us make the decision. We have

$$95\%\ \text{CI for}\ \mu = \bar{x} \pm t_{49,\ 0.975}\frac{s}{\sqrt{n}} = 150 \pm 2.01\frac{20}{\sqrt{50}} = 150 \pm 5.7 = (144.3,\ 155.7)$$

where we obtain $t_{49,\ .975}$ from Excel. Since this interval excludes the national average (140), we conclude that the true mean SBP for 65+-year-olds in this community is above the national average.

$$\text{In general, a } 100\% \times (1-\alpha)\ \text{CI for}\ \mu = \bar{x} \pm t_{n-1,\ 1-\alpha/2}\frac{s}{\sqrt{n}}.$$

6.8.2 Factors Influencing the Length of a Confidence Interval

$$\text{Length} = 2t_{n-1,\ 1-\alpha/2}\frac{s}{\sqrt{n}}$$

1. α: level of confidence; as α gets smaller, CI gets longer
2. s: as $s \uparrow$ length $\uparrow$
3. n: as $n \uparrow$ length $\downarrow$

Also, if n is very large (>200), a $100\% \times (1-\alpha)$. CI can be obtained approximately from $\bar{x} \pm z_{1-\alpha/2}\, s/\sqrt{n}$. In addition, this interval can be used for $n \le 200$ if σ is known by replacing s by σ.

SECTION 6.9 Estimation of the Variance of a Distribution

When are we interested in estimating σ^2? In reproducibility studies, estimation of σ^2 is often of primary interest. Suppose we are planning a study of the role of alpha tocopherol (vitamin E) on heart disease. We want to assess the reproducibility of serum alpha-tocopherol levels because we are setting up a new laboratory to perform this test.

We recruit 15 volunteers who provide two blood samples one week apart. We compute $d_i = x_{1i} - x_{2i}$, where $x_{1i} =$ 1st sample ith subject, $x_{2i} =$ 2nd sample ith subject, and find a mean difference $= \bar{d}$ of 0 and a standard deviation $= s$ of 0.25. In the literature, we find that $\sigma = 0.20$ for a similar experiment based on a much larger sample. Can we say that our lab is less reproducible than the literature laboratory? We need to obtain a 95% CI for σ^2. We will base our interval on s^2, which we use as a point estimate of σ^2. A 95% CI for σ^2 is given by the interval

$$\frac{(n-1)s^2}{\chi^2_{n-1,\ .975}} < \sigma^2 < \frac{(n-1)s^2}{\chi^2_{n-1,\ .025}}$$

where $\chi^2_{n-1,\ p}$ = the pth percentile of a chi-square distribution with $n-1$ *df*.

The chi-square distribution is a family of distributions indexed by a parameter d known as the *degrees of freedom* (or *df*). The pth percentile of a chi-square distribution with d *df* is denoted by $\chi^2_{d,\ p}$ and is given in Table 6 (see the Appendix in the main text). We can also obtain percentiles of a chi-square distribution from Excel or MINITAB (see Appendix).

Example: Compute a 95% CI for σ^2 for the vitamin-E reproducibility study. Since $n = 15$, $s = 0.25$, the 95% CI is given by

$$\left[\frac{14(0.25)^2}{\chi^2_{14,\ .975}},\ \frac{14(0.25)^2}{\chi^2_{14,\ .025}}\right] = \left[\frac{14(0.0625)}{26.12},\ \frac{14(0.0625)}{5.63}\right] = (0.0335,\ 0.1554)$$

$$95\%\text{ CI for }\sigma = \left(\sqrt{0.0335},\ \sqrt{0.1554}\right) = (0.183,\ 0.394)$$

This interval includes 0.20 and indicates that our lab is comparable in reproducibility to the laboratory in the literature.

SECTION 6.10 Estimation for the Binomial Distribution

6.10.1 Large-Sample Method

Some studies have suggested a relationship between exposure to anesthetic gases and incidence of breast cancer. To examine this relationship, a study was set up among 10,000 female operating-room nurses ages 30–49. Suppose that the 5-year risk of breast cancer in the general population in this age group = $p_0 = .005$.

We find that among 10,000 female operating-room nurses, 60 women have developed disease over 5 years. Is this a significant excess over the expected risk based on national rates? Our best estimate of the risk = $\hat{p} = 60/10{,}000 = .006$. This is a 20% excess over the national average. In this example,

p = true rate for 30-49-year-old female operating-room nurses, which is unknown

$p_o = .005$ = general-population rate for $40-59$-year-old women

$\hat{p}$ = estimate of $p = .006$

We will use $\hat{p}$ as an estimate of p. We need to obtain a 95% CI for p based on $\hat{p}$.

Based on the central-limit-theorem, an approximate 95% CI for p is given by the interval

$$\left(\hat{p} - 1.96\sqrt{\frac{\hat{p}\hat{q}}{n}} < p < \hat{p} + 1.96\sqrt{\frac{\hat{p}\hat{q}}{n}}\right)$$

We will only use this interval if $n\hat{p}\hat{q} \geq 5$ (this is the condition under which the normal approximation to the binomial distribution is valid).

In this example, the 95% CI for p is given by

$$\left(.006 - 1.96\sqrt{\frac{.006(0.994)}{10{,}000}},\ .006 + 1.96\sqrt{\frac{.006(.994)}{10{,}000}}\right) = (.006 - .0015,\ .006 + .0015)$$

$$= (.0045,\ .0075)$$

Since this interval includes $p_0 = .005$, we would conclude that there is not a significant excess risk of breast cancer among female operating-room nurses. It would still warrant setting up a better study based on a control group of unexposed nurses (e.g., female general-duty nurses who do not work in operating rooms).

6.10.2 Small-Sample Method

The drug erythromycin has been postulated to prevent infections in the last trimester of pregnancy and as a result prevent low-birthweight deliveries. A *matched-pair* study was set up comparing erythromycin (*E*) versus placebo (*P*). The patients were matched on age, parity, (i.e., number of children), race, smoking, and number of miscarriages. Fifty matched pairs of women were formed. One member of the matched pair randomly received (*E*) and the other member of the matched pair received (*P*). The results were as follows:

Outcome for *E* member of matched pair	+	+	–	–
Outcome for *P* member of matched pair	+	–	+	–
Number of matched pairs	1	2	6	41

where + indicates a low-birthweight delivery (i.e., ≤ 88 oz) and – indicates a normal birthweight delivery (i.e., > 88 oz). Thus, there are 2 (+, –) pairs, or in other words for two matched pairs, the erythromycin-treated woman had a low-birthweight infant, while the placebo-treated woman had a normal birthweight infant, 6 (–, +) pairs, …, etc.

How can we compare the *E* treatment versus the *P* treatment for the prevention of low-birthweight deliveries? We will focus on the discordant pairs (i.e., the pairs of women who had a differential response). If *E* and *P* are equally effective, then the number of (+, –) pairs should be about the same as the number of (–, +) pairs. Another way of saying this is that p = proportion of discordant pairs that are (+, –) should be 1/2. In this case, there are 8 (2 + 6) discordant pairs of which 2 are (+, –) pairs. Therefore, our best estimate of p is $\hat{p} = 2/8 = .25$. We also wish to compute a 95% CI for p. Since $n\hat{p}\hat{q} = 8(1/4)(3/4) = 1.5 < 5$, we can't use the large-sample method. Instead, we need to compute exact binomial confidence limits for p.

To obtain exact binomial confidence limits requires a computer or use of a nomogram such as the one given in Table 7 (in the Appendix of the main text). Nomograms are provided in Tables 7a and 7b, giving exact 95% and 99% confidence limits for p, respectively. To use these nomograms,

If $\hat{p} \le 0.50$,

1. Find $\hat{p}$ on the lower horizontal axis.
2. Draw a vertical line and note the points of intersection with the appropriate n curves.
3 For each point draw a horizontal line across to the left vertical scale. The lower value = p_1 and the upper value = p_2.

If $\hat{p} > .50$,

1. Find $\hat{p}$ on the upper horizontal axis.
2. Same as step 2 in first list.
3. For each point draw a horizontal line to the right vertical scale. The upper value = p_1, and the lower value = p_2.

In this case, we use Table 7a with $n = 8$ and $\hat{p} = 2/8 = .25$. The exact 95% CI for p is obtained from the left vertical axis and is (.03, .65). Since this interval includes .50, we conclude that *E* and *P* are equivalent; that is, erythromycin is no better than placebo for the prevention of low-birthweight deliveries.

SECTION 6.11 Estimation for the Poisson Distribution

Suppose we consider the data in Example 4.38 (text, Chapter 4). In this example, we described the cancer-mortality experience from 161 white male employees of two plants in Texas and Michigan who were

exposed to ethylene dibromide (EDB) during the time period 1940–1975. Seven deaths due to cancer were observed, while only 5.8 were expected based on U.S. white male cancer-mortality rates. Is the observed number of cases excessive?

We answered this question by calculating $Pr(X \geq 7|\mu = 5.8)$ where $X =$ observed number of cancer deaths, which was assumed to follow a Poisson distribution with mean $= \mu = 5.8$. Under this assumption, the probability was .36 and we concluded that the number of cancer deaths was not excessive in this group.

Another approach to this problem is to compute a 95% CI for μ based on the observed number of deaths $(x = 7)$. To obtain this CI, we will usually use exact methods based on Table 8 (see the Appendix in the text). We refer to the $x = 7$ row and find that the 95% CI for $\mu = (2.81, 14.42)$. Since this interval includes 5.8, we can again conclude that the observed number of deaths are not excessive in this group. We can also obtain a corresponding 95% CI for the standardized mortality ratio (SMR) given by $(2.81/5.8,\ 14.42/5.8) = (0.48,\ 2.49)$. The SMR is an index of (mortality rate in an exposed group)/(mortality rate in the general population). Since the 95% CI for the SMR includes 1, it again indicates that there is no excess cancer mortality risk among the workers in these plants.

SECTION 6.12 One-Sided Confidence Limits

Suppose it is hypothesized that anesthetic gases can only potentially be harmful and not protective with regard to breast cancer incidence. Instead of a CI of the form $Pr(c_1 < p < c_2) = 1 - \alpha$ we might want a CI for breast cancer incidence of the form $Pr(p > c_1) = 1 - \alpha$. That is, what is the minimum plausible value for the probability of breast cancer among exposed nurses? This is referred to as an *upper one-sided CI*. This interval is given by

$$p > \hat{p} - z_{1-\alpha}\sqrt{\frac{\hat{p}\hat{q}}{n}}$$

In this case, if we refer to Section 6.10.1, we have $\alpha = 0.05$, $z_{1-\alpha} = z_{.95} = 1.645$, $\hat{p} = .006$, $\hat{q} = .994$, $n = 10{,}000$. The interval is

$$p > .006 - 1.645\sqrt{\frac{.006(.994)}{10{,}000}} = .006 - 1.645(.0008) = .006 - .0013 = .0047$$

Thus, $p > .0047$ is an upper 95% CI. This includes $p_0 = .005$ and we again conclude that breast cancer incidence is not significantly elevated among female operating room nurses.

Similarly, a lower one-sided CI is an interval of the form $Pr(p < c_2) = 1 - \alpha$.

This might be useful in the birthweight example (Section 6.10.2) if erythromycin is seen as only being possibly protective. Such an interval is given by

$$p < \hat{p} + z_{1-\alpha}\sqrt{\frac{\hat{p}\hat{q}}{n}}$$

Suppose in the birthweight example that there were actually 100 discordant pairs instead of 8, which would enable us to use the normal approximation to the binomial distribution. In this case, $\hat{p} = .25$, $n = 100$ and the interval would be

$$p < .25 + 1.645\sqrt{\frac{.25(.75)}{100}} = .25 + 1.645(.0433) = .25 + .0712 = .321$$

Thus, $p < .321$ is a lower 95% CI for p, and because this interval excludes $p_0 = .5$ we would conclude that erythromycin prevents low birthweight deliveries.

Similar methods can be used to obtain one-sided confidence limits for other parameters (e.g., the mean of a normal distribution).

PROBLEMS

6.1 What is the upper 10th percentile of a chi-square distribution with 5 *df*?

6.2 What is the upper 1st percentile of a chi-square distribution with 3 *df*?

Pathology
The data in Table 6.1 are measurements from a group of 10 normal males and 11 males with left-heart disease taken at autopsy at a particular hospital. Measurements were made on several variables at that time, and the table presents the measurements on total heart weight (THW) and total body weight (BW). Assume that the diagnosis of left-heart disease is made independently of these variables.

6.3 Using the data in Table 6.1 construct a 99% confidence interval for the variance of total heart weight for left-heart disease males.

Table 6.1 Autopsy data of total heart weight (THW) and body weight (BW)

Left-heart disease males			Normal males		
Observation number	THW (g)	BW (kg)	Observation number	THW (g)	BW (kg)
1	450	54.6	1	245	40.8
2	760	73.5	2	350	67.4
3	325	50.3	3	340	53.3
4	495	44.6	4	300	62.2
5	285	58.1	5	310	65.5
6	450	61.3	6	270	47.5
7	460	75.3	7	300	51.2
8	375	41.1	8	360	74.9
9	310	51.5	9	405	59.0
10	615	41.7	10	290	40.5
11	425	59.7			

6.4 Using the data in Table 6.1 answer Problem 6.3 for normal males.

Cancer
A case-control study of the effectiveness of the Pap smear in preventing cervical cancer (by identifying precancerous lesions) was performed [1]. It was found that 28.1% of 153 cervical-cancer cases and 7.2% of 153 age-matched (within 5 years) controls had never had a Pap smear prior to the time of the case's diagnosis.

6.5 Provide a 95% CI for the percentage of cervical-cancer cases who never had a Pap test.

6.6 Provide a 95% CI for the percentage of controls who never had a Pap test.

6.7 Do you think the Pap test is helpful in preventing cervical cancer?

Hypertension
Hypertensive patients are screened at a neighborhood health clinic and are given methyl dopa, a strong antihypertensive medication for their condition. They are asked to come back 1 week later and have their blood pressures measured again. Suppose the initial and follow-up systolic blood pressures of the patients are given in Table 6.2.

To test the effectiveness of the drug, we want to measure the difference (D) between initial and follow-up blood pressures for each person.

6.8 What is the mean and sd of D?

6.9 What is the standard error of the mean?

6.10 Assume that D is normally distributed. Construct a 95% CI for μ.

6.11 Do you have any opinion on the effectiveness of methyl dopa from the results of these 10 patients?

Table 6.2 Initial and follow-up SBP (mm Hg) for hypertensive patients given methyl dopa

Patient number	Initial SBP	Follow-up SBP
1	200.0	188.0
2	194.0	212.0
3	236.0	186.0
4	163.0	150.0
5	240.0	200.0
6	225.0	222.0
7	203.0	190.0
8	180.0	154.0
9	177.0	180.0
10	240.0	225.0

Cancer
Data from U.S. cancer-tumor registries suggest that of all people with the type of lung cancer where surgery is the recommended therapy, 40% survive for 3 years from the time of diagnosis and 33% survive for 5 years.

6.12 Suppose that a group of patients who would have received standard surgery are assigned to a new type of surgery. Of 100 such patients, 55 survive for 3 years and 45 survive for 5 years. Can we say that the new form of surgery is better in any sense than the standard form of surgery?

Cardiovascular Disease
A recent hypothesis states that vigorous exercise is an effective preventive measure for subsequent cardiovascular death. To test this hypothesis, a sample of 750 men aged 50–75 who report that they jog at least 10 miles per week is ascertained. After 6 years, 64 have died of cardiovascular disease.

6.13 Compute a 95% CI for the incidence of cardiovascular death in this group.

6.14 If the expected death rate from cardiovascular disease over 6 years in 50–75-year-old men based on large samples is 10%, then can a conclusion be drawn concerning this hypothesis from these data?

Cardiovascular Disease
A group of 60 men under the age of 55 with a prior history of myocardial infarction are put on a strict vegetarian diet as part of an experimental program. After 5 years, 3 men from the group have died.

6.15 What is the best point estimate of the 5-year mortality rate in this group of men?

6.16 Derive a 95% confidence interval for the 5-year mortality rate.

6.17 Suppose that from a large sample of men under age 55 with a prior history of myocardial infarction, we know that the 5-year mortality rate is 18%. How does the observed mortality rate obtained in Problem 6.15 compare with the large sample rate of 18%?

Pulmonary Disease
A spirometric tracing is a standard technique used to measure pulmonary function. These tracings represent plots of the volume of air expelled over a 6-second period and tend to look like Figure 6.1. One quantity of interest is the slope at various points along the curve. The slopes are referred to as **flow rates**. A problem that arises is that the flow rates cannot be accurately measured, and some observer error is always introduced. To quantify the observer error, an observer measures the flow at 50% of forced vital capacity (volume as measured at 6 seconds) twice on tracings from 10 different people. A machine called a *digitizer* can trace the curves automatically and can estimate the flow mechanically. Suppose the digitizer is also used to measure the flow twice on these 10 tracings. The data are given in Table 6.3.

6.18 Find a 95% CI for the standard deviation of the difference between the first and second replicates using the manual method.

6.19 Answer Problem 6.18 for the difference between the first and second replicates using the digitizer method.

Suppose we want to compare the variability of the two methods within the same person. Let x_{i1}, x_{i2} represent the 2 replicates for the ith person using the manual method, and let y_{i1}, y_{i2} represent the 2 replicates for the ith person using the digitizer method. Let

$$d_i = |x_{i1} - x_{i2}| - |y_{i1} - y_{i2}|.$$

Then, d_i is a measure of the difference in variability using the two methods. Assume that d_i is normally distributed with mean μ_d and variance σ_d^2.

FIGURE 6.1 A typical spirometric tracing

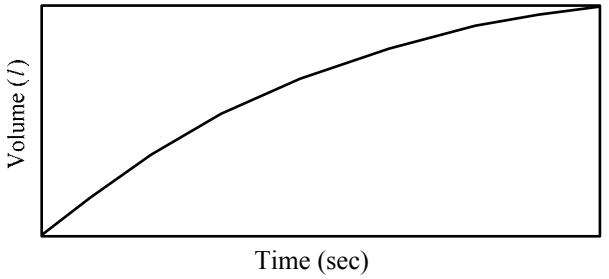


Table 6.3 Estimation of flow rates (L/sec) at 50% of forced vital capacity by a manual and a digitizer method

Person	Manual method replicate 1	Manual method replicate 2	Digitizer replicate 1	Digitizer replicate 2
1	1.80	1.84	1.82	1.83
2	2.01	2.09	2.05	2.04
3	1.63	1.52	1.62	1.60
4	1.54	1.49	1.49	1.45
5	2.21	2.36	2.32	2.36
6	4.16	4.08	4.21	4.27
7	3.02	3.07	3.08	3.09
8	2.75	2.80	2.78	2.79
9	3.03	3.04	3.06	3.05
10	2.68	2.71	2.70	2.70

6.20 Find a 95% CI for μ_d.

6.21 What is your opinion as to the relative variability of the two methods?

Turn to the table of random digits (Table 4 in the Appendix of the main text). Start at the top left. Reading across, record for each of the first 11 sets of 10 consecutive digits (1) the second digit, X_i, and (2) the mean of the 10 digits, Y_i.

6.22 Derive the theoretical mean and variance of the distribution of second digits.

6.23 How do the sample properties of the X_i compare with the results of Problem 6.22?

6.24 From your answer to Problem 6.22, what should the mean and variance of the Y_i be?

6.25 How do the actual Y_i's compare with the results of Problem 6.24?

6.26 Do the actual Y_i's relate to the central-limit theorem in any way? Elaborate.

Bacteriology
Suppose a group of mice are inoculated with a uniform dose of a specific type of bacteria and all die within 24 days, with the distribution of survival times given in Table 6.4.

Table 6.4 Survival time of mice after inoculation with a specific type of bacteria

Survival time (days)	Number of mice
10	5
11	11
12	29
13	30
14	40
15	51
16	71
17	65
18	48
19	36
20	21
21	12
22	7
23	2
24	1

6.27 Assume that the underlying distribution of survival times is normal. Estimate the probability p that a mouse will survive for 20 or more days.

6.28 Suppose we are not willing to assume that the underlying distribution is normal. Estimate the probability p that a mouse will survive for 20 or more days.

6.29 Compute 95% confidence limits for the parameter estimated in Problem 6.28.

6.30 Compute 99% confidence limits for the parameter estimated in Problem 6.28.

Cardiovascular Disease
In Table 2.1 (Chapter 2, Study Guide) data on serum-cholesterol levels of 24 hospital employees before and after they adopted a vegetarian diet were provided.

6.31 What is your best estimate of the effect of adopting a vegetarian diet on change in serum-cholesterol levels?

6.32 What is the standard error of the estimate given in Problem 6.31?

6.33 Provide a 95% CI for the effect of adopting the diet.

6.34 What can you conclude from your results in Problem 6.33?

Some physicians consider only changes of at least 10 mg/dL (the same units as in Table 2.1) to be clinically significant.

6.35 Among people with a clinically significant change in either direction, what is the best estimate of the proportion of subjects with a clinically significant decline in cholesterol?

6.36 Provide a 95% CI associated with the estimate in Problem 6.35.

6.37 What can you conclude from your results in Problem 6.36?

Pulmonary Disease
Wheezing is a common respiratory symptom reported by both children and adults. A study is conducted to assess whether either personal smoking or maternal smoking are associated with wheezing in children ages 6–19.

6.38 Suppose that 400 children in this age group whose mothers smoke are assessed in 1985, and it is found that 60 report symptoms of wheezing. Provide a point estimate and a 95% CI for the underlying rate of wheezing in the population of children whose mothers smoke.

6.39 Suppose the rate of wheezing among children ages 6–19 in the general population is 10%. What can you conclude from the results in Problem 6.38?

6.40 Suppose we have a subgroup of 30 children where both the mother and the child are smokers. Six of these 30 children have wheezing symptoms. Provide a 95% CI for the true rate of wheezing in this population.

6.41 How do you assess the results in Problem 6.40?

6.42 Suppose 6 events are realized over a 1-year period for a random variable that is thought to be Poisson-distributed. Provide a 95% CI for the true expected number of events over 1 year.

SOLUTIONS

6.1 9.24

6.2 11.34

6.3 A 99% CI for the variance of total heart weight is given by

$$\left[\frac{(n-1)s^2}{\chi^2_{n-1,\ .995}},\ \frac{(n-1)s^2}{\chi^2_{n-1,\ .005}}\right]$$

We have $s^2 = 19{,}415$, $n = 11$. Therefore, the 99% CI is

$$\left[\frac{10(19{,}415)}{\chi^2_{10,\ .995}},\ \frac{10(19{,}415)}{\chi^2_{10,\ .005}}\right] = \left(\frac{194{,}150}{25.19},\ \frac{194{,}150}{2.16}\right)$$
$$= (7707,\ 89{,}884)$$

6.4 We have that $s^2 = 2{,}217.8$, $n = 10$. Thus, the 99% CI is

$$\left[\frac{9(2{,}217.8)}{\chi^2_{9,\ .995}},\ \frac{9(2{,}217.8)}{\chi^2_{9,\ .005}}\right] = \left(\frac{19{,}960}{23.59},\ \frac{19{,}960}{1.73}\right)$$
$$= (846,\ 11{,}538)$$

6.5 A 95% CI is given by

$$\hat{p} \pm 1.96\sqrt{\frac{\hat{p}\hat{q}}{n}} = .281 \pm 1.96\sqrt{\frac{.281(.719)}{153}}$$
$$= .281 \pm .071 = (.210,\ .352).$$

6.6 A 95% CI is given by

$$\hat{p} \pm 1.96\sqrt{\frac{\hat{p}\hat{q}}{n}} = .072 \pm 1.96\sqrt{\frac{.072(.928)}{153}}$$
$$= .072 \pm .041 = (.031,\ .113).$$

6.7 Since the confidence intervals do not overlap, it is clear that the true percentage of cervical-cancer cases who do not have a Pap smear is larger than the comparable percentage for controls.

6.8 We have the following data:

Patient number	$d_i(I-F)$	d_i^2
1	12.0	144
2	−18.0	324
3	50.0	2500
4	13.0	169
5	40.0	1600
6	3.0	9
7	13.0	169
8	26.0	676
9	−3.0	9
10	15.0	225
Total	151.0	5825

The summary statistics are obtained as follows:

$$\bar{d} = \frac{\sum d_i}{n} = \frac{151}{10} = 15.1$$

$$s = \sqrt{\frac{\sum(d_i - \bar{d})^2}{n-1}} = \sqrt{\frac{3544.9}{9}} = \sqrt{393.878} = 19.85$$

6.9 $se = \dfrac{s}{\sqrt{n}} = \dfrac{19.85}{\sqrt{10}} = 6.28$

6.10 A 95% CI for μ is provided by

$$\left(\bar{d} - \frac{t_{9,\ .975}s}{\sqrt{n}},\ \bar{d} + \frac{t_{9,\ .975}s}{\sqrt{n}}\right)$$
$$= [15.1 - 2.262(6.28),\ 15.1 + 2.262(6.28)]$$
$$= (15.1 - 14.2,\ 15.1 + 14.2) = (0.9,\ 29.3).$$

6.11 Since the 95% CI in Problem 6.10 does not contain 0, the results allow us to state that the true mean drop in blood pressure is significantly greater than 0. We would need a control group followed over the same time period to conclude from the sample evidence that methyl dopa is effective.

6.12 We want to construct an upper one-sided confidence interval for the new-surgical-method 3-year survival rates. Since $n\hat{p}\hat{q} = 100 \times .55 \times .45 = 24.75$, we use the normal approximation to the binomial, whereby we obtain the following upper one-sided 95% CI:

$$p > \hat{p} - z_{1-\alpha}\sqrt{\frac{\hat{p}\hat{q}}{n}} = .55 - 1.645\sqrt{\frac{.55(.45)}{100}}$$
$$= .55 - .082 = .468$$

Since this confidence interval does *not* include .40, we can say that the 3-year survival rate is improved with the new surgery.

For 5-year survival, we have $\hat{p} = .45$, $\hat{q} = .55$. Thus, the upper one-sided 95% CI is given by

$$p > \hat{p} - z_{1-\alpha}\sqrt{\frac{\hat{p}\hat{q}}{n}} = .45 - 1.645\sqrt{\frac{.45(.55)}{100}} = .368$$

Since this interval does *not* contain .33, we can say that the 5-year survival rate is also improved with the new method.

6.13 A 95% CI is given by

$$\hat{p} \pm 1.96\sqrt{\frac{\hat{p}\hat{q}}{n}} = \frac{64}{750} \pm 1.96\sqrt{\frac{(\frac{64}{750})(1-\frac{64}{750})}{750}}$$
$$= .085 \pm 1.96\sqrt{\frac{.085(.915)}{750}}$$
$$= .085 \pm .020 = (.065,\ .105)$$

6.14 The rate of .10 is compatible with these data, because it falls within the 95% CI in Problem 6.13. Thus, we *cannot* conclude from these data that jogging 10 miles per week prevents death from cardiovascular disease.

6.15 The best point estimate of the 5-year mortality rate is 3/60 = .05.

6.16 We must use the exact method for obtaining the confidence interval because

$$n\hat{p}\hat{q} = 60 \times .05 \times .95 = 2.85 < 5.$$

We refer to Table 7a (Appendix, text), $\alpha = .05$ and note the two curves with $n = 60$. Since $\hat{p} = .05 < .5$, we refer to the lower horizontal axis and draw a vertical line at $\hat{p} = .05$. We then locate where the vertical line intersects the two $n = 60$ curves and refer to the left vertical axis for the confidence limits. The confidence limits are given by (.01, .14).

6.17 Since 18% is not within the confidence interval obtained in Problem 6.16, we can exclude the possibility that the large-sample mortality rate holds for the group on a vegetarian diet and conclude that the vegetarian-diet group has a lower mortality rate.

6.18 The concern here is with the variability of two methods of measuring flow, a manual method and a digitizer method. A 95% C1 must be constructed for

$$\sigma^2_{\text{manual differences}}$$

We have

$$s_{\text{diff}} = 0.0779,\ n = 10$$

The interval is given by

$$\left(\frac{(n-1)s^2}{\chi^2_{n-1,\ .975}},\ \frac{(n-1)s^2}{\chi^2_{n-1,\ .025}}\right) = \left(\frac{9(0.0779)^2}{\chi^2_{9,\ .975}},\ \frac{9(0.0779)^2}{\chi^2_{9,\ .025}}\right)$$
$$= \left(\frac{9(0.0779)^2}{19.02},\ \frac{9(0.0779)^2}{2.70}\right)$$
$$= (0.0029,\ 0.0202)$$

So, the 95% CI for

$$\sigma^2_{\text{manual}} = (0.0029,\ 0.0202)$$

and thus the 95% CI for

$$\sigma_{\text{manual}} = (0.054,\ 0.142)$$

6.19 The same method can be used for the digitizer differences. We have that $s_{\text{diff}} = 0.0288$. Thus, the 95% CI is given by

$$\left(\frac{(n-1)s^2}{\chi^2_{9,\ .975}},\ \frac{(n-1)s^2}{\chi^2_{9,\ .025}}\right) = \left(\frac{9(0.0288)^2}{19.02},\ \frac{9(0.0288)^2}{2.70}\right)$$
$$= (0.00039,\ 0.0276)$$

Finally, the 95% CI for

$$\sigma_{\text{dig}} = (0.020,\ 0.053)$$

6.20 $d_i = |\text{manual diff}_i| - |\text{digitizer diff}_i|$. A 95% CI for μ_d comes from the *t* statistic approach:

$$\bar{d} \pm t_{9,\ .975}\left(\frac{s_d}{\sqrt{10}}\right) = 0.0440 \pm 2.262\left(\frac{0.0353}{\sqrt{10}}\right)$$
$$= (0.0187,\ 0.0693)$$

6.21 The digitizer method appears to be significantly less variable than the manual method, as shown in Problem 6.20, since both ends of the 95% CI for μ_d are above 0.

6.22 $Pr(X_i = j) = .1$, $j = 0,\ 1,\ \ldots,\ 9$. Therefore,

$$\mu = E(X) = \frac{0 + 1 + \cdots + 9}{10} = 4.5$$
$$\sigma^2 = E(X-\mu)^2 = E(X^2) - \mu^2$$
$$= \frac{0^2 + 1^2 + \cdots + 9^2}{10} - 4.5^2 = 8.25$$

6.23 Since $\bar{X}$ and S^2 are unbiased estimators of μ and σ^2, respectively, the calculated mean and variance of the 11 sample X_i's should approximate the theoretical values in Problem 6.22. Indeed, we have that $X_i = 2$, 8, 4, 4, 8, 7, 1, 1, 0, 7, 0. Thus, $\bar{x} = 3.82 \cong \mu$, $s_x^2 = 10.36 \cong \sigma^2$.

6.24 We have

$$Y_i = \frac{\sum_{j=1}^{10} X_j}{10}$$
$$E(Y_i) = \frac{\sum E(X_j)}{10} = E(X_j) = 4.5$$
$$\text{Var}(Y_i) = \frac{\text{Var}(X_j)}{10} = \frac{8.25}{10} = 0.825$$

6.25 The calculated mean and variance of the actual Y_i should approximate the theoretical values in Problem

6.24. Since $\text{Var}(Y_i) < \text{Var}(X_i)$, we would expect the Y_i to be more closely clustered than the X_i about the common mean of 4.5. Indeed, we have that $Y_i = 3.3$, 4.0, 4.3, 5.0, 4.0, 4.9, 3.4, 3.2, 6.2, 6.0, 4.0. Furthermore, $\bar{y} = 4.39 \cong \mu$, $s_y^2 = 1.05 \cong \sigma_y^2$.

6.26 Since the Y_i's are means of random samples selected from a population with

$$\text{mean} = \mu = 4.5 \text{ and variance} = \sigma^2 = 8.25,$$

the central-limit theorem implies that their distribution will be approximately normal, with mean $\mu_y = 4.5$ and variance $\sigma_y^2 = 8.25/10 = 0.825$. The number of samples of size 10(11) is too small to confirm whether the central-limit theorem is applicable to samples of size 10 drawn from this population of random digits.

6.27 We first estimate the parameters μ and σ^2 of the normal distribution. We use $\bar{x}$ as an estimate of μ and s^2 as an estimate of σ^2. We have $\bar{x} = 16.13$, $s^2 = 7.07$. If we assume that the underlying distribution of survival times is $N(16.13, 7.07)$, then an estimate of the probability that a mouse will survive for 20 or more days is given by

$$\begin{aligned} Pr(X \geq 19.5) &= 1 - \Phi\left(\frac{19.5 - 16.13}{\sqrt{7.07}}\right) \\ &= 1 - \Phi(1.27) = 1 - .8977 = .102 \end{aligned}$$

6.28 We can treat this random variable as binomial; that is,

$X_i = 1$ if mouse i survives for 20 or more days
$= 0$ otherwise

where

$X_i = 1$ with probability p
$= 0$ with probability $1 - p$

Then an estimate of p is given by the sample proportion of successes, namely

$$\hat{p} = \frac{21 + 12 + 7 + 2 + 1}{5 + 11 + \cdots + 1} = \frac{43}{429} = .100$$

6.29 95% CI's for p are given by

$$\begin{aligned} \hat{p} \pm 1.96\sqrt{\frac{\hat{p}\hat{q}}{n}} &= .100 \pm 1.96\sqrt{\frac{.100(.900)}{429}} \\ &= .100 \pm .028 = (.072,\ .129) \end{aligned}$$

6.30 99% CI's for p are given by

$$\begin{aligned} \hat{p} \pm 2.576\sqrt{\frac{\hat{p}\hat{q}}{n}} &= .100 \pm 2.576\sqrt{\frac{.100(.900)}{429}} \\ &= .100 \pm .037 = (.063,\ .138) \end{aligned}$$

6.31 We compute

$$\bar{d} = \frac{\sum d_i}{n} = \frac{469}{24} = 19.54 \text{ mg/dL},$$

which is our best estimate of the mean decline in serum-cholesterol levels.

6.32 $sem = s/\sqrt{n} = 16.81/\sqrt{24} = 3.43$

6.33 A 95% CI for μ is provided by

$$\begin{aligned} \bar{d} \pm t_{n-1,\ .975}\frac{s}{\sqrt{n}} &= 19.54 \pm t_{23,\ .975}(3.43) \\ &= 19.54 \pm 2.069(3.43) \\ &= 19.54 \pm 7.10 = (12.4,\ 26.6) \end{aligned}$$

6.34 We can conclude that cholesterol levels have declined after adopting the diet because 0 is not within the 95% CI in Problem 6.33.

6.35 Among people with an absolute change of at least 10 mg/dL in either direction, 18 have declined, while 2 have increased. Thus, the proportion who have declined = 18/20 = .90. Therefore, 90% of people with a clinically significant change have declined.

6.36 Since $n\hat{p}\hat{q} = 20 \times .9 \times .1 = 1.8 < 5$, we must use the exact method. We refer to Table 7a (Appendix, text), $\alpha = .05$ and locate .9 on the upper horizontal axis. We then draw a vertical line and find the point of intersection with the n = 20 curves. Referring to the right vertical axis, we find that $c_1 = .68$, $c_2 = .99$. Thus, the 95% CI $= (.68,\ .99)$.

6.37 Since 50% is not within the 95% CI in Problem 6.36, we can conclude that more than 50% of people with a clinically significant change have declined. Thus, among people with a clinically significant change, a decline in cholesterol levels is more likely than an increase.

6.38 The best point estimate is $\hat{p} = 60/400 = .15$. Since $n\hat{p}\hat{q} = 400(.15)(.85) = 51.0 \geq 5$, we can use the normal approximation method to obtain an interval estimate. The estimate is given by

$$\begin{aligned} \hat{p} \pm 1.96\sqrt{\frac{\hat{p}\hat{q}}{n}} &= .15 \pm 1.96\sqrt{\frac{.15(.85)}{400}} \\ &= .15 \pm .035 = (.115,\ .185) \end{aligned}$$

6.39 Since 10% is not included in the preceding 95% CI, we can conclude that the true rate of wheezing among children whose mothers smoke is significantly higher than the rate of wheezing in the general population.

6.40 The best point estimate is $\hat{p} = 6/30 = .20$. Since $n\hat{p}\hat{q} = 30(.20)(.80) = 4.8 < 5$, we must use the exact binomial confidence limits. We refer to Table 7a (Appendix, text) for $n = 30$, $\hat{p} = .20$ and obtain the 95% CI $= (.08, .38)$.

6.41 Since the 95% CI in Problem 6.40 includes 10%, we cannot be certain that the true rate of wheezing is higher among smoking children whose mothers smoke than in the general population. We would need a larger population of smoking children whose mothers also smoke to make a more definitive assessment.

6.42 Referring to Table 8 (Appendix, text) with $x = 6$, we obtain the 95% CI $= (2.20, 13.06)$.

REFERENCE

[1] Celentano, D. D., Klassan, A. C., Weisman, C. S., & Rosenshein, N. S. (1988). Cervical-cancer screening practices among older women: Results from the Maryland cervical-cancer case-control study. *Journal of Clinical Epidemiology*, **41**(6), 531–541.