Suggested problems ch6

The data in Table 6.10 concern the mean triceps skin-fold thickness in a group of normal men and a group of men with chronic airflow limitation [5].

TABLE 6.10 **Triceps skin-fold thickness in normal men and men with chronic airflow limitation**

Group	Mean	*sd*	*n*
Normal	1.35	0.5	40
Chronic airflow limitation	0.92	0.4	32

Source: Adapted from *Chest*, *85*(6), 58S–59S, 1984.

***6.5** What is the standard error of the mean for each group?

6.6 Assume that the central-limit theorem is applicable. What does it mean in this context?

6.7 Find the upper 1st percentile of a *t* distribution with 16 *df*.

6.8 Find the lower 10th percentile of a *t* distribution with 28 *df*.

6.9 Find the upper 2.5th percentile of a *t* distribution with 7 *df*.

6.10 What are the upper and lower 2.5th percentiles for a chi-square distribution with 2 *df*? What notation is used to denote these percentiles?

Refer to the data in Table 2.13. Regard this hospital as typical of Pennsylvania hospitals.

6.11 Compute a 95% CI for the mean age.

6.12 Compute a 95% CI for the mean white blood count following admission.

6.13 Answer Problem 6.12 for a 90% CI.

6.14 What is the relationship between your answers to Problems 6.12 and 6.13?

***6.15** What is the best point estimate of the percentage of males among patients discharged from Pennsylvania hospitals?

***6.16** What is the standard error of the estimate obtained in Problem 6.15?

***6.17** Provide a 95% CI for the percentage of males among patients discharged from Pennsylvania hospitals.

6.5 sem $= \frac{0.5}{\sqrt{40}} = 0.079$ for normal men and $\frac{0.4}{\sqrt{32}} = 0.071$ for men with chronic airflow limitation.

6.6 It means that the distribution of mean triceps skin-fold thickness from repeated samples of size 40 drawn from the population of normal men can be considered to be normal with mean μ and variance $\frac{\sigma^2}{n} \cong \frac{s^2}{n} = \frac{0.5^2}{40} = 0.0063$. A similar statement holds for men with chronic airflow limitation.

6.7 2.583

6.8 −1.313

6.9 2.365

6.10 We refer to Table 6. The lower 2.5th percentile is 0.0506 and is denoted by $\chi^2_{2,.025.}$ The upper 2.5th percentile is 7.38 and is denoted by $\chi^2_{2,.975}$.

6.11 We have that $\bar{x} = \frac{1031}{25} = 41.24$ years. Therefore, a 95% confidence interval for μ is given by

$$\bar{x} \pm \frac{t_{24,.975}s}{\sqrt{n}} = 41.24 \pm \frac{2.064(20.10)}{\sqrt{25}}$$
$$= 41.24 \pm 8.30 = (32.94, 49.54)$$

6.12 The 95% confidence interval is computed from $\bar{x} \pm t_{n-1,.975}\frac{s}{\sqrt{n}}$. We have that $\bar{x} = 7.84$, $s = 3.21$. Therefore, we have the following 95% confidence interval

$$7.84 \pm t_{24,.975} \times \frac{3.21}{\sqrt{25}} = 7.84 \pm 2.064 \times \frac{3.21}{5}$$
$$= 7.84 \pm 1.33 = (6.51,\ 9.17)$$

6.13 A 90% confidence interval is given by

$$\bar{x} \pm t_{n-1,.95} \frac{s}{\sqrt{n}} = 7.84 \pm t_{24,.95} \times \frac{3.21}{\sqrt{25}}$$
$$= 7.84 \pm 1.711 \times \frac{3.21}{5}$$
$$= 7.84 \pm 1.10 = (6.74,\ 8.94)$$

6.14 The 90% confidence interval should be shorter than the 95% confidence interval, since we are requiring less confidence. This is indeed the case.

6.15 Our best estimate is given by $\hat{p} = \frac{11}{25} = .44$.

6.16 The standard error = $\sqrt{\frac{\hat{p}\hat{q}}{n}} = \sqrt{\frac{.44 \times .56}{25}} = .099$.

6.17 Since $n\hat{p}\hat{q} = 25 \times .44 \times .56 = 6.16 \ge 5$, we can use the normal theory method. Therefore, a 95% confidence interval for the percentage of males discharged from Pennsylvania hospitals is given by

$$\hat{p} \pm 1.96\sqrt{\frac{\hat{p}\hat{q}}{n}} = .44 \pm 1.96(.099)$$
$$= .44 \pm .195$$
$$= (.25,\ .63)$$

Suppose a clinical trial is conducted to test the efficacy of a new drug, spectinomycin, for treating gonorrhea in females. Forty-six patients are given a 4-g daily dose of the drug and are seen 1 week later, at which time 6 of the patients still have gonorrhea.

***6.27** What is the best point estimate for p, the probability of a failure with the drug?

***6.28** What is a 95% CI for p?

***6.29** Suppose we know penicillin G at a daily dose of 4.8 megaunits has a 10% failure rate. What can be said in comparing the two drugs?

6.27 The best point estimate is $\hat{p} = \frac{6}{46} = .130$.

6.28 Since $n\hat{p}\hat{q} = 46(6/46)(40/46) = 5.2 \geq 5$, we can use the normal approximation to the binomial distribution. If a normal approximation is used, then the lower confidence limit is

$$c_1 = \hat{p} - 1.96\sqrt{\frac{\hat{p}\hat{q}}{n}}$$

$$= .130 - 1.96\sqrt{\frac{.130(.870)}{46}} = .033$$

The upper confidence limit is

$$c_2 = \hat{p} + 1.96\sqrt{\frac{\hat{p}\hat{q}}{n}}$$

$$= .130 + 1.96\sqrt{\frac{.130(.870)}{46}} = .228$$

6.29 Since 10% is within the 95% confidence interval, we would conclude that it is possible that the two drugs are equally effective (i.e., have the same failure rate), or at least are not "significantly different".

Suppose we want to estimate the concentration (μg/mL) of a specific dose of ampicillin in the urine after various periods of time. We recruit 25 volunteers who have received ampicillin and find they have a mean concentration of 7.0 μg/mL with a standard deviation of 2.0 μg/mL. Assume the underlying population distribution of concentrations is normally distributed.

***6.30** Find a 95% CI for the population mean concentration.

***6.31** Find a 99% CI for the population variance of the concentrations.

***6.32** How large a sample would be needed to ensure that the length of the CI in Problem 6.30 is 0.5 μg/mL assuming the sample standard deviation remains at 2.0 μg/mL?

6.30 We assume that $x_1, \ldots, x_{25} \sim N(\mu, \sigma^2)$, where μ, σ^2 are unknown, and find that $\bar{x} = 7.0$, $s^2 = 4.0$. Thus, a two-sided 95% confidence interval for the mean is given by

$$\left(\bar{x} - t_{24,.975}\frac{s}{\sqrt{n}}, \bar{x} + t_{24,.975}\frac{s}{\sqrt{n}}\right) = \left[7.0 - \frac{2.064(2)}{5}, 7.0 + \frac{2.064(2)}{5}\right]$$

$$= (6.17, 7.83)$$

6.31 A two-sided 99% confidence interval for the unknown variance σ^2 is given by

$$\left(\frac{(n-1)s^2}{\chi^2_{24,.995}}, \frac{(n-1)s^2}{\chi^2_{24,.005}}\right) = \left(\frac{24(4)}{45.56}, \frac{24(4)}{9.89}\right)$$

$$= (2.11, 9.71)$$

6.32 The length of the 95% confidence interval in Problem 6.41 is given by

$$2t_{n-1,.975}\frac{s}{\sqrt{n}}$$

Figure 6.4b (p. 172) plotted the sampling distribution of the mean from 200 samples of size 5 from the population of 1000 birthweights given in Table 6.2. The mean of the 1000 birthweights in Table 6.2 is 112.0 oz with standard deviation 20.6 oz.

***6.52** If the central-limit theorem holds, what proportion of the sample means should fall within 0.5 lb of the population mean (112.0 oz)?

***6.53** Answer Problem 6.52 for 1 lb rather than 0.5 lb.

***6.54** Compare your results in Problems 6.52 and 6.53 with the actual proportion of sample means that fall in these ranges.

***6.55** Do you feel the central-limit theorem is applicable for samples of size 5 from this population? Explain.

five sample points.

6.52 If the central-limit theorem holds, then $\bar{X} \sim N\left(\mu, \frac{\sigma^2}{n}\right) = N\left(112, \frac{20.6^2}{5}\right) = N(112, 84.87)$. Therefore, it follows that

$$
\begin{aligned}
Pr(104 < \bar{X} < 120) &= \Phi\left(\frac{120-112}{\sqrt{84.87}}\right) - \Phi\left(\frac{104-112}{\sqrt{84.87}}\right) \\
&= \Phi(0.87) - \Phi(-0.87) \\
&= \Phi(0.87) - [1 - \Phi(0.87)] \\
&= 2\Phi(0.87) - 1 \\
&= 2(.8074) - 1 = .615
\end{aligned}
$$

6.53 We have

$$
\begin{aligned}
Pr(96 < \bar{X} < 128) &= \Phi\left(\frac{128-112}{\sqrt{84.87}}\right) - \Phi\left(\frac{96-112}{\sqrt{84.87}}\right) \\
&= \Phi(1.74) - \Phi(-1.74) \\
&= 2\Phi(1.74) - 1 \\
&= 2(.9588) - 1 = .918
\end{aligned}
$$

6.54 The percentage of points in the 104–120 range (i.e., corresponding to the 104, 106, . . . , 118 base) $= 5\% + 8\% + 12.5\% + 8.5\% + 6.5\% + 5\% + 7\% + 9\% = 61.5\%$. Similarly, the bars in the 96–128 range comprise 93.5% of the points. These observed proportions compare very well with the theoretical proportions in Problems 6.52 and 6.53.

6.55 The central-limit theorem seems to be applicable here since the agreement in Problem 6.54 between observed and expected proportions is excellent.