Answers to Selected Exercises

CHAPTER 1

1.1. A p-value is the probability that the value of the test statistic is more extreme than the value that was computed if the null hypothesis were true.

1.2. A very small difference in two population means is generally not going to be of any practical interest.

1.6. The statement is misleading because the significance level is the probability of rejecting the null hypothesis if the null hypothesis were true.

CHAPTER 2

2.5. Using Eq. (2.5), we obtain Power $= 1 - \Phi\ [(Z_\alpha - \sqrt{n}(|\mu - \mu_0|)/\sigma)] = 1 - \Phi\ [(1.645 - \sqrt{100}(5)/10)]$ since σ would be estimated as $(70 - 10)/6 = 10$. Thus, Power $= 1 - \Phi\ [(1.645 - \sqrt{100}(5)/10)] = 1 - \Phi\ [-3.355] = 1 - .00039 = .9996$. The experimenter should not have arbitrarily selected the number of observations because the power is far higher than necessary or desirable. Next time the experimenter should use software to determine the sample size to use or, if preferred, perform the computation by hand.

2.8. No, because sigma might have been overestimated by at least enough to make up the difference.

Sample Size Determination and Power, First Edition. Thomas P. Ryan.

CHAPTER 3

3.1. Using the PASS software, the required minimum sample size is 56. If $n = 100$ is used, the power is .95434. Since such a power value would generally be considered excessive, this illustrates why sample size should not be arbitrarily chosen.

3.2. It is never logical or practical to use hand computation when iteration is required, as is the case if Eq. (3.4) were used to solve for a sample size. So the use of software is imperative.

3.8. The power will probably be less because it is more difficult to detect a small difference than a large difference. A definite answer cannot be given, however, because the standard deviations are almost certainly unknown and will not be estimated precisely, so it depends on the magnitudes and direction of the misestimation.

CHAPTER 4

4.1. Using

$$n = \left[\frac{z_\alpha\sqrt{(p_1 + p_2)(1 - p_1 + 1 - p_2)/2} + z_\beta\sqrt{p_1(1 - p_1) + p_2(1 - p_2)}}{p_1 - p_2}\right]^2$$

we have

$$n = \left[\frac{1.645\sqrt{(.55 + .65)(1 - .55 + 1 - .65)/2} + 0.84\sqrt{.55(1 - .55) + .65(1 - .65)}}{.55 - .65}\right]^2$$

so that $n = 295.365$ and $n = 296$ would thus be used. It is desirable to see how sample sizes are obtained, but once that understanding is achieved, it is preferable to use software to determine sample size when the necessary hand calculation would be somewhat mathematically laborious.

4.5. Using the PASS software, we obtain $n = 35$. Of course, the use of software is almost imperative, because there is no equation for the sample size, as stated in Section 4.5.

CHAPTER 5

5.2. Using the Demidenko applet, a sample size of $n = 275$ is obtained.

5.3. Since the minimum acceptable R^2 value was stated, let's start with that and the relationship between R^2 and the calculated t-statistic for testing the null hypothesis that $\beta_1 = 0$, with that relationship being $R^2 = t^2/(n-2+t^2)$. We know that we need to do more than just barely reject the null hypothesis in order to have a good model, so let's combine this expression with the Wetz criterion. If the latter is just barely met, then $R^2 = 4t^2_{\alpha/2,n-2}/(n-2+4t^2_{\alpha/2,n-2})$. Substituting .75 for R^2 and solving for n gives $n = \frac{4}{3}\,t^2_{\alpha/2,n-2} + 2$. Unfortunately, this doesn't give us a simple solution since n is on both sides of the equation. We can, however, write a small program to produce the value of each side of the equation for a range of n values and see where approximate equality occurs. This will give us, at the very least, a starting point for determining n. Doing so and using $\alpha = .05$ produces the following results.

n	$\frac{4}{3}t^2_{.025,n-2}+2$
5	15.5040
6	12.2782
7	10.8105
8	9.9832
9	9.4553
10	9.0902
11	8.8231
12	8.6195

This leads us to zero in on $n = 9$ or 10 because if a fractional value for n were possible, equality between the left side and the right side would occur between 9 and 10. If we used $n = 10$ and the Wetz criterion is just barely met, then $R^2 = .73$. Similarly, $n = 9$ gives $R^2 = .76$. Thus, $n = 9$ or 10 is a reasonable choice for this problem, in general accordance with the Draper and Smith rule-of-thumb and in recognition of the cost of sampling. Of course, this says nothing about whether the Wetz criterion will actually be met, as that will depend on the relationship between X and Y.

5.4. The use of a sample size of 10 may be acceptable, depending on the expected variability of the data, but the use of retrospective power is not acceptable.

CHAPTER 6

6.4. This would make it impossible to determine the sample size. If MINITAB is used, for example, the user must enter the "values of the maximum difference between means." If Power and Precision were used, the user would enter the "effect size *f*." The user of nQuery Advisor must enter either the effect size or both the variance of the means and the common standard deviation. And so on.

6.12. **(a)** The required sample sizes are 97, 81, and 50 for $a = 19$, 20, and 22, respectively.

(b) The required sample sizes are 693, 709, and 740 for $a = -19$, -20, and -22, respectively.

CHAPTER 7

7.1. Using PASS 11, the required sample size for the clinical trial is 55. The null hypothesis will be rejected if the number of responses is at least 10.

7.2. There should be concern because some survey articles have shown that a very high percentage of clinical trials have been underpowered. This should be alarming since human subjects are involved in the experimentation.

CHAPTER 8

8.1. These values are obtained by first computing

$$\begin{aligned} Z &= \left(\mu + 3\sigma/\sqrt{n} - (\mu + a\sigma)\right) / \left(\sigma/\sqrt{n}\right) \\ &= 3 - a\sqrt{n} \end{aligned}$$

with $a = 1$. Thus, for $n = 4$, $Z = 1$ and $P(Z > 1) = .15866$, as can be obtained using statistical software. The other values in Table 8.1 would be obtained by substituting each of the other values of n in the expression for Z.

CHAPTER 9

9.1. PASS gives the sample size for each group as 328 and the power as .8001. The number of events is given as 214 for the first group and 213 for the second group. This is important information as these numbers of events must occur in order for the desired power to be attained.

CHAPTER 10

10.2. In PASS, for example, the routine for a test of one mean with a Wilcoxon test is the same routine as is used when normality is assumed. The difference is that for the Wilcoxon test the user specifies a "nonparametric adjustment" and selects either the uniform, double exponential, or logistic distribution. This is necessary for determining the power of the test. If the true distribution were known, however, then there would not be a need to use a nonparametric test, as a user could work directly with that distribution, as indicated in Section 10.1. Thus, determining sample size for nonparametric tests can very easily be a less-than-satisfying exercise.