Exercise 10.8

A two-stage clinical trial is planned for testing $H_0 : p = .10$ versus $H_a : p > .10$, where p is the proportion of responders among patients who were treated by the protocol treatment. At the first stage, 15 patients are accrued and treated. If 4 or more responders are observed among the (first) 15 patients, H_0 is rejected, the study is terminated, and no more patients are accrued. Otherwise, another 15 patients will be accrued and treated in the second stage. If a total of 6 or more responders are observed among the 30 patients accrued in the two stages (15 in the first stage and 15 more in the second stage), then H_0 is rejected. For example, if 5 responders are found among the first-stage patients, H_0 is rejected and the study is over. However, if 2 responders are found among the first-stage patients, 15 second-stage patients are accrued, and an additional 4 or more responders (for a total of 6 or more among the 30) are identified, H_0 is rejected and the study is over.

- **a** Use the binomial table to find the numerical value of α for this testing procedure.
- **b** Use the binomial table to find the probability of rejecting the null hypothesis when using this rejection region if $p = .30$.
- **c** For the rejection region defined above, find β if $p = .30$.

Exercise 10.17

A survey published in the *American Journal of Sports Medicine* reported the number of meters (m) per week swum by two groups of swimmers-those who competed exclusively in breaststroke and those who competed in the individual medley (which includes breaststroke). The number of meters per week practicing the breaststroke was recorded for each swimmer, and the summary statistics are given below. Is there sufficient evidence to indicate that the average number of meters per week spent practicing breaststroke is greater for exclusive breaststrokers than it is for those swimming individual medley?

- **a** State the null and alternative hypotheses.

b What is the appropriate rejection region for an $\alpha = .01$ level test?

c Calculate the observed value of the appropriate test statistic.

d What is your conclusion?

e What is a practical reason for the conclusion you reached in part (**d**)?

	Specialty	
	Exclusively Breaststroke	Individual Medley
Sample size	130	80
Sample mean (m)	9017	5853
Sample standard deviation (m)	7162	1961
Population mean	μ_1	μ_2

Exercise 10.27

The state of California is working very hard to ensure that all elementary age students whose native language is not English become proficient in English by the sixth grade. Their progress is monitored each year using the California English Language Development test. The results for two school districts in southern California for the 2003 school year are given in the accompanying table. Do the data indicate a significant difference in the 2003 proportions of students who are fluent in English for the two districts? Use $\alpha = .01$.

District	Riverside	Palm Springs
Number of students tested	6124	5512
Percentage fluent	40	37

Exercise 10.8

Let Y_1 and Y_2 have binomial distributions with parameters $n = 15$ and p.

a

$$\alpha = P(\text{ reject } H_0 \text{ in stage } 1 \mid H_0 \text{ true }) + \text{P}(\text{ reject } H_0 \text{ in stage } 2 \mid H_0 \text{ true })$$

$$= P(Y_1 \geq 4) + P(Y_1 + Y_2 \geq 6, Y_1 \leq 3) = P(Y_1 \geq 4) + \sum_{i=0}^{3} P(Y_1 + Y_2 \geq 6, Y_1 \leq i)$$

$$= P(Y_1 \geq 4) + \sum_{i=0}^{3} P(Y_2 \geq 6 - i) P(Y_1 \leq i) = .0989(\text{ calculated with } p = .10).$$

b Similar to part a with $p = .3 : \alpha = .9321$.

c $\beta = P(\textit{fail to reject } H_0 \mid p = .3) = \sum_{i=0}^{3} P(Y_1 = i, Y_1 + Y_2 \leq 5) = \sum_{i=0}^{3} P(Y_2 = 5 - i) P(Y_1 = i) = .0679$.

Exercise 10.17

a $H_0 : \mu_1 = \mu_2, H_a : \mu_1 > \mu_2$.

b Reject if $Z > 2.326$, where Z is given in Example 10.7 $(D_0 = 0)$.

c $Z = \dfrac{\bar{Y_1} - \bar{Y_2}}{\sqrt{\frac{\sigma_1^2}{n_1} + \frac{\sigma_2^2}{n_2}}} \approx 4.76$.

d Reject H_0.

Exercise 10.27

Define: p_1 = proportion of English-fluent Riverside students, and p_2 = proportion of English-fluent Palm Springs students. To test $H_0 : p_1 - p_2 = 0$, versus $H_a : p_1 - p_2 \neq 0$, we can use the large-sample test statistic

$$Z = \frac{\hat{p}_1 - \hat{p}_2}{\sqrt{\frac{p_1 q_1}{n_1} + \frac{p_2 q_2}{n_2}}}$$

However, this depends on the (unknown) values p_1 and p_2. Under $H_0, p_1 = p_2 = p$ (i.e. they are samples from the same binomial distribution), so we can "pool" the samples to estimate p:

$$\hat{p}_p = \frac{n_1 \hat{p}_1 + n_2 \hat{p}_2}{n_1 + n_2} = \frac{Y_1 + Y_2}{n_1 + n_2}$$

So, the test statistic becomes

$$Z = \frac{\hat{p}_1 - \hat{p}_2}{\sqrt{\hat{p}_p \hat{q}_p \left(\frac{1}{n_1} + \frac{1}{n_2}\right)}} \approx 3.32.$$