Sign test for matched pairs

Example 6

Time Browsing the Internet or Watching TV

Picture the Scenario

Which do most students spend more time doing—browsing the Internet or watching TV? Let's consider the students surveyed at the University of Georgia whose responses are in the Georgia Student Survey data file. The results for the first three students in the data file (in minutes per day) were

Student	Internet	TV
1	60	120
2	20	120
3	60	90

All three spent more time watching TV. For the entire sample, 35 students spent more time watching TV and 19 students spent more time browsing the Internet. (The analysis ignores the 5 students who reported the same time for each.)

Question to Explore

Let p denote the population proportion who spent more time watching TV. Find the test statistic and P-value for the sign test of H_0: $p = 0.50$ against H_a: $p \neq 0.50$. Interpret.

Think It Through

Here, $n = 35 + 19 = 54$. The sample proportion who spent more time watching TV was $35/54 = 0.648$. For testing that $p = 0.50$, the *se* of the sample proportion is

$$se = \sqrt{(0.50)(0.50)/n} = \sqrt{(0.50)(0.50)/54} = 0.068.$$

The test statistic is

$$z = (\hat{p} - 0.50)/se = (0.648 - 0.50)/0.068 = 2.18.$$

From the normal distribution table (Table A or software), the two-sided P-value is 0.03. This provides considerable evidence that most students spend more time watching TV than browsing the Internet. The conclusion must be tempered by the fact that the data resulted from a convenience sample (students in a class for a statistics course) rather than a random sample of all college students.

Insight

The sign test uses merely the information about *which* response is higher and *how many* responses are higher, not the quantitative information about *how much* higher. This is a disadvantage compared to the corresponding parametric test, the matched pairs t test of Section 10.4, which analyzes the mean of the differences between the two responses. The sign test is most appropriate when we can order the responses but do not have quantitative information, such as in the next example.

Try Exercise 15.10