Example 1 Shoe Print and Height: Finding a Prediction Interval

If we use the 40 paired shoe print lengths (cm) and heights (cm) from Data Set 2 in Appendix B, we find that there is sufficient evidence to support the claim of a linear correlation between those two variables. Also, the regression equation is $\hat{y} = 80.9 + 3.22x$. If we measure a suspect's shoe print length and obtain a value of 29.0 cm, we can substitute that value into the regression equation to find the predicted height of 174.3 cm. Construct a 95% prediction interval for the height, given that the shoe print length is 29.0 cm (so that $x = 29.0$ cm).

Solution

The accompanying Minitab display automatically provides the 95% prediction interval, which is 162 cm $< y <$ 186 cm when rounded. The same prediction interval could be manually calculated using these components:

$x_0 = 29.0$ (given)

$s_e = 5.943761$ (provided by many technologies, including STATDISK, Minitab, Excel, and TI-83/84 Plus)

$\hat{y} = 174.3$ (predicted value of y found by substituting $x = 29.0$ into the regression equation)

$t_{\alpha/2} = 2.024$ (from Table A-3 with $df = 38$ and an area of 0.05 in two tails)

$n = 40$ $\quad \bar{x} = 29.0175$ $\quad \Sigma x = 1160.7$ $\quad \Sigma x^2 = 33{,}933.17$

Interpretation

The 95% prediction interval is 162 cm $< y <$ 186 cm. This means that if we measure a shoe print and find a length of 29.0 cm, we have 95% confidence that the limits of 162 cm and 186 cm (or 64 in. and 73 in.) contain the height of the suspect. That is a large range of values, so the single shoe print clue doesn't give us very good information about the suspect's height. The prediction interval would be much narrower and our estimated height would be much better if we were using a much larger set of sample data instead of using only the 40 pairs of values listed in Data Set 2 from Appendix B.

MINITAB

```
Predicted Values for New Observations

New Obs      Fit   SE Fit         95% CI              95% PI
      1  174.269    0.940  (172.366, 176.171)  (162.087, 186.451)

Values of Predictors for New Observations

          Shoe
New Obs  Print
      1   29.0
```

Football for Predicting the Stock Market and the Presidency

According to the "Super Bowl omen," the New York Stock Exchange index will rise in a year that the Super Bowl is won by a team with NFL origins; otherwise, it will fall. (The old NFL and AFL merged into the current NFL in 1970.) After the first 29 Super Bowl games, the prediction was correct 90% of the time, but it has been less successful in recent years. As of this writing, it has been correct in 80% of Super Bowl games. Forecasting and predicting are important goals of statistics, but common sense suggests that it would be foolish to base investments on the outcome of a football game. Other somewhat exotic indicators used to forecast stock market performance include rising skirt hemlines (rising hemlines followed by rising market), aspirin sales (higher sales followed by lower market), limousines on Wall Street, orders for cardboard boxes, sales of beer versus wine, and elevator traffic at the New York Stock Exchange.

The "Redskins Rule" is that in years of presidential elections, the candidate from the incumbent party will win if the Redskins win their last home game. If the Redskins lose their last home game, the candidate from the incumbent party will lose. As of this writing, the Redskins Rule has been correct in 17 of the 18 presidential elections that were held after the Redskins moved to Washington.