1

Worked Examples for Smoothers

These examples use the ASA software integrated into Excel or SPSS (www.asastat.com). ASA is, in part, a point-and-click interface to R but analyses can be conducted from within SPSS or Excel. All data in the examples are hypothetical. We assume you have read the primer on smoothers.

The first example predicts student test performance on an exam from a measure of test anxiety. The outcome is measured on a 0 to 100 metric, with higher scores indicating better performance. Test anxiety is a multi-item scale that ranges from -5 to +5, where 0 is a normed average or typical score from a broader student population. Scores greater than 0 reflect increasingly above average test anxiety and scores less than 0 reflect increasingly below average test anxiety. The program is "Lowess smoothing" in the folder "Robust Statistics > Smoothers and Exploratory Curve Analysis > Single Predictor" of the ASA software. It interfaces with the R programs by Rand Wilcox at the University of Southern California. The value for the span is 0.75. For the execution of the program, watch the video on our website. Figure 1.1 presents the smooth:

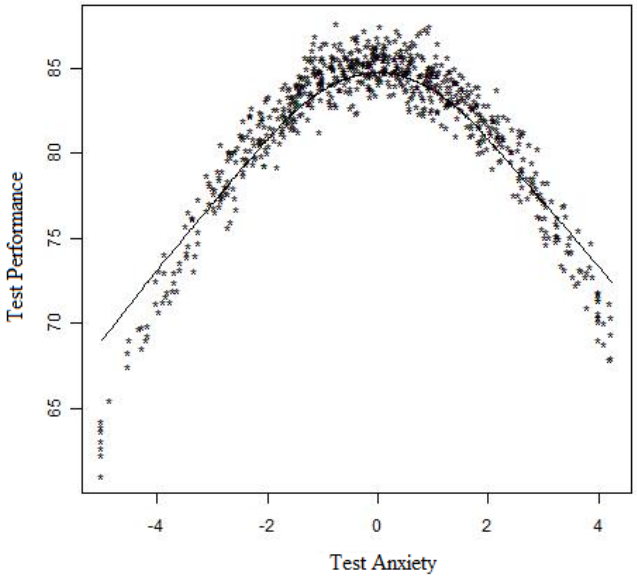

FIGURE 1.1. Smoother for Test Performance and Test Anxiety

The relationship is decidedly non-linear. At the low end of the test anxiety dimension, increases in anxiety are associated with increases in test performance, probably by motivating students to study more. At some point (near a score of 0), however, test anxiety begins to interfere with test performance such that increases in anxiety are associated with decreases in performance. To analyze these data more formally, one would need to use some form of non-linear regression, such as regression with polynomials or spline regression.

Although ASA is point-and-click, it makes use of the R function lplot in the WRS2 package from R. Wilcox's program generates a statistic analogous to a squared R in multiple regression called "Explanatory Power." Here is the output:

```
EXPLANATORY POWER OF SMOOTH

   Explanatory power: 0.9149
```

Test anxiety accounts for about 91% of the variation in test performance. For technical details of this index, see Wilcox (2017).

In addition, the output provides a predicted conditional mean for each value of the predictor observed in the data. Here are some examples showing the first 10 cases:

```
SAMPLE Xs AND PREDICTED PARAMETER VALUES

Param is the predicted mean conditional on the predictor value
Value of -9999 indicates parameter could not be estimated

Case    Param          X1
1      83.9746         -0.862
2      83.0042         1.314
3      84.101          -0.79
4      72.5274         4.2
5      83.3736         -1.153
6      83.7117         -0.995
7      84.6661         -0.229
8      79.7973         2.274
9      77.928          -2.768
10     80.2316         2.156
```

For example, for people who have a test anxiety score of 1.314, the predicted mean test performance is 83.0042. For people who have a test anxiety score of -1.153, the predicted mean test performance is 83.3736.

Smoothers are rarely published in articles. Instead, they tend to be used in preliminary analyses when exploring data. However, if we were to write-up the test anxiety and test performance example, the write-up might appear as follows:

"A LOWESS (locally weighted scatterplot smoothing) smoother was applied to the data using a span of 0.75, the results of which are shown in Figure 1. The relationship is decidedly non-linear. At the low end of the test anxiety dimension, increases in anxiety are associated with increases in test performance, probably by motivating students to study more. At some point (near a score of 0), however, test anxiety begins to interfere with test performance such that increases in anxiety are associated with decreases in performance. Test anxiety accounts for about 91% of the variation in test performance (Wilcox, 2017). As examples of the curve, for people who have a low test anxiety score of -2, the predicted mean test performance is 80.5; for people who have a high test anxiety score of 2, the predicted mean test performance is 80.6; for people who have a high test anxiety score of 0, the predicted mean test performance is 84.7."

ASA provides different types of smoothers for continuous and binary outcomes.

MULTIPLE SMOOTHERS ON THE SAME PLOT

As discussed in Chapter 11, it is possible to plot multiple smooths on the same graph for different groups to provide insights into moderated relationships. The program in the ASA software to accomplish this is "Lowess smoothing for multiple groups" in the folder Robust Statistics > Smoothers and Exploratory Curve Analysis > Single Predictor. In this example, we examine the effects of stress on anxiety for males and females separately. Stress is a multi-item scale that ranges from 1 to 5 with higher scores indicating greater stress. A score of 3 represents typical levels of stress experienced by people. Anxiety is measured on a 0 to 100 scale with higher scores indicating higher felt worry and anxiety. A score of 50 represents typical anxiety levels in people. Figure 1.2 presents a plot where the smooth for males is the dashed line and the smooth for females is the solid line.

Mean gender differences in anxiety are indicated by the degree of separation between the two lines at any given value of stress. At low levels of stress, there is little difference in anxiety for males and females. As stress increases, the gender difference becomes larger, with males experiencing increasingly more anxiety than females at any given value of stress.

Another way of thinking about the plot is in terms of the function relating stress to anxiety for males as opposed to the function relating stress to anxiety for females. For females, the function is basically linear and relatively flat: Increases in stress lead to some increases in anxiety, but not by much. For males, the function is non-linear and has the shape of a power function (see Chapter 8 on Mathematical Modeling in the main textbook): As stress increases at its low end, anxiety increases somewhat but at a score of around 2 or 2.5, the effect of stress on anxiety accelerates considerably.

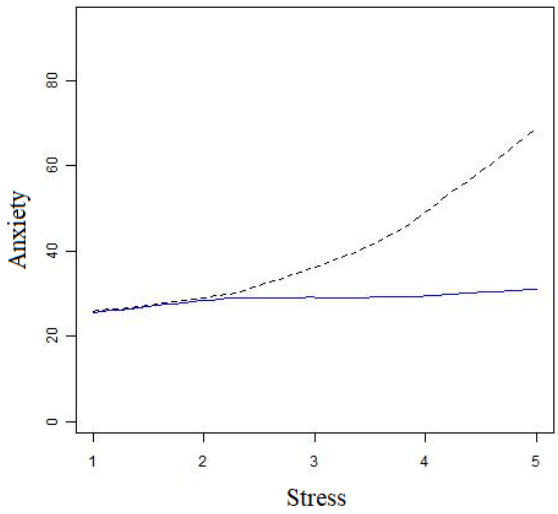

FIGURE 1.2. Smoother for Males and Females

Based on these smooths, there is an obvious interaction (moderated relationship) between gender and stress relative to their influence on anxiety. The data analyst might then formally model these differences using interaction terms in non-linear regression.

CONCLUDING COMMENTS

When working with quantitative variables, we find it instructive to always examine the relationship between them on an exploratory basis using smoothers.

REFERENCES

Wilcox, R. (2017). *Introduction to robust estimation and hypothesis testing*. New York: Academic Press (fourth edition).