Quiz #2 (Ch. 7) – Solutions

1. Match the definitions to the following extensions of the simple linear regression model for predicting a response variable, Y, from a single predictor variable, X.

a) Polynomial regression.
b) Step functions.
c) Regression splines.
d) Smoothing splines.
e) Local regression.

- Fit a smooth curve to the data by minimizing a weighted sum of the residual sum of squares and a measure of the total change in the curve's slope over the range of X.
- For different non-overlapping intervals of X fit separate low-degree polynomials that have continuity constraints at each interval break point.
- Fit a multiple linear regression model to X and sequential powers of X: X^2, X^3, ...
- For each unique value of X fit a weighted least squares regression to the points closest to X where the weights reflect the distance from X.
- Divide X into non-overlapping intervals and fit a different constant to each interval.

Solution: D, C, A, E, B

2. Select the true statements from the following. (Select all that apply.)

a) Polynomial regression often predicts independent test data better than regression splines because it can flexibly adapt to highly nonlinear relationships in the training data.
b) A natural spline is a regression spline with more stable, linear-constrained estimates at the boundaries.
c) An objective method for determining the number of knots for a regression spline is to minimize the cross-validated residual sum of squares.
d) For regression splines, it makes sense to place more knots where the regression function is stable and fewer knots where the function varies most rapidly.

Solution: B, C

3. True or false? Smoothing splines can be characterized as shrunken versions of natural cubic splines with knots at each of the unique predictor values.

Solution: True

4. True or False? The larger the value of the span for a local regression model, the more flexible and wiggly the fitted line.

Solution: False

5. Select the true statements from the following statements about generalized additive models (GAMs) for a quantitative response variable.

a) The fitted or predicted values for a GAM are linear combinations of functions of the predictor variables.
b) Examples of the predictor functions in a GAM include polynomial regression, regression splines, smoothing splines, and local regression.
c) GAMs are fit iteratively to partial residuals, which model the association between a particular predictor and the remaining unexplained variation in the response.
d) Since GAMs are additive models, it is not possible to include interaction terms of the form XjX_k.

Solution: A, B, C