

Regression Splines

Constraints and Degrees of Freedom

- In the previous example, we started with a cubic piecewise polynomial with 8 unconstrained parameters, so we started with 8 **degrees of freedom** (df)
- We initially imposed one constraint, which restricted one parameter, so we lost a degree of freedom $8 - 1 = 7$
- With the further two constraints: $8 - 3 = 5$ df
- In general, a cubic spline with K knots has $4 + K$ degrees of freedom. In R we can specify either the number of knots or just the degrees of freedom.

A degree- d regression spline is a piecewise degree- d polynomial with continuity in derivatives up to degree $d - 1$ at each knot

Natural Splines

- Regression splines have high variance at the outer range of the predictor (the tails)
- The confidence intervals at the tails can be wiggly (especially for small samples)

Natural splines are extensions of regression splines which remedy these problems

Two additional constraints at each boundary region:

1. The spline function is constrained to be close to linear when $X < \text{smallest knot}$
2. The spline function is constrained to be close to linear when $X > \text{largest knot}$

