Figure 1: Plot of sin(x) with random noise added to the data for x between 0 and 2pi (scatter plots). Different fitting polynomials are used to fit the data (solid curves).

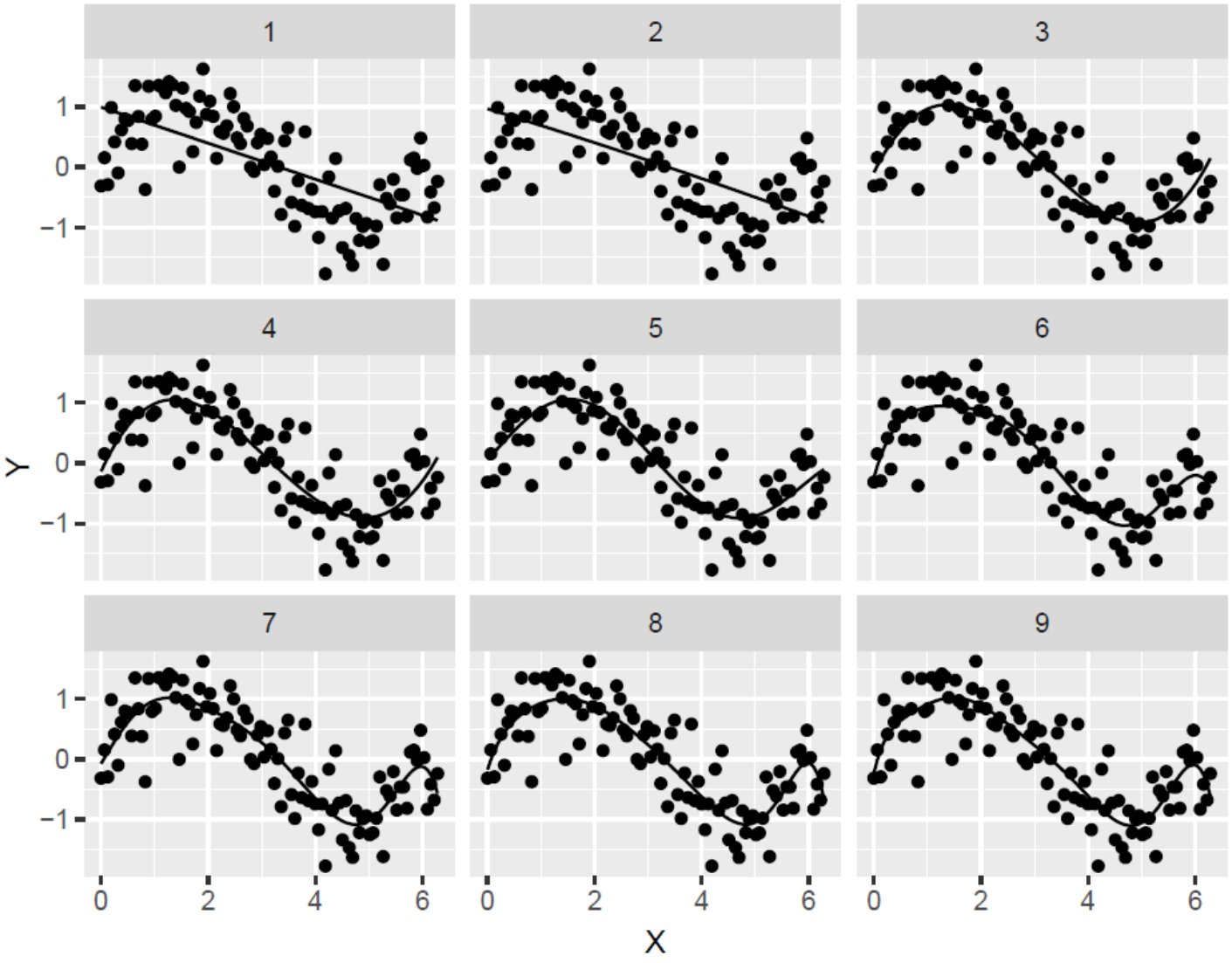
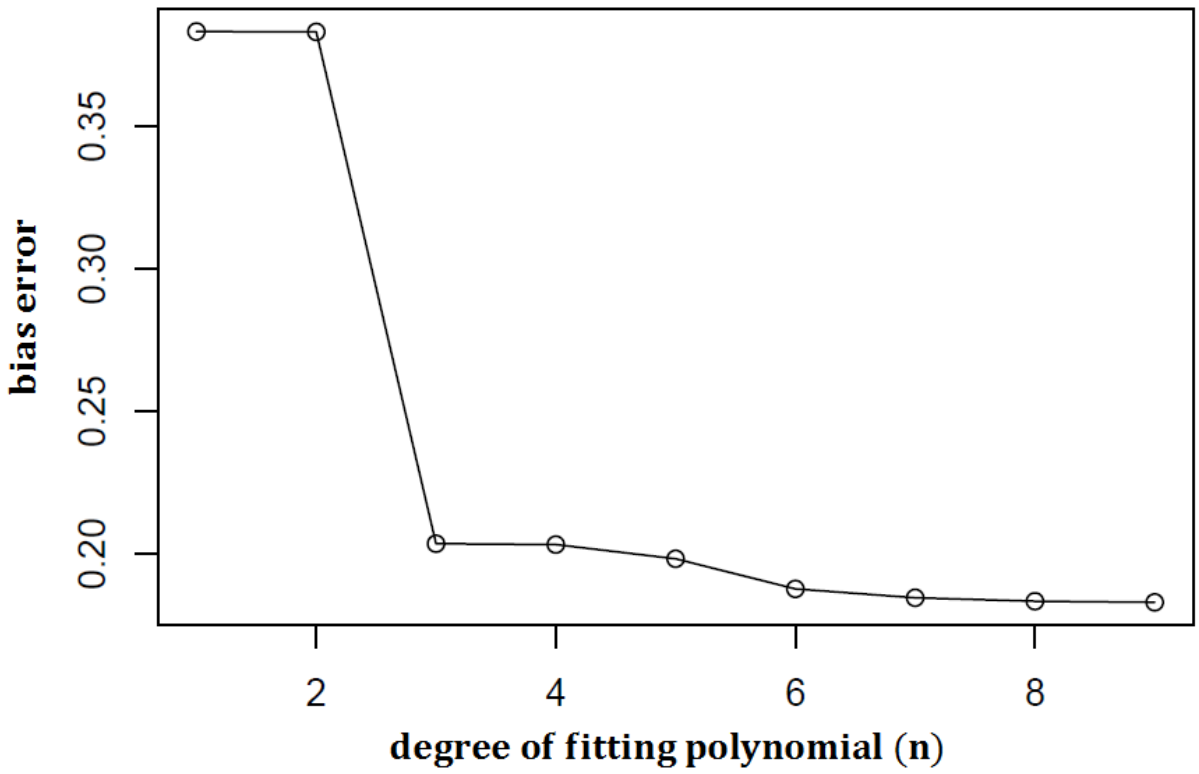
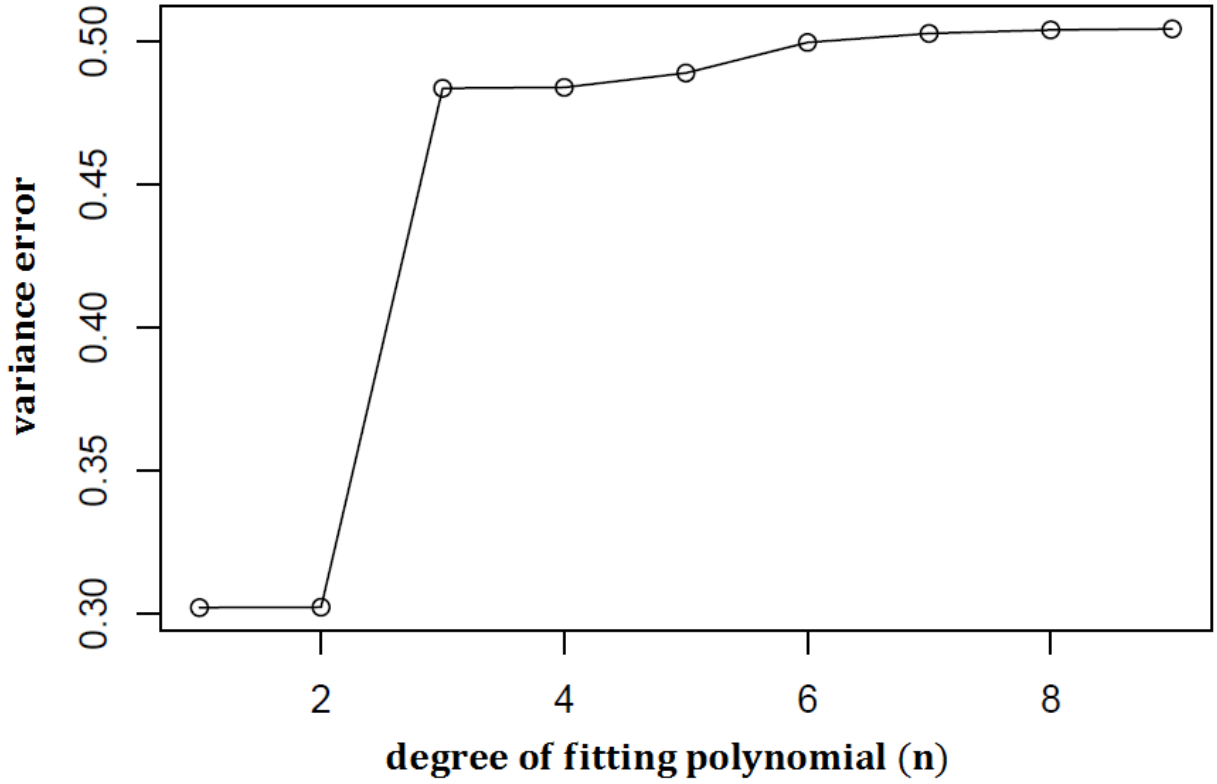


Figure shows that a good compromise between the **bias-variance tradeoff** is when the degree of the fitting polynomial is n = 5.

When n < 5, most of the error is due to high bias (bias error), as the model is overly simple.

When n > 5, the error is dominated by variance error, as the model is too complex so that it captures random noise in the data.


bias error
0.20
0.25
0.30
0.35
2
4
6
8
degree of fitting polynomial (n)


variance error
0.30
0.35
0.40
0.45
0.50
2
4
6
8
degree of fitting polynomial (n)