STAT452/652 Solution to HW02d

Due on Sept 25, 2020

1 Lecture 2d: How the Universe Works

1.1 Question 1

Looking at the bottom two plots in Figures 1 and 2, we see that the quartic model tends to be closer to the true structure when $n = 25$. This difference is largest in regions with strong curvature. The difference is smallest at the edges of the data, where neither model does very well (although the quartic has the most extreme deviations).

Results for $n = 50$ are given in Figures 3 and 4. The two plots show the same overall pattern, but in both cases the spread of lines/curves is lower (they have lower variance).

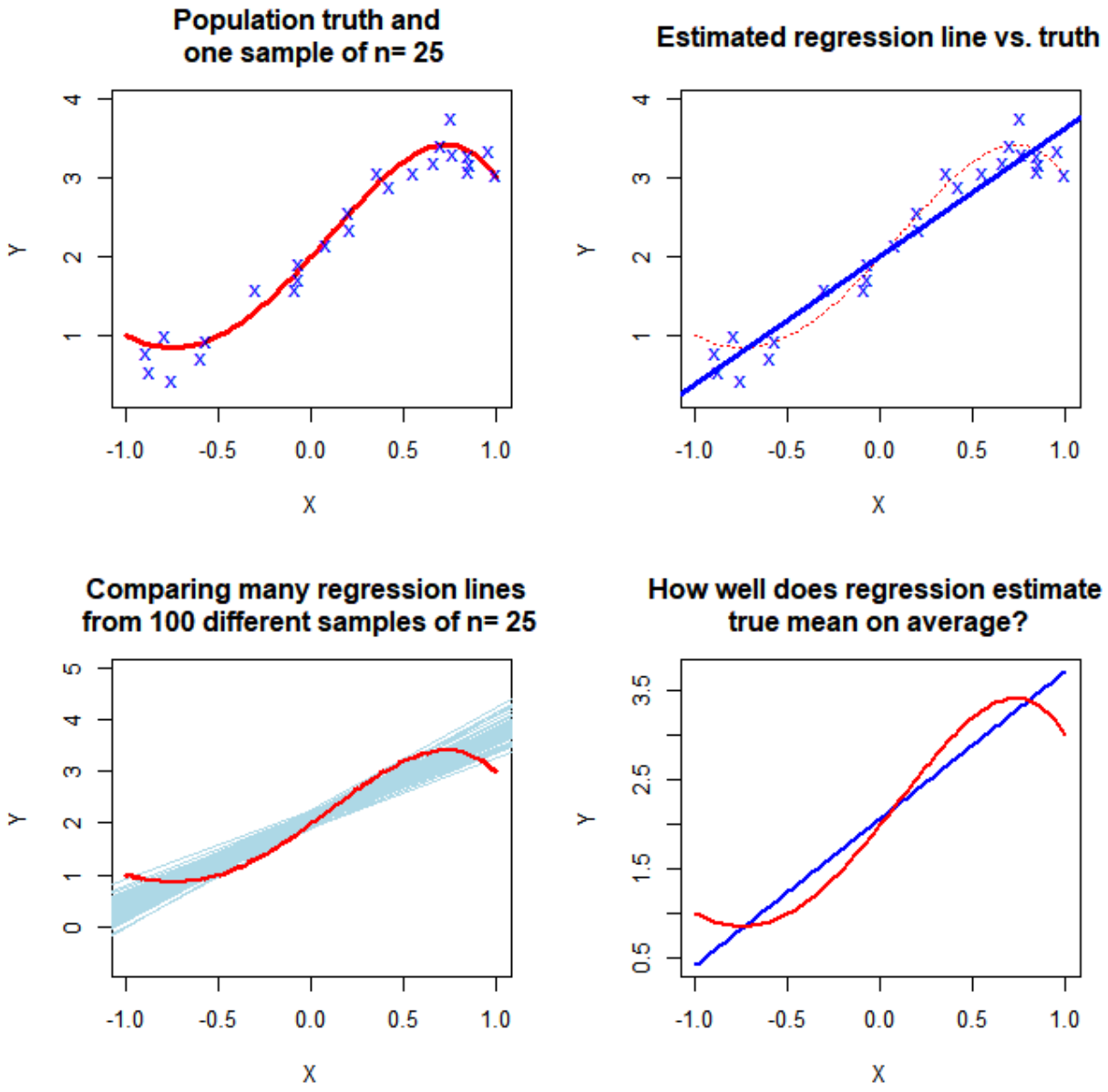

Figure 1: Plots for linear models with n=25.

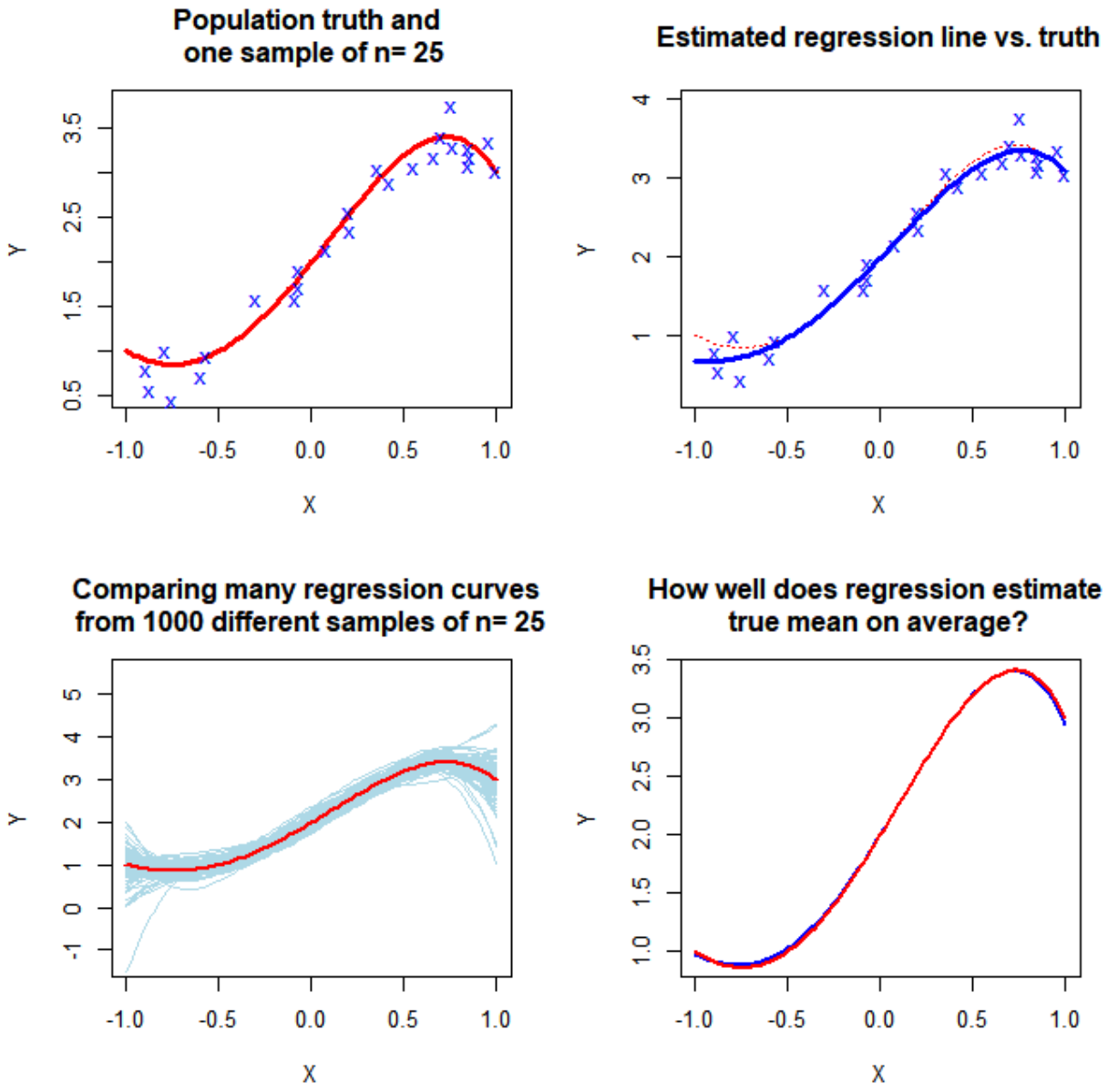

Figure 2: Plots for quartic models with n=25.

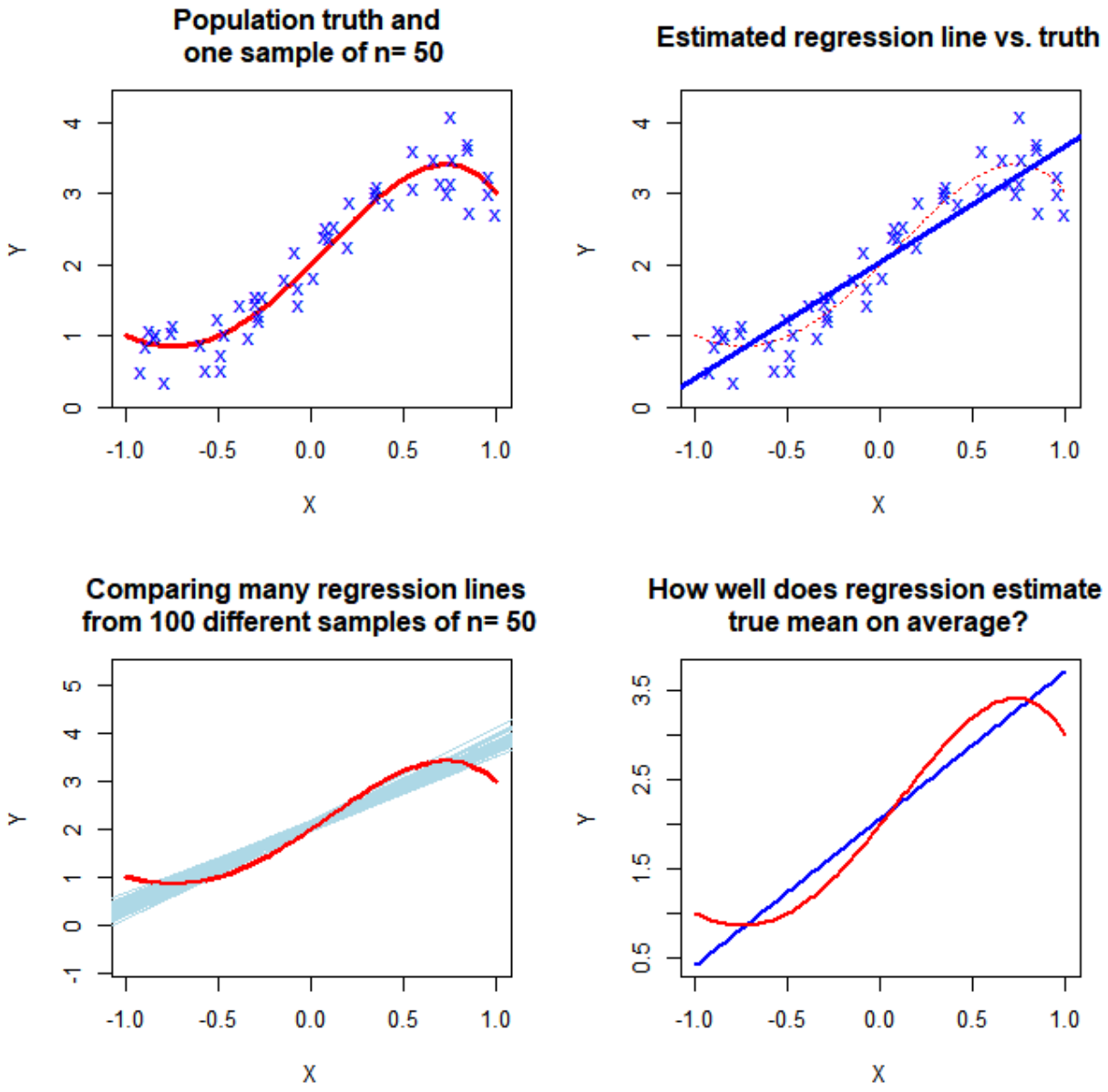

Figure 3: Plots for linear models with n=50.

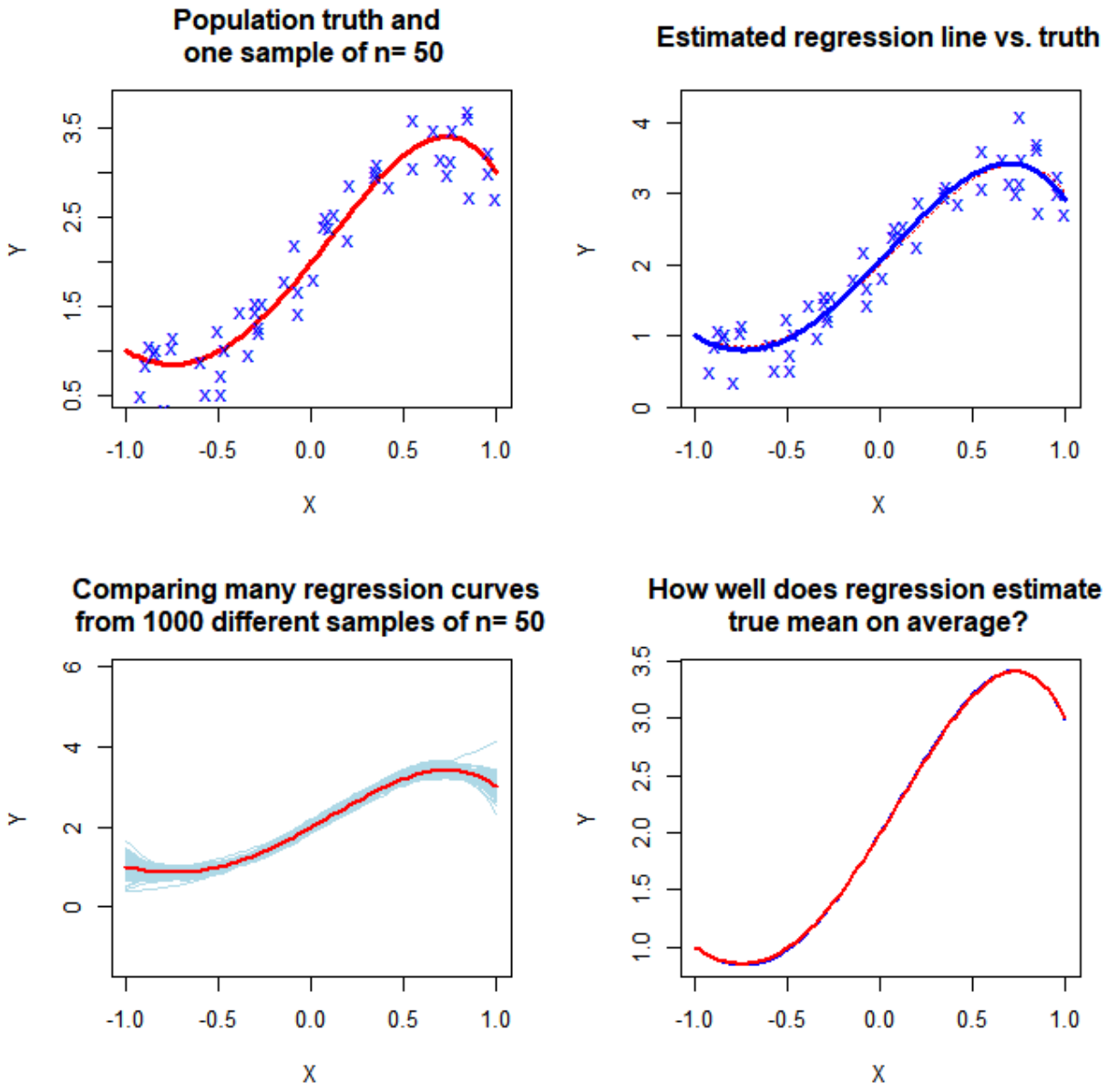

Figure 4: Plots for quartic models with n=50.

1.2 Question 2

(a) i. For a similar structure to the one in Figure 2 of Lecture 2d but with less curvature, we would expect to see less bias (because the true structure is not as different from a line as it was before)

ii. We expect to see the same variance, because variance comes from the sample errors, which change from sample to sample, not the shape of the underlying structure, which is not changing.

(b) As we increase the sample size, the variance will decrease, but the bias will not change.