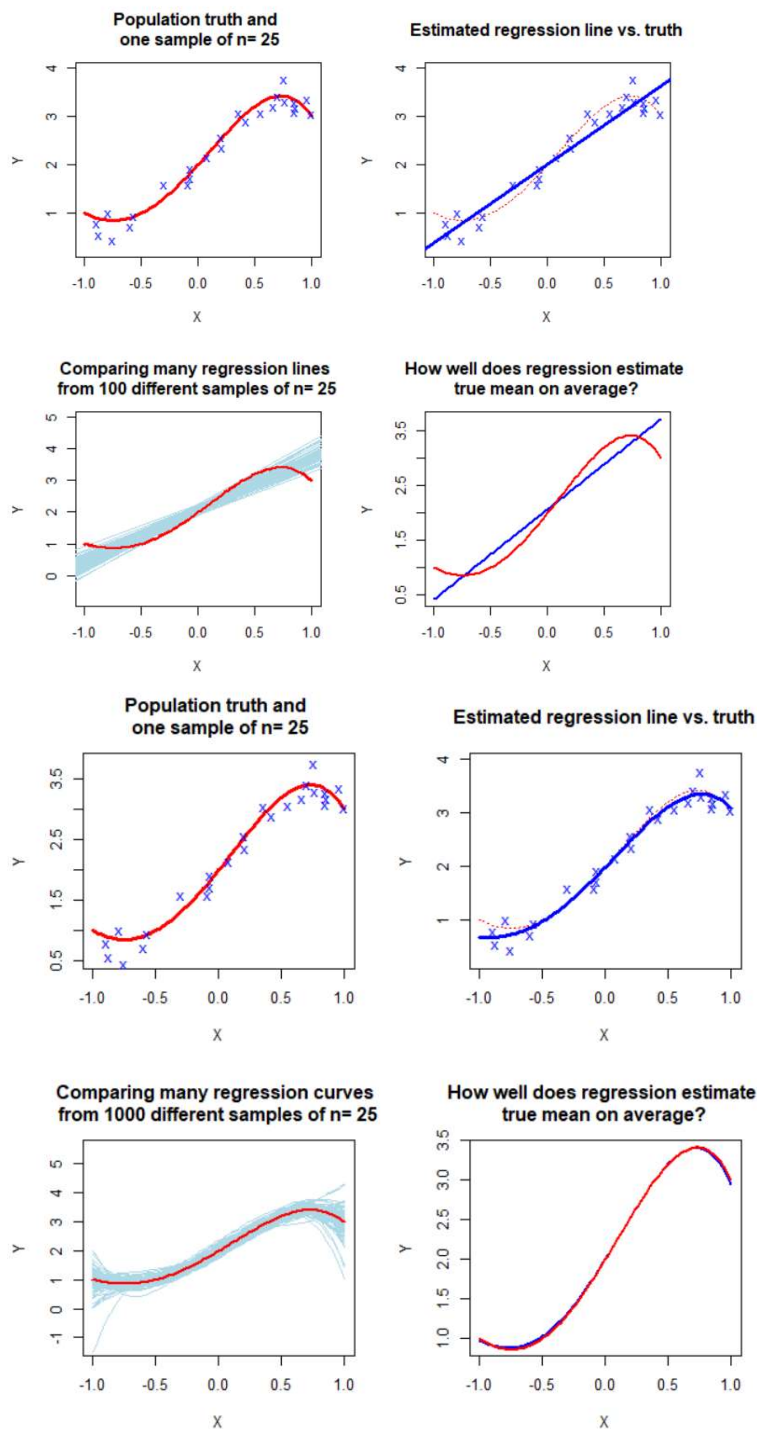


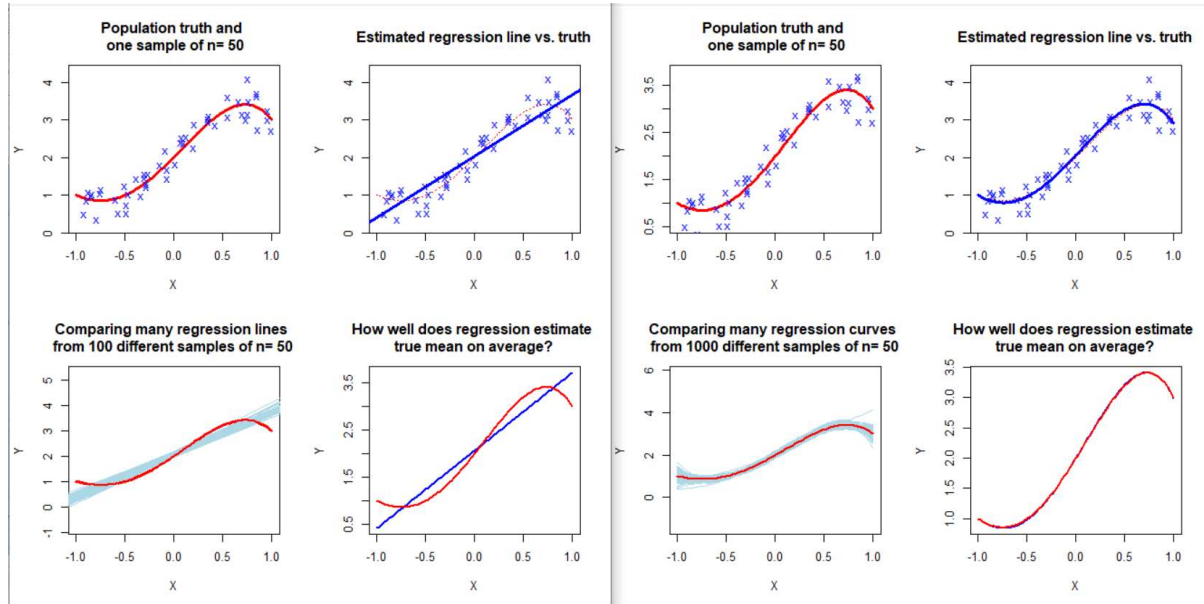
1. Here look a little more at the effect of sample size on fitting regression models. Use the R program **RegressionBiasVarianceHighBias.R** to do the following. This code produces the two plots as seen in Figure 4 when all the code is run. You do not need to understand how the program works or rewrite any code. A few lines from the top is a line that says “ $n = 10$ ”, which controls the sample size. We saw plots for $n = 10$ and $n = 100$ above.

(a) Change n to be 25 and rerun the entire code. Report these figures. Focusing on the bottom left plots, which model seems to get closer to the true structure across the whole range of X : the linear or the 4th-order polynomial?



➔ the 4th-order polynomial fits better than linear.

(b) Repeat for $n = 50$, reporting the figures and commenting on how the bottom left plots change.



➔ The blue lines get slimmer compare to blue lines from $n=25$, and also blue lines in 4th order polynomial get closer to population.