



**FIGURE 3.3.** A simulated data set. Left: The red line represents the true relationship,  $f(X) = 2 + 3X$ , which is known as the population regression line. The blue line is the least squares line; it is the least squares estimate for  $f(X)$  based on the observed data, shown in black. Right: The population regression line is again shown in red, and the least squares line in dark blue. In light blue, ten least squares lines are shown, each computed on the basis of a separate random set of observations. Each least squares line is different, but on average, the least squares lines are quite close to the population regression line.