Olivia Guswiler / Caleb Milford

Lab Assignment 2

2024-09-24

Part I: Effects on parameter estimate precision

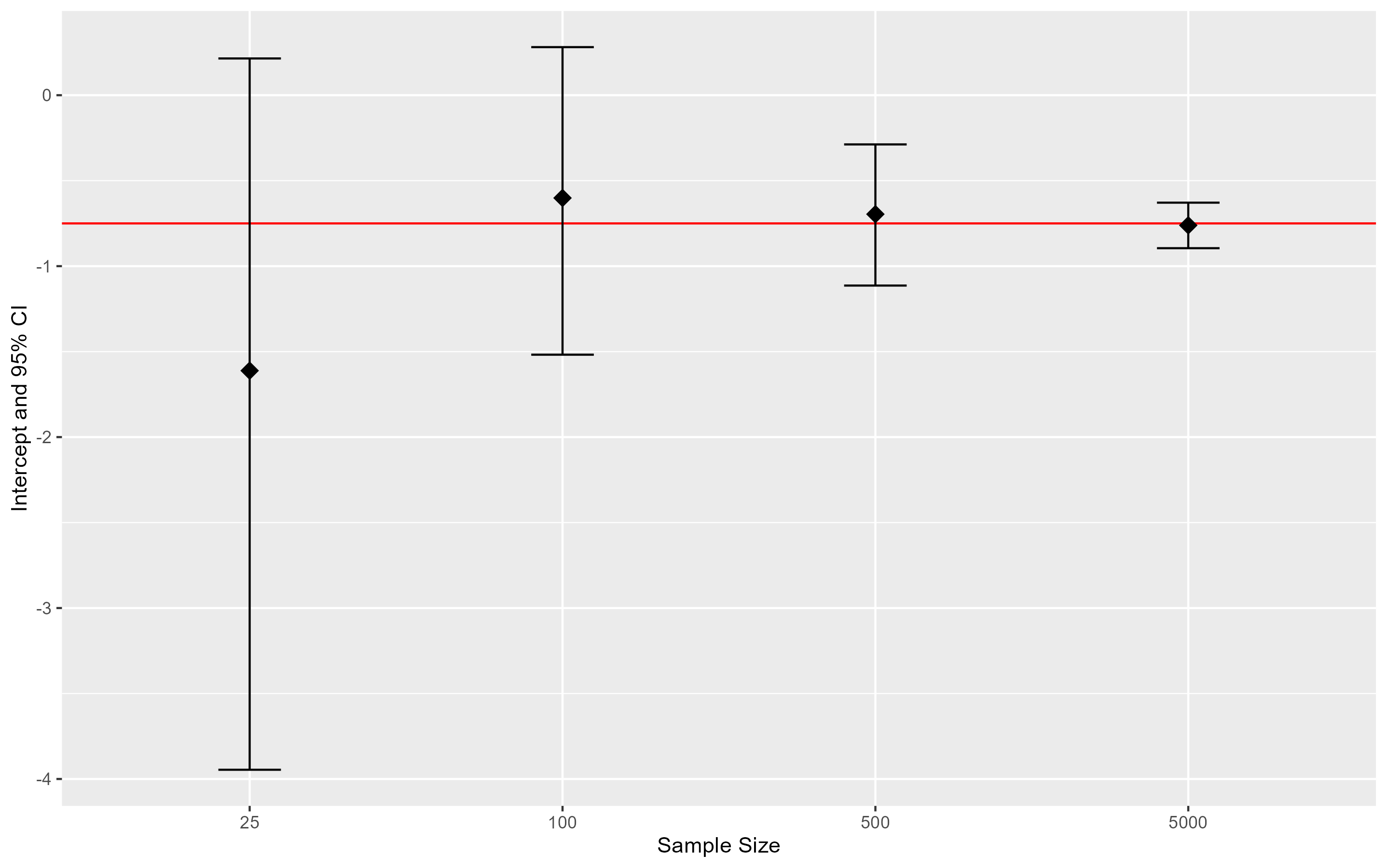
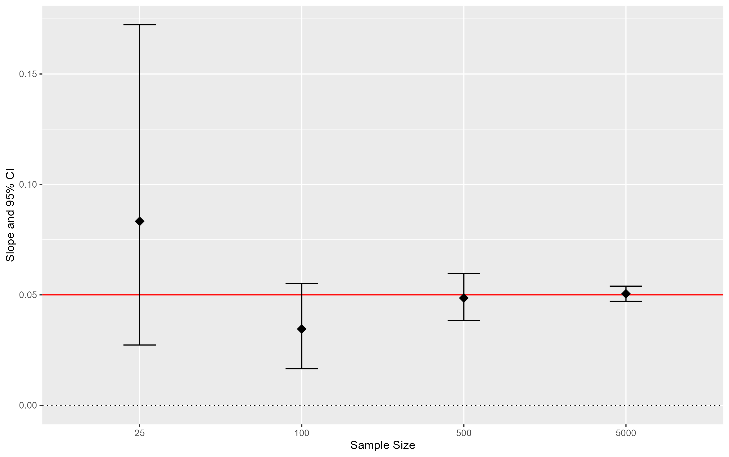
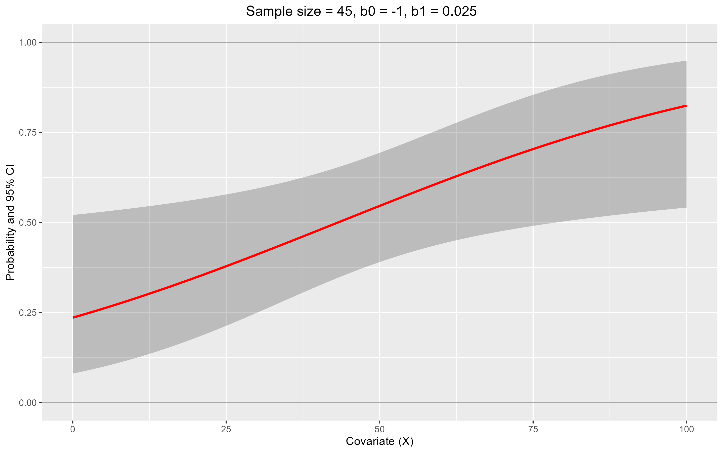
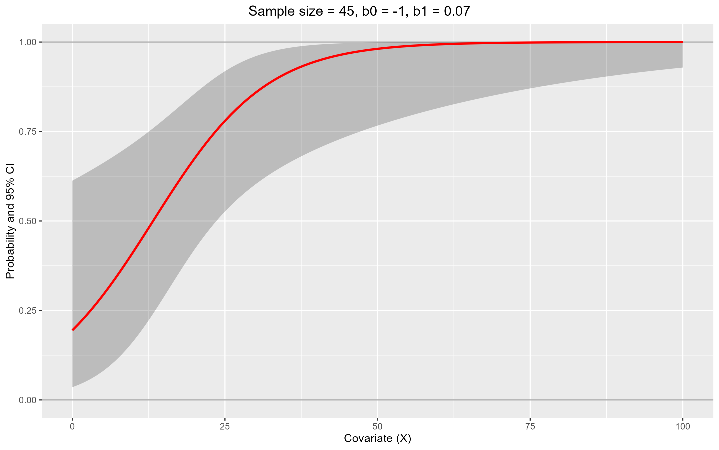
We observed how expected values and confidence intervals changed as sample size increased (*n* = 25, 100, 500, and 5000) while intercept and slope remain constant (β0 = -0.75, β1 = 0.05). As our sample size increased our estimates became more precise and accurate (see Figure 1). Increasing precision is evident by decreasing width of 95% CIs as sample size increases and increasing accuracy is evident by the estimated means generally falling closer to the true value as sample size increases. Because smaller sample size is correlated with less precise estimates (greater width of confidence intervals), it is more likely that a slope estimate’s 95% CI from smaller sample sizes would include zero. Therefore, using the criterion that inferences should only be made when slope does not include zero, making inferences from slope estimates are less likely to be possible at small sample sizes.

Figure 1. Intercept (left) and slope (right) estimates with 95% confidence intervals calculated from simulated data at sample sizes of 25, 100, 500, and 5000. The red line on each graph indicates the true value of the parameter (β0 = -0.75, β1 = 0.05).

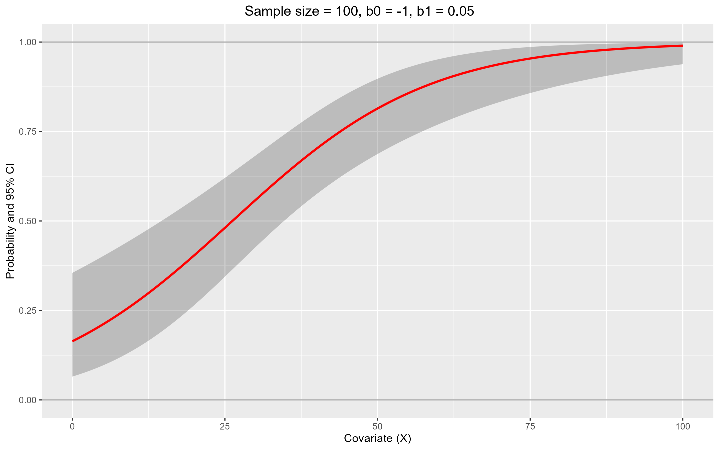
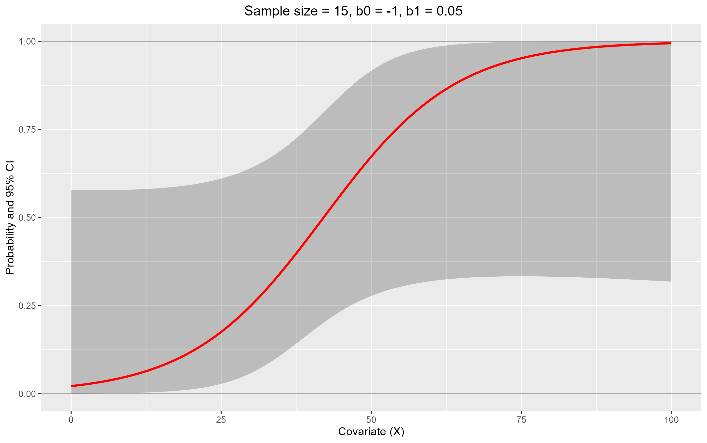
Part II: Plotting expected values from a binomial GLM

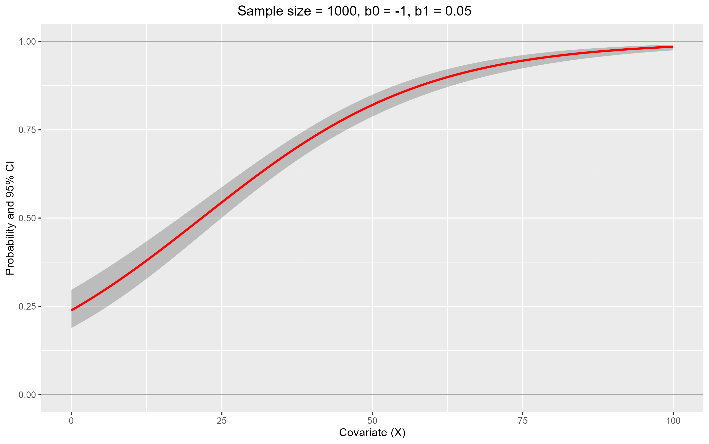
We observed the expected values calculated using a binomial GLM where sample size and intercept remain constant (*n* = 45, β0 = -1), but slope is changed (β1 = 0.025 and β1 = 0.07; see Figure 2). At a slope of 0.025, the expected probabilities appear near linear in relation to the covariate (X). At a slope of 0.07, expected probabilities have a long-tailed distribution, reaching a threshold near 100% probability between 25 and 50 units of X.

Figure 2. Plots of expected values calculated from a binomial GLM where *n* = 45 and β0 = -1, but slope is either 0.025 (left) or 0.07 (right). The red line indicates the expected values with the shaded area displaying 95% confidence intervals of the values.

Part III: Effect of sample size on expected value precision

We observed the expected values calculated from a binomial GLM where slope and intercept remained constant (β0 = -1, β1 = 0.05), but sample size was increased (*n* = 15, 100, 1,000, and 10,000; see Figure 3). Similar to what we discussed in Part I, as sample size increased our estimates became more precise and accurate. The slope of the expected values deviated most drastically at *n* = 15 when compared to our greater sample sizes. At *n* = 15, the slope displayed a distinct concave-convex shape, at *n* = 100 the slope was slightly concave before shifting to convex, and at *n* = 1,000 and *n* = 10,000 both slopes were near identical convex shapes. 95% CIs were greatest at *n* = 15, covering nearly half the graph at any given point. 95% CIs were much tighter at *n* = 100 and nearly indistinguishable from the expected values at *n* = 10,000.



A graph with a red line

Description automatically generated

Figure 3. Plots of expected values calculated from a binomial GLM where β0 = -1 and β1 = 0.05, but sample size was set to a value of either 15 (top left), 100 (top right), 1,000 (bottom left), or 10,000 (bottom right). The red line indicates the expected values with the shaded area displaying 95% confidence intervals of the values.