Lab: Priors & The Power of Fake Data

1. Posterior Credible Intervals via Free Throws

Recall the beta binomial model:

$$\begin{cases} W \sim \text{Binomial}(W+L, p) & \longrightarrow \text{likelihood} \\ p \sim \text{Beta}(\alpha, \beta) & \longrightarrow \text{prior} \end{cases}$$

By Bayes Rule, the posterior distribution of $p|W,L$ is proportional to $\mathbb{P}(p|W,L) \propto p^{W+\alpha-1}(1-p)^{L+\beta-1}$

so is $\boxed{p|W,L \sim \text{Beta}(\alpha+W, \beta+L)}$ [the symbol $\propto$ means "proportional to"]

Which reflects our Bayesian beliefs about p in a distribution.

α acts like a synthetic number of successful trials and β acts like a synthetic number of unsuccessful trials.

Since we now have a full distribution (a Beta distribution) encoding our beliefs (and uncertainty) about p given the data W, L, we can create a Bayesian posterior credible interval CI to summarize our uncertainty in p:

Find $CI = [A, B]$ so that

$$\mathbb{P}(A \leq p \mid W, L \leq B) = \mathbb{P}(A \leq \text{Beta}(\alpha + W, \beta + L) \leq B) = 0.95.$$

The CI is a function of the prior hyperparameters α, β.

Using the free throw data, compare these credible interval for various combos of α, β to the Wald, Agresti, and Bootstrapped confidence intervals.

- Note: Bayesian posterior credible intervals are NOT frequentist confidence intervals, even if they turn out similar! Credible intervals are the result of a full probability model.