STA 6352, Report 10.1

Carson Slater *Baylor University*

10.3.2 Logistic Regression

We now extend the example of the previous section to the logistic regression case. For the ith response y_i, let x_{1i} be a treatment indicator where $x_{1i} = 1$ for a new treatment of interest and $x_{1i} = 0$ for a standard treatment. Let $x_{2i}, \ldots, x_{qi}$ be other covariates. The model is

$$y_i \sim \text{Bernoulli}(\pi_i),$$
$$\text{logit}(\pi_i) = \theta_0 + \theta_1 x_{1i} + \sum_{j=2}^{q} \theta_j x_{ji}.$$

The parameter θ_1 represents the effect of the new treatment compared to placebo. The other covariates are captured in $\theta_2, \ldots, \theta_q$.

Sample Size Algorithm

1. Specify design/analysis prior parameters, coverage probability, desired interval width, significance level, and power.
2. For $m = 1, \ldots, M$ simulation replications at each n:
   (a) Generate parameter values from the design priors.
   (b) Generate covariates from their distributions.
   (c) Compute response probabilities using:
   $$\text{logit}(\pi_i) = \theta_0^{(m)} + \theta_1^{(m)} x_{1i} + \sum_{j=2}^{q} \theta_j^{(m)} x_{ji}.$$
   (d) Generate $y_i \sim \text{Bernoulli}(\pi_i)$.
   (e) Fit Bayesian logistic regression using MCMC and analysis priors.
   (f) Store interval width and whether $P(\theta_1 > 0 \mid y) > 1 - \alpha$.
3. For each n, compute average width and power across M simulations.
4. Fit a curve to $(n_i, w_A(n_i))$ or $(n_i, \hat{p}_n)$.

Parameter	Design Prior	Analysis Prior
θ_0	$\mathcal{U}(-0.5, 0.0)$	$\mathcal{N}(0, 10)$
θ_1	Fixed at 0.4055 (log OR = 1.5)	$\mathcal{N}(0, 10)$
θ_2	$\mathcal{U}(-0.6, -0.3)$	$\mathcal{N}(0, 10)$
θ_3	$\mathcal{U}(0.2, 0.4)$	$\mathcal{N}(0, 10)$
x_1	Bernoulli(0.5)	N/A
x_2	Bernoulli(0.4)	N/A
x_3	$\mathcal{N}(0, 1)$	N/A

Table 1: Design and analysis priors used in the simulation.

Design and Analysis Priors

Computational Considerations

In practice, we often use $M = 100$, 200, 500, or 1000 replications. The Average Length Criterion (ALC) stabilizes quickly, but the Posterior Probability Criterion (PPC) requires more replications to reduce variability.

Report 10.1

1. Reproduce the results in the example, including graphical summaries.
2. Approximate the sample size needed to obtain:
    - Average interval width of 1.5
    - Power of 0.6
3. Study the influence of design and analysis priors by varying:
    - SD of Normal priors
    - Bounds of Uniform priors
4. Use $M = 1000$ and original priors to zoom into sample sizes $n \in \{300, 310, \ldots, 500\}$ and refine power estimate for 0.6.

Parallelized