(11) Case Studies: Adaptive clinical trials

ST440/540: Applied Bayesian Analysis
Use of Bayesian methods in clinical trials

- Bayesian methods are becoming more common in clinical trials analysis

- We will study how to compute the sample size for a Bayesian clinical trial

- We will then explore adaptive clinical trials with interim analyses and Bayesian stopping rules

- This provides a clever application of prior distributions and uncertainty quantification
Sample size computation

- Say \( Y_1, \ldots, Y_n \sim \text{Normal}(\mu, \sigma^2) \)

- Our objective is to test \( H_0 : \mu = 0 \) versus \( H_A : \mu \neq 0 \)

- The power of the study is the probability (over random future datasets) of rejecting \( H_0 \) when in fact \( \mu \neq 0 \)

- “Everyone is Bayesian at the design stage”

- That is, we must specify values of \( \mu \) and \( \sigma \) so that the trial can be designed to have reasonable power if these values are true
The classic approach is to set true values $\mu_0$ and $\sigma_0$ based on literature review or the smallest interesting difference.

The power is then determined based on data being generated using these parameter values.

Rather than single values for $\mu_0$ and $\sigma_0$, we could represent our uncertainty using “design priors” for these parameters.

Denote this as $\mu_0 \sim \text{Normal}(\gamma_o, \tau_o^2)$ and $\sigma_o^2 \sim \text{InvG}(a_o, b_o)$.

We design the trial by averaging over our uncertainty in the true values.
Data analysis

- The data will be analyzed using Bayesian methods

- The “analysis prior” and design prior need not be the same

- We will use the Jeffreys priors as the analysis prior

- Perhaps an informative analysis prior could be used for $\sigma$ and an uninformative prior could be used for $\mu$?
Power simulation

- Fix $n$ and repeat the following simulation $N$ times:
  1. Generate $\mu_0$ and $\sigma_0$ from the design priors
  2. Generate $Y_1, \ldots, Y_n \sim \text{Normal}(\mu_0, \sigma_0^2)$
  3. Compute the posterior 95% intervals for $\mu$ using $Y_1, \ldots, Y_n$ and the analysis priors
  4. Reject the null if the interval excludes zero

- The power is approximated as the proportion of the $N$ intervals that exclude zero

- [Website](http://www4.stat.ncsu.edu/~reich/ABA/Code/OneStage)
Two-stage trials

- Now consider a two-stage trial

- In the first stage we collect \( n_1 \) samples

- After this stage we conduct an interim analysis and decide whether or not to collect an additional \( n_2 \) samples

- Two-stage (or generally adaptive) trials are designed to be more efficient and ethical than traditional static studies

- We will stop the trial if the likelihood of eventual success is near either zero or one
Interim analysis

► We will make our decision to continue or discontinue the trial based on the power for rejection the null if the additional $n_2$ samples are collected

► We compute the power using the same steps as before

► The only difference is the design prior

► We use the posterior distribution after collecting the first $n_1$ observations as the design prior

► The accurately conveys our uncertainty about the true parameters at this point in the trial

► [http://www4.stat.ncsu.edu/~reich/ABA/Code/TwoStage](http://www4.stat.ncsu.edu/~reich/ABA/Code/TwoStage)
Utilizing this method

- In reality, you could have to conduct a power study to select $n_1$ and $n_2$ before the trial begins

- This would require a costly simulation

- You would generate $N$ datasets and conduct an interim analysis for each simulated dataset

- You would then obtain estimates of frequentist Type I error and power for each $n_1$ and $n_2$

- If Type I error control is desired you could tune the criteria for rejecting the null
Quote from the FDA

- In traditional frequentist clinical trial design, the sample size is determined in advance.

- As an alternative to specifying a fixed sample size, the Bayesian approach (and some modern frequentist methods) may specify a particular criterion to stop the trial.

- Appropriate stopping criteria may be based on a specific amount of information about the parameter (e.g., a sufficiently narrow credible interval) or an appropriately high probability for a pre-specified hypothesis.
Quote from the FDA

- At any point before or during a Bayesian clinical trial, you can obtain the predictive distribution for the sample size.

- Therefore, at any point in the trial, you can compute the expected additional number of observations needed to meet the stopping criterion.

- In other words, the predictive distribution for the sample size is continuously updated as the trial goes on.

- Because the sample size is not explicitly part of the stopping criterion, the trial can be ended at the precise point where enough information has been gathered to answer the important questions.