Bayesian Monitoring for Group Sequential Trials

Will Gertsch

May 29, 2020

5-FU Trial

The 5-FU trial investigated the effect of 5-Fluorouracil and levamisole on the survival of patients with colorectal cancer.

- *Goal of the trial*: see if the drug improves patient survival.
- The usual trial design is to enroll N patients and randomize into treatment and control groups.
- At the end of the trial, we see if the patients in the treatment group had longer survival times than the patients in the control group.

Sequential Trials

What if N is not fixed?

- In the 5-FU trial, patients and data accumulate throughout the trial.
- At preset numbers of patients, conduct analysis using current data.
- If there is enough evidence (for or against the treatment), stop the trial.
- *Monitoring*: the process of sequentially analyzing the accumulating data in order to decide if the trial should stop.

Monitoring Terms

A few technical terms:

- *Interim analyses*: statistical analyses at preset numbers of patients to determine the effect of the treatment given available evidence.
- *Stopping rule*: Stop the trial once a preset level of significance has been reached.

The Frequentist Approach

The frequentist approach to analyzing sequential trials is based on adjusting for the effect of multiple interim tests on the overall type I error rate, α.

- *Example*: a trial with 1 interim test and one final test, both at $\alpha = 0.05$.
- Because of multiple testing, the overall α will be greater than 0.05.
- Using a Bonferroni correction to adjust, the new α's for each test should be 0.025.

Problems with the Frequentist Approach

There are variety of frequentist methods to monitor sequential trials, but these methods are philosophically awkward.

- This is easy to see with the 1 interim test and one final test trial.
- The trial has 200 subjects and ends with a statistically significant analysis ($P = 0.03$) in favor of the treatment.
- The interpretation of this result depends on whether or not there is more data to accumulate.
- If there are more analyses, then this interim result might not be significant.

The Bayesian Approach

The Bayesian approach to monitoring uses the posterior distribution of the treatment effect.

- For each interim analysis, compute a posterior distribution for the treatment effect based on current data.
- Estimation and hypothesis testing are conditional *only* on current data and not on future plans for data collection and analysis.
- Stopping rules are based on the current posterior at an interim analysis.

Bayesian Stopping Rules

Bayesian stopping rules work by dividing the possible values of the outcome into regions.

- Given the treatment effect δ, we can use a fixed value δ_A as a reference point.
- This leads to the regions $\delta \leq \delta_A$ and $\delta > \delta_A$.
- Use the posterior of δ to compute $P(\delta > \delta_A \mid Y)$ where Y is the observed data.
- If $P(\delta > \delta_A \mid Y)$ is greater than some cut-off probability, stop the trial and conclude $\delta > \delta_A$.

Bayesian Approach Applied to the 5-FU Trial

Example: Analyzing the 5-FU trial with a Bayesian approach.

- We want compare the time to death in the treatment group to time to death in the control group.
- This is survival data, so the natural choice of outcome is the log-hazard ratio, δ.
- The trial was originally designed to detect the alternative $\delta_A = 0.3$ with the null hypothesis $H_0 : \delta = 0$.

Choosing a Prior

We can model δ using a normal-normal model, but how do we specify the prior?

- There are several possible normal priors to choose from.
- Choose a *skeptical prior*, centered on the null hypothesis value δ_0.
- Choose a prior standard deviation, σ_0, such that there is a 5% a priori probability of exceeding δ_A.
- Therefore, $\sigma_0 = \frac{\delta_A}{Z_{.95}} = 0.182$

Therefore, a skeptical prior for the 5-FU trial is

$$\delta \sim N\left(\delta_0, \sigma_0^2\right) = N\left(0, 0.182^2\right)$$

Posterior Parameter Estimates

For the first stage of the trial:

- We observe $N = 192$ patients and estimate $\hat{\delta} = 0.4$.
- Compute normal likelihood using N, $\hat{\delta}$, and the standard error of $\hat{\delta}$.

Therefore, we obtain a normal posterior with mean δ_p and standard deviation σ_p.

$$\delta \sim N(\delta_p, \sigma_p^2) = N(0.246, 0.113^2)$$

5-FU Conclusions

The posterior distribution is $\delta \sim N(0.246, 0.113)$.

- Suppose we want to see if treatment is better than control.
- We define "better" as corresponding to $\delta > 0.285$.
- Using the posterior, we compute $P(\delta > 0.285) = 0.365$.
- Therefore, the probability that the treatment is strictly better than the control is 0.365.

Check stopping rule:

- Stopping rule: $P(\delta > 0.285) \geq 0.75$
- Therefore, we do not have sufficient evidence for the superiority of the treatment and we continue the trial.

Summary

- The analysis for group sequential trials needs to account for the accumulation of data and interim analyses.
- There is a Frequentist approach, but it is philosophically awkward.
- The Bayesian approach is ideal because the key components of analyzing sequential trials are all reduced to facts about the posterior distribution.

References

- Freedman, L.S., Spiegelhalter, D.J. and Parmar, M.K.B. (1994), The what, why and how of bayesian clinical trials monitoring. Statist. Med., 13: 1371-1383.
- Whitehead, J. (2014). Sequential Methods for Clinical Trials.
- O'Brien, Peter C., and Thomas R. Fleming. "A Multiple Testing Procedure for Clinical Trials." Biometrics 35, no. 3 (1979): 549-56. Accessed May 15, 2020. doi:10.2307/2530245.