STAT 778 Final Project: Implementation of Igeta, Takahashi, and Matsui 2018

Tom Wallace

April 19, 2018

1 Introduction

This paper documents an implementation of the technique of Igeta, Takahashi, and Matsui 2018. It gives some background on their method; presents a fast software program in C implementing their method; and conducts verification and validation of this program.

2 Methods and Code

2.1 Background

Overdispersion refers to a situation in which the variance of a dataset exceeds that which would be expected under the assumed statistical distribution. Overdispersion is common in count data. As a motivating example, consider the Poisson distribution. A single parameter determines both the mean and variance (i.e., for $X \sim \text{Poisson}(\lambda)$, $\mu = \sigma^2 = \lambda$). As a consequence, if the mean and variance differ in a dataset, the researcher cannot adjust parametric assumptions for without also affecting the other. If not addressed, overdispersion results in distorted test statistics and estimated standard errors. Clinical trials often feature count data: e.g., a trial of an anti-epsilepsy treatment may use *number of seizures over study period* as the outcome variable. The serious consequences of statistical error in such settings demands a rigorous method for dealing with overdispersion. Igeta, Takahashi, and Matsui 2018 is a new entry to the large literature on this topic. In particular, it presents methods for calculating statistical power and sample size in the presence of misspecified variance.

2.2 Methods

Consider a randomized control trial featuring n subjects. Subjects are randomly assigned to a treatment or control group. Let n_i be the sample size of the ith group, with $i \in \{1, 2\}$. Let $q_i = \frac{n_i}{n}$ be the proportion of subjects assigned to the

ith group. Let π_i be the allocation ratio to the ith group, where $\lim_{n\to\infty} q_i = \pi_i$ Let X_{ij} be an indicator of the treatment assignment for the jth subject ($j = 1, 2 \ldots n$) in the ith group. $X_{ij} = 0$ indicates assignment to the control group. $X_{ij} = 1$ indicates assignment to the treatment group. Let Y_{ij} be the number of events associated with the subject during the period $[0, T_{ij}]$, from the point of treatment assignment to the end of follow-up.

Igeta, Takahashi, and Matsui 2018 consider a statistical test for group comparison. The test is based on a quasi-likelihood using a sandwich-type robust estimator; for further details, consult their paper. The goal is to estimate coefficients in a Poisson model:

$$\lambda_i = \exp\left(\beta_0 + \beta_1 X_{ij}\right)$$

Igeta, Takahashi, and Matsui 2018 propose that the asymptotic power of the test is

$$\Pr\left(Z > z_{1-\alpha/2}\right) = 1 - \Phi\left(z_{1-\alpha/2}\sqrt{\frac{W_0}{W_1}} - \sqrt{n}\frac{\beta_1}{\sqrt{W_1}}\right) \tag{1}$$

and the sample size that provides power greater than or equal to $1 - \beta$ is

$$n \geq \frac{(z_{1-\alpha/2}\sqrt{W_0} + z_{1-\beta}\sqrt{W_1})^2}{\left(\log(\lambda_2/\lambda_1)\right)^2} \tag{2}$$

where:

- z_a is the lower a point of the standard normal distrbution
- Φ is the CDF of the standard normal distribution
- When a constant follow-up period τ

3 Simulation Study

4 Conclusion

5 References

[1] Masataka Igeta, Kunihiko Takahashi, and Shigeyuki Matsui. "Power and sample size calculation incorporating misspecifications of the variance function in comparative clinical trials with over-dispersed count data". In: *Biometrics* (2018). DOI: 10.111/biom.12878.