Example Statistical Analysis Plan – Parallel-Group Trial with a Continuous Outcome
CDER Center for Clinical Trial Innovation (C3TI) Bayesian Statistical Analysis Demonstration Project

Overview

When using a parametric statistical test to compare two means, it is often difficult to know whether the data are normally enough distributed to trust p-values and confidence limits from a t-test. It is also difficult to know whether to use the standard equal-variance t-test or the Welch t-test. Examining the data to make decisions about transformations to normality and equal variance will distort frequentist operating characteristics, but it is difficult to completely pre-specify the statistical analysis. A Bayesian approach has a parameter for each known aspect of the statistical model that is important. This makes Bayesian uncertainty intervals slightly wider by a proper amount, to take model uncertainty into account.

Study Design

Study XY-01 is a double-blind, parallel-group, two-treatment, randomized controlled trial of drug (group B) vs. placebo (group A) focused on acute hypertension in an emergency department setting. The primary analysis is intent-to-treat and the primary response variable is systolic blood pressure (SBP) measured at 2 hours post-randomization.

Statistical Analysis

Note: In practice the test below would be replaced with analysis of covariance adjusted for baseline SBP. For simplicity in this example, baseline SBP is not included in the statistical model.

The Bayesian t test will be used to quantify the evidence for effectiveness of drug B in reducing SBP, specifically by computing the posterior probability that the B-A SBP difference is negative. As a secondary assessment, the posterior probability that the reduction exceeds 3 mmHg will be computed. The analysis does not assume that the SBP variances are equal in the two treatment groups, and puts a prior on their ratio, favoring equality but allowing for inequality. The analysis does not assume that SBP is normally distributed, instead assuming the raw data come from a t distribution with ν degrees of freedom. This allows the data to be heavier tailed than the normal distribution. When $\nu > 20$ the distribution is effectively normal. A gamma prior distribution is assumed for ν.

The prior distribution for the difference Δ in mean SBP was developed during a meeting among the sponsor and FDA statistical and medical reviewers. The chosen prior is such that a reduction and an increase are equally likely (the mean of the prior is zero) and such that it is very unlikely to achieve a reduction > 15 mmHg or < -15 mmHg. Specifically, the prior is chosen as normal with standard deviation (SD) σ chosen such that $\Pr(\Delta < -15) = \Pr(\Delta > 15) = 0.025$. This implies $\sigma = \frac{15}{\Phi^{-1}(0.975)} = \frac{15}{1.95996} = 7.653$.

Letting r denote the ratio of variances for the two treatment groups, assume a mean-zero normal prior for $\log(r)$ with standard deviation chosen so that $\Pr(r > 1.5) = \Pr\left(r < \frac{1}{1.5}\right) = 0.05$. The required SD is $\frac{\log(1.5)}{\Phi^{-1}(0.95)} = 0.2465$.

The prior for the degrees of freedom ν in the normal data distribution is gamma (2, 0.1), which is the default in the software setup below.

When the analysis is completed, secondary parameters will be summarized as well as the treatment effect parameter. This will assist in future study planning. Secondary assessments will include the posterior distribution for the variance ratio, and $\Pr(\nu > 20)$. The latter posterior probability is essentially the probability of normality.

A secondary analysis will be done to assess evidence for an age-dependent treatment effect. Age and age-squared and their interactions with treatment will be added to the model, with flat priors for the associated parameters. Evidence for heterogeneity of treatment effect due to varying age will be taken as a high posterior probability that the treatment effect evaluated at the lower quartile of age is greater than the treatment effect evaluated at the upper age quartile (this is not a subgroup analysis). This posterior probability is computed by the proportion of posterior draws for which the above difference is positive.

The decision criteria for the primary analysis are based solely on the posterior distribution for Δ given the analysis prior, data, and data model. Sufficient evidence for efficacy will exist if the posterior probability of positive benefit, Pr(Δ> 0), exceeds a context-specific threshold and the probability that the benefit is more than 5 mmHg, Pr(Δ> 5) exceeds a smaller context-specific threshold. Typical posterior probability thresholds for a conclusion of an effect in the right direction are between 0.9 and 0.99.

Software

The R brms package[1] brm function will be used for the analysis. Since the mean and variance are allowed to change with treatment, a double model is specified. Prototypical code is below.

```
# flat (non-informative) prior for intercept
pr0 <- set_prior("", class="Intercept")
# normal(0, 7.653) prior for difference in mean SBP
pr1 <- set_prior("normal(0, 7.653)", class="b", coef="txB")
# normal(0, 0.2465) for log SD ratio
pr2 <- set_prior("normal(0, 0.2465)", class="b", coef="txB",
            dpar="sigma")

brm(bf(sbp   ~ tx,
```

[1] Paul-Christian Bürkner (2017). brms: An R Package for Bayesian Multilevel Models Using Stan. Journal of Statistical Software, 80(1), 1-28. doi:10.18637/jss.v080.i01

```
    sigma ~ tx),
    family=student, prior=c(pr0, pr1, pr2))
```

Bayesian Markov Chain Monte Carlo simulation will be run in four independent chains with 4000 iterations per chain. Diagnostics including trace plots and Rhat value will be used to check for convergence of posterior distributions.