SEMINAR 2 REPORT

Speaker:  
Dr. Jackson Barth, PhD  
Assistant Professor  
Department of Statistical Science  
Baylor University

Seminar Topic:  
A Meta-Analysis-Based Hierarchical Variance Model for Powering One and Two-Sample t-tests

Abstract:  
Sample size determination (SSD) is a critical component of statistical inference and hypothesis testing, directly influencing the accuracy and statistical power of the analysis. Dr. Barth introduced a novel SSD methodology for one- and two-sample t-tests, designed to ensure clinical relevance by utilizing a predetermined unstandardized effect size. The approach employs Bayesian meta-analysis to address uncertainty in variance estimation—an essential factor often challenging in SSD. By leveraging prior knowledge from related studies through a Bayesian gamma-inverse gamma model, an informative posterior predictive distribution for the variance is generated, resulting in more accurate and reliable sample size decisions.

Dr. Barth also introduced an empirical Bayes approach to improve computational efficiency in posterior sampling, which is further enhanced by a quantile simulation method. Simulations and empirical studies demonstrated the superiority of this method over traditional aggregation techniques—such as simple average, weighted average, and median—in variance estimation for SSD, particularly in meta-analyses with large disparities in sample sizes and moderate variance levels. Overall, this methodology provides a robust and practical solution for SSD in t-tests.

Key Assumptions and Considerations:

1. Addressing μ Assumptions:
   * Often considered as the minimal effect.
   * Same response may lead to different solutions.
2. Addressing σ Assumptions:
   * Typically guessed from previous studies, leading to potential power-hacking.
   * Simple average or weighted average methods may lose critical information about σ.
   * Each σ from related studies is assumed to be exchangeable, not identical, allowing for more reliable synthesis of information.

Literature Review:

1. *Harmonizing Frequentist & Bayesian Analysis* (2004)
2. Point estimation of unstandardized effect size (*What if* scenario analysis)

Proposed Methodology:

1. Fix μ to a particular significance level.
2. Collect information on σ².
3. Create a predictive posterior distribution for σ² using an empirical Bayes approach.
4. Perform simulations to validate the method.

Methodology for One-Sample t-test:

1. Hierarchy Construction  
   A hierarchical Bayesian structure is built to model the distribution of variance across studies.
2. Distribution of θ/y  
   The empirical Bayes approach is used to estimate α and β by maximizing the marginal likelihood:  
   P(y/α, β) = l(α, β|y), where ll = log likelihood. Empirical distance is then calculated.
3. Dealing with Trivial Distributions:  
   In some cases, when the data strongly point to a single estimate or the sample size is small, the distribution may become trivial, with α, β tending to infinity but their ratio β/α remaining stable. This can be simplified by using power.t.test in R, where θ\* = β/α is treated as the variance.

Power Calculation for Sample Size Determination:

1. Method: One-sided t-test.
2. Requirements:  
   a. A rejection region.  
   b. P(A) that t falls into the rejection region.  
   Power = 1 - β^t = P(T ≥ R).
3. Steps:
   * Generate M draws of θ.
   * For each draw, calculate P(T/θ ≥ R).
   * The average of these probabilities approximates the power.

Sample Size Selection:

1. Choose an initial value of n, then iterate backward or forward based on power calculations.
2. Use systematic sampling to reach approximately 80-90% power, though tail probabilities might be missed.

Handling α < 1 Cases:

* When α < 1, the data fit poorly into the inverse gamma distribution, and the variance becomes less stable. Variance is finite only when α > 2, and skewness is finite only when α > 3. As skewness decreases, α increases, showing an inverse relationship.

Comparison to Other Approaches:

* The proposed method is compared against simple average, weighted average, and median S² approaches. The new method showed distinct advantages, especially in cases with varying sample sizes across studies.

Validation and Application:

1. Real-World Data Validation:  
   The methodology was validated using two large data sets, consisting of 17 and 19 studies, respectively. The goal was to establish the accuracy of sample size estimation by mimicking the meta-analysis process.
2. Applications:  
   a. Unified Parkinson’s Disease Rating Scale (UPDRS): Measures disability levels in patients with Parkinson’s Disease. If UPDRS = 0, means no disability.

b. Cognitive Behavioural Therapy (CBT): Applied to meta-analyses of anxiety in children.

In summary, Dr. Barth’s seminar provided insightful advancements in sample size determination for t-tests, demonstrating the effectiveness of using Bayesian meta-analysis to improve variance estimation. His method offers a practical solution for researchers facing challenges in power analysis, particularly when handling variance uncertainty in small or disparate data sets.