ON THE SMALL SAMPLE BEHAVIOR OF GENERALIZED LINEAR MIXED MODELS WITH COMPLEX EXPERIMENTS

Julie Couton

Walt Stroup

ON THE SMALL SAMPLE BEHAVIOR OF GENERALIZED LINEAR MIXED MODELS WITH COMPLEX EXPERIMENTS

Julie Couton, Department of Statistics, University of Nebraska- Lincoln
Walt Stroup, Department of Statistics, University of Nebraska- Lincoln

ABSTRACT

Generalized linear mixed models (GLMMs), regardless of the software used to implement them (R, SAS, etc.), can be formulated as conditional or marginal models and can be computed using pseudo-likelihood, penalized quasi-likelihood, or integral approximation methods. While information exists about the small sample behavior of GLMMs for some cases- notably RCBDs with Binomial or count data- little is known about GLMMs for continuous proportions (e.g. Beta) or time-to-event (e.g. Gamma) data or for more complex designs such as the split-plot. In this presentation we review the major model formulation and estimation options and compare their small sample performance for cases listed above.

INTRODUCTION

Agricultural experiments commonly use blocked designs or designs with split plot features. Analysis of variance (ANOVA), which assumes normality, is the standard tool for analyzing these experiments. Frequently, response variables of interest are not normally distributed. Historically when confronted with non-normal data, statisticians either relied on the Central Limit Theorem as justification to assume ANOVA's robustness or when ANOVA assumptions were clearly violated, transformations became a potential option. As computational power improved and theory advanced, generalized linear models and generalized linear mixed models (GLMM) became practical alternatives. Statisticians now face a dilemma. Which method is appropriate and what are the costs of choosing an inappropriate one? Specifically, how do these methods compare with sample sizes typical of agricultural research?

This paper aims to shed light on the small sample behavior of untransformed and transformed ANOVA and applicable versions of the GLMM for analysis of complex experiments with non-normal data. Research, including Stroup (2013b), has examined the small sample behavior of analyses with Binomial, Poisson, and negative Binomial data. This paper extends Stroup (2013b) to include response variables assumed to have Gamma and Beta distributions, discussed in context of a randomized complete block and split-plot experiments.

The four resulting scenarios will be discussed including the process of simulation, methods of analysis and results. Several models are considered for data analysis including untransformed ANOVA, transformed ANOVA, and various forms of GLMM. The models will be compared using Type I error rate, power, average mean, average confidence interval bounds and coverage probability. Discussion will focus on performance relative to current GLMM conventional wisdom and on any apparent advantages or disadvantages.

As with Stroup (2013b), the intention of this paper is to demonstrate a method for performing realistic comparisons. Only a limited number of scenarios were considered for this research;

this paper is not meant to be exhaustive. Readers are encouraged to extend these methods to other scenarios involving the distributions considered here, where applicable to their research.

SCENARIOS CONSIDERED

We considered two different response variable distributions, the Gamma and Beta, and two designs, a randomized complete block design (RCBD) and split-plot, resulting in four total scenarios. Each distribution will be discussed with relevant background information, data generation, models considered for analysis, and methods of comparison.

Gamma Distribution

1. *Relevant Background*

The Gamma distribution is a right skewed, non-negative function; the probability density function (PDF) can take on a variety of forms depending on the shape and scale parameters. The PDF can be written in two ways: the first is the form typically shown in mathematical statistics textbooks, e.g. Casella and Berger (2002), and the second is a GLM-friendly form following McCullagh and Nelder (1989).

$$f(x|\alpha,\beta) = \frac{1}{\Gamma(\alpha)\beta^{\alpha}} x^{\alpha-1} e^{-\frac{x}{\beta}}, \qquad\qquad x,\alpha,\beta > 0 \quad mean = \alpha\beta \quad var = \alpha\beta^2 \qquad (1)$$

$$f(x|\mu,\varphi) = \frac{1}{\Gamma\left(\frac{\mu}{\varphi}\right)\varphi^{\frac{\mu}{\varphi}}} x^{\frac{\mu}{\varphi}-1} e^{-\frac{x}{\varphi}}, \qquad\qquad x,\mu,\varphi > 0 \quad mean = \mu \quad var = \mu\varphi \qquad (2)$$

Note: $\alpha = \mu/\varphi$ and $\beta = \varphi$. One common application of the Gamma distribution is to model time to event data. In addition, the chi-square distribution is a special case of the Gamma distribution. Because estimates of the variance use quadratic forms with approximate or exact chi-square distributions, the Gamma distribution is a common choice for modeling when variance is the response of primary interest.

2. *Randomized Complete Block Design*

The first scenario involves a randomized complete block design (RCBD) with 8 blocks and 2 treatments. Stroup (2013b) presents an approach to data generation in which the researcher specifies the sources of variation that affect the response mean and whether the sources affect the mean on the data scale or link scale. Stroup considered both data scale and link scale generation with count data. In this study, only data generation on the link scale was considered. Following the notation of Stroup we have

$$y_{ij}|b_j \sim Gamma(\mu_{ij},\varphi); \;\; link = \eta_{ij} = \log(\mu_{ij}) = \eta + \tau_i + {}_j, \;\; b_j \sim Gaussian(0,\sigma_b^2) \qquad (3)$$

Note: Since the Gamma is a two parameter distribution, having a unit level effect term (τb_{ij}) is not appropriate. Notice that (3) implies that μ_{ij} is perturbed by block and treatment on the link scale.

a. *Data Generation and Simulation*

Two cases were considered: one with equal and another with unequal treatment means. Figure 1 shows the associated PDFs of $y_{ij}|b_j$. One thousand experiments were generated using the following process:

- **Step 1:** 8 block effects $, b_2, \dots, b_8$ are generated from a $N(0, \sigma_b^2)$ distribution with $\sigma_b^2 = 0.2$.
- **Step 2:** Block-perturbed unit level η_{ij} were computed for each experimental unit. In the equal treatment case, $\mu = \mu_2 = 25$ and for the unequal $\mu = 25, \mu_2 = 35$ selected based on the example of flowering time given in Stroup (2013a). Then $\eta_{ij} = \log(\mu_i) + {}_j$ as follows from equation (3).
- **Step 3:** Means were computed as $\mu_{ij} = e^{\eta_{ij}}$
- **Step 4:** Rather than the equation (2) method of parameterization considered in this scenario, the SAS® rangam function uses traditional $\alpha\beta$ notation as in equation (1). Recall in equation (2) that $\alpha_{ij} = \mu_{ij}/\varphi$ and $\beta = \varphi$. φ (or β) $= 1.5$ and $\alpha_{ij} = e^{\eta_{ij}}/\varphi$.
- **Step 5:** Then observable values y_{ij} were computed using the rangam function in SAS® $y_{ij} = \varphi \cdot rangam(\alpha_{ij})$.

b. *Analysis Models*

Several analysis models could be used for a randomized complete block design with Gamma data. Four common options include:
- **Option 1:** Untransformed ANOVA. In this model, we assume that the sample mean has an approximate normal distribution and fit the model: $y_{ij} = \mu + \tau_i + b_j + e_{ij}$ where ${}_j \sim NI(0, \sigma_b^2), \ e_{ij} \sim NI(0, \sigma_e^2)$
- **Option 2:** Transformed ANOVA.
  - The log transformation, $log(y_{ij}) = \mu + \tau_i + {}_j + e_{ij}$ where ${}_j \sim NI(0, \sigma_b^2), \ e_{ij} \sim NI(0, \sigma_e^2)$
  - The square root transformation, $\sqrt{y_{ij}} = \mu + \tau_i + {}_j + e_{ij}$ where $b_j \sim NI(0, \sigma_b^2), e_{ij} \sim NI(0, \sigma_e^2)$
- **Option 3:** Basic GLMM. $y_{ij}|b_j \sim Gamma(\mu_{ij}, \varphi)$; link $\eta_{ij} = \log(\mu_{ij}) = \eta + \tau_i + {}_j$ where ${}_j \sim N(0, \sigma_b^2)$

  Many methods of estimation exist for this model. In this paper, we considered restricted subject specific pseudo-likelihood (RSPL) and two forms of integral approximation, quadrature and Laplace.

- **Option 4:** Compound symmetry model alternatively known as generalized estimating equation (GEE) or R-side model. $\begin{pmatrix} y_{\ j} \\ y_{2j} \end{pmatrix} \sim quasi - Gamma \left(\begin{bmatrix} \mu_{\ j} \\ \mu_{2j} \end{bmatrix}, \phi \begin{bmatrix} 1 & \rho \\ \rho & 1 \end{bmatrix} \right)$ link $\eta_i = log(\mu_i) = \eta + \tau_i$ where ϕ denotes the Gamma quasi-likelihood scale parameter and ρ denotes the working compound symmetry correlation coefficient.

As mentioned in Stroup (2013b), the basic GLMM (Option 3) targets the expected value of the conditional distribution of $y_{ij}|b_j$ i.e. μ_{ij}, whereas the untransformed ANOVA and GEE (Options 1 and 4 respectively) target the expected value of the marginal distribution of y_{ij}. Also note that the target of estimation with the transformed ANOVA (Option 2) considered here is neither μ_{ij} or the marginal mean. A more in depth discussion of the aforementioned statements can be found in Stroup (2013b).

3. *Split Plot Design*

The second scenario involves a split plot experiment with 12 blocks where A (the whole plot factor) and B (the subplot factor) both have two levels. Thus we have 4 treatment combinations: A_1B_1, A_1B_2, A_2B_1, A_2B_2. For data generation purposes, following Milliken and Johnson (2009) we can view data generation as a 4-treatment incomplete block design: $A_1B_1 =$ Treatment 1, $A_1B_2 =$ Treatment 2, $A_2B_1 =$ Treatment 3, and $A_2B_2 =$ Treatment 4. In 6 blocks, Treatments 1 and 2 appear together and in the other 6 blocks, Treatments 3 and 4 appear together. At this point, we can use the same generation model presented in equation (3) and simulate 1000 experiments following this process:

- **Step 1:** 12 block effects $_1, b_2, \ldots, b_{12}$ are generated from a $N(0, \sigma_b^2)$ distribution with $\sigma_b^2 = 0.2$.
- **Step 2:** Block-perturbed unit level η_{ij} were computed for each experimental unit. Means were set to be 25 for treatments 1 and 4 and 35 for treatments 2 and 3, chosen so that the average over factor A is the same (30) and the average over factor B is the same (30). This allows us to measure the Type I error rate for the main effects of A and B and power for the test of interaction. Then $\eta_{ij} = \log(\mu_i) + b_j$ as follows from equation (3).
- **Step 3:** Means were computed as $\mu_{ij} = e^{\eta_{ij}}$.
- **Step 4:** The scale parameter was set to $\varphi = 1.5$ thus $\alpha_{ij} = e^{\eta_{ij}}/\varphi$.
- **Step 5:** y_{ij} were computed using the rangam function in SAS; $y_{ij} = \varphi \cdot rangam(\alpha_{ij})$.

Analysis methods considered for the RCBD scenario were adapted for the split-plot.

Beta Distribution

1. *Relevant background*

The Beta distribution, commonly used for continuous proportions (values bounded between 0 and 1), has a PDF that can take on a variety of shapes depending on the values of the shape and scale parameters. The function can take on two parameterizations: one as in Casella and Berger (2002), and the second a form presented by Ferrari and Cribari-Neto (2004).

$$f(x|\alpha,\beta) = \frac{1}{B(\alpha,\beta)} x^{\alpha-1}(1-x)^{\beta-1}, \quad mean = \frac{\alpha}{\alpha+\beta} \quad var = \frac{\alpha\beta}{(\alpha+\beta)^2(\alpha+\beta+1)} \quad (4)$$

$$f(x|\mu,\varphi) = \frac{1}{B(\mu\varphi, \varphi(1-\mu))} x^{\mu\varphi-1}(1-x)^{\varphi(1-\mu)-1}, \quad mean = \mu \quad var = \frac{\mu(1-\mu)}{1+\varphi} \quad (5)$$

Note: $0 < \mu < 1, \varphi > 0, \alpha = \mu\varphi, \beta = \varphi(1-\mu)$ and $\varphi = \alpha + \beta$. The Beta distribution can also be obtained from a ratio Gamma random variables $\frac{\Gamma(\alpha,1)}{\Gamma(\alpha,1)+\Gamma(1,\beta)}$ resulting in a Beta(α, β). For consistency, the form presented in Cribari-Neto (2004) will be used for the duration of paper.

2. *Randomized Complete Block Design*

The third scenario is a randomized complete block design with 8 blocks and 2 treatments. Following the notation of Stroup (2013b) we have the following data generation model:

$$y_{ij}|b_j \sim Beta(\mu_{ij}, \varphi); \quad \text{link} = \eta_{ij} = \text{logit}(\mu_{ij}) = \eta + \tau_i + b_j, \quad b_j \sim Gaussian(0, \sigma_b^2) \qquad (6)$$

The Beta distribution is also a two parameter family, and thus having a unit level effect (τb_{ij}) is not appropriate. Notice that the proportions are perturbed by the block and treatment effects on the link scale as in generation model (3).

a. *Data Generation and Simulation*

Two cases were considered: one with equal and another with unequal treatment means. Figure 2 shows the associated PDFs; the combined distributions form a bathtub shape, a common occurrence with Beta data. One thousand experiments were generated using the following process:

- **Step 1:** 8 block effects $b_1, b_2, \ldots, b_8$ are generated from a $N(0, \sigma_b^2)$ distribution with $\sigma_b^2 = 0.15$.
- **Step 2:** Block-perturbed unit level η_{ij} were computed for each experimental unit. In the equal treatment case, $\mu_1 = \mu_2 = 0.2$ and for the unequal $\mu_1 = 0.2, \mu_2 = 0.8$. Then $\eta_{ij} = \text{logit}(\mu_i) + b_j$ as follows from (6).
- **Step 3:** Proportions were computed as $\mu_{ij} = \frac{1}{1+e^{-\eta_{ij}}}$.
- **Step 4:** For the equal treatment case $\alpha = 0.25, \beta = 1$ hence $\mu_1 = \mu_2 = \frac{0.25}{1+0.25} = 0.2$. For the unequal treatment case, $\alpha_1 = 0.25, \beta_1 = 1$ which implies $\mu_1 = 0.2$ and $\alpha_2 = 1, \beta_2 = 0.25$ which implies $\mu_2 = 0.8$. Note that this ensure $\varphi = 1.25$ in all cases satisfying a key GLIMMIX restriction that the scale parameter be constant over treatments. For data generation, β was set to the values shown above, where as α reflected the block effects on μ_{ij} that is $\mu_{ij} = \frac{\alpha_{ij}}{\alpha_{ij}+\beta}$ which implies $\alpha_{ij} = \frac{\mu_{ij}}{(1-\mu_{ij})}\beta$.
- **Step 5:** y_{ij} were computed using the SAS rangam function $y_{ij} = rangam(\alpha_{ij})/[rangam(\alpha_{ij}) + rangam(\beta)]$.

b. Analysis Models

Several analysis models could be used for a randomized complete block design with Beta data. Four common options include:

- **Option 1:** Untransformed ANOVA; in this model, we assume the sample proportion has an approximate normal distribution. $y_{ij} = \mu + \tau_i + b_j + e_{ij}$ where $b_j \sim NI(0, \sigma_b^2)$, $e_{ij} \sim NI(0, \sigma_e^2)$

- **Option 2:** Transformed ANOVA. $\sin^{-1}\sqrt{y_{ij}} = \mu + \tau_i + b_j + e_{ij}$ where $b_{ij} \sim NI(0, \sigma_b^2)$, $e_{ij} \sim NI(0, \sigma_e^2)$

- **Option 3:** Basic GLMM. $y_{ij}|b_j \sim Beta(\mu_{ij}, \varphi)$; link $\eta_{ij} = \text{logit}(\mu_{ij}) = \eta + \tau_i + b_j$ where $b_j \sim N(0, \sigma_b^2)$

    This paper considers restricted subject specific pseudo-likelihood (RSPL) and two forms of integral approximation, quadrature and Laplace for estimation methods.

- **Option 4:** Generalized estimating equation (GEE).
$$\binom{y_{1j}}{y_{2j}} \sim quasi - Beta\left(\begin{bmatrix}\mu_{1j}\\\mu_{2j}\end{bmatrix}, \phi\begin{bmatrix}1 & \rho\\\rho & 1\end{bmatrix}\right) \quad \text{link } \eta_i = logit(\mu_i) = \eta + \tau_i \text{ where } \phi$$
denotes the Gamma quasi-likelihood scale parameter and ρ denotes the working compound symmetry correlation coefficient.

The target of estimation for each option here is synonymous with those mentioned in the Gamma section of this paper.

3. *Split Plot Design*

The fourth scenario involves a split plot experiment with 12 blocks where A (the whole plot factor) and B (the subplot factor) both have two levels. The four treatment combinations will be considered as a 4 treatment incomplete block design, as mentioned above. The same generation model presented in equation (6) and 1000 experiments will be simulated:

- **Step 1:** 12 block effects $b_1, b_2, \ldots, b_{12}$ are generated from a $N(0, \sigma_b^2)$ distribution with $\sigma_b^2 = 0.15$.

- **Step 2:** Block-perturbed unit level η_{ij} were computed for each experimental unit. Proportions were set to be 0.2 for treatments 1 and 4 and 0.8 for treatments 2 and 3. Then $\eta_{ij} = \text{logit}(\mu_i) + b_j$ as follows from (6).

- **Step 3:** Proportions were computed as $\mu_{ij} = \frac{1}{1+e^{-\eta_{ij}}}$.

- **Step 4:** For treatments 1 and 4, $\alpha_1 = 0.25$, $\beta_1 = 1$ which implies $\mu_1 = \mu_4 = 0.2$; for treatments 2 and 3 $\alpha_2 = 1$, $\beta_2 = 0.25$ which implies $\mu_2 = \mu_3 = 0.8$. This ensures

$\varphi = 1.25$ in all cases satisfying a key GLIMMIX restriction that the scale parameter be constant over treatments. For data generation, β was set to the values shown above, where as α reflected the block effects on μ_{ij} that is $\mu_{ij} = \frac{\alpha_{ij}}{\alpha_{ij}+\beta}$ which implies

$$\alpha_{ij} = \frac{\mu_{ij}}{1-\mu_{ij}}\beta.$$

- **Step 5:** y_{ij} were computed using the SAS rangam function
$$y_{ij} = rangam(\alpha)/(rangam(\alpha) + rangam(\beta)).$$

Analysis methods considered for the RCBD scenario were adapted for the split-plot.

RESULTS AND DISCUSSION

Stroup (2013b) listed several items widely assumed to be true about ANOVA with non-normal data and about GLMMs. Stroup referred to these as articles of conventional wisdom, and they are often based on limited information or special cases. Relevant to this paper, little information exists about Gamma or Beta response variables. Following Stroup, simulation results here are discussed in the context of information they provide to support or contradict conventional wisdom.

The expected Type I error rate in the equal treatment case is 0.05 if α = 0.05. Using standard margin of error methods, with 1000 simulated experiments we would expect to see rejection rates between 0.03 and 0.07; if rejection rates falls into this range, Type I error rate is considered to be adequately controlled for the associated method. Among the methods that adequately controlled Type I error, power loss is calculated by comparing the power of a given analysis method to the analysis method with the highest power; this loss in efficiency is reported as a percentage. Nominal confidence interval coverage is considered to be between 0.93 and 0.97 again using standard margin of error methods with 1000 simulated experiments. Reported lower confidence interval limits are the average of the lower limits found in the 1000 simulated experiments; similarly for the reported means and upper limits. For cases considering equal treatments, reported coverage probabilities are the average of the resulting coverage probabilities for the two treatments; due to the similarity in results, reporting both values is extraneous. For the split-plot (SP) experiments 2 of 4 treatments were set to the same value (e.g. treatments 1 and 4 were set to the same value and treatments 2 and 3 were set to the same value). One coverage probability is reported for treatments 1 and 4, and is calculated by averaging the coverage probabilities for each of the respective treatments individually; similarly for treatments 2 and 3. All of the following conventional wisdoms reference Figures 3-22.

- **Conventional Wisdom Item 1 (CW1):** Standard ANOVA should provide sufficient Type I error control and confidence interval coverage in the equal treatment. If there are problems, they will arise in the unequal treatment case.

Standard ANOVA for the Gamma RCBD scenario showed acceptable Type I error and a slight power loss (4.8%). Confidence interval coverage was nominal in the equal treatment case, however with unequal treatments coverage probability fell outside of the nominal range (0.985)

for treatment 1; also confidence intervals are clearly targeting the marginal mean. With the Gamma SP, all coverage probabilities fell outside of the nominal range (0.978 for treatments 1 and 4, 0.911 for treatments 2 and 3), also intervals are targeting the marginal mean. For the RCBD Beta Type I error is acceptable and showed minimal power loss (1.5%). All coverage probabilities fell within the nominal range, however with the equal treatment case the lower confidence bound was -0.016. With the unequal treatment case, the lower limit for treatment 1 fell below 0 (-0.014) and the upper limit for treatment 2 rose above 1 (1.019). For the Beta SP, Type I error falls in the acceptable range and only slight power losses occur for determining simple effects (2-3%). Coverage probabilities fall in the nominal range, but confidence interval bounds fall below zero and above one as in the RCBD scenario. In summary, the principal concern with the equal treatment case was interval estimates outside of the plausible range; the majority of issues arise with the unequal treatment case.

- **CW2:** Transformations should help to stabilize the variance issues that arise with standard ANOVA in the unequal treatment case.

For RCBD Gamma, both transformations (log and square root) have Type I errors within the acceptable range.   Both the log and square root show power losses of 7.1% and 3.3% respectively. Coverage probability for the log fall within the nominal range, but treatment 1 for the square root has a coverage probability that falls outside of the nominal range (0.976). Preliminary study of the transformations for the SP Gamma shows that the power losses and coverage probability issues that surface in the RCBD were amplified, and thus results will not be discussed. For the RCBD Beta the arcsine square root transformation shows acceptable Type I error and a slight power advantage. However coverage probability in the equal treatment case falls outside of the nominal range (0.904); in the unequal treatment case coverage probability falls outside of the nominal range for both treatments (0.911, 0.911). Due to the aforementioned issues with the arcsine square root transformation in the RCBD Beta scenario, this transformation was not considered for analysis in the SP Beta scenario. All interval estimates with the Beta distribution in both scenarios fall within the range of the data (0,1).

- **CW3:** GEEs are robust to model misspecification.

With the RCBD Gamma, the GEE has acceptable Type I error rate and had the largest power of all estimation methods considered. Coverage probabilities for both the equal and unequal treatment cases fall in the nominal range. It is evident that the GEE is estimating the marginal means. Figure 1 shows the PDFs of the treatments considered for the RCBD Gamma case. Because the GEE is targeting the marginal means, a power advantage is expected due to the fact that the marginal means are spread apart more than the conditional means. The GEE has acceptable Type I Error with the SP Gamma; it shows a slight power loss with determining the simple effects of factor A (4.2%). With the RCBD Beta, the GEE has acceptable Type I error but shows a power loss (11.0%). Coverage probabilities fall out of the nominal range in the equal treatment case (0.978) and unequal case (0.979, 0.982). A power loss is expected; the GEE is estimating the marginal means, which in this case are closer together than the conditional means (Figure 2). The SP Beta reveals that the Type I error rate for the main effect of factor A fall outside of the acceptable range (0.0202) which may explain the power loss present for the simple effects of A (15.8%) and for the simple effects of B (18.3%).

- **CW4:** Pseudo-likelihood has accuracy issues with certain GLMMs and quadrature should help address this.

For the Gamma distribution scenarios, all Type I error rates fall in the acceptable range and coverage probabilities are nominal. Only slight power losses occur (4.2% in RCBD and 4.9% in SP for simple effect of factor A). RSPL appears to have variance component estimation issues with the Beta distribution. For RCBD Beta, RSPL has an inflated Type I error rate (0.1024) and thus was discounted from other analyses including power and confidence interval coverage. For SP Beta, RSPL had similar issues with Type I error with both the main effect of factor A (0.0157) and main effect of factor B (0.0741) following outside of the acceptable range.

- **CW5:** Quadrature is preferable to Laplace.

In almost all instances, quadrature and Laplace had the same results; any estimates that differed were within 0.001 of each other. Due to this equivalence, Laplace results will be reported. Type I error for RCBD Gamma falls within the acceptable range and power loss of 3.2%. All coverage probabilities are nominal. For SP Gamma, Type I error falls outside of the nominal range for the main effect of factor A (0.0720). There appears to be a slight power advantage for the simple effect of factor A (4.2%) however this may be due to the inflated Type I error. Coverage probability falls outside of the nominal range (0.927 for Treatments 1 and 4). For the Beta distribution with both the RCBD and SP all Type I Errors fall within the acceptable range. Coverage probabilities are nominal with the RCBD but with the SP fall outside of the nominal range for treatments 1 and 4 (0.924).

CONCLUSIONS:

This research focuses only on 1 set of specified means or proportions; before firm conclusions can be determined, the study should be expanded to include additional scenarios with varying means and proportions.

CW1: Preliminary findings of this experiment suggest that CW1 was accurate in assuming that issues would arise when treatments were unequal, however consequences also occur with the equal treatment case (e.g. interval estimates below 0). If the interval were truncated to take this issue into account, it is unclear if the interval is truly estimating a 95% confidence interval.

CW2: Transformations did not clearly improve issues that arose with standard ANOVA in the unequal cases. With the Gamma transformations, power loss and coverage probability issues occurred as with the standard ANOVA. All interval estimates did fall within the plausible range (e.g. between 0 and 1 with Beta), however coverage probability was a large issue with the arcsine square root. Finally, with standard ANOVA, it was apparent that procedures were estimating the marginal mean; with the transformations, it is unclear which parameter (marginal or conditional mean) the models are estimating. A clear advantage is not apparent with the transformations.

CW3: It is imperative to reiterate that the GEE is targeting the marginal means. It certain circumstances, this could result in a power advantage (e.g. RCBD Gamma) or a power loss (e.g. RCBD Beta). Power and coverage probability issues increased as the structure of the experiment

increased. Potentially the GEE may be sufficient with simple designs, but major issues appeared as the experiment grew in complexity.

CW4: There does not appear to be a major difference between RSPL and quadrature with the Gamma distribution scenarios. However, RSPL did have issues with the Beta distribution. Further study is needed to determine if the issue stems from design complications or from the distribution itself (potentially values near the extremes). Quadrature did perform better with the Beta scenarios; Type I error was not an issue as it was with RSPL. Quadrature and Laplace did have issues with coverage probability in the SP Beta case. Further investigation may reveal that a form of pseudo-likelihood may be preferred, but with this study it is unclear.

CW5: There were no meaningful differences between quadrature and Laplace; in almost all instances, results were identical. Any differences that occurred were slight. Computing time was much faster for Laplace and further study may reveal that it is a suitable replacement for quadrature.

After reviewing the conventional wisdoms hopefully it has become apparent to the reader that rarely is a "one size fits all" methodology appropriate. All methods studied produced irregular results, some in multiple instances. Thus, before any model can be chosen to analyze the data it is the obligation of the researcher to first consider what process is giving rise to the data. Potentially a simulation study can be a first step in assessing power or estimation accuracy as a way to determine which model or method is appropriate.

Further investigations need to evaluate the effects of different mean values (or proportions), specified scale parameters, or larger sample sizes in order to determine if the results found in this study are pervasive.

REFERENCES

Casella, G. and R. Berger. 2002. *Statistical Inference*, *2nd ed.* Pacific Grove, CA. Duxbury.

Ferrari, S. and F. Cribari-Neto. 2004. Beta regression for modeling rates and proportions. *J. Applied Statist.* **31**: 799-815.

McCullagh, P. and J.A Nelder. 1989. *Generalized Linear Models, 2nd ed.* London: Chapman & Hall.

Milliken, G.A. and D.E. Johnson. 2009. *Analysis of Messy Data, Vol. 1, 2nd Ed.* New York: Chapman and Hall.

Stroup, W.W. 2013a. *Generalized Linear Mixed Models.* Boca Raton, FL. CRC Press.

Stroup, W.W. 2013b. Non-normal data in agricultural experiments. In *Proceedings of the 25th Conference on Applied Statistics in Agriculture.* Manhattan, KS: Kansas State University Department of Statistics, in text.
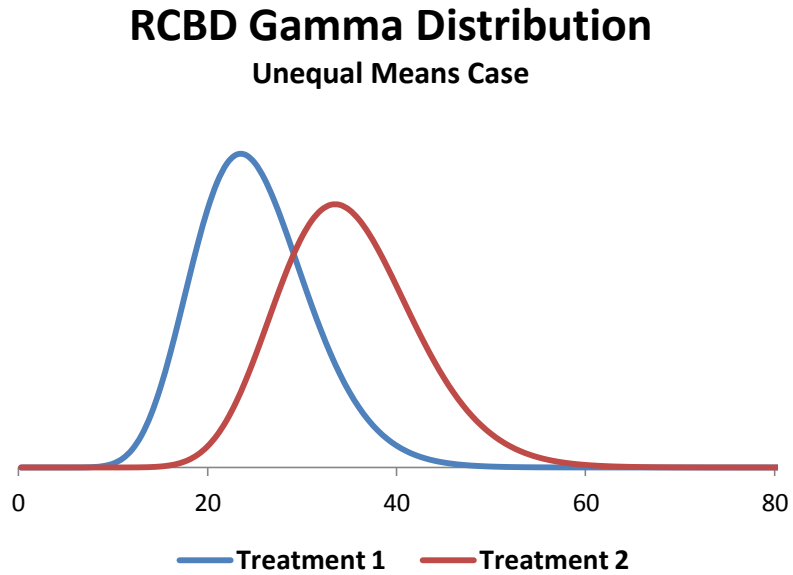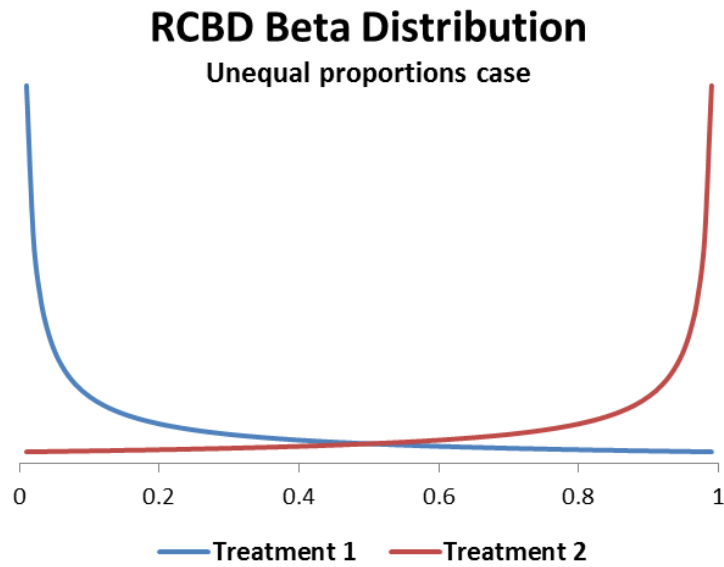
APPENDIX

Figure 1:

RCBD Gamma Distribution
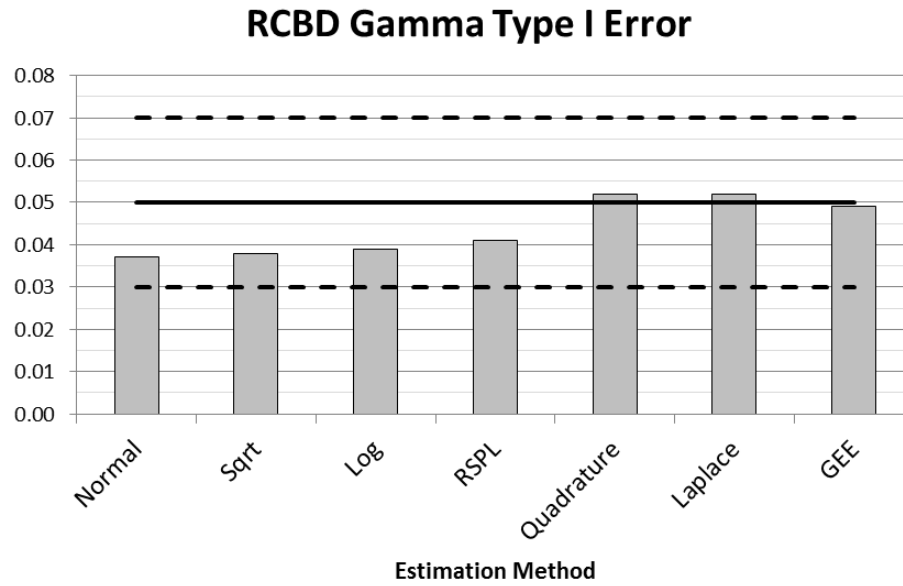### Unequal Means Case

Figure 2:

RCBD Beta Distribution
Unequal proportions case

Figure 3:

RCBD Gamma Type I Error

Figure 4:

RCBD Gamma Power

Figure 5:

RCBD Gamma- Equal Treatments Case

Figure 6:

RCBD Gamma- Unequal Treatments Case

Figure 7:

Split Plot Gamma- A Main Effect

Figure 8:

Split Plot Gamma- B Main Effect

Figure 9:

Split Plot Gamma- A x B Interaction

Figure 10:

Split Plot Gamma- Simple Effects of A

Figure 11:

Split Plot Gamma- Simple Effects of B

Figure 12:

Split Plot Gamma

Figure 13:

RCBD Beta Type I Error

Figure 14:

RCBD Beta Power

Figure 15:

RCBD Beta- Equal Treatments Case

Figure 16:

RCBD Beta- Unequal Treatment Case

Figure 17:

Split Plot Beta- A Main Effect

Figure 18:

Split Plot Beta- B Main Effect

Figure 19:

Split Plot Beta- A x B Interaction

Figure 20:

Split Plot Beta- Simple Effects of A

Figure 21:

Split Plot Beta- Simple Effects of B

Figure 22:

Split Plot Beta