One-sided tolerance interval in a two-way balanced nested model with mixed effects

Adilson de Jesus Martins da Silva
University of Cape Verde*(Uni-CV)
Praia - Cape Verde

PhD Student & Researcher at CMA
New University of Lisbon† Portugal
Mathematics - Theoretical Statistic

April 13, 2013

In this paper we approach the construction of the both upper and lower tolerance limit in a two-way nested model with mixed effects in balanced data. In order to do so we proceed as Fonseca et al [3] did in order to derive the upper tolerance limit in a two-way nested model with mixed effects in unbalanced data, by using the generalized confidence interval idea earlier used by Krishnamoorthy and Mathew [4] to perform the construction of the upper tolerance limit in a one-way nested model with mixed or random effects model in balanced and unbalanced data. The underlying idea goes through the construction of an approximation for the quantile of the general pivotal quantity for a convinient parametric function.

Keywords: *Mixed model; Balanced data; Upper tolerance limit; Lower tolerance limit; Confidence limit; Generalized pivotal quantity.*

1 Introduction

In many research areas (such as public health, environmental contamination, and others) one deals with the necessity of using data to infer whether some proportion (%) of a population of interest is (or one wants it to be) below and/or over some threshold, through the computation of *tolerance interval*. The idea is, once a threshold is given, one computes the *tolerance interval* or *limit* (which might be one or two - sided bounded) and then to check if it satisfies the given threshold.

Since in this work we deal with the computation of one - sided *tolerance interval*, for the two-sided case we recomend, for instance, Krishnamoorthy and Mathew [5].

Krishnamoorthy and Mathew [4] performed the computation of upper *tolerance limit* in balanced and unbalanced one-way random effects models, whereas Fonseca

*Corresp. e-mail: adilson.dasilva@docente.unicv.edu.cv
†Corresp. e-mail: ad.silva@campus.fct.unl.pt

et al [3] performed it based in a similar ideas but in a tow-way nested mixed or random effects model. In case of random effects model, Fonseca et al [3] performed the computation of such interval only for the balanced data, whereas in the mixed effects case they dit it only for the unbalanced data. For the computation of two-sided *tolerance interval* in models with mixed and/or random effects we recomend, for instance, Sharma and Mathew [7].

The purpose of this paper is the computation of upper and lower *tolerance interval* in a two-way nested mixed effects models in balanced data. For the case of unbalanced data, as mentioned above, Fonseca et al [3] have already computed upper *tolerance interval.* Hence, using the notions persented in Fonseca et al [3] and Krishnamoorthy and Mathew [4], we present some results on the construction of one-sided *tolerance interval* for the balanced case. Thus, in order to do so at first instance we perform the construction for the upper case, and then the construction for the lower case.

2 Nested Model Design - Basics Notions

A statistical model is said to have a *mixed effects* if it consists of a mixture of fixed and random effects factors. Such a model is said to be a two-way nested one, if it consists of two factors, say A and B, where levels of factor B are nested within levels of the factor A. There is many publications abording nested models. For example, in Ukaegbu and Smaila [8], the authors deal with the problem of performing the ANOVA for the thee-way nested model, i.e., with thee factors nested. More over, for complete notion about nested models we suggest Montegomery [6] and/or Dowdy and Chilko [2].

The effects associated with any factor in a nested model are the effects which the levels have on the interest response variable.

Let suppose that the factor B has $b_i, i = 1, ..., a$, levels nested within the *ith* level of the factor A. Thus, the *two-way nested mixed effects model* is given by

$$Y_{ijk} = \mu + \tau_i + \beta_{j(i)} + \epsilon_{k(ij)},\ k = 1, \dots, n_{ij},\ j = 1, \dots, b_i, \tag{1}$$

where μ is the general mean, τ_i is considered to be the fixed effect term due to the *ith*, $i = 1, ..., a$, level of the factor A, $\beta_{j(i)}$, $j = 1, ..., b_i$, the random effect term due to the *jth* level of the factor B nested within the *ith* level of the factor A, and $\epsilon_{k(ij)}$ the error term associated to the observed value Y_{ijk}. It is assumed that $\beta_{j(i)} \sim N(0, \sigma^2_\beta)$, and $\epsilon_{k(ij)} \sim N(0, \sigma^2_\epsilon)$ are independent from each other.

Recalling that our model considers the data to be balanced, that is, $b_i = b$, $i = 1, \dots, a$, and $n_{ij} = n$, $i = 1, \dots, a$, $j = 1, \dots, b$, and letting $\bar{Y}_{ij\bullet} = \frac{1}{n}\sum_{k=1}^{n} Y_{ijk}$, the sums of squares are given by

$$SS_\beta = \sum_{i=1}^{a}\sum_{j=1}^{b}\left(\bar{Y}_{ij\bullet} - \frac{1}{b}\sum_{j=1}^{b}\bar{Y}_{ij\bullet}\right)^2 \text{ and } SS_\epsilon = \sum_{i=1}^{a}\sum_{j=1}^{b}\sum_{k=1}^{n}(Y_{ijk} - \bar{Y}_{ij\bullet})^2.$$

Now ($\bar{Y}_{ij\bullet}$, SS_β and SS_ϵ are independent distributed variables.) we define the independent variables

$$U_\epsilon = \frac{SS_\epsilon}{\sigma_\epsilon^2} \text{ and } U_\beta = \frac{SS_\beta}{\sigma_\beta^2 + n^{-1}\sigma_\epsilon^2}.$$

Thus (a generalization of the Theorem 5.3.1 of Casella and Berger [1] together with Fonseca et al [3] may be very useful here!)

$$U_\beta \sim \chi^2_{a(b-1)} \text{ and } U_\epsilon \sim \chi^2_{ab(n-1)},$$

with χ_r^2 the chi-square distribution with r degrees of freedom.

In this paper we approach the two following problems:

1) The construction of both upper and lower *tolerance limit* for the observable random variable Y, where $Y \sim N(\mu_i, \sigma_\beta^2 + \sigma_\epsilon^2)$, with $\mu_i = \mu + \tau_i$;

2) The construction of both upper and lower *tolerance limit* for the *unobserved* "true effect" $Y^* = \mu_i + \tau_i + \beta_{j(i)} \sim N(\mu_i, \sigma_\beta^2)$.

The next section is addressed to the construction of upper *tolerance limit* for both cases $Y \sim N(\mu_i, \sigma_\beta^2 + \sigma_\epsilon^2)$ and $Y^* = \mu_i + \tau_i + \beta_{j(i)} \sim N(\mu_i, \sigma_\beta^2)$. The first one is the subsection of 3.1, and the later one of subsection 3.2.

3 Upper Tolerance Limit

Let $\mathbf{Y} = \{Y_{ij1}, \ldots, Y_{ijk}\}$ be a sample of the random variable $Y \sim N(\mu_i, \sigma_\beta^2 + \sigma_\epsilon^2)$. A statistic C is a (p, γ)-upper *tolerance limit* for Y, if the equation

$$P_{\mathbf{Y}}[P_Y(Y \le C \| \mathbf{Y}) \ge p] = \gamma \Longleftrightarrow P_{\mathbf{Y}}[q_p \le C] = \gamma, \tag{2}$$

holds for $0 < \gamma < 1,$ and $0 < p < 1$, where q_p is the *pth* quantile of $N(\mu_i, \sigma_\beta^2 + \sigma_\epsilon^2)$. Thus, clearly, C is also a γ-upper *confidence limit* for the quantile q_p. More over, C is a γ-upper *confidence limit* for the parametric function

$$\mu_i + z_p\left(\sqrt{\sigma_\beta^2 + \sigma_\epsilon^2}\right),$$

where z_p denotes the *pth* quantile of the $N(0, 1)$ (the standard normal distribution), as we are about to show:

$$\begin{aligned}
&P_{\mathbf{Y}}[P_Y(Y \le C \| \mathbf{Y}) \ge p] = \gamma \Longleftrightarrow \\
&P_{\mathbf{Y}}\left[P_Y\left(\frac{Y - \mu_i}{\sqrt{\sigma_\beta^2 + \sigma_\epsilon^2}} \le \frac{C - \mu_i}{\sqrt{\sigma_\beta^2 + \sigma_\epsilon^2}} \| \mathbf{Y}\right) \ge p\right] = \gamma \Longleftrightarrow \\
&P_{\mathbf{Y}}\left[z_p \le \frac{C - \mu_i}{\sqrt{\sigma_\beta^2 + \sigma_\epsilon^2}}\right] = \gamma \Longleftrightarrow \\
&P_{\mathbf{Y}}\left[\mu_i + z_p\left(\sqrt{\sigma_\beta^2 + \sigma_\epsilon^2}\right) \le C\right] = \gamma,
\end{aligned}$$

where z_p is the standard normal distribution *ith* quantile, since $\frac{Y-\mu_i}{\sqrt{\sigma_\beta^2+\sigma_\epsilon^2}} \sim N(0,1)$.

Now, with out lost of generality, let $\mathbf{Y} = \{Y_{ij1}, \ldots, Y_{ijk}\}$ be a sample of the random variable $Y^* \sim N(\mu_i, \sigma_\beta^2)$. Then, proceeding identically to the case of $Y \sim N(\mu_i, \sigma_\beta^2 + \sigma_\epsilon^2)$, since the unobserved "true effect" variable $Y^* \sim N(\mu_i, \sigma_\beta^2)$, the

(p, γ)-upper *tolerance limit* for Y^*

is simply the

γ-upper *confidence limit* for the parametric function $\mu_i + z_p\sigma_\beta$.

3.1 The Observable Random Variable $Y \sim N(\mu_i, \sigma_\beta^2 + \sigma_\epsilon^2)$

Here our concern is the the construction of the upper *tolerance limit* for $Y \sim N(\mu_i, \sigma_\beta^2 + \sigma_\epsilon^2)$.

Let $W_i = \hat{\mu_i} = \frac{1}{b}\sum_{j=1}^{b} \bar{Y}_{ij\bullet}$. Then easy computation shows that

$$W_i \sim N\left(\mu_i, \frac{\sigma_\beta^2 + n^{-1}\sigma_\epsilon^2}{b}\right),$$

and

$$Z = \frac{\sqrt{b}(W_i - \mu_i)}{\sqrt{\sigma_\beta^2 + n^{-1}\sigma_\epsilon^2}} \sim N(0,1).$$

Since a (p, γ)-upper *tolerance limit* for $Y \sim N(\mu_i, \sigma_\beta^2 + \sigma_\epsilon^2)$ is simply a γ-upper *confidence limit* for the parametric function $\mu_i + z_p\left(\sqrt{\sigma_\beta^2 + \sigma_\epsilon^2}\right)$, with z_p the *pth* quantile of the standard normal distribution, we consctruct an aproximation for the *γth* quantile, say $L_{1p}(\gamma)$, for the *general pivotal quantity*, say L_{1p}, for the function $\mu_i + z_p\left(\sqrt{\sigma_\beta^2 + \sigma_\epsilon^2}\right)$ to obtain an approximation for such γ-upper *confidence limit* and, consequentely, for such a (p, γ)-upper *tolerance limit* for Y.

Definition 3.1. Let $x = \{x_1, \ldots, x_n\}$ be an observed value of any random vector $X = \{X_1, \ldots, X_n\}$, $n > 0$, whose distribution depends on the parameter of interest θ and a nuisance parameter η. A random function $T(X; x, \theta, \eta)$, which is a function of the random vector X, its observed value x, and the parameters θ and η, is said to be a *generalized pivotal quantity* for the parameter of interest θ, if it satisfies the following two conditions:

- For fixed value of x, the distribution of $T(X; x, \theta, \eta)$ is free of unknown parameters.
- The observed value of $T(X; x, \theta, \eta)$, which is obtained by replacing X with its observed value x, is simply the parameter of interest θ.

The definition of a *general pivotal quantity* we presented above may be founded (in a slightly different way) at Krishnamoorthy and Mathew [5].

Let w_i, ss_β, and ss_ϵ be the observable values of the random variables W_i, SS_β, and SS_ϵ, respectively (or $(w_i, ss_\beta, ss_\epsilon)$ the observed vector of the random vector $(W_i, SS_\beta, SS_\epsilon)$). Now, let L_{1p} be given by

$$\begin{aligned}
L_{1p} &= w_i - \frac{\sqrt{b}(W_i-\mu_i)}{\sqrt{SS_\beta}}\sqrt{\frac{ss_\beta}{b}} + z_p\left[\frac{\left(\sigma_\beta^2+n^{-1}\sigma_\epsilon^2\right)}{SS_\beta}ss_\beta + \left(1-n^{-1}\right)\frac{\sigma_\epsilon^2}{SS_\epsilon}ss_\epsilon\right]^{\frac{1}{2}} \qquad (3)\\
&= w_i - \frac{\sqrt{b}(W_i-\mu_i)}{\sqrt{SS_\beta}}\sqrt{\frac{ss_\beta}{b}} + z_p\left[\frac{ss_\beta}{U_\beta} + \left(1-n^{-1}\right)\frac{ss_\epsilon}{U_\epsilon}\right]^{\frac{1}{2}} \qquad (4)\\
&= w_i + \sqrt{\frac{ss_\beta}{ab(b-1)}}\left[\frac{-Z + z_p\left(b + a(b-1)\frac{ss_\epsilon}{ss_\beta}\frac{[U_\beta/a(b-1)]}{[U_\epsilon/b(1-n^{-1})]}\right)^{\frac{1}{2}}}{\sqrt{U_\beta/a(b-1)}}\right] \qquad (5)\\
&= w_i + \sqrt{\frac{ss_\beta}{ab(b-1)}}\left[\frac{-Z + z_p\left(b + \frac{a(b-1)}{ab(n-1)}\frac{ss_\epsilon}{ss_\beta}\frac{[U_\beta/a(b-1)]b\left(1-n^{-1}\right)}{[U_\epsilon/ab(n-1)]}\right)^{\frac{1}{2}}}{\sqrt{U_\beta/a(b-1)}}\right] \qquad (6)\\
&= w_i + \sqrt{\frac{ss_\beta}{ab(b-1)}}\left[\frac{Z + z_p\left(b + \frac{(b-1)\left(1-n^{-1}\right)}{(n-1)}\frac{ss_\epsilon}{ss_\beta}F\right)^{\frac{1}{2}}}{\sqrt{U_\beta/a(b-1)}}\right] \qquad (7)\\
&\approx_d w_i + \sqrt{\frac{ss_\beta}{ab(b-1)}}\left[\frac{Z + z_p\left(b + \frac{(b-1)\left(1-n^{-1}\right)}{(n-1)}\frac{ss_\epsilon}{ss_\beta}F_{a(b-1);ab(n-1)}(\theta)\right)^{\frac{1}{2}}}{\sqrt{U_\beta/a(b-1)}}\right] \qquad (8)
\end{aligned}$$

where $F = \frac{[U_\beta/a(b-1)]}{[U_\epsilon/ab(n-1)]}$ has a F (Fisher) distribution with $(a(b-1), ab(n-1))$ degrees of freedon (df), $\theta = 1-\gamma$ and $F_{r;s}(m)$ denotes the *mth* quantile of the F distribution with (r,s) *df*. Looking at the equation (3) one sees that replacing the random variables W_i, SS_β, and SS_ϵ with their observed values w_i, ss_β, and ss_ϵ, respectively, the resulting value is precisely $\mu_i + z_p\left(\sqrt{\sigma_\beta^2+\sigma_\epsilon^2}\right)$, that is, the parameter of the interest, and looking equation (4) one sees that given a sample (w_i, ss_β, and ss_ϵ are fixed) L_{1p} is free of unknown parameters. So we just showed that L_{1p} is a *general pivotal quantity.*

Remark 3.1. Let $M \sim N(0,1)$, and $V = -M$. Then, since

$$E(V) = -E(M) = 0 \text{ and } Var(V) = (-1)^2 Var(M) = 1,$$

$$V \sim N(0,1).$$

This **Remark** justifies the transition from the equation (6) to the equation (7).

Remark 3.2. The transition from the equation (7) to the approximation (8) is due to the replacement of the F distribution with its $(1-\gamma)th$ quantile. Indeed, according to Krishnamoorthy and Mathew [5], the idea is to replace F with a suitable quantile and for that, the other option would be its γth quantile, but by doing so, the γth quantile of the term

$$\sqrt{\frac{ss_\beta}{ab(b-1)}}\left[\frac{Z + z_p\left(b + \frac{(b-1)\left(1-n^{-1}\right)}{(n-1)}\frac{ss_\epsilon}{ss_\beta}F_{a(b-1);ab(n-1)}(\theta)\right)^{\frac{1}{2}}}{\sqrt{U_\beta/a(b-1)}}\right]$$

with $F_{a(b-1);ab(n-1)}(\theta)$ replaced with $F_{a(b-1);ab(n-1)}(\gamma)$ is much larger than the γth quantile of the original random variable

$$\sqrt{\frac{ss_\beta}{ab(b-1)}}\left[\frac{Z + z_p\left(b + \frac{(b-1)\left(1-n^{-1}\right)}{(n-1)}\frac{ss_\epsilon}{ss_\beta}F\right)^{\frac{1}{2}}}{\sqrt{U_\beta/a(b-1)}}\right].$$

Thus, as we see, it is necessary to replace F with a quantity smaller than $F_{a(b-1);ab(n-1)}(\gamma)$. Furthermore, according to Krishnamoorthy and Mathew [5], simulations show that $F_{a(b-1);ab(n-1)}(\theta)$ is the appropriate choice.

Now, once $Z \sim N(0,1)$, and $U_\beta \sim \chi^2_{a(b-1)}$, the term

$$T_{1p} = \left[\frac{Z + z_p\left(b + \frac{(b-1)\left(1-n^{-1}\right)}{(n-1)}\frac{ss_\epsilon}{ss_\beta}F_{a(b-1);ab(n-1)}(\theta)\right)^{\frac{1}{2}}}{\sqrt{U_\beta/a(b-1)}}\right]$$

(in approximation (8) of the L_{1p} development) has a noncentral t (student) distribution with $a(b-1)$ *df* and the noncentrality parameter $z_p\left(b + \frac{(b-1)\left(1-n^{-1}\right)}{(n-1)}\frac{ss_\epsilon}{ss_\beta}F_{a(b-1);ab(n-1)}(\theta)\right)^{1/2}$. Then, an approximation for the γth quantile of L_{1p}, say $L_{1p}(\gamma)$, is

$$L_{1p}(\gamma) = w_i + t_{a(b-1);\gamma}(\delta_1)\sqrt{\frac{ss_\beta}{ab(b-1)}}, \tag{9}$$

with

$$\delta_1 = z_p\left(b + \frac{(b-1)\left(1-n^{-1}\right)}{(n-1)}\frac{ss_\epsilon}{ss_\beta}F_{a(b-1);ab(n-1)}(\theta)\right)^{1/2}, \tag{10}$$

where $t_{r;s}(m)$ denotes the *sth* quantile of a noncentral t distribution with r *df*, and the noncentrality parameter m.

3.2 The Unobserved "True Effects" Variable $Y^* \sim N(\mu_i, \sigma_\beta^2)$

Now our concern is the construction of an upper *tolerance limit* for the unobserved "true effect" $Y^* = \mu_i + \tau_i + \beta_{j(i)} \sim N(\mu_i, \sigma_\beta^2)$.

As proven in section 3, it is easily shown that the (p, γ)-upper *tolerance limit* for Y^* is simply the γ-upper *confidence limit* for the parametric function $\mu_i + z_p\sigma_\beta$, with z_p the *pth* quantile of the standard normal distribution. Hence, we consider the following *general pivotal quantity* for the parametric function $\mu_i + z_p\sigma_\beta$:

$$
\begin{aligned}
L_{2p} &= w_i - \sqrt{b}\left(\frac{W_i - \mu_i}{\sqrt{SS_\beta}}\right)\sqrt{\frac{ss_\beta}{b}} + z_p\left[\frac{\sigma_\beta^2 + n^{-1}\sigma_\epsilon^2}{SS_\beta}ss_\beta - n^{-1}\frac{\sigma_\epsilon^2}{SS_\epsilon}ss_\epsilon\right]_+^{\frac{1}{2}} && (11)\\
&= w_i - \sqrt{b}\left(\frac{W_i - \mu_i}{\sqrt{SS_\beta}}\right)\sqrt{\frac{ss_\beta}{b}} + z_p\left[\frac{ss_\beta}{U_\beta} - n^{-1}\frac{ss_\epsilon}{U_\epsilon}\right]_+^{\frac{1}{2}} && (12)\\
&= w_i + \sqrt{\frac{ss_\beta}{ab(b-1)}}\left[\frac{Z + z_p\left\{b + a(b-1)\frac{ss_\epsilon}{ss_\beta}\frac{[U_\beta/a(b-1)]}{[U_\epsilon/(-n^{-1})]}\right\}_+^{\frac{1}{2}}}{\sqrt{U_\beta/a(b-1)}}\right] && (13)\\
&= w_i + \sqrt{\frac{ss_\beta}{ab(b-1)}}\left[\frac{Z + z_p\left\{b - n^{-1}\frac{(b-1)}{b(n-1)}\frac{ss_\epsilon}{ss_\beta}F\right\}_+^{\frac{1}{2}}}{\sqrt{U_\beta/a(b-1)}}\right] && (14)\\
&\approx_d w_i + \sqrt{\frac{ss_\beta}{ab(b-1)}}\left[\frac{Z + z_p\left\{b - n^{-1}\frac{(b-1)}{b(n-1)}\frac{ss_\epsilon}{ss_\beta}F_{a(b-1);ab(n-1)}(\theta)\right\}_+^{\frac{1}{2}}}{\sqrt{U_\beta/a(b-1)}}\right] && ,(15)
\end{aligned}
$$

where $c_+ = max(0, c)$, and all parameters as defined above at section 3.

Proceeding in some way as in the section 3, for the term

$$
T_{2p} = \left[\frac{Z + z_p\left\{b - n^{-1}\frac{(b-1)}{b(n-1)}\frac{ss_\epsilon}{ss_\beta}F_{a(b-1);ab(n-1)}(\theta)\right\}_+^{\frac{1}{2}}}{\sqrt{U_\beta/a(b-1)}}\right]
$$

of the approximation (15) (the L_{2p} approximation), since $Z \sim N(0,1)$ and $U_\beta \sim \chi^2_{a(b-1)}$,

$$
T_{2p} \sim t_{a(b-1)}(\delta_2),
$$

with

$$
\delta_2 = z_p\left\{b - n^{-1}\frac{(b-1)}{b(n-1)}\frac{ss_\epsilon}{ss_\beta}F_{a(b-1);ab(n-1)}(\theta)\right\}_+^{\frac{1}{2}}.
$$

That is, T_{2p} as a t distribution with $a(b-1)$ *df*, and the noncentrality parameter δ_2.

Thus, an aproximation, say $L_{2p}(\gamma)$, for the the (p, γ)-upper *tolerance limit* is given by

$$L_{2p}(\gamma) = w_i + t_{a(b-1);\gamma}(\delta_2)\sqrt{\frac{ss_\beta}{ab(b-1)}}. \tag{16}$$

with

$$\delta_2 = z_p \left\{ b - n^{-1}\frac{(b-1)}{b(n-1)}\frac{ss_\epsilon}{ss_\beta}F_{a(b-1);ab(n-1)}(\theta) \right\}_+^{\frac{1}{2}}. \tag{17}$$

4 Lower Tolerance Limit

Let $\mathbf{Y} = \{Y_{ij1}, \ldots, Y_{ijk}\}$ be a sample of the random variable $Y \sim N(\mu_i, \sigma_\beta^2 + \sigma_\epsilon^2)$. A statistic D is a (p, γ)- lower *tolerance limit* for Y, if

$$P_{\mathbf{Y}}[P_Y(Y \geq D \| \mathbf{Y}) \geq p] = \gamma \Longleftrightarrow P_{\mathbf{Y}}[D \leq q_{1-p}] = \gamma, \tag{18}$$

holds for $0 < \gamma < 1,$ and $0 < p < 1$, where q_{1-p} is the *(1-p)th* quantile of $N(\mu_i, \sigma_\beta^2 + \sigma_\epsilon^2)$. Like at the section 3, it is readily shown that D is a γ-lower *confidence limit* for the parametric function $\mu_i + z_{1-p}\sqrt{\sigma_\beta^2 + \sigma_\epsilon^2}$, where z_{1-p} denotes the *(1-p)th* quantile of the $N(0, 1)$, as we are about to show:

$$\begin{aligned}
&P_{\mathbf{Y}}[P_Y(Y \geq D \| \mathbf{Y}) \geq p] = \gamma \Longleftrightarrow \\
&P_{\mathbf{Y}}\left[P_Y\left(\frac{Y-\mu_i}{\sqrt{\sigma_\beta^2+\sigma_\epsilon^2}} \geq \frac{D-\mu_i}{\sqrt{\sigma_\beta^2+\sigma_\epsilon^2}}\|\mathbf{Y}\right) \geq p\right] = \gamma \Longleftrightarrow \\
&P_{\mathbf{Y}}\left[z_{1-p} \geq \frac{D-\mu_i}{\sqrt{\sigma_\beta^2+\sigma_\epsilon^2}}\right] = \gamma \Longleftrightarrow \\
&P_{\mathbf{Y}}\left[D \leq \mu_i + z_{1-p}\sqrt{\sigma_\beta^2+\sigma_\epsilon^2}\right] = \gamma,
\end{aligned}$$

where z_{1-p} is the standard normal distribution, since $\frac{Y-\mu_i}{\sqrt{\sigma_\beta^2+\sigma_\epsilon^2}} \sim N(0, 1)$.

Now, again with out lost of generality, let $\mathbf{Y} = \{Y_{ij1}, \ldots, Y_{ijk}\}$ be a sample of the random variable $Y^* \sim N(\mu_i, \sigma_\beta^2)$. Then, proceeding in a simillar way as at section 3, one concludes that the

(p, γ)-lower *tolerance limit* for the “true effects” variable
$Y^* = \mu_i + \tau_i + \beta_{j(i)} \sim N(\mu_i, \sigma_\beta^2)$

is simply the

γ-lower *confidence limit* for the parametric function $\mu_i + z_{1-p}\sigma_\beta$.

4.1 The Observable Random Variable $Y \sim N(\mu_i, \sigma_\beta^2 + \sigma_\epsilon^2)$

This section is devoted to the construction of the (p,γ)-lower *tolerance limit* for the random variable $Y \sim N(\mu_i, \sigma_\beta^2+\sigma_\epsilon^2)$. In order to do so, we consider a *general pivotal quantity*, say L_{3p}, for the parametric function $\mu_i + z_{1-p}\sqrt{\sigma_\beta^2+\sigma_\epsilon^2}$ and then, like as at the section 3.1, we construct an appoximation, say $L_{3p}(\gamma)$, to its γth quantile.

Following the procedure performed at the section 3.1,

$$
\begin{aligned}
L_{3p} &= w_i - \frac{\sqrt{b}(W_i-\mu_i)}{\sqrt{SS_\beta}}\sqrt{\frac{ss_\beta}{b}} + z_{1-p}\left[\frac{\left(\sigma_\beta^2+n^{-1}\sigma_\epsilon^2\right)}{SS_\beta}ss_\beta + \left(1-n^{-1}\right)\frac{\sigma_\epsilon^2}{SS_\epsilon}ss_\epsilon\right]^{\frac{1}{2}} \\
&= w_i - \frac{\sqrt{b}(W_i-\mu_i)}{\sqrt{SS_\beta}}\sqrt{\frac{ss_\beta}{b}} + z_{1-p}\left[\frac{ss_\beta}{U_\beta} + \left(1-n^{-1}\right)\frac{ss_\epsilon}{U_\epsilon}\right]^{\frac{1}{2}} && (19) \\
&\approx_d w_i + \sqrt{\frac{ss_\beta}{ab(b-1)}}\left[\frac{-Z + z_{1-p}\left(b + \frac{(b-1)\left(1-n^{-1}\right)}{(n-1)}\frac{ss_\epsilon}{ss_\beta}F_{a(b-1);ab(n-1)}(\theta)\right)^{\frac{1}{2}}}{\sqrt{U_\beta/a(b-1)}}\right] && (20) \\
&= w_i - \sqrt{\frac{ss_\beta}{ab(b-1)}}\left[\frac{Z + z_p\left(b + \frac{(b-1)\left(1-n^{-1}\right)}{(n-1)}\frac{ss_\epsilon}{ss_\beta}F_{a(b-1);ab(n-1)}(\theta)\right)^{\frac{1}{2}}}{\sqrt{U_\beta/a(b-1)}}\right] && (21) \\
&= w_i - \sqrt{\frac{ss_\beta}{ab(b-1)}}\left[\frac{Z + z_p\left(b + \frac{(b-1)\left(1-n^{-1}\right)}{(n-1)}\frac{ss_\epsilon}{ss_\beta}F_{a(b-1);ab(n-1)}(\theta)\right)^{\frac{1}{2}}}{\sqrt{U_\beta/a(b-1)}}\right], && (22)
\end{aligned}
$$

and therefore, proceeding in same way, the approximation referred to at the previous paragraph is given by

$$L_{3p}(\gamma) = w_i - t_{a(b-1);\gamma}(\delta_3)\sqrt{\frac{ss_\beta}{ab(b-1)}}, \tag{23}$$

with

$$\delta_3 = z_p\left(b + \frac{(b-1)\left(1-n^{-1}\right)}{(n-1)}\frac{ss_\epsilon}{ss_\beta}F_{a(b-1);ab(n-1)}(\theta)\right)^{1/2}, \tag{24}$$

where $t_{r;s}(m)$ denotes the sth quantile of a noncentral t distribution with r *df*, and the noncentrality parameter m, with the random variables W_i, SS_β, SS_ϵ, Z, U_ϵ, and U_β, and the quantities z_p, w_i, ss_β, ss_ϵ, $F_{r;s}(m)$, and θ defineds at section 3. L_{3p} is clearly a *general pivotal quantity* as shown at section 3.

The transition from approximation (20) to the equation (21) in the L_{3p} development is due to the standard normal distribution symmetry around 0.

4.2 The Unobserved "True Effects" Variable $Y^* \sim N(\mu_i, \sigma_\beta^2)$

Here our concern is the construction of an lower *tolerance limit* for the unobserved "true effect" $Y^* = \mu_i + \tau_i + \beta_{j(i)} \sim N(\mu_i, \sigma_\beta^2)$.

As proven at the section 4.1, it is easily shown that the (p, γ)-lower *tolerance limit* for Y^* is simply the γ-lower *confidence limit* for the parametric function $\mu_i + z_{1-p}\sigma_\beta$, with z_{1-p} the *(1-p)th* quantile of the standard normal distribution. Hence, the following quantity (L_{4p}) is a *general pivotal quantity* for the parametric function $\mu_i + z_p\sigma_\beta$:

$$
\begin{aligned}
L_{4p} &= w_i - \sqrt{b}\left(\frac{W_i - \mu_i}{\sqrt{SS_\beta}}\right)\sqrt{\frac{ss_\beta}{b}} + z_{1-p}\left[\frac{\sigma_\beta^2 + n^{-1}\sigma_\epsilon^2}{SS_\beta} ss_\beta - n^{-1}\frac{\sigma_\epsilon^2}{SS_\epsilon} ss_\epsilon\right]_+^{\frac{1}{2}} && (25)\\
&= w_i - \sqrt{b}\left(\frac{W_i - \mu_i}{\sqrt{SS_\beta}}\right)\sqrt{\frac{ss_\beta}{b}} + z_{1-p}\left[\frac{ss_\beta}{U_\beta} - n^{-1}\frac{ss_\epsilon}{U_\epsilon}\right]_+^{\frac{1}{2}} && (26)\\
&= w_i + \sqrt{\frac{ss_\beta}{ab(b-1)}}\left[\frac{Z + z_{1-p}\left\{b + a(b-1)\frac{ss_\epsilon}{ss_\beta}\frac{[U_\beta/a(b-1)]}{[U_\epsilon/(-n^{-1})]}\right\}_+^{\frac{1}{2}}}{\sqrt{U_\beta/a(b-1)}}\right] && (27)\\
&= w_i + \sqrt{\frac{ss_\beta}{ab(b-1)}}\left[\frac{Z + z_{1-p}\left\{b - n^{-1}\frac{(b-1)}{b(n-1)}\frac{ss_\epsilon}{ss_\beta}F\right\}_+^{\frac{1}{2}}}{\sqrt{U_\beta/a(b-1)}}\right] && (28)\\
&\approx_d w_i + \sqrt{\frac{ss_\beta}{ab(b-1)}}\left[\frac{Z + z_{1-p}\left\{b - n^{-1}\frac{(b-1)}{b(n-1)}\frac{ss_\epsilon}{ss_\beta}F_{a(b-1);ab(n-1)}(\theta)\right\}_+^{\frac{1}{2}}}{\sqrt{U_\beta/a(b-1)}}\right] && (29)\\
&= w_i - \sqrt{\frac{ss_\beta}{ab(b-1)}}\left[\frac{Z + z_p\left\{b - n^{-1}\frac{(b-1)}{b(n-1)}\frac{ss_\epsilon}{ss_\beta}F_{a(b-1);ab(n-1)}(\theta)\right\}_+^{\frac{1}{2}}}{\sqrt{U_\beta/a(b-1)}}\right], && (30)
\end{aligned}
$$

where $c_+ = max(0, c)$, and all parameters as defined at section 3.

Recalling from the section 3.2 that the term

$$
T_{4p} = \left[\frac{Z + z_p\left\{b - n^{-1}\frac{(b-1)}{b(n-1)}\frac{ss_\epsilon}{ss_\beta}F_{a(b-1);ab(n-1)}(\theta)\right\}_+^{\frac{1}{2}}}{\sqrt{U_\beta/a(b-1)}}\right]
$$

of equation (30) (of the L_{4p} development) has a t distribution with $a(b-1)$ *df*, and the noncentrality parameter δ_4, i.e,

$$
T_{4p} \sim t_{a(b-1)}(\delta_4),
$$

with $\delta_4 = z_p\left\{b - n^{-1}\frac{(b-1)}{b(n-1)}\frac{ss_\epsilon}{ss_\beta}F_{a(b-1);ab(n-1)}(\theta)\right\}_+^{\frac{1}{2}}$, an aproximation, say $L_{4p}(\gamma)$, for the the (p, γ)-lower *tolerance interval* is given by,

$$L_{4p}(\gamma) = w_i - t_{a(b-1);\gamma}(\delta_4)\sqrt{\frac{ss_\beta}{ab(b-1)}}. \tag{31}$$

with

$$\delta_4 = z_p \left\{ b - n^{-1} \frac{(b-1)}{b(n-1)} \frac{ss_\epsilon}{ss_\beta} F_{a(b-1);ab(n-1)}(\theta) \right\}_+^{\frac{1}{2}}. \tag{32}$$

5 Concluding Remarks

In spite of the problem of deriving the one-sided *tolerance limit* presented here and so at Fonseca et al [3], and Krishnamoorthy and Mathew [4], by using the *generalized confidence limit*, *confidence limit* and *tolerance limit* solve different problems in the statistical inferences context. The first one only provides bounds for unknowns scalar parameters of the population such as its mean, its variance, its quantile, its tail probability, etc, having the sample available. The later one, the *tolerance limit*, provides bounds for the entire populations, based on populations data, i.e., such limit is expected to contain a specified proportion or more of the sample population. Both *tolerance limit* and *confidence limit* regions are similarly defined for a multivariate population.

The approximation for the upper *tolerance limit* for both the observed random variable Y and the unobserved true effects Y^* presented here have a very satisfactory performance in the context of unbalanced data as shown by Fonseca et al [3] by numerically investigating both of them. In this sense, and once the lower *tolerance limit* we have constructed here is based on the same *generalized confidence limit* idea we do not investigate its numerical performance once it would be exactly the same one.

References

[1] G. Casella and R.L. Berger. *Statistical Inference.* 2002.

[2] S.M. Dowdy and D. Chilko. *Statistics for Research.* 2004.

[3] M. Fonseca, T. Mathew, J. T. Mexia, and R. Zmyslony. Tolerance intervals in a two-way nested model with mixed or random effects. *Statistics: A Journal of Theoretical and Applied Statistics*, 2007.

[4] K. Krishnamoorthy and T Mathew. One-sided tolerance limits in balanced and unbalanced one-way random models based on generalized confidence intervals. *American Statistical Association*, 2004.

[5] K. Krishnamoorthy and T. Mathew. *Statistical Tolerance Regions: Theory, Applications, and Computation.* 2009.

[6] D. C. Montgomery. *Design and Analysis of Experiments.* 2008.

[7] G. Sharma and T. Mathew. One-sided and two-sided tolerance intervals in general mixed and random effects models using small-sample asymptotics. *American Statistical Association*, 2012.

[8] E. Ukaegbu and S. Sanni. On three-way unbalance nested analysis of variance. *Journal of Mathematics and Statistics*, 2012.