

# CLASSROOM SIMULATION: UNDERSTANDING ONE-WAY RANDOM-EFFECT ANOVA

Eric A. SUESS, Bruce E. TRUMBO, and Yun JIANG  
*California State University, East Bay*

Department of Statistics; California State University, East Bay (formerly CSU Hayward); Hayward, CA 94542 USA  
 eric.suess@csueastbay.edu or bruce.trumbo@csueastbay.edu

**Abstract:** The one-way random-effect ANOVA model is presented, and two simulated datasets are analyzed. and discussed from three points of view: (1) The standard ANOVA table, F test, and method-of-moments estimates of variance components, which can lead to negative estimates. (2) Maximum likelihood estimates of variance components. (3) Bayesian probability intervals for variance components based on flat priors and computed using a Gibbs sampler. Computations are done in R and WinBUGS. Level and methods are appropriate for presentation to advanced undergraduate and first year MS students.

**Key Words:** Analysis of variance, random effect model, estimation of variance components, interval estimates, Bayesian estimation, flat (noninformative) prior, Gibbs sampler, teaching, S-Plus / R.

**1. Introduction.** A common use of the one-way random-effect analysis of variance model is in manufacturing situations where a product is made in two stages: first, batches of a precursor are made or selected; second, the final items are produced and measured. The central question is often what relative contributions these two stages make to the observed variability of the final items.

Examples are batches of concrete cast into individual concrete beams, which are then tested for breaking strength, wafers from which individual computer chips are cut and performance tested, batches of a precursor protein from which individual doses of a drug are derived and tested for potency, and plants from which individual leaves or fruits are selected and assayed for chemical composition.

**The model.** Here we consider the balanced model

$$Y_{ij} = \mu + A_i + e_{ij},$$

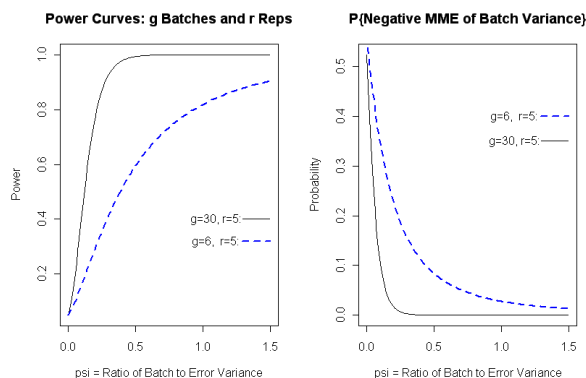
where  $i = 1, \dots, g$  batches,  $j = 1, \dots, r$  replications within each batch,  $A_i$  are independent identically distributed (iid)  $\text{NORM}(0, \sigma_A^2)$ , and  $e_{ij}$  are iid  $\text{NORM}(0, \sigma^2)$ . Thus the variance reflected in the measurement of an individual item is  $V(Y_{ij}) = \sigma_A^2 + \sigma^2$ . We focus mainly on inferences about the variance components  $\sigma_A^2$  (among batches) and  $\sigma^2$  (within batches).

**2. Traditional ANOVA.** The traditional analysis of variance tests  $H_0: \sigma_A^2 = 0$  against  $H_a: \sigma_A^2 > 0$ . The test statistic is  $F = \text{MS}_B/\text{MS}_E$ , where  $\text{MS}_B = \text{SS}_B/\text{df}_B$ ,  $\text{SS}_B = r \sum_i (\bar{Y}_i - \bar{Y}_{..})^2$ ,  $\text{df}_B = g - 1$ ,  $\text{MS}_E = \text{SS}_E/\text{df}_E$ ,  $\text{SS}_E = \sum_i \sum_j (Y_{ij} - \bar{Y}_i)^2$ , and  $\text{df}_E = g(r - 1)$ . Under  $H_0$ , the

statistic  $F = \text{MS}_B/\text{MS}_E$  has the F distribution with  $\text{df}_B$  and  $\text{df}_E$  degrees of freedom and we reject  $H_0$  for sufficiently large values of  $F$ .

In practice,  $H_0$  is rarely precisely true, so that the salient issue in testing  $H_0$  is by how much the ratio  $\psi = \sigma_A^2/\sigma^2$  must exceed 0 in order for the F test to have a reasonable chance of detecting that it does exceed 0.

**Power.** In contrast with the one-way fixed-effect ANOVA, where power is found using a *noncentral* F distribution, the power for a random-effect ANOVA is found using the (ordinary) F distribution. Specifically, the power of a test of level  $\alpha$  is  $P\{(1 + r\psi)F > F^*\}$ , where  $F^*$  cuts off area  $\alpha$  from the right tail of the distribution  $F(g - 1, g(r - 1))$ .



**Fig. 1. Power (left) and  $P\{\text{MME}(\sigma_A^2) < 0\}$  for  $g = 6, 30; r = 5$ .**

Many textbook examples use small values of  $g$  to simplify computation. For  $\alpha = 5\%$ ,  $g = 6$ ,  $r = 5$ ,  $\psi = 0.1$ , the power of the F test is  $P\{F > 2.62/1.5\} = 16.2\%$ . Even if  $\psi = 1$ , the power is only 81.8%. [In R: `1 - pf(2.62/6, 5, 24)` returns 0.8184741.] As shown in the left side of Fig. 1, the power is much better for a larger experiment with  $g = 30$  and  $r = 5$  (solid line).

**Method of moments estimates.** In real applications, it is often more to the point to estimate variance components than to test  $H_0$ . Because  $E(\text{MS}_E) = \sigma^2$  and  $E(\text{MS}_B) = r\sigma_A^2 + \sigma^2$ , the method-of-moments estimators (MMEs) are  $\text{MS}_E$  for  $\sigma^2$  and  $(\text{MS}_B - \text{MS}_E)/r$  for  $\sigma_A^2$ . Unfortunately, the latter estimate takes absurd negative values whenever it happens that  $F < 1$ . If  $\sigma_A^2$  is relatively close to 0, then negative estimates are more than a theoretical possibility as shown in the right side of Fig. 1.

**Confidence intervals.** A confidence interval (CI) for  $\sigma^2$  is based on the chi-squared distribution with  $\text{df}_E = g(r - 1)$  degrees of freedom. For  $\text{df}_E = 120$ , the ratio of the upper

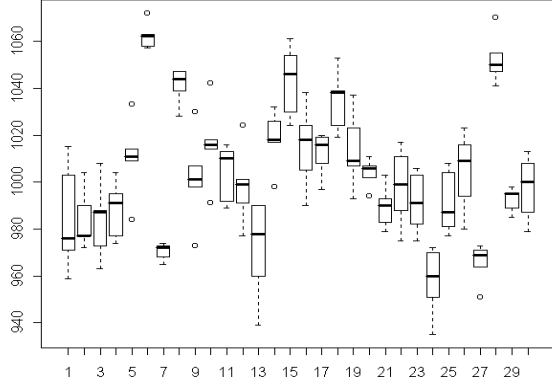


Fig. 2. Boxplots for 30 batches of Dataset 1:  $\sigma_A = 25$ ,  $\sigma = 15$ .

Source	DF	SS	MS	F	P
Batch	29	91365	3151	15.0	0.000
Error	120	25215	210		

Fig. 3. ANOVA table for Dataset 1: F significant.

to lower 95% confidence limits for  $\sigma^2$  can be computed as  

$$qchisq(.975, 120)/qchisq(.025, 120)$$
which returns 1.66.

No exact confidence interval for  $\sigma_A^2$  is known, and unless  $g$  and  $r$  are both large enough, the accuracy of approximate intervals may be in doubt. Also, a confidence interval based on the method-of-moments estimate of  $\sigma_A^2$  is even more likely to include negative values than the estimate itself is to be negative.

Furthermore, if the number  $g$  of batches is small, then any correct CI for  $\sigma_A^2$  may be so long as to be of no practical use. For example, with a simple random sample of size  $n = 6$  from a normal population the ratio of the upper to the lower 95% confidence limit will be about 15.6 to 1 (for  $n = 30$ , about 2.85 to 1). And in a one-way random effects ANOVA there is the additional complication that information on  $\sigma_A^2$  must be disentangled from information on  $\sigma^2$ . Thus the examples with small  $g$  that are common in textbooks may not be ideal for giving students an realistic view of the applications of this design.

However, after a little algebra, one can see that an exact 95% confidence interval for  $\psi$  has the form  $([F/U^* - 1]/r, [F/L^* - 1]/r)$ , where  $F$  is the F-statistic in the ANOVA table, and  $L^*$  and  $U^*$  are the 2.5% and 97.5% points, respectively, of the F distribution with degrees of freedom  $df_B$  and  $df_E$ . Sometimes it is useful to think in terms of the *interclass correlation coefficient* (ICC)  $\rho_I = \sigma_A^2/(\sigma_A^2 + \sigma^2) = \psi/(\psi + 1)$ . The confidence interval  $([F - U^*]/[F + (r - 1)U^*], [F - L^*]/[F + (r - 1)L^*])$  for  $\rho_I$  can be derived from the interval for  $\psi$  given above. [3, 4]

It is natural to estimate  $\mu$  as the grand mean of all  $gr$  observations,  $\bar{Y}_{..} = (1/gr)\sum_i \sum_j Y_{ij}$ , distributed normally with mean  $\mu$  and variance  $(\sigma_2 + r\sigma_A^2)/gr$ . But  $\sigma^2 + r\sigma_A^2$

is estimated by  $MS_B$ , which has  $df_B = g - 1$  degrees of freedom. Thus the t distribution can be used to obtain  $\bar{Y}_{..} \pm t^*(MS_B/gr)^{1/2}$  as a 95% confidence interval for  $\mu$ , where  $t^*$  cuts area 2.5% from the upper tail of the t distribution with  $g - 1$  degrees of freedom.

### 3. Two simulated datasets.

In what follows we refer to two simulated datasets, each with  $g = 30$  and  $r = 5$ . Thus there are enough batches to get stable results with a variety of methods and we have the advantage of knowing the true parameter values so we can judge whether our estimates are reasonable. Printouts of these datasets are available online [5].

Dataset 1, with  $\mu = 1000$ ,  $\sigma_A = 25$ , and  $\sigma = 15$ , was generated batchwise, using the default random number generator in R 2.1.1 with `set.seed(1237)` and rounding to integer values. Fig. 2 shows boxplots of the  $g = 30$  batches and Fig. 3 shows the ANOVA table with a very highly significant F statistic. The MMEs are 1003 for  $\mu$  with 95% CI (994.2, 1012.3); 588.1 for  $\sigma_A^2$  (or 24.25 for  $\sigma_A$ ); and 14.5 for  $\sigma$  with CI (12.87, 16.59).

Dataset 2, with  $\mu = 1000$ ,  $\sigma_A = 1$ , and  $\sigma = 25$  was generated similarly with `set.seed(12)`. Fig. 4 shows boxplots of the batches. The F ratio in the ANOVA table of Fig. 5 is far from significant. Here the MMEs are 1000 for  $\mu$  with CI (996.4, 1003.7); *negative value*  $-5.96$  for  $\sigma_A^2$ ; and 23.34 for  $\sigma$  with CI (20.72, 26.72).

Below we look at methods that give similar results to those above for the “well-behaved” Dataset 1, but also provide more useful results for Dataset 2, which is in some respects analytically “badly-behaved.”

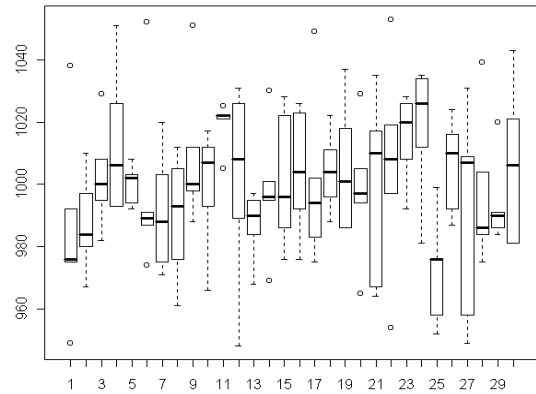


Fig. 4. Boxplots for 30 batches of Dataset 2:  $\sigma_A = 1$ ,  $\sigma = 25$ .

Source	DF	SS	MS	F	P
Batch	29	14931	515	0.95	0.552
Error	120	65362	545		

Fig. 5. ANOVA table for Dataset 2: F not significant.

### 3. Maximum likelihood estimators.

Observations from different batches are independent. For example,  $Y_{11} = \mu + A_1 + e_{11}$  and  $Y_{21} = \mu + A_2 + e_{21}$  are independent because  $A_1, A_2, e_{11}$  and  $e_{12}$  are mutually independent. However, different observations from the *same* batch, say  $Y_{11}$  and  $Y_{12}$ , are correlated because they have  $A_1$  in common. Specifically,  $\text{Cov}(Y_{11}, Y_{12}) = \text{Cov}(A_1 + e_{11}, A_1 + e_{12}) = \text{Cov}(A_1, A_1) = \text{V}(A_1) = \sigma_A^2$ . Because  $Y_{11}$  and  $Y_{12}$  both have the same variance  $\text{V}(Y_{11}) = \text{V}(Y_{12}) = \sigma_A^2 + \sigma^2$ , the correlation of these (or any other) two different observations from the same batch is  $\rho_I = \sigma_A^2 / (\sigma_A^2 + \sigma^2)$ .

```
require(nlme)
ml.fit <- lme(x ~ 1, random = ~1|s,
             method="ML"); ml.fit

Linear mixed-effects model fit
  by maximum likelihood
Data: NULL
Log-likelihood: -654.0247
Fixed: x ~ 1
(Intercept)
 1003.253

Random effects:
Formula: ~1 | s
(Intercept) Residual
StdDev:      23.81335 14.49575

Number of Observations: 150
Number of Groups: 30
```

**Fig. 6. Maximum likelihood estimates for Dataset 1.**

**Joint density function.** From the results in the previous paragraph we see that the joint density function of the observations  $Y_{ij}$  is multivariate normal of dimension  $br$ :

$f(\mathbf{y} | \boldsymbol{\mu}, \mathbf{V}) = (2\pi)^{-br/2} |\mathbf{V}|^{1/2} \exp[-(1/2)(\mathbf{y} - \boldsymbol{\mu})' \mathbf{V}^{-1} (\mathbf{y} - \boldsymbol{\mu})]$ , where  $\mathbf{y} = (Y_{11}, Y_{12}, \dots, Y_{1r}, \dots, Y_{br})$  and the mean vector  $\boldsymbol{\mu}$  has all  $gr$  components equal to  $\mu$ . The  $gr \times gr$  correlation matrix  $\mathbf{C} = [1/(\sigma_A^2 + \sigma^2)] \mathbf{V}$  is composed of  $b^2$  submatrices, each  $r \times r$ . Of these submatrices,  $g^2 - g$  have all elements 0. But along the principal diagonal of  $\mathbf{C}$  are  $b$  identical submatrices, each with elements 1 along its principal diagonal and elements  $\rho_I$  elsewhere.

**Computing MLEs.** If the density function  $f(\mathbf{y} | \boldsymbol{\mu}, \mathbf{V})$  is viewed as a function of  $\boldsymbol{\mu}$  and  $\mathbf{V}$  for  $\mathbf{y}$  as given by the data, then it is called a likelihood function. The values of  $\mu, \sigma_A^2$ , and  $\sigma^2$ , that maximize this likelihood function are called the maximum likelihood estimates (MLEs) of these three parameters of the model. One thing about the MLE for  $\sigma_A^2$  is clear without computation: it *cannot be negative* because  $f(\mathbf{y} | \boldsymbol{\mu}, \mathbf{V}) = 0$ , except where  $\sigma_A^2, \sigma^2 > 0$ . Specifically, the MLE of  $\sigma_A^2$  is  $[(1 - 1/g)\text{MS}_A - \text{MS}_E] / r$ , provided this quantity is positive; and the MLE of  $\sigma$  is  $\min[\text{MS}_E, \text{SS}_E + \text{SS}_B/gr]$ ; see [2].

Fig. 6 shows R code (in **bold**) and output for Dataset 1. Also, the additional command `intervals(ml.fit)` gives 95% CIs for the parameters. Approximate MLEs are: 1003 for  $\mu$  with CI (994.3, 1012.2); 23.81 for  $\sigma_A$  with CI (18.14, 31.26); and 14.50 for  $\sigma$  with CI (12.77, 16.45). These results are similar to the ones shown earlier for MMEs, and all intervals easily cover the known true parameter values. (S-Plus can also find CIs, but the implementation is a little different; see [5].)

MLEs obtained similarly for Dataset 2 are: 1000 for  $\mu$ , 0.56 for  $\sigma_A$ , (as approximated in R), and 23.13 for  $\sigma$ . However, R does not provide confidence intervals for this dataset because the sample variance-covariance matrix is ill-conditioned. This is not surprising: If we had  $\sigma_A = 0$  instead of 1, the dimensionality of the model would have been smaller.

**4. The Gibbs sampler using R.** The Gibbs sampler is a computational method that involves Markov chains and a Bayesian framework. See [6] for an elementary general introduction to Gibbs sampling and [1] for the details of the specific situation discussed in this section.

With brackets denoting density functions in order to avoid multiple-level subscripts, the joint distribution of  $Y_{ij}, \mu, A_i, \tau_A = 1/\sigma_A^2$  and  $\tau = 1/\sigma^2$  (all random variables in the Bayesian model) can be written in terms of conditional distributions as follows [1]:

$$[\mathbf{Y}, \mathbf{A}, \mu, \tau_A, \tau] \propto [\mathbf{Y} | \mathbf{A}, \tau_A] \times [\mathbf{A} | \mu, \tau_A] \times [\mu] \times [\tau_A] \times [\tau].$$

In the Gibbs sampler we use the flat normal prior  $\mu \sim \mathbf{N}(a_1 = 0, b_1 = 10^{10})$  and the flat conjugate gamma priors  $\tau_A \sim \text{GAMMA}(a_2 = 0.001, b_2 = 0.001)$  and  $\tau \sim \text{GAMMA}(a_3 = 0.001, b_3 = 0.001)$ . Also, we assign starting values to parameters:  $\mu = 1500$ ,  $\sigma_A^2 = 1/\tau_A = 1$ ,  $\sigma^2 = 1/\tau = 1$ , and  $A_i = \bar{Y}_i = (1/r) \sum_j Y_{ij}$ .

*At step one* the data  $\mathbf{Y}$  and these starting values and priors are used to generate a value of  $\sigma_A^2$  according to

$$[\sigma_A^2 | \mathbf{Y}, \mathbf{A}, \mu, \sigma^2] = [\sigma_A^2 | \mathbf{A}, \mu] = 1/\text{GAMMA}(a_2 + g/2, b_2 + (1/2) \sum_i (A_i - \mu)^2)$$

and a value of  $\sigma^2$  according to

$$[\sigma^2 | \mathbf{Y}, \mathbf{A}, \mu, \sigma_A^2] = [\sigma^2 | \mathbf{Y}, \mathbf{A}] = 1/\text{GAMMA}(a_3 + gr/2, b_3 + (1/2) \sum \sum (Y_{ij} - A_i)^2)$$

Next, these values of  $\sigma_A^2$  and  $\sigma^2$  are used to generate a new value of  $\mu$

$$[\mu | \mathbf{Y}, \mathbf{A}, \sigma_A^2, \sigma^2] = [\mu | \sigma_A^2, \mathbf{A}] = \mathbf{N}([a_1 \sigma_A^2 + b_1 \sum A_i] / d_1, b_1 \sigma_A^2 / d_1)$$

and new values of  $A_i$  are generated from  $[\mathbf{A} | \mathbf{Y}, \mu, \sigma_A^2, \sigma^2]$ , or, more simply, as

$$A_i \sim \mathbf{N}([r \sigma_A^2 / d_2] \bar{Y}_i + \mu \sigma^2 / d_2, \sigma^2 \sigma_A^2 / d_2)$$

where  $d_1 = \sigma_A^2 + rb_1$  and  $d_2 = g \sigma_A^2 + \sigma^2$ .

Then recursively, *at step n* the simulated values from step  $n - 1$  are used to find the  $n^{\text{th}}$  value of  $\sigma_A^2$ . In each normal distribution used here, the mean can be viewed as a precision-weighted average of information

from appropriate resampled data and prior, and the precision (reciprocal variance) can be viewed as a sum of precisions from appropriate resampled data and prior.

```
# Assumes g x r matrix Y of observations
m <- 100000 # iterations
b <- m/4 # burn-in

va <- numeric(m); ve <- numeric(m)
mu <- numeric(m); a <- rowMeans(Y)

## Prior parameters, G/S Scenario I
a1 <- 0 # mean of prior on mu
b1 <- 10^(10) # var of prior on mu
a2 <- .001 # shape of prior on batch var
b2 <- .001 # rate of prior on batch var
a3 <- .001 # shape of prior on error var
b3 <- .001 # rate of prior on error var

## Initial values of model parameters
mu[1] <- 1500
va[1] <- 1
ve[1] <- 1

for (n in 2:m){
  va[n] <- 1/rgamma(1, a2 + g/2,
    b2 + sum((a - mu[n-1])^2)/2)
  ve[n] <- 1/rgamma(1, a3 + r*t/2,
    b3 + sum((Y - matrix(a, g, r))^2)/2)
  mu[n] <- rnorm(1, (va[n]*a1 +
    b1*sum(a))/(va[n] + g*v1),
    sqrt(b1*vb[n]/(va[n] + t*b1)))
  a <- rnorm(g, (r*va[n]*X.bar +
    ve[n]*mu[n])/(r*va[n] + ve[n]),
    sqrt((va[n]*ve[n])/(r*va[n]+ve[n])))
}
```

**Fig. 7. R code for our Gibbs Sampler.**

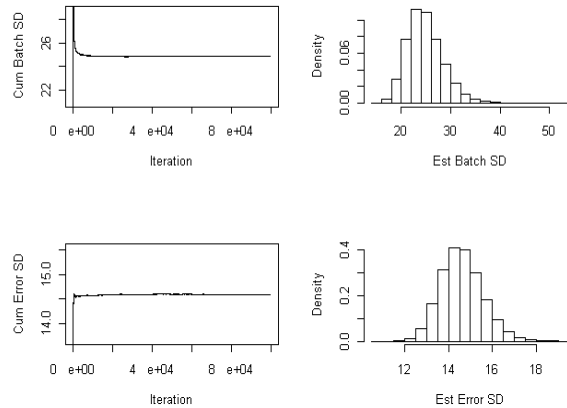
The R code to implement this Gibbs sampler is shown in Fig. 7. For Dataset 1, the first column in Fig. 8 shows the stability of the point estimates of  $\sigma_A$  (upper) and  $\sigma$ , and the second column shows the simulated densities (after burn in) from which the interval estimates are obtained. For Dataset 2, the simulated densities for  $\sigma_A$  (and thus  $\rho_I$ ) are strongly right-skewed. Diagnostic graphics show that the Gibbs Sampler is performing properly; see [5]. Bayesian point and interval estimates obtained from our Gibbs sampler are shown in Fig. 10. The Gibbs sampler used flat priors, so it is not surprising that the resulting Bayesian estimates are in relatively close agreement with the MLEs.

#### 4. Bayesian Results From WinBUGS.

WinBUGS [8] is free software widely used in Bayesian analysis. We used two parameterizations to estimate the mean  $\mu$ , the variance components  $\sigma_A$  and  $\sigma$  and the intraclass correlation  $\rho_I$ . The first parameterization is the same as in Section 3, and the second puts a uniform prior on  $\rho_I$  instead of the flat gamma prior on  $\tau_A = 1/\sigma_A^2$ .

An advantage of WinBUGS is that it that the user does not have to program the partial conditional distributions as was done in Fig. 7, so that the code is relatively simple. Fig. 9 shows the code for the second parameterization, which would have been especially difficult in R.

WinBUGS results are shown in Fig. 10; those based on the same parameterization as in Section 3 are very similar to the results obtained there. (See [5] for more-detailed results and code for the first parameterization.)



**Fig. 8. Graphics for Gibbs Sampler: Dataset 1.**

```
data
list(batches = 30, samples = 5,
y = structure(.Data =
  c(959, 976, 1015, 1003, 971,
    [28 Batches omitted to save space]
    1000, 987, 979, 1013, 1008),
  .Dim = c(30, 5)))

inits
list(theta=1500, tau.with=1, ICC=0.5)

model {
for( i in 1 : batches ) {
mu[i] ~ dnorm(theta, tau.btw)
for( j in 1 : samples ) {
y[i , j] ~ dnorm(mu[i], tau.with) } }
theta ~ dnorm(0.0, 1.0E-10)

# prior for within-variation
sigma2.with <- 1 / tau.with
tau.with ~ dgamma(0.001, 0.001)
mu.grand <- mean(mu[])
sigma.with <- sqrt(sigma2.with)
sigma.btw <- sqrt(sigma2.btw)

# prior for ICC
ICC ~ dunif(0,1)
sigma2.btw <- sigma2.with *ICC/(1-ICC)
tau.btw <- 1/sigma2.btw
}
```

**Fig. 9. WinBUGS Code for Second Parameterization**

## 5. Summary Comments.

We have seen that all methods—MME, MLE, and Bayesian—give good results for Dataset 1, in which  $\sigma_A^2$  is relatively large. Exact confidence intervals are available for  $\mu$ ,  $\sigma$ , and  $\rho_I$  (see Fig. 10).

However, for Dataset 2, in which  $\sigma_A^2$  is too small to detect with a standard ANOVA, the MME is negative. A standard interpretation of a negative MME is that  $\sigma_A^2 > 0$  is “very small,” but then confidence intervals for  $\sigma_A^2$  or  $\rho_I$  are problematic. (In certain biological applications a negative value of  $\rho_I$  may be legitimate, but not for the ANOVA model assumed here, see [4].) Dataset 2 also illustrates that an ill-conditioned sample variance-covariance matrix prevents R from finding confidence intervals for the ML estimates.

For Dataset 2 the Bayesian approach of Section 3 yields reasonable point estimates and interval estimates that cover the known parameter values. But in practice, the interval estimates of  $\sigma_A$  may too long to be useful. The second parameterization of Section 4 (WinBUGS) gives less-satisfactory interval estimates for  $\sigma_A$  and  $\rho_I$ .

Except for Bayes estimates with Dataset 2 based on this alternative parameterization, all of the methods give similar results—when they are able to give sensible results at all. However, this good agreement among methods may not hold for smaller datasets (typical in some applications) where there is less information to help determine the relative sizes of  $\sigma$  and  $\sigma_A$ .

## 6. References

- Gelfand, Alan E.; and Smith, A.F.M.: “Sampling-based approaches to calculating marginal densities,” *JASA* **85**, 410 (1990) pp 398-409. [Partial conditional distributions for Gibbs sampler in Sect. 3.4 pp 405-406.]
- Miller, R.G., *Beyond ANOVA*, Chapman-Hall (1997).
- Oehlert, Gary W.: *A first course in design and analysis of experiments*. W.H. Freeman (2000). [Chapter 11 discusses ANOVA models for random effects models including power and MME point and interval estimation.]
- Snedecor, G.W.; and Cochran, W.G.: *Statistical Methods*, 7th ed. (1980). [Chapter 13 discusses models used here and also an alternative model where the intraclass correlation can legitimately be negative.]
- Suess, Eric A.: [www.sci.csueastbay.edu/~esuess/onewayrand](http://www.sci.csueastbay.edu/~esuess/onewayrand) [Full code for MLE, Gibbs and WinBUGS approaches. Auxiliary instructional materials.]
- Trumbo, Bruce E.; Suess, Eric A.; and Fraser, C.M.: “Elementary uses of the Gibbs Sampler: Applications to medical screening tests,” *STATS*, Winter 2000, pp 3-10. [An elementary exposition of Gibbs samplers.]
- University of Montana, ST412 website: [www.math.montana.edu/~jarrett/st412/lec17.pdf](http://www.math.montana.edu/~jarrett/st412/lec17.pdf). [How to use R to get MLEs for the one-way random effect model.]
- WinBugs website [www.mrc-bsu.cam.ac.uk/bugs/welcome.shtml](http://www.mrc-bsu.cam.ac.uk/bugs/welcome.shtml). [Under Examples for Volume 1 is a description of the one-way random effects model (Dyes data); specifically, see the PDF file [www.mrc-bsu.cam.ac.uk/bugs/winbugs/Vol11.pdf](http://www.mrc-bsu.cam.ac.uk/bugs/winbugs/Vol11.pdf).]

### Dataset 1

Method	$\mu = 1000$	“Among” $\sigma_A = 25$	“Within” $\sigma = 15$	ICC = .735
MME (ANOVA)	994 <b>1003</b> 1012	<b>24.25</b>	12.9 <b>14.5</b> 16.6	.610 <b>.737</b> .845
MLE (R)	994 <b>1003</b> 1012	18.1 <b>23.8</b> 31.3	12.8 <b>14.5</b> 16.5	<b>.729</b>
Bayes (R, Gibbs)	994 <b>1003</b> 1012	18.9 <b>24.9</b> 33.1	12.9 <b>14.6</b> 16.6	.609 <b>.736</b> .845
Bayes* (WinBUGS, MCMC)	994 <b>1003</b> 1013	18.8 <b>24.9</b> 33.0	12.9 <b>14.6</b> 16.6	.607 <b>.736</b> .844
	994 <b>1003</b> 1012	18.6 <b>24.4</b> 32.1	12.9 <b>14.6</b> 16.7	.600 <b>.727</b> .836

### Dataset 2

Method	$\mu = 1000$	“Among” $\sigma_A = 1$	“Within” $\sigma = 25$	ICC = .0016
MME (ANOVA)	996 <b>1000</b> 1004	<b><math>[\sigma_A^2 = -5.96]</math></b>	20.7 <b>23.3</b> 26.7	<b><math>[-.097 \ -0.11 \ .137]</math></b>
MLE † (R)	<b>1000</b>	<b>0.56</b>	<b>23.1</b>	<b>.0006</b>
Bayes (R, Gibbs)	996 <b>1000</b> 1004	.033 <b>1.41</b> 6.63	20.8 <b>23.3</b> 26.1	2 e-6 <b>.009</b> † .078
Bayes*(WinBUGS, MCMC)	996 <b>1000</b> 1004	.033 <b>1.28</b> 6.28	20.8 <b>23.3</b> 26.1	2 e-6 <b>.008</b> † .070
	996 <b>1000</b> 1004	1.01 <b>5.13</b> 10.4**	20.4 <b>22.8</b> 25.7	.002 <b>.07</b> .17**

\* R-Gibbs results and results in the top row for WinBUGS use flat priors on  $\mu$ ,  $\sigma_A$ ,  $\sigma$  as R-Gibbs; the bottom row of WinBUGS uses uniform prior on ICC, and flat priors on  $\mu$  and  $\sigma$ .

‡ Because of an ill-conditioned sample variance-covariance matrix, confidence intervals are not available in R.

† Means of simulated values after burn-in are shown; the corresponding medians are .0007 (Gibbs), .0005 (BUGS).

\*\* An interval estimate that does not cover the known true parameter value.

Fig. 10. Summary of Results. Bold numbers are point estimates (when possible, with 95% intervals in plain type.)