A Tolerance Interval for Assessing the Quality of Glucose Monitoring Meters

C. T. Liao
Department of Agronomy
National Taiwan University
Taipei, Taiwan, R. O. C.

H. K. Iyer
Department of Statistics
Colorado State University
Fort Collins, CO 80523

Abstract

We propose a method for constructing a two-sided tolerance interval for the distribution $N(\theta, \sigma_1^2 - \sigma_2^2)$ based on a statistic $\hat{\theta}$ whose distribution is $N(\theta, \sum_{i=1}^{q} c_i\sigma_i^2)$, where $c_1, c_2, \cdots, c_q$ are known constants, and mutually independent scaled chi-squared estimators of $\sigma_1^2, \sigma_2^2, \cdots, \sigma_q^2$. For the special case of $q = 2$, this problem has been considered by Wang and Iyer (1994). The problem was motivated by an experimental study in which data were collected to assess the quality of blood glucose monitoring meters. We describe this study and show how the tolerance interval procedure suggested here can be used to assess the performance of the glucose monitoring meters. We also report the results of a simulation study we conducted to evaluate its frequentist coverage probability. The results indicate that the proposed method may be recommended for use in practical applications.

Keywords: chi-squared approximation; confidence intervals; measurement errors; linear models; standard reference materials; variance components.

1 Introduction

A manufacturer of health monitoring devices has developed a glucose monitoring meter for in-home use by patients with diabetes. These meters must meet certain quality requirements before they can be marketed. To test the accuracy and precision of the meters, data are collected in a gage study using the design described below.

A random sample of m meters (called test meters) is selected along with a random sample of n meters (called reference meters). The reference meters are generally very accurate and precise. Consequently, they are also expensive. The meter readings from the test meters are to be compared with corresponding readings from the reference meters in order to estimate the accuracy of the test meters. Each of the $m+n$ meters will be used to measure the glucose level in each of B blood samples selected from a pool of available blood samples. A random sample of L lots of specially coated paper strips is chosen. These strips will be dipped in the blood sample and placed in the meter. The meter will then produce an optical measurement which, after proper calibration, will be translated into a glucose concentration reading. A strip can be used to obtain only a single reading. Let X denote a measurement using a test meter and Y denote a measurement using a reference meter. Then X and Y are modeled as follows.

$$X_{ijkl} = \mu_T + M_i + B_j + L_k + e_{ijkl} \tag{1.1}$$

for $i = 1, \cdots, m$, $j = 1, \cdots, B$, $k = 1, \cdots, L$ and $l = 1, \cdots, R$, where μ_T denotes the expected reading when using a test meter, M_i the effect of test meter i, B_j the effect of the j^{th} blood sample, L_k the effect of the k^{th} strip-lot and e_{ijkl} measurement error. Likewise,

$$Y_{ijkl} = \mu_R + M_i^{'} + B_j + L_k + e_{ijkl}^{'} \tag{1.2}$$

for $i = 1, \cdots, n$, $j = 1, \cdots, B$, $k = 1, \cdots, L$ and $l = 1, \cdots, R$, where μ_R denotes the expected reading when using a reference meter, $M_i^{'}$ the effect of reference meter i, B_j the effect of the j^{th} blood sample, L_k the effect of the k^{th} strip-lot and $e_{ijkl}^{'}$ measurement error. The effects $M_i, M_i^{'}, B_j, L_k, e_{ijkl}, e_{ijkl}^{'}$ are random effects, normally distributed with zero mean and standard deviations equal to σ_T, σ_R, σ_B, σ_L, σ_e and σ_e, respectively (the variances of e_{ijkl} and $e_{ijkl}^{'}$ are assumed to be equal).

The theoretical mean for the i^{th} test meter when using blood sample j and strip-lot k is equal to $\mu_T + M_i + B_j + L_k$. The theoretical mean reading, averaging over *all* reference meters, for the same blood sample and strip-lot is equal to $\mu_R + B_j + L_k$. This theoretical mean reading is used as the reference value against which the readings from individual test meters will be compared to assess their accuracy. The deviation of the reading obtained using a single test meter from the mean over all reference meters is thus equal to $D_i = \mu_T - \mu_R + M_i$. It is the distribution of the D_i's that is of interest. These D_i have a normal distribution with mean $\mu_T - \mu_R$ and standard deviation σ_T. For the quality control objective, a tolerance interval is required for the distribution of D_i. Sample statistics L and U are sought such that with confidence level γ we can say that the area under the density of D_i between L and U is no less than β ($\gamma = 0.90$ and $\beta = 0.95$ are commonly used values in many applications). The control limits for the distribution of D_i, determined based on regulatory considerations, are $-C$ to C. A batch of test meters is deemed to have met the quality requirements if the tolerance interval $[L, U]$ is completely contained in the interval $[-C, C]$.

Discussion of tolerance intervals in fixed-effects linear models can be found in Odeh and Owen (1980) and Hahn and Meeker (1991); and for the case of one-way random-effects model in Lemon (1977), Hahn (1982), Jaech (1984), Mee and Owen (1983), Mee (1984), Wang (1988), Beckman and Tietjen (1989) and Vangel (1992). But it turns out that no satisfactory tolerance interval procedure is available in the literature for the problem being considered in this paper. Wang and Iyer (1994) developed a tolerance interval procedure for a related, but simpler, problem. Brown et al. (1997) applied the results of Wang and Iyer

(1994) to evaluate the bioequivalence of two formulations of a drug using various cross-over designs for data collection.

Although the blood glucose monitor example was the motivating reason for considering the tolerance interval problem stated above, we formulate the problem in a slightly more general setting and obtain an approximate procedure based on the Wang-Iyer method in the next section. Then we revisit the blood glucose monitor example in section 3. In section 4, we report the results of a simulation study we conducted to assess the coverage properties of the proposed tolerance interval procedure.

2 A Tolerance Interval Problem in the Presence of Several Variance Components

As mentioned earlier, Wang and Iyer (1994) developed a (β, γ)-tolerance interval for the distribution $N(\theta, \sigma_1^2 - \sigma_2^2)$ based on the mutually independent statistics $\hat{\theta}$, S_1^2 and S_2^2, where $\hat{\theta} \sim N(\theta, c\sigma_1^2)$, c is a known constant, $n_1 S_1^2/\sigma_1^2$ is a chi-squared random variable with n_1 df (degrees of freedom) and $n_2 S_2^2/\sigma_2^2$ is a chi-squared random variable with n_2 df. We generalize their problem as follows.

We seek a (β, γ)-tolerance interval for a random variable W whose distribution is $N(\theta, \sigma_1^2 - \sigma_2^2)$. Suppose there are mutually independent statistics $\hat{\theta}$, S_1^2, S_2^2, $\cdots$, S_q^2, where $\hat{\theta}$ is normally distributed with mean θ and variance σ^2, with σ^2 given by

$$\sigma^2 = c_1\sigma_1^2 + c_2\sigma_2^2 + \cdots + c_q\sigma_q^2 = \sum_{i=1}^{q} c_i\sigma_i^2,$$

c_i are known constants, and $n_i S_i^2/\sigma_i^2$ are independent chi-squared random variables with n_i df, for $i = 1, 2, \cdots, q$. The problem considered by Wang and Iyer (1994) is a special case with $q = 2$.

Employing arguments similar to those in Wang and Iyer (1994), we obtain the following (β, γ)-tolerance interval for W.

$$\hat{\theta} \pm max\{\hat{k}\sqrt{max(0, S_1^2 - S_2^2 F_0)}, t_{\frac{1+\gamma}{2};\hat{f}} S_p\} \tag{2.1}$$

where

$$F_0 = F_{\frac{1-\gamma}{3};n_1,n_2} \qquad \hat{k} = \hat{u}\hat{\lambda}\sqrt{\frac{\hat{D}}{\chi^2_{1-\gamma;\hat{D}}}} \qquad \hat{u} = Z_{\frac{1+\beta}{2}}\sqrt{1+\frac{1}{\hat{\psi}}}$$

$$\hat{\lambda} = \sqrt{\frac{\hat{\phi}}{1-(1-\hat{\phi})F_0}} \qquad \hat{D} = \frac{(1-(1-\hat{\phi})F_0)^2}{\frac{1}{n_1}+\frac{(1-\hat{\phi})^2F_0^2}{n_2}} \qquad \hat{\psi} = \frac{\hat{\phi}}{c_1+c_2(1-\hat{\phi})+\sum_{i=3}^{q}\frac{c_iS_i^2}{S_1^2}}.$$

$$\hat{\phi} = max(0, \frac{S_1^2-S_2^2F_0}{S_1^2}) \quad \hat{f} = \frac{(c_{12}S_{12}^2+\sum_{i=3}^{q}c_iS_i^2)^2}{\frac{(c_{12}S_{12}^2)^2}{n_{12}}+\sum_{i=3}^{q}\frac{(c_iS_i^2)^2}{n_i}} \quad S_p = \sqrt{c_{12}S_{12}^2+\sum_{i=3}^{q}c_iS_i^2}$$

and

$$c_{12} = c_1 + c_2, \quad n_{12} = n_1 + n_2, \quad S_{12}^2 = \frac{n_1S_1^2 + n_2S_2^2}{n_{12}}.$$

Here $F_{\alpha;n_1,n_2}$ is the α percentile of an F distribution with n_1 and n_2 df; Z_α is the α percentile of the standard normal distribution; $\chi^2_{\alpha;\hat{D}}$ is the α percentile of a chi-squared distribution with $\hat{D}$ df and $t_{\alpha;\hat{f}}$ is the α percentile of a Student-t distribution with $\hat{f}$ df.

3 The Blood Monitor Example

We now apply the tolerance interval given in Section 2 to the glucose monitor problem. Let $\bar{X}$, MS_T and MSE_T denote the sample mean, the mean square of the test meter effect and the error mean square, respectively, for model (1.1), i.e.,

$$\bar{X} = \frac{1}{mBLR}\sum_{i=1}^{m}\sum_{j=1}^{B}\sum_{k=1}^{L}\sum_{l=1}^{R}X_{ijkl}$$

$$MS_T = B\,L\,R\sum_{i=1}^{m}(\bar{X}_{i...} - \bar{X})^2$$

$$MSE_T = \sum_{i=1}^{m}\sum_{j=1}^{B}\sum_{k=1}^{L}\sum_{l=1}^{R}(X_{ijkl} - \bar{X}_{i...} - \bar{X}_{.j..} - \bar{X}_{..k.} + 2\bar{X})^2.$$

Similarly, let $\bar{Y}$, MS_R and MSE_R denote the corresponding sample mean, mean square for the reference meter effect and error mean square for model (1.2).

Then the statistics $\bar{X}$, $\bar{Y}$, MS_T, MSE_T, MS_R and MSE_R are mutually independent. Let $\sigma_{1T}^2 = k_0\sigma_T^2 + \sigma_e^2$ and $\sigma_{1R}^2 = k_0\sigma_R^2 + \sigma_e^2$ where $k_0 = BLR$. It follows that $(m-1)MS_T/\sigma_{1T}^2 \sim \chi^2_{m-1}$, $(n-1)MS_R/\sigma_{1R}^2 \sim \chi^2_{n-1}$, $v_1MSE_T/\sigma_e^2 \sim \chi^2_{v_1}$ and $v_2MSE_R/\sigma_e^2 \sim \chi^2_{v_2}$, where $v_1 =$

$mk_0 - m - B - L + 2$ and $v_2 = nk_0 - n - B - L + 2$. We further pool MSE_T with MSE_R to get

$$MSE = \frac{v_1 MSE_T + v_2 MSE_R}{v_1 + v_2}.$$

So we have $vMSE/\sigma_e^2 \sim \chi_v^2$ with $v = v_1 + v_2$. The distribution of D_i we are interested in is $N(\mu_T - \mu_R, \sigma_T^2)$. Also observe that $\bar{X} - \bar{Y} \sim N(\mu_T - \mu_R, \frac{\sigma_{1T}^2}{mk_0} + \frac{\sigma_{1R}^2}{nk_0})$.

In the tolerance interval framework of section 2, $W = D_i$, $\theta = \mu_T - \mu_R$, $\sigma_1^2 = \sigma_T^2 + \sigma_e^2/k_0 = \sigma_{1T}^2/k_0$, $\sigma_2^2 = \sigma_e^2/k_0$, $\hat{\theta} = \bar{X} - \bar{Y}$ and $\sigma^2 = \sigma_{1T}^2/mk_0 + \sigma_{1R}^2/nk_0 = \sigma_1^2/m + \sigma_3^2/n$, where $\sigma_3^2 = \sigma_{1R}^2/k_0$. Therefore, we have $\sigma^2 = c_1\sigma_1^2 + c_2\sigma_2^2 + c_3\sigma_3^2$, with $c_1 = 1/m$, $c_2 = 0$ and $c_3 = 1/n$. The estimators for the variances σ_1^2, σ_2^2 and σ_3^2 are $S_1^2 = MS_T/k_0$, $S_2^2 = MSE/k_0$ and $S_3^2 = MS_R/k_0$, respectively. Also $n_1 = m - 1$, $n_2 = v$ and $n_3 = n - 1$.

An experimental data set was provided to us by a manufacturer of health monitoring devices. For proprietary reasons we are unable to make the data set publicly available. However, the essential summary statistics are given below. For this data set, $m = 44$, $n = 10$, $B = L = R = 3$. Analysis of the data results in $\hat{\theta} = -1.13654$ and the required variance estimates are $S_1^2 = 0.61928$, $S_2^2 = 0.19052$ and $S_3^2 = 0.63132$.

For $\beta = 0.95$ and $\gamma = 0.90$, we get $F_{0.10/3;43,1362} = 0.63974$, $\hat{\phi} = 0.80319$ and $\hat{k} = 2.41427$. The (95% content, 90% confidence)-tolerance interval for the distribution of D_i is obtained as $[-2.83923, 0.56615]$. We can then state that we are 90 percent confidence that at least 95 percent of the differences between an individual test meter and the average of all reference meters (which is used as the reference value relative to which the accuracy of an individual meter is assessed) fall between these two limits. Thus if one agrees that a batch of test meters meets the quality requirement if a (95% content, 90% confidence)-tolerance interval is completely contained in the interval [-5,5], then, for this example, one concludes that the batch of test meters has satisfied this quality (accuracy and precision) requirement.

4 A Simulation Study

To evaluate the true confidence coefficient of the constructed tolerance intervals, the following simulation study was carried out based on the glucose monitor example. Without loss of generality we may assume that $\theta = 0$ and $\sigma_R = 1$. For fixed $B = L = R = 3$ and specified values of m, n, σ_T and σ_e, we generated a normal random deviate Z from the distribution $N(0, \sigma^2)$ and three chi-squared random deviates U_1, U_2 and U_3 with n_1, n_2, n_3 df, using the functions RNORM and RCHISQ, respectively, in the statistical package S-PLUS. The corresponding sample statistics $S_1^2 = n_1 U_1/\sigma_1^2$, $S_2^2 = n_2 U_2/\sigma_2^2$, $S_3^2 = n_3 U_3/\sigma_3^2$ are then generated. We then computed the quantity

$$max\{\hat{k}\sqrt{max(0, S_1^2 - S_2^2 F_0)}, t_{\frac{1+\gamma}{2};\hat{f}} S_p\}$$

which, for ease of notation, we shall denote by ME (*margin of error*). Let

$$p = \Phi(Z + ME) - \Phi(Z - ME)$$

where $\Phi(\cdot)$ is the standard normal distribution function. The procedure was repeated 10,000 times and the fraction of times that p was greater than or equal to β was computed. The results are presented in the following tables.

Table 1. Simulated confidence coefficients (times 10^4) for the ($\beta = 0.95, \gamma = 0.90$) tolerance interval, with $m = 5, 10$; $B = L = R = 3$ and $\sigma_R = 1$.

		m											
		5						10					
		n						n					
σ_T	σ_e	5	10	20	40	60	80	5	10	20	40	60	80
	0.5	9732	9737	9591	9412	9242	9183	9369	9480	9402	9229	9131	9162
	1	9701	9714	9541	9294	9242	9191	9330	9459	9302	9168	9145	9133
0.5	2	9622	9607	9431	9310	9270	9197	9339	9383	9302	9164	9077	9103
	4	9601	9536	9408	9300	9305	9249	9410	9362	9301	9197	9191	9247
	8	9718	9677	9654	9595	9560	9572	9588	9548	9467	9316	9308	9320
	0.5	9518	9403	9205	9091	9054	9086	9316	9148	9129	9077	9057	9005
	1	9558	9372	9220	9075	9059	9099	9334	9255	9109	9086	8979	9014
1	2	9531	9378	9196	9118	9111	9136	9274	9236	9097	9049	9016	8993
	4	9500	9386	9258	9148	9150	9068	9279	9214	9128	9094	9070	9059
	8	9515	9403	9368	9314	9209	9230	9361	9249	9259	9199	9175	9163
	0.5	9186	9034	9017	9031	9005	9049	9110	9094	9073	9008	9005	9065
	1	9179	9086	9060	9003	9050	9033	9125	9086	9020	8966	8985	8987
2	2	9246	9134	9113	9002	9054	9053	9148	9053	9022	9033	8987	9007
	4	9248	9191	9098	9090	9072	9080	9164	9110	9078	9025	8983	9030
	8	9336	9229	9138	9126	9119	9070	9237	9126	9095	9095	9101	9024
	0.5	9077	9051	9003	9006	9029	9014	9056	9012	8994	9006	8991	9026
	1	9046	9050	9008	9025	9017	9013	8976	9010	8949	9007	9050	9000
4	2	9092	9005	8977	9090	9045	9003	9009	8957	9011	9002	9040	9004
	4	9127	9064	9078	9012	9047	9019	9081	9050	8963	8995	8968	9011
	8	9120	9092	9090	9035	9064	9035	9056	9044	9013	9028	9015	9022

Table 2. Simulated confidence coefficients (times 10^4) for the ($\beta = 0.95, \gamma = 0.90$) tolerance interval, with $m = 25, 50$; $B = L = R = 3$ and $\sigma_R = 1$.

		m											
		25						50					
		n						n					
σ_T	σ_e	5	10	20	40	60	80	5	10	20	40	60	80
	0.5	8689	9012	9077	9073	9068	9112	8139	8517	8876	8956	8969	8983
	1	8747	8970	9069	9101	9019	9067	8229	8636	8884	8970	8982	9050
0.5	2	8835	9093	9109	9098	9098	9036	8314	8748	8970	9018	8998	9026
	4	9012	9179	9176	9095	9147	9116	8686	8998	8977	9043	9094	9092
	8	9381	9417	9276	9279	9302	9218	9111	9245	9176	9234	9198	9239
	0.5	8976	9052	9010	9082	8997	8986	8669	8994	9024	8996	8994	9000
	1	9002	9079	9030	9007	9043	8991	8711	8957	8996	8980	8981	8950
1	2	9037	9106	9082	8998	8966	8995	8708	8943	9004	8989	9016	8990
	4	9037	9091	9073	9078	9024	9010	8670	8962	8991	9000	9056	8958
	8	9063	9207	9106	9118	9116	9081	8858	9005	9090	9060	9047	9078
	0.5	9054	9014	8988	8979	8989	8963	9012	9002	8962	8991	8978	9017
	1	8997	8945	9014	9016	8985	9031	8906	8984	9062	8960	8944	8955
2	2	9031	9026	9064	9015	9019	8942	8978	8969	9025	9020	9003	8945
	4	9030	9034	9020	9017	8998	8946	8927	9010	9004	8944	8998	8952
	8	9055	9099	9100	9033	9015	9033	8862	8933	8990	8999	9012	8978
	0.5	9008	8985	9022	8992	9030	8975	8985	8969	9015	8993	8961	9044
	1	8990	9008	8945	9010	8991	9036	8957	9028	8968	8991	9017	8992
4	2	8999	9025	9023	9021	8993	8964	9022	9013	9037	8989	8988	8944
	4	9020	9018	8963	9031	8978	8933	9025	8953	8976	8997	8955	8996
	8	9034	9011	8983	9025	8955	8986	8996	9003	9035	8930	8979	8972

Table 3. Simulated confidence coefficients (times 10^4) for the ($\beta = 0.95, \gamma = 0.90$) tolerance interval, with $m = 75, 100$; $B = L = R = 3$ and $\sigma_R = 1$.

		m											
		75						100					
		n						n					
σ_T	σ_e	5	10	20	40	60	80	5	10	20	40	60	80
	0.5	8009	8388	8698	8897	8953	8934	7795	8216	8518	8808	8887	8996
	1	8051	8451	8742	8854	8894	8958	7843	8213	8514	8809	8896	8954
0.5	2	8174	8565	8723	8969	8984	8977	7999	8337	8666	8861	8882	8976
	4	8457	8767	8924	8978	9022	9038	8296	8626	8908	8981	9052	8993
	8	8935	9070	9096	9189	9178	9173	8841	9017	9055	9127	9141	9154
	0.5	8500	8865	9001	8973	8994	9015	8334	8766	8875	8993	9018	8988
	1	8485	8814	8957	8954	8999	9000	8272	8767	8949	8935	8941	8975
1	2	8484	8792	8941	8973	9035	9010	8319	8738	8925	8962	9027	9049
	4	8511	8759	8974	8976	9032	9057	8369	8734	8861	9015	8988	8999
	8	8646	8928	9044	9077	9081	9078	8595	8821	8977	9062	9004	9030
	0.5	8876	8978	8963	9016	8957	9069	8840	9027	9044	9002	9022	9003
	1	8924	9006	8933	9021	9017	9007	8845	9021	8958	8983	8979	8972
2	2	8925	9035	9037	9021	8999	8978	8808	8957	9000	8981	9025	9009
	4	8844	8986	8994	8963	8990	8962	8756	8960	8986	8952	9016	8995
	8	8749	8951	8983	9007	8983	8990	8640	8912	8961	9022	8981	8998
	0.5	8970	8997	8997	9031	9007	8990	8981	9016	9008	8985	8980	8996
	1	9048	9011	8952	8983	8976	8981	8983	9015	9037	9015	8995	8962
4	2	9000	9038	8954	9008	8967	9011	8970	8937	9000	8978	8998	9001
	4	8981	9045	9011	9024	8997	8990	8967	8994	9009	8978	9000	9005
	8	8928	8991	9005	9005	8967	8976	8927	9011	9018	8992	9039	9054

For most parameter combinations, the constructed tolerance intervals are successful in maintaining the confidence level close to the stated value of $\gamma = 0.90$. Nonetheless, the results indicate that when σ_T is smaller than σ_R, and both n and m are small, the proposed tolerance interval can be conservative. On the other hand, when σ_T is smaller than σ_R, and n is small but m is large, the proposed tolerance interval appears to be somewhat liberal. Fortunately, in most practical situations, σ_T is usually larger than σ_R because the reference meters tend to have much higher precision than the test meters. If the number of reference meters is at least 5, then the results indicate that the proposed approach can be satisfactory for practical use.

5 Concluding Remarks

In this paper we have generalized the tolerance interval method of Wang and Iyer (1994) and broadened the range of its application. Whereas the performance of the proposed interval method appears to be satisfactory for many practical situations, both sets of simulation results – those given in this paper and those given in the Wang and Iyer (1994) paper – indicate that there is room for improvement of these methods. In particular, it may be possible to modify the tolerance factor k and get a tolerance interval whose confidence coefficient is closer to the stated confidence level for a wider range of values of the model parameters. This is a topic for future research.

References

[1] Beckman, R. J. and Tietjen, G. L. (1989). Two-sided tolerance limits for balanced random-effects ANOVA models. *Technometrics*, **31**, 185-197.

[2] Brown, E. B., Iyer, H. K. and Wang, C. M.(1997). Tolerance intervals for assessing individual bioequivalence. *Statistics in Medicine*, **16**, 803-820.

[3] Hahn, G. J. (1982). Removing measurement error in assessing conformance to specifications. *Journal of Quality Technology*, **14**, 117-121.

[4] Hahn, G. J. and Meeker, W. Q. (1991). Statistical intervals, New York: John Wiley.

[5] Jaech, J. L. (1984). Removing the effects of measurement errors in constructing statistical tolerance intervals. *Journal of Quality Technology*, **16**, 69-73.

[6] Lemon, G. H. (1977). Factors for one-sided tolerance limits for balanced one-way ANOVA random-effects model. *Journal of the American Statistical Association*, **72**, 676-680.

[7] Mee, R. W.(1984). β-expectation and β-content tolerance limits for balanced one-way ANOVA random-effects model. *Technometrics*, **26**, 251-254.

[8] Mee, R. W. and Owen, D. W. (1983). Improved factors for one-sided tolerance limits for balanced one-way ANOVA random-effects model. *Journal of the American Statistical Association*, **78**, 901-905.

[9] Vangel, M. G. (1992). New methods for one-sided tolerance limits for a one-way balanced ANOVA random-effects model. *Technometrics*, **34**, 176-186.

[10] Wang, C. M.(1988). β-expection tolerance limits for balanced one-way random-effects model. *Probability and Statistics: Essays in Honor of Franklin A. Graybill, ed. J. N. Srivastava*, Amsterdam: North-Holland, 285-295.

[11] Wang, C. M. and Iyer, H. K. (1994). Tolerance interval for the distribution of true values in the presence of measurement errors. *Technometrics*, **36**, 162-170.

sets of predictor values, i.e., each of the N! possible order associations has the same chance for each of the L shuffles (H_0). Then an approximate P-value based on the L randomly obtained T's under H_0 is given by

$$\frac{\text{number of L T's} \geq T_0}{L}$$

The fact that the response and predictor values are decoupled is obvious for the Manly variation and is also true for the other two variations. The Cade and Richards (1996) variation is similar to the Manly variation except that the N y's used in the L shuffles are replaced with the N residuals resulting from the reduced model (H_0) for the observed data. Finally the ter Braak (1992) variation is also similar to the Manly variation except that the N y's used in the L shuffles are replaced with the N residuals resulting from the full model (H_1) for the observed data. While Freedman and Lane (1983) anticipated the Manly (1991) and Cade-Richards (1996) variations, they did not suggest utilization of the ter Braak (1992) variation.