Conference on Applied Statistics in Agriculture 1992 - 4th Annual Conference Proceedings

CO-EFFECT ANALYSIS OF VARIANCE: A NEW METHOD FOR UNBALANCED DATA

Andre Plante

CO-EFFECT ANALYSIS OF VARIANCE: A NEW METHOD FOR UNBALANCED DATA

André PLANTE

Département de mathématiques et d'informatique, *Université du Québec à Montréal*

ABSTRACT

For fixed-effect models one can always, according to the Gauss-Markov Theorem, uniquely determine independent variables called source identifiers, each corresponding to a source of variation. When linearly combined, source identifiers can generate all possible expected values for the response variable. The co-effect method uses regression of the response variable on source identifiers. Corresponding regression coefficients are, by definition, unbiased estimates of co-effects, and satisfy the same restrictions as those imposed on main effects and interaction effects in standard analysis of variance. With balanced data, co-effect analysis gives results identical to those of the standard method; with unbalanced data, however, results can be significantly different.

An example is given where predicted genetic interaction can be easily observed using the co-effect method ($P \approx 10^{-14}$) while Yates' weighted-squares-of-means method does not detect any interaction effects ($P > 0.1$).

Key words and phrases: Unbalanced data, analysis of variance, interaction, co-effects, genetic experiment.

1. INTRODUCTION

The co-effect method suggested can be used to supplement Yates' weighted-squares-of-means analysis of variance for fixed-effect models since with balanced data it gives results identical to those of standard methods, whereas with unbalanced data this technique can be significantly more sensitive because the cell-size-dependent extended parameters used, the co-effects, differ from the usual effects. This new method will be illustrated using complete two-way layouts *i.e.* without missing treatment combinations.

The Gautschi-Scheffé proof for the Gauss-Markov Theorem (Scheffé 1959, Section 1.4) associates each estimable parameter δ with a unique and distinct independent variable $\mathbf{V}$ lying in the subspace of possible values for the expected response $\boldsymbol{\mu}$. ($\mathbf{V}$, by definition, satisfies the equation $\boldsymbol{\mu}'\mathbf{V}=\delta$.) In a fixed effect model, a *source identifier variable* $\mathbf{V}_j^*$ is a scaled

Gautschi-Scheffé independent variable $\mathbf{V}_j$ associated with a given effect δ_j. An algorithm to obtain the $\mathbf{V}_j$'s is given in Section 3. The corresponding *co-effect* δ_j^* is the coefficient of $\mathbf{V}_j^*$ when μ is expressed as a linear combination

$$\mu=\Sigma_j \delta_j^* \mathbf{V}_j^*$$

of the source identifier variables for all the parameters of the model. It is assumed that the usual sum-to-zero ANOVA restrictions are imposed on the δ_j's, which implies that the same restrictions hold for the δ_j^*'s and that both effects and co-effects are uniquely defined and are estimable. The source identifier variable $\mathbf{V}_j^*$ corresponding to a Gautschi-Scheffé independent variable $\mathbf{V}_j$ and to an effect δ_j is, by definition, $\mathbf{V}_j^*=\|\mathbf{V}_j\|^{-2}\mathbf{V}_j$, where $\|\mathbf{V}_j\|$ is the length of $\mathbf{V}_j$. This scaling process ensures that the projection of μ on $\mathbf{V}_j^*$ is $\delta_j\mathbf{V}_j^*$ and that effects and corresponding co-effects are identical in an orthogonal design. Orthogonality between source identifier variables, as used here, implies non-correlation between associated BLU estimators.

In the foster-nursing genetic experiment example discussed in Section 2, which uses Bailey's data listed in Scheffé (1959, p.140), one might expect strong positive interaction effects due to genotype correspondence between foster-mothers and litters, but Yates' weighted-squares-of-means method (SAS GLM Type III analysis) does not detect any such interaction effect (P>0.1). However, the co-effect method shows that extremely strong interaction co-effects are present ($P\approx10^{-14}$). Through resampling, 95% confidence intervals are established for all the co-effects, and, as might be expected, interaction co-effects are seen to be higher when foster-mother and litter share the same genotype. Moreover, if an almost-outlier found in Bailey's data is pulled back, interaction effects then become significant when Yates' method is used. It is even possible to modify the almost-outlier and to inflate intra-cell variances in such a way that the models themselves are no longer significant and Yates' method will detect no effect, while interaction co-effects remain highly significant (P<10^{-8}). This genetic example is followed, in Section 3, by an elementary formal example for illustrative purposes.

Yates' weighted-squares-of-means method (Yates 1934, Federer and Zelen 1966) can be deduced from the general likelihood ratio test (Mann 1949, Chapter X, Scheffé 1959, Section 1.5 and Section 4.4, Graybill 1976, Section 14.8) for testing cell-size-independent null hypotheses concerning main effects and interaction effects. This method is recommended by Francis (1972), Kutner (1974), Nelder (1974), and by Milliken and Johnson (1984, p.158), is reviewed by Steel, Hocking, and Hackney (1978), and is already implemented in SAS, in SPSS, and in BMDP. The co-effect method can provide a new and efficient stepwise regression algorithm (not given in this paper) for applying Yates' method to various unbalanced designs. Also, since the co-effect method is as general as Yates' weighted-squares-of-means method, both methods are theoretically applicable to all fixed effect non-orthogonal designs using blocks of unequal sizes, for example, the designs advocated by Mead (1990).

2. A GENETIC EXAMPLE

Scheffé (1959, p. 140) reports data displayed in a complete two-way layout from an unbalanced foster-nursing genetic experiment (Bailey 1953) with four hybrid female rat genotypes. The factors are foster-mother genotype and litter genotype. Each of the 61 responses indicates an *average litter weight i.e.* the average weight of female baby rats in each foster-mother's litter at 28 days. Cell sizes, which vary from 2 to 5, are given in Table 1.

Yates' weighted-squares-of-means analysis of variance results for Bailey's data, given in Table 2, are based on the SAS GLM procedure with Type III sums-of-squares (SAS Institute Inc. 1988, Chapter 9 and Chapter 20). Connections between SAS Type III analysis and Yates' weighted-squares-of-means analysis of variance are given in Speed, Hocking, and Hackney (1978) and in Milliken and Johnson (1984, Chapter 10). No genetic interaction effects are detected ($P \approx 0.120$), and foster-mother genotype main effects seem to exist ($P \approx 0.011$).

Table 3 gives the results of a formal co-effect analysis of variance for Bailey's data supplementing the results of Yates' weighted-squares-of-means analysis. The sums of squares reported for the various groups of co-effects are calculated using backward stepwise regression over the source identifier variables for each group in turn. The computation method is described in Plante (1992). Both the co-effect method and Yates' weighted-squares-of-means method yield the same error-sum-of-squares. However, every co-effect group is seen to contain non-zero members. The very strong interaction co-effects detected ($P \approx 1.01 \times 10^{-14}$) were further examined. *P*-values were checked by resampling. The null distribution for the interaction co-effects *F*-statistic is approximately the same when the distribution of residuals is used as error distribution. When the error distribution in any given cell is assumed to be normally distributed with a standard deviation equal to the observed standard deviation for that cell in Bailey's experiment, we find that the expected value for the interaction co-effect *F*-statistic, under the associated null hypothesis, is increased by about 20%. Therefore, the extreme co-effects detected are not an artifax resulting from breaking model assumptions.

Table 4 is made up of four subtables giving the estimated main co-effects for litter genotype, the main co-effects for foster-mother genotype, the genotype interaction co-effects, and the general mean co-effect. Conservative approximate 95% confidence intervals are given. These intervals are mutually consistent estimates (Plante 1991) for each co-effect. They are based on two small resampling experiments. The first experiment used the distribution of adjusted residuals as error distribution with resampling size 200; the second used resampling size 1000, and normally distributed error random variables with the standard deviation varying from cell to cell according to the observed standard deviations in Bailey's experiment. Interaction co-effects are positive and important when the foster-mother and the litter genotype are the same — just as one might expect in Bailey's experiment.

The data point y=68.0 in the cell "Litter Genotype=I and Foster-Mother Genotype=A" is significant at the 1% level according to the Studentized Residual Test for a single outlier from a normal distribution (Lund 1975). Since both the normality assumption and the equality of variance assumption used in that test are unwarranted here, we can conclude only that y=68.0 is an almost-outlier. If this almost-outlier is pulled back to the value y=48.0 near the mean 47.10 of the three original values in that cell, one obtains, using SAS GLM Type III sums-of-squares, the results summarized in Table 5 which show that genetic interaction effects are now apparent ($P\approx 0.005$). Co-effect analysis of variance results for the same data are given in Table 6 where we can see that F-statistics for co-effects, as expected, now have quite extreme values. One interpretation of these results is that the presence of easily observable co-effects is indirect evidence of the existence of real effects.

Tables 7, 8, and 9 report results from a confirmatory analysis — performed with modified Bailey data — aimed at completely concealing effects while leaving co-effects visible. The almost-outlier is pulled back as before, while intra-cell variances are inflated. As a result, every effect is now masked from a SAS GLM Type III ANOVA. Residuals do not give the impression that effects could be masked by outliers. Co-effect analysis, however, still indicates significant interaction co-effects ($P\approx 3.90\text{X}10^{-9}$). (This significance of co-effects might be found surprising, since neither model is significant — $P\approx 0.391$). Co-effect ANOVA is probably the only technique that would lead one to suspect, from the data shown in Table 7, that there might be some hidden effect. This confirmatory analysis leads me to believe that co-effect analysis is useful to supplement Yates' weighted-squares-of-means test for real effects.

3. AN ELEMENTARY FORMAL EXAMPLE

The 2X2 layout used in *SAS/STAT™ User's Guide* (1988, p.556) to illustrate the SAS GLM procedure is used here to explain how to construct source identifier variables on which the co-effect method is based.

Consider the two-way layout

		B = 1	B = 2
A	1	12 14	11 9
	2	20 18	17

which, when displayed in serial form, is

$$\mathbf{y} = \begin{bmatrix} y_{111} \\ y_{112} \\ y_{121} \\ y_{122} \\ y_{211} \\ y_{212} \\ y_{221} \end{bmatrix} = \begin{bmatrix} 12 \\ 14 \\ 11 \\ 9 \\ 20 \\ 18 \\ 17 \end{bmatrix} .$$

Cell expected responses and their estimates are

$$\mu = \begin{bmatrix} \mu_{11} \\ \mu_{11} \\ \mu_{12} \\ \mu_{12} \\ \mu_{21} \\ \mu_{21} \\ \mu_{22} \end{bmatrix} \quad \text{and} \quad \mathbf{y}_{\cdot} = \begin{bmatrix} y_{11\cdot} \\ y_{11\cdot} \\ y_{12\cdot} \\ y_{12\cdot} \\ y_{21\cdot} \\ y_{21\cdot} \\ y_{22\cdot} \end{bmatrix} = \begin{bmatrix} 13 \\ 13 \\ 10 \\ 10 \\ 19 \\ 19 \\ 17 \end{bmatrix} .$$

The variables

$$\mathbf{J}^{*}_{\cdot\cdot} = 0.8 \begin{bmatrix} 1 \\ 1 \\ 1 \\ 1 \\ 1 \\ 1 \\ 2 \end{bmatrix}, \quad \mathbf{U}^{*}_{1\cdot} = 0.8 \begin{bmatrix} 1 \\ 1 \\ 1 \\ 1 \\ -1 \\ -1 \\ -2 \end{bmatrix}, \quad \mathbf{U}^{*}_{\cdot 1} = 0.8 \begin{bmatrix} 1 \\ 1 \\ -1 \\ -1 \\ 1 \\ 1 \\ -2 \end{bmatrix}, \quad \mathbf{U}^{*}_{11} = 0.8 \begin{bmatrix} 1 \\ 1 \\ -1 \\ -1 \\ -1 \\ -1 \\ 2 \end{bmatrix},$$

$$\mathbf{U}^{*}_{2\cdot} = -\mathbf{U}^{*}_{1\cdot}, \quad \mathbf{U}^{*}_{\cdot 2} = -\mathbf{U}^{*}_{\cdot 1}, \quad \mathbf{U}^{*}_{12} = -\mathbf{U}^{*}_{11}, \quad \mathbf{U}^{*}_{21} = -\mathbf{U}^{*}_{11}, \quad \mathbf{U}^{*}_{22} = \mathbf{U}^{*}_{11}$$

are source identifier variables. $\mathbf{J}^{*}_{\cdot\cdot}$ identifies the general mean co-effect; $\mathbf{U}^{*}_{1\cdot}$ and $\mathbf{U}^{*}_{2\cdot}$ identify factor A main co-effects; $\mathbf{U}^{*}_{\cdot 1}$ and $\mathbf{U}^{*}_{\cdot 2}$ identify factor B main co-effects. $\mathbf{U}^{*}_{11}$, $\mathbf{U}^{*}_{12}$, $\mathbf{U}^{*}_{21}$, $\mathbf{U}^{*}_{22}$ are interaction co-effect identifiers. Co-effects are defined by the equation

$$\mu = \mu^{*}\mathbf{J}^{*}_{\cdot\cdot} + \alpha^{*}_{1}\mathbf{U}^{*}_{1\cdot} + \alpha^{*}_{2}\mathbf{U}^{*}_{2\cdot} + \beta^{*}_{1}\mathbf{U}^{*}_{\cdot 1} + \beta^{*}_{2}\mathbf{U}^{*}_{\cdot 2} + \tau^{*}_{11}\mathbf{U}^{*}_{11} + \tau^{*}_{12}\mathbf{U}^{*}_{12} + \tau^{*}_{21}\mathbf{U}^{*}_{21} + \tau^{*}_{22}\mathbf{U}^{*}_{22}$$

with the restrictions

$$\alpha^{*}_{1} + \alpha^{*}_{2} = 0, \quad \beta^{*}_{1} + \beta^{*}_{2} = 0, \quad \tau^{*}_{11} + \tau^{*}_{12} = 0,$$

$$\tau^{*}_{21} + \tau^{*}_{22} = 0, \quad \tau^{*}_{11} + \tau^{*}_{21} = 0, \quad \tau^{*}_{12} + \tau^{*}_{22} = 0 .$$

Source identifier variables, according to a corollary of the Gauss-Markov Theorem (Scheffé 1959, Chapter 1), are the only variables remaining constant within each experimental group such that the regression of $\mathbf{y}$ on each source identifier variable separately is $y_{\cdots}\mathbf{J}^{*}_{\cdot\cdot}$, $(y_{i\cdot\cdot} - y_{\cdots})\mathbf{U}^{*}_{i\cdot}$, $(y_{\cdot j\cdot} - y_{\cdots})\mathbf{U}^{*}_{\cdot j}$, and $(y_{ij\cdot} - y_{i\cdot\cdot} - y_{\cdot j\cdot} + y_{\cdots})\mathbf{U}^{*}_{ij}$ $(i, j = 1, 2)$ respectively.

To obtain these identifier variables, we can proceed in stages. First, we define experimental group averaging variables

$$\mathbf{J}_{11} = \begin{bmatrix} \frac{1}{2} \\ \frac{1}{2} \\ 0 \\ 0 \\ 0 \\ 0 \\ 0 \end{bmatrix}, \quad \mathbf{J}_{12} = \begin{bmatrix} 0 \\ 0 \\ \frac{1}{2} \\ \frac{1}{2} \\ 0 \\ 0 \\ 0 \end{bmatrix}, \quad \mathbf{J}_{21} = \begin{bmatrix} 0 \\ 0 \\ 0 \\ 0 \\ \frac{1}{2} \\ \frac{1}{2} \\ 0 \end{bmatrix}, \quad \mathbf{J}_{22} = \begin{bmatrix} 0 \\ 0 \\ 0 \\ 0 \\ 0 \\ 0 \\ 1 \end{bmatrix},$$

such that $y_{11\cdot}=\mathbf{y'J}_{11}=13$, $y_{12\cdot}=\mathbf{y'J}_{12}=10$, $y_{21\cdot}=\mathbf{y'J}_{21}=19$, and $y_{22\cdot}=\mathbf{y'J}_{22}=17$. Then, we define line, column, and general averaging variables $\mathbf{J}_{1\cdot}=\frac{1}{2}(\mathbf{J}_{11}+\mathbf{J}_{12})$, $\mathbf{J}_{2\cdot}=\frac{1}{2}(\mathbf{J}_{21}+\mathbf{J}_{22})$, $\mathbf{J}_{\cdot 1}=\frac{1}{2}(\mathbf{J}_{11}+\mathbf{J}_{21})$, $\mathbf{J}_{\cdot 2}=\frac{1}{2}(\mathbf{J}_{12}+\mathbf{J}_{22})$, $\mathbf{J}_{\cdot\cdot}=\frac{1}{2}(\mathbf{J}_{1\cdot}+\mathbf{J}_{2\cdot})=\frac{1}{2}(\mathbf{J}_{\cdot 1}+\mathbf{J}_{\cdot 2})$, with $y_{1\cdot\cdot}=\mathbf{y'J}_{1\cdot}=11.5$, $y_{2\cdot\cdot}=\mathbf{y'J}_{2\cdot}=18$, $y_{\cdot 1\cdot}=\mathbf{y'J}_{\cdot 1}=16$, $y_{\cdot 2\cdot}=\mathbf{y'J}_{\cdot 2}=13.5$, and $y_{\cdot\cdot\cdot}=\mathbf{y'J}_{\cdot\cdot}=14.75$. Next, we define variables for calculating main effects and interaction effects: $\mathbf{U}_{1\cdot}=\mathbf{J}_{1\cdot}-\mathbf{J}_{\cdot\cdot}$; $\mathbf{U}_{2\cdot}=\mathbf{J}_{2\cdot}-\mathbf{J}_{\cdot\cdot}$; $\mathbf{U}_{\cdot 1}=\mathbf{J}_{\cdot 1}-\mathbf{J}_{\cdot\cdot}$; $\mathbf{U}_{\cdot 2}=\mathbf{J}_{\cdot 2}-\mathbf{J}_{\cdot\cdot}$; $\mathbf{U}_{11}=\mathbf{J}_{11}-\mathbf{U}_{1\cdot}-\mathbf{U}_{\cdot 1}-\mathbf{J}_{\cdot\cdot}$; $\mathbf{U}_{12}=\mathbf{J}_{12}-\mathbf{U}_{1\cdot}-\mathbf{U}_{\cdot 2}-\mathbf{J}_{\cdot\cdot}$; $\mathbf{U}_{21}=\mathbf{J}_{21}-\mathbf{U}_{2\cdot}-\mathbf{U}_{\cdot 1}-\mathbf{J}_{\cdot\cdot}$; $\mathbf{U}_{22}=\mathbf{J}_{22}-\mathbf{U}_{2\cdot}-\mathbf{U}_{\cdot 2}-\mathbf{J}_{\cdot\cdot}$. This yields $\hat{\alpha}_1=\mathbf{y'U}_{1\cdot}=-3.25$, $\hat{\alpha}_2=\mathbf{y'U}_{2\cdot}=3.25$, $\hat{\beta}_1=\mathbf{y'U}_{\cdot 1}=1.25$, $\hat{\beta}_2=\mathbf{y'U}_{\cdot 2}=-1.25$, $\hat{\tau}_{11}=\mathbf{y'U}_{11}=0.25$, etc. Source identifier variables are $\mathbf{J}^*_{\cdot\cdot}=\|\mathbf{J}_{\cdot\cdot}\|^{-2}\mathbf{J}_{\cdot\cdot}$, $\mathbf{U}^*_{1\cdot}=\|\mathbf{U}_{1\cdot}\|^{-2}\mathbf{U}_{1\cdot}$, etc, where $\|\cdot\|^2$ designates a sum of squares of components. The resulting estimated co-effects are coefficients in the equation

$$\begin{aligned} \mathbf{y}_{\cdot} \approx\ & -0.586\mathbf{U}^*_{11} + 0.586\mathbf{U}^*_{12} - 0.703\mathbf{U}^*_{1\cdot} \\ & +0.586\mathbf{U}^*_{21} - 0.586\mathbf{U}^*_{22} + 0.703\mathbf{U}^*_{2\cdot} \\ & +2.109\mathbf{U}^*_{\cdot 1} - 2.109\mathbf{U}^*_{\cdot 2} + 15.781\mathbf{J}^*_{\cdot\cdot} \quad . \end{aligned}$$

SUMMARY

Source-identifier variables in a General Linear Model are defined using the Gautschi-Scheffé proof for the Gauss-Markov Theorem, and co-effects are defined as coefficients of corresponding source-identifier variables when the expected observation vector is expressed as a unique linear combination of source-identifier variables subject to certain restrictions. The resulting co-effect analysis of variance can supplement Yates' weighted-squares-of-means analysis of variance for fixed effect models; since, with balanced data, it gives results identical to those of standard methods, whereas with unbalanced, data this technique can be significantly more sensitive — because the cell-size-dependent extended parameters used, the co-effects, differ from the usual effects.

ACKNOWLEDGEMENT

I would like to thank Michael Plante for improving the presentation of this paper.

REFERENCES

Bailey, D.W. (1953). *The Inheritance of Maternal Influences on the Growth of Rat*, Ph. D. Thesis, University of California.

Federer, W.T. and Zelen, M. (1966). Analysis of Multifactor Classifications with Unequal Numbers of Observations, *Biometrics*, 22, 525-552.

Francis, I, (1973). Comparison of Several Analysis of Variance Programs, *J. Amer. Statist. Assoc.*, 68, 860-865.

Graybill, F.A. (1976). *Theory and Application of the Linear Model*, Belmont, California: Wadsworth.

Kutner, M.H. (1974). Hypothesis Testing in Linear Models (Eisenhart Model 1), *The Americam Statistician*, 28, 98-100.

Lund, R.E. (1975). Tables for an Approximate Test for Outliers in Linear Models, *Technometrics*, 17, 473-476.

Mann, H.B. (1949). *Analysis and Design of Experiments*, New York: Dover.

Mead, R. (1990). The Non-orthogonal Design of Experiment, *J. R. Statist. Soc.* A, 153, 151-201.

Milliken, G.A. and Johnson, D.E. (1984). *Analysis of Messy Data Vol.I.* New York: Van Nostrand Reinhold.

Nelder, J.A. (1974). Letter to the Editor, *J. R. Statist. Soc.* C, 23, 232.

Plante, A. (1991). An inclusion-consistent solution to the problem of absurd confidence statements: 1. Consistent exact confidence interval estimation, *Canad. J. Statist.*, 19, 389-387.

Plante, A. (1992). Utilisation conjointe de l'opérateur SWEEP et d'une méthode de calcul recommandée par Scheffé pour l'estimation sous contraintes en analyse de la variance, *Actes du colloque sur les méthodes et domaines d'application de la statistique 1991*, 91-100, Bureau de la statistique du Québec, Québec, Canada.

SAS Institute Inc. (1988). *SAS/STAT User's Guide, Release 6.03 Edition*, SAS Institute Inc., Cary, North Carolina.

Scheffé, H. (1959). *The Analysis of Variance*. New York: John Wiley & Sons.

Speed, F.M., Hocking, R.R., and Hackney, O.P. (1978). Methods of Analysis of Linear Models with Unbalanced Data, *J. Amer. Statist. Assoc.*, 73, 105-112.

Yates, F. (1934). The Analysis of Multiple Classifications with Unequal Numbers in the Different Classes, *J. Amer. Statist. Assoc.* 29, 52-66.

TABLES

Table 1. Cell Sizes in Bailey's Foster-Nursing Experiment

Litter Genotype	Foster-Mother Genotype A	F	I	J
A	5	3	4	5
F	4	5	4	2
I	3	3	5	3
J	4	3	3	5

Table 2. Yates' Weighted-Squares-of-Means Analysis of Variance for Bailey's Foster-Nursing Data (based on SAS GLM Procedure)

Source of variation	Degrees of Freedom	Sum of Squares	Mean Square	*F*-Ratio	*P*-Value
Litter Genotype	3	27.65592	9.21864	0.17	0.9161
Foster-Mother Genotype	3	671.73765	223.91255	4.13	0.0114
Genotype Interaction	9	824.07251	91.56361	1.69	0.1201
Error	45	2440.81650	54.24040	———	———

Table 3. Co-effect Analysis of Variance for Bailey's Foster-Nursing Data (based on an *ad hoc* GAUSS program)

Source of Variation	Degrees of Freedom	Sum of Squares	Mean Square	*F*-Ratio	*P*-Value
Litter Genotype	3	1295.1696	431.7232	7.96	0.00023
Foster-Mother Genotype	3	884.1750	294.725	5.43	0.00283
Genotype Interaction	9	12258.9025	1362.100	25.11	$1.008X10^{-14}$
Error	45	2440.8165	54.240	———	———

Table 4. Estimation of Co-effects in Bailey's Foster-Nursing Experiment, With Approximate Conservative 95% Resampling Confidence Intervals

Interaction Co-effect					Litter Genotype Main Co-Effect
Litter Genotype	Foster-Mother Genotype				
	A	F	I	J	
A	9.5±3.3	-12.3±3.6	-3.3±3.3	6.1±3.7	5.7±2.5
F	-0.4±3.3	15.3±3.5	1.5±3.5	-16.3±3.7	0.0±2.6
I	-9.8±6.0	1.8±3.7	9.0±4.0	-0.9±3.5	-4.4±3.5
J	0.8±3.2	-4.8±3.3	-7.1±3.3	11.1±3.2	-1.3±2.5
Foster-Mother Main Co-Effect	3.3±2.9	0.1±2.5	1.7±2.6	-5.1±2.7	General Mean Co-Effect 57.9±2.1

Table 5. Yates' Weighted-Squares-of-Means Analysis for Bailey's Data Modified by Pulling Back an Almost-Outlier.

Source of variation	Degrees of Freedom	Sum of Squares	Mean Square	*F*-Ratio	*P*-Value
Litter Genotype	3	82.1520	27.3840	0.66	0.5819
Foster-Mother Genotype	3	669.4835	223.1612	5.37	0.0030
Genotype Interaction	9	1193.7732	132.6415	3.19	0.0046
Error	45	1871.4832	41.5885	———	———

Table 6. Co-Effect Analysis of Variance With Modified Bailey's Data Obtained By Pulling Back an Almost-Outlier

Source of variation	Degrees of Freedom	Sum of Squares	Mean Square	*F*-Ratio	*P*-Value
Litter Genotype	3	1537.3545	512.45	12.32	$2.76X10^{-6}$
Foster-Mother Genotype	3	741.0788	247.03	5.94	0.0017
Genotype Interaction	9	13082.659	1453.63	34.95	$6.86X10^{-19}$
Error	45	1871.4832	41.59	———	———

Table 7. Modified Bailey's Data Obtained by Pulling Back an Almost-Outlier and by Inflating Intra-Cell Variances

Litter Genotype	Foster-Mother Genotype A		F		I		J	
A	60.0	65.9	56.8	65.7	51.4	37.1	47.1	47.7
	71.4	56.9	34.7		67.2	51.7	57.5	33.0
	64.2						69.4	
F	65.9	47.2	43.9	66.4	58.3	42.5	55.1	36.7
	51.3	45.0	67.5	63.0	62.5	52.3		
			62.4					
I	27.0	58.0	50.7	68.8	31.4	57.9	50.4	58.0
	36.3		73.6		42.1	58.6	39.9	
					68.1			
J	62.3	53.8	61.9	55.9	38.7	66.2	41.8	37.1
	59.5	41.9	50.5		58.7		53.2	57.5
							55.8	

Table 8. Analysis of Variance for the Data in Table 7 Using Yates' Weighted-Squares-of-Means Method.

Source of variation	Degrees of Freedom	Sum of Squares	Mean Square	*F*-Ratio	*P*-Value
Litter Genotype	3	82.0243	27.3414	0.22	0.8839
Foster-Mother Genotype	3	669.3106	223.1035	1.77	0.1659
Genotype Interaction	9	1193.6570	132.6286	1.05	0.4142
Error	45	5662.5622	125.8347	———	———

Table 9. Co-Effect Analysis of Variance for the Data in Table 7.

Source of variation	Degrees of Freedom	Sum of Squares	Mean Square	*F*-Ratio	*P*-Value
Litter Genotype	3	1535.5571	511.85	4.07	0.0122
Foster-Mother Genotype	3	742.7836	247.59	1.97	0.1320
Genotype Interaction	9	13086.1458	1454.02	11.56	$3.90X10^{-9}$
Error	45	5662.5622	125.83	———	———