On Means Estimated from Fixed and Mixed Linear Models

Shayle R. Searle
Friedrich Pukelsheim

Preprint Nr. 58

Februar 1985

BU-868-M

On Means Estimated from Fixed and Mixed Linear Models

Shayle R. Searle * and Friedrich Pukelsheim

Institut für Mathematik
Universität Augsburg, Federal Republic of Germany

ABSTRACT

Means estimated from fixed and mixed models of the 1-way classification are compared in terms both of sampling variances and of weights given to the class means. Extensions to other models are indicated.

*On leave from Cornell University, Ithaca, N.Y., supported by a U. S. Senior Scientist Award from the Alexander von Humboldt-Stiftung.

1. Introduction

Subclasses of data that contain differing numbers of observations can have their means combined linearly with a variety of different weights: weighting by the number of observations leads to the overall mean, weighting equally yields the mean of the subclass means, and a third possibility is weighting inversely according to variances when those variances are unequal. Several aspects of these means are considered.

2. Fixed Effects Models

Suppose y_{ij} is the j'th observation of the i'th class of a 1-way classification, with $i = 1, \ldots, a$ and $j = 1, \ldots, n_i$; i.e., a classes and n_i observations in the i'th class. Then the model equation for y_{ij} can be taken as

$$y_{ij} = \mu_i + e_{ij} \tag{1}$$

where μ_i is the population mean of the i'th class, and e_{ij} is a random error term. In the fixed effects model the e_{ij} are assumed to be independent random variables identically distributed with zero mean and variance σ_e^2; and the covariance between any pair of (different) e_{ij} terms is assumed to be zero. Under these conditions the BLUE (best linear unbiased estimator) of μ_i and the sampling variance of that estimator are, respectively,

$$\hat{\mu}_i = \bar{y}_i = \sum_{j=1}^{n_i} y_{ij}/n_i \text{ and } v(\bar{y}_i) = \sigma^2/n_i, \tag{2}$$

similar to Searle (1971, pages 325 and 339).

We consider three different weighted means of the μ_is. First is μ_n, in which the number of observations are used as weights, $\mu_n = \sum n_i\mu_i / \sum n_i$. (All summations are with respect to i, over the range $i = 1, 2, \ldots, a$.) Second is μ_e, based on equal weights, $\mu_e = \sum \mu_i/a$; and third is a general weighted average, $\mu_w = \sum w_i\mu_i / \sum w_i$ using arbitrary weights w_i. The BLUEs of these and their sampling variances are as follows:

$$\hat{\mu}_n = \sum n_i\bar{y}_i \Big/ \sum n_i = \bar{y}_{..}, \text{ with } v(\hat{\mu}_n) = \sigma_e^2 \Big/ \sum n_i, \tag{3}$$

$$\hat{\mu}_e = \sum \bar{y}_i/a, \text{ with } v(\hat{\mu}_e) = \sigma_e^2 (\sum 1/n_i)/a^2, \tag{4}$$

and

$$\hat{\mu}_w = \sum w_i\bar{y}_i \Big/ \sum w_i, \text{ with } v(\hat{\mu}_w) = \sigma_e^2 (\sum w_i^2/n_i) \Big/ (\sum w_i)^2. \tag{5}$$

Some elementary properties can be noted. $\hat{\mu}_n$ is the grand mean $\bar{y}_{..}$, whereas $\hat{\mu}_e$ is the mean of class means, $\sum \bar{y}_i/a$. They are equal when all n_i are the same, as are μ_n and μ_e. Also, μ_w for $w_i = n_i$ is μ_n, and μ_w for $w_i = 1$ is μ_e. Of the three estimators, $\hat{\mu}_n$ has the smallest variance as evidenced by an application of the Cauchy- Schwarz inequality:

$$\sum n_i \sum w_i^2/n_i \geq (\sum \sqrt{n_i}\sqrt{w_i^2/n_i})^2 = (\sum w_i)^2$$

so that

$$1/\sum n_i \leq (\sum w_i^2/n_i)\Big)/(\sum w_i^2), \text{ i.e. }, v(\hat{\mu}_n) \leq v(\hat{\mu}_w). \tag{6}$$

Thus no weighted mean of the μ_is has a BLUE with smaller variance than that of μ_n. This is an attractive property for μ_n even though defining an overall mean as μ_e seems more natural than does μ_n because of the dependence of μ_n on the numbers of observations in the classes.

3. Mixed Models

What is usually known as the random effects model for the 1-way classification has model equation $y_{ij} = \mu + \alpha_i + e_{ij}$ for e_{ij} as in the fixed effects model and for the α_is being uncorrelated random effects with zero means and variance σ_α^2; and with the covariance between every α_i and every e_{hk} being zero. Since μ is a fixed effect this model is strictly a mixed model and we think of it in this manner because of being interested in estimating μ in the presence of the random effects. Its BLUE, to be denoted $\hat{\mu}_r$ is, similar to Searle (1971, page 463),

$$\hat{\mu}_r = \sum \frac{n_i}{n_i\sigma_\alpha^2 + \sigma_e^2}\bar{y}_i \Big/ \sum \frac{n_i}{n_i\sigma_\alpha^2 + \sigma_e^2} \text{ with } v(\hat{\mu}_r) = \sigma^2 \Big/ \sum \frac{n_i}{n_i\sigma_\alpha^2 + \sigma_e^2}. \tag{7}$$

A comparison of variances is of interest. That of $\hat{\mu}_w$ in the mixed model, to be denoted $v_M(\hat{\mu}_w)$, is simply $v(\hat{\mu_w})$ of (5) with σ_e^2/n_i replaced by $\sigma_\alpha^2 + \sigma_e^2/n_i$; and by the same reasoning as used in deriving (6) one can show that $v(\hat{\mu}_r) < v_M(\hat{\mu}_w)$, of which $v(\hat{\mu}_r) < v_M(\hat{\mu}_n)$ is then but a special case. Nevertheless, $v(\hat{\mu}_n) = \sigma_e^2/\sum n_i$ of (3) is less than $v(\hat{\mu}_r)$ of (7), as may be seen by observing that

$$1/v(\hat{\mu}_n) - 1/v(\hat{\mu}_r) = \sum n_i[1/\sigma_e^2 - 1/(n_i\sigma_\alpha^2 + \sigma_e^2)] > 0$$

and so

$$v(\hat{\mu}_n) < v(\hat{\mu}_r) < v_M(\hat{\mu}_n). \tag{8}$$

Thus in the mixed model no linear combination of the $\bar{y}_i$s has a smaller variance than does μ_r (as is to be expected because $\hat{\mu}_r$ is the BLUE of μ), but $v(\hat{\mu}_n)$ in the fixed effects model has smaller variance than $v(\hat{\mu}_r)$ in the mixed model. The inequality chain in (8) can also be extended to

$$v(\hat{\mu}_n) < v(\hat{\mu}_e) < v(\hat{\mu_r}) < v_M(\hat{\mu}_n) < v_M(\hat{\mu}_e).$$

4. Relationships Among the Means

The intra-class correlation in the mixed model is $\rho = \sigma_\alpha^2/(\sigma_\alpha^2 + \sigma_e^2)$. To emphasize dependence on ρ we now write $\hat{\mu}_r$ of (7) as

$$\hat{\mu}_{r,\rho} = \sum \frac{n_i}{n_i\rho + 1 - \rho}\bar{y}_i \Big/ \sum \frac{n_i}{n_i\rho + 1 - \rho}. \tag{9}$$

Immediately we see, by comparison with (3) and (4) that

$$\hat{\mu}_{r,0} = \hat{\mu}_n \quad \text{and } \hat{\mu}_{r,1} = \hat{\mu}_e.$$

This is not surprising. $\rho = 0$ is equivalent to $\sigma_\alpha^2 = 0$ as is true of the fixed effects model and so $\hat{\mu}_{r,0} = \hat{\mu}_n$, its BLUE counterpart in that model. And $\rho = 1$, although equivalent to $\sigma_e^2 = 0$, is more interestingly the case of observations within each class being perfectly correlated — in effect, identical. Hence, no matter what the value of n_i is, $\bar{y}_i$ has variance σ_α^2 and so the linear combination of $\bar{y}_i$s that has minimum variance is $\hat{\mu}_e = \sum \bar{y}_i / a$.

Despite these consequences of putting $\rho = 0$ and $\rho = 1$ in $\hat{\mu}_r$, it is nevertheless surprising how quickly the weights given to each $\bar{y}_i$ change from being proportional to n_i in $\hat{\mu}_{r,0} = \mu_n$ to approaching being equal in $\hat{\mu}_{r,1} = \mu_e$ as ρ increases from 0 to 1. Consider two classes, one described as having a large number of observations, n_L, and the other having a small number, n_S, with, of course, $n_L > n_S$. In $\hat{\mu}_r$ the ratio of the weight given $\bar{y}_S$ to that given to y_L is τ_ρ say, where, from (9)

$$\tau_\rho = \frac{\text{coefficient of } \bar{y}_S \text{ in } \hat{\mu}_{r,\rho}}{\text{coefficient of } \bar{y}_L \text{ in } \hat{\mu}_{r,\rho}} = \frac{n_S(n_L\rho + 1 - \rho)}{n_L(n_S\rho + 1 - \rho)}. \tag{10}$$

Now $\tau_0 = n_S/n_L$, corresponding to $\hat{\mu}_{r,0} = \hat{\mu}_n$, and as ρ increases from zero to unity τ_ρ increases from $\tau_0 = n_S/n_L$ to $\tau_1 = 1$. Thus as $\rho \to 1$ we see that y_S, the mean of the smaller sized class, gets increasingly larger weights in $\hat{\mu}_{r,\rho}$, relative to $\bar{y}_L$. What is interesting about this is that this increase can, depending on the magnitudes of n_L and n_S be quite appreciable, even for very small values of ρ. The accompanying table shows values of τ_ρ for three pairs of n_L, n_S values and a range of values of ρ.

(show table)

5. Extensions

Consider a 2-way nested classification of a main classes the i'th of which has b_i sub-classes, in the j'th of which there are n_{ij} observations y_{ijk} for $k = 1, \ldots, n_{ij}$, with $i = 1, \ldots, a$ and $j = 1, \ldots, b_i$. A mixed model for this situation can be taken as $y_{ijk} = \mu_i + \beta_{ij} + e_{ijk}$ where μ_i is a fixed effect and β_{ij} and e_{ijk} are random effects with zero means, variances σ_β^2 and σ_e^2, respectively, and with all covariances zero. Then, similar to $\hat{\mu}_r$ of (7), the BLUE of μ_i is

$$\hat{\mu}_i = \sum_{j=1}^{b} \frac{n_{ij}}{n_{ij}\sigma_\beta^2 + \sigma_e^2} \bar{y}_{ij.} \Big/ \sum_{j=1}^{b} \frac{n_{ij}}{n_{ij}\sigma_\beta^2 + \sigma_e^2}. \tag{11}$$

Discussions of this and of linear combinations of the $\hat{\mu}_i$s, can be made similar to those of Sections 2 and 3. Analogous extensions could also be made for a 2-way crossed classification for combining BLUEs $\hat{\mu}_{ij} = \bar{y}_{ij.}$ in situations where $v(\bar{y}_{ij.}) = \sigma_\gamma^2 + \sigma_e^2/n_{ij}$.

Reference

Searle, S.R. (1971): Linear Models. *Wiley, New York.*

EXAMPLES OF THE RELATIVE WEIGHTS GIVEN TO TWO SAMPLE MEANS IN THE ESTIMATOR

$$\hat{\mu}_{r;\rho} = \sum \frac{n_i}{n_i\sigma_\alpha^2+\sigma_e^2}\,\bar{y}_i \Big/ \sum \frac{n_i}{n_i\sigma_\alpha^2+\sigma_e^2}$$

$$\tau_\rho = \frac{\text{coefficient of } \bar{y}_S \text{ in } \hat{\mu}_{r,\rho}}{\text{coefficient of } \bar{y}_L \text{ in } \hat{\mu}_{r,\rho}} = \frac{n_S(n_L\rho+1-\rho)}{n_L(n_S\rho+1-\rho)}$$

$\rho = \frac{\sigma_\alpha^2}{\sigma_\alpha^2+\sigma_e^2}$	Three sets of n_L, n_S values: n_L, n_S = 20, 4	n_L, n_S = 100, 20	n_L, n_S = 100, 5
0 ($\hat{\mu}_{r,0}=\hat{\mu}_n$)	.20	.20	.05
.05	.33	.61	.28
.1	.45	.75	.38
.3	.71	.92	.70
.5	.840	.962	.842
.7	.923	.983	.925
.9	.978	.996	.979
1.0 ($\hat{\mu}_{r,1}=\hat{\mu}_e$)	1.00	1.00	1.00

Bestellungen nimmt entgegen:

Institut für Mathematik
der Universität Augsburg
Memminger Str. 6
8900 Augsburg

Verzeichnis der erschienenen Preprints

1. Hoffmann, K.-H.; Sprekels, J.: Real-time Control of the Free Boundary in a Two-phase Stefan Problem (1982), Numer. Funct. Anal. and Optimiz. 5, 47-76 (1982)

2. Schert, R.: Eine Bemerkung zu einer Arbeit von C. Meyer über Imaginär-quadratische Körper mit der Klassenzahl 2 (1982)

3. Brokate, M.: A Hyperbolic Free Boundary Problem: Existence, Uniqueness and Discretization (1982)

4. Grötschel, M.; Padberg, M.: Polyhedral Aspects of the Traveling Salesman Problem I: Theory (1983)

5. Grötschel, M.; Padberg, M.: Polyhedral Aspects of the Traveling Salesman Problem II: Computation (1983)

6. Ritter, J.; Sehgal, S. K.: On a Conjecture of Zassenhaus on Torsion Units in Integral Group Rings (1983), Math. Ann. 264, 257-270 (1983)

7. Barahona, F.; Grötschel, M.; Mahjoub, A. R.: Facets of the Bipartite Subgraph Polytope (1983)

8. Grötschel, M.; Jünger, M.; Reinelt, G.: A Cutting Plane Algorithm for the Linear Ordering Problem (1983)

9. Grötschel, M.; Jünger, M.; Reinelt, G.: Optimal Triangulation of Large Real World Input-output Matrices (1983), Statistische Hefte, 25 (1984), 261-295

10. Jünger, M.; Pulleyblank, W. R.; Reinelt, G.: On Partitioning the Edges of Graphs into Connected Subgraphs (1983)

11. Grötschel, M.: On Intersections of Longest Cycles (1983), "Graph Theory and Combinatorics", B. Bollobás, London, Academic Press, 1984, 171-189

12. Hoffmann, K.-H.; Sprekels, J.: Real-time Control in a Free Boundary Problem Connected with the Continuous (1983) Casting of Steel, Intern. Ser. Numer. Math. 68, Birkhäuser-Verlag (1983)

13. Brokate, M.: Numerical Solution of a Hyperbolic Free Boundary Problem with a Method of Characteristics (1983)

14. Hoffmann, K.-H.; Sprekels, J.: Exact Bounds for the Radially Symmetric Shape of Confined Plasma in the Unit Circle (1983), Math. Meth. in Appl. Sci. 6 (1984) 496-511

15. Grötschel, M.: Polyhedral Combinatorics: An Annotated Bibliography (1983)

16. Euler, R.; Jünger, M.; Reinelt, G.: Generalizations of Cliques, Odd Cycles and Anticycles and their Relation to Independence System Polyhedra (1983)

17. Hoffmann, K.-H.; Sprekels, J.: On the Identification of Coefficients of Elliptic Problems by Asymptotic Regularization (1983)

18. Hoffmann, K.-H.; Sprekels, J.: A Note on the Thermostat Control of Heating Processes in Materials with Thermal Memory (1983), Mat. Aplic. Comp., V. 3, n°. 3, pp. 257 a 264, 1984

19. Alt, H. W.; Hoffmann, K.-H.; Sprekels, J.: A Numerical Procedure to Solve Certain Identification Problems (1983), Intern. Ser. Numer. Math. 68, Birkhäuser-Verlag

20. Brüning, J.: Heat Equation Asymptotics for Singular Sturm-Liouville Operators (1983), Math. Ann. 268, 173-196 (1984)

21. Knabner, P.: A Free Boundary Problem Arising from the Leaching of Saline Soils (1983)

22. Knabner, P.: Regularization of the Cauchy Problem for the Heat Equation by Norm Bounds (1984)

23. Brüning, J.; Heintze, E.: Spektrale Starrheit gewisser Drehflächen (1984), Math. Ann. 269, 95-101 (1984)

24. Hoffmann, K.-H.; Sprekels, J.: On the Identification of Parameters in General Variational Inequalities by Asymptotic Regularization (1984)

25. Hoffmann, K.-H.; Kornstaedt, H.-H.; Sprekels, J.: Automatische Steuerung freier Ränder bei Stefan Problemen durch Thermostatkontrollen am festen Rand (1984)

26. Grötschel, M.: Developments in Combinatorial Optimization (1984), "Persepctives in Mathematics" (Anniversary of Oberwolfach 1984), W. Jäger, J. Moser & R. Remmert, Basel, Birkhäuser-Verlag, 1984, 249-294

27. Knabner, P.: Control of Stefan Problems by Means of Linear-Quadratic Defect Minimization (1984)

28. Grötschel, M.; Jünger, M.; Reinelt, G.: Acyclic Subdigraphs and Linear Orderings: Polytopes, Facets and a Cutting Plane Algorithm (1984)

29. Müller-Funk, U.; Pukelsheim, F.; Witting, H.: On the Duality between Locally Optimal Tests and Optimal Experimental Designs (1984)

30. Jürgens, U.: The Estimation of a Nonstationary Random Coefficient AR(1) Process (1984)

31. Brüning, J.: Spectral Analysis of Singular Sturm-Liouville Problems with Operator Coefficients (1984)

32. Schröder, H.: Some Continuous Index Theorems (1984)

33. Zehendner, E.: A Non-Existence Theorem for Cyclic MDS-Codes (1984)

34. Christof, K.; Pukelsheim, F.: Approximate Design Theory for a Simple Block Design with Random Block Effects (1984)

35. Grötschel, M.; Lovász, L.; Schrijer, A.: Relaxations of Vertex Packing (1984)

36. Barahona, F.; Grötschel, M.: On the Cycle Polytope of a Binary Matroid (1984)

37. Grötschel, M.; Holland, O.: Solving Matching Problems with Linear Programming (1984)

38. Grötschel, M.: Polyedrische Kombinatorik und Schnittebenenverfahren (1984), Bericht Nr. 231 (1984) der Mathematisch-Statistischen Sektion, Forschungszentrum Graz, Österreich

39. Gurgel, M. A.; Wakabayashi, Y.: On k-leaf connected graphs (1984)

40. Hoffmann, K.-H.; Sprekels, J.: The Method of Asymptotic Regularization and Restricted Parameter Identification Problems in Variational Inequalities (1984)

41. Kielhöfer, H.: Multiple Eigenvalue Bifurcation for Fredhom Operators (1984)

42. Baksalary, J. K.; Pukelsheim, F.: A Note on the Matrix Ordering of Special C-Matrices (1984)

43. Pawlow, I.; Niezgódka, M.: Discrete Approximation of Multi-Phase Stefan Problems with Possible Degeneration (1984)

44. Pawlow, I.: Boundary Control of Degenerate Two-Phase Stefan Problems (1984)

45. Kielhöfer, H.: Interaction of Periodic and Stationary Bifurcation from Multiple Eigenvalues (1984)

46. Grötschel, M.: Fortschritte in der polyedrischen Kombinatorik (1984)

47. Schröder, H.: On the Homotopy type of the regular group of a real W*-algebra (1984)

48. Pukelsheim, F.: A Note on Predictable Criteria for Absolute Continuity and Singularity of Two Probability Measures (1984)

49. Knabner, P.: Global Existence in a General Stefan-like Problem (1984)

50. Hoffmann, K.-H.; Sprekels, J.: On the Identification of Heat Conductivity and Latent Heat in a One-Phase Stefan Problem (1985)

51. Appell, J.: Deux Méthodes Topologiques pour la Résolution des Équations Elliptiques Non linéaires sans Compacité

52. Appell, J.; de Pascale Espedito: Théorèmes de Bornage pour l'Opérateur de Nemyckii dans les Espaces Idéaux

53. Appell, J.: Misure di Non Compattezza in Spazi Ideali