SORENSON, H. W. (1970). Least-squares estimation from Gauss to Kalman. *IEEE Spectrum* **7** 63–68.
STEIN, C. (1945). A two-sample test for a linear hypothesis whose power is independent of the variance. *Ann. Math. Statist.* **16** 243–258.
STEIN, C. (1956). Inadmissibility of the usual estimator for the mean of a multivariate normal distribution. *Proc. Third Berkeley Symp. Math. Statist. Probab.* **1** 197–206. Univ. California Press, Berkeley.
THOMPSON, R. (1973). The estimation of variance and covariance components with an application when records are subject to culling. *Biometrics* **39** 527–550.
TURNER, H. N. and YOUNG, S. S. Y. (1973). *Quantitative Genetics in Sheep Breeding*. Macmillan, New York.
YATES, F. (1949). Systematic sampling. *Philos. Trans. Roy. Soc. London Ser. A* **241** 345–377.

Comment

Katherine Campbell

It has been a pleasure to read about the long history of Best Linear Unbiased Prediction, and especially about its uses in traditional statistical areas of application such as agriculture. My own experience with BLUP is in the context of ill-posed inverse problems, and I would like to discuss this paper from this point of view, where the random effects are generated by hypothesized superpopulations, in contrast with the identifiable populations considered by Robinson.

MODEL-BASED ESTIMATION FOR ILL-POSED INVERSE PROBLEMS

The author mentions two examples of superpopulation approaches to estimation: image restoration and geostatistics. The same ideas are also used in model-based estimation for finite populations, function approximation and many other inference problems. These problems concern inference about a reality that is in principle completely determined, but whose observation is limited by the number and/or resolution of the feasible measurements, as well as by noise. In geophysics, x-ray imaging and many other areas of science and engineering these are known as inverse problems (O'Sullivan, 1986; Tarantola, 1987).

The unknown reality we may consider to be a function $\mathbf{m}$ defined on some domain $\mathbf{T}$. The data typically consist of noisy observations on a finite number n of functionals of $\mathbf{m}$. We can write the data vector $\mathbf{y}$ in terms of a transformation L mapping $\mathbf{m}$ into an n-dimensional vector:

$$\mathbf{y} = L\mathbf{m} + \mathbf{e}.$$

Katherine Campbell is a staff member of the Statistics Group, Los Alamos National Laboratory, Los Alamos, New Mexico 87545.

In the sequel, we will assume that L is a linear transformation, i.e., that the observed functionals are linear. In particular, if the cardinality $|\mathbf{T}|$ of $\mathbf{T}$ is finite, L can be represented by an $n \times |\mathbf{T}|$ matrix.

BLUP arises when we embed this problem in a superpopulation model, under which $\mathbf{m}$ is one realization (albeit the only one of interest) of a stochastic process $\mathbf{M}$ indexed by $\mathbf{T}$. This superpopulation model has two components, corresponding to the "fixed" and "random" effects in Robinson's discussion. The fixed effects define the mean of the superpopulation, which is here assumed to lie in a finite-dimensional subspace of functions on $\mathbf{T}$. We denote this subspace by $\mathbf{R}(F)$, the range of the linear operator F that maps a p-vector $\mathbf{b}$ into the function

$$F\mathbf{b} = \sum b_i \mathbf{f}_i.$$

where $\{\mathbf{f}_1, \ldots, \mathbf{f}_p\}$ is a basis for the subspace.

Any realization of $\mathbf{M}$ can then be written as a sum $F\beta + \mathbf{u}$, where β is an unknown vector of p fixed effects and $\mathbf{u}$ is a realization of a stochastic "random effects" process with mean zero and covariance P. As we are interested in the realized $\mathbf{m}$, we need to estimate both the fixed and random effects. Among estimates that are linear functions of the data vector

$$(1) \qquad \mathbf{y} = LF\beta + L\mathbf{u} + \mathbf{e},$$

the BLUP $\hat{\mathbf{m}} = F\hat{\beta} + \hat{\mathbf{u}}$ is the optimal choice: under the assumed superpopulation model $\hat{\mathbf{m}}$ is unbiased in the sense of Section 7.2 (i.e., $E\hat{\mathbf{m}} = E\mathbf{m}$) and it minimizes the variance of any linear functional of $\hat{\mathbf{m}} - \mathbf{m}$. (To make the correspondence between equation 1 and Robinson's equation (1.1) explicit, $X \approx LF$, $Z \approx L$, $G \approx LPL^T$, and $\mathbf{e}$ is a realization of a random n-vector with mean zero and covariance R. $\hat{\beta}$ and $\hat{\mathbf{u}}$ are then provided by the BLUP formulas.)