Estimation and variance in block designs

1 Background

Assume that we have a finite set Ω of plots, partitioned into blocks. Write $B(\omega)$ for the block containing plot ω. There is also a set $\mathcal{T}$ of treatments. The design consists of an allocation of one treatment to each plot. Write $T(\omega)$ for the treatment allocated to plot ω.

Let n be the number of plots, b the number of blocks and v the number of treatments. The incidence of plots in blocks can be recorded by an $n \times b$ matrix X_B whose (ω, m) entry is equal to 1 if $B(\omega) = m$ and equal to zero otherwise. The allocation of plots to treatments is shown in an analogous $n \times v$ matrix X_T. The $v \times v$ matrix $X_T' X_B X_B' X_T$ is called the *concurrence* matrix Λ of the design: its (i, j)-entry λ_{ij} is equal to the number of ordered pairs of plots (α, ω) for which (i) α and ω are in the same block (ii) $T(\alpha) = i$ and (iii) $T(\omega) = j$. If the design is binary then λ_{ii} is equal to the *replication* of treatment i: that is, the number of plots which are allocated treatment i. Also, if the design is binary then, for $i \neq j$, the concurrence λ_{ij} is equal to the number of blocks in which treatments i and j both appear.

2 Statistical model

The experimenter measures a response Y_ω on each plot ω, hence a vector Y in $\mathbb{R}^\Omega$. We assume that the Y_ω are real random variables satisfying the following two conditions.

(1) There are (unknown) constants τ_i, for i in $\mathcal{T}$, and β_m, for $m = 1, \ldots, b$, such that if $T(\omega) = i$ and $B(\omega) = m$ then

$$\mathbb{E}(Y_\omega) = \tau_i + \beta_m.$$

(2) The Y_ω are independent, each having variance σ^2 (which is also probably unknown).

We want to estimate the treatment parameters τ_i: the block parameters are merely a nuisance.

Note that we could add a constant to each τ_i, and subtract it from each β_m, without changing the model for expectation, so we cannot in fact estimate the treatment parameters. The best we can hope to do is to estimate all the differences $\tau_i - \tau_j$ and, more generally, linear combinations of the form $\sum_{i \in \mathcal{T}} x_i \tau_i$ for known real numbers x_i satisfying $\sum_{i\in\mathcal{T}} x_i = 0$. Such a linear combination is called a *contrast*. The word 'contrast' is also used for a vector x in $\mathbb{R}^{\mathcal{T}}$ for which $\sum_{i\in\mathcal{T}} x_i = 0$.

3 Connectedness

Let i and j be distinct treatments. Suppose that there are plots α and ω in block m such that $T(\alpha) = i$ and $T(\beta) = j$. Then

$$\mathbb{E}(Y_\alpha - Y_\omega) = \tau_i - \tau_j,$$

so we can estimate $\tau_i - \tau_j$. If l is a further treatment and there is a block m' containing both j and l then we can estimate $\tau_j - \tau_l$. Hence we can estimate $\tau_i - \tau_l$.

The *Levi graph* of the block design has the treatments and the blocks as vertices. Each plot gives an edge of the graph: plot ω gives an edge from $T(\omega)$ to $B(\omega)$. The above argument shows that we can estimate $\tau_i - \tau_j$ whenever there is a path from i to j in the Levi graph. Thus all contrasts are estimable if the Levi graph is connected. In this case the design is also called *connected*.

4 Estimation

The model for expectation can be written in matrix form as

$$\mathbb{E}(Y) = X_T\tau + X_B\beta.$$

Let Q be the $n \times n$ matrix which gives orthogonal projection onto the orthogonal complement of the column space of X_B. Thus $QX_B = O$, $Q' = Q$ and $Q^2 = Q$. It can be shown that

$$Q = I - X_B K^{-1} X_B',$$

where K is the diagonal matrix whose entries are the block sizes.

The *information matrix* L of the design is defined by

$$L = X_T' Q X_T.$$

Note that the row sums of L are all zero, so that the all-1 vector in $\mathbb{R}^{\mathcal{T}}$ is in the kernel of L. Moreover Lx is a contrast if x is. The design is connected if and only if the column space of L contains all contrasts.

The information matrix can often be simplified. If all blocks have size k then

$$Q = I - \frac{1}{k} X_B X_B'$$

and so

$$L = X_T' X_T - \frac{1}{k}\Lambda.$$

The matrix $X_T' X_T$ is diagonal: if the design is binary then its entries are the replications of the treatments. In particular, if every treatment has replication r then

$$L = rI_v - \frac{1}{k}\Lambda.$$

Suppose that x and z are contrasts with $Lz = x$. Consider the linear combination $z' X_T' QY$ of the responses. Expectation is linear, so

$$\mathbb{E}(z' X_T' QY) = z' X_T' Q \mathbb{E}(Y) = z' X_T' Q X_T \tau = z' L' \tau = x' \tau = \sum_{i \in \mathcal{T}} x_i \tau_i;$$

in other words, this linear combination is an *unbiased* estimator of $\sum_{i \in \mathcal{T}} x_i \tau_i$ in the sense that the expectation is equal to the value which we are trying to estimate.

The following theorem, known as the Gauss–Markov theorem, is one of the most important results from the theory of linear models.

Theorem 1 *Under the assumptions in Section 2, the above estimator for $x'\tau$ has minimum variance among linear unbiased estimators.*

5 Variance

What is this minimum variance? The assumptions about independence and variance in Section 2 imply that

$$\begin{aligned}
\mathrm{Var}(z' X_T' QY) &= (z' X_T Q)(z' X_T Q)' \sigma^2 \\
&= z' X_T' Q Q X_T z \sigma^2 \\
&= z' X_T' Q X_T z \sigma^2 \\
&= z' L z \sigma^2.
\end{aligned}$$

We would really prefer to express this variance in terms of x. We can do this by using the *generalized inverse* of L.

Since L is symmetric, it has a spectral decomposition

$$L = \sum \mu_i E_i,$$

where the μ_i are the eigenvalues of L and the matrices E_i are the orthogonal projectors onto the corresponding eigenspaces. Then the generalized inverse L^- of L is defined by

$$L^- = \sum_{\mu_i \neq 0} \frac{1}{\mu_i} E_i.$$

(Such a generalized inverse can be defined for any real symmetric matrix. If the matrix is invertible then the generalized inverse coincides with the inverse.)

We already know that L has an eigenvalue zero, so L^- is not a true inverse of L. However,

$$LL^- = L^-L = \begin{cases} \text{the identity on the column space of } L \\ \text{zero on the kernel of } L \end{cases}$$

and so we may treat L^- as a true inverse if we are dealing with a vector x of the form Lz. Hence the above variance $z'Lz\sigma^2$ is equal to $x'L^-x\sigma^2$.

(Note that if the design is connected then

$$LL^- = L^-L = I - \frac{1}{v}J,$$

where J is the all-1 matrix.)

For the elementary contrast $\tau_i - \tau_j$, we have $x_i = 1$, $x_j = -1$ and $x_l = 0$ if $l \neq i$ and $l \neq j$, so the variance is equal to

$$(L^-_{ii} - L^-_{ij} - L^-_{ji} + L^-_{jj})\sigma^2.$$

Theorem 2 *If design is connected and the information matrix has exactly two non-zero eigenvalues then there are positive constants c_1 and c_2 such that the variance of the estimator of $\tau_i - \tau_j$ is equal to $c_1 - c_2\lambda_{ij}$.*

Usually there is no such simple relationship between concurrence and variance.

6 Efficiency

We want the variances of estimators to be as small as possible. For reference, we compare our block design with a complete-block design using the same number of plots. This makes sense if our block design has equal replication. Then the complete-block design has $k = v$ and $b = r$, so its concurrence matrix is rJ and its information matrix is

$$L_{\text{CBD}} = r(I - \frac{1}{v}J).$$

If x is a contrast then $L_{\text{CBD}}x = rx$ and so the variance of the estimator of $x'\tau$ is equal to

$$\frac{x'x}{r}\sigma^2_{CBD},$$

where σ^2_{CBD} is the variance of each response in the hypothetical complete-block design.

The relative *efficiency* for contrast $x'\tau$ in our block design is defined to be the ratio of the variances in the two designs, which is

$$\frac{\frac{x'x}{r}\sigma^2_{CBD}}{x'L^-x\sigma^2}.$$

Thus low variance in our block design corresponds to high efficiency.

In practice, neither σ^2 nor σ^2_{CBD} is known, so we define the *efficiency factor* for the contrast $x'\tau$ to be

$$\frac{x'x}{rx'L^-x}.$$

It can be shown that all efficiency factors lie between 0 and 1. The value zero is conventionally given if $Lx = 0$, which happens only if the design is disconnected.

Theorem 3 *For a binary balanced block design, every efficiency factor is equal to*

$$\frac{v}{v-1}\frac{k-1}{k}.$$

In the equireplicate case, efficiency factors have a particularly simple form for eigenvectors of L. If $Lx = r\varepsilon x$ then $L^-x = (r\varepsilon)^{-1}x$ and so

$$\frac{x'x}{rx'L^-x} = \varepsilon.$$

In this case, ε is called a *canonical efficiency factor*.

Counting according to multiplicities, there are $v-1$ canonical efficiency factors. Their harmonic mean is denoted A. The following theorem shows why this is important.

Theorem 4 *In a connected equireplicate block design, the average of the variances of the estimators of elementary contrasts is equal to*

$$\frac{2\sigma^2}{rA}.$$

The proofs of Theorems 3 and 4 are given in [1].

7 Optimality

A binary equireplicate block design with equal block sizes is said to be *A-optimal* if it maximizes the value of A over all binary designs with the same values of v, b, r and k.

Theorem 5 *The following block designs are A-optimal:*

(a) balanced designs;

(b) duals of A-optimal designs;

(c) square lattice designs;

(d) group divisible designs in which the concurrence between pairs in different groups is one more than the concurrence between pairs in the same group, and either

  (i) there are two groups, or

  (ii) each block contains equal numbers of treatments from each group.

See [2] for more about optimality.

References

[1] J. A. John, *Cyclic Designs*, Chapman and Hall, London, 1987.

[2] K. R. Shah and B. K. Sinha, *Theory of Optimal Designs*, Springer,New York, 1989.

R. A. Bailey

September 26, 2002