3.3 Random variables

This section is a very brief recall of the facts we need about random variables. I will not attempt to explain what they are.

Given a set Γ, suppose that there are random variables $Y(\gamma)$, for γ in Γ, with a joint distribution. Then functions of two or more of these, such as $Y(\alpha)Y(\beta)$, are also random variables. The components $Y(\gamma)$ can be assembled into a random vector Y.

The random variable $Y(\gamma)$ has an *expectation* $\mathbb{E}(Y(\gamma))$ in $\mathbb{R}$. (I assume that the sum or integral which defines the expectation does converge for all the random variables that we consider.) Then we define $\mathbb{E}(Y)$ in $\mathbb{R}^\Gamma$ by

$$\mathbb{E}(Y)(\gamma) = \mathbb{E}(Y(\gamma)).$$

The main result we need about expectation is the following.

Proposition 3.3 *Expectation is* affine *in the sense that if* $M \in \mathbb{R}^{\Delta\times\Gamma}$ *and* $f \in \mathbb{R}^\Delta$ *then*

$$\mathbb{E}(MY + f) = M\,\mathbb{E}(Y) + f.$$

The *covariance* of random variables $Y(\alpha)$ and $Y(\beta)$ is defined by

$$\mathrm{cov}\,(Y(\alpha), Y(\beta)) = \mathbb{E}\left[(Y(\alpha) - \mathbb{E}(Y(\alpha)))\,(Y(\beta) - \mathbb{E}(Y(\beta)))\right].$$

The covariance of $Y(\alpha)$ with itself is called the *variance* of $Y(\alpha)$, written $\mathrm{Var}(Y(\alpha))$. The *covariance matrix* $\mathrm{Cov}(Y)$ of the random vector Y is defined by

$$\mathrm{Cov}(Y)(\alpha,\beta) = \mathrm{cov}\,(Y(\alpha), Y(\beta)).$$

Lemma 3.4 *Covariance is* bi-affine *in the sense that*

(i) *if* $g \in \mathbb{R}^\Gamma$ *then* $\mathrm{Cov}(Y+g) = \mathrm{Cov}(Y)$;

(ii) *if* $M \in \mathbb{R}^{\Delta\times\Gamma}$ *then* $\mathrm{Cov}(MY) = M\,\mathrm{Cov}(Y)M'$.

Proof (i) Put $Z = Y + g$. The $\mathbb{E}(Z) = \mathbb{E}(Y) + g$, by Proposition 3.3, so $Z - \mathbb{E}(Z) = Y - \mathbb{E}(Y)$, in particular $Z(\alpha) - \mathbb{E}(Z(\alpha)) = Y(\alpha) - \mathbb{E}(Y(\alpha))$.

(ii) By (i), we can assume that $\mathbb{E}(Y) = 0$. Then

$$\begin{aligned}
\mathrm{Cov}(MY)(\alpha,\beta) &= \mathrm{cov}((MY)(\alpha),(MY)(\beta)) \\
&= \mathbb{E}[((MY)(\alpha))((MY)(\beta))] \\
&= \mathbb{E}\left[\left(\sum_{\gamma\in\Gamma} M(\alpha,\gamma)Y(\gamma)\right)\left(\sum_{\delta\in\Gamma} M(\beta,\delta)Y(\delta)\right)\right]
\end{aligned}$$

$$\begin{aligned} &= \sum_{\gamma}\sum_{\delta} M(\alpha,\gamma)\left[\mathbb{E}(Y(\gamma)Y(\delta))\right]M'(\delta,\beta) \\ &= \sum_{\gamma}\sum_{\delta} M(\alpha,\gamma)\operatorname{Cov}(Y)(\gamma,\delta)M'(\delta,\beta) \\ &= \left(M\operatorname{Cov}(Y)M'\right)(\alpha,\beta). \quad \blacksquare \end{aligned}$$

3.4 Estimation and variance

Put

$$V_B = \left\{v \in \mathbb{R}^\Omega : v(\alpha) = v(\beta) \text{ if } \alpha \text{ and } \beta \text{ are in the same block}\right\}.$$

Then the characteristic functions of the blocks form an orthogonal basis for V_B, so $\dim V_B = b$. Also

$$w \in V_B^\perp \iff \sum_{\alpha\in\delta} w(\alpha) = 0 \quad \text{for each block } \delta.$$

Let P and Q be the orthogonal projectors onto V_B and $V_B^\perp$. It can be easily checked that $P = k^{-1}B$ and $Q = I - P$.

Let $Y(\omega)$ be the response on plot ω when our incomplete-block design is used for an experiment. We assume that

$$\mathbb{E}(Y) = X\tau + h,$$

where τ is an unknown vector in $\mathbb{R}^\Theta$ and h is an unknown vector in V_B, and

$$\operatorname{Cov}(Y) = I\sigma^2,$$

where σ^2 is an unknown positive constant. That is, the expectation of $Y(\omega)$ is the sum of two parts, one depending on the treatment applied to ω and the other depending on the block containing ω; and the responses on different plots are uncorrelated and all have the same variance.

We want to use the observed values of the $Y(\omega)$ from the experiment to estimate τ.

In $\mathbb{R}^\Theta$, let U_0 be the space spanned by χ_Θ. Now

$$X\chi_\Theta = \chi_\Omega \in V_B,$$

so we cannot estimate τ: the best we can hope to do is to estimate τ up to a multiple of χ_Θ. Then we could estimate differences such as $\tau(\theta) - \tau(\eta)$.

Definition A vector x in $\mathbb{R}^\Theta$ is a *contrast* if $x \in U_0^\perp$. It is a *simple* contrast if there are θ, η in Θ such that $x(\theta) = 1$, $x(\eta) = -1$ and $x(\zeta) = 0$ for ζ in $\Theta \setminus \{\theta,\eta\}$.

We want to estimate linear combinations such as $\sum_\theta x(\theta)\tau(\theta)$ for x in $U_0^\perp$. In order to use the results of the previous section in a straightforward way, it is convenient to make a slight shift of perspective on our vectors. I have defined x to be a function from Θ to $\mathbb{R}$. However, the definitions of the action of a matrix on a vector, and of matrix multiplication, are consistent with the idea that x is a column vector, that is, an element of $\mathbb{R}^{\Theta\times\{1\}}$. So we can define the transpose of x as a matrix x' in $\mathbb{R}^{\{1\}\times\Theta}$. Then

$$\sum_{\theta\in\Theta} x(\theta)\tau(\theta) = \langle x,\tau\rangle = x'\tau.$$

Definition An *unbiased estimator* for $x'\tau$ is a function of Y and of the design (but not of τ, h or σ^2) whose expectation is equal to $x'\tau$.

Theorem 3.5 *If there is a vector z in $\mathbb{R}^\Theta$ with $X'QXz = x$ then $z'X'QY$ is an unbiased estimator for $x'\tau$ and its variance is $z'X'QXz\sigma^2$.*

Proof

$$\begin{aligned}
\mathbb{E}(z'X'QY) &= z'X'Q\mathbb{E}(Y), \qquad \text{by Proposition 3.3,}\\
&= z'X'Q(X\tau+h)\\
&= z'X'QX\tau, \qquad \text{because } Qh=0,\\
&= x'\tau
\end{aligned}$$

because $Q' = Q$. Then, by Lemma 3.4,

$$\begin{aligned}
\operatorname{Var}(z'X'QY) &= z'X'Q(I\sigma^2)Q'Xz\\
&= (z'X'Q^2Xz)\sigma^2\\
&= z'X'QXz\sigma^2
\end{aligned}$$

because Q is idempotent. ■

Theorem 3.6 *The kernel of $X'QX$ is spanned by the characteristic functions of the connected components of the treatment-concurrence graph.*

Proof Let z be in $\mathbb{R}^\Theta$. If $z \in \ker X'QX$ then $\langle QXz, QXz\rangle = z'X'Q'QXz = z'X'QXz = 0$; but $\langle\ ,\ \rangle$ is an inner product, so $QXz = 0$. Thus

$$\begin{aligned}
z \in \ker X'QX &\iff QXz = 0\\
&\iff PXz = Xz\\
&\iff Xz \in V_B\\
&\iff z(\theta) = z(\eta) \text{ whenever } \Lambda(\theta,\eta) > 0\\
&\iff z \text{ is constant on each component of the treatment-concurrence graph.}
\end{aligned}$$

■

Corollary 3.7 *If an incomplete-block design is connected then* $\mathrm{Im}(X'QX) = U_0^\perp$.

Proof If the design is connected then $\ker X'QX = U_0$. But $X'QX$ is symmetric, so $\mathrm{Im}(X'QX) = (\ker X'QX)^\perp$. ■

Definition The matrix $X'QX$ is the *information matrix* of the design. Write $L = X'QX$.

The matrix kL is sometimes called the *Laplacian*, particularly when $k = 2$.

Theorem 3.5 says that if $Lz = x$ then $z'X QY$ is an unbiased estimator of $x'\tau$ with variance $z'Lz\sigma^2$. Recall from Section 2.2 that L has a generalized inverse L^- such that $LL^-L = L$. Thus $z'Lz\sigma^2 = z'LL^-Lz\sigma^2 = x'L^-x\sigma^2$ because L is symmetric, so we obtain an expression for the variance of the estimator of $x'\tau$ in terms of x rather than z. In particular, if the design is connected then we can estimate every difference $\tau(\theta) - \tau(\eta)$ and the variance of the estimator is

$$\mathrm{Var}(\widehat{\tau(\theta) - \tau(\eta)}) = \left(L^-(\theta,\theta) - L^-(\theta,\eta) - L^-(\eta,\theta) + L^-(\eta,\eta)\right)\sigma^2, \qquad (3.2)$$

where we have used the statisticians' notation $\widehat{\quad}$ for an estimator.

In an equi-replicate block design,

$$\begin{aligned} L = X'QX &= X'(I-P)X \\ &= X'X - k^{-1}X'BX \\ &= rI_\Theta - k^{-1}\Lambda. \end{aligned} \qquad (3.3)$$

Example 3.1 revisited Let $x = \chi_1 - \chi_2$. Since treatments 1 and 2 always occur together in a block, $Nx = 0$. We say that x is "orthogonal to blocks" to indicate that $\sum_{\alpha\in\delta}(Xx)(\alpha) = 0$ for every block δ; that is, $Xx \in V_B^\perp$ and $PXx = 0$. Equivalently, the 1- and 2-rows of Λ are identical, so $\Lambda x = 0$.

Now Equation (3.3) gives $Lx = rx = 2x$ so we may take $z = \frac{1}{2}x$ in Theorem 3.5. Then $z'X'QY = \frac{1}{2}x'X'QY = \frac{1}{2}x'X'(I-P)Y = \frac{1}{2}x'X'Y$ so we estimate $\tau(1) - \tau(2)$ by

$$\frac{Y(\omega_1) - Y(\omega_2) + Y(\omega_5) - Y(\omega_6)}{2}.$$

In other words, we take the difference between the response on treatments 1 and 2 in each block where they occur, and average these differences. The variance of this estimator is $(\sigma^2 + \sigma^2 + \sigma^2 + \sigma^2)/2^2 = \sigma^2$.

Now put $x = \chi_1 - \chi_3$. This contrast is not orthogonal to blocks, so we will have to do some explicit calculations. We have

$$L = 2I - \frac{1}{4}\begin{bmatrix} 2 & 2 & 1 & 1 & 1 & 1 \\ 2 & 2 & 1 & 1 & 1 & 1 \\ 1 & 1 & 2 & 2 & 1 & 1 \\ 1 & 1 & 2 & 2 & 1 & 1 \\ 1 & 1 & 1 & 1 & 2 & 2 \\ 1 & 1 & 1 & 1 & 2 & 2 \end{bmatrix} = \frac{1}{4}\begin{bmatrix} 6 & -2 & -1 & -1 & -1 & -1 \\ -2 & 6 & -1 & -1 & -1 & -1 \\ -1 & -1 & 6 & -2 & -1 & -1 \\ -1 & -1 & -2 & 6 & -1 & -1 \\ -1 & -1 & -1 & -1 & 6 & -2 \\ -1 & -1 & -1 & -1 & -2 & 6 \end{bmatrix}.$$

Put $z = \frac{1}{12}(7\chi_1 + \chi_2 - 7\chi_3 - \chi_4)$. Direct calculation shows that $Lz = x$, so we estimate $\tau(1) - \tau(3)$ by $z'X'QY$. Now, the effect of Q is to subtract the block average from every entry in a block, so

$$\begin{aligned} 12z'X'QY &= 7Y(\omega_1) + Y(\omega_2) - 7Y(\omega_3) - Y(\omega_4) \\ &\quad + 5Y(\omega_5) - Y(\omega_6) - 2Y(\omega_7) - 2Y(\omega_8) \\ &\quad - 5Y(\omega_9) + Y(\omega_{10}) + 2Y(\omega_{11}) + 2Y(\omega_{12}). \end{aligned}$$

Now the response on *every* plot contributes to the estimator of $\tau(1) - \tau(3)$, whose variance is

$$\frac{\sigma^2}{12^2}(7^2 + 1^2 + 7^2 + 1^2 + 5^2 + 1^2 + 2^2 + 2^2 + 5^2 + 1^2 + 2^2 + 2^2) = \frac{7\sigma^2}{6}.$$

Since so many more responses are involved, it is, perhaps, not surprising that this variance is greater than the variance of the estimator of $\tau(1) - \tau(2)$. (This issue will be discussed in Section **??**.)

Finally, we look at a non-simple contrast. Put $x = \chi_1 + \chi_2 - \chi_3 - \chi_4$. Direct calculation shows that $Lx = \frac{3}{2}x$, so the estimator of $\tau(1) + \tau(2) - \tau(3) - \tau(4)$ is $(2/3)x'X'QY$, which is

$$\frac{1}{3}\begin{pmatrix} 2Y(\omega_1) + 2Y(\omega_2) - 2Y(\omega_3) - 2Y(\omega_4) \\ + Y(\omega_5) + Y(\omega_6) - Y(\omega_7) - Y(\omega_8) \\ - Y(\omega_9) - Y(\omega_{10}) + Y(\omega_{11}) + Y(\omega_{12}) \end{pmatrix}.$$

This is the sum of the estimators of $\tau(1) - \tau(3)$ and $\tau(2) - \tau(4)$, which is no surprise, because estimation is linear in x. Its variance is $(8/3)\sigma^2$. ■