NOTES

AN EXTENSION OF THE OPTIMUM PROPERTY OF THE SEQUENTIAL PROBABILITY RATIO TEST

By M. A. Girshick[1]

Stanford University

Let $f(x, \theta)$ be a family of densities or discrete probability functions depending on the parameter θ. Let H_0 be the hypothesis $\theta = \theta_0$ and H_1 the hypothesis that $\theta = \theta_1$. A sequential probability ratio test of H_0 versus H_1 is defined by two numbers A and B. After drawing the mth observation, sampling is continued if

$$B < \prod_{i=1}^{m} \frac{f(x_i, \theta_1)}{f(x_i, \theta_0)} < A, \tag{1}$$

where $x_1, \cdots, x_m$ are the first m observations. If the probability ratio is at least equal to A, H_1 is accepted, and if it is not greater than B, H_0 is accepted.

For any sequential procedure T, let the operating characteristic be

$$L(\theta, T) = \Pr\{\text{Accepting } H_0 \mid \theta, T\}, \tag{2}$$

and let $\mathcal{E}_\theta(n \mid T)$ be the expected number of observations required by T when sampling from $f(x, \theta)$. The so-called optimum property (see [5], for instance) of a sequential probability ratio test, say T^*, is that if $L(\theta_0, T) \geqq L(\theta_0, T^*)$ and $L(\theta_1, T) \leqq L(\theta_1, T^*)$, then

$$\mathcal{E}_{\theta_0}(n \mid T) \geqq \mathcal{E}_{\theta_0}(n \mid T^*), \qquad \mathcal{E}_{\theta_1}(n \mid T) \geqq \mathcal{E}_{\theta_1}(n \mid T^*).$$

In many cases this optimum property can be extended to all values of the parameter. Suppose $\theta_0 < \theta_1$, and let $\bar{\theta}$ be a number to be defined later such that $\theta_0 < \bar{\theta} < \theta_1$. Under conditions stated below, we give the extended optimum property. If

$$\begin{aligned} L(\theta, T) &\geqq L(\theta, T^*), \qquad \theta < \bar{\theta}, \\ L(\theta, T) &\leqq L(\theta, T^*), \qquad \theta > \bar{\theta}, \end{aligned} \tag{3}$$

for all $\theta \neq \bar{\theta}$, then

$$\mathcal{E}_\theta(n \mid T) \geqq \mathcal{E}_\theta(n \mid T^*) \tag{4}$$

Received May 9, 1956; revised November 15, 1957.

[1] The result reported in this note was mentioned by the late M. A. Girshick to several of his colleagues, but was unpublished at the time of his death. Since I think the result is of sufficient interest to be in the literature, I have taken the liberty of writing this note in Girshick's name. T. W. Anderson.

The research was sponsored by the Office of Naval Research.

for all θ. Inequalities (3) indicate the premise that T is everywhere as good as T^* in the sense that the operating characteristic for T is at least as high as for T^* for θ on one side of $\bar{\theta}$ and is as least as low as for T^* on the other side of $\bar{\theta}$. Then T^* is everywhere as good as T in terms of expected number of observations.

To demonstrate the property we assume that for $\theta \neq \bar{\theta}$ there is a unique nonzero root, say $h(\theta)$, of

$$\mathcal{E}_\theta \left[\frac{f(x, \theta_1)}{f(x, \theta_0)}\right]^h = 1, \tag{5}$$

and that $h(\theta) > 0$ for $\theta < \bar{\theta}$ and $h(\theta) < 0$ for $\theta > \bar{\theta}$. (See [4] for discussion of the assumption and of the technique used here.) This implies that given θ_0 and θ_1 the value of $\bar{\theta}$ for which the assumption holds is unique. We make the further assumption that for each θ there is a θ' such that

$$\left[\frac{f(x, \theta_1)}{f(x, \theta_0)}\right]^{h(\theta)} f(x, \theta) = f(x, \theta'). \tag{6}$$

We now prove (4) for $\theta < \bar{\theta}$ by assuming (3) for θ and θ'. Since

$$h(\theta') = -h(\theta),$$

we have $\theta' > \bar{\theta}$. The sequential probability ratio test T^* defined by (1) can also be defined by

$$B^{h(\theta)} < \prod_{i=1}^{m} \left[\frac{f(x_i, \theta_1)}{f(x_i, \theta_0)}\right]^{h(\theta)} < A^{h(\theta)} \tag{7}$$

or by

$$B^{h(\theta)} < \prod_{i=1}^{m} \frac{f(x_i, \theta')}{f(x_i, \theta)} < A^{h(\theta)}. \tag{8}$$

Then (4) follows by the usual optimum property because T^* is a sequential probability ratio test for testing hypothesis θ versus the hypothesis θ'. For $\theta > \bar{\theta}$ a similar argument can be used.

The conditions assumed for this extended property are satisfied by many distributions. In particular the existence of such so-called conjugate pairs for distributions of the Koopman-Darmois form has been shown [2]. Savage [3] has shown that the assumptions restrict the families to have a certain exponential form (which includes the Koopman-Darmois form). This note makes explicit Blasbalg's statement [1] that a sequential probability ratio test is optimum at an infinity of parameter points.

REFERENCES

[1] H. Blasbalg, "Sequential analysis," *Ann. Math. Stat.*, Vol. 28 (1957), pp. 1024–1028.
[2] M. A. Girshick, "Contributions to the theory of sequential analysis II, III," *Ann. Math. Stat.*, Vol. 17 (1946), pp. 282–298.

[3] L. J. Savage, "When different pairs of hypotheses have the same family of likelihood-ratio test regions," *Ann. Math. Stat.*, Vol. 28 (1957), pp. 1028–1032.
[4] A. Wald, *Sequential Analysis*, John Wiley and Sons, New York, 1947.
[5] A. Wald and J. Wolfowitz, "Optimum character of the sequential probability ratio test," *Ann. Math. Stat.*, Vol. 19 (1948), pp. 326–339.

A NOTE ON BALANCED DESIGNS

By V. R. Rao

University of Bombay

0. Summary. It is proved that a necessary and sufficient condition for a general design to be balanced is that the matrix of the adjusted normal equations for the estimates of treatment effects has $v - 1$ equal latent roots other than zero.

1. Estimates and their properties. We consider a design whose incidence matrix is $N_{v \times b} = [n_{ij}]$ in which the ith treatment is replicated r_i times and the blocks are of sizes $k_1, \cdots, k_b$. With the usual assumptions, the adjusted normal equations for the treatment effects are

$$Q = C\hat{\tau}, \tag{1.1}$$

where

$$Q = T - N \operatorname{diag}\left(\frac{1}{k_1}, \cdots, \frac{1}{k_b}\right) B \tag{1.2}$$

and

$$C = \operatorname{diag}(r_1, \cdots, r_v) - N \operatorname{diag}\left(\frac{1}{k_1}, \cdots, \frac{1}{k_b}\right) N' \tag{1.3}$$

with the condition

$$E_{1v}\hat{\tau} = 0 \tag{1.4}$$

(where E_{pq} denotes a $p \times q$ matrix with all its elements as unity).

It is well known that if rank $C = v - t$, a set of $t - 1$ independent treatment contrasts are not estimable. But if rank $C = v - 1$ every contrast is estimable and in this case the design is said to be connected.

If the design is connected there are $v - 1$ non-zero latent roots, say, $\lambda_1, \lambda_2, \cdots, \lambda_{v-1}$. As the rows of C add to zero, $(v^{-1/2}, \cdots, v^{-1/2})$ is the latent vector corresponding to the root zero.

Let

$$L = \left[\frac{L_1}{v^{-\frac{1}{2}}E_{1v}}\right] = \left[\frac{(l_{ij})}{v^{-\frac{1}{2}}E_{1v}}\right] \tag{1.5}$$

be an orthogonal matrix transforming C into diagonal form.

Received February 28, 1957; revised July 15, 1957.