

# A NOTE ON ROBUST TESTING BY INDEXING THE PARAMETER SPACE

by

Erling Sverdrup

A preliminary test is sometimes performed to obtain a simple model with few parameters. This should ostensibly result in a better performance in the subsequent treatment of the observations. For obvious reasons such a procedure is questionable. It is felt that one should stick to the original a priori model with the original parameter vector  $\theta$ . In this note a third method is proposed. Stick to original parameters  $\theta$ , but fix attention on an index  $\phi(\theta)$  which sums up some essential features of the model. (This principle has been advocated by Goldstein [1981].)

Thus in a one-way lay-out with population means  $\xi_1, \dots, \xi_s$ , it may not be realistic to assume that they depend linearly on  $t_1, \dots, t_s$  even if there is a "simple" relation between  $\xi_i$  and  $t_i$ . Hence we may be justified in concentrating attention on a "regression index"

$$\beta(\xi) = \frac{\sum n_i \xi_i (t_i - \bar{t})}{\sum n_i (t_i - \bar{t})^2}$$

( $n_i$  = number of observation in each group,  $\bar{t}$  the weighted average).

Such a procedure is familiar in the case of a two-way lay-out with means  $\xi_{ij}$ . The dependence on the factor "j" is often studied by means of the "main effect" index

$$\beta_j(\xi) = \bar{\xi}_{+j} - \bar{\xi}_{++} = \frac{1}{r} \sum_i \xi_{ij} - \frac{1}{rs} \sum_{i,j} \xi_{ij}$$

( $r$  and  $s$  are the number of levels of the factors).

This is justified if the interaction is not too dominating.

The method is certainly familiar in the case of non-parametric models with unspecified distribution function  $F$ . With little a priori knowledge one should expect poor performance, but obtains high power by fixing attention on a very simple index, viz. the median  $\mu(F)$ .

The following computations show that such is also the case in the parametric situations mentioned above.

2. The problem of non-linearity

( $\alpha$ ) Rich model with strict hypothesis

$X_{ij}$ ;  $j = 1, 2, \dots, n_i$ ;  $i = 1, 2, \dots, s$ ; are independent normal  
 $EX_{ij} = \xi_i$ ;  $\text{var } X_{ij} = \sigma^2$ . The hypothesis is  $\xi_1 = \dots = \xi_s$ . Let

$$Q_0 = \sum (X_{ij} - \bar{X}_i)^2, \quad Q_1 = \sum n_j (\bar{X}_j - \bar{X})^2 \quad (1)$$

( $\bar{X}_i$  is a group mean,  $\bar{X}$  is the total mean.)

The hypothesis is rejected if

$$F = Q_1(n-s)/Q_0(s-1) > f_\epsilon(s-1, n-s), \quad (n = \sum n_i), \quad (2)$$

where  $f_\epsilon$  is adjusted to a level  $\epsilon$ . The power function is given by the eccentricity

$$\kappa = \sum n_i (\xi_i - \bar{\xi})^2 / \sigma^2 \quad (3)$$

where  $\bar{\xi}$  is the weighed mean,

( $\beta$ ) Strict (parsimonious) model

To the  $s$  groups are attached numbers  $t_1, \dots, t_s$  and it is assumed (perhaps after a preliminary test) that

$$\xi_i = EX_{ij} = \alpha + \beta(t_i - \bar{t}) \quad (4)$$

The hypothesis then becomes  $\beta = 0$ . Let

$$\hat{\beta} = \sum n_i \bar{X}_i (t_i - \bar{t}) / M, \quad M = \sum n_i (t_i - \bar{t})^2, \\ Q_2 = \sum n_i (\bar{X}_i - \bar{X} - \hat{\beta}(t_i - \bar{t}))^2. \quad (5)$$

The hypothesis is rejected, with level  $\epsilon$ , if

$$F = \hat{\beta}^2 M(n-2) / (Q_0 + Q_2) > f_\epsilon(1, n-2) \quad (6)$$

The power function is given by the eccentricity

$$\kappa = M \beta^2 / \sigma^2 \tag{7}$$

(neglecting the effect of the preliminary testing).

( $\gamma$ ) Rich model with generous hypothesis

We use the same rich model as in ( $\alpha$ ), but we find that an essential feature of the parameters  $\xi_1, \dots, \xi_s$  can be described by

$$\beta(\xi) = \sum n_i \xi_i (t_i - \bar{t}) / M \tag{8}$$

The null-hypothesis is  $\beta(\xi) = 0$ . This hypothesis is rejected if

$$F = \hat{\beta}^2 M(n-s) / Q_0 > f_\epsilon(1, n-s) \tag{9}$$

The power function has eccentricity

$$\kappa = M \beta^2 / \sigma^2 \tag{10}$$

It is seen that the powers in ( $\beta$ ) and ( $\gamma$ ) differ only in the degrees of freedom  $n-s$  and  $n-2 = (n-s) + (s-2)$ . This difference is unimportant for moderately large  $n$ .

Let us now see how the three tests behave if the strict regression model (4), holds. We set all  $n_i = 5$ ,  $s = 6$ ,  $n = 30$ . Furthermore  $t_1 = 1$ ,  $t_2 = 3$ ,  $t_3 = 4$ ,  $t_4 = 8$ ,  $t_5 = 9$ ,  $t_6 = 11$ . We compute the powers for  $\beta/\sigma = 0.2$  and level  $\epsilon = 0,05$ . We get in all three cases  $\kappa = 15.2$ .

|   | Power  |
|---|--------|
| ( $\alpha$ ) Rich model, strict hypothesis                    | 0.7742 |
| ( $\beta$ ) Strict model                                      | 0.9642 |
| ( $\gamma$ ) <u>Robust method</u> : Rich model, generous hyp. | 0.9623 |

Thus the test ( $\gamma$ ) is robust, very little is lost by playing safe and using the rich, i.e. unrestrictive model.

### 3. The problem of interaction

#### ( $\alpha$ ) Rich model with strict hypothesis

The  $X_{ijk}$ ;  $i = 1, 2, \dots, r$ ;  $j = 1, 2, \dots, s$ ;  $k = 1, 2, \dots, m$ ; are independent normal,  $EX_{ijk} = \xi_{ij}$ ,  $\text{var } X_{ijk} = \sigma^2$ . We are interested in the dependence of  $X_{ij}$  on the factor represented by "j". The hypothesis is  $\xi_{i1} = \xi_{i2} = \dots = \xi_{is}$ ;  $i = 1, \dots, r$ . The Fisher F-testing is in this case with  $r(s-1)$  and  $v = rs(m-1)$  degrees of freedom and the power function has eccentricity

$$\kappa = m \sum (\xi_{ij} - \bar{\xi}_{+j})^2 \quad (11)$$

#### ( $\beta$ ) Strict (parsimonious) model

We assume (perhaps after a preliminary test) that we have no interaction, i.e.

$$\xi_{ij} = \xi + \alpha_i + \beta_j; \quad \sum \alpha_i = \sum \beta_j = 0 \quad (12)$$

and the null-hypothesis is  $\beta_1 = \dots = \beta_s = 0$ . The Fisher test is with  $s-1$  and  $\mu = v + (r-1)(s-1)$  degrees of freedom and the power function has eccentricity

$$\kappa = m r \sum \beta_j^2 \quad (13)$$

(neglecting the effect of the preliminary testing).

#### ( $\gamma$ ) Rich model with generous hypothesis

We use the same model as in ( $\alpha$ ), but we find that an essential feature of the dependence of  $X_{ij}$  on  $j$  can be described by

$$\beta_j(\xi) = \bar{\xi}_{+j} - \bar{\xi}_{++}; \quad j = 1, 2, \dots, s \quad (14)$$

This means that we admit that the interactions may be different from 0, but they are not dominating. The null hypothesis is  $\beta_1(\xi) = \beta_2(\xi) = \dots = 0$ . We use the test

$$v m r \sum (\bar{X}_{+j} - \bar{X}_{++})^2 / (s-1) \sum_{i,j,k} (X_{ijk} - \bar{X}_{ij})^2 > f_{\epsilon}(s-1, v)$$

with  $s-1$  and  $v = rs(m-1)$  degrees of freedom. The power function has eccentricity

$$\kappa = m r \sum_j (\bar{\xi}_{+j} - \bar{\xi}_{++})^2$$

Let us now see how the three tests behave if the strictly additive model (12) is true. Let  $r = 4$ ,  $s = 3$ ,  $m = 5$ . We compute the powers for  $\sum \beta_j^2 / \sigma^2 = 0.6$  with level 0.05. We get in all cases

$$\kappa = m r \sum \beta_j^2 / \sigma^2 = 12$$

|   | Power  |
|---|--------|
| ( $\alpha$ ) Rich model, strict hypothesis                    | 0.6023 |
| ( $\beta$ ) Strict model                                      | 0.8639 |
| ( $\gamma$ ) <u>Robust method</u> ; Rich model, generous hyp. | 0.8612 |

Hence, again, testing the index parameters  $\beta_j(\xi)$  results in a robust procedure, i.e. it is based on an unrestrictive and hence "safe", but the power is almost as high as the power of the conventional method, if the restrictive model is true.

Of course, whether indexing should be used must be based on a priori considerations. Do the indices reflect an essential feature? If not, the methods should not be used. Other methods are available (e.g. using "smooth" dependence of the interactions on the main effects, which will certainly result in high power [See Scheffé (1959) p. 131-132]).

Any statistical method must be constructed on the basis of the statistical situation which the statistician is faced with in a special case.

It may be objected that the interpretation of the regression index  $\beta(\xi)$  and main effect indices  $\beta_j(\xi)$  may be diffuse. That may be true in the sense that most aggregate measures of essentials are arbitrary. However, the interpretation of the tests 2( $\beta$ ) and 3( $\beta$ ) are diffuse if the models are unrealistic.

Goldstein, Harald (1981). Robust Inference in Contingency Tables, Oslo, University of Oslo, Institute of Economics.

Scheffé, Henry (1959). The analysis of Variance, New York, John Wiley & Sons, Inc.