29 Tensor Methods in Statistics *P. McCullagh* (1987)
30 Transformation and Weighting in Regression *R.J. Carroll and D. Ruppert* (1988)
31 Asymptotic Techniques for Use in Statistics *O.E. Barndoff-Nielson and D.R. Cox* (1989)
32 Analysis of Binary Data, 2nd edition *D.R. Cox and E.J. Snell* (1989)
33 Analysis of Infectious Disease Data *N.G. Becker* (1989)
34 Design and Analysis of Cross-Over Trials *B. Jones and M.G. Kenward* (1989)
35 Empirical Bayes Method, 2nd edition *J.S. Maritz and T. Lwin* (1989)
36 Symmetric Multivariate and Related Distributions *K.-T. Fang, S. Kotz and K. Ng* (1989)
37 Generalized Linear Models, 2nd edition *P. McCullagh and J.A. Nelder* (1989)
38 Cyclic Designs *J.A. John* (1987)
39 Analog Estimation Methods in Econometrics *C.F. Manski* (1988)
40 Subset Selection in Regression *A.J. Miller* (1990)
41 Analysis of Repeated Measures *M. Crowder and D.J. Hand* (1990)
42 Statistical Reasoning with Imprecise Probabilities *P. Walley* (1990)
43 Generalized Additive Models *T.J. Hastie and R.J. Tibshirani* (1990)
44 Inspection Errors for Attributes in Quality Control *N.L. Johnson, S. Kotz and X. Wu* (1991)
45 The Analysis of Contingency Tables, 2nd edition *B.S. Everitt* (1992)
46 Analysis of Quantal Response Data *B.J.T. Morgan* (1992)

(Full details concerning this series are available from the Publishers.)

Analysis of Quantal Response Data

B. J. T. MORGAN
Professor of Applied Statistics
Institute of Mathematics and Statistics
University of Kent
Canterbury

EXHIBIT
23
Rausser
depobook.com

CHAPMAN & HALL
London · Glasgow · New York · Tokyo · Melbourne · Madras

EXHIBIT
D
Blumberg No. 5208

the model to the data. Why does this model result in under-dispersion?

6.35 (Ridout and Morgan, 1991.) Close examination of the data of Table 6.12 suggests that responses may favour 6 months and 12 months. Such digit preference is indeed very strongly indicated in more extensive sets of fecundability data. Consider how you might model digit preference in this application.

6.36 Maruani and Schwartz (1983) consider the following mixture model for **prospective** fecundability data:

$$f(p) = (1-\eta)\frac{p^{\alpha-1}(1-p)^{\beta-1}}{B(\alpha,\beta)}, \quad \text{for } 0 \leqslant p \leqslant 1$$

with an additional mass of η at $p = 0$.

Why is this particular model inappropriate for **retrospective** data, as in fact have been presented in Table 6.12? For one set of data they estimate $\hat{\eta} = 0.00117\,(0.00023)$. Is it likely that the model requires $\eta > 0$ for these data? Discuss potential problems with fitting this model to data.

6.37 Discuss how the $\tilde{X}^2$ statistic of equation (6.24) has adjusted the Cochran/Armitage statistic. When will the historical data have the greatest impact on $\tilde{X}^2$?

6.38 (Hoel and Yanagawa, 1986.) The analysis of section 6.7.4 is based upon a logit model. Obtain the corresponding test statistic when an exponential model is used.

CHAPTER 7

Non-parametric and robust methods

7.1 Introduction

The history of the statistical analysis of quantal response data is documented by Finney (1971, pp 38–42). The maximum-likelihood fitting of the probit model stems from the papers of Fisher (1935) and Bliss (1938). However, non-parametric approaches to estimating the ED_{50} have a generally earlier origin, in papers such as Behrens (1929), Kärber (1931) and Spearman (1908). Computationally, the early non-parametric measures are far simpler than those which result from an iterative maximum-likelihood fit, and this was clearly a matter of great importance before the computer era.

As we saw in Chapter 4, the maximum-likelihood fitting of extended models can be difficult. However, today access to computers makes the fitting of simple models, such as the logit or probit, to 'well-behaved' data a formality. Nevertheless, modern developments in the areas of robustness and influence have resulted in a re-examination of the relative merits of parametric and non-parametric procedures.

The emphasis in this chapter is partly on measurement of the ED_{50} of substances presented for assay as part of a routine testing program. Such experiments may well be seen just as pilot studies, and doses may be set conservatively to cover a wide range, quite often resulting in a number of instances of 0% and 100% response. If too high doses are employed then there may be additional problems arising from high dose toxicity effects. Small numbers of individuals may be allotted to each dose level. In such an experimental context, data may well cease to be 'well behaved', and indeed standard parametric procedures may fail. Examples of this have already been noted in Exercise 2.27. If there is either no, or

only one, dose which results in partial response (i.e. other than 0% or 100%) then maximum-likelihood fitting of standard models is either not possible (Silvapulle, 1981), or results in a fit with infinite slope parameter, β. The failure of standard procedures is normally an indication that further detailed investigation is necessary. In the context of routine testing, when only rough estimates are required, it can be a nuisance. In such a framework, more robust methods are needed. Robustness has already been encountered in Chapter 6, through the use of quasi-likelihood for describing over-dispersed data. In this chapter we start with a review of non-parametric procedures, with particular reference to estimation of the ED_{50}. We then proceed to consider robust parametric and non-parametric procedures, and describe their efficiency. The chapter concludes with a review of a number of alternative distribution-free procedures. As we shall see, some of the techniques which are proposed are very simple to operate, and of great practical importance. Others are of more academic interest, and unlikely to be used in a routine situation. The basis of much non-parametric work is a distribution-free estimate of the cumulative distribution function of the tolerance distribution, to be described in the next section.

7.2 The pool-adjacent-violators algorithm; ABERS estimate

Let $P(d)$ denote the probability of response to dose d, and let $P_i = P(d_i)$, for $1 \leqslant i \leqslant k$. If the only assumption made in the model is that $P(d)$ is a monotonic increasing function of d, then the maximum-likelihood estimate of $\mathbf{P} = (P_1, P_2, \ldots, P_k)$, corresponding to the log-likelihood,

$$l = \text{constant} + \sum_{i=1}^{k} r_i \log P_i + \sum_{i=1}^{k} (n_i - r_i) \log (1 - P_i)$$

results as an example of optimization under the order restriction:

$$P_1 \leqslant P_2 \leqslant \cdots \leqslant P_k$$

The necessary theory is to be found in Barlow *et al.* (1972, p 13). As it was originally propounded by Ayer *et al.* (1955), the resulting estimate is sometimes called the ABERS estimate. It is obtained by a straightforward algorithm called the pool-adjacent-violators algorithm, which proceeds as follows.

For any i we can form the maximum-likelihood estimator of P_i: $\hat{P}_i = r_i/n_i$. Naturally, in most cases we would like the $\hat{P}_i$ to be monotonically increasing as i increases. (An exception might arise if high-dose toxicity were anticipated.) If any two adjacent $\hat{P}_i$ are in the wrong order, i.e. $r_i/n_i > r_{i+1}/n_{i+1}$, then at the doses d_i and d_{i+1} we replace each of the responses $\hat{P}_i$ and $\hat{P}_{i+1}$ values by $(r_i + r_{i+1})/(n_i + n_{i+1})$. These two doses may now be considered as a **block**, and a number of such blocks may need to be formed. After the process of block formation is completed, it may be that adjacent block, or block and dose proportions may be found to be out of order. In such a case the relevant blocks and/or blocks and doses are pooled to give larger blocks, as appropriate, proceeding as in the original manner for proportions, until no further pooling is necessary. A simple illustration is given in Example 7.1.

Example 7.1 (Egger, 1979)

An illustration of the ABERS estimate, formed for a data set from Reed and Muench (1938). This is data set 3 of Exercise 1.3. The two non-monotonic proportions before adjustment are indicated by the arrows. Thus here there are just two blocks of doses, each of size two.

	Before adjustment			*After adjustment*		
Dose	*No. of subjects*	*No. responding*	*Proportion*	*No. of subjects*	*No. responding*	*Proportion*
1	6	0	0	6	0	0
2	6	0	0	6	0	0
4	6	1	↑0.1667	12	1	0.0833
8	6	0	↓0	12	1	0.0833
16	6	2	0.3333	6	2	0.3333
32	6	4	0.6667	6	4	0.6667
64	6	4	0.6667	6	4	0.6667
128	6	6	↑1.0	12	11	0.9167
256	6	5	↓0.8333	12	11	0.9167

In this case the adjustments are completed after just one stage. If additional adjustments are needed, then the order in which they are carried out is immaterial, as shown by Ayer *et al.* (1955) – see Exercise 7.1 Lexicographic confidence limits for the probabilities, illustrated for this data set, are given in Morris (1988). □

The ABERS estimate is a natural preliminary to computing the Spearman–Kärber and related estimates of the ED_{50}. It has also been found to be useful when employed with the method of minimum logit chi-square (Hamilton, 1979; James *et al.*, 1984). As we can see from the illustration of Example 7.1, by itself the ABERS estimate is not particularly useful, and some additional smoothing is usually considered advantageous. A simple approach is to use linear interpolation. This is employed by Hamilton (1979) as a preliminary to constructing a trimmed Spearman–Kärber estimate of the ED_{50}, and it is used also in section 7.5.3 in constructing an influence curve. Other ways of proceeding have been suggested by Schmoyer (1984) and Glasbey (1987), and we consider their work in section 7.6. A FORTRAN program for the pool-adjacent violators algorithm is given in Cran (1980).

In mathematical terms, if $\tilde{P}_1, \ldots, \tilde{P}_k$ denote the distribution-free maximum-likelihood estimators of $P(d_1), \ldots, P(d_k)$ under the order-restriction, $P_1 \leqslant P_2 \leqslant \cdots \leqslant P_k$, then (Exercise 7.2),

$$\tilde{P}_i = \max_{1 \leqslant u \leqslant i} \min_{i \leqslant v \leqslant k} \left\{ \sum_{j=u}^{v} r_j \Big/ \sum_{j=u}^{v} n_j \right\}$$

In the special case when $n_j = n$ for all j, then

$$\tilde{P}_i = \max_{1 \leqslant u \leqslant i} \min_{i \leqslant v \leqslant k} \frac{\sum_{j=u}^{v} r_j}{\{n(v-u+1)\}}$$

7.3 The Spearman–Kärber estimate of the ED_{50}

If the tolerance distribution has cumulative distribution function $F(x)$, then its mean may be written as:

$$\mu = \int_{-\infty}^{\infty} x\,dF(x)$$

Let

$$\hat{P}_i = \frac{r_i}{n_i}, \quad 1 \leqslant i \leqslant k$$

and

$$\Delta_i = (d_{i+1} - d_i), \quad 1 \leqslant i \leqslant k-1$$

then the Spearman–Kärber estimate for μ is given by:

$$\hat{\mu} = \hat{P}_1\left(d_1 - \frac{\Delta_1}{2}\right) + \sum_{i=1}^{k-1} (\hat{P}_{i+1} - \hat{P}_i)\left(d_i + \frac{\Delta_i}{2}\right) + (1 - \hat{P}_k)\left(d_k + \frac{\Delta_{k-1}}{2}\right) \tag{7.1}$$

There are two features to note in connection with this expression. First, the effect of expression (7.1) is equivalent to introducing additional artificial doses, d_0 and d_{k+1}, where $d_0 = d_1 - \Delta_1$, and $d_{k+1} = d_k + \Delta_{k-1}$, with respectively, 0% and 100% responses. If in fact $\hat{P}_1 = 0$ and $\hat{P}_k = 1$, then, as can be seen from equation (7.1) this device is not used. The arbitrariness of this procedure if either $\hat{P}_1 \neq 0$ or $\hat{P}_k \neq 1$ has understandably led to criticism of the Spearman–Kärber approach. It has, additionally, been argued that efficient design of an experiment aimed at estimating the ED_{50} should in fact result in $\hat{P}_1 \neq 0$ and $\hat{P}_k \neq 1$. However, as we have seen in the discussion above, in the observational study of Table 1.4, and in Exercises 2.27 and 2.30, data encompassing both 0% and 100% response are frequently encountered, especially in the case of routine assessment of potentially toxic or beneficial substances, when a wide range of doses is typically adopted. In any case, this feature of the Spearman–Kärber estimate is avoided if suitable trimming is employed, as will be discussed in section 7.4.

Secondly, if $P_1 = 0$, $P_k = 1$ and $P_i = F(x_i)$, $1 \leqslant i \leqslant k-1$, $\hat{\mu}$ estimates μ indirectly, through estimating the mean, η, of the discrete random variable X, with distribution,

$$\Pr(X = d_i + \Delta_i/2) = P_{i+1} - P_i, \quad 1 \leqslant i \leqslant k-1$$

The estimation of the probability $(P_{i+1} - P_i)$ by $(\hat{P}_{i+1} - \hat{P}_i)$ suggests that it may be useful to smooth the $\{\hat{P}_i\}$ before calculating the Spearman–Kärber estimate, to ensure that the $\{\hat{P}_i\}$ are monotonically increasing. The ABERS estimate may be used for this purpose if necessary. As we shall now see, this adjustment does not necessarily affect $\hat{\mu}$.

If we find that $\hat{P}_1 = 0$ and $\hat{P}_k = 1$, then equation (7.1) reduces to:

$$\hat{\mu} = \sum_{i=1}^{k-1} (\hat{P}_{i+1} - \hat{P}_i)\left(d_i + \frac{\Delta_i}{2}\right) \equiv \frac{1}{2} \sum_{i=1}^{k-1} (\hat{P}_{i+1} - \hat{P}_i)(d_i + d_{i+1})$$

which is a form in which the Spearman–Kärber estimate is often given. A simple rearrangement of terms gives the alternative

expression:

$$\hat{\mu} = \frac{1}{2}\left\{\sum_{j=2}^{k-1} \hat{P}_j(d_{j-1} - d_{j+1})\right\} + (d_{k-1} + d_k) \tag{7.2}$$

again making use of $\hat{P}_1 = 0$ and $\hat{P}_k = 1$. This form is convenient for evaluating the properties of $\hat{\mu}$. Further simplification results if the doses are equally spaced, i.e. $\Delta_i = \Delta$, for all i, leading to the expression:

$$\hat{\mu} = d_k + \frac{\Delta}{2} - \Delta \sum_{i=1}^{k} \hat{P}_i \tag{7.3}$$

If additionally the same number, n, of individuals is employed at each dose, then if the ABERS estimate is formed this does not change $\hat{\mu}$, since

$$\sum_{i=1}^{k} \hat{P}_i = \frac{1}{n}\sum_{i=1}^{k} r_i, \quad \text{and} \quad \sum_{i=1}^{k} r_i$$

does not change under the ABERS averaging procedure. This is true also in the slightly more general case of an equal weight design (Exercise 7.4).

Working directly from the expression of equation (7.2), assuming $P_1 = 0$, $P_k = 1$, and without reference to any possible ABERS correction, we can immediately write down the variance of $\hat{\mu}$ as:

$$V(\hat{\mu}) = \frac{1}{4}\sum_{j=2}^{k-1} P_j(1 - P_j)(d_{j-1} - d_{j+1})^2/n_j$$

which simplifies in the special case of equally spaced doses and equal numbers at each dose to give

$$V(\hat{\mu}) = \frac{\Delta^2}{n}\sum_{j=2}^{k-1} P_j(1 - P_j) \tag{7.4}$$

In general, for the case when $P_1 = 0$ and $P_k = 1$, we obtain an unbiased estimate of $V(\hat{\mu})$ if we evaluate

$$\frac{1}{4}\sum_{j=2}^{k-1} \hat{P}_j(1 - \hat{P}_j)(d_{j-1} - d_{j+1})^2/(n_j - 1)$$

as long as $n_j \geqslant 2$, for all j.

The explicit forms for $\hat{\mu}$ and its variance were clearly attractive when the iterations necessary for probit analysis, say, were onerous,

prior to the arrival of computers. It is still attractive to have a simple expression for $V(\hat{\mu})$ as an aid to the design of experiments (Brown (1966), and Chapter 8). As the expressions for $\hat{\mu}$ are weighted sums of independent binomial random variables, confidence regions for μ may be based on the assumption of normality of the estimator $\hat{\mu}$. We note, however, that $E[\hat{\mu}] = \eta$ and in general, $\eta \neq \mu$, since η is the mean of the discrete random variable X with distribution specified in terms of the particular dose levels of an experiment. Thus $\hat{\mu}$ is a biased estimator of μ. In fact, as $\hat{\mu}$ is a function of the sufficient statistic, $(r_1, \ldots, r_k)$, which is complete, then $\hat{\mu}$ is the uniform minimum variance unbiased estimator of η. Additionally, $\hat{\mu}$ is the unrestricted non-parametric maximum likelihood estimator of η, as it is a linear combination of $P_1, \ldots, P_k$, and the unrestricted maximum likelihood estimator of $(P_1, \ldots, P_k)$ is $(\hat{P}_1, \ldots, \hat{P}_k)$. Several authors have demonstrated the close affinity between the logit model and the Spearman–Kärber estimate, which we shall encounter again in section 7.5. See, for example, Finney (1950, 1952), and Anscombe (1956). Cornfield and Mantel (1950) showed the asymptotic equivalence of the logistic maximum likelihood estimator of the ED_{50} and the Spearman–Kärber estimate.

Of course the Spearman–Kärber estimate estimates the **mean** of the tolerance distribution, and if we use $\hat{\mu}$ to estimate the ED_{50}, which is the **median** of the tolerance distribution, then we are additionally making the assumption of symmetry. The same is true of the Reed–Meunch and Dragstedt–Behrens estimators (Miller, 1973 – see Exercises 7.6 and 7.7).

The logic which results in the Spearman–Kärber estimator applies equally well to estimation of higher order moments of the tolerance distribution. Thus, for example, the Spearman–Kärber estimate of the jth non-central moment estimator is:

$$\hat{\mu}_j = \hat{P}_1\left(d_1 - \frac{\Delta_1}{2}\right)^j + \sum_{i=1}^{k-1} (\hat{P}_{i+1} - \hat{P}_i)\left(d_i + \frac{\Delta_i}{2}\right)^j$$

$$+ (1 - \hat{P}_k)\left(d_k + \frac{\Delta_{k-1}}{2}\right)^j \quad \text{for } j \geqslant 1$$

An interesting use is made of these estimators by Egger (1979). She replaced d by its Box–Cox transformed value, $(d^\lambda - 1)/\lambda$, and sought

Table 7.1 *Asymptotic efficiency values for the estimators of μ and $\mu_{(2)}$, resulting from the Spearman–Kärber approach (Chmiel, 1976) when a variety of distributional forms are assumed for the tolerance distribution. The angular distribution has cumulative distribution function: $F(x) = \sin^2(x + \pi/4)$, for $-\pi/4 \leqslant x \leqslant \pi/4$*

Distribution	$\hat{\mu}$	$\hat{\mu}_{(2)}$
Logistic	1.0000	1.0000
Normal	0.9814	0.9637
Laplace	0.9618	0.9928
Angular	0.8106	0.6677

λ as a root of the equation,

$$\hat{\mu}_{(3)} = 0$$

where now $\hat{\mu}_{(3)}$ estimates the third central moment, and is given by:

$$\hat{\mu}_{(3)} = \hat{\mu}_3 - 3\hat{\mu}_2\hat{\mu}_1 + 2\hat{\mu}_1^3$$

The resulting value for λ could then be used to symmetrize the data in a non-parametric way, in comparison with the approach of section 4.4, based on maximum-likelihood. This procedure was not in fact found to be straightforward in practice, and is discussed further in Exercise 7.8. The properties of $\hat{\mu}_{(2)}$ have been investigated in detail by Chmiel (1976) who presented the asymptotic efficiency values shown in Table 7.1. In the quantal response context, care is needed in how one takes asymptotic limits, and we return to these results in our discussion in section 7.5. Additionally, Chmiel (1976) indicates how to use $\hat{\mu}_{(2)}$ to estimate percentiles.

Example 7.2

As an illustration of computation of the Spearman–Kärber estimate of the ED_{50} we may use the data from Exercise 1.14 on the detonation of explosives. Here, $d_i = 4$ for all i, and $n_i = 16$ for all i. Thus formula (7.3) applies and we obtain $\hat{\mu} = 46.5$, with an estimated standard error of 0.85. This contrasts with the corresponding maximum likelihood estimated values under the logit model of 46.64 (0.95).

7.4 Trimming

7.4.1 *Trimmed Spearman–Kärber*

If one's aim is to analyse routinely and automatically many sets of quantal response data, and in each case obtain an estimate of the ED_{50} and of its error, then the Spearman–Kärber method may be made more robust by the addition of trimming, which is performed by analogy with the use of trimming to estimate the mean in direct samples from univariate distributions (Andrews *et al.* 1972, p. 31). The procedure for quantal response data was propounded by Hamilton *et al.* (1977) and has subsequently been evaluated by Hamilton (1979, 1980), Miller and Halpern (1980), James *et al.* (1984), Sanathanan *et al.* (1987) and Hoekstra (1989). We shall now state how the procedure operates, and present illustrations of the results, leaving efficiency considerations until section 7.5.

The first stage is to form the ABERS estimate, if necessary. What follows next is best illustrated first of all by example.

Example 7.3

Calculation of a 10% trimmed Spearman–Kärber estimate: we shall use the data of Example 7.1.

We start by forming the ABERS estimate $\{\tilde{P}_i\}$ from Example 7.1. The points on the graph of $\tilde{P}_i$ versus $x_i = \log_2(d_i)$ are joined by straight lines to give the polygon of Figure 7.1(a). The points marked ⊗ are then trimmed away, and the points marked • are added. The axis for proportions is then rescaled to give $\tilde{\tilde{P}}_i = (\tilde{P}_i - 0.1)/0.8$, with the result shown in Figure 7.1(b), and finally the Spearman–Kärber estimate is formed for the $\{\tilde{\tilde{P}}_i\}$. For the case of this example we obtain the 10% trimmed Spearman–Kärber estimate, which we denote by $\hat{\mu}_{10}$, as $\hat{\mu}_{10} = 27.7$. An estimated 95% confidence interval, obtained as shown in Figure 7.1, is (15.37, 49.93).

In mathematical terms, the $100\alpha\%$ trimmed Spearman–Kärber estimate is defined as follows (the notation is illustrated in Figure 7.1).

For $0 \leqslant \alpha \leqslant 0.5$, let

$$L(\alpha) = \max\{i : \hat{P}_i \leqslant \alpha, \quad 1 \leqslant i \leqslant k\} \text{ and}$$
$$U(\alpha) = \min\{i : \hat{P}_i \geqslant (1 - \alpha), \quad 1 \leqslant i \leqslant k\}$$