Isotonic Inference

C. Hirotsu
in

Encyclopedia of Biostatistics
(ISBN 0471 975761)

Edited by

Peter Armitage and Theodore Colton

Isotonic Inference

Isotonic inference concerns situations in which a set of parameters is assumed, a priori, to satisfy certain order restrictions. In the most common case, where data are arranged in ordered groups, the **mean** value of a **random variable** is assumed to change monotonically with the ordering of the groups. It is then reasonable to take account of the order restrictions in making inferences about the group means, such as point or interval estimations or significance tests. Isotonic inference extends more generally to situations where there are various shape constraints on response curves, such as convexity, concavity, or sigmoidicity.

One approach to such problems is to assume a parametric model that incorporates those order or shape constraints such as a **linear regression** equation or a particular **dose–response** function. The inference based on the parametric model can, however, be considerably **biased** and variable when the specified model is incorrect. It has been pointed out in environmental toxicology applications, for example, that no parametric dose–response model can be assumed to hold generally at very low doses of interest, and yet a monotone and convex relationship might reasonably, and more reliably, be assumed. We are therefore concerned in this article mainly with methods of inference that avoid the need to

specify a rigid parametric model, but nevertheless allow for those order restrictions.

There is a large literature on **estimation** and testing (*see* **Hypothesis Testing**) in the areas of isotonic and order-restricted inferences, and comprehensive surveys of these areas include [3] and [37].

One general approach to the isotonic inference is **maximum likelihood** estimation. The problem of finding order-restricted maximum likelihood estimates is often solved by using **isotonic regression**. In its simplest case an explicit solution is obtained by the pool-adjacent-violators method, but in more general cases it is solved only by some nonlinear programming, see [40] and [9], for example, or by the aid of a formal **Bayesian** approach, as in [36]. For a restricted **likelihood ratio test** the usual asymptotic **chi-square distribution** theory does not apply. In some cases, the resulting distributions are known to be a mixture of χ^2 distributions, but in other cases some **computer-intensive methods** such as parametric **bootstrap** tests [9], or an asymptotic conservative approximation method [40] may be used. The maximum likelihood approach is outlined in another article (*see* **Isotonic Regression**). Here we are concerned mainly with other approaches to isotonic inference. As a natural method of incorporating prior knowledge in particular applications, a Bayesian approach is also briefly mentioned.

Table 1 Half life of an antibiotic in rats

Dose (mg/kg)	Data (h)					Average
5	1.17	1.12	1.07	0.98	1.04	1.076
10	1.00	1.21	1.24	1.14	1.34	1.186
25	1.55	1.63	1.49	1.53		1.550
50	1.21	1.63	1.37	1.50	1.81	1.504
200	1.78	1.93	1.80	2.07	1.70	1.856

The Case for Isotonic Inference

The data in Table 1 are measurements of the half-life of an antibiotic drug in relation to the dose administered. The usual **analysis of variance** (ANOVA) is obviously inappropriate, because of the ordering of the doses, and one possible approach is to assume a parametric model. The simplest model for the monotone relationship is linear regression. However, it is generally difficult to assume that a linear relationship holds over a wide range of an **explanatory variable**. For the dose–response relationship there are of course more natural response curves, such as a sigmoid function, but it is still often difficult to assume a particular model for the given set of data. Furthermore, it also sometimes suffices to show an overall upward trend or to detect a steep **change-point** in the responses. It is then unnecessary to assume a rigid parametric model, and a nonparametric trend test or some **multiple comparisons** procedure is more appropriate

(*see* **Simultaneous Inference**). We need assume only a monotone relationship in the mean half-life,

$$H_1: \mu_1 \leq \cdots \leq \mu_a, \tag{1}$$

where at least one inequality is strong, so that the null model H_0: $\mu_1 = \cdots = \mu_a$ is excluded.

The data in Table 2 show ordinal (i.e. **ordered**) **categorical data** typical of a Phase III comparative **clinical trial**. Assuming a **multinomial** model with cell probabilities p_{ij}, the **null hypothesis** that the two treatments are equal can be expressed as $p_{1j} = p_{2j}$, $j = 1, \ldots, 4$, or equivalently as

$$p_{ij} = p_{i\cdot} p_{\cdot j}. \tag{2}$$

Eq. (2) is the familiar independence hypothesis for a two-way **contingency table**. However, the usual **goodness-of-fit chi-square test** is inappropriate, since we are interested in a more restricted alternative

$$p_{11}/p_{21} \leq \cdots \leq p_{14}/p_{24}$$

Table 2 Efficacy in a phase III trial of antibiotics

Drug	Not effective	Slightly effective	Effective	Excellent
AMPC	3	8	30	22
S6472	8	9	29	11

$$H_2 : \text{or}$$

$$p_{11}/p_{21} \geq \cdots \geq p_{14}/p_{24},$$

where at least one inequality is strong, implying that treatment 1 is superior to treatment 2 in efficacy or vice versa. **Ordered categorical data** are a special case of rank data with many ties, and any method for rank data can be applied to ordered categorical data, and vice versa.

If ordered categorical data are obtained at several doses, as in Table 3, then we are interested in testing the two-way **ordered alternative**:

$$H_3 : p_{i+1,j}/p_{i,j} \leq p_{i+1,j+1}/p_{i,j+1},$$

$$i = 1, \ldots, a-1;$$

$$j = 1, \ldots, b-1,$$

which implies that higher doses are superior to lower doses in efficacy.

Table 3 Usefulness in a dose-finding clinical trial

Drug	Undesirable	Slightly undesirable	Not useful	Slightly useful	Useful	Excellent
Placebo	3	6	37	9	15	1
AF 3 (mg/kg)	7	5	33	21	10	1
AF 6 (mg/kg)	5	6	21	16	23	6

A similar hypothesis

$$H_4 : \mu_{i+1,j+1} - \mu_{i+1,j} - \mu_{i,j+1} + \mu_{i,j} \geq 0,$$

$$i = 1, \ldots, a-1;\ j = 1, \ldots, b-1,$$

has been considered for normal means from a two-way layout experiment, which implies that the differences, $\mu_{ij} - \mu_{i'j}$, tend upwards as the level j increases for any $i > i'$; see [13].

Various Extensions of the Monotone Relationship

A monotone dose–response relationship may be disturbed by toxicity at higher doses, and a **nonparametric** testing procedure for the downturn (or "umbrella") hypothesis,

$$H_5 : \mu_1 \leq \cdots \leq \mu_{\tau+1} \geq \mu_{\tau+2} \geq \cdots \geq \mu_a,$$

$$\tau = 1, \ldots, a-1,$$

has been proposed in [44]; here τ is an unknown turning point.

Some other extensions arise when responses show monotone relationships with the passage of time. Frequently encountered examples include the monotonic change of occurrence probabilities of some events, increasing treatment effects, and increasing **hazard rates** with time. For instance, the hypothesis

$$H_6 : \mu_2 - \mu_1 \leq \mu_3 - \mu_2 \leq \cdots \leq \mu_a - \mu_{a-1}$$

arises from the analysis of the age–period–cohort effects model (*see* **Age–Period–Cohort Analysis**) where only the second-order differences are estimable in each effect along with the time axis; see [19]. Hypothesis H_6 is equivalent to $H_6' : \mu_i - 2\mu_{i+1} + \mu_{i+2} \geq 0$, and may be called the "convexity hypothesis". Hypothesis H_6 is of course mathematically different from hypothesis H_5, but there are some similarities in their shapes, and with a slight modification the test for H_6 performs well also as a test for H_5. Convexity, concavity and sigmoidicity constraints are commonly employed also in the field of bioassay as reasonable shape constraints on a dose–response relationship (*see* **Biological Assay, Overview**; **Quantal Response Models**).

As seen from the above examples, isotonic inference is closely related to **change-point** analysis. Actually, a one-sided change-point model may be formulated as a set of particular monotone relationships,

$$H_7 : \mu_1 = \cdots = \mu_\tau < \mu_{\tau+1} = \cdots = \mu_a,$$

$$\tau = 1, \ldots, a-1, \tag{3}$$

with τ an unknown change-point parameter, so that a useful statistic for change-point analysis is useful also for isotonic inference. Interestingly, (3) defines $a - 1$ edges of the convex cone defined by the simple

ordered alternative (1).

For other extensions, including **tree-structured**, star-shaped, unimodality, and symmetry models, the reader is referred to [3] and [37].

Testing a Simple Ordered Alternative in Normal Means

We wish to test a simple ordered alternative H_1 in the one-way layout model

$$y_{ij} = \mu_i + \varepsilon_{ij}, \quad i = 1, \ldots, a; \; j = 1, \ldots, n_i,$$

where the ε_{ij} are assumed to be independently distributed as **normal** $N(0, \sigma^2)$ with known **variance** σ^2. Then there are two major streams of overall trend tests and multiple contrast type tests. Most cases of unknown variance can be dealt with similarly, if an **unbiased** variance estimator distributed as a multiple of χ^2 is available.

Overall Trend Tests

One possible approach is the restricted likelihood ratio test developed extensively in [3] (*see* **Isotonic Regression**). The approach does not, however, possess any obvious optimal property for such restricted alternatives, and is rather difficult to extend to higher-way problems.

Abelson & Tukey [1] proposed a linear score statistic which maximizes the minimum **power** in the region defined by H_1 within the class of linear tests. This has been extended to the most stringent and somewhere most powerful (MSSP) test for a more general restricted

alternative by Schaafsma [38, 39]. In the balanced case, Abelson & Tukey's score is determined by equalizing powers at all the $a-1$ edges of H_1 and is given by

$$c_i \propto -[i(1-i/a)]^{1/2} + \left[(i-1)\left(1-\frac{i-1}{a}\right)\right]^{1/2},$$
$$i = 1, \ldots, a.$$

Extending Taguchi's idea [46], the cumulative χ^2 test was introduced in [14], and its power has been compared with that of the previous two approaches. The test statistic χ^{*2} is the sum of squares of the standardized accumulated statistics

$$y_i^* = \frac{1}{\sigma}\left(\frac{1}{N_i}+\frac{1}{N_i^*}\right)^{-1/2}(\overline{Y}_i^* - \overline{Y}_i),$$
$$i = 1, \ldots, a-1, \tag{4}$$

where $N_i = n_1 + \cdots + n_i$, $N_i^* = n_{i+1} + \cdots + n_a$, and $\overline{Y}_i = (y_{1\cdot} + \cdots + y_{i\cdot})/N_i$, $\overline{Y}_i^* = (y_{i+1\cdot} + \cdots + y_{a\cdot})/N_i^*$ with $y_{i\cdot} = (y_{i1} + \cdots + y_{in_i})$, $i = 1, \ldots, a$. The χ^{*2} statistic is characterized by the strong positive **correlations** between the serial components y_i^*, and in particular by the expansion for the balanced case in a series of independent χ^2 variables,

$$\chi^{*2} = \frac{1}{1\cdot 2}\chi_{(1)}^2 + \frac{a}{2\cdot 3}\chi_{(2)}^2 + \cdots + \frac{a}{(a-1)\cdot a}\chi_{(a-1)}^2,$$

where $\chi_{(l)}^2$ is the 1 df χ^2 statistic for detecting the departure from the null model in the direction of Chebyshev's lth order orthogonal

polynomial. Hence χ^{*2} tests mainly, but not exclusively, a linear trend; see [18] and [31] for details.

Multiple Contrast Type Tests

Several multiple comparison procedures have been proposed for ordered parameters. Williams [49] proposed a closed testing procedure based on the maximum likelihood estimator for defining the maximal noneffective dose level. Marcus [26] modified the method by changing the estimator at the control level from $\overline{y}_1$. to $\hat{\mu}_1$, the maximum likelihood estimator of μ_1, so that his statistic is the maximal component of Bartholomew's $\overline{\chi}^2$. The limiting distribution of the latter statistic is obtained in [50], where upper percentiles are tabulated, including the case of unknown variance. The maximal component of χ^{*2} has been proposed also for this purpose, and is called the "max t" method, where t stands for the y_i^* of (4). The statistic is characterized as the likelihood ratio test for the change-point hypothesis H_7, and an exact and very efficient algorithm for calculating the ***P* value** has been obtained by Hawkins [11] (*see* **Change-Point Problem**). The power functions of these closed multiple testing procedures have been compared in [26], [43], and [25]. More general multiple tests for ordered parameters are obtained in [28].

Confidence Interval

A **confidence interval** taking advantage of order restrictions can be

obtained by inverting an appropriate test for order restricted alternatives. For example, Marcus & Peritz [27] and Schoenfeld [41] obtain confidence intervals for normal means by inverting a multiple contrast type test and the restricted likelihood ratio test, respectively. Wynn [53] gives a general methodology for obtaining one-sided confidence intervals, and Hayter [12] obtains confidence intervals based on the one-sided **studentized range** test. In particular, in the bioassay problem, Schmoyer [40] obtains improved upper confidence bounds for the responses at very low doses by assuming sigmoidicity in the dose–response curve.

There is no extensive work on the design of experiments on the ordered parameters, although optimal allocation has been discussed in [24] (*see* **Optimal Design**).

Applications

A test of Abelson & Tukey, the cumulative χ^2 test, and some of the multiple comparison procedures, are now applied to the data in Table 1. Since the variance σ^2 is unknown, it is replaced by the usual unbiased estimate of variance, $\hat{\sigma}^2 = \sum\sum(y_{ij} - \overline{y}_{i\cdot})^2/(24-5) = 0.020741$.

The linear score statistic of Abelson & Tukey is calculated as

$$(-c_1\overline{y}_{1\cdot} - c_2\overline{y}_{2\cdot} + c_3\overline{y}_{3\cdot} + c_4\overline{y}_{4\cdot} + c_5\overline{y}_{5\cdot})/\hat{\sigma}$$

$$= 9.1206,$$

with scores $c_1 = c_5 = (\sqrt{6}+1)/\sqrt{5} = 1.543$, $c_2 = c_4 = (4-\sqrt{6})/\sqrt{20} = 0.3467$, and $c_3 = 0$, giving a P value of 2.2×10^{-8} as evaluated by the t distribution with 19 df.

The null distribution of cumulative χ^2 statistic $\sum y_i^{*2}$ is well approximated by $d\chi_f^2$, a multiple of the χ^2 variable with df f, where the constants d and f are given by

$$d = 1 + \frac{2}{a-1} \times \left(\frac{\lambda_1}{\lambda_2} + \frac{\lambda_1+\lambda_2}{\lambda_3} + \cdots + \frac{\lambda_1+\cdots+\lambda_{a-2}}{\lambda_{a-1}} \right),$$

$$f = (a-1)/d, \tag{5}$$

with $\lambda_i = N_i/N_i^*$. An even better approximation based on the expansions by Laguerres' orthogonal polynomials, and also the approximation under the alternative hypothesis, are given in [15]. Then the P value of the statistic

$$F^* = (a-1)^{-1} \chi^{*2}|_{\sigma^2=\hat{\sigma}^2} = 54.739$$

can be evaluated as 1.1×10^{-8} by the ***F* distribution** with df $(f, \sum n_i - a)$, where $f = 2.067$ from (5).

The maximal component of the $\chi^{*2}|_{\sigma^2=\hat{\sigma}^2}$ is obtained at the partition between levels 2 and 3:

$$\max t = \left[\left(\frac{1}{10} + \frac{1}{14} \right) (0.02741) \right]^{-1/2}$$

$$\times \left(\frac{23.00}{14} - \frac{11.31}{10} \right) = 8.584,$$

the one-sided P value of which is evaluated as 1.1×10^{-7} by the recurrence formula based on the Markov property of $y_i^* s$. According to the closed testing procedure of [28], the process proceeds to the final step where the t statistic between levels 1 and 2 shows a nonsignificant result at the one-sided significance level 0.10, thus suggesting finally the difference between the dose levels (1,2) and (3,4,5).

In applying the Williams [49] procedure and the modified Williams procedure of [26], we need the maximum likelihood estimaters of the μ_i, which are

$$\hat{\mu}_1 = 1.076, \qquad \hat{\mu}_2 = 1.186,$$

$$\hat{\mu}_3 = \hat{\mu}_4 = 1.524, \qquad \hat{\mu}_5 = 1.856,$$

by the pool-adjacent-violators method. Since $\hat{\mu}_1 = \overline{y}_{1.}$, both statistics coincide and equal

$$w = \max \sqrt{m}(\hat{\mu}_i - \overline{y}_{1.})/\hat{\sigma}$$

$$= \sqrt{m}(\overline{y}_{5.} - \overline{y}_{1.})/\hat{\sigma} = 8.357,$$

where we take the repetition number m as the harmonic mean of the n_is for referring approximately to the tables for upper percentiles in the balanced case by [49] and [50], respectively. In any case, the statistic w is highly significant and the closed testing procedure stops with the

nonsignificant result between levels 1 and 2, thus again suggesting a difference between the dose levels (1, 2) and (3, 4, 5).

For a more general likelihood $L(\mathbf{y}, \boldsymbol{\theta}, \boldsymbol{\nu})$ with the ordered parameter $\boldsymbol{\theta}$, and possibly with the **nuisance parameter** $\boldsymbol{\nu}$, arguments similar to those used above apply if the asymptotic normality of the likelihood estimators is assured. In particular, the cumulative χ^2 and the max t statistics can be based on the cumulative efficient scores evaluated at the null hypothesis and extended easily to two-way problems; see [16], [17], and [7] for details.

Testing Ordered Alternatives in Binomial Probabilities

The data in Table 4 are from a dose–response **clinical trial**. Assuming that the y_i are independently distributed as **binomial** $B_{\text{in}}(n_i, p_i)$ we are interested in testing the simple **ordered alternative**

$$\mathrm{H}_1\colon\ p_1 \leq \cdots \leq p_a.$$

Table 4 Dose finding trial for a heart disease drug

Dose (mg/day)	Improved	Not improved
100	20	16
150	23	18
200	27	9
225	26	9
300	9	5

If the quantitative measures $d_1 < \cdots < d_a$ are attached to the y_i, then the locally most powerful test against a wide range of monotone relationships of p_i to d_i is obtained by Cochran [6] and Armitage [2] (*see* **Trend Test for Counts and Proportions**). For the case where there is no information on d_i, the likelihood ratio test has been developed by Chacko [5]. The tests based on the cumulative χ^2 and its maximal component have also been extended as follows:

$$\begin{aligned} &\text{the cumulative } \chi^2 : \chi^{*2} = \sum y_i^{*2}, \\ &\text{the maximal component of } \chi^{*2} : \max t = \max y_i^*, \end{aligned} \tag{6}$$

where y_i^* is given by (4) with σ replaced by $[\overline{Y}(1-\overline{Y})]^{1/2}$, $\overline{Y} = \sum y_i / \sum n_i$, and $y_{i\cdot}$ replaced by y_i in defining Y_i. Formula (5) is also valid for the χ^{*2} to give a two-sided P value of 0.113 when applied to Table 4. For max t, another exact algorithm is available based on the Markov property of the y_i^* to give a one-sided P value of 0.044 for Table 2 at the partition between levels (1, 2) and (3, 4, 5); see [51], [52], and [25] for the algorithm. The Cochran–Armitage test gives a slightly larger one-sided P value of 0.049 since there is a slight downturn tendency in this example.

Analyzing the Two-Way Contingency Table with Ordered Column Categories

Two-Sample Problem

First consider the two-sample problem presented by Table 2. A popular approach to the analysis is to use a nonparametric test based on a linear score statistic such as Wilcoxon's (*see* **Wilcoxon–Mann–Whitney Test**). Now, for the two-sided alternative H_2 the two statistics,

$$\text{the cumulative } \chi^2 : \chi^{*2} = \chi_1^2 + \cdots + \chi_{b-1}^{*2}, \tag{7}$$

the maximal component of χ^{*2} :

$$\max \chi^2 = \max \chi_j^2, \tag{8}$$

can be defined in terms of the accumulated efficient scores, where χ_j^2 is the goodness-of-fit χ^2 statistic for the ***2 x 2* table** formed by accumulating the first j and the remaining $b - j$ columns. The χ_j^2 is, however, identical to the y_i^{*2} of (6) if the binomial data are arranged in a $2 \times b$ table in an obvious way and exactly the same distribution theory applies also to this case. The two-sided P values are 0.039 for χ^{*2} and 0.154 for max χ^2, whereas it is 0.025 for the Wilcoxon test. In this case the Wilcoxon test shows the smallest P value, since approximately a linear trend is observed in p_{1j}/p_{2j}, $j = 1, \ldots, 4$. If these tests are applied to the last two rows of Table 3 for comparing AF 3 mg and AF 6 mg, then the two-sided P values are 0.0128, 0.0096, and

0.0033 for the Wilcoxon, the χ^{*2}, and max χ^2 methods, respectively. It has been verified by **simulation** that, when evaluated as two-sample nonparametric tests, the Wilcoxon method is useful for the location shift of the underlying symmetrical and light-tailed distributions such as the **logistic** or normal, the max χ^2 method is useful for skewed or heavy-tailed distributions, and the cumulative χ^2 method is characterized by its **robustness**, having relatively high **power** over a wide range of underlying distributions – normal, heavy-tailed, or skewed.

Another important approach to the problem is to assume an underlying continuous distribution for each treatment and to compare the parameters describing those distributions. The **proportional-odds** and **proportional-hazards** models are important examples; see [30] for details.

General a-*Sample Problem*

For a general a-sample problem the Wilcoxon test is extended to the Kruskal–Wallis test. The same type extensions are available for the χ^{*2} and its maximal component by defining the χ_j^2 in (7) and (8) as the goodness-of-fit χ^2 statistic for the accumulated $a \times 2$ table for the partition between columns j and $j+1$. The constants for the χ^2 approximation of χ^{*2} are obtained by

$$d = 1 + \frac{2}{b-1}$$

$$\times \left(\frac{\gamma_1}{\gamma_2} + \frac{\gamma_1 + \gamma_2}{\gamma_3} + \cdots + \frac{\gamma_1 + \cdots + \gamma_{b-2}}{\gamma_{b-1}} \right),$$

$$f = (a-1)(b-1)/d,$$

with $\gamma_j = C_j / C_j^*$, $C_j = y_{\cdot 1} + \cdots + y_{\cdot j}$, and $C_j^* = y_{\cdot j+1} + \cdots + y_{\cdot b}$. The max χ^2 can be evaluated by the calculation **algorithm** based on the Markov property of the subsequent χ_j^2s [25].

For the row-wise multiple comparisons based on the cumulative χ^2, the statistic

$$S = \max ||(\mathbf{a}' \otimes \mathbf{C}^{*'})\mathbf{z}||^2$$

is defined where $\otimes$ is a Kronecker product, $\mathbf{z}$ a vector of $\sqrt{y_{\cdot\cdot} y_{ij} / (y_{i\cdot} y_{\cdot j})}^{1/2}$ arranged in dictionary order, $\mathbf{C}^{*'}$ a $b-1 \times b$ matrix defined so that the (j, j') th element of $\mathbf{C}^{*'}\mathbf{C}^*$ is $(\gamma_j / \gamma_{j'})^{1/2}$ for $j \leq j'$ and the maximum is taken over all $\mathbf{a}$ that satisfy $\mathbf{a}'\mathbf{a} = 1$, and $(\sqrt{y_{1\cdot}}, \ldots, \sqrt{y_{a\cdot}})\,\mathbf{a} = 0$. When $a \geq b$ and under the null model, the statistic S is asymptotically distributed as the largest root of the Wishart matrix $W(\mathbf{C}^{*'}\mathbf{C}^*, a-1)$, which is well approximated by $\gamma_{(1)}\chi^2(a-1)$ with $\gamma_{(1)}$ the largest root of $\mathbf{C}^{*'}\mathbf{C}^*$. The statistic S gives the Scheffé-type multiple comparison test, and has been applied to taste-testing data of five foods in five ordered categorical responses of [4] to obtain the significant classification of rows (foods) (1, 2), (3, 4) and (5). The max χ^2 is also applied to the data for multiple comparisons of the columns

to obtain a highly significant classification (1, 2, 3) and (4, 5). The resulting block interaction model is expressed as

$$p_{ij} = p_{i\cdot}\, p_{\cdot j} q_{\mu\nu}, \quad \mu = 1, 2, 3, \nu = 1, 2,$$

if i belongs to the μth subgroup of rows and j to the νth subgroup of columns. The goodness-of-fit χ^2 has been compared with the fitting of the proportional-odds model [45] and its extension [29]; see [20] and [22] for details. The Scheffé-type multiple comparison method is applied to the normal distribution model in [21], for classifying subjects based on the similarity of the time series profiles defined by repeated measurements.

Two-Way Contingency Table with Natural Orderings in Both Rows and Columns

Assuming a multinomial model $M(y_{\cdot\cdot}, p_{ij})$ for the data y_{ij} in Table 3 the cumulative χ^2 statistic and its maximal component are defined for testing H_3 using the cumulative efficient scores evaluated at the null hypothesis. These are

$$\text{the doubly cumulative } \chi^2 : \chi^{**2} = \sum\sum \chi_{ij}^2,$$

$$\text{the maximal component of } \chi^{**2} : \max\max \chi_{ij}^2,$$

with the χ_{ij}^2 being the goodness-of-fit χ^2 for the 2×2 tables obtained from partitioning and accumulating rows and columns at $i = 1, \ldots, a -$

1, and $j = 1, \ldots, b-1$, respectively. The χ^{**2} is for the two-sided version of H_3, and max max χ^2 is applicable to both one- and two-sided problems. When applied to Table 3 the two-sided P values are approximately 0.0065 for the χ^{**2} and exactly 0.0142 for max max χ^2. The details of the P value calculations and other variations of the test statistics are given in [22] and [23]. As a semiparametric model for the ordered two-way table the constant-**odds ratio** model has been proposed by Wahrendolf [48] based on Plackett's [32] coefficient of association for **bivariate distributions** (*see* **Association, Measures of**).

As an example of the higher-way layouts, a three-way contingency table with age at four levels, existence of the metastasis into the lymph node at two levels, and the soaking grade at three levels, is analyzed in [22]. An example of highly fractional factorial experiments with ordered categorical responses is given in [10]; see also the discussion following that article (*see* **Factorial Experiments**).

Bayesian Approach to Isotonic Inference

Since the purpose of an isotonic inference is to make use of the prior knowledge to enhance the efficiency of test and estimation, it is natural to consider a Bayesian approach. For example, an essentially complete class of tests for orderly constrained hypothesis is

obtained as the whole set of Bayes tests with a **prior distribution** defined on those constrained supports. The cumulative χ^2 and max t methods are derived from this idea; see [47], [16], and [7]. More specifically, in bioassay problems, the Dirichlet prior has been introduced for the successive differences of the responses for doses d_i, $p(d_i) - p(d_{i-1}), i = 1, \ldots, a+1$; $p(d_0) = 0$, $p(d_{a+1}) = 1$, reflecting the nondecreasing nature of the dose–response relationship, see [34] and [35], for example. Shaked & Singpurwalla [42] discuss the defect of the Dirichlet prior, and introduce concavity constraints on the shape of a dose–response curve, reflecting a situation encountered in practice. Because of computational difficulties, however, they are unable to compute posteriors beyond modal estimates. The computational problem was overcome only recently by Gelfand & Kuo [8], who showed how a sampling-based approach could be used to develop the desired marginal posterior distributions and their features, for Dirichlet and product–beta priors. Ramgopal et al. [33] consider convex, concave, and ogive constraints to specify the shape of dose–response curves, and extend the sampling-based approach to calculating any posterior feature of interest in these generalized constrained problems.

References

[1] Abelson, R.P. & Tukey, J.W. (1963). Efficient utilization of non-numerical information in quantitative analysis: general theory and the case of the simple order, *Annals of Mathematical Statistics* **34**, 1347–1369.

[2] Armitage, P. (1955). Test for linear trends in proportions and frequencies, *Biometrics* **11**, 375–386.

[3] Barlow, R.E., Bartholomew, D.J., Bremner, J.M. & Brunk, H.D. (1972). *Statistical Inference Under Order Restrictions*. Wiley, Chichester.

[4] Bradley, R.A., Katti, S.K. & Coons, I.J. (1962). Optimal scaling for ordered categories, *Psychometrika* **27**, 355–374.

[5] Chacko, Y.C. (1966). Modified chi-square test for ordered alternatives, *Sankhyā, Series B* **28**, 185–190.

[6] Cochran, W.G. (1954). Some methods for strengthening the common χ^2 test, *Biometrics* **10**, 417–451.

[7] Cohen, A. & Sackrowitz, H.B. (1991). Tests for independence in contingency tables with ordered categories, *Journal of Multivariate Analysis* **36**, 56–67.

[8] Gelfand A.E. & Kuo, L. (1991). Nonparametric Bayesian bioassay including ordered polytomous response, *Biometrika* **78**, 657–666.

[9] Geyer, C.J. (1991). Constrained maximum likelihood exemplified by isotonic convex logistic regression, *Journal of the American Statistical Association* **86**, 717–724.

[10] Hamada, M. & Wu, C.F.J. (1990). A critical look at accumulation analysis and related methods (with discussion), *Technometrics* **32**, 119–130.

[11] Hawkins, D.M. (1977). Testing a sequence of observations for a shift in location, *Journal of the American Statistical Association* **72**, 180–186.

[12] Hayter, A.J. (1990). A one-sided Studentized range test for testing against a simple ordered alternative, *Journal of the American Statistical Association* **85**, 778–785.

[13] Hirotsu, C. (1978). Ordered alternatives for interaction effects, *Biometrika* **65**, 561–570.

[14] Hirotsu, C. (1979). The cumulative chi-squares method and Studentized maximal contrast method for testing an ordered alternative in a one-way analysis of variance model, *Reports of Statistical Application Research, Union of Japanese Scientists and Engineers* **26**, 12–21.

[15] Hirotsu, C. (1979). An F-approximation and its application, *Biometrika* **66**, 577–584.

[16] Hirotsu, C. (1982). Use of cumulative efficient scores for testing ordered alternatives in discrete models, *Biometrika* **69**, 567–577.

[17] Hirotsu, C. (1983). Defining the pattern of association in two-way contingency tables, *Biometrika* **70**, 579–589.

[18] Hirotsu, C. (1986). Cumulative chi-squared statistic as a tool for testing goodness of fit, *Biometrika* **73**, 165–173.

[19] Hirotsu, C. (1988). A class of estimable contrasts in an age-period-cohort model, *Annals of Institute of Statistical Mathematics* **40**, 451–465.

[20] Hirotsu, C. (1990). Discussion on Hamada and Wu's paper, *Technometrics* **32**, 133–136.

[21] Hirotsu, C. (1991). An approach to comparing treatments based on repeated measures, *Biometrika* **78**, 583–594.

[22] Hirotsu, C. (1992). *Analysis of Experimental Data, Beyond Analysis of Variance* (in Japanese). Kyoritsu-shuppan, Tokyo.

[23] Hirotsu, C. (1997). Two-way change-point model and its application, *Australian Journal of Statistics* **39**, 205–218.

[24] Hirotsu, C. & Herzberg, A.M. (1987). Optimal allocation of observations for inference on k ordered normal population means, *Australian Journal of Statistics* **29**, 151–165.

[25] Hirotsu, C., Kuriki, S. & Hayter, A.J. (1992). Multiple comparison procedure based on the maximal component of the cumulative chi-squared statistic, *Biometrika* **79**, 381–392.

[26] Marcus, R. (1976). The powers of some tests of the equality of normal means against an ordered alternative, *Biometrika* **63**, 177–183.

[27] Marcus, R. & Peritz, E. (1976). Some simultaneous confidence bounds in normal models with restricted alternatives, *Journal of the Royal Statistical Society, Series B* **38**, 157–165.

[28] Marcus, R., Peritz, E. & Gabriel, K.R. (1976). On closed testing procedure with special reference to ordered analysis of variance, *Biometrika* **63**, 655–660.

[29] McCullagh, P. (1980). Regression models for ordinal data, *Journal of the Royal Statistical Society, Series B* **42**, 109–142.

[30] McCullagh, P. & Nelder, J.A. (1989). *Generalized Linear Models*, 2nd Ed. Chapman & Hall, London.

[31] Nair, V.N. (1986). On testing against ordered alternatives in analysis of variance models, *Biometrika* **73**, 493–499.

[32] Plackett, R.L. (1965). A class of bivariate distributions, *Journal of the American Statistical Association* **60**, 516–522.

[33] Ramgopal, P., Laud, P.W. & Smith, A.F.M. (1993). Nonparametric Bayesian bioassay with prior constraints on the shape of the potency curve, *Biometrika* **80**, 489–498.

[34] Ramsey, F.L. (1972). A Bayesian approach to bioassay, *Biometrics* **28**, 841–858.

[35] Ramsey, F.L. (1973). Correction, *Biometrics* **29**, 830.

[36] Robert, C.P. & Hwang, J.T.G. (1996). Maximum likelihood estimation under order restrictions by the prior feedback method, *Journal of the American Statistical Association* **91**, 167–172.

[37] Robertson, T., Wright, F.T. & Dykstra, R.L. (1988). *Order Restricted Statistical Inference*. Wiley, New York.

[38] Schaafsma, W. (1966). Hypothesis testing problems with the alternative restricted by a number of inequalities, *Doctoral dissertation.* University of Groningen, Noordhoff, Groningen.

[39] Schaafsma, W. (1968). A comparison of the most stringent and the most stringent somewhere most powerful tests for certain problems with restricted alternatives, *Annals of Mathematical Statistics* **39**, 531–546.

[40] Schmoyer, R.L. (1984). Sigmoidally constrained maximum likelihood estimation in quantal bioassay, *Journal of the American Statistical Association* **79**, 448–453.

[41] Schoenfeld, D.A. (1986). Confidence bounds for normal means under order restrictions, with application to dose-response curves, toxicology experiments, and low-dose extrapolation, *Journal of the American Statistical Association* **81**, 186–195.

[42] Shaked M. & Singpurwalla, N.D. (1990). A Bayesian approach for quantile and response probability estimation with applications to reliability, *Annals of the Institute of Statistical Mathematics* **42**, 1–19.

[43] Shirley, E. (1979). The comparison of treatment with control group means in toxicological studies, *Applied Statistics* **28**, 144–151.

[44] Simpson, D.G. & Margolin, B.H. (1986). Recursive nonparametric testing for dose response relationships subject to downturn at high doses, *Biometrika* **73**, 589–596.

[45] Snell, E.J. (1964). A scaling procedure for ordered categorical data, *Biometrics* **20**, 592–607.

[46] Taguchi, G. (1966). *Statistical Analysis* (in Japanese). Maruzen, Tokyo.

[47] Takeuchi, K. (1979). Test and estimation problems under restricted null and alternative hypotheses (in Japanese), *Journal of Economics* **45**, 2–10.

[48] Wahrendolf, J. (1980). Inference in contingency tables with ordered categories using Plackett's coefficient of association for bivariate distributions, *Biometrika* **76**, 15–21.

[49] Willams, D.A. (1971). A test for differences between treatment means when several dose levels are compared with a zero dose control, *Biometrics* **27**, 103–117.

[50] Williams, D.A. (1977). Some inference procedures for monotonically ordered normal means, *Biometrika* **64**, 9–14.

[51] Worsley, K.J. (1983). The power of likelihood ratio and cumulative sum tests of a change in a bionomial probability, *Biometrika* **70**, 455–464.

[52] Worsley, K.J. (1986). Confidence regions and tests for a change point in a sequence of exponential family of random variables, *Biometrika* **73**, 91–104.

[53] Wynn, H.P. (1975). Integrals for one-sided confidence bounds: A general result, *Biometrika* **62**, 393–396.