Electronic Journal of Statistics
Vol. 3 (2009) 912–931
ISSN: 1935-7524
DOI: 10.1214/09-EJS392

Multiple testing in ordinal data models

Chuanwen Chen*, Arthur Cohen† and Harold B. Sackrowitz†

Department of Statistics,
Rutgers University, Piscataway, NJ 08854
e-mail: chuanwen@stat.rutgers.edu; artcohen@rci.rutgers.edu
sackrowi@rci.rutgers.edu

Abstract: Consider an $R \times C$ contingency table in which the categories are ordered. Multiple testing of the hypotheses that each local odds ratio is one is carried out. The methodology to perform the multiple tests is an extension of the MRDSS method of Chen, Cohen, and Sackrowitz (2009). The MRDSS method extends the MRD method of Cohen, Sackrowitz, and Xu (2009) by adding a screen stage and a sign stage to MRD. The MRDSS method as well as the extension here is admissible and consistent. Both Fisher-type statistics and Chi-square statistics are used. Examples and a simulation study are included.

AMS 2000 subject classifications: Primary 62H17; secondary 62C15, 62F03.
Keywords and phrases: Admissibility, chi-square test, consistency, contingency table, Fisher's exact test, full multinomial model, independent Poisson model, ordered categories.

Received March 2009.

Contents

1. Introduction

$R \times C$ contingency tables with ordered categories are studied. In particular we study multiple testing of the $(R-1) \times (C-1)$ local log odds ratios. That is, the null hypotheses are that these log odds ratios are zero against the alternative that each is not zero. One sided alternatives are also studied. For two sided alternatives we show that oftentimes the usual step-up and step-down

*This represents part of the Ph.D thesis of C. Chen.
†Supported by NSF Grant DMS 0804547 and NSA GRANT H-98230-06-0076.

procedures based on P-values derived from Fisher's exact test statistics or from Pearson's Chi-square statistics for local 2×2 subtables, have an undesirable property. Namely, that relevant acceptance sections of individual hypotheses are not always convex. This renders these multiple testing procedures (MTPs) inadmissible in terms of expected number of type I errors and expected number of type II errors.

New admissible and consistent MTPs are offered. The new procedures represent an extension of the MRDSS method developed in the companion paper by Chen, Cohen, and Sackrowitz (2009). Precise definitions of admissibility and consistency are given in that reference. MRDSS is a three stage multiple testing method. The first stage called maximum residual down was developed by Cohen, Sackrowitz, and Xu (2009). It is a step-down method with several desirable practical and theoretical properties. However, MRD is not always consistent in the sense that asymptotically (as sample size tends to infinity) it can sometimes make mistakes. A screen stage added to MRD plus a sign stage added to MRD yields MRDSS which is both admissible and consistent.

The MRD stage of MRDSS is based on statistics called residuals. These statistics arise naturally in multivariate normal models and in some exponential family models. The residuals have certain monotonicity properties that are essential for admissibility of individual tests. In the contingency table models analogues of the residuals, with the desirable properties, are found. The new, so called residuals, can be determined as Fisher statistics or Chi-square statistics for certain 2×2 tables obtained by collapsing $R \times C$ tables and certain subtables. Sometimes this involves segmenting the overall table into smaller tables that subsequently are collapsed into 2×2 tables. Systems of equations need to be derived and solved to provide estimated expected frequencies and estimated variances of frequencies for the Chi-square statistics. This all must be done so that the analogues of the residuals have the necessary monotonicity properties enjoyed by the residuals of MRD. Precise definitions are forthcoming.

The screen stage and sign stage of MRDSS also must be developed for the $R \times C$ model. The new screen stage must depend on statistics for local 2×2 tables that yield consistency. For MRDSS applied in normal models the choice of statistic for the screen stage is obvious. For the $R \times C$ case there are several choices of statistics. The screen stage statistic is formed for local 2×2 tables. The sign stage for the $R \times C$ table is not literally a comparison of signs. This time "signs" match or not, depending whether a rejection by MRD is for a left or right side rejection as compared to whether the screen stage statistic would reject on the right side or left side. Again precise definitions are forthcoming.

Multiple testing of odd ratios in 2×2 subtables of $R \times C$ tables is a problem discussed in Gabriel (1966), Hirotsu (1978, 1983) and Hochberg and Tamhane (1987). The procedures recommended are all single step and are very conservative. Shaffer (1986) recommends stepwise procedures for testing 2×2 subtables. We regard testing all local odds ratios as a meaningful problem for $R \times C$ tables with both factors having ordered categories. Such testing enables detections of change points in directions of each factor, enables examination of likelihood ratio order (all log odds non-negative one-sided testing) between pairs of rows or

pairs of columns, and helps identify the adjacent levels of each factor where a true association exists.

Textbook discussions on stepwise methods for multiple testing in general (not necessarily for categorical data) can be found in Hochberg and Tamhane (1987), Lehmann and Romano (2005) and Dudoit and Van Der Laan (2008). Another useful reference for these methods is Dudoit, Shaffer, and Boldrick (2003).Cohen and Sackrowitz (2005b, 2005a, 2007, 2008) and Cohen, Kolassa, and Sackrowitz (2007) demonstrate the inadmissibility of the usual step-up and step-down procedures for a wide variety of models, loss functions and conditions for some one-sided and some two-sided hypotheses. Cohen, Sackrowitz, and Xu (2009) and Chen, Cohen, and Sackrowitz (2009) offer new methodologies in models concerned with means of normal variables.

In the next section we pose several discrete models for an $R \times C$ contingency table and give some other preliminaries. In Section 3 we cite examples for the $2 \times C$ case in which step-up or step-down procedures based on Fisher exact test statistics or based on Pearson's chi-squared statistics are inadmissible. We give examples since discreteness and other conditions prevent a general result. Nevertheless the examples are such that it will be clear that inadmissibility will often be the case. Examples of inadmissibility of typical procedures for the $R \times C$ case will be similar to the $2 \times C$ case.

New methodology, first for $2 \times C$ tables will be given in Section 4. An example will be given. New methodology will be given in Section 5 for the $R \times C$ case. Pearson's chi-squared statistics will be used for the $R \times C$ case since Fisher's statistic becomes less computationally feasible. An example is given in Section 6 for the $R \times C$ case and a simulation study is offered in Section 7.

2. Models and preliminaries

Consider an $R \times C$ contingency table with ordered categories in both rows and columns. Let X_{ij} be the frequency in the ijth cell, $i = 1, \ldots, R$, $j = 1, \ldots, C$. Assume either the full multinomial model or the product multinomial model with cell probabilities p_{ij}. In either case, if we condition on both the column totals t_j, $j = 1, \ldots, C$ and the row totals $R_i, i = 1, \ldots, R$, the conditional distribution of X_{ij}, $i = 1, \ldots, R-1$, $j = 1, \ldots, C-1$ given all R_i, t_j is multivariate hypergeometric. Let $n = \sum_{i=1}^{R} R_i = \sum_{j=1}^{C} t_j$.

The conditional distribution expressed in exponential family form is

$$f_{\mathbf{X}}(\boldsymbol{x}|\boldsymbol{\nu}) \propto \exp\left(\sum_{i=1}^{R-1}\sum_{j=1}^{C-1} x_{ij}\nu_{ij}\right) \tag{2.1}$$

where ν_{ij} are log odds ratios, i.e., $\nu_{ij} = \log\frac{p_{ij}p_{RC}}{p_{iC}p_{Rj}}$ with $\boldsymbol{x} = (\boldsymbol{x}_1, \ldots, \boldsymbol{x}_R)'$, $\boldsymbol{x}_i = (x_{i1}, \ldots, x_{iC})'$, $\boldsymbol{\nu} = (\boldsymbol{\nu}_1, \ldots, \boldsymbol{\nu}_R)'$. See Cohen and Sackrowitz (2000). If we let $S_{ij} = \sum_{k=1}^{i}\sum_{l=1}^{j} X_{kl}$ and let $\mu_{ij} = \log[p_{ij}p_{(i+1)(j+1)}/p_{i(j+1)}p_{(i+1)j}]$ then

with $\boldsymbol{S} = (S_{11}, \ldots, S_{RC})'$

$$f_{\mathbf{S}}(\mathbf{s}|\boldsymbol{\mu}) \propto \exp\left(\sum_{i=1}^{R-1}\sum_{j=1}^{C-1} \mu_{ij} S_{ij}\right). \tag{2.2}$$

Note (2.2) would also ensue for the model in which X_{ij} are independent Poisson variables with parameter λ_{ij}. Then upon conditioning on R_i, $\boldsymbol{X}_i$ is multinomial with parameter $p_{ij} = \lambda_{ij}/\sum_{j=1}^{C}\lambda_{ij}$. Conditioning next on t_j yields the multivariate hypergeometric of (2.1).

Hypotheses of interest for the $R \times C$ table are

$$H_{ij} : \mu_{ij} = 0 \quad \text{vs} \quad K_{ij} : \mu_{ij} \neq 0 \tag{2.3}$$

or,

$$H_{ij} : \mu_{ij} = 0 \quad \text{vs} \quad K_{ij}^{*} : \mu_{ij} > 0 \tag{2.4}$$

A local 2×2 table of frequencies is

x_{ij}	$x_{i(j+1)}$
$x_{(i+1)j}$	$x_{(i+1)(j+1)}$

(2.5)

and let T_{ij} be a statistic to test H_{ij}. Typical step-down and step-up procedures are based on statistics T_{ij} that depend only on the cell frequencies in (2.5). Such statistics are often Fisher's exact test statistics or Pearson's chi-square statistics.

To describe a class of step-down procedures based on P-values let $0 < \alpha_1 < \cdots < \alpha_q$, where $q = (R-1)(C-1)$, be a sequence of critical values. Let $P_{ij}(T_{ij})$ be the P-value for testing H_{ij} and let $P_{(1)} \leq \cdots \leq P_{(q)}$ correspond to the ordered P-values.

(i) If $P_{(1)} < \alpha_1$, reject $H_{(1)}$ and continue to step 2. Otherwise stop and accept $H_{(i)}$, $i = 1, \ldots, q$.

(ii) If $H_{(1)}$ is rejected, reject $H_{(2)}$ if $P_{(2)} < \alpha_2$. Otherwise stop and accept $H_{(2)}, \ldots, H_{(q)}$.

(iii) In general, at step m, if $P_{(m)} < \alpha_m$, reject $H_{(m)}$. Otherwise stop and accept $H_{(m)}, \ldots, H_{(q)}$.

The critical values α_i can be chosen in variety of ways.

Often times α_1 is chosen at step 1 to control the weak familywise error rate (FWER). That is, the probability of at least one false rejection when all nulls are true. The other α's are often chosen to control strong FWER or the false discovery rate (FDR). See for example, Dudoit, Shaffer, and Boldrick (2003).

To describe a class of step-up procedures again consider a sequence of critical values $0 < \alpha_1 < \cdots < \alpha_q$, not necessarily the same as in the step-down case.

(i) If $P_{(q)} < \alpha_q$, stop and reject all $H_{(i)}$. If $P_{(q)} \geq \alpha_q$ accept $H_{(q)}$ and go to step (ii).

(ii) If $P_{(q-1)} < \alpha_{q-1}$, stop and reject $H_{(1)}, \ldots, H_{(q-1)}$. Otherwise accept $H_{(q-1)}$ and continue.

(iii) In general, at step m, if $P_{(q-m+1)} < \alpha_{q-m+1}$, stop and reject $H_{(1)}, \ldots, H_{(q-m+1)}$. Otherwise accept $H_{(q-m+1)}$ and continue.

The critical constants are often chosen to control FDR (or FWER). See for example, Benjamini and Hochberg (1995).

We evaluate the collection of q tests by evaluating each individual test by its expected type I error and expected type II error. If $\mathbf{T}$ represents the collection of test statistics $T_{ij}(\boldsymbol{x})$ and $\psi_{ij}(\boldsymbol{x})$ represents the test function based on T_{ij} for testing H_{ij} then the risk for the individual test is

$$E_{\boldsymbol{\mu}}\psi_{ij}(\boldsymbol{x}) \text{ when } \mu_{ij} = 0 \text{ and } E_{\boldsymbol{\mu}}(1 - \psi_{ij}(\boldsymbol{x})) \text{ when } \mu_{ij} \neq 0 \tag{2.6}$$

One can consider a q-vector risk function consisting of q components, each representing the risk for an individual test. In terms of admissibility, multiple testing procedures that are inadmissible for this vector risk would remain so if the risk is any non-decreasing function of the collection of individual components of the vector risk.

3. $2 \times C$ tables — examples of non-convex acceptance sections

Poisson, multinomial, or independent binomial models

Our starting point in this section is to assume the model where upon conditioning on row and column totals we have the multivariate hypergeometric distribution given in (2.2). We start with a result for $R \times C$ tables.

Lemma 3.1. *A necessary and sufficient condition for a test $\psi_{ij}(\boldsymbol{x})$ to be admissible is that for all $S_{i'j'}(i' \neq i, j' \neq j)$, except S_{ij}, fixed, the acceptance sections of the test are convex in S_{ij}.*

Proof. See Matthes and Truax (1967). □

Note that convex in the discrete case is defined for sample points whose coordinates take on integer values. A discrete set B is convex if whenever $\mathbf{s}_1 \in B$, $\mathbf{s}_2 \in B$, all sample points between $\mathbf{s}_1$ and $\mathbf{s}_2$ are in B. Note that when $R = 2$, $\boldsymbol{S}$ depends only on $S_{11}, \ldots, S_{1(C-1)}$ and these depend only on $X_{11}, \ldots, X_{1(C-1)}$.

Furthermore when $R = 2$ an increase in S_{1j} by one unit while $S_{1j'}$ remain fixed $j' \neq j$ is accomplished with $(X_{11}, \ldots, X_{1j}, X_{1(j+1)}, \ldots, X_{1(C-1)})' + \boldsymbol{g}$ where $\boldsymbol{g} = (0, \ldots, 1, -1, \ldots, 0)'$ where 1 is in position j.

Whereas one cannot demonstrate that step-down and step-up procedures based on Fisher exact tests or Pearson's chi-square tests are always inadmissible, we give examples (which are more typical than not) to indicate that those procedures often are inadmissible. It sufficies to work with 2×3 tables since extensions to $2 \times C$ tables would easily follow.

Example 3.1. Consider a 2×3 table as follows

			109
			100
90	60	59	

Test $H_{1j} : \mu_{1j} = 0$ vs $K_{1j} : \mu_{1j} \neq 0$, $j = 1, 2$. For step-down using Fisher's P-value (with FWER controlled at $\alpha = 0.1$). The critical values are $\alpha_1 = 0.05$, $\alpha_2 = 0.1$. Now suppose the frequencies are:

$$\begin{array}{ccc} 59 & 31 & 19 \\ 31 & 29 & 40 \end{array}$$

Then the P-values are 0.093, 0.041 which leads to rejecting both H_{11} and H_{12}. However for frequencies

$$\begin{array}{ccc} 60 & 30 & 19 \\ 30 & 30 & 40 \end{array}$$

the P-values are 0.061, 0.063 which means both hypotheses are accepted. For a first row of 55 35 19, H_{11} would be accepted which demonstrates that the acceptance section for H_{11} would not be an interval in S_{11}:

S_{11}	55	59	60
Decision	A	R	A

For the two-sided case using the same critical values of α, using Pearson's chi-squared statistic the following tables of frequencies also provide an example where the acceptance section is not an interval. Note in the following pairs of tables, the first table leads to reject and the second table to accept. It is understood that a third extreme table exists which would lead to rejection.

65	51	38	
39	49	66	P-values are 0.0973, 0.0373
66	50	38	
38	50	66	P-values are 0.0523, 0.0523

Similarly, using the same critical value, for one-sided Fisher, $\alpha = .1$

65	52	26	
40	48	44	P-values are 0.0983, 0.0392
66	51	26	
39	49	44	P-values are 0.0577, 0.0513

For chi-square one-sided, we find that

65	51	27	
41	47	43	P-values are 0.0906, 0.0427
66	50	27	
40	48	43	P-values are 0.0526, 0.0552

For step-up, Fisher 2-sided with $\alpha_1 = 0.05$, $\alpha_2 = 0.1$, consider

59	31	13	
31	29	26	P-values are 0.0931, 0.0980
60	30	13	
30	30	26	P-values are 0.0608, 0.1461

For step-up Fisher 1-sided

65	52	33	
40	48	47	P-values are 0.0983, 0.0992
66	51	33	
39	49	47	P-values are 0.0577, 0.1245

Examples for step-up and chi-square are easily obtained. Examples for $R \times C$ tables are also easily obtained.

4. $2 \times C$ tables. New methodology

Our starting point for this section is the case $R = 2$ and we observe $\boldsymbol{S}$ with distribution given in (2.2).

The methods offered are related to the MRDSS method of Chen, Cohen, and Sackrowitz (2009) which extends the MRD method of Cohen, Sackrowitz, and Xu (2009). The new method also features a 3-stage process in the spirit of the original MRDSS. Although the statistics used at each stage are quite different we will still refer to the new method as MRDSS. Here it involves using Fisher's test statistics (or Pearson's chi-square statistics) for various sometimes collapsed tables that wind up as 2×2 tables. We give the method using Fisher's statistics P_{1j} although Pearson's chi-square statistics could also be used.

Recall there are $(C-1)$ local log odds ratios to be tested; namely $H_{1j} : \mu_{1j} = 0$ vs $K_{1j} : \mu_{1j} \neq 0$, $j = 1, \ldots, C-1$. Note when $R = 2$, the approach makes sense if only the columns are ordered. The main method involves 3 stages. Stage 1 involves a step-down testing method; stage 2 involves screening and stage 3 involves a possible change due to sign differences. The method is as follows: Let $0 < \alpha_1 < \cdots < \alpha_{C-1}$.

At stage 1,

(i) consider the $(C-1)$ 2×2 tables

S_{1j}	$R_1 - S_{1j}$
$\sum_{k=1}^{j} t_k - S_{1j}$	$R_2 - \sum_{k=1}^{j} t_k + S_{1j}$

(4.1)

Compute $P_{1j}^{(1)}$ for Table j, $j = 1, \ldots, C-1$ and find $P_{(1)}^{(1)}$, the smallest among $P_{1j}^{(1)}$. If $P_{(1)}^{(1)} = P_{1j_1}^{(1)} < \alpha_1$, reject $H_{(1)}$ and continue. Otherwise stop and accept H_{1j}, $j = 1, \ldots, C-1$.

(ii) If $j_1 = 1$ consider the $2 \times (C-2)$ table with column 1 left out. Proceed as in step (i). That is, form (C-2) 2×2 tables

$S_{1j} - S_{11}$	$R_1 - S_{1j}$
$\sum_{k=2}^{j} t_k - S_{1j} + S_{11}$	$R_2 - \sum_{k=1}^{j} t_k + S_{1j} - S_{11}$

(4.2)

Compute $P_{1j}^{(2)}$, $j = 2, \ldots, C-1$ and find $P_{1j_2}^{(2)} = \min_{2 \le j \le C-1} P_{1j}^{(2)}$. If $P_{1j_2}^{(2)} < \alpha_2$, reject H_{1j_2} and continue. Otherwise stop and accept $H_{(2)}, \ldots, H_{(C-1)}$.

If $1 < j_1 < C-1$, segment the $2 \times C$ table into 2 tables. The first table is the $2 \times j_1$ table consisting of the first j_1 columns of the original table and the second table is $2 \times (C - j_1)$ consisting the last $(C - j_1)$ columns of the table. Now treat each table as in the $2 \times C$ table case and form 2×2 tables as in step (i). That is, form $(j_1 - 1)$ P-values for the $2 \times j_1$ table and $C - j_1 - 1$ P-values for the $2 \times (C - j_1)$ table and then get $P^{(2)}_{1j_2}$ as the minimum P-value and reject H_{1j_2} if $P^{(2)}_{1j_2} < \alpha_2$ and continue. Otherwise stop and accept $H_{(2)}, \ldots, H_{(C-1)}$.

If $j_1 = C - 1$ then do essentially what was done when $j_1 = 1$.

(iii) If the process continues beyond step (ii), then continue for each table at step (ii) by further segmentation if necessary at all future steps. Compute P-values as in previous stages and get the min P-value and compare it with the appropriate critical value, i.e., at step m, if $P^{(m)}_{1j_m} < \alpha_m$, reject H_{1j_m} and continue. Otherwise stop and accept $H_{(m)}, \ldots, H_{(C-1)}$.

At stage 2, screen the results of stage 1, as follows: Let $\alpha_L < \alpha_U$ be two critical values between 0 and 1. Typically $\alpha_L < \alpha_{C-1} \leq \alpha_U$. Find Fisher's statistic for each local 2×2 table. If the Fisher statistic is less than α_L and at stage 1 the test accepted, switch to rejection. If the statistic exceeds α_U and at stage 1 the test rejected, switch to accept. Otherwise maintain the results of stage 1.

At stage 3, switch a reject at stage 1 to a final accept at stage 3, if when screening, the statistic lies between α_L and α_U and if the Fisher statistic used in stages 1 and 2 were compiled using opposite sides of their respective 2 sided rejection regions. For example, suppose we are testing $H_{11} : \mu_{11} = 0$ vs $K_{11} : \mu_{11} \neq 0$. Suppose that at stage 1, step 1 the table

S_{11}	$S_{iC} - S_{11}$
$t_1 - S_{11}$	$t_2 - S_{1C} + S_{11}$

yielded the minimum P-value $< \alpha_1$ and thus H_{11} was rejected at stage 1. Note that (1 - P-value) based on Fisher's statistic for this table as a function of S_{11} is decreasing and then increasing. See Cohen (1987). Suppose further that the observed value of S_{11} was on the increasing part of the function. Now for the table

x_{11}	x_{12}
x_{21}	x_{22}

assume the Fisher test statistic P-value was between α_L and α_U. Further suppose for the observed value of x_{11} that (1 - P-value) as a function of x_{11} was on the decreasing part of the function. This would call for a switch of reject H_{11} to accept H_{11}.

Remark 4.1. The critical values $0 < \alpha_1 < \cdots < \alpha_{C-1}$ in stage 1 and $\alpha_L \leq \alpha_U$ in stage 2 can be chosen according to some considerations. They can be chosen so that the overall procedure satisfies an error rate control such as FWER or FDR. They can be chosen so that FDR is controlled only when the numbers of non-nulls is large. This can sometimes be accomplished by simulation. The degree of conservativeness, total sample size, namely S_{RC} and the total number of hypotheses to be tested influence the choice of constants.

TABLE 4.1
Change in Size of Ulcer Crater

Treatment	Larger	$< 2/3$ Healed	$\geq 2/3$ Healed	Healed	Total
A	12	10	4	6	32
B	5	8	8	11	32
Total	17	18	12	17	64

At stage 2 α_L and α_U would be chosen so that changes are made only if the evidence is compelling. In other words α_L is usually small and α_U is relatively large. Sometimes critical values are obtained via simulation. In this case we seek FDR control for all situations except perhaps when the number of true alternations is small.

We now illustrate the method for $2 \times C$ tables. The data are from Agresti (1984a) and are provided in Table 4.1.

Stage 1- step 1. Form 3 2×2 tables where the entries in the first row-first column are S_{11}, S_{12} and S_{13} respectively. These 3 tables and their corresponding P-values based on Fisher statistics are:

	12 20	22 10	26 6
	5 27	23 19	21 11
$P-values$	.0879	.0437	.2474

The minimum P-value among the 3 tables occurs for the second table. Using critical values (.05, .075, .1) we reject $H_{12} : \mu_{12} = 0$.

Stage 1- step 2. Since $\mu_{12} = 0$ was rejected the 2×4 table is segmented into two 2×2 tables. Namely

$$\begin{array}{cc} 12 & 10 \\ 5 & 8 \end{array} \quad \text{and} \quad \begin{array}{cc} 4 & 6 \\ 8 & 11 \end{array}$$

The P-values for these are .4887 and 1 respectively. Since P-values are less than .075, both H_{11} and H_{13} are not rejcted and Stage 1 is complete.

Stage 2 (screen stage): The 3 local 2×2 tables are considered. These include the two considered at Stage 1 - step 2 and the third is

$$\begin{array}{cc} 10 & 4 \\ 8 & 8 \end{array}$$

with a P-value of .2839. If $\alpha_L = .01$ and $\alpha_U = .3$ no change is made at this stage.

Stage 3 (sign stage): The only rejection was for H_{12}. Both at stage 1 and stage 2 the statistics were on the same side (right hand side) of rejection regions so again no change occurs at the sign stage.

Thus there is evidence that the ratio of the proportion of larger craters to less than 2/3 healed craters is different for the two treatments.

Note if in stage 2 $\alpha_U = .2$, the decision to reject H_{12} at stage 1 would be reversed and all 3 null hypotheses would have been accepted.

5. $R \times C$ tables

Recall that in an $R \times C$ table we wish to test $(R-1)(C-1)$ individual hypotheses $H_{ij} : \mu_{ij} = 0$ vs $K_{ij} : \mu_{ij} \neq 0$, $i = 1, \ldots, R-1$; $j = 1, \ldots, C-1$. Let $\Omega_1 > \Omega_2 > \cdots > \Omega_{(R-1)(C-1)} > 0$ be a set of critical values. We will do as follows at stage 1, MRD stage:

Step 1: Form $(R-1)(C-1)$ 2×2 tables:

S_{ij}	$S_{iC} - S_{ij}$
$S_{Rj} - S_{ij}$	$S_{RC} - S_{iC} - S_{Rj} + S_{ij}$

Find the statistics:

$$U_{ij}^{(1)} = (S_{ij} - \hat{E}_{ij}^{(1)}) / \sqrt{\hat{V}_{ij}^{(1)}} \tag{5.1}$$

where $\hat{E}_{ij}^{(1)}$ and $\hat{V}_{ij}^{(1)}$, along with $\hat{E}_{ij}^{(m)}$ and $\hat{V}_{ij}^{(m)}$ for step m, $m = 1, 2, \ldots, (R-1)(C-1)$ will be determined below. $U_{ij}^{(1)}$ will be used to test hypothesis H_{ij}. If $|U_{i_1 j_1}^{(1)}| = \max_{(i,j)} |U_{ij}^{(1)}| > \Omega_1$, reject $H_{i_1 j_1}$, record the sign of $U_{i_1 j_1}^{(1)}$ and continue. Note that $(U_{ij}^{(1)})^2$ is equivalent to Pearson's chi-square statistic for the above 2×2 table.

At step 2, find the statistics

$$U_{ij}^{(2)} = (S_{ij} - \hat{E}_{ij}^{(2)}) / \sqrt{\hat{V}_{ij}^{(2)}} \tag{5.2}$$

If $|U_{i_2 j_2}^{(2)}| = \max_{(i,j)} |U_{ij}^{(2)}| > \Omega_2$, reject $H_{i_2 j_2}$, record the sign of $U_{i_2 j_2}^{(2)}$ and continue.

The quantities $\hat{E}_{ij}^{(1)}$, $\hat{V}_{ij}^{(1)}$ will be estimated under two conditions:

(I) all null hypotheses are assumed true
(II) Row sums, column sums are fixed

The quantities $\hat{E}_{ij}^{(2)}$, $\hat{V}_{ij}^{(2)}$ will be estimated under two conditions:

(I) all null hypotheses except $H_{i_1 j_1}$ are assumed true
(II) Row sums, column sums, and $S_{i_1 j_1}$ are fixed

In general, at step m, suppose $H_{i_1 j_1}, \ldots, H_{i_{m-1} j_{m-1}}$ have been rejected. Then $\hat{E}_{ij}^{(m)}$ and $\hat{V}_{ij}^{(m)}$ will be estimated under two conditions:

(I) all null hypotheses except $H_{i_1 j_1}, \ldots, H_{i_{m-1} j_{m-1}}$ are assumed true
(II) Row sums, column sums, and $S_{i_1 j_1}, \ldots, S_{i_{m-1} j_{m-1}}$ are fixed

Now we indicate exactly how $\hat{E}_{ij}^{(m)}$ and $\hat{V}_{ij}^{(m)}$ are estimated.

At step 1 let λ_{ij} denote the mean of cell(i, j). Then estimate λ_{ij} by $\hat{\lambda}_{ij}$, where $\hat{\lambda}_{ij}$ is the solution to the following set of equations

$$\sum_{k,l} \lambda_{kl} = n \tag{5.3}$$

$$\sum_{k=1}^{R} \lambda_{kj} = t_j \qquad \text{for } j = 1, \ldots, C-1 \tag{5.4}$$

$$\sum_{l=1}^{C} \lambda_{il} = R_i \qquad \text{for } i = 1, \ldots, R-1 \tag{5.5}$$

$$\lambda_{ij}\lambda_{(i+1)(j+1)} = \lambda_{(i+1)j}\lambda_{i(j+1)} \qquad \text{for } 1 \le i \le R-1, 1 \le j \le C-1 \tag{5.6}$$

There are a total of $1 + (C-1) + (R-1) + (R-1)(C-1) = RC$ equations in RC variables. The unique non-negative solution is $\lambda_{ij} = R_i t_j / n$.

At step 2, suppose $H_{i_1 j_1}$ is rejected by the end of step 1. Estimate the means of cells again by using the same $R \times C$ equations, except now replace the equation

$$\lambda_{i_1 j_1}\lambda_{(i_1+1)(j_1+1)} = \lambda_{(i_1+1)j_1}\lambda_{i_1(j_1+1)} \tag{5.7}$$

with equation

$$\sum_{k \le i_1, l \le j_1} \lambda_{kl} = S_{i_1 j_1} \tag{5.8}$$

In general at step m, we remove $(m-1)$ equations of the type (5.6) and add $(m-1)$ equations of the type (5.8).

To get the estimate $\hat{V}_{ij}^{(m)}$ we think of the cell frequencies x_{ij} as independent normal variables with mean λ_{ij} and variance λ_{ij}. Recall $S_{ij} = \sum_{l=1}^{i}\sum_{k=1}^{j} x_{lk}$, so $\boldsymbol{S} = A\boldsymbol{X}$ for the appropriate A, It would follow that $\boldsymbol{S} \sim N(A\boldsymbol{\lambda}, A\Sigma A')$ with $\Sigma = diag(\lambda_{11}, \ldots, \lambda_{RC})$ being the covariance matrix of $\boldsymbol{X}$. Having estimated λ_{ij} by $\hat{\lambda}_{ij}$ we thus have an estimator of $\Sigma_{\mathbf{S}} = A\Sigma A'$. However to find $\hat{V}_{ij}^{(m)}$, we compute the conditional covariance matrix of $\boldsymbol{S}$ under conditions:

(i) S_{iC}, $i = 1, \ldots, R-1$ are known
(ii) S_{Rj}, $j = 1, \ldots, C-1$ are known
(iii) $S_{i_1 j_1}, \ldots, S_{i_{m-1} j_{m-1}}$ are known.

The quantity $V_{ij}^{(m)}$ represents the conditional variance of S_{ij}. $V_{ij}^{(m)}$ is a function of λ_{ij} and is estimated accordingly. The computation of $\hat{\lambda}_{ij}$, $\hat{E}_{ij}^{(m)}$ and $\hat{V}_{ij}^{(m)}$ is done as follows:

First, we solve the equations for estimating λ_{ij}. This is done by transforming the equation-solving problem into a minimization problem.

At step m, let

$$f_m(\boldsymbol{\lambda}) = \left(\sum_{k,l} \lambda_{kl} - n\right)^2 + \sum_{j=1}^{C-1}\left(\sum_{k=1}^{R} \lambda_{kj} - t_j\right)^2 + \sum_{i=1}^{R-1}\left(\sum_{l=1}^{C} \lambda_{il} - R_i\right)^2 + \sum_{i=1}^{R-1}\sum_{j=1}^{C-1} d_{ij}^2 \tag{5.9}$$

where

$$d_{ij} = \begin{cases} \lambda_{ij}\lambda_{(i+1)(j+1)} - \lambda_{(i+1)j}\lambda_{i(j+1)} & \text{if } H_{ij} \text{ is not rejected before step m} \\ \sum_{k\le i_1, l\le j_1} \lambda_{kl} - S_{ij} & \text{otherwise} \end{cases}$$

It is clear that $f_m(\boldsymbol{\lambda}) \ge 0$ for all $\boldsymbol{\lambda}$.

Let $\hat{\boldsymbol{\lambda}}$ be the unique non-negative solution of equations (5.3)–(5.8). This implies $f_m(\hat{\boldsymbol{\lambda}}) = 0$. Hence $\hat{\boldsymbol{\lambda}}$ is the point that minimizes f_m. So by minimizing f_m, we can get the solution of equations (5.3)–(5.8). It is obvious that f_m is a convex function for $\boldsymbol{\lambda} \ge \mathbf{0}$, hence the minimization of f_m can be done very easily and efficiently. In our study, we use the *nlm* function in the popular **R**(2009) software to minimize f_m.

After we solve the equations, i.e, we get $\hat{\boldsymbol{\lambda}}$, we compute the conditional variance of S_{ij}. This enables us to complete the computation of $U_{ij}^{(m)}$.

After the MRD stage, we proceed to stage 2, the screen stage. This time we compute the statistics for each local 2×2 table and find the P-values for these statistics, comparing them to α_L and α_U. Then we proceed to stage 3, the sign stage. Here we proceed as done in the $2 \times C$ case.

We conclude this section with:

Theorem 5.1. *For an $R \times C$ contingency table, assuming either the product multinomial model, full multinomial model or independent Poisson model, the MRDSS method using Pearson's Chi-square statistic or Fisher's statistic in the $2 \times C$ case is admissible.*

Proof. Chen, Cohen, and Sackrowitz (2009) prove that for some exponential family models adding a screen and sign stage to an MRD procedure maintains its admissibility property. A similar argument in this case can be used to demonstrate that the screen and sign stage will enable an admissible MRD procedure to maintain its admissibility property. Thus for MRDSS to be admissible it suffices to prove that MRD is admissible. In Cohen, Sackrowitz, and Xu (2009) MRD was shown to be admissible in a multivariate normal model where residuals were defined as functions of the coefficients of the parameters to be tested. For the $R \times C$ contingency table model the S_{ij}, $i = 1, \ldots, R-1$; $j = 1, \ldots, C-1$ are the coefficients of the parameters to be tested and suitably centered and normalized they have the identical properties as the residuals U_{mj} in Cohen, Sackrowitz, and Xu (2009). That is, suppose we are examing the admissibility of the individual test of $H_{ij} : \mu_{ij} = 0$ vs $K_{ij} : \mu_{ij} \ne 0$. Then $|U_{ij}^{(m)}|$ given in (5.1) and (5.2) and given implicitly for $m > 2$ decrease and then increase as a function of S_{ij} while all other S_{ij} are fixed. Also all other $|U_{ij}^{(m)}|$ do not change. These are the properties needed to prove MRD is admissible. □

Remark 5.2. In the case of a $2 \times C$ table Fisher's test statistic can be used instead of Pearson's Chi-Square statistic and it too has the same required properties; namely the statistic for testing H_{ij} is decreasing and then increasing while other $S_{i'j'}$ remain fixed and all other Fisher statistics for hypotheses $H_{i'j'}$ remain unchanged.

6. $R \times C$ example

In this section we illustrate the method with an 3×3 table example. The data to be used are shown below in Table 6.1.

The data are from Agresti (1984b):

TABLE 6.1
Cross-Classification of Attitude Toward Abortion by Amount of Schooling

	Generally Disapprove	Middle Position	Generally Approve
Less than high school	209	101	237
High school	151	126	426
More than high school	16	21	138

There are a total of $2 \times 2 = 4$ local log odds ratios, i.e., a total of 4 hypotheses to be tested.

For the MRD stage, we use critical values obtained from the significance level $1-(1-\alpha)^{1/k}$, $k = 1, \ldots, 4$ for our test statistics at step 4,3,2 and 1, respectively, where $\alpha = 0.05$. Hence the critical values are:

$$(\Omega_1, \ldots, \Omega_4) = (2.49, 2.39, 2.24, 1.96)$$

At stage 1 step 1, we first solve the equations (5.3)–(5.6) to get the estimated mean of each cell given the row and column totals, under the assumption all local log odds ratios equal to 0. The mean value of each cell is shown below in Table 6.2.

TABLE 6.2
Step 1 – Expected cell frequencies

	1	2	3
1	144.33	95.20	307.47
2	185.49	122.35	395.16
3	46.18	30.45	98.37

After this, we can compute the conditional variance of S_{ij}. The conditional variance is shown below in Table 6.3.

TABLE 6.3
Step 1 – Conditional variances of S_{ij}

	1	2	3
1	65.46	82.96	0.00
2	29.82	37.79	0.00
3	0.00	0.00	0.00

Note the zeros in Table 6.3 reflect the fact that the corresponding S_{ij} variables are fixed.

Now, we can compute the test statistics $U_{ij}^{(1)}$, shown in Table 6.4.

The maximum of these is 7.99, which is greater than the critical value 2.49, hence we reject the hypothesis H_{11} and continue to the next step. We also record the sign of the test: $+$.

TABLE 6.4
Step 1 – Test statistics $U_{ij}^{(1)}$

	1	2
1	7.99	7.74
2	5.53	6.45

At step 2, we estimate the mean of each cell conditioned on the row and column summation being fixed, and conditioned on all local log odds ratios except μ_{11} are 0, and also conditioned on S_{11} fixed. Solve the subset of equations (5.3)–(5.8) as explained earlier to get the estimated mean of each cell: shown in Table 6.5.

TABLE 6.5
Step 2 – Expected cell frequencies

	1	2	3
1	209	79.91	258.09
2	133.71	134.59	434.70
3	33.29	33.50	108.21

As in step 1, we can compute the conditional variance of S_{ij}: conditioned on S_{iC}, S_{Rj}, S_{11} are known and fixed. The conditional variance of S_{ij} are shown below in Table 6.6:

TABLE 6.6
Step 2 – Conditional variances of S_{ij}

	1	2	3
1	0.00	41.36	0.00
2	21.58	34.71	0.00
3	0.00	0.00	0.00

Now we can compute the test statistics $U_{ij}^{(2)}$:

TABLE 6.7
Step 2 – Test statistics $U_{ij}^{(2)}$

	1	2
1		3.28
2	3.72	5.06

The maximum of these is 5.06, which is greater than 2.39, so we reject H_{22} and continue to the next step. Also we record the sign of the test: +.

Step 3: We estimate the mean of each cell conditioned on the row and column sum fixed, conditioned on all local log odds ratios (except μ_{11} and μ_{22}) are 0, and also conditioned on S_{11} and S_{22} are fixed. As in step 2, solve the subset set of equations (5.3)–(5.8) to get the estimated mean of each cell.

Now, we can compute the conditional variance of S_{ij}: conditioned on S_{iC}, $i = 1, \ldots, R-1, S_{Rj}$, $j = 1, \ldots, C$, S_{11} and S_{22} are all fixed. The conditional variances of S_{ij} are in Table 6.9.

The test statistics $U_{ij}^{(3)}$ are in Table 6.10.

Table 6.8
Step 3 – Expected cell frequencies

	1	2	3
1	209	87.01	250.99
2	148.16	142.83	412.01
3	18.84	18.16	138

Table 6.9
Step 3 – Conditional variances of S_{ij}

	1	2	3
1	0.00	40.81	0.00
2	8.34	0.00	0.00
3	0.00	0.00	0.00

Table 6.10
Step 3 – Test statistics $U_{ij}^{(3)}$

	1	2
1		2.19
2	0.98	

The maximum of these is 2.19, which is smaller than the critical value 2.24, So MRD stops here, and all remaining hypotheses are accepted.

Now the screening stage. The P-values based on Pearson's Chi-square test statistics for the 4 local 2×2 tables are shown below:

Table 6.11
Screen – local Pearson's Chi-square P-values

	1	2
1	0.002	0.024
2	0.265	0.011

If we choose $\alpha_L = 0.005$, $\alpha_U = 0.05$, then it is clear that screening stage does not have any effect. So MRD+screening gives the same decisions as MRD.

It is clear that sign stage does not have any action either. So the decision of MRDSS is: reject H_{11}, H_{22} and accept H_{12}, H_{21}.

Thus there is evidence of opinion differences only at the extremes. That is, the "generally disapprove" and "middle position" categories when paired with "less than high school" and "high school" as well as "middle position" and "generally approve" categories when paired with "high school" and "more than high school".

7. Simulation for $R \times C$ table

We did some simulations to compare our method, MRDSS, to Holm's step down method. FWER is controlled at level $\alpha = 0.05$ for Holm's step down (SD). For the MRD stage of MRDSS we use critical values obtained from the normal distribution based on significance levels $1-(1-\alpha)^{\frac{1}{M+1-i}}$ in step i. Here $\alpha = .05$.

TABLE 7.1
Comparison of MRDSS to Holm's step down (SD) for 3×3 tables

	θ_{11}	θ_{21}	θ_{12}	θ_{22}	Type I		Type II		FWER		% Power Incr
					MRDSS	SD	MRDSS	SD	MRDSS	SD	
1	1.0	1.0	1.0	1.0	0.038	0.036	0.000	0.000	0.036	0.031	
2	1.0	1.0	1.0	3.0	0.044	0.034	0.433	0.587	0.043	0.031	37.0
3	1.0	1.0	1.0	0.3	0.044	0.033	0.483	0.618	0.041	0.028	35.1
4	1.0	1.0	3.0	3.0	0.032	0.027	0.994	1.238	0.031	0.025	31.9
5	3.0	1.0	1.0	3.0	0.044	0.027	0.736	0.983	0.044	0.026	24.2
6	1.0	3.0	3.0	1.0	0.041	0.028	1.104	1.325	0.040	0.026	32.7
7	0.3	1.0	1.0	0.3	0.041	0.026	1.086	1.302	0.039	0.025	30.9
8	1.0	0.3	0.3	1.0	0.044	0.027	0.725	0.972	0.043	0.026	24.0
9	1.0	1.0	0.3	3.0	0.029	0.025	1.208	1.185	0.028	0.023	-2.9
10	1.0	1.0	0.3	0.3	0.030	0.024	0.992	1.227	0.029	0.022	30.4
11	1.0	3.0	3.0	3.0	0.022	0.013	1.651	1.979	0.022	0.013	32.1
12	1.0	0.3	0.3	0.3	0.020	0.015	1.338	1.677	0.020	0.015	25.6
13	0.3	3.0	3.0	3.0	0.000	0.000	2.379	2.304	0.000	0.000	-4.4
14	3.0	3.0	3.0	3.0	0.000	0.000	2.075	2.437	0.000	0.000	23.2
15	3.0	0.3	0.3	0.3	0.000	0.000	2.298	2.205	0.000	0.000	-5.2
16	0.3	0.3	0.3	0.3	0.000	0.000	2.073	2.412	0.000	0.000	21.4
17	1.0	1.0	1.0	0.5	0.037	0.028	0.812	0.869	0.035	0.024	43.1
18	0.5	1.0	1.0	0.5	0.035	0.026	1.630	1.757	0.034	0.024	52.5
19	1.0	0.5	0.5	1.0	0.034	0.022	1.510	1.682	0.034	0.021	53.9
20	1.0	1.0	0.5	3.0	0.025	0.022	1.414	1.450	0.024	0.021	6.6
21	1.0	1.0	0.5	0.5	0.026	0.018	1.549	1.741	0.025	0.017	74.3
22	1.0	0.5	0.5	0.5	0.019	0.011	2.264	2.554	0.019	0.011	65.1
23	0.5	3.0	3.0	3.0	0.000	0.000	2.695	2.711	0.000	0.000	1.3
24	3.0	0.5	0.5	0.5	0.000	0.000	3.120	3.154	0.000	0.000	4.0
25	0.5	0.5	0.5	0.5	0.000	0.000	3.059	3.453	0.000	0.000	72.0
26	1.0	1.0	1.0	2.0	0.043	0.032	0.787	0.860	0.042	0.028	51.5
27	1.0	1.0	2.0	2.0	0.029	0.023	1.554	1.740	0.028	0.021	71.6
28	2.0	1.0	1.0	2.0	0.035	0.024	1.511	1.675	0.035	0.023	50.7
29	1.0	2.0	2.0	1.0	0.036	0.024	1.632	1.757	0.035	0.023	51.6
30	1.0	2.0	2.0	2.0	0.018	0.012	2.387	2.626	0.018	0.012	63.9
31	2.0	2.0	2.0	2.0	0.000	0.000	3.061	3.463	0.000	0.000	74.9
32	1.0	1.0	0.5	2.0	0.024	0.024	1.726	1.718	0.024	0.021	-2.9
33	0.5	2.0	2.0	2.0	0.000	0.000	3.369	3.451	0.000	0.000	15.0
34	2.0	0.5	0.5	0.5	0.000	0.000	3.291	3.429	0.000	0.000	24.2

For the screen stage we use $\alpha_L = 1 - (1-\alpha)^{1/M}$, $\alpha_U = \alpha = .05$. For a 3×3 table, M = 4; for a 3×4 table, M = 6.

The type I and type II error columns reflect the average number of type I and type II errors respectively.

We define the power of a MTP as:

power = 1 - E { # of type II errors } / # of true non-null hypotheses.

The percentage power increase of MRDSS relative to SD is defined as:

power incr = $100 \times$(power$_{MRDSS}$ - power$_{SD}$)/power$_{SD}$

We first list 34 sets of local odds ratios. For the 3×3 table, each set contains 4 local odds ratios.

TABLE 7.2
Comparison of MRDSS to Holm's step down (SD) for 3×4 *tables*

	θ_{11}	θ_{21}	θ_{12}	θ_{22}	θ_{13}	θ_{23}	Type I		Type II		FWER		% P.Inc
							MRDSS	SD	MRDSS	SD	MRDSS	SD	
1	1.0	1.0	1.0	1.0	1.0	1.0	0.042	0.037	0.000	0.000	0.038	0.031	
2	1.0	1.0	1.0	1.0	1.0	2.0	0.040	0.030	0.799	0.894	0.038	0.027	90.1
3	1.0	1.0	2.0	1.0	1.0	1.0	0.049	0.033	0.788	0.887	0.047	0.029	86.7
4	2.0	1.0	1.0	1.0	1.0	1.0	0.046	0.032	0.798	0.890	0.044	0.028	84.6
5	1.0	1.0	1.0	1.0	2.0	2.0	0.036	0.024	1.571	1.793	0.034	0.021	107.5
6	1.0	1.0	2.0	2.0	1.0	1.0	0.044	0.024	1.557	1.788	0.042	0.022	108.7
7	2.0	1.0	1.0	1.0	1.0	2.0	0.051	0.027	1.544	1.761	0.049	0.024	90.4
8	1.0	2.0	1.0	1.0	1.0	2.0	0.045	0.025	1.600	1.793	0.043	0.023	93.3
9	1.0	1.0	2.0	1.0	1.0	2.0	0.049	0.025	1.535	1.748	0.048	0.023	84.9
10	1.0	1.0	1.0	2.0	1.0	2.0	0.037	0.028	1.545	1.789	0.036	0.026	115.7
11	2.0	1.0	1.0	1.0	2.0	2.0	0.039	0.021	2.320	2.660	0.038	0.019	99.7
12	1.0	2.0	1.0	1.0	2.0	2.0	0.040	0.022	2.428	2.706	0.039	0.019	94.9
13	1.0	1.0	2.0	1.0	2.0	2.0	0.032	0.019	2.315	2.675	0.031	0.017	110.7
14	1.0	1.0	1.0	2.0	2.0	2.0	0.035	0.018	2.419	2.716	0.034	0.018	104.3
15	2.0	1.0	2.0	2.0	1.0	1.0	0.039	0.022	2.316	2.655	0.038	0.020	98.4
16	2.0	2.0	2.0	2.0	1.0	1.0	0.024	0.013	3.113	3.593	0.023	0.013	117.8
17	1.0	1.0	2.0	2.0	2.0	1.0	0.038	0.019	2.431	2.728	0.037	0.018	108.9
18	1.0	1.0	2.0	2.0	1.0	2.0	0.038	0.022	2.305	2.645	0.037	0.021	95.8
19	1.0	1.0	2.0	2.0	2.0	2.0	0.025	0.014	3.114	3.593	0.024	0.013	117.7
20	2.0	2.0	1.0	1.0	2.0	2.0	0.027	0.013	3.140	3.589	0.027	0.012	109.1
21	1.0	2.0	1.0	2.0	2.0	2.0	0.024	0.013	3.242	3.640	0.024	0.012	110.2
22	1.0	2.0	2.0	1.0	2.0	2.0	0.033	0.016	3.166	3.599	0.033	0.016	107.7
23	1.0	2.0	2.0	2.0	1.0	2.0	0.031	0.015	3.218	3.583	0.031	0.015	87.7
24	1.0	2.0	2.0	2.0	2.0	1.0	0.032	0.013	3.280	3.667	0.031	0.013	116.6
25	2.0	2.0	2.0	2.0	2.0	2.0	0.000	0.000	4.694	5.392	0.000	0.000	114.7
26	1.0	1.0	1.0	1.0	1.0	3.0	0.042	0.030	0.422	0.621	0.040	0.026	52.7
27	1.0	1.0	3.0	1.0	1.0	1.0	0.046	0.029	0.430	0.647	0.045	0.026	61.5
28	3.0	1.0	1.0	1.0	1.0	1.0	0.045	0.033	0.424	0.626	0.042	0.029	54.1
29	1.0	1.0	1.0	1.0	3.0	3.0	0.036	0.029	1.044	1.368	0.034	0.025	51.3
30	1.0	1.0	3.0	3.0	1.0	1.0	0.047	0.025	0.965	1.350	0.046	0.024	59.2
31	3.0	1.0	1.0	1.0	1.0	3.0	0.053	0.032	0.763	1.149	0.051	0.029	45.4
32	1.0	3.0	1.0	1.0	1.0	3.0	0.047	0.030	1.011	1.342	0.045	0.027	50.2
33	1.0	1.0	3.0	1.0	1.0	3.0	0.048	0.025	0.768	1.134	0.046	0.022	42.2
34	1.0	1.0	1.0	3.0	1.0	3.0	0.040	0.029	0.954	1.321	0.038	0.026	54.1
35	3.0	1.0	1.0	1.0	3.0	3.0	0.038	0.021	1.446	1.919	0.036	0.020	43.7
36	1.0	3.0	1.0	1.0	3.0	3.0	0.039	0.021	1.715	2.143	0.038	0.019	50.0
37	1.0	1.0	3.0	1.0	3.0	3.0	0.035	0.024	1.471	1.951	0.033	0.022	45.7
38	1.0	1.0	1.0	3.0	3.0	3.0	0.038	0.020	1.686	2.149	0.037	0.019	54.4
39	3.0	1.0	3.0	3.0	1.0	1.0	0.042	0.024	1.348	1.850	0.041	0.023	43.7
40	3.0	3.0	3.0	3.0	1.0	1.0	0.025	0.014	2.211	2.755	0.024	0.013	43.8
41	1.0	1.0	3.0	3.0	3.0	1.0	0.045	0.022	1.723	2.188	0.044	0.021	57.2
42	1.0	1.0	3.0	3.0	1.0	3.0	0.039	0.022	1.351	1.844	0.039	0.022	42.6
43	1.0	1.0	3.0	3.0	3.0	3.0	0.026	0.016	2.216	2.738	0.026	0.016	41.3
44	3.0	3.0	1.0	1.0	3.0	3.0	0.032	0.016	2.213	2.724	0.031	0.016	40.1
45	1.0	3.0	3.0	1.0	3.0	3.0	0.036	0.020	2.252	2.793	0.036	0.019	44.8

For each set, we generate N = 10,000 3×3 tables from independent Poisson distributions with parameters in a 3×3 matrix $\boldsymbol{\lambda}$. The matrix $\boldsymbol{\lambda}$ is chosen such that the true local odds ratios are listed in the left 4 columns of Table 7.1, and the row and column sum of λ_{ij} are 60. Hence local odds ratios and row/col sum of λ_{ij} uniquely define $\boldsymbol{\lambda}$, $\boldsymbol{\lambda}$ itself is not shown in the table.

TABLE 7.3
(Continue)Comparison of MRDSS to Holm's step down (SD) for 3×4 *tables*

	θ_{11}	θ_{21}	θ_{12}	θ_{22}	θ_{13}	θ_{23}	Type I		Type II		FWER		% P.Inc
							MRDSS	SD	MRDSS	SD	MRDSS	SD	
46	1.0	3.0	3.0	3.0	1.0	3.0	0.035	0.015	2.197	2.762	0.035	0.015	45.6
47	1.0	3.0	3.0	3.0	3.0	1.0	0.036	0.015	2.519	3.091	0.036	0.015	62.9
48	3.0	3.0	3.0	3.0	3.0	3.0	0.000	0.000	3.473	4.241	0.000	0.000	43.7
49	1.0	1.0	1.0	1.0	0.5	2.0	0.033	0.029	1.781	1.784	0.032	0.026	1.6
50	1.0	1.0	1.0	2.0	0.5	1.0	0.038	0.031	1.688	1.789	0.037	0.028	47.8
51	1.0	1.0	2.0	1.0	0.5	1.0	0.034	0.028	1.789	1.782	0.032	0.025	-3.4
52	1.0	2.0	1.0	1.0	0.5	1.0	0.040	0.029	1.659	1.778	0.038	0.026	53.6
53	2.0	1.0	1.0	1.0	0.5	1.0	0.031	0.024	1.717	1.784	0.029	0.022	31.1
54	1.0	1.0	2.0	0.5	1.0	1.0	0.038	0.025	1.743	1.773	0.036	0.023	13.5
55	1.0	2.0	1.0	0.5	1.0	1.0	0.031	0.025	1.766	1.768	0.029	0.022	0.6
56	2.0	1.0	1.0	0.5	1.0	1.0	0.034	0.024	1.677	1.783	0.033	0.022	48.9
57	0.5	1.0	1.0	1.0	2.0	2.0	0.026	0.021	2.437	2.675	0.025	0.019	72.9
58	1.0	0.5	1.0	1.0	2.0	2.0	0.027	0.020	2.452	2.688	0.026	0.019	75.6
59	1.0	1.0	0.5	1.0	2.0	2.0	0.029	0.021	2.526	2.676	0.028	0.019	46.1
60	1.0	1.0	1.0	0.5	2.0	2.0	0.027	0.022	2.571	2.691	0.026	0.020	39.1
61	2.0	1.0	1.0	1.0	0.5	2.0	0.029	0.020	2.602	2.661	0.028	0.018	17.4
62	1.0	2.0	1.0	1.0	0.5	2.0	0.033	0.021	2.573	2.683	0.031	0.020	34.8
63	1.0	1.0	2.0	1.0	0.5	2.0	0.036	0.019	2.601	2.654	0.035	0.017	15.4
64	1.0	1.0	1.0	2.0	0.5	2.0	0.028	0.020	2.505	2.695	0.026	0.018	62.3
65	2.0	2.0	2.0	2.0	0.5	0.5	0.000	0.000	4.914	5.360	0.000	0.000	69.6
66	2.0	2.0	0.5	0.5	2.0	2.0	0.000	0.000	5.135	5.322	0.000	0.000	27.5
67	0.5	0.5	2.0	2.0	2.0	2.0	0.000	0.000	4.897	5.357	0.000	0.000	71.7
68	2.0	0.5	2.0	0.5	2.0	2.0	0.000	0.000	5.056	5.361	0.000	0.000	47.6
69	0.5	2.0	2.0	2.0	2.0	0.5	0.000	0.000	5.166	5.355	0.000	0.000	29.3
70	2.0	0.5	2.0	2.0	0.5	2.0	0.000	0.000	5.026	5.296	0.000	0.000	38.3

We compare the expected number of type I errors, expected number of type II errors and FWER of MRDSS and SD. We also give the percentage power increase of MRDSS relative to SD.

For the 3×3 table, note that both MRDSS and SD control FWER at $\alpha = .05$. Also note that in 29 out of 33 cases MRDSS had an increase in power over SD and in many cases the increase was substantial.

In Table 7.2 and 7.3, we list the simulation results for a 3×4 table. There are a total of 70 configurations of local odds ratios. For each row, the row sum of the Poisson parameter $\boldsymbol{\lambda}$ is 80, and for each column, the column sum is 60.

Note that MRDSS strongly controls FWER at level $\alpha = .05$ in all 70 cases, except one (where MRDSS has a FWER = 0.051). We also compare the percentage power increase of MRDSS relative to stepdown method. MRDSS had an increase in power with only one exception. The increase was substantial in most cases.

References

AGRESTI, A. (1984a). *An Introduction to Categorical Data Analysis*. Wiley-Interscience. MR1394195

AGRESTI, A. (1984b). *Analysis of Ordinal Categorical Data*. Wiley-Interscience. MR0747468

Benjamini, Y. and Hochberg, Y. (1995). Controlling the false discovery rate: a practical and powerful approach to multiple testing. *Journal of the Royal Statistical Society, Ser. B 57*, 289–289. MR1325392

Chen, C., Cohen, A., and Sackrowitz, H. (2009). Admissible,consistent multiple testing with applications including variable selection. *Electronic Journal of Statistics 3*, 633–650.

Cohen, A. (1987). Optimality properties of some tests for homogeneity in contingency tables. *Statistics 18*, 591–597. MR0906797

Cohen, A., Kolassa, J., and Sackrowitz, H. (2007). A smooth version of the step-up procedure for multiple tests of hypotheses. *Journal of Statistical Planning and Inference* **137**, 11, 3352–3360. MR2363261

Cohen, A. and Sackrowitz, H. (2000). Testing whether treatment is "better" than control with ordered categorical data: Definition and complete class theorems. *Statistics and Decisions 18*, 1–25. MR1763329

Cohen, A. and Sackrowitz, H. (2005a). Characterization of Bayes procedures for multiple endpoint problems and inadmissibility of the step-up procedure. *The Annals of Statistics* **33**, 1, 145–158. MR2157799

Cohen, A. and Sackrowitz, H. (2005b). Decision theory results for one-sided multiple comparison procedures. *The Annals of Statistics* **33**, 1, 126–144. MR2157798

Cohen, A. and Sackrowitz, H. (2007). More on the inadmissibility of step-up. *Journal of Multivariate Analysis* **98**, 3, 481–492. MR2293009

Cohen, A. and Sackrowitz, H. (2008). Multiple testing to two-sided alternatives with dependent data. *Statistica Sinica 18*, 1593–1602. MR2469325

Cohen, A., Sackrowitz, H., and Xu, M. (2009). A new multiple testing method in the dependent case. *The Annals of Statistics 37*. MR2509082

Dudoit, S., Shaffer, J., and Boldrick, J. (2003). Multiple Hypothesis Testing in Microarray Experiments. *Statistical Science* **18**, 1, 71–103. MR1997066

Dudoit, S. and Van Der Laan, M. (2008). *Multiple Testing Procedures with Applications to Genomics*. Springer Verlag. MR2373771

Gabriel, K. R. (1966). Simultanesous test procedures for multiple comparisons on categocal data. *Journal of the American Statistical Association 61*, 1081–1096. MR0207104

Hirotsu, C. (1978). Multiple comparisons and grouping rows in a two-way contingency table. *Rep. Statis, Appl. Res., Union of Japanese Scientists and Engineers* **25**, 1, 1–12. MR0510359

Hirotsu, C. (1983). Defining the pattern of association in two-way contingency tables. *Biometrika 70*, 579–589. MR0725371

Hochberg, Y. and Tamhane, A. (1987). *Multiple comparison procedures*. John Wiley & Sons. MR0914493

Lehmann, E. and Romano, J. (2005). *Testing Statistical Hypotheses*, Third ed. Springer. MR2135927

Matthes, T. and Truax, D. (1967). Tests of composite hypotheses for the multivariate exponential family. *The Annals of Mathematical Statistics 38*, 681–697. MR0208745

R Development Core Team. (2009). *R: A Language and Environment for Statistical Computing*. R Foundation for Statistical Computing, Vienna, Austria. ISBN 3-900051-07-0, http://www.R-project.org.

Shaffer, J. (1986). Modified sequentially rejective multiple test procedures. *Journal of the American Statistical Association 81*, 826–831.