Testing for Conditional Multiple Marginal Independence

Christopher R. Bilder
Department of Statistics
Oklahoma State University
www.chrisbilder.com
bilder@okstate.edu

Thomas M. Loughin
Department of Statistics
Kansas State University

March 26, 2001

What is CMMI?

- Conditional multiple marginal independence (CMMI)
- Survey of 239 college women
  - Contraceptive use and first time urinary tract infection (UTI)

Age>=24		Contraceptive					
		Oral	Condom	L. Cond.	Spermicide	Diaphragm	Total
UTI	No	18	9	8	7	0	24
	Yes	8	9	2	3	2	14
							38

Age<24		Contraceptive					
		Oral	Condom	L. Cond.	Spermicide	Diaphragm	Total
UTI	No	55	41	37	27	0	85
	Yes	75	68	33	22	5	116
							201

  - LogXact 4 manual (p. 198-9) and Foxman et al. (1997)

What is CMMI?

- Contraceptive use is a “pick any/c variable”
  - Coombs (1964)
  - Subjects may pick any out of the 5 types of contraception
  - Each choice is referred to as an “item” (Agresti and Liu, 1999)
- UTI is referred to as the “group variable”
- Age is referred to as the “stratification variable”

Age>=24		Contraceptive					
		Oral	Condom	L. Cond.	Spermicide	Diaphragm	Total
UTI	No	18	9	8	7	0	24
	Yes	8	9	2	3	2	14
							38

Age<24		Contraceptive					
		Oral	Condom	L. Cond.	Spermicide	Diaphragm	Total
UTI	No	55	41	37	27	0	85
	Yes	75	68	33	22	5	116
							201

What is CMMI?

- Hypothesis test for CMMI
  - Are the contraception practices of college women marginally independent of their UTI history, controlling for age?

Age>=24		Contraceptive					
		Oral	Condom	L. Cond.	Spermicide	Diaphragm	Total
UTI	No	18 (0.75)	9 (0.38)	8 (0.33)	7 (0.29)	0 (0.00)	24
	Yes	8 (0.57)	9 (0.64)	2 (0.14)	3 (0.21)	2 (0.14)	14
							38

Observed proportion of women selecting an item is in parentheses

Age<24		Contraceptive					
		Oral	Condom	L. Cond.	Spermicide	Diaphragm	Total
UTI	No	55 (0.65)	41 (0.48)	37 (0.44)	27 (0.32)	0 (0.00)	85
	Yes	75 (0.65)	68 (0.59)	33 (0.28)	22 (0.19)	5 (0.04)	116
							201

What is CMMI?

- Hypothesis test for CMMI
  - Let $\pi_{j|ik}$=P(subject picks item j|subject is in group i and stratum k)
    - i=1,…,r denotes the group (row)
    - j=1,…,c denotes the item (column)
    - k=1,…,q denotes the stratum
  - Hypotheses (in general)
    $H_o:\pi_{j|1k}=\pi_{j|2k}=\ldots=\pi_{j|rk}$ for j=1, …,c and k=1,…,q
    H_a:At least one of the equalities does not hold
- Cochran (1954) and Mantel and Haenszel (1959) tests should not be used
  - Need the pick any/c variable to be pick 1/c

Purpose

- Derive test for CMMI
  - Extend Cochran's statistic to include $r \times 2 \times q$ tables
    - Test for conditional independence for each item
  - Sum c "extended Cochran" statistics to form a new statistic to test for CMMI
    - Approximations to its sampling distribution

Multiple Marginal Independence (MMI)

- Special case of CMMI when the number of strata is 1
- Previous research
  - Umesh (1995)
  - Loughin and Scherer (1998)
  - Agresti and Liu (1998, 1999)
  - Decady and Thomas (2000)
  - Bilder, Loughin, and Nettleton (2000)
    - Conclude the best MMI testing methods are:
      - Bootstrapping the naïve chi-squared statistic
      - Bootstrapping p-value combination methods
    - Most consistently hold the correct size while providing power against various alternatives

Notation

- For subject s in row i, $Y_{s(ik)j}=1$ if a positive response is given for item j and $Y_{s(ik)j}=0$ for a negative response
  - $s=1,\ldots,n$
  - All responses by subject s can be viewed as an item response vector, $Y_{s(ik)} = (Y_{s(ik)1}, Y_{s(ik)2}, \ldots, Y_{s(ik)c})'$.
    - $h=1,\ldots,2^c$ different item response vectors
- Let n_{ihk} = number of observed subjects for the h^{th} item response vector
  - Let $n_{ik} = (n_{i1k},\ldots,n_{i2^c k})' \sim \text{Multinomial}(\tau_{1|ik},\ldots,\tau_{2^c|ik})$
    - Independent multinomial sampling within each row of each stratum
- Let m_{ijk} = number of positive responses to item j in group i and stratum k
  - Then $m_{ijk} = \sum_s Y_{s(ik)j} = \sum_{\{h:Y_j=1\}} n_{ihk}$

Notation

- Example 2×2×q table

		k=1	
	j=	1	2
i=	1	$\pi_{1\|11}$	$\pi_{2\|11}$
	2	$\pi_{1\|21}$	$\pi_{2\|21}$

k=2	
1	2
$\pi_{1\|12}$	$\pi_{2\|12}$
$\pi_{1\|22}$	$\pi_{2\|22}$

…

k=q	
1	2
$\pi_{1\|1q}$	$\pi_{2\|1q}$
$\pi_{1\|2q}$	$\pi_{2\|2q}$

		k=1	
	j=	1	2
i=	1	m_{111}	m_{121}
	2	m_{211}	m_{221}
		m_{+11}	m_{+21}

k=2	
1	2
m_{112}	m_{122}
m_{212}	m_{222}
m_{+12}	m_{+22}

…

k=q	
1	2
m_{11q}	m_{12q}
m_{21q}	m_{22q}
m_{+1q}	m_{+2q}

Extended Cochran Statistic - C_j^2

- Consider the j[th] item and examine the responses over the strata
  - Develop a test for conditional independence

- Let

$$m_{j+} = \sum_{k=1}^{q} (m_{1jk}, \ldots, m_{r-1,jk})'$$

$$m_{j+}^{\circ} = \sum_{k=1}^{q} \left(\frac{n_{1+k} m_{+jk}}{n_{++k}}, \ldots, \frac{n_{r-1,+k} m_{+jk}}{n_{++k}} \right)'$$

Extended Cochran Statistic - C_j^2

- Let $\hat{V}_j$ denote the estimated covariance matrix of $\mathbf{m}_{j+} - \mathbf{m}^{\circ}_{j+}$ under conditional independence with elements of

$$\sum_{k=1}^{q} \widehat{Cov}\left(m_{ijk} - \frac{n_{i+k}m_{+jk}}{n_{++k}}, m_{i'jk} - \frac{n_{i'+k}m_{+jk}}{n_{++k}} \right)$$

$$= \sum_{k=1}^{q} \frac{m_{+jk}(n_{++k} - m_{+jk})n_{i+k}(\delta_{ii'}n_{++k} - n_{i'+k})}{n_{++k}^3}$$

$$\text{where } \delta_{ii'} = \begin{cases} 1 \text{ if } i = i' \\ 0 \text{ if } i \neq i' \end{cases}$$

Extended Cochran Statistic - C_j^2

- The "Extended Cochran Statistic" is

$$C_j^2 = (m_{+j} - m^{\circ}_{+j})' \hat{V}_j^{-1} (m_{+j} - m^{\circ}_{+j})$$

- Under conditional independence as $n \to \infty$:
  - $n^{-1}\hat{V}_j \xrightarrow{p} V_j$
  - $n^{-1/2}(m_{+j} - m^{\circ}_{+j}) \xrightarrow{d} N(0, V_j)$
  - Then $C_j^2 \xrightarrow{d} \chi^2_{r-1}$
- Notes:
  - C_j^2 simplifies to Cochran's original statistic for $2\times2\times q$
  - Similar to the Generalized Mantel-Haenszel test statistic
    - Landis, Heyman, and Koch (1978)
    - Difference: Assumes a multiple hypergeometric distribution in each stratum.

Modified Cochran Statistic - C_M^2

- To test CMMI, the c C_j^2 are summed to form the "modified Cochran" statistic

- $$C_M^2 = \sum_{j=1}^{c} C_j^2 = (\mathbf{m}_+ - \mathbf{m}_+^\circ)'\hat{\mathbf{Q}}^{-1}(\mathbf{m}_+ - \mathbf{m}_+^\circ)$$
  $$\mathbf{m}_+ = (\mathbf{m}_{1+}', \ldots, \mathbf{m}_{c+}')'$$
  $$\mathbf{m}_+^\circ = (\mathbf{m}_{1+}^{\circ\prime}, \ldots, \mathbf{m}_{c+}^{\circ\prime})'$$
  $$\hat{\mathbf{Q}} = \text{Diag}(\hat{\mathbf{V}}_j)$$

- If item responses for each subject are independent, then
  $$C_M^2 \xrightarrow{d} \chi^2_{c(r-1)} \text{ under CMMI}$$
  - Reject CMMI if $C_M^2 > \chi^2_{c(r-1),1-\alpha}$

Modified Cochran Statistic - C_M^2

- In most situations, the item responses are **not** independent
- Note that under CMMI:
  - $n^{-1}\hat{Q} \xrightarrow{p} Q = \text{Diag}(V_j)$
  - $n^{-1/2}(m_+ - m_+^\circ) \xrightarrow{d} N(0, P_o)$

  Note: P_o is excluded from here due to time constraints
- Then under CMMI

$$C_M^2 \xrightarrow{d} \sum_{p=1}^{c(r-1)} \lambda_p X_p^2$$

where

λ_p's are the eigenvalues of $Q^{-1}P_o$

X_p^2's are independent χ_1^2 random variables

Modified Cochran Statistic - C_M^2

- Nonparametric bootstrap (bootstrap C_M^2)
  - Approximate sampling distribution of C_M^2
  - Algorithm
    - Take B resamples of size n by randomly selecting $Y_{s(i)k} = (Y_{s(ik)1}, Y_{s(ik)2}, \ldots, Y_{s(ik)c})'$ and group (row) designation independently with replacement from the original data within strata
    - For each resample, calculate the test statistic, $C_{M,b}^{2*}$, for b=1,…,B
    - Calculate the p-value as

$$\frac{1}{B}\sum_{b=1}^{c(r-1)} I\left(C_{M,b}^{2*} > C_M^2\right)$$

where I(A)=1 if event A occurs, 0 otherwise

Other CMMI testing methods

- Bootstrap p-value combination methods
  - Combine the p-values from C_j^2 (using a χ^2_{r-1} app.) for j=1,…,c to form a "new" test statistic, $\tilde{p}$
  - Product of the p-values or minimum p-value
  - Algorithm
    - Take B resamples of size n by randomly selecting $Y_{s(i)k} = (Y_{s(i)1k}, Y_{s(i)2k}, \ldots, Y_{s(i)ck})'$ and group (row) designation independently with replacement from the original data
    - For each resample, calculate the test statistic, $\tilde{p}_b^*$, for b=1,…,B
    - Calculate the p-value as $\frac{1}{B}\sum_{b=1}^{B} I(\tilde{p}_b^* < \tilde{p})$
- Bonferroni adjustment to the c C_j^2
  - Reject CMMI if $C_j^2 > \chi^2_{r-1,1-\alpha/c}$

UTI Data

- Evidence against CMMI
  - 5,000 resamples
  - Rejection may be due to Diaphragm and L.Cond.
    - $\chi_1^2(0.99) = 6.63$

CMMI Testing Method	P-value
C_M^2 using $\chi^2_{c(r-1)}$ app.	0.0005
Bootstrap C_M^2	0.0058
Bootstrap prod. p-values	0.0330
Bootstrap min. p-values	0.0060
Bonferroni	0.0408

	Contraceptive				
	Oral	Condom	L. Cond.	Spermicide	Diaphragm
$C_j^2 =$	0.20	3.87	6.42	4.51	7.00

CMMI Type I Error Simulations

- Estimated type I error rate
  - Proportion of data sets in which CMMI is incorrectly rejected
- Data generated using an algorithm by Gange (1995)
  - Specify $\pi_{j|ik}$'s
    - Under CMMI
  - Specify odds ratios (ORs)

$$OR_{AB} = \frac{Odds(Y_B = 1 \mid Y_A = 1)}{Odds(Y_B = 1 \mid Y_A = 0)}$$

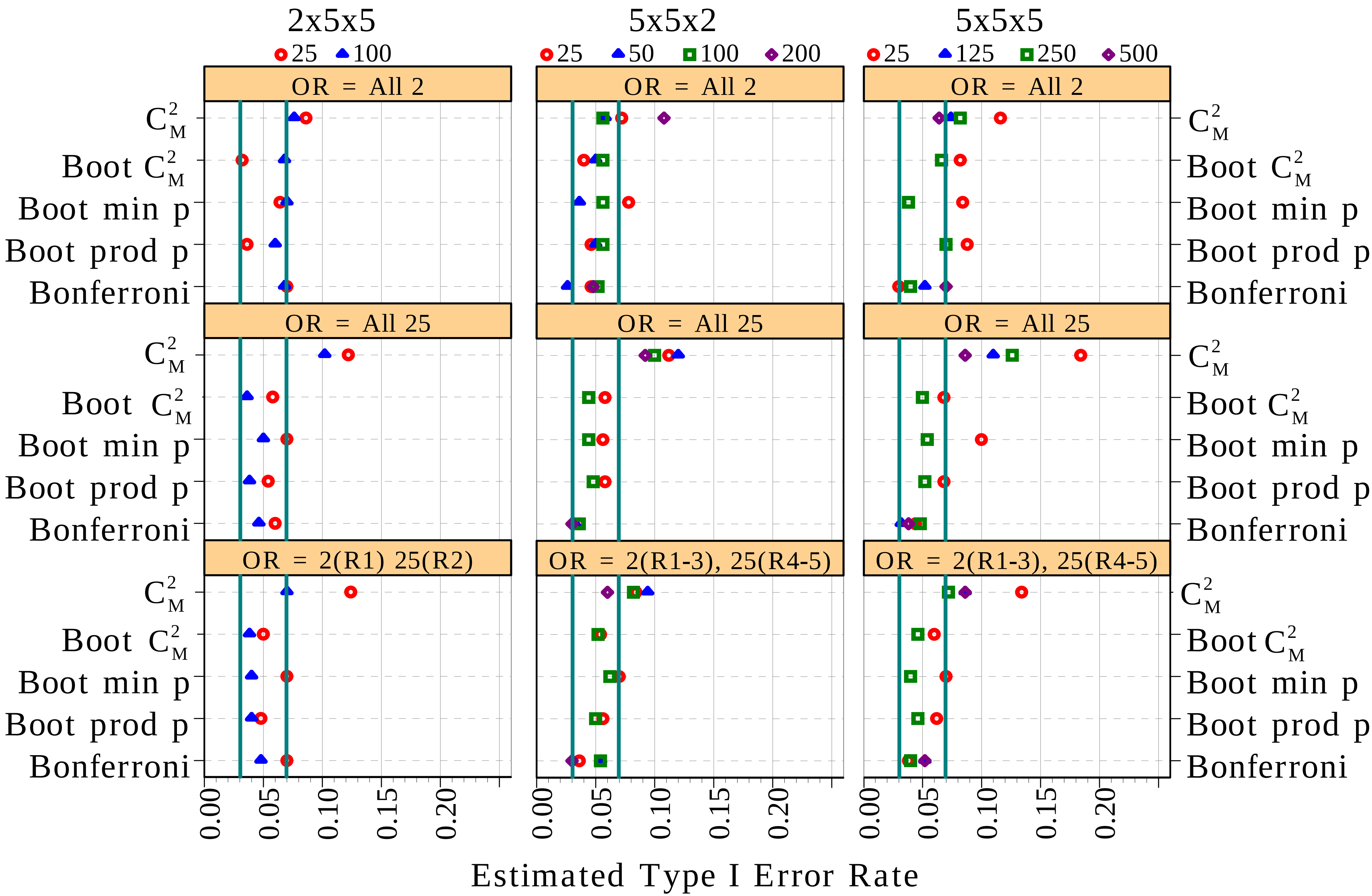
- For each data set generated, the CMMI testing methods are applied

CMMI Type I Error Simulations

- Settings held constant for each simulation
  - Nominal type I error rate=0.05
  - 500 data sets generated
  - 1,000 resamples for bootstrap methods
  - Expected range of estimated type I error rates for methods holding the nominal level:

$$0.05 \pm 2\sqrt{\frac{(0.05)(1-0.05)}{500}} = 0.05 \pm 0.0195$$

- Trellis plots on next slide shows estimated type I error rates
  - Includes only the $\pi_{1|ik}=0.1$, $\pi_{2|ik}=0.2$, $\pi_{3|ik}=0.3$, $\pi_{4|ik}=0.4$, and $\pi_{5|ik}=0.5$ cases for $i=1,\ldots,r$ and $k=1,\ldots,q$
  - Results generalize to other cases



•Bootstrap simulations are only run at n=25 and n=100 for 5×5×2 and n=25 and 250 for 5×5×5

•OR=2(R1) 25(R2) means all ORs are 2 for row 1 and 25 for row 25

CMMI Type I Error Simulations

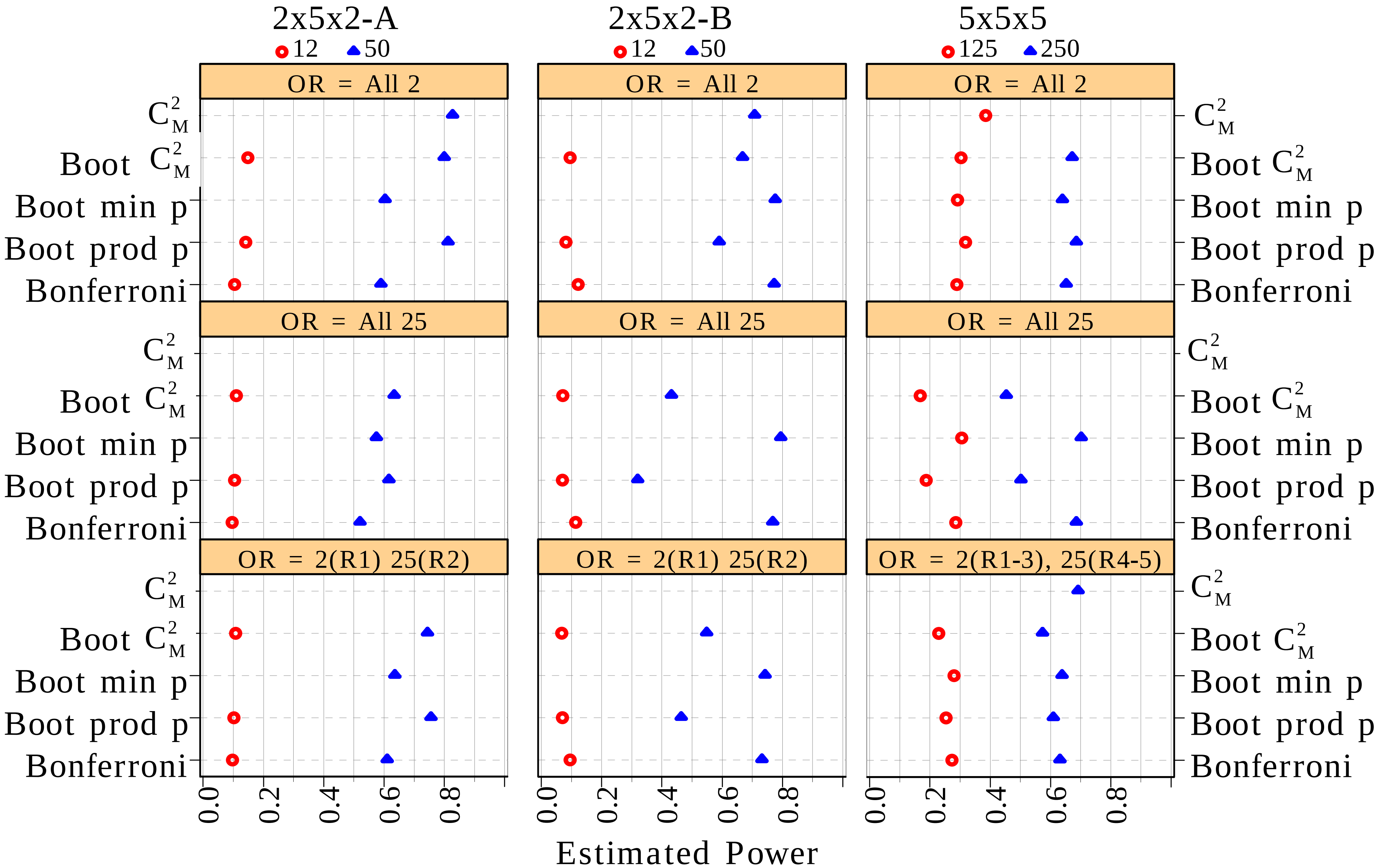
- Summary
  - Bootstrap C_M^2, bootstrap product of p-values, and Bonferroni testing methods most consistently hold the correct size
  - Bootstrap minimum p-value holds the correct size, provided sample size is not too small
  - C_M^2 with a $\chi^2_{c(r-1)}$ approximation does not hold the correct size with large pairwise association between item responses
    - Corresponds to when the variation among the eigenvalues is the greatest.

CMMI Power Simulations

- Proportion of data sets in which MMI is correctly rejected
- Data generated same way as in the type I error simulation study except that marginal probabilities differ across the rows
- Trellis plot on next slide shows the estimated power
  - Includes only a few of the cases examined
  - Some estimated powers are excluded for the plot
    - Do not hold the correct size for comparable marginal probabilities, ORs, sample sizes, and marginal tables sizes
  - Marginal probabilities used (same across strata):

Row	2x5x2-A	2x5x2-B
1	0.1,0.2,0.3,0.4,0.5	0.1,0.2,0.3,0.4,0.5
2	0.3,0.4,0.1,0.2,0.3	0.5,0.2,0.3,0.4,0.5

Row	5x5x5
1-3	0.1,0.2,0.3,0.4,0.5
4-5	0.1,0.3,0.4,0.2,0.5



Some estimated powers are excluded from the plot for methods that do not hold the correct size for comparable marginal probabilities, ORs, …

CMMI Power Simulations

- Summary
  - There is not one method with uniformly largest power
  - Bootstrap C_M^2, bootstrap product of p-values have comparable powers when plotted
  - Bootstrap minimum p-value and Bonferroni have comparable powers when plotted
  - Some p-value combination methods are better at detecting certain types of alternative hypotheses (Loughin, 2000)
    - Deviation from CMMI for only a few items - minimum p-value has higher power
    - Deviation from CMMI for most items by the same degree - product of p-values has higher power

CMMI Testing Recommendations

- Bootstrap C^2_M, Bootstrap product of p-values, and Bonferroni
  - Most consistently hold the correct size
  - Provide power against detecting various alternatives

Testing for Conditional Multiple Marginal Independence

Christopher R. Bilder
Department of Statistics
Oklahoma State University
www.chrisbilder.com
bilder@okstate.edu

Thomas M. Loughin
Department of Statistics
Kansas State University

March 26, 2001

References

- Agresti, A. and Liu, I.-M. (1998). *Modelling Responses to a Categorical Variable Allowing Arbitrarily Many Category Choices*. Univ. of Florida Dept. of Statistics Technical Report no.575.
- Agresti, A. and Liu, I.-M. (1999). Modeling a Categorical Variable Allowing Arbitrarily Many Category Choices. *Biometrics* 55, 936-943.
- Bilder, C. R., Loughin, T. M., Nettleton, D. (2000). Multiple Marginal Independence Testing for Pick Any/c Variables. To appear in *Communications in Statistics: Simulation and Computation* 29(4).
- Cochran, W. G. (1954). Some Methods for Strengthening the Common χ^2 Test. *Biometrics* 10, 417-451.
- Coombs, C. H. (1964). *A Theory of Data*. New York: John Wiley & Sons, Inc.
- Decady, Y. J. and Thomas, D. R. (2000). A simple test of association for contingency tables with multiple column responses. *Biometrics* 56, 893-896..
- Foxman, B., Marsh, J. Gellespie, B., Rubin, N. Kopman, J., and Spear, S. (1997). Condom Use and First-Time Urinary Tract Infection. *Epidemiology* 8, 637-641.
- Gange, S. J. (1995). Generating Multivariate Categorical Variables Using the Iterative Proportional Fitting Algorithm. *The American Statistician* 49, 134-138.
- Landis, J. R., Heyman, E. R., Koch, G. G. (1978). Average Partial Association in Three-way Table Contingency Tables: a Review and Discussion of Alternative Tests. *International Statistical Review* 46, 237-254.
- Loughin, T. M. and Scherer, P. N. (1998). Testing for Association in Contingency Tables with Multiple Categorical Responses. *Biometrics* 54, 630-637.
- Loughin, T. M. (2000). A systematic comparison of methods for combining independent tests. Technical Report, Kansas State University, Department of Statistics, Manhattan, KS.
- Mantel, A. M. and Haenszel, W. (1959). Statistical Aspects of the Analysis of Data from Retrospective Studies on Disease. *Journal of the National Cancer Institute* 22, 719-748.