

The Geometry of Statistical Models for Two-Way Contingency Tables with Fixed Odds Ratios

ENRICO CARLINI AND FABIO RAPALLO (*)

Contribution to “School (and Workshop) on Computational Algebra for Algebraic Geometry and Statistics”, Torino, September 2004.

SUMMARY. - *We study the geometric structure of the statistical models for two-by-two contingency tables. One or two odds ratios are fixed and the corresponding models are shown to be a portion of a ruled quadratic surface or a segment. Some pointers to the general case of two-way contingency tables are also given and an application to case-control studies is presented.*

1. Introduction

A two-way contingency table gives the joint distribution of two random variables with a finite number of outcomes. If we denote by $\{0, \dots, I-1\}$ and $\{0, \dots, J-1\}$ the outcomes of X_1 and X_2 respectively, the contingency table is represented by a matrix $P = (p_{ij})$, where p_{ij} is the probability that $X_1 = i$ and $X_2 = j$. The table P is also called an $I \times J$ contingency table, in order to emphasize that the variable X_1 has I outcomes and the variable X_2 has J outcomes.

(*) Authors' addresses: Enrico Carlini, Dipartimento di Matematica, Politecnico di Torino, Corso Duca degli Abruzzi, 24, I-10129 Torino, ITALY, e-mail: enrico.carlini@polito.it

Fabio Rapallo, Dipartimento di Matematica, University of Genova, via Dodecaneso, 35, I-16146 Genova, ITALY, e-mail: rapallo@dima.unige.it

AMS Subject Classification: 14Q99, 62H17.

In the analysis of contingency tables odds ratios, or cross-product ratios, are major parameters, and their use in the study of 2×2 tables goes back to the 1970's. For an explicit discussion on this approach see, e.g., [6].

For a 2×2 table of the form:

$$\begin{pmatrix} p_{00} & p_{01} \\ p_{10} & p_{11} \end{pmatrix} \quad (1)$$

there is only one cross-product ratio, namely:

$$r = \frac{p_{00}p_{11}}{p_{01}p_{10}}.$$

In the general $I \times J$ case, there is one cross-product ratio for each 2×2 submatrix of the table. Thus, they have the form

$$\frac{p_{ij}p_{kh}}{p_{ih}p_{kj}}$$

for $0 \leq i < k \leq I - 1$ and $0 \leq j < h \leq J - 1$, see [2, Chapter 2]. In this paper we will consider the cross-product ratio and other ratios naturally defined.

Odds ratios are used in a wide range of applications, and in particular in case-control studies in pharmaceutical and medical research. Following the theory of log-linear models, the statistical inference for the odds ratios is made under asymptotic normality, see for example [3]. More recently, some methods for exact inference have been introduced, see [2] and [1] for details and further references. For the theory about the Bayesian approach, see [13].

From the point of view of Probability and Mathematical Statistics, different descriptions of the geometry of the statistical models for contingency tables are presented in [5, Chapter 2], and in [3, Section 2.7], using vector space theory. An earlier approach to the geometry of contingency tables with fixed cross-product ratio can be found in [7]. In the last few years, the introduction of techniques from Commutative Algebra gave a new flavor to the geometrical representation of statistical models, as shown in, e.g., [15, Chapter 6], [16], [9] and [18].

While the present research was in progress, we became aware of some relevant works in the same direction. Stephen Fienberg

had the idea of examining questions $Q1$ and $Q2$ that we analyze in Section 2. He also presented some results on the characterization of statistical models for 2×2 tables in his lectures slides, publicly posted at <http://www.niss.org/>. Moreover, A. Slavkovic in her Ph.D. dissertation [18] presented many results similar or even coinciding with ours. We remark here that the coinciding results were obtained independently and we acknowledge A. Slavkovic's priority. [18] also contains further investigations which we do not pursue here. For example about general $I \times J$ tables (following and generalizing Fienberg's results) and multi-way tables.

The present paper and [18] mainly differ in the parameterizations used and in the expository style. We use odds ratios while Slavkovic mainly uses conditional and marginal probabilities. Our choice is more adapted to the application to case-control studies discussed in Section 4 which is not given in [18]. More precisely, the use of odds ratios allows a simple representation of relevant parameters such as the Error Odds Ratio and the Diagnostic Odds Ratio which play a key role in case control studies (see Section 4). As far as the exposition is concerned, we decided to be as essential and synthetic as possible following the statement-proof paradigm.

After a careful reading of [18] we also notice that our Proposition 2.6 is essentially [18, Proposition 3.1, page 43]. In order to translate our results in Slavkovic's notation it is enough to apply the well known formulae relating odds ratios and conditional probabilities, see e.g. [18, page 62].

In this paper we use Algebraic and Geometric techniques in order to describe the structure of some models for two-way contingency tables described through odds ratios. The main contributions of this paper to the topic are: a complete and clear description of 2×2 tables having one, two or three fixed odds ratios; the explicit discussion of some of the pathologies appearing in the study of 2×3 tables; the application of the results to case-control studies.

We first consider the case of 2×2 contingency tables of the form (1) with the constraints $p_{ij} > 0$ for all $i, j = 0, 1$ and $p_{00} + p_{01} + p_{10} + p_{11} = 1$. If we allow some probabilities to be zero, notice that the ratios are either zero or undefined. Thus we restrict the analysis to the strictly positive case.

In a 2×2 table we consider the three odds ratios:

$$r_{\times} = \frac{p_{00}p_{11}}{p_{01}p_{10}},$$

$$r_{||} = \frac{p_{00}p_{10}}{p_{01}p_{11}},$$

$$r_{=} = \frac{p_{00}p_{01}}{p_{10}p_{11}}.$$

The meaning of the three odds ratios above will be fully explained in Section 4.

Let $r_{\times} = \alpha^2$, $r_{||} = \beta^2$ and $r_{=} = \gamma^2$. For further use, it is useful to make explicit the following identities. Considering $r_{=}$ and $r_{||}$, it is easy to check that:

$$\beta\gamma = \frac{p_{00}}{p_{11}}, \quad (2)$$

and

$$\frac{\beta}{\gamma} = \frac{p_{01}}{p_{10}}. \quad (3)$$

In Section 2, we study the geometric properties of some statistical models for 2×2 contingency tables. We consider models obtained by fixing two odds ratios, showing that the model is represented by a segment in the probability simplex and studying the behavior of the third ratio. In particular, an expression for tables with three fixed ratios is derived. We also recover classical results about models with a fixed odds ratio. In Section 3, we give a glimpse of the general situation of $I \times J$ contingency tables. We focus our attention on 2×3 tables and we present some of the difficulties arising in the general case. An application to case-control studies is presented in Section 4.

2. Odds Ratios

In this section, we use basic geometric techniques to study the 2×2 tables having two out of the three ratios r_{\times} , $r_{=}$ and $r_{||}$ fixed.

We consider a 2×2 matrix as a point in the real affine 4-space \mathbb{A}^4 . In particular, with the notation of Equation (1), the p_{ij} 's are

coordinates in \mathbb{A}^4 . A 2×2 table is a matrix in the open *probability simplex*

$$\Delta = \left\{ P = (p_{ij}) \in \mathbb{A}^4 : \sum p_{ij} = 1, p_{ij} > 0, i, j = 0, 1 \right\}.$$

As our goal is to describe odds ratios for tables, we may assume the ratios to be non-zero positive numbers.

Fixed the first two ratios

$$r_{\times} = \alpha^2 \quad \text{and} \quad r_{\parallel} = \beta^2,$$

we want to answer the following question:

Q1: How can we describe the locus of tables having the assigned two ratios?

and also

Q2: What are the possible values of the third ratio?

These questions were posed in the AIM computational algebraic statistics plenary lecture by Stephen Fienberg in 2003. In that situation, some interesting comments about treating questions Q1 and Q2 were also made.

Consider the quadratic hypersurfaces of \mathbb{A}^4 :

$$Q_{\alpha} : \alpha^2 p_{01} p_{10} = p_{00} p_{11}$$

and

$$Q_{\beta} : \beta^2 p_{01} p_{11} = p_{00} p_{10}.$$

Notice that a matrix in $Q_{\alpha} \cap Q_{\beta}$ is such that $r_{\times} = \alpha^2$ and $r_{\parallel} = \beta^2$ as soon as the ratios are defined. Hence, to answer the first question, it is enough to study

$$Q_{\alpha} \cap Q_{\beta} \setminus Z,$$

where $Z = \{P = (p_{ij}) \in \mathbb{A}^4 : p_{00} p_{01} p_{10} p_{11} = 0\}$.

We readily see that $Q_{\alpha} \cap Q_{\beta}$ contains the 2-dimensional skew linear spaces

$$p_{00} = p_{01} = 0 \quad \text{and} \quad p_{10} = p_{11} = 0$$

and by general facts on quadrics (see [11, page 301]) we know that there exist two more 2-dimensional skew linear spaces, R and S , such that

$$Q_\alpha \cap Q_\beta = \{p_{00} = p_{01} = 0\} \cup \{p_{10} = p_{11} = 0\} \cup R \cup S.$$

Manipulating equations we notice that a point in $Q_\alpha \cap Q_\beta \setminus Z$ is such that

$$\frac{p_{00}}{p_{01}} = \alpha^2 \frac{p_{10}}{p_{11}} = \beta^2 \frac{p_{11}}{p_{10}}$$

and

$$\frac{p_{10}}{p_{11}} = \beta^2 \frac{p_{01}}{p_{00}} = \frac{1}{\alpha^2} \frac{p_{00}}{p_{01}}.$$

Hence, R and S lie in the intersection of the two 3-dimensional spaces

$$(\alpha p_{10} - \beta p_{11})(\alpha p_{10} + \beta p_{11}) = 0$$

and

$$(\beta p_{01} - \frac{1}{\alpha} p_{00})(\beta p_{01} + \frac{1}{\alpha} p_{00}) = 0,$$

where α and β are chosen to be positive. Only two out of the four resulting 2-dimensional linear spaces lie in both Q_α and Q_β and these are R and S :

$$R : \alpha p_{10} - \beta p_{11} = \beta p_{01} - \frac{1}{\alpha} p_{00} = 0,$$

$$S : \alpha p_{10} + \beta p_{11} = \beta p_{01} + \frac{1}{\alpha} p_{00} = 0,$$

which have parametric presentations

$$R = \{(\beta u, \frac{1}{\alpha} u, \beta v, \alpha v) : u, v \in \mathbb{R}\},$$

$$S = \{(\beta u, -\frac{1}{\alpha} u, \beta v, -\alpha v) : u, v \in \mathbb{R}\}.$$

Summing all these facts up, we get

PROPOSITION 2.1. *Fix the ratios $r_\times = \alpha^2$ and $r_{||} = \beta^2$. Then, a matrix has the given ratios if and only if it has the form*

$$\begin{pmatrix} \beta u & \frac{1}{\alpha} u \\ \beta v & \alpha v \end{pmatrix} \text{ or } \begin{pmatrix} \beta u & -\frac{1}{\alpha} u \\ \beta v & -\alpha v \end{pmatrix}$$

with u, v non-zero real parameters.

Finally, we have to intersect R and S with the probability simplex. As we can choose α and β to be positive, we immediately see that $S \cap \Delta = \emptyset$ (there is always a non-positive coordinate).

To determine $R \cap \Delta$, notice that $R \cap \{\sum p_{ij} = 1\}$ is obtained by taking

$$u = \frac{1 - (\beta + \alpha)v}{\beta + \frac{1}{\alpha}}$$

in the parametric presentation of R . Hence, we get

PROPOSITION 2.2. *Fix the ratios $r_{\times} = \alpha^2$ and $r_{\parallel} = \beta^2$. Then, a table has the given ratios if and only if it has the form*

$$\begin{pmatrix} \frac{\beta}{\beta + \frac{1}{\alpha}}[1 - (\beta + \alpha)v] & \frac{1}{\alpha\beta + 1}[1 - (\beta + \alpha)v] \\ \beta v & \alpha v \end{pmatrix}$$

where $0 < v < \frac{1}{\alpha + \beta}$.

This answers question Q1: fixed the two ratios, the tables with those ratios describe a segment in the probability simplex.

REMARK 2.3. *In [3, Section 2.7], a parametric description of the tables with $r_{\times} = 1$ is written in the form*

$$\begin{pmatrix} st & s(1-t) \\ (1-s)t & (1-s)(1-t) \end{pmatrix}. \quad (4)$$

Let us check that our parametrization contains this as a special case. In order to do this, we will compute the marginal sums

$$\left(\begin{array}{cc|c} \beta[\frac{1}{\beta+1} - v] & [\frac{1}{\beta+1} - v] & 1 - (\beta + 1)v \\ \beta v & v & (\beta + 1)v \\ \hline \frac{\beta}{\beta+1} & \frac{1}{\beta+1} & 1 \end{array} \right). \quad (5)$$

Hence, the parametrizations in Equations (4) and (5) are just the same, simply let $t = \frac{\beta}{\beta+1}$ and $s = 1 - (\beta + 1)v$.

REMARK 2.4. *Suppose to fix r_{\times} and to ask for a geometric description of the locus of tables with this ratio. Using Proposition 2.2 we can easily get an answer. For each value of r_{\parallel} we get a segment of*

tables, and making $r_{||}$ to vary this segment describes a portion of a ruled quadratic surface. Notice that, for $r_{\times} = 1$, this is the result contained in [3, Section 2.7]. In particular, we recall that matrices such that r_{\times} is fixed form a so called Segre variety (i.e., in this case, a smooth quadric surface in the projective three space). For more on this see, e.g., [8].

Answering question Q2 is just a computation, and we see that

$$r_{=} = \frac{1}{\alpha\beta + 1} \frac{[1 - (\beta + \alpha)v]^2}{v},$$

where $r_{\times} = \alpha^2$ and $r_{||} = \beta^2$. Notice that, fixed r_{\times} and $r_{||}$, the third ratio can freely vary in $(0, +\infty)$.

REMARK 2.5. We expressed $r_{=}$ as a function $r_{=}(\alpha, \beta, v)$, and standard computations show that this is an invertible function of v . In particular, we get

$$v = \frac{1}{\alpha + \beta + \sqrt{(\alpha\beta + 1)r_{=}}}.$$

Thus, given $r_{\times} = \alpha^2$, $r_{||} = \beta^2$ and $r_{=}$, we have an explicit description of the unique table with these ratios (use Proposition 2.2).

Clearly, completely analogous results hold if we fix the ratios r_{\times} and $r_{=}$.

If we fix the ratios $r_{||} = \beta^2$, $r_{=} = \gamma^2$ and we argue as above, we get the following:

PROPOSITION 2.6. Fix the ratios $r_{||} = \beta^2$ and $r_{=} = \gamma^2$. Then, a table has the given ratios if and only if it has the form

$$\begin{pmatrix} \frac{\beta}{\beta + \frac{1}{\gamma}} [1 - (\beta + \gamma)v] & \gamma v \\ \beta v & \frac{1}{\beta\gamma + 1} [1 - (\beta + \gamma)v] \end{pmatrix}$$

where $0 < v < \frac{1}{\beta + \gamma}$.

Again, a trivial computation yields:

$$r_{\times} = \left(\frac{\beta}{\beta\gamma + 1} \right)^2 \frac{[1 - (\beta + \gamma)v]^2}{v^2},$$

and hence, fixed $r_{=}$ and $r_{||}$, the third ratio can freely vary in $(0, +\infty)$, see Remark 2.5.

REMARK 2.7. *As mentioned in the Introduction, the above results can also be found in [18, Chapter 3], but expressed in terms of conditional and marginal distributions. Both representations allow the graphical visualization of the variety in the affine three-space. Another approach to get this visualization is presented in [14].*

3. The 2×3 case

The study of tables with more than two rows and columns would be of great interest, but the complexity of the problem readily increases as we show in the 2×3 case.

Consider the 2×3 matrix

$$\begin{pmatrix} p_{00} & p_{01} & p_{02} \\ p_{10} & p_{11} & p_{12} \end{pmatrix}$$

and define odds ratios as above for each 2×2 submatrix. We complete our previous notation by adding a superscript to denote the deleted column, e.g.

$$r_{\underline{=}}^{(1)} = \frac{p_{00}p_{02}}{p_{10}p_{12}}.$$

Again, we consider a matrix as a point in a real affine space, in this case \mathbb{A}^6 . Notice that the ratios are well defined for matrices in $\mathbb{A}^6 \setminus Z$, where Z denotes the set of matrices with at least a zero entry.

Relations among the ratios are the cause of the increased complexity of the higher dimensional cases. For example, as we will see, two of the ratios can always be freely fixed. But, as soon as three ratios are considered, constraints come in the picture.

Easy calculations show that the following relations hold:

$$r_{\parallel}^{(0)} r_{\parallel}^{(2)} = r_{\parallel}^{(1)},$$

$$r_{\times}^{(0)} r_{\times}^{(2)} = r_{\times}^{(1)}$$

and also

$$r_{\times}^{(0)} = r_{\underline{=}}^{(2)} (r_{\underline{=}}^{(1)})^{-1},$$

$$r_{\times}^{(1)} = r_{\underline{=}}^{(2)} (r_{\underline{=}}^{(0)})^{-1},$$

$$r_{\times}^{(2)} = r_{\underline{=}}^{(1)} (r_{\underline{=}}^{(0)})^{-1}.$$

These relations, beside producing constraints on the numerical choice of the ratios, lead to a much more complex geometric situation. We illustrate this by exhibiting some explicit examples (worked out with the Computer Algebra systems **Singular** and **CoCoA**). As references for the software, see [4] and [10].

More precisely, we fix some of the ratios and we describe the locus of matrices satisfying these relations in

$$\Sigma^\circ = \{P = (p_{ij}) \in \mathbb{A}^6 : \sum p_{ij} = 1\} \setminus Z,$$

i.e. the space of matrices with non-null entries of sum one. For the sake of simplicity, we do not consider the positivity conditions defining the simplex.

In our geometric descriptions, we will slightly abuse terminology, e.g. we will call a line in Σ° a line in \mathbb{A}^6 not contained in Z ; notice that our lines may have some holes (i.e. the points of intersection with Z).

We start by considering the easiest case where two of the ratios are fixed. Already at this stage, a dichotomy arises and we have two different situations, as shown in the following examples:

$$r_{\times}^{(1)} = r_{\times}^{(2)} = 1, \quad (6)$$

$$r_{\times}^{(1)} = r_{\equiv}^{(2)} = 1 \text{ or } r_{\equiv}^{(1)} = r_{\parallel}^{(2)} = 1 \text{ or } r_{\parallel}^{(1)} = r_{\parallel}^{(2)} = 1 \text{ or } r_{\equiv}^{(1)} = r_{\equiv}^{(2)} = 1 \quad (7)$$

The locus of matrices in Σ° satisfying one of conditions (7) is a 3-dimensional variety of degree 4, while condition (6) describes a 3-dimensional variety of degree 3. Roughly speaking, the degree (see [11, page 16] and [17, page 41]) is a measure of the complexity of the variety. For a surface in 3-space, for example, the degree bounds the number of intersections with a line and, in a certain sense, measures how the surface is folded.

Next, we try to fix three of the ratios, for example:

$$r_{\times}^{(0)} = r_{\equiv}^{(1)} = r_{\parallel}^{(2)} = 1 \text{ or } r_{\times}^{(0)} = r_{\times}^{(1)} = r_{\parallel}^{(2)} = 1, \quad (8)$$

$$r_{\times}^{(0)} = r_{\times}^{(1)} = r_{\equiv}^{(2)} = 1, \quad (9)$$

$$r_{\times}^{(0)} = 4, r_{\times}^{(1)} = 3, r_{\equiv}^{(2)} = 2. \quad (10)$$

The locus of matrices in Σ° satisfying one of conditions (8) is the union of two quadratic surfaces, while condition (9) gives a plane. Moreover, if we consider the same ratios but we vary their values, as in (10), the locus of matrices is now described by a single quadratic surface.

Finally, a glimpse of the case of four fixed ratios:

$$r_{\times}^{(0)} = r_{\times}^{(1)} = r_{\parallel}^{(1)} = r_{\parallel}^{(2)} = 1, \tag{11}$$

$$r_{\times}^{(0)} = r_{\times}^{(1)} = r_{\underline{=}}^{(1)} = r_{\underline{=}}^{(2)} = 1, \tag{12}$$

In both cases, the locus is described by a curve as expected. But, condition (11) produces the union of four lines, while condition (12) is satisfied by a single line in Σ° .

The Computer Algebra systems **Singular** and **CoCoA** were used to compute primary decompositions (giving the irreducible components of the loci) and Hilbert functions (giving the dimension and the degree of the loci).

4. An application. The case-control studies

Two-by-two contingency tables are natural models for a large class of problems known, in medical literature, as case-control studies. Let us consider a table coming, e.g., from the study of a new pharmaceutical product, or clinical test, designed for the detection of a disease. This is an example of a case-control study.

In a case-control study there are two random variables. The first variable X_1 encodes the presence (level 1) or absence (level 0) of the disease. The second variable X_2 encodes the result of the clinical test (level 1 if positive, level 0 if negative).

The joint variable (X_1, X_2) has 4 outcomes, namely:

$$(0, 0), (0, 1), (1, 0), (1, 1).$$

Its probabilities form a 2×2 contingency table:

$$\begin{pmatrix} p_{00} & p_{01} \\ p_{10} & p_{11} \end{pmatrix}.$$

The probabilities p_{00} and p_{11} are called the probability of true negative and of true positive, respectively. They correspond to the cases of correct answer of the clinical test. The probabilities p_{10} and p_{01} are called the probability of false positive and of false negative, respectively. They correspond to the two types of error which can show in a case-control study. For example, the probability of false negative is the probability that a diseased subject is incorrectly classified as not diseased.

A perfect clinical test which correctly classifies all the subjects would have p_{01} and p_{10} as low as possible, implying a large value of the odds ratio r_{\times} . Therefore, the odds ratio r_{\times} measures the validity of the clinical test. In particular, when $r_{\times} = 1$, the random variables are statistically independent. In our framework this means that, when $r_{\times} = 1$, the result of the clinical test is independent from the presence or absence of the disease. Unless one obtains a large value of r_{\times} , the clinical test is judged as non efficient. The odds ratio r_{\times} is also called Diagnostic Odds Ratio (DOR) in medical literature.

In such a case-control study, two essential indices are the specificity and the sensitivity, defined as:

$$\text{specificity} = \frac{p_{00}}{p_{00} + p_{01}}$$

and

$$\text{sensitivity} = \frac{p_{11}}{p_{10} + p_{11}}.$$

Specificity is the proportion of true negative among the diseased subjects, while sensitivity is the proportion of true positive among the non-diseased subjects.

Straightforward computations show that

$$r_{\times} = \frac{\text{specificity}/(1 - \text{specificity})}{(1 - \text{sensitivity})/\text{sensitivity}}.$$

In view of the definition above, it is easy to show that the relative magnitude of the sensitivity and specificity is measured by the odds ratio r_{\parallel} . In fact one can show that

$$\frac{\text{sensitivity}/(1 - \text{sensitivity})}{\text{specificity}/(1 - \text{specificity})} = \frac{1}{r_{\parallel}}.$$

The ratio above is called Error Odds Ratio (EOR).

In recent literature, the DOR and the EOR are relevant parameters for the assessment of the validity of a clinical test. They have received increasing attention in the last few years and a huge amount of literature has been produced. Hence, we refrain from any tentative description and refer the interested reader to, for example, [12].

The meaning of the third ratio $r_{=}$ is not straightforward as explained in [3, Page 21]. However its statistical meaning can be derived using Equations (2) and (3) shown in Section 1.

Finally, we remark that the geometrical structure of the statistical models for case-control studies is very simple. From the results in Section 2, one readily sees that the models are segments or portions of ruled quadratic surfaces. Moreover, from a Bayesian point of view, Propositions 2.2 and 2.6 allow to compute the exact range of the free odds ratio.

REFERENCES

- [1] A. AGRESTI, *Exact inference for categorical data: Recent advances and continuing controversies*, Statist. Med. **20** (2001), 2709–2722.
- [2] A. AGRESTI, *Categorical data analysis*, 2 ed., Wiley, New York, 2002.
- [3] Y.M. BISHOP, S. FIENBERG AND P.W. HOLLAND, *Discrete multivariate analysis: theory and practice*, MIT Press, Cambridge, 1975.
- [4] COCoATEAM, *CoCoA: a system for doing Computations in Commutative Algebra*, Available at <http://cocoa.dima.unige.it>, 2004.
- [5] D. COLLOMBIER, *Récherches sur l'analyse des tables de contingence*, PhD thesis, Université Paul Sabatier de Toulouse, Julliet 1980.
- [6] S. FIENBERG, *The analysis of cross-classified categorical data*, MIT Press, Cambridge, 1980.
- [7] S.E. FIENBERG AND J.P. GILBERT, *The geometry of a two by two contingency table*, J. Amer. Statist. Assoc. **65** (1970), 694–701.
- [8] L. GARCIA, M. STILLMAN AND B. STURMFELS, *Algebraic geometry of Bayesian networks*, J. Symb. Comput. **39** (2005), 331–355.
- [9] D. GEIGER, D. HECKERMAN, H. KING AND C. MEEK, *Stratified exponential families: Graphical models and model selection*, Ann. Statist. **29** (2001), no. 3, 505–529.
- [10] G.-M. GREUEL, G. PFISTER, AND H. SCHÖNEMANN, *SINGULAR 2.0*, A Computer Algebra System for Polynomial Computations, Centre for Computer Algebra, University of Kaiserslautern, 2001, <http://www.singular.uni-kl.de>.

- [11] J. HARRIS, *Algebraic geometry, a first course*, Graduate Texts in Math., Springer-Verlag, New York, 1992.
- [12] A. KNOTTNERUS, *The evidence base of clinical diagnosis*, British Medical Association, Maastricht, The Netherlands, 2001.
- [13] D.V. LINDLEY, *The bayesian analysis of contingency tables*, Ann. Math. Statist. **35** (1964), 1622–1643.
- [14] D. LUO, G. R. WOOD AND G. JONES, *Visualising contingency tables data*, Austral. Math. Soc. Gaz. **31** (2004), no. 4, 258–262.
- [15] G. PISTONE, E. RICCOMAGNO AND H.P. WYNN, *Algebraic statistics: Computational commutative algebra in statistics*, Chapman & Hall/CRC, Boca Raton, 2001.
- [16] G. PISTONE, E. RICCOMAGNO AND H.P. WYNN, *Computational commutative algebra in discrete statistics*, Algebraic Methods in Statistics and Probability (Marlos A. G. Viana and Donald St. P. Richards, eds.), Contemporary Mathematics, vol. 287, American Mathematical Society, 2001, pp. 267–282.
- [17] I. SHAFAREVICH, *Basic algebraic geometry*, 2nd ed., Springer Verlag, Berlin, 1995.
- [18] A.B. SLAVKOVIC, *Statistical disclosure limitation beyond the margins: Characterization of joint distributions for contingency tables*, Phd thesis, Carnegie Mellon University, August 2004.

Received December 6, 2005.