A Fourfold Display for 2 by 2 by k Tables

Michael Friendly

Psychology Department
York University

January 24, 1995
Report Number: 217

Abstract

This paper describes the fourfold display, a graphic designed to display the frequencies in a $2 \times 2 \times k$ contingency table. In this display the frequency in each cell is shown by a quarter circle, whose area is proportional to the cell count in a given 2×2 layer, in a way that depicts the odds ratios for k strata. Confidence rings for the odds ratio can be superimposed to provide a visual test of the hypothesis of no association in each stratum.

Résumé

Cet article décrit un visuel quadruple, une méthode graphique conçu pour visualiser les fréquences dans une table de contingence $2 \times 2 \times k$. Dans ce visuel, la fréquence de chaque cellule est representée par un quart de cercle, qui a une superficie proportionnelle au compte de la cellule dans un couche 2×2 donée, afin de décrire les rapports de chance pour k strates. Des anneaux qui démonstrant la confiance des rapports de chance peuvent être superposés afin de fournir un test visuel de l'hypothèse d'aucune association dans chaque strate.

Keywords: Categorical data; Fourfold display; Graphical methods; Mosaic display; Parquet diagram

AMS classification: 65S05

Introduction

For the display of proportions, frequencies or percentages the pie chart has been highly favored in popular media and presentation graphics nearly since its invention by Playfair in 1786. Among statisticians and those concerned with statistical graphics, however, the pie chart is generally deprecated in favor of the bar chart, usually on the belief that judgments of areas or angles are more difficult or error-prone than judgments of length (e.g., MacDonald-Ross, 1977; Tufte, 1983; Wainer & Thissen, 1981). In Tufte's (1983) view, for example, "the only design worse than a pie chart is several of them ... pie charts should never be used". Although a critical review of the research on graphical perception by Spence and Lewandowski (1990) concluded that "this prejudice against the pie chart is unfounded", the prevailing views and the results of influential studies by Cleveland & McGill (1984, 1986) continue to lead statisticians to hold the pie chart in disfavor. Can the reputation of the pie chart in statistical graphics be salvaged?

This paper describes a relative of the pie chart, termed the "fourfold display" designed for the display of $2 \times 2 \times k$ tables. Several versions of this display (called a floating four-fold circular display) appear in Fienberg (1975), Wainer & Reiser (1976) and Bertin (1983, p.53). In this display the frequency n_{ij} in each cell of a fourfold table is shown by a quarter circle, whose radius is proportional to $\sqrt{n_{ij}}$, so the area is proportional to the cell count. This use of circular wedges which vary in radii to depict parts of a whole has a history which goes back to the "coxcomb" (or rose chart) used by Florence Nightingale (1858; see Funkhauser, 1937, p. 344) to display deaths due to preventable causes in hospitals among British soldiers in the Crimean War.

For a single 2×2 table the fourfold display described here also shows the frequencies by area, but scaled in a way that depicts the sample odds ratio, $\hat{\theta} = (n_{11} / n_{12}) / (n_{21} / n_{22})$. An association between the variables ($\theta \neq 1$) is shown by the tendency of diagonally opposite cells in one direction to differ in size from those in the opposite direction, and the display uses color or shading to show this direction. Confidence rings for the observed θ allow a visual test of the hypothesis H_0: $\theta = 1$. They have the property that the rings for adjacent quadrants overlap iff the observed counts are consistent with the null hypothesis.

In the case of a $2 \times 2 \times k$ table, the last dimension typically corresponds to "strata" or populations, and it is typically of interest to see if the association between the first two variables is homogeneous across strata. The fourfold display is designed allow easy visual comparison of the pattern of association between two dichotomous variables across two or more populations.

The next section presents an example to make the graphical ideas concrete. The few details which are not obvious from the example are presented in the following section. A final section compares the fourfold display to other graphs suggested recently for categorical data.

An example: Berkeley admissions data

Figure 1 shows aggregate data on applicants to graduate school at Berkeley for the six largest departments in 1973 classified by admission and sex. At issue is whether the data show evidence of sex bias in admission practices (Bickel et al., 1975). The figure shows the cell frequencies numerically in the corners of the display. Thus there were 2691 male applicants, of whom 1193 (44.4%) were admitted, compared with 1855 female applicants of whom 557 (30.0%) were admitted. Hence the sample odds ratio, Odds (Admit|Male) / (Admit|Female) is 1.84 indicating that males were almost twice as likely to be admitted.

The frequencies displayed graphically by shaded quadrants in Figure 1 are not the raw frequencies. Instead, the frequencies have been standardized (by iterative proportional fitting) so that all table margins are equal, while preserving the odds ratio. Each quarter circle is then drawn to have an area proportional to this

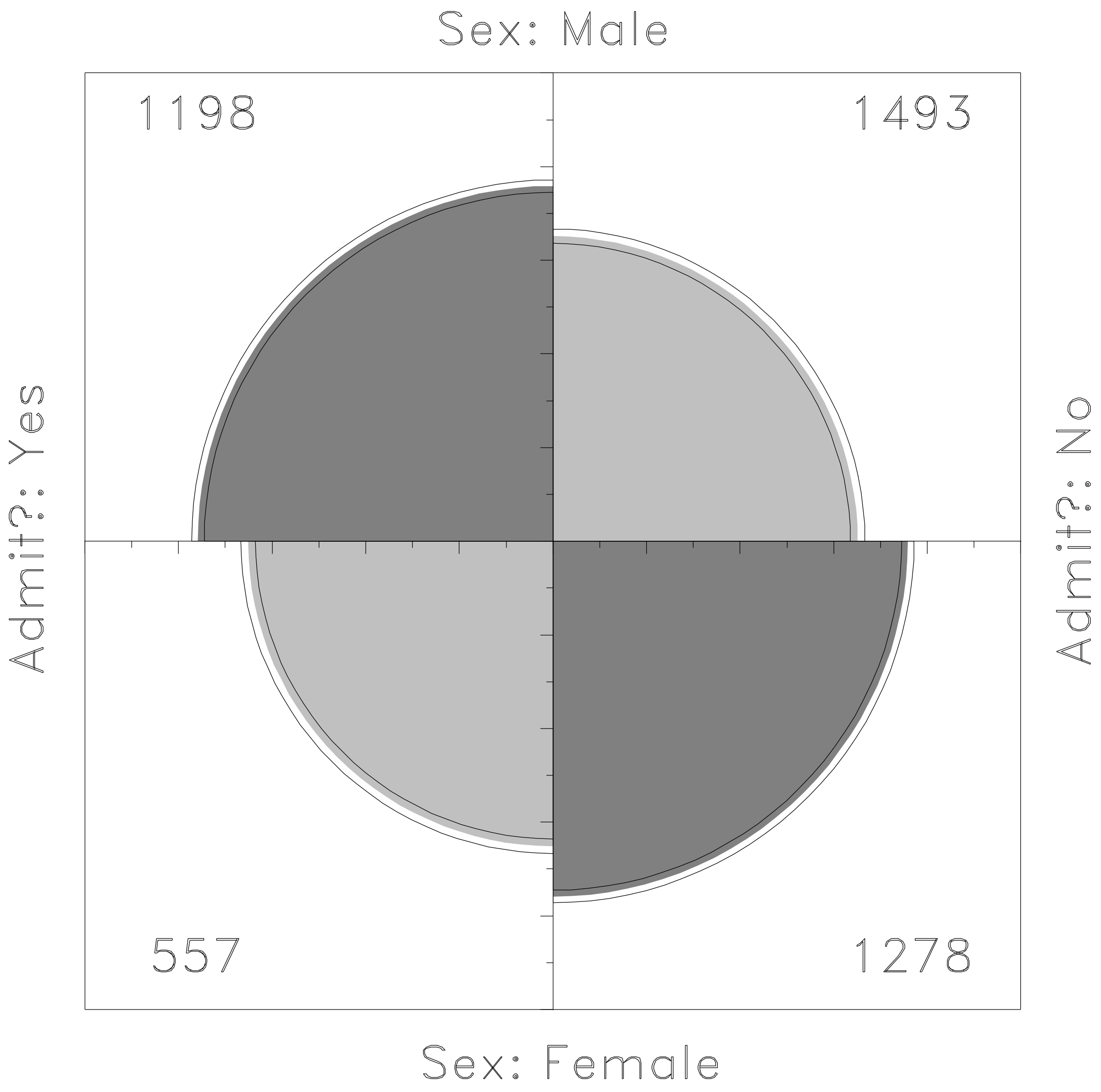

Figure 1: Fourfold display for Berkeley admissions data: Evidence for sex bias?. The area of each shaded quadrant shows the frequency, standardized to equate the margins for sex and admission. Circular arcs show the limits of a 99% confidence interval for the odds ratio.

standardized cell frequency. This makes it easier to see the association between admission and sex without being influenced by the overall admission rate or the differential tendency of males and females to apply. With this standardization the four quadrants will align when the odds ratio is 1, regardless of the marginal frequencies.

The shaded quadrants in Figure 1 do not align and the 99% confidence rings around each quadrant do not overlap, indicating that the odds ratio differs significantly from 1. The width of the confidence rings gives a visual indication of the precision of the data.

Both the pie chart and the fourfold display depict frequency by area. However, the pie chart varies the angle of each slice (keeping the radius fixed), so that it is hard to compare a given slice whose location varies across two or more pies (which prompted Tufte's negative comment). The fourfold display holds the angles and locations constant (varying the radius), and so allows alignment to be judged along the horizontal and vertical scales. Note that while frequency is displayed by area (so that the odds ratio corresponds visually to the ratio of areas of adjacent quadrants) the radii in the fourfold display depict frequency on a square root scale, so that differences along the linear scale are deemphasized. The effect is that adjacent quadrants appear to differ when there is a true difference in the odds; the confidence rings make this appearance precise.

The admissions data shown in Figure 1 were obtained from a sample of six departments, so to determine the source of the apparent sex bias in favor of males, we make a new plot, Figure 2, stratified by department. The departments are labelled so that the overall acceptance rate is highest for Department A and decreases steadily to Department F. Again each panel is standardized to equate the marginals for sex and admission. This standardization also equates for the differential total applicants across departments, facilitating visual comparison.

Surprisingly, Figure 2 shows that, for five of the six departments, the odds of admission is approximately the same for both men and women applicants. Department A appears to differs from the others, with women approximately 2.86 ($= (313/19) \div (512/89)$) times as likely to gain admission. This appearance is confirmed by the confidence rings, which in Figure 2 are joint 99% intervals for θ_c, $c = 1, \ldots, k$.

This result, which contradicts the display for the aggregate data in Figure 1, is a classic example of Simpson's paradox. The resolution of this contradiction can be found in the large differences in admission rates among departments. Men and women apply to different departments differentially, and in these data women apply in larger numbers to departments that have a low acceptance rate. The aggregate results are misleading because they falsely assume men and women are equally likely to apply in each field. (This explanation ignores the possibility of structural bias against women, e.g., lack of resources allocated to departments that attract women applicants.)

The differential rates of application by men and women across departments can be seen in Figure 3, a fourfold display of the data standardized to equate the margins for admission, but not sex. Note that the number of female applicants is quite small in Departments A and B where the admissions rate is relatively higher, and greater in Departments C and E with lower admissions rate. Also note that, with this standardization, the confidence rings overlap across the vertical axis only, when the odds of admission are the same for males and females.

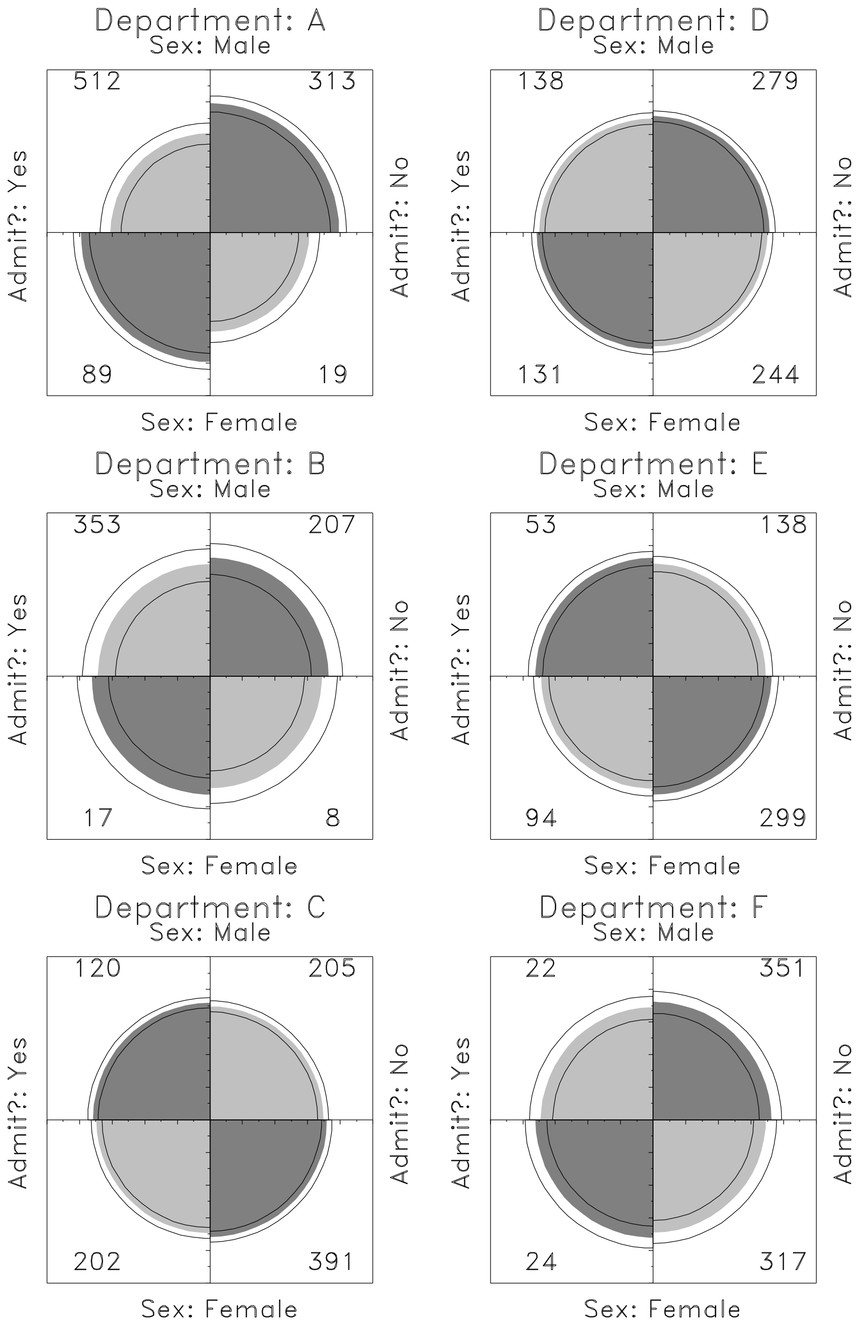

Figure 2: Fourfold display of Berkeley admissions, by department. In each panel the confidence rings for adjacent quadrants overlap if the odds ratio for admission and sex does not differ significantly from 1. The data in each panel have been standardized as in Figure 1.

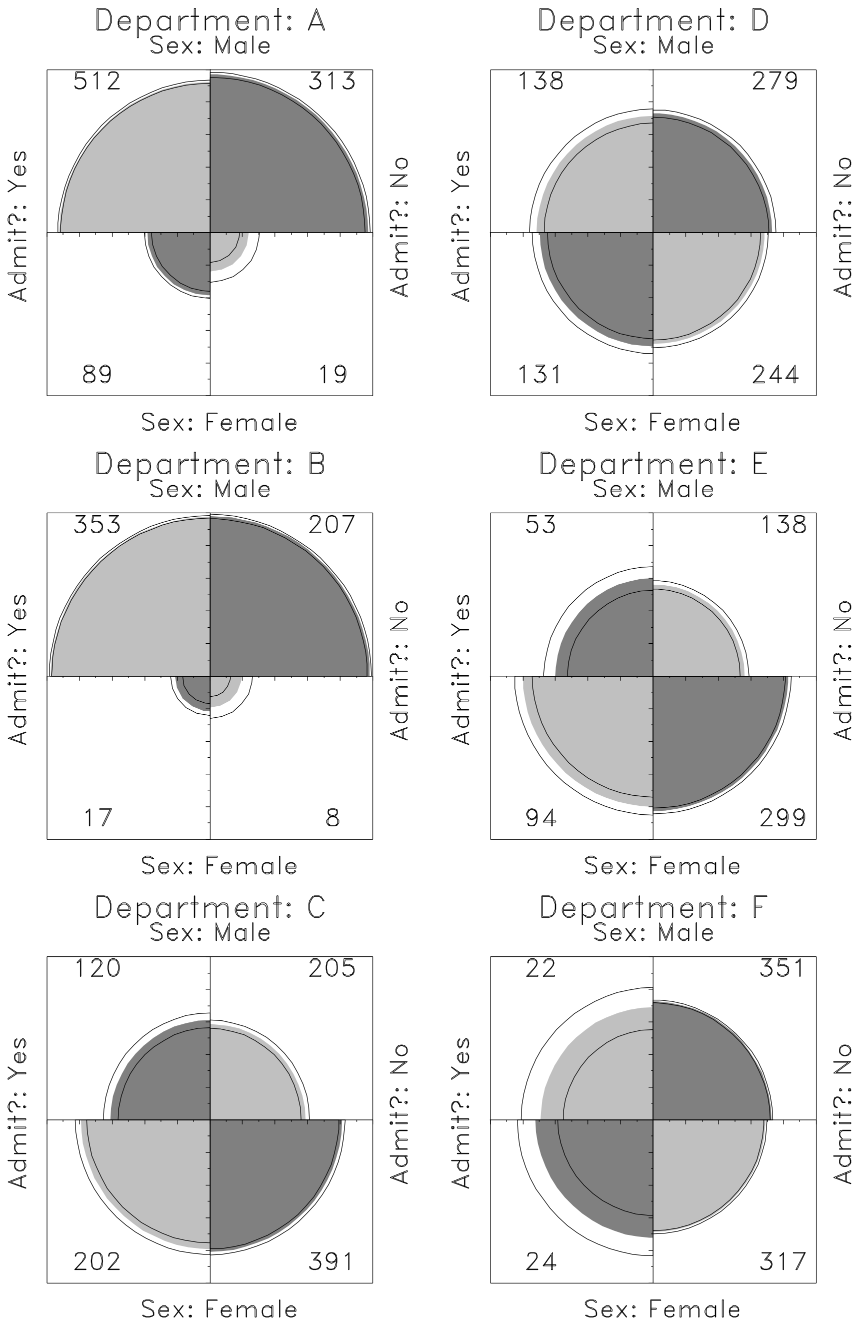

Figure 3: Fourfold display of Berkeley admissions, by department. The data in each panel have been standardized to equate the margins for admissions, but not for sex.

Details

The construction details of the fourfold display are straight forward and need no further elaboration here. A program written using SAS/IML software (SAS Institute, 1989) has been described by Friendly (1994b). This program (`fourfold.sas`) and examples of its use (`fourdemo.sas`) are available by anonymous FTP transfer from the site FTP.SAS.COM. Login as user "anonymous", change to the proper directory with the command `cd /observations/3q94/friendly`, and use the `get` command to retrieve each of the files.

Confidence rings

Confidence rings for the fourfold display are computed from a confidence interval for θ, whose endpoints can each be mapped into a 2×2 table. Each such table is then drawn in the same way as the data.

The interval for θ is most easily found by considering the distribution of $\hat{\psi} = \ln \hat{\theta}$, whose sampling variance may be estimated by $\hat{s}^2(\hat{\psi}) = \Sigma\Sigma\, n_{ij}^{-1}$. Then an approximate $1 - \alpha$ confidence interval for ψ is given by $\hat{\psi} \pm \hat{s}(\hat{\psi})\; z_{1-\alpha/2} = \{\hat{\psi}_l, \hat{\psi}_u\}$, and the corresponding limits for the odds ratio θ are $\{\exp(\hat{\psi}_l), \exp(\hat{\psi}_u)\}$. Thus, for the aggregate data shown in Figure 1, $\hat{\psi} = \ln \hat{\theta} = .6104$, and $\hat{s}(\hat{\psi}) = 0.0639$, so the 99% limits for θ are $\{1.5617, 2.1704\}$.

Now consider how to find a 2×2 table whose frequencies correspond to the odds ratios at the limits of the confidence interval. A table standardized to equal row and column margins can be represented by the 2×2 matrix with entries $\begin{bmatrix} p & (1-p) \\ (1-p) & p \end{bmatrix}$, whose odds ratio is $\theta = p^2 / (1-p)^2$, so $p = \sqrt{\theta} / (1 + \sqrt{\theta})$. The corresponding frequencies can then be found by adjusting the standardized table to have the same row and column margins as the data. The results of these computations which generate the confidence rings in Figure 1 are shown in Table 1.

Table 1: Odds ratios and equivalent tables for confidence rings

	Odds Ratio	Standardized Table		Frequencies	
Lower	1.562	0.555	0.445	1157.2	1533.8
		0.445	0.555	597.8	1237.2
Data	1.841	0.576	0.424	1198.	1493.
		0.424	0.576	557.	1278.
Upper	2.170	0.596	0.404	1237.8	1453.2
		0.404	0.596	517.2	1317.8

Other displays for contingency tables

It is of interest to compare the fourfold display with two other graphical methods recently suggested for contingency tables.

The effects described earlier can be seen somewhat differently in Figure 4, a mosaic display presented by Friendly (1994a). In this display observed frequencies are shown by the areas of tiles, while shading is used to show the residuals from a given log-linear model. Figure 4 shows the results of fitting the log-linear model [AdmitDept] [SexDept] which asserts that admission and sex are conditionally independent (that is, no sex-bias), given department. Cells with absolute adjusted residuals less than 2.0 are unfilled, while cells whose adjusted residuals exceed 2.0 in absolute value are shaded, with cross-hatching for positive residuals and grayscale for negative residuals.

The four large blocks corresponding to admission by sex show the greater overall acceptance of males than females. Among admitted applicants, however, there are larger proportions of women in the departments (C-F) with low admission rates. The lack of fit of model [AD] [GD] (likelihood ratio $G^2 = 20.2$ with 5 df) is concentrated entirely in Department A, where a greater proportion of females is admitted.

A final display of these data appears in Figure 5, called a parquet diagram by Riedwyl & Schuepbach (1994). This graph is strictly applicable to two-way tables, so in the Berkeley data admission and sex were combined to form a single variable. In the parquet diagram rectangles are drawn whose dimensions are proportional to the marginal frequencies, so the area of each rectangle is proportional to expected frequency in that cell; observed frequency is shown by the number of squares in each rectangle. Hence, the difference between observed and expected frequency appears as variations in the density of shading, using color to indicate whether the deviation from independence is positive or negative. (In monochrome versions, positive deviations are shown by solid lines, negative by broken lines.)

In this display the widths of the columns show the greater number of male applicants than female; the greater overall admission rate for males can be seen by comparing the ratio od widths (M:Yes / M:No) to that of (F:Yes / F:No). Cells with many small squares correspond to those whose observed frequencies are greater than expected under independence. Figure 5 shows the greater numbers of male applicants in departments A and B and greater numbers of female applicants in the remaining departments.

Note that Figure 2 to Figure 5 all use the graphic elements of area and visual density to depict cell frequencies in a contingency table. A conceptual rationale for this graphic representation is provided by a physical model for categorical data (Friendly, 1993) which likens observations to gas molecules in a pressure chamber wherein fitting a statistical model by maximum likelihood corresponds to mimimizing energy or balancing of forces.

Statistical graphs, like other forms of communication, serve different purposes, the most common of which are summarization and exposure. The fourfold display is tuned to depict the odds ratios in 2×2 tables so that these values can be readily compared across strata; the doubly standardized version in Figure 2 is more effective for this purpose, but conceals differences in the marginal frequencies which are shown in Figure 3. The mosaic display is designed as a data-analytic graph to show both counts (by area) and the pattern of departure from a specified log-linear model (by visual density). It applies to tables of any number of dimensions whereas the fourfold display and parquet diagram do not generalize readily. For simple two-way tables, however, the display of frequencies by density in the parquet diagram is more direct.

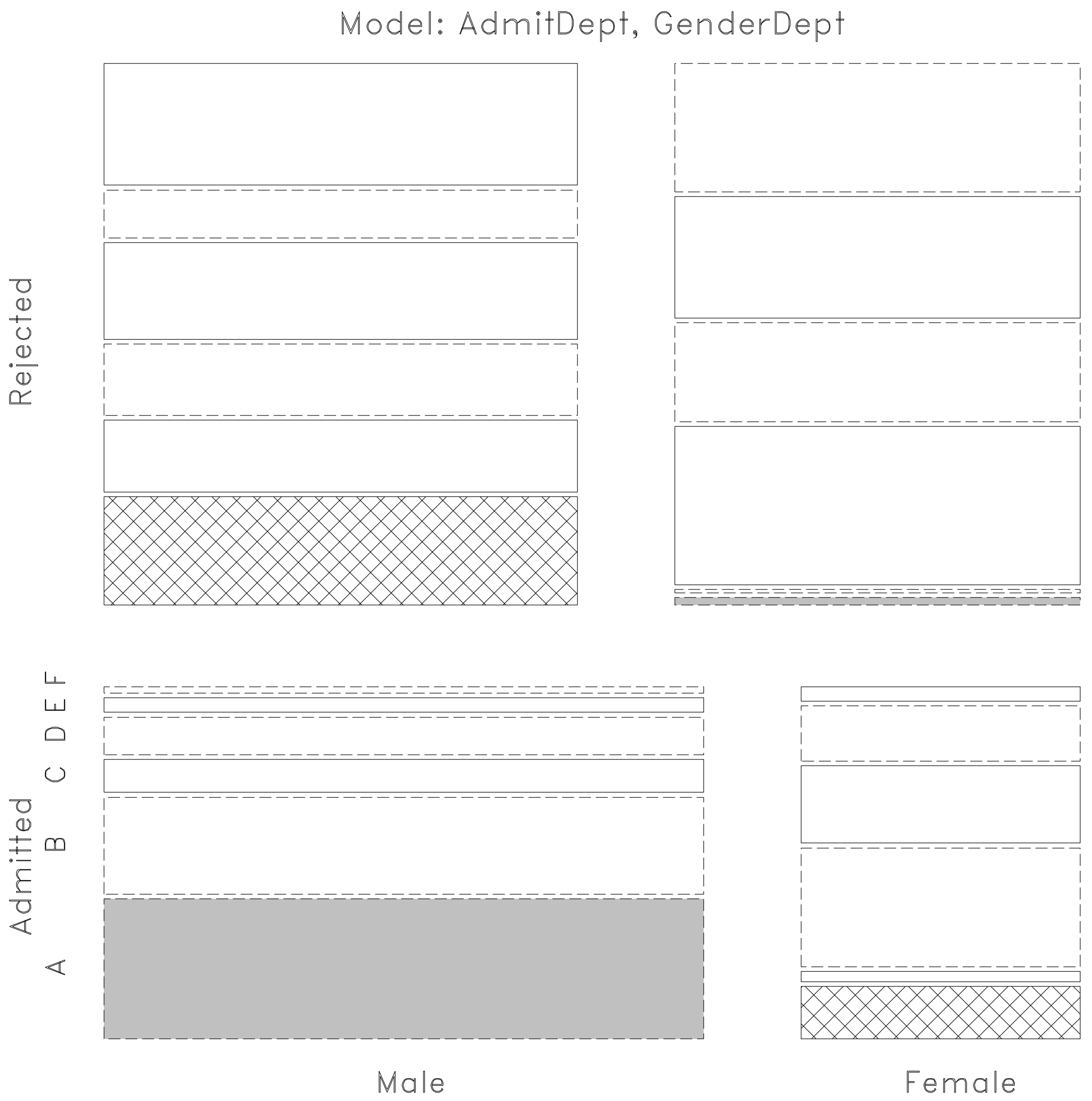

Figure 4: Mosaic display of Berkeley admissions data. The area of each tile is proportional to cell frequency. Shading and line style show adjusted residuals from the log-linear model [AdmitDept] [SexDept] which asserts that admission and sex are conditionally independent, given department.

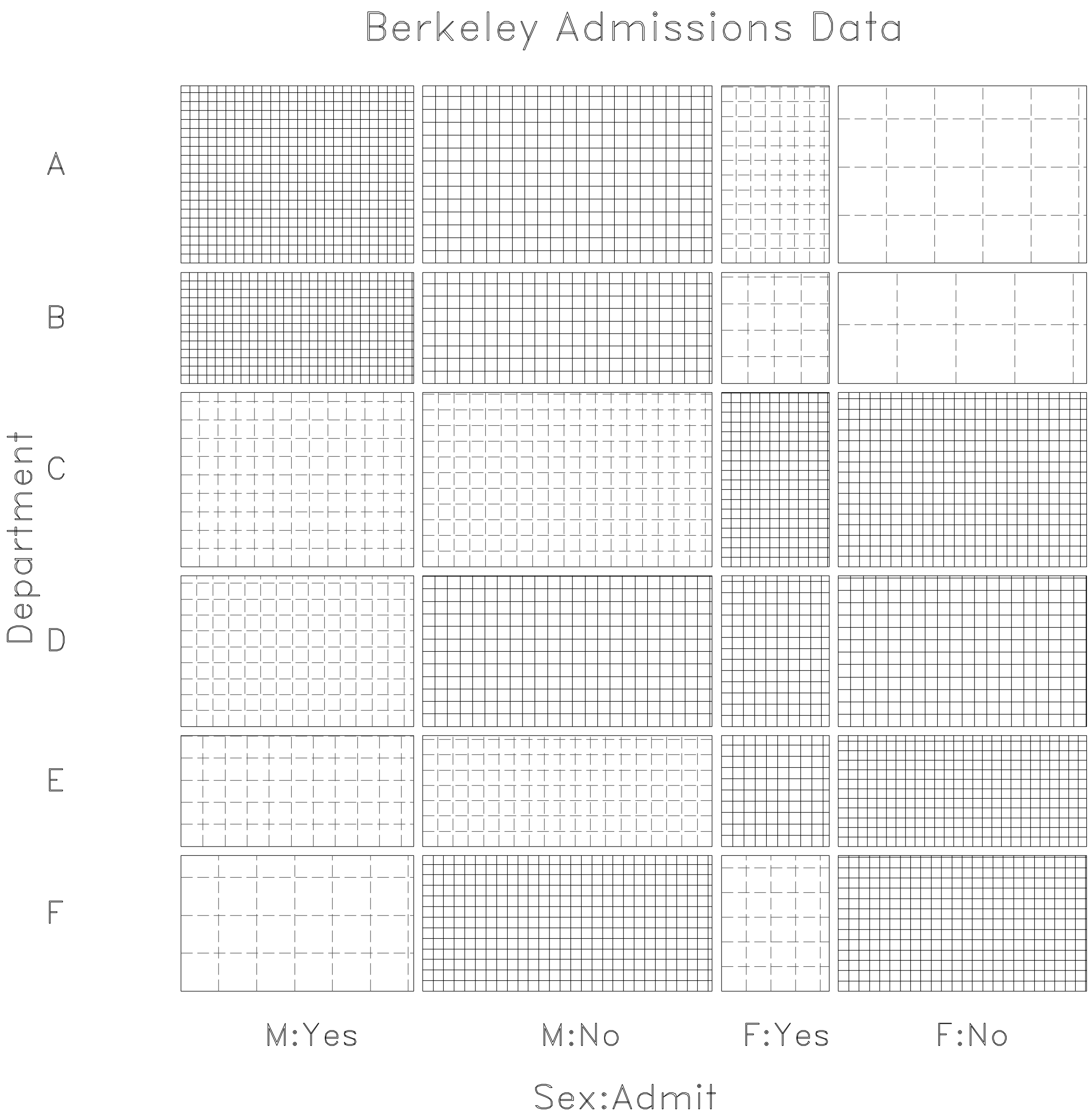

Figure 5: Parquet diagram. In this display observed counts are shown by the number of squares within each rectangle. The area of each rectangle is proportional to the expected frequency under independence of departments and the categories of sex and admission, treated jointly.

Acknowledgments

This work was supported by Grant OGP0138748 from the National Science and Engineering Research Council of Canada. I am grateful to Howard Wainer for helpful comments.

References

Agresti, A. (1990), *Categorical Data Analysis*, New York: Wiley-Interscience.

Bertin, J. (1983). *Semiology of Graphs* Madison, WI: University of Wisconsin Press.

Bickel, P. J., Hammel, J. W. & O'Connell, J. W. (1975). Sex bias in graduate admissions: data from Berkeley. *Science*, *187*, 398-403.

Cleveland, W. S., and McGill, R. (1984). Graphical perception: Theory, experimentation and application to the development of graphical methods. *Journal of the American Statistical Association*, *79*, 531-554.

Cleveland, W. S., and McGill, R. (1986). An experiment in graphical perception. *International Journal of Man-Machine Studies*, *25*, 491-500.

Fienberg, S. E. (1975). Perspective Canada as a social report. *Social Indicators Research*, *2*, 153-174.

Friendly, M. (1992). Graphical methods for categorical data, *Proceedings of the SAS User's Group International Conference*, *17*, 1367-1373.

Friendly, M. (1993). Conceptual and visual models for categorical data. Under review, *American Statistician*.

Friendly, M. (1994a). Mosaic displays for multi-way contingency tables. *Journal of the American Statistical Association*, *89*, 190-200.

Friendly, M. (1994b). SAS/IML graphics for fourfold displays. *Observations*, *3(4)*, 47-56.

Funkhouser, H. G. (1937). Historical development of the graphical representation of statistical data. *Osiris*, *3*, 269-404.

Macdonald-Ross, M. (1977). How numbers are shown: A review of research on the presentation of quantitative information in texts. *Audio-Visual Communication Review*, *25*, 359-407.

Nightingale, F. (1858). *Notes on Matters Affecting the Health, Efficiency, and Hospital Administration of the British Army*. London: Harrison and Sons.

SAS Institute (1989). *SAS/IML Software: Usage and Reference, Version 6, First Edition*, Cary, NC: SAS Institute.

Tufte, E. R. (1983). *The Visual Display of Quantitative Information*. Cheshire, CT: Graphics Press.

Riedwyl, H., and Schuepbach, M. (1994). Parquet diagram to plot contingency tables. In *Softstat '93: Advances in Statistical Software*, F. Faulbaum (Ed.). New York: Gustav Fischer, 293-299.

Spence, I., and Lewandowski, S. (1990). Graphical perception. In *Modern Methods of Data Analysis*, J. Fox, & J. S. Long (Eds.), Sage Publications, 13-57.

Wainer, H., and Reiser, M. (1976). Assessing the efficacy of visual displays. *Graphic presentation of statistical information*, 83-88. American Statistical Association.

Wainer, H., and Thissen, D. (1981). Graphical data analysis. *Annual Review of Psychology*, *32*, 191-241.