Analysis of Opinion Swing: Comparison of Two Correlated Proportions

Telba Z. IRONY, Carlos A. de B. PEREIRA, and Ram C. TIWARI

An important problem that arises in introductory courses of applied statistics and categorical data analysis is the evaluation of opinion swing. Suppose that a group of individuals is surveyed on their support for the President. After the State of the Union Address the individuals are consulted again. The objective is to analyze whether or not there has been a change—a swing—in their opinion. Due to the longitudinal nature of data, only tests of hypothesis are presented to the students—for instance the McNemar test. The estimation of the parameters of interest is left to the advanced literature. In this article we suggest simple solutions for the estimation problem, to be presented in introductory courses of applied statistics.

KEY WORDS: Bayes factor; Credible interval; McNemar test; Partial likelihood.

1. INTRODUCTION

Panel survey samples are used to study changes in behavior or opinion swings. The statistical significance of the swing is verified by the standard McNemar test. Since the same group is studied longitudinally, however, the measurements are dependent, making it difficult to quote confidence intervals for the appropriate parameters. Consequently, the estimation of those parameters is done in the advanced literature and is not included in basic courses of data analysis. In this article we present simple solutions for this problem to be introduced in a course of applied statistics or categorical data analysis.

Suppose that a group of individuals is surveyed on their support for the President. After the State of the Union Address the individuals are consulted again. The objective is to analyze whether or not there has been a change—a swing—in their opinion.

In studies of this type, the data consist of n independent observations of the pair (X, Y) of random quantities taking values (x, y) in the set $\{(0,0),(0,1),(1,0),(1,1)\}$. X and Y represent opinions before and after the address. π_{xy} and n_{xy} denote, respectively, the probability and the sample frequency of the event $\{X = x, Y = y\}$.

The data can be arranged in a 2×2 table in a usual form, using "dots" in the marginal subscripts as follows:

	$Y=0$	$Y=1$	Total
$X=0$	π_{00} (n_{00})	π_{01} (n_{01})	$\pi_{0\cdot}$ ($n_{0\cdot}$)
$X=1$	π_{10} (n_{10})	π_{11} (n_{11})	$\pi_{1\cdot}$ ($n_{1\cdot}$)
Total	$\pi_{\cdot 0}$ ($n_{\cdot 0}$)	$\pi_{\cdot 1}$ ($n_{\cdot 1}$)	1 (n)

The objective is to estimate the following parameters:

$\theta = \pi_{01}/(\pi_{01} + \pi_{10})$—it measures the proportion of swings from 0 to 1;

$\rho = \pi_{01}/\pi_{10}$—it is the odds ratio of the swing type; and

$\delta = \pi_{01} - \pi_{10}$—it measures the difference of the two types of swing.

In addition, it is desired to test whether or not $\pi_{01} = \pi_{10}$.

In practice, these parameters measure the same effect. We discuss all of them in this article because all three forms appear in the literature. Furthermore, it shows the students that alternative parameterizations can change the level of difficulty of a problem.

The background is presented in Section 2 that will guide the instructor through several nuances of the problem, comment on important references, and explain the theory behind the technique. Section 3 provides a brief discussion of hypothesis testing to be presented to the students. Section 4 is devoted to the estimation of the parameters of interest. Section 5 discusses examples to be used by the instructors to compare Bayesian and classical techniques. Section 6 presents a case study that gave rise to heated class discussions and motivated us to attack the problem from several different perspectives.

2. BACKGROUND

The statistical model for the data $(n_{00}, n_{01}, n_{10}, n_{11})$ is multinomial with parameter $\boldsymbol{\pi} = (\pi_{00}, \pi_{01}, \pi_{10}, \pi_{11})$, where $0 \le \pi_{00}, \pi_{01}, \pi_{10}, \pi_{11} \le 1$ and $\sum_{x,y=0}^{1} \pi_{xy} = 1$.

The family of Dirichlet distributions is the natural conjugate prior for $\boldsymbol{\pi}$. For properties see Wilks (1968, chap. 7). Since the prior distribution we consider for $\boldsymbol{\pi}$ is Dirichlet with parameter $\mathbf{a} = (a_{00}, a_{01}, a_{10}, a_{11})$, its posterior distribution is also Dirichlet with updated parameter $\mathbf{A} = (A_{00}, A_{01}, A_{10}, A_{11})$, where $A_{xy} = a_{xy} + n_{xy}$.

Testing $H_0 : \pi_{01} = \pi_{10}$ is equivalent to testing $H_0 : \pi_{\cdot 0} = \pi_{0\cdot}$, which is a test often used in controlled clinical trials to study the effectiveness of a certain treatment. Since it is usually difficult to obtain the responses with and without treatment for the same subject, the procedure usually employed is to match pairs of subjects on the basis of characteristics that are associated with the response being

Telba Z. Irony is Mathematical Statistician, Division of Biostatistics, Center for Devices and Radiological Health, Food and Drug Administration, 1350 Piccard Drive, Rockville, MD 20850 (E-mail: tzi@cdrh.fda.gov). Carlos A. de B. Pereira is Professor, Department of Statistics, University of Sao Paulo, Sao Paulo, Brazil. Ram C. Tiwari is Professor, Department of Mathematics, University of North Carolina, Charlotte, NC 28223. The views expressed are those of the author and not necessarily those of the FDA.

studied, and to randomize the treatment assignments within each matched pair. The objective is to find out whether or not the probabilities of the response—effectiveness of the treatment—are the same for both subjects of the pair.

The problems above are instances of the general problem of estimating the difference between two correlated proportions. It was first attacked by McNemar (1947) and his hypothesis test solution is widely used in fields such as marketing, psychology, and medicine.

Several authors have also considered the problem of estimating the difference between two correlated proportions from a classical perspective. See, for example, Cox (1970, p.55), Fleiss (1981, p. 112), Santner and Snell (1980), and Armitage and Berry (1987) among others. More recently, Lloyd (1990) constructed approximate and conservative confidence intervals for the difference δ.

Agresti (1990) presented a good review of statistical techniques to compare dependent proportions, many of them based upon asymptotic results. However, the asymptotic methods are appropriate only when the diagonal total ($d = n_{01} + n_{10}$) is sufficiently large to allow the use of the normal approximation. Furthermore, the unknown parameter of interest, θ, should not be far from .5.

All of these studies offer methods to test for no difference between the *correlated* proportions. However, the presence of nuisance parameters prevents the derivation of an exact classical method to quote confidence intervals for the difference, δ, or the ratio ρ. Consequently, the intervals are either approximate or conservative. As pointed out by Lloyd (1990), the coverage properties of such confidence intervals are lower than the nominal level and sometimes the results are inaccurate (see also Casella 1986).

Broemeling and Gregurich (1996) performed a Bayesian analysis of matched categorical responses. They generated a sample from the posterior distribution of $(\pi_{00}, \pi_{01}, \pi_{10}, \pi_{11})$ and used resampling in order to make inferences about δ.

In this article, we start focusing on the parameter $\theta = \pi_{01}/(\pi_{01} + \pi_{10})$. Testing $H_0 : \pi_{01} = \pi_{10}$ against $H_1 : \pi_{01} \neq \pi_{10}$ is equivalent to testing $H : \theta = .5$ against $A : \theta \neq .5$. Our interest is only on the parameters π_{01} and π_{10} and, consequently, only the trinomial data $(n_{01}, n_{10}, n_{00} + n_{11})$ is considered, as in Altham (1971). Since $d = n_{01} + n_{10}$ and $\eta = \pi_{01} + \pi_{10}$ the likelihood becomes:

$$
\begin{aligned}
&f(n_{01}, n_{10}, n - d | \pi_{01}, \pi_{10}, \eta) \\
&= \binom{n}{n_{01}, n_{10}} \pi_{01}^{n_{01}} \pi_{10}^{n_{10}} (1 - \eta)^{n-d} \\
&= \binom{n}{d} \eta^d (1 - \eta)^{n-d} \binom{d}{n_{01}} \theta^{n_{01}} (1 - \theta)^{n_{10}} \\
&= f_0(d|\eta) f_1(n_{01}|\theta, d).
\end{aligned}
$$

Note that this likelihood is written as a product of two factors, one depending only on the parameter of interest, θ, and the other depending only on the nuisance parameter, η. In order to test H against A only the first factor is needed.

The statistic d is called specific sufficient to the nuisance parameter, η, and specific ancillary to the parameter of interest, θ. This means that there is no loss of information by considering only f_1 when estimating θ (Basu 1977).

Since the posterior distribution of $(\pi_{00}, \pi_{01}, \pi_{10}, \pi_{11})$ is Dirichlet with parameter $(A_{00}, A_{01}, A_{10}, A_{11})$, it can be shown (see Appendix A.1) that

a. $(\pi_{01}, \pi_{10}, 1 - \eta) \sim \text{Dirichlet}(A_{01}, A_{10}, A_{00} + A_{11})$;
b. $\theta \sim \text{Beta}(A_{01}, A_{10})$;
c. $\eta \sim \text{Beta}(A_{01} + A_{10}, A_{00} + A_{11})$;
d. θ and η are independent; and
e. $\rho = \frac{\theta}{1-\theta}$ is a Beta of second kind with parameters (A_{01}, A_{10}).

3. THE HYPOTHESIS TEST

To test $H : \theta = .5$ against $A : \theta \neq .5$, we consider the partial likelihood

$$f_1(n_{01}|\theta, d) = \binom{d}{n_{01}} \theta^{n_{01}} (1 - \theta)^{n_{10}}.$$

From the classical point of view, an exact significance test may be obtained by computing a p value using the above binomial model. Furthermore, this model will generate the McNemar test when n is large and the normal approximation is used (Agresti 1990).

A Bayesian test consists of evaluating the posterior odds ratio which is the ratio of the posterior probabilities of H and A.

If the prior probability of H is $\xi = \Pr(H) = 1 - \Pr(A)$, then the posterior odds ratio becomes $O = \frac{\xi}{1-\xi} \frac{f_H}{f_A}$, where f_H and f_A are the average likelihoods under H and A, respectively.

$R = \frac{f_H}{f_A}$ is called Bayes's factor (Press 1989).

Since H is simple, f_H is obtained by replacing θ by .5 in f_1. f_A is obtained by integrating f_1 weighted by the prior density of θ—an interesting exercise for the students. The resulting expression for the posterior odds ratio is

$$O = \frac{\xi B(a_{01}, a_{10})}{2^d (1 - \xi) B(A_{01}, A_{10})},$$

where $B(x, y)$ is the Beta function evaluated at the point (x, y).

The null hypothesis, $H : \theta = .5$, is rejected if $O < c$, where c is a positive constant related to the losses and gains associated with the decisions of accepting or rejecting the null hypothesis. When $\xi = 1 - \xi = .5$, $O = R$, and the Bayesian test is based on the Bayes's factor R.

4. INTERVAL ESTIMATION

4.1 Credible Intervals

An interval with credibility $(1 - \alpha)100\%$, $0 \leq \alpha \leq 1$, is an interval with posterior probability equal to $(1 - \alpha)$. Here, we shall consider the smallest $(1 - \alpha)100\%$ credible interval.

This article deals only with Beta distributions or its transformations. In such cases, if g and G are, respectively, the density and the distribution functions of the parameter, the interval (a, b) is the smallest credible interval with credibility $(1 - \alpha)100\%$ for the parameter if $G(b) - G(a) = 1 - \alpha$

and $g(b) = g(a)$. Hence, to obtain the interval (a, b), one should solve, numerically, the system of two equations and two unknowns.

First we consider the situation in which the null hypothesis H is not rejected. When this happens, it is desired to construct a $(1-\alpha)100\%$ credible interval for the common parameter $\pi_c = \pi_{01} = \pi_{10}$. Recalling that $\pi_{01} + \pi_{10} = \eta \sim \text{Beta}(A_{01} + A_{10}, A_{00} + A_{11})$, we first construct a credible interval for η using the procedure described earlier. Next, we obtain a credible interval for π_c by dividing the limits of that interval by 2.

The credible interval for θ is also obtained by using the fact that θ has a Beta distribution.

To construct a credible interval for ρ, we recall that its density is a Beta of second kind (Khatri and Rao 1968), given by

$$f(\rho|n_{01}, n_{10}) = \frac{1}{B(A_{01}, A_{10})} \frac{\rho^{A_{01}-1}}{(1+\rho)^{A_{01}+A_{10}}} \quad \text{for} \quad \rho \geq 0,$$

and also use the procedure described above.

The mean and the variance of ρ are:

$$\bar{\rho} = E\{\rho|n_{01}, n_{10}\} = \frac{A_{01}}{A_{10}-1},$$

and

$$V\{\rho|n_{01}, n_{10}\} = \bar{\rho}\frac{A_{01}+A_{10}-1}{(A_{10}-1)(A_{10}-2)}, \quad \text{for} \quad A_{10} > 2.$$

In case $A_{10} \leq 2$, a credible interval for $\rho^{-1} = (1-\theta)/\theta$ may be constructed instead (see Example 3). This is done by interchanging the roles of A_{01} and A_{10} in the expressions above.

Finally, the estimation of δ is a bit problematic because the analytical form of its density cannot be obtained. Consequently, the credible interval for δ is constructed by generating a large number of observations from the posterior distribution of $(\pi_{01}, \pi_{10}, 1-\eta)$; that is, a Dirichlet$(A_{01}, A_{10}, A_{00} + A_{11})$.

If $A = A_{00} + A_{01} + A_{10} + A_{11}$, the posterior mean and variance of δ are:

$$\bar{\delta} = E\{\delta|n_{00}, n_{01}, n_{10}, n_{11}\} = \frac{A_{01} - A_{01}}{A},$$

and

$$V\{\delta|n_{00}, n_{01}, n_{10}, n_{11}\} = \frac{4A_{01}A_{10} + (A_{01}+A_{10})(A_{00}+A_{11})}{(A+1)A^2}.$$

For this last expression see Appendix A.2.

4.2 Confidence Intervals

Classical confidence intervals for $\pi_c = \eta/2$ and θ are obtained as standard binomial confidence intervals from the partial likelihoods f_0 and f_1, respectively.

The confidence interval for δ is obtained by using the procedure suggested by Lloyd (1990). According to him, an approximate $(1-\alpha)100\%$ confidence interval for δ is given by

$$\hat{\delta}^* \pm \left(q \left(\frac{\hat{\eta}^* - \hat{\delta}^{*2}}{n} \right)^{1/2} + \frac{1}{n} \right),$$

where $D = n_{01} - n_{10}$, $d = n_{01} + n_{10}$, $\hat{\delta} = \frac{D}{n}$, and $\hat{\eta} = \frac{d}{n}$. Here we are estimating η by $\hat{\eta}$.

Moreover,

$$\hat{\delta}^* = \frac{\hat{\delta}}{1+\frac{q^2}{n}}, \quad \hat{\eta}^* = \frac{\hat{\eta}}{1+\frac{q^2}{n}},$$

and q is the $(1-\frac{\alpha}{2})100\%$ quantile of a standard normal distribution.

Santner and Snell (1980) provided a procedure to construct a confidence interval for ρ when π_{01} and π_{10} are independent. However, there is no exact classical procedure to obtain a confidence interval for ρ when π_{01} and π_{10} are correlated.

For large values of $d = n_{10} + n_{01}$, an approximate $(1-\alpha)100\%$ confidence interval for $\rho = \frac{\theta}{1-\theta}$ may be obtained using the Delta method (Sen and Singer 1993, p. 131). When d is large, $\hat{\theta} = \frac{n_{01}}{d}$ is asymptotically normal and an approximate $(1-\alpha)100\%$ confidence interval for $g(\theta)$, a continuous function of θ, is given by $g(\hat{\theta}) \pm q\{[g'(\hat{\theta})]^2\hat{V}(\hat{\theta})\}^{1/2}$, where g' is the derivative of g and $\hat{V}(\hat{\theta})$ is the estimate of the asymptotic variance of $\hat{\theta}$.

Consequently, when d is large, an approximate confidence interval for ρ is given by $\hat{\rho} \pm q\frac{\{\hat{V}(\hat{\theta})\}^{1/2}}{\hat{\theta}^2}$, where $\hat{V}(\hat{\theta}) = \frac{\hat{\theta}(1-\hat{\theta})}{d}$.

For θ and ρ, both classical and Bayesian methods of estimation provide similar results when sample sizes are large. However, for small sample sizes, the Bayesian estimation may detect differences between the proportions that were not detected by the classical counterparts.

5. EXAMPLES

This section presents three examples to illustrate Bayesian and classical techniques. To show the students that the two approaches may disagree, we analyze three cases:

Example 1: both techniques accept the null hypothesis.
Example 2: both techniques reject the null hypothesis.
Example 3: the techniques disagree.

The instructor may follow the examples in class or assign them as a course project. The students will have the opportunity to explore integration techniques in order to construct the credible intervals and simulation techniques to generate the distribution of δ. A standard mathematical package or even a spreadsheet may be chosen to perform the task.

5.1 Example 1: Both Techniques Accept the Null Hypothesis

A sample of individuals are surveyed on their support for the President. After the President's address they are consulted again. Has there been a swing in opinion?

The survey results may be displayed in the following table:

	After		
Before	*No*	*Yes*	*Total*
No	20	17	37
Yes	10	53	63
Total	30	70	100

The objective is to check whether or not the proportion of individuals who did not support the President before the address ($\pi_{0\cdot}$) is the same as after the address ($\pi_{\cdot 0}$); that is, $H_0 : \pi_{0\cdot} = \pi_{\cdot 0}$. This is equivalent to testing whether the proportion of swings from "Yes" to "No" is the same as the proportion of swings from "No" to "Yes"; that is, $H : \theta = .5$ against $A : \theta \neq .5$.

To compare Bayesian credible intervals and classical confidence intervals, we consider a uniform prior for π (corresponding to $a_{00} = a_{01} = a_{10} = a_{11} = 1$). This is a straightforward way to express no preference for any of the possible parameter values. There are other choices to represent noninformative prior opinions available in the literature. We believe, however, that the uniform prior above is the most intuitive.

We will compare the Bayesian and classical hypotheses tests by considering $\xi = 1 - \xi = .5$ and $c = 1$, in all examples. This will slightly change the prior distribution used above by assessing a positive probability to H_0. In this case, the Bayes factor is $R = .5^{27}\{B(18, 11)\}^{-1} = 1.76 > 1$ and the p value computed from a Binomial distribution with parameters $d = 27$ and $\theta = .5$ is .248.

Consequently, H is not rejected and one may estimate the common value $\pi_c = \pi_{01} = \pi_{10}$, knowing that $\eta \sim$ Beta(29, 75). Note also that $\theta \sim$ Beta(18, 11).

The estimates of the relevant parameters are given in the following table.

Parameter	*Post. mean*	*Post. std.*	*95% Credible interval*	*95% Confidence interval*
π_c	.140	.022	[.095; .181]	[.092; .179]
θ	.621	.089	[.447; .790]	[.430; .837]
ρ	1.800	.748	[.808; 3.762]	[.371;3.035]
δ	.067	.051	[−.030; .165]	[−.042; .177]

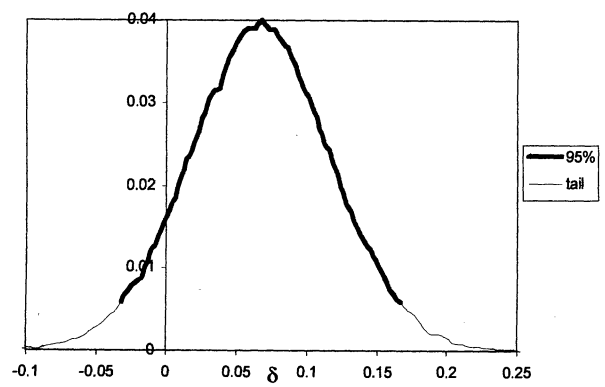

Figure 1. Simulated density of δ in Example 1: 100,000 observations.

The credible interval for δ was constructed using 100,000 values, simulated as described in Section 4 (see Figure 1). The confidence interval for ρ is approximate, given by the Delta method. Since $d = 27$ is not very large, the approximation is not very good.

5.2 Example 2: Both Techniques Reject the Null Hypothesis

Suppose that the survey results were as follows:

	After		
Before	*No*	*Yes*	*Total*
No	20	21	41
Yes	9	50	59
Total	29	71	100

In this case we obtain $R = .5^{30}\{B(22, 10)\}^{-1} = .413 < 1$, and the p value computed from the Binomial($d = 30, \theta = .5$) is .043 resulting in the rejection of H.

The estimates of the relevant parameters are given in the following table.

Parameter	*Post. mean*	*Post. std.*	*95% Credible interval*	*95% Confidence interval*
θ	.688	.081	[.529; .840]	[.536; .864]
ρ	2.444	1.026	[1.123; 5.250]	[.500; 4.160]
δ	.115	.053	[.015; .220]	[.003; .228]

In this case the credible interval for δ was also constructed using a simulated sample (see Figure 2). Again, we have a crude approximation for the confidence interval for ρ because $d = 29$ is not sufficiently large.

Next, we present an example with a small sample size ($n = 14$), showing a difference between the Bayesian and the classical results.

5.3 Example 3: The Techniques Disagree

Suppose that the survey results were as follows:

	After		
Before	*No*	*Yes*	*Total*
No	1	7	8
Yes	1	5	6
Total	2	12	14

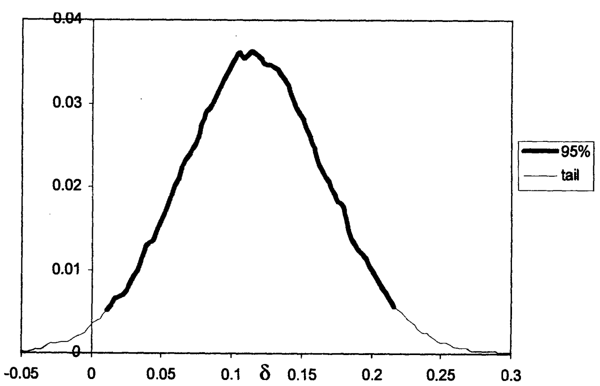

Figure 2. Simulated density of δ in Example 2: 100,000 observations.

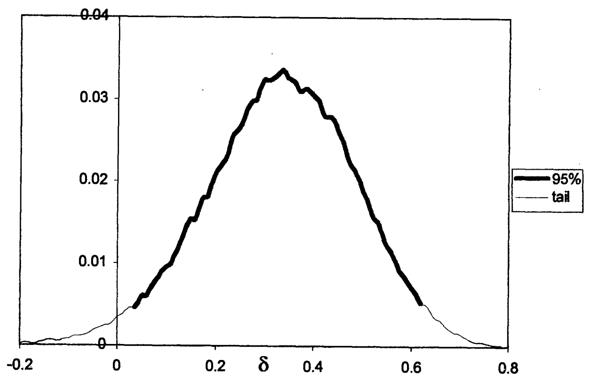

Figure 3. Simulated density of δ in Example 3: 100,000 observations.

Here, the Bayes factor, $R = .5^8\{B(8,2)\}^{-1} = .281 < 1$, leading to the rejection of H. On the other hand, the p value (Binomial($d = 8, \theta = .5$)) is .070 and H is not rejected at a significance level $\alpha = .05$.

The estimates of the relevant parameters are given in the following table.

Parameter	Post. mean	Post. std.	95% Credible interval	95% Confidence interval
θ	.727	.129	[.567; .991]	[.473; .968]
ρ^{-1}	.286	.247	[.009; .764]	
δ	.333	.148	[.025; .615]	[−.035; .711]

In this case $d = 8$, and there is no classical procedure to compute a confidence interval for ρ or ρ^{-1}.

The classical 95% confidence interval for the difference δ encompasses 0, whereas the Bayesian credible interval does not (see Figure 3). In addition, the confidence interval for θ encompasses .5, whereas the Bayesian credible interval does not. This means that the classical analysis is in disagreement with the Bayesian analysis, not rejecting the hypothesis that the proportions π_{01} and π_{10} are equal.

6. A CASE STUDY IN THE APPLIED STATISTICS COURSE

The Department of Statistics offers a course on applied statistics to upper level undergraduate and master's students. In this course, the students solve consulting problems in the Consulting Center. We guided the students to solve the following problem using the ideas in this article.

Two professors in the Department of Dentistry were judging the skills of 224 students in preparing dental fillings for 224 dental models. The evaluation results are presented in the following table:

	Professor A		
Professor B	Disapprove	Approve	Total
Disapprove	62	41	103
Approve	25	96	121
Total	87	137	224

The department decided to verify if Professor A was as tough as Professor B.

Our students identified the problem and decided to test $H_0 : \pi_{.0} = \pi_{0.}$ against $H_1 : \pi_{.0} \neq \pi_{0.}$. The Bayes factor is $R = .95 < 1$ and H is rejected. On the other hand, the p value calculated from a Binomial($d = 66, \theta = .5$) is $.064 > .05$ and H was not rejected at a 5% significance level.

The estimates of the parameters are shown in the following table. See also Figure 4.

Parameter	Post. mean	Post. std.	95% Credible interval	95% Confidence interval
θ	.618	.032	[.503; .731]	[.494; .721]
ρ	1.680	.433	[1.012; 2.717]	[.818; 2.446]
δ	.070	.021	[.001; .139]	[−.003; .146]

There is a disagreement between the classical and the Bayesian decisions and this gave rise to heated class discussions. The students noticed that, although the decisions were different, the difference between the intervals was very slight. If the intervals were computed at the level of 90%, both procedures would reject H. On the other hand, if the level were 99%, both procedures would not reject H. However, at the 95% level, the procedures disagree.

The conclusion is that whenever an interval's limit is too close to the value of a sharp hypothesis, one cannot blindly follow the hypothesis test rules in order to make a decision. Further thought must be given to the problem and one must give emphasis to the estimation process.

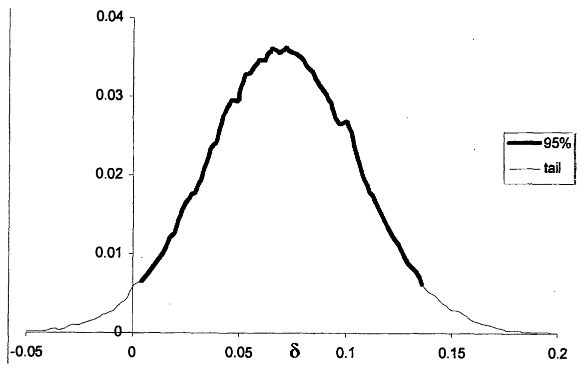

Figure 4. Simulated density of δ for the case study: 100,000 observations.

7. SUMMARY

A complete analysis of opinion swing to be introduced in a course of applied statistics or categorical data analysis is suggested. We suggest the use of a partial likelihood that transforms a difficult problem into a mere exercise of probability and statistics. In particular, it simplifies the inferences about θ under both classical and Bayesian approaches. Since ρ is a function of θ, Bayesian inferences for ρ are derived from the posterior distribution of θ. As far as we know, a classical interval for ρ is not available, when d is not large, which is often the case. Inferences about δ are more difficult. Lloyd (1990) presented a classical method to build a confidence interval for δ. We present a Bayesian solution to estimate δ using simulation. This provides a good opportunity for the instructor to illustrate the use of this pervasive numerical technique.

APPENDIXES

A.1 The Dirichlet Distribution Properties

To prove statements (a), (b), (c), (d), and (e) at the end of Section 2, we recall the characterization of Dirichlet distributions by transformations of independent Gamma random variables (Basu 1982). Let X, Y, Z, and W be independent Gamma random variables with parameters (A_{01}, β), (A_{10}, β), (A_{00}, β), and (A_{11}, β), respectively. Here, β is any positive real number and $A_{ij}, i, j = 0$ or 1 are as defined in the text. If $T = X + Y + Z + W$, then $(\frac{X}{T}, \frac{Y}{T}, \frac{Z}{T}, \frac{W}{T}) \sim (\pi_{01}, \pi_{10}, \pi_{00}, \pi_{11}) \sim \text{Dirichlet}(A_{01}, A_{10}, A_{00}, A_{11})$.

Note that this distribution does not depend on the value of β. In addition, $Z+W$ is distributed as $\text{Gamma}(A_{00}+A_{11}, \beta)$ and is independent of (X, Y). To show statement (a) we write that

$$(\pi_{01}, \pi_{10}, \pi_{00} + \pi_{11}) \sim (\frac{X}{T}, \frac{Y}{T}, \frac{Z+W}{T}) \sim \text{Dirichlet}(A_{01}, A_{10}, A_{00} + A_{11}).$$

Statements (b) and (c) are shown by noticing that $\theta = \frac{\pi_{01}}{\eta} \sim \frac{X/T}{(X+Y)/T} = \frac{X}{X+Y} \sim \text{Beta}(A_{01}, A_{10})$ and $\eta = (\pi_{01} + \pi_{10}) \sim \frac{X+Y}{T} \sim \text{Beta}(A_{01} + A_{10}, A_{00} + A_{11})$.

Now, we use the fact that T is a complete sufficient statistic for β and that $(\frac{X}{T}, \frac{Y}{T}, \frac{Z}{T}, \frac{W}{T})$ is ancillary for β.

As a consequence of Basu's Theorem (Boos and Hughes-Oliver 1998), $(\frac{X}{T}, \frac{Y}{T}, \frac{Z}{T}, \frac{W}{T})$ and T are independent.

Because $\frac{X/T}{(X+Y)/T} = \frac{X}{X+Y}$, then $\frac{X}{X+Y}$, $\frac{Z+W}{T}$, and T are mutually independent.

Because $\theta \sim \frac{X}{X+Y}$ and $\eta \sim (1 - \frac{Z+W}{T})$, statement (d) is proved.

Statement (e) follows from the fact that $\rho = \frac{\theta}{1-\theta} = \frac{\pi_{01}}{\pi_{10}} = \frac{X}{Y}$.

Note also that $\delta = \frac{X-Y}{T}$.

A.2 The Variance of δ

To compute the variance of δ we recall that π_{01}, π_{10}, and η are Beta random variables with parameters $(A_{01}, A_{00} + A_{10} + A_{11})$, $(A_{10}, A_{00} + A_{01} + A_{11})$, and $(A_{01} + A_{10}, A_{00} + A_{11})$, respectively. Using the fact that $2\text{cov}(\pi_{01}, \pi_{10}) = V(\eta) - V(\pi_{01}) - V(\pi_{01})$, we obtain $V(\delta) = 2V(\pi_{01}) + 2V(\pi_{10}) - V(\eta)$. The expression for the variance of δ presented in the text follows directly.

[Received February 1998. Revised July 1999.]

REFERENCES

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

Altham, P. M. E. (1971), "Exact Bayesian Analysis of an Intraclass 2 × 2 Table," *Biometrika*, 58, 679–680

Armitage, P., and Berry, G. (1987), *Statistical Methods in Medical Research*, London: Blackwell Scientific Publications.

Basu, D. (1977), "On the Elimination of Nuisance Parameters," *Journal of the American Statistical Association*, 72, 355–366.

——— (1982), "Basu Theorems," in *The Encyclopedia of Statistical Sciences*, eds. N. L. Johnson and S. Kotz, 1, pp. 193–196.

Boos, D. D., and Hughes-Oliver, J. M. (1998), "Applications of Basu's Theorem," *The American Statistician*, 52, 218–221.

Broemeling, L. D., and Gregurich, M. A. (1996), "A Bayesian Alternative to the Analysis of Matched Categorical Responses," *Communications in Statistics—Theory and Methods*, 25, 1429–1445.

Casella, G. (1986), "Refining Binomial Intervals," *The Canadian Journal of Statistics*, 14, 113–129.

Cox, D. R. (1970), *The Analysis of Binary Data*, London: Methuen.

Fleiss, J. L. (1981), *Statistical Methods for Rates and Proportions*, New York: Wiley.

Khatri, C. G., and Rao, C. R. (1968), "Some Characterizations of the Gamma Distribution," *Sankhyã*, A, 30, 157–166.

Lloyd, C. J. (1990), "Confidence Intervals From the Difference Between Two Correlated Proportions," *Journal of the American Statistical Association*, 85, 1154–1158.

McNemar, Q. (1947), "Note on the Sampling Error of the Difference Between Correlated Proportions," *Psychometrika*, 12, 1154–1158.

Press, S. J. (1989), *Bayesian Statistics: Principles, Models, and Applications*, New York: Wiley.

Santner, T. J., and Snell, M. K. (1980), "Small-Sample Confidence Intervals for $p_1 - p_2$ and p_1/p_2 in 2 × 2 Contingency Tables," *Journal of the American Statistical Association*, 75, 386–394.

Sen, P. K., and Singer, J. M. (1993), *Large Sample Methods in Statistics*, New York: Chapman & Hall.

Wilks, S. S. (1968), *Mathematical Statistics*, New York: Wiley.