A Standard Error Estimator for the Polarization Index: Assessing the Measurement Error in One Approach to the Analysis of Loyalty

Cam Rungie, Bruce Brown, Gilles Laurent and Suma Rudrapatna

The Polarization Index, developed by Sabavala, Morrison and Kalwani (1977), is a well-known measure of behavioral loyalty. It is of considerable use in analyzing purchase data and repeated choice revealed preference data. Until now, it has only been possible to evaluate the accuracy of estimates or polarization through simulations. We present a closed form estimator of its asymptotic standard error. Using simulations, we show that this estimator is sufficiently accurate for the typical sample sizes used in marketing studies.

Keywords: Polarization Index, Behavioural loyalty, preference data, repeat choice

Introduction

The Polarization Index, φ, appears in several guises in the marketing literature. It is a measure of loyalty for repeated discrete choices by the same person from the same binary choice set, such as repeated purchases from a product category with two brands, or repeated choices between a specified brand and all other brands within a product category.

The Polarization Index φ is first described as a characteristic of the beta distribution by Sabavala and Morrison (Sabavala & Morrison 1977), and is further investigated by Kalwani and Morrison (Kalwani 1980; Kalwani & Morrison 1980). It relates to other measures of loyalty including Hendry's k (Kalwani & Morrison 1977b), Bass's θ (Bass, Jeuland & Wright 1976), Chatfield and Goodhardt's work on the beta binomial distribution (Chatfield & Goodhardt 1970) and Goodhardt and Ehrenberg's Dirichlet S statistic (Goodhardt, Ehrenberg & Chatfield 1984). Under beta distribution conditions $\varphi = \theta = 1/(1+S) = 1-k$. Note that the Polarization Index refers to choices between two alternatives, whereas some of these other models refer to a selection from multiple alternatives in a choice set, such as the study of excess loyalty in big brands (Fader & Schmittlein 1993). Still the Index is useful because it can describe loyalty in the marginal choice between one specified alternative and all others in the choice set.

Typically, the Polarization Index is estimated from repeated discrete choice data using the beta binomial distribution. The data records the repeated discrete choices from a binary choice set by a sample of decision makers where the underlying choice probabilities, across the population of decision makers, has a beta distribution. Thus, the manifest repeated discrete choices have a binomial distribution mixed by a latent beta distribution. The resultant distribution for the data is the beta binomial but with the same parameters as the beta distribution. Polarization is best interpreted as a property of the underlying beta distribution. But, the data is observations from a beta binomial distribution and the estimation of the parameters is based on the beta binomial.

A common approach is to fit the distribution to the data using maximum likelihood. The accuracy of such estimates is of considerable interest, so much so that several authors use simulations to estimate it (Kalwani & Morrison 1977a; Kalwani & Morrison 1980; Tripathi & Gupta 1994). The contribution of this paper is to present a closed form estimator of the standard error of φ for maximum likelihood estimates.

The paper starts with a brief reminder on the Polarization Index, its mathematical definition, and its ability to measure loyalty. We then introduce a new estimator of its standard error. Finally, since this is an asymptotic estimator, we use simulations to demonstrate that it is valid for smaller samples, and specially for sample sizes typical of consumer research.

The Polarization Index

The Polarization Index, φ, is derived from the beta distribution. Consider a choice set of two alternatives, One and Two. Each decision maker in the population has a probability of choosing alternative One. Over the population, this choice probability is a bounded random variable P, where $0<P<1$. We assume it has a beta distribution (Johnson, Kotz & Balakrishnan 1994a; Johnson, Kotz & Balakrishnan 1994b). The beta distribution has two parameters, α_1 and α_2, and the following properties:

(1) $$\mu = E[P] = \frac{\alpha_1}{\alpha_1 + \alpha_2}$$

(2) $$\sigma^2 = Var(P) = \left(\frac{\alpha_1}{\alpha_1 + \alpha_2}\right)\left(\frac{\alpha_2}{\alpha_1 + \alpha_2}\right)\left(\frac{1}{1 + \alpha_1 + \alpha_2}\right)$$

The mean μ is just the average frequency (as a proportion) with which alternative One will be selected, its market share. The variance σ^2 is a measure of the differences between decision makers. Because P is bounded, $0<P<1$, then for a given μ, the variance is also bounded. The maximum possible variance is: $\mu(1-\mu) = \left(\frac{\alpha_1}{\alpha_1 + \alpha_2}\right)\left(\frac{\alpha_2}{\alpha_1 + \alpha_2}\right)$. When the variance is at its greatest, one group of decision makers has a choice probability P of zero and a second group a P of one. Each of these two groups is totally homogeneous and loyal. Thus the variance of P is an indicator of loyalty.

The Polarization Index is defined to be (Kalwani 1980; Kalwani & Morrison 1980; Sabavala & Morrison 1977):

(3) $$\varphi = \frac{1}{1 + \alpha_1 + \alpha_2} \quad \text{where } 0 < \varphi < 1.$$

The Polarization Index can be seen as a standardized form of the variance: the variance divided by its upper bound, $\varphi = \sigma^2 / [\mu(1-\mu)]$. It is this property which gives the Polarization Index its ability to measure loyalty. The closer φ is to one, the more the probabilities P concentrate on zero and one and the more each individual repeatedly chooses the same alternative. Conversely, if the Polarization Index is close to zero then all decision makers have about the same P and will switch regularly. There will be no loyalty. Consider three examples. First, if respondents are asked their gender, they will give repeatedly the same answer. This leads to φ = 1. Second, assume respondents are asked to toss a coin and to report the result. If the question is asked repeatedly, there is no 'loyalty' and this leads to φ = 0. Third, consider a brand with 20% market share. It might conceivably achieve this through 20% of all buyers always choosing it, and the remaining 80% never doing so (φ = 1), or through all buyers buying the brand at random, 20% of the time (φ = 0). In the former case there is considerable loyalty and in the latter considerable switching.

Also from (1) to (3):

(4) $$\alpha_1 = \frac{(1-\varphi)\mu}{\varphi} \quad \text{and} \quad \alpha_2 = \frac{(1-\varphi)(1-\mu)}{\varphi}$$

The Polarization Index, as a measure of loyalty, is of high practical importance. Different marketing strategies will apply when loyalty is high and when it is low. The very basis of product differentiation, segmentation and the selection of the marketing mix is in the variance over the population of preferences for individual alternatives and in the directly associated loyalty. An over statement of this loyalty by brand managers would lead to an over emphasis on segmentation. Conversely an understatement would lead to a failure to recognize a marketing opportunity. The Polarization Index measures this variation and loyalty. It would be most useful to regularly publish typical loyalty levels for product categories and brands, and to analyze marketing activities on that basis.

There is evidence that the Polarization Index is constant for all brands in a given product category (Ehrenberg 1988; Ehrenberg, Uncles & Goodhardt 2003; Goodhardt, Ehrenberg & Chatfield 1984; Uncles, Ehrenberg & Hammond 1995), but conversely there is also evidence that bigger brands have greater loyalty, with higher values for the Polarization Index (Fader & Schmittlein 1993). The Index is a sensible measure of possible variations in loyalty between brands, it can identify 'niche' brands with excess loyalty and 'change of pace' brands with less loyalty (Kahn, Kalwani & Morrison 1988).

Estimating the Polarization Index and its standard error

The estimator

Consider data where each decision maker makes q repeated independent choices, with replacement, from the same binary choice set. Each decision maker's q choices follow a binomial process based on her or his individual value of the choice probability P. Over the population of decision makers, the choices of Alternative One are a random variable R which follow a binomial process mixed by the beta distribution, i.e. the choices have a beta binomial distribution with the same parameters α_1 and α_2 (Johnson, Kotz & Kemp 1993). Thus, R is a discrete variable, with $0 \le R \le q$. Provided q>1, the beta binomial distribution can be fitted to data, and the parameters α_1 and α_2 can be estimated (Tripathi & Gupta 1994). Then, φ is easily derived, using (3). Note that for some estimation procedures, and in particular for maximum likelihood estimation, the observed number of choices by each decision maker, q, may vary from one decision maker to the next (Edwards 1976; Eliason 1993; Kalwani 1980; Kalwani & Morrison 1977a).

For at least two reasons, maximum likelihood theory provides one of the better methods for estimating the population parameters of a distribution from sample data (Edwards 1976;

Eliason 1993). First, generally, the estimates generated by maximizing the likelihood function will have less sampling error than estimates generated by other methods. Furthermore, the likelihood function leads to an estimate of the standard errors. The expected matrix of second partial differentials of the log-likelihood function, differentiated with respect to the parameters, gives what is known as the Hessian Matrix. In general, the variance covariance matrix for the maximum likelihood estimators (MLEs) of the parameters is approximately the inverse of the Hessian matrix but enumerated at the MLEs (Eliason 1993). For the beta binomial distribution, this inverse matrix includes an estimate of the variance of the maximum likelihood estimate of the Polarization Index. An outline of the proof is given in the Appendix. Taking the square root gives the standard error of the estimate of the Polarization Index. The main result is:

(5) $$\lim_{n\to\infty}\{n\,\mathrm{var}(\hat{\varphi})\} = \hat{\varphi}^4\left\{\frac{m_1 m_2}{m_1+m_2} - m_0\right\}^{-1}$$

with:

(6) $$m_0 = \sum_{j=0}^{q-1}\frac{1}{(j+\alpha_1+\alpha_2)^2}$$

(7) $$m_1 = E\left[\sum_{j=0}^{R-1}\frac{1}{(j+\alpha_1)^2}\right]$$

(8) $$m_2 = E\left[\sum_{j=0}^{q-R-1}\frac{1}{(j+\alpha_2)^2}\right]$$

where the expectation E is taken over the beta binomial distribution of R.

From (5), we derive the primary result of this paper, an estimate of the standard error for the likelihood estimate of the Polarization Index:

(9) $$\sigma_{\hat{\varphi}} = \sqrt{\mathrm{var}(\hat{\varphi})} = \sqrt{\frac{\hat{\varphi}^4}{n}\left\{\frac{m_1 m_2}{m_1+m_2} - m_0\right\}^{-1}}$$

In practice, the estimated standard error is evaluated by replacing unknown parameter values in (6), (7) and (8) by the values of their MLEs.

may still be sufficiently low.

Table 1 shows the standard errors calculated from (9) when each decision maker makes two selections, q=2, and for a range of situations. When the mean μ is not close to zero, the standard errors are small. The Polarization Index, φ, varies between zero and one. Relative to this range of values for φ, some of the standard errors are large. For a mean μ close to zero, a large sample size is required in order to generate an estimate of the Polarization Index which is even close to accurate. Intuitively, for a small mean, very few choices will be for Alternative One, and any estimate based on these data will have a high sampling error. For example if μ=.01, q=2 and n=1000, then over 2,000 choices Alternative One will only be selected 20 times. This is very little data to estimate loyalty accurately. Such situations are not unusual: some brands have market shares of the order of 1%. However, data bases for purchases often include several thousand buyers. Thus, in practice, standard errors for the Polarization Index for small brands may still be sufficiently low.

Table 1. Standard Errors for the Likelihood Estimate of the Polarization Index when *q* = 2

Coefficient of Polarization	n=500	n=1000	n=2000	n=4000
	Mean μ = 0.5			
φ = 0.1	0.044	0.031	0.022	0.016
φ = 0.5	0.039	0.027	0.019	0.014
φ = 0.9	0.019	0.014	0.010	0.007
	Mean μ = 0.1 (or μ = 0.9)			
φ = 0.1	0.057	0.040	0.028	0.020
φ = 0.5	0.065	0.046	0.032	0.023
φ = 0.9	0.033	0.023	0.016	0.012
	Mean μ = 0.01 (or μ = 0.99)			
φ = 0.1	0.136	0.096	0.068	0.048
φ = 0.5	0.195	0.138	0.097	0.069
φ = 0.9	0.100	0.071	0.050	0.035

A key conclusion from may still be sufficiently low.

Table 1 is that the standard errors vary considerably from situation to situation. In practice, the standard error should be checked for each data set. The estimator in (9) has four inputs; n, q, μ and φ. Generating a table which covers all possible values is not feasible given space constraints. Rather, we provide readers with a computer function which calculates the standard error, in Figure 1.

The properties of maximum likelihood estimators are asymptotic. In principle, (9) only gives the standard error as n tends to infinity, i.e. for very large sample sizes. In practice, the relevant question is how small can the sample size be before this estimator starts to err? How small can n be for it to be still safe to use it?

We run simulations on the following scenarios:

- Number of choices: q = 2 & 5,
- Polarization Index: φ = 0.1, 0.5 & 0.9 and
- Mean: μ = 0.5, 0.1 & 0.01 (because of symmetry this also covers the cases where μ = 0.9 and 0.99).
- Sample sizes: 50, 100, 200, … 2000.

For each scenario, we run 1,000 replications, entailing creating 400 million simulated choices. For each replication, the observed likelihood estimate of the Polarization Index is calculated. Then over the 1,000 replications the standard error of the observed likelihood estimate is calculated. Figure 2 plots the deviation between this observed standard error for the simulations and the standard error estimated by (9). In the figure, a deviation near zero validates (9).

```
%pise(n,q,mu,pi)
%   e.g.                pise(1000,2,.4,.5)
%                       result=0.028
%   Description         Standard error for likelihood estimate of Polarization Index.
%   Inputs
%       n               Sample size.
%       q               Number of trials in the beta binomial distribution.
%       mu              Mean of the choice probabilities.
%       pi              Polarization Index.

function pise=pise(n,q,mu,pi)
alpha1=(1-pi)*mu/pi;
alpha2=(1-pi)*(1-mu)/pi;
m0=0;
m1=0;
m2=0;
for j=0:q-1
   m0=m0+1/(j+alpha1+alpha2)^2;
end
for r=0:q
    % the next 4 lines give the pdf for the beta binomial distribution.
    bb=gammaln(alpha1+alpha2)+gammaln(q+1)+gammaln(alpha1+r)+gammaln(alpha2+q-r);
    bb=bb-gammaln(alpha1+alpha2+q)-gammaln(alpha1)-gammaln(r+1)-gammaln(alpha2);
    bb=bb-gammaln(q-r+1);
    bb=exp(bb);
    mf1=0;
    for j=0:r-1
        mf1=mf1+1/(j+alpha1)^2;
    end
    mf2=0;
    for j=0:q-r-1
        mf2=mf2+1/(j+alpha2)^2;
    end
    m1=m1+mf1*bb;
    m2=m2+mf2*bb;
end
pise=((1/(1+alpha1+alpha2))^4*(m1*m2/(m1+m2)-m0)^(-1)/n)^.5;
```

Figure 1. MATLAB Code for calculating the standard error for the Polarization Index.

Sample sizes for panel surveys and purchase data bases are often in excess of 1,000, and often more than two choices are recorded for many buyers. Figure 2 shows that in such cases (9) is generally accurate. It can be accurate for quite small sample sizes, provided the mean is large. As anticipated above, for smaller means and smaller sample sizes the accuracy of (9) diminishes. The figure indicates the point at which (9) substantially deviates from the simulated results.

One specific use for the Polarization Index is to identify ‘niche’ brands, with excess loyalty, or ‘change of pace’ brands, with less loyalty, if they exist. This may require measuring the Polarization Index for brands with small market shares. Figure 2 shows that for market shares as low as 1% the accuracy of (9) can be a problem. For sample sizes between 250 and 1000:

- If the Polarization Index is high (0.9) then (9) understates the standard error.
- If the Polarization Index is low (0.1) then (9) overstates the standard error.
- If the Polarization Index is 0.5 then (9) is approximately valid.

Fortunately, for many product categories, typical values of the Polarization Index can be of the order of 0.2 to 0.6. Nevertheless the real lesson from Figure 2 is that larger sample sizes are desirable. To study loyalty amongst brands with about 1% market share, the sample size should be over 1000 in order for the estimate of the standard error of the Polarization Index to be accurate.

Overall, the standard errors given by (9) are generally accurate for samples with a mean greater than 1% (and less than 99%) and a sample size greater than 1,000. If the mean is larger, the sample size can be smaller.

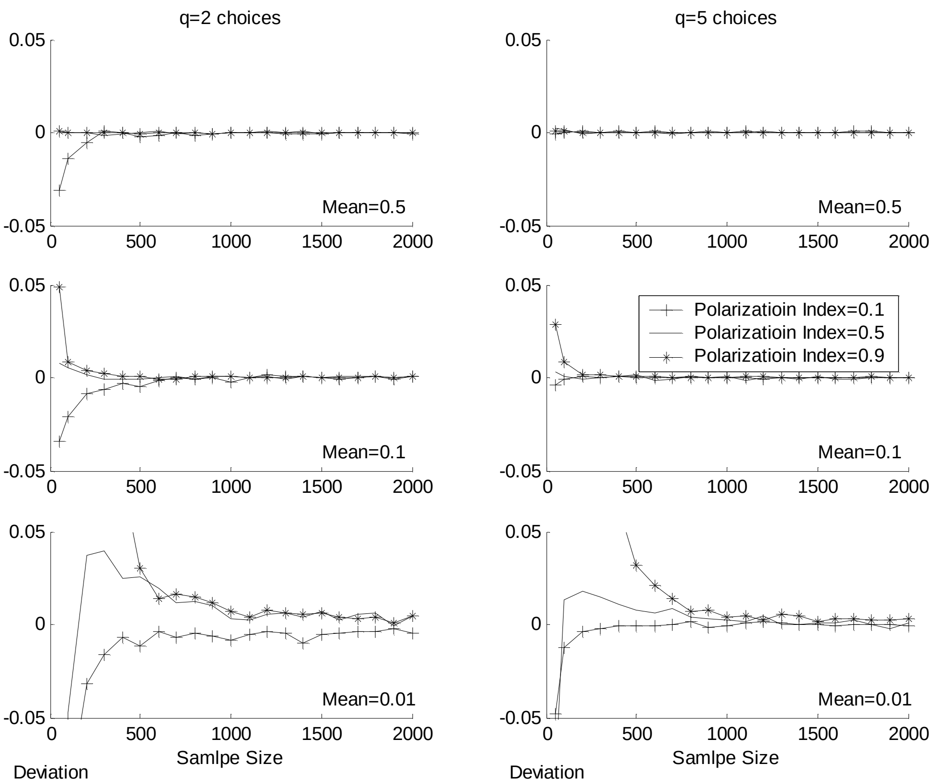

Figure 2. Deviation between the standard error observed in the simulations and the standard error estimated from (9).

Note. A small deviation indicates that (9) has high accuracy. Thus, if the sample size is greater than 1,000 then the estimator for the standard error from (9) is valid.

Discussion

Polarization is a well known measure of behavioral loyalty. So much so that formerly papers were published discussing its estimation through likelihood theory and establishing the accuracy of such estimates only using simulation. We have taken the discussion one step further. We have derived the formula for the standard error for large samples. We have then used simulations to show (1) that the formula is accurate for large samples (i.e. we use simulation to further validate the formula) and (2) how small the samples need be for the formula to be no longer accurate. Thus, the contribution of this technical paper is specific and narrow. It improves the methods for assessing the standard errors for polarization. It shows that for the sample sizes often found in the analysis of scanner and panel data the formula for the standard error is appropriate.

Further Research

The approach presented here, while generating a narrow and specific contribution, has the potential for expansion into useful further research. Polarization is just one measure of buyer behavior used in the analysis of panel data. Other measures include purchase rate, penetration, share of category requirements, 100% loyals and the Dirichlet S statistic. The last of these has been discussed by Chickamenahalli (2000) but the rest, while being routinely used in market research have no recognized procedures for estimating the standard errors. These measures are estimated directly from data or indirectly from the parameters of models such as the Dirichlet Model; see Ehrenberg, Uncles and Goodhardt (2004) and Rungie and Goodhardt (2004).

This paper considers the large sample standard errors from likelihood theory and has shown these to be appropriate in general for panel research. However, for smaller samples the formula derived hare is not appropriate. There would be value in considering bootstrapping and comparing its results with simulations (as is done here) and with the formula derived here.

Conclusion

The Polarization Index, φ, is a standardized measure of loyalty for repeated choices from a choice set containing two alternatives. The need for an assessment of the accuracy of its estimates is evident from papers in leading journals that address the question using simulations (Kalwani & Morrison 1977a; Kalwani & Morrison 1980; Tripathi & Gupta 1994). Our contribution is to present a closed form estimator for the standard error of maximum likelihood estimates of the Polarization Index. Its formula takes into consideration (i) the sample size, n, (ii) the number of repeated discrete choices recorded for each decision maker, q, (iii) the mean, μ, i.e. the proportion of decision makers selecting each alternative and (iv) the value of the Index, φ. We show that the standard error varies considerably depending on these factors. Further, since our estimator is based on the asymptotic properties of likelihood theory, we evaluate using simulations the accuracy of the formula for smaller sample sizes. The result indicates that for means of 1% or more and sample sizes of 1,000 or more the estimator for the standard error is accurate. For larger means, smaller sample sizes are also accurate.

References

Bass, FM, Jeuland A & Wright GP (1976). Equilibrium stochastic choice and market penetration theories: Derivations and comparisons. Management Science, 22 (10), 1051-1063.

Chatfield C & Goodhardt GJ (1970). The Beta-Binomial Model for consumer purchasing behaviour. Journal of the Royal Statistical Society: Series C (Applied Statistics), 19 (3), 240-250.

Chickamenahalli S (2000). Estimation of Heterogeneity in Brand Choice Models. Masters dissertation, University of South Australia.

Edwards A & Fairbank W (1976). Likelihood - An Account of the Statistical Concepts of Likelihood and its Application to Scientific Inference. Bath, UK: Cambridge University Press.

Ehrenberg ASC (1988). Repeat-Buying: Facts, Theory and Applications. New York: New York: Oxford University Press.

Ehrenberg ASC, Uncles MD & Goodhardt GG (2004). Understanding brand performance measures: Using Dirichlet benchmarks. Journal of Business Research, 57 (Issue 12), 1307-1325.

Eliason SR (1993). Maximum Likelihood Estimation - Logic and Practice. California: Sage Publications Inc.

Fader PS & Schmittlein DC (1993). Excess behavioral loyalty for high-share brands: Deviations from the Dirichlet Model for repeat purchasing. Journal of Marketing Research, 30 (4 November), 478-493

Goodhardt, GG, Ehrenberg ASC & Chatfield C (1984). The Dirichlet: A comprehensive model of buying behaviour. Journal of the Royal Statistical Society, Series A (General), 147 (Part 5), 621-655.

Johnson NL, Kotz S & Balakrishnan N (1994a). Continuous Univariate Distributions. Volume 1 (Second ed.). New York: John Wiley & Sons, Inc.

Johnson NL, Kotz S & Balakrishnan N (1994b), Continuous Univariate Distributions. Volume 2 (Second ed.). New York: John Wiley & Sons.

Johnson NL, Kotz S & Kemp AW (1993). Univariate Discrete Distributions (Second ed.). New York: John Wiley & Sons, Inc.

Kahn B, Kalwani MU & Morrison DG (1988). Niching versus change-of-pace brands: Using purchase frequencies and penetration rates to infer brand positionings. Journal of Marketing Research, 25 (Nov.), 384-390.

Kalwani MU (1980). Maximum likelihood estimation of zero-order models given variable numbers of purchases per household. Journal of Marketing Research, XVII, 547-551.

Kalwani MU & Morrison DG (1977a). Estimating the proportion of "always buy" and "never buy" consumers: A likelihood ratio test with sample size implications. Journal of Marketing Research, XIV, 601-606.

Kalwani MU & Morrison DG (1977b). A parsimonious description of the Hendry System. Management Science, 23 (January), 467-477.

Kalwani MU & Morrison DG (1980). Sample size requirements for zero-order models. Journal of Marketing Research, XVII, 221-227.

Rungie C & Goodhardt G (2004). Calculation of theoretical brand performance measures from the parameters of the Dirichlet Model. Marketing Bulletin, 15, Technical Note 2.

Sabavala DJ & Morrison DG (1977). A model of TV show loyalty. Journal of Advertising Research, 17 (6), 35-43.

Tripathi RC & Gupta RC (1994). Estimation of parameters in the Beta Binomial Model. Annals of the Institute of Statistical Mathematics, 46 (2), 317-331.

Uncles MD, Ehrenberg ASC & Hammond K (1995). Patterns of buyer behavior: Regularities, models and extensions. Marketing Science, 14 (3, 2 of 2), G71-G78.

Cam M. Rungie is a Senior Lecturer, School of Marketing, University of South Australia, Bruce M. Brown is a Professor at the School of Mathematics and Statistics, University of South Australia, Gilles Laurent is the Carrefour Professor of Marketing, Marketing Department, HEC School of Management, and Suma Rudrapatna, formerly known as Suma Chickamenahalli, is a Tutor in the School of Mathematics and Statistics, University of South Australia.

Appendix A. The large-sample variance of the MLE $\hat{\varphi}$. The derivation of (5)

The beta-binomial likelihood is $L = \prod_{i=1}^{n} \Pr\left(R = r_i\right)$, where the n observed values of R are $\{r_i\}$, and the beta-binomial probability is

(10) $$\Pr\left(R = r\right) = \frac{q!\,\Gamma(\alpha_1 + r)\,\Gamma(\alpha_2 + q - r)\,\Gamma(\alpha_1 + \alpha_2)}{r!\,(q-r)!\,\Gamma(\alpha_1 + \alpha_2 + q)\,\Gamma(\alpha_1)\,\Gamma(\alpha_2)}.$$

The large-sample asymptotic variance of the MLE $\hat{\varphi}$ is approximated by the bottom diagonal element of the inverse of the second derivative matrix of the log-likelihood, with respect to (μ, φ), evaluated at the MLE $(\hat{\mu}, \hat{\varphi})$.

Replace α_1, α_2 by $\mu = \alpha_1 / (\alpha_1 + \alpha_2)$ and $\varphi = (1 + \alpha_1 + \alpha_2)^{-1}$, then differentiate $\ln L$ twice with respect to (μ, φ). This is a routine but arduous task. There is some simplification from noting that the first derivative = 0 at the MLE. The negative of the resulting 2×2 matrix has the following elements, where $\hat{m}_1, \hat{m}_2$ denote versions of m_1, m_2 given in (7), (8) with expectation E is replaced by the sample average over observed values $\{r_i\}$. Also note that $\psi'(x) = \sum_{j=0}^{\infty} (j + x)^{-2}$, where $\psi(x) = \Gamma'(x) / \Gamma(x)$ is the digamma function.

Matrix (1,1) element: $\dfrac{n\,(1 - \hat{\varphi})^2}{\hat{\varphi}^2}\left(\hat{m}_1 + \hat{m}_2\right)$

Matrix (1,2), (2,1) elements: $\dfrac{n\,(1 - \hat{\varphi})}{\hat{\varphi}^3}\left\{(1 - \hat{\mu})\hat{m}_2 - \hat{\mu}\,\hat{m}_1\right\}$

Matrix (2,2) element: $\frac{n}{\hat{\varphi}^4}\left\{\mu^2\hat{m}_1 + (1-\hat{\mu})^2\hat{m}_2 - m_0\right\}$

From these expressions, the inverse matrix may be calculated, and the version of the bottom diagonal element coincides with (9). In the standard practice for likelihood theory the sample averages are converted back to the expected values which they estimate and so $\hat{m}_1, \hat{m}_2$ are replaced with m_1, m_2 (Chickamenahalli 2000) .