A Comparison of Two Elicitation Methods for a Prior Distribution for a Binomial Parameter
Author(s): Umesh Gavasakar
Source: *Management Science*, Vol. 34, No. 6 (Jun., 1988), pp. 784-790
Published by: INFORMS
Stable URL: http://www.jstor.org/stable/2632131
Accessed: 15/02/2015 19:07

MANAGEMENT SCIENCE
Vol. 34, No. 6, June 1988
Printed in U.S.A.

A COMPARISON OF TWO ELICITATION METHODS FOR A PRIOR DISTRIBUTION FOR A BINOMIAL PARAMETER*

UMESH GAVASAKAR
formerly at the University of Wisconsin at Stevens Point

This paper compares two methods of elicitation for the hyperparameters of the conjugate beta distribution for a binomial sampling model. The PM (posterior mode) method of Chaloner and Duncan appears to be more sensitive to elicitation errors than does a new method proposed here. This new method uses the device of imaginary results, due to I. J. Good (1967). An explicit model for elicitation errors is used in a simulation.

(ELICITATION; IMAGINARY RESULTS; BETA PRIOR DISTRIBUTION; BINOMIAL SAMPLING DISTRIBUTION; PM METHOD)

Introduction

For the Bayesian approach to be operational, reliable methods of elicitation are essential. The study of elicitation methods is just beginning. Some of the current effort is directed toward constructing methods of elicitation for popular models (Kadane et al. 1980, Garthwaite 1984, Chaloner and Duncan 1983), while some is aimed at comparing such methods (Dickey 1980, Lindley, Tversky, and Brown 1979). This paper aims to make a contribution in both senses, by proposing an elicitation method for the binomial sampling model, and by using an explicit model of elicitation error to compare it to the method of Chaloner and Duncan.

The binomial is an especially simple sampling model to study, and hence a natural choice of model to compare methods of prior elicitation. Additionally, deFinetti's Theorem (1937) says that every exchangeable sequence of Bernoulli random variables can be represented as a binomial distribution with respect to some prior.

Among families of prior distributions, the beta family is an obvious first choice because it is the natural conjugate (Raiffa and Schlaiffer 1961). It contains a wide variety of shapes, including both unimodal and bimodal distributions. For these reasons, this paper considers only members of the beta family as possible priors, that is, priors of the form

$$f(p\,|\,\alpha, \beta) = \frac{\Gamma(\alpha+\beta)}{\Gamma(\alpha)\Gamma(\beta)}\, p^{\alpha-1}(1-p)^{\beta-1}, \qquad 0 < p < 1.$$

Having made this choice, the problem of elicitation reduces to that of estimating the parameters of the beta distribution α and β, which act as hyperparameters for the binomial sampling model.

The PM Method of Chaloner and Duncan

Chaloner and Duncan (1983) give a method for eliciting these hyperparameters. (See also Chaloner and Duncan 1987.) Following others who endorse the use of the predictive distribution for elicitation (see Kadane et al. 1980, Geisser 1980, Kadane 1980, and Winkler 1980), they concentrate on the predictive distribution, which here is the beta-binomial or compound binomial distribution with parameters n, α, and β. Because of the asymmetry of the beta-binomial distribution when α and β are unequal, the mode,

* Accepted by Robert L. Winkler; received June 17, 1985. This paper was with the author and Joseph B. Kadane 22 months for 2 revisions.

0025-1909/88/3406/0784$01.25

m, is more attractive to elicit than a moment. (Another possibility, the predictive median, is not explored here because of the difficulty of computing it under the beta-binomial distribution.)

The steps of the PM (posterior mode) method of Chaloner and Duncan are as follows:

(1) specify a value for n, the number of hypothetical trials to be considered in the elicitation ($n = 20$ is suggested by Chaloner and Duncan)

(2) ask for m, the most likely number of successes in n trials

(3) display a bar chart of a binomial $(n, m/n)$ distribution

(4) ask for dropoffs d_l and d_u defined as follows: $d_l = p(m - 1)/p(m)$ and $d_u = p(m + 1)/p(m)$, where $p(\ )$ is the predictive probability of the subject. The quantities d_l and d_u are bounded above by 1 and below by the analogous quantities for the binomial distribution calculated in step (3). Additionally, these quantities as elicited are required to satisfy the following relationship:

$$d_l d_u > \frac{m(n - m)}{(m + 1)(n - m + 1)}. \tag{1}$$

(Chaloner and Duncan give this inequality in the opposite sense, but this appears to have been a typographical error, for reasons demonstrated in the Appendix.)

(5) Using these values, and conditioning on m, the program solves for α and β in the equations

$$d_l = \frac{f(m - 1)}{f(m)} = \frac{(n - m)(m + \alpha)}{(m + 1)(n - m + \beta - 1)} \qquad \text{and}$$

$$d_u = \frac{f(m + 1)}{f(m)} = \frac{m(n - m + \beta)}{(n - m + 1)(m + \alpha - 1)} \tag{2}$$

where $f(\ )$ is the probability mass function of the beta-binomial distribution. (Chaloner and Duncan give the relations in (2), but reverse them.) These equations yield two linear equations in α and β. Call the solutions α_1 and β_1.

(6) In the ensuing steps, the mode of the elicited prior is held fixed at the value $\gamma = (\alpha_1 - 1)/(\alpha_1 + \beta_1 - 2)$, which is the mode of the beta distribution with parameters α_1 and β_1. Now calculate the shortest interval having probability at least 50%, and display its constituent points, their probabilities, and the sum of the probabilities. The subject is asked if this interval is too long, in which case $h = -1$, just right, in which case $h = 0$, or too short, in which case $h = 1$. Then new values for α and β are defined by

$$\alpha_{i+1} = 1 + 2^h(\alpha_i - 1), \qquad \beta_{i+1} = 1 + 2^h(\beta_i - 1). \tag{3}$$

Note these new values do not change the ratio of $(\alpha - 1)$ to $(\beta - 1)$ so that the mode γ does not change.

(7) If h is not zero, continue with step (6) until h changes sign. Suppose this change of sign happens at stage k. Then $(\alpha_k - 1, \beta_k - 1)$ and $(\alpha_{k-1} - 1, \beta_{k-1} - 1)$ form upper and lower bounds on $(\alpha - 1, \beta - 1)$ which are tightened with bisection until $h = 0$ is achieved.

(8) Chaloner and Duncan suggest that this procedure might be repeated with different values of n, and that the resulting estimates of the hyperparameters might be amalgamated in some way.

Critique of the PM Method

In view of the subjective difficulty in eliciting d_l and d_u, steps (4) and (5) are a natural subject of inquiry. These steps might be regarded in two lights; either as a way of approximating the mode γ of the prior, or as a way of eliciting α and β. Each of these objectives is considered in turn below.

1. Dropoffs to elicit the prior mode. It is natural to ask why the estimate m/n for the prior mode is not adequate. It should be noted that the mode m as elicited is the mode of the subject's subjective predictive distribution, not his prior. Nonetheless, m/n does seem a reasonable estimate. As will soon be seen, something very close to this is implied by the PM method.

The appendix proves the following inequalities for the estimate, γ, of the prior mode used in the PM method:

$$\frac{m}{n+1} < \gamma < \frac{m+1}{n+1}. \tag{4}$$

The implication of equation (4) for the PM method is that γ is quite insensitive to the values of d_l and d_u. One might take $(m + \frac{1}{2})/(n + 1)$ for the value of γ without making a serious error in estimating it (less than 0.025 if $n = 20$), and not do steps (4) and (5) at all. This would be equivalent to taking $\alpha'_1 = 1 + k(m + \frac{1}{2})$ and $\beta'_1 = 1 + k(n - m + \frac{1}{2})$ for some k. Since steps (6) and (7) amount to the estimation of k, this seems feasible and would eliminate a difficult elicitation from the PM method. However, this idea has the disadvantage that it uses a single elicited datum, m, to estimate the hyperparameter γ. It is better practice (see Kadane, et al. 1980 for a discussion) to average the responses from several elicitations to estimate a single hyperparameter.

2. Dropoffs to estimate α and β. A second way of looking at steps (4) and (5) is as a way of estimating α and β. To see how sensitive the Chaloner-Duncan dropoff method is to perturbations in the assessed dropoffs, consider the following derivatives, found by differentiating (A2) in the Appendix:

$$\frac{\partial\beta}{\partial d_l} = \frac{-(n-m+1)(n-m)(m+1)}{n+1}, \qquad \frac{\partial\alpha}{\partial d_l} = \frac{-m(n-m)(m+1)}{n+1},$$

$$\frac{\partial\beta}{\partial d_u} = \frac{-m(n-m)(n-m+1)}{n+1}, \qquad \frac{\partial\alpha}{\partial d_u} = \frac{-m(m+1)(n-m+1)}{n+1},$$

all evaluated at $d_u = d_l = 1$. Supposing that n approaches infinity in such a way that m/n approaches a constant neither zero nor one, each of these derivatives is seen to be quadratic in n. This suggests extreme sensitivity, a matter to be taken up later by simulation.

An Alternative Method

The mode of the predictive distribution consists of all integers in the interval

$$\left[\frac{na-b}{a+b}, \frac{(n+1)a}{a+b}\right]$$

where $a = \alpha - 1$ and $b = \beta - 1$. Since this interval has length one, only in the situation in which the endpoints are integers does the mode fail to be unique. If the endpoints are integers, then there are exactly two modes, namely the two endpoints. These facts can be summarized by saying that the mode(s) are the closest integer(s) to $((n + 1)a/(a + b)) - (1/2)$. The simplicity of the mode suggests repeated modal-elicitations might be used to estimate the hyperparameters. This would proceed as follows:

Modified PM Elicitation Procedure

1. Envision n_0 independent Bernoulli trials. Specify your mode, m_0, i.e., the number of successes most likely, in your opinion.

2. For $i = 1, \ldots, I$, suppose k_i Bernoulli trials were performed and s_i successes were obtained. Now envision n_i further trials. Specify your mode m_i.

3. Find estimates of a and b to minimize

$$\sum_{i=0}^{I}\left[m_i - \left(\frac{(n_i+1)(a+s_i)}{a+b+k_i} - \frac{1}{2}\right)\right]^2$$

where $k_0 = s_0 = 0$.

TABLE 1

Values of n, k, and s used for simulating the alternative procedure, and the correct predictive mode that results

Index	n (further trials)	k (conducted in past)	s (# successes in k past trials)	m (new predictive mode)
0	20	0	0	7
1	20	25	7	6
2	20	25	14	10
3	20	25	21	14
4	20	40	8	4
5	20	40	16	8
6	20	40	26	11
7	20	40	35	16

The numbers n_i, k_i and s_i are design parameters, controlled by the person doing the elicitation to help the person being elicited to describe most fully his/her views. In the work that follows, I have adopted the Chaloner-Duncan suggestion of $n_i = 20$. The parameters k_i (number of hypothetical trials conducted in the past) should be sufficient to have a real impact on the expert's view, but not so many as to overwhelm it. In the simulation described later, I use $k = 25$ and 40. For s_i the hypothetical number of successes in i trials, the most information is gathered by presenting the expert with several rather different situations to respond to. Thus, in the simulation when k is 25, I use 7, 14 and 21 successes, and when k is 40, I use 8, 16, 26 and 35 successes. These values effectively span the space of reasonable possible outcomes.

The sum of squared deviation between the elicited modes m_i and the targets (center point of the interval in which the true mode should fall) is used because it is heuristically reasonable and easy to compute.

A Simulation to Compare Methods for Estimating α and β

The purpose of this section is to compare steps (1), (2), (4) and (5) of the PM method to the modified procedure stated above. While there is some interest in the comparison of the procedures *per se*, of perhaps greater interest is the explicit model of elicitation errors required for the simulation.

Throughout, the prior is assumed to have a beta distribution with parameters $\alpha = 5.0$ and $\beta = 8.0$. These were chosen to have a moderately skewed example. Other parameter values have been studied, and the overall results are much the same.

To simulate the PM method, suppose that the expert is asked to envision 15, 25, and 40 trials ahead, with modal number of successes 5, 9, and 14, respectively. However, the expert cannot assess these modal numbers exactly. Suppose that the expert has the following error structure in eliciting the mode: with probability 0.40, he gives the true mode. With probability 0.25 he errs by giving a mode too high by one, and symmetrically with probability 0.25 he errs by giving a mode too low by one. Finally, the probability of giving a mode too high or too low by two is 0.05 each.

TABLE 2

Results for Base Run Simulation

(Uncorrelated Error with 0.02 *Standard Deviation)*

	$\bar{\hat{\alpha}}$	$s_{\hat{\alpha}}$	$\bar{\hat{\beta}}$	$s_{\hat{\beta}}$
PM Procedure	9.86	0.62	17.12	1.17
Modified Procedure	4.78	0.15	7.72	0.20
Target	5		8	

TABLE 3

Results for the PM Procedure with Elicitation Standard Error 0.002. *The Modified Procedure is Unchanged*

	$\bar{\hat{\alpha}}$	$s_{\hat{\alpha}}$	$\bar{\hat{\beta}}$	$s_{\hat{\beta}}$
PM Procedure	5.10	0.022	8.51	0.061
Target	5		8	

The problem in simulating an error process for the dropoffs is that they may not satisfy (1) or the condition that neither dropoff be larger than one. The solution adopted is to calculate the intended dropoffs using (2). If both are less than or equal to 1, no modification is made. If the larger is greater than one, the smaller is replaced by the product of the intended dropoffs and the larger by one. This replacement guarantees intended dropoffs satisfying (1) and that neither intended dropoff should exceed one. Simulated dropoffs are then formed by adding normal noise to the intended dropoffs, in the base case, independent normals with mean zero and standard deviation 0.02. If the simulated dropoffs are greater than one or fail to satisfy (1) they are rejected and new simulated dropoffs are found. Note that the standard deviations are such that the expert is credited with quite accurate elicitation of the dropoffs.

The above method requires nine elicited quantities—three modes and six dropoffs. To compare, eight elicited medians are used in the alternative method. They are given in Table 1.

To each of the predictive modes, error of the same sort as above is added (-2, -1, 0, 1, 2) with probabilities (0.05, 0.25, 0.40, 0.25, 0.05), respectively. The results obtained are shown in Table 2.

These results show that the modification is substantially better. One question they raise is how small the standard deviation on the normal noise would have to be before the performance of the two methods is roughly the same. Reducing the normal noise to a standard deviation of 0.002, the results obtained are shown in Table 3. Note that this case requires truly remarkable ability to assess dropoffs.

Values of σ between 0.02 and 0.002 show a gradual improvement in the PM procedure. Next, suppose that the errors in the dropoffs are positively correlated, with correlation 0.7. (The standard deviations are returned to 0.02.) Then Table 4 records the results. The PM procedure is degraded very substantially by this change.

Finally, consider the case in which the modes for the PM procedure are considered to be easier to elicit than those for the modified procedure. Thus, suppose for the PM procedure the error added to the mode is (-1, 0, 1) with respective probabilities (0.25, 0.50, 0.25), while for the modified procedure the error is (-3, -2, -1, 0, 1, 2, 3) with respective probabilities (0.05, 0.1, 0.2, 0.3, 0.2, 0.1, 0.05). The normal errors in the dropoffs are returned to the base case: uncorrelated with standard errors of 0.02. Table 5 gives the results.

This change does not help the PM method to become more accurate. For the modified procedure, the parameter estimates are essentially unchanged, and are very accu-

TABLE 4

Results for the PM Procedure with Noise Correlated 0.7, *and Standard Errors* 0.02. *The Modified Procedure is Unchanged*

	$\bar{\hat{\alpha}}$	$s_{\hat{\alpha}}$	$\bar{\hat{\beta}}$	$s_{\hat{\beta}}$
PM Procedure	15.3	1.92	26.4	3.39
Target	5		8	

TABLE 5

Results for Both Procedures with Revised Modal Elicitation

	$\bar{\hat{\alpha}}$	$s_{\hat{\alpha}}$	$\bar{\hat{\beta}}$	$s_{\hat{\beta}}$
PM Procedure	10.02	0.61	17.36	1.10
Modified Procedure	5.03	0.24	7.92	0.32
Target	5		8	

rate. The simulation study discussed above suggests that the estimates of the hyperparameters using the modified procedure may be more accurate than those using the Chaloner-Duncan PM procedure. But note that the PM procedure makes use of computer graphics to display the fitted distributions, giving feedback that helps the expert. This feature is absent in the modified procedure. Also the modified procedure assumes that the expert uses Bayes Theorem to modify his judgements after being given data on the number of successes in previous trials. This may not be true in practice. The modified procedure has the advantage, however, of using only modal estimates, which arguably are easier to elicit than are dropoffs.

Conclusions

The use of dropoffs to elicit the hyperparameters in the beta-binomial model can be regarded in two lights: either as a way to estimate $\gamma = a/(a+b) = (\alpha-1)/(\alpha+\beta-2)$, or as a way to estimate α and β. The inequality (4) shows that for practical purposes the estimate of γ does not depend on the dropoffs in the PM method. The simulation study shows that the modified procedure is more accurate as an estimate of α and β. It would seem, then, that dropoffs may be dominated by other ideas for elicitation in this model.

The comparison of methods given here relies partly on analysis, and partly on a simulation. The latter requires explicit statement of the suspected error process in elicitation.[1]

[1] This paper was part of the author's Ph.D. dissertation in statistics at Carnegie-Mellon University. After his untimely death, the paper was completed by his advisor, Joseph B. Kadane. Requests for reprints should be sent to Professor Joseph B. Kadane, Dept. of Statistics, Carnegie-Mellon University, Pittsburgh Pa. 15213.

Appendix: Bounds on γ

For convenience, slightly different notation is used here. Let $s = d_l$ and $t = d_u$. Also let $a = \alpha_1 - 1$ and $b = \beta_1 - 1$. Then equations (2) can be expressed in the form

$$\begin{bmatrix} s(m+1) & -(n-m) \\ -m & t(n-m+1) \end{bmatrix} \begin{bmatrix} b \\ a \end{bmatrix} = \begin{bmatrix} (1-s)(n-m)(m+1) \\ (1-t)m(n-m+1) \end{bmatrix}. \tag{A1}$$

Hence

$$\begin{bmatrix} b \\ a \end{bmatrix} = \frac{1}{D} \begin{bmatrix} t(n-m+1) & n-m \\ m & s(m+1) \end{bmatrix} \begin{bmatrix} (1-s)(n-m)(m+1) \\ (1-t)m(n-m+1) \end{bmatrix} \tag{A2}$$

where $D = st(m+1)(n-m+1) - m(n-m)$. Since all other quantities are positive, (A1) will have positive solutions in a and b provided $D > 0$, i.e., provided (1) in the text is correct.

Now

$$\gamma = (a_1 - 1)/(a_1 + \beta_1 - 2) = a/(a+b) = \frac{1}{1 + (b/a)}. \tag{A3}$$

Hence bounds on γ can be obtained from bounds on (b/a). From (A2),

$$\frac{b}{a} = \frac{(n-n)(n-m+1)}{m(m+1)} B \qquad \text{where} \tag{A4}$$

$$B = \frac{t(m+1) + wm}{(n-m) + sw(n-m+1)} \tag{A5}$$

and $w = (1-t)/(1-s)$, $0 < w < \infty$.

Now $B(s, t, w)$ in the domain $0 < w < 1, 0 < t < 1, 0 < w < \infty$ is relevant only in a two-dimensional subspace of the domain, since the values of any two parameters determine the third. Nonetheless, (A5) may be formally treated as if B were a function of three variables. Any bound on B in this three-dimensional domain is a bound on B in the relevant (curved) two-dimensional domain.

$$B(s, t, 0) = \frac{t(m+1)}{n-m} < \frac{m+1}{n-m}, \tag{A6}$$

$$B(s, t, \infty) = \frac{m}{s(n-m+1)} > \frac{m}{n-m+1}. \tag{A7}$$

Now B is shown to be decreasing in w, by computing its first derivative:

$$\frac{\partial B(s, t, w)}{\partial w} = \frac{[(n-m) + sw(n-m+1)]m - [t(m+1) + wm]s(n-m+1)}{[(n-m) + sw(n-m+1)]^2}$$

$$= \frac{m(n-m) - st(m+1)(n-m+1)}{[(n-m) + sw(n-m+1)]^2} \leq 0 \tag{A8}$$

using (1).

Putting together (A6), (A7), and (A8),

$$\frac{m+1}{n-m} > B(s, t, 0) \geq B(s, t, w) \geq B(s, t, \infty) > \frac{m}{n-m+1} \tag{A9}$$

for all s and t, $0 < s < 1$, and $0 < t < 1$. Hence B is bound .n the relevant two-dimensional space by $(m+1)/(n-m)$ and $m/(n-m+1)$. These bounds on B imply bounds for (b/a) in (A4), and hence for γ in (A3), as follows:

$$\frac{n-m+1}{m} > b/a > \frac{n-m}{m+1}; \tag{A10}$$

$$\frac{m+1}{n+1} > \gamma > \frac{m}{n+1}. \tag{A11}$$

References

CHALONER, K. AND G. T. DUNCAN, "Assessment of a Beta Prior Distribution: PM Elicitation," *The Statistician*, 27 (1983), 174–80.

——— AND ———, "Some properties of the Dirichlet-Multinomial Distribution and Its Use in Prior Elicitation," *Comm. Statist.—Theory Methods*, 16, 2 (1987), 511–523.

DEFINETTI, B., "Foresight: Its Logical Laws, Its Subjective Sources," Translated and reprinted in *Studies in Subjective Probability*, H. Kyburg and H. Smokler (Eds.), 1964, Wiley, New York, 1937, 93–158.

DICKEY, J. M., "Beliefs about Beliefs: A Theory of Stochastic Assessments of Subjective Probabilities," In *Bayesian Statistics*, J. M. Bernardo, M. H. DeGroot, D. V. Lindley and A. F. M. Smith (Eds.), University Press, Valencia, 1980.

GARTHWAITE, P. M., "Quantifying Expert Opinion About a Linear Regression," Department of Statistics Technical Report No. 7, University of Aberdeen, Scotland, 1984.

——— AND J. M. DICKEY, "Double- and Single-Bisection Methods for Subjective Probability Assessment in a Location Scale Family," *J. Econometrics*, 29 (1985), 149–163.

GAVASAKAR, U., "A Study of Elicitation Procedures by Modelling the Errors in Responses," unpublished dissertation, Department of Statistics, Carnegie-Mellon University, 1984.

GEISSER, S., "A Predictivist Primer," In *Bayesian Analysis in Econometrics and Statistics*, A. Zellner (Ed.), North-Holland, Amsterdam, 1980, 363–381.

GOOD, I. J., "A Bayesian Significance Test for Multinomial Distributions," *JRSSB* 29 (1967), 399–431 (with discussion).

KADANE, J. B., "Predictive and Structural Methods for Eliciting Prior Distributions," In *Bayesian Analysis in Econometrics and Statistics*, A. Zellner (Ed.), North-Holland, Amsterdam, 1980, 89–93.

———, J. M. DICKEY, R. L. WINKLER, W. SMITH AND S. C. PETERS, "Interactive Elicitation of Opinion for a Normal Linear Model," *J. Amer. Statist. Assoc.*, 75 (1980), 845–854.

LINDLEY, D. V., A. TVERSKY AND R. V. BROWN, "On Reconciliation of Probability Assessments," *JRSSA*, 142 (1979), 146–180.

RAIFFA, H. AND R. SCHLAIFER, *Applied Statistical Theory*, Division of Research, Graduate School of Business Administration, Harvard University, Boston, 1961.

WINKLER, R., "Prior Information, Predictive Distributions and Bayesian Model Building," In *Bayesian Analysis in Econometrics and Statistics*, A. Zellner (Ed.), North-Holland, Amsterdam, (1980), 95–109.