

UNCLASSIFIED

SECURITY CLASSIFICATION OF THIS PAGE

DTIC

SELECTED REPORT DOCUMENTATION PAGE

DTIC FILE COPY

7

1a. REPORT SECURITY CLASSIFICATION
UNCLASSIFIED

2a. SECURITY CLASSIFICATION AUTHORITY
N/A

2b. DECLASSIFICATION/DOWNGRADING SCHEDULE
N/A

1b. RESTRICTIVE MARKINGS

3. DISTRIBUTION/AVAILABILITY OF REPORT
Approved for Public Release; Distribution Unlimited.

4. PERFORMING ORGANIZATION REPORT NUMBER(S)
FSU Technical Report No. M-760

5. MONITORING ORGANIZATION REPORT NUMBER(S)
AFOSR-TR- 87 - 1041

6a. NAME OF PERFORMING ORGANIZATION
Florida State University

6b. OFFICE SYMBOL
(If applicable)

7a. NAME OF MONITORING ORGANIZATION
AFOSR/nm

6c. ADDRESS (City, State and ZIP Code)
Department of Statistics
Tallahassee, FL 32306-3033

7b. ADDRESS (City, State and ZIP Code)
Building 410
Bolling, AFB, DC 20332-6448

8a. NAME OF FUNDING/SPONSORING ORGANIZATION
AFOSR

8b. OFFICE SYMBOL
(If applicable)
NM

9. PROCUREMENT INSTRUMENT IDENTIFICATION NUMBER
F49620-85-C-0007
AFOSR No. 85-C-0007

8c. ADDRESS (City, State and ZIP Code)
Building 410
Bolling AFB, DC 20332-6448

10. SOURCE OF FUNDING NOS.

PROGRAM ELEMENT NO.	PROJECT NO.	TASK NO.	WORK UNIT NO.
6.1102F	2304	AS	

11. TITLE (Include Security Classification)
A Three-Parameter Generalisation of the Beta-Binomial Distribution with Applications

12. PERSONAL AUTHOR(S)
Peter Danaher

13a. TYPE OF REPORT
Technical Journal

13b. TIME COVERED
FROM _____ TO _____

14. DATE OF REPORT (Yr., Mo., Day)
July 1987

15. PAGE COUNT
20

16. SUPPLEMENTARY NOTATION

17. COSATI CODES

FIELD	GROUP	SUB. GR.

18. SUBJECT TERMS (Continue on reverse if necessary and identify by block number)
binomial; beta-binomial; modified beta-binomial; empirical Bayes; magazine reach; effective reach.

19. ABSTRACT (Continue on reverse if necessary and identify by block number)
A three-parameter generalisation of the beta-binomial distribution (BBD) is derived and examined. ^{The author} We obtain the maximum likelihood estimates of the parameters and show that the regularity conditions for asymptotic efficiency are satisfied. To exhibit the applicability of the generalised distribution ^{it is shown} we show how it gives an improved fit over the BBD for magazine exposure and consumer purchasing data. Finally we derive an empirical Bayes estimate of a binomial proportion based on the generalised beta distribution used in this study.

20. DISTRIBUTION/AVAILABILITY OF ABSTRACT
UNCLASSIFIED/UNLIMITED SAME AS RPT. DTIC USERS

21. ABSTRACT SECURITY CLASSIFICATION
UNCLASSIFIED

22a. NAME OF RESPONSIBLE INDIVIDUAL
Maj. Crowley

22b. TELEPHONE NUMBER
(Include Area Code)
202-767-5025

22c. OFFICE SYMBOL
AFOSR/nm1

AD-A185 733

AFOSR-TR- 87-1041

**A Three-Parameter Generalisation
of the Beta-Binomial Distribution
with Applications**

by

Peter J. Danaher

**Department of Statistics
Florida State University
Tallahassee, Florida 32306**

**FSU Technical Report M760
AFOSR Technical Report 87-208**

July 1987

**Research supported by the Air Force Office of Scientific
Research Grant AFOSR 85-C-0007.**

A THREE-PARAMETER GENERALISATION OF THE BETA-BINOMIAL DISTRIBUTION WITH APPLICATIONS

Peter J. Danaher
Department of Statistics
Florida State University
Tallahassee, Florida.

Key Words and Phrases: binomial; beta-binomial; modified beta-binomial; empirical Bayes; magazine reach; effective reach.

Summary

A three-parameter generalisation of the beta-binomial distribution (BBD) is derived and examined. We obtain the maximum likelihood estimates of the parameters and show that the regularity conditions for asymptotic efficiency are satisfied. To exhibit the applicability of the generalised distribution we show how it gives an improved fit over the BBD for magazine exposure and consumer purchasing data. Finally we derive an empirical Bayes estimate of a binomial proportion based on the generalised beta distribution used in this study.



Accession For	
NTIS CRA&I	<input checked="" type="checkbox"/>
DTIC TAB	<input type="checkbox"/>
Unannounced	<input type="checkbox"/>
Justification	
By	
Distribution	
Availability Codes	
Dist	Avail and/or Special
A-1	

1. Introduction

Suppose an advertiser is about to launch an advertising campaign by placing one advertisement in k successive issues of a magazine. His objective is to estimate the proportion of the population which sees none, one, or up to all k of the advertisements. We say that an individual reading a magazine is *exposed* to an advertisement when he or she sees the ad. Let X be the number of exposures and P the probability an individual is exposed to any one ad. The distribution of X is called the *exposure distribution* (e.d.). It is reasonable to assume that $X|P \sim \text{bin}(k, p)$ if, given $P = p$, readership of successive issues is independent (Ehrenberg 1975 discusses this assumption in some detail). Now let P be a random variable having a beta distribution. The beta distribution is particularly attractive since it can assume so many shapes. By compounding the binomial with the beta distribution we obtain the BBD. It is also known as the negative hypergeometric distribution (Johnson and Kots 1969). The mass function of the BBD is

$$f^B(X = x) = \binom{k}{x} \frac{\Gamma(\alpha + \beta)}{\Gamma(\alpha + \beta + k)} \frac{\Gamma(k - x + \beta)}{\Gamma(\beta)} \frac{\Gamma(x + \alpha)}{\Gamma(\alpha)},$$
$$x = 0, 1, \dots, k, \alpha > 0, \beta > 0,$$

where $\Gamma(l + 1) = l\Gamma(l)$, the usual gamma function.

The BBD was first derived by Skellam (1948). Multivariate generalizations of the BBD were studied by Ishii and Hayakawa (1960). The BBD has been successfully applied to estimating the e.d. of media schedules with one magazine (Metheringham 1964; Chandon 1976), television schedules (Rust and Klompaker 1981), consumer purchasing behaviour (Morrison 1979) and as an indicator of television loyalty (Sabavala and Morrison 1977). In addition the BBD has been used to estimate the distribution of household disease (Griffiths 1973) and proportions with extraneous variance (Kleinman 1973; Moore 1987).

Many people subscribe to one or more magazines. Among them is a proportion

which *always* reads a particular magazine. Chandon (1976) suggests a modification of the BBD which he calls the "two segment beta-binomial model". One segment is definite readers and the other is probable readers. In Section 2 we will derive a generalisation of the BBD based on the two segment model introduced by Chandon and develop it further. In Section 3 we will look at maximum likelihood estimation of the parameters of this new distribution then prove that regularity conditions for these estimates to be asymptotically efficient are satisfied. In Section 4 we will give some examples where the new distribution gives an improved fit over the BBD for magazine exposure data and consumer purchasing behaviour. Finally, in Section 5 we will use the generalised BBD to obtain an empirical Bayes estimate of a binomial proportion. This estimate will be applied to simulation of magazine e.d.s

2. Generalisation of the BBD

Let ω and $1 - \omega$ represent the proportion of definite and probable readers, respectively. The parameter ω may be viewed as a loyalty factor since a high value of ω indicates an appreciable reading loyalty whilst a low value indicates little or no loyalty to a particular magazine.

To incorporate this reading loyalty proportion we change the distribution of P from a beta distribution to a beta distribution mixed with a distribution degenerate at $p = 1$. The cumulative distribution function of P is now

$$F_P^M(p) = (1 - \omega) \int_0^p \frac{\Gamma(\alpha + \beta)}{\Gamma(\alpha)\Gamma(\beta)} y^{\alpha-1} (1 - y)^{\beta-1} dy + \omega I_{\{p=1\}}, \quad (2.1)$$

$$0 < p < 1, \alpha > 0, \beta > 0, 0 \leq \omega \leq 1.$$

When $F_P^M(p)$ is compounded with the $\text{bin}(k, p)$ distribution we obtain the modified BBD (MBBD) with mass function

$$f^{MB}(X = x) = (1 - \omega) \binom{k}{x} \frac{\Gamma(\alpha + \beta)}{\Gamma(\alpha + \beta + k)} \frac{\Gamma(k - x + \beta)}{\Gamma(\beta)} \frac{\Gamma(x + \alpha)}{\Gamma(\alpha)} + \omega I_{\{X=k\}},$$

$$x = 0, \dots, k. \quad (2.2)$$

When $\omega=0$ the MBBD reduces to the BBD.

We derive the factorial moments of the MBBD using knowledge of the factorial moments of the BBD and linearity of the expectation operator. They are,

$$m_l = (1 - \omega) \frac{k!}{(k-l)!} \frac{\Gamma(\alpha+l)}{\Gamma(\alpha)} \frac{\Gamma(\alpha+\beta)}{\Gamma(\alpha+\beta+l)} + \frac{k!}{(k-l)!} \omega, \quad l = 1, 2, \dots$$

In particular,

$$E(X) = k \frac{\alpha + \omega\beta}{\alpha + \beta} \tag{2.3}$$

and

$$\text{var}(X) = (1 - \omega) \frac{[k\alpha\beta(\alpha + \beta + k) + k^2\omega\beta^2(\alpha + \beta + 1)]}{(\alpha + \beta)^2(\alpha + \beta + 1)}.$$

3. Parameter Estimation

Chandon (1976) estimates α , β and ω by equating the sample proportion of nonreaders to the proportion of nonreaders given by the MBBD model for $k = 1, 2, 3$. In Section 4 it will be seen that we do not have these data at our disposal. Furthermore, Chandon's method does not produce estimates having any optimality properties such as being BAN and, in addition, suffers from inconsistencies which sometimes force him to set $\omega = 0$, thereby losing any advantage of using the MBBD.

We will estimate α , β and ω using maximum likelihood.

Let n_i be the number of people in the sample (of size $n = \sum_i n_i$) which see i out of k issues of a magazine, $i = 0, 1, \dots, k$. We will see in Section 4 how these data are obtained from a media sample survey.

Define $c = c(\alpha, \beta) = [\Gamma(\alpha + k)\Gamma(\alpha + \beta)] / [\Gamma(\alpha)\Gamma(\alpha + \beta + k)]$.

Then the likelihood equations are

$$\begin{aligned}\frac{\partial \log L}{\partial \alpha} &= \sum_{i=1}^{k-1} \Delta_1(\alpha, i) n_i - (n - n_k) \Delta_1(\alpha + \beta, k) \\ &\quad + \frac{(1 - \omega)c[\Delta_1(\alpha, k) - \Delta_1(\alpha + \beta, k)] n_k}{\omega + (1 - \omega)c}, \\ \frac{\partial \log L}{\partial \beta} &= \sum_{i=0}^{k-1} \Delta_1(\beta, k - i) n_i - (n - n_k) \Delta_1(\alpha + \beta, k) \\ &\quad - \frac{(1 - \omega)c \Delta_1(\alpha + \beta, k) n_k}{\omega + (1 - \omega)c}, \\ \frac{\partial \log L}{\partial \omega} &= \frac{-(n - n_k)}{1 - \omega} + \frac{(1 - c)n_k}{\omega + (1 - \omega)c},\end{aligned}$$

where $\Delta_1(\gamma, l) = \sum_{j=0}^{l-1} 1/(\gamma + j)$.

By equating $\partial \log L / \partial \omega$ to zero we get

$$\hat{\omega} = \frac{n_k/n - c(\hat{\alpha}, \hat{\beta})}{1 - c(\hat{\alpha}, \hat{\beta})}$$

Substitution of $\hat{\omega}$ into the first order partial derivatives results in the following second order partial derivatives

$$\begin{aligned}\frac{\partial^2 \log L}{\partial \alpha^2} &= \frac{(n - n_k)c}{(1 - c)} [(\Delta_1(\alpha, k) - \Delta_1(\alpha + \beta, k))^2 / (1 - c) + \Delta_2(\alpha + \beta, k) \\ &\quad - \Delta_2(\alpha, k)] - \sum_{i=1}^{k-1} \Delta_2(\alpha, i) n_i + (n - n_k) \Delta_2(\alpha + \beta, k), \\ \frac{\partial^2 \log L}{\partial \beta \partial \alpha} &= \frac{(n - n_k)c}{(1 - c)} [(\Delta_1(\alpha + \beta, k) - \Delta_1(\alpha, k)) \Delta_1(\alpha + \beta, k) / (1 - c) \\ &\quad + \Delta_2(\alpha + \beta, k)] + (n - n_k) \Delta_2(\alpha + \beta, k), \\ \frac{\partial^2 \log L}{\partial \beta^2} &= \frac{(n - n_k)c}{(1 - c)} [(\Delta_1(\alpha + \beta, k))^2 / (1 - c) + \Delta_2(\alpha + \beta, k)] \\ &\quad - \sum_{i=0}^{k-1} \Delta_2(\beta, n - i) n_i + (n - n_k) \Delta_2(\alpha + \beta, k),\end{aligned}$$

where $\Delta_2(\gamma, l) = -\frac{\partial}{\partial \gamma} \Delta_1(\gamma, l)$.

The MBBD likelihood equations have no closed-form solution but may be solved by the Newton-Raphson method using the above partial derivatives. Since $\hat{\omega}$ is an explicit function of $\hat{\alpha}$ and $\hat{\beta}$ the numerical work is considerably reduced.

In Section 4 we will define some criteria by which the effectiveness of an advertising campaign is judged. These criteria are functions of $\vec{\theta} = (\alpha, \beta, \omega)$. To obtain asymptotic variances for the criteria estimates we first need the asymptotic joint distribution of $\vec{\theta} = (\hat{\alpha}, \hat{\beta}, \hat{\omega})$.

Consistency and asymptotic efficiency of the MLEs can be established using the multiparameter discrete distribution version of theorems utilized in Giesbrecht and Kempthorne (1976) and proved in Kulldorff (1957). The statement of Kulldorff's theorem has been tailored somewhat to suit the three-parameter, discrete distribution case.

Theorem (Kulldorff 1957). Let $f_i = f(X = i)$. The parameter space of $\vec{\theta}$ is denoted Ω and is an open ball. If the following regularity conditions are satisfied:

- i) $\frac{\partial \log f_i}{\partial \theta_j}$, $\frac{\partial^2 \log f_i}{\partial \theta_j^2}$ and $\frac{\partial^3 \log f_i}{\partial \theta_j^3}$ exist for every $\vec{\theta} \in \Omega$, $i = 0, \dots, k$, $j = 1, 2, 3$;
- ii) $\sum_{i=0}^k \frac{\partial f_i}{\partial \theta_j} = 0$, $\sum_{i=0}^k \frac{\partial^2 f_i}{\partial \theta_j^2} = 0$ for all θ_j ;
- iii) The information matrix $I(\vec{\theta})$ is positive definite;
- iv) There exist numbers $\{H_i\}$ (independent of the parameters except possibly the true parameter values) and a positive, twice differentiable function $g(\vec{\theta})$ such that

$$\left| \frac{\partial^2}{\partial \theta_{i_1} \partial \theta_{i_2}} \left[g(\vec{\theta}) \frac{\partial \log f_{i_3}}{\partial \theta_{i_3}} \right] \right| < H_{i_3}, \quad 1 \leq i_1 \neq i_2 \neq i_3 \leq 3,$$

for all parameter values $\in \Omega$ where $\sum_i f_i H_i < \infty$;

then $\vec{\theta}$ is unique and $\sqrt{n}(\hat{\vec{\theta}} - \vec{\theta}) \xrightarrow{d} MVN(\vec{0}, [I(\vec{\theta})]^{-1})$ as $n \rightarrow \infty$. \diamond

For the MBBD let $\vec{\theta}^T = (\alpha, \beta, \omega)$, $\Omega = (0, \infty) \times (0, \infty) \times (0, 1)$ and $f_i = f_i^{MB}$.

We will now show that the above conditions are satisfied.

- i) As f_i^{MB} is a ratio of polynomials in θ_j , it is clear that all third order partial derivatives of $\log f_i^{MB}$ exist on Ω .

ii) We have $\sum_{i=0}^k f_i^{MB} = 1$ so these two conditions are satisfied if we can interchange the order of differentiation (w.r.t. θ_j) and summation. As the sum is over only a finite number of terms and the second order derivatives exist (by (i)) this interchange is valid.

iii) The 3×3 information matrix $I(\vec{\theta})$ has elements

$$I_{11}(\vec{\theta}) = (1-\omega) \sum_{i=0}^{k-1} [\Delta_1(\alpha, i) - \Delta_1(\alpha + \beta, k)]^2 f_i^B + \frac{c^2(1-\omega)^2 [\Delta_1(\alpha, k) - \Delta_1(\alpha + \beta, k)]^2}{(1-\omega)c + \omega},$$

$$I_{22}(\vec{\theta}) = (1-\omega) \sum_{i=0}^{k-1} [\Delta_1(\beta, k-i) - \Delta_1(\alpha + \beta, k)]^2 f_i^B + \frac{c^2(1-\omega)^2 [\Delta_1(\alpha + \beta, k)]^2}{(1-\omega)c + \omega},$$

$$I_{33}(\vec{\theta}) = \frac{1-c}{[(1-\omega)c + \omega](1-\omega)},$$

$$I_{12}(\vec{\theta}) = (1-\omega) \sum_{i=0}^{k-1} [\Delta_1(\alpha, i) - \Delta_1(\alpha + \beta, k)][\Delta_1(\beta, k-i) - \Delta_1(\alpha + \beta, k)] f_i^B - \frac{c^2(1-\omega)^2 [\Delta_1(\alpha, k) - \Delta_1(\alpha + \beta, k)] \Delta_1(\alpha + \beta, k)}{(1-\omega)c + \omega},$$

$$I_{13}(\vec{\theta}) = \frac{c[\Delta_1(\alpha, k) - \Delta_1(\alpha + \beta, k)]}{(1-\omega)c + \omega},$$

$$I_{23}(\vec{\theta}) = -\frac{c\Delta_1(\alpha + \beta, k)}{(1-\omega)c + \omega}.$$

The proof that $I(\vec{\theta})$ is positive definite when $k \geq 3$ is somewhat messy. The full details are given by Danaher (1987). The information matrix is singular when $k \leq 2$ as this corresponds to the situation where we have more parameters to estimate than we have data.

iv) The only problem points are when $\alpha + \beta$ is near 0 or when ω is near 1. One possible g which fulfills this requirement is $g(\vec{\theta}) = (1-\omega)e^{-(1/\alpha+1/\beta)}$. This function is suitable since $\frac{1}{\alpha}g(\vec{\theta})$, $\frac{1}{\beta}g(\vec{\theta})$, $\frac{1}{\alpha+\beta}g(\vec{\theta})$ and all requisite derivatives tend to zero as α , β , and $\alpha + \beta$ tend to 0. In addition $g(\vec{\theta})$ eliminates any problems when $\omega = 1$.

It follows from Kulldorff (1957) that the MLEs are best asymptotically normal with covariance matrix $I^{-1}(\vec{\theta})$.

Kulldorff's (1957) results can also be applied to the BBD to show the MLEs of α and β are consistent and asymptotically normal, something Kleinman (1973) stated he was unable to do. Minor modifications to the above method are necessary; for example, let $\omega = 0$ in $g(\vec{\theta})$.

4. Applications

The survey data we will use here comes from the AGB: McNair Surveys New Zealand Ltd. "National Media Survey" of 5201 residents of New Zealand conducted in 1985. In the survey two of the questions asked of the respondents were (for weekly magazines);

Q1) "Have you personally read or looked into any issue of ...(magazine name) in the last seven days - it doesn't matter where?" (Has a Y/N answer).

Q2) "How many different issues of ...(magazine name), if any, do you personally read or look into in an average month - it doesn't matter where?" (Has answer 0,1,2,3,4 issues).

The wording of Q1 and Q2 are modified appropriately for fortnightly, monthly and two-monthly magazines. These questions were asked for forty different magazines.

An implicit assumption in the magazine advertising field is that a person who reads a magazine is exposed to all the advertisements in that magazine. This is unlikely to be true for people who meet the criterion of "read" in Q1 and Q2. However, it is usually impractical to ask respondents which advertisements they have been exposed to so we cannot avoid making this assumption for the available data.

The parameter estimates for the *National Business Review* are given in Table 1.

The estimate of ω tells us that 1.7% of the respondents always read this magazine. From Q1 we can estimate that in any particular week 2.6% of the population will read the *National Business Review*. This implies that of the 2.6% who read this magazine in any particular week $1.7/2.6 = 65.4\%$ read it every week. This gives the *National Business Review* a high readership loyalty, something well known by its publishers.

Let e_i be the estimated number of people who have i exposures. The e_i 's in Table 1 come from substituting the estimated parameters of Table 1 into (2.2). Then the Pearson χ^2 goodness of fit statistic is defined to be $\chi^2 = \sum_{i=0}^4 (n_i - e_i)^2/e_i = \sum_{i=0}^4 c_i$ (say). We can interpret c_i as the contribution to χ^2 from the i^{th} exposure. In this case the χ^2 goodness of fit statistic for the BBD is significant (p-value < 0.001) but for the MBBD it is not significant (p-value > 0.1). The c_i 's for the two distributions (values are in parentheses next to expected frequencies) show that a considerable improvement in accuracy has been made, particularly for three exposures.

Table 1 also gives the likelihood ratio test for $H_0 : \omega = 0$ vs. $H_1 : \omega > 0$. It shows that ω is significantly nonzero (there is 1 df for this test).

The important goodness of fit criterion to an advertising agency is not the χ^2 statistic (Naples 1979). They measure the closeness of the fit by three other criteria.

The first is *reach*, which is defined as the proportion of the population which is exposed to at least one of the advertisements, i.e., $1 - f^{MB}(X = 0)$. The second criterion is *effective reach*, the mean of the e.d. The third criterion is *single issue reach*, the proportion of the population exposed to any one issue of a magazine.

It can be seen in Table 1 that for the three criteria above the MBBD produces estimates closer to the observed values than the BBD.

Table 1: Readership data for the *National Business Review* showing the fits for the BBD and the MBBD. Sample size = 5201.

Number of Exposures	Observed Frequency	Expected Frequency	
		BBD (c_i)	MBBD (c_i)
0	4961	4961.3 (0.00)	4960.9 (0.00)
1	90	67.9 (7.19)	95.5 (0.32)
2	43	43.4 (0.00)	35.3 (1.68)
3	12	42.5 (12.89)	15.3 (0.71)
4	95	85.8 (0.99)	95.0 (0.00)
Parameter Estimates		$\hat{\alpha}=0.012$ $\hat{\beta}=0.372$	$\hat{\alpha}=0.024$ $\hat{\beta}=2.113$ $\hat{\omega}=0.017$
χ^2 Goodness of Fit		30.0	2.7
d.o.f.		2	1
Likelihood Ratio Test Statistic			18.8
Reach %	4.614	4.609	4.616
Effective Reach	0.114	0.120	0.114
Single Issue Reach %	2.59	3.01	2.84

Of course, it is expected that the addition of a parameter will make the model more flexible and hence improve the fit. Notice, however, that the shape of the distribution of P for the MBBD is different from that of the BBD, as shown in Figure 1. The essence of P^{MB} is a reverse J-shape distribution with a jump at $p = 1$. This, empirically and intuitively, is a better distribution to allow for magazine reading loyalty. On the other hand P^M assumes a U-shape which puts too much weight on the probability of three exposures, a property not consistent with the data.

Assuming our MBBD model is correct we can use (2.1) and (2.3) to write down expressions for reach (ρ):

$$\rho = 1 - f^{MB}(X = 0) = 1 - (1 - \omega) \frac{\Gamma(\alpha + \beta)}{\Gamma(\alpha + \beta + k)} \frac{\Gamma(\beta + k)}{\Gamma(\beta)} ;$$

and effective reach (ρ_e):

$$\rho_e = k \frac{\alpha + \omega\beta}{\alpha + \beta}.$$

Since both ρ and ρ_e are differentiable functions of (α, β, ω) we can obtain asymptotic variances for $\hat{\rho}$ and $\hat{\rho}_e$ using the delta method and the information matrix $I(\hat{\theta})$ given in Section 3. The asymptotic 95% confidence interval for ρ for the data in Table 1 is [4.03, 5.20]% and the 95% asymptotic confidence interval for ρ_e is [0.098, 0.130].

Just because the MBBD fit is better for one magazine it does not mean that it is always better than the BBD. We compared the fit for all forty magazines in the survey which covered the entire spectrum of entertainment magazines through to computer magazines. The average absolute error between estimated reach and sample reach for the forty magazines is 0.030% for the BBD and 0.036% for the MBBD. The average absolute error between estimated effective reach and sample effective reach is 1.36% for the BBD and 0.28% for the MBBD. The average absolute error between estimated single issue reach and sample single issue reach is 0.44% for the BBD and 0.36% for the MBBD. Summing up, the MBBD error is marginally

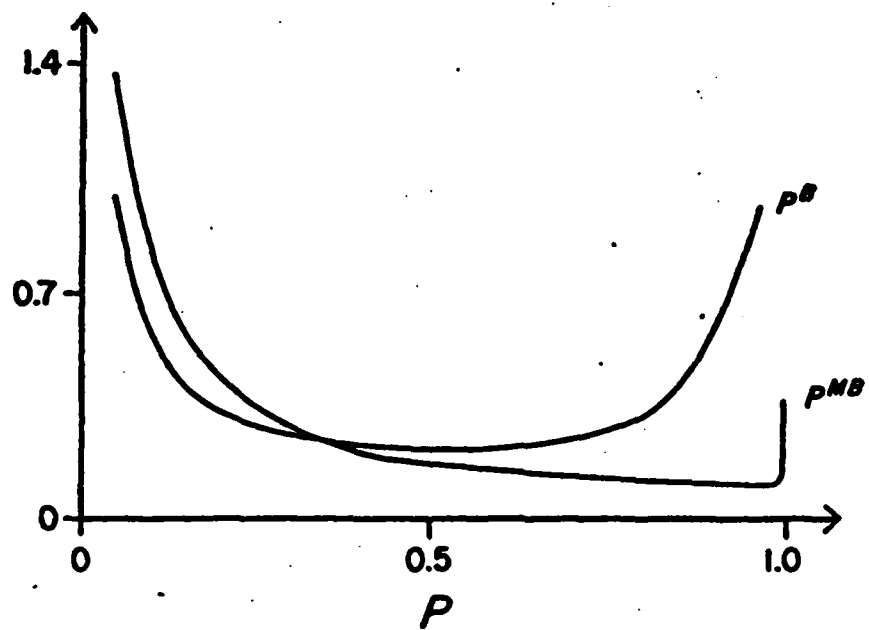


Figure 1: Shapes of the distribution of P for the BBD (P^B) and the MBBB (P^{MB}) models.

worse than the BBD error when estimating reach but the MBBD is clearly better at estimating the effective reach and single issue reach. Overall it is fair to say that the MBBD gives an improved fit over the BBD for magazine e.d. data.

The MBBD can also be utilized as a marginal distribution when estimating the reach and effective reach of advertising schedules in higher dimensions (Danaher 1987). In addition the MBBD has applications not only for magazine readership but also to television viewership and to newspapers, whose readership exhibits a high level loyalty.

The MBBD need not be restricted to fitting media exposure data. If we have a proportion of the population which *always* behaves in a specified way whilst the rest of the population has a certain *probability* of behaving in the specified way then the MBBD should be considered as a possible model instead of (say) the BBD. Such a situation arises in consumer purchasing behaviour, as the following example shows.

Morrison (1979) uses some purchase intention data of Juster (1966) in which Juster asks respondents to rate their purchase intentions for autos and appliances on a scale from 0 to 1 in 0.1 gradations. Zero is for no intention and 1 is for an almost certain purchase (see Morrison (1979) for details). A follow-up study was conducted in which Juster asked the respondents if they actually bought an auto or appliance. Morrison constructs a model to predict actual purchase behaviour from stated purchase intention in which he uses the BBD to fit the intention data.

The data have the following characteristics: a large group have no purchase intention, some people have a probable purchase intention and some people are certain to purchase in the future. Owing to the nature of people's intentions as revealed by Juster's data the MBBD is a good distribution to use to fit the data instead of the BBD.

In Table 2 we see that both the BBD and MBBD give an excellent fit to the

Table 2: Purchase intention data for appliances to be bought in the next 12 months. Sample size = 2688.

Intention Scale	Observed Frequency	Expected Frequency	
		BBD	MBBD
0.0	2377	2373.5	2377.0
0.1	87	85.4	90.7
0.2	57	45.8	47.7
0.3	29	32.3	32.9
0.4	23	25.6	25.5
0.5	22	21.8	21.0
0.6	21	19.5	18.1
0.7	14	18.3	16.1
0.8	11	17.9	14.8
0.9	17	19.0	14.1
1.0	30	24.9	30.0
Parameter Estimates		$\hat{\alpha}=0.035$ $\hat{\beta}=0.687$	$\hat{\alpha}=0.038$ $\hat{\beta}=0.873$ $\hat{\omega}=0.006$
χ^2 Goodness of Fit		8.43	5.02
d.o.f.		8	7
Likelihood Ratio Test Statistic			3.5
p-value			0.065

data as neither of their χ^2 goodness-of-fit statistics is significant. The χ^2 for the MBBD is smaller than that of the BBD, for which the likelihood ratio test (with $8 - 7 = 1$ df.) gives a marginally significant p-value of 0.065. Hence, the MBBD gives a better fit than the BBD for these data. A parallel study to Morrison's could, therefore, be done using his method but replacing the BBD with the MBBD.

5. Empirical Bayes Estimate of Single Issue Reach

We can think of $F_P^M(p)$ of (2.1) as the prior distribution of an individual's exposure to a single issue of a magazine. The posterior distribution of $P|X = x$ is $\frac{f(X=x|P=p)f_P^M(p)}{f_{MB}(X=x)}$. Under squared error loss the Bayes estimate of p is the mean of the posterior distribution (Berger 1980). If we estimate the parameters α , β , and ω from the data we get the empirical Bayes estimate (Casella 1985). The empirical Bayes estimate of single issue reach under the distribution of P in (2.1) is

$$\hat{p}_{MB}(x) = \begin{cases} \frac{(x + \hat{\alpha})}{(\hat{\alpha} + \hat{\beta} + k)}, & x = 0, \dots, k-1; \\ \frac{\frac{(\hat{\alpha} + k)}{(\hat{\alpha} + \hat{\beta} + k)}(1 - \hat{\omega})c(\hat{\alpha}, \hat{\beta}) + \hat{\omega}}{(1 - \hat{\omega})c(\hat{\alpha}, \hat{\beta}) + \hat{\omega}}, & x = k. \end{cases} \quad (5.1)$$

This estimator makes a great deal of sense since if $x = 0$ in the survey the MLE is $\hat{p}_{MLE}(0) = 0$, which implies that a person will *never* read the magazine, whereas a person may buy the magazine on impulse or glance at it in a doctor's surgery, for example. On the other hand, if $x = 4$ in the survey for a weekly magazine, the MLE implies a person *always* reads the magazine which is unlikely to be true since various reasons may prevent a person from reading a particular issue of a magazine. That is, this estimator tends to moderate, from the extreme, an individual's exposure probability.

We calculated the empirical Bayes estimates of single issue reach for the *Na-*

tional Business Review thus:

$$\hat{p}_{MB}(x) = \begin{cases} 0.163(x + 0.024), & x = 0, \dots, 3; \\ 0.9799, & x = 4. \end{cases} \quad (5.2)$$

Such an estimator is useful in simulation studies, for instance, if it were required to estimate the audience for a schedule which combined different media types and no exposure model were available. Suppose, for example, a schedule has 4 insertions in particular magazine and three insertions in a particular television time slot. Extract the response to Q2 (call it x_{mag}) for the magazine, then calculate $\hat{p}_{MB}(x_{mag})$. Now simulate 4 Bernoulli trials with probability of success $\hat{p}_{MB}(x_{mag})$. Keeping with the same individual extract that person's probability of viewing television in the desired time slot and conducted 3 Bernoulli trials as before. The total number of successes for the 7 trials is the individual's total exposure. Repeat this procedure for each individual in the survey. Gifford (personal communication) used simulation with the MLE ($\hat{p}_{MLE}(x) = x/k$) of single issue reach to estimate the audience for a combined magazine/television schedule in such a way. Table 3 shows the results of averaging 50 simulated e.d.'s for the *National Business Review* using the personal probabilities given by the MLE, the BBD empirical Bayes estimator ($\hat{p}_B(x) = (\hat{\alpha} + x)/(\hat{\alpha} + \hat{\beta} + k) = .228(x + .012)$) and the MBBD empirical Bayes estimator (given by (5.2)).

All the χ^2 statistics are significant (p-value < 0.01) whereas the χ^2 for the *National Business Review* using the MBBD model (from Table 1) is 2.6 (p-value > 0.1) so that simulation methods are inferior to fitting a model in this example. Nonetheless, the empirical Bayes estimate of personal probability, based on the MBBD model, gives the best results of the three $\hat{p}(x)$'s used. From this example we infer it is best to use simulation only when it is impossible to construct an e.d. model, and that the empirical Bayes estimate, based on the MBBD in (5.1), is likely to give better results than simulations based on the MLE.

Table 3: E.d.'s obtained by averaging 50 simulated e.d.'s using $\hat{p}_{MLE}(x)$, $\hat{p}_B(x)$ and $\hat{p}_{MB}(x)$ as estimates of personal probability.

	Exposures					χ^2
	0	1	2	3	4	
Observed	4961	90	47	12	95	
MLE	4991.9	48.4	36.6	21.9	102.3	42.2
BBD	4945.8	101.7	39.4	43.9	70.6	33.4
MBBD	4935.1	132.3	29.1	16.2	88.3	21.9

Acknowledgement

This paper is based on part of the author's Ph.D. dissertation for which the author gratefully acknowledges the assistance of Professor Duane Meeter. The author would also like to thank Ron Stroeven and Clive Gifford of AGB: McNair Surveys New Zealand Ltd. Ron provided the data and Clive gave helpful insights into techniques for simulating magazine exposure distributions.

References

- Be-ger, J.O. (1980). *Statistical Decision Theory*. Springer-Verlag, New York.
- Casella, G. (1985). An introduction to empirical Bayes data analysis. *The American Statistician*. 39, 83-90.
- Chandon, J-L. J. (1976). A comparative study of media exposure models. Ph.D. dissertation, Northwestern University.
- Danaher, P.J. (1987). Estimating multidimensional tables from survey data: Predicting magazine audiences. Ph.D. dissertation, Florida State University.
- Ehrenberg, A.S.C. (1975). *Data Reduction Analysis and Interpretation of Statistical Data*. Wiley, New York.
- Giesbrecht, F. and Kempthorne, O. (1976). Maximum likelihood estimation in the three-parameter lognormal distribution. *J. R. Statist. Soc. B*, 38, 257-264.
- Griffiths, D.A. (1975). Likelihood estimation for the beta-binomial distribution and an application to the household distribution of the total number of cases of a disease. *Biometrics*. 29, 637-648.
- Ishii, G. and Hayakawa, R. (1960). On the compound binomial distribution. *Annals of the Institute of Statistical Mathematics*. 12, 69-80.
- Johnson, N.L. and Kotz, S. (1969). *Discrete Distributions*. John Wiley and Sons, New York.
- Juster, F.T. (1966). Consumer buying intentions and purchase probability. *J. American Statist. Assoc.* 61, 658-696.
- Kleinman, J.C. (1973). Proportions with extraneous variance: single and indepen-

- dent samples. *J. American Statist. Assoc.* 68, 46-54.
- Kulldorff, G. (1957). On conditions for consistency and asymptotic efficiency of maximum likelihood estimates. *Skandinavisk Aktuarietidskrift.* 40, 129-144.
- Metheringham, R.A. (1984). Measuring the net cumulative coverage of a print campaign. *J. of Advertising Research* 4, 23-28.
- Moore, D.F. (1987). Modelling the extraneous variance in the presence of extra-binomial variation. *Applied Statistics.* 31, 1, 8-14.
- Morrison, D.G. (1979). Purchase intentions and purchasing behavior. *J. of Marketing.* 43, 65-74.
- Naples, M.J. (1979). *Effective Frequency: The Relationship between Frequency and Advertising Awareness.* Association of National Advertisers, New York.
- Rust, R.T. and Klompmaker, J.E. (1981). Improving the estimation procedure for the beta binomial t.v. exposure model. *Journal of Marketing Research.* 18, 442-448.
- Sabavala, D.J. and Morrison, D.G. (1977). Television show loyalty: a beta- binomial model using recall data. *Journal of Advertising Research.* 17, 35-43.
- Skellam, J.G. (1948). A probability distribution derived from the binomial distribution by regarding the probability of success as a variable between sets of trials. *J. R. Statist. Soc. B,* 10, 257-265.