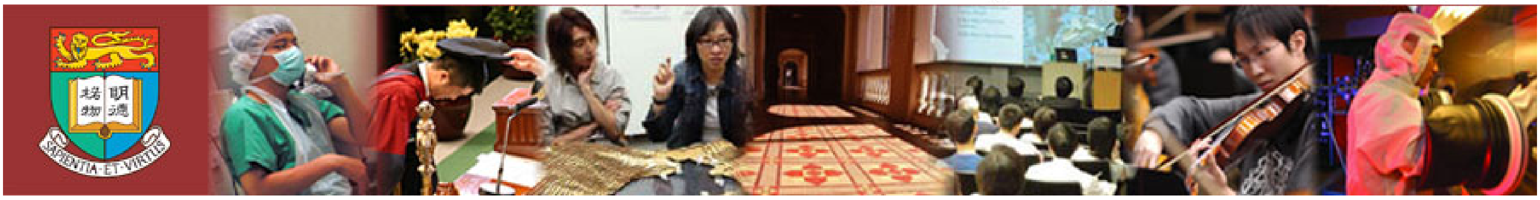

Title	Bayesian analysis of errors-in-variables in binary regression models
Author(s)	Tang, PK; Bacon-Shone, J
Citation	Research Report, n. 12, p. 1-32
Issued Date	1992-05
URL	http://hdl.handle.net/10722/60986
Rights	

RESEARCH REPORT

Serial No. 12

May 1992

BAYESIAN ANALYSIS OF ERRORS-IN-VARIABLES

IN BINARY REGRESSION MODELS

by

P.K. Tang and J. Bacon-Shone

THE UNIVERSITY OF HONG KONG
DEPARTMENT OF STATISTICS

Pokfulam Road, Hong Kong. Tel: (852) 859 2466 Fax: (852) 858 9041
E-Mail: STATIST@HKUCC.HKU.HK or STATIST@HKUCC.BITNET

BAYESIAN ANALYSIS OF ERRORS-IN-VARIABLES IN BINARY REGRESSION MODELS

P.K. Tang and John Bacon-Shone
Department of Statistics, University of Hong Kong

Keywords: errors-in-variables, binary regression, Bayesian inference, loss function, non-informative prior.

Abstract

There has been considerable research done on the problems of errors-in-variables for linear regression, including a Bayesian solution by Lindley and El-Sayyad (1968). Recently, interest has extended to binary regression and in particular probit regression. Burr (1985) performed frequentist analysis of Berkson's error in probit regression and found that the MLE does not always exist in finite samples. In this paper, we show that it is the tail behaviour of the likelihood that causes the problem and this in turn makes Bayesian estimation inadmissible if improper priors are used. Two non-informative priors are derived and simulation results indicate that the Bayesian solutions are generally superior to various likelihood based estimates, including the modified MLE proposed by Burr. It is further shown that the estimation problem vanishes if there are replicates and that the logistic model has the same behaviour as the probit model.

1 Introduction

The probit model is given by

$$p(Y=1|x) = \int_{-\infty}^{\alpha+\beta x} \frac{1}{\sqrt{2\pi}}\exp(-\frac{t^2}{2})\, dt \qquad -\infty < \alpha < \infty, \quad \beta > 0$$

$$= \Phi(\alpha+\beta x)$$

where Y is a binary success/failure variable and x is the known covariate. For the Berkson error model (Berkson, 1950), a surrogate of x, z, is observed instead of x and

$$X = Z + U$$

where

$$U \sim N(0,\sigma_u^2)\ .$$

Then

$$p(Y=1\,|\,z) = \int \Phi[\alpha + \beta x]\, f_{x|z}(x\,|\,z)\, dx$$

$$= \Phi\left[\alpha_r + \beta_r z\right]$$

where

$$\alpha_r = \frac{\alpha}{\sqrt{1 + \beta^2 \sigma_u^2}}$$

and

$$\beta_r = \frac{\beta}{\sqrt{1 + \beta^2 \sigma_u^2}}$$

Note that

$$\gamma = -\frac{\alpha}{\beta} = -\frac{\alpha_r}{\beta_r}$$

is not affected by the error. Therefore, the probit model considered in the sequel is given by

$$p(Y=1\,|\,z) = \Phi\left(\frac{\beta}{\sqrt{1+\beta^2\sigma_u^2}}(z-\gamma)\right).$$

There is a problem in that

$$\lim_{\beta\to\infty} \Phi\left(\frac{\beta}{\sqrt{1+\beta^2\sigma_u^2}}(z-\gamma)\right) = \Phi\left(\frac{z-\gamma}{\sigma_u}\right)$$

which does not go to 1 or 0 for fixed z ,γ and σ_u.

For binary data, the contribution of an observation to the likelihood is either the probability of success or the probability of failure, which is 1 minus the probability of success, depending on whether the observation is a success or a failure. In the current situation, the fact that the probability of success does not go to 1 or 0 implies that the probability of failure is also bounded within 0 and 1. This means that the likelihood for β does not decay to zero in the tail for any finite sample. We define a likelihood to be improper if the likelihood function, when treated as a function of a parameter, does not decay to zero as the parameter goes to infinity for a fixed sample. Note further that the larger the variance of the error distribution is, the more serious the problem it brings.

2 Probit Model

Burr (1985) noted that the problem is unidentifiable in β and σ_u. In the classical approach, it is assumed that σ_u is known to avoid the problem of unidentifiability (Carroll, Spiegelman, Lan, Bailey and Abbott, 1984 and Burr, 1985). For this problem, we follow Burr to assume σ_u is known. Let $\theta = (\gamma, \beta)^T$. The simplest approach is to ignore the error and treat the observed z's as true x's. The ordinary probit regression MLE is called the naive estimate. That is, assuming the underlying model is

$$p(Y=1\,|z) = \Phi\big(\beta_0(z - \gamma_0)\big).$$

2.1 Maximum Likelihood Estimators

The naive estimate of γ is consistent but that of β is inconsistent as

$$\hat{\beta}_r \rightarrow \frac{\beta}{\sqrt{1 + \beta^2\sigma_u^2}}$$

in probability.

Note that

$$\beta = \frac{\beta_r}{\sqrt{1 - \beta_r^2\sigma_u^2}} .$$

Thus, using the invariance property of the MLE, we obtain the error-in-x MLE, or the true MLE, $\hat{\beta}_1$, as

$$\hat{\beta}_1 = \begin{cases} \dfrac{\hat{\beta}_0}{\sqrt{1 - \hat{\beta}_0^2\sigma_u^2}}, & \text{if } \hat{\beta}_0 < \dfrac{1}{\sigma_u} \\ \infty & \text{if } \hat{\beta}_0 \geq \dfrac{1}{\sigma_u} \end{cases}$$

The major drawback of $\hat{\beta}_1$ is that it does not exist when $\hat{\beta}_0 \geq \sigma_u^{-1}$ as the likelihood attains its maximum at infinity.

Burr (1985) proposed a modified MLE which always exists and reduces the skewness in the distribution of $\hat{\beta}_1$. The modified MLE, $\hat{\beta}_2$, is given by

$$\hat{\beta}_2 = \frac{\hat{\beta}_0}{\exp(-.5\,\hat{\beta}_0^2\sigma_u^2)} .$$

The major drawback of this estimator is that

$$\hat{\beta}_2 \rightarrow \frac{\beta_r}{\exp(-.5\,\beta_r^2\sigma_u^2)}$$

in probability. Therefore it is inconsistent and biased downward in absolute value. To summarise, for β, both the naive MLE and the modified MLE are inconsistent, and the true MLE may not exist in finite samples.

2.2 Bayesian Approach

In order to compare the various maximum likelihood approaches and the Bayesian approach, it is assumed that σ_u is known though this is unnecessary in a strict Bayesian sense. Given that γ is little affected by error, we concentrate on the prior distribution for β. Recall that the likelihood for β is improper. Hence an improper prior for β would result in an improper posterior for β. Therefore if a uniform prior or an inverse prior for β is used, the posterior distribution of β will be improper. The dangers of using improper priors are discussed in detail in Dawid, Stone and Zidek (1973). If an unbounded loss function is used, an infinite Bayes estimate will incur an infinite loss. This implies that the use of an improper prior and a square loss function is inadmissible, because the Bayes estimate resulting from a squared loss function is the posterior mean and the mean of an improper distribution is undefined.

When more prior knowledge of β is available, a class of proper priors is the half normal family

$$\beta \sim N\left(0\ ,\ \frac{1}{k^2 \sigma_u^2}\right), \qquad \beta > 0.$$

where the hyperparameter k corresponds to $Z_{p/2}$ where p is the prior probability that $\beta > \sigma_u^{-1}$.

When p is small, i.e. it is less likely that $\beta > \sigma_u^{-1}$, then $Z_{p/2}$ would be large so that $1/k^2\sigma_u^2$ will be small. One the other hand, the prior will tend to the uniform distribution when the possibility that $\beta > \sigma_u^{-1}$ is very likely. However, it is true that it is difficult to determine k in many practical problems.

A simulation study of the performance of the posterior mean resulting from the half normal priors and various other priors can be found in Tang (1992).

Empirical Bayes analysis (Berger, 1985, Sec. 4.5) suggests estimating k from the marginal probability of data given k

$$m(data\,|\,k) \;=\; \int_\gamma\int_\beta L(data\,|\,\gamma,\beta)\ \frac{k\sigma_u}{\sqrt{2\pi}}\ \exp(-\frac{1}{2}k^2\beta^2\sigma_u^2)\ d\gamma\ d\beta\ .$$

k is found by taking the value that maximizes m(data | k). Unfortunately, preliminary simulation studies show that the variability in k is so large, particularly when β is large, as to make this method useless. In fact, it is expected that this method will not work well since the data do not provide much information about β because the likelihood is improper. Any effort to extract information about β from the likelihood is thus futile. Therefore, empirical Bayes procedures do not make any sense in such situation.

We need to find a proper prior for β by other means. If we have no further information about β, it is difficult to find a proper prior for it. One way to get around the difficulty is to put a noninformative uniform prior on β_r. This is a standard Bayes solution for models with a restricted parameter space (Berger, 1980). Since β_r is bounded between 0 and 1, this gives

$$f(\beta \mid \sigma_u) \propto f(\beta_r) \left|\frac{\partial \beta_r}{\partial \beta}\right| \tag{1}$$

$$= (1 + \beta^2 \sigma_u^2)^{-3/2} \quad .$$

which is $O(\beta^{-3})$.

This is a proper prior in β given that σ_u is known. Moreover, it is a uniform prior in β when σ_u is 0. When there is no measurement error, the probit model can be solved using an improper prior for β. Recall that the density function for y conditional on z is a constant when $\beta \to \infty$ for given z and γ. Then the conditional likelihood, as a function of β, also tends to a constant in the tail. Hence, any proper prior distribution of β that decays to 0 at the tail or any prior with a finite range will make the resulting posterior distribution of β proper. Obviously, the prior distribution given in (1) has such a property.

Another approach is to consider the Jeffreys prior (Jeffreys, 1946) for γ and β which is proportional to the square root of the determinant of the expected Fisher information matrix. Given a design measure putting n_i weight at k distinct design points z_i, $i=1,\ldots,k$, $\sum n_i = 1$. Then the Jeffreys prior is

$$p(\gamma,\beta) \propto \frac{\beta\sqrt{ts}}{(1 + \beta^2\sigma_u^2)^2} \tag{2}$$

where

$$w_i = \frac{\phi_i^2}{\Phi_i(1-\Phi_i)} \; ; \qquad t = \sum_{i=1}^{k} n_i\, w_i \; ;$$

$$\bar{z} = \frac{\sum_{i=1}^{k} n_i\, w_i\, z_i}{\sum_{i=1}^{k} n_i\, w_i} \; ; \qquad s = \sum_{i=1}^{k} n_i\, w_i\, (z_i - \bar{z})^2 \; ;$$

$$\phi_i = \phi\left(\frac{\beta(z-\gamma)}{\sqrt{1 + \beta^2\sigma_u^2}}\right) ; \quad \Phi_i = \Phi\left(\frac{\beta(z-\gamma)}{\sqrt{1 + \beta^2\sigma_u^2}}\right) .$$

Note that this is a proper prior in γ and in β. In particular, (2) is $O(\beta^{-3})$ in the tail for fixed γ.

It is desirable to perform Bayesian inference for parameters based on some type of noninformative prior. In situations such that a uniform prior for the parameters or the log of the parameters would yield improper posterior densities, a commonly used alternative is to use Bernardo's reference prior (Bernardo 1979, Berger and Bernardo 1989, and Ye and Berger 1991). If there is no nuisance parameter, Bernardo's reference prior is simply Jeffreys prior. For the current model, since σ_u is assumed to be known, there is no nuisance parameter and hence the Jeffreys prior can also be viewed as Bernardo's reference prior.

Another issue regarding the point estimate is the form of the loss function. Since there is substantial uncertainty about the correctness of the informative prior (1) and of the Jeffreys prior (2), it will be a wise choice to find a loss function that is loss robust (Kadane and Chuang, 1978, Berger, 1984, 1985) so that we will not get a very different estimate when the prior is changed slightly. According to Kadane and Chuang (1978), the squared loss function is unstable and they recommended to use a bounded loss function. One of the most common Bayesian point estimates is the posterior mode. Note that the posterior mode is derived from the bounded loss function

$$L(\theta,a) = \begin{cases} 0 & \text{if} \quad \| \theta - a \| \le \epsilon \\ 1 & \text{if} \quad \| \theta - a \| > \epsilon \end{cases} .$$

where θ is the true parameter, a is the point estimate of θ and ϵ is a small positive constant. When we try to select a to minimize the loss w.r.t. the above loss function, it is easily found that the decision we take is to use the posterior mode as the point estimate (Box and Tiao, 1973, Lee, 1989). Recall that a Bayes estimate resulting from a bounded loss function and a proper prior is admissible (Lindley and Smith, 1972). Therefore, the posterior mode, if it exists, resulting from using (1) as the prior for β in conjunction with any proper prior for γ is admissible. Bayes decisions arising from an unbounded loss function can change enormously when the distribution of the random variable changes slightly (Kadane and Chuang, 1978). Further discussion of the dangers of using an unbounded loss function can be found in Kadane and Chuang (1978) and Bayesian estimation using bounded loss function is also discussed by Smith (1980).

One must be reminded that though the likelihood is bounded below by some non-zero bound for fixed sample size, the bound decreases to 0 as the sample size gets large which implies that the problem of improper likelihood decreases as the sample size increases. In practice, when the sample size is very large, the chance of getting an infinite MLE of β becomes small.

Aitchison and Lauder (1979) noted the improper behaviour of the likelihood of the current model. Their concern was to find the posterior distribution of the parameters and their approach was to make use of the Bayesian form of large sample maximum likelihood theory (Lindley, 1961). We feel that their approach is rather ad hoc and may not work well for small samples for two reasons. The first reason is that the large sample property only holds for very large samples and the second reason is that small samples are likely to give an infinite maximum likelihood estimate of the slope parameter.

3 Logistic Model

Another common binary regression model is the logistic model which specifies the probability of success conditional on the covariate x as

$$p(Y=1\,|x) = \frac{\exp(\beta(x-\gamma))}{1+\exp(\beta(x-\gamma))} \qquad 0 < \beta < \infty, \quad -\infty < \gamma < \infty \quad .$$

The probability of success conditional on z becomes

$$p(Y=1\,|\,z) = \int_{-\infty}^{\infty} \frac{\exp(\beta(z-\gamma)+\beta u)}{1 + \exp(\beta(z-\gamma)+\beta u)}\, p(u)\, du \quad .$$

If u ~ N(0, σ_u^2), the probability of success conditional on z can be approximated (Abramowitz and Stegun, 1970) by

$$\int_{-\infty}^{\infty} e^{-u^2} f(u)\, du \approx \sum_{i=1}^{N} w_{N,i}\, f(s_{N,i})$$

where $s_{N,i}$ ($i=1,\ldots,N$) are the roots of the Hermite polynomial equation $H_N(x)=0$ and

$$w_{N,i} = \frac{2^{N-1} N! \sqrt{\pi}}{N^2 H_{N-1}(s_{N,i})} .$$

The approximation of the integral will be exact, (Davis and Rabinowitz, 1984, p. 228), if N $\rightarrow \infty$. i.e.

$$\int_{-\infty}^{\infty} e^{-u^2} f(u)\, du = \lim_{N\rightarrow\infty} \sum_{i=1}^{N} w_{N,i}\, f(s_{N,i}) .$$

if for all sufficiently large values of $|u|$, f(u) satisfies the inequality

$$|f(u)| \leq \frac{e^{u^2}}{|u|^{1+\rho}}, \qquad \rho > 0 . \tag{3}$$

In the current situation,

$$|f(u)| = \left|\frac{1}{\sqrt{\pi}}\right| \left|\frac{\exp(\beta(z-\gamma)+\beta\sigma_u\sqrt{2}u)}{1+\exp(\beta(z-\gamma)+\beta\sigma_u\sqrt{2}u)}\right| .$$

Note that

$$\left|\frac{\exp(\beta(z-\gamma)+\beta\sigma_u\sqrt{2}u)}{1+\exp(\beta(z-\gamma)+\beta\sigma_u\sqrt{2}u)}\right| \leq 1, \qquad -\infty < u < \infty .$$

Therefore, condition (3) is satisfied because r.h.s. of (3) is unbounded in u. Then we have

$$p(Y=1\,|\,z) = \lim_{N\rightarrow\infty} \frac{1}{\sqrt{\pi}} \sum_{i}^{N} \left[w_{N,i} \frac{\exp(\beta(z-\gamma)+\beta\sigma_u\sqrt{2}s_{N,i})}{1+\exp(\beta(z-\gamma)+\beta\sigma_u\sqrt{2}s_{N,i})}\right] .$$

When $z = \gamma$, then

$$p(Y=1\,|\,z) = \lim_{N\rightarrow\infty} \frac{1}{\sqrt{\pi}} \sum_{i=1}^{N} \left(w_{N,i} \frac{\exp(\beta\sigma_u\sqrt{2}s_{N,i})}{1+\exp(\beta\sigma_u\sqrt{2}s_{N,i})}\right) .$$

The following properties of the Hermite polynomial equation are useful:

1. The roots are symmetric about zero so that the sum of the roots equals 0.
2. Sum of the $w_{N,i}$ equals $\sqrt{\pi}$.
3. $w_{N,i}$ at skew-symmetric roots are equal.

Then, we have

$$p(Y=1\,|\,z) = \lim_{N\to\infty} \frac{1}{\sqrt{\pi}} \sum_{i=1}^{N} \left(w_{N,i} \frac{\exp(\beta\, \sigma_u \sqrt{2} s_{N,i})}{1+\exp(\beta\, \sigma_u \sqrt{2} s_{N,i})} \right)$$

$$= \lim_{N\to\infty} \frac{1}{\sqrt{\pi}} \sum_{i=1}^{N} \left(\frac{w_{N,i}}{2} \right)$$

$$= \lim_{N\to\infty} \frac{1}{\sqrt{\pi}} \frac{\sqrt{\pi}}{2}$$

$$= \frac{1}{2} \;.$$

That is, γ is little affected by the measurement error.
Note that in the case of probit model, we have, for $z = \gamma$,

$$p(Y=1\,|\,z) = \Phi(0) = \frac{1}{2}$$

Now, consider $z - \gamma = c > 0$. The general idea can be easily illustrated by using an even grid integration rule though it is the same idea if an odd grid rule is used. Using the skew-symmetric property of the roots, the $s_{N,i}$ are all positive in the following derivation.

$$p(Y=1 \mid z) = \lim_{N\to\infty} \frac{1}{\sqrt{\pi}} \sum_{i=1}^{\frac{N}{2}} w_{N,i} \left[\frac{\exp(\beta c + \beta\, \sigma_u \sqrt{2} s_{N,i})}{1 + \exp(\beta c + \beta\, \sigma_u \sqrt{2} s_{N,i})} + \frac{\exp(\beta c - \beta\, \sigma_u \sqrt{2} s_{N,i})}{1 + \exp(\beta c - \beta\, \sigma_u \sqrt{2} s_{N,i})} \right] .$$

As $N \to \infty$, some $s_{N,i}$ will be so large as to make the value of the difference in the exponent of the second term become negative. As a result, as $\beta \to \infty$, the second term goes to 0 so that the conditional probability of success is less than 1. Following the argument given in the probit model, it can be seen that the likelihood in the logit model is also bounded within 0 and 1 for the given sample and is improper for β. Furthermore, the conditional probability goes to zero at an exponential rate as α goes to $\pm\infty$ or as γ goes to $\pm\infty$ for fixed β and z. This is additional evidence that the behaviour of the logistic model is similar to that of the probit model.

4 Binomial Model

In the Binomial Model, the random experiment consists of n repeated independent Bernoulli Trials when the probability of success at each individual trial is the same. Let r be the number of successes in n independent Bernoulli Trials. Then

$$p(Y=r \mid x) = \binom{n}{r} p^r \, (1-p)^{n-r}$$

where p is the probability of success at each trial.

For the probit model, the probability of getting r successes out of n trials conditional on x is given by

$$p(Y=r \mid x) = \binom{n}{r} p^r (1-p)^{n-r}$$

$$= \binom{n}{r} [\Phi(\beta(x-\gamma))]^r [1-\Phi(\beta(x-\gamma))]^{n-r} .$$

The probability of getting r successes out of n trials conditional on z is given by

$$p(Y=r \mid z)$$

$$= \int_{-\infty}^{\infty} \binom{n}{r} [\Phi(\beta(z-\gamma)+\beta u)]^r [1-\Phi(\beta(x-\gamma)+\beta u)]^{n-r} \times$$

$$\frac{1}{\sqrt{2\pi}\,\sigma_u} e^{-\frac{u^2}{2\sigma_u^2}} du .$$

Note that

$$| f(u) | = \left| \frac{\binom{n}{r}}{\sqrt{\pi}} [\Phi(\beta(z-\gamma)+\beta\sigma_u\sqrt{2}u)]^r \times \right.$$

$$\left. [1-\Phi(\beta(z-\gamma)+\beta\sigma_u\sqrt{2}u)]^{n-r} \right|$$

and

$$\left| [\Phi(\beta(z-\gamma)+\beta\sigma_u\sqrt{2}u)]^r [1-\Phi(\beta(z-\gamma)+\beta\sigma_u\sqrt{2}u)]^{n-r} \right| \le 1$$

for $-\infty < u < \infty$.

Therefore, $|f(u)|$ is bounded and condition (3) is satisfied and we can use the Gauss Hermite rule to find

$$p(Y=r \mid z) = \lim_{N\to\infty} \frac{\binom{n}{r}}{\sqrt{\pi}} \sum_{i=1}^{N} w_{N,i} \left\{ \left[\Phi(\beta(z-\gamma)+\sqrt{2}\sigma_u\beta s_{N,i})\right]^r \right.$$

$$\left. \times \left[1-\Phi(\beta(z-\gamma)+\sqrt{2}\sigma_u\beta s_{N,i})\right]^{n-r} \right\} .$$

Consider $z - \gamma = c > 0$ and $s_{N,i}$'s > 0. Using an even grid rule, we have

$$p(Y=r \mid z)$$

$$= \lim_{N \to \infty} \frac{\binom{n}{r}}{\sqrt{\pi}} \sum_{i=1}^{\frac{N}{2}} w_{N,i} \Big\{ \left[\Phi(\beta c + \sqrt{2}\sigma_u \beta s_{N,i})\right]^r \times \left[1 - \Phi(\beta c + \sqrt{2}\sigma_u \beta s_{N,i})\right]^{n-r} + \left[\Phi(\beta c - \sqrt{2}\sigma_u \beta s_{N,i})\right]^r \left[1 - \Phi(\beta c - \sqrt{2}\sigma_u \beta s_{N,i})\right]^{n-r} \Big\} \quad (4)$$

We consider the behaviour of (4) when $\beta \to \infty$.

1. If $r > 0$ and $r < n$, the first term inside the curly brackets tends to 0. $\beta c - \sqrt{2}\sigma_u\beta s_{N,i}$ goes to $-\infty$ or ∞, depending on the sign and magnitude of c and $s_{N,i}$. In either case, the second term inside the curly brackets goes to 0 at an exponential rate. This will make the conditional probability go to 0 when $\beta \to \infty$.

2. If $r = 0$, the first term inside the curly brackets tends to 0. The second term will go either to 0, if $\beta c - \sqrt{2}\sigma_u\beta s_{N,i} \to \infty$, or to 1, if $\beta c - \sqrt{2}\sigma_u\beta s_{N,i} \to -\infty$. This implies that the conditional probability does not go to 0 and it is less than 1.

3. If $r = n$, the first term inside the curly brackets tends to 1. The second term will go either to 0, if $\beta c - \sqrt{2}\sigma_u\beta s_{N,i} \to -\infty$, or to 1, if $\beta c - \sqrt{2}\sigma_u\beta s_{N,i} \to \infty$. This implies that the conditional probability does not go to 0 and it is less than 1.

Hence, if the outcomes are not all successes and not all failures, the likelihood will decay to 0 when β goes to infinity at an exponential rate, making the problem of improper likelihood of β in the Bernoulli Trial disappears. Furthermore, when γ goes to $\pm\infty$, the conditional probability goes to zero when other parameters are fixed under the 3 situations. Hence, using an improper prior together with the quadratic loss will not incur infinite loss.

It can be shown that the logistic model has the same property.

5 Simulation

Following Burr, γ and σ_u^2 were assigned to 1.0 and 0.26 respectively. In order to assess the behaviour of the estimators, a series of simulations, for β = 0.4, 1.4, 2.0 and 5.0, were run. For small values of β, e.g. β=0.4, the effect of the measurement error on the parameters estimation is small. β=1.4 is the most common setting in simulation studies (such as Carroll *et al*, 1984 and Burr, 1985) as this is close to the findings of the Framingham Heart Study (Gordon and Kannel, 1968). The boundary value of the naive MLE such that the true MLE becomes unbounded is about 2.0 for σ_u^2 = 0.26. A large value of β, such as 5.0, provides evidence regarding the performances of the estimators in more difficult situations. For each value of β, 1000 simulations with sample size = 40, 80, 160, 320, 480 and 960 were run. The design points were derived from a two-point locally D-optimal design. The sampling properties of the estimators are examined. Since γ is little affected by error so we concentrate on the estimation of β.

Two Bayesian solutions are considered. The first one corresponds to the joint posterior mode of γ and β. The other employs the marginal posterior mode of β and of γ. Both the joint mode and the marginal modes can be used as Bayes point estimates (DeGroot, 1970). In the case of large samples, the posterior joint mode is close to the posterior marginal modes.

6 Results

Tables 1(a)-4(b) provide a summary of the simulation results. Figures 1-4 show the plots of the bias of β estimates against the log of the sample size for the estimators and figures 5-8 show the plot of the log of the MSE of β estimates against the log of the sample sizes. As the performances of the two Bayesian solutions are very similar, only the joint modes are used in the plots. The plots for γ estimates only show that all methods are very similar so they are not shown. It can be seen that both the bias and the MSE of the Bayes estimates decrease as the sample size increases while the biases of both the naive MLE and the modified MLE increase as the sample size gets large. The performance of the Bayesian approaches are very similar. The number of finite true MLE increases while the bias and the MSE of finite true MLE decreases as sample size increases. The Bayesian estimates perform better than the true MLE, except under very limited conditions, i.e. when both sample size and β are moderate. When β is 0.4, the naive MLE seems to be slightly better than others except when the sample size is 40. In the case of small samples, (n=40), or large samples, (n ranges from 320 to 960), or when β is large, ($\beta = 5.0$), the Bayesian estimates are the best. The modified MLE is the best when sample size is moderate, in the range of 80 to 480, and moderate values of β, $\beta = 1.4$ and 2.0. Furthermore, the Bayes estimates have the smallest bias in most cases.

7 Conclusion

The benefits of the Bayesian approach, with a proper prior, to this problem can be stated as:

1. the Bayesian solutions are always finite.
2. the Bayesian solutions resulting from a bounded loss function and a proper prior will be admissible and consistent. Also, it is better than the frequentist approaches under nearly all conditions.
3. any further information about β can be specified through a suitably defined prior.
4. when substantial prior information on β is obtained so that the prior puts sufficient probability near the true value, the Bayes estimate must work well.

References

Abramowitz, M. and Stegun, I. (1970). Handbook of Mathematical Functions. Dover Publications, Inc, New York.

Aitchison, J. and Lauder, I.J. (1979). Statistical Diagnosis from Imprecise Data. Biometrika, 66, 475-483.

Berger, J.O. (1980). Statistical Decision Theory: Foundations, Concepts and Methods. Springer-Verlag, New York Inc.

Berger, J.O. (1985). Statistical Decision Theory and Bayesian Analysis. Springer-Verlag, New York Inc.

Berger, J.O. and Bernardo, J. (1989). Estimating a Product of Means: Bayesian Analysis with Reference Priors. Journal of the American Statistical Association, 84, 200-207.

Berkson, J. (1950). Are There Two Regressions? Journal of the American Statistical Association, 45, 164-180.

Bernardo, J.M. (1979). Reference Posterior Distribution for Bayesian Inference. Journal of the Royal Statistical Society, 41, 113-147.

Box, G.E.P. and Tiao, G.C. (1973). Bayesian Inference in Statistical Analysis. Addison-Wesley.

Burr, D. (1985). On Errors-in-variables in Binary Regression - Berkson Case. Ph.D. thesis, Stanford University.

Carroll, R.J., Spiegelman, C.H., Lan, K.K., Bailey, K.T., Abbott, R.D. (1984). On Errors In Variables for Binary Regression Models. Biometrika, 71, 19-25.

Davis, P.J. and Rabinowitz, P. (1984). Methods of Numerical Integration, Second Edition. Academic Press Inc.

Dawid, A.P., Stone, M. and Zidek, J.V. (1973). Marginalization Paradoxes in Bayesian and Structural Inference. Journal of the Royal Statistical Society, Series B, 35, 189-233.

DeGroot, M.H. (1970). Optimal Statistical Decisions. McGraw-Hill, New York.

Gordon, T. and Kannel, W.E. (1968). The Framingham Study, introduction and general background in the Framingham study, §§ 1,2. Bethesda, Maryland: National Heart, Lung, and Blood Institute.

Jeffreys, H. (1946). An Invariant Form for the Prior Probability in Estimation Problems. Proceedings of the Royal Statistical Society of London, Series A, 196, 453-461.

Kadane, J.B. (1984). Robustness of Bayesian Analysis. North-Holland, Amsterdam.

Kadane, J.B. and Chuang, D.T. (1978). Stable Decision Problems. Annals of Statistics, 6, 1095-1111.

Lee, P.M. (1989). Bayesian Statistics. Oxford University Press.

Lindley, D.V. (1961). The Use of Prior Probability Distributions in Statistical Inference and Decisions. Proceedings of the 4th Berkeley Symposium, 453-468.

Lindley, D.V. and Smith, A.F.M. (1972). Bayes Estimates for the Linear Model. Journal of the Royal Statistical Society, Series B, 34, 1-41.

Smith, J.Q. (1980). Bayes Estimates Under Bounded Loss. Biometrika, 67, 629-638.

Tang P.K. (1992). Bayesian Analysis of Errors-in-variables in Generalized Linear Models. Ph.D. Thesis, University of Hong Kong.

Ye, K. and Berger J.O. (1991). Noninformative Priors for Inferences in Exponential Regression Models. 78, 645-656.

Tables

Table 1a

Monte Carlo study with 1000 simulations, and

$\sigma_u^2 = 0.26$, $\gamma = 1.0$, $\beta = 0.4$,

sample sizes = 40, 80, 160,

$x_1 = -1.904$, $x_2 = 3.904$

Sample sizes			n=40	n=80	n=160
Join Post. Mode $\beta_r \sim U(0, \sigma_u^{-1})$	$\hat{\gamma}$	Bias	0.002621	-0.020355	-0.011718
		MSE	0.218277	0.108167	0.052221
		Variance	0.218271	0.107752	0.052084
	$\hat{\beta}$	Bias	0.013679	0.007661	0.003832
		MSE	0.005412	0.002374	0.001134
		Variance	0.005225	0.002315	0.001120
Marginal Post. Mode $\beta_r \sim U(0, \sigma_u^{-1})$	$\hat{\gamma}$	Bias	-0.000292	-0.020714	-0.011780
		MSE	0.252341	0.110700	0.052775
		Variance	0.252341	0.110271	0.052637
	$\hat{\beta}$	Bias	0.006679	0.004292	0.002161
		MSE	0.005369	0.002375	0.001134
		Variance	0.005324	0.002356	0.001130
Naive MLE	$\hat{\gamma}$	Bias	-0.000280	-0.020344	-0.011666
		MSE	0.232053	0.108218	0.052218
		Variance	0.232053	0.108218	0.052218
	$\hat{\beta}$	Bias	0.006740	-0.000553	-0.004319
		MSE	0.005608	0.002042	0.001008
		Variance	0.005608	0.002042	0.001008

cont....

Table 1a (continued)

Monte Carlo study with 1000 simulations, and

$\sigma_u^2 = 0.26$, $\gamma = 1.0$, $\beta = 0.4$,

sample sizes = 40, 80, 160,

$x_1 = -1.904$, $x_2 = 3.904$

Sample sizes			n=40	n=80	n=160
True MLE	$\hat{\beta}$	Bias	0.016961	0.008359	0.004159
		MSE	0.007208	0.002404	0.001141
		Variance	0.007208	0.002404	0.001141
	# finite		1000	1000	1000
Mod. MLE	$\hat{\beta}$	Bias	0.016630	0.008153	0.003975
		MSE	0.007064	0.002389	0.001135
		Variance	0.007064	0.002389	0.001135
Join Post. Mode Jeffreys' Prior	$\hat{\gamma}$	Bias	-0.001334	-0.019991	-0.011633
		MSE	0.208758	0.104985	0.051478
		Variance	0.208756	0.104585	0.051343
	$\hat{\beta}$	Bias	0.011518	0.007246	0.003736
		MSE	0.004420	0.002232	0.001103
		Variance	0.004288	0.002180	0.001089
Marginal Post. Mode Jeffreys' Prior	$\hat{\gamma}$	Bias	-0.001296	-0.020295	-0.011694
		MSE	0.218119	0.107222	0.052010
		Variance	0.218118	0.106810	0.051873
	$\hat{\beta}$	Bias	0.004799	0.003930	0.002077
		MSE	0.004431	0.002232	0.001103
		Variance	0.004408	0.002217	0.001099

Table 1b

Monte Carlo study with 1000 simulations, and

$\sigma_u^2 = 0.26$, $\gamma = 1.0$, $\beta = 0.4$,

sample sizes = 320, 480, 960,

$x_1 = -1.904$, $x_2 = 3.904$

Sample sizes			n=320	n=480	n=960
Join Post. Mode $\beta_r \sim U(0,\sigma_u^{-1})$	$\hat{\gamma}$	Bias	-0.008038	-0.004337	-0.003272
		MSE	0.026503	0.016746	0.008248
		Variance	0.026438	0.016727	0.008238
	$\hat{\beta}$	Bias	0.002367	0.001669	0.000612
		MSE	0.000543	0.000366	0.000199
		Variance	0.000537	0.000363	0.000199
Marginal Post. Mode $\beta_r \sim U(0,\sigma_u^{-1})$	$\hat{\gamma}$	Bias	-0.008059	-0.004340	-0.003277
		MSE	0.026644	0.016803	0.008263
		Variance	0.026579	0.016784	0.008252
	$\hat{\beta}$	Bias	0.001535	0.001115	0.000335
		MSE	0.000542	0.000366	0.000199
		Variance	0.000540	0.000364	0.000199
Naive MLE	$\hat{\gamma}$	Bias	-0.007966	-0.004252	-0.003169
		MSE	0.026500	0.016742	0.008244
		Variance	0.026500	0.016742	0.008244
	$\hat{\beta}$	Bias	-0.005776	-0.006461	-0.007485
		MSE	0.000508	0.000363	0.000231
		Variance	0.000508	0.000363	0.000231

cont....

Table 1b (continued)

Monte Carlo study with 1000 simulations, and

$\sigma_u^2 = 0.26$, $\gamma = 1.0$, $\beta = 0.4$,

sample sizes = 320, 480, 960,

$x_1 = -1.904$, $x_2 = 3.904$

Sample sizes			n=320	n=480	n=960
True MLE	$\hat{\beta}$	Bias	0.002520	0.001765	0.000650
		MSE	0.000545	0.000367	0.000199
		Variance	0.000545	0.000367	0.000199
	# finite		1000	1000	1000
Mod. MLE	$\hat{\beta}$	Bias	0.002346	0.001594	0.000483
		MSE	0.000542	0.000365	0.000198
		Variance	0.000542	0.000365	0.000198
Join Post. Mode Jeffreys' Prior	$\hat{\gamma}$	Bias	-0.008009	-0.004327	-0.003268
		MSE	0.026317	0.016669	0.008230
		Variance	0.026253	0.016650	0.008219
	$\hat{\beta}$	Bias	0.002340	0.001657	0.000609
		MSE	0.000536	0.000363	0.000198
		Variance	0.000530	0.000360	0.000197
Marginal Post. Mode Jeffreys' Prior	$\hat{\gamma}$	Bias	-0.008033	-0.004328	-0.003159
		MSE	0.026454	0.016726	0.007706
		Variance	0.026390	0.016707	0.007696
	$\hat{\beta}$	Bias	0.001516	0.001104	0.009419
		MSE	0.000535	0.000362	0.000283
		Variance	0.000533	0.000361	0.000194

Table 2a

Monte Carlo study with 1000 simulations, and

$\sigma_u^2 = 0.26$, $\gamma = 1.0$, $\beta = 1.4$,

sample sizes = 40, 80, 160,

$x_1 = 0.0$, $x_2 = 2.0$

Sample sizes			n=40	n=80	n=160
Join Post. Mode $\beta_r \sim U(0,\sigma_u^{-1})$	$\hat{\gamma}$	Bias	-0.000076	-0.007110	-0.003430
		MSE	0.024971	0.013345	0.006225
		Variance	0.024971	0.013294	0.006213
	$\hat{\beta}$	Bias	0.030038	0.021345	0.010025
		MSE	0.112651	0.061087	0.029458
		Variance	0.111749	0.060632	0.029357
Marginal Post. Mode $\beta_r \sim U(0,\sigma_u^{-1})$	$\hat{\gamma}$	Bias	-0.000202	-0.007375	-0.003468
		MSE	0.026741	0.013768	0.006298
		Variance	0.026740	0.013714	0.006286
	$\hat{\beta}$	Bias	-0.009052	0.004544	0.001634
		MSE	0.117254	0.060407	0.029306
		Variance	0.117172	0.060387	0.029304
Naive MLE	$\hat{\gamma}$	Bias	-0.000639	-0.007225	-0.003423
		MSE	0.027679	0.013447	0.006231
		Variance	0.027679	0.013447	0.006231
	$\hat{\beta}$	Bias	-0.260140	-0.236256	-0.249500
		MSE	0.100839	0.073609	0.070811
		Variance	0.100797	0.073553	0.070749

cont....

Table 2a (continued)

Monte Carlo study with 1000 simulations, and

$\sigma_u^2 = 0.26, \gamma = 1.0, \beta = 1.4,$

sample sizes = 40, 80, 160,

$x_1 = 0.0, x_2 = 2.0$

Sample sizes			n=40	n=80	n=160
True MLE	$\hat{\beta}$	Bias	0.192623	0.073347	0.032193
		MSE	3.252274	0.087337	0.033560
		Variance	3.252236	0.087332	0.033559
	# finite		989	1000	1000
Mod. MLE	$\hat{\beta}$	Bias	0.090746	-0.001153	-0.028335
		MSE	0.426147	0.049280	0.023140
		Variance	0.426139	0.049280	0.023140
Join Post. Mode Jeffreys' Prior	$\hat{\gamma}$	Bias	0.000012	-0.006971	-0.003402
		MSE	0.023581	0.012948	0.006137
		Variance	0.023581	0.012899	0.006126
	$\hat{\beta}$	Bias	0.017305	0.019127	0.009510
		MSE	0.101731	0.056966	0.028613
		Variance	0.101431	0.056600	0.028523
Marginal Post. Mode Jeffreys' Prior	$\hat{\gamma}$	Bias	-0.000137	-0.007206	-0.003439
		MSE	0.024749	0.013314	0.006206
		Variance	0.024794	0.013262	0.006194
	$\hat{\beta}$	Bias	-0.014778	0.002628	0.001206
		MSE	0.101209	0.056358	0.028468
		Variance	0.100990	0.056351	0.028466

Table 2b

Monte Carlo study with 1000 simulations, and

$\sigma_u^2 = 0.26$, $\gamma = 1.0$, $\beta = 1.4$,

sample sizes = 320, 480, 960,

$x_1 = 0.0$, $x_2 = 2.0$

Sample sizes			n=320	n=480	n=960
Join Post. Mode $\beta_r \sim U(0,\sigma_u^{-1})$	$\hat{\gamma}$	Bias	-0.002648	-0.001252	-0.001079
		MSE	0.003049	0.002067	0.000972
		Variance	0.003042	0.002065	0.000970
	$\hat{\beta}$	Bias	0.003700	0.003293	0.000310
		MSE	0.014436	0.009507	0.004812
		Variance	0.014423	0.009497	0.004812
Marginal Post. Mode $\beta_r \sim U(0,\sigma_u^{-1})$	$\hat{\gamma}$	Bias	-0.002659	-0.001255	-0.001081
		MSE	0.003066	0.002074	0.000973
		Variance	0.003059	0.002073	0.000972
	$\hat{\beta}$	Bias	-0.000503	0.000489	-0.001093
		MSE	0.014409	0.009491	0.004812
		Variance	0.014408	0.009490	0.004810
Naive MLE	$\hat{\gamma}$	Bias	-0.002625	-0.001223	-0.001044
		MSE	0.003050	0.002067	0.000971
		Variance	0.003050	0.002067	0.000971
	$\hat{\beta}$	Bias	-0.255878	-0.257014	-0.259559
		MSE	0.069660	0.068824	0.068774
		Variance	0.069594	0.068758	0.068707

cont....

Table 2b (continued)

Monte Carlo study with 1000 simulations, and

$\sigma_u^2 = 0.26$, $\gamma = 1.0$, $\beta = 1.4$,

sample sizes = 320, 480, 960,

$x_1 = 0.0$, $x_2 = 2.0$

Sample sizes			n=320	n=480	n=960
True MLE	$\hat{\beta}$	Bias	0.014098	0.010088	0.003592
		MSE	0.015316	0.009889	0.004894
		Variance	0.015316	0.009888	0.004894
	# finite		1000	1000	1000
Mod. MLE	$\hat{\beta}$	Bias	-0.041154	-0.043807	-0.048651
		MSE	0.012395	0.008930	0.005899
		Variance	0.012393	0.008928	0.005897
Join Post. Mode Jeffreys' Prior	$\hat{\gamma}$	Bias	-0.002638	-0.001249	-0.001079
		MSE	0.003028	0.002058	0.000969
		Variance	0.003021	0.002056	0.000968
	$\hat{\beta}$	Bias	0.003573	0.003229	0.000294
		MSE	0.014238	0.009422	0.004791
		Variance	0.014225	0.009412	0.004791
Marginal Post. Mode Jeffreys' Prior	$\hat{\gamma}$	Bias	-0.002649	-0.001252	-0.001080
		MSE	0.003045	0.002065	0.000971
		Variance	0.003038	0.002063	0.000970
	$\hat{\beta}$	Bias	-0.000589	0.000459	-0.001084
		MSE	0.014210	0.009405	0.004791
		Variance	0.014100	0.009405	0.004790

Table 3a

Monte Carlo study with 1000 simulations, and

$\sigma_u^2 = 0.26$, $\gamma = 1.0$, $\beta = 2.0$,

sample sizes = 40, 80, 160,

$x_1 = 0.1873$, $x_2 = 1.8127$

Sample sizes			n=40	n=80	n=160
Join Post. Mode $\beta_r \sim U(0,\sigma_u^{-1})$	$\hat{\gamma}$	Bias	-0.002831	-0.007201	-0.002720
		MSE	0.016792	0.008853	0.004274
		Variance	0.016784	0.008801	0.004267
	$\hat{\beta}$	Bias	-0.094029	-0.027770	-0.007964
		MSE	0.240126	0.177616	0.103445
		Variance	0.231285	0.176845	0.103382
Marginal Post. Mode $\beta_r \sim U(0,\sigma_u^{-1})$	$\hat{\gamma}$	Bias	-0.002903	-0.007360	-0.002753
		MSE	0.017370	0.009092	0.004329
		Variance	0.017361	0.009038	0.004321
	$\hat{\beta}$	Bias	-0.147255	-0.057234	-0.023542
		MSE	0.250940	0.177316	0.102722
		Variance	0.229256	0.174040	0.102168
Naive MLE	$\hat{\gamma}$	Bias	-0.002994	-0.007338	-0.002721
		MSE	0.018282	0.008960	0.004284
		Variance	0.018282	0.008960	0.004284
	$\hat{\beta}$	Bias	-0.535302	-0.572539	-0.587193
		MSE	0.366153	0.354983	0.358070
		Variance	0.365867	0.354656	0.357726

cont....

Table 3a (continued)

Monte Carlo study with 1000 simulations, and

$\sigma_u^2 = 0.26$, $\gamma = 1.0$, $\beta = 2.0$,

sample sizes = 40, 80, 160,

$x_1 = 0.1873$, $x_2 = 1.8127$

Sample sizes			n=40	n=80	n=160
True MLE	$\hat{\beta}$	Bias	0.396836	0.283131	0.089786
		MSE	1.581492	1.988306	0.164424
		Variance	1.581329	1.988226	0.164416
	# finite		967	999	1000
Mod. MLE	$\hat{\beta}$	Bias	0.050727	-0.115587	-0.157357
		MSE	1.513346	0.131704	0.078146
		Variance	1.513344	0.131690	0.078121
Join Post. Mode Jeffreys' Prior	$\hat{\gamma}$	Bias	-0.002725	-0.007090	-0.002699
		MSE	0.015879	0.008603	0.004215
		Variance	0.015872	0.008553	0.004207
	$\hat{\beta}$	Bias	-0.095133	-0.029422	-0.008635
		MSE	0.220638	0.168166	0.100542
		Variance	0.211587	0.167300	0.100467
Marginal Post. Mode Jeffreys' Prior	$\hat{\gamma}$	Bias	-0.002794	-0.007228	-0.002731
		MSE	0.016299	0.008814	0.004265
		Variance	0.016291	0.008761	0.004258
	$\hat{\beta}$	Bias	-0.146534	-0.058482	-0.024102
		MSE.	0.229955	0.167918	0.099853
		Variance	0.208483	0.164498	0.099272

Table 3b

Monte Carlo study with 1000 simulations, and

$\sigma_u^2 = 0.26$, $\gamma = 1.0$, $\beta = 2.0$,

sample sizes = 320, 480, 960,

$x_1 = 0.1873$, $x_2 = 1.8127$

Sample sizes			n=320	n=480	n=960
Join Post. Mode $\beta_r \sim U(0,\sigma_u^{-1})$	$\hat{\gamma}$	Bias	-0.002053	-0.001174	-0.000679
		MSE	0.002097	0.001383	0.000627
		Variance	0.002093	0.001382	0.000627
	$\hat{\beta}$	Bias	-0.004475	-0.002032	-0.003021
		MSE	0.054787	0.036747	0.018582
		Variance	0.054767	0.036743	0.018573
Marginal Post. Mode $\beta_r \sim U(0,\sigma_u^{-1})$	$\hat{\gamma}$	Bias	-0.002062	-0.001176	-0.000679
		MSE	0.002109	0.001388	0.000628
		Variance	0.002105	0.001387	0.000628
	$\hat{\beta}$	Bias	-0.012446	-0.007386	-0.005710
		MSE	0.054552	0.036629	0.018563
		Variance	0.054397	0.036574	0.018530
Naive MLE	$\hat{\gamma}$	Bias	-0.002036	-0.001151	-0.000650
		MSE	0.002098	0.001383	0.000627
		Variance	0.002098	0.001383	0.000627
	$\hat{\beta}$	Bias	-0.594484	-0.596063	-0.598710
		MSE	0.360005	0.359685	0.360664
		Variance	0.359652	0.359329	0.360305

cont....

Table 3b (continued)

Monte Carlo study with 1000 simulations, and

$\sigma_u^2 = 0.26$, $\gamma = 1.0$, $\beta = 2.0$,

sample sizes = 320, 480, 960,

$x_1 = 0.1873$, $x_2 = 1.8127$

Sample sizes			n=320	n=480	n=960
True MLE	$\hat{\beta}$	Bias	0.038846	0.025803	0.010288
		MSE	0.066372	0.041475	0.019599
		Variance	0.066371	0.041474	0.019599
	# finite		1000	1000	1000
Mod. MLE	$\hat{\beta}$	Bias	-0.177429	-0.182380	-0.189374
		MSE	0.057258	0.050226	0.044305
		Variance	0.057227	0.050193	0.044269
Join Post. Mode Jeffreys' Prior	$\hat{\gamma}$	Bias	-0.002045	-0.001171	-0.000678
		MSE	0.002082	0.001377	0.000626
		Variance	0.002078	0.001375	0.000625
	$\hat{\beta}$	Bias	-0.004667	-0.002123	-0.003034
		MSE	0.054025	0.036415	0.018500
		Variance	0.054003	0.036410	0.018491
Marginal Post. Mode Jeffreys' Prior	$\hat{\gamma}$	Bias	-0.002054	-0.001173	-0.000678
		MSE	0.002094	0.001382	0.000627
		Variance	0.002090	0.001380	0.000626
	$\hat{\beta}$	Bias	-0.012612	-0.007462	-0.005717
		MSE	0.053792	0.036294	0.018479
		Variance	0.053633	0.036239	0.018446

Table 4a

Monte Carlo study with 1000 simulations, and

$\sigma_u^2 = 0.26$, $\gamma = 1.0$, $\beta = 5.0$,

sample sizes = 40, 80, 160,

$x_1 = 0.3766$, $x_2 = 1.6234$

Sample sizes			n=40	n=80	n=160
Join Post. Mode $\beta_i \sim U(0, \sigma_u^{-1})$	$\hat{\gamma}$	Bias	-0.005333	-0.004455	-0.002803
		MSE	0.010176	0.005242	0.002718
		Variance	0.010148	0.005222	0.002710
	$\hat{\beta}$	Bias	-2.241226	-1.774272	-1.284754
		MSE	5.422818	3.671449	2.323108
		Variance	0.399725	0.523408	0.672516
Marginal Post. Mode $\beta_i \sim U(0, \sigma_u^{-1})$	$\hat{\gamma}$	Bias	-0.005289	-0.004439	-0.002802
		MSE	0.010017	0.005221	0.002719
		Variance	0.009989	0.005202	0.002711
	$\hat{\beta}$	Bias	-2.331887	-1.838137	-1.330165
		MSE	5.836318	3.894984	2.432167
		Variance	0.398622	0.516236	0.662828
Naive MLE	$\hat{\gamma}$	Bias	-0.005196	-0.004480	-0.002801
		MSE	0.011082	0.005296	0.002730
		Variance	0.011082	0.005296	0.002730
	$\hat{\beta}$	Bias	-3.073353	-3.133056	-3.150456
		MSE	9.587374	9.860868	9.947629
		Variance	9.577929	9.851052	9.937703

cont....

Table 4a (continued)

* Monte Carlo study with 1000 simulations, and

$\sigma_u^2 = 0.26, \gamma = 1.0, \beta = 5.0,$

sample sizes = 40, 80, 160,

$x_1 = 0.3766, x_2 = 1.6234$

Sample sizes			n=40	n=80	n=160
True MLE	$\hat{\beta}$	Bias	0.996723	0.485574	1.160758
		MSE	28.702300	15.862030	34.280980
		Variance	28.700690	15.861690	34.279250
	# finite		616	683	782
Mod. MLE	$\hat{\beta}$	Bias	-1.363584	-1.992234	-2.081423
		MSE	25.402840	4.451098	4.540688
		Variance	25.400980	4.447129	4.536355
Join Post. Mode Jeffreys' Prior	$\hat{\gamma}$	Bias	-0.005199	-0.004396	-0.002784
		MSE	0.009660	0.005107	0.002683
		Variance	0.009633	0.005087	0.002675
	$\hat{\beta}$	Bias	-2.212540	-1.760087	-1.278419
		MSE	5.280353	3.610916	2.299336
		Variance	0.385020	0.513009	0.664980
Marginal Post. Mode Jeffreys' Prior	$\hat{\gamma}$	Bias	-0.005152	-0.004377	-0.002782
		MSE	0.009497	0.005079	0.002681
		Variance	0.009470	0.005060	0.002673
	$\hat{\beta}$	Bias	-2.301695	-1.823584	-1.324136
		MSE	5.680628	3.831280	2.407991
		Variance	0.382829	0.505820	0.654654

Table 4b

Monte Carlo study with 1000 simulations, and

$\sigma_u^2 = 0.26$, $\gamma = 1.0$, $\beta = 5.0$,

sample sizes = 320, 480, 960,

$x_1 = 0.3766$, $x_2 = 1.6234$

Sample sizes			n=320	n=480	n=960
Join Post. Mode $\beta_r \sim U(0,\sigma_u^{-1})$	$\hat{\gamma}$	Bias	-0.002648	-0.001662	-0.000918
		MSE	0.001271	0.000844	0.000382
		Variance	0.001263	0.000841	0.000381
	$\hat{\beta}$	Bias	-0.884142	-0.646836	-0.372232
		MSE	1.530906	1.235380	0.861404
		Variance	0.749199	0.816983	0.722847
Marginal Post. Mode $\beta_r \sim U(0,\sigma_u^{-1})$	$\hat{\gamma}$	Bias	-0.002648	-0.001662	-0.000917
		MSE	0.001272	0.000845	0.000382
		Variance	0.001265	0.000843	0.000382
	$\hat{\beta}$	Bias	-0.915700	-0.672207	-0.388229
		MSE	1.578128	1.258682	0.867023
		Variance	0.739622	0.806820	0.716302
Naive MLE	$\hat{\gamma}$	Bias	-0.002643	-0.001645	-0.000894
		MSE	0.001273	0.000845	0.000382
		Variance	0.001273	0.000845	0.000382
	$\hat{\beta}$	Bias	-3.165291	-3.167592	-3.171970
		MSE	10.029700	10.041210	10.064970
		Variance	10.019680	10.031180	10.054900

cont....

Table 4b (continued)

Monte Carlo study with 1000 simulations, and

$\sigma_u^2 = 0.26$, $\gamma = 1.0$, $\beta = 5.0$,

sample sizes = 320, 480, 960,

$x_1 = 0.3766$, $x_2 = 1.6234$

Sample sizes			n=320	n=480	n=960
True MLE	$\hat{\beta}$	Bias	1.971829	0.930198	0.816134
		MSE	128.725100	15.095600	30.051630
		Variance	128.720700	15.094670	30.050950
	# finite		899	927	989
Mod. MLE	$\hat{\beta}$	Bias	-2.142563	-2.153862	-2.172300
		MSE	4.684757	4.704171	4.749118
		Variance	4.680167	4.699532	4.744399
Join Post. Mode Jeffreys' Prior	$\hat{\gamma}$	Bias	-0.002639	-0.001658	-0.000917
		MSE	0.001262	0.000840	0.000381
		Variance	0.001255	0.000838	0.000380
	$\hat{\beta}$	Bias	-0.881446	-0.645564	-0.372105
		MSE	1.521164	1.229275	0.856719
		Variance	0.744218	0.812523	0.718257
Marginal Post. Mode Jeffreys' Prior	$\hat{\gamma}$	Bias	-0.002641	-0.001659	-0.000917
		MSE	0.001263	0.000841	0.000381
		Variance	0.001256	0.000839	0.000381
	$\hat{\beta}$	Bias	-0.913660	-0.671537	-0.388960
		MSE	1.568185	1.252587	0.862432
		Variance	0.733410	0.801625	0.711142

Fig. 1 Bias of estimates of β vs log(sample size) for $\beta = 0.4$

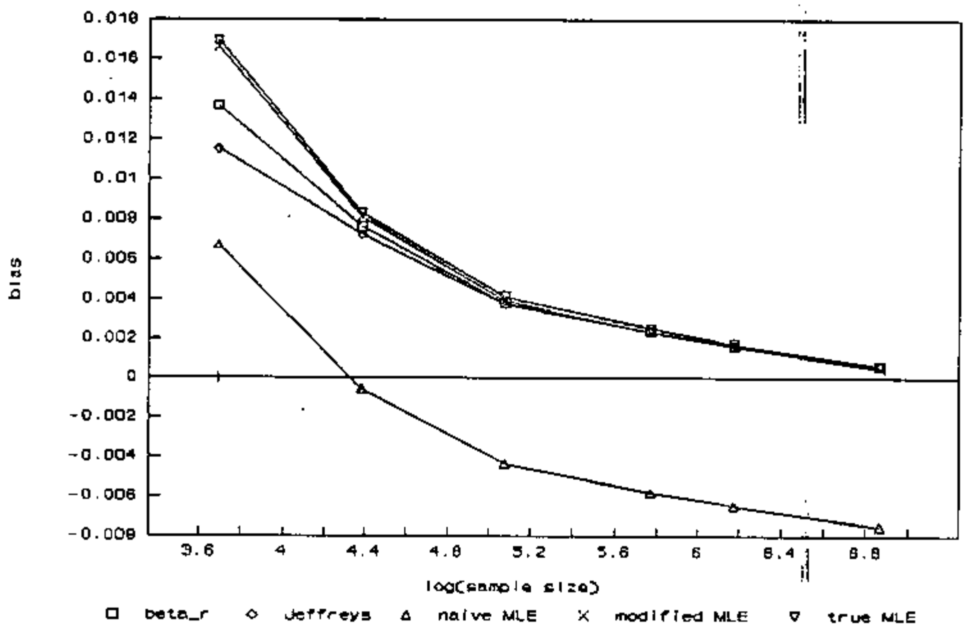

Fig. 2 Bias of estimates of β vs log(sample size) for $\beta = 1.4$

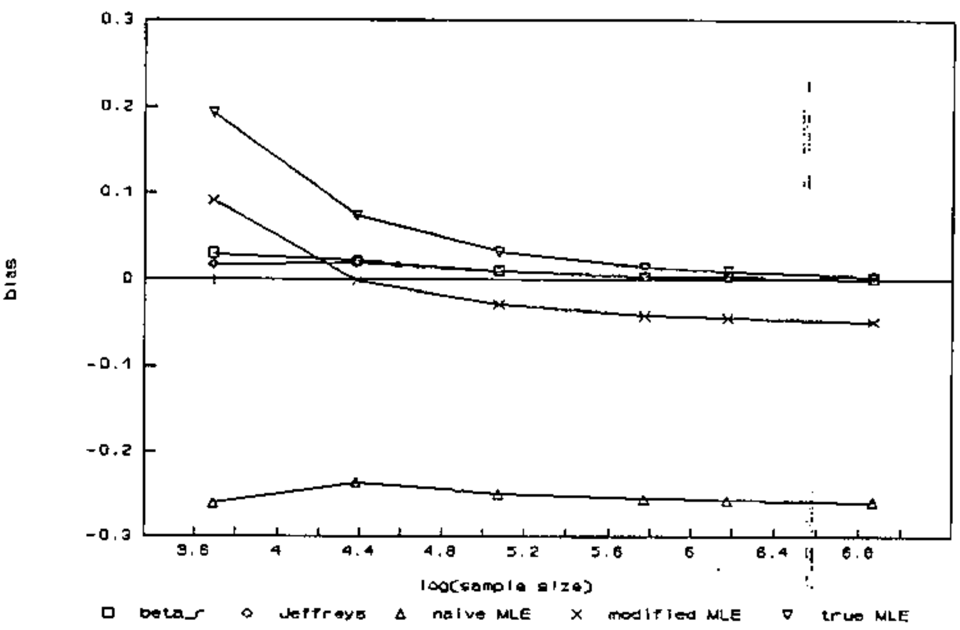

Fig. 3 Bias of estimates of β vs log(sample size) for β = 2.0

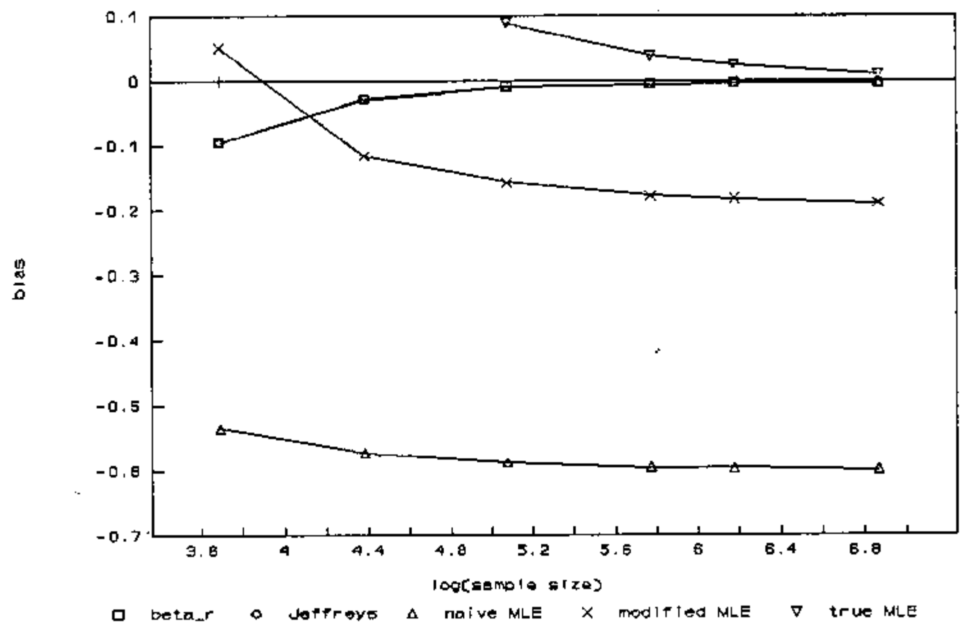

Fig. 4 Bias of estimates of β vs log(sample size) for β = 5.0

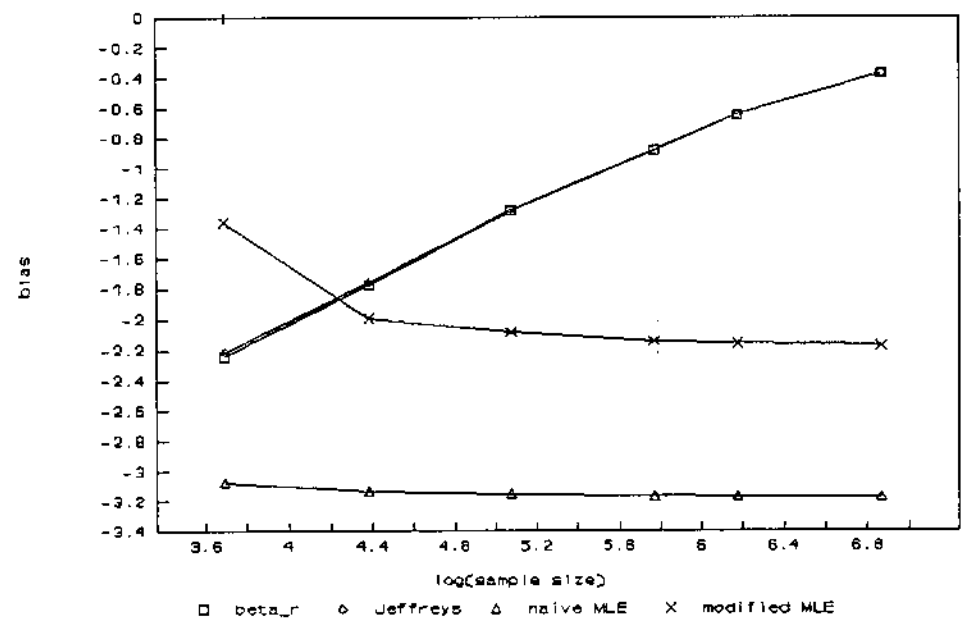

Fig. 5 log(MSE) of estimates of β vs log(sample size) for $\beta = 0.4$

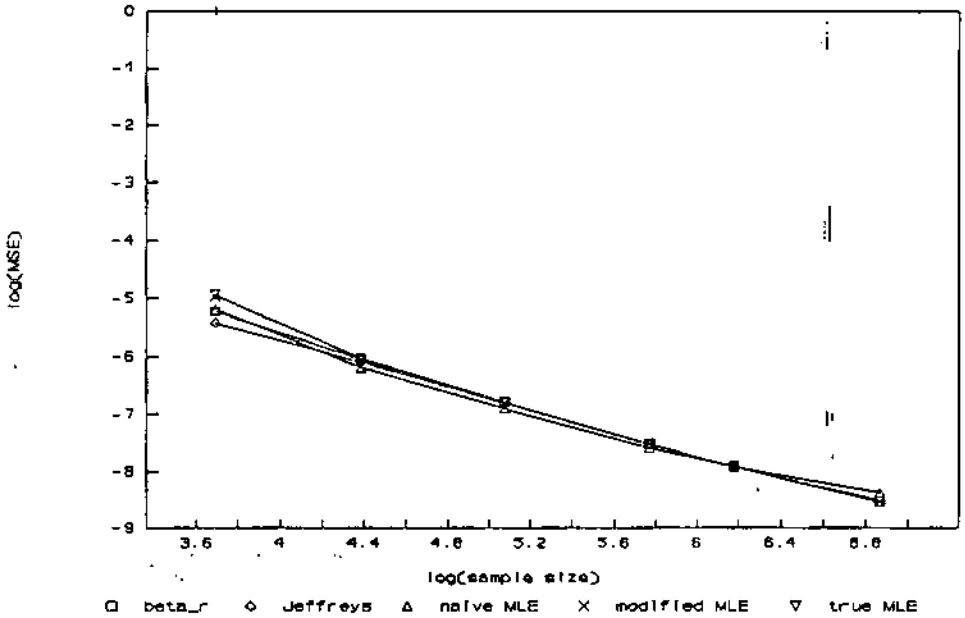

Fig. 6 log(MSE) of estimates of β vs log(sample size) for $\beta = 1.4$

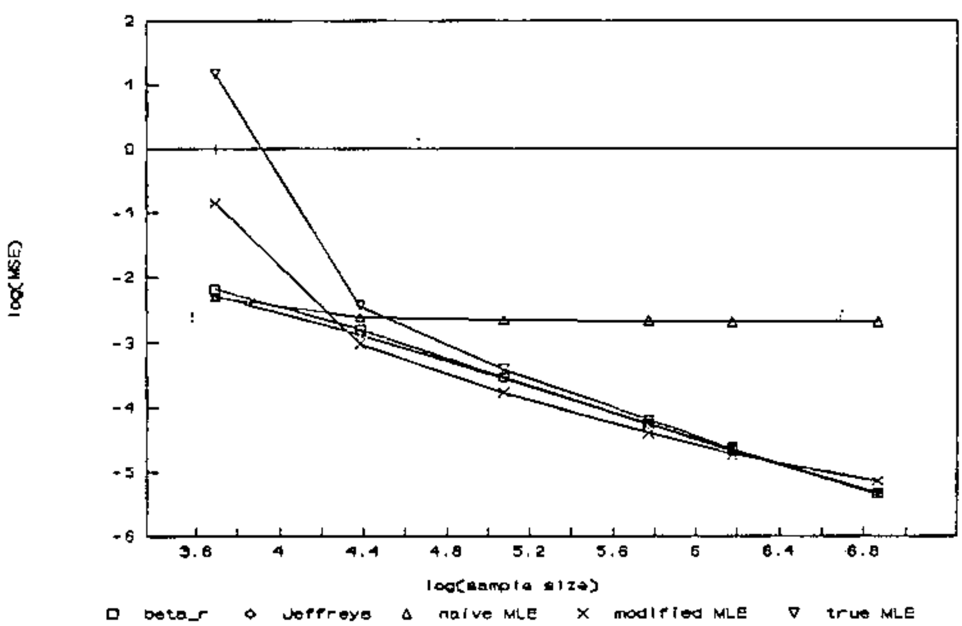

Fig. 7 log(MSE) of estimates of β vs log(sample size) for $\beta = 2.0$

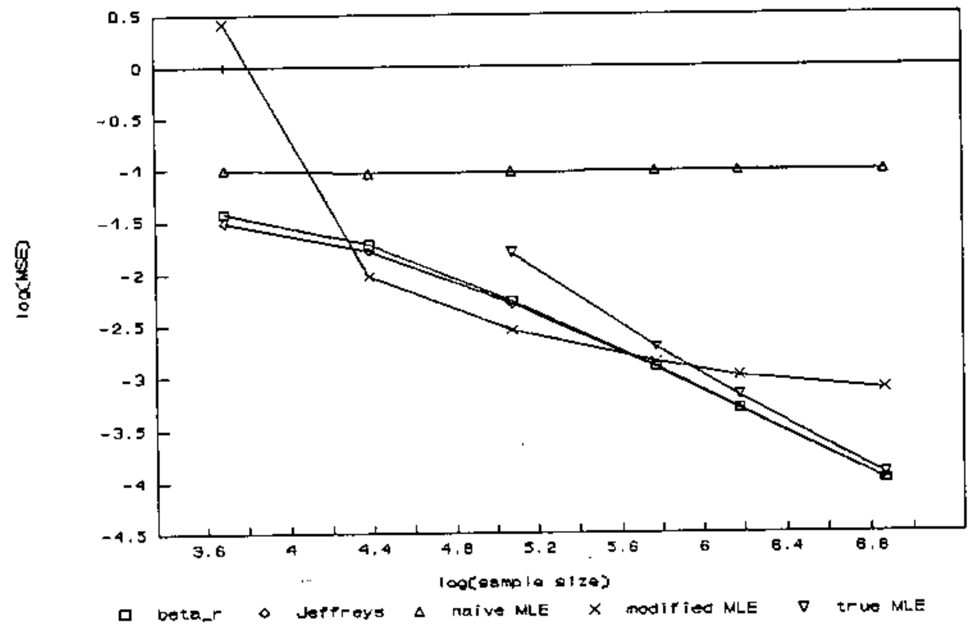

Fig. 8 log(MSE) of estimates of β vs log(sample size) for $\beta = 5.0$