Hierarchical Bayes Analysis of the Log-normal Distribution

Fabrizi Enrico
DISES, Università Cattolica
Via Emilia Parmense 84
29122 Piacenza,Italy
E-mail: enrico.fabrizi@unicatt.it

Trivisano Carlo
Dipartimento di Statistica 'P. Fortunati', Università di Bologna
Via Belle Arti 41
40126 Bologna, Italy
E-mail: carlo.trivisano@unibo.it

1 Introduction

The log tranformation is one of the most commonly used when analyzing skewed data to make them approximately normal. In this paper we consider Bayesian inference on the expected value and other moments and measures of central tendency in the log-normal model. Suppose that a random variable X with mean ξ and variance σ^2 is normally distributed, such that $\exp(X) \sim LogN(\xi, \sigma^2)$. Let's consider functionals of (ξ, σ^2) of the form $\theta_{a,b} = \exp(a\xi + b\sigma^2)$ with $a, b \in \Re$ based on a random sample $(X_1, \ldots, X_n)$. We may obtain the mean, the median, the mode and various non-central moments of the log-normal distribution for different choices of a, b. More specifically $a = 1$ and $b = 0$ yields the median $(\theta_{1,0})$, $a = 1$ and $b = -1$ yields the mode $(\theta_{1,-1})$ and $a = 1$ and $b = .5$ yields the mean $(\theta_{1,0.5})$.

In the Bayesian literature, an important reference is Zellner (1971). He considers diffuse priors of the type $p(\xi, \sigma) \propto \sigma^{-1}$ and shows, for the log-normal median, that $p(\theta_{1,0}|data)$ is a log-t distribution. Summarizing the log-t distribution is challenging using popular loss functions, such as the quadratic, because moments of all orders do not exist. Similarly, for the log-normal mean Zellner (1971) shows that $p(\theta_{1,0.5}|data)$ follows a distribution without finite moments.

Moreover, considering inference conditional on σ, Zellner (1971) notes that, within the class of estimators of the form $k\exp(\bar{X})$ with $\bar{X} = n^{-1}\sum_{i=1}^{n} X_i$ and where k a is constant, the estimator for $\theta_{1,0.5}$ with minimum mean square error (MSE) is given by $\breve{\theta}_{1,0.5} = \exp(\bar{X} + \sigma^2/2 - 3\sigma^2/2n)$. From a Bayesian point of view, this estimator may be justified as the minimizer of the posterior expected loss, provided that the relative quadratic loss function $L_{RQ} = [(\theta - \hat{\theta})/\theta]^2$ is adopted.

Another important reference is Rukhin (1986). Rukhin proposes the following 'generalized' prior: $p(\xi, \sigma) = p(\sigma) \propto \sigma^{-2\nu+n-2} \exp\left(-\sigma^2\left[\gamma^2/2 - 2\left(b - a^2/n\right)\right]\right)$, with $\gamma^2 > 4(b - a^2/n)$. Assuming the relative quadratic loss function L_{RQ}, he obtains an estimator for $\theta_{a,b}$ of the form $\hat{\theta}^{Ru}_{a,b} = \exp(a\bar{X})g(Y)$ that is given by

$$(1) \quad \hat{\theta}^{Ru}_{a,b} = \exp(a\bar{X})\left(\frac{\beta}{\gamma}\right)^{\nu} \frac{K_\nu(\beta Y)}{K_\nu(\gamma Y)},$$

$\beta = \gamma^2 - 2c$, $c = b - 3a^2/(2n)$ and $K_\nu(\,)$ is the modified Bessel function of the third kind (the Bessel-K function from now on). For a general introduction to Bessel functions, see Abramowitz and Stegun (1968), chapters 9 and 10. To obtain the values for the hyperparameters ν, γ, Rukhin (1986) chooses to minimize the frequentist MSE of $\hat{\theta}^{Ru}_{a,b}$. As the $K_\nu(\,)$ are quite difficult to handle, Rukhin uses a 'small arguments' approximation to $\hat{\theta}^{Ru}_{a,b}$ to propose a value for ν and a 'large arguments' approximation to propose a value for γ. Rukhin does not recognize that, with a simple change of variable, the prior he proposes may be seen as the product of a flat prior over the real line for ξ and the following prior on

$\sigma^2\ p(\sigma^2) \propto (\sigma^2)^{-\nu+n/2-3/2} \exp(-\sigma^2[\psi^2/2 - 2(b - a^2/n)])$ which is the limit of a generalized inverse gamma distribution, $GIG(\lambda, \delta, \gamma)$ as $\delta \to 0$. The other parameters are given by $\lambda = -\nu + n/2 - 1/2$ and $\gamma^2 = \psi^2/2 - 2(b - a^2/n)$ (see section 2 for more details and notation). He does not provide the posterior distribution, so his proposal is inadequate for many inferential purposes (i.e., calculating of posterior variances or posterior probability intervals).

In this paper, we derive the posterior distribution of $\theta_{a,b}$ assuming a proper generalized inverse gamma prior on σ^2 (and a flat prior over the real line for ξ). We show that this posterior is a log-generalized hyperbolic distribution and state the conditions on the hyperparameters that guarantee the existence of posterior moments of a given order. Once these conditions are met for the first two non-central moments, we discuss the Bayes estimators with the ordinary quadratic loss function $L_Q = (\theta - \hat{\theta})^2$. Moreover, we show that, given our choice of the prior distributions, Bayes estimators associated with the relative quadratic loss function L_{RQ} can be reconducted to posterior expectations provided that b is properly modified. Adopting a 'small arguments' approximation to the Bessel-K functions we propose a choice of the hyparameters aimed at minimizing the MSE. We show using simulation that our Bayes estimator of the mean, i.e. $\theta_{1,0.5}$ is substantially equivalent to the estimator proposed in Shen et al. (2006), which has been proven to be superior to many of the alternatives previously proposed in the literature.

The paper is organized as follows. In Section 2 we briefly present the generalized inverse Gaussian and generalized hyperbolic distributions. In Section 3, posterior distributions for σ^2 and $\theta_{a,b}$ are derived, and Bayes estimators under quadratic and relative quadratic losses are introduced. Section 4 is devoted to the choice of values to be assigned to the hyperparameters in order to obtain Bayes estimators with the minimum frequentist MSE. Section 5 offers some conclusions and outlines the results of simulation exercises not reported here for brevity.

2 The generalized inverse Gaussian and generalized hyperbolic distributions

In this section we briefly introduce the generalized inverse Gaussian (GIG) and generalized hyperbolic (GH) distributions, establish the notation and mention some key properties that will be used later. For more details on these distributions, see Bibby and Sørensen (2003) and Eberlein and von Hammerstein (2004) among others.

The density of the GIG distribution may be written as follows:

$$(2) \quad p(x) = \left(\frac{\gamma}{\delta}\right)^{\lambda} \frac{1}{2K_\lambda(\delta\gamma)} x^{\lambda-1} \exp\left\{-\frac{1}{2}\left(\delta^2 x^{-1} + \gamma^2 x\right)\right\} \mathbf{1}_{\Re+}$$

If $\delta > 0$ the permissible values for the other parameters are $\gamma \geq 0$ if $\lambda < 0$ and $\gamma > 0$ if $\lambda = 0$. If $\delta \geq 0$ then γ, λ should be strictly positive.

Many important distributions may be obtained as special cases of the GIG. For $\lambda > 0$ and $\gamma > 0$, the gamma distribution emerges as the limit when $\delta \to 0$. The inverse-gamma is obtained when $\lambda < 0$, $\delta > 0$ and $\gamma \to 0$ and an inverse Gaussian distribution is obtained when $\lambda = -\frac{1}{2}$.

Barndorff-Nielsen (1977) introduces the generalized hyperbolic (GH) distribution as a normal variance-mean mixture where the mixing distribution is GIG. That is, if $(X|W = w) \sim N(\mu + \beta w, w)$ and $W \sim GIG(\lambda, \delta, \gamma)$ then the marginal distribution of X will be GH (i.e., $X \sim GH(\lambda, \alpha, \beta, \delta, \mu)$, where $\alpha^2 = \beta^2 + \gamma^2$). The probability density function of the GH is given by

$$(3) \quad f(x) = \frac{\left(\frac{\gamma}{\delta}\right)^{\lambda}}{\sqrt{2\pi} K_\lambda(\delta\gamma)} \frac{K_{\lambda-1/2}\left(\alpha\sqrt{\delta^2 + (x-\mu)^2}\right)}{\left(\sqrt{\delta^2 + (x-\mu)^2}/\alpha\right)^{1/2-\lambda}} \exp\left(\beta(x-\mu)\right) \mathbf{1}_{\Re},$$

where $\gamma^2 = \alpha^2 - \beta^2$. The parameter domain is defined by the following conditions: *i)* $\delta \geq 0$, $\alpha > 0$, $\alpha^2 > \beta^2$ if $\lambda > 0$; *ii)* $\delta > 0$, $\alpha > 0$, $\alpha^2 > \beta^2$ if $\lambda = 0$; *iii)* $\delta > 0$, $\alpha \geq 0$, $\alpha^2 \geq \beta^2$ if $\lambda < 0$. The parameter α determines the shape, β determines the skewness (the sign of the skewness is consistent with that of β), μ is a location parameter, δ serves for scaling and λ influences the size of mass contained in the tails. The class of GH distributions is closed under affine transformations i.e if $X \sim GH(\lambda, \alpha, \beta, \delta, \mu)$ and $Z = b_0 X + b_1$ then $Z \sim GH(\lambda, \alpha/|b_0|, \beta/|b_0|, |b_0|\delta, b_0\mu + b_1)$. An essential tool in what follows is the moment generating function of the GH distribution:

$$(4) \quad M_{GH}(t) = \exp(\mu t)\left(\frac{\gamma^2}{\alpha^2 - (\beta + t)^2}\right)^{\lambda/2} \frac{K_\lambda(\delta\sqrt{\alpha^2 - (\beta + t)^2})}{K_\lambda(\delta\gamma)}$$

which exists provided that $|\beta + t| < \alpha$.

3 Bayes estimators

The representation of the GH distribution as a normal mean-variance mixture with the GIG as mixing distribution introduced in previous section provides the basis for obtaining the posterior distribution of $\eta_{a,b} = log(\theta_{a,b})$ when assuming a GIG prior for σ^2. More specifically we can prove the following result.

Theorem 3.1. *Assume the following: i)* $p(\eta_{a,b}|\sigma^2, X) \sim N(\eta_{a,b}, a^2\sigma^2/n)$ *and ii)* $p(\xi, \sigma^2) = p(\xi)p(\sigma^2)$, *with* $p(\sigma^2) \sim GIG(\lambda, \delta, \gamma)$ *and* $p(\xi)$ *an improper distribution uniform over the real line.*
It follows that

$$(5) \quad p(\sigma^2|data) \sim GIG(\bar{\lambda}, \bar{\delta}^\star, \gamma),$$

$$(6) \quad p(\eta_{a,b}|data) \sim GH(\bar{\lambda}, \bar{\alpha}, \bar{\beta}, \bar{\delta}, \bar{\mu})$$

where $\bar{\delta}^\star = \sqrt{Y^2 + \delta^2}$, $Y^2 = \sum_{i=1}^n (X_i - \bar{X})^2$, $\bar{\lambda} = \lambda - \frac{n-1}{2}$, $\bar{\alpha} = \sqrt{\frac{n}{a^2}(\gamma^2 + \frac{nb^2}{a^2})}$ *and* $\bar{\beta} = n\frac{b}{a^2}$. *Let* $\bar{\gamma}^2 = \bar{\alpha}^2 - \bar{\beta}^2$. *As a consequence* $\bar{\gamma}^2 = \frac{n}{a^2}\gamma^2$, $\bar{\delta} = \sqrt{\frac{a^2}{n}(Y^2 + \delta^2)}$ *and* $\bar{\mu} = a\bar{X}$.

We are not primarily interested in $p(\eta_{a,b}|data)$, but rather in $\theta_{a,b} = \exp(\eta_{a,b})$ which is distributed as a log-GH, a distribution that has not, to our knowledge, received any attention in the literature. In any case, we can calculate the moments of $p(\theta_{a,b}|data)$ that we need for summarizing the posterior distribution with a quadratic loss function by using the moment-generating function of the GH distribution ($M_{GH}(t)$) and more specifically the fact that $E(\theta|data) = M_{GH}(1)$ and $V(\theta|data) = M_{GH}(2) - [M_{GH}(1)]^2$.

If we are able to generate samples from the GH distribution, moreover, we may obtain a sample from its exponential transformation. The quantiles and probability intervals may then be calculated using MC techniques. From among the variety of software available for generating random GH numbers, we mention the `ghyp` package running under `R` (Breymann and Lüthi, 2010).

$M_{\eta|data}(t)$ exists only if $|\bar{\beta} + t| < \bar{\alpha}$, or equivalently if $\bar{\alpha}^2 - (\bar{\beta} + t)^2 > 0$ (i.e., $\bar{\gamma}^2 > t^2 + 2n\frac{b}{a^2}t$). This condition implies the following constraint on the prior parameter γ:

$$(7) \quad \gamma^2 > \frac{a^2}{n}t^2 + 2bt.$$

The existence of posterior moments requires that γ is above a positive threshold when $a \neq 0, b > 0$ (as for the expected value). The threshold is asymptotically 0 for the median, (i.e., $\theta_{0,1}$), and it is negative for the mode (whenever $n > t/2$): so, it does not represent a restriction. With respect to the

inference on the expected value ($\theta_{1,0.5}$) note that the popular inverse gamma prior on σ^2, a special case of the GIG for $\lambda < 0$, $\delta > 0$ when $\gamma \to 0$, does not respect condition (7) thereby leading to a posterior distribution with non-existent moments. Note that this result is consistent with the following remark from Zellner (1971) concerning the inference about $\theta_{1,0}$: posterior moments exist only for the limit as $n \to \infty$ (that is, when the log-t posterior converges to the log-normal).
Similarly, the uniform prior over the range $(0, A)$ for σ (Gelman , 2006) implies that $p(\sigma^2) \propto \frac{1}{\sigma}\mathbf{1}_{(0,A)}$, which may be seen as an approximation to a $Gamma(\frac{1}{2}, \epsilon)$ (where $\epsilon = (4A^2)^{-1}$) truncated at A^2. For $\lambda > 0$, $\gamma > 0$ and $\delta \to 0$, $GIG(\lambda, \delta, \gamma) \to Gamma(\lambda, \gamma^2/2)$. If we let $A \to \infty$, therefore, $p(\sigma) \propto 1$ is equivalent to a GIG prior with $\gamma \to 0$ and thus implies non-existent posterior moments.

Consistent with intuition, condition (7) implies that, in practice, to obtain a posterior distribution of $\theta_{a,b}$ with finite moments, a prior with short tails should be chosen. If we summarize $p(\theta|data)$ using the ordinary quadratic loss function we obtain $\hat{\theta}^{QB}_{a,b} = E(\theta_{a,b}|data)$ or

$$\text{(8)} \quad \hat{\theta}^{QB}_{a,b} = \exp(\bar{\mu})\left(\frac{\bar{\gamma}^2}{\bar{\alpha}^2 - (\bar{\beta}+1)^2}\right)^{\bar{\lambda}/2} \frac{K_{\bar{\lambda}}(\bar{\delta}\sqrt{\bar{\alpha}^2 - (\bar{\beta}+1)^2})}{K_{\bar{\lambda}}(\bar{\delta}\bar{\gamma})}$$

$$\text{(9)} \quad = \exp(a\bar{X})\left(\frac{\gamma^2}{\gamma^2 - (\frac{a^2}{n} + 2b)}\right)^{(\lambda - \frac{n-1}{2})/2} \frac{K_{\{\lambda - \frac{n-1}{2}\}}\left(\sqrt{(Y^2+\delta^2)(\gamma^2 - (\frac{a^2}{n} + 2b))}\right)}{K_{\{\lambda - \frac{n-1}{2}\}}\left(\sqrt{(Y^2+\delta^2)\gamma^2}\right)}.$$

We provided two alternative expressions for $\hat{\theta}^{QB}_{a,b}$: (8) is indexed on the posterior parameters, and (9) highlights the role of the prior parameters, which and will be useful for studying the choice of hyperparameters that is discussed in the next section.

Under a relative quadratic loss function, the Bayes estimator is defined as $\hat{\theta}^{RQB}_{a,b} = E(\theta^{-1}_{a,b})/E(\theta^{-2}_{a,b})$ (see Zellner (1971)). The following result shows that $\hat{\theta}^{RQB}$ may be reconducted to a Bayes predictor under a quadratic loss function with a different choice of b and modified prior parameters.

Theorem 3.2. *For the Bayes estimator under relative quadratic loss function, we have that* $\hat{\theta}^{RQB}_{a,b} = \hat{\theta}^{QB}_{a,b^\star}$ *with* $b^\star = b - 2a^2/n$ *provided that the prior* $p(\sigma^2) \sim GIG(\lambda, \delta, \gamma_\star)$ *with* $\gamma^2_\star = \gamma^2 - 4a^2/n + 4b$ *is assumed.*

4 Prior elicitation

We can easily see that $\hat{\theta}^{QB}_{a,b}$ is sensitive to the choice of the prior parameters; therefore a careful choice of λ, δ, γ is an essential part of the inferential procedure. Following Rukhin (1986), our aim is to choose the hyperameters to minimize the frequentist MSE of the Bayes estimators. In practice this choice is a complicated task because expression (9) contains a ratio of Bessel-K functions that is quite intractable. Following Rukhin (1986) again, we will use a 'small argument' approximation to obtain the MSE-'optimal' values of the hyperparameters. Unfortunately this method is viable only for λ and δ because the small value approximation is free of γ. We also proved, using simulations not reported here for brevity, that the parameters determined in this manner also leads to good estimators performance when the arguments of the Bessel-K functions are no longer small. Consider first the following approximation of $\hat{\theta}^{QB}_{a,b}$ using the 'small argument' approximation of the Bessel-K functions.

Theorem 4.1. *Under the assumption that* $(Y^2+\delta^2)\gamma^2 < 1$, $(Y^2+\delta^2)(\gamma^2 - (\frac{a^2}{n} + 2b)) < 1$ *and* $\bar{\lambda} < -1$ *then*

$$\text{(10)} \quad \hat{\theta}^{QB} \cong \exp(a\bar{X}) \exp\left\{\frac{-(Y^2+\delta^2)(a^2+2nb)}{4n(\lambda - \frac{n-3}{2})}\right\}$$

Denote the 'small argument' approximation of $\hat{\theta}^{QB}$ by $\hat{\theta}^{qb}$. Note that, because $\bar{\lambda} = \lambda - (n-1)/2$ the assumption $\bar{\lambda} < 1$ is consistent (at least for large samples) for all choices of λ as a constant independent of n. As anticipated, this approximation is free of γ, therefore we will postpone the discussion of an effective choice of the value to assign to that parameter.

Rukhin (1986) proves that to minimize the frequentist MSE of estimators in the form $\exp(aX)g(Y)$ such as $\hat{\theta}^{qb}$, and $\hat{\theta}^{QB}$,

$$(11) \quad E\Big[g(Y) - \exp(c\sigma^2)\Big]^2,$$

where $c = b - 3a^2/(2n)$, should be minimized. Unfortunately, minimization of (11) with respect to (λ, δ) does not lead to a unique minimum. The optimum MSE is reached for a set of (λ, δ) pairs that are described by equation (12).

Theorem 4.2. *The value of λ in (10) that minimizes (11) (i.e., its frequentist MSE) is given by*

$$(12) \quad \lambda_{opt} = \frac{n-3}{2} - \frac{(n-1)(a^2+2nb)}{4nc} - \frac{(a^2+2nb)}{4nc}\frac{\delta^2}{\sigma^2}$$

for any δ in $\Re^+$.

The λ_{opt} in (12) is a function not only of δ, as anticipated, but also of the unknown σ^2; therefore, an optimal choice of (δ, λ) should depend, at least in principle, on a prior guess for σ^2. A method for circumventing the problem that is implicitly suggested by the generalized prior proposed by Rukhin (1986), is to let $\delta \to 0$. This condition may be approximated in practice by a δ that is much smaller than σ so as to make the third addend in (12) negligible. This approximation can be justified by noting that, in the GH distribution, δ has the same order of magnitude of the expectation and the mode of σ^2; therefore, choices of the type $\delta = k\sigma_0^2$ for some constant k and prior guess of the variance σ_0^2, imply a negligible third addend in (12) in the 'small value setting' we assumed for the derivation of $\hat{\theta}^{qb}$.

The estimator $\hat{\theta}^{qb}_{1,0.5}$ is connected to popular estimators that have already been discussed in the literature. Note that if $\delta = 0$, $a = 1$, $b = 0.5$ and we replace λ_{opt} into (3.1), we obtain $\hat{\theta}^{qb}_{1,0.5} = \exp\left((\bar{X}) + \frac{S^2(n-3)}{2n}\right)$, which is the MSE-optimal estimator (3.9) proposed by Zellner (1971) with the assumed known σ^2 replaced by $S^2 = Y^2/(n-1)$. Similarly, for the log-normal median (i.e., for $a = 1$, $b = 0$), we obtain $\hat{\theta}^{qb}_{1,0} = \exp\left((\bar{X}) - \frac{S^2 3}{2n}\right)$, again the MSE-optimal estimator (3.3) of Zellner (1971) with σ^2 replaced by its unbiased estimator.

As far as γ is concerned, we propose choosing a value close to the minimum value to assure for the existence of the first two posterior moments. Therefore, we specify the GIG with heaviest possible tail among those yielding $p(\theta|data)$ with finite variance:

$$(13) \quad \gamma_0^2 = 4\Big(\frac{a^2}{n} + b\Big) + \epsilon.$$

Note that γ_0 depends on n. In any case, we found in our simulations that $\hat{\theta}^{QB}_{1,0.5}$ are not particularly sensitive to alternative choices of γ_0 that are close to (i.e., of the same order of magnitude of) the γ_0 we propose. Much larger values lead to inefficient $\hat{\theta}^{QB}_{1,0.5}$ with far larger frequentist MSEs.

5 Conclusions

In this paper, we considered the popular log-normal model and the specific problems associated with estimating many of its parameters (including mean, median and mode). These problems are caused by the fact that $log-t$ and other distributions that can be met the analysis of the log-normal model have no finite moments. Our approach is Bayesian but parallel problems arise from a frequentist perspective. Specifically we wanted to continue using the popular quadratic loss function to summarize the posterior distribution. We found that a generalized inverse gamma prior for the population variance allows formally stating the conditions on the prior parameters that lead to posterior distributions with finite moments.

We complemented the results outlined in previous sections with a simulation exercise based on the same setting considered by Shen et al. (2006), Section 5. The main result of this simulation is that, adopting our prior specification the Bayes estimator of the log-normal mean based on quadratic loss is substantially equivalent to the estimator proposed by Shen et al. (2006), which has been proved to be superior to many of the alternatives previously proposed in the literature. Moreover our Bayes estimator of the mean has also smaller frequentist MSE than the Bayes estimators based on relative quadratic loss proposed by Rukhin (1986) and the one introduced in theorem 3.2.

We also considered alternative choices of the hyperameters when a guess of σ^2 is available. This leads to different choices of δ and λ. As it may be expected the use of a guess of σ^2 has a great potential for improving the MSE of the resulting Bayes estimators; anyway it may be detrimental when the guess is grossly wrong. We studied how to use the available prior guess of σ^2 to obtain a prior that is reasonably robust with respect to wrong guessing.

References

Abramowitz, A., and Stegun, I.A. (1968), Handbook of mathematical functions, (5^th edition), Dover, New York.

Barndorff-Nielsen, O.E. (1977), Exponentially decreasing distributions for the logarithm of particle size, *Proceedings of the Royal Statistical Society*, series A, 353, 401–419.

Bibby, B.M., and Sørensen M. (2003), Generalized hyperbolic and inverse Gaussian distributions: limiting cases and approximation of processes, in Rachev, S.T. (ed.) *Handbook of Heavy Tailed Distributions in Finance*, Elsevier Science B.V., 211–248.

Breymann, W., and Lüthi, D. (2010), `ghyp`: a package on Generalized Hyperbolic distributions, availalble at `cran.r-project.org/web/packages/ghyp/index.html`.

Eberlein, E., and von Hammerstein E.A. (2004), Hyperbolic processes in finance, in Dalang, R.C., Dozzi,M., Russo F. (eds.) (ed.) *Seminar on Stochastic Analysis, Random Fields and Applications IV, Progress in Probability*, Birkhäuser Verlag, 221–264.

Gelman, A., (2006) Prior distributions for variance parameters in hierarchical models, *Bayesian Analysis*, 1, 515–533.

Rukhin, A.L. (1986), Improved estimation in lognormal models, *Journal of the American Statistical Association*, 81, 1046–1049.

Shen, H., Brown, L.D., and Zhi H. (2006), Efficient estimation of log-normal means with application to pharmacokinetic data, *Statistics in Medicine*, 25, 3023–3038.

Zellner, A. (1971), Bayesian and non-Bayesian analysis of the log-normal distribution and log-normal regression, *Journal of the American Statistical Association*, 66, 327–330.