BAYESIAN ESTIMATION AND TESTING FOR CONTINUOUS ZERO-MODIFIED MODELS

FÉLIX LABRECQUE-SYNNOTT AND JEAN-FRANÇOIS ANGERS
DÉPARTEMENT DE MATHÉMATIQUES ET DE STATISTIQUE
UNIVERSITÉ DE MONTRÉAL
C.P. 6128 SUCC. "CENTRE-VILLE", MONTRÉAL, QC., CANADA H3C 3J7

Abstract. While discrete zero-inflated and zero-deflated models are fairly well known, continuous zero-modified models are mostly limited to the zero-inflated case in the literature, where they are often represented as a mixture between a probability density function and a Dirac mass at zero. An alternative formulation of zero-inflation for continuous models which can be extended to the zero-deflated case was introduced in Synnott & Angers (2010), where likelihood-based estimation was discussed. We consider here estimation and testing in a Bayesian framework. In particular, the choice of priors, posterior maximum and expectation estimators are discussed. Simulation results and an application to a real data set demonstrate the estimation and testing procedures in the context of continuous zero-deflated models.

1. Introduction

In the present paper, we consider continuous zero-modified models within a Bayesian framework. The parameter p controlling whether the model is zero-inflated, zero-deflated, or unmodified, is considered unknown; we are therefore interested in estimation or testing procedures providing good results in any of those cases.

Discrete zero-modified models, introduced in [14], are frequently encountered in the literature. Over the years, various and more complex models have been proposed, notably in [11], [8], [13]. For this class of models, the score test is the preferred method to test hypotheses on p; and difficulties in testing within the more common zero-inflated models are often related to the fact that the value of p under the null hypothesis lies at the boundary of the parameter space ([17], [9], and [7]). Discrete zero-deflated and zero-modified models where the sign of p is unknown are discussed in [2] and [3].

In the continuous case, until recently, only zero-inflated models were considered. Such models, and their estimation, are introduced in [1], and further discussed in [5] [15], [4], [6], and [16]. Zero-deflated continuous models were introduced in [10], where maximum likelihood estimation was considered. Here, we consider estimation within a Bayesian context, as well as hypothesis testing under the classical and Bayesian paradigms.

In section 2, we review the zero-modified continuous model and present appropriate priors. In section 3, we discuss parameter estimation and their asymptotic properties. Section 4 focuses on testing hypotheses on p. We present simulation

Date: March 15, 2010.

results and an application to a real data set in section 5. Section 6 contains some concluding remarks.

2. The model

Let X be a non-negative random variable such that it can be modelled by a probability density of the form

$$f(x|\theta, p, x_0) = p \times f_0(x) + (1-p) \times f_1(x|\theta), \tag{1}$$

where $f_0(x)$ is a probability density function on $[0, x_0]$, $f_1(x|\theta)$ is a probability density function on $\mathbb{R}_+$, and p is the zero-modification parameter, which can take both positive and negative values. For positive values of p, this probability density function can therefore be viewed as a mixture of f_0 and f_1. We cannot generally suppose that p or θ will be known; we therefore note $\omega = (p, \theta)$, the set of parameters of interests.

As discussed in [10], some constraints on $f_0(x)$ and p must be respected in order for $f(x|\omega)$ to be a continuous probability density function. In particular, we must have $\lim_{x\to x_0} f_0(x) = 0$, and

$$1 \geqslant p \geqslant \max_{x\in[0,x_0]} \frac{-f_1(x|\theta)}{f_0(x) - f_1(x|\theta)} = p_{min}(\theta). \tag{2}$$

The lower bound $p_{min}(\theta)$ depends on the parameter θ, creating a dependance between the parameters. It might sometimes be useful to work with

$$\tilde{p} = \begin{cases} p/|p_{min}(\theta)| & \text{if } p < 0, \\ p & \text{otherwise.} \end{cases} \tag{3}$$

In the absence of information on p, we consider non-informative priors on p and θ. Jeffreys prior is a poor choice here, as the form of the likelihood will usually necessitate the use of numerical integration or Monte-Carlo methods to obtain parameter estimates. The prior must therefore be evaluated at a large number of points. Since the Jeffreys prior is based on the Fisher information matrix, which must also be obtained numerically, its use leads to nested numerical procedures, which is much more computationally intensive than other non informative priors. Instead, we can use a suitable non-informative prior on θ, and either

$$\begin{aligned} \pi_1(p|\theta) &\sim \text{uniformly on} [p_{min}(\theta), 1], \\ \pi_2(p|\theta) &\sim \text{uniformly on} [p_{min}(\theta), 0] \text{ with probability 0.5,} \\ &\quad \text{and uniformly on} [0, 1] \text{with probability 0.5,} \end{aligned}$$

where choosing of $\pi_2(p|\theta)$ is equivalent to choosing a uniform distribution on $[-1, 1]$ for $\tilde{p}$.

In particular, we're interested in the zero-modified location-scale Laplace family of distributions truncated at 0. Let $f_0(x) \propto (x_0 - x)^\tau$ and

$$f_1(x|\mu, \lambda) = \frac{1}{\lambda\left[1 + \text{sign}(\mu)(1 - e^{-|\mu|/\lambda})\right]} \exp\left\{\frac{-|x-\mu|}{\lambda}\right\}, x \geqslant 0. \tag{4}$$

For negative values of μ, we can write

$$f_1(x|\mu,\lambda) = \frac{1}{\lambda\left[1-(1-e^{\mu/\lambda})\right]}\exp\left\{\frac{\mu-x}{\lambda}\right\}$$
$$f_1(x|\mu,\lambda) = \frac{1}{\lambda}\exp\left\{\frac{-x}{\lambda}\right\},$$

which is the exponential distribution and corresponds to the case $\mu = 0$. We will therefore restrict μ to non-negative values in order to have a identifiable model. We can then write (4) as:

$$f_1(x|\mu,\lambda) = \frac{1}{\lambda\left[2-e^{-\mu/\lambda})\right]}\exp\left\{\frac{-|x-\mu|}{\lambda}\right\}.$$

In this case, as μ and λ were originally a location and a scale parameter, respectively, improper non informative priors could be

$$\begin{aligned}\pi_3(\mu)&\propto 1, \forall\mu \geqslant 0,\\ \pi_4(\lambda)&\propto 1/\lambda, \forall\lambda > 0.\end{aligned} \tag{5}$$

3. Estimation

Let $x = (x_1,\ldots,x_n)$ be a sample of n i.i.d. random variables from a continuous zero-modified model $f(x_i|\omega)$. Within a Bayesian framework, inference on the parameters $(p,\theta) = \omega$ will be based on the posterior distribution $\pi(\omega|x)$, which is proportional to the product of the prior $\pi(\omega)$ and the likelihood $L(\omega|x) = \Pi_{i=1}^n f(x_i|\omega) = \Pi_{i=1}^n\left[pf_0(x)+(1-p)f_1(x|\theta)\right]$. Due to the form of the likelihood function, neither the posterior maximum nor the posterior expectations can generally be obtained analytically.

It is nonetheless rather easy to maximize the posterior, as most statistical software include routines to quickly maximize the likelihood of an arbitrary distribution. Such routines can easily be adapted to maximize the product of the likelihood and the prior. As in the frequentist framework, the main issue with maximum posterior estimates is that of multimodality: if the function to be maximized is not unimodal, there is a possibility that the algorithm used to compute the estimate converges to a local, rather than global, mode. In a Bayesian framework, possible prior-data conflicts can introduce multimodality in the posterior distribution. While non-informative priors are likely to be dominated by the likelihood function, and therefore carry a lesser risk of multimodality than informative priors, using the posterior expectation of the parameters as estimators completely sidesteps the multimodality issue.

However, the posterior expectation also cannot be explicitly obtained, and must be approximated by numerical quadrature or a Monte-Carlo estimate. The dependence between p and θ will complicate matters; in particular, since the support of p depends on the value of θ, direct numerical integration is not trivial. When using numerical quadrature, it is therefore useful to work with $\tilde{p}$ instead, as its support is always $[-1,1]$. Another problem is the curse of dimensonality: as the dimension of θ increases, the space where the posterior is nearly 0 grows much faster than the space where most of the posterior probability is concentrated. This is worsened by the parameter p, as typically a narrow range of values of θ will be compatible with zero-deflation, while another will be compatible with zero-inflation.

It is therefore possible that Monte-Carlo estimation with importance sampling require a very large random sample to converge, making it a slow and unattractive choice. An adaptive computation method, such as the quadrature proposed by [12] or the Monte-Carlo method proposed by [2], can produce estimates more quickly under a given convergence criterion. The method proposed in [2] is used with a transformation of the parameters $g : \Omega \to \mathbb{R}^k$. A random sample from the multivariate Student's t distribution with mean 0 and covariance matrix identity is generated. At each iteration, the importance weights and Monte-Carlo estimates for the mean vector and covariance matrix of $g(\omega)$ are computed. We then left-multiply the random sample by the matrix square root of the current estimate of the covariance matrix and add the current estimate of the mean vector, the goal of the method being to put the bulk of the random sample in the region of the (transformed) parameter space where the bulk of the probability density can be found. The method detailed in [12] is analogous, but based on Gauss-Hermite rather than Monte-Carlo integration; the quadrature node and re-scaled and translated at each iteration in a similar way.

If we choose to work with $\tilde{p}$, however, the integral

$$\int_\Omega g(\omega)\pi(\omega)L(\omega|x)d\omega$$

can be expressed as the iterated integral

$$\int_0^\infty \left[\int_0^\infty \left(\int_{-1}^1 g(\omega)\pi(\omega)L(\omega|x)dp \right) d\lambda \right] d\mu.$$

Gauss-Legendre quadrature can then be used to evaluate the innermost integral, while Gauss-Laguerre quadrature is used to integrate with respect to μ and λ. An advantage to using Gaussian quadrature is that the points ω_i where the integrand is evaluated depend only on the number of nodes, and not on the data x, the function of interest g, nor the prior π. The weights and nodes for Gaussian quadratures can be obtained from tables and are not recalculated for every integral; similarly, the bound p_{min} does not depend on the data and can also be tabulated, since a given node on Gauss-Legendre-Laguerre iterated quadrature will always have the same p_{min}. Since the numerical computation of p_{min} is the task with the highest computational cost for the evaluation of the integrand, this can save quite a bit of time, especially when a large number of integrals have to be computed (e.g., when analyzing multiple samples, or in simulation studies). This approach - tabulating the bound $p_{min}(\mu, \lambda)$ for different values of μ and λ - cannot be used when using adaptive quadrature methods (as the nodes will then be adjusted to better fit the data) or Monte-Carlo methods.

4. Hypothesis testing

Testing hypotheses of the form $H_i : p \in P_i$ within a Bayesian framework requires computing the posterior probability

$$\int_0^\infty \left[\int_0^\infty \left(\int_{P_i} \pi(\omega)L(\omega|x)dp \right) d\lambda \right] d\mu. \tag{6}$$

If multiple hypotheses are to be confronted, the hypothesis with the highest posterior probability is selected. When comparing two hypotheses, Bayesian testing can also be carried out through the use of Bayes factors, which can be obtained from

the prior and posterior probabilities. When considering a point hypothesis, such as $H_0 : p = 0$, it is necessary to add a point mass to our prior to have a nonzero posterior probability for the point hypothesis.

Equation (6) can be rewritten as $\int_\Omega g(\omega)\pi(\omega)L(\omega|x)d\omega$, with $g(\omega) = \mathbb{I}_{H_i}(p)$, where $\mathbb{I}$ is the indicator function. The problem of Bayesian hypothesis testing can therefore be reduced to the computation of the posterior expectation of a certain function $g(\omega)$, and will therefore face the same issues as described in the previous sections. The solutions considered are also the same, namely adaptive Monte-Carlo (as in [2]), adaptive quadrature (as in [12]), and Gauss-Legendre-Laguerre quadrature with tabulated bounds $p_{min}(\mu, \lambda)$.

5. Simulation and application results

The performance of Bayesian tests and point estimators can be assessed through a simulation study. Let $\tau = 0.05$, and $x_0 = 10$. For samples sizes $n = 10, 20, 30, 60,$ and 100, 500 random samples are generated with $p = -0.1, \mu = 20, \lambda = 20$. Table 1 lists the mean and variance of parameter estimates as a function of sample size. We consider both maximum posterior and posterior mean estimates; the latter are computed using an adaptive Monte-Carlo method, adaptive Gauss-Hermite quadrature, and Gauss-Legendre-Laguerre (GLL) iterated quadrature. The convergence criterion used for the Monte-Carlo and numerical quadrature methods is a maximal relative error of at most 1×10^{-3}. GLL quadrature fulfills this criterion with $20 \times 20 \times 20$ or fewer nodes for more than half the random samples. Adaptive Gauss-Hermite quadrature requires less nodes ($12 \times 12 \times 12$) for the same precision, but is slower as more complex computations are required (as the grid of nodes is iteratively translated and scaled) and since tabulated values of $p_{min}(\mu, \lambda)$ cannot be used. To obtain the adaptive Monte-Carlo estimates, samples of 750 triplets (x, y, z) are generated from the multinormal $(0, I_3)$ distribution and are used to estimate the transformed set of parameters $\left(\log(\mu), \log(\lambda), \log\left[-\log\left(\frac{\tilde{p}+1}{2}\right)\right]\right)$. These triplets are iteratively translated and scaled, and convergence is usually attained after 20 to 80 iterations. Non-adaptive importance sampling was also considered with a variety of importance functions and was mostly outperformed by the aforementioned estimation methods; these results are not reported here for the sake of brevity.

TABLE 1. Mean and variance of parameter estimates as a function of n, with $p = -0.1$, $\mu = \lambda = 20$

n		ML			MAP			Monte-Carlo			Gauss-Hermite			GLL		
		p	μ	λ	p	μ	λ	p	μ	λ	p	μ	λ	p	μ	λ
10	mean	-0.052	23.06	18.69	-0.046	24.54	16.01	-0.087	24.99	15.73	-0.111	23.28	21.78	-0.056	23.57	20.16
	var	0.008	54.23	58.93	0.0074	55.91	45.20	0.0059	52.72	42.05	0.002	70.96	35.16	0.0029	55.52	37.1
20	mean	-0.082	21.1	19.28	-0.079	21.38	17.98	-0.11	22.81	17.86	-0.126	21.25	19.64	-0.071	22.03	17.67
	var	0.0048	27.5	24.6	0.0058	30.4	23.39	0.0033	23.22	29.45	0.0017	25.76	28.31	0.0024	24.78	30.32
30	mean	-0.092	20.61	18.52	-0.086	21.08	17.58	-0.112	20.53	16.41	-0.135	23.57	19.91	-0.084	20.48	18.61
	var	0.0035	23.45	21.81	0.0033	23.15	20.2	0.0043	21.17	23.33	0.0023	27.43	24.61	0.0028	22.92	26.13
60	mean	-0.102	19.88	19.76	-0.1	20.01	19.33	-0.093	20.2	18.88	-0.123	22.23	17.37	-0.106	21.11	19.18
	var	0.0026	15.82	11.72	0.0025	15.33	11.94	0.0018	12.2	15.75	0.0019	13.3	18.62	0.0026	12.97	22.82
100	mean	-0.095	20.09	19.7	-0.094	20.23	19.4	-0.095	20.08	19.91	-0.112	20.93	19.85	-0.096	20.51	19.62
	var	0.0016	9.537	7.38	0.0016	9.14	7.103	0.0011	7.576	7.911	0.0013	8.21	9.42	0.0018	7.54	8.94

TABLE 2. Mean and variance of parameter estimates as a function of p
$\mu = \lambda = 20$, $n = 50$

p		ML			MAP			Monte-Carlo			Gauss-Hermite			GLL		
		p	μ	λ	p	μ	λ	p	μ	λ	p	μ	λ	p	μ	λ
-0.075	mean	-0.068	21.08	18.59	-0.065	21.35	17.67	-0.072	20.25	18.23	-0.069	20.23	18.25	-0.063	20.48	18.1
	var	0.0057	31.28	24.61	0.0067	34.01	22.94	0.0066	20.66	41.3	0.0061	20.72	42.26	0.0038	26.13	35.49
-0.05	mean	-0.047	20.51	19.23	-0.042	20.96	18.3	-0.056	20.75	17.18	-0.062	20.82	17.08	-0.056	21.18	18.62
	var	0.0067	22.48	19.77	0.0069	24.09	17.87	0.0088	19.78	37.8	0.0078	19.71	37.81	0.0022	18.85	26.46
-0.025	mean	-0.037	19.1	19.69	-0.034	19.38	18.75	-0.031	19.29	18.51	-0.032	19.23	18.47	-0.032	21.54	18.02
	var	0.0084	24.61	19.28	0.0086	24.34	18.42	0.0083	13.9	38.44	0.011	16.94	38.48	0.0054	19.28	22.88
-0.01	mean	-0.028	19.09	20.15	-0.024	19.43	19.17	-0.011	20.17	19.05	-0.0092	20.32	19.03	-0.013	20.92	18.32
	var	0.0061	33.65	30.55	0.0061	32.72	28.34	0.0073	16.24	32.87	0.0093	16.19	32.76	0.0046	28.88	30.29
-0.005	mean	-0.015	20.07	18.38	-0.015	20.54	17.44	-0.009	19.6	18.98	-0.011	19.66	19.03	-0.007	20.15	18.31
	var	0.0058	23.65	18.98	0.0048	24.14	17.53	0.012	15.52	46.65	0.009	15.57	44.24	0.003	18.47	23.32
0	mean	-0.0015	19.81	18.86	0.0046	20.34	17.83	-0.0018	19.9	18.22	-0.0039	19.79	18.52	-0.0036	20.64	17.82
	var	0.006	25.72	19.9	0.006	23.44	17.8	0.0105	15.39	43.79	0.0135	16.44	41.75	0.0035	18.82	27.28
0.05	mean	0.0091	16.87	20.64	0.0322	18.77	19.16	0.0621	19.68	18.15	0.0351	18.75	18.17	0.0057	21.86	17.51
	var	0.0077	57.98	25.62	0.0089	53.25	25.57	0.0122	13.74	36.84	0.0179	13.7	37.45	0.0075	26.15	31.02
0.1	mean	0.0191	15.66	20.27	0.019	16.41	19.03	0.0958	20.15	18.78	0.1292	20.13	18.75	0.092	21.24	18.44
	var	0.0114	65.34	23.66	0.0118	75.01	21.79	0.0084	17.19	34.09	0.0103	17.2	34.18	0.0091	19.98	33.42
0.2	mean	0.0601	9.97	21.41	0.0567	8.53	20.28	0.166	20.83	17.84	0.162	20.76	17.91	0.18	21.72	17.9
	var	0.0126	73.80	22.09	0.0122	78.38	20.78	0.0116	16.51	40.85	0.0103	16.54	39.77	0.0151	19.13	22.77
0.3	mean	0.109	5.258	22.12	0.114	3.186	20.57	0.271	20.31	18.30	0.286	20.28	18.34	0.281	21.12	18.02
	var	0.025	36.70	20.00	0.0235	41.88	19.52	0.011	17.68	36.2	0.0145	17.76	36.26	0.0197	21.01	28.27

While the estimates of μ and λ appear to be highly variable regardless of the estimation method, these variances do not seem out of place when considering the inverse of the Fisher information matrix:

$$I^{-1}(\omega) = \begin{bmatrix} 0.1302 & 6.8851 & -2.8064 \\ 6.8851 & 660.6605 & -256.2556 \\ -2.8064 & -256.2556 & 598.5605 \end{bmatrix}.$$

We next consider a fixed sample size $n = 60$ and let p take values between -0.1 and 1. In this scenario, the lower bound on p is $p_{min} = -0.1202$. For each value of p, 500 random samples have been generated; the mean and variance of the parameter estimates are reported in Table 2.

An especially interesting point is that simple maximization of the likelihood or posterior density simply breaks down for positive values of p, as both the bias and variance of the estimates of p and μ become large as p grows. This is most likely due to multimodality; for positive values of p, the zero-modified model is essentially a mixture model, and there is a possibility that the maximization algorithms converge to a local mode. Estimates obtained with these methods should be considered carefully when the estimate of p is positive; re-analyzing the data using a method well-suited to mixture models (e.g., based on the EM algorithm) is a possible solution. Posterior expectation estimates do not suffer from this disadvantage and remain reliable as p varies.

Also of interest is the fact that the adaptive Monte-Carlo, adaptive Gauss-Hermite quadrature and GLL quadrature all produce different estimates, and those differences can be important in some cases. This can be surprising, as all three methods give an estimate of the same quantities - the posterior expectation of p, μ, and λ. This difference can be partly explained by the use of iterative methods, which can sometimes converge to local rather than global modes. We also note that all three methods integrate transformed parameters, and then inversely transform the resulting estimates to obtain estimates of (p, μ, λ). In the case of the GLL, this transformation is simple –working with $\tilde{p}$ instead of p to ensure a constant domain of integration– but the adaptive methods take the logarithm of the parameters so that the transformed parameter space is $\mathbb{R}^3$.

We next consider testing hypotheses on p. As mentioned in section 4, posterior probabilities can be viewed as expectations of indicator functions – the testing methods considered will therefore be the three methods for obtaining posterior expectation estimates. Let $p = -0.1$. Figure 1 plots the posterior probabilities, obtained with the three estimation methods considered, of $H_- : p < 0$, $H_0 : p = 0$, and $H_+ : p > 0$ as a function of the sample size. The numerical quadratures lead to a more conservative test than adaptive Monte-Carlo. To test if $p = 0$, we can easily obtain $P(p \neq 0) = P(p < 0) + P(p > 0)$.

Under the classical paradigm, we consider two method for testing $H_0 : p = 0$ against $H_1 : p \neq 0$: the likelihood ratio test, and Rao's score test. Tables 3 and 4 compare the power of all testing procedures discussed in this paper (likelihood ratio, score, and posterior probabilities computed using adaptive Monte-Carlo, adaptive Gauss-Hermite quadrature and Gauss-Legendre-Laguerre quadrature) as a function of n and p, respectively. While there is no testing procedure uniformly better (or worse) than the others, we can note that the score test performs poorly when p is negative. The Bayesian test with adaptive Gauss-Hermite quadrature seems

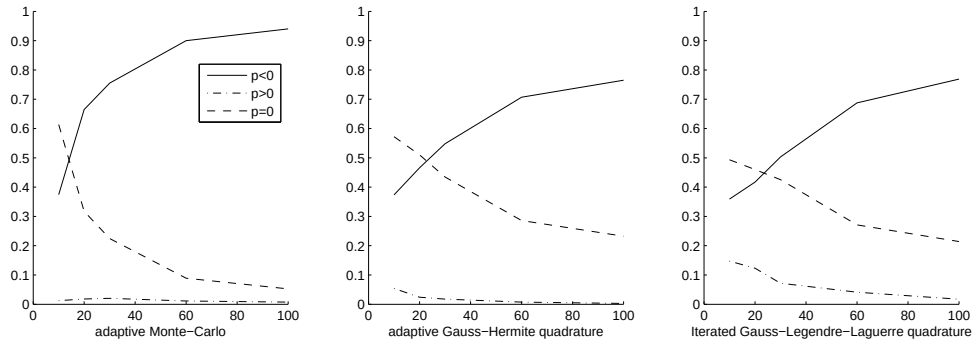

FIGURE 1. Posterior probabilities for tests as a function of n.

TABLE 3. Power as a function of n, with $p = -0.1$, $\mu = \lambda = 20$

n	LK	Score	Monte-Carlo	Gauss-Hermite	GLL
10	0.00	0.02	0.00	0.00	0.02
20	0.18	0.03	0.14	0.08	0.26
30	0.30	0.04	0.20	0.18	0.35
60	0.52	0.05	0.51	0.46	0.63
100	0.61	0.06	0.64	0.64	0.73

TABLE 4. Power as a function of p
$\mu = \lambda = 20$, $n = 50$

p	LK	Score	Monte-Carlo	Gauss-Hermite	GLL
-0.1	0.45	0.10	0.36	0.28	0.34
-0.075	0.32	0.08	0.27	0.19	0.21
-0.05	0.19	0.07	0.24	0.18	0.18
-0.025	0.07	0.07	0.12	0.13	0.15
-0.01	0.05	0.06	0.06	0.11	0.13
-0.005	0.03	0.05	0.09	0.09	0.08
0	0.05	0.06	0.06	0.06	0.05
0.05	0.01	0.09	0.13	0.17	0.09
0.1	0.07	0.19	0.24	0.29	0.16
0.2	0.15	0.37	0.51	0.65	0.41
0.3	0.22	0.43	0.70	0.85	0.78

the most powerful for positive values of p, but it is outperformed by both other Bayesian tests and by the likelihood ratio test when p is negative. Both classical tests are much faster than all three Bayesian tests; however, should Bayesian testing be preferable, the GLL approach is the fastest. This, combined with its generally good performance, can make it an attractive choice when working in a Bayesian context.

Finally, we apply these estimations methods to the data set analyzed in [10]. Parameter estimates, as can be seen in Table 5, are mostly consistent amongst the

TABLE 5. Estimation of a real data set of aggregate rainfall data

	p	μ	λ
ML	-0.0569	22.7	22.31
MAP	-0.0569	22.7	22.43
Monte-Carlo	-0.0557	22.61	22.45
Gauss-Hermite	-0.0551	22.61	22.48
GLL	-0.0562	22.58	22.74

TABLE 6. Testing a real data set of aggregate rainfall data

	χ^2_1	p-val	$P(p < 0\|x)$	$P(p = 0\|x)$	$P(p > 0\|x)$
LR	6.131	0.013	n/a	n/a	n/a
Score	8.696	0.003	n/a	n/a	n/a
Monte-Carlo	n/a	n/a	0.992	0.0063	0.0015
Gauss-Hermite	n/a	n/a	0.962	0.0321	0.0059
GLL	n/a	n/a	0.88	0.116	0.004

different methods; in particular, the MAP and ML estimates are nearly identical. This is not surprising, given our use of a non-informative prior and the large sample size ($n = 217$). The three Bayesian and two classical tests (using $\alpha = 0.05$), as seen in Table 6, are in agreement that p is negative, although the iterated GLL quadrature is still the most conservative.

6. Concluding remarks

In this paper, we have considered the use of continuous zero-modified models within a Bayesian framework. The choice of non-informative priors was discussed, and several methods for obtaining posterior expectation estimates have been proposed. The performance of these estimators has been assessed through a simulation study. While the three proposed methods of estimation of the posterior expectation are generally in agreement, they also outperform maximum likelihood and maximum posterior estimators when p is positive. Testing hypotheses on p in this kind of model was also considered, and can be viewed as the posterior expectation of an indicator function. Re-analysis of the aggregate rainfall data presented in [10] confirms that this data set is best modelled by a zero-deflated model.

References

[1] J. Aitchison. On the distribution of a positive random variable having a discrete probability mass at the origin. *Journal of the american statistical association*, 50(271):901–908, 1955.

[2] J.F. Angers and A. Biswas. A Bayesian analysis of zero-inflated generalized Poisson model. *Computational statistics & data analysis*, 42(1-2):37–46, 2003.

[3] E. Dietz and D. Böhning. On estimation of the Poisson parameter in zero-modified Poisson models. *Computational statistics & data analysis*, 34(4):441–459, 2000.

[4] M.V.M. Fernandes, A.M. Schmidt, and H.S. Migon. Modelling zero-inflated spatio-temporal processes. *Statistical Modelling*, 9(1):3, 2009.

[5] A. Feuerverger. On some methods of analysis for weather experiments. *Biometrika*, 66(3):655–658, 1979.

[6] D. Fletcher. Confidence intervals for the mean of the delta-lognormal distribution. *Environmental and Ecological Statistics*, 15(2):175–189, 2008.

[7] P.L. Gupta, R.C. Gupta, and R.C. Tripathi. Score test for zero inflated generalized Poisson regression model. *Communications in Statistics-Theory and Methods*, 33(1):47–64, 2005.
[8] D.B. Hall. Zero-inflated Poisson and binomial regression with random effects: a case study. *Biometrics*, 56(4):1030–1039, 2000.
[9] N. Jansakul and JP Hinde. Score tests for zero-inflated Poisson models. *Computational statistics & data analysis*, 40(1):75–96, 2002.
[10] F. Labrecque-Synnott and J.F. Angers. An Extension of zero-modified models to the continuous case. 2009.
[11] D. Lambert. Zero-inflated Poisson regression, with an application to defects in manufacturing. *Technometrics*, 34(1):1–14, 1992.
[12] JC Naylor and AFM Smith. Applications of a method for the efficient computation of posterior distributions. *Applied Statistics*, 31(3):214–225, 1982.
[13] M. Ridout, J. Hinde, and C.G.B. Demétrio. A score test for testing a zero-inflated Poisson regression model against zero-inflated negative binomial alternatives. *Biometrics*, 57(1):219–223, 2001.
[14] SN Singh. A note on inflated Poisson distribution. *Journal of the Indian Statistical Association*, 1(3):140–144, 1963.
[15] G. Stefansson. Analysis of groundfish survey abundance data: combining the GLM and delta approaches. *ICES Journal of Marine Science*, 53(3):577, 1996.
[16] L. Tian. Inferences on the mean of zero-inflated lognormal data: the generalized variable approach. *Statistics in medicine*, 24(20):3223–3232, 2005.
[17] J. Van den Broek. A score test for zero inflation in a Poisson distribution. *Biometrics*, 51(2):738–743, 1995.