EIEF WORKING PAPER SERIES

EIEF

Einaudi Institute for Economics and Finance

EIEF Working Paper 19/11

June 2019

Posterior moments and quantiles for the normal location model with Laplace prior

by

Giuseppe De Luca

(University of Palermo)

Jan R. Magnus

(Vrije Universiteit Amsterdam and Tinbergen Institute)

Franco Peracchi

(Georgetown University and EIEF)

Posterior moments and quantiles for the normal location model with Laplace prior

Giuseppe De Luca
University of Palermo, Italy

Jan R. Magnus
Vrije Universiteit Amsterdam and Tinbergen Institute, The Netherlands

Franco Peracchi*
Georgetown University, USA, and EIEF, Italy

June 2, 2019

Abstract.
We derive explicit expressions for arbitrary moments and quantiles of the posterior distribution of the location parameter η in the normal location model with Laplace prior, and use the results to approximate the posterior distribution of sums of independent copies of η.

Keywords.
Normal location model; Laplace priors; Reflected generalized gamma priors; Posterior moments and cumulants; Posterior quantiles; Cornish–Fisher approximation.

*Corresponding author. Email address: fp211@georgetown.edu (Franco Peracchi).

1 Introduction

Empirical data are often characterized by skewed distributions with fat tails, and a Bayesian analysis should explicitly account for these two features; see Fernández and Steel (1998) for an early analysis in this direction. Higher moments of posterior distributions are, however, difficult to obtain in general. In this note we derive explicit formulae for posterior moments and quantiles of any order in the important special case of a normal location model with Laplace (double-exponential) prior.

Let us consider a sample $x = (x_1, \ldots, x_n)$ from a univariate normal distribution with unknown mean η and unit variance, so the density of x given η (the likelihood) is

$$L(x;\eta) = (2\pi)^{-n/2} \exp\left[-\frac{1}{2}\sum_{i=1}^{n}(x_i - \eta)^2\right].$$

Suppose, in addition, that some information on η is available in the form of a prior density $\pi(\eta)$. The posterior density of η given x is then given by

$$p(\eta;x) = \frac{e^{-z^2/2}\pi(\eta)}{\int e^{-z^2/2}\pi(\eta)\,d\eta},$$

where

$$z = \sqrt{n}\left(\frac{1}{n}\sum_{i=1}^{n} x_i - \eta\right).$$

This intriguingly simple model is known as the normal location model with location parameter η; see Pericchi and Smith (1992), Mitchell (1994), and Magnus (2002). Because the distribution of z given η is $\mathcal{N}(0,1)$ for all values of n, there is no loss of generality in setting $n = 1$, which is the case commonly considered in the literature, so we shall also confine ourselves to this case.

The following two questions arise: How do we treat prior information or the lack of it (ignorance), that is, how do we choose the prior? And, what can we say about the posterior distribution? In particular, under what conditions can we obtain explicit formulae for posterior moments and quantiles?

The first question has received considerable attention. Dawid (1973) provided sufficient conditions for the posterior distribution of η to approach the prior distribution as x tends to infinity, so that an outlier has bounded and vanishing influence on the posterior distribution. Pericchi and Sansó (1995) presented a result closely related to Dawid's theorem, which enabled them to consider priors and likelihoods with bounded but non-vanishing influence on the posterior moments. The use of the Laplace prior as a basis for constructing a location invariant family for 'near ignorant' robust Bayesian analysis of credible intervals was considered by Pericchi and Walley (1991). Magnus (2002) and Kumar and Magnus (2013) discussed the choice of prior within the framework of neutrality and robustness, while O'Hagan and Pericchi (2012) focussed on how to handle conflicting information sources using Bayesian heavy-tailed models. The robustness of the normal location parameter was also considered by Choy and Smith (1997) and Choy and Walker (2003).

The second question has only received a partial answer, as authors have concentrated on the mean and variance of the posterior distribution of η; see Pericchi and Smith (1992), Mitchell (1994), and some generalizations in Pericchi and Sansó (1995). Meeden and Isaacson (1977) obtained some results on the rate of convergence of the posterior mean to the prior mean. Pericchi, Sansó, and Smith (1993), building on results of Meeden and Isaacson (1977), investigated the relationship between the posterior mean and the posterior mode; see also Kumar and Magnus (2013).

Pericchi, Sansó, and Smith (1993) also derived relationships for posterior moments and cumulants of η, but their results are restricted to the mean and variance only. These results are relevant in the areas of Bayesian robustness and approximation. In particular, results are obtained on the behavior of the posterior distribution for a large observation, generalizing Meeden and Isaacson (1977).

The purpose of this note is to derive closed-form expressions for any posterior moment or quantile of η. Higher-order moments (in fact, cumulants) are of interest in themselves, but they are also useful when we wish to obtain posterior quantiles for sums (or, more generally, linear combinations) of the parameters of interest, as for example in weighted-average least squares (WALS); see Magnus and De Luca (2016).

We shall concentrate primarily on Laplace priors, whose role in robust Bayesian inference was established by Pericchi and Smith (1992). Their use in the Bayesian representation of the LASSO (Tibshirani, 1996) was discussed by Li and Goel (2006), Park and Casella (2008), and Hans (2009).

In Section 2 we discuss our framework in a general context. Then, restricting ourselves to Laplace priors, we derive posterior quantiles and moments in Sections 3 and 4 respectively. Knowledge of the moments leads to knowledge of the cumulants, which have an important advantage over moments because the cumulant of sum of independent random variables is equal to the sum of the corresponding cumulants. Using the Cornish-Fisher expansion to obtain an approximation to the posterior distribution of this sum is discussed in Section 5.

2 Priors and posteriors

Although in this note we shall mostly be concerned with the Laplace prior, let us consider first a more general prior density π which satisfies the following three conditions: (i) is symmetric around zero, $\pi(-\eta) = \pi(\eta)$ for all $\eta > 0$; (ii) is positive and non-increasing on $(0, \infty)$; and (iii) is differentiable, except possibly at 0.

A flexible and mathematically tractable three-parameter class of priors satisfying these conditions is the family of reflected generalized gamma distributions, with densities of the form

$$\pi(\eta) = \frac{qc^\delta}{2\Gamma(\delta)} |\eta|^{-\alpha} e^{-c|\eta|^q}, \tag{1}$$

where $c > 0$, $q > 0$, $0 \leq \alpha < 1$, and $\delta = (1 - \alpha)/q$. This class includes as special cases the two-parameter families of Subbotin ($\alpha = 0$) and reflected Weibull distributions ($\alpha + q = 1$); and the one-parameter families of Laplace ($\alpha = 0$, $q = 1$) and normal distributions ($\alpha = 0$, $q = 2$, with zero mean and variance $1/(2c)$). A Laplace prior is thus a special case of both a Subbotin and a reflected Weibull prior.

Given the prior density (1), the resulting posterior density of η given x is

$$p(\eta; x) = \frac{|\eta|^{-\alpha}}{A(x)} \exp\left[-\frac{1}{2}\left((\eta - x)^2 + 2c|\eta|^q\right)\right], \tag{2}$$

where

$$\begin{aligned} A(x) &= \int_{-\infty}^{\infty} |\eta|^{-\alpha} \exp\left[-\frac{1}{2}\left((\eta - x)^2 + 2c|\eta|^q\right)\right] d\eta \\ &= \int_0^\infty \eta^{-\alpha} e^{-c\eta^q} \left(e^{-(\eta-x)^2/2} + e^{-(\eta+x)^2/2}\right) d\eta \end{aligned} \tag{3}$$

is a normalizing constant which depends only on x. Note that $A(-x) = A(x)$.

Since the prior distribution is symmetric, we have that $\Pr(\eta < 0) = \Pr(\eta > 0) = 1/2$. If, in addition, $\Pr(|\eta| < 1) = \Pr(|\eta| > 1) = 1/2$, we say that the prior is 'neutral', a concept attempting to capture the vague notion of ignorance in an explicit and applicable form. Neutrality occurs if and only if

$$\frac{1}{\Gamma(\delta)} \int_0^c t^{\delta-1} e^{-t} dt = 1/2.$$

In the case of the normal distribution we have $\delta = 1/2$, so that neutrality holds for $c = 0.2275$. In the case of the Laplace and reflected Weibull distributions we have $\delta = 1$, and neutrality holds for $c = \log 2 = 0.6931$. In the case of Subbotin no explicit solution exists.

Next, we define

$$A(a;x) = \int_{-\infty}^{a} |\eta|^{-\alpha} \exp\left[-\frac{1}{2}\left((\eta - x)^2 + 2c|\eta|^q\right)\right] d\eta,$$

so $A(x) = A(\infty; x)$ and the posterior distribution function of η may be written as

$$F(a;x) = \Pr(\eta \leq a|x) = \frac{A(a;x)}{A(x)}.$$

If $F(a;x) = p$, then the posterior quantile function Q is defined by the equation $Q(p;x) = a$. We are interested in the quantiles of the posterior distribution, but we might also be interested in the quantiles of the sum of independent random variables, each of which follows the distribution in (2). For this we need the moments (in fact, the cumulants) of each of the posteriors; see Section 5.

The moments of the posterior distribution are given by

$$\begin{aligned} \mathrm{E}(\eta^k|x) &= \frac{1}{A(x)} \int_{-\infty}^{\infty} \eta^k |\eta|^{-\alpha} \exp\left[-\frac{1}{2}\left((\eta - x)^2 + 2c|\eta|^q\right)\right] d\eta \\ &= \frac{1}{A(x)} \int_0^{\infty} \eta^{k-\alpha} e^{-c\eta^q} \left(e^{-(\eta-x)^2/2} + (-1)^k e^{-(\eta+x)^2/2}\right) d\eta, \end{aligned} \quad (4)$$

and the associated cumulants follow from these moments.

For the general class of priors considered in this section we don't obtain explicit formulae for the posterior distribution function and the posterior moments and quantiles. There is one prior however which allows such explicit formulae, namely the Laplace prior.

3 Posterior quantiles under a Laplace prior

In the special case of the Laplace prior we have $\alpha = 0$ and $q = 1$ and hence, as a special case of (1),

$$\pi(\eta) = \frac{c}{2} e^{-c|\eta|} \qquad (c > 0).$$

The posterior density of η given x, as a special case of (2), then takes the form

$$p(\eta;x) = \frac{1}{A(x)} \exp\left[-\frac{1}{2}\left((\eta - x)^2 + 2c|\eta|\right)\right]. \quad (5)$$

This distribution has received considerable attention (Pericchi and Smith, 1992; Mitchell, 1994; Magnus, 2002), and in this note we present expressions for any moment or quantile of the posterior distribution of η under a the Laplace prior.

We discuss posterior quantiles first. We do so by deriving the posterior distribution function of η and then, by inversion, its posterior quantile function. Defining

$$G(x) = e^{-cx}\Phi(x-c) + e^{cx}\Phi(-x-c),$$

where Φ denotes the cumulative distribution function of the standard normal distribution, as a special case of (3) we obtain the following expression for the normalizing constant in (4)

$$\begin{aligned}
A(x) &= \int_0^\infty e^{-c\eta}\left(e^{-(\eta-x)^2/2} + e^{-(\eta+x)^2/2}\right) d\eta \\
&= e^{c^2/2}e^{-cx}\int_0^\infty e^{-(\eta-x+c)^2/2}\,d\eta + e^{c^2/2}e^{cx}\int_0^\infty e^{-(\eta+x+c)^2/2}\,d\eta \\
&= e^{c^2/2}e^{-cx}\int_{-x+c}^\infty e^{-t^2/2}\,dt + e^{c^2/2}e^{cx}\int_{x+c}^\infty e^{-t^2/2}\,dt \\
&= \sqrt{2\pi}\,e^{c^2/2}G(x).
\end{aligned}$$

For $a \leq 0$ we can write the posterior distribution function of η as

$$\begin{aligned}
F(a;x) &= \frac{1}{A(x)}\int_{-\infty}^a \exp\left[-\frac{1}{2}\left((\eta-x)^2 - 2c\eta\right)\right] d\eta \\
&= \frac{e^{c^2/2}e^{cx}}{A(x)}\int_{-\infty}^a e^{-(\eta-x-c)^2/2}\,d\eta \\
&= \frac{e^{cx}\Phi(-x-c+a)}{G(x)}.
\end{aligned}$$

Similarly, for $a \geq 0$, we have

$$\begin{aligned}
F(a;x) &= F(0;x) + \frac{1}{A(x)}\int_0^a \exp\left[-\frac{1}{2}\left((\eta-x)^2 + 2c\eta\right)\right] d\eta \\
&= 1 - \frac{e^{-cx}\Phi(x-c-a)}{G(x)}.
\end{aligned}$$

Given the distribution function $F(a;x)$ we have, for $a \leq 0$,

$$F(a;x) = p \iff \Phi(-x-c+a) = pe^{-cx}G(x),$$

and similarly, for $a \geq 0$,

$$F(a;x) = p \iff \Phi(x-c-a) = (1-p)e^{cx}G(x).$$

Now using the fact that

$$a \leq 0 \iff F(a;x) \leq F(0;x) = \frac{e^{cx}\Phi(-x-c)}{G(x)},$$

we obtain the posterior quantile function of η as

$$Q(p;x) = \begin{cases} x + c + \Phi^{-1}\left(pe^{-cx}G(x)\right) & (p \leq F(0;x)), \\ x - c - \Phi^{-1}\left((1-p)e^{cx}G(x)\right) & (p > F(0;x)). \end{cases} \tag{6}$$

All posterior quantiles based on the Laplace prior can thus be computed explicitly. Note, however, that this applies only to one variable, not to a sum of independent variables. Quantiles for such a sum will be discussed in Section 5.

4 Posterior moments under a Laplace prior

Let us now consider the posterior moments of η under a Laplace prior. Using (4), these moments are given by

$$\begin{aligned} \mu_k = \mathrm{E}(\eta^k|x) &= \frac{1}{A(x)}\int_0^\infty \eta^k e^{-c\eta}\left(e^{-(\eta-x)^2/2} + (-1)^k e^{-(\eta+x)^2/2}\right) d\eta \\ &= \frac{(-1)^k e^{-cx}P_k(-c+x) + e^{cx}P_k(-c-x)}{G(x)}, \end{aligned} \tag{7}$$

with

$$P_k(a) = \int_{-\infty}^{a} (t-a)^k \phi(t)\,dt,$$

where ϕ denotes the density function of the standard normal distribution and we used the fact that

$$\begin{aligned} \int_0^\infty \eta^k e^{-c\eta} e^{-(\eta-x)^2/2}\,d\eta &= e^{c^2/2}e^{-cx}\int_0^\infty \eta^k e^{-(\eta+c-x)^2/2}\,d\eta \\ &= (-1)^k e^{c^2/2}e^{-cx}\int_{-\infty}^{-c+x}(t+c-x)^k e^{-t^2/2}\,dt \end{aligned}$$

and, similarly,

$$\int_0^\infty \eta^k e^{-c\eta}e^{-(\eta+x)^2/2}\,d\eta = (-1)^k e^{c^2/2}e^{cx}\int_{-\infty}^{-c-x}(t+c+x)^k e^{-t^2/2}\,dt.$$

The moments in (7) can be computed if we can find explicit expressions for the $P_k(a)$, which are closely related (but not identical) to the Hermite polynomials defined by

$$H_k(x) = (-1)^k e^{x^2/2} \frac{d^k}{d\,x^k} e^{-x^2/2}.$$

Let $p_k(t) = (t-a)^k \phi(t)$, so $P_k(a) = \int_{-\infty}^{a} p_k(t)\,dt$. Then,

$$\begin{aligned} p'_{k-1} &= (k-1)(t-a)^{k-2}\phi(t) + (t-a)^{k-1}\phi'(t) \\ &= (k-1)p_{k-2} - p_k - a p_{k-1}, \end{aligned}$$

using the fact that $\phi'(t) = -t\,\phi(t)$. Integrating then gives, for $k \geq 2$,

$$0 = p_{k-1}(a) = \int_{-\infty}^{a} p'_{k-1}(t)\,dt = (k-1)P_{k-2}(a) - P_k(a) - aP_{k-1}(a),$$

which leads to the recursion

$$P_k(a) = -aP_{k-1}(a) + (k-1)P_{k-2}(a) \qquad (k \geq 2),$$

with $P_0(a) = \Phi(a)$, $P_1(a) = -a\Phi(a) - \phi(a)$ as starting values. In contrast, the Hermite polynomials satisfy the recursion

$$H_k(a) = aH_{k-1}(a) - (k-1)H_{k-2}(a) \qquad (k \geq 2),$$

with $H_0(a) = 1$ and $H_1(a) = a$ as starting values.

It is now easy to obtain closed-form expressions for all posterior moments (and hence all posterior cumulants) of η. For example, the posterior mean of η is

$$\begin{aligned} \mu_1 &= \frac{-e^{-cx}P_1(-c+x) + e^{cx}P_1(-c-x)}{G(x)} \\ &= \frac{-e^{-cx}(c-x)\Phi(-c+x) + e^{cx}(c+x)\Phi(-c-x)}{e^{-cx}\Phi(-c+x) + e^{cx}\Phi(-c-x)} \\ &= x - h(x)c, \end{aligned}$$

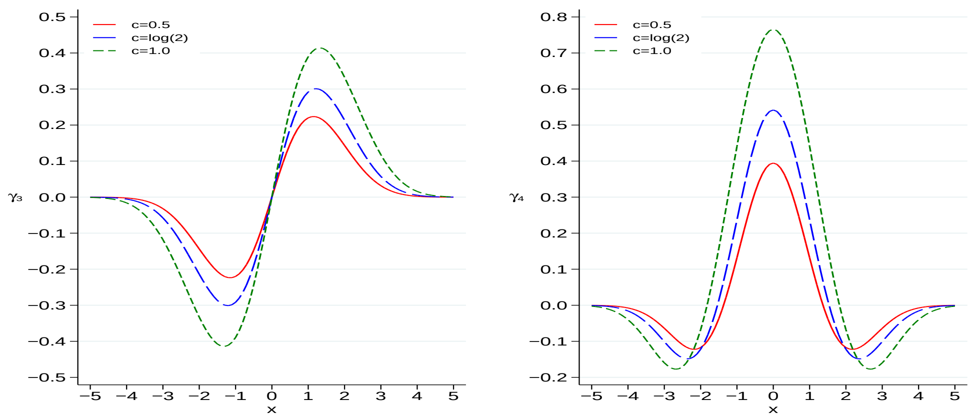
where

$$h(x) = \frac{e^{-cx}\Phi(-c+x) - e^{cx}\Phi(-c-x)}{e^{-cx}\Phi(-c+x) + e^{cx}\Phi(-c-x)}$$

and we used the fact that

$$e^{-cx}\phi(-c+x) = e^{cx}\phi(-c-x) = e^{-c^2/2}\phi(x),$$

Figure 1: Posterior skewness (left) and posterior excess kurtosis (right) of η



in agreement with Pericchi and Smith (1992, Eq. (6)). Similarly, from (7) and

$$
\begin{aligned}
P_2(a) &= (a^2+1)\Phi(a) + a\phi(a), \\
P_3(a) &= -(a^3+3a)\Phi(a) - (a^2+2)\phi(a), \\
P_4(a) &= (a^4+6a^2+3)\Phi(a) + (a^3+5a)\phi(a),
\end{aligned}
$$

we can easily derive closed-form expressions for the posterior variance of η, its posterior skewness

$$\gamma_3 = \frac{\mu_3 - 3\mu_1\mu_2 + 2\mu_1^3}{(\mu_2 - \mu_1^2)^{3/2}},$$

and its posterior excess kurtosis

$$\gamma_4 = \frac{\mu_4 - 4\mu_1\mu_3 + 6\mu_1^2\mu_2 - 3\mu_1^4}{(\mu_2 - \mu_1^2)^2} - 3.$$

The left panel of Figure 1 presents the posterior skewness of η as a function of x and c. The posterior distribution of η is negatively skewed for $x < 0$, positively skewed for $x > 0$, and symmetric for $x = 0$ (the prior mean). For $x > 0$, the skewness reaches a maximum of 0.22 at $x = 1.15$ when $c = 0.5$, a maximum of 0.30 at $x = 1.21$ when $c = \log(2)$, and a maximum of 0.41 at $x = 1.32$ when $c = 1$. As noted by Mitchell (1994), these results have implications for the ordering of the posterior mean, median and mode. Using similar arguments in a regression setup, Hans (2009) showed that the skewed posterior distribution of the regression parameters may lead to sizeable differences in Bayesian LASSO predictions based on alternative loss functions.

The right panel of Figure 1 presents instead the posterior excess kurtosis of η as a function of x and c. The excess kurtosis is positive (fatter tails) when $|x|$ is 'small', and negative (thinner tails) when $|x|$ is 'large'. There is no excess kurtosis at $|x| = 1.37$ when $c = 0.5$, $|x| = 1.54$ when $c = \log(2)$, and $|x| = 1.82$ when $c = 1$.

While posterior moments are more commonly employed than cumulants, working with cumulants has the advantage that the cumulant of a sum of independent random variables is just the sum of the corresponding cumulants. We employ this feature in the next section.

5 The Cornish–Fisher expansion

Consider a random variable z with cumulants $c_1, c_2, \dots$. Let $\gamma_k = c_k/c_2^{k/2}$, so γ_1 is the signal-to-noise ratio, $\gamma_2 = 1$, γ_3 is the skewness, and γ_4 is the (excess) kurtosis. Given the first five cumulants, define the function

$$\begin{aligned} C(t) = t &+ (t^2-1)\gamma_3/6 + (t^3-3t)\gamma_4/24 \\ &- (2t^3-5t)\gamma_3^2/36 + (t^4-6t^2+3)\gamma_5/120 \\ &- (t^4-5t^2+2)\gamma_3\gamma_4/24 + (12t^4-53t^2+17)\gamma_3^3/324. \end{aligned}$$

Given a particular value of p (e.g., $p = 0.95$), let t_p be the unique value of t for which $\Phi(t) = p$, where as before Φ denotes the standard normal cumulative distribution function. Thus, t_p denotes the pth quantile (or $100p$th percentile) of the standard normal distribution.

The pth quantile q_p of the distribution of z, defined implicitly by $\Pr(z \leq q_p) = p$, is then approximated by

$$q_p \approx c_1 + c_2^{1/2} C(t_p).$$

This is the Cornish–Fisher expansion (Cornish and Fisher, 1938; Fisher and Cornish, 1960; Chernozhukov, Fernández-Val, and Galichon, 2010), here based on five cumulants. Using a version with more than five cumulants will not necessarily produce a better approximation.

In our case, the expansion is useful when we wish to obtain quantiles of the sum $z = z_1 + \cdots + z_n$ of n independent random variables, each of which follows the posterior distribution (5). We know the quantiles of each of the random variables from (6), but we don't know the quantiles of their sum. Let $c_{i,k}$ denote the kth cumulant of z_i. Then the kth cumulant of z is given by $c_k = \sum_i c_{i,k}$. Given the cumulants of the individual components z_i we can thus approximate the quantiles of their sum z through the Cornish–Fisher expansion.

References

Chernozhukov, V., I. Fernández-Val, and A. Galichon (2010). Rearranging Edgeworth–Cornish–Fisher expansions, *Economic Theory*, 42, 419–435.

Choy, S. T. B. and A. F. M. Smith (1997). On robust analysis of a normal location parameter, *Journal of the Royal Statistical Society, Series B*, 59, 463–474.

Choy, S. T. B. and S. G. Walker (2003). The extended exponential power distribution and Bayesian robustness, *Statistics & Probability Letters*, 65, 227–232.

Cornish, E. A. and R. A. Fisher (1938). Moments and cumulants in the specification of distributions, *Revue de l'Institut International de Statistique*, 5, 307–320.

Dawid, A. P. (1973). Posterior expectations for large observations, *Biometrika*, 60, 664–666.

Fernández, C. and M. F. J. Steel (1998). On Bayesian modeling of fat tails and skewness, *Journal of the American Statistical Association*, 93, 359–371.

Fisher, R. A. and E. A. Cornish (1960). The percentile points of distributions having known cumulants, *Technometrics*, 2, 209–225.

Hans, C. (2009). Bayesian lasso regression, *Biometrika*, 96, 835–845.

Kumar, K. and J.R. Magnus (2013). A characterization of Bayesian robustness for a normal location parameter, *Sankhya (Series B)*, 75, 216–237.

Li, B. and P. K. Goel (2006). Regularized optimization in statistical learning: A Bayesian perspective, *Statistica Sinica*, 16, 411–424.

Magnus, J. R. (2002). Estimation of the mean of a univariate normal distribution with known variance, *Econometrics Journal*, 5, 225–236.

Magnus, J. R. and G. De Luca (2016). Weighted-average least squares: A review, *Journal of Economic Surveys*, 30, 117–148.

Meeden, G. and D. Isaacson (1977). Approximate behavior of the posterior distribution for a large observation, *Annals of Statistics*, 5, 899–908.

Mitchell, A. F. S. (1994). A note on posterior moments for a normal mean with double-exponential prior, *Journal of the Royal Statistical Society, Series B*, 56, 605–610.

O'Hagan, A. and L. Pericchi (2012). Bayesian heavy-tailed models and conflict resolution: A review, *Brazilian Journal of Probability and Statistics*, 26, 372–401.

Park, T. and G. Casella (2008). The Bayesian lasso, *Journal of the American Statistical Association*, 103, 681–686.

Pericchi, L. R. and B. Sansó (1995). A note on bounded influence in Bayesian analysis, *Biometrika*, 82, 223–225.

Pericchi, L. R., B. Sansó, and A. F. M. Smith (1993). Posterior cumulant relationships in Bayesian inference involving the exponential family, *Journal of the American Statistical Association*, 88, 1419–1426.

Pericchi, L. R. and A. F. M. Smith (1992). Exact and approximate posterior moments for a normal location parameter, *Journal of the Royal Statistical Society, Series B*, 54, 793–804.

Pericchi, L. R. and P. Walley (1991). Robust Bayesian credible intervals and prior ignorance, *International Statistical Review / Revue Internationale de Statistique*, 59, 1–23.

Tibshirani, R. (1996). Regression shrinkage and selection via the lasso, *Journal of the Royal Statistical Society, Series B*, 58, 267–288.