A Note on Bayes Estimates for Exponential Families

Eduardo Gutiérrez-Peña*

Departamento de Probabilidad y Estadística, IIMAS-UNAM

Apdo. Postal 20-726, Admón. 20, México D.F. 01000, Mexico

Manuel Mendoza

Departamento de Estadística, ITAM

Río Hondo 1, San Angel, México D.F. 01000, Mexico

January 4, 1999

Key words: Bayesian inference; Conjugate analysis; Quadratic variance function; Unbiased estimator

Short title: Bayes Estimates for Exponential Families

*Corresponding author

Abstract

For regular exponential families, Diaconis and Ylvisaker (1979) have shown that the posterior expectation of the mean parameter μ is linear in the canonical statistic, provided that a conjugate prior is used on the corresponding canonical parameter. In particular, the hyperparameters of the prior can be chosen so that this posterior expectation corresponds to an unbiased estimator of μ. In this paper we explore an analogue of this latter property for certain other parametrisations for which an unbiased estimator exists. The result is then used to find Bayes estimates under quadratic loss for such parametrisations.

Resumen

Diaconis e Ylvisaker (1979) han demostrado que, en el caso de las familias exponenciales regulares, el valor esperado final del parámetro medio μ es lineal como función de la estadística canónica siempre y cuando se utilice una distribución inicial conjugada relativa al parámetro canónico. En particular, los hiperparámetros de esta distribución pueden seleccionarse de manera que el valor esperado final de μ corresponda a un estimador insesgado. En este artículo se explora una propiedad análoga de ciertas parametrizaciones alternativas para las que existe un estimador insesgado. El resultado obtenido es usado entonces para encontrar estimadores de Bayes (bajo pérdida cuadrática) para dichas parametrizaciones.

1 Introduction

We first review some basic results concerning exponential families. See Barndorff-Nielsen (1978) for a comprehensive account of the properties of these models. Let η be a σ-finite positive measure on the Borel sets of $\mathbb{R}$, and consider the family $\mathcal{F}$ of probability measures whose density with respect to η is of the form

$$p(x|\theta) = b(x)\,\exp\{x\theta - M(\theta)\}, \qquad \theta \in \Theta, \tag{1}$$

for some function $b(\cdot)$, where $M(\theta) = \log \int b(x)\exp(x\theta)\,\eta(dx)$ and $\Theta = \operatorname{int}\Xi$, with $\Xi = \{\theta \in \mathbb{R} : M(\theta) < +\infty\}$. We assume that Θ is not empty. The family $\mathcal{F}$ is called a natural exponential family (NEF), and is said to be regular if Ξ is an open subset of $\mathbb{R}$. The function $M(\theta)$, called the cumulant transform of $\mathcal{F}$, is convex and analytic. A well-known result then states that

$$E[\,X \mid \theta\,] = \frac{dM(\theta)}{d\theta} \quad \text{and} \quad \mathrm{Var}[\,X \mid \theta\,] = \frac{d^2M(\theta)}{d\theta^2}.$$

The mapping $\mu = \mu(\theta) = \frac{dM(\theta)}{d\theta}$ is one-to-one and differentiable, with inverse $\theta = \theta(\mu)$, and provides an alternative parametrisation of $\mathcal{F}$, called the mean parametrisation. The set $\Omega = \mu(\Theta)$ is termed the mean parameter space. The function

$$V(\mu) = \frac{d^2M\{\theta(\mu)\}}{d\theta^2}, \qquad \mu \in \Omega,$$

is called the *variance function* of $\mathcal{F}$. An important property of $V(\cdot)$ is that the pair $(V(\cdot), \Omega)$ characterises $\mathcal{F}$ (see, for example, Morris, 1982).

Consider now a sample $\boldsymbol{X} = (X_1, \ldots, X_n)$ of independent observations from (1) and let $S = \sum_i X_i$. Let η_n denote the n-fold convolution of the measure η. Then the distribution of S has a density with respect to η_n of the form

$$p_n(s|\theta) = b(s,n)\,\exp\{s\theta - n\,M(\theta)\}, \qquad \theta \in \Theta,$$

where $b(\cdot, n)$ denotes the n-fold convolution of $b(\cdot)$.

In this paper we shall only be concerned with natural exponential families having a quadratic variance function (QVF), *i.e.* such that $V(\mu) = A\mu^2 + B\mu + C$ $(\mu \in \Omega)$ for some real constants A, B and C. Morris (1982) has shown that there exist only six different types of NEF-QVF models. Specifically, any real natural exponential family having a quadratic variance function can be obtained, via a nonsingular affine transformation and a power of convolution, from one of the following six basic families: *Normal:* $V(\mu) = 1$, with $\Omega = \mathbb{R}$; *Poisson:* $V(\mu) = \mu$, with $\Omega = \mathbb{R}_+$; *Binomial:* $V(\mu) = \mu(1-\mu)$, with $\Omega = (0,1)$; *Negative-Binomial:* $V(\mu) = \mu(1+\mu)$, with $\Omega = \mathbb{R}_+$; *Gamma:* $V(\mu) = \mu^2$, with $\Omega = \mathbb{R}_+$; and *Hyperbolic-Secant:* $V(\mu) = (1+\mu^2)$, with $\Omega = \mathbb{R}$. Thus, the NEF-QVF class includes some of the most widely used families of distributions. These families are all regular. See Morris (1982, 1983) for a thorough discussion of the properties of these models.

Consider the family $\mathcal{C}_\theta(\mathcal{F})$ of probability measures on Θ whose density function with respect to Lebesgue measure takes the form

$$\pi(\theta|\tilde{s}, \tilde{n}) = h(\tilde{s}, \tilde{n}) \exp\{\tilde{s}\theta - \tilde{n}\, M(\theta)\}, \qquad (\tilde{s}, \tilde{n}) \in \mathcal{H},$$

where $\tilde{s} \in \mathbb{R}$, $\tilde{n} \in \mathbb{R}$,

$$h(\tilde{s}, \tilde{n}) = \left\{ \int \exp\{\tilde{s}\theta - \tilde{n}\, M(\theta)\}\, d\theta \right\}^{-1},$$

and $\mathcal{H} = \operatorname{int}\{(\tilde{s}, \tilde{n}) \in \mathbb{R}^2 : h(\tilde{s}, \tilde{n})^{-1} < \infty\}$.

The family $\mathcal{C}_\theta(\mathcal{F})$ is essentially characterised by the property that the posterior expected value of the mean parameter μ is linear in the canonical sufficient statistic s (Diaconis and Ylvisaker, 1979), and is closed under sampling. We shall refer to $\mathcal{C}_\theta(\mathcal{F})$ as DY-conjugate to $\mathcal{F}$ with respect to θ. Provided $\mathcal{H}$ is not empty, $\mathcal{C}_\theta(\mathcal{F})$ is a 2-parameter exponential family with canonical 'statistic' $\boldsymbol{\tau}(\theta) = (\theta, -M(\theta))$ and canonical (hyper-)parameter $\tilde{\boldsymbol{z}} = (\tilde{s}, \tilde{n})$. Note that $h(\tilde{s}, \tilde{n})$ defines the normalising constant of the densities in $\mathcal{C}_\theta(\mathcal{F})$. We note in passing that Theorem 1 of Diaconis and Ylvisaker (1979) implies that $\mathcal{H}_0 = \{(\tilde{s}, \tilde{n}) \in \mathbb{R}^2 : \tilde{n} \in \mathbb{R}_+ \text{ and } \tilde{s} \in \tilde{n}\,\Omega\}$ is contained in $\mathcal{H}$. For a recent review of the theory of conjugate families for exponential family likelihoods, the reader is referred to Gutiérrez-Peña and Smith (1997).

Suppose now that θ has a standard conjugate distribution in $\mathcal{C}_\theta(\mathcal{F})$. Then it follows from the theorem in Section 5 of Jorgensen, Letac and Seshadri (1989) that the corresponding normalising constant can be written as

$$h(\tilde{s}, \tilde{n}) = \tilde{n}\, b(\tilde{s}, \tilde{n})\, V\left(\tilde{s}/\tilde{n}\right), \qquad (\tilde{s}, \tilde{n}) \in \mathcal{H}.$$

Let $\lambda = \lambda(\theta)$ be a one to one transformation of θ defining a reparametrisation of $\mathcal{F}$. Then $\mathcal{C}_\theta(\mathcal{F})$ induces a family of distributions on λ. We shall denote this family by $\mathcal{C}_\lambda^\theta(\mathcal{F})$ and call it DY-conjugate to $\mathcal{F}$ with respect to λ. Note that $\mathcal{C}_\theta^\theta(\mathcal{F}) = \mathcal{C}_\theta(\mathcal{F})$. The densities in $\mathcal{C}_\lambda^\theta(\mathcal{F})$ will be denoted by $\pi_\lambda^\theta(\lambda \mid \tilde{s}, \tilde{n})$.

We now give the definition of a conjugate parametrisation for natural exponential families, as this will be used in the sequel. A thorough account of the concept and properties of conjugate parametrisations can be found in Gutiérrez-Peña and Smith (1995).

Let $L_\theta(\theta \mid s, n)$ denote the likelihood function for θ given a sample of size n yielding the value s for the canonical statistic. Now consider an alternative parameterisation $\lambda = \lambda(\theta)$. The likelihood for λ is then given by $L_\lambda(\lambda \mid s, n) = L_\theta(\theta(\lambda) \mid s, n)$, where $\theta(\cdot)$ is the inverse of the transformation $\lambda(\cdot)$. A new parameterisation $\phi = \phi(\lambda)$ is said to be *conjugate* for λ, denoted $\phi \smile \lambda$, if and only if

$$|\dot{\phi}(\lambda)| \propto L_\lambda(\lambda \mid s_c, n_c)$$

for some real constants s_c and n_c, where the dot denotes differentiation. In other words, $\phi \smile \lambda$ if and only if the Jacobian of the transformation $\phi = \phi(\lambda)$ has the form of the likelihood of λ for some 'canonical statistic' s_c and some 'sample size' n_c.

Further, the relation '$\smile$' is an equivalence relation. Consonni and Veronese (1992) have proved that the mean parametrisation is conjugate for the canonical parameter if and only if the natural exponential family (1) has a quadratic variance function.

2 Unbiased estimators

2.1 Posterior Unbiasedness

Let $\mathcal{F}$ be a regular natural exponential family on $\mathbb{R}$. Suppose $X_1, \ldots, X_n$ is a sample of independent observations from $\mathcal{F}$ and recall $S = \sum_i X_i$. It is well-known that the canonical sufficient statistic S is complete (see, for example, Barndorff-Nielsen 1978, Lemma 8.2). Let $\mathcal{F}$ be indexed by $\lambda = \lambda(\theta)$ and suppose there exists an unbiased estimator $\hat{\lambda} = \hat{\lambda}(s, n)$ of λ, where $\hat{\lambda}(\cdot, n)$ is a one to one transformation of s for all n. It then follows from the Rao-Blackwell theorem that $\hat{\lambda}$ is the minimum variance unbiased estimator of λ (see, for example, Cox and Hinkley 1974, Chapter 8). We shall call $\hat{\lambda}$ the *sufficient unbiased estimator* of λ. In the particular case of the mean parameter μ, such an estimator not only exists but also attains the Cramer-Rao lower bound for the variance of unbiased estimators, μ being the only parametrisation with this property (up to affine transformations).

Recall that the sufficient unbiased estimator of μ is given by

$$\hat{\mu} = \hat{\mu}(s, n) = \frac{s}{n}.$$

Thus, for a DY-conjugate prior $\pi^{\theta}_{\mu}(\mu \mid s', n')$, we can write the posterior expected value of μ as

$$E[\, \mu \mid \boldsymbol{x} \,] = \hat{\mu}(s + s', n + n'). \tag{2}$$

Note that $\hat{\mu}$ corresponds to the case $s' = 0$ and $n' = 0$. We shall show that a condition analogous to (2) holds for other parametrisations of the family $\mathcal{F}$ for which an unbiased estimator exists. Consider the following example.

Example 1 Let $X_1, \ldots, X_n$ be an i.i.d. random sample from a negative exponential distribution with mean μ, and let $S = \sum_i X_i$. Then S follows a gamma distribution with density $Ga(s|n, \mu^{-1})$.

(a) Consider the class of transformations $\lambda_r(\mu) = \mu^r$, where r is a nonzero real constant, and let $\lambda = \lambda_r(\mu)$. It is then easy to see that the sufficient unbiased estimator of λ is given by

$$\hat{\lambda} = \hat{\lambda}(s, n) = \frac{\Gamma(n)}{\Gamma(n+r)}\, s^r,$$

provided that $n > \max(0, -r)$. The DY-conjugate prior for μ is inverse-gamma, with density $IGa(\mu|n' + 1, s')$. Thus, the posterior expected value of λ is

$$\begin{aligned} E[\,\lambda \mid \boldsymbol{x}\,] &= \frac{\Gamma(n'' - r + 1)}{\Gamma(n'' + 1)}\, s''^{\,r} \\ &= \hat{\lambda}(s'', n'' - r + 1), \end{aligned}$$

where $s'' = s + s'$ and $n'' = n + n'$. Note that if we set $s' = 0$ and $n' = r - 1$, then the posterior expected value of λ is precisely $\hat{\lambda}$, the sufficient unbiased estimator of λ. Moreover, the kernel of this prior density is then proportional to $V(\mu)^{-1}\, |\dot{\lambda}(\mu)|^{-1}$, where $\dot{\lambda}(\mu) = r\, \mu^{r-1}$ is the Jacobian of the transformation $\lambda = \lambda_r(\mu)$. We note that $\lambda \smile \mu$ in this case.

(b) Let $\lambda = \log \mu$. This transformation can be regarded as the limiting case of $\lambda_r(\cdot)$ as $r \to 0$. The sufficient unbiased estimator of λ is then given by

$$\hat{\lambda} = \hat{\lambda}(s, n) = \log s - \psi(n),$$

where $\psi(\cdot)$ denotes the digamma function. On the other hand, the posterior expected value of λ is

$$E[\,\lambda \mid \boldsymbol{x}\,] \;=\; \log s'' - \psi(n'' + 1)$$

$$= \hat{\lambda}(s'', n'' + 1).$$

As in the previous case, there exists a prior in the DY-conjugate family (namely that with $s' = 0$ and $n' = -1$), such that the posterior expected value of λ is precisely the sufficient unbiased estimator $\hat{\lambda}$. The kernel of this prior is, again, proportional to $V(\mu)^{-1} \, |J_\lambda(\mu)|^{-1}$. In this case $|J_\lambda(\mu)| = \mu^{-1}$, and so $\lambda \smile \mu$. $\triangleleft$

The phenomenon shown in the previous example holds for many of the most common exponential families and parametrisations for which a sufficient unbiased estimator exists. Whenever it occurs we shall say that the DY-conjugate family attains *posterior unbiasedness.* In Section 2.2 we show how to find unbiased estimators for certain parametrisations of an exponential family. Then, in Section 3, we explore conditions under which the corresponding posterior expectations coincide with these estimators.

2.2 Unbiased estimators for certain parametrisations of an exponential family

Here we briefly review a simple method for obtaining sufficient unbiased estimators for certain parametrisations of a given probability model (Guenther, 1978). As we shall see, such a method is particularly fruitful when applied to exponential family models.

By definition, $\hat{\lambda}(s, n)$ is an unbiased estimator of $\lambda = \lambda(\theta)$ if

$$\int \hat{\lambda}(s, n) \, p_n(s \mid \theta) \, \eta_n(ds) = \lambda(\theta).$$

This identity can be rewritten in the form

$$\int \hat{\lambda}(s, n) \, \frac{p_n(s \mid \theta)}{\lambda(\theta)} \, \eta_n(ds) = 1. \tag{3}$$

The method discussed by Guenther (1978) is based on inspection of (3). Specifically, if (3) can be written as

$$\int \hat{\lambda}(s,n)\,\rho(s,n)\,p_{n^*}(s \mid \theta_*)\,\eta_n(ds) = 1,$$

where $p_{n^*}(s \mid \theta_*)$ is a density from the same family as $p_n(s \mid \theta)$, θ_* is some value in the parameter space Θ and $n^* \in \mathbb{R}_+$, then, by virtue of the completeness of the sufficient statistic S, $[\hat{\lambda}(s,n)\,\rho(s,n) - 1] \equiv 0$ for all θ_* and

$$\hat{\lambda}(s,n) = 1/\rho(s,n).$$

We now apply this idea to the exponential family density (1). Consider transformations of the form

$$\lambda(\theta) = \exp\{r\theta - qM(\theta)\}, \tag{4}$$

for some real constants r and q such that $|r| + |q| \neq 0$. For (4) to be a reparametrisation of model (1), we require that $\dot{\lambda}(\theta)$ does not change sign as θ varies across Θ. In other words, r and q must be such that $[r - q\dot{M}(\theta)] \neq 0$ for all $\theta \in \Theta$. In particular, since $\dot{M}(\theta) = \mu$, it follows that q must be equal to zero if $\Omega = \mathbb{R}$.

Now, the sufficient unbiased estimator $\hat{\lambda}(s,n)$ is such that

$$\int_{\mathcal{S}} \hat{\lambda}(s,n)\,b(s,n)\,\exp\{s\theta - nM(\theta)\}\,\eta_n(ds) = \lambda(\theta),$$

where $\mathcal{S}$ denotes the support of the distribution of S. Let $\mathcal{T}_r = (\mathcal{S} + r) \cap \mathcal{S}$ and suppose $\mathcal{T}_r \neq \emptyset$. Then

$$\begin{aligned}
1 &= \int_{\mathcal{S}} \hat{\lambda}(s,n)\,b(s,n)\,\exp\{(s-r)\theta - (n-q)M(\theta)\}\,\eta_n(ds) \\
&= \int_{\mathcal{T}_r} \hat{\lambda}(s,n)\,\frac{b(s,n)}{b(s-r,n-q)}\,b(s-r,n-q)\,\exp\{(s-r)\theta - (n-q)M(\theta)\}\,\eta_n(ds) \\
&= \int_{\mathcal{T}_r - r} \hat{\lambda}(u+r,n)\,\frac{b(u+r,n)}{b(u,n-q)}\,b(u,n-q)\,\exp\{u\theta - (n-q)M(\theta)\}\,\nu(du),
\end{aligned}$$

where the convolution operator $b(\cdot,\cdot)$ is considered in the generalised version described in Section 3.5 of Gutiérrez-Peña and Smith (1997), ν is the dominating measure with respect to which the density of u is defined, and $\hat{\lambda}(s,n)$ is set to zero for $s \notin \mathcal{T}_r$. Thus, provided $0 < q < n$,

$$\int_{\mathcal{T}_r - r} \hat{\lambda}(u+r,n)\,\frac{b(u+r,n)}{b(u,n-q)}\,p_{n-q}(u \mid \theta)\,\nu(du) = 1,$$

from which it follows that

$$\hat{\lambda}(u+r,n) = \frac{b(u,n-q)}{b(u+r,n)}\,I_{\mathcal{T}_r - r}(u).$$

Therefore

$$\hat{\lambda}(s,n) = \frac{b(s-r,n-q)}{b(s,n)}\,I_{\mathcal{T}_r}(s), \tag{5}$$

where $I_{\mathcal{T}_r}(\cdot)$ denotes the indicator function of the set $\mathcal{T}_r$, is the sufficient unbiased estimator of λ.

3 Main result

Proposition 1 *Let $\mathcal{F}$ be a natural exponential family on $I\!R$ having a quadratic variance function, and let $\lambda = \lambda(\theta)$ be a reparametrisation of the form (4) for the family $\mathcal{F}$. Suppose that $\lambda \smile \theta$ and that θ has the (improper) prior distribution*

$$\pi(\theta) \propto |\dot{\lambda}(\theta)|^{-1} |\dot{\mu}(\theta)|. \tag{6}$$

Then the posterior expected value of λ is given by

$$E[\,\lambda \mid \boldsymbol{x}\,] = \hat{\lambda}(s,n).$$

Proof. Since the variance function is quadratic, we have

$$|\dot{\mu}(\theta)| \propto \exp\{B\theta + 2AM(\theta)\}.$$

(This follows from Example 3.5 of Gutiérrez-Peña and Smith, 1997). Similarly, since $\lambda \smile \theta$, then $\dot{\lambda}(\theta) \propto \exp\{k_1\theta - k_2 M(\theta)\}$ for some real constants k_1 and k_2. Hence the posterior distribution of θ under the prior (6) is given by

$$\pi(\theta) = h(s + B - k_1, n - 2A - k_2)\, \exp\{(s + B - k_1)\theta - (n - 2A - k_2)M(\theta)\}.$$

Now,

$$\begin{aligned} E[\,\lambda \mid \boldsymbol{x}\,] &= h(s + B - k_1, n - 2A - k_2) \\ &\times \int_\Theta \exp\{(s + B - k_1 + r)\theta - (n - 2A - k_2 + q)M(\theta)\}\, d\theta. \end{aligned}$$

In other words,

$$E[\,\lambda \mid \boldsymbol{x}\,] = \frac{h(s + B - k_1, n - 2A - k_2)}{h(s + B - k_1 + r, n - 2A - k_2 + q)}. \tag{7}$$

The proof now proceeds by direct comparison of (7) with (5) for each of the six natural exponential families having a quadratic variance function. The information provided in Table I is useful for carrying out such comparisons.

Table I about here

Normal.

$$p(x \mid \mu) = (2\pi)^{-1/2} \exp\left\{-\frac{1}{2}(x - \mu)^2\right\}, \qquad x \in \mathbb{R};\ \mu \in \mathbb{R}.$$

Written in terms of the usual parametrisation, μ, the transformation (4) takes the form $\lambda = \exp\{r\mu\}$. In this case,

$$b(s, n) = h(s, n) = (2\pi n)^{-1/2} \exp\left\{-\frac{s^2}{2n}\right\}.$$

Therefore

$$\begin{aligned} \hat{\lambda}(s, n) &= E[\,\lambda \mid \boldsymbol{x}\,] \\ &= \exp\left\{\frac{rs}{n} - \frac{r^2}{2n}\right\}. \end{aligned}$$

Poisson.

$$p(x \mid \mu) = \frac{\mu^x e^{-\mu}}{x!}, \qquad x = 0, 1, \ldots;\ \mu \in \mathbb{R}_+.$$

Written in terms of the usual parametrisation, μ, the transformation (4) takes the form $\lambda = \mu^r e^{q\mu}$. Here $b(s,n) = n^s/\Gamma(s+1)$ and $h(s,n) = n^s/\Gamma(s)$, whence

$$\begin{aligned} \hat\lambda(s,n) &= E[\,\lambda \mid \boldsymbol{x}\,] \\ &= \begin{cases} \frac{(n-q)^s}{n^s}\, I_{\mathcal{T}_r}(s) & \text{if } r = 0 \text{ and } 0 < q < n \\ \\ \frac{\Gamma(s+1)}{n^r\,\Gamma(s-r+1)}\, I_{\mathcal{T}_r}(s) & \text{if } r \in \mathbb{N} \text{ and } q = 0. \end{cases} \end{aligned}$$

Gamma.

$$p(x \mid \mu) = \mu^{-1} e^{-x/\mu}, \qquad x \in \mathbb{R}_+;\ \mu \in \mathbb{R}_+.$$

Written in terms of the usual parametrisation, μ, the transformation (4) takes the form $\lambda = \mu^{-q} e^{-r/\mu}$. In this case $b(s,n) = s^{n-1}/\Gamma(n)$ and $h(s,n) = s^{n+1}/\Gamma(n+1)$. Therefore

$$\begin{aligned} \hat\lambda(s,n) &= E[\,\lambda \mid \boldsymbol{x}\,] \\ &= \begin{cases} \frac{\Gamma(n)}{s^q\,\Gamma(n-q)}\, I_{\mathcal{T}_r}(s) & \text{if } r = 0 \text{ and } 0 < q < n \\ \\ \frac{(s-r)^{n-1}}{s^{n-1}}\, I_{\mathcal{T}_r}(s) & \text{if } r \in \mathbb{R}_+ \text{ and } q = 0. \end{cases} \end{aligned}$$

Binomial.

$$p(x \mid \mu) = \mu^x (1-\mu)^{1-x}, \qquad x = 0, 1;\ \mu \in (0,1).$$

Written in terms of the usual parametrisation, μ, the transformation (4) takes the form $\lambda = \mu^r (1-\mu)^{q-r}$. Here

$$b(s,n) = \frac{\Gamma(n+1)}{\Gamma(s+1)\,\Gamma(n-s+1)}$$

and

$$h(s,n) = \frac{\Gamma(n)}{\Gamma(s)\,\Gamma(n-s)}.$$

Therefore

$$\begin{aligned}\hat{\lambda}(s,n) &= E[\,\lambda \mid \boldsymbol{x}\,] \\ &= \begin{cases} \frac{\Gamma(n-q+1)\,\Gamma(n-s+1)}{\Gamma(n-q-s+1)\,\Gamma(n+1)}\, I_{\mathcal{T}_r}(s) & \text{if } r=0 \text{ and } q \in J_{n-1} \\ \\ \frac{\Gamma(n-r+1)\,\Gamma(s+1)}{\Gamma(s-r+1)\,\Gamma(n+1)}\, I_{\mathcal{T}_r}(s) & \text{if } r \in J_{n-1} \text{ and } q = r. \end{cases}\end{aligned}$$

Negative Binomial.

$$p(x \mid \nu) = \nu^x (1-\nu), \qquad x = 0, 1, \ldots;\ \nu \in (0,1).$$

Written in terms of the usual parametrisation, ν, the transformation (4) takes the form $\lambda = \nu^r(1-\nu)^q$. Here

$$b(s,n) = \frac{\Gamma(n+s)}{\Gamma(n)\,\Gamma(s+1)}$$

and

$$h(s,n) = \frac{\Gamma(n+s+1)}{\Gamma(n+1)\,\Gamma(s)},$$

from which it follows that

$$\begin{aligned}\hat{\lambda}(s,n) &= E[\,\lambda \mid \boldsymbol{x}\,] \\ &= \begin{cases} \frac{\Gamma(n-q+s)\,\Gamma(n)}{\Gamma(n-q)\,\Gamma(n+s)}\, I_{\mathcal{T}_r}(s) & \text{if } r=0 \text{ and } q \in J_{n-1} \\ \\ \frac{\Gamma(n+s-r)\,\Gamma(s+1)}{\Gamma(s-r+1)\,\Gamma(n+s)}\, I_{\mathcal{T}_r}(s) & \text{if } r \in J_{n-1} \text{ and } q = 0. \end{cases}\end{aligned}$$

Hyperbolic Secant.

$$p(x \mid \theta) = \{2\,\cosh(\pi x/2)\}^{-1}\,\cos(\theta)\,\exp\{x\theta\} \qquad x \in \mathbb{R};\ -\pi/2 < \theta < \pi/2.$$

Written in terms of the parametrisation θ, the transformation (4) takes the form $\lambda = \exp\{r\theta\}$. Here

$$b(s,n) = \frac{2^{n-2}}{\Gamma(n)\,\Gamma(n/2)} \left| \Gamma\left(\frac{n}{2} + \frac{\boldsymbol{i}s}{2}\right) \right|^2$$

and

$$h(s,n) = \frac{2^{n-2}\,(n^2+s^2)}{\Gamma(n+1)\,\Gamma(n/2)} \left| \Gamma\left(\frac{n}{2} + \frac{\boldsymbol{i}s}{2}\right) \right|^2,$$

where $\boldsymbol{i} = \sqrt{-1}$.

Therefore

$$\begin{aligned} \hat{\lambda}(s,n) &= E[\,\lambda \mid \boldsymbol{x}\,] \\ &= \frac{\left| \Gamma\left(\frac{n}{2} + \frac{\boldsymbol{i}(s-r)}{2}\right) \right|^2}{\left| \Gamma\left(\frac{n}{2} + \frac{\boldsymbol{i}s}{2}\right) \right|^2}, \end{aligned}$$

since $|\Gamma(u+\boldsymbol{i}v+1)|^2 = (u^2+v^2)\,|\Gamma(u+\boldsymbol{i}v)|^2$ (see Abramowitz and Stegun, 1965). $\diamond$

4 Bayes estimates under quadratic loss

It is well known that the Bayes estimate of a parameter λ under a quadratic loss function is given by the posterior expected value of λ. Proposition 1 therefore allows us to easily find Bayes estimates for parametrisations of the form (4), provided a DY-conjugate prior is used.

Corollary 1 *Let $\lambda = \lambda(\theta)$ be a parametrisation of the form (4) for the exponential family (1), and suppose λ has a distribution $\pi^{\theta}_{\lambda}(\lambda \mid s', n')$ in the DY-conjugate family $\mathcal{C}^{\theta}_{\lambda}(\mathcal{F})$. Then*

$$E[\,\lambda \mid \boldsymbol{x}\,] = \hat{\lambda}(s+s'-B+k_1, n+n'+2A+k_2), \tag{8}$$

where $\hat{\lambda}(s,n)$ is the sufficient unbiased estimator of λ. $\diamond$

Note that (8) generalises expression (2) to other parametrisations of an exponential family for which a sufficient unbiased estimator exist.

Acknowledgements. The authors are indebted to Raúl Rueda for his comments. This work was partially supported by the Sistema Nacional de Investigadores, Mexico.

References

1. Barndorff-Nielsen, O. (1978). *Information and Exponential Families in Statistical Theory*. Wiley, Chichester.

2. Morris, C.N. (1982). Natural Exponential Families with Quadratic Variance Functions. *The Annals of Statistics*, **10**, 65-80.

3. Morris, C.N. (1983). Natural Exponential Families with Quadratic Variance Functions: Statistical Theory. *The Annals of Statistics*, **11**, 515-529.

4. Diaconis, P. and Ylvisaker, D. (1979). Conjugate Priors for Exponential Families. *The Annals of Statistics*, **7**, 269-281.

5. Gutiérrez-Peña, E. and Smith, A.F.M. (1997). Exponential and Bayesian Conjugate Families: Review and Extensions (with discussion). *Test*, **6**, 1-90.

6. Jorgensen, B., Letac, G. and Seshadri, V. (1989). Sur une Propriété des Familles Exponentielles Naturelles de Variance Quadratique. *Canadian Journal of Statistics*, **17**, 1-8.

7. Gutiérrez-Peña, E. and Smith, A.F.M. (1995). Conjugate Parameterizations for Natural Exponential Families. *Journal of the American Statistical Association*, **90**, 1347-1356.

8. Consonni, G. and Veronese, P. (1992). Conjugate Priors for Exponential Families Having Quadratic Variance Functions. *Journal of the American Statistical Association*, **87**, 1123-1127.

9. Cox, D.R., and Hinkley, D.V. (1974). *Theoretical Statistics.* Chapman and Hall, London.

10. Guenther, W.C (1978). Some Easily Found Minimun Variance Estimators. *The American Statistician*, **32**, 29-34.

11. Abramowitz, M. and Stegun, I.A. (1965). *Handbook of Mathematical Functions.* Dover Publications, New York.

Table 1: Properties of the NEF-QVF class

Family	$M(\theta)$	A	B	Admissible values of r and q	k_1	k_2
Normal	$\theta^2/2$	0	0	$r \in \mathbb{R},\ q = 0$	r	0
Poisson	e^θ	0	1	$r = 0,\ 0 < q < n$	1	q
				$r \in \mathbb{N},\ q = 0$	r	0
Gamma	$-\log(-\theta)$	1	0	$r = 0,\ 0 < q < n$	0	$q-1$
				$r \in \mathbb{R}_+,\ q = 0$	r	0
Binomial	$\log(1+e^\theta)$	-1	1	$r = 0,\ q \in J_{n-1}$	1	$q+1$
				$r \in J_{n-1},\ q = r$	r	$q+1$
Negative Binomial	$-\log(1+e^\theta)$	1	1	$r = 0,\ q \in J_{n-1}$	1	$q-1$
				$r \in J_{n-1},\ q = 0$	r	0
Hyperbolic Secant	$-\log\{\cos(\theta)\}$	1	0	$r \in \mathbb{R},\ q = 0$	r	0

$J_n = \{1, 2, \ldots, n\}$