Abstract

We investigate a logarithmic-link generalized linear model, whose underlying sampling density is in an exponential family distribution with power variance function. The multiple-strata case is studied with stratum-dependent intercepts and a common slope. We prove that there exists a conjugate prior density on the intercept parameter, and the conjugate analysis is discussed. An estimation procedure is given, which includes the optimal estimating function of the parameters other than the intercept, and an empirical Bayes estimation of the hyper-parameters of the prior density. As an example, rainfall data for Queensland, Australia, is analyzed.

Key Words: common slope, conjugate analysis, empirical Bayesian estimation, estimating function, generalized linear model, logarithmic link, Pythagorean relationship, posterior mode, Tweedie distribution

1 Introduction

The generalized linear model (GLM, McCullagh & Nelder [14]) plays an important role in data analysis, enjoying wide application in fields such as insurance, climatology, economics and biostatistics. Their popularity is partially because GLMs are based on the exponential dispersion model (EDM) family of distributions, which includes common distributions such as the Normal, binomial, Poisson and gamma distributions as special cases. In addition, inference for the GLM has a minimax property: an exponential family distribution minimizes the Fisher information for the mean parameter under a given mean–variance relationship (Tsubaki [21]; Ohnishi & Tsubaki [15]).

If the response variable $Y$ follows an EDM distribution with mean $\mu$, the variance is $\text{var}(Y) = V(\mu)/\tau$, where $\tau > 0$ is a precision parameter, and $V(\mu)$ is the variance function, some function of the mean.

A special subset of EDMs are the Tweedie distributions, with variance function $V(\mu) = \mu^p$ for some $p$. Special cases include the Normal ($p = 0$), Poisson ($p = 1$ with $\tau = 1$), gamma ($p = 2$) and inverse Gaussian ($p = 3$) distributions. Tweedie distributions exist for all $p \notin (0, 1)$.

The Tweedie distributions with $p \in (1, 2)$ are of interest here. These distributions have support on the non-negative reals, and are continuous for $Y > 0$ with a positive probability $P(Y = 0)$. 
2 Generalized linear models, the exponential family
and location-dispersion models

The EDM family of distributions have probability functions

\[ f(y; \mu, \tau) = \exp\{\tau[c(\mu)y - M(\mu)]\} \alpha(y; \tau), \]

(1)

where \( \mu \) is the mean, and \( \tau > 0 \) is the precision parameter. The known function \( c(\mu) \) is the canonical parameter; the known function \( M(\mu) \), when written in terms of the canonical parameter, is the cumulant function; \( \alpha(y; \tau) \) is the supporting measure. The variance of \( Y \) is \( \text{var}[Y] = V(\mu)/\tau \) where the variance function \( V(\mu) \) is

\[ V(\mu) = \left( \frac{dc(\mu)}{d\mu} \right)^{-1}, \]

which uniquely identifies the distribution in the EDM family (Jørgensen [12, §2.3.1]). Use the notation ED(\( \mu, \tau \)) to denote a random variable \( Y \) has the EDM distribution in (1). As indicated earlier, common distributions such as the Normal, binomial, Poisson and gamma distributions are in the EDM family. EDMS are important as they are the response distributions for GLMS.

Closely related to the EDM family is the location-dispersion family (Jørgensen [12]). Distributions in this family have the form

\[ p(y - \mu, \tau) = \exp\{-\tau d(y - \mu) + b(\tau)\}, \]

(2)

for a normalizing constant \( \exp\{b(\tau)\} \), where \( \mu \) and \( \tau > 0 \) are the location (but not the mean) and precision parameters respectively. The function \( d(t) \) is called the unit deviance function, where \( d(0) = 0 \), and \( d(t) > 0 \) when \( t \neq 0 \); that is, \( d(t) \) is a distance measure.

GLMS consist of two components (McCullagh and Nelder [14]):
1. The response variable, $Y_i$, follows an EDM family distribution, with mean $\mu$ and precision parameter $\tau$ such that $Y_i \sim EDM(\mu_i, \tau w_i)$ where $w_i > 0$ are known prior weights; and

2. The expected values of the $Y_i$, say $\mu_i$, are related to the covariates $x_i$ through a monotonic, differentiable link function $h(\cdot)$ so that $h(\mu_i) = \alpha + x_i^T \beta$, where $\beta$ is a vector of unknown regression coefficients, and $\alpha$ represents a constant term.

Often the linear component, or linear predictor, $\alpha + x_i^T \beta$ is denoted by $\eta_i$, when $h(\mu_i) = \eta_i = \alpha + x_i^T \beta$.

3 Tweedie distributions

The EDMs with variance function $V(\mu) = \mu^p$ for some real $p$ are of interest here. For these EDMs,

$$f(y; \mu, \tau, p) = \exp\{\tau c(\mu, p)y - M(\mu, p)\} a(y; \tau, p),$$

with

$$V(\mu) = \left\{ \frac{\partial c(\mu, p)}{\partial \mu} \right\}^{-1} = \mu^p.$$

These EDM distributions are called the Tweedie distributions by Jørgensen [12] in honor of Tweedie [22]. Use the notation $ED_p(\mu, \tau)$ to denote a random variable $Y$ follows the Tweedie EDM in (3).

Jørgensen [12] shows distributions exist for all values of $p \not\in (0, 1)$. The Tweedie family includes important distributions, such as the Normal ($p = 0$), the Poisson ($p = 1$ with $\tau = 1$), the gamma ($p = 2$) and the inverse Gaussian ($p = 3$) distributions. The binomial distribution is a notable exception.

When $1 < p < 2$, the density (3) can be represented as a Poisson sum of gamma distributions, and is sometimes called the Poisson-gamma distribution. Suppose $N$ random variables $X_i$ (for $i = 1, \ldots, N$) are observed, where $N$ follows the Poisson distribution Po($m$) with mean (and variance) $m$. Also, suppose each random variable $X_i$ follows a gamma distribution with shape parameter $\theta$ and scale parameter $\lambda$ such that the mean is $\theta \lambda$ and variance $\theta \lambda^2$. Then the distribution of $\sum_{i=1}^{N} X_i$, where $N, X_1, X_2, \ldots$ are mutually independent, corresponds to an $ED_p(\mu, \tau)$ distribution where

$$m = \tau \frac{\mu^{2-p}}{2-p}, \quad \lambda = \frac{(p-1)\mu^{p-1}}{\tau}, \quad \theta = \frac{2-p}{p-1}.$$

The probability function for the Tweedie distribution with a power parameter $p \in (1, 2)$ is

$$f(y; \mu, \tau, p) = \exp\left\{ -\tau \left( -\frac{\mu^{1-p}}{1-p} y + \frac{\mu^{2-p}}{2-p} \right) \right\} a(y; \tau, p)$$

for $y > 0$ (Jørgensen [12, Chapter 4]), where

$$a(y; \tau, p) = \frac{1}{y} \sum_{j=1}^{\infty} \frac{\tau^j}{2-p} \left( \frac{\tau y}{p-1} \right)^{(2-p)/(p-1)} \Gamma((2-p)j/(p-1))j!.$$

The Tweedie distribution with $1 \leq p < 2$ has the positive probability of zero

$$P(Y = 0) = \exp\left( -\tau \frac{\mu^{2-p}}{2-p} \right).$$

The normalizing constant $a(y; \tau, p)$ cannot be written in closed form apart from the four special cases indicated earlier. To evaluate $a(y; \tau, p)$, many authors directly evaluate the
summation (5), but Dunn & Smyth [3] presented the first rigorous study of the series expansion, and found it is not possible to accurately evaluate this expansion for all parts of the parameter space. Dunn & Smyth [4] then developed a method for inverting the simple form of the moment generating function of the Tweedie densities, producing accurate computations in the parts of the parameter space where the series is not accurate. These algorithms are implemented in the \texttt{tweedie} package (Dunn [2]) for the \texttt{R} statistical environment [18], and we use these programs to derive our numerical results.

In Tweedie GLMs with \( p > 1 \), a logarithm link function is commonly used, since it ensures \( \mu > 0 \) as required. For later convenience, write the Tweedie density in (4) as

\[
f(y; \mu, \tau, p) = \exp \left[ -\tau \left\{ u(\mu; 2-p) - yu(\mu; 1-p) \right\} \right] \tilde{a}(y; \tau, p),
\]

for \( \tilde{a}(y; \tau, p) = \exp \left[ -\tau \left\{ -(1-p)^{-1}y + (2-p)^{-1} \right\} \right] a(y; \tau, p) \), where

\[
u(t; \kappa) = \begin{cases} \frac{\log t}{t^\kappa - 1} & \text{for } \kappa = 0, \\ \frac{\log t}{\kappa} & \text{otherwise}. \end{cases}
\]

In this form, \( u \) is continuous in \( \kappa \), and is equivalent to the log-limit form used by Dunn & Smyth [3]. The function \( u(t; \kappa) \) proves crucial later, based on the following properties. The proof is a straightforward calculation, and is omitted.

**Lemma 1** Consider a function \( u(t; \kappa) \) as defined in (7).

(i) For any \( s, t \) and \( \kappa \), \( u(st; \kappa) = t^\kappa u(s; \kappa) + u(t; \kappa) \).

(ii) Suppose \( \kappa \) and \( \nu \) are non-negative and \( \kappa + \nu > 0 \). Then, for any \( q > 0 \) and \( r > 0 \)

\[
q u(t; \kappa) - ru(t; -\nu) = \delta_* \left\{ u(t/t_*; \kappa) - u(t/t_*; -\nu) \right\} + q u(t_*; \kappa) - r u(t_*; -\nu),
\]

where \( t_* = (r/q)^{\frac{\kappa}{\kappa+\nu}} \) and \( \delta_* = q^{\frac{\kappa}{\kappa+\nu}} r^{\frac{\nu}{\kappa+\nu}} \).

4 The likelihood function for the Tweedie glm

Although an extension to the vector slope parameter is straightforward, we now focus on the scalar case for simplicity.

In this paper, interest is in the logarithmic link Tweedie GLM with linear predictor. Suppose \( Y_i \) is distributed according to \( \text{ED}_p(\mu_i, \tau) \) with \( \log \mu_i = \alpha + \beta x_i (1 \leq i \leq n) \); then using the expression (6), the density function is given as

\[
\prod_{i=1}^{n} f(y_i; e^{\alpha+\beta x_i}, \tau, p) = \exp \left[ -\tau \sum_{i=1}^{n} \left\{ u(e^{\alpha+\beta x_i}; 2-p) - y_i u(e^{\alpha+\beta x_i}; 1-p) \right\} \right] \prod_{i=1}^{n} \tilde{a}(y_i; \tau, p).
\]

For use later, use this density function to form the likelihood function, and re-write separating \( \alpha \) and \( \beta \), as shown in the following Proposition.

**Proposition 1** The Tweedie likelihood in (8) can be written as

\[
\prod_{i=1}^{n} f(y_i; e^{\alpha+\beta x_i}, \tau, p) = \exp \left[ -n \tau \left\{ A(\beta, p) u(e^{\alpha}; 2-p) - B(\beta, p) u(e^{\alpha}; 1-p) \right\} \right] \prod_{i=1}^{n} f(y_i; e^{\beta x_i}, \tau, p).
\]
where

\[ A(\beta, p) = \frac{1}{n} \sum_{i=1}^{n} e^{(2-p)\beta z_i} \quad \text{and} \quad B(\beta, p) = \frac{1}{n} \sum_{i=1}^{n} y_i e^{(2-p)\beta z_i}. \]

As well as noting the separation between \( \alpha \) and \( \beta \), note the symmetry between the two components involving \((1 - p)\) and \((2 - p)\) in (9). This proves useful when the conjugate analysis of the intercept \( \alpha \) is studied in Sect. 5.

**Proof**  Apply Lemma 1(i) to the likelihood function based on the density function in (8); the summation term in the exponent of the likelihood function becomes

\[
\sum_{i=1}^{n} \left\{ u(e^{\alpha+\beta z_i}; 2 - p) - y_i u(e^{\alpha+\beta z_i}; 1 - p) \right\}
\]

\[
= \sum_{i=1}^{n} \left\{ e^{(2-p)\beta z_i} u(e^{\alpha}; 2 - p) - y_i e^{(1-p)\beta z_i} u(e^{\alpha}; 1 - p) \right\}
+ u(e^{\beta z_i}; 2 - p) - y_i u(e^{\beta z_i}; 1 - p) \right\}
\]

\[
= n \left\{ A(\beta, p) u(e^{\alpha}; 2 - p) - B(\beta, p) u(e^{\alpha}; 1 - p) \right\}
+ \sum_{i=1}^{n} \left\{ u(e^{\beta z_i}; 2 - p) - y_i u(e^{\beta z_i}; 1 - p) \right\}.
\]

From (8), the middle expression is equivalent to the summation term in the exponent of the product of the density function in (10);

\[
\prod \left\{ u(e^{\alpha+\beta z_i}; 2 - p) - y_i u(e^{\alpha+\beta z_i}; 1 - p) \right\}
\]

then using the given definitions of \( A(\beta, p) \) and \( B(\beta, p) \), the result follows. \( \square \)

Now, we consider a specific Tweedie GLM using a logarithmic link function, where the data consist of \( K \) strata with \( n_k \) observations in stratum \( k \). For each stratum, assume separate intercepts \( \alpha_k \) but common slope \( \beta \) for covariate \( x_{ki} \); primary interest is in the slope \( \beta \).

The density function is

\[
f(y; \alpha, \beta, \tau, p) = \prod_{k=1}^{K} f(y_k; \alpha_k, \beta, \tau, p)
= \prod_{k=1}^{K} \prod_{i=1}^{n_k} f(y_{ki}; \mu_{ki}, \tau, p),
\]

where \( \log \mu_{ki} = \alpha_k + \beta x_{ki} \). In the above, \( y = (y_1, \ldots, y_K)^T \) with \( y_k = (y_{k1}, \ldots, y_{kn_k})^T \), \( \alpha = (\alpha_1, \ldots, \alpha_K)^T \) is the intercept parameter vector, \( \beta \) is the common slope parameter, and \( x_{ki} \) are the covariates.

To simplify the notation, we define several quantities. We extend \( A(\beta, p) \) and \( B(\beta, p) \) in Proposition 1 to the multi-strata case as

\[
A_k(\beta, p) = \frac{1}{n_k} \sum_{i=1}^{n_k} e^{(2-p)\beta z_{ki}} \quad \text{and} \quad B_k(\beta, p) = \frac{1}{n_k} \sum_{i=1}^{n_k} y_{ki} e^{(1-p)\beta z_{ki}},
\]

respectively. We also introduce the following quantities:

\[
\hat{\delta}_{M_k} = \hat{\delta}_{M_k}(\beta, p) = \log \frac{B_k(\beta, p)}{A_k(\beta, p)} \quad \text{and} \quad \
\delta_{M_k} = \delta_{M_k}(\beta, \tau, p) = \tau \left\{ A_k(\beta, p) \right\}^{p-1} \left\{ B_k(\beta, p) \right\}^{2-p}.\]
The former will be shown to be the maximum likelihood estimator (MLE) for $\alpha$ given $(\beta, p)$.

Interestingly, the likelihood function with respect to the intercept parameter $\alpha$ is similar to that of the location–dispersion family, as shown in the following proposition.

**Proposition 2** The likelihood function corresponding to (10) is

$$f(y; \alpha, \beta, \tau, p) = f(y; 0, \beta, \tau, p) \exp \left[ - \sum_{k=1}^{K} \delta_{Mk} n_{k} L(\alpha_{k} - \hat{\alpha}_{Mk}; p) \right] \times \prod_{k=1}^{K} \exp \{ -n_{k} C_{k}(\beta, \tau, p) \},$$

where 0 is the $K$-dimensional zero vector,

$$L(t; p) = u(e^{t}; 2 - p) - u(e^{t}; 1 - p), \quad C_{k}(\beta, \tau, p) = \tau \{ A_{k}(\beta, p) u(e^{\alpha}); 2 - p) - B_{k}(\beta, p) u(e^{\alpha}; 1 - p) \}.$$

Notice the form of (12) is like that of the location–dispersion family (2), where $\tau \equiv \delta_{Mk} n_{k}$ and $d(y - \mu) \equiv L(\alpha_{k} - \hat{\alpha}_{Mk}; p)$. Note that $L(t; p) > 0$ for $t \neq 0$, and $L(t; p) = 0$ if $t = 0$, as required for a deviance function.

**Proof** As an application of Proposition 1 to the $k$th stratum, we have

$$f(y_{k}; \alpha_{k}, \beta, \tau, p) = \exp \left[ -\tau n_{k} \left\{ A_{k}(\beta, p) u(e^{\alpha_{k}}); 2 - p) - B_{k}(\beta, p) u(e^{\alpha_{k}}); 1 - p) \right\} \right] \times f(y_{k}; 0, \beta, \tau, p).$$

We apply Lemma 1(ii) to the linear combination of the $u(\cdot; \cdot)$ terms in the exponent. Recalling the definitions of $L(t; p)$ and $C_{k}(\beta, \tau, p)$, we see that

$$A_{k}(\beta, p) u(e^{\alpha_{k}}); 2 - p) - B_{k}(\beta, p) u(e^{\alpha_{k}}); 1 - p) = \frac{1}{\tau} \{ \delta_{Mk} L(\alpha - \hat{\alpha}; p) + C_{k}(\beta, \tau, p) \}.$$

Combining (14) and (15), the likelihood function for stratum $k$ is

$$f(y_{k}; \alpha_{k}, \beta, \tau, p) = f(y_{k}; 0, \beta, \tau, p) \exp \{ -\delta_{Mk} n_{k} L(\alpha_{k} - \hat{\alpha}_{Mk}; p) - n_{k} C_{k}(\beta, \tau, p) \},$$

which completes the proof. □

The function $L(t; p)$ is used as a loss function in the conjugate analysis discussed in Sect. 5. Proposition 2 shows that $\hat{\alpha}_{Mk}(\beta, p)$ is the MLE for $\alpha$ given $\beta$ and $p$.

5 Conjugate analysis of the intercept

Motivated by the result of Proposition 2, assume the following prior density on $\alpha_{k}$, which is in the location–dispersion family (2):

$$\pi(\alpha_{k} - \alpha_{0}; p, \delta_{n_{k}}) = \frac{1}{K(p, \delta_{n_{k}})} \exp \{ -\delta_{n_{k}} L(\alpha_{k} - \alpha_{0}; p) \}.$$
Here $\alpha_0$ and $\delta > 0$ are hyper-parameters, $L(t; p)$ is the deviance function defined in (13), and

$$K(p, t) = \int_{-\infty}^{\infty} \exp\{-tL(s; p)\} \, ds$$  \hfill (18)

is the normalizing constant. When $p = 1$ or $p = 2$, the density (17) is a log-transformed gamma density.

The prior density (17) may be derived in two different ways.

1. The first is based on the likelihood approach, related to the notion of the power prior proposed by Ibrahim & Chen [10]. To see this, consider (16) as a likelihood function of $\alpha_k$, supposing the other parameters are known. Replace $\hat{\alpha}_{Mk}$ and $\delta_{Mk}$ with the hyper-parameters $\alpha_0$ and $\delta$, respectively, and the assumed prior density is obtained.

2. The second is an application of the method in Yanagimoto & Ohnishi [23]. The Kullback-Leibler divergence between two Tweedie densities is

$$KL(f(y; \mu_1, \tau, p), f(y; \mu_2, \tau, p)) = \tau \mu_1^{2-p}L(\log \mu_2 - \log \mu_1; p).$$

Thus the Kullback-Leibler divergence from model $f(y_k; \alpha_{k1}, \beta, \tau, p)$ to $f(y_k; \alpha_{k2}, \beta, \tau, p)$ is

$$KL(\alpha_{k1}, \alpha_{k2}; \beta, \tau, p) = \tau L(\alpha_{k2} - \alpha_{k1}; p) \sum_{i=1}^{n_k} e^{(2-p)(\alpha_{0} + \beta g_{ki})}.$$  

Consider the prior density proportional to $\exp\{-\tilde{\delta}KL(\alpha_0, \alpha_k; \beta, \tau, p)\}$. Substitution of $\tilde{\delta} \tau \sum e^{(2-p)(\alpha_{0} + \beta g_{ki})}$ with $\delta n_k$ gives the assumed prior density.

Prove the conjugacy of the assumed prior density (17) using Lemma 1.

**Proposition 3** The posterior density corresponding to the prior density (17) under the sampling density $f(y_k; \alpha_k, \beta, \tau, p)$ is $\pi(\alpha_k, \beta, \tau, p)$, where

$$\hat{\alpha}_{Bk} = \hat{\alpha}_{Bk}(\beta, \tau, p, \alpha_0, \delta) = \log \frac{\tau B_k(\beta, \tau, p, \alpha_0, \delta) + \delta}{\tau A_k(\beta, \tau, p, \alpha_0, \delta) + \delta},$$

$$\delta_{Bk} = \delta_{Bk}(\beta, \tau, p, \alpha_0, \delta)$$

$$= \tau A_k(\beta, \tau, p, \alpha_0, \delta)^{2-p} \tau B_k(\beta, \tau, p, \alpha_0, \delta)^{p-1}.$$  \hfill (19)

Therefore, the prior density (17) is conjugate.

**Proof** From Lemma 1(i),

$$L(\alpha_k - \alpha_0; p) = u(e^{a_k - a_0}; 2 - p) - u(e^{a_k - a_0}; 1 - p)$$

$$= e^{-(2-p)a_0} u(e^{a_k}; 2 - p) - e^{-(1-p)a_0} u(e^{a_k}; 1 - p) + L(-\alpha_0; p).$$

This, together with Lemma 1(ii), gives

$$\tau \left\{ A_k(\beta, p) u(e^{a_k}; 2 - p) - B_k(\beta, p) u(e^{a_k}; 1 - p) \right\} + \delta L(\alpha_k - \alpha_0; p)$$

$$= \left\{ \tau A_k(\beta, p) + \delta e^{-(2-p)a_0} \right\} u(e^{a_k}; 2 - p)$$

$$- \left\{ \tau B_k(\beta, p) + \delta e^{-(1-p)a_0} \right\} u(e^{a_k}; 1 - p) + \delta L(-\alpha_0; p)$$

$$= \delta_{Bk} L(\alpha_k - \alpha_{Bk}; p) + D_k(\beta, p, \alpha_0, \delta).$$  

141
where

\[
D_k(\beta, \tau, p, \alpha_0, \delta) = \left\{ \tau A_k(\beta, p) + \delta e^{-(2-p)\alpha_0} \right\} u(e^{\delta n_k}; 2-p) - \left\{ \tau B_k(\beta, p) + \delta e^{-(1-p)\alpha_0} \right\} u(e^{\delta n_k}; 1-p) + \delta L(-\alpha_0; p).
\]

It follows from (14), (17) and (19) that

\[
f(y_k; \alpha_k, \beta, \tau, p) \pi(\alpha_k - \alpha_0; p, \delta n_k) = \frac{f(y_k; 0, \beta, \tau, p)}{K(p, \delta n_k)} \exp\{-\delta_{Bk} n_k L(\alpha_k - \hat{\alpha}_{Bk}; p) - n_k D_k(\beta, \tau, p, \alpha_0, \delta)\}.
\]

Thus, using (18), the posterior density is calculated as

\[
\pi(\alpha_k - \hat{\alpha}_{Bk}; p, \delta_{Bk} n_k) = \frac{1}{K(p, \delta_{Bk} n_k) \pi(\alpha_0; p, \delta n_k)} \exp\{-\delta_{Bk} n_k L(\alpha_k - \hat{\alpha}_{Bk}; p)\},
\]

which is of the form of the location-dispersion family (2). This completes the proof. \(\Box\)

Other properties of the assumed prior density (17) are given in the following Lemma, which will be applied in discussing the conjugate analysis.

**Lemma 2** Set \(\xi(p, t) = 1 - (2-p)(p-1)(\partial/\partial t)\log K(p, t)\).

(i) Both the following are true:

\[
E_\pi[e^{(2-p)\alpha_k}] = \xi(p, \delta n_k) e^{(2-p)\alpha_0},
\]

\[
E_\pi[e^{(1-p)\alpha_k}] = \xi(p, \delta n_k) e^{(1-p)\alpha_0},
\]

where \(E_\pi[\cdot]\) denotes the expectation with respect to the prior density (17).

(ii) The Kullback-Leibler separator from \(\pi(\alpha_k - \alpha_{01}; p, \delta n_k)\) to \(\pi(\alpha_k - \alpha_{02}; p, \delta n_k)\) is

\[
KL(\pi(\alpha_k - \alpha_{01}; p, \delta n_k), \pi(\alpha_k - \alpha_{02}; p, \delta n_k)) = \delta n_k \xi(p, \delta n_k) L(\alpha_{01} - \alpha_{02}).
\]

**Proof**

(i) The required result is obvious when \(p = 1\) or \(p = 2\) since the density (17) is a log-transformed gamma density, so suppose \(1 < p < 2\). Differentiating both sides of the equality

\[
K(p, t) = \int_{-\infty}^{\infty} \exp[-t L(\alpha_k - \alpha_0; p)] d\alpha_k
\]

with respect to \(\alpha_0\) and \(t\) gives

\[
\int_{-\infty}^{\infty} \{e^{(2-p)(\alpha_k - \alpha_0)} - e^{(1-p)(\alpha_k - \alpha_0)}\} \exp[-t L(\alpha_k - \alpha_0; p)] d\alpha_k = 0,
\]

\[
\int_{-\infty}^{\infty} L(\alpha_k - \alpha_0; p) \exp[-t L(\alpha_k - \alpha_0; p)] d\alpha_k = -\frac{\partial}{\partial t} K(p, t).
\]

Multiplying both sides of the latter equality with \((2-p)(p-1)\) and using (18),

\[
K(p, t) - (2-p)(p-1) \frac{\partial}{\partial t} K(p, t)
\]

\[
= \int_{-\infty}^{\infty} \{(p-1)e^{(2-p)(\alpha_k - \alpha_0)} - (2-p)e^{(1-p)(\alpha_k - \alpha_0)}\} \exp[-t L(\alpha_k - \alpha_0; p)] d\alpha_k.
\]
Replace $t$ with $\delta n_k$. Then the above two equalities form a set of linear equations

$$E_\pi \left[ e^{(2-p)(\alpha_k - \alpha_0)} \right] = E_\pi \left[ e^{(1-p)(\alpha_k - \alpha_0)} \right],$$

$$(p-1)E_\pi \left[ e^{(2-p)(\alpha_k - \alpha_0)} \right] + (2-p)E_\pi \left[ e^{(1-p)(\alpha_k - \alpha_0)} \right] = \xi(p, \delta n_k).$$

The results are the solution to this set of equations.

(ii) From Lemma 1(i),

$$L(\alpha_k - \alpha_{02}; p) - L(\alpha_k - \alpha_{01}; p) = e^{(2-p)(\alpha_k - \alpha_{01})}u(e^{\alpha_{01}-\alpha_{02}}; 2-p) - e^{(1-p)(\alpha_k - \alpha_{01})}u(e^{\alpha_{01}-\alpha_{02}}; 1-p).$$

The definition of the Kullback-Leibler divergence, together with (i), yields the required result.

Now we discuss the conjugate analysis for $\alpha_k$ assuming on a temporary basis that $\beta$, $\alpha_0$ and $\delta$ are known. As stated in Sect. 4, the loss function $L(\alpha_k - \hat{\alpha}_k; p)$ is adopted. This is a Kullback-Leibler loss function, which follows from Lemma 2(ii). The conjugate analysis of $\alpha_k$ is summarized in the following proposition.

**Proposition 4** A modified Pythagorean relationship

$$E_{\text{post}} \left[ L(\alpha_k - \hat{\alpha}_k; p) - L(\alpha_k - \hat{\alpha}_{Bk}; p) - \xi(p, \delta_{Bk}n_k) L(\hat{\alpha}_{Bk} - \hat{\alpha}_k; p) \right] = 0$$

holds for any estimator $\hat{\alpha}_k$ where $E_{\text{post}}[\cdot]$ stands for the posterior expectation. Therefore, the estimator $\hat{\alpha}_{Bk}$ is optimal under the loss function $L(\alpha_k - \hat{\alpha}_k; p)$.

**Proof** Consider the Kullback-Leibler divergence from the posterior density $\pi(\alpha_k - \hat{\alpha}_{Bk}; p, \delta_{Bk}n_k)$ to another density $\pi(\alpha_k - \hat{\alpha}_k; p, \delta_{Bk}n_k)$. The latter is obtained by substituting $\alpha_k$ with an arbitrary estimator $\hat{\alpha}_k$. Note that the two densities have the same normalizing constant. It follows from (17) that

$$KL \left( \pi(\alpha_k - \hat{\alpha}_{Bk}; p, \delta_{Bk}n_k), \pi(\alpha_k - \hat{\alpha}_k; p, \delta_{Bk}n_k) \right) = \delta_{Bk}n_k E_{\text{post}} \left[ L(\alpha_k - \hat{\alpha}_k; p) - L(\alpha_k - \hat{\alpha}_{Bk}; p) \right].$$

Apply Lemma 2(ii) to the left-hand side of this equality. □

Note that $\hat{\alpha}_{Bk}$ and $\delta_{Bk}$ (in the Bayesian context) coincide with $\hat{\alpha}_M$ and $\delta_M$ (in the maximum likelihood context) respectively in Proposition 2 when $\delta$ is zero.

The family of distributions to which the conjugate prior density (17) belongs was first derived by Ohnishi & Yanagimoto [16] in the context of seeking members of the location-dispersion family having a conjugate prior. They sought location-dispersion densities $f(y - \mu)$ with a conjugate prior density of the form $\pi(\mu - m; \delta) \propto \{f(m - \mu)\}^{\delta}$. This requisition also yields the Normal and the von Mises distributions.

### 6 The optimal estimating function

We now investigate the following estimating equation of ($\beta$, $\tau$, $p$):

$$E_{\text{post}}[\nabla l(y; \alpha, \beta, \tau, p)] = 0 \quad (22)$$
where \( l(y; \alpha, \beta, \tau, p) = \log f(y; \alpha, \beta, \tau, p) \) and \( \nabla = (\partial/\partial \beta, \partial/\partial \tau, \partial/\partial p)^T \). This estimating equation is proved to be optimal under a certain Bayesian criterion. The hyperparameters \( \alpha_0 \) and \( \delta \) are assumed to be known in this Section.

The following proposition gives another expression of the estimating function (22). The proof is straightforward and is omitted.

**Proposition 5** It holds that

\[
\nabla \log f_{\text{marg}}(y; \beta, \tau, p, \alpha_0, \delta) = E_{\pi}[\nabla l(y; \alpha, \beta, \tau, p)]
\]

where \( f_{\text{marg}}(y; \beta, \tau, p, \alpha_0, \delta) \) is the marginal density.

An optimality of the estimating function (22) is shown in the following proposition. The criterion function in the proposition was adopted by Ferreira [6], which is an extended version of the one in Godambe & Kale [9] adapted to the Bayesian framework.

**Proposition 6** Suppose \( \alpha_0 \) and \( \delta \) are known, and consider an estimating function \( g(y; \beta, \tau, p) \) which is unbiased in the sense that

\[
E_{\pi}[E_{f}[g(y; \beta, \tau, p)]] = 0,
\]

where \( E_{f}[\cdot] \) denotes the expectation with respect to the sampling density. Then, the estimating function in (22) is optimal with respect to the criterion

\[
\mathcal{M}[g] = Tr \left( B^{-1}A(B^T)^{-1} \right)
\]

where \( A = E_{\pi}[E_{f}[gg^T]] \) and \( B = E_{\pi}[E_{f}[\nabla g^T]] \).

**Proof** Since \( g(y; \beta, \tau, p) \) does not depend on \( \alpha \), write the unbiasedness condition (23) as

\[
E_{\pi}[E_{f}][g(y; \beta, \tau, p)] = 0,
\]

where \( E_{\text{marg}}[\cdot] \) is the expectation with respect to the marginal density. Similarly, the matrices \( A \) and \( B \) in the criterion function (24) can be expressed as \( A = E_{\text{marg}}[gg^T] \) and \( B = E_{\text{marg}}[\nabla g^T] \). Using criterion (ii) in Godambe & Kale [9, Section 1.7], the optimal estimating function is \( \nabla \log f_{\text{marg}}(y; \beta, \tau, p, \alpha_0, \delta) \), which Proposition 5 proves to be equivalent to the estimating function in (22). \( \square \)

An interesting relationship holds between the first element of the above optimal estimating function and the optimal estimator derived in Sect. 5. The optimal estimating function of \( \beta \) is expressed in terms of the optimal estimator \( \hat{\alpha}_{Bk} \). The score function of \( \beta \) is expressed as

\[
\begin{align*}
\frac{d}{d\beta} \log f(y_k; \alpha_k, \beta, \tau, p) & = \sum \frac{d}{d\beta} l_k(y_k; \alpha_k, \beta, \tau, p) \\
& = -\tau \left\{ e^{(2-p)\alpha_k} \sum_{i=1}^{n_k} x_{ki} e^{(2-p)\beta u_{ki}} - e^{(1-p)\alpha_k} \sum_{i=1}^{n_k} x_{ki} y_{ki} e^{(1-p)\beta u_{ki}} \right\}.
\end{align*}
\]

This is shown by noting that

\[
\begin{align*}
\log f(y_k; \alpha_k, \beta, \tau, p) & = -\tau \sum_{i=1}^{n_k} \left\{ u(e^{\alpha_k+\beta u_{ki}}; 2 - p) - y_{ki} u(e^{\alpha_k+\beta u_{ki}}; 1 - p) \right\} + F_k,
\end{align*}
\]

where \( F_k \) is the term constant in \( \beta \).
Proposition 7 For any $k \in \{1, \ldots, K\}$,

$$E_{\text{post}}[l_{k\beta}(y_{k}; \alpha_{k}, \beta, \tau, p)] = \xi(p, \delta_{Bk}n_{k})l_{k\beta}(y_{k}; \hat{\alpha}_{Bk}, \beta, \tau, p).$$

Therefore, the optimum estimating function is

$$E_{\text{post}}[l_{\beta}(y; \alpha, \beta, \tau, p)] = \sum_{k=1}^{K} \xi(p, \delta_{Bk}n_{k})l_{k\beta}(y_{k}; \hat{\alpha}_{Bk}, \beta, \tau, p).$$

Proof Proposition 3 and Lemma 2 yield that

$$E_{\text{post}}[e^{(2-p)\alpha_{k}}] = \xi(p, \delta_{Bk}n_{k})e^{(2-p)\alpha_{k}},$$

$$E_{\text{post}}[e^{(1-p)\alpha_{k}}] = \xi(p, \delta_{Bk}n_{k})e^{(1-p)\alpha_{k}}.$$

This, together with (25), completes the proof. □

7 Estimation procedure

We propose to estimate $\beta$, $\tau$, $p$ and $\delta$ by maximizing the marginal likelihood function, although these parameters are assumed to be known in Sects. 4 and 5. Here set $\alpha_{0} = 0$ for simplicity.

Proposition 8 The marginal likelihood function with $\alpha_{0} = 0$ is

$$f_{\text{marg}}(y; \beta, \tau, p, 0, \delta) = f(y; 0, \beta, \tau, p) \prod_{k=1}^{K} \left[ \frac{K(p, \delta_{Bk}n_{k})}{K(p, \delta n_{k})} \exp\{-n_{k}D_{k}(\beta, \tau, p, 0, \delta)\} \right],$$

where $\delta_{Bk} = \delta_{Bk}(\beta, \tau, p, 0, \delta)$ and $D_{k}(\beta, \tau, p, \alpha_{0}, \delta)$ is defined by (20).

Proof Using the expression (21) with $\alpha_{0} = 0$ and (18),

$$\int_{-\infty}^{\infty} f(y_k; \alpha_k, \beta, \tau, p) \pi(\alpha_k; p, \delta n_{k}) d\alpha_k = f(y_k; 0, \beta, \tau, p) \frac{K(p, \delta_{Bk}n_{k})}{K(p, \delta n_{k})} \exp\{-n_{k}D_{k}(\beta, \tau, p, 0, \delta)\},$$

which completes the proof. □

Our estimation procedure consists of the following two steps:

Step 1. Maximize the marginal likelihood $f_{\text{marg}}(y; \beta, \tau, p, 0, \delta)$ with respect to $\beta$, $\tau$, $p$ and $\delta$. 

Step 2. Estimate $\alpha_{k}$ by plugging the estimates in Step 1 into $\hat{\alpha}_{Bk}(\beta, \tau, p, 0, \delta)$ in Proposition 3.

In practice, the marginal likelihood for given values of $p$ is maximized and the maximizer $\hat{\beta}$ is found through a cubic spline curve computed over a suitable range.

It is interesting to compare the Bayesian estimation procedure with the maximum likelihood (ML) procedure. It follows from Proposition 2 that the ML procedure maximizes

$$f(y; 0, \beta, \tau, p) \prod_{k=1}^{K} \exp\{-n_{k}C_{k}(\beta, \tau, p)\}$$
with respect to \((\beta, \tau, p)\). Although \(\lim_{\delta \to +0} D_h(\beta, \tau, p, 0, \delta) = C_h(\beta, \tau, p)\), a positive estimate for \(\delta\) results since \(K(p, \delta n_k)\) tends to infinity as \(\delta\) approaches to zero. Thus, our Bayesian procedure is different from the ML procedure.

An R function is available, on request, for fitting the models proposed in this paper. Evaluation of the Tweedie density function \(f(y; 0, \beta, \tau, p)\) in the marginal likelihood function given in Proposition 8 is performed using numerical algorithms [Dunn & Smyth [3, 4]] as implemented in the tweedie package [Dunn [2]]. Evaluation of the \(K(p, \cdot)\) functions, defined in Eq. (18), in the marginal likelihood, is performed using the integrate function in R, which is based on QUADPACK routines [Piesens [17]]; these programs routinely accommodate infinite limits of integration. The optimization in Step 1. is performed using constrained multivariate minimization as found in the R function nlmintb, based on the PORT routines http://netlib.bell-labs.com/netlib/port/.

### 8 Example

Forty-five stations were selected for this study (Fig. 1), all in the same climatic region as identified by the Australian Bureau of Meteorology at http://www.bom.gov.au/cgi-bin/climate/cgi-bin_scripts/clim.classification.cgi (based on Gaffney [7]). All stations are considered subtropical (Cfa in the classification of Köppen [13]).

The total July rainfall from 1970 to 2006 (37 observations per station) is used here (Table 1). A plot of the rainfall at selected stations (Fig 2) shows the extreme skewness of the distributions. Also recorded, but not shown, is the average monthly southern oscillation index, or SOI [20], for the corresponding months. The SOI is defined as difference between Tahiti and Darwin air pressures, and has been linked to Australian rainfall (Stone & Auliciems [19]).

We consider a Tweedie GLM with a logarithmic link function such that \(h(\mu_{ki}) = \log \mu_{ki} = \alpha_k + \beta x_{ki}\), where \(x_{ki} = x_i\) represents the SOI. In this example, the intercepts \(\alpha_k\) represent specific features of the observation stations and the slope \(\beta\) represents the common effect of the SOI in the region. Thus our primary interest is placed on the common slope; that is, the effect of the SOI on rainfall. The intercepts are parameters of secondary interest.

After initially using a coarse grid to determine \(p\), the final model was fitted to the data considering a finer grid of values from \(p = 1.66\) to \(p = 1.73\) in steps of 0.01. A smooth cubic spline interpolation may be fitted through these computed points for a more accurate estimate. A nominal confidence interval for \(p\) is found using that \(2 [\log L(\beta) - \log L(p_0)]\) has, asymptotically, a \(\chi^2\) distribution, where \(p_0\) is the true parameter value. The profile likelihood plot (Fig. 3) indicates an estimate of \(\hat{\beta} \approx 1.69\), with a nominal 95% confidence interval from 1.663 to 1.719 approximately. In practice, any value of \(p\) in the (say) 95% confidence interval produces very similar estimates, models and residual plots. At this empirical Bayesian estimate of \(p\), compute \(\hat{\beta} = 0.0372\), \(\hat{\tau} = 0.376\) and \(\hat{\delta} = 0.00173\).

For each candidate value of \(p\), numerical methods are used to maximize the log-likelihood over the \((\beta, \delta)\) space, then the optimum value of \(\tau\) found for each \((\beta, \delta)\) combination. The procedure iterates until convergence, and the entire process repeated for the next value of \(p\) under consideration.

In particular, the value of \(\beta\) is of interest. The profile likelihood plot for \(\beta\) (Fig. 4) shows a nominal 95% confidence interval for \(\beta\), computed similarly to that for \(p\), is from 0.032 to 0.0427 approximately. This interval certainly does not contain zero, so the effect of SOI is statistically significant, though the value is small so may not be of any practical significance. In this case, the maximum likelihood estimates are very similar to those computed using the empirical Bayesian approach (Table 2); this is expected since \(\delta\) is small.
References


Figure 1: The location of the stations in the study

Figure 2: Boxplots of total July rainfall for selected stations
Figure 3: The profile likelihood plot for $p$. The points represent the actual log-likelihoods computed at the given values of $p$. The thick solid line is a cubic spline smooth through the computed points. The dotted horizontal line represents the height of the nominal 95% confidence interval. The vertical line is the location of the empirical Bayesian estimate of $p$.

Figure 4: The profile likelihood plot for $\beta$. The thick solid line is the log-likelihood. The dotted horizontal line represents the height of the nominal 95% confidence interval. The vertical line is the location of the empirical Bayesian estimate of $p$. 
Table 1: Summaries of the total July rainfall (in millimetres) at each station

<table>
<thead>
<tr>
<th>Station</th>
<th>Min</th>
<th>Median</th>
<th>Mean</th>
<th>95% quantile</th>
<th>Max</th>
<th>Std dev</th>
<th>IQR</th>
<th>Percent zeros</th>
</tr>
</thead>
<tbody>
<tr>
<td>40014</td>
<td>1.4</td>
<td>34.0</td>
<td>36.8</td>
<td>96.8</td>
<td>138.8</td>
<td>31.2</td>
<td>35.2</td>
<td>0.0</td>
</tr>
<tr>
<td>40024</td>
<td>0.0</td>
<td>29.8</td>
<td>35.7</td>
<td>99.9</td>
<td>159.6</td>
<td>35.0</td>
<td>39.8</td>
<td>8.0</td>
</tr>
<tr>
<td>40047</td>
<td>0.1</td>
<td>51.0</td>
<td>72.4</td>
<td>178.6</td>
<td>520.2</td>
<td>90.4</td>
<td>54.6</td>
<td>0.0</td>
</tr>
<tr>
<td>40075</td>
<td>0.0</td>
<td>30.7</td>
<td>37.7</td>
<td>118.3</td>
<td>226.6</td>
<td>42.8</td>
<td>23.9</td>
<td>5.0</td>
</tr>
<tr>
<td>40079</td>
<td>0.0</td>
<td>0.0</td>
<td>27.7</td>
<td>94.8</td>
<td>258.0</td>
<td>45.7</td>
<td>25.3</td>
<td>8.0</td>
</tr>
<tr>
<td>40082</td>
<td>0.0</td>
<td>0.8</td>
<td>29.9</td>
<td>89.6</td>
<td>306.4</td>
<td>52.4</td>
<td>28.0</td>
<td>3.0</td>
</tr>
<tr>
<td>40083</td>
<td>0.0</td>
<td>0.9</td>
<td>28.2</td>
<td>93.4</td>
<td>224.2</td>
<td>41.9</td>
<td>29.7</td>
<td>3.0</td>
</tr>
<tr>
<td>40094</td>
<td>0.8</td>
<td>2.1</td>
<td>34.5</td>
<td>35.2</td>
<td>90.2</td>
<td>150.8</td>
<td>29.7</td>
<td>27.8</td>
</tr>
<tr>
<td>40096</td>
<td>0.0</td>
<td>0.3</td>
<td>26.2</td>
<td>107.4</td>
<td>142.5</td>
<td>33.7</td>
<td>35.1</td>
<td>3.0</td>
</tr>
<tr>
<td>40104</td>
<td>0.0</td>
<td>0.0</td>
<td>34.8</td>
<td>107.9</td>
<td>137.7</td>
<td>33.5</td>
<td>32.4</td>
<td>8.0</td>
</tr>
<tr>
<td>40110</td>
<td>0.0</td>
<td>1.7</td>
<td>23.1</td>
<td>43.6</td>
<td>112.9</td>
<td>365.4</td>
<td>64.7</td>
<td>34.3</td>
</tr>
<tr>
<td>40117</td>
<td>1.0</td>
<td>4.0</td>
<td>55.4</td>
<td>84.9</td>
<td>228.4</td>
<td>879.9</td>
<td>145.5</td>
<td>42.7</td>
</tr>
<tr>
<td>40120</td>
<td>0.8</td>
<td>3.1</td>
<td>27.8</td>
<td>36.1</td>
<td>101.0</td>
<td>243.7</td>
<td>43.3</td>
<td>24.0</td>
</tr>
<tr>
<td>40124</td>
<td>0.0</td>
<td>0.0</td>
<td>24.5</td>
<td>34.5</td>
<td>114.2</td>
<td>204.8</td>
<td>40.3</td>
<td>29.1</td>
</tr>
<tr>
<td>40157</td>
<td>0.0</td>
<td>1.1</td>
<td>53.4</td>
<td>90.9</td>
<td>245.2</td>
<td>835.4</td>
<td>142.6</td>
<td>54.8</td>
</tr>
<tr>
<td>40158</td>
<td>0.0</td>
<td>2.1</td>
<td>32.2</td>
<td>42.8</td>
<td>134.1</td>
<td>273.3</td>
<td>50.4</td>
<td>25.7</td>
</tr>
<tr>
<td>40166</td>
<td>1.2</td>
<td>5.5</td>
<td>50.5</td>
<td>66.8</td>
<td>178.5</td>
<td>367.3</td>
<td>67.8</td>
<td>61.8</td>
</tr>
<tr>
<td>40171</td>
<td>0.0</td>
<td>2.3</td>
<td>32.3</td>
<td>53.9</td>
<td>137.6</td>
<td>481.7</td>
<td>80.7</td>
<td>32.6</td>
</tr>
<tr>
<td>40184</td>
<td>0.8</td>
<td>1.4</td>
<td>29.0</td>
<td>37.9</td>
<td>104.8</td>
<td>228.1</td>
<td>42.0</td>
<td>38.5</td>
</tr>
<tr>
<td>40190</td>
<td>1.0</td>
<td>3.5</td>
<td>58.8</td>
<td>72.7</td>
<td>153.9</td>
<td>406.2</td>
<td>71.1</td>
<td>62.0</td>
</tr>
<tr>
<td>40196</td>
<td>0.7</td>
<td>3.8</td>
<td>54.2</td>
<td>71.8</td>
<td>178.3</td>
<td>363.4</td>
<td>71.2</td>
<td>69.0</td>
</tr>
<tr>
<td>40197</td>
<td>3.8</td>
<td>4.2</td>
<td>54.0</td>
<td>69.4</td>
<td>184.7</td>
<td>427.3</td>
<td>79.3</td>
<td>44.4</td>
</tr>
<tr>
<td>40224</td>
<td>0.1</td>
<td>2.6</td>
<td>32.4</td>
<td>51.9</td>
<td>149.2</td>
<td>404.0</td>
<td>72.0</td>
<td>27.8</td>
</tr>
<tr>
<td>40229</td>
<td>0.2</td>
<td>1.4</td>
<td>28.0</td>
<td>48.6</td>
<td>162.2</td>
<td>329.6</td>
<td>64.3</td>
<td>28.6</td>
</tr>
<tr>
<td>40231</td>
<td>0.0</td>
<td>3.5</td>
<td>36.1</td>
<td>54.9</td>
<td>153.5</td>
<td>435.6</td>
<td>77.0</td>
<td>44.0</td>
</tr>
<tr>
<td>40237</td>
<td>2.0</td>
<td>6.0</td>
<td>30.0</td>
<td>54.1</td>
<td>161.8</td>
<td>430.2</td>
<td>76.2</td>
<td>28.2</td>
</tr>
<tr>
<td>40242</td>
<td>0.4</td>
<td>3.4</td>
<td>31.5</td>
<td>51.0</td>
<td>142.9</td>
<td>423.0</td>
<td>71.9</td>
<td>27.7</td>
</tr>
<tr>
<td>40244</td>
<td>0.2</td>
<td>4.9</td>
<td>31.2</td>
<td>52.5</td>
<td>159.4</td>
<td>378.8</td>
<td>68.8</td>
<td>27.3</td>
</tr>
<tr>
<td>40245</td>
<td>0.2</td>
<td>3.8</td>
<td>29.8</td>
<td>48.5</td>
<td>154.5</td>
<td>318.8</td>
<td>60.8</td>
<td>28.4</td>
</tr>
<tr>
<td>40257</td>
<td>1.2</td>
<td>5.5</td>
<td>58.9</td>
<td>88.7</td>
<td>234.5</td>
<td>867.1</td>
<td>145.6</td>
<td>51.6</td>
</tr>
<tr>
<td>40382</td>
<td>0.0</td>
<td>1.8</td>
<td>40.2</td>
<td>42.1</td>
<td>109.0</td>
<td>212.1</td>
<td>41.1</td>
<td>34.2</td>
</tr>
<tr>
<td>41001</td>
<td>0.0</td>
<td>0.1</td>
<td>30.4</td>
<td>38.8</td>
<td>107.0</td>
<td>122.8</td>
<td>33.6</td>
<td>42.4</td>
</tr>
<tr>
<td>41011</td>
<td>0.0</td>
<td>0.4</td>
<td>30.8</td>
<td>31.9</td>
<td>75.7</td>
<td>128.5</td>
<td>28.3</td>
<td>30.4</td>
</tr>
<tr>
<td>41018</td>
<td>0.4</td>
<td>0.8</td>
<td>31.0</td>
<td>34.9</td>
<td>87.4</td>
<td>115.0</td>
<td>28.6</td>
<td>37.8</td>
</tr>
<tr>
<td>41022</td>
<td>2.3</td>
<td>4.3</td>
<td>45.2</td>
<td>49.3</td>
<td>141.3</td>
<td>158.3</td>
<td>40.2</td>
<td>48.6</td>
</tr>
<tr>
<td>41053</td>
<td>0.0</td>
<td>0.0</td>
<td>21.1</td>
<td>31.5</td>
<td>83.3</td>
<td>138.4</td>
<td>29.8</td>
<td>35.8</td>
</tr>
<tr>
<td>41056</td>
<td>1.6</td>
<td>2.6</td>
<td>31.6</td>
<td>39.2</td>
<td>104.3</td>
<td>184.3</td>
<td>38.9</td>
<td>41.8</td>
</tr>
<tr>
<td>41063</td>
<td>0.0</td>
<td>0.5</td>
<td>30.8</td>
<td>37.7</td>
<td>93.0</td>
<td>131.4</td>
<td>31.6</td>
<td>39.9</td>
</tr>
<tr>
<td>41072</td>
<td>0.0</td>
<td>0.0</td>
<td>25.0</td>
<td>33.3</td>
<td>76.1</td>
<td>130.9</td>
<td>29.2</td>
<td>45.3</td>
</tr>
<tr>
<td>41082</td>
<td>0.0</td>
<td>1.2</td>
<td>31.2</td>
<td>36.1</td>
<td>98.6</td>
<td>137.4</td>
<td>32.6</td>
<td>36.2</td>
</tr>
<tr>
<td>41083</td>
<td>0.0</td>
<td>0.6</td>
<td>31.0</td>
<td>40.1</td>
<td>104.2</td>
<td>124.9</td>
<td>33.6</td>
<td>47.6</td>
</tr>
<tr>
<td>41095</td>
<td>1.3</td>
<td>1.8</td>
<td>39.7</td>
<td>45.2</td>
<td>128.3</td>
<td>143.8</td>
<td>36.4</td>
<td>39.7</td>
</tr>
<tr>
<td>41103</td>
<td>0.2</td>
<td>1.3</td>
<td>39.0</td>
<td>45.7</td>
<td>121.7</td>
<td>158.7</td>
<td>39.3</td>
<td>53.2</td>
</tr>
<tr>
<td>41118</td>
<td>2.0</td>
<td>3.1</td>
<td>34.4</td>
<td>45.0</td>
<td>112.9</td>
<td>144.1</td>
<td>34.4</td>
<td>43.4</td>
</tr>
<tr>
<td>41126</td>
<td>0.0</td>
<td>0.0</td>
<td>30.2</td>
<td>33.2</td>
<td>78.9</td>
<td>122.2</td>
<td>28.7</td>
<td>37.1</td>
</tr>
</tbody>
</table>
Table 2: Empirical Bayesian estimates and the maximum likelihood estimates for fitting the model to July rainfall

<table>
<thead>
<tr>
<th>Parameter</th>
<th>Empirical Bayesian estimate</th>
<th>Maximum likelihood estimate</th>
</tr>
</thead>
<tbody>
<tr>
<td>$p$</td>
<td>1.69</td>
<td>1.68</td>
</tr>
<tr>
<td>$\beta$</td>
<td>0.0372</td>
<td>0.0371</td>
</tr>
<tr>
<td>$\tau$</td>
<td>0.376</td>
<td>0.369</td>
</tr>
<tr>
<td>$\delta$</td>
<td>0.00173</td>
<td></td>
</tr>
</tbody>
</table>