Statistica Sinica: Supplement

# SINGULAR PRIOR DISTRIBUTIONS AND ILL-CONDITIONING IN BAYESIAN $D$-OPTIMAL DESIGN FOR SEVERAL NONLINEAR MODELS

Timothy W. Waite

*University of Manchester*

**Supplementary Material**

## S1 Further discussion of alternative priors

In Section 4.1, it is stated that 'one must avoid selecting prior distributions for analytical convenience if they do not accurately represent the available expert belief or knowledge'. Here we provide further discussion of this point.

As an example, in Section 2.1 we found that $\mathcal{P} : \theta \sim U(0, a)$ is singular for the exponential regression model. A natural question is whether it is sufficient to find designs for the non-singular prior $\mathcal{P}_\epsilon : \theta \sim U(\epsilon, a)$ for some small value of $\epsilon$ (e.g. $10^{-3}$ or $10^{-6}$). The adequacy of $\mathcal{P}_\epsilon$ as a representation of the expert's beliefs will depend substantially on the specifics of the application. For small $\epsilon$, the quartiles of $\mathcal{P}$ and $\mathcal{P}_\epsilon$ are similar, thus for example it is possible for both distributions to fit expert statements obtained by the bisection method (Garthwaite et al. (2005)). However, the implication of $\mathcal{P}_\epsilon$ that there is zero probability that $\theta < \epsilon$ is too strong unless the expert is certain that $\theta \geq \epsilon$. The fidelity of the representation $\mathcal{P}_\epsilon$ would be less important if the resulting design decision were insensitive to the choice of $\epsilon$. Unfortunately this is not the case, as shown by the proposition below and its proof. Intuitively, as $\epsilon \to 0$, some points in the Bayesian $D$-optimal design for $\mathcal{P}_\epsilon$ will converge to zero (while never being equal to zero).

**Proposition 1.** *For the exponential model, if $\xi$ does not vary with $\epsilon$ then*

$$\text{Bayes-eff}(\xi; \mathcal{P}_\epsilon) \to 0 \quad \textit{as} \quad \epsilon \to 0\,.$$

Thus, even if one were to compute the Bayesian $D$-optimal design for $\mathcal{P}_{\epsilon'}$, with say $\epsilon' = 10^{-6}$, the resulting design would be highly inefficient when evaluated under $\mathcal{P}_\epsilon$ for $\epsilon \ll \epsilon'$.

The situation above is somewhat similar to problems in the objective Bayesian approach with improper uninformative priors (e.g. Berger (1985, Ch.3); Berger (2006)), which one may need to modify in order to obtain a proper posterior. For example, if an improper prior, say $U(10, \infty)$, does not give a proper posterior, one might attempt to replace it with $U(10, M)$, with $M$ large, e.g. $10^5$ or $10^6$. However, the results would often be highly sensitive to the value chosen for $M$, which is arbitrary and typically has no objective justification.

## S2 Proofs of analytical results

*Proof of Proposition 2.* Assume that at least one $x_i > 0$. For the $\theta$ parameterization, we demonstrate two implications: (i) if $E_{\mathcal{P}}(1/\theta) < \infty$ and $E_{\mathcal{P}}(\log\theta) < \infty$, then $\phi(\xi;\mathcal{P}) > -\infty$; and (ii) if $E_{\mathcal{P}}(\log\theta) = \infty$ or $E_{\mathcal{P}}(1/\theta) = \infty$, then $\phi(\xi;\mathcal{P}) = -\infty$. Here, $\phi(\xi;\mathcal{P}) = E\{\log|M_\theta(\xi;\theta)|\}$, where $\log|M_\theta(\xi;\theta)|$ is given by (2.2).

For (i), observe that $-\infty \leq E_{\mathcal{P}}\left\{(2/\theta)\max_{i=1,\ldots,n}\{x_i\} + 4\log\theta\right\} < \infty$. Considering the left hand side of (2.1) and the reparameterization (2.2),

$$-\infty < \log\sum_{i=1}^{n} x_i^2 - E_{\mathcal{P}}\left\{(2/\theta)\max_{i=1,\ldots,n}\{x_i\} + 4\log\theta\right\} \leq \phi(\xi;\mathcal{P})\,,$$

as required. For (ii), note that in addition to (2.1), the following weaker inequality holds:

$$\phi(\xi;\mathcal{P}) \leq \log\sum_{i=1}^{n} x_i^2 - 4\log\theta\,.$$

Taking expectations of both sides, if $E_{\mathcal{P}}(\log\theta) = \infty$ then $\phi(\xi;\mathcal{P}) = -\infty$.

For the other case, let

$$b(\theta) = \frac{1}{\theta}\left\{2\min_{i=1,\ldots,n}\{x_i : x_i > 0\} + 4\theta\log\theta\right\}\,.$$

Since $\theta\log\theta \to 0$ as $\theta \to 0$, there is some $\delta > 0$ such that, for $\theta < \delta$,

$$b(\theta) \geq (1/\theta)\min_{i=1,\ldots,n}\{x_i : x_i > 0\}\,.$$

Hence, with $\mathbb{I}$ denoting an indicator function,

$$\begin{aligned} E_{\mathcal{P}}\{b(\theta)\} &\geq E_{\mathcal{P}}\{b(\theta)\mathbb{I}(\theta < \delta) + \inf_{\theta \geq \delta} b(\theta)\,\mathbb{I}(\theta \geq \delta)\} \\ &\geq \min_{i=1,\ldots,n}\{x_i : x_i > 0\} E_{\mathcal{P}}\{(1/\theta)\mathbb{I}(\theta < \delta)\} + (4\log\delta)\Pr(\theta \geq \delta) \end{aligned} \tag{S2.1}$$

If $E_{\mathcal{P}}(1/\theta) = \infty$, then $E_{\mathcal{P}}\{(1/\theta)\mathbb{I}(\theta < \delta)\} = \infty$, and so by (S2.1), we have that $E_{\mathcal{P}}\{b(\theta)\} = \infty$, regardless of whether $E_{\mathcal{P}}(\log\theta) = -\infty$. Recall from (2.1) that

$$\phi(\xi; \mathcal{P}) \leq \log \sum_{i=1}^{n} x_i^2 - E_{\mathcal{P}}\{b(\theta)\}\,.$$

Hence if $E_{\mathcal{P}}(1/\theta) = \infty$, we have $\phi(\xi; \mathcal{P}) = -\infty$. This is sufficient to establish the proposition. □

*Proof of Lemma 1.* Observe that $M(x_i; \theta) = e^{-2\theta_1 x_i}\tilde{M}^{(i)}_{\delta,\theta_3}$, where $\tilde{M}^{(i)}_{\delta,\theta_3}$ is defined in the statement of the lemma. Moreover, for $i = 1, \ldots, n$, either (i) $x_i = 0$ or (ii) $x_i \geq x_{\min}$. In (ii), we have

$$e^{-2\theta_1 x_{\max}}\tilde{M}^{(i)}_{\delta,\theta_3} \preceq M(x_i; \theta) \preceq e^{-2\theta_1 x_{\min}}\tilde{M}^{(i)}_{\delta,\theta_3}\,. \tag{S2.2}$$

Moreover, the above holds also in (i) since then $M(x_i; \theta)$ and $M^{(i)}_{\delta,\theta_3}$ are matrices of zeroes. Summing (S2.2) over $i = 1, \ldots, n$, we obtain:

$$e^{-2\theta_1 x_{\max}}\tilde{M}_{\delta,\theta_3} \preceq M(\xi; \theta) \preceq e^{-2\theta_1 x_{\min}}\tilde{M}_{\delta,\theta_3}\,. \tag{S2.3}$$

Taking log-determinants throughout (S2.3) yields the result, when combined with the fact that $|\tilde{M}_{\delta,\theta_3}| = \theta_3^4|\tilde{M}_{\delta,1}|$. □

Define $g_\xi(\delta) = |\tilde{M}_{\delta,1}|$. The following is needed to establish Lemma 2.

**Lemma 5.** *Suppose that $\xi$ contains at least three distinct $x_i > 0$. Then the derivatives of $g_\xi(\delta)$ satisfy: (i) $g_\xi^{(k)}(0) = 0$, $k = 1, \ldots, 7$, (ii) $g_\xi^{(8)}(0) > 0$.*

*Proof of Lemma 5.* Part (i) can be verified using symbolic computation, e.g. Mathematica. It can also be shown that

$$g_\xi^{(8)}(0) = 280\{S_2S_4S_6 - S_2S_5^2 - S_3^2S_6 + S_3S_4S_5 + S_3S_4S_5 - S_4^3\}\,,$$

where $S_l = \sum_{i=1}^n x_i^l$. Define the following,

$$K = \begin{pmatrix} S_2 & S_3 & S_4 \\ S_3 & S_4 & S_5 \\ S_4 & S_5 & S_6 \end{pmatrix}, \qquad K' = \sum_{i:x_i>0} \begin{pmatrix} 1 & x_i & x_i^2 \\ x_i & x_i^2 & x_i^3 \\ x_i^2 & x_i^3 & x_i^4 \end{pmatrix},$$

and $x_{\min} = \min\{x_i : x_i > 0\}$. Note that $K \succeq x_{\min}^2 K'$. We have

$$g_\xi^{(8)}(0) = 280|K| \geq 280 x_{\min}^6 |K'|\,.$$

Observe also that $K'$ is the information matrix of the design $\xi' = (x_i : x_i > 0)$ under the linear model with regressors $1, x, x^2$. By the assumption that there are at least three distinct $x_i > 0$, the above linear model is estimable and so $|K'| > 0$. This establishes part (ii). $\square$

*Proof of Lemma 2.* If $\xi$ has fewer than three distinct $x_i > 0$, then we have $\text{rank}(\tilde{M}_{\delta,1}) \leq 2$ and $E_{\mathcal{P}}(\log|\tilde{M}_{\delta,1}|) = -\infty$ for any prior $\mathcal{P}$. Thus we may assume that $\xi$ has at least three distinct $x_i > 0$. From Lemma 5, it is clear that $g_\xi(\delta) \approx (\kappa/2)\delta^8$ for small $\delta$, where $\kappa > 0$. We show that the approximation is sufficiently close that $E_{\mathcal{P}}(\log|\tilde{M}_{\delta,1}|) = -\infty$ if $\int_{\delta<1} \log\delta \, d\mathcal{P}(\theta) = -\infty$. By Taylor's theorem, there is an $\epsilon_1 > 0$ and $\lambda > 0$ such that, for $\delta \in (0, \epsilon_1)$,

$$|g_\xi(\delta) - (\kappa/2)\delta^8| \leq \lambda\delta^9 \,.$$

Hence, for $\delta \in (0, \epsilon_1)$,

$$|2g_\xi(\delta)/(\delta^8\kappa) - 1| \leq (2\lambda/\kappa)\delta \,.$$

As the logarithm function has derivative 1 at argument 1, there exists $0 < \epsilon_2 \leq \epsilon_1$ such that for $\delta \in (0, \epsilon_2)$,

$$\left|\log\frac{2g_\xi(\delta)}{\delta^8\kappa} - \log 1\right| \leq 2|2g_\xi(\delta)/(\delta^8\kappa) - 1| \leq (4\lambda/\kappa)\delta \,.$$

Thus, for $\delta \in (0, \epsilon_2)$,

$$|\log g_\xi(\delta) - \log(\kappa\delta^8/2)| \leq (4\lambda/\kappa)\delta \,,$$

so that

$$\left|\int_{\delta<\epsilon_2} \log g_\xi(\delta) d\mathcal{P}(\theta) - \int_{\delta<\epsilon_2} \{8\log\delta + \log(\kappa/2)\} d\mathcal{P}(\theta)\right| \leq (2\lambda/\kappa)\epsilon_2^2 \,.$$

Hence it is clear that $\int_{\delta<\epsilon_2} \log g_\xi(\delta) d\mathcal{P}(\theta) = -\infty$ if and only if $\int_{\delta<\epsilon_2} \log\delta \, d\mathcal{P}(\theta) = -\infty$. Further, $g_\xi(\delta)$ is bounded above, and

$$\int \log g_\xi(\delta) d\mathcal{P}(\theta) = \int_{\delta<\epsilon_2} \log g_\xi(\delta) d\mathcal{P}(\theta) + \int_{\delta>\epsilon_2} \log g_\xi(\delta) d\mathcal{P}(\theta) \,.$$

Thus, $\int \log g_\xi(\delta) d\mathcal{P}(\theta) = -\infty$ when $\int_{\delta<\epsilon_2} \log \delta \, d\mathcal{P}(\theta) = -\infty$. The result is finally established by observing that $\int_{\delta<\epsilon_2} \log \delta \, d\mathcal{P}(\theta) = -\infty$ if we have $\int_{\delta<1} \log \delta \, d\mathcal{P}(\theta) = -\infty$.

□

*Proof of Proposition 3.* From Lemma 1,

$$\log |M(\xi;\theta)| \leq -6\theta_1 x_{\min} + 4\log\theta_3 + \log|\tilde{M}_{\delta,1}| \,. \tag{S2.4}$$

It can be shown that $|\tilde{M}_{\delta,1}| \leq 2S_0S_2^2 + 4S_2S_1^2$, thus $\int \log|\tilde{M}_{\delta,1}| \, d\mathcal{P}(\theta) < \infty$. As $\theta_1 > 0$, $\int -6\theta_1 x_{\min} d\mathcal{P}(\theta) \leq 0 < \infty$. If $\int_{\theta_3>1} \log\theta_3 \, d\mathcal{P}(\theta) < \infty$, as assumed by the lemma, then all terms on the right hand side of (S2.4) have integral $< \infty$ and

$$\int \log|M(\xi;\theta)| \, d\mathcal{P}(\theta) \leq \int -6\theta_1 x_{\min} \, d\mathcal{P}(\theta) + 4\int \log\theta_3 \, d\mathcal{P}(\theta) + \int \log|\tilde{M}_{\delta,1}| \, d\mathcal{P}(\theta) \,.$$

Hence if, in addition to $\int_{\theta_3>1} \log\theta_3 \, d\mathcal{P}(\theta) < \infty$, we have that at least one of $\int \log|\tilde{M}_{\delta,1}| \, d\mathcal{P}(\theta) = -\infty$, $\int -6\theta_1 x_{\min} \, d\mathcal{P}(\theta) = -\infty$, or $\int_{\theta_3<1} \log\theta_3 \, d\mathcal{P}(\theta) = -\infty$ holds, then also $\int \log|M(\xi;\theta)| \, d\mathcal{P}(\theta) = -\infty$. Using Lemma 2, the condition $\int \log|\tilde{M}_{\delta,1}| \, d\mathcal{P}(\theta) = -\infty$ in the preceding statement may be replaced by $\int_{\delta<1} \log\delta \, d\mathcal{P}(\theta) = -\infty$. This establishes the result. □

*Proof of Theorem 1.* It follows from Lemma 3 that

$$\log|M(\xi;\beta)| \geq \log|F^{\mathrm{T}}F| + p\min_i \log w_i\,.$$

From (2.6), $w(\eta) \geq (1/4)e^{-|\eta|}$. Thus,

$$\begin{aligned}\log|M(\xi;\beta)| &\geq \log|F^{\mathrm{T}}F| + p\log\left[(1/4)e^{-\max_i|\eta_i|}\right]\\ &\geq \log|F^{\mathrm{T}}F| - p\max_i|\eta_i| - p\log 4\,.\end{aligned}$$

Moreover, by the triangle inequality, $\max_i|\eta_i| \leq \sum_j \max_i|f_j(x_i)||\beta_j|$, and hence

$$\log|M(\xi;\beta)| \geq \log|F^{\mathrm{T}}F| - p\log 4 - p\sum_j \max_i|f_j(x_i)||\beta_j|\,. \tag{S2.5}$$

The right hand side of (S2.5) has expectation greater than $-\infty$ due to the assumptions that $E_{\mathcal{P}}(|\beta_j|) < \infty$ and $|F^{\mathrm{T}}F| > 0$. Therefore we have that $E_{\mathcal{P}}\{\log|M(\xi;\beta)|\} > -\infty$. □

*Proof of Proposition 4.* From Lemma 3,

$$\log|M(\xi;\beta)| \leq \log|F^{\mathrm{T}}F| + p\max_i \log w_i\,.$$

It can also be shown that $w(\eta)$ is a decreasing function of $|\eta|$ and, from (2.6), that $w(|\eta|) \leq \exp(-|\eta|)$. Hence,

$$\begin{aligned}\log|M(\xi;\beta)| &\leq \log|F^{\mathrm{T}}F| + p\log w(\min_i|\eta_i|)\\ &\leq \log|F^{\mathrm{T}}F| - p\min_i|\eta_i|\,.\end{aligned}$$

It remains to prove $E_{\mathcal{P}}(\min_i |\eta_i|) = \infty$ to establish that $E_{\mathcal{P}}\{\log |M(\xi;\beta)|\} = -\infty$. This is achieved by conditioning on an event where the parameter $\beta_j$ dominates. Given $j \in \{0, \ldots, p-1\}$, let $\mathcal{E} \in \Sigma$ be an event such that (a) $\beta_j > 1$, and (b) $\sum_{k \neq j} |f_k(x_i)||\beta_k| < \epsilon$ for all $i$, where $\epsilon > 0$ is such that

$$||f_j(x_i)| - |f_j(x_{i'})|| > 2\epsilon \qquad \text{for any } i, i' \text{ with } |f_j(x_i)| \neq |f_j(x_{i'})| \,.$$

We can guarantee (a) and (b), for example by taking

$$\mathcal{E} = \{\beta : \beta_j > 1,\ |\beta_k| < \delta\,,\ \text{for } k \neq j\} \in \Sigma\,, \tag{S2.6}$$

with $\delta = \epsilon / [(p-1) \max_{i,l} |f_l(x_i)|]$. The above satisfies

$$\Pr(\mathcal{E}) = \Pr(\beta_j > 1) \Pr(|\beta_k| < \delta \text{ for all } k \neq j \mid \beta_j > 1) > 0\,,$$

by assumptions (i) and (ii) of the proposition.

By the reverse triangle inequality and from (b), on event $\mathcal{E}$,

$$||\eta_i| - |f_j(x_i)|\beta_j| \leq \sum_{k \neq j} |f_k(x_i)||\beta_k| \leq \epsilon\,. \tag{S2.7}$$

Since on $\mathcal{E}$ the term from $\beta_j$ dominates, the minimum of $|\eta_i|$ is found by minimizing the $\beta_j$ term. To see this formally, observe that if $|f_j(x_i)|\beta_j > |f_j(x_{i'})|\beta_j$, then by the definition of $\epsilon$,

$$|f_j(x_i)|\beta_j - |f_j(x_{i'})|\beta_j > 2\epsilon\beta_j > 2\epsilon\,.$$

and, by also using (S2.7),

$$|\eta_{i'}| < |f_j(x_{i'})|\beta_j + \epsilon < |f_j(x_i)|\beta_j - \epsilon < |\eta_i|\,.$$

Thus, on $\mathcal{E}$, if $|f_j(x_i)|\beta_j > |f_j(x_{i'})|\beta_j$ then $|\eta_i| > |\eta_{i'}|$. This can be used to show that on $\mathcal{E}$, if $i^* \in \arg\min_i |\eta_i|$ then $i^* \in \arg\min_i |f_j(x_i)|$, as follows. Suppose that $i^* \notin \arg\min_i |f_j(x_i)|$, then there would exist some $i$ such that $|f_j(x_{i^*})|\beta_j > |f_j(x_i)|\beta_j$. By the above, we would have that $|\eta_{i^*}| > |\eta_i|$, which contradicts the definition of $i^*$ as a member of $\arg\min_i |\eta_i|$. Hence,

$$\begin{aligned}\min_i |\eta_i| = |\eta_{i^*}|\,, \quad & i^* \in \arg\min_i |f_j(x_i)| \\ & \geq |f_j(x_{i^*})|\beta_j - \epsilon\,.\end{aligned}$$

Consequently,

$$\begin{aligned}E_{\mathcal{P}}(\min_i |\eta_i| \mid \mathcal{E}) &\geq |f_j(x_{i^*})|\, E_{\mathcal{P}}(\beta_j \mid \mathcal{E}) - \epsilon \\ &= \infty \quad \text{by assumptions (iii) and (iv) of the proposition.}\end{aligned}$$

For the marginal expectation, note that $\Pr(\mathcal{E}) > 0$, and hence

$$E_{\mathcal{P}}(\min_i |\eta_i|) \geq \Pr(\mathcal{E}) E_{\mathcal{P}}(\min_i |\eta_i| \mid \mathcal{E}) = \infty\,.$$

□

*Proof of Proposition 5.* Case (i): assume that $f_j(x_i) > 0$. On the event $\mathcal{E}$ defined in the previous proof, we have that $\eta_i \geq f_j(x_i)\beta_j - \epsilon$. Hence $E(\eta_i \mid \mathcal{E}) = \infty$. Let $P = \Pr(y_i = 1 \mid \beta)$, noting that $1 - \Pr(y_i = 1 \mid \beta) \leq e^{-\eta_i}$. Then,

$$E^{\mathcal{G}}(1 - P \mid \mathcal{E}) = \exp E \log\{1 - P \mid \mathcal{E}\} \leq \exp E(-\eta_i \mid \mathcal{E}) = 0\,,$$

where $E^{\mathcal{G}}$ denotes the geometric mean. Hence the conditional geometric mean of $1-P$ is zero.

Case (ii): assume $f_j(x_i) < 0$. On $\mathcal{E}$, $\eta_i \le f_j(x_i)\beta_j + \epsilon$ and so $E(\eta_i \,|\, \mathcal{E}) = -\infty$. However, $P \le e^{\eta_i}$ and so $E^{\mathcal{G}}(P \,|\, \mathcal{E}) \le \exp E(\eta_i \,|\, \mathcal{E}) = 0$. □

*Proof of Theorem 2.* Assume $\xi$ is such that $|F^{\mathrm{T}}F| > 0$. Note that $w(\eta)$ is decreasing in $|\eta|$ and so, by Lemma 3,

$$\log|M(\xi;\beta)| \ge \log|F^{\mathrm{T}}F| + p\log w(\max_i |\eta_i|)\,.$$

We split $E\log|M(\xi;\beta)|$ into two components,

$$\begin{aligned} E\log|M(\xi;\beta)| &= E\left[\,\log|M(\xi;\beta)|\,\mathbb{I}(\max|\eta_i| \le \kappa)\,\right] \\ &\quad + E\left[\,\log|M(\xi;\beta)|\,\mathbb{I}(\max|\eta_i| > \kappa)\,\right]\,, \end{aligned} \tag{S2.8}$$

where $\kappa > 0$, and then show that both components are $> -\infty$.

Note that for $|\eta| \le \kappa$, $w(\eta)$ is bounded below by a constant, $\lambda > 0$. Thus, if $\max_i |\eta_i| \le \kappa$, then $\log|M(\xi;\beta)| \ge \log|F^{\mathrm{T}}F| + p\log\lambda$ and so

$$E\left[\log|M(\xi;\beta)|\,\mathbb{I}(\max_i |\eta_i| \le \kappa)\right] > -\infty\,. \tag{S2.9}$$

For $|\eta| > \kappa$, with $\kappa$ sufficiently large, by the asymptotic approximation (2.8),

$$w(\eta) \ge L|\eta|e^{-\eta^2/2}\,,$$

for some $L > 0$. Hence if $\max_i |\eta_i| > \kappa$, then

$$\begin{aligned}\log|M(\xi;\beta)| &\geq \log|F^{\mathrm{T}}F| + p\log L + p\log\max_i|\eta_i| - p\max_i \eta_i^2/2 \\ &\geq \log|F^{\mathrm{T}}F| + p\log L + p\log\kappa - p\max_i \eta_i^2/2\,.\end{aligned}$$

However, it is straightforward to show that if $E\beta_k^2 < \infty$ and $E|\beta_k\beta_l| < \infty$ for $k, l = 0, \ldots, p-1$, then $E\max_i \eta_i^2 < \infty$. This is sufficient to prove that

$$E\left[\log|M(\xi;\beta)|\,\mathbb{I}(\max_i|\eta_i| > \kappa)\right] > -\infty\,. \tag{S2.10}$$

Combining (S2.8), (S2.9) and (S2.10), we find that overall $E\log|M(\xi;\beta)| > -\infty$, and so $\mathcal{P}$ is non-singular.

□

**Lemma 6.** *Let $X$ be a random variable taking values in $A \subseteq \mathbb{R}$, with $A$ unbounded above, and let $s, t : A \to \bar{\mathbb{R}}$ be measurable extended real-valued functions that satisfy (i) for all $k \in \mathbb{R}$, $\sup_{\{x\in A\,|\,x\leq k\}} |s(x)| < \infty$ (ii) $t(x)$ is increasing, and (iii) $r(x) = t(x)/s(x) \to 0$ as $x \to \infty$. Given the above, if $E[s(X)] = \infty$ then $E[s(X) - t(X)] = \infty$.*

*Proof.* Note that from (iii) there exists $k' > 0$ such that when $X > k'$, we have $s(X) - t(X) \geq (1/2)s(X)$. When $X \leq k'$, by (ii) we have $t(X) \leq t(k')$.

Hence,

$$E[s(X) - t(X)] = E\{[s(X) - t(X)]\,\mathbb{I}(X \le k') + [s(X) - t(X)]\,\mathbb{I}(X > k')\}$$
$$\ge E\{[s(X) - t(k')]\,\mathbb{I}(X \le k') + (1/2)s(X)\,\mathbb{I}(X > k')\}\,.$$

By condition (i), $s(x)$ is bounded on $\{x \le k'\}$. Therefore the first term inside the expectation above is also bounded, and since $E\{s(X)\} = \infty$ we must have that $E\{s(X)\mathbb{I}(X > k')\} = \infty$. Hence the right hand side of the above inequality has infinite expectation, and so $E[s(X) - t(X)] = \infty$.

□

*Proof of Proposition 7.* Similar to the proof of Theorem 2, we split the integral $E \log|M(\xi;\beta)|$ into two components,

$$E \log|M(\xi;\beta)| = E\left[\log|M(\xi;\beta)|\,\mathbb{I}(\min_i |\eta_i| \le \kappa)\right] + E\left[\log|M(\xi;\beta)|\,\mathbb{I}(\min_i |\eta_i| > \kappa)\right], \quad \text{(S2.11)}$$

where $\kappa > 0$. Note that $w(\eta)$ is symmetric and decreasing in $|\eta|$. Thus, by Lemma 3, for $\min_i |\eta_i| \le \kappa$ we have $\log|M(\xi;\beta)| \le \log|F^{\mathrm{T}}F| + \log w(0)$, so

$$E\left[\log|M(\xi;\beta)|\,\mathbb{I}(\min_i |\eta_i| \le \kappa)\right] \le \Pr(\min_i |\eta_i| \le \kappa)[\log|F^{\mathrm{T}}F| + \log w(0)] < \infty\,,$$

i.e. the first term in (S2.11) is always $< \infty$.

We now consider the second term in (S2.11). For $\min_i |\eta_i| > \kappa$, provided

$\kappa$ is sufficiently large then by (2.8),

$$\max_i w(\eta_i) = w(\min_i |\eta_i|) \leq L \min_i |\eta_i| \, e^{-\min_i \eta_i^2/2} \,,$$

for some $L > 0$. By Lemma 3, for $\min_i |\eta_i| > \kappa$,

$$\log |M(\xi;\beta)| \leq \log |F^{\mathrm{T}} F| + p \log L + p \log \min_i |\eta_i| - (p/2) \min_i \eta_i^2 \,.$$

Assume that $|F^{\mathrm{T}}F| > 0$. Let $X_1 = \min_i |\eta_i|$, $A_1 = [0, \infty)$, $s_1(X_1) = (p/2) X_1^2 \, \mathbb{I}(X_1 > \kappa)$, and $t_1(X_1) = p \log X_1 \, \mathbb{I}(X_1 > \kappa)$. We have that

$$\begin{aligned} &E \left[ \log |M(\xi;\beta)| \, \mathbb{I}(\min_i |\eta_i| > \kappa) \right] \\ &\quad \leq E \left[ t_1(X_1) - s_1(X_1) + (\log |F^{\mathrm{T}}F| + p \log L) \, \mathbb{I}(X_1 > \kappa) \right] \,. \qquad \text{(S2.12)} \end{aligned}$$

We may assume that $\kappa > 1$, in which case $t_1$ is increasing and so $X_1$, $A_1$, $s_1$, $t_1$ satisfy the conditions of Lemma 6. Hence, if $E[s_1(X_1)] = \infty$ then $E[t_1(X_1) - s_1(X_1)] = -\infty$, in which case, from (S2.12),

$$E[\log |M(\xi;\beta)| \, \mathbb{I}(\min_i |\eta_i| > \kappa)] = -\infty \,,$$

and so by (S2.11) we have that $E \log |M(\xi;\beta)| = -\infty$. Hence to prove the proposition it is sufficient to show that $E[s_1(X_1)] = \infty$; we demonstrate that this holds in the next paragraph.

Recall that on the event $\mathcal{E}$, defined in (S2.6), we have that $\min_i |\eta_i| \geq$

$\min_i |f_j(x_i)||\beta_j| - \epsilon$. Thus, on $\mathcal{E}$,

$$\begin{aligned}\min_i \eta_i^2 &\geq \min_i |f_j(x_i)|^2|\beta_j|^2 - 2\epsilon \min_i |f_j(x_i)||\beta_j| + \epsilon^2 \\ &\geq s_2(X_2) - t_2(X_2) + \epsilon^2\,, \qquad \text{(S2.13)}\end{aligned}$$

where above $X_2 = \min_i |f_j(x_i)||\beta_j|$, $A = [0,\infty)$, with $s_2(X_2) = X_2^2$ and $t_2(X_2) = 2\epsilon X_2$. From assumptions (ii) and (iii) of the proposition we have that $E[|\beta_j|^2 \,\big|\, \mathcal{E}] = \infty$ and $\min_i |f_j(x_i)| > 0$, thus

$$E[s_2(X_2) \,|\, \mathcal{E}] = E[\min_i |f_j(x_i)|^2|\beta_j|^2 \,|\, \mathcal{E}] = \infty\,.$$

Hence, applying Lemma 6 we see that $E[s_2(X_2) - t_2(X_2) \,|\, \mathcal{E}] = \infty$ and so, by (S2.13), $E[\min_i \eta_i^2 \,|\, \mathcal{E}] = \infty$. To complete the proof we must consider the marginal expectation of $s_1(X_1) = (p/2)X_1^2\,\mathbb{I}(X_1 > \kappa)$, where $X_1 = \min_i |\eta_i|$. Note that by assumption (i), $\Pr(\mathcal{E}) > 0$, thus

$$EX_1^2 = E\min_i \eta_i^2 \geq \Pr(\mathcal{E})E(\min_i \eta_i^2 \,|\, \mathcal{E}) = \infty\,.$$

Finally, observe that $X_1^2 = X_1^2\,\mathbb{I}(X_1 \leq \kappa) + X_1^2\,\mathbb{I}(X_1 > \kappa)$ and

$$0 \leq E\{X_1^2\,\mathbb{I}(X_1 \leq \kappa)\} \leq \kappa^2\,.$$

Since $EX_1^2 = \infty$, we therefore have that $E\{X_1^2\,\mathbb{I}(X_1 > \kappa)\} = \infty$. Hence $E[s_1(X_1)] = \infty$. As shown in the previous paragraph, this is enough to establish that $E\log|M(\xi;\beta)| = -\infty$, and the proposition is proved.

□

*Proof of Theorem 3.* From Lemma 3 and the fact that $w(\eta) = \exp(\eta)$,

$$\begin{aligned} \log|M(\xi;\beta)| &\geq p \min_i \log w_i + \log|F^{\mathrm{T}}F| \\ &\geq p \min_i \eta_i + \log|F^{\mathrm{T}}F|\,. \end{aligned}$$

However, $\min_i \eta_i \geq -\max_i |\eta_i|$ and so

$$\log|M(\xi;\beta)| \geq -p \max_i |\eta_i| + \log|F^{\mathrm{T}}F|\,.$$

We know from the proof of Theorem 1 that $E \max_i |\eta_i| < \infty$ under the conditions given, and so we also have that $E \log|M(\xi;\beta)| > -\infty$ for the Poisson model. □

*Proof of Proposition 8.* First note from Lemma 3 that

$$\log|M(\xi;\beta)| \leq p \max_i \eta_i + \log|F^{\mathrm{T}}F|\,.$$

Thus, to establish that $E \log|M(\xi;\beta)| = -\infty$, it is sufficient to show that $E \max_i \eta_i = -\infty$. Similar to the proof of Proposition 4, the strategy is to find an event where $\eta_i$ is well approximated by $f_j(x_i)\beta_j$. Let $\mathcal{E}_2$ be an event such that $\beta_j < -1$ and $\sum_{k \neq j} |f_k(x_i)||\beta_k| < \epsilon$ for all $i$, where $\epsilon > 0$ satisfies

$$||f_j(x_i)| - |f_j(x_{i'})|| > 2\epsilon \qquad \text{for any } i, i' \text{ with } |f_j(x_i)| \neq |f_j(x_{i'})|\,.$$

For example, one possible definition is

$$\mathcal{E}_2 = \{\beta : \beta_j < -1\,,\, |\beta_k| < \delta \text{ for all } k \neq j\}\,,$$

with $\delta = \epsilon/[(p-1)\max_{i,k}|f_k(x_i)|]$. On $\mathcal{E}_2$, by arguments similar to those in the proof of Proposition 4,

$$\begin{aligned}\max_i \eta_i &\leq \max_i\{f_j(x_i)\beta_j\} + \epsilon \\ &\leq \beta_j \min_i f_j(x_i) + \epsilon\,, \end{aligned} \tag{S2.14}$$

where the second line follows since, by assumptions (i) and (v), we have $\max_i\{f_j(x_i)\beta_j\} = \beta_j \min_i f_j(x_i)$. By condition (iii), $E[\beta_j \,|\, \mathcal{E}_2] = -\infty$ and so, from (S2.14) and condition (v), we have $E[\max_i \eta_i \,|\, \mathcal{E}_2] = -\infty$. Moreover, by condition (ii), $\Pr(\mathcal{E}_2) > 0$ and so

$$E[\max_i \eta_i\, \mathbb{I}(\mathcal{E}_2)] = \Pr(\mathcal{E}_2)\, E[\,\max_i \eta_i \,|\, \mathcal{E}_2\,] = -\infty\,. \tag{S2.15}$$

Note that

$$\max_i \eta_i = \max_i \eta_i\, \mathbb{I}(\mathcal{E}_2) + \max_i \eta_i\, \mathbb{I}(\mathcal{E}_2^C)\,. \tag{S2.16}$$

By assumptions (i) and (v), $f_j(x_i)\beta_j$ is negative and so we have $\max_i \eta_i \leq \sum_{k\neq j} \max_i |f_k(x_i)||\beta_k|$. Thus by assumption (iv), $E\{\max_i \eta_i\, \mathbb{I}(\mathcal{E}_2^C)\} < \infty$. Hence, by (S2.16),

$$E \max_i \eta_i = \Pr(\mathcal{E}_2)\, E[\,\max_i \eta_i \,|\, \mathcal{E}_2\,] + \Pr(\mathcal{E}_2^C)\, E[\,\max_i \eta_i \,|\, \mathcal{E}_2^C\,]\,,$$

and, by (S2.15), we have $E\max_i \eta_i = \infty$ and hence $E\log|M(\xi;\beta)| = -\infty$.

□

*Proof of Proposition 1.* Let $\zeta_\epsilon$ denote a one-run exact design with design point $x_\epsilon = -1/\log\epsilon$, with $x_\epsilon \to 0$ as $\epsilon \to 0$. We show that, compared to $\zeta_\epsilon$, the relative Bayesian $D$-efficiency under $\mathcal{P}_\epsilon$ of a fixed (exact) design $\xi$ tends to zero as $\epsilon \to 0$. It is not claimed that $\zeta_\epsilon$ is Bayesian $D$-optimal. However, the relative Bayesian $D$-efficiency of $\xi$ is an upper bound for the absolute Bayesian $D$-efficiency of $\xi$, and so this argument is sufficient to establish that Bayes-eff$(\xi; \mathcal{P}_\epsilon) \to 0$ as $\epsilon \to 0$.

First note that under $\mathcal{P}_\epsilon$

$$E(\beta) = \int_\epsilon^a \frac{1}{\theta}\frac{1}{a-\epsilon}d\theta = \frac{\log a - \log\epsilon}{a-\epsilon} \to \infty \text{ as } \epsilon \to 0\,. \tag{S2.17}$$

Observe that, for the $\beta$-parameterization, using (2.1),

$$\phi(\xi;\mathcal{P}_\epsilon) - \phi(\zeta_\epsilon;\mathcal{P}_\epsilon) \leq \log S_{xx} - 2\min_{i:x_i>0}\{x_i\}E\beta - 2\log x_\epsilon + 2E\beta x_\epsilon$$
$$\leq \log S_{xx} - 2\left\{\min_{i:x_i>0} x_i - x_\epsilon\right\}E\beta - 2\log x_\epsilon\,.$$

Using (S2.17) and the definition of $x_\epsilon$, for $\epsilon$ sufficiently small,

$$\phi(\xi;\mathcal{P}_\epsilon) - \phi(\zeta_\epsilon;\mathcal{P}_\epsilon) \leq \log S_{xx} + 2K\log\epsilon - 2\log\left(\frac{-1}{\log\epsilon}\right)\,,$$

for some $K > 0$. Hence, provided $\epsilon$ is sufficiently small, the relative Bayesian $D$-efficiency satisfies

$$\exp\{\phi(\xi;\mathcal{P}_\epsilon) - \phi(\zeta_\epsilon;\mathcal{P}_\epsilon)\} \leq S_{xx}(\epsilon^K\log\epsilon)^2 \to 0 \text{ as } \epsilon \to 0\,,$$

which is sufficient to prove the claim. The limit above can be found using L'Hospital's rule. Using (2.2), the same result for the Bayesian $D$-efficiency also holds under the $\theta$-parameterization. $\square$

## Additional references

Berger, J. O. (1985). *Statistical decision theory and Bayesian analysis*. Springer, New York.

Berger, J. O. (2006). 'The case for objective Bayesian analysis'. *Bayesian Anal.* **1**, 385–402.

Garthwaite, P. H., Kadane, J. B., and O'Hagan, A. (2005). 'Statistical methods for eliciting probability distributions'. *J. Amer. Statist. Assoc.* **100**, 680–701.