A note on Bayesian credible sets in restricted parameter space problems and lower bounds for frequentist coverage [1]

Éric Marchand[a], William E. Strawderman[b]

a Université de Sherbrooke, Département de mathématiques, Sherbrooke Qc, CANADA, J1K 2R1 (e-mail: eric.marchand@usherbrooke.ca)

b Rutgers University, Department of Statistics, 501 Hill Center, Busch Campus, Piscataway, N.J., USA, 08855 (e-mail: straw@stat.rutgers.edu)

Summary

For estimating a lower bounded parametric function in the framework of Marchand and Strawderman (2006), we provide "through" a unified approach a class of Bayesian confidence intervals with credibility $1-\alpha$ and frequentist coverage probability bounded below by $\frac{1-\alpha}{1+\alpha}$. In cases where the underlying pivotal distribution is symmetric, the findings represent extensions with respect to the specification of the credible set achieved through the choice of a *spending function*, and including Marchand and Strawderman's HPD procedure result. More significantly for non-symmetric cases, the lower bound $\frac{1-\alpha}{1+\alpha}$ finding, although not applicable to the HPD procedure, is novel. Several examples are presented demonstrating wide applicability.

AMS 2000 subject classifications: 62C10, 62F15, 62F25, 62F30

Key words and phrases: Bayesian methods, Credible sets, Frequentist coverage probability, Lower bound, Restricted Parameter, Spending function

[1] August 2, 2012

1. Introduction

Bayesian credible sets are not designed (e.g., Robert, 2011) and are far from guaranteed (Fraser, 2011) to have satisfactory, exact or precise frequentist coverage but it is nevertheless of interest to investigate (Wasserman, 2011) to what extent there is convergence or divergence in various situations, including the choice of the prior and of the credible set. A historically resonating example where there is exact convergence arises for estimating the mean of a $N(\mu, \sigma^2)$ distribution, and where the use of the non-informative prior leads to a $(1-\alpha)\times 100\%$ HPD credible set (i.e. the z or t confidence interval) with exact frequentist coverage. This, however, is very much the exception. Even, in the simple presence of a lower bound on the mean parameter μ (e.g., Mandelkern, 2002), with the prior taken to be the truncation of the non-informative prior onto the restricted parameter space, the frequentist coverage of the $(1-\alpha)\times 100\%$ HPD credible set fluctuates from its credibility (or nominal coverage) $1-\alpha$. However, the HPD procedure does not fare poorly as a frequentist procedure for large $1-\alpha$ as witnessed by the lower bound $\frac{1-\alpha}{1+\alpha}$ on its frequentist coverage due to Roe and Woodroofe (2000, known σ^2) and Zhang and Woodroofe (2003, unknown σ^2), as well as the better lower bound $1-\frac{3\alpha}{2}$ (for $\alpha < 1/3$, known σ^2) obtained by Marchand et al. (2008).

Marchand and Strawderman (2006) showed that the lower bound $\frac{1-\alpha}{1+\alpha}$ is applicable in a vast number of situations. Their findings relate to the following setup.

Assumption 1. *We have a model density $f(x;\theta)$; $x \in \mathcal{X}$, $\theta \in \Theta$; for an observable X, with both X and θ being vectors, and we seek to estimate a parametric function $\tau(\theta)$ ($\Re^p \to \Re$) with the constraint $\tau(\theta) \geq 0$. We assume there exists a pivot $T(X,\theta) = \frac{a_1(X)-\tau(\theta)}{a_2(X)}$; $a_2(\cdot) > 0$; such that $-T(X,\theta)$ has cdf G and Lebesgue density g. We consider prior measures π and π_0, where $\pi_0(\theta) = \pi(\theta) I_{[0,\infty)}(\tau(\theta))$, and π is the Haar right invariant measure. We further assume that the decision problem is invariant under a group $\mathcal{G}$ of transformations and that the pivot satisfies the* *invariance requirement* $T(x,\theta) = T(gx, \bar{g}\theta)$, *for all* $x \in \mathcal{X}$, $\theta \in \Theta$, $g \in \mathcal{G}$, $\bar{g} \in \bar{\mathcal{G}}$, *with* $\mathcal{X}$, Θ, G, *and* $\bar{G}$ *being isomorphic. Referring to Marchand and Strawderman (2006, Lemma 2 and Corollary 1) for further details, we recall a key feature of the above structure, which is that the posterior distribution under* π *of* $T(x,\theta)$ *coincides, for all* x, *with the frequentist distribution of* $T(X,\theta)$ *(which is free of* θ*) and thus has cdf* G.

For symmetric and unimodal g, Marchand and Strawderman (2006) show that the $(1-\alpha)\times 100\%$ HPD credible set has frequentist coverage $C(\theta)$ greater than $\frac{1-\alpha}{1+\alpha}$ for all θ such that $\tau(\theta) \geq 0$. For non-symmetric densities g, Marchand and Strawderman (2006) provide proof in some cases, and evidence in others, of satisfactory performance for the HPD credible set. More precisely, the proof of the validity of the lower bound $\frac{1-\alpha}{1+\alpha}$ is achieved with an additional assumption on the type of asymmetry (skewness) of the distribution of the pivot $T(X,\theta)$, while the evidence is based on numerical evaluations of a theoretical and unexplicit lower bound for frequentist coverage (e.g., Example 2 of Marchand and Strawderman, 2006 concerning a lower bounded Gamma scale parameter). Nevertheless, a clear result or lower bound for frequentist coverage in such cases is lacking, and it is our motivation here to try to fill this gap.

Indeed, the main finding here is as follows. For prior measure π_0, we obtain a class of $(1-\alpha)\times 100\%$ credible sets for $\tau(\theta)$ for which

$$C(\theta) > \frac{1-\alpha}{1+\alpha}, \quad \text{for all } \theta \text{ such that } \tau(\theta) \geq 0. \tag{1}$$

This is achieved irrespectively of G and is thus totally general in the context described above. Our class of credible sets includes an "equal-tailed modification" of the HPD procedure which we describe below, and which coincides with the HPD procedure whenever g is symmetric and unimodal. Hence, our findings as applied to the equal tailed procedure extends those of the symmetric case.

It is by considering a class of Bayesian credible sets, rather than only the HPD procedure, that we are able to establish clear results on frequentist coverage. It is indeed with the analysis of the spending function related to credible set that we can pinpoint conditions for which the lower bound (1) holds. The main findings are elaborated upon in Section 2 and various applications are presented in Section 3. To facilitate the presentation of the results, here is a list of definitions and notations used.

Cheklist

- $1-\alpha$: credibility or posterior coverage or nominal frequentist coverage ($\alpha \in (0,1)$)
- $T(X,\theta) = \frac{a_1(X)-\tau(\theta)}{a_2(X)}$: pivot
- π: unrestricted prior density chosen as the right Haar invariant measure
- π_0 : prior density given by the truncation of π onto the restricted parameter space
- G: cumulative distribution function (cdf) of $-T(X,\theta)|x$ and of $-T(X,\theta)|\theta$ under π (which coincide for all x,θ)
- $g = G'$: probability density function (pdf) of $-T(X,\theta)$
- G^{-1}: inverse cdf
- $\alpha(\cdot)$: spending function
- $I_{\pi_0,\alpha(\cdot)}(X) = [l(X), u(X)]$: Bayesian credible set of credibility $1-\alpha$ associated with the prior π_0 and the spending function $\alpha(\cdot)$
- $C(\theta)$: the frequentist coverage at θ of the confidence interval $I_{\pi_0,\alpha(\cdot)}(X)$ given by $C(\theta) = P_\theta(I_{\pi_0,\alpha(\cdot)}(X) \ni \tau(\theta))$
- $y_0 = -G^{-1}(\frac{\alpha}{1+\alpha})$
- $t(x) = \frac{a_1(x)}{a_2(x)}$
- $\Delta_0(x) = (1-\alpha)(1-G(-t(x)))$

2. Main results

We begin this section by describing two different, yet equivalent, and instructive approaches to constructing a credible set for $\tau(\theta)$.

(A) (Spending function approach)

A $(1-\alpha)\times 100\%$ credible interval for $\tau(\theta)$ associated with prior π_0 can be generated by a *spending* function $\alpha(\cdot) : \Re^p \to [0,\alpha]$, such that $I_{\pi_0}(X) = [l(X), u(X)]$ [2] with $P_{\pi_0}(\tau(\theta) \geq u(x)|x) = \alpha(x)$ (and consequently $P_{\pi_0}(\tau(\theta) \leq l(x)|x) = \alpha - \alpha(x)$). More precisely, we have the following under Assmption 1.

Lemma 1. *For a given spending function* $\alpha(\cdot)$*, we have* $l_{\alpha(\cdot)}(x) = a_1(x)+a_2(x)G^{-1}\{G(-t(x))+(\alpha-\alpha(x))(1-G(-t(x))\}$ *and* $u_{\alpha(\cdot)}(x) = a_1(x)+a_2(x)G^{-1}\{1-\alpha(x)(1-G(-t(x))\}$*, with* $t(x) = \frac{a_1(x)}{a_2(x)}$.

Proof. Since the posterior cdf of $-T(X,\theta)$ under π is given by G, we have by definition of π_0, for $y \geq 0$:

$$P_{\pi_0}(\tau(\theta) \geq y|\, x) \;=\; \frac{P_\pi(\tau(\theta) \geq y|\, x)}{P_\pi(\tau(\theta) \geq 0|\, x)} \;=\; \frac{1-G(\frac{y-a_1(x)}{a_2(x)})}{1-G(-t(x))}\,. \tag{2}$$

Hence, we obtain for $\beta \in (0,1), y > 0$, $P_{\pi_0}(\tau(\theta) \geq y|\, x) \;=\; \beta \;\Leftrightarrow y = a_1(x)+a_2(x)G^{-1}\{1-\beta+\beta G(-t(x))\}$, and the result follows with the choices $\beta = \alpha(x)$ and $\beta = 1-(\alpha-\alpha(x))$ for $u(x)$ and $l(x)$ respectively. $\square$

Example 1. *The HPD procedures studied by Marchand and Strawderman (2006) for symmetric and unimodal* g *are given by the bounds* $l(x) = \max\{0, a_1(x)+a_2(x)G^{-1}(\frac{1-(1-\alpha)G(t(x))}{2})\}$ *and* $u(x) = a_1(x)+a_2(x)\min\{G^{-1}(1-\alpha G(t(x))), G^{-1}(\frac{1+(1-\alpha)G(t(x))}{2})\}$. *With these given bounds,*

[2] we suppress the dependence on $\alpha(\cdot)$ unless required

one may verify directly from (2) that the corresponding spending function is equal to

$$\min\{\alpha, \frac{\alpha}{2} + \frac{G(-t(x)}{2(1-G(-t(x)))}\}, \tag{3}$$

with $\alpha(x) = \alpha$ *if and only if* $t(x) \leq -G^{-1}(\frac{\alpha}{1+\alpha}) = G^{-1}(\frac{1}{1+\alpha})$ *since* g *is symmetric about* 0. *Conversely, applying Lemma 1 with the spending function choice* $\alpha(\cdot)$ *in (3) leads to the HPD procedure above (using the equality of* $G(\cdot)$ *and* $1 - G(-\cdot)$ *for symmetric about* g*).*

(B) (Approach based on quantiles of the pivot)

Alternatively, a second approach for cases where $l(x) > 0$ begins with choices γ_1 and γ_2, which will be made for each x, such that $G(\gamma_2) - G(-\gamma_1) = \Delta$, for a given $\Delta \in (0, 1)$. Since, for any x, we require $1 - \alpha = P_{\pi_0}(l(x) \leq \tau(\theta) \leq u(x)|x)$, we must have by (2):

$$G(\frac{u(x) - a_1(x)}{a_2(x)}) - G(\frac{l(x) - a_1(x)}{a_2(x)}) = (1-\alpha)(1 - G(-t(x))),$$

and this can be achieved with choices $-\gamma_1$ and γ_2 above for $\Delta = \Delta_0(x) = (1-\alpha)(1-G(-t(x)))$ yielding $\frac{u(x)-a_1(x)}{a_2(x)} = \gamma_2(\Delta_0(x))$ and $\frac{l(x)-a_1(x)}{a_2(x)} = -\gamma_1(\Delta_0(x))$, in other words

$$l(x) = a_1(x) - a_2(x)\,\gamma_1(\Delta_0(x)), \text{ and } u(x) = a_1(x) + a_2(x)\,\gamma_2(\Delta_0(x))\,, \tag{4}$$

whenever $l(x) > 0$. In view of the lower bound restriction on $\tau(\theta)$ (i.e., $\tau(\theta) \geq 0$), and the corresponding requirement that $l(X) \geq 0$, observe that not all choices of $-\gamma_1$ (and hence of γ_2) are feasible in (4) and that we must have

$$-\gamma_1(\Delta_0(x)) \geq -\frac{a_1(x)}{a_2(x)}\,.$$

Example 2. *With the above construction in (4), an equal-tailed choice of* $-\gamma_1$ *and* γ_2*, that is* $-\gamma_1(\Delta) = G^{-1}(\frac{1-\Delta}{2})$ *and* $\gamma_2(\Delta) = G^{-1}(\frac{1+\Delta}{2})$*, leads to the credible set bounds*

$$l(x) = a_1(x) + a_2(x)\,G^{-1}(\frac{1-\Delta_0(x)}{2}), \text{ and } u(x) = a_1(x) + a_2(x)\,G^{-1}(\frac{1+\Delta_0(x)}{2}), \tag{5}$$

when $l(x) > 0$. These above bounds coincide with those of the HPD procedure (when $l(x) > 0$) in the symmetric case of Example 1, as well as the spending function given in (3) as can be verified directly from (2) (Please note that the terminology "equal tails" does not mean $\alpha(x) = \alpha/2$ but rather to the choice of the quantiles $-\gamma_1$ and γ_2 under G.)

We define $y_0 = -G^{-1}(\frac{\alpha}{1+\alpha})$ and $t(x) = \frac{a_1(x)}{a_2(x)}$. Observing that $l(x) \leq 0$ in (5) if and only if $t(x) \leq y_0$, we are lead to investigate the frequentist performance of Bayesian credible intervals which have $l(x) = 0$, or equivalently $\alpha(x) = \alpha$ whenever $t(x) \leq y_0$, and which include the equal-tailed procedure of Example 2. We now have the following.

Theorem 1. *Under the conditions of Theorem 1 of Marchand and Strawderman (2006), that is Assumption 1, consider Bayesian credible intervals $I_{\pi_0,\alpha(\cdot)}$ associated with prior π_0 and a spending function $\alpha(\cdot)$ such that $\alpha(x) = \alpha$ for all x with $t(x) \leq y_0$. For the frequentist coverage $C(\theta) = P_\theta(I_{\pi_0,\alpha(\cdot)}(X) \ni \tau(\theta))$, we then have*

(a) $C(\theta) = \frac{1}{1+\alpha}$ for all θ such that $\tau(\theta) = 0$;

(b) Moreover, we have $C(\theta) > \frac{1-\alpha}{1+\alpha}$ for all θ such that $\tau(\theta) \geq 0$ as long as $\alpha(x)$ satisfies, for all x,

$$\frac{(1-\alpha)\,G(-t(x)) + \frac{\alpha^2}{1+\alpha}}{1 - G(-t(x))} \leq \alpha(x) \leq (\frac{\alpha}{1+\alpha})\frac{1}{1 - G(-t(x))}. \tag{6}$$

Proof.

(a) For θ such that $\tau(\theta) = 0$, we have

$$P_\theta(I_{\pi_0,\alpha(\cdot)}(X) \ni 0) = P_\theta(\alpha(X) = 0) = P_\theta(t(X) \leq y_0) \; = 1 - G(-y_0) = \frac{1}{1+\alpha},$$

since $-t(X)$ has cdf G whenever $\tau(\theta) = 0$.

(b) As in Marchand and Strawderman (2006), first observe that the confidence interval $I_1(X) = [l_1(X), u_1(X)] = \max\{0, a_1(X) + a_2(X)G^{-1}(\frac{\alpha}{1+\alpha})\}, a_1(X) + a_2(X)G^{-1}(\frac{1}{1+\alpha})\}$ has the same frequentist coverage as $I_1^*(X) = [a_1(X) + a_2(X)G^{-1}(\frac{\alpha}{1+\alpha})\}, a_1(X) + a_2(X)G^{-1}(\frac{1}{1+\alpha})\}]$ equal to $P_\theta(G^{-1}(\frac{\alpha}{1+\alpha}) \leq \frac{\tau(\theta)-a_1(X)}{a_2(X)} \leq G^{-1}(\frac{1}{1+\alpha})) = G(G^{-1}(\frac{1}{1+\alpha})) - G(G^{-1}(\frac{\alpha}{1+\alpha})) = \frac{1-\alpha}{1+\alpha}$. Now, we show that the given conditions on $\alpha(\cdot)$ imply that $I_{\pi_0,\alpha(\cdot)} \supseteq I_1$ which will lead to the result directly. Indeed, we have by the upper bound in (6) and Lemma 1: $u_{\alpha(\cdot)}(x) \geq a_1(x) + a_2(x)G^{-1}(1 - \frac{\alpha}{1+\alpha}) = u_1(x)$. Similarly, from the lower bound (6) and Lemma 1 we obtain $l(x) \leq a_1(x) + a_2(x)G^{-1}\{G(-t(x) + \alpha(1 - G(-t(x))) - \frac{\alpha^2}{1+\alpha} - (1-\alpha)G(-t(x))\} = a_1(x) + a_2(x)G^{-1}(\frac{\alpha}{1+\alpha}) = l_1(x)$. □

Corollary 1. *Under Assumption 1, the equal-tails credible interval $I_{\pi_0,\alpha(\cdot)}$ defined by the spending function $\alpha(\cdot)$ in (3) has minimum frequentist coverage $C(\theta)$ greater than $\frac{1-\alpha}{1+\alpha}$ for all θ such that $\tau(\theta) \geq 0$.*

Proof. It suffices to show directly that (6) is satisfied for the selection $\alpha(x) = \alpha_{\text{eqt}}(x)$ given in (3) for x such that $t(x) \geq y_0$. Indeed, we have for such x's:

$$\alpha_{\text{eqt}}(x)(1 - G(-t(x))) = \frac{\alpha}{2} + \frac{1-\alpha}{2}\,G(-t(x)) \leq \frac{\alpha}{2} + \frac{1-\alpha}{2}\,G(-y_0) = \frac{\alpha}{1+\alpha}\,,$$

and

$$\begin{aligned}\alpha_{\text{eqt}}(x)(1 - G(-t(x))) - (1-\alpha)\,G(-t(x)) - \frac{\alpha^2}{1+\alpha} &= \frac{\alpha(1-\alpha)}{2(1+\alpha)} - \frac{1-\alpha}{2}\,G(-t(x)) \\ &\geq \frac{\alpha(1-\alpha)}{2(1+\alpha)} - \frac{1-\alpha}{2}\,G(-y_0)) = 0.\end{aligned}$$

□

Remark 1. *In cases where the underlying pivotal distribution is non-symmetric, Corollary 1 is a new result, generalizing Theorem 1 of Marchand and Strawderman (2006), by virtue of Example 1 and is widely applicable given the lack of assumptions on g. Also, the bounds of the equal tails* *procedure are easier to evaluate than that of the HPD credible interval. And the findings of Theorem 1 go beyond a single procedure, even in the symmetric case, by providing a class of credible sets, as specified by a spending function, with frequentist coverage bounded below by* $\frac{1-\alpha}{1+\alpha}$.

3. Examples

At the risk of some redundancy with the examples provided by Marchand and Strawderman (2006), it is still beneficial here to present various applications with accompanying commentary. Assumption 1 is satisfied in all of the examples below with the underlying family of transformations (distributions) being either the location family, the scale family, or the location-scale family.

(A) (**a single location parameter**) $X \sim f_0(x-\theta)$; $\tau(\theta) = \theta \geq 0$; $T(X,\theta) = X - \theta$; $\pi_0(\theta) = 1_{[0,\infty)}(\theta)$. In such cases, all Bayes credible sets $I_{\pi_0,\alpha(\cdot)}$ (with credibility $1-\alpha$), with the spending function $\alpha(\cdot)$ satisfying the conditions of Theorem 1 and the bounds in (6), have necessarily minimum frequentist coverage bounded below by $\frac{1-\alpha}{1+\alpha}$. Through the transformations $X \to X - a$ and $X \to -X + a$, one can reduce all lower bounded restrictions $\theta \geq a$ and upper bounded restrictions $\theta \leq a$ to the case $\theta \geq 0$ considered here and we will not make further explicit mention of such transformations below.

Remark 2. *Results such as those in (**A**) are applicable as well for several observations by conditioning on a maximal invariant. Indeed, suppose that* $X = (X_1, \ldots, X_n) \sim f_0(x_1 - \theta, \ldots, x_n - \theta)$, *where* f_0 *is known and the* X_i*'s are not necessarily independently distributed. Set* $V = (X_2 - X_1, \ldots, X_n - X_1)$ *as a maximal invariant. One can then proceed, for a given value* v *of* V, *with an interval estimate* $I_{\pi_0,\alpha(\cdot,v)}(X_1, v)$ *as given in Lemma 1 with* $G \equiv G_v$ *representing the cdf of the pivot* $X_1 - \theta$ *conditional on* $V = v$, *and* $\alpha(x,v)$ *satisfying the conditions of Theorem 1 and (6). In such a case,* *Theorem 1 applies to the conditional frequentist coverage* $C(\theta, v) = P_\theta(I_{\pi_0,\alpha(\cdot,v)}(X, v) \ni \tau(\theta)|V = v)$ *yielding the inequality* $C(\theta, v) > \frac{1-\alpha}{1+\alpha}$ *for all* $\theta \geq 0$*. Since this is true for all* v*, the unconditional frequentist coverage* $C(\theta)$ *of the Bayes credible set* $I_{\pi_0,\alpha(\cdot,\cdot)}(X, V)$ *will also exceed* $\frac{1-\alpha}{1+\alpha}$ *for all* $\theta \geq 0$ *(see Marchand and Strawderman, 2006, for more details related to a multivariate Student model) In the same vein, all the scenarios below (B to G), although presented for simplicity in the single observation case, are also applicable in presence of a sample by conditioning on a maximal invariant.*

(B) (**a lower bounded scale parameter**) $X \sim \frac{1}{\theta} f_1(\frac{x}{\theta})\, 1_{(0,\infty)}(x)$ with $\theta \geq a$; $\tau(\theta) = \log(\theta) - \log(a) \geq 0$; $T(X, \theta) = \log(X) - \log(a) - \tau(\theta)$; $\pi_0(\theta) = \frac{1}{\theta} 1_{[0,\infty)}(\tau(\theta))$. Here, an interval estimate of $\tau(\theta)$ provides an interval estimate of θ. Important models include Gamma, Weibull, Fisher, among others. A familiar setup where the results can be applied arises in random effects analysis of variance models with a Fisher distributed pivot (see Zhang and Woodroofe, 2002, for details).

Remark 3. *Further applications consist of power parameter families where we have a scale family for an observable* Y *and the model of interest are the distributions for* $X = e^Y$*. As as simple illustration, consider the Pareto model for* X *with densities* $\frac{\gamma}{x^{\gamma+1}} 1_{(1,\infty)}(x)$ *and the parametric constraint* $\gamma \in (1, \gamma_0)$*. In such cases, we have that* $\gamma_0 \log(X) \sim$ *Exp*(θ) *with* $\theta = \frac{\gamma_0}{\gamma} \geq 1$ *and the results in (B) apply.*

(C) (**location-scale families**) $(X_1, X_2) \sim f_0(\frac{x_1-\theta_1}{\theta_2}, \frac{x_2}{\theta_2})\, 1_{(0,\infty)}(x_2)$; $\tau(\theta) = \theta_1 \geq 0$; $T(X, \theta) = \frac{X_1-\theta_1}{X_2}$; $\pi_0(\theta) = \frac{1}{\theta_2} 1_{(0,\infty)}(\theta_2) 1_{[0,\infty)}(\theta_1)$. This setup encompasses the basic normal case: $X_1, \ldots X_n$ ind. $\mathrm{N}(\mu, \sigma^2)$ with σ^2 unknown and $\mu \geq 0$, and more generally linear models $Y = Z\beta + \epsilon$, $\epsilon \sim N(0, \sigma^2 I_n)$ where the objective is to estimate a lower-bounded linear combination $\tau(\theta) = l'\beta$, by setting $X_1 = \hat{\beta}(Z'Z)^{-1}Z'Y$, $X_2 = \|Y - Z\beta\|^2$, $\theta_1 = \beta$, $\theta_2 = \sigma$. Here, the pivot $T(X, \theta)$ has a Student distribution. Alternatively, if the objective is to estimate a lower bounded scale θ_2, one can proceed as in (B).

(D) (**linear combination of several location parameters**) $X = (X_1, \ldots, X_p) \sim f_0(x_1 - \theta_1, \ldots, x_p - \theta_p)$; $\tau(\theta) = \sum_{i=1}^p a_i\theta_i; \pi_0(\theta) = 1_{[0,\infty)}(\tau(\theta))$, $T(X,\theta) = (\sum_{i=1}^p a_i X_i) - \tau(\theta)$. This setup includes, for instance, estimating a difference $\theta_1 - \theta_2$ with an order constraint $\theta_1 \geq \theta_2$.

(E) (**multivariate location-scale families with homogeneous scale**)

In (D), we can incorporate a common scale and apply the results of this paper for estimating a lower bounded linear combination with $X = (X_1, \ldots, X_p\,, X_{p+1}) \sim f_0(\frac{x_1-\theta_1}{\theta_{p+1}}, \ldots, \frac{x_p-\theta_p}{\theta_{p+1}}, \frac{x_{p+1}}{\theta_{p+1}})$, $\tau(\theta) = \sum_{i=1}^p a_i\theta_i$, $T(X,\theta) = \frac{(\sum_{i=1}^p a_i X_i)-\tau(\theta)}{X_{p+1}}$, and $\pi_0(\theta) = \frac{1}{\theta_{p+1}} 1_{(0,\infty)}(\theta_{p+1}) 1_{[0,\infty)}(\tau(\theta))$.

(F) (**several scale parameters**)

$(X_1, \ldots, X_p) \sim (\Pi_{i=1}^p \frac{1}{\theta_i})\ f_1(\frac{x_1}{\theta_1}, \ldots, \frac{x_p}{\theta_p})$; $\tau(\theta) = \sum_{i=1}^p a_i \log(\theta_i)$. This can consist, for instance, of estimating a lower bounded ratio $\frac{\theta_2}{\theta_1} \geq a$ of two scale parameters.

(G) (**quantiles in location-scale families**) $X_i \sim N(\mu, \sigma)$ iid; $\tau(\theta) = \mu + \eta\sigma \geq 0$, $\pi_0(\mu,\sigma) = \frac{1}{\sigma} 1_{(0,\infty)}(\sigma) 1_{(0,\infty)}(\mu + \eta\sigma)$. $T(X,\theta) = \frac{\bar{X}-\mu-\eta\sigma}{S}$. Here, $T(X,\theta)$ is distributed as non-central Student. The applications are not restricted to normality and are applicable in general for location-scale families as in (C).

Concluding remarks

For a large variety of situations with a lower bounded parametric constraint, we have obtained a class of Bayesian $(1-\alpha) \times 100\%$ credible sets which provide minimal frequentist probability coverage exceeding $\frac{1-\alpha}{1+\alpha}$. These Bayesian confidence intervals includes an equal tailed modification or approximation of the HPD credible set which coincides with the latter when the distribution of the underlying pivot is symmetric. We have made use of the spending function interpretation of Bayesian confidence intervals. In seeking to evaluate the frequentist performance of Bayesian confidence intervals, our results illustrate that the choice of bounds matters, so that there does not necessarily exist a single universal assessment of their frequentist performance even in a given specific problem.

Acknowledgements

Eric Marchand's research is supported in part by a grant from the Natural Sciences and Engineering Research Council of Canada, and William Strawderman's research is partially supported by a grant from the Simons Foundation (#209035).

References

Efron, B. (2004). President's Corner: Life in a Random Universe. AMSTATNews, December 2004, 2-3.

Feldman, G.J. and Cousins, R. (1998). Unified approach to the classical statistical analysis of small signals. *Physical Review D*, **57**, 3873-3889.

Fraser, D.A.S. (2011). Is Bayes posterior just quick and dirty confidence? *Statistical Science*, **26**, 299-316.

Mandelkern, M. (2002). Setting Confidence Intervals for Bounded Parameters with discussion. *Statistical Science*, **17**, 149-172.

Marchand, É., Strawderman, W. E., Bosa, K., and Lmoudden, A. (2008). On the frequentist coverage of Bayesian credible intervals for lower bounded means. *Electronic Journal of Statistics*, **2**, 1028-1042.

Marchand, É. and Strawderman, W. E. (2006). On the behaviour of Bayesian credible intervals for some restricted parameter space problems. *Recent Developments in Nonparametric Inference and Probability : A Festschrift for Micheal Woodroofe,*, IMS Lecture Notes-Monograph Series, **50**, pp. 112-126.

Marchand, É. and Strawderman, W. E. (2004). Estimation in restricted parameter spaces: A review. *Festschrift for Herman Rubin*, IMS Lecture Notes-Monograph Series, **45**, pp. 21-44.

Robert, C.P. (2011). Discussion of "Is Bayes posterior just quick and dirty confidence?" by D.A.S. Fraser. *Statistical Science*, **26**, 317-318.

Roe, B. and Woodroofe, M. (2000). Setting confidence belts. *Physical Review D*, **63**, 013009/01-09.

Wasserman, L. (2011). Frasian inference. *Statistical Science*, **26**, 322-325.

Zhang, T. and Woodroofe, M. (2003). Credible and confidence sets for restricted parameter spaces. *Journal of Statistical Planning and Inference*, **115**, 479-490.

Zhang, T. and Woodroofe, M. (2002). Credible and confidence sets for the ratio of variance components in the balanced one-way model. *Sankhyā: Special issue in memory of D. Basu*, **64**, 545-560.