

A Note on the Sensitivity of the Posterior Distribution to the Likelihood Function

Arnaud Doucet

Department of Statistics, University of Oxford, UK

Abstract

Assume you are given a prior distribution μ and two different likelihood functions G_0 and G_1 which result in two posteriors distributions π_0 and π_1 . We show here that the total variation norm between π_0 and π_1 can be upper bounded by half the oscillation of $\log G_0 - \log G_1$.

Disclaimer: I do not think the result presented here is original but I have not been able to find it in the literature.

Main Result

Let μ be a probability measure on $(E, \mathcal{E})$ and let $G_0 : E \rightarrow \mathbb{R}^+$ and $G_1 : E \rightarrow \mathbb{R}^+$ be two measurable μ -integrable functions. We write for $i = 0, 1$

$$\pi_i := \frac{\mu \cdot G_i}{\mu(G_i)} \quad (1)$$

where we assume that $\mu(G_i) := \int_E G_i(x) \mu(dx) > 0$. In a Bayesian context, μ would correspond to a prior distribution, G_i is the likelihood function and π_i the corresponding posterior distribution. In a non-Bayesian context, the measure π_i is often referred to as the Boltzmann-Gibbs transformation of μ through the potential function G_i [1]. Finally we write for any measurable function f

$$\text{osc}(f) = \text{ess sup } f - \text{ess inf } f$$

where the essential supremum and infimum are w.r.t μ . We have the following result.

Proposition 1. *The total variation norm between π_0 and π_1 satisfies*

$$\|\pi_0 - \pi_1\|_{\text{TV}} \leq \frac{1}{2} \text{osc}(\log G_0 - \log G_1).$$

Proof of the Proposition

For $i = 0, 1$, we write $g_i := \log G_i$ and we denote $\Delta g := g_1 - g_0$. We assume $\text{osc}(\Delta g) < \infty$, otherwise the result is trivial. For any $\lambda \in (0, 1)$, we introduce the measure

$$\pi_\lambda := \frac{\mu \cdot G_\lambda}{\mu(G_\lambda)}$$

where

$$\log G_\lambda := g_0 + \lambda \Delta g.$$

We have

$$\|\pi_0 - \pi_1\|_{TV} = \sup_{|f| \leq 1} \pi_1(f) - \pi_0(f).$$

Let f be a measurable bounded function such that $|f| \leq 1$ then

$$\pi_1(f) - \pi_0(f) = \int_0^1 \frac{d\pi_\lambda(f)}{d\lambda} d\lambda$$

where

$$\begin{aligned} \frac{d\pi_\lambda(f)}{d\lambda} &= \frac{d}{d\lambda} \frac{\mu(G_\lambda f)}{\mu(G_\lambda)} = \frac{\mu(G_\lambda \Delta g f) \mu(G_\lambda) - \mu(G_\lambda f) \mu(G_\lambda \Delta g)}{\mu^2(G_\lambda)} \\ &= \pi_\lambda(\Delta g f) - \pi_\lambda(f) \pi_\lambda(\Delta g) \\ &= \pi_\lambda(\{\Delta g - \pi_\lambda(\Delta g)\} f). \end{aligned}$$

Hence, we have

$$\begin{aligned} \pi_1(f) - \pi_0(f) &= \int_0^1 \pi_\lambda(\{\Delta g - \pi_\lambda(\Delta g)\} f) d\lambda \\ &\leq \int_0^1 \pi_\lambda(|\Delta g - \pi_\lambda(\Delta g)|) d\lambda \text{ as } |f| \leq 1 \\ &\leq \int_0^1 \pi_\lambda(|\Delta g - \pi_\lambda(\Delta g)|^2)^{1/2} d\lambda \end{aligned}$$

Now we have for any $\varsigma \in \mathbb{R}$

$$\pi_\lambda(|\Delta g - \pi_\lambda(\Delta g)|^2) \leq \pi_\lambda(|\Delta g - \varsigma|^2)$$

so

$$\begin{aligned} \pi_1(f) - \pi_0(f) &\leq \int_0^1 \pi_\lambda(|\Delta g - \varsigma|^2)^{1/2} d\lambda \\ &\leq \text{ess sup } |\Delta g - \varsigma| \end{aligned}$$

As it is true for any $\varsigma \in \mathbb{R}$, we have

$$\|\pi_0 - \pi_1\|_{TV} \leq \inf_{\varsigma \in \mathbb{R}} \text{ess sup } |\Delta g - \varsigma| = \frac{1}{2} \text{osc}(\Delta g).$$

References

- [1] P. Del Moral. *Feynman-Kac Formulae*. New York: Springer, 2004.