Statistics and Probability Letters 133 (2018) 38–41

Contents lists available at ScienceDirect

Statistics and Probability Letters

journal homepage: www.elsevier.com/locate/stapro

The median of an exponential family and the normal law

Gérard Letac [a,*], Lutz Mattner [b], Mauro Piccioni [c]

[a] Laboratoire de Statistique et Probabilités, Université Paul Sabatier, Toulouse, France
[b] Universität Trier, Fachbereich IV - Mathematik, 54286 Trier, Deutschland, Germany
[c] Dipartimento di Matematica, Sapienza Università di Roma, 00185 Roma, Italy

ARTICLE INFO

Article history:
Received 9 June 2017
Received in revised form 3 October 2017
Accepted 4 October 2017
Available online 23 October 2017

MSC 2010:
62E10
60E05
45E10

Keywords:
Characterization of the normal laws
Real exponential families
Median of a distribution
Choquet–Deny equation

ABSTRACT

Let P be a probability on the real line generating a natural exponential family $(P_t)_{t\in\mathbb{R}}$. We show that the property that t is a median of P_t for all t characterizes P as the standard Gaussian law $N(0, 1)$.

1. Introduction

Let P be a probability on the real line and assume that

$$L(t) = \int_{-\infty}^{+\infty} e^{tx} P(dx) < \infty \quad \text{for } t \in \mathbb{R}. \tag{1}$$

Such a probability generates the natural exponential family

$$\mathcal{F}_P = \{P_t(dx) = \frac{e^{tx}}{L(t)} P(dx),\ t \in \mathbb{R}\}.$$

Then it might happen that the natural parameter t of $\mathcal{F}_P$ is always a median of P_t, in the sense of

$$P_t((-\infty, t)) \leq \frac{1}{2} \leq P_t((-\infty, t]) \quad \text{for } t \in \mathbb{R}. \tag{2}$$

In the sequel we denote by $\mathcal{P}$ the set of probabilities P such that (1) and (2) are fulfilled. A noteworthy example of an element of $\mathcal{P}$ is the standard normal distribution $N(0, 1)$, for which $L(t) = e^{t^2/2}$ and $P_t = N(t, 1)$. It will turn out that it is the only one. The following preliminary lemmas simplify the study of $\mathcal{P}$.

* Corresponding author.
E-mail addresses: gerard.letac@math.univ-toulouse.fr (G. Letac), mattner@uni-trier.de (L. Mattner), piccioni@mat.uniroma1.it (M. Piccioni).

https://doi.org/10.1016/j.spl.2017.10.002

Lemma 1. *If $P \in \mathcal{P}$, then P is absolutely continuous with respect to Lebesgue measure. As a consequence, we have equality throughout in (2).*

Lemma 2. *If $P \in \mathcal{P}$, then its distribution function is strictly increasing.*

If $P \in \mathcal{P}$, then Lemma 1 allows us to write

$$P(\mathrm{d}x) = g(x)\varphi(x)\mathrm{d}x, \tag{3}$$

where g is some measurable non-negative function and $\varphi(x) = \mathrm{e}^{-x^2/2}/\sqrt{2\pi}$ denotes the standard normal density, and we will show that then $g(x) = 1$ a.e. to get:

Theorem 1. $\mathcal{P} = \{N(0, 1)\}$.

The proofs of the above results are contained in Section 2, followed by a conjecture and a further theorem.

2. Proofs

Proof of Lemma 1. The next paragraph shows that the distribution function of P is locally Lipschitz, and this implies the claimed absolute continuity, even with a locally bounded density, compare for example Royden and Fitzpatrick (2010, pp. 120–124).

For $t \in \mathbb{R}$, multiplying in assumption (2) by $L(t)$ yields

$$h(t) := \int_{(-\infty,t]} \mathrm{e}^{tx}P(\mathrm{d}x) \geq \frac{1}{2}L(t) \geq \int_{(-\infty,t)} \mathrm{e}^{tx}P(\mathrm{d}x) = h(t-). \tag{4}$$

Hence, if $A > 0$ is given, then for s, t with $-A \leq s < t \leq A$, we get

$$\begin{aligned} P\big((s,t)\big) &= \int_{(s,t)} \mathrm{e}^{-tx}\mathrm{e}^{tx}P(\mathrm{d}x) \;\leq\; \mathrm{e}^{A^2}\int_{(s,t)} \mathrm{e}^{tx}P(\mathrm{d}x) \\ &= \mathrm{e}^{A^2}\left(h(t-) - h(s) + \int_{(-\infty,s]} (\mathrm{e}^{sx} - \mathrm{e}^{tx})P(\mathrm{d}x)\right) \\ &\leq \mathrm{e}^{A^2}\left(\frac{1}{2}(L(t) - L(s)) + (t-s)\int_{\mathbb{R}} |x|\mathrm{e}^{A|x|}P(\mathrm{d}x)\right) \\ &\leq c_A \cdot (t-s) \end{aligned}$$

for some finite constant c_A. We have been using (4) and $|\mathrm{e}^u - \mathrm{e}^v| \leq |u-v|\mathrm{e}^w$ for $|u|, |v| \leq w$ at the penultimate step. Using assumption (1), we rely at the ultimate step on local Lipschitzness of L, due to its analyticity, and on finiteness of $\int_{\mathbb{R}} |x|\mathrm{e}^{A|x|}P(\mathrm{d}x)$. □

Proof of Lemma 2. Assume to the contrary that there exist $a, b \in \mathbb{R}$ with $a < b$ and $P((a, b)) = 0$. Then, for $t \in (a, b)$, Lemma 1 and (2) yield

$$\int_{-\infty}^{a} \mathrm{e}^{tx}P(\mathrm{d}x) = \int_{-\infty}^{t} \mathrm{e}^{tx}P(\mathrm{d}x) = \int_{t}^{+\infty} \mathrm{e}^{tx}P(\mathrm{d}x) = \int_{b}^{\infty} \mathrm{e}^{tx}P(\mathrm{d}x).$$

Thus the two measures $\mathbf{1}_{(-\infty,a]}(x)P(\mathrm{d}x)$ and $\mathbf{1}_{[b,+\infty)}(x)P(\mathrm{d}x)$ have finite and identical Laplace transforms on some non-empty interval. Hence the two measures coincide, and hence P must be the zero measure, which is absurd. □

Proof of Theorem 1. With the representation (3) for $P \in \mathcal{P}$, assumption (2) is rewritten as

$$\int_{-\infty}^{t} \mathrm{e}^{tx-\frac{x^2}{2}} \frac{1}{\sqrt{2\pi}} g(x)\,\mathrm{d}x = \frac{1}{2}\int_{-\infty}^{+\infty} \mathrm{e}^{tx-\frac{x^2}{2}} \frac{1}{\sqrt{2\pi}} g(x)\,\mathrm{d}x. \tag{5}$$

We multiply both sides by $\mathrm{e}^{-t^2/2}$:

$$\int_{-\infty}^{t} \mathrm{e}^{-\frac{(t-x)^2}{2}} \frac{1}{\sqrt{2\pi}} g(x)\,\mathrm{d}x = \frac{1}{2}\int_{-\infty}^{+\infty} \mathrm{e}^{-\frac{(t-x)^2}{2}} \frac{1}{\sqrt{2\pi}} g(x)\,\mathrm{d}x. \tag{6}$$

In other terms the unknown function g satisfies

$$\int_{-\infty}^{+\infty} \mathrm{sign}\,(t-x)\varphi(t-x)g(x)\,\mathrm{d}x = 0 \tag{7}$$

for all $t \in \mathbb{R}$. A formal derivation of (7) in t, using the product rule under the integral, and with one derivative being twice a delta function, leads to the equation

$$g(t) = \int_{-\infty}^{+\infty} q(t-x)g(x)\,\mathrm{d}x \tag{8}$$

a.e. in t, where $q(y) := \frac{1}{2}|y|e^{-\frac{y^2}{2}}$ is a probability density, but instead of justifying this formal differentiation, it seems easier to start by computing the derivative of

$$h(t) := \int_{-\infty}^{t} e^{tx} P(dx).$$

By Lemma 2 the distribution function F of P has a continuous inverse F^{-1}. Using the quantile transform we have

$$h(t) = \int_0^1 \mathbf{1}_{\{F^{-1}\le t\}}(u) e^{tF^{-1}(u)}\, du = \int_0^{F(t)} e^{tF^{-1}(u)}\, du = H(F(t), t)$$

with $H(s,t) := \int_0^s e^{tF^{-1}(u)}\, du$ for $s \in (0,1)$ and $t \in \mathbb{R}$. Now H has continuous partial derivatives $H_1(s,t) = e^{tF^{-1}(s)}$ and $H_2(s,t) = \int_0^s F^{-1}(u) e^{tF^{-1}(u)}\, du$, due to the continuity of F^{-1}, and hence H is differentiable. Let f be a Lebesgue density of P. Then, at every t where $F'(t) = f(t)$, and hence at Lebesgue-a.e. t, the chain rule yields

$$h'(t) = H_1(F(t), t) f(t) + H_2(F(t), t) \;=\; e^{t^2} f(t) + \int_0^{F(t)} F^{-1}(u) e^{tF^{-1}(u)}\, du$$

$$= e^{t^2} f(t) + \int_{-\infty}^{t} x e^{tx} f(x)\, dx.$$

Thus differentiating the identity (5) and observing that $f(x) = g(x)\varphi(x)$ we obtain the following a.e.-identity

$$\frac{1}{\sqrt{2\pi}} e^{t^2/2} g(t) + \int_{-\infty}^{t} x e^{tx-\frac{x^2}{2}} \frac{1}{\sqrt{2\pi}} g(x)\, dx = \frac{1}{2}\int_{-\infty}^{+\infty} x e^{tx-\frac{x^2}{2}} \frac{1}{\sqrt{2\pi}} g(x)\, dx,$$

and multiplying the latter by $\sqrt{2\pi}\, e^{-t^2/2}$ gives

$$g(t) = \frac{1}{2}\left(\int_t^{+\infty} x e^{-(t-x)^2/2} g(x)\, dx - \int_{-\infty}^{t} x e^{-(t-x)^2/2} g(x)\, dx\right).$$

Adding to the right hand side above the quantity

$$0 = \frac{t}{2}\left(\int_{-\infty}^{t} e^{-(t-x)^2/2} g(x)\, dx - \int_t^{+\infty} e^{-(t-x)^2/2} g(x)\, dx\right)$$

(recall (6)) yields the desired (8).

Next, with the (positive) Radon measures $\mu(dx) := g(x)dx$ and $\sigma(dx) := q(x)dx$, Eq. (8) can be rewritten as the so-called Choquet–Deny equation $\mu = \mu * \sigma$. Observe that $t \mapsto \int_{-\infty}^{+\infty} e^{tx}\sigma(dx)$ is even and strictly convex, and is therefore equal to 1 only at $t = 0$. We can now use the results in section 6 of Deny (1960), where "$n > 1$" is evidently a misprint for "$n \ge 1$", to conclude that μ has to be a positive scalar multiple of the Lebesgue measure. Since g is a probability density with respect to a probability measure, we have $g = 1$ a.e., and the theorem is proved. □

A trivial characterization of the standard normal exponential family is to say that the mean is equal to the parameter t. But it is worthwhile to mention a natural conjecture about exponential families which seems harder to establish:

Conjecture. *Suppose that the probability P satisfies (1), and denote $m(t) := \int_{\mathbb{R}} xP_t(dx)$. If for all t real $m(t)$ is a median of P_t, then $P = N(m, \sigma^2)$ for some m and σ.*

This conjecture, which is probably more meaningful from a methodological point of view than the result established in the paper, does not translate in a neat harmonic analysis statement as (7) and (8) and as such it seems harder to establish. The next simple result offers some support to the conjecture. A probability Q on $\mathbb{R}^n$ is said to be symmetric if there exists some $m \in \mathbb{R}^n$ such that $X - m \sim m - X$ when $X \sim Q$.

Theorem 2. *Let P be a probability on $\mathbb{R}^n$ such that*

$$L(t) = \int_{\mathbb{R}^n} e^{\langle t,x\rangle} P(dx)$$

is finite for all $t \in \mathbb{R}^n$. Assume that for all $t \in \mathbb{R}^n$ the probability $P_t(dx) = e^{\langle t,x\rangle} P(dx)/L(t)$ is symmetric. Then P is normal.

Proof. Clearly $m(t) = \int_{\mathbb{R}^n} xP_t(dx) = L'(t)/L(t)$ exists and, since P_t is symmetric, $X_t - m(t) \sim m(t) - X_t$ when $X_t \sim P_t$. Therefore its Laplace transform

$$s \mapsto \mathbb{E}(e^{\langle s, X_t - m(t)\rangle}) = e^{-\langle s, m(t)\rangle} \frac{L(t+s)}{L(t)}$$

does not change when we replace s by $-s$. Considering the logarithm and taking the derivative in s we get $2m(t) = m(t+s) + m(t-s)$. Taking again the derivative in s we get $m'(t+s) = m'(t-s)$ for all $t, s \in \mathbb{R}^n$, which means that m' is constant, hence $\log L$ is polynomial of degree at most 2, and hence P is normal. □

References

Deny, J., 1960. Sur l'équation de convolution $\mu = \mu * \sigma$. Séminaire Brelot-Choquet-Deny (Théorie du Potentiel) 4e année, 1959–60, Exposé numéro 5. http://www.numdam.org/article/SBCD_1959-1960__4__A5_0.pdf.
Royden, H.L., Fitzpatrick, P.M., 2010. Real Analysis, fourth ed. Prentice-Hall.