Journal de la société statistique de Paris

Maurice Dumas

Sénèque et les évidences à posteriori

Journal de la société statistique de Paris, tome 126, nº 4 (1985), p. 135-138
<http://www.numdam.org/item?id=JSFS_1985__126_4_135_0>

I

ARTICLES

SÉNÈQUE ET LES ÉVIDENCES A POSTERIORI

Maurice DUMAS
Ancien président de la Société de statistique de Paris (*)

Sénèque, repris par Laplace, constate que, dans les sciences, on rencontre des propositions à ce point évidentes, que l'on s'étonne qu'elles n'aient pas été dégagées bien avant qu'elles le furent effectivement.

L'auteur, qui les appelle des « évidences a posteriori *», donne quelques exemples de telles propositions, appartenant au domaine des probabilités* a priori *des paramètres de lois de probabilité; ces exemples conduisent en particulier aux lois de Lhoste, qui, rappelées dans le corps de l'article, trouvent des applications dans la plupart des raisonnements bayésiens.*

Seneca, and then Laplace, notice that, in sciences, we meet so evident propositions that we are surprised they have not been found for before they have effectively been.

The author, who calls them "a posteriori evidences", gives some examples of such propositions, in the field of a priori probabilities of parameters of probability laws; these examples lead particularly to the Lhoste laws, which, as the paper reminds, have applications in most bayesian reasonings.

> *« Le jour viendra où, par une étude suivie de plusieurs siècles, les choses actuellement cachées paraîtront avec évidence, et la postérité s'étonnera que des vérités si claires nous aient échappé. »*
>
> *SÉNÈQUE. Formule citée par LAPLACE, au début de son « Essai philosophique sur les probabilités ».*

1. Je me propose de faire état ici de différentes propositions relatives aux probabilités *a priori* de paramètres, à savoir exclusivement de quatre paramètres très usuels :

— p, pour la proportion de Blancs dans une population infinie;
— m, pour la valeur moyenne de lois de Laplace-Gauss;
— σ, pour l'écart-type des mêmes lois;
— h, pour le module de précision, lié à σ par $h\,\sigma\sqrt{2} = 1$.

Dans mon jargon personnel, j'appelle *« évidence a posteriori »* une proposition qui est devenue évidente pour moi, seulement après que je l'aie lue, ou que je l'aie trouvée moi-même à la suite soit d'une réflexion plus ou moins longue, soit de calculs parfois laborieux. Je garde toute ma considération aux auteurs qui n'ont pas eu comme moi la sorte de révélation qui est à la base de toute évidence *a posteriori*, mais je leur souhaite de l'avoir bientôt, bien avant la fin des siècles dont parle Sénèque.

(*) Maurice DUMAS, 10, boulevard Jourdan, 75014 Paris.
Journal de la Société de statistique de Paris, tome 126, n° 4, 1985.

2. *1923*. LA LOI dh/h DE LHOSTE

2.1. Au début du présent siècle, les ouvrages classiques faisaient état uniquement, en cas de connaissance nulle, de lois de probabilités *a priori* caractérisées respectivement par dp, par dm, par $d\sigma$ et par dh. En 1923, un modeste officier d'artillerie, Edmond Lhoste (1880-1948), a remarqué comme particularité pouvant affecter la loi de probabilité *a priori* à retenir, le fait que p était borné des deux côtés, tandis que m ne l'était d'aucun côté et que h, comme σ, l'était d'un seul. Alors, il a écrit que les deux intervalles (h_1, h_2) et (σ_1, σ_2) contenaient la même quantité d'information lorsque l'on avait $h_1 \sigma_1 \sqrt{2} = 1$ et $h_2 \sigma_2 \sqrt{2} = 1$. Par un raisonnement — qui serait certainement à revoir du point de vue de la rigueur mathématique — il est arrivé à ceci, qu'il fallait abandonner les lois $d\sigma$ et dh, au profit respectivement de $d\sigma/\sigma$ et dh/h, les deux lois n'en faisant évidemment qu'une seule.

2.2. Après Lhoste, plusieurs auteurs ont vu cette loi et l'ont admise pour des raisons diverses, dont aucune ne me semble avoir la valeur qu'aurait une bonne démonstration.

Ainsi, en *1937*, j'ai retenu dh/h sous un prétexte vraiment secondaire (elle conduit à des expressions de probabilité *a posteriori* correspondant harmonieusement à des expressions de probabilité *a priori*) et je l'ai très largement développée, d'ailleurs sans rencontrer aucune difficulté mathématique.

De son côté, H. Jeffreys, en *1939*, a retenu dh/h, en avançant — gratuitement — que c'était $\log h$, et non : h, qui devait être distribué uniformément : $d \log h = dh/h$.

Je cite enfin le travail de G. Hansel et D. Grouchko, *1965*, basé sur des considérations de « lois conjuguées naturelles ». Ils ont développé la loi $d\sigma/\sigma$, et ont retrouvé bon nombre de mes résultats de 1937.

2.3. Dès *1923*, Lhoste a remarqué que les lois dh/h et $d\sigma/\sigma$ pouvaient être employées l'une et l'autre dans un même calcul sans que le résultat en soit influencé. C'est là une remarque qui ne s'impose pas à première vue; ainsi, A. Kolmogoroff a bien remarqué que les lois dh et $d\sigma$ conduisaient à des résultats différents; mais sans chercher à porter remède à cela, il s'est borné à constater que les deux résultats coïncidaient à l'infini. Lorsque l'on se contente de cela, bien des choses se simplifient...

3. *1923*. LES TROIS LOIS DE LHOSTE

3.1. Lhoste a encore été plus loin. Il a remarqué que la loi dh/h permettait de passer aux lois convenant aux autres paramètres considérés ici. Voici les trois lois auxquelles il est arrivé :

— pour le paramètre p, borné par 0 et par 1 : $dp/p(1-p)$;
— pour le paramètre h(ou : σ), borné par 0 : dh/h (ou $d\sigma/\sigma$);
— pour le paramètre m, aucunement borné : dm.

3.2. *Les trois lois de Lhoste* sont cohérentes en ce sens que l'on passe de l'une à l'autre sans qu'aucune hypothèse particulière soit à faire.

Par exemple, si $U = p/(1-p)$ remplace p comme paramètre, on a bien $dU/U = dp/p(1-p)$.

Les trois lois de Lhoste correspondent à des « probabilités « virtuelles » », telles que celle dont la densité de probabilité, pour $0 \leqslant p \leqslant 1$, est $1/p(1-p)$, et dont l'intégrale est par suite infinie, au lieu d'être égale à 1. Mais la théorie récente (1955) de A. Renyi sur les « probabilités conditionnelles », où cette dernière condition n'est pas remplie, est de nature à rassurer les hésitants ». Les deux phrases entre guillemets ont été écrites en 1956, par notre ancien Président, le professeur M. Fréchet; remarquer le mot « virtuelles », mis par lui entre guillemets.

Je vais plus loin : en cas de connaissance nulle, la loi de probabilité *a priori* ne peut être que virtuelle, car si elle était réelle, cela reviendrait à avoir à agir comme si l'on possédait une connaissance non pas nulle, mais très réelle, celle que représente la loi réelle en cause — ce qui est manifestement absurde.

3.3. *Les trois lois de Lhoste* sont les lois limites de familles de lois de probabilité réelles. Le cas de la connaissance nulle n'est pas un de ces cas n'intéressant que les seuls théoriciens des probabilités. D'une façon générale, une famille de lois de probabilité *a priori*, si elle est vraie lorsqu'elle correspond à une connaissance non nulle, doit présenter un caractère de continuité (à défaut, elle ne pourrait, au mieux, qu'être une loi approchée) et, de plus, elle doit présenter le même caractère de continuité entre les cas de connaissance très faible et de connaissance nulle. Cette condition s'impose dans tous les cas : *les trois lois de Lhoste* remplissent cette condition.

4. Le paramètre p

4.1. *Des trois lois de Lhoste,* l'une, la loi dm, ne surprend aucunement : on est habitué à elle depuis Bayes et Laplace, et son caractère de loi virtuelle n'a vraiment gêné aucun auteur; la loi dh/h, déjà admise par certains, comme il a été dit, va, à n'en pas douter, l'être bientôt par tous; reste la loi $dp/p(1-p)$.

La pratique courante concernant p est de recourir à la loi dp, si aisée à manier; mais cette loi dp est inadmissible puisqu'elle est une loi réelle; une conséquence en est qu'elle conduit à un résultat que j'estime choquant : là où l'expérience a donné 1 sur 5, la loi dp dit d'espérer 2 sur 7, tandis que le bon sens dit d'espérer 1 sur 5.

4.2. Cependant, je ne connais guère d'auteur qui, d'après une connaissance chiffrée par n en N, n'ait pas adopté comme loi *a priori* une loi β, caractérisée, donc, par

$$p^{r-1}\,(1-p)^{s-1}\,dp$$

Les paramètres r et s de cette loi β sont choisis au mieux par chacun, mais un choix s'impose, celui de $r = n$ et $s = N-n$, car, avec ce choix, la loi β a pour espérance mathématique n/N, ce qui est conforme au bon sens. En cas de connaissance nulle, c'est-à-dire pour $n = N = 0$, cette loi β conduit précisément au $dp/p(1-p)$ de Lhoste. Ainsi, l'auteur qui utilise cette loi β sans connaître la loi $dp/p(1-p)$ de Lhoste, obéit sans le savoir à cette loi, et lui ajoute la sanction de la pratique.

4.3. Je crois devoir ajouter quelques mots, correspondant au cas où les données sont 0 blanche en N essais, N non nul. Le probabiliste conclut, justement, après avoir appliqué la loi $dp/p(1-p)$, à la probabilité 1 de la non-arrivée d'une blanche au prochain essai. Or, la probabilité 1, c'est la certitude, et la certitude a un caractère manifestement trop absolu en l'occurrence. Faut-il donc s'en prendre à la loi $dp/p(1-p)$? Nullement, car tout résultat de calcul est à interpréter avec bon sens, et le statisticien remarque naturellement qu'un tirage supplémentaire aurait pu suffire à détruire cette certitude : au résultat 0 en N, il fait correspondre un degré de certitude non pas égal à 1, mais *au moins égal* à celui qui aurait correspondu à 1 en $(N+1)$ et qui se chiffre par $N/(N+1)$.

Inversement, soit le cas d'un auteur qui, partant des mêmes données 0 en N, se montrerait satisfait de trouver une probabilité *a posteriori* bien déterminée, égale ni à 0 ni à 1. Je trouve qu'il aurait tort : il ne s'est pas trompé dans ses calculs, mais il s'est servi comme loi de probabilité *a priori*, d'une loi réelle laquelle a introduit dans les calculs, une connaissance qui n'était pas dans les données.

5. Récapitulation

Soit maintenant à passer en revue les propositions qui m'apparaissent comme évidentes, compte tenu des remarques faites plus haut, et donc comme *évidentes a posteriori* pour moi et pour les lecteurs qui ont bien voulu me suivre.

— Les paramètres des lois de probabilité sont à répartir en *trois classes* d'après leurs intervalles de variation; en ce qui concerne les paramètres que j'ai considérés, aucune sous-classe n'est à envisager.

— Une loi de probabilité *a priori* correspondant à *une connaissance nulle* doit nécessairement être une *loi virtuelle;* elle doit non moins nécessairement se présenter comme *loi limite* d'une famille de lois réelles.

— Tout calcul fait à partir des données 0 en N (N non nul) ne peut aboutir qu'à une *certitude a posteriori;* il est aisé pour le statisticien de tempérer le caractère trop absolu, attaché à toute certitude.

— *Les lois de Lhoste,* caractérisées respectivement par dm, par dh/h (ou par : $d\sigma/\sigma$) et par $dp/p(1-p)$ s'imposent à moi autant que s'impose une évidence.

En est-il de même pour le lecteur?

De toute façon, vivent les lois de Lhoste!

Et que le nom de leur inventeur leur soit conservé!

N.B. *Pour permettre au chercheur de se faire une idée plus complète sur les questions soulevées par mon exposé, je dépose à la Bibliothèque de notre Société :*

— *une photocopie du texte de 1923 de E. Lhoste;*

— *un tiré à part de mon mémoire « Lois de probabilité à priori de Lhoste ». Sciences et Techniques de l'armement. Mémorial de l'Artillerie française, 4e fasc. 1982, 29 p.*