Memos on densities and metrics on simplices

Luca <piero.mana@ntnu.no>

Draft of 15 January 2019 (first drafted 11 January 2019)

Some memos on densities and metrics on simplices.

1 Setup and base density

Consider $N+1$ mutually exclusive and exhaustive propositions. The distributions of their relative frequencies form an N-dimensional simplex. Label the propositions $\{0,\dots,N\}$, denote a relative-frequency distribution by $(q_0,\dots,q_N) =: \boldsymbol{q}$. In the rest of this memo it's always implicitly assumed that $q_i \geqslant 0$, also in the integration domains. Denote

$$\Delta_N \coloneqq \{(x_1,\dots,x_N) \mid x_i \geqslant 0, \textstyle\sum_i x_i \leqslant 1\}, \tag{1}$$

which is the $(N+1)$-simplex asymmetrically embedded in $\mathbf{R}^{N+1}$. As basic density (that is, volume element) we can take either

$$\mathrm{d}q_1 \cdots \mathrm{d}q_N, \qquad (q_1,\dots,q_N) \in \Delta_N, \tag{2}$$

or

$$\mathrm{d}q_0 \cdots \mathrm{d}q_N \; \delta(1 - \textstyle\sum \boldsymbol{q}), \qquad (q_0,\dots,q_N) \in [0,+\infty[^{N+1}, \tag{3}$$

which is borrowed from a Euclidean volume element of the embedding space. The latter leads to more symmetric formulae. The two densities are equivalent, and their integration gives $1/N!$, as can be proven inductively (Δ_k is the base of Δ_{k+1}: multiply its k-volume by a unit height and divide by $k+1$) or as shown in Jaynes (2003 § 18.10). Let's denote either density by $\mathrm{d}\boldsymbol{q}$. When (2) is intended, any q_0 that appears in the integral must be understood as $q_0 \equiv 1 - \sum_{i=1}^{N} q_i$.

2 Flat prior

The N-simplex has a natural convex structure. Thus the ratio of two N-volumes is well-defined. There's only one normalized density that assigns the same degree of belief to any two N-volumes having unit ratio:

$$N!\,\mathrm{d}\boldsymbol{q}, \tag{4}$$

called the flat prior.

3 Jeffreys prior

It is also possible to embed the N-simplex into a hyperspherical surface in $[0, +\infty[^{N+1}$ via

$$(q_1, \dots, q_N) \mapsto (x_0, x_1, \dots, x_N) = \left(\sqrt{1 - q_1 - \dots - q_N}, \sqrt{q_1}, \dots, \sqrt{q_N}\right). \tag{5}$$

The density induced by the Euclidean one is in this case given by

The relative-frequency distribution $\boldsymbol{q}$ in m observations can be obtained in $\binom{m}{m\boldsymbol{q}}$ ways, where

$$\binom{m}{m\boldsymbol{q}} \coloneqq \frac{m!}{(mq_0)! \cdots (mq_N)!} \approx \exp[m\mathrm{H}(\boldsymbol{q})] \tag{6}$$

is the multinomial coefficient. If we have equal beliefs in the occurrence of these ways and m is very large, our belief about the relative frequency $\boldsymbol{q}$ can be approximated by the density

$$\frac{\exp[m\mathrm{H}(\boldsymbol{q})]}{\int \mathrm{d}\boldsymbol{q}\ \exp[m\mathrm{H}(\boldsymbol{q})]}\, \mathrm{d}\boldsymbol{q}, \tag{7}$$

called the entropy prior.

4 Metric

A density on the simplex doesn't induce any canonical density on a lower-dimensional subset. One way to induce a density on every lower-dimensional subset is to equip the simplex with a metric. Let's consider the flat metric. We find its expression in the coordinates $(q_1, \dots, q_N)$ by symmetrically embedding the simplex in $[0, +\infty[^{N+1}$ as a flat hypersurface and pulling back the Euclidean metric onto it. We use the embedding

$$(q_1, \dots, q_N) \mapsto (x_0, x_1, \dots, x_N) = (1 - q_1 - \dots - q_N, q_1, \dots, q_N), \tag{8}$$

which has tangent map

$$\begin{pmatrix} -\boldsymbol{u} \\ I_N \end{pmatrix} \tag{9}$$

with

$$\boldsymbol{u} \coloneqq \underbrace{\begin{pmatrix}1 & \dots & 1\end{pmatrix}}_{N} \tag{10}$$

and I is the identity matrix.

In coordinates $\boldsymbol{x}$ the metric is represented by the identity matrix I_{N+1}. The representation of its pull-back is therefore

$$\begin{pmatrix}-\boldsymbol{u}^\top & I_N\end{pmatrix} I_{N+1} \begin{pmatrix}-\boldsymbol{u}^\top \\ I_N\end{pmatrix} = I_N + \boldsymbol{u}^\top \boldsymbol{u}. \tag{11}$$

This is a matrix with all unit elements outside the diagonal and 2 on the diagonal. Its determinant can be found with Sylvester's theorem (Sylvester 1851; Akritas et al. 1996):

$$\det(I_N + \boldsymbol{u}^\top \boldsymbol{u}) = \det(I_1 + \boldsymbol{u}\boldsymbol{u}^\top) = 1 + N. \tag{12}$$

Bibliography

('de *X*' is listed under D, 'van *X*' under V, and so on, regardless of national conventions.)

Akritas, A. G., Akritas, E. K., Malaschonok, G. I. (1996): *Various proofs of Sylvester's (determinant) identity*. Math. Comput. Simulat. **42**$^{4–6}$, 585–593. http://www.inf.uth.gr/~akritas/articles/38.pdf.

Jaynes, E. T. (2003): *Probability Theory: The Logic of Science*. (Cambridge University Press, Cambridge). Ed. by G. Larry Bretthorst. First publ. 1994. https://archive.org/details/XQUHIUXHIQUHIQXUIHX2, http://www-biba.inrialpes.fr/Jaynes/prob.html.

Sylvester, J. J. (1851): *On the relation between the minor determinants of linearly equivalent quadratic functions*. Phil. Mag. **1**^{4}, 295–305.