A geometric tutorial on exchangeability and related topics

P.G.L. Porta Mana
<piero.mana@ntnu.no>

23 June 2018; updated 3 September 2018

Application of the concept of sufficiency and the Pitman-Koopman theorem to the definition and measurement of 'cooperativity' among neurons

> Most books on modern physics assume a knowledge of probability theory and claim to derive probability distributions. A casual observer might be led into supposing that the authors would be reasonably up to date in the subject. In fact, however, they seldom show any knowledge of it that is less than fifty years old
>
> H. Jeffreys (1955 p. 275)

> I believe that the greatest unsolved problem in statistics is communicating the subject to others.
>
> N. Davies (in Copas et al. 1995 p. 445)

1 Introduction

Exchangeable models are a family of statistical models that appear over and over in all sciences. Many scientists probably use them routinely without knowing the technical name 'exchangeable model', used mainly in the statistics literature – and possibly also without knowing of some important properties of these statistical models. This ignorance is mainly caused by a language barrier: most literature that discusses exchangeable models uses a very specialized technical language strongly connected with topology, measure theory, differential geometry, and other subjects that are usually only superficially touched in the teaching of other sciences.

Such subjects are of course necessary for a precise foundation of exchangeable models and other statistical notions. So I don't think that scientists are justified in shying away from them. But I don't think that statisticians are justified in adopting an over-technical jargon either, wherever a simpler one can suffice. They often seem to willingly insulate themselves within linguistic walls, instead of building bridges that could make their beautiful work more accessible to other sciences. Most technical concepts underlying statistics have, in fact, a very intuitive and even visual meaning at bottom.

The present note tries to build a bridge, explaining the essentials of exchangeable models and some related topics in an intuitive and geometric way, hoping to make the main concepts more accessible to scientists in other disciplines. I'm not a statistician or a specialized scientist, so my attempt will likely leave all parties unsatisfied. I hope it will at least give a hunch that a bridge is worthwhile constructing.

The other topics discussed in this note include:

- the choice of so-called 'parameter priors'
- so-called 'non-parametric models'
- model comparison
- sufficiency and its relation to some toy physical models, like Ising models
- partial exchangeability and its relation to machine learning and neural networks
- the phenomenon of 'overtraining'

The mathematics I'll use is very simple: functional analysis and combinatorics. The geometry will rely on convex spaces. These spaces appear in all of statistics. Their basics are very intuitive, and their visualization often allows us to guess mathematical properties or solutions to problems that are harder to guess algebraically. For this reason this note starts with a summary about them.

2 Exchangeability

There's a deep reason why exchangeable models appear in all sciences: they are connected with *reproducibility*, and therefore also with *induction*.

Reproducibility and circularity.

Exchangeability expresses the way we do induction (doesn't justify it).

Example with pattern: we discard exchangeability (dead hypothesis resurrection) – just as we do with reproducibility.

3 Convex spaces

4 Exchangeable models and their update

5 Probability distributions for the limit frequencies

6 Model comparison

7 Models with sufficient statistics

8 Partially exchangeable models

✠ to be continued

Bibliography

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

Chatfield, C. (1995): *Model uncertainty, data mining and statistical inference*. J. Roy. Stat. Soc. A **158**3, 419–444. See also discussion in Copas, Davies, Hand, Lunneborg, Ehrenberg, Gilmour, Draper, Green, et al. (1995).

Copas, J. B., Davies, N., Hand, D. J., Lunneborg, C. E., Ehrenberg, A. S., Gilmour, S. G., Draper, D., Green, P. J., et al. (1995): *Discussion of the paper by Chatfield [Model uncertainty, data mining and statistical inference]*. J. Roy. Stat. Soc. A **158**3, 444–466. See Chatfield (1995).

Jeffreys, H. (1955): *The present position in probability theory*. Brit. J. Phil. Sci. **5**20, 275–289.

Moore, G. P., Perkel, D. H., Segundo, J. P. (1966): *Statistical analysis and functional interpretation of neuronal spike data*. Annu. Rev. Physiol. **28**, 493–522.