Quantifying uncertainty about the contents of a magic urn

P.G.L. Porta Mana
<pgl@portamana.org>

Draft of 24 June 2019 (first drafted 22 June 2019)

***

Note: Dear Reader & Peer, this manuscript is being peer-reviewed by you. Thank you.

1 A magic urn

Quantifying our uncertainty about the contents of a urn that contains, say, white and black balls in an unknown proportion is a standard textbook topic. Our uncertainty depends on our data about the urn and on our state of knowledge about the urn before we knew those data; for example we know the colours of five balls drawn from the urn, and how or why the urn was prepared. Jaynes's book (2003 chs 3 & 6) gives a superb analysis and discussion of this problem and of the probability formulae that typically arise there; and even of more realistic variants, such as a tendency of drawing and replacing from the top layer of the urn (ibid. § 3.9).

In this note I derive some probability formulae for a more complex version of the urn problem, and some approximations of these formulae. They can be useful in several scientific questions that, sufficiently simplified, can be mapped to the present problem. examples

The problem we consider is the following. There's an urn with N balls, each of which can be white or black. At regular time intervals every ball can change its colour: a white can turn black, or vice versa; or it can keep the colour it had. We can observe n out of the N balls for T time intervals. The n observed balls are initially chosen in a way unknown to us. Note that we observe the *same* n balls at all times. Some variations of this set-up will be discussed later on. We ask several questions about this magic urn:

Q1 What was the proportion of white and black balls in the urn at some specific time, among the times we observed?

Q2 What was (or will be) the proportion of white and black balls in the urn at some specific time, among the times we did not observe?

Q3 How frequently did every possible proportion of white and black balls appear, during the times we observed?

Q4 How frequently will every possible proportion of white and black balls appear, during all times the urn exists?

And our goal is to quantify our belief in each of their possible answers.

The set-up above is an abstract picture; in a real situation our uncertainty and knowledge would be more specific – although their details are often difficult to specify in words. The quantification of our beliefs strongly depends on such details; it ought to be faced differently in each specific situation, just like every probability problem should. The formulae I present below apply only to an abstract, unrealistic situation. But they may be taken as an approximation or as a starting point for formulae more suited to some particular, real situations.

Bibliography

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

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.