Comment on Glenn Shafer's paper

Vladimir Vovk

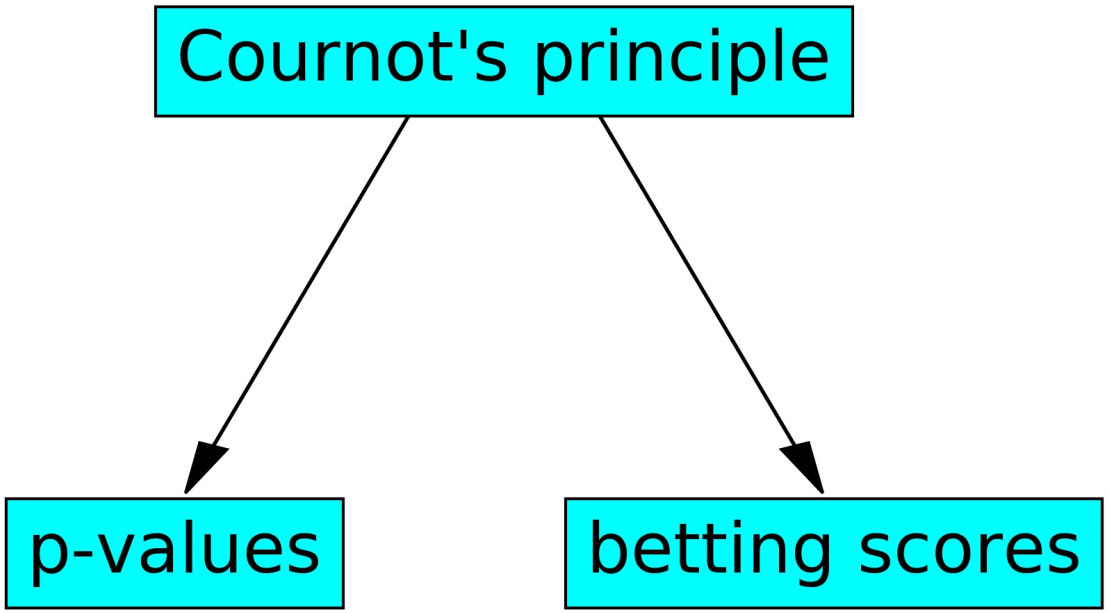
Department of Computer Science
Royal Holloway, University of London

RSS discussion meeting
9 September 2020

Betting scores and p-values

- This paper is a powerful appeal for a wider use of betting ideas and intuitions in statistics.
- Shafer admits that p-values will never be completely replaced by betting scores.
- Both p-values and betting scores generalize Cournot's principle (a fixed event of a small probability is not expected to happen), but they do it in their different ways, and both ways are interesting and valuable.

Two generalizations



Bayes factors (1)

- Other authors have referred to betting scores as Bayes factors (Shafer, Shen, Vereshchagin, Vovk,...) and e-values (Wang, Grünwald, de Heide, Koolen, Vovk,...).
- For simple null hypotheses, betting scores and usual Bayes factors coincide, but for composite null hypotheses they diverge, and using "Bayes factor" to mean "betting score" is confusing (especially to Bayesians), and I am sure should be avoided.

Bayes factors (2)

- But the Bayesian connection still allows us to apply Jeffreys's rule of thumb to betting scores: a p-value of 5% is roughly equivalent to a betting score of $\sqrt{10}$, and a p-value of 1% to a betting score of 10.
- This agrees beautifully with Shafer's rule

$$\frac{1}{\sqrt{p}} - 1,$$

which gives, to two decimal places:

  - for $p = 5\%$, 3.47 instead of Jeffreys's 3.16;
  - for $p = 1\%$, 9 instead of Jeffreys's 10.

E-values

- The term “e-values” emphasizes the fundamental role of expectation in the definition of betting scores.
- The natural habitat for “betting scores” is game-theoretic while for “e-values” it is measure-theoretic.

Combining betting scores

- Betting scores are easier to combine in general (e.g., the average of betting scores is a betting score).
- Useful in multiple hypothesis testing (Ruodu Wang will say more in his written comment).
- In my online appendix to this comment I explain that averaging betting scores allows us to achieve "inferential reproducibility" in David Cox's flexible and efficient way of testing statistical hypotheses via data splitting.

Technical report


Vladimir Vovk.
A note on data splitting with e-values: online appendix to my comment on Glenn Shafer's "Testing by betting".
arXiv:2008.11474 [stat.ME], August 2020.

Thank you for your attention!