coherence assessed via gambling: a discussion

Christian P. Robert

Université Paris-Dauphine (CEREMADE)
& University of Warwick (Dept. of Statistics)

http://xianblog.wordpress.com,seriesblog.net

4th Bayes, Frequentism, Fiducial Conference, May 1 2017

DAUPHINE UNIVERSITÉ PARIS

Validation of statistical procedures

"One of the most compelling arguments for Bayesian decision theory and the use of probability to model uncertainty..."

- what does make a procedure coherent? [relative to a given Universe]
- why should we play by the [new] rules set by the authors? [why should we play at all?]
- how comes the reasoning involves averaging over observations? [frequentism]...
- ...and validates all prior procedures?

In which sense does it relate to statistical practice?

how not to gamble if you must not...

[Freedman and Purves, 1969; Schervish, Seidenfeld, and Kadane, 2002]

"Coherence" is defined in terms of gambling

- who should be forced to bet and why?
- gambling is ethically disputable
- equating statisticians with bookmakers = stain on the profession [which one?!]
- leads to a highly formal(ised) framework
- [value-laden terms (e.g., incoherent, irrational)]
- "compelling", really??? [I never understood the argument]

O herence, where art thou?

Weak notion: excludes only predictions that do not enter the range of the random variable, i.e. sure loss for bookmaker

$$\sup_{\omega \in \Omega} \alpha_i (X_i(\omega) - p_i) < 0$$

Self-fulfilling: involves a probability measure over the space of predictions *[Freedman, 2003]* [and takes betting for granted]
Doom-monger: does not weight down extremes and unlikely events [another nuance of minimax?!]
a-scientific: Science does not advance by gambling or voting
a-decisional: turns all losses into additive monetary losses

O herence, where art thou?

Weak notion: does not exclude any proper prior distribution [what of improper priors?]
Self-fulfilling: involves a probability measure over the space of predictions *[Freedman, 2003]* [and takes betting for granted]
Doom-monger: does not weight down extremes and unlikely events [another nuance of minimax?!]
a-scientific: Science does not advance by gambling or voting
a-decisional: what's wrong with admissibility?

the elephant in the room...

Based on

- assumption of a probabilistic model generating the random [??] data
- and of repeatability of experiments
- and infinite precision of prior

testing example[s]

Disclaimer: I am not defending classical procedures *per se*!

In simple vs. simple hypothesis testing, UMP test at level α is Bayes procedure as well for specific 0–c loss function

testing example[s]

Disclaimer: I am not defending classical procedures *per se*!

Classical statistician does

- produce test at every α level
- associate to a different α, a different loss function
- change decision rule if faced with different loss function [hopefully!]
- ignore Bayesian interpretation

testing example[s]

Disclaimer: I am not defending classical procedures *per se*!

Classical statistician does not

- mix values of σ
- "prefer level $\alpha = 0.05$ test whenever available" [SSK, 99]
- wish to bet or gamble
- see the $0-1$ loss as connecting α level UMP tests

Bernoulli example

Criticism of minimax estimator of θ

$$\delta^0(n, X) = \frac{X + \sqrt{n}/2}{n + \sqrt{n}} \qquad X \sim \mathcal{B}(n, \theta)$$

[Bayes estimator under $\text{Be}(\sqrt{n}/2, \sqrt{n}/2)$ prior]

Summarised by existence of two Bayes estimators $\delta_1(n_1, X_1)$ and $\delta_2(n_2, X_2)$ such that

$$\alpha_1 R(\delta_1(n_1, \cdot), \theta) + \alpha_2 R(\delta_2(n_2, \cdot), \theta) \leqslant \alpha_1 R(\delta^0(n_1, \cdot), \theta) + \alpha_2 R(\delta^0(n_2, \cdot), \theta)$$

Amounts to $(\delta^0(n_1, X_1), \delta^0(n_2, X_2))$ being inadmissible under this new loss

Bernoulli a-example

But...

- minimax estimator is proper Bayes and admissible under original loss
- dominating estimator is mix of two Bayes estimators against different priors, hence not a Bayes estimator...
- ...although there exists a Bayes minimax estimator, suffering from similar defaults [associated with problem induced prior]
- new loss artificial in adding errors for two sample sizes when estimating same probability θ
- when estimating two values of θ, no Stein effect possible: aggregation of admissible estimators remains admissible

ultimate issue with statistics

"...the ultimate inaccessibility of a reality that is truly independent of observers is a basic human condition." A. Gelman & C. Hennig, RSS, April 2017

Except for the most [basic] scientific settings, there is not reality behind statistical models [Box], hence an inaccessible consensus is the rule

ultimate issue with statistics

"...the ultimate inaccessibility of a reality that is truly independent of observers is a basic human condition." A. Gelman & C. Hennig, RSS, April 2017

Except for the most [basic] scientific settings, there is not reality behind statistical models [Box], hence an inaccessible consensus is the rule

Where is the $\mathcal{M}$-open version of coherence?

Relativity

Coherence of soft?

Pertains to awareness of multiple perspectives but also to stability

Fundamental dependence of inferential output and efficiency assessment on statistical framework(s)

- difference in outcomes perfectly acceptable
- inner assessment of model feasible by producing pseudo data
- comparison of frameworks only to the extent of predictive characteristics

Relativity

Coherence of soft?

Pertains to awareness of multiple perspectives but also to stability

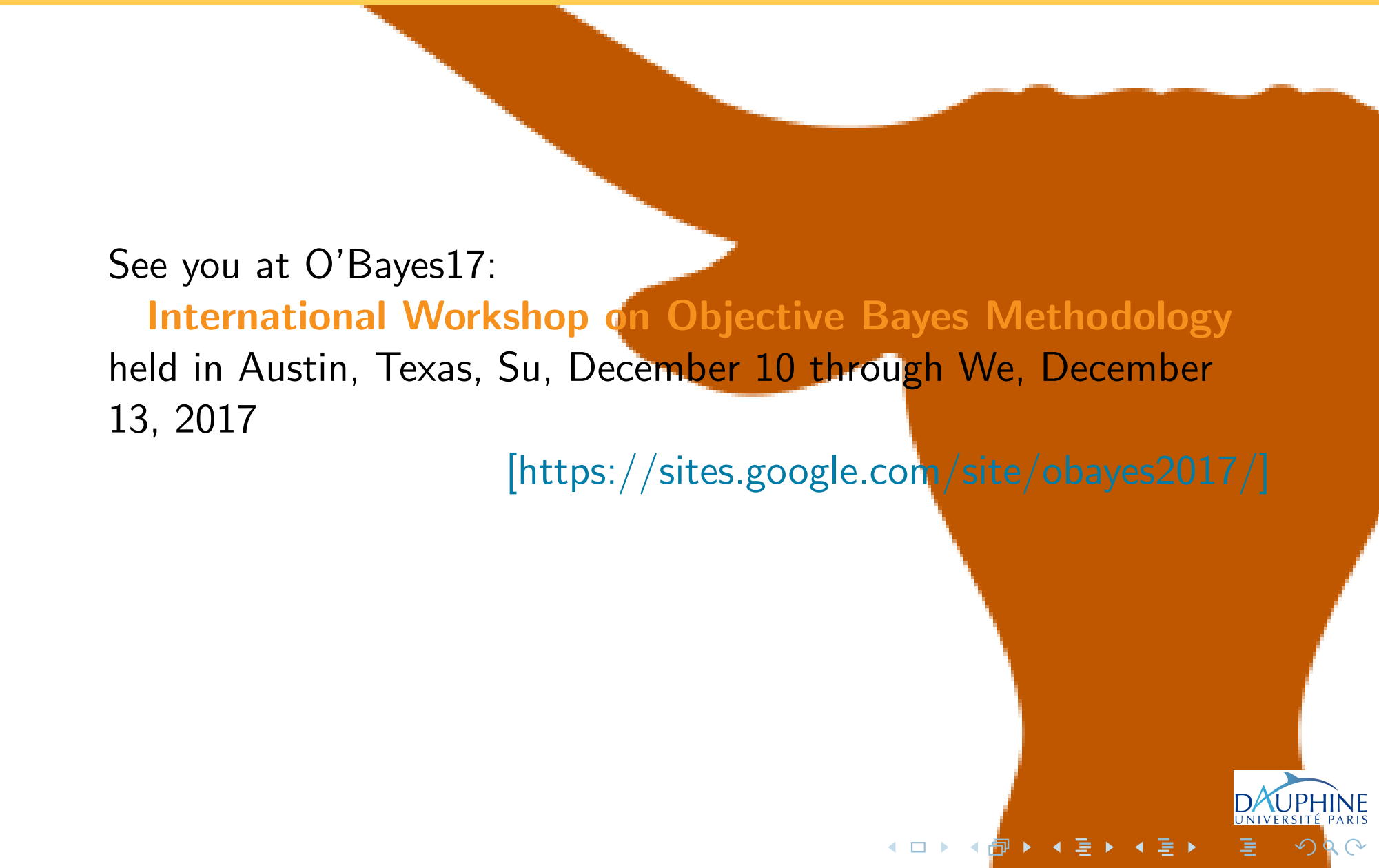
- Bayesian analysis [both objective and subjective] well-suited to this purpose / rescued by relativity
- reproducibility fraught with danger: production of similar results depends on rigid framework, relates to fact that statistics not experimental science

embracing uncertainty with a subjective hug

basic realism and uncertain nature of data call for an absence of hard decisions like tests and model choices, but rather for descriptive performances of the suggested procedures, accepting imperfection and variability in the answer(s) produced locally

> *"It's not so hard to move away from hypothesis testing and toward a Bayesian approach of embracing variation and accepting uncertainty."* A. Gelman, April 2017

a debate to be continued, hopefully



See you at O'Bayes17:

International Workshop on Objective Bayes Methodology

held in Austin, Texas, Su, December 10 through We, December 13, 2017

[https://sites.google.com/site/obayes2017/]

DAUPHINE UNIVERSITÉ PARIS