

Commentary on Descartes Lectures 2014

# **Statistics and Full Belief**

Or: How to Judge if You Must

★

Jan-Willem Romeijn  
University of Groningen

Joint work with Hannes Leitgeb  
LMU Munich

## Ring-ring

You're called up by a decision maker who says: what statistical facts can I rely on?



What do you report? Can you advise acceptance of your favourite statistical hypothesis?

# This talk

This talk offers a perspective on an important domain of application of the stability-based belief concept.

- The problem of how statistics relates to full belief is in need of a principled solution.
- One attractive solution employs a notion of belief based on stability, as developed by Leitgeb.
- This solution makes full belief highly sensitive to the context of the decision maker.

# **Contents**

<b>1 Beliefs in statistics</b>	<b>5</b>
<b>2 Statistics and the stability approach</b>	<b>10</b>
<b>3 Example: Neyman-Pearson tests</b>	<b>16</b>
<b>4 Contextual beliefs</b>	<b>22</b>
<b>5 Summing up</b>	<b>25</b>

# 1 Beliefs in statistics

In the foundational debates, classical and Bayesian statistics offer only limited room for a notion of belief.



Both sides work from empiricist and behaviorist presuppositions. Central notions are decision and action.

## Belief vs decision

There is a pragmatic reason for accommodating belief in a statistical methodology.

(Concept)



**Begrippenkader**

**'Gepaste zorg en praktijkvariatie'**

Informatie ten behoeve van de **tweede Invitational Conference**  
op 18 juni 2014 (13:00–17:30) in de Domus Medica te Utrecht

Often a decision is political or ethical, rendering a decision-theoretic treatment infeasible. A statistical procedure should then be able to offer epistemic guidance.

## **Subjective Bayesian statistics**

Belief may sensibly be conceived and reported in a radically probabilistic way. However...

- Scientists encounter practical problems in communicating finicky probability assignments.
- People may not want to commit to probability assignments as an expression of their beliefs.

So there are reasons for accommodating a categorical or at least more coarse-grained notion of belief as well.

## **Earlier approaches**

How to introduce categorical belief states into statistics? Some earlier approaches:

- Kyburg's logic of defeasible commitments to statistical hypotheses.
- Levi's partition-sensitive acceptance rules as add-on to Bayesian analyses.
- Mayo's attempts to epistemicize classical statistics, primarily Neyman-Pearson testing.

Our approach is similar to Levi's approach in the sense that beliefs are rendered context-sensitive.



## **Our approach**

Notice that classical and Bayesian statistics relate to categorical belief in complementary ways.

- Classical statistics relates naturally to categorical judgments but not in the format of belief.
- Bayesian statistics relates naturally to belief but not in the categorical format.

Leitgeb offers a middle ground: a notion of acceptance that maintains a systematic relation to credal sets, i.e., sets of probability assignments.

## 2 Statistics and the stability approach

The stability view offers a systematic non-reductive relation between categorical and probabilistic epistemic attitudes.

### *Stability-based belief*

A probability assignment warrants commitment to a categorical belief in a proposition  $X$  iff for any proposition  $A \in \mathcal{A}$  we have  $P(X|A) > r$ , where  $\mathcal{A}$  is the set of propositions whose negation is not included in the full belief state.

As we have seen, there are several ways of arriving at this belief concept.

## **Properties of stability-based beliefs**

Numerous properties make stability-based belief an attractive option for the application in statistics.

- The belief states have pleasing logical properties, e.g., they respect Hintikka's doxastic logic and Lehmann and Magidor's belief revision theory.
- It has an appealing semantics of Grove spheres, so that the corresponding constraints on probability assignments are easily computed.
- The representation works both ways: a belief state generates a credal set, and is warranted by elements of it.

## **From beliefs to statistics and back**

The stability-based notion of belief offers a systematic way of reconstructing credal sets from full belief states, and conversely.

- A classical statistician who wants to accept  $H_0$  when she learns  $R^0$  but who considers  $R^1$  possible can be represented with a particular prior credal set.
- A Bayesian statistician whose probability assignment falls within a particular credal set is warranted to commit to the set of accepted propositions associated with that credal set.

## **Centrality of the decision maker**

Our proposal is to apply this belief concept to statistics, to offer advice to our decision maker.



The things deemed salient by the decision maker determine the categorical judgments that we offer as advice.

## **Isolating epistemic advice?**

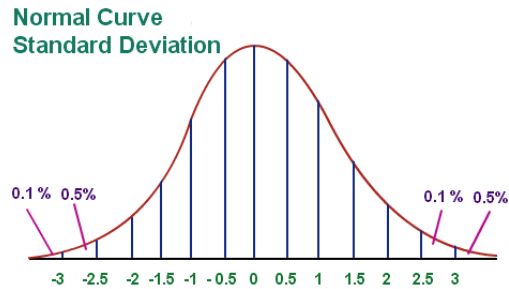
It is debatable whether we will manage to properly isolate the epistemic components of the decision situation.



Against the principled objections, we advance the necessity of epistemic advice as encountered in scientific practice.

## Coarse-grained sample space

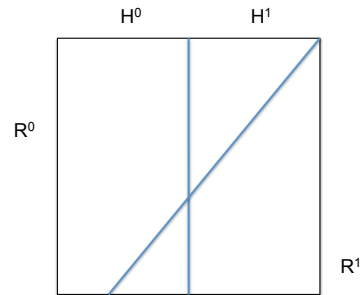
The requirement that potential defeaters are possible, i.e., that their negation is not believed, is particularly salient here.



Sample space needs to be sufficiently coarse-grained. Alternatively, some corners of sample space may be ignored in the analyses.

### 3 Example: Neyman-Pearson tests

Consider a surgeon who asks our opinion on a null hypothesis, and who will subsequently incorporate additional tests in her decision.



Qualitative and quantitative features of the decision situation will allow us to derive non-trivial bounds on error probabilities and thresholds.



## **Exploiting the bridge principles**

Specifically, we may adapt the setup of our statistical tests to the constraints dictated by qualitative considerations concerning outcomes. The credal sets associated with these outcomes determine the constraints.

- What sample size will warrant the report of a full belief, given a decision context?
- Which tests, and hence values of significance and power, are needed for reaching stable beliefs?
- How do the belief states depend on the threshold value  $t$ , and how do our commitments change if we raise the threshold?

### Example on NP-testing

To illustrate, we re-represent Neyman-Pearson testing as an inference. We assume that the test results are simply  $R^1$  and  $R^0$ , and we employ a threshold  $t$ . We have

$$P(R^1|H^0) = \alpha, \quad P(R^1|H^1) = \kappa,$$

$$P(H^0|R^0) \in [l^+, u^+], \quad P(H^0|R^1) \in [l^-, u^-].$$

The requirements that  $l^+ > t$  and  $u^- < t$ , or even  $u^- < 1 - t$ , then determine an interval-valued prior  $P(H^0) \in [l, u]$ .

## **Further constraints**

There are at least two other types of constraint that tie together the credal sets, the thresholds and the test specifics.

- The interval-valued prior is narrowed down further by the fact that the belief must be stable under a designated set of possible further tests.
- We must suppose that the decision maker considers disconfirmation of the null and alternative hypothesis possible. This pushes the bounds on the intervals away from the extremes.

Relative to test specifics and threshold, it may so happen that no interval-valued prior can match all the constraints.

### Example of a conflict

Based on the constraints that evidence  $R^0$  must lead us to accept the null and that evidence  $R^1$  should be possible, we can derive the following:

$$\frac{\kappa t}{\alpha(1-t)} < \frac{P(H^0)}{P(H^1)} < \frac{\kappa + t - 1}{1 - t - \alpha}.$$

Using  $\alpha = a(1-t)$  and  $\kappa = k(1-t)$ , we find a requirement for the existence of an interval-valued prior:

$$t < \frac{a(k-1)}{k-a}.$$

## **Radical context sensitivity**

The foregoing does not yet employ constraints that derive from the requirement of stability under further tests. Adding these will shrink the interval-valued prior further.



Moreover, these constraints are highly sensitive to context, almost ridiculously so. Should we bite the bullet?

## 4 Contextual beliefs

The set of beliefs depends on threshold, test specifics, priors, and crucially, the set of possibilities  $\mathcal{A}$ .



We adapt our report to the surgeon according to what is salient to her decision situation. A Cartesian surgeon might even nullify her full belief set.

## **Contextualism in statistics**

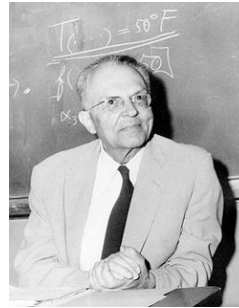
We think it fitting that the reported full beliefs depend on what is deemed salient by the decision maker.

- In classical statistics, acceptance and rejection are relative to the choice of a sample space.
- In Bayesian statistics, acceptance rules are often taken to be relative to a partition of the algebra.

In other words, categorical judgments in statistics are already seen as contextual, or even “holistic”, by many.

## **In terms of “isms”**

The context sensitivity of belief mirrors a widely accepted dependence of decisions on context.



We shed behaviorism and strict empiricism but we retain pragmatism.



## 5 Summing up

We argued for the following.

- The problem of how full belief relates to statistics is in need of a principled solution.
- An attractive solution employs a notion of belief based on stability, as developed by Leitgeb.
- This solution illuminates how full belief is sensitive to the context of the decision maker.

## **Future work**

We will employ this view on statistics and belief in an analysis of various statistical techniques.

- Offering a qualitative understanding of the constraints on reporting hypothesis tests.
- A similar analysis of confidence intervals using the closely related credal intervals.
- Robustness analysis of models via full beliefs based on Bayesian model selection.

# **Thank you**

The slides for this talk will be available at <http://www.philos.rug.nl/romeyn>.  
For comments and questions, email [j.w.romeijn@rug.nl](mailto:j.w.romeijn@rug.nl).