Objective Bayesianism and Unfair Coins

Bert Leuridan

Centre for Logic and Philosophy of Science
Ghent University

Bert.Leuridan@UGent.be

1-4 September 2009
Foundations of Uncertainty
Prague

Outline

1. The virtues of Objective Bayesianism
2. Principles of Objective Bayesianism
3. A Thought Experiment
   - Tossing an Unfair Coin
   - Long Run Degrees of Belief
   - A Dutch Book Argument
4. Suggestions to Solve the Problem

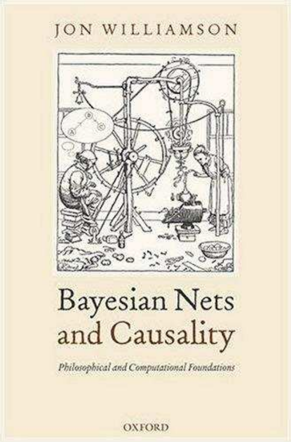
Outline

1. The virtues of Objective Bayesianism
2. Principles of Objective Bayesianism
3. A Thought Experiment
   - Tossing an Unfair Coin
   - Long Run Degrees of Belief
   - A Dutch Book Argument
4. Suggestions to Solve the Problem

Objective Bayesianism (OB)

- probability as *mental* (not physical)
  - → probabilities as an agent's degrees of belief
- yet *objective* (not subjective)
  - → probabilities are not arbitrary, but fixed by an agent's background knowledge. Two agents with the same background knowledge must adopt the same probabilities as their rational degrees of belief.

Williamson (2005): *Bayesian Nets and Causality*



The *philosophical* virtues of OB (Williamson 2005)

Philosophical virtues of OB:

- allows for probabilities over single-case (non-repeatable) variables (contra frequentism)
- provides a mechanism for attributing probabilities (contra chance-interpretation)
- allows for objective (non-arbitrary) probabilities (contra subjective Bayesianism)

Practical/scientific importance of OB

OB has promising scientific applications: breast cancer prognosis

- Background knowledge for breast cancer prognosis is very diverse and complex
  - clinical databases, molecular databases, quantitative data from the literature, . . .
- Problem: how to integrate such diverse sources of knowledge?
- Nagl *et al*. (2006), Nagl *et al*. (in press): with Objective Bayesianism

Outline

1. The virtues of Objective Bayesianism
2. Principles of Objective Bayesianism
3. A Thought Experiment
   - Tossing an Unfair Coin
   - Long Run Degrees of Belief
   - A Dutch Book Argument
4. Suggestions to Solve the Problem

Outline of OB

Intuitively, an agent's degrees of belief should satisfy (Williamson 2005, 65–84):

- **Empirical Constraints:** knowledge about the world ought to constrain degrees of belief
- **Logical Constraints:** lack of information about the world ought to constrain degrees of belief

$\mathbb{P}$ set of all probability functions

$\downarrow$ (empirical information)

$\mathbb{P}_\pi$ set of probability functions satisfying empirical constraints

$\downarrow$ (logical principles)

$P \in \mathbb{P}_\pi$ (normally) one single, 'objective' probability function

A note on notation

$p(u)$: agent's degree of *belief* in u

$p^*(u)$: physical *chance* of u

Empirical Principles of Objective Bayesianism

- **Truth Principle (T):** If an agent knows u to be true then she should have maximum degree of belief in u, $p(u) = 1$

  more generally,

- **Mental-Physical Calibration Principle (MPC1):** If an agent knows the chance $p^*(u)$ of u then she should set her degree of belief in u to that probability, $p(u) = p^*(u)$.

  still more generally,

- **Mental-Physical Calibration Principle (MPC2)**. If an agent knows that $f(p^*_{\downharpoonright U}) \in X$ for $U \subseteq V$ then her belief function p should satisfy the constraint $p_{\downharpoonright U} \in Y$ where Y is the smallest closed convex set of probability functions on U that contains $f^{-1}X$.

(Williamson 2005, 70–73)

Mental-Physical Calibration Principle (MPC2)

P(H)=?

Mental-Physical Calibration Principle (MPC2)

P(H)=?

Mental-Physical Calibration Principle (MPC2)

$P(H) \in [0{,}2\ ,\ 0{,}8]$

Logical Principles of Objective Bayesianism

- **Maximum Entropy Principle (ME):** An agent ought to adopt, out of all probability functions that satisfy the constraints imposed by her background knowledge, a function p that maximizes entropy,

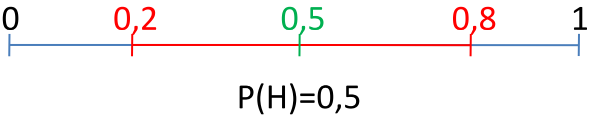
$$H = -\sum_{v@V} p(v) \log p(v)$$

motivation: this probability function is maximally non-committal (or uncertain) to what we do not know

cf. Principle of Indifference

(Williamson 2005, 80)

Maximum Entropy (ME)



Outline

1 The virtues of Objective Bayesianism

2 Principles of Objective Bayesianism

3 A Thought Experiment
- Tossing an Unfair Coin
- Long Run Degrees of Belief
- A Dutch Book Argument

4 Suggestions to Solve the Problem

Introduction

Suppose: unfair coin – unfair in that it is biased to at least an extent δ (but we do not know in which direction):

$$p^*(H) \in X = [0, \frac{1}{2} - \delta] \cup [\frac{1}{2} + \delta, 1]$$

(where $p^*(H)$ is the chance of throwing heads and δ some fixed number)

P(H) ∈ [0, ½-δ] ∪ [½+δ,1]

Question (single case)

- Question: What should your degree of belief $p(H)$ be that on the next toss, this coin will land up heads?
- Objective Bayesianism's answer:

step 1 MPC2: p should lie in the convex hull of $X = [0, \frac{1}{2} - \delta] \cup [\frac{1}{2} + \delta, 1]$,
in this case: $p(H) \in [0, 1]$

step 2 ME: from this convex hull, select the probability function that maximizes entropy,
in this case: $p(H) = \frac{1}{2}$

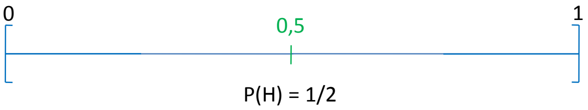
Question (single case)

$P(H) \in [0, \frac{1}{2}-\delta] \cup [\frac{1}{2}+\delta, 1]$

Question (single case)

$P(H) \in [0,1]$

Question (single case)



Question (infinite sequence of tosses)

- Suppose that our unfair coin is tossed infinitely many times.
- Suppose, moreover, that we are never told the outcome of a toss.
- Question: What should, for each toss, the objective bayesian's degree of belief $p(H)$ be?
- Answer: Given that we have no more information than we had in the case of the first toss, $p(H)$ should each time be $\frac{1}{2}$.
  - 1st toss: $p(H) = \frac{1}{2}$
  - 2nd toss: $p(H) = \frac{1}{2}$
  - 3rd toss: $p(H) = \frac{1}{2}$
  - 4th toss: $p(H) = \frac{1}{2}$
  - ... (infinitely many times)

Long run degrees of belief

Goal: to arrive at an inconsistency by applying the weak law of large numbers.

$freq^n = \frac{x}{n}$: the relative frequency of heads in some sequence of n tosses

Theorem (weak law of large numbers)

However small $\epsilon > 0$ is, as n increases, the probability approaches 1 that $freq^n$, the relative frequency of heads in n trials, differs by less than ϵ from its expected value q.

$$\lim_{n \to \infty} p(|freq^n - q| < \epsilon) = 1$$

note: independently and identically distributed (i.i.d.)

Law of large numbers

For every toss, Objective Bayesianism states that $p(H) = \frac{1}{2}$ (*degree of belief*). So the expected value $q = \frac{1}{2}$. It follows that for any $0 < \epsilon$,

$$\lim_{n \to \infty} p(|freq^n - \frac{1}{2}| < \epsilon) = 1 \quad (1)$$

For every toss, the *physical chance* of heads was $p^*(H) \in X = [0, \frac{1}{2} - \delta] \cup [\frac{1}{2} + \delta, 1]$. So the expected value q is some unknown $r \in X$. It follows that for any $0 < \epsilon$,

$$\lim_{n \to \infty} p^*(|freq^n - r| < \epsilon) = 1 \quad (2)$$

$$\Downarrow \text{(MPC1)}$$

$$\lim_{n \to \infty} p(|freq^n - r| < \epsilon) = 1 \quad (3)$$

Law of large numbers

For every toss, Objective Bayesianism states that $p(H) = \frac{1}{2}$ (*degree of belief*). So the expected value $q = \frac{1}{2}$. It follows that for any $0 < \epsilon$,

$$\lim_{n\to\infty} p(|freq^n - \frac{1}{2}| < \epsilon) = 1 \quad (1)$$

For every toss, the *physical chance* of heads was $p^*(H) \in X = [0, \frac{1}{2} - \delta] \cup [\frac{1}{2} + \delta, 1]$. So the expected value q is some unknown $r \in X$. It follows that for any $0 < \epsilon$,

$$\lim_{n\to\infty} p^*(|freq^n - r| < \epsilon) = 1 \quad (2)$$

$\Downarrow$ (MPC1)

$$\lim_{n\to\infty} p(|freq^n - r| < \epsilon) = 1 \quad (3)$$

Law of large numbers

For every toss, Objective Bayesianism states that $p(H) = \frac{1}{2}$ (*degree of belief*). So the expected value $q = \frac{1}{2}$. It follows that for any $0 < \epsilon$,

$$\lim_{n\to\infty} p(|freq^n - \frac{1}{2}| < \epsilon) = 1 \quad (1)$$

For every toss, the *physical chance* of heads was $p^*(H) \in X = [0, \frac{1}{2} - \delta] \cup [\frac{1}{2} + \delta, 1]$. So the expected value q is some unknown $r \in X$. It follows that for any $0 < \epsilon$,

$$\lim_{n\to\infty} p^*(|freq^n - r| < \epsilon) = 1 \quad (2)$$

$$\Downarrow \text{(MPC1)}$$

$$\lim_{n\to\infty} p(|freq^n - r| < \epsilon) = 1 \quad (3)$$

Law of large numbers

For every toss, Objective Bayesianism states that $p(H) = \frac{1}{2}$ (*degree of belief*). So the expected value $q = \frac{1}{2}$. It follows that for any $0 < \epsilon$,

$$\lim_{n\to\infty} p(|freq^n - \frac{1}{2}| < \epsilon) = 1 \tag{1}$$

For every toss, the *physical chance* of heads was $p^*(H) \in X = [0, \frac{1}{2} - \delta] \cup [\frac{1}{2} + \delta, 1]$. So the expected value q is some unknown $r \in X$. It follows that for any $0 < \epsilon$,

$$\lim_{n\to\infty} p^*(|freq^n - r| < \epsilon) = 1 \tag{2}$$

$$\Downarrow$$

$$\lim_{n\to\infty} p(|freq^n - r| < \epsilon) = 1 \tag{3}$$

Equations 1 and 3 are INCONSISTENT!!!

(Inconsistency)

- If $r \in [0, \frac{1}{2} - \delta]$, then for any $freq^n$, $|freq^n - \frac{1}{2}| < |freq^n - r|$.
  Hence there is an ϵ such that
  $$\lim_{n\to\infty} p(|freq^n - \frac{1}{2}| < \epsilon < |freq^n - r|) = 1 \tag{4}$$
  But this contradicts equation 3.

- If $r \in [\frac{1}{2} + \delta, 1]$, then for any $freq^n$, $|freq^n - r| < |freq^n - \frac{1}{2}|$.
  Hence there is an ϵ such that
  $$\lim_{n\to\infty} p(|freq^n - r| < \epsilon < |freq^n - \frac{1}{2}|) = 1 \tag{5}$$
  But this contradicts equation 1.

Dutch Book Argument

This inconsistency leads to a Dutch Book
$\rightarrow$ bets on outcomes of *finite* sequences of coin tosses (outcomes in principle verifiable)

From infinite to finite sequences of tosses

observation: for all $\epsilon > 0$ there is a finite n_ϵ such that for all $n \geq n_\epsilon$,

$$p(|freq^n - \frac{1}{2}| > \epsilon) \leq \epsilon \quad (6)$$

observation: for all $\epsilon > 0$ there is a finite n'_ϵ such that for all $n \geq n'_\epsilon$,

$$p(|freq^n - r| > \epsilon) \leq \epsilon \quad (7)$$

note: these observations follow from equations (1) and (3)
note: we can in principle calculate the numbers n_ϵ and n'_ϵ

From infinite to finite sequences of tosses

observation: for all $\epsilon > 0$ there is a finite n_ϵ such that for all $n \geq n_\epsilon$,

$$p(|freq^n - \frac{1}{2}| > \epsilon) \leq \epsilon \tag{6}$$

observation: for all $\epsilon > 0$ there is a finite n'_ϵ such that for all $n \geq n'_\epsilon$,

$$p(|freq^n - r| > \epsilon) \leq \epsilon \tag{7}$$

note: these observations follow from equations (1) and (3)
note: we can in principle calculate the numbers n_ϵ and n'_ϵ

From infinite to finite sequences of tosses

Given these observations (and the strong law of large numbers), we know that for all $0 < \epsilon$ it is the case that for all $n \geq \max(n_\epsilon, n'_\epsilon)$,

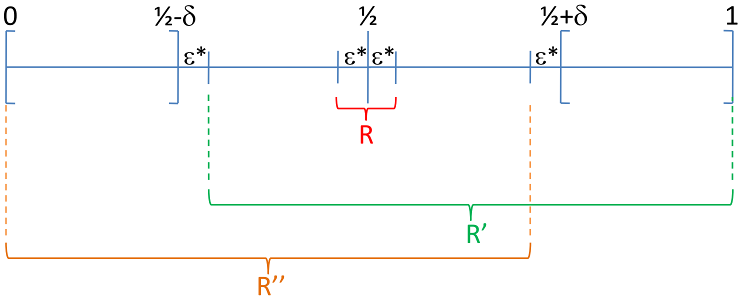
$$p(|freq^n - \frac{1}{2}| > \epsilon) \leq \epsilon \text{ and} \tag{8}$$

$$p(|freq^n - r| > \epsilon) \leq \epsilon \tag{9}$$

note: these are the objective bayesian's degrees of belief

Particular finite sequence of tosses

Choose some fixed $\epsilon^* < \frac{\delta}{2}$ and ask the objective bayesian to bet on the value of $freq^{n^*}$, where $n^* = \max(n_{\epsilon^*}, n'_{\epsilon^*})$.



Betting quotients for bets on $freq^{n^*}$

Given that the inequalities (8) and (9) hold for all ϵ, we may derive that for ϵ^* and $freq^{n^*}$ $(n^* = \max(n_{\epsilon^*}, n'_{\epsilon^*}))$:

$$p(|freq^{n^*} - \frac{1}{2}| > \epsilon^*) \leq \epsilon^* \text{ and } \quad (10)$$

$$p(|freq^{n^*} - r| > \epsilon^*) \leq \epsilon^* \quad (11)$$

RECALL: these are the objective bayesian's degrees of belief

↓

BETTING QUOTIENTS

Betting quotients for bets on $freq^{n^*}$

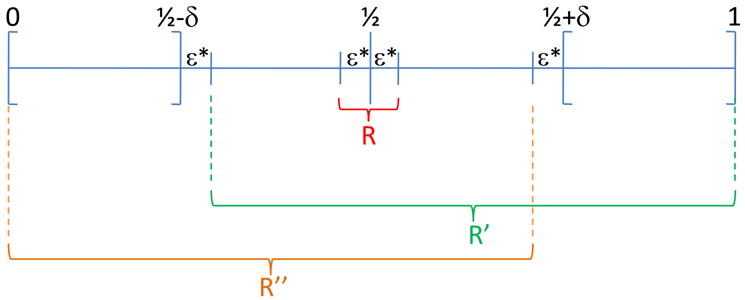
- Given (10), the objective bayesian is prepared to bet that $freq^{n^*} \in R$ with betting quotient $1 - \epsilon^*$.

  That is, she is prepared to pay $(1 - \epsilon^*)Q$ if she would win Q in case $|freq^{n^*} - \frac{1}{2}| < \epsilon^*$.

- Given (11), she is also prepared to bet that $freq^{n^*} \notin R$ with betting quotient $1 - \epsilon^*$.

  That is, she is prepared to pay $(1 - \epsilon^*)Q$ if she would win Q in case $|freq^{n^*} - \frac{1}{2}| > \epsilon^*$.

Betting quotients for bets on $freq^{n^*}$



(Justification for second betting quotient)

- Either $r \in [0, \frac{1}{2} - \delta]$ or $r \in [\frac{1}{2} + \delta, 1]$.
  - In the first case, $p(freq^{n^*} \in R') \leq \epsilon^*$ and a fortiori $p(freq^{n^*} \in R) \leq \epsilon^*$ (since $R \subset R'$).
  - Analogously, in the second case, $p(freq^{n^*} \in R'') \leq \epsilon^*$ and a fortiori $p(freq^{n^*} \in R) \leq \epsilon^*$ (since $R \subset R''$).
- Hence by dilemma, $p(freq^{n^*} \in R) \leq \epsilon^*$.
- Thus $p(freq^{n^*} \notin R) \geq 1 - \epsilon^*$.

Dutch Book

But these bets together form a Dutch Book:

	$freq^{n^*} \in R$	$freq^{n^*} \notin R$
bet on $freq^{n^*} \in R$ for $(1-\epsilon^*)Q$	$Q-(1-\epsilon^*)Q$	$-(1-\epsilon^*)Q$
bet on $freq^{n^*} \notin R$ for $(1-\epsilon^*)Q$	$-(1-\epsilon^*)Q$	$Q-(1-\epsilon^*)Q$
	$(2\epsilon^*-1)Q$	$(2\epsilon^*-1)Q$

No matter what the outcome of the experiment is, the objective bayesian wins $(2\epsilon^*-1)Q$. But given that $2\epsilon^* < \delta < 1$, this is a sure loss.

Outline

1. The virtues of Objective Bayesianism
2. Principles of Objective Bayesianism
3. A Thought Experiment
   - Tossing an Unfair Coin
   - Long Run Degrees of Belief
   - A Dutch Book Argument
4. Suggestions to Solve the Problem

Suggestions to solve the problem

- restrict the scope of application of Objective Bayesianism:
  - no single case objective degrees of belief if outcome of previous tosses is unknown
  - justified in practice,
  - but an ad hoc solution in general
- hierarchical model
  - second order probabilities $P(p^*(H) = x)$ for all $x \in X$
  - applying OB to second order probabilities
  - problem: objectivity not guaranteed for infinite domains
- imprecise probabilities
  - use credal sets (= imprecise probabilities)
  - instead of a single probability function (= precise)
  - problem: computationally harder than Bayesian theory
- change the theory of Objective Bayesianism:
  - inconsistency handling mechanism
  - prioritizing degrees of belief

Suggestions to solve the problem

- restrict the scope of application of Objective Bayesianism:
  - no single case objective degrees of belief if outcome of previous tosses is unknown
  - justified in practice,
  - but an ad hoc solution in general
- hierarchical model
  - second order probabilities $P(p^*(H) = x)$ for all $x \in X$
  - applying OB to second order probabilities
  - problem: objectivity not guaranteed for infinite domains
- imprecise probabilities
  - use credal sets (= imprecise probabilities)
  - instead of a single probability function (= precise)
  - problem: computationally harder than Bayesian theory
- change the theory of Objective Bayesianism:
  - inconsistency handling mechanism
  - prioritizing degrees of belief

Suggestions to solve the problem

- restrict the scope of application of Objective Bayesianism:
  - no single case objective degrees of belief if outcome of previous tosses is unknown
  - justified in practice,
  - but an ad hoc solution in general
- hierarchical model
  - second order probabilities $P(p^*(H) = x)$ for all $x \in X$
  - applying OB to second order probabilities
  - problem: objectivity not guaranteed for infinite domains
- imprecise probabilities
  - use credal sets (= imprecise probabilities)
  - instead of a single probability function (= precise)
  - problem: computationally harder than Bayesian theory
- change the theory of Objective Bayesianism:
  - inconsistency handling mechanism
  - prioritizing degrees of belief

Suggestions to solve the problem

- restrict the scope of application of Objective Bayesianism:
  - no single case objective degrees of belief if outcome of previous tosses is unknown
  - justified in practice,
  - but an ad hoc solution in general
- hierarchical model
  - second order probabilities $P(p^*(H) = x)$ for all $x \in X$
  - applying OB to second order probabilities
  - problem: objectivity not guaranteed for infinite domains
- imprecise probabilities
  - use credal sets (= imprecise probabilities)
  - instead of a single probability function (= precise)
  - problem: computationally harder than Bayesian theory
- change the theory of Objective Bayesianism:
  - inconsistency handling mechanism
  - prioritizing degrees of belief

Prioritized Objective Bayesianism

- which rules are to be blamed?
  - T, MPC1: within the context, background knowledge should not be doubted
  - MPC2: addition of probability functions not directly warranted by available evidence
  - ME: at best a ‘best guess’, certainly not infallible
- basic idea:
  - do not dispense with the rules MPC2 and ME
  - only dispense with fallacious applications
- general framework
  - prioritized adaptive logics
  - cf. prioritized Rescher-Manor inconsistency handling mechanisms

Conclusion

- Objective Bayesianism is interesting (philosophical reasons, scientific reasons)
- But it leads to inconsistency / incoherent degrees of belief / Dutch Book
- And hence should be adjusted

References

- P. Billingsley (1995). *Probability and Measure*. New York, John Wiley and Sons, 3rd edition.
- W. Feller (1950, 1968). *An Introduction to Probability Theory and Its Applications,* Volume I, 3rd edition. New York, Wiley.
- W. Feller (1966). *An Introduction to Probability Theory and Its Applications,* Volume II. Chichester, Wiley.
- C. Howson and Peter Urbach (1993). *Scientific Reasoning. The Bayesian Approach*. Chicago and La Salle, Illinois, Open Court, 2nd edition.
- S. Nagl *et al*. (2006). "Objective Bayesian nets for integrating cancer knowledge: a systems biology approach." In: J. Rousu *et al*. (eds.), *Proceedings of the Workshop on Probabilistic Modeling and Machine Learning in Structural and Systems Biology*, Helsinki, Helsinki University Printing House, pp. 44–49.

References

- S. Nagl *et al.* (in press). “Objective Bayesian nets for systems modelling and prognosis in breast cancer.” In: D. Holmes and L.C. Jain (eds.), *Innovations in Bayesian Networks: Theory and Applications*, Springer, in press.
- H. Reichenbach (1949). *The Theory of Probability. An Inquiry into the Logical and Mathematical Foundations of the Calculus of Probability.* Berkeley and Los Angeles, University of California Press.
- R. von Mises (1964). *Mathematical Theory of Probability and Statistics*. New York, Academic Press.
- J. Williamson (2005). *Bayesian Nets and Causality. Philosophical and Computational Foundations*. New York, Oxford University Press.