university of groningen

NWO
Netherlands Organisation for Scientific Research

Philosophy Colloquium
Salzburg 2015

What are the chances?

or: How to do Things with Frequentism

★

Jan-Willem Romeijn
University of Groningen

Why care?

Chances play a central role in the cognitive, social, and life sciences.


(Concept)

Zorginstituut Nederland

ZonMw

NFU

Begrippenkader

‘Gepaste zorg en praktijkvariatie’

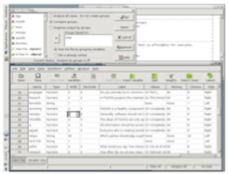
Informatie ten behoeve van de **tweede Invitational Conference**

op **18 juni 2014 (13:00–17:30) in de Domus Medica** te Utrecht

Surely the debates on disease models and health policy are about *something*.

The easy answer

Chances are described mathematically by probability distributions. They are theoretical terms in scientific theory.

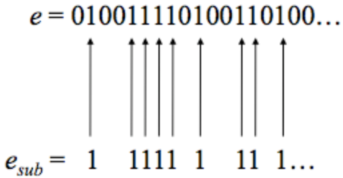


But how are they related to scientific data? And can we provide a realist interpretation for them?

Contents

1 Von Mises' frequentism

An important starting point in the discussion on chance is the frequentist definition of probability by von Mises.



It is based on the notion of a *collective*: a binary series that has a limiting relative frequency and for which there is no gambling system.

Problems with frequentism

Frequentism has been severely criticised, primarily for being unempirical.

- ↯ The theory requires infinitely long sequences of events.
- ↯ The randomness of sequences depends on the observer.

In response, we might develop a finite version of frequentism.

- ↯ Events that do not occur may still have a definite probability.
- ↯ This theory introduces biases and spurious correlations.

I do not claim that von Mises cannot defend his theory against all these criticisms. Despite that. . .

Alternative frequentisms

In this talk I employ frequentist ideas in the semantics and metaphysics of chance.

Semantic frequentism provides a formal semantics for statistical inference by specifying the nature of statistical hypotheses.

Metaphysical frequentism fosters an interpretation of single-case chances that escapes the reference class problem.

Importantly, frequentism is thereby detached from strict empiricism.

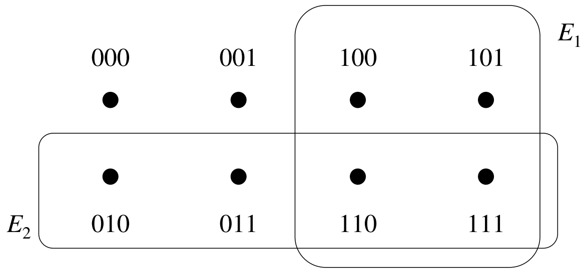
2 Statistical hypotheses

Gaifman and Snir have become famous for the convergence results: the priors will always wash out.

I will however focus on their wonderful idea of rich languages: they show how to express statistical hypotheses in a space of possible observations.

Sample space

A statistical analysis is always based on a set of possible observations, a sample space. For tossing a coin N times, the sample space is $\{0, 1\}^N$.



Samples, written in lowercase as e_t, can be represented as sets E_t in this space.

Hypotheses as distributions

We may also construct an idealised sample space consisting of infinitely long samples: $\{0,1\}^{\Omega}$.

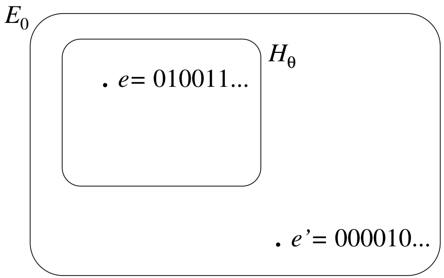
$$e = 010011011001010\ldots$$

$$P_{\theta}(E_t|E_{t'}) = \theta$$

The statistical hypothesis is a distribution over this infinite sample space, written h_{θ}.

...and as events

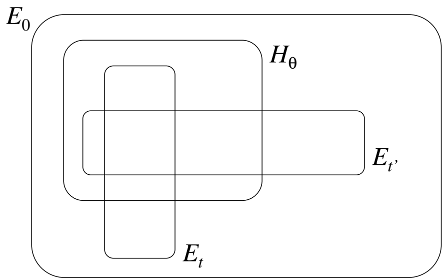
Note that some elements $e = 00110111\ldots$ of this sample space are collectives in the sense of von Mises.



We can identify the statistical hypothesis h_θ with the set of all collectives that instantiate the probability distribution P_θ.

Events as distributions

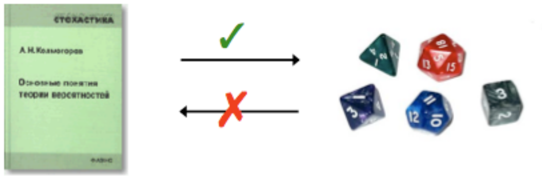
The distribution P_θ of hypothesis h_θ is thus associated with a particular set H_θ in the sample space.



Each set H_θ intersects with every observation E_t that is assigned some probability by P_θ.

Reversed frequentism

Von Mises presented frequentism as a theory on what probability is, grounding it in empirical phenomena.



Instead, I presuppose a notion of probability and use frequentism to relate it to a model of empirical fact.

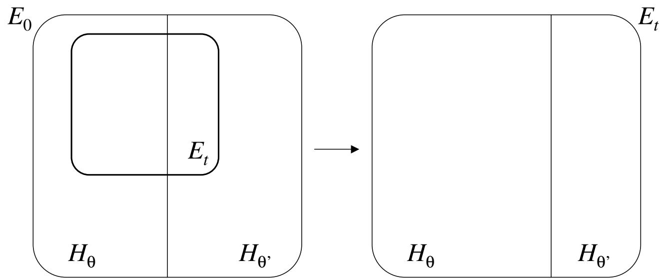
3 Frequentist Bayesian inference

You can be both a Bayesian and a frequentist, much in line with Jeffrey's mixed Bayesianism.

The association of hypotheses and events offers a fresh perspective on statistical inference.

The probability of a distribution?

It becomes natural to assign a probability to a probability distribution, $P(H_\theta)$.



Consequently Bayesian conditionalisation can be located in sample space, as a “zooming in” operation.

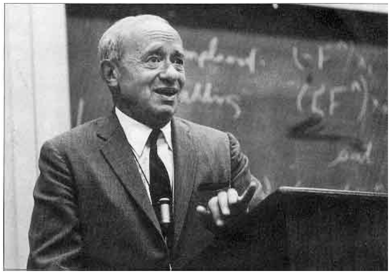
Hypotheses fix inductive dependence

Carnapian inductive logic fixes inductive relations between observations by constraints on the probability over sample space.

Statistical hypotheses provide an alternative way of fixing these constraints. They extend the language of inductive logic, for free!

Formal semantics for statistics

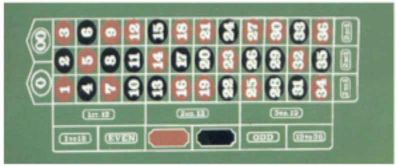
The conceptual clarity of deductive logic is partly due to a clear separation of syntax and semantics.



The semantics of statistical hypotheses may be a step towards clearing up inductive logic, in particular statistical inference.

4 The reference class problem

We have located chances within the epistemic domain, as components of a formal semantics for statistical inference.



What about chances as pertaining to physical events? Can we give chance a realist interpretation?

Reference classes

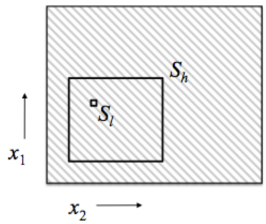
A central problem for physical probability is the reference class problem: different descriptions of an event lead to different chances.



Chances thereby become description-relative, and thus partly epistemic.

Deterministic systems

A special case of this problem occurs for deterministic systems: a complete description of such systems trivializes their chances.



Chances thus become entirely epistemic: they are determined by a lack of information on deterministic states.

Perspectival chance

We might stipulate (cf. Glynn and Strevens) that the correct chances are those that follow from our preferred theory or our explanatory ideals.

- ↯ Chances become tied up with our theoretical perspective.

One alternative is that we reduce all chances to the theory that is most fundamental, namely quantum mechanics.

- ↯ Those chances do not relate to events that are intuitively chancy.

I will try and provide an objective motivation for choosing a theoretical perspective that features chance.

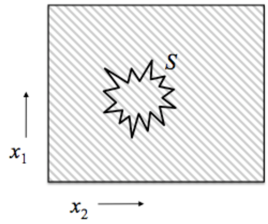
5 Irreducible chances

We can use ideas from the debate on reductionism and emergentism to establish chances at a level of description.

The key concept is a radical form of multiple realizability: some concepts principally resist translation to a different theoretical level.

Multiply realizable states

Sometimes the macro-level description of a physical state, denoted $\mathcal{S}$, cannot be defined in terms of sets of micro-level states, $\mathcal{X}$.



Even stronger: the micro-level descriptions might not offer the conceptual means to refine or alter macro-level probabilities. In such cases we call the events *random* relative to the micro-level descriptions.

Random events

An event S is random relative to some algebra $\mathcal{X}$ iff for all refinements X_i from this algebra we have

$$P(S|R \cap X_i) = P(S|R).$$

This echoes ideas from Lewis and Skyrms on the robustness of chance: adding further information from $\mathcal{X}$ to a reference class R does not refine or alter the chances for S. Here robustness rests on the intricacy of the events S and R in relation to the refinements of $\mathcal{X}$.

6 Defining robust chances

We can provide a formal underpinning of randomness and robustness that relies on ideas from frequentism. Recall the definition of the limiting relative frequency F of a series s:

$$F(s) = \lim_{n\to\infty} \frac{\Sigma_{t=1}^{n} s(t)}{n},$$

where $s(t)$ is the digit at position t in the series. Note: the denotation of the series by s, while uppercase S is a set, is not a coincidence.

Place selections

We can use a second series x_i to select elements from the series x, and construct the relative frequency of a subseries:

$$F(s; x_i) = \lim_{n \to \infty} \frac{\sum_{t=1}^{n} s(t) x_i(t)}{\sum_{t=1}^{n} x_i(t)}.$$

The series s is random relative to a set $\mathcal{X}$ of series or selection rules if for all i we have:

$$F(s) = F(s; x_i).$$

We might say that the set of series $\mathcal{X}$ contains "no information" about the original series s.

Series and place selections as sets

Series and place selections can be viewed as sets of natural numbers. For binary sequences we have:

$$S_s = \{t \mid t \in \mathbb{N} \text{ and } s(t) = 1\}.$$

Similar relations between a series s and a set S can be given when we consider S to be a set in a continuous space. This is easiest if S is a countable set of points in the space.

Randomness and robustness

We can now apply the same notion of invariance under place selection in the richer context of events: the set S is random relative to a lower-level algebra $\mathcal{X}$ iff for all X_i we have

$$P(S|X_i) = P(S).$$

We thus employ the mathematical machinery of place selections to arrive at a notion of randomness for an event S, and hence at the robustness of the probability assignment $P(S)$.

Infinite intricacy

The ultimate version of robustness has the events S distributed uniformly over the space X. Consider the σ-algebra $\mathcal{X}$ generated by all sets of the form

$$B[x,\delta] = \{x' : x' \in (x, x+\delta)\}$$

where the δ can be arbitrarily small, and stipulate that

$$P(S|B[x,\delta]) = P(S).$$

That is, no amount of fine-graining will offer additional information on the event S, making the chances fully robust or *autonomous*.

Random set: example

We can construct a random set by employing a simple ergodic dynamical system. Consider a set R with elements indexed by t,

$$R(t) = 2^t R(0) \pmod 1.$$

where $R(0) \in [0,1]$ is the starting position generating the set. For a given initial state $R(0)$, we can label all the points $R(t)$ by 0 if $R(t-1) \leq 1/2$ and by 1 if $R(t-1) > 1/2$. The set of points labelled 1 we call S.

Autonomous chance: example

Assuming that $R(0)$ is a collective under a certain notion of randomness and with relative frequency σ of 1's, the limiting relative frequency of points within $R \cap S$ is

$$P(S|R) = \sigma.$$

Notice that the countable set R of points in $[0,1]$ is dense everywhere. Moreover, for any set X_i selecting out an interval within $[0,1]$, however small, we have

$$P(S|R \cap X_i) = P(S|R).$$

Objective chance

My proposal is to call a chance ascription $P(S|R)$ objective if it is suitably close to being autonomous in the above sense. Some remarks:

- There are many variations on the randomness of the event S depending on the details of the algebra $\mathcal{X}$. This runs parallel to the randomness of sequences.
- A natural line is drawn by the algebra $\mathcal{X}_S$ that corresponds to “Schnorr randomness”: whether or not a point x is a member of an element $X \in \mathcal{X}$ must be computable.
- This leaves room for the random event S to be the result of an effective procedure.

Objective chance?

There are many loose ends in this picture of objective chances on the macro-level. Some considerations:

- Every point x in the space X, or in the class C, is labelled as either $f_S(x) = 0$, or $f_S(x) = 1$. We need not violate determinism.
- Nothing guarantees that random events like S or C actually obtain. The foregoing offers an extreme case of objective chance but the reality of chances may fall short of this.
- The identification of a macro-level that features chances is a semi-objective matter. It does not hinge fully on our theoretical perspective.

7 Conclusion

I have argued that the frequentist theory of chance can be used to our advantage in two separate philosophical projects.

- It elucidates statistical inference by specifying the semantics of statistical hypotheses and thereby clarifying statistical inference.

- It fosters an interpretation of single-case chance that escapes the reference class problem. The autonomy of chance is an objective matter.

- This may explain what the ultimate aim of statistical model selection is: to find autonomous chances.

Ironically, frequentism was motivated by strict empiricism but finds promising applications in metaphysics and semantics.

Thanks!

This talk will be available at http://www.philos.rug.nl/~romeyn. For comments and questions, email j.w.romeijn@rug.nl.