Von Mises, Church, and the Birth of Algorithmic Randomness

Christopher P. Porter
Université Paris 7
LIAFA

HaPoC 2013
ENS, Paris
29 October 2013

Introduction

The general goal of my talk is to highlight certain circumstances under which computability theory was first applied to the task of defining randomness.

The computability-theoretic study of randomness is nowadays known as *algorithmic randomness*, an area of research that has been very active as a sub-branch of computability theory since the late 1990s.

The main figures we will discuss are

- the applied mathematician Richard von Mises, and
- the logician Alonzo Church.

Key points

The three main points that I want to emphasize today are:

1. the role of randomness in von Mises' theory of probability;
2. Church's computability-theoretic modification of von Mises' definition of randomness; and
3. Church's view that this computable modification could serve as a foundation for von Mises' theory of probability.

Von Mises' definition of random sequence

Introducing von Mises

Richard von Mises is considered to be one of the founding fathers of algorithmic randomness (although, strictly speaking, his definition was not algorithmic).

In 1919, von Mises published a definition of random sequences, which served as the basis of his theory of probability.

There have been a number of misconceptions about the actual role that von Mises' definition played in his larger theory of probability.

Von Mises' theory of probability

Von Mises held that an scientifically adequate theory of probability is one according to which the probability of a given event is the limiting relative frequency of that event in a sequence of relevant events.

However, in von Mises' view, not just any sequences were appropriate for determining the probability of an event.

Such a sequence had to be sufficiently *random*.

He also formulated his theory in terms of *infinite* sequences of events.

Thus, for his theory of probability to get off the ground, he needed a definition of random infinite sequences.

Von Mises' definition of randomness

Random sequences, which von Mises referred to as *collectives*, are those sequences that satisfy two axioms:

- The first axiom guarantees that the limiting relative frequencies in collectives exist.
- The second axiom guarantees that these limiting relative frequencies are invariant under "admissibly selected subsequences".

Von Mises referred to his second axiom as *the principle of the impossibility of a gambling system*.

Limiting relative frequencies

For simplicity, let's restrict our attention to sequences in 2^ω.

Given $X \in 2^\omega$, the relative frequency of the occurrence of 1 in the first n values of X is

$$\frac{\#\{i < n : X(i) = 1\}}{n}.$$

The limiting relative frequency of the occurrence of 1 in X is

$$\lim_{n \to \infty} \frac{\#\{i < n : X(i) = 1\}}{n}.$$

Von Mises' first axiom

Axiom 1: The limiting relative frequency

$$\lim_{n \to \infty} \frac{\#\{i < n : X(i) = 1\}}{n}$$

of $X \in 2^\omega$ exists.

Admissibly selected subsequences

For a fixed sequence $X \in 2^\omega$, the sequence $Y \in 2^\omega$ is *admissibly selected* from X if Y is a subsequence of X that is selected from X by means of an *admissible place selection*.

A *place selection* $S : 2^{<\omega} \to \{0, 1\}$ is a map that determines whether or not we are to include a given element of our sequence in the selected subsequence.

A place selection is *admissible* if the choice to select a given element from a sequence is independent of the value of that element.

Given an admissible place selection S and a sequence $X \in 2^\omega$, X_S will denote the subsequence selected from X by S.

Some examples

Admissible place selections:

- Select every odd-indexed place.
- Select every place that follows a 0.
- Select every odd-indexed place that follows a 0.

A non-admissible place selection:

- Select every place that contains a 1.

Von Mises' second axiom

Axiom 2: If

$$\lim_{n\to\infty} \frac{\#\{i < n : X(i) = 1\}}{n} = p$$

for some $p \in [0,1]$, then for any admissible place selection S, we have

$$\lim_{n\to\infty} \frac{\#\{i < n : X_S(i) = 1\}}{n} = p.$$

Why is randomness required?

Without the von Mises' second axiom of randomness, we cannot guarantee that certain calculations in the probability calculus can be carried out.

For example, if we hold that probabilities are limiting relative frequencies, and we accept the product rule for probabilities, we can prove that any sequence that yields probabilities satisfying the product rule must be invariant under

- the place selection that selects at the odd-indexed places,
- the place selection that selects at the even-indexed places, and
- the place selection that selects at the even-indexed places that are preceded by the occurrence of 1.

A serious problem

Von Mises' contemporaries raised the following question:

Which place selections are the admissible ones?

In von Mises' original formulation, he appears to allow every place selection to be counted as admissible.

This is a legitimate worry, since for each $X \in 2^\omega$ that contains infinitely many 0s, there is some place selection S such that $X_S = 0^\omega$.

How do we rule out such place selections as inadmissible?

Wald's proof of the consistency of collectives

In response to von Mises' critics, in 1937, the statistician Abraham Wald gave what he referred to as a proof of the consistency of collectives.

Specifically, Wald proved that for any *countable* collection C of place selections, there are continuum many sequences that are invariant under every place selection in C.

Von Mises' contextual approach

Inspired by Wald's result, von Mises defined collectives *contextually*.

The collection of place selections $\mathscr{S}$ in von Mises' second axiom is determined by the problem of probability that we are trying to solve:

> *We obtain a concrete idea of the set G of place selections which are supposed not to change the frequency limits if we visualize G, for example, as follows: in G are contained all those place selections which present themselves in a particular problem under consideration. (Mathematical Theory of Probability and Statistics, pg. 12)*

An ideal of completeness

Why did von Mises take a contextual approach rather than working with one fixed collection of place selections?

> *It is not possible to build a theory of probability on the assumption that the limiting values of the relative frequencies should remain unchanged only for a certain group of place selections, predetermined once and for all. (Probability, Statistics, and Truth, pg. 91)*

Von Mises wanted a definition of probability by means of which he could solve *all* problems of the probability calculus, an ideal I call the *resolutory ideal of completeness*.

Moreover, von Mises held that restricting his definition to some fixed collection of place selections would prevent him from attaining this ideal.

Church's modification of von Mises' definition

Church's contribution

In his 1940 paper, "On the concept of a random sequence," Church suggested a modification of von Mises' definition.

In order to make von Mises' definition well-defined, Church suggested that we take the admissible place selections to be the *computable* ones.

That is, every admissible place selection is given by a computable function $f : 2^{<\omega} \to \{0, 1\}$, and every computable function of this type determines an admissible place selection.

The rationale behind Church's suggestion

Why make this restriction?

> *To a player who would beat the wheel at roulette a system is unusable which corresponds to a mathematical function known to exist but not given by explicit definition; and even the explicit definition is of no use unless it provides a means of calculating the particular values of the function. ("The Concept of a Random Sequence", pg. 133)*

Church's definition, the first definition of an algorithmically random sequence, can thus be seen as one of the earliest applications of the Church-Turing thesis.

Rejecting the resolutory ideal?

Earlier we saw that von Mises' held that any theory of probability defined in terms of one fixed collection of place selections would be an incomplete theory.

Thus one might suspect that Church, in restricting the admissible place selections to the computable place selections, was either unaware of von Mises' intention or explicitly rejected it.

On the contrary, like von Mises, Church was motivated by producing a definition of probability that attained the resolutory ideal of completeness.

Church and the resolutory ideal of completeness, I

Concerning previous attempts to make von Mises' definition precise by restricting the class of place selections, Church writes,

> *Their use [i.e the use of other definitions of randomness] for this purpose, however, is open to certain objections from the point of view of completeness of the theory, as has been forcibly urged by von Mises, and it is therefore desirable to consider further the question of finding a satisfactory form for the definition of a random sequence. ("The Concept of a Random Sequence", pg. 133)*

Church and the resolutory ideal of completeness, II

Further, Church recognized that defining collectives in terms of computable place selection rules yields the result that the existence proof of a collective is necessarily non-constructive.

However, Church did not consider a weaker collection of place selections, such as the primitive recursive place selections, for he explicitly states that the resulting definition would not attain the resolutory ideal.

But how did Church address the worry that his restricted definition would yield an incomplete theory of probability?

A source of further insight

We find a number of helpful insights in the 1966 correspondence between Church and Hilda Geiringer, von Mises' wife, who edited his 1946 Harvard lecture notes and the 3rd edition of von Mises' book *Probability, Statistics, and Truth*.

Church's hypothesis

In Church's first letter to Geiringer, he explicitly states the hypothesis that von Mises would have given defined randomness in terms of the computable place selections if the definition had been available to him.

> *[I]t seems to me very plausible to say (though of course no proof of such a proposition can be offered) that the definition of "collective" which results from the approach of this paper is the one which von Mises in some sense actually intended when he wrote in 1931, but that it was impossible for him to make the definition in this way because at that date the precise mathematical definition of effective calculability did not yet exist (Letter to Geiringer, March 16, 1966)*

Is Church's definition too restrictive?

In her second letter to Church, Geiringer questions whether Church's definition is more restrictive than an alternative formulated by Wald.

Church responds that a "good reconstitution of the notion of collective" need not "cover all selection that anybody has ever made or claimed to make in a probability problem or probability proof."

So now it appears that Church is conceding that a theory of probability founded on his definition is an incomplete theory.

Church on the prospects of his definition

However, Church held that even such an incomplete account would suffice as long as it had "signficant gaps in neither in the internal logical structure of the theory nor in its applications."

He adds,

> *I have every reasonable expectation that the criterion in my paper results in a class of selections for which this is true. (Letter to Geiringer, June 11, 1966)*

A striking difference with von Mises

Church held that his definition would yield a theory of probability with no "significant gaps". In his view, the remaining gaps are in a certain sense irrelevant.

> *And I would hold that if it is true that no such calculation procedure exists [to implement a given place selection], then the indicated method of selection is an unreasonable one to use in any probability problem. (Has any one ever used it in a probability problem? I don't know, but I would think it unlikely.) (Letter to Geiringer, June 11, 1966)*

A restricted ideal of completeness

Thus, in Church's view, the theory of probability based on his definition of randomness only satisfies a restricted version of the resolutory ideal of completeness.

Any problem of the probability calculus the solution of which requires a non-effective place selection is *not* the sort of problem one encounters in actual practice.

It is somewhat puzzling for Church to take this approach.

Church is fully aware of the significance of problems that cannot be effectively solved (after all, he proved the undecidability of the *Entscheidungsproblem*).

Why is it appropriate here to ignore these sorts of problems?

Several questions

There are a number of questions to ask of Church's approach:

- Should we restrict our attention to problems that can only be solved effectively?
- Why are problems that are effectively solvable privileged over problems that are not?
- Is there anything lost by ignoring those problems that cannot be solved effectively?

A trade-off

Church's introduction of computability in this setting thus yields a trade-off:

- On the one hand, Church's definition yields a restricted theory of probability, and it is not clear that such a theory is complete.
- On the other hand, one gains *uniformity*, as one need not vary the underlying definition of randomness from problem to problem.

  Rather, on Church's approach there is only one fixed collection of place selections to which we must appeal to solve every problem that we encounter in practice.

Thank you for your attention!