

ON THE NEED TO ADAPT DE FINETTI'S PROBABILITY INTERPRETATION TO QM

GÉRARD G. EMCH

*Department of Mathematics, University of Florida
Gainesville, Florida 32611-8105, U.S.A.
E-mail: gge@math.ufl.edu*

Abstract. von Neumann's reliance on the von Mises frequentist interpretation is discussed and compared with the Dutchbook approach proposed by de Finetti.

1. Introduction. In a meeting devoted to Quantum Probability it seems reasonable to glance occasionally backwards and to review some of the circumstances that presided over the birth of our science. In particular, the beginning of the 20th century saw very diverse approaches towards the foundations of classical probability theory. The problem had been squarely posed by Hilbert [Hi1]: “*6. Mathematical Treatment of the Axioms of Physics. The investigations on the foundations of geometry suggest the problem: To treat in the same manner, by means of axioms, those physical sciences in which mathematics plays an important role; in the first rank are the theory of probabilities and mechanics. As to the axioms of the theory of probabilities, it seems to me desirable that their logical investigation should be accompanied by a rigorous and satisfactory development of the method of mean values in mathematical physics, and in particular in the kinetic theory of gases*”.

Note that Hilbert made two distinct demands: (1) the call for *mathematical* definitions (an axiomatics that makes precise the *syntax*), quite likely in a spirit similar to his [Hi2]; (2) the need for an interpretation (*semantic*) of the mathematical structures, even in then controversial (e.g. [Blz]) applications to *physical* world situations: “*...for mathematics is the foundation of all exact knowledge of natural phenomena.*”

Thus, it may appear somewhat strange today that, while Hilbert quoted for mechanics some of the perennial classics (Mach, 1901; Hertz, 1894; Boltzmann, 1897; Volkman, 1900), his only reference pertaining to probability proper was to a text by Bohlmann, 1900, on the “Mathematics of Insurance”.

1991 *Mathematics Subject Classification*: Primary 81P10; Secondary 60A05.

The paper is in final form and no version of it will be published elsewhere.

An authoritative response to Hilbert's first demand came in the form of the concise essay by Kolmogorov [Ko2] who, for generations to come, established the language of probability to be measure theory, with expectation values defined as

$$\langle f \rangle = \int f(x) d\mu(x). \quad (1.1)$$

This is the form that Born [Brn] carried over to quantum mechanics as

$$\langle f \rangle = \int_X f(x) |\Psi(x)|^2 dx \quad (1.2)$$

where the “probability amplitude” $|\Psi(x)|^2$ at x is obtained from the square-integrable “wave function” Ψ . Since $|\Psi|^2$ is \mathcal{L}^1 , the integral is defined for all \mathcal{L}^∞ -functions f . The space X of the variable x was thought of, originally, as the configuration space of the position, but it was learned quickly that this applies as well to the spectrum (discrete or continuous) of a complete set of compatible observables (Dirac's CSCO), i.e. a maximal abelian subalgebra of the algebra of observables. Until the existence of superselection rules was recognized, and the theory of representations was better established, it was natural to identify the observables with the self-adjoint operators affiliated to $\mathcal{B}(\mathcal{H})$. One of the achievements of von Neumann's pioneering book [vN1] was to recognize that the natural extension of (1.1) is thus to identify a *state* as a countably additive, positive linear functional, normalized to 1, hence giving the expectation values

$$\langle A \rangle = \text{Tr } \rho A \quad (1.3)$$

where the “density matrix” ρ , replacing the classical probability measure, is a positive self-adjoint operator belonging to the trace class $\mathcal{T}(\mathcal{H})$, the predual of the von Neumann algebra $\mathcal{B}(\mathcal{H})$ in the same way as \mathcal{L}^1 is the predual of the von Neumann algebra \mathcal{L}^∞ .

We emphasize that up to this point, we are still at the *syntactic* level. Already for classical probabilities, Carnap had complained about – but, in spite of [Car], unfortunately not cured – the condition that the theory suffers not from lacking a semantic but from having too many.

Hilbert's view that probability theory should account for its use in physics, especially as it relates to predictions made with only partial control of the initial conditions, was also expressed by several of the greatest among his contemporaries [Had, HP2, Bo1]; for a later extension of these concerns to the instability of small quantum systems under minute perturbations, see [Zeh]. *Semantic* concerns, moreover, had unescapably intruded the scene with the consideration of the quantum mechanical measuring process, see e.g. [Wig]. This is where von Neumann, von Mises, and later de Finetti, come into the play.

2. von Neumann's probabilistic interpretation of QM states. In [vN1], von Neumann immediately makes clear his position on how to interpret probabilities, namely as limiting frequencies, and why: “*Classical mechanics is a causal discipline ... there also exists a statistical method of treatment of classical mechanics. But this is, as it were, a luxury ... The statistical statements which we found in quantum mechanics have a different character ... [We observe] statistical ensembles which consist of many systems $\mathcal{S}_1, \dots, \mathcal{S}_N$ (i.e. N models of \mathcal{S} , N large.) Such ensembles, called collectives, are in general necessary for establishing probability theory as the theory of frequencies. They were introduced by R.*

v. Mises, who discovered their meaning for probability theory, and who built up a complete theory on this foundation.”

Remark in particular that von Neumann’s presentation strongly presupposes a frequentist interpretation of classical probability theory (an interpretation to which Kolmogorov was to pay barely more than lip service in [Ko2].) Thus two comments must be made here regarding (a) the manufacture and (b) the testing of these “collectives”.

(a) One likes to say today that the systems are identically prepared, as it were by some state factory; these systems are not necessarily in the same pure state, yet they all carry some information characteristic of the ensemble. In the classical realm we can think of such a preparation as letting a gas reach its thermal equilibrium: we do not focus on the instantaneous velocity of the individual molecules, but the temperature of the gas is reflected in their mean speed. A similar situation prevails when discussing the beam of particles produced for the Stern-Gerlach experiment, for instance. We do not need to say that the state depends on the observer, but we need to specify the information we have (and that we don’t have), in order to be able to make reliable predictions.

(b) How do we test all that, in both the quantum and the classical realms ? In spite of the classical character of von Mises’ theory [vM2], von Neumann could import into quantum theory the essential, but as yet vague, concept of subsequence selection: “...with an ensemble of N elements it suffices to carry out the statistical inspections, relative to the distribution of values of the quantity \mathcal{R} not on all elements $\mathcal{S}_1, \dots, \mathcal{S}_N$ but only on any subset of $M < N$ elements ... provided that M and N are both large, and that M is very small compared to N ... This follows from the so-called law of large numbers, the theorem of Bernoulli.”

Von Neumann is less precise than one should want to be today, and perhaps than he could have been. To start with, one should distinguish (see e.g. [Fel]) two “laws of large numbers”: the weak law ((which is Bernoulli’s, and is the only one in [Bo2]) and the strong law [Bo3, Can] (which is what we actually need); moreover there is the law of iterated logarithms [Ko1, Khi] lurking in the background; all were known by 1932.

Let ξ_1, ξ_2, \dots be an infinite sequence of independent Bernoulli trials, each with $\langle \xi_n \rangle = p$. The weak law asserts simply that

$$\lim_{n \rightarrow \infty} \text{Prob} \left\{ \left| \left(\frac{1}{n} \sum_{k=1}^n \xi_k \right) - p \right| \geq \epsilon \right\} = 0. \quad (1.4)$$

The strong law is much stronger (and its proof requires countable additivity):

$$\text{Prob} \left\{ \lim_{n \rightarrow \infty} \left(\frac{1}{n} \sum_{k=1}^n \xi_k \right) = p \right\} = 1. \quad (1.5)$$

This strong law did generate quite some controversy; referred to for a while as “Borel paradox” [Ste, Hau], it goes against the grain: to require that a sequence converges is a severe demand; yet the strong law asserts that almost all of them do. The situation was made worse by the law of iterated logarithms, supporting the naively opposite sentiment that in the long run improbable things *do* happen: it even quantifies how fast they do.

Furthermore, sets of measure zero presented their own semantic problems. Two attitudes were indeed still fighting for the command of the classical field in the years just

preceding von Neumann’s explorations into quantum probability. Bernstein had proposed an axiom [Ber] amounting to say that in the natural sciences, one can safely “*exclude any set of measure zero.*” Yet, Weyl was skeptical (in spite of, or perhaps because of [Wey]): *one should not evaluate highly the value of theorems in which an unspecified set of exceptions of measure zero appears.* This ambiguity must have been one of the motivations of von Mises [vM1,2], whose profound positivist persuasion is on record [vM3]. The same philosophical standpoint may also have contributed to von Neumann’s choice in focusing on von Mises’ “collectives” , which we review in the next section.

3. von Mises revisited. von Mises’ foundations of probability are grounded in the necessity to have an interpretation of the theory in terms of frequencies; he introduced the concept of “collective” specifically for that purpose. These are ensembles of sequences $\xi : n \in \mathcal{N} \mapsto \xi_n \in \Omega$ that satisfy two axioms. For the time-being, let Ω be unspecified.

The first axiom requires that, for *all* sequences in a given collective K , the limiting relative frequencies exist and converge, giving the “limiting frequencies” characteristic of the collective. Specifically, for *every* part $A \subseteq \Omega$, and every $N \in \mathcal{N}$, let $N_A = \#\{n \mid 1 \leq n \leq N, \xi_n \in A\}$:

$$W_A = \lim_{N \rightarrow \infty} \frac{N_A}{N} \text{ exists.} \quad (3.1)$$

The second axiom amounts to a rule that selects those subsequences ξ' of $\xi \in K$ that form a new collective K' with the same limiting frequencies:

$$W'_A = W_A \text{ for all } A \subseteq \Omega. \quad (3.2)$$

For half-a-century, several objections to this axiomatization were raised against each and both of these axioms (for a review, see [ML2]). The first line of attack is that Ω must be specified, lest one runs into difficulties with the requirement that limiting frequencies exist “for all” parts $A \subseteq \Omega$ especially if one wants to avoid restrictions to measurable subsets in any of the usual senses (e.g. Peano–Jordan sets would have to be included, even in some favorable circumstances [Cop]). A second line of criticism is linked to the apparent unavoidability of some form of the law of iterated logarithms: large deviations are hard to proscribe [Vil]. The third line questions the selection of admissible subsequences. It is along this line [Wal] that a liberating path was opened [Chu] in the maze of objections that had all but discredited von Mises’ approach with many probabilists. Note Wald’s affiliation with the Vienna colloquium of Gödel, Menger and Tarski’s fame; logics (specifically recursive function theory) had entered the foundations of probability theory, an entry that was to be confirmed much later, and from a different point of view, by Kolmogorov’s conversion [Ko3] from infinite sequences to finite strings.

Much of the confusion seems to have stemmed from the difficulty inherent in capturing mathematically what constitutes a *random sequence*. Intuitively, a random sequence is one that can *not* be predicted, a sequence such that the observation of its first entries can *not* be used to devise a winning strategy for betting on the subsequent entries, or, still in another formulation, a sequence that presents *no* regularity.

While the pioneering role of von Neumann in the elaboration of quantum probability was based on von Mises’ frequentist approach, it is interesting to note that when faced

with actually producing the kind of (classical) randomness necessary for Monte–Carlo computations, von Neumann took an attitude that is quite in contrast with the smooth confidence of his 1932 urtext: “*Anyone who considers arithmetical methods of producing random digits is, of course, in a state of sin ... there is no such thing as a random number – there are only methods to produce random numbers ... and a strict arithmetic procedure ... is not such a method*” [vN2].

This might have suggested a mathematical definition of randomness along the pragmatic line that a sequence is to be reputed random if it fails to satisfy regularity (or arithmetic) tests. But a Popperian *falsification* approach [Pop] to a mathematical definition was apparently not in the cards at the time. Yet, such an approach, albeit not with this philosophical parti–pris, was to be successfully followed by Martin–Löf [ML1] who adapted the notion of Kolmogorov complexity [Ko3], making essential use of recursive theory (see e.g. [Hod]). As the physics community does not seem to have grasped Martin–Löf’s argument, it is sketched here; for further discussion, see [LiV]. To explain what goes on, it is sufficient to consider binary strings $\{x_1, x_2, \dots, x_n\}$ and sequences $\{\xi_1, \xi_2, \dots, \xi_n, \dots\} : x_k$ and ξ_k belong to $\{0, 1\}$. We denote by X the (countable) set of all binary strings.

DEFINITION 3.1. A test is a subset

$$V = \bigcup_{m \in \mathcal{N}} V_m \subset \mathcal{N} \times X \quad \text{with} \quad x \in V_m \text{ iff } (m, x) \in V$$

such that

- (1) V is recursively enumerable ;
- (2) $V_m \supseteq V_{m+1}$;
- (3) $\#\{A^n \cap V_m\} \leq 2^{n-m}$.

The index of a string x of length $l(x)$ is $m_V(x) = \max\{m \in \mathcal{N} \mid x \in V_m\}$.

Note that (3) gives the key to the interpretation of Def. 3.1: there are 2^n strings in A^n and at most a fraction 2^{-m} of those have the regularity necessary to pass the filter V_m ; this gives a meaning to the fact that the successive filters test for more and more regularity. $m_V(x)$ is thus the maximal level of regularity exhibited by the string, in the test V ; and one has $m_V(x) \leq l(x)$. Evidently, how regular a string is reputed to be will in general depend of the test V ; this situation would improve if one could prove that there is at least one test U that satisfies the following condition.

DEFINITION 3.2. A test U is universal if for every test V there is a $c \in \mathcal{N}$ such that

$$V_{m+c} \subseteq U_m \quad \forall \quad m \in \mathcal{N}.$$

That such a test exists is where recursive theory enters: Since every test V is recursively enumerable, the set of all these tests is also recursively enumerable, i.e. there exists a subset $T \subset \mathcal{N} \times \mathcal{N} \times X$ such that to every test V corresponds precisely one $e \in \mathcal{N}$ with

$$(m, x) \in V \quad \text{iff} \quad (e, m, x) \in T. \quad (3.3)$$

One then verifies that the test U defined, for fixed $n \in \mathcal{N}$, by

$$(m, x) \in U \quad \text{iff} \quad (n, m+n, x) \in T \quad (3.4)$$

is universal. Moreover, for every test V , there exists then an element $c \in \mathcal{N}$ such that

$$m_V(x) \leq m_U(x) + c \quad \forall x \in X. \quad (3.5)$$

Next, for any string x of length $l(x)$, Martin–Löf introduces a conditional complexity $K(x|l(x))$ that measures the minimal length of an algorithmic description of x given its length $l(x)$; and he proves that there is a $k \in \mathcal{N}$ such that, given K and U

$$m_U(x) - k \leq l(x) - K(x|l(x)) \leq m_U(x) + k \quad \forall x \in X. \quad (3.6)$$

This provides a link between Kolmogorov complexity and the lack of regularity in a universal test. The gist of Martin–Löf’s contribution is that his notion of test, and its main properties, can be extended from strings to sequences by the following adjustments.

DEFINITION 3.3. A sequential test is a subset

$$V = \bigcup_{m \in \mathcal{N}} V_m \subset \mathcal{N} \times X \quad \text{with} \quad x \in V_m \text{ iff } (m, x) \in V$$

such that

- (1) V is recursively enumerable ;
- (2) $\{x \in V_m, n \leq m, y \geq x\} \Rightarrow y \in V_n$;
- (3) $\#\{A^n \cap V_m\} \leq 2^{n-m}$

where $y \geq x$ means $l(y) \geq l(x)$ and $y_k = x_k \forall 1 \leq k \leq l(x)$.

Note that a sequential test is indeed a test in the sense of Def. 3.1, and that one can again define the index of a string, now with respect to a sequential test. Universal sequential tests exist, again by an appeal to recursive theory. The interest of Def. 3.3 is that it allows to consider any sequence ξ through its initial strings $\xi^{(n)}$ of length n : $\xi_k^{(n)} = \xi_k \forall 1 \leq k \leq n$. Since, for any sequence ξ and any universal sequential test U , the index $m_U(\xi^{(n)})$ is a non-decreasing function of n , the following limit exists (finite or infinite):

$$m_U(\xi) = \lim_{n \rightarrow \infty} m_U(\xi^{(n)}). \quad (3.7)$$

DEFINITION 3.4. A binary sequence ξ is said to be random in the sense of Martin–Löf iff $m_U(\xi) < \infty$.

Because of (3.5), the condition that a sequence be random is independent of the universal test used in Def. 3.4. Moreover, we have as a consequence of (3.6):

$$m_U(\xi) - k \leq n - K(\xi^{(n)}|n) \leq m_U(x) + k \quad \forall x \in X \text{ and } n \in \mathcal{N} \quad (3.8)$$

which forces the complexity of the initial segment $\xi^{(n)}$ of any random sequence ξ to increase sufficiently fast with n to remain within a certain finite range of n . Some converse of this also holds [ML3], in the sense that any sequence ξ with

$$\#\{n \in \mathcal{N} \mid n - K(\xi^{(n)}|n) \leq c\} = \infty \quad \text{for some } c \in \mathcal{N} \quad (3.9)$$

must be random in the sense of Martin–Löf.

The term “random sequence” has therefore been consistently described at the theoretical level, even if one must admit [Knu] that “*From the standpoint of practical random number generation ... this is ... the worst definition of randomness that can be imagined.*” Suffice it to say here that Martin–Löf also succeeded in [ML1], by a modification

of condition (3) of Definition 3.3, to characterize Bernoulli sequences and that “*these are precisely the sequences for which von Mises introduces the term Kollektiv.*” It is somewhat doubtful that von Neumann would have accepted lightly such a long detour to justify in elementary terms what he meant by an ensemble of identically prepared systems. Hence the urge to turn to some more simple-minded semantic, one that would put the emphasis on the information one has on the manner the system has been prepared.

4. Objective underpinnings in de Finetti’s subjective scheme. Several contributions to classical probability are due to de Finetti (e.g. infinite divisibility, and his representation theorem for exchangeable distributions [Fel]; note in passing that the simplicial structure at the heart of the latter has been extended most remarkably to quantum many-body systems of identical bosons [Hud]). We want to concentrate here on the foundational setting in [deF].

From a (maximal) abelian subalgebra of the observables algebra, we retain a Boolean lattice of events:

$$\mathcal{P} = \{E\} \quad \text{with} \quad \cup, \cap, \subseteq, \neg \quad \text{and} \quad \emptyset, I.$$

The interpretation formalizes, as usual, the idea that events either “occur” or “do not occur”; e.g. $\neg E$ occurs exactly when E does not occur; $E \cup F$ occurs exactly when at least one of E or F occurs; $E \cap F$ occurs exactly when both E and F occur; $E \subseteq F$ means $E \cap F = E$, or E entails F ; E and F are mutually exclusive whenever $E \cap F = \emptyset$, i.e. $E \subseteq \neg F$; $\{E_i \mid i = 1, 2, \dots, n\}$ is exhaustive whenever $\bigcup_{i=1}^n E_i = I$; etc ...

There are two actors: the Bookie and the Bettor; the setting is given by a finite, exhaustive collection $\{E_i \mid i = 1, 2, \dots, n\}$ of mutually exclusive events $E_i \in \mathcal{P}$; the action is governed only by the requirements that for each E_i , the Bookie be free to choose the “odds” p_i , but is obligated to accept bets of any sum (or “stake”) S_i offered by the Bettor, subject to the agreement that when the event E_i is tested, the Bettor’s net gain will be:

$$G_i = \begin{cases} (1 - p_i)S_i & \text{if } E_i \text{ occurs} \\ -p_i S_i & \text{if } E_i \text{ does not occur} \end{cases}$$

DEFINITION 4.1. A Bookie’s assignment $\{p_i\}$ is said to be coherent, whenever the Bettor cannot put his stakes $\{S_i\}$ in such a manner that his net gain

$$G = \sum_{i=1}^n G_i$$

be always positive, independently of the outcomes of the test of the events $\{E_i\}$.

The beauty of this model resides in its semantic simplicity; in effect, the above definition does state that a Bookie makes coherent assignments *if and only if* his assignments conform to the measure-theoretical syntax of [Ko2], i.e. satisfy the following three conditions:

- (1) $0 \leq p_i \leq 1$
- (2) $p_i = 0$ (resp. 1) if $E_i = \emptyset$ (resp. I)
- (3) $\sum_{i=1}^n p_i = 1$.

Indeed, note the equivalence of (3) above with

$$\begin{vmatrix} (1-p_1) & -p_2 & \cdots & -p_n \\ -p_1 & (1-p_2) & \cdots & -p_n \\ \cdots & \cdots & \cdots & \cdots \\ -p_1 & -p_2 & \cdots & (1-p_n) \end{vmatrix} = 0; \quad (4.1)$$

note also that when E_k occurs, the net gain is

$$G^{(k)} = S_k - \sum_{i=1}^n p_i S_i. \quad (4.2)$$

The standard probabilistic interpretation in terms of the expectation value $\langle G \rangle = \sum_{k=1}^n p_k G^k$ is that the condition of coherence is $\langle G \rangle = 0$. No such interpretation however is necessary to make sense of Definition 4.1. The latter is only a statement to the effect that the Bookie has checked his informations and that they are consistent. It makes no statement on which specific values the Bookie should attribute to the particular “odds” he offers for each event. Any such attribution, beyond coherence, requires a specification of the information content of the game. This is what Shannon’s entropy allows to do when the method of Lagrange multipliers is used to optimize this information content, subject to some supplementary knowledge. Using this method, the classical physicist computes such things as canonical equilibrium states. *So does his quantum colleague*; even when doing quantum statistical mechanics, the optimization is usually subject only to the knowledge one has on some set of compatible observables. The propriety of the specific assignments is then to be judged by the predictions that are drawn from them.

5. Conclusions. According to de Finetti’s foundational position [deF], neither these assignments, nor the concomitant predictions, necessarily need to be formulated in a statistical language of the kind that von Mises proposed, although habitually they have been so interpreted. The view of the state of a system, as it emerges from de Finetti’s emphasis on the relation between preparations and predictions, does not presuppose any commitment to some philosophical view that the state of a quantum system depends on the observer, only that the observer makes predictions on the basis of the information he has on the preparation of the system; for a pursuit of this Bayesian line in physics, see [Jay]. Even more urgently, a view such as de Finetti’s seems to be necessary if one is to speak of the behavior of an “isolated” system, a task which is now demanded by actual laboratory situations where individual quantum systems (rather than ensembles) are claimed to be prepared and sustained.

References

- [Ber] F. Bernstein, *Ueber eine Anwendung der Mengenlehre auf ein aus der Theorie des säkularen Störungen herrührendes Problem*, Math. Annalen, 71 (1912), 417–439.
- [Blz] L. Boltzmann, *Referat Ueber die Abhandlung von J.C. Maxwell “Ueber Boltzmanns Theorie betreffend mittlere Verteilung der Lebendige Kraft in einem System materieller Punkte”*, (1881), WA 2 582-595.

- [Bo1] E. Borel, *Sur les principes de la théorie cinétique des gaz*, Ann. Ecole Norm. 23 (1906), 9–32.
- [Bo2] E. Borel, *Eléments de la théorie des probabilités*, Hermann, Paris, 1909.
- [Bo3] E. Borel, *Les probabilités dénombrables et leurs applications arithmétiques*, Rend. Cir. mat. Palermo 27 (1909), 247–271.
- [Brn] M. Born, *Quantenmechanik und Statistik*, Die Naturwissenschaften 15 (1927) 238–242; *Statistical Interpretation of Quantum Mechanics (Nobel Lecture)*, Sciences 122 (1955), 675–679; reprinted in M. Born *Physics in my Generation*, Pergamon Press, London, 1969.
- [Can] F. Cantelli, *Sulla legge dei grandi numeri*, Mem. Acad. Lincei 11 (1916), 330–349; *Sulla probabilità come limite della frequenza*, Rend. Acad. Lincei 26 (1917), 39–45.
- [Car] R. Carnap, *Logical Foundations of Probability*, 2nd ed., University of Chicago Press, Chicago, 1962.
- [Chu] A. Church, *On the Concept of a Random Sequence*, Bull. Amer. Math. Soc. 47 (1940), 130–135.
- [Cop] A.H. Copeland, *Admissible Numbers in the Theory of Geometrical Probability*, Amer. J. Math. 53 (1931), 153–162; *The theory of Probability from the point of view of Admissible Numbers*, Ann. Math. Statist. 3 (1932), 1435–1456.
- [Fel] W. Feller, *An Introduction to Probability Theory and its Applications*, vols. 1 (3rd ed.) and 2 (2nd ed.), Wiley, New York, 1968, 1966.
- [deF] B. de Finetti, *La prévision, ses lois logiques et ses sources subjectives*, Annales de l'Institut Henri Poincaré 7 (1937), 1–68; *Probability, Induction, and Statistics*, Wiley, New York, 1972.
- [Had] J. Hadamard, *Les surfaces à courbures opposées et leurs lignes géodésiques*, J. math. pures et appl. 4 (1898), 27–73.
- [Hau] F. Hausdorff, *Grundzüge der Mengenlehre*, De Gruyter, Leipzig, 1914.
- [Hi1] D. Hilbert, *Sur les problèmes futurs des mathématiques* in: *Comptes Rendus du Deuxième Congrès International des Mathématiciens, Paris, 1900*, Gauthier-Villars, Paris, 1902, 58–114; see also: *Mathematical Developments arising from Hilbert problems*, F.E. Browder (ed.), Proc. Symp. Pure Math. XXVIII, 1976.
- [Hi2] D. Hilbert, *Grundlagen der Geometrie*, Teubner, Leipzig, 1899 (1st ed.), 1922 (5th ed.); *Foundations of Geometry* (L.Unger transl. from the 10th ed.), Open Court, LaSalle IL, 1971.
- [Hod] R.E. Hodel, *An Introduction to Mathematical Logic*, PWS Publ., Boston, 1995.
- [Hud] R.L. Hudson, *Analogs of de Finetti's Theorem and the Interpretative Problems of Quantum Mechanics*, Found. Phys. 11 (1981), 805–808; see also: R.L. Hudson and G.R. Moody, *Locally Normal Symmetric States and an Analogue of de Finetti's Theorem*, Z. Wahrscheinlichkeitstheorie verw. Gebiete 33 (1976), 343–351.
- [Jay] E.T. Jaynes, *Papers on Probability, Statistics, and Statistical Physics*, 2nd ed., Kluwer, 1989.
- [Khi] A. Khintchine, *Ueber einem Satz der Wahrscheinlichkeitsrechnung*, Fundam. Math. 6 (1924), 9–20.
- [Knu] D. E. Knuth, *The Art of Computer Programming, vol. 2*, 2nd ed., Addison-Wesley, Reading, Mass., 1981.
- [Ko1] A.N. Kolmogorov, *Das Gesetz des iterierten Logarithmus*, Math. Annalen 101 (1929), 126–135.

- [Ko2] A.N. Kolmogorov, *Grundbegriffe der Wahrscheinlichkeitsrechnung*, Springer, Berlin, 1933; *Foundations of the Theory of Probability* (N. Morrison, transl.), Chelsea, 1956.
- [Ko3] A. Kolmogorov, *On Tables of Random Numbers*, Sankhya, The Indian Journal of Statistics, Ser. A 25 (1963), 369–376; *Logical Basis for Information Theory and Probability Theory*, IEEE Trans. Inform. Theory 14 (1968), 662–664.
- [LiV] M. Li and P. Vitanyi *An Introduction to Kolmogorov Complexity and its Applications*, 2nd ed., Springer, New York, 1997.
- [ML1] P. Martin-Löf, *The Definition of Random Sequences*, Information and Control, 9 (1966), 602–619.
- [ML2] P. Martin-Löf, *The Literature on von Mises' Kollektivs Revisited*, Theoria, 35 (1969), 12–37.
- [ML3] P. Martin-Löf, *Complexity Oscillations in Infinite Binary Sequences*, Z. Wahrsch. verw. Geb. 19 (1971), 225–230.
- [vM1] R. von Mises, *Grundlagen der Wahrscheinlichkeitsrechnung*, Math. Z. 5 (1919) 52-99.
- [vM2] R. von Mises, *Wahrscheinlichkeit, Statistik und ihre Wahrheit*, Springer, Wien, 1928; *Probability, Statistics and Truth*, MacMillan, New York, 1939.
- [vM3] R. von Mises, *Kleines Lehrbuch der Positivismus, Einführung in die empiristische Wissenschaftsauffassung*, W. P. van Stockum, Den Haag, and University of Chicago Press, Chicago, Ill., 1939; transl. by J. Bernstein and R.G. Newton, as: *Positivism*, Harvard University Press, Cambridge, Mass., 1951.
- [vN1] J. von Neumann, *Mathematische Grundlagen der Quantenmechanik*, Springer, Berlin, 1932; *Mathematical Foundations of Quantum Mechanics* (R.T. Beyer, transl.) Princeton University Press, Princeton NJ, 1955.
- [vN2] J. von Neumann, *Various Techniques used in Connection with Random Digits*, J. Res. Nat. Bur. Stand. Appl. Math Series 3 (1951), 36-38; *Collected Works, vol. V*, MacMillan, New York, 1963, 768-770.
- [HP2] H. Poincaré, *La science et l'hypothèse*, Flammarion, Paris, 1902.
- [Pop] K.R. Popper, *Logik der Forschung*, Vienna, 1935; *The Logic of Scientific Discovery*, Basic Books, New York, 1959; Harper & Row, New York, 1968.
- [Ste] H. Steinhaus, *Les probabilités dénombrables et leur rapport à la théorie de la mesure*, Fundam. Math. 4 (1923), 286–310.
- [Vil] J. Ville, *Sur la notion de collectif*, C.R. Acad. Sci. Paris, 203 (1936), 26–27; *Etude critique de la notion de collectif*, Gauthier-Vilars, Paris, 1939.
- [Wal] A. Wald, *Sur la notion de collectif dans le calcul des probabilités*, C.R. Acad. Sci. Paris, 202 (1936), 180–183; *Die Widerspruchsfreiheit des Kollektivbegriffes der Wahrscheinlichkeitsrechnung*, Ergebnisse eines Kolloquium (Vienna) 8 (1937), 38–72; *Die Widerspruchsfreiheit des Kollektivbegriffes*, Actualités Sci. Indust. 735 (1938), 79–99.
- [Wig] E.P. Wigner, *Review of the Quantum-Mechanical Measurement Problem*, in: *Science, Information Onslaught*, D.M. Kerr et al., eds., Academic Press, New York, 1984, 63–82. *Collected Works, vol 6*, Springer, Heidelberg, 1994, 225-244.
- [Wey] H. Weyl, *Ueber die Gleichverteilung von Zahlen mod Eins*, Math. Annalen, 77 (1916) 313–352.
- [Zeh] H.D. Zeh, *On the Interpretation of Measurement in Quantum Theory*, Found. of Phys. 1 (1970), 69–76.