Dispute and agreement between two pioneers of modern probability and mathematical statistics: Andrej Kolmogorov and Richard von Mises in 1932

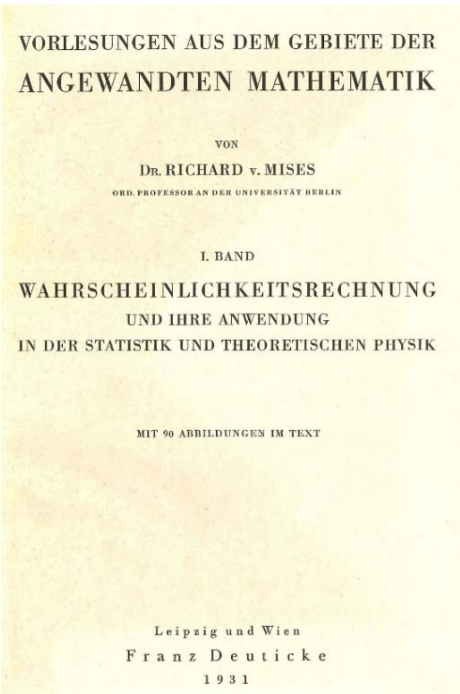
Reinhard Siegmund-Schultze (University of Agder, Kristiansand, Norway)

From a one- hour lecture held during the 64th ISI WSC in Ottawa 2023, invited by the President of the Bernoulli Society

1.) Introduction

The two famous reformers of probability theory Richard von Mises (1883-1953) and Andrej Kolmogorov (1903-1987) had a correspondence of five letters in 1932, triggered by Kolmogorov's review of von Mises' book on probability and applications of 1931.

А.Н. Колмогоров (1930 г.)



VORLESUNGEN AUS DEM GEBIETE DER
ANGEWANDTEN MATHEMATIK

VON
DR. RICHARD v. MISES
ORD. PROFESSOR AN DER UNIVERSITÄT BERLIN

I. BAND
WAHRSCHEINLICHKEITSRECHNUNG
UND IHRE ANWENDUNG
IN DER STATISTIK UND THEORETISCHEN PHYSIK

MIT 90 ABBILDUNGEN IM TEXT

Leipzig und Wien
Franz Deuticke
1931

"Probability Theory and its application in statistics and theoretical physics" (1931)

In this voluminous book of 574 pp. the author von Mises goes rarely into philosophical or foundational issues, because "Anwendungen" (Applications) had priority.

Further volumes of the "Vorlesungen"-series never appeared due to von Mises' persecution by the Nazis and emigration to Istanbul in 1933.

Hilda Geiringer (1893-1973)

In his Berlin institute for applied mathematics Richard von Mises had guided his assistant and future wife Hilda Geiringer (prev. married to the statistician F. Pollaczek) to applied mathematics and stochastics. She wrote the first draft of his probability book of 1931 after his lectures of 1925. Geiringer finally became an international recognized specialist in statistics and plasticity.

My talk at the Ottawa 64th World Congress of Statistics in July 2023 had two main goals:

- To correct the distorted view of the contributions of Richard von Mises (henceforth RvM) which often focuses too much on his partly unsuccessful axioms for probability (his "collectives") and neglects his positive impact on foundations (distributions), in Markov chains (ergodic theory), mathematical statistics, and generalisations
- To indicate the influence of RvM on Kolmogorov in foundations (see also Shafer/Vovk 2006), Markov chains, interpretation of relation to reality, and mathematical statistics

In a German article on the history of mathematical statistics of 1990, Hermann Witting says the following about RvM's contribution:

> "R. VON MISES was far more important [than Felix Bernstein] for the development of mathematical statistics. Admittedly, the professorship established (ultimately for him) in Berlin in 1920 was dedicated to the entire field of applied mathematics. His name is perhaps also more associated with (classical) applied mathematics and mechanics as well as the foundation of the concept of probability. VON MISES nevertheless made decisive contributions to at least four areas of mathematical

> statistics: extreme value theory, the [Sergei] Bernstein - von Mises theorem, the Cramér-von Mises test and the von Mises functionals." (Witting 1990, 792)

Speaking about von RvM's much disputed two axioms for the foundations of probability theory (1919) and their influence on the future axiomatics of probability in (Kolmogoroff 1933a) the former Croatian-Austrian, later American William Feller wrote in 1950 in his *An Introduction to Probability Theory and Its Applications*:

> "The statistical, or empirical, attitude toward probability has been developed mainly by R. A. Fisher and R. von Mises. The notion of sample space [footnote by Feller: "The German word is *Merkmalraum* (label space)"] comes from von Mises. This notion made it possible to build up a strictly mathematical theory of probability based on measure theory. Such an approach emerged gradually in the twenties under the influence of many authors. An axiomatic treatment representing the modern development was given by A. Kolmogorov." (Feller 1968, 6)

As is well known, Kolmogorov used RvM's rudimentary notion of randomness much later in 1963 within his fundamental work on algorithmic complexity. Referring to his own pioneering German book of 1933 Kolmogorov said:

> "I have already expressed the view … that the basis for the applicability of the results of the mathematical theory of probability to real 'random phenomena' must depend on some form of the *frequency concept of probability*, the unavoidable nature of which has been established by von Mises in a spirited manner." (Kolmogorov 1963, 369)

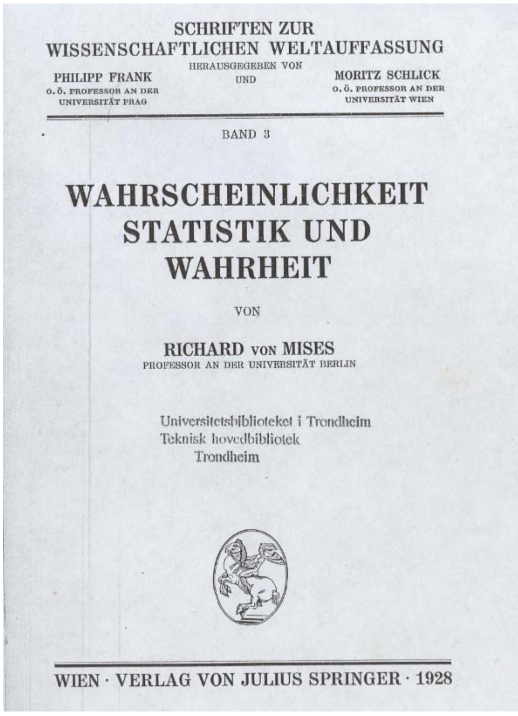
2. Kolmogorov's interest in von Mises' work before 1931

Already in the 1920s Russian and Soviet mathematicians had been keenly interested in RvM's concept of probability of 1919 (Mises 1919a/b). This had apparently to do with RvM's competence in applied mathematics which also informed his view of probability theory.

Historian of mathematics A. P. Yushkevich reported on seminars in Moscow around 1928 which he himself, but also leading experts such as A. Ya. Khinchin and A. N. Kolmogorov, attended:

“Lectures were given and were followed by discussions. I remember the heated arguments that arose in connection with R. von Mises’ frequency theory of the foundations of probability.” (Zdravkovska/Duren 1993, 18)

From the beginning this Russian interest was mixed with philosophical concerns because RvM was known as a positivist and a staunch supporter of Ernst Mach, who had been severely criticized for his “idealism” in Lenin’s book *Materialism and Empirio-Criticism* of 1909. In 1930 the Russians translated RvM’s semi-popular book of 1928 “Probability, statistics, and truth”, however not without distorting and mistranslating the title:


SCHRIFTEN ZUR
WISSENSCHAFTLICHEN WELTAUFFASSUNG
HERAUSGEGEBEN VON
PHILIPP FRANK UND MORITZ SCHLICK
O. Ö. PROFESSOR AN DER UNIVERSITÄT PRAG — O. Ö. PROFESSOR AN DER UNIVERSITÄT WIEN

BAND 3

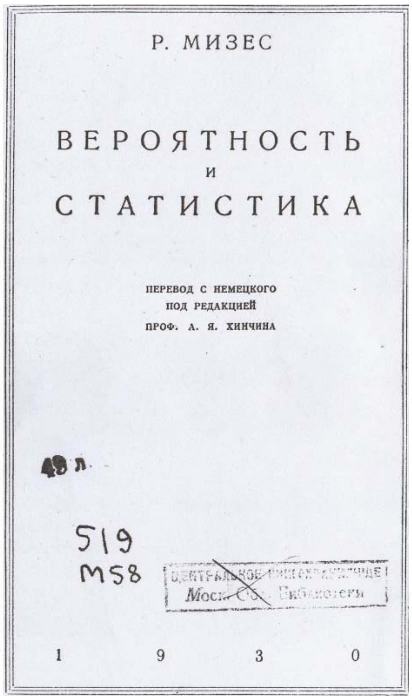
WAHRSCHEINLICHKEIT
STATISTIK UND
WAHRHEIT

VON

RICHARD VON MISES
PROFESSOR AN DER UNIVERSITÄT BERLIN

Universitetsbiblioteket i Trondheim
Teknisk hovedbibliotek
Trondheim

WIEN · VERLAG VON JULIUS SPRINGER · 1928


Р. МИЗЕС

ВЕРОЯТНОСТЬ
И
СТАТИСТИКА

ПЕРЕВОД С НЕМЕЦКОГО
ПОД РЕДАКЦИЕЙ
ПРОФ. А. Я. ХИНЧИНА

519
M58

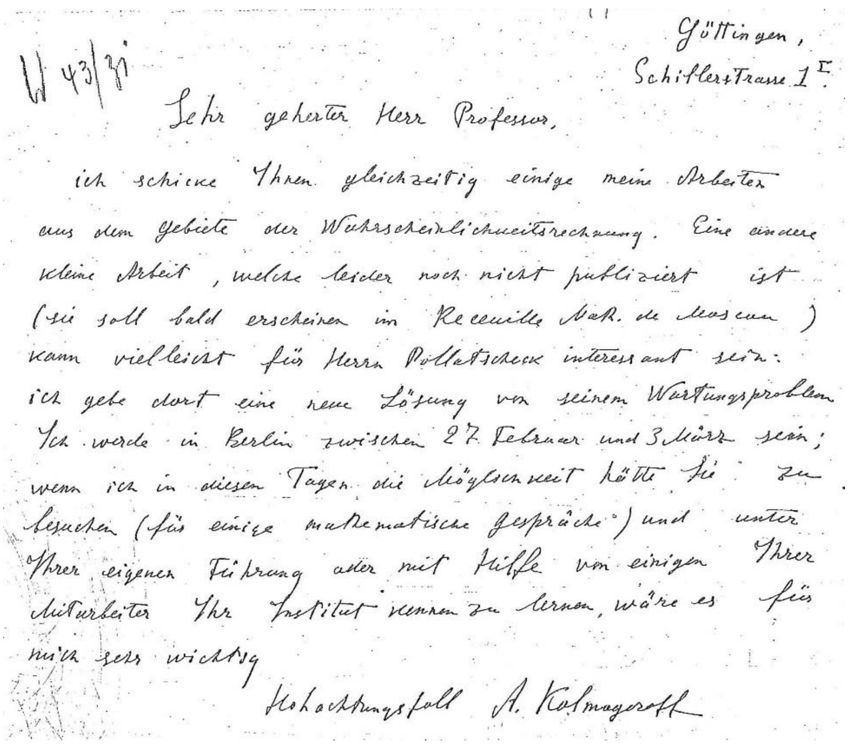
1 9 3 0

The Russian (pirate-) translation, organised by A. Ya. Khinchin in 1930, eliminates the word “Wahrheit” (= truth = istina), probably out of ideological fear: von Mises was a positivist and supporter of Ernst Mach! (Siegmund-Schultze 2004)

3. Von Mises’ book (1931), and Kolmogorov’s review (1932)

In early 1931 the 28years old Kolmogorov was visiting Göttingen University (Germany). From there he sent an undated letter to RvM in Berlin, which is kept – as the other correspondence between the two – at the Harvard University Archives, Cambridge, MA.

Enclosed within the letter Kolmogorov sent separata of his work and asked RvM – 20 years his senior – for an invitation to RvM's institute in Berlin and for "einige mathematische Gespräche" (some mathematical conversations).



W 43/31

Göttingen,
Schillerstrasse 1 I.

Sehr geehrter Herr Professor,

ich schicke Ihnen gleichzeitig einige meine Arbeiten aus dem Gebiete der Wahrscheinlichkeitsrechnung. Eine andere kleine Arbeit, welche leider noch nicht publiziert ist (sie soll bald erscheinen im Recueille Math. de Moscou) kann vielleicht für Herrn Pollatscheck interessant sein: ich gebe dort eine neue Lösung von seinem Wartungsproblem. Ich werde in Berlin zwischen 27 Februar und 3 März sein; wenn ich in diesen Tagen die Möglichkeit hätte Sie zu besuchen (für einige mathematische Gespräche) und unter Ihrer eigenen Führung oder mit Hilfe von einigen Ihrer Mitarbeiter Ihr Institut kennen zu lernen, wäre es für mich sehr wichtig

Hochachtungsvoll A. Kolmogoroff

The meeting took place in March 1931. To his intimate friend, the topologist P. S. Aleksandrov, Kolmogorov reported on it 2 March 1931 with the following words:

> "Was at Mises' institute today from 10 to 12. It was interesting, although his exceptionalism (in the sense of not wanting to know and understand the work of other directions) did surprise me. Before he arrived, a certain assistant was demonstrating a Galton Board [? вечную модель с шариками]. By the way, Mises is writing a book on probability theory for some Viennese publisher. Saw the proofs: the principles of Kollektivslehre are somewhat fluffy [пухлы] (100 pages), the real theory (in the sense of Chebyshev and Lyapunov) is very reduced, but the physical applications will probably be interesting." (Shiryaev 2003, volume II, 409)[1]

Early in 1932 Kolmogorov wrote a review of von RvM's book (Mises 1931) for *Zentralblatt* (Kolmogoroff 1932). In the review Kolmogorov mentions as particularly valuable RvM's work on the "Ergodenhypothese," but he utters doubt about RvM's handling of "conditional probabilities." Kolmogorov insists on the need for "pure axiomatics," apparently as opposed

---

[1] Translations in this paper, mostly from German and Russian, are mine.

to RvM’s more intuitive and empirical axioms of 1919 which continued to be the foundation of his 1931 book. Kolmogorov points, in particular, to the impossibility to ascribe “probabilities for all subsets” of given sets (events). Another point of Kolmogorov’s criticism was RvM’s proof of the central limit theorem (See Siegmund-Schultze 2006).

Wahrscheinlichkeitsrechnung, Versicherungsmathematik:

● **Mises, Richard v.: Vorlesungen aus dem Gebiete der angewandten Mathematik. Bd. 1. Wahrscheinlichkeitsrechnung und ihre Anwendung in der Statistik und theoretischen Physik.** Leipzig u. Wien: Franz Deuticke 1931. X, 574 S. u. 90 Abb. RM. 30.—.

Das Buch hat nach dem Vorwort des Verf. 2 verschiedene Zwecke: einen wissenschaftlichen — eine vollständige Darstellung der Wahrscheinlichkeitsrechnung auf Grund der Theorie der Kollektive (diese Darstellung soll „nicht nur die Möglichkeit, sondern auch die Notwendigkeit der veränderten Grundauffassung“ nachweisen) und zweitens einen pädagogischen, d. h. als Lehrbuch zu dienen. Im letzten Paragraphen des Buches findet man überdies eine längere interessante Untersuchung des Verf. über die Ergodenhypothese. Für die Grundlagenauffassung des Verf. ist der erste Abschnitt („Die Elemente der Wahrscheinlichkeitsrechnung, § 1. Das Kollektiv und die Wahrscheinlichkeit, § 2. Die Verteilungen, § 3. Die Grundoperationen“) der wesentlichste. Der Kollektivbegriff ist hier gründlich erklärt als eine eigenartige Idealisierung des empirischen Begriffs der Massenerscheinung. Auf die Möglichkeit einer rein mathematischen axiomatischen Behandlung dieses Begriffes [vgl. Karl Dörge, Math. Zeitschr. **32**, 232—258 (1930)] wird dabei nur hingewiesen. Nach der Postulierung der Grundeigenschaften des Kollektivs sind die Wahrscheinlichkeiten als die Grenzwerte der relativen Häufigkeiten definiert. Die bedingten Wahrscheinlichkeiten sind nur bei der positiven Wahrscheinlichkeit der Bedingung definiert (§ 3,3, S. 86), obwohl man sie weiter auch im Falle der Wahrscheinlichkeit Null der Bedingung braucht (z. B. § 6, 1, S. 152). Eine andere Unvollständigkeit der Darstellung besteht darin, daß man eigentlich die Existenz der Wahrscheinlichkeiten für alle Untermengen der Merkmalmenge des Kollektivs — wie es ausdrücklich in der Definition des Kollektivs steht — nicht verlangen kann. Dieser Umstand wird übrigens vom Verf. selbst erwähnt (§ 1, 3, S. 17). Als Lehrbuch ist das Buch außerordentlich

Kolmogorov refers in some detail to a diagram given by RvM on page 198 of his book, which follows (Mises 1919a), basing probability theory on two “fundamental theorems”, relating to direct (Bernoullian) and inverse (Bayesian) probability.

I. Bernoullischer Problemkreis Summen- bzw. Durchschnittsbildung		II. Bayesscher Problemkreis Rückschluß nach Beobachtungen
1. Direkte Lösung (Newtons Formel) $\mathfrak{w}_n(x) = \binom{n}{x} p^{n-x} q^x$	aus n gleichen Alternativen bzw. aus n-maliger Beobachtung an einer Alternative	1. Direkte Lösung (Bayes) $w_n(x) = \text{konst.} \times v(x)\left[\left(\frac{x}{\alpha}\right)^{\alpha}\left(\frac{1-x}{1-\alpha}\right)^{1-\alpha}\right]^n$
2. Grenzübergang (Laplace) $\mathfrak{w}_n(x) \to \text{konst.}\, e^{-u^2}$ für $n \to \infty$		2. Grenzübergang (Laplace) $w_n(x) \to \text{konst.}\, e^{-u^2}$ für $n \to \infty$
3. Bernoullisches Theorem $W'_n(q-\varrho, q+\varrho) \to 1$ für $n \to \infty$ bei jedem ϱ		3. Bayessches Theorem $W_n(\alpha-\varrho, \alpha+\varrho) \to 1$ für $n \to \infty$ bei jedem ϱ
4. Erstes Gesetz der großen Zahlen $W'_n(b'_n-\varrho, b'_n+\varrho) \to 1$ für $n \to \infty$	aus n beliebigen Kollektivs bzw. aus n-maliger Beobachtung an einem beliebigen Kollektiv	4. Zweites Gesetz der großen Zahlen $W_n(\alpha_\varkappa-\varrho, \alpha_\varkappa+\varrho) \to 1$ für $n \to \infty$
5. Erster Fundamentalsatz $w_n(x)$ bzw. $\mathfrak{w}_n(x) \to$ $\to \text{konst.}\, e^{-u^2}$ für $n \to \infty$		5. Zweiter Fundamentalsatz $w_n(x) \to \text{konst.}\, e^{-u^2}$ für $n \to \infty$

“The second section (§§ 5-8) is devoted to the limit theorems of probability theory. These limit theorems are arranged in 2 columns (cf. the table on p.198). The first column begins with Bernoulli’s theorem and leads to the first fundamental theorem as its final generalisation (§ 8, 7, p. 224). This theorem was first proved by Liapounoff in 1900 (only for discrete distributions, but under much less restrictive conditions for the dispersions and the third moments). The second column begins with Bayes’ theorem (also called the inverse Bernoullian theorem). As a generalisation of this theorem, one obtains the second fundamental theorem: ‘If the n times observation of a collective has yielded n results which have the mean α and the mean square of deviation σ^2, then the probability density that the expected value of the distribution lies at x is given for sufficiently large n, whatever the initial probability, by Gauss’s law with the mean α and the dispersion $\sigma^2 : n$.’” (Kolmogoroff 1932, 277)

Interestingly enough, Kolmogorov stressed the “inverse” Bayesian theorem as particularly noteworthy.

RvM reacted to Kolmogorov’s review of his book with a letter to Moscow, dated 25 January 1932. He asked him three questions, of which we quote here only the third:[2]

“You note that the first fundamental theorem was already proved in 1900 by Liapounoff under partly more general conditions, even if only for arithmetic [i.e. discrete] distributions. You cite the formulation of the theorem by me on p. 224 as the ‘final generalisation’; however, it is a fact that what is derived by me in § 8.4 as the first fundamental theorem for arithmetic distributions and stated on p. 212 has a much more far-reaching content than the passage cited by you on p. 224.”

RvM’s local limit theorem with conditions for moments (Mises 1931, 212/13, also referring to Mises 1919a, 28) was the following:

“We now express the first fundamental theorem of probability theory [for ‘arithmetic’, i.e., discrete distributions, R.S.] as follows:

The summation of n collectives with arithmetic distributions leads with large n to Gauss’ law; mean and dispersion are here the sums of the means and the dispersions

[2] I restrict the discussion here to this question, which is closest to mathematical statistics. The other two questions concerned conditional probabilities and RvM’s axioms for probability (collectives). The Ottawa talk touched upon these two as well. The entire correspondence will be discussed in a forthcoming publication, presumably in book form.

of the individual distributions, respectively. The conditions under which this theorem was proved here are:

a) The sum of the dispersions of the given distributions grows like n to infinity.

b) The absolute third-degree moments of the distributions have an upper bound.

c) In the given distributions – except for at most finitely many among them – occurs at least one pair of immediately consecutive integers with non-vanishing probabilities."

Kolmogorov responded on 10 February 1932:

> "The theorem p. 212 was unknown before your work, as far as I know. That I do not consider this theorem but the integral theorem p. 224 as the most important final result in this direction is really a deviation from your view. The reasons are twofold: first, I believe that not only the arithmetic distributions with integer values of the label (Merkmal) x are of interest to statistics, for example, when in heritability theory a quantitative trait depends on several Mendelian factors. Secondly, the assertion of the integral theorem, which is therefore valid under much more general conditions, is completely sufficient in most practical cases: how the distribution is constructed in very small intervals is less interesting, especially since one usually cannot directly observe this fine structure of the distribution. You yourself show, for example, that one can determine the distribution function with arbitrary accuracy on the basis of observations (ω^2 - method) [Cramér-Mises test]; this would not be correct for the density function."

RvM insisted on the importance of his local limit theorem with the following letter of 19 February 1932:

> "I want to explain why I consider my theorem of p. 212 to be more important, or at least just as important, as L[iapunov]'s result. If one only says something about the integrals or sums of probabilities, then one does not even know whether, for example, in the simplest Bernoullian case, even and odd results occur with equal frequency. If one assumes, for example, that in the urn from which it has been n times drawn, the even numbers from 0 to 200 are all present and, in addition, only the one. It then does not seem very plausible that (assuming equal probability of the numbers) even and odd sums are drawn equally often. Also, knowing only Liapounoff's theorem, it

would still be possible that individual sum-values appear with relatively very high probability. What is more, the ratio of the probability of two immediately neighbouring numbers could be arbitrarily different from one, and this would be obviously something quite different from what the statistician believes and would like to have proven."

In the correspondence about the review (Kolmogoroff 1932), which ends with a letter by Kolmogorov on 26 March 1932, one does not find an admission by Kolmogorov of the importance of RvM's theorem. However, one finds such admission in a publication one year later in an important Italian paper "Sulla determinazione empirica di una legge di distribuzione" (Kolmogoroff 1933b) which connected to the ω^2 - method (Cramér-Mises test) in the book (Mises 1931, 316). Although the latter method was originally related to the more global level of distributions, not to densities (as Kolmogorov had emphasized in his letter), Kolmogorov had now realized the importance of RvM's local limit theorem for discrete distributions, connecting to condition c) in his theorem (see above):

> "The condition c) which is essential in our new theorem, has already been used by von Mises in similar considerations." (Kolmogoroff 1933b, 145)

E.V. Khmaladze (1986) described this paper as one of the most important among Kolmogorov's statistical works, among other things introducing the 'Kolmogorov distribution'.

The further development of the relationship between the two probability theorists and statisticians Richard von Mises and Andrej Kolmogorov is characterized by continued mutual appreciation although RvM would never accept modern measure theory as a foundation for probability. This has already been indicated in the introduction above and will be specified in a forthcoming detailed study.

Acknowledgments:

I thank the former president of the Bernoulli Society, Adam Jakubowski (Torun), for the invitation to the Ottawa talk, and the Archives of Harvard University for the permission to use the correspondence between Richard von Mises and A. N. Kolmogorov which is contained in the Richard von Mises Papers HUG 4574.

References:

Feller, W. (1968): *An Introduction to Probability Theory and Its Applications,* 3. ed. New York: Wiley

Fischer, G., F.Hirzebruch, W.Scharlau und W.Toernig (Hrg., 1990): *Ein Jahrhundert Mathematik 1890-1990*, Braunschweig: Vieweg.

Khmaladze, E.V. (1986). Empirical Distributions [commentary on (Kolmogorov 1933b)], in Shiryaev (ed. 1986/1992), 574-583.

Kolmogoroff, A. (1932): Review of Mises (1931), *Zentralblatt* 2, no. 5, 277/78.

Kolmogoroff, A. (1933a): *Grundbegriffe der Wahrscheinlichkeitsrechung*, Berlin: Springer.

Kolmogoroff, A. (1933b): Sulla determinazione empirica di una legge di distribuzione, *Giorn. Istit. Ital. Attuari*., v. 4, № 1, 1933; 83-91. (Quote following Engl translation as "On the empirical determination of a distribution law", in Shiryaev (ed. 1986/1992), 139-146.

Kolmogorov, A.N. (1963): On Tables or Random numbers, *Sankhyā. Indian Journal of Statistics*. Ser. A. 1963, vo1.25, N 4, 369-376.

Mises, R. von (1919a): Fundamentalsätze der Wahrscheinlichkeitsrechnung, *Mathematische Zeitschrift* 4, 1-97.

Mises, R. von (1919b): Grundlagen der Wahrscheinlichkeitsrechnung; *Mathematische Zeitschrift* 5, 52-99.

Mises, R. von (1931): *Wahrscheinlichkeitsrechnung und ihre Anwendungen in der Statistik und der theoretischen Physik*, Leipzig and Vienna: Franz Deuticke.

Shafer, G. and V. Vovk (2006): The Sources of Kolmogorov's Grundbegriffe, *Statistical Science* 21, No. 1, 70–98.

Shiryaev, A.N. (ed. 1986/1992): *Selected Works of A. N. Kolmogorov*, Volume II *Probability Theory and Mathematical Statistics* (Russian 1986, Moscow: Nauka, English 1992, Dordrecht: Kluwer).

Shiryaev, A.N. (ed. 1987/1993): *Selected Works of A. N. Kolmogorov*, Volume III, *Information Theory and the Theory of Algorithms* (Russian 1987, Moscow: Nauka, English 1993, Dordrecht: Kluwer).

Shiryaev, A.N. (ed. 2003): КОЛМОГОРОВ юбилейное издание в трех книгах, МОСКВА: ФИЗМАТЛИТ (Kolmogorov. Jubilee Edition in three volumes, Russian, Moscow: Fizmatlit)

Siegmund-Schultze, R. (2004): Mathematicians forced to philosophize: an introduction to Khinchin's paper on von Mises' theory of probability, *Science in Context* 17 (2004), 373-390.

Siegmund-Schultze, R. (2006): Probability in 1919/20: the von Mises-Pólya-controversy, *Archive for History or Exact Sciences* 60, 431–515.

Witting, H. (1990): Mathematische Statistik, in Fischer et al. (eds. 1990), 781-815.

Zdravkovska, S. and P. L. Duren (eds. 1993): *Golden Years of Moscow Mathematics*, Providence: American Mathematical Society.