From Markov to Doeblin : events in chain

Laurent MAZLIAK

RMR-2010
Rouen- 1er juin 2010

The aim of my talk is to stress the connection between the first studies about random events in chain by Markov and Poincaré and the explosive development of these questions since the end of the 1920s. An interesting hypothesis considered by several authors, such as Bru or Von Plato, is that the International Congress of Mathematicians in Bologna, Italy in 1928 allowed the confluence of several currents of study of markovian type phenomena, which were going to join as a common river. A central character in this story, quite forgotten today, is the Czech mathematician Hostinský who played the role of a hub and center of gravity between the Eastern and Western parts of Europe during the 1930s.
My presentation will therefore be divided along the following main lines.
1) Markov and Poincaré
2) Hostinský
3) The consequences of the Bologna congress
In fact, I shall probably only briefly comment on part 3 though it should obvisouly be the longest one. And I shall not comment at all on a virtual part 0 which would have consisted in the description of what Sheynin has called the *prehistory* of chains : one may indeed observe Markovian situations in scientific litterature long before the 20th century, in particular in Bernoulli or Laplace.

1. THE FOUNDERS : MARKOV AND POINCARÉ.

a) Markov

Credit where credit is due. Markov, in 1907, introduced the first systematic model of events linked in chain which would lead to the future developments, well known to us now. Before giving some details about this model and the subsequent studies of the Russian mathematician, it is a good idea to have a look at the couloured life of this overexcited character. Eugene Seneta has devoted several papers to him, and especially to his recurrent polemics with the Moscow mathematician Nekrasov which will play a role in our story.

Markov was born in 1856 in Riazan. Quite early in his adolescence, he seems to have been noticed for two characteristics which accompanied him all his life : his mathematical talent and his bad-tempered character incapable of compromise. He entered Saint Petersburgh university in 1874 where he followed lectures by Korkin, Zolotarev and above all Chebyshev of whom he became the favorite pupil. In 1883, Chebyshev retired and proposed Markov to replace him, in particular for his lectures on Probability. At the same time, Markov defended his thesis *On several applications of algebraic continued fractions* where one does not find probability explicitly but where many results are based on inequalities obtained by Chebyshev as consequence of his method of moments.

Markov's probabilistic career really began with his extension of Chebyshev's incomplete proof of the central limit theorem in 1898. In fact, from this moment, Markov constantly worked on generalizations and studies of limit theorems (Law of Large Numbers and Central Limit Theorem), in parallel with the other main Chebyshev's disciple Lyapunov who in 1901 proposed a proof of the Central Limit theorem based on characteristic functions.

As mentioned before, Markov had a tough character, and very young, he proclaimed his hostility to the czarist regime. He was nicknamed *Andrei neistovy* (*André the enraged*) and is was only by the constant support of Chebyshev that he avoided major troubles. His protestations were directed against academic nominations of people linked to the czar, against the Academy's servility to the regime, against discriminatory measures against Jews. In 1910, when the Holy Sinod excommunicated Tolstoy, Markov worked hard to obtain the same fate, and eventually obtained it. Besides, it is on a background of religion and of philosophical rivalry between Moscow and Saint Petersburgh mathematical schools that Markov developed an interesting polemic with the mathematician Pavel Alekseevitch Nekrasov (1853-1924). He was a gifted analyst from Moscow, who was fond of philosophical interpretation of mathematics - and also a strong supporter of the czarist regime and of the most conservative tendency of the Orthodox church. In 1902, Nekrasov published a paper, called *Philosophy and Logic in the study of mass phenomena in human activity. Considerations about the foundations of Quetelet's social physics*. He asserted in the paper that the frequent observation of relative average stability in social phenomena (such as criminalty figures), along with the *necessity* of the independence of random events for the law of large numbers to be satisfied, can be interpreted as a proof of the existence of human free-will. Moreover, the assertion is mixed with a gabble of religion and nationalism. Markov could not resist the temptation of affronting Nekrasov on this battlefield and he seems to have taken this opportunity to look frantically for an example of *dependent* variables satisfying the law of Large Numbers.

And in 1907 was published in the Bulletin of Kazan Physical and Mathematical Society a paper by Markov called *Extension of the* *law of large numbers to quantities depending on each other*. Nekrasov was never explicitely quoted, either as a mathematical reference (only Chebyshev was mentioned) or as a philosophical point of view but it transparent that the last sentence of the paper is dedicated to him.

Markov's paper is a series of examples of sequences of random variables for which, despite the dependence between the variables, the law of large numbers remains true. The second example given by Markov can be seen as the archetype of Markov chains. In a sequence of random experiments, the variable x_i is the indicator of an event A occuring at the i-th run, and one supposes that the probability that x_{i+1} is 0 or 1 depends only on the value 0 or 1 of x_i, *whatever could have been the results of the previous experiments*. Precisely written in the text, this sentence therefore represents the first explicit formulation of Markov property in a formal framework. In the last section of the paper, Markov mentioned that it would be possible to generalize this example by considering not the indicator of an event, but a sequence of random variables whose conditional distributions satisfy this property of independence with respect to the previous runs.

During the following years, Markov constantly returned to his model of chain for generalization and studies of the validity of limit theorems in this framework (especially the Central Limit theorem). Among the generalizations, he proposed in 1910 a first model of an inhomogeneous chain.

Two years later, Markov realized a very original work. Having always been a great lover of litterature, he decided to choose as a source of randomness the novel in verses Eugene Oneguin by Pushkin . More precisely, he considered the 20000 first letters of the text (after having exluded the hard and soft signs whose systematic appearance is determined by the previous letter) and took them as

20000 random experiments with the result either vowel or consonant. He ordered them in groups of one hundred (in ten lines of ten letters), the columns of the table two by two (1 and 6, 2 and 7 etc.) and counted the vowels in each one ; he eventually ordered the obtained figures in 40 tables with 5 columns plus a column and a row corresponding to the sums. Every final table therefore represented 500 letters. In the last row, we read the total of vowels in successive hundreds. The last column contains the total of vowels when one groups columns 1 and 6, 2 and 7 of every hunded of the five-hundred. Markov realized a statistical test on the totals in rows and then in columns to show that the letters (that is the fact of being a vowel or a consonant) chosen to form the last rows may be roughly considered as independent (from one hundred to the other, the links are not tight) whereas those taken for the last columns are not independent (there is a strong connection between the status of a letter and the status of the following letter). More precisely, Markov showed that the results are quite close to those obtained if one supposed that the letters' status followed a Markov chain distribution with a given conditional probability for a vowel to follow a vowel or a consonant.

Oddly enough, despite Markov's good reputation at the beginning of the Soviet regime, his works were not especially widespread ; only Bernstein in 1926 published a long but unnoticed paper in French in the *Matematische Annalen* about extensions of theorems obtained by Markov and Liapounoff for the chains he calls *of A.Markoff*. The paper remained unnoticed and so did this first attempt of naming Markov chains : an important notion cannot seriously be saddled with a first name !

b) Poincaré

Let's turn now to Poincaré. His attitude towards probability was always somewhat ambiguous and it is reasonable to think that for him probability was never considered as real mathematics. But, the

increasing presence of probability in modern physics - especially statistical mechanics - forced him to accept to use randomness and to consider that the calculus of probability derives its scientific value from the fact that randomness sometimes rules over the world by producing fortuitous phenomena, a conception he found in Cournot. In this spirit, Poincaré wrote a textbook for his students of Mathematical Physics in 1896. This textbook had a famous second edition in 1912 , few months before his premature death. In this second edition, in the chapter XVI called *Various questions*, Poincaré developed an elementary example of ergodic evolution towards a uniform distribution in card shuffling .
Poincaré presents the successive orders of the cards as events in chain through the transition probabilities. He proves that the transition matrix P is diagonalizable, and that the eigenvalues are all with modulus less than 1 except 1 itself, which enables him to prove the convergence of the powers of P towards an eigenvector associated to the eigenvalue 1, therefore the uniform distribution. Let us observe that Poincaré learnt about Perron-Frobenius theory of nonegative matrices which includes his results only after the book was finished (he himself made a footnote about the fact).
The immediate posterity of Poincaré was not very interested with these results. Only Borel quickly reacted by proposing an interesting extension of Poincaré's results as early as 1912 ; but Borel's note remained unnoticed and it seems that these considerations on card shuffling stayed under an eclipse for 15 years.

2. An outsider : Bohuslav Hostinsky.

It is from another direction that a decisive impulse was given from which emerged the modern theory of Markov chains. One of its major characters is a Czech, Bohuslav Hostinsky, about whom one can say that his geographical localization between German science and the French sphere of influence (especially after World War 1, due to

the political proximity of the new Czechoslovakia and France) allowed him to realize a junction between several scattered elements. Born in Prague in 1884, Hostinsky is representative of a certain class of Czech intellectuals who, in 1919, wanted to participate and exalt in the creation of the new Czechoslovakia. I shall not have enough time here to give details about his life, so I shall mention several papers we have devoted to him. Let me now just say some words to prove that Hostinsky had the luck to be present at this very particular moment of the 1920s when the center of gravity of probability was sliding towards the East, from Paris to Moscow.
After having devoted his efforts to differential geometry, Hostinsky was progressively interested in several physical models, in particular through Borel's geometric approach (Borel published in 1914 a book on the subject). Hostisnky himself mentioned how from 1915 on he began to study Boltzmann's works and to be interested in the efforts made to give precise mathematical bases to the kinetic theory of gases. Hostisnky was particularly impressed by Poincaré's writings and began in 1917 to think about probabilistic questions.
His first work concerns a solution of Buffon's needle problem by use of Poincaré's method of arbitrary functions in order to avoid the usual physically unrealistic hypothesis of the needle falling anywhere on the infinite plane. He proved that the classical result can be seen as an asymptotical result when the distance between two grooves of the parquet tends to 0. Hostinsky published this paper in Czech in 1917 but had the fortunate idea to send its translation into French to Picard in 1920 for publication in a French journal. That was a good intuition because a wave of czechophilia wave was submerging French society at this moment, especially in Strasbourg where several individuals, such as Fréchet, had begun to look for possible contacts in Eastern and Central Europe. Picard immediately published the paper in the *Bulletin des Sciences Mathematiques*. Hostinsky was moreover present in Strasbourg in

September 1920 among the important Czechoslovakian delegation for the International Congress of Mathematicians. At this occasion, he began a huge scientific correspondence with Fréchet. During the 1920s, Hostinsky worked on generalization of the method of arbitrary functions.

Hadamard in 1927 began to teach probability at the Ecole Polytechnique in Paris. Not completely satisfied by Poincaré's considerations on card shuffling, he wrote a short note on the subject to the Academy of Science, where he reobtained more or less the same results as Borel fifteen years before. When he read this note, Hostisnky himself sent in January 1928 a note to the Academy where he extended the ergodic result to the case of a chain of events with continuous states. Apart from Bachelier's works, which we know to have been completely forgotten by mathematicians, it seems that this was the first occasion for the formulation of a continuous state chain before Kolmogorov's formalization some years after. Hadamard was immediately interested by Hostinsky's note and both began several months of exchanges on card shuffling and ergodic theorems, resulting in a series of notes on the subject. This ergodic spring was in fact the unique occasion for Hadamard to be involved in probability and marked the beginning of Hostisnky's international career.

3. Bologna congress and its consequences

We therefore are at Summer 1928 and on September 3rd opens the International Congress of Mathematicians in Bologna with his Excellency, the Head of Government Benito Mussolini as honorary president.

It is a mythical congress. First, the Germans are back after the long parenthesis following First World War. Pincherle, as president of the Italian mathematical society, had unceasingly worked to achieve this goal. This did not go without sulkings and Picard, for instance,

remained in Paris. Second, the Soviet mathematicians were still there and it was to be the last time (at least at this large scale) before Stalin's death in 1953. Moreover, the Italians, who were obviously well represented, had a specificity : among them, specialists in probability (such as Cantelli) were closely related to the actuars. Now, it seems to have existed what can be called a specific culture of the actuaries, which for instance had kept memories of Bachelier's works.
In brief, in 1928, everybody was in Bologna. And the encounter happened ! At least it is romantically seductive to see it like that. The French school and its adherence (including Hostinsky for instance) brought up the results about ergodicity for events in chain, Bernstein and Polya brought up Markov's works and their own extensions, actuaries brought up Bachelier's memory.
Naturally, we do not know to what extent the encounter was as powerful as this coulourful description. But what is sure is that after Bologna, Markov chains became a subject of major interest in probabilistic studies. In 1929, Romanovsky and Hostinský imposed the appellation. In 1931, Hostinsky published a volume of the Mémorial des Sciences Mathématiques where he exposed all the results he obtained about the the theory of events in chain and the method of arbitrary functions. The same year was published Kolmogorov's masterpiece about the analytical treatment of Markov process. Fréchet, whom Hostinsky had very early convinced of the interest of Markovian studies as we have seen, began to study them with care : in 1932, he obtained a first classification of chains through the status of the states. Next step seems to have been when a young German refugee, Wolfgang Döblin , took the subject over in 1935 and achieved the definitive classification of discrete chains, introducing in particular the impressive method of coupling. But this is already another story and I shall now stop on this opening. Thank you for your attention.

RÉFÉRENCES

[1] B.BRU : Souvenirs de Bologne, J.Soc.Fran.Stat, 144, 1-2 ; 2003

[2] V.HAVLOVA, L.MAZLIAK et P.ŠIŠMA : Le début des relations mathématiques franco-tchécoslovaques vu à travers la correspondance Hostinský-Fréchet, *Electronic Journal for History of Probability and Statistics* (www.jehps.net), Vol.1, 1, 2005

[3] C.HEYDE and E.SENETA (Ed) : Statisticians of the Centuries, Springer, 2001

[4] L.MAZLIAK : On the exchanges between Wolfgang Doeblin and Bohuslav Hostinský , Rev.Hist.Maths, 13, 155-180, 2007

[5] M.PETRUSZEWYCZ : Les chaînes de Markov dans le domaine linguistique, Slatkine, 1981

[6] E.SENETA : Statistical Regularity and Free Will : Quetelet and Nekrasov, Int.Stat.Review, 71, 319-334, 2003

[7] O.B.SHEYNIN : Markov's Work on Probability, Arch.Hist.Ex.Sci., 39, 337-377, 1988