BRUNO DE FINETTI: THE MATHEMATICIAN, THE STATISTICIAN, THE ECONOMIST, THE FORERUNNER♣

> *"A Mathematician who is not also 'a poet' is not a good Mathematician."*
>
> KARL WEIERSTRASS

CARLA ROSSI

University of Rome "Tor Vergata",

Department of Mathematics,

Via Ricerca Scientifica

00133 Rome, Italy

Tel. +39 06 72594676, Fax +390672594699

E-MAIL ADDRESS: rossi@mat.uniroma2.it

♣ I want to thank the Scientific Committee of ISCB2000 for inviting me to talk about Bruno de Finetti's work in Trento, where he spent years during his childhood and adolescence, and Fulvia de Finetti for allowing me to read and use her father's letters and for our very interesting and useful conversations.

Summary

Bruno de Finetti is possibly the best known Italian applied mathematician of the XX^{th} century. But, was he really just a mathematician? Looking at his papers it is always possible to find original and pioneering contributions to the various fields he was interested in, where he always put his mathematical "forma mentis" and skills at the service of the applications, often extending standard theories and models in order to achieve more general results. Many contributions are also devoted to educational issues, in mathematics in general and in probability and statistics in particular.

He really thought that mathematics and, in particular, those topics related to uncertainty, should enter in everyday life as a useful support to everyone's decision making. He always imagined and lived mathematics as a basic tool both for better understanding and describing complex phenomena and for helping decision makers in assuming coherent and feasible actions. His many important contributions to the theory of probability and to mathematical statistics are well known all over the world, thus, in the following, minor, but still pioneering, aspects of his work, related both to theory and to applications of mathematical tools, and to his work in the field of education and training of teachers are presented.

1. Introduction: Who was he?.

Bruno de Finetti is possibly the best known Italian applied mathematician of the XX^{th} century. But, was he really just a mathematician? Looking for quotations I have found him mentioned in journals, magazines and even, recently, in some newspapers as a "statistician", as an "economist", as the "founder" of the subjective theory of probability and several other definitions. Philosophers also took an interest in his fundamental contributions to the basic principles of the theory of probability and devoted several publications[1] to this issue, whilst the Department of History of Philosophy of the University of Pittsburgh is presently organising a section of the Archives of Scientific Philosophy devoted to Bruno de Finetti. Thus, who was Bruno de Finetti? First of all, he was a very impressive and complex figure. There was something harmonious in his personality and in his dealing with people, either with his students, as I was at the beginning, or with his colleagues, as I became later.

Reading his papers, it is always possible to find original and pioneering contributions to the various fields he was interested in, where he always put his mathematical "forma mentis" and skills at the service of the applications, often extending standard theories and models in order to achieve better and more general results. Many of his papers are also devoted to educational issues, both mathematics in general and probability and statistics in particular.

He actually believed that mathematics, specifically those aspects related to uncertainty, should be part of everyday life as a useful support for decision making:

"Our world, our life, our destiny, are dominated by Uncertainty; this is perhaps the only statement we may assert without uncertainty" he wrote in 1979[2].

He always imagined and lived mathematics as a basic tool both for better understanding and describing complex phenomena and for helping decision makers in assuming coherent and feasible actions. His personal view about Mathematics was critical in comparison to most mathematicians of his time, placing, as he did, an extremely high value on intuition rather than reason. In 1965 he wrote:

"Another preconception and another reason for abstract reasoning for many people is the wish "to ban intuition, because it may sometimes lead one into error". This might be justified in questions related to fundamental principles. However, these exceptional situations apart, a far greater risk of

error comes from lack of control of intuition, rather than from its imperfections when it is used. To abolish it would be rather like blinding yourself because of some "optical illusions", without considering that being blind is not really convenient" [3].

And later, in 1979[4]:

"Reason is but a pole, apt to aid the tree of our intuitive thinking to grow upright, or- in a different image - the salt giving some flavour to bread; however, the pole itself in not a tree, and no bread may be prepared with salt only"

His important contributions to the theory of probability and to mathematical statistics are well known all over the world, thus, I prefer to deal with "minor", but nevertheless pioneering, aspects of his work related both to theory and to the application of mathematical tools and to his work in the field of education and training teachers. In particular, I will deal with his first papers in population genetics, with some "minor" contributions to statistics and with his fundamental approach to mathematical education which has informed all my professional life.

2. The ugly duckling.

Bruno de Finetti started his academic life as a student of engineering in Milan for 2 years. Then, after a hard battle with his mother, who preferred him to become an engineer like his father and his grandfather, he succeeded in moving to the new Faculty of Mathematics. The echoes of the painful battle, just won, can be found in a very impressive letter written to his mother just after the final step was completed[°]. In this letter his passion for Mathematics is brightly expressed and the programme for all his academic life, already identified, appears in a vivid and distinct description:

"… Mathematics is not by now a field already explored, just to learn and pass on to the posterity as it is. It is always progressing, it is enriching and lightening itself, it is a lively and vital creature, in full development and just by these reasons I love it, I study it and I wish to devote my life to it, I wish to understand what professors teach not just to teach to others, as this can be done by anyone who is not completely stupid, I swear that well further I will proceed because I know that I can and I want to do it…."

And describing his feelings he wrote in the same letter (he was just 19 years old):

"Just now I feel to be as the ugly duckling, among my true companions, now I see forward a free, quiet and bright horizon where reality is more beautiful than any dream."

He worked hard, both in study and in early research and in 1926, whilst still an undergraduate student, he published his first paper in applied mathematics (population genetics). This developed a mathematical model, based on a system of non-linear differential equations, to mirror the diffusion process of hereditary traits, determined by a pair of autosomal alleles in a locus, in a mendelian "panmictic" (random mixing) population. The paper, written in Italian, appeared in METRON[5]. It is actually the first example of a model with overlapping generations in the field. In the following year he generalised the model to take into account non-random mixing based on phenotypic assortative mating[6]. It was then necessary to wait some 45-50 years to find further examples of models with overlapping generations in the field of population genetics. The first comparable contributions, though actually less general and elegant, appeared only during the '70s[7], but, as a matter of fact, in the same period Bruno de Finetti, in a joint paper[8], presented new generalisations of his first models incorporating age structure and thus obtaining a system of integral equations. Again "innovation"!.
All his life is a long series of scientific pioneering creations. He wrote more than 250 papers and books in various fields, mainly, but not only, in applied (and pure) Mathematics. They were, however, always in the framework of a "problem solving" approach[°], with the goal of developing new general and elegant tools useful to obtain both a global and a pragmatic view of the field of interest. His theorem, the concept of coherence, the definitions of exchangeability and partial exchangeability, are just examples of his skill which may be defined as "global mathematical intuition". Always he put passion in his work. Here is a short description, written by him to his mother, of the "agony and the ecstasy" which is typical in creating new mathematics (the work mentioned is actually his first work of 1926).

"Every word, every formula in the work I have done is blood of my blood, is the fruit of strong-willed inebriation and deep and creative pain. And if I, who have no patience and hate patience, patiently wrote and thought, it means that I was persuaded that it was worth making the effort, that I really felt the desire, even the need, to see all that was germinating inwardly and was making me suffer completed. So my patience was made of impatience and this intimate pain was purifying and converted into an indefinite joy in which I felt myself dreamily swimming."

[°] The letters are in press in Nuncius, Annali di Storia della Scienza, XV, 2, 2000.

His curiosity for the world and for its inhabitants and his ability to express in logical and mathematical terms, always within a unitary view, quite complex situations, permitted him to become a distinguished and original expert in many fields of application: economics, social sciences, statistics, insurance sciences and to be always and everywhere considered as a forerunner and innovator. He was well aware of this and, speaking about "the innovators", in 1976 he wrote[9]:

"The new voices are those which are able to see the problems globally, that means substantially as a whole, and not cut into pieces like anatomical sections."

What de Finetti wrote about the Bayesian approach to inference[10] highlights this approach:

"In order to avoid ambiguousness, it is necessary to remark at once and firmly an essential distinction between what may be respectively called Bayesian standpoint and Bayesian techniques.

Both apply to the problem of statistical (or probabilistic) inference, and are, of course, perfectly connected together as being the logical framework and the mathematical tool of the same theory: namely, of the theory concerning the way in which our opinions (or beliefs) must be modified (according to Bayes' theorem) when new information is attained. Nevertheless, in practice, the overlap of the fields of the published applications inspired to the Bayesian standpoint and of those making use of Bayesian techniques seems rather narrow.

In fact, most applications of Bayesian standpoint in everyday life, in scientific guessing, and often also in statistics, do not require any mathematical tool nor numerical evaluations of probabilities; a qualitative adjustment of beliefs to changes in the relevant information is all that may be meaningfully performed. And conversely, Bayesian techniques, more or less developed into imposing mathematical machinery, are often applied as such, using standardized "models" and standardized "prior distributions", instead of carefully keeping realistic adherence to the specific features of each particular case and to the true opinion of the person concerned (the statistician himself, or the decision maker, or somebody else).

Thus, the given distinction, fundamental per se, is also necessary for a preliminary explanation of the thesis maintained in the present paper, and of the succession of the aspects that will be considered and discussed here in order to clarify the point of view defended.

• His work experience in the insurance field, at the beginning of his career, is crucial for the development of this attitude and also of the subjective view as a basis for the theory of probability.

To begin with, let me express it very roughly as follows.

Bayesian standpoint is no-ways one among many possible theories, but is an almost self-evident truth, simply and univocally relying on the indisputable coherence rules for probabilities. It should be always applied in the most natural and naive form, paying attention-whenever the recourse to a more sophisticated machinery seems unavoidable-that its introduction should not induce to loose sight of the true situation and opinion.

At the contrary, Bayesian techniques, if considered as merely formal devices, are no more trustworthy than any other tool (or "ad hoc" method, or "Adhockery" to use the word introduced by Good) of the plentiful arsenal of the "objectivist Statistics". It is true that Bayesian techniques give rise to all (and only all) the <u>*admissible*</u> *decision rules (according to Abraham Wald) but each one is valid with reference to a particular initial opinion; therefore, any conclusion is arbitrary if the choice (purposedly or unadvertedly) responds to arbitrary formal criteria (inspired e.g. to simplicity, or to mathematical convenience) rather than to the personal advice. And, moreover, two or more admissible decisions may constitute, together, an* <u>*inadmissible*</u> *compound decision when based on incompatible initial opinions instead of the same one."*

This concern for the problems of the real world and society did not allow him to stay in an ivory tower as most mathematicians of his time used to do. He was also involved in politics. During the '70s he was arrested for antimilitarist positions and he was also a candidate in several elections. This concern also generated his involvement in the problem of mathematical education to which he devoted attention and interest, with more than 30 publications specifically dealing with this issue.

A very "special" peculiarity of Bruno de Finetti was his ability and delight in inventing new and evocative words and sentences aimed at enlightening. He used jokes to mock incoherence, biases or falsities and unpleasant features of several kinds in various fields. Some examples from economic or social works are:

- "Press of deformation" (instead of press of information) to highlight the peculiarity of most Italian journalists to "interpret" the facts as they like them instead of reporting as they are.

- "Bureau-phrenia" to mean the gigantism and deformation of our national bureaucracy.

- "Idiot-cracy" to mean the government of idiot politicians or managers.

Many others can easily be found reading here and there. To justify this habit he wrote in 1976[11]:

"Things said in a too colourless, painless, discrete, polite way, have no influence at all when they collide with preconceived ideas and prejudices. By contrast, a "strong" word can be useful for destroying the mental defences of fossilised and adulterated ideas which constitute the background of the so called "established order" (more properly called "established disorder" by genuine right-thinking people). Thus, I think that creating mocking and funny words or sentences is a powerful, necessary and praiseworthy weapon for discrediting, fighting and pulling down the despicable conformity which privileges maintenance of the worst raging institutions and ideas now in force."

Even in many scientific pages some special terms from the common language are used by him as "strong" words to highlight some special features of approaches he did not agree with, joking with language and concepts and mocking. In the appendix one of the most impressive examples (dealing with the "mass effect" of statistical units in a frequentist framework), which I will try to do my best to translate without missing either the meaning or the joke[12], is reported[13].

Another peculiarity of Bruno de Finetti was his ability to use significant, easy to understand and evocative graphical representations. Taking a simple example, he often used the triangular diagram (sometimes actually mentioned in the literature as de Finetti's diagram) to represent the probability (or statistical distribution) of a characteristic with 3 outcomes (Figure 1). This graphical representation originates indeed from chemistry as it is used to represent the alloys of 3 metals. Bruno de Finetti used it first during his studies in engineering and presented it in his first paper to show the genotype distribution in the population under study. The triangular diagram is an equilateral triangle (A_1,A_2,A_3) with unit height; if we put a mass proportional to p_1 at A_1, proportional to p_2 at A_2, proportional to p_3 at A_3 and take the centre of gravity of the three masses, we get a point which has a distance p_1 from the side opposite to the vertex A_1, p_2 from the side opposite to the vertex A_2, p_3 from the side opposite to the vertex A_3. Inside the triangle such a point is the graphical representation of the distribution (A_1,A_2,A_3:p_1,p_2,p_3). He continued using the diagram in all his papers on population genetics and also in many other situations.

Triangular Diagram

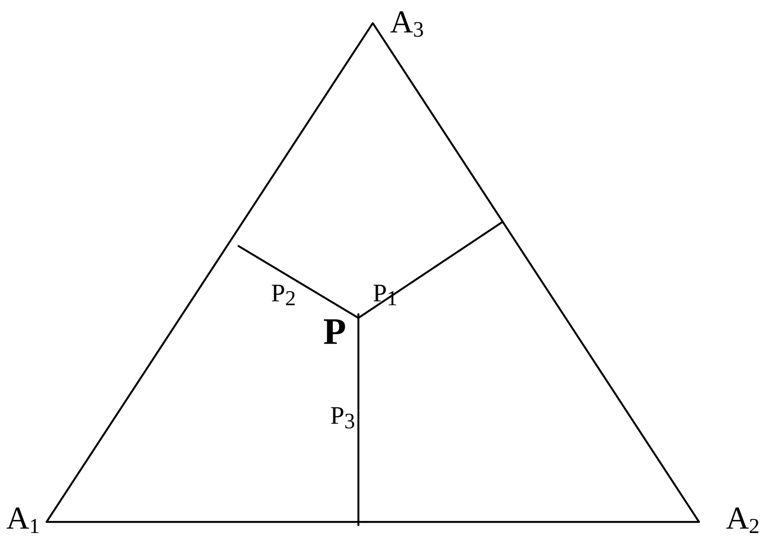


Figure 1. The triangular diagram which allows to represent the distribution of a random variable with 3 outcomes (A_1, A_2, A_3) with probabilities p_1, p_2, p_3 by the point P with a distance p_i from each side opposite to the vertex A_i.

Another example can be found in Figure 2 where the graphical representation, which he prepared for our 1976 joint paper to explain the differences between models with overlapping and separated generations, appears. His intuition in this case was to project backwards and look at the ancestors of a fixed individual, instead of projecting forwards to study the successors. This allowed a quite simple, though complete and significant, representation. Through it, the unique relevant difference between the two approaches can be shown and the direction of the time axis has no influence.

— 30 —

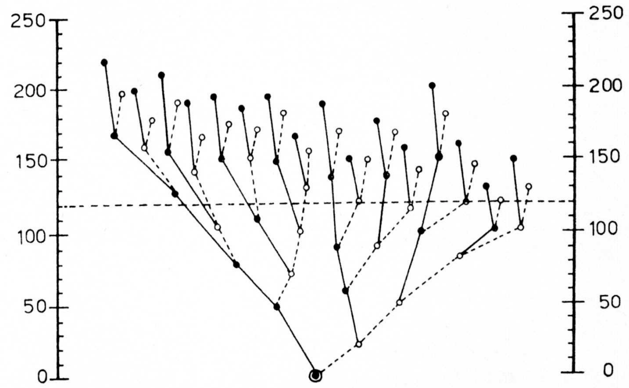


Fig. 1 – Gli ascendenti di un individuo: maschi • e femmine o, nel tempo.

— 31 —

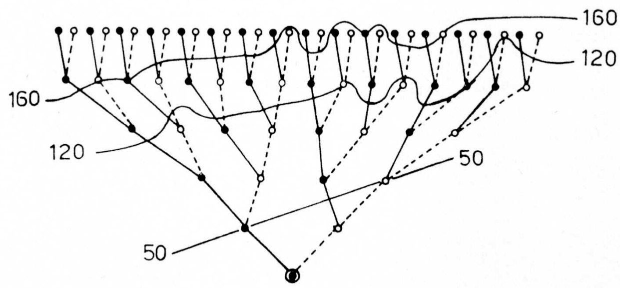


Fig. 2 – Gli ascendenti di un individuo: maschi • e femmine o, per generazioni.

Figure 2. Different approaches to modelling in population genetics: the models with overlapping generations and the models with separated generations (backward projections).

3. Early papers in population genetics: the curiosity of a young mathematician.

His 4 papers devoted to population genetics were published in 1926 and 1927[14]. He started working on the subject after reading a book by P. Enriques and a paper published by the biologist Carlo Foà. In this paper the mendelian laws and the diffusion of hereditary traits, determined by a single pair of autosomal alleles A and a with genotypes AA, aa, Aa, in a general population were presented from a biological point of view. Bruno de Finetti set up proper mathematical models (deterministic models based on systems of ordinary differential equations) to mirror the biological phenomenon both for the simpler case of random mixing and for the case of assortative mating. In particular, he considered in depth the case with A dominant, thus assortative mating based on the phenotypes: D=AA+Aa (dominant) and R=aa (recessive). For the random mixing model, he obtained many analytical and elegant (from a mathematical point of view) results. These concerned, in particular, equilibria and dynamical aspects, such as the Hardy-Weinberg law[♥] (already established for models with separated generations), the law of the constant difference (the difference in the proportion of homozygotes remains constant), the law of non-reversibility (each trajectory of the process cannot pass twice through the same points), just to mention the most significant. The 3 laws just mentioned allow us to say that all the equilibrium points belong to the Hardy-Weinberg parabola $p_3^2=4p_1p_2$, which can be represented inside the triangular diagram and contains the two vertices A_1 (homozygotes AA with coordinates 1,0,0) and A_2 (homozygotes aa with coordinates 0,1,0) and the internal polymorphic point with barycentrical coordinates (1/4,1/4,1/2). All the trajectories are vertical segments, parallel to the height of the triangle, originating from any initial point and directed towards the parabola.

The general case is much more complex and some analytical and some numerical results are available.

In particular, denoting by λ_D the assortative mating coefficient (relative probability) between dominants, λ_R the assortative mating coefficient between recessives and λ_M the assortative mating coefficient between dominants and recessives ($\Lambda=\lambda_D, \lambda_R, \lambda_M$), the following results were obtained:

- For any given Λ there exist only a finite set of equilibrium points.

[♥] The Hardy-Weinberg law says that: if $p_A°$ is the initial proportion of the allele A in the population and $p_a°$ is the initial proportion of the allele a and if the diffusion process is originated by random mixing, then the equilibrium point for the genotypes has baricentrical coordinates: $p_{AA}= (p_A°)^2$, $p_{aa}= (pa°)^2$, $p_{Aa}= 2p_A°p_a°$.

- If we want a given point of the triangle to be an equilibrium point there is only one way of choosing Λ.

- Some points of the diagram above the Hardy-Weinberg parabola correspond to $\lambda_R<0$, so they are unfeasible equilibrium points.

From numerical results some conjectures were set up:

- The feasible points above the Hardy-Weinberg parabola are stable equilibrium points for some choices of Λ. In particular they correspond to the case with $\lambda_M>\max(\lambda_D, \lambda_R)$ (bifurcation).

- The points under the Hardy-Weinberg parabola are unstable equilibrium points for some choice of Λ.

It was only after some 50 years that the above mentioned conjectures were proved, and even then not completely, and a characterisation of the stable and unstable equilibria was derived[15].

4. Predictive induction and prequential probability: the forecast-meter.

"A theory, in the sense of a law, is not a statement whose truth or falsity is objectively decidable. To find the correct answer would only be possible by opening the safe where the absolute truth of the platonic world is locked and checking whether that law is there or not. In fact, we can only observe observable outcomes (we can also possibly experimentally produce outcomes) and check whether they always appear in agreement with the theory. But, even if we could look at the past, the present and the future all at the same time, we could never decide whether a possible full agreement is due to the law or to chance…..Thus, it is just about observable outcomes, single observable outcomes, that we can consistently discuss. Only through observable outcomes, not yet known, we can consistently assess the probability.[16]"

In this paper the theory of predictive inference is clearly outlined, starting from the general considerations already mentioned. In simple words, it stresses that the only way to make consistent inferences is to focus on predictive distribution of observables rather than on posterior distribution of the "unknown" parameters. Intrinsic in these ideas was the fact that parameters (as laws) can never be observed. No one will be able to observe the mean of a normal density N(θ,1). Therefore, we can never be certain how well a construct, such as a mean, has been estimated or approximated. Much more relevant are observables or facts, that is quantities or events capable of being observed. Philosophically, this means that in the field of statistics we should focus solely on observables and use only predictive distributions to make inferences. From a technical point of view the procedure is based on de Finetti's theorem which allows easy expression of any predictive distribution as a mixture on the parameter space.

Let $X_1, X_2, \ldots X_n$ denote a sample of exchangeable quantities from a distribution depending on a parameter θ (possibly a vector parameter), with (generalized) density $f(x \mid \theta)$. If the prior density on the parameter space Θ is $p(\theta)$, then the posterior distribution is given by:

$$g(\theta \mid x_1, x_2, ..., x_n) \propto \prod_{i=1}^{n} f(x_i \mid \theta) p(\theta)$$

and the predictive density for a new observation Y is given by:

$$h(y \mid x_1, x_2, ..., x_n) \propto \int_{\Theta} f(y \mid \theta) g(\theta \mid x_1, x_2, ..., x_n) d\theta$$

Thus, even if parameters, such as hypotheses, are, strictly speaking, meaningless as events or observables, reference to them may be a valuable aid to express the probability of the real meaningful events or observables.

This predictive framework is the basis for the development of the prequential probability approach. This is based on the idea that we can judge the quality of an inference method by converting it into a forecasting system and assessing the empirical success of the sequence of one-step-ahead forecasts that it implies[17].

5. Complexity, mixture-models and non-parametric statistics: blurred and unfinished suggestions ….and recent developments.

Mixture distributions have always been presented by Bruno de Finetti as very powerful tools for modelling complex situations. As mentioned above, his theorem specifically sought to find the mixture representation of exchangeable observations. Mixture models were widely used in many of his papers and each year some of his lectures were devoted to this issue. Examples of applications ranged from compound Poisson processes to the bayesian approach to the rejection of outliers[18]. One of the most interesting applications of mixture models, which has found important developments recently, is the study of non homogeneous populations in various fields, in particular, but not only, in the framework of *survival analysis*. The most popular approach uses "Frailty models" which, in the original well known formulation, cannot be dealt with using mixture distributions. The alternative approach uses mixture models both in a parametric and in a non-parametric framework. In the first case, using the machinery set up for bayesian parametric estimation, namely the conjugate families, it is possible to take into account heterogeneity in various situations. In particular, considering one or more parameters, characterising the likelihood function, as random variables with "prior" probability distribution function belonging to the conjugate family, the model can be expressed in terms of the predictive distribution obtained by integrating the likelihood with mixing distribution given by the prior distribution. Thus, generalised models, accounting for overdispersion due to heterogeneity, can be easily set up. Just as an example Table 1 contains some of the simplest models obtained this way and used in the literature, mostly in bio-medical applications.

Table 1. Some parametric mixture model for overdispersion.

Likelihood	***Prior***	***Predictive (heterogeneity model)***
Poisson(θ)	Gamma(α,λ)	Negative binomial
Exponential(θ)	Gamma(α,λ)	Pareto
Binomial(θ)	Beta(α,β)	Beta-binomial
Weibull(c,θ)	Gamma(α,λ)	Weibull-gamma

Even more interesting is its application for the detection of cured individuals in survival analysis[19]. In such a case, we can consider the mixture of two components with likelihoods representing survival functions of risk and cured individuals respectively. These latter correspond to a survival function S(t) identically equal to one, whereas the others are represented by some parametric or non parametric model with S(t) tending to zero for t increasing. The possible situations which can be modelled this way have been extensively reported in recent papers.

Another well known application of the general mixture approach is the "Back-Calculation" methodology. This has broadly developed in the framework of HIV/AIDS modelling, estimation and forecast, where the random variable of the mixture is the time interval between HIV infection and the development of AIDS and the mixing distribution is the incubation period distribution of AIDS[20].

Further suggestions relate to *non parametric statistics* and I will mention just one. During a seminar in Rome in 1970, Lester Dubins presented his method for assessing non parametric priors, that is, constructing a random distribution function, i.e. a stochastic process with increasing trajectories going from 0 to 1. His approach was based on random subdivision of the interval (0,1) and was not completely satisfactory[21]. At the end of the seminar Bruno de Finetti suggested trying to use conditional stable stochastic processes with characteristic exponent $\alpha<1$. In the following years two of his students extensively studied the case with $\alpha=1/2$, conditional to the maximum level, say a. Such a process, suitably normalised, is a "random distribution function" in the sense of Dubins. Thus a different approach to the problem can be outlined. They found some interesting general results which appear in a paper[22] published in 1972. Unfortunately no one has continued this work and the approach is still awaiting developers. I can't say whether the alternative approach by means of neutral to the right distributions, developed in the same period, is more attractive and elegant, but it would be interesting to continue investigation of the application of stable processes. I could continue to mention further late, but pioneering, suggestions resulting in finished and unfinished studies by some of Bruno de Finetti's students but it would take too much time and would mean neglecting another important aspect of his work.

6. The approach to Mathematical education: non scholae sed vitae discimus.

The substance of Bruno de Finetti's approach and ideas on teaching can be found in each scientific paper even more than in the many works specifically devoted to this issue. Even more could be understood and learnt by observing his teaching or in discussion with him about any issue. An indication of his attitude can be discovered in the few words he used to describe Oscar Chisini's contribution, highlighting the meaning and importance of the concept of "mean"[23]:

"…he gave a quite suitable answer, that is the one which does not highlight the formal properties, but rather the conceptual nature and the operational meaning of the idea."

And about problems to take into consideration he used to say:

"..before approaching a problem to solve you need to see it…"

suggesting that it is only possible to solve a problem satisfactorily by strictly linking the technical tools to intuition in such a way as to preserve the global view of both: the rigorous reasoning and the core of the matter in developing the solution. These concepts are specifically dealt with in one of his best contribution to education, where his fundamental approach appears also in the title "Il saper vedere in matematica"[24]. There he wrote:

"In order that a subject of study, specifically Mathematics, does not appear sterile, obscure and useless, it should always be presented so that studying it is fully and genuinely justified"

and in 1954 he wrote[25]:

"One of the major faults (in mathematical education) lies in separating formal lecturing from the conceptual meaning and objectives…From a methodological viewpoint, just a misunderstood abstract approach pushes teachers to start with formal concepts. Genuine mathematical abstraction can only derive from concrete ideas and it aims at showing the common abstract concepts behind them…Abstracting means synthesising several notions having something, no matter what, in common. It does not mean using solemn foolish words and technicalities based on nothing, aimed at nothing…..A word which arouses a concrete image can help in understanding a theorem…the translation of that theorem, expressed in abstract language, into concrete images may be preferred as it can make it easier to understand its importance…."

Some of his late students cannot forget his lecture on Poisson distribution starting with the image of the probability distribution of the number raisins in a Panettone (it was Christmas time 1969).

These concepts are also linked to his approach to problem solving in general and to his criticism with respect to abstraction and technicalities per se, without clear objectives and motivations, in a word his "fusionist" approach both to scientific work and to teaching♠. The attitude of most teachers of his time (and even now, unfortunately) to breaking up mathematical thinking into small technical pieces reminds me of something which Oliver Sacks wrote in his delightful and interesting book "The man who mistook his wife for a hat". He talks about his patient, who had lost the ability to recognise faces or objects as a whole but still had the ability to analyse correctly the details. I am

♠ Several strong and mocking words were invented to express such criticism but they are actually too Italian to be properly translated.

sure that Bruno de Finetti would highly appreciate reading the book, which was also mentioned by one of his collaborators in a paper about teaching[26].

Describing the situation of his patient, Sacks wrote:

"…'What is it?' I asked, holding up a glove,

'May I examine it?' he asked, and, taking it from me, he proceeded to examine it as he had examined the geometrical shapes.

'A continuous surface', he announced at last, 'infolded on itself. It appears to have – he hesitated – five outpouchings, if this is the word'.

'Yes', I said cautiously. 'You have given me a description. Now tell me what it is'

'A container of some sort?'

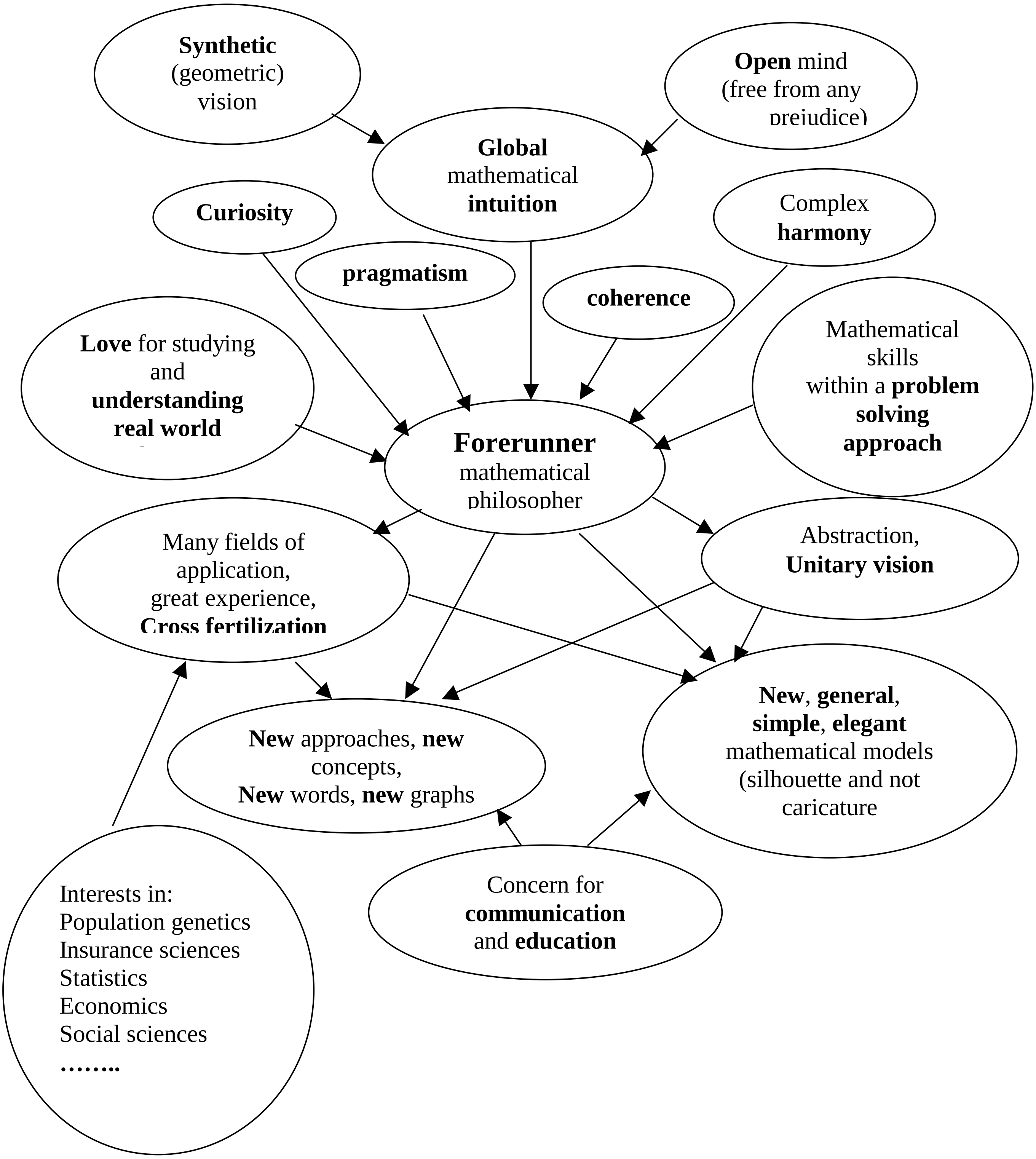
'Yes,' I said, 'and what would it contain?'

'It would contain its contents' …

...No child would have the power to see and speak of 'a continuous surface…infolded on itself', but any child, any infant would immediately know a glove as a glove, see it as familiar, as going with a hand."

In conclusion, Bruno de Finetti's approach to teaching can be summarised in few words: if you want to be a good teacher you should act in such a way that a student feels that abstracting, setting up axiomatic systems, formalising and deducing by logical rules is just the endpoint of his (her) experience which is needed to better highlight and simplify what he (she) has already learned, not to introduce useless technical complications. It is the avenue for discovering the unity which is behind the apparent diversity. It is the arrival point, as always happened throughout the history of mathematics, not the starting point.
These last sentences also summarise his philosophy of life and research. At the end of this story, a graph representation of his complex personality and experience is reported in short in Figure 3, through key words and links.

Figure 3. Who was he?
Key words and links



REFERENCE NOTES

[1] M.C. Galavotti "Sulla filosofia della probabilità di Bruno de Finetti", Rivista di Filosofia, LXXXIII, 1, 1991, 120-138; Erkenntnis, Special issue on "Bruno de Finetti's Philosophy of Probability", 31, 2-3, 1989.

[2] B. de Finetti "Insurance and the views about probability", Giornale dell'Istituto Italiano degli Attuari, 1979, 71-83.

[3] B. de Finetti "Programmi e criteri per l'insegnamento della matematica alla luce della diverse esigenze", Periodico di Matematiche, 1965, 119-143.

[4] B. de Finetti "Insurance and the views about probability", Giornale dell'Istituto Italiano degli Attuari, 1979, 71-83.

[5] B. de Finetti "Considerazioni matematiche sull'eredità mendeliana", METRON, 1926, 3-41.

[6] B. de Finetti: "Conseguenze statistiche delle leggi di Mendel", Rivista di Biologia, 1927, 525-553; B. de Finetti "Conservazione e diffusione dei caratteri mendeliani. Nota I. Il caso panmittico", Rendiconti della R. Accademia del Lincei, 1927, 913-921"; B. de Finetti "Conservazione e diffusione dei caratteri mendeliani. Nota II. Il caso generale", Rendiconti della R. Accademia del Lincei, 1927, 1024-1029.

[7] B. Charlesworth "Selection in populations with overlarlapping generations. I. The use of Malthusian parameters in population genetics", Theoretical Population Biology, 1970, 352-370; B. Charlesworth "Selection in populations with overlarlapping generations. II. Conditions for genetic equilibrium ", Theoretical Population Biology, 1972, 377-395; J.L. Cornette "Some basic elements of continuous selection models", Theoretical Population Biology, 1975, 301-313.

[8] B. de Finetti & C. Rossi "Diffusione di caratteri mendeliani in relazione a fattori variabili" Atti dei Convegni Lincei, 1976, 29-43.

[9] B. de Finetti "Dall'utopia all'alternativa", Franco Angeli, 1976.

[10] B. de Finetti “Bayesianism: its Unifying Role for Both the Foundations and the Application of Statistics”, Proceedings of the 39th Session, Bulletin of the International Statistical Institute, International Statistical Institute, 1973, 349-368.

[11] B. de Finetti “Dall’utopia all’alternativa”, Franco Angeli, 1976.

[12] B. de Finetti “Teoria delle Probabilità”, Einaudi, 1970, vol. 2.

[13] B. de Finetti “Theory of Probability”, Wiley, 1975.

[14] B. de Finetti “Considerazioni matematiche sull’eredità mendeliana”, METRON, 1926, 3-41; B. de Finetti “Conseguenze statistiche delle leggi di Mendel”, Rivista di Biologia, 1927, 525-553; B. de Finetti “Conservazione e diffusione dei caratteri mendeliani. Nota I. Il caso panmittico”, Rendiconti della R. Accademia del Lincei, 1927, 913-921”; B. de Finetti “Conservazione e diffusione dei caratteri mendeliani. Nota II. Il caso generale”, Rendiconti della R. Accademia del Lincei, 1927, 1024-1029.

[15]C. Rossi “Asymptotic behaviour of the mean of a phenotypic diffusion process”, Annals of Human Genetics, 1976, 245-251; C. Rossi “Equilibrium stability of a diffusion process in population genetics”, Annals of Human Genetics, 1979, 187-192.

[16] B. de Finetti “Probabilità di una teoria e probabilità dei fatti”, in Studi di probabilità, statistica e ricerca operativa in onore di Giuseppe Pompilj, Oderisi, Gubbio, 1971, 86-101.

[17] A.P. Dawid & V.G. Vovk “Prequential probability: principles and properties”, Bernoulli, 1999, 125-162; B. de Finetti “Il Previsiometro: presentazione e spiegazione in forma elementare”, in Studi di Mercato, 1968, 3-12.

[18] B. de Finetti “The Bayesian approach to the rejection of outliers”, Proceedings of the Fourth Berkeley symposium on Mathematical Statistics and Probability, University of California Press, 1960, 199-210.

[19] Y. Peng & K.B.G. Dear “A nonparametric mixture model for cure rate estimation”, Biometrics, 2000, 237-243; G. Schinaia & C. Rossi “EM estimation of diagnosis”, Biometrical Journal, 2000, in press.

[20] Rosemberg P., Gail M.H., Backcalculation of Flexible Linear Models of the Human Immunodeficiency Virus Infection Curve. Applied Statistics 40, 269-282, 1991.

[21] L. Dubins & D.A. Freedman “Random distribution functions” Fifth Berkeley Symposium, 1963.

[22] G. Del Grosso & F. Fabi “Processi stabili condizionati e una loro applicazione”, Giornale dell’Istituto italiano degli Attuari, 1972, 112-123.

[23] B. de Finetti “Oscar Chisini e il suo insegnamento”, Periodico di Matematiche, 1968, 26-33.

[24] B. de Finetti “Il saper vedere in matematica”, Loescher, 1967.

[25] B. de Finetti “E’ difficile capire la matematica?”, Archimede, 1954, 6-12.

[26] D. Fürst “de Finetti e l’insegnamento della matematica”, in ‘Probabilità e Induzione’, edited by Paola Monari & Daniela Cocchi, Editrice Clueb, Bologna, 1993, 17-28.

APPENDIX: DOES STATISTICAL EVIDENCE NEED A HEAP[12]?

"…si tratterebbe di una proprietà legata all'esistenza di un mucchio: finché si hanno pochi oggetti essi non costituiscono un mucchio e nulla si potrebbe concludere, ma se sono molti il mucchio c'è e allora, ma soltanto allora, tutto il ragionamento fila. Se si pensa di aggiungere un oggetto per volta, nulla si potrà dire finché il numero è insufficiente per formare un mucchio, e la conclusione balzerà fuori (d'improvviso? Passando da 99 a 100? o da 999 a 1000?…!) quando finalmente il nonmucchio si trasforma in mucchio. No, si dirà; questa versione è caricaturale; non c'è un salto netto, bensì sfumato; il nonmucchio attraverserà una fase di forsechesìforsechenomucchio da piùforsechenocheforsechesìmucchio a piùforsechesìcheforsechenomucchio e solo poi diverrà gradualmente un vero mucchio. Ma ciò non toglie il difetto d'origine, cioè la distinzione, concettualmente posta come fondamentale, tra "effetto di massa" e "effetto dei singoli elementi"; il riconoscere che non può esistere una separazione netta, se elimina forse, apparentemente, una circostanza paradossale, non ne estirpa la radice ed anzi mette in luce la debolezza e contraddittorietà del concetto di partenza.
Non se ne esce se non negando ogni distinzione del genere. La conclusione cui si giunge sulla base di una massa di dati è determinata non globalmente, come effetto di massa, bensì come risultante, come effetto cumulativo, dell'apporto di ogni singolo dato. Conoscere l'esito di un certo numero di prove, grande o piccolo che sia, conduce dall'opinione iniziale all'opinione finale esattamente nello stesso modo che si otterrebbe pensando di venire a conoscere l'esito delle singole prove, una per volta, e di modificare ogni volta l'opinione conformemente al (piccolo in genere) influsso di una singola informazione."

The translation below is mostly due to Adrian Smith and Antonio Machì[13], just few sentences have been slightly changed.

"…we would be dealing with a property connected with the existence of a heap♦. So long as we are dealing with just a few objects, they do not form a heap and no conclusions can be reached. If, however, we have a large number of objects, then we do have a heap, and then, and only then, does the argument go through. If we add an object one at a time, nothing can be said until the number of objects becomes sufficient to be considered a heap, then the conclusion appears (just like that? In passing from 99 to 100? or from 999 to 1000?….!) as that which is not yet a heap at last becomes one. Now it will be objected that this version is a travesty: there is no sharp break of this kind, but

♦ In technical terms it would be better denoted using the word "aggregate" (see the translation by Smith and Machì) but this way the joke is missing.

rather a gentle transition. The non-heap passes through fuzzy phases: not-to-be-or-to-be-a-heap phase, to-be-or-not-to-be-a-heap phase, and only subsequently does it gradually transform itself into a real and genuine heap. But this does not answer the original objection raised against the distinction, here put forward as being of fundamental conceptual importance, between the "mass effect" and the "effect of individual elements". To recognise that a clear-cut separation cannot exist, even though this admission may perhaps, just apparently, eliminate the evident paradox, does not eliminate the origin of it, and, indeed, highlights the weakness and the contradictory nature of the basic approach.

The problem is only resolved by acknowledging that distinctions of this kind have no significance. The conclusions one arrives at on the basis of a large quantity of data are not the consequences of some mass effect, but simply the cumulative effects of the contributions of each single object. The modification of the prior opinion into the posterior opinion through knowing the outcome of some given set of trials is precisely the same as that obtained by considering each item of data separately, and effecting the appropriate modifications (in general, minor) one at a time. This is so no matter whether the number of trials is large or small."