



---

# The Mathematics of Subjective Probability

Università Milano Bicocca, 3-5 September 2018

Book of Abstracts

---

Argenziano R., Gilboa I.,

*Learning What is Similar: Precedents and Equilibrium Selection*

We argue that a precedent is important not only because it changes the relative frequency of a certain event, making it positive rather than zero, but also because it changes the way that relative frequencies are weighed. Specifically, agents assess probabilities of future events based on past occurrences, where not all of these occurrences are deemed equally relevant. More similar cases are weighed more heavily than less similar ones. Importantly, the similarity function is also learnt from experience by "second-order induction". The model can explain why a single precedent affects beliefs above and beyond its effect on relative frequencies, as well as why it is easier to establish reputation at the outset than to re-establish it after having lost it. We also apply the model to equilibrium selection in a class of games dubbed "Statistical Games", suggesting the notion of Similarity-Nash equilibria, and illustrate the impact of precedents on the play of coordination games.

---

Burzoni M., Riedel F., Soner H. M.,

*Ambiguity, Knightian Uncertainty, Equivalent Martingale Measures*

We reconsider the microeconomic foundations of financial economics under Knightian Uncertainty. In a general framework, we discuss the absence of arbitrage, its relation to economic viability, and the existence of suitable non-linear pricing expectations. Classical financial markets under risk and no ambiguity are contained as special cases, including various forms of the Efficient Market Hypothesis. For Knightian uncertainty, our approach unifies recent versions of the Fundamental Theorem of Asset Pricing under a common framework.

---

De Blasi P.

*Asymptotics of the Number of Groups in a Sample from the Geometric Stick-breaking Process*

Random discrete probability measures define prior distributions which are used for estimation in mixture models and species sampling problems. The law of the frequencies can be described in terms of the random partition featured by a sample taken from the prior and, in particular, by its large sample behaviour as the sample size increases. By a result of Karlin, the asymptotics of the number of groups of the partition is determined by the rate of decay of the small frequencies, so that in the random frequency case it can be studied by a deconditioning argument whenever the frequencies in decreasing order admit a tractable form. The geometric stick-breaking process has been recently introduced as a simple yet effective alternative to the Dirichlet process, with frequencies that take the form of geometric probabilities with random probability of success. In this work we investigate the effect of the prior of this parameter on the asymptotic behaviour of the number of groups, and we show that a whole range of logarithmic behaviour can be achieved by appropriately tuning the prior. We also extend the analysis to a large class of priors that includes the geometric stick breaking process and features an additional parameter representing the number of failures of the negative binomial distribution.

---

Durante D.

*Global Mean-Field Variational Bayes for Density Regression*

Quadratic approximations of logistic log-likelihoods are fundamental to ease approximate Bayesian inference for binary data. Although the expansions underlying Newton–Raphson and Fisher scoring methods have attracted much of the interest, there has been also a recent focus on tangent quadratic bounds that uniformly minorize the logistic log-likelihood. A relevant contribution, within this class of approximations, relies on a convex duality argument to derive a tractable family of tangent quadratic expansions indexed by a location parameter. This approximation is still being successfully implemented to facilitate approximate Bayesian inference in several models, but less attempts have been made to understand the formal reasons underlying its excellent performance. To address this gap we provide a novel connection between the above bound and a recent Polya-gamma data augmentation for logistic regression. This result places the computational methods associated with the aforementioned bound within a more general and well-established framework (i.e. conjugate global mean-field variational Bayes) having desirable theoretical and computational properties. The practical advantages of such global routines are highlighted in a predictor-dependent mixture of Gaussians with logit stick-breaking weights for nonparametric density regression.

Flesch J., Vermeulen D., Zseleva A.,

*On the Equivalence of Mixed and Behaviour Strategies in Finitely Additive Decision Problems*

We consider decision problems with arbitrary action spaces, deterministic transitions and infinite time horizon. We assume that the decision maker has perfect recall. In the usual setup when probability measures are countably additive, a fundamental theorem (a general version of Kuhn's theorem, see. Aumann (1964)) implies under fairly general conditions that for every mixed strategy of the decision maker there exists an equivalent behaviour strategy, i.e., they induce the same probability measure on the set of plays. In this paper we examine to which extent this remains valid when probability measures are only assumed to be finitely additive.

The answer to this question depends on how we define the finitely additive probability measure that a behaviour strategy induces on the set of plays. In the classical approach by Dubins and Savage (2014), this is defined on the algebra of all clopen subsets of plays. Under this approach, we prove the following statements: (1) if the action space is finite, every mixed strategy has an equivalent behaviour strategy, and (2) even if the action space is infinite, at least one optimal mixed strategy has an equivalent behaviour strategy. The approach by Dubins and Savage turns out to be essentially maximal: roughly speaking, these two statements are no longer valid if we take any extension of the clopen algebra that includes all singleton plays. Our results suggest that mixed strategies may be more suitable to study finitely additive decision problems.

---

Argenziano R., Gilboa I.,

*Second-Order Induction: Uniqueness and Complexity*

Agents make predictions based on similar past cases, while also learning the relative importance of various attributes in judging similarity. We ask whether the resulting "empirical similarity" is unique, and how easy it is to find it. We show that with many observations and few relevant variables, uniqueness holds. By contrast, when there are many variables relative to observations, non-uniqueness is the rule, and finding the best similarity function is computationally hard. The results are interpreted as providing conditions under which rational agents who have access to the same observations are likely to converge on the same predictions, and conditions under which they may entertain different probabilistic beliefs.

Ghirardato, P., Pennesi D.,

*Foundations of Subjective Probabilities, Ambiguity*

We provide a general – fully subjective – foundation to the notion of mixture of acts introduced by Anscombe and Aumann (1963), and successfully employed in many axiomatic models of decision making under uncertainty.

---

F. Massari,

*Learning and Ambiguity*

We present evidence of non-trivial interactions between learning and ambiguity. We find that ambiguous averse decision functional do not produce robust out-of-sample predictions when matched with learning models. There are situations in which the illusion of learning induces an ambiguous averse DM to optimally choose a sequence of ambiguous acts over a sequence of risky acts which would deliver a higher average utility.

---

Canale A., Corradin R., Nipoti D.,

*Augmented Conditional Sampler for Nonparametric Mixture Models*

Nonparametric mixture models based on the Pitman-Yor (PY) process are a flexible tool for density estimation and clustering. Two main classes of algorithms, namely marginal and conditionals, have been considered in literature. We propose a new algorithm, named Augmented Conditional Sampler (ACS), which, although technically conditional, is closely reminiscent of the Pólya urn marginal scheme and features the same degree of interpretability. Unlike its most popular conditional competitors, the ACS does not rely on the stick-breaking representation of the underlying PY process and turns out to be more robust to the choice of the parameters characterising the distribution of the underlying PY process. The performance of the ACS is investigated and compared with popular competitors, by means of an extensive simulation study. Finally, the proposed sampler is used as the building block of a new algorithm for carrying out posterior inference based on a class of dependent nonparametric priors.

---

M. Pahlke,

*Dynamic Consistency in Incomplete Information Games with Multiple Priors*

This paper generalizes the concept of Sequential Equilibrium to allow for ambiguous incomplete information about types or states. We characterize conditions that ensure existence of Sequential Equilibria under ambiguous incomplete information. Under these conditions players form subjective prior belief sets that satisfy a rectangularity condition which leads to dynamic consistent behaviour. Furthermore, we give an example which shows that ambiguity can introduce new Sequential Equilibria.

---

Y. Permana,

*Decision Models under Risk and Uncertainty*

A key assumption of mainstream economics is that people have well-defined preferences, and can state them precisely. This view is being challenged, often through experiments, leading to the increasingly-accepted notion that people may be unsure about their preferences, or at least are unable to state them with precision. A number of experiments have been conducted investigating the nature of, and the reasons for, such imprecision. One interesting design is that of Cubitt et al (2015) that allows subjects to state that they are unsure between two alternatives. I follow their design, but use a different payment scheme, which I feel helps to understand better the reasons why people might state that they are unsure. The key innovation is the way that a statement of unsureness is played out in the experiment: if a subject states that he or she is unsure whether he or she prefers  $A$  or  $B$ , then the experimenter will toss a fair coin to decide. I have four distinct stories and I compare their goodness-of-fit in explaining the data. The first story is that the decision-maker (DM) has convex preferences within Marschak-Machina Triangle (MMT) and actually prefers a mixture of  $A$  and  $B$ ; the second is that the DM simply makes a mistake; the third is that the DM cannot distinguish between  $A$  and  $B$  unless their difference exceeds some threshold; the fourth is that the DM actually prefers to delegate the choice (to the coin), shifting the “responsibility” to the coin. My results show that the first and the third have the most empirical support. Further research is necessary to disentangle these two.

---

Coletti G., Petturiti D., Vantaggi B.,

*Decisions on Generalized Anscombe-Aumann Acts under Possibly “Unexpected” Scenarios*

We consider decisions on generalized Anscombe-Aumann acts, mapping states of the world to belief functions over a set of consequences. Preference relations on these acts are given by a decision maker under different scenarios (conditioning events). Then, we provide a system of axioms which are necessary and sufficient for the representability of these “conditional preferences” through a conditional functional  $CEU_{P,u}$ , parametrized by a unique full conditional probability  $P$  on the algebra of events and a cardinal utility function  $u$  on consequences. The model is able to manage also “unexpected” (i.e., “null”) conditioning events. We finally provide an elicitation procedure that reduces to a Quadratically Constrained Linear Problem (QCLP).

---

Bartl D., Hellman Z., M. Pinter,

*Characterization of Common Prior.*

Samet (1998) characterized the common prior in finite knowledge-belief spaces. Feinberg (2000) by a counterexample showed that Samet’s characterization did not work in general, even in countable type spaces. Lehrer and Samet (2014) considered the countable type spaces and provided three characterizations of common prior. One of Lehrer and Samet’s characterization says that weak trade consistency is equivalent with weak believe consistency. In this paper we slightly modify the notion of weak belief consistency and show in the new setting Lehrer and Samet’s above mentioned characterization works in any type spaces. Moreover, we can reformulate our result as it shows that if the common prior is a probability charge and not a measure, then Samet’s characterization works in any type spaces

---

Al-Najjar N., Pomatto L.,

*Expected Utility; Decision Theory; Finitely Additive Probabilities*

In the evaluation of public policies, a crucial distinction is between plans that involve purely idiosyncratic risk and policies that generate aggregate, correlated risk. While natural, this distinction is not captured by standard utilitarian aggregators. In this paper we revisit Harsanyi’s (1955) celebrated theory of preferences aggregation and develop a parsimonious generalization of utilitarianism. The theory we propose can capture sensitivity to aggregated risk, it is apt for studying large populations and is characterized by two simple axioms of preferences aggregation. The analysis is based on a new abstract law of large numbers for finitely additive probabilities.

Flesch J., A. Predtetchinski,

*Game Theory, Finitely Additive Probabilities*

This work is motivated by the ubiquitous use of “trembles” or “slight mistakes” in game and decision theory. Thus, for instance, solution concepts such as trembling hand perfect, proper, quasi-perfect, and sequential equilibrium all invoke trembles. Trembles are usually introduced by imposing that each action be taken with a positive probability; often times a lower bound on the probabilities is used. We consider a setup where each period of time an action is to be chosen from a finite set  $A$ . A behavioural strategy then gives rise to a probability measure on infinite histories. If we insist that every action in every period be taken with a probability of least  $\varepsilon > 0$ , then the resulting probability measure is a doubling measure. The paper builds on the large body of literature on doubling measures and thin sets. This draft focusses on a particular set of questions related to the behaviour of continuous functions  $f$  mapping  $A^N$  surjectively onto itself. We give a sufficient condition to guarantee that  $f$  preserves doubling measures. We obtain a characterization of quasi symmetric homeomorphisms. We provide an application in dynamic game theory. We also discuss the possibility of introducing finitely additive doubling measures, and suggest a number of open problems that whence arise.

---

Pressacco F., Ziani L.,

*Decision Models under Risk and Uncertainty in Economics*

The economic vision had a pervasive influence on de Finetti’s research and in particular in his approach to the theory of subjective probability. In this paper, we intend to deepen this, up to now relatively little-known, feature of Finetti’s thought. It turns out that the idea of defining probability as a fair price reveals to be an ideal bridge between the classic setting and the recent evolution of modern quantitative finance. This allowed de Finetti to use original reading keys and innovative ideas on various economic and financial issues. There is also an interesting isomorphism between the behaviour of the de Finetti’s evaluator and the Walrasian auctioneer, able to clarify the strategic choices of some economic operators such as market makers and bookmakers.

The standard (Kolmogorovian) notion of conditional probability is compared with the (de Finettian) notion based on the coherence principle. Each notion has both merits and drawbacks, and the talk aims to highlight them with special attention to the connection points. Various results, both old and new, are discussed and some open problems are mentioned. The talk is split into four parts. (i) Classical (Kolmogorovian) conditional probabilities: after recalling the basic definitions, a few critical examples are discussed. Also, a 0 – 1 law for regular conditional distributions is stated and applied to some meaningful sub- $\sigma$ -fields (like the tail and the symmetric). (ii) Disintegrability: Disintegrability makes sense in both the classical and the coherent frameworks. It usually holds in the first, at least under some conditions, but not in the second. Here, after discussing the definition, we give conditions for the existence of  $\sigma$ -additive disintegrations. (iii) Coherent (de Finettian) conditional probabilities: the basic definition is discussed and its merits and drawbacks (relative to the classical definition) are listed. (iv) Some statistical implications: two results are stated. Both are in line with the subjective view of probability and both apply in a finitely additive setting. The first states that, for any prior  $\pi$  and any statistic  $T$ , there is a posterior for  $\pi$  which makes  $T$  sufficient. By “a posterior for  $\pi$ ”, we mean a posterior which is connected with  $\pi$  and the statistical model by a certain disintegrability condition. The second states that, for any estimator  $d$ , if the prior  $\pi$  vanishes on compacta there is a posterior for  $\pi$  which makes  $d$  optimal (i.e., the posterior Bayes rule) under quadratic loss.

---

Costantini C., De Blasi P., Ethier S.N., Ruggiero M., Spanò D.,

*Wright-Fisher construction of the Two-parameter Poisson-Dirichlet Diffusion*

The two-parameter Poisson-Dirichlet diffusion, introduced in 2009 by Petrov, extends the infinitely-many-neutral-alleles diffusion model, related to Kingman’s one-parameter Poisson-Dirichlet distribution and to certain Fleming-Viot processes. The additional parameter has been shown to regulate the clustering structure of the population, but is yet to be fully understood in the way it governs the reproductive process. Here, we shed some light on these dynamics by formulating a  $K$ -allele Wright-Fisher model for a population of size  $N$ , involving a uniform mutation pattern and a specific state-dependent migration mechanism. Suitably scaled, this process converges in distribution to a  $K$ -dimensional diffusion process as  $N$  goes to infinity. Moreover, the descending order statistics of the  $K$ -dimensional diffusion converge in distribution to the two-parameter Poisson-Dirichlet diffusion as  $K$  goes to infinity.

Gilio A., Sanfilippo G.,

*Generalized Logical Operations among Conditional Events*

We generalize, by a progressive procedure, the notions of conjunction and disjunction of two conditional events to the case of  $n$  conditional events. In our coherence-based approach, conjunctions and disjunctions are suitable conditional random quantities. We define the notion of negation, by verifying De Morgan's Laws. We also show that conjunction and disjunction satisfy the associative and commutative properties, and a monotonicity property. Then, we give some results on coherence of prevision assessments for some families of compounded conditionals; in particular we examine the Fréchet-Hoeffding bounds. Moreover, we study the reverse probabilistic inference from the conjunction  $\mathcal{C}_{n+1}$  of  $n + 1$  conditional events to the family  $\{\mathcal{C}_n, E_{n+1}|H_{n+1}\}$ . We consider the relation with the notion of quasi-conjunction and we examine in detail the coherence of the prevision assessments related with the conjunction of three conditional events. Based on conjunction, we also give a characterization of  $p$ -consistency and of  $p$ -entailment, with applications to several inference rules in probabilistic nonmonotonic reasoning. Finally, we examine some non  $p$ -valid inference rules; then, we illustrate by an example two methods which allow to suitably modify non  $p$ -valid inference rules in order to get inferences which are  $p$ -valid.

---

Spizzichino F.

*Relations between Stochastic Ageing and Dependence from the Viewpoint of Subjective Probability and Theory of Risk*

Let  $\mathbf{X} \equiv (X_1, \dots, X_n)$  be a vector of non-negative random variables to be interpreted as *lifetimes* of different individuals. In such a frame, attention is often concentrated on the two different phenomena of *stochastic dependence* and of *stochastic ageing*. A very rich literature in applied probability has been devoted to this field and, from a technical viewpoint, we remind that several different notions of dependence and of ageing have been considered. Dependence and ageing are strictly related one another and, at a time, they are heavily affected by the actual state of information about  $(X_1, \dots, X_n)$ .

In the first part of the talk, I aim to reviewing and clarifying some related aspects in the perspective of Subjective Probability and Bayesian Statistics. From a technical point of view, the discussion will be based on the concepts of copula, of *semi-copula* (a rather natural extension of the notion of copula) and, in particular, on Archimedean copulas and "Archimedean" semi-copulas. An important issue is the possibility of formally extending from copulas to semi-copulas the definitions of some properties of stochastic dependence. This approach permits properties of ageing to be formally described as properties of dependence for semi-copulas and provides us with a method for analysing some relations between dependence and ageing.

The second part will be dealing with connections with problems in Risk Theory. Actually, in some past papers, some risk-related properties of a single-attribute utility function have been respectively related to the dependence properties of an Archimedean copula and to the one-dimensional ageing property of a survival function. I will discuss some implications that may arise when applying to these topics the method reviewed in the first part.