Courses

Antonio Cuevas – UAM – A short, partial overview of Functional Data Analysis.

The aim of this mini-course is to provide a partial survey of the theory and practice of statistical methods with functional data [often referred to as Functional Data Analysis (FDA)]. Among the different topics of interest in this field I will try to consider the following ones:

- Introduction and motivation. A real-data example.
- How to handle in practice the functional data.
- Some basic hints on the underlying probability theory: functional probability models, functional expectation, the Law of Large Numbers and the Central Limit Theorem in function spaces.
- The notions of median and mode in FDA.
- Depth measures for functional data.
- The Random Projections (RP) method. RP-based depth measures.
- The linear regression model with functional regressor and a scalar response.
- An introduction to classification with functional data.

No previous background on FDA is assumed. The style will be expository: the proofs as well as many technicalities will be omitted. However, the focus will be on mathematical aspects (rather than on computational problems or specific applications). The goal is discussing topics with a rich mathematical structure (quite different to that of their finite-dimensional counterparts) where many interesting problems remain open.

Milton Jara – IMPA – An introduction to the KPZ equation.

The aim of this course is to give an introduction to what is called the KPZ equation. The KPZ equation was introduced in [3] as a continuous model for stochastic growth of interfaces and corresponds to the stochastic PDE

$$\partial_t h = \nu \partial_{xx} h + a(\partial_x h)^2 + \sigma \mathcal{W},$$

where $h : R \times [0, \infty) \to R$ is a space-time stochastic process and $\mathcal{W}$ is a standard white noise in $R \times [0, \infty)$. Up to date, existence and uniqueness of solutions, as well as well-poseness of the Cauchy problem are still missing. In the original work of [3], a notion of solution, known as the Cole-Hopf solution was proposed. The idea is use the Cole-Hopf transformation at a formal level to linearize the KPZ equation. This notion of solution leads to existence, uniqueness and well-posedness, although the validity of the formal transformation is still to be justified in a rigorous way. It was not until ten years ago that it was proven that this notion of solution arises in a natural way from a particular lattice dynamics: the so-called asymmetric simple exclusion process. Recently,

deep connections between the KPZ equation, random polymers and also random matrices were discovered, which gathered a revival of the research around the KPZ equation. However, existence, uniqueness and well-posedness still remain as one of the main open problems of the field.

This course will consists of three lectures; a temptative program is the following:

- Lecture 1: The KPZ equation and the stochastic Burgers equation; the Cole-Hopf transformation and the stochastic heat equation. Existence and uniqueness of solutions of the stochastic heat equation.
- Lecture 2: The asymmetric simple exclusion process; Gartner transformation and convergence to the stochastic heat equation. Bertini-Giacomin theorem [1].
- Lecture 3: De

  nition of energy solutions of the KPZ equation. Universality of stationary energy solutions of the KPZ equation.

References

1 L. Bertini and G. Giacomin, Stochastic Burgers and KPZ equations from particle systems, Comm. Math. Phys. 183(3), 571–607 (1997).

2 Patricia Gonzalcalves and Milton Jara, Universality of KPZ equation, (2010).

http://arxiv.org/abs/1003.4478

2 M. Kardar, G. Parisi and Y.-C. Zhang, Dynamic Scaling of Growing Interfaces, Phys. Rev. Lett. 56(9), 889–892 (Mar 1986).

Invited Talks

Miguel Abadi – IME USP.

The non-central limit theorem for the auto-correlation function

We consider a stochastic process over a finite or countable alphabet. We introduce the auto-correlation function that makes occurrences of a given observable (in space or time), appear in clusters or isolated. When the observables are words, we compute the exact distribution and the limiting distribution of this function when the sequence is generated by iid random variables. We also give a point-wise upper bound for the velocity of this convergence. Further, we show how this function gives a measure of the complexity of the process and can be used to determine the complexity of graphs. We illustrate with some examples.

Renato Assuncao – Universidade Federal de Minas Gerais

A general association index with applications to spatio-temporal data

We propose an association index for random variables that measures the strength of their dependency and that is able to capture complex non-linear associations. We apply the index to several spatial and spatio-temporal problems, including association between space and time in point processes and preferential sampling in geostatistics. We show how to estimate our index avoiding the curse of dimensionality.

Hubert Lacoin – Universit Paris Dauphine

Approximate Lifshitz law for the mixing time of the zero-temperature stochastic Ising model with + boundary conditions in any dimension

It has been noticed that below the critical temperature the mixing properties of the Stochastic Ising model are strongly dependent on boundary condition. Indeed if one considers Heat-Bath Dynamics for Ising in the cube of side L, the mixing time is exponential in L whereas it is believed that for all + boundary condition, it behaves like L^2 (conjecture called "Lifshitz law"). What we present here is a new step toward the verification of the conjecture, showing that in all dimension, the mixing time with + boundary condition is $\mathrm{O}(L^2 \log L^c)$ for soma appropriate c in any dimension. This generalizes a recent result by Caputo, Martinelli, Simenhaus and F.L.Toninelli (who proved it in two and three dimension). We will present the key results obtained by Caputo et al. for the three dimensional model and explain how they can be used to obtained the result for dimension larger than three.

Rob Morris – IMPA

Noise Sensitivity in Continuum Percolation

In critical bond percolation on $\mathbb{Z}^2$ (i.e., with $p = 2$, consider the event that there is a horizontal crossing of the box $[n]^2$. Suppose that this event occurs for a particular (random) configuration $x \in \{0,1\}^E$, and let x^ε be obtained by re-randomizing each edge with probability $\varepsilon > 0$. What is the probability that there is a crossing in x^ε? This question was first asked by Benjamini, Kalai and Schramm [1], who proved that the probability converges to $\frac{1}{2}$ as $n \to \infty$, and termed this phenomenon *noise sensitivity*. Their proof used techniques from discrete Fourier analysis, and built on ideas introduced in the famous paper of Kahn, Kalai and Linial [3]. In recent years much more

precise results have been obtained about the Fourier spectrum of percolation, and about dynamical percolation on the triangular lattice (see [2, 5]). In this talk we consider the corresponding question for Continuum Percolation, and in particular for the Poisson Boolean model (also known as the Gilbert disc model). Let η be a Poisson process of density λ in the plane, and connect two points of η by an edge if they are at distance at most 1. We prove that, at criticality, the event that there is a crossing of the box rns2 is noise sensitive. The proof is based on two extremely general tools: a version of the BKS Theorem for product measure, and a new extremal result on hypergraphs. The former result was first proved in [4]; we shall describe how it may be easily deduced from the uniform version. This is joint work with Daniel Ahlberg, Erik Broman and Simon Griffiths.

1. Itai Benjamini, Gil Kalai and Oded Schramm,Noise sensitivity of Boolean functions and applications to percolation, *Inst. Hautes Etudes Sci. Publ. Math.* **90** ,pp. 5–43, 1999.
2. Christophe Garban, Gabor Pete and Oded Schramm,The Fourier spectrum of critical percolation, *Acta Math.* **205**, pp. 19–104, 2010.
3. Jeff Kahn, Gil Kalai and Nathan Linial,The influence of variables on boolean functions,29th Annual Symposium on Foundations of Computer Science, pp. 68–80, 1988.
4. N. Keller and G. Kindler, Quantitative Relation Between Noise Sensitivity and Influences, preprint, 2010.
5. Oded Schramm and Jeffrey Steif, Quantitative noise sensitivity and exceptional times for percolation,*Ann. Math.* **171**, pp. 619–672,2010.

Matias Salibian-Barrera – University of British Columbia.

On the stability of bootstrap estimators for support vector machines

Let T be an estimator based on a continuous operator from the space of probability measures over a compact metric space into a complete separable metric space. In this talk we show that bootstrap approximations to the distribution of T are stable ("qualitatively robust"). An application of this result that is of particular interest is the bootstrap approximation to the distribution of support vector machines. This is a joint work with Andreas Christmann and Stefan van Aelst.

Víctor Yohai – Conicet and UBA.

Weak Continuity and Differentiability of Regression M Estimating Functionals

This talk deals with the Fisher consistency, weak continuity and differentiability of estimating functionals corresponding to a class of both linear and nonlinear regression high breakdown M estimates, which includes S and MM estimates. A restricted type of differentiability, called weak differentiability, is defined, which suffices to prove the asymptotic normality of estimates based on the functionals. This approach allows to prove the consistency, asymptotic normality and qualitative robustness of M estimates under more general conditions than those required in standard approaches. In particular we prove that regression MM-estimates are asymptotically normal when the observations are alpha-mixing.

This is a joint work with María V. Fasano (Universidad Nacional de La Plata), Ricardo A.Maronna (Universidad Nacional de La Plata), Mariela Sued (Universidad de Buenos Aires and CONICET) and Víctor J. Yohai (Universidad de Buenos Aires and CONICET)

Generalized Linear Model with Missing Responses: a Robust Approach.

Speaker: Ana M. Bianco - Universidad de Buenos Aires and CONICET.

When dealing with situations in which the responses are discrete or show some type of asymmetry, the linear model is not appropriate to establish the relation between the responses and the covariates. Generalized linear models (GLM) serve this purpose, since they allow to model the mean of the responses through a link function, linearly on the covariates.

When atypical observations are present in the sample, robust estimators are useful to provide fair estimations and also to build outlier detection rules. We focus on robust inference procedures involving the regression parameter of a GLM when missing data possibly occur in the responses.

First, we propose robust general M-estimators. We study the asymptotic behaviour of the proposed estimators: they turn out to be consistent under mild assumptions and asymptotically normal, as well. Besides, outlier detection rules are defined using the influence function. Secondly, robust Wald–type tests are derived. The Poisson and Gamma regression models are considered as special cases.

By means of a simulation study we compare the behaviour of the classical and robust procedures, under different contamination and missing probability schemes. We also illustrate the usefulness of the proposals through some real data sets. This is a joint work with Graciela Boente y Isabel Rodrigues.

Strong Law of Large Numbers in the Supremum Norm for Independent RW's with Explosions

Speaker: Tertuliano Franco - Universidad de Buenos Aires

We present a strong law of large numbers in the supremum norm for the density of particle system consisting of independent random walks which disappear at the boundary of the box and particles are created/destroyed in each site according to quantity there. It is also presented some estimates on the explosion time and how it is related to the explosion time of the associated partial differential equation.

The overlapping function of words generated by correlated random variables

Speaker: Erika Alejandra Rada Mora - IME USP Brazil

We consider the set of finite sequences of length n over a finite or countable alphabet χ. We consider the function defined over χ^n which gives the side maximum overlap of a given sequence with a (shifted) copy of itself.

Abadi and Lambert computed the exact distribution and the limiting distribution of the overlapping function when the sequence is chosen according to a product measure with marginals identically

distributed. Further, they gave a point-wise upper bound for the velocity of the convergence.

Our work consists in a generalization of the work done by Abadi and Lambert, without the independence hypothesis to generate the word.

Interpretable Clustering using Unsupervised Binary Trees

Speaker: Marcela Svarc - Universidad de San Andrés

We herein introduce a new method of interpretable clustering that uses unsupervised binary trees. It is a three-stage procedure, the first stage of which entails a series of recursive binary splits to reduce the heterogeneity of the data within the new subsamples. During the second stage (pruning), consideration is given to whether adjacent nodes can be aggregated. Finally, during the third stage (joining), similar clusters are joined together, even if they do not share the same parent originally. Consistency results are obtained, and the procedure is used on simulated and real data sets. (joint work with R. Fraiman and B. Ghattas)

Multivariate Probabilistic Context Trees using Probabilistic Suffix automata.

Speaker: Denise Duarte - UFMG.

In the literature there exists some methodology concerning multivariate Markov models for discrete sequences where the number of parameters in the model is $O(ns2|A|^2)$ (Ching et al, 2008), where s is the number od sequences. But in such models the order r of the chain is fixed.

In a univariate Markov model the number of parameters grows exponentially with the order of the chain, $O(|A|^r)$, where $|A|$ is the size of the alphabet. This could be intractable even for not very large orders. If we want to describe a model for more than one sequence of Markov chains, say r sequences things tends to become worst. In this case the number of parameters in the model to be estimated is $O(|A|^{sr})$.

In this work we want to propose estimators for multivariate Markov models following the approach proposed by (Ching et al, 2008), but instead working with full Markov Chain models we deal with Variable Length Memory Chains (VLMC), also known by probabilistic Contex Trees (PCT), introduced by Rissanen in 1983, extending the results to a more general class of processes.

Estimation of Regression Models with Locally Stationary Long-Memory Errors.

Speaker: Guillermo Ferreira - Universidad de Concepcion- Chile

This paper addresses the statistical analysis of regression models with locally stationary disturbances. This methodology allows for the fitting of non-stationary time series data displaying both trends and time-varying long-range dependent errors.

In order to deal with the non-stationary behavior of the regression errors, a locally stationary approach is proposed. This statistical framework allows for the modeling of a time-varying autocovariance structure. In this context, the parameters of the non-stationary model are allowed to vary smoothly over time so that it can be locally approximated by stationary processes.

The study conducted in this work focuses on the analysis of some statistical properties of the least squares estimates (LSE) of the regression models described above. These estimators are widely used in practice because they can be readily calculated. Observe that other techniques such as, for example the best linear unbiased estimators (BLUE), make the unrealistic assumption that the dependence structure of the errors in known a priori. This critical assumption is even harder to justify in practice since the dependence structure of the errors is not necessarily stationary.

Regular g measures are not always gibbsian.

Speaker: Sandro Gallo - UFRJ.

We present a simple example of g-function for which the unique stationary compatible g-measure is not gibbsian. This shows that the results of unidimensional Gibbs states cannot (in general) be

imported directly to the study of g-measures. This is a joint work with R. Fernández and G. Maillard.

Threshold selection for extremes under a semiparametric model

Speaker: Juan D. Gonzalez - UBA

We propose semiparametric likelihood estimator for the threshold under a semiparametric model that assume that the excedent distribution belong the Generalized Paretto family, above the threshold. A simulation study is performed to show empirically the properties of the proposal. We also compare our results with those obtained by some of the most common existing methods.

Estimación en modelos aditivos con respuestas faltantes

Speaker: Alejandra Martínez - FCEyN, UBA - CONICET

El modelo de regresión noparamétrico aditivo supone que se tienen observaciones independientes $(\mathbf{x}_i^T, y_i)$, $1 \leq i \leq n$, $\mathbf{x}_i \in \mathbb{R}^d$ tales que $E(y_i|\mathbf{x}_i) = m(\mathbf{x}_i)$ con

$$m(\mathbf{x}) = \mu + \sum_{\alpha=1}^{d} g_\alpha(x_\alpha)$$

Las funciones $g_\alpha : \mathbb{R} \to \mathbb{R}$ y el parámetro de posición marginal μ son las cantidades a estimar. Estimadores para este modelo han sido ampliamente estudiados en la literatura. Presentaremos estimadores para las componentes de un modelo aditivo cuando las respuestas pueden ser faltantes, es decir, cuando observamos $(\mathbf{x}_i^T, y_i, \delta_i)$, $1 \leq i \leq n$ donde $\delta_i = 1$ si y_i es observada y $\delta_i = 0$ si y_i es faltante. Para ello, suponemos un mecanismo de pérdida de observaciones ignorable (MAR). Bajo estas hipótesis proponemos dos familias de estimadores basados en la muestra completa, o sea, eliminando todos los pares incompletos (aquellos con $\delta_i = 0$) y que llamaremos simplificados. Una familia se basa en el estimador de núcleos de Nadaraya–Watson y la otra en estimadores internamente corregidos.

Describiremos el comportamiento asintótico de ambas propuestas y presentaremos los resultados obtenidos al comparar ambos estimadores mediante un estudio de simulación.

Counting the crossings of random segments

Speaker: J.C.S. de Miranda University of Sao Paulo

In this note we will be interested in studying the number of intersections of segments whose extremes lie on the boundary of convex bounded subsets of $\mathbb{R}^2$. Let us denote by $\mathcal{C}$ the convex set and by $\partial\mathcal{C}$ its boundary. We will suppose that the endpoints of the segments follow a uniform law with respect to arc length on this boundary which will be assumed to be differentiable by parts when arc length is used for its parametrization. Given an ordered collection of uniformly and independently distributed points $P_1, ..., P_{2N}$ on $\partial\mathcal{C}$ we will form the N line segments $\overline{P_1P_2}, ..., \overline{P_{2i-1}P_{2i}}, ..., \overline{P_{2N-1}P_{2N}}$. Only the crossings that lie in the interior of the convex set will be counted. Generalizations to unbounded convex sets and non uniform distributions are also

presented.

A review about negative integer moments in Moment generating function and Probability generating function.

Speaker: Bruno Monte de Castro - Universidade de Sao Paulo.

Usually we see in many books an alternative way of calculating the positive integer moments of a random variable, using the moment generating function, $M_X(t)$ and probability generating function $\phi_X(t)$, the latter only in the discrete case.

However we did not find a way to calculate the negative integer moments, which can be calculated as

$$\exp(X^{-1}) = \int_{-\infty}^{0} M_X(t)dt$$

or

$$\exp[(X+1)^{-1}] = \int_{0}^{1} \phi_X(t)dt$$

in the general case, we can calculate any order such as

$$\exp(X^{-k}) = \int_{-\infty}^{0}\int_{-\infty}^{t_1}\cdots\int_{-\infty}^{t_{k-1}} M_X(t_k)dt_k\cdots dt_2dt_1$$

$$\exp\left(\frac{1}{(X+1)(X+2)\cdots(X+k)}\right) = \int_{0}^{1}\int_{0}^{t_1}\cdots\int_{0}^{t_{k-1}} \phi_X(t_k)dt_k\cdots dt_2dt_1$$

Besides being easier to calculate the negative integer moments like this, we know that doesn't exist closed form to moment generating function of $Y \sim IG(\alpha, \beta)$ (*Inverse Gamma distribution*), but using this method, all negative integer moments can be obtained through the moment-generating function of $X \sim G(\alpha, \beta)$ (Gamma distribution).

Segmentation of Symbolic Sequences based on Variable Length Markov Chains

Speaker: Bruno Monte de Castro - Universidade de Sao Paulo.

We consider the problem of segmenting symbolic sequences over a finite alphabet into a finite number of independent blocks. This problem arises in many different application problems, as for example when analysing the structure of single nucleotide polymorphisms (SNPs) maps [1] or in the case of linguistic sequences [2]. Assuming the sequences where generated by variable length Markov chain models [3], we propose to use a penalized maximum likelihood estimator to infer the number of blocks and the location of the boundaries for each block. Based on independent sequences of the same length (as in the case of SNPs sequences) we study the behaviour of the estimator on simulated data.

1. Bianchi, A.; Giolo, S.; Leonardi, F. and Pavan, J. Finding variable dependence structure in Single Nucleotide Polymorphism maps. Manuscript, 2011.

2. Galves, A.; Galves, C.; Garcia, J.; Garcia, N. and Leonardi, F. Context tree selection and linguistic rhythm retrieval from written texts. *Annals of Applied Statistics* (to appear), 2011.

3. Rissanen, J. A universal data compression system. *IEEE Trans. Inform. Theory* **29**(5), pp. 656–664, 1983.

Study of conformal maps using measure harmonic

Speaker: Jose Javier Cerda Hernandez - University of Sao Paulo

In this poster we show an application of harmonic measure to study conformal maps using the conformal invariance of brownian motion.

Interactions on Ising Model: A Result About Coupling Between Long Range and Finite Range Glauber Dynamics

Speaker: Estefano Alves de Souza- IME-USP

This is a PHD Thesis, supervised by Antonio Galves (IME-USP). We present the classic ferromagnetic Ising Model μ on $\mathbb{Z}^d$ and its version with truncation level $L > 0$, $\mu^{[L]}$. Our main goal is to estimate the interaction function $\{J(i,j) \in \mathbb{R}, (i,j) \in \mathbb{Z}^d \times \mathbb{Z}^d\}$ inside a finite box $F \subset \mathbb{Z}^d$. To do this, our first step is to define a coupling between two Glauber Dynamics on $\{-1,1\}^{\mathbb{Z}^d}$ with reversible (invariant) measures μ and $\mu^{[L]}$, respectively, and give a bound to the probability of the proportion of time which the projections of the processes at a site i are different in the interval $[0,t]$ is greater than a constant $\delta \in (0,1)$.

The second step is: Given that both processes are equal with high probability on the time interval $[0,t]$, we are interested about to construct estimators for the quantities $\{J(i,j) \in \mathbb{R}, (i,j) \in F^2\}$ using some known method (e.g., Maximum Likelihood Estimation). This will be done at the following of the work of the Thesis.

Probabilistic Forests

Speaker: Karina Yuriko Yaginuma - IME-USP

Let A be a finite alphabet Let Γ be a finite set of context trees in A and $\pi(\tau)_{\tau\in\Gamma}$ a probabilistic distribution over the set Γ.

Suppose that $A = 0, 1, *$ and let $\Gamma = \tau_1, \tau_2$ be a set of context trees such that for all tree in Γ has at least one context that ends with the symbol *. Suppose also that the symbol * only appears in the end of the contexts. For each tree we associate $p_i = p_i(a|w) : a \in A, w \in \tau_i$ for $i = 1, 2$, the probabilities transition associate with each tree. Suppose also that $\pi(\tau_1) = 1 - \epsilon$ and $\pi(\tau_2) = \epsilon$.

Let $I_1, I_2, ...$ be iid random variables with $P(I_n = 1) = 1 - \epsilon$ and $P(I_n = 2) = \epsilon$. Consider the process $\eta_1^{I_1}, \eta_2^{I_2}$,...where $\eta_i^{I_i}$ is the outcomes of the process generate by (τ_{I_i}, p_{I_i}) until find the first symbol *.

Given a sample sequence $\eta_1^{I_1}, ..., \eta_n^{I_n}$ the probability of each tree in the set can be estimated by

the procedure in Eskin et al. (2002). If ϵ is small, we want to know if the estimator $\hat{\pi}(\tau)$ converge in distribution to $\delta_{(\tau_1,p_1)}$.

The Monge-Kantorovich Problem for the Quadratic Cost

Speaker: Guilherme Ost de Aguiar - Universidade de Sao Paulo

We analyze the Monge-Kantorovich optimal transportation problem in the case where the cost function is given by the square of the Euclidean norm. Such cost has a structure which allow us to get more interesting results than the general case. Our main purpose is to determine if there are solutions to such problem and characterize them.

We also give an informal treatment to the optimal transportation problem in the general case.

Keywords: Monge-Kantorovich. Transference plans. The Kantorovich duality. Subdifferential.

The Kantorovich Duality Theorem for the Optimal Transport

Speaker: Aline Duarte de Oliveira - Universidade de Sao Paulo

We analyze the optimal transport theory proving the Kantorovich duality theorem for a wide class of cost functions. Such result plays an extremely important role in the optimal transport theory. We also prove the Kantorovich-Rubinstein duality theorem, which deals with the particular case of cost function given by the distance.

Title: Maximal Spacing for non-homogeneous d-dimensional Poisson Process

Speaker: Rafael Grisi - UFABC

Let K be a compact subset of $\mathbb{R}^d$ and N_f a Poisson process on K with intensity function f. The maximal spacing of N_f on K is defined as the radius of the largest open ball in K, avoiding the points of N_f.

Considering f constant, Jason (1987) studied the behaviour of the maximal spacing of $N_t f$, as t goes to infinity.

In this ongoing work, we try to generalize Jasons results, considering f non-constant. We show that as t goes to infinity, the behavior of the maximal spacing is defined by the minimum value of f on K.

Asymthotic properties of the exit time probability

Speaker: Mariana Pereira de Melo - Universidade de Sao Paulo (IME/USP)

For an α-mixing stochastic process it has been proved in [1] the following results: For any fixed n-string A: (1) The hitting time of A has approximately exponential law. (2) The return time of A has approximately a convex combination between a Dirac measure at the origin and

an exponential law. In both cases the parameter of the exponential law is $\lambda(A)P(A)$, where $P(A)$ is the measure of the n-string and $\lambda(A)$ is the probability of the exit time of A, Speaker: ly, $\lambda(A) = P(X_\tau^{\tau+n-1} \neq x_0^{n-1} | X_0^{n-1} = x_0^{n-1})$, where $\tau = \tau(A)$ is the first return time to A, i.e. $\tau = \inf\{k \geq 1 : \{X_k^{k+n-1} = x_0^{n-1}\} \cap \{X_0^{n-1} = x_0^{n-1}\} \neq \emptyset\}$. Moreover, the weight of the aforementioned convex combination is $\lambda(A)$. We consider $\rho(A) = 1 - \lambda(A)$. We are interested in the asymptotic behaviors of $\rho(A)$. We present some preliminary results of $\rho(A)$, and compare them with those of $P(A)$, as the Shannon-Mc Millan Breiman Theorem and its normal fluctuations.

[1] Abadi, M.: Hitting, returning and the short correlation function, Bulletin of the Brazilian Mathematical Society 37(4), 593-609.

Ising model on Lorentzian triangulations. Annealed case.

Speaker: Anatoli Iambartsev - IME-USP

We study the Ising model on Lorentzian triangulations annealed case. The transfer matrix technique provides the uniqueness area on the set of parameters β and μ, where β is traditionally inverse temperature and μ is a chemical potential. This a joint work with S.Zohen and Y.Suhov.

K-Nearest neighbor density estimation on Riemannian Manifolds.

Speaker: Andrés Muñoz - UBA

We consider a k-nearest neighbor kernel type estimator when the random variables belong in a Riemannian manifold. We study asymptotic properties such as the consistency and the asymptotic distribution. A simulation study is also consider to evaluate the performance of the proposal. Finally, to illustrate the potential applications of the proposed estimator, we analyzed two real example where two different manifolds are considered. This a joint work with Guillermo Henry and Daniela Rodriguez.

Deteccion de bordes via wavelets

Speaker: Javier Gimenez - Famaf, UNC.

La idea del trabajo es formalizar matematicamente y explicar la simulacion de un detector de bordes basado en la Transformada Wavelet Haar Discreta modificada de un salto, para luego comparar el resultado obtenido con el tradicional Algoritmo de Canny. En el camino a tal objetivo, exponemos una intepretacion grafica e intuitiva de los coeficientes Haar. Es sabido que dada una imagen de la realidad, su digitalizacion no es mas que una discretizacion de la misma, en la que se asigna un determinado valor en cada sitio de una maya discreta colocada sobre el dominio de la imagen. Este valor tiene diferentes interpretaciones segun la naturaleza del sensor. Nosotros supondremos que este valor es la media de la intensidad sobre la region que representa el pixel, ponderando uniformemente dicha region. Esta suposicion permite ver a la digitalizacion de una señal, como la proyeccion de la misma sobre un determinado nivel de resolucion del sistema de multiresolucion asociado a la Wavelet Haar. Sin embargo, si esta suposicion se elimina, podemos formalizar matematicamente el detector de bordes, pero fuera del contexto de proyecciones

ortogonales sobre subespacios vectoriales de $L^2(\mathbb{R}^2)$.

El trabajo consta de un recopilado de resultados consernientes a la base Haar bidimensional de $L_2(\mathbb{R}^2)$ con la interpretación gráfica de los coeficientes resultantes al aproximar una determinada imagen. Estos coeficientes se clasifican en coeficientes de ajuste y coeficientes de detalle, entre estos últimos podemos encontrar los coeficientes de detalles verticales, horizontales y diagonales. Los coeficientes de ajuste nos dan una idea sobre la intensidad promedio que contiene la imagen en una determinada zona de la misma, mientras que los coeficientes de detalle, nos dan un claro indicio de la existencia de bordes en la zona.

Definimos y simulamos la Transformada Wavelet Haar Discreta, con su respectiva aplicacion inversa; y finalizaremos con una explicacion detallada de un procedimiento que proponemos para extraer de una forma eficiente la mayor cantidad de informacion posible de los coeficientes de detalle, con el fin de hallar los bordes de una determinada imagen.

Rubik's Cube: modeled probabilistically

Speaker: Felipe Silva de lima - Universidade Federal de Pernambuco

This work is the result of a survey on applications of the concepts of probability theory in the Rubik's Cube, in order to transform it into an illustration to make your probability. With the assumption that the probability theory is not restricted to the investigation of a likelihood of an event occur, but also in the creation, development and research models that can be used to study random phenomena or experiments (Morgado & CIA, 2004). In this work it is possible to understand how the Rubik's Cube can be a tool to demonstrate some of the items and concepts of Probability Theory with the help of combinatorics, like how to calculate the whole universe, ie, all possible configurations of cube; Calculate the probability of turning the Rubik's Cube using the method of Freidrik; assemble the tree of decisions, all to the magic cube as a model to facilitate understanding of teaching, making this area of study more attractive. There are several common and instruments used to perform experiments to identify whether the phenomena are random or deterministic, such as currency - has a sample space of two possibilities, heads or tails, the data - the sample space has six possibilities, faces 1, 2, 3, 4, 5 or 6, since the Rubik's Cube stands out among those cited, because it has a huge sample space of 43,252,003,274,489,856,000 possible configurations, making it a veritable galaxy of possibilities and also a great tool for exploration and probabilistic statement. According to Google's computers, someone would take trillions of years 1400 to run all the settings, so if you can do each of the movements every second. Finally the Rubik's Cube is not limited to the title of "the most popular toy in the world or a great exercise for the reasoning," but a tool that can be very useful for the study and teaching of probability, because we can explore whole universe, its tree-making, possible permutations of its vertices, optimization in the solution of the cube using the theory of groups, ie, calculations that reduce the amount of movements that lead in solution. Finally, from the simple creation of your tree to the most complex decisions, which in this case would be the optimization of their solutions, the Rubik's Cube out of the class to be toy and becomes an object of research, modeling and teaching.

Estimación de parámetros en un modelo de difusión generado por un proceso Gamma y el movimiento Browniano.

Speaker: Nicolas Moreno - Universidad de Valparaíso

A partir del trabajo de [1], se usan los procesos en tiempo continuo para modelar el comportamiento del precio de un activo. [2] desarrollaron la fórmula que es la base de todo el mercado de opciones. Una opción se entiende como el pago para poseer el derecho pero no la obligación con fecha de expiración fija, en el cual el contratante, posee la opción de comprar o vender activos en un precio de ejercicio. En el modelo de Black-Scholes, el precio de una opción, S_t, queda determinado por la solución de la siguiente ecuación diferencial estocástica

$$dS_t = \mu S_t dt + \sigma S_t dB_t, \quad t > 0, \mu \in R, \sigma \in R^+,$$

donde B_t representa el movimiento Browniano. El modelo de Black-Scholes no siempre es apropiado, por el hecho que las variaciones en el precio de una opción no es totalmente capturado por la distribución normal, puesto que esta posee 'colas más livianas' y los retornos históricos poseen 'colas más pesadas '. Es por esta razón que en [3] consideraron una generalización del modelo de Black-Scholes, en la cual se reemplaza el movimiento Browniano por un ruido generado por el movimiento Browniano (B_t) y el proceso Gamma (γ_t), $\tau_t = B_t \gamma_t^{\alpha} t^{\beta}$ con $\alpha, \beta \in R$,

$$dS_t = \mu S_t dt + \sigma S_t d\tau_t, \quad t > 0, \mu \in R, \sigma \in R^+,$$

En este trabajo se realiza un estudio de la estimación, en el modelo discretizado, de la tasa esperada de retornos, μ, y la volatilidad implícita, σ, por medio de una variación del método de máxima verosimilitud. Realizado en conjunto con Soledad Torres.

[1] Bachelier, L. Theorie de la speculation. Annales Scientifiques de l'Ecole Normale Superieueure, 3 (17), 21-86, (1900).
[2] Black, F.; Scholes M. The pricing of options and corporate liabilities. Journal of Political Economy 81, 635-654, (1973).
[3] Iglesias P., San Martin J., Torres, S. and Viens, F. Option pricing under a Gamma-modulated diffusion process, Annals of finance 7, 199-219, (2011).

Deviation Inequalities for the Historical Estimator of AVR.

Speaker: Hector Olivero Quinteros - Mathematical Engineering Department, Universidad de Chile.

The average value at risk of an asset X, AVR(X), is a widely used financial risk measure and its estimation constitutes a very important practical problem. In this poster we show some bounds for the probability of having a large estimation error for the historical estimator, and we do so considering different assumptions on the distribution of X. In particular, we show bounds for random variables with finite square-exponential moments and with finite polynomial moments. All of these results are part of a paper in preparation. This is a joint work wit Joaquín Fontbona.

Caracterizacion de incerteza en detectores de bordes: Caso Canny y Wavelet Maxima.

Speaker: Georgina Flesia - Famaf, UNC

Existe una gran numero de campos de aplicacion donde mediciones derivadas de imagenes digitales tienen gran relevancia. Sin embargo, para realizar esas mediciones, es indispensable alcanzar

un control cuantitativo y cualitativo de las incertezas que son introducidas por pasos previos a la generacion de la medicion. En este trabajo se discute este problema en el caso la caracterizacion del mapa de bordes generado por dos detectores de bordes muy usados, Canny y Wavelet Maxima. En un primer paso se genera un modelo de imagen simplificado para testeo, y se deriva analiticamente la incerteza introducida en los mapas producidos por los detector es Canny y Wavelet Maxima. Dicha incerteza es luego calculada y comparada en imagenes simuladas, y estimada en imagenes reales.

Graphical tools for high dimensional and functional data

Speaker: Ricardo Fraiman - Universidad de San Andrés

Modelos de Regresión con ruido derivado del movimiento Browniano Asimétrico.

Speaker: Manuel Pereira - Universidad de Valparaíso

A partir de los trabajos de Bachelier, L. (1900), se utilizan los procesos en tiempo continuo para modelar el comportamiento del precio de un activo. Black, F. y Scholes M. (1973) desarrollaron la fórmula que es la base de todo el mercado de opciones. En este modelo, el precio de un activo, S_t , queda determinado por la solución de la siguiente ecuación diferencial estocástica:

$$dS_t = \mu S_t dt + \sigma S_t dB_t, \quad t > 0.$$

La cual si se considera una discretización en el tiempo e $Y = ln(S)$ obten- emos $Y_i = \mu\Delta t + \Delta B_i$; donde B_i representa el movimiento Browniano standard, μ es el retorno medio y σ la volatilidad. Una generalización del modelo de Black Scholes, consiste en utilizar el movimiento Browniano asimétrico (SBM) en reemplazo del movimiento Browniano standard. Esta generalización, nos da una amplia gama de posibles ruidos, desde el mismo movimiento Browniano hasta el movimiento Browniano con reflexión (es de- cir totalmente asimétrico). En esta tesis consideraremos el modelo discreto $Y_i = \mu\Delta t + \Delta SB_i$; donde SB_i es el ruido generado a partir de un SBM. Estimaremos los parámetros μ por medio de Mínimos Cuadrados y estudi- aremos vía simulación las propiedades asintóticas de este. Se realizará un análisis de datos reales.

Analysis of queueing networks under dynamic and optimal static routing policies.

Speaker: Heloisa Maria de Oliveira - Unicamp-IMECC

We consider N servers and K classes of customers (non-empty sets of servers $S_i \subset \{1, \ldots, N\}$). The customers (or jobs) arrive at the system accordingly to independent Poisson process with rate λ_i, $i = 1, \ldots, K$. For each arriving customer a fixed subset of servers (or queues) is presented, then the customer is routed to a server accordingly to some routing policy.

Assume that service times at each server j are exponentially distributed with rate μ_{ij} which depend on the customer and the server chosen. This queueing network model was analysed according to optimal static routing policy, *create smallest workload* (CSW) and *join the smallest workload* (JSW) routing policies. Our goal is to present some theoretical and practical results for this proposed model. This is a joint work with Marina Vachkovskaia.

d-DISTANCE BETWEEN TWO RENEWAL PROCESS BY A RENEWAL COUPLING

Speaker: Walter Augusto Fonseca- Unicamp

In this work we are worried with the calculation of the discrepance between two renewal processes by a renewal coupling. In our analysis we use the probability measures: Partition-distance, Markovian distance, Renewal distance and d-distance. initially we verify that is ever possible to build a renewal process with states space f0; 1g from a markovian process with states space f1; 2; 3g. The motivation to this proceedmant is the great diculty , and until the impossibility, of to calculate the invariants measures of a generic renewal process and therefore the renewal distance R and so the d-distance, which are our especially interest. As for markovian process this method alread is well-known and of calculation more easy, we will use the proceedmant: We will calculate the invariants measures to markovian case then we will build the renewal process of way to preserve, at least partially, markovian process's measures.

Stabilized integrators for stochastic differential equations driven by small additive noise.

Speaker: Hugo de la Cruz Cansino - IMPA

Abstract: We propose a new class of stable integrators for the effective integration of small additive noise Stochastic Differential Equations (SDEs) (i.e., SDEs with small diffusion term). We study the mean-square order of convergence and the asymptotic stability of the proposed methods. Computer simulations illustrate the theoretical findings and the advantages of the proposed integrators in comparison with some conventional and commonly-used methods. (joint work with Jorge Zubelli - IMPA)

Generation of sub hourly Rainfall Fields from Stochastic Cascades

Speaker: Eddy Herrera Daza - Javeriana University

The construction of a multiplicative cascade process for synthesising sub-hourly rainfall sequences and its applications to modelling of rainfall fields are detailed in this work .Data were used the 15-min rainfall data from FOPAE raingauge station in Bogota . The results entail that the proposed method has the ability to produce the sub-daily rainfall sequences having more similar patterns.

Random Perturbations of Stochastic Processes

Speaker: Lucas Moreira - IMECC.

We consider binary infinite order stochastic chains perturbed by a random noise. This means that at each time step, the value assumed by the chain can be randomly and independently flipped with a small fixed probability. We show that the transition probabilities of the perturbed chain are uniformly close to the corresponding transition probabilities of the original chain. As a consequence, in the case of stochastic chains with unbounded but otherwise finite variable length memory, we show that it is possible to recover the context tree of the original chain, using a suitable version of the algorithm Context, provided that the noise is small enough

Overlapping hypercubes.

Speaker: Rodrigo Lambert - University of Sao Paulo

In this poster, we will try to present the higher-dimension problem of overlapping and studying the overlapping function and its properties, as was done by Abadi & Lambert (2011), for sequences (one-dimensional) generated by IID processes.

Variable length markov chains and dynamical sources

Speaker: Andrés Arias Rodríguez

Variable Length Markov Chains were first introduced by Rissanen (1983) as flexible and par-

simonious models for data compression. Originally called by Rissanen finite memory source or probabilistic tree, this class of models recently became popular in the statistics literature under the name of Variable Length Markov Chains by Buhlman and Wyner (1999). This communication shows the relationship between a VLMC and dynamical system like is presented by Cénac et.al (2010) in the paper Variable Length Markov Chains and dynamical sources. They considered a finite alphabet A and the final letter process $(X_n)_{n\in N}$, which is obtained from a VLMC. On the other hand, a dynamical source produces a A-valued random process $(Y_n)_{n\in N}$. They proved that the letter process $(X_n)_{n\in N}$ and the process $(Y_n)_{n\in N}$ are symmetrically distributed, which means that for any finite word w of length $N+1$, $P(X_0, ..., X_n = w) = P(Y_0, ..., Y_n = \bar{w})$, where $\bar{w}$ denotes the reversed word.

A Bayesian Approach to Aggregate Functional Data: Applications towards Electric Load Monitoring and Chemometrics.

Speaker: Ronaldo Dias - UNICAMP- University of Campinas

In many areas of science one aims to estimate latent sub-population mean curves based only on observations of aggregated population curves. By aggregated curves we mean linear combination of functional data that cannot be observed individually. We assume that several aggregated curves with linear independent coefficients are available. More specifically, we assume each aggregated curve is an independent partial realization of a Gaussian process with mean modeled through a weighted linear combination of the disaggregated curves. We model the mean of the Gaussian processes as a smooth function approximated by a function belonging to a finite dimensional space HK which is spanned by K B-splines basis functions. We explore two different specifications of the covariance function of the Gaussian process: one that assumes a constant variance across the domain of the process, and a more general variance structure which is itself modelled as a smooth function, providing a nonstationary covariance function. Inference procedure is performed following the Bayesian paradigm allowing experts opinion to be considered when estimating the disaggregated curves. Moreover, it naturally provides the uncertainty associated with the parameters estimates and fitted values. Our model is suitable for a wide range of applications. We concentrate on two different real examples: calibration problem for NIR spectroscopy data and an analysis of distribution of energy among different type of consumers.

A study of context tree estimation via BIC

Speaker: Douglas Rodrigues Pinto - Universidade de Sao Paulo

Stochastic chains with memory of variable length were introduced by Rissanen (1983) and are currently an important subject of research due to its applications in various fields such as linguistics and genetics, for example. Given a sample originated from a probabilistic context tree, our main objective is to estimate such tree using the Bayesian Information Criterion (BIC), which consists in calculating the tree $\tau_{BIC}\left(X_1^n, c\right)$ that maximizes the log-likelihood function, penalized by a function $g\left(\tau, n, c\right) = c \cdot df\left(\tau\right) \cdot \log n$, where n is the sample size, $df\left(\tau\right)$ is the function that returns the degrees of freedom of τ and c is any positive constant. Although Csizár and Talata (2006) have proved that the value of the constant c is asymptotically irrelevant, for a fixed sample the result is affected by the choice of c. Our objective is to calculate the constant value c^* which optimizes the estimation of the probabilistic context tree that generated the sample.