

RESUMENES

Miguel Abadi

Aproximación de Poisson y aplicación a secuencias biológicas

Gracias al trabajo acerca de la estadística de instantes de ocurrencia de una secuencia de símbolos en un proceso estocástico con propiedades mezcladoras, ofrecemos dos nuevos métodos para la búsqueda de "palabras raras" en secuencias biológicas. Nuestro objetivo es obtener una cota superior en el error cometido al aproximar la ley exacta del número de apariciones de una palabra y una ley de Poisson, que sea apropiada para secuencias biológico. Un método alternativo ya clásico es el de Chen-Stein. Usamos este error para analizar cual método resulta más sensible para detectar este tipo de palabras. Para comparar los métodos de Chen-Stein, ϕ -mezclador y ψ -mezclador usamos un programa para la implementación de las fórmulas de cada método y presentamos algunos resultados preliminares. Este es un trabajo en conjunto con Nicolas Vergne.

Vladimir Belitsky

*The Influence of Heterogeneity of Individual Preferences of Consumers
on the Formation of Price-Demand Relation*

Our talk concerns certain properties of collective behaviour in economic and social systems. More precisely, we analyse how demand for a good or a service in a population of consumers depends on interactions between consumers and on heterogeneity of their individual preferences. We construct a stochastic model of the evolution of consumers' decisions (in respect to whether to consume or not); in this model, each decision is ruled by the following factors: (1) the price of the consumed good or service; (2) the current demand for this good or service; (3) the individual preference of the consumer whose decision is being modelled. The factor (2) expresses in our model the fact that social interactions affect the formation of demand; it makes our model be similar to processes studied in the Statistical Mechanics and in the Interacting Particle Systems. We analyse our model and conclude that there may be either one or two equilibria for the demand; which one will be the case, depends on the strength of social interaction and on the heterogeneity of the distribution of the individual preferences over the population. This result explains why in the reality, there may be more than one demand for the same value of price (for example, why two similar goods may have similar prices, but, nevertheless, one may be much more popular than another one). This result enhances the arguments of Gary S. Becker that show that the non-uniqueness relation between demand and price is caused by social interaction (see G. S. Becker, "A Note on Restaurant Pricing and Other Examples of Social Influences on Price" in Journal of Political Economy, Vol. 99, Nr. 5 (1991), or G. S. Becker and K. M. Murphy, "Social Economics: market behavior in a social environment" by Belknap Press of Harvard University Press (2000)).

Ana M. Bianco

Tests robustos en modelos de regresión parcialmente lineales

Sean $(y_i, x_i', t_i)'$ observaciones independientes tales que $y_i \in R$, $t_i \in R$, $x_i = (x_{i1}, \dots, x_{ip})' \in R^p$ y

$$y_i = \beta' x_i + g(t_i) + \epsilon_i \quad 1 \leq i \leq n, \quad (1)$$

donde los errores ϵ_i son independientes e independientes de $(x_i', t_i)'$. Supondremos que para $1 \leq j \leq p$, $x_{ij} = \phi_j(t_i) + z_{ij}$, donde los errores z_{ij} son independientes e independientes de t_i . Este modelo, que ha sido estudiado por varios autores, generaliza al modelo lineal y es más flexible, dado que incluye una componente no paramétrica. Como en los casos de regresión lineal y regresión no paramétrica, los estimadores de mínimos cuadrados en este modelo pueden verse seriamente afectados por la presencia de datos anómalos. Esta charla comenzará introduciendo un procedimiento de estimación robusto de 3 pasos para este modelo y describiendo las propiedades asintóticas del estimador $\hat{\beta}$.

Además de la importancia de desarrollar estimadores robustos, también merece atención el problema de test de hipótesis. En muchas situaciones, estamos interesados en determinar el impacto de una covariable sobre la variable de respuesta. O sea, deseamos hacer inferencia sobre el parámetro e o sobre alguna de sus componentes. Sea $\beta = (\beta'_1, \beta'_2)'$, donde $\beta_1 \in R^q$. Nos interesa testear las siguientes hipótesis paramétricas:

$$i) H_{0\beta} : \beta = \beta_0 \quad ii) H_{0\beta_1} : \beta_1 = \beta_{10}, \beta_2 \text{ no especificado}$$

y propondremos una clase de tests basados en el procedimiento de estimación de 3 pasos mencionado antes. Se obtendrá el comportamiento asintótico del estadístico del test bajo la hipótesis nula y bajo la alternativa y, a través de un estudio Monte Carlo, se ilustrará la performance de los tests propuestos.

Finalmente, teniendo en cuenta que en muchas situaciones es necesario hacer inferencia sobre la función de regresión, nos interesará el problema

$$i) H_{0g} : g = g_0$$

Presentaremos una alternativa robusta al test clásico (González Manteiga y Aneiros Pérez (2003)), cuyo comportamiento asintótico bajo la hipótesis nula y bajo alternativas contiguas se obtuvo en forma heurística. Finalmente, se presentarán los resultados de un estudio de simulación que permite comparar el comportamiento en nivel y potencia del test propuesto y del test clásico, para muestras normales y para distintas contaminaciones. (Trabajo conjunto con Graciela Boente y Elena Martínez)

Bernardo de Lima

Dependent Bernoulli Percolation on Z^2

Let $\delta, p_1, p_2 \in (0, 1)$. In Z^2 , the set of vertical edges whose ends have the same y -coordinate will be a horizontal stair; analogously, the set of horizontal edges whose ends have the same x -coordinate will be a vertical stair. Consider the following bond percolation model: each stair will be good or bad, independently, with probability $1 - \delta$ or δ respectively; in the sequel, given the configuration of bad and good stairs, we declare each edge as open or closed, independently, where the probability of each edge is open is p_1 or p_2 if the edge belongs to a bad or a good stair, respectively. We show that $\forall p_1 > 0$ and p_2 close to 1, we can take δ small enough, such that there is percolation in this model. This result implies that there is percolation on the random stretched lattice in $d = 2$, a conjecture proposed by Jonasson, Mossel and Peres. Joint work with H. Kesten, V. Sidoravicius and M.E. Vares.

Pablo de Napoli

Sobre problemas inversos en finanzas

Se presentara un panorama del llamado "problema inverso en finanzas": la determinación del valor de la volatilidad implícita a partir de los precios de mercados de las opciones. Se discutirá su carácter de problema mal planteado, y los diferentes métodos propuestos para su abordaje. Se discutirá el caso de las opciones europeas y asiáticas (trabajo en desarrollo).

Gilles Didier

An introduction to some mathematical approaches of genomics

The goal of this course is to present an overview of several mathematical and statistical approaches used to solved genomic questions. The introduction will be devoted to a brief description of biological notions involved in genomics and of the various types of data available (DNA, proteins, microarrays data...) will be introduced. The first part concerns the field of sequence analysis. In particular some methods applied to the following topics will be discussed: - sequence modelling: Markov Chains and Hidden Markov Model - genome annotaion, - evolution of sequences: models, phylogeny. - sequence comparison: pairwise and multiple alignment. Finally, some problems arising from analysis of transcriptom data will be presented, including: - microarrays data analysis, - regulation network.

Denise Duarte

Bootstrapping Chains of infinite order

Apresentamos uma nova metodologia de reamostragem para estimar a variancia da mdia de uma funo de uma cadeia de ordem infinita. Aplicamos esta metodologia para o caso especial das cadeias esparsas onde sabemos calcular a variancia verdadeira a ttulo de comparao e validao do mtodo de reamostragem que propomos e a seguir apresentamos uma aplicao em Lingustica, construindo intervalos de confiana para a sonoridade.

Pablo A. Ferrari

Limit theorems for sequences of random trees

We consider a random tree and introduce a metric in the space of trees to define the “mean tree” as the tree minimizing the average distance to the random tree. When the resulting metric space is compact we show laws of large numbers and central limit theorems for sequences of independent identically distributed random trees. As application we propose tests to check if two samples of random trees have the same law. (Joint work with David Balding, Ricardo Fraiman and Mariela Sued.)

Diana Divanilda Ferreira

Testing VLMC Trees Modeling Stress Contours in Brazilian and Portuguese Written Texts

In this working paper we test the goodness of fit of two trees supporting VLMC models for brazilian and european portuguese written texts.

Renato Fontes

Aging and scaling limits

Aging is a memory effect of disordered systems, like spin glasses, in low temperature, where the system, after starting off at high temperature and being cooled down abruptly, subsequently spends ever longer times at a sequence of metastable configurations. This is usually described formally in terms of a scaling relation obeyed by a two-time correlation function of an appropriate stochastic dynamics. Our approach is to obtain a suitable scaling limit for the dynamics, yielding the relation which signals aging as a corollary, thus shedding light at a deeper scaling relation behind the phenomenon. In this talk, we present briefly some processes for which this program has been/is being carried out, including one dimensional Glauber dynamics at zero/very low temperature (like the voter model) and Bouchaud models.

Daniel Fraiman

Estimacion y tests en Procesos Autoregulados

Existen muchos ejemplos donde la hipótesis Markoviana no es adecuada, algunas veces por desconocimiento de alguna variable y otras porque el verdadero proceso tiene memoria. En este trabajo introducimos una nueva clase de procesos con memoria, llamados autorregulados, estimadores de los parámetros y una manera de testear si una realización de un proceso estocástico proviene de un determinado proceso autorregulado. Por ultimo presentamos dos casos particulares: 1) Estimación del alcance, 2) Estimación del orden.

Ricardo Fraiman

Identifying rhythmic classes through the sonority of languages: a Kolmogorov-Smirnov approach

In the linguistic literature it has been conjectured that languages are divided into different classes according to their rhythmic properties (see for instance, Lloyd 1940, Pike 1945, Abercromie 1967). More recently Ramus et al (1999) considered this problem using sentences in different languages. Recently Galves et al (2004) propose to use a rough measure of sonority defined directly from the spectrogram of the signal in order to find acoustic correlates of the rhythmic classes of speech signals. Following this approach we will try to answer the question of how can languages be divided into classes according to their rhythmic properties using 20 sonority curves for each of 7 different languages. The main tool we will use in to perform Kolmogorov type test for the two sample problem when data are trajectories of stochastic processes, based on the results in Cuesta-Albertos, Fraiman and Ransford (2004). (Joint work with: Juan Cuesta, Antonio Galves, Jesus Garcia and Marcela Svarc.)

Alejandro Frery

Agrupamento de dados com o modelo de Potts Superparamagnético

Recentemente foi proposta uma técnica de agrupamento de dados que utiliza o modelo de Potts de maneira inovadora: ao invés de considerá-lo a distribuição "a priori" em um procedimento bayesiano, as dissimilaridades entre observações são utilizadas como as interações entre partículas. Cada observação é associada a uma partícula e deixa-se, então, o sistema evoluir em uma região da temperatura conhecida como "região superparamagnética". A cada temperatura é estimada a correlação entre os estados das partículas, e os grupos são formados com aquelas partículas que exibem correlação superior a um certo limiar. Embora a técnica tenha demonstrado o seu potencial em uma série de aplicações, restam ainda muitas perguntas a serem respondidas: quais são as propriedades teóricas do procedimento? qual ou quais as dinâmicas mais adequadas para alcançar resultados relevantes? qual é a relação entre as métricas de dissimilaridade e a qualidade dos agrupamentos? Esta palestra pretende apresentar a técnica, mostrar alguns resultados de aplicação, fazer considerações computacionais e levantar outras questões de relevância para a compreensão e o aprimoramento do método. Referências BLATT, M.; WISEMAN, S.; DOMANY, E. Superparamagnetic clustering of data. *Physical Review Letters*, v.76, n.18, p.3251-3254, 1996. DOMANY, E.; BLATT, M.; GDALYAHU, Y.; WEINSHALL, D. Superparamagnetic clustering of data: application to computer vision. *Computer Physics Communications*, v.121-122, p.5-12, 1999. GEMAN, D. e GEMAN, S. Stochastic relaxation, Gibbs distributions and the Bayesian restoration of images. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, v.6, n.6, p.721-741, 1984. KAUFMAN, L.; ROUSSEEUW, P.J. Finding groups in data: an introduction to cluster analysis. New York: Wiley, 1990. WISEMAN, S.; BLATT, M.; DOMANY, E. Superparamagnetic clustering of data. *Physical Review E*, v.57, n.4, p.3767-3783, 1998.

Sacha Friedli

Retrieving rhythmic features in linguistic random sequences

We define the problem of truncation as follows: Is it that the occurrence of phase transitions, in two-dimensional systems with long range interactions, depends only on some bounded part of the interaction? We answer affirmatively to this question in some long range percolation and spin models with heavy tails.

Antonio Galves

Retrieving rhythmic features in linguistic random sequences

In this talk I will present some recent results which aim to rigorously address the following central problems in Linguistics:

-the question of the existence of rhythmic patterns in natural languages;

- the existence of a discrete typology characterized by well-defined critical points, as opposed to a rhythmic continuum;

-the existence of rhythmic features in the acoustic signal and in written texts.

Besides obtaining new mathematical results which are interesting by themselves, the goal is to use the conceptual framework of Probability Theory to effectively interpret linguistic data.

I believe that the probabilistic models and results developed to address linguistic issues could also be of some interest for people studying genetic data and I will say a few words about this.

The key words of this tutorial are: the rhythmic class hypothesis, speech sonority, chains of infinite order, variable length Markov chains and tied quantized chains.

Nancy García

Finding correlates of rhythm in written texts using variable length Markov chains

We propose a Variable Length Markov Chain approach to detect the rhythm of written texts. Texts from 20th century from Brazilian and Portuguese authors (henceforth respectively EP and BP) were codified in such way that all the syllables are marked as stressed or not, beginning of a phonological word (a lexical word plus all non stressed words on its left) and the final period. The goal is to estimate patterns for each text that discriminate between EP and BP.

Jesús García

*A stochastic model for the speech sonority:
tied quantized chains and cross-linguistic estimation of the cut-points*

A rough measure of the speech sonority was introduced in Galves et al. (2002). In the present paper we introduce the notion of tied quantized chains which provides an adequate model for the data. The quantized chains corresponding to different languages are tied together by the assumption that the distribution of the sonority, conditioned to the fact that it belongs to certain fixed intervals are language independent. We define cross-linguistic estimators for the cut-points separating these intervals and prove their asymptotic consistency. We check the predictions of the model on a corpus with 1667 sentences from eight different languages. We show that a model with four cut-points is in good agreement with the data. This is a joint work with Marzio Cassandro, Pierre Collet, Denise Duarte and Antonio Galves.

Marta García Ben

Estimación robusta para el modelo de regresión con errores asimétricos

Se propone una clase de estimadores robustos para el modelo de regresión con errores que siguen una distribución logaritmo-gamma. Estos estimadores se extienden a otros modelos de regresión con errores asimétricos y continuos. Existen varias clases de estimadores robustos para el modelo de regresión lineal que son simultáneamente muy resistentes a valores atípicos (alto punto de ruptura) y altamente eficientes si el modelo se cumple y los errores tienen distribución gaussiana. Una clase de estimadores son los MM propuestos por Yohai (1987). Tanto los MM estimadores como otros estimadores robustos propuestos para el modelo de regresión lineal le dan menor peso a los residuos que son "grandes" en valor absoluto. Por esta razón, cuando estos estimadores se aplican a datos que siguen un modelo de regresión con errores asimétricos, el estimador de la función de regresión tiene sesgo y no es consistente. En este trabajo estudiamos el comportamiento de los M-estimadores para modelos de regresión con errores con distribución asimétrica, continua y unimodal y generalizamos los MM-estimadores para algunos modelos de regresión que incluyen el modelo con errores logaritmo-gamma. (Trabajo conjunto con Ana M. Bianco y Victor J. Yohai).

Pablo Groisman

Aproximación numérica de ecuaciones estocásticas con soluciones que explotan

Las soluciones de ecuaciones estocásticas pueden explotar en tiempo finito si los coeficientes no son globalmente Lipschitz. Es decir que existe un tiempo finito (aleatorio) T , tal que la solución de la ecuación tiende a infinito al acercarse a este tiempo. En estos casos los métodos numéricos tradicionales no son aplicables ya que están basados en hipótesis de regularidad de las soluciones. En esta charla desarrollaremos un método numérico diseñado especialmente para este tipo de situaciones y mostraremos sus propiedades asintóticas.

Milton David Jara Valenzuela

Estabilidad del coeficiente de difusión para el proceso de exclusion simple

En 1985, C. Kipnis y S.R.S Varadhan probaron que la posición de una partícula marcada en un proceso de exclusión simple simétrico, cuando debidamente re-escalada, converge para un movimiento Browniano de matriz de difusión $D(a)$, dependiente de la densidad de partículas a . Posteriormente este resultado fue extendido para el caso de media cero, y el caso asimétrico en dimensión mayor o igual que 3. El método utilizado en la demostración de este resultado es válido tanto en el caso de la red entera, como en el caso de una red periódica finita. Será probado, entonces, que el coeficiente de difusión en el caso infinito puede ser aproximado por su contraparte en el caso periódico

Claudio Landim

Macroscopic current fluctuations in stochastic lattice

We study current fluctuations in lattice gases in the macroscopic limit extending the dynamic approach to density fluctuations developed in previous articles. More precisely, we derive large deviation estimates for the space-time fluctuations of the empirical current which include the previous results. Large time asymptotic estimates for the fluctuations of the time average of the current, recently established by Bodineau and Derrida, follow easily and are improved and generalized in several respects. Some explicit examples are discussed.

Regina Liu

Mining Massive Text Data and Developing Tracking Statistics

We present a systematic data mining procedure for exploring large free-style text datasets to discover useful features and develop tracking statistics, often referred to as performance measures or risk indicators. The procedure includes text classification, inference under error measurements and risk analysis. An aviation safety report repository from the FAA is used to illustrate applications of our research to aviation risk management and general decision-support systems. Some specific text analysis methodologies and tracking statistics are discussed. Several approaches for incorporating misclassified data or error measurements into the inference for tracking statistics are proposed and evaluated.

This is joint work with Daniel Jeske, Department of Statistics, UC Riverside.

Regina Liu

Clustering Genes Based on P-Values

In many microarray experiments, gene expression data consist of repeated measures in multiple conditions. However, this experimental setup is often not fully accounted for in the existing approaches for clustering genes. We present in this talk a new clustering methodology which can account for the exact experimental setup. This methodology compares genes by testing the equality of the condition-mean vectors and condition-variances across experimental conditions, and views the resulting P-value as a standardized measure of similarity between genes for the clustering purpose. Using P-value as a

measure of similarity between genes is supported by the statistical framework of hypothesis testing, and is thus less arbitrary than the existing choices such as Euclidean distance, correlation, etc. In addition, the P-value based clustering can be easily calibrated for different separating criteria by adopting different tests. We also use our P-value based clustering method to combine gene selection and gene clustering into one step. Thus, clusters of genes are selected rather than individual genes, which increases power of detection.

The proposed P-value based clustering is applied to both simulated and real data sets. The preliminary findings appear to be quite supportive.

This is joint work with Rebecka Jornsten and Jun Li, Department of Statistics, Rutgers University.

Alejandro Maass

Re-written of genome sequences

En esta charla se presenta un modelo matemático basado en elementos de dinámica simbólica que permiten reinterpretar secuencias biológicas en función de la dependencia espacial de sus símbolos.

Fabio Machado

On an epidemic model over finite and infinite graphs

We study a system of simple symmetric random walks on the vertices set of finite and infinite connected graphs. At the beginning there is one particle at each vertex of a graph G , being all particles inactive, except for the one which is placed at the origin, a chosen site of G . Inactive particles are waiting to be activated. The active particle performs a simple random walk with (random) lifetime X on the set of vertices of G , moving along its edges. When an active particle hits an inactive particle, the latter becomes active. One can also consider a random amount of particles per vertex. The graph structure, the lifetime (distribution) of active particles and the (random) amount of particles per vertex are ingredients whose combination will guide the model to different behaviors. Our goal is to present rigorous results followed by simulations that either illustrate them or guide us to open questions. The model is meant to be both simple enough to allow for a good level of rigorous tractability, and flexible enough to be applied to real modeling situations. Just to mention an example, it can be considered for modeling the spread of a virus over a net of computers or for the transmission of a rumors trough a large group of people. In this talk, questions (many) and answers (some) on shape theorems, phase transition results, recurrence/transience, cover times and extinction times on finite and infinite connected graphs will be delivered.

Ricardo Maronna

Componentes principales robustas

Consideremos una muestra multivariada de n puntos en dimensión m . El análisis de Componentes Principales se puede describir como: encontrar una variedad lineal de dimension dada que minimice la media cuadrática de las distancias ortogonales de los datos a la variedad. La solución se obtiene a través de los vectores propios de la matriz de covarianzas. Se sabe que basta que una pequeña proporción de los n puntos sean atípicos para que las componentes sean seriamente alteradas. Los dos enfoques principales para robustificar las Componentes Principales son: reemplazar la matriz de covarianza por una matriz robusta de dispersion, y buscar p direcciones ortogonales, tales que las correspondientes proyecciones maximicen o minimicen una medida robusta (univariada) de dispersión ("projection pursuit"). El presente trabajo es próximo al segundo enfoque. Pero en vez de encontrar las componentes una a una, se busca, para un p dado, una variedad de dimensión p que minimice una escala robusta de las distancias ortogonales de los datos a la variedad. La escala usada es un M-estimador de escala, que puede considerarse como una media cuadrática con pesos que dependen de los datos. Se presenta un algoritmo iterativo, que converge a un mínimo local. Se propone una estrategia de búsqueda aleatoria para aproximar el mínimo global. Se muestra que este método es más rápido que sus competidores, especialmente para dimensión m grande. Se propone un método eficiente para determinar un p adecuado. Las comparaciones basadas en datos

reales y simulados muestran que este método es más robusto que sus competidores. También se considera la situación -frecuente en análisis de imágenes y de datos funcionales- en que cada uno de los n valores observados puede independientemente tener un error grosero.

Elena Julia Martinez

Estimadores robustos para el modelo lineal multivariado basados en escalas tau

Se presenta una clase de estimadores robustos para modelos lineales multivariados. Los coeficientes de regresión y la matriz de covarianza de los errores son estimados simultáneamente minimizando el determinante de la matriz de covarianza, sujeto a una restricción en una escala robusta de la norma de Mahalanobis de los residuos. Eligiendo como escala robusta un s -estimador, los estimadores combinan buenas propiedades de robustez y eficiencia asintótica bajo normalidad de los errores. Los estimadores propuestos tienen distribución asintótica normal y, si los errores tienen distribución elíptica, su matriz de covarianza asintótica difiere sólo en un factor escalar de la correspondiente al estimador de máxima verosimilitud. Se deriva la curva de influencia y se prueba que el punto de ruptura es próximo a 0.5. Se describe un algoritmo iterativo de cálculo basado en estimadores de máxima verosimilitud ponderada y se presentan los resultados de un estudio Monte Carlo cuyo objetivo fue analizar la robustez y la eficiencia de los estimadores propuestos. Finalmente, se discute un ejemplo con casos reales. (Trabajo conjunto con Marta García Ben y Víctor Yohai).

Servet Martinez

Un modelo Markoviano y un criterio de selección simple de anotación en procariontes

Proponemos y analizamos un modelo simple de selección de candidatos a genes basado en un modelo Markoviano, y en un criterio de maximización sugerido por las estadísticas derivadas de la anotación de procariontes. El análisis de este modelo prueba que existen fuertes correlaciones entre las hebras de DNA para la selección de candidatos a genes, a pesar que en su inicio ellas se suponen independientes (hipótesis no realista). Este trabajo es fruto de nuestras observaciones de anotación en el LBMG, colaboración con Andrew Hart (Núcleo Milenio), y en el ajuste a datos de anotación con Leonardo Videla (LBMG).

Ernesto Mordecki

Modelos Estocásticos en Finanzas

En esta charla, de carácter introductorio haremos una revisión de los modelos más utilizados en la modelación matemática de la dinámica de los mercados financieros, en la cual hay dos posibilidades de inversión, una determinista, los bonos B (o bonds), y otra estocástica, aleatoria o contingente, las acciones (stock) S , junto con los problemas de valuación de opciones (y eventualmente futuros) en estos modelos. El plan es el siguiente: 1- Modelos en general: Stocks (aleatorio) y Bonds (determinista) - opciones - futuros. 2- Modelos en tiempo discreto: Cox-Ross-Rubinstein, nociones de arbitraje y de mercados completos. 3- Modelos en tiempo continuo continuo: Black-Scholes 4- Generalizaciones: Modelos de volatilidad estocástica (difusiones) y Modelos con procesos de Lévy. 5- Algunos resultados nuevos en modelos con procesos de Lévy.

Leandro Pinto Rodrigueus Pimentel

The time constant and critical probabilities in percolation models on Voronoi tilings

We consider the first-passage percolation and the percolation problems on the graph $G = (V, E)$, where the vertex set V is a Poisson process in R^2 and the edge set E is composed by pairs (v, v') such that the Voronoi tile located at v , $C_v := \{x \in R^2; |x - v| < |x - v''| \text{ for all } v'' \text{ different from } v\}$, and the Voronoi tile located at v' , $C_{v'}$, share a one dimensional boundary. The partition of the plane $\{C_v; v \in V\}$ is called the Voronoi tiling and graph G is called the Delaunay triangulation. The difficulty with an approach of this type is that there necessarily exist arbitrarily large Voronoi tiles and hence the model

has long-range dependence. By using the static renormalization method we “control” this long-range dependence and relate the behavior of the time constant, as a function of the passage time distribution, with the value of critical probabilities in standard percolation models on G .

Ana Maria Sfer

Estimación del Efecto Causal en Modelos no Lineales

En el modelo lineal normal, la aleatorización produce un estimador insesgado para el efecto tratamiento aditivo, aún cuando importantes covariables sean omitidas o no medidas. (Cox, 1958). Sin embargo en modelos no lineales, este resultado no es claro (Gail, Wiand y Piantadosi, 1984). En esta presentación se muestra el efecto de la aleatorización en modelos no lineales con covariables omitidas. Se revisa las propiedades de las estimaciones por máxima verosimilitud y se propone una nueva forma de definir el Efecto Causal del Tratamiento de manera que sea posible extender el resultado para conocido para modelos no lineales.

Marcela Svarc

*Curvas de máximo sesgo
para estimadores de regresión basados en S-estimadores*

Se considera el modelo de regresión lineal con p regresores y intercept. Supongamos que tenemos n observaciones (x_i, y_i) que satisfacen

$$y_i = \alpha + \beta^t \mathbf{x}_i + \mathbf{e}_i \quad \mathbf{1} \leq \mathbf{i} \leq \mathbf{n}, \quad (2)$$

donde $\mathbf{x}_i \in \mathbf{R}^p$ es el vector de covariables correspondiente a la observación i , $\beta \in \mathbf{R}^p$ y α son los coeficientes de regresión (pendientes y ordenada en el origen) y e_i es el error de regresión. Suponemos que (\mathbf{x}_i, u_i) son i.i.d y u_i es independiente de x_i . Llamemos H_0 la distribución conjunta de (\mathbf{x}, \mathbf{y}) bajo este modelo. Para permitir una fracción ϵ de contaminación en los datos supondremos que la verdadera distribución H de (y_i, \mathbf{x}_i) pertenece al entorno de contaminación de H_0 dado por

$$\mathcal{V}_\epsilon = \{\mathcal{H} : (1 - \epsilon)\mathcal{H}_0 + \epsilon\mathcal{H}^* \text{ donde } \mathcal{H}^* \text{ es una distribución arbitraria}\}$$

Sea T en \mathbf{R}^p un estimador funcional equivariante de (α, β) , definido en un subconjunto de funciones de distribución H en \mathbf{R}^{p+1} . Una medida invariante de robustez está dada por el máximo sesgo de T en \mathcal{V}_ϵ .

Berrendero-Zamar proponen modificar la ordenada en el origen de los S-estimadores de manera que la mediana sea 0. En este trabajo se propone otra modificación de los S-estimadores que tiene la misma propiedad.

El objetivo del trabajo es comparar los máximos sesgos de tres estimadores: el S-estimador, la variante de Berrendero-Zamar y la nueva propuesta. (Trabajo conjunto con V.J.Yohai)

G. Tomás Tetzlaff

Variantes estocásticas de los métodos matriciales en biología poblacional

Los modelos que utilizan las denominadas matrices de proyección se usan extensamente tanto para pronosticar el estado futuro de una población como para evaluar hipótesis sobre su historia. Las propiedades de estas matrices, incluyendo las asintóticas, son susceptibles de interpretación biológica. Por otro lado las matrices de Leslie, pionero de estos modelos, resultaron ser objetos de interés matemático. Aquí describimos el modelo determinístico original de Leslie y los procesos de ramificación que corresponden a su versión estocástica. Exponemos problemas y posibles aplicaciones de la extensión de estos últimos al caso de infinitos tipos.

Julia Villar

Estimadores robustos para el modelo logístico multinomial

Para el modelo logístico multinomial se propusieron M-estimadores de clase Mallows y se demostró que éstos eran consistentes, asintóticamente normales y con función de influencia acotada. Se estudió su comportamiento para pequeñas muestras, con y sin contaminación, mediante simulaciones y se comparó su performance con la de los estimadores de máxima verosimilitud. (Trabajo conjunto con V. J. Yohai)

Mario Wschebor

*De los campos vectoriales aleatorios a la complejidad algorítmica:
algunos ejemplos de aplicación de la fórmula de Rice*

Las llamadas "Fórmulas de Rice" permiten calcular los momentos del número de preimágenes de un punto por un campo vectorial aleatorio. La exposición estará destinada a mostrar cómo es posible aplicar estas fórmulas a la resolución de algunos temas vinculados con la complejidad de algoritmos. Más precisamente, después de presentar esas fórmulas, mencionaremos dos tipos de aplicaciones: 1) El teorema de Shub-Smale sobre los sistemas polinomiales y sus extensiones. 2) El condicionamiento de ciertos problemas de Análisis Numérico y su relación con el espectro de matrices aleatorias.

Victor Jaime Yohai

Robust regression with censored data

We extend the MM-estimates, S-estimates, least median of squares and tau-estimate to the case of a linear model where the responses are censored. These estimates are Fisher consistent, equivariant and highly robust. We will also present the results of a Monte Carlo study.

Flavio Augusto Ziegelmann

*A Robust Nonparametric Estimator of Volatility Functions
for Heavy-Tailed Distributions*

We apply the local linear least-absolute-deviations (LAD) method to estimate conditional volatility functions, having its asymptotic properties established. The importance of the LAD estimator arises from the need to have a robust estimator for the case where heavy-tailed distributions occur. It is well known that least-squares (LS) estimators tend to fail in this case. In our numerical examples, when we simulate from a model in which the fourth conditional moment of the observed variable does not exist, the LAD estimator shows superior performance. In addition, these numerical implementations strongly suggest the asymptotic adaptiveness of the LAD estimator of the volatility to the unknown conditional location quantity (mean or median). Also, insightful results are found when studying SP500 and IBOVESPA real data sets.