

Resúmenes:

XVIII Escuela de Probabilidad y Estadística EPE2020

<http://epe2020.eventos.cimat.mx/>

Mini cursos:

Octavio Arizmendi (CIMAT)

“Energía de Gráficas”

Resumen:

En esta serie de charlas se dará una introducción al tema de energía de gráficas. un concepto que se origina en química matemática pero que es interesante cómo un objeto matemático. Explicaremos también un concepto nuevo que denominamos energía de vértices y que ayuda a entender mejor la energía total de la gráfica..

Andrés Christen (CIMAT)

“Cuantificación de Incertidumbre (UQ)”

Resumen:

Recientemente se ha involucrado la estadística en los problemas inversos, y su solución usando estadística bayesiana se le llama *Uncertainty Quantification* o UQ. Esta área ha tomado bastante fuerza con muchos investigadores sumándose al tema. En resumidas cuentas, se trata de inferencia estadística que involucra parámetros desconocidos en sistemas de ecuaciones diferenciales, ordinarias o parciales, que típicamente provienen de modelos de la física. En este mini curso daremos una introducción al tema, mostraremos sus retos e interés, y como el tema se encuentra en la intersección de matemáticas aplicadas, análisis numérico, computación y estadística.

Ehyter Martín (DEMAT)

“Extremos en teoría de riesgo”

Resumen:

La teoría de extremos cuenta con diversas aplicaciones en problemas relacionados con hidrología, ecología, finanzas y riesgo. En particular, es una teoría sumamente útil para estimar valores de cuantiles grandes cuando la verdadera distribución de los datos de interés cumple algunas condiciones no muy restrictivas. En este curso veremos algunos resultados clásicos de la teoría de extremos y presentaremos las aplicaciones de tales resultados al cálculo del VaR (Value at Risk, Valor en Riesgo) y estimación de probabilidades de ruina, entre otras cantidades de interés.

Pláticas:

Mario Díaz, IIMAS, UNAM

“Privacidad en Machine Learning: Retos y Oportunidades”

Resumen:

Detrás de muchos servicios digitales contemporáneos hay una gama amplia de técnicas de Machine Learning que sistemáticamente capturan y procesan información. Si bien el éxito de algunas de estas técnicas es indiscutible, estas representan amenazas potenciales a la privacidad de los usuarios. En esta plática presentaremos una noción de privacidad conocida como 'privacidad diferencial', así como algunas herramientas matemáticas de utilidad en su cuantificación y análisis.

Emilien Joly (CIMAT)

“How to understand geometry with Gaussian processes?”

Resumen:

The subject of Gaussian processes is very rich even though these objects are very simple to define. It turns out to be the central object of astonishingly good theories such as Brownian motions, Isoperimetry or Concentration of measure. These very diverse fields form a variety of leads to understand deep properties of the geometry of some complex sets. As a basic introduction to the field, we will expose these links on very simple concrete contexts and we will expose the consequence in terms of concentration of measure. As time permits, we will expose some arising conjectures on the form of convex sets in high dimensions that would be tackled by Gaussian processes!

Juan Carlos Pardo (CIMAT)

“¿Cuántas veces hay que barajar un mazo de cartas para que esté aproximadamente ordenado de forma aleatoria?”

Resumen:

La respuesta es 7, pero ¿cómo es que se obtiene este resultado? En esta plática vamos a explorar las matemáticas detrás de esta respuesta y deducir por qué es cierta. El fenómeno que ocurre al responder esta pregunta es interesante per se y es conocido como la convergencia abrupta, algo que ocurre en muchos modelos aleatorios que convergen al equilibrio.

José Luis Pérez Garmendia

“Ansioso por ejercer tu seguridad? Calma, a veces es óptimo esperar”

Resumen:

En esta charla veremos algunos resultados clásicos de la teoría de paro óptimo para valorar seguridades americanas. En particular se discutirán resultados clásicos cuando el precio del activo está modelado por un proceso de Lévy asimétrico. Finalmente se discutirán generalizaciones a estrategias de paro periódicas y resultados recientes cuando la tasa de descuento en el mercado deja de ser positiva.

Rogelio Ramos (CIMAT-CONACYT)

“Métodos Bayesianos empíricos y control FDR”

Resumen:

El material de esta plática está basado en Efron (2010). Su relevancia actual es crítica pues aborda el problema del control de tasas de descubrimientos falsos, en el contexto de pruebas de hipótesis múltiples, donde "múltiples" significa cientos o miles de pruebas de hipótesis, lo cual es importante de considerar, por ejemplo, en Genómica. Veremos cómo el procedimiento clásico Benjamini-Hochberg (1995) puede justificarse desde un punto de vista Bayesiano.

Carlos Vargas (CIMAT)

"Construyendo estadísticos para contrastar la forma de los datos"

Resumen:

Una de las ideas centrales en el Análisis Topológico de Datos es la incorporación de nuevos parámetros estadísticos que tomen en cuenta propiedades de la forma que tienen los datos, como periodicidad o la topología de su espacio ambiente. La plática pretende presentar, a grandes rasgos, la heurística detrás de los métodos de clasificación basados en estos nuevos parámetros, así como los conceptos elementales que estos requieren para su definición. Nos enfocamos en dos ejemplos prácticos, uno de ellos propone mejoras a métodos de clasificación no invasivos de enfermedades cardíacas, a partir de electrocardiogramas cortos.

Leticia Ramírez (CIMAT)

"Sobre los modelos epidemiológicos en redes y su inferencia"

Resumen:

En esta charla se presenta los modelos epidemiológicos compartimentales de transmisión directa más comunes (SIR y SEIR) que evolucionan en una red de contactos. Esta red describe el tipo de contactos entre individuos que pueden resultar en la transmisión de la infección. Abordamos el problema de inferencia de este modelo y se presentan algunos modelos auxiliares que se estudian para realizar la inferencia en forma más eficiente.

Carteles:

Aldo Ernesto Escobedo

“Evaluación del desempeño de jugadores de basquetbol mediante gráficos de tiro”

Resumen:

Los gráficos de tiros (shot charts) son una herramienta de visualización de datos utilizada en el basquetbol, que permite ubicar en la cancha la posición en la que se realiza cada tiro, así como su resultado (anotación o fallo), de modo que a través de este tipo de gráficos es posible determinar las regiones de la cancha en las que ciertos jugadores o equipos son más efectivos en sus tiros. Este trabajo se enfoca en presentar algunas variaciones de los gráficos de tiros utilizando el paquete

estadístico R para su implementación, y se muestra la utilidad de estos gráficos en el análisis del desempeño ofensivo de un jugador de basquetbol para generar posibles estrategias de juego. Se toman, a manera de ejemplo, los datos del jugador Kobe Bryant durante toda su carrera profesional.

Ana Victoria Flores

“Brecha Salarial de Género en México”

Resumen:

Ante la ley hombres y mujeres somos iguales. Desearíamos que esto ocurriera en todos los aspectos importantes de la vida diaria; sin embargo, esto no sucede, por ejemplo, en el aspecto laboral.

Para darnos una idea, en 2018 la participación económica de los hombres fue de 77.5 % y la de las mujeres de 43.7 %.

En esta investigación se busca probar por medio de la estadística que existe una brecha salarial en el país con base al género de las personas. Para lo cual, se buscó información de los salarios en México, de acuerdo con el género, los sectores de trabajo y el porcentaje de la población que es económicamente activa.

Posteriormente, se busca clasificar los salarios en los siguientes: salario bajo, medio y alto. Finalmente, probar que, en México, el salario depende del género.

Arely Maldonado Azcona.

“Modelado y predicción del PIB trimestral mediante modelos VAR”

Resumen:

A menudo se desea conocer el comportamiento futuro de distintos fenómenos de interés basándose en observaciones del pasado. Para ello es posible hacer uso del análisis de series de tiempo, el cual nos permite entender y extraer información que; con ayuda de distintos métodos matemáticos, se utilizará para realizar un pronóstico a través de un modelo.

Los métodos para hacer dicha predicción, se usan según las características de las variables. Cuando se tiene más de una serie temporal, y cada una no sólo está relacionada con observaciones del pasado, sino que además depende de otras variables, es necesario hacer otro tipo de análisis; similar al del caso univariado. El objeto de interés de este trabajo es el estudio y aplicación de los modelos VAR (análogo multivariado de los Procesos Auto-resesivos). En este trabajo se hace el análisis del PIB trimestral de actividades primarias, secundarias y terciarias a través de los modelos VAR y la metodología de Box-Jenkins, así como del software Python.

Bernardo Rodríguez Corona

“Algoritmos de simulación para cadenas de Markov en tiempo continuo”

Resumen:

Como estudiantes de ciencias, sabemos que las matemáticas, en particular los procesos estocásticos, son una herramienta muy poderosa en el estudio y análisis de enfermedades tales como el VIH y el ERV, ya que podemos estudiar el comportamiento y evolución de los padecimientos, de este modo nos dimos a la tarea de investigar algunos métodos de simulación que pudieran servir al estudio de dichas enfermedades, para esto usamos como herramientas las cadenas de Markov. Es sabido que existe una plétora de situaciones en donde se aplican las cadenas de Markov, por ejemplo: en finanzas, medicina, química, biología etc. En general se usan para determinar las transiciones (aleatorias) que existen de un estado a otro,

ya sea en su versión a tiempo continuo o a tiempo discreto.

Dichas situaciones se pueden modelar mediante una simulación que consiste en la utilización de ciertas técnicas matemáticas, empleadas en computadores, las cuales permiten imitar el funcionamiento de prácticamente cualquier tipo de operación o proceso. El proceso de simulación se hace mediante el uso de algoritmos, es decir, una serie de pasos de forma sistematizada que se emplea para hallar la solución de un problema en particular.

Nosotros estamos interesados en tres algoritmos para poder simular las cadenas de Markov. Por lo que la finalidad del presente informe es mencionar las características más importantes de los tres diferentes algoritmos que nos permitirán analizar dos modelos estocásticos sobre infecciones, el VIH (Virus de Inmunodeficiencia Humana) y el ERV (Enterococos Resistentes a la Vancomicina), y con esto poder dar conclusiones en específico de cada método, poder compararlos entre sí, poder afirmar si existe alguno mejor y bajo que circunstancia resultaría más pertinente el uso de uno que de otro.

Cesar Andrés Cárdenas Alvarado

“Abuso en la Significancia Estadística en la Ciencia”

Resumen:

¿Qué es la Ciencia? La Real Academia Española define a la ciencia como: “Conjunto de conocimientos obtenidos mediante la observación y el razonamiento, sistemáticamente estructurados y de los que se deducen principios y leyes generales con capacidad predictiva y comprobables experimentalmente”. Independientemente de la definición de Ciencia que adoptemos, al ser conocimientos adquiridos mediante la experiencia o la razón, puede que no sean ciertos... desde un punto de vista filosófico NUNCA sabremos que conocimientos son verdades absolutas, citando al Antiguo Problema del Criterio.

¿Qué es la Estadística? El Cambridge Dictionary of Statistics nos la presenta de la siguiente manera: “La Estadística es un método intelectual y general que aparece donde sea que hayan datos, variación o aleatoriedad”. El pensamiento estadístico es un proceso natural en el ser humano, de hecho, existen estudios que comprueban que los bebés y niños pequeños tienen ciertas nociones e intuiciones estadísticas. La Estadística nos proporciona herramientas para “medir” la plausibilidad de una hipótesis en concordancia o correspondencia con las observaciones. Esto lo podemos hacer mediante las pruebas de hipótesis y de significancia.

Informalmente las pruebas de significancia “miden” que tan razonable es que el azar no haya influido en los resultados de nuestro experimento. Podemos explicar la idea de este tipo de pruebas, por medio del ejemplo de la Dama del Té. Cierta dama afirma que puede distinguir si al té se le ha diluido leche o si a la leche se le ha diluido té. Proponemos el siguiente experimento para corroborar si posee dicha capacidad: La dama debe escoger 4 tazas de té (de un solo tipo elegido al comienzo del experimento) entre 8 tazas de té (donde hay 4 y 4 tazas de cada tipo). Supongamos que la dama acierta 3 de 4 tazas. ¿Es eso evidencia suficiente para que sea plausible asumir que posea dicha habilidad? La probabilidad que ocurra dicho evento suponiendo que escogió las tazas al azar (es decir, de manera uniforme) es de $16/70$ lo cual es relativamente alto. Con lo cual, pudo haber acertado 3 de 4 tazas por la influencia del azar.

Dentro de las pruebas de hipótesis, el p-valor es una herramienta muy usada. El p-valor es la probabilidad bajo un modelo estadístico completamente especificado sea igual o más extremo que un valor observado. Se ha abusado de sobremanera del p-valor, ya que se han construido paradigmas en torno a él, sin ningún fundamento científico, el más conocido se trata de rechazar la hipótesis nula si nuestro p-valor es menor que 0.05 (sin considerar el conocimiento previo o las particularidades del fenómeno).

Los abusos anteriormente mencionados, conllevan varias consecuencias en nuestra sociedad actual. Por ejemplo, la fácil manipulación de los datos para rechazar una hipótesis

(un caso muy impactante es el del analgésico Vioxx, que probablemente sea el envenenamiento más grande en la historia de la humanidad y la retirada más importante de un medicamento a nivel mundial), obtener conclusiones poco “creíbles” y el sesgo de publicación.

Aún podemos remediar ésta situación, y de ser posible se mencionarán algunas propuestas que se han dado para lidiar con los abusos anteriormente mencionados.

Cynthia Calderón García

“Clusters con distintas métricas”

Resumen:

El Análisis de Clusters es un método estadístico para agrupar elementos de un conjunto de datos según sus características, buscando la homogeneidad en los elementos de cada grupo y la mayor diferencia entre grupos. La clasificación se realiza tomando en cuenta distintas variables para la selección de los elementos que pertenecerá a cada grupo. De esta manera, mediante el cálculo de distancias o similitudes entre los elementos, se busca el mejor agrupamiento de éstos. La distancia y similitudes son conceptos que utilizan funciones matemáticas como herramienta para hacer dichos cálculos. En la distancia se tiene la función llamada métrica, mientras que en el caso de la similitud se utiliza la función que lleva el mismo nombre.

Algunos ejemplos de las métricas más utilizadas son: Distancia Euclídea, Métrica de Chebychev, Distancia de Manhattan. Por otro lado, la función de similitud más utilizada es el coeficiente de correlación de Pearson.

Hacer agrupaciones permite identificar las tendencias y necesidades de cada grupo, aprovechando las condiciones de cada uno para mejorar el rendimiento de los datos. También ayuda a realizar modelos que describan su comportamiento y con esto hacer estimaciones a futuro.

Así, ésta técnica es utilizada en diversos campos con el propósito de hacer agrupaciones de los objetos de estudio para optimizar su análisis, por ejemplo: en ciencias de la vida, ciencias de la tierra, medicina, ingeniería, ciencias sociales, ciencias de la información y política.

Gabriel Cao Romero Martínez, Bernardo Rodríguez Corona, Roberto Capitaine Venegas

“El impacto del narcotráfico en la economía mexicana”

Resumen:

Estamos convencidos que el objetivo principal de ser universitarios es cultivar nuestros conocimientos y trabajar nuestros talentos para ponerlos al servicio de la sociedad. Al ser actuarios desarrollamos habilidades en la probabilidad, estadística, economía, entre otros, y al ir adentrándonos entre estas materias, fueron saliendo interrogantes en nuestro camino. Una de ellas fue, ¿Cuál es el impacto del narcotráfico en la economía mexicana? Y quisimos abordar esta situación ya que representa un gran problema en cuestión económica, social, seguridad, salud, etc. Entonces a partir de lo que hemos aprendido en la licenciatura, hicimos un análisis macroeconómico sobre cómo afecta el narcotráfico a un indicador económico muy importante, el producto interno bruto (PIB), usando herramientas como la matriz de correlación y regresiones lineales principalmente.

Nuestro objetivo es explicar de manera concisa y detallada los impactos del narcotráfico en la economía de México, hacer una aproximación para determinar que tanto podría aumentar o disminuir nuestros indicadores económicos. Analizar la matriz de correlación generada por las variables y ver si estadísticamente el impacto es significativo.

Así como, enunciar algunos ejemplos de países donde se ha legalizado el consumo de drogas, y observar que implicaciones ha tenido en su economía, para de esta manera hacer una comparación objetiva y brindar nuestras conclusiones sobre nuestros resultados.

Se analizaron estadísticas del PIB tomadas del INEGI de cada trimestre desde el año 2000 hasta el segundo trimestre del año 2018. Se realizó una matriz de 2x2 de correlación con 140 observaciones, correspondientes a cada trimestre respecto a todos los años analizados. Además, mediante el uso de elementos gráficos, trabajamos en determinar si estadísticamente el impacto es significativo o no.

Al ser el trabajo una aplicación, principalmente macroeconómica, fue de suma importancia hacer una investigación teórica sobre los expertos que ya han trabajado este tema. Tras el análisis, obtuvimos principalmente dos vertientes, por un lado, se sustenta que el impacto económico es favorable por la cantidad de dinero que mueve el narcotráfico, directa o indirectamente. Por otro lado, se argumenta que el impacto es desfavorable por los problemas sociales que esto implica, creando más gasto en el sector salud o materias de seguridad, ya sea, policía o ejército, entre otros.

En nuestros resultados obtuvimos que, en la matriz de correlaciones, el narcotráfico como variable está fuertemente ligado con los ingresos que se producen anualmente en México. Finalmente, este también genera violencia lo cual produce inestabilidad en la sociedad, y esto trae numerosas consecuencias como la incertidumbre, que a su vez propicia una disminución en las inversiones locales y extranjeras, y esto a su vez afecta la economía en términos negativos produciendo fuga de capital.

Isidro Vargas-Morenoa, Juan Francisco Rodríguez-Landa c, Héctor Gabriel Acosta-Mesad, Rafael Fernández-Demeneghia, Socorro Herrera-Mezab

“Reconocimiento de patrones conductuales relacionados con la ansiedad y la actividad motora producidos por el consumo de aceite estercúlico en la rata Wistar utilizando aprendizaje estadístico”

Resumen:

En la zona de los Tuxtles, México, las semillas de *Sterculia apetala* (orden: Malvales, familia: Sterculiaceae) son consumidas por los pobladores; estas contienen ácidos grasos ciclopropenoicos como el malválico y el estercúlico, cuyo consumo se ha relacionado con efectos benéficos y adversos a la salud. El aceite estercúlico (AE) se ha propuesto como posible agente terapéutico contra la obesidad, debido a que genera reducción en peso corporal, adiposidad, triglicéridos y presión arterial, no obstante, se desconoce si su consumo, ejerce efectos sobre la actividad motora y la ansiedad. Por otro lado, en el área de neurofarmacología la técnica más utilizada para el análisis de datos es el ANOVA o la t de Student, sin embargo, algunas técnicas de aprendizaje estadístico como árboles de decisión y gráficos de comportamiento transicional implementadas mediante cadenas de Markov, permiten reconocer y visualizar relaciones entre variables, con el fin de observar patrones conductuales producidos por la administración de algunos tratamientos. El objetivo del presente estudio fue evaluar el efecto del consumo de aceite estercúlico sobre la ansiedad y la locomoción en la rata Wistar, utilizando técnicas de aprendizaje estadístico. En este estudio se realizaron dos experimentos: 1) Identificación de la dosis efectiva del AE, donde se utilizaron 26 ratas Wistar divididas en un grupo control y tres grupos tratados con dosis de 0.06 g, 0.12 g, y 0.24 g de AE; 2) Identificación de patrones conductuales producidos por la dosis efectiva del AE, donde se utilizaron 16 ratas Wistar, divididas en un grupo control y un experimental. En ambos experimentos se utilizaron las pruebas de conducta de laberinto de brazos elevados (LBE) y campo abierto. Los resultados del primer experimento mostraron que la dosis efectiva fue de 0.12 g de AE, ya que produjo una disminución del tiempo empleado en brazos abiertos, de la frecuencia y del tiempo en la conducta de head-dipping, así como del tiempo empleado en el acalamiento respecto al grupo control. Se concluye de este estudio que el AE produce

un efecto de tipo ansiogénico a nivel conductual. En el segundo experimento se identificaron patrones conductuales característicos del consumo de AE, mostrando cambios en la incidencia de una conducta a otra, relacionados con un estado de ansiedad. En general, se concluye que la administración de AE ejercen un efecto ansiogénico en el LBE, por lo que el consumo de esta semilla debe limitarse en la alimentación. Aunado a lo anterior, el análisis de datos se ve complementado con las técnicas de aprendizaje estadístico, debido a que mostraron información adicional sobre los cambios de conducta que en el primer experimento no fueron evidentes; tales como las interacciones temporales entre las conductas inducidas por el consumo del aceite estercúlico.

Julia Trinidad Reyes

“Estudio del sistema compartido de bicicletas: Ecobice”

Resumen:

En el trabajo realizado se presenta un estudio de la red de bicicletas compartidas ECOBICI, sistema de bicicletas públicas de la Ciudad de México, el cual ha integrado a la bicicleta como parte esencial de la movilidad como modo de transporte dirigido a los habitantes de la capital, de sus alrededores y turistas. El estudio se centra en la disponibilidad de bicicletas y lugares de estacionamiento para éstas en las estaciones de la colonia Polanco. No se sabe (con la base de datos que proporciona ECOBICI) si existen este tipo de problemas en la red, ni cuales son las estaciones más afectadas si es que las hay. Se modeló la entrada y salida de bicicletas en el sistema usando Teoría de Colas (algún tipo de proceso de migración cerrado), Estadística y Simulación Estocástica, con el objetivo de identificar problemas en el servicio y proponer mejoras a éste, esto es, se estimó la frecuencia con la que se solicitan préstamos en cada estación o se guardan bicicletas que ya fueron ocupadas, para poder brindar suficientes equipos a cada cicloestación (que no se queden sin bicis) y al mismo tiempo tener lugares disponibles para guardarlos. Al final se propuso un sistema de autorregulación “Two-Choice” y se analizaron los posibles resultados de su implementación en ECOBICI (Polanco).

Juan Francisco Muñoz Elguezabal, Riemann Ruiz Cruz

“Clustering subsecuencial de series de tiempo: Evidencia de patrones temporales en el tipo de cambio UsdMxn”

Resumen:

Una gran cantidad de procesos naturales, financieros, tecnológicos y sociales producen grandes cantidades de datos ordenados en el tiempo, es decir, datos del tipo series de tiempo. El análisis de este tipo de datos ha captado la atención de una gran parte de la comunidad de científicos de datos y, argumentablemente, técnicas de “clustering” o “agrupamiento” son unas de las principales para el análisis de este tipo de datos. Este trabajo es sobre Clustering SubSecuencial de Series de Tiempo, una técnica que busca agrupar subsecuencias contenidas dentro de una misma serie de tiempo, por medio del cálculo de un término de distancia euclidiana a manera de medida de similitud entre los datos. Trabajos de investigadores en la Universidad de California Riverside han mostrado la carencia de sentido en el uso de ciertas técnicas de clustering para series de tiempo, particularmente K-Means con “ventanas móviles” puesto que utilizando conjuntos de datos completamente diferentes se obtenían centroides estadísticamente similares. Los autores de UC Riverside mencionan que otros investigadores han conducido las pruebas sugeridas y han manifestado la reproducibilidad de los resultados en sus propios trabajos, los cuales han sido principalmente series de tiempo generadas por procesos biológicos y físicos. Algo a

destacar es que, de acuerdo a lo escrito por los autores, estos hallazgos obtenidos invalidan los resultados de decenas de trabajos previamente publicados.

Hacemos uso del algoritmo propuesto por el grupo de investigación de UC Riverside, MASS (Mueen's Algorithm for Similarity Search), para la identificación de patrones temporales en la subsecuencia de precios intradía del tipo de cambio Dólar americano Vs Peso Mexicano (UsdMxn). La hipótesis principal es que existen condiciones bajo las cuales se puede clasificar el comunicado de un indicador macroeconómico en México o Estados Unidos como un evento disparador que genera patrones temporales en la serie de tiempo del tipo de cambio UsdMxn.

Una búsqueda quasi-exhaustiva de evidencia es conducida utilizando 10 años de información, 14.5 Millones de precios (OHLC de cada minuto), 36,000 mediciones de indicadores macroeconómicos. Debido a la cantidad de datos a procesar, el uso de un cluster de 64 núcleos ha sido planteado como parte fundamental del proyecto, tomando como ventaja que el algoritmo MASS soporta la paralelización y velocidad de convergencia de su proceso de búsqueda y ajuste.

Finalmente, con estos resultados aportamos evidencia del uso de técnicas modernas para análisis de datos del tipo series de tiempo financieras, desde un punto de estadística para el clustering subsecuencial de series de tiempo, y evidencia sobre el comportamiento de mercados financieros. Los resultados que mostramos son consistentes y documentamos las condiciones bajo las cuales, el algoritmo K-means con “ventanas móviles” produce resultados inferiores al MASS, y también, condiciones bajo las cuales no se cumple la Hipótesis del Mercado Eficiente.

Luis Ramón Munive Hernández

“Spatio-temporal Point Process Analysis of Forest Fires in Mexico State”

Resumen:

Naturally forest fires can be studied and analyzed with spatial statistics methods, specially with a focus from spatio-temporal point processes, in this contribution we explore if complete spatial randomness is present, clustering or a systematic distribution on the fires.

For a most informative data exploration, some mark variables are considered such as fire cause, afectation kind, impact level on environment, duration.

A data exploration considering only spatial point processes is made for a good interpretation of forest fires spatial pattern behavior. The process intensity is a essential parameter because through this give us a idea of regions with the highest fires occurrences, other intensity type is considered , called second order intensity and it is the analogous to a correlation function as well as first order intensity is to a mean.

There is an interesting detail in the data set, there are missing data, i.e. occurrences without his coordinates, the treatment for this situation is estimating the missing values, as occur in the spatial analysis and the other statistics areas, sometimes data isn't complete and we must fix it, thus in the spatial case the values of neighboring can be used for our objective, and in the spatiotemporal case neighbors in space and time can be used to fill it data missing.

Also a complete analysis is made integrating the spatial and temporal process components, in this we can decide if the process is separable and thus compute the correlation level between this parts as well as modelling the process for forecast future fires, among other things that we can derive. By last, we conclude some recommendations for avoid forest fires, based on obtained result from the analysis.

References

- [1] Camarero J. J., Pompa-García Marín and Rodríguez-Trejo D. A. (2017), Drought and spatiotemporal variability of forest fires across Mexico. Chinese Geographical Science.
- [2] Díaz-Avalos C., Juan P. and Mateu J. (2010), Characterizing spatial-temporal forest fire patterns. 5th International Workshop on Spatio-Temporal Modelling.
- [3] Diggle P. J. (2005) Spatio-temporal Point Processes: Methods and Applications. Johns Hopkins University, Dept. of Biostatistics Working Papers. Working Paper 78.
- [4] Illian J., Penttinen A., Stoyan H. and Stoyan D. (2008), Statistical analysis and modelling of spatial point patterns. John Wiley & Sons.

- [5] Lestari A., Rumantir G., and Tapper N. (2016), A Spatio-temporal Analysis on the Forest Fire Occurrence in Central Kalimantan, Indonesia. PACIS 2016 Proceedings. 90.
- [6] Møller J. and Waagepetersen P. (2007), Modern Statistics for Spatial Point Processes. 21st Nordic Conference on Mathematical Statistics.
- [7] Møller J. and Waagepetersen P. (2003), Statistical Inference and Simulation for Spatial Point Processes. CRC Press.
- [8] Stoyan D. (2006), Fundamentals of Point Processes Statistics en Case Studies in Spatial Point Processes Modeling. Springer Science & Business Media.

Manuela Torres

“Un ejercicio de análisis de Caries dental usando R, en pacientes entre 25 y 45 años de edad la Clínica de la universidad El Bosque, Bogotá (Colombia)”

Hoy en día, hay una gran variedad de programas que permiten el tratamiento de datos para que puedan ser analizados. R es una herramienta valiosa y de fácil acceso para los estadísticos. Para reconocer su efectividad y presencia en las ciencias de la salud se tomarán datos de la clínica ubicada en la universidad El Bosque y se estimaran los índices epidemiológicos Índice CPO-D, Índice de Clune, para las caries dentales que permiten evaluar el estado bucal de los pacientes, además a través de la ANOVA se comparara el promedio de dientes sanos CPO-D y sus componentes según la edad.

Bibliografía

- Maindonald, J., & Braun, J. (2006). Data Analysis and Graphics Using R: An Example- based Approach. Cambridge University Press.
- Mayra, P., & Cardoso, D. (2014). Indicadores de prevalencia y de predicción. Revista Especializada en Ciencias de la Salud, 61-72.

Roberto Capitaine Venegas, Marypaz Navarrete González, Gabriel Cao Romero Martínez, Bernardo Rodríguez Corona

“Autonomía del Banco Central y la inflación en México”

Resumen:

Un banco central es una institución pública que gestiona la moneda de un país o grupo de países y controla la oferta monetaria, es decir, la cantidad de dinero que está en circulación. El objetivo principal de muchos bancos centrales es la estabilidad de precios.

Ahora bien, en la estabilización de los precios entra en juego la tasa de inflación, es decir, el aumento anual porcentual del nivel general de precios. Muchos economistas sostienen que la independencia del banco central favorece el control de la inflación, donde juega un papel vital, y así, ayudan a generar una estabilidad en los precios. Por esto, aseguran que el banco central debe de separarse del Estado. No obstante, esta independencia del banco central no se da en todos los países, hoy en día, por ejemplo, bancos como el de Venezuela siguen siendo controlados por el Estado.

La independencia del ahora llamado BANXICO comenzó en 1993 y a partir de esa fecha, nos hemos enfrentado como país a innumerables cambios y retos en materia económica, como lo fueron el “Efecto Tequila” en 1994 o como la crisis del 2008. Hoy, a 26 años de la otorgación de la autonomía al Banco de México, es momento de ponernos a pensar ¿Realmente ha sido benéfica la decisión de separar al Banco de México del gobierno?

La investigación tiene como objetivo analizar el desarrollo de las tasas inflacionarias que se han presentado en México dentro del periodo de 1970 a

2018. Se han utilizado pruebas estadísticas, una regresión discontinua, con el objetivo de rechazar o no rechazar la hipótesis de la existencia de una correlación entre la autonomía del Banco de México y los movimientos de las tasas inflacionarias. De la misma forma, se han generado intervalos de confianza para conocer la variabilidad de las tasas de inflación y conocer el nivel de correlación que existe entre la tasa de inflación y el PIB.

La importancia de la estadística en el trabajo de investigación es principalmente resumir una gran cantidad de datos, en datos representativos y significantes para un análisis más preciso.

Víctor Miguel García Sánchez

“La volatilidad de un activo financiero como movimiento Browniano fraccionario”

Resumen:

Un proceso gaussiano W^H es un Movimiento Browniano fraccionario (fBm) si es autosimilar de parámetro de Hurst $H \in (0,1)$ y tiene incrementos estacionarios. Podemos probar que tiene función de covarianza

$$\begin{equation} \mathbb{E}(W^H_t W^H_s) = \frac{1}{2}(t^{2H} + s^{2H} + |t-s|^{2H}). \end{equation}$$

Sean $t < s \in \mathbb{R}$ y $h > 0$ el tamaño de los incrementos de tiempo, de forma que $t+h \leq s$.

Denotemos la covarianza entre incrementos con la notación de [biagini_fractional_2008](#), $\rho_H(n) = \text{Cov}[W^H_{s+h} - W^H_s, W^H_{t+h} - W^H_t]$.

Si $H > \frac{1}{2}$, entonces W^H tiene incrementos correlacionados positivamente y decimos que W^H posee memoria larga. Por otro lado, si $H < \frac{1}{2}$, entonces

W^H tiene incrementos correlacionados negativamente. Tal caso puede usarse para describir sistemas que cuentan con memoria y persistencia, aunque no tendrá la

suavidad que caracteriza al modelo con $H > \frac{1}{2}$, por el contrario, diremos que es áspero. Gatheral, et al. 2018, definen con base en lo anterior al modelo de volatilidad

estocástica fraccionaria áspera (RFSV) por fBm con $H < \frac{1}{2}$, por lo tanto, tiene el potencial de ser coherente con las propiedades observadas empíricamente de las

series temporales de volatilidad. Estimando la volatilidad con datos recientes de alta frecuencia, revisamos la cuestión de la suavidad del proceso de volatilidad. Y observamos que el modelo RFSV es

notablemente consistente con los datos financieros de series de tiempo; Una aplicación es que nos permite obtener pronósticos mejorados de volatilidad realizada. Además,

encontramos que aunque la volatilidad no es un proceso de memoria larga en el modelo RFSV, los procedimientos estadísticos clásicos que apuntan a detectar la persistencia

de la volatilidad tienden a concluir la presencia de memoria larga en los datos generados a partir de ella.