

Colaboración

Software

Actividades da SGAPEIO

Novas do IGE

Traballos de Estatística

Coñecéndonos

Sabías que?

Editorial

TOMÁS COTOS YÁÑEZ – VICEPRESIDENTE DA SGAPEIO

Estimados compañeiros e compañeiras:

É un pracer para min dirixirme de novo a todos vós con motivo da presente edición dun número novo do boletín da nosa sociedade INFORMEST. Como xa saberedes na reunión do Consello Executivo da SGAPEIO do 21 de Novembro de 2013 constituíse un novo Consello Executivo presidido por D^a M^a Esther López Vizcaíno. Neste consello acordouse nomear a:

- Secretaria da sociedade: D^a. Balbina Casas Méndez
- Vicepresidente de Estatística: D. Tomás R. Cotos Yáñez
- Vicepresidenta de Investigación de Operacións: D^a. Silvia María Lorenzo Freire
- Directora do Boletín: D^a. Paula Raña Míguez
- Xestoras da web: D^a. Soly Santiago Pérez e D^a. María José Ginzo Villamayor

Agradecer aos membros saíntes do consello executivo anterior D. Antonio Vaamonde Liste (Presidente), D. Pedro Faraldo Roca (Vicepresidente de Estatística), D^a. M^a Amalia Jácome Pumar (codirectora de INFORMEST), D^a. M^a José Lombardía Cortiña e D^a. M^a del Carmen Iglesias Pérez polo seu desinteresado e continuado traballo pola sociedade.

Accións destacables neste primeiro ano foron, entre outras, a nova edición (IV) do Concurso Incubadora de Sondaxes e Experimentos, o Acto de clausura do Ano Internacional da Estatística, a publicación do número 37 do boletín INFORMEST e o I Encontro Luso-Galaico de Estatística no Medio Ambiente e na Ecoloxía que cando se faga público este boletín xa estará realizado. Tamén ao longo deste ano fixéronse moitas actualizacións e melloras na web da sociedade. No que queda de ano teremos unha Xornada de formación para elaborar proxectos de estatística e o Curso de introdución a R: Software libre para PEMES. No vindeiro ano 2015 seguiremos cunha nova edición do Concurso Incubadora de Sondaxes e Experimentos tanto a fase local como a nacional. Para rematar dicirvos que todos estes eventos están dispoñibles na web da sociedade e que esperamos foran/sexan do voso agrado.

Tomás R. Cotos Yáñez

Vicepresidente de Estatística. SGAPEIO

Dirección:

Paula Raña Míguez

Comunicación coa SGAPEIO:

www.sgapeio.es
secretaria@sgapeio.es
boletin@sgapeio.es
Twitter @sgapeio
LinkedIn Sgapeio

Depósito Legal:

LU-191-1995 - I.S.S.N.:1695-7083



SGAPEIO

Sociedade Galega para a
Promoción da Estatística
e da Investigación de
Operacións

Facultade de Matemáticas
15706 Santiago de Compostela





Aplicacións e técnicas da Estatística na Investigación Social

por Rebeca Villaverde López

Rebeca Villaverde é estatística, licenciada en Matemáticas. Toda a súa carreira profesional está vinculada ao sector da Investigación de Mercados e Opinión. Actualmente é a responsable de Anova Multiconsulting. Dende Novembro de 2013 é vogal no Consello Executivo da SGAPEIO.

Introdución

Ante o gran abano de empresas que fan un uso frecuente da Estatística vou priorizar naquelas que se dedican fundamentalmente á investigación social. En boa medida, por ser o ámbito máis estendido e divulgado nesta era na que o marketing cobra unha especial relevancia.

Os comportamentos adquiridos pola poboación son obxecto de estudo e análise na investigación social e empréganse principalmente para tomar decisións de índole comercial. Aplicacións como poden ser a previsión de demanda dun determinado produto ou o estudo de correlación entre dous produtos substitutivos en potencia son algúns dos exemplos prácticos. Preguntas como: ¿Que queren os meus clientes? ¿Están satisfeitos co servizo ou produto que lles ofrezco? ¿Como podo captar novos clientes? ¿Será aceptado positivamente este novo produto como substituto doutro? E un longo etcétera de cuestións que xeran respostas de gran interese para a empresa. Ao redor deste conto expoño a argumentación do artigo.

En futuras edicións desta sección do boletín ampliaremos a información a outros ámbitos empresariais nos que se emprega cada vez máis a Estatística. Neste punto ten gran importancia o sector industrial, sen esquecermos doutros de recente implantación como é o das novas tecnoloxías. O control de calidade é unha das aplicacións estatísticas máis habituais no mundo da empresa dedicada á produción, pero veremos outras técnicas e sen deixar de lado obviamente as técnicas de mostraxe máis utilizadas en cada ocasión.

Asociacións de empresas e profesionais

No ámbito que nos ocupa, o da investigación social e comercial, ademais de empresas especializadas como é o Centro de Investigacións Sociolóxicas (CIS), gustaríame destacar a existencia de asociacións, que se atopan na nosa contorna máis próxima, e que verdadeiramente

son as que velan polos intereses do sector e as que nos achegan datos do mercado actual das empresas e dos profesionais que desenvolven estas investigacións. Estou a falar de ANEIMO (Asociación Nacional de Empresas de Investigación de Mercados e Opinión pública), AEDEMO (Asociación Española de Estudos de Mercado, Marketing e Opinión) e ESOMAR (European Society for Opinion and Marketing Research).

AEDEMO agrupa a profesionais que desenvolven a súa actividade nestas áreas. Esta asociación confórmana ao redor de 800 profesionais e o seu obxectivo fundamental réxese en promover o crecemento da profesión, mediante a difusión de metodoloxías que facilitan a correcta obtención de datos a través de estudos e investigacións.

ANEIMO está formada polas empresas máis importantes en investigación de mercados e actualmente as súas empresas asociadas representan o 60 % do volume de negocio do sector. A entidade traballa en asegurar que cada traballo se realiza seguindo os máis altos estándares de calidade e ademais en velar polo respecto dos códigos de ética profesional.

ESOMAR é unha organización internacional con preto de 5.000 membros pertencentes a 130 países. Promove o valor da investigación de mercados e estudos de opinión, resaltando os temas de verdadeiro interese para unha toma de decisións eficaz.

Tanto as cifras de ingresos do sector de investigación de mercados como as previsións son alentadoras a pesar do contexto económico presente.

Facturación do sector da investigación social

O recente VI ENCONTRO “La investigación a análisis”, convocado de forma conxunta por AEDEMO, ANEIMO e ESOMAR, conclúe que as previsións do sector de investigación de mercados son optimistas, dentro da moderación, no ámbito español. Pola súa banda o “Estudio sobre el Sector de la Investigación de Mercados 2013” realizado conxuntamente por Aneimo e Aedemo estima que a facturación deste sector en España ascende a 438.208.000 €.

Unha vez reflectidas as cifras de ingresos do sector e a previsión para un futuro próximo, que nos permiten ter datos de referencia sobre o peso do sector da investigación social e comercial, continúo enumerando os principais tipos de métodos empregados na actualidade dentro deste contexto.

DESAGREGACIÓN POR MÉTODO DE INVESTIGACIÓN DATOS MUNDIAIS 2013

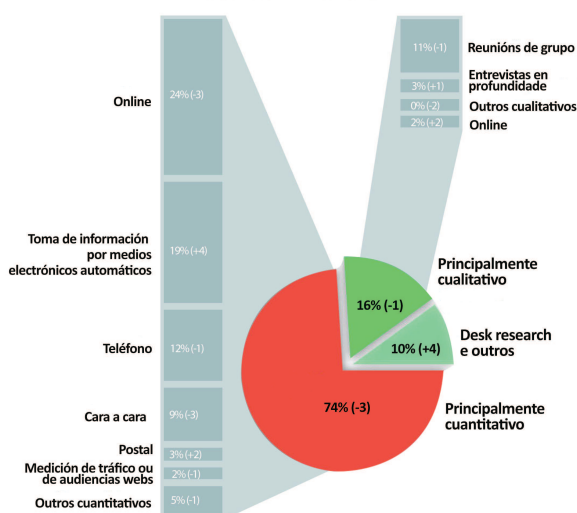


FIGURA 1: Datos de facturación mundiais según método de investigación.

Técnicas de investigación social. A enquisa:

Na investigación social, o tipo de estudos que máis se desenvolven divídense en dous tipos: os cualitativos e os cuantitativos.

Os estudos cualitativos baséanse en principios teóricos filosóficos e o seu fin é investigar o comportamento humano, por iso úsanse basicamente no ámbito social. Entre eles, cabe mencionar técnicas como o Focus Group ou grupo de discusión, o Mystery shopper, as Entrevistas en profundidade e a técnica Delphi.

Os estudos cuantitativos pola súa banda empregan información numérica que se pode extrapolar á poboación obxecto de estudo xa que teñen unha base científica. Son dous, os censos e as populares enquisas.

As enquisas, tan recorrentes e coñecidas como consecuencia do uso que se fai delas nos medios (sobre todo as famosas enquisas electorais e as de estudos de audiencias) pódense clasificar segundo distintos criterios tales como os métodos de recollida de información ou pola súa propia temática. Atendendo á clasificación segundo o método de recollida, existen as enquisas persoais,

enquisas telefónicas a fogares, enquisas por correo postal e as en liña. Cada unha delas presenta importantes vantaxes e inconvenientes. Existe suficiente bibliografía respecto diso das vantaxes e inconvenientes que presenta cada unha das súas tipoloxías pero o que si é importante é recoñecer a enquisa como método máis empregado, sendo a técnica cuantitativa por excelencia.

A utilización de técnicas cuantitativas en España están 10 puntos porcentuais por enriba do resto do mundo. Isto vai en detrimento das técnicas cualitativas, ás que se recorre menos aquí que a nivel mundial.

Ao realizar unha radiografía da facturación de ambos os tipos de técnicas maioritarias a nivel mundial e nacional, pódese apreciar que o 84,1 % das investigacións sociais de España realízanse con técnicas cuantitativas, fronte a un 74 % a nivel mundial (datos do ano 2013). Ademais, a partir do ano 2011 as técnicas cualitativas experimentaron unha evolución negativa en España perdendo cada vez máis presenza, mesmo en maior medida que as cuantitativas.

	2010 vs 2009	2011 vs 2010	2012 vs 2011	2013 vs 2012
CUANTITATIVA	-3,1 %	-1,7 %	-6,5 %	-4,9 %
CUALITATIVA	1,1 %	-11,7 %	-22,4 %	-4,6 %

Fonte: ANEIMO/AEDEMO.

A pesar de que existen defensores e detractores dos dous tipos de técnicas, segundo os datos da táboa anterior en España demándanse cada vez menos investigacións cualitativas. De todos os xeitos non é conveniente defender exclusivamente unha única vía para realizar investigacións. En moitas ocasións complementanse e a miúdo unha técnica mixta que usa ambas as técnicas, "cualis" e "cuantis" como se coñecen na xerga deste gremio, é a máis recomendable para lograr un coñecemento máis completo e alcanzar os obxectivos perseguidos.

Algunhas aplicacións prácticas:

Para concluír este artigo introductorio, gustárame destacar que as enquisas adquiriron un gran protagonismo no ámbito da comunicación con importantes funcionalidades como son as de:

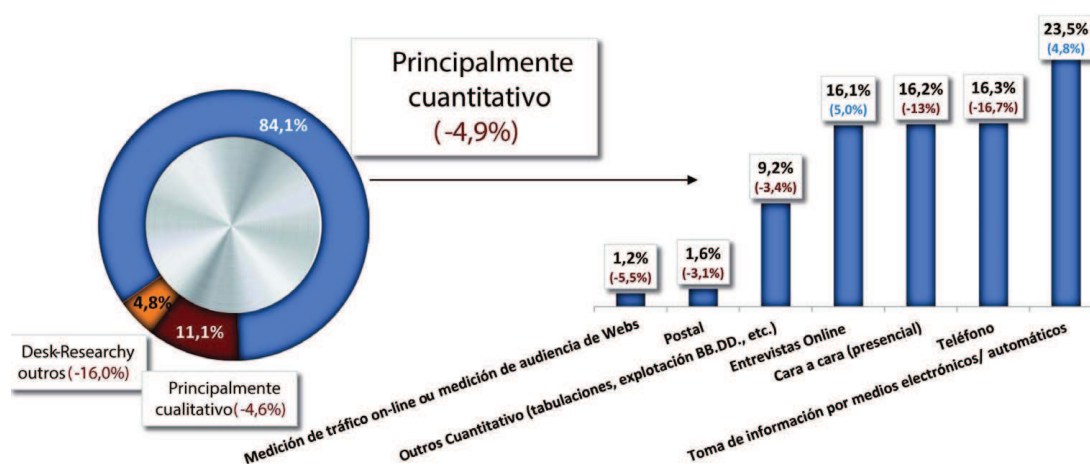
- Informar sobre hábitos de vida e consumo dunha poboación.

- Diseñar e avaliar estratexias de ofertas comunicativas no caso dos medios e de comercialización no caso das empresas.
- Realizar pre-test e post-test publicitarios para facer os balances de determinadas campañas, así como facer avaliacións continuas das mesmas (tracking).
- Realizar auditorías tanto dentro como fóra das empresas. Desta forma coñécese a imaxe corporativa da compañía, entre outras informacións de interese empresarial.
- Informar do clima laboral dunha empresa.

- Medir audiencias de determinados medios de comunicación.
- Obter información da situación financeira.

Os usos máis xeneralizados das enquisas fanse no ámbito da comunicación social e están articuladas ou son demandadas polos medios de comunicación, as axencias de publicidade ou os anunciantes. Mesmo en ocasións o seu uso esténdese a outros organismos que están vinculados con estes últimos. Por tanto, esta metodoloxía despunta no desenvolvemento comunicativo e forman parte do seu conxunto alcanzando unha situación de protagonismo.

CIFRA DE NEGOCIO POR MÉTODO DE INVESTIGACIÓN



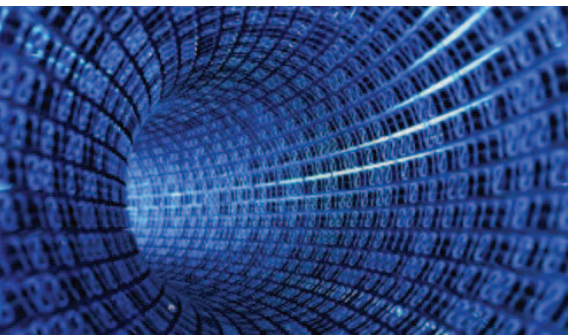
Entre paréntese, % de diferencia 2013 vs 2012.

Cifra neta de negocio: 438.202.000€. Fonte Aneimo / Aedemo. Estudio sobre o Sector da Investigación de Mercados 2013

FIGURA 2: Datos de facturación nacionais segundo o método de investigación.

Referencias

1. Juan Antonio Gaitán Moya e José Luis Piñuel Raigada. "Técnicas de investigación en Comunicación Social".
2. [Web AEDEMO](#)
3. [Web ANEIMO](#)
4. [Web ESOMAR](#)



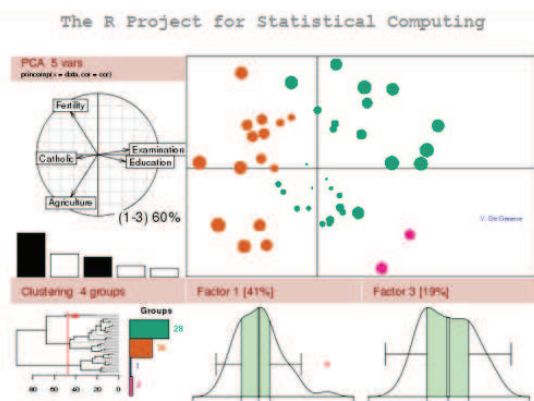
O programa R

por Javier Roca-Pardiñas

Programa R

O recente desenvolvemento da informática permitiu que na actualidade existan potentes ordenadores cunha gran capacidade de cálculo o que fixo que se puidesen automatizar os procesos numéricos subxacentes aos métodos estatísticos. No mercado existen distintos paquetes de software que para o tratamento estatístico de datos que a cotío conteñen unha grande cantidade de información. Algúns dos programas estatísticos máis coñecidos son o SPSS, Statgraphics, Minitab, Statistics, SAS, S-Plus, etc.

Fronte a estas opcións propietarias, recentemente xurdiu con forza o entorno informático R para a análise estatística, como alternativa de software libre, distribuído baixo a licenza GLP de GNU. R naceu no eido da estatística nos anos noventa, como variante libre da linguaxe propietaria S. O nome S ven de *Statistics* (Estatística) e R pronúnciase como *our* (noso en inglés) para destacar o seu carácter público. Nos últimos anos R volveuse extremadamente popular no ámbito científico e académico.



Esta popularidade vén motivada polas grandes vantaxes que presenta o programa; pode ser descargado e instalado de forma gratuíta dende a páxina web www.r-project.org; é un software libre e funciona baixo os principais sistemas operativos (Windows, MAC OS, Linux e Ubuntu); é unha potentísima ferramenta informática para o manexo, procesado, representación gráfica e análises de datos de calquera tipo; está moi ben documentado con numerosos foros de axuda; conta cunha amplísima comunidade de programadores moi dinámica que posibilita unha multitude de paquetes adicionais; permite traballar de maneira interactiva con outros programas (SPSS, Microsoft Excel, etc.) e linguaxes de programación (Fortran, C ou Matlab).

Inicialmente R foi deseñado como unha interface de programador ou de comandos, onde se teclean as instrucións que se pretenden executar. Iso pode facer que aumente o tempo de aprendizaxe daqueles usuarios acostumados ás interfaces gráficas (Excel, SPSS, etc.) onde as operacións máis comúns se realizan a través dunha serie de menús ou mediante a interacción cos obxectos gráficos que aparecen na pantalla, a través de periféricos coma o rato.

Dada esa posible dificultade, existen varias interfaces gráficas que permiten un manexo máis “amigable” de R. Unha listaxe destas interfaces encóntrase na web www.sciviews.org. Seguramente a interface máis popular para R sexa Rcommander (www.rcommander.com) que se distribúe tamén de maneira libre e gratuíta baixo a licenza GPL de GNU. Outras interfaces moi coñecidas son Rkward ([web](#)), ou RExcel e R00o ([web](#)). Estas últimas son interfaces entre o programa R e os paquetes ofimáticos Microsoft Office e OpenOffice.org, en concreto, coas súas respectivas follas de cálculo (Excel e Calc). Finalmente, en relación a interfaces de programación para R, pódese facer mención de RWinEdt e Tinn-R.

Como xa se comentou, o programa R conta cada vez con maior presenza tanto no campo académico-investigador como no campo empresarial. Proba disto foron as VI Jornadas de Usuarios de R ([web](#)) celebradas recentemente no Centro de Novas Tecnoloxías de Galicia co obxectivo de promover o coñecemento de R e as súas aplicacións, e fomentar a colaboración interdisciplinaria entre os asistentes.

Por outro lado, hai unha ampla oferta de cursos de R para distintos colectivos profesionais. A Sociedade Galega para a Promoción da Estatística e a Investigación de Operacións (SGAPEIO) en colaboración coa Facultade de Matemáticas, organiza o curso “Introducción a R: Software Estatístico Libre para PEMES” dirixido a empresas pequenas e medianas. Outros cursos ofertados son “Estatística con R para profesorado de Enseñanza Media”, “Control Estadístico de Calidade con R” e “Estatística Avanzada con R para Profesionais de la Salud”. Os tres cursos anteriores están sendo ofertados dende a Universidade de Vigo (máis información en <http://bubela.uvigo.es/>).

Paquetes de R

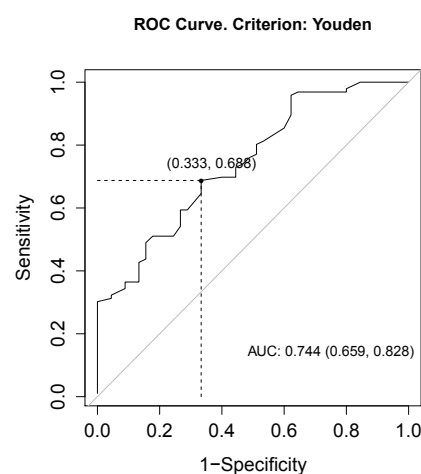
R é un programa moi flexible que permite ao usuario facer as súas propias funcións e integrais en librerías, tamén chamadas paquetes, que se poderán facer públicas ao resto da comunidade de usuarios mediante a súa publicación no CRAN. Dende este [link](#) pódese acceder e instalar os distintos paquetes dispoñibles que na actualidade son máis de 6000. Esta sección está orientada á revisión de paquetes realizados por membros do Sistema Universitario de Galicia. En particular, o presente número estará centrado no paquete `OptimalCutpoints`: Computing optimal cutpoints in diagnostic tests realizado por Mónica López Ratón (Universidade de Santiago) e María Xosé Rodríguez Álvarez (Universidade de Vigo).

`OptimalCutpoints`: Selección de puntos de corte óptimos nos tests diagnósticos

Os tests diagnósticos continuos son empregados frecuentemente para discriminar entre as poboacións de sans e enfermos a partir dos valores dun determinado marcador. Asumindo que os valores altos do marcador son indicativos da enfermidade, a clasificación da proba dará un valor positivo cando o marcador supere un determinado valor de corte. A proba tomará valor negativo en caso contrario. Na práctica clínica, na aplicación deste tipo de tests, será necesario seleccionar o punto de corte que defina os resultados positivos e negativos do test. A elección do punto de corte apropiado será fundamental para evitar conclusións clinicamente erróneas.

`OptimalCutpoints` é un paquete deseñado para seleccionar puntos de corte óptimos que pode ser instalado directamente dende este [link](#). Incorpora un total de 34 criterios empregados con frecuencia na práctica clínica baseados nas medidas de precisión diagnóstica do test (sensibilidade, especificidade, valores predictivos e razóns de verosimilitude). Algúns destes criterios tamén teñen en conta os costes das diferentes decisións do diagnóstico, e a prevalencia da enfermidade no estudo. Con este paquete, os usuarios poden obter facilmente resultados numéricos e gráficos. A saída numérica inclúe o punto de corte óptimo de acordo co método seleccionado e as medidas de precisión diagnóstica asociadas. A saída gráfica inclúe a curva características operacionais do receptor (curva ROC nas súas siglas en inglés) e a curva ROC predictiva do test diagnóstico.

Para ilustrar a aplicación do paquete considerouse un estudo levado a cabo no Servicio de Cardiología do Hospital Clínico de Santiago de Compostela cuxo obxectivo foi a selección de puntos de corte óptimos da taxa de leucocitos como marcador diagnóstico da enfermidade coronaria arterial. A continuación móstrase unha saída numérica e unha saída gráfica correspondente a este estudo.



Call:

```
optimal.cutpoints.default(X="elas", status=
"status", tag.healthy=0, methods="Youden",
data=elas, pop.prev=NULL, control=control.cutpoints(),
ci.fit =TRUE)
```

Area under the ROC curve (AUC): 0.744 (0.659, 0.828)

CRITERION: Youden

Number of optimal cutoffs: 1

Estimate	Lower Limit	Upper Limit	
cutoff	37.00	-	-
Se	0.69	0.58	0.78
Sp	0.67	0.51	0.80
PPV	0.81	0.69	0.87
NPV	0.50	0.39	0.67
DLR.Positive	2.06	1.33	3.18
DLR.Negative	0.47	0.32	0.67
FP	15.00	-	-
FN	30.00	-	-
Opt. crit.	0.35	-	-

Para máis información sobre este paquete véxase o artigo López-Ratón *et al.* (2014) publicado na prestixiosa revista de carácter computacional [Journal of Statistical Software](#).

Referencias

López-Ratón, M., Rodríguez-Alvarez, M.X., Cadarso-Suarez, C. and Gude-Sampedro, F. (2014) `OptimalCutpoints`: An R Package for Selecting Optimal Cutpoints in Diagnostic Test. *Journal of Statistical Software*, 61(8), 136.



Actividades da SGAPEIO

Xornada de formación para elaborar proxectos de Estatística.

Soly Santiago Pérez

Por segundo ano consecutivo, a SGAPEIO e AGAPEMA organizaron conxuntamente, co apoio de educaBarrié, unha xornada de formación dirixida ao profesorado de Educación Secundaria co obxectivo de mellorar a súa formación estatística. Este ano, a xornada enfocouse á elaboración de proxectos estatísticos na aula con aplicación á Incubadora de Sondaxes e Experimentos e, igual que o ano pasado, celebráronse dúas edicións nas sedes da Fundación Barrié na Coruña, o día 15 de novembro, e en Vigo, o 29 do mesmo mes.

A xornada comezou coa ponencia “Como realizar un traballo para a Incubadora”, impartida por María Jesús Casado, que é profesora no IES Daviña Rey de Monforte de Lemos. Na súa charla, María Jesús empezou facendo un repaso á historia do concurso para centrarse, a continuación, nos distintos aspectos que se deben ter en conta á hora de elaborar un proxecto estatístico e na estrutura que debe ter o informe; tamén puxo exemplos de traballos gañadores, tanto da fase local como da nacional.

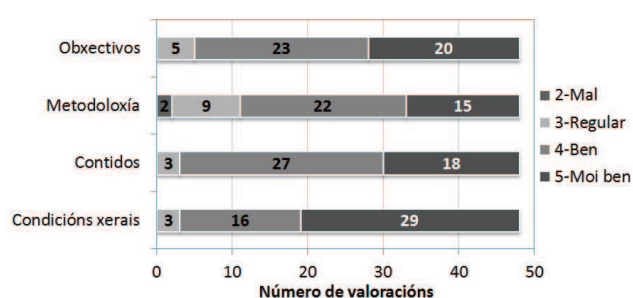
A primeira experiencia foi exposta por Mari Carmen Quiroza, profesora do IES Sánchez Cantón de Pontevedra, e que foi a titora do traballo que gañou a III edición da Incubadora na categoría de 3º–4º da ESO, presentado por un grupo de 4 alumnos do IES a Xunqueira I de Pontevedra. O traballo titulábase “E ti, como empezas o día?” e, na súa charla, Mari Carmen falounos da súa experiencia na realización do mesmo.

Os obradoiros de Epidat e GeoGebra permitiron mostrar as posibilidades destas dúas ferramentas como recursos didácticos para o seu uso na aula. [Epidat](#) é un programa de libre distribución, multiplataforma e multilingüe, orientado á análise estatística e epidemiolóxica de datos. O obradoiro, que eu impartín, estivo centrado na análise estatística das respostas á enquisa realizada no traballo “E ti, como empezas o día?”. Pola súa banda, [Geogebra](#) é un programa multiplataforma de matemáticas dinámicas para todos os niveis educativos que reúne xeometría, álgebra, folla de cálculo, gráficos, estatística e cálculo nun so programa fácil de usar. O obradoiro deste programa foi impartido por Esperanza Gesteira, Fernando Zacarías, Ignacio Larrosa e Enrique de la Torre, do grupo

[Xeodín](#), que realizaron algúns exercicios de estatística descritiva e probas de hipóteses.

Gonzalo Temperán falounos pola tarde da súa experiencia na Incubadora. El é profesor no IES Monelos de A Coruña e presentouse á I e á II edición do concurso. Na primeira o traballo que titorizaba resultou gañador na categoría de 3º–4º da ESO (“Podemos predicir a estatura á que vai chegar unha persoa?”), e representou a Galicia na I fase nacional, onde tamén resultaron gañadores. Na II edición presentouse como titor de 5 traballos, e o titulado “Cómo cae unha chincheta?”, levou unha mención na categoría de 1º–2º da ESO. Na súa exposición fixo un repaso dos distintos traballos que presentou á Incubadora.

Para rematar a xornada contamos coa intervención de Ignacio García Jurado, profesor do Departamento de Matemáticas da UDC, que impartiu a conferencia “Aplicacións da programación lineal na planificación de proxectos”. Na súa charla, Ignacio explicou algúns conceptos e métodos da planificación de proxectos e puxo como exemplo un proxecto que se podería levar a cabo nun centro educativo.



Respecto á participación, para a xornada da Coruña inscribíronse 55 persoas das que asistiron 36 (67 %); para Vigo inscribíronse 36 persoas e asistiron 27 (75 %). Ao remate de cada xornada pediuse aos participantes que realizasen unha valoración cuantitativa (de 0 a 5) e cualitativa sobre o cumprimento de catro aspectos da actividade. Os resultados cuantitativos das 48 enquisas recollidas resúmense na figura, que mostra en xeral un alto grao de satisfacción por parte dos asistentes. Dende SGAPEIO agradecemos a AGAPEMA e á Fundación Barrié a súa colaboración na organización destas xornadas, aos ponentes por aportarnos a súa experiencia, e a todos os asistentes polo seu interese e esforzo.

I Encontro sobre a Estatística no Medio Ambiente e na Ecoloxía en Portugal (EES2014).

M^a José Ginzo Villamayor

O I Encontro Luso-Galaico de Estatística no Medio Ambiente e na Ecoloxía ([EES2014](#)) celebrouse do 6 ao 8 de Novembro de 2014 en Vila Real (Portugal). Este encontro foi organizado conxuntamente pola Sociedade Portuguesa de Estatística ([SPE](#)) e a Sociedade Galega para a Promoción da Estatística e da Investigación de Operacións ([SGAPEIO](#)) e contou co apoio do Departamento de Matemática da Universidade de Trás-os-Montes e Alto Douro ([UTAD](#)) e do Departamento de Matemática e Aplicacións da Universidade do Minho ([UMinho](#)).

A realización conxunta deste encontro vén reforzar e solidificar as colaboracións xa existentes entre ambas sociedades, para a promoción e divulgación da Estatística e da Investigación Operativa. Os encontros temáticos realizados entre as dúas sociedades son unha excelente oportunidade para que os investigadores presenten os seus novos resultados, fomenten novas discusións e xurdan proxectos novidosos. A extensión destes obxectivos a toda a comunidade científica nacional e internacional fixo que os encontros fosen moi frutíferos.

O Encontro de Estatística no Medio Ambiente e na Ecoloxía estivo patrocinado polo Centro de Matemática da Universidade do Minho ([CMAT](#)), Centro de Matemática da UTAD ([CM-UTAD](#)), Centro de Investigación e de Tecnologías Agro-Ambientais e Biológicas ([CITAB](#)), Centro de Estatística e Aplicacións da Universidade de Lisboa ([CEAUL](#)), o Instituto Nacional de Estatística de Portugal ([INE](#)), o [Concello de Vila Real](#), [PSE-Produtos e Servizos de Estatística](#), o [Banco Português de Investimento](#), o [Hotel Miracorgo](#) e o [Proxecto Seivacorgo](#).



O Comité Organizador estivo formado por investigadores de Portugal e de Galicia:

Portugal: Irene Oliveira (CITAB, UTAD) (Presidenta); Argentina Leite (INESC, UTAD); Manuela Gonçalves (CMat, UMinho); Cecília Azevedo (CMat, UMinho); Elisete Correia (CM-UTAD, UTAD); Eva Moraes (CM-UTAD, UTAD); M. Fátima Ferreira (CM-UTAD, UTAD); Sandra Dias (CM-UTAD, UTAD).

Galicia: Esther López (Instituto Galego de Estatística); María Isabel Borrajo (Universidade de Santiago de Com-

postela (USC)); María José Ginzo (USC); Miguel Branco (UVigo); Sofía Debén (USC).

O Comité Científico estivo integrado por oito investigadores de recoñecido prestixio no eido do medio ambiente e da ecoloxía de distintas universidades e presidido por Raquel Menezes (UMinho). O resto de membros foron os seguintes: Jesús Aboal (USC); Manuela Neves (CEAUL, ISA, Universidade Técnica de Lisboa (ULisboa)); Marta Ferreira (CMat, UMinho); María José Lombardía (Universidade da Coruña); Rosa M^a Crujeiras (USC); Russell Alpizar Jara (CIMA-UE, Universidade de Évora); Tiago Marques (CREEM, Universidade de St. Andrews); Wenceslao González (USC).

No programa científico, a maiores das sesións de comunicacións orais e de presentación de pósters, houbo 4 sesións plenarias para as que foron convidados os poñentes: Wenceslao González Manteiga, Modelos de predición semiparamétrica. Aplicación a indicadores medioambientais; Janine Illian (Universidade de St. Andrews), Developing flexible spatial models for the real world - a multi-disciplinary symbiosis; Alexandra M. Schmidt (Universidade Federal do Rio de Janeiro), Modelagem da densidade relativa de peixes numa rede fluvial e Kamil Feridun Turkman (ULisboa), Ventos extremos. Como usar dados observados para calibrar dados de simuladores de vento. Ademais das conferencias plenarias, houbo catro sesións invitadas cos poñentes: Jesús Aboal Viñas, Empleo de herramientas estadísticas durante una década de investigación en biomonitorización ambiental; Rosa M^a Crujeiras Casais, Variogramas suavizados e contrastes de hipóteses en procesos espaciais João Cabral (UTAD), Dando vida às relações estatísticas: Uma Metodologia Estocástico-Dinâmica aplicada à Ecologia e María José Lombardía Cortiña, Predicción de incendios con modelos mixtos de Poisson. Tamén houbo un minicurso sobre o tema Métodos para a estimación da abundancia animal, impartido por Russell Alpizar Jara e Tiago Marques e, xa por último, unha mesa redonda titulada: “A Estatística como ferramenta nas Ciencias Ambientais e na Ecoloxía”, moderada por Rosa M^a Crujeiras Casais e na que participaron Tiago Marques, Juan Picos Martín (UVigo), Isabel Natário (ULisboa) e Miguel Branco. Os resumos das contribucións publicáronse nun libro de resumos que contén as comunicacións orais / póster. Este encontro contou coa participación de 56 persoas, fundamentalmente de Portugal e Galicia, con investigadores sénior e júnior.

O “feedback” recibido por parte dos participantes foi moi positivo, tanto no relativo á organización como á calidade das contribucións seleccionadas para a súa presentación e ademais a raíz do evento, establecéronse colaboracións entre distintos grupos de investigación e entre as dúas sociedades.

Últimas novidades en publicacións e documentos do IGE

Observatorio transfronteirizo Galicia–Norte de Portugal

O Observatorio transfronteirizo Galicia–Norte de Portugal é o resultado das relacións de traballo entre o IGE e a Comissão de Coordenação e Desenvolvimento Regional do Norte, que conta coa colaboración do Instituto Nacional de Estatística de Portugal, no que se refire á selección de datos estatísticos e na preparación da metainformación asociada. Tendo en conta as diferenzas existentes entre os sistemas estatísticos dos respectivos países, tratouse de seleccionar as variables para as cales é posible asegurar a necesaria harmonización. Só a comparabilidade da información permite a lectura conxunta das dúas rexións e, en consecuencia, o afondamento no estudo das relacións transfronteiras. A información do Observatorio está dispoñible na páxina web do IGE na seguinte ligazón: [web](#)

Tríptico Galicia–Norte de Portugal

O IGE publicou en agosto de 2014 o tríptico Galicia–Norte de Portugal, cun resumo de resultados da información máis destacada e nova do Observatorio Galicia–Norte de Portugal. Este tríptico, ademais de contar coa información básica sobre a Eurorrexión, tamén ten información sobre indicadores de competitividade e unha proxección de poboación da zona formada polos concellos galegos e portugueses, que se encontran a unha distancia de 35 km ao norte e a o sur da fronteira que separa aos dous países. Este tríptico está dispoñible na páxina web do IGE na seguinte ligazón: [web](#)

Consello Galego de Estatística

O día 23 de xullo de 2014 tivo lugar a reunión da Comisión Permanente do Consello Galego de Estatística, coa seguinte orde día:

1. Aprobación da acta da sesión anterior
2. Informe de seguimento do Programa estatístico 2013
3. Revisión da variable grao de urbanización
4. Novos proxectos técnicos

5. Rolda de intervencións

Nesta reunión aprobouse a acta da sesión anterior e presentouse o informe de seguimento do Programa estatístico 2013. Tamén se presentou a revisión da variable grao de urbanización co obxectivo de adaptala á revisión metodolóxica proposta por EUROSTAT. Por último, presentáronse novos proxectos técnicos das seguintes operacións estatísticas: Explotación estatística do Rexistro de Parellas de Feito, Táboas de mortalidade, Cooperativas e Sociedades laborais.

Convenios

- Protocolo de colaboración no ámbito do Gabinete de información da Eurorrexión Galicia–Norte de Portugal (GIE–GNP) da Agrupación Europea de Cooperación Territorial Galicia–Norte de Portugal.
- Convenio de colaboración entre o Instituto Nacional de Estatística, o IGE e a Axencia para a modernización tecnolóxica de Galicia, para eliminar duplicidades estatísticas no ámbito do uso das tecnoloxías da información no sector empresas.

Cursos

O IGE en colaboración coa Consellería de Cultura, Educación e Ordenación Universitaria impartirá un curso non presencial no 1º trimestre de 2015, dirixido aos profesores de ESO e Bacharelato da nosa Comunidade Autónoma e co título “Recursos didácticos do portal educativo do Instituto Galego de Estatística”. A data límite para inscribirse é o 12/12/2014. Máis información sobre este curso en: [web](#)

Por outra parte, o IGE ten previsto impartir, a través da Escola Galega de Administración Pública (EGAP), os seguintes cursos ao longo do 2015:

- Cálculos estatísticos básicos co software libre R.
- A poboación como factor clave para a planificación.
- Uso e interpretación de datos para elaborar informes e indicadores de seguimento das políticas públicas.

Traballos de Estatística no SUG

Teses

Estimación non-paramétrica da regresión no contexto de poboacións finitas (pp.ff.)

Carlos L. Iglesias Patiño.
Director: Wenceslao González Manteiga.
Data da defensa: Xullo, 2014. USC.

Trata da regresión non-paramétrica na mostraxe en pp.ff. tradicional. Estúdase Y como función de X nun universo. Para aproximala, considéranse tres estatísticos, os regresograma, cerne e suavizador lineal local (SLL), que constitúen tres parámetros da p.f.. Para estimalos realízase unha sobrevisión seguindo un deseño probabilístico de tamaño fixo non informativo $p(\cdot)$.

Entre os estimadores introducidos destaca unha nova definición do SLL para superar os denominadores nulos. Resulta natural, intuitiva, comunicable a un non-especialista, mais rigorosa e operativa.

Tres métodos diferentes conseguen, sen a linearización, expresar o p -nesgo e a p -varianza como desenvolvemento finito. As expresións resultantes permiten aproximalos considerando un marco asintótico onde X segue un deseño regular fixo dentro dun mesmo segmento, para todos os universos dunha cadea.

Estúdase a orde dos termos distinguindo entre dominantes e resto. Comporta probabilidades de inclusión de orde superior a dúas que conduce a restrinxir o $p(\cdot)$ a mostraxe aleatoria simple sen reposición. Obtéñense expresións asintóticas con restos analíticos evitando ordes en probabilidade.

Posteriormente, admítese que Y se pode expresar, mediante un modelo de regresión de localización e escala, en función de X . Polo tanto, cada estimador serve para estimar o parámetro da p.f. correspondente e a función de regresión.

Estas dúas fontes de aleatoriedade conducen a estudar, con técnicas novedosas, en que condicións os sumandos herdados de $p(\cdot)$ son, en media, de menor orde que os típicos de pp.infinitas. Obtéñense o nesgo, a varianza totais e a varianza media, cun modelo pouco esixente.

Unha metodoloxía heurística orixinal de selección da xanela que aproveita as definicións e resultados dos estimadores e se axeita á mostraxe en pp.ff. permite tratar con mostras pequenas e na depuración.

Aplicase á relación entre a proporción de habitantes en núcleos e a densidade de poboación no universo de parroquias de Galicia. Conxecturábase que era crecente e resultou non-monótona.

Traballos fin de Máster

Aplicacións e extensións do Filtro de Kalman a series temporais.

Iván Gómez Costa.
Directores: Manuel Febrero Bande e Wenceslao González Manteiga.
Data de defensa: Xullo, 2014. USC.

No contexto actual onde a información nos chega de todos lados, é moi habitual que esta información sexa tomada en forma secuencial ao longo do tempo; exemplos disto son os estudos da evolución diaria de activos financeiros, o control diario da temperatura en meteoroloxía, etc. Neste contexto de estudo de series temporais, a representación en espazo de estados é un importante paso para a análise estatística, debido a que ofrece unha notación unificada para unha gran diversidade de

modelos de series temporais, así como facilitar o emprego de técnicas estatísticas como a metodoloxía do Filtro de Kalman, moi útil para a estimación de procesos non observables de forma simple e directa. Nesta liña, faise unha revisión da aplicación da metodoloxía do Filtro de Kalman e o formato en espazo de estados de varios modelos de series temporais, concluíndo con unha aplicación novedosa desta metodoloxía ao modelo en tempo continuo do CKLS con volatilidade estocástica, aplicado ao conxunto de datos correspondente aos tipos de interese do mercado Europeo, EURIBOR.

Optimización da Transferencia de Inventarios. Estudo de caso sobre unha empresa de distribución de moda da Coruña, Galicia, España.

Pablo Montero Souto.
Director: Julio González Díaz.
Data de defensa: Setembro, 2014. USC.

Este proxecto ofrece un novo enfoque para realizar as transferencias de inventarios nunha rede de puntos de venda. Abórdase así o problema de enviar e repoñer a mercadoría mediante unha serie de tres modelos de programación lineal e enteira: (1) un problema de fluxo en redes a custo mínimo, (2) un problema de redondeo de matrices, e (3) un problema de embalaxe. Preséntase tamén unha nova ferramenta para a optimización das transferencias de inventario en R. O emprego desta ferramenta ilústrase cunha aplicación no sector da venda minorista. En concreto, realízase un estudo de caso nunha empresa de moda infantil da cidade da Coruña, Galicia, España.

Optimal trees: Programación en R de problemas de busca de árbores óptimas

Manuel Fontenla Cadavid.
Directoras: Maria Luisa Carpena Rodríguez e Silvia María Lorenzo Freire.
Data de defensa: Setembro, 2014. USC.

Neste traballo abórdanse un conxunto específico de problemas de Investigación Operativa, e máis concretamente do ámbito da Teoría de Grafos, que teñen por obxectivo a determinación de árbores óptimas en grafos con pesos. O principal cometido do traballo foi o desenvolvemento no linguaxe R do paquete Optrees, no cal se implementaron os algoritmos máis coñecidos para resolver os problemas da árbore de custo mínima, da arborescencia de custo mínima, da árbore do camiño máis curto e da árbore do corte minimal. Nalgúns destes pro-

blemas cobran especial relevancia aspectos cooperativos, como aqueles relativos ao reparto dos costes, que levaron á programación dun segundo paquete de R, Cooprees, que implementa as principais regras de reparto para os problemas da árbore e da arborescencia de custo mínima. Para rematar, construíuse a interface web OptreesGUI, baseada en R e na librería Shiny, que permite resolver estes problemas directa e gráficamente dende a páxina optrees.net.

Aplicación de algoritmos heurísticos para optimizar os custos da dobraxe de películas

Alberto Caldas Lima.
Directoras: Silvia María Lorenzo Freire e Maria Luisa Carpena Rodríguez
Data de defensa: Setembro, 2014. UDC.

As películas que van a ser dobradas divídense en pequenos anacos denominados takes, onde un ou máis actores gravan un pequeno conxunto de frases. Durante un día de traballo o número de takes gravados é limitado, e é traballo do director do estudio organizar o proceso de dobraxe para minimizar o custo para o estudio. Ese custo vén determinado polo número de días que un actor é convocado para dobrar a película. Co obxectivo de mellorar o proceso de planificación e abaratar o custo para o estudio, tanto monetario como de tempo de traballo, aplicáronse varios algoritmos heurísticos ao problema. Ademais, construíuse unha aplicación de escritorio que permita aos traballadores do estudio executar os devanditos algoritmos. Empregando esta aplicación poderán obter solucións con custos para o estudio próximos ao mínimo, nun tempo moi reducido.

Predicción de la demanda de electricidad en España: un estudio comparativo

Sonia Rodríguez Hermida.
Director: Germán Aneiros Pérez.
Data de defensa: Setembro, 2014. UDC.

O Traballo Fin de Mestrado céntrase na predicción da demanda horaria de electricidade en España durante o ano 2012 (a base histórica comenza no 2011). Na súa primeira parte, presenta os modelos estatísticos a utilizar: modelos ARIMA e modelos de regresión parcialmente lineal (PL). Tamén, un método naïve que consiste en considerar como predicción o valor histórico que, sendo observado en condicións similares ao que se pretende predecir, esté máis próximo temporalmente a éste. A continuación obtéñense, a través de cada un dos tres

métodos, prediccións para a demanda de electricidade en totadas horas do ano 2012. Nos modelos ARIMA e máis nos modelos PL contémlase a posibilidade de incluír información procedente da temperatura. Un estudio com-

parativo das prediccións mostra similitud entre o comportamento das predicción obtidas a través dos modelos ARIMA e PL, efecto non significativo da temperatura e mal comportamento do predictor naïve.

Os traballos fin de máster presentados nesta sección están dispoñibles na páxina web do máster: [Web](#).

Traballos fin de Grao en Matemáticas na área de Estatística

Os traballos presentados nesta sección foron defendidos ao longo do curso 2013–2014 na facultade de Matemáticas da USC como Traballos Fin de Grao dirixidos dende o Departamento de Estatística e Investigación de Operacións.

Inferencia sobre proporcións

Inés Barbeito Cal.
Director: César A. Sánchez Sellero.

Nos textos básicos de Estatística fórmase e resólvese o problema da inferencia sobre a proporción, para obter intervalos de confianza e para a realización de contrastes de hipóteses, facendo uso da aproximación normal e estimando o erro típico de maneira directa. Sen embargo, este procedemento presenta graves inconvenientes, como erros de cobertura dos intervalos de confianza ou mesmo intervalos non contidos no $[0, 1]$, no cal debe atoparse calquera proporción. Neste traballo revísanse varias alternativas para a construción de intervalos de confianza sobre unha proporción e sobre a diferenza de proporcións poboacionais independentes, comparando as propiedades de cada unha delas.

Exemplos notables de probabilidade e estatística

María Jesús Cobián Martínez.
Director: Manuel Febrero Bande.

Ao igual que calquera rama da ciencia, as Matemáticas tamén describen contrastes do mundo no que vivimos. É natural, pois, que a historia das Matemáticas, revelou moitos paradoxos interesantes algúns dos cales serviron como puntos de partida para grandes cambios. Neste traballo diferenciaremos entre os distintos paradoxos que se nos poden presentar realizando un especial estudo naqueles relacionados co campo da probabilidade e da estatística, paradoxos que aclararon crises da teoría da aleatoriedade, paradoxos que non foron considerados

orixinariamente como tal, etc, e concluiremos facendo un estudo simulado dos paradoxos co software estatístico R.

RANOVA, unha introdución aos modelos multinivel

Marta Cousido Rocha.
Directora: Rosa M. Crujeiras Casais.

As estruturas xerárquicas de datos, as cales xorden cando as unidades elementais se agrupan ou aniñan en unidades maiores, son frecuentes nas ciencias sociais, na medicina ou na bioloxía. Como exemplos, podemos considerar o estudo de individuos clasificados en unidades agrupadas, como rexións xeográficas ou administrativas. Nesta situación, cabe esperar que dous individuos elixidos ao chou no mesmo grupo tendan a ser máis similares que aqueles de grupos diferentes. Neste contexto, resulta de interese analizar as relacións a dous niveis: entre os grupos e dentro dos mesmos. As estruturas xerárquicas son un exemplo de estrutura complexa, e a análise tanto descritiva como inferencial de poboacións con estruturas complexas é o obxectivo dos modelos multinivel. O modelo multinivel máis sinxelo é o modelo de análise da varianza con efectos aleatorios (RANOVA, do inglés Random Effects Analysis of Variance), que será o obxecto de estudo deste traballo.

Modelos autorregresivos de series de tempo

Miguel Fernández Castiñeira.
Directora: Rosa M. Crujeiras Casais.

Os modelos autorregresivos permiten modelar o comportamento dunha serie de observacións recollidas ao longo do tempo e que presentan unha dependencia feble entre elas. Neste traballo centrarémonos no estudo da modelización de tipo Box-Jenkins para series autorregresivas. No traballo preséntase inicialmente unha introdución ás series de tempo e aos modelos autorregresivos, introducindo as funcións de autocorrelación e de densidade espectral, abordando posteriormente as tarefas de estimación e diagnose e presentando unha introdución das técnicas anteriores ao caso de modelos autorregresivos multivariantes.

Verosimilitude empírica.

Andrea Meilán Vila.
Director: César A. Sánchez Sellero.

A verosimilitude empírica é un método que permite realizar tarefas de inferencia, adaptándose de maneira moi flexible á distribución do estimador ou do estatístico de contraste. Ademais, permite a aplicación dunha distribución asintótica chi-cadrado moi sinxela, o cal facilita as labores de inferencia. Neste traballo revisáronse os elementos básicos do método de verosimilitude empírica, o cal implica a utilización de multiplicadores de Lagrange. Ilustráronse no traballo a inferencia por verosimilitude empírica con parámetros vectoriais. Os métodos ilustráronse con exemplos simulados, onde se poderon ver as vantaxes da verosimilitude empírica en distribucións asimétricas.

Funcións xeneratrices e aplicacións ao cálculo en Teoría de Xogos

Guido Ignacio Novoa Flores.
Directora: Balbina V. Casas Méndez.

Unha función xeradora ou función xeneratriz é unha serie formal de potencias cuxos coeficientes codifican información sobre unha sucesión a_n cuxo índice percorre os enteiros non negativos. Esta técnica da análise combinatoria vén mostrando a súa utilidade para o cálculo de solucións na teoría de xogos cooperativos desde que David G. Cantor en 1962 a utilizou para obter o denominado valor de Shapley. Neste traballo preténdese revisar a literatura máis relevante nesta liña de traballo, desde a contribución de David G. Cantor ata a actualidade. Isto inclúe a presentación de diferentes algoritmos, o estudo da súa complexidade e algunhas das súas aplicacións ás ciencias sociais. As habilidades informáticas (programación en R) poden enriquecer de xeito significativo o traballo. Por outra banda, existe a posibilidade de ini-

ciarse á investigación comezando o estudo e proposta de posibles solucións a problemas abertos neste campo.

Contando eventos no tempo: o proceso de Poisson

Laura Pérez Vilarelle.
Directora: Rosa M. Crujeiras Casais.

Este traballo céntrase no estudo dos procesos de Poisson en tempo, con especial atención á súa caracterización. Comézase introducindo as distribucións de Poisson e exponencial e analizando as súas propiedades para despois caracterizar os procesos de Poisson. Enuméranse algúns resultados relevantes destes procesos e coméntase brevemente a súa extensión a situacións non homoxéneas.

Modelos de regresión con efectos aleatorios

Gabriel Rodríguez González.
Director: Wenceslao González Manteiga.

Trátase de revisar e describir os modelos de regresión con efectos aleatorios con resposta cualitativa e continua que xurdiron nos últimos anos. En segunda instancia, a partir dunha base de datos de incendios forestais, ilustrar mediante macros no software estatístico R parte da metodoloxía anteriormente descrita, coa conseguinte explicación dos resultados acadados.

Introdución á elección social. O teorema de Arrow

Manuel Vila Tosar.
Directora: Balbina V. Casas Méndez.

Este traballo consta dunha parte inicial de introdución á Teoría da Decisión relativa ao estudo de funcións de preferencia e á representación por medio de funcións de utilidade. En segundo lugar, efectuarase unha introdución á Teoría da Elección Social: esta disciplina ocúpase do estudo dos procedementos e institucións que permiten a grupos de individuos, pertencentes por exemplo a unha determinada organización política ou social, realizar unha toma de decisións baseándose nas preferencias de todos os membros do grupo. O tema central do que se ocupa a Teoría da Elección Social é o de elaborar regras de votación. Ao comezar os estudos neste campo, Arrow probou o seu coñecido “teorema de imposibilidade”.

Este establecía que ao impoñérense unhas poucas condicións, totalmente sinxelas e naturais, as únicas regras que as cumpren son as ditatoriais. Dende os traballos de Arrow, a Teoría da Elección Social experimentou un gran desenvolvemento. Cabe mencionar os denominados resultados de posibilidade e as regras de decisión colectivas.

Inferencia estatística con datos circulares

Carla Díaz Louzao.
Director: Alberto Rodríguez Casal.

Cando medimos a dirección do vento estamos rexistrando un dato circular. Esta dirección pode representarse como un punto na circunferencia. Outros exemplos poden ser a dirección do polo magnético, a hora en que empeoran os síntomas dunha determinada enfermidade ou a dirección en que rompen os implantes de cadeira. Dada a súa natureza circular, estes datos non teñen escala e son periódicos, polo que as técnicas clásicas utilizadas para modelar e analizar a súa distribución deben ser convenientemente adaptadas. O obxectivo deste traballo é revisar as técnicas estatísticas existentes para analizar datos circulares. Realízase unha breve introdución sobre as técnicas de tipo descritivo (tanto numéricas como gráficas), para posteriormente estudar os principais modelos paramétricos de datos circulares. Finalmente, abórdase o problema da estimación nonparamétrica da densidade con datos circulares.

Resolución de problemas de programación lineal mediante o algoritmo de punto interior de Karmarkar

Alejandro Blanco Romón.
Director: Julio González Díaz.

Neste traballo de fin de grao introdúcense en primeiro lugar algunhas nocións de complexidade computacional e clasificación de algoritmos. Máis adiante abórdase o tema da complexidade exponencial do simplex, repasando a clase de problemas do hipercubo de Klee-Minty. A continuación detállase paso a paso o algoritmo de punto interior de Karmarkar para a resolución de problemas de programación lineal. Tamén se explican a rutina de redondeo, o procedemento que permite obter a solución exacta a partir dunha solución óptima obtida mediante Karmarkar, e a transformación que poden requirir os problemas de programación lineal para ser resolto con Karmarkar. Finalmente engádense as demostracións que garanten a converxencia e a orde do algoritmo, e a súa programación en R.

Modelos de regresión cuantil

Marta Santos Romero.
Director: César A. Sánchez Sellero.

Os modelos de regresión cuantil describen o comportamento de certo cuantil da variable resposta condicionado ás variables explicativas. Estes modelos gozan de boas propiedades de robustez e flexibilidade fronte aos modelos clásicos, baseados na esperanza condicionada e o método de mínimos cadrados. Neste traballo revísanse os modelos de regresión cuantil e efectúase unha avaliación comparativa das súas propiedades mediante datos reais e simulados.

Coñecéndonos

Coñecendo a unha socia da SGAPEIO: M^a José Ginzo Villamayor.

M^a José Ginzo Villamayor é licenciada en Matemáticas pola USC. Realizou un Posgrao en Estatística da Universidade de Porto e o Máster interuniversitario en Técnicas Estatísticas na USC. Actualmente traballa no Departamento de Estatística e Investigación Operativa da USC, dende 2008. Dende o 2013 é vogal do consello executivo da SGAPEIO.



Cantos anos levas traballando en estatística ou IO?

Dende outubro do 2006, fun bolseira no Instituto Galego de Estatística (IGE). Antes traballei no sector TIC, como programadora e analista de aplicacións de software con IBM para Caixanova e no CIXTEC.

Descríbenos o teu traballo

Traballo no departamento de Estatística e I.O. da USC, no grupo Modelos de optimización, decisión, estatística e aplicacións (Modestya). Son técnico de consultoría no servizo de Consultoría Estatística do grupo. Prestamos asesoramento estatístico de calquera nivel, dende análise de datos para traballos fin de grao/máster, teses, ata desenvolver proxectos que requiren técnicas máis sofisticadas e tamén a empresas, isto permíteme participar en proxectos no sector industrial e produtivo. Agora participo nun liderado por INAER relacionado coa loita contra o lume. E colaboro na xestión do Máster en Técnicas Estatísticas.

Dende cando es socio/a da SGAPEIO?

Dende o 2010 e membro da actual executiva dende o 2013. Coñezo a SGAPEIO dende a miña época de estudante na facultade, nas clases nos teñen falado da sociedade e polos actos que organizaba.

Cal foi o 1º congreso da SGAPEIO ao que asistiches?

Ao VIII en Santiago. E logo fun aos 3 seguintes en Ourense, Pontevedra e Coruña. Formo parte da organización do XII Congreso da SGAPEIO (Lugo, outubro 2015).

O tema do teu próximo proxecto/artigo?

Estou traballando en varias cousas, dende medidas de biodiversidade aplicadas ás aves de Galicia, resultados preliminares presentáronse no I Encontro Luso-Galaico de Estatística no Medio Ambiente e na Ecoloxía (Vila Real-Portugal, 6-8 de novembro), ata temas relacionados coa miña tese en Xeolingüística.

Un curso que che gustaría facer?

Moitos, un que me interesa moito é sobre demografía.

Que libro de estatística ou I.O. recomendarías?

Esta é unha pregunta complicada, depende de quen sexa o lector e cal sexa o obxectivo que pretenda.

Se non foses estatístico/a, que che gustaría ser?

Non sei, esta foi a profesión que elixín, para a que me formei, continúo formándome, e por sorte podo dedicarme. En calquera caso, seguiría sendo matemática.

Se oes “estatística”, que é o primeiro que se che vén á cabeza?

Depende, se a escoito en foros da universidade venme o meu traballo e compañeiros. Se o que escoito é paro, IPC, ... venme a época no IGE, entre outros.

Que é o que máis valoras da labor da SGAPEIO?

A SGAPEIO, ademais da difusión que leva a cabo de forma divulgativa acercando a nosa labor á sociedade en xeral, é un punto de encontro entre os que nos dedicamos á Estatística e á I.O., dinamizando o colectivo de estatísticos, acadando unha Sociedade Viva.

En que cres que deberíamos mellorar?

Acercando máis a Estatística, o que nós sabemos facer, a distintos sectores da sociedade: estudantes, docentes, empresas, cidadáns en xeral. É necesario que saiban ao que nos dedicamos. Transmitir a importancia que a Estatística e a I.O. supoñen para o progreso da sociedade, incrementar as relacións coa Administración e integrar a universitarios usuarios da Estatística e a I.O. e empresarios.



Sabías que?

A Estatística no fútbol. por Manuel Febrero Bande

Podemos predicir con moita exactitude a cantidade de puntos acadados por un equipo de fútbol no final de liga a partires simplemente da diferenza de goles. O curioso do asunto é que o tipo de relación destas dúas variables é basicamente lineal con coeficientes de determinación por riba do 90%. Así, se consideramos as variables X ="Diferenza promedio de goles por partido xogado" e Y ="Promedio de Puntos obtidos por partidos xogados" na clasificación acumulada dos últimos 25 anos na 1ª división, obtense a relación lineal $Y = 1,3657 + 0,6294X$ cun coeficiente de determinación de 0,9591. Relacións moi semellantes téñense para as outras ligas europeas (Inglaterra: $R^2 = 0,9368$, Italia: $R^2 = 0,9527$, Alemaña: $R^2 = 0,906$). Esta relación fortemente lineal mantense asemade tempada a tempada. A tempada pasada na liga española a relación entre as dúas variables mencionadas foi: $Y = 1,3868 + 0,5578X$ con un $R^2 = 0,9516$. Polo tanto, este exemplo de fútbol sería bo para unha clase adicada á recta de regresión onde, por outra banda, non abundan os exemplos estatísticos. Como se denuncia no libro *Soccernomics*, a aplicación de métodos estatísticos no fútbol, tanto en xestión coma en análise de xogo, vai

moi por detrás do desenvolvemento destes procedementos nos outros grandes deportes profesionais: baloncesto, fútbol americano ou béisbol. Aínda así, neste libro recóllese algunha aplicación curiosa como a da aplicación de métodos estatísticos e do equilibrio de Nash ao lanzamento de tandas de penalties que fixo o catedrático de economía da London School of Economics Natxo Palacios Huerta e que foi aplicada exitosamente na final da liga de campións de 2008 entre o Chelsea e o Manchester United. Este exemplo tamén se conta nun recente programa de Informe Robinson ([Web](#)) adicado ao tema onde se descubre que este profesor asesoraba á selección nacional de Holanda durante o mundial de Sudáfrica en previsión de que na final do campionato se chegase á tanda de penalties. Ufff, menos mal que Iniesta se encargou de que non tiveramos que xogar tamén en contra da Estatística.

Soccernomics. Why England loses, why Spain, Germany, and Brazil win, and why the US, Japan, Australia –and even Iraq– are destined to become the kings of the world's most popular sport. (2014) Simon Kuper & Stefan Szymanski. Nation Books, New York.

Dirección:

Paula Raña Míguez

Equipo editorial:

Esther Calvo Ocampo, M^a Jesús Casado Barrio, Balbina Casas Méndez, Tomás Cotos Yáñez, Rosa Crujeiras Casais, M^a Esther López Vizcaíno, Silvia M^a Lorenzo Freire, María Loureiro García, Javier Roca Pardiñas, Rebeca Villaverde López.

Coa colaboración de:

Soly Santiago Pérez,
M^a José Ginzo Villamayor,
Manuel Febrero Bande.

Comunicación coa SGAPEIO:

Facultade de Matemáticas, Campus Universitario Sur,
15706-Santiago de Compostela, A Coruña.

www.sgapeio.es
secretaria@sgapeio.es
boletin@sgapeio.es
Twitter @sgapeio
LinkedIn SGAPEIO

COLABORA !!!

Informest é unha publicación cuatrimestral da SGAPEIO que podes atopar na web da sociedade, www.sgapeio.es, no apartado PUBLICACIÓNs.

Se queres participar cun artigo de divulgación para a sección "Colaboracións" ou cunha curiosidade estatística para o "Sabías que?", contacta con nós no correo "boletin@sgapeio.es" onde te informaremos das normas de edición.

Anímate!!



SGAPEIO

Sociedade Galega para a
Promoción da Estatística
e da Investigación de
Operacións

Depósito Legal: LU-191-1995 - I.S.S.N.:1695-7083