

Colaboración

Recensión bibliográfica

Actividades da SGAPEIO

Novas do IGE

Traballos de Estatística

Coñecéndonos

Sabías que?

Editorial

SILVIA M^A LORENZO FREIRE – VICEPRESIDENTA DE INVESTIGACIÓN DE OPERACIÓNS

Estimados socios e socias:

É para min un pracer poder dirixirme por primeira vez a todos vós con motivo da edición deste novo número de Informest. Nestes últimos meses, a nosa sociedade organizou diferentes actividades entre as que cabe destacar a celebración do I Encontro Luso-Galaico de Estatística no Medio Ambiente e na Ecoloxía (celebrada en colaboración coa Sociedade Portuguesa de Estatística), o curso “Introdución a R: Software estatístico libre para PEMES” e as dúas xornadas de formación para elaborar proxectos de Estatística (levadas a cabo en colaboración coa Asociación Galega de Profesorado de Educación Matemática e a Fundación Barrié). Como xa saberedes, este mes de Xaneiro tamén tivo lugar a Asemblea Xeral da nosa sociedade, antes da cal se celebrou unha xornada co obxectivo de mostrar posibles aplicacións da Estatística e a Investigación Operativa.

Ao longo deste ano desenvolveranse novas actividades. Unha destas actividades é o XII Congreso Galego de Estatística e Investigación de Operacións, organizado polo Departamento de Estatística e Investigación Operativa da Universidade de Santiago de Compostela, que terá lugar no Pazo de San Marcos, en Lugo, do 22 ao 24 de outubro. A páxina web do congreso é xiicongreso.sgapeio.es.



Tamén haberá unha nova edición do “Concurso Incubadora de Sondaxes e Experimentos”, que se vén celebrando con periodicidade ao longo dos últimos cinco anos. O período de inscrición para o mesmo estará aberto do 1 ao 15 de abril. Toda a información sobre o concurso está dispoñible no seguinte enderezo web.



Outra futura actividade que temos previsto levar a cabo é unha xornada sobre Big Data e Estatística. Esta xornada, organizada en colaboración co grupo MODES e a Rede de Tecnoloxías Cloud e Big Data para HPC, terá lugar o 19 de xuño na Facultade de Informática da Universidade da Coruña. A inscrición xa está aberta na páxina web bigdata.sgapeio.es

Destas e doutras actividades iremos informando a través da páxina web da SGAPEIO, www.sgapeio.es, e nas próximas edicións de Informest.

Por último, gustárame agradecer a colaboración de todos os membros que participan nas actividades que organiza esta sociedade xa que sen a súa participación todo isto non sería posible e, en especial, a aqueles que colaboraron nesta e noutras edicións deste boletín.

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



Cartografando a incertidume e facendo predicións en tempos desafiantes.

por Juan Picos

Juan Picos é Enxeñeiro de Montes pola Universidade Politécnica de Madrid e Doutor pola Universidade de Vigo. Actualmente é Profesor contratado doutor do Departamento de Enxeñería dos Recursos Naturais e Medio Ambiente da Universidade de Vigo na Área de Enxeñería Agroforestal. Foi Director da Asociación Galega Monte Industria e actualmente é membro do Team of Specialists on Forest Sector Outlook, da Comisión Económica das Nacións Unidas para Europa (UNECE) e a Organización para Alimentación e a Agricultura das Nacións Unidas (FAO), vocal da Xunta Directiva da Sociedade Española de Ciencias forestais e decano autonómico do Colexio de Enxeñeiros de Montes.

O pasado mes de novembro fun convidado pola Sociedade Galega para a Promoción da Estatística e da Investigación de Operacións (SGAPEIO) a aportar, no marco do I Encontro Luso-Galaico de Estatística no Medio Ambiente e na Ecoloxía (EES2014), a miña visión sobre o papel que estas disciplinas están a desenvolver neste eido. Centrei a miña intervención, sobre a que baseo este texto, na miña visión sobre a xeración, análise e emprego de Estatísticas Públicas referidas aos recursos naturais e o seu efecto na toma de decisións para políticas públicas e estratexias empresariais.

Vivimos tempos desafiantes. A aparente estabilidade da segunda metade do século XX deu paso a unha sensación de vertixe. Nun planeta globalizado, no que a información flúe dunha punta a outra á velocidade da luz, as fontes de incerteza multiplícanse e as autoridades vense enfrontadas a numerosos desafíos, pois teñen que ser quen de conxugar a credibilidade das súas accións coa flexibilidade necesaria para superar situacións imprevistas. Nun mundo global non hai lugar para regras estritas e concepcións ríxidas, como tampouco para a improvisación e a incapacidade de reacción. Nun ambiente económico que cambalea entre recesión e crecemento, os negocios afrontan volatilidade aumentada de exixencia. As previsións e plans pasaron de anuais a trimestrais, logo, a mensuais e, despois, a semanais.

Ao mesmo tempo, o cambio climático proporciona evidencia da variabilidade inherente do clima da Terra e a súa resposta á influencia humana. As implicacións

do cambio climático para o ambiente e a sociedade dependerán da resposta do sistema terrestre pero, tamén, de cómo a Humanidade afronte dito reto con cambios tecnolóxicos, económicos, no estilo de vida e nas políticas.

Os recursos naturais forman o eixo vertebrador de moita da actividade económica mundial e dos modos de vida da práctica totalidade da nosa sociedade. Ademais os espazos menos alterados conteñen a maior parte dos lugares de importancia cultural e histórica. Os espazos naturais son tamén uns que requiren a xestión máis clara, participada, aínda que, á vez, altamente a máis complexa e incerta.



Durante os últimos anos véñense potenciando un proceso de participación pública, un proceso de identificación e incorporación das preocupacións, necesidades e valores dos distintos axentes na toma de decisións. Por exemplo, O Convenio de Aarhus (1998) sobre acceso á información, participación pública na toma de decisións e acceso á xustiza en materia de medio ambiente, así como a normativa comunitaria derivada do mesmo, tentan implantar o concepto da Administración pública aberta e transparente no eido medioambiental. O público goza así do dereito a acceder á información ambiental que as autoridades públicas posúen, e obrígase á difusión de ampla información ambiental, como é, por exemplo, información sobre a lexislación, sobre o estado do medio ambiente, sobre proxectos, plans e programas ou sobre decisións que se adopten que poidan afectar ao medio ambiente.

Cada vez máis cidadáns son usuarios da información estatística, pero, ao mesmo tempo, cada vez máis cidadáns son “produtores” de información con potencial valor estatístico.

Tendo en conta esta situación, cales son os desafíos no campo da estatística relativa ao Medio Ambiente?

A primeira barreira a superar é entender qué tería que ser medido e prognosticado e con qué precisión e periodicidade. Ás veces o que necesita ser prognosticado non son os procesos, senón os “drivers” dos mesmos. Outras veces a solución máis eficiente pasa por deslindar o que non se sabe. A mellora non pasa sempre por aumentar a precisión do que se mide, senón ter unha medida da incerteza de dita medición e das consecuencias que a mesma implica.

A segunda barreira a superar é aceptar que prognosticando algúns comportamentos isolados, aínda que sexa cunha alta precisión, non se consegue nada, cando ese proceso é susceptible de vir influenciado por outros dos que non temos información.

A terceira barreira a superar é o determinismo de crer que un bo pronóstico será sempre exacto. É imprescindible poder valorar diversos escenarios e ser quen de mudar as hipóteses de partida e, rápido, ver o efecto daqueles cambios no modelo. Á parte de mellorar a adaptabilidade, isto contribuirá a construír consenso ao redor das decisións a tomar ou plans a levar a cabo.

A cuarta barreira refírese aos novos paradigmas na información susceptible de ser tratada estatisticamente. Xa vivimos nun mundo no que os sensores fornecen acotío grandes datasets, estamos nos tempos da Big Data. Porén, moita desta información é numerosa pero incompleta, a facilidade con que se adquiren datos nuns lugares contrasta coa inexistencia de datos noutros.

Finalmente, se as tendencias non son ben comunica-

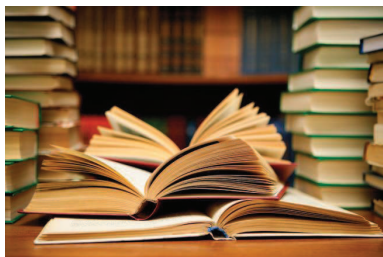
das, non haberá ningún diálogo, ningún compromiso e ningunha acción.

As axencias gobernamentais normalmente comunican as estatísticas oficiais ao público sen facer practicamente ningunha mención á incerteza das estimacións incluídas nelas. Obviamente hai publicacións técnicas que documentan a orixe dos datos e os métodos empregados que proporcionan información sobre os erros de mostraxe, pero poucas veces se aborda a existencia de erros alleos ás mostraxes. Débese facer un maior esforzo por manifestar de xeito explícito as fortalezas e limitacións dos datos fornecidos, tomando tanto coidado para entender e explicar o que se sabe como o que non se sabe. Que as estatísticas oficiais só se transmitan como estimacións puntuais, sen unha adecuada referencia aos erros, proxecta una falsa sensación de certeza. Tamén é certo que os policy makers e, incluso, o público aparentemente desexan recibir información que transmita certeza aínda que dita certeza posúa unha total falta de credibilidade.

Deste xeito, os usuarios desta información (cada vez máis numerosos, e, cada vez, máis fonte de prescripción) acaban asumindo inocentemente que os erros son pequenos e sen consecuencias.

Tamén é certo que temos moi pouco coñecemento do motivo e o xeito nos que os usuarios das estatísticas as interpretan. Algúns poden erroneamente tomalas ao pé da letra. Non obstante, outros poden asumir a existencia dos erros máis variados en magnitude ou dirección. É necesario afondar en como os resultados estatísticos son empregados e reelaborados na toma de decisións. Vivimos tempos incertos con necesidades certas. Anticipar, impulsar, coordinar e facilitar a participación, son os desafíos aos que a Estatística, como disciplina, terá que contribuír. Teño o convencemento de que unicamente cunhas axeitadas Estatísticas Públicas e unha axeitada explotación das mesmas é posible desenvolver políticas no que atinxe ao Medio Natural e aos aproveitamentos/valores do mesmo.





Betanzos magic town

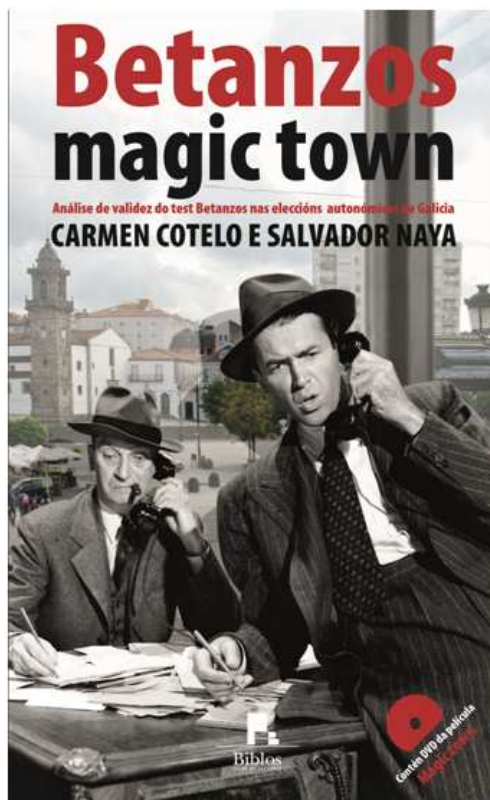
por Guido I. Novoa Flores e Balbina V. Casas Méndez.

Betanzos magic town, de Carmen Coteló e Salvador Naya: unha recensión

Magic town (Cidade máxica) foi unha película de 1947, dirixida por William Wellman e protagonizada por James Stewart e Jane Wyman. Trata dunha empresa dedicada ás sondaxes de opinión que busca con entusiasmo unha especie de “milagre estatístico”: que a opinión dunha pequena cidade reflicta a de todo o país. Cando as cousas se complican e a empresa parece estar abocada á bancarrota, publícanse os resultados dunha enquisa realizada a milleiros de persoas en todo o país. Sorprendentemente, eses resultados coinciden cos da sondaxe realizada pola empresa na pequena vila de Grandview.

O resumo desta película pódenos achegar ao libro “Betanzos magic town”, escrito pola xornalista e funcionaria da Xunta de Galicia, Carmen Coteló e polo matemático e profesor de estatística da Universidade da Coruña, Salvador Naya. Está prologado polo político e politólogo Xosé Luís Barreiro Rivas, quen afirma que se trata “dunha achega necesaria para os estudos de márketing e comunicación política”. Publicouse en 2013, que foi declarado Ano Internacional da Estatística.

Naturalmente, o equivalente de Grandview, no libro de Coteló e Naya, é a cidade de Betanzos e remóntanos aos tempos, non tan afastados, cando Manuel Fraga, o político ex-presidente da Xunta de Galicia, pedía os resultados de Betanzos, xunto cos de Vilalba (a súa vila natal) e Miño (municipio no que dende sempre pasaba o tempo de descanso) ao terminar un proceso electoral. Pero, por que tiña Fraga tanto interese nos resultados de Betanzos?...



O libro ten unha primeira parte na que se fai un percorrido ata a actualidade de autores e estudos, achega do que hoxe coñecemos por magic towns, desde os estudos iniciais dos sociólogos norteamericanos Robert Lynd Staughton e Helen Lynd Merrell, profesores da Columbia University e o Sarah Lawrence College, respectivamente. Eles publicaron en 1929 o primeiro traballo sobre este tema. Nese traballo acuñouse o nome de middletown, é dicir, cidade media, en referencia a Muncie (Indiana), que é a cidade na que focalizaron os seus primeiros estudos, e que no libro de Carmen e Salvador se describe económica e socialmente con moitísimo detalle. Pódese mencionar que un orixinal dun dos libros dos Lynd, sobre esta temática, atópase na Biblioteca América da Universidade de Santiago de Compostela.

Nesta parte, revísanse as técnicas de enquisas aplicadas desde as eleccións en España de 1977 ata hoxe; fála-se da empresa demoscópica do matemático e estatístico estadounidense George Gallup, que ademais exercía o xornalismo, ou das primeiras longas noites electorais e



Salvador Naya na presentación do libro.

os avances electorais e técnicos que permitiron acurtalas. Atopamos middletowns alternativas, chegamos ao chamado teorema do votante medio, proposto polo matemático e politólogo francés N. de Condorcet, e que se asocia a senllos artigos de H. Hotelling e D. Black, e de aí aos chamados “home de Essex” ou “mamá do hockey sobre xeo”, entre outros, e a súa relación cos resultados de determinadas eleccións.



Carmen Coteló, coautora do libro.

A segunda parte ocúpase do caso de Betanzos e da demostración estatística de que os resultados das eleccións autonómicas en Betanzos son similares aos obtidos ao nivel de toda a comunidade galega. Para iso, fanse uso de medidas de correlación e de distancia e execútanse contrastes de hipóteses. Fálanos do denominado “test Betanzos”, alusivo ao interese que xa se mencionou de Manuel Fraga acerca dos resultados electorais de Betanzos. Introdúcese a relación existente entre a estatística e as eleccións e abórdase a resposta a diferentes cuestións

como o problema da estimación dos votos mediante enquisas ou os motivos polos cales as enquisas electorais poden errar. Preséntanse as denominadas forcadadas electorais e outras estimacións dos erros ou cómo elixir o tamaño da mostra.

Todo o anterior vai dando resposta á gran cuestión do libro respecto de Betanzos, e como apoio faise uso dos resultados nas eleccións autonómicas de Galicia en 1997, 2001, 2005 e 2009. Para explicar as características que fan de Betanzos unha middletown, máis especificamente, que algunha das súas seccións censuais son “mesas test” sostidas no tempo, realízase unha aproximación á súa realidade social, facendo uso de datos oficiais ofrecidos polos Institutos Galego e Nacional de Estatística (IGE e INE).

A última parte establece e estuda o paralelismo entre Betanzos e Muncie, a middletown de Robert e Hellen Lynd, e para iso resulta de grande importancia a ampla descrición demográfica, social e económica de Betanzos que se fai.

O Anexo que se ofrece revela a ampla documentación considerada para a preparación do libro. Inclúe unha enumeración de fontes e recursos de información sobre magic towns e “mesas test”, artigos en revistas especializadas, libros, artigos xornalísticos e noticias en medios de comunicación; describe o soporte audiovisual que inclúe a película Magic town en DVD e, finalmente, ofrece bibliografía estatística.

Indubidablemente, o tema tratado neste libro é moi suxestivo, está tratado de forma rigorosa e amplamente documentado, presentado de forma moi amena e por todo iso, na nosa opinión, é de recomendable lectura para un amplo sector de lectores, en particular para estudantes e profesionais da estatística e das ciencias sociais.



Praza dos Irmáns García Naveira, en Betanzos.



Actividades da SGAPEIO

Introdución a R: software estatístico libre para PEMES.

M^a José Ginzo e Rebeca Villaverde

A Sociedade Galega para a Promoción da Estatística e a Investigación de Operacións (SGAPEIO) en colaboración coa Facultade de Matemáticas, organizou o curso "INTRODUCCIÓN A R: SOFTWARE ESTATÍSTICO LIBRE PARA PEMES". O curso foi impartido por María José Ginzo da USC e Rebeca Villaverde de Anova Multiconsulting.

Este curso entra dentro do obxectivo xeral da SGAPEIO de conseguir, dende a Sociedade, un maior acercamento entre a universidade e as empresas galegas. Este obxectivo xeral concretouse, neste curso, en ofrecer formación en materia estatística ás empresas galegas que realizan estudos de mercado, tratamento de datos de enquisas, análise de datos do propio funcionamento da empresa,... Centrouse principalmente naquelas ferramentas que necesitan no día a día. Para facilitar a asistencia, fíxose na modalidade semipresencial. O curso contou coa asistencia de 11 alumnos. Aínda que, a priori, poidan parecer poucos, hai que ter en conta que é a primeira vez que se fai un curso destas caracterís-

ticas dende SGAPEIO e que, ademais, é a primeira vez que a SGAPEIO se achega ao mundo da empresa. Por iso valoramos esta asistencia positivamente, aínda que consideramos que debemos melloralala en futuras edicións.

Na última sesión pediuse ao alumnado que realizase unha serie de valoracións cualitativas en relacións aos obxectivos, motivos para participar, información, poñentes, contidos, tarefas levadas a cabo, apuntamentos e outros intereses do curso así como unha valoración cuantitativa sobre a satisfacción xeral do curso.

O grao de satisfacción medio do curso foi de 7.

Tras os bos resultados acadados nas valoracións do curso e a boa acollida entre o alumnado asistente, é probable que se celebren vindeiras edicións do mesmo e doutra formación relacionada con fundamentos teóricos estatísticos.

Dende SGAPEIO agradecemos á Facultade de Matemáticas a súa colaboración na organización deste curso; ás relatoras e a todos os alumnos, o seu interese e esforzo.

A continuación móstranse os resultados principais das valoracións.

Avalía segundo a túa opinión as metas e obxectivos deste curso en función de:

	Moi alta	Alta	Media	Baixa	Moi baixa
Relevancia para o meu traballo	11.11 %	44.44 %	44.44 %	0.00 %	0.00 %
Realismo e practicidade	0.00 %	44.44 %	55.56 %	0.00 %	0.00 %
Claridade, estruturación dos obxectivos	0.00 %	55.56 %	44.44 %	0.00 %	0.00 %

Estou interesado en adquirir máis coñecementos en:

Estatística descritiva	23.08 %
Estatística avanzada	46.15 %
Linguaxe de programación R	23.08 %
Outra linguaxe de programación ou software estatístico	7.69 %
Outros	0.00 %

Asemblea Xeral da SGAPEIO 2014

Balbina Casas Méndez

O día 29 de xaneiro de 2015 tivo lugar, na Facultade de Matemáticas de Santiago de Compostela, a Asemblea Xeral da SGAPEIO do ano 2014. Antes de comezar, impartíronse dúas conferencias: “Solucións estatísticas para retos actuais en enxeñaría e ciencia de materiais” por Javier Tarrío, profesor da Universidade da Coruña, e “Votando secuencialmente” por Julio González, investigador da Universidade de Santiago de Compostela.



Os conferenciantes: Javier Tarrío e Julio González.

Tras aprobar a acta da Asemblea anterior e informar do número de socios, que actualmente é de 180 individuais e 7 institucionais, realizouse unha recompilación das principais actividades de SGAPEIO en 2014.

Unha delas foi o IV Concurso “Incubadora de Sondaxes e Experimentos” que contou con 36 equipos participantes. Fíxose tamén unha Xornada de Formación a profesores de Ensino Secundario en colaboración con AGAPEMA en novembro. Houbo 90 inscritos. Constou dun taller de *epidat* e outro de *geogebra*, e senllas conferencias de M^a Jesús Casado e Ignacio García Jurado.

En 2014 saíron dous números de INFORMEST e a súa directora, Paula Raña, xa ten planificado o contido dos tres números previstos para 2015. Tamén en novembro, María José Ginzo e Rebeca Villaverde, impartiron un curso de R para PEMES. Contou, este curso, con 11 asistentes e as enquisas realizadas revelaron un alto grao de satisfacción. Para finalizar esta recompilación, informouse da colaboración de SGAPEIO coa Sociedade Portuguesa de Estatística (SPE) na organización do I Encontro Luso-Galaico de Estatística no Medio Ambiente e na Ecoloxía, celebrado en novembro en Vila Real.

En canto ás actividades futuras, a Presidenta de SGAPEIO informa de que se están a dar os primeiros pasos para a realización do II Congreso Luso-Galaico de Biometría, a finais do curso 2015-16 en Galicia, co-organizado coa SPE. En xuño de 2015, celebrárase na Facultade de Informática da Coruña unha Xornada sobre Big Data, en colaboración co grupo de investigación MODES da Universidade da Coruña. Polo que se refire ás axudas económicas para 2015, obtívose unha do Ministerio de Economía de 7.900 euros e outra da Fundación Barrié de 3.000, para a realización do V Concurso “Incubadora” e para unha nova Xornada Formativa con AGAPEMA.

Outros asuntos dos que se informou foron a nova conta bancaria da Sociedade para facilitar a xestión económica do XII Congreso SGAPEIO e a asistencia de Soly Santiago á reunión do Consello Galego de Estatística en decembro de 2014, na que se trataron asuntos de trámite.

Aprobouse o balance económico de 2013 e 2014 e o orzamento do ano 2015. Informouse de que se realizou xa a convocatoria e a difusión do V Concurso “Incubadora”.

Finalmente, Luís Ramil falou do XII Congreso Galego de Estatística e Investigación de Operacións, que se celebrará no Pazo de San Marcos de Lugo os días 22, 23 e 24 de outubro de 2015. Constituíronse os comités organizador e científico e creouse a páxina web para ir incorporando toda a información, <http://xiicongreso.sgapeio.es>.

Haberá catro conferencias plenarias: Philippe Vieu, da Universidade Paul Sabatier (Toulouse), especialista en métodos estadísticos non paramétricos para a análise de datos funcionais; Mónica Bécue, da Universidade Politécnica de Cataluña, especialista na análise estatística de datos textuais; Jorge López Muñiz, responsable da área de Intelixencia de Negocio da empresa Hijos de Rivera, S.A. e Julio González, investigador na Universidade de Santiago de Compostela, con experiencia en transferencia científica con empresas do sector da enerxía.

Están programadas dúas mesas redondas, unha delas sobre os servizos de consultoría estatística, e tamén catro obradoiros, dous deles complementarios ás conferencias de Philip Vieu e Mónica Bécue, respectivamente. Un, sobre Control de Calidade e outro estará destinado a profesores de Ensino Secundario. Prevense tamén actividades sociais e culturais.

A acta da Asemblea atópase a disposición dos socios na páxina web da SGAPEIO, no apartado A Sociedade, Actas e Informes ([web](#)).

Novas do IGE

Últimas novidades en publicacións e documentos do IGE

Análise do sector da pesca

A Análise do sector da pesca é unha publicación do Instituto Galego de Estatística (IGE) que dá cumprimento á Actividade de Interese Estatístico (AIE13) “Análise estatística de sectores produtivos e da estrutura económica en xeral” recollida no Programa estatístico da comunidade autónoma de 2014.

Esta publicación é a terceira dunha serie de análises iniciadas coa da cadea forestal-madeira, difundida en 2013, e coa análise do sector da automoción, difundida en 2014. O obxectivo destas publicacións é dobre: por unha banda, afondar no coñecemento da economía galega a través do estudo de sectores destacados na nosa estrutura económica e, por outra, dar unha mellor difusión ás estatísticas de síntese económica elaboradas polo IGE.

Consello Galego de Estatística

O día 12 de decembro de 2014 tivo lugar a reunión do Pleno do Consello Galego de Estatística, coa seguinte orde día:

1. Aprobación da acta da sesión anterior
2. Aprobación da memoria de actividades do 2013
3. Informe do proxecto de Decreto do Programa Estatístico do 2015
4. Proxectos técnicos das operacións estatísticas
5. Rolda de intervencións

Nesta reunión aprobouse a acta do pleno de novembro de 2013 e tamén se aprobou a Memoria de actividades do ano 2013. Neste ano celebráronse dous plenos do Consello Galego de Estatística nos que se aprobaron 16 proxectos técnicos, os resultados do Programa Estatístico Anual do 2012 e o proxecto do Programa Estatístico Anual do 2014.

Tamén se presentou o informe do proxecto de Decreto do Programa Estatístico de 2015. Este proxecto está dotado cun orzamento de 1,4 millóns de euros e destacan os seguintes puntos:

- O calendario é máis esixente e recolle o día de difusión das estatísticas conxunturais.
- Prográmanse novos indicadores sobre plans de infancia e adolescencia, fondos europeos, ordenación territorial e desenvolvemento rural.

- Xerárase información para cubrir metas non tratadas sobre capital social, participación cidadá e prácticas culturais.
- Necesidade de avanzar no mandato de aproveitar información administrativa e conseguir a colaboración da administración tributaria mediante acordos como os establecidos coa Seguridade Social.

Por último, presentáronse os proxectos técnicos das operacións estatísticas:

- Indicadores de seguimento das Directrices de Ordenación do Territorio (DOT) e Plan de Ordenación do Litoral (POL).
- Sistema de indicadores complementarios do Plan de Desenvolvemento Rural (PDR).

Convenios

- Firma de documento de cooperación entre a Universidade de Santiago de Compostela (USC) e o Instituto Galego de Estatística (IGE) no marco do proxecto: FCT-14-8405 “Bioestadística para Vivir y cómo Vivir de la Estadística”.
- Convenio de colaboración entre a Universidade de Santiago de Compostela e o Instituto Galego de Estatística para a realización de prácticas de alumnos/as do Máster universitario en Planificación e Xestión do Desenvolvemento Territorial e o Grao en Xeografía e Ordenación do Territorio da Facultade de Xeografía e Historia.

Cursos

O IGE impartirá, a través da Escola Galega de Administración Pública (EGAP) e por teleformación, o curso de “Cálculos estatísticos básicos co software libre R”

- Datas: 04/05/2015 a 03/06/2015
- Destinatarios: Persoal ao servizo da Administración pública da Comunidade Autónoma de Galicia
- Máis información na seguinte ligazón: [web](#).

Oposicións

Publicouse a lista dos aspirantes que superaron o proceso selectivo para o ingreso no corpo superior da Xunta de Galicia, subgrupo A1 da escala superior de estatísticos, que se pode consultar na seguinte ligazón: [web](#).

Traballos de Estatística no SUG

Teses

Nonparametric inference with directional and linear data

Eduardo García Portugués.

Directores: Wenceslao González Manteiga e Rosa M. Crujeiras Casais.

Data da defensa: 12/12/2014. USC.

Esta tese introduce novas ferramentas metodolóxicas para a inferencia nonparamétrica con datos direccionais (o seu soporte é a circunferencia, a esfera ou a hiperesfera de dimensión arbitraria) e lineais (i.e., datos Euclídeos usuais). Os datos direccionais aparecen en multitude de campos aplicados, como, por exemplo, proteómica ou ciencias medioambientais. A tese presenta novos métodos nonparamétricos para a estimación e o contraste de hipóteses, para as curvas de densidade e de regresión, en situacións nas que aparecen variables aleatorias direccionais. As principais contribucións desta tese recóllense en seis capítulos procedentes de seis artigos.

No Capítulo 2 empréganse cópulas para estimar densidades circular-lineais a partir de métodos paramétricos, semiparamétricos ou nonparamétricos e amósase unha aplicación medioambiental. O Capítulo 3 introduce un estimador núcleo para densidades direccional-lineais e as expresións exactas para o erro cadrático medio integrado. Tres novas propostas de selectores do parámetro ventana para o estimador núcleo direccional son comparadas no Capítulo 4 fronte a propostas previas nun extenso estudo de simulación. No Capítulo 5 constrúese un test de independencia, con calibrado por permutacións, que se aplica ao estudo da relación entre a orientación dos incendios e o seu tamaño. No Capítulo 6 desenvólvese un teorema central do límite para o erro cadrático integrado do estimador direccional-lineal, que serve para obter a distribución asíntótica do test de independencia e dun contraste de bondade de axuste para densidades paramétricas direccional-lineais. O Capítulo 7 dedícase á proposta dun novo estimador local con predictor direccional e resposta lineal, ó estudo das súas propiedades e á construción dun contraste de bondade de axuste para funcións de regresión paramétricas. Este contraste é aplicado a un conxunto de datos de text mining de alta dimensión.

Advances in Multiple Hypothesis Testing: The Sequential Goodness-of-Fit Procedure Revisited and Expanded

Irene Castro Conde.

Director: Jacobo de Uña Álvarez.

Data da defensa: 18/12/2014. UVIGO.

Os contrastes de hipóteses múltiples ocúpanse da toma de decisións en situacións onde numerosas hipóteses nulas son consideradas simultaneamente. Isto ocorre, por exemplo, cando se compara a expresión xénica en dúas mostras ao longo dun conxunto de xenes. Os problemas de contrastes múltiples adquiriron grande importancia nos últimos anos debido á gran cantidade de información dispoñible en campos como as ciencias ómicas, a neuroimaxe ou a farmacoloxía.

No ano 2009, nun artigo publicado en BMC Bioinformatics, os xenetistas vigueses Antonio Carvajal e Emilio Rolán introduciron, en colaboración co matemático Jacobo de Uña, o método Sequential Goodness-of-Fit (SGoF).

SGoF decide qué hipóteses son falsas mediante a comparación do número de p -valores por debaixo dun umbral inicial de significación e a cantidade esperada baixo a hipótese nula completa. Este método tivo pronto moita penetración na investigación aplicada ao garantir unha grande potencia estatística baixo unha taxa de falsos positivos moderada.

Nesta tese, o obxectivo é estudar máis a fondo as propiedades do método SGoF, e estender esta metodoloxía a escenarios como o Bayesiano ou o discreto. Ademais de achegas metodolóxicas, que inclúen a definición e estudo de p -valores axustados, esta tese presenta un novo paquete de R, chamado *sgof*, que implementa estes novos procedementos. Varios conxuntos de datos reais da bioloxía, a biomedicina ou a farmacoloxía son usados como ilustración.

A tese, realizada na modalidade de Doutoramento Internacional, foi defendida o día 18 de decembro de 2014 na Facultade de Económicas de Vigo, obtendo a cualificación de Sobresaliente “Cum Laude”. O tribunal estivo formado polos recoñecidos expertos Daniel Yekutieli (Universidade de Tel Aviv), Frank Bretz (Novartis e Universidade de Hannover) e María Concepción Ausín (Universidade Carlos III de Madrid). Parte dos resultados da tese están xa publicados en revistas de prestixio internacional, como Biometrical Journal, The R Journal e Computational Statistics.

Métodos automáticos non paramétricos de estimación de conxuntos

Paula Saavedra Nieves.

Directores: Wenceslao González Manteiga e Alberto Rodríguez Casal.

Data da defensa: 06/03/15. USC.

A estimación de conxuntos abre un capítulo relativamente recente da estatística matemática onde a xeometría xoga un papel moi relevante. Esta teoría ten como finalidade estimar un conxunto no espazo Euclidiano a partir dunha mostra aleatoria de puntos cuxa distribución está intimamente relacionada con el. Neste traballo centrámonos nos problemas de estimación do soporte e de conxuntos de nivel dunha función de densidade.

Existen distintas alternativas na literatura dependendo das condicións de forma asumidas sobre o conxunto a reconstruír. Se non se dispón de ningunha información a priori, será preciso considerar estimadores flexibles que permitan abordar de xeito eficiente a maior cantidade de situacións posibles. Noutro caso, se a estimación se restrinxe a unha familia de conxuntos, pódese traballar con estimadores máis sofisticados, que se adapten mellor ás restricións xeométricas establecidas. A maioría destes estimadores dependen fortemente da elección de parámetros de suavizado ao igual que sucede no contexto da estimación funcional non paramétrica. No Capítulo 2, revísanse en profundidade as metodoloxías automáticas existentes para a estimación do soporte e de conxuntos de nivel. Para os últimos, móstranse tamén os resultados obtidos da realización dun extenso estudo de simulación onde se analiza o comportamento práctico dos métodos considerados.

Nos Capítulos 3 e 4, propóñense dous novos algoritmos para reconstruír soportes e conxuntos de nivel r -convexos. A r -convexidade é unha condición de forma flexible que xeneraliza a noción de convexidade onde o parámetro de suavizado r pode ser visto tamén como un índice de forma do conxunto. O seu valor real é, por suposto, descoñecido. O obxectivo principal deste traballo de investigación pasa por estimalo de xeito automático e consistente a partir da mostra de puntos para, a continuación, analizar o comportamento teórico dos estimadores resultantes do soporte e dos conxuntos de nivel. Para analizar o comportamento práctico dos métodos propostos, realizáronse varios estudos de simulación e analizáronse dous problemas con datos reais.

Traballos fin de máster

Selección de variables: unha revisión de métodos existentes

Ana González Vidal.
Director: Germán Aneiros Pérez.
Data da defensa: 22/01/15.
UDC, modalidade académica.

O tópico a tratar neste traballo de fin de máster é o da selección de variables explicativas a ser incorporadas nun modelo de regresión lineal. En primeiro lugar, introducirase o modelo de regresión lineal e motivarase a necesidade da selección de variables. Considerarase tanto o caso en que o total de covariables, p , é inferior ao tamaño muestral, n , como o caso correspondente a p moito maior que n . En primeiro lugar, presentaranse diferentes métodos de selección de variables así como as propiedades que cumpren cada un deles. Posteriormente compararanse, a través dun estudo de simulación, o comportamento dos devanditos métodos de selección de variables. O traballo de fin de máster finalizará aplicando os devanditos métodos a algún conxunto de datos reais. Na implementación (simulación e aplicación a datos reais) utilizaranse rutinas xa existentes en distintos paquetes de R.

Comparación de técnicas de estimación de datos xenéticos: imputación xenética vs interpolación espacial

Adrián Lema Casal.
Directora: Rosa Mª Crujeiras Casais (USC)
Titora: Raquel Cruz Guerrero (CIMUS).
Data da defensa: 22/01/15.
USC, modalidade prácticas en empresa.

En numerosos estudos de asociación xenética con enfermidades avalíase a diferenza en frecuencia relativa dos xenotipos en pacientes e poboación sa nun elevado número de SNPs empregando bases de datos comerciais. Habitualmente existen diferenzas entre os SNPs xenotipados en cada base de datos, feito que dificulta a replicación ou combinación de distintos estudos individuais. Debido a isto desenvóléronse diversos métodos de imputación xenética que estiman o xenotipo para un SNP non xenotipado nunha serie de individuos, empregando técnicas asociadas aos modelos de Markov Ocultos. Non obstante, a frecuencia relativa dun SNP pode verse modificada ao longo dunha rexión xeográfica ampla (en función da poboación estudada). Nestes casos, i.e., cando o que interesa é maximizar a información xenética en diferentes puntos espaciais resultado da combinación de diferentes estudos en diversos puntos xeográficos (aínda non sendo todos os SNPs coincidentes), pódese recorrer á realización dunha interpolación espacial da frecuencia poboacional observada neses puntos, utilizando técnicas xeoestadísticas como o kriging. Este TFM pretende, empregando unha base de datos reais, comparar ambas técnicas de estimación en contextos que poidan resultar de utilidade en posteriores aplicacións reais.

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

Coñecéndonos

Coñecendo a un socio da SGAPEIO: Andrés E. Cabana González.

Andrés E. Cabana González é Licenciado en Matemáticas (especialidade de Estatística e IO) pola USC. Profesor de matemáticas de Ensino Secundario, funcionario de carreira dende 1996. Actualmente director do IES Salvaterra de Miño, no que ten praza definitiva dende o curso 1999-2000.



Cantos anos levas traballando en estatística ou IO?

En estatística ou IO, como tal, non traballo especificamente, pero levo explicando dende hai vinte anos cousas de estatística dentro do currículo de matemáticas. É obvio que son cousas elementais do currículo de secundaria e bacharelato, pero, a fin de contas, tamén é traballar coa estatística.

Descríbenos o teu traballo

Son profesor de matemáticas de ensino secundario. Na actualidade, director do IES Salvaterra de Miño, pero isto último é algo temporal. O que realmente me gusta é estar na aula co alumnado.

Dende cando es socio/a da SGAPEIO?

Pois creo que practicamente dende a súa fundación. Igual estou equivocado, pero, se non recordo mal, son socio dende o ano 1993.

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

Non sei se o primeiro foi un que se celebrou na Coruña, a verdade é que levo tempo sen asistir. Creo que fun aos tres primeiros e logo abandonei.

O tema do teu próximo proxecto/artigo?

Non teño ningún proxecto a curto prazo en mente. Tal vez o próximo ano faga algún traballinho con alumna-

do de 4º da ESO, de Diversificación Curricular, para presentalo ao concurso de “Incubadora de Sondaxes”. Levo anos intentado facelo xa que anualmente sigo con detalle o concurso, de feito, no primeiro concurso da “Incubadora” fun un dos membros do xurado.

Un curso que che gustaría facer?

Gustaríame que se seguisen organizando, coma nestes dous últimos anos, as Xornadas sobre o ensino da estatística. Asistín ás do ano 2013, pero non puideron asistir ás do 2014. Espero que continúe esta iniciativa.

Que libro de estatística ou I.O. recomendarías?

Paréceme moi interesante para unha aprendizaxe divertida da estatística “La Estadística en Cómic” de Larry Gonick e Woolcott Smith. Imaxino que para alguén experto na materia é demasiado sinxelo, pero para os máis profanos, que non nos dedicamos única e exclusivamente a este campo, pódenos resultar moi útil.

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

A resposta ideal para publicar en Informest sería: estatístico. Pero tirando un pouco pola liña marcada por Ignacio García Jurado na súa entrevista, direi que me sinto máis matemático ca estatístico. Dende a miña época de estudante de BUP no instituto quixen ser profesor de matemáticas, algo que lle teño que agradecer a don Enrique Pujales.

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

Gráficos e análise de datos. Lamentablemente, tamén se me vén á cabeza o abuso interesado, partidista e parcial que se fai da estatística a través da manipulación dos datos.

Que é o que máis valoras do labor da SGAPEIO?

O seu labor divulgativo da estatística.

En que cres que deberíamos mellorar?

Non creo que sexa labor só da SGAPEIO, senón de todo o profesorado de matemáticas actuando dende idades temperás, a de formar á cidadanía nunha cultura estatística que nos permita ser críticos cos datos estatísticos con que, día a día, somos bombardeados polos medios de comunicación.



Sabías que?

A Investigación Operativa na industria da dobraxe.

por Luisa Carpente e Silvia Lorenzo.

Na actualidade, a industria da dobraxe ten que axustar o máximo posible os presupostos de produción de películas. Para levar a cabo o proceso de dobraxe cómpre dividir o guión en unidades coñecidas como “takes”. Os “takes” corresponden a un pequeno número de frases que poden corresponder a diferentes personaxes. Estes personaxes son interpretados por un ou varios actores de dobraxe.

Cando a produtora decide os actores que deben dobrar os personaxes, xa sabe cal é o custo fixo relativo ás aparicións dese actor na película. Agora ben, a produtora tamén paga unha cantidade polas veces que convoca ao actor para facer a dobraxe. Tendo en conta que os rangos de corte da película e o número de actores son relativamente elevados e que, tamén, cada sesión de dobraxe debe conter un número máximo de “takes”, é necesario obter unha planificación de dobraxe que cumpra estas condicións e sexa o máis barata posible.

É aí onde ten sentido introducir as técnicas de investigación operativa para obter a planificación óptima. Nunha primeira etapa bótase man dun modelo de programación lineal enteira e despois, para aforrar en tempos de computación, faise uso de algoritmos heurísticos que atopan, na maioría dos casos, os esquemas máis baratos. Por exemplo, na realidade, na dobraxe da película Donkey Xote, a produtora pagou un total de 43 sesións mentres que, co modelo de programación lineal e co algoritmo heurístico, tería pagado un total de 35. Así mesmo, mentres o problema de programación lineal pode tardar varias horas en dar unha solución nalgúns casos, o algoritmo nunca excede dun minuto.

Nieves R. Brisaboa, Luisa Carpente, Ana Cerdeira-Pena e Silvia Lorenzo-Freire. (2015). “Optimization in dubbing scheduling”. TOP, DOI: 10.1007/s11750-015-0361-4

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.

Revisión lingüística: María Miragaya Pereira

Coa colaboración de: Juan Picos, Guido I. Novoa Flores, M^a José Ginzo Villamayor, Andrés E. Cabana González, Luisa Carpente.

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