

Colaboración

Recensión bibliográfica

Novas

Novas do IGE

Traballos de Estatística e IO

Coñecéndonos

Sabías que?

Editorial

Juan José Vidal Puga – Presidente da SGAPEIO

Recentemente renovouse o Consello Executivo da SGAPEIO. Habitualmente esta renovación é parcial, mais nesta ocasión produciuse unha renovación case total, con nove cambios sobre dez membros, incluído quen escribe estas liñas. A única persoa que continúa no Consello Executivo é Leticia Lorenzo, que asume agora a Secretaría da Sociedade, garantindo así unha necesaria continuidade institucional.

Tomar a responsabilidade da presidencia supón para min unha grande ilusión, pero tamén unha fonda responsabilidade, especialmente por suceder a César A. Sánchez Sellero, quen liderou a SGAPEIO nos últimos anos cunha dedicación e un compromiso exemplares. Quero expresar o meu sincero agradecemento ao Consello Executivo saínte polo excelente traballo realizado, que deixa unha Sociedade activa, recoñecida e con múltiples liñas de actuación consolidadas. Un agradecemento moi especial merece Covadonga Rodríguez-Moldes, recentemente xubilada, pola súa traxectoria e pola súa contribución decisiva á sección de Sondaxe de Experimentos, hoxe renomeada como AVATAR.

Precisamente, o concurso AVATAR continúa a ser unha das iniciativas máis salientables da SGAPEIO. Os resultados da representación galega na última edición foron sobresaíntes, acadando a vitoria en dúas das tres categorías e unha mención especial na terceira. Este ano, ademais de ser a primeira edición baixo a nova denominación, AVATAR ten un significado especial desde o punto de vista loxístico, ao celebrarse en Las Palmas de Gran Canaria, o que supón un reto adicional en termos de financiamento dos desprazamentos das persoas gañadoras da fase galega organizada pola SGAPEIO.

Queremos tamén manter e reforzar colaboracións que consideramos estratéxicas. Continúa a cooperación coa Asociación Galega do Profesorado de Educación Matemática (AGAPEMA), coa Sociedade Portuguesa de Estatística (SPE), que inclúe a celebración do II Simpósio de Biometría Luso-Galaico, así como co Instituto Galego de Estatística (IGE) e o Instituto Nacional de Estatística (INE). A estas relacións súmase a colaboración coa Sociedade Española de Estatística, Investigación Operativa e Ciencia de Datos (SEIO), coa que a SGAPEIO mantén un acordo de cooperación que consideramos fundamental para fortalecer os vínculos co ámbito estatal da nosa disciplina. Todas estas colaboracións constitúen unha base sólida que queremos conservar e ampliar no futuro.

Son moitos os retos e oportunidades que se presentan para a Estatística e a Investigación de Operacións. Confiamos en afrontalos coa colaboración activa das socias e socios, para que a SGAPEIO continúe a ser un espazo de encontro, participación e impulso para todas as persoas que ensinan, investigan, traballan ou simplemente desfrutan da Estatística e da Investigación de Operacións en Galicia.

Dirección:

Laura Davila Pena
Marta Sestelo Pérez

Comunicación coa SGAPEIO:

www.sgapeio.es
secretaria@sgapeio.es
boletin@sgapeio.es
[@sgapeio](https://twitter.com/sgapeio)
[sgapeio](https://facebook.com/sgapeio)
[sgapeio](https://instagram.com/sgapeio)
[Sgapeio](https://linkedin.com/company/sgapeio)

Depósito Legal:

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

A análise de supervivencia como ferramenta para mitigar o abandono laboral no sector TIC

por Nora M. Villanueva e Marta Sestelo

Nora M. Villanueva é doutora en Estatística e Investigación Operativa pola Universidade de Vigo onde actualmente traballa como Profesora Axudante Doutora no Departamento de Estatística e Investigación Operativa. É membro do Grupo SiDOR e persoal investigador colaborador do CITMAga.

Marta Sestelo é doutora en Estatística e Investigación Operativa pola Universidade de Vigo onde actualmente traballa como Profesora Titular de Universidade no Departamento de Estatística e Investigación Operativa. É membro do Grupo SiDOR e persoal investigador vinculado do CITMAga.

Canto tempo transcorre dende unha cirurxía ata que xorde unha complicación? Que tempo de progresión existe ata que unha persoa con prediabetes desenvolve diabeite tipo 2 baixo un tratamento específico? Ou que marxe separa a extirpación dun tumor da posible aparición dunha metástase? Estas interrogantes, xunto coa necesidade de identificar factores de risco, foron o motor da análise de supervivencia. Aínda que o seu nome delata unha orixe puramente médica, esta disciplina naceu cun obxectivo tan vital como estatístico: modelar o tempo ata que ocorre un evento de interese —xa sexa a curación, a recaída ou o falecemento—, permitíndonos comprender mellor a dinámica do tempo que nos queda.

Hai xa tempo que a análise de supervivencia transcendeu o ámbito clínico para converterse nun piar estratéxico en sectores como a Economía ou as Tecnoloxías da Infor-

mación e as Comunicacións (TICs). Na actualidade, o seu potencial vai moito máis alá dos hospitais: permite a unha entidade bancaria anticipar o momento dunha falta de pagamento, a unha consultora tecnolóxica estimar a taxa de rotación do seu talento humano ou a un operador de telefonía calcular o “tempo de vida” dun cliente antes de que decida marchar á competencia.

Nesta nova contorna, as variables cambian de nome pero manteñen a súa esencia: o paciente transfórmase en cliente ou empregado, e o evento clínico cede o seu lugar ao incumprimento dun crédito, á baixa laboral ou ao abandono dun servizo (*churn*). Esta transposición conceptual permite ás organizacións pasar dunha cultura reactiva a unha proactiva. Grazas a ela, as empresas poden anticiparse aos eventos críticos, optimizando os seus recursos e deseñando estratexias de fidelización ou retención de talento plenamente personalizadas.

A gran singularidade da análise de supervivencia reside na censura que afecta á variable tempo. Na práctica, só coñecemos o tempo exacto de resposta daqueles individuos que experimentan o evento antes de que finalice o seu seguimento. Para o resto, a información é parcial ou incompleta: só sabemos que o tempo ata o evento é maior que o seu período de seguimento. Estes tempos denomínanse censurados (Figura 1), e a análise de supervivencia é, precisamente, a ferramenta encargada de xestionar esta información incompleta sen nesgar os resultados.

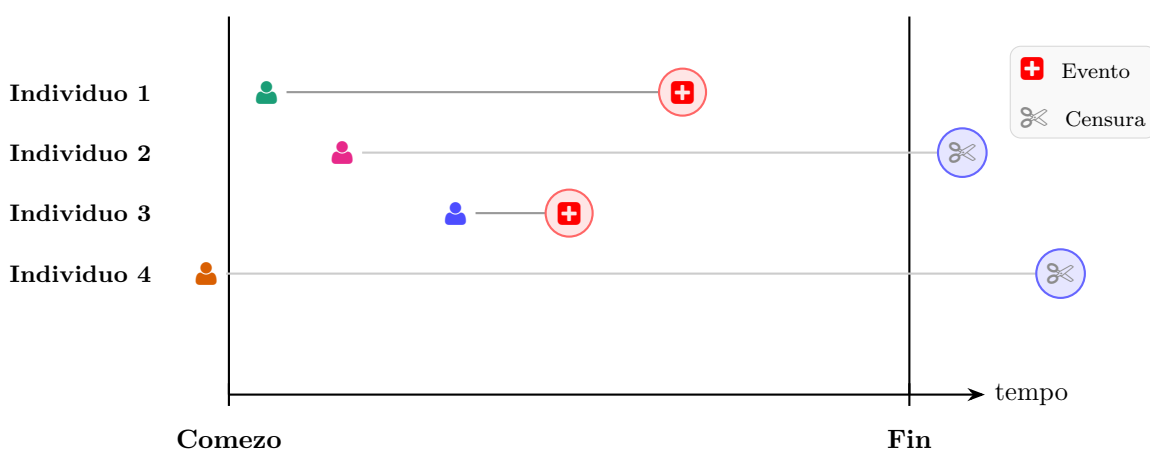


Figura 1: Representación dos tempos de vida para catro individuos.

Análise do abandono laboral

Hoxe en día, a toma de decisións estratéxicas nas organizacións está a experimentar unha transformación profunda grazas á analítica de Recursos Humanos, tamén coñecida como *People Analytics*. Esta disciplina utiliza datos e técnicas avanzadas de análise co obxectivo de comprender e optimizar o rendemento e a xestión do talento, aplicando un enfoque científico aos procesos e decisións relacionados co capital humano.

Nas empresas do sector TIC, a xestión do capital humano enfrenta un desafío constante: mitigar a alta rotación nun mercado globalizado e altamente competitivo, onde a oferta de profesionais cualificados é insuficiente para cubrir a demanda existente. Neste escenario, a analítica de Recursos Humanos convértese nunha ferramenta predictiva clave para desvelar patróns de comportamento ao longo do tempo, permitindo identificar os factores que precipitan o abandono voluntario e anticipar estas situacións en función de variables como o departamento, o rol ou a experiencia previa do profesional.

Baixo este enfoque, o obxectivo desta colaboración foi estudar o abandono do persoal dunha empresa galega do sector TIC a través da comparación das curvas de supervivencia, co fin de proporcionar evidencias que permitan adoptar medidas eficaces de retención de talento.

A base de datos utilizada integra información de 288 individuos. Para cada suxeito, rexistrouse o tempo de permanencia na empresa (en anos) desde a súa incorporación ata o cese, xunto coa variable categórica departamento, que consta de sete niveis ($j = 1, \dots, 7$). Así mesmo, inclúese un indicador da situación laboral: esta variable toma o valor 1 se o individuo abandonou

a compañía e 0 se permanece en activo ao concluír o estudo, considerándose neste último suposto como unha observación censurada.

Tal como se reflicte na Figura 3, as curvas de supervivencia estimadas para os sete departamentos amosan traxectorias de permanencia diverxentes. Esta observación preliminar formula a cuestión central da nosa análise: Existen diferenzas significativas na fidelización dos empregados en función do departamento ao que pertencen?

Para dar resposta á cuestión formulada, aplicouse a metodoloxía descrita anteriormente. Tras a execución do algoritmo, os resultados determinaron que o número óptimo de grupos é $K = 4$. Na Figura 4 amósase a asignación definitiva das curvas aos seus respectivos clústeres.

Esta agrupación permite observar que os departamentos 4, 6 e 7 (Grupo I, en vermello) presentan os niveis de supervivencia máis elevados, consolidándose como as áreas con maior capacidade de retención de talento. Séguenlles os departamentos 1 e 3 (Grupo II, en azul), que amosan un patrón de abandono intermedio con traxectorias moi similares entre si. Pola súa banda, os departamentos 2 (Grupo III, en amarelo) e 5 (Grupo IV, en verde) presentan comportamentos singulares ao conformar clústeres independentes. É especialmente notable o caso do departamento 5, cunha curva de supervivencia que cae abruptamente ata alcanzar o cero antes dos catro anos. Este fenómeno reflicte unha situación crítica na que a unidade quedou sen persoal activo debido ao abandono sucesivo dos seus traballadores. Como consecuencia desta perda de capital humano, a empresa optou pola extinción definitiva de dito departamento.

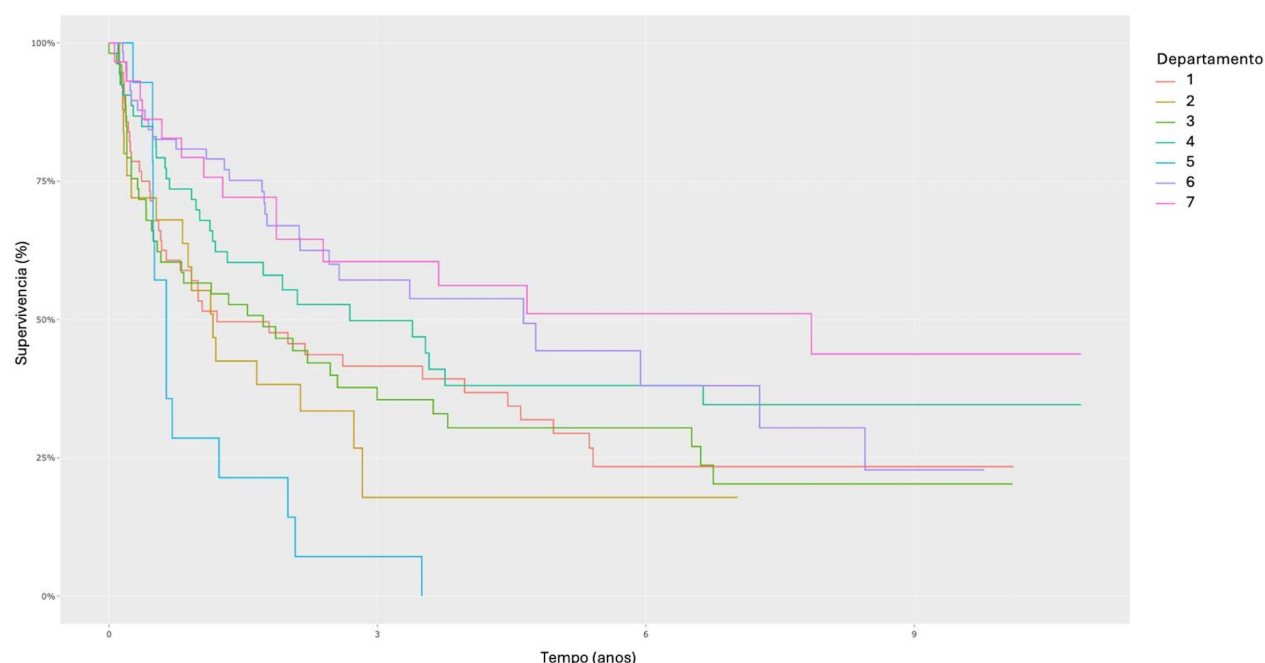


Figura 3: Curvas de supervivencia estimadas para cada un dos sete departamentos.

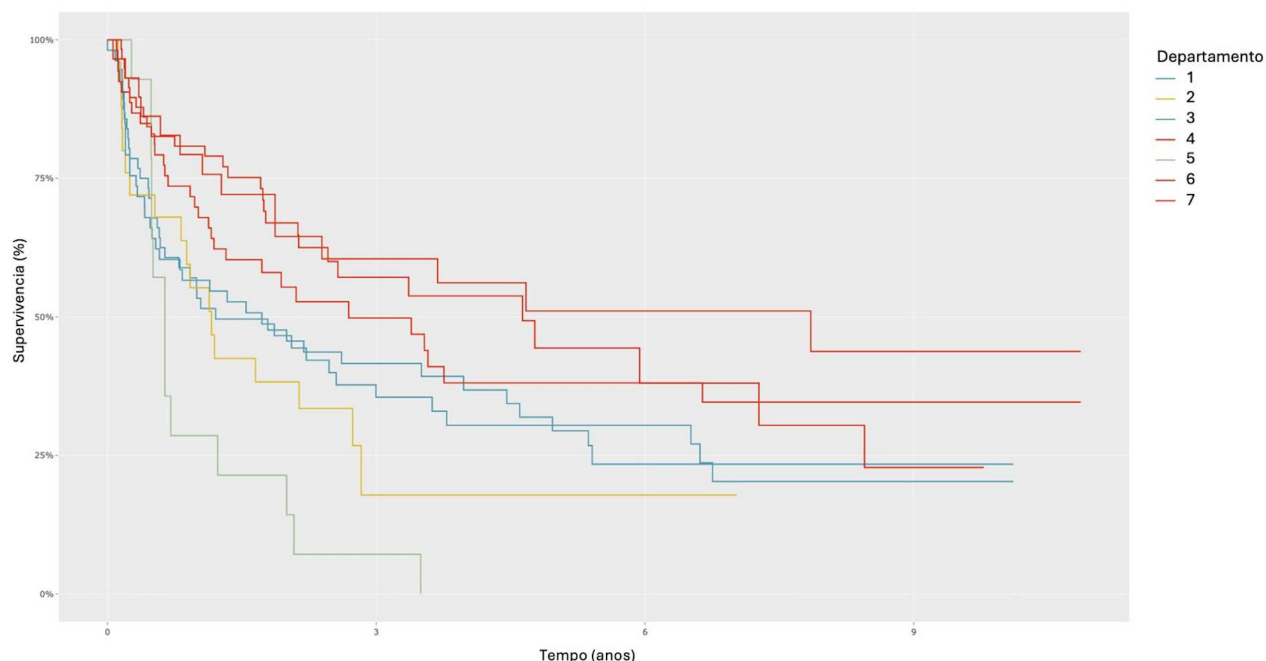


Figura 4: Curvas de supervivencia estimadas para cada un dos departamentos e coloreadas en función do grupo ao que pertencen, $K = 4$.

Conclusións

A integración das técnicas estatísticas de agrupamento ou *clustering* coas de análise de supervivencia proporciona unha ferramenta útil para a xestión do talento em-

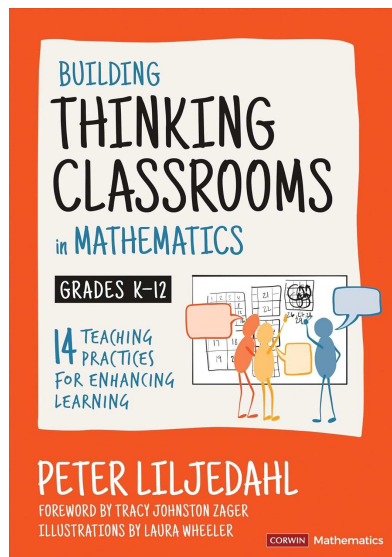
presarial. Esta metodoloxía permite entender os tempos de permanencia de cada departamento, transformando os resultados da análise en estratexias reais para manter o talento e polo tanto, mellorar a xestión de equipos.

Referencias

- Efron, B. (1981). Censored data and the bootstrap. *Journal of the American Statistical Association*, 76(374):312–319.
- Kaplan, E. L. and Meier, P. (1958). Nonparametric estimation from incomplete observations. *Journal of the American Statistical Association*, 53(282):457–481.
- Kaufman, L. and Rousseeuw, P. J. (1990). *Finding groups in data: an introduction to cluster analysis*. Wiley, Hoboken, NJ.
- Macqueen, J. (1967). Some methods for classification and analysis of multivariate observations. In *Proceedings of the Fifth Berkeley Symposium on Mathematical Statistics and Probability*, volume 1, Berkeley, CA. University of California Press.
- Mantel, N. (1966). Evaluation of survival data and two new rank order statistics arising in its consideration. *Cancer Chemother Reports*, 50(3):163–170.
- Villanueva, N. M., Sestelo, M., and Meira-Machado, L. (2019). A method for determining groups in multiple survival curves. *Statistics in Medicine*, 38(5):866–877.
- Villanueva, N. M., Sestelo, M., Meira-Machado, L., and Roca-Pardiñas, J. (2021). clustcurv: An R package for determining groups in multiple curves. *The R Journal*, 13(1):164–183.

Building Thinking Classrooms in Mathematics

por Fabio José González Eiras



En outubro de 2023, no marco do programa PIALE (Programa Integral de Aprendizaxe de Linguas Estranxeiras) da Consellería de Educación, tiven a oportunidade de realizar unha estadía formativa na cidade de Toronto (Ontario, Canadá), concretamente no Forest Hill Collegiate Institute. Este programa ten como obxectivo principal a inmersión do profesorado en sistemas educativos estranxeiros, co fin de mellorar tanto as competencias lingüísticas como as pedagóxicas. Profesorado de distintas especialidades fomos asignados a diversos centros educativos da cidade; no noso caso, Álvaro Crego Deán e mais eu exercemos como representantes do profesorado galego de matemáticas de ensino medio.

Un dos aspectos que máis nos chamou a atención desde o primeiro momento foi o clima humano e pedagóxico xerado no departamento de matemáticas, dirixido polo entrañable Mr. Ahamed Basheer. Cada sesión comezaba cun xesto sinxelo pero significativo: o profesor recibía o alumnado á entrada da aula, interesándose polo seu estado anímico e benestar. Esta práctica, lonxe de ser anecdótica, parecía sentar as bases dun ambiente de confianza e disposición positiva cara á aprendizaxe.

Desde o punto de vista metodolóxico, o funcionamento da aula resultaba especialmente salientable. O alumnado traballaba en grupos aleatorios de tres persoas, de pé e en pizarras verticais distribuídas polo espazo. As tarefas propostas estaban deseñadas de xeito que permitían múltiples niveis de profundidade, desde aproximacións iniciais ata desenvolvementos máis elaborados. O docente desprazábase continuamente entre os grupos, ofrecendo orientacións puntuais, formulando preguntas

estratéxicas e realizando avaliación formativa en tempo real. A primeira impresión era clara: o alumnado afrontaba problemas matemáticos de certa complexidade con seguridade, curiosidade e unha actitude claramente favorable ao desafío intelectual.

Ante a novidade do enfoque, pregunteille ao Mr. Basheer se se trataba dunha proposta persoal. A súa resposta foi reveladora: “*Nope, it isn't, my friend. This all comes from here!*”, mentres amosaba un pequeno libro de cuberta laranxa. A súa resposta remitiu a *Building Thinking Classrooms in Mathematics*, de Peter Liljedahl (304 páxinas), traducido ao castelán como *Diseñando aulas para pensar en matemáticas*. A obra estrutúrase en 14+1 capítulos, cada un centrado nunha práctica docente orientada a transformar a aula tradicional nun espazo que favoreza a activación do pensamento. Abórdanse cuestións que van desde a configuración do espazo físico ata os procedementos de avaliación ou a xestión da toma de apuntamentos. A lectura non precisa ser lineal: o profesorado pode acceder aos capítulos segundo as súas necesidades ao iniciar ou adaptar a metodoloxía. Con todo, se se desexa implementar progresivamente as propostas, resultan especialmente nucleares o capítulo 1 (tipoloxía de tarefas), o 6 (formulación de tarefas), o 9 (uso de pistas e extensións) e o 13 (avaliación formativa). O capítulo 15 ofrece, ademais, orientacións para ensamblar coherentemente as catorce prácticas descritas.

Como exemplo de tarefa destinada a “detonar” o pensamento no alumnado, propónse o seguinte problema: teño un temporizador de catro minutos e outro de 7, dos que se poden voltear e deixar que a area caia. Podo utilizalos para cocer un ovo durante, exactamente, 9 minutos? Se é así, canto tempo terá que esperar alguén para comer o ovo?

Considero que este libro resulta de grande interese para o profesorado de tódolos niveis educativos, dende primaria ata a universidade. A súa proposta metodolóxica é innovadora, flexible e facilmente adaptable a contextos moi diversos. O seu eixo central é a mobilización efectiva do pensamento matemático no alumnado, promovendo condicións que favorecen o chamado estado de fluxo descrito por Mihaly Csikszentmihalyi.

En particular, os bloques de estatística e probabilidade no ensino secundario e bacharelato presentan un enorme potencial para a aplicación desta metodoloxía, ao permitir o deseño de tarefas ricas arredor do reconto, o azar ou a inferencia estatística, ámbitos especialmente propicios para fomentar o razoamento, a discusión matemática e a toma de decisións fundamentadas.

Xubilación de Antonio Sampayo Flores

os compañeiros do Campus de Lugo (USC)



Antonio Sampayo

Antonio foi un referente para os seus compañeiros, por todo o que aportou, aporta e seguro que seguirá aportando á promoción da Estatística e da Investigación de Operacións. Falar da súa carreira como docente é falar de centos e centos de horas nas aulas, na xa desaparecida Escola de Peritos onde empezou, no antigo Colexio Universitario, e nos últimos anos, na Facultade de Ciencias, na Facultade de Veterinaria e na Escola Politécnica Superior de Enxeñaría. Foi así que impartiu clases nas diversas titulacións de grao, e nas antigas enxeñarías e licenciaturas nas que a nosa área ten ou tivo encargo docente. Antonio sempre estivo disposto a innovar e a colaborar cos seus compañeiros en calquera iniciativa que xurdiu no campus de Lugo. Ao botar a vista atrás, podemos lembrar as andainas desde o despacho ás aulas doutros centros cargando co primeiro computador portátil da nosa área, “o do millón de pesetas” e os seus moitos quilos de peso, o despregue das transparencias de acetato, e máis adiante, as prácticas nas aulas de informática co reparto de disquetes para os exames con Statgraphics, SPSS ou Excel.

A súa pegada non se limita ás aulas, xa que desempeñou un papel relevante na vida académica e científica, participando activamente en moitas actividades: reunións científicas, cursos e diversos eventos. En todas elas deixouse notar a súa experiencia nas súas outras facetas persoais, polas profesións que desempeñou, no Instituto Nacional de Estatística primeiro, e posteriormente na administración pública galega, así como pola súa pertenza ao grupo de teatro Achádego e o seu coñecemento do mundo da cultura da cidade de Lugo. Como suma de todo isto, fomos quen de gozar da súa gran capacidade organizativa.

Mención especial merecen as súas contribucións á SGAPEIO. É un dos 32 socios fundadores da sociedade, crea-

da no 1992, inicialmente referida como SGPEIO. Colaborou na organización de varios congresos, entre os que cabe destacar os que tiveron lugar en Lugo, o III Congreso SGAPEIO (1997), no que foi o presidente do comité organizador, e o XII Congreso SGAPEIO (2015). O congreso do 97 foi o que acadou a maior participación de todos os organizados ata agora pola sociedade, con 227 congresistas inscritos, que xunto cos especialistas invitados e outros participantes superou os 250 asistentes. Este congreso é recordado non só polas súas actividades científicas, senón tamén polas súas actividades sociais e culturais. Outras colaboracións destacadas son a súa participación na tradución ao galego do cómic “A Estatística ¡en caricaturas!”, a organización e a impartición de cursos promovidos pola SGAPEIO, e as súas achegas a moitos números de Informest.

A súa presenza na nosa área achegou equilibrio e moito criterio, sempre aberto aos cambios e presentándose como voluntario para calquera nova tarefa. Entre as moitas virtudes de Antonio, queremos destacar a súa coherencia e a ética en tódalas actividades que realizou. Os seus compañeiros do campus de Lugo só podemos darlle as grazas por todo o que fixo con nós e para nós. Consideramos a Antonio non só coma un gran compañeiro senón que despois de todos estes anos xuntos, Antonio forma parte deste grupo de amigos (case que poderíamos dicir, a nosa segunda familia) que formamos a área de Estatística e Investigación Operativa no campus de Lugo. Ademais, a partir de agora, Antonio terá máis tempo para organizar as ceas de matemáticos do campus de Lugo e para colaborar connosco en actividades extra académicas para promover a estatística e a investigación operativa, entre as que, seguindo a secuencia 1997, 2015, . . . , haberá outro congreso da SGAPEIO en Lugo.

Os profesores de Lugo da área de Estatística e Investigación Operativa:

José M^a. Alonso, José Manuel Colmenero, Nieves Muñoz e Luis Ramil.



De menos a máis lonxe de Antonio: Nieves, Pepe, José Manuel e Luis no XII Congreso SGAPEIO

Novidades na información estatística publicada

Análise lonxitudinal das afiliacións ao Sistema da Seguridade Social en Galicia

A actividade *Análise lonxitudinal das afiliacións ao Sistema da Seguridade Social en Galicia*, elaborada por primeira vez polo IGE, ten como obxectivo ofrecer unha visión lonxitudinal das afiliacións. Estúdanse o número e a duración media dos períodos de afiliación, entendidos como períodos de tempo que transcorre entre a data de alta e a data de baixa. Tamén se estuda a intensidade do traballo, definida como a proporción de días que unha persoa está inscrita en alta laboral á Seguridade Social no ano en estudo. Por último, analízase a permanencia temporal das afiliacións dadas de alta no ano de estudo.

Ao longo do ano 2024, 1.163.908 persoas que residen e traballan en Galicia estiveron dadas de alta nalgún momento. O 59,7% destas persoas traballaron todo o ano en Galicia, mentres que o 9,9% traballaron menos da cuarta parte do ano, o que suxire a presenza de empregos de curta duración ou de baixa intensidade temporal.

Nese mesmo período, as persoas tiveron de media 2 afiliacións, e cada afiliación durou de media 164,3 días. O número medio de afiliacións foi parecido entre os homes e mulleres, pero a duración media foi maior nos homes (175,1 días) que nas mulleres (154,6 días). [+ info](#)

Enquisa de ciberseguridade

A falta de seguridade dixital pon en risco a información sensible e pode afectar á reputación e á viabilidade económica dunha empresa. Diseñouse esta operación estatística cos seguintes fins: identificar o nivel de preocupación que suscita a ciberseguridade nas empresas, e a regulación e normativa que lles afecta nesta materia; coñecer as medidas adoptadas, os incidentes sufridos e as súas consecuencias na actividade da empresa; coñecer os recursos investidos e usados na protección contra ameazas cibernéticas e as razóns destes investimentos; estudar a actuación das empresas ante ameazas ou incidentes cibernéticos a través da asistencia externa, accións de mellora, actividades formativas e persoal especialista.

A enquisa vai dirixida ás empresas con sede social en Galicia que contan con un ou máis asalariados no réxime xeral en polo menos un establecemento en Galicia e con actividade incluída nas seccións: B a J e L, M e N da CNAE-2009.

En Galicia, en 2025, o 15,49% das empresas sufriron algún incidente de ciberseguridade. A porcentaxe de incidentes de ciberseguridade aumenta en todos os sectores a medida que crece o número de empregados da empresa, e son as empresas do “Comercio” as que presentan unhas porcentaxes máis elevadas.

Entre as ameazas e vulnerabilidades á seguridade das

empresas galegas, os ataques de *phishing* (intentos de obter información confidencial suplantando a identidade dun terceiro) son os máis comúns. En 2025, o 10,19% das empresas galegas estudadas sufriron un ataque de *phishing*. [+ info](#)

Inserción laboral das persoas egresadas no Sistema Universitario Público de Galicia

O IGE elabora por segundo ano a actividade estatística *Inserción laboral das persoas egresadas no SUG*, co obxectivo de ofrecer información sobre a súa incorporación ao mercado laboral. Para iso empréganse os datos de matrícula facilitados polas universidades galegas e a información sobre emprego e desemrego da Base de Datos Sociodemográfica do IGE.

O estudo abrangue só as persoas tituladas nas universidades do SUG que residen ou traballan en Galicia. Estudarase a súa situación en relación coa actividade rexistrada cun período de seguimento de ata tres anos desde o egreso, tomando como referencia o último día hábil do período de seguimento.

No curso que remata en 2022 tituláronse en Galicia 11.936 persoas: arredor do 40% na USC, o 32,7% na UVigo e o 27,4% na UDC. Neste traballo analízanse as 11.242 persoas tituladas que residen ou traballan en Galicia (94,2% do total).

O 57,1% dos titulados universitarios no curso 2021/22 traballa ao ano de graduarse, acadando o 70,5% aos tres anos. Tres anos despois de titularse, máis da metade das persoas que traballan (56,4%) teñen un emprego acorde coa súa formación universitaria. [+ info](#)

Outras novas

- En decembro/2025 publicouse unha modificación da Lei 6/2024 de Estatística de Galicia que modifica a composición do Consello Galego de Estatística coa finalidade de reforzar a súa pluralidade e independencia, consistente en incrementar o número de persoas de relevancia profesional no campo da estatística de tres a oito.
- En xaneiro/2026 publicouse o Decreto 128/2025, do 29 de decembro, polo que se aproba o Programa estatístico anual da Comunidade Autónoma de Galicia para o ano 2026. Neste programa inclúense as actuacións estatísticas que se levarán a cabo polo IGE, os órganos superiores da Administración xeral e as entidades públicas instrumentais do sector público autonómico.
- O 21 de xaneiro reuniuse o Consello Reitor do IGE para aprobar a modificación da relación de postos de traballo do IGE.

Traballos defendidos no SUG

Traballos Fin de Mestrado

Aplicación de modelos de series temporais para a previsión de demanda e prezos nunha empresa de transporte internacional

Diego Airas Fernández
Directores: María José Ginzo Villamayor e Jose Ameijeiras Alonso
Titira: Lucía Tajuelo López (Trucksters) USC. Modalidade prácticas

Este Traballo Fin de Mestrado presenta unha metodoloxía para a predición da demanda loxística en Trucksters mediante técnicas de series temporais. O obxectivo é anticipar o volume de envíos nun mercado volátil e estacional.

A investigación desenvolve unha abordaxe comparativa entre a metodoloxía clásica ARIMA e os modelos Prophet para determinar cal captura con maior precisión a complexidade da demanda. A estratexia artículase mediante unha metodoloxía dual: contrástase a utilidade da segmentación estratéxica de clientes mediante clústeres e avalíase o rendemento dos modelos sobre a base de datos orixinal fronte a unha serie depurada de atípicos.

A análise demostra que o tratamento previo da información e a modelización desagregada reducen o ruído estatístico e mitigan o impacto das anomalías. Como resultado, obtense un marco predictivo máis robusto que optimiza a planificación operativa e a rendibilidade comercial de Trucksters.

Optimización de rutas de camión eléctrico: comparación de métodos exactos e heurísticos

Natalia Álvarez Tejero
Directores: Alejandro Saavedra Nieves e Laura Davila Pena
Titira: Lucía Tajuelo López (Trucksters) UDC. Modalidade prácticas

Os problemas de rutas de vehículos constitúen un campo amplamente estudado no ámbito da optimización matemática, que conta con diversas formulacións segundo as necesidades dos clientes. Ademais, existen múltiples métodos de resolución, xa que non resulta sinxelo abordar estes problemas de maneira directa.

Neste Traballo Fin de Mestrado estúdase o caso particu-

lar dos problemas de rutas de vehículos eléctricos formulado pola startup española Trucksters. Neste contexto, ademais das restricións habituais de capacidade, tempo ou distancia, engádense novas limitacións relacionadas coa carga da batería, o que implica revisar e adaptar os métodos tradicionais de resolución. Para iso, analízanse métodos exactos e diferentes heurísticas que incorporan estas novas restricións.

Finalmente, emprégase a linguaxe Python para obter solucións ao problema formulado, implementando un modelo exacto co solver SCIP e desenvolvendo as heurísticas do veciño máis próximo e 2-opt, permitindo comparar rendemento, calidade e escalabilidade.

Aplicación de técnicas estatísticas para a redución do tempo de cómputo en simulacións de sistemas fotovoltaicos

Adrián Blanco Aguiar
Director: Brais González Rodríguez
Titor: Miguel Sánchez de León Peque (ieco.io) USC. Modalidade prácticas

Neste traballo aplícanse diferentes técnicas estatísticas para reducir o tempo necesario para realizar simulacións de sistemas fotovoltaicos. Un exemplo destas técnicas é a utilización dun algoritmo de aprendizaxe automática non supervisada (*clustering*) para a agrupación dos datos de entrada co obxectivo de reducir a súa dimensionalidade. O obxectivo é reducir o máximo posible o tempo necesario para levar a cabo as simulacións, mantendo controlado o erro que se comete ao realizar as simplificacións.

Optimización matemática no ámbito da transición enerxética

Pablo Lagoa Loureiro
Directores: Julio González Díaz e Brais González Rodríguez
Titores: Raúl Baigorri Martínez e Marta Rodríguez Barreiro (CITMAga) USC. Modalidade prácticas

Este Traballo Fin de Mestrado, froito da colaboración entre CITMAga e Repsol, optimiza o aprovisionamento de materias BIO para biocombustibles e combustibles de

baixa pegada de carbono, baixo o marco das normativas europeas de transición enerxética. A investigación fundaméntase na modelización e integración dunha planta de biocombustibles no problema xeral de optimización.

O proxecto aborda a transición desde a formulación matemática de restricións cara á súa implementación técnica nun contorno industrial. Mediante unha simplificación funcional do modelo global, lévase a cabo unha análise profunda da planta sen a complexidade inabarcable do

sistema completo. Este desenvolvemento integra teoría de optimización con enxeñaría de datos, abrangendo código fonte, xestión de bases de datos e visualización de resultados en Power BI.

Finalmente, avalíanse as novas funcionalidades e as limitacións técnicas presentes, propoñendo alternativas e futuras liñas de desenvolvemento que aseguren a operatividade e eficacia da ferramenta no contexto actual de sostibilidade.

Traballos Fin de Grao

Métodos de Gradiente Estocástico

Martín Alonso Gamarra

Titor: Julio González Díaz

Titulación: Grao en Matemáticas. USC

Este traballo analiza os métodos de gradiente estocástico (SGD) como peza fundamental na resolución de problemas de optimización a grande escala, especialmente no ámbito da aprendizaxe automática. O descenso de gradiente tradicional (GD) presenta limitacións computacionais prohibitivas cando o volume de datos é elevado, xa que require procesar o conxunto completo de observacións en cada iteración. Como alternativa, o SGD utiliza estimacións ruidosas pero eficientes do gradiente a partir de observacións aleatorias.

Ao longo do traballo, estúdanse as condicións de regularidade necesarias para garantir a estabilidade do método, tales como a suavidade en esperanza e o control da varianza. Analízanse resultados de converxencia para funcións convexas, fortemente convexas e aquelas que satisfán a condición de Polyak-Łojasiewicz. Finalmente, explóranse variantes de alto impacto práctico, como o *Minibatch* SGD e o método de *Momentum*, xustificando como estas estratexias mitigan o ruído e melloran a dinámica de optimización en escenarios mal condicionados ou sobreparametrizados.

O problema do viaxante: formulación, resolución e aplicacións

Carmen Bello García

Titores: Ángel Manuel González Rueda e Brais González Rodríguez

Titulación: Grao en Matemáticas. USC

Neste traballo desenvólvese o estudo do problema do viaxante de comercio (TSP) mediante a descrición de tres formulacións alternativas: a de Dantzig-Fulkerson-Johnson (DFJ), a de Miller-Tucker-Zemlin (MTZ) e a formulación de fluxo de mercancía única (SCF); reparando nas características particulares de cada unha. Tamén se leva a cabo a análise de métodos de resolución mediante algoritmos aproximados para un caso particular, o TSP métrico. Previamente, realízase unha introdución aos problemas de optimización e á teoría de grafos, así

como aos problemas de fluxo en redes a custo mínimo, resaltando a súa baixa complexidade en comparación co problema do viaxante. Para finalizar, como complemento práctico ás propiedades expostas de xeito teórico, preséntase un estudo computacional sobre as tres formulacións abordadas, observando o seu comportamento numérico ante instancias de distintos tamaños.

Optimización semidefinida en algoritmos de optimización polinómica

Marta Ceán Romero

Titores: Julio González Díaz e Brais González Rodríguez

Titulación: Grao en Matemáticas. USC

A optimización polinómica é de alta relevancia na actualidade debido á ampla variedade de aplicacións e ámbitos nos que se emprega. Dado que o problema xeral de optimización polinómica global é NP-duro, o desenvolvemento de técnicas que permitan obter solucións óptimas en tempos razoables constitúe un reto de especial interese. Neste contexto, a optimización semidefinida adquiriu unha grande importancia. Caracterízase por empregar restricións convexas que facilitan a resolución eficiente dos problemas e, ademais, esta técnica permite obter boas aproximacións de restricións polinómicas non lineais, o que a converte nunha ferramenta clave neste ámbito.

Este traballo estrutúrase de maneira progresiva, avanzando desde os fundamentos teóricos da optimización polinómica e da optimización semidefinida ata un estudo computacional do impacto da técnica denominada cortes semidefinidos positivos no optimizador global RAPOSA. Ademais, preséntase un refinamento da implementación existente que permite mellorar o rendemento do optimizador.

Curvas ROC tempo-dependentes: Aplicación á prognose da colanxite biliar primaria

Marta González López

Titor: Juan Carlos Pardo Fernández

Titulación: Grao en Enxeñaría Biomédica. UVigo

As curvas ROC tempo-dependentes son unha ferramenta estatística que permite analizar a evolución da capacidade discriminativa dun biomarcador ao longo do tempo. Neste traballo empréganse co obxectivo de avaliar o valor prognóstico de catro biomarcadores (tempo de protrombina, albumina, bilirrubina e idade) na supervivencia de pacientes con colanxite biliar primaria (PBC). Os datos analizados proceden dun estudo da Mayo Clinic dispoñibles na base de datos “*pbcr*” do paquete de R *survival*. Mediante os estimadores implementados no paquete *timeROC*, xéranse as curvas ROC tempo-dependentes, tanto de maneira global como distinguindo por sexo e por estadio da enfermidade. Os resultados permiten determinar a evolución da capacidade discriminativa de cada biomarcador ao longo do tempo, identificar diferenzas por sexo e estadio da enfermidade e establecer cales presentan maior utilidade como indicadores da supervivencia a curto e longo prazo.

Introdución ós Problemas de Empaquetado Óptimo

José Manuel Otero Castro

Titor: Julio González Díaz
Titulación: Grao en Matemáticas. USC

Este traballo desenvolve unha introdución aos conceptos, modelos e algoritmos fundamentais da teoría de Problemas de Corte e Empaquetado. Para iso, establécese inicialmente un marco teórico común. Tras esta base, analízanse en profundidade tres problemas unidimensionais clásicos.

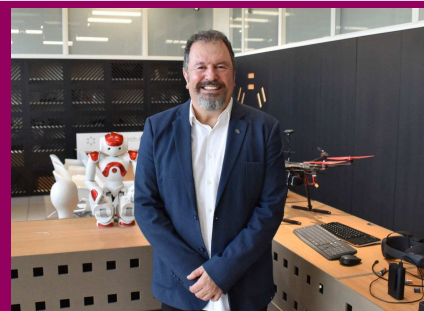
En primeiro lugar, estúdase o Problema da Mochila, un problema de maximización de valor para o que se presentan dous métodos de resolución exacta. A continuación, abórdase o Problema de *Bin Packing* (empaquetado de contedores), centrado na minimización do uso de recursos. Para este, expóñense os principais algoritmos heurísticos e introdúcese unha análise do seu rendemento. En terceiro lugar, preséntase o Problema de *Cutting Stock* (corte de material), de gran relevancia industrial, cuxa resolución se enfoca en obter a solución da relaxación lineal do problema mediante a técnica de xeración de columnas.

Finalmente, realízase unha breve incursión nos problemas bidimensionais, destacando como método de resolución a redución aos seus análogos unidimensionais.

Se dirixes/titorizas traballos académicos no eido da Estatística e da Investigación de Operacións e queres que se publique o seu resumo no **informest**, contacta connosco.

Coñecendo un socio: Salvador Naya Fernández

Salvador Naya Fernández é Catedrático da área de Estatística e Investigación de Operacións do Departamento de Matemáticas da Universidade da Coruña (UDC). Pertence ao grupo de investigación MODES (Grupo de Modelización, Optimización e Inferencia Estatística) e é investigador vinculado do Centro de Investigación en Tecnoloxías da Información e as Comunicacións (CITIC) da UDC.



Descríbenos o teu traballo

Como profesor de Estatística, imparto a materia de Estatística no Grao de Enxeñaría Naval e diversas materias noutros mestrados, como Control Estatístico da Calidade no Mestrado en Técnicas Estatísticas, no de Enxeñaría Téxtil e no Erasmus Mundus SEAS 4.0, coa Universidade Federico II de Nápoles. Como investigador do grupo MODES, estou implicado en varios proxectos. Ademais, dirixo tres teses de doutoramento e participo na elaboración de artigos que serven para difundir os nosos traballos.

Cantos anos levas traballando en estatística?

Levo xa máis de 30 anos na Universidade da Coruña, despois de pasar polo ensino medio durante 11 cursos en distintos institutos, exercendo como profesor de Matemáticas, sempre cun especial interese polos temas de estatística. Propiamente na investigación estatística empecei a traballar co inicio da miña tese de doutoramento, hai arredor de 25 anos.

Dende cando es socio da SGAPEIO?

Son socio dende a creación da SGAPEIO, participando activamente nos congresos e nas distintas actividades organizadas pola sociedade.

Cal foi o teu 1º Congreso da SGAPEIO?

O primeiro foi na Coruña, no ano 1993. Lembro especialmente a conferencia do profesor Miguel Bernardo, da Universidade de Valencia, sobre predicións electorais, que logo tomei como referencia para un traballo sobre as mesas testemuña en Galicia.

O tema do teu próximo proxecto ou artigo?

Teño recentemente solicitados dous novos proxectos ligados ao ámbito da enxeñaría naval e da defensa: un liderado por Navantia e outro cun consorcio europeo sobre datos sintéticos en defensa. Ademais, vou comezar coa dirección dunha nova tese coa empresa Hijos de Rivera, en co-dirección con Javier Tarrío e co apoio dunha axuda de doutoramento industrial do Ministerio de Ciencia.

Un curso que che gustaría facer?

Gustaríame facer un curso sobre Edge Computing, pola súa relevancia actual na análise de datos en tempo real e en contornas distribuídas.

Que libro de estatística ou I.O. recomendarías?

Recomendaría *An Introduction to Statistical Learning with Applications in R*, de Gareth James, Daniela Witten, Trevor Hastie e Robert Tibshirani; paréceme unha obra excelente para comprender métodos estatísticos modernos cun enfoque práctico.

Se non traballases no ámbito da estatística ou da I.O., que che gustaría ser?

Gustaríame ser diplomático; sempre me fascinou o reto de resolver problemas complexos, mediar entre posicións distintas e buscar puntos de encontro.

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

Cando escoito “estatística” penso na ciencia que analiza datos para comprender mellor a realidade, formular boas preguntas e dar respostas fundamentadas. É a disciplina que nos permite separar sinal de ruído, identificar patróns que non se aprecian a primeira ollada e cuantificar a incerteza coa que convivimos cada día.

Que é o que máis valoras do labor da SGAPEIO?

Valoro especialmente a creación dunha rede de sinerxías entre os grupos galegos de Estatística e I.O., así como as pontes con outras sociedades de España e Portugal. Tamén aprecio moito o concurso de deseño de experimentos para ensino secundaria.

En que cres que deberíamos mellorar?

Aínda que estou moi satisfeito co labor da SGAPEIO, penso que poderíamos aproveitar aínda máis as relacións que algúns socios temos con Portugal e, especialmente, con Hispanoamérica. Sería interesante fortalecer esa ponte atlántica para xerar proxectos conxuntos, intercambios e actividades compartidas.

O sudoku: cando un pasatempo se converte nun problema (de optimización)

por Alejandro Saavedra Nieves

Os sudokus, ese pasatempo tan popular que moitos resollen durante o tempo libre, teñen unha conexión profunda coa teoría da optimización e a teoría das redes. No fondo, un sudoku é un problema de asignación de valores, no que buscamos colocar números nunha cuadrícula 9x9 onde cada fila, cada columna e cada unha das nove subcuadrículas 3x3 contén números do 1 ao 9 sen repetirse. Detrás desa simplicidade, hai un desafío de optimización sorprendentemente complexo.

Un sudoku pode verse como buscar a mellor solución para un problema de asignación ou de distribución de recursos, onde as restricións (non repetir números en filas, columnas e subcuadrículas) limitan as opcións dispoñibles. Polo tanto, buscamos unha solución que maximice a “utilidade” de cada decisión tomada (colocar un número nunha cela) respectando as restricións.

Unha das formas máis eficaces de abordar un sudoku usando a optimización é utilizar algoritmos de *backtracking*, que exploran as solucións dunha forma sistemática. Este enfoque “proba” números nunha cela e, se a decisión tomada leva a un conflito máis adiante (por exemplo, dous números repetidos na mesma fila), retrocede e proba outra opción, continuando ata a solución correcta. Ademais, o sudoku tamén pode verse como un problema de optimización en redes, onde cada cela é un nodo e cada restrición (como non ter dous números iguais nunha fila) é unha aresta que conecta diferentes nodos, para os que os algoritmos xenéticos ou as redes neuronais tamén permiten resolver sudokus.

Aínda que o sudoku parece un simple xogo de lóxica, é un exemplo práctico de como os métodos de optimización poden ser aplicados na resolución de problemas, a simple vista, triviais. A beleza do sudoku non está só na súa resolución manual, senón tamén no feito de que reflicte perfectamente como as teorías matemáticas, a computación e a optimización poden ser utilizadas para abordar e resolver problemas en tempo real.

informest é unha publicación cuatrimestral da SGapeio. Podes atopar os números anteriores xa publicados en www.sgapeio.es/, no apartado PUBLICACIONES.

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.

Equipo editorial:

Laura Davila Pena e Marta Sestelo Pérez.
Paula Blanco Mosquera, M^a Esther Calvo Ocampo,
M^a Erika Francesch Domenech, Fabio José González
Eiras, Ana López Cheda, Leticia Lorenzo Picado,
Manuel Oviedo de la Fuente e Juan José Vidal Puga.

Revisión lingüística:

Trinidad Pazos Celis.

Coa colaboración de:

Nora M. Villanueva, José M^a Alonso, José Manuel Colmenero, Nieves Muñoz, Luis Ramil, Esther López Vizcaíno, Salvador Naya Fernández e Alejandro Saavedra Nieves.

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

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