(#10). EL DESCUBRIMIENTO DEL ARGÓN

[MONOTEMA] El descubrimiento del Argón es un artículo del economista de Virginia Tech, Aris Spanos, una de las personalidades más destacadas en su campo sin lugar a dudas. Este es uno de los artículos científicos que más me ha gustado en los últimos años porque explica, a través de la historia del descubrimiento del gas Argón, varios elementos fundamentales en la metodología científica que siempre debemos tener presentes los investigadores aplicados. En este post, voy a comentar detalladamente este artículo, con el fin de exponer sus puntos más relevantes e ilustrar de la manera más sencilla posible cómo Spanos propone proceder a la hora de analizar datos, basándose en la perspectiva que él mismo, junto a la filósofa Deborah G. Mayo, también de Virgina Tech, llevan defendiendo desde hace dos décadas: Error-statistics, y que sorprendentemente no está diseminada de la forma que quizá debiera. Intentaré emplear un lenguaje lo más cercano posible, aunque ello haga quizá que se pierda cierto rigor conceptual, pero creo que merece la pena hacerlo así para que este post se entienda mejor.

Aprender de los datos

Uno de los aspectos sobre los que más se discute en metodología es acerca de la aproximación deductiva frente a la inductiva. Básicamente la distinción entre estas dos aproximaciones se sustenta en teorizar primero y analizar después (deductiva), frente a analizar los datos y teorizar después (inductiva). Por ejemplo, si tengo una teoría sobre el movimiento de los planetas, debo poner a prueba esa teoría con las observaciones empíricas con el fin de contrastar si los datos son consistentes con la teoría. En este caso, opero de manera deductiva. Para generalizar esta teoría debo replicarla con múltiples observaciones en múltiples experimentos. Si esos datos no son consistentes con la teoría, entonces debo aprender de ellos, realizar más experimentos para buscar el porqué de esas inconsistencias y, si es necesario, replantear la teoría. Aquí se está produciendo un

proceso de inducción, en el que los datos nos “hablan” para que confirmemos o modifiquemos el conocimiento existente. En realidad, muchos vemos este proceso de una forma similar a como lo hacía el filósofo C. S. Pierce, donde deducción e inducción son parte secuencial del mismo mecanismo de generación de conocimiento, precedido de una etapa de abducción o generación de las teorías o hipótesis en base al cuerpo de conocimiento existente. Esa secuencia de abducción->deducción–>inducción se realimenta constantemente. En definitiva, se trata de poner a prueba las teorías con los datos empíricos, y si los datos me dicen que no son compatibles con la teoría, entonces debo buscar explicaciones alternativas que sean a su vez confirmadas por los datos (en sucesivas replicaciones), y que puedan, de este modo, generalizarse. El cómo se generan esas ideas/teorías/hipótesis puede fundamentarse en otros datos empíricos, teorías alternativas, etc.

Las discrepancias empíricas

Spanos comienza su artículo explicando cómo Lord Rayleigh y su colaborador Sir William Ramsay, en la última década del siglo XIX realizaron una serie de experimentos que permitieron encontrar una discrepancia en la medida de la densidad del gas nitrógeno producido por dos procedimientos diferentes. Rayleigh intentaba medir el peso atómico del nitrógeno a través de conocer su densidad. Para ello lo extrajo del aire, el cual se creía entonces que estaba compuesto de oxígeno, nitrógeno, dióxido de carbono y vapor de agua, con trazas de amoniaco, y empleó el conocido método del amoniaco. Rayleigh estaba dispuesto a publicar sus resultados, pero inteligentemente pensó que debía utilizar otros métodos distintos de medición que, considerando siempre el error experimental, dieran el mismo resultado. Rayleigh empleó en este segundo experimento un procedimiento más tradicional; tratar aire con cobre caliente. Para su sorpresa, el peso del nitrógeno obtenido era 1/1000 superior que el obtenido por el procedimiento anterior.

Aquí ya tenemos dos importantes elementos a destacar que son esenciales metodológicamente hablando. Primero, Rayleigh implementó un procedimiento de triangulación metodológica, en el cual, si existen

dos métodos válidos para estudiar un fenómeno, los resultados del primero deben coincidir estadísticamente con los del segundo. Al fin y al cabo esa triangulación es una replicación que, en lugar de hacerse al aplicar el mismo método varias veces, se ejecuta empleando diferentes métodos una vez. Segundo, Rayleigh evaluó el tamaño del efecto de la discrepancia como grande, a pesar de que en realidad es un número muy pequeño (una milésima). Esto nos recuerda que es esencial valorar los tamaño de efecto (las discrepancias sustantivas de las hipótesis nulas) de manera cualitativa en función del contexto y problema de investigación que se esté estudiando.

Como esa discrepancia era relevante, Rayleigh realizó otro experimento con otro procedimiento; esta vez no empleó aire, sino oxígeno puro para producir nitrógeno, es decir, una producción "química" del nitrógeno. La discrepancia fue en esta ocasión 5 veces mayor. Llegados a este punto Rayleigh se planteó ya muy seriamente que los datos obtenidos hasta ahora eran inconsistentes con la teoría imperante sobre la composición del aire atmosférico. Pero Rayleigh decidió seguir investigando con el fin de dar robustez a sus resultados.

El siguiente paso fue de nuevo replicar sus resultados empleando 4 métodos distintos para obtener nitrógeno del aire y otros 4 métodos distintos para obtenerlo químicamente. Para cada uno de esos 8 métodos empleó diversas muestras (de tamaño entre 1 y 4), cuyas mediciones promedió. De nuevo tenemos que detenernos aquí, ya que lo que hizo Rayleigh fue equivalente a obtener diferentes mediciones de un mismo fenómeno con el fin de disminuir el error experimental, o más concretamente, mejorar la fiabilidad de las mediciones. Es decir, si no hay ningún error sistemático de medición, cuantas más observaciones individuales se promedien (se incremente el tamaño de la muestra), la media muestral (que es un estimador insesgado de la media poblacional) tendrá menor varianza. Por tanto, se obtiene una estimación más precisa de la media poblacional. Ciertamente, Rayleigh empleó muestras extremadamente pequeñas (como he dicho, entre 1 y 4 observaciones por cada método), lo que sería una posible limitación de su procedimiento. De hecho, en alguna de sus observaciones la diferencia entre sus

mediciones era mayor que esa milésima que él consideraba un tamaño de efecto importante. De este modo (y esto no lo comenta Spanos), Rayleigh quizá debería haber incrementado el tamaño de sus muestras con el fin de aumentar la precisión de los valores medios estimados de cada uno de sus 8 métodos. Conviene matizar también que muestras muy pequeñas podrían ser suficientes para observaciones muy homogéneas, es decir, con dispersiones muy bajas.

Finalmente, y tras el proceso de obtención de los valores medios de los 8 métodos, Rayleigh volvió a agregarlos (4 en el grupo de "aire" y 4 en el grupo de "químico), obteniendo un valor medio para el método del aire y un valor medio para el método químico. La resta de ambos promedios fue de 0.010515, es decir, una diferencia en peso de unos 11 miligramos, lo que llevó al investigador a concluir que, dado que el nitrógeno obtenido químicamente y el obtenido a través del aire atmosférico diferían de ese modo, debía haber una razón desconocida que explicara esa divergencia. Rayleigh basó esa conclusión en que las diferencias intra-métodos eran prácticamente despreciables, y esos 11 miligramos debían ser reflejo de una diferencia real entre la obtención de nitrógeno por ambos métodos. Como vemos, aquí sí que Rayleigh valoró las diferencias intra-método, es decir, las ocho diferentes observaciones (promedios) de cada método diferente empleado, que eran menores de 1 milésima.

Rayleigh obtuvo unos datos que eran inconsistentes con la teoría establecida. El investigador, en aras de dar robustez a sus resultados, replicó sus experimentos empleando diferentes métodos. Una vez que se aseguró (en la medida de lo posible) que sus datos eran fiables y que no había errores sistemáticos, entonces ejecutó un último paso: planteó dos hipótesis que podrían explicar esa divergencia; (1) que el nitrógeno atmosférico fuera demasiado pesado debido a la eliminación imperfecta de oxígeno, y; (2) que el nitrógeno obtenido por el método químico fuera menos pesado debido a la contaminación con otros gases, como el hidrógeno. Ambas hipótesis fueron descartadas dado el conocimiento existente, como bien explica Spanos. Rayleigh incluso llegó a realizar más experimentos en aras de eliminar otras hipótesis similares relativas a la imperfección de los

métodos empleados, es decir, trató de llevar al extremo las hipótesis que podrían explicar esa discrepancia asociadas a la mala ejecución de los experimentos, pero no encontró ninguna evidencia de que así fuera.

Nuevas hipótesis

Llegados a este punto, el investigador tenía una discrepancia de la teoría existente que necesitaba de nuevas hipótesis para ser explicada. Aquí pasamos de nuevo al proceso de abducción, donde Rayleigh se planteó de nuevo dos alternativas; (1) que el nitrógeno atmosférico fuera demasiado pesado porque efectivamente contenía un gas más pesado, y; (2) que el nitrógeno químico fuera demasiado ligero porque contenía un gas más ligero. Fue aquí cuando pidió el consejo del eminente químico William Ramsay para tratar de arrojar luz sobre la primera hipótesis, ya que la segunda de ellas estaba prácticamente descartada por sus experimentos. Ambos investigadores realizaron de nuevo otra serie de experimentos, con otros métodos diferentes a los de los experimentos originales, y de nuevo se mostró esa discrepancia, esta vez de 0.011167, prácticamente idéntica a la anterior. De este modo, y ya con mucho más fundamento, se atrevieron a exponer que había un nuevo gas en el aire, al que llamaron Argón, del griego “inactivo”, ya que este nuevo elemento debía ser químicamente inerte.

Como tantas veces ocurre en la ciencia, este descubrimiento fue visto con desconfianza por otros reputados investigadores, entre ellos Dimitri Mendeleev, el creador de la tabla periódica porque, entre otras razones, no cuadraba dentro de su clasificación de elementos (posteriormente Moseley solucionaría este problema). Como comenta Spanos, no fue hasta el descubrimiento de otros gases nobles (helio, neón, kriptón, xenón y radón) entre 1895 y 1900 cuando la comunidad científica en pleno aceptó los resultados de Ramsay y Rayleigh.

El punto de vista estadístico

Hay que recordar que los procedimientos estadísticos de contraste de hipótesis e inferencia, tal y como hoy los conocemos, no aparecieron hasta la década de 1930. Pero, como bien indica Spanos, lo interesante de los procedimientos descritos anteriormente es que mimetizan, en cierta forma, los estándares que hoy tenemos en estadística aplicada. De este modo, hoy en día seguiríamos los siguientes pasos:

1. Establecer un modelo estadístico

La fiabilidad de una inferencia depende de la validez de las asunciones probabilísticas, es decir, del modelo estadístico. Esto es muy importante recalcarlo porque un modelo estadístico no es sólo establecer una relación entre variables sino también una serie de asunciones que deben cumplirse, y que son parte intrínseca del modelo. Pensad en la cantidad de artículos que leemos en revistas académicas que son incluso de gran nivel y que sólo inciden en ver la significación de la relación entre variables sin atender al cumplimiento de las asunciones.

El modelo estadístico es pues un conjunto de asunciones probabilísticas. Este tipo de asunciones condicionan el siguiente punto.

2. Formalizar el test para detectar la discrepancia

Como Rayleigh quería estudiar si verdaderamente existía una diferencia sustantiva entre los dos métodos de medición (nitrógeno atmosférico frente al nitrógeno químico), esto llevaría actualmente a la formalización de un test de diferencia de medias empleando la prueba T de Student, si las asunciones sobre los datos son las de normalidad, independencia y varianzas constantes (Spanos añade también la asunción de medias constantes dentro de cada método, pero esta es una asunción particular para este diseño de Rayleigh). Así, se establece una hipótesis nula en la que esa diferencia es cero, y una alternativa en la que esa diferencia sea mayor que cero. Es importante resaltar aquí que la hipótesis alternativa es unidireccional, o “de una cola”, ya que firmemente estamos proponiendo que el peso de nitrógeno atmosférico es mayor por la presencia del gas inerte.

Si nuestras asunciones son otras entonces el test elegido será consistente con esas asunciones. Por ejemplo, la no existencia de normalidad hace que el test T no deba implementarse (y aquí, ciertamente, podríamos discutir acerca de la cantidad de artículos en los últimos años que ha habido sobre la robustez de este test ante el incumplimiento de las asunciones, pero lo dejaremos para otro post).

3. Establecer una discrepancia estadística

Llegados a este punto aplicamos el test y vemos si existe significación estadística, es decir, si se rechaza la hipótesis nula. Y, como de nuevo excelentemente bien matiza Spanos, si un resultado es estadísticamente significativo no quiere decir necesariamente que exista una diferencia sustantiva o que exista un valor particular de la hipótesis alternativa (falacia del rechazo), mientras que si el resultado no es estadísticamente significativo tampoco significa necesariamente que no haya evidencia de que realmente lo sea (falacia de la aceptación), ya que en este último caso el test puede no ser lo suficientemente sensible como para detectar una discrepancia sustantiva (suele ocurrir con tamaños de muestra pequeños o con dispersiones muy grandes).

4. Establecer una discrepancia sustantiva

Spanos, en base al trabajo de Mayo (1996), propone establecer un análisis de severidad, que no es más que una especie de criba que el test elegido tiene que pasar. En realidad es la hipótesis la que tiene que pasar un test severo empleando los datos existentes, es decir, cuando se afirma que una hipótesis es rechazada o aceptada en función de los datos empíricos, se debe realizar esa afirmación siempre que esté fundamentada en un test severo. Conviene asimismo recordar que, desde el punto de vista frecuentista, se está analizando P(D|H0), es decir, se estima la probabilidad de los datos (D) asumiendo que la hipótesis nula (H0) es cierta. Cuando esa probabilidad es menor que un determinado umbral alfa (usualmente 0.05), entonces la probabilidad de los datos es tan baja que se suele rechazar H0, admitiéndose una discrepancia (tamaño de efecto). Es importante recalcar que ese "p-valor" no es equivalente a la probabilidad de H0, sino a la de los datos (D). Nótese que la perspectiva bayesiana computa P(H0|D), lo que resulta mucho más intuitivo.

La cuestión relevante es cuál es la discrepancia de la hipótesis nula garantizada por el test elegido, dados los datos de la muestra. Así, un test provee evidencia para una hipótesis en la medida en que los datos no sólo concuerdan con la hipótesis sino, además, ese resultado habría sido aún más probable si la hipótesis fuera falsa. Dicho de otro modo, si encuentro evidencia a favor de mi hipótesis (lo que se

llama “resultado no significativo”), entonces debería analizar la probabilidad de obtener ese mismo resultado si la hipótesis fuera falsa. Es muy importante señalar que la severidad puede ser alta y la potencia baja. Esto ocurre porque la potencia se evalúa en un punto de corte de la distribución del estadístico independientemente de cuál sea el valor muestral obtenido. Mayo y Spanos (2006) lo explican perfectamente. Esto no quiere decir que el análisis de potencia haya que obviarse, ni mucho menos, pero reportar la severidad en conjunción con la potencia nos puede dar una visión mucho más completa de la medida en que la hipótesis nula se acepta pasando un test severo. En cualquier caso, este es un punto un tanto polémico en el empleo del test de severidad.

De más interés si cabe resulta el caso en el que la hipótesis nula sea rechazada; el test de severidad se refiere a la máxima discrepancia garantizada por los datos (tamaño de efecto “lambda”), es decir, una vez fijado un valor de severidad (que puede ser, por ejemplo, 0.95), puedo conocer el máximo valor del tamaño de efecto garantizado por los datos. Por tanto, teniendo simplemente la estimación muestral y la hipótesis nula, la severidad me indica una forma de evaluar el tamaño del efecto.

En el caso de los experimentos de Rayleigh, esas 0.01 unidades de discrepancia corresponden con un nivel de severidad de 0.85 (bastante alto), lo que indica que la hipótesis alternativa de que el nitrógeno atmosférico pesaba 0.01 unidades más que el nitrógeno químico pasa un test de severidad con valor de 0.85, es decir, esa discrepancia que para Rayleigh era sustantiva (tamaño de efecto importante) está garantizada por los datos empíricos.

Fiabilidad de la inferencia y adecuación estadística

Pero, como he dicho, aquí no acaba el trabajo, ahora hay que garantizar la fiabilidad de la inferencia a través de la verificación del cumplimiento de las asunciones del modelo. Spanos llama a este paso la adecuación estadística del modelo. En palabras llanas: Hay que evaluar la validez de las asunciones.

La forma en la que Spanos propone realizar esta validación es a través

de procedimientos formales (tests) e informales (análisis de gráficos) empleando los llamados test de mala especificación, basados en los residuos y en los datos brutos. Recordemos que, de manera general, los residuos reflejan las discrepancias entre las predicciones del modelo ajustado y los valores nominales. Es decir, los residuos nos comunican la bondad de ajuste del modelo estadístico. Para el caso específico de los experimentos de Rayleigh:

1. La asunción de normalidad se cumple tras emplear el test de Shapiro-Wilks.

2. La asunción de independencia se cumple tras emplear el test de rachas, que recordemos es un test no paramétrico.

3. La asunción de homogeniedad de las medias dentro de cada método se cumple tras aplicar ANOVAs para cada uno de los dos grupos.

4. La asunción de homogeniedad de varianzas no se cumple tras aplicar el test F.

Spanos comenta, sin embargo, que aplicando el test de Welch, que es una forma de corregir el test T cuando las varianzas de ambos grupos no son iguales, los resultados son muy similares.

Por tanto, el proceso de análisis de datos concluye aquí, verificando el cumplimiento de las asunciones (en el caso de la homogeneidad de varianzas modificando el test), por lo que los resultados pueden escribirse e interpretarse con garantías.

Conclusión

El excepcional investigador Aris Spanos nos explica fundamentos básicos de estadística aplicada a través de la revisión de uno de los descubrimientos más importantes del siglo XIX en química; el gas Argón. Es admirable cómo Rayleigh siguió de manera intuitiva un proceso metodológico que no se formalizaría hasta 30 años después, y que le permitió sustentar su descubrimiento en base a los datos empíricos obtenidos.

Este artículo nos hace reflexionar sobre el dinamismo de los procesos

de abducción, deducción e inducción, y nos conmueve cuando nos damos cuenta de cómo los investigadores antaño cuidaban y mimaban sus estudios hasta asegurarse de que sus resultados eran válidos. La diferencia de cómo se opera en muchas ocasiones en la ciencia actual (especialmente en ciencias sociales) es patente, donde la prisa, la falta de replicación, y las meteduras de pata estadísticas son constantes. Casi ninguno de nosotros estamos al margen de esta falta de rigor, que conste, todos podemos equivocarnos alguna vez y la presión por publicar es muy alta. La diferencia entre unos investigadores y otros radica precisamente en que algunos tratan de buscar mejorar día a día, aunque ello haga que vayan más despacio a la hora de publicar, y ser honestos con los datos con el perjuicio que eso supone también de cara a la publicación. Otros investigadores, sin embargo, no se preocupan por ello, cogen unos datos, los tabulan y les dan al botón de análisis "a ver qué pasa", y si ven un p-valor que se ajuste a lo que quieren pues ya está…y a publicar. A estos últimos investigadores, bajo mi punto de vista, habría que, de vez en cuando, ponerles en el brete que el genial Nassim Taleb recordaba en "Antifragile": que apuesten su propio dinero a que sus resultados son válidos…o mejor aún, que al igual que los romanos hacían dormir a los ingenieros que construían puentes debajo de los mismos con toda su familia durante los primeros días tras la construcción, estos investigadores de "a ver qué pasa" durmieran con una espada de Damocles representando la validez de sus estudios. Entonces, seguramente, el mundo de la investigación social cambiaría radicalmente.

Spanos nos introduce, además, en el análisis de severidad, una forma de cribar la calidad del test para que no realicemos afirmaciones a la ligera sobre aceptación o rechazo, y nos recalca la importancia de los test de mala especificación para garantizar la idoneidad del modelo.

Muchos de vosotros pensaréis que, si para este caso simplísimo del análisis de diferencia de medias entre dos grupos hemos de realizar todo este laborioso proceso, cuando las hipótesis y los modelos se compliquen la situación será mucho más difícil. Y es cierto. Pero, afortunadamente, los investigadores aplicados contamos con

herramientas que nos pueden facilitar un poco el trabajo (test robustos, test no paramétricos, procedimientos de remuestreo, etc.). En cualquier caso, la filosofía de Spanos (basada en su trabajo conjunto con Deborah Mayo) es pertinente en todas las situaciones de análisis de datos.

El descubrimiento del Argón, una apasionante historia de metodología de investigación. Ojalá la profesora que tuve cuando empezaba en esto, en mis curso de doctorado de “Análisis de datos”, me hubiera explicado esta historia con este nivel de detalle, en lugar de decirme (a mí y a todos los demás doctorandos) textualmente: “Cuando le deis al botón de análisis en un Anova y una prueba T únicamente tenéis que mirar que el p-valor sea menor que 0.05”.

Todos los posts relacionados

(#424) TEORÍA DE PROBABILIDAD E INFERENCIA ESTADÍSTICA SEGÚN ARIS SPANOS (I)
(#422). JUGADORES ACB AL NIVEL DE NBA EN LA TEMPORADA 2018/19
(#421). LOS JUGADORES MÁS PRODUCTIVOS DE LA ACB 2018/19
(#420). LA PARADOJA DE CHEVALIER DE MERE Y LOS ESTUDIANTES DE MARKETING
(#412). LOS JUGADORES MÁS PRODUCTIVOS DE LA NBA 2018/19
(#411). LOS ROOKIES MÁS PRODUCTIVOS DE LA NBA 2018/19
(#410) ¿QUIÉN HA SIDO EL MEJOR ANOTADOR DE LA FASE REGULAR DE LA EUROLIGA 2018/19?
(#409). ALGUNAS CONSIDERACIONES EN LA ESTIMACIÓN DEL RENDIMIENTO POR MINUTO EN BALONCESTO (II)
(#408). ALGUNAS CONSIDERACIONES EN LA ESTIMACIÓN DEL RENDIMIENTO POR MINUTO EN BALONCESTO (I)
(#407). AVANCE DE LOS NOVATOS MÁS PRODUCTIVOS DE LA NBA 2018/19
(#398). AJUSTE DE FUNCIONES (II): CAPACIDAD EXPLICATIVA Y PARTICIÓN DE LA FUNCIÓN
(#397). AJUSTE DE FUNCIONES (I): SPLINES VS MÍNIMOS CUADRADOS
(#396). INTERPOLACIÓN (IV): MÉTODO DE SPLINE CÚBICO
(#395). INTERPOLACIÓN (III): MÉTODO DE DIFERENCIAS DIVIDIDAS
(#394). INTERPOLACIÓN (II): MÉTODO DE NEVILLE
(#393). INTERPOLACIÓN (I): POLINOMIO DE LAGRANGE
(#391). BÚSQUEDA DE SOLUCIONES (V): MÉTODO DE LA POSICIÓN FALSA
(#390). BÚSQUEDA DE SOLUCIONES (IV): MÉTODO DE LA SECANTE
(#389). BÚSQUEDA DE SOLUCIONES (III): METODO DE NEWTON
(#388). BÚSQUEDA DE SOLUCIONES (II): METODO DEL PUNTO FIJO
(#387). BÚSQUEDA DE SOLUCIONES (I): MÉTODO DE LA BISECCIÓN
(#384). MEJOR UN SÓLO ITEM QUE VARIOS PARA MEDIR ACTITUDES
(#381). TABLAS DE TUMORES TOTALES EN EL ESTUDIO DEL NTP SOBRE EFECTOS DE LA RADIACIÓN DE TELÉFONO MÓVIL
(#380). ÍNDICES APROXIMADOS FLEXIBLES EN ECUACIONES ESTRUCTURALES
(377). AVANZANDO EN LA COMPLEJIDAD; DOBLE DEPENDENCIA
(#376). CAMPOS DIRECCIONALES, COMENZANDO A REPRESENTAR LA COMPLEJIDAD
(#373). PREDICCIONES CON BUENAS APROXIMACIONES
(#366). MODELO DE HOMOGENEIDAD EN TABLAS DE 2X2
(#365). BÁSICOS DE ECUACIONES ESTRUCTURALES (V): MATRIZ DE DATOS BRUTOS
(#362). ¿HORMESIS ENTRE EL VOLUMEN DE ENTRENAMIENTO Y LA HIPERTROFIA?
(#358). BÁSICOS DE ECUACIONES ESTRUCTURALES (IV): REGRESIÓN LINEAL SIMPLE CON FIABILIDADES DIVERSAS
(#354). BÁSICOS DE ECUACIONES ESTRUCTURALES (III): COVARIANZA Y CORRELACIÓN ENTRE VARIABLES LATENTES
(#353). BÁSICOS DE ECUACIONES ESTRUCTURALES (II): VARIABLES LATENTES Y FIABILIDAD
(#352). TESTOSTERONA Y CONSUMO CONSPICUO EN HOMBRES
(#350). BÁSICOS DE ECUACIONES ESTRUCTURALES (I): COVARIANZAS Y DESVIACIONES SOBRE LA MEDIA
(#346). MODELO NO LINEAL EN LAS VARIABLES PERO LINEAL EN LOS PARÁMETROS CON STATA Y MAXIMA
(#337). EXPLORANDO LA PREDICCIÓN LINEAL EN MARKETING CON MAXIMA
(#336) EXPLORANDO EL DETERMINISMO Y EL AZAR CON C Y MAXIMA
(#335). EXPLORANDO EL COMPORTAMIENTO CAÓTICO CON MAXIMA
(#334). EXPLORANDO LA NO LINEALIDAD CON MAXIMA
(#331). LOS CONSTRUCTOS FORMATIVOS NO REFLEJAN ESTADOS PSICOLÓGICOS
(#323). EL ATAQUE CON MICROONDAS A LA EMBAJADA DE EEUU EN MOSCÚ DURANTE LA GUERRA FRÍA
(#314). REDUCIR EL SESGO EN ESTUDIOS SOBRE EFECTIVIDAD DE VACUNAS
(#313). EL SISTEMA PLANETARIO DE TRAPPIST-1
(#310). MÁS ALLÁ DEL REDUCCIONISMO EN BIOLOGÍA
(#307). LA CRISIS DE LA INVESTIGACIÓN EN PSICOLOGÍA
(#302). CONSTRUCTOS EN PSICOLOGÍA E IMPLICACIONES ESTADÍSTICAS
(#301). CENSURA EN EL LENGUAJE CIENTÍFICO
(#298). ¿CAMBIAR DE ENTRENADOR A MITAD DE TEMPORADA?
(#294). NEUROLOGÍA DE LA MALDAD: PISTAS SOBRE LAS RAÍCES DE LA MALDAD INDIVIDUAL
(#287). MEDIAS VERDADES Y OCULTACIÓN: ACEPTACIÓN DEL ENGAÑO DE CORPORACIONES
(#283). INTERPRETACIÓN FRECUENTISTA DE LA PROBABILIDAD: EL PAPEL DE LOS MODELOS
(#276). LA IRRACIONALIDAD, SCHRÖDINGER, Y EL EFECTO TÚNEL
(#265). DEFENDIENDO EL P-VALOR A TRAVÉS DE SU CORRECTA INTERPRETACIÓN
(#264). MÁS ALLÁ DE LOS ENSAYOS CONTROLADOS ALEATORIZADOS
(#263). CAUSALIDAD EN CASOS INDIVIDUALES; ALGUNAS RECOMENDACIONES PARA SU ESTUDIO
(#260). PROPONEN SER 10 VECES MÁS EXIGENTES CON EL P-VALOR
(#257). ERRORES DE CODIFICACIÓN DE EMBOLIAS PULMONARES AL USAR EL ICD-10
(#252). LA PREGABALINA NO SE ASOCIA A MALFORMACIONES EN EL EMBARAZO
(#243). LOS DEPORTISTAS DE ÉLITE VIVEN MÁS, PERO CON MATICES
(#232). LA OTRA HISTORIA DE NIKE
(#191). LA TRIADA OSCURA, LOS ANTICONCEPTIVOS Y LA PREFERENCIA SEXUAL
(#185). VENTA DE ENTRADAS EN LA NBA Y POTENCIAL DE LOS EQUIPOS
(#163). CRECEN LAS DUDAS SOBRE PLS (PARTIAL LEAST SQUARES)
(#159). ENTENDER LA NATURALEZA DE LOS DATOS Y DE LOS TEST ESTADÍSTICOS
(#128). INDICADORES FORMATIVOS EN ECUACIONES ESTRUCTURALES
(#127). AJUSTE EXACTO EN ECUACIONES ESTRUCTURALES
(#104). CORRUPCIÓN Y DESIGUALDAD ENTRE LOS MUY RICOS
(#100). DESCREÍDOS DE LA EPIDEMIOLOGÍA
(#90). REDES EN PSICOPATOLOGÍA Y EN MARKETING
(#76). EL 85-90% DE LOS NUEVOS MEDICAMENTOS SON INNECESARIOS
(#33). CLUSTERS DE CÁNCER PEDIÁTRICO
(#31). TRÁFICO Y CÁNCER DE PULMÓN
(#26). FOTOVOLTAICAS HIPERGEOMÉTRICAS
(#24). ROMPER EL CRISTAL
(#22). COMPLEJIDAD DE LOS FENÓMENOS SOCIALES
(#20). DECISIONES HEURÍSTICAS
(#17). RADICALES DE LAS VACUNAS
(#13). 24755 CACAS EN EL PARQUE
(#12). ESTADÍSTICAS SIN ESTADÍSTICA
(#10). EL DESCUBRIMIENTO DEL ARGÓN
(#6). ¿VIVIMOS MÁS Y CON MEJOR SALUD?
(#5). ¿ES LA ECONOMÍA UNA CIENCIA?
(#3). HEZONJA Y LAS PREDICCIONES ESTADÍSTICAS
(#1) ¿CÓMO JUGARÁ MARIO HEZONJA SU PRÓXIMO PARTIDO?