



LA ESPADA DE DAMOCLES DE LA ESTADÍSTICA

¿Existe una espada de Damocles en la Estadística?. Expondremos distintas razones, y siempre que se opte por potenciar el cálculo numérico sin acompañarlo con el razonamiento estadístico, la respuesta no puede dejar de ser afirmativa. La decisión de optimizar la rapidez de cálculo y el ahorro de tiempo derivado, junto a facilitar una herramienta técnica y favorecer su difusión, conlleva que la espada de Damocles esté también sobre nuestras cabezas.

En última instancia es una decisión humana aventurarse en el uso de una tecnología sin los conocimientos necesarios. Quizá también aquellos que participan en la formación estadística, no hemos sido muy eficaces a la hora de explicar que una cosa es la instrucción estadística, necesaria para todos, y otra, muy distinta, que de ésta se derive la capacitación para realizar análisis estadísticos. También es posible que algunos docentes hayan actuado más como mercaderes de la Estadística que como difusores del razonamiento estadístico. Eso sería cierto si, durante el proceso de enseñanza, se ha puesto más énfasis en cómo se pulsan botones que en por qué se realizan las acciones en el entorno de una aplicación informática, asemejándose más a una venta de estadística que a una formación sobre la misma.

Pero el problema es mucho mayor y con muchos más agentes incidiendo en el futuro de la Estadística. Las revistas científicas requieren análisis estadístico en los artículos y, por descontado, la mayoría de revistas no son de contenido estadístico ni sus autores tiene como campo de trabajo la Estadística; pero se nota un apreciable incremento en la cantidad de análisis estadísticos que se mencionan en la publicaciones científicas (Smith, 1996).

Una buena parte de docentes universitarios transmiten a sus colegas la necesidad de utilizar técnicas estadísticas, aunque la mayoría de los alumnos dedican a su actividad profesional a cuestiones escasamente relacionadas con la Estadística. Su función profesional no será llevar a cabo investigaciones o estudios ni para el desarrollo científico ni para obtener nuevo conocimiento.

También se exige con mayor énfasis a las administraciones públicas que justifiquen sus acciones y eso se traduce en una extensa cantidad de resultados estadísticos. Las empresas privadas destinadas al análisis de datos insisten en la



importancia de extraer información y las empresas cuantifican sus procesos para maximizar beneficio y reducir costos. No olvidemos que algunas empresas han formado a sus operarios en el control estadístico de procesos, implicando a los mismos en la mejora del sistema productivo.

Todas estas fuerzas, entre otras, colaboran en transmitir la idea según la cual la estadística debe ser conocida por todos. Y, como parece potenciarse el saber cómo se hacen las cosas más que el comprender aquello que se realiza, resulta difícil exigir a las personas que se limiten sólo a un conocimiento teórico e intelectualmente enriquecedor. Por tanto, conocer es aplicar. En consecuencia, y como la cantidad de datos disponibles crece sin cesar, la mayoría de personas que tengan posibilidad y necesidad de utilizar técnicas estadísticas recurrirán a las aplicaciones informáticas. Pero, y para un uso riguroso y responsable del análisis estadístico en la investigación y la docencia, la utilización de los ordenadores y las aplicaciones informáticas no dejan de constituir una espada de Damocles sobre nuestras cabezas. Al igual que en la metáfora, ordenadores y las aplicaciones informáticas pueden o no caer, ello dependerá de que seamos conocedores de los posibles peligros que de su utilización pueden derivarse y de que sepamos evitarlos.

La primera razón a la cual haremos referencia entre aquellos problemas derivados de las aplicaciones informáticas es que favorecen la automatización del análisis estadístico. Podemos establecer un símil entre las actuales aplicaciones informáticas para el análisis de datos y las máquinas. No sería nada complejo instruir a una persona sin conocimientos de estadística para que se responsabilizara del control estadístico de procesos indicándole que cuando ocurra uno de entre una serie de hechos avisará a un técnico. No se advierte una diferencia entre el manejo de una máquina reducido a una serie de rutinas y el análisis estadístico mediante estas aplicaciones.

Una serie de acciones, se comprenda o no lo que implica, lleva a un resultado. ¿cuál es el problema? simplemente que no se requiere un especialista. Mientras un experto sigue un proceso de formación continuamente, además de los conocimientos que ya posee, en la mayoría de los casos no podemos esperar que aquellas personas, que operan sobre estas aplicaciones de forma rutinaria tengan un interés especial por mejorar sus conocimientos estadísticos. Por mucho que a quienes estamos en el campo de la estadística nos apasione esa temática, no podemos olvidar que otras personas contemplan esos análisis como un trabajo rutinario. En consecuencia, la automatización



no es una garantía para extraer toda información relevante de los datos con? la correcta utilización de las técnicas.

No abundaremos en la idea, pues se ha tratado ya, pero el uso masivo de las aplicaciones informáticas no permite poseer un mínimo de garantías sobre la calidad de los análisis realizados, una opinión con la cual no creemos que los diferenciamos en exceso de otros autores (Abraira, Cadalso, Gómez, Martín y Pita 2001). La tecnología estadística está, y su facilidad de manejo también, en una sociedad fuertemente competitiva la ventaja diferencial es fundamental y, por esa razón, quien pueda demostrar un dominio de la tecnología que otros no poseen tiene un valor añadido en su perfil profesional. La estrategia óptima consiste en disminuir la ventaja del competidor y, si es posible, obtener otra que no posea. En esta dinámica está la Estadística y sus aplicaciones informáticas, pues son dos aspectos del conocimiento que pueden permitir alcanzar un perfil profesional.

Consecuencia de la automatización y la masiva disusión y utilización, se detecta una creciente escasez de rigor estadístico. Se advierte la carencia de conocimientos teóricos y flagrantes errores. De nuevo la razón es la facilidad de acceso a ese tipo de tecnología en ausencia de firmes conocimientos estadísticos. En parte la responsabilidad de las empresas productoras y distribuidoras de estos programas, pues el diseño del producto se favorece que cualquier persona no experta pueda operar ese sistema, mientras los distribuidores enfatizan la facilidad con que se pueden obtener interesantes resultados estadísticos. Pero también existe un error en el enfoque docente, si no se consigue transmitir que conocer algo de Estadística no es suficiente para realizar análisis Estadísticos. Por supuesto utilizar esas aplicaciones e interpretar los resultados es una decisión libre, pero no sería desaconsejable una dosis de autocrítica.

Un problema muy importante que se deriva de las aplicaciones informáticas es que el usuario, en general, se ciñe a la oferta de técnicas estadísticas que proporciona un programa. Eso se debe seguramente a dos motivos. Primero, es muy probable que sólo disponga de una única aplicación informática para el análisis de datos. Segundo, muchos usuarios considerarán que un programa de renombre debe contener todas las técnicas estadísticas que se corresponden con los distintos problemas de investigación. La segunda razón es especialmente preocupante porque, otra vez, aparece la pertinaz idea de la existencia de una única vía entre el problema de investigación y la técnica estadística. Ninguna aplicación informática para el análisis de datos es una Biblioteca de



Alejaría en la cual el usuario hallará la totalidad de técnicas que puedan ser necesarias para realizar un análisis estadístico, y eso sin entrar en los diferentes coeficientes e índices que pudieran ser precisos en esa investigación.

No es muy habitual que en los resultados que proporcionan los programas se advierta al usuario del incumplimiento de los supuestos requeridos por las técnicas estadísticas, aunque su violación resulta en general, crítica (Hartmann,1974). Estas advertencias son más la excepción que la regla. Si el usuario posee conocimientos estadísticos suficientes, el problema queda mitigado. No afirmamos que se elimina el problema debido a que ningún estadístico es un profundo conocedor de la totalidad de técnicas estadísticas. Cuando quien realiza el análisis posee escasos conocimientos de estadística, las consecuencias en la estimación y la decisión estadística pueden resultar catastróficas.

Pero el punto crítico que queremos destacar en este momento es que resulta sorprendente esta aparente falta de interés por los supuestos en las aplicaciones informáticas. ¿Por qué estos programas no proporcionan mensajes escritos sobre los supuestos en los listados de resultados? ¿Cuál es la razón por la cual no se advierte al usuario sobre la violación de los supuestos?. No es tan complejo ni supone constreñir al analista de datos; al contrario, se agradecerían los mensajes de aviso. La explicación reside en los inicios de estas aplicaciones informáticas. Al principio eran programas cuyo mercado era única y exclusivamente el ámbito universitario, donde la mayoría de investigadores que los utilizaban tenían conocimientos suficientes para realizar los análisis estadísticos. Para acabar de complicar la situación, esas iniciales aplicaciones requerían conocer un complejo conjunto de comandos, siendo necesario elaborar algo parecido a un programa informático para obtener un análisis estadístico, siempre bajo entornos no gráficos, como eran por ejemplo, los sistemas operativos UNIX y DOS de la época. Incluso podemos remontarnos unos años antes y recordar que los análisis estadísticos se obtenían mediante una tarjeta perforada. De hecho eran unas pocas personas las que realizaban sus propios análisis estadísticos y los de los demás. En, salvo algunas técnicas muy simples y de fácil cálculo, el resto de investigadores no utilizaban complejas técnicas multivariantes.

Retomando la pregunta, la mayoría de aplicaciones que hoy en día se comercializan para el análisis de datos en entornos o interfaces gráficas de usuario arrancan de aquellos tiempos que parecen remotos. Un notable número de sus actuales



procedimientos se remontan a aquellos días. Ciertamente han sido retocadas las salidas de resultados, aprovechando las mejoradas posibilidades que ofrecen los entornos gráficos. Eso implica modificar el código informático, ya sea cambiando valores de variables de salida o mejorando la salida gráfica con nuevo código. Este es el tipo de modificación que ha primado, junto a desarrollar un código nuevo para dotar a las aplicaciones de procedimientos totalmente nuevos. Ahora bien, completar el viejo código o dotar al nuevo código informático de líneas de instrucciones para advertir al usuario sobre la violación de los supuestos no ha sido una prioridad. Es más, en muchos casos los resultados para poder realizar el análisis de los supuestos no se halla en las opciones por defecto de los procedimientos. Mientras el análisis estadístico sea realizado por expertos el problema es relativamente poco importante. Pero no olvidemos que para estas empresas informáticas los consumidores ya no son sólo expertos estadísticos y, como era de esperar, el escaso énfasis en los supuestos que arrastra desde los inicios tienen efectos descontrolados sobre la decisión estadística. En general, el programa proporciona el resultado se cumplan o no los supuestos.

Hemos definido que deben entenderse las aplicaciones informáticas para el análisis de datos como productos. Desde esta perspectiva, las empresas que desarrollan estos programas están obligadas a mantener procedimientos y algoritmos durante un tiempo para asegurar su rentabilidad no siendo evidente que estas compañías respondan con rapidez a los avances teóricos e incorporen con prontitud estos desarrollos en sus programas. Pero un problema más importante es que se mantienen algunos procedimientos anticuados o cuestionados, que seguramente se utilizan por la seguridad que transmite a muchos usuarios el aval que supone su inclusión en una aplicación informática de renombre punto aparte?

(Bartholomew, ¿1997). Dionisio el Tirano (S.IV a. de J.C.) quiso que uno de sus cortesanos, Damocles, comprendiera cual es la felicidad de un Rey pleno de grandeza. Ordenó a sus servidores que tratasen al cortesano como a él mismo e invitó a Damocles a un fastuoso festín, donde ocupó el lugar del Rey. Se encontraba el cortesano embriagado por tanta felicidad cuando, al alzar su vista, vio una pesada espada colgada sobre su cabeza y sujeta a la crin de un caballo. Todavía llena, cayó la copa de las manos de Damocles, entendiendo el cortesano cual era la felicidad de un tirano. La espada de Damocles es una alegoría que se refiere al peligro que amenaza a un hombre en medio de una aparente prosperidad. En este sentido tomamos la alegoría para referirnos a los



inconvenientes, desventajas y peligros de una utilización indiscriminada de las aplicaciones informáticas destinadas al análisis de datos.

Autor Joan Manuel Batista- Foguet

Análisis Estadístico con Aplicaciones Informáticas: ¿Ayuda o sustituye el Razonamiento?

----- O -----

El Mundo Estadístico

Normalmente se piensa que los aviones con cuatro motores son más seguros que los que sólo tienen dos. Esto es totalmente falso, como se indica en la pagina 14 de Air & Space, agosto y septiembre 1993: “cuantos menos motores, menor probabilidad de que alguno de ellos se estropee”. Por tanto, los aviones más seguros son los que tienen un solo motor e INCLUSO NINGUNO.

En realidad, volar en avión es muy seguro. La práctica totalidad de los fallecidos en accidentes aéreos han muerto al llegar al suelo.

Un hombre tenía miedo de tomar un avión por aquello de los secuestros aéreos. Mirando unas estadísticas, encontró que la probabilidad de que hubiese una bomba en su vuelo era de una entre mil, mientras que la probabilidad de que hubiese dos era de una entre 100000 por lo tanto, tomó el avión llevando él mismo una bomba.

Durante la Segunda Guerra Mundial, a alguien se le ocurrió la idea de mirar dónde habían sido tocados los aviones al volver de sus misiones y-reforzar esos puntos. Así que se empezaron a hacer estadísticas acerca de qué zonas del avión estaban más expuestas. Al analizar los resultados, se dieron cuenta de un pequeño detalle: lo que había que reforzar eran las zonas que recibían más balazos de los aviones que NO volvían de sus misiones.

