

EL SENTIDO DE LA ESTADÍSTICA: ESTADÍSTICA CON SENTIDO

José Rodríguez Avi, Universidad de Jaén, Jaén *jravi@ujaen.es*

RESUMEN.

El estudio de la estadística en secundaria y bachillerato presenta a veces ciertos problemas derivados de la densidad de los temarios y de la manera de abordar la materia, la cual es importante para los alumnos, no solo para los que pretenden proseguir estudios universitarios sino para cualquier tipo de alumnos debido a que es imprescindible una cierta formación estadística para movernos por el mundo actual.

En esta ponencia se proponen diferentes aspectos para contribuir a aumentar el interés del alumnado por la materia, así como ciertos consejos a la hora de facilitar su comprensión

Nivel educativo: Secundaria y Bachillerato

1. INTRODUCCIÓN.

Los conocimientos estadísticos se están haciendo cada vez más importantes para la población debido a la influencia que las técnicas estadísticas presentan en aspectos relevantes para nuestra vida. Por esa razón es deseable que los alumnos que cursan las enseñanzas obligatorias adquieran al menos un conocimiento, quizás rudimentario pero eficaz de los términos y conceptos estadísticos básicos.

Podemos decir que la estadística es a los datos lo que la medicina es a nuestra salud. De hecho, la estadística es la medicina de los datos y al igual que ella, a partir de los síntomas hace análisis y pruebas variadas, a partir de las pruebas establece un diagnóstico, a partir del diagnóstico, prescribe un tratamiento y si el tratamiento es correcto el paciente sana y si no,...EXITUS. Por tanto, al igual que en la medicina: es conveniente que la población tenga conocimientos previos para tratamiento y prevención de enfermedades leves (Medicina preventiva) y es muy peligrosa la automedicación: No vale que te tomes tú lo que alguien ha recomendado a la vecina. En caso de duda, acudir al "médico". Así es conveniente que el "paciente" no se automedique para enfermedades graves -problemas complejos- pero si debe tener conocimientos adecuados de medicina preventiva -problemas simples, leer un periódico- ya que si no puede tener dificultades -ser fácilmente manipulable- Pues bien, yo entiendo que ha de ser en los institutos donde se impartan esos conocimientos de estadística preventiva necesarios.

Desde un punto de vista curricular, y pensando en los alumnos que desean continuar realizando estudios superiores en la Universidad, la Estadística, en mayor o menor medida, será uno de los *huesos* con los que toparán, como puede verse en cualquier anuncio de clases particulares dirigida a universitarios. Eso es así tanto



XV CONGRESO DE ENSEÑANZA Y APRENDIZAJE DE LAS
MATEMÁTICAS: EL SENTIDO DE LAS MATEMÁTICAS.
MATEMÁTICAS CON SENTIDO



para los alumnos de Ciencias Sociales, los cuales tienen una mayor carga de estadística en el programa de Matemáticas aplicadas a las CCSS, como para los alumnos del científico tecnológico. De hecho, los alumnos que posteriormente cursen estudios de medicina o enfermería no volverán a ver el álgebra o la geometría, pero sí tendrán asignaturas de estadística.

Si nos centramos en los alumnos que desean seguir estudiando, la estadística puede proporcionar a los actuales bachilleres una interesante vía de acceder al mundo del trabajo. En nuestra sociedad de la información generamos muchos, muchísimos datos. Y ese exceso de información puede asfixiarnos si no sabemos manejarlo adecuadamente. Y ese es el valor añadido de los estadísticos, el poder organizar, resumir, interpretar y tomar decisiones a partir de esa nube de datos, el poder desbrozar el camino del conocimiento en el bosque impenetrable de los datos. Por tanto, el descubrirla a los alumnos de manera atractiva puede contribuir a que la tomen en consideración para su futuro profesional.

Pero además, el impartir conocimientos adecuados de estadística es necesario para los alumnos que no desean seguir sus estudios. La estadística nos rodea en las noticias, en radio prensa y televisión, en los anuncios, lo que comemos o cómo compramos. Es por eso necesario proporcionarles una base para que puedan comprender e incluso ser un poco críticos cada vez que en las noticias se diga eso de "*un grupo de científicos han descubierto que...*" seguido de algo que cause espanto al sentido común. Baste con recordar lo que ocurrió con el aceite de oliva y el pescado azul en los años ochenta, en donde se consideraban *científicamente* casi veneno para pasar a hora a considerarlos, también *científicamente*, lo más parecido al néctar y la ambrosía que el garzón de Ida suministraba a los dioses en el Olimpo, tanto que hasta se vende -mucho más cara, por supuesto- leche con sardina (las leches con ácidos grasos Omega 3) como el summun de la salubridad.

ESTADÍSTICA Y FUTURO DEL ALUMNADO

La Estadística puede ser presentada como una materia de especial interés para el alumnado que desea seguir estudiando en la Universidad, por dos razones fundamentales:

2. ESTADÍSTICA EN LA UNIVERSIDAD.

Asignaturas con contenido estadístico aparecen con frecuencia en múltiples títulos de grado que se imparten en diferentes áreas. Así, la Tabla 1 muestra el catálogo de titulaciones en donde el Departamento de Estadística e I. O. de la Universidad de Jaén imparte, al menos, una asignatura. A esto hay que añadir otros grados con asignaturas de contenido estadístico, tales como psicología o pedagogía, o la estadística que aparece en los planes de estudio de los grados de Medicina, Odontología, por ejemplo, sin olvidar los grados en Estadística que se ofertan en diferentes universidades españolas.

Tabla 1.- Grados con asignaturas de Estadística en la UJA

Centro	Plan de Estudios
F. CC. De la Salud	Grado en Enfermería

	Grado en Fisioterapia
F. CC. Experimentales	Grado en Biología
	Grado en Ciencias Ambientales
F. Trabajo Social	Grado en Trabajo Social
F. CC. SS. Y JJ.	Grado en Finanzas y Contabilidad
	Grado en Administración y dirección de empresas
	Grado en Gestión y administración pública
	Grado en Relaciones laborales y recursos humanos
	Grado en Turismo
	Grado en Estadística y Empresa
	Grado conjunto en Derecho y Administración y Dirección de Empresas
E. Politécnica Superior de Jaén	G. en Ingeniería de Organización industrial
	Grado en Ingeniería Electrónica
	Grado en Ingeniería Eléctrica
	Grado en Ingeniería Mecánica
	Grado en Ingeniería geomática y topográfica
	Grado en Ingeniería informática
E. Politécnica Superior de Linares	Grado en Ingeniería civil
	G. en Ingeniería de Recursos Energéticos
	Grado en Ingeniería de tecnologías mineras
	Grado en Ingeniería de tecnologías de telecomunicación
	Grado en Ingeniería química industrial
	Grado en Ingeniería mecánica
	Grado en Ingeniería eléctrica
F. Humanidades y CC. Educación	Grado en Arqueología

2.2 ESTADÍSTICA Y TRABAJO

Además, la Estadística puede ofrecerse a los alumnos como una posible salida profesional, atractiva para aquellos alumnos con inquietudes matemáticas pero que les guste más los aspectos prácticos. De esta manera, y según el Instituto Nacional de Estadística, la tasa de paro por sector de formación más baja en España en el 2013 corresponde al sector de Matemáticas y Estadística, con sólo un 6.9% (Periódico 5 días, 23 de mayo de 2014).

Esa actitud es compartida también en el extranjero. Así, por ejemplo, el Bureau of Labour Statistics de Estados Unidos (2013) publicó un estudio en su página web en el que se dice "*Employment of statisticians is projected to grow 27 percent from*

2012 to 2022, much faster than the average for all occupations. Growth is expected to result from more widespread use of statistical analysis to make informed business, healthcare, and policy decisions”, además de mencionar datos muy halagüeños sobre su salario en comparación de otras profesiones. Otras referencias interesantes se encuentran en la revista EIEconomista (16-10-2013) sobre la masiva demanda de especialistas en análisis de datos a nivel mundial, y la preocupación de no encontrar suficientes especialistas en la materia.

3. EN LA ENSEÑANZA: MENOS FÓRMULAS Y MÁS CONCEPTOS.

A la hora de explicar estadística muchas veces nos centramos en demasía en que el alumno aprenda fórmulas raras en vez de centrarnos en los conceptos. Vamos a ver algunos ejemplos:

3.1. Gráficos.

Los gráficos son herramientas muy útiles a la hora de explicar datos y se emplean profusamente en numerosas asignaturas, distintas de la Estadística. Así podemos ir desde gráficos simples, como los que se muestran en la Figura 1

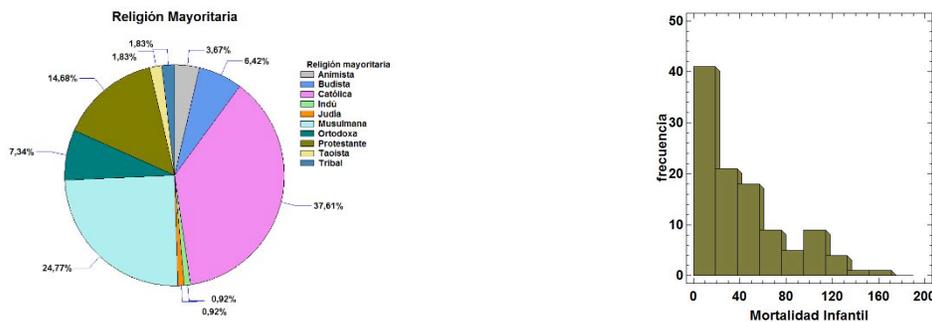


Figura 1. Gráficos de barras y sectores para datos sociodemográficos.

Hasta gráficos más complejos, como el gráfico de burbuja para representar conjuntamente tres variables que aparece en la Figura 2.

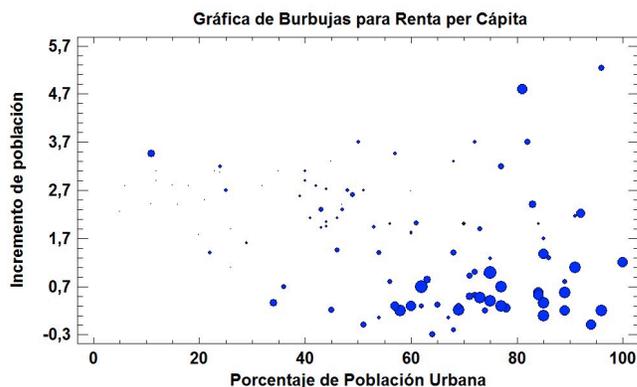


Figura 2. Gráfico de burbujas.

O el Gráfico por perfiles individuales que proporciona el procedimiento de las Caras de Chesnoff que se muestra en la Figura 4, que permite comparar individuos de manera que los más parecidos desde el punto de vista estadístico se corresponderán con aquellos de caras más parecidas.

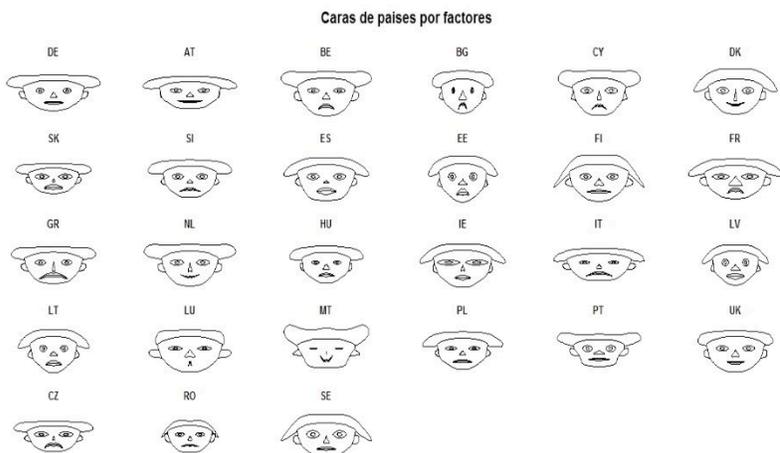


Figura 3. Caras de Chesnoff.

Sin embargo, hay que tener cuidado con la interpretación. Por regla general hay que huir de interpretar diferencias simplemente con un gráfico, dado que éstos pueden engañar. El procedimiento más adecuado es realizar primero pruebas estadísticas que demuestren la diferencia, y sólo entonces, utilizar el gráfico para interpretar o representar la diferencia detectada. Así, las Figura 4, sobre el nivel de ventas de una empresa muestra que en el primer caso las tres empresas van más o menos igual, mientras que el segundo destaca sobremanera los éxitos de la empresa A, que parece incluso triplicar a la empresa C.

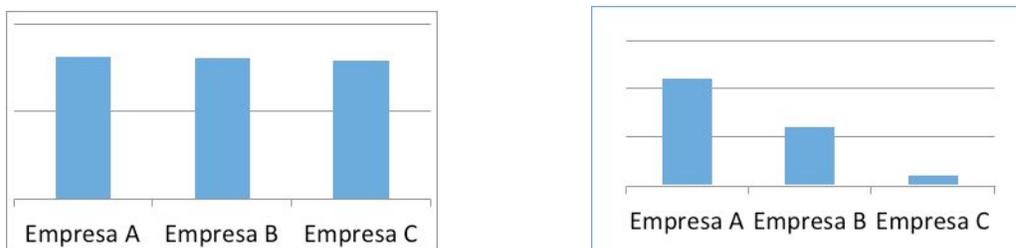


Figura 4. Gráficos de empresas I.

Sin embargo, esa diferencia no es real, ya que sólo se debe a una "trampa" en las escalas del eje de ordenadas, como puede verse en la figura 5:

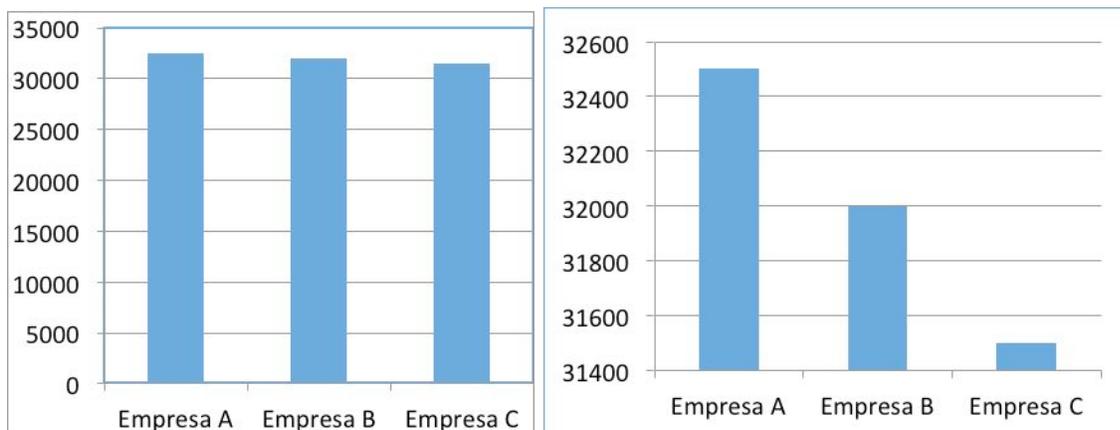


Figura 5. Gráficos de empresas II.

Este hecho es muy frecuente y también ocurre cuando hay datos atípicos que, de alguna manera, cambian la escala. Sólo en determinados casos, en los que el gráfico lleva asociado su propio estudio estadístico, un gráfico es directamente interpretable. Por ejemplo, el gráfico de Shewhart de control de un proceso que aparece en la Figura 6, en donde puede afirmarse que el proceso está fuera de control por desplazamiento de la media.

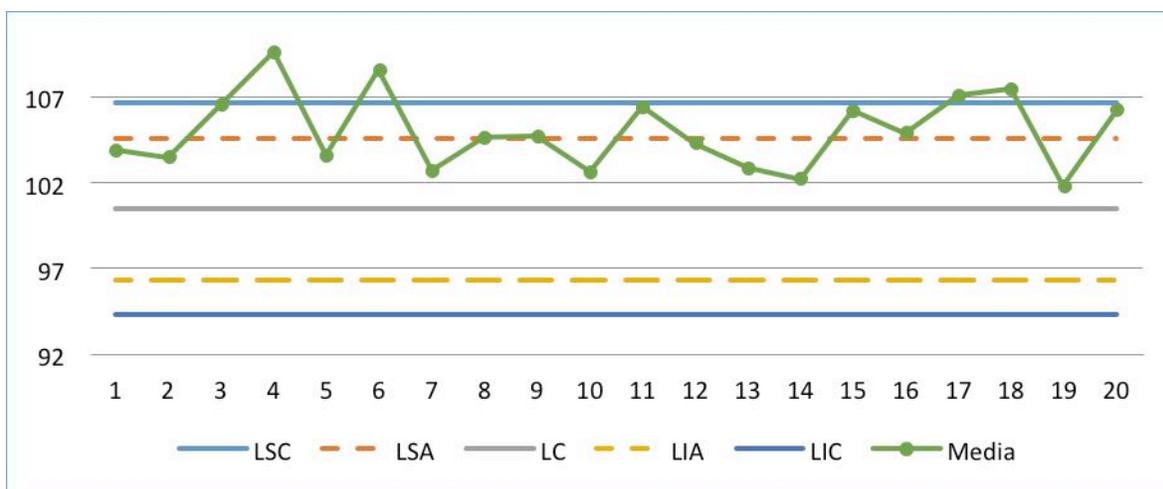


Figura 6. Gráfico de control para la media.

3.2 LA VARIANZA

1. A la hora de explicar la varianza nos centramos en la expresión

$$V(X) = \frac{1}{n} \sum_{i=1}^n x_i^2 - \bar{x}^2$$

Dado que es la expresión que se utiliza para el cálculo –a mano– de la varianza y a lo mejor se deja un poco arrinconada a la definición

$$V(X) = \frac{1}{n} \sum_{i=1}^n (x_i - \bar{x})^2$$

Cuando es ésta la que puede permitirnos interpretar su significado como medida de la incertidumbre de la distribución, entendida como los valores observados van a diferir de lo que debería salir si el fenómeno no fuese aleatorio (valor que representa la media). Así, cuanto menor sea la varianza, los valores observados están más cerca de la media y nuestra incertidumbre, por ejemplo a la hora de hacer un pronóstico, sería menor. De hecho sería preferible incluso prescindir del n y centrarnos en explicar el valor W (llamado así porque, en el caso multivariante sigue una distribución de Wishart)

$$W = \sum_{i=1}^n (x_i - \bar{x})^2$$

Que corresponde a la suma de cuadrados de las desviaciones a la media, es decir la medida de la incertidumbre presente en las observaciones de la variable x que no se puede conocer cuando lo que se utiliza como representante es la media. Aquí se observa como, si todos los valores son iguales, entonces $W = 0$, y ese valor será mayor cuanto mayor sea la distancia entre los valores observados y la media, como valor teórico. Esta introducción de W además resuelve un problema conceptual que los alumnos tiene después, a la hora de hablar de varianza (W/n) o de cuasivarianza ($W/(n-1)$ o $W/(n-k-1)$ en el modelo de regresión lineal múltiple con k variables explicativas). Incluso sería conveniente prestar atención a la expresión

$$\sum_{i=1}^n (x_i - a)^2 = W - n(\bar{x} - a)^2$$

dado que puede plantearse como un ejercicio de identidades notables y con la idea de que la media y varianza permiten descomponer la incertidumbre respecto a cualquier valor en dos partes: Una explicada por la media y otra que no se puede explicar o error residual. Esta es la cuestión clave en estadística, Variando los valores de referencia, a , obtenemos las expresiones base del ANOVA, la Regresión o incluso del Análisis Factorial y los modelos de ecuaciones estructurales.

3.3 INTERVALOS DE CONFIANZA

Los intervalos de confianza aparecen en Matemáticas aplicadas a las CC.SS. En este caso debe quedar claro la interpretación del IC: *No podemos hablar de que confiamos que la media (o cualquier parámetro) poblacional caiga dentro del intervalo*. El parámetro está, tanto si construimos un intervalo o no. La idea es que *confiamos* (con un nivel de confianza) en que el intervalo recubra al verdadero valor del parámetro.

Dado que hablamos de confianza, la manera de calcular un IC siempre es la misma:

$$IC(\theta) = \hat{\theta} \pm K_{\alpha} \sigma_{\hat{\theta}}$$

Es decir: el I.C. para un parámetro se construye teniendo en cuenta el estimador del parámetro, $\bar{\theta}$, su desviación típica, $\sigma_{\bar{\theta}}$, y una constante, K_{α} , asociada al nivel de confianza, independiente de la muestra y que se calculará en alguna distribución. De hecho, y de manera general, K suele valer 2 o 3. A partir de aquí se puede construir e interpretar cualquier IC, en concreto los que se emplean en estos niveles, sin más que considerar que

$$\bar{x} \rightarrow \mathcal{N}\left(\mu, \frac{\sigma}{\sqrt{n}}\right)$$

Un interesante ejemplo de aplicación de los IC está en comentar con los alumnos las horquillas de las encuestas electorales.

Podemos comparar los I. C para un parámetro con la caza. El parámetro es como un pajarillo que está libre en el espacio paramétrico. Para cazar al pajarillo, y si somos muy buenos tiradores (tipo John Wayne en las pelis del oeste), podemos ir con una bala (estimación puntual). Es lo mejor si acertamos, pero si nos desviamos, aunque sea unos milímetros del objetivo, el tiro será errado y no cazaremos al parámetro (digo, pajarillo). Por eso, otra opción es ir de caza con escopeta de perdigones. Al dispararlo, el cartucho se abrirá y ocupará más espacio, por lo que es más fácil que atinemos con nuestro objetivo, siempre y cuando el pajarillo –el parámetro– esté volando en el radio de dispersión. Si somos muy malos tiradores podremos recurrir a escopetas con cañones recortados –máximo rango, mínima precisión–, mientras que cuanto más confiemos en nuestra puntería podemos utilizar cartuchos con menor rango de dispersión (amplitud del intervalo).

Por último, una observación personal. Creo peligroso explicar al alumno contrastes de hipótesis e Intervalos de confianza como si fueran dos maneras similares de hacer lo mismo, ya que puede inducirlo a confusión. Por su propia filosofía son dos técnicas totalmente diferentes: El I.C. es un método de estimación en el que lo que nos importa es conocer el verdadero valor del parámetro, o, al menos, donde confiamos que esté –y como contrapartida, donde confiamos que no esté–. El contraste es una técnica de toma de decisiones en donde queremos tomar una decisión pero en la que no nos importa realmente cual es el verdadero valor del parámetro, sino sólo si podemos decidir a favor de la hipótesis alternativa o no tenemos más remedio que “tragarnos” la hipótesis nula.

4. EN LA HISTORIA: LO QUE LA ESTADÍSTICA HA HECHO POR LA HUMANIDAD.

Otra manera de hacer atractivo el estudio de la estadística puede encontrarse en explicar ejemplos reales en los que la estadística ha colaborado a “cambiar el mundo”. Algunas propuestas son las siguientes:

- **La Estadística y la Escritura:** La necesidad de contar y apuntar lo contado. Existen estudios en los que se reconoce que probablemente los primeros signos escritos fuesen números, para indicar el número de reses que se poseía o que se enviaban a otro lugar. Esto relaciona la Estadística con el Estado, por lo que puede datarse el origen de la Estadística en el mismo

momento en que la humanidad pasa de nómada a sedentaria y aparecen las primeras ciudades estado. De hecho, una herramienta que surge en esos momentos y aún perdura con extraordinaria importancia son los censos, de los que se pueden citar ejemplos como el Libro de los números en la Biblia, los primeros censos chinos o el Domesday Book, recopilación censal que ordenó Guillermo el conquistador tras la batalla de Hastings, en 1066. Así mismo puede ser de interés explicar la relación entre censo y censura a partir de la institución del censorado romano.

- **La Estadística y la Salud:** Si bien la relación entre ambas es continua – baste considerar los estudios del genoma humano a través del análisis cluster o la aprobación de un medicamento contra el cáncer- se puede incidir en momentos históricos en los que el uso de herramientas estadísticas introdujeron cambios radicales en la atención sanitaria. En este aspecto se puede hablar de Florence Nightingale, matemática y madre de la Enfermería moderna, y el papel que el análisis estadístico de los datos de mortalidad en la guerra de Crimea (1855-56) desempeñaron para la revolución en el cuidado de los pacientes y de las medidas higiénicas que aún se utilizan. O bien, el estudio de dispersión espacial que el médico John Snow desarrolló en 1854 para demostrar que el cólera se contagiaba a través de agua contaminada y para detectar el origen específico de la epidemia de cólera de Londres en ese año. Este caso puede servir también para ilustrar el modo de trabajar con valores outliers.
- La Estadística y la guerra. Aquí se puede destacar el papel que la Estadística desempeñó como una de las armas más poderosas con la que contaron los aliados para ganar la Segunda Guerra Mundial, tanto en su utilización para desentrañar los códigos alemanes y japoneses (en este caso fue lo que permitió la victoria decisiva de Midway, del 2 al 4 de junio del 44, en el peor momento de la flota americana y el mejor momento de la flota japonesa), como por el desarrollo de la investigación operativa y el control de calidad, lo que permitió la fabricación masiva, y repartida por todo Estados Unidos, de aviones, tanques Sherman y sobre todo, los

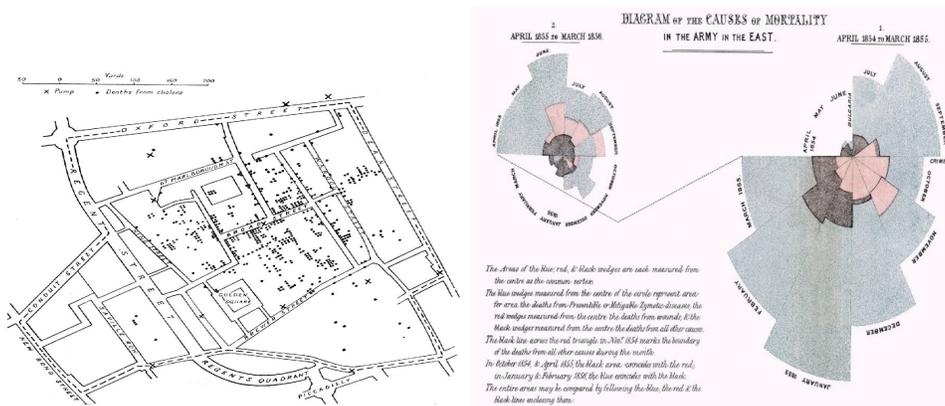


Figura 7. Gráfico de John Snow (izqda) y de Florence Nightingale (Dcha).

barcos Liberty que 40 años después, seguían navegando. Este es el origen del Control Estadístico de Calidad y de la primera norma estandar, la MIL-STD 105 D, de la que luego se desarrollaron todas las normas ISO (a nivel internacional) o de AENOR (España).

- La Estadística y el Cine. Desgraciadamente la estadística no ha tenido el mismo papel protagonista que otras titulaciones a la hora de ser ensalzadas en la pantalla –pequeña o grande-. Todos nos hemos emocionado e incluso identificado con abnegados médicos –desde el doctor Gannon hasta House, Anatomía de Grey o el doctor Vilches de Hospital central-. La biología también ha tenido sus momentos gloriosos, con el doctor Rodríguez de la fuente, los documentales de la dos, etc. Incluso ahora los físicos teóricos tienen su momento de gloria en Big Bang –aunque no creo que haga mucho proselitismo, la verdad- Sin embargo, los estadísticos no hemos tenido tanta suerte, si bien podemos hablar de series como Numbers, en la que uno de los protagonistas era un genio en estadística y que resolvía crímenes espeluznantes con modelos multivariantes, y que pasó sin enganchar demasiada audiencia (aunque métodos más exóticos e irreales como los de Imborrable, El Mentalista o Medium si han tenido mucho éxito). En pantalla grande podemos citar películas como Una Mente Maravillosa, sobre de John Forbes Nash Jr (interpretado por Rusell Crowe) uno de los pocos matemáticos ganadores de un premio Nobel. Especial atención, desde el punto de vista didáctico presenta la película Moneyball rompiendo las reglas. Trata sobre un entrenador de béisbol, Billy Beane, que revolucionó las Ligas Mayores mediante la recopilación y análisis de datos estadísticos que le permitió fichar un equipo nuevo, con jugadores específicamente adaptados a concretas situaciones de juego y mucho más barato que el resto dado que no iba buscando estrellas “galácticas”. Este equipo batió records de victorias consecutivas y, aunque no consiguió ganar el campeonato, su procedimiento fue seguido por otros equipos, como los Red Socks de Boston con mayor éxito. Este método se exportó luego a la NFL y más tarde a la NBA e incluso se está utilizando a la hora de determinar estrategias en tandas de penalties –recordemos Van Gaal y el cambio de portero en cuartos de final del pasado mundial.

Existen numerosas publicaciones, como Tamur et al (1992) o Llopis (1996) en donde aparecen casos similares a los mencionados, y que pueden ser utilizados a modo de ayuda para este tipo de motivaciones.

5. CONCLUSIÓN

La enseñanza de la estadística, sobre todo en lo referente a proporcionar al alumno de secundaria unas bases sólidas que despierten su espíritu crítico a la hora de darse cuenta de que “no es oro todo lo que reluce” en los mensajes que recibe, en especial si van soportados por expresiones como “grupo de científicos han



XV CONGRESO DE ENSEÑANZA Y APRENDIZAJE DE LAS
MATEMÁTICAS: EL SENTIDO DE LAS MATEMÁTICAS.
MATEMÁTICAS CON SENTIDO



descubierto que...” se ha convertido en una necesidad básica. Por eso es conveniente replantarse la manera en que abordamos esta materia ante ellos. El mundo que nos rodea, la sociedad de la información, se basa en datos, y la única herramienta científica para tratarlos es la estadística. A principios del siglo XX un escritor visionario, que incluso “viajó en el tiempo”, H. G. Wells escribió que

“Llegará el día en que el pensamiento estadístico será una condición tan necesaria para la convivencia eficiente como la capacidad de leer y escribir”.

Pues bien, podemos decir, casi seguramente, que ese día ha llegado ya.

REFERENCIAS.

¿Qué debo estudiar para sortear el paro?. Revista 5 días, 26-05-2014. Dirección wbb:

http://cincodias.com/cincodias/2014/05/23/economia/1400840398_735678.html

BUREAU OF LABOUR STATISTICS (United States Department of Labour) 2013. Occupational Outlook Handbook, Statisticians.

<http://www.bls.gov/ooh/math/statisticians.htm#tab-1>

GARCIA, N (2013). La gestión de datos supondrá más de 4 millones de empleos en 2015. ElEconomista.es, 16-10-2013. Dirección web:

<http://ecoaula.eleconomista.es/interstitial/volver/Nuezoct13/primer-empleo/noticias/5230158/10/13/La-gestion-de-datos-supondra-mas-de-4-millones-de-empleos-en-2015.html>

LLOPIS PÉREZ, J (1996). La Estadística: Una orquesta hecha instrumento. Ariel Ciencia.

TAMUR, J.M; MOSTELLER, W.H.; KRUSKAL, W.H. (1992). La Estadística. Una guía a lo desconocido. Alianza Editorial