ESTRATEGIAS DEL ANÁLISIS ESTADÍSTICO PARA EL TRATAMIENTO DE CUESTIONES DE DIDÁCTICA

Eduardo Lacasta Zabalza
Universidad Pública de Navarra (UPNA-NUP)

Resumen

El bagaje científico de los investigadores en didáctica de las matemáticas incluye en muchos casos algunos conocimientos de estadística matemática. Sin embargo, la distancia entre la estadística matemática y la estadística aplicada dificulta la aplicación de esos conocimientos. El objetivo de este seminario es facilitar la incorporación de herramientas estadísticas al trabajo de los investigadores en didáctica de las matemáticas.

A partir del establecimiento de unos conceptos básicos: modelo, variable, contingencia, significación de una medida, condiciones del rechazo o aceptación de una hipótesis, etc., se presentan de una manera razonada algunos de los principales métodos estadísticos aplicables.

En particular se describen algunas aportaciones que, dentro del análisis multivariante de datos, se han realizado desde el campo de la didáctica de las matemáticas, y más concretamente el análisis implicativo.

Se exponen con detalle la construcción de un sencillo test de hipótesis para tratar un caso puntual surgido en el tratamiento de una encuesta y el análisis de un resultado concreto, obtenido a través del análisis implicativo.

Abstract

The scientific background of researchers in Mathematics Education includes some knowledge of Mathematical Statistics. Nevertheless, the distance betwen Mathematical Statistics and Applied Statistics is a handicap in order to apply that knowledge. The purpose of this lecture is to provide researchers in Mathematics Education with tools form Applied Statistics.

Based upon some basic concepts such as model, variables, contingency, significance of a measure, conditions of rejection or acceptance of a hypothesis, etc., some of the main techniques in Applied Statistics will be presented in a reasoned sequence.

Contributions from the field of Mathematics Education to the Multivariate Analysis of Data, such as implicative analysis, will be specially described.

Furthermore, the lecture includes the detailed designing of a simple test of hypothesis for a case which appears in the treatement of a particular item and the analysis of a particular result obtained by Implicative Analysis.

1. La estadística en la Investigación en Didáctica de la Matemática

¿Es necesaria la estadística?

La didáctica de las matemáticas no dispone en general actualmente de respuestas para la mayor parte de las preguntas que se le plantean. Sin embargo, para fundamentarse científicamente, superando la mera opinión, tiene que poder formular hipótesis correspondientes a esas preguntas.

Dependiendo del problema de investigación y de los datos disponibles, será adecuado o no el uso de la estadística. Por tanto, no se puede pretender que sea imprescindible. Sin entrar en el fondo del debate metodológico, sí que pensamos que en didáctica de matemáticas se dan las condiciones necesarias para asegurar al máximo que, cuando se utilice, se haga con un control y un conocimiento de los métodos estadísticos utilizados.

¿Existen barreras también en didáctica de las matemáticas para una utilización ágil de la estadística?

Quien hace una investigación o más en concreto una tesis en didáctica de las matemáticas, tiene la posibilidad de recurrir a expertos que le proporcionen las técnicas estadísticas a aplicar para completar su trabajo. Pero creemos que, aún siendo a veces muy necesaria la ayuda de expertos en estadística aplicada –que es deseable que tengan una buena formación matemática–, el investigador debe responsabilizarse al menos del conocimiento y control de los métodos empleados.

A primera vista pudiera parecer que la formación matemática que normalmente posee el investigador en didáctica de las matemáticas –que ha podido estudiar en su carrera probabilidad y estadística matemática con un alto nivel de formalización– le ha de permitir abordar directamente el uso de los métodos estadísticos aplicados que le son de interés. Sin embargo, la experiencia nos muestra que hay una gran distancia entre el conocimiento teórico y su aplicación a problemas particulares y que normalmente el investigador no cuenta con una formación práctica en la aplicación de la estadística. Además, el crecimiento y la diversificación de los métodos estadísticos en los últimos años ha sido tan espectacular, que es difícil para el no especialista mantenerse al día en las nuevas técnicas y posibilidades del análisis de datos.

Por otro lado, la presentación clásica de la estadística inferencial, apoyada en una construcción previa de la teoría de la probabilidad –presente en muchos manuales de estadística aplicada, como el excelente libro de Amón (1990)– se aleja a veces de las bases intuitivas y de las fuentes históricas de la estadística, tan necesarias para la comprensión de los conceptos y para sus aplicaciones prácticas.

Sin embargo, como veremos, es posible dar una introducción intuitiva de las ideas básicas en inferencia, necesaria para la comprensión de los conceptos y para su aplicación práctica. Así pues, no es absolutamente necesario seguir una exposición formal para enseñar métodos estadísticos, existiendo buenos libros de estadística aplicada sin este enfoque.

Aportaciones desde la didáctica de las matemáticas

Régis Gras (1992) propone "poner en marcha un dispositivo particular de recogida y tratamiento de datos, susceptible de confirmar o invalidar las hipótesis correspondientes a las preguntas que se plantean en didáctica de las matemáticas y de avanzar conclusiones". Él mismo reconoce que es una estrategia muy ambiciosa y que no se puede aplicar efectivamente desde las primeras aproximaciones de los fenómenos a observar y de los que surgen las cuestiones, pero "acaba imponiéndose posteriormente si se quiere que las decisiones didácticas se apoyen en una estabilidad y una pertinencia de las respuestas, ganando así precisión, validez y capacidad de predecir".

Varias aportaciones en esta línea estratégica, como el análisis *a priori* de la contingencia ideado por Brousseau o los estudios de Pluvinage sobre los "cuestionarios con modalidades" han enriquecido el bagaje científico de la didáctica de las matemáticas. Un caso particular es el del análisis implicativo de Gras (1996) al que nos referiremos más adelante con cierto detenimiento.

Objetivo de este Seminario

Más que exponer los pros y contras de los distintos métodos estadísticos, nuestro objetivo es el de exponer de una manera sistemática un abanico metodológico lo más amplio posible. En el fondo, lo más importante es que el investigador tenga un criterio para decidir los métodos adecuados y conocer su alcance y sus límites.

Hemos dicho que la exposición de métodos estadísticos será "sistemática", pero no seguiremos la presentación formal –como se ha visto, no sólo por razones de espacio–, sino que, entre las diferentes aproximaciones posibles, nos apoyaremos en Brousseau (1993a).

2. Algunos principios generales del análisis estadístico

Con el fin de establecer un mínimo repertorio de conceptos necesarios para desarrollar nuestros objetivos, describiremos algunos de ellos, siguiendo el trabajo que acabamos de citar.

El primer objetivo del análisis estadístico es el de describir un conjunto de datos de manera *fiel y económica,* es decir, dar un resumen. Pero estas dos condiciones son bastante contradictorias, porque para ser fiel hay que dar toda la información y entonces no hay tal economía.

La contingencia

La contingencia es todo lo que es y podría no ser o podría ser de otra manera. Considerar un hecho como contingente supone pues tener en cuenta que podrían haber ocurrido otros hechos en su lugar y supone por tanto situarlo en un conjunto de hechos posibles, es decir, en una estructura. En la práctica, la contingencia está constituída por observaciones, datos recogidos, etc.

Las variables

Los datos recogidos tienen que estar estructurados a priori; es decir, tienen que reconocerse, antes de cualquier tratamiento, como un objeto matemático de un cierto tipo. Cada hecho que se produzca se considerará como un "valor" tomado en un cierto conjunto, llamado variable, en el que se han definido ciertas operaciones.

Las variables son pues estructuras y, por tanto, es importante identificar a priori la estructura pertinente más completa posible. Ejemplos de estructuras son las variables nominales, ordinales, de intervalo, de razón…

Los resúmenes o modelos

En la práctica, la contingencia, los datos, se representan mediante resúmenes. Los mejores resúmenes son aquéllos que, a igual longitud, permiten reconstituir la información inicial de la manera más precisa. Esta reconstitución exige que el receptor conozca el código utilizado por el autor del resumen. Por otra parte, será útil indicar la fidelidad del resumen mediante una medida de su parecido a la contingencia inicial.

Pero representar la contingencia mediante sistemas matemáticos es también el objeto de la modelización. Una misma contingencia puede modelizarse de distintas formas; por ejemplo, mediante una distribución de probabilidad, una serie de reúmenes o de características estadísticas. Los modelos habrá que escogerlos de modo que representen los datos lo mejor posible.

Para Brousseau (1993a), "la *estadística matemática* es el estudio de estos códigos, de las distintas modelizaciones de la contingencia, de sus propiedades y de los métodos de elección".

Propiedades de los modelos

En una primera aproximación, el modelo debe:

- Representar bien las observaciones (pertinencia).
- Constituir un resumen más simple que las observaciones (comunicabilidad).
- Permitir lo mejor posible la reproducción del conjunto de las observaciones (fidelidad).
- Permitir la comprensión de los datos, es decir, poder situarlos con relación a modelos familiares, universales y, por tanto, poder compararlos con otros modelos (inteligibilidad).
- Ser accesibles al control matemático (consistencia).

Ejemplos de modelos pueden ser una función, una ley, etc.

3. Primera estrategia estadística: situar una medida con relación a una distribución

Significación de una medida

Con frecuencia vemos en artículos de investigación (aún en revistas de reconocido prestigio) hacer comparaciones de valores, de medias o de porcentajes y establecer conclusiones sin tomar precauciones. También se puede ver cómo se adjudica el carácter de "significativa" a una diferencia de valores, sin que se describa claramente el criterio empleado. A lo más, se suele recurrir al tamaño de la muestra para relativizar la solidez de la argumentación, sin ir más allá de la intuición. En lo que sigue, haremos una aproximación de la significación estadística, que permite tomar una decisión sobre si las diferencias encontradas se pueden explicar fácilmente por la variabilidad aleatoria

del muestreo, o bien si es muy *raro* que esos valores se hayan obtenido debido a esa variabilidad (azar).
La significación estadística indica la *rareza* de un valor, o más bien la rareza de los valores mayores (o menores) que él; es decir, lo que indica si ese valor es grande o pequeño en las condiciones en las que se ha medido y bajo unas hipótesis determinadas (o sea, en relación con un cierto modelo). La rareza relativa de un valor será la proporción de valores mayores (o menores) que él en un cierto modelo. Habitualmente se expresa en procentaje y viene a ser la "significación" de un valor; es decir, en una perspectiva clásica, la probabilidad de obtener ese valor u otro mayor (alternativamente menor) debido a la variabilidad aleatoria, en las condiciones de medición dadas y bajo un modelo supuesto.
Evidentemente, para poder determinar la rareza relativa de un valor, la variable tiene que ser al menos de carácter ordinal. En la figura 1 hemos reproducido un histograma de frecuencias de una variable en porcentajes.

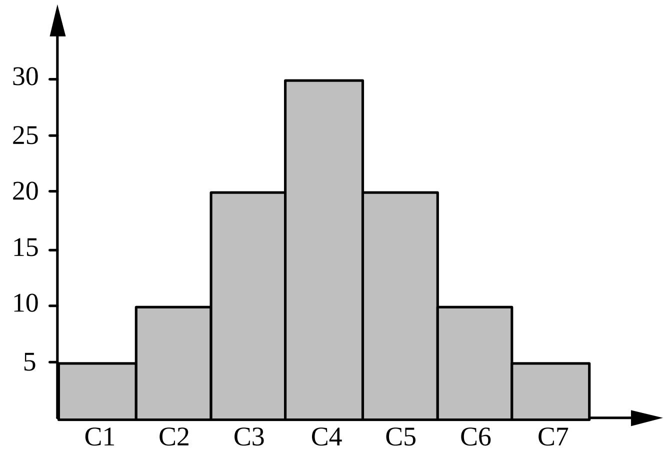

Fig. 1

Frecuencias (%)
En el caso del histograma de la figura 1, un valor que esté en la clase C1 es bastante raro, puesto que a lo sumo hay un 5% más pequeños que él; también un valor que estuviese en la clase C7 sería bastante raro, porque a lo sumo habría un 5% mayores que él. En ambos casos podemos medir la *rareza* de ese valor, diciendo que ese valor es *raro* o no se puede explicar por el azar, con una probabilidad del 95%. En psicología, por ejemplo, se puede rechazar un modelo que proporciona una distribución en la que el valor encontrado (la contingencia) es raro con una probabilidad del 95%, pero en medicina, las probabilidades habitualmente utilizadas son por lo menos del 99,9%.
Para determinar la significación de un valor o una diferencia de valores encontrados, existen en la mayoría de los casos pruebas o tests, que nos dicen cuál es la probabilidad de encontrar esos valores, para el tamaño de la muestra utilizado.

Los tests de hipótesis

Para explicar un hecho dado, se elige un modelo de distribución. La *hipótesis nula* consiste en la declaración "el hecho observado está de acuerdo con el modelo". En la literatura estadística existen numerosos tests de hipótesis aplicables a lo distintos casos que el investigador se pueda encontrar. En estos tests se calcula el valor de un estadístico de contraste cuya distribución es conocida y que nos permite rechazar o aceptar la hipótesis nula, según que dicho valor sea "raro" en esa distribución o no, con una probabilidad dada.

Condiciones del rechazo o de la aceptación de una hipótesis

El rechazo y la aceptación no constituyen una alternativa; no rechazar un modelo no quiere decir que se acepte. En efecto, puede ser que varios modelos estén lo suficientemente próximos de la contingencia como para no rechazar ninguno, pero al mismo tiempo sería absurdo aceptarlos todos con sus contradicciones. Por el contrario, no hay ningún inconveniente en rechazar varios modelos, ya sean compatibles o no entre ellos.

El conocimiento científico debe pues fundamentarse en el rechazo de modelos y no en su aceptación.

Los riesgos de rechazar o no un modelo

Desde el momento en que se acepta la idea de que la contingencia observada no es más que un ejemplar extraído de un conjunto con muchos ejemplares parecidos y que se fija una probabilidad arbitraria para la decisión, aparecen dos riesgos:

- el de rechazar un buen modelo del que la contingencia se aleja accidentalmente, que se llama *error de tipo I.*
- el de no rechazar un mal modelo, porque ocasionalmente está cerca de la contingencia, que es el *error de tipo II.*

Todo modelo que no es rechazado puede "aceptarse" de manera hipotética, aunque no se pueda establecer como verdadero. Hay pues que poner en pie una jerarquía de preferencias en el conjunto de los modelos, en función de su verosimilitud. Pero será preciso controlar estas decisiones y estos niveles de confianza arbitrarios bajo el control de una teoría para poder controlar esos riesgos; ese es el objeto de la teoría de la decisión.

Los procedimientos de la estadística inferencial

Régis Gras (1992) resume la manera de proceder en la estadística inferencial: a partir de un muestreo de una población, que hay que suponer que se ha efectuado mediante pruebas independientes y de la misma distribución:

- se ajusta a la variable o a dos variables […] una variable predictiva con una finalidad de predicción o de explicación;
- se decide validar o invalidar juicios e hipótesis de modelos mediante los llamados tests de hipótesis (esto último también en el caso de la estadística no paramétrica). »

Las condiciones que legitiman estos procedimientos propios de lo que hemos llamado "primera estrategia" son:

- presuponer en la mayoría de los casos hipótesis de normalidad de las variables, lo que es difícilmente verificable y restrictivo (no es el caso para los métodos no paramétricos);
- aceptar métodos largos y fastidiosos de cruce de variables 2 a 2;
- saber formular las hipótesis nulas y
- diluir la individualidad de los sujetos en el muestreo. »

Sin embargo, estas condiciones no se satisfacen fácilmente en el estado actual de la didáctica de las matemáticas. Parece indispensable poder recurrir a otros métodos distintos para sintetizar y estructurar los datos y poder identificar las variables, los factores implicados, sus relaciones, su jerarquía, etc.

El caso de la estadística no paramétrica

Dentro de la estadística clásica, las técnicas no paramétricas se prestan especialmente a tratar los datos de didáctica y en general de ciencias de la conducta, puesto que en ellas no se supone que los datos que se están analizando se hayan sacado de una población distribuída de una determinada manera; por ejemplo, de una población distribuída normalmente. Sidney Siegel (1990) hace una exposición amplia de los métodos no paramétricos, acompañada de ejemplos, en la que se analizan las exigencias de la estadística paramétrica y no paramétrica. Guy Brousseau (1993b) desarrolla en un volumen los tests no paramétricos más utilizables en nuestro caso, en la misma perspectiva que aquí hemos adoptado y con abundantes ejemplos, todos ellos referidos a la didáctica de las matemáticas.

La estadística no paramétrica dispone de tests aplicables a distintos casos (como los de Wilcoxon, Cochram, Kendall, etc.). Todos los paquetes estadísticos estándar (Statgraphics, SPSS, BMDP, etc.) tienen incorporados los cálculos no paramétricos.

Un test de hipótesis para un caso particular

Si no se dispone de tests ya confeccionados, o si no existen, se pueden crear pruebas *ad hoc*, que a veces pueden ser muy simples, como la que exponemos a continuación.

En el ejemplo correspondiente a la tabla 1, tomado de Lacasta (1995), se pasó un cuestionario a 23 profesores de instituto, sobre diversas cuestiones relacionadas con el papel de los gráficos cartesianos de funciones en la enseñanza. Se observó que en algunas preguntas abiertas, se repetían algunos tipos de respuesta, que en la tabla 1 se han codificado con las letras a, b, c, etc.

Nos encontramos por ejemplo con que el 71,4% de los profesores (15 profesores de los 21 que contestaron a esta pregunta) responden de la manera "a" (en concreto, declaran que "al acabar el COU, los alumnos utilizan correcta y suficientemente los gráficos cartesianos").

Con una muestra de 23 profesores de instituto, no podemos pretender establecer conclusiones para el conjunto de todos los profesores de instituto, pero sí que nos podemos plantear si ese conjunto de profesores se decanta claramente ("significativamente") por el "sí" o por el "no". En particular queremos saber si ese

71,4% puede haberse obtenido por azar o se aleja tanto del 50% que es demasiado raro que ocurra por azar y, por tanto, refleja más bien una tendencia clara por el "sí".

	Tipos de respuesta					
Profesores	a	b	c	d	e	f
p01	1	0	–	–	0	1
p02	0	1	1	0	0	0
p03	–	–	–	–	–	–
p04	0	1	1	1	0	1
p05	1	1	1	1	–	–
p06	0	1	–	–	–	–
p07	0	1	–	–	–	–
p08	0	0	1	1	0	1
p09	0	1	0	0	0	0
p10	1	0	1	0	0	0
p11	–	–	–	–	–	–
p12	1	0	1	1	1	1
p13	1	0	1	1	0	1
p14	1	1	–	–	–	–
p15	1	0	1	0	0	0
p16	1	0	1	0	1	0
p17	1	0	0	1	0	0
p18	1	0	0	1	1	0
p19	1	0	1	0	0	0
p20	1	0	1	1	0	0
p21	1	1	1	0	0	0
p22	1	1	1	0	–	–
p23	1	1	1	1	1	1
Total	15	10	14	9	4	6
Nº de respuestas	21	21	17	17	16	16
Porcentajes	71,4	47,6	82,4	52,9	25	37,5

Tabla 1

Tipos de respuestas de profesores a preguntas de un cuestionario

Naturalmente, si no hay una preferencia de estos 23 profesores por el "sí" o por el "no", el porcentaje de síes no tiene por qué ser exactamente del 50%; podrá ser un poco más (por ejemplo, 52,9%), un poco menos (37,5%)… ¿Cómo definir un criterio de decisión para saber si existe o no un acuerdo entre los 21 profesores que han contestado a la cuestión?

En este caso se tomó el proceso de decisión como una distribución binomial en la que existían n = 21 decisiones "sí" o "no", cuyas probabilidades, de no haber acuerdo, eran p = q = 0,50. La media de esta distribución es = np = 10,5 y la varianza es ^{2} = npq = 5,25. Aproximando la distribución a la normal, la probabilidad de obtener al menos 15 veces la misma respuesta es: P(X 15) = 0,025… Así pues, no hay más que un

2,5% de probabilidad de obtener este valor por azar, luego este resultado *es significativo* con una probabilidad del 95%.

Análogamente se puede calcular que el 37,5% de respuestas afirmativas (6 respuestas de 16 que contestan) del tipo"f" (que consiste en una declaración explícita de la superioridad o de la necesidad del gráfico en la enseñanza, del tipo "una imagen vale más que mil palabras" o similares) se produce por azar con una probabilidad: P(X   6) = 0,158… O sea, que hay aproximadamente hasta un 16% de probabilidad de obtener por azar este valor, luego este resultado *no es significativo* con una probabilidad del 95% y no podremos argumentar que exista una tendencia clara de esos profesores a hacer o a no hacer ese tipo de declaraciones.

La primera estrategia consiste en comparar la estadística observada (datos, modelo, medida de calidad) a la distribución de la cual ha sido extraída y, cuando sea preciso, decidir acerca de su interés en tanto que resumen, según la rareza relativa.

4. Segunda estrategia estadística: la elección del mejor modelo

Los métodos propios de esta segunda estrategia se suelen conocer como *análisis de datos,* aunque ésta es una expresión que tiene distintos significados para algunos autores. En ella se pueden llegar a englobar el análisis multivariante, el análisis exploratorio o el análisis cualitativo de datos (que a su vez puede ser también exploratorio e inferencial, multivariante o univariante); incluso hay autores que consideran la inferencia estadística dentro del análisis de datos. En lo que sigue, nosotros nos referiremos preferentemente al análisis multivariante.

Para Gras (1992) el análisis de datos (que sitúa en una "ruptura epistemológica" de la estadística) se distingue a la vez de la estadística inferencial y de la estadística descriptiva clásica. Su nacimiento histórico se debe a la formalización del álgebra lineal, de la geometría y de las probabilidades y por otro, al hecho de que los ordenadores permiten cálculos rápidos en estructuras complejas, sin tener que disminuir el tamaño de las tablas de datos a tratar, y permitiendo al mismo tiempo representaciones variadas de la información obtenida.

En el análisis de datos se asigna al conjunto de observaciones un conjunto de puntos, entre los que se define una distancia. Esta idea permite adoptar un punto de vista geométrico, lo que explica la potencia de este tipo de análisis. El establecimiento de una distancia abre pues nuevas posibilidades; entre otras, la de conseguir representaciones gráficas que permiten visualizar los datos y, como vemos a continuación, la de controlar la calidad de un modelo.

Las distancias de un modelo a la contingencia

Para saber si un modelo es un buen representante de la contingencia, la segunda estrategia estadística consiste en representar la calidad de la representación mediante un número, una medida de la calidad y, más concretamente, mediante una distancia entre el modelo y los datos.

Para un tipo de distancia dado (euclídea, euclídea relativa, ^{2}, etc.) el mejor modelo será aquél para el que se obtiene el valor más pequeño. Aunque muchas veces no es posible encontrar un modelo que sea con certeza óptimo, podemos decir en general que el resultado de un análisis estadístico será el mejor modelo obtenido en su categoría, junto con la distancia de este modelo a la contingencia. Esta distancia es indispensable para indicar la calidad del modelo propuesto.

El modelo se puede interpretar como la parte explicada de la contingencia y la distancia como una parte que no se ha tomado en cuenta; por ello se la llama distancia "residual".

¿Qué tipo de mejor modelo propone cada distancia en el estudio de la representación de un conjunto de números por un número?

La elección de las distancias

Según las condiciones, existen distancias más apropiadas que otras. Ya sea porque tienen propiedades matemáticas más interesantes, ya sea porque nos sitúan en un lenguaje familiar al analista o al destinatario del análisis.

Ejemplo: la distancia euclídea relativa

La distancia euclídea es la de nuestro espacio familiar y estamos acostumbrados a ella para hacer comparaciones, pero la dimensión del espacio es en este caso invariante. Sin embargo, en estadística, la dimensión de los vectores cambia con el tamaño del conjunto de los datos. Como la distancia euclídea es una suma de términos positivos, crece con el número de términos. Así pues se hace difícil comparar, por ejemplo, la calidad de un modelo para representar un conjunto de 10 valores, con la calidad de un modelo que representa un conjunto de 1.000 valores.

Por tanto, se prefiere utilizar la distancia euclídea relativa, en la que la suma de las diferencias cuadráticas está dividida por el tamaño de la muestra (es decir, la dimensión del espacio). En el caso particular de la distancia entre un número y una variable numérica (por ejemplo, un conjunto de números E = $\{x_1, x_2, x_3, \ldots x_i, \ldots, x_n\}$ y un número "a"), se pueden considerar esos "n" números como las componentes de un vector X = $(x_1, x_2, x_3, \ldots x_i, \ldots, x_n)$. Entonces, buscar un número que represente al vector X consiste en buscar un vector constante A = (a, a, a, …, a, …, a) que esté a la menor distancia posible de X. En estas condiciones, la distancia euclídea relativa es:

$$d\,(A, X) = \sqrt{\frac{1}{n}\left(\Sigma\,(a - x_i)^2\right)}$$

Lo mismo que con la distancia euclídea, el modelo que minimiza la distancia "d" es también la media aritmética. A la distancia euclídea relativa se le llama *desviación típica* y a su cuadrado *varianza.*

Ejemplo: la distancia de ^{2}

Supongamos ahora que se tiene como modelo un conjunto de frecuencias teóricas o esperadas, que proporcionan el vector E = $(E_1, E_2, E_3, \ldots E_i, \ldots, E_n)$ y que el conjunto de las frecuencias observadas empíricamente (es decir, en la contingencia) da el vector O = $(O_1, O_2, O_3, \ldots O_i, \ldots, O_n)$. Si se considera en cambio que la aproximación de "E_i" a "O_i" es tanto mejor cuanto pequeño sea $(E_i - O_i)^2$ respecto a "E_i", se obtiene la distancia[1] de ^{2} del vector "E" al vector "O":

$$d (E, O) = \frac{(O_i - E_i)^2}{E_i}$$

La modelización estadística

En forma simplificada, en la filosofía del análisis de datos, a la *contingencia* (lo que es, pero podría no ser, la variabilidad) se opone la *necesidad* (las tendencias). Idealmente las tendencias serían consecuencia lógica obligada de los hechos realizados.

Teóricamente, las tendencias se expresarían mediante reglas, leyes, teoremas, relaciones establecidas, deducciones, explicaciones, razonamientos, cálculos de previsión…

Para construir un resumen de un conjunto de hechos, parece razonable y económico identificar todos los que ya están determinados o casi (se les llama ligados) por otros (llamados libres), con el fin de no comunicar más que estos últimos, acompañados de la ley de producción o de dependencias de los primeros.

Cuando la estadística debe rendir cuenta de diversas variables, nos hemos planteado si la mejor estrategia es la de dar un modelo para cada una de ellas, incluso si es localmente el mejor modelo posible. Nos planteamos ahora la cuestión de saber si no sería mejor tener en cuenta las informaciones que algunas de las variables pueden dar sobre otras, para disminuir el tamaño de un modelo único más general o aumentar su inteligibilidad o su precisión.

Así pues, siempre hay que tratar de buscar modelos y clasificarlos. Siguiendo siempre la primera estrategia, un estudio estadístico consistiría entonces, a la vista de la contingencia, en proponer no sólo un objeto matemático, o varios, encargados de representarla –un resumen–, sino también un conjunto de relaciones organizadas en un verdadero modelo, acompañado de parámetros que los determinen y que sirva para discriminar en las observaciones lo que es explicado por una regularidad o, metafóricamente, por una "necesidad".

Presentación y tratamiento de los datos

La elección de la estructura a priori tiene su importancia, como se ha visto para determinar el tipo de distancias utilizables.

En un caso bastante general, un conjunto de datos puede presentarse en forma de una tabla (ver tabla 2) en la que las "n" líneas son los individuos observados y las "m" columnas representan variables cuyos valores son las observaciones hechas en relación con esos individuos. Cada elemento de esa matriz es por ejemplo un número que puede representar una medida, un rango, una frecuencia absoluta, una presencia o una ausencia. En sentido estricto, esta matriz es la matriz de contingencia si los valores que la constituyen son frecuencias absolutas.

---

[1] Hay que advertir que, de hecho, esta fórmula no satisface todos los axiomas de las distancias; es más bien un índice de distancia

	Variables				
Individuos	V_1	V_2	V_3	………..	V_m
01					
02					
03					
…	…	…	…		…
n					

Tabla 2

El análisis factorial (en las dos versiones que presentaremos brevemente más adelante) permite el tratamiento de un cierto número de variables y/o de individuos llamados *suplementarios,* que no entran a formar parte de la contingencia, pero cuya situación se puede determinar en relación a la representación de los individuos y variables contingentes. El análisis *a priori* antes citado consiste básicamente en el control de las dependencias de variables y/o individuos, independientemente de la contingencia. De esta manera se puede saber si las dependencias puestas de manifiesto por el análisis factorial son debidas a la contingencia o si estaban ya determinadas previamente. Un desarrollo de este análisis se encuentra en Brousseau y Lacasta (1995).

Entonces se puede considerar esa tabla ("n" individuos x "m" variables) de maneras muy diferentes, por ejemplo como:

a) Un conjunto amorfo de n x m puntos en R.

b) Un punto único de $R^{n \times m}$ (el espacio sobre R de dimensión n x m).

c) m vectores de R^n que representan las p variables.

d) n vectores de R^m que representan las n observaciones.

e) una aplicación de R^{n-1} en R

f) una aplicación vectorial de R^k en R^{m-k}, etc.

Aplicando nuestro primer principio a la modelización de esta tabla, siguiendo estas diferentes estructuras a priori, se engendran tipos de análisis estadísticos distintos.

Algunos métodos del análisis de datos

<u>El Análisis en Componentes Principales</u> (ACP)

En el caso general del punto d), los sujetos son puntos en el espacio R^m, cuyos ejes representan cada uno caracteres observados. El método consiste en buscar los planos principales determinados por los ejes de inercia y el centro de gravedad de la nube, para representarla según las mejores proyecciones, realizadas con la métrica euclídea.

Este análisis permite además precisar en qué medida las variables están correlacionadas o ligadas entre ellas. En la interpretación estadística los ejes principales se llamarán "factores principales".

En teoría no se excluye utilizar el ACP para estudiar el punto de vista c), pero las correlaciones entre los sujetos son más delicadas de interpretar.

La calidad de la representación se define en este método como el coseno del ángulo formado por una observación (un vector de R^m) con su proyección sobre el subespacio de los ejes factoriales que se consideren. Esto hace que las observaciones que están próximas al origen estén mal representadas.

<u>El Análisis Factorial de Correspondencias</u> (AFC)

Sigue el mismo principio, pero la distancia utilizada es la de ^{2}, lo que permite tratar mejor el caso de una matriz de incidencia (variables dicotómicas).

Además los sujetos y las variables se pueden situar en un mismo espacio. De esta manera es posible interpretar las proximidades de los sujetos entre ellos, la de las variables entre ellas y las de los sujetos con las variables. Este análisis trata a la vez los puntos de vista c) y d).

El AFC utiliza como datos tablas de contingencia (análisis de correspondencias simple) disyuntivas completas o de datos ordinales (análisis de correspondencias múltiples). Para el caso de tablas de contingencia se estudia la estructura de asociación entre dos variables cualitativas, permitiendo apreciar qué valores de cada una de las variables está asociado con cada valor particular de la otra variable.

En el caso del análisis de correspondencias múltiple, se propone dar una representación geométrica, en general en un espacio de dimensión grande, de una distribución conjunta de dos conjuntos: E (en general el conjunto de los sujetos) y V (en general el de variables o el de las modalidades de las variables).

Cada uno de estos conjuntos E y V admite un espacio de representación en el que los puntos son los "perfiles" de uno sobre el otro, traduciendo así las relaciones, las correspondencias entre E y V. Al estar estos espacios provistos de una distancia ^{2}, diremos que dos sujetos están tanto más próximos cuanto más parecido es su comportamiento en relación con las variables y, en particular, en relación con las variables menos frecuentes y por tanto las menos triviales.

El método analítico consiste entonces en entonces en extraer, a partir de los espacios de representación, subespacios de dimensión reducida, pero tales que, en ellos, la nube de puntos de E o de V se represente de manera óptima mediante sus diferentes proyecciones. A los ejes de estos subespacios corresponden factores discriminantes en los conjuntos E y V: el primero es el que más información suministra, el segundo algo menos, etc. El investigador en didáctica debe entonces determinar la significación de estos factores, mediante la observación de la oposición y la proximidad de las proyecciones de los puntos. Esta significación le servirá para analizar y después interpretar las informaciones más escondidas que se deducen de ella. Asimismo se interesará en las contribuciones de ciertos puntos a estos factores y, por ejemplo, en las posiciones relativas de los subgrupos de la población estudiada, entre otras cuestiones.

El Análisis Jerárquico de Similitudes según I. C. Lerman

Como en todos los métodos de clasificación, se busca constituir particiones cada vez menos finas, en el conjunto V de las variables, construídas de manera ascendente en árbol, con la ayuda de un criterio de similitud entre variables. El investigador en didáctica se interesa en este tipo de análisis que le permite estudiar e interpretar en términos de tipología y de proximidad (y de lejanía) decreciente de los núcleos de variables, constituídos significativamente a ciertos niveles del árbol, oponiéndose a otros núcleos en esos mismos niveles.

El criterio de similitud se expresa de la manera siguiente:

- Dos variables "a" y "b" están tanto más próximas, cuanto más importante es el número de sujetos que las satisfacen simultáneamente (A B) con relación a los cardinales de A y B y con relación a los que habría habido si las variables "a" y "b" no hubiesen estado ligadas a priori (ver figura 2). Esta proximidad se mide a través de la probabilidad de su inverosimilitud.

De esta manera, el índice definido entre las variables no está sesgado por las frecuencias absolutas y puede servir posteriormente para definir un índice de similitud entre clases de variables.

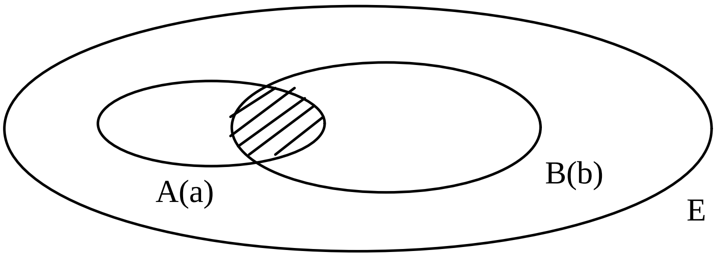

Figura 2

Así pues, para construir un árbol de clasificación, primeramente se reúnen en una clase al nivel más bajo las dos variables que más se parecen según este índice, después se unen otras dos variables o, como en el ejemplo de la figura 3, una variable a la clase ya formada, después otras variables o clases de variables, etc.

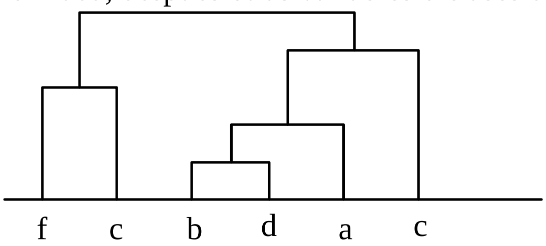

Figura 3

El Análisis Implicativo

Al contrario de todos los métodos anteriores, en los que la distancia o el índice de similitud son simétricos, el método implicativo no es simétrico. La problemática general que introdujo el método es la siguiente (ver figura 4), en el caso en el que las variables consideradas son binarias (un individuo satisface o no una variable):

Si A y B son las subpoblaciones de los sujetos que han satisfecho respectivamente las variables "a" y "b", ¿en qué medida se puede decir a entonces b, sin que la implicación sea debida al azar?

Si B A, la proposición se verifica, pero normalmente lo que se tiene es una intersección A $\bar{B}$ no vacía.
El índice de implicación mide, de una manera comparable al de similitud, hasta qué punto es la pequeñez de A $\bar{B}$ con relación a la independencia a priori y a las frecuencias absolutas observadas. Así pues, se dirá por ejemplo que, siendo X e Y dos partes aleatorias de E, que tienen respectivamente los mismos cardinales que A y B, es admisible con una probabilidad igual a 0,95 si y solamente si:

Prob [card (X $\bar{Y}$) < card (A $\bar{B}$)] < 0,05.

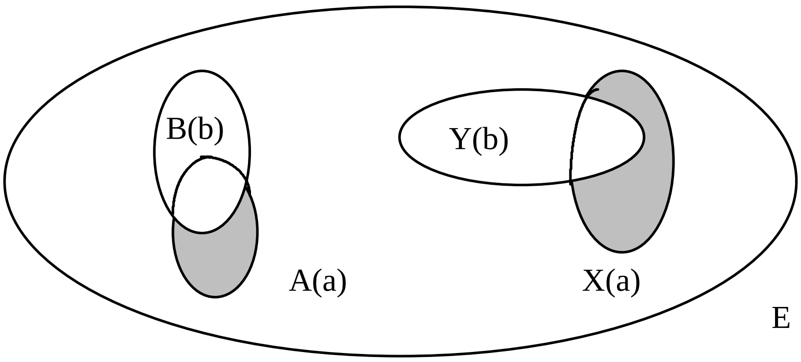

Figura 4
Un árbol implicativo rinde cuenta del orden parcial inducido por esta intensidad de implicación.
La implicación entre variables no resulta suficientemente sintética en algunos casos y Régis Gras creó a partir de ella (y para satisfacer una demanda surgida de la didáctica de las matemáticas, en concreto de Gérard Vergnaud) la implicación entre clases de variables.
En una reciente obra (Gras 1996), el autor expone en detalle el método, su fundamento teórico y sus aplicaciones didácticas.
De una manera general, los dos métodos basados en la similitud y la implicación han permitido en diversos trabajos definir procedimientos e incluso concepciones estables y consistentes, a partir de clases de comportamientos; Batanero C., Godino J. D. y Navarro-Pelayo V. (1995), aplican el análisis implicativo, en combinación con otros métodos, a la evaluación del razonamiento combinatorio de alumnos secundaria.
Procedimiento para la utilización de los métodos del análisis de datos
Para Gras (1992), las nuevas perspectivas que ofrece el análisis de datos crean la *ilusión* de que los datos recogidos y tratados sin una elección oportuna del método a seguir y sin hipótesis previas, van a proporcionar informaciones claras y resultados organizados.
El procedimiento que propone Gras (1992) es el siguiente:

- elegir un método de análisis;

- elaborar una tabla de datos exhaustiva, pertinente, homogénea y compatible con el método elegido;

- poseer un conocimiento matemático que controle y facilite la interpretación e

- interpretar los resultados numéricos y gráficos de manera sintética y con una cierta distancia y saber restringir o extender los datos para un segundo análisis que confirme o critique las primeras interpretaciones. Finalmente y, llegado el caso, practicar un método inferencial».

El "segundo análisis" puede convertirse en varios análisis sucesivos, puesto que las observaciones efectuadas en didáctica de las matemáticas tienen en general un carácter dinámico y para dar cuenta de él no basta con un solo análisis, que vendría a ser como una instantánea –una fotografía– de la contingencia.

5. Un resultado del análisis implicativo

En un cuestionario sobre proporcionalidad (Lacasta 1995) en el que se planteaban a una muestra de 110 alumnos problemas paralelos, redactados de diversas formas (a través de un gráfico, de un texto y de una tabla numérica), entre otros análisis multivariantes, se realizó un análisis implicativo[1] y se encontró que el éxito en los problemas planteados a través de un texto o de una tabla numérica (variables "Tx" y "Tb"), implicaba el éxito en los problemas planteados gráficamente (variable "G"):

{Tx, Tb}   {G}

Se podría pensar que será mejor plantear primero los problemas mediante textos o tablas, porque su resolución implica la resolución gráfica.

Ahora bien, en términos conjuntistas, el conjunto de los alumnos que resuelven los problemas planteados a través de texto o tabla (llamémosle "T") está "estadísticamente incluído" en el de los alumnos que resuelven los problemas de proporcionalidad planteados gráficamente (llamémosle "G"):

G   T

Se podría pues pensar también que, puesto que prácticamente sólo entre los alumnos que resuelven gráficamente esos problemas, se pueden encontrar aquéllos que continuarán resolviéndolos planteados de otra forma, será mejor empezar por el planteamiento gráfico.

Esta especie de paradoja pone de manifiesto que, si bien el análisis implicativo es un buen instrumento para el investigador, el establecimiento de conclusiones no se puede hacer, llevados por la novedad o por la pertinencia del método (éste u otro), de manera simplista o meramente empirista. François Pluvinage (1986), sin referirse en particular al análisis implicativo, advierte que no existe un "método canónico" que pueda practicarse con la exclusión de los demás y que .

6. Tercera estrategia estadística: ordenar los modelos sucesivos, separar la variabilidad explicada de la variabilidad aleatoria

---

[1] Se realizó a través de una de las versiones del programa informático CHIC (Classification Hiérarchique Implicative et Cohésive), Université de Rennes 1, IRMAR, 35042 Rennes Cedex - FRANCIA.

Las dos estrategias anteriores ponen de manifiesto que, ante un conjunto de datos, se pueden proponer distintas leyes matemáticas y modelos estocásticos y se pueden aplicar varios métodos para probar el valor de esas leyes como modelos para esos datos. Tras esta aplicación de los métodos, finalmente se retendrán algunas de esas leyes y se rechazarán otras.

Pero en este caso hay que establecer una estrategia de conjunto para elegir los modelos que se examinarán sucesivamente.

Se trata de examinar las fluctuaciones de los valores en torno a la media, o más generalmente en torno al modelo y la ley que se le puede atribuir: es decir, analizar la varianza.

La tercera estrategia consiste en descomponer la varianza en varianza explicada por el modelo determinista y varianza residual explicada por el azar y comparar las proporciones de una y otra.

7. A modo de conclusión

En la búsqueda de la validación de una hipótesis, coincidimos con Régis Gras en que se trata de encontrar un justo equilibrio entre el desprecio de los métodos estadísticos, el rechazo a invertir tiempo y esfuerzo en este campo y la "estadísticomanía", que conduce a una gran cantidad de resultados inaprovechables. Si el investigador en didáctica de las matemáticas conoce los principios, alcance y limitaciones de esos métodos, podrá usarlos o no, pero estará en mejores condiciones para tomar las decisiones oportunas en su trabajo y para juzgar la pertinencia del aparato estadístico utilizado en otras investigaciones.

La existencia de numerosos programas estadísticos (SPSS, STATITCF, BMDP, Excel, CHIC, etc…) puede facilitar el que el investigador se lance a una vorágine estéril de representaciones y cálculos, utilizando el máximo de métodos, lo más sofisticados posible, despreciando algunas veces las aplicaciones estadísticas más sencillas (como la mera observación de un histograma de frecuencias), que le pueden ayudar mucho más eficazmente en los primeros pasos de su investigación.

Reconocimientos

Agradezco a Carmen Batanero el interés mostrado por este escrito y sus acertadas sugerencias, que he procurado seguir. Asimismo agradezco a Guy Brousseau sus críticas y el ánimo recibido.

Bibliografía

AMON, J. (1990): *Estadística para psicólogos,* Ed. Pirámide. Madrid.

BATANERO C., GODINO J. D. y NAVARRO-PELAYO V. (1995): 'The Use of Implicative and Correspondence Analysis for assessing Pupil's Combinatorial Reasoning', Actes du Colloque *Méthodes d'analyses statistiques multidimensionnelles en didactique des mathématiques,* pp. 245-256, IUFM Caen, R. Gras éd. U. de Rennes.

BROUSSEAU G. (1993a): *Stratégies de l'analyse statistique,* LADIST. Université Bordeaux I.

BROUSSEAU G. (1993b): *Fiches de Statistiques non paramétriques pour la didactique,* LADIST. Université Bordeaux I.

BROUSSEAU G. y LACASTA E.(1995): 'L'analyse statistique des situations didactiques', Actes du Colloque *Méthodes d'analyses statistiques multidimensionnelles en didactique des mathématiques,* pp. 53-90, IUFM Caen, R. Gras éd. Université de Rennes.

GRAS R. (1992): 'L'analyse des données : une méthodologie de traitements de questions de didactique' *Recherches en Didactique des Mathématiques* Vol 12 nº 1, pp. 59-72.

GRAS R. (1996): *L'implication statistique. Nouvelle méthode exploratoire de données. Applications à la didactique.* La Pensée Sauvage, Grenoble.

LACASTA ZABALZA E. (1995): *Les graphiques cartésiens de fonctions dans l'enseignement secondaire des mathématiques: illusions et contrôles,* LADIST, Université Bordeaux I.

PLUVINAGE F. (1986): 'Codages et analyses de réponses à des questionnaires de mathématiques'. *Séminaire de Didactique des Mathématiques et de l'Informatique, Année 1985-86,* pp. 211-228, LSD-IMAG, Université Joseph Fourier, Grenoble.

SIEGEL S. (1990): *Estadística no paramétrica. Aplicada a las ciencias de la conducta,* Trillas, México.