



## Análisis estratificado

Preparado por Luis M. Molinero (Alce Ingeniería) Junio 2004

CorreoE: [bioestadistica@alceingenieria.net](mailto:bioestadistica@alceingenieria.net)

 [Artículo en formato PDF](#)

[www.seh-lelha.org/stat1.htm](http://www.seh-lelha.org/stat1.htm)

### Introducción

En algunos trabajos de investigación, en los que se comparan dos tratamientos, se utiliza el procedimiento denominado **estratificación** al objeto de obtener una estimación de la diferencia entre tratamientos no sesgada y más eficiente. Para ello una variable, o más de una, que son consideradas como factores pronóstico importantes, y por tanto pueden afectar al resultado, se dividen en categorías o estratos. Los pacientes que entran en el estudio quedan clasificados en uno de esos estratos, de acuerdo con sus valores de las variables pronóstico, y el procedimiento de asignación aleatoria a uno de los tratamientos se realiza dentro de cada estrato.

Una situación típica de estratificación se puede producir en los ensayos multicéntricos, donde el procedimiento de asignación aleatoria del tratamiento al paciente se efectúa dentro de cada centro. Pudiera ser que los pacientes de un mismo centro tiendan a tener respuestas al tratamiento más parecidas entre sí, debido a variables influidas por las características del centro que son difíciles de cuantificar.

Otras variables utilizadas en la estratificación pueden ser la edad, el sexo, el grado de severidad de la patología, etc.

En epidemiología en estudios en los que se busca determinar la posible asociación entre un factor de riesgo y una determinada patología, la estratificación se emplea para controlar la influencia de otras variables que pueden afectar al resultado.

En un análisis estratificado nos encontramos con los siguiente pasos:

1. Seleccionar las variables que vamos a controlar.
2. Si las variables no son cualitativas habrá que convertirlas, definiendo los correspondientes intervalos.
3. Los estratos quedarán definidos de acuerdo a todas las combinaciones posibles de las diferentes categorías de las variables que se controla.
4. Dentro de cada estrato se asigna aleatoriamente a la mitad de los pacientes un tratamiento y a la otra mitad otro tratamiento (o de acuerdo a la proporción marcada por el diseño del estudio).

Aunque el objetivo del trabajo suele ser la estimación de una medida global de la diferencia de efecto entre los dos tratamientos, parece lógico que previamente se investigue si existe o no homogeneidad entre los resultados en los diferentes estratos, que justifique la utilización de esa medida global.

En un análisis estratificado la variable respuesta que se determina puede ser de cualquier tipo, por ejemplo numérica como una presión arterial o cualitativa como infarto (*no,sí*). Inicialmente vamos a centrarnos en el caso de una variable binaria o dicotómica, con dos posibles respuestas como puede ser *éxito/fracaso*.

### Análisis estratificado para resultados binarios

Cuando se compara un resultado binario para dos tratamientos en un conjunto de K estratos, tenemos una relación de K tablas de contingencia 2x2. En esta situación podemos utilizar diferentes medidas para comparar los tratamientos: puede ser una medida absoluta como la **diferencia de proporciones** de éxitos entre los dos grupos, o una medida relativa como el **odds ratio**. En cualquier caso, la prueba de homogeneidad que se

aplique deberá efectuarse sobre la medida que luego se vaya a emplear para presentar el resultado promedio global, ya que una homogeneidad de los odds ratio no garantiza una homogeneidad de diferencia de proporciones, y viceversa.

Existen diferentes procedimientos para verificar la homogeneidad del resultado entre los diferentes estratos. Los más utilizados se basan en ponderar los resultados para cada estrato.

Supongamos que tenemos K estratos. Denominamos  $n_{i1}$  al número de pacientes en el estrato i que han sido asignados al primer tratamiento y  $n_{i2}$  los asignados al segundo tratamiento. El tamaño total del estrato i es por tanto  $N_i = n_{i1} + n_{i2}$

Designemos como  $\delta_i$  al parámetro que compara los tratamientos en el estrato i (puede ser la diferencia de proporciones, el odds ratio, el riesgo relativo).

La medida global (desconocida) del efecto del tratamiento viene dada por  $\delta = \sum_{i=1}^K f_i \delta_i$ , donde  $f_i$  es la fracción

de pacientes en la población con las características de ese estrato, que normalmente es desconocida y se estimará como

$$f_i = \frac{n_{i1} + n_{i2}}{\sum N_i}$$

La hipótesis nula que interesa verificar es  $\delta = 0$ , y se desea obtener un intervalo de confianza para el efecto global. Vamos a comentar el análisis para la diferencia de riesgo y para el odds ratio.

### Análisis estratificado para diferencias de riesgo

En cada estrato podemos calcular una diferencia de proporciones entre los grupos

$$d_i = p_{i1} - p_{i2}$$

Si llamamos  $X_{i1}$  al número de sucesos de interés para el tratamiento 1 en el estrato i, la proporción  $p_{i1}$  se estima como  $p_{i1} = X_{i1}/n_{i1}$ , y de forma análoga la proporción en el grupo 2. La varianza de la diferencia de proporciones  $d_i$  se estima como:

$$\text{Var}(d_i) = \frac{p_{i1}(1-p_{i1})}{n_{i1}-1} + \frac{p_{i2}(1-p_{i2})}{n_{i2}-1}$$

Si se cumple que la diferencia de proporciones se mantiene en los diferentes estratos (hipótesis nula)  $d_1 = d_2 = \dots = d_k = D$ , se puede obtener una estimación de D como una media ponderada

$$D = \frac{\sum w_i d_i}{\sum w_i} \quad [1]$$

donde  $w_i$  es el **peso** que se asigna a cada estudio en el cálculo de esa media ponderada.

Existen diferentes alternativas a la hora de escoger esos pesos, siendo las dos más sencillas la basada en el tamaño de los estratos y la basada en la varianza.

Si utilizamos la **varianza** el peso es  $w_i = 1/\text{Var}(d_i)$ .

La prueba que permite verificar la homogeneidad de las diferencias de proporciones entre los estratos (Fleiss, 1981) es

$$Q = \sum w_i (d_i - D)^2 \quad [2]$$

que si se cumple la hipótesis nula se distribuye como una  $\chi^2$  con  $K-1$  grados de libertad.

Utilizando como pesos la inversa de la varianza se está dando mayor peso a los estratos en los que la diferencia de proporciones se estima con mayor precisión.

Pero si en lugar de utilizar como peso la inversa de la varianza se utiliza la **media armónica del tamaño de los estratos**, los pesos se calculan ponderando según la siguiente expresión

$$\frac{n_{i1}n_{i2}}{n_{i1} + n_{i2}}$$

El peso para el estrato  $i$  viene dado por

$$w_i = \frac{(n_{i1}n_{i2}) / (n_{i1} + n_{i2})}{\sum_{j=1}^K (n_{j1}n_{j2}) / (n_{j1} + n_{j2})}$$

Estos pesos se conocen con el nombre de **pesos de Cochran–Mantel–Haenszel**, debido a que fueron propuestos por Cochran y son idénticos a los pesos que se utilizan en la difundida **prueba de Mantel–Haenszel** para comparar dos proporciones en un diseño estratificado. Con estos pesos se está dando mayor importancia a los estratos que tienen mayor número de pacientes.

Las fórmulas anteriores son análogas a las utilizadas en un [meta-análisis](#) para combinar los diferentes estudios (estratos), lo que resulta lógico.

### **Análisis estratificado para el odds ratio**

En ocasiones es necesario o se prefiere utilizar el odds ratio como medida de comparación entre los grupos. En estos casos tanto la fórmula [\[1\]](#) como la [\[2\]](#) son aplicables substituyendo  $d_i$  por el  $\ln OR_i$  y siendo ahora  $D$  el OR ponderado, y utilizando como peso en este caso la inversa de la varianza del  $\ln OR$ .

Existen otras pruebas para verificar la homogeneidad de los odds ratios entre varios estratos, siendo una de las más conocidas la **prueba de Breslow–Day**.

Para combinar los odds ratio en una medida global se emplea con frecuencia el **método de Mantel–Haenszel**, pero también podemos utilizar el [modelo logístico](#) como herramienta alternativa, tanto para calcular un odds ratio ajustado como para verificar si existe o no heterogeneidad entre los estratos.

Veamos mediante un ejemplo como emplearíamos la regresión logística. En un ensayo clínico aleatorizado se estudia la mortalidad después de un infarto agudo de miocardio en dos grupos, un grupo control con placebo y otro de pacientes tratados con metoprolol. El análisis se estratificó en tres categorías de edad. Los resultados se indican en la siguiente tabla [\(\\*\)](#):

Edad	Mortalidad		Diferencia	Odds Ratio
	Placebo (P)	Metoprolol (M)	P-M	P:M
40-64	0.057 (26/453)	0.045 (21/464)	0.012	1.28
65-69	0.144 (25/174)	0.067 (11/165)	0.077	2.35
70-74	0.157 (11/70)	0.116 (8/69)	0.041	1.42

Si se aplica las fórmula [1] con el odds ratio se obtiene un OR global de 1.58, que es muy similar al que se obtiene por el procedimiento de Mantel-Haenszel 1.59.

La prueba de heterogeneidad [2] nos da un valor de 1.61 que acudiendo a una chi2 con 2 grados de libertad proporciona un nivel de probabilidad de 0.447, por lo que, de acuerdo a ese resultado, no rechazamos la hipótesis de homogeneidad entre los estratos.

Si ajustamos un modelo logístico para la mortalidad con dos variables: el tratamiento y la edad

$$\text{Mortalidad} = \text{Modelo logístico (Tratamiento, Edad)}$$

y exponenciamos el coeficiente correspondiente al tratamiento obtenemos una estimación del odds ratio de 1.59, análoga a la que se obtenía por el procedimiento de Mantel-Haenszel. (Nota: recuérdese que al tratar la edad como variable cualitativa hay que convertirla en variables diseño "dummy").

Si queremos verificar la hipótesis de homogeneidad contrastamos el modelo anterior con otro modelo en el que se añade un término de interacción, correspondiente al producto *Tratamiento x Edad*

$$\text{Mortalidad} = \text{Modelo logístico (Tratamiento, Edad, Tratamiento x Edad)}$$

y mediante el cociente de verosimilitud se contrasta este modelo con el modelo anterior (sin el término de interacción), obteniendo en nuestro ejemplo un valor de chi2 de 1.65, con 2 grados de libertad, muy similar por tanto al valor obtenido mediante la fórmula [2] y análogo también al que se obtiene mediante la prueba de Breslow-Day.

Pero si lo que pretendemos estimar es la diferencia de proporciones entre los grupos, aplicaríamos las fórmulas [1] y [2] a la diferencia de proporciones y en ese caso la regresión logística no nos serviría. En nuestro ejemplo obtenemos una diferencia global ajusta de 0.02, mientras que el resultado de la chi2 de la prueba de heterogeneidad es ahora de 3.34, lo que para 2 grados de libertad corresponde a un nivel de probabilidad de 0.19, mucho menor que el valor obtenido cuando se verificaba la homogeneidad de los odds ratio. Vemos que la heterogeneidad es más acusada para las diferencias de proporciones que para los odds ratio. Luego en este ejemplo los estratos son menos homogéneos en cuanto a la diferencia de proporciones que en cuanto al odds ratio. La diferencia de proporciones sin ajustar es de 0.03 frente a 0.02 ajustando según los estratos (edad), sin embargo el odds ratio sin ajustar es 1.61 frente a 1.58 ajustado.

## Análisis estratificado para variables cuantitativas

Si la variable que se analiza es cuantitativa, por ejemplo una presión arterial, el método de análisis que primero nos viene a la mente, si tenemos dos factores: el grupo de tratamiento y el estrato, es el [análisis de la varianza para dos factores](#), donde podemos evaluar el efecto del tratamiento ajustado y verificar si se puede aceptar la hipótesis de homogeneidad (no hay interacción con el estrato). Sin embargo, dado que es muy difícil en los estudios reales que los tamaños de cada combinación tratamiento/estrato sean iguales, o incluso que guarden una relación de proporcionalidad, es raro que se den las condiciones para la aplicación de un análisis de la varianza estándar, por lo que lo habitual será utilizar un [modelo de regresión lineal](#) y aplicarlo de forma similar a como se comentó anteriormente para un modelo logístico.

## Modelos de efectos aleatorios

En el [ejemplo](#) que se ha planteado más arriba hemos dividido la edad en 3 estratos concretos, que son los que nos interesan, quedando fijos una vez especificados. Sin embargo, si el estudio se ha llevado a cabo en diferentes hospitales, dentro de los cuales se ha efectuado la asignación aleatoria del tratamiento a los pacientes, podemos imaginar que esos hospitales constituyen una muestra de todos los hospitales, aunque probablemente no hayan sido escogidos de forma aleatoria. Mientras que los estratos de edad son **fijos**, decimos que los estratos hospital son **aleatorios**, y así mientras que en el primer caso hablamos de un **modelo de efectos fijos**, en el segundo hablamos de un **modelo de efectos aleatorios**. Es lógico dudar de que en cada hospital las observaciones sean independientes, y cabe sospechar que pueda existir correlación entre las observaciones de cada hospital, de tal manera que las mayores diferencias entre los tratamientos aparezcan con más frecuencia en los mismos centros. Para tener en cuenta esa posible correlación entre las observaciones del mismo hospital (o hablando en términos generales de cada **bloque** o "**cluster**") se utilizan modelos de regresión que introducen términos "aleatorios" y por ello se denominan **modelos de efectos aleatorios**. Si estos modelos sólo tienen estratos fijos, como era el caso de la edad, hablamos de **modelos de efectos fijos**. Y si combinan estratos de ambos tipos, por ejemplo edad y hospital, se habla de **modelos de efectos mixtos**.

Estos modelos de regresión son más complicados tanto conceptualmente como de estimación, pero razonando de forma similar a como se ha explicado para el modelo de regresión logística de efectos fijos, nos permiten estimar parámetros ajustados teniendo en cuenta la estratificación y verificar la hipótesis de homogeneidad de una forma más precisa.

Pero de todo eso ya hemos hablado en un [artículo anterior](#).

## Enlaces



[Stratified analysis: introduction to confounding and interaction](#)

Gerstman, B. B. (2003). StatPrimer: Statistics for Public Health Practice. Available: [www.sjsu.edu/faculty/gerstman/StatPrimer](http://www.sjsu.edu/faculty/gerstman/StatPrimer)



[Cochran–Mantel–Haenszel Statistics](#)

[SAS/STAT User's Guide](#)

---

(\*) [Ejemplo](#) tomado de Pocock SJ. Clinical Trials: A practical Approach. Wiley: New York, 1983

---



[Índice de artículos](#)

[Principio de la página](#) 