Asociación de la Sociedad Española de Hipertensión
Liga Española para la lucha contra la Hipertensión Arterial

# La paradoja de Simpson

*Preparado por Luis M. Molinero (Alce Ingeniería)* **Artículo en formato PDF**
CorreoE: **bioestadistica@alceingenieria.net**

*Octubre 2001*

Se denomina paradoja de Simpson al cambio en el sentido de una asociación entre dos variables (numéricas o cualitativas) cuando se controla el efecto de una tercera variable.

Un ejemplo que se presenta habitualmente para ilustrar esa situación es la comparación de las tasas de mortalidad de dos hospitales, que pueden favorecer de forma global al hospital A frente al B, y sin embargo al analizarlas por procedimentos se descubre que cambia el signo de la diferencia, debido a que los pacientes con peor prónostico y patologías más graves son internados en el hospital B con mayor frecuencia.

Vamos a plantear un ejemplo concreto: en un estudio comparativo sobre tolerancia de dos fármacos antihipertensivos se determina la presencia o no de efectos secundarios leves y se obtiene los siguientes datos

| | | Tratamiento | | |
|---|---|---|---|---|
| | | **A** | **B** | |
| **Efecto secundario** | **NO** | 410 | 434 | 844 |
| | **SI** | 115 | 91 | 206 |
| | | 525 | 525 | 1050 |

Donde vemos que el 21.9 % de pacientes tiene algún efecto adverso en el grupo A, frente a 17.3 % en el grupo B, diferencia importante pero que no llega al nivel de significación estadística habitualmente aceptado, ya que p = 0.07.

Pero si se separa el estudio en pacientes ancianos (≥ 75 años) y el resto (< 75 años) se obtienen las siguientes tablas

**Pacientes < 75 años**

| | | Tratamiento | | |
|---|---|---|---|---|
| | | **A** | **B** | |
| **Efecto secundario** | **NO** | 122 | 351 | 473 |
| | **SI** | 8 | 54 | 62 |
| | | 130 | 405 | 535 |

ahora la proporción de efectos adversos, en los pacientes de menos de 75 años, es 6.2 % en el grupo A, frente a 13.3 % en el grupo B, diferencia que es importante y además estadísticamente significativa, p = 0.027, y con signo contrario a la observada en el estudio completo.

En el otro grupo de pacientes de más edad:

**Pacientes ≥ 75 años**

| | | Tratamiento | | |
|---|---|---|---|---|
| | | **A** | **B** | |
| **Efecto secundario** | **NO** | 288 | 83 | 371 |
| | **SI** | 107 | 37 | 144 |
| | | 395 | 120 | 515 |

tenemos un 27.1 % pacientes con reacciones adversas en A, frente a 30.8 % en el grupo B, p = 0.42

Vemos que al considerar la edad (según la clasificación escogida), la relación cambia de signo: en el estudio global era superior la tasa de efectos adversos en el grupo A, pero al estratificar por edad en ambos casos es menor en el grupo A que en el B. Lo resumimos en la siguiente tabla

| | **A** | **B** | **p** |
|---|---|---|---|
| **Global** | 21.9 % | 17.3 % | 0.07 |
| **< 75** | 6.2 % | 13.3 % | 0.03 |
| **≥ 75** | 27.1 % | 30.8 % | 0.42 |

La interpretación de una paradoja de este tipo no siempre es fácil, sobre todo cuando hay más de dos estratos; incluso es posible que en ocasiones no tenga interpretación, y en cualquier caso ésta depende siempre de las características de cada estudio.

En este ejemplo está claro que, para ese punto de corte en la edad (75 años), los dos grupos de tratamiento están muy desequilibrados: la proporción de ancianos es de 75.2 % en el grupo A, frente a 22.9 % en el grupo B, y la tasa de efectos adversos en el grupo de pacientes con menos de 75 años es del 11.6 % frente al 28 % en los pacientes ancianos.

Puesto que los datos de este ejemplo son ficticios no tiene ningún sentido buscar una explicación, pero en una situación real la interpretación está condicionada a cómo se diseñó el estudio, si se trata de un ensayo aleatorio, es decir si a los pacientes les fue asignado el tratamiento de forma aleatoria o si se trata de un estudio observacional sin esa característica de diseño.

Una pregunta que nos viene a la mente enseguida es ¿qué ocurriría si el punto de corte para la edad se hubiera fijado en otro valor?. Por ello en el caso de variables continuas como la edad se sugiere utilizar un modelo de regresión logística en el que intervenga esa variable como tal, permitiendo así ajustar su efecto sin necesidad de fijar un punto de corte que siempre será, en cierta medida, arbitrario.

En los enlaces que se presentan al final del artículo se pueden ver ejemplos, reales y ficticios de esta paradoja.

En el primero de ellos "Confounding and Simpson's paradox", se nos presenta, entre otros, un ejemplo real, a partir de los datos del artículo de Charig et all "Comparison of treatment of renal calculi by operative surgery, percutaneous nephrolithotomy, andextracorporeal shock wave lithotripsy". En dicho artículo se compara en un estudio retrospectivo las tasas de éxito en la eliminación de cálculos renales mediante cirugía abierta o mediante nefrolitotomía percutánea:

| | | Tratamiento | |
|---|---|---|---|
| | | **Cirugía** | **Nefrolitotomía** |
| **Exito** | **NO** | 77 | 61 |
| | **SI** | 273 | 289 |
| | | 350 | 350 |

lo que supone un 78 % de éxito en la cirugía frente a un 83 % en nefrolitotomía (p = 0.13). Peso si se estratifica teniendo en cuenta el tamaño del cálculo, el panorama que obtenemos cambia:

**Menores de 2 cm**

| | | Tratamiento | |
|---|---|---|---|
| | | **Cirugía** | **Nefrolitotomía** |
| **Exito** | **NO** | 6 | 36 |
| | **SI** | 81 | 234 |
| | | 87 | 270 |

93 % en cirugía frente a 87 % en nefrolitotomía (p = 0.13)

**Igual o mayor de 2 cm**

| | | Tratamiento | |
|---|---|---|---|
| | | **Cirugía** | **Nefrolitotomía** |
| **Exito** | **NO** | 71 | 25 |
| | **SI** | 192 | 55 |
| | | 263 | 80 |

73 % en cirugía frente a 69 % en nefrolitotomía (p = 0.48).

Resultados que se resumen en la tabla siguiente

| | **Cirugía** | **Nefrolitotomía** | **p** |
|---|---|---|---|
| **Global** | 78 % | 83 % | 0.13 |
| **< 2 cm** | 93 % | 87 % | 0.13 |
| **≥ 2 cm** | 73 % | 69 % | 0.48 |

Los autores del artículo sobre la paradoja hacen una posible interpretación de las causas de estos resultados, que por cierto no parecen satisfacer al autor del artículo original, Charing, tal y como se puede juzgar por su un tanto airada respuesta.

En uno de los enlaces que se indican al final (Simpson's paradox masks the good news about american schools...)se presenta un ejemplo interesante y real, que destroza las conclusiones obtenidas por otros investigadores respecto al descenso en el nivel de conocimientos de los estudiantes de enseñanza primaria en USA respecto a periodos anteriores, conclusión que cambia completamente si se efectúa el estudio por grupos raciales.

¿Qué conocimiento podemos extraer de todo esto? En su conocido libro de epidemiología Rothman afirma que "*aunque este curioso cambio es conocido como la paradoja de Simpson, el fenómeno no es realmente una paradoja; no se contravienen ni la lógica ni ninguna de las premisas*". La verdad es que, aunque resulte desalentador no soy tan optimista y creo que el término paradoja está bien aplicado, y que por ello no existen normas generales al respecto, salvo que no es una buena práctica evaluar hipótesis que no hayan sido expuestas y fundamentadas en el diseño del estudio, así como tampoco lo es la elección de puntos de corte arbitrarios en el control de variables numéricas.

## Enlaces de interés

- Confounding and Simpson's paradox
  Steven A J,Mark A M. BMJ 1994;309:1480−1481
- Simpson's paradox masks the good news about american schools and undoes Stephan Thernstrom, Winthrop professor of History at Harvard Universiy
  Gerald Bracey, Education Disinformation Detection and Reporting Agency
- Simpson's Paradox: A real example from a longitudinal study in South Africa
  Morrell C H, Journal of Statistics Education v.7, n.3 (1999)
- Simpson's Paradox Worksheet
- Introduction to Research Design and Statistics:Simpson's Paradox
  Ejemplo con medias
- What is Simpson's Paradox?

## Bibliografía seleccionada


- Simpson, E. H. (1951), "**The Interpretation of Interaction in Contingency Tables**," Journal of the Royal Statistical Society, Ser. B, 13, 238−241
- Rothman, J.K, **Epidemiología moderna** Ed. Díaz de Santos, 9187


Indice de artículos

Principio de la página