

Comparación de un resultado de tipo ordinal entre dos muestras independientes

Preparado por Luis M. Molinero (Alce Ingeniería)

CorreoE: bioestadistica@alceingenieria.net

[Artículo en formato PDF](#)

Marzo 2001

En medicina es muy corriente el uso de variables de tipo **ordinal**; se trata de variables cualitativas cuyas categorías, sin un significado numérico preciso, guardan una relación de orden. Las encontramos con frecuencia cuando se trata de valorar el grado de curación, o el estadio de una enfermedad.

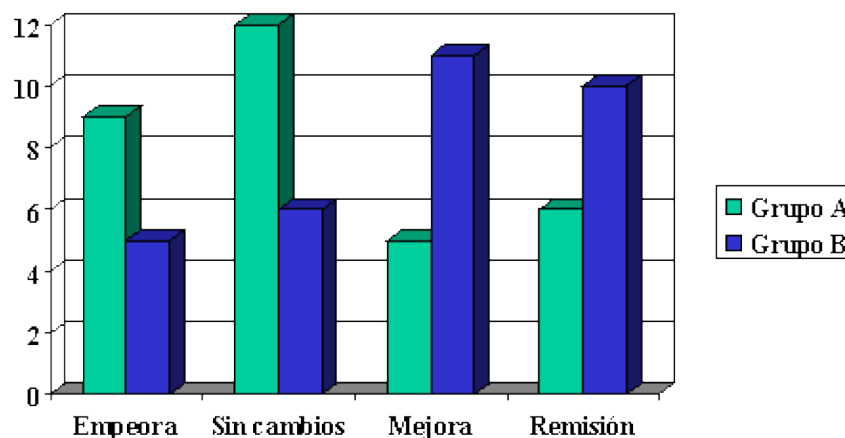
Las variables cualitativas que no tienen una relación de orden entre sus categorías se las designa como **nominales**, o también simplemente cualitativas. Variables nominales son, por ejemplo, el estado civil, o el grupo sanguíneo.

En el caso de las variables cualitativas nominales, cuando se desea valorar si existen diferencias en las proporciones de cada categoría observadas en dos grupos diferentes, la prueba estadística habitualmente empleada es un contraste utilizando la **prueba de χ^2** . Así, por ejemplo, si queremos determinar si existe asociación entre el buen o mal cumplimiento de una pauta terapéutica para HTA y el estado civil (con 4 categorías); se construye una tabla 2 x 4, a partir de la cual se calcula el correspondiente valor del χ^2 .

Pero veamos ahora una situación diferente, en la que para valorar dos pautas terapéuticas el resultado que se empleará es una variable ordinal con cuatro posibles respuestas: *empeora*, *sin cambios*, *mejora* y *remisión*, y que en un ensayo aleatorizado con 32 pacientes por grupo se observa la situación descrita en la tabla siguiente:

| | Empeoran | Sin cambios | Mejoran | Remisión |
|---------|----------|-------------|---------|----------|
| Grupo A | 9 | 12 | 5 | 6 |
| Grupo B | 5 | 6 | 11 | 10 |

A simple vista parece que en el grupo B hay una tendencia a valores mayores de frecuencia hacia la derecha de la tabla (mejores resultados), mientras que la situación es la contraria en el caso del grupo A, tal y como podemos ver en la gráfica



Si el lector interesado realiza un contraste de asociación para esta tabla obtendrá un valor de χ^2 de 6.39 que con 3 grados de libertad nos da un valor de probabilidad de 0.094; por lo que, si aceptamos como punto de

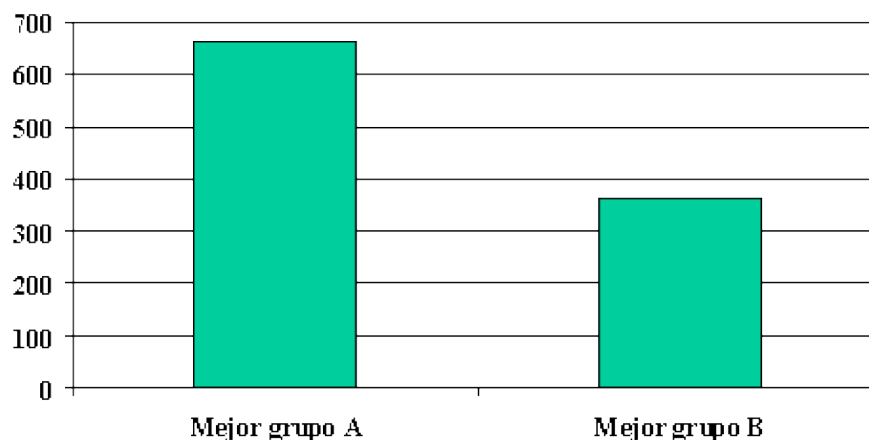
corte el nivel de $p < 0.05$, a pesar de lo que nos parecía intuir, no hay suficiente evidencia estadística para rechazar la hipótesis de no asociación.

¿Qué es lo que ha ocurrido? Sencillamente que en el contraste efectuado no se ha tenido en cuenta el tipo de variable medida y ha sido considerada como nominal sin apreciar que existe una relación de orden entre las respuestas. En este caso sería más adecuada la utilización de la prueba descrita por **Mann–Whitney** y que consiste en comparar cada individuo del primer grupo con cada individuo del segundo grupo, registrándose cuántas veces sale favorecido en esa comparación. Básandose en ese recuento se construye una medida que es la que se contrasta para ver si la diferencia con el resultado esperado, en el caso de que hubiera diferencias entre los grupos, puede o no ser atribuido al azar.

En nuestro ejemplo si comparamos cada uno de los 32 pacientes del primer grupo con cada uno de los 32 del segundo podemos formar $32 \times 32 = 1024$ parejas o comparaciones. Si no hubiera diferencias entre los grupos es razonable esperar que alrededor de la mitad (512) de esas comparaciones favorezcan a un grupo y el resto al otro.

Si en caso de empate se asigna la mitad de las comparaciones a un grupo y la otra mitad al otro grupo se puede comprobar que en el ejemplo que nos ocupa, el primer grupo sale favorecido en 661 de las 1024 comparaciones (el 65 %) frente a 363 veces en las que lo es el grupo B, casi el doble de veces. Esos números constituyen precisamente el estadístico de Mann–Whitney y se denominan habitualmente **U**.

En la figura se refleja gráficamente este resultado



Si utilizamos un programa de cálculo estadístico y solicitamos la prueba de Mann–Whitney, obtendremos los valores de **U** para cada muestra, y si los dividimos entre $n1 \cdot n2$, siendo **n1** el tamaño de la primera muestra y **n2** el de la segunda, habremos calculado la proporción de veces en las que los individuos de esa muestra salen favorecidos en las comparaciones, siempre que una categoría más baja suponga una mejor situación ya que, obviamente, si fuera una peor situación habremos calculado el número de veces en los que sale desfavorecido.

Los programas de estadística nos dan además un nivel de significación para esos valores de **U**. En nuestro caso el valor de probabilidad obtenido es de 0.038, por lo que ahora sí rechazaríamos la hipótesis nula (con punto de corte $p < 0.05$).

Bajo la condición de hipótesis nula, es decir tratamientos equivalentes, **U** se distribuye según una normal, de media

$$\frac{n1 \cdot n2}{2}$$

y desviación típica

$$\sigma = \sqrt{\frac{n1 \cdot n2 \cdot (n1 + n2 + 1)}{12}}$$

La aproximación mediante la distribución a la normal se puede considerar válida en el caso de muestras con más de 10 observaciones por grupo, y en las que en ninguna de las categorías se acumula más de la mitad de las observaciones. Cuando no es válida esta aproximación es necesario calcular el valor exacto de p, existiendo valores tabulados .



[Indice de artículos](#)

[Principio de la página](#) ▲