

El uso correcto de las pruebas estadísticas

ULISES ALBERTO QUESTA

Trabajo recibido para su publicación: 3/91. Aceptado: 4/91

Dirección para separatas: Arenales 2396, 2º "8", Buenos Aires, Argentina

Al proceder al análisis de los datos numéricos de una investigación, se debe recordar que es perentorio utilizar una prueba de significación estadística acorde con las características matemáticas de los datos obtenidos.

Un gran número de pruebas estadísticas exige que los datos presenten distribución normal gaussiana, es decir, que la mayor parte de ellos se encuentren ubicados alrededor de los valores medios, adoptando así la conocida forma acampanada. Muchos fenómenos biológicos se distribuyen de esa manera, o lo suficientemente cercana a ella como para poder ser procesados con las pruebas clásicas.

Cuando esos datos son "duros", o sea confiables en su precisión y tipo de distribución, pues proceden de valores realmente cuantificables (segundos, milímetros, gramos), se cumplen las condiciones necesarias para utilizar las "pruebas paramétricas" (t de Student, análisis de varianza). Si se trabaja con valores cualitativos (mucho, poco, nada, por ejemplo) a los que luego se otorga un seudovalor convencional (mucho = 2, poco = 1, nada = 0), no podemos precisar con seguridad su tipo de distribución, porque nos está vedado el uso de esas pruebas, debiendo utilizar otra de la gran familia de pruebas llamadas "no paramétricas" (Tabla 1).

Una segunda característica de los datos a procesar es la manera como fueron recogidos. Tanto los datos paramétricos como los no paramétricos pueden haber sido obtenidos (y procesados) individualmente, caso por caso. El valor de **cada** unidad experimental, cualquiera que sea la cantidad de unidades, es recogida y procesada de manera individual, contrastando así los valores del grupo experimental y del grupo control.

Otro modo de proceder es analizar los valores agrupados, estudiándolos mediante las medias aritméticas y el error estándar (en los casos paramétricos) o con la conocida técnica de chi cuadrado o por la diferencia entre porcentajes.

Hay que tener en cuenta si las unidades expe-

rimentales (pacientes o animales sometidos al experimento) conforman grupos independientes o dependientes. Se dice que dos grupos son independientes si los individuos que los integran pertenecen a grupos distintos (un grupo de pacientes tratados con una droga bajo estudio y otro tratado con placebo); se los llama dependientes si ambos grupos están integrados por los mismos individuos (se estudia en el sujeto una variable dada antes y después de haberle proporcionado el estímulo).

Si los datos son paramétricos —y se trata de dos grupos— la prueba t de Student es útil, pues pequeñas variaciones en su formulación permiten utilizarla en ambas situaciones.

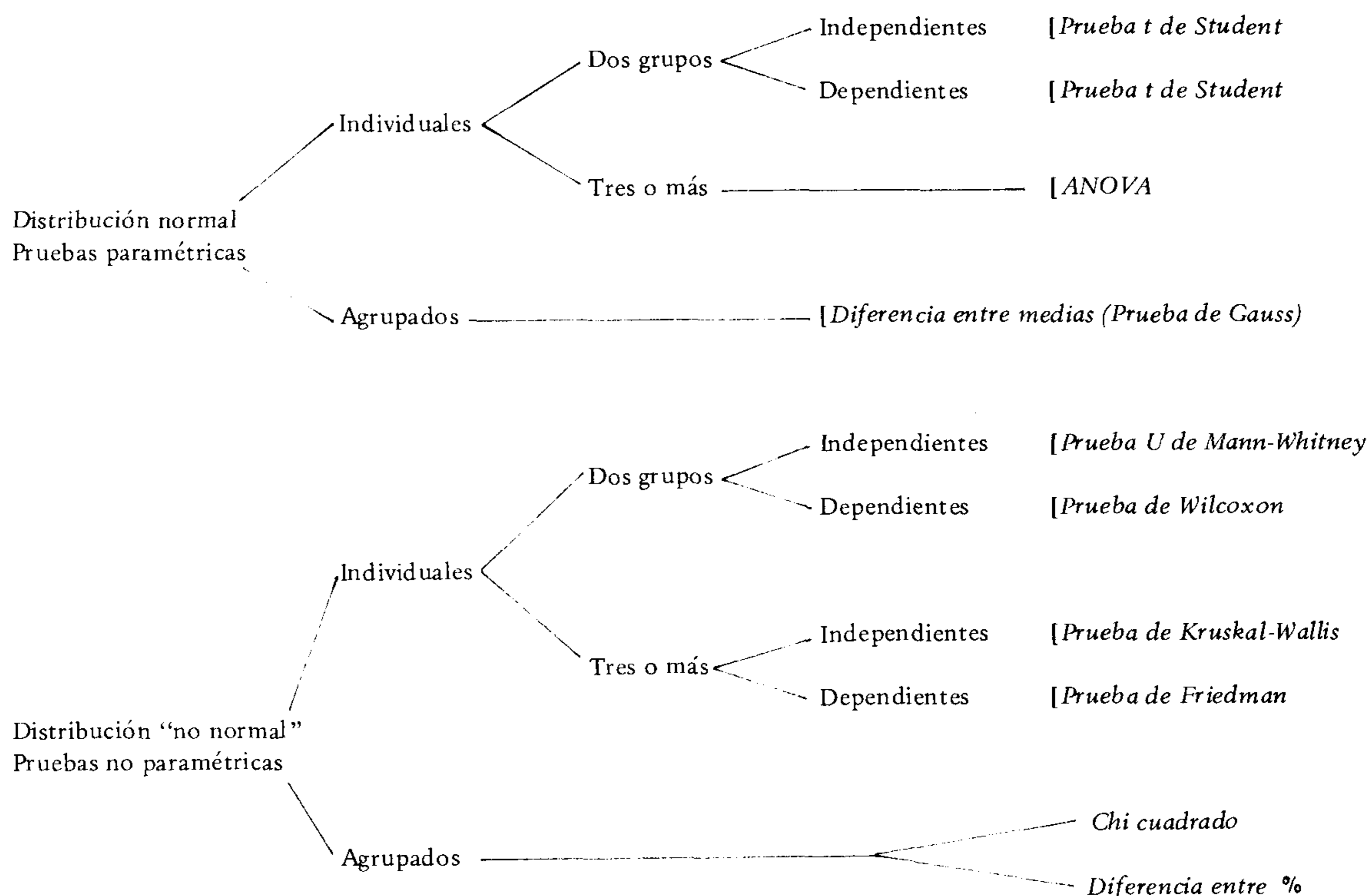
La mayor limitación de esa técnica es que su diseño no permite contrastar más de dos muestras simultáneamente (grupo A-grupo B, antes-después), no siendo lícito utilizarla en múltiples comparaciones sucesivas.

En esta situación se halla indicado el análisis de varianza, la prueba de mayor potencia estadística, que toma en consideración no sólo las variaciones que existen **entre los grupos** (como hace la prueba de t) sino también las existentes **dentro de los grupos**. Las características tan particulares de esta técnica permiten tratar los datos de cualquier número de grupos experimentales.

Al considerar numerosos grupos simultáneamente y recibir —al finalizar el análisis— una información global sobre la importancia de las diferencias, no conocemos con precisión la significación que existe entre los grupos. Ello obliga a efectuar un análisis posterior (si los resultados son significativos) mediante pruebas accesorias tales como las de Tukey o Duncan para determinar las diferencias existentes entre parejas de grupos.

Algo similar ocurre al analizar datos no paramétricos, ya sean de origen biológico, social o educativo. Estos datos también pueden originarse en estudios individuales o agrupados, y

Tabla 1
Selección de la prueba estadística más adecuada



ser independientes o dependientes. En la Tabla 1 se pueden observar cuáles son las pruebas estadísticas indicadas para cada circunstancia.

Debemos tener muy en cuenta, entonces, que si estamos analizando datos obtenidos en una investigación biológica que comparó dos grupos y esos valores son realmente cuantitativos (mmHg, por ejemplo), utilizaremos la prueba *t* de Student en alguna de sus dos versiones, pero si procesamos datos biológicos con características cualitativas (dolor, ansiedad) y les hemos otorgado algún valor de escala, ya no podríamos hacerlo. Para esas circunstancias, si son dos grupos dependientes empleare-

mos la prueba de Wilcoxon (datos antes-después) y si son independientes, la de Mann-Whitney. Lo mismo ocurrirá si estudiamos muchos grupos de valores no paramétricos (Friedman o Kruskal-Wallis).

Se debe recordar que los datos paramétricos pueden ser analizados mediante pruebas no paramétricas, pero no a la inversa. ¿Por qué, entonces, no utilizar siempre estas últimas, independizándonos del problema de la distribución de los datos? Porque, entre otras razones, ellas son menos potentes que las paramétricas y pueden proporcionar resultados falsamente no significativos.