

¹Juan O. Talavera,
²Rodolfo Rivas-Ruiz

Investigación clínica IV. Pertinencia de la prueba estadística

¹Centro de Adiestramiento en Investigación Clínica,
Coordinación de Investigación en Salud,

²Servicio de Escolares y Adolescentes, Hospital de Pediatría

Autores 1 y 2, Centro Médico Nacional Siglo XXI,
Instituto Mexicano del Seguro Social, Distrito Federal, México

Comunicación con: Juan O. Talavera

Correo electrónico: jotalaverap@uaemex.mx

Resumen

Cuando observamos la diferencia entre dos terapias o la asociación de un factor de riesgo o indicador pronóstico con su desenlace, tenemos que evaluar la certeza del resultado. Esta evaluación se basa en un juicio que utiliza información relacionada con el diseño del estudio y el manejo estadístico de la información. En este artículo se menciona específicamente la pertinencia de la prueba estadística seleccionada. Las pruebas estadísticas se eligen a partir de dos características: el objetivo del estudio y el tipo de variables. El objetivo se puede dividir en tres grupos de pruebas: a) en las que se desea mostrar diferencias entre grupos, o de un mismo grupo antes y después de una maniobra; b) en las que se busca mostrar la relación entre las variables; c) en las que se pretende predecir un desenlace. En cuanto a los tipos de variables tenemos dos: las cuantitativas (continuas y discontinuas) y las cualitativas (ordinales y dicotómicas). Por ejemplo, si se busca demostrar diferencias en edad (variable cuantitativa), entre pacientes con lupus eritematoso sistémico, con y sin afección neurológica (dos grupos), la prueba apropiada es la *t* de Student para muestras independientes; pero si lo que se compara entre esos mismos grupos es la frecuencia de mujeres (variable binomial), entonces la prueba estadística pertinente es la χ^2 .

Palabras clave

investigación biomédica
proyectos de investigación
estadística y datos numéricos

Summary

When we look at the difference between two therapies or the association of a risk factor or prognostic indicator with its outcome, we need to evaluate the accuracy of the result. This assessment is based on a judgment that uses information about the study design and statistical management of the information. This paper specifically mentions the relevance of the statistical test selected. Statistical tests are chosen mainly from two characteristics: the objective of the study and type of variables. The objective can be divided into three test groups: a) those in which you want to show differences between groups or inside a group before and after a maneuver, b) those that seek to show the relationship (correlation) between variables, and c) those that aim to predict an outcome. The types of variables are divided in two: quantitative (continuous and discontinuous) and qualitative (ordinal and dichotomous). For example, if we seek to demonstrate differences in age (quantitative variable) among patients with systemic lupus erythematosus (SLE) with and without neurological disease (two groups), the appropriate test is the "Student *t* test for independent samples". But if the comparison is about the frequency of females (binomial variable), then the appropriate statistical test is the χ^2 .

Key words

biomedical research
research design
statistics and numerical data

Introducción

Cuando observamos la diferencia entre dos terapias o la asociación de un factor de riesgo o indicador pronóstico con su desenlace, surge un cuestionamiento: ¿el resultado es real? Decidir si es real requiere dos juicios que se complementan:

1. La planeación y el desarrollo del proceso que documentan dicha diferencia o asociación están libres de errores, o cuando menos éstos son de una magnitud menor que no modifican la orientación de la diferencia o de la asociación (es decir, diseño apropiado y ejecución adecuada).

2. El tamaño de la muestra es suficiente para mantener la estabilidad de los datos y la prueba estadística es apropiada al objetivo.

La planeación y el desarrollo del proceso se han mencionado en los tres capítulos previos de esta serie. Por otra parte, la estabilidad de los datos será discutida a detalle en un artículo posterior al hablar de tamaño de muestra y valor de *p*.

En este artículo discutiremos la pertinencia de la prueba estadística seleccionada. Sin duda, este conocimiento nos permitirá comprender con mayor precisión los resultados obtenidos en estudios de investigación clínica y, por supuesto, incrementará nuestra capacidad para hacer un uso adecuado de los mismos.

Objetivo del estudio y tipo de variable

Las pruebas estadísticas se eligen a partir de dos características: el objetivo del estudio y el tipo de variables. Dentro de los objetivos del estudio podemos identificar tres:

1. Demostrar diferencias entre grupos o diferencias en un mismo grupo antes y después de una maniobra (por ejemplo, el tratamiento con el fármaco A reduce en mayor proporción la presión sanguínea que el tratamiento con el fármaco B).

Cuadro I | Peso de sujetos en estudio bajo dos esquemas terapéuticos

Grupo A			Grupo B		
77			65		
78			69		
80			77		
82			78		
85	83.5	Promedio	85	83.5	Promedio
85	85.0	Mediana	85	85.0	Mediana
85	85.0	Moda	89	85.0	Moda
86			93		
88			96		
89			98		

Las medidas de tendencia central son iguales, pero la dispersión de los datos es distinta

2. Mostrar relación (correlación) entre variables (por ejemplo, la creatinina sérica se eleva conforme disminuye la función renal).
3. Predecir un desenlace (por ejemplo, la probabilidad de que el sujeto con vida sedentaria y sobrepeso desarrolle diabetes mellitus tipo 2).

Frecuentemente se superponen los modelos, y es así que en ocasiones para demostrar diferencias entre dos grupos se hace uso de modelos inicialmente identificados para prede-

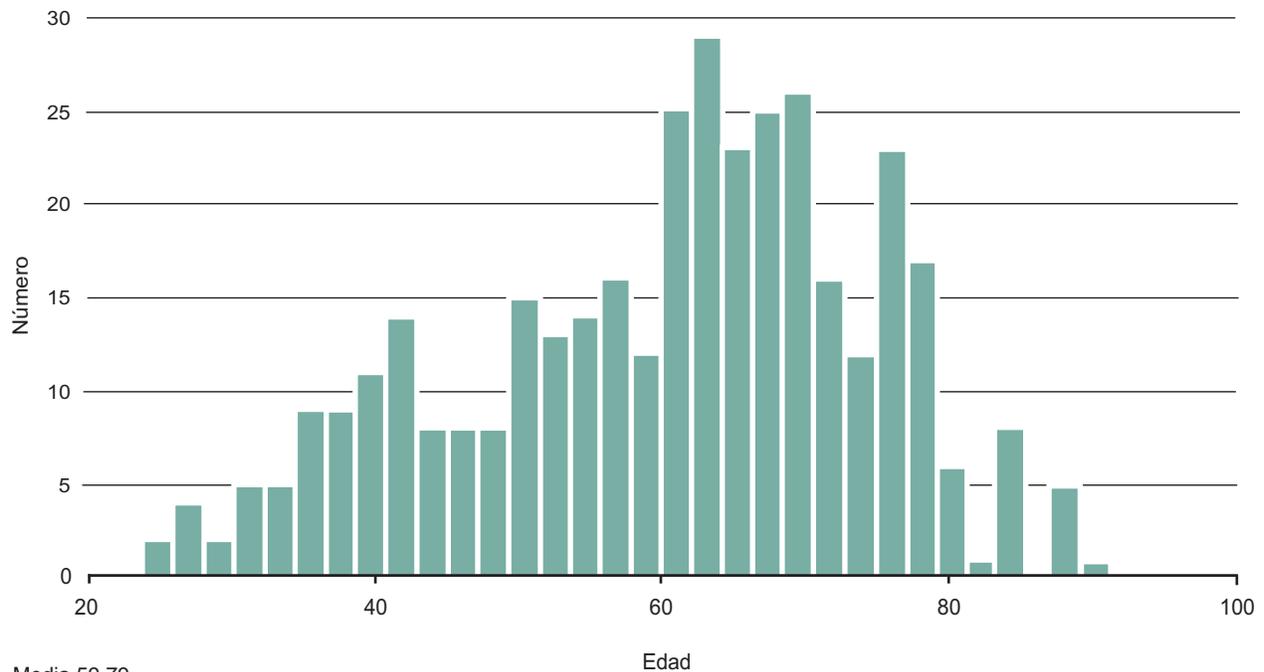


Figura 1 | Histograma

cir un desenlace; esto ocurre sobre todo cuando se requiere ajustar la maniobra principal (fármaco A versus fármaco B) por múltiples factores (edad, sexo, índice de masa corporal, etcétera). Pero también acontece el fenómeno opuesto, cuando se busca predecir un suceso que ocurrirá a futuro pero solo se cuenta con una o dos variables predictivas, en este caso se utiliza una prueba para demostrar diferencias.

Es importante aclarar que la correlación básicamente sirve para ver la magnitud de la asociación entre variables, aunque debe quedar claro que no determina causalidad, de hecho ninguna prueba estadística puede hacerlo, para ello se requiere cubrir una serie de principios descritos por sir Austin Bradford Hill.

Definir el *tipo de variable* resulta relevante debido a que es el eje para seleccionar la prueba apropiada dependiendo del objetivo buscado. Dentro del tipo de variable existen dos grupos:

1. *Cuantitativas continuas y discontinuas o discretas*: las primeras se caracterizan porque pueden tomar cualquier valor a lo largo de un continuo (por ejemplo, talla 1.75 m). Por su lado, las discontinuas o discretas utilizan solo números enteros (número de partos, 1, 2, 3...). En ambas, la distancia entre una unidad y otra a través de toda su escala es equidistante.
2. *Cualitativas*: dentro de éstas se encuentran la ordinal y la dicotómica. La *ordinal* permite dar un orden a la característica en estudio, y a diferencia de las cuantitativas la distancia entre dos categorías no es equidistante (por ejemplo, falla cardíaca grados I a IV). Las *dicotómicas*, como su nombre lo dice, son en las que solo hay dos categorías, que pueden ser binomiales (es una opción u otra, por ejemplo, hombre o mujer) o nominales (se refiere a la presencia o no de la característica, por ejemplo, vivo a seis meses, sí o no).

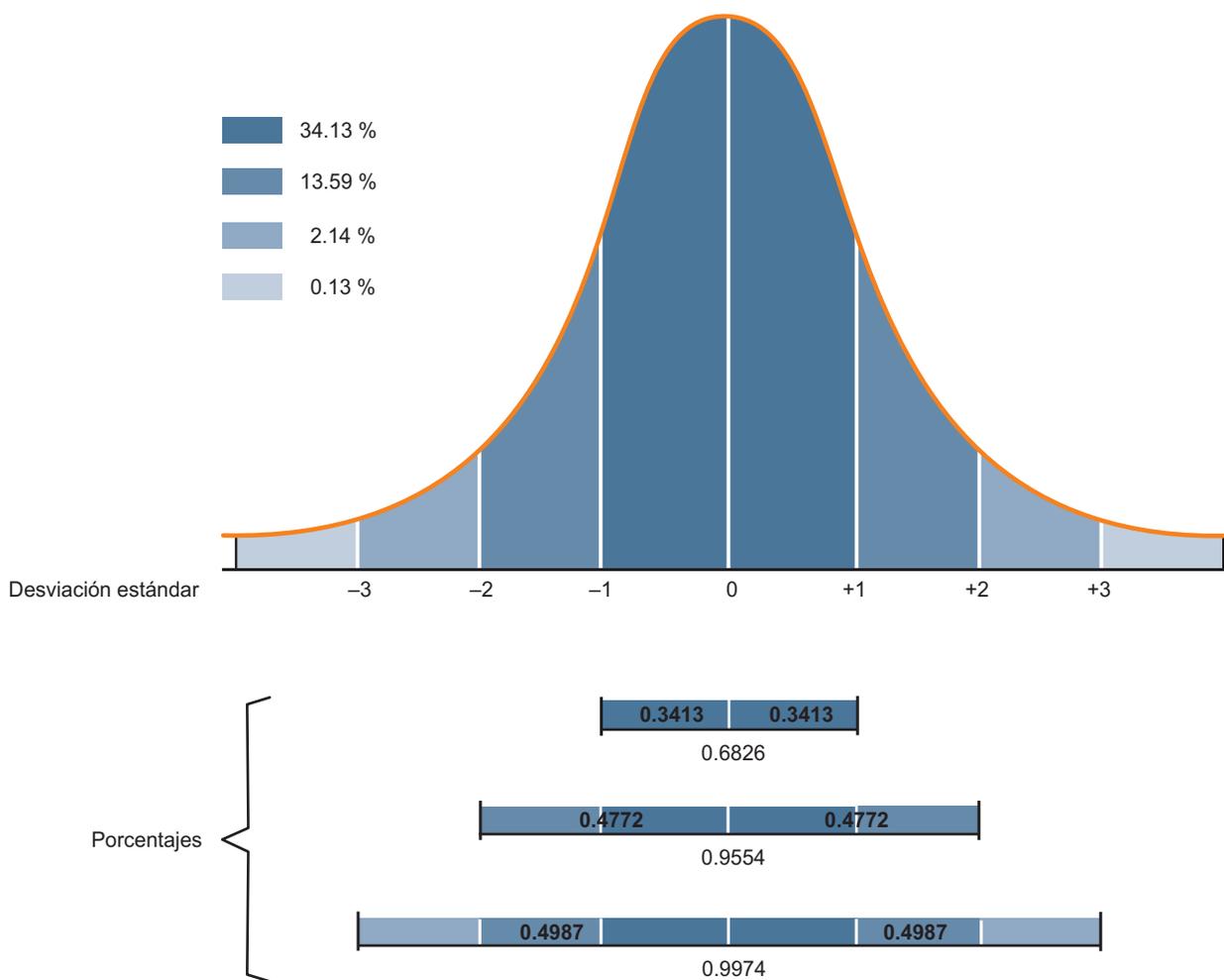


Figura 2 | Curva de distribución normal

Cuadro II Selección de la prueba estadística conforme el objetivo y el tipo de variable

Tipo variable	Tipo muestra	Demostrar diferencia		Mostrar relación ^{&}	Predecir 1 variable [‡]
		Dos grupos	Tres grupos	Dos variables	Variable desenlace
Cuantitativa (distribución normal)	NR	t Student*	Anova 1 factor	Pearson	Regresión lineal
	R	t Student**	Anova 1 factor		
Cualitativa ordinal	NR	U Mann-Whitney	Kruskal-Wallis	Spearman	
	R	Wilcoxon	Friedman		
Cualitativa dicotómica	NR	χ^2 (Fisher)	χ^2	Coeficiente phi	Regresión logística/ Curvas de supervivencia
	R	McNemar			

NR = no relacionada, R = relacionada, R = medición de la variable en el mismo sujeto en dos tiempos distintos
 *t de Student para muestras independientes
 **t de Student para muestras relacionadas
[&]Para la correlación entre 2 variables se utiliza la prueba de aquella en la escala inferior (en realidad ninguna escala es inferior, sin embargo, se han ordenado de las cuantitativas continuas a las dicotómicas, pasando por las cuantitativas discontinuas y las ordinales).
[‡]La variable predictiva puede ser cuantitativa, dicotómica u ordinal (estas últimas transformadas a variables tipo dummy)

Es importante mencionar el manejo que sufrirá el tipo de variable durante el proceso de análisis, empezando por la recolección de datos en forma “cruda”, lo que significa que solo se trata de una compilación de información de un grupo de sujetos. Para que estos datos tengan un significado de utilidad se necesita organizarlos y resumirlos. La forma más simple de organización son las tablas de distribución de frecuencia, sin embargo, en ocasiones es más fácil entender su representación gráfica a través de un histograma o polígono de frecuencias. No obstante la utilidad de esta información, se necesita información cuantitativa de los datos recolectados, es decir, se requieren *índices numéricos* que reflejen las distintas distribuciones de probabilidad, cuya función primordial es *modelar* el comportamiento de una gran variedad de fenómenos biológicos. Dentro de estos índices numéricos encontramos las medidas de tendencia central y las medidas de dispersión.

1. Medidas de tendencia central (cuadro I y figura 1).

- a) *Media*: es la suma de un conjunto de datos divididos por el número total de ellos. El símbolo para representar la media de una población es la letra griega μ , y la media de una muestra está representada por \bar{x} . Es la medida de resumen más utilizada para variables cuantitativas.
- b) *Mediana*: es el valor que se encuentra exactamente a la mitad del total de los datos. La mediana divide una distribución de datos ordenados en exactamente dos partes iguales. La ventaja de la mediana como medida de tendencia central es que ésta no se ve afectada por el valor de datos extremos, fenómeno que sí ocurre con la media. Es el tipo de medida de resumen más utilizado para

las variables cuantitativas que no siguen una distribución normal y para variables ordinales.

- c) *Moda*: se refiere al valor que más se repite en una distribución.

2. Medidas de dispersión más comunes.

- a) *Desviación estándar*: refleja la variación entre todo el conjunto de datos y se utiliza cuando éstos siguen una distribución normal.
- b) *Percentil*: describe la posición de un valor de la distribución. Se utiliza para variables cuantitativas que no siguen una distribución normal y para variables de tipo ordinal.
- c) *Rango*: diferencia entre el valor más alto y el valor más bajo de la distribución.

En investigación clínica, como en muchos otros fenómenos de la vida real, los datos más frecuentemente analizados son los cuantitativos, que en el más de los casos muestran una distribución gaussiana, también conocida como distribución normal, que se caracteriza por tener forma acampanada, ser simétrica respecto a su media, tener valores de frecuencia decrecientes conforme se alejan de la media y que nunca llegan a cero (asintótica), la moda y la mediana son iguales a la media, aproximadamente 68 % de los datos se encuentran en ± 1 desviación estándar respecto a su media y 95 % en ± 2 desviaciones estándar (figura 2). Es así que si la agrupación de datos son de tipo cuantitativo con distribución normal, su medida de resumen será la media y su medida de dispersión, la desviación estándar. Sin embargo, si su distribución no es gaussiana, al igual que para una variable de tipo ordinal, su

medida de resumen será la mediana y su medida de dispersión, el percentil o rango. Finalmente, habrá que mencionar que para las variables dicotómicas su medida de resumen es la frecuencia y su medida de dispersión la desviación estándar (su distribución se acerca a la normal).

Pertinencia de la prueba estadística

Una vez que conocemos nuestro objetivo y las características de nuestros datos (tipo de variable) podemos considerar la pertinencia de la prueba estadística (cuadro II). Sin embargo, existen dos consideraciones más cuando el objetivo sea demostrar diferencia:

1. Si se trata de un estudio donde se compara el valor de un dato antes y después de una maniobra ya sea observacional o experimental, se le conoce como de muestras relacionadas, pero si se trata de la comparación de datos entre distintos grupos se le denomina de muestras no relacionadas.
2. Si se trata de comparación entre distintos grupos es necesario determinar si será entre dos o más grupos.

Ya con la información completa, con el cuadro II podemos verificar si la selección de la prueba estadística fue la adecuado conforme la variable y el objetivo. Por ejemplo, si se compara la edad (variable cuantitativa con distribución normal en este caso), entre pacientes con lupus eritematoso sistémico, con y sin afección neurológica (dos grupos), la

prueba apropiada es la *t* de Student para muestras independientes. Pero si lo que se compara entre esos mismos es la frecuencia de mujeres (variable binomial), entonces la prueba estadística pertinente es la χ^2 . Si lo que se compara entre los dos grupos es su grado de actividad lúpica (escala ordinal), la prueba estadística pertinente es la *U* de Mann-Whitney. Por otro lado, si lo que nos muestran es la magnitud de la asociación (relación) entre edad (variable cuantitativa con distribución normal) y el grado de actividad lúpica (variable ordinal), la prueba pertinente es la *r* de Spearman. Finalmente, si lo que se busca es predecir el peso de un niño (variable cuantitativa) a partir de la edad (variable cuantitativa), tipo de alimentación (variable ordinal: buena, regular o mala) y el sexo (dicotómica), la prueba pertinente es la regresión lineal. Pero si lo que se quiere es predecir la probabilidad de infarto (dicotómica nominal) en los siguientes 10 años a partir de la edad (cuantitativa), el riesgo aterogénico (ordinal, bajo, medio y alto) y el sexo (dicotómica binomial), la prueba pertinente es la regresión logística múltiple.

Por último, esperamos que el presente artículo permita entender la razón de la selección de las pruebas estadísticas más empleadas en la investigación en salud y que, de igual manera, sirva de guía a quienes se inician en la estadística. No es suficiente para definir si los resultados obtenidos son reales, será necesario tener en consideración el diseño y ejecución del estudio y la estabilidad de los datos, pero este último aspecto merece ser analizado en otro apartado.

Bibliografía

1. Armitage P, Berry G, Matthews JNS. Statistical methods in medical research. 4th ed. Massachusetts, MA: Blackwell Publishing; 2002.
2. Bland M. Introduction to medical statistics. 3rd ed. Oxford: Oxford University Press; 2003.
3. Feinstein AR. Clinical epidemiology. The architecture of clinical research. Philadelphia, PA: W.B. Saunders; 1985.
4. Feinstein AR. Multivariable analysis: an introduction. New Haven, CT: Yale University Press; 1996.
5. Feinstein AR. Principles of medical statistics. New York, NY: Chapman and Hall/CRC; 2002.
6. Le Chap T. Introductory biostatistics. Hoboken, NJ: New Jersey: John Wiley and Sons; 2003.
7. Peat J, Barton B. Medical statistics. A guide to data analysis and critical appraisal. Malden, MA: Blackwell Publishing; 2005.
8. Portney LG, Watkins MP. Foundations of clinical research: applications to practice. 3rd ed. Saddle River, NJ: Pearson/Prentice Hall; 2009.