

METODOS ESTADISTICOS APLICADOS A LA INVESTIGACION PSICOLOGICA

Martha Vázquez Villazón y María Eugenia de Bernard Cerezo. Facultad de Psicología, Universidad de La Habana.

RESUMEN

En el presente artículo se discuten de manera general algunos aspectos importantes para la aplicación de los métodos estadísticos en la investigación psicológica. Se propone además, una clasificación útil para la selección del método estadístico más adecuado a un problema concreto.

ABSTRACT

In the present paper we discuss important aspects in the use of statistical methods for psychological research. We also offer a useful classification for the selection of the most adequate statistical method to analyse an specific problem.

En Psicología, el investigador utiliza de una manera particular, las palabras de uso común. Desde que se es estudiante se sabe que el término *personalidad* o *capacidad*, por ejemplo, implica mucho más que cuando alguien, desdeñosamente, afirma que tal persona *no tiene personalidad*, o cuando un jefe de Brigada le dice a un trabajador *tu eres un incapaz*. De manera semejante también es necesario comprender que el campo de la Estadística es completamente diferente de la idea común que se tiene de él.

Todos estamos familiarizados con la idea de que el estadístico se encarga de determinar, por ejemplo, *el promedio de niños con problemas de conducta en padres alcohólicos o no* o *el salario promedio de trabajadores con capacidad de rendimiento diferente*.

El investigador en Psicología debe saber ya, de una vez, que la DESCRIPCION es solamente una de las funciones de la estadística y, por cierto, no la más importante en el campo de la investigación psicológica. La otra, la INFERENCIA ESTADISTICA es la que permite que la investigación cobre un verdadero carácter científico.

La estadística descriptiva o deductiva, es la parte de la estadística que trata de la obtención y compendio de datos.

Ejemplo de esto son las encuestas sobre la base de cuestionarios y métodos gráficos.

La estadística inferencial o inductiva, es la parte de la Estadística que trata de la obtención de conclusiones, respecto a la fuente de datos.

La Inferencia Estadística estudia distintos problemas: la estimación de los parámetros de la población y las pruebas de hipótesis.

Como el objetivo final es hacer inferencia, o sea, obtener conclusiones, la parte descriptiva de la Estadística debe considerarse como preliminar en el análisis de los datos.

A la inferencia estadística le interesa sacar conclusiones sobre un gran número de acontecimientos, fundamentados en la observación de parte de ellos.

La estadística proporciona herramientas que formalizan y uniforman nuestros procedimientos para sacar conclusiones. Por ejemplo, podemos querer determinar cual es el mejor procedimiento de aprendizaje laboral de dos que se usan en las escuelas de capacitación para elevar la capacidad de rendimientos a obreros que no cumplen la norma de producción y, entonces, poder implantar el mejor en todas las escuelas de capacitación del país.

Podríamos reunir la información necesaria por distintas vías pero, decidir cuál de los dos métodos es el mejor, dependerá de los resultados de la prueba estadística utilizada. Sólo este resultado es el que indicará si las diferencias observadas entre estos dos procedimientos son tan grandes que podemos confirmar que representarán diferencias, respecto al grupo del que fueron muestradas sólo unos pocos acontecimientos y que verdaderamente estas no se deben al azar.

El psicólogo cuenta con un número grande de métodos estadísticos para resolver los distintos problemas que aborda la psicología, en la práctica.

Estos métodos los podemos dividir en paramétricos y no paramétricos, los primeros hacen suposiciones acerca de la naturaleza de la población, de la que se obtuvieron los puntajes, por ejemplo, una técnica de inferencia puede basarse en la suposición de que los puntajes se sacaron de una población distribuida normalmente o de que ambos conjuntos de puntajes se sacaron de poblaciones que tienen la misma varianza o dispersión de los puntajes.

Los segundos no hacen suposiciones tan severas, acerca de los parámetros de la población, a algunas de estas técnicas no paramétricas se les llama a menudo *PRUEBAS DE RANGO* o *PRUEBAS DE ORDEN*.

Cabría preguntarnos ahora cómo elegir una de estas técnicas como la más adecuada para una investigación concreta:

Un investigador dispone de numerosas pruebas estadísticas para un diseño de investigación particular; por tanto, es necesario emplear un criterio de elección adecuado si queremos estar seguros de plantear siempre una buena verificación de la hipótesis en cuestión.

Un criterio de elección es el *criterio de potencia*, que no es más que conocer la certeza de rechazar H_0 con una pequeña probabilidad de que sea verdadera y una probabilidad grande de rechazar H_0 siendo falsa.

Cuando tenemos un problema de decisión entre dos posibilidades (únicas y excluyentes), podemos tratarlo con las técnicas de las pruebas de hipótesis o docimasia de hipótesis, donde se definen:

- H_0 como la hipótesis nula
- H_1 como la hipótesis alternativa.

Una vez determinado que nuestro problema puede tratarse por docimasia y realizar la técnica, es posible equivocarnos en el resultado obtenido. Pueden cometerse dos tipos de errores, que son:

Error tipo I: que sea cierta H_0 y nos decidamos por H_1 , o sea, rechazamos una hipótesis que es verdadera.

Error tipo II. que sea cierta H_1 y nos decidamos por H_0 , o sea, aceptamos una hipótesis que es falsa.

Evidentemente, para que nuestra decisión sea lo más correcta posible, debemos tratar de eliminar o por lo menos minimizar estos errores:

En función de los tamaños de los dos tipos de errores, es posible introducir un sencillo principio a seguir en la determinación de buenas pruebas de hipótesis:

Entre todas las pruebas que posean el mismo tamaño de error de tipo I, se elegirá aquella para la cual el tamaño del error de tipo II, es lo más pequeño posible.

Sin embargo, hay otros criterios no menos importantes y que también hay que considerar en el momento de elegir la prueba estadística. Ellos son:

- la manera en que la muestra fue obtenida
- la naturaleza de la población de la que se sacó la muestra y
- la clase de medición o escala que se utilizó.

La muestra tiene que sustituir la población, tiene que representarla para los fines de la investigación. Por eso, en primer lugar, hay que eliminar del proceso de elección de la muestra, todos los efectos personales, impedir que las preferencias personales de quien va a realizar el muestreo, causen la elección de un número excesivo de elementos atípicos de la población. Además, para que los métodos de la teoría de probabilidades sean aplicables, hay que procurar que el proceso de la selección de la muestra sea comparable con un experimento aleatorio.

El método de elección de una muestra es un factor importante para saber qué uso puede hacerse de la muestra. Pueden obtener resultados muy erróneos si las muestras no se toman correctamente. Además, es necesario valorar el requisito de adecuación de los datos de la investigación a la prueba seleccionada, pues cada prueba tiene sus propias condiciones en las que es más poderosa y sin las cuales no se puede tener confianza en los resultados que se obtengan. Por ejemplo, si el tamaño de una muestra es tan pequeño como $n = 16$, se debe elegir una prueba estadística no paramétrica.

Por otra parte, las pruebas estadísticas que puedan usarse en una investigación, dependen de qué clase específica de población proviene la muestra. Hay poblaciones que se conciben formadas solamente por dos clases, digamos, masculino y femenino; otras que se clasifican en varias categorías; otras con valores continuos.

El estudio de pruebas que verifiquen si la muestra extraída viene de una población con una distribución especificada, puede ayudar a escoger la mejor prueba para los datos de una investigación.

Es igualmente importante también, la clase de medida que se utiliza en el momento de coleccionar los datos, porque según las características de la escala utilizada, se tendrá un criterio para la elección de una prueba estadística: Por ejemplo, los datos medidos por escalas de intervalos, deben analizarse mediante métodos paramétricos.

Considerar todos estos criterios asegurará que la elección de la prueba, sea la adecuada para el análisis de los datos de la investigación particular.

La naturaleza de la población y el método de muestras, establece el MODELO ESTADISTICO; cada prueba estadística implica un modelo y un requisito de medida.

A veces las condiciones del modelo se pueden verificar si satisfacen la prueba estadística elegida; pero a menudo las condiciones del modelo estadístico son SUPOSICIONES.

Por eso, cuando los datos de la investigación satisfacen las suposiciones de las pruebas paramétricas, las conclusiones serán más válidas y confiables.

Las condiciones en que una prueba paramétrica es más poderosa y sin las cuales no se puede tener confianza de cualquier conclusión obtenida con ella, son:

- Las observaciones deben ser independientes entre sí. La selección de un caso cualquiera de la población, incluido en la muestra, no debe afectar las posibilidades de incluir cualquier otro y el puntaje que se le asigna a un caso cualquiera, no debe influir en el puntaje que se le asigne a cualquier otro.
- Las observaciones deben hacerse en poblaciones distribuidas normalmente.
- Estas poblaciones deben tener las mismas varianzas (en casos especiales, deben tener una proporción de varianza conocida). Esto es importante ya que, por ejemplo, el problema de enseñar en un aula de niños con rendimientos muy homogéneos, es muy diferente al de enseñar en una clase en la que los rendimientos individuales son muy heterogéneos, aún cuando ambas clases sean de rendimiento promedio parecido. En este caso, para hacer investigaciones con ambas clases, es necesario que ellas sean bastante parecidas en cuanto a la diversidad de los rendimientos de los niños.
- En el caso del análisis de varianza (prueba F), las medias de las poblaciones deberán ser combinaciones lineales de efectos debido a las columnas y a las filas o ambas. Esto se debe a que cada observación tomada al azar se representa de la siguiente forma:

$$Y_j = X_{ij}\beta_1 + X_{2j}\beta_2 + \dots + X_{pj}\beta_p + e_j \quad \begin{matrix} i = 1, \dots, p \\ j = 1, \dots, n \end{matrix}$$

$$\text{donde: } \{ X_{ij} \} = \begin{cases} 1 & \text{si aparece } \beta_i \\ 0 & \text{si no aparece } \beta_i \end{cases}$$

β_i : son los efectos producidos por el tratamiento, son medias.

e_j : error no controlado

El modelo anterior permite representar cada observación como una combinación lineal en las constantes desconocidas $\{\beta_j\}$ más el error.

- Las variables que se analizan deben haberse medido, por lo menos en una escala de intervalo, de manera que sea posible hacer las operaciones de aritmética con los puntajes.

Cuando hay razones para creer que estas condiciones se satisfacen con los datos de la investigación particular, se escoge una prueba paramétrica para el análisis, pero cuando no se tiene la certeza de que los datos de la investigación cumplen esos requisitos se debe elegir una prueba no paramétrica.

De los requisitos vistos anteriormente, pensamos que se hace necesario detenernos en el análisis de la MEDICION DE LAS ESCALAS.

MEDIR es el proceso de asignar números o poner en correspondencia de uno a uno a objetos u observaciones.

La clase de medida que se obtiene es una función de las reglas bajo los cuales fueron asignados los números y, por supuesto, si se escoge una u otra medida, ellas definen las manifestaciones y operaciones que se puedan hacer con los puntajes.

Existen cuatro escalas generales de medición:

NOMINAL, ORDINAL, DE INTERVALO Y DE PROPORCION.

El nivel más bajo de medición es la clasificación de los objetos de acuerdo con una característica dada. Cuando los objetos que se estudian quedan clasificados por medio de una *partición*, se dice que tenemos una **escala nominal**.

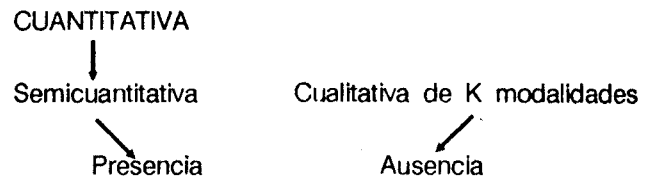
Cuando entre los objetos clasificados por medio de una partición existe, además, una relación de orden, se dice que tenemos una **escala ordinal**.

Si entre los objetos medidos, según una escala ordinal, esta es tal, que distancias iguales en la escala representan distancias iguales en la propiedad que se está midiendo, se dice, entonces que tenemos una **escala de intervalo**. Si esta posee un cero absoluto, con sentido empírico, o sea, si algo es cero en la escala, y esto quiere decir que el sujeto medido no tiene la propiedad que se observa, entonces se dice que tenemos una **escala de razón o de proporción**.

Las dos primeras son las más corrientemente utilizadas en la práctica psicológica.

El problema de las escalas de medición a veces se simplifican al tratarlo como un problema de codificación. Recordamos brevemente los diferentes tipos de codificaciones: cuantitativa, por una variable continua -- semi cuantitativa, por una variable cuantitativa repartida en clases o una variable observada de modalidades diferentes entre las que existe una relación de orden -- cualitativa o varias modalidades -- cualitativa en todo o nada: presencia (1) ausencia (0).

Se puede transformar una codificación cuantitativa en una codificación semicuantitativa, una codificación semicuantitativa en todo o nada, una codificación cualitativa de varias modalidades en todo o nada. El esquema siguiente resume estas posibilidades:



En ciertos problemas es útil trabajar con una codificación homogénea, a pesar de disminuir las variables cuantitativas en variables de presencia o ausencia.

METODO ESTADISTICO UNIVARIADO: PARAMETRICOS Y NO PARAMETRICOS

1. EL CASO DE UNA MUESTRA:

Cuando se desea probar una hipótesis en la se requiere una sola muestra, las pruebas estadísticas utilizadas indicarán si la muestra particular proviene de una población específica. Con ello se prueba la hipótesis de que la muestra fue extraída de una población con una distribución específica.

Entonces, la prueba estadística para el caso de una muestra, puede contestar a preguntas, como:

- ¿Hay una diferencia significativa de ubicación (tendencia central) w entre la muestra y la población?
- ¿Hay una diferencia significativa entre las frecuencias observadas y las frecuencias esperadas apoyándose en algún principio?
- ¿Es razonable creer que esta muestra fue sacada de una población, de forma o aspecto específico, por ejemplo, normal?

- ¿Es razonable creer que esta muestra fue sacada al azar de alguna población específica?

Los métodos estadísticos no paramétricos para este caso, son:

Escala de Medición	El caso de una muestra
Nominal	La prueba binomial La prueba χ^2
Ordinal	La prueba de Kolmogorov-Smirnov La prueba de rachas.

Tres de estas pruebas son del tipo de bondad de ajuste y la cuarta es de aleatoriedad de la serie de eventos en una muestra.

Al probar la hipótesis, acerca de si el origen de una muestra en una población con distribución específica, el investigador puede usar una de las tres pruebas de bondad de ajuste.

La técnica paramétrica correspondiente en el caso de una muestra, consiste en aplicar una prueba t o la diferencia entre la media observada (muestra) y la media esperada (población).

2. EL CASO DE DOS MUESTRAS:

Las pruebas estadísticas de dos muestras se usan cuando el investigador desea establecer la diferencia entre dos tratamientos o si un tratamiento es mejor que otro.

Por ejemplo, adiestramiento, uso de psicofármaco, en cada caso el grupo que ha sufrido el tratamiento es comparado con el que no lo ha experimentado o que ha sufrido un tratamiento diferente.

En la comparación de estos grupos, a veces se observan diferencias significativas que no son el resultado del tratamiento, por ejemplo, en el estudio de los trabajadores que se someten a un entrenamiento diferente para determinar cual es el mejor para elevar su calificación, puede ser que la diferencia no se deba, realmente, a uno u otro tratamiento, sino que uno de los grupos estaba más motivado por elevar rápidamente su calificación y, de esta forma, no se refleja verdaderamente la efectividad del procedimiento de enseñanza.

Una forma de eliminar esta dificultad, es usar MUESTRAS RELACIONADAS estas se pueden lograr:

- Cuando el propio sujeto es su propio control.
- Con parejas de sujetos en las que se asignan los miembros de cada pareja, a las dos condiciones.

La técnica paramétrica usual para analizar datos provenientes de dos muestras relacionadas es aplicar la prueba t a los puntajes; estos se pueden obtener de los dos puntajes de cada pareja igualada o de los puntajes de cada sujeto bajo las dos condiciones.

Quando estudiamos las diferencias entre dos grupos, podemos también usar grupos independientes, las pruebas estadísticas para dos grupos independientes, pueden utilizarse en diseños de este tipo.

Estas pruebas determinan la medida en que las diferencias de las muestras indican, de forma convincente, una diferencia en el proceso aplicado en ellos.

A pesar de las ventajas de usar muestras relacionadas, a veces es poco práctico, frecuentemente la naturaleza de la variable dependiente impide usar a los sujetos como sus propios controles, por ejemplo, si la ORIENTACION VOCACIONAL INFLUYE en el rendimiento docente universitario.

Tampoco se puede relacionar las muestras cuando el investigador no sabe cómo confeccionar los pares, adecuadamente, o porque sencillamente no hay buenos pares disponibles.

En el caso de dos MUESTRAS INDEPENDIENTES, ellas pueden obtenerse:

- Tomando al azar sujetos de dos poblaciones.
- Asignando al azar ambos tratamientos a miembros de algunas muestras de orígenes arbitrarios.

No es necesario que la muestra sea del mismo tamaño.

En este caso, la prueba t es la técnica paramétrica indicada para analizar los datos de las dos muestras independientes.

Los métodos estadísticos no paramétricos adecuados para estos casos, son:

Escala de medición	Caso de dos muestras: relacionadas MCNEMAR	Independientes
Nominal	La prueba de MCNEMAR	La prueba de la probabilidad exacta de Fisher. La prueba de X^2 .
Ordinal	La prueba de los signos La prueba de los rangos señalados y pares iguales de Wilcoxon.	La prueba de la mediana. La prueba U de Mann-Witney. La prueba de Kolmogorov-Smirnov. La prueba de rachas de Wald.
Intervalo	La prueba de Walsh Prueba de aleatoriedad para pares igualados.	Prueba de aleatoriedad para dos muestras independientes.

3. EL CASO DE K MUESTRA:

Hasta aquí hemos visto las pruebas estadísticas idóneas para probar la significación de las diferencias.

- Entre una sola muestra y una población determinada.
- Entre dos muestras relacionadas o independientes.

Ahora veremos las pruebas que determinan la significación de las diferencias entre 3 o más grupos, relacionados o independientes.

A veces las circunstancias requieren de diseños experimentales de más de dos muestras o condiciones

que puedan estudiarse simultáneamente y entonces es necesario usar una prueba estadística que indique si existe una diferencia total entre las k muestras o condiciones, ya que no es posible tener confianza en una decisión acerca de k muestras, en la que el análisis se haga probando las muestras, 2 a 2.

La técnica paramétrica para probar si varias muestras proceden de una misma población, es el análisis de varianza o prueba F. La misma facilita que no haya pérdida de precisión alestimar la varianza por separado, pues se utiliza una varianza combinada.

En el caso no paramétrico, tenemos:

Escala de medición	Caso de K muestra Relacionadas	Independientes
Nominal	La prueba Q de Cocran	X^2
Ordinal	Análisis de varianza de dos clasificaciones por rango, de Friedman.	Extensión de la prueba de la mediana. Análisis de varianza, de una clasificación por rangos, de Kruskal-Wallis.

4. LAS MEDIDAS DE CORRELACION Y SUS PRUEBAS DE SIGNIFICACION

En la investigación psicológica, frecuentemente se desea saber si dos conjuntos de puntajes están relacionados, o conocer el grado de su relación.

Establecer la correlación entre dos variables, puede ir encaminada en dos sentidos:

- Ser la meta última de una investigación, por ejemplo, si hay relación entre la asistencia y participación a los círculos de interés y los resultados docentes en los estudios universitarios.
- O ser un paso dentro de otra investigación que tiene otros fines, como es el caso de cuando aplicamos la correlación en el psicodiagnóstico laboral, solamente para conocer la concordancia de los juicios emitidos por 5 jueces, al evaluar la eficiencia laboral en un grupo de trabajadores. Esto sólo es un paso dentro de la investigación total, pues solo se refiere a la variable *Criterio de validez*.

Cada una de las medidas de correlación no paramétricas que más abajo se mencionan, llevan

relacionadas las pruebas estadísticas que determinan la probabilidad asociada con la ocurrencia de una correlación entre dos conjuntos de atributos, es tan grande como la observada en la muestra, conforme a una H_0 que supone que las variables no están relacionadas en la población.

Si de interés para el psicólogo es indicar el grado de asociación entre dos conjuntos de puntajes de un grupo dado de sujeto, mucho más importante aún es afirmar si alguna relación en una muestra de puntaje, indica que las variables en estudio están asociadas muy probablemente en la población de la que se tomó la muestra.

El coeficiente de correlación representa el grado de relación entre dos variables, las pruebas de significación del coeficiente determinan en un nivel de probabilidad declarada, si la asociación verdaderamente existe en la población, de la que se tomó la muestra.

En el caso no paramétrico, los coeficientes de correlación, son:

Escala de Medición	Coefficiente de correlación
Nominal	El coeficiente de Contingencia C.
Ordinal	El coeficiente de correlación de rango, de Sperman. El coeficiente de correlación de rango, de Kendall. El coeficiente de correlación parcial de rango, de Kendall. El coeficiente de concordancia, de Kendall.

REFERENCIAS

1. Dixon - Marsey (1971). **Introducción al Análisis Estadístico**. Ed. R.
2. Garret (1974). **Estadística en Psicología y Educación**. Ed. Paidés 33 Edición.
3. Houli, P. C. (1971). **Estadística Elemental**. E.D., R.
4. Siegal, S. (1970). **Diseño Experimental no Paramétrico**. Ed. R.