

INFERENCIA ESTADÍSTICA: CONTRASTE DE HIPÓTESIS

Objetivo: Seleccionar la prueba de contraste de hipótesis en ejemplos dados

INTRODUCCIÓN:

Al realizar una investigación, el investigador usualmente se plantea una hipótesis; si estudia a toda la población podrá aceptar o rechazar dicha hipótesis con toda certeza, pero si no puede estudiar la población total sino una muestra, entonces deberá seguir un proceso por medio del cual decidirá si aceptar o no su hipótesis con un riesgo conocido de estar equivocado. En eso consiste el proceso de Contraste de Hipótesis, aspecto cuyo aprendizaje reviste gran importancia no solo para quienes realizan investigación sino también para los profesionales que, para actualizar sus conocimientos, requieren de lecturas de artículos científicos en su especialidad. En este material se hace una breve explicación de los fundamentos del contraste de hipótesis y se exponen los elementos que deben ser considerados para seleccionar la prueba estadística adecuada a los datos, acompañando la lectura de ejercicios que deberán ser resueltos y discutidos con el facilitador para asegurar el aprendizaje deseado. Se anexa a la guía un material complementario donde se presentan varios casos de investigaciones con sus datos respectivos y se detallan para cada uno de ellos, todos los pasos realizados para verificar si la hipótesis planteada puede o no ser aceptada.

EL PROCESO DE CONTRASTE DE HIPÓTESIS:

El primer paso de este proceso consiste en formular las hipótesis estadísticas: hipótesis nula (H_0) y alternativa (H_A).

La hipótesis nula representa la negación de la hipótesis de investigación, como se verá posteriormente. Luego se hacen las pruebas de significación estadística (otro paso del proceso), el investigador acepta o rechaza la hipótesis nula formulada y en consecuencia, estará rechazando o aceptando su hipótesis de investigación.

En este proceso pueden cometerse dos tipos de error:

1. Que la hipótesis nula sea verdadera y el investigador la considere falsa (Error Tipo I ó α)
2. Que la hipótesis nula sea falsa y el investigador la considere verdadera (Error tipo II ó β)

Sin embargo, el trabajar con una muestra probabilística le permite al investigador conocer la probabilidad de equivocarse al aceptar su hipótesis.

Esto se logra cuando la persona fija el nivel de confianza con el cual desea trabajar (el P_k), pues como ya habíamos visto en la lectura anterior, al fijar el P_k automáticamente se está determinando el nivel de significación α el cual no es más que la probabilidad de equivocarnos al rechazar una hipótesis nula que es verdadera (α es la probabilidad de cometer el error tipo I), es decir, aceptar una hipótesis de investigación falsa.

En Medicina, generalmente se acostumbra trabajar con P_k que van de 0.95 a 0.99, por lo tanto los niveles de significación oscilan entre 0.05 y 0.01.

Si se rechaza una hipótesis a un nivel $\alpha = 0,05$, existe un 5% de posibilidades de que esté equivocado. Ese 5% puede disminuirse pero a expensas de un aumento en la probabilidad de cometer el error tipo II (aceptar como cierta una hipótesis que no lo es).

Pasos del Proceso de Contraste de Hipótesis

- a) Formular las hipótesis estadísticas
- b) Fijar los niveles de confianza
- c) Decidir el tipo de prueba de significación a emplear y ejecutarla
- d) Tomar la decisión de aceptar o rechazar la hipótesis.

a) Formular las hipótesis estadísticas:

- ***Hipótesis Alternativa (H_A):*** es la misma hipótesis de investigación, pero planteada en términos estadísticos, por lo tanto debe decidirse de antemano qué medidas se van a usar para resumir las variables en estudio. Por ejemplo, si la hipótesis de investigación plantea que el tiempo que tienen trabajando los controladores aéreos influye en la aparición de hipertensión arterial diastólica (HAD), la hipótesis alternativa podría ser:

H_A : "El promedio de años de trabajo en los controladores aéreos que padecen de HAD es significativamente mayor que el promedio de años de trabajo en aquellos que no tienen HAD".

En términos estadísticos esta hipótesis se representa de la siguiente manera:

$$H_A: \mu_1 > \mu_2$$

- **Hipótesis Nula (H_0)**: es la negación de las diferencias planteadas en la hipótesis alternativa.

H_0 : "No existe diferencia significativa entre el promedio de años de trabajo en los controladores aéreos que padecen de HAD y el promedio de años de trabajo en aquellos que no tienen HAD"

$$H_0: \mu_1 = \mu_2 \quad \text{o también} \quad H_0: \mu_1 - \mu_2 = 0$$

Con las pruebas de contraste de hipótesis, el investigador lo que trata es de rechazar la hipótesis nula, pues así confirma su hipótesis de investigación.

EJERCICIO 1:

- 1- Supóngase que se desea comparar la efectividad de los fármacos A y B en el tratamiento del cáncer. Los resultados se medirán en tiempo de sobrevivencia.

La hipótesis de investigación podría ser:

La hipótesis alternativa sería:

La hipótesis nula sería:

- 2- En un hospital se lleva a cabo una experiencia en la cual a un grupo de pacientes se trata con la psicoterapia clásica y a otro grupo con una nueva técnica. Luego se comparan los porcentajes de pacientes curados en cada grupo. Para ese ejemplo hipotético, la hipótesis de investigación sería:

La hipótesis alternativa sería:

La hipótesis nula sería:

b) Fijar los niveles de confianza y de significación a emplear.

Este paso ya fue expuesto anteriormente. El nivel de significación no es más que la probabilidad de rechazar una hipótesis nula que es verdadera.

EJERCICIO 2:

- 1- Para los ejercicios resueltos anteriormente, ¿Cómo interpretaría usted el hecho de que se rechazara la hipótesis nula a un nivel de significación de 0.01?

c) Escoger la prueba de significación a emplear e interpretarla:

Muchas de las investigaciones que se realizan pretenden comparar un valor o medida obtenido de una muestra, con un valor teórico; o bien, comparar dos o más grupos para saber si existe diferencia entre ellos.

Se toma una muestra de cada grupo poblacional y a cada uno se le calculan las medidas de resumen que se desean comparar.

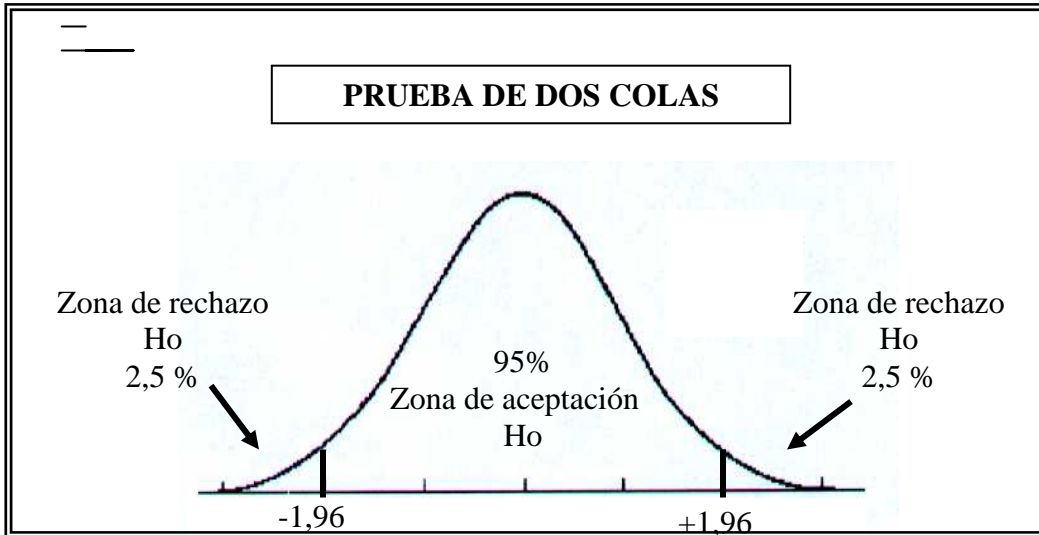
Si la hipótesis de investigación plantea que los grupos difieren en cuanto a la variable estudiada, la hipótesis nula expresará que no existe diferencia significativa entre los grupos, resumidos bien sea por promedios, porcentajes, etc; es decir: la hipótesis nula plantea que la diferencia observada en la muestra obedece al azar y no se da igualmente en la población (la diferencia no es significativa).

Para poder rechazar esa hipótesis se debe realizar una prueba estadística de significación.

Muchas de estas pruebas se basan en el supuesto de que la población estudiada tiene una distribución normal y, por lo tanto, las propiedades de la curva normal son aplicables.

Este tipo de prueba consiste, en líneas generales, en lo siguiente:

- 1) Se calcula un valor con la (s) muestra (s) y ese valor se lleva a puntaje Z usando las fórmulas indicadas
- 2) Se fija un nivel de significación α el cual determina en la curva normal dos zonas: una zona de aceptación de la hipótesis nula y una zona de rechazo de la hipótesis nula.



Para un nivel de significación α : 0.05, la zona de aceptación de la H_0 está comprendida entre los valores Z: -1,96 y +1,96.

- 3) Se ve en qué zona cae el valor Z calculado con los datos de la muestra. Si dicho valor cae en la zona de aceptación, entonces no se puede rechazar la hipótesis nula. Ejemplo:

Supóngase que se desea saber con un 95% de certeza si el promedio de 74 pulsaciones por minuto observado en una muestra de 36 individuos normales se diferencia significativamente del valor de 70 pulsaciones, considerado como normal. El error estándar de la media es 2,11.

Nótese que lo que se desea saber es si esa muestra formada procede de la población cuyo promedio se conoce, o por el contrario, procede de otra población.

Datos: $\bar{x} = 74$
 $\mu_0 = 70$
 $n = 36$
 $\sigma_{\bar{x}} = 2,11$

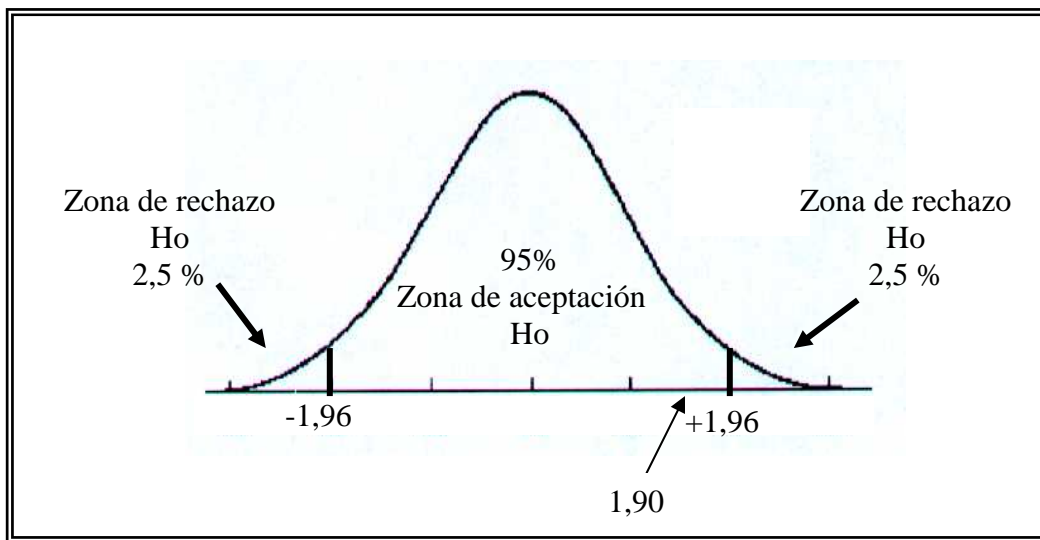
Procedimiento:

- Se calcula la diferencia entre los 2 valores y se transforma en puntaje Z, utilizando la fórmula:

$$Z_c = \frac{\bar{x} - \mu_0}{\sigma_{\bar{x}}} \qquad Z_c = \frac{74 - 70}{2,11} = 1,90$$

- Se definen las zonas de aceptación y rechazo de la hipótesis nula:

Como se desea una certeza de un 95%, el nivel de significación es 0,05 y para este nivel Z_0 : 1,96.

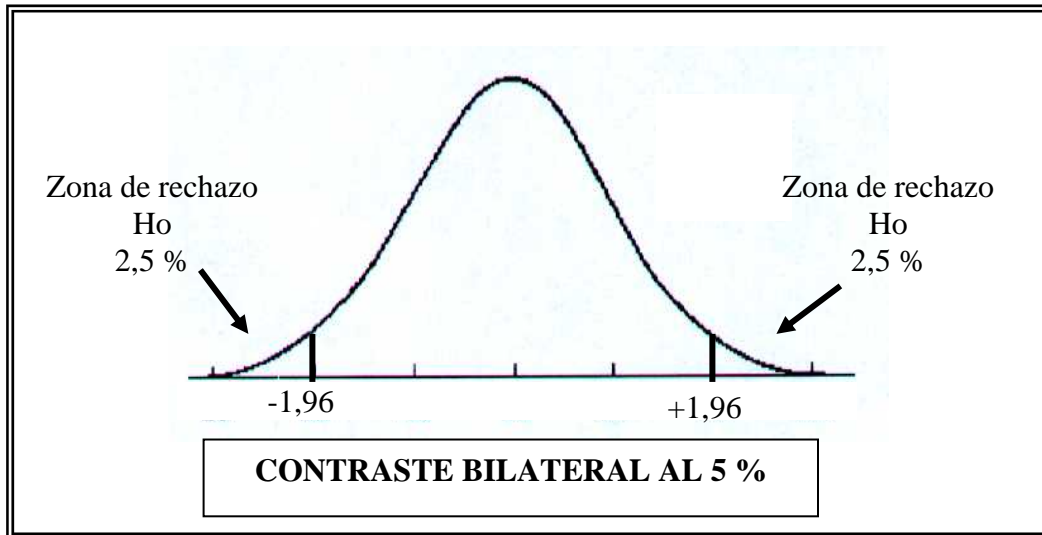


- Puesto que el valor de Z calculada cae dentro de los límites de la zona de aceptación de H_0 , esta hipótesis no puede ser rechazada → no existe diferencia significativa entre el promedio y el valor teórico con el cual se compara.

PRUEBAS UNILATERALES O BILATERALES (de una o dos colas)

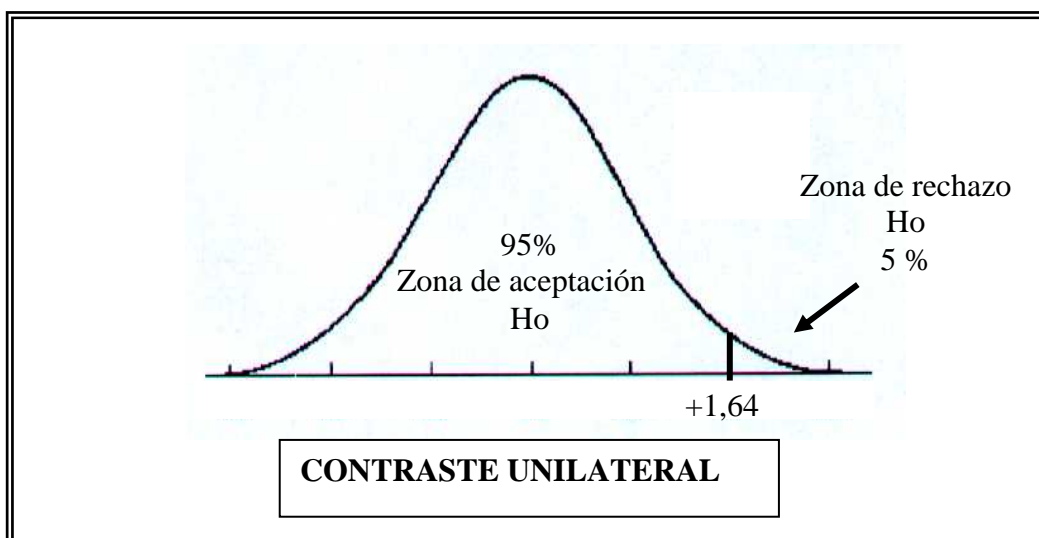
Como ya se vio, el área a ambos lados de la curva normal es una de las cantidades que participan en las pruebas de decisión o contraste. Esta área está definida por el nivel de significación α fijado. A cada α fijado le corresponde un valor Z ó t. Si $\alpha = 0,05$, $Z = 1,96$.

Cuando se realiza un contraste de significación, generalmente se trata de comparar si existe o no una diferencia significativa entre dos medias; si la primera media es mayor o menor que la segunda no importa, porque lo importante, solamente, es la diferencia. Esto es entonces, un contraste bilateral.



Cuando se puede establecer la dirección de la diferencia, es decir, cuando se puede establecer la hipótesis de que una media es mayor (o menor) que otra, pues se tiene conocimiento de la materia, entonces se decide trabajar con pruebas de una sola cola (unidireccionales).

La diferencia está en que ahora la zona de rechazo de la hipótesis nula se ubica en un solo extremo de la curva y, por lo tanto, el valor Z es diferente.



Valores de Z en pruebas unilaterales y bilaterales para diferentes niveles de significación

Nivel de Significación	0.10	0.05	0.01	0.005	0.002
Valores de Z en pruebas de 1 cola	1.28	1.64	2.33	2.58	2.88
Valores de Z en pruebas de 2 colas	1.64	1.96	2.58	2.81	3.08

SELECCIÓN DE LA PRUEBA DE SIGNIFICACIÓN:

Para escoger la prueba de significación adecuada, se deben considerar los siguientes aspectos:

a) **Nivel de medición de las variables:** El nivel de medición de las variables determina el tipo de estadístico a emplear (promedio, porcentajes, etc) y éste a su vez determina la prueba de significación.

b) **El número de grupos que se comparan:**

Existen pruebas para comparar 2 grupos y pruebas para comparar más de 2 grupos. Así como también, para comparar un valor de una muestra con un valor teórico.

c) **La independencia de los datos:**

Cuando a una misma persona se le realizan 2 medidas para compararlas (Ejemplo tensión arterial antes y después de un medicamento) se dice que son datos no independientes o correlacionados. También cuando se usan pares de sujetos (Se pueden usar sujetos apareados naturalmente como por ejemplo, 2 ratones de la misma camada y uno de ellos es el control. También si se aparean artificialmente los sujetos del estudio como cuando se usan pares de la misma edad y sexo).

Existen pruebas para datos correlacionados y para datos no correlacionados.

d) **Si se conoce o no la desviación estándar de la población (o la varianza).**

e) **Tamaño de la(s) muestra(s):**

Para muestras grandes la distribución de muchos estadísticos es normal y su interpretación se hace como en los ejemplos expuestos (utilizando la tabla de áreas de la curva normal para hallar los valores Z).

En muestras pequeñas la distribución de los estadísticos no es igual, por lo tanto no puede hablarse de valores Z ni utilizarse éstos en las pruebas; para estos casos existe una tabla, la tabla de la “ t ” de Student.

A continuación se presenta un esquema con las pruebas a usar en cada caso.

EJERCICIO 3:

A continuación se presentan varios ejemplos de investigación. Para cada uno de ellos construya su sistema de hipótesis e indique cuál es la prueba de significación estadística apropiada.

- 1) El promedio de pulsaciones por minuto para una muestra de alumnos varones del primer semestre de Medicina, fue 68,7 y la desviación estándar fue 8,67. Las normas clínicas indican que la media de pulsaciones en varones adultos es de 72 por minuto. Se desea saber si esa diferencia es significativa.
- 2) Para verificar el valor de la hidroclorotiacida sobre la presión sanguínea sistólica en pacientes hipertensos, se administró a un grupo de 40 pacientes un placebo, y un mes más tarde recibieron la hidroclorotiacida, registrándose los datos de presión arterial con ambos productos.
- 3) De entre todos individuos susceptibles expuestos a un determinado agente infeccioso 36% generalmente contrae un estado de enfermedad clínica. En un grupo de escolares compuesto por 144 individuos de los que se sospecha estuvieron expuestos a ese agente, sólo 35 contrajeron la enfermedad. Se desea saber si este resultado obedece a una variación fortuita.

BIBLIOGRAFÍA QUE PUEDE CONSULTAR PARA AMPLIAR SOBRE EL TEMA

- Sampieri, R, Collado, C, Lucio, P.. Metodología de la Investigación, 2da. Edición. Mc Graw Hill. 1991
- Dawson, S. Bioestadística Médica. Manual Moderno, México DF, 1993.
- Scheffler. Bioestadística. Fondo Educativo Interamericano, SA. México DF, 1989.
- Rieglman, R; Hirsch, R. Cómo Estudiar un Estudio y Probar una Prueba. Lectura Crítica de la Literatura Médica. OPS, 1992.