

**(1994-
2024)**

30 años de la
Consagración Constitucional
de la Autonomía y Autarquía
Universitaria en Argentina.



UNIVERSIDAD NACIONAL DEL LITORAL

FACULTAD DE INGENIERÍA Y CIENCIAS HÍDRICAS

ESTADÍSTICA

NOTAS DE INTRODUCCIÓN A LA TEORÍA

UNIDAD 6 - PARTE C - ESTADÍSTICA INFERENCIAL PRUEBAS DE HIPÓTESIS

Responsable de cátedra: Prof. Juan Pablo Taulamet

Equipo de cátedra: **Auxiliares:** Lic. María José Llop (JTP) - Ing. Ana Lisa Eusebi (JTP) - Prof. Fátima Bolatti (JTP) - Ing. Franco Nardi (Ay. 1°) **Ayudantes:** AIA Cristian Bottazzi - Téc. Eliana García

Carreras: Ingenierías en: Recursos Hídricos - Ambiental - Agrimensura

AÑO ACADÉMICO 2024

TIPOS DE PRUEBAS

Las pruebas de hipótesis, son una herramienta para la toma de decisiones referidas a una población a partir de una muestra. En este material se abordarán sintéticamente dos tipos de pruebas: paramétricas y no paramétricas. Las primeras se encargarán verificar afirmaciones acerca de algún parámetro de una determinada población, mientras que las segundas realizarán una afirmación referida a la bondad del ajuste de los datos con respecto a alguna distribución de probabilidades.

PRUEBAS PARAMÉTRICAS

Se trata de un tipo de prueba que propone un par de afirmaciones referidas a un parámetro, es decir al valor a alguna característica de la población, de la cual se dispone una muestra. De esta forma, se utilizará lo estudiado acerca de la distribución por muestreo de los estimadores para formular un par de afirmaciones, -habitualmente denominadas hipótesis-, de las cuales se intentará probar la veracidad con un cierto nivel de significación.

LA HIPÓTESIS NULA

El primer paso para la construcción de una prueba de hipótesis paramétrica, consiste en formular un supuesto, conocido como hipótesis nula acerca de un parámetro, a modo de igualdad o desigualdad con un valor de la variable, de tal forma que si se toma como ejemplo el parámetro μ y el valor 42, podría formularse la hipótesis que afirma que μ es igual a 42, simbólicamente:

Ejemplo 1: $H_0 : \mu = 42$

De manera similar al ejemplo anterior, podrían proponerse las siguientes hipótesis nulas, en relación al valor 4:

Ejemplo 2: $H_0 : \mu \leq 42$

Ejemplo 3: $H_0 : \mu \geq 42$

Es importante que la hipótesis nula incluya siempre la oportunidad de que la VA tome un valor en particular y ese valor puede simbolizarse como μ_0 , de tal forma que en cualquiera de los casos anteriores $\mu_0 = 42$.

LA HIPÓTESIS ALTERNATIVA

Luego de formular la H_0 referida a algún parámetro, formulará una segunda hipótesis, de tal forma que se antagónica con la anterior. Se busca que las mismas se comporten como eventos mutuamente excluyentes. Esta nueva hipótesis recibe el nombre de alternativa y

se denota con el símbolo H_1 . A partir del primer ejemplo 1, planteado de H_0 , se ofrecen a su vez los tres ejemplos posibles de hipótesis alternativas que se podrían proponer:

Ejemplo 4: $H_1 : \mu \neq 42$ Ejemplo 5: $H_1 : \mu < 42$ Ejemplo 6: $H_1 : \mu > 42$

Si se analizan los ejemplos 3 y 4, podrá concluirse que si se buscan alternativas que sean antagónicas, sólo podrán considerarse las vinculadas al ejemplo 6 y 5 respectivamente. En definitiva los pares de hipótesis podrían ser:

$$H_0 : \mu \leq 42$$

$$H_1 : \mu > 42$$

o bien,

$$H_0 : \mu \geq 42$$

$$H_1 : \mu < 42$$

Dependiendo de la expresión que tome la H_1 , la prueba podrá ser considerada bilateral (cuando la misma incluya el símbolo distinto) o unilateral cuando utilice los símbolos mayor o menor. De esta forma si se considera el ejemplo 4, se trata de un caso bilateral mientras que en los ejemplos 5 y 6 se trata de casos unilaterales. A su vez, los casos unilaterales recibirán una denominación dependiendo del símbolo, cuando se trate de un símbolo mayor, se denominará prueba unilateral por cola derecha y en caso contrario prueba unilateral por cola izquierda.

NIVEL DE SIGNIFICACIÓN

El nivel de significación de la prueba, será la probabilidad elegida asociada al error cometido si se rechaza H_0 cuando en realidad es cierta. Se suele utilizar valores como 1%, 5% y 10% y se denota con el símbolo α .

OBJETIVO DE LA PRUEBA

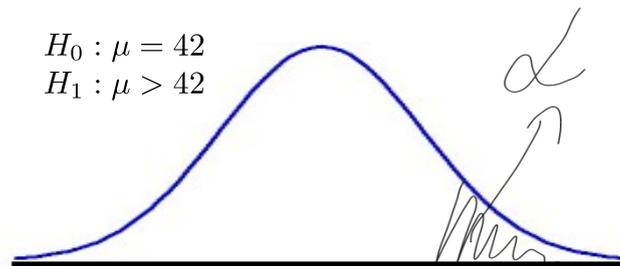
Una vez establecidas las hipótesis nula y alternativa de la prueba, el objetivo será establecer cuál de las afirmaciones vinculadas con las mismas puede ser considerada cierta. Naturalmente, por tratarse de afirmaciones que representan eventos excluyentes, no podrán ser ambas ciertas a la vez. Esto es: si H_0 es cierta, H_1 no lo será y viceversa.

El procedimiento que se propone para alcanzar el objetivo es el siguiente: Se supone que la H_0 es cierta, pero se intentará encontrar evidencia para rechazarla, y dar por cierta entonces la H_1 .

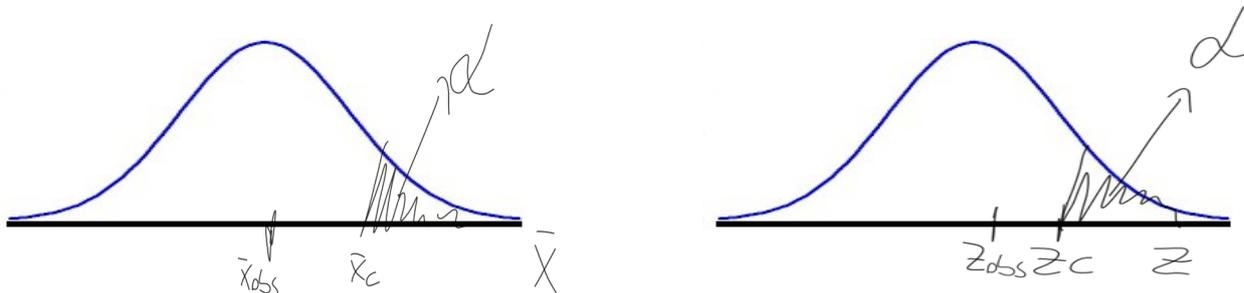
REPRESENTACIÓN GRÁFICA DE LA PRUEBA

La representación gráfica de la prueba se realiza esquematizando la función de densidad del estimador y rayando la probabilidad correspondiente al nivel de significación del lado

correspondiente a lo que representa la H_1 . En el caso de de las pruebas bilaterales se representará la mitad de α en cada cola y en las unilaterales α estará de un solo lado. Es por ello que a menudo se utiliza el nombre de pruebas de dos colas para el caso bilateral y pruebas de una cola para los casos unilaterales. La zona rayada que se corresponde con el nivel de significación recibe el nombre de zona de rechazo y la zona contraria zona de no rechazo. El valor que marca el inicio de cada cola rayada recibe el nombre de valor crítico y se encuentra asociado a un percentil de la VA estimadora en función del nivel de significación y lo desarrollaremos a continuación. Otro valor que suele representarse en el esquema es el valor del estimador puntual que recibe el nombre de estadístico de la prueba. A continuación se presenta un ejemplo de representación gráfica, para una prueba de igual contra mayor, con un nivel de significación del 5%.



La línea azul del gráfico anterior, representa la función de densidad de la VA estimadora y tal como hemos visto, en el caso de la \overline{X} , podría corresponder a la Distribución Normal o a la Distribución t de Student, dependiendo de los supuestos y el tamaño de la muestra n . A su vez, si se supone que la media muestral sigue una distribución Normal con esperanza igual a μ_0 y desvío $s/\sqrt{(n)}$. A continuación se presentan dos representaciones gráficas que consideran las referencias correspondientes a la VA estimadora, en dos versiones, una con la VA \overline{X} y otra con la versión estandarizada Z .



En los esquemas anteriores el símbolo X_{obs} refiere al valor de la media muestral observado en la muestra, es decir el estimador puntual y el símbolo Z_{obs} refiere a su versión estandarizada. Por otra parte el símbolo X_c refiere al valor crítico de la prueba, es decir a aquel valor a partir del cual comienza la zona de rechazo y Z_c es su versión estandarizada.

TOMA DE DECISIÓN DE LA PRUEBA

La toma de decisión de la prueba, tiene el objetivo de determinar si se ha encontrado evidencia para rechazar la H_0 , que había sido supuesta como cierta. Para realizar la toma de la decisión existen varias formas posibles de comparación. Una forma sencilla puede realizarse a partir de la representación gráfica de la prueba y consiste en evaluar si el valor puntual del estimador se encuentra en la zona de rechazo. Otra forma de tomar la decisión es calcular una probabilidad denominada P-valor y compararla con α . trabajaremos esta cuestión en detalle más adelante.

DETERMINACIÓN DEL VALOR CRÍTICO

El valor crítico de una prueba se encuentra vinculado con un percentil cuyo orden depende de la cantidad de colas, de la ubicación de la zona de rechazo y del nivel de significación. Por ejemplo si se plantea un nivel de significación del 5% y una prueba por cola derecha, el valor crítico será el percentil de orden 95.

DETERMINACIÓN DEL P-VALOR

El p-valor es uno de los valores más importantes de una prueba de hipótesis, ya que sintetiza la información correspondiente a la VA estimadora, considerando la distribución de probabilidades y la hipótesis alternativa. El p-valor representa una probabilidad que se encuentra del mismo lado que el nivel de significación α y se vincula con el valor del estadístico de la prueba observado en la muestra. A modo de ejemplo, si se plantea una prueba unilateral por cola derecha, el p-valor representa la probabilidad de superar el valor de la media muestral observada en la muestra, simbólicamente $p\text{-valor} = P(\bar{X} > \bar{x}_{obs})$. De manera análoga, si se presenta una prueba por cola izquierda, tanto el valor de α como el p-valor se encuentran del lado izquierdo y la probabilidad se puede calcular haciendo $p\text{-valor} = P(\bar{X} < \bar{x}_{obs})$. Por último, si la prueba fuera bilateral, tanto alfa como el p-valor se encuentran repartidos en dos partes mitad en cada cola. De esta forma puede obtenerse el valor de manera similar a lo realizado para el caso de una sola cola y luego multiplicarlo por dos.