INTRODUCCIÓN

En esta unidad veremos algunos conceptos referidos a la prueba de hipótesis acerca de parámetros poblacionales.

Si la distribución de frecuencias de las observaciones puede asimilarse a la distribución de probabilidad teórica en la cual está basada la aplicación de la metodología inferencial elegida, entonces el investigador podrá estar seguro de que el error que informa en el proceso de prueba de hipótesis estadísticas planteadas es correcto.

Aplicar cualquier metodología estadística inferencial sin estudiar a fondo el cumplimiento de los supuestos en los cuales ella está basada, lleva irremediablemente a conclusiones erróneas.

Este curso apunta, como uno de sus objetivos fundamentales, a fortalecer la capacidad de discernimiento del investigador respecto a la correcta utilización de software estadístico, no sólo por el mero hecho de tenerlos a mano, sino sobre la base del conocimiento de sus fundamentos teóricos y prácticos.

Desarrollaremos también el caso en que se comparan solamente dos muestras, tanto las sus situaciones teóricas y prácticas teniendo en cuenta que las muestras que se analizan pueden ser muestras obtenidas independientemente una de otra (muestras independientes) o pueden estar relacionadas experimentalmente y conformar una situación de dependencia entre ellas (muestras dependientes o apareadas).

También distinguiremos situaciones en las cuales es posible aplicar pruebas paramétricas y otras donde es necesario apelar a pruebas no paramétricas o de libre distribución.

De ahora en más, pondremos mayor énfasis en la discusión de la correcta aplicación de las metodologías estadísticas inferenciales que en los cálculos aritméticos que estas requieran. Los cálculos pueden resolverse utilizando cualquier programa de computación, mientras que, la responsabilidad insoslayable del investigador radica en la elección de la metodología adecuada para analizar sus datos y en la correcta interpretación de los resultados.

3. PRUEBAS DE HIPÓTESIS
EL CONCEPTO DE PRUEBA DE HIPÓTESIS

A través de un ejemplo veremos ciertas consideraciones que nos aproximarán al concepto de prueba de hipótesis.

Supongamos que se establece la siguiente hipótesis:

"La droga A mata a las bacterias x"

A partir de esta hipótesis, el investigador formula la siguiente deducción:

Si se exponen colonias de bacterias x a la droga A, las bacterias morirán.

Si la hipótesis es verdadera, se sabrá que la deducción también lo es, pues si la droga realmente mata, las bacterias morirán.

Ahora bien, ¿si al exponer las bacterias x a la droga A, las bacterias no mueren, se puede afirmar que la hipótesis es falsa?

Si se denomina A a la hipótesis y B a la deducción se puede resumir que:

- Si la deducción no se cumple, la hipótesis es falsa.

B $\rightarrow$ A

- Si la deducción se cumple, decir que la hipótesis es verdadera:

B $\rightarrow$ A no es totalmente cierto

(las bacterias pueden haber muerto por muchas otras causas: temperatura moderada, medio de cultivo mal preparado, pH incorrecto, etc. y no sólo por la droga).

En resumen, cuando se analiza sólo una parte de la población (una muestra), siempre se puede rechazar una hipótesis pero no podemos aceptarla. (La evidencia únicamente nos llevará a no rechazarla).

Otro ejemplo muy simple puede ayudarnos a clarificar estas ideas.

Supongamos que establecemos la siguiente hipótesis y su correspondiente deducción con respecto a un aula de la Facultad que está con las puertas cerradas:

A: la habitación está vacía (hipótesis)

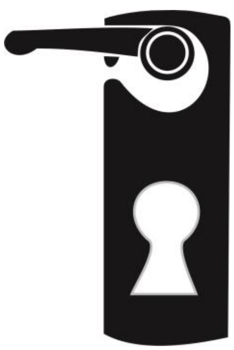
B: no se ve a nadie (deducción).

Si miramos por el ojo de la cerradura y vemos un alumno sentado en un banco, podemos sentirnos seguros al decir que la hipótesis A es falsa.

Deducción no se cumple $\rightarrow$ Hipótesis falsa

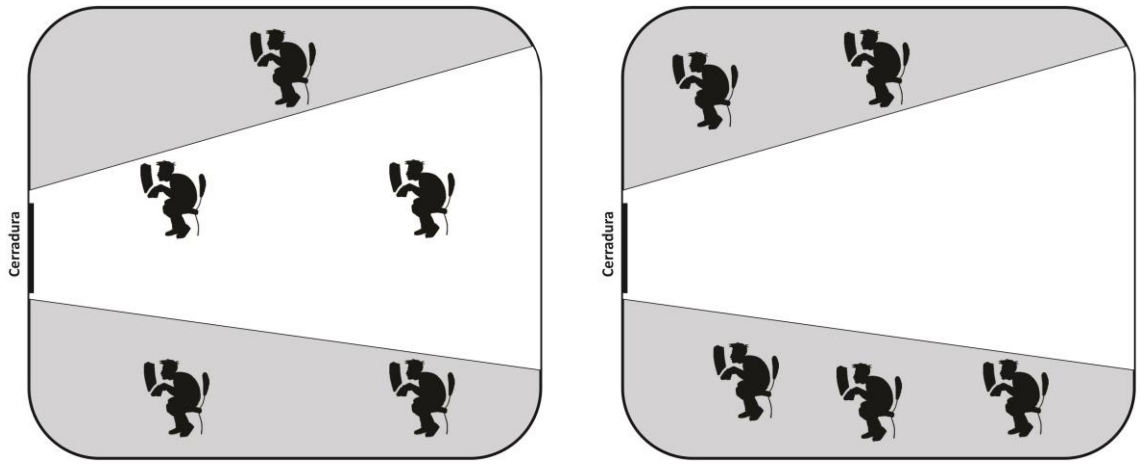
B $\rightarrow$ A

Si miramos por el ojo de la cerradura y no vemos a nadie en el aula, decir que la hipótesis es verdadera es incorrecto.



Deducción se cumple $\rightarrow$ Aceptar la hipótesis es una afirmación incierta

B $\rightarrow$ A es una afirmación incierta.



El hecho de mirar por el ojo de la cerradura equivale a observar sólo una parte del aula. Podemos asimilar esta situación con la inferencia estadística por cuanto debemos tomar una decisión en base a una información limitada (muestra) respecto a la totalidad del aula (población).

Generalmente, en una investigación científica se trabaja con muestras obtenidas a partir de una cierta población. Por ello se debe tener mucho cuidado al establecer conclusiones a partir de ellas. Si la evidencia de la muestra lleva a la decisión de rechazar la hipótesis planteada estaremos mucho más seguros de nuestra determinación que cuando no existan evidencias para su rechazo.

Hipótesis nula y alternativa

En el Módulo I definimos el concepto de hipótesis estadísticas desde un contexto científico muy general. Ahora estamos en condiciones de ser más específicos.

Una **hipótesis estadística** es un supuesto que se establece sobre las características de una distribución poblacional.

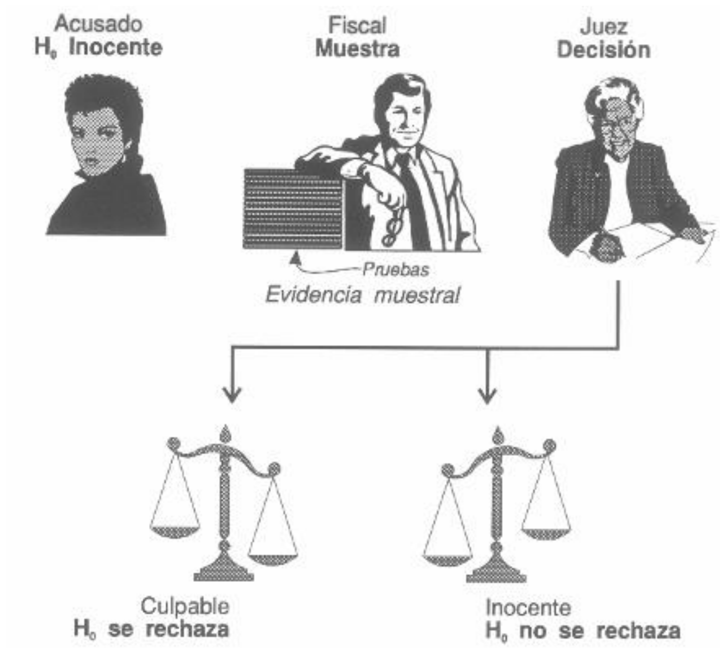
Pasaremos ahora a comentar el razonamiento que se sigue al efectuar una prueba de hipótesis.

Este razonamiento es muy similar al que se utiliza en una corte de justicia cuando se debe tomar la decisión de declarar o no culpable a una persona acusada de haber cometido un delito.

Los actores de un juicio pueden resumirse en un acusado, un fiscal y un juez, que es quien debe tomar la decisión de declarar o no culpable al acusado.

El acusado será considerado inocente hasta tanto las pruebas presentadas por el fiscal demuestren lo contrario.

Haciendo una analogía con la prueba de hipótesis, la hipótesis nula asume el papel del acusado en el juicio y el investigador el papel de fiscal.



La **hipótesis nula** es un supuesto acerca de uno o más parámetros de la población que debe ser rechazado o no en base a la evidencia muestral.

Simbolizaremos la hipótesis nula como H_0.

Se denomina nula en el sentido que supone que no existe diferencia real entre el verdadero valor del parámetro de la población de la que hemos obtenido la muestra y el valor hipotetizado (supuesto de inocencia).

Ahora bien, si la evidencia presentada al juez no es contundente, éste decidirá por la inocencia del acusado (no rechazo de la hipótesis nula). En caso contrario, cuando la evidencia condene al acusado, el juez tiene a mano una alternativa, la acusación de culpabilidad (lo que equivale al rechazo de la hipótesis nula planteada).

Si la hipótesis nula es falsa, deberá existir otra hipótesis que sea verdadera. Esta hipótesis recibe el nombre de hipótesis alternativa.

La hipótesis alternativa será simbolizada con H_1.

Aplicaremos los conceptos estudiados a un ejemplo real.

El Instituto Pasteur de París y el Laboratorio Merieux bajo la rectoría de la Agencia Nacional de Investigación contra el SIDA (ANRS) han elaborado una vacuna contra tal enfermedad. En estos días comenzó a probarse en seres humanos, luego de que los ensayos con chimpancés resultaran positivos. Cincuenta voluntarios, hombres y mujeres de 18 a 55 años seleccionados con minuciosidad entre más de 600 candidatos, serán los cobayos en esta experiencia histórica que acerca un paso más a la prevención de la enfermedad. Los candidatos debían ser todos sero-negativos, y con un alto riesgo de contraer la enfermedad (prácticamente se conoce que todos van a enfermarse a causa de su situación de alto riesgo).

Previamente a esta experiencia, el grupo francés ensayó la inmunidad lograda en chimpancés inoculados con la vacuna, obteniendo una respuesta favorable en un 75% de los casos. Por este motivo se decidió que la vacuna podrá ser considerada efectiva en seres humanos cuando ella prevenga de la enfermedad a más del 80% de los pacientes tratados.

Como dijimos que la hipótesis nula asume el rol del acusado y que éste es inocente (no efectiva) hasta que se demuestre su culpabilidad (efectiva), la hipótesis nula deberá expresar:

$$H_0)\ P \leq 0.80$$

donde P es la proporción poblacional, es decir, el parámetro del cual hacemos un juicio provisorio.

Los 50 voluntarios de la muestra serán inoculados con la vacuna y se les efectuará un seguimiento de un año luego de lo cual se obtendrían las primeras conclusiones.

Podemos comparar al equipo de investigación con el fiscal, pues cuando realizan su experimento están buscando pruebas respecto a la falsedad de la hipótesis nula especificada. En otras palabras, están tratando de demostrar que el acusado es realmente culpable. La muestra de personas a las que se les dio la vacuna será la evidencia que el equipo de investigación llevará ante el tribunal encargado de dictar sentencia (ANSR).

También dijimos que en el procedimiento de prueba de una hipótesis nula debe existir lo que se denomina una hipótesis alternativa.

Si en nuestro ejemplo la hipótesis nula especificaba que $P \leq 0.80$, evidentemente, H_1 deberá especificar que $P > 0.80$.

$$H_0)\ P \leq 0.80 \qquad y \qquad H_1)\ P > 0.80$$

Hasta aquí hemos resuelto el primer paso a seguir en el planteamiento de hipótesis, o sea, el establecimiento de la hipótesis nula y de la hipótesis alternativa.

3.1 Error Tipo I y Error Tipo II

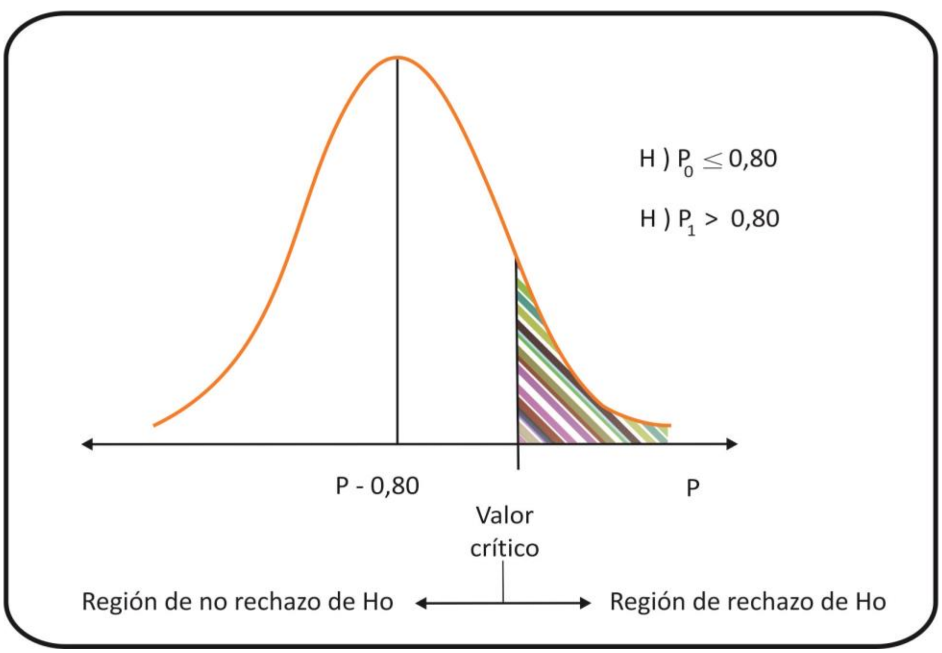
Ahora bien, si en la muestra considerada, el 15% de las personas no contrae la enfermedad, es muy probable que la vacuna sea declarada como no efectiva. Por el contrario, si el 95% de las personas no contraen la enfermedad, con seguridad ella será considerada efectiva.

Pero, ¿Qué decisión se tomará si, por ejemplo, el porcentaje de personas que no se enferman oscila entre el 70% y el 85%?

Evidentemente, debería establecerse un criterio que especificara, de manera objetiva, a partir de qué valor de p (proporción muestral), se tomará la decisión de rechazar o no la hipótesis nula. Este criterio, conocido como criterio de prueba, se calcula en base a la información aportada por la muestra.

Como consecuencia del criterio de prueba adoptado se podrá establecer un valor crítico que determinará una región de rechazo y una región de no rechazo de la hipótesis nula.

En el ejemplo de la vacuna, el estadístico para efectuar la prueba será la proporción muestral p de personas que quedaron inmunizadas por su acción y las reglas de decisión determinarán en cuál de las dos regiones se posiciona el valor muestral.



Al llegar a este punto debemos recordar que se está decidiendo algo con respecto al parámetro poblacional P a partir de la información proporcionada por una muestra. En realidad, estamos mirando por el ojo de la cerradura y, en consecuencia, el investigador podrá cometer dos tipos de errores.

Por una parte, podría rechazar la hipótesis nula diciendo que la vacuna es efectiva cuando en realidad no lo es. Es fácil imaginarse la peligrosidad de esta aseveración en cuanto a la salud de los pacientes que creen estar inmunizados y no toman las precauciones necesarias para no contraer la enfermedad.

Otro error que puede cometer consiste en no rechazar la hipótesis cuando en realidad es falsa. Estaría diciendo que la vacuna no es efectiva cuando en realidad lo es. La ciencia médica perdería la posibilidad de erradicar el SIDA cuando tenía todo a su alcance para hacerlo.

Pasaremos a definir estos dos tipos de errores que son: error de Tipo I y error de Tipo II.

El **error de Tipo I** es aquel que se comete al rechazar una hipótesis nula cuando ella es verdadera.

La probabilidad de cometer un error de Tipo I se simboliza con α, conocida como p o nivel de significación de una prueba.

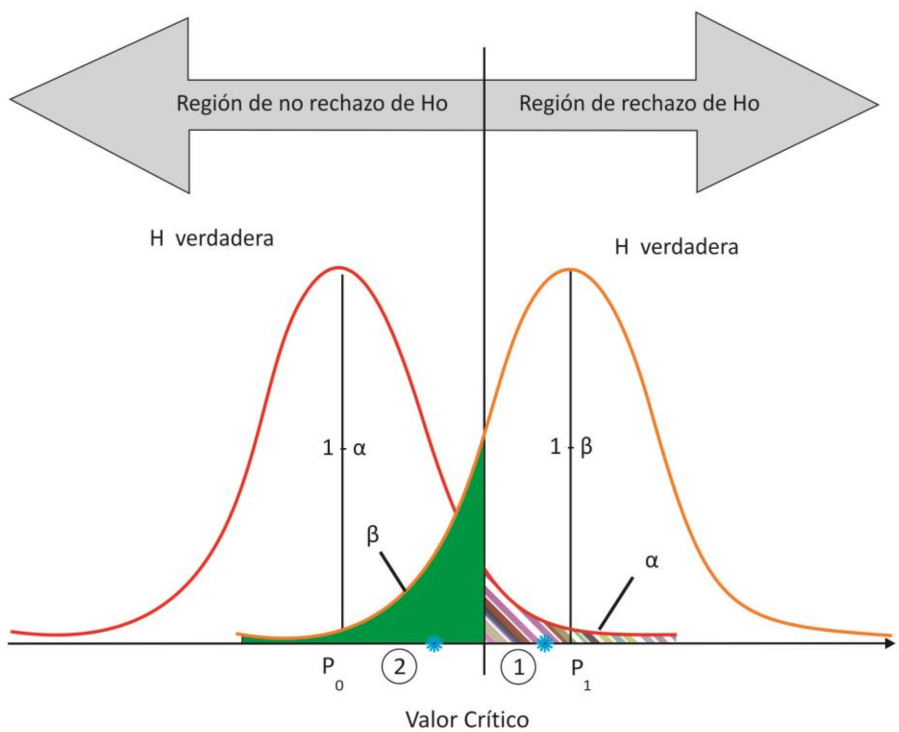
El **error de Tipo II** es aquel que se comete al no rechazar una hipótesis nula cuando en realidad es falsa.

La probabilidad de cometer un error de Tipo II se simboliza con β. El siguiente cuadro resume todas las acciones que se pueden llevar a cabo en un procedimiento de prueba de hipótesis.

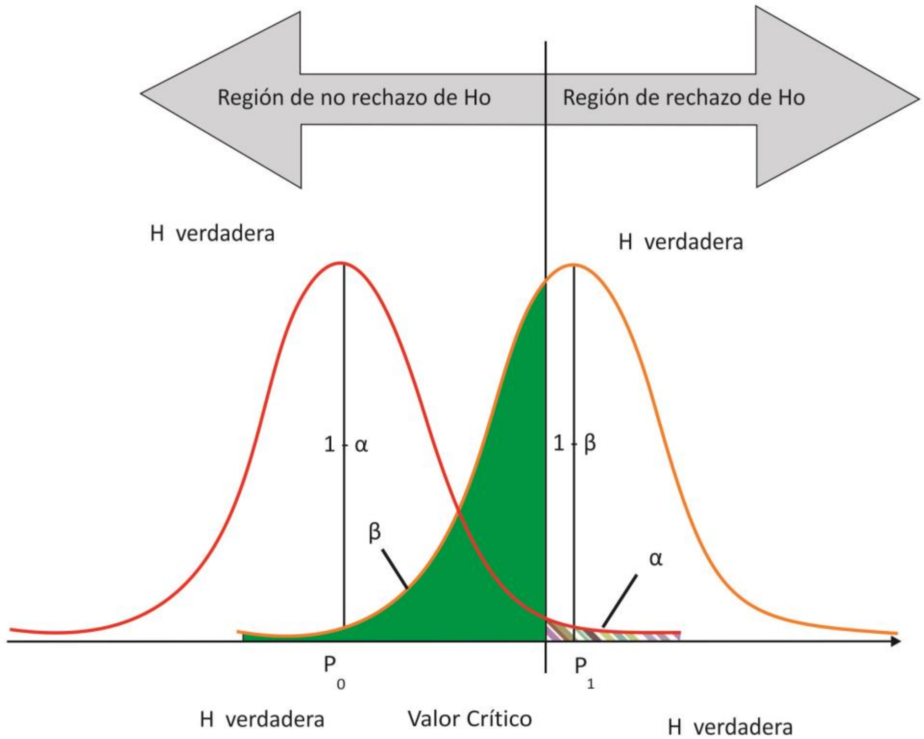
Decisión	Hipótesis nula	
	Verdadera	Falsa
Rechazar	**Error tipo I** $P(E_I) = \alpha$	**Decisión correcta** $(1 - \beta)$
No rechazar	**Decisión correcta** $(1 - \alpha)$	**Error tipo II** $P(E_{II}) = \alpha$

También podemos observarlo gráficamente:

a)



b)



En los gráficos a) y b), la curva de la izquierda representa la distribución de p cuando H_0 es verdadera es decir, la muestra proviene de una población con media P_0. La curva de la derecha representa la distribución del estadístico cuando se cumple una hipótesis alternativa H_1 (una alternativa podría ser $P = P_1$ con $P_1 > P_0$). Las regiones de rechazo y no rechazo de H_0 siempre se definen con referencia a la distribución que supone que H_0 es verdadera.

Las reglas de decisión de una prueba especifican que H_0 será rechazada cuando el estadístico muestral utilizado como criterio para efectuar el correspondiente prueba (p), cae en la región de rechazo de la misma (cae en el punto (*1*) del gráfico). La probabilidad de que un estadístico "caiga" en esta región es igual a α cuando H_0 es verdadera.

La probabilidad del error de Tipo II, es decir la probabilidad de no rechazar H_0 cuando es falsa, está localizada bajo la curva donde se cumple H_1 pero a la izquierda del valor crítico, es decir, en la frontera de la zona de aceptación de H_0.

Comparando ambos gráficos, vemos que en el b), la probabilidad del error de Tipo I es menor que en el a) pero, el área correspondiente al error de Tipo II, se ha incrementado.

Entonces, cuando tratamos de reducir la probabilidad de cometer el error Tipo I (α), aumenta la probabilidad de cometer el error de Tipo II (β).

La única manera que tiene el investigador de disminuir ambos tipos de errores a la vez consiste en aumentar el tamaño de la muestra (n).

3.2 Potencia de una prueba

La elección de un criterio de prueba que defina reglas de decisión adecuadas a la hipótesis que queremos probar, depende de la potencia que tenga la prueba para detectar cuando rechazar o no una hipótesis nula.

En otras palabras, se debe buscar una prueba que sepa discriminar correctamente cuándo las diferencias entre los valores muestrales realmente observados por el investigador y el valor verdadero que toma el parámetro a ser probado se deban solamente al azar (no se rechaza H_0) y cuándo la divergencia es tan grande que la evidencia muestral no pueda sostener al valor que toma el parámetro en la hipótesis nula (rechazo de H_0).

Esto se resume diciendo que se debe buscar la prueba más potente. Por ello, pasamos a definir, a continuación, lo que se entiende por potencia de una prueba estadística.

La **potencia de una prueba** se define como la función que establece la probabilidad de rechazar la hipótesis nula cuando es falsa.

De acuerdo a la definición anterior, la potencia de una prueba se calcula tomando $1 - \beta$, o sea, uno menos la probabilidad de cometer un error de Tipo II.

Cada uno de los posibles valores que puede tomar el parámetro en la hipótesis alternativa tiene su propia potencia. Así, cuanto más próximo se encuentre un valor del parámetro definido por la hipótesis alternativa al valor que toma en la hipótesis nula, mayor será la superposición de las correspondientes distribuciones y menor será la potencia de la prueba con respecto a esta alternativa.

3.3 Significancia estadística y valor-P

Cuando las pruebas estadísticas son realizadas mediante el uso de software, resulta muy útil tomar las decisiones en las pruebas de hipótesis, con el valor p o valor de significación (p-value).

Este valor proporciona una información interesante acerca de la verdadera probabilidad de cometer el error tipo I cuando se realiza una prueba de hipótesis. El valor "*p*" se define como la probabilidad de obtener una discrepancia mayor o igual que la observada en la muestra cuando se cumple la hipótesis nula.

Esta medida de discrepancia, en general, puede definirse como:

$$d_i = \left| \frac{\hat{\theta} - \theta}{\sigma_{\hat{\theta}}} \right|$$

Donde:
θ: parámetro a estimar
$\hat{\theta}$: estimador
$\sigma_{\hat{\theta}}$: error estándar de la estimación

Es decir, se trata del cálculo del "valor observado" del estadístico para una muestra en particular, y tiene una distribución de probabilidad conocida.

La medida de discrepancia depende de la dirección de la hipótesis alternativa. Si la prueba de hipótesis es bilteral, no se considera el signo de la desviación que se produce entre el estimador y el parámetro especificado en la hipótesis nula.

Si la prueba es unilateral se debe tener en cuenta si la dirección de la hipótesis ula es hacia la izquierda o hacia la derecha de la distribución de probabilidad utilizada.

Se denomina $\hat{d}$ al valor observado de la discrepancia, se tiene:

$$p = P(d \geq \hat{d} / H_o)$$

Es decir, *p* es la probabilidad de obtener una discrepancia igual o mayor a la observada en la muestra. Se trata de la probabilidad de las "colas" (si la prueba es bilateral); o la "cola" (si es unilateral) que ocurran valores iguales o superiores al observado.

Este valor de p debe compararse con el α que el investigador está dispuesto a aceptar. Si p es menor que α, entonces debemos rechazar la H_o, de lo contrario no podemos rechazarla. Dicho de otro modo, cuanto menor sea *p*, menor es la probabilidad de aparición de una discrepancia como la observada y menor será la credibilidad de H_o.