

PRUEBA DE HIPÓTESIS

Error Tipo I y tipo II

Como hemos visto ya, nuestras pruebas de hipótesis están estrechamente relacionadas con la forma en que establecemos la hipótesis nula y la alternativa, y estas mismas están formadas de tal forma en que sabemos que una de ellas si o si debe ser verdadera, y por lo tanto la otra falsa. Esto último nos hace llevar a plantear una idea: nosotros aceptaremos o rechazaremos nuestra hipótesis nula contra la alternativa basados en los resultados de la muestra que tenemos, por lo tanto, existe la posibilidad de que la decisión nos lleve a aceptar la hipótesis nula siendo que esta es falsa, o rechazar la hipótesis nula siendo que esta es correcta. Este tipo de errores se conocen como error Tipo I y Tipo II.

Errores Tipo I: Probabilidad de rechazar la hipótesis nula cuando esta es verdadera, se denota como un valor α , donde $0 \leq \alpha \leq 1$. (Este valor α también se conoce como nivel de significancia)

Error Tipo II: Probabilidad de aceptar la hipótesis nula (o no rechazarla) cuando esta es falsa, se denota con un valor β , donde $0 \leq \beta \leq 1$.

Cometer cualquiera de estos dos errores nos traería problemas al momento de inferir y tomar decisiones en base a los datos con los que contamos, por eso es importante darles la relevancia, ya que Rechazar H_0 no significa necesariamente que esta hipótesis sea falsa, si no que dado los datos que manejamos, no tenemos evidencia para probar que sea verdadera.

Explicación

Téngase presente también que solo es posible cometer uno de estos errores a la vez, es decir, si rechazamos la hipótesis nula solo corremos el riesgo de cometer el error tipo I, mientras que al aceptarla, solo podríamos estar cometiendo el error tipo II. También notemos que dichos errores son probabilidades condicionales, ya que no sabemos con certeza cuál de las hipótesis es la verdadera, y por lo tanto solo podemos realizar una estimación de dichos errores asumiendo que estamos en una de aquellas situaciones, esto es:

$$P(\text{rechazar } H_0 | H_0 \text{ es verdadera}) = \alpha$$

$$P(\text{no rechazar } H_0 | H_0 \text{ es falsa}) = \beta$$

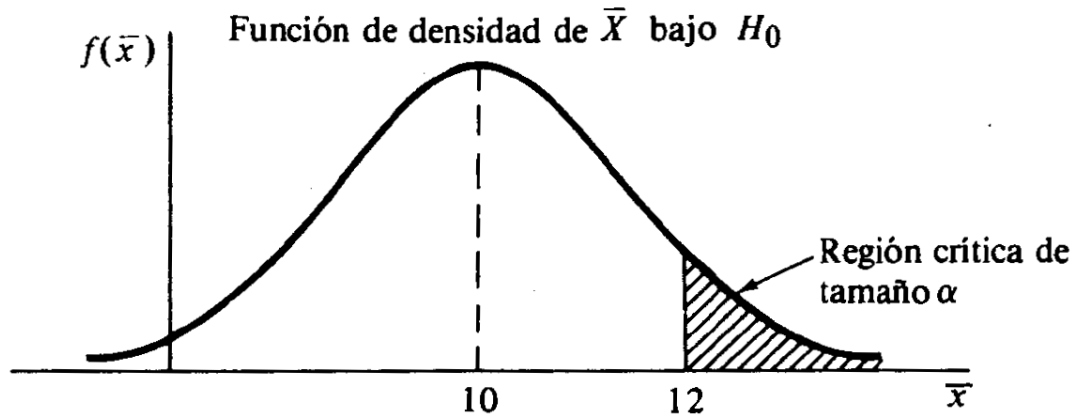
Donde H_0 , es la hipótesis nula. Por lo tanto no podemos obtener en un sentido absoluto las probabilidades de dichos errores.

Para entender de mejor manera la importancia de estos errores supongamos que el gerente de producción de una empresa decide que continuara con el proceso de ensamblaje del área de armado solo si el tiempo promedio de armado por unidad es de 10 minutos. En este caso es claro que la hipótesis nula es $H_0: \mu = 10$

Ahora nosotros queremos probar si esto es efectivamente cierto o no, pero para ello debemos primero formular una regla que nos permita inferir sobre nuestra hipótesis nula. A esta regla es lo que se le llama prueba de hipótesis. Luego con un proceso estadístico apropiado (Que recibe el nombre de estadístico de prueba) podremos inferir finalmente si podemos o no rechazar la hipótesis.

Para seguir con nuestro ejemplo supongamos que dada una muestra de tamaño n nosotros decidimos que rechazaremos la hipótesis nula si el valor de la media muestral \bar{x} es mayor a 12 (suponga por ahora esto). Por lo tanto tenemos que \bar{x} es el estadístico de prueba, $\bar{x} = 12$ el valor crítico de la prueba, y el conjunto de los valores de \bar{x} mayor a 12 el área o región crítica de la prueba.

Vamos a suponer que n es lo suficientemente grande para asumir que la prueba que estamos haciendo sobre \bar{x} esta sobre una distribución normal.



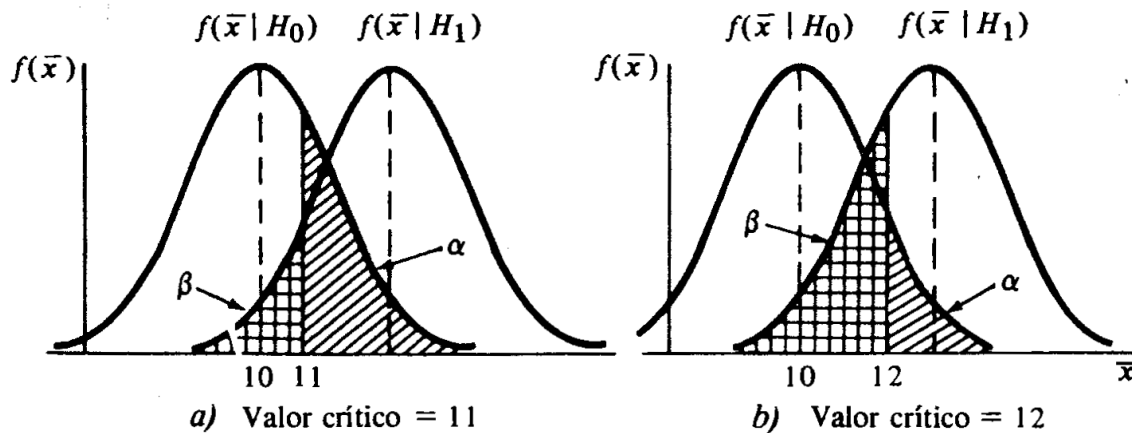
La figura muestra lo que hemos descrito anteriormente. Asumimos que la muestra se distribuye normal, bajo nuestra hipótesis de que $\mu = 10$, donde la región sombreada es todo lo que está a la derecha del valor 12, y representa la región crítica. El área de la región crítica es igual al tamaño del error tipo I.

Para continuar enfoquémonos en la importancia de definir bien las hipótesis nulas y alternativas. Ya dijimos anteriormente que dadas ciertas circunstancias muchas veces es conveniente identificar lo que queremos probar como hipótesis nula y otras como alternativa, esta idea tiene directa relación con el hecho de que lo principal de dicha hipótesis es enfatizar en que es más grave cometer el error tipo I que el error tipo II. En nuestro ejemplo, supongamos que el gerente sospecha la media está por sobre 10, y por lo tanto una buena hipótesis alternativa sería por lo tanto

$$H_1: \mu > 10$$

Ahora ¿Por qué sería importante plantearlo así y no de otra manera? Recordemos que el error tipo I es rechazar la hipótesis nula cuando esta es verdadera, y por tanto cometer dicho error es no tener la evidencia suficiente para "apoyar" dicha hipótesis. En nuestro ejemplo, si los datos que maneja el gerente lo llevan a inferir que la media del proceso es mayor a los 10 minutos, el detendría el proceso de producción en ese punto, provocando tomar medidas que es probable no sean necesarias, como cambiar las líneas de ensamblaje, nueva tecnología, inversión, etc. Por lo mismo, es que existe un consenso de que cometer el error tipo I es peor que cometer el error tipo II, y con ello la mayoría de las pruebas de hipótesis que se realizan están basadas en primero encontrar las pruebas que garanticen un cierto nivel de probabilidad del error tipo I (una probabilidad baja, por lo general del 1% o 5%) y que luego de eso presenten la menor probabilidad para el error tipo II.

Ahora bien hay que tener claro que no es posible disminuir el error tipo I sin que haga aumentar el error tipo II, lo que podemos notar en la siguiente gráfica:



Lo que aquí se representa es similar a lo que veíamos en el primer gráfico, donde la distribución de la izquierda en ambos gráficos representa aquella donde la media efectivamente es 10, mientras que la de la derecha representa una donde la media es mayor a 10. Se ve de manera clara que para el gráfico de la izquierda el área que representa el error tipo I en comparación con el gráfico de la derecha es mayor, mientras que lo contrario pasa con las áreas que representan el error tipo II en ambos gráficos.

Ejercicio

- Según un estudio, los hombres jóvenes en EEUU ven 52,6 minutos de televisión en la hora de mayor audiencia. Un investigador cree que en Alemania a la hora de mayor audiencia se ve más televisión. Dicho investigador toma los datos de una muestra de jóvenes y con los resultados busca contrastar las siguientes hipótesis:

$$H_0: \mu \leq 52,6$$

$$H_1: \mu > 52,6$$

Donde H_0 es la hipótesis nula, H_1 la alternativa, y μ representa los valores a contrastar.

- ¿Cuál sería el error tipo I (conceptualmente)? ¿Qué consecuencia tiene cometer dicho error?
 - ¿Cuál es el error tipo II (conceptualmente)? ¿Qué consecuencia tiene cometer dicho error?
- La información nutricional de un jugo de naranjas de litro dice que dicho jugo contiene 1 gramo o menos de grasa, lo que usted quiere probar
 - Formule las hipótesis para este caso
 - ¿Cuáles son los errores tipo I y II en este caso? ¿Cuáles serían sus consecuencias?

Esto nos hace poner énfasis en que por lo general las pruebas de hipótesis se realizan tomando en cuenta ambas probabilidades, pero donde primero se enfatiza en el error tipo I para luego encontrar cual es la que presenta un menor error tipo II.

Para mostrar mejor esto, sigamos con el mismo ejemplo de la explicación, seguimos suponiendo que la distribución del tiempo es normal, y que tiene una desviación estándar de 1,4 minutos. Además, se observa el tiempo de armado de 25 unidades. Se desea medir y comparar las siguientes zonas críticas:

Prueba A: rechazar H_0 si $\bar{X} = 10,65$

Prueba B: rechazar H_0 si $\bar{X} = 10,45$

Prueba C: rechazar H_0 si $\bar{X} = 10,25$

Queremos saber cuál es la prueba que presenta una menor probabilidad de error tipo II, dado que cumplen con alguna condición para el error tipo I, en este caso no se tolerará un error tipo I superior a 0,06.

Para poder determinar la probabilidad del error tipo I, debemos asumir que H_0 es cierta, y por lo tanto:

$$P(\bar{X} > c | \mu = 10) = \alpha$$

Donde c es el valor crítico, y según los datos tenemos una distribución para $\bar{X} \sim N(10, \frac{1,4}{\sqrt{25}})$ por lo tanto:

$$P(\bar{X} > 10,65 | \mu = 10) = \alpha$$

$$P(\bar{Z} > (10,65 - 10)/0,28 | \mu = 10) = \alpha$$

$$P(\bar{Z} > 2,32 | \mu = 10) = \alpha$$

$$\alpha = 0,0102$$

Análogamente para las otras dos pruebas queda:

$$P(\bar{X} > 10,45 | \mu = 10) = \alpha$$

$$\alpha = 0,0537$$

$$P(\bar{X} > 10,25 | \mu = 10) = \alpha$$

$$\alpha = 0,1867$$

Como sabemos el valor de α representa la probabilidad de cometer el error tipo I, por lo tanto, dada la condición de que queremos escoger las pruebas que tengan a lo más una probabilidad de 0,06, solo las pruebas A y B cumplen con dicha condición, por lo que descartamos el C.

Ahora podemos ver cuál de las dos pruebas que cumplían con la condición sobre el error tipo I tiene el error tipo II más bajo.

Para poder encontrar el valor de la probabilidad del error tipo II debemos trabajar ahora con la hipótesis alternativa, la que está formada como:

$$\beta(\mu) = P(\bar{X} \leq c | \mu > 10)$$

Por lo tanto, para las pruebas:

$$\beta(10.65) = P(\bar{X} \leq 10.65 \mid \mu = 10.4)$$

$$\beta(10.65) = P(\bar{Z} \leq 0.89 \mid \mu = 10.4)$$

$$\beta = 0.8133$$

Para la prueba A, mientras que para la prueba B:

$$\beta(10.4) = P(\bar{X} \leq 10.45 \mid \mu = 10.4)$$

$$\beta(10.4) = P(\bar{Z} \leq 0.18 \mid \mu = 10.4)$$

$$\beta = 0.5714$$

En este caso el error tipo II de la prueba B es la más pequeña, por lo tanto, de las pruebas que se hicieron, la mejor, porque presenta un error tipo I aceptado y el menor error tipo II es la prueba B. Nótese que en este caso para la misma muestra, solo suponiendo distintas pruebas, pudimos notar que la prueba con menor error tipo I (la prueba A) fue la que mayor error tipo II tenía, mostrando que efectivamente existe una especie de tira y afloja entre el error tipo I y tipo II.

La siguiente tabla muestra los valores del error tipo II para las pruebas A y B para diferentes niveles dados de μ :

μ	10.2	10.4	10.6	10.8	11.0	11.2	11.4
Prueba A	0.9463	0.8133	0.5714	0.2946	0.1056	0.0250	0.0037
Prueba B	0.8133	0.5714	0.2946	0.1056	0.0250	0.0037	0.0003

Cabe destacar que a medida que el valor de μ aumenta, la probabilidad de cometer el error tipo II disminuye para ambos casos, pero aun así la probabilidad de cometer el error tipo II para la prueba B sigue siendo más pequeños.

Visto de otra forma, a cada valor dado de μ , obtenemos distintos valores para β , lo que nos permite mirar dichos valores como una función $\beta(\theta)$, donde θ es un parámetro que toma el valor de μ . Dicha función se conoce como Función de Característica de Operación, la cual al graficarse para los distintos valores de θ , nos entrega la curva característica de operación.

Dado que $\beta(\theta)$ es la probabilidad de que un valor no se encuentre en la región crítica cuando la hipótesis nula es falsa (probabilidad del error tipo II), su complemento, es decir $1 - \beta(\theta)$, nos entrega la probabilidad de que un valor estadístico de prueba se encuentre dentro de la región crítica cuando la hipótesis es falsa. Dicho valor, $1 - \beta(\theta)$, es conocido como la función Potencia

Función Potencia: $P(\theta) = 1 - \beta(\theta)$, es la probabilidad de rechazar la hipótesis nula cuando esta es falsa, es decir, cuando la hipótesis alternativa es la correcta.

Ejercicio (tarea entregar próxima clase) (20 minutos):

Calcule los datos entregados en la tabla de más arriba