AWAKELAB

BASECAMP

Ciencia de Datos

Inferencia Estadística

Objetivo de la jornada

- Realizar una prueba de hipótesis para probar la validez de una aseveración acerca de un parámetro de la población

Test de Significancia

Qué es una prueba de hipótesis

Una prueba de hipótesis es una regla que especifica si se puede aceptar o rechazar una afirmación acerca de una población dependiendo de la evidencia proporcionada por una muestra de datos.

Una prueba de hipótesis examina dos hipótesis opuestas sobre una población: la **hipótesis nula y la hipótesis alternativa**. La hipótesis nula es el enunciado que se probará. Por lo general, la hipótesis nula es un enunciado de que "*no hay efecto*" o "*no hay diferencia*". **La hipótesis alternativa** es el enunciado que se desea poder concluir que es verdadero de acuerdo con la evidencia proporci

Con base en los datos de muestra,
rechazar la hipótesis nula. Usted utiliza el valor p para tomar esa decisión. Si el valor p es menor que el nivel de significancia (denotado como α o alfa), entonces puede rechazar la hipótesis nula.

Un error común de percepción es que las pruebas estadísticas de hipótesis están diseñadas para seleccionar la más probable de dos hipótesis. Sin embargo, al diseñar una prueba de hipótesis, establecemos la hipótesis nula como lo que queremos desaprobar. Puesto que establecemos el nivel de significancia para que sea pequeño antes del análisis (por lo general, un valor de 0.05 funciona adecuadamente), cuando rechazamos la hipótesis nula, tenemos prueba estadística de que la alternativa es verdadera. En cambio, si no podemos rechazar la hipótesis nula, no tenemos prueba estadística de que la hipótesis nula sea verdadera. Esto se debe a que no establecimos la probabilidad de aceptar equivocadamente la hipótesis nula para que fuera pequeña.

Entre las preguntas que se pueden contestar con una prueba de hipótesis están las siguientes:

- ¿Tienen las estudiantes de pregrado una estatura media diferente de 66 pulgadas?
- ¿Es la desviación estándar de su estatura igual a o menor que 5 pulgadas?
- ¿Es diferente la estatura de las estudiantes y los estudiantes de pregrado en promedio?
- ¿Es la proporción de los estudiantes de pregrado significativamente más alta que la proporción de las estudiantes de pregrado?

Ejemplo de cómo realizar una prueba de hipótesis básica

Usted puede seguir seis pasos básicos para configurar y realizar correctamente una prueba de hipótesis. Por ejemplo, el gerente de una fábrica de tuberías desea determinar si el diámetro promedio de los tubos es diferente de 5 cm. El gerente sigue los pasos básicos para realizar una prueba de hipótesis.

NOTA: Debe determinar los criterios para la prueba y el tamaño de muestra necesario antes de recolectar los datos.

1. Especificar las hipótesis.

En primer lugar, el gerente formula las hipótesis. La hipótesis nula es: la media de la población de todos los tubos es igual a 5 cm. Formalmente, esto se escribe como: $H0: \mu = 5$

Luego, el gerente elige entre las siguientes hipótesis alternativas:

Condición que se probará	**Hipótesis alternativa**
La media de la población es menor que el objetivo.	unilateral: $\mu < 5$
La media de la población es mayor que el objetivo.	unilateral: $\mu > 5$
La media de la población es diferente del objetivo.	bilateral: $\mu \neq 5$

Como tiene que asegurarse de que los tubos no sean más grandes ni más pequeños de 5 cm, el gerente elige la hipótesis alternativa bilateral, que indica que la media de la población de todos los tubos no es igual a 5 cm. Formalmente, esto se escribe como $H_1: \mu \neq 5$.

2. Elegir un nivel de significancia (también denominado alfa o α).

El gerente selecciona un nivel de significancia de 0.05, que es el nivel de significancia más utilizado.

3. Determinar la potencia y el tamaño de la muestra para la prueba.

El gerente utiliza un cálculo de potencia y tamaño de la muestra para determinar cuántos tubos tiene que medir para tener una buena probabilidad de detectar una diferencia de 0.1 cm o más con respecto al diámetro objetivo.

4. **Recolectar los datos**

   Recoge una muestra de tubos y mide los diámetros.

5. **Comparar el valor p de la prueba con el nivel de significancia**

   Después de realizar la prueba de hipótesis, el gerente obtiene un valor p de 0.004. El valor p es menor que el nivel de significancia de 0.05.

6. **Decidir si rechazar o no rechazar la hipótesis nula**

   El gerente rechaza la hipótesis nula y concluye que el diámetro medio de todos los tubos no es igual a 5 cm.

Hipótesis Nula y alternativa

Las hipótesis nula y alternativa son dos enunciados mutuamente excluyentes acerca de una población. Una prueba de hipótesis utiliza los datos de la muestra para determinar si se puede rechazar la hipótesis nula.

Hipótesis Nula (H_0)

La hipótesis nula indica que un parámetro de población (tal como la media, la desviación estándar, etc.) es igual a un valor hipotético. La hipótesis nula suele ser una afirmación inicial que se basa en análisis previos o en conocimiento especializado.

Hipótesis alternativa (H_1)

La hipótesis alternativa indica que un parámetro de población es más pequeño, más grande o diferente del valor hipotético de la hipótesis nula. La hipótesis alternativa es lo que usted podría pensar que es cierto o espera probar que es cierto.

La hipótesis alternativa puede ser unilateral o bilateral.

Bilateral

Utilice una hipótesis alternativa bilateral (también conocida como hipótesis no direccional) para determinar si el parámetro de población es mayor que o menor que el valor hipotético. Una prueba bilateral puede detectar cuándo el parámetro de población difiere en cualquier dirección, pero tiene menos potencia que una prueba unilateral.

Ejemplo:

Un investigador tiene los resultados de una muestra de estudiantes que presentaron un examen nacional en una escuela secundaria. El investigador desea saber si las calificaciones de esa escuela difieren del promedio nacional de 850. Una hipótesis alternativa bilateral (también conocida como hipótesis no direccional) es adecuada porque el investigador está interesado en determinar si las calificaciones son menores que o mayores que el promedio nacional. ($H0: \mu = 850$ vs. $H1: \mu \neq 850$)

Unilateral

Utilice una hipótesis alternativa unilateral (también conocida como hipótesis direccional) para determinar si el parámetro de población difiere del valor hipotético en una dirección específica. Usted puede especificar la dirección para que sea mayor que o menor que el valor hipotético. Una prueba unilateral tiene mayor potencia que una prueba bilateral, pero no puede detectar si el parámetro de población difiere en la dirección opuesta.

Ejemplo:

Un investigador tiene los resultados de una muestra de estudiantes que tomaron un curso de preparación para un examen nacional. El investigador desea saber si los estudiantes preparados tuvieron puntuaciones por encima del promedio nacional de 850. Una hipótesis alternativa unilateral (también conocida como hipótesis direccional) se puede utilizar porque el investigador plantea la hipótesis de que las puntuaciones de los estudiantes preparados son mayores que el promedio nacional. ($H0: \mu = 850$ vs. $H1: \mu > 850$)

Significancia Estadística

La diferencia entre un estadístico de muestra y un valor hipotético es estadísticamente significativa si una prueba de hipótesis indica que es muy poco probable que la misma haya ocurrido en virtud de las probabilidades. Para evaluar la significancia estadística, examine el valor p de la prueba. Si el valor p está por debajo de un nivel de significancia (α) especificado (generalmente 0.10, 0.05 o 0.01), usted puede decir que la diferencia es estadísticamente significativa y rechazar la hipótesis nula de la prueba.

Por ejemplo, supongamos que usted desea determinar si el grosor de unos parabrisas de vehículo supera los 4 mm, tal como lo exigen las normas de seguridad. Usted toma una muestra de parabrisas y realiza una prueba t de 1 muestra con un nivel de significancia (α) de 0.05 y plantea las hipótesis siguientes:

- H_0: $\mu = 4$
- H_1: $\mu > 4$

Si la prueba produce un valor p de 0.001, usted declara significancia estadística y rechaza la hipótesis nula, porque el valor p es menor que α. Usted concluye a favor de la hipótesis alternativa: que el grosor de los parabrisas es mayor que 4 mm.

Sin embargo, si el valor p es igual a 0.50, usted no puede declarar significancia estadística. No tiene suficiente evidencia para afirmar que el grosor promedio de los parabrisas es mayor de 4 mm.

"Un resultado estadísticamente significativo podría no ser significativo desde el punto de vista práctico"

La significancia estadística por sí sola no implica que los resultados tengan una consecuencia práctica. Si utiliza una prueba con una potencia muy alta, podría concluir que una pequeña diferencia con respecto al valor hipotético es estadísticamente significativa. Sin embargo, esa pequeña diferencia podría ser insignificante para su situación. Debe usar su conocimiento especializado para determinar si la diferencia es significativa desde el punto de vista práctico.

Por ejemplo, supongamos que usted está evaluando si la media de la población (μ) para las horas trabajadas en una planta de manufactura es igual a 8. Si μ no es igual a 8, la potencia de la prueba se acercará a 1 a medida que aumente el tamaño de la muestra y el valor p se acerque a 0.

Con suficientes observaciones, es probable que incluso las diferencias triviales entre los valores hipotéticos y reales de los parámetros se vuelvan significativas. Por ejemplo, supongamos que el valor real de mu es 7 horas, 59 minutos y 59 segundos. Con una muestra lo suficientemente grande, lo más probable es que usted rechace la hipótesis nula de que μ es igual a 8 horas, aunque la diferencia no tenga importancia práctica.

Los intervalos de confianza (si corresponde) suelen ser más útiles que las pruebas de hipótesis, porque ofrecen una manera de evaluar la importancia práctica, además de la significancia estadística. Ayudan a determinar cuál *es* el valor de un parámetro en lugar de determinar cuál *no* es.

El valor P

El valor-p, del inglés, *p-value*, es el nivel de significación mínimo no arbitrario con el que podemos rechazar la hipótesis nula (H0) dada una función de distribución y un estadístico de contraste.

En otras palabras, el valor-p, es la probabilidad mínima definida por la distribución con la que podemos rechazar la hipótesis nula (H0) sin necesidad de definir *a priori* el nivel de significación para el contraste.

Recordar que el área bajo la curva de la función de distribución es una probabilidad. Entonces, desde este punto de vista, el valor-p será la probabilidad de observar un estadístico de contraste tan extremo para que la hipótesis nula sea cierta.

Dominio

Dado que el valor-p es una probabilidad, este valor estará comprendido entre 0 y 1.

No arbitrario

A diferencia de los niveles de significación que estamos más acostumbrados a ver, tales como 1%, 5% y 10%, el valor-p depende de la función de distribución que tenga el estadístico de contraste. Entonces, los niveles de 1%, 5% y 10% los decidimos al principio del contraste. A esta selección se le llama arbitraria.

Fórmula del p-valor

El valor-p no es un valor único como el valor crítico, sino que dependerá del estadístico. Para diferentes valores del estadístico de contraste, el valor crítico será el mismo. En cambio, para diferentes valores del estadístico de contraste, el valor-p también será distinto, porque el valor-p depende del valor que tome el estadístico de contraste.

$$P(|D| > |d|)$$

$$Valor - p$$

Donde,

- D, es una variable aleatoria que sigue una distribución determinada.
- d, es el valor del estadístico de contraste.

Cálculo

Es posible calcular el valor-p a mano pero se tendría que disponer de tablas de distribuciones muy precisas, es decir, con muchos decimales porque el valor-p tiende a ser pequeño. La mayoría de programas estadísticos llevan ya incorporado el valor-p y normalmente aparece en el output de resultados de las estimaciones por Mínimos Cuadrados Ordinarios (MCO). Puede parecer difícil de utilizar pero con la práctica es una herramienta muy útil.

Para calcular el valor-p necesitamos:

- Estadístico de contraste.
- La distribución del estadístico de contraste y conocer sus parámetros.

- Tabla tstudent:

  https://cms.dm.uba.ar/academico/materias/1ercuat2015/probabilidades_y_estadistica_C/tabla_tstudent.pdf

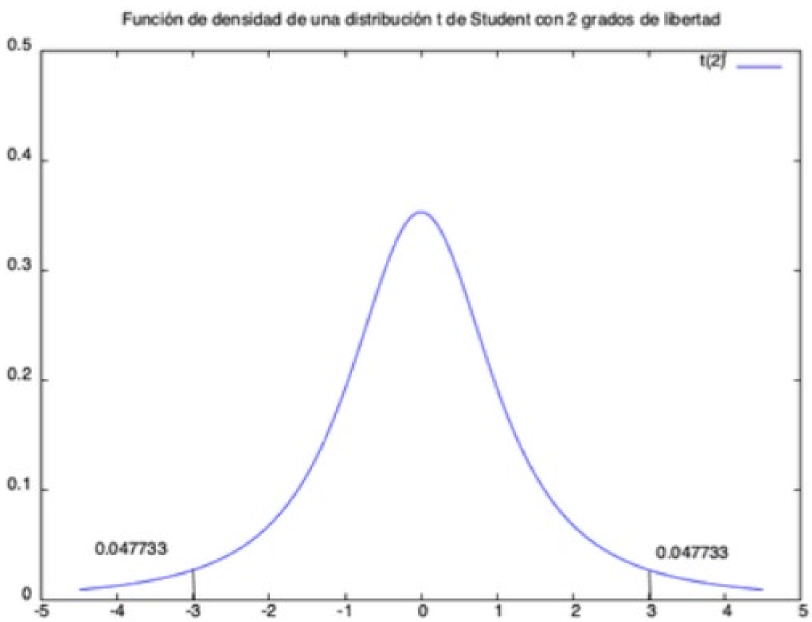
Regla de rechazo

Si valor–p **<** nivel de significación => Rechazo H0.

Si valor–p **>** nivel de significación => No rechazo H0.

Representación

En el caso de una distribución t de Student con 2 grados de libertad y un estadístico de contraste igual a 3, la probabilidad de encontrar un estadístico tan extremo cuando la hipótesis nula (H0) es cierta es de 4,77%.



Valor-p de una distribución t de Student con 2 grados de libertad

En otras palabras, si la hipótesis nula (H0) fuera verdadera, un estadístico tan grande como 3, solo sería observado un 4,77% de las veces.

¿Por qué se llama valor-p?

El nombre del valor-p tiene su origen en la definición que hace referencia a ser el área bajo la curva de la función de distribución fuera del intervalo de confianza. Entonces, cómo esa área es la probabilidad mínima de rechazar la hipótesis nula, la “p” de valor-p se refiere a la **probabilidad**. Y, como el valor-p corresponde a un número, y por tanto, a un valor, la palabra “valor” de valor-p se atribuye a la cifra numérica. En algunos libros podemos encontrar “valor de la probabilidad” refiriéndose al valor-p. A lo mejor, decirle “probabilidad mínima para rechazar la hipótesis nula” era demasiado largo y no albergaba ningún misterio para los estudiantes…

Pruebas sobre una proporción de la población

Este tipo de prueba, también conocida la Prueba Z para una proporción poblacional, se utiliza para una variable cualitativa X que representa el estado de algo, por ejemplo, defectuoso y no defectuoso, vendido o no vendido etc, y de la cual se quiere demostrar una hipótesis con la proporción de determinada categoría de una variable (Millones, Barreno, Vásquez y Castillo, 2016).

El estadístico de prueba está dado por:

$$z0=\frac{x-np0}{\sqrt{np0(1-p0)}}$$

En donde:

n: Tamaño de la muestra

x: Cantidad de elementos de la muestra que cumplen con el criterio deseado

p0: Proporción a probar

Se supone que x tiene distribución binomial (lo que significa que solo puede tener dos posibles resultados, por ejemplo defectuoso o no defectuoso) y n es un tamaño relativamente grande.

Las hipótesis son las siguientes:

- Prueba de dos colas:
    - H0: p=p0
    - H1: p≠p0

- Pruebas de una cola.
    - H0: p=p0
    - H1: p>p0
    - H0: p=p0
    - H1: p<p0

Es decir, se quiere verificar si el valor de una cierta proporción poblacional p es igual a p0.

Los criterios de rechazo de H0 son los siguientes:

- Si es de dos colas se rechaza $H0$ si $|z0|>z(\alpha/2)$
- Si $H1: p>p0$ se rechaza $H0$ si $z0>z\alpha$
- Si H1: p

Ejemplo

En Ciudad Gótica se quiere saber si los jóvenes de 12 a 19 años tienen más o menos problemas de exceso de peso corporal que en el conjunto del país. De los datos nacionales se sabe que el 35% de los jóvenes de esas edades tiene sobrepeso u obesidad, si p es la proporción de jóvenes de 12 a 19 años con exceso de peso entonces se desea probar lo siguiente:

$H0: p=0,35$
$H1: p\neq0,35$
$\alpha: 0,05$

Se ha tomado una muestra de 300 jóvenes en ese rango de edad y 90 de ellos tienen sobrepeso.

Solución:

n: 300

x: 90

p0: 0,35

$$z0=\frac{x-np0}{\sqrt{np0(1-p0)}}$$

$$z0=\frac{90-300*0,35}{\sqrt{300*0.35(1-0.35)}}=-1,812$$

Si z(α/2)=1,96 entonces 1,812<1,96 por lo tanto no cumple con el criterio de rechazo entonces se acepta H0, es decir, no hay evidencia suficiente como para afirmar que la situación del peso de Ciudad Gótica es diferente a la del resto del país.

Pruebas sobre una media poblacional

Una de las partes que más confunden al hacer una prueba de hipótesis sobre la media es ¿qué estadístico de prueba hay que usar?

Para saber el estadístico a utilizar, se deben ver tres situaciones:

Caso 1:

Si las condiciones son:

- La variable X tiene distribución normal
- Conocemos el desvío estándar poblacional σ

Entonces el estadístico que se usa es:

$$\frac{\bar{X}-\mu}{\sigma/\sqrt{n}} \sim N(0,1)$$

Caso 2:

Si las condiciones son (lo único que cambia es que no conocemos el desvío poblacional):

- La variable X es normal
- No conocemos el desvío estándar poblacional σ, así que lo estimamos usando el desvío estándar muestral S

Usamos:

$$\frac{\bar{X} - \mu}{S/\sqrt{n}} \sim t_{n-1}$$

Caso 3:

Por último si las condiciones son:

- La variable X no sabemos qué distribución tiene (puede ser cualquier distribución)
- Conocemos el desvío estándar poblacional σσ
- El tamaño de la muestra debe ser grande n≥30n≥30

Usamos:

$$\frac{\bar{X}-\mu}{\frac{\sigma}{\sqrt{n}}} \approx N(0,1)$$

La distribución en este caso no es exactamente normal, sino APROXIMADAMENTE normal.

¿Por qué? Porque tenemos que usar el teorema central del límite para conocer la distribución de $\bar{X}$. Y el teorema dice que $\bar{X}$ tiende a la distribución normal en la medida en que n crece… pero no que tiene "exactamente" la distribución normal.

También se suele usar esta distribución aproximada si no se conoce σ:

$$\frac{\bar{X}-\mu}{\frac{\sigma}{\sqrt{n}}} \approx N(0,1)$$

Resumen:

Condiciones	Estadístico
• X normal • σ conocido	$\frac{\bar{X}-\mu}{\sigma/\sqrt{n}} \sim N(0,1)$
• X normal • σ desconocido	$\frac{\bar{X}-\mu}{S/\sqrt{n}} \sim t_{n-1}$
• X distribución cualquiera • σ conocido • $n \geq 30$	$\frac{\bar{X}-\mu}{\sigma/\sqrt{n}} \approx N(0,1)$ O bien si no se conoce σ: $\frac{\bar{X}-\mu}{S/\sqrt{n}} \approx N(0,1)$

Ejemplo:

La duración de las bombillas de **100** watt que fabrica una empresa sigue una distribución normal con una desviación de **120** horas. Su vida media está garantizada durante un mínimo de 800 horas.

Se escoge al azar una muestra de **50** bombillas de un lote y, después de comprobarlas, se obtiene una vida media de **750** horas.

a) Con un nivel de significación de 0,**01**, ¿habría que rechazar el lote por no cumplir la garantía?

b) ¿Cuál es la probabilidad de cometer el error tipo II si el tiempo medio de vida de las bombillas es **790** horas?

Solución:

Queremos hacer una prueba de hipótesis sobre la media de la duración de las bombillas. ¿Durán en promedio más de 800 horas o menos?

La variable es X: duración (en horas) de una bombilla de **100** watts, fabricada por cierta empresa.

Se sabe que:

$$X \sim N(\mu = ?; \sigma = 120)$$

No conocemos el valor de la media. Pero sí conocemos la media muestral de una muestra de tamaño 50:

$$n = 50 \quad ; \quad \bar{x} = 750$$

A primera vista parecería que las bombillas están durando menos que lo prometido por el fabricante. (El fabricante garantiza que duran en promedio **800** horas o más y obtuvimos una media muestral de **750** horas.)

Pero no podemos tomar la decisión "a ojo".

Tenemos que realizar una prueba de hipótesis.

Vamos a hacer la prueba de hipótesis realizando los pasos recomendados.

No es necesario escribir todos estos pasos, pero lo hacemos porque lo hace mucho más fácil de entender.

***Paso 1**: Definir la variable.*

X: duración (en horas) de una bombilla de **100** watts, fabricada por cierta empresa.

$$X \sim N(\mu = ?; \sigma = 120)$$

Paso 2: *Plantear las hipótesis estadísticas*

El fabricante afirma que duran **800** horas o más:

$$H_0 : \mu \geq 800$$

Queremos contrastar esa hipótesis con:

$$H_1 : \mu < 800$$

Paso 3: *Establecer un estadístico de prueba.*

En este caso hay dos posibles. Que son equivalentes.

Cómo X~N y σ es conocida, conocemos la distribución de la variable media muestral:

$$\bar{X} \sim N\left(\mu, \frac{\sigma}{\sqrt{n}}\right)$$

Este es un estadístico de prueba adecuado.

Pero también se puede estandarizar esta variable, y obtener:

$$\frac{\bar{X} - \mu}{\frac{\sigma}{\sqrt{n}}} \sim N(0,1)$$

Cualquiera de los dos sirve.

(Son básicamente el mismo. **En un caso está estandarizada la variable normal** y **en el otro no está estandarizada**).

Vamos a usar los dos para poder explicar cómo se hace con ambos.

Pero no es necesario que usen los dos.

Paso 4*: Seleccionar un nivel de significación*

El enunciado determina que:

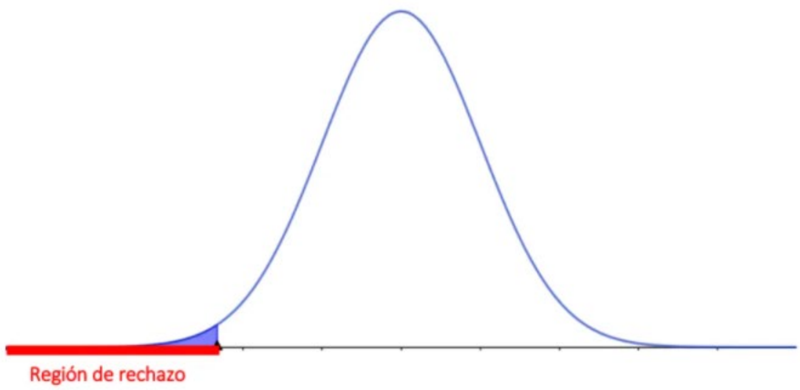
$$\alpha = 0,01$$

__Paso 5__: Determinar la zona de rechazo y la regla de decisión

Cómo la hipótesis alternativa afirma que μ es **menor que** un cierto valor, entonces decimos que la prueba es unilateral izquierda: la zona de rechazo queda ubicada a la izquierda.

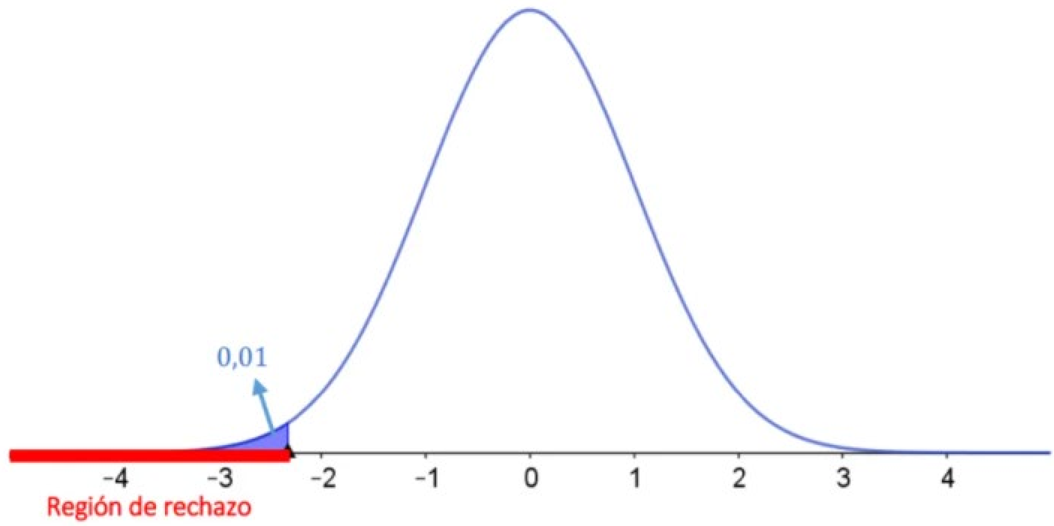
La distribución de ambos estadísticos es normal.

Así que el diagrama con la distribución del estadístico y la zona de rechazo a izquierda es así:



Pero ahora queremos determinar exactamente cuál es la región de rechazo. ¿Cuál es el valor de la abscisa que define la región de rechazo?

Si usamos el estadístico de prueba $\frac{\bar{X}-\mu}{\frac{\sigma}{\sqrt{n}}} \sim N(0,1)$ el diagrama con la zona de rechazo unilateral izquierda y el nivel de significación sería así:



¿Cuál es el valor de la variable normal estándar que acumula una probabilidad de **0,01** a su izquierda?

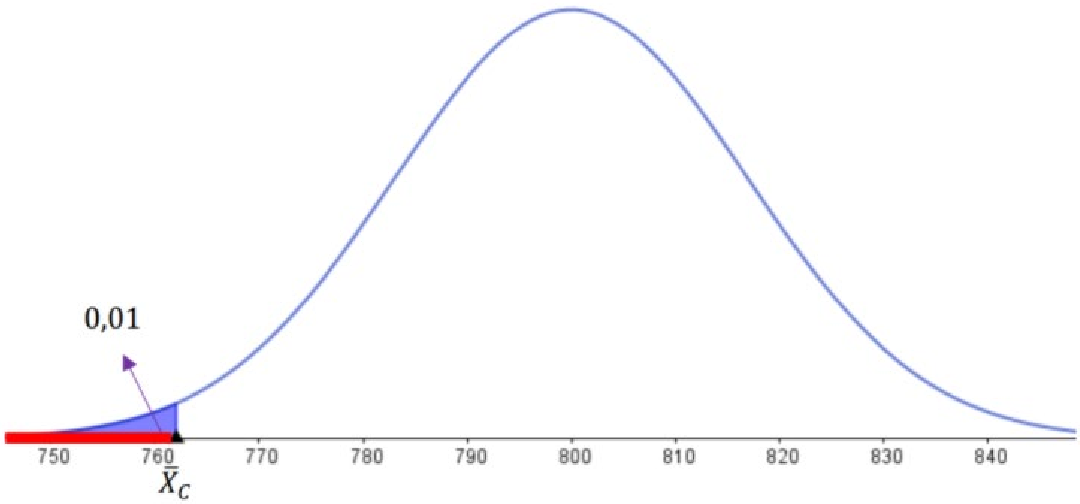
$$z_{0,01} = -2,33$$

(Eso se busca en la tabla de la normal estándar o usando software)

Entonces la regla de decisión es:

- Rechazo H0 si ep≤–2,33.
- No rechazo H0 si ep>–2,33.

Si quisiéramos usar el otro estadístico de prueba posible $\bar{X} \sim N\left(800, \frac{120}{\sqrt{50}} \cong 16,97\right)$, la lógica es exactamente la misma:



Pero ¿cómo averiguamos el valor crítico **$\bar{X}_C$** que acumula una probabilidad de **0,01** a su izquierda?

- **Opción 1**: usando software cómo Probability Distributions (para Android) o GeoGebra (para Windows).
- **Opción 2**: usando la tabla de probabilidad normal estándar. En general en un examen no se permite usar software. Así que no nos queda otra que esta opción. Veamos cómo es.

Si **$\bar{X}_c$** es aquel valor que acumula una probabilidad de **0,01** a su izquierda, entonces al estandarizarlo obtendremos **$z_{0,01}=-2,33$:**

$$-2,33 = \frac{\bar{X}_C - 800}{\frac{120}{\sqrt{50}}}$$

De acá podemos despejar ¯XC:

$$\Rightarrow \bar{X}_C = -2,33.\frac{120}{\sqrt{50}} + 800 \cong 760,46$$

Obtenemos que el valor crítico es **¯XC=760,46**. Luego la regla de decisión es:

- Rechazo H_0 si ¯X≤760,46.
- No rechazo $_0$ si ¯X>760,46.

Paso 6: *Calcular el valor observado del estadístico de prueba*

Usando el estadístico de prueba **¯X**:

El valor observado **¯X=750** pertenece a la zona de rechazo **(–∞;760,46**).

Si usamos el estadístico estandarizado tenemos que realizar el siguiente cálculo:

$$e_{p,obs} = \frac{750 - 800}{\frac{120}{\sqrt{50}}} \cong -2,95$$

También ocurre que el valor observado **(–2,95)** pertenece a la zona de rechazo **(–∞;–2,33)**.

Conclusión:

Decidimos rechazar la hipótesis nula.

La conclusión podría ser:

"Con un nivel de significación del **1%** hay evidencias suficientes para afirmar que la media de la duración de las bombillas es inferior a **800** horas."

Errores tipo I y tipo II

Ninguna prueba de hipótesis es 100% cierta. Puesto que la prueba se basa en probabilidades, siempre existe la posibilidad de llegar a una conclusión incorrecta. Cuando usted realiza una prueba de hipótesis, puede cometer dos tipos de error: tipo I y tipo II. Los riesgos de estos dos errores están inversamente relacionados y se determinan según el nivel de significancia y la potencia de la prueba. Por lo tanto, usted debe determinar qué error tiene consecuencias más graves para su situación antes de definir los riesgos.

Error de tipo I

Si usted rechaza la hipótesis nula cuando es verdadera, comete un error de tipo I. La probabilidad de cometer un error de tipo I es α, que es el nivel de significancia que usted establece para su prueba de hipótesis. Un α de 0.05 indica que usted está dispuesto a aceptar una probabilidad de 5% de estar equivocado al rechazar la hipótesis nula. Para reducir este riesgo, debe utilizar un valor menor para α. Sin embargo, usar un valor menor para alfa significa que usted tendrá menos probabilidad de detectar una diferencia si esta realmente existe.

Error de tipo II

Cuando la hipótesis nula es falsa y usted no la rechaza, comete un error de tipo II. La probabilidad de cometer un error de tipo II es β, que depende de la potencia de la prueba. Puede reducir el riesgo de cometer un error de tipo II al asegurarse de que la prueba tenga suficiente potencia. Para ello, asegúrese de que el tamaño de la muestra sea lo suficientemente grande como para detectar una diferencia práctica cuando ésta realmente exista.

La probabilidad de rechazar la hipótesis nula cuando es falsa es igual a 1–β. Este valor es la potencia de la prueba.

	Verdad acerca de la población	
Decisión basada en la muestra	H_0 es verdadera	H_0 es falsa
No rechazar H_0	Decisión correcta (probabilidad = 1 - α)	**Error tipo II** - no rechazar H_0 cuando es falsa (probabilidad = β)
Rechazar H_0	**Error tipo I** - rechazar H_0 cuando es verdadera (probabilidad = α)	Decisión correcta (probabilidad = 1 - β)

Ejemplo de error de tipo I y tipo II

Para entender la interrelación entre los errores de tipo I y tipo II, y para determinar cuál error tiene consecuencias más graves para su situación, considere el siguiente ejemplo.

Un investigador médico desea comparar la efectividad de dos medicamentos. Las hipótesis nula y alternativa son:

- Hipótesis nula (H_0): $\mu_1 = \mu_2$
- Los dos medicamentos tienen la misma eficacia.
- Hipótesis alternativa (H_1): $\mu_1 \neq \mu_2$
- Los dos medicamentos no tienen la misma eficacia.

Un error de tipo I se produce si el investigador rechaza la hipótesis nula y concluye que los dos medicamentos son diferentes cuando, en realidad, no lo son. Si los medicamentos tienen la misma eficacia, el investigador podría considerar que este error no es muy grave, porque de todos modos los pacientes se beneficiarían con el mismo nivel de eficacia independientemente del medicamento que tomen. Sin embargo, si se produce un error de tipo II, el investigador no rechaza la hipótesis nula cuando debe rechazarla. Es decir, el investigador concluye que los medicamentos son iguales cuando en realidad son diferentes. Este error puede poner en riesgo la vida de los pacientes si se pone en venta el medicamento menos efectivo en lugar del medicamento más efectivo.

Cuando realice las pruebas de hipótesis, considere los riesgos de cometer errores de tipo I y tipo II. Si las consecuencias de cometer un tipo de error son más graves o costosas que cometer el otro tipo de error, entonces elija un nivel de significancia y una potencia para la prueba que reflejen la gravedad relativa de esas consecuencias.

AWAKELAB

Referencias

[1] Prueba de hipótesis
https://probafacil.com/prueba-de-hipotesis-estadistica/

[2] Hipótesis nula y alternativa
https://tesisdeceroa100.com/aprende-a-diferenciar-entre-hipotesis-nula-y-alternativa-facil/

[3] Significancia Estadística.
https://www.fisterra.com/formacion/metodologia-investigacion/significancia-estadistica-relevancia-clinica/

[4] El valor P
https://conceptosclaros.com/que-es-el-p-valor/
https://es.docpid.com/calculadoras/valor-de-p

[5] Prueba de hipótesis para la proporción poblacional
https://www.youtube.com/watch?v=hm6CkL-Y8vY

[6] Prueba de Hipótesis Para la Media Poblacional
https://www.youtube.com/watch?v=kA8lY_kP4sE