Cómo calcular la potencia de pruebas estadísticas fundamentales

García García Sacbe, Martinez Tizoc Guillermo

Centro de Investigación en Matemáticas
Monterrey, Nuevo León
E-mail: guillermo.martinez@cimat.mx, sacbe.garcia@cimat.mx

7 de diciembre, 2022

I. Introducción

Una prueba de hipótesis consiste en hacer inferencia sobre alguna población de interés. Para el estudio de esta población se proponen dos hipótesis, una nula y su complemento, la alternativa.

La hipótesis nula, escrita como H_0, representa aquel supuesto que se trata de rechazar mediante la prueba que se realiza.

La hipótesis alternativa por el otro lado, es la conjetura que realmente se piensa como la causa del fenómeno que se esté estudiando. También puede expresarse como la hipótesis con la que busca demostrar la falsedad de la hipótesis nula. Suele expresarse como H_1. Durante la prueba, se asume que la hipótesis nula H_0 es verdadera hasta demostrar lo contrario.

Además, existen ciertos valores para los cuales se acepta H_0 como verdadera y para cuales H_1 es verdadera. La región formada por el conjunto de valores que provocan que la hipótesis nula sea rechazada se llama región de rechazo y la región donde estos valores hacen que no sea rechazada se identifica como región de aceptación.

La calidad de una prueba de hipótesis se mide por su capacidad para tomar decisiones correctas sobre la hipótesis nula. Esto se logra mediante el cálculo de dos parámetros importantes: la potencia de la prueba y el nivel de significación. En resumen, una prueba de hipótesis de alta calidad es aquella que tiene una alta potencia y un bajo nivel de significación.

1. Error en las pruebas

Al realizar una prueba de hipótesis, existe el riesgo de aceptar incorrectamente alguna de las dos hipótesis. Este error se divide en dos tipos. El error tipo I ocurre cuando se rechaza la hipótesis nula y ésta era verdadera. El tipo II sucede cuando no se rechaza la hipótesis nula y esta era falsa.

	No Rechazar H_0	Rechazar H_0
H_0 verdadera	✓	Error Tipo I
H_1 verdadera	Error Tipo II	✓

Cuadro 1. Posibles resultados de una prueba de hipótesis.

En la Figura 1 se muestran los dos tipos de errores, el área sombreada a la derecha de 121.9 representa la probabilidad de de cometer un error de tipo I mientras que el área sombreada a la izquierda representa la probabilidad de cometer un error de tipo II. Para seleccionar ambas hipótesis es muy importante hacerlo de manera apropiada, de tal manera que al rechazar o no alguna de ellas podamos obtener conclusiones contundentes sobre la prueba.

Los errores tipo II pueden tener un impacto significativo en diversas áreas, ya que pueden llevar a decisiones incorrectas basadas en la aceptación de la hipótesis nula cuando ésta es falsa.

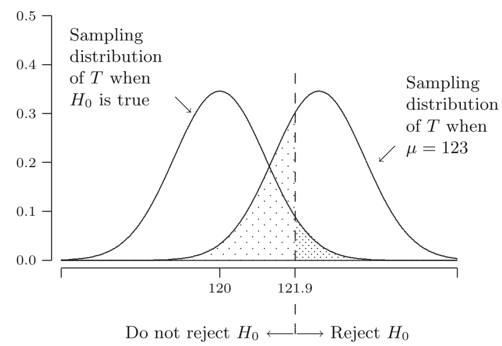

Figura 1. Ejemplo de errores tipo I y tipo II.

Por ejemplo, en el ámbito judicial, un falso negativo puede llevar a la absolución de un acusado cuando en realidad es culpable, lo que puede poner en peligro la seguridad de la comunidad.

2. Entonces, ¿por qué es importante la potencia?

La potencia de las pruebas de hipótesis es importante porque determina la capacidad de la prueba para detectar un efecto o diferencia significativa cuando realmente existe. Cuanto mayor sea la potencia de una prueba, menor será la probabilidad de cometer un error tipo II, es decir, de concluir que no hay diferencia cuando en realidad sí la hay. Esto garantiza una mayor precisión y confiabilidad en los resultados obtenidos. Además, una prueba de hipótesis con alta potencia permite utilizar muestras más pequeñas, lo que reduce costos y tiempos en los estudios.

II. Potencia de una prueba

1. Función de potencia

La potencia de una prueba de hipótesis es la probabilidad de que la prueba rechace correctamente la hipótesis nula cuando la hipótesis alternativa es cierta. La función de potencia de una prueba estadística con región de rechazo R está definida por

$$\beta(\theta) = P_\theta(X \in R)$$

Donde θ denota un parámetro de población. El tamaño de la prueba se define como

$$\alpha = \sup_{\theta \in \Theta_0} \beta(\theta).$$

Decimos que una prueba tiene un nivel de significación α si el tamaño es menor o igual a α. También es usual encontrar la potencia de la siguiente manera.

$$\text{Potencia} = 1 - \beta = P(\text{ Rechazar } H_0 | H_A),$$

Donde β es la probabilidad de cometer un error de tipo II.

Ejemplo 1 Sea $X_1, ..., X_n \sim N(\mu, \sigma)$ donde σ es conocido. Se busca probar H_0: $\mu \leq 0$ contra H_1: $\mu > 0$. Por lo tanto, $\Theta_0 = (-\infty, 0]$ y $\Theta_1 = (0, \infty)$.

Considérese la prueba: Rechazar H_0 si $T > c$. Donde $T = \overline{X}$. La región de rechazo entonces queda definida por:

$$R = \{(x_1, ..., x_n) : T(x_1, ..., x_n) > c\}$$

Denotando como Z a una variable aleatoria normal estándar, la función de potencia queda de la siguiente manera:

$$\begin{aligned}
\beta(\mu) &= P_\mu(\overline{X} > c) \\
&= P_\mu\left(\frac{\sqrt{n}(\overline{X} - \mu)}{\sigma} > \frac{\sqrt{n}(c - \mu)}{\sigma}\right) \\
&= P\left(Z > \frac{\sqrt{n}(c - \mu)}{\sigma}\right) \\
&= 1 - \Phi\left(\frac{\sqrt{n}(c - \mu)}{\sigma}\right)
\end{aligned}$$

Ya que la función es creciente con respecto a μ, entonces el nivel de significación se alcanza cuando $\mu = 0$ ya que por la hipótesis nula tenemos H_0: $\mu \leq 0$. Entonces para tener un nivel de significación α igualamos como sigue

$$\sup_{\mu \leq 0} \beta(\mu) = \beta(0) = 1 - \Phi\left(\frac{\sqrt{n}c}{\sigma}\right) = \alpha.$$

Despejando c, el valor crítico está dado por:

$$c = \mu + \frac{\sigma \Phi^{-1}(1 - \alpha)}{\sqrt{n}}$$

Se rechaza cuando:

$$\frac{\sqrt{n}(\overline{X} - 0)}{\sigma} > z_\alpha$$

Donde:

$$z_\alpha = \Phi^{-1}(1 - \alpha)$$

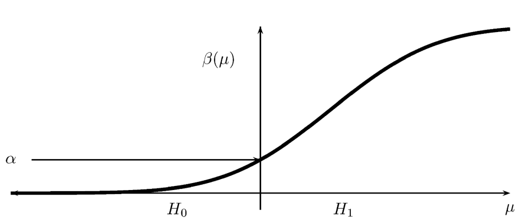

Figura 2. Ejemplo de una curva de potencia

Ejemplo 2 Asúmase que se ha desarrollado un nuevo proceso químico que puede incrementar el rendimiento en comparación con el proceso que se tiene actualmente. Se sabe que el proceso actual tiene un rendimiento medio de 80 y una desviación estándar de 5, donde las unidades son el porcentaje de un máximo teórico. Si el rendimiento medio del nuevo proceso se muestra mayor a 80, este será implementado. Sea μ la media de rendimiento del nuevo proceso. Se ha propuesto correr el nuevo proceso 50 veces para luego probar la hipótesis.

$$H_0 : \mu \leq 80 \qquad \text{vs} \qquad H_1 : \mu > 80$$

Se toma en cuenta un nivel de significación de 5 %. Si H_0 es rechazada, se concluye que $\mu > 80$ y el nuevo proceso será puesto en producción. Se asume, además, que si el nuevo proceso tiene un rendimiento medio de 81, se considera un beneficio sustancial para usar este nuevo proceso. Si se cumple que en realidad $\mu = 81$, ¿cuál es la potencia de la prueba?

Primero, es necesario establecer un valor específico para μ, en este caso $\mu = 81$ para la hipótesis alternativa. La razón de esto es que la potencia es diferente para diferentes valores de μ. Si μ está cerca de H_0, la potencia será baja, mientras que si μ está alejada de H_0, la potencia será alta.

El cálculo de la potencia involucra dos pasos:

1. Calcular la región de rechazo.
2. Calcular la probabilidad de que el estadístico de prueba caiga en la región de rechazo si la hipótesis alternativa es verdadera, es decir, calcular la potencia.

Se tiene que $\overline{X}$ es el estadístico de prueba y está distribuido normalmente con media μ y una desviación estándar $\sigma_{\overline{x}} = \frac{\sigma}{\sqrt{n}}$, donde el tamaño de muestra es $n = 50$. Bajo la hipótesis nula se toma a $\mu = 80$. Hallar una aproximación para la desviación estándar es complicado ya que no se ha corrido ninguna prueba. Supóngase que se tiene un historial de una prueba parecida donde la desviación estándar es 5. Entonces

$$\sigma_{\overline{X}} = \frac{5}{\sqrt{50}} = 0{,}707$$

En la Figura 3 se muestra la distribución de $\overline{X}$. Ya que para este caso $\mu = 80$, los valores mayores a 80 comprenderán la región de rechazo, es decir, se ubicará sobre la derecha de la curva. Además, se tiene que el nivel de significación $\alpha = 5\,\%$, indicando que el 5 % superior de la curva será la región de rechazo.

Para obtener el valor crítico, se sigue la ecuación:

$$c = \mu + [(z)(\sigma_{\overline{X}})]$$

Donde el valor z se obtiene usando tablas de distribución normal estándar o con la ayuda de un software directamente. Entonces, para el caso particular se tiene que:

$$\begin{aligned}
c &= 80 + (1{,}645)(0{,}707) \\
&= 81{,}16
\end{aligned}$$

Tenemos pues, que H_0 se va a rechazar si $\overline{X} \geq 81{,}16$.

Ahora, para calcular la potencia, debemos obtenerla probabilidad de que $\overline{X}$ caiga en la región de rechazo si la hipótesis alternativa $\mu = 81$ es verdadera. Las consideraciones que se tienen para la hipótesis alternativa es que $\overline{X}$ tiene una distribución normal con media de 81 y desviación estándar de 0.707. Se tiene la misma curva que con la hipótesis nula, solo que ahora

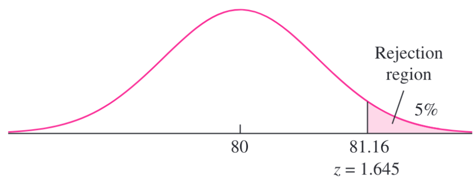

Figura 3. Región de rechazo para la distribución nula

la media está ubicada en 81. La Figura 4 muestra gráficamente esto.

Tenemos pues, que el valor crítico para este estadístico es 81.16, por lo que z está dado por:

$$z = \frac{(c-\mu)}{\sigma_{\overline{X}}} = \frac{(81{,}16-81)}{0{,}707} = 0{,}23$$

Usando una vez más la tabla o el software, se indica que el área a la derecha de $z = 0{,}23$ es 0,4090. Por lo tanto:

$$\text{Potencia} = 0{,}4090$$

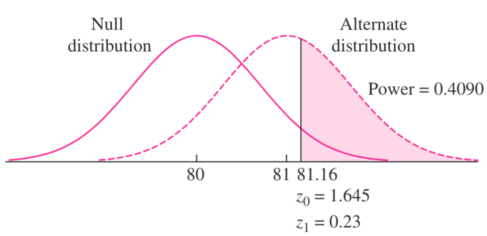

Figura 4. Región de rechazo para ambas distribuciones

Esto significa que si realmente el nuevo proceso tiene un rendimiento medio de 81, existe una probabilidad del 41 % de que el experimento propuesto vaya a detectar la mejora del nuevo proceso respecto al anterior. Esto es es muy bajo y no valdría la pena invertir tiempo ni dinero en una prueba que tiene altas probabilidades de fallar. De manera general, se consideran aceptables pruebas con potencias mayores al 80 %.

Para calcular la potencia de una prueba, es necesario especificar la hipótesis alternativa, el nivel de significación α deseado y el tamaño de la muestra.

Es importante tener en cuenta que la potencia de una prueba no se puede calcular sin especificar la hipótesis alternativa, ya que la potencia depende de la diferencia entre la hipótesis nula y la hipótesis alternativa. Además, a medida que aumenta el nivel de significación de la prueba, disminuye la potencia de la prueba. Por lo tanto, es importante equilibrar el nivel de significación y la potencia de una prueba para obtener resultados confiables y significativos.

Los cálculos de la potencia generalmente son realizados antes de la recolección de datos ya que su propósito es determinar si la prueba de hipótesis va a tener la capacidad de rechazar la hipótesis nula cuando esta sea falsa.

III. Curvas de potencia

La hipótesis alternativa H_1 suele depender de un parámetro, lo que hace que $1-\beta$ sea una función de ese parámetro. La relación que comparten puede representarse dibujando una curva de potencia, que no es más que un gráfico de $1-\beta$ frente al conjunto de todos los valores posibles del parámetro.

En el ejemplo anterior:

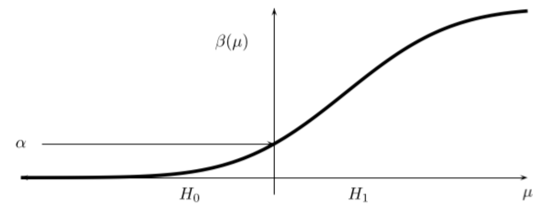

Figura 5. Curva de potencia una v.a. $N(\theta, \sigma)$ con las hipótesis dadas.

Las curvas de potencia también son útiles para comparar un procedimiento de inferencia con otro. Para cada situación concebible de comprobación de hipótesis, se dispondrá de una variedad de procedimientos para elegir entre H_0 y H_1. ¿Cómo sabemos cuál utilizar? La respuesta a esta pregunta no siempre es sencilla. Algunos procedimientos serán computacionalmente más convenientes que otros; algunos harán suposiciones ligeramente diferentes sobre la pdf. La figura muestra las curvas

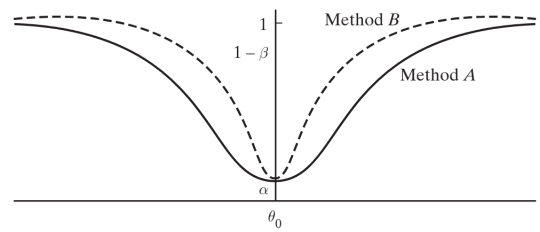

Figura 6. Ejemplo de curva de potencia.

de potencia para dos métodos hipotéticos A y B, cada uno de los cuales prueba $H_0 : \theta = \theta_0$ frente a $H_1 : \theta \neq \theta_0$ al nivel de significación α. Desde el punto de vista de la potencia, el método B es claramente el mejor de los dos: siempre tiene una mayor probabilidad de rechazar correctamente H_0 cuando el parámetro θ no es igual a θ_0.

IV. Factores que influyen en la potencia de una prueba

1. Aspectos generales

La potencia es una función de α, σ y n, por lo que el cambio de cualquiera de ellas afectará el resultado.

El cambiar el nivel de significación α representa el alterar el área de rechazo de la distribución nula, por lo tanto, se tiene que aumentar α incrementa la potencia de la prueba y viceversa.

En el caso de σ se tiene que si se reduce, la potencia de la prueba aumenta ya que se concentra sobre la media de la distribución.

Cuanto mayor es el tamaño de muestra, mayor es la potencia de la prueba. Esto se debe a que una muestra más grande proporciona más información y permite una mayor precisión en la estimación de la población. Como resultado, una prueba con un tamaño de muestra más grande tiene una mayor probabilidad de rechazar correctamente la hipótesis nula cuando la hipótesis alternativa es cierta.

2. El efecto de α sobre $1-\beta$

Sea $X_1, ..., X_n \sim N(\mu, \sigma)$ donde σ es conocido. Se busca probar H_0: $\theta = 25{,}0$ vs H_1: $\theta > 25$.

Si establecemos $\alpha = 0{,}05$, $\sigma = 2{,}4$, $n = 30$ y digamos que H_0 es rechazada si $\bar{y} > 25{,}718$. La potencia en este caso es 0,53, sin embargo, si mantenemos n, σ, μ contantes pero $\alpha = 0{,}10$ incrementa entonces la regla de decisión cambia a $\bar{y} > 25{,}561$ y el poder aumenta de 0,53 a 0,67

$$\begin{aligned}
1-\beta &= P(\text{ Rechazar } H_0 \mid H_1 \text{ es verdadero })\\
&= P(\bar{Y} \geq 25{,}561 \mid \mu = 25{,}75)\\
&= P\left(\frac{\bar{Y}-25{,}75}{2{,}4/\sqrt{30}} \geq \frac{25{,}561-25{,}75}{2{,}4/\sqrt{30}}\right)\\
&= P(Z \geq -0{,}43)\\
&= 0{,}6664
\end{aligned}$$

En la practica los experimentadores no aumentan α para lograr un $1-\beta$ deseado.

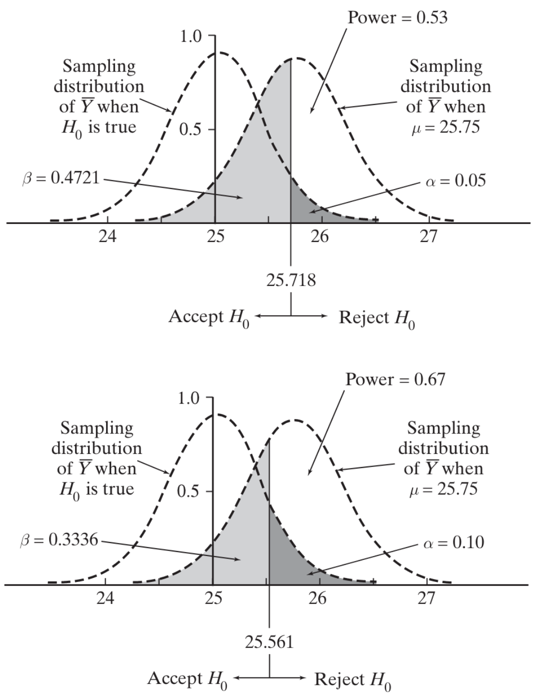

Figura 7. Incrementar α disminuye β, es decir aumenta su potencia.

3. El efecto de σ y n sobre $1-\beta$

A pesar de que no siempre será factible (o incluso posible), disminuir la desviación estándar σ va a aumentar la potencia. Tomemos el ejemplo de la diapositiva anterior. Se tiene que la desviación estándar es de 2.4 millas por galón. ¿Cuál es el efecto sobre la potencia si σ baja de 2.4 a 1.2?

Se observa que la distribución nula de $\overline{Y}$ se concentra más sobre $\mu_0 = 25$ y que la distribución de $\overline{Y}$ bajo H_1 se concentra alrededor de $\mu = 25{,}75$

El nuevo valor crítico, dado por:

$$c = 25 + (1{,}64)\left(\frac{1{,}2}{\sqrt{30}}\right) = 25{,}359$$

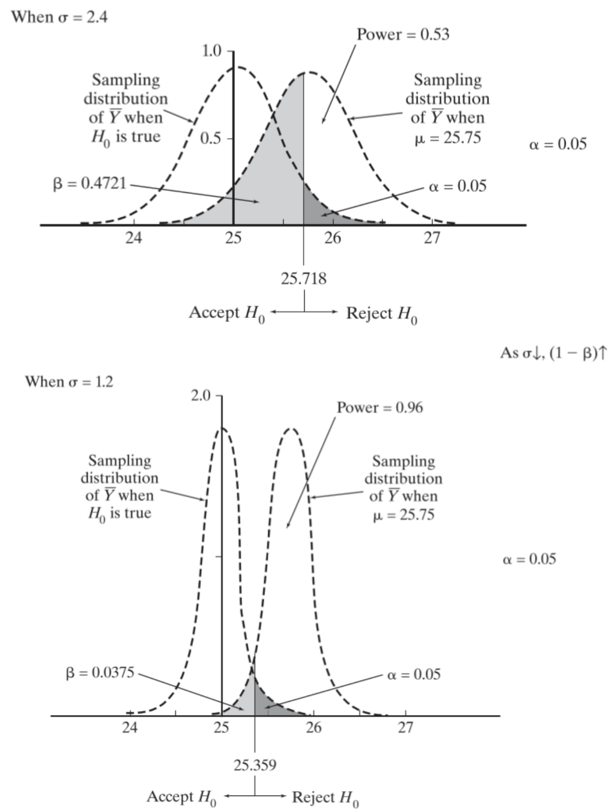

Figura 8. Efecto de disminuir σ sobre la potencia

Ocasiona que la potencia sea:

$$\begin{aligned}
1-\beta &= P(\overline{Y} \geq 25{,}359 \mid \mu = 25{,}75)\\
&= P\left(Z \geq \frac{25{,}359-25{,}75}{1{,}2/\sqrt{30}}\right)\\
&= P(Z \geq -1{,}78) = 0{,}9625
\end{aligned}$$

El incremento en el poder se logró al disminuir el denominador en el estadístico de prueba

$$z = \frac{\bar{y}-25}{\sigma/\sqrt{30}}$$

Si ahora tomamos en cuenta el otro elemento que compone el denominador, que es $\sqrt{n}$, se intuye que puede ocasionar el mismo efecto si cambiamos su tamaño. Es decir, si aumentamos n de 30 a 120, lograríamos el mismo valor de z. Ya que el valor de n puede ser más fácilmente disminuido o aumentado, los investigadores usan este parámetro como mecanismo para que la prueba de hipótesis tenga una mayor potencia.

V. Tamaño de una muestra

1. Importancia del tamaño de la muestra

Aunque en general aumentar el tamaño de la muestra puede mejorar la precisión y la potencia de una prueba de hipótesis, no siempre es posible o conveniente aumentar el tamaño de la muestra. Una de las principales razones por las que no siempre se aumenta el tamaño de la muestra es el costo y el tiempo que implica recolectar datos adicionales. Por ejemplo, si se está realizando un estudio clínico con pacientes, aumentar el tamaño de la muestra puede implicar incluir a más pacientes en el estudio, lo que puede aumentar el costo y el tiempo del estudio. Además, también es posible que aumentar el tamaño de la muestra no aporte un beneficio significativo en términos de precisión o potencia de la prueba, lo que puede hacer que no sea necesario o conveniente realizar un esfuerzo adicional para recolectar datos adicionales.

2. Cálculo del tamaño de la muestra

Supongamos que el experimentador desea tener un error tipo I máximo de 0,1. Supongamos en adición que el experimentador desea tener un error tipo II máximo de 0,2 si $\theta > \theta_0 + \sigma$.

Se mostrará como escoger c y n para alcanzar ese objetivo. Usando una prueba que rechaza H_0: $\theta \leq \theta_0$ si $(\overline{X} - \theta_0)/(\sigma/\sqrt{n}) > c$. Como se mostró en el cálculo de la función de potencia para una Normal(θ, σ) tenemos que

$$\beta(\theta) = 1 - \Phi(c + \frac{\theta_0 - \theta}{\sigma/\sqrt{n}})$$

Dado que $\beta(\theta)$ es creciente en θ los requisitos se cumplirán si

$$\beta(\theta_0) = 0{,}1 \qquad \text{y} \qquad \beta(\theta_0 + \sigma) = 0{,}8$$

Eligiendo $c = 1{,}28$ logramos $\beta(\theta_0 = 0{,}1) = P(Z > 1{,}28) = 0{,}1$ sin importar el n.

Ahora deseamos escoger un n tal que $\beta(\theta_0 + \sigma) = P(Z > 1{,}28 - \sqrt{n}) = 0{,}8$, pero ya que

$$P(Z > -{,}84) = 0{,}8$$

entonces igualamos $1{,}28 - \sqrt{n} = -0{,}84$ y despejamos n. Obtenemos que $n = 4{,}49 \sim 5$. Así, eligiendo $c = 1{,}28$ y $n = 5$ nos lleva a una prueba con probabilidades de error controlados por el experimentador.

VI. Potencia de diferentes pruebas

1. Aspectos generales

La prueba F de Fisher y la prueba t de Student son dos pruebas estadísticas que se utilizan para comparar dos grupos de datos y determinar si hay una diferencia significativa entre ellos. La principal diferencia entre ambas pruebas es que la prueba F se utiliza para comparar dos o más varianzas, mientras que la prueba t se utiliza para comparar dos o más medias.

Por último, la prueba F se utiliza cuando se conocen las varianzas de los grupos, mientras que la prueba t se utiliza cuando solo se conocen las medias y se asume que las varianzas son iguales. Esto significa que la prueba F es más precisa que la prueba t cuando se conocen las varianzas, pero menos precisa cuando solo se conocen las medias. La prueba F también es útil cuando se desconoce la forma de la distribución de la población.

Para calcular la potencia de una prueba t o una prueba F, es necesario especificar la hipótesis alternativa, el nivel de significación deseado y el tamaño de la muestra. Con esta información, se puede calcular el valor crítico de la prueba y luego utilizar una distribución de probabilidad t o F para calcular la probabilidad de que la prueba rechace correctamente la hipótesis nula.

A. Potencia de una prueba t con muestra única

Las expresiones para el estadístico t depende del tipo de hipótesis a realizar. Entonces la ecuación explícita para la función de potencia de prueba t con muestra única es la siguiente:

$$\text{Potencia} = P(\text{Rechazar } H_0|H_A) = 1 - \beta = \int_{\frac{\bar{x}-\mu_0}{\frac{s}{\sqrt{n}}}}^{+\infty} t_{n-1,\alpha}(x)dx$$

donde $\bar{x}$ es la media muestral, s es la desviación estándar muestral y n es el tamaño de la muestra. Los grados de libertad utilizados en esta prueba se corresponden al valor $n-1$.

Las siguientes manipulaciones algebraicas ayudan a poner la densidad necesaria en una forma reconocible.

$$\frac{\overline{Y} - \mu_0}{S/\sqrt{n}} = \frac{\overline{Y} - \mu_1 + (\mu_1 - \mu_0)}{S/\sqrt{n}} = \frac{\frac{\overline{Y}-\mu_1}{\sigma} + \frac{(\mu_1-\mu_0)}{\sigma}}{\frac{S/\sqrt{n}}{\sigma}} = \frac{\frac{\overline{Y}-\mu_1}{\sigma/\sqrt{n}} + \frac{(\mu_1-\mu_0)}{\sigma/\sqrt{n}}}{S/\sigma}$$

$$= \frac{\frac{\overline{Y}-\mu_1}{\sigma/\sqrt{n}} + \frac{(\mu_1-\mu_0)}{\sigma/\sqrt{n}}}{\sqrt{\frac{(n-1)S^2/\sigma^2}{n-1}}} = \frac{\frac{\overline{Y}-\mu_1}{\sigma/\sqrt{n}} + \delta}{\sqrt{\frac{(n-1)S^2/\sigma^2}{n-1}}} = \frac{Z + \delta}{\sqrt{\frac{U}{n-1}}}$$

donde $Z = \frac{\overline{Y}-\mu_1}{\sigma/\sqrt{n}}$ es normal, $U = \frac{(n-1)S^2}{\sigma^2}$ es una variable chi cuadrado con $n-1$ grados de libertad, y $\delta = \frac{(\mu_1-\mu_0)}{\sigma/\sqrt{n}}$ es una constante (desconocida). Una distribución de esta forma se dice que tiene una distribución t no central.

B. Potencia de una prueba t con dos muestras independientes

Esta prueba se utiliza solo si

- Los dos tamaños muestrales (esto es, el número, n, de participantes en cada grupo) son iguales.
- Se puede asumir que las dos distribuciones poseen la misma varianza.

La ecuación explícita para la función de potencia de prueba t con muestra única es la siguiente:

$$\text{Potencia} = P(\text{Rechazar } H_0|H_A)$$
$$= 1 - \beta = \int_{\frac{\bar{X}_1 - \bar{X}_2}{s_{X_1X_2}\sqrt{\frac{2}{n}}}}^{+\infty} t_{2n-2,\alpha}(x)dx$$

donde

$$S_{X_1X_2} = \sqrt{\frac{1}{2}(S_{X_1}^2 + S_{X_1}^2)},$$

es la desviación estándar combinada.

C. Potencia de una prueba t diferentes tamaños muestrales, iguales varianzas

Esta prueba se puede utilizar únicamente si se puede asumir que las dos distribuciones poseen la misma varianza

La ecuación explícita para la función de potencia de prueba t con muestra única es la siguiente:

$$\text{Potencia} = P(\text{Rechazar } H_0|H_A)$$
$$= 1 - \beta = \int_{\frac{\bar{X}_1 - \bar{X}_2}{s_{X_1X_2}\sqrt{\frac{1}{n_1}+\frac{1}{n_2}}}}^{+\infty} t_{n_1+n_2-2,\alpha}(x)dx$$

donde

$$S_{X_1X_2} = \sqrt{\frac{(n_1-1)(S^2_{X_1} + (n_2-1)S^2_{X_1})}{n_1+n_2-2}},$$

es la desviación estándar combinada.

D. Potencia de una prueba t diferentes tamaños muestrales, diferentes varianzas

Esta prueba es también conocida como prueba t de Welch y es utilizada únicamente cuando se puede asumir que las dos varianzas poblacionales son diferentes (los tamaños muestrales pueden o no ser iguales) y por lo tanto deben ser estimadas por separado.

La ecuación explícita para la función de potencia de prueba t con muestra única es la siguiente:

$$\begin{aligned}\text{Potencia} &= P(\text{Rechazar } H_0|H_A)\\ &= 1-\beta = \int_{\frac{\bar{X}_1-\bar{X}_2}{S_{\bar{X}_1-\bar{X}_2}}}^{+\infty} t_{g.l.,\alpha}(x)dx\end{aligned}$$

donde

$$S_{\bar{X}_1-\bar{X}_2} = \sqrt{\frac{s_1^2}{n_1}+\frac{s_2^2}{n_2}},$$

donde s_1^2, s_2^2 son los estimadores sin sesgo de las varianzas de la muestra. Además

$$\text{g. l. } = \frac{\left(s_1^2/n_1+s_2^2/n_2\right)^2}{\left(s_1^2/n_1\right)^2/(n_1-1)+\left(s_2^2/n_2\right)^2/(n_2-1)}$$

son los grados de libertad.

Ejemplo 2 Supongamos que queremos probar $H_0 : \mu = \mu_0$ frente a $H_1 : \mu > \mu_0$ al nivel de significación $\alpha = 0,05$. Sea $n = 20$. En este caso, la prueba consiste en rechazar H_0 si el estadístico de prueba $\frac{\bar{y}-\mu_0}{s/\sqrt{n}}$ es mayor que $t_{,05,19} = 1{,}7291$. ¿Cuál será el error de tipo II si la media se ha desplazado 0,5 desviación típica a la derecha de μ_0?

Decir que la media se ha desplazado 0,5 desviación típica a la derecha de μ_0 equivale a establecer $\frac{\mu_1-\mu_0}{\sigma} = 0,5$. En ese caso, el parámetro de no centralidad es $\delta = \frac{\mu_1-\mu_0}{\sigma/\sqrt{n}} = (0,5)\cdot\sqrt{20} = 2,236$. La probabilidad de un error de tipo II es

$$P\left(T_{19,2,236} \leq 1{,}7291\right) = 0{,}304$$

donde $T_{19,2,236}$ es una variable no central t con 19 grados de libertad y parámetro de no centralidad 2,236.

2. Potencia de una prueba F

Si queremos probar H_0: $\sigma_1^2 = \sigma_2^2$ vs H_0: $\sigma_1^2 > \sigma_2^2$ La ecuación explícita para la función de potencia de prueba F es la siguiente:

$$\begin{aligned}\text{Potencia} &= P(\text{Rechazar } H_0|H_A)\\ &= 1-\beta = 1-\int_{\frac{\sigma_1^2}{\sigma_2^2}}^{+\infty} F_{n_1-1,n_2-1,\alpha}(x)dx\end{aligned}$$

En esta ecuación, σ_1^2 y σ_2^2 son las varianzas de las dos poblaciones bajo la hipótesis alternativa, n_1 y n_2 son los tamaños de las muestras de cada población y α es el nivel de significación de la prueba. La función $F_{n_1-1,n_2-1,\alpha}(x)$ es la función de distribución F de Fisher-Snedecor con n_1-1 y n_2-1 grados de libertad y un nivel de significación de α.

Si queremos probar H_0: $\sigma_1^2 = \sigma_2^2$ vs H_0: $\sigma_1^2 < \sigma_2^2$ La ecuación explícita para la función de potencia de prueba F es la siguiente:

$$\begin{aligned}\text{Potencia} &= P(\text{Rechazar } H_0|H_A)\\ &= 1-\beta = \int_{\frac{\sigma_1^2}{\sigma_2^2}}^{+\infty} F_{n_1-1,n_2-1,\alpha}(x)dx\end{aligned}$$

Por último si queremos probar H_0: $\sigma_1^2 = \sigma_2^2$ vs H_0: $\sigma_1^2 \neq \sigma_2^2$ La ecuación explícita para la función de potencia de prueba F es la siguiente:

$$\begin{aligned}\text{Potencia} &= P(\text{Rechazar } H_0|H_A)\\ &= 1-\beta = \int_{\frac{\sigma_1^2}{\sigma_2^2}}^{+\infty} F_{n_1-1,n_2-1,1-\alpha/2}(x)dx.\end{aligned}$$

El estadístico σ_1^2/σ_2^2 sigue una distribución llamada distribución F no central con densidad

$$p(f) = \sum_{k=0}^{\infty} \frac{e^{-\lambda/2}(\lambda/2)}{B(\frac{\nu_2}{2},\frac{\nu_1}{2}+k)k!}\left(\frac{\nu_1}{\nu_2}\right)^{\frac{\nu_1}{2}+k}\left(\frac{\nu_2}{\nu_2+\nu_1 f}\right)^{\frac{\nu_1+\nu_2}{2}+k} f^{\nu_1/2-1+k}$$

donde $f > 0$. Los grados de libertad ν_1 y ν_2 son positivos y $B(x, y)$ es la función beta. Si $\lambda = 0$ la distribución F no central se convierte en la distribución F.

VII. Ejercicio de tarea

Siguiendo el ejemplo 2 encuentre la potencia de la prueba de nivel del 5 % de H_0: $\mu \leq 80$ frente a H_1: $\mu > 80$ para el rendimiento medio del nuevo proceso bajo la alternativa $\mu = 82$, suponiendo $n = 50$ y $\sigma = 5$.

Referencias

[1] R. J. Larsen y M. L. Marx, An introduction to mathematical statistics and its applications, 5th ed. Boston: Prentice Hall, 2012.
[2] G. Casella y R. L. Berger, Statistical inference, 2nd ed. Australia Pacific Grove, CA: Thomson Learning, 2002.
[3] M. Dekking, Ed., A modern introduction to probability and statistics: understanding why and how. London: Springer, 2005.
[4] L. Wasserman, All of Statistics: A Concise Course in Statistical Inference. New York, NY: Springer New York, 2004. doi: 10.1007/978-0-387-21736-9.
[5] W. Navidi, W.C. Navidi, Statistics for Engineers and Scientists. New York, NY: McGraw-Hill, 2011. ISBN: 978-0-07-337633-2.
[6] J. A. Rice, Mathematical Statistics and Data Analysis. Belmont, CA: Duxbury, 2007. ISBN: 0-534-39942-8.