

Tests de hipótesis

Introducción

En el capítulo anterior vimos cómo estimar un parámetro con un solo número (una estimación puntual) o un intervalo completo de valores plausibles (un intervalo de confianza), a partir de los datos muestrales. Con frecuencia, sin embargo, el objetivo de una investigación no es estimar un parámetro sino decidir cuál de dos pretensiones contradictorias sobre él es la correcta. Los métodos para lograr esto constituyen la parte de la inferencia estadística llamada pruebas o tests de hipótesis.

En muchos aspectos, el procedimiento formal para pruebas de hipótesis es semejante al método científico. Observar la naturaleza, formular una teoría y confrontarla con lo observado.

En nuestro contexto, tenemos una hipótesis sobre la población (por ejemplo acerca de su media o de su varianza) y queremos saber si es cierta o no, para verificarla tomamos una muestra aleatoria, y en función de la misma, decidimos si la aceptamos o no.

Ahora bien, ¿cómo se utiliza la estadística en este procedimiento? Probar una hipótesis requiere tomar una decisión cuando se compara la muestra observada contra la teoría propuesta. ¿Cómo decidimos si la muestra concuerda o no con la hipótesis del científico? ¿Cuándo debemos rechazar la hipótesis y cuándo debemos aceptarla? ¿Cuál es la probabilidad de que tomemos una mala decisión? Y, en particular, ¿qué función de las mediciones muestrales debe emplearse para llegar a una decisión? En este capítulo trataremos de dar respuesta a estas preguntas.

Test de hipótesis para la media de una distribución normal con varianza conocida

Para presentar las ideas de test de hipótesis, comencemos desarrollando un ejemplo:

Una concentración de mercurio (Hg) en el agua igual a 1 *mcg/l*, se considera un riesgo para la salud. Se realizan 6 determinaciones de concentración de Hg en una muestra de agua y se obtienen los siguientes valores (*mcg/l*):

0.86 0.89 0.93 1.02 0.96 0.83

Suponemos que el método tiene un error de medición con $\sigma = 0.08$ *mcg/l*. Queremos determinar si esa muestra de agua puede considerarse libre de contaminación. Para ello, debemos decidir entre estas dos hipótesis:

- La concentración de Hg en la muestra de agua es admisible.
- La concentración de Hg en la muestra de agua es un riesgo para la salud.

Podemos modelizar esta situación como sigue: tenemos una m.a. X_1, X_2, \dots, X_6 donde cada X_i es el resultado de la i -ésima determinación y tiene distribución $N(\mu, 0.08^2)$. El problema a resolver es decidir cuál de las siguientes hipótesis es cierta:

$$H_0 : \mu = 1 \quad (\text{el agua es un riesgo para la salud})$$

$$H_1 : \mu < 1 \quad (\text{el agua es admisible})$$

donde H_0 se llama **hipótesis nula** y H_1 **hipótesis alternativa**.

Debemos notar que al decidirnos por una de las dos hipótesis, pueden ocurrir alguna de estas cuatro situaciones:

Decisión a partir del test	H_0 es cierta	H_0 no es cierta
Se rechaza H_0	<i>error tipo I</i>	correcto
No se rechaza H_0	correcto	<i>error tipo II</i>

Como se aprecia en la tabla anterior, pueden cometerse dos tipos de errores, que se los distingue con los nombres de **error de tipo I** y **error de tipo II**.

En nuestro ejemplo el error de tipo I sería afirmar que la muestra de agua tiene una concentración de mercurio aceptable ($\mu < 1$), cuando en realidad la concentración es alta. Y el error de tipo II sería afirmar que la muestra de agua tiene una concentración alta ($\mu = 1$), cuando en realidad es más baja. En este caso, el error de tipo I es más grave que el error de tipo II.

Los procedimientos que vamos a ver, nos permiten acotar la probabilidad de cometer un error de tipo I.

Recordemos que nunca conocemos cuánto vale μ y sólo podemos hacer inferencias, acerca de su valor, basadas en la muestra. Sabemos que \bar{X} es un estimador de μ y el valor \bar{x} observado es una estimación del verdadero μ . Entonces, si \bar{x} resulta mucho más chico que 1, tendremos motivos para pensar que en realidad $\mu < 1$ (cuanto menor sea \bar{x} , mayor será la evidencia contra H_0 a

favor de H_1). Debemos decidir cuándo consideramos que \bar{x} es lo suficientemente pequeño como para rechazar la hipótesis nula, manteniendo acotada la probabilidad de error de tipo I. Para esto debemos considerar un estadístico (estadístico de prueba) con distribución conocida cuando H_0 es verdadera y definir una zona de rechazo. En nuestro ejemplo, usaremos el estadístico de prueba:

$$Z = \frac{\bar{X} - 1}{0.08/\sqrt{6}} = \frac{\sqrt{6}(\bar{X} - 1)}{0.08}$$

que bajo el modelo propuesto y cuando H_0 es verdadera, tiene distribución $N(0, 1)$.

Supongamos que queremos que la probabilidad de error de tipo I sea 0.05. Se puede establecer una regla de decisión como la siguiente: rechazar H_0 cuando el valor del estadístico de prueba sea menor que -1.645 ; de este modo nos aseguramos que:

$$P(\text{error de tipo I}) = P\left(\frac{\sqrt{6}(\bar{X} - 1)}{0.08} < -1.645 \mid H_0 \text{ es verdadera}\right) = \Phi(-1.645) = 0.05$$

En general, para que la probabilidad de error de tipo I sea α la regla será:

$$\text{rechazar } H_0 : \mu = 1 \text{ cuando } z = \frac{\sqrt{6}(\bar{x} - 1)}{0.08} < -z_\alpha$$

ya que:

$$P(\text{error de tipo I}) = P\left(\frac{\sqrt{6}(\bar{X} - 1)}{0.08} < -z_\alpha \mid H_0 \text{ es verdadera}\right) = \Phi(-z_\alpha) = \alpha$$

Este valor α se llama **nivel de significación**. Al fijar un nivel de significación $\alpha = 0.05$, nos aseguramos que la probabilidad de cometer un error de tipo I, no puede ser mayor que 0.05. Para este test, llamamos zona de rechazo a los valores del estadístico de prueba menores que $-z_\alpha$, puede verse que el área bajo la curva de densidad en la zona de rechazo es igual a α .

Con los datos del ejemplo, al reemplazar \bar{X} por $\bar{x} = 0.915$, el valor que toma el estadístico es: $\frac{\sqrt{6}(0.915-1)}{0.08} = -2.60$, este valor cae en la zona de rechazo, de modo que podemos rechazar H_0 con nivel $\alpha = 0.05$. Es decir, podemos afirmar que la concentración de Hg en la muestra de agua está dentro de los niveles admisibles, y la probabilidad de equivocarnos al hacer esta afirmación es a lo sumo 0.05. También se dice que el resultado es significativo al 5 %.

También podemos razonar de esta manera: si $\mu = 1$, ¿cuál es la probabilidad de obtener una media muestral que fuera tanto o más pequeña que la observada?, o lo que es equivalente, ¿cuál es la probabilidad de que el estadístico de prueba sea menor o igual que -2.60 , si $\mu = 1$? Esta probabilidad puede calcularse ya que el estadístico de prueba tiene distribución $N(0, 1)$ cuando $\mu = 1$ y es:

$$P(Z < -2.60) = \Phi(-2.60) = 0.0047$$

Esto es lo que se llama el “valor- p ”, cuanto menor sea este p , más evidencia tengo contra H_0 . En nuestro ejemplo, valor- $p = 0.0047$ es una probabilidad muy pequeña, podemos rechazar H_0 y afirmar la hipótesis alternativa, es decir que $\mu < 1$ con fuerte convicción.

El valor- p se puede definir como el menor nivel de significación (el más exigente) para el cual se puede rechazar H_0 con los valores observados. Otra manera de expresar la regla de decisión es: se rechaza H_0 cuando el valor- p es menor que el α elegido.

En este caso el valor- $p = 0.0047$, esto significa que podemos rechazar H_0 con cualquier $\alpha \geq 0.0047$ que es equivalente a decir que el resultado es significativo al 0.47 %.

Cualquier prueba de hipótesis estadística funciona de la misma forma y está compuesta de los mismos elementos esenciales.

Un **test o prueba de hipótesis** es un proceso de decisión que, en función de los datos de una m.a. X_1, X_2, \dots, X_n , nos permite decidir entre la validez de dos hipótesis contradictorias.

Los elementos de un test de hipótesis son:

1. Las hipótesis: hipótesis nula e hipótesis alternativa.
2. El estadístico de prueba.
3. La región de rechazo.

Definiremos a continuación formalmente estos elementos:

Definición:

Las hipótesis **nula** y **alternativa** son hipótesis contradictorias sobre el parámetro de interés. Se decide cuál es la hipótesis alternativa, de modo que el error de tipo I sea el más grave (ya que la probabilidad de cometer dicho error quedará acotada). Habitualmente, en los trabajos de investigación, la hipótesis alternativa es la que el investigador desea demostrar.

Definición:

El **estadístico de prueba** es una función de la muestra aleatoria, y debe tener una distribución conocida cuando el valor del parámetro sea el que indica la hipótesis nula. La decisión estadística estará basada en el valor del estadístico de prueba.

Definición:

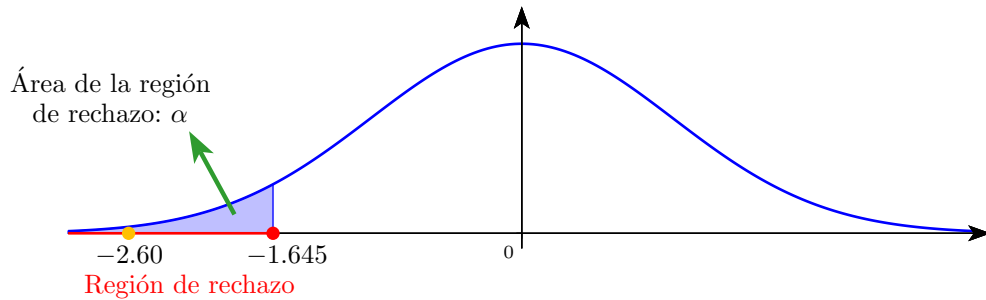
La **región de rechazo** especifica los valores del estadístico de prueba para los cuales la hipótesis nula ha de ser rechazada a favor de la hipótesis alternativa. Se define de modo que la probabilidad de cometer error de tipo I sea igual a α . Este valor α se llama **nivel de significación** del test, y es elegido “a priori” por el investigador (generalmente 0.05). El área de la zona de rechazo es igual al nivel de significación (α) elegido.

Entonces con todos estos elementos, para una muestra particular, si el valor calculado del estadístico de prueba cae en la región de rechazo, rechazamos la hipótesis nula H_0 y aceptamos la hipótesis alternativa H_1 , sabiendo que la probabilidad de equivocarnos (error de tipo I) es α . Si el valor del estadístico de prueba no cae en la región de rechazo, concluimos que a partir de los datos, no tenemos evidencia suficiente para rechazar H_0 ; si nos equivocamos estaríamos cometiendo un error de tipo II, la probabilidad de error de tipo II se suele llamar β y es más difícil de calcular.

Ejemplo 6.1

Para el ejemplo de la introducción, los elementos del test de hipótesis son:

1. Las hipótesis: $H_0 : \mu = 1$ vs. $H_1 : \mu < 1$. Este ejemplo corresponde a un test unilateral pues la alternativa sólo puede ocurrir en una dirección (hacia la izquierda).
2. El estadístico de prueba: $Z = \frac{\sqrt{6}(\bar{X} - 1)}{0.08}$ y cuando H_0 es verdadera, $Z \sim N(0, 1)$.
3. La región de rechazo: $z = \frac{\sqrt{6}(\bar{x} - 1)}{0.08} < -z_{0.05} = -1.645$, gráficamente:



En el gráfico se puede ver que el estadístico de prueba $z = -2.60$ cae dentro de la región de rechazo.

Definición:

Se define el **valor- p** como la probabilidad de que el estadístico de prueba tome un valor “tan extremo” como el obtenido, si fuera cierta la hipótesis nula. O también puede definirse como el menor nivel de significación α , para el cual se puede rechazar la hipótesis nula con los valores observados.

Observación:



Las dos definiciones de anteriores de valor- p son equivalentes. Es importante observar que, el valor- p y el nivel de significación son cosas diferentes: el valor- p depende de los valores observados, mientras que el nivel de significación se elige a priori.

Si el valor- p es menor o igual a un nivel de significancia α asignado previamente, entonces la hipótesis nula puede ser rechazada con dicho α y se puede afirmar que los resultados son estadísticamente significativos a ese nivel.

El Ejemplo 6.1 es un test unilateral, analicemos ahora un ejemplo donde la hipótesis alternativa es bilateral.

Ejemplo 6.2

Una importante propiedad de un método analítico es que no tenga error sistemático. Ya vimos que el resultado de una medición puede modelizarse como una variable aleatoria con distribución normal, cuya media es igual al verdadero valor del analito, cuando el método no tiene error sistemático. Un error sistemático implicaría que esa media difiere del valor verdadero del analito. Para ver si se verifica esta propiedad, se analiza con dicho método una muestra estándar con valor conocido del analito. En este ejemplo, se quiere determinar si un método para medir selenourea en agua tiene error sistemático. Para ello, se analizaron muestras de agua de grifo adicionadas con 50 *ng/ml* de selenourea. Se sabe que la desviación típica del error de medición es 0.9 *ng/ml*. Se obtuvieron los siguientes valores (medidos en *ng/ml*):

48.8 50.2 50.1 48.1 48.7 49.4

En este ejemplo queremos decidir entre estas dos hipótesis:

- el método de medición no tiene error sistemático
- el método de medición tiene error sistemático

Podemos plantear el siguiente modelo probabilístico:

Sea X_1, X_2, \dots, X_n una m.a. donde X_i es el resultado de la i -ésima medición del contenido de selenourea en la muestra de agua y la suponemos con distribución $N(\mu, 0.9^2)$. En este caso $n = 6$. Si el método no tiene error sistemático, la media μ es igual al verdadero contenido de selenourea en esa muestra. Con este modelo las hipótesis se escriben:

$$H_0: \mu = 50 \text{ (el método de medición no tiene error sistemático)}$$

$$H_1: \mu \neq 50 \text{ (el método de medición tiene error sistemático)}$$

Usaremos el estadístico de prueba:

$$Z = \frac{\sqrt{6}(\bar{X} - 50)}{0.9}$$

que, bajo el modelo propuesto y cuando H_0 es verdadera, tiene distribución $N(0, 1)$.

En este caso, parece lógico rechazar H_0 cuando \bar{x} sea mucho mayor o mucho menor que 50. La regla de decisión será: rechazar H_0 cuando el valor absoluto del valor del estadístico de prueba sea mayor que 1.96, de este modo nos aseguramos que:

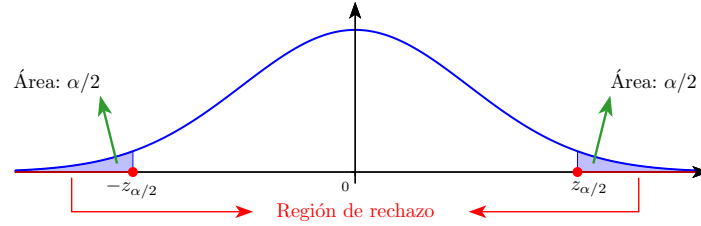
$$P(\text{error de tipo I}) = P\left(\frac{\sqrt{6}|\bar{X} - 50|}{0.9} > 1.96 \mid H_0 \text{ es verdadera}\right) = 2(1 - \Phi(1.96)) = 0.05$$

En general, en un test bilateral, para que la probabilidad de error de tipo I sea α la regla será:

$$\text{rechazar } H_0: \mu = \mu_0 \text{ cuando } \frac{\sqrt{n}|\bar{x} - \mu_0|}{\sigma} > z_{\alpha/2}$$

entonces:

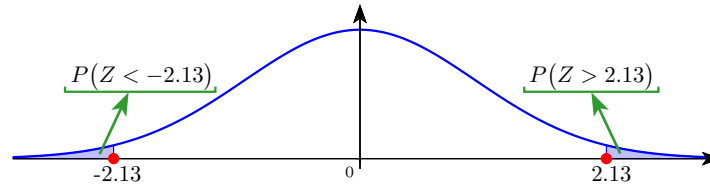
$$P(\text{error de tipo I}) = P\left(\frac{\sqrt{n}|\bar{X} - \mu_0|}{\sigma} > z_{\alpha/2} \mid H_0 \text{ es verdadera}\right) = 2(1 - \Phi(z_{\alpha/2})) = \alpha$$



En nuestro ejemplo, con las 6 mediciones se obtuvo $\bar{x} = 49.217$, reemplazando \bar{X} por $\bar{x} = 49.217$, el estadístico de prueba toma el valor $\frac{\sqrt{6}(49.217-50)}{0.9} = -2.13$. Como este valor cae en la región de rechazo, pues $|-2.13| > 1.96$, podemos rechazar la hipótesis nula con nivel 0.05. Esto significa que podemos afirmar que $\mu \neq 50$, es decir, el método tiene error sistemático y la probabilidad de equivocarnos al hacer esta afirmación es a lo sumo 0.05.

También podemos calcular el valor- p :

$$\text{valor-}p = P\left(\frac{\sqrt{6}|\bar{X} - 50|}{0.9} > |-2.13| \mid H_0 \text{ es verdadera}\right) = P(|Z| > 2.13)$$



$$= P(Z > 2.13) + P(Z < -2.13) = 2(1 - \Phi(2.13)) = 0.0332$$



Observación:

Los test bilaterales son más conservadores que los unilaterales, ya que para un mismo valor del estadístico de prueba, el valor- p es mayor para un test bilateral que para un test unilateral.

Podemos resumir lo que hemos visto sobre tests para la media μ de la siguiente manera:

RESUMEN 6.1

Para la m.a. X_1, X_2, \dots, X_n con distribución normal con σ_0 conocido.			
Hipótesis nula	$H_0 : \mu = \mu_0$		
Valor del estadístico de prueba	$z = \frac{\sqrt{n}(\bar{x} - \mu_0)}{\sigma_0}$		
Hipótesis alternativa	$H_1 : \mu < \mu_0$	$H_1 : \mu > \mu_0$	$H_1 : \mu \neq \mu_0$
Región de rechazo con nivel α	$z < -z_{\alpha}$	$z > z_{\alpha}$	$ z > z_{\alpha/2}$

EJERCICIO 6.1

Se extrajo una m.a. de 16 informes de urgencias de los archivos de un servicio de ambulancias. El tiempo medio para que las ambulancias llegaran a sus destinos fue de 13 min. Suponga que la población de tiempos sigue una distribución normal con varianza 9. ¿Es posible concluir, con un nivel de significancia de 0.05, que la media de la población es menor que 10 min?

Test de hipótesis para la media de una distribución normal con varianza desconocida

Cuando la distribución de los datos es normal, pero desconocemos el valor de σ , no podemos usar el mismo estadístico de prueba que en el caso anterior. Recordemos lo que vimos al construir intervalos de confianza, en ese caso usamos una función pivote con distribución de Student. En el caso de un test, cuando $\mu = \mu_0$, el estadístico:

$$T = \frac{\bar{X} - \mu_0}{S/\sqrt{n}}$$

tiene distribución de Student con $n - 1$ grados de libertad. Usaremos entonces este estadístico de prueba, del mismo modo que antes usamos el Z .

Ejemplo 6.3

El zooplankton se compone de todos los animales oceánicos que se dejan arrastrar pasivamente por el movimiento del agua. Se extrajeron 9 muestras de agua en las proximidades de una isla, se determinó el número de individuos por m^3 y se obtuvieron los siguientes valores:

5000 5700 4450 4500 4825 4025 3700 4900 3750

Suponemos que el número de individuos por m^3 de agua está normalmente distribuido.

El investigador sospecha, por tener experiencia en el tema, que en esa zona la cantidad media de zooplankton por m^3 de agua supera los 4200 individuos y quiere verificar esa teoría. ¿Proveen los datos suficiente evidencia para apoyar esta suposición? Usar $\alpha = 0.05$.

El modelo en este ejemplo es: X_1, X_2, \dots, X_9 una m.a. donde cada X_i es el número de individuos por m^3 de agua de la muestra i , con $i = 1, \dots, 9$. Suponemos que las X_i tienen distribución $N(\mu, \sigma^2)$. Podemos enunciar el problema como:

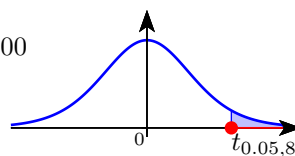
$$H_0 : \mu = 4200 \quad \text{vs.} \quad H_1 : \mu > 4200$$

Como no conocemos la varianza usaremos el siguiente estadístico de prueba:

$$T = \frac{\bar{X} - 4200}{S/\sqrt{9}}$$

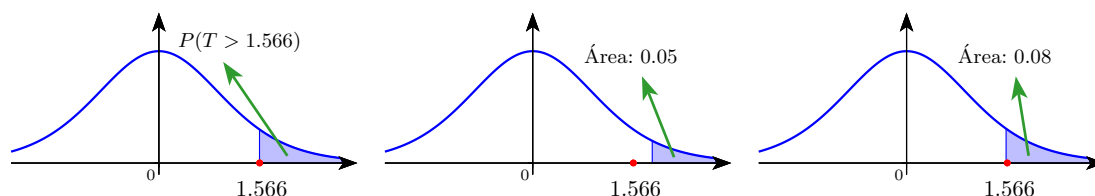
que, bajo el modelo supuesto y cuando la hipótesis nula H_0 es verdadera, tiene distribución de Student con $n - 1 = 8$ grados de libertad. Luego, se puede definir la regla de decisión del siguiente modo:

rechazar $H_0 : \mu = 4200$ a favor de $H_1 : \mu > 4200$
cuando $t > t_{0.05,8}$



Se busca en la Tabla de Student para 8 grados de libertad, resulta $t_{0.05,8} = 1.8595$. Esto significa que la región de rechazo está a la derecha de 1.86. Reemplazando por los valores de la media muestral $\bar{x} = 4538.9$ y la desviación típica muestral $s = 649.29$, calculamos el valor del estadístico y obtenemos $t = 3(4538.9 - 4200)/649.29 = 1.566$. Este valor no cae en la región de rechazo, entonces no podemos afirmar a nivel 0.05 que el número medio de individuos de la población de zooplancton sea mayor que 4200 por m^3 .

Como no tenemos una tabla que nos permita calcular la probabilidad $P(T > 1.566)$, no podemos calcular exactamente el valor- p , pero teniendo en cuenta que es el mínimo nivel de significación con el que rechazamos H_0 con los datos observados, ya podemos afirmar que valor- $p > 0.05$ (ya que con nivel $\alpha = 0.05$ no pudimos rechazar H_0), entonces vemos si es posible rechazar H_0 con un nivel menos exigente. Podemos ver en la Tabla de Student con 8 grados de libertad que el valor crítico correspondiente a $\alpha = 0.08$ es 1.5489 esto significa que con nivel 0.08 podemos rechazar H_0 , entonces el valor- $p < 0.08$. Resumiendo $0.05 < \text{valor-}p < 0.08$, como se muestra en la siguiente gráfica.



Ejemplo 6.4

Volviendo a la situación de determinar si un método de medición tiene error sistemático, consideremos un ejemplo más realista donde no se conoce el desvío estándar. Se hacen 10 determinaciones del contenido de níquel para una aleación estándar preparada de modo que se conoce el verdadero valor del contenido, que es 4.44 %. Se obtienen los siguientes valores:

4.32 4.31 4.50 4.12 4.43 4.36 4.48 4.28 4.18 4.42

La pregunta que nos formulamos es: ¿con estos 10 datos podemos afirmar que el método de medición tiene error sistemático?

Tenemos una m.a. X_1, X_2, \dots, X_{10} , donde X_i es la i -ésima determinación del contenido de

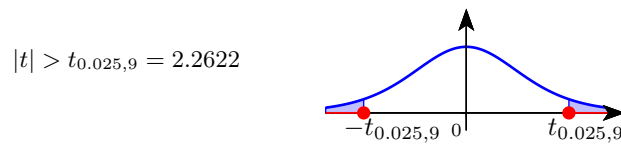
níquel en la aleación, y suponemos que las X_i tiene distribución $N(\mu, \sigma^2)$, si el método no tiene error sistemático, la media es igual al verdadero valor. Las hipótesis a contrastar en este caso son:

$$H_0: \mu = 4.44 \text{ vs. } H_1: \mu \neq 4.44,$$

el estadístico de prueba es:

$$T = \frac{\bar{X} - 4.44}{S/\sqrt{10}}$$

que tiene distribución de Student con 9 grados de libertad cuando H_0 es verdadera. Luego, la zona de rechazo con nivel $\alpha = 0.05$ es:



Haciendo los cálculos se obtiene: $\bar{x} = 4.34$ y $s = 0.1243$ y reemplazando en la fórmula del estadístico de prueba, se obtiene: $t = -2.5441$. Vemos que el valor del estadístico de prueba cae en la región de rechazo, y en consecuencia, podemos afirmar con nivel $\alpha = 0.05$ que el método tiene error sistemático.

Como hemos rechazado H_0 con $\alpha = 0.05$, sabemos que $\text{valor-}p < 0.05$. Podemos ver si aún rechazamos con un nivel más exigente. Si eligiéramos $\alpha = 0.02$, el valor crítico es $t_{0.01,9} = 2.8214$ y por lo tanto el valor observado $t = -2.5441$ no cae en la correspondiente región de rechazo, entonces $\text{valor-}p > 0.02$. Es decir:

$$0.02 < \text{valor-}p < 0.05$$



Podemos resumir los diferentes tests para la media μ de la siguiente manera:

RESUMEN 6.2

Para la m.a. X_1, X_2, \dots, X_n con distribución normal con σ desconocido.			
Hipótesis nula	$H_0: \mu = \mu_0$		
Valor del estadístico de prueba	$t = \frac{\sqrt{n}(\bar{x} - \mu_0)}{s}$		
Hipótesis alternativa	$H_1: \mu < \mu_0$	$H_1: \mu > \mu_0$	$H_1: \mu \neq \mu_0$
Región de rechazo con nivel α	$t < -t_{\alpha, n-1}$	$t > t_{\alpha, n-1}$	$ t > t_{\alpha/2, n-1}$

EJERCICIO 6.2

Los siguientes datos son medias de consumo de oxígeno (en ml) durante la incubación de una m.a. de 14 suspensiones celulares:

14.0 14.1 14.5 13.2 11.2 14.1 12.2 11.1 13.7 13.2 16.0 12.8 14.4 12.9 ¿Proporcionan estos datos suficiente evidencia, en un nivel de significación de 0.05, de que la media del consumo de oxígeno es menor a 14 ml ?, ¿qué supuestos se deben cumplir?

Test de hipótesis para la media de una distribución desconocida (muestras grandes)

Del mismo modo que en la construcción de un intervalo de confianza, cuando no conocemos la distribución de los datos, si la muestra es suficientemente grande, podemos utilizar el Teorema Central del Límite.

Ejemplo 6.5

Recordemos el Ejemplo 5.14. ¿Se puede afirmar, en base a estos datos, que los peces de esa región tienen niveles medios de zinc mayores a 8.2 $\mu g/g$?

Aquí tenemos una m. a. X_1, X_2, \dots, X_{56} donde cada X_i es la concentración de zinc en el hígado ($\mu g/g$) del i -ésimo pez examinado y desconocemos su distribución, pero suponemos que tiene media $\mu = E(X_i)$ y varianza $\sigma^2 = V(X_i)$.

El problema puede plantearse como:

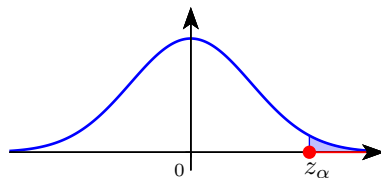
$$H_0 : \mu = 8.2 \quad \text{vs.} \quad H_1 : \mu > 8.2$$

y en este caso, como n es grande, podemos aplicar el resultado del Teorema Central del Límite y usar el estadístico de prueba:

$$Z = \sqrt{56} \frac{(\bar{X} - 8.2)}{S}$$

ya que, según este teorema, cuando $\mu = 8.2$ tiene una distribución aproximadamente $N(0, 1)$. Entonces podemos definir, como siempre, una regla de decisión:

rechazar $H_0 : \mu = 8.2$ a favor de $H_1 : \mu > 8.2$ cuando $z > z$



con los datos del ejemplo, $\bar{x} = 9.15 \mu g/g$ y $s = 1.27 \mu g/g$, reemplazando en el estadístico, obtenemos un valor $z = \sqrt{56} (9.15 - 8.2)/1.27 = 5.598$, vemos en la Tabla de la distribución Normal que el valor- $p = P(Z > 5.598) < 0.0001$, esto significa que hay muy fuerte evidencia para rechazar H_0 , y se puede rechazar con cualquier nivel de significación razonable. O sea, podemos concluir que la concentración media de zinc en el hígado de los peces de esa región es superior a $8.2 \mu g/g$.



Podemos resumir los casos de tests para la media de una distribución desconocida, cuando n es grande, de la siguiente manera:

RESUMEN 6.3

Para la m.a. X_1, X_2, \dots, X_n con distribución desconocida y n grande.			
Hipótesis nula	$H_0 : \mu = \mu_0$		
Valor del estadístico de prueba	$z = \frac{\sqrt{n}(\bar{x} - \mu_0)}{s}$		
Hipótesis alternativa	$H_1 : \mu < \mu_0$	$H_1 : \mu > \mu_0$	$H_1 : \mu \neq \mu_0$
Región de rechazo con nivel aproximado α	$z < -z_\alpha$	$z > z_\alpha$	$ z > z_{\alpha/2}$



Observación:

En este caso el nivel es aproximado, porque no conocemos la distribución exacta del estadístico de prueba y estamos utilizando una aproximación.

EJERCICIO 6.3

En una muestra de 49 adolescentes que se prestaron como sujetos para un estudio inmunológico, una variable de interés fue la prueba del diámetro de reacción de la piel a un antígeno. La media y la desviación estándar de la muestra fueron de 21 y 11 mm, respectivamente. ¿Es posible concluir a partir de estos datos que la media poblacional es diferente a 30? Tomar $\alpha = 0.01$.

Test de hipótesis para una proporción

Ya vimos al tratar el tema intervalo de confianza para una proporción, en el Capítulo 5, que si tenemos una v.a. X con distribución $B(n, p)$ entonces, si n es suficientemente grande, la distribución de:

$$Z = \frac{\hat{p} - p}{\sqrt{\frac{p(1-p)}{n}}}$$

es aproximadamente $N(0, 1)$ por el Teorema Central del Límite.

Ejemplo 6.6

Una de las célebres leyes de Murphy establece que si se deja caer al suelo una tostada untada con dulce, la probabilidad de que caiga del lado del dulce es mayor que la de que caiga del lado del pan. Para verificarla, se realizó un experimento: se dejaron caer 1000 tostadas untadas con mermelada, de las cuales 540 cayeron del lado del dulce. ¿Qué podría concluir a nivel $\alpha = 0.05$?

Para modelizar este problema definimos $X =$ “el número de tostadas que caen del lado del dulce, entre las 1000 del experimento”. Esta v.a. tiene distribución $B(1000, p)$, según la ley de Murphy este valor de p es mayor que 0.5. Entonces, para verificar esta ley planteamos el siguiente test:

$$H_0 : p = 0.5 \quad \text{vs.} \quad H_1 : p > 0.5$$

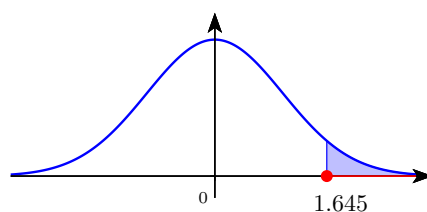
Si H_0 es verdadera ($p = 0.5$), según el TCL,

$$Z = \frac{\hat{p} - 0.5}{\sqrt{\frac{0.5(1-0.5)}{1000}}}$$

tiene distribución aproximadamente $N(0, 1)$, este será nuestro estadístico de prueba.

Si usamos un nivel $\alpha = 0.05$, la regla de decisión será:

rechazar $H_0 : p = 0.5$ a favor de $H_1 : p > 0.5$ cuando $z > z_{0.05} = 1.645$



En el experimento se obtuvo $\hat{p} = 0.54$, reemplazando en el estadístico de prueba se obtiene $z = 2.53$, que cae en la zona de rechazo. Conclusión, en base a este experimento, podemos afirmar (con nivel 0.05) que la ley de Murphy es verdadera.

Si queremos conocer el valor- p , calculamos $P(Z > 2.53) \cong 1 - \Phi(2.53) = 0.0057$. Esto significa que podemos afirmar que esta ley de Murphy es verdadera con cualquier nivel ≥ 0.0057 .

Podemos resumir los casos de tests para una proporción como sigue:

RESUMEN 6.4

Para la v.a. X con distribución $B(n, p)$ y n grande.			
Hipótesis nula	$H_0 : p = p_0$		
Valor del estadístico de prueba	$z = \frac{\hat{p} - p_0}{\sqrt{\frac{p_0(1-p_0)}{n}}}$		
Hipótesis alternativa	$H_1 : p < p_0$	$H_1 : p > p_0$	$H_1 : p \neq p_0$
Región de rechazo con nivel aproximado α	$z < -z_\alpha$	$z > z_\alpha$	$ z > z_{\alpha/2}$



Observación:

En este caso también el nivel es aproximado, porque no conocemos la distribución exacta del estadístico de prueba y estamos utilizando una aproximación.

EJERCICIO 6.4

1. En un estudio se encontró que 66 % de los niños en una muestra de 670 completaron toda la serie de vacunas contra la hepatitis B. ¿Es posible concluir que, con base a estos datos, en la población muestreada, más del 60 % tienen la serie completa de vacunas contra la hepatitis B? Tomar $\alpha = 0.01$.
2. En una investigación de consumidores de drogas intravenosas en una ciudad grande, encontraron a 18 de 423 individuos con HIV positivo. Se pretende saber si es posible concluir que los consumidores de drogas intravenosas en la población muestreada tienen HIV positivo en un porcentaje diferente del 5 %.