

Problemas. Intervalos de Confianza y Contrastes de Hipótesis

Ejemplos resueltos y propuestos

Intervalos de Confianza

Variable Normal en la población

Se selecciona una muestra de tamaño n de una población

$$X \rightarrow N(\mu, \sigma)$$

1. Intervalo de confianza para la media μ con σ conocida

$$\bar{X} \pm z_{\alpha/2} \frac{\sigma}{\sqrt{n}}$$

2. Intervalo de confianza para la media μ con σ desconocida

$$\bar{X} \pm t_{\alpha/2} \frac{s}{\sqrt{n}}$$

3. Intervalo de confianza para la varianza σ^2

$$\left(\frac{(n-1)s^2}{\chi_{sup}^2}, \frac{(n-1)s^2}{\chi_{inf}^2} \right)$$

Variable Bernoulli en la población

Se selecciona una muestra de tamaño n de una población

$$X \rightarrow B(1, p)$$

1. Intervalo de confianza para la proporción

$$u_p \pm z_{\alpha/2} \sqrt{\frac{u_p(1-u_p)}{n}}$$

Ejemplos Resueltos de Intervalos de Confianza

Ejemplo 1. Intervalo de confianza para la media con σ conocida

Se ha obtenido una muestra de 25 alumnos de una Facultad para estimar la calificación media de los expedientes de los alumnos en la Facultad. Se sabe por otros cursos que la desviación típica de las puntuaciones en dicha Facultad es de 2.01 puntos.

La media de la muestra fue de 4.9.

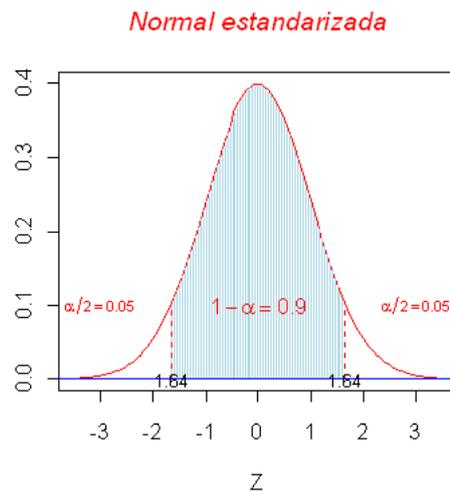
1. Intervalo de confianza al 90 %.
2. Intervalo de confianza al 99 %.

Solución ejemplo 1.

1. Intervalo de confianza al 90 %. Usamos la fórmula:

$$\bar{X} \pm z_{\alpha/2} \frac{\sigma}{\sqrt{n}}$$

Los cuantiles de orden 0.05 y 0.95, que encierran en el centro de la distribución normal un área igual a 0.9 se muestran en el gráfico siguiente:



Por último, sustituyendo los datos en la fórmula del intervalo, tenemos:

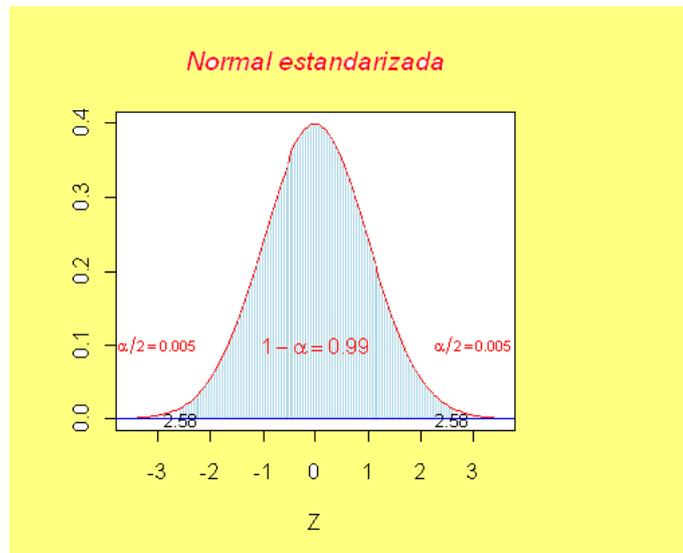
$$4,9 \pm 1,64 \frac{2,01}{\sqrt{25}} \equiv 4,9 \pm 0,66$$

(4,24, 5,56)

2. Intervalo de confianza al 99 %.

$$\bar{X} \pm z_{\alpha/2} \frac{\sigma}{\sqrt{n}}$$

De modo similar obtenemos los cuantiles de orden 0.005 y 0.995 que describen en el modelo normal una confianza del 99 %



Por último, sustituyendo los datos en la fórmula del intervalo, tenemos:

$$4,9 \pm 2,58 \frac{2,01}{\sqrt{25}} = 4,9 \pm 1,04$$

(3,86, 5,94)

Ejemplo 2. Intervalo de confianza para la media con σ desconocida

Se ha obtenido una muestra de 15 vendedores de una Editorial para estimar el valor medio de las ventas por trabajador en la Empresa.

La media y varianza de la muestra (en miles de euros) son 5 y 2, respectivamente.

1. Intervalo de confianza para la venta media por trabajador en la Editorial al 90 %.
2. Intervalo de confianza para la varianza de las ventas por trabajador en la Editorial al 90 %.

Solución ejemplo 2.

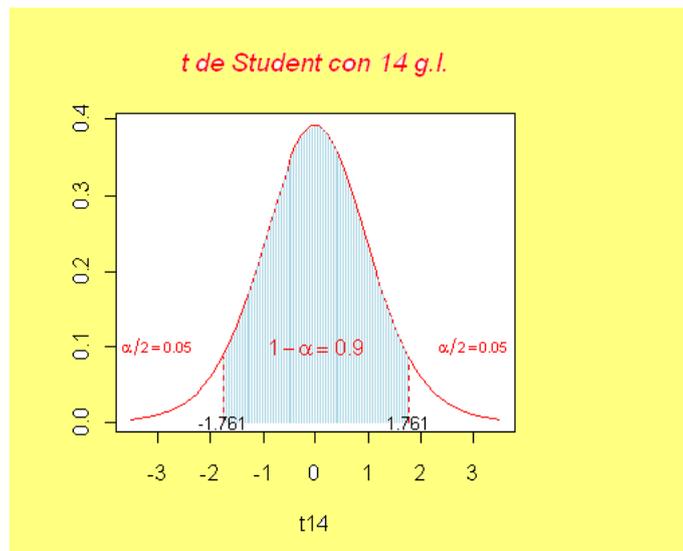
1. Intervalo de confianza para la media de ventas por vendedor al 90%.
Usamos la fórmula:

$$\bar{X} \pm t_{\alpha/2} \frac{s}{\sqrt{n}}$$

Se conoce la varianza de la muestra $V(X) = 2$, pero necesitamos la cuasi-varianza s^2

$$s^2 = \frac{n}{n-1} V(x) = \frac{15}{14} 2 = 2,143$$

Y, por tanto, una cuasi-desviación típica igual a $s = \sqrt{s^2} = 1,464$. Los cuantiles de orden 0.05 y 0.95, que encierran en el centro de la distribución t de Student con 14 g.l. un área igual a 0.9 se muestran en el gráfico siguiente:



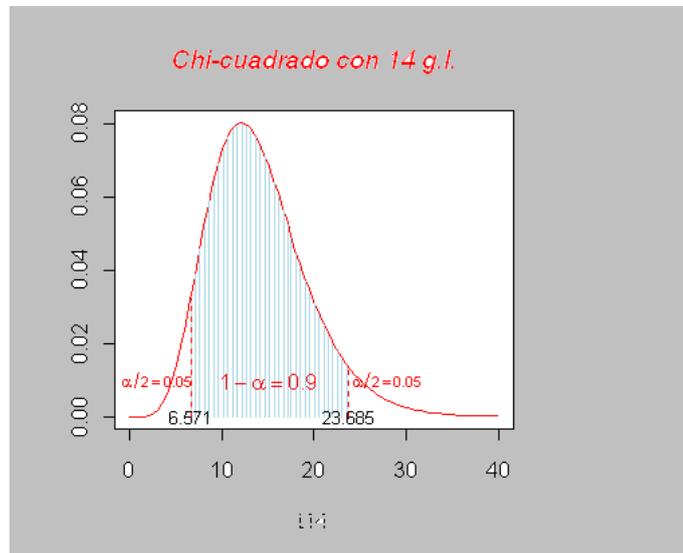
Por último, sustituyendo los datos en la fórmula del intervalo, tenemos:

$$5 \pm 1,761 \frac{1,464}{\sqrt{15}} \equiv 5 \pm 0,666$$
$$(4,334, 5,666)$$

2. Intervalo de confianza para la varianza al 90%.

$$\left(\frac{(n-1)s^2}{\chi_{sup}^2}, \frac{(n-1)s^2}{\chi_{inf}^2} \right)$$

De modo similar obtenemos los cuantiles de orden 0.05 y 0.95 que describen en el modelo chi-cuadrado con 14 g.l. una confianza del 90%



Por último, sustituyendo los datos en la fórmula del intervalo, tenemos:

$$\left(\frac{(15 - 1)2,143}{23,685}, \frac{(15 - 1)2,143}{6,571} \right)$$

$$(1,27, 4,58)$$

Ejemplo 3. Intervalo de confianza para la proporción

Se ha obtenido una muestra al azar de 150 vendedores de una Editorial para estimar la proporción de vendedores en la Editorial que no alcanza un límite de ventas mínimo establecido por la dirección.

De entre los seleccionados, 50 no han conseguido llegar al límite de ventas mínimo establecido.

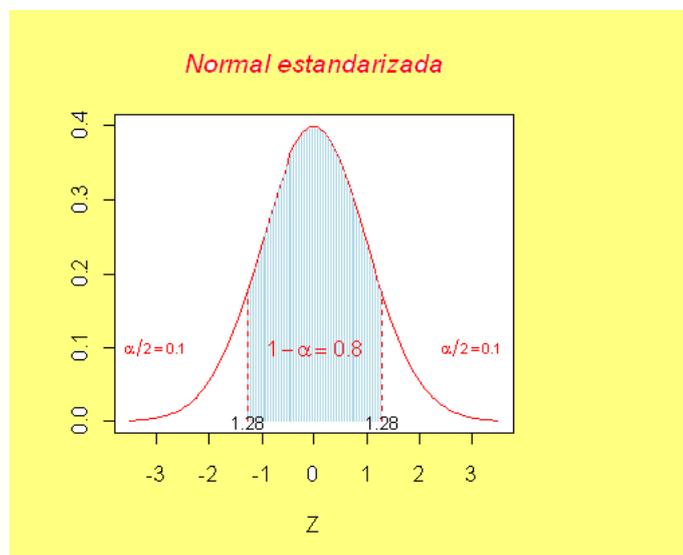
1. Intervalo de confianza para la proporción de trabajadores en la Editorial que no alcanza el límite al 80 %.
2. Intervalo de confianza para la proporción de trabajadores en la Editorial que no alcanza el límite al 99 %.
3. Interprete los intervalos obtenidos.

Solución ejemplo 3.

1. Intervalo de confianza para la proporción al 80 %. Usamos la fórmula:

$$u_p \pm z_{\alpha/2} \sqrt{\frac{u_p(1 - u_p)}{n}}$$

La proporción de la muestra es $u_p = \frac{50}{150} = 0,333$. Los cuantiles de orden 0.1 y 0.9 para el nivel de confianza dado son -1.28 y 1.28, respectivamente.



Sustituyendo en la expresión del intervalo:

$$0,333 \pm 1,28 \sqrt{\frac{0,333(1 - 0,333)}{150}}$$

(0,28, 0,38)

Ejemplos propuestos de intervalos de Confianza

Ejemplo 1. Estimación de intervalo de confianza para la media

Una muestra de 26 personas seleccionadas al azar de una población de un barrio, tiene una media salarial de 1800 euros y una varianza de

10.000 euros. Estime la media salarial en el barrio a un nivel confianza de 90

Ejemplo 2. Estimación de intervalo de confianza para la varianza

Con los datos del ejemplo anterior estime la varianza salarial en el barrio a un nivel de confianza del 80

Ejemplo 3. Estimación de intervalo de confianza para la proporción

En una muestra al azar de 120 empresas inspeccionadas de entre las visitadas un año determinado por los inspectores de trabajo de una provincia se ha sancionado a 90 de ellas. Estime a un nivel de confianza del 80 % la proporción de empresas que sanciona la Inspección de Trabajo.

Ejemplo 4. Estimación de intervalo de confianza para la media y para la varianza

Se desea estimar la demanda diaria de un producto que registra una empresa. Para ello se seleccionan 10 días al azar con los siguientes valores en miles

35, 44, 38, 55, 33, 56, 67, 45, 48, 40

- Obtenga el intervalo de confianza para la demanda media diaria a un nivel de confianza del 90 %.
- Obtenga el intervalo de confianza para la varianza un nivel de confianza del 90 %

Contrastes de Hipótesis

Ejemplos resueltos de contrastes de hipótesis para la media y la proporción

Dada una población sobre la que se observa una variable X , tal que

$$X \rightarrow N(\mu, \sigma)$$

Se desea contrastar a un nivel de significación α la hipótesis nula:

$$H_0 : \mu = \mu_0$$

frente a la alternativa

$$H_1 : \mu \neq \mu_0$$

Con los datos observados se determina la media de la muestra, \bar{X} y el estadístico de contraste Z . Se sabe que el estadístico media muestral sigue un modelo normal de media $E(\bar{X}) = \mu$ y desviación típica $\sigma_{\bar{X}} = \frac{\sigma}{\sqrt{n}}$

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

por tanto, el estadístico de contraste, z , definido como

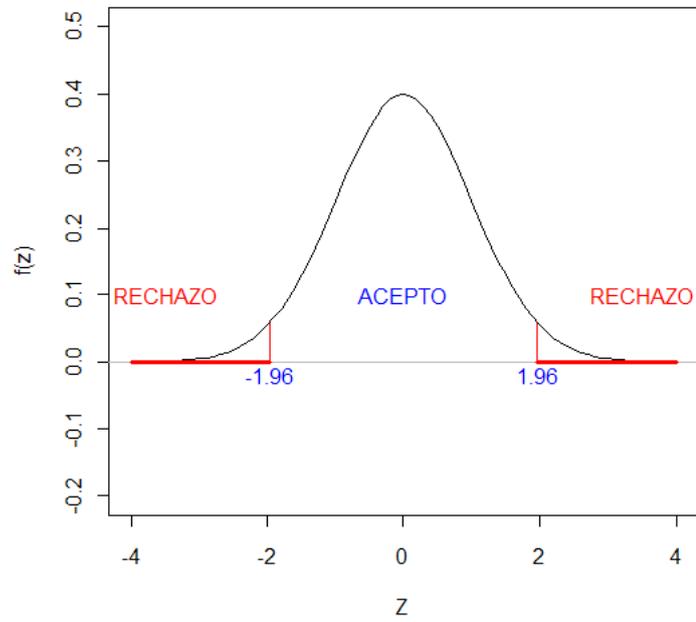
$$z = \frac{\bar{X} - \mu}{\frac{\sigma}{\sqrt{n}}}$$

que sigue un modelo $N(0, 1)$.

En general, se desconoce la desviación típica, σ , de la población, por lo que suele usarse un estimador usando los datos de la muestra. La cuasi-varianza de la muestra, s^2 , es un estimador insesgado de la varianza, σ^2 , de la población. Para tamaños de muestra $n \geq 30$ se puede tomar la cuasi-desviación típica de la muestra, $s = \sqrt{s^2}$, en sustitución del parámetro σ del estadístico de contraste z , que se aproxima a una $N(0, 1)$

$$z = \frac{\bar{X} - \mu}{\frac{s}{\sqrt{n}}} \hookrightarrow N(0, 1)$$

Contraste de hipótesis con alfa 5%



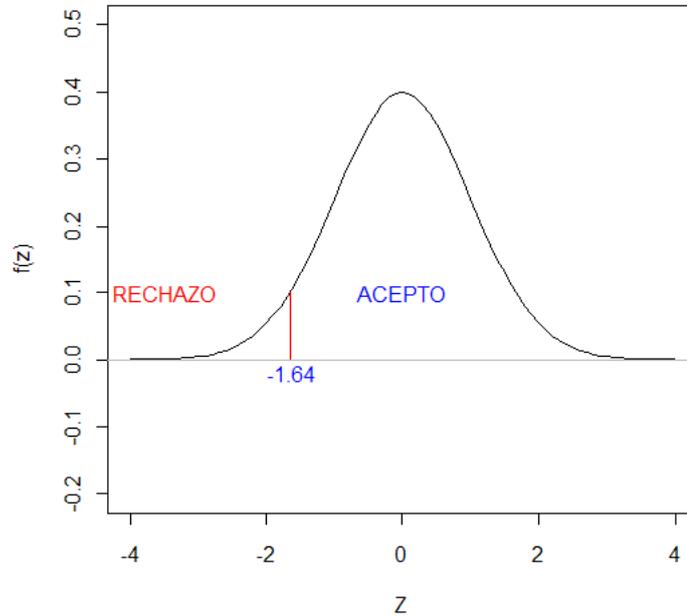
Otra forma de plantear el contraste es estableciendo únicamente una cola para la región de rechazo. Como por ejemplo:

$$H_0 : \mu \geq \mu_0$$

frente a la alternativa

$$H_1 : \mu < \mu_0$$

Contraste de hipótesis con alfa 5%



Nota: Recuerde que

$$\bar{X} = \frac{\sum x_i n_i}{N}$$
$$s^2 = \frac{\sum (x_i - \bar{X})^2 n_i}{N - 1}$$

Ejemplo 1. Contraste de hipótesis para la media

Un auditor desea contrastar a un nivel de significación de 0.05 la hipótesis nula de que la media de las deudas por cobrar de una empresa es de 150.000 euros. Para ello se selecciona una muestra al azar de 50 de dichas deudas con un valor medio y cuasi-desviación típica muestrales 189.000 y de 80.000, respectivamente. Compruebe si se rechaza o acepta dicha hipótesis

$$H_0 : \mu = 150,000$$

$$H_1 : \mu \neq 150,000$$

Solución ejemplo 1. Contraste de hipótesis para la media

$$\alpha = 0,05$$

Para $\alpha = 0,05$ los límites de confianza en la normal estandarizada son $\pm 1,96$

Se determina el estadístico de contraste Z con el estadístico muestral transformándolo en un valor Z , usando media de la muestra y desviación típica muestral), y se observa si cae en la zona de rechazo

El estadístico muestral Z

$$z = \frac{\bar{X} - \mu}{\frac{s}{\sqrt{n}}}$$

sigue un modelo $N(0, 1)$, dado que el tamaño de muestra es mayor que 30.

Estadístico de contraste z es igual a

$$z = \frac{189000 - 150000}{\frac{80000}{\sqrt{50}}} = 3,447$$

$Z = 3,447$ cae en zona de rechazo, fuera de los límites $(-1,96, 1,96)$, por lo que se rechaza H_0 .

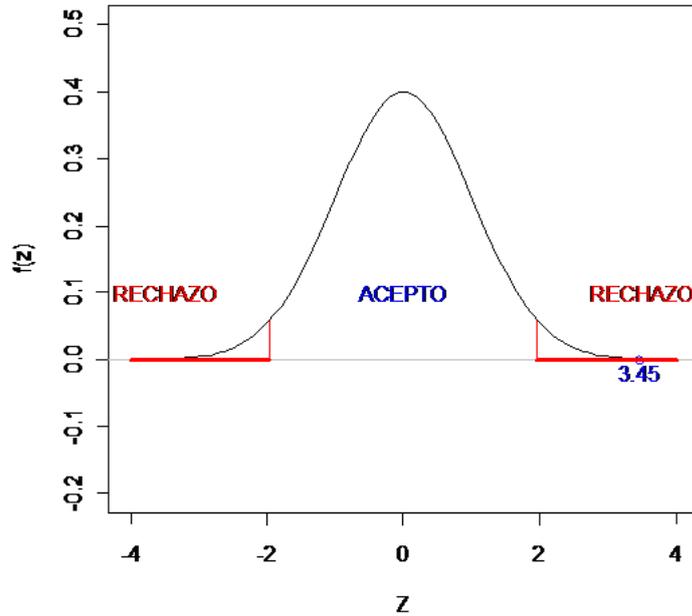
Otra forma equivalente de decisión es viendo si el p - valor es inferior al nivel de significación 0,05

$$p - valor < \alpha \implies RECHAZO$$

$$p - valor \geq \alpha \implies ACEPTO$$

$$p - valor = P(|Z| > 3,45) = 0,00056 < 0,05$$

Por tanto, se rechaza H_0 .



Ejemplo 2. Contraste de hipótesis para la media

Un fabricante afirma que la vida media de un componente electrónico supera las 1500 horas. Se selecciona una muestra de 900 componentes de la producción para establecer un control de calidad, obteniéndose los siguientes resultados. Media y cuasi-varianza muestrales iguales a $\bar{X} = 1450$ y $s^2 = 650$, respectivamente.

Contraste la hipótesis

$$H_0 : \mu > 1500$$

frente a la alternativa

$$H_1 : \mu < 1500$$

a un nivel de significación α del 5%.

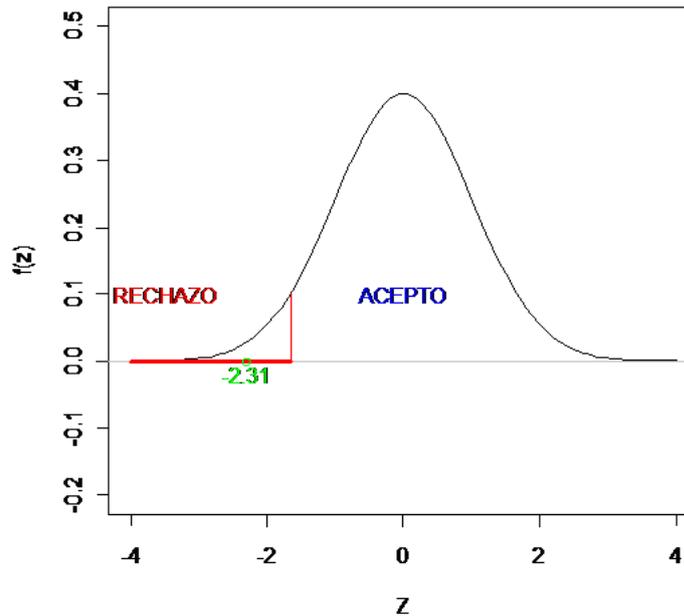
Solución ejemplo 2. Contraste de hipótesis para la media

El estadístico muestral $Z = \frac{\bar{X} - \mu}{\frac{s}{\sqrt{n}}}$ sigue un modelo $N(0, 1)$, dado que el tamaño de muestra es mayor que 30.

Estadístico de contraste $z = (1450 - 1500) / \sqrt{650/900}$

$Z = -2,31$ cae en zona de rechazo (fuera de los límites $(-1,645, \infty)$)

La región de rechazo ahora es



Otra forma de decidir es viendo si el p – *valor* es inferior al nivel de significación 0,05

$$p\text{-valor} = P(Z < -1,645) = 0,0105 < 0,05$$

Por tanto se rechaza H_0 .

Ejemplo 3. Contraste de hipótesis para la media

Se piensa que el tiempo medio que está en paro un tipo de profesional de un determinado sector es de 13,5 meses. Para contrastar esta hipótesis al nivel $\alpha = 5\%$ frente a la alternativa (que no es cierto) se tomó una muestra de 45 profesionales que estuvieron en paro en ese sector y se obtuvo una media de 17,2 meses y una cuasi-desviación típica de 15,3 meses.

Compruebe si se rechaza o acepta dicha hipótesis.

Solución ejemplo 3. Contraste de hipótesis para la media

$$H_0 : \mu = 13,5$$

$$H_1 : \mu \neq 13,5$$

$\alpha = 0,05$. La hipótesis alternativa define dos zonas de rechazo.

Para $\alpha/2 = 0,05/2 = 0,025$ los límites de confianza en la normal estandarizada son $\pm 1,96$.

Se determina el estadístico de contraste Z con los datos de la muestra y se observa si cae o no en la zona de rechazo.

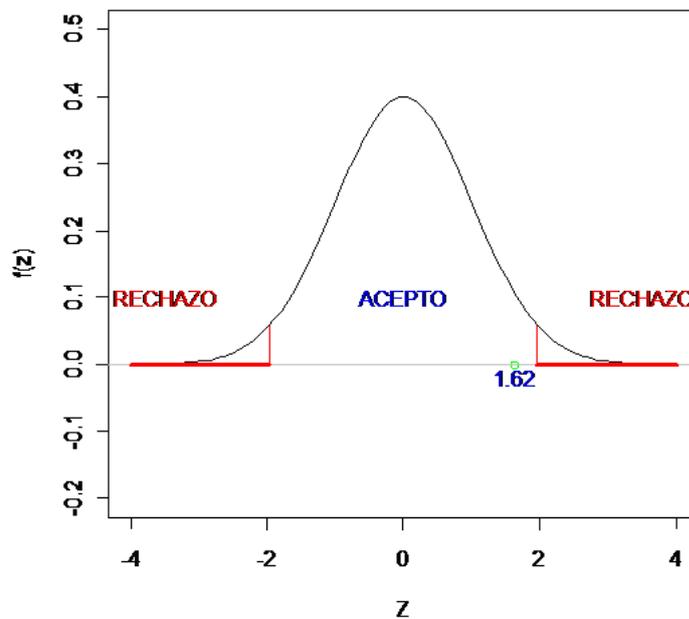
El estadístico muestral

$$Z = \frac{\bar{X} - \mu}{\frac{s}{\sqrt{n}}}$$

sigue un modelo $N(0, 1)$, dado que el tamaño de muestra (45) es mayor que 30.

Estadístico de contraste $z = (17,2 - 13,5)/(15,3/\sqrt{45})$

$Z = 1,62$ cae en zona de aceptación, dentro de los límites $(-1,96, 1,96)$, por tanto no se puede rechazar H_0 .



Otra forma: viendo si el p -valor es o no inferior al nivel de significación del 5

$$p\text{-valor} = P(|Z| > 1,62) = P(Z < -1,62) + P(Z > 1,62) = 2 * P(Z < -1,62) = 0,1048$$

$$p\text{-valor} = 0,1048 > 0,05$$

por tanto, se acepta la hipótesis nula.

Ejemplo 4. Contraste de hipótesis para la proporción

En una muestra de 250 personas en edad laboral de una zona determinada se encuentra que el 14 % esta en paro. Los datos recogidos del INE para el año anterior fueron de una tasa de paro del 10 %. ¿Puede asumirse la hipótesis de estabilidad de la tasa de paro?. Contraste la hipótesis de mantenimiento de la tasa en el 10 %, dada por la Administración a un nivel α del 5 %

Solución ejemplo 4. Contraste de hipótesis para la proporción

$$H_0 := 0,1$$

$$H_1 := \neq 0,1$$

$\alpha = 0,05$. La hipótesis alternativa define dos zonas de rechazo.

Para $\alpha/2 = 0,05/2 = 0,025$ los límites de confianza en la normal estandarizada son $\pm 1,96$.

Se determina el estadístico de contraste Z con los datos de la muestra y se observa si cae o no en la zona de rechazo.

El estadístico muestral

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

sigue un modelo $N(0, 1)$, dado que el tamaño de muestra ($n=250$) es grande.

Estadístico de contraste $z = (0,14 - 0,1)/(0,10,9/\sqrt{250})$

$Z = 7,027$ cae en zona de rechazo (fuera de los límites $(-1,96, 1,96)$), por tanto se rechaza H_0 .

Otra forma: viendo si el p -valor es o no inferior al nivel de significación del 5 %

$$p\text{-valor} = P(|Z| > 7,03) = P(Z < -7,03) + P(Z > 7,03) = 2 * P(Z < -7,03) = 0$$

$$p\text{-valor} = 0 < 0,05$$

por tanto, se RECHAZA la hipótesis nula. La tasa de paro no es la misma del año anterior.

Ejemplos propuestos

Ejemplo 1. Contraste de hipótesis para media

Los resultados de una muestra seleccionada al azar de las calificaciones obtenidas por los 120 estudiantes de una academia de opositores en un examen ha sido media=35 puntos varianza=25 puntos

El director de la academia establece que la media alcanzada por sus alumnos supera lo 40 puntos. Contraste la hipótesis a un nivel de significación del 1%.

Ejemplo 2. Contraste de hipótesis para la proporción

Un partido político realiza un sondeo para conocer la intención de voto. En una muestra de 300 votantes se encuentra que solo 105 son favorables al partido. Contraste la hipótesis a un nivel de significación del 10% de que al menos el 40% de la población lo votará.