

INFERENCIA ESTADÍSTICA CON GEOGEBRA 4

José Luis Álvarez (IES N°5 de Avilés)

1. INTERVALOS CARACTERÍSTICOS

PROBLEMA 1

Sabemos que las merluzas de un caladero tienen un peso medio de 2450 gramos y una desviación típica de 410 gramos. Considerando las posibles muestras de tamaño 100, calcula el intervalo característico de la media muestral para una probabilidad del 95%.

Seleccionamos la herramienta **Cálculo de Probabilidades** y se nos abre una ventana emergente. Elegimos la pestaña **Estadísticas**. Seleccionamos ahora, en la lista desplegable, la opción **Estimación de Z de una media**. Y no queda más que introducir los datos y obtener los resultados.

Cálculos de Probabilidad

Distribución **Estadísticas**

Estimación de Z de una Media

Nivel de Confianza 0.95

Muestra

Media 2450

σ 410

N 100

Resultado

Estimación de Z de una Media

Media	2450
σ	410
ES	41
N	100
Límite Inferior	2369.6415
Límite Superior	2530.3585
Intervalo	2450 \pm 80.3585

Resultados

PROBLEMA 2

Tras repetidas experiencias se admite que el tratamiento con un fármaco da un 80% de curaciones. Calcula e interpreta el intervalo característico de la proporción muestral para una probabilidad del 99%, en muestras de tamaño 200.

Si hacemos los cálculos, el intervalo sería:

$$\left(0,8 - 2,58 \sqrt{\frac{0,8 \cdot 0,2}{52}} \leq pr \leq 0,8 + 2,58 \sqrt{\frac{0,8 \cdot 0,2}{52}} \right)$$

En este caso, en la lista desplegable de la pestaña Estadísticas elegimos la opción **Z estimada de una proporción**. Hay que tener cuidado al introducir los datos, dado que no nos pide la proporción, en este caso 0,8, sino el número de éxitos y el valor de N. Por tanto, en nuestro caso, el número de éxitos sería $0,8 \cdot 200 = 160$. De modo que los datos y resultados que obtendríamos serían los que veis en la siguiente ventana:

Datos

Nivel de Confianza: 0.99

Muestra

Éxitos: 160

N: 200

Resultado

Z Estimada de una Proporción

Éxitos	160
N	200
ES	0.0283
Límite Inferior	0.7271
Límite Superior	0.8729
Intervalo	0.8 ± 0.0729

Resultados

2. CONTRASTE DE HIPÓTESIS

En los libros de texto de bachillerato, este tipo de problemas se suelen resolver hallando el intervalo de aceptación y valorando, a continuación, si se acepta la hipótesis nula o la alternativa en función de que el valor que se compara esté en la zona de aceptación o en la de rechazo. Sin embargo esos problemas pueden resolverse de otras formas. Por ejemplo, tipificando el valor muestral y comparándolo con el intervalo de aceptación expresado en valores de la distribución normal $N(0,1)$ o calculando el valor exacto del grado de significación correspondiente al resultado muestral. Pues bien, GeoGebra lo que hace son precisamente estas dos cosas.

Veamos un par de ejemplos, uno bilateral con una proporción y otro unilateral con una media.

PROBLEMA 1

Un dentista afirma que el 40% de los niños de 10 años muestran indicios de caries dental. Tomada una muestra de 100 niños se observó que 32 tenían estos indicios. Se quiere saber si este resultado permite rechazar la afirmación del dentista con un nivel de significación del 5%.

Primero veamos cómo haríamos el problema utilizando los procedimientos que antes citábamos, para luego comparar con lo que haremos con GeoGebra.

En este caso las hipótesis nula y alternativa serían:

$$H_0: p = p_0 = 0,4$$

$$H_1: p \neq p_0 \neq 0,4$$

La distribución normal de referencia sería:

$$pr \rightarrow N\left(p_0, \sqrt{\frac{p_0 \cdot q_0}{n}}\right) = N\left(0,4; \sqrt{\frac{0,4 \cdot 0,6}{100}}\right) = N(0,4; 0,049)$$

Si calculamos el intervalo de aceptación, del modo habitual:

$$\left(p_0 - z_{0,025} \cdot \sqrt{\frac{p_0 \cdot q_0}{n}}, p_0 + z_{0,025} \cdot \sqrt{\frac{p_0 \cdot q_0}{n}}\right) = (0,4 - 1,96 \cdot 0,049; 0,4 + 1,96 \cdot 0,049) = (0,304; 0,496)$$

Así que obtenemos el intervalo: (0,304; 0,496). Por tanto, la proporción muestral, 0,32, está dentro del intervalo de aceptación y aceptamos la hipótesis nula.

Ahora bien, también podríamos haber calculado el nivel de significación que corresponde a la proporción muestral 0,32:

$$z_m = \frac{pr - p_0}{\sqrt{\frac{p_0 \cdot q_0}{n}}} = \frac{0,32 - 0,4}{\sqrt{\frac{0,4 \cdot 0,6}{100}}} = \frac{-0,08}{0,049} \approx -1,63$$

Ese valor está en el intervalo de aceptación que, en la normal (0,1) para un nivel de significación del 5% sería: (-1,96, 1,96), de modo que aceptaríamos la hipótesis nula.

Y también podríamos calcular el valor exacto del nivel de significación que corresponde a -1,63:

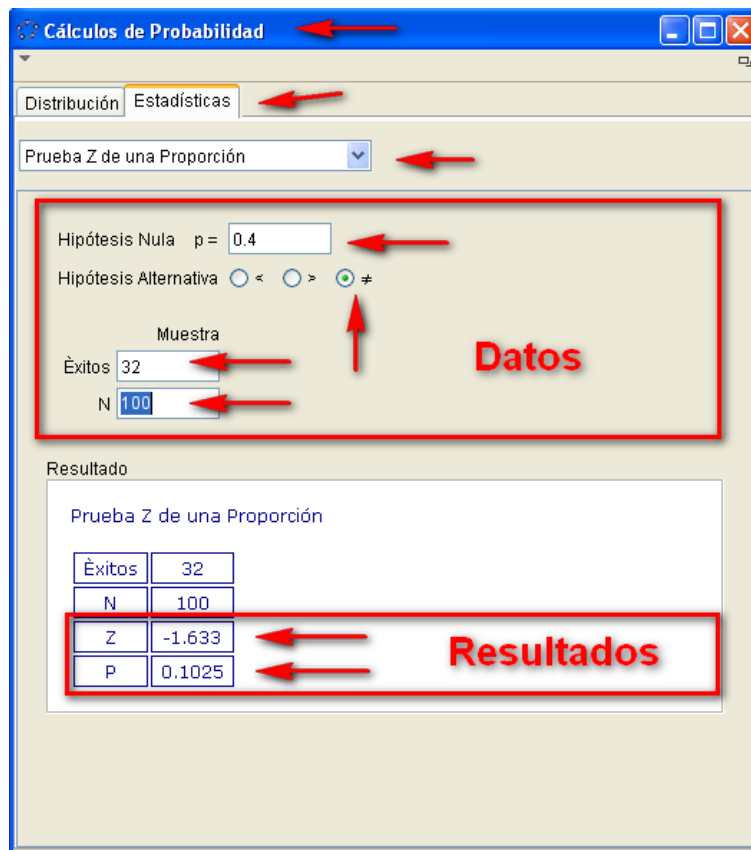
$$z_m = -1,63; 1 - \frac{\alpha}{2} = 0,9484 \text{ (en tablas). } \alpha \approx 0,1$$

El nivel de significación de la proporción muestral 0,1 es mayor de 0,05: el resultado no es significativo y, en consecuencia, no se puede rechazar H_0 .

Pues bien: **estos dos últimos valores, -1,63 y 0,1, son los que nos proporciona GeoGebra.**

¿Cómo resolveríamos el problema con GeoGebra? Elegimos la herramienta **Cálculo de probabilidades** y seleccionamos la pestaña de **Estadísticas**.

Ahora elegimos, en la lista desplegable, **Prueba Z de una proporción**. Tenemos que introducir los datos e, inmediatamente, obtendremos los resultados. Os copio la imagen:



Obviamente, la interpretación es nuestra porque, como se puede apreciar, en ningún momento nos pregunta por el nivel de significación del contraste. Lo que está claro es que con los resultados que nos proporciona GG el problema está resuelto, como hemos visto anteriormente.

PROBLEMA 2

La duración de las bombillas de 100 w de una empresa sigue una distribución normal con una desviación típica 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 comprobadas, se obtiene una vida media de 750 horas. Con un nivel de significación de 0,01, ¿habría que rechazar el lote por no cumplir la garantía?

Ahora tenemos un test para una media, en este caso unilateral. Las hipótesis nula y alternativa serían:

$$H_0: \mu \geq \mu_0 = 800$$

$$H_1: \mu < \mu_0 = 800$$

La distribución normal de referencia:

$$\bar{x} \rightarrow N\left(\mu_0, \frac{\sigma_0}{\sqrt{n}}\right) = N\left(800, \frac{120}{\sqrt{50}}\right) = N(800; 17,0)$$

Si calculamos el intervalo de aceptación:

$$\left(\mu_0 - z_{0,01} \cdot \frac{\sigma_0}{\sqrt{n}}; +\infty\right) = (800 - 2,33 \cdot 17,0; +\infty) = (760; +\infty)$$

Como 750 no pertenece al intervalo de aceptación, rechazamos la hipótesis nula y aceptamos la alternativa, con un nivel de significación del 1%.

También podríamos calcular, como en el problema anterior, el nivel de significación que corresponde a la media muestral.

$$z_m = \frac{\bar{x} - \mu_0}{\frac{\sigma_0}{\sqrt{n}}} = \frac{750 - 800}{\frac{120}{\sqrt{50}}} = \frac{-50}{17} \approx -2,94$$

El intervalo de aceptación en $N(1, 0)$ es $(-z_\alpha, +\infty) = (-2,33; +\infty)$

Como -2,94 no pertenece al intervalo $(-2,33; +\infty)$: la media muestral está incluida en la región crítica o de rechazo y, por tanto, se rechaza H_0 y se acepta H_1 .

También podemos calcular el valor exacto del nivel de significación que corresponde a -2,94, que sería:

$$z_m = -2,94 \quad 1 - \alpha = 0,9984 \text{ (en tablas). } \alpha \approx 0,0016$$

Estos dos valores, -2,94 y 0,0016, son los que nos proporciona GeoGebra.

¿Cómo lo hacemos con GeoGebra? Una vez abierta la ventana de Cálculo de probabilidades, ahora seleccionamos **Test Z de una media**. Introducimos los datos y nos dará de inmediato los resultados:

The screenshot shows the 'Cálculos de Probabilidad' window in GeoGebra. The 'Estadísticas' tab is active, and 'Test Z de una Media' is selected. The configuration is as follows:

- Hipótesis Nula $\mu = 800$
- Hipótesis Alternativa: $<$ $>$ \neq
- Muestra:
 - Media: 750
 - σ : 120
 - N: 50

The results are displayed in a table:

Test Z de una Media	
Media	750
σ	120
ES	16.9706
N	50
Z	-2.9463
P	0.0016

Red arrows in the image point to the input fields for μ , the alternative hypothesis selection, and the sample statistics (Media, σ , N). Another set of red arrows points to the Z and P values in the results table. The word 'Datos' is written in red next to the input fields, and 'Resultados' is written in red next to the Z and P values.