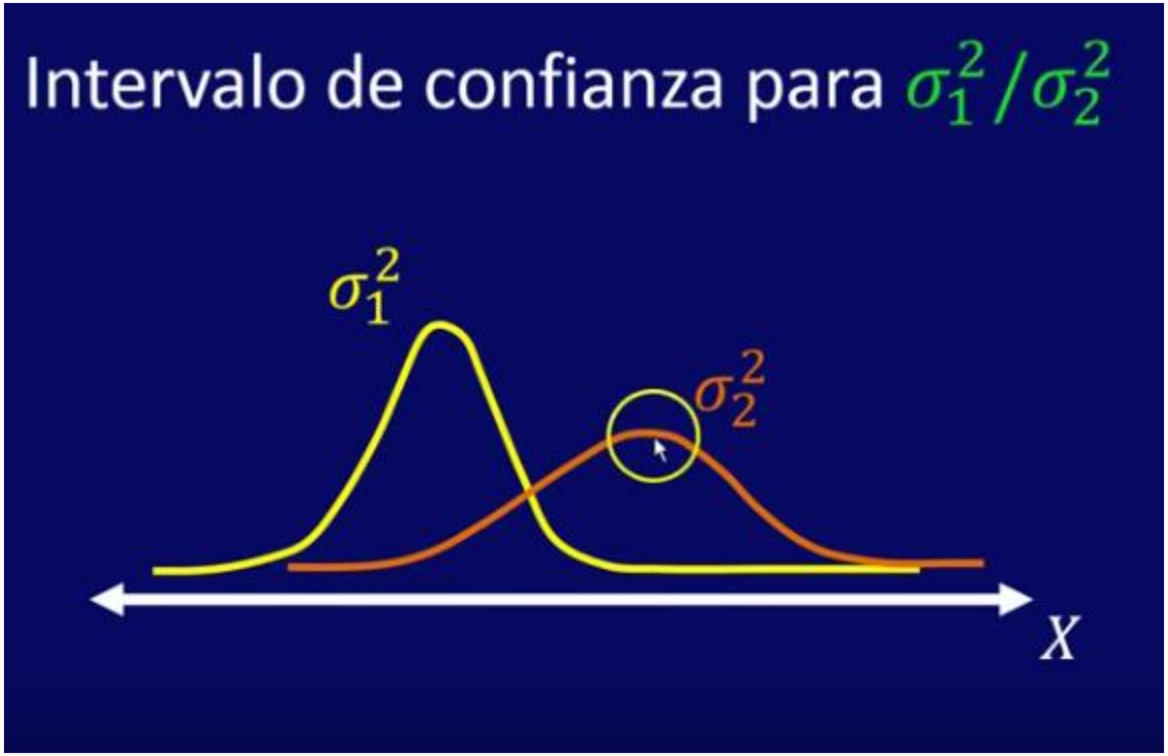
Intervalo de confianza para la razón de varianzas σ_1^2/σ_2^2



La pregunta de interés consiste en saber si son iguales las varianzas poblacionales?

El intervalo de confianza corresponde a:

$$\frac{{S_1}^2}{{S_2}^2} * \frac{1}{f_{\frac{\alpha}{2},n_1-1,n_2-1}} < \frac{\boldsymbol{\sigma_1}^{\mathbf{2}}}{\boldsymbol{\sigma_2}^{\mathbf{2}}} < \frac{{S_1}^2}{{S_2}^2} * f_{\frac{\alpha}{2},n_2-1,n_1-1}$$

Se observa si el intervalo de confianza contiene al 1:

- Si lo contiene se dice que hay evidencia de que las varianzas poblacionales son iguales

- Si el intervalo es inferior a 1 la varianza de la segunda población es superior a la varianza de la población 1

- Si el intervalo es superior a 1 la varianza de la primera población es superior a la varianza de la segunda población.

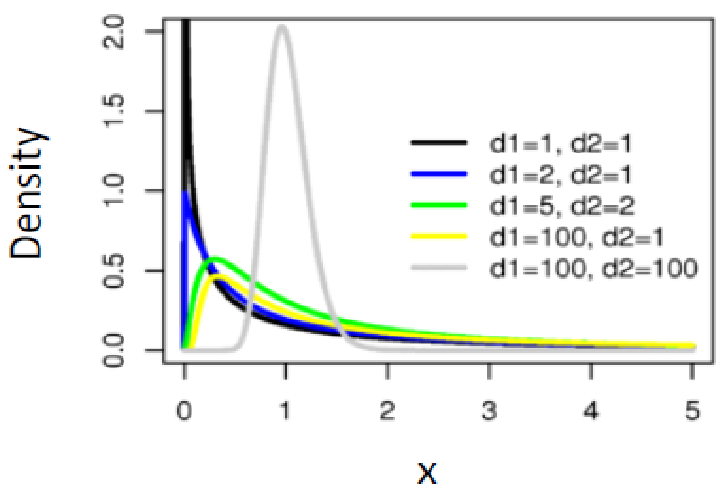
F distribution

This distribution is used in analysis of variance.

The probability density function (fdp) is given by

$$f(x;d_1,d_2) = \frac{\sqrt{\frac{(d_1\,x)^{d_1}\,\,d_2^{d_2}}{(d_1\,x+d_2)^{d_1+d_2}}}}{x\,\mathrm{B}\!\left(\frac{d_1}{2},\frac{d_2}{2}\right)}$$

with x ∈ $[0,\infty)$ and $d_1, d_2 > 0$

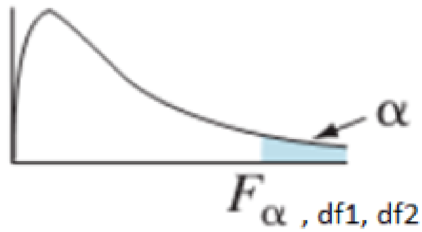


The parameter d_1, d_2 are called degrees of freedom (df)

Where $\mathrm{B}(x,y) = \frac{\Gamma(x)\,\Gamma(y)}{\Gamma(x+y)}$

F table

Areas under curve



		df₁			
df_2	**α**	**1**	**2**	**3**	**4**
1	.25	5.83	7.50	8.20	8.58
	.10	39.86	49.50	53.59	55.83
	.05	161.4	199.5	215.7	224.6
	.025	647.8	799.5	864.2	899.6
	.01	4052	5000	5403	5625
2	.25	2.57	3.00	3.15	3.23
	.10	8.53	9.00	9.16	9.24
	.05	18.51	19.00	19.16	19.25
	.025	38.51	39.00	39.17	39.25
	.01	98.50	99.00	99.17	99.25
	.005	198.5	199.0	199.2	199.2
	.001	998.5	999.0	999.2	999.2

Degrees of freedom df1

Quantiles

Degrees of freedom df2

EJEMPLO

- Se requiere determinar si dos tipos de motores, uno con gas y el otro con gasolina, tiene la misma variabilidad en millas por galón de gasolina o su equivalente en gas. Se tomaron un conjunto de autos de cada tipo y midió el rendimiento del combustible en un recorrido estándar, dando los siguientes resultados:

Gasolina	34	36	39	31	33	26	45	34	39	38	37	36		
Gas	33	41	39	32	29	28	33	34	25	28	36	33	35	35

- Construya un intervalo de confianza del 90% para la razón de varianzas

Solución

$$NC = 0.9 \qquad \alpha = 0.1 \qquad \frac{\alpha}{2} = 0.05$$

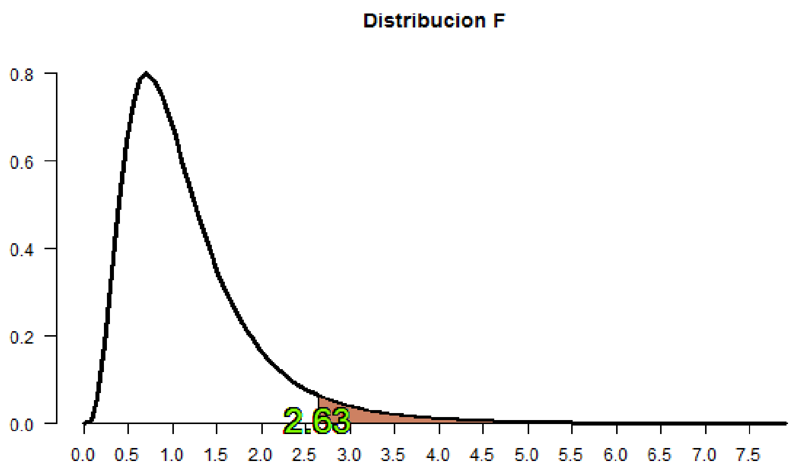
Limite inferior: $f_{\frac{\alpha}{2},n_1-1,n_2-1}$=$f_{0.05,11,13=2.634}$

Limite superior: $f_{\frac{\alpha}{2},n_2-1,n_1-1}$=$f_{0.05,13,11=2.7614}$

	Gasolina 1	**Gas 2**
n	12	14
S^2	22.24	19.15

Limite inferior:

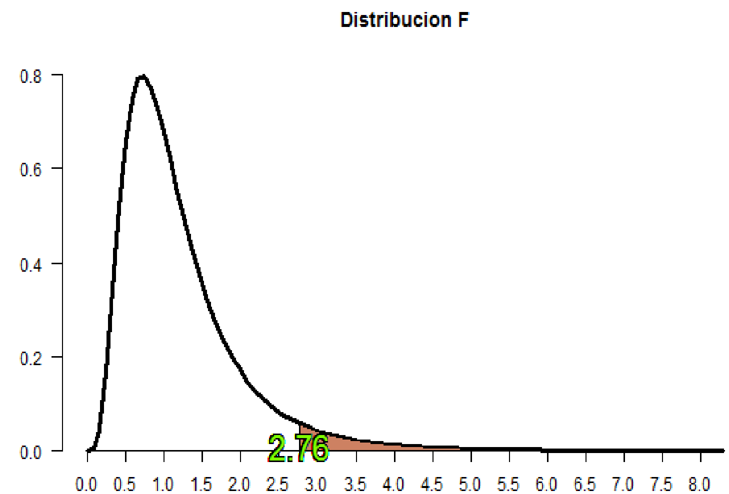
$$f_{\frac{\alpha}{2}, n_1-1, n_2-1} = f_{0.05,11,13=2.634}$$



$P(F > 2.63465) = 0.05$

Limite superior:

$$f_{\frac{\alpha}{2}, n_2-1, n_1-1} = f_{0.05,13,11=2.7614}$$



$P(F > 2.761417) = 0.05$

- Reemplazando en la fórmula:

$$\frac{{S_1}^2}{{S_2}^2} * \frac{1}{f_{\frac{\alpha}{2},n_1-1,n_2-1}} < \frac{\boldsymbol{\sigma_1}^{\mathbf{2}}}{\boldsymbol{\sigma_2}^{\mathbf{2}}} < \frac{{S_1}^2}{{S_2}^2} * f_{\frac{\alpha}{2},n_2-1,n_1-1}$$

$$\frac{22.24}{19.15} * \frac{1}{2.634} < \frac{\boldsymbol{\sigma_1}^{\mathbf{2}}}{\boldsymbol{\sigma_2}^{\mathbf{2}}} < \frac{22.24}{19.15} * 2.76$$

$$0.45 < \frac{\boldsymbol{\sigma_1}^{\mathbf{2}}}{\boldsymbol{\sigma_2}^{\mathbf{2}}} < 3.2$$

Otro ejemplo

Un fabricante de automóviles pone a prueba dos nuevos métodos de ensamblaje de motores respecto al tiempo en minutos. Los resultados se muestran en la siguiente tabla

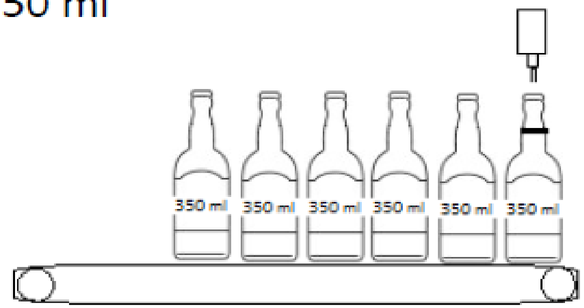
Método	**1**	**2**
n	31	25
S^2	50	24

Construya un intervalo de confianza del 90% para la razón de varianzas

Pruebas de hipótesis

Motivación 1

Proceso de llenado de cervezas de 350 ml

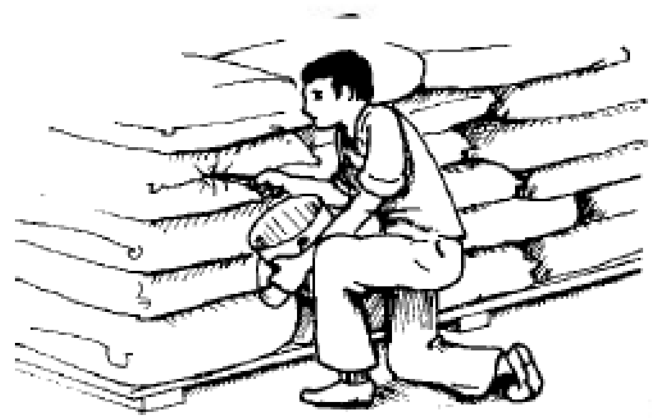


Muestra de cervezas

¿Está el proceso de llenado cumpliendo con lo prometido en la etiqueta?

¿Será el contenido medio de cerveza μ igual a 350ml?

Motivación 2



¿Está el saco de café cumpliendo con lo porcentaje de granos defectuosos?

¿Será el porcentaje p de granos defectuosos igual al 3%?

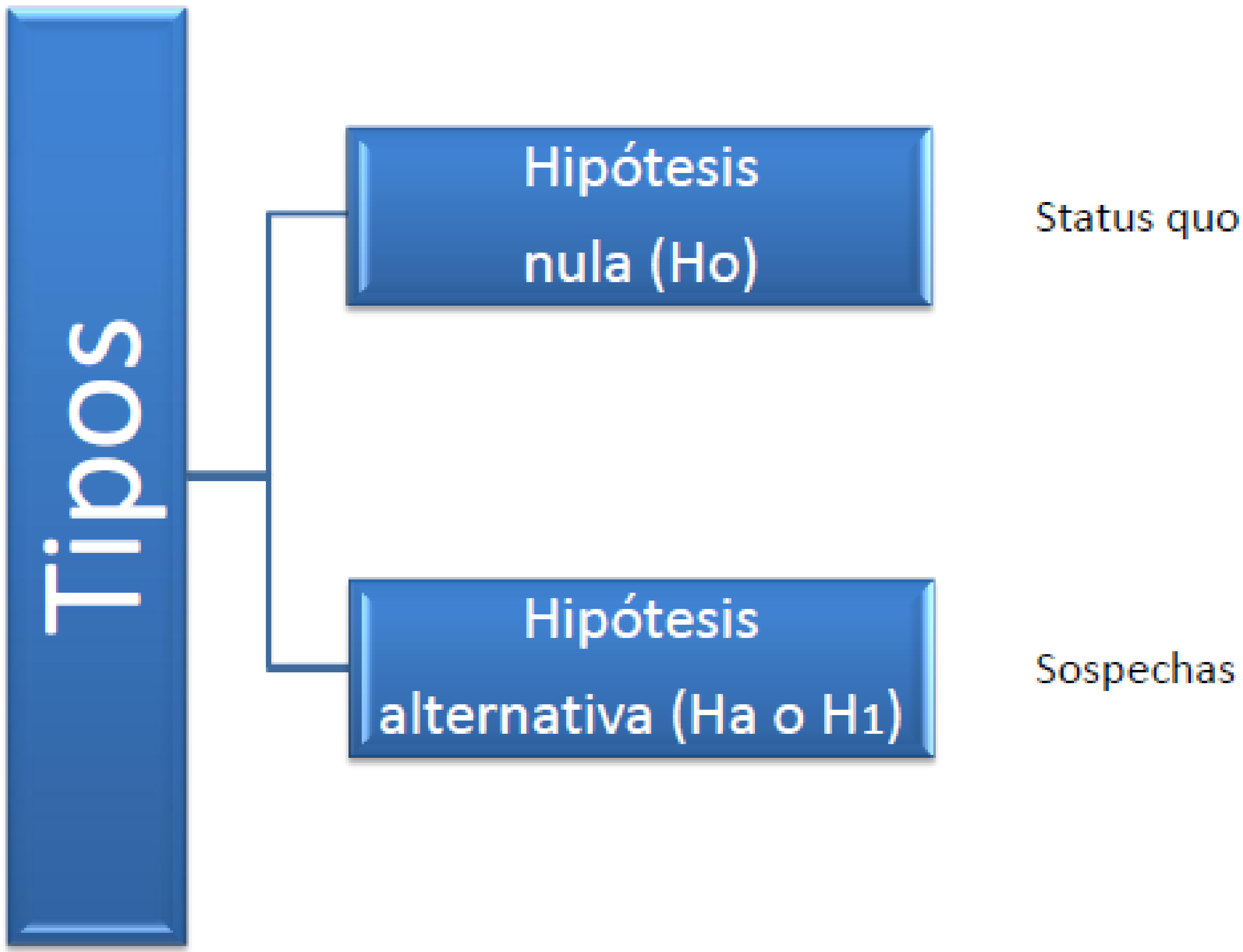
Problemas de interés en inferencia

- Problemas
  - Estimación
    - Puntual
    - Por intervalo
  - Prueba de hipótesis

Prueba de hipótesis

Una hipótesis estadística es una aseveración o conjetura con respecto a una o más poblaciones.

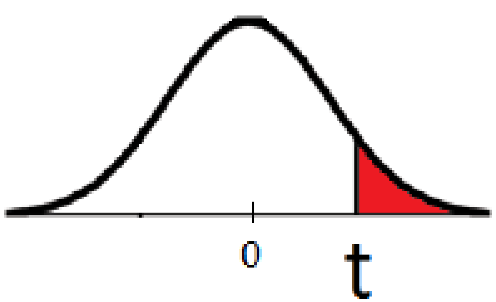
Tipos de hipótesis



Tipos de pruebas

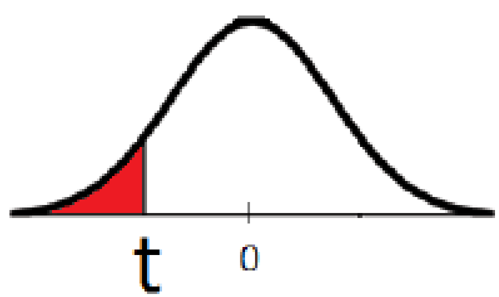
$$H_0: \quad \theta = \theta_0$$
$$H_1: \quad \theta > \theta_0$$

Prueba unilateral derecha



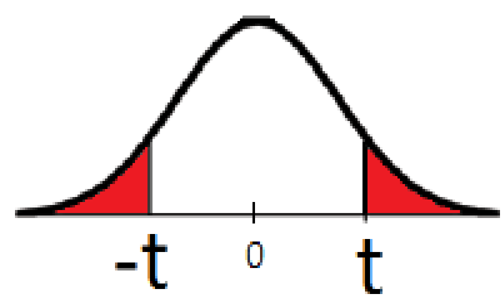
$$H_0: \quad \theta = \theta_0$$
$$H_1: \quad \theta < \theta_0$$

Prueba unilateral izquierda

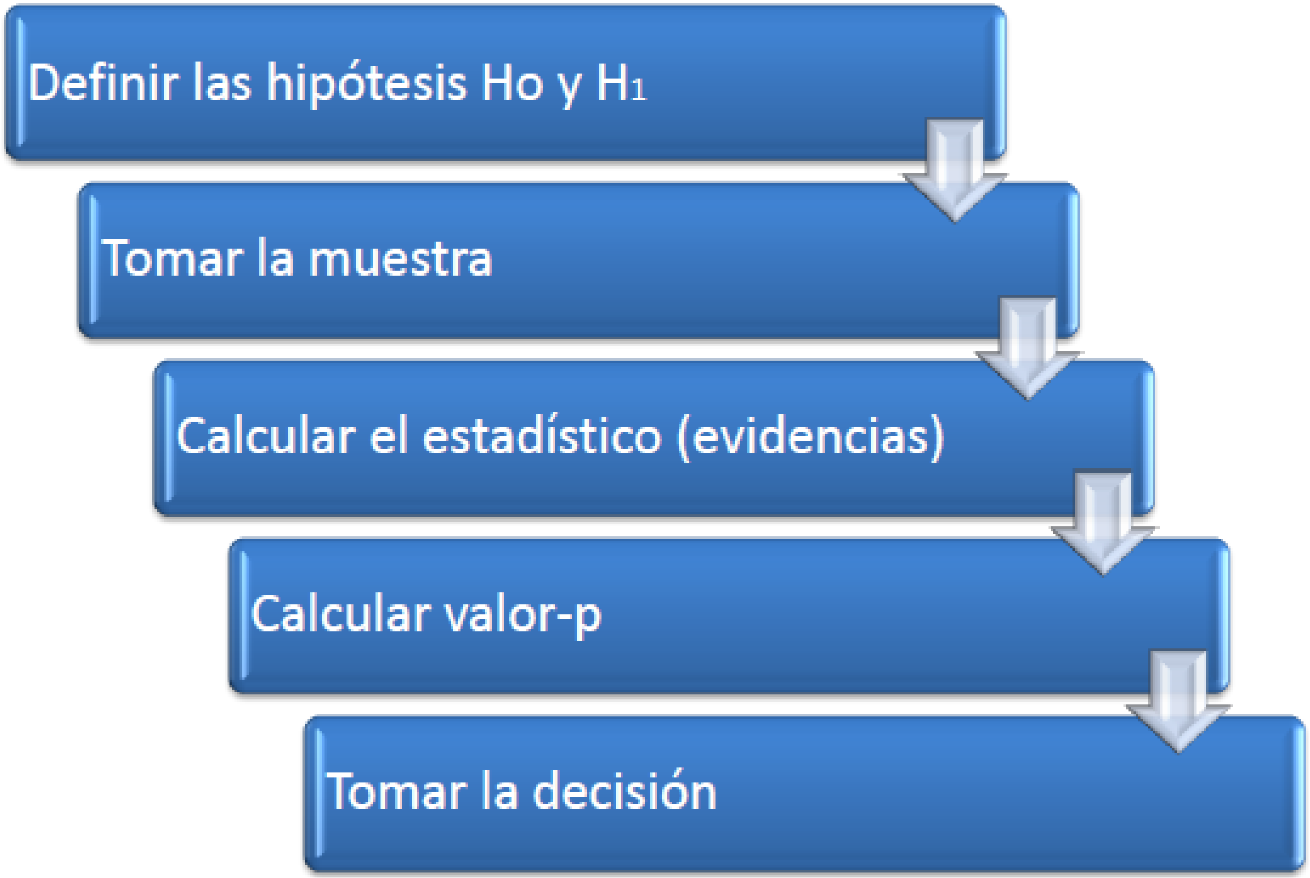


$$H_0: \quad \theta = \theta_0$$
$$H_1: \quad \theta \neq \theta_0$$

Prueba bilateral



Proceso de prueba de hipótesis



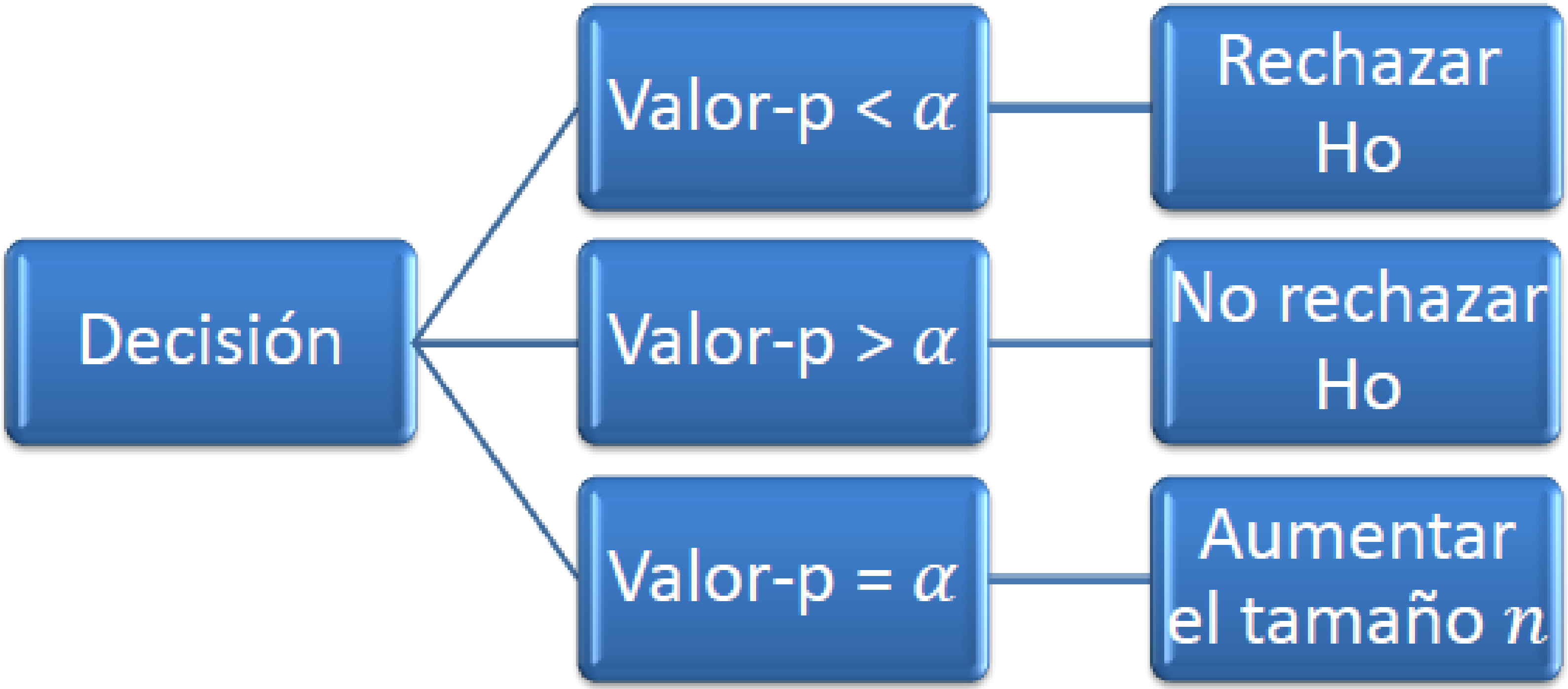
Tipos de errores en prueba de hipótesis

	Situación real de Ho	
Decisión	Ho es verdadera	Ho es falsa
Aceptar Ho	👍	Error tipo II (β)
Rechazar Ho	Error tipo I (α)	👍

Valor p

- El valor-p de una prueba de hipótesis es la probabilidad de obtener un estadístico (evidencias) igual al que se obtuvo o más extremo.
- Probabilidad de equivocarse rechazando la hipótesis nula.

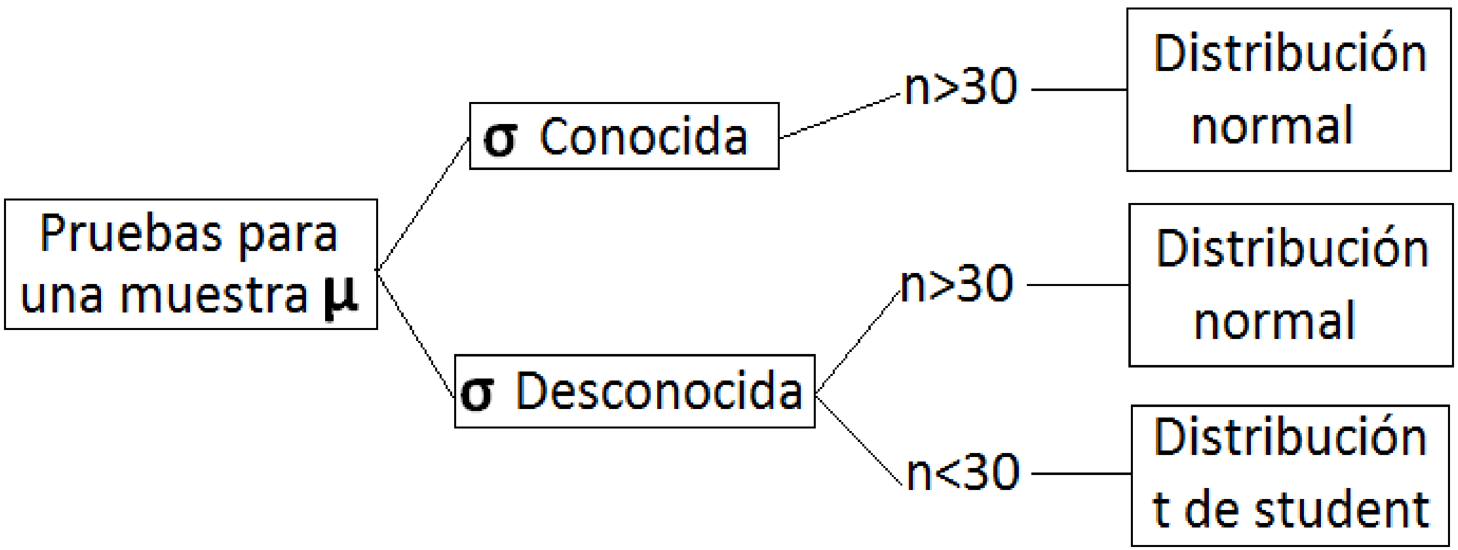
Toma de decisión en pruebas de hipótesis



Problemas de pruebas de hipótesis

- Pruebas
  - Para una muestra
    - μ
    - p
    - σ^2
  - Para dos muestras
    - $\mu_1 - \mu_2$
    - $\mu_1 - \mu_2$ pareadas
    - $p_1 - p_2$
    - σ_1^2 / σ_2^2

Pruebas para la media μ



1. Prueba de hipótesis para μ

Paso 1. Definir Ho, H1 y el tipo de prueba

$$\begin{aligned} H_0&: \ \mu = \mu_0 \\ H_1&: \ \mu > \mu_0 \end{aligned} \qquad \begin{aligned} H_0&: \ \mu = \mu_0 \\ H_1&: \ \mu < \mu_0 \end{aligned} \qquad \begin{aligned} H_0&: \ \mu = \mu_0 \\ H_1&: \ \mu \neq \mu_0 \end{aligned}$$

Paso 2. Calcular el estadístico

Media muestral

$$t = \frac{\bar{x} - \mu_0}{s/\sqrt{n}}$$

Valor de referencia

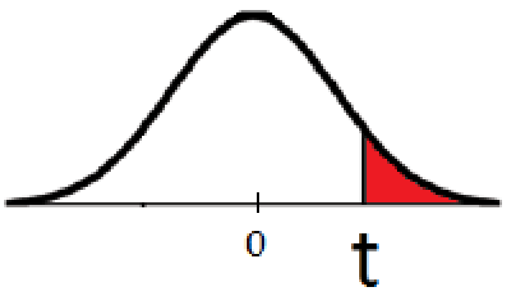
Tamaño de muestra

Desviación muestral

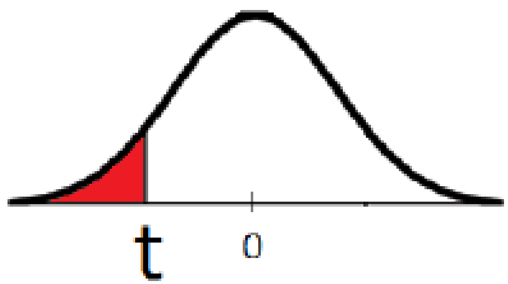
Paso 3. Definir el error tipo I denotado por α

Paso 4. Calcular el valor-p en una distribución t-student con $n-1$ grados de libertad

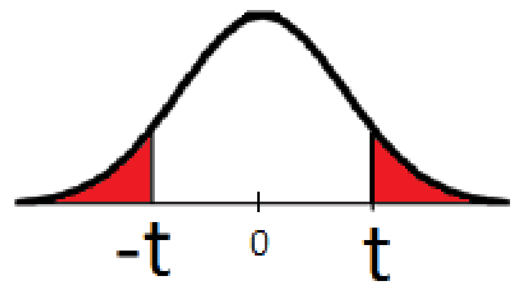
$$H_0: \quad \mu = \mu_0$$
$$H_1: \quad \mu > \mu_0$$



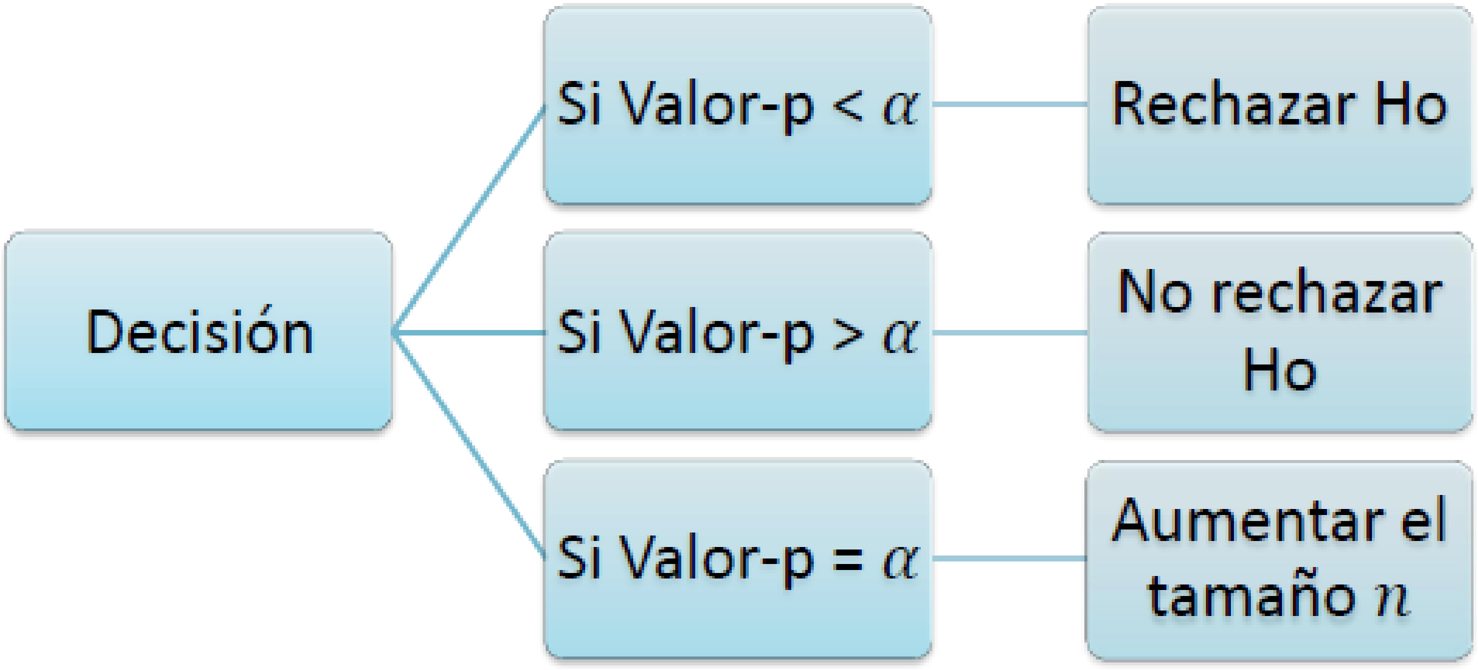
$$H_0: \quad \mu = \mu_0$$
$$H_1: \quad \mu < \mu_0$$



$$H_0: \quad \mu = \mu_0$$
$$H_1: \quad \mu \neq \mu_0$$



Paso 5. Comparar valor-p con α y tomar la decisión

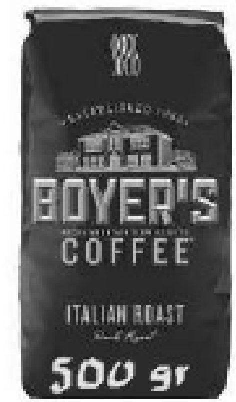


Ejemplo 1

Para verificar si el proceso de llenado de bolsas de café con 500 gramos está operando correctamente se toman aleatoriamente muestras de tamaño diez cada cuatro horas. Una muestra de bolsas está compuesta por las siguientes observaciones:

502, 501, 497, 491, 496, 501, 502, 500, 489, 490.

¿Está el proceso llenando bolsas conforme lo dice la envoltura?



Paso 1. Definir Ho y H1

$$H_0: \quad \mu = 500\ gr$$
$$H_1: \quad \mu \neq 500\ gr$$

Paso 2. Calcular el estadístico

De la muestra tenemos que:

$$\bar{x} = 496.9\ gr$$
$$s = 5.174\ gr$$
$$n = 10$$

Por tanto el estadístico es:

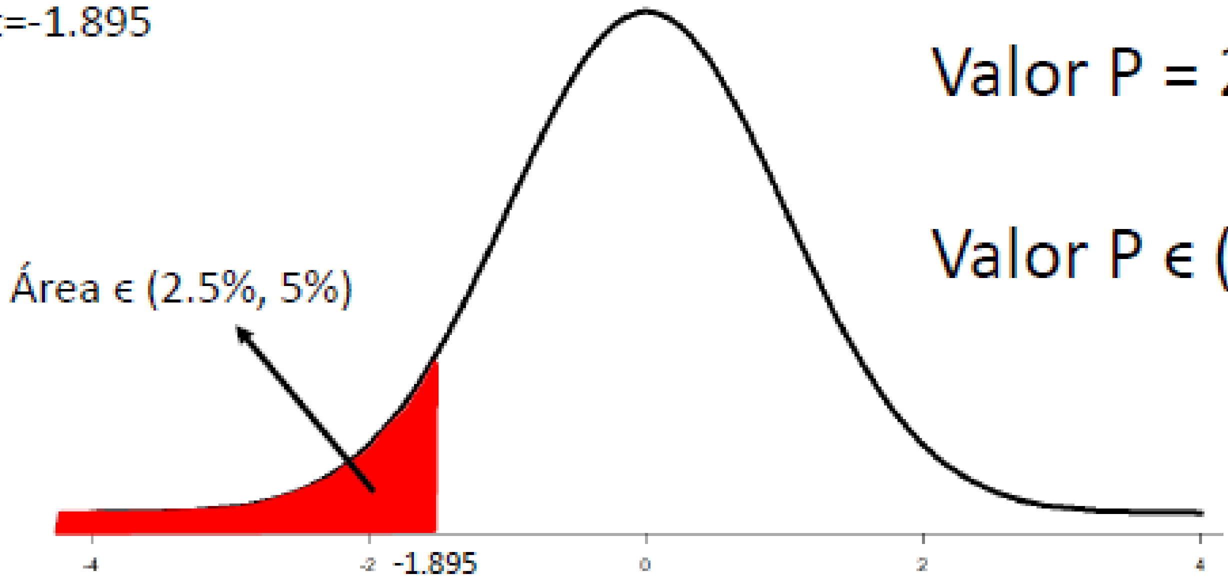
$$t = \frac{\bar{x} - \mu_0}{s/\sqrt{n}} = \frac{496.9 - 500}{5.174/\sqrt{10}} = -1.895$$

Paso 3. Definir el error tipo I denotado por α

Usemos $\alpha = 0.05$

Paso 4. Calcular el valor-p en una distribución t-student con $n-1$ grados de libertad

t=-1.895



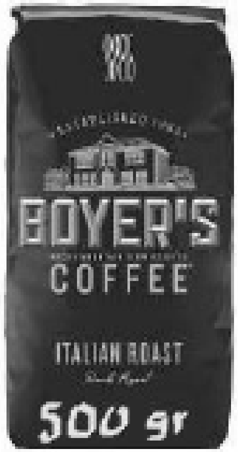
Valor P = 2 x Área roja

Valor P ϵ (5%, 10%)

Paso 5. Comparar valor-p con α y tomar la decisión

Como Valor P $>$ $\alpha=0.05$ entonces NO RECHAZAMOS Ho

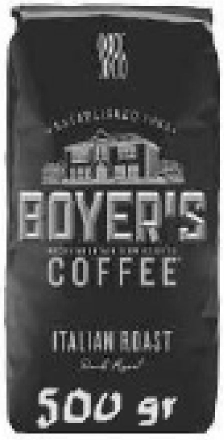
El proceso está llenando las bolsas conforme al aviso en la envoltura.



Tiempo después se lleva a cabo otra verificación del proceso y se obtiene la siguiente muestra:

500, 495, 494, 498, 495, 500, 500, 496, 498, 493.

¿Está el proceso llenando bolsas conforme lo dice la envoltura?



Paso 1. Definir Ho y H1

$$H_0\colon \quad \mu = 500\ gr$$
$$H_1\colon \quad \mu \neq 500\ gr$$

Paso 2. Calcular el estadístico

De la muestra tenemos que:

$$\bar{x} = 496.9\ gr$$
$$s = 2.643\ gr$$
$$n = 10$$

Por tanto el estadístico es:

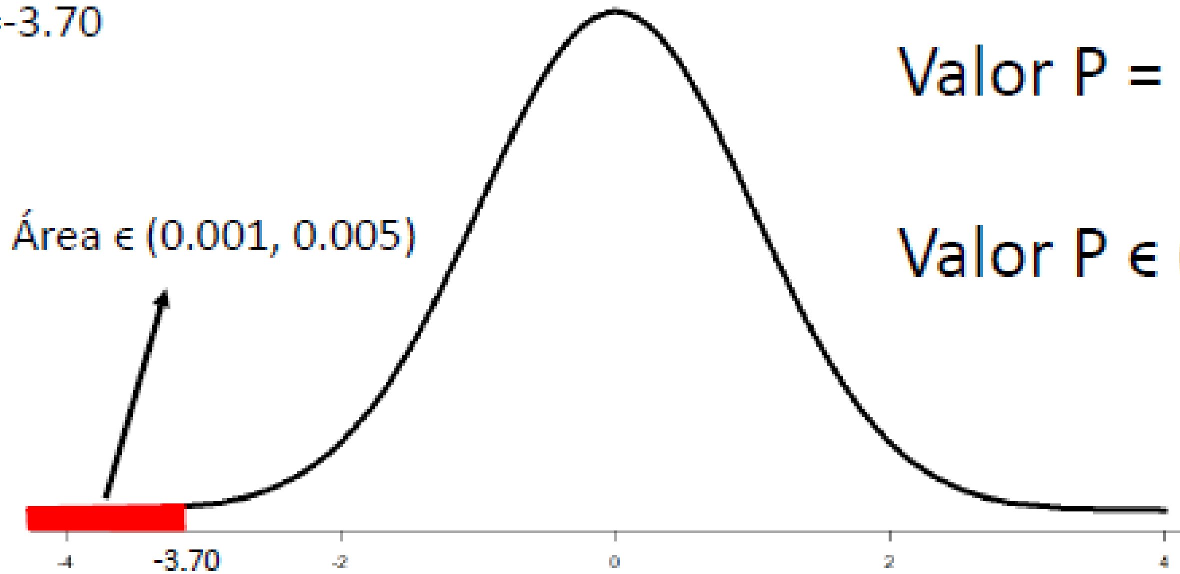
$$t = \frac{\bar{x} - \mu_0}{s/\sqrt{n}} = \frac{496.9 - 500}{2.643/\sqrt{10}} = -3.70$$

Paso 3. Definir el error tipo I denotado por α

Usemos $\alpha = 0.05$

Paso 4. Calcular el valor-p en una distribución t-student con $n-1$ grados de libertad

t=-3.70



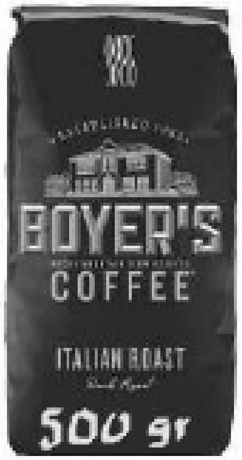
Valor P = 2 x Área roja

Valor P ∈ (0.002, 0.01)

Paso 5. Comparar valor-p con α y tomar la decisión

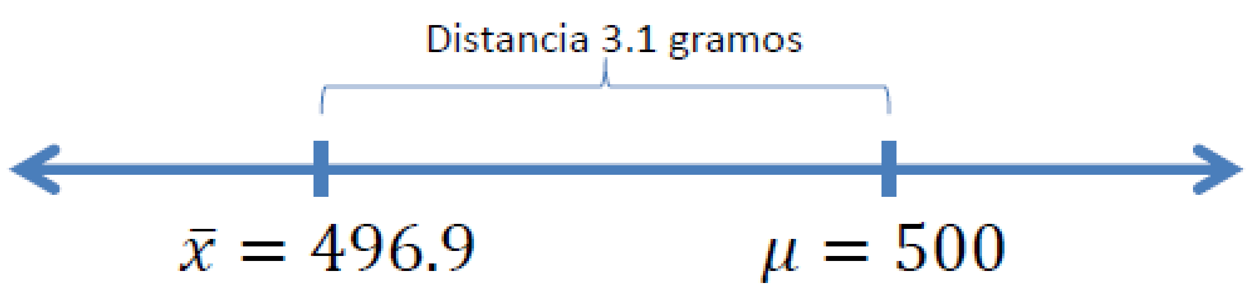
Como Valor P $<$ α=0.05 entonces RECHAZAMOS Ho

El proceso NO está llenando las bolsas conforme al aviso en la envoltura.

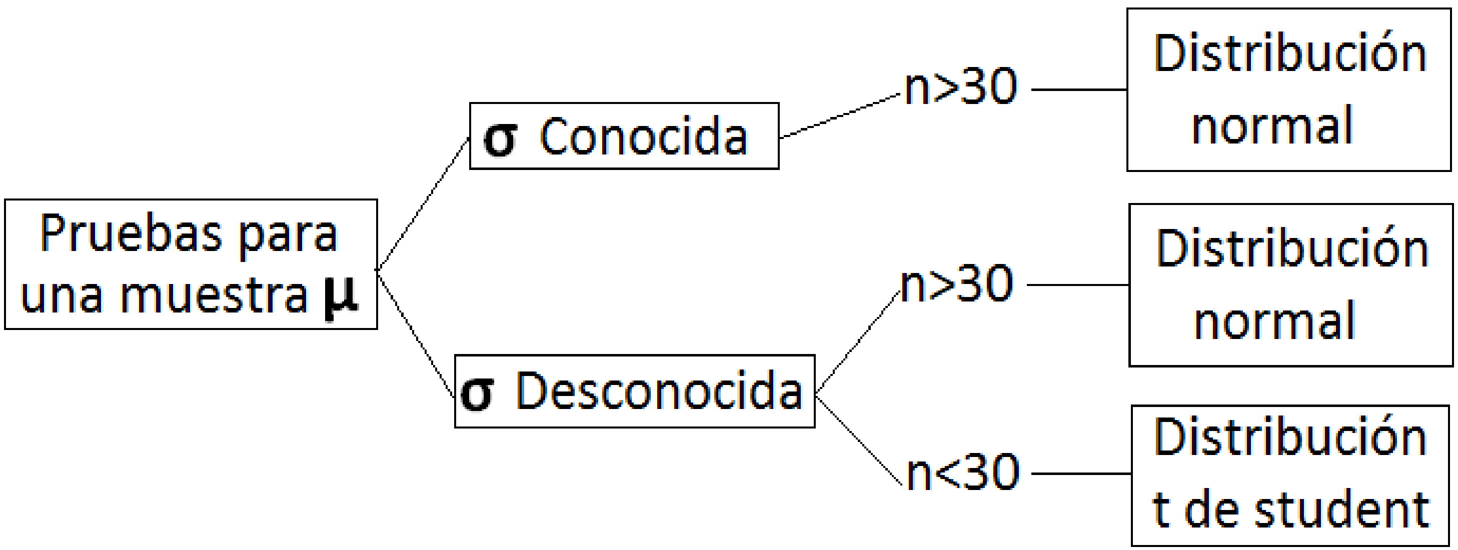


Resumen de los ejemplos 1 y 2

	Ejemplo 1	Ejemplo 2
Muestra	502, 501, 497, 491, 496, 501, 502, 500, 489, 490	500, 495, 494, 498, 495, 500, 500, 496, 498, 493
Media muestral	496.9 gramos	496.9 gramos
Desviación muestral	5.174 gramos	2.643 gramos
Tamaño de muestra	10	10
Hipótesis nula	$H_0: \quad \mu = 500\ gr$	$H_0: \quad \mu = 500\ gr$
Decisión	Aceptamos Ho	Rechazamos Ho



Pruebas para la media μ



Población normal con σ conocida

Hipótesis nula: $H_0\colon\ \mu = \mu_0$

Valor del estadístico de prueba: $z = \dfrac{\bar{x} - \mu_0}{\sigma/\sqrt{n}}$

Hipótesis alternativa		Región de rechazo para la prueba de nivel α
$H_a\colon\ \mu > \mu_0$		$z \geq z_\alpha$ (prueba de cola superior)
$H_a\colon\ \mu < \mu_0$		$z \leq -z_\alpha$ (prueba de cola inferior)
$H_a\colon\ \mu \neq \mu_0$	o	$z \geq z_{\alpha/2}$ o $z \leq -z_{\alpha/2}$ (prueba de dos colas)

EJEMPLO

Un fabricante de rociadores utilizados como protección contra incendios en edificios de oficinas, afirma que la temperatura de activación del sistema promedio verdadera es de 130°F. Una muestra de n=9 sistemas, cuando se somete a prueba, da una temperatura de activación promedio muestral de 131.08°F. Si la distribución de los tiempos de activación es normal con desviación estándar de 1.5°F, ¿contradicen los datos la afirmación del fabricante a un nivel de significancia 0.01?

PASOS A REALIZAR

Parámetro de interés:

μ=temperatura de activación promedio verdadera

1. Definir Ho y H1:

Ho: $\mu=130$ H1: $\mu\neq130$

2. Calcule el estadístico

De la muestra se tiene:

$\bar{X}$=131.08

σ=1.5

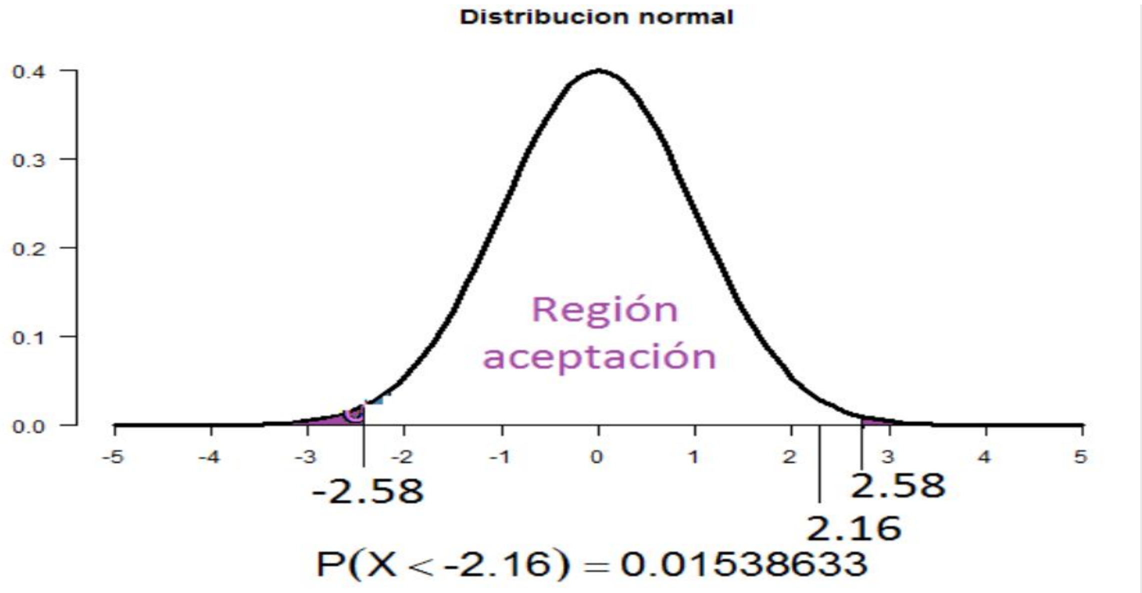
n=9

$$z = \frac{\bar{x} - \mu}{\sigma/\sqrt{n}} = \frac{131.08 - 130}{1.5/\sqrt{9}} = 2.16$$

3. Definir el error tipo 1 denotado por α

NC=0.99 α=0.01 α/2=0.005 $Z_{\alpha/2} = 2.58$

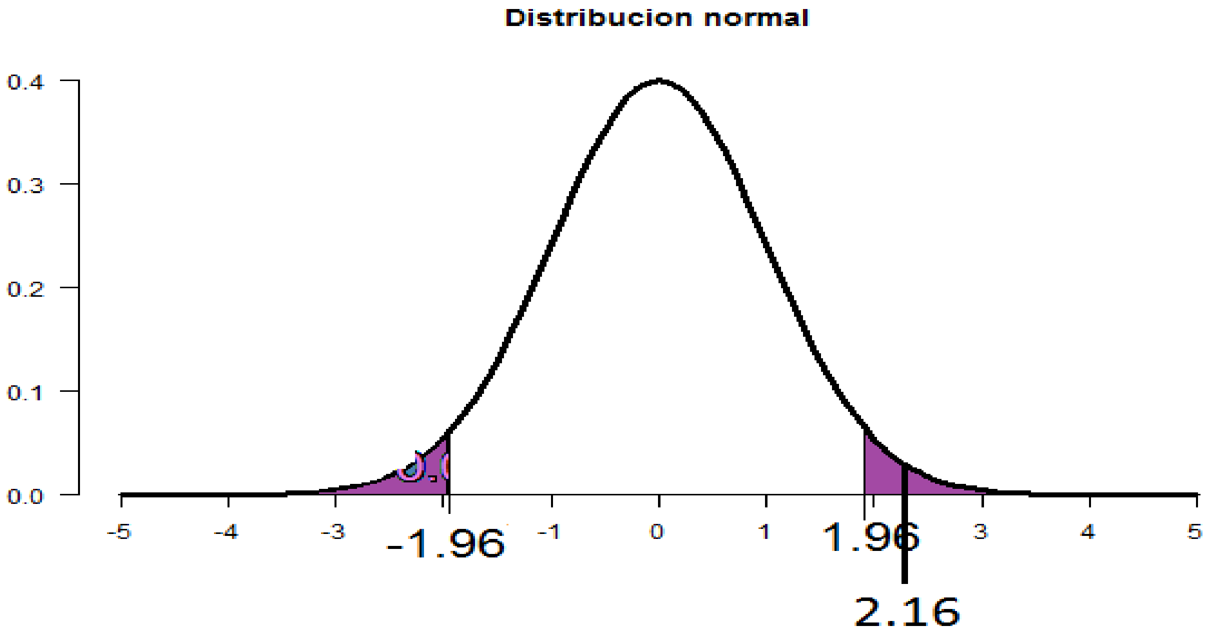
4. Defina la zona de aceptación y de rechazo



Con un nivel de significancia del 1% se acepta la hipótesis nula,
no existe evidencia para rechazar la evidencia de que la media difiera de 130

¿Qué pasa si se cambia el nivel de significancia?

NC=0.95 α=0.05 α/2= 0.025 $Z_{\alpha/2} = 1.96$



Con un nivel de significancia del 5%, se rechaza la hipótesis nula, existe evidencia para afirmar de que la temperatura media de los rociadores difiere de 130

2. Prueba de hipótesis para p

Paso 1. Definir Ho, H1 y el tipo de prueba

$$\begin{array}{lll} H_0:\ \ p = p_0 & \qquad H_0:\ \ p = p_0 & \qquad H_0:\ \ p = p_0 \\ H_1:\ \ p > p_0 & \qquad H_1:\ \ p < p_0 & \qquad H_1:\ \ p \neq p_0 \end{array}$$

Paso 2. Calcular el estadístico

Proporción muestral

$$Z_0 = \frac{\hat{p} - p_0}{\sqrt{\frac{p_0(1-p_0)}{n}}}$$

Valor de referencia

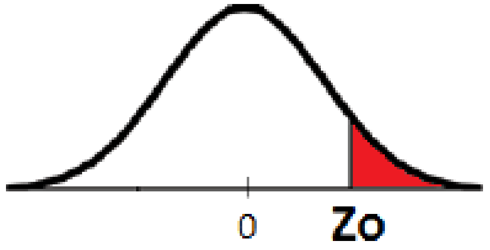
Tamaño de muestra

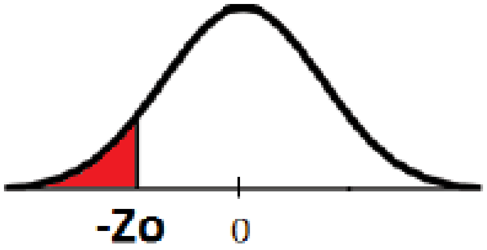
2. Prueba de hipótesis para p

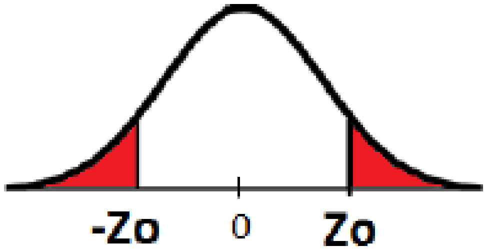
Paso 3. Definir el error tipo I denotado por α

Paso 4. Calcular el valor-p en una distribución normal

$$H_0: \quad p = p_0 \qquad H_0: \quad p = p_0 \qquad H_0: \quad p = p_0$$
$$H_1: \quad p > p_0 \qquad H_1: \quad p < p_0 \qquad H_1: \quad p \neq p_0$$

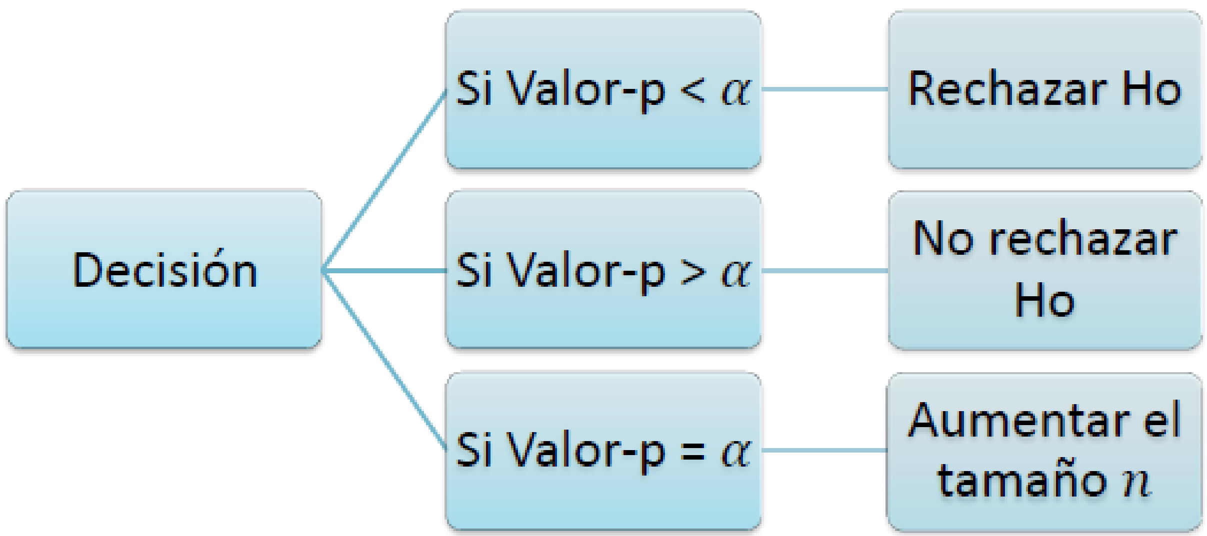






2. Prueba de hipótesis para p

Paso 5. Comparar valor-p con α y tomar la decisión



EJEMPLO

Un fabricante de un quitamanchas afirma que su producto quita 90 por ciento de todas las manchas. Para poner a prueba esta afirmación se toman 200 camisetas manchadas de las cuales a solo 174 les desapareció la mancha. Pruebe la afirmación del fabricante a un nivel $\alpha = 0.05$

Paso 1. Definir Ho y H1

$$H_0: \quad p = 0.90$$
$$H_1: \quad p < 0.90$$

Paso 2. Calcular el estadístico

De la muestra tenemos que:

$$n = 200$$
$$x = 174$$
$$\hat{p} = \frac{174}{200} = 0.87$$

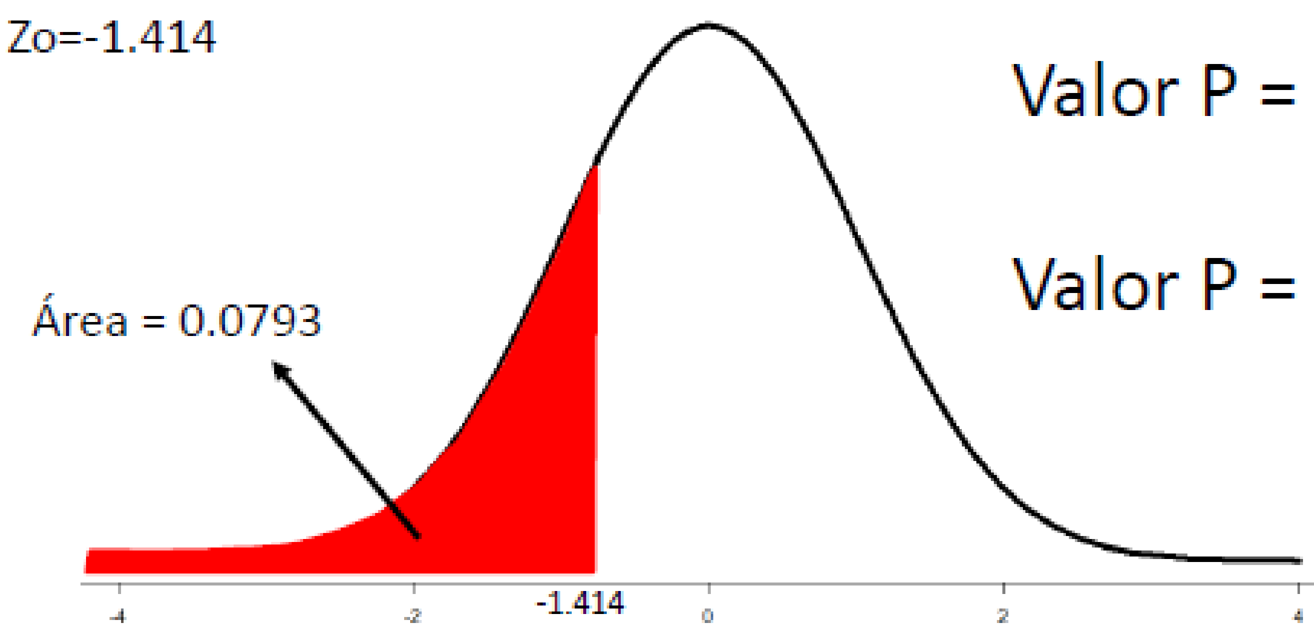
Por tanto el estadístico es:

$$Z_0 = \frac{\hat{p} - p_0}{\sqrt{\frac{p_0(1 - p_0)}{n}}} = \frac{0.87 - 0.90}{\sqrt{\frac{0.90(1 - 0.90)}{200}}} = -1.414$$

Paso 3. Definir el error tipo I denotado por α

Del enunciado tenemos que $\alpha = 0.05$

Paso 4. Calcular el Valor P en una normal



Valor P = Área roja

Valor P = 7.93%

Paso 5. Comparar valor-p con α y tomar la decisión

Como valor-p= 0.0793 > α=0.05 entonces NO RECHAZAMOS Ho

No hay evidencias suficientes para afirmar que el detergente no limpia el 90% de todas las manchas.

Pruebas de hipótesis para la varianza de una distribución normal

Con frecuencia surgen problemas que requieren inferencias acerca de la variabilidad, por ejemplo considere la variabilidad de las calificaciones otorgadas por cierto profesor en determinado examen. Se esperaría que las puntuaciones tuvieran una varianza pequeña y que además su media fuera mayor o igual al promedio mínimo aprobatorio.

En otras palabras, interesa la prueba de hipótesis relacionada con la uniformidad de una población. Se parte bajo el supuesto de la muestra proviene de una población que se distribuye normal.

Hipótesis de interés

- H_o: $\sigma^2 = \sigma^2_o$. En oposición a H_a: $\sigma^2 \neq \sigma^2_o$.

- H_o: $\sigma^2 = \sigma^2_o$. En oposición a H_a: $\sigma^2 > \sigma^2_o$.

- H_o: $\sigma^2 = \sigma^2_o$. En oposición a H_a: $\sigma^2 < \sigma^2_o$.

donde σ^2_o es el valor propuesto de σ^2.

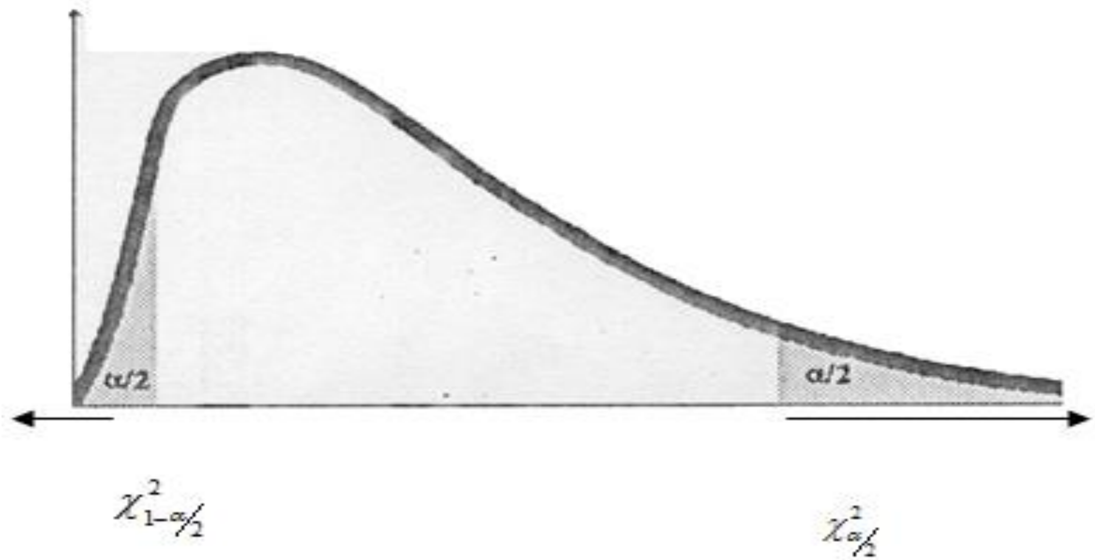
El estadístico

El estadístico para probar estas hipótesis se basa en la varianza muestral S^2.

$$\frac{(n-1)S^2}{\sigma^2} \approx \chi^2_{n-1}$$

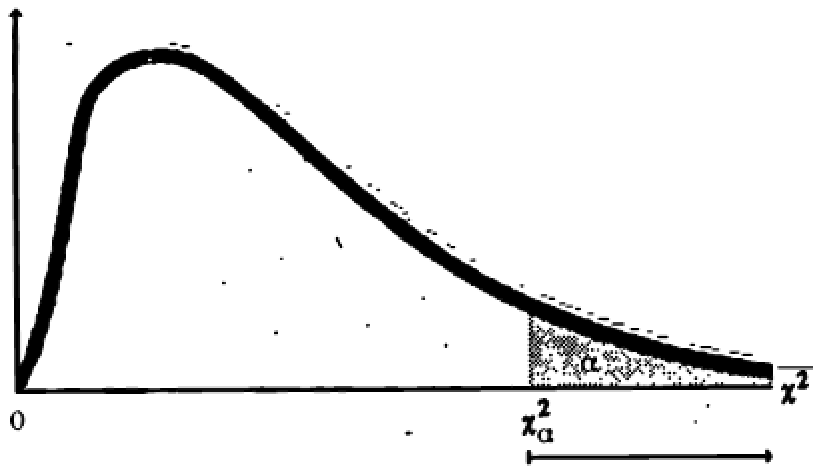
$$\chi^2_c = \frac{(n-1)S^2}{\sigma^2_o}$$

Regiones de rechazo



Región de rechazo para una prueba de hipótesis bilateral para σ^2

Región de rechazo de cola derecha



Región de rechazo de cola izquierda

EJEMPLO

Datos de archivo indican que la varianza de las mediciones efectuadas sobre lámina metálica grabada, las cuales fueron obtenidas por inspectores expertos en control de calidad es de 0.18 pulgadas cuadradas.

Las mediciones realizadas por un inspector sin experiencia podrían tener una varianza mayor o menor.

Si un nuevo inspector mide 101 laminas grabadas con una varianza de 0.13 pulgadas cuadradas, pruébese con un nivel de significancia de 0.05 si el inspector realiza mediciones satisfactorias.

Solución

1. Definir Ho y H1:

$$H_0\text{: } \sigma^2 = 0.18$$
$$H_1\text{: } \sigma^2 \neq 0.18$$

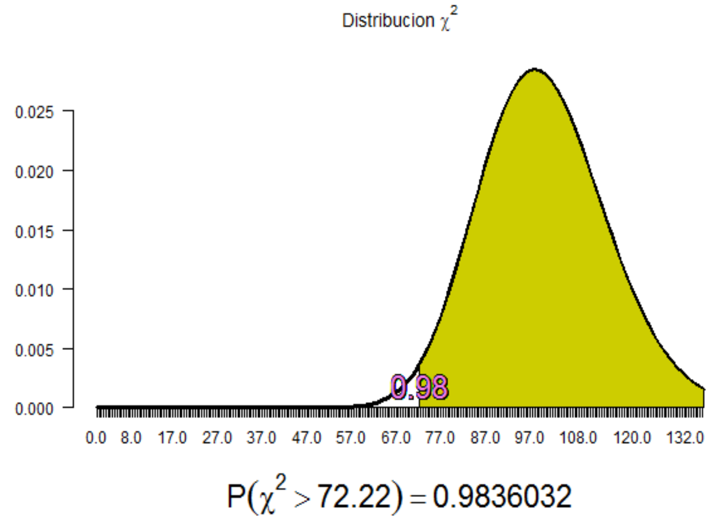
2. Calcule el estadístico

De la muestra se tiene:

σ^2=0.13

n=101

$$\chi^2 = \frac{(n-1)S^2}{\sigma^2}$$
$$= \frac{(100)0.130}{0.18} = 72.22$$

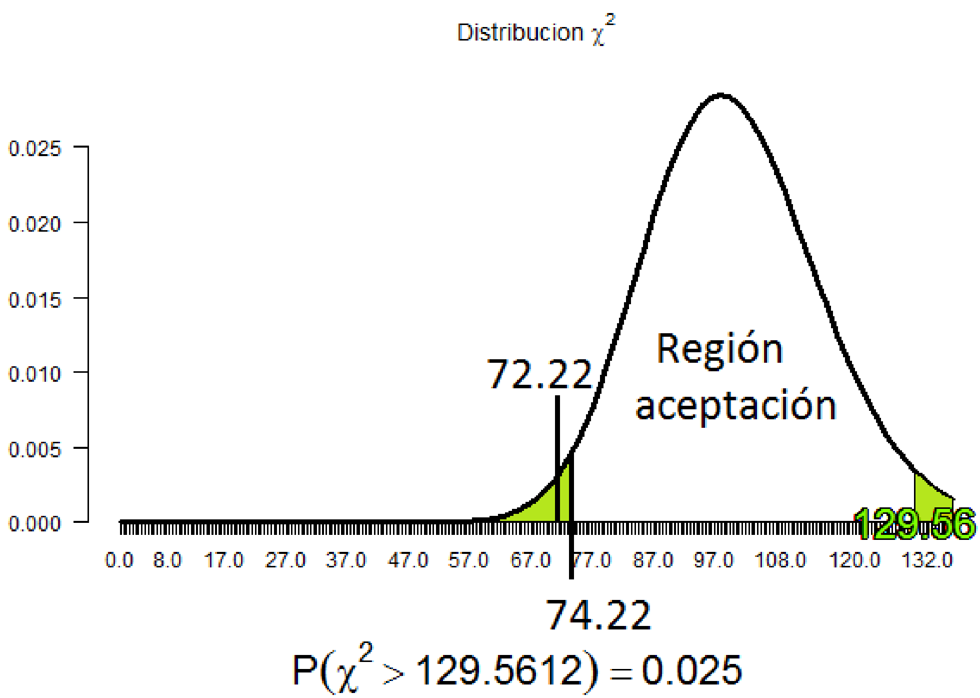


3. Definir el error tipo 1 denotado por α

$NC=95\% \quad \alpha=5\% \quad \alpha/2=0.025$

4. Calcular el valor p

$Vp=0.98$

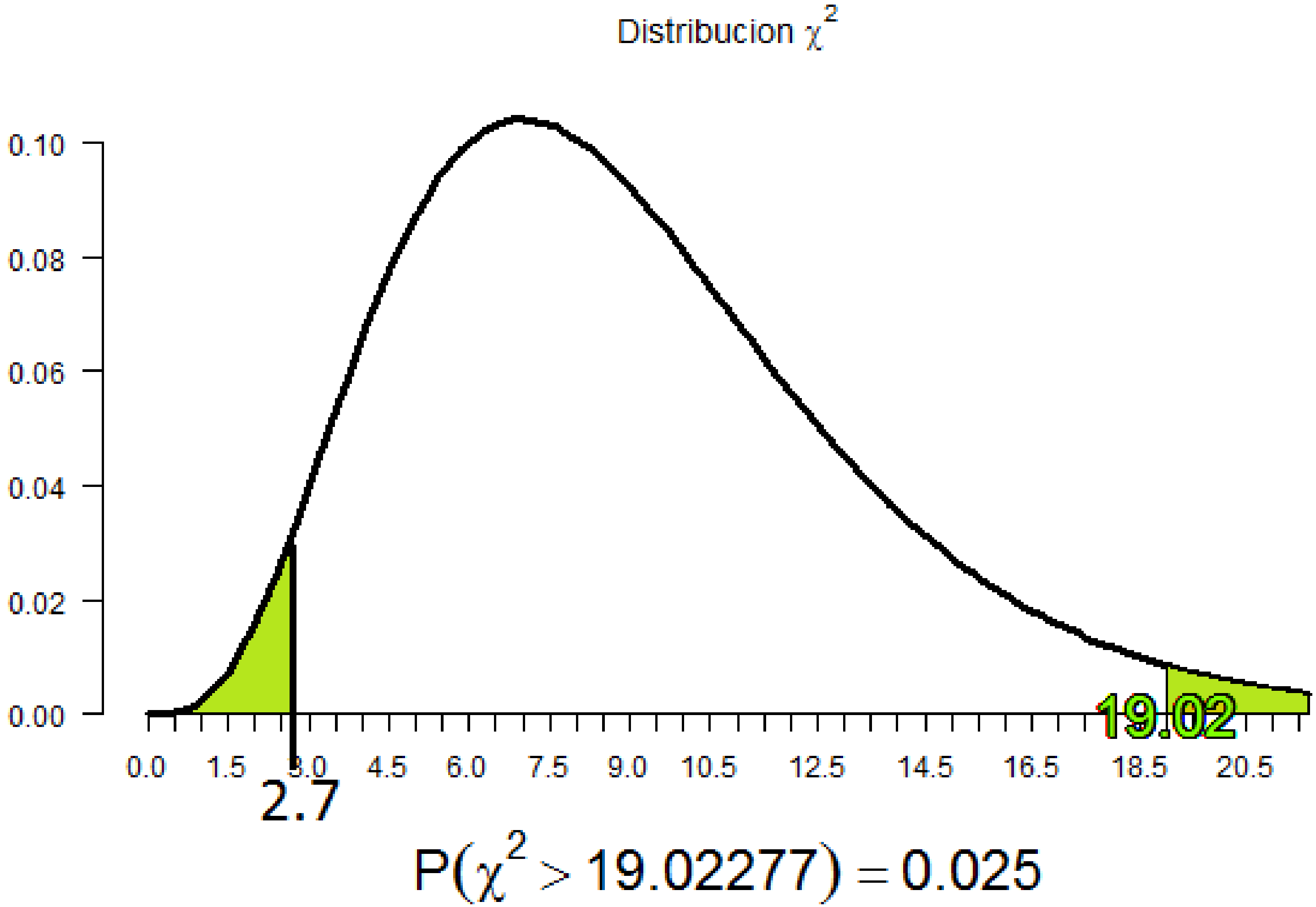


Decisión:

Rechazar H_o en favor de H_a con un nivel de significancia $\alpha=0.05$. Es decir existe evidencia muestral suficiente para indicar que el nuevo inspector no toma satisfactoriamente sus mediciones.

Otro ejemplo

El gerente de una planta sospecha que el número de piezas que produce un trabajador en particular por día, fluctúa más allá del valor normal esperado. El gerente decide observar el número de piezas que produce este trabajador durante 10 días, seleccionados estos al azar. Los resultados son 15, 12, 8, 13, 12, 15, 16, 9, 8 y 14. Si se sabe que la desviación estándar para todos los trabajadores es de 2 unidades y si el numero de estas que se produce diariamente se encuentra modelada en forma adecuada por una distribución normal, aun nivel de a = 0.05, ¿ Tiene apoyo la sospecha del gerente?



Distribucion χ^2
0.10
0.08
0.06
0.04
0.02
0.00
0.0
1.5
3.0
4.5
6.0
7.5
9.0
10.5
12.5
14.5
16.5
18.5
20.5
2.7
19.02
$P(\chi^2 > 19.02277) = 0.025$