

## **EJERCICIOS INFERENCIA SOBRE UNA DISTRIBUCIONES**

### **NORMALES 5.2**

#### **Ejercicio 1:**

**A partir de la transparencia 35 UD5 inferencia parte 1-2.pdf**

La media muestral de una muestra con 10 medidas vale 360,89. La desviación típica muestral vale 1,78.

Se plantea la hipótesis simple nula de que  $m=350$ .  $H_0: m=350$ ,  $H_1: m \neq 350$

Resolución según test de hipótesis:

Test de hipótesis para la media (Test t)		
$H_0: m = m_0$ $H_1: m \neq m_0$	Si $\left  \frac{\bar{X} - m_0}{S/\sqrt{N}} \right  \leq t_{N-1}^{\alpha/2} \Rightarrow$ Aceptar $H_0$	Si <b>p-value</b> $\geq \alpha \Rightarrow$ Aceptar $H_0$
	Si $\left  \frac{\bar{X} - m_0}{S/\sqrt{N}} \right  > t_{N-1}^{\alpha/2} \Rightarrow$ Rechazar $H_0$	Si <b>p-value</b> $< \alpha \Rightarrow$ Rechazar $H_0$
	$t_{N-1}^{\alpha/2} \Rightarrow$ valor en tabla	$\alpha =$ Riesgo de 1ª especie

$t_{N-1}^{\alpha/2}$  con  $N-1 = 9$ . En tablas para riesgo de primera especie del 5%, se obtiene el valor de t:

Contiene los valores de  $t$  tales que  $\frac{\alpha}{2} = P(t_v \geq t)$ , donde  $v$  son los Grados de Libertad

		$\alpha/2$							
		0,0005	0,001	0,005	0,01	0,025	0,05	0,1	0,2
	1	636,619	318,309	63,657	31,821	12,706	6,314	3,078	1,37
	2	31,599	22,327	9,925	6,965	4,303	2,920	1,886	1,06
	3	12,924	10,215	5,841	4,541	3,182	2,353	1,638	0,97
	4	8,610	7,173	4,604	3,747	2,776	2,132	1,533	0,94
	5	6,869	5,893	4,032	3,365	2,571	2,015	1,476	0,92
	6	5,959	5,208	3,707	3,143	2,447	1,943	1,440	0,90
	7	5,408	4,785	3,499	2,998	2,365	1,895	1,415	0,89
	8	5,041	4,501	3,355	2,896	2,306	1,860	1,397	0,88
	9	4,781	4,287	3,250	2,821	2,262	1,833	1,383	0,88
	10	4,587	4,144	3,169	2,764	2,228	1,812	1,372	0,87
	11	4,437	4,025	3,106	2,718	2,201	1,796	1,363	0,87

$$\left| \frac{\bar{X} - m_0}{S/\sqrt{N}} \right|$$

El valor de es el siguiente:  $ABS((360,89-350)/(1,78/3,16228))=3,1945$   
 $3,1945 > 2,262$  por lo que se rechaza la hipótesis nula.

Resolución según intervalo de confianza para la media:

Intervalo de confianza para la media (IC <sub>m</sub> )	
$IC_m \Rightarrow \left[ \bar{X} - t_{N-1}^{\alpha/2} \frac{S}{\sqrt{N}}, \bar{X} + t_{N-1}^{\alpha/2} \frac{S}{\sqrt{N}} \right]$	$t_{N-1}^{\alpha/2} \Rightarrow \text{valor en tabla}$

El valor del término  $t_{N-1}^{\alpha/2} \frac{S}{\sqrt{N}}$  es:  $2,262 \cdot (1,78/3,16228)=1,2732$

El resultado del intervalo es:  $[360,89-1,2732, 360,89+1,2732] = [359,6168, 362,1632]$ . El valor 350 queda fuera del intervalo, por lo que no es aceptable la hipótesis  $m_0 = 350$ .

## **Ejercicio 2:**

### **Transparencia 41 UD5 inferencia parte 1-2.pdf**

Var. aleatoria  $X$ : N° de defectos en una pieza.  $X \sim \text{Ps}(\lambda)$

$\lambda$  = número medio de defectos en una pieza

A partir de una muestra de tamaño 10, se desea contrastar la hipótesis de que el parámetro  $\lambda$  de una distribución de Poisson es 2 frente a la alternativa de que es  $> 2$ . Se aceptará  $H_0$  si la media muestral es  $\leq 2,5$  y se rechazará en caso contrario.

Recordemos que:

$X_i \sim \text{Ps}(\lambda_i)$  ;  $Y = \sum X_i$  (independientes) :

$Y \sim \text{Ps}(\lambda = \sum \lambda_i)$

Por lo tanto la variable aleatoria media muestral tendrá de media  $\lambda$  igualmente.

A) Calcular el riesgo de tipo I de este test

El riesgo de tipo I es el riesgo de rechazar ( $X > 2,5$ ) la hipótesis nula cuando esta es cierta (cuando realmente  $X \sim P_s(2)$ ), es decir  $P(X > 2,5, \text{ Con } X \sim P_s(2))$ :

## Distribuciones de Probabilidad

Distribución: Poisson

<i>Parámetros:</i>	<i>Media</i>
Dist. 1	2
Dist. 2	3
Dist. 3	4
Dist. 4	
Dist. 5	

El StatAdvisor


Área Cola Superior (>)

<i>Variable</i>	<i>Dist. 1</i>	<i>Dist. 2</i>	<i>Dist. 3</i>	<i>L</i>
2,5	0,323323	0,57681	0,761897	

Por lo tanto el riesgo de primera especie es 0,323323, o del 32,3323%

B) Calcular el riesgo de tipo II si  $\lambda$  realmente vale 3.

El de tipo I sería:  
0,57681, o del 57,681%

 Distribuciones de Probabilidad

## Distribuciones de Probabilidad

Distribución: Poisson

<i>Parámetros:</i>	<i>Media</i>
Dist. 1	2
Dist. 2	3
Dist. 3	4
Dist. 4	
Dist. 5	

### El StatAdvisor

Distribución Acumulada

Distribución: Poisson

Área Cola Inferior (<)

<i>Variable</i>	<i>Dist. 1</i>	<i>Dist. 2</i>	<i>Dist. 3</i>	<i>Dist. 4</i>	<i>Dist. 5</i>
2,5	0,676677	0,42319	0,238103		

Probabilidad de Masa (=)

<i>Variable</i>	<i>Dist. 1</i>	<i>Dist. 2</i>	<i>Dist. 3</i>	<i>Dist. 4</i>	<i>Dist. 5</i>
2,5	0,0	0,0	0,0		



El de tipo II es aceptar la hipótesis cuando es falsa, lo que sería  $P(X \leq 2,5, \text{ Con } X \sim P_s(3)) = 0,42319$

C) Calcular el riesgo de tipo II si realmente vale 4

El de tipo I sería:

0,761897, o del 76,1897%

El de tipo II es aceptar la hipótesis cuando es falsa, lo que sería  $P(X \leq 2,5, \text{ Con } X \sim P_s(4)) = 0,238103$

### **Ejercicio 3:**

La empresa **ECOLINE S.A.** ha instalado placas solares **JUNKERS** en sus últimas construcciones. Éstas conseguían almacenar más o menos unas 200 unidades de energía por término medio, con una variabilidad de 4 unidades de energía<sup>2</sup>. Como las placas aún están en garantía la empresa quiere comprobar si los

valores iniciales siguen siendo válidos o por el contrario existen cambios y hay que tomar medidas al respecto. Por ello, durante 20 días mide la energía acumulada por término medio en cinco edificaciones obteniendo las siguientes cantidades de energía media generada diariamente:

Día	1	2	3	4	5	6	7	8	9	10	11	12	13	14	15	16	17	18	19	20
Energía	196	197	198	200	202	199	198	197	195	194	198	202	199	202	201	199	200	197	196	199

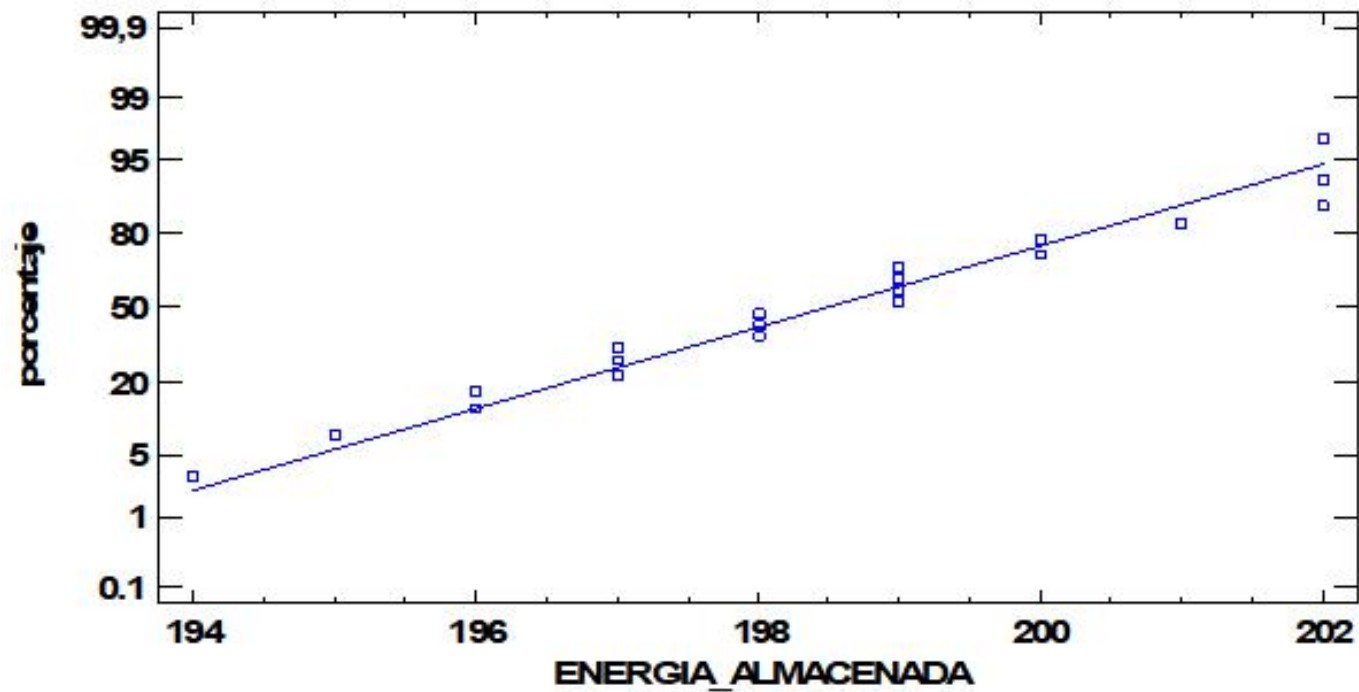
A partir de la muestra anterior la empresa ha elaborado la siguiente información estadística:

### **Resumen Estadístico para ENERGIA\_ALMACENADA**

Recuento	20
Promedio	198,45
Mediana	198,5
Varianza	5,31316
Desviación Estándar	2,30503

Sesgo	-0,0434765
Sesgo Estandarizado	-0,0793769
Curtosis	-0,596594
Curtosis Estandarizada	-0,544613

Gráfico de Probabilidad Normal



1) Identifica la variable y la población sobre la que se realiza el análisis.

La variable, es la medida de la energía media en 5 edificaciones almacenada por las placas solares y es una variable cuantitativa continua. Se toman 20 muestras (días), siendo la población las edificaciones con la instalación solar.

**2) ¿Qué hipótesis debemos asumir sobre la distribución de la variable objeto de estudio de tal manera que podamos utilizar las herramientas de inferencia analizadas en clase?**

Para poder utilizar las herramientas de inferencia es necesario que la distribución de la población tenga un comportamiento normal.

**3) Teniendo en cuenta la información estadística anterior, ¿consideras que se cumplen las hipótesis del apartado anterior? Justifica razonadamente tu respuesta.**

Tras observar el PPN, se puede asumir que la distribución de la población tiene un comportamiento normal. Los puntos se ajustan a la recta razonablemente y los parámetros de asimetría y curtosis también están dentro de los valores que corresponden a una distribución normal.

**4) Identifica los parámetros poblacionales e indica el valor que éstos tenían cuando, inicialmente, se instalaron las placas solares. ¿Conoce actualmente la empresa el valor de esos parámetros poblacionales? Si tu respuesta es “no” proporciona una estimación puntual para ellos.**

Según el enunciado, la empresa indica que la media declarada por Junkers para el producto instalado debía ser de 200 unidades de energía y una desviación típica de 4 unidades de energía. En caso de no tener dicha

información un estimador podría haber sido la media total de todas las muestras, que sería 198,45.

**5) Calcula los intervalos de confianza para la media poblacional con riesgos de primera especie del 10, 5 y 1%.**

Fórmula para calcular los intervalos:

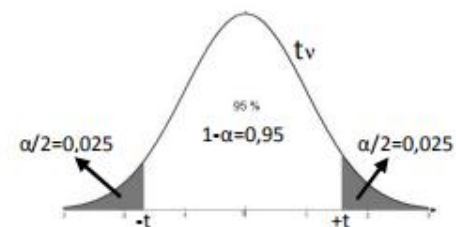
Intervalo de confianza para la media (IC <sub>m</sub> )	
$IC_m \Rightarrow \left[ \bar{x} - t_{N-1}^{\alpha/2} \frac{S}{\sqrt{N}}, \bar{x} + t_{N-1}^{\alpha/2} \frac{S}{\sqrt{N}} \right]$	$t_{N-1}^{\alpha/2} \Rightarrow$ valor en tabla

Donde  $t_{n-1}^{\alpha/2}$  = valor de un t de Student con n-1 grados de libertad que deja a la derecha alfa partido por dos. En el caso del intervalo con una confianza del 95%, el alfa es igual al 5%, por tanto, deberíamos buscar el valor de una t-Student con 19 grados de libertad que deja a la derecha una probabilidad de 0,025. En este caso,  $t_{19}^{0,005}=2,861$ ,  $t_{19}^{0,025}=2,093$ ,  $t_{19}^{0,05}=1,729$ .



## Distribución t de Student

Contiene los valores de  $t$  tales que  $\frac{\alpha}{2} = P(t_v \geq t)$ , donde  $v$  son los Grados de Libertad



		$\alpha/2$												
		0,0005	0,001	0,005	0,01	0,025	0,05	0,1	0,2	0,25	0,3	0,4	0,45	0,475
v grados de libertad	1	636,619	318,309	62,657	31,821	12,706	6,314	3,078	1,376	1,000	0,727	0,325	0,158	0,079
	2	31,599	22,327	9,925	6,965	4,303	2,306	1,886	1,061	0,816	0,617	0,289	0,142	0,071
	3	12,924	10,215	5,841	4,541	3,182	2,353	1,638	0,978	0,765	0,584	0,277	0,137	0,068
	4	8,610	7,173	4,604	3,747	2,776	2,132	1,533	0,941	0,741	0,569	0,271	0,134	0,067
	5	6,869	5,893	4,032	3,365	2,571	2,015	1,476	0,920	0,727	0,559	0,267	0,132	0,066
	6	5,959	5,208	3,707	3,143	2,447	1,943	1,440	0,906	0,718	0,553	0,265	0,131	0,065
	7	5,408	4,785	3,499	2,998	2,365	1,895	1,415	0,896	0,711	0,549	0,263	0,130	0,065
	8	5,041	4,501	3,355	2,896	2,306	1,860	1,397	0,889	0,706	0,546	0,262	0,130	0,065
	9	4,781	4,297	3,250	2,821	2,262	1,833	1,383	0,883	0,703	0,543	0,261	0,129	0,064
	10	4,587	4,144	3,169	2,764	2,228	1,812	1,372	0,879	0,700	0,542	0,260	0,129	0,064
	11	4,437	4,025	3,106	2,718	2,201	1,796	1,363	0,876	0,697	0,540	0,260	0,129	0,064
	12	4,318	3,930	3,055	2,681	2,179	1,782	1,356	0,873	0,695	0,539	0,259	0,128	0,064
	13	4,221	3,852	3,012	2,650	2,160	1,771	1,350	0,870	0,694	0,538	0,259	0,128	0,064
	14	4,140	3,787	2,977	2,624	2,145	1,761	1,345	0,868	0,692	0,537	0,258	0,128	0,064
	15	4,073	3,733	2,947	2,602	2,131	1,753	1,341	0,866	0,691	0,536	0,258	0,128	0,064
	16	4,015	3,686	2,921	2,583	2,119	1,746	1,337	0,865	0,690	0,535	0,258	0,128	0,064
	17	3,965	3,646	2,898	2,567	2,109	1,740	1,333	0,863	0,689	0,534	0,257	0,128	0,064
	18	3,922	3,610	2,878	2,552	2,101	1,734	1,330	0,862	0,688	0,534	0,257	0,127	0,064
	19	3,883	3,579	2,861	2,539	2,093	1,729	1,328	0,861	0,688	0,533	0,257	0,127	0,064
	20	3,850	3,552	2,848	2,528	2,088	1,725	1,325	0,860	0,687	0,533	0,257	0,127	0,063
	21	3,819	3,527	2,831	2,518	2,080	1,721	1,323	0,859	0,686	0,532	0,257	0,127	0,063
	22	3,792	3,505	2,819	2,508	2,074	1,717	1,321	0,858	0,686	0,532	0,256	0,127	0,063

Con los valores correspondientes a  $t_{19}$  para cada nivel de confianza se obtienen los siguientes intervalos:

Intervalos de confianza del 90,0% para la media:  $198,45 \pm 0,891232$   
[197,559; 199,341]

Intervalos de confianza del 95,0% para la media:  $198,45 \pm 1,07879$   
[197,371; 199,529]

Intervalos de confianza del 99,0% para la media:  $198,45 \pm 1,47459$   
[196,975; 199,925]

Estos intervalos se comprueban fácilmente trabajando con el SG, y en el panel se pueden modificar los valores de nivel de confianza.

**Intervalos de Confianza para Col\_1**

Intervalos de confianza del 95,0% para la media: 198,45 +/- 1,07879 [197,371; 199,529]

Intervalos de confianza del 95,0% para la desviación estándar: [1,75295; 3,36666]

**El StatAdvisor**

Este panel muestra los intervalos de confianza del 95,0% para la media y la desviación estándar de Col\_1. La interpretación clásica de estos intervalos es que, en muestreos repetidos, estos intervalos contendrán la media verdadera ó la desviación estándar verdadera de la población de la que fueron extraídas las muestras, el 95,0% de las veces. En términos prácticos, puede establecerse con 95,0% de confianza, que la media verdadera de Col\_1 se encuentra en algún lugar entre 197,371 y 199,529, en tanto que la desviación estándar verdadera está en algún lugar entre 1,75295 y 3,36666.

Ambos intervalos asumen que la población de la cual proviene la muestra puede representarse por la distribución normal. Mientras que el intervalo de confianza para la media es bastante robusto y no muy sensible a violaciones de este supuesto, los intervalos de confianza para la desviación estándar son muy sensibles. Si los datos no provienen de una distribución normal, el intervalo para la desviación estándar puede ser incorrecto. Para verificar si los datos provienen de una distribución normal, seleccione Resumen Estadístico de la lista de Opciones Tabulares, ó escoja Gráfica de Probabilidad Normal de la lista de Opciones Gráficas.

**6) ¿Qué interpretación tiene cada uno de esos intervalos?.**

Estos intervalos representan los valores que para esa muestra dada (media muestral, desviación típica muestral y tamaño de muestra), recogen con el nivel de confianza indicado (99, 95, 90%) el valor de la media poblacional  $\mu$ .

**7) ¿Está dentro del intervalo el valor inicial de la media poblacional?. En base a la respuesta dada, ¿qué crees que debería hacer la empresa?, ¿se estaría equivocando?.**

En ningún caso, el valor de 200 entra en los intervalos de confianza, por lo que la empresa estaba equivocada en su afirmación y se debería revisar el proceso si desea alcanzar el valor afirmado (se debe incrementar el valor poblacional mediante mejoras en la calidad de paneles, de regulación, de instalación....).

**8) Realiza ahora el cálculo del estadístico de contraste para los siguientes tests de hipótesis y verifica los resultados obtenidos con los intervalos de confianza para la media. Razona el cumplimiento de las hipótesis alternativas.**

Tests de hipótesis 1	Tests de hipótesis 2	Tests de hipótesis 3
$H_0: m=200$ $H_1: m \neq 200$	$H_0: m \leq 200$ $H_1: m > 200$	$H_0: m \geq 200$ $H_1: m < 200$

Test de hipótesis para la media (Test t)		
$H_0: m = m_0$ $H_1: m \neq m_0$	Si $\left  \frac{\bar{X} - m_0}{S/\sqrt{N}} \right  \leq t_{N-1}^{\alpha/2} \Rightarrow$ Aceptar $H_0$	Si <b>p-value</b> $\geq \alpha \Rightarrow$ Aceptar $H_0$
	Si $\left  \frac{\bar{X} - m_0}{S/\sqrt{N}} \right  > t_{N-1}^{\alpha/2} \Rightarrow$ Rechazar $H_0$	Si <b>p-value</b> $< \alpha \Rightarrow$ Rechazar $H_0$
	$t_{N-1}^{\alpha/2} \Rightarrow$ valor en tabla	$\alpha =$ Riesgo de 1ª especie

## PRIMER TEST DE HIPÓTESIS

El valor del estadístico  $\left| \frac{\bar{X} - m_0}{S/\sqrt{N}} \right|$  para el contraste del test de hipótesis tiene el siguiente valor:  $ABS((198,45-200)/(2,30503/4,4721))=3,00725$ .



Como ya se vió,  $t^{0,025}_{19}=2,093$ .

El primer test de hipótesis es un test de hipótesis simple donde  $3,00725 > 2,093$  por lo que se rechaza  $H_0$  y se acepta  $H_1$ .

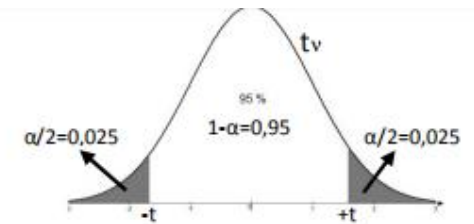
Para calcular el p-value, debemos conceptualmente suponer que tuviéramos un valor  $\alpha$  que dejara en el límite de aceptación el valor obtenido con el valor

del estadístico de contraste  $\left| \frac{\bar{X} - m_0}{S/\sqrt{N}} \right|$  para nuestra hipótesis y muestra ( $\left| \frac{\bar{X} - m_0}{S/\sqrt{N}} \right| = 3,00725$ ). Voy a la tabla de dos colas y busco este valor en  $t_{19}$  para y el p-value/2 (valor  $\alpha/2$  que fuera límite de aceptación) es aproximadamente 0,0035:

## Distribución t de Student

Contiene los valores de  $t$  tales que  $\frac{\alpha}{2} = P(t_v \geq t)$ , donde  $v$  son los Grados de Libertad

0,0035 aprox



		$\alpha/2$												
		0,0005	0,001	0,005	0,01	0,025	0,05	0,1	0,2	0,25	0,3	0,4	0,45	0,475
v grados de libertad	1	636,619	318,309	12,706	31,821	12,706	6,314	3,078	1,376	1,000	0,727	0,325	0,158	0,079
	2	31,599	22,327	9,925	6,965	4,303	2,920	1,886	1,061	0,816	0,617	0,289	0,142	0,071
	3	12,924	10,215	5,841	4,541	3,182	2,353	1,638	0,978	0,765	0,584	0,277	0,137	0,068
	4	8,610	7,173	4,604	3,747	2,776	2,132	1,533	0,941	0,741	0,569	0,271	0,134	0,067
	5	6,869	5,893	4,032	3,365	2,571	2,015	1,476	0,920	0,727	0,559	0,267	0,132	0,066
	6	5,959	5,208	3,707	3,143	2,447	1,943	1,440	0,906	0,718	0,553	0,265	0,131	0,065
	7	5,408	4,785	3,499	2,998	2,365	1,895	1,415	0,896	0,711	0,549	0,263	0,130	0,065
	8	5,041	4,501	3,355	2,896	2,306	1,860	1,397	0,889	0,706	0,546	0,262	0,130	0,065
	9	4,781	4,297	3,250	2,821	2,262	1,833	1,383	0,883	0,703	0,543	0,261	0,129	0,064
	10	4,587	4,144	3,169	2,764	2,228	1,812	1,372	0,879	0,700	0,542	0,260	0,129	0,064
	11	4,437	4,025	3,106	2,718	2,201	1,796	1,363	0,876	0,697	0,540	0,260	0,129	0,064
	12	4,318	3,930	3,055	2,681	2,179	1,782	1,356	0,873	0,695	0,539	0,259	0,128	0,064
	13	4,221	3,852	3,012	2,650	2,160	1,771	1,350	0,870	0,694	0,538	0,259	0,128	0,064
	14	4,140	3,787	2,977	2,624	2,145	1,761	1,345	0,868	0,692	0,537	0,258	0,128	0,064
	15	4,073	3,733	2,947	2,602	2,131	1,753	1,341	0,866	0,691	0,536	0,258	0,128	0,064
	16	4,015	3,686	2,921	2,583	2,120	1,746	1,337	0,865	0,690	0,535	0,258	0,128	0,064
	17	3,965	3,646	2,898	2,567	2,110	1,740	1,333	0,863	0,689	0,534	0,257	0,128	0,064
	18	3,922	3,612	2,878	2,552	2,101	1,734	1,330	0,862	0,688	0,534	0,257	0,127	0,064
	19	3,883	3,579	2,861	2,539	2,093	1,729	1,328	0,861	0,688	0,533	0,257	0,127	0,064
	20	3,850	3,552	2,845	2,528	2,086	1,725	1,325	0,860	0,687	0,533	0,257	0,127	0,063
	21	3,819	3,527	2,831	2,518	2,080	1,721	1,323	0,859	0,686	0,532	0,257	0,127	0,063
	22	3,792	3,505	2,819	2,508	2,074	1,717	1,321	0,858	0,686	0,532	0,256	0,127	0,063
	23	3,768	3,485	2,807	2,500	2,069	1,714	1,319	0,858	0,685	0,532	0,256	0,127	0,063
	24	3,745	3,467	2,797	2,492	2,064	1,711	1,318	0,857	0,685	0,531	0,256	0,127	0,063



Por lo tanto el p-value es para este test de hipótesis aproximadamente 0,007. Este valor es inferior a  $\alpha=0,05$  y por tanto se rechaza la hipótesis, se rechaza  $H_0$  y se acepta  $H_1$ .

Los tests de hipótesis pueden igualmente realizarse en SG:

**Prueba de Hipótesis para Col\_1**

Media Muestral = 198,45

Mediana Muestral = 198,5

Desviación Estándar de la Muestra = 2,30503

Prueba t

Hipótesis Nula: media = 200,0

Alternativa: no igual

|

Estadístico t = -3,00726

Valor-P = 0,00724458

Se rechaza la hipótesis nula para alfa = 0,05.

## SEGUNDO TEST DE HIPÓTESIS

El segundo test de hipótesis es una hipótesis compuesta donde  $H_0$  supone que  $m \leq 200$  y por tanto incluye  $m=200$ , pero también admite todos los valores a la

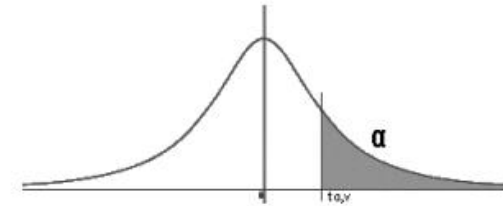
$$\left| \frac{\bar{X} - m_0}{S/\sqrt{N}} \right|$$

izquierda de 200. Sabemos que se comporta como una  $t_{N-1}$ , pero en este caso no recurrimos al valor absoluto porque no busco el intervalo que me da el 95% centrado alrededor de la media de dicha distribución, sino que debo tomar ese 95% tomando los valores que incluyen la cola izquierda de

$\frac{\bar{X} - m_0}{S/\sqrt{N}}$  (se acepta la cola izquierda del estadístico  $\frac{\bar{X} - m_0}{S/\sqrt{N}}$  pues serán válidos los valores de media muestral hacia la izquierda si los valores de la media poblacional válidos en la hipótesis incluyen los valores a la izquierda de

$m_0=200$ ) o lo que es lo mismo es válida también la cola izquierda de  $t_{N-1}$  (pues  $t$  se comporta como el estadístico y dichos valores son válidos según la hipótesis). De otro modo, la hipótesis no se ciñe a un intervalo centrado (por ser una igualdad en el caso de hipótesis simple) que recoge el 95% de probabilidad de  $t_{19}$ , sino que también recoge los valores inferiores, añadiendo dentro de lo válido la cola izquierda (la cola derecha, de rechazo, debería alcanzar el 5%, en este caso no se reparte entre dos colas).

El valor de nuestro estadístico  $\frac{\bar{X} - m_0}{S/\sqrt{N}}$  a partir de nuestra hipótesis y muestra debería quedar por debajo del valor límite derecho de aceptación que marca dicha cola derecha (para un 5% de riesgo de primera especie o alfa). Dicho valor se obtiene de la tabla de  $t_{19}$  para una sola cola:



**Distribución t de Student** Contiene los valores de  $t$  tales que  $\alpha = P(t_{\nu}^2 \geq t)$ , donde  $\nu$  son los Grados de Libertad

$\nu/\alpha$	0,0005	0,001	0,005	0,01	0,025	0,05	0,1	0,2	0,25	0,3	0,4	0,45	0,475
1	636,619	318,309	63,657	31,821	12,706	6,314	3,078	1,376	1,000	0,727	0,325	0,158	0,079
2	31,599	22,327	9,925	6,965	4,303	2,920	1,886	1,061	0,816	0,617	0,289	0,142	0,071
3	12,924	10,215	5,841	4,541	3,182	2,353	1,638	0,978	0,765	0,584	0,277	0,137	0,068
4	8,610	7,173	4,604	3,747	2,776	2,132	1,533	0,941	0,741	0,569	0,271	0,134	0,067
5	6,869	5,893	4,032	3,365	2,571	2,015	1,476	0,920	0,727	0,559	0,267	0,132	0,066
6	5,959	5,208	3,707	3,143	2,447	1,943	1,440	0,906	0,718	0,553	0,265	0,131	0,065
7	5,408	4,785	3,499	2,998	2,365	1,895	1,415	0,896	0,711	0,549	0,263	0,130	0,065
8	5,041	4,501	3,355	2,896	2,306	1,860	1,397	0,889	0,706	0,546	0,262	0,130	0,065
9	4,781	4,297	3,250	2,821	2,262	1,833	1,383	0,883	0,703	0,543	0,261	0,129	0,064
10	4,587	4,144	3,169	2,764	2,228	1,812	1,372	0,879	0,700	0,542	0,260	0,129	0,064
11	4,437	4,025	3,106	2,718	2,201	1,796	1,363	0,876	0,697	0,540	0,260	0,129	0,064
12	4,318	3,930	3,055	2,681	2,179	1,782	1,356	0,873	0,695	0,539	0,259	0,128	0,064
13	4,221	3,852	3,012	2,650	2,160	1,771	1,350	0,870	0,694	0,538	0,259	0,128	0,064
14	4,140	3,787	2,977	2,624	2,145	1,761	1,345	0,868	0,692	0,537	0,258	0,128	0,064
15	4,073	3,733	2,947	2,602	2,131	1,753	1,341	0,866	0,691	0,536	0,258	0,128	0,064
16	4,015	3,686	2,921	2,583	2,120	1,746	1,337	0,865	0,690	0,535	0,258	0,128	0,064
17	3,965	3,646	2,898	2,567	2,110	1,740	1,333	0,863	0,689	0,534	0,257	0,128	0,064
18	3,922	3,610	2,878	2,552	2,101	1,734	1,330	0,862	0,688	0,534	0,257	0,127	0,064
19	3,885	3,578	2,861	2,539	2,093	1,729	1,328	0,861	0,688	0,533	0,257	0,127	0,064
20	3,850	3,552	2,845	2,528	2,086	1,724	1,325	0,860	0,687	0,533	0,257	0,127	0,063
21	3,816	3,527	2,831	2,519	2,080	1,721	1,322	0,859	0,686	0,532	0,257	0,127	0,063

El valor es 1,729 que queda bien a la derecha del valor del estadístico de

contraste  $\frac{\bar{X} - m_0}{S/\sqrt{N}}$  para nuestra hipótesis y muestra que resulta ser  $\frac{\bar{X} - m_0}{S/\sqrt{N}} =$

-3,00725. Esto implica que el valor -3,00725 es perfectamente probable pues queda bien por debajo (y por tanto en el 95 % de los datos válidos para nuestra hipótesis) de 1,729. Por lo tanto se aceptaría  $H_0$ .

Para calcular el p-value, debemos conceptualmente suponer que tuviéramos un valor  $\alpha$  que dejara en el límite de aceptación el valor obtenido para

nuestra muestra e hipótesis del valor del estadístico de contraste  $\frac{\bar{X} - m_0}{S/\sqrt{N}}$  (  $\frac{\bar{X} - m_0}{S/\sqrt{N}}$  = -3,00725). Es decir, el p-value equivale al valor de probabilidad que marca el estadístico de contraste para muestra e hipótesis, si dicho valor fuera el límite de aceptación y por tanto el p-value el riesgo de primera especie establecido. Calculando por tanto que valor de probabilidad (área de la función de densidad) se tiene de que el valor de la  $t_{19}$  quede a la derecha (cola derecha) del valor del estadístico obtenido -3,00725 , estamos calculando dicho p-value.

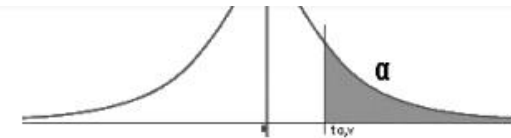


Buscamos este valor en tablas, vemos que la gráfica solo muestra valores de cola a la derecha de la media (que es cero) para la t de student.



0,004 aprox.

**Distribución t de Student** Contiene los valores de  $t$  tales que  $\alpha = P(t_v^2 \geq t)$ , donde  $v$  son los Grados de Libertad



$v/\alpha$	0,0005	0,001	0,005	0,01	0,025	0,05	0,1	0,2	0,25	0,3	0,4	0,45	0,475
1	636,619			31,821	12,706	6,314	3,078	1,376	1,000	0,727	0,325	0,158	0,079
2	31,599	22,327	9,925	6,965	4,303	2,920	1,886	1,061	0,816	0,617	0,289	0,142	0,071
3	12,924	10,215	5,841	4,541	3,182	2,353	1,638	0,978	0,765	0,584	0,277	0,137	0,068
4	8,610	7,173	4,604	3,747	2,776	2,132	1,533	0,941	0,741	0,569	0,271	0,134	0,067
5	6,869	5,893	4,032	3,365	2,571	2,015	1,476	0,920	0,727	0,559	0,267	0,132	0,066
6	5,959	5,208	3,707	3,143	2,447	1,943	1,440	0,906	0,718	0,553	0,265	0,131	0,065
7	5,408	4,785	3,499	2,998	2,365	1,895	1,415	0,896	0,711	0,549	0,263	0,130	0,065
8	5,041	4,501	3,355	2,896	2,306	1,860	1,397	0,889	0,706	0,546	0,262	0,130	0,065
9	4,781	4,297	3,250	2,821	2,262	1,833	1,383	0,883	0,703	0,543	0,261	0,129	0,064
10	4,587	4,144	3,169	2,764	2,228	1,812	1,372	0,879	0,700	0,542	0,260	0,129	0,064
11	4,437	4,025	3,106	2,718	2,201	1,796	1,363	0,876	0,697	0,540	0,260	0,129	0,064
12	4,318	3,930	3,055	2,681	2,179	1,782	1,356	0,873	0,695	0,539	0,259	0,128	0,064
13	4,221	3,852	3,012	2,650	2,160	1,771	1,350	0,870	0,694	0,538	0,259	0,128	0,064
14	4,140	3,787	2,977	2,624	2,145	1,761	1,345	0,868	0,692	0,537	0,258	0,128	0,064
15	4,073	3,733	2,947	2,602	2,131	1,753	1,341	0,866	0,691	0,536	0,258	0,128	0,064
16	4,015	3,686	2,921	2,583	2,120	1,746	1,337	0,865	0,690	0,535	0,258	0,128	0,064
17	3,965	3,646	2,898	2,567	2,110	1,740	1,333	0,863	0,689	0,534	0,257	0,128	0,064
18	3,922	3,610	2,878	2,552	2,101	1,734	1,330	0,862	0,688	0,534	0,257	0,127	0,064
19	3,883	3,579	2,861	2,539	2,093	1,729	1,328	0,861	0,688	0,533	0,257	0,127	0,064
20	3,850	3,552	2,847	2,528	2,086	1,725	1,325	0,860	0,687	0,533	0,257	0,127	0,063
21	3,819	3,528	2,831	2,518	2,080	1,721	1,323	0,859	0,686	0,532	0,257	0,127	0,063
22	3,792	3,506	2,816	2,508	2,074	1,717	1,321	0,858	0,686	0,532	0,256	0,127	0,063
23	3,768	3,485	2,807	2,500	2,069	1,714	1,319	0,858	0,685	0,532	0,256	0,127	0,063

Al ser perfectamente simétrica buscamos el valor que queda a la derecha de 3,00725 y nos da un valor aproximado de 0,004. Por lo tanto el área a la izquierda será aproximadamente 0,996. Invirtiendo la gráfica sobre su eje de simetría, vemos que acabamos de calcular el área a la derecha de -3,00725 que es el valor buscado. Por lo tanto p-value será aproximadamente 0,996 muy superior a 0,05 y por lo tanto se aceptaría  $H_0$ .

### **Prueba de Hipótesis para Col\_1**

Media Muestral = 198,45

Mediana Muestral = 198,5

Desviación Estándar de la Muestra = 2,30503

#### Prueba t

Hipótesis Nula: media = 200,0

Alternativa: mayor que

Estadístico t = -3,00726

Valor-P = 0,996378

No se rechaza la hipótesis nula para alfa = 0,05.

Recuerda que en un contraste de hipótesis entre las dos hipótesis se tiene que abarcar todas las posibilidades, con respecto al parámetro poblacional sobre el que se está planteando el contraste. Así pues, aunque se escribe (por ejemplo en SG) que la hipótesis nula es que  $m=200$  frente a la hipótesis



alternativa de  $m > 200$ . Sin embargo, realmente se está contrastando la hipótesis nula  $m \leq 200$  frente a  $m > 200$  aunque no se explicita.

### **TERCER TEST DE HIPÓTESIS (análogo al segundo).**

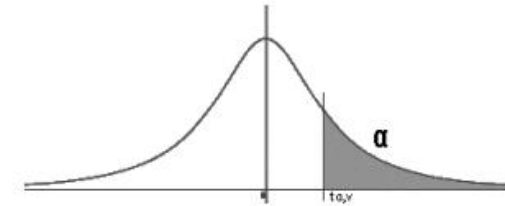
El tercer test de hipótesis es una hipótesis compuesta donde  $H_0$  supone que  $m \geq 200$  y por tanto incluye  $m = 200$ , pero también admite todos los valores a la

derecha de 200. Sabemos que  $\left| \frac{\bar{X} - m_0}{S/\sqrt{N}} \right|$  se comporta como una  $t_{N-1}$ , pero en este caso no recurrimos al valor absoluto porque no busco el intervalo que me da el 95% centrado alrededor de la media de dicha distribución, sino que

debo tomar ese 95% tomando los valores que incluye la cola derecha de  $\frac{\bar{X} - m_0}{S/\sqrt{N}}$

o lo que es lo mismo la cola derecha de  $t_{N-1}$  (pues dichos valores son válidos según la hipótesis). Es decir, la hipótesis no se ciñe a un intervalo centrado (por ser una igualdad en el caso de hipótesis simple) que recoge el 95% de probabilidad de  $t_{19}$ , sino que también recoge los valores superiores, hay que añadir la cola derecha como zona de aceptación. En este caso la cola izquierda de rechazo debería recoger el 5% de área o probabilidad, no se reparte entre dos colas.

Para aceptar la hipótesis, el valor de nuestro estadístico debería quedar por encima del valor límite que marca dicha cola izquierda (para un 5% de riesgo de primera especie o alfa), que equivaldría a un 95% de área a la derecha de dicho valor límite. Buscamos dicho valor en la tabla de  $t_{19}$  para una sola cola, pero vemos que solo se muestran valores de área a la derecha de la media, por lo que sucede algo similar al caso anterior: buscamos el valor de  $t$  que deja el 5% a su derecha (cola derecha) e invirtiendo la curva, el valor leído será el mismo con el signo negativo ( $t$  simétrica entorno a 0).



**Distribución t de Student** Contiene los valores de  $t$  tales que  $\alpha = P(t_{\nu}^2 \geq t)$ , donde  $\nu$  son los Grados de Libertad

$\nu/\alpha$	0,0005	0,001	0,005	0,01	0,025	0,05	0,1	0,2	0,25	0,3	0,4	0,45	0,475
1	636,619	318,309	63,657	31,821	12,706	6,314	3,078	1,376	1,000	0,727	0,325	0,158	0,079
2	31,599	22,327	9,925	6,965	4,303	2,920	1,886	1,061	0,816	0,617	0,289	0,142	0,071
3	12,924	10,215	5,841	4,541	3,182	2,353	1,638	0,978	0,765	0,584	0,277	0,137	0,068
4	8,610	7,173	4,604	3,747	2,776	2,132	1,533	0,941	0,741	0,569	0,271	0,134	0,067
5	6,869	5,893	4,032	3,365	2,571	2,015	1,476	0,920	0,727	0,559	0,267	0,132	0,066
6	5,959	5,208	3,707	3,143	2,447	1,943	1,440	0,906	0,718	0,553	0,265	0,131	0,065
7	5,408	4,785	3,499	2,998	2,365	1,895	1,415	0,896	0,711	0,549	0,263	0,130	0,065
8	5,041	4,501	3,355	2,896	2,306	1,860	1,397	0,889	0,706	0,546	0,262	0,130	0,065
9	4,781	4,297	3,250	2,821	2,262	1,833	1,383	0,883	0,703	0,543	0,261	0,129	0,064
10	4,587	4,144	3,169	2,764	2,228	1,812	1,372	0,879	0,700	0,542	0,260	0,129	0,064
11	4,437	4,025	3,106	2,718	2,201	1,796	1,363	0,876	0,697	0,540	0,260	0,129	0,064
12	4,318	3,930	3,055	2,681	2,179	1,782	1,356	0,873	0,695	0,539	0,259	0,128	0,064
13	4,221	3,852	3,012	2,650	2,160	1,771	1,350	0,870	0,694	0,538	0,259	0,128	0,064
14	4,140	3,787	2,977	2,624	2,145	1,761	1,345	0,868	0,692	0,537	0,258	0,128	0,064
15	4,073	3,733	2,947	2,602	2,131	1,753	1,341	0,866	0,691	0,536	0,258	0,128	0,064
16	4,015	3,686	2,921	2,583	2,120	1,746	1,337	0,865	0,690	0,535	0,258	0,128	0,064
17	3,965	3,646	2,898	2,567	2,110	1,740	1,333	0,863	0,689	0,534	0,257	0,128	0,064
18	3,922	3,610	2,878	2,552	2,101	1,734	1,330	0,862	0,688	0,534	0,257	0,127	0,064
19	3,886	3,579	2,861	2,539	2,093	1,729	1,328	0,861	0,688	0,533	0,257	0,127	0,064
20	3,850	3,552	2,845	2,528	2,086	1,725	1,325	0,860	0,687	0,533	0,257	0,127	0,063
21	3,816	3,527	2,831	2,519	2,080	1,721	1,322	0,859	0,686	0,532	0,257	0,127	0,063

El valor límite es -1,729 que queda bien a la derecha del valor del estadístico

de contraste  $\frac{\bar{X} - m_0}{S/\sqrt{N}}$  para nuestra hipótesis y muestra que resulta ser  $\frac{\bar{X} - m_0}{S/\sqrt{N}} =$

-3,00725. Esto implica que el valor -3,00725 es muy improbable que suceda pues queda muy a la izquierda de -1,729(y por tanto en el 5 % de los datos válidos para nuestra hipótesis). Por lo tanto se rechazaría  $H_0$ .

Para calcular el p-value, debemos conceptualmente suponer que tuviéramos un valor  $\alpha$  que dejara en el límite de aceptación el valor obtenido con el valor

del estadístico de contraste  $\frac{\bar{X} - m_0}{S/\sqrt{N}}$  para nuestra hipótesis y muestra ( $\frac{\bar{X} - m_0}{S/\sqrt{N}} = -3,00725$ ). Es decir, el p-value equivale al valor que sería el límite de aceptación si él mismo fuera el riesgo de primera especie establecido. Calculando por tanto que valor de probabilidad (área de la función de densidad) que hay de que el valor de la  $t_{19}$  quede a la izquierda (cola izquierda) del valor -3,00725 , estamos calculando dicho p-value.

El caso es análogo al previo, buscamos este valor en tablas, vemos que la gráfica solo muestra valores de cola a la derecha de la media que es cero para la  $t$  de student.





0,004 aprox.

**Distribución t de Student** Contiene los valores de  $t$  tales que  $\alpha = P(t_v^2 \geq t)$ , donde  $v$  son los Grados de Libertad



$v/\alpha$	0,0005	0,001	0,005	0,01	0,025	0,05	0,1	0,2	0,25	0,3	0,4	0,45	0,475
1	636,619			31,821	12,706	6,314	3,078	1,376	1,000	0,727	0,325	0,158	0,079
2	31,599	22,327	9,925	6,965	4,303	2,920	1,886	1,061	0,816	0,617	0,289	0,142	0,071
3	12,924	10,215	5,841	4,541	3,182	2,353	1,638	0,978	0,765	0,584	0,277	0,137	0,068
4	8,610	7,173	4,604	3,747	2,776	2,132	1,533	0,941	0,741	0,569	0,271	0,134	0,067
5	6,869	5,893	4,032	3,365	2,571	2,015	1,476	0,920	0,727	0,559	0,267	0,132	0,066
6	5,959	5,208	3,707	3,143	2,447	1,943	1,440	0,906	0,718	0,553	0,265	0,131	0,065
7	5,408	4,785	3,499	2,998	2,365	1,895	1,415	0,896	0,711	0,549	0,263	0,130	0,065
8	5,041	4,501	3,355	2,896	2,306	1,860	1,397	0,889	0,706	0,546	0,262	0,130	0,065
9	4,781	4,297	3,250	2,821	2,262	1,833	1,383	0,883	0,703	0,543	0,261	0,129	0,064
10	4,587	4,144	3,169	2,764	2,228	1,812	1,372	0,879	0,700	0,542	0,260	0,129	0,064
11	4,437	4,025	3,106	2,718	2,201	1,796	1,363	0,876	0,697	0,540	0,260	0,129	0,064
12	4,318	3,930	3,055	2,681	2,179	1,782	1,356	0,873	0,695	0,539	0,259	0,128	0,064
13	4,221	3,852	3,012	2,650	2,160	1,771	1,350	0,870	0,694	0,538	0,259	0,128	0,064
14	4,140	3,787	2,977	2,624	2,145	1,761	1,345	0,868	0,692	0,537	0,258	0,128	0,064
15	4,073	3,733	2,947	2,602	2,131	1,753	1,341	0,866	0,691	0,536	0,258	0,128	0,064
16	4,015	3,686	2,921	2,583	2,120	1,746	1,337	0,865	0,690	0,535	0,258	0,128	0,064
17	3,965	3,646	2,898	2,567	2,110	1,740	1,333	0,863	0,689	0,534	0,257	0,128	0,064
18	3,922	3,610	2,879	2,552	2,101	1,734	1,330	0,862	0,688	0,534	0,257	0,127	0,064
19	3,883	3,579	2,861	2,539	2,093	1,729	1,328	0,861	0,688	0,533	0,257	0,127	0,064
20	3,850	3,552	2,845	2,528	2,086	1,725	1,325	0,860	0,687	0,533	0,257	0,127	0,063
21	3,819	3,527	2,831	2,518	2,080	1,721	1,323	0,859	0,686	0,532	0,257	0,127	0,063
22	3,792	3,503	2,817	2,508	2,074	1,717	1,321	0,858	0,686	0,532	0,256	0,127	0,063
23	3,768	3,485	2,807	2,500	2,069	1,714	1,319	0,858	0,685	0,532	0,256	0,127	0,063

3,00725

Al ser perfectamente simétrica buscamos el valor que queda a la derecha de 3,00725 y nos da un valor aproximado de 0,004. Por lo tanto, por la simetría

de  $t$  entorno a su media de valor 0, el área a la izquierda del valor -3,00725 será igualmente 0,04. Por lo tanto  $p$ -value es aproximadamente 0,004 muy inferior a 0,05 y por lo tanto se rechaza  $H_0$ .

---

### Prueba de Hipótesis para Col\_1

Media Muestral = 198,45

Mediana Muestral = 198,5

Desviación Estándar de la Muestra = 2,30503

### Prueba t

Hipótesis Nula: media = 200,0

Alternativa: menor que

Estadístico  $t = -3,00726$

Valor-P = 0,00362229

Se rechaza la hipótesis nula para  $\alpha = 0,05$ .

**9) Si la empresa ha fijado un alfa o nivel de significación o probabilidad de cometer un error de tipo 1 igual al 5%, interpreta cada uno de los contrastes y justifica si EcoLine tiene evidencia desde el punto de vista estadístico para reclamar a Junkers.**

Como ya se ha comentado, el contraste que resulta verdadero es el que indica que la media poblacional es menor o igual que 200, pero además el contraste de igualdad ha fallado. La conclusión es que la media es menor que 200. La posibilidad de reclamar a Junkers está fundada, siempre que la pérdida de rendimiento respecto a lo declarado sea imputable a la calidad de los paneles e instalación atribuible a Junkers, pues las causas podrían ser una mala instalación si la hizo ECO LINE S.A. o un mal mantenimiento, etc.



10) La empresa desea analizar la variabilidad (varianza declarada de 4) en la cantidad de energía que las placas pueden almacenar, para ello, calcula intervalos de confianza de la desviación típica para 1, 5 y 10% de riesgos de error de primera especie.

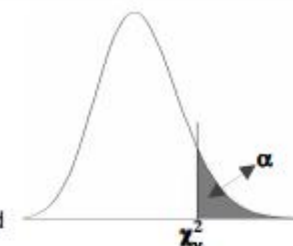
Intervalo de confianza para la varianza	
$IC_{\sigma^2} \Rightarrow \left[ \frac{(N-1)S^2}{g_2}, \frac{(N-1)S^2}{g_1} \right]$	$g_1 \text{ y } g_2 \Rightarrow \text{valores en tabla}$ $g_1 / P(\chi_{N-1}^2 > g_1) = 1 - \frac{\alpha}{2} \quad \text{y} \quad g_2 / P(\chi_{N-1}^2 > g_2) = \frac{\alpha}{2}$
Test de hipótesis para la varianza mediante IC	
$\left. \begin{array}{l} H_0: \sigma^2 = \sigma_0^2 \\ H_1: \sigma^2 \neq \sigma_0^2 \end{array} \right\}$	<p>Si <math>\sigma_0^2 \in IC_{\sigma^2} \Rightarrow \text{Aceptar } H_0</math></p> <p>Si <math>\sigma_0^2 \notin IC_{\sigma^2} \Rightarrow \text{Rechazar } H_0</math></p>

El estadístico para el contraste de las hipótesis de varianza es:

$$(N-1)\frac{s^2}{\sigma^2} \sim \chi_{N-1}^2$$

Su valor es  $19 \cdot 5,31316 / 4 = 25,23751$

Para la varianza debemos recurrir a la distribución  $\chi^2$  y obtener los valores  $g_1$  y  $g_2$  que dejan a su izquierda y a su derecha respectivamente  $\alpha/2$ . Por ejemplo para el 5% de  $\alpha$ :



**Distribución Chi Cuadrado  $\chi^2$**  Contiene los valores de  $\chi^2$  tales que  $\alpha = P(\chi_v^2 \geq \chi)$ , donde  $v$  son los Grados de Libertad

$v/\alpha$	0,005	0,01	0,025	0,05	0,1	0,25	0,5	0,75	0,9	0,95	0,975	0,99	0,995
1	7,879	6,635		3,842	2,706	1,323	0,455	0,102	0,016	0,004		0,000	0,000
2	10,597	9,210	7,378	5,992	4,605	2,773	1,386	0,575	0,211	0,103	0,051	0,020	0,010
3	12,838	11,345	9,348	7,815	6,251	4,108	2,366	1,213	0,584	0,352	0,16	0,115	0,072
4	14,860	13,277	11,143	9,488	7,779	5,385	3,357	1,923	1,064	0,711	0,34	0,297	0,207
5	16,750	15,086	12,833	11,071	9,236	6,626	4,352	2,675	1,610	1,146	0,631	0,554	0,412
6	18,548	16,812	14,449	12,592	10,645	7,841	5,348	3,455	2,204	1,635	1,037	0,872	0,676
7	20,278	18,475	16,013	14,067	12,017	9,037	6,346	4,255	2,833	2,167	1,390	1,239	0,989
8	21,955	20,090	17,535	15,507	13,362	10,219	7,344	5,071	3,490	2,733	2,180	1,647	1,344
9	23,589	21,666	19,023	16,919	14,684	11,389	8,343	5,899	4,168	3,325	2,700	2,088	1,735
10	25,188	23,209	20,483	18,307	15,987	12,549	9,342	6,737	4,865	3,940	3,147	2,558	2,156
11	26,757	24,725	21,920	19,675	17,275	13,701	10,341	7,584	5,578	4,575	3,516	3,054	2,603
12	28,300	26,217	23,337	21,026	18,549	14,845	11,340	8,438	6,304	5,226	4,004	3,571	3,074
13	29,819	27,688	24,736	22,362	19,812	15,984	12,340	9,299	7,042	5,892	5,009	4,107	3,565
14	31,319	29,141	26,119	23,685	21,064	17,117	13,339	10,165	7,790	6,571	5,529	4,660	4,075
15	32,802	30,578	27,488	24,996	22,307	18,245	14,339	11,037	8,547	7,261	6,162	5,229	4,601
16	34,267	32,000	28,845	26,296	23,542	19,369	15,339	11,912	9,312	7,962	6,808	5,812	5,142
17	35,718	33,409	30,191	27,597	24,769	20,489	16,338	12,792	10,085	8,672	7,464	6,408	5,697
18	37,156	34,805	31,526	28,899	25,989	21,605	17,338	13,675	10,865	9,390	8,231	7,015	6,265
19	38,582	36,191	32,852	30,143	27,204	22,718	18,338	14,558	11,651	10,117	8,907	7,633	6,844
20	39,997	37,566	34,178	31,410	28,412	23,828	19,337	15,452	12,443	10,851	9,591	8,260	7,434
21	41,401	38,932	35,479	32,671	29,615	24,935	20,337	16,344	13,240	11,591	10,283	8,897	8,034
22	42,796	40,289	36,781	33,925	30,813	26,039	21,337	17,240	14,042	12,338	10,982	9,543	8,643

El extremo izquierdo del intervalo para la varianza queda:  $19 \cdot 5,31316 / 8,907 = 11,3338$ . El extremo izquierdo del intervalo para la desviación típica será la raíz, de valor 3,3666.

Los intervalos para los tres valores de nivel de confianza son:

Intervalos de confianza del 90,0% para la desviación estándar: [1,83002; 3,15884]

Intervalos de confianza del 95,0% para la desviación estándar: [1,75295; 3,36666]

Intervalos de confianza del 99,0% para la desviación estándar: [1,61756; 3,8406]

**11) Indica que conclusiones sacas sobre la variabilidad declarada**

En este caso el valor de desviación estándar  $\sigma = 2$  está incluido en los tres intervalos, por lo que se acepta dicho valor como hipótesis de valor poblacional para cualquiera de los tres niveles de confianza.

**12) Si la empresa ha fijado un alfa o nivel de significación o probabilidad de cometer un error de tipo 1 igual al 5%, interpreta cada uno de los contrastes siguientes y justifica si EcoLine tiene evidencia desde el punto de vista estadístico para reclamar a Junkers.**

Tests de hipótesis 1	Tests de hipótesis 2	Tests de hipótesis 3
$H_0: \sigma = 2$ $H_1: \sigma \neq 2$	$H_0: \sigma \geq 2$ $H_1: \sigma < 2$	$H_0: \sigma \leq 2$ $H_1: \sigma > 2$

Los resultados de los P-value se muestran a continuación y se observa como todos ellos son superiores a 0,05 y no se puede rechazar ninguna.

Hipótesis Nula:  $\sigma = 2,0$

Alternativa: no igual

Chi-cuadrado calculado = 25,2375

Valor-P = 0,305744

No se rechaza la hipótesis nula para  $\alpha = 0,05$ .

Prueba chi-cuadrada

Hipótesis Nula:  $\sigma = 2,0$

Alternativa: menor que

Chi-cuadrado calculado = 25,2375

Valor-P = 0,847128

No se rechaza la hipótesis nula para  $\alpha = 0,05$ .



Prueba chi-cuadrada

Hipótesis Nula:  $\sigma = 2,0$

Alternativa: mayor que

Chi-cuadrado calculado = 25,2375

Valor-P = 0,152872

No se rechaza la hipótesis nula para  $\alpha = 0,05$ .

Recuerda que en un contraste de hipótesis entre las dos hipótesis se tiene que abarcar todas las posibilidades, con respecto al parámetro poblacional sobre el que se está planteando el contraste. Así pues, se puede ver planteada la hipótesis como  $H_0$  igual y después mayor o menor, por ejemplo aunque se escribe que la hipótesis nula es que  $\sigma=2$  frente a la hipótesis

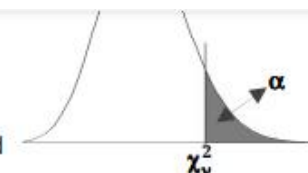


alternativa de  $\sigma < 2$ . Sin embargo, realmente se está contrastando la hipótesis nula  $\sigma \geq 2$  frente a  $\sigma < 2$  aunque no se explicita.

Para el cálculo del P-value manualmente, en el caso de hipótesis simple, se toma el valor del estadístico 25,23751 y tras observar las tablas, se ve que a su derecha deja aproximadamente el 15% de probabilidad.



**Distribución Chi Cuadrado  $\chi^2$**  Contiene los valores de  $\chi^2$  tales que  $\alpha = P(\chi_v^2 \geq \chi)$ , donde  $v$  son los Grados de Libertad



0,15 aprox

$v/\alpha$	0,005	0,01	0,025	0,05	0,1	0,25	0,5	0,75	0,9	0,95	0,975	0,99	0,995
1	7,879	6,635	5,024	3,842	2,706	1,323	0,455	0,102	0,016	0,004	0,001	0,000	0,000
2	10,597	9,210	7,378	5,992	4,605	2,773	1,386	0,575	0,211	0,103	0,051	0,020	0,010
3	12,838	11,345	9,348	7,815	6,251	4,108	2,366	1,213	0,584	0,352	0,216	0,115	0,072
4	14,860	13,277	11,143	9,488	7,779	5,385	3,357	1,923	1,064	0,711	0,484	0,297	0,207
5	16,750	15,086	12,833	11,071	9,236	6,626	4,352	2,675	1,610	1,146	0,831	0,554	0,412
6	18,548	16,812	14,449	12,592	10,645	7,841	5,348	3,455	2,204	1,635	1,237	0,872	0,676
7	20,278	18,475	16,013	14,067	12,017	9,037	6,346	4,255	2,833	2,167	1,690	1,239	0,989
8	21,955	20,090	17,535	15,507	13,362	10,219	7,344	5,071	3,490	2,733	2,180	1,647	1,344
9	23,589	21,666	19,023	16,919	14,684	11,389	8,343	5,899	4,168	3,325	2,700	2,088	1,735
10	25,188	23,209	20,483	18,307	15,987	12,549	9,342	6,737	4,865	3,940	3,247	2,558	2,156
11	26,757	24,725	21,920	19,675	17,275	13,701	10,341	7,584	5,578	4,575	3,816	3,054	2,603
12	28,300	26,217	23,337	21,026	18,549	14,845	11,340	8,438	6,304	5,226	4,404	3,571	3,074
13	29,819	27,688	24,736	22,362	19,812	15,984	12,340	9,299	7,042	5,892	5,009	4,107	3,565
14	31,319	29,141	26,119	23,685	21,064	17,117	13,339	10,165	7,790	6,571	5,629	4,660	4,075
15	32,802	30,578	27,488	24,996	22,307	18,245	14,339	11,037	8,547	7,261	6,262	5,229	4,601
16	34,267	32,000	28,845	26,296	23,542	19,369	15,339	11,912	9,312	7,962	6,908	5,812	5,142
17	35,718	33,409	30,191	27,587	24,769	20,489	16,338	12,792	10,085	8,672	7,564	6,408	5,697
18	37,156	34,805	31,526	28,869	25,989	21,605	17,338	13,675	10,865	9,390	8,231	7,015	6,265
19	38,582	36,191	32,852	30,144	27,204	22,718	18,338	14,562	11,651	10,117	8,907	7,633	6,844
20	39,997	37,566	34,170	31,410	28,412	23,828	19,337	15,452	12,443	10,851	9,591	8,260	7,434
21	41,401	38,932	35,479	32,671	29,615	24,935	20,337	16,344	13,240	11,591	10,283	8,897	8,034

25,23751

Vemos que 25,23751 queda cerca de  $g_1$  (extremo derecho del intervalo) y por tanto, al ser la hipótesis simple, se trata de la cola a la derecha ( $p\text{-value}/2$ ) y falta incluir la cola izquierda. Por tanto  $p\text{-value}$  será aproximadamente  $2 \cdot 0,15 = 0,3$ .

En el caso del segundo test, 25,23751 representa el límite que deja a su izquierda todo el  $p\text{-value}$  (una única cola), por lo que su valor es aproximadamente 0,85.

En el caso del tercer test, 25,23751 representa el límite que deja a su derecha todo el  $p\text{-value}$  (una única cola), por lo que su valor es aproximadamente 0,15.