Cómo evitar errores de exactitud en los cálculos: Uso de calculadoras y reglas de redondeo

Contenidos

Presentación

Este documento está dirigido a los alumnos que empiezan los estudios de postgrado con el curso de *Fundamentos de Diseño y Estadística*. El curso está diseñado para iniciarse en Bioestadística y aprender a interpretar los resultados de los análisis realizados con *Stata*.

Las autoevaluaciones, las pruebas de evaluación y los ejercicios de los textos están diseñados para realizarlos con calculadoras estadísticas, de manera que no es necesario disponer de *Stata*.

Se aconseja a **todos** los alumnos, dispongan o no de *Stata*, usar una **Hoja de cálculo** (Microsoft Excel o la del paquete LibreOffice (https://es.libreoffice.org), o la de Google disponible *on-line* de forma gratuita en https://www.google.com/intl/es_es/sheets/about/) como herramienta de cálculo para realizar este primer curso; aprender a trabajar con hojas de cálculo es muy formativo. Sólo se recomienda usar la calculadora del Windows para operaciones muy elementales.

Los alumnos que dispongan de *Stata* deben usar la calculadora de *Stata* (o el Excel) para realizar los ejercicios que requieren aplicar fórmulas y las evaluaciones. Una vez resueltos los ejercicios y las evaluaciones se les aconseja replicar todos los ejemplos del texto con *Stata* ya que es la mejor forma de consolidar los conocimientos adquiridos en el curso de *Stata*.

Ejemplos de error de redondeo

1) Calcular el valor del RR dando el resultado con 2 decimales:

- Cálculo exacto: $RR = \frac{R_1}{R_0} = \frac{11/14}{2/31} = \frac{0.7857143...}{0.0645161...} = \mathbf{12.18} \leftarrow$ *Valor exacto redondeado a 2 decimales*
- Cálculo con los riesgos redondeados a 2 y 3 decimales: $RR = \frac{0.79}{0.06} = \mathbf{13.17}$; $RR = \frac{0.786}{0.065} = \mathbf{12.09}$

2) Redondeo de un valor redondeado:

Sea el valor **X=2.745739** que redondeado a 2 decimales es $\mathbf{X_{R2} = 2.75}$ y a 1 decimal es: $\mathbf{X_{R1} = 2.7}$.
Pero si se redondea $X_{R2} = 2.75$ a 1 decimal da: $X_R = 2.8$ que no coincide con $X_{R1} = 2.7$.

¿Sabe usar la calculadora? Reglas para no cometer errores de redondeo

Los errores de redondeo se cometen cuando, en una operación compleja, los cálculos intermedios se realizan utilizando valores redondeados en lugar de introducir los valores exactos.

Ejercicio. Para aprender a realizar cálculos correctos, **efectúe las siguientes operaciones** (con su calculadora o con la de Windows) guardando la razón de prevalencias (PR) y su error estándar (SE) en las memorias para evitar ir acumulando errores de redondeo.

$$\text{PR}=\frac{9/1100}{1/1000}=8.1818.....;\quad \text{SE}(\ln \text{PR})=\sqrt{\frac{1}{1}-\frac{1}{1000}+\frac{1}{9}-\frac{1}{1100}}=1.0531....$$

El siguiente intervalo de confianza (IC 95%) debe obtenerlo recuperando los datos de memoria e introduciendo una sola vez el valor 1.96 (no debe introducir resultados parciales en la calculadora). Para ello, primero debe multiplicar SE por 1.96 y volverlo a guardar en la memoria. Los puntos suspensivos en las siguientes fórmulas indican que se trata de valores guardados con todos los decimales:

$$\text{IC 95\% de PR:}\quad \text{PR}\times e^{\pm 1.96\times \text{SE}}=8.1818...\times e^{\pm 2.0642...}$$

$$\text{Límite inferior: } 8.1818...\times 0.1269...=1.038$$

$$\text{Límite superior: } 8.1818...\times 7.8793...=64.47$$

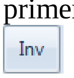
Calculadora de Windows

Si para calcular el anterior intervalo usa la **calculadora de Windows** (que sólo dispone de una memoria), deberá guardar PR=8.18... en la memoria (pulsando la tecla MS) y el exponente (1.96×SE) en el portapapeles (Edición | Copiar ó Ctrl+C); de esta forma tendrá los valores PR y SE almacenados con gran exactitud ya que la calculadora los obtiene y guarda con 32 cifras.

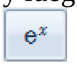
El valor PR de la memoria se recupera pulsando la tecla MR. El exponente guardado en el portapapeles se recupera con el menú Edición | Pegar ó Ctrl+V.

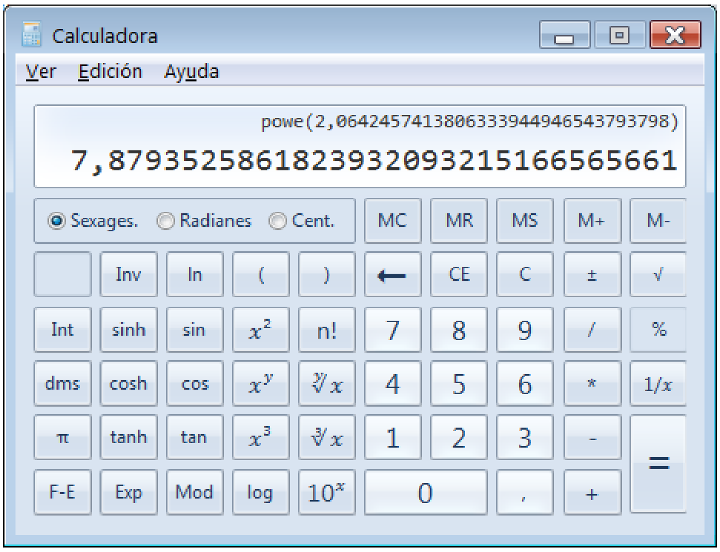
Imagen de la calculadora de Windows para hallar $e^{2.0642}$:

Puesto que la función exponencial es la inversa del logaritmo neperiano, primero se pulsa la tecla



para que aparezcan las funciones inversas y luego se pulsa la tecla

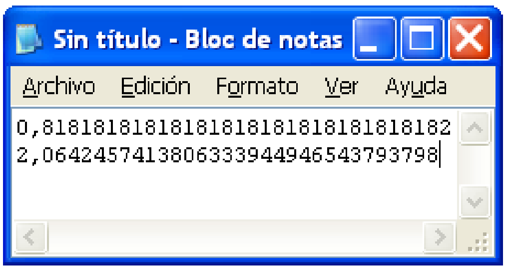




¿Cómo realizar con la calculadora de Windows cálculos que requieran más memorias?

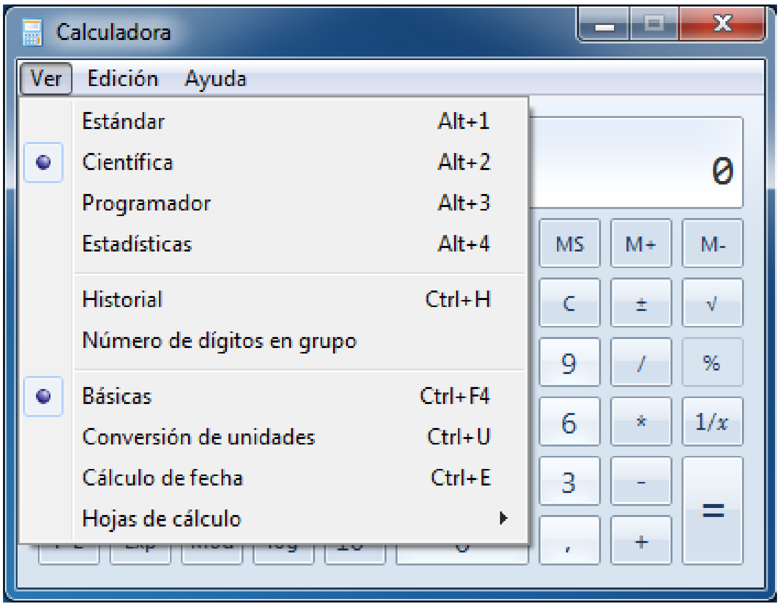
Abra el **Bloc de notas** junto a la calculadora, copie cada cálculo parcial pulsando Ctrl+C y péguelo en el Bloc de notas pulsando Ctrl+V.

Para recuperar los cálculos se copian del Bloc de notas (Ctrl+C) y se pegan en la calculadora (Ctrl+V).



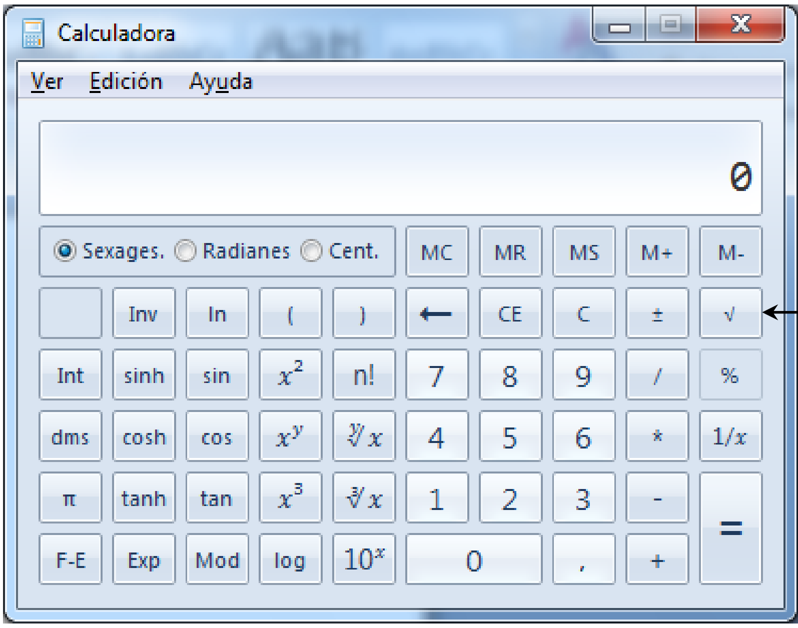
Calculadora de Windows 7

Está formada por un conjunto de calculadoras que se eligen desplegando el menú **Ver**.

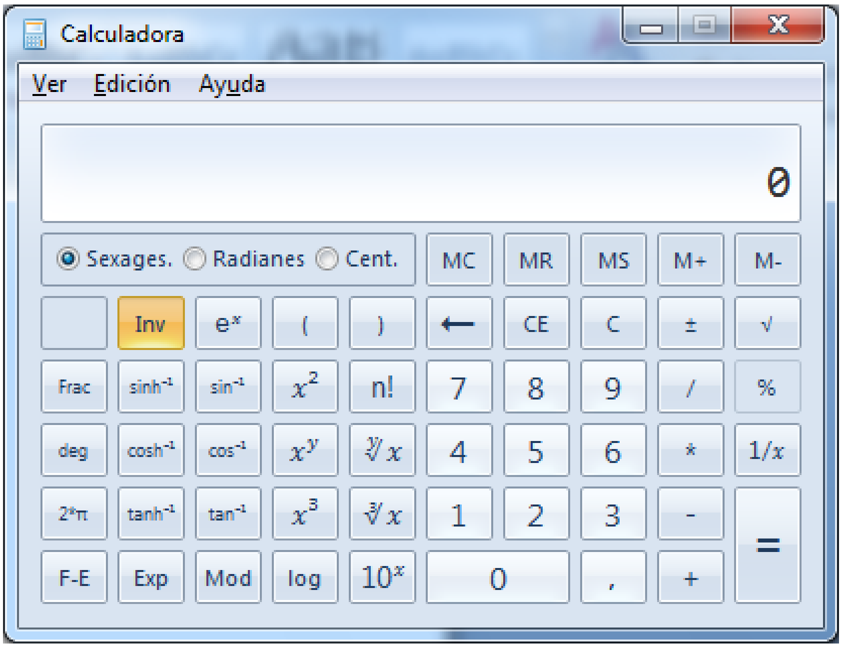


Se elige la **Científica**.

Puede comprobar en esta imagen que la calculadora ya incorpora algunas funciones inversas; por ejemplo $\boxed{x^2}$ y raíz cuadrada $\boxed{\sqrt{\ }}$



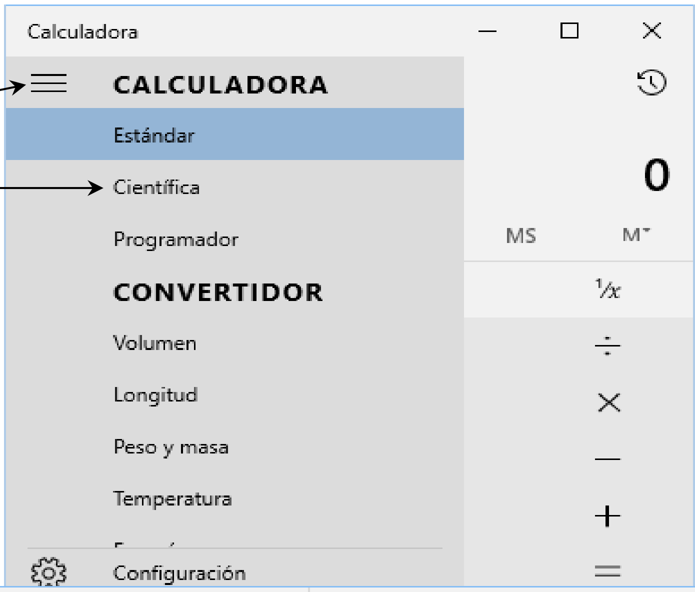
Para trabajar con las funciones inversas no disponibles en el teclado se pulsa la tecla $\boxed{\text{Inv}}$ y las inversas aparecen en el teclado. Por ejemplo, si se pulsa $\boxed{\text{Inv}}$ en el lugar $\boxed{\text{ln}}$ aparece $\boxed{e^x}$



Calculadora de Windows 10

Está formada por varias calculadoras que se eligen desplegando el menú **Ver**

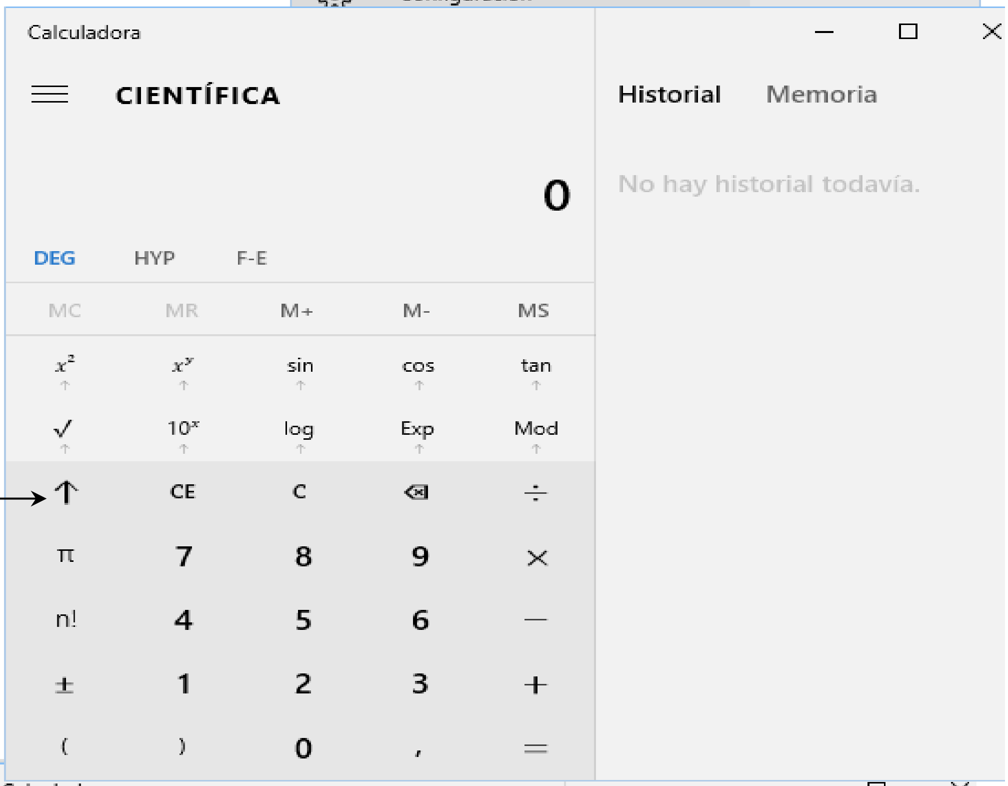
Se elige la **Científica**



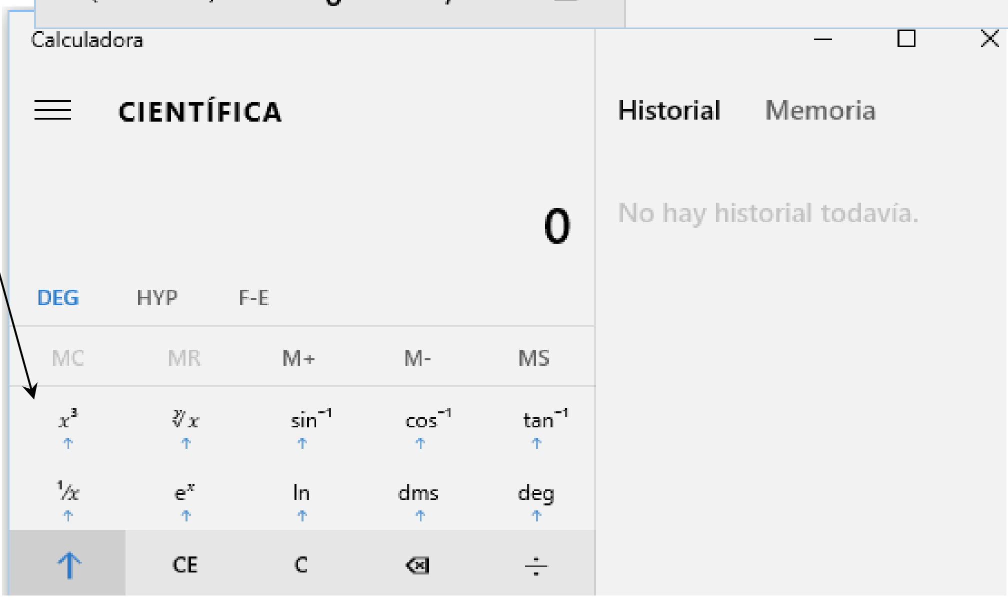
1) La columna de la derecha aparece al ampliar la anchura de la ventana

2) Para borrar un resultado pulse la tecla **Esc.**

Esta imagen indica las funciones básicas que incorpora.



Pulsa la tecla ↑ para trabajar con el resto de funciones que incorpora



Usar Excel o Calc de LibreOffice como calculadora (recomendado)

Se recomienda realizar los cálculos con Excel en vez de la calculadora porque:

1) Los valores que se introducen en el Excel quedan permanentemente guardados y pueden verse.
2) Los cálculos complejos se realizan con la máxima exactitud y sin problemas de redondeo.
3) Permite realizar cálculos repetitivos ya que se guarda la fórmula introducida.
4) Dispone de funciones matemáticas y estadísticas.

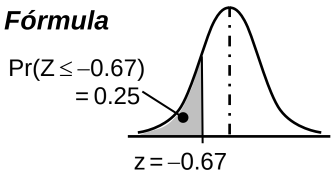
Nota: Estos ejemplos se han efectuado con el Excel 2010 de Microsoft, pero pueden realizarse de forma muy similar con la Hoja de cálculo gratuita del paquete LibreOffice, o la de Google disponible *on-line*.

Ejemplo 1: Distribución Normal

B4 | =INV.NORM.ESTAND(A4)

	A	B	C	D	E
1	**Ley Normal N(0; 1)**				
2					
3	P.Acum →	z			
4	0,25	-0,674490			
5	0,975	1,959964			

← ***Fórmula***

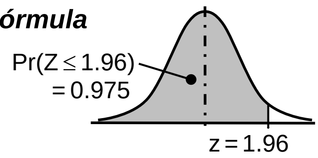


A partir del área acumulada 0.25 da el valor z ≈ −0.67

E5 | =DISTR.NORM.ESTAND.N(D5;1)

	A	B	C	D	E
1	**Ley Normal N(0; 1)**				
2					
3	P.Acum →	z		z →	P.Acum
4	0,25	-0,674490		-0,674490	0,250000
5	0,975	1,959964		1,959964	0,975000

← ***Fórmula***

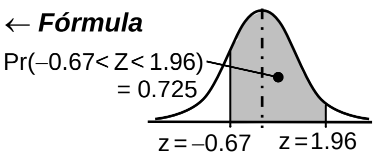


A partir del valor z≈1.96 da el área acumulada 0.975

F4 | =E5-E4

	A	B	C	D	E	F
1	**Ley Normal N(0; 1)**					
2						
3	P.Acum →	z		z →	P.Acum	P(-0.67< Z< 1.96)
4	0,25	-0,674490		-0,674490	0,250000	0,725000
5	0,975	1,959964		1,959964	0,975000	

← ***Fórmula***



Área acumulada entre los valores z ≈ −0.67 y z≈1.96: Área acumulada = 0.725

Equivalencias con las funciones disponibles en Microsoft Excel 2003:

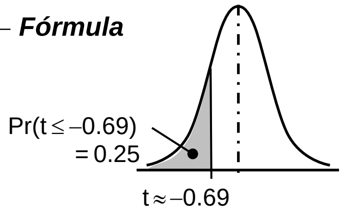
Excel 2010	Excel 2003
DISTR.NOR.ESTAND.N(1,96;1)	DISTR.NOR.ESTAND(1,96)
INV.NOR.ESTAND(0,95)	DISTR.NOR.ESTAND.INV(0,95)

Ejemplo 2: Distribución t de Student (df=20)

B3 | f_x =INV.T(A3;20)

	A	B	C	D	E
1	Distribución t de Student (con df=20)				
2	P.Acum →	t(20)		t(20) →	P.Acum
3	0,25	-0,686954		-0,686954	0,250000
4	0,975	2,085963		2,085963	0,975000
5	P.2Colas				P.ColaDcha
6	0,05	2,085963		2,085963	0,025000

← ***Fórmula***

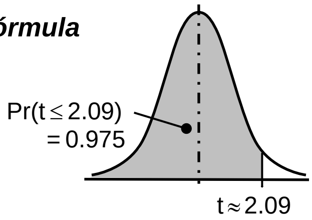


A partir del área acumulada 0.25 da el valor t ≈ –0.67

E4 | f_x =DISTR.T.N(D4;20;1)

	A	B	C	D	E
1	Distribución t de Student (con df=20)				
2	P.Acum →	t(20)		t(20) →	P.Acum
3	0,25	-0,686954		-0,686954	0,250000
4	0,975	2,085963		2,085963	0,975000
5	P.2Colas				P.ColaDcha
6	0,05	2,085963		2,085963	0,025000

← ***Fórmula***

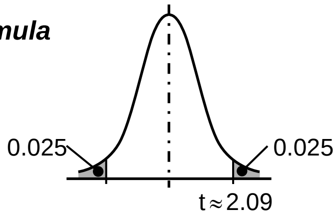


A partir del valor t ≈ 2.09 da el área acumulada 0.975

B6 | f_x =INV.T.2C(A6;20)

	A	B	C	D	E
1	Distribución t de Student (con df=20)				
2	P.Acum →	t(20)		t(20) →	P.Acum
3	0,25	-0,686954		-0,686954	0,250000
4	0,975	2,085963		2,085963	0,975000
5	P.2Colas				P.ColaDcha
6	0,05	2,085963		2,085963	0,025000

← ***Fórmula***

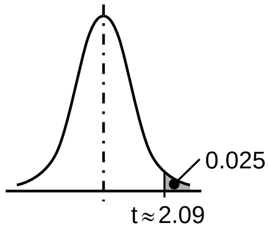


A partir del área de las dos colas 0.025+0.025 = 0.05 da el valor t ≈ 2.09

E6 | f_x =DISTR.T.CD(D6;20)

	A	B	C	D	E
1	Distribución t de Student (con df=20)				
2	P.Acum →	t(20)		t(20) →	P.Acum
3	0,25	-0,686954		-0,686954	0,250000
4	0,975	2,085963		2,085963	0,975000
5	P.2Colas				P.ColaDcha
6	0,05	2,085963		2,085963	0,025000

← ***Fórmula***



A partir del valor t≈2.09 da el área de la cola derecha 0.025

Equivalencias con las funciones disponibles en Microsoft Excel 2003:

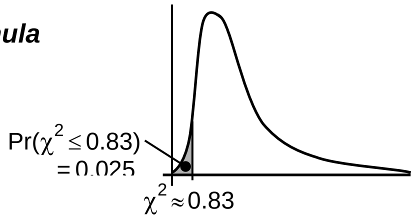
Excel 2010	Excel 2003
DISTR.T(2,09;20;1)	DISTR.T.CD(2,09;20)
INV.T.2C(0,05)	DISTR.T.INV(0,05)

Ejemplo 3: Distribución Ji cuadrado (df=5)

B3 | =INV.CHICUAD(A3;5)

	A	B	C	D	E
1	**Distribución Ji cuadrado (con df=5)**				
2	P.Acum →	$\chi^2(5)$		$\chi^2(5)$ →	P.Acum
3	0,025	0,831212		0,831212	0,025000
4	0,975	12,832502		12,832502	0,975000
5	P.ColaDcha				P.ColaDcha
6	0,05	11,070498		11,070498	0,050000

← ***Fórmula***

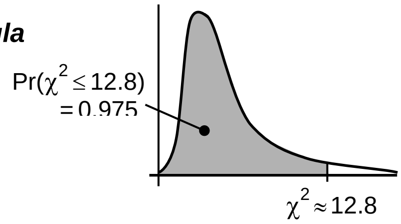


A partir del área acumulada 0.025 da el valor $\chi^2 \approx 0.83$

E4 | =DISTR.CHICUAD(D4;5;1)

	A	B	C	D	E
1	**Distribución Ji cuadrado (con df=5)**				
2	P.Acum →	$\chi^2(5)$		$\chi^2(5)$ →	P.Acum
3	0,025	0,831212		0,831212	0,025000
4	0,975	12,832502		12,832502	0,975000
5	P.ColaDcha				P.ColaDcha
6	0,05	11,070498		11,070498	0,050000

← ***Fórmula***

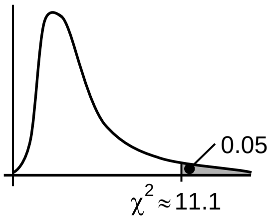


A partir del valor $\chi^2 \approx 12.8$ da el área acumulada 0.975

B6 | =INV.CHICUAD.CD(A6;5)

	A	B	C	D	E
1	**Distribución Ji cuadrado (con df=5)**				
2	P.Acum →	$\chi^2(5)$		$\chi^2(5)$ →	P.Acum
3	0,025	0,831212		0,831212	0,025000
4	0,975	12,832502		12,832502	0,975000
5	P.ColaDcha				P.ColaDcha
6	0,05	11,070498		11,070498	0,050000

← ***Fórmula***

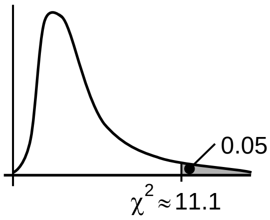


A partir del área de la cola derecha 0.05 da el valor $\chi^2 \approx 11.1$

E6 | =DISTR.CHICUAD.CD(D6;5)

	A	B	C	D	E
1	**Distribución Ji cuadrado (con df=5)**				
2	P.Acum →	$\chi^2(5)$		$\chi^2(5)$ →	P.Acum
3	0,025	0,831212		0,831212	0,025000
4	0,975	12,832502		12,832502	0,975000
5	P.ColaDcha				P.ColaDcha
6	0,05	11,070498		11,070498	0,050000

← ***Fórmula***



A partir del valor $\chi^2 \approx 11.1$ da el área de la cola derecha 0.05

Equivalencias con las funciones disponibles en Microsoft Excel 2003:

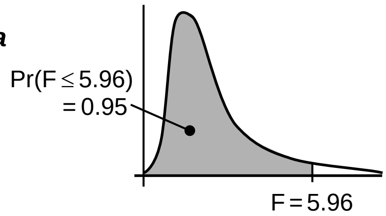
Excel 2010	Excel 2003
DISTR.CHICUAD.CD(11,07;5)	DISTR.CHI(11,07;5)
INV.CHICUAD.CD(0,05;5)	*No dispone de distribución inversa*

Ejemplo 4: Distribución F de Snedecor (df1=10 y df2=4)

B3 =INV.F(A3;10;4)

	A	B	C	D	E
1	**Ley F de Snedecor (df1=10; df2=4)**				
2	P.Acum →	F(10;4)		F(10;4) →	P.Acum
3	0,95	5,964371		5,964371	0,950000
4	0,05	0,287517		0,287517	0,050000
5	P.ColaDcha				P.ColaDcha
6	0,05	5,964371		5,964371	0,05

← ***Fórmula***

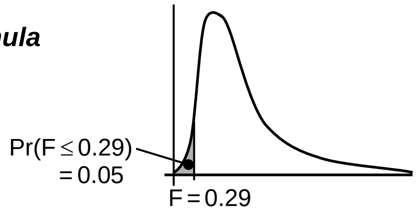


A partir del área acumulada 0.95 da el valor F≈5.96

E4 =DISTR.F.N(D4;10;4;1)

	A	B	C	D	E
1	**Ley F de Snedecor (df1=10; df2=4)**				
2	P.Acum →	F(10;4)		F(10;4) →	P.Acum
3	0,95	5,964371		5,964371	0,950000
4	0,05	0,287517		0,287517	0,050000
5	P.ColaDcha				P.ColaDcha
6	0,05	5,964371		5,964371	0,05

← ***Fórmula***

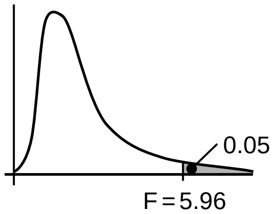


A partir del valor F≈0.29 da el área acumulada 0.05

B6 =INV.F.CD(A6;10;4)

	A	B	C	D	E
1	**Ley F de Snedecor (df1=10; df2=4)**				
2	P.Acum →	F(10;4)		F(10;4) →	P.Acum
3	0,95	5,964371		5,964371	0,950000
4	0,05	0,287517		0,287517	0,050000
5	P.ColaDcha				P.ColaDcha
6	0,05	5,964371		5,964371	0,05

← ***Fórmula***

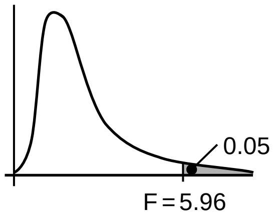


A partir del área de la cola derecha 0.05 da el valor t ≈2.09

E6 =DISTR.F.CD(D6;10;4)

	A	B	C	D	E
1	**Ley F de Snedecor (df1=10; df2=4)**				
2	P.Acum →	F(10;4)		F(10;4) →	P.Acum
3	0,95	5,964371		5,964371	0,950000
4	0,05	0,287517		0,287517	0,050000
5	P.ColaDcha				P.ColaDcha
6	0,05	5,964371		5,964371	0,05

← ***Fórmula***



A partir del valor F≈5.96 da el área de la cola derecha 0.05

Equivalencias con las funciones disponibles en Microsoft Excel 2003:

Excel 2010	Excel 2003
DISTR.F.CD(5,96;10;4)	DISTR.F(5,96;10;4)
INV.F.CD(0,05;10;4)	DISTR.F.INV(0,05;10;4)

Ejemplo 5: Distribuciones Binomial y Poisson

Las casillas B1 y B2 contienen los parámetros de la ley Binomial y la B3 el parámetro λ de la ley de Poisson.

`D3` *fx* `=DISTR.BINOM.N(C3;B$2;B$1;0)`

	A	B	C	D	E	F	G	H	I	J
1	p =	**0,05**		**Ley de Binomial**				**Ley de Poisson**		
2	n =	**100**	k	Pr(S=k)	Pr(S<=k)	Pr(S>k)		Pr(S=k)	Pr(S<=k)	Pr(S>k)
3	L =	**5**	0	0,005921	0,005921	0,994079		0,006738	0,006738	0,993262
4			1	0,031161	0,037081	0,962919		0,033690	0,040428	0,966310
5			2	0,081182	0,118263	0,881737		0,084224	0,124652	0,915776

← ***Fórmula***

*Esta función con el último parámetro igual a **0** da la probabilidad de encontrar en la muestra k de sujetos con la característica objeto de estudio.*

`E5` *fx* `=DISTR.BINOM.N(C5;B$2;B$1;1)`

	A	B	C	D	E	F	G	H	I	J
1	p =	**0,05**		**Ley de Binomial**				**Ley de Poisson**		
2	n =	**100**	k	Pr(S=k)	Pr(S<=k)	Pr(S>k)		Pr(S=k)	Pr(S<=k)	Pr(S>k)
3	L =	**5**	0	0,005921	0,005921	0,994079		0,006738	0,006738	0,993262
4			1	0,031161	0,037081	0,962919		0,033690	0,040428	0,966310
5			2	0,081182	0,118263	0,881737		0,084224	0,124652	0,915776

← ***Fórmula***

*Esta función con el último parámetro igual a **1** da la probabilidad acumulada de encontrar en la muestra k o menos de k de sujetos con la característica objeto de estudio.*

`F3` *fx* `=1-E3`

	A	B	C	D	E	F	G	H	I	J
1	p =	**0,05**		**Ley de Binomial**				**Ley de Poisson**		
2	n =	**100**	k	Pr(S=k)	Pr(S<=k)	Pr(S>k)		Pr(S=k)	Pr(S<=k)	Pr(S>k)
3	L =	**5**	0	0,005921	0,005921	0,994079		0,006738	0,006738	0,993262
4			1	0,031161	0,037081	0,962919		0,033690	0,040428	0,966310
5			2	0,081182	0,118263	0,881737		0,084224	0,124652	0,915776

← ***Fórmula***

Esta columna es el complemento de la anterior: da la probabilidad de encontraren la muestra más de k de sujetos con la característica objeto de estudio.

Este bloque realiza los mismos cálculos con la ley de Poisson. Se obtienen resultados similares porque se cumple el supuesto de muestra grande (n≥100) y baja probabilidad (π≤0.05).

`H3` *fx* `=POISSON.DIST(C3;B$3;0)`

	A	B	C	D	E	F	G	H	I	J
1	p =	**0,05**		**Ley de Binomial**				**Ley de Poisson**		
2	n =	**100**	k	Pr(S=k)	Pr(S<=k)	Pr(S>k)		Pr(S=k)	Pr(S<=k)	Pr(S>k)
3	L =	**5**	0	0,005921	0,005921	0,994079		0,006738	0,006738	0,993262
4			1	0,031161	0,037081	0,962919		0,033690	0,040428	0,966310
5			2	0,081182	0,118263	0,881737		0,084224	0,124652	0,915776

← ***Fórmula***

*Esta función con el último parámetro igual a **0** da la probabilidad de encontrar en la muestra k de sujetos con la característica objeto de estudio.*

`I5` *fx* `=POISSON.DIST(C5;B$3;1)`

	A	B	C	D	E	F	G	H	I	J
1	p =	**0,05**		**Ley de Binomial**				**Ley de Poisson**		
2	n =	**100**	k	Pr(S=k)	Pr(S<=k)	Pr(S>k)		Pr(S=k)	Pr(S<=k)	Pr(S>k)
3	L =	**5**	0	0,005921	0,005921	0,994079		0,006738	0,006738	0,993262
4			1	0,031161	0,037081	0,962919		0,033690	0,040428	0,966310
5			2	0,081182	0,118263	0,881737		0,084224	0,124652	0,915776

← ***Fórmula***

*Esta función con el último parámetro igual a **1** da la probabilidad acumulada de encontrar en la muestra k o menos de k de sujetos con la característica objeto de estudio.*

`J3` *fx* `=1-H3`

	A	B	C	D	E	F	G	H	I	J
1	p =	**0,05**		**Ley de Binomial**				**Ley de Poisson**		
2	n =	**100**	k	Pr(S=k)	Pr(S<=k)	Pr(S>k)		Pr(S=k)	Pr(S<=k)	Pr(S>k)
3	L =	**5**	0	0,005921	0,005921	0,994079		0,006738	0,006738	0,993262
4			1	0,031161	0,037081	0,962919		0,033690	0,040428	0,966310
5			2	0,081182	0,118263	0,881737		0,084224	0,124652	0,915776

← ***Fórmula***

Esta columna es el complemento de la anterior: da la probabilidad de encontraren la muestra más de k de sujetos con la característica objeto de estudio.

Ejemplo 6: Distribución Hipergeométrica

Las casillas B1, B2 y B3 contienen los parámetros de la ley Hipergeométrica.
La diferencia con la ley Binomial es que la población es finita (está formada por N sujetos de los cuales K poseen la característica objeto de estudio) y el muestreo es exhaustivo (sin remplazamiento).

D3 =DISTR.HIPERGEOM.N(C3;B$3;B$2;B$1;0)

	A	B	C	D	E	F	G
1	Poblac. N=	**10**		**Ley Hipergeométrica**			
2	Poblac. K=	**6**	k	Pr(S=k)	Pr(S<=k)	Pr(S>k)	
3	Muestra n=	**3**	0	0,033333	0,033333	0,966667	
4			1	0,300000	0,333333	0,666667	
5			2	0,500000	0,833333	0,166667	
6			3	0,166667	1,000000	0,000000	

← ***Fórmula***

*Esta función con el último parámetro igual a **0** da la probabilidad de encontrar en la muestra k de sujetos con la característica objeto de estudio.*

E5 =DISTR.HIPERGEOM.N(C5;B$3;B$2;B$1;1)

	A	B	C	D	E	F	G
1	Poblac. N=	**10**		**Ley Hipergeométrica**			
2	Poblac. K=	**6**	k	Pr(S=k)	Pr(S<=k)	Pr(S>k)	
3	Muestra n=	**3**	0	0,033333	0,033333	0,966667	
4			1	0,300000	0,333333	0,666667	
5			2	0,500000	0,833333	0,166667	
6			3	0,166667	1,000000	0,000000	

← ***Fórmula***

*Esta función con el último parámetro igual a **1** da la probabilidad acumulada de encontrar en la muestra k o menos de k de sujetos con la característica objeto de estudio.*

F4 =1-E4

	A	B	C	D	E	F	G
1	Poblac. N=	**10**		**Ley Hipergeométrica**			
2	Poblac. K=	**6**	k	Pr(S=k)	Pr(S<=k)	Pr(S>k)	
3	Muestra n=	**3**	0	0,033333	0,033333	0,966667	
4			1	0,300000	0,333333	0,666667	
5			2	0,500000	0,833333	0,166667	
6			3	0,166667	1,000000	0,000000	

← ***Fórmula***

Esta columna es el complemento de la anterior: da la probabilidad de encontraren la muestra más de k de sujetos con la característica objeto de estudio.

Ejemplo 7: Intervalo de confianza de una razón de proporciones

$$\text{PR}=\ \frac{9/1100}{1/1000} = 8.1818\ .....;\quad \text{SE(ln PR)}\ = \sqrt{\frac{1}{1}-\frac{1}{1000}+\frac{1}{9}-\frac{1}{1100}}\ = \ 1.0532\ ....$$

$$\text{IC 95\% de PR:}\quad \text{PR}\times e^{\pm 1.96\times SE} = \ 8.1818\times e^{\pm 1.96\times 1.0532}$$

$$\text{Límite inferior:}\ \ 8.1818\times e^{-1.96\times 1.0532} \ = \ 8.1818\times e^{-2.0643} = 8.1818\times 0.1269 = \ 1.04$$

$$\text{Límite superior:}\ 8.1818\times e^{1.96\times 1.0532} \ = \ 8.1818\times e^{2.0643} \ = 8.1818\times 7.8798 = \ 64.5$$

Las siguientes imágenes contienen el cálculo con Excel. Los datos se han introducido en las casillas con fondo blanco. Se han señalado con fondo de colar las casillas que contienen las fórmulas de cálculo y que se actualizan automáticamente al cambiar los datos.

En la primera fila de la imagen aparece la fórmula que contiene la casilla señalada.

E2 | f_x =A2*EXP(-D2*B2) ← ***Fórmula***

	A	B	C	D	E	F
1	PR	SE	Alfa	Z	**Lím.Inf**	**Lím.Sup**
2	8,1818	1,0532	0,05	1,96	1,03835754	64,4689798

Las columnas A a D contienen los valores necesarios para realizar el cálculo.
En la casilla E2 se introduce el signo = para escribir la fórmula del límite inferior que usa la función exponencial EXP( )

D5 | f_x =INV.NORM.ESTAND(1-C5/2) ← ***Fórmula***

	A	B	C	D	E	F
1	PR	SE	Alfa	Z	**Lím.Inf**	**Lím.Sup**
2	8,1818	1,0532	0,05	1,96	1,03835754	64,4689798
3						
4	PR	SE	Alfa	Z	**Lím.Inf**	**Lím.Sup**
5	8,1818	1,0532	0,05	1,95996	1,03839693	64,4665345

La fila 5 contiene una mejora del cálculo. En lugar de introducir 1.96 este valor se calcula de forma exacta a partir de α=0.05 con la función INV.NORM.ESTAND( ). Esta función necesita el valor del área acumulada, que para una confianza del 95% deja el 2.5% por cada lado, y vale: 1 – 0.05/2 = 0.975.

B10 | f_x =RAIZ(1/A8-1/B8+1/C8-1/D8) ← ***Fórmula***

	A	B	C	D	E	F
1	PR	SE	Alfa	Z	**Lím.Inf**	**Lím.Sup**
2	8,1818	1,0532	0,05	1,96	1,03835754	64,4689798
3						
4	PR	SE	Alfa	Z	**Lím.Inf**	**Lím.Sup**
5	8,1818	1,0532	0,05	1,95996	1,03839693	64,4665345
6						
7	A1	N1	A0	N0		
8	9	1100	1	1000		
9	PR	SE	Alfa	Z	**Lím.Inf**	**Lím.Sup**
10	8,1818	1,0532	0,05	1,95996	1,0384265	64,464985

En la fila 9 se sitúan los números de casos en los expuestos y en los no expuestos.
A partir de estos casos, la fila 10 contiene el cálculo completo.
La casilla A10 contiene la fórmula de la razón de prevalencias PR y la casilla B10 la del error estándar que se calcula con la función raíz cuadrada RAIZ( )

Ejemplo 8: Momentos S1 a S4 de una distribución

Los pesos (x_i) de una muestra de n=5 sujetos son: 58; 60; 48; 60 y 54 kg.

Calcular los cuatro primeros momentos (m_1 a m_4) de esta distribución con las siguientes fórmulas:

$$S1 = \sum x_i \quad \rightarrow \quad m_1 = S1/n = \bar{x}$$

$$S2 = \sum (x_i - \bar{x})^2 \quad \rightarrow \quad m_2 = S2/n$$

$$S3 = \sum (x_i - \bar{x})^3 \quad \rightarrow \quad m_3 = S3/n$$

$$S4 = \sum (x_i - \bar{x})^4 \quad \rightarrow \quad m_4 = S4/n$$

B7 =SUMA(B2:B6) ← ***Fórmula***

	A	B	C	D	E
1	Id	Peso	S2	S3	S4
2	1	58			
3	2	60			
4	3	48			
5	4	60			
6	5	54			
7	S:	280			
8	m:				

Situarse en la casilla B7, pulsar el botón fx *y elegir la función suma* SUMA( ). *Señalar con el ratón el conjunto de los 5 pesos (casillas B2 a B6) para introducirlos en el cuadro de diálogo de la función* SUMA *y pulsar* Aceptar.

B8 =B7/$A$6 ← ***Fórmula***

	A	B	C	D	E
1	Id	Peso	S2	S3	S4
2	1	58			
3	2	60			
4	3	48			
5	4	60			
6	5	54			
7	S:	280			
8	m:	56			

En la casilla B8 se introduce el signo = y se divide la suma de la casilla B7 por el total de sujetos de la casilla A6. Para lograr que la casilla con el total de sujetos siempre sea la misma cuando se expanda la fórmula, delante de A y 6 se añade el símbolo $.

E2 =(B2-$B$8)^4 ← ***Fórmula***

	A	B	C	D	E
1	Id	Peso	S2	S3	S4
2	1	58	4	8	16
3	2	60	16	64	
4	3	48	64	-512	
5	4	60	16	64	
6	5	54	4	-8	
7	S:	280	104	-384	
8	m:	56	20,8	-76,8	

En la casilla E2, introducir el signo = y la diferencia entre el peso (casilla B2) y la media (casilla B8) elevada a 4. Se añade el signo $ delante de B y de 8 para fijar la casilla al expandir la fórmula. Situarse en el ángulo inferior derecho y cuando aparece ✚ *pulsar el botón izquierdo del ratón y arrastrar hacia abajo para expandir la fórmula al resto de casillas. Esta acción introducirá:*

- *en la casilla E3:* =(B3-$B$8)
- *en la casilla E4:* =(B4-$B$8)

etc.

D8 =D7/$A$6 ← ***Fórmula***

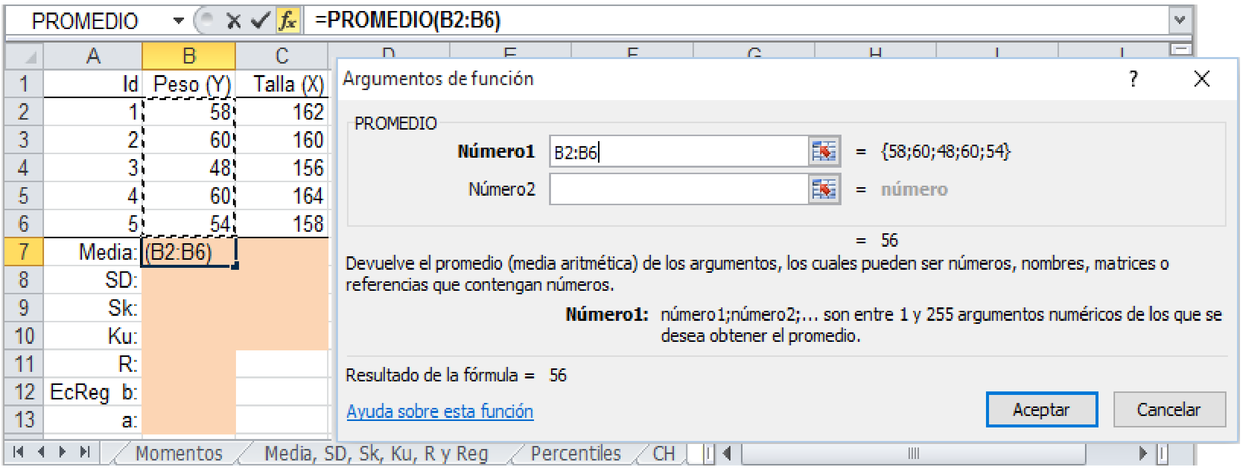
	A	B	C	D	E
1	Id	Peso	S2	S3	S4
2	1	58	4	8	16
3	2	60	16	64	256
4	3	48	64	-512	4096
5	4	60	16	64	256
6	5	54	4	-8	16
7	S:	280	104	-384	4640
8	m:	56	20,8	-76,8	

Situarse en el ángulo inferior derecho y cuando aparece ✚ *pulsar el botón izquierdo del ratón y arrastrar hacia la derecha para expandir la fórmula a la casilla contigua. La casilla D7 pasará a E7, pero la casilla A6 con el número de casos no cambiará porque delante de A y de 6 se ha situado el símbolo $. Se obtendrá el valor* **928.**

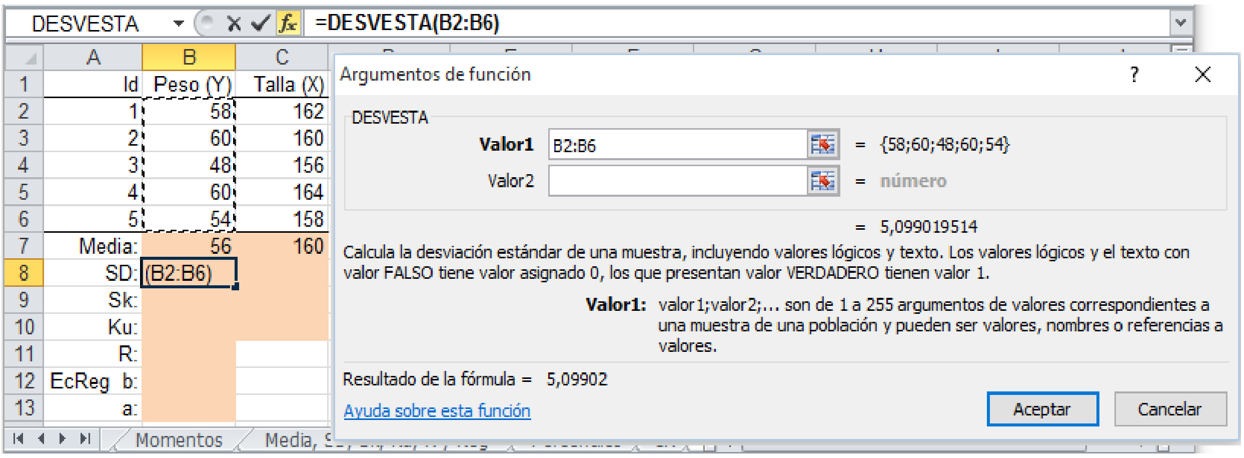
Ejemplo 9: Media, desviación estándar, asimetría, curtosis, correlación y regresión

Calcular la correlación y la ecuación de regresión entre el peso (y_i) y la talla (x_i) de una muestra de n=5 sujetos cuyos pares de valores son: (58; 162), (60; 160), (48; 156), (60; 164) y (54; 158). Calcular también la media, la desviación estándar (SD) y los coeficientes de asimetría (Sk) y curtosis (Ku) de las variables Peso y Talla.

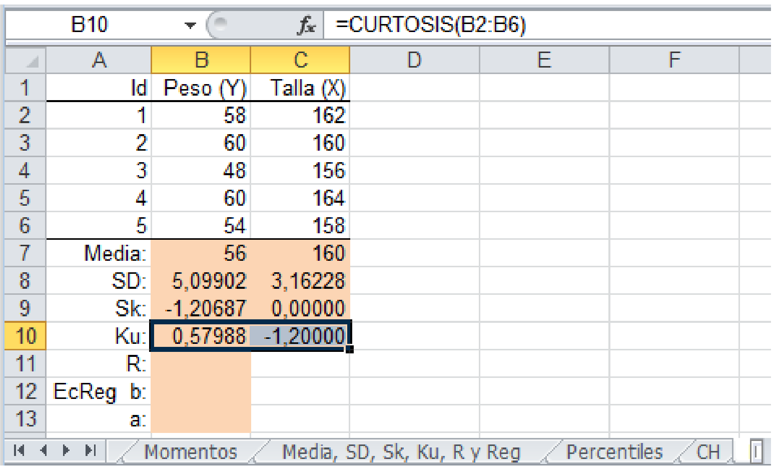
1) Situarse en la casilla B7, pulsar el botón f_x y elegir la función PROMEDIO. Señalar con el ratón el conjunto de pesos (casillas B2 a B6) y pulsar Aceptar.
La media aparece en la casilla B7 y luego se expande a la casilla C7 para obtener la media de la talla.



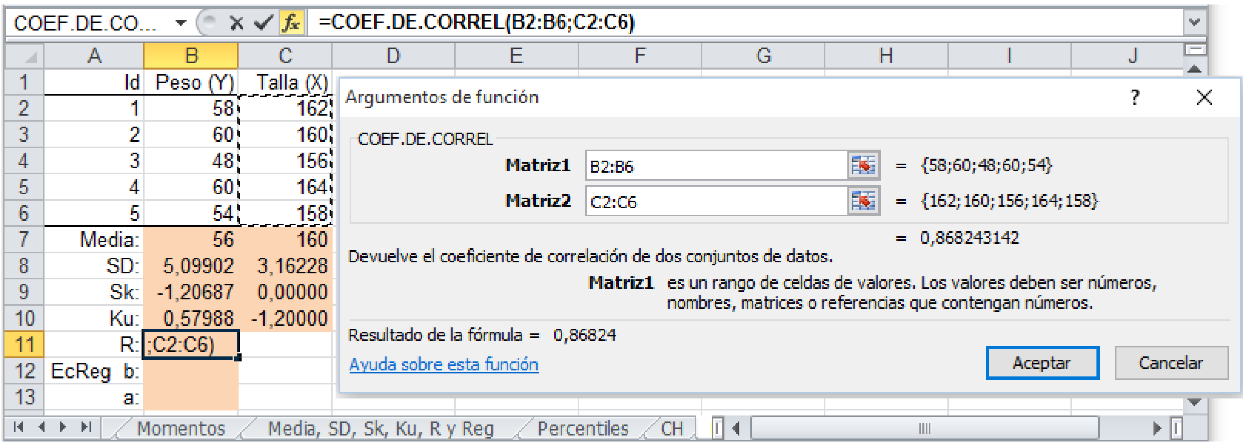
2) Situarse en la casilla B8, pulsar el botón f_x y elegir la función DESVESTA. Señalar con el ratón el conjunto de pesos (casillas B2 a B6) y pulsar Aceptar. Expandir la casilla B8 a la C8.



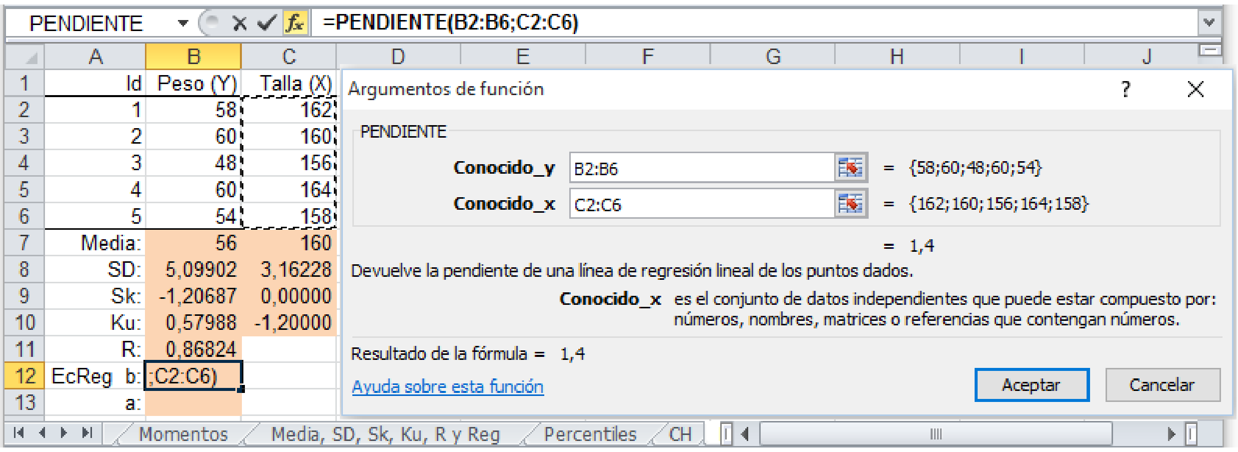
3) Repetir el proceso anterior eligiendo las funciones COEFICIENTE.ASIMETRIA y CURTOSIS



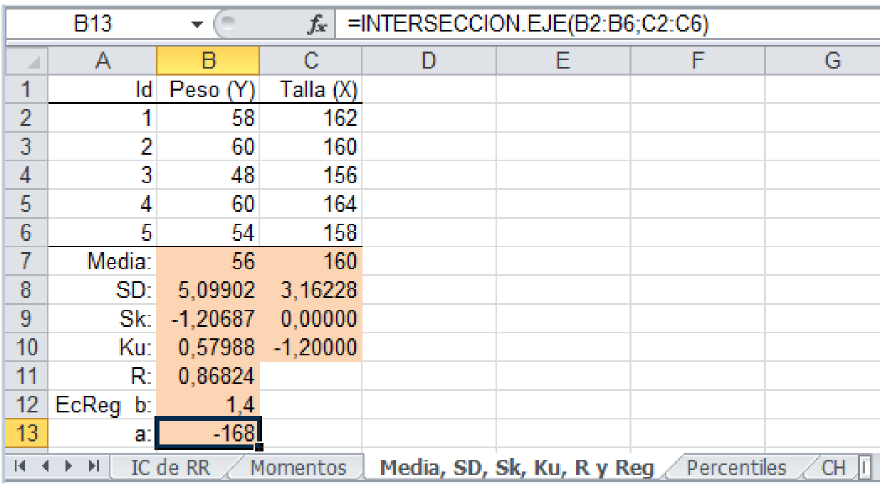
4) Situarse en la casilla B11, pulsar el botón y elegir la función COEF.DE.CORREL.
Con el cursor en **Matriz1** señalar con el ratón el conjunto de pesos (casillas B2 a B6);
con el cursor en **Matriz2** señalar con el ratón el conjunto de tallas (casillas C2 a C6). Pulsar Aceptar.



5) Situarse en la casilla B12, pulsar el botón y elegir la función PENDIENTE
Con el cursor en **Conocido_y** señalar con el ratón el conjunto de pesos (casillas B2 a B6);
con el cursor en **Conocido_x** señalar con el ratón el conjunto de tallas (casillas C2 a C6). Pulsar Aceptar.



6) Repetir el proceso anterior eligiendo la función INTERSECCION.EJE

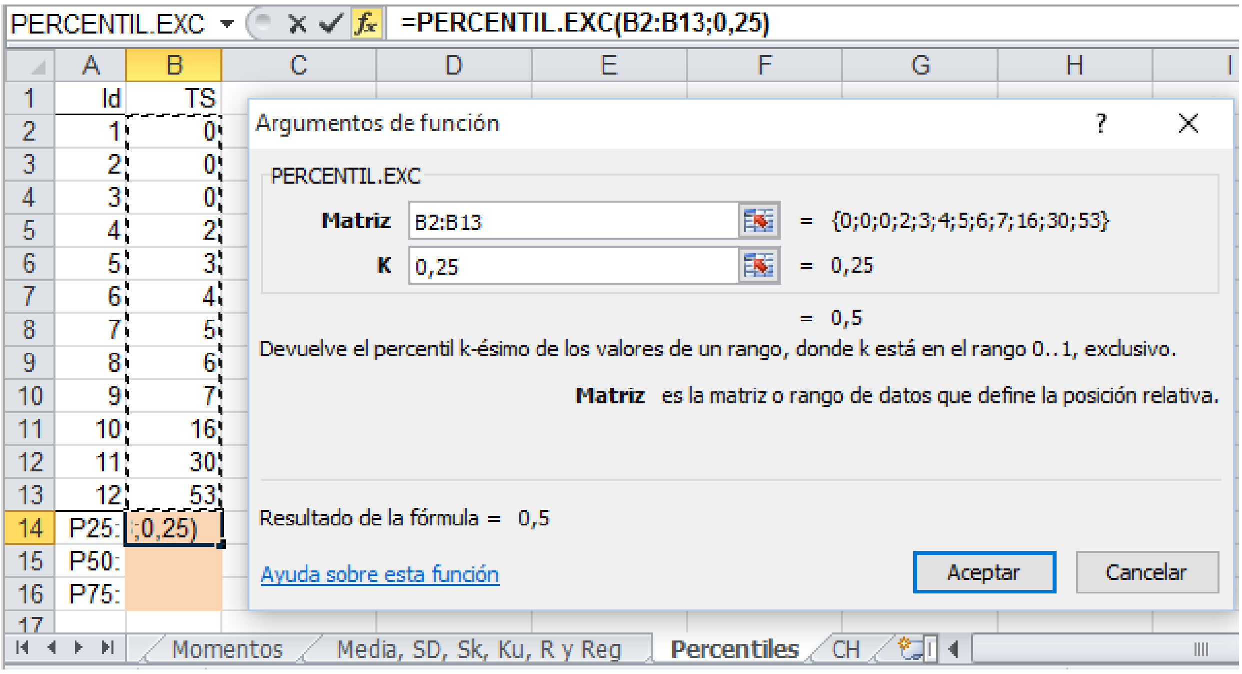


Ecuación de regresión: Peso = –168 + 1.4×Talla

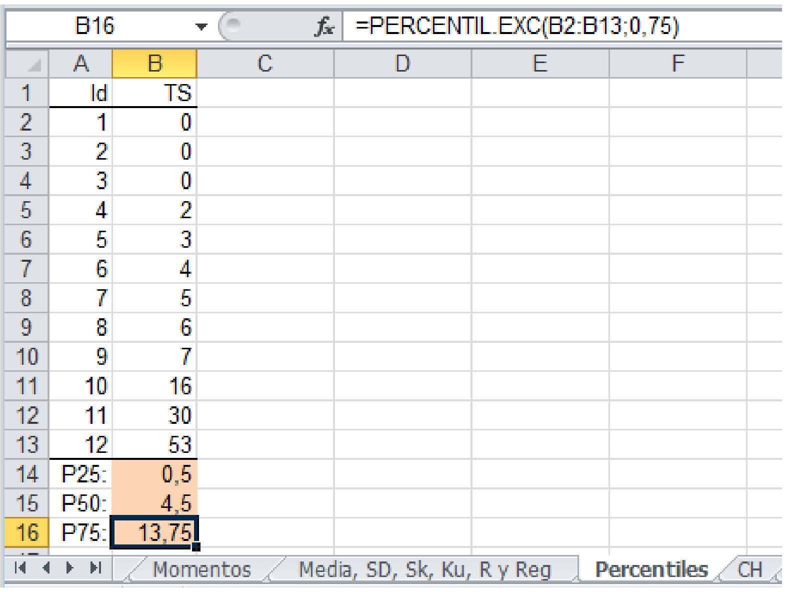
Ejemplo 10: Percentiles por el método del promedio ponderado

Calcular los percentiles 25, 50 y 75 de la distribución del tiempo de supervivencia (meses cumplidos) de 12 pacientes. Los percentiles en Excel se calculan por el **método del promedio ponderado**.

1) Situarse en la casilla B14, pulsar el botón *fx* y elegir la función PERCENTIL.EXC. Con el cursor en **Matriz** señalar con el ratón el conjunto de tiempos (casillas B2 a B13). Con el cursor en **K** indicar el rango del percentil en tantos por 1 (0,25). Pulsar Aceptar.



2) Repetir el proceso anterior indicando los rangos de percentil 0,50 y 0,75.



Nota: Los valores obtenidos deben incrementarse en 0.5 meses porque la variable TS se ha registrado en meses cumplidos: P25 = 1 mes; P50 = 5 meses; P75 = 14,25 meses.

Cálculos con la calculadora de *Stata* - Opción recomendada

Es la forma **más recomendable de realizar los cálculos** porque dispone del comando **`display`** (**`di`** de forma abreviada) al que se accede con el menú: **Data > Other utilities > Hand calculator** y que permite realizar todo tipo de cálculos aritméticos usando, si es necesario, el amplio conjunto de funciones de *Stata*. Este comando se explica en el curso de *Stata*. Además:

1) Los valores y las fórmulas que se introducen quedan permanentemente guardados en el *Review* de *Stata* y se pueden recuperar en cualquier momento.
2) Los cálculos complejos se realizan con la máxima exactitud y sin problemas de redondeo
3) Los valores de los resultados se recuperan sin necesidad de introducirlos manualmente asegurando la máxima precisión.

Ejemplo 1: Cálculo del coeficiente de variación

```
use "C:\...\CH.dta"
summarize CH
```

Variable	Obs	Mean	Std. Dev.	Min	Max
Edad	28	40.44394	13.34519	20.02856	66.13665

Imprime el nombre y el valor de los escalares que guarda el comando **`summarize`** →

```
return list
```

```
scalars:
                  r(N) =  28
              r(sum_w) =  28
               r(mean) =  40.44393525804792
                r(Var) =  178.0939697402266
                 r(sd) =  13.34518526436507
                r(min) =  20.02855682373047
                r(max) =  66.13665008544922
                r(sum) =  1132.430187225342
```

Calcula el coeficiente de variación en %. →

```
display "CV = " 100*r(sd) / r(mean)
```

```
CV = 32.996753
```

Ejemplo 2: Cálculo de los coeficientes de asimetría del 50% y del 80%

Cálculo de los percentiles con el comando **`tabstat`** *y guardados con la opción* **`save`** *en* **`r(StatTotal)`**

```
tabstat Edad, statistics( p10 p25 p50 p75 p90 ) save
```

variable	p10	p25	p50	p75	p90
Edad	21.58447	31.31098	35.68253	51.87022	59.5551

Fórmulas de cálculo: $S_{50} = \dfrac{P_{75} + P_{25} - 2\times Md}{P_{75} - P_{25}}$; $S_{80} = \dfrac{P_{90} + P_{10} - 2\times Md}{P_{90} - P_{10}}$

Visualizar los estadísticos almacenados con opción **`save`** *de* **`tabstat`** *en la matriz* **`r(StatTotal)`** →

```
matrix list r(StatTotal)
```

```
r(StatTotal)[5,1]
          Edad
p10  21.584465
p25  31.310979
p50  35.682529
p75  51.870218
p90  59.555103
```

Para poder operar con estos valores primero se deben copiar en la matriz R →

```
matrix R = r(StatTotal)
```

Fórmulas de cálculo →

```
display "S50 = " ( R[4,1]+R[2,1]-2*R[3,1] )/( R[4,1]-R[2,1] )
display "S80 = " ( R[5,1]+R[1,1]-2*R[3,1] )/( R[5,1]-R[1,1] )
```

```
S50 = 0
S80 = .375
```

Ejemplo 3. Cálculo del % de cambio entre el coeficiente *b* no ajustado y ajustado estimados con regresión lineal

La siguiente sintaxis *Stata* realiza los cálculos a partir de la variable de sistema **`_b[Tabaco]`** que contiene el coeficiente de regresión ***b*** de la variable Tabaco correspondiente al último modelo estimado.

El cálculo final consiste en obtener el porcentaje de cambio entre el coeficiente de regresión b=1.87 de la variable Tabaco (efecto no ajustado) que se guarda en el escala *b*, y el coeficiente de regresión b=1.049 de la variable Tabaco ajustado por la variable de confusión Edad.

```
use "C:\...\CH.dta"
regress CH Tabaco      //Estima el modelo de regresión con la variable Tabaco
```

```
          CH |      Coef.   Std. Err.      t    P>|t|     [95% Conf. Interval]
-------------+----------------------------------------------------------------
      Tabaco |   1.870833   .6346351     2.95   0.007     .5663221    3.175345
       _cons |   2.041667   .4797391     4.26   0.000     1.055549    3.027784
------------------------------------------------------------------------------
```

```
matrix list e(b)       //Visualizar nombres y valores de los coeficientes almacenados
```

```
e(b)[1,2]
       Tabaco      _cons
y1  1.8708333  2.0416667
```

```
scalar b = _b[Tabaco]    //Guarda en el escalar b el coeficiente b=1.87 sin ajustar
```

```
regress CH Tabaco Edad   //Estima el modelo añadiendo la variable de confusión Edad
```

```
          CH |      Coef.   Std. Err.      t    P>|t|     [95% Conf. Interval]
-------------+----------------------------------------------------------------
      Tabaco |    1.04924   .4564543     2.30   0.030     .1091553    1.989326
        Edad |    .096133    .017237     5.58   0.000     .0606327    .1316334
       _cons |   -1.37685   .6945274    -1.98   0.059    -2.807256    .0535557
------------------------------------------------------------------------------
```

```
display "Cambio = " 100*abs(b - _b[Tabaco]) / _b[Tabaco] "%"  //Calcula el % de cambio
```

```
Cambio = 78.303591%
```

Redondear números

Redondear un número X (por ejemplo, 1.73506) a D decimales (por ejemplo, 2) significa escribir el valor con D decimales **que más próximo esté** del verdadero valor X.
En este ejemplo el valor redondeado es $X_R = 1.74$.

Comprobación: Comprobar que X=1.73506 redondeado a dos decimales es 1.74 y no 1.73:

Distancia de 1.73 a X: $1.73506 - 1.73 = 0.00506$

Distancia de 1.74 a X: $1.74 - 1.73506 = 0.00494$

Caso particular: ¿Qué ocurre si hay dos valores que están a igual distancia? **Por convenio** se elige el mayor (en valor absoluto).

Ejemplo 1: Redondear X= 1.835 a d=2 decimales. Puesto que 1.83 y 1.84 están a igual distancia (0.005) de X, por convenio se elige X_R =1.84 (el mayor en valor absoluto).

Ejemplo 2: Redondear X= –1.835 a d=2 decimales. Puesto que –1.83 y –1.84 están a igual distancia (0.005) de X, por convenio se elige X_R = –1.84 (el mayor en valor absoluto).

¡Atención! El valor redondeado 0.17 no es lo mismo que el valor redondeado 0.170

El valor 0.17 es el resultado de redondear un número a 2 decimales mientras que el valor 0.170 es el resultado de redondear un número a 3 decimales.

División: 517/3026 = 0.17085261... Redondeo a 2 decimales: 0.17
Redondeo a 3 decimales: 0.171

¡Atención! Nunca redondear un valor redondeado.

Sea el valor X=2.745739 que redondeado a 2 decimales es X_{R2} = 2.75 y a 1 decimal es: X_{R1} = 2.7.
Pero si se redondea X_{R2} = 2.75 a 1 decimal da: X_R = 2.8 que no coincide con X_{R1} = 2.7.

¿Cómo realizarlo en la práctica?

Los ejemplos 1 y 2 ilustran el proceso de redondeo de un número cuyo **valor exacto** tiene 5 decimales:

1) Se "extrae" el bloque de decimales que sobran (en este ejemplo, 3) y el bloque se considera como un número entero (en este ejemplo, de 3 cifras).
2) Si este número es **menor de 500**, el número resultante ya está redondeado.
3) Si este número es **mayor o igual a 500**, se redondea incrementa la última cifra en una unidad

Nota: Si el bloque extraído tiene 1 cifra, el punto de corte es 5; si tiene 2 cifras el punto de corte es 50; si tiene 3 cifras el punto de corte es 500; y así sucesivamente.

Ejemplo 1: Redondear los tres siguientes números a 2 cifras decimales.

X= 1.73|499 → $499 < 500 \rightarrow X_R = \mathbf{1.73}$

X= 1.73|500 → $500 \geq 500 \rightarrow X_R = \mathbf{1.74}$ (se incrementa la última cifra en 1 unidad)

X= 1.79|501 → $501 \geq 500 \rightarrow X_R = \mathbf{1.80}$ (se incrementa la última cifra en 1 unidad)

Ejemplo 2: Redondear los tres siguientes números a 1 cifra decimal.

X= 1.9|499 → $499 < 500 \rightarrow X_R = \mathbf{1.9}$

X= 1.9|500 → $500 \geq 500 \rightarrow X_R = \mathbf{2.0}$ (se incrementa la última cifra en 1 unidad)

X= 1.9|501 → $500 \geq 500 \rightarrow X_R = \mathbf{2.0}$ (se incrementa la última cifra en 1 unidad)

Ejemplo 3 Redondear el número exacto X=1.7445 a 3, 2 y 1 cifra decimal.

X= 1.744|5 → $5 \geq 5 \rightarrow X_R = \mathbf{1.745}$ (se incrementa la última cifra en 1 unidad)

X= 1.74|45 → $45 < 50 \rightarrow X_R = \mathbf{1.74}$

X= 1.7|445 → $445 < 500 \rightarrow X_R = \mathbf{1.7}$