Estadística con GuadaLinex Edu

Índice

1. Introducción

GNULinux y Matemáticas, un binomio casi perfecto. Si nos centramos en Estadística y Probabilidad, el "matrimonio de hecho" funciona igualmente bien.

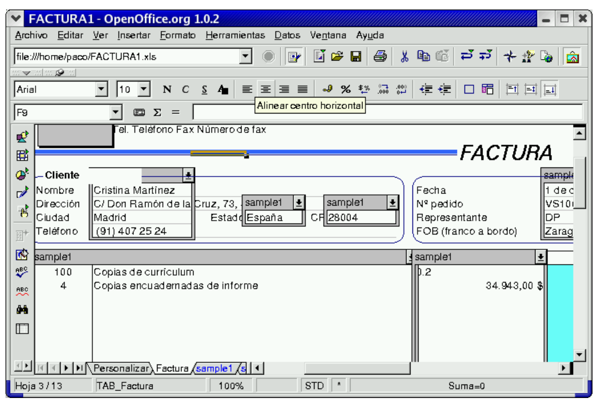
Entre la multitud de herramientas del software libre, vamos a comentar las que siguen[1]:

- Hojas de cálculo: Gnumeric y calc del paquete OpenOffice
- Programas específicos:
  - De gráficos: grace
  - El genuino: R

2. Hojas de cálculo

2.1. OpenOffice Calc

Esta hoja de cálculo es muy similar a la Excel. En sus celdas podemos introducir texto, números o fórmulas con referencias a otras celdas para que la aplicación realice los cálculos establecidos. El programa incorpora también una amplia gama de funciones para análisis estadísticos y puede importar hojas de cálculo externas.



Las hojas de cálculo pueden servir como fuente de datos para generar informes o cartas en serie y para la elaboración de gráficos y diagramas. El programa puede utilizar también datos procedentes de una base de datos externa.

En el apéndice C en la página 87 se listan las funciones de esta hoja de cálculo. Las posibiliades gráficas de Calc sun superiores que las de Gnumeric, sin embargo la facilidad de uso de esta última me hace decantarme por ella, así que:

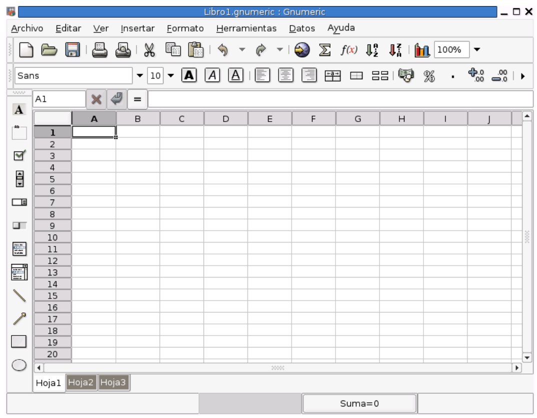
2.2. Gnumeric

Si lo que necesita es una hoja de cálculo, **Gnumeric** es magnífica. Forma parte de la suite ofimática de Gnome Office y es fácil de utilizar.

Para cargar el programa **Aplicaciones→Oficina→Gnumeric** o desde una xterm

[1]
- Aquellos que están a nuestra disposición en los Centros TIC. Se quedan en el tintero KChart (muy simple y sólo para representar gráficos) así como los programas de cálculo simbólico: maxima, octave, ...
- Todos ellos están disponibles para entornos Windows.

```
$ gnumeric &.
```



La ventana es muy intuitiva, semejante a la de otros programas de estas características y no nos vamos a parar a describirla. La página oficial del programa es `http://www.gnome.org/projects/gnumeric` en ella encontrareis un manual, FAQ, ... (en inglés).

Comentar que ofrece compatibilidad con los formatos de Excel, Lotus y por supuesto con la hoja de cálculo de OpenOffice.

El trabajo lo vamos a realizar, siempre que se pueda, sin echar mano de las funciones de esta hoja de cálculo. Un listado de las funciones estadísticas está en este documento en el apéndice B en la página 72.

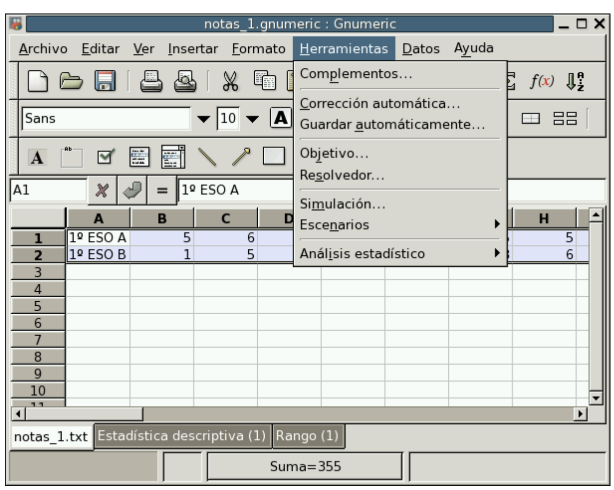
2.2.1. Estadística descriptica

➥ **Para practicar:** Las notas de matemáticas de 1º ESO A y B han sido:

A	5, 6, 4, 5, 6, 5, 5, 8, 6, 3, 2, 1, 5, 5, 7, 6, 7, 3, 2, 3, 8, 9, 4, 4, 8, 9, 1, 5, 5, 4, 7, 6, 3, 4, 5, 5
B	1, 5, 7, 4, 3, 8, 6, 5, 4, 7, 3, 7, 1, 2, 4, 8, 9, 7, 8, 1, 2, 4, 5, 6, 8, 2, 3, 2, 8, 4, 5, 6, 5, 6, 8

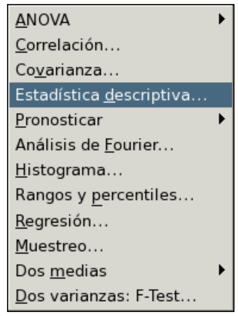
¿Qué clase obtuvo mejores resultados? Realiza un estudio estadístico de las notas de ambas clases.

Introducimos los datos en Gnumeric (mejor en columnas y no tal cual está la captura[2]) y tras marcar el bloque de ambas listas de datos optamos por pulsar sobre:



[2]Para transponer filas a columnas sólo hemos de seleccionar las filas, cortarlas y tras seleccionar pegar especial, optar por transponer.

y analicemos qué nos ofrece para nuestras clases de secundaria[3]:



De las disponibles vamos a ver:

Correlación Calcula los coeficientes de correlación de Pearson de las variables dadas.

Covarianza Calcula la matriz de covarianzas de las variables.

Estadística descriptiva Calcula diversos parámetros estadísticos: media, error estándar, mediana, moda, desviación estándar, varianza simple, curtosis, asimetría, rango, mínimo, máximo, suma y conteo.

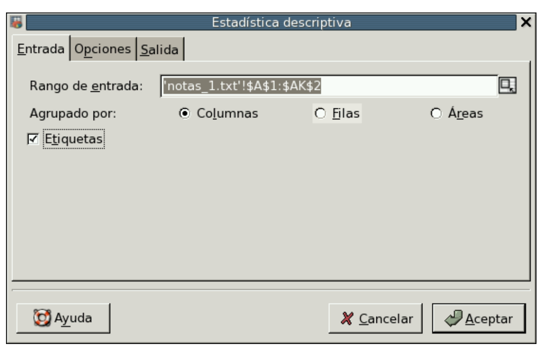
Histograma Diversos tipos de tablas de frecuencias para una o más variables

Rangos y percentiles Calcula el rango y percentiles de los datos.

Regresión Análisis de regresión múltiple: coeficiente de correlación lineal, pendiente y ordenada en el origen de las rectas de regresión ...

Muestreo Permite obtener una muestra de un conjunto de datos. El muestreo puede ser aleatorio o periódico.

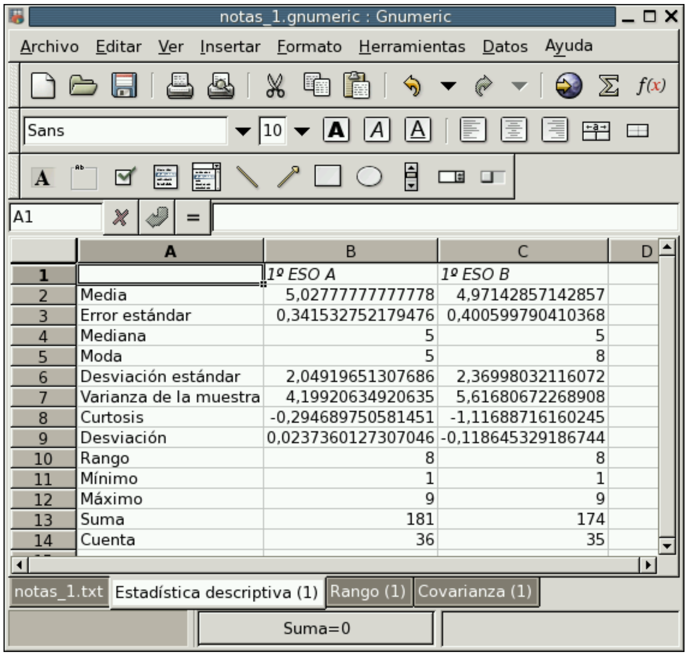
Comenzaremos marcando el bloque del que realizar el análisis (etiquetas incluidas) y una vez seleccionada la opción de **Estadística Descriptiva** optamos por marcar la casilla de etiquetas y si nuestros datos están en Columnas, Filas o Áreas. El resto de opciones no es necesario tocarlas ya que los valores por defecto son por ahora más que suficientes:



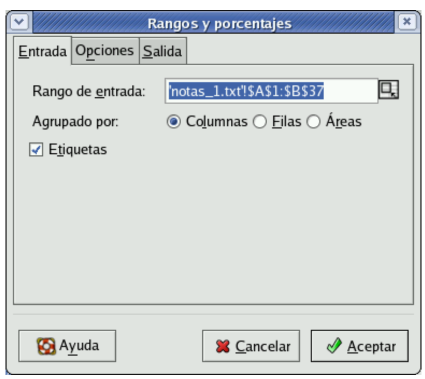
La salida da bastante información, nos aparece la media, error estándar ($\frac{S}{\sqrt{N}}$), mediana, moda, desviación estándar (cuasidesviación típica), varianza de la muestra (cuasivarianza)[4], curtosis (coeficiente de apuntamiento), desviación (coeficiente de asimetría), rango, valores mínimo y máximo, suma y nº de datos.

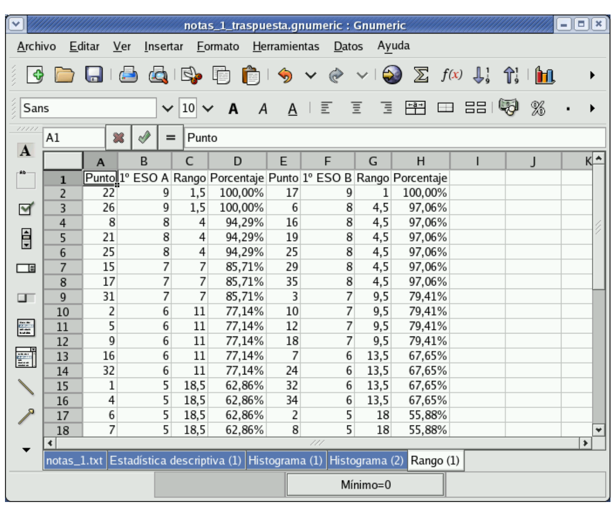
[3]Merece la pena pararse en la parte inferior de Gnumeric.

[4]Para obtener la varianza y la desviación típica tendremos que usar las funciones `varp` y `stdevp` respectivamente.



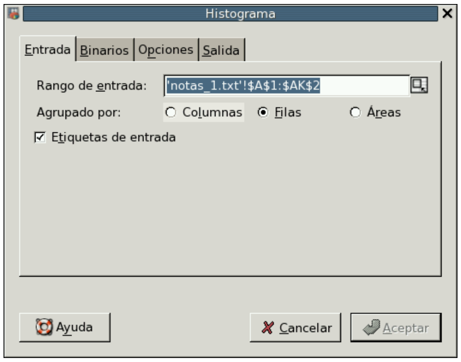
Antes de seguir: ¿qué clase obtuvo mejores resultados? ¿podemos responder ya la pregunta? Más posibilidades:

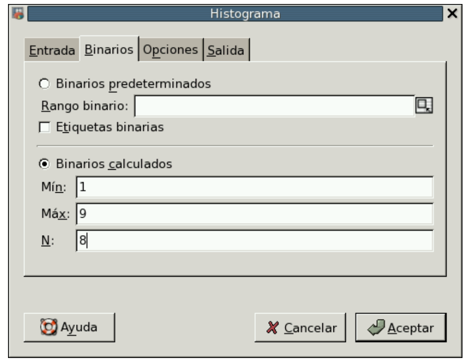




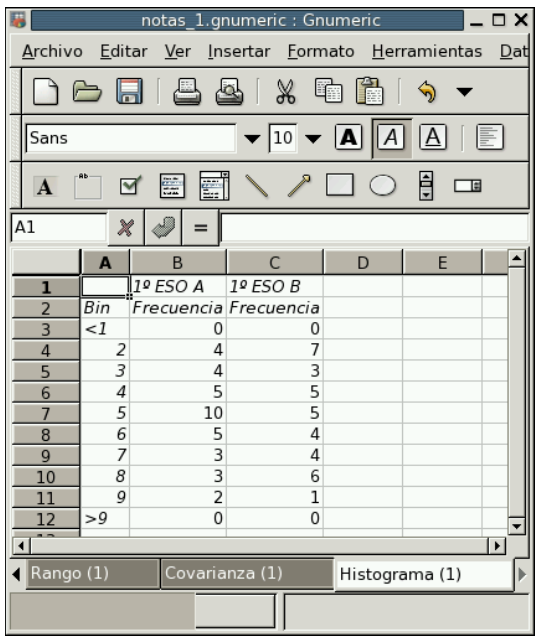
La interpretación de la captura es: en la primera columna la posición que ocupa el dato en la lista original, en la segunda los datos ordenados de mayor a menor, en la tercera la media aritmética de la posición de los valores ordenados, en la cuarta el percentil que ocupa el dato.

Otra opción, optemos por histograma[5]





La salida



	A	B	C
1		1º ESO A	1º ESO B
2	Bin	Frecuencia	Frecuencia
3	<1	0	0
4	2	4	7
5	3	4	3
6	4	5	5
7	5	10	5
8	6	5	4
9	7	3	4
10	8	3	6
11	9	2	1
12	>9	0	0

Pasa algo raro verdad, ¿qué ha pasado con la nota 1 y 2?, esto no está bien, la frecuencia absoluta de la nota 1 no está bien calculada. Vamos a intentar arreglarlo:

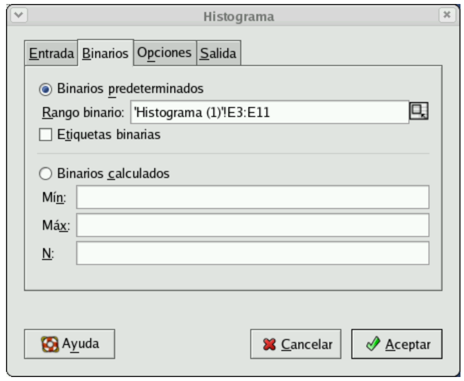
1. Podemos ampliar los valores de Mínimo y Máximo, por ejemplo: Min=0, Max= 9, y establecer en 9 el valor de N

$$N^o\, valores = \frac{Max - Min}{N} = \frac{9-0}{9} = 9$$

2. Introduzcamos en la hoja de cálculo las etiquetas que vamos a usar (del 1 al 9[6]) y ahora, tras seleccionar **Histograma**, optemos por escribir:

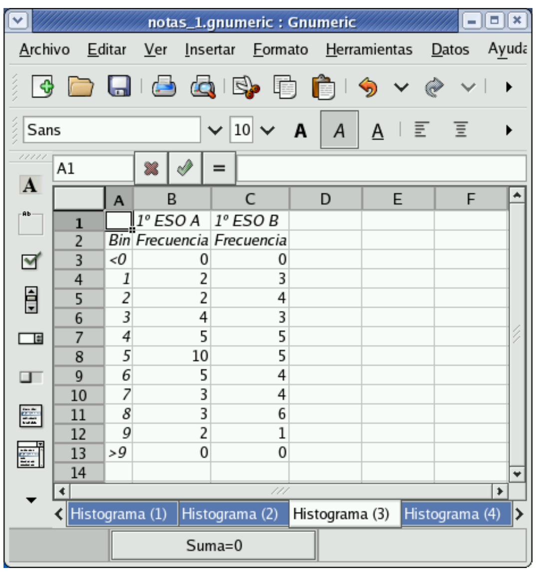
[5] Cuidado en marcar bien las casillas de **Columnas/Filas** y en su caso **Etiquetas de entrada.**

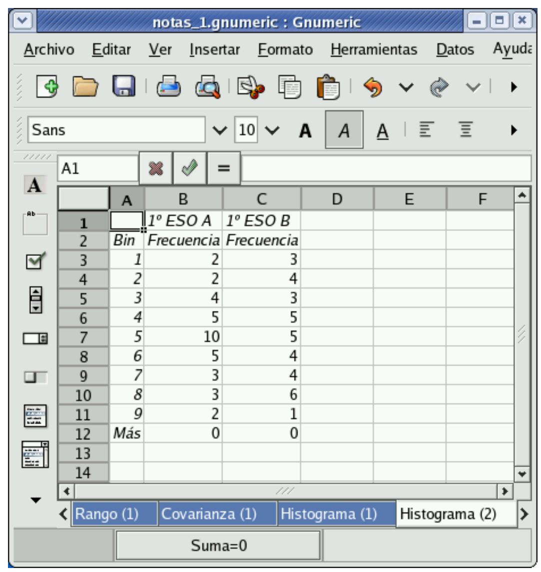
[6] Notar que se puede hacer de forma automática desde **Editar⟶Llenar⟶Series**



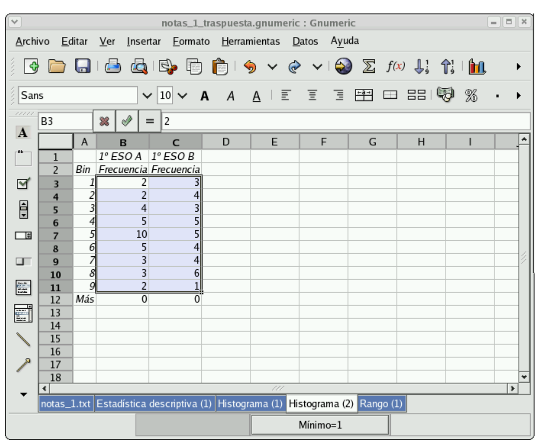
Lo único a tener en cuenta es que en rango Binario hemos de seleccionar el conjunto de celdas que contiene las 9 etiquetas a usar como notas.

El resultado, con cualquiera de los dos métodos anteriores ya sí es el esperado:



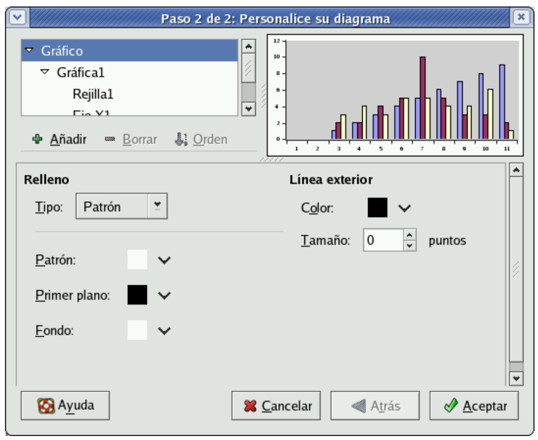


Hagamos nuestro diagrama de barras:

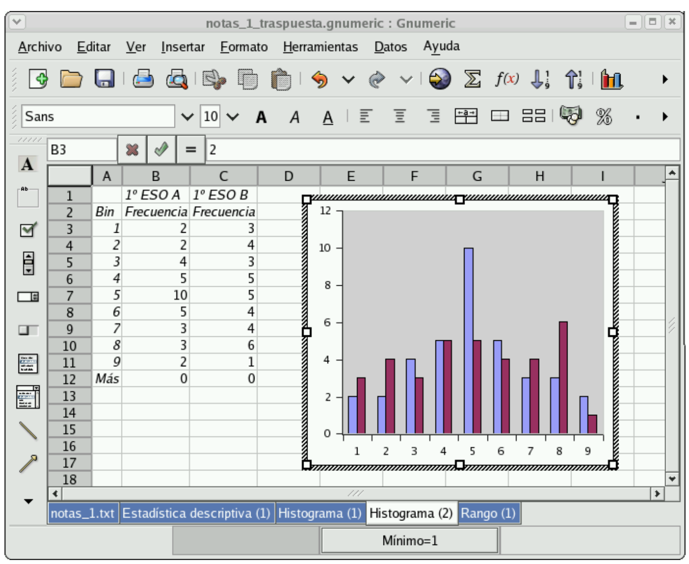


y pulsamos sobre 

Un comentario: una vez que optemos por un tipo de gráfico, no podemos optar por otro. Sí que podremos adecuarlo a nuestro gusto con la ventana

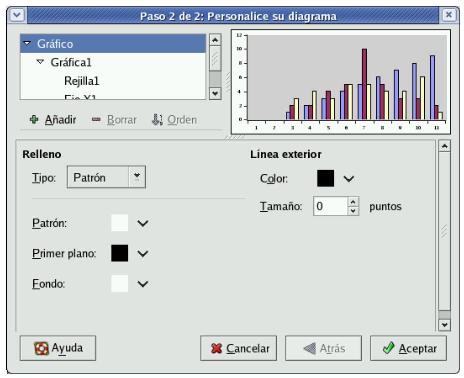


Para que se vea el gráfico, marcamos sobre Gnumeric la zona rectangular en la que deseamos que aparezca y

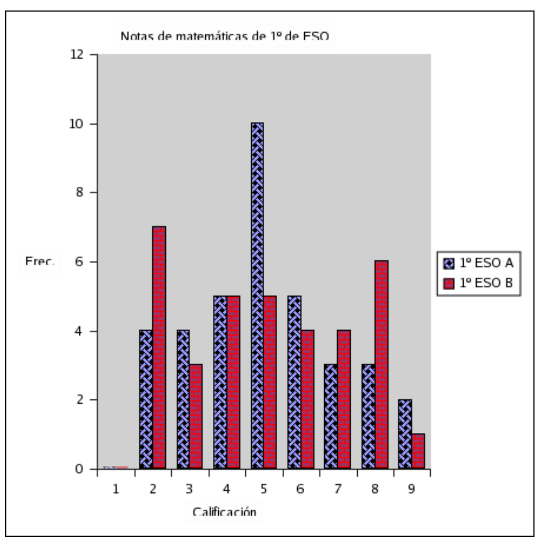


Para modificar nuestro gráfico, pulsaremos sobre él con el botón derecho del ratón y tras optar por

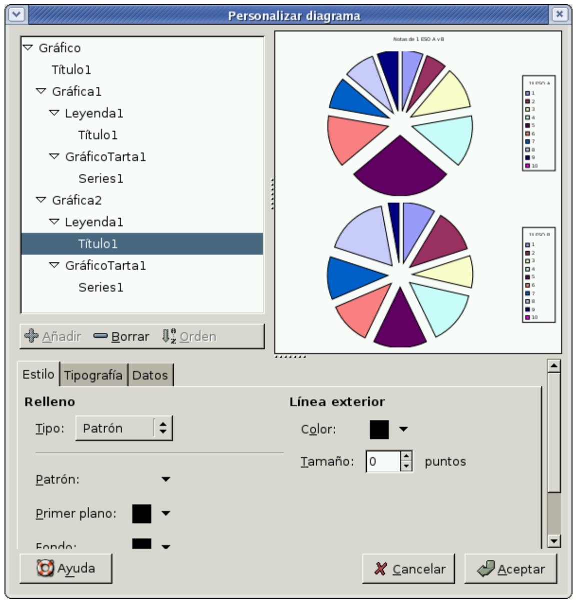
obtendremos de nuevo



Desde este punto y añadiendo los elementos necesarios, hemos de conseguir que quede de la forma:



Ejercicio: Realizar un gráfico de sectores como el que sigue:



Ejercicio La edad de 130 personas que realizaron un experimento se da a continuación:

15	17	19	23	22	17	19	24	15	17	23	19	23	18	17	16	24	19	17	18
15	24	19	17	22	19	19	20	21	22	17	17	21	16	16	18	24	21	19	19
19	20	21	21	23	18	23	18	17	17	20	23	17	21	23	17	16	19		
21	17	18	21	19	18	21	22	17	19	24	20	21	21	16	17	19	19		
22	17	21	18	23	18	16	23	18	17	22	19	15	17	19	17	16	22		
16	23	18	17	21	24	21	21	18	18	19	24	16	22	23	19	17	20		
20	22	22	17	17	19	19	15	20	19	18	20	17	15	17	18	18	21		

1. Agrupa las edades por años.
2. Da una tabla de frecuencias simple (para cada año) y acumulada.
3. Realiza un diagrama de barras verticales y uno de sectores.

2.2.2. Correlación

➥ **Para practicar:** La tabla siguiente muestra las respectivas alturas X e Y de una muestra de 12 padres y sus hijos primogénitos (en pulgadas):

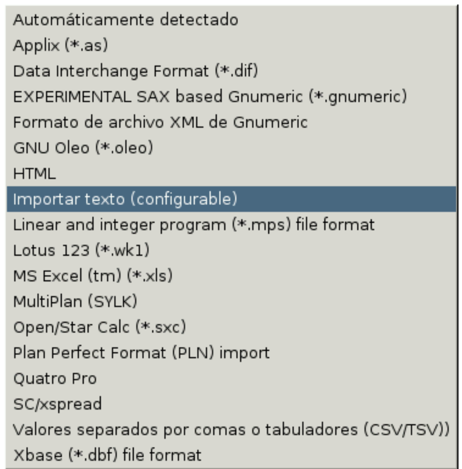
Altura X del padre (en pulg)	65	63	67	64	68	62	70	66	68	67	69	71
Altura Y del hijo (en pulg)	68	66	68	65	69	66	68	65	71	67	68	70

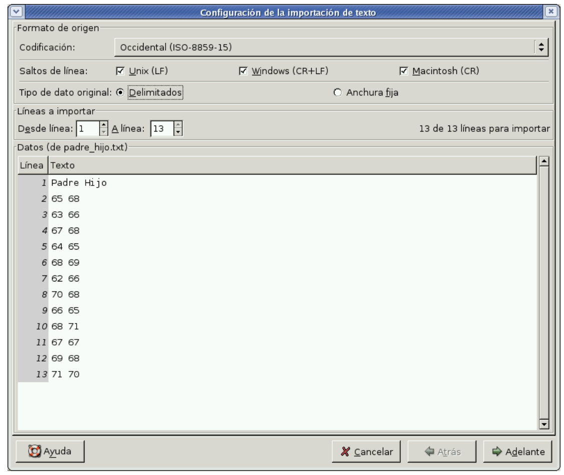
1. Construye el diagrama de dispersión.
2. Estudia la correlación entre ambas variables.
3. Halla la recta de regresión de Y sobre X.

Solución:

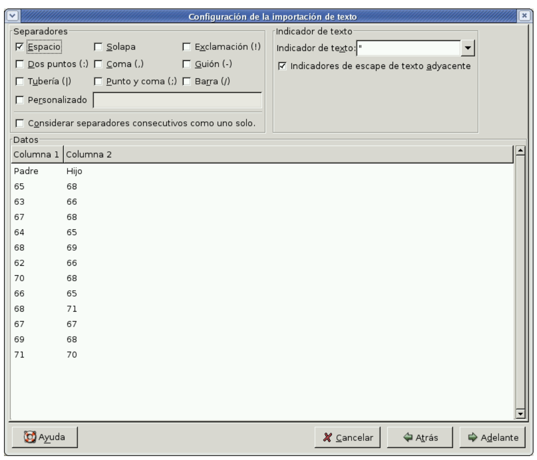
Lo mejor es importar los datos desde texto plano, de esa forma los podemos usar en todos los programas[7]. Para leer un fichero en texto plano con Gnumeric, pulsamos sobre **Abrir Archivo** y en la ventana que aparece optamos por seleccionar el **Tipo de Archivo**

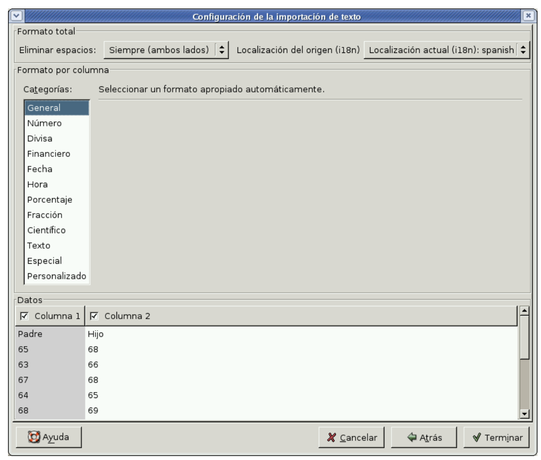
[7] Dependiendo de la versión de Gnumeric puede que tengamos que hacerlo desde **Datos→ Obtener datos externos→ Obtener archivo de texto**



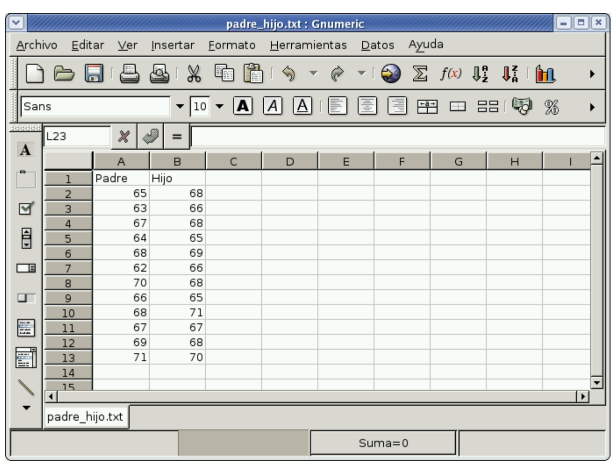


Sólo tenemos que ajustar el formato de los datos que estamos exportando en las ventanas que vamos obteniendo

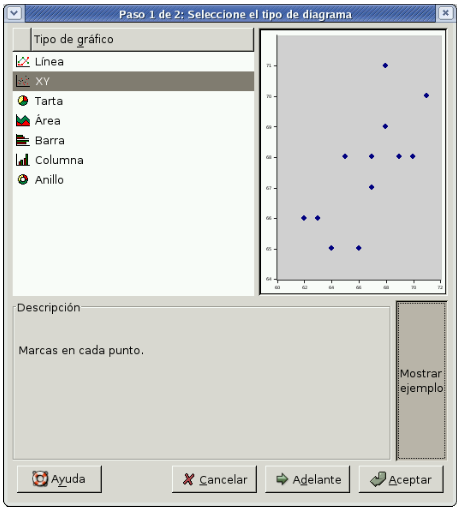


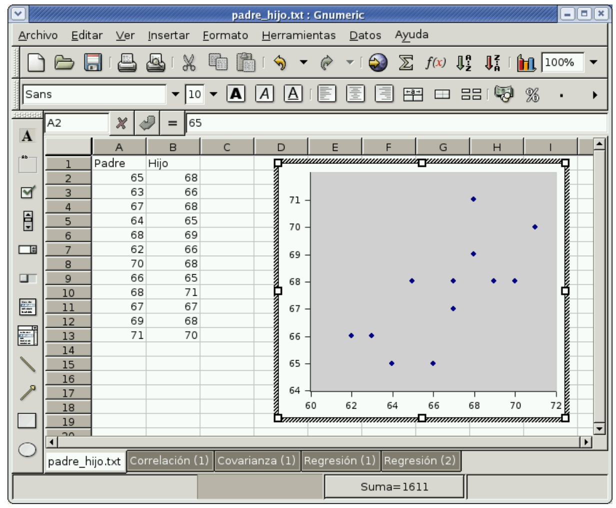


El resultado final, el mismo que si los introducimos desde Gnumeric:

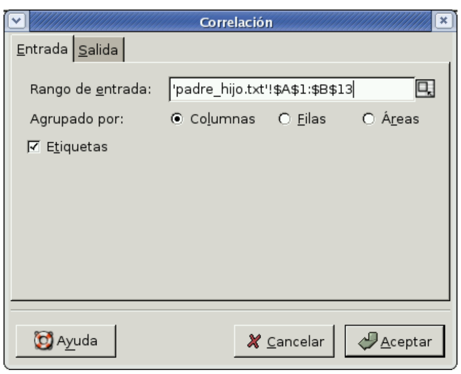


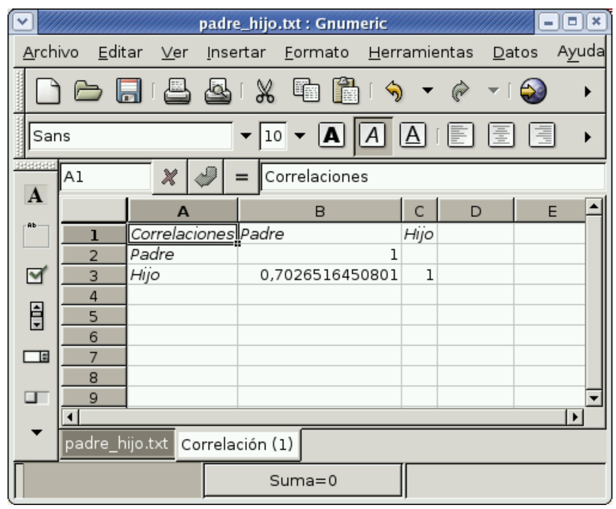
Nos toca el apartado gráfico, seleccionamos las columnas de datos (sin los nombres de las variables) y pulsamos sobre , sólo hay que optar por la segunda opción



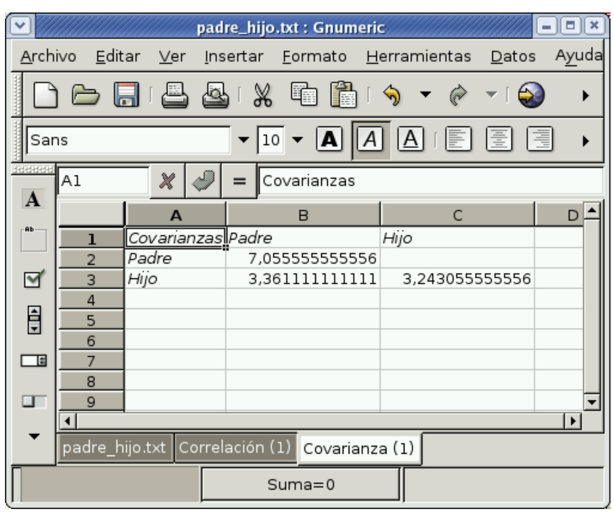


Ahora seleccionamos de nuevo las columnas de datos, pero esta vez con los nombres de las variables, y pulsamos en **Herramientas→Análisis Estadístico→Correlación**

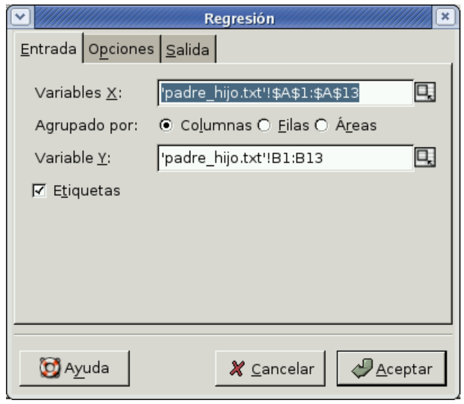




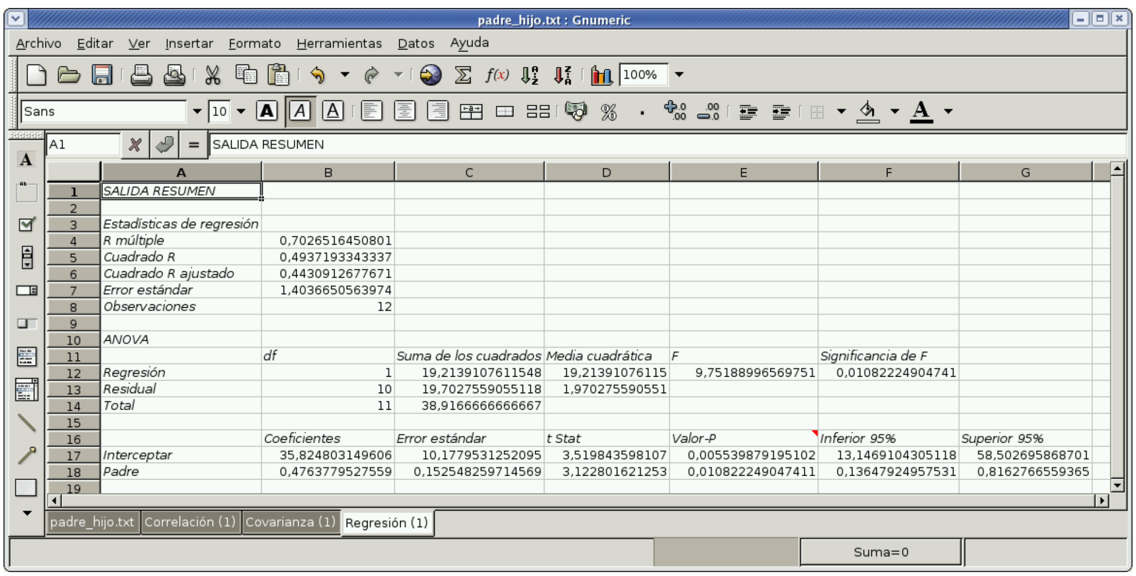
Si ahora pulsamos sobre **Herramientas→Análisis Estadístico→Covarianza**, obtendremos



Con **Herramientas→Análisis Estadístico→Regresión** y, tras ajustar los valores de entrada,



obtendremos:



De esta hoja sólo nos centraremos en aquellos resultados interesantes para Secundaria o Bachillerato. Se trata de las celdas

B4 Coeficiente de correlación lineal

B5 Coeficiente de determinación

B8 Tamaño de la muestra analizada

B17 Ordenada en el origen de la recta de regresión[8] de Y (Talla del hijo) sobre X (Talla del Padre)

B18 Pendiente de la recta de regresión

Es decir $r_{y/x}\,:\, y = 0,476\,x + 35,824$. Si ahora deseamos hacer estimaciones, usaremos la función:

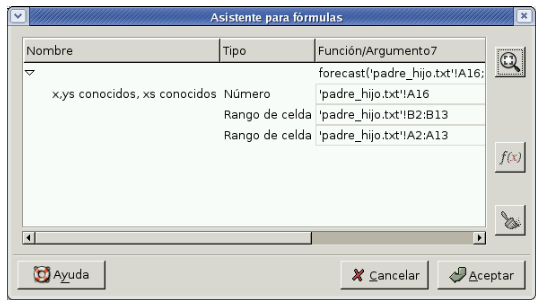
FORECAST(x;y_conocidos;x_conocidos)

La función FORECAST estima un valor futuro, según los valores existentes usando una regresión linear simple. El valor futuro estimado es el valor de y para un valor de x dado (x).

[8]Este valor y el que sigue se pueden obtener usando la función LINEST

Vamos a comprobar que la recta de regresión pasa por el punto $(\bar{x}, \bar{y})$. Para eso, escribamos el valor de $\bar{x}$ en una casilla vacía[9], después nos situaremos en donde obtendremos el resultado, pulsaremos sobre el icono[10] 

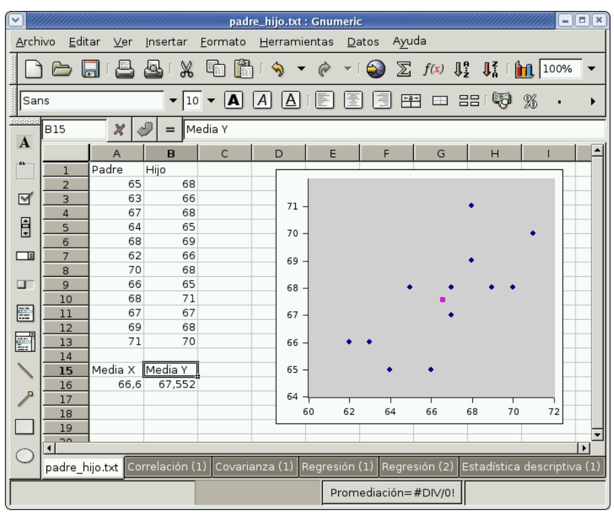
y seleccionaremos la función anterior de entre las funciones **Estadísticas**



Los valores a introducir se explican en la definición de la función, se trata de: $\bar{x}$, Y y X.

¡OJO!, tras seleccionar los valores sobre la hoja de cálculo hay que pulsar sobre la tecla **INTRO** para introducirlos.

El resultado, con el punto dibujado sobre la gráfica es:

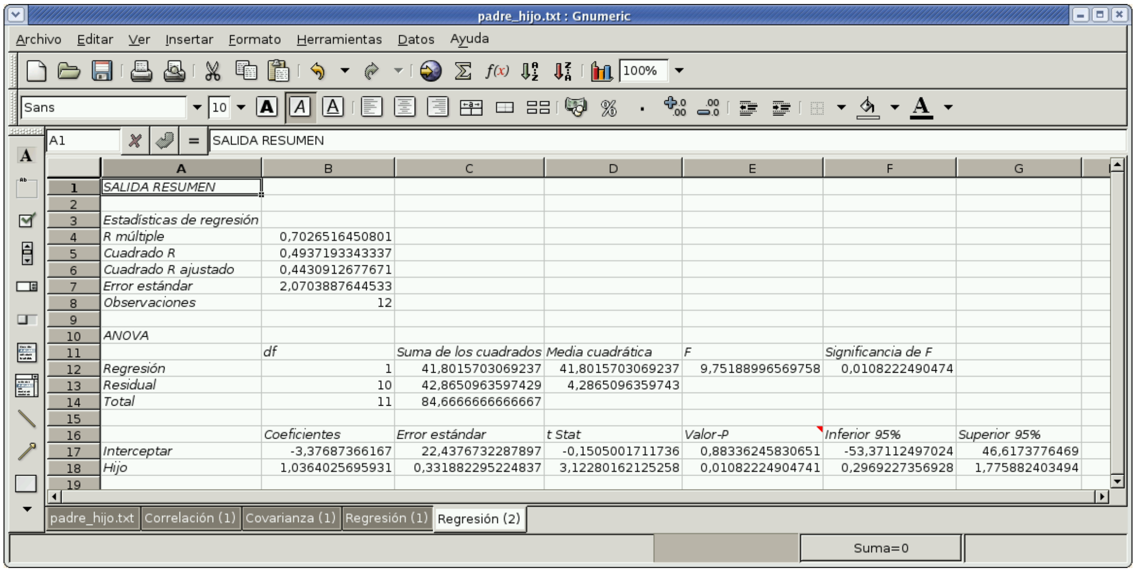


Si ahora intercambiamos las variables

[9] ¿Lo calculamos de nuevo?, ¿desde la barra de estado?, ¿con estadística descriptiva?, ¿con la función average?

[10] Este icono tiene doble funcionalidad:

- Si la celda no contiene una función se nos abre el asistente de funciones.
- Si contiene una función nos permite editar la función de la celda activa



	A	B	C	D	E	F	G
1	SALIDA RESUMEN						
2							
3	Estadísticas de regresión						
4	R múltiple	0,7026516450801					
5	Cuadrado R	0,4937193343337					
6	Cuadrado R ajustado	0,4430912677671					
7	Error estándar	2,0703887644533					
8	Observaciones	12					
9							
10	ANOVA						
11		df	Suma de los cuadrados	Media cuadrática	F	Significancia de F	
12	Regresión	1	41,8015703069237	41,8015703069237	9,75188996569758	0,0108222490474	
13	Residual	10	42,8650963597429	4,2865096359743			
14	Total	11	84,6666666666667				
15							
16		Coeficientes	Error estándar	t Stat	Valor-P	Inferior 95%	Superior 95%
17	Interceptar	-3,37687366167	22,4376732287897	-0,1505001711736	0,88336245830651	-53,37112497024	46,6173776469
18	Hijo	1,0364025695931	0,331882295224837	3,12280162125258	0,01082224904741	0,2969227356928	1,775882403494
19							

Ejercicios:

1. Obtener la Talla estimada de un hijo cuyo padre mida:

   a) 60
   b) 70
   c) 80
   d) 100

   Comenta los resultados

2. Obtener los valores estimados de la Talla de un Padre si el Hijo mide:

   a) 60
   b) 70
   c) 80
   d) 100

3. Representa la nube de puntos así como las dos rectas de regresión[11].

4. En la tabla siguiente damos la evolución del récord del mundo de salto de longitud masculina, en metros.

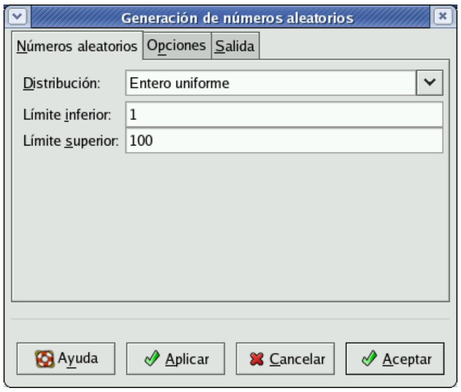
Años	1901	1921	1925	1931	1935	1961	1962	1965	1968	1991
Marca	7,61	7,69	7,89	7,98	8,13	8,28	8,31	8,35	8,90	8,95

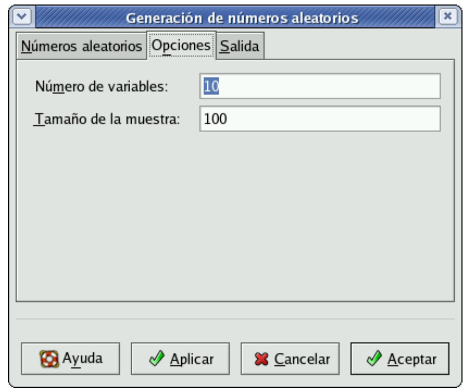
   a) Representa la nube de puntos asociada a estos datos.
   b) ¿Puede estimarse la marca del récord por regresión lineal? Si es así, ¿cuánto se saltará en los años 2000 y 2005?

[11] Que no se nos olvide que para pintar una recta sólo necesitamos dos puntos por los que pasa.

2.2.3. Muestras y números aleatorios

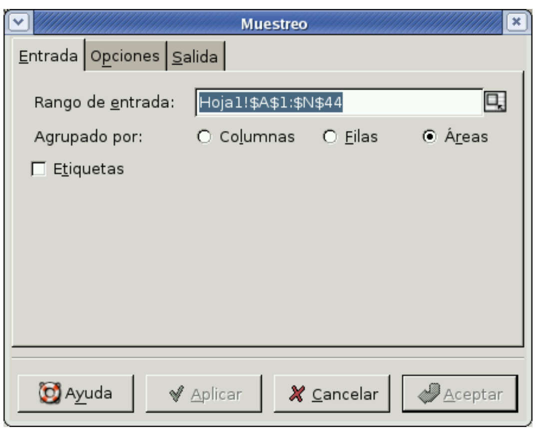
Generamos 1000 números aleatorios enteros dentro del rango 1.,100 siguiendo una distribución uniforme. Pulsamos sobre **Editar→Llenar→Generación de números aleatorios**





Como lo queremos en 10 columnas de 100 datos cada una escribiremos los valores adecuados en **Número de variables** y **Tamaño de la muestra**.

Optar por rellenar los valores de la tercera pestaña es opcional y si la dejamos como está obtendremos una tabla de 10x100 números aleatorios. Obtengamos una muestra aleatoria de tamaño 100, pulsemos sobre **Herramientas→Análisis Estadístico→Muestreo**



En una nueva Hoja obtendremos una muestra aleatoria de tamaño 100.

De esa muestra hay que hacer un análisis estadístico y realizar algunas gráficas de ella. Después, usar la muestra y lo visto antes para realizar el estudio de esos datos.

La carrera El juego que sigue se basa en las siguientes reglas. En una pista de carreras de 12 calles numeradas del 1 al 30 tenemos 12 caballos numerados del uno al doce. Cada uno de ellos corre por una de las pistas y gana el juego el que llegue primero al final.

	1	2	...	12
1				
2				
...				
30				

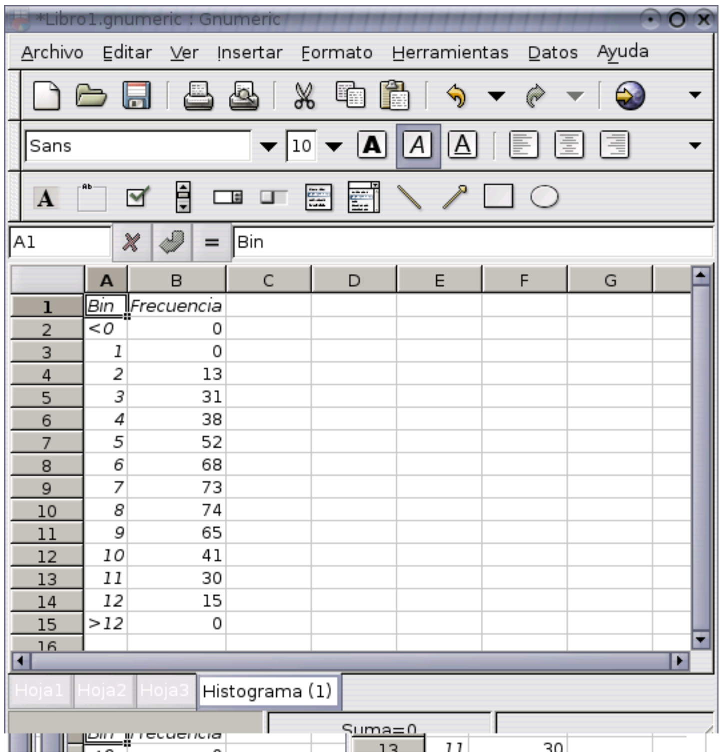
El juego consiste en lanzar dos dados y sumar las caras superiores. Al número obtenido significa que ese caballo avanzará una casilla, y así hasta que uno llegue a la meta.

¿Qué caballo elegirías?

Solución: Simulemos el juego con Gnumeric, para eso introduzcamos en la casilla **A1** la fórmula `=int(6*rand()+1)`, con ella generamos un número aleatorio comprendido entre 1 y 6. El resto es fácil, copiemos la fórmula a la casilla **B1** y sumemos el contenido de las celdas anteriores en la celda **C1**.

Si ahora copiamos el contenido del las celdas 400 o 500 veces prodremos similar ese número de lanzamientos.

Ya sólo tenemos que hacer un estudio estadístico (por ejemplo con la herramienta histograma) de la tercera columna y hallar la distribución de frecuencias, como muestra de un posible resultado:



2.2.4. Combinatoria

FACT(n)

Devuelve el factorial de n ($n!$)

PERMUT(n;k)

La función PERMUT devuelve el número de variaciones de k elementos tomados de entre n. n es el número de objetos, k es el número de objetos en cada variación.

- Si n = 0, PERMUT devuelve el error #NÚM!.
- Si n <k, PERMUT devuelve el error #NÚM!.
- Esta función es compatible con Excel.

Y dentro de **Funciones Matemáticas**

COMBIN(n;k)

Calcula la cantidad de combinaciones, C_n^k

- Efectuar esta función con números no enteros o negativos devuelve un error #NÚM!.
- Si n es menor que k devuelve un error #NÚM!.
- Esta función es compatible con Excel.

Ejercicios

1.
    a) Halla 17!
    b) Seis amigos corren en una carrera, ¿de cuántas formas pueden llegar?
2. Halla:
    a) $V_{10,\,3}$
    b) ¿Cuántos números de 5 cifras se pueden formar con los dígitos 1 al 9?
3.
    a) Halla $C_{10,\,4}$
    b) Comprueba con varios números que
        1) $C_{m,\,m-n} = C_{m,\,n}$
        2) $C_{m,\,n} + C_{m,\,n+1} = C_{m+1,\,n+1}$
    c) ¿Cuántos grupos de 4 pueden formarse en una clase de 30 alumnos?

2.2.5. Distribuciones de probabilidad

Distribución binomial La función a usar es:

BINOMDIST(n;intentos;p;acumulado) La función BINOMDIST devuelve la distribución binomial.

n es el número de éxitos,

intentos es el número total de intentos independientes,

p es la probabilidad de éxito en intentos, y

acumulado describe si debe devolver la suma de las funciones binomiales de 0 a n. Toma los valores 0 (no es acumulado) o 1 (acumulado)

Lo único a tener en cuenta es que pulsaremos en el botón  y, tras seleccionar la función anterior de entre las funciones estadísticas, tenemos que rellenar los datos adecuados: ¡ Que no se nos olvide pulsar sobre la tecla **Intro**!

Ejercicios

1. Un examen de preguntas con respuestas múltiples consta de 8 preguntas con 4 opciones de contestación. Si un alumno responde al azar, halla, con la ayuda de la tabla binomial[12], la probabilidad de que:

[12] Perdón, con la ayuda de Gnumeric

a) Responda correctamente a 6.

b) Responda menos de 6 correctamente.

2. La probabilidad de que en una empresa minera de 10.000 empleados haya uno enfermo es de 0,2. Cuál será la probabilidad de que en un momento determinado haya en la empresa:

   a) 1.980 enfermos.

   b) Más de 2.080 enfermos.

Distribución normal Gnumeric implementa cuatro funciones relacionadas con la distribución Normal:

NORMDIST(x;media;desvest;acumulado) devuelve la distribución normal acumulada.

- **x** es el valor para el que calculamos la probabilidad
- **media** es la media de la distribución,
- **desvest** es la desviación estándar,
- **acumulado** es un valor booleano (0 ó 1) que indica si se va a obtener la probabilidad acumulada hasta ese valor.

NORMINV(p;media;desvest) devuelve la inversa de la distribución normal acumulada.

- **p** es la probabilidad dada correspondiente a la distribución normal,
- **media** es la media aritmética de la distribución y
- **desvest** es la desviación estándar de la distribución.

NORMSDIST(x) devuelve la distribución normal estándar acumulativa.

- **x** es el valor para el que calculamos la probabilidad

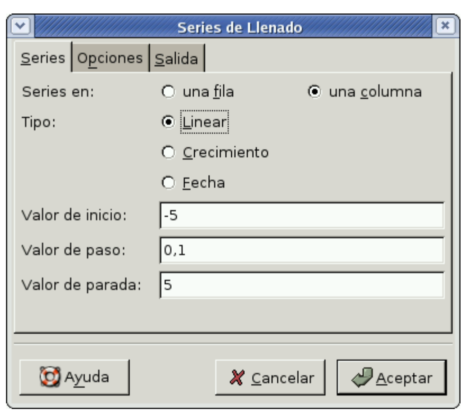
NORMSINV(p) devuelve el inverso de la distribución acumulativa normal estándar.

- **p** es la probabilidad dada correspondiente a la distribución normal.

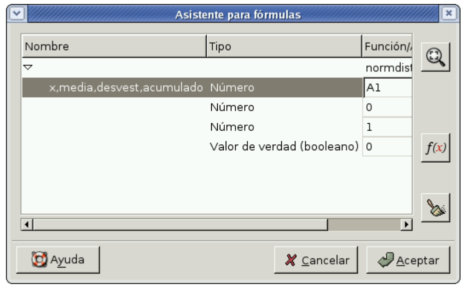
Ejercicio

1. Representar una curva normal.

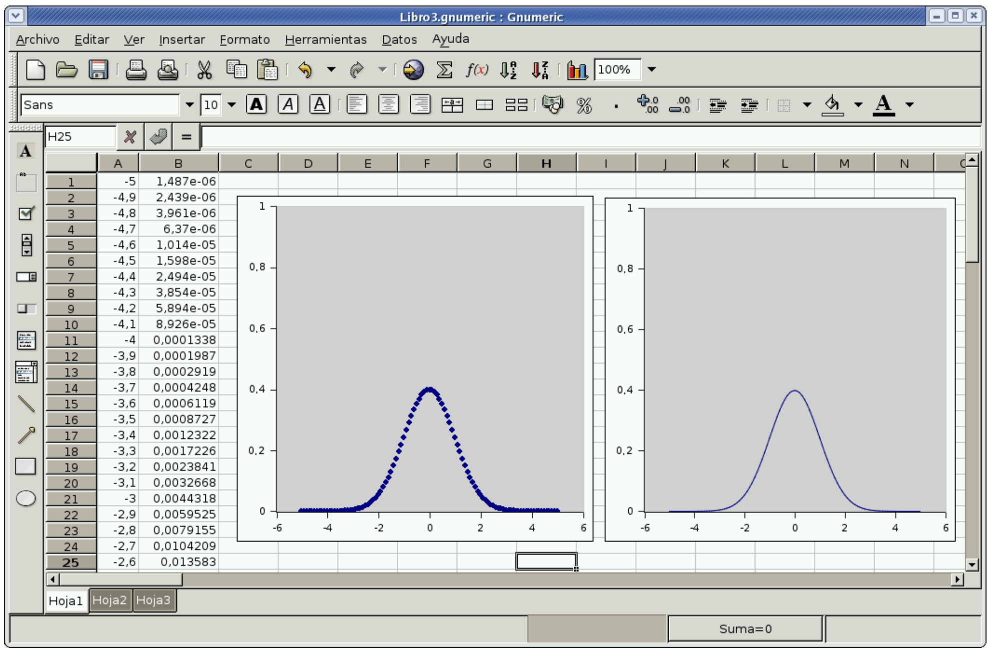
   a) Rellenar los valores de x (100)



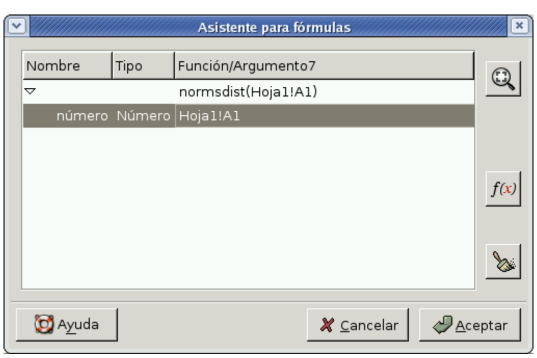
   b) Obtener los valores en esos puntos

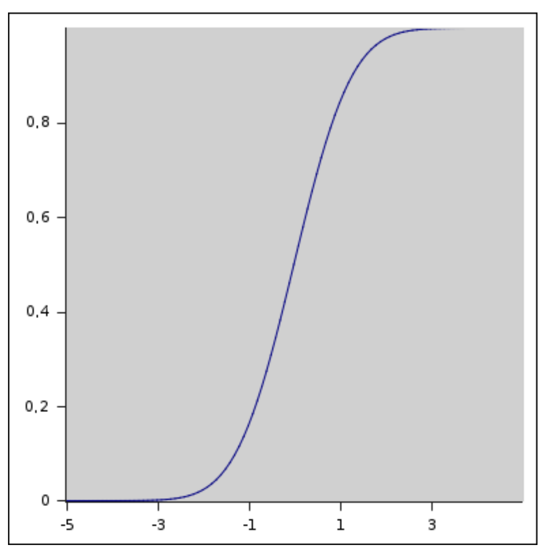


c) Representar las gráficas (discreta y continua), adecuando los ejes a nuestros intereses



2. Representa la gráfica de la distribución normal acumulada





3. Se fabrican unas pilas alcalinas cuya duración en horas sigue una distribución normal de media 60 h y desviación típica 5 h. Si se elige una pila al azar, ¿qué probabilidad hay de que dure?:

   a) Menos de 50 h.

   b) Entre 52 y 65 h.

4. El peso (P) de los socios del Club de Amigos del Buen Comer ha resultado que se distribuye conforme a una normal $N(92, 20)$. ¿Qué peso máximo tienen el 10 % de los individuos menos obesos de tan opíparo club?

3. Grace

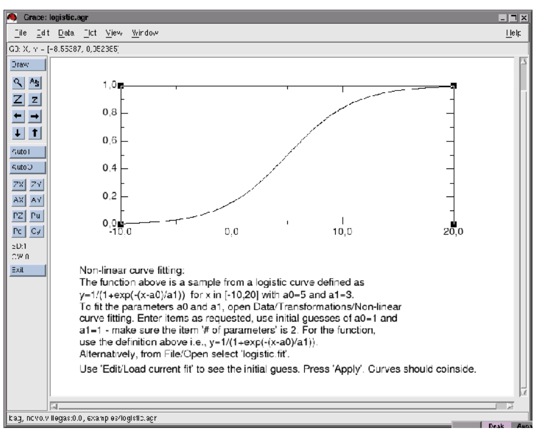
Grace es un programa para representar gráficos en dos dimensiones. No sólo permite representar funciones, sino que además es muy bueno para gráficos estadísticos. Destaca porque permite realizar análisis estadísticos sobre los datos, integrales y derivadas, interpolaciones, etc.

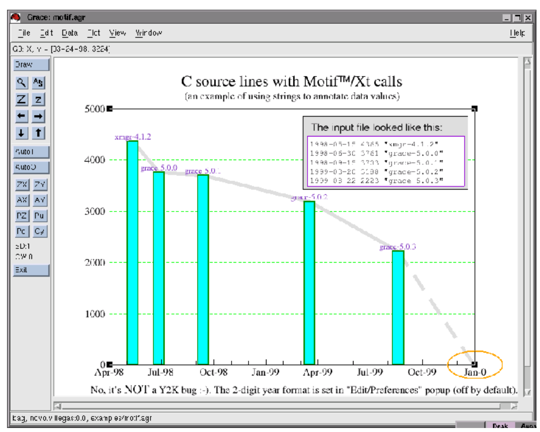
La página principal del programa es `http://plasma-gate.weizmann.ac.il/Grace/`.

Para ejecutar el programa escribiremos desde una xterm

`$ xmgrace`[13]

Si pulsamos sobre **Help** podremos comprobar que tenemos a nuestra disposición un tutorial, una guía de uso, una FAQ y además podemos acceder a ejemplos. Un par de capturas de ellos son:





Veamos una serie de ejemplos de sus posibilidades:

3.1. Estadística descriptiva

➥ **Para practicar:** Las notas de matemáticas de 1º ESO A han sido:

A	5, 6, 4, 5, 6, 5, 5, 8, 6, 3, 2, 1, 5, 5, 7, 6, 7, 3, 2, 3, 8, 9, 4, 4, 8, 9, 1, 5, 5, 4, 7, 6, 3, 4, 5, 5

Haz un estudio estadístico de ellas.

Solución

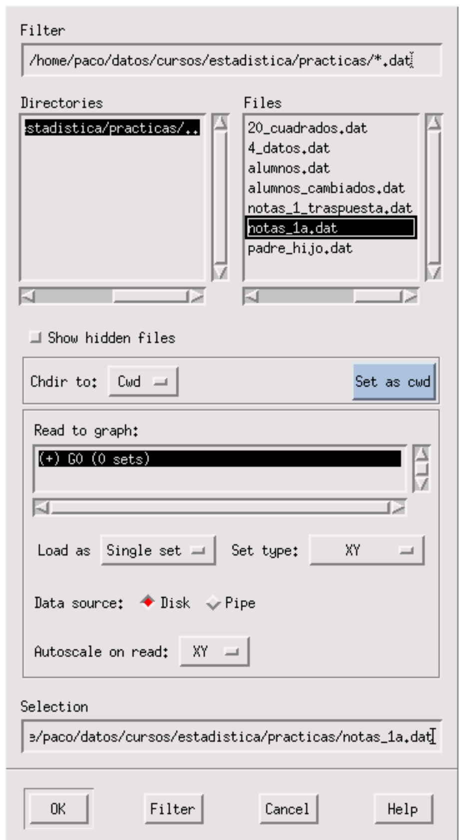
[13]Con
`$ grace`
se ejecuta en modo alfanumérico.

Las introducimos en un fichero de texto plano[14], pero hay que introducir dos columnas, la primera para enumerar la nota y la segunda la nota en sí. Nuestro fichero tendrá la forma:

```
 1  5
 2  6
...
36  5
```

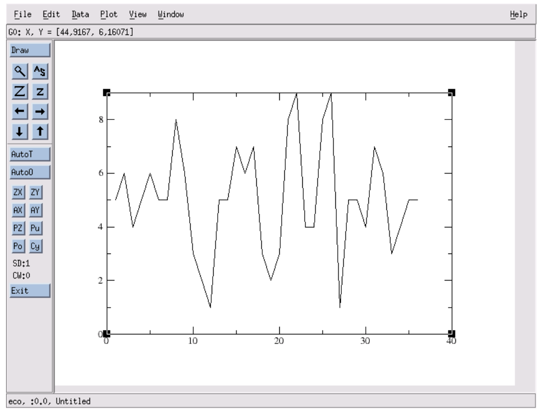
Este fichero lo llamaremos `notas_1a.dat`. Para leer el archivo de datos (tiene que contener los datos a dos columnas y sólo los números, sin encabezados) pulsamos sobre

Data → Import → ASCII

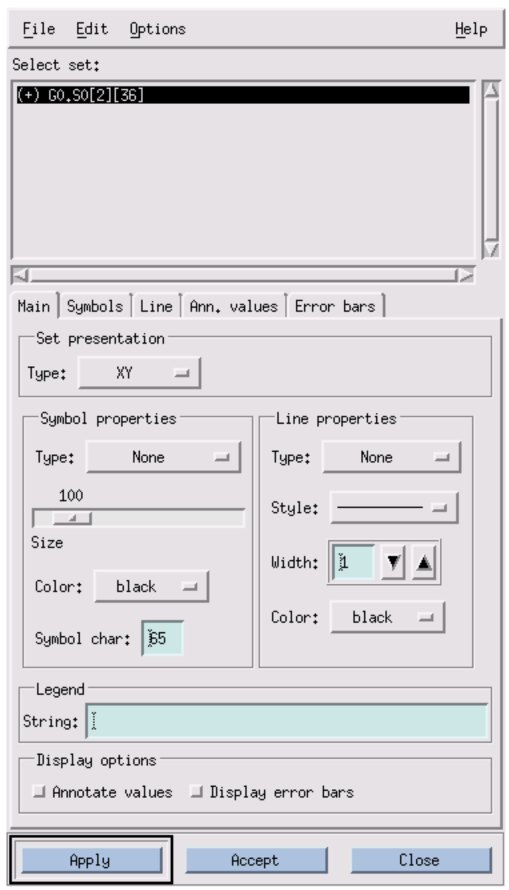


No se cerrará la ventana y tendremos que optar por cerrarla pulsando en la esquina superior derecha. Quedará entonces:

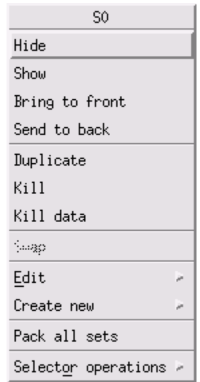
[14] Una buena opción es usar Gnumeric para generar la primera columna, introducir las notas en la segunda y exportar a Texto Plano.



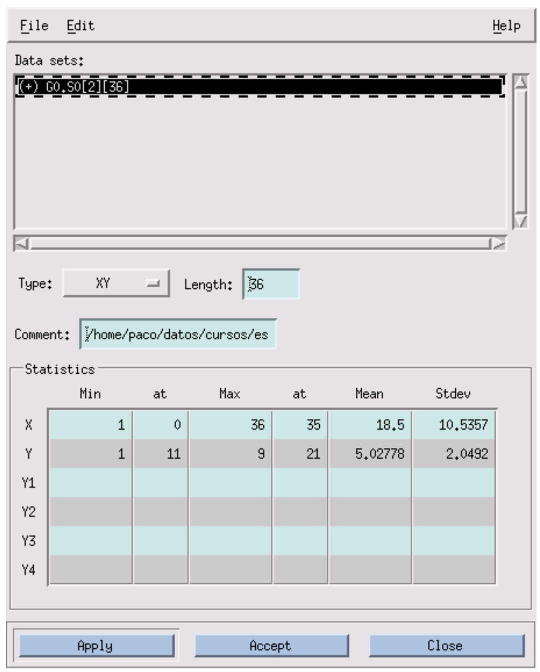
El gráfico es un poco malo, ¿verdad?. Vamos a mejorarlo:
Hagamos doble clic sobre el gráfico



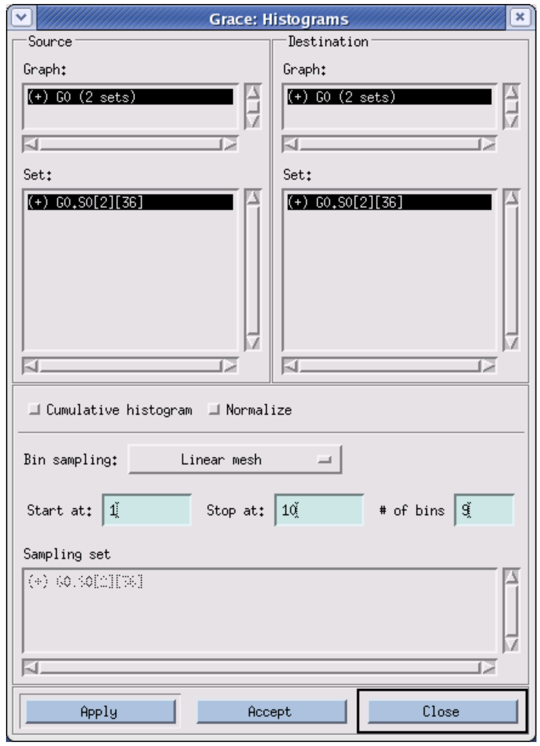
y optemos por dejar en limpia la zona de dibujo. Para eso pulsamos con el botón derecho sobre la ventana anterior



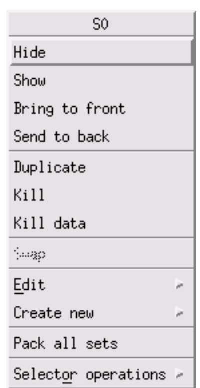
y ocultamos este gráfico (**Hide**). Todo limpio ¿verdad?. Vamos a obtener algunos datos sobre nuestra clase, para ello **Edit→ Data set** y señalamos el conjunto de datos



El gráfico es autoexplicativo. Ahora optemos por la cadena de menús **Data→Transformations→Histograms**

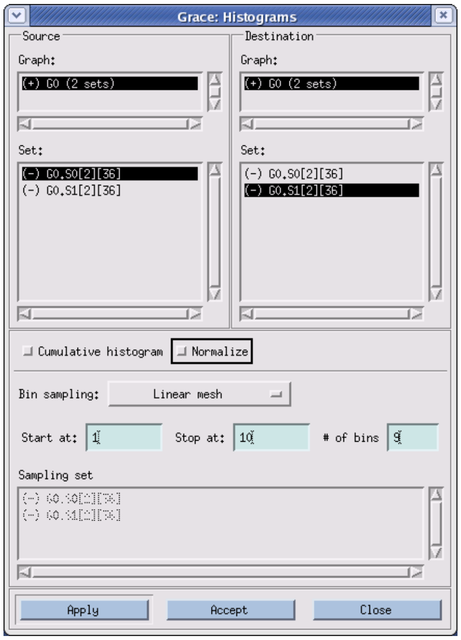


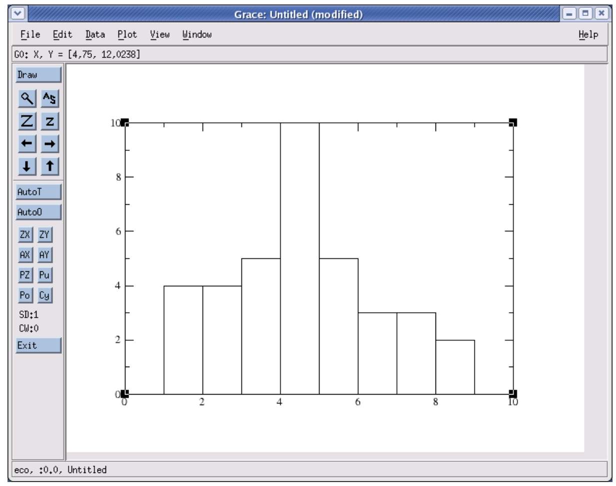
Para obtener el "histograma" tenemos que crear el gráfico de salida (distinto del que contiene los datos). Así que pulsamos con el botón derecho sobre la ventana anterior y optamos por seleccionar **Duplicate**



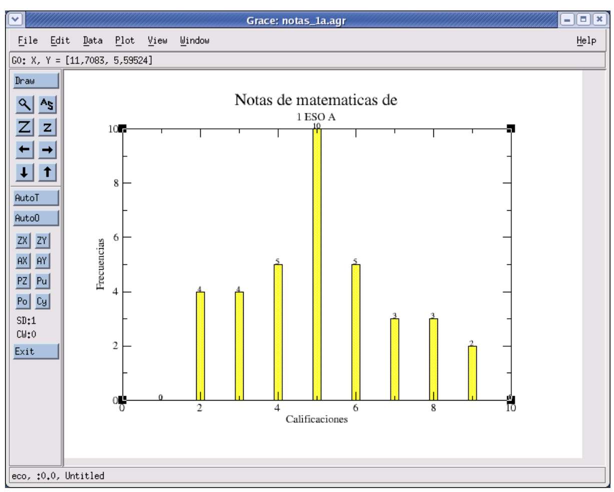
Es el momento, marquemos el gráfico de entrada[15] (G0.S0), el de salida (G0.S1) y rellenemos los campos tal cual aparece en el gráfico. Cuando apliquemos y aceptemos saldrá

[15]Esta nomenclatura indica que se trata del gráfico 0 (G0), curva dibujada o serie S0, S1, S2 ....



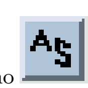


Esto mejora[16]. Pero podemos hacer que salga mejor

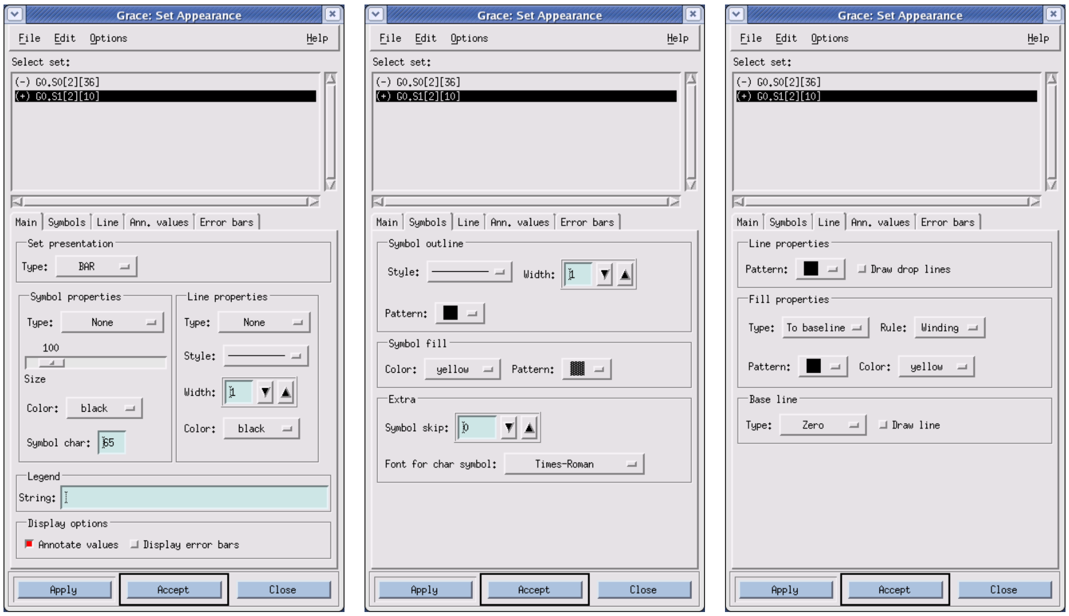


Para conseguir este gráfico he modificado[17]:

- Desde **Plot→Graph Appearance...** los títulos
- Haciendo doble clic sobre los ejes he introducido los títulos para los ejes.
- Desde **Plot→Set Appearance...** he modificado aspectos de las pestañas:

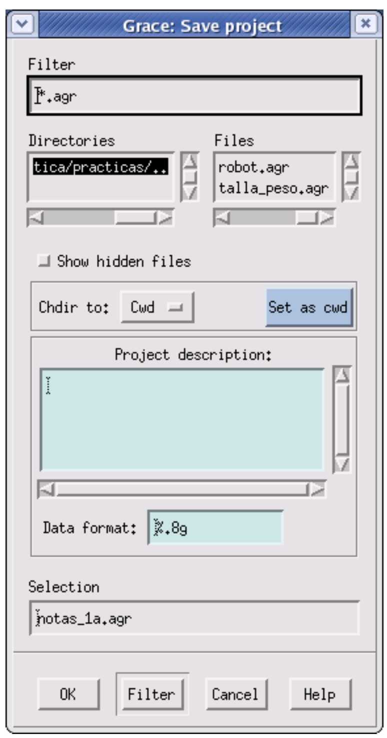
[16]Si la escala no es la adecuada, no se verá bien el gráfico. En esa caso sólo debemos pulsar sobre el icono  de la barra de la derecha.

[17]Pulsando en la zona adecuada de la zona de gráficos se accede a las mismas funcionalidades que desde el menú **Plot**

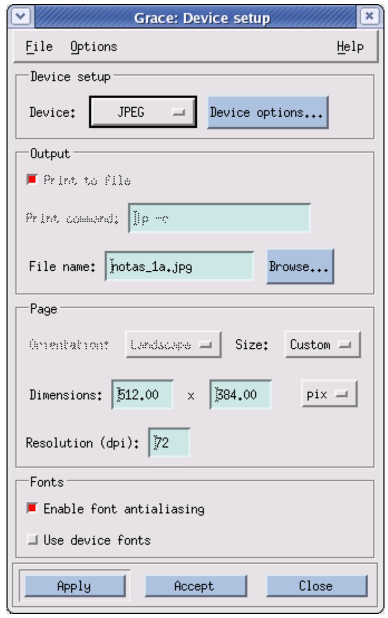


Se trata de conseguir que salga igual que el de la captura anterior.

Está claro que no debemos desperdiciar el tiempo, así que **File→Save**



Por último, si deseamos imprimir el trabajo **File→Print**, si lo que deseamos es exportarlo a algún formato gráfico, optaremos por **File→Print Setup**



y seleccionaremos el formato gráfico de salida adecuado. Que no se nos olvide, con esto configuramos la salida al fichero de formato adecuado, pero no tenemos que imprimirlo, así que falta **File→Print.**

Ejercicio La edad de 130 personas que realizaron un experimento se da a continuación:

	15	22	19	17	22	19	19	15	15	19	23	20	23	15	17	18	24	21	17
17	15	23	19	17	22	24	19	17	21	19	17	18	21	16	16	19	24	18	19
24	19	17	21	19	23	20	23	22	17	17	20	16	17	18	23	21	16	19	
20	21	21	18	18	19	18	21	17	17	23	24	21	21	17	16	19	19		
17	22	21	21	18	23	22	16	19	18	20	22	21	15	17	19	19	16		
17	16	18	18	18	21	23	21	17	18	19	19	17	16	17	23	22	17		
23	20	17	22	24	17	21	19	18	20	24	18	22	17	19	17	20	18		

1. Introduce los datos en un fichero de texto a dos columnas (separadas por un espacio). En la primera columna se escribirá el número de orden del dato y en la segunda el dato (Por ejemplo: 1 15 hasta 130 19). Para eso puedes usar
   a) Un editor cualquiera y guardarlos en texto plano.
   b) Rellenar con Gnumeric una columna con una serie del 1 al 130, en la segunda introducir los datos. Por último hay que exportarlos a texto plano.
   c) Usar el editor del programa Grace.
2. Agrupa las edades por años.
3. Da una tabla de frecuencias simple (para cada año).

3.2. Correlación

➥ **Para practicar:** La tabla siguiente muestra las respectivas alturas X e Y de una muestra de 12 padres y sus hijos primogénitos (en pulgadas):

Altura X del padre (en pulg)	65	63	67	64	68	62	70	66	68	67	69	71
Altura Y del hijo (en pulg)	68	66	68	65	69	66	68	65	71	67	68	70

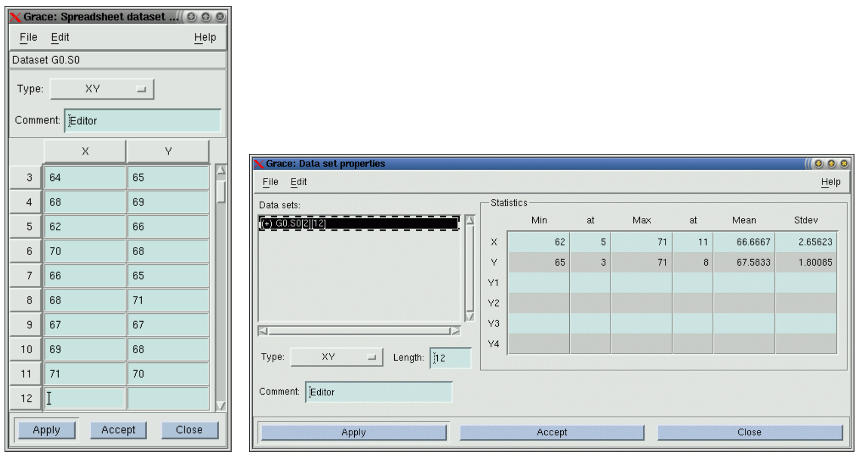
1. Construye el diagrama de dispersión.
2. Estudia la correlación entre ambas variables.
3. Halla la recta de regresión de Y sobre X.

Solución:

Ejecutemos

```
# xmgrace
```

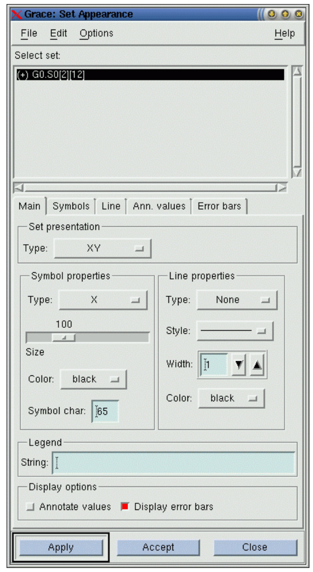
y en la ventana resultante, pulsemos sobre[18] **Edit→Data Sets...→Edit→Create new→In spreadseht**, después pasemos a introducir los datos:



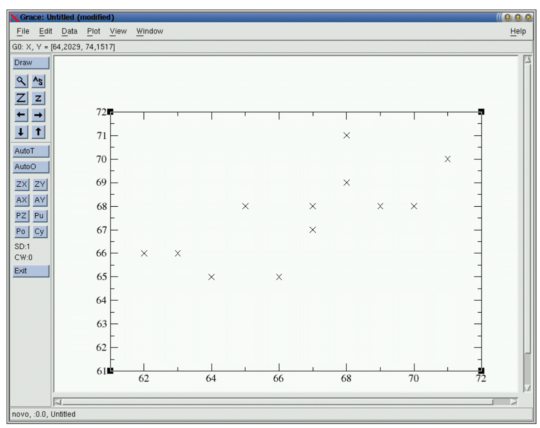
Tras aceptar, veremos en la ventana anterior que ya tenemos calculados algunos parámetros estadísticos

1. Si aceptamos en la última ventana y pulsamos sobre el botón  de la pantalla principal de la aplicación, nos aparecerán una serie de líneas que hemos de "eliminar":
   a) Pulsemos sobre los ejes con el ratón y adecuemos la escala (desde 61 a 72).
   b) Después pulsemos sobre las líneas con el ratón y, en la ventana que aparece, optemos porque no se unan los puntos con líneas y que los puntos se muestren como X.

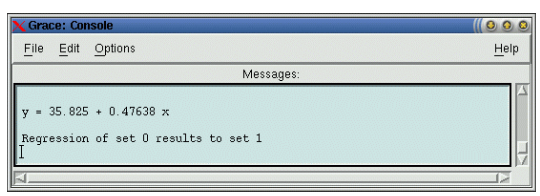
[18]También podemos leer el archivo de datos `datos.dat` (tiene que contener los datos a dos columnas y sólo los números, sin encabezados) usando
Data → Import → ASCII



Si aceptamos tendremos:



2. En el menú principal pulsemos sobre **Data→Transformations→Regression** y aceptemos en la ventana resultante, saldrá



Además, la recta de regresión se dibujará en la nube de puntos.

Ejercicios

1. En la tabla siguiente damos la evolución del récord del mundo de salto de longitud masculina, en metros.

Años	1901	1921	1925	1931	1935	1961	1962	1965	1968	1991
Marca	7,61	7,69	7,89	7,98	8,13	8,28	8,31	8,35	8,90	8,95

Nota: Puedes trabajar a partir de los datos ya introducidos en Gnumeric. Hay que exportarlos a texto plano, y después, con un editor sustituir la coma decimal por un punto.

a) Representa la nube de puntos asociada a estos datos.

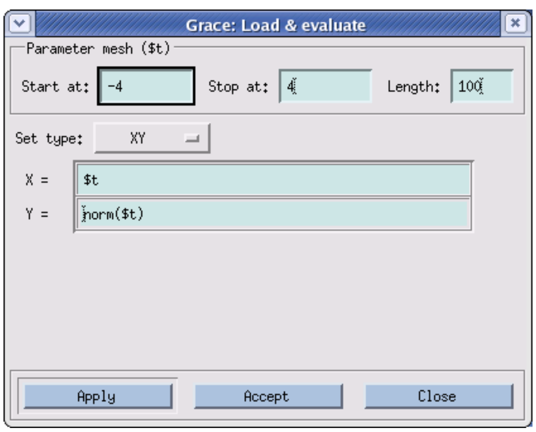
b) ¿Puede estimarse la marca del récord por regresión lineal? Si es así, ¿cuánto se saltará en los años 2000 y 2005?

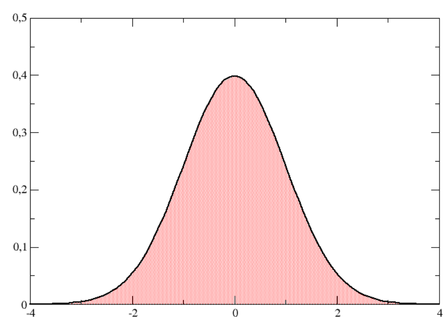
c) Usa la Web: `http://tux.iesmurgi.org/matematicas/materiales/correlacion/` para comprobar los resultados del ejercicio anterior y haz 10 estimaciones de la marca que se obtendrá en años venideros y otras 10 del año en que se alcanzarán las marcas (pon 5 valores "extremos" y 5 dentro del rango de las variables). Comenta los resultados.

3.3. Un añadido: gráficas, integrales y derivadas

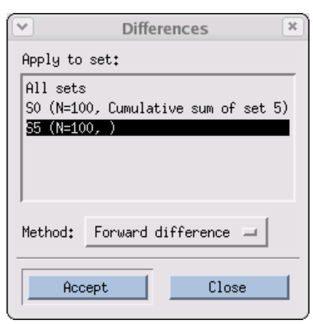
Por último veamos otra posibilidad interesante de este programa. Vamos a representar la función de densidad de la distribución normal, su derivada y su integral.

Los valores de las variables los podemos introducir de forma calculada, para ese menester, optamos por **Edit→Data sets...** y en la ventana resultante **Edit→Create New→By Formula**. Las opciones[19] y la salida[20] obtenida:





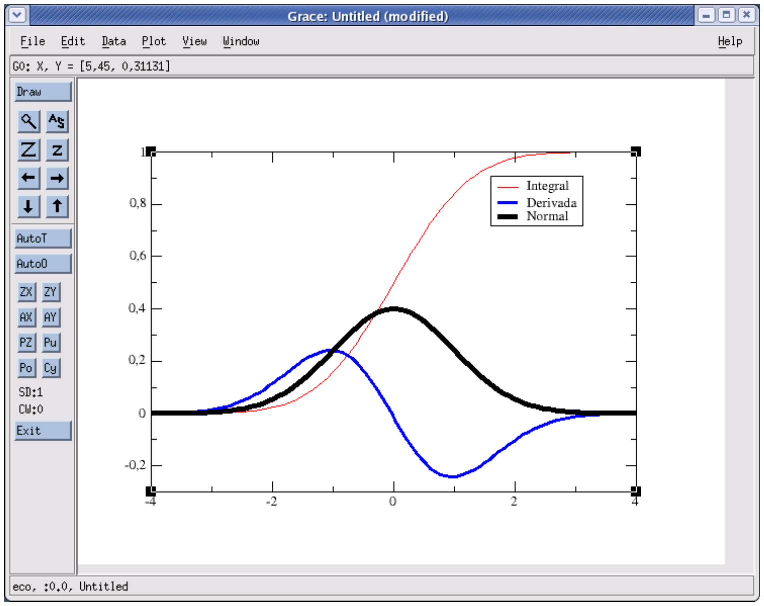
Para obtener la gráfica de la función integral **Data→Transformations→Integration...** y marcar el gráfico adecuado (S0).



En el caso de derivada **Data→Transformations→Differences...**

[19] El listado de las funciones que admite Grace se puede consultar en [1], se instala con el programa en `/usr/share/doc/grace` o en `http://plasma-gate.weizmann.ac.il/Grace/doc/UsersGuide.html`

[20] Un poco mejorada



Ejercicicio:

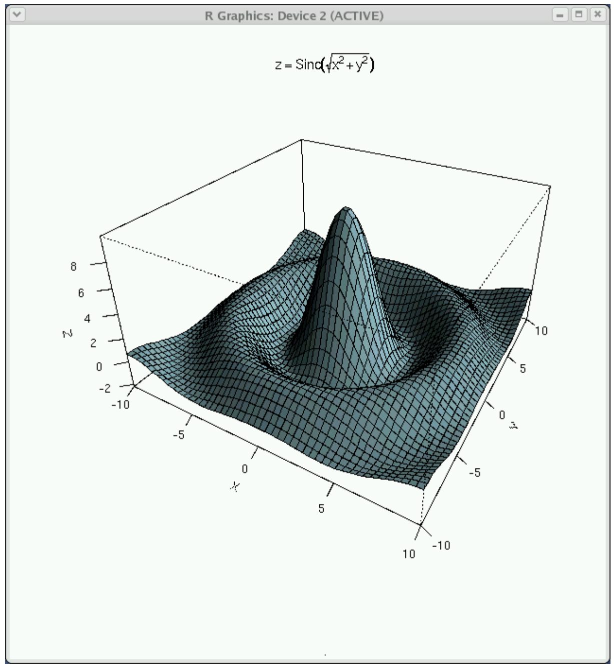
1. Representar la parábola $f(x) = x^2 - 2x$ en el intervalo $[-1, 4]$, la gráfica de su derivada y su integral.

2. Transformar la curva anterior usando **Data Trasformations** → **Evaluate expresion**. Seleccionamos la curva a transformar (S1, o S2 o S3... en **source**). No seleccionamos ninguna curva en **Destination** si deseamos crear una nueva curva. Introducimos la fórmula de la transformación (por ejemplo $y = y + 1$). Seleccionamos en **restriction** la región de datos que queremos transformar (ver apartado definir regiones). Por último **Apply** y luego **Close**

4. R

R ó "GNU S" es un programa libre, disponible para entornos Linux, Mac y Windows (`http://cran.r-project.org/`[21]) que nos permite realizar cálculos y gráficos (de alta calidad) estadísticos. R permite trabajar las técnicas estadístcas más básicas, pero llegando a las más avanzadas. Además, permite que le añadamos nuevas funcionalidades, ya que podemos programar nuevas funciones o instalar nuevos paquetes (`http://cran.r-project.org/src/contrib/PACKAGES.html`).

[21] Desde esta página es posible acceder a la completa documentación disponible para el programa, mucha de ella en castellano (véase [4, 5, 6, 7])



4.1. Comencemos

Para ejecutar el programa, lo mejor es situarnos en el directorio de trabajo y desde un terminal gráfico escribimos:

```
$ mkdir pracR
$ cd pracR
$ R
```

Aparecerá el mensaje que reproducimos debajo y al final de éste el *promp* de entrada de órdenes del programa, en el mensaje se nos indica que para salir hay que escribir `q()`.

```
R : Copyright 2004, The R Foundation for Statistical Computing
Version 1.9.0 (2004-04-12), ISBN 3-900051-00-3

R is free software and comes with ABSOLUTELY NO WARRANTY.
You are welcome to redistribute it under certain conditions.
Type 'license()' or 'licence()' for distribution details.

R is a collaborative project with many contributors.
Type 'contributors()' for more information and
'citation()' on how to cite R in publications.

Type 'demo()' for some demos, 'help()' for on-line help, or
'help.start()' for a HTML browser interface to help.
Type 'q()' to quit R.

>
```

Si deseamos acceder a la completa ayuda que trae podemos ejecutar:

```
> help()
```

Para salir de la ayuda

```
: q
```

La ayuda en modo html

```
>help.start()
```

Para obtener ayuda sobre un comando

```
> ?mean
```

o para buscar qué comando contiene en su ayuda la cadena "mean"

```
> help.search("mean")
```

- Las teclas de cursor nos permiten modificar o movernos por el histórico de comandos. R también acepta cortar y pegar.
- [7] nos permite tener en un folio una referencia rápida de R. Se adjunta en el apéndice D en la página 111.
- Casi todos los comandos que se van a estudiar están resumidos en el apéndice E en la página 112.

Con

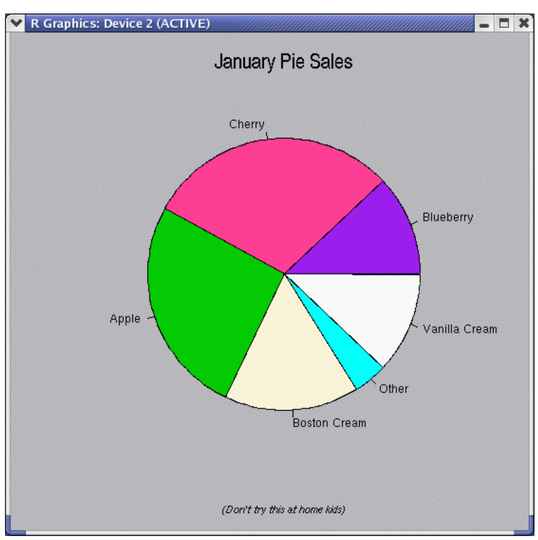
```
> apropos("mean")
[1] "kmeans"          "weighted.mean"   "mean"            "mean.Date"
[5] "mean.POSIXct"    "mean.POSIXlt"    "mean.data.frame" "mean.default"
[9] "mean.difftime"
```

obtenemos todas las funciones cuyo nombre contiene la palabra pasada como argumento.

Para ver algunos ejemplos de los gráficos que podemos hacer con el programa, escribiremos:

```
>demo(graphics)
```

y pulsaremos, teniendo activa la ventana del terminal, sobre la tecla **Intro** varias veces hasta que finalice la presentación.



Para salir:

```
> q()
Save workspace image? [y/n/c]:
```

Se nos pregunta si deseamos guardar la sesión (y), salir y no guardar la lista de comandos (n), o cancelar, de esa forma seguimos en el programa. Optemos por esta última opción e iniciemos un paseo por el programa:

```
>#Si se pone una almohadilla, es un comentario
```

Operaciones Matemáticas básicas

```
> 4+10*10
[1]  104
>(5+2)*3
[1] 21
```

Disponemos de las funciones matemáticas más usuales:

```
>#sqrt, exp, log, sin, cos, tan, ...
> sqrt (16)
[1]  4
> exp(1)
[1] 2.718282
> exp(5/0)
[1] Inf
> exp(-5/0)
[1] 0
```

Podemos ver cómo están definidas las variables de entorno con

```
> options()
```

Para trabajar con más dígitos significativos

```
> pi
[1] 3.141593
> options(digits=22)
> pi
[1] 3.141592653589793115998
> #Dejemos el valor por defecto
> options(digits=7)
```

Variables y asignaciones

```
> resultado <- (4+10)*3
> resultado
[1] 42
> nombre <- "Pepe"
> nombre
[1] Pepe
```

R es casesensitive, esto implica que si escribimos

```
> Nombre
```

no encontrará la variable

Para saber qué objetos están en memoria, podemos usar (indistintamente)

```
>ls()
>objects()
>#Para borrar el objeto x
> rm(x)
```

Vectores de datos

```
>#Para crearlos: Comando c -> Concatenar
> x<- c(1,2,3,4,5)
> x
[1] 1 2 3 4 5
> x<- c(6,7,x)
> x
[1] 6 7 1 2 3 4 5
> x[3]
[1] 1
>almnos<-c("Pepe","Cristina","Marta")
>alumnos
[1] "Pepe"     "Cristina" "Marta"
```

Introducir datos con la función `scan()`

```
> datos<-scan()
1: 5
2: 4
3: 3
4: 2
5: 1
6:
Read 5 items
> datos
[1] 5 4 3 2 1
```

Secuencias de números, se introducen con la función:

```
seq(mínimo,máximo,[incremento,longitud]))
```

Los comandos que siguen obtienen el mismo resultado

```
seq(1,10,1)=seq(1,10)=seq(10)=seq(1:10)=1:10
>1:10
[1] 1 2  3  4 5  6  7  8  9  10
```

Comando Réplica

```
>rep(2,5)
[1] 2 2 2 2 2
```

Se pueden escribir varias sentencias separadas por ;

```
>rep(1:10,3); rep(c("Sí","No"),5)
 [1]  1  2  3  4  5  6  7  8  9 10  1  2  3  4  5  6  7  8  9 10  1  2  3  4  5
[26]  6  7  8  9 10
 [1] "Si" "No" "Si" "No" "Si" "No" "Si" "No" "Si" "No"
```

Podemos crear secuencias aleatorias (sin repetición) con:

```
> # Una lista de 1:10 ordenada aleatoriamente
> sample(10)
 [1]  8  5  9  6  7 10  2  4  1  3
> # 6 números naturales menores de 100
> sample(100,6)
[1] 49 84 13 36  9 31
> # Dentro de una distribución normal
> rnorm(4)
[1] -0.2502104 -2.1177462 -1.2384110 -0.6614978
> #Un poco de "mezclas"
> expand.grid(c(1:6),c(1:6))
   Var1 Var2
1     1    1
...
36    6    6
```

EJERCICIO: Halla todas las posibles formas de mezclar los números 4 al 7 con las letras "a","b" y "c" y con los caracteres "+" y "-"

Todas las operaciones aritméticas y funciones matemáticas pueden ser aplicadas a vectores y se realizan componente a componente.

```
> a <- 1:10
> a
> a^2
> b <- sqrt(a)+5
> b
> #¿Qué os parece esto?
> plot(a,b)
> plot(a,b, type="l")
```

Matrices El comando "básico" para crear matrices[22] es: `matrix(datos,nfilas,ncolumnas)`

```
> matrix(1:4)
     [,1]
[1,]    1
[2,]    2
[3,]    3
[4,]    4
> matrix(1:10, ncol=2)
     [,1] [,2]
[1,]    1    6
[2,]    2    7
[3,]    3    8
[4,]    4    9
[5,]    5   10
> matrix(1:10, ncol=2, byrow=TRUE)
     [,1] [,2]
[1,]    1    2
[2,]    3    4
[3,]    5    6
```

[22]Por defecto las matrices se crean por columnas. Para crearlas por filas hemos de usar el parámetro `byrow=TRUE`

```
[4,]    7    8
[5,]    9   10
```

Podemos nombrar las filas y columnas[23]:

```
> datos<-matrix(c(8,3,5,7),ncol=2,
+ dimnames=list(c("Matematicas","Lengua"),c("Pepe","Marta")))
> datos
                   Pepe  Marta
Matematicas          8     5
Lengua               3     7
```

- Operaciones elementales con Matrices

  A ± B suma/resta de matrices

  A %* % B producto de matrices

  t(A) transpuesta de la matriz A

  solve(A,b) solución del sistema de ecuaciones $A \cdot x = b$.

  solve(A) inversa de la matriz A

  diag(A) diagonal principal de la matriz A

  det(A) para obtener el determinante de A

Operemos con un par de matrices[24]:$A = \begin{pmatrix} 1 & 2 & 4 \\ 2 & -1 & 0 \\ 1 & 0 & 1 \end{pmatrix}$ y $B = \begin{pmatrix} 1 & 2 & 3 \\ 1 & 4 & 9 \\ 1 & 8 & 27 \end{pmatrix}$

```
> A <-matrix(c(1,2,1,2,-1,0,4,0,1),ncol=3)
> A
[,1] [,2] [,3]
[1,] 1  2 4
[2,] 2 -1 0
[3,] 1  0 1
> diag(A)
[1]  1 -1  1
> solve(A)
     [,1] [,2] [,3]
[1,]    1    2   -4
[2,]    2    3   -8
[3,]   -1   -2    5
> det(A)
[1] -1
> t(A)
     [,1] [,2] [,3]
[1,]    1    2    1
[2,]    2   -1    0
[3,]    4    0    1
> x <- 1:3
> B <- matrix(c(x,x^2,x^3),ncol=3,byrow=T)
> B
```

[23]El símbolo + aparece automáticamente cuanto pulsamos intro y con él R nos indica que espera más instrucciones.
[24]También podemos introducir la matriz con

```
> A <- matriz(scan(),ncol=3)
```

```
> B
     [,1] [,2] [,3]
[1,]    1    2    3
[2,]    1    4    9
[3,]    1    8   27
> A + B
     [,1] [,2] [,3]
[1,]    2    4    7
[2,]    3    3    9
[3,]    2    8   28
```

Cuidado con el producto, no es `*` es `%*%`

```
> A*B
 A*B
     [,1] [,2] [,3]
[1,]    1    4   12
[2,]    2   -4    0
[3,]    1    0   27
> A %*% B
     [,1] [,2] [,3]
[1,]    7   42  129
[2,]    1    0   -3
[3,]    2   10   30
```

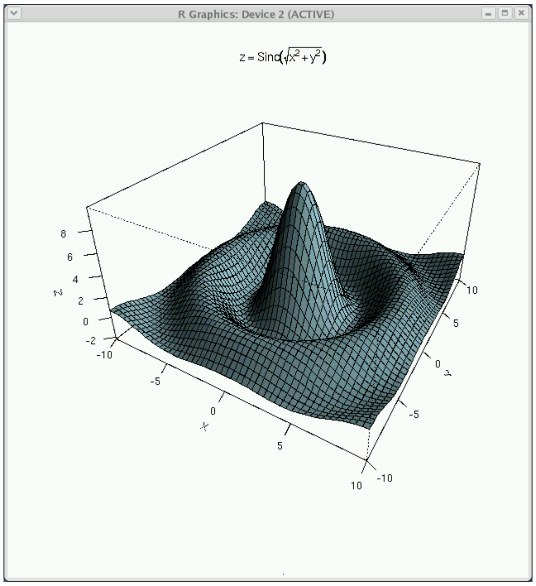
Resolución de sistemas de ecuaciones

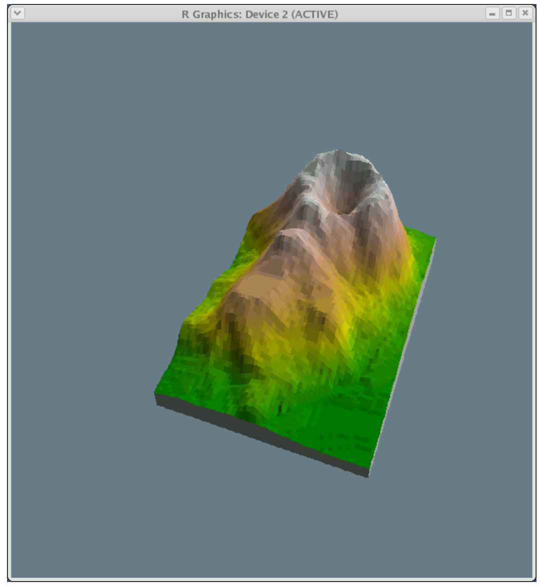
```
> C <- matrix(1:3)
> C
> solve(A,C)
     [,1]
[1,]   -7
[2,]  -16
[3,]   10
```

Un *impass* por favor:

```
> demo(persp)
```

Entre los resultados dos:






```
> q()
```

Y guardamos la sesión

4.2. Para practicar: alumnos.txt

Vamos a ver algunas cuestiones básicas que podemos hacer desde R. Creamos un directorio en donde poner el trabajo que vamos a realizar. Por ejemplo:

```
$mkdir alumnos
```

y ponemos en él el fichero `alumnos.txt`. El fichero `alumnos.txt` lo podemos bajar de `http://www.iesmurgi.org/~pvillegas/estadistica/alumnos.txt`, contiene la talla, estudios, habitantes y peso de una población de 50 individuos. Los distintos valores se separan con espacios[25].

```
Talla Estudios Habitantes Peso
1.63 bachiller 2 54.88
1.44 fp 3 33.5
1.9 bachiller 3 89.38
1.58 bachiller 3 53.28
1.38 bachiller 5 36.62
1.8 bachiller 6 79.47
1.64 superior 3 72.47
1.93 bachiller 5 85.26
1.95 bachiller 2 102
1.59 bachiller 2 62.28
1.78 primario 4 81.17
1.96 primario 4 100.61
1.89 bachiller 5 94.43
1.52 bachiller 4 57.96
1.74 diplomado 4 67.86
1.68 diplomado 5 63.2
2 diplomado 6 105.01
1.67 diplomado 3 66.57
1.46 primario 5 43.46
1.98 primario 4 96.4
1.47 primario 2 42.38
1.74 primario 3 80.41
1.9 diplomado 4 92.7
1.65 fp 5 62.81
1.34 fp 8 39707
1.75 diplomado 2 78.951
1.68 diplomado 3 67.15
1.72 bachiller 4 71.94
1.65 bachiller 4 64.1
1.8 bachiller 6 76.68
1.94 bachiller 4 96.18
1.99 primario 4 103.029
1.36 primario 3 44.105
1.69 primario 5 70.16
1.69 bachiller 4 66.79
1.51 bachiller 3 57.27
1.98 bachiller 5 94.71
1.84 bachiller 3 81.25
2.02 bachiller 6 101.3
1.76 primario 4 73.68
1.96 fp 3 90.7
1.78 fp 3 84.59
1.54 superior 4 53.97
1.8 diplomado 5 82.13
1.53 diplomado 4 52.8
1.74 diplomado 5 77.22
1.7 primario 3 74.7
1.66 primario 4 69.34
1.83 primario 5 92.24
1.52 bachiller 4 61.05
```

4.2.1. Estadística descriptiva

Desde el directorio `alumnos` comencemos una sesión del programa.

[25] No se escriben en tres columnas, se trata de introducirlos en una sola. En los apuntes, para no gastar demasiado papel, se han puesto en tres columnas y letra diminuta.

```
$R
> list.files()
[1] "alumnos.txt"
> file.show("alumnos.txt")
Talla Estudios Habitantes Peso
1.63 bachiller 2 54.88
...
1.52 bachiller 4 61.05
> datos.alumnos<-read.table("alumnos.txt",header=TRUE)
> names(datos.alumnos)
[1] "Talla" "Estudios" "Habitantes" "Peso"
```

Con la sentencia `read.table` leemos el contenido del fichero `alumnos.txt` y lo ponemos en la variable `datos.alumnos`. Notar que hemos añadido la opción de que la primera línea del fichero es el nombre de los campos contenidos en él. Podemos ver cuáles son con el comando de la 2ª línea.

Veamos el contenido del caracter Talla

```
> datos.alumnos$Talla
[1] 1.63 1.44 1.90 1.58 1.38 1.80 1.64 1.93 1.95 1.59 1.78 1.96 1.89 1.52 1.74
[16] 1.68 2.00 1.67 1.46 1.98 1.47 1.74 1.90 1.65 1.34 1.75 1.68 1.72 1.65 1.80
[31] 1.94 1.99 1.36 1.69 1.69 1.51 1.98 1.84 2.02 1.76 1.96 1.78 1.54 1.80 1.53
[46] 1.74 1.70 1.66 1.83 1.52
```

Para no tener que acceder a los nombres de los campos con la nomenclatura `variable.$nombrecampo`, ejecutamos:

```
> attach(datos.alumnos)
```

a partir de ahora podemos referenciar un campo sólo por su nombre.

```
> Talla
[1] 1.63 1.44 1.90 1.58 1.38 1.80 1.64 1.93 1.95 1.59 1.78 1.96 1.89 1.52 1.74
[16] 1.68 2.00 1.67 1.46 1.98 1.47 1.74 1.90 1.65 1.34 1.75 1.68 1.72 1.65 1.80
[31] 1.94 1.99 1.36 1.69 1.69 1.51 1.98 1.84 2.02 1.76 1.96 1.78 1.54 1.80 1.53
[46] 1.74 1.70 1.66 1.83 1.52
```

Obtengamos algunos datos sobre la variable `Talla`

```
> summary(Talla)
  Min.   1st Qu.  Median  Mean   3rd Qu.   Max.
  1.340   1.600   1.730   1.721   1.878   2.020
```

Obtenemos el valor máximo, mínimo, la media y los cuartiles de esta variable.[26]

Obtengamos la media, mediana, cuasivarianza[27], cuasidesviación típica, rango y algunos percentiles de la variable `Talla`:

```
> mean(Talla)
[1] 1.7212
> median(Talla)
```

[26]Si no hubiésemos ejecutado el comando `attach`, tendríamos que haber escrito
`>summary(datos.alumnos$Talla)`

[27]Podemos crear una función que calcule la varianza (y después la desviación típica como sigue):
```
>varianza <- function(x){(length(x)-1)/length(x)*var(x)}
>dt<- funcion(x){sqrt(varianza(x))}
```
Para obtenerlas, ahora sólo hemos de escribir
```
>varianza(datos)
>dt(datos)
```

```
[1] 1.73
> var(Talla)
[1] 0.03368016
> sqrt(var(Talla))
[1] 0.1835216
> sd(Talla)
[1] 0.1835216
> range(Talla)
[1] 1.34 2.02
> quantile(Talla, .01)
    1%
1.3498
> quantile(Talla, c(.05,.15))
   5%   15%
1.407 1.520
```

Por último obtengamos las frecuencias absolutas de la variable Habitantes

```
> table(Habitantes)
Habitantes
2 3 4 5 6 8
5 13 16 11 4 1
```

salimos (y guardamos la sesión) con[28]:

```
>q()
```

4.2.2. Algunos gráficos con R

Para saber casi todo sobre este tema, y las opciones disponibles para cada uno véase [6].

Los gráficos más usuales son

`piechart(datos, opciones)` Diagrama de sectores

`hist(datos,nclass=n, opciones)` Histograma, `nclass` es opcional y con él establecemos el número de intervalos

`barplot(datos, opciones)` Diagrama de barras

`boxplot(datos, opciones)` Diagrama de cajas y bigotes

`stem(datos)` Diagrama de tallo y hojas

Estas funciones gráficas tienen muchas posibilidades más y para ampliar lo mejor es mirar la ayuda del programa, algunas de las opciones más usuales son:

`main="titulo"` Título del gráfico

`xlab="etiqueta eje X"` Etiqueta del eje X

`ylab="etiqueta eje Y"` Etiqueta del eje Y

`type="l"` Une los puntos con líneas

`col="nombre"` Color del gráfico

Podemos añadir varias separadas por comas

[28] Antes de salir nos preguntará si queremos guardar la imagen del espacio de trabajo (se guardan todas las órdenes introducidas en la sesión), lo mejor sería decir que sí. Podemos guardar el espacio de trabajo en cada directorio en donde ejecutamos R, así, según el trabajo que estemos realizando disponemos de la posibilidad de tener distintas sesiones de trabajo guardadas en función del directorio desde donde lo hayamos llamado.

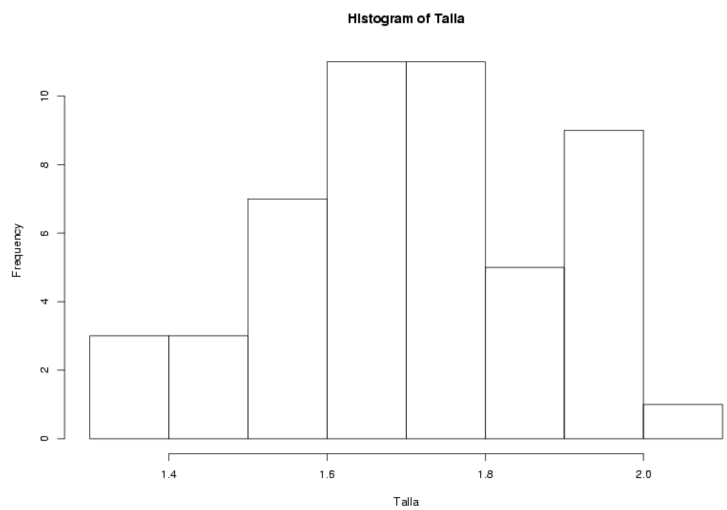
Árboles de tallo y hoja

```
> stem(Talla)
The decimal point is 1 digit(s) to the left of the |
13 | 468
14 | 467
15 | 1223489
16 | 3455678899
17 | 024445688
18 | 000349
19 | 0034566889
20 | 02
```

Para practicar Comprueba la salida de `stem(Talla,0.1)` y `stem(Talla,10)`

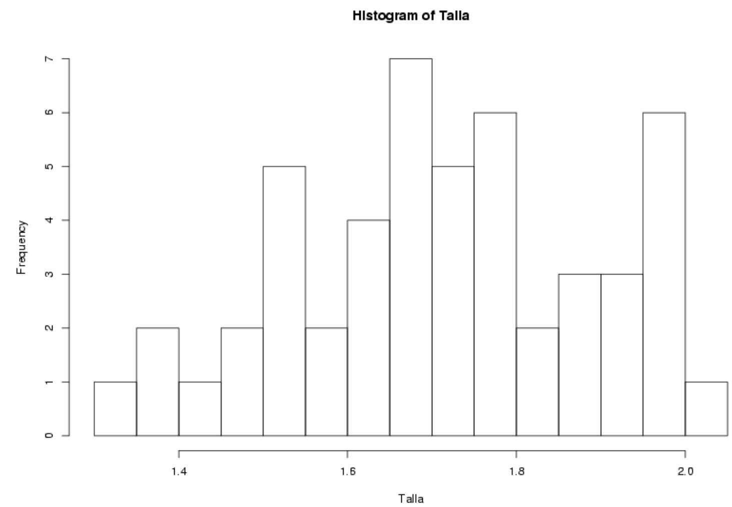
Histogramas Podemos hacer un histograma con:

```
> hist(Talla)
```

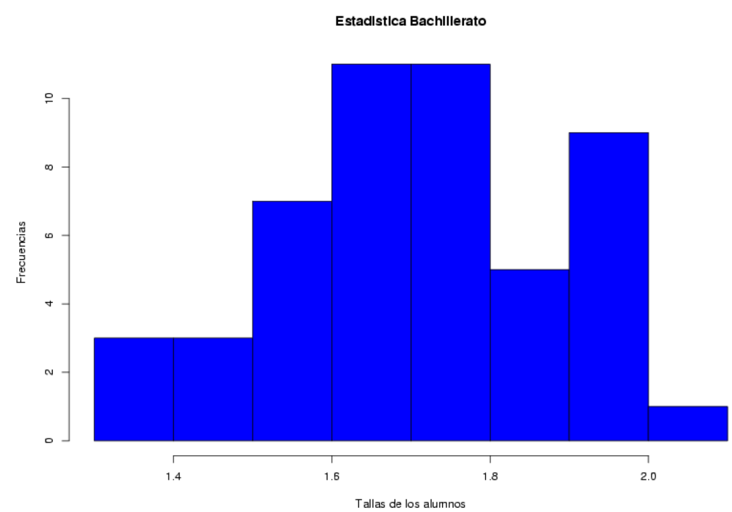


Podemos decirle que intente ajustar el número de intervalos:

```
> hist(Talla,nclass=15)
```



Pero nos interesa adecuarlo a nuestro gusto y para eso ejecutamos:



Si deseamos guardar el fichero en formato `png` escribiremos:

```
> jpeg()
> hist(Talla, main="Estadistica 2°", xlab="Tallas de los alum-
nos", ylab="Frecuencias", col="blue")
> dev.off()
```

El programa guardará el gráfico en el directorio de trabajo en un fichero de nombre `Rplotxxx.jpeg`.[29]
Con

[29] Si queremos dar un nombre particular al fichero, debemos especificarlo previamente con la orden `jpeg("nombre_fichero")`

```
>x11()
```

dejamos las cosas como estaban.

NOTA: Para saber qué formatos gráficos hay a nuestra disposición debemos ejecutar

```
> help(Devices)
```

Para practicar El comando `hist` tiene más posibilidades que merece la pena estudiar, y se trata de usar la opción `plot=F`, es decir, que no lo represente gráficamente.

Por ejemplo, si usamos

```
> hist(Talla,plot=F)
$breaks
[1] 1.3 1.4 1.5 1.6 1.7 1.8 1.9 2.0 2.1

$counts
[1]  3  3  7 11 11  5  9  1

$intensities
[1] 0.5999975 0.6000000 1.4000000 2.2000000 2.2000000 1.0000000 1.8000000
[8] 0.2000000

$density
[1] 0.5999975 0.6000000 1.4000000 2.2000000 2.2000000 1.0000000 1.8000000
[8] 0.2000000

$mids
[1] 1.35 1.45 1.55 1.65 1.75 1.85 1.95 2.05

$xname
[1] "Talla"

$equidist
[1] TRUE

attr(,"class")
[1] "histogram"
```

Nos agrupará los datos en intervalos[30]. Podemos conseguir intervalos de distinta amplitud usando `br`, por ejemplo

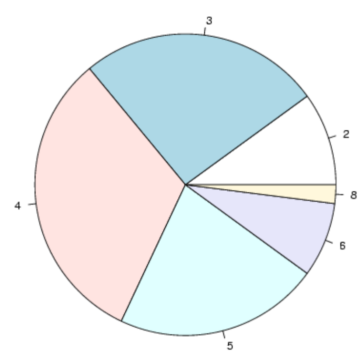
```
> hist(Talla,plot=F,br=c(1,1.2,1.5,2,2.1))
$breaks
[1] 1.0 1.2 1.5 2.0 2.1
$counts
[1]  0  6 43  1
$intensities
[1] 0.00 0.40 1.72 0.20
$density
[1] 0.00 0.40 1.72 0.20
$mids
[1] 1.10 1.35 1.75 2.05
```

[30] Por defecto, los intervalos son de la forma]a,b]. Este valor se puede cambiar, para ver cómo se hace revisar la ayuda del comando `hist`.

```
$xname
[1] "Talla"
$equidist
[1] FALSE
attr(,"class")
[1] "histogram"
```

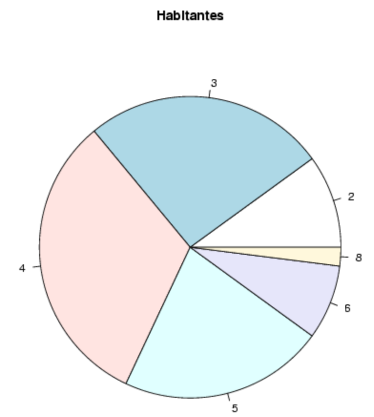
Diagramas de sectores Primero hemos de obtener las frecuencias absolutas, y después pintamos el gráfico

```
> table(Habitantes)
Habitantes
2 3 4 5 6 8
5 13 16 11 4 1
> frec.Habitantes<-table(Habitantes)
> pie(frec.Habitantes)
```



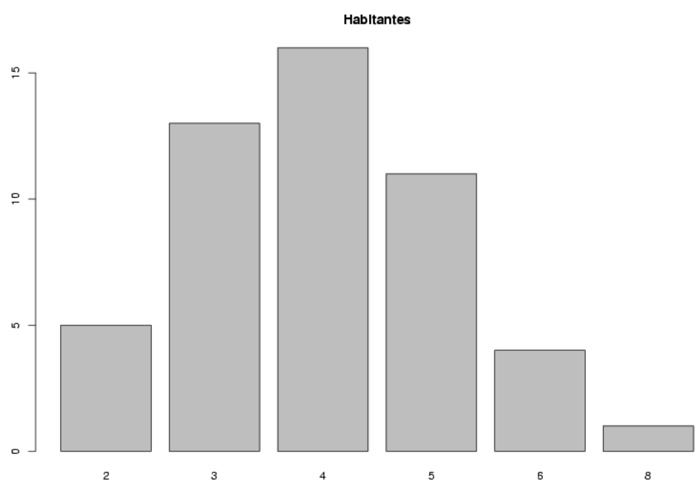
Podemos mejorarlo un poco con:

```
> pie(frec.Habitantes,main="Habitantes")
```



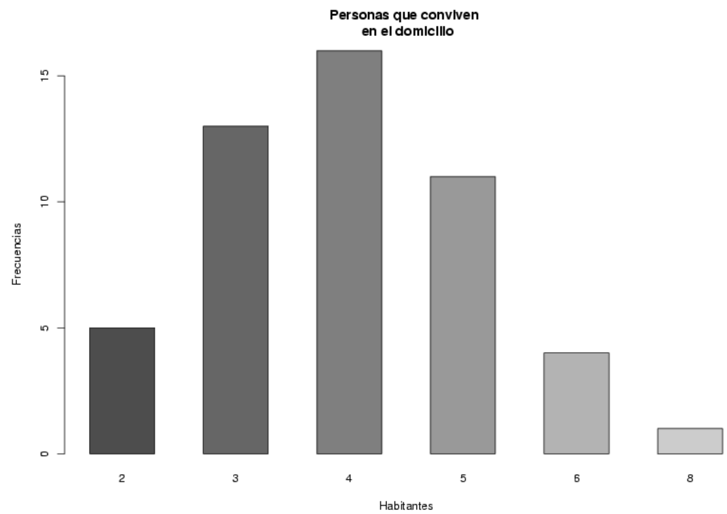
Diagramas de barras Realicemos ahora varios diagramas de barras

```
>barplot(frec.Habitantes,main="Habitantes")
```



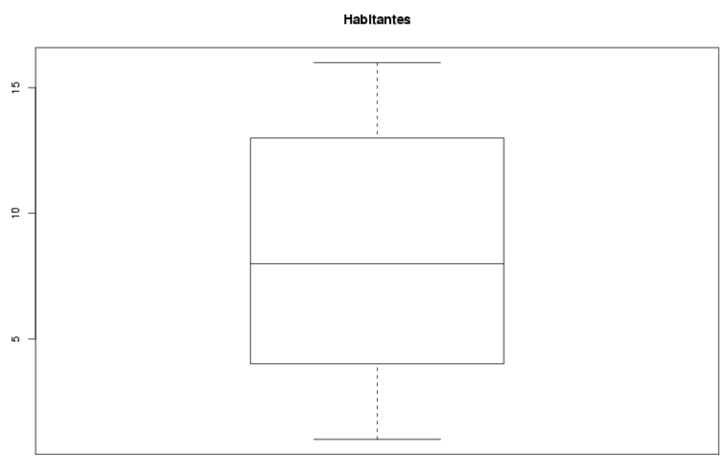
Mejor con algunos retoques:

```
> barplot(frec.Habitantes,main="Personas que conviven \n en el domici-
lio", xlab="Habitantes", ylab="Frecuencias", col=c("gray30", "gray40", "gray50", "gray60", "gra
ce=.75)
```

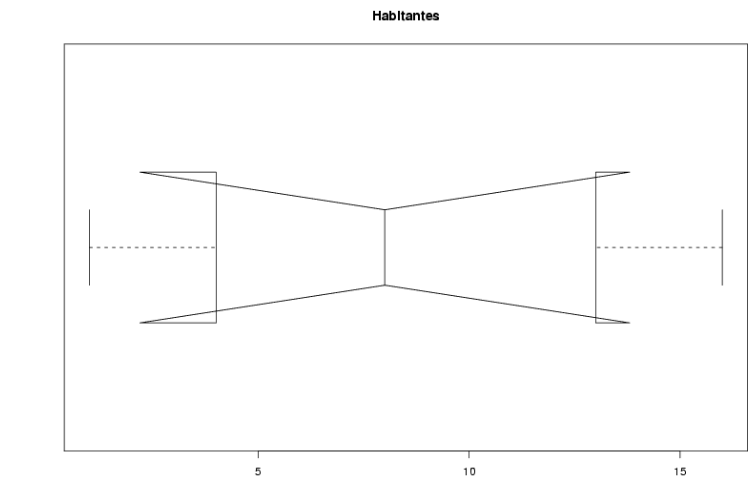


Gráfica de caja y bigotes

```
> boxplot(frec.Habitantes,main="Habitantes")
```



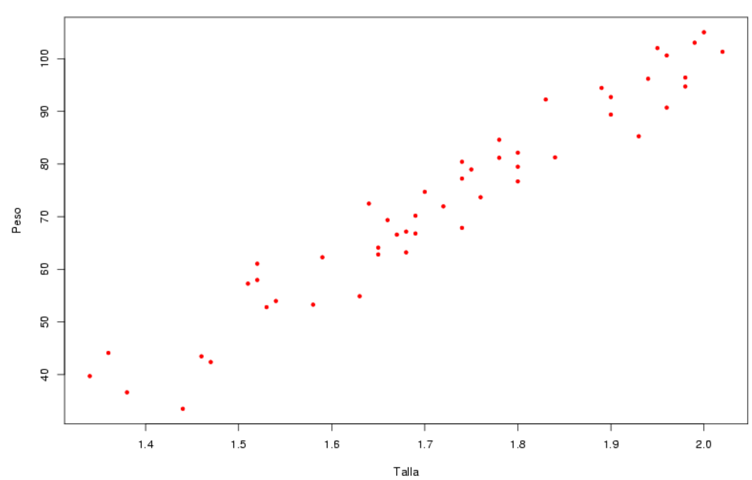

```
> boxplot(frec.Habitantes,main="Habitantes",horizontal=TRUE,notch=TRUE)
```



4.2.3. Estadística bidimensional:

Obtengamos la covarianza, coeficiente de correlación y la nube de puntos (modificando el "punto" por defecto a relleno y color rojo) de las dos variables cuantitativas

```
> cov(Talla,Peso)
[1] 3.370751
> cor(Talla,Peso)
[1] 0.9676415
> plot(Talla, Peso, pch=20, col="red")
```



`lm` (*lineal model*) función que realiza los modelos lineales en R. No da ninguna salida por pantalla, con ella creamos un objeto que va a ser un modelo de regresión lineal. Tiene múltiples opciones, para verlas podéis escribir en la línea de comandos `?lm` y accederéis a la ayuda de R sobre el procedimiento. Su sintaxis básica es:

```
lm(formula)
```

En `formula` ponemos el modelo expresado de la forma `y ~ x1+x2+...+xn`

Con:

```
> lm(Talla ~ Peso)

Call:
lm(formula = Talla ~ Peso)

Coefficients:
(Intercept)         Peso
   1.038261     0.009356
```

otenemos la ordenada en el origen y la pendiente de la recta de regresión de `Talla` sobre `Peso`, y con:

```
> lm(Peso ~ Talla)

Call:
lm(formula = Peso ~ Talla)

Coefficients:
(Intercept)        Talla
     -99.26       100.08
```

la ordenada en el origen y la pendiente de la recta de regresión de `Peso` sobre `Talla.`

Si lo almacenamos en un objeto (hacemos un ajuste de la variable `Talla` sobre la variable `Peso` y lo almacenamos en `regrePT`), podemos trabajar mejor

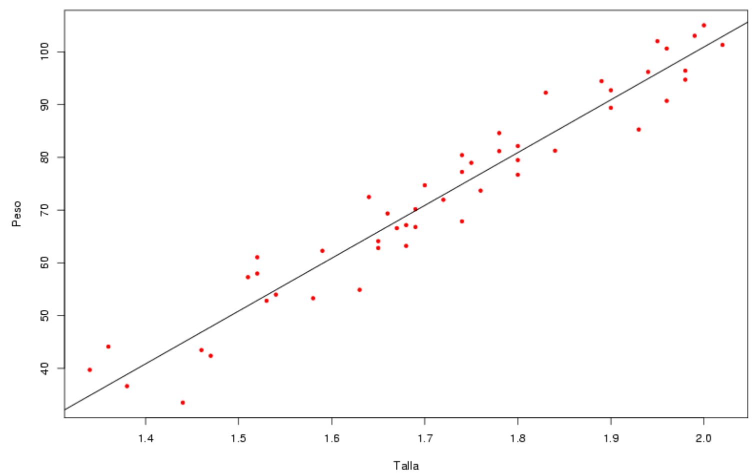
```
> regrePT <-lm(Peso~Talla)
```

La función `abline` permite añadir una línea, definida por la ordenada en el origen y su pendiente, a un gráfico que esté activo. La secuencia de órdenes[31] que siguen producirán el gráfico de abajo.

```
> plot(Talla, Peso, pch=20, col="red")
> abline(lm(Peso~Talla))
```

[31] Podemos conseguir "casi" lo mismo con:

```
> plot(formula(regre))
> abline(regre)
```



Para practicar Ambas rectas de regresión sobre los mismos ejes

La función `abline`, lo que hace es representar una recta en el gráfico activo a partir de dos números: la ordenada en el origen y la pendiente de la recta.

↬ Por ejemplo, para representar la recta de ecuación $y = -2x + 3$ escribiremos la sentencia

```
> abline(3,-2)
```

Hay que tener en cuenta que debemos tener alguna ventana gráfica activa. Si no es así, se iniciará una pero no obtendremos nada, nos dará error.

Para hacer este apartado sólo hemos de tener en cuenta que la recta de regresión de x sobre y, será de la forma:

$$r_{x/y} : x = m \cdot y + n{:}y = \frac{1}{m} \cdot x - \frac{n}{m} \tag{1}$$

Con la relación anterior, y teniendo en cuenta que

```
> coef(regrePT)
(Intercept)       Talla
   -99.2627    100.0812
```

nos permite acceder a la pendiente y ordenada en el origen de la recta de regresión, creemos la varibale que almacenará el otro ajuste lineal

```
> regreTP<-lm(Talla~Peso)
```

y veamos qué valores toma la pendiente y ordenada en el origen

```
> coef(regreTP)
(Intercept) Peso
1.038261217 0.009355705
```

Para que el proceso quede más “claro”, almacenemos esos valores en un par de variables de nombre descriptivo

```
> pendiente <- coef(regre)[2]
> ordenada <- coef(regre)[1]
```

Hagamos las transformaciones necesarias deducidas de 1

```
> m <- 1/pendiente
> n <- (-1)*ordenada/pendiente
```

Con esa información ya es fácil, solo hemos de pintar en el gráfico la nueva recta a partir de los datos anteriores, es decir:

1. Reperesentemos la nube de puntos

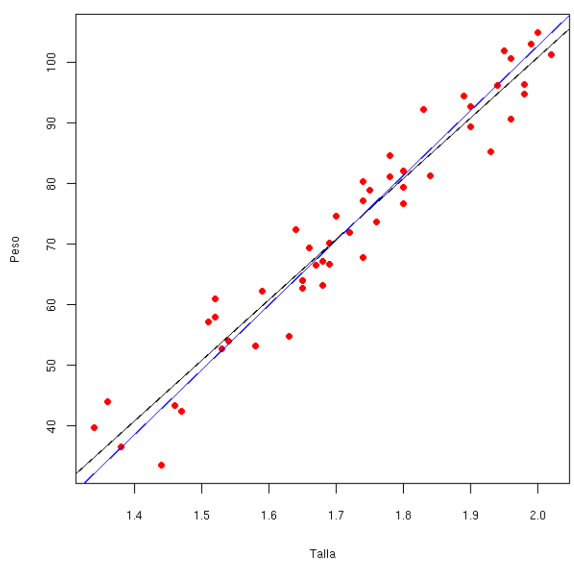
   ```
   > plot(Talla, Peso, pch=20, col=■red■)
   ```

2. Representemos la recta de regresión asociada a esa nube de puntos (Peso sobre Talla)

   ```
   > abline(regrePT)
   ```

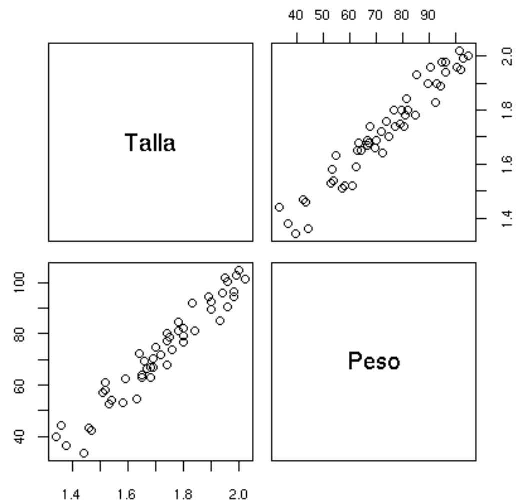
3. Representemos la segunda recta de regresión (Talla sobre Peso)

   ```
   > abline(n,m,col="blue")
   ```



Continuemos, veamos la salida gráfica del comando:

```
> pairs(formula(regre))
```



Por último, con[32]

```
> summary(regre)
Call:
lm(formula = Talla ~ Peso)
Residuals:
  Min        1Q       Median      3Q     Max
-0.090895 -0.029947 -0.001555 0.033607 0.094071
Coefficients:
             Estimate Std. Error t value Pr(>|t|)
(Intercept) 1.0382612 0.0265428 39.12 <2e-16 ***
Peso        0.0093557 0.0003521 26.57 <2e-16 ***
---
Signif. codes: 0 ‘***’ 0.001 ‘**’ 0.01 ‘*’ 0.05 ‘.’ 0.1 ‘ ’ 1
Residual standard error: 0.04679 on 48 degrees of freedom
Multiple R-Squared: 0.9363, Adjusted R-squared: 0.935
F-statistic: 705.9 on 1 and 48 DF, p-value: < 2.2e-16
```

mostramos el resultado. Es un excelente modelo lineal ya que el `Multiple R-Squared: 0.9363` (coeficiente de determinación) se acerca mucho a 1.

Con:

```
> predict(regre)
    1        2        3        4        5        6        7        8
1.551702 1.351677 1.874474 1.536733 1.380867 1.781759 1.716269 1.835929
    9       10       11       12       13       14       15       16
1.992543 1.620935 1.797664 1.979539 1.921720 1.580518 1.673139 1.629542
   17       18       19       20       21       22       23       24
2.020704 1.661071 1.444860 1.940151 1.434756 1.790553 1.905535 1.625893
```

[32] Una vez definido el objeto `regre`, para pintar la línea también podíamos haber escrito

```
> abline(regre)
```

Merece la pena probar

```
>plot(regre)
```

```
   25       26       27       28       29       30       31       32
1.409748 1.776903 1.666497 1.711311 1.637962 1.755657 1.938093 2.002170
   33       34       35       36       37       38       39       40
1.450895 1.694657 1.663129 1.574062 1.924340 1.798412 1.985994 1.727590
   41       42       43       44       45       46       47       48
1.886824 1.829660 1.543189 1.806645 1.532242 1.760709 1.737132 1.686986
   49       50
1.901231 1.609427
```

hallamos los valores esperados[33] de `Talla` para los 50 valores de la variable `Peso`.

Como no he encontrado la función que me permite hacer prediciones he optado por construírmela yo, sólo hay que tener en cuenta que:

```
> coef(regre)
(Intercept)        Peso
1.038261217 0.009355705
```

Así que:

```
> pendiente <- coef(regre)[2]
> ordenada  <- coef(regre)[1]
> estima    <- function(x){pendiente*x + ordenada}
```

Para estimar la estatura de una persona que pese $100Kg$ escribiremos

```
> estima(100)
    Peso
1.973832
> #Podemos comprobar que la fórmula está bien
> estima(Peso)
 [1] 1.551702 1.351677 1.874474 1.536733 1.380867 1.781759 1.716269 1.835929
 [9] 1.992543 1.620935 1.797664 1.979539 1.921720 1.580518 1.673139 1.629542
[17] 2.020704 1.661071 1.444860 1.940151 1.434756 1.790553 1.905535 1.625893
[25] 1.409748 1.776903 1.666497 1.711311 1.637962 1.755657 1.938093 2.002170
[33] 1.450895 1.694657 1.663129 1.574062 1.924340 1.798412 1.985994 1.727590
[41] 1.886824 1.829660 1.543189 1.806645 1.532242 1.760709 1.737132 1.686986
[49] 1.901231 1.609427
```

Y deben salir los mismos resultados anteriores.

Si lo que deseamos es la tabla del análisis de la varianza usamos la instrucción `anova(objeto)`:

```
> anova(regre)
Analysis of Variance Table

Response: Talla
          Df  Sum Sq Mean Sq F value    Pr(>F)
Peso       1 1.54525 1.54525  705.89 < 2.2e-16 ***
Residuals 48 0.10508 0.00219
---
Signif. codes:  0 ‘***’ 0.001 ‘**’ 0.01 ‘*’ 0.05 ‘.’ 0.1 ‘ ’ 1
```

Ejercicios:

Realiza con R las actividades realizadas con **Gnumeric** sobre estadística descriptiva y correlación. Varias notas a tener en cuenta:

[33]Para conocer la sintaxis completa `?predict.lm`

- El ejercicio sobre las notas de alumnos trabaja con dos grupos de diferente tamaño. Para que R no dé problemas al leer los datos de un fichero con columnas de diferente tamaño, hay que conseguir que lo sean escribiendo un último valor `Na` en la columna del menor número de datos (para que el tamaño de la muestra sea igual).
- R trabaja por defecto con punto decimal, si tenemos datos separados por coma decimal, podemos:
  - Usar un editor (**gedit**) para sustituir la , por .
  - Cuando usamos el comando `read.table`, usar la opción `dec=","`

4.3. Combinatoria y numeros aleatorios

choose(m,n) $\binom{m}{n}$ combinaciones

factorial(n) $n!$ como su nombre indica

```
> choose(5, 2)
[1] 21
> for (n in 0:10) print(choose(n, k = 0:n))
[1] 1
[1] 1 1
[1] 1 2 1
[1] 1 3 3 1
[1] 1 4 6 4 1
[1]  1  5 10 10  5  1
[1]  1  6 15 20 15  6  1
[1]  1  7 21 35 35 21  7  1
[1]  1  8 28 56 70 56 28  8  1
[1]   1   9  36  84 126 126  84  36   9   1
[1]   1  10  45 120 210 252 210 120  45  10   1
> factorial(10)
[1] 3628800
```

Ejercicios:

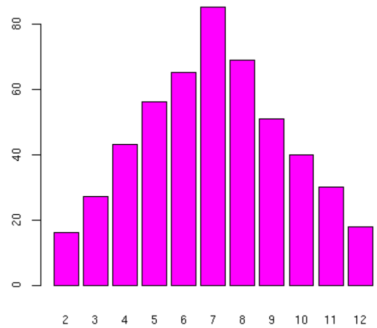
1. Realiza con R las actividades planteadas con **Gnumeric** sobre combinatoria (apartado 2.2.4 en la página 19).
2. Para resolver el ejercicio 2.2.3 en la página 18 con R podemos usar las funciones

   `runif(n,minimo,máximo)` devuelve `n` valores aleatorios distribuidos uniformemente comprendidos entre `mínimo` y `máximo`.

   `floor(x)` devuelve la parte entera del número pasado como argumento

   Usando ambas funciones, el ejercicio con 500 tiradas se reduce a

   ```
   > x<-floor(runif(500,1,7))
   > y<-floor(runif(500,1,7))
   > table(x+y)
    2  3  4  5  6  7  8  9 10 11 12
   16 27 43 56 65 85 69 51 40 30 18
   ```



Si a esto añadimos la posibilidad de obtener las combinaciones:

```
>expand.grid(c("vaqueros","chandal"),c("camisa","camiseta","polo"),c("deportivos","zapatos"))
Var1 Var2 Var3
1 vaqueros camisa deportivos
2 chandal camisa deportivos
3 vaqueros camiseta deportivos
4 chandal camiseta deportivos
5 vaqueros polo deportivos
6 chandal polo deportivos
7 vaqueros camisa zapatos
8 chandal camisa zapatos
9 vaqueros camiseta zapatos
10 chandal camiseta zapatos
11 vaqueros polo zapatos
12 chandal polo zapatos
```

R no está mal ¿verdad?

4.4. Distribuciones Binomial y Normal

Las funciones son:

dbinom(k, n, p) función de probabilidad ($P\left(X=k\right)=\left(\begin{array}{c} n \\ k \end{array}\right) p^{k}\left(1-p\right)^{n-k}$)

pbinom(k, n, p) función de distribución ($P\left(X\leq k\right)=\sum_{i=0}^{k}P\left(X=i\right)$)

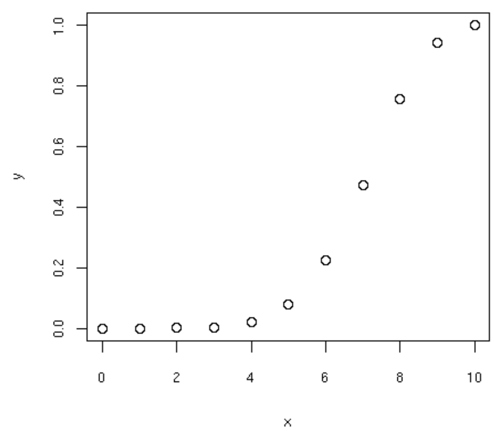
qbinom(P, n, p) calcula $x \mid P(X\leq x)=P$

rbinom(x, n, p) para obtener x números aleatorios de una distribución $B(n,\, p)$

```
> pbinom(10,10,0.75)
[1] 1
> pbinom(5,10,0.5)
[1] 0.6230469
> pbinom(5,10,0.5)-pbinom(4,10,0.5)
[1] 0.2460937
> dbinom(5,10,0.5)
[1] 0.2460938
```

Y una posible gráfica gráfica:

```
> x <- 0:10
> y <- pbinom(x,10,0.75)
> x <- 0:10
> y <- pbinom(x,10,0.75)
> plot(x,y)
```



La “normal”

dnorm(x, μ, σ**)** función de densidad $f(x) = \frac{1}{\sqrt{2\pi}\sigma} e^{-\frac{1}{2}\left(\frac{x-\mu}{\sigma}\right)^2}$

pnorm(x, μ, σ**)** función de distibución $F(x) = \int_{-\infty}^{x} f(t)\, dt$

qnorm(p, μ, σ**)** calcula $x \mid P(X \leq x) = p$

rnorm(x, μ, σ**)** para obtener x números aleatorios siguiendo una distribución normal

Si no escribimos ni la media ni la desviación típica, R trabaja con una $N\,(0,\,1)$

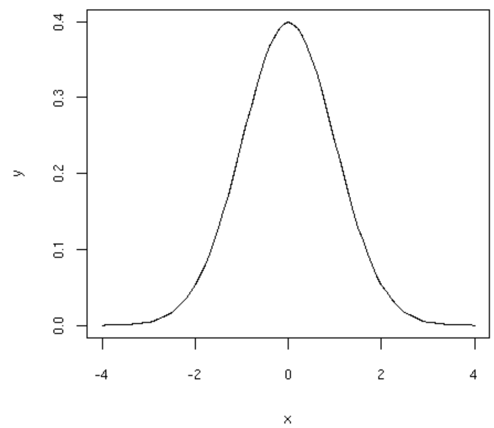
```
> pnorm(0)
[1] 0.5
> dnorm(0)
[1] 0.3989423
> rnorm(5)
[1] -0.63707195  0.94687649  1.68808897  0.05409749 -0.76185058
```

Si nuestro objetico es trabajar con una $N(\mu,\,\sigma)$, escribiremos $dnorm(x,\,\mu,\,\sigma)$:

```
> dnorm(0,10,4)
[1] 0.004382075
```

El gráfico obligado

```
> x <- seq(-4,4,by=0.1)
> y <- dnorm(x)
> plot(x,y,type="l")
```



Ejercicios:

1. Dibujar la función de distribución de la $N(0,1)$. El gráfico resultante estará en formato `pdf` y tendrá de nombre `acum_gauss.pdf`

2. Realiza con R las actividades planteadas sobre Binomial y Normal (ejercicios 3 y 4) para Gnumeric (2.2.5 en la página 20).

4.5. Inferencia

Nada mejor que un par de ejercicios "retocados" de selectividad:

1. Un fabricante de baterías de automóviles sabe que la vida de éstas sigue una ley normal y quiere informar sobre el promedio de vida en meses para sus baterías. Si elegida una muestra de seis de estas se tienen las siguientes duraciones (en meses):

   33 44 37 31 39 42

   a) Determine un intervalo de confianza del 95 % para dicho promedio de vida en meses.

   b) Determine un intervalo de confianza del 90 % para dicho promedio de vida en meses.

   c) ¿Qué podemos afirmar sobre la hipótesis de que $\mu = 35$?

   - Solución

     a) Se trata de obtener un intervalo de confianza para la media poblacional con σ conocida, es decir hay que hallar:

     $$\left(\bar{X} - Z_{\frac{\alpha}{2}}\,\sigma_{\bar{x}},\ \ \bar{X} + Z_{\frac{\alpha}{2}}\,\sigma_{\bar{x}}\right)$$

```
> baterias <- scan()
1: 33
2: 44
3: 37
4: 31
5: 39
6: 42
7:
Read 6 items
```

```
> t.test(baterias)
One Sample t-test
data: baterias
t = 18.283, df = 5, p-value = 9e-06
alternative hypothesis: true mean is not equal to 0
95 percent confidence interval:
32.37074 42.96259
sample estimates:
mean of x
37.66667
```

Ya es fácil: (32,37074, 42,96259)

a) Sólo hemos de escribir

```
> t.test(baterias,con.level=0.90)

        One Sample t-test

data:  baterias
t = 18.283, df = 5, p-value = 9e-06
alternative hypothesis: true mean is not equal to 0
95 percent confidence interval:
 32.37074 42.96259
sample estimates:
mean of x
 37.66667

> help(t.test)
> t.test(baterias,conf.level=0.90)

        One Sample t-test

data:  baterias
t = 18.283, df = 5, p-value = 9e-06
alternative hypothesis: true mean is not equal to 0
90 percent confidence interval:
 33.51525 41.81808
sample estimates:
mean of x
    37.66667
```

b) Con R la estadística es otra cosa

```
> t.test(baterias,mu=35)

        One Sample t-test

data:  baterias
t = 1.2944, df = 5, p-value = 0.2521
alternative hypothesis: true mean is not equal to 35
95 percent confidence interval:
 32.37074 42.96259
sample estimates:
mean of x
 37.66667
```

El `p-valor` es mayor que el nivel de significancia utilizado ($\alpha = 0{,}05$). Luego, no existen evidencias para rechazar la hipótesis de que la media no sea 35 (Hipótesis Nula Ho:).

Para casa

1. En 10 talleres elegidos al azar se ha pedido presupuesto para efectuar una cierta revisión a un mismo automóvil, obteniéndose los siguientes precios en euros:

   160.25; 162.15; 158.75; 169.50; 170.10

   165.20; 172.35; 175.00; 168.30; 167.10

   a) Determine cuáles están incluidos en el intervalo de confianza de nivel 95 % que se obtiene con esta muestra para la citada ley normal.

   b) ¿Podemos afirmar que la media del coste de la revisión de todos los talleres es de $\mu = 1{,}60$?

Apéndices

A. Actividades

Las que están hechas en los apuntes se marcan con ★

A.1. Gnumeric

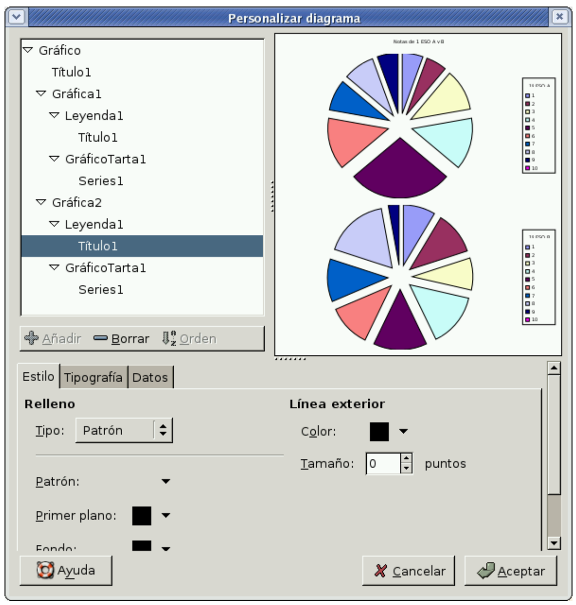
A.1.1. Estadística descriptica

1. Las notas de matemáticas de 1º ESO A y B han sido:

A	5, 6, 4, 5, 6, 5, 5, 8, 6, 3, 2, 1, 5, 5, 7, 6, 7, 3, 2, 3, 8, 9, 4, 4, 8, 9, 1, 5, 5, 4, 7, 6, 3, 4, 5, 5
B	1, 5, 7, 4, 3, 8, 6, 5, 4, 7, 3, 7, 1, 2, 4, 8, 9, 7, 8, 1, 2, 4, 5, 6, 8, 2, 3, 2, 8, 4, 5, 6, 5, 6, 8

★ ¿Qué clase obtuvo mejores resultados? Realiza un estudio estadístico de las notas de ambas clases.

a) Realizar un gráfico de sectores como el que sigue:



2. La edad de 130 personas que realizaron un experimento se da a continuación:

15	17	19	23	22	17	19	24	15	17	23	19	23	18	17	16	24	19	17	18
15	24	19	17	22	19	19	20	21	22	17	17	21	16	16	18	24	21	19	19
19	20	21	21	23	18	23	18	17	17	20	23	17	21	23	17	16	19		
21	17	18	21	19	18	21	22	17	19	24	20	21	21	16	17	19	19		
22	17	21	18	23	18	16	23	18	17	22	19	15	17	19	17	16	22		
16	23	18	17	21	24	21	21	18	18	19	24	16	22	23	19	17	20		
20	22	22	17	17	19	19	15	20	19	18	20	17	15	17	18	18	21		

a) Agrupa las edades por años.

b) Da una tabla de frecuencias simple (para cada año) y acumulada.

c) Realiza un diagrama de barras verticales y uno de sectores.

A.1.2. Correlación

1. La tabla siguiente muestra las respectivas alturas X e Y de una muestra de 12 padres y sus hijos primogénitos (en pulgadas):

Altura X del padre (en pulg)	65	63	67	64	68	62	70	66	68	67	69	71
Altura Y del hijo (en pulg)	68	66	68	65	69	66	68	65	71	67	68	70

   a) ★ Construye el diagrama de dispersión.
   b) ★Estudia la correlación entre ambas variables.
   c) ★ Halla la recta de regresión de Y sobre X.
   d) Obtener la Talla estimada de un hijo cuyo padre mida:
      1) 60
      2) 70
      3) 80
      4) 100

      Comenta los resultados
   e) Obtener los valores estimados de la Talla de un Padre si el Hijo mide:
      1) 60
      2) 70
      3) 80
      4) 100
   f) Representa la nube de puntos así como las dos rectas de regresión[34].

2. En la tabla siguiente damos la evolución del récord del mundo de salto de longitud masculina, en metros.

Años	1901	1921	1925	1931	1935	1961	1962	1965	1968	1991
Marca	7,61	7,69	7,89	7,98	8,13	8,28	8,31	8,35	8,90	8,95

   a) Representa la nube de puntos asociada a estos datos.
   b) ¿Puede estimarse la marca del récord por regresión lineal? Si es así, ¿cuánto se saltará en los años 2000 y 2005?

A.1.3. Muestras y números aleatorios

1.

   a) ★ Generar 1000 números aleatorios enteros dentro del rango 1.,100 siguiendo una distribución uniforme. En una nueva Hoja obtener una muestra aleatoria de tamaño 100.

      Para esto pulsamos sobre **Editar→Llenar→Generación de números aleatorios** y rellenamos los campos de la pestaña **Números aleatorios.** Como lo queremos en 10 columnas de 100 datos cada una escribiremos los valores adecuados en **Número de variables** y **Tamaño de la muestra** (en la pestaña **Opciones**). Los valores de la tercera pestaña son opcionales y si la dejamos como está obtendremos una tabla de 10x100 números aleatorios.

      Para obtener la muestra aleatoria de tamaño 100, pulsemos sobre **Herramientas→Análisis Estadístico→Muestreo**
   b) De esa muestra hay que hacer un análisis estadístico y realizar algunas gráficas de ella. Después, usar la muestra y lo visto antes para realizar el estudio de esos datos.

[34]Que no se nos olvide que para pintar una recta sólo necesitamos dos puntos por los que pasa.

A.1.4. Combinatoria

1.
   a) Halla 17!
   b) Seis amigos corren en una carrera, ¿de cuántas formas pueden llegar?
2. Halla:
   a) $V_{10,\,3}$
   b) ¿Cuántos números de 5 cifras se pueden formar con los dígitos 1 al 9?
3.
   a) Halla $C_{10,\,4}$
   b) Comprueba con varios números que
      1) $C_{m,\,m-n} = C_{m,\,n}$
      2) $C_{m,\,n} + C_{m,\,n+1} = C_{m+1,\,n+1}$
   c) ¿Cuántos grupos de 4 pueden formarse en una clase de 30 alumnos?

A.1.5. Distribuciones de probabilidad

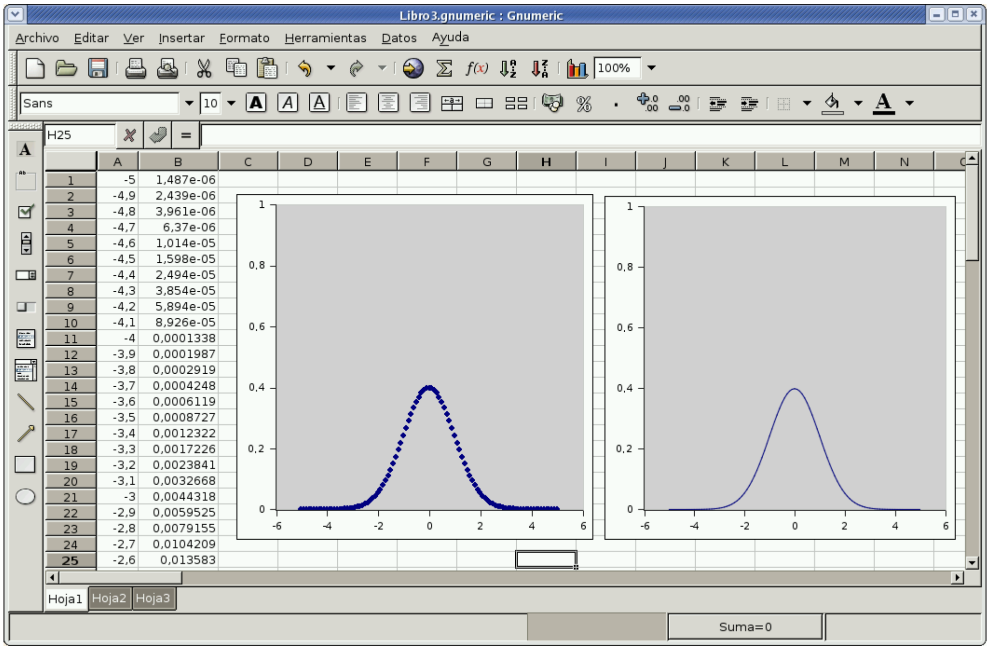
Distribución binomial

1. Un examen de preguntas con respuestas múltiples consta de 8 preguntas con 4 opciones de contestación. Si un alumno responde al azar, halla, con la ayuda de la tabla binomial[35], la probabilidad de que:
   a) Responda correctamente a 6.
   b) Responda menos de 6 correctamente.
2. La probabilidad de que en una empresa minera de 10.000 empleados haya uno enfermo es de 0,2. Cuál será la probabilidad de que en un momento determinado haya en la empresa:
   a) 1.980 enfermos.
   b) Más de 2.080 enfermos.

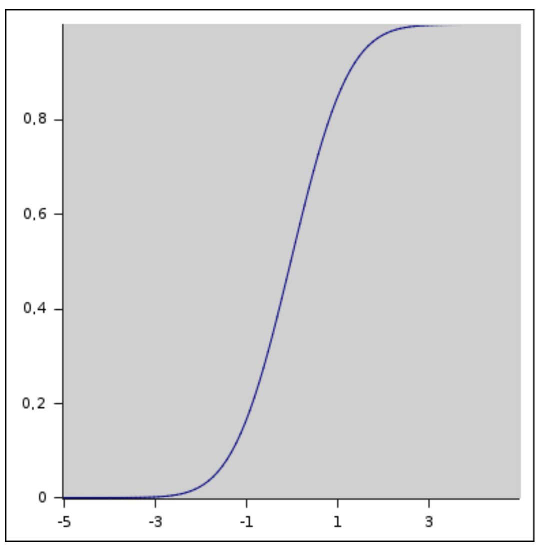
Distribución normal

1. ★ Representar una curva normal.

[35] Perdón, con la ayuda de Gnumeric



2. Representa la gráfica de la distribución normal acumulada



3. Se fabrican unas pilas alcalinas cuya duración en horas sigue una distribución normal de media 60 h y desviación típica 5 h. Si se elige una pila al azar, ¿qué probabilidad hay de que dure?:

   a) Menos de 50 h.

   b) Entre 52 y 65 h.

4. El peso (P) de los socios del Club de Amigos del Buen Comer ha resultado que se distribuye conforme a una normal $N(92, 20)$. ¿Qué peso máximo tienen el 10 % de los individuos menos obesos de tan opíparo club?

A.2. Grace

A.2.1. Estadística descriptiva

1. ★ Las notas de matemáticas de 1^{o} ESO A han sido:

A	5, 6, 4, 5, 6, 5, 5, 8, 6, 3, 2, 1, 5, 5, 7, 6, 7, 3, 2, 3, 8, 9, 4, 4, 8, 9, 1, 5, 5, 4, 7, 6, 3, 4, 5, 5

Haz un estudio estadístico de ellas.

2. La edad de 130 personas que realizaron un experimento se da a continuación:

15	17	19	23	22	17	19	24	15	17	23	19	23	18	17	16	24	19	17	18
15	24	19	17	22	19	19	20	21	22	17	17	21	16	16	18	24	21	19	19
19	20	21	21	23	18	23	18	17	17	20	23	17	21	23	17	16	19		
21	17	18	21	19	18	21	22	17	19	24	20	21	21	16	17	19	19		
22	17	21	18	23	18	16	23	18	17	22	19	15	17	19	17	16	22		
16	23	18	17	21	24	21	21	18	18	19	24	16	22	23	19	17	20		
20	22	22	17	17	19	19	15	20	19	18	20	17	15	17	18	18	21		

a) Introduce los datos en un fichero de texto a dos columnas (separadas por un espacio). En la primera columna se escribirá el número de orden del dato y en la segunda el dato (Por ejemplo: 1 15 hasta 130 19). Para eso puedes usar

1) Un editor cualquiera y guardarlos en texto plano.
2) Rellenar con Gnumeric una columna con una serie del 1 al 130, en la segunda introducir los datos. Por último hay que exportarlos a texto plano.
3) Usar el editor del programa Grace.

b) Agrupa las edades por años.

c) Da una tabla de frecuencias simple (para cada año).

A.2.2. Correlación

1. ★ La tabla siguiente muestra las respectivas alturas X e Y de una muestra de 12 padres y sus hijos primogénitos (en pulgadas):

Altura X del padre (en pulg)	65	63	67	64	68	62	70	66	68	67	69	71
Altura Y del hijo (en pulg)	68	66	68	65	69	66	68	65	71	67	68	70

a) Construye el diagrama de dispersión.

b) Halla la recta de regresión de Y sobre X.

2. En la tabla siguiente damos la evolución del récord del mundo de salto de longitud masculina, en metros.

Años	1901	1921	1925	1931	1935	1961	1962	1965	1968	1991
Marca	7,61	7,69	7,89	7,98	8,13	8,28	8,31	8,35	8,90	8,95

Nota: Puedes trabajar a partir de los datos ya introducidos en Gnumeric. Hay que exportarlos a texto plano, y después, con un editor sustituir la coma decimal por un punto.

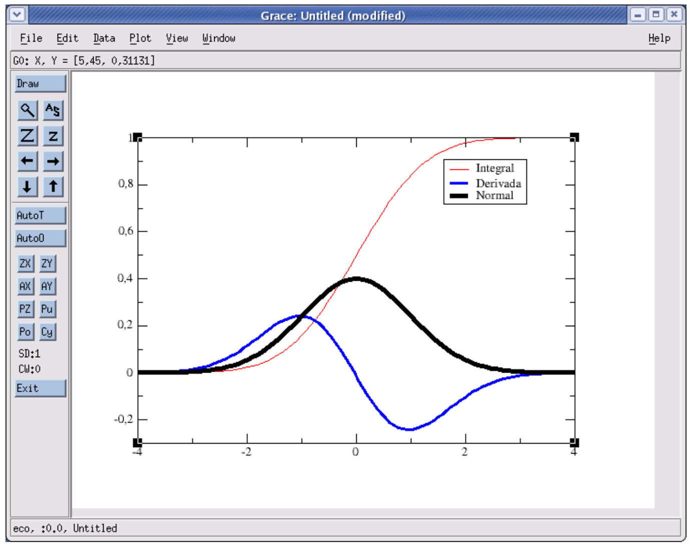
a) Representa la nube de puntos asociada a estos datos.

b) ¿Puede estimarse la marca del récord por regresión lineal? Si es así, ¿cuánto se saltará en los años 2000 y 2005?

c) Usa la Web: `http://tux.iesmurgi.org/matematicas/materiales/correlacion/` para comprobar los resultados del ejercicio anterior y haz 10 estimaciones de la marca que se obtendrá en años venideros y otras 10 del año en que se alcanzarán las marcas (pon 5 valores "extremos" y 5 dentro del rango de las variables). Comenta los resultados.

A.2.3. Un añadido: gráficas, integrales y derivadas

1. ★ Representar la función de densidad de la distribución normal, su derivada y su integral.



2. Representar la parábola $f(x) = x^2 - 2x$ en el intervalo $[-1, 4]$, la gráfica de su derivada y su integral.

3. Transformar la curva anterior usando **Data Trasformations** $\rightarrow$ **Evaluate expresion**. Seleccionamos la curva a transformar (S1, o S2 o S3... en **source**). No seleccionamos ninguna curva en **Destination** si deseamos crear una nueva curva. Introducimos la fórmula de la transformación (por ejemplo $y = y + 1$). Seleccionamos en **restriction** la región de datos que queremos transformar (ver apartado definir regiones). Por último **Apply** y luego **Close**

A.3. R

A.3.1. Varios

1. Halla todas las posibles formas de mezclar los números 4 al 7 con las letras "a","b" y "c" y con los caracteres "+" y "-"

2. ★ Dadas las matrices: $A = \begin{pmatrix} 1 & 2 & 4 \\ 2 & -1 & 0 \\ 1 & 0 & 1 \end{pmatrix}$ y $B = \begin{pmatrix} 1 & 2 & 3 \\ 1 & 4 & 9 \\ 1 & 8 & 27 \end{pmatrix}$ y $C = \begin{pmatrix} 1 \\ 2 \\ 3 \end{pmatrix}$. Halla:

   a) La diagonal principal de A. A^{-1}, A^t y $|A|$.

   b) $A + B$ y $A \cdot B$

   c) Resuelve el sistema $A \cdot x = C$

A.3.2. Descriptiva y Correlación.

1. ★ La talla, estudios, habitantes y peso de una población de 50 individuos está recogida a continuarción.

Talla Estudios Habitantes Peso	1.58 bachiller 3 53.28	1.93 bachiller 5 85.26
1.63 bachiller 2 54.88	1.38 bachiller 5 36.62	1.95 bachiller 2 102
1.44 fp 3 33.5	1.8 bachiller 6 79.47	1.59 bachiller 2 62.28
1.9 bachiller 3 89.38	1.64 superior 3 72.47	1.78 primario 4 81.17

1.96 primario 4 100.61	1.34 fp 8 39707	1.84 bachiller 3 81.25
1.89 bachiller 5 94.43	1.75 diplomado 2 78.951	2.02 bachiller 6 101.3
1.52 bachiller 4 57.96	1.68 diplomado 3 67.15	1.76 primario 4 73.68
1.74 diplomado 4 67.86	1.72 bachiller 4 71.94	1.96 fp 3 90.7
1.68 diplomado 5 63.2	1.65 bachiller 4 64.1	1.78 fp 3 84.59
2 diplomado 6 105.01	1.8 bachiller 6 76.68	1.54 superior 4 53.97
1.67 diplomado 3 66.57	1.94 bachiller 4 96.18	1.8 diplomado 5 82.13
1.46 primario 5 43.46	1.99 primario 4 103.029	1.53 diplomado 4 52.8
1.98 primario 4 96.4	1.36 primario 3 44.105	1.74 diplomado 5 77.22
1.47 primario 2 42.38	1.69 primario 5 70.16	1.7 primario 3 74.7
1.74 primario 3 80.41	1.69 bachiller 4 66.79	1.66 primario 4 69.34
1.9 diplomado 4 92.7	1.51 bachiller 3 57.27	1.83 primario 5 92.24
1.65 fp 5 62.81	1.98 bachiller 5 94.71	1.52 bachiller 4 61.05

Realiza un estudio estadístico de ellos. Representa gráficamente los datos. Estudia la correlación entre las variables Peso y Talla.

2. Realiza con R las actividades realizadas con **Gnumeric** sobre estadística descriptiva (notas de matemáticas, edad de 130 personas) y correlación (altura Padre e Hijo, récord del mundo de salto). Varias notas a tener en cuenta:

   - El ejercicio sobre las notas de alumnos trabaja con dos grupos de diferente tamaño. Para que R no dé problemas al leer los datos de un fichero con columnas de diferente tamaño, hay que conseguir que lo sean escribiendo un último valor `Na` en la columna del menor número de datos (para que el tamaño de la muestra sea igual).
   - R trabaja por defecto con punto decimal, si tenemos datos separados por coma decimal, podemos:
     - Usar un editor (**gedit**) para sustituir la , por .
     - Cuando usamos el comando `read.table`, usar la opción `dec=","`

A.3.3. Combinatoria

1. ★ Halla un triángulo de Tartaglia hasta la décima fila.

2. Realiza con R las actividades planteadas con **Gnumeric** sobre combinatoria (apartado A.1.4 en la página 66).

3. ★ Halla todas las combinaciones posibles con la que Marta se puede vestir si dispone de pantalones vaqueros y de chandal, camisa, camiseta y polo y zapatos y zapatillas de deporte.

A.3.4. Distribuciones Binomial y Normal

1. ★ Halla:

   a) $P(X \leq 10),\ X \to B(10,\ 0{,}75)$
   b) $P(X \leq 5),\ X \to B(10,\ 0{,}5)$
   c) $P(X = 5),\ X \to B(10,\ 0{,}75)$
   d) La gráfica de la función de distribución de $X \to B(10,\ 0{,}75)$

2. ★ Halla

   a) $P(X \leq 0),\ X \to N(0,\ 1)$
   b) $f(0),\ X \to N(0,\ 1)$

   c) Halla 5 números aleatorios según una $N(0,\ 1)$
   d) $f(0)$, $X \to N(10,\ 4)$

3. ★ Representa las gráficas de la función de densidad de una $N(0,1)$
4. Dibujar la función de distribución de la $N(0,1)$. El gráfico resultante estará en formato `pdf` y tendrá de nombre `acum_gauss.pdf`
5. Realiza con R las actividades planteadas sobre Binomial y Normal (ejercicios 3 y 4) para Gnumeric (A.1.5 en la página 66).

A.3.5. Inferencia

1. ★ Un fabricante de baterías de automóviles sabe que la vida de éstas sigue una ley normal y quiere informar sobre el promedio de vida en meses para sus baterías. Si elegida una muestra de seis de estas se tienen las siguientes duraciones (en meses):

   33 44 37 31 39 42

   a) Determine un intervalo de confianza del 95 % para dicho promedio de vida en meses.
   b) Determine un intervalo de confianza del 90 % para dicho promedio de vida en meses.
   c) ¿Qué podemos afirmar sobre la hipótesis de que $\mu = 35$?

2. En 10 talleres elegidos al azar se ha pedido presupuesto para efectuar una cierta revisión a un mismo automóvil, obteniéndose los siguientes precios en euros:

   160.25; 162.15; 158.75; 169.50; 170.10

   165.20; 172.35; 175.00; 168.30; 167.10

   a) Determine cuáles están incluidos en el intervalo de confianza de nivel 95 % que se obtiene con esta muestra para la citada ley normal.
   b) ¿Podemos afirmar que la media del coste de la revisión de todos los talleres es de $\mu = 1{,}60$?

B. Funciones Estadísticas de Gnumeric[36]

AVEDEV(n1; n2; ...)

AVEDEV devuelve la desviación media. (E)

AVERAGE(valor1; valor2;...)

AVERAGE calcula la media de todos los valores y celdas referenciadas en la lista de argumentos. Es equivalente a la suma de los argumentos dividido por el número de argumentos. (E)

AVERAGEA(valor1;valor2;...)

Calcula la media de los valores y celdas referenciados en la lista de argumentos. Se incluyen en el cálculo los números y también los textos y expresiones lógicas. Si la celda contiene texto o la expresión lógica FALSO, se contará como un cero (0). Si contiene un valor lógico VERDADERO se contará como un uno (1). Note que las celdas vacías no son contadas. (E)

RANDBERNOULLI(k;p)

BERNOULLI devuelve la probabilidad p(k) de obtener k de una distribución de Bernoulli con parámetro de probabilidad p.

- Si k != 0 y k != 1, BERNUOLLI devuelve un error #NÚM!
- Si p <0 o p >1 BERNOULLI devuelve un error #NÚM!

BETADIST(x;alfa;beta[,a,b])

La función BETADIST devuelve la distribución beta acumulada. a es el límite inferior opcional de x y b es el límite superior opcional de x. (E)

- Si no se da a, BETADIST usa 0.
- Si no se da b, BETADIST usa 1.
- Si x <a ó x >b, BETADIST devuelve el error #NÚM!.
- Si alfa <= 0 ó beta <= 0, BETADIST devuelve el error #NÚM!.
- Si a >= b BETADIST devuelve el error #NÚM!.

BETAINV(p;alfa;beta[;a;b])

La función BETAINV devuelve la inversa de la distribución beta acumulada. a es el límite inferior opcional de x y b es el límite superior opcional de x. (E)

- Si no se de a, BETAINV usa 0.
- Si no se da b, BETAINV usa 1.
- Si p <0 ó p >1, BETAINV devuelve el error #NÚM!.
- Si alfa <= 0 o beta <= 0, BETAINV devuelve el error #NUM.
- Si a >= b, BETAINV devuelve el error #NÚM!.

[36] Las que son compatibles con Excel se marcan con (E)

BINOMDIST(n;intentos;p;acumulado)

La función BINOMDIST devuelve la distribución binomial. n es el número de éxitos, intentos es el número total de intentos independientes, p es la probabilidad de éxito en intentos, y acumulado describe si debe devolver la suma de las funciones binomiales de 0 a n. (E)

- Si n ó intentos no son enteros, se truncan.
- Si n <0 ó intentos <0, BINOMIST devuelve el error #NÚM!.
- Si n >intentos, BINODIST devuelve el error #NÚM!.
- Si p <0 ó p >1, BINOMIST devuelve el error #NÚM!.

CAUCHY(x;a;cum)

CAUCHY devuelve la distribución de Cauchy con parámetro de escala a. Si cum es VERDADERO, CAUCHY devuelve la distribución acumulada.

- Si a <0 CAUCHY devuelve un error #NÚM!.
- Si cum != VERDADERO y cum != FALSO CAUCHY devuelve un error #VALOR!.

CHIDIST(x;gdl)

La función CHIDIST devuelve la probabilidad con una cola de la función chi cuadrado. gdl es el número de grados de libertad. (E)

- Si gdl no es entero, se trunca.
- Si gdl <1, CHIDIST devuelve el error #NÚM!

CHIINV(p;gdl)

La función CHIINV devuelve la inversa de la probabilidad de una cola de la distribución chi cuadrado. (E)

- Si p <0 ó p >1 ó gdl <1, CHIINV devuelve el error #NÚM!.

CHITEST(rango_actual;rango_teórico)

La función CHITEST devuelve las pruebas para la independencia de la distribución chi cuadrado. (E)

rango_actual es un rango que contiene los puntos de datos observados. rango_teórico es un rango que contiene los valores esperados de los puntos de datos.

CONFIDENCE(x;desvstd;tamaño)

CONFIDENCE devuelve el intervalo de confianza de la media. x es el nivel de significancia, desvstd es la desviación estándar de la población y tamaño corresponde al tamaño de la muestra. (E)

- Si tamaño no es un número entero, entonces se trunca.
- Si tamaño <0, CONFIDENCE devuelve el error #NÚM!.
- Si tamaño es 0, CONFIDENCE devuelve el error #DIV/0!.

CORREL(array1;array2)

Devuelve el coeficiente de correlación de dos conjuntos de datos. (E)

- Las celdas vacías o con texto son ignoradas.

COUNT(valor1; valor2; ...)

Devuelve el número total de argumentos enteros o de coma flotante pasados. (E)

COUNTA(valor1; valor2; ...)

Calcula el número de argumentos pasados sin incluir celdas vacías. (E)

COVAR(array1;array2)

Devuelve la covarianza de dos conjuntos de datos. (E)

- Las celdas vacías o con texto siempre son ignoradas.

CRITBINOM(intentos;p;alfa)

La función CRITBINOM devuelve el valor más pequeño para el que el acumulado es mayor o igual que un valor dado. n es el número de intentos, p es la probabilidad de éxito en intentos, y alpha es el valor de criterio. (E)

- Si intentos no es entero, se trunca.
- Si intentos <0, CRITBINOM devuelve un error #NÚM!.
- Si p <0 ó p >1, CRITBINOM devuelve el error #NÚM!.
- Si alfa <0 ó alfa >1, CRITBINOM devuelve el error #NÚM!.

CRONBACH(ref1;ref2;...)

CRONBACH devuelve el alfa de Cronbach para los casos dados. ref1 es un conjunto de datos, ref2 el segundo conjunto de datos, etc..

DEVSQ(n1; n2; ...)

DEVSQ devuelve la suma de los cuadrados de las desviaciones de un conjunto de datos de la media de su muestra. (E)

- Las cadenas y las celdas vacías simplemente se ignoran.

EXPONDIST(x;y;acumulado)

La función EXPONDIST devuelve la distribución exponencial (E). Si el booleano acumulado es falso devuelve:
$y \cdot e^{-y \cdot x}$, en otro caso, devuelve $1 - e^{-y \cdot x}$
Si x <0 ó y <= 0 esto devolverá un error.

EXPPOWDIST(x;a;b)

EXPPOWDIST devuelve la densidad de probabilidad p(x) en x para la distribución exponencial con el parámetro de escala a y exponente b.

FDIST(x;gdl1;gdl2)

La función FDIST devuelve la distribución de probabilidad F. gdl1 son los grados de libertad del numerador y gdl2 son los grados de libertad del denominador. (E)

- Si x <0, FDIST devuelve el error #NÚM!.
- Si gdl1 <1 ó gdl2 <1, FDIST devuelve el error #NÚM!.

FINV(p;gdl1;gdl2)

La función FINV devuelve la inversa de la distribución de probabilidad F. (E)

- Si p <0 ó p >1, FINV devuelve el error #NÚM!.
- Si gdl1 <1 ó gdl2 <1, FINV devuelve el error #NÚM!.

FISHER(x)

La función FISHER devuelve la transformación de Fisher en x. (E)

- Si x no es un número, FISHER devuelve el error #VALOR!.
- Si x <= -1 ó x >=1, FISHER devuelve el error #NÚM!.

FISHERINV(x)

La función FISHERINV devuelve la inversa de la transformación de Fisher en x. (E)

- Si x no es un número, FISHERINV devuelve el error VALOR!.

FORECAST(x;y_conocidos;x_conocidos)

La función FORECAST estima un valor futuro, según los valores existentes usando una regresión linear simple. El valor futuro estimado es el valor de y para un valor de x dado (x). (E)

- Si x_conocidos o y_conocidos no contienen datos o contienen cantidades de datos diferentes, FORECAST devuelve el error #N/D! .
- Si la varianza de x_conocidos es cero, FORECAST devuelve el error #DIV/0 .

FREQUENCY(array_datos;array_intervalos)

La función FREQUENCY cuenta cuantas veces aparecen los valores dados en un rango de valores. Los resultados se devuelven en un array. (E)

array_datos es un array de los datos de los que queremos contar las frecuencias. array_intervalos es un array que contiene los intervalos en los que reagrupar los valores de array_datos. Si array_intervalos está vacío, FREQUENCY devuelve la cantidad de puntos de datos en array_datos.

FTEST(array1;array2)

La función FTEST devuelve la probabilidad de que las varianzas de los dos conjuntos de datos dados no se diferencien significativamente. (E)

GAMMADIST(x;alfa;beta;acum)

La función GAMMADIST devuelve la distribución gamma. Si acum es VERDADERO, GAMMADIST devuelve la función gamma incompleta. En otro caso, devuelve la función de probabilidad de masa. (E)

- Si x <0, GAMMADIST devuelve el error #NÚM!.
- Si alfa <= 0 ó beta <= 0, GAMMADIST devuelve el error #NÚM!

GAMMAINV(p;alfa;beta)

La función GAMMAINV devuelve la inversa de la distribución gama acumulada. (E)

- Si p <0 ó p >1, GAMMAINV devuelve el error #NÚM!.
- Si alfa <= 0 ó beta <= 0, GAMMAINV devuelve el error #NÚM!.

GAMMALN(x)

La función GAMMALN devuelve el logaritmo natural de la función gamma. (E)

- Si x no es un número, entonces GAMMALN devuelve el error #VALOR!.
- Si x <= 0, entonces GAMMALN devuelve el error #NÚM!.

GEOMDIST(k;p;cum)

GEOMDIST devuelve la probabilidad p(k) de obtener k de una distribución geométrica con parámetro de probabilidad p.

- Si k <0 GEOMDIST devuelve un error #NÚM!
- Si p <0 o p >1 GEOMDIST devuelve un error #NÚM!
- Si cum != VERDADERO y cum != FALSO GEOMDIST devuelve #NÚM!

GEOMEAN(valor1; valor2; ...)

Devuelve la media geométrica de los argumentos, es decir, la raíz n-ésima del producto de los términos. (E)

GROWTH(y_conocido[;x_conocido;x_nuevo;const])

La función GROWTH aplica el método de los mínimos cuadrados para que se ajuste una curva exponencial a sus datos y prediga el crecimiento exponencial usando esta curva.

GROWTH devuelve un array que tiene una columna y una fila por cada punto de datos de nuevo_x.

- Si se omite x_conocido, se usa un array {1, 2, 3, ...}.
- Si se omite x_nuevo, se asume que es el mismo que x_conocido.
- Si y_conocido y x_conocido tienen distinto número de puntos de datos, GROWTH devuelve el error #NÚM!.
- Si const es FALSO, se forzará a que la línea pase por el origen, por ejemplo, b sería cero. Por omisión es VERDADERO.

HARMEAN(b1; b2; ...)

HARMEAN devuelve la media armónica de los N puntos de datos (esto es, N dividido por la suma de las inversas de los puntos de datos). (E)

HYPGEOMDIST(x;n;M;N)

La función HYPGEOMDIST devuelve la distribución hipergeométrica. x es el número de éxitos en la muestra, n es el número de intentos, M es el número de éxitos en total y N es el tamaño de la población. (E)

- Si x, n, M ó N no son enteros, se truncan.
- Si x, n, M ó N <0, HYPGEOMDIST devuelve el error #NÚM!.
- Si x >M ó n >N, HYPGEOMDIST devuelve el error #NÚM!.

INTERCEPT(y_conocidos;x_conocidos)

La función INTERCEPT calcula el punto en el cual la regresión lineal corta el eje y. (E)

- Si x_conocidos o y_conocidos no contienen datos o contienen diferente número de entrada de datos, INTERCEPT devuelve el error #N/D!.
- Si la varianza de x_conocidos es cero, INTERCEP devuelve el error #DIV/0.

KURT(n1; n2; ...)

KURT devuelve una estimación imparcial de la kurtosis de un conjunto de datos. (E)
Note que esto sólo tiene significado pleno si la distribución subyacente tiene realmente un cuarto momento. La kurtosis se desplaza por tres así que una distribución normal tendrá kurtosis cero.

- Las cadenas y las celdas vacías, simplemente se ignoran.
- Si se dan menos de cuatro números o todos son iguales, KURT devuelve el error #DIV/0!.

KURTP(n1; n2; ...)

KURTP devuelve la población kurtosis de un conjunto de datos.

- Las cadenas y las celdas vacías simplemente se ignoran.
- Si se dan menos de dos números o todos son iguales, KURTP devuelve el error #DIV/0!.

LANDAU(x)

LANDAU devuelve la densidad de probabilidad p(x) en x para la distribución de Landau usando un método de aproximación.

LAPLACE(x;a)

LAPLACE devuelve la densidad de probabilidad p(x) en x para la distribución de Laplace con media a.

LARGE(n1;n2;...;k)

Devuelve el k-ésimo mayor valor en un conjunto de datos. (E)

- Si el conjunto de datos está vacío, devuelve el error #NÚM!.
- Si k <= 0 o k es mayor que el número de términos, LARGE devuelve el error #NÚM!.

LINEST(y_conocido[;x_conocido[;const[;estad]]])

La función LINEST calcula la línea de mínimos cuadrados que mejor se ajuste a sus datos de y_conocido. x_conocido contiene las x correspondientes donde y=mx+b.

LINEST devuelve un array con dos columnas y una fila. La pendiente (m) de la línea de regresión y=mx+b se da en la primera columna y la y intercepta a (b) en la segunda.

Si estad es VERDADERO, se devolverá información estadística extra. Debajo de la los coeficientes de la línea de regresión se escribe información estadística extra en el array resultado. La información estadística extra consiste en cuatro filas de datos. En la primera fila se representan los valores estándar de error para los coeficientes m1, (m2, ...), b. La segunda fila contiene el cuadrado de R y el error estándar para la y estimada. La tercera fila contiene el valor observado F y los grados de libertad. La última fila contiene la suma de regresión de los cuadrados y la suma residual de los cuadrados.

- Si se omite x_conocido, se usa un array {1, 2, 3, ...}.
- Si y_conocido y x_conocido tienen distinto número de puntos de datos, LINEST devuelve un error #NÚM!.
- Si const es FALSO, se forzará que la línea pase por el origen, por ejemplo, b será cero. De forma predeterminada, es VERDADERO.
- De forma predeterminada, estad es FALSO.

LOGEST(y_conocido[;x_conocido;const;estad])

La función LOGEST aplica el método de mínimos cuadrados para ajustar una curva exponencial de la forma:

$$y = b \cdot m_1^{x_1} \cdot m_2^{x_2} \ldots$$

a sus datos.

Si estad es VERDADERO, se devuelve información estadística adicional. La información estadística adicional se escribe bajo los coeficientes de regresión lineal en el array de resultados. La información estadística adicional consiste en cuatro líneas de datos. En la primera línea se representan los valores de error estándar para los coeficientes m_1, (m_2, ...), b. La segunda fila contiene el cuadrado de R y el error estándar para la estimación de y. La tercera línea el valor observado por F y los grados de libertad. La última fila contiene la suma de los cuadrados de la regresión y la suma de los cuadrados residual.

- Si y_conocida se omite, se usa un array {1, 2, 3, ...}. LOGEST devuelve un array { m_n, m_{n-1}, ..., m_1, b }.
- Si alguna y_conocida y x_conocida tienen distinto número de puntos de datos, LOGEST devuelve el error #NÚM!
- Si const es FALSO, la línea será forzada a ir a través de (0,1), o sea, b será uno. Por omisión es VERDADERO.
- El valor predeterminado de estad es FALSO.

LOGFIT(known_y's;known_x's)

La función LOGFIT aplica el método de mínimos cuadrados para ajustar la ecuación logarítmica

y = a + b * ln(sign * (x - c)) , sign = +1 o -1

a sus datos. La gráfica de la ecuación es una curva logarítmica desplazada horizontalmente por c y posiblemente reflejada a través del eje y (si sign = -1).

LOGFIT devuelve un array conteniendo cinco columnas y una fila. 'Signo' se da en la primera columna, 'a', 'b', y 'c' se dan el las columnas 2 a 4. La columna 5 contiene la suma de los residuos al cuadrado.

Se devuelve un error cuando hay menos de 3 x's o y's diferentes, o cuando la forma de la nube de puntos es demasiado diferente de una forma logarítmica.

Puede usar la fórmula
= a + b * ln(sign * (x - c))
o reordenarla a
= (exp((y - a) / b)) / sign + c
para calcular y's o x's desconocidas respectivamente.

Técnicamente, esto es un ajuste no linear mediante prueba y error.. La precisión de c es: el ancho de x-rango ->redondeado al siguiente más pequeño (10ˆenteros), 0.000001 veces. Quizá haya casos en los que el ajuste devuelto no sea el mejor posible.

LOGINV(p;media;desvest)

La función LOGINV devuelve la inversa de la distribución acumulativa lognormal. p es la probabilidad dada correspondiente a la distribución normal. media es la media aritmética de la distribución y desvest es la desviación estándar de la distribución. (E)

- Si p <0 ó p >1 ó desvest <= 0, LOGINV devuelve el error #NÚM!.

LOGISTIC(x;a)

LOGISTIC devuelve la densidad de probabilidad p(x) en x para una distribución logística con parámetro de escala a

LOGNORMDIST(x;media;desvest)

La función LOGNORMDIST devuelve la distribución lognormal. x es el valor para el que quiere la distribución, media es la media de la distribución y desvest es la desviación estándar de la distribución. (E)

- Si desvest = 0, LOGNORMDIST devuelve el error #DIV/0!.
- Si x <= 0, significado <0 ó desvest <0, LOGNORMDIST devuelve el error #NÚM!.

LOGREG(y_conocido[;x_conocido[;const[;estad]]])

La función LOGREG transforma las x's a $n = \ln(x)$ y aplica el método de los mínimos cuadrados para ajustar a la ecuación lineal
$y = m \cdot z + b$
a sus y's y z's —equivalente a ajustar la ecuación
$y = m \cdot \ln(x) + b$
a las y's y a las x's

Si se omite x_conocido, se usa un array {1, 2, 3, ...}. LOGREG devuelve un array conteniendo dos columnas y una fila. m se da en la primera columna y b en la segunda.

Si y_conocido y x_conocido tienen distintos números de puntos de datos, LOGREG devuelve el error #NÚM!.

Si const es FALSO, la curva se forzará para que pase por [0,1], por ejemplo, b sería cero. De forma predeterminada, es VERDADERO.

Si estad es VERDADERO, se devolverá información estadística extra. La información estadística extra se aplica al estado ANTES de la transformación a z.. La información estadística extra consiste en cuatro filas de datos. En la primera fila, los valores de error estándar para los coeficientes m, b, son representados La segunda fila contiene el cuadrado de R y el error estándar para el y estimado. La tercera fila contiene el valor observado de F y los grados de libertad. La última fila contiene la suma de regresión de los cuadrados y la suma residual de los cuadrados.

El valor predeterminado de estad es FALSO.

MAX(b1;b2;...)

Devuelve el valor del elemento de los valores pasados que es mayor a todos los demás. Los valores negativos se consideran menores que los números positivos. (E)

MAXA(número1;número2;...)

MAXA devuelve el mayor valor de los argumentos dados. También se incluyen en el cálculo números, texto y valores lógicos. Si la celda contiene texto o el argumento se evalúa a FALSO, se cuenta como valor cero (0). Si el argumento se evalúa a VERDADERO, se cuenta como uno (1). Note que las celdas vacías no se cuentan. (E)

MEDIAN(n1;n2;...)

Devuelve la mediana del conjunto de datos. (E)

- Las cadenas y celdas vacías son ignoradas.
- Si el número de datos es par, MEDIAN devuelve el media de los dos números del medio.

MIN(b1;b2;...)

Devuelve el valor del elemento pasado en argumento cuyo valor es menor a todos los demás. Los valores negativos se consideran menores a los positivos. (E)

MINA(número1;número2;...)

MINA devuelve el menor valor de los argumentos dados. También se incluyen en el cálculo números, texto y valores lógicos. Si la celda contiene texto o el argumento se evalúa a FALSO, se cuenta como valor cero (0). Si el argumento se evalúa a VERDADERO, se cuenta como uno (1). Note que las celdas vacías no se cuentan. (E)

MODE(n1; n2; ...)

MODE devuelve el número más común del conjunto de datos. Si el conjunto de datos tiene muchos números comunes, MODE devuelve el primero de ellos. (E)

- Las cadenas y las celdas vacías, simplemente se ignoran.
- Si el conjunto de datos no contiene ningún duplicado, MODE devuelve el error #N/D.

NEGBINOMDIST(f;t;p)

La función NEGBINOMDIST devuelve la distribución binomial negativa. f es el número de fallos, t es el número umbral de éxitos y p es la probabilidad de éxito. (E)

- Si f ó t no son enteros, se truncan.
- Si (f + t -1) <= 0, NEGBINOMDIST devuelve el error #NÚM!.
- Si p <0 ó p >1, NEGBINOMDIST devuelve el error #NÚM!.

NORMDIST(x;media;desvest;acumulado)

La función NORMDIST devuelve la distribución normal acumulada. x es el valor para el que se va a obtener la distribución, media es la media de la distribución, desvest es la desviación estándar. (E)

- Si desvest es 0, NORMDIST devuelve el error #DIV/0!.

NORMINV(p;media;desvest)

La función NORMINV devuelve la inversa de la distribución normal acumulada. p es la probabilidad dada correspondiente a la distribución normal, media es la media aritmética de la distribución y desvest es la desviación estándar de la distribución. (E)

- Si p <0 ó p >1 ó desvest <= 0, NORMINV devuelve el error #NÚM!.

NORMSDIST(x)

La función NORMSDIST devuelve la distribución normal estándar acumulativa. x es el valor para el que se va a obtener la distribución. (E)

NORMSINV(p)

La función NORMSINV devuelve el inverso de la distribución acumulativa normal estándar. p es la probabilidad dada correspondiente a la distribución normal. (E)

- Si p <0 ó p >1, NORMSINV devuelve el error #NÚM!.

PARETO(x;a;b)

PARETO devuelve la densidad de probabilidad p(x) en x para una distribución Pareto con exponente a y escala b.

PEARSON(array1;array2)

Devuelve el coeficiente de correlación de Pearson para dos conjuntos de datos. (E)

- Se ignoran las cadenas de texto y las celdas vacías.

PERCENTILE(array;k)

La función PERCENTILE devuelve el 100*k-ésimo percentil de los puntos de datos dados (esto es, un número x tal que una fracción k de los puntos de datos es menor que x.). (E)

- Si array está vacío, PERCENTILE devuelve el error #NÚM!.
- Si k <0 o k >1, PERCENTILE devuelve el error #NÚM!.

PERCENTRANK(array;x[;significatividad])

La función PERCENTRANK devuelve el rango de un punto de datos en un conjunto de datos. array es el rango de valores numéricos, x es el punto de datos que quiere clasificar, y la significatividad opcional identifica el número de dígitos con significatividad para el valor devuelto, truncando lo restante. Si se omite significatividad, PERCENTRANK usa tres dígitos.

- Si array no contiene puntos, PERCENTRANK devuelve el error #NÚM!.
- Si significatividad <1, PERCENTRANK devuelve el error #NÚM!.

- Si x supera el valor más grande o es menor que el más pequeño del array, PERCENTRANK devuelve el error #NÚM!.
- Si x no coincide con ninguno de los valores de array ó x coincide con más de uno, PERCENTRANK interpola el valor devuelto.

PERMUT(n;k)

La función PERMUT devuelve el número de permutaciones. n es el número de objetos, k es el número de objetos en cada permutación. (E)

- Si n = 0, PERMUT devuelve el error #NÚM!.
- Si n <k, PERMUT devuelve el error #NÚM!.

POISSON(x;media;acumulativo)

Devuelve la distribución de Poisson, x es la cantidad de eventos, media es el valor numérico esperado, acumulativo describe si se debe devolver la suma de la función de Poisson de 0 a x. (E)

- Si x no es un entero será truncado.
- Si x <= 0 POISSON devuelve el error #NÚM!
- Si media <= 0 POISSON devuelve el error #NÚM!

PROB(rango_x;rango_prob;límite_inf[;límite_sup])

La función PROB devuelve la probabilidad de que los valores de un rango o un array estén entre dos límites. Si límite_sup no se da, PROB devuelve la probabilidad de que los valores de rango_x sean iguales a límite_inf. (E)

- Si la suma de las probabilidades de rango_prob no es igual a 1, PROB devuelve el error #NÚM!.
- Si cualquier valor de rango_prob es <= 0 ó >1, PROB devuelve el error #NÚM!.
- Si rango_x y rango_prob contienen un número diferente de entradas de datos, PROB devuelve el error #N/D.

QUARTILE(array;cuart)

La función QUARTILE devuelve el cuartil de los puntos de datos dados. (E)
Si cuart es igual a: QUARTILE devuelve:
0 el valor más pequeño del array.
1 el primer cuartil.
2 el segundo cuartil.
3 el tercer cuartil.
4 el valor más grande del array.

- Si array está vacío, QUARTILE devuelve el error #NÚM!.
- Si cuart <0 ó quart >4, QUARTILE devuelve el error #NÚM!.
- Si cuart no es un entero, se trunca.

RANK(x;ref[;orden])

RANK devuelve la clasificación de un número en una lista de números. x es el número cuya clasificación quiere encontrar,ref es la lista de números y orden especifica cómo clasificar los números. Si orden es 0, los números se clasifican en orden descendente. En otro caso, se clasifican en orden ascendente. (E)

RAYLEIGH(x;sigma)

RAYLEIGH devuelve la densidad de probabilidad p(x) en x para una distribución de Rayleigh con parámetro de escala sigma.

RAYLEIGHTAIL(x;a;sigma)

RAYLEIGHTAIL devuelve la densidad de probabilidad p(x) en x para una distribución en cola Rayleigh con parámetro de escala sigma y límite inferior a.

RSQ(array1;array2)

RSQ devuelve el cuadrado del coeficiente de correlación de Pearson de dos conjuntos de datos. (E)

- Las cadenas y las celdas vacías simplemente se ignoran.

SKEW(n1; n2; ...)

SKEW devuelve una estimación de la desviación de una distribución. (E)
Note que esto sólo tiene significado si la distribución subyacente realmente tiene un tercer momento. La desviación de una distribución simétrica (ej., normal) es cero.

- Las cadenas y las celdas vacías se ignoran.
- Si se dan menos de tres números, SKEW devuelve un error #DIV/0!

SKEWP(n1; n2; ...)

SKEWP devuelve la desviación de la población de un conjunto de datos.

- Las cadenas y las celdas vacías se ignoran.
- Si se dan menos de dos números, SKEWP devuelve un error #DIV/0!

SLOPE(y_conocidos;x_conocidos)

Devuelve la pendiente de la línea de regresión linear. (E)

SMALL(n1;n2;...;k)

Devuelve el k-ésimo menor valor en un conjunto de datos. (E)

- Si el conjunto de datos está vacío, entonces devuelve el error #NÚM!.
- Si k <= 0 o k es mayor que el número de términos, SMALL devuelve el error #NÚM!.

SSMEDIAN(array[;intervalo]))

La función SSMEDIAN devuelve la media para datos agrupados como los que se determinan comúnmente en las ciencias sociales. Los puntos de datos dados en array se asume que son el resultado de la agrupación de datos en intervalos de longitud intervalo.

- Si no se da intervalo, SSMEDIAN usa 1.
- Si array está vacío, SSMEDIAN devuelve el error #NÚM!.
- Si intervalo <= 0, SSMEDIAN devuelve el error #NÚM!.
- ssmedian no comprueba si los puntos de datos son al menos un intervalo a parte.

STANDARDIZE(x;media;desvest)

La función STANDARDIZE devuelve un valor normalizado. x es el número a ser normalizado, media es la media de la distribución, desvest es la desviación estándar de la distribución. (E)

- Si desvest es 0, STANDARDIZE devuelve el error #DIV/0!.

STDEV(b1; b2; ...)

STDEV calcula la desviación estándar de un conjunto de números tratando estos números como miembros de una población. (E)

STDEVA(número1;número2;...)

STDEVA devuelve la desviación estándar basada en una muestra. También se incluyen en el cálculo números, texto y valores lógicos. Si la celda contiene texto o el argumento se evalúa a FALSO, se cuenta como valor cero (0). Si el argumento se evalúa a VERDADERO, se cuenta como uno (1). Note que las celdas vacías no se cuentan. (E)

STDEVP(b1; b2; ...)

STDEVP devuelve la desviación estándar de un conjunto de números tratando esos números como miembros de una población completa. (E)

STDEVPA(número1;número2;...)

STDEVPA devuelve la desviación estándar basada en una la población entera. También se incluyen en el cálculo números, texto y valores lógicos. Si la celda contiene texto o el argumento se evalúa a FALSO, se cuenta como valor cero (0). Si el argumento se evalúa a VERDADERO, se cuenta como uno (1). Note que las celdas vacías no se cuentan. (E)

STEYX(y_conocido;x_conocido)

La función STEYX devuelve el error estándar de los valores de y previstos por cada x en la regresión. (E)

- Si y_conocido y x_conocido están vacíos o tienen diferente número de argumentos, STEYX devuelve el error #N/D.

SUMIF(función_num;ref1;ref2;...)

La función SUBTOTAL devuelve el subtotal de la lista de argumentos. función_num es el número que especifica la función a usar para calcular el subtotal. (E)

Las siguientes funciones están disponibles:

1. AVERAGE
2. COUNT
3. COUNTA
4. MAX
5. MIN
6. PRODUCT
7. STDEV
8. STDEVP
9. SUM
10. VAR
11. VARP

TDIST(x;gol;colas)

La función TDIST devuelve la distribución t de Student. gdl es el grado de libertad y colas es 1 ó 2 dependiendo de si quiere una distribución de 1 ó 2 colas. (E)

- Si gdl <1, TDIST devuelve el error #NÚM!.
- Si colas no es ni 1 ni 2, TDIST devuelve el error #NÚM!.

TINV(p;gdl)

La función TINV devuelve la inversa de la distribución t de Student de dos colas. (E)

- Si p <0 ó p >1 ó gdl <1, TINV devuelve el error #NÚM!.

TREND(y_conocido[;x_conocido[;x_nuevo[;const]]])

La función TREND estima los valores futuros de un conjunto de datos dado usando la línea de mínimos cuadrados que mejor se ajuste a sus datos. y_conocido son los valores de y donde y=mx+b, y x_conocido contiene los valores correspondientes de x. x_nuevo contiene los valores de x para los que quiere estimar los valores de y.

- Si se omite x_conocido, se usa un array {1, 2, 3, ...}.
- Si se omite x_nuevo, se asume que es el mismo que x_conocido.
- Si y_conocido y x_conocido tienen distinto número de puntos de datos, TREND devuelve el error #NÚM!.

TRIMMEAN(ref;fracción)

TRIMMEAN devuelve la media del interior de un conjunto de datos. ref es la lista de los números cuya media se quiere calcular y fracción es la fracción del conjunto de datos excluido de la media. Por ejemplo, si fracción=0.2 y el conjunto de datos contiene 40 números, se recortan 8 números del conjunto de datos (40 x 0.2), 4 de arriba y 4 de abajo del conjunto. (E)

TTEST(array1;array2;colas;tipo)

La función TTEST devuelve la probabilidad de una prueba t de Student. (E)
array1 es el primer conjunto de datos y array2 es el segundo conjunto de datos. Si colas es uno, TTEST usa la dissVaribución de una cola y si colas es dos, TTEST usa la distribución de dos colas. tipo determina la clase de la prueba:

1. Test de paridad
2. Vcrianza de igualdad de dos muestras
3. Varianza de desigualdad de dos muestras

- Si los conjuntos de datos contienen un número diferente de puntos de datos y la prueba es de paridad (tipo uno), TTEST devuelve el error #N/A.
- Si colas no es ni uno ni dos, TTEST devuelve el error #NÚM!.
- Si tipo es distinto de uno, dos o tres, TTEST devuelve el error #NÚM!.

VAR(b1; b2; ..,)

VAR estima la varianza de una muestra de población. Para obtener la varianza real de una población completa, use VARP. (E)

- VAR también se conoce como varianza-N-1. En condiciones razonables, es el estimador de bondad máximo de la verdadera varianza.

VARA(número1;número2;...)

VARA devuelve la varianza basada en una muestra. También se incluyen en el cálculo números, texto y valores lógicos. Si la celda contiene texto o el argumento se evalúa a FALSO, se cuenta como valor cero (0). Si el argumento se evalúa a VERDADERO, se cuenta como uno (1). Note que las celdas vacías no se cuentan. (E)

VARP(b1; b2; ...)

VARP calcula la varianza de un conjunto de números, cada número es el miembro de una población y el conjunto es la población entera.

- VARP también se conoce como N-varianza.

VARPA(número1;número2;...)

VARPA devuelve la varianza basada en la población entera. También se incluyen en el cálculo números, texto y valores lógicos. Si la celda contiene texto o el argumento se evalúa a FALSO, se cuenta como valor cero (0). Si el argumento se evalúa a VERDADERO, se cuenta como uno (1). Note que las celdas vacías no se cuentan. (E)

WEIBULL(x;alfa;beta;acumulado)

La función WEIBULL devuelve la distribución Weibull. (E)
Si el booleano acumulado es cierto, devolverá:

$$1 - e^{-\left(\frac{x}{\beta}\right)^{\alpha}}$$

en otro caso, devolverá

$$\frac{\alpha}{\beta^{\alpha}} \cdot x \cdot (\alpha - 1) \cdot e^{-\frac{x}{\beta^{\alpha}}}$$

- Si x <0, WEIBULL devuelve el error #NÚM!.
- Si alfa <= 0 ó beta <= 0, WEIBULL devuelve el error #NÚM!.

ZTEST(ref;x)

ZTEST devuelve la probabilidad de dos colas de una prueba Z. (E)
ref es el conjunto de datos y x es el valor a ser probado.

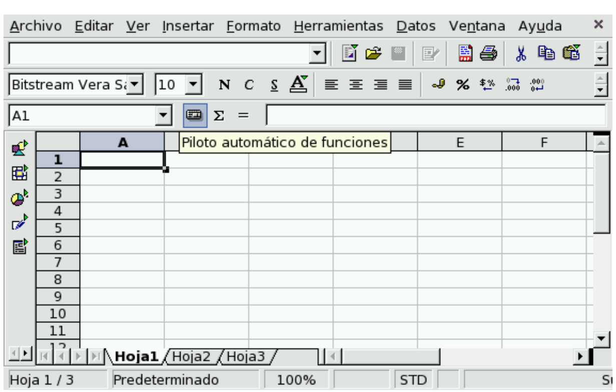
- Si ref contiene menos de dos elementos de datos, ZTEST devuelve el error #DIV/0!.

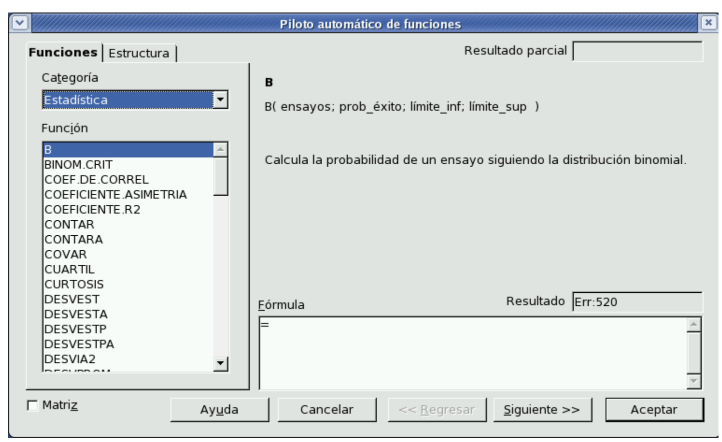
C. Funciones Estadísticas de Calc[37]

Aparecen las funciones disponibles en la categoría **Estadística** junto con un ejemplo. Se trata de las siguientes funciones:

BDISTR.BETA
BINOM.CRIT
COEF.DE.CORREL
COEFICIENTE.ASIMETRIA
COEFICIENTE.R2
CONTAR, CONTARA
COVAR
CUARTIL
CURTOSIS
DESVEST
DESVESTA
DESVESTP
DESVESTPA
DESVIA2
DESVPROM
DISTR.BETA.INV
DISTR.BINOM
DISTR.EXP
DISTR.F
DISTR.F.INV
DISTR.GAMMA
DISTR.GAMMA.INV
DISTR.HIPERGEOM
DISTR.JI
DISTR.LOG.NORM
DISTR.NORM
DISTR.NORM.ESTAND
DISTR.NORM.ESTAND.INV
DISTR.NORM.INV
DISTR.T
DISTR.T.INV
DIST.WEIBULL
ERROR.TÍPICO.XY
FISHER
GAMMA.LN
GAUSS
INTERSECCIÓN.EJE
INTERVALO.CONFIANZA
INV.LOG
JERARQUÍA
K.ESIMO.MAYOR
K.ESIMO.MENOR
MÁX
MÁXA
MEDIA.ACOTADA
MEDIA.ARMO
MEDIA.GEOM
MEDIANA
MÍN
MÍNA
MODA
NEGBINOMDIST
NORMALIZACI\'{O}N
PEARSON
PENDIENTE
PERCENTIL
PERMUTACIONES
PERMUTACIONESA
PHI
POISSON
PROBABILIDAD
PROMEDIO
PROMEDIOA
PRONÓSTICO
PRUEBA.F
PRUEBA.FISHER.INV
PRUEBA.JI
PRUEBA.JI.INV
PRUEBA.T
PRUEBA.Z
RANGO.PERCENTIL
VAR
VARA
VARP
VARPA

Cómo acceder a esta función...
AutoPiloto de funciones - Categoría **Estadística**





Algunos de los ejemplos emplean la siguiente tabla de datos:

[37] Tomado de la ayuda de Calc

	C	D
2	Valor x	Valor y
3	-5	-3
4	-2	0
5	-1	1
6	0	3
7	2	4
8	4	6
9	6	8

Las funciones estadísticas se describen en las subsecciones siguientes.

Funciones estadísticas, primera parte

INTERSECCIÓN.EJE

En este campo se puede definir el punto de corte de la línea de regresión con el eje Y.

Sintaxis INTERSECCIÓN.EJE(Datos_Y; Datos_X)
Datos_Y es el grupo de las lecturas o datos subordinados.
Datos_X es el grupo de las lecturas o datos subordinados.
Se deben utilizar nombres, matrices o referencias que contengan números. También se pueden escribir números directamente.

Ejemplo Para calcular el eje de intersección se utilizan como valor Y las celdas D3:D9 y como valor X, las celdas C3:C9 de la hoja de ejemplo. La entrada queda como sigue:
Intersección.eje(D3:D9;C3:C9) = 2,15. El resultado es entonces 2,15.

- Vea también las siguientes funciones:

  COEFICIENTE.R2, **PRONÓSTICO**, **CRECIMIENTO**, **ESTIMACIÓN.LINEAL**, **ESTIMACIÓN.LOGARÍTMICA**, **PEARSON**, **PENDIENTE**, **ERROR.TÍPICO.XY**, **TENDENCIA**.

CONTAR

Se puede calcular la cantidad de números que contiene la lista de argumentos. Al definirse la cantidad no se tienen en cuenta las entradas de texto. Los valores individuales se añaden.

Sintaxis CONTAR(valor 1; Valor 2; ...)
valor 1; valor 2,... son valores a partir de los cuales se calcula la cantidad de argumentos.

Ejemplo Se deben contar las entradas **2,4,6** y **ocho** del cuadro de texto **valor 1-4**.
CONTAR(2;4;6;ocho) = 3. La cantidad de entradas es **3**.

- Vea también las siguientes funciones:

  PROMEDIO, **CONTARA**, **BDCONTAR**, **BDCONTARA**, **SUMA**

CONTARA

Calcula la cantidad de números que contiene una lista de argumentos. Al definirse la cantidad se tienen en cuenta las entradas de texto. Los valores individuales se añaden. Los argumentos vacíos no se tienen en cuenta.

Sintaxis CONTARA(valor 1; Valor 2; ...)
valor 1; valor 2,... son valores a partir de los cuales se calcula la cantidad de argumentos.

Ejemplo Se cuentan las entradas **2,4,6** y **ocho** del cuadro de texto **valor1-4**.
Contara(2;4;6;ocho) = 4. La cantidad de entradas es **4**.

- Vea también las siguientes funciones:
  CONTAR, **DBCONTAR**, **DBCONTARA**, **PROMEDIO**, **PRODUCTO**, **SUMA**

B

Se calcula la probabilidad de los resultados de prueba con distribuciones por binomios.

Sintaxis B(ensayos; prob_éxito; límite_inf; límite_sup)
ensayos define la cantidad de intentos.
prob_éxito determina la probabilidad individual del resultado del intento.
límite_inf define el límite inferior de la cantidad de intentos.
S2 (opcional) define el límite superior de la cantidad de intentos.

Ejemplo ¿Cuál debe ser la probabilidad si al tirar un dado 10 veces sale dos veces el seis? La probabilidad para un seis (o para cualquier otro número) es 1/6, luego el resultado es la siguiente fórmula:
=B(10; 1/6; 2) el resultado es un 29 % de probabilidad.

- Vea también las siguientes funciones:
  DISTR.BINOM

COEFICIENTE.R2

Para definir el cuadrado del coeficiente de correlación de Pearson, hay que introducir los valores correspondientes en el cuadro de texto. El coeficiente R2 es una medida para obtener un buen ajuste, que puede obtener una regresión, y también se conoce como Coeficiente de determinación.

Sintaxis COEFICIENTE.R2(**Datos_Y; Datos_X**)
Datos_Y. Punto de datos de una matriz o área.
Datos_X. Punto de datos de una matriz o área.

Ejemplo =COEFICIENTE.R2(A1:A20; B1:B20) calcula el coeficiente de determinación para los dos registros de datos de las columnas A y B.

- Vea también las siguientes funciones:
  COEF.DE.CORREL, **COVAR**, **INTERSECCIÓN.EJE**, **ESTIMACIÓN.LINEAL**, **ESTIMACIÓN.LOGARÍTMICA**, **PEARSON**, **PENDIENTE**, **ERROR.TÍPICO.XY**, **TENDENCIA**.

DISTR.BETA.INV

Proporciona los valores de una variable aleatoria con distribuciones beta invertidas.

Sintaxis DISTR.BETA.INV(probabilidad; alfa; beta; comienzo; fin)
Probabilidad es el valor con el cual se debe calcular la función a través del intervalo **comienzo** hasta el **fin**.
Alfa es un parámetro de distribución.
Beta es un parámetro de distribución.
Comienzo (opcional) es el límite inferior de **número**.
Fin (opcional) es el límite superior de **número**.

Ejemplo =DISTR.BETA.INV(0,5; 5; 10) da el valor **0,33**.

- Vea también las siguientes funciones:
  DISTR.BETA

DISTR.BETA

Calcula la distribución del intervalo de probabilidad de una variable aleatoria con distribución beta.

Sintaxis DISTR.BETA(x; alfa; beta; comienzo; fin)
x es el valor con el cual se debe calcular la función a través del intervalo **principio** hasta el **fin**.
Alfa es un parámetro de distribución.
Beta es un parámetro de distribución.
Comienzo (opcional) es el límite inferior de **x**.
Fin (opcional) es el límite superior de **número**.

Ejemplo =DISTR.BETA(0,75; 3; 4) da el valor **0,96**.

- Vea también las siguientes funciones:
  DISTR.BETA.INV

DISTR.BINOM

Calcula el intervalo de probabilidad a partir de una variable aleatoria con distribución binomial.

Sintaxis DISTR.BINOM(X; ensayos; prob_éxito; acumulado)
X es la cantidad de éxitos de una sucesión de intentos.
ensayos es el total de intentos.
prob_éxito es el intervalo de probabilidad de éxito de un intento.
acumulado = 0 calcula la probabilidad individual, **acumulado** = 1 la probabilidad acumulada.

Ejemplo =DISTR.BINOM(A1; 12; 0,5; 0) muestra los intervalos de probabilidad cuando se introducen para A1 valores comprendidos entre el 0 y el 12, que resultan de tirar 12 veces una moneda en la cantidad A1 llamada **encabezamiento**.
=DISTR.BINOM(A1; 12; 0,5; 1) muestra los intervalos de probabilidad acumulados para la misma serie, esto es, por ejemplo, para A1 = 4 el intervalo de probabilidad para 0, 1, 2, 3 o 4 veces **encabezamiento** (el O no es excluyente).

- Vea también las siguientes funciones:
  B, **FACT**, **DISTR.HIPERGEOM**, **COMBINAR**, **BINOM.CRIT**, **NEGBINOMDIST**, **PERMUTACIONES**, **PROBABILIDAD**

PRUEBA.JI.INV

Calcula, para el intervalo de probabilidad de error, el valor (teórico) de la distribución del cuadrado de JI que no debe superar la distribución observada, para que la hipótesis que se va a examinar sea verdadera.

Sintaxis PRUEBA.JI.INV(probabilidad; grados_libertad)

Probabilidad es el valor del intervalo de probabilidad de error para el que se debe calcular el tamaño PRUEBA.JI.IN crítico, es decir, el intervalo de probabilidad con el que se cumple la hipótesis.

Grados libertad es la cantidad de grados libertad del experimento.

Ejemplo Se tira un dado 1020 veces. Los números de las caras del 1 al 6 aparecen 195, 151, 148, 189, 183 y 154 veces (valores observados). Se debe verificar la hipótesis de si el dado es real.

La distribución del cuadrado de ji de la muestra se calcula con la fórmula anterior. Como el valor previsto para cada uno de los números de las caras en n dados n veces es 1/6, entonces 1020/6 = 170, la fórmula da un valor de cuadrado de ji de 13,27.

Si el cuadrado de ji (observado) es mayor o igual al cuadrado PRUEBA.JI.INV (teórico), entonces se descarta la hipótesis, pues la desviación entre teoría y práctica es demasiado grande. Si el cuadrado ji observado es inferior a PRUEBA.JI.INV, entonces la hipótesis cumple el intervalo de probabilidad de error dado.

=PRUEBA.JI.INV(0,05; 5) da 11,07.

=PRUEBA.JI.INV(0,02; 5) da 13,39.

Con un intervalo de probabilidad de error del 5 % el dado no es de verdad; si el intervalo de error es del 2 % no hay razón para cuestionar su veracidad.

- Vea también las siguientes funciones:

  PRUEBA.JI, **DISTR.JI**

PRUEBA.JI

Con ayuda de las pruebas del cuadrado de Ji se obtiene, directamente a partir de los datos de medida, el valor del intervalo de probabilidad, que constituye una hipótesis. De este modo se comparan los tamaños observados y los esperados de una muestra: PRUEBA.JI compara las dos series de datos y calcula el valor del cuadrado de Ji. A partir de este valor del cuadrado de ji se determina el intervalo de probabilidad de error de la hipótesis que se va a demostrar.

El intervalo de probabilidad calculado mediante PRUEBA.JI también se puede determinar mediante DISTR.JI; en este caso en lugar de una serie de datos, el cuadrado de ji de la muestra se debe presentar como parámetro.

Sintaxis PRUEBA.JI(datos B; datos E)

datos_B es la matriz de las observaciones.

Datos E es la matriz de los valores previstos.

Ejemplo

	A	B
1	195	170
2	151	170
3	148	170
4	189	170
5	183	170
6	154	170

=PRUEBA.JI(A1:A6; B1:B6) da 0,02. Es el intervalo de probabilidad con el que se cumple la distribución teórica del cuadrado de ji.

- Vea también las siguientes funciones:

  PRUEBA.JI.INV, **DISTR.JI**

DISTR.JI

A partir del cuadrado de ji determinado se obtiene el intervalo de probabilidad que constituye una hipótesis. DISTR.JI compara el valor del cuadrado de ji de una muestra, que se calcula a partir de la suma de (valor observado-valor previsto)ˆ2/valor previsto en todos los valores, con la distribución teórica del cuadrado de ji y origina el intervalo de probabilidad de error de la hipótesis que se debe demostrar.

El intervalo de probabilidad calculado mediante DISTR.JI también se puede determinar mediante PRUEBA.JI; en este caso, en lugar del cuadrado de ji de la muestra, los datos observados y previstos se deben suministrar como parámetros.

Sintaxis DISTR.JI (X; grados_libertad)

x es el valor del cuadrado de ji de la muestra sobre el que se debe calcular el intervalo de probabilidad de error.

Grados libertad es la cantidad de grados libertad del experimento.

Ejemplo =DISTR.JI(13,27; 5) da como resultado 0,02.

Si el valor del cuadrado de ji de la muestra asciende a 13,27 y el experimento tiene 5 grados libertad, entonces la hipótesis se cumple con un intervalo de probabilidad de error del 2 %.

- Vea también las siguientes funciones:

  PRUEBA.JI.INV, **PRUEBA.JI**

DISTR.EXP

Calcula el intervalo de probabilidad de variables aleatorias de distribución exponencial.

Sintaxis DISTR.EXP(x; lambda; acum)

x es el valor sobre el que se calcula la distribución exponencial.

Lambda es el parámetro de la distribución exponencial.

acum acumulado = 0 calcula la función de densidad, **acum** = 1 la distribución.

Ejemplo =DISTR.EXP(3; 0,5; 1) da 0,78.

- Vea también las siguientes funciones:

  DISTR.GAMMA, **POISSON**

Funciones estadísticas, segunda parte

DISTR.F.INV

Calcula el percentil de la distribución F. La distribución F se utiliza en pruebas F y dispersa un par de grupos de datos para definir su comportamiento.

Sintaxis DISTR.F.INV(probabilidad; grados_libertad1; grados_libertad2)

Probabilidad es el valor de probabilidad para el que se calcula la distribución F inversa.

Grados_libertad1 es la cantidad del grado de libertad del numerador de la distribución F.

Grados_libertad2 es la cantidad del grado de libertad del denominador de la distribución F.

Ejemplo =DISTR.F.INV(0,5; 5; 10) resulta 0,93.

- Vea también las siguientes funciones:

  PRUEBA.F, **DISTR.F**

FISHER

Calcula la transformación Fisher para x y da como resultado una función que se distribuye casi de forma normal y tiene un cociente cero, aproximadamente.

Sintaxis FISHER(número)
Número es el valor que se debe transformar.

Ejemplo =FISHER(0,5) resulta 0,55.

- Vea también las siguientes funciones:
  PRUEBA.FISHER.INV, **COEF.DE.CORREL**, **COVAR**

PRUEBA.FISHER.INV

Calcula la transformación Fisher para X invertida y da como resultado una función que se distribuye de forma casi normal y tiene un coeficiente cero, aproximadamente.

Sintaxis PRUEBA.FISHER.INV(número)
Número es el valor que se debe reconvertir.

Ejemplo

=PRUEBA.FISHER.INV(0,5) resulta 0,46.

- Vea también las siguientes funciones:
  FISHER, **COEF.DE.CORREL**, **COVAR**

PRUEBA.F

Realiza una prueba de variable F y calcula la estadística.

Sintaxis PRUEBA.F(datos 1; datos 2)
Datos 1 es la matriz del primer registro de datos.
Datos 2 es la matriz del segundo registro de datos.

Ejemplo =PRUEBA.F(A1:A30; B1:B12) calcula si ambas series de datos difieren en la varianza y, como resultado, genera la probabilidad de que ambas filas se originen a partir de la misma causa.

- Vea también las siguientes funciones:
  DISTR.F.INV, **DISTR.F**

DISTR.F

Calcula el valor de la función de distribución F.

Sintaxis DISTR.F(x; grados_libertad 1; grados_libertad 2)
x es el valor para el que se debe calcular la distribución F.
Grados_libertad 1 es el grado de libertad del numerador de la distribución F.
Grados_libertad 2 es el grado de libertad del denominador de la distribución F.

Ejemplo =DISTR.F(0,8; 8; 12) resulta 0,61.

- Vea también las siguientes funciones:
  DISTR.F.INV, **PRUEBA.F**

DISTR.GAMMA.INV

Calcula el percentil de la distribución gamma inversa. Con esta función se pueden buscar variables con una distribución que puede no ser homogénea.

Sintaxis DISTR.GAMMA.INV(probabilidad; alfa; beta)
Probabilidad es el valor de probabilidad para el que se calcula la distribución gamma inversa.
Alfa es el parámetro alfa de la distribución gamma.
Beta es el parámetro beta de la distribución gamma.

Ejemplo =DISTR.GAMMA.INV(0,8; 1; 1) resulta 1,61.

- Vea también las siguientes funciones:
  DISTR.GAMMA

GAMMA.LN

Calcula el logaritmo lógico de la función gamma, G(x).

Sintaxis GAMMA.LN(X)
X es el valor para el que se calcula el logaritmo lógico de la función gamma.

Ejemplo =GAMMA.LN(2) resulta 0.

- Vea también las siguientes funciones:
  Factorial

DISTR.GAMMA

Calcula las probabilidades de una variable aleatoria de distribución gamma.

Sintaxis DISTR.GAMMA(x; alfa; beta; acum)
x es el valor para el que se debe calcular la distribución gamma.
Alfa es el parámetro alfa de la distribución gamma.
Beta es el parámetro alfa de la distribución gamma.
acum = 0 calcula la función de densidad, **K** = 1 la distribución.

Ejemplo =DISTR.GAMMA(2; 1; 1; 1) resulta 0,86.

- Vea también las siguientes funciones:
  DISTR.JI, **DISTR.EXP**, **DISTR.GAMMA.INV**

GAUSS

Fija el valor integral de la distribución estándar normal.

Sintaxis GAUSS(z)
z es el valor para el que se calcula el valor integral de la distribución estándar normal.

Ejemplo GAUSS(0,19) = 0,08
GAUSS(0,0375) = 0,01

- Vea también las siguientes funciones:
  PRUEBA.Z, **DISTR.NORM**, **DISTR.NORM.ESTAND**

MEDIA.GEOM

Calcula la media geométrica de una cantidad de números positivos.

Sintaxis MEDIA.GEOM(número 1; número 2; ...)
Número 1, número 2,... son argumentos numéricos que representan una muestra.

Ejemplo Al introducir los valores 23, 46 y 69 en el valor de cuadro de texto 1, 2 y 3 se obtiene como resultado 41,79.
MEDIA.GEOM(23; 46; 69) = 41,79. Así pues, la media geométrica de esta muestra es 41,79.

- Vea también las siguientes funciones:
  MEDIA.ACOTADA, **MEDIA.ARMO**, **MEDIANA**, **PROMEDIO**, **MODA**

MEDIA.ACOTADA

Calcula el promedio de un grupo de datos sin tener en cuenta el valor en los márgenes.

Sintaxis MEDIA.ACOTADA(datos; alfa)
datos es la matriz de datos dentro de la muestra.
Alfa es el porcentaje de los datos marginales que no se deben tener en cuenta.

Ejemplo =MEDIA.ACOTADA(A1:A50; 0,1) calcula el promedio de los números en A1:A50, sin tener en cuenta el 5 por ciento de valores más bajos y el 5 % de valores más altos. Los porcentajes se aplican a la cantidad del promedio no recortado, no a la cantidad de los sumandos.

- Vea también las siguientes funciones:
  MEDIA.GEOM, **MEDIA.ARMO**, **MEDIANA**, **PROMEDIO**, **MODA**

PRUEBA.Z

Calcula una estadística de dos páginas de una prueba z con distribución normal.

Sintaxis PRUEBA.Z(datos; x; STD)
datos es la matriz de datos.
X es el valor a probar.
sigma (opcional) es la desviación estándar de la totalidad de fondo. Si falta este argumento se procesa con la desviación estándar de la muestra en cuestión.

Ejemplo =PRUEBA.Z(A1:A50; 12) da como resultado la probabilidad de que el valor 12 pertenezca a la totalidad distribuida normalmente de los datos en A1:A50.

- Vea también las siguientes funciones:
  INTERVALO.CONFIANZA, **DISTR.NORM.INV**, **DISTR.NORM**, **NORMALIZACIÓN**, **DISTR.NORM.ESTAND.INV**, **DISTR.NORM.ESTAND**

MEDIA.ARMO

Calcula la media armonizada de una cantidad de datos.

Sintaxis

MEDIA.ARMO(número 1; número 2; ...)
Número 1, número 2,... son hasta 30 argumentos que se emplean para calcular la media armonizada.

Ejemplo Al introducir el valor 23, 46 y 69 en el cuadro de texto número 1, 2 y 3 se obtiene como resultado 37,64.

MEDIA.ARMO(23;46;69) = 37,64. El promedio armonizado de esta muestra es, por tanto, 7,64.

- Vea también las siguientes funciones:

  MEDIA.GEOM, **MEDIA.ACOTADA**, **MEDIANA**, **PROMEDIO**, **MODA**

DISTR.HIPERGEOM

Calcula intervalos de probabilidad en variables aleatorias distribuidas de forma geométrica.

Sintaxis DISTR.HIPERGEOM(núm_éxito; n_muestra; pobl_éxito; n_total)
núm_éxito es la cantidad obtenida en el resultado de la muestra.
n_muestra es el tamaño de la muestra.
pobl_éxito es la cantidad de los resultados posibles en la totalidad de fondo.
n_total es el tamaño de la totalidad de fondo.

Ejemplo

=DISTR.HIPERGEOM(2; 2; 90; 100) tiene como resultado 0,81. Si, de cada 100 tostadas con mantequilla que se caen de la mesa, 90 llegan al suelo por el lado de la mantequilla, la probabilidad de que 2 tostadas que se caen de la mesa lleguen al suelo por el lado de la mantequilla es del 81 %.

- Vea también las siguientes funciones:

  DISTR.BINOM, **COMBINAR**, **FACT**, **NEGBINOMDIST**, **PERMUTACIONES**

Funciones estadísticas, tercera parte

K.ESIMO.MAYOR

Calcula el valor k superior de un grupo de datos.

Sintaxis K.ESIMO.MAYOR(datos; rango_K)
Datos es la matriz de los datos de la muestra.
Rango_K es la jerarquía del valor.

Ejemplo =K.ESIMO.MAYOR(A1:C50; 2) da el segundo valor superior en A1:C50.

- Vea también las siguientes funciones:

  K.ESIMO.MENOR, **MÁX**, **MEDIANA**, **PERCENTIL**, **RANGO.PERCENTIL**, **CUARTIL**

K.ESIMO.MENOR

Calcula el valor k inferior de un grupo de datos.

Sintaxis

K.ESIMO.MENOR(datos; rango_K)
Datos es la matriz de los datos en la muestra.
Rango_K es el rango del valor.

Ejemplo =K.ESIMO.MENOR(A1:C50; 2) da el segundo valor menor en A1:C50.

- Vea también las siguientes funciones:

  K.ESIMO.MAYOR, **MEDIANA**, **MÍN**, **PERCENTIL**, **RANGO.PERCENTIL**, **CUARTIL**

INTERVALO.CONFIANZA

Calcula un intervalo confianza (1 alfa) para distribución normal.

Sintaxis INTERVALO.CONFIANZA(alfa; desv_estándar; tamaño)
Alfa es el nivel del intervalo de confianza.
STD es la desviación estándar de la totalidad base.
tamaño es el tamaño de la totalidad base.

Ejemplo =INTERVALO.CONFIANZA(0,05; 1,5; 100) da 0,29.

- Vea también las siguientes funciones:

  PRUEBA.Z

COEF.DE.CORREL

Calcula el coeficiente de correlación de un tamaño aleatorio bidimensional.

Sintaxis COEF.DE.CORREL(datos_1; datos_2)
Datos_1 es la matriz del primer registro de datos.
Datos_2 es la matriz del segundo registro de datos.

Ejemplo =COEF.DE.CORREL(A1:A50; B1:B50) calcula el coeficiente de correlación como medida de la conexión linear entre dos series de datos.

- Vea también las siguientes funciones:

  FISHER, **PRUEBA.FISHER.INV**, **COVAR**

COVAR

Calcula la covarianza en todos los pares de datos de los productos representados.

Sintaxis COVAR(datos 1; datos 2)
Datos 1 es la matriz del primer registro de datos.
Datos 2 es la matriz del segundo registro de datos.

Ejemplo =COVAR(A1:A30; B1:B30)

- Vea también las siguientes funciones:

  FISHER, **PRUEBA.FISHER.INV**, **COEF.DE.CORREL**

BINOM.CRIT

Calcula el valor menor para el que el intervalo de probabilidad acumulada de la distribución binominal es igual o superior a uno de los intervalos de probabilidad límite.

Sintaxis

BINOM.CRIT(ensayos; prob_éxito; alfa)
ensayos es el total de intentos.
prob_éxito es el intervalo de probabilidad de éxito de un intento.
Alfa es el intervalo de probabilidad límite que se debe alcanzar o superar.

Ejemplo =BINOM.CRIT(100; 0,5; 0,1) da como resultado 44.

- Vea también las siguientes funciones:

  DISTR.BINOM, **FACT**, **DISTR.HIPERGEOM**, **COMBINAR**, **NEGBINOMDIST**, **PERMUTACIONES**, **PROBABILIDAD**

CURTOSIS

Calcula la curtosis (exceso) de un grupo de datos. Se deben introducir como mínimo 4 valores.

Sintaxis CURTOSIS(número 1; número 2; ...)
Número1, número2, ... son argumentos numéricos que representan una muestra de la distribución.

Ejemplo =CURTOSIS(A1;A2;A3;A4;A5;A6)

- Vea también las siguientes funciones:

  COEFICIENTE.ASIMETRIA, **DESVEST**, **DESVESTP**, **VAR**, **VARP**

INV.LOG

Calcula la función de retorno de la distribución normal logarítmica.

Sintaxis INV.LOG(probabilidad; media; desv_estándar)
Probabilidad es el valor del intervalo de probabilidad para el cual se debe calcular la distribución normal logarítmica inversa.
Media es el promedio de la distribución normal logarítmica.
Desv_estándar es la desviación estándar de la distribución normal logarítmica.

Ejemplo

=INV.LOG(0,05; 0; 1) da 0,19.

- Vea también las siguientes funciones:

  EXP, **LN**, **LOG**, **LOG10**, **DISTR.LOG.NORM**

DISTR.LOG.NORM

Calcula el valor de la función de distribución de una variable aleatoria con distribución normal logarítmica.

Sintaxis DISTR.LOG.NORM(x; media; desv_estándar)
x es el valor del intervalo de probabilidad para el cual se debe calcular la distribución normal logarítmica.
Media es el promedio de la distribución normal logarítmica.
Desv_estándar es la desviación estándar de la distribución normal logarítmica.

Ejemplo =DISTR.LOG.NORM(0,1; 0; 1) da 0,01.

- Vea también las siguientes funciones:
  EXP, **LN**, **LOG**, **LOG10**, **INV.LOG**

Funciones estadísticas, cuarta parte

MÁX

Calcula el mayor valor dentro de una lista de argumentos.

Sintaxis MÁX(número 1; número 2; ...)
Número 1; número 2;... son argumentos numéricos cuyo número máximo se debe determinar. Cada número se puede reemplazar por una referencia.

Ejemplo =MÁX(A1;A2;A3;50;100;200) da como resultado el valor mayor de la lista.
=MÁX(A1:B100) da como resultado el valor mayor de la lista.

- Vea también las siguientes funciones:
  BSMAX, **MÍN**

MÁXA

Calcula el valor máximo de una lista de argumentos. A diferencia de MÁX, la función MÁXA admite texto, que se valora como 0.

Sintaxis MÁXA(Valor 1; Valor 2; ...Valor 30)
Valor 1; Valor 2;...Valor 30 son argumentos cuyo valor máximo debe calcularse. En lugar de valores, también se pueden utilizar referencias. Cuando se inserta un texto, éste se valora como 0.

Ejemplo =MÁXA(A1;A2;A3;50;100;200;texto) da como resultado el valor máximo de esta lista.
=MÁXA(A1:B100) produce como resultado el valor máximo de esta lista.

- Vea también las siguientes funciones:
  MÍNA

MEDIANA

Calcula la mediana (punto medio) de los números introducidos. Es el valor que, en una cantidad de números impar, se encuentra en medio de la lista. Si la cantidad de valores es par, se determinan los dos valores medios.

Sintaxis MEDIANA(número 1; número 2; ...)
Número 1; número 2;... son argumentos que representan una muestra. Cada número se puede reemplazar por una referencia.

Ejemplo (cantidad impar): =MEDIANA(1, 5, 9, 20, 21) da como resultado el valor medio de esta lista, es decir 9.
(número par): =MEDIANA(1, 5, 9, 20) devuelve como resultado el promedio de los dos valores intermedios 5 y 9, es decir, 7.

- Vea también las siguientes funciones:
  CONTAR, **CONTARA**, **DBPROMEDIO**, **BDPROMEDIO**, **MODA**, **SUMA**

MÍN

Calcula el número menor introducido como argumento.

Sintaxis MÍN(número 1; número 2; ...)
Número 1; número 2;... son argumentos numéricos de los cuales se debe calcular el número más pequeño. Cada número se puede reemplazar por una referencia.

Ejemplo =MÍN(A1:B100) da como resultado el valor más pequeño de la lista.

- Vea también las siguientes funciones:
  BDMIN, **MÁX**

MÍNA

Devuelve el valor mínimo de una lista de argumentos. También puede introducir texto. El valor del texto es 0.

Sintaxis MÍNA(Valor 1; Valor 2; ...Valor 30)
Valor 1; Valor 2;...Valor 30 son argumentos cuyo valor mínimo debe calcularse. En lugar de valores también se pueden utilizar referencias. El texto se valora como 0.

Ejemplo =MÍNA(1; texto; 20) devuelve el valor mínimo de esta lista.
=MÍNA(A1:B100) devuelve el valor mínimo de esta lista.

- Vea también las siguientes funciones:
  MÁXA

DESVPROM

Calcula la desviación absoluta media de una secuencia de características y sus promedios. Muestra la dispersión de un grupo de datos.

Sintaxis PROMEDIO(número 1; número 2; ...número 30)
Número 1, número 2, ...número 30 son argumentos que representan una muestra aleatoria. En lugar de números también se pueden utilizar referencias.

Ejemplo =PROMEDIO(A1:A50)

- Vea también las siguientes funciones:
  DESVEST, **DESVESTP**, **DESVIA2**, **VAR**, **VARP**

PROMEDIO

Calcula el promedio aritmético de los argumentos.

Sintaxis PROMEDIO(número 1; número 2; ...número 30)
Número 1; número 2;...número 30 son argumentos numéricos que representan una muestra aleatoria extraída de una determinada población. En lugar de números también se pueden utilizar referencias.

Ejemplo =PROMEDIO(A1:A50)

- Vea también las siguientes funciones:
  MEDIA.GEOM, **MEDIA.ACOTADA**, **MEDIA.ARMO**, **MEDIANA**, **MODA**

PROMEDIOA

Calcula la media aritmética de los argumentos. El texto se valora como 0.

Sintaxis PROMEDIOA(Valor 1; Valor 2; ...Valor 30)
Valor 1; Valor 2;...Valor 30 son argumentos que representan una muestra aleatoria extraída de una población. En lugar de valores también se pueden utilizar referencias. El texto se valora como 0.

Ejemplo =PROMEDIOA(A1:A50)

- Vea también las siguientes funciones:
  PROMEDIO,

MODA

Calcula el valor más frecuente en una matriz o en un conjunto de datos. Si existen varios valores que se repiten con la misma frecuencia, se devuelve como resultado el menor de ellos. Si ningún valor se repite dos veces, se muestra un mensaje de error.

Sintaxis MODA(número 1; número 2; ...número 30)
Número 1; número 2,... son argumentos numéricos que representan una muestra aleatoria. En lugar de números también se pueden utilizar referencias.

Ejemplo =MODA(A1:A50)

- Vea también las siguientes funciones:
  MEDIA.GEOM, **MEDIA.ACOTADA**, **MEDIA.ARMO**, **MEDIANA**, **PROMEDIO**

NEGBINOMDIST

Calcula la probabilidad de una variable aleatoria negativa de dirección binomial.

Sintaxis NEGBINOMDIST(núm_fracasos; núm_éxitos; prob_éxito)
Núm_fracasoses el número de fracasos en la serie de ensayos.
Núm_éxitos es el número de éxitos en la serie de ensayos.
Prob_éxito es la probabilidad de éxito de un intento.

Ejemplo NEGBINOMDIST(1; 1; 0,5) da como resultado 0,25.

- Vea también las siguientes funciones:
  DISTR.BINOM, **FACT**, **DISTR.HIPERGEOM**, **COMBINAR**, **PERMUTACIONES**

DISTR.NORM.INV

Calcula el percentil de distribución normal correspondiente al promedio y desviación estándar indicados.

Sintaxis DISTR.NORM.INV(probabilidad; media; desv_estándar)
Probabilidad es el valor probable en función del cual debe calcularse la distribución normal inversa.
Media es el promedio de la distribución normal.
Desv_estándar es la desviación estándar de la distribución normal.

Ejemplo DISTR.NORM.INV(0,9; 63; 5) da como resultado 69,41. Si un huevo de gallina estándar pesa 63g, con una desviación estándar de 5, la probabilidad de que un huevo no pese más de 69,41g es del 90 %.

- Vea también las siguientes funciones:

  PRUEBA.Z, **DISTR.NORM**, **NORMALIZACIÓN**, **DISTR.NORM.ESTAND.INV**, **DISTR.NORM.ESTAND**

DISTR.NORM

Calcula la probabilidad de una variable aleatoria de distribución normal en función del promedio y desviación estándar indicados.

Sintaxis DISTR.NORM(x; media; desv_estándar; acum)
X es el valor de distribución en función del cual debe calcularse la distribución normal.
Media es el promedio aritmético de la distribución.
Desv_estándar es la desviación estándar de la distribución.
Acum = 0 calcula la densidad de probabilidad; **Acum** = 1 calcula la distribución.

Ejemplo =DISTR.NORM(70; 63; 5; 0) da como resultado 0,03.
=DISTR.NORM(70; 63; 5; 1) da como resultado 0,92.

- Vea también las siguientes funciones:

  PRUEBA.Z, **DISTR.NORM.INV**, **NORMALIZACION**, **DISTR.NORM.ESTAND.INV**, **DISTR.NORM.ESTAND**

PEARSON

Calcula el coeficiente de correlación producto o momento r de Pearson.

Sintaxis PEARSON(Datos 1; Datos 2)
Datos 1 es la matriz del primer registro de datos.
Datos 2 es la matriz del segundo registro de datos.

Ejemplo =PEARSON(A1:A30; B1:B30) devuelve el coeficiente de correlación de Pearson correspondiente a ambas series de datos.

- Vea también las siguientes funciones:

  INTERSECCIÓN.EJE, **COEFICIENTE.R2**, **ESTIMACIÓN.LINEAL**, **PENDIENTE**, **ERROR.TÍPICO.XY**

PHI

Calcula los valores de la función de distribución para una distribución normal estándar.

Sintaxis PHI(Número)
Número es el valor en función del cual se calcula la distribución normal estándar.

Ejemplo PHI(2,25) = 0,03
PHI(-2,25) = 0,03
PHI(0) = 0,4

- Vea también las siguientes funciones:

  DISTR.NORM.ESTAND

POISSON

Calcula la probabilidad de una variable aleatoria según la distribución de Poisson.

Sintaxis POISSON(x; media; acumulado)
X es el valor en función del cual debe calcularse una distribución de Poisson.
Media es el promedio de la distribución de Poisson.
Acumulado = 0 calcula la densidad de probabilidad; **Acumulado** = 1 calcula la distribución.

Ejemplo

=POISSON(60;50;1) da como resultado 0,93.

- Vea también las siguientes funciones:
  DISTR.EXP

PERCENTIL

Para una muestra aleatoria se calcula el percentil alfa. Un percentil devuelve el valor de escala respecto a una serie de datos, que se halla ubicado en un porcentaje **alfa** de la escala desde el valor mínimo hasta el valor máximo de la serie de datos. Para **alfa** = 25 %, el percentil recibe el nombre de primer cuartil; para **alfa** = 50 % se denomina MEDIANA.

Sintaxis PERCENTIL(Datos;K)
Datos es la matriz de los datos.
K determina el porcentaje del percentil, situado entre 0 y 1.

Ejemplo =PERCENTIL(A1:A50; 0,1) da como resultado el valor de la serie de datos que constituye el 10 % de la escala de todos datos contenidos en A1:A50.

- Vea también las siguientes funciones:
  K.ESIMO.MAYOR, **K.ESIMO.MENOR**, **MÁX**, **MEDIANA**, **MÍN**, **RANGO.PERCENTIL**, **CUARTIL**

RANGO.PERCENTIL

Calcula el rango porcentual (alfa) de un valor de una muestra aleatoria.

Sintaxis RANGO.PERCENTIL(Datos; x)
Datos es la matriz de los datos de la muestra aleatoria.
X es el valor cuyo rango porcentual debe calcularse.

Ejemplo =RANGO.PERCENTIL(A1:A50; 50) produce como resultado el rango porcentual del valor 50 en el área completa de todos los valores de A1:A50. Si 50 queda fuera del área, la función produce un mensaje de error.

- Vea también las siguientes funciones:
  K.ESIMO.MAYOR, **K.ESIMO.MENOR**, **MÁX**, **MEDIANA**, **MÍN**, **PERCENTIL**, **CUARTIL**

CUARTIL

Calcula el cuartil de un conjunto de datos.

Sintaxis CUARTIL(Datos; Cuartil)
Datos es la matriz de los datos de la muestra aleatoria.
Cuartil es el tipo de cuartil. (0 = MÍN, 1 = 25 %, 2 = 50 % (MEDIANA), 3 = 75 % y 4 = MÁX.)

Ejemplo =CUARTIL(A1:A50; 2) da como resultado el valor correspondiente al 25 % de la escala comprendida entre el valor mínimo y el máximo del área A1:A50.

- Vea también las siguientes funciones:
  K.ESIMO.MAYOR, **K.ESIMO.MENOR**, **MÁX**, **MEDIANA**, **MÍN**, **PERCENTIL**, **RANGO.PERCENTIL**

Funciones estadísticas, quinta parte

JERARQUÍA

Calcula la jerarquía de una cifra dentro de una lista de cifras.

Sintaxis JERARQUÍA(Número; Datos; Orden)
Número es el número para el que se define la jerarquía.
Datos es la matriz de los datos de la muestra.
Orden (opcional) es el orden de la sucesión jerárquica. = 0 significa ascendente, = 1 significa descendente.

Ejemplo =JERARQUÍA(A10; A1:A50) indica la posición jerárquica del valor en A10 respecto a los valores comprendidos entre A1 y A50. Si **Valor** no existe dentro del área, se muestra un mensaje de error.

- Vea también las siguientes funciones:
  RANGO.PERCENTIL

COEFICIENTE.ASIMETRIA

Calcula la asimetría de una distribución.

Sintaxis COEFICIENTE.ASIMETRIA(Número 1; Número 2; ...Número 30)
Número 1, Número 2... son argumentos numéricos que representan una muestra de la distribución. También es posible indicar áreas.

Ejemplo =COEFICIENTE.ASIMETRIA(A1:A50) calcula el valor de la asimetría para los datos de la referencia indicada.

- Vea también las siguientes funciones:
  CURTOSIS, **DESVEST**, **DESVESTP**, **VAR**, **VARP**

PRONÓSTICO

Calcula un valor a partir de la regresión lineal.

Sintaxis PRONÓSTICO(x; Datos_Y; Datos_X)
X es el valor X para el que se ha de calcular el valor Y de la regresión lineal.
Datos_Y es la matriz de los datos Y.
Datos_X es la matriz de los datos X.

Ejemplo =PRONÓSTICO(50; A1;A50; B1;B50) devuelve el valor Y que se espera para el valor X 50 cuando los valores X y los valores Y de las dos referencias están vinculados a través de una regresión lineal.

- Vea también las siguientes funciones:

  ESTIMACIÓN.LINEAL, **ESTIMACION.LOGARÍTMICA**, **TENDENCIA**, **CRECIMIENTO**

DESVEST

Calcula una estimación de la desviación estándar a partir de una muestra.

Sintaxis DESVEST(Número 1;Número 2;...Número 30)
Número 1, Número 2,... son argumentos numéricos que representan una muestra de un total base.

Ejemplo =DESVEST(A1:A50) devuelve la desviación estándar calculada a partir de los datos de la referencia.

- Vea también las siguientes funciones:

  DESVPROM, **PROMEDIO**, **MODA**, **DESVESTP**, **VAR**

DESVESTA

Calcula una estimación de la desviación estándar a partir de una muestra. El texto adopta el valor 0.

Sintaxis DESVESTA(Valor 1;Valor 2;...Valor 30)
Valor 1, Valor 2,...Valor 30 son argumentos que representan una muestra tomada de un total base. También se admiten textos, a los que se asignará el valor 0.

Ejemplo =DESVESTA(A1:A50) devuelve la desviación estándar calculada a partir de los datos de la referencia.

- Vea también las siguientes funciones:

  DESVEST, **DESVESTP** y **DESVESTPA**.

DESVESTP

Calcula la desviación estándar a partir del total base.

Sintaxis DESVESTP(Número 1; Número 2;...Número 30)
Número 1, Número 2,... son argumentos numéricos que representan una muestra sobre un total.

Ejemplo =DESVESTP(A1:A50) devuelve la desviación estándar de los datos de la referencia.

- Vea también las siguientes funciones:

  DESVPROM, **PROMEDIO**, **MODA**, **DESVEST**, **VARP**

DESVESTPA

Calcula la desviación estándar a partir del total base. El texto adopta el valor 0.

Sintaxis DESVESTPA(Valor 1;Valor 2;...Valor 30)
valor 1,Valor 2,...Valor 30 son argumentos que representan una muestra tomada de un total. Al texto se le asigna el valor 0.

Ejemplo =DESVESTPA(A1:A50) devuelve la desviación estándar de los datos de la referencia.

- Vea también las siguientes funciones:
  DESVESTP y **DESVESTA**.

NORMALIZACION

Calcula el valor normalizado de una distribución caracterizada por un promedio y una desviación estándar.

Sintaxis NORMALIZACIÓN(x; media ;desv_estándar)
X es el valor que va a normalizarse.
Media es el promedio de desplazamiento que se desea realizar.
Desv_estándar es la desviación estándar en la que hay que escalar.

Ejemplo =NORMALIZACIÓN(11; 10; 1) devuelve 1. En una distribución normal con una media de 10 y una desviación estándar de 1, el valor 11 está a la misma distancia por encima del promedio 10 que el valor 1 sobre el promedio 0 de la distribución normal estándar.

- Vea también las siguientes funciones:
  PRUEBA.Z, **DISTR.NORM.INV**, **DISTR.NORM**, **DISTR.NORM.ESTAND.INV**, **DISTR.NORM.ESTAND**

DISTR.NORM.ESTAND.INV

Calcula el percentil de la distribución normal estándar.

Sintaxis DISTR.NORM.ESTAND.INV(probabilidad)
Probabilidad es el valor de probabilidad para el que se va a calcular la distribución normal estándar inversa.

Ejemplo DISTRIB.NORM.ESTAND.INV(0,908789) devuelve 1,3333.

- Vea también las siguientes funciones:
  PRUEBA.Z, **DISTR.NORM.INV**, **DISTR.NORM**, **NORMALIZACIÓN**, **DISTR.NORM.ESTAND**

DISTR.NORM.ESTAND

Calcula los valores de la función de distribución de las variables aleatorias de distribución normal estándar.

Sintaxis DISTR.NORM.ESTAND(z)
Z es el valor para el que se calcula la distribución normal estándar.

Ejemplo =DISTR.NORM.ESTAND(1) devuelve 0,84. La superficie situada debajo de la curva de la distribución normal estándar a la izquierda del valor X 1 es el 84 % de la superficie total.

- Vea también las siguientes funciones:
  PRUEBA.Z, **DISTR.NORM.INV**, **DISTR.NORM**, **PHI**, **NORMALIZACIÓN**, **DISTR.NORM.ESTAND.INV**

PENDIENTE

Calcula la pendiente de la línea de regresión lineal. La pendiente se adapta a los puntos de datos definidos en los valores Y y los valores X.

Sintaxis PENDIENTE(Datos_Y; Datos_X)
Datos_Y es la matriz de los datos Y.
Datos_X es la matriz de los datos X.

Ejemplo =PENDIENTE(A1:A50; B1:B50)

- Vea también las siguientes funciones:
  INTERSECCIÓN.EJE, **COEFICIENTE.R2**, **PEARSON** ;**ESTIMACIÓN.LINEAL**, **ESTIMACIÓN.LOGARÍTMICA**, **ERROR.TÍPICO.XY**, **TENDENCIA**

ERROR.TÍPICO.XY

Calcula el error típico de los valores Y calculados para todos los valores X de la regresión.

Sintaxis ERROR.TÍPICO.XY(Datos_Y; Datos_X)
Datos_Y es la matriz de los datos Y.
Datos_X es la matriz de los datos X.

Ejemplo =ERROR.TÍPICO.XY(A1:A50; B1:B50)

- Vea también las siguientes funciones:
  INTERSECCIÓN.EJE, **COEFICIENTE.R2**, **PEARSON**, **ESTIMACIÓN.LINEAL**, **ESTIMACIÓN.LOGARÍTMICA**, **PENDIENTE**

DESVIA2

Efectúa la suma de las desviaciones cuadradas de datos a partir del valor medio de la muestra.

Sintaxis DESVIA2(Número 1; Número 2; ...Número 30)
Número 1, 2,...Número 30 son argumentos numéricos que representan una muestra. También es posible definir referencias.

Ejemplo =DESVIA2(A1:A50)

- Vea también las siguientes funciones:
  DESVPROM, **DESVEST**, **DESVESTP**, **VAR**, **VARP**

DISTR.T.INV

Calcula percentiles de la distribución t de Student para los grados de libertad indicados.

Sintaxis DISTR.T.INV(probabilidad; grados_libertad)
Probabilidad es el valor de probabilidad para el que se pretende calcular la distribución t inversa.
Grados_libertad es la cantidad de grados de libertad de la distribución t.

Ejemplo =DISTR.T.INV(0,1; 6) devuelve 1,94

- Vea también las siguientes funciones:
  PRUEBA.T, **DISTR.T**

PRUEBA.T

Calcula la estadística de una prueba T de Student.

Sintaxis PRUEBA.T(Datos_1; Datos_2; Modo; Tipo)
Datos_1 es la matriz del primer registro de datos.
Datos_2 es la matriz del segundo registro de datos.
Modo = 1 calcula la prueba en una cola, **Modo** = 2 calcula la prueba de dos colas.
Tipo muestra el modelo de prueba t que debe realizarse. El tipo 1 significa por pares; 2 significa dos muestras con varianzas iguales (homoscedástico); 3 quiere decir que se toman dos muestras con varianzas distintas (heteroscedástico).

Ejemplo =PRUEBA.T(A1:A50; B1:B50; 2; 2)

- Vea también las siguientes funciones:
  DISTR.T.INV, **DISTR.T**

DISTR.T

Calcula los valores de la función de distribución (1-alfa) de una variable aleatoria de distribución t (Student).

Sintaxis DISTR.T(x; grados_libertad; Modo)
X es el valor para el que se quiere calcular la distribución t.
Grados_libertad es el grado de libertad de la distribución t.
Modo = 1 calcula la prueba en una cola, **Modo** = 2 calcula la prueba de dos colas.

Ejemplo =DISTR.T(12; 5; 1)

- Vea también las siguientes funciones:
  DISTR.T.INV, **PRUEBA.T**

VAR

Realiza una estimación de la varianza a partir de una muestra.

Sintaxis VAR(Número 1; Número 2; ...Número 30)
Número 1,Número 2,... son argumentos numéricos que representan una muestra sobre un total base. También es posible indicar referencias.

Ejemplo =VAR(A1:A50)

- Vea también las siguientes funciones:
  DESVPROM, **PROMEDIO**, **MODA**, **DESVEST**

VARA

Realiza una estimación de la varianza a partir de una muestra. El texto adopta el valor 0.

Sintaxis VARA(Valor 1; Valor 2; ... Valor 30)
Valor 1, Valor 2,...Valor 30 son argumentos que representan una muestra tomada de un total base. También es posible indicar referencias. El texto adopta el valor 0.

Ejemplo =VARA(A1:A50)

- Vea también las siguientes funciones:
  VAR, **DESVEST** y **DESVESTA**

VARP

Calcula la varianza a partir de un total base.

Sintaxis VAR(Número 1; Número 2; ...Número 30)
Número 1, Número 2,... son argumentos numéricos que representan un total base.

Ejemplo =VAR(A1:A50)

- Vea también las siguientes funciones:
  DESVPROM, **PROMEDIO**, **MODA**, **DESVESTP**

VARPA

Calcula la varianza a partir del total base. El texto adopta el valor 0.

Sintaxis VARPA(Valor 1; Valor 2; ... Valor 30)
Valor 1, Valor 2,...Valor 30 son argumentos que representan un total base.

Ejemplo =VARPA(A1:A50)

- Vea también las siguientes funciones:
  DESVESTP y **DESVESTPA**

PERMUTACIONES

Devuelve el número de permutaciones para un número determinado de objetos.

Sintaxis PERMUTACIONES(Cantidad_1; Cantidad_2)
Cantidad_1 es la cantidad total de elementos.
Cantidad_2 es la cantidad seleccionada de elementos.

Ejemplo =PERMUTACIONES(6; 3) devuelve 120. Existen 120 posibilidades distintas de obtener series de 3 cartas a partir de 6 naipes.

- Vea también las siguientes funciones:
  DISTR.BINOM, **FACT**, **DISTR.HIPERGEOM**, **COMBINAR**, **BINOM.CRIT**, **NEGBINOMDIST**

PERMUTACIONESA

Calcula el número de variaciones de elementos con repetición.

Sintaxis PERMUTACIONESA(Cantidad_1; Cantidad_2)
Cantidad_1 es la cantidad total de objetos.
Cantidad_2 define la cantidad seleccionada de objetos.

Ejemplo ¿Cuántas veces se pueden extraer 2 elementos de un conjunto de 11 elementos?
=PERMUTACIONESA(11;2) da como resultado 121.
=PERMUTACIONESA(6; 3) da como resultado 216. Si se toman 6 naipes y se saca una carta a la vez que se devuelve antes de sacar de la siguiente, hay 216 posibilidades distintas de obtener una sucesión de tres cartas.

- Vea también las siguientes funciones:
  DISTR.BINOM, **FACT**, **DISTR.HIPERGEOM**, **COMBINAR2**, **BINOM.CRIT**, **NEGBINOMDIST**

PROBABILIDAD

Calcula la probabilidad de obtener un intervalo cerrado que conste de dos valores. Si no se indica el límite superior, esta función calcula la probabilidad a partir del supuesto de que los valores que pertenecen a los datos son iguales al valor del límite inferior.

Sintaxis PROBABILIDAD(intervalo_x; intervalo_probabilidad; límite_inf; límite_sup)
Intervalo_x es la matriz de los datos de la muestra.
Intervalo_probabilidad es la matriz de las correspondientes probabilidades.
Límite_inf es el inicio del intervalo de valor cuyas probabilidades se quieren sumar.
Límite_sup (opcional) es el final del intervalo de valores cuyas probabilidades se quieren sumar. Si falta este parámetro se calcula la probabilidad de que exista exactamente el valor **Límite_inf**.

Ejemplo =PROBABILIDAD(A1:A50; B1:B50; 50; 60) calcula la probabilidad de que un valor del área A1:A50 esté contenido entre los límites de 50 y 60. Para cada valor del sector A1:A50 existe un valor de probabilidad asignado en el área B1:B50.

- Vea también las siguientes funciones:
  DISTR.BINOM, **BINOM.CRIT**

DIST.WEIBULL

Calcula las probabilidades de variables aleatorias con distribución de Weibull.

Sintaxis DIST.WEIBULL(x; alfa ; beta ; acumulado)
X es el valor para el que se calcula la distribución de Weibull.
Alfa es el parámetro alfa de la distribución de Weibull.
Beta es el parámetro beta de la distribución de Weibull.
Acumulado indica el tipo de función. Si Acumulado es igual a 0 se calcula la función acumulada; si Acumulado es igual a 1 se calcula la distribución.

Ejemplo =DIST.WEIBULL(2; 1; 1; 1) da como resultado 0,86.

- Vea también las siguientes funciones:
  DISTR.EXP

D. Referencia rápida de R

R reference card, by Jonathan Baron

Parentheses are for functions, brackets are for indicating the position of items in a vector or matrix. (Here, items with numbers like `x1` are user-supplied variables.)

Miscellaneous

`q()`: quit
`<-`: assign
`INSTALL package1`: install package1
`m1[,2]`: column 2 of matrix `m1`
`m1[,2:5]` or `m1[,c(2,3,4,5)]`: columns 2–5
`m1$a1`: variable `a1` in data frame `m1`
`NA`: missing data
`is.na`: true if data missing
`library(mva)`: load (e.g.) the `mva` package

Help

`help(command1)`: get help with `command1` (NOTE: USE THIS FOR MORE DETAIL THAN THIS CARD CAN PROVIDE.)
`help.start()`: start browser help
`help(package=mva)`: help with (e.g.) package `mva`
`apropos("topic1")`: commands relevant to topic1
`example(command1)`: examples of `command1`

Input and output

`source("file1")`: run the commands in `file1`.
`read.table("file1")`: read in data from file1
`data.entry()`: spreadsheet
`scan(x1)`: read a vector `x1`
`download.file(url1)`: from internet
`url.show(url1), read.table.url(url1)`: remote input
`sink("file1")`: output to `file1`, until `sink()`
`write(object, "file1")`: writes an object to `file1`
`write.table(dataframe1,"file1")`: writes a table

Managing variables and objects

`attach(x1)`: put variables in `x1` in search path
`detach(x1)`: remove from search path
`ls()`: lists all the active objects.
`rm(object1)`: removes object1
`dim(matrix1)`: dimensions of `matrix1`
`dimnames(x1)`: names of dimensions of `x1`
`length(vector1)`: length of `vector1`
`1:3`: the vector `1,2,3`
`c(1,2,3)`: creates the same vector
`rep(x1,n1)`: repeats the vector `x1` `n1` times
`cbind(a1,b1,c1), rbind(a1,b1,c1)`: binds columns or rows into a matrix
`merge(df1,df2)`: merge data frames
`matrix(vector1,r1,c1)`: make `vector1` into a matrix with `r1` rows and `c1` columns
`data.frame(v1,v2)`: make a data frame from vectors `v1` and `v2`
`as.factor(), as.matrix(), as.vector()`: conversion
`is.factor(), is.matrix(), is.vector()`: what it is
`t()`: switch rows and columns
`which(x1==a1)`: returns indices of `x1` where `x1==a1`

Control flow

`for (i1 in vector1)`: repeat what follows
`if (condition1) ...else ...`: conditional

Arithmetic

`%*%`: matrix multiplication
`%/%, ^, %%, sqrt()`: integer division, power, modulus, square root

Statistics

`max(), min(), mean(), median(), sum()`, var(): as named
`summary(data.frame)`: prints statistics
`rank(), sort()` rank and sort
`ave(x1,y1)`: averages of `x1` grouped by factor `y1`
`by()`: apply function to data frame by factor
`apply(x1,n1,function1)`: apply `function1` (e.g. mean) to x by rows (n1=1) or columns (`n2=2`)
`tapply(x1,list1,function1)`: apply `function` to `x1` by `list1`
`table()`: make a table
`tabulate()`: tabulate a vector

basic statistical analysis

`aov(), anova(), lm(), glm()`: linear and nonlinear models, anova
`t.test()`: t test
`prop.test(), binom.test()`: sign test
`chisq.test(x1)`: chi-square test on matrix `x1`
`fisher.test()`: Fisher exact test
`cor(a)`: show correlations
`cor.test(a,b)`: test correlation
`friedman.test()`: Friedman test

some statistics in `mva` package

`prcomp()`: principal components
`kmeans()`: kmeans cluster analysis
`factanal()`: factor analysis
`cancor()`: canonical correlation

Graphics

`plot(), barplot(), boxplot(), stem(), hist()`: basic plots
`matplot()`: matrix plot
`pairs(matrix)`: scatterplots
`coplot()`: conditional plot
`stripplot()`: strip plot
`qqplot()`: quantile-quantile plot
`qqnorm(), qqline()`: fit normal distribution

E. Comandos de R

Comencemos

`R` iniciar el programa

`q()` para salir

`#comentario` para escribir comentarios

`help()` ayuda, para salir `q`

`help.start()` ayuda en modo html

`?comando` ayuda sobre un comando

`help.search("comando")` busca en la ayuda de los comandos de R la palabra "cadena"

`apropos("comando")` funciones cuyo nombre contiene la palabra pasada como argumento.

`demo(graphics)` ejemplo de las posibilidades gráficas

- Operaciones Matemáticas básicas: +, - ,*, /, ^
- Disponemos de las funciones matemáticas más usuales: `sqrt`, `exp`, `log`, `sin`, `cos`, `tan`, ...

`options()` cómo están definidas las variables de entorno

`options(digits=22)` para trabajar con el máximo de dígitos significativos

`<-` asignaciones

`dato` muestra el contenido de la variable `dato`

`ls()` qué objetos están en memoria

`objects()` qué objetos están en memoria

`rm(x)` borra el objeto x

`c(datos)` soncatenar

`x[i]` dato i-esimo del vector `x`

`datos<-scan()` introducir datos

`seq(mínimo,máximo,[incremento,longitud]))` secuencias de números

`a:b` secuencias de números

`rep(datos,num_veces)` comando réplica

`sample(máximo,num_datos)` secuencias aleatorias

Matrices

`matrix(datos,nfilas,ncolumnas)` crear matrices

- Operaciones elementales con Matrices

  `A ± B` suma de matrices

  `A%*% B` producto de matrices

  `t(A)` transpuesta de la matriz A

  `solve(A,b)` solución del sistema de ecuaciones $A \cdot x = b$.

  `solve(A)` inversa de la matriz A

  `diag(A)` diagonal principal de la matriz A

  `det(A)` para obtener el determinante de A

Estadística descriptiva

`list.files()` lista los ficheros del directorio de trabajo

`file.show("fichero.txt")` muestra el contenido del fichero

`datos<-read.table("fichero.txt",header=TRUE)` leemos el contenido de `fichero.txt` y lo ponemos en la variable `datos`

`names(datos)` lista el nombre de las variables del objeto datos

`datos$variable` muestra el contenido de `variable`

`attach(datos)` permite acceder a las variables del objeto `datos` con el nombre `variable`

`summary(variable)` resumen

`mean(variable)` media

`median(variable)` mediana

`var(variable)` cuasivarianza

`sd(variable)` cuasidesviación típica

`range(variable)` rango

`quantile(variable, n)` percentil `n` (n se escribe en tanto por uno)

`table(variable)` frecuencias absolutas

Algunos gráficos con R

`piechart(datos, opciones)` diagrama de sectores

`hist(datos,nclass=n, opciones)` histograma, `nclass` es opcional y con él establecemos el número de intervalos

`barplot(datos, opciones)` diagrama de barras

`boxplot(datos, opciones)` diagrama de cajas y bigotes

`stem(datos)` diagrama de tallo y hojas

- Estas funciones gráficas tienen muchas posibilidades (podemos añadir varias separadas por comas), algunas de las opciones más usuales son:

  `main="titulo"` título del gráfico

  `xlab="etiqueta eje X"` etiqueta del eje X

  `ylab="etiqueta eje Y"` etiqueta del eje Y

  `type="l"` une los puntos con líneas

  `col="nombre"` color del gráfico

`help(Devices)` formatos gráficos que hay a nuestra disposición

`jpeg()` guardará el gráfico, en formato `jpeg`, en el directorio de trabajo en un fichero de nombre `Rplotxxx.jpeg`

`jpeg("fichero")` guardará el gráfico, en formato `jpeg`, en el directorio de trabajo en un fichero de nombre `fichero.jpeg`

`dev.off()` devuelve la salida a la salida estándar

`x11()` muestra el gráfico en pantalla

Estadística bidimensional

`cov(variable,Peso)` covarianza

`cor(variable,Peso)` coeficiente de correlación

`plot(x,y,type="l")` representa la nube de puntos, con la opción `l` unimos los puntos

`lm(y~x)` función que realiza los modelos lineales

`abline(lm(y~x))` añade una línea (definida por la ordenada en el origen y su pendiente) a un gráfico,

`pairs(formula(objeto_lm))` nubes de puntos

`summary(objeto lm)` sumario o resumen del objeto

`predict(objeto lm)` valores esperados

`coef(objeto)` coeficientes de la recta de regresión

`anova(objeto)` tabla del análisis de la varianza

Combinatoria

`choose(m,n)` $\binom{m}{n}$ combinaciones

`factorial(n)` $n!$ como su nombre indica

`for (n in 0:10) print(choose(n, k = 0:n))` un bucle para el triángulo de Tartaglia

`expand.grid(datos1,datos2,...)` mezcla las series de datos

Distribuciones Binomial y Normal

- La binomial

`dbinom(k, n, p)` función de probabilidad $\left(P\left(X=k\right)=\binom{n}{k}p^k\left(1-p\right)^{n-k}\right)$

`pbinom(k, n, p)` función de distribución $\left(P\left(X\leq k\right)=\sum_{i=0}^{k}P\left(X=i\right)\right)$

`qbinom(P, n, p)` calcula $x \mid P(X\leq x)=P$

`rbinom(x, n, p)` para obtener x números aleatorios de una distribución $B(n,p)$

- La "normal"

`dnorm(x,` μ,σ`)` función de densidad $f(x)=\frac{1}{\sqrt{2\pi}\sigma}e^{-\frac{1}{2}\left(\frac{x-\mu}{\sigma}\right)^2}$

`pnorm(x,` μ,σ`)` función de distibución $F(x)=\int_{-\infty}^{x}f(t)\,dt$

`qnorm(p,` μ,σ`)` calcula $x \mid P(X\leq x)=p$

`rnorm(x,` μ,σ`)` para obtener x números aleatorios siguiendo una distribución normal

Inferencia

`t.test(datos,conf.level=x,mu=`μ`)` test de hipótesis

Referencias

[1] *Grace User's Guide (for Grace-5.1.17)*, Grace Team. http://plasma-gate.weizmann.ac.il/Grace/doc/UsersGuide.html

[2] *Grace Tutorials*, Edward Vigmond. http://plasma-gate.weizmann.ac.il/Grace/doc/Tutorial.html

[3] *The Gnumeric Manual, version 1.2*. http://www.gnome.org/projects/gnumeric/doc/index.html

[4] *R para Principiantes*, the Spanish version of *R for Beginners*, translated by Jorge A. Ahumada. http://cran.r-project.org/other-docs.html

[5] A Spanish translation of *An Introduction to R* by Andrés González and Silvia González. http://cran.r-project.org/other-docs.html

[6] *Gráficos Estadísticos con R* by Juan Carlos Correa and Nelfi González. http://cran.r-project.org/other-docs.html

[7] *R reference card* by Jonathan Baron. http://cran.r-project.org/other-docs.html

[8] Documentación de OpenOffice. http://es.openoffice.org/servlets/ProjectDocumentList

F. Licencia de Documentación Libre GNU (traducción)

Versión 1.1, Marzo de 2000

Esta es la GNU Free Document License (GFDL), versión 1.1 (de marzo de 2.000), que cubre manuales y documentación para el software de la Free Software Foundation, con posibilidades en otros campos. La traducción[38] no tiene ningún valor legal, ni ha sido comprobada de acuerdo a la legislación de ningún país en particular. Vea el original `http://www.gnu.org/copyleft/fdl.html`

Los autores de esta traducción son:

- Igor Támara `mailto:ikks@bigfoot.com`
- Pablo Reyes `mailto:reyes_pablo@hotmail.com`
- Revisión : Vladimir Támara P. `mailto:vtamara@gnu.org`

0. Preámbulo

El propósito de esta licencia es permitir que un manual, libro de texto, u otro documento escrito sea ”libre” en el sentido de libertad: asegurar a todo el mundo la libertad efectiva de copiarlo y redistribuirlo, con o sin modificaciones, de manera comercial o no. En segundo término, esta licencia preserva para el autor o para quien publica una manera de obtener reconocimiento por su trabajo, al tiempo que no se consideran responsables de las modificaciones realizadas por terceros.

Esta licencia es una especie de ”copyleft” que significa que los trabajos derivados del documento deben a su vez ser libres en el mismo sentido. Esto complementa la Licencia Pública General GNU, que es una licencia de copyleft diseñada para el software libre.

Hemos diseñado esta Licencia para usarla en manuales de software libre, ya que el software libre necesita documentación libre: Un programa libre debe venir con los manuales que ofrezcan la mismas libertades que da el software. Pero esta licencia no se limita a manuales de software; puede ser usada para cualquier trabajo textual, sin tener en cuenta su temática o si se publica como libro impreso. Recomendamos esta licencia principalmente para trabajos cuyo fin sea instructivo o de referencia.

1. Aplicabilidad y definiciones

Esta Licencia se aplica a cualquier manual u otro documento que contenga una nota del propietario de los derechos que indique que puede ser distribuido bajo los términos de la Licencia. El ”Documento”, en adelante, se refiere a cualquiera de dichos manuales o trabajos. Cualquier miembro del público es un licenciatario, y será denominado como ”Usted”.

Una ”Versión Modificada” del Documento significa cualquier trabajo que contenga el Documento o una porción del mismo, ya sea una copia literal o con modificaciones y/o traducciones a otro idioma.

Una ”Sección Secundaria” es un apéndice titulado o una sección preliminar al prólogo del Documento que tiene que ver exclusivamente con la relación de quien publica o, los autores del Documento o, el tema general del Documento (o asuntos relacionados) y cuyo contenido no entra directamente en este tema general. (Por ejemplo, si el Documento es en parte un texto de matemáticas, una Sección Secundaria puede no explicar matemáticas.) La relación puede ser un asunto

[38] N. del T. Derechos Reservados en el sentido de GNU `http://www.gnu.org/copyleft/copyleft.es.html`

de conexión histórica, o de posición legal, comercial, filosófica, ética o política con el tema o la materia del texto.

Las ”Secciones Invariantes” son ciertas Secciones Secundarias cuyos títulos son denominados como Secciones Invariantes, en la nota que indica que el documento es liberado bajo esta licencia.

Los ”Textos de Cubierta” son ciertos pasajes cortos de texto que se listan, como Textos de Portada o Textos de Contra Portada, en la nota que indica que el documento es liberado bajo esta Licencia.

Una copia ”Transparente” del Documento, significa una copia para lectura en máquina, representada en un formato cuya especificación está disponible al público general, cuyos contenidos pueden ser vistos y editados directamente con editores de texto genéricos o (para imágenees compuestas por pixeles) de programas genéricos de dibujo o (para dibujos) algún editor gráfico ampliamente disponible, y que sea adecuado para exportar a formateadores de texto o para traducción automática a una variedad de formatos adecuados para ingresar a formateadores de texto. Una copia hecha en un formato de un archivo que no sea Transparente, cuyo formato ha sido diseñado para impedir o dificultar subsecuentes modificaciones posteriores por parte de los lectores no es Transparente. Una copia que no es ”Transparente” es llamada ”Opaca”.

Como ejemplos de formatos adecuados para copias Transparentes están el ASCII plano sin formato, formato de Texinfo, formato de LaTeX, SGML o XML usando un DTD disponible ampliamente, y HTML simple que sigue los estándares, diseñado para modificaciones humanas. Los formatos Opacos incluyen PostScript, PDF, formatos propietarios que pueden ser leídos y editados unicamente en procesadores de palabras propietarios, SGML o XML para los cuáles los DTD y/o herramientas de procesamiento no están disponibles generalmente, y el HTML generado por máquinas producto de algún procesador de palabras solo para propósitos de salida.

La ”Portada” en un libro impreso significa, la portada misma, más las páginas siguientes necesarias para mantener la legibilidad del material, que esta Licencia requiere que aparezca en la portada. Para trabajos en formatos que no tienen Portada como tal, ”Portada” significa el texto cerca a la aparición más prominente del título del trabajo, precediendo el comienzo del cuerpo del trabajo.

2. Copia literal

Puede copiar y distribuir el Documento en cualquier medio, sea en forma comercial o no, siempre y cuando esta Licencia, las notas de derecho de autor, y la nota de licencia que indica que esta Licencia se aplica al Documento se reproduzca en todas las copias, y que usted no adicione ninguna otra condición a las expuestas en en esta Licencia. No puede usar medidas técnicas para obstruir o controlar la lectura o copia posterior de las copias que usted haga o distribuya. Sin embargo, usted puede aceptar compensación a cambio de las copias. Si distribuye un número suficientemente grande de copias también deberá seguir las condiciones de la sección 3.

También puede prestar copias, bajo las mismas condiciones establecidas anteriormente, y puede exhibir copias publicamente.

3. Copiado en cantidades

Si publica copias impresas del Documento que sobrepasen las 100, y la nota de Licencia del Documento exige Textos de Cubierta, debe incluir las copias con cubiertas que lleven en forma clara y legible, todos esos textos de Cubierta: Textos Frontales en la cubierta frontal, y Textos Posteriores de Cubierta en la Cubierta Posterior. Ambas cubiertas deben identificarlo a Usted clara y legiblemente como quien publica tales copias. La Cubierta Frontal debe mostrar el título completo con todas las palabras igualmente prominentes y visibles. Además puede adicionar otro material en la cubierta. Las copias con cambios limitados en las cubiertas, siempre que preserven el título del Documento y satisfagan estas condiciones, puede considerarse como copia literal.

Si los textos requeridos para la cubierta son muy voluminosos para que ajusten legiblemente, debe colocar los primeros (tantos como sea razonable colocar) en la cubierta real, y continuar el resto en páginas adyacentes.

Si publica o distribuye copias Opacas del Documento cuya cantidad exceda las 100, debe incluir una copia Transparente que pueda ser leída por una máquina con cada copia Opaca, o entregar en o con cada copia Opaca una dirección en red de computador publicamente-accesible conteniendo una copia completa Transparente del Documento, sin material adicional, a la cual el público en general de la red pueda acceder a bajar anónimamente sin cargo usando protocolos de standard público. Si usted hace uso de la última opción, deberá tomar medidas necesarias, cuando comience la distribución de las copias Opacas en cantidad, para asegurar que esta copia Transparente permanecerá accesible en el sitio por lo menos un año después de su última distribución de copias Opacas (directamente o a través de sus agentes o distribuidores) de esa edición al público.

Se solicita, aunque no es requisito, que contacte a los autores del Documento antes de redistribuir cualquier gran número de copias, para permitirle la oportunidad de que le provean una versión del Documento.

4. Modificaciones

Puede copiar y distribuir una Versión Modificada del Documento bajo las condiciones de las secciones 2 y 3 anteriores, siempre que usted libere la Versión Modificada bajo esta misma Licencia, con la Versión Modificada haciendo el rol del Documento, por lo tanto licenciando la distribución y modificación de la Versión Modificada a quienquiera que posea una copia de este. En adición, debe hacer lo siguiente en la Versión Modificada:

- A. Uso en la Portada (y en las cubiertas, si hay alguna) de un título distinto al del Documento, y de versiones anteriores (que deberían, si hay alguna, estar listados en la sección de Historia del Documento). Puede usar el mismo título que versiones anteriores al original siempre que quién publicó la primera versión lo permita.

- B. Listar en la Portada, como autores, una o más personas o entidades responsables por la autoría o las modificaciones en la Versión Modificada, junto con por lo menos cinco de los autores principales del Documento (Todos sus autores principales, si hay menos de cinco).

- C. Estado en la Portada del nombre de quién publica la Versión Modificada, como quien publica.

- D. Preservar todas las notas de derechos de autor del Documento.

- E. Adicionar una nota de derecho de autor apropiada a sus modificaciones adyacentes a las otras notas de derecho de autor.

- F. Incluir, immediatamente después de la nota de derecho de autor, una nota de licencia dando el permiso público para usar la Versión Modificada bajo los términos de esta Licencia, de la forma mostrada en la Adición (LEGAL)abajo.

- G. Preservar en esa nota de licencia el listado completo de Secciones Invariantes y en los Textos de las Cubiertas que sean requeridos como se especifique en la nota de Licencia del Documento

- H. Incluir una copia sin modificación de esta Licencia.

- I. Preservar la sección llamada ”Historia”, y su título, y adicionar a esta una sección estableciendo al menos el título, el año,los nuevos autores, y quién publicó la Versión Modificada como reza en la Portada. Si no hay una sección titulada ”Historia” en el Documento, crear una estableciendo el título, el año, los autores y quien publicó el Documento como reza en la Portada, añadiendo además un artículo describiendo la Versión Modificada como se estableció en el punto anterior.

- J. Preservar la localización en red, si hay , dada en la Documentación para acceder públicamente a una copia Transparente del Documento, tanto como las otras direcciones de red dadas en el Documento para versiones anteriores en las cuáles estuviese basado. Estas pueden ubicarse en la sección ”Historia”. Se puede omitir la ubicación en red para un trabajo que sea publicado por lo menos 4 años antes que el mismo Documento, o si quien publica originalmente la versión da permiso explícitamente.

- K. En cualquier sección titulada ”Agradecimientos” o ”Dedicatorias”, preservar el título de la sección, y preservar en la sección toda la sustancia y el tono de los agradeimientos y/o dedicatorias de cada contribuyente que estén incluídas.

- L. Preservar todas las Secciones Invariantes del Documento, sin alterar su texto ni sus títulos. Números de sección o el equivalente no son considerados parte de los títulos de la sección. M. Borrar cualquier sección titulada ”Aprobaciones”. Tales secciones no pueden estar incluidas en las Versiones Modificadas.

- M. Borrar cualquier sección titulada ”Aprobaciones”. Tales secciones no pueden estar incluidas en las Versiones Modificadas.

- N. No retitular ninguna sección existente como ”Aprobaciones” o conflictuar con título con alguna Sección Invariante.

Si la Versión Modificada incluye secciones o apendices nuevos o preliminares al prólogo que califican como Secciones Secundarias y contienen material no copiado del Documento, puede opcionalmente designar algunas o todas esas secciones como invariantes. Para hacerlo, adicione sus títulos a la lista de Secciones Invariantes en la nota de licencia de la Versión Modificada. Tales títulos deben ser distintos de cualquier otro título de sección.

Puede adicionar una sección titulada ”Aprobaciones”, siempre que contenga unicamente aprobaciones de su Versión Modificada por varias fuentes–por ejemplo, observaciones de peritos o que el texto ha sido aprobado por una organización como un standard.

Puede adicionar un pasaje de hasta cinco palabras como un Texto de Cubierta Frontal, y un pasaje de hasta 25 palabras como un texto de Cubierta Posterior, al final de la lista de Textos de Cubierta en la Versión Modificada. Solamente un pasaje de Texto de Cubierta Frontal y un Texto de Cubierta Posterior puede ser adicionado por (o a manera de arreglos hechos por) una entidad. Si el Documento ya incluye un texto de cubierta para la misma cubierta, previamente adicionado por usted o por arreglo hecho por la misma entidad, a nombre de la cual está actuando, no puede adicionar otra; pero puede reemplazar la anterior, con permiso explícito de quien publicó anteriormente tal cubierta.

El(los) autor(es) y quien(es) publica(n) el Documento no dan con esta Licencia permiso para usar sus nombres para publicidad o para asegurar o implicar aprobación de cualquier Versión Modificada.

5. Combinando documentos

Puede combinar el Documento con otros documentos liberados bajo esta Licencia, bajo los términos definidos en la sección 4 anterior para versiones modificadas, siempre que incluya en la combinación todas las Secciones Invariantes de todos los documentos originales, sin modificar, y listadas todas como Secciones Invariantes del trabajo combinado en su nota de licencia.

El trabajo combinado necesita contener solamente una copia de esta Licencia, y múltiples Seccions Invariantes Idénticas pueden ser reemplazadas por una sola copia. Si hay múltiples Secciones Invariantes con el mismo nombre pero con contenidos diferentes, haga el título de cada una de estas secciones único adicionándole al final de este, en paréntesis, el nombre del autor o de quien publicó originalmente esa sección, si es conocido, o si no, un número único. Haga el mismo ajuste a los títulos de sección en la lista de Secciones Invariantes en la nota de licencia del trabajo combinado.

En la combinación, debe combinar cualquier sección titulada ”Historia” de los varios documentos originales, formando una sección titulada ”Historia”; de la misma forma combine cualquier seción titulada ”Agradecimientos”, y cualquier sección titulada ”Dedicatorias”. Debe borrar todas las secciones tituladas ”Aprobaciones.”

6. Colecciones de documentos

Puede hacer una colección consistente del Documento y otros documentos liberados bajo esta Licencia, y reemplazar las copias individuales de esta Licencia en los varios documentos con una sola copia que esté incluida en la colección, siempre que siga las reglas de esta Licencia para una copia literal de cada uno de los documentos en cualquiera de todos los aspectos.

Puede extraer un solo documento de una de tales colecciones, y distribuirlo individualmente bajo esta Licencia, siempre que inserte una copia de esta Licencia en el documento extraido, y siga esta Licencia en todos los otros aspectos concernientes a la copia literal de tal documento.

7. Agregación con trabajos independientes

Una recopilación del Documento o de sus derivados con otros documentos o trabajos separados o independientes, en cualquier tipo de distribución o medio de almacenamiento, no como un todo, cuenta como una Versión Modificada del Documento, teniendo en cuenta que ninguna compilación de derechos de autor sea clamada por la recopilación. Tal recopilación es llamada un ”agregado”, y esta Licencia no aplica a los otros trabajos auto-contenidos y por lo tanto compilados con el Documento, o a cuenta de haber sido compilados, si no son ellos mismos trabajos derivados del Documento.

Si el requerimiento de la sección 3 del Texto de la Cubierta es aplicable a estas copias del Documento, entonces si el Documento es menor que un cuarto del agregado entero, Los Textos de la Cubierta del Documento pueden ser colocados en cubiertas que enmarquen solamente el Documento entre el agregado. De otra forma deben aparecer en cubiertas enmarcando todo el agregado.

8. Traducción

La Traducción es considerada como una clase de modificación, Así que puede distribuir traducciones del Documento bajo los términos de la sección 4. Reemplazar las Secciones Invariantes con traducciones requiere permiso especial de los dueños de derecho de autor, pero puede incluir traducciones de algunas o todas las Secciones Invariantes adicionalmente a las versiones originales de las Secciones Invariantes. Puede incluir una traducción de esta Licencia siempre que incluya también la versión Inglesa de esta Licencia. En caso de un desacuerdo entre la traducción y la versión original en Inglés de esta Licencia, la versión original en Inglés prevalecerá.

9. Terminación

No se puede copiar, modificar, sublicenciar, o distribuir el Documento excepto por lo permitido expresamente bajo esta Licencia. Cualquier otro intento de copia, modificación, sublicenciamiento o distribución del Documento es nulo, y serán automáticamente terminados sus derechos bajo esa licencia. De todas maneras, los terceros que hayan recibido copias, o derechos, de su parte bajo esta Licencia no tendrán por terminadas sus licencias siempre que tales personas o entidades se encuentren en total conformidad con la licencia original.

10 Futuras revisiones de esta licencia

La Free Software Foundation puede publicar nuevas, revisadas versiones de la Licencia de Documentación Libre GNU de tiempo en tiempo. Tales nuevas versiones serán similares en espíritu a la presente versión, pero pueden diferir en detalles para solucionar problemas o intereses. Vea `http://www.gnu.org/copyleft/`.

Cada versión de la Licencia tiene un número de versión que la distingue. Si el Documento especifica que una versión numerada particularmente de esta licencia o "cualquier versión posterior" se aplica a esta, tiene la opción de seguir los términos y condiciones de la versión especificada o cualquiera posterior que ha sido publicada(no como un borrador)por la Free Software Foundation. Si el Documento no especifica un número de versión de esta Licencia, puede escoger cualquier versión que haya sido publicada(no como un borrador) por la Free Software Foundation.