

# Tema 7

## Estadística

- Introducción.
- Estadística unidimensional. Cálculo de la media, varianza y desviación típica:
  - Con datos simples.
  - Con datos agrupados por frecuencias.
- Cálculo de probabilidades en una distribución  $N(\mu, \sigma)$
- Estadística bidimensional.
- Actividades.

## INTRODUCCIÓN

Este tema explica como usar la calculadora científica para aplicarla al estudio de Estadística ya que cuenta con herramientas para realizar diversos cálculos estadísticos de variables estadísticas unidimensionales y bidimensionales.

Además, la calculadora lleva incorporada funciones que permiten generar fácilmente la tabla de frecuencias de una distribución estadística.

## ESTADÍSTICA UNIDIMENSIONAL. Cálculo de la media, varianza y desviación típica

Recordemos que la media es una medida de tendencia central que se calcula dividiendo la suma de todos los valores de la variable por el número de valores del conjunto de datos. La desviación típica es una medida de dispersión que es la raíz cuadrada positiva de la varianza.

La interpretación de la media y la desviación típica para un determinado conjunto de datos es más fácil cuando la tecnología, como la calculadora, está presente; así los cálculos de estas medidas son más rápidos y más sistemáticos.

Preparamos la calculadora para que opere en modo estadístico.

Con **MODE** accedemos a la configuración para establecer distintos modos de trabajo.

```

1:COMP   2:CMPLX
3:STAT   4:BASE-N
5:EQN    6:MATRIX
7:TABLE  8:VECTOR

```

Pulsamos **3**: **STAT** (modo estadístico).

```

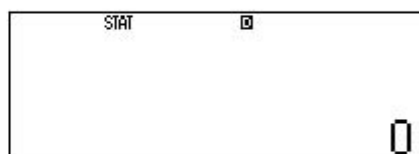
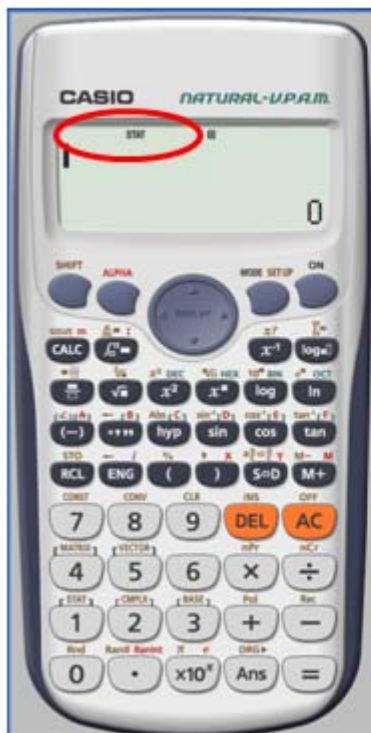
1:1-VAR  2:A+BX
3:_+CX2 4:ln X
5:eX    6:A·BX
7:A·XB 8:1/X

```

**1:1-VAR (una variable) 2:A+BX (dos variables, regresión lineal)**

Las demás opciones son para realizar otro tipo de regresiones: exponencial, logarítmica, potencial, cuadrática, inversa, etc.

Así debemos tener nuestra calculadora

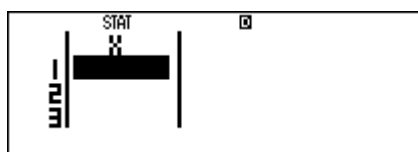
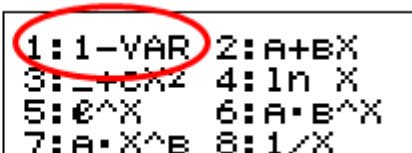


**Ejemplo 1. Calculo con datos simples**

Estas son las edades de los siete miembros de una familia. Calcula la media, varianza y desviación típica de esas edades.

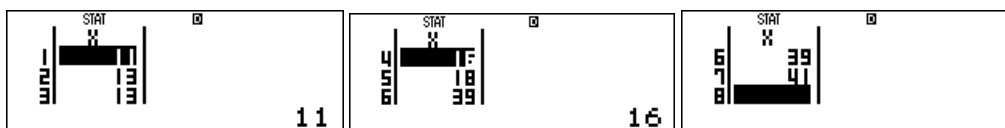
11 13 13 16 18 39 41

Pulsamos **MODE** **3** **1**



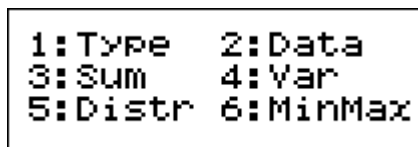
El editor **STAT** dispone de 80 filas para ingresar datos si hay una sola columna **X**

Introducimos los datos uno a uno, pulsando el signo  $\boxed{=}$

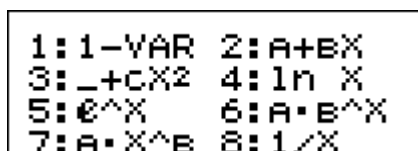


Una vez introducidos los datos pulsamos  $\boxed{AC}$  para salir de la tabla.

Pulsamos  $\boxed{SHIFT}$   $\boxed{1}$  para la edición de la tabla y calculo de parámetros.

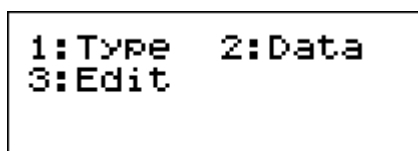


**1:Type.** Cambia el tipo de variable y vuelve a la pantalla:

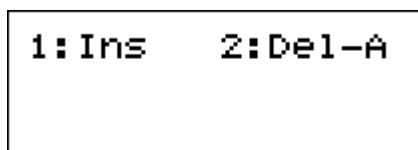


**2:Data.** Muestra la tabla de datos.

Si volvemos a pulsar  $\boxed{SHIFT}$   $\boxed{1}$  aparece un menú contextual



en el que al pulsar  $\boxed{3}$



tenemos la opción de  $\boxed{1}$ : **Insertar** una fila en cualquier sitio de la tabla o bien  $\boxed{2}$ : **Del -A**, borrar todo el contenido de la tabla.

**3:Sum.** Proporciona la suma de los cuadrados de los valores introducidos (1) y la suma de todos los valores (2).

1: $\Sigma x^2$	2: $\Sigma x$
-----------------	---------------

**4: Var.** Devuelve el número de valores (1), la media aritmética (2), la desviación típica tomada como población (3) y la desviación típica tomada como muestra (4).

1: n	2: $\bar{x}$
3: $\sigma_x$	4: $s_x$

**5: Distr.** Realiza cálculos con la distribución normal

1: P(	2: Q(
3: R(	4: $\mu t$

**6: MinMax.** Devuelve los valores mínimo (1) y máximo (2) de la distribución.

1: minX	2: maxX
---------	---------

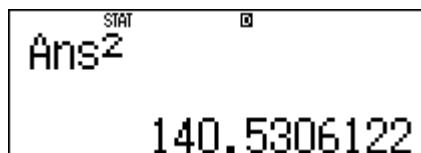
Para calcular la media aritmética pulsamos **SHIFT** **1** **4** **2** **=**

STAT	□
$\bar{x}$	
	21.57142857

Y para la desviación típica **SHIFT** **1** **4** **3** **=**

STAT	□
$\sigma_x$	
	11.85456082

Si queremos la varianza pulsamos  **$x^2$**  **=**



**Ejemplo 2. Calculo con datos estadísticos agrupados por frecuencias**

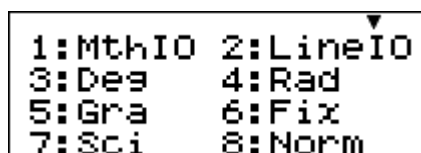
Al preguntar en 50 familias por el número de personas que forman el hogar familiar, hemos obtenido la información que se recoge en la siguiente tabla:


Nº de personas	1	2	3	4	5	6
Nº de familias	3	10	23	9	3	2

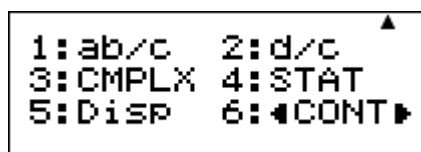
Calcula la media aritmética del número de personas que forman el hogar así como su desviación típica y su varianza.

Lo primero que hacemos es acceder a la configuración **SETUP** para activar la columna de frecuencias.

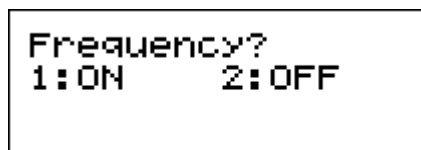
Para ello, pulsamos **SHIFT** **MODE** y cuando aparezca la pantalla:



pulsamos el cursor que marca la dirección hacia abajo  para que aparezcan más opciones:



A continuación pulsamos en la opción **4** accediendo a la pantalla:



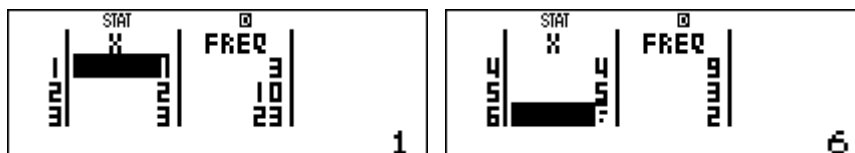
donde pulsaremos **1: ON**, para que se active la columna de frecuencias. Si queremos desactivarla, seguimos los mismos pasos y pulsamos **2:OFF**.

Para realizar el ejercicio, seguimos los pasos indicados a continuación.

**SHIFT** **1** **2** o bien **MODE** **3** **1**

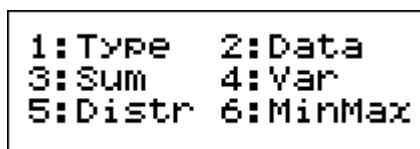


Introducimos los valores de la variable y de sus frecuencias por columnas, moviéndonos con la tecla de cursor.

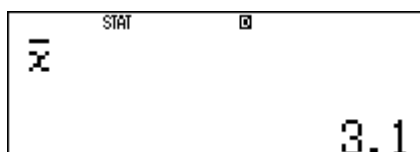


Una vez introducidos los datos pulsamos **AC** para salir de la tabla.

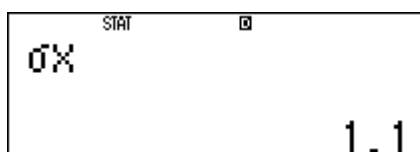
A continuación pulsamos **SHIFT** **1** para la edición de la tabla y calculo de parámetros.



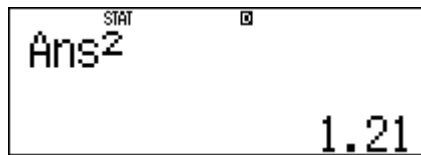
Para calcular la media aritmética pulsamos **SHIFT** **1** **4** **2** **=**



Y para la desviación típica **SHIFT** **1** **4** **3** **=**



Por último, si queremos la varianza pulsamos las teclas  $\boxed{x^2}$   $\boxed{=}$



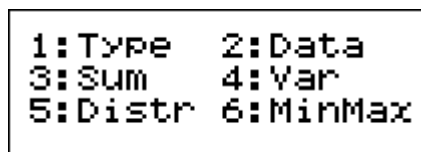
**Nota:** Si los datos están agrupados en intervalos se toma como valor de la variable  $x_i$  la marca de clase o punto medio del intervalo.

### CÁLCULO DE PROBABILIDADES EN UNA DISTRIBUCIÓN $N(\mu, \sigma)$

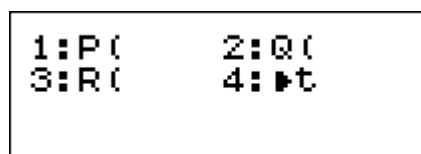
**Ejemplo 3.** En una  $N(6,4)$ , calcular las probabilidades siguientes:

- a)  $P[x \leq 3]$    b)  $P[x \geq 12]$    c)  $P[5 \leq x \leq 8]$

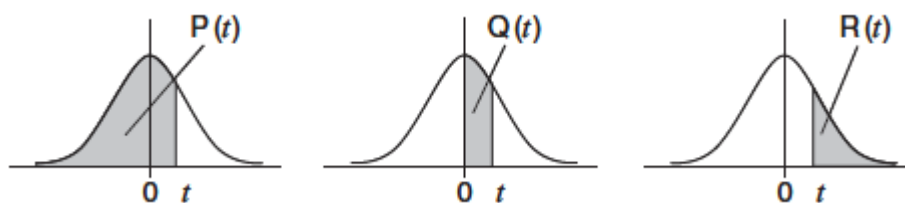
Entramos en el modo **STAT** :  $\boxed{\text{MODE}}$   $\boxed{3}$



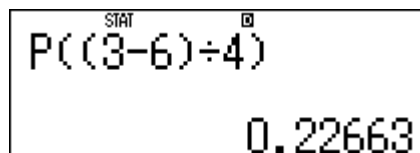
Luego pulsamos  $\boxed{\text{SHIFT}}$   $\boxed{1}$  y a



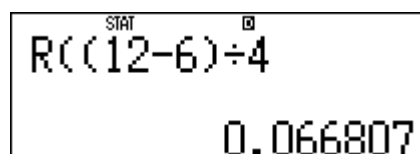
continuación **5: Distr** donde



a)  $\boxed{\text{SHIFT}}$   $\boxed{1}$   $\boxed{5}$   $\boxed{1}$



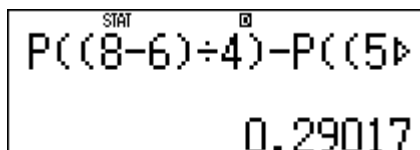
b)  $\boxed{\text{SHIFT}}$   $\boxed{1}$   $\boxed{5}$   $\boxed{3}$





c)

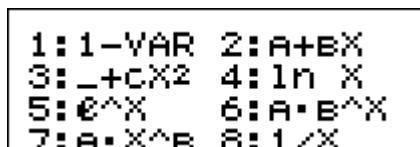
SHIFT 1 5 1 ( 8 - 6 ) ÷ 4 ) - SHIFT 1 5 1  
( 5 - 6 ) ÷ 4 ) =



P((8-6)÷4) - P((5-6)÷4)  
 0.29017

## ESTADÍSTICA BIDIMENSIONAL

La calculadora lleva incorporadas herramientas que permiten generar fácilmente la tabla de frecuencia absolutas de una distribución estadística bidimensional, así como estudiar la correlación entre variables. También es posible ver los distintos modelos de regresión para ajustar los tipos de correlación como son:



1: 1-VAR    2: A+BX  
 3: \_+CX<sup>2</sup>    4: ln X  
 5: e<sup>X</sup>    6: A·B<sup>X</sup>  
 7: A·X<sup>B</sup>    8: 1/X

- 2: Regresión lineal
- 3: Regresión cuadrática
- 4: Regresión logarítmica
- 5: Regresión exponencial
- 6: Regresión exponencial
- 7: Regresión potencial

Recordemos que al estudiar distribuciones bidimensionales el objetivo perseguido es determinar si existe relación estadística entre las dos variables consideradas; es decir, ver si los cambios en una de las variables influyen en los cambios de la otra. Cuando sucede ésto, diremos que ambas variables están correlacionadas o que hay correlación entre ellas.

Si las variables crecen conjuntamente, la correlación es directa. Si, por el contrario, al aumentar una de ellas disminuye la otra, la correlación será

inversa. La correlación puede calificarse como fuerte cuando el grado de dependencia es alto; y como débil en caso contrario.

En los cálculos que podemos hacer con esta calculadora no aparece ningún parámetro estadístico conjunto como la covarianza, que recordemos es la media aritmética de los productos de las diferencias de los valores de cada variable respecto de su media marginal; esto no tiene importancia ya que se puede calcular, y sobre todo porque la covarianza no da una medida objetiva (comparable) de la correlación entre variables.

El criterio que se utiliza para medir la fuerza de la correlación entre dos variables es el coeficiente de correlación lineal,  $r$  que es la razón entre la covarianza de las variables  $X$  e  $Y$  y el producto de sus desviaciones típicas marginales. Recordemos alguna de sus propiedades fundamentales:

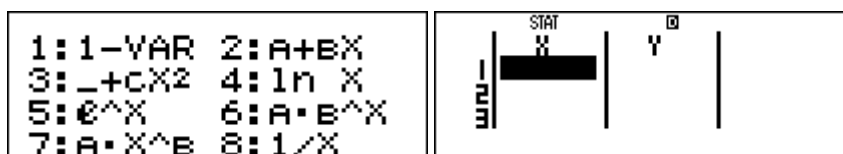
- El valor de  $r$  no cambia al hacerlo la escala de medición.
- Si  $r > 0$ , la correlación es directa.
- Si  $r < 0$ , la correlación es inversa.
- El valor de  $r$  está entre  $-1$  y  $+1$ .
- Si  $r$  toma valores cercanos a  $-1$  la correlación es fuerte (e inversa).
- Si  $r$  toma valores cercanos a  $+1$  la correlación es fuerte (y directa).
- Si  $|r|=1$ , la correlación es perfecta. Hay dependencia lineal entre las variables  $X$  e  $Y$ .
- Si  $r$  toma valores cercanos a  $0$ , la correlación es débil.

**Ejemplo 4.** *En una empresa se ha estudiado la relación entre los gastos (en millones de €) y los ingresos (también en millones de €). Los datos se resumen en la tabla. Halla las medias y desviaciones típicas de las dos variables. Calcula la recta de regresión de los ingresos en función de los gastos y halla el coeficiente de correlación lineal. Si la empresa gasta 6 millones de euros, ¿cuánto debería de ingresar?*

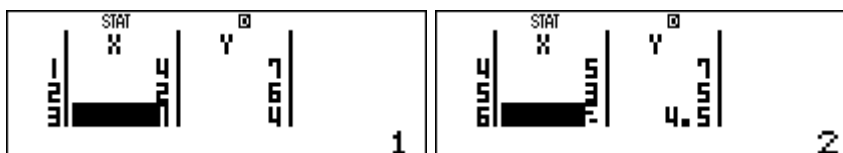
X: Gastos	4	2	1	5	3	2
Y: Ingresos	7	6	4	7	5	4.5

Para la realización de este ejercicio las frecuencias deben estar desactivadas.

En la pantalla del modo estadístico pulsamos **2: A+BX**

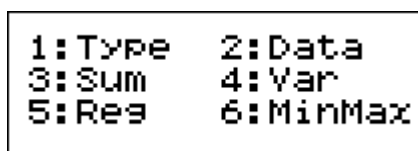


Introducimos los datos de la variable X e Y por columnas y con ayuda de los cursores.



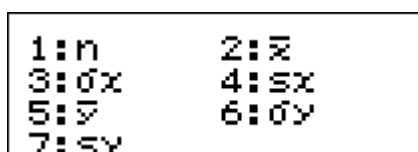
Una vez introducidos los datos pulsamos **AC** para salir de la tabla.

Pulsamos **SHIFT 1** para la edición de la tabla y calculo de parámetros.



Observamos que aparece en pantalla nuevos datos **5:Reg** (para hacer el análisis de regresión).

Para calcular las medias aritméticas así como las desviaciones típicas de las variables X e Y pulsamos: **SHIFT 1 4**



Vamos a hacer el análisis de regresión; en la pantalla principal pulsamos

**SHIFT** **1** **5**

1:A	2:B
3:r	4: $\hat{x}$
5: $\hat{y}$	

La recta de regresión es  $y=A+BX$ .

**1:A** (Coeficiente A de la recta de regresión).

**2:B** (Coeficiente B de la recta de regresión).

**3:r** (Coeficiente de correlación lineal).

**4:**  $\hat{x}$  (estimación del valor de "x" conocido la "y").

**5:**  $\hat{y}$  (estimación del valor de "y" conocido la "x").

**Nota:** La estimación del valor de "x" conocido la "y" no es un buen valor a utilizar pues sustituye la "y" en la recta de regresión  $y=A+BX$ , cuando debería sustituirla en  $x=A'+B'Y$ .

Para calcular A pulsamos **SHIFT** **1** **5** **1** **=**

STAT	□
A	
	3.469230769

Para calcular B pulsamos **SHIFT** **1** **5** **2** **=**

STAT	□
B	
	0.7461538462

Por tanto la recta de regresión será:  $y = 0.75x + 3.47$

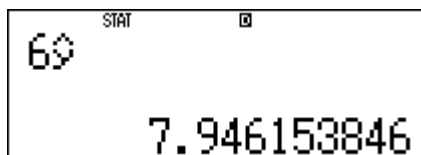
Vamos a calcular ahora el coeficiente de regresión r. Para ello pulsamos

**SHIFT** **1** **5** **3** **=**

STAT	□
r	
	0.8571995667

Observamos que la correlación es fuerte y directa.

Si ahora pulsamos **AC**, escribimos el valor 6 y volvemos a entrar en el cuadro de Regresión **6** **SHIFT** **1** **5** **5** **=** nos estima el valor buscado.



Por tanto, si  $x=6 \Rightarrow y=7,95$

### ACTIVIDADES

1. Las calificaciones de 90 opositores en el primer ejercicio han sido:

$x_i$	0	1	2	3	4	5	6	7	8	9	10
$f_i$	4	10	13	11	13	10	9	7	7	4	2

- a) Calcula la nota media de dichos opositores.
- b) ¿Cuál es su desviación típica?

2. Un empresario desea repartir unas bonificaciones entre sus empleados en base a la categoría y productividad de los mismos. Dicha distribución quedó de la siguiente forma:

Bonificaciones (cientos de euros)	10-15	15-20	20-25	25-30	30-35	35-40
Nº Empleados	3	8	12	15	7	5

- a) Calcula la bonificación media.
- b) Halla la desviación típica y varianza de dicha actividad.

3. A unas oposiciones de auxiliar administrativo se han presentado un total de 1000 aspirantes. Éstos han realizado un primer ejercicio que se ha calificado en una escala entera de 0 a 100 puntos. Los resultados se muestran en la siguiente tabla:

Intervalo	x	f
[40,45)	42.5	1
[45,50)	47.5	2
[50,55)	52.5	10
[55,60)	57.5	46
[60,65)	62.5	120
[65,70)	67.5	202
[70,75)	72.5	242
[75,80)	77.5	205
[80,85)	82.5	113
[85,90)	87.5	45
[90,95)	92.5	12
[95,100)	97.5	2

Construye la tabla con la frecuencia y calcula la media aritmética, la varianza y la desviación típica.

4. Una zapatería ha vendido en un día 25 pares de zapatos de caballero de las siguientes tallas:

40	41	40	42	40
41	43	43	45	44
45	42	40	41	45
45	44	40	44	41
42	43	43	46	41

Construye una tabla con la frecuencia y halla la media aritmética, la varianza y la desviación típica.

5. El coeficiente de variación de una distribución,  $CV$ , es el cociente entre la desviación típica y la media aritmética.

La siguiente tabla muestra las calificaciones obtenidas por Paco y Eva en diez controles de matemáticas:

Notas de Paco	4	5	5	4	6	7	8	9	3	9
Notas de Eva	6	6	5	6	7	7	6	5	7	5

Halla sus medias y desviaciones típicas. ¿Quién es más regular?

6. Las puntuaciones obtenidas por 30 alumnos de 4° de ESO en una prueba de inteligencia han sido

100	102	98	95	92	105	121	110	84	87
94	99	98	112	123	145	116	93	89	85
86	97	114	127	103	104	135	128	109	110

Agrupar los datos en intervalos de clase y halle la media aritmética, varianza y la desviación típica.

7.- De la distribución de una variable sabemos que el recorrido es 60, y la distribución está dividida en seis intervalos de amplitud constante. Las frecuencias de cada intervalo son, por orden, 7, 11, 15, 10, 5, 2. La media aritmética es  $35,2$ . Una vez hallada la distribución (su tabla con la frecuencia), halla la media aritmética, la varianza y la desviación típica.

8.- En una región de España se observó el precio del vino y la cantidad de producción durante algunos años, obteniéndose los siguientes datos:

X	35	31	42	60	52	49	61	50	55	58
Y	100	140	120	110	200	200	110	160	160	200

donde X es el precio por litro del vino en céntimos de Euro e Y es la cantidad producida en miles de litros.

- Calcula la media aritmética de las dos variables, así como su desviación típica.
- Halla la recta de regresión de Y sobre X y el coeficiente de correlación. Interpreta los resultados.
- Si el precio del vino es de 45 céntimos de euro, ¿Cuánto se estima que habrá de producción?

9. En una competición de patinaje artístico por parejas se otorgan dos notas: una a los ejercicios obligatorios (X) y otra a los ejercicios libres (Y).

Las seis parejas que se disputan la final han obtenido los siguientes resultados:

X	5	5	6	7	7	7
Y	5	7	7	7	7	8

- ¿Qué tipo de correlación hay entre las variables?
- Calcula el coeficiente de regresión.
- Ensayá los diferentes tipos de regresión que hemos visto en el tema.

10. En la siguiente tabla se dan las edades  $x_i$  y las presiones sanguíneas  $y_i$  de las 12 personas de una muestra estadística extraída de un cierto colectivo.

$x_i$	58	42	50	44	33	37	63	55	51	37	67	31
$y_i$	144	123	145	140	110	135	147	150	142	120	155	113



Halla las medias y desviaciones típicas de las dos variables.

Calcula la recta de regresión de la presión en función de la edad y halla el coeficiente de correlación lineal. Haz un comentario al respecto.

Estima la presión sanguínea de una persona de 40 años del citado colectivo.

11.- La estatura de los individuos de una población se distribuyen normalmente con media 175 cm y desviación típica 10 cm. Calcula la probabilidad de que:

a) Calcula la probabilidad de que un individuo tenga una estatura mayor que 180 cm.

b) Calcula la probabilidad de que un individuo tenga una estatura menor que 170 cm.

c) ¿Qué proporción de individuos tiene una estatura comprendida entre 170 cm y 180 cm?