

# **TALLER DE ESTADÍSTICA**

## **1. ESTADÍSTICA DESCRIPTIVA.**

---

**MAURICIO CONTRERAS**

## EL CURRÍCULUM DE ESTADÍSTICA EN ESO Y BACHILLERATO

### CARACTERÍSTICAS GENERALES DEL NUEVO

#### CURRÍCULUM DE E.S.O. (DECRETO 112/2007, de 20 de julio (2007/9717))

- OBJETIVOS

2. Aplicar con soltura y adecuadamente las herramientas matemáticas adquiridas a situaciones de la vida diaria.

4. Detectar los aspectos de la realidad que sean cuantificables y que permitan interpretarla mejor: utilizar técnicas de recogida de la información y procedimientos de medida y realizar el análisis de los datos mediante el uso de distintas clases de números y la selección de los cálculos apropiados, todo ello de la forma más adecuada, según la situación planteada.

5. Identificar los elementos matemáticos (datos estadísticos, geométricos, gráficos, cálculos, etc.) presentes en los medios de comunicación, Internet, publicidad u otras fuentes de información, analizar críticamente las funciones que desempeñan estos elementos matemáticos y valorar su aportación para una mejor comprensión de los mensajes.

7. Utilizar de forma adecuada los distintos medios tecnológicos (calculadoras, ordenadores, etc.) tanto para realizar cálculos como para buscar, tratar y representar informaciones de índole diversa y también como ayuda en el aprendizaje.

12. Valorar las Matemáticas como parte integrante de nuestra cultura: tanto desde un punto de vista histórico como desde la perspectiva de su papel en la sociedad actual y aplicar las competencias matemáticas adquiridas para analizar y valorar fenómenos sociales como la diversidad cultural, el respeto al medio ambiente, la salud, el consumo, la igualdad entre los sexos o la convivencia pacífica.

ESTOS OBJETIVOS HACEN NECESARIO INTRODUCIR EN LAS AULAS EL ESTUDIO DE CONCEPTOS, PROCEDIMIENTOS Y TÉCNICAS ESTADÍSTICAS Y PROBABILÍSTICAS PARA EL TRATAMIENTO DE LA INFORMACIÓN.

EL OBJETIVO 7 INDICA LA NECESARIA INTRODUCCIÓN DE LAS NUEVAS TECNOLOGÍAS EN LA PRÁCTICA COTIDIANA DE LAS CLASES, COMO INSTRUMENTOS QUE FACILITAN EL APRENDIZAJE.

**LOS CONTENIDOS DEL PRIMER CICLO DE E.S.O.****PRIMER CURSO****Bloque 6. Estadística y probabilidad**

- Diferentes formas de recogida de información. Organización en tablas de datos recogidos en una experiencia. Frecuencias absolutas y relativas.
- Diagramas de barras, de líneas y de sectores. Análisis de los aspectos más destacables de los gráficos estadísticos.
- Formulación de conjeturas sobre el comportamiento de fenómenos aleatorios sencillos y comprobación mediante la realización de experiencias repetidas.
- Reconocimiento y valoración de las matemáticas para interpretar y describir situaciones inciertas.

**Criterios de evaluación**

12. Organizar e interpretar informaciones diversas mediante tablas y gráficas, e identificar relaciones de dependencia en situaciones cotidianas.
13. Hacer predicciones sobre la posibilidad de que un suceso ocurra a partir de información previamente obtenida de forma empírica.

**SEGUNDO CURSO****Bloque 6. Estadística y probabilidad**

- Organización de los datos. Estadística unidimensional. Población y muestra. Distribuciones discretas. Recuento de datos.
- Frecuencia absoluta y relativa. Frecuencias acumuladas.
- Construcción e interpretación de tablas de frecuencias y diagramas de barras y de sectores. Análisis de los aspectos más destacables de los gráficos estadísticos.
- Cálculo e interpretación de la media aritmética, la mediana y la moda de una distribución discreta con pocos datos.
- Utilización conjunta de la media, la mediana y la moda para realizar comparaciones y valoraciones.
- Utilización de la hoja de cálculo para organizar los datos, realizar los cálculos y generar los gráficos más adecuados.

**Criterios de evaluación**

14. Formular las preguntas adecuadas para conocer las características de una población y recoger, organizar y presentar datos relevantes para responderlas, utilizando los métodos estadísticos apropiados y las herramientas informáticas adecuadas.
15. Obtener e interpretar la tabla de frecuencias y el diagrama de barras o de sectores, así como la moda y la media aritmética, de una distribución discreta sencilla, con pocos datos, sirviéndose, si es preciso, de una calculadora de operaciones básicas.

**TERCER CURSO****Bloque 6. Estadística y probabilidad**

- Estadística descriptiva unidimensional. Variables discretas y continuas.
- Interpretación de tablas de frecuencias y gráficos estadísticos.
- Agrupación de datos en intervalos. Histogramas y polígonos de frecuencias.
- Construcción de la gráfica adecuada a la naturaleza de los datos y al objetivo deseado.
- Cálculo e interpretación de los parámetros de centralización (media, moda, cuartiles y mediana) y dispersión (rango y desviación típica).
- Interpretación conjunta de la media y la desviación típica.
- Utilización de las medidas de centralización y dispersión para realizar comparaciones y valoraciones. Análisis y crítica de la información de índole estadístico y de su presentación.
- Utilización de la calculadora y la hoja de cálculo para organizar los datos y realizar cálculos.
- Experimentos aleatorios. Sucesos y espacio muestral. Utilización del vocabulario adecuado para describir y cuantificar situaciones relacionadas con el azar.
- Frecuencia y probabilidad de un suceso.
- Cálculo de probabilidades mediante la Ley de Laplace.
- Cálculo de la probabilidad mediante simulación o experimentación.
- Formulación y verificación de conjeturas sobre el comportamiento de fenómenos aleatorios sencillos.
- Utilización de la probabilidad para tomar decisiones fundamentadas en diferentes contextos.
- Reconocimiento y valoración de las Matemáticas para interpretar, describir y predecir situaciones inciertas.

**Criterios de evaluación**

1. Elaborar e interpretar tablas y gráficos estadísticos (diagramas de barras o de sectores, histogramas, etc.), así como los parámetros estadísticos más usuales (media, moda, mediana y desviación típica), correspondientes a distribuciones sencillas y utilizar, si es necesario, una calculadora científica.
2. Hacer predicciones cualitativas y cuantitativas sobre la posibilidad de que un suceso ocurra a partir de información previamente obtenida de forma empírica o como resultado del recuento de posibilidades, en casos sencillos.
3. Determinar e interpretar el espacio muestral y los sucesos asociados a un experimento aleatorio sencillo y asignar probabilidades en situaciones experimentales equiprobables, utilizando adecuadamente la Ley de Laplace y los diagramas de árbol.

**CUARTO CURSO OPCIÓN A****Bloque 6. Estadística y probabilidad**

- Estadística descriptiva unidimensional. Identificación de las fases y tareas de un estudio estadístico a partir de situaciones concretas cercanas a la alumna y al alumno.
- Análisis elemental de la representatividad de las muestras estadísticas.
- Variable discreta: elaboración e interpretación de tablas de frecuencias y de gráficos estadísticos: gráficos de barras, de sectores, diagramas de caja y polígonos de frecuencias. Uso de la hoja de cálculo.
- Cálculo e interpretación de los parámetros de centralización y dispersión para realizar comparaciones y valoraciones.
- Variable continua: intervalos y marcas de clase. Elaboración e interpretación de histogramas. Uso de la hoja de cálculo.
- Azar y probabilidad. Idea de experimento aleatorio y suceso. Frecuencia y probabilidad de un suceso.
- Experiencias compuestas. Utilización de tablas de contingencia y diagramas de árbol para la asignación de probabilidades.
- Utilización del vocabulario adecuado para describir y cuantificar situaciones relacionadas con el azar.

**Criterios de evaluación**

14. Elaborar e interpretar tablas y gráficos estadísticos, así como los parámetros estadísticos más usuales, correspondientes a distribuciones discretas y continuas, y valorar cualitativamente la representatividad de las muestras utilizadas.
15. Aplicar los conceptos y técnicas de cálculo de probabilidades para resolver diferentes situaciones y problemas de la vida cotidiana.

**CUARTO CURSO OPCIÓN B****Bloque 6. Estadística y probabilidad**

- Estadística descriptiva unidimensional. Identificación de las fases y tareas de un estudio estadístico.
- Análisis elemental de la representatividad de las muestras estadísticas.
- Variable discreta: elaboración e interpretación de tablas de frecuencias y de gráficos estadísticos: gráficos de barras, de sectores, diagramas de caja y polígonos de frecuencias.
- Cálculo e interpretación de los parámetros de centralización y dispersión: media, mediana, moda, recorrido y desviación típica para realizar comparaciones y valoraciones.
- Representatividad de una distribución por su media y desviación típica o por otras medidas ante la presencia de descentralizaciones, asimetrías y valores atípicos. Valoración de la mejor representatividad, en función de la existencia o no de valores atípicos.
- Variable continua: intervalos y marcas de clase. Elaboración e interpretación de histogramas.
- Análisis crítico de tablas y gráficas estadísticas en los medios de comunicación, Detección de falacias.
- Experimentos aleatorios. Espacio muestral asociado a un experimento aleatorio. Sucesos.
- Técnicas de recuento. Introducción a la combinatoria: combinaciones, variaciones y permutaciones. Aplicación al cálculo de probabilidades.
- Experiencias compuestas. Utilización de tablas de contingencia y diagramas de árbol para la asignación de probabilidades.
- Probabilidad condicionada.
- Utilización del vocabulario adecuado para describir y cuantificar situaciones relacionadas con el azar.

**Criterios de evaluación**

14. Elaborar e interpretar tablas y gráficos estadísticos, así como los parámetros estadísticos más usuales en distribuciones unidimensionales y valorar cualitativamente la representatividad de las muestras utilizadas.
15. Determinar e interpretar el espacio muestral y los sucesos asociados a un experimento aleatorio, simple o compuesto, y utilizar la Ley de Laplace, los diagramas de árbol, las tablas de contingencia u otras técnicas combinatorias para calcular probabilidades simples o compuestas.
16. Aplicar los conceptos y técnicas de cálculo de probabilidades para resolver diferentes situaciones y problemas de la vida cotidiana.

**MATEMÁTICAS APLICADAS A LAS CIENCIAS SOCIALES I y II**  
**(Bachillerato de Humanidades y Ciencias Sociales)**

**Objetivos generales**

4. Utilizar los conocimientos matemáticos adquiridos para interpretar críticamente los mensajes, datos e informaciones que aparecen en los medios de comunicación y otros ámbitos sobre cuestiones económicas y sociales de la actualidad.
5. Elaborar juicios y formar criterios propios sobre fenómenos sociales y económicos mediante actitudes propias de la actividad matemática como son la visión crítica, la necesidad de verificación, la justificación de las afirmaciones, la valoración de la precisión, el gusto por el rigor, la necesidad de cuestionar las apreciaciones intuitivas y la apertura a nuevas ideas.
6. Establecer relaciones entre las matemáticas y el entorno social, cultural y económico, apreciando su lugar como parte de nuestra cultura.
7. Servirse de los medios tecnológicos que se muestran a su disposición, haciendo un uso racional de ellos y descubriendo las enormes posibilidades que nos ofrecen.
8. Aprovechar los cauces de información facilitados por las nuevas tecnologías, seleccionando aquello que pueda ser más útil para resolver los problemas planteados.

EL CUMPLIMIENTO DE LOS OBJETIVOS 4, 5 y 6 PUEDE SER FAVORECIDO POR EL ESTUDIO DE LOS CONCEPTOS, PROCEDIMIENTOS Y TÉCNICAS ESTADÍSTICAS.

EL CUMPLIMIENTO DE LOS OBJETIVOS 7, 8 Y 10 EXIGE EL USO REITERADO DE NUEVAS TECNOLOGÍAS EN LA CLASE DE MATEMÁTICAS O QUE, AL MENOS, LOS ESTUDIANTES EXPERIMENTEN CON HERRAMIENTAS TECNOLÓGICAS APROPIADAS

**MODIFICACIONES CURRICULARES EN EL BACHILLERATO DE**  
**HUMANIDADES Y CIENCIAS SOCIALES**

**MATEMÁTICAS APLICADAS A LAS CIENCIAS SOCIALES I**

4. Estadística.
  - Terminología y conceptos básicos de la Estadística: individuo, población, muestra, variable estadística. Organización de los datos: gráficos y tablas de frecuencias. Distribución de frecuencias. Parámetros estadísticos. Significado y cálculo.
  - Estadística bidimensional. Elaboración e interpretación de tablas de frecuencias de doble entrada y nubes de puntos.
  - Cálculo e interpretación de los parámetros estadísticos bidimensionales usuales.
  - Coeficiente de correlación lineal. Interpretación y cálculo.
  - Regresión lineal. Rectas de regresión. Utilización de las rectas de regresión para interpolar. Predicciones estadísticas.

#### 5. Probabilidad.

- Medida de la incertidumbre. Asignación de probabilidades. Leyes de la probabilidad. Experiencias aleatorias compuestas. Tablas de contingencia y diagramas de árbol. Probabilidad condicionada. Probabilidad total. Probabilidad a posteriori.
- Distribuciones de probabilidad binomial y normal. Utilización de tablas de la distribución binomial y de la distribución normal en la resolución de problemas que requieran cálculos probabilísticos.

#### **Criterios de evaluación**

6. Interpretar el grado de correlación existente entre las variables de una distribución estadística bidimensional y obtener las rectas de regresión para hacer predicciones estadísticas en un contexto de resolución de problemas relacionados con fenómenos económicos o sociales.
7. Utilizar técnicas estadísticas elementales para tomar decisiones ante situaciones que se ajustan a una distribución de probabilidad binomial o normal, determinando las probabilidades de uno o varios sucesos, sin necesidad de cálculos combinatorios.

- LOS CONTENIDOS DE ESTE CURSO NO HAN EXPERIMENTADO CAMBIOS SIGNIFICATIVOS.

### **MATEMÁTICAS APLICADAS A LAS CIENCIAS SOCIALES II**

#### 4. Estadística y Probabilidad.

- Experimentos aleatorios. Sucesos. Operaciones con sucesos.
- Profundización en las leyes de la probabilidad. Probabilidad condicionada. Probabilidad total.
- Técnicas de muestreo. Parámetros de una población. Distribución de probabilidad de la media muestral. Teorema central del límite.
- Intervalo de confianza de la media de la población. Nivel de confianza.

#### **Criterios de evaluación**

6. Asignar probabilidades a sucesos aleatorios simples y compuestos, dependientes e independientes, relacionadas con fenómenos sociales o naturales e interpretarlas; utilizar técnicas de conteo directo, diagramas de árbol, cálculos simples o tablas de contingencia.
7. Planificar y realizar estudios concretos de una población, a partir de una muestra bien seleccionada, asignar un nivel de significación, para inferir sobre la media poblacional y estimar el error cometido.
8. Analizar de forma crítica informes estadísticos presentes en los medios de comunicación y otros ámbitos, y detectar posibles errores y manipulaciones en la presentación de determinados datos.

LA INFERENCIA ESTADÍSTICA QUEDA REDUCIDA A LA ESTIMACIÓN DE PARÁMETROS POR INTERVALOS DE CONFIANZA Y DESAPARECE EL CONTRASTE DE HIPÓTESIS.

## **MODIFICACIONES CURRICULARES EN EL BACHILLERATO CIENTÍFICO**

### **Objetivos generales**

2. **Aplicar los conocimientos matemáticos para plantear y resolver problemas en diversas situaciones de la actividad cotidiana, científica y tecnológica.**
6. **Servirse de los medios tecnológicos que se muestran a su disposición, apreciando las ventajas y las limitaciones que comporta su uso, seleccionando aquello que pueda ser más útil para resolver los problemas planteados y descubriendo las enormes posibilidades que nos ofrecen a la hora de realizar investigaciones o ejecutar cálculos o resolver problemas.**

EL CUMPLIMIENTO DEL OBJETIVO 2 PUEDE SER FAVORECIDO POR EL ESTUDIO DE LOS CONCEPTOS, PROCEDIMIENTOS Y TÉCNICAS ESTADÍSTICAS.

EL CUMPLIMIENTO DEL OBJETIVO 6 EXIGE EL USO REITERADO DE NUEVAS TECNOLOGÍAS EN LA CLASE DE MATEMÁTICAS O QUE, AL MENOS, LOS ESTUDIANTES EXPERIMENTEN CON HERRAMIENTAS TECNOLÓGICAS APROPIADAS

### **MATEMÁTICAS I**

#### **V. Estadística y Probabilidad.**

- Terminología y conceptos básicos de la Estadística. Conceptos básicos en el tratamiento de datos muestrales. Distribuciones unidimensionales. Medida de la dispersión.
- Estadística descriptiva bidimensional. Relaciones entre dos variables estadísticas. El coeficiente de correlación lineal. Regresión lineal. Rectas de regresión.
- Aplicaciones de las rectas de regresión a la resolución de problemas. Interpolación y predicción en las distribuciones estadísticas bidimensionales.
- Terminología y conceptos básicos de la Probabilidad.
- Medida de la incertidumbre. Asignación de probabilidades.
- Experiencias aleatorias compuestas. Independencia de sucesos.
- Tablas de contingencia. Diagramas de árbol.
- Leyes de la probabilidad. Probabilidad condicionada. Probabilidad total y a posteriori.
- Distribuciones de probabilidad binomial y normal. Utilización de tablas de la distribución binomial y de la distribución normal en la resolución de problemas de cálculo probabilístico.

#### **Criterios de evaluación**

7. Interpretar el grado de correlación existente entre las variables de una distribución estadística bidimensional sencilla y obtener las rectas de regresión para hacer predicciones estadísticas.
8. Utilizar técnicas estadísticas elementales para tomar decisiones ante situaciones que se ajusten a una distribución binomial o normal, calculando las probabilidades de uno o varios sucesos.

- NO SE INCLUYE LA COMBINATORIA, PERO EN EL BLOQUE DE ÁLGEBRA SE ESTUDIAN LOS NÚMEROS COMBINATORIOS Y EL BINOMIO DE NEWTON.
- EN ESTADÍSTICA DESCRIPTIVA, EL FOCO DE ATENCIÓN ES EL ESTUDIO DE LOS PARÁMETROS DE DISPERSIÓN.
- SE ESTUDIAN VARIABLES BIDIMENSIONALES, INCLUYENDO AJUSTES POR RECTAS DE REGRESIÓN, PREDICCIONES E INTERPOLACIONES.
- SE ESTUDIA LA ASIGNACIÓN DE PROBABILIDADES EN EXPERIMENTOS COMPUESTOS MEDIANTE DIAGRAMAS DE ÁRBOL Y TABLAS DE CONTINGENCIA.
- SE INCLUYE EL ESTUDIO DE LA PROBABILIDAD CONDICIONADA, TOTAL Y A POSTERIORI.
- SE ESTUDIA LA RESOLUCIÓN DE PROBLEMAS UTILIZANDO LOS MODELOS PROBABILÍSTICOS BINOMIAL Y NORMAL.
- EN 2º CURSO DESAPARECE EL BLOQUE DE ESTADÍSTICA Y PROBABILIDAD, QUE PASA ÍNTEGRAMENTE A 1º. ES DECIR, SÓLO HAY ESTADÍSTICA EN 1º CURSO.



# GENERACIÓN DE DATOS ESTADÍSTICOS A PARTIR DE DATOS Y RULETAS

## Introducción

Podemos obtener series de datos estadísticos con generadores aleatorios, tales como dados (cúbicos, poliédricos, con las caras en blanco, simuladores de quinielas, etc) y ruletas de distintos tipos (estáticas y dinámicas, con sectores de la misma o distinta amplitud, etc). Posteriormente, representaremos los datos obtenidos usando diagramas de distintos tipos y obtendremos parámetros estadísticos que permitirán resolver algunos problemas concretos relacionados con el azar. Esta fase experimental es fundamental en el trabajo estadístico, ya que pone de manifiesto algunos problemas interesantes, como por ejemplo, recogida de la información, técnicas de agrupación y recuento, tabulación, etc. En las siguientes actividades tenemos algunos ejemplos de este tipo que se han experimentado en la ESO.

## 1. Lanzamiento de dados

- **Efectúa 120 lanzamientos de un dado cúbico y anota la cifra que aparece en la cara superior.**

Por ejemplo, nosotros hemos efectuado una serie de 120 lanzamientos y hemos obtenido los siguientes resultados:

3	6	4	1	4	4	2	5	4	6	5	2	2	4	5	2	3	3	3	3
3	6	4	1	3	4	6	1	3	4	4	1	3	5	3	6	5	6	6	4
1	5	5	5	5	5	2	3	6	2	5	3	5	6	5	2	5	1	1	3
3	1	5	5	3	6	4	5	1	2	6	2	2	6	5	6	3	1	4	1
6	2	4	5	4	4	4	1	2	2	1	2	1	4	6	6	1	3	2	4
4	1	2	5	6	1	3	5	3	1	4	2	1	3	4	6	6	1	3	4

Como puedes observar, los datos están desordenados y es difícil trabajar con ellos. Por ello hemos de contar cuántas veces ha salido cada uno y presentar los datos en una tabla. Podemos hacer el recuento de unos de la siguiente forma:

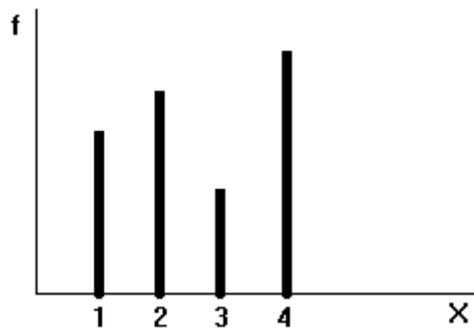
Resultados	Recuento	Frecuencia absoluta
1	☑ ☑ ☑ ☑	20
2		
3		
4		
5		
6		

- **Completa el recuento con los datos obtenidos entre toda la clase. Con los datos obtenidos construye la siguiente tabla:**

Resultado	Frecuencia absoluta f	$\% = \frac{f}{N} \times 100$	Frecuencia relativa $f_r = \frac{f}{N}$
1			
2			
3			
4			
5			
6			
Totales	N=		

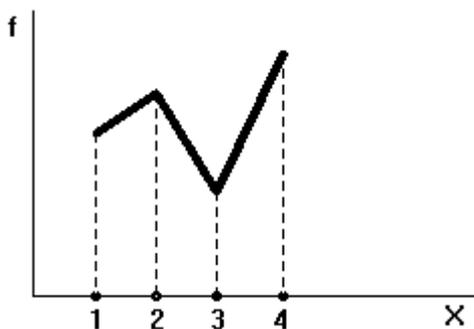
- **¿Cuál es el total de la columna % de porcentajes?. ¿Y de la columna de frecuencias relativas?.**

Los datos de la columna de frecuencias absolutas se pueden representar en un diagrama de barras como el siguiente:



- **Dibuja el diagrama de barras correspondiente a los datos de tu clase.**

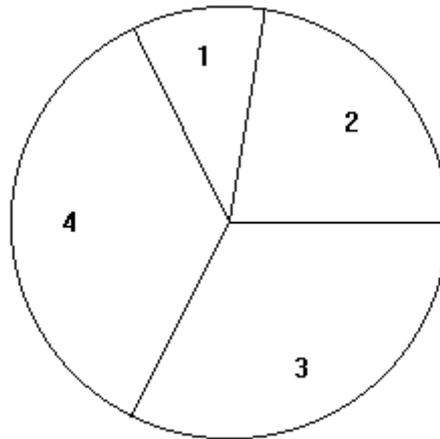
Si las parejas de puntos de la tabla (1, 20), ... las representamos por puntos del plano, al unir todos los puntos obtenemos un polígono de frecuencias.



- **Dibuja el polígono de frecuencias correspondiente a los datos de tu clase.**
- **¿Cuántas veces ha salido un número menor que 4?. ¿Cuántos lanzamientos han dado resultados mayores o iguales que 3?. ¿Cuál es el porcentaje de lanzamientos con resultados menores o iguales a 5?.**
- **Teniendo en cuenta los resultados obtenidos, ¿crees que el dado está bien construido o que, por el contrario, "tiene trampas"?**

## 2. Giro de ruletas

- Disponemos de una ruleta como la de la figura. ¿Piensas que será igual de fácil obtener un 1, un 2, un 3 o un 4 con esta ruleta?. Efectúa 100 giros de la ruleta y anota el número que sale cada vez.



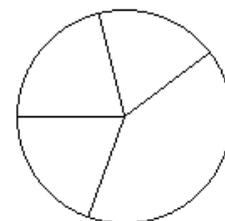
Con los resultados obtenidos, podemos efectuar un recuento y construir una tabla de frecuencias absolutas como la siguiente:

Resultados	Recuento	Frecuencia absoluta
1		
2		
3		
4		

- Efectúa el recuento de los datos obtenidos entre toda la clase. Con los datos obtenidos completa la siguiente tabla:

Resultado	Frecuencia absoluta $f$	$\% = \frac{f}{N} \times 100$	Frecuencia relativa $f_r = \frac{f}{N}$
1			
2			
3			
4			
Totales			

Podemos representar gráficamente estos resultados dibujando un diagrama de sectores, que consiste en un círculo cuyos sectores tienen una amplitud en grados proporcional a las frecuencias respectivas. Puedes obtener la amplitud de cada sector multiplicando la frecuencia relativa por 360 o utilizando otras técnicas. Para hacer el dibujo deberás ayudarte de compás, regla y transportador de ángulos.



- Construye el diagrama de sectores correspondiente a los resultados de tu clase y compáralo con la ruleta original. ¿Observas algo interesante?.
- ¿Crees que es igual de fácil obtener uno cualquiera de los sectores de esta ruleta?.

## ESTADÍSTICA DESCRIPTIVA EN ESO

### Introducción.

En las siguientes páginas se muestran algunos ejemplos de actividades de Estadística Descriptiva experimentadas en el segundo ciclo de E.S.O. En todos los casos se utilizan datos reales procedentes de fuentes como Anuarios Estadísticos de Instituciones, Anuarios de periódicos, revistas, publicaciones del INE, televisión, Internet, etc. Con estas actividades se pretende que los estudiantes puedan:

- buscar información, eligiendo la fuente más adecuada
- interpretar la información contenida en tablas, gráficos y diagramas
- utilizar herramientas matemáticas y recursos tecnológicos adecuados para describir y comunicar la información previamente obtenida.

Es conveniente que los estudiantes se enfrenten desde el primer momento con datos reales (aunque ofrezcan cierto nivel de dificultad), porque el objetivo último de estas actividades no es la enseñanza de algoritmos de cálculo de parámetros, sino que los alumnos aprendan a leer e interpretar la información contenida en tablas y diagramas y aprovechen este conocimiento para criticar la información o hacer predicciones.

Por esta razón, se incluyen datos reales procedentes de diversas publicaciones, tales como:

- Revista DADES (Ayuntamiento de Valencia)
- Anuario estadístico de la ciudad de Valencia
- Anuario de El País.
- Boletines del INE.

### 1. Diagramas estadísticos.

- CONTAMINACIÓN ACÚSTICA

El nivel sonoro se mide a partir de un mínimo que se establece en el umbral de percepción del oído humano, este es de 0 dB (decibelios), entorno a 130 dB se sitúa el nivel del dolor. La O.C.D.E. establece como recomendable valores inferiores a los 65 dB.

Aquí tienes algunos niveles sonoros de referencia:

Sonido de fondo en el campo: entre 15 y 20 dB. Sonido en una biblioteca: en torno a 35 dB. Sonido de una conversación: en torno a 65 dB. Sonido del tráfico: en torno a 70 dB. Sonido de un avión despegando: en torno a 120 dB.
----------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------

Los siguientes datos proceden del Servicio de Medio Ambiente del Ayuntamiento de Valencia. Las medidas del nivel sonoro se toman en tres puntos de la ciudad: Nuevo Centro, Plaza de España y Pista de Silla. La información indica niveles promedios durante un día para las distintas franjas horarias. Las mediciones se han hecho en el cuarto trimestre de 1997.

## NIVEL SONORO EN dB. Media por hora del día

Hora	Nuevo Centro	Plaza de España	Pista de Silla
0	69'1	67'8	66'5
1	67'5	67'1	65'7
2	66'5	65'9	65'4
3	66'1	65'4	65'0
4	68'1	66'1	65'8
5	71'2	68'1	67'4
6	73'2	70'8	69'8
7	73'7	72'6	70'8
8	73'8	73'3	71'2
9	73'6	73'7	71'3
10	73'5	73'9	71'5
11	73'5	74'1	71'5
12	73'5	74'3	71'4
13	73'3	74'0	70'6
14	73'4	73'7	70'0
15	73'9	73'5	70'7
16	73'9	73'9	71'5
17	74'1	74'2	71'8
18	73'8	74'2	71'8
19	73'7	74'1	71'2
20	73'2	73'6	70'8
21	72'4	72'3	69'4
22	71'2	70'7	68'0
23	70'6	69'2	67'1

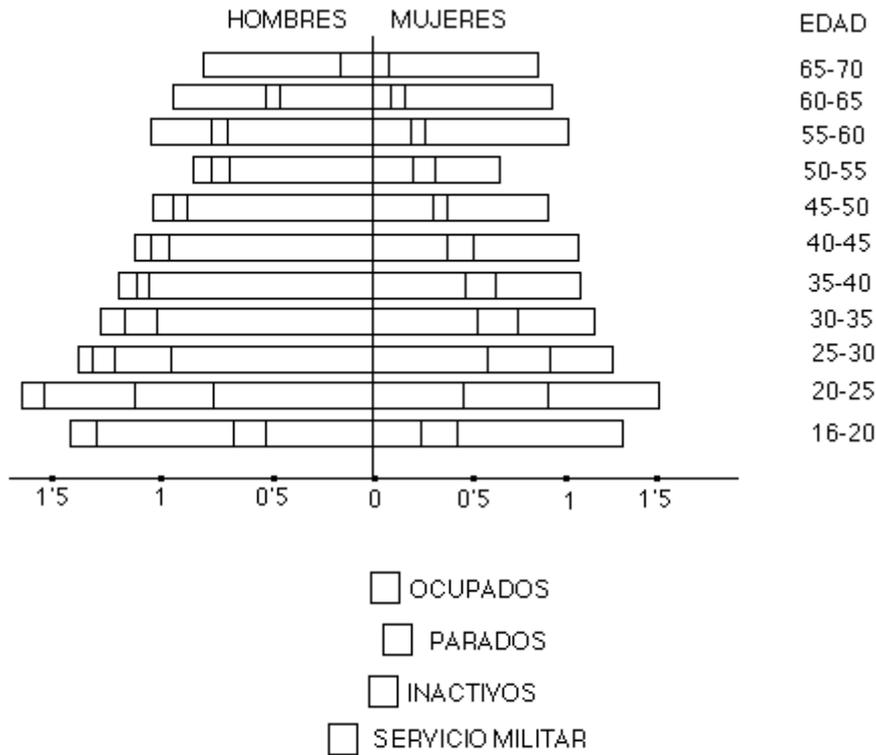
- a) Dibuja, en unos mismos ejes, tres gráficas que muestren la evolución, durante el día, del nivel sonoro en cada punto de referencia (Nuevo Centro, Plaza de España y Pista de Silla).

*Cada una de las gráficas obtenidas se llama serie temporal, porque muestra la evolución en el tiempo de una variable aleatoria. Las hemos dibujado en los mismos ejes para poder establecer comparaciones entre ellas.*

- b) ¿Cuál es el nivel de ruido en cada uno de los puntos de referencia a las 6 de la mañana?. ¿Y a las 14 horas?. ¿A qué horas del día hay un nivel de ruido superior a 72 dB?. ¿En qué lugares?.
- c) ¿Cómo evoluciona el nivel sonoro a lo largo del día?. Redacta un pequeño informe.
- d) ¿Cuál es el mayor nivel de ruido durante el día?. ¿A qué hora?. ¿Dónde?. ¿Y el menor?. ¿A qué hora?. ¿Dónde?. ¿Se cumple la recomendación de la O.C.D.E.?.
- e) ¿Dónde hay más contaminación acústica, en Nuevo Centro o en la plaza de España?. Investiga según las diferentes horas del día.
- f) ¿Habrá mucha variación en la gráfica en otras épocas del año?. ¿Y en otros años?.

• **PIRÁMIDE DE POBLACIÓN**

La siguiente gráfica es una **pirámide de población**. Nos informa de la distribución de población de un país según su relación con la actividad económica en 1992 (en millones de personas).



- a) Averigua aproximadamente el número de hombres y mujeres en edad de trabajar que había en 1992.
- b) Halla aproximadamente el número, así como el porcentaje, de hombres y mujeres ocupados.
- c) Compara la población de hombres y mujeres parados. ¿En qué intervalo de edades se encuentra el mayor número de parados en el grupo de hombres y en el de mujeres?.

• **DIAGRAMAS CIRCULARES**

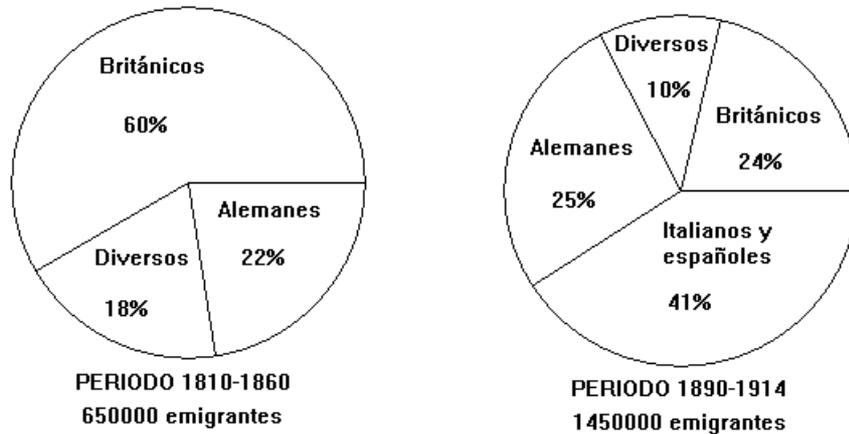
En la tabla siguiente aparecen los datos de las superficies (en millones de km<sup>2</sup> ) y de las poblaciones (en millones de habitantes) de las cinco partes del mundo:

	SUPERFICIE	POBLACIÓN
Europa	10'5	686
África	30'3	452
Australia y Oceanía	8'5	24
Asia	44'4	2580
América	42	586

- a) Sobre un diagrama circular, representa las cinco partes del mundo por sectores directamente proporcionales a su superficie.
- b) Sobre otro diagrama circular, representa las cinco partes anteriores por sectores directamente proporcionales a su población.

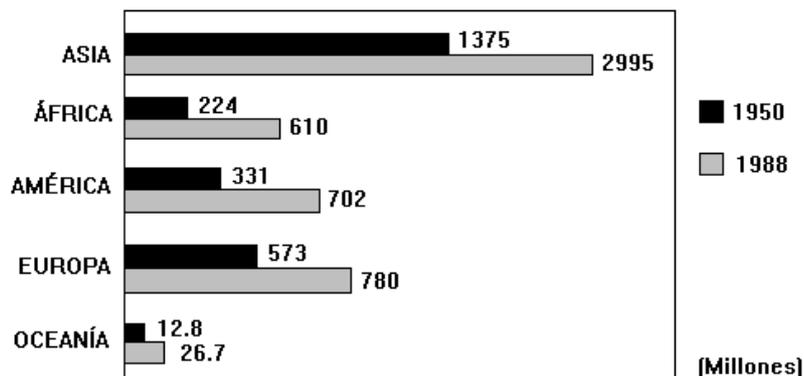
### • EMIGRACIÓN EUROPEA

Dados los siguientes diagramas de sectores, que representan gráficamente la emigración europea a ultramar, elabora con la mayor aproximación posible, las tablas de frecuencias que se han utilizado para la obtención de dichos diagramas. Extrae de dichas tablas las conclusiones pertinentes.



### • POBLACIÓN MUNDIAL

En la siguiente figura se da la distribución de la población mundial en los años 1950 y 1988.



Con estos datos construye un diagrama de sectores para cada año. ¿Cómo ha variado el peso de cada continente en la población total ?.

## 2. Media – Desviación típica. Intervalos de normalidad.

### • TU CLASE

Te proponemos que recojas información sobre tu clase en los siguientes aspectos:

- ✓ número de hermanos.
- ✓ tiempo invertido para desplazarse de casa al instituto.
- ✓ tallas y pesos.

a) Con los datos obtenidos, dibuja:

- ✓ un diagrama de barras para la información referida al número de hermanos.
- ✓ un diagrama de rectángulos para el tiempo de desplazamiento.
- ✓ un diagrama de rectángulos para los pesos.

En los dos últimos casos te será útil agrupar los datos en intervalos. ¿Cómo puedes hacerlo?.

- b) Un histograma es un diagrama de rectángulos en el que el área de cada rectángulo es igual a la frecuencia del intervalo correspondiente. Dibuja el histograma correspondiente a las tallas de tu clase.
- c) Analiza la información obtenida:
  - ✓ ¿Cuál es el número de hermanos más frecuente?.
  - ✓ ¿Cuántos hermanos tienen tus compañeros por término medio?.
  - ✓ ¿Cuál es el tiempo medio de desplazamiento al instituto?.
  - ✓ ¿Cuál es la talla media de la clase?. ¿Y el peso medio?.
  - ✓ ¿Hay mucha dispersión en el número de hermanos?. ¿Y en la talla?. ¿Y en el peso?. ¿Y en el tiempo de desplazamiento?.
  - ✓ ¿Cuál de las cuatro magnitudes consideradas te parece más dispersa?.

*El dato más frecuente en una estadística se llama moda. La media es un valor tal que si todos los datos fueran iguales a él, sumarían lo que realmente suman. Si disponemos de una tabla de frecuencias:*

DATOS	X	X <sub>1</sub>	X <sub>2</sub>	X <sub>3</sub>	... ..	X <sub>N</sub>
FRECUENCIAS	f	f <sub>1</sub>	f <sub>2</sub>	f <sub>3</sub>	... ..	f <sub>N</sub>

entonces, la media  $\bar{X}$  de los datos cumple que  $\sum_{i=1}^N X_i \cdot f_i = N \cdot \bar{X}$ . Por tanto, la media se

calcula mediante la fórmula:  $\bar{X} = \frac{\sum_{i=1}^N X_i \cdot f_i}{\sum_{i=1}^N f_i}$ . Además, una forma de medir la dispersión

de los datos es mediante la varianza, que es la media de los cuadrados de las desviaciones de cada dato a la media y viene dada por la fórmula:

$$V = \frac{\sum_{i=1}^N (X_i - \bar{X})^2 \cdot f_i}{\sum_{i=1}^N f_i}$$

Se toman los cuadrados de las desviaciones, porque la media

de las desviaciones siempre está próxima a cero, puesto que unas desviaciones son positivas y otras negativas, por lo que unas se compensan con otras.

La desviación típica,  $\sigma$ , es la raíz cuadrada positiva de la varianza, es decir:  $\sigma = \sqrt{V}$ .

Un procedimiento más rápido para hallar la media y la desviación típica consiste en utilizar el modo SD de la calculadora:

- ✓ Coloca la calculadora en modo SD o STAT.
- ✓ Introduce los datos y las frecuencias del siguiente modo:

X <sub>1</sub>	×	f <sub>1</sub>	DATA
X <sub>2</sub>	×	f <sub>2</sub>	DATA
....	....	....	....
X <sub>N</sub>	×	f <sub>N</sub>	DATA

- ✓ Para hallar la media, activa la función  $\bar{X}$ .
- ✓ Para hallar la desviación típica, activa la función  $\sigma_n$ .

✓ Para borrar los datos estadísticos de la memoria, pulsa **SHIFT AC**.

### • **BALONCESTO**

En la siguiente tabla tienes los puntos totales conseguidos por cada uno de los jugadores de dos equipos de baloncesto en la pasada liga:

EQUIPO A	315	355	420	392	457	480	387	340
EQUIPO B	444	432	416	388	368	367	362	360

- a) Calcula la media y desviación típica de cada equipo. ¿Qué equipo es mejor?  
 b) ¿Qué equipo tiene puntuaciones menos dispersas en torno a la media?

*El **coeficiente de variación** es el cociente entre la desviación típica y la media de un conjunto de datos estadísticos.*

$$CV = \frac{\sigma}{\bar{X}}$$

*Se suele expresar en porcentaje y sirve para comparar y medir la dispersión relativa de distintas poblaciones.*

*En general, no es menos dispersa la población que presenta menos desviación típica, sino la que presenta un menor coeficiente de variación, ya que la dispersión depende también del valor de la media. No es lo mismo una desviación de 5 frente a una media de 10 que una desviación de 5 frente a una media de 100. Siendo las desviaciones típicas iguales, en el segundo caso hay menor dispersión relativa, lo que se traduce en un menor valor del coeficiente de variación.*

### • **DOS EMPRESAS**

Los gastos mensuales de una empresa A tienen una media de 100000 euros y una desviación típica de 12500 euros. En otra empresa más pequeña B, la media es 15000 euros y la desviación típica 2500 euros. Calcula mediante el coeficiente de variación, cuál de las dos tiene más variación relativa.

### • **EXPORTACIÓN E IMPORTACIÓN**

El volumen de exportaciones de una empresa tiene una media mensual de 650000 dólares, con desviación típica de 92500 dólares. La misma empresa vende mensualmente, en el mercado interior, un promedio de 50 millones de dólares, con desviación típica de 4'3 millones. ¿Qué mercado es más estable, el mercado interior o el exterior?

### • **SOCIOLOGÍA**

En un grupo de sociología se han obtenido estas puntuaciones en un test de habilidad mental:

50	23	45	36	56	34	56	67	45	34	23	45	23	67	54	21	34	43	12	78	36
49	53	27	66	31	45	22	33	44	48	53	57	77	31	23	47	52	33	37	64	21

Comprueba si en el intervalo  $(m-\sigma, m+\sigma)$  se encuentra aproximadamente el 68% de los datos, siendo  $m$  la media y  $\sigma$  la desviación típica.

### • **SOLDADOS**

Las estaturas aproximadas de 4350 soldados son las siguientes:

Estatura (en m)	1'52	1'56	1'60	1'64	1'68	1'72	1'76	1'80	1'84	1'88
Nº de soldados	62	186	530	812	953	860	507	285	126	29

Decimos que los soldados que tienen su estatura entre  $m+\sigma$  y  $m+3\sigma$  son altos; si la tienen entre  $m-3\sigma$  y  $m-\sigma$  son bajos, y son *normales* si la tienen entre  $m-\sigma$  y  $m+\sigma$ .

Calcula, aproximadamente, el porcentaje de bajos, normales y altos que hay en la muestra.

En los ejercicios anteriores habrás observado que en el intervalo  $[m-\sigma, m+\sigma]$  se encuentran, aproximadamente el 67 % de los datos. También se cumple que:

En el intervalo  $[m-\sigma, m+\sigma]$  se encuentra el 67% de los datos.  
 En el intervalo  $[m-2\sigma, m+2\sigma]$  se encuentra el 75% de los datos.  
 En el intervalo  $[m-3\sigma, m+3\sigma]$  se encuentra el 89% de los datos.  
 En el intervalo  $[m-4\sigma, m+4\sigma]$  se encuentra el 94% de los datos.  
 En el intervalo  $[m-5\sigma, m+5\sigma]$  se encuentra el 96% de los datos.

En general, es muy poco probable que un dato se aparte de la media más de tres desviaciones típicas.

Este resultado es el Teorema de Tchebycheff y puedes utilizarlo para obtener intervalos centrados en la media que contengan un cierto % de los datos.

### • MATERNIDAD

En una maternidad se han tomado los pesos (en kg) de 50 recién nacidos, obteniendo estos resultados:

2'8	3'2	3'8	2'5	2'7	3'7	1'9	2'6	3'5	2'3
3'0	2'6	1'8	3'3	2'9	2'1	3'4	2'8	3'1	3'9
2'9	3'5	3'0	3'1	2'2	3'4	2'5	1'9	3'0	2'9
2'4	3'4	2'0	2'6	3'1	2'3	3'5	2'9	3'0	2'7
2'9	2'8	2'7	3'1	3'0	3'1	2'8	2'6	2'9	3'3

- Calcula la media y la desviación típica.
- Halla intervalos, centrados en la media, que contengan el 75 %, 89 %, 94 % y 96 % de los datos, respectivamente.

### • RITMO CARDÍACO

A un grupo de 30 personas se le ha tomado el número de pulsaciones por minuto (ritmo cardíaco) y se han obtenido los siguientes resultados:

87	85	61	51	64	75	80	70	69	82	80	79	82	74	90
76	72	73	63	65	67	71	88	76	68	73	70	76	71	86

- Calcula la media y la desviación típica.
- Halla intervalos, centrados en la media, que contengan el 75 %, 89 % y 94 % de los datos, respectivamente.

### • DENSIDAD DE POBLACIÓN

La densidad de población, en habitantes por  $\text{km}^2$ , de las provincias españolas, en el año 1996, era de:

50	141	54	63	43	48	158	110	13	10	48	105	136
184	226	99	23	24	13	12	32	23	25	34	24	29
22	10	60	21	597	83	29	83	208	65	193	31	21
141	41	59	201	598	89	50	88	345	532	52	3620	3742

En ocasiones no resulta útil usar intervalos de la misma amplitud, como ocurre en este caso. Utiliza los intervalos  $[0, 20)$ ,  $[20, 40)$ ,  $[40, 80)$ ,  $[80, 160)$ ,  $[160, 3800]$  y elabora la tabla

estadística completa, incluyendo marcas de clase y frecuencias absolutas y relativas acumuladas.

### 3. Medidas de posición.

#### • FRECUENCIAS ACUMULADAS

La *frecuencia acumulada* correspondiente a un determinado valor es la suma de todas las frecuencias de los valores anteriores a él más su propia frecuencia.

Por ejemplo, hemos efectuado 100 lanzamientos de cuatro monedas obteniendo estos resultados:

$X$	Nº DE CARAS	0	1	2	3	4
$f$	FRECUENCIA	6	26	37	24	7

La frecuencia acumulada de  $X=2$  es  $F(2) = f(0) + f(1) + f(2) = 6 + 26 + 37 = 69$ .

La frecuencia acumulada de  $X=4$  es  $F(4) = f(0) + f(1) + f(2) + f(3) + f(4) = 100$ .

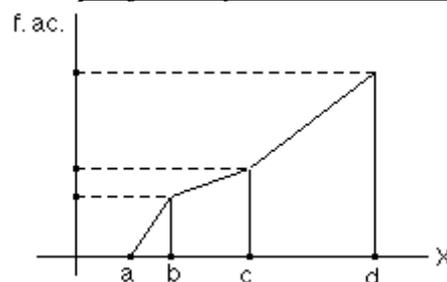
La frecuencia absoluta acumulada correspondiente al último dato coincide con el tamaño muestral  $N$ .

La frecuencia relativa acumulada correspondiente al último dato es igual a 1.

Si los datos están agrupados en intervalos, las frecuencias acumuladas, absolutas o relativas, se representan se la siguiente forma:

- ✓ Por el extremo derecho de cada intervalo se traza una vertical de altura igual a la frecuencia acumulada, absoluta o relativa, correspondiente.
- ✓ Se unen por una poligonal los extremos de estas verticales.
- ✓ Se completa la poligonal uniendo el extremo izquierdo del primer intervalo con el extremo superior de la primera vertical.

La gráfica así obtenida se llama polígono de frecuencias acumuladas.



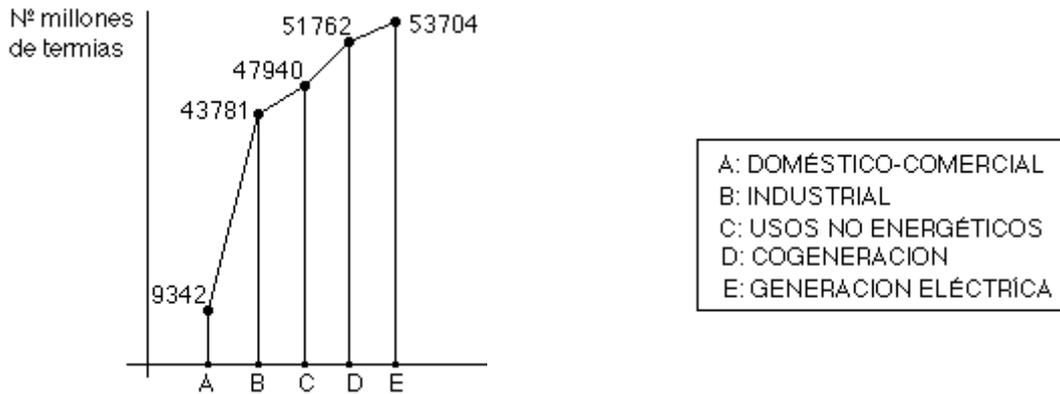
- 1) Construye la tabla de frecuencias absolutas y relativas acumuladas correspondiente a los 100 lanzamientos de las cuatro monedas.
- 2) La distribución de trabajadores de una empresa según su edad es la indicada en esta tabla:

EDADES (años)	FRECUENCIA (nº trabajadores)
16 – 20	10
20 – 25	15
25 – 35	22
35 – 45	24
45 – 55	18
55 – 65	11

Construye la tabla de frecuencias absolutas y relativas acumuladas y dibuja el polígono de frecuencias absolutas acumuladas.

• **GAS NATURAL**

La siguiente gráfica es un polígono de frecuencias acumuladas que muestra los distintos usos del gas natural en 1990.



**Termia** es la cantidad de calor necesaria para aumentar la temperatura de una tonelada de agua en 1° C.

- a) ¿Qué cantidad de energía procedente del gas natural se consumió en total durante el año 1990?.
- b) ¿A qué se destinó la mayor cantidad de gas natural en 1990?.
- c) La siguiente tabla recoge las previsiones para el año 2020 de los usos del gas natural. Dibuja el polígono de frecuencias acumuladas correspondiente.

USOS	PREVISIÓN DEMANDA GAS NATURAL AÑO 2020 (millones de termias)
Doméstico – comercial	13400
Industrial	56700
Usos no energéticos	7200
Cogeneración	24900
Generación eléctrica convencional	31000

• **MEDIANA**

La *mediana* de una distribución estadística es el valor que deja a su izquierda un número de datos igual a los que deja a su derecha; es decir, se trata del valor central de la distribución. Se representa por *M*.

*Ejemplo 1.-* Las calificaciones de 7 alumnos en Lengua y de otros 8 en Matemáticas han sido las siguientes:

2 4 5 6 8 9 10

1 3 4 5 6 7 9 10

En las notas de Lengua, la nota 6 deja tres notas a su izquierda y tres a su derecha; es el valor central, porque el número de dados es impar. En las de Matemáticas, como no hay

una calificación central, tomamos la media aritmética de las dos centrales:

$$M = \frac{5+6}{2} = 5.5.$$

Si el número de datos es impar, la mediana coincide con el valor central. Si el número de datos es par, hay dos valores centrales; entonces, la mediana es la media aritmética de los dos valores centrales.

Ejemplo 2.- Las notas que obtuvieron 32 alumnos de una clase, en Matemáticas, en la primera evaluación fueron las siguientes:

X	1	2	3	4	5	6	7	8	9	10
f	2	2	3	5	7	5	3	2	2	1

En primer lugar hay que ordenar todos los datos en orden creciente, teniendo en cuenta que hay dos 1, dos 2, tres 3, cinco 4, siete 5, etc.

1, 1, 2, 2, 3, 3, 3, 4, 4, 4, 4, 4, 5, 5, 5, 5, 5, 5, 5, 5, 6, 6, 6, 6, 6, 6, 7, 7, 7, 8, 8, 9, 9, 10

Como hay 32 datos (que es par), hay dos valores centrales, cuyas posiciones son  $\frac{32}{2} = 16$  y 17. En la posición 16 hay un 5 y en la posición 17 hay un 5. Luego la mediana

$$\text{es } M = \frac{5+5}{2} = 5.$$

También podemos calcular la mediana en este caso utilizando las frecuencias acumuladas:

X	1	2	3	4	5	6	7	8	9	10
f	2	2	3	5	7	5	3	2	2	1
F	2	4	7	12	19	24	27	29	31	32

Como el número de datos  $N=32$  es par, la mediana es media aritmética de los dos valores centrales que ocupan las posiciones  $\frac{N}{2} = \frac{32}{2} = 16$  y  $\frac{N}{2} + 1 = 17$ , que a la vista de la tabla de frecuencias acumuladas, son 5 y 5, respectivamente, ya que en dicha tabla se observa que desde la posición 12 hasta la posición 18 los datos son siempre 5. Por lo tanto la mediana es:  $M = \frac{5+5}{2} = 5$ .

Si el número de datos  $N$  es par, hay dos valores centrales, cuyas frecuencias acumuladas son  $\frac{N}{2}$  y  $\frac{N}{2} + 1$ . Entonces la mediana es la media aritmética de estos dos valores centrales.

Si el número de datos  $N$  es impar, hay un valor central, cuya frecuencia acumulada es  $\frac{N+1}{2}$ . Entonces la mediana es este valor central.

Si los datos están agrupados en intervalos, para determinar la mediana dibujaremos el polígono de frecuencias acumuladas, trazaremos una horizontal por la frecuencia acumulada  $\frac{N}{2}$  hasta alcanzar a la gráfica y la proyección sobre el eje horizontal será la mediana.

- 1) Las temperaturas medias en las capitales de 12 países europeos, en grados centígrados, son:

Amsterdam	13	Atenas	25	Berlín	14
-----------	----	--------	----	--------	----

Bruselas	15	Copenaghe	11	Dublín	14
Lisboa	20	Londres	15	Luxemburgo	15
Madrid	20	París	15	Roma	22

Calcula en esta distribución estadística la media, la moda y la mediana.

- 2) Esta tabla recoge las medidas de las cinturas de 100 personas. Calcula la mediana.

Cintura (cm)	40-50	50-60	60-70	70-80	80-90	90-100
Frecuencias	2	11	26	40	20	1

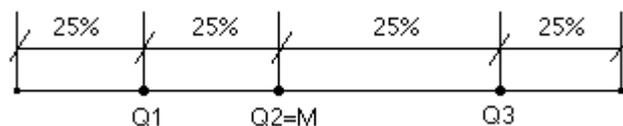
- 3) Los resultados obtenidos al lanzar un dado 200 veces vienen reflejados en la siguiente tabla:

Número de puntos	1	2	3	4	5	6
Repeticiones	x	32	35	33	y	35

- a) Determina las frecuencias que faltan, x e y, sabiendo que la puntuación media es 3'6.  
 b) Calcula la mediana y determina un intervalo centrado en la mediana que contenga al 50% de los datos.

### • CUARTILES

Los cuartiles  $Q_1$ ,  $Q_2$  y  $Q_3$  son los valores que dividen la distribución en cuatro partes iguales, es decir, de igual porcentaje.



El primer cuartil deja por debajo un 25% de los datos y por encima un 75%. El segundo cuartil deja un 50% de datos por debajo y un 50% por encima. El tercer cuartil deja por debajo un 75% de los datos y un 25% por encima.

El segundo cuartil coincide con la mediana.

Decir, por ejemplo, que el precio de un coche está en el tercer cuartil significa que dicho coche es más caro que el 75% de los coches del mercado y que es más barato que el 25% restante.

Para obtener los cuartiles a partir de una tabla de frecuencias acumuladas basta tener en cuenta que:

- ✓ La frecuencia acumulada asociada al primer cuartil  $Q_1$  es igual a  $\frac{N}{4}$ .
- ✓ La frecuencia acumulada asociada al segundo cuartil  $Q_2$  es igual a  $\frac{N}{2}$ .
- ✓ La frecuencia acumulada asociada al tercer cuartil  $Q_3$  es igual a  $\frac{3N}{4}$ .

Si los datos están agrupados en intervalos, para calcular los cuartiles basta dibujar el polígono de frecuencias acumuladas y trazar rectas horizontales por las frecuencias acumuladas  $\frac{N}{4}$ ,  $\frac{N}{2}$  y  $\frac{3N}{4}$ , hasta alcanzar a la gráfica. Las proyecciones sobre el eje horizontal son los cuartiles.

Llamamos rango intercuartílico a la diferencia entre el primer y el tercer cuartil, es decir se cumple que:  $R = Q_3 - Q_1$ . El rango intercuartílico contiene el 50% de los datos de la muestra.

El uso de los cuartiles y del rango intercuartílico permite determinar intervalos centrados en la mediana que contengan la mitad de los datos muestrales.

- 1) En una clase de 30 alumnas y alumnos se hace una encuesta para conocer el número de hermanos que tienen cada uno y los resultados son los siguientes:

Nº HERMANOS	0	1	2	3	4
FRECUENCIA	6	10	8	5	1

- a) Amplia la tabla con dos columnas que contengan los datos correspondientes a las frecuencias y porcentajes acumulados.
- b) Observando los datos de la tabla, ¿sabrías hallar la mediana y los cuartiles?.
- 2) En una clase se recogen datos sobre el tamaño de las viviendas en las que residen los estudiantes. Los resultados se indican en la siguiente tabla:

SUPERFICIE (m <sup>2</sup> )	[40, 60)	[60, 80)	[80, 100)	[100, 120)	[120, 140)
FRECUENCIA	2	12	3	5	8

- a) Representa gráficamente el polígono de frecuencias acumuladas.
- b) Trazando paralelas al eje X a la altura de los siguientes porcentajes: 25 %, 50 % y 75 %, obtén a partir de la gráfica los valores de los cuartiles y la mediana. Interpretalos.

### • ANDAR

Se ha observado en una muestra de 30 niños la edad, en meses, a la que empiezan a andar, obteniéndose los siguientes resultados:

Meses	9	10	11	12	13	14	15
Frecuencia	1	2	4	1	3	6	3

- a) Justifica por qué la media es superior a la mediana.
- b) Halla un intervalo de edades centrado en la mediana que contenga al 50% de los datos.
- c) ¿A partir de qué edad se encuentra el 25% de los niños mas tardíos en comenzar a andar?.

### • UN TEST

En un test, compuesto de 10 preguntas, a un grupo de 40 personas, se dan los siguientes resultados:

Nº de respuestas	[0, 2)	[2, 4)	[4, 6)	[6, 8)	[8, 10)
Nº de personas	4	9	15	7	5

- a) Representa gráficamente la distribución.
- b) ¿A partir de qué valor se encuentra el 75% de las personas que han obtenido mejor resultado?.

- c) ¿A partir de qué valor se encuentra el 25% de las personas que han obtenido mejor resultado?.

## DESCRIPCIÓN DE DATOS ESTADÍSTICOS CON LA CALCULADORA GRÁFICA TI-83

### Introducción

La calculadora gráfica es una herramienta de trabajo muy útil que presenta la gran ventaja de la rapidez en los cálculos y en las representaciones gráficas, lo que permite que se pueda dedicar más tiempo al análisis de los conceptos y a la interpretación de los resultados. Además, es muy sencillo analizar los efectos que producen en los resultados las variaciones efectuadas en los datos iniciales. Permite, en suma, hacer conjeturas, ponerlas a prueba, analizar y discutir los resultados obtenidos. De esta forma el foco de atención deja de ser el algoritmo de cálculo para centrarse en los conceptos.

El proceso de trabajo estadístico con una calculadora gráfica sería el siguiente:

1. Selección de los datos y preparación de los cálculos. (Introducción de los datos en la calculadora).
2. Selección y realización de los cálculos pertinentes. (Obtención de los parámetros estadísticos).
3. Representación gráfica.
4. Análisis de los resultados.
5. Conclusiones (y posible regreso a los pasos anteriores 1, 2, 3 y 4).

Vamos a ver algunos ejemplos de datos estadísticos procedentes del Instituto Nacional de Estadística (INE, España en cifras, 2000) y del Anuari Estadístic de la Ciutat de València, 1999. Con ellos analizaremos el uso de la calculadora gráfica para analizar los datos publicados, hacer comparaciones y extraer conclusiones a partir de los resultados obtenidos.

### NUPCIALIDAD Y NATALIDAD. 1998 (España en cifras, INE, [www.ine.es](http://www.ine.es))

	<b>Matrimonios</b>	<b>Nacimientos</b>
<b>ESPAÑA</b>	<b>202.494</b>	<b>361.930</b>
Andalucía	40.022	75.854
Aragón	5.973	9.184
Asturias (Principado de)	4.683	6.429
Balears (Illes)	4.380	8.161
Canarias	7.346	16.966
Cantabria	2.686	3.824
Castilla y León	12.097	17.000
Castilla-La Mancha	9.711	15.789
Cataluña	32.046	56.314
Comunidad Valenciana	21.412	37.190
Extremadura	5.705	9.863

Galicia	12.344	18.310
Madrid (Comunidad de)	23.164	49.034
Murcia (Región de)	6.193	12.703
Navarra (C. Foral de)	2.591	4.953
País Vasco	10.079	16.168
Rioja (La)	1.357	2.163
Ceuta	296	969
Melilla	409	1.056

### 3. Selección de los datos y preparación de los cálculos.

Pulsamos ON para iniciar la calculadora. Si aparecen datos en pantalla, pulsamos CLEAR para borrarlos.

Pulsamos STAT. Elegimos la opción 1: Edit, para lo cual pulsamos la tecla 1 o ENTER. Se mostrará el editor de listas estadísticas.

EDIT CALC TESTS
1: Edit...
2: SortA(
3: SortD(
4: ClrList
5: SetUpEditor

L1	L2	L3	1
----	----	----	----
L1=			

L5	L6	----	7
----	----	----	----
Name=			

Pulsamos **▶** hasta que aparezca una lista sin nombre. Se muestra el indicador Name= en la línea de introducción y se activa el bloqueo alfabético. Pulsamos las letras o números para el introducir el nombre. El primer carácter no puede ser un número.

Pulsamos ENTER ó **▼** para almacenar el nombre de la lista en la columna actual del editor de listas estadísticas. El nombre de la lista ahora será un elemento del menú LIST NAMES.

Ahora comenzamos a introducir los datos en las listas MAT y NAC, en la lista NUM introducimos 1, 2, 3, 4, ... que representan a las distintas comunidades autónomas.

Para comparar la nupcialidad con la natalidad en cada comunidad, podemos hacer una lista COC con los cocientes entre el número de nacimientos y el número de matrimonios; de esta forma sabremos cuántos nacimientos se producen, por término medio por cada nuevo matrimonio. Para ello, situamos el cursor en la etiqueta COC y tecleamos LNAC / LMAT y pulsamos ENTER.

NUM	MAT	NAC	9
1	40022	75854	
2	5973	9184	
3	4683	6429	
4	4380	8161	
5	7346	16966	
75854			

MAT	NAC	COC	10
40022	75854	----	
5973	9184		
4683	6429		
4380	8161		
7346	16966		
COC=			

MAT	NAC	COC	10
40022	75854	1.8953	
5973	9184	1.5376	
4683	6429	1.3728	
4380	8161	1.8632	
7346	16966	2.3096	
COC(1)=1.895307580...			

Volveremos más tarde a la edición de datos, pero vamos ahora a ver algunas opciones de cálculos estadísticos. Concretamente veremos cómo obtener con la calculadora gráfica los parámetros estadísticos más significativos (de centralización, de dispersión, de posición).

#### 4. Selección y realización de los cálculos.

Pulsamos STAT opción CALC, opción 1: 1-Var Stats. Con ello aparece en la pantalla principal el texto 1-Var Stats. A continuación escribimos el nombre de la lista 2nd STAT y flechas de cursor hasta seleccionar MAT. Finalmente pulsamos ENTER. En la ventana siguiente se muestran los parámetros estadísticos más significativos. Podemos pulsar la tecla de cursor ▼ para visualizar el resto de parámetros.

EDIT CALC TESTS 1: 1-Var Stats 2: 2-Var Stats 3: Med-Med 4: LinReg(ax+b) 5: QuadReg	NAMES OPS MATH 1: COC 2: NAC 3: NUM 4: MAT	1-Var Stats LMAT
----------------------------------------------------------------------------------------------------	--------------------------------------------------------	------------------

1-Var Stats $\bar{x} = 10657.57895$ $\Sigma x = 202494$ $\Sigma x^2 = 4336032578$ $S_x = 10999.84037$ $\sigma_x = 10706.45842$ ↓ n = 19	1-Var Stats ↑ n = 19 minX = 296 $Q_1 = 2686$ Med = 6193 $Q_3 = 12344$ maxX = 40022
-----------------------------------------------------------------------------------------------------------------------------------------------------------	------------------------------------------------------------------------------------------------------

Hacemos ahora lo mismo pero seleccionando como datos los contenidos en la lista NAC:

1-Var Stats LNAC	1-Var Stats $\bar{x} = 19048.94737$ $\Sigma x = 361930$ $\Sigma x^2 = 1.46321 E10$ $S_x = 20733.42751$ $\sigma_x = 20180.43646$ ↓ n = 19	1-Var Stats ↑ n = 19 minX = 969 $Q_1 = 4953$ Med = 12703 $Q_3 = 18310$ maxX = 75854
------------------	------------------------------------------------------------------------------------------------------------------------------------------------------------	-------------------------------------------------------------------------------------------------------

Con los resultados obtenidos podemos comparar la media del número de matrimonios y del número de nacimientos, así como las medianas y los rangos intercuartílicos.

El número medio de matrimonios es de 10657, con  $Q_1=2686$  y  $Q_3=12344$ , con una mediana de 6193. El número medio de nacimientos es de 19048, con  $Q_1=4953$  y  $Q_3=18310$ , con una mediana igual a 12703.

Estos resultados podemos representarlos gráficamente.

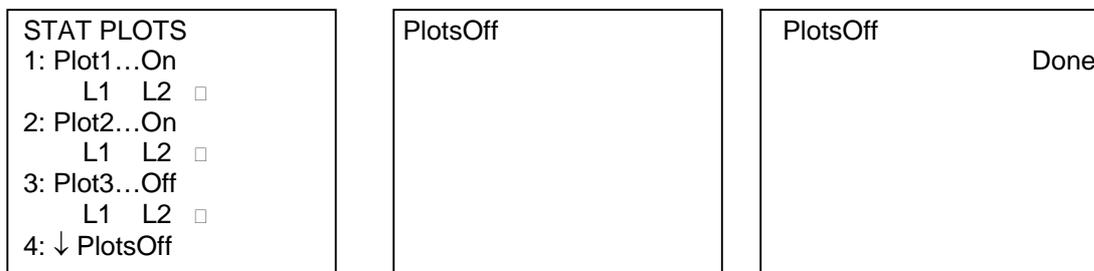
## 5. Representación gráfica.

Para visualizar el menú STAT PLOT hay que pulsar las teclas  $2^{nd}$   $Y=$ . La pantalla STAT PLOTS nos muestra 5 posibles opciones. La quinta se puede ver pulsando la tecla de cursor  $\blacktriangledown$ .

Las tres primeras indican el estado de las tres posibles representaciones estadísticas que puede realizar simultáneamente la calculadora. Si está visible la opción ON, significa que ese gráfico está activado; si está visible la opción OFF, el gráfico está desactivado.

Las opciones 4 y 5 sirven para poner todas las representaciones en ON o en OFF. Para desactivar todos los gráficos estadísticos elegimos la opción 4: PlotsOff. De esta forma aparece este comando en la pantalla principal. Al pulsar ENTER se ejecuta el comando apareciendo en pantalla el mensaje "Done".

La TI-83 puede mostrar simultáneamente tres representaciones gráficas estadísticas, además de las funcionales.



Cuando volvemos a la pantalla de representaciones estadísticas (pulsando  $2^{nd}$   $Y=$ ) observamos que las tres opciones están en OFF.

Elegimos la primera opción 1: Plot1... Aparece una pantalla en la que podemos indicar el tipo de representación gráfica de las seis posibles y qué listas actuarán como Xlist, Ylist o Freq, así como el tipo de trazado para los puntos.

En primer lugar seleccionamos la opción ON, para activar el gráfico. Para ello basta pulsar ENTER cuando el cursor esté situado sobre la opción. Para moverse de una a otra opción hay que utilizar las teclas de cursor.

A continuación en la línea Type: elegimos el tipo de representación. El primero de ellos es el de nube de puntos (diagrama de dispersión) en la que hay que decir qué lista es la X, qué lista es la Y y la marca para la representación (hay tres opciones).

La siguiente representación posible es como la anterior pero uniendo los puntos. Hay que indicar también quién es Xlist, Ylist y la marca Mark.

El tercer tipo es el Histograma. En el caso de tomar este tipo es necesario que determinemos Xlist y la frecuencia Freq, que puede ser cualquiera de las listas o la lista constante igual a 1. El valor de la variable de ventana Xscl determina el ancho de cada barra, empezando en Xmin. ZoomStat ajusta Xmin, Xmax y Xscl, de forma que siempre se cumple la desigualdad

$$(X_{\max} - X_{\min}) / X_{\text{scl}} \leq 47$$

El cuarto tipo es el diagrama de caja con valores alejados (outliers). Los outliers son aquellos datos que distan de los cuartiles más allá de 1'5 veces el rango intercuartílico. El rango intercuartílico es la diferencia entre el tercer cuartil  $Q_3$  y el primer cuartil  $Q_1$ .

El quinto tipo es el diagrama de caja sin valores alejados. En este diagrama se representan los valores mínimo, máximo y rango intercuartílico. En este caso hay que indicar Xlist y Freq.

El sexto tipo es el diagrama de probabilidad normal. Representa cada observación de X en Data List frente al cuartil correspondiente z de la distribución normal estándar. Si los puntos representados se aproximan a una línea recta, el gráfico indicará que los datos siguen una distribución normal.

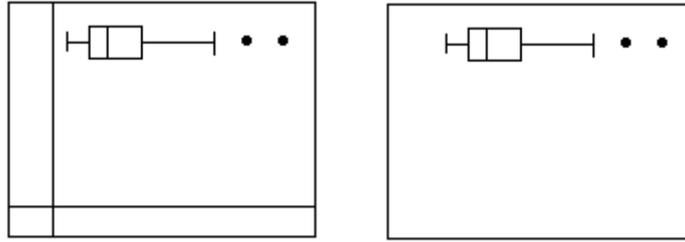
En nuestro caso elegimos el diagrama de caja con valores alejados, tomando como Xlist la lista MAT con Freq igual a 1 y marca  $\square$ .

Volvemos a la pantalla resumen de las opciones de representación estadística, pulsando 2nd Y= y observamos que efectivamente está seleccionado el Plot1 con la opción elegida de caja para la lista MAT.

Para obtener la representación gráfica en pantalla, pulsamos Y= para comprobar que no hay ninguna función en el menú de funciones. En caso de existir alguna, basta desplazarse hasta ella con las teclas de cursor y borrarla, pulsando CLEAR. A continuación, pulsamos la tecla ZOOM.

ZOOM	MEMORY
1: Zbox	
2: Zoom In	
3: Zoom Out	
4: Zdecimal	
5: Zsquare	
6: Zstandard	
7↓ Ztrig	

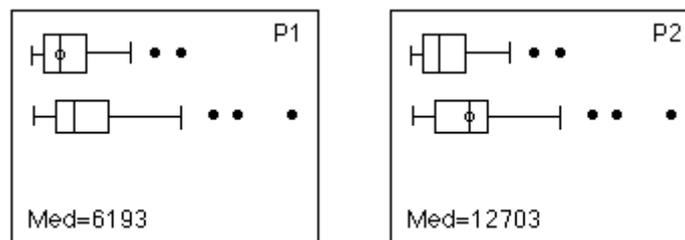
En esta pantalla se muestran distintas opciones de visualización. Podemos utilizar las teclas de cursor y pulsar ENTER cuando esté seleccionada la opción 9: ZoomStat, o bien directamente pulsar la tecla 9 para activar dicha opción. Con ello se establecerán las dimensiones adecuadas de la ventana gráfica y se representará el diagrama de caja.



Si no queremos visualizar los ejes coordenados, pulsamos 2nd ZOOM para activar el menú FORMAT y con ayuda de las teclas de cursor, seleccionamos la opción AxesOff y pulsamos ENTER. Posteriormente, regresamos a la pantalla gráfica, pulsando la tecla GRAPH.

Pulsamos la tecla TRACE y al mover el cursor con las teclas de desplazamiento se sitúa (y nos lo indica en el extremo izquierdo) en minX, Q1, Med, Q3,  $X=1.5 \cdot (\text{rango intercuartílico})$  y los valores alejados. Con ello podemos visualizar el rango intercuartílico (que contiene el 50% de los datos) y comparar su gran concentración (caja muy estrecha) frente a los datos comprendidos entre Q3 y el maxX (que suponen el 25% de los datos), así como lo separado que está este valor máximo de los 50% centrales. En realidad este valor máximo es un valor alejado, ya que se encuentra a más de 1.5 veces el rango intercuartílico, con lo que para un estudio de la evolución del número de matrimonios debería eliminarse por la distorsión que produce en la media).

Vamos a comparar los datos de matrimonios con los de nacimientos. Para ello vamos a activar el mismo tipo de diagrama de caja en el Plot2 para la lista NAC: Pulsamos 2nd Y=, elegimos la opción 2: Plot2... y activamos la opción ON. En Xlist tomamos la lista NAC (que hay que elegir del menú de listas LIST, pulsando 2nd STAT y utilizando las teclas de cursor). En Freq elegimos 1. A continuación pulsamos ZOOM 9 para activar la ventana de gráficos estadísticos. En pantalla se muestran los dos diagramas de caja.

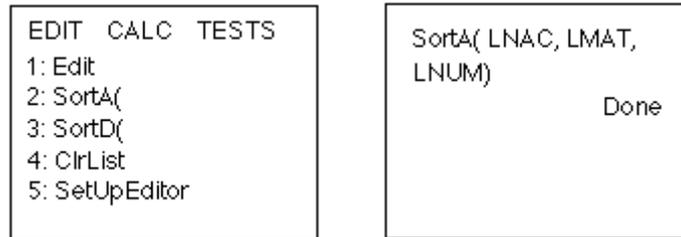


Pulsamos la tecla TRACE. Aparece en el extremo superior derecho de la pantalla la indicación P1 y el cursor se sitúa en la barra correspondiente a la mediana de la lista MAT y en la parte inferior de la pantalla se indica el valor numérico de la mediana.

Pulsando las teclas ◀ y ▶, el cursor se desplaza a lo largo del diagrama Plot1 correspondiente a los datos de matrimonios. Cuando pulsamos las teclas ▲ y ▼, el cursor se traslada al segundo diagrama de caja, el correspondiente al número de nacimientos. Al desplazar el cursor a la izquierda y a la derecha recorreremos los valores minX, Q1, Med, Q3, etc de dicho diagrama.

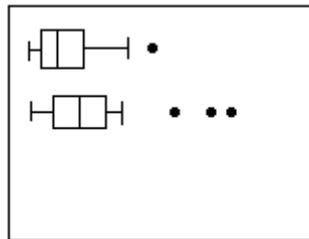
De esta forma hemos localizado los valores outliers, tanto del número de matrimonios como de nacimientos. Se trata de los correspondientes a Cataluña y Andalucía. Si eliminamos dichos valores, obtendremos una media más ajustada, menos distorsionada, y los diagramas de caja serán más representativos.

Pulsamos STAT y elegimos la opción 2: SortA( para ordenar los datos ascendentemente (de menor a mayor). En la pantalla principal aparece SortA(. Pulsamos 2nd STAT y las flechas de cursor para seleccionar las listas NAC, MAT y NUM. Después pulsamos la tecla ) y ENTER.

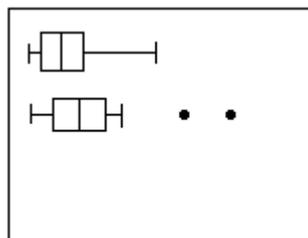


De esta forma se ordenan los datos de las tres listas ascendentemente, arrastrando los datos aparejados de MAT y NUM. Al ejecutar el comando aparece la respuesta “Done” en pantalla. El comando SortA( LNAC ) ordena sólo la lista NAC y no las demás, por lo que perderíamos la relación que existe entre las listas. Podemos comprobar en el editor de listas que, efectivamente éstas se han ordenado de menor a mayor. Desplazamos el cursor hasta el final de las listas y con DEL eliminamos la última fila, correspondiente a los datos de Andalucía.

Pulsando ZOOM 9 obtenemos un nuevo diagrama de caja. También podemos obtener los nuevos parámetros estadísticos mediante los comandos 1-Var Stats LMAT y 1-Var Stats LNAC.



Podemos quitar ahora los datos correspondientes a Cataluña para ver como evolucionan los parámetros estadísticos y los diagramas de caja. Pulsamos STAT y elegimos la opción 1: Edit... Desplazamos el cursor hasta la última fila de las variables NUM, MAT y NAC y la eliminamos con la tecla DEL. Al igual que antes, pulsando ZOOM 9 podemos visualizar los nuevos diagramas de caja y con los comandos correspondientes obtener los nuevos parámetros estadísticos.



Volvemos a introducir los valores eliminados en las listas MAT y NAC (introduciéndolos, por supuesto, al final de las listas de forma emparejada, para que se conserve la relación de orden). Vamos a estudiar la relación entre el número de nacimientos y el de matrimonios. Para ello

analizaremos la lista COC. Como esta lista ha quedado desordenada, hemos de definirla de nuevo. Pasamos a la etiqueta COC y la definimos de nuevo como  $COC = LNAC / LMAT$  y pulsamos ENTER.

MAT	NAC	COC 10
296	969	3.2736
409	1056	2.5819
1357	2163	1.594
2686	3824	1.4237
2591	4953	1.9116
COC(1)=3.273648648...		

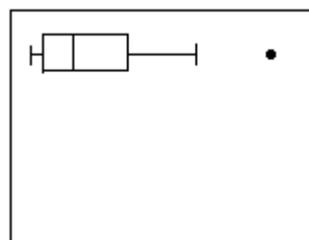
En esta ventana observamos que en Ceuta se han producido 3'27 nacimientos por cada matrimonio, en Melilla 2'58 nacimientos por cada matrimonio, etc. Ahora podemos obtener los parámetros estadísticos de esta lista COC:

1-Var Stats LCOC

1-Var Stats  
 $\bar{x} = 1.856472405$   
 $\Sigma x = 35.2729757$   
 $\Sigma x^2 = 69.4254186$   
 $S_x = .4679810387$   
 $\sigma_x = .455499295$   
 $\downarrow n = 19$

1-Var Stats  
 $\uparrow n = 19$   
 $\min X = 1.3728379$   
 $Q_1 = 1.537585803$   
 $\text{Med} = 1.73687652$   
 $Q_3 = 2.051186824$   
 $\max X = 3.2736486$

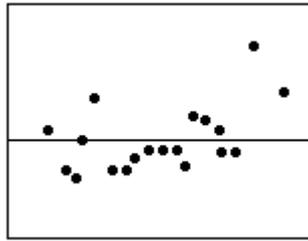
Y también podemos obtener el diagrama de caja correspondiente, para lo que primero desactivamos todos los gráficos estadísticos, pulsando 2nd Y= y eligiendo el comando 4: PlotsOff. A continuación activamos el Plot1, introduciendo como variable Xlist la lista COC. Al pulsar de nuevo ZOOM 9 obtenemos el siguiente diagrama:



O también podemos formar el gráfico de puntos o diagrama de dispersión, introduciendo como variable Xlist la lista NUM y como variable Ylist la lista COC. Al pulsar ZOOM 9 obtenemos:

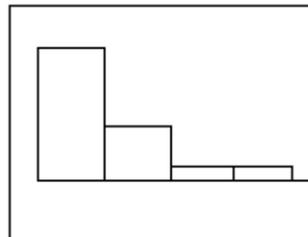


Si queremos tener representada la media de la variable COC, introducimos en el editor de funciones  $Y =$  la expresión  $Y_1 = \bar{x}$ , pulsando VARS , 5: Statistics... , 2:  $\bar{x}$ . Al acceder a la pantalla gráfica, pulsando GRAPH, obtenemos:



quedando por encima de la media claramente las comunidades de Ceuta, Melilla y Canarias, que ocupan los números de orden 18, 19 y 5, respectivamente. Resulta curioso que estas comunidades estén situadas en las zonas más sureñas del país, muy próximas a territorio africano.

Cuando la opción de representación es HISTOGRAMA, en la definición de la ventana de representación WINDOW, Xscl representa la amplitud de los rectángulos, por lo que si damos una amplitud, por ejemplo, de 0.5 y definimos Ymax = 13, al pulsar GRAPH, obtenemos (una vez borrada la expresión  $Y_1 = \bar{x}$ ):



Al pulsar TRACE y las teclas de cursor, observamos que hay 12 comunidades autónomas con entre 1.3 y 1.8 nacimientos por matrimonio, 5 comunidades que tienen entre 1.8 y 2.3 nacimientos por matrimonio, una comunidad con un número de nacimientos por matrimonio entre 2.3 y 2.8 y una sola comunidad entre 2.8 y 3.3.

## 6. Series temporales.

Hemos realizado un tratamiento de los datos en sentido amplio, obteniendo consecuencias de los parámetros estadísticos y de las representaciones gráficas. Todo ello ha sido posible gracias

a la potencia y versatilidad de la calculadora gráfica, con la que es muy sencillo realizar cálculos, obtener diagramas y efectuar cambios en los datos.

Veamos otro ejemplo:

La siguiente tabla, procedente de la Oficina d'Estadística del Ajuntament de València, muestra la evolución de la tasa de paro durante los años 1997 y 1998. Vamos a utilizar estos datos para comparar la evolución del paro en la ciudad de Valencia con la de la Comunidad Valenciana y la estatal.

**EVOLUCIÓN DE LA TASA DE PARO. 1997–1998**  
(Institut Nacional d'Ocupació, Delegació Provincial de València. Oficina d'Estadística Ajuntament de València)

		<b>Ciutat de València</b>	<b>Comunitat Valenciana</b>	<b>Espanya</b>
<b>1997</b>	<b>Gener</b>	16,1%	14,3%	14,1%
	<b>Febrer</b>	16,4%	14,4%	14,1%
	<b>Març</b>	16,1%	14,2%	13,9%
	<b>Abril</b>	16,0%	14,2%	13,6%
	<b>Maig</b>	15,6%	13,9%	13,3%
	<b>Juny</b>	15,4%	13,8%	13,1%
	<b>Juliol</b>	15,0%	13,2%	12,5%
	<b>Agost</b>	14,7%	13,3%	12,4%
	<b>Setembre</b>	15,0%	13,3%	12,7%
	<b>Octubre</b>	15,0%	13,1%	12,9%
	<b>Novembre</b>	14,8%	12,8%	12,9%
	<b>Desembre</b>	14,2%	12,5%	12,8%
<b>1998</b>	<b>Gener</b>	14,4%	12,6%	12,9%
	<b>Febrer</b>	15,0%	12,6%	12,8%
	<b>Març</b>	14,7%	12,4%	12,6%
	<b>Abril</b>	14,3%	12,2%	12,1%
	<b>Maig</b>	14,1%	12,1%	11,8%
	<b>Juny</b>	13,9%	11,9%	11,5%
	<b>Juliol</b>	13,5%	11,4%	11,0%
	<b>Agost</b>	13,4%	11,7%	10,9%
	<b>Setembre</b>	13,3%	11,3%	11,0%
	<b>Octubre</b>	13,3%	11,0%	11,1%
	<b>Novembre</b>	12,8%	10,5%	11,0%
	<b>Desembre</b>	12,0%	10,1%	10,9%

Introducimos los datos en las listas: TEM, CIU, COM y EST. La primera indicará el mes y tomará valores de 1 a 24 (los 12 primeros son los meses de 1997 y el resto los de 1998). La segunda, tercera y cuarta listas indican, respectivamente la tasa de paro en Valencia ciudad, la de la Comunidad Valenciana y la tasa estatal.

Podemos obtener los parámetros estadísticos de cada una de las listas que indican las tasas de paro por el procedimiento habitual:

1-Var Stats LCIU	1-Var Stats $\bar{x}=14.54166667$ $\Sigma x=349$ $\Sigma x^2=5104.06$ $S_x=1.123239363$ $\sigma_x=1.09958957$ $\downarrow n = 24$	1-Var Stats $\uparrow n = 24$ $\min X=12$ $Q_1=13.7$ $\text{Med}=14.7$ $Q_3=15.2$ $\max X=16.4$
1-Var Stats LCOM	1-Var Stats $\bar{x}=12.61666667$ $\Sigma x=302.8$ $\Sigma x^2=3854.44$ $S_x=1.217862228$ $\sigma_x=1.192220151$ $\downarrow n = 24$	1-Var Stats $\uparrow n = 24$ $\min X=10.1$ $Q_1=11.8$ $\text{Med}=12.6$ $Q_3=13.55$ $\max X=14.4$
1-Var Stats LEST	1-Var Stats $\bar{x}=12.4125$ $\Sigma x=297.9$ $\Sigma x^2=3722.99$ $S_x=1.048938387$ $\sigma_x=1.026853$ $\downarrow n = 24$	1-Var Stats $\uparrow n = 24$ $\min X=10.9$ $Q_1=11.3$ $\text{Med}=12.65$ $Q_3=13$ $\max X=14.1$

Observamos que la tasa de paro en la ciudad de Valencia, durante el período 97–98, es, por término medio, superior a la de la Comunidad Valenciana y ésta es ligeramente superior a la estatal. Las dispersiones en torno a la media son bastante parecidas en los tres casos.

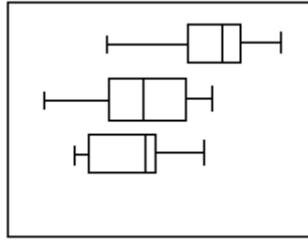
Para comparar las tres tasas en cuanto a sus medianas y rangos intercuartílicos es mejor utilizar los diagramas de caja. Para ello definimos los tres gráficos estadísticos posibles de la siguiente forma:

Plot1: On. Type: diagrama de caja con outliers; Xlist= CIU; Freq=1.

Plot2: On. Type: diagrama de caja con outliers; Xlist= COM; Freq=1.

Plot3: On. Type: diagrama de caja con outliers; Xlist= EST; Freq=1.

Una vez eliminadas todas las expresiones del menú de funciones Y =, pulsamos ZOOM 9 y obtenemos lo siguiente:



Pulsando TRACE y utilizando las teclas de cursor, observamos que, efectivamente, el primer diagrama de caja, correspondiente a la tasa de paro en la ciudad de Valencia presenta muchas diferencias con los otros dos diagramas. Esto parece confirmar que la tasa de paro en la ciudad es superior.

Como en realidad cada una de las tres listas CIU, COM y EST es una serie temporal, interesa otro tipo de diagrama que muestre claramente las diferencias entre ellas. El gráfico más apropiado es el diagrama de líneas, que también podemos obtener con la calculadora gráfica.

Pulsamos 2nd Y= y elegimos la opción 4: PlostsOff. A continuación activamos cada uno de los tres gráficos estadísticos con las siguientes definiciones:

Plot1: On. Type: diagrama de líneas. Xlist=TEM; Ylist=CIU. Mark = •

Plot2: On. Type: diagrama de líneas. Xlist=TEM; Ylist=COM. Mark = •

Plot3: On. Type: diagrama de líneas. Xlist=TEM; Ylist=EST. Mark = •

Pulsando ZOOM 9 obtenemos lo siguiente:



Observamos claramente una tendencia decreciente en las tres series, con independencia de que localmente se rompa esta tendencia en algunos meses. Pero hay una serie que se mantiene siempre por encima de las otras dos.

Al pulsar TRACE y las teclas de desplazamiento del cursor, observamos que, efectivamente, la gráfica de la tasa de paro en la ciudad de Valencia se mantiene siempre por encima de las otras dos. Por tanto, en el período 97–98, la tasa de paro de la ciudad es superior a la media de la Comunidad y a la media estatal. Otro problema diferente sería averiguar las causas de estas diferencias en las tasas de paro. Pero para ello necesitaríamos otro tipo de datos...

## ESTADÍSTICA DESCRIPTIVA CON LA CALCULADORA GRÁFICA CLASSPAD 300 CASIO

### Introducción

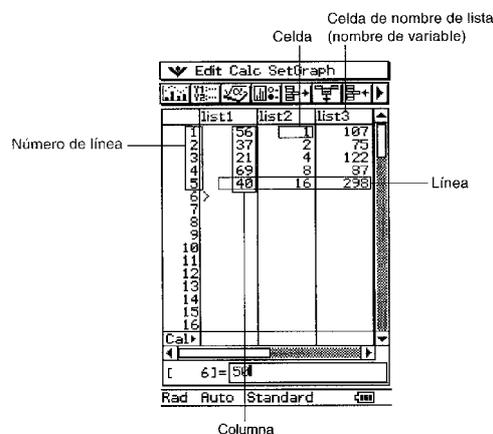
Para adquirir una visión crítica de las fuentes de información y desarrollar capacidades como la de búsqueda, selección, control y evaluación de la información, es necesario realizar experiencias que permitan conocer y aplicar técnicas matemáticas adecuadas y recursos tecnológicos apropiados. La ClassPad 300 permite generar números aleatorios y hacer simulaciones, seleccionar muestras, representar los datos, calcular medidas de centralización y dispersión, ajustar una curva a un conjunto de datos estadísticos, etc.

A continuación estudiaremos algunas de las posibilidades de la ClassPad 300 para el estudio de la Estadística en ESO y Bachillerato.

### 1. Estadística descriptiva y análisis de regresión

#### a) LA APLICACIÓN ESTADÍSTICA

- En el menú de aplicaciones toca el botón . Se inicia la aplicación Estadística y se muestra la ventana del editor de listas.



- En la siguiente tabla tienes los comandos disponibles en los menús y botones del editor de listas:

Función	Botón	Acción
Edit /Abrir lista	—	Abre una lista existente
Edit / Cerrar lista	—	Cierra la lista seleccionada
Edit / Ir al principio	—	Salta a la línea 1 de la lista actual
Edit / Ir al final	—	Salta a la línea siguiente de la última línea en la lista actual.

Edit / Ordenar (Asc.)		Ordena los datos de una lista en orden ascendente
Edit / Ordenar (Desc.)		Ordena los datos de una lista en orden descendente
Edit / Borrar / Celda		Borra una celda
Edit / Borrar / Columna		Borra todos los datos de una lista
Edit / Borrar / Variable lista	—	Borra una lista de la memoria
Edit / Insertar celda		Inserta una celda en una lista

Función	Botón	Acción
		Convierte una expresión matemática a un valor.
		Dibuja un gráfico estadístico.
		Muestra la ventana del editor de gráficos.
		Muestra la ventana de la aplicación Principal
/Preferencias / Ventana vis.		Muestra el cuadro de diálogo de la ventana de visualización.
/Preferencias / Adm. de variable		Muestra el administrador de variables.
ConfGraf / Opciones		Muestra el cuadro de diálogo de configuración de gráficos estadísticos.
		Muestra dos columnas en la ventana del editor de listas
		Muestra tres columnas en la ventana del editor de listas
		Muestra cuatro columnas en la ventana del editor de listas

- La barra de estado muestra la configuración del editor de listas, indicando las unidades de ángulos, el tipo de ventana de estadística y la configuración del cálculo decimal.



**b) TRABAJANDO CON LISTAS**

- En la ventana del editor de listas, toca la celda del nombre de lista en la parte superior de la lista list1. De esta forma se selecciona la celda de nombre de lista.
- Introduce el nombre de lista (hola), de ocho bytes como máximo, y pulsa [EXE]. Si introduces un nombre de lista existente, aparecerán los datos de dicha lista. Si introduces un nombre de lista sin indicar carpeta, el nombre de variable se almacena en la carpeta actual.
- En la ventana del editor de listas, selecciona la celda “Cal” de la lista “hola” e introduce la lista {1,2,3} utilizando para ello el teclado virtual [mth]. Toca el botón [EXE].

	list1	list2	list3
1	1		
2	2		
3	3		
4	4		
5	5		
Cal=	"{1,2,3}"		
Cal=	{1,2,3}		

- Multiplica el valor de cada celda de la lista "hola" por 2 y escribe los resultados en la lista list2. Para ello:
  - Selecciona la celda "Cal" de la lista list2.
  - En el cuadro "Cal=" introduce la fórmula  $hola \times 2$  y pulsa [EXE] para efectuar el cálculo. Aparecen en la list2 los valores de la lista hola multiplicados por 2.

	list1	list2	list3
1	1		
2	2		
3	3		
4	4		
5	5		
Cal=	"{1,2,3}"		
Cal=	{1,2,3}		

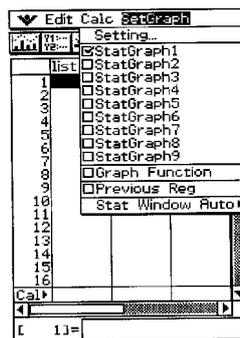
	list1	list2	list3
1	1	2	
2	2	4	
3	3	6	
4	4		
5	5		
Cal=	"{1,2,3}"	"list1"	
Cal=		list1×2	

- Selecciona la celda de nombre de la lista list3 y cambia el nombre por "num3", utilizando el teclado virtual. A continuación, toca la celda "Cal" de dicha lista e introduce la lista {1, 2, 3, 4, 5, 6, 7, 8, 9, 10} utilizando el teclado virtual [mth].
- Selecciona la celda 6 de la lista num3. Selecciona el comando Edit/Ir al principio.
- Selecciona la celda 4 de la lista num3. Selecciona el comando Edit/Ir al final.
- Selecciona la celda 5 de la lista num3. Selecciona el comando Edit / Borrar / Celda. Se borra dicha celda y las situadas por debajo se desplazan hacia arriba.
- Selecciona la celda 6 de la lista num3. Selecciona el comando Edit / Insertar celda. En la nueva celda introduce el valor 12 y pulsa [EXE].
- De la misma forma, selecciona la celda 9 de la lista num3 e inserta una nueva celda. Introduce en ella el valor 14 y pulsa [EXE].
- Selecciona el comando Edit / Ordenar (Asc.). En el siguiente cuadro de diálogo selecciona 1 lista y pulsa el botón [Acep.]. En la lista desplegable selecciona la lista main/num3 y toca el botón [Acep.]. La lista se ordena de menor a mayor.
- De la misma forma, selecciona el comando Edit / Ordenar (Desc.) y en el siguiente cuadro de diálogo selecciona 1 lista y toca el botón [Acep.]. En la lista desplegable, selecciona la lista main/num3 y toca el botón [Acep.] para ordenar la lista en orden decreciente.
- Con el cursor en una celda cualquiera de la lista num3, selecciona el comando Edit / Cerrar lista. La lista desaparece de pantalla y se reemplaza por una columna en blanco.
- Con el cursor en cualquier celda de la lista list1, selecciona el comando Edit / Abrir lista. En el cuadro List= situado en la parte inferior de la pantalla introduce el nombre de la lista, num3, usando para ello el menú virtual [abc]. Toca el botón [EXE] para ver la lista.
- Selecciona la celda de nombre de la lista num3 y toca el cuadro "list=" en la parte inferior de la ventana del editor de listas. Pulsa la tecla [CLEAR] para que se borre la lista. De esta forma, se cierra la lista seleccionada.

- Selecciona la celda de nombre de la lista list3, pulsa [KEYBOARD] y, con ayuda del teclado virtual [abc], escribe el nombre de la lista num3 en la celda de nombre de lista. Toca el botón [EXE] y observa como se abre la lista num3.
- Selecciona la celda de nombre de la lista list5. En el cuadro “Calc=” introduce la fórmula  $3 \cdot \text{num3} + 4$  y pulsa [EXE] para efectuar el cálculo. Aparecen los resultados en la lista list5.
- Selecciona el comando Edit / Ordenar (Asc.). En el siguiente cuadro de diálogo, selecciona 2 listas y pulsa el botón [Acep.]. En la lista desplegable que aparece a continuación del mensaje “Seleccionar lista de base”, selecciona el nombre num3 y pulsa el botón [Acep.]. En la lista desplegable que aparece a continuación del mensaje “Seleccionar segunda lista”, selecciona list5 y pulsa el botón [Acep.]. Observa como se ordenan las listas seleccionadas.
- Pulsa sucesivamente los botones para cambiar el número de listas visibles en la pantalla del editor de listas. Observa que se pueden visualizar dos, tres o cuatro listas. Pulsa el botón [▶] o el cuadro de desplazamiento horizontal situado en la parte inferior de la pantalla para visualizar más listas. Pulsa el botón [▼] o el cuadro de desplazamiento vertical situado a la derecha de la pantalla para visualizar más valores de cada una de las listas.
- Selecciona la lista list5 y elige el comando Edit / Borrar / Columna. En el siguiente cuadro de diálogo toca el botón [Acep.]. De esta forma se borran los datos de la lista y ésta permanece vacía en la memoria.
- Selecciona la lista num3 y elige el comando Edit / Borrar / Variable Lista. En el siguiente cuadro de diálogo toca el botón [Acep.]. De esta forma se borra la lista de la memoria.
- Selecciona el comando Edit / Borrar todo. En el siguiente cuadro de diálogo, toca el botón [Acep.]. De esta forma se borran todos los datos del editor de listas y aparecen las listas 1 a 6 en blanco.

**c) CONFIGURANDO GRÁFICOS ESTADÍSTICOS**

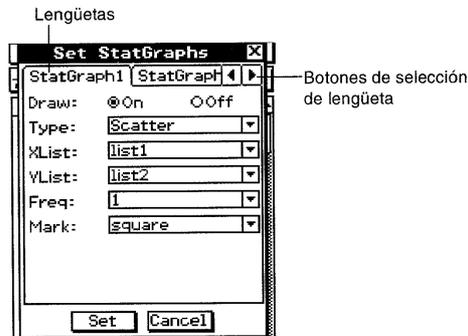
- Toca el menú ConfGraf en la barra de menús del editor de listas. Aparece el menú de configuración de gráficos estadísticos. En la siguiente tabla tienes las opciones de dicho menú.



Función	Acción
Opciones...	Muestra un cuadro de diálogo para indicar el tipo de gráfico y la lista de datos para cada configuración de gráfico estadístico
Opciones... / <input checked="" type="checkbox"/> Gráfico estadís. 1 a <input checked="" type="checkbox"/> Gráfico estadís. 9	Selecciona una configuración de gráfico estadístico
<input checked="" type="checkbox"/> Función gráfica	Superpone un gráfico de función sobre un gráfico estadístico
<input type="checkbox"/> Función gráfica	Desactiva la superposición de un gráfico de función.
<input checked="" type="checkbox"/> Regresión previa	Representa gráficamente los resultados del último cálculo de regresión realizado.
<input type="checkbox"/> Regresión previa	Desactiva la representación gráfica de los últimos resultados de

	cálculo de regresión.
Vent. estad. autom. / On	Obtiene una configuración automática de las opciones de la ventana de visualización de estadísticas.
Vent. estad. autom. / Off	Configura las opciones de la ventana de visualización de estadísticas manualmente.

- En la ventana del editor de listas, selecciona el comando ConfGraf / Opciones... Aparece el cuadro de diálogo de configuración de gráficos estadísticos. En él hay una ficha para cada gráfico estadístico, denominadas Gráfico estadís. 1 a Gráfico estadís. 9.



- Toca la solapa de cada ficha para configurar cada gráfico estadístico. Toca los botones de selección On Off para activar o desactivar cada uno de los gráficos estadísticos.
- En la ListaX, selecciona el nombre de la lista (list1 a list6, o un nombre de lista) a usar como datos del eje X. La opción por defecto es list1.
- En la ListaY, selecciona el nombre de la lista (list1 a list6, o un nombre de lista) a usar como datos del eje Y. La opción por defecto es list2.
- En la lista Frec., selecciona 1 si todos las parejas (x, y) tienen frecuencia 1, o selecciona la lista (list1 a list6 o nombre de lista) que contiene las frecuencias de cada pareja (x, y).
- Configura las opciones de gráficos estadísticos, eligiendo el tipo de gráfico de la lista desplegable Tipo. Los tipos disponibles se muestran en la siguiente tabla:

Tipo	Gráfico
Disper.	Dispersión
Línea xy	Línea xy
GrafPNormal	Gráfico de probabilidad normal
Histogr.	Histograma
RegrLin	Gráfico de regresión lineal
RegrCuadr	Gráfico de regresión cuadrática
RegrCuart	Regresión de orden cuatro
RegrExp.	Regresión exponencial ( $y = a \cdot e^{b \cdot x}$ )
RegrPot	Regresión potencial
RegrLogis	Regresión logística

Tipo	Gráfico
CajaMed	Diagrama de caja Mediana
CajaMod	Diagrama de caja modificado
DistNor	Curva de distribución normal
L. Trazos	Línea a trazos
MedMed	Gráfico Med-Med
RegrCúbic	Regresión cúbica
RegrLog	Regresión logarítmica
Rexp.ab	Regresión exponencial ( $y = a \cdot b^x$ )
RegrSin	Regresión sinusoidal

- En la lista Marca, selecciona la forma que deben tener los puntos del gráfico. Pueden ser: cuadrado (■), cruz (×), punto grueso (●) y punto (•).
- Finalmente, toca el botón [Def.] para definir la configuración de los gráficos estadísticos.

**d) GRÁFICOS ESTADÍSTICOS DE UNA VARIABLE**

- Una compañía de seguros quiere realizar un estudio sobre la esperanza de vida de los españoles para ajustar sus cuotas de seguros. Para ello contrata los servicios de una

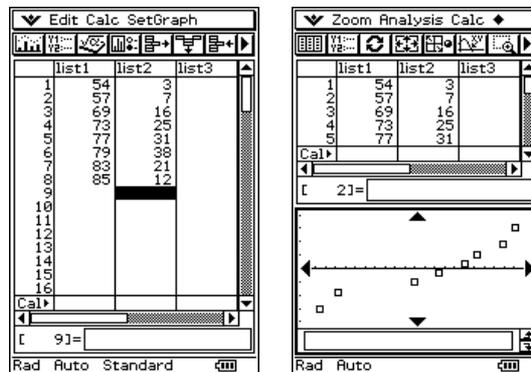
empresa de investigación, que inicia el estudio con una muestra de 153 individuos, que resultaron tener una esperanza de vida  $X$  en años expresada en la siguiente tabla:

$X$	54	57	69	73	77	79	83	85	
$F$	3	7	16	25	31	38	21	12	$\sum F = 153$

Dibuja el gráfico de probabilidad normal, el histograma, el diagrama de cajas Med, el diagrama de cajas modificado, la curva de distribución normal y el gráfico de línea a trozos.

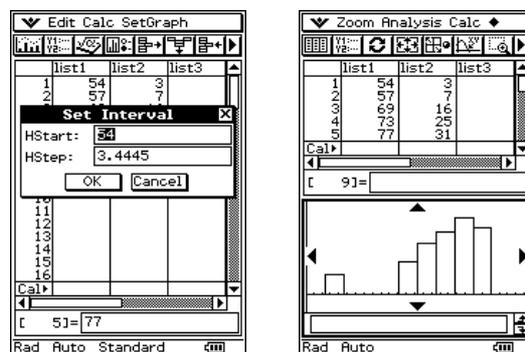
• **Gráfico de probabilidad normal**

- En el editor de listas, introduce los valores de  $X$  en la lista list1 y las frecuencias en la lista list2.
- A continuación, selecciona el comando ConfGraf / Opciones... En la siguiente ventana selecciona el tipo de gráfico GrafPNormal, ListaX = list1, Marca =  cuadrado. Toca el botón [Def.]. Toca el botón  para dibujar el gráfico. Observa el resultado.



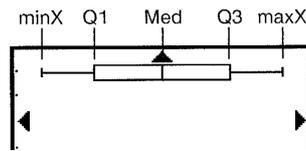
• **Histograma**

- Selecciona el comando ConfGraf / Opciones... En la siguiente ventana selecciona el tipo de gráfico Histograma, ListaX = list1, Freq = list2. Toca el botón [Def.]. Toca el botón  para dibujar el gráfico. Pulsa el botón [Acep.] para aceptar las dimensiones del intervalo. Observa el resultado.

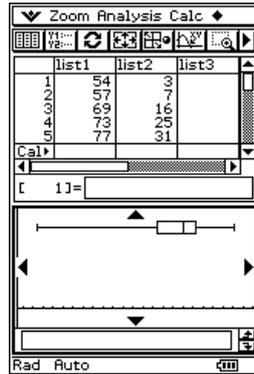


• **Diagrama de cajas Med**

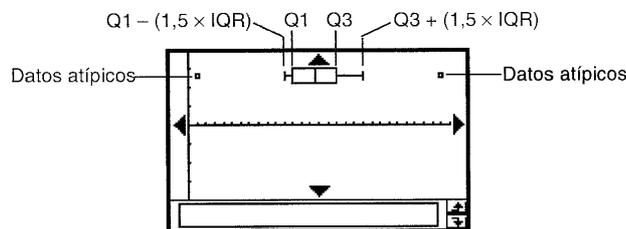
- Este gráfico se llama también “de cajas y bigotes” y representa los valores mínimo (minX) y máximo (maxX), así como los cuartiles  $Q_1$ ,  $Q_3$  y la mediana Med, tal como se indica en la siguiente figura:



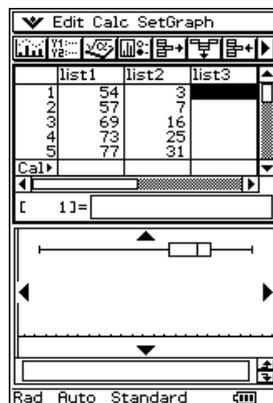
- Selecciona el comando ConfGraf / Opciones... En la siguiente ventana selecciona el tipo de gráfico CajaMed., ListaX = list1, Freq = list2. Toca el botón [Def.]. Toca el botón  para dibujar el gráfico.



- **Diagrama de cajas modificado**
- Este gráfico utiliza la fórmula  $1,5 \times RI$ , siendo  $RI = \text{rango intercuartílico} = Q_3 - Q_1$ , para mostrar los valores alejados (outliers) que son aquellos valores que no siguen el mismo patrón que el resto de datos. Si un valor es menor que  $Q_1 - 1,5 \times RI$ , se dice que es un valor alejado por la izquierda. Si un valor es mayor que  $Q_3 + 1,5 \times RI$ , se dice que es un valor alejado por la derecha. Los valores alejados se representan por círculos en el diagrama.



- Selecciona el comando ConfGraf / Opciones... En la siguiente ventana selecciona el tipo de gráfico CajaMod., ListaX = list1, Freq = list2. Toca el botón [Def.]. Toca el botón  para dibujar el gráfico.

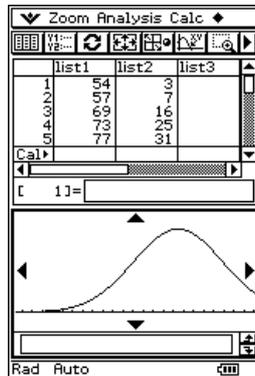


- **Curva de distribución normal**

- La curva de distribución normal se representa gráficamente utilizando la función de

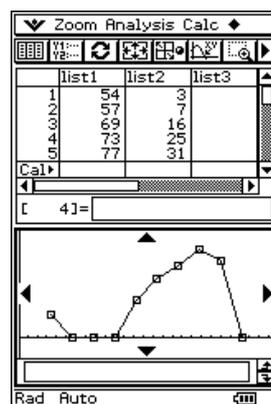
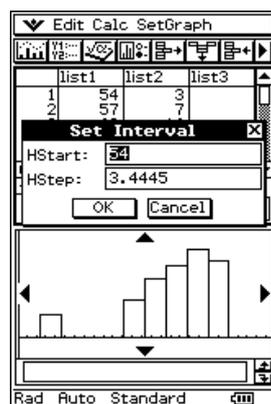
distribución normal siguiente:  $y = \frac{1}{\sqrt{2\pi}\sigma} \cdot e^{-\frac{(x-\mu)^2}{2\sigma^2}}$

- Selecciona el comando ConfGraf / Opciones... En la siguiente ventana selecciona el tipo de gráfico DistNor., ListaX = list1, Freq = list2. Toca el botón [Def.]. Toca el botón  para dibujar el gráfico.



- **Gráfico de Línea a trazos**

- En este gráfico se conectan mediante segmentos los puntos centrales de cada barra del histograma.
- Selecciona el comando ConfGraf / Opciones... En la siguiente ventana selecciona el tipo de gráfico L.Trazos, ListaX = list1, Freq = list2. Toca el botón [Def.]. Toca el botón  para dibujar el gráfico. Toca el botón [Acep.] para aceptar las dimensiones del intervalo. Observa el resultado.



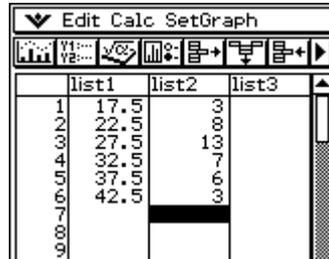
e) **CÁLCULOS ESTADÍSTICOS**

- **Cálculos estadísticos de una variable**
- Se ha aplicado un test de inteligencia a 40 estudiantes, obteniéndose los siguientes resultados, agrupados en intervalos:

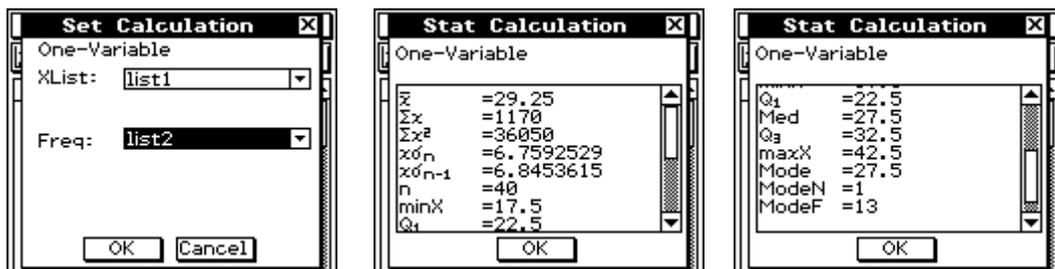
Puntuaciones	[15-20)	[20-25)	[25, 30)	[30, 35)	[35, 40)	[40, 45)
Nº de alumnos	3	8	13	7	6	3

Calcula los parámetros estadísticos: media, varianza, desviación típica, mínimo, máximo, primer cuartil, mediana, tercer cuartil, moda. Sigue los siguientes pasos:

- 1) En la ventana del editor de listas, introduce las marcas de clase en la lista list1 y las frecuencias en la list2.



- 2) En la barra de menús, selecciona el comando Calc / Una variable.
- 3) En el cuadro de diálogo que aparece, selecciona el nombre list1 como ListaX y selecciona el nombre list2 como Frec y toca el botón [Acep.].



- 4) Aparece una ventana con los valores de los parámetros estadísticos, que son los siguientes:

$\bar{x}$	Media	$Q_1$	Primer cuartil
$\Sigma x$	Suma de datos	Med	Mediana
$\Sigma x^2$	Suma de cuadrados	$Q_3$	Tercer cuartil
$x\sigma_n$	Desviación típica poblacional	MaxX	Máximo
$x\sigma_{n-1}$	Desviación típica muestral	Mode	Moda
N	Tamaño muestral	ModeN	Número de elementos iguales a la moda
MinX	mínimo	ModeF	Frecuencia de la moda