MATEMÁTICAS B

4.º de Educación Secundaria Obligatoria

Unidad 9

Estadística unidimensional y bidimensional

Manual del alumnado

Curso 2026–2027

Edición V01 · Índice cerrado el 30 de octubre de 2026

Índice

1 El estudio estadístico 3
1.1 De una pregunta a una conclusión . . . 3
1.2 Variables y selección . . . 3
1.3 Practica y razona . . . 3

2 Tablas de frecuencias 4
2.1 Frecuencias y comprobación . . . 4
2.2 Intervalos y marcas de clase . . . 4
2.3 Practica y razona . . . 4

3 Representaciones gráficas 5
3.1 Barras, sectores e histogramas . . . 5
3.2 Pictogramas: símbolos con escala constante . . . 5
3.3 Mapas por comunidades autónomas . . . 5
3.4 Lectura crítica . . . 7
3.5 Practica y razona . . . 7

4 Medidas de centralización 9
4.1 Media, mediana y moda . . . 9
4.2 Aproximaciones con clases . . . 9
4.3 Practica y razona . . . 9

5 Medidas de posición y diagramas de caja 10
5.1 Convenciones explícitas . . . 10
5.2 Cajas y posibles valores atípicos . . . 10
5.3 Practica y razona . . . 10

6 Medidas de dispersión 12
6.1 Describir la variabilidad . . . 12
6.2 Seleccionar la salida de la calculadora . . . 12
6.3 Comparar con sentido . . . 12
6.4 Practica y razona . . . 12

7 Distribuciones bidimensionales **13**
7.1 Pares, tablas y marginales . . . 13
7.2 Centro de gravedad y tipo de relación . . . 13
7.3 Practica y razona . . . 13

8 Covarianza y coeficiente de Pearson **15**
8.1 Resultados de calculadora e interpretación . . . 15
8.2 Excepciones y causalidad . . . 15
8.3 Practica y razona . . . 15

9 Recta de regresión y predicciones **16**
9.1 Obtener la recta con CASIO . . . 16
9.2 Fiabilidad orientativa según Pearson . . . 16
9.3 Practica y razona . . . 16

10 Calculadora y herramientas digitales **17**
10.1 Una variable y frecuencias . . . 17
10.2 Dos variables y regresión lineal . . . 17
10.3 Hoja de cálculo y gráficos . . . 17
10.4 Practica y razona . . . 17

11 Problemas integrados y pequeño estudio estadístico **18**
11.1 Una investigación pequeña y defendible . . . 18
11.2 Pictograma o mapa como producto . . . 18
11.3 Comunicar con prudencia . . . 18
11.4 Practica y razona . . . 18

A Formulario y convenciones de consulta **19**
A.1 Frecuencias y datos agrupados . . . 19
A.2 Centralización y posición . . . 19
A.3 Dispersión . . . 19
A.4 Dos variables: consulta de calculadora . . . 19

B Práctica global **20**

C Autoevaluación de 20 preguntas **21**

Capítulo 1

El estudio estadístico

1.1 De una pregunta a una conclusión

Un estudio comienza con una pregunta concreta, identifica población y variable, elige cómo recoger datos, los organiza, los representa y analiza, y concluye indicando limitaciones. La población es el conjunto de interés; cada elemento es un individuo y una muestra es el subconjunto observado. Describir una muestra no garantiza que represente a toda la población.

1.2 Variables y selección

Las cualitativas nominales son categorías sin orden; las ordinales tienen orden. Las cuantitativas discretas cuentan valores aislados y las continuas miden magnitudes que pueden variar en intervalos. Un código numérico no convierte una categoría en cantidad.

Ejemplo resuelto

Código postal: cualitativa nominal; satisfacción baja/media/alta: ordinal; número de hermanos: discreta; tiempo de carrera: continua.

Una muestra grande puede estar sesgada si se selecciona solo a quienes están disponibles o responden voluntariamente. Define a quién excluye la recogida, evita preguntas que sugieran la respuesta y registra las unidades. En investigaciones escolares se pueden usar datos anónimos o conjuntos didácticos; no hacen falta nombres.

1.3 Practica y razona

U09-01

Se pregunta por hábitos de lectura solo a quienes están en la biblioteca. Identifica el posible sesgo.
(Respuesta: Selección de un grupo posiblemente más lector.)

U09-02

Clasifica deporte favorito, talla S/M/L, número de faltas y estatura.
(Respuesta: Nominal, ordinal, discreta y continua.)

U09-03

Se encuestan 80 estudiantes de un centro de 600. Distingue población, muestra e individuo.
(Respuesta: Población: 600 estudiantes; muestra: 80; individuo: cada estudiante.)

Resumen del capítulo

Precisar población, variable, recogida y límites de la conclusión antes de calcular.

Capítulo 2

Tablas de frecuencias

2.1 Frecuencias y comprobación

La frecuencia absoluta f_i cuenta observaciones, $h_i = f_i/N$ es la relativa y $100h_i$ el porcentaje. En variables ordenables, F_i y H_i acumulan las frecuencias hasta ese valor. Se verifica $\sum f_i = N$, $\sum h_i = 1$. No se acumulan categorías nominales arbitrariamente ordenadas.

x_i	f_i	h_i	F_i	H_i
0	2	0,2	2	0,2
1	4	0,4	6	0,6
2	3	0,3	9	0,9
3	1	0,1	10	1

2.2 Intervalos y marcas de clase

Todos los intervalos se escriben $[L, U)$: incluyen L y excluyen U. Un dato situado en una frontera interior pertenece al intervalo que comienza allí. Si un valor coincide con el extremo final, se amplía la cobertura añadiendo otra clase; no se cierra por la derecha la última clase. La amplitud es $U - L$ y la marca de clase su punto medio $c = (L + U)/2$. Para $[0, 10)$, $[10, 20)$ y $[20, 30)$, las marcas son $5, 15, 25$. Las frecuencias $3, 5, 2$ suman diez y las acumuladas son $3, 8, 10$. Las marcas permiten aproximar parámetros; no revelan los datos exactos de cada clase.

2.3 Practica y razona

U09-04

En la tabla del capítulo, ¿cuántos datos son menores o iguales que 1 y qué porcentaje supera 1?
(Respuesta: 6 y 40%.)

U09-05

Obtén marcas y amplitudes de $[0, 10)$, $[10, 20)$ y $[20, 30)$. Asigna 10 y 20.
(Respuesta: Marcas $5, 15, 25$; amplitudes 10; 10 en la segunda y 20 en la tercera.)

U09-06

¿Dónde colocarías un dato de valor 30 si las clases terminan en $[20, 30)$?
(Respuesta: En una nueva clase, por ejemplo $[30, 40)$.)

Resumen del capítulo

Intervalos siempre cerrados a la izquierda y abiertos a la derecha; marcas como puntos medios.

Capítulo 3

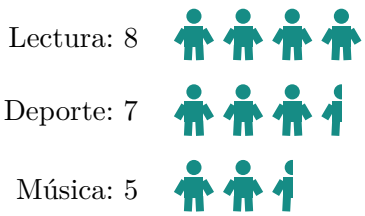
Representaciones gráficas

3.1 Barras, sectores e histogramas

Las barras separadas representan categorías o valores discretos. En sectores, cada ángulo es $360°h_i$. Para variables continuas se usan histogramas con rectángulos contiguos; con amplitudes diferentes, la altura correcta es $f_i/(U_i - L_i)$ para que el área represente frecuencia. Con amplitudes iguales puede usarse directamente frecuencia como altura, indicando el eje. El polígono une puntos situados sobre las marcas de clase.

3.2 Pictogramas: símbolos con escala constante

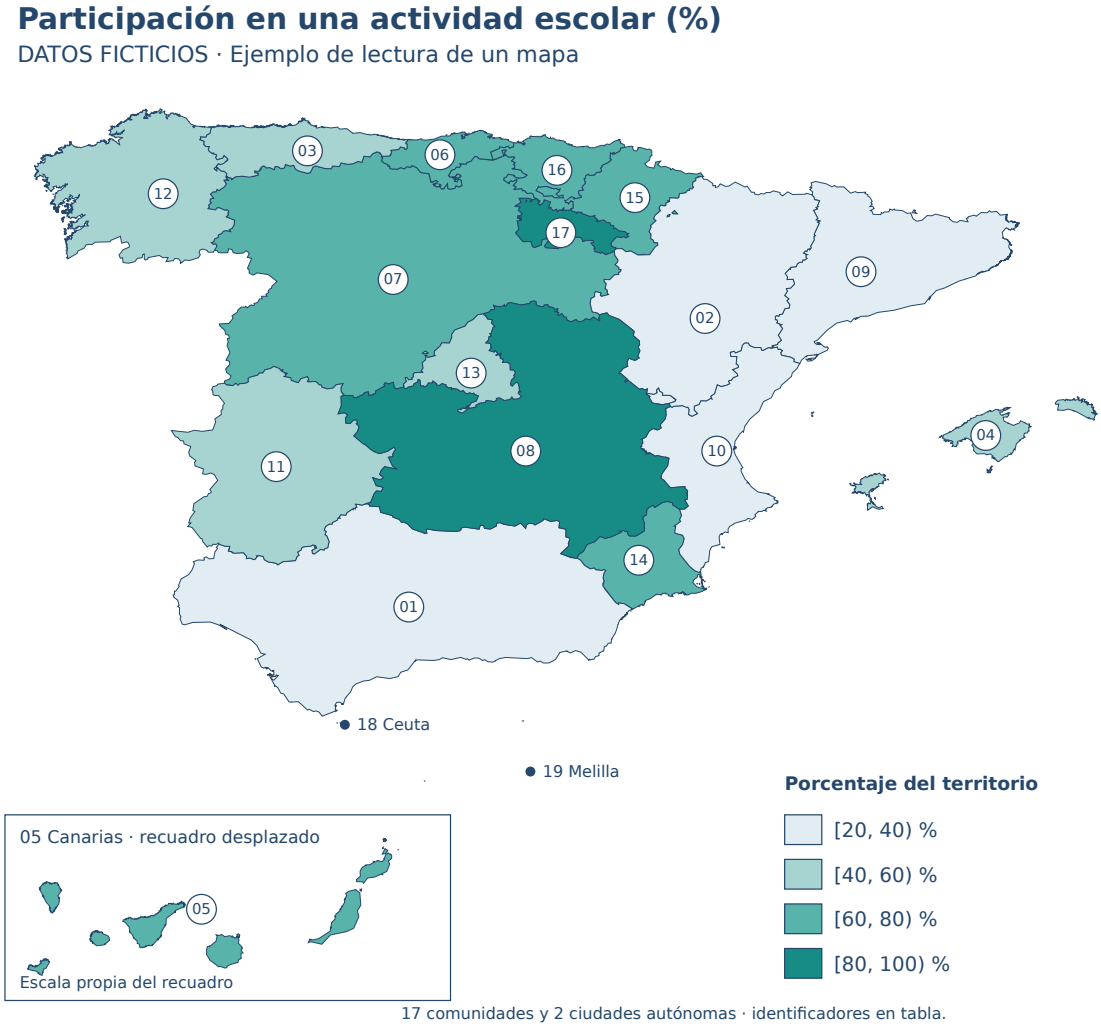
Cada icono debe tener un valor indicado y conservar el mismo tamaño. Si una persona dibujada equivale a dos observaciones, media figura equivale a una. No amplíes a la vez alto y ancho para representar el doble: el área se multiplicaría por cuatro. El gráfico incluido representa ocho preferencias por lectura, siete por deporte y cinco por música; son datos didácticos.



Una figura completa representa 2 personas; media figura, 1 persona.

3.3 Mapas por comunidades autónomas

El mapa incluido representa un porcentaje ficticio de participación en un proyecto escolar por comunidad. No son resultados reales ni datos del INE. Los límites geográficos proceden de una base cartográfica identificada; los valores y colores están en la tabla adjunta. Las ciudades de Ceuta y Melilla se distinguen de las diecisiete comunidades. Canarias se muestra en un recuadro desplazado, con escala propia.



Geometría: click_that_hood / spain-communities.geojson. Datos elaborados para estos apuntes.

Porcentajes territoriales: no son recuentos ni se suman para obtener el total de España.

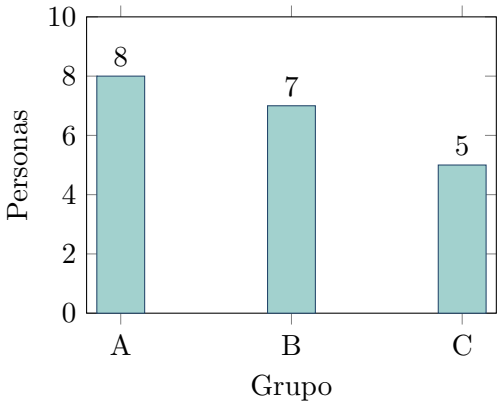
Código	Territorio	Participación (%)
01	Andalucía	32
02	Aragón	39
03	Asturias	46
04	Baleares	53
05	Canarias	60
06	Cantabria	67
07	Castilla y León	74
08	Castilla-La Mancha	81
09	Cataluña	27
10	Comunitat Valenciana	34
11	Extremadura	41
12	Galicia	48
13	Madrid	55
14	Murcia	62
15	Navarra	69
16	País Vasco	76
17	La Rioja	83
18	Ceuta	29
19	Melilla	36

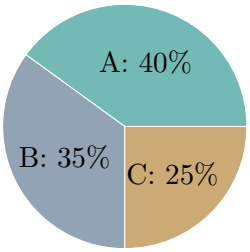
Un mapa de recuentos puede reflejar el tamaño de la población. Para comparar participación, utiliza

porcentaje de la población de referencia; para una tasa, identifica claramente el denominador. El color de una comunidad resume un agregado, no describe a cada habitante.

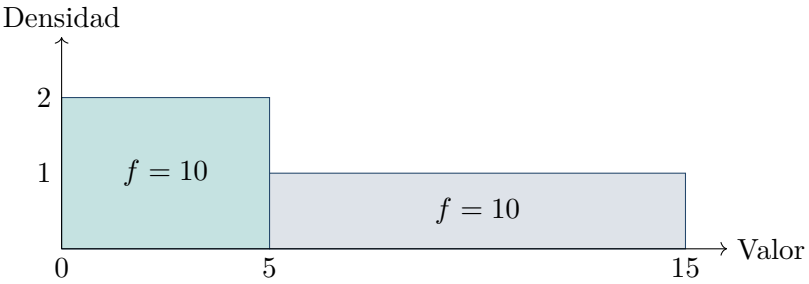
3.4 Lectura crítica

Comprueba título, fuente, fecha o carácter ficticio, unidades, ejes y leyenda. Escalas truncadas e iconos de tamaños distintos pueden exagerar diferencias. Elige el gráfico por la pregunta, no solo por su aspecto.





Barras y sectores del mismo conjunto: $8 + 7 + 5 = 20$.



Clases de amplitudes distintas: área = frecuencia; alturas $10/5 = 2$ y $10/10 = 1$.

3.5 Practica y razona

U09-07

Una encuesta tiene frecuencias $8, 7, 5$ sobre 20. Calcula porcentajes y ángulos de sectores.
(Respuesta: $40\%, 35\%, 25\%$; $144°, 126°, 90°$.)

U09-08

En un pictograma con una figura igual a dos personas, ¿cómo representas siete?
(Respuesta: Tres figuras completas y media.)

U09-09

Clases $[0, 5)$ y $[5, 15)$ tienen diez datos cada una. Calcula alturas de densidad y áreas.
(Respuesta: Alturas 2 y 1; áreas 10 y 10.)

U09-10

Una comunidad registra 500 participantes entre 1000 personas y otra 800 entre 4000. Compara recuentos y porcentajes.

(Respuesta: La segunda tiene más participantes; la primera mayor porcentaje: 50% frente a 20%.)

Resumen del capítulo

Cada representación necesita escala y leyenda; en mapas distinguir cantidades, porcentajes y datos agregados.

Capítulo 4

Medidas de centralización

4.1 Media, mediana y moda

La media aritmética es $\bar{x} = \sum x_j/N$; con tabla, $\bar{x} = \sum f_i x_i/N$. Usar frecuencias es resumir datos repetidos, no un bloque de ponderación por porcentajes o calificaciones. La mediana se obtiene ordenando: dato central si N es impar y promedio de los dos centrales si es par. La moda es el valor o valores de mayor frecuencia; en una cualitativa nominal puede hallarse moda, pero no una media numérica de sus etiquetas.

Ejemplo resuelto

En $2, 2, 4, 8$, media 4, mediana 3 y moda 2. Un valor extremo afecta mucho a la media y menos a la mediana.

4.2 Aproximaciones con clases

Sustituye las observaciones de cada intervalo por su marca y usa la frecuencia: para marcas $5, 15, 25$ y frecuencias $3, 5, 2$, $\bar{x} \approx 14$. Con clases de igual amplitud, el intervalo modal es el de frecuencia máxima; tomamos su marca como moda aproximada. En el ejemplo, $[10, 20)$ y moda aproximada 15. Para evitar ambigüedades, las actividades de moda agrupada usan amplitudes iguales. Los histogramas de amplitud desigual se reservan para lectura de densidades. No se interpolan valores dentro de un intervalo.

4.3 Practica y razona

U09-11

Calcula media, mediana y moda de $2, 2, 4, 8$.
(Respuesta: 4, 3, 2.)

U09-12

Para $x_i = 0, 1, 2, 3$ con $f_i = 2, 4, 3, 1$, halla media, mediana y moda.
(Respuesta: 1,3, 1, 1.)

U09-13

Con clases $[0, 10)$, $[10, 20)$, $[20, 30)$ y frecuencias $3, 5, 2$, estima media y moda.
(Respuesta: Media 14; moda 15.)

Resumen del capítulo

Media usa todos los valores; mediana ordena; moda identifica máxima frecuencia; agrupados dan aproximaciones.

Capítulo 5

Medidas de posición y diagramas de caja

5.1 Convenciones explícitas

$Q_1 = P_{25}$, $Q_2 = P_{50}$, $Q_3 = P_{75}$. Para datos individuales ordenados y $0 < k < 100$, calcula $r = kN/100$: si no es entero, elige el dato de posición $\lceil r \rceil$; si es entero, promedia los datos de posiciones r y $r+1$. Es una convención declarada; otras herramientas pueden dar resultados distintos. En datos agrupados, busca la primera clase cuya acumulada cumpla $F_i \geq kN/100$ y toma su marca como percentil aproximado. Si coincide exactamente con una acumulada, conservamos esa primera clase: no saltamos a la siguiente ni promediamos marcas. Esta convención agrupada se aplica también a la mediana P_{50}. El intervalo percentílico no es necesariamente el modal.

Ejemplo resuelto

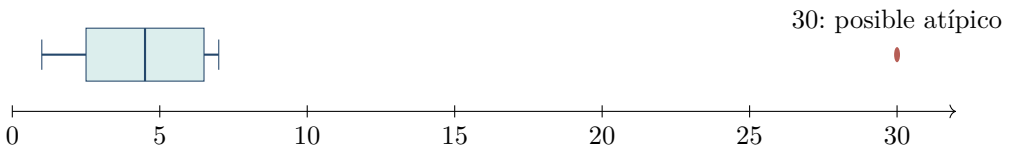
Con frecuencias $3, 5, 2$ en las clases de marcas $5, 15, 25$: $P_{30} \approx 5$ porque la primera acumulada es tres; $P_{50} \approx 15$, $P_{75} \approx 15$ y $P_{90} \approx 25$. Son aproximaciones elegidas, no posiciones exactas de datos desconocidos.

5.2 Cajas y posibles valores atípicos

El rango intercuartílico es $RIC = Q_3 - Q_1$. La caja va de Q_1 a Q_3 y contiene la línea de mediana. En el diagrama modificado, los bigotes llegan a los datos extremos dentro de $[Q_1 - 1{,}5RIC, Q_3 + 1{,}5RIC]$; los exteriores se marcan como posibles atípicos, no se eliminan sin justificación.

Ejemplo resuelto

En $1, 2, 3, 4, 5, 6, 7, 30$, $Q_1 = 2{,}5$, mediana $4{,}5$, $Q_3 = 6{,}5$, $RIC = 4$; barreras $-3{,}5$ y $12{,}5$. Los bigotes llegan a uno y siete; treinta es posible atípico.



5.3 Practica y razona

U09-14

Para $1, 2, 3, 4, 5, 6, 7, 30$, calcula cuartiles, RIC y posibles atípicos.
(Respuesta: 2,5, 4,5, 6,5; RIC 4; posible atípico 30.)

U09-15

En las clases con marcas $5, 15, 25$ y frecuencias $3, 5, 2$, calcula P_{30}, P_{50} y P_{90} según nuestra convención. (Respuesta: 5, 15, 25, aproximadamente.)

U09-16

Dos grupos tienen la misma mediana, pero RIC 2 y 8. ¿Qué indica y qué no permite asegurar? (Respuesta: El segundo tiene mayor amplitud del tramo intercuartílico; no basta para ordenar todos los demás indicadores.)

Resumen del capítulo

Distinguir datos individuales y agrupados; fijar la convención antes de comparar percentiles o cajas.

Capítulo 6

Medidas de dispersión

6.1 Describir la variabilidad

Rango: máximo menos mínimo. Varianza descriptiva $\sigma^2 = \sum f_i(x_i - \bar{x})^2/N$ y desviación típica $\sigma = \sqrt{\sigma^2}$. La varianza tiene unidades al cuadrado; la desviación, las mismas que los datos. Para clases, sustituir x_i por marcas y señalar que es aproximado.

Ejemplo resuelto

En la tabla $x_i = 0, 1, 2, 3$, $f_i = 2, 4, 3, 1$, CASIO debe dar $\bar{x} = 1{,}3$, $\sigma^2 = 0{,}81$, $\sigma = 0{,}9$. En las clases con marcas $5, 15, 25$ y frecuencias $3, 5, 2$, las estimaciones son $\bar{x} = 14$, $\sigma^2 = 49$, $\sigma = 7$.

6.2 Seleccionar la salida de la calculadora

Para describir este conjunto usamos divisor N: salida σ_x o equivalente descriptiva. La salida s_x usa corrección con $N - 1$ y no es intercambiable. Una discrepancia requiere revisar frecuencias y opción seleccionada antes de atribuirla a redondeo.

6.3 Comparar con sentido

Para medias positivas de magnitudes donde tenga sentido comparar razones, $CV = \sigma/\bar{x}$ o $100\sigma/\bar{x}\%$. Con media cero no está definido y con escalas sin cero absoluto o medias negativas puede ser inadecuado. Interpretar media y dispersión conjuntamente; no se desarrolla distribución normal.

6.4 Practica y razona

U09-17

Comprueba con calculadora media, varianza y desviación de $0, 0, 1, 1, 1, 1, 2, 2, 2, 3$.
(Respuesta: 1,3, 0,81, 0,9.)

U09-18

Con marcas $5, 15, 25$ y frecuencias $3, 5, 2$, calcula dispersión aproximada y CV.
(Respuesta: Varianza 49, desviación 7, CV 50%.)

U09-19

Dos grupos tienen medias 20 y 50 y desviaciones 4 y 5, en una magnitud de razón positiva. Compara dispersión absoluta y relativa.
(Respuesta: El segundo tiene mayor desviación absoluta; el primero mayor CV: 20% frente a 10%.)

Resumen del capítulo

Usar desviación descriptiva con divisor N y relacionar dispersión con unidades y contexto.

Capítulo 7

Distribuciones bidimensionales

7.1 Pares, tablas y marginales

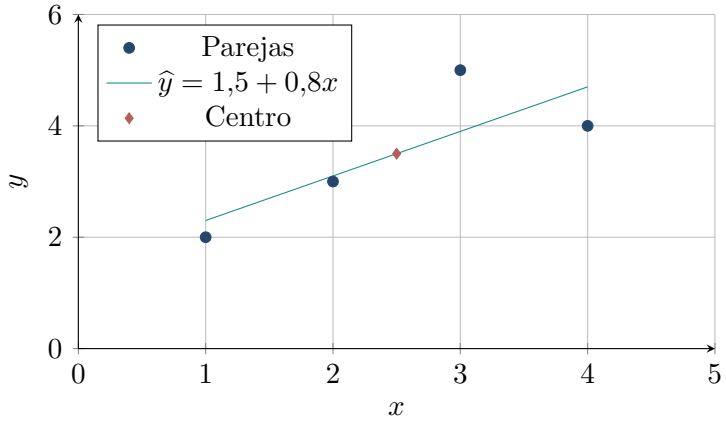
Cada individuo aporta un par (x, y); no se pueden reordenar por separado las columnas porque se perdería el emparejamiento. Una tabla de doble entrada cuenta coincidencias y las marginales suman filas o columnas. En una nube de puntos se conserva cada pareja, incluidas sus repeticiones si las hay.

7.2 Centro de gravedad y tipo de relación

El centro es $(\bar{x}, \bar{y})$. Una relación funcional asigna una única salida a cada entrada; una asociación estadística puede presentar dispersión para la misma entrada. Una nube sugiere relación positiva, negativa, no lineal o poca relación visible; no prueba causalidad.

Ejemplo resuelto

Para $(1,2), (2,3), (3,5), (4,4)$, la calculadora da $\bar{x} = 2{,}5$, $\bar{y} = 3{,}5$. La nube tiene tendencia ascendente con dispersión: no está exactamente sobre una recta.



El trabajo manual se centra en introducir y comprobar cuatro parejas y leer el gráfico. Los cálculos repetitivos de productos y cuadrados se realizan con calculadora.

7.3 Practica y razona

U09-20

Introduce $(1,2), (2,3), (3,5), (4,4)$ en CASIO. Obtén tamaño y centro de gravedad.
(Respuesta: $N = 4$; $(2{,}5, 3{,}5)$.)

U09-21

¿Es válido ordenar la columna y sin reordenar simultáneamente sus parejas de x?
(Respuesta: No.)

U09-22

Una tabla tiene filas $(4, 6)$ y $(3, 7)$. Calcula marginales de filas, columnas y total.
(Respuesta: Filas $10, 10$; columnas $7, 13$; total 20.)

Resumen del capítulo

Mantener los pares, obtener las marginales y leer la nube antes de calcular correlación.

Capítulo 8

Covarianza y coeficiente de Pearson

8.1 Resultados de calculadora e interpretación

La covarianza descriptiva es $\sigma_{xy} = \sum(x_i - \bar{x})(y_i - \bar{y})/N$. También $\sum x_i y_i/N - \bar{x}\bar{y}$. Su signo informa del sentido de la asociación lineal; su magnitud depende de las unidades. No se pide completar tablas largas a mano. Pearson es $r = \sigma_{xy}/(\sigma_x \sigma_y)$ si ambas desviaciones son positivas, y cumple $-1 \leq r \leq 1$. r positivo indica asociación ascendente; negativo, descendente. El módulo próximo a uno indica fuerte ajuste lineal. Próximo a cero indica débil asociación lineal, no necesariamente ausencia de cualquier relación.

Ejemplo resuelto

En las cuatro parejas del capítulo anterior, CASIO da $r = 0{,}8$, desviaciones descriptivas de ambas columnas $\sqrt{1{,}25} \approx 1{,}1180$ y covarianza 1. Si la calculadora no presenta la covarianza directamente, se obtiene como $r\sigma_x\sigma_y$.

8.2 Excepciones y causalidad

Si una columna es constante, su desviación es cero y Pearson no está definido. $r = -0{,}95$ indica mayor intensidad lineal que $r = 0{,}6$: el signo no mide fiabilidad. Una correlación alta puede deberse a factores comunes; no demuestra que modificar una variable cause el cambio de la otra.

8.3 Practica y razona

U09-23

Obtén con CASIO Pearson y covarianza para $(1, 2), (2, 3), (3, 5), (4, 4)$.
(Respuesta: $r = 0{,}8$ y covarianza 1.)

U09-24

Compara $r = -0{,}95$, $r = 0{,}6$ y $r = 0{,}1$ en intensidad y sentido.
(Respuesta: Intensidad mayor en $-0{,}95$, después 0,6, después 0,1; primera descendente y otras ascendentes.)

U09-25

En $(1, 1), (0, 0), (-1, 1)$, Pearson es cero. ¿Implica ausencia de relación?
(Respuesta: No; existe la relación no lineal $y = x^2$.)

Resumen del capítulo

Interpretar el signo y el módulo de Pearson por separado; revisar columnas constantes y no afirmar causalidad.

Capítulo 9

Recta de regresión y predicciones

9.1 Obtener la recta con CASIO

Elegimos regresión lineal $y = a + bx$: a es término independiente, b pendiente. La calculadora produce ambos a partir de parejas. Si una herramienta usa $y = mx + n$, identifica el papel de cada letra: no copies el orden sin leer el modelo. Para comprobar, $b = \sigma_{xy}/\sigma_x^2$ y $a = \bar{y} - b\bar{x}$; la recta pasa por $(\bar{x}, \bar{y})$. Estas fórmulas explican el resultado y permiten una comprobación breve; no se exige el desarrollo manual completo.

Ejemplo resuelto

En $(1,2), (2,3), (3,5), (4,4)$ se obtiene $a = 1{,}5$, $b = 0{,}8$, $r = 0{,}8$. Para $x = 3$, la estimación es $\widehat{y} = 3{,}9$. No tiene que coincidir con el dato observado cinco: la recta resume el conjunto.

9.2 Fiabilidad orientativa según Pearson

Cuanto más próximo esté $|r|$ a uno, mejor ajuste lineal y mayor confianza orientativa en el modelo, consideradas condiciones comparables. El signo indica dirección. No fijamos fronteras universales entre correlación fuerte y débil; una asociación alta no garantiza cada predicción ni prueba causalidad. Aquí no se desarrollan los apartados de interpolación y extrapolación.

9.3 Practica y razona

U09-26

Con las cuatro parejas habituales, obtén recta y estima y para $x = 3$. Comprueba el centro.
(Respuesta: $\widehat{y} = 1{,}5 + 0{,}8x$; 3,9; pasa por $(2{,}5, 3{,}5)$.)

U09-27

La calculadora da $y = 9 - 2x$, $r = -1$. Estima en $x = 2{,}5$ e interpreta el ajuste.
(Respuesta: 4; ajuste lineal perfecto descendente en los datos.)

U09-28

Dos modelos comparables tienen $r = 0{,}2$ y $r = -0{,}9$. ¿Cuál tiene mayor ajuste lineal?
(Respuesta: El segundo.)

Resumen del capítulo

Leer a, b y r en el modelo elegido; predecir con la recta y valorar su ajuste por el módulo de Pearson.

Capítulo 10

Calculadora y herramientas digitales

10.1 Una variable y frecuencias

En fx-991SP CW: HOME, Estadística y una variable; en fx-570SP X II: MENU, Estadística y una variable. Activa la columna de frecuencia cuando introduzcas una tabla. Escribe valores o marcas en x y frecuencias en la columna correspondiente; verifica N antes de leer media y desviación descriptiva. No introduzcas cada marca una sola vez si sus frecuencias son distintas.

10.2 Dos variables y regresión lineal

Elige el modelo lineal $y = a + bx$ e introduce cada par en su fila. En CW, utiliza las opciones de resultados o cálculo de regresión; en X II, las opciones de estadística y regresión. Localiza a, b, r; para la covarianza usa $r\sigma_x\sigma_y$ con desviaciones descriptivas. Conserva las cifras internas hasta el resultado final. Si aparece error, revisa columnas constantes y datos mal emparejados.

10.3 Hoja de cálculo y gráficos

Una hoja con columnas de valores y frecuencias permite organizar sumas, porcentajes y gráficos; una tabla de comunidades debe enlazar el código o nombre correcto con su valor. Para un pictograma conserva escala de los símbolos; para un mapa, documenta límites geográficos, valores y leyenda. GeoGebra permite contrastar nube y recta; las diferencias de cuartiles se revisan según la convención declarada, no se corrigen a ciegas.

10.4 Practica y razona

U09-29

Para marcas $5, 15, 25$ con frecuencias $3, 5, 2$, CASIO devuelve $N = 3$. ¿Qué debes revisar?
(Respuesta: La columna de frecuencias.)

U09-30

Una calculadora muestra s_x y σ_x. ¿Cuál usamos para describir el conjunto de esta unidad?
(Respuesta: σ_x, con divisor N.)

Resumen del capítulo

Introducir bien los datos y seleccionar el estadístico correcto es más importante que repetir cuentas largas.

Capítulo 11

Problemas integrados y pequeño estudio estadístico

11.1 Una investigación pequeña y defendible

En parejas, elegid una pregunta con variable no sensible, definid población y recoged u organizad una muestra anónima. Entregad tabla, representación adecuada, medidas pertinentes y conclusión con una limitación. Cada integrante explicará una decisión y comprobará un cálculo con CASIO.

11.2 Pictograma o mapa como producto

Podéis reutilizar los datos ficticios incluidos o sustituirlos por una fuente pública identificada con fecha y unidades. Conservad tabla y gráfico coherentes; no presentéis los porcentajes ficticios del mapa como estadísticas reales. Justificad por qué elegís tasa o recuento y qué no permite concluir un dato autonómico.

11.3 Comunicar con prudencia

No generalicéis a la población por el mero tamaño muestral ni convirtáis una correlación en causalidad. Distinguid dato, cálculo, interpretación y decisión. Los ejercicios cerrados tienen solución; la investigación abierta se valora por evidencias y justificación.

11.4 Practica y razona

U09-31

Un titular afirma que un programa causa mejora porque participación y resultados tienen $r = 0{,}85$. Señala qué conclusión es excesiva.
(Respuesta: La causalidad no queda demostrada por la correlación.)

U09-32

¿Puede el color alto de una comunidad en el mapa demostrar que cada municipio tiene ese mismo porcentaje?
(Respuesta: No.)

Resumen del capítulo

El informe une pregunta, datos, representación y conclusión con sus límites.

Apéndice A

Formulario y convenciones de consulta

A.1 Frecuencias y datos agrupados

$N = \sum f_i$, $h_i = f_i/N$, $F_i = \sum_{j \leq i} f_j$. Intervalos siempre $[L_i, U_i)$, con marca $c_i = (L_i + U_i)/2$. En un histograma con amplitudes distintas, altura $= f_i/(U_i - L_i)$ y área $= f_i$.

A.2 Centralización y posición

$\bar{x} = \sum f_i x_i / N$; agrupados, sustituir por marcas y expresar como aproximación. Moda agrupada: marca de la clase modal en las actividades con amplitudes iguales. No hay bloque de media ponderada. En datos individuales ordenados y percentil de rango $r = kN/100$: si r no es entero se toma la posición superior; si es entero se promedian las posiciones r y $r+1$ (para $0 < k < 100$). En agrupados, primera clase con $F_i \geq kN/100$ y su marca; una coincidencia queda en esa primera clase. Mediana agrupada: P_{50}. No se realiza ajuste dentro del intervalo.

A.3 Dispersión

$\sigma^2 = \sum f_i x_i^2 / N - \bar{x}^2$, $\sigma = \sqrt{\sigma^2}$; $CV = 100\sigma/\bar{x}$ cuando la escala y una media positiva lo hagan adecuado. Usar desviación descriptiva, divisor N, y distinguir de la corregida con $N - 1$. Rango $= x_{\max} - x_{\min}$; $RIC = Q_3 - Q_1$; posibles atípicos fuera de $[Q_1 - 1{,}5RIC, Q_3 + 1{,}5RIC]$. Bigotes en extremos observados no atípicos.

A.4 Dos variables: consulta de calculadora

Introducir parejas sin romper filas. Leer $\bar{x}, \bar{y}, \sigma_x, \sigma_y, r, a, b$ para modelo $\widehat{y} = a + bx$. Covarianza descriptiva $= r\sigma_x\sigma_y$. Centro $(\bar{x}, \bar{y})$. $-1 \leq r \leq 1$; signo: sentido, módulo: intensidad lineal orientativa. No definido si una variable es constante. Correlación no demuestra causalidad; un módulo elevado no garantiza predicciones.

Apéndice B

Práctica global

Resuelve con el método acordado, explica cada paso y comprueba tus resultados.

U09-G01

Con valores $0, 1, 2, 3$ y frecuencias $2, 4, 3, 1$, obtén N, media, mediana, moda y desviación descriptiva.
(Respuesta: 10; 1,3; 1; 1; 0,9.)

U09-G02

Con clases $[0, 10)$, $[10, 20)$, $[20, 30)$ y frecuencias $3, 5, 2$, estima media, moda, mediana, P_{30} y P_{90}.
(Respuesta: 14, 15, 15, 5, 25.)

U09-G03

Representa una caja modificada de $1, 2, 3, 4, 5, 6, 7, 30$.
(Respuesta: Caja [2,5, 6,5], mediana 4,5, bigotes $1, 7$ y punto atípico 30.)

U09-G04

Un pictograma tiene una figura por cuatro participantes y muestra tres figuras y media. ¿Cuántos representa? Explica cómo representar diez.
(Respuesta: 14; dos figuras y media para diez.)

U09-G05

En el mapa didáctico, ¿por qué no se suman los porcentajes autonómicos para obtener el porcentaje de España?
(Respuesta: Tienen denominadores territoriales distintos.)

U09-G06

Para $(1, 2), (2, 3), (3, 5), (4, 4)$ obtén con calculadora centro, covarianza, Pearson y regresión.
(Respuesta: (2,5, 3,5); 1; 0,8; $\widehat{y} = 1{,}5 + 0{,}8x$.)

U09-G07

Con $(1, 7), (2, 5), (3, 3)$, obtén regresión y estima en $x = 2{,}5$.
(Respuesta: $\widehat{y} = 9 - 2x$, $r = -1$; estimación 4.)

U09-G08

Un grupo tiene media 40 y desviación 6; otro media 60 y desviación 6. Compara dispersión relativa cuando el CV sea adecuado.
(Respuesta: 15% y 10%; mayor en el primero.)

Apéndice C

Autoevaluación de 20 preguntas

Una sola respuesta correcta. Anota tus elecciones y justifica las que te hayan generado dudas. La clave está exclusivamente en el documento docente.

Pregunta 1

El código postal es una variable:

A) Nominal

B) Continua

C) Discreta de recuento

D) Ordinal necesariamente

Pregunta 2

En $[10, 20)$ se incluye:

A) 20

B) 10

C) Ambos extremos

D) Ningún extremo

Pregunta 3

La marca de $[10, 20)$ es:

A) 10

B) 20

C) 15

D) 30

Pregunta 4

Con frecuencias $3, 5, 2$, el total es:

A) 3

B) 5

C) 8

D) 10

Pregunta 5

Una figura igual a dos personas: siete personas se representan con:

A) 3,5 figuras

B) 7 figuras

C) 14 figuras

D) 2 figuras

Pregunta 6

En histogramas con amplitudes distintas, la altura debe ser:

A) La marca

B) La densidad de frecuencia

C) El total N

D) El porcentaje siempre sin ajustar

Pregunta 7

La media de $2, 2, 4, 8$ es:

A) 2

B) 3

C) 4

D) 8

Pregunta 8

La mediana de $2, 2, 4, 8$ es:

A) 2

B) 4

C) 8

D) 3

Pregunta 9

En clases de marcas $5, 15, 25$ y frecuencias $3, 5, 2$, la moda aproximada es:

A) 15

B) 14

C) 5

D) 25

Pregunta 10

En esas mismas clases, P_{30} según nuestra convención es:

A) 15

B) 5

C) 25

D) 10

Pregunta 11

RIC significa:

A) Media menos moda

B) Rango total

C) $Q_3 - Q_1$

D) Desviación al cuadrado

Pregunta 12

La desviación típica tiene:

A) Unidades al cuadrado

B) Siempre porcentaje

C) Ninguna unidad nunca

D) Las mismas unidades que los datos

Pregunta 13

Para describir los datos usamos divisor:

A) N

B) N-1 siempre

C) N+1

D) 2N

Pregunta 14

El centro de una nube es:

A) (r, r)

B) $(\bar{x}, \bar{y})$

C) (a, b) siempre

D) El origen siempre

Pregunta 15

Pearson $r = -0{,}9$ indica:

A) Ausencia de relación

B) Relación positiva fuerte

C) Relación lineal descendente fuerte

D) Causalidad negativa demostrada

Pregunta 16

Un r próximo a cero:

A) Prueba independencia siempre

B) Prohíbe relación no lineal

C) Prueba relación perfecta

D) Indica débil asociación lineal

Pregunta 17

Si una columna es constante, Pearson:

A) No está definido

B) Vale siempre cero

C) Vale siempre uno

D) Vale siempre menos uno

Pregunta 18

En $y = a + bx$, el término independiente es:

A) b

B) a

C) r

D) N

Pregunta 19

Para $\widehat{y} = 1{,}5 + 0{,}8x$, estimar en 3 da:

A) 5

B) 2.4

C) 3.9

D) 4.5

Pregunta 20

Un mapa autonómico de porcentajes:

A) Describe a cada habitante

B) Siempre se suma para hallar el total nacional

C) No necesita leyenda

D) Resume agregados territoriales

Revisión personal

Selecciona dos respuestas y explica el procedimiento. Anota un error detectado y una acción concreta para mejorar. No cambies la respuesta sin explicar qué has revisado.