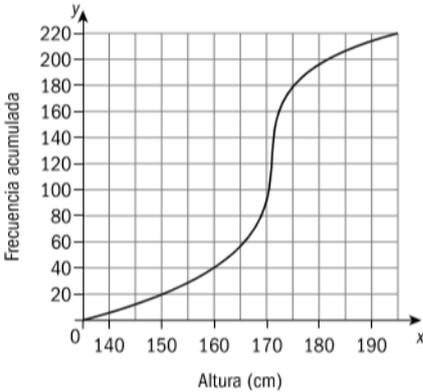


SERIE DE ESTADÍSTICA Y PROBABILIDAD (I)

| | |
|---|----------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------|
| 1 | <p>La suma de las edades, en el momento actual, de un padre y sus dos hijos, es 73 años. Dentro de 10 años la edad del padre será el doble de la edad del hijo menor y hace doce años la edad del hijo mayor era doble de la edad de su hermano. Halla la edad de cada uno.</p> <p>Plantea un sistema de ecuaciones y resuélvelo.</p> |
| 2 | <p>Representa la región factible o conjunto de soluciones del siguiente sistema de inecuaciones:</p> $\begin{cases} x - y \geq 0 \\ y - 2 \leq 0 \\ 2x + y \leq 10 \\ y \geq 0 \end{cases}$ |
| 3 | <p>PREGUNTA TIPO EXAMEN</p> <p>El siguiente diagrama de frecuencia acumulada muestra las alturas de 220 girasoles.</p> <ol style="list-style-type: none"> Halle la mediana de las alturas de los girasoles. El 25% de menor altura se envían a tiendas de florería. ¿Cuántos van a esas tiendas? ¿Entre qué alturas están? El 10% de mayor altura se destinan a decoración de hoteles. ¿Cuántos van a los hoteles? ¿Cuál es el girasol más bajo que se enviará a los hoteles para decoración? La mitad del medio de los girasoles se venden inmediatamente. ¿Cuántos son? La altura del girasol más alto es 195 cm y la altura del más bajo es 136 cm. Dibuje un diagrama de caja y bigotes para representar las alturas de los girasoles.  |
| 4 | <p>PREGUNTA TIPO EXAMEN</p> <p>La edad media de un grupo de amigos al terminar la escuela es de 17,5 años y la desviación típica es de 0,4 años. Todos se reencuentran en una reunión escolar después de 10 años.</p> <p>¿Cuál es ahora la media y la desviación típica de sus edades?</p> |

5

PREGUNTAS TIPO EXAMEN

Unos investigadores estudiaron el comportamiento de un grupo de 10 adolescentes. Evaluaron una variable de la personalidad llamada “agradabilidad”, que es una medida de cuán agradable resulta una persona para los demás. Se preguntó cuán alegre, terca, amable, mandona y cooperativa era la persona. La tabla registra las medias de las puntuaciones obtenidas por cada adolescente en estas características.

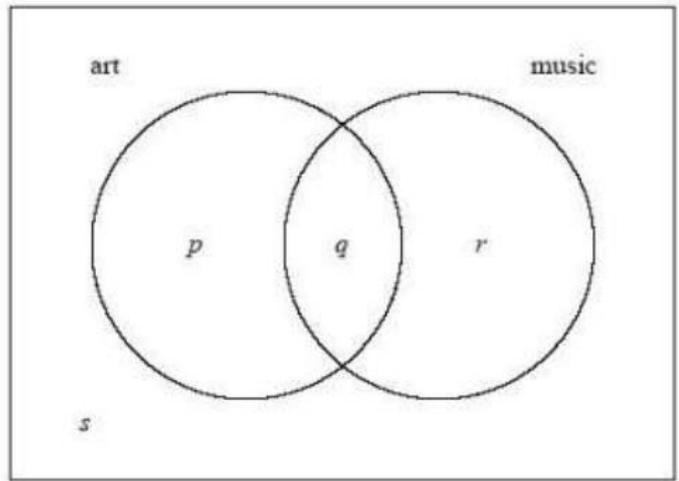
Los investigadores también crearon una forma de medir los problemas de conducta. Los jóvenes reportaron varios problemas de conducta en los últimos seis meses, tales como el engaño, el lenguaje vulgar, el hurto y las peleas. La tabla registra la suma obtenida por cada adolescente en la medición de estos problemas.

| Participante | Factor de agradabilidad | Problemas de conducta |
|--------------|-------------------------|-----------------------|
| Jorge | 4,3 | 5 |
| Guillermo | 3,0 | 22 |
| Oscar | 3,4 | 10 |
| Juan | 3,3 | 12 |
| Gerardo | 2,9 | 23 |
| Laura | 4,0 | 21 |
| Pilar | 4,7 | 2 |
| Nancy | 2,4 | 35 |
| Nora | 2,9 | 12 |
| Elizabeth | 4,7 | 4 |

- a** Elabore un diagrama de dispersión y muestre la recta de regresión.
- b** ¿Qué ocurre a medida que aumenta el factor de agradabilidad?
- c** Halle el coeficiente de correlación.
- d** Describa la correlación.
- e** Copie y complete la oración “Los adolescentes más agradables tendieron a tener _____ problemas de conducta”.
- f** Escriba la ecuación de la recta de regresión.
- g** Michelle estuvo ausente para las preguntas referidas a los problemas de conducta pero tuvo una puntuación de 4,5 en agradabilidad. Estime su puntuación para los problemas de conducta.

6

In a group of 16 students, 12 take art and 8 take music. One student takes neither art nor music. The Venn diagram below shows the events art and music. The values p , q , r and s represent numbers of students.



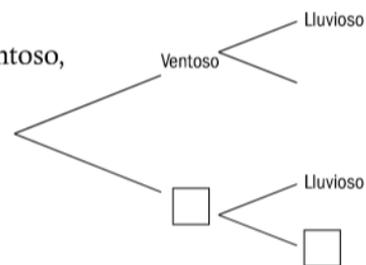
- (a)
 - (i) Write down the value of s .
 - (ii) Find the value of q .
 - (iii) Write down the value of p and of r . (5)
- (b)
 - (i) A student is selected at random. Given that the student takes music, write down the probability the student takes art.
 - (ii) **Hence**, show that taking music and taking art are **not** independent events. (4)
- (c) Two students are selected at random, one after the other. Find the probability that the first student takes **only** music and the second student takes **only** art. (4)

7

PREGUNTA TIPO EXAMEN

La probabilidad de que un día sea ventoso es 0,6. Si está ventoso, la probabilidad de que llueva es 0,4. Si no está ventoso, la probabilidad de que llueva es 0,2.

- a Copie y complete el diagrama de árbol.
- b ¿Cuál es la probabilidad de que un día dado llueva?
- c ¿Cuál es la probabilidad de que **no** llueva dos días consecutivos?



8

PREGUNTAS TIPO EXAMEN

Para los sucesos C y D se sabe que:
 $P(C) = 0,7$ $P(C' \cap D') = 0,25$ $P(D) = 0,2$.

- a Halle $P(C \cap D')$.
- b Explique por qué C y D no son sucesos independientes.

Proyecto de investigación

El proyecto de investigación, consiste en hacer un **estudio estadístico UNIDIMENSIONAL y otro BIDIMENSIONAL**. Al final tenéis un ejemplo con los dos estudios que pido.

Los datos deben de ser reales, y debéis especificar de dónde los habéis sacado, es decir debe estar la dirección de la página web de donde los sacasteis en el proyecto, para que yo pueda comprobarlos. Por ejemplo, yo los saqué de la página del instituto nacional de estadística.:

<http://www.ine.es/jaxiBD/tabla.do?per=01&type=db&divi=EPOB&idtab=3>

y copiáis el enlace para que yo lo vea.

Los dos estudios deben contener:

- Variables objeto del estudio y dirección de la página web de los datos
- Tabla con los datos
- Medidas de centralización
- Etc:
- Exactamente los mismos apartados que los dos ejemplos que encontrareis más adelante.

La página con los datos por supuesto, no debe ser la misma bajo ningún concepto para dos o más alumnos, de tal manera que antes de empezar el estudio, me diréis de donde sacaréis los datos, para que nadie más pueda sacar los mismos. Si hay alguien que tiene los mismos datos, consideraré que han copiado uno del otro y los dos tendrán 0 en el proyecto y serie

Los datos pueden provenir del periódico o de otra fuente, no es necesario que sean de una página web.

Mejor si los datos provienen de algún tema en que tengas interés: deporte, cine, música, etc.

Cualquier pregunta se me puede hacer en clase y los 5-10 primeros minutos de la clase los dedicaremos a los proyectos si es necesario.

Las gráficas no es necesario hacerlas al ordenador, se pueden hacer a mano.

Ejemplos de los proyectos: ➔

EJEMPLO DE ESTUDIO ESTADÍSTICO DE UNA VARIABLE UNIDIMENSIONAL

Variable objeto de estudio

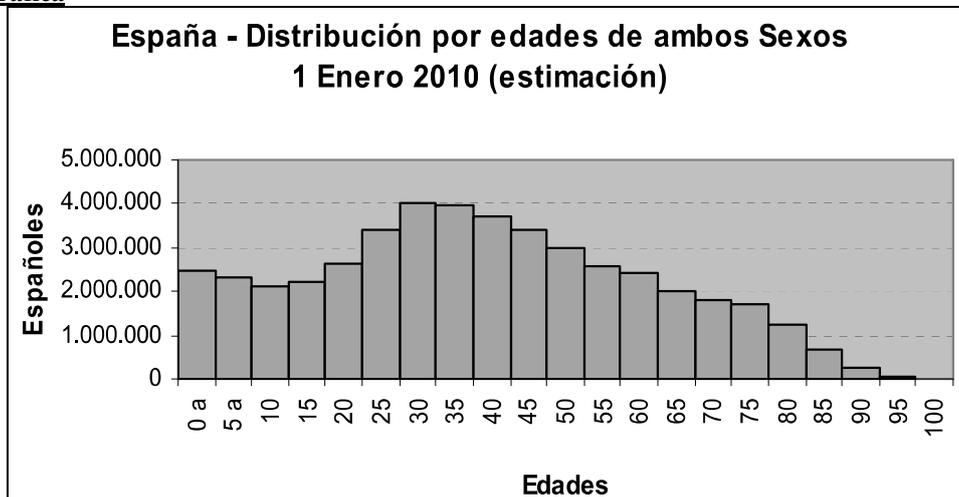
Vamos a estudiar la distribución de población de ambos sexos por segmentos de edad, de cuatro en cuatro años estimada para el 1 de enero de 2010. Están sacados de la web del Instituto Nacional de Estadística español.

(<http://www.ine.es/jaxiBD/tabla.do?per=01&type=db&divi=EPOB&idtab=3>)

Tabla con los datos

| Total Nacional | x | f | F | f·x | f·x ² |
|----------------|-----|-------------------|------------|----------------------|-----------------------|
| 0 a 4 años | 2 | 2.461.227 | 2.461.227 | 4.922.454 | 9.844.908 |
| 5 a 9 años | 7 | 2.297.895 | 4.759.122 | 16.085.265 | 112.596.855 |
| 10 a 14 años | 12 | 2.113.106 | 6.872.228 | 25.357.272 | 304.287.264 |
| 15 a 19 años | 17 | 2.240.229 | 9.112.457 | 38.083.893 | 647.426.181 |
| 20 a 24 años | 22 | 2.619.905 | 11.732.362 | 57.637.910 | 1.268.034.020 |
| 25 a 29 años | 27 | 3.379.985 | 15.112.347 | 91.259.595 | 2.464.009.065 |
| 30 a 34 años | 32 | 4.030.195 | 19.142.542 | 128.966.240 | 4.126.919.680 |
| 35 a 39 años | 37 | 3.945.775 | 23.088.317 | 145.993.675 | 5.401.765.975 |
| 40 a 44 años | 42 | 3.711.892 | 26.800.209 | 155.899.464 | 6.547.777.488 |
| 45 a 49 años | 47 | 3.423.294 | 30.223.503 | 160.894.818 | 7.562.056.446 |
| 50 a 54 años | 52 | 3.013.424 | 33.236.927 | 156.698.048 | 8.148.298.496 |
| 55 a 59 años | 57 | 2.595.058 | 35.831.985 | 147.918.306 | 8.431.343.442 |
| 60 a 64 años | 62 | 2.411.466 | 38.243.451 | 149.510.892 | 9.269.675.304 |
| 65 a 69 años | 67 | 2.015.458 | 40.258.909 | 135.035.686 | 9.047.390.962 |
| 70 a 74 años | 72 | 1.786.020 | 42.044.929 | 128.593.440 | 9.258.727.680 |
| 75 a 79 años | 77 | 1.706.879 | 43.751.808 | 131.429.683 | 10.120.085.591 |
| 80 a 84 años | 82 | 1.226.794 | 44.978.602 | 100.597.108 | 8.248.962.856 |
| 85 a 89 años | 87 | 691.867 | 45.670.469 | 60.192.429 | 5.236.741.323 |
| 90 a 94 años | 92 | 248.325 | 45.918.794 | 22.845.900 | 2.101.822.800 |
| 95 a 99 años | 97 | 63.071 | 45.981.865 | 6.117.887 | 593.435.039 |
| 100 y más años | 102 | 7.152 | 45.989.017 | 729.504 | 74.409.408 |
| | | 45.989.017 | | 1.864.769.469 | 98.975.610.783 |

Representación gráfica



Medidas de centralización y dispersión

Aplicadas las fórmulas, hemos obtenido los siguientes resultados:

Vemos que la población está bastante envejecida, ya que el primer cuarto de la población ocupa aproximadamente 22 años, el segundo cuarto 15, el tercero 20 y el último 47 años. Se observa bien en el diagrama de cajas y bigotes.

Un porcentaje estimable de la población pertenece al intervalo: (18, 63)

Diagrama de Cajas y Bigotes



| | |
|---------------------------|----------|
| moda: | 32 |
| media: | 40,5481 |
| Varianza: | 508,0053 |
| desviacion típica: | 22,5390 |

| | |
|-------------------------|----|
| Primer Cuartil: | 22 |
| Segundo Cuartil: | 37 |
| Tercer Cuartil: | 57 |

EJEMPLO DE ESTUDIO ESTADÍSTICO DE UNA VARIABLE BIDIMENSIONAL

Variables objeto de estudio

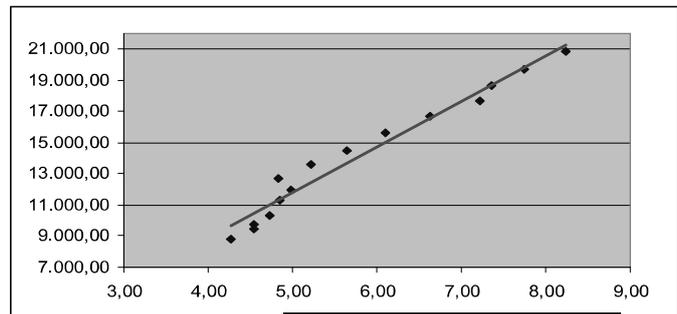
Vamos a estudiar la posible relación entre el crecimiento económico en España, concretamente el PIB medido en euros por habitante, y la natalidad medida con el número de nacimientos por 1.000 mujeres. Disponemos de datos de ambas variables desde el año 1991 hasta el 2005 y los hemos obtenido de la web del Instituto Nacional de Estadística español. (<http://www.ine.es/daco/daco42/sociales06/sociales.htm>)

Tabla con los datos

| | nº nacimientos/1000 mujeres | PIB (€/hab) | | | |
|------|-----------------------------|-------------------|----------------|--------------------------|----------------------|
| | x | y | x ² | y ² | x·y |
| 1991 | 4,28 | 8.798,10 | 18,308 | 77.406.511,778 | 37.644,880 |
| 1992 | 4,54 | 9.444,55 | 20,571 | 89.199.582,147 | 42.836,376 |
| 1993 | 4,54 | 9.740,83 | 20,642 | 94.883.792,185 | 44.256,399 |
| 1994 | 4,74 | 10.332,14 | 22,422 | 106.753.090,237 | 48.924,406 |
| 1995 | 4,85 | 11.353,85 | 23,570 | 128.909.811,061 | 55.121,414 |
| 1996 | 4,99 | 12.002,66 | 24,905 | 144.063.894,161 | 59.899,300 |
| 1997 | 4,83 | 12.730,62 | 23,356 | 162.068.697,274 | 61.524,286 |
| 1998 | 5,23 | 13.581,69 | 27,337 | 184.462.365,617 | 71.012,176 |
| 1999 | 5,64 | 14.524,98 | 31,831 | 210.974.949,376 | 81.948,044 |
| 2000 | 6,11 | 15.653,20 | 37,327 | 245.022.685,310 | 95.634,092 |
| 2001 | 6,62 | 16.715,47 | 43,858 | 279.406.834,138 | 110.699,296 |
| 2002 | 7,22 | 17.650,33 | 52,077 | 311.534.090,172 | 127.372,078 |
| 2003 | 7,36 | 18.629,66 | 54,156 | 347.064.275,451 | 137.097,637 |
| 2004 | 7,74 | 19.678,42 | 59,908 | 387.240.023,131 | 152.310,933 |
| 2005 | 8,24 | 20.863,89 | 67,898 | 435.301.845,973 | 171.918,442 |
| | 86,93 | 211.700,38 | 528,165 | 3.204.292.448,010 | 1.298.199,759 |

Representación gráfica

La nube de puntos y la recta de regresión *y sobre x* están representadas en el siguiente diagrama:



Medidas de centralización, dispersión y regresión

Aplicadas las fórmulas, hemos obtenido los resultados de la tabla adjunta

El coeficiente de correlación lineal tiene un valor muy cercano a 1, por lo que se puede calificar de **positiva y muy fuerte** la relación existente entre el PIB y la natalidad. Es decir, que al ir aumentando la riqueza en España también va aumentando la natalidad.

| | |
|--------------------|------------|
| media_x : | 5,795 |
| des_tip_x : | 1,275 |
| media_y : | 14.113,359 |
| des_tip_y : | 3.799,027 |
| covar : | 4.755,921 |
| coef_corr : | 0,982 |

Rectas de regresión

Si escribimos la ecuación de la recta de regresión *y sobre x* en la forma: $y = m \cdot x + n$ los coeficientes que hemos hallado son:

| |
|-------------------------------------|
| Recta de regresión y sobre x |
| m: 2925,1674 |
| n: 10383,4973 |

Si escribimos la ecuación de la recta de regresión *x sobre y* en la forma: $x = m \cdot y + n$ los coeficientes que hemos hallado son:

| |
|-------------------------------------|
| Recta de regresión x sobre y |
| m: 0,0003 |
| n: 1,1446 |

Estimación de resultados

Vamos a estimar el nº de nacidos por 1.000 mujeres para un PIB de 25.000€ por habitante.

Sustituimos $y = 25.000$ en la recta de regresión *x sobre y* obteniendo $x = 9,38$

Es decir, que para dicho PIB habría 9,38 nacidos por 1.000 mujeres en un año.