

ESTADÍSTICA. REGRESIÓN

ACTIVIDADE 1. TAXA BRUTA DE MORTALIDADE E ÍNDICE DE ENVELLECIMENTO

Obxectivos:

- ◆ Familiarizar ao estudante cos indicadores demográficos.
- ◆ Interpretación dun gráfico de dispersión.
- ◆ Analizar a dependencia entre dúas variables e obtención do coeficiente de correlación como medida da relación lineal.
- ◆ Cálculo da recta de regresión e uso dela para predicir.

Coñecementos previos necesarios:

- ◆ Para realizar esta actividade unicamente é necesario ter coñecementos básicos de estatística así como saber interpretar gráficos. A maior parte dos cálculos que se fan nesta actividade pódense realizar cunha calculadora.

Cuestións a resolver:

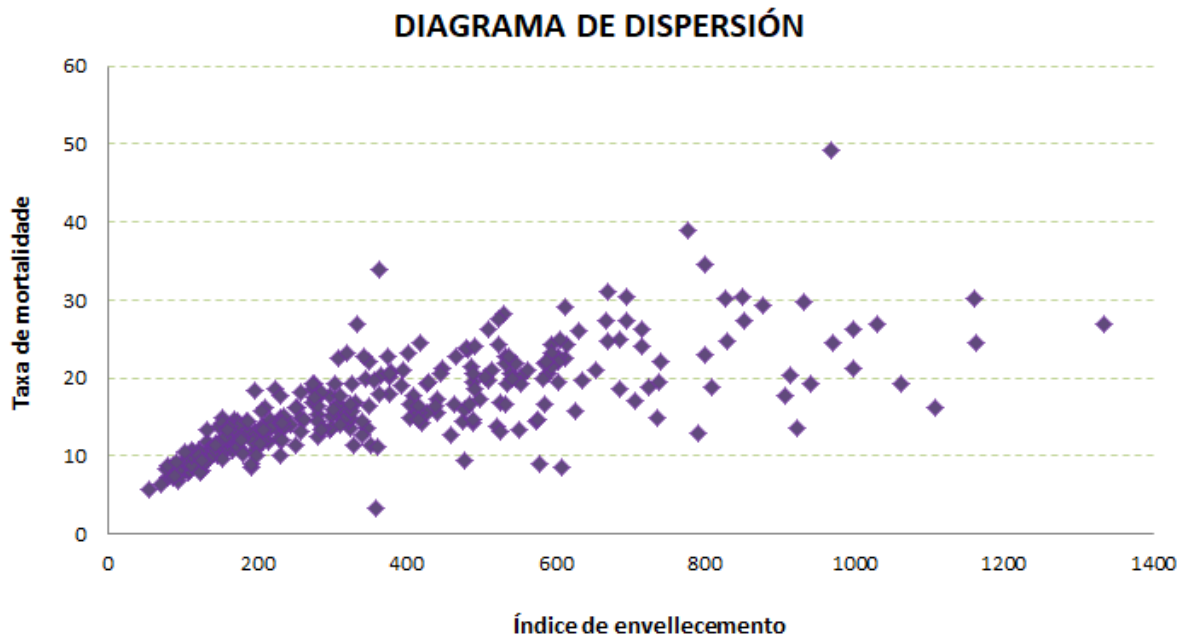
Nota para resolver as cuestións: non teñas en conta os concellos de Negueira de Muñiz, Triacastela, Beade, Pontedeira, Vilariño de Conso e Mondariz-Balneario. Polo tanto son un total de 307 concellos.

- 1) Completa a seguinte táboa cos datos da taxa bruta de mortalidade e o índice de envellecemento do ano 2019:

	Taxa bruta de mortalidade	Índice de envellecemento
Galicia
A Coruña
Lugo
Ourense
Pontevedra

- ¿Cal é a provincia que ten maior taxa de mortalidade?.....
- ¿Coincide coa provincia con maior índice de envellecemento?.....
- ¿Pensas que existe unha relación entre a taxa bruta de mortalidade e o índice de envellecemento? ¿Cal?.....

2) O seguinte gráfico de dispersión representa a taxa bruta de mortalidade e o índice de envellecemento de todos os concellos de Galicia no ano 2018:



- Neste gráfico ¿cal é a variable dependente?
- ¿E a variable independente?.....
- ¿Podes dicir se a taxa de mortalidade e o índice de envellecemento presentan algún tipo de relación?.....
- Se cres que existe relación, ¿é esta estatística ou funcional?
- ¿Podería dicirse que a relación é positiva?.....

- 3) Se calculamos a covarianza entre a taxa de mortalidade e o índice de envelhecimento para os concellos de Galicia no ano 2019, temos que :

$$\text{COV}(X,Y)= 1133,60$$

- ¿Axúdanos este dato a decidir se existe relación entre as variables?.....
- ¿Permite dicir se a relación é positiva ou negativa?.....
- ¿Poderíamos dicir que a relación entre as dúas variables é forte ou débil?.....
- ¿Que medida deberemos empregar para medir a relación entre as dúas variables?.....

- 4) O coeficiente de correlación lineal de Pearson entre a taxa de mortalidade e o índice de envelhecimento para os concellos de Galicia no ano 2019 toma o valor:

$$r = r(X,Y)= 0,74$$

- ¿Que tipo de relación existe entón entre as dúas variables?.....
- ¿É forte ou débil?.....
- Se tivéssemos $r(X,Y)= 0$, ¿poderíamos dicir que non existe relación entre as dúas variables?.....

- 5) Se calculamos os coeficientes da recta de regresión para os concellos de Galicia no ano 2019 temos que :

$$b = r \frac{S_y}{S_x} = 0,019 \quad e \quad a = \bar{y} - b\bar{x} = 9,26$$

e a recta de regresión ten a seguinte expresión:

$$y = 0,019 x + 9,26$$

A recta de regresión pode empregarse para predicir o valor da taxa bruta de mortalidade a partir do índice de envelhecimento.

Completa a seguinte táboa:

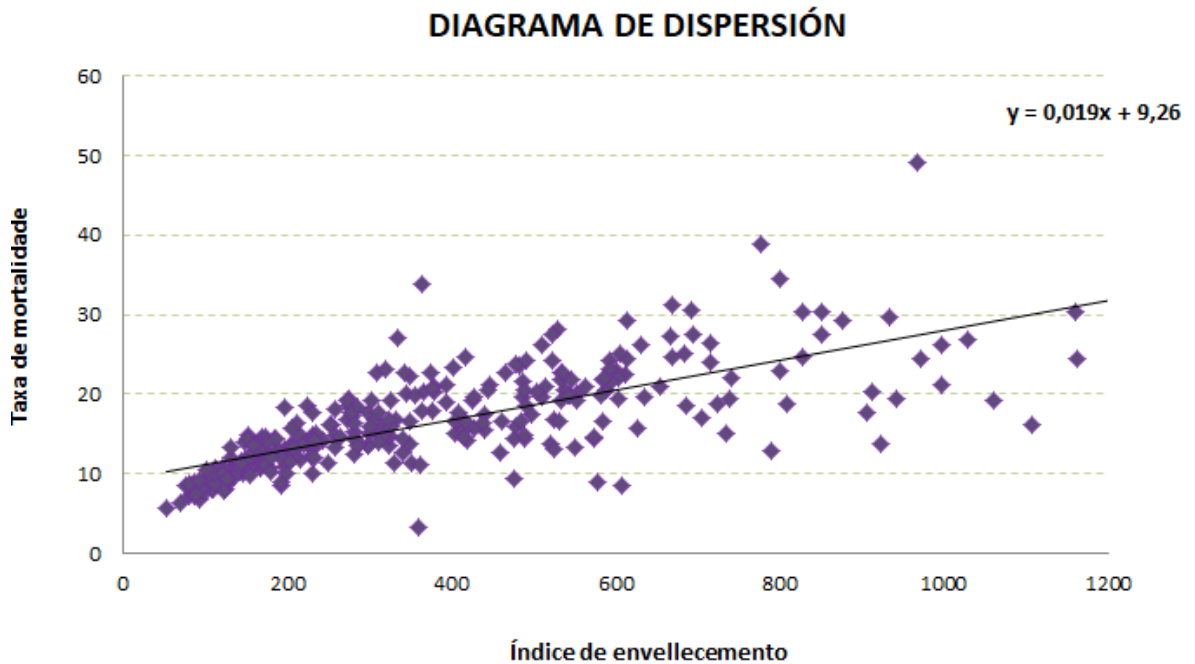
Concello	Índice de envellecemento	Taxa bruta de mortalidade	Taxa de mortalidade predita
Moeche	346,70	20,01
Guitiriz	259,26	13,22
Ourol	609,15	8,51
Lobeira	1.333,06	26,97
Catoira	154,51	11,89

¿Para cal de estes concellos é mellor a predición obtida a partir da recta de regresión?

.....

¿Cal ten maior erro de precisión ou residuo?.....

No seguinte gráfico amósase o diagrama de dispersión xunto coa recta de regresión que acabamos de calcular





Intenta identificar os concellos da táboa no gráfico,

¿Que representa o residuo?.....

¿Indica algo o signo do residuo?.....

¿Cando é maior o residuo (en valor absoluto)? ¿Canto máis cerca estea o punto da recta ou ao revés?.....