

Resumo tema 11

Variables estadísticas dobres

Ao estudo conxunto de dúas características ou variables estadísticas unidimensionais X e Y sobre unha mesma poboación acostúmase chamarlle **variable estatística bidimensional**

Diagrama de dispersión ou nube de puntos

No eixe de abscisas levamos os valores da variable X , que consideramos como variable independente, e sobre o eixe de ordenadas levamos os valores da variable Y , que consideramos como dependente. Debe quedar claro que as dúas variables non xogan o mesmo papel, a que denominamos independente é a que permite explicar o comportamento da outra, a denominada variable Y .

La representación gráfica dunha distribución bidimensional denomínase **diagrama de dispersión ou nube de puntos**.

Axuste da nube de puntos por unha recta. Recta de regresión.

Á recta que mellor se axusta á nube de puntos chámase **recta de regresión**.

$$y - \bar{y} = \frac{\frac{\sum x_i y_i}{n} - \bar{x}\bar{y}}{s_x^2} \cdot (x - \bar{x})$$

O punto (\bar{x}, \bar{y}) é chamado **centro de gravidade** da nube de puntos.

Concepto de correlación

O grao de dependencia lineal entre dúas variables mídese co **coeficiente de correlación** lineal, e cando a dependencia lineal é débil a recta de regresión carece de interese.

Covarianza

A **covarianza** é un parámetro que mide este tipo de relación e está definida como a media aritmética dos produtos da desviacións de cada un dos valores das variables respecto ás súas medias.

$$s_{xy} = \frac{\sum x_i y_i}{n} - \bar{x}\bar{y}$$

Coeficiente de correlación

A medida precisa da relación de dúas variables estadísticas proporciónaa o coeficiente de correlación lineal, representado pola letra r , e que está definido pola expresión seguinte:

$$r = \frac{s_{xy}}{s_x s_y}$$

Covarianza positiva indica correlación directa.

Covarianza negativa indica correlación inversa.

Covarianza nula indica que non hai correlación entre as variables