

ESTADISTICA

Practica 6

Problema de la dependencia entre variables medibles

Grado en ingeniería informática

Francisco Joaquín Murcia Gómez 48734281H

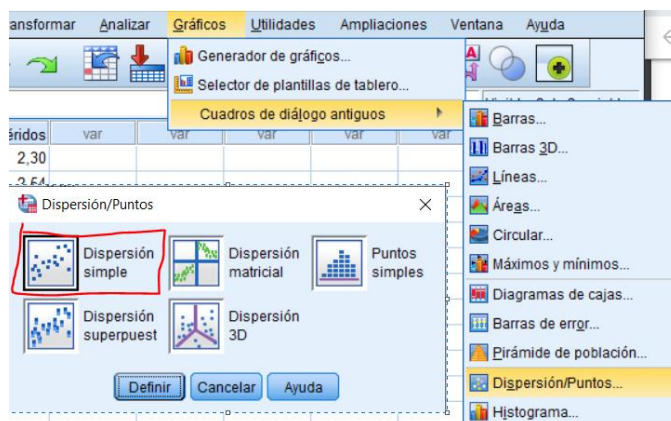
Grupo 1

- 1) En un estudio llevado a cabo en Italia, 10 pacientes con Hipertrigliceridemia se sometieron a una dieta baja en grasas y alta en carbohidratos para investigar si había alguna relación entre estas variables. La tabla muestra los valores antes de comenzar la dieta:

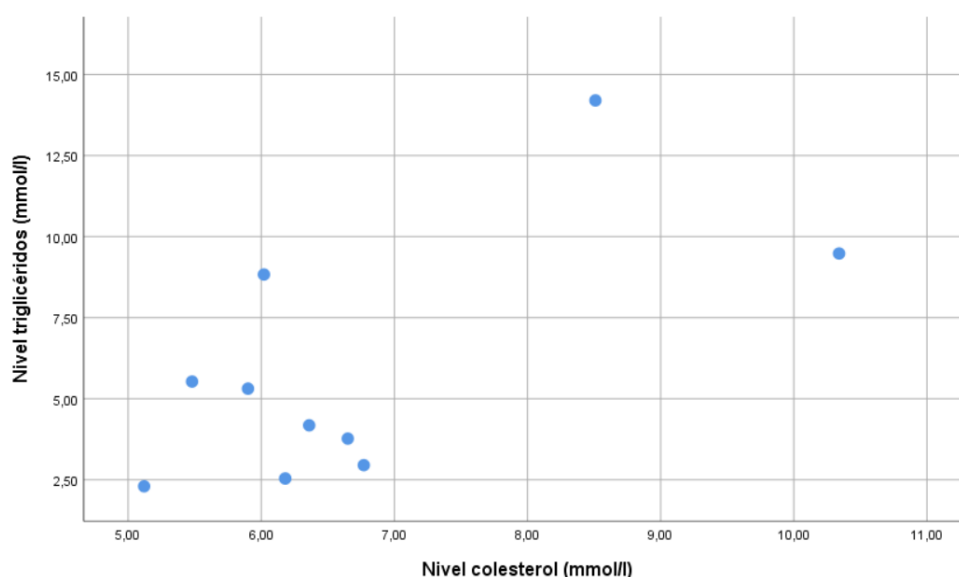
Nivel colesterol (mmol/l)	5,12	6,18	6,77	6,65	6,36	5,9	5,48	6,02	10,34	8,51
Nivel triglicéridos (mmol/l)	2,3	2,54	2,95	3,77	4,18	5,31	5,53	8,83	9,48	14,2

a) Construye un diagrama de dispersión para estos datos.

Introducimos los datos de la tabla a dos variables en el SPSS, una vez creadas las variables nos vamos a la ruta "gráficos>Cuadro de diálogos antiguos>Dispersión/Puntos" allí seleccionamos "simple" (como se ve en la imagen).

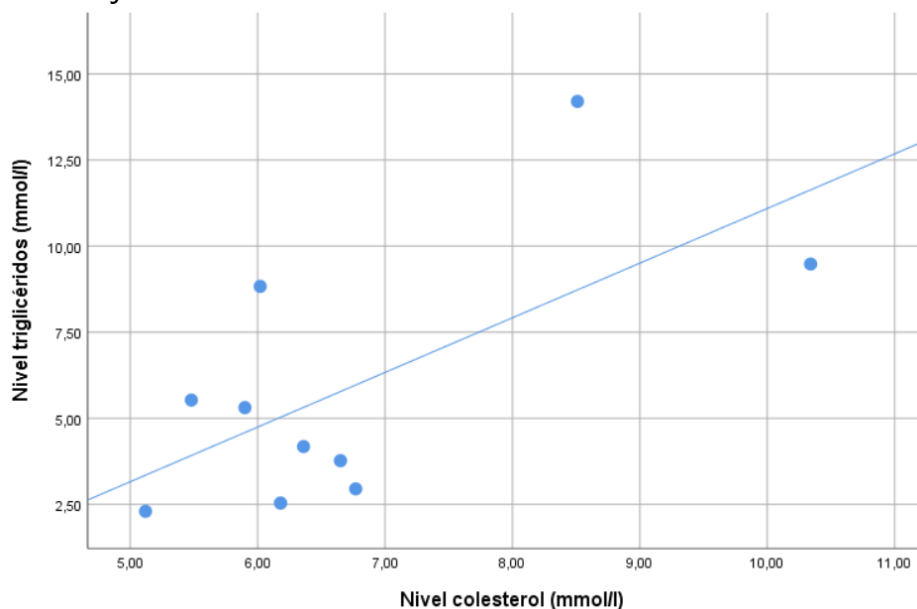


Una vez seleccionado nos aparecerá un menú, allí seleccionamos la variable del eje x y la del eje y, para este ejercicio es indiferente cual seleccionemos. El resultado es el siguiente.



b) ¿Existe evidencia de relación lineal entre los niveles de colesterol y triglicéridos antes de la dieta?

Para hacer la línea de ajuste damos doble clic al gráfico y clicamos añadir línea de ajuste total" en el menú de arriba



Al crecer los índices de colesterol cuando carecen los de triglicéridos podemos decir que es un ajuste lineal y, además, al haber mucha separación con la línea podemos decir que tienen mal ajuste.

c) Estimar el nivel de triglicéridos cuando el nivel de colesterol es de 6,10 mmol/l.

Para calcular la estimación vamos a la ruta "analizar>regresión>estimación curvilínea" ponemos dependiente colesterol que es el dato que nos dan, y la variable triglicéridos en la variable, nos da como resultado el grafico de arriba y la siguiente tabla:

Resumen de modelo y estimaciones de parámetro							
Variable dependiente: Nivel triglicéridos (mmol/l)							
Ecuación	R cuadrado	Resumen del modelo				Estimaciones de parámetro	
		F	gl1	gl2	Sig.	Constante	b1
Lineal	,422	5,842	1	8	,042	-4,775	1,587
La variable independiente es Nivel colesterol (mmol/l).							

Utilizando la ecuación de la pendiente de una recta " $y=m*x+n$ " podemos calcular el nivel de triglicéridos, ya que en le margen derecho del tabal tenemos la pendiente seria "b1" y el valor independiente seria "la constante "y "x" el nivel de colesterol. Por lo tanto, la operación sería la siguiente: nivel de triglicéridos= $1,587*6,10-4,775$.

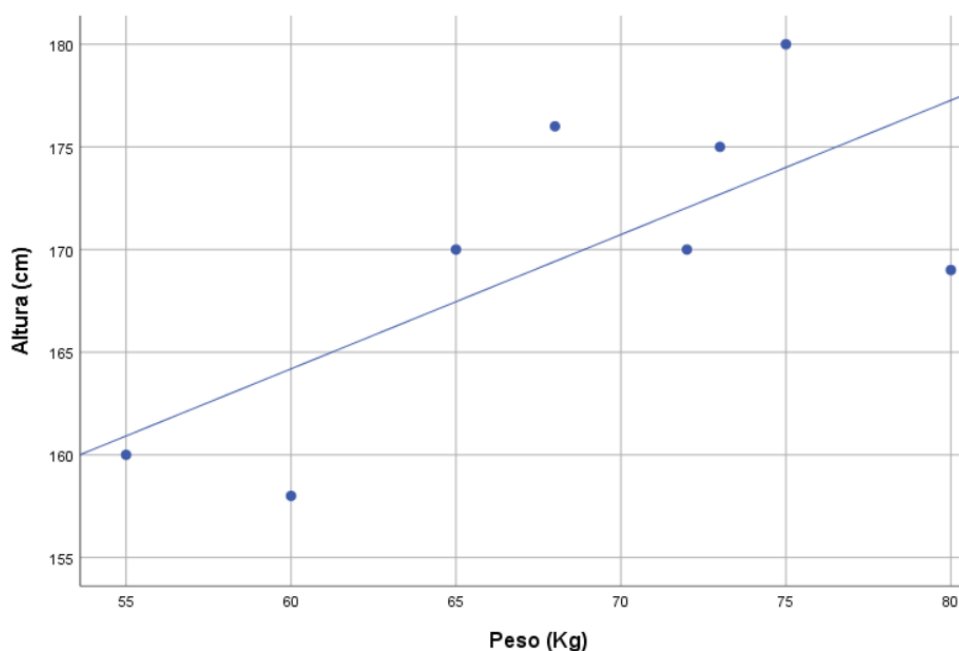
El resultado es 4,9057 mmol/l.

2) Las mediciones de peso y altura de una muestra de ocho estudiantes se en la siguiente tabla:

Peso (Kg)	75	60	72	80	65	68	73	55
Altura (cm)	180	158	170	169	170	176	175	160

- a) **Construye el gráfico de dispersión (nube de puntos) de los datos. ¿Parece plausible ajustar una recta de regresión? ¿Cómo debe salir el coeficiente de correlación? Razona la respuesta.**

Construimos el grafico con el procedimiento explicado en el ejercicio 1.a, esta vez he colocado el va variable peso en el eje "x" y la altura en el "y", aunque en este caso da lo mismo los ejes. El resultado es el siguiente:



Podemos decir, que si se puede representar con una recta de regresión debido a que los puntos siguen una tendencia ascendente similar a una recta ascendente, este grafico seguiría un modelo lineal con mal ajuste, por lo dicho el coeficiente de correlación tendería a 1

b) Calcular la covarianza existente entre ambas variables así como el coeficiente de correlación.

Para obtener el índice de correlación hacemos la ruta "analizar> correlación> bivariadas", añadimos las dos variables y seleccionamos "pearson", además, para sacar la covarianza le damos a opciones y seleccionamos "desviaciones de productos vectoriales y covarianzas".

El resultado es el siguiente:

Correlaciones			
		Peso (Kg)	Altura (cm)
Peso (Kg)	Correlación de Pearson	1	,707*
	Sig. (bilateral)		,050
	Suma de cuadrados y productos vectoriales	474,000	310,000
	Covarianza	67,714	44,286
	N	8	8
Altura (cm)	Correlación de Pearson	,707*	1
	Sig. (bilateral)	,050	
	Suma de cuadrados y productos vectoriales	310,000	405,500
	Covarianza	44,286	57,929
	N	8	8

*. La correlación es significativa en el nivel 0,05 (bilateral).

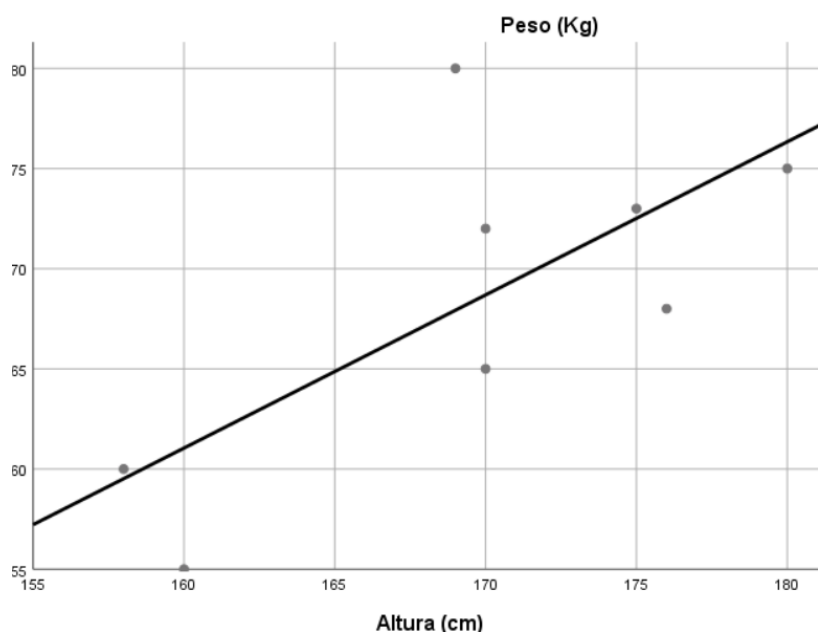
Covarianza=44.286

Coeficiente de correlación=0.707

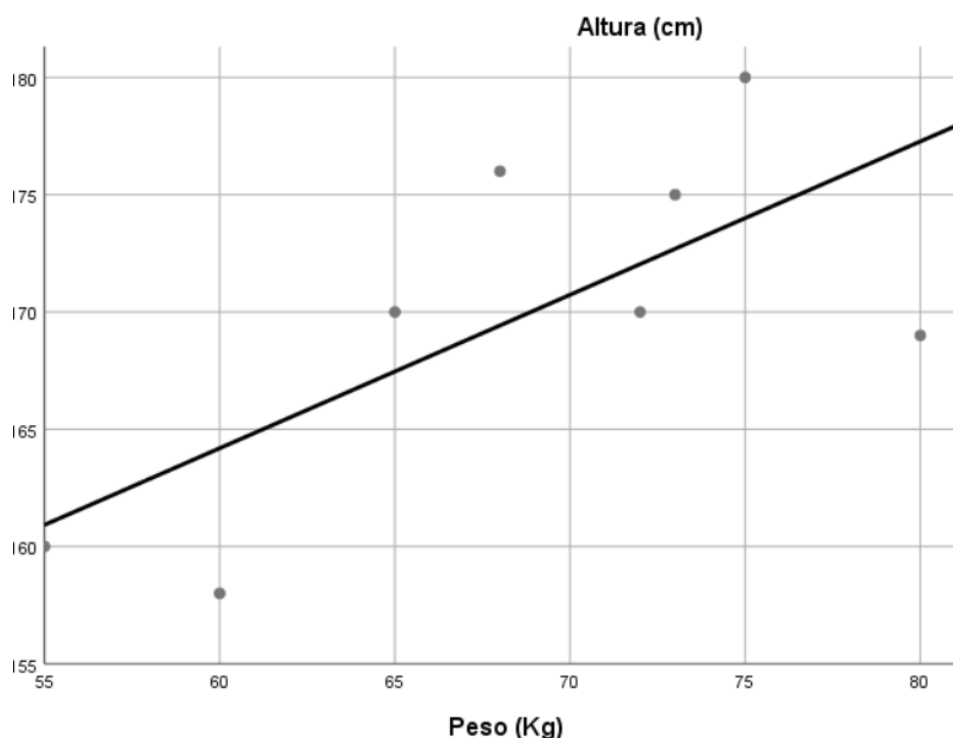
c) Calcular las dos rectas de regresión.

El procedimiento es como hicimos en el ejercicio 1.c a la hora de calcular la estimación. Esta vez haremos un gráfico de cada.

Primero teniendo como variable dependiente el peso:



El segundo con la variable dependiente la altura:



d) ¿Qué altura cabe esperar para un estudiante que pese 70 kg?

Para calcular la altura, utilizaremos la formula del ejercicio 1.c, para ello necesitaremos la tabla que nos genera SPSS ala hacer la recta de regresión del ejercicio 2.c con la altura siendo la variable dependiente, que es la siguiente:

Variable dependiente: Altura (cm)

Ecuación	R cuadrado	Resumen del modelo				Estimaciones de parámetro	
		F	gl1	gl2	Sig.	Constante	b1
Lineal	,500	6,000	1	6	,050	124,950	,654

La variable independiente es Peso (Kg).

En este caso m será 0,654 y n 124,950, dando la siguiente ecuación:

Altura=0,654*70+124,950, que da como resultado una altura estimada de 170,73 cm