



ANALIZANDO LA CORRELACIÓN

La correlación es muy utilizada como herramienta descriptiva como ya has visto en el bloque anterior.

Y se utiliza para encontrar las variables relacionadas entre sí para poder después entender mucho mejor qué variables contienen información similar.

Pero también se puede utilizar desde el punto de vista analítico y poder aportar conclusiones importantes en estudios científicos.

En esta hoja de trabajo practicarás las tres situaciones más comunes paso a paso para que entiendas cómo hacerlo en la práctica.

¡Let's go!

LA CORRELACIÓN DE DOS MEDIDAS

Con los datos Espalda como siempre vamos a tratar de ver si están relacionadas dos variables. El peso y la altura y el ODImes0.

Las variables a trabajar son:

- Peso
- Altura
- ODImes0

Los pasos serán los siguientes.

Paso 1 – Define la preguntas u objetivo

El objetivo lo acabamos de definir pero escríbelo aquí para seguir el orden de la hoja de trabajo de los test estadísticos.

Determinar si existe una correlación entre las variables peso, altura y ODImes0

Paso 2 – Describe la pregunta con estadística descriptiva

Antes de nada, es muy interesante visualizar la información que queremos analizar.

En el test de normalidad nos interesa dibujar:

- El matrixplot mixto

ANALIZANDO LA CORRELACIÓN

- El diagrama de correlación

```
> #COMPROBAR NORMALIDAD
>
> # para el peso
> dvectorPeso <- df[,1]
> shapiro.test(dvectorPeso)

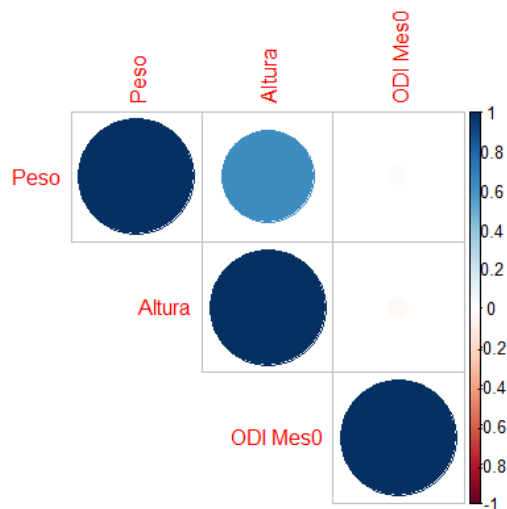
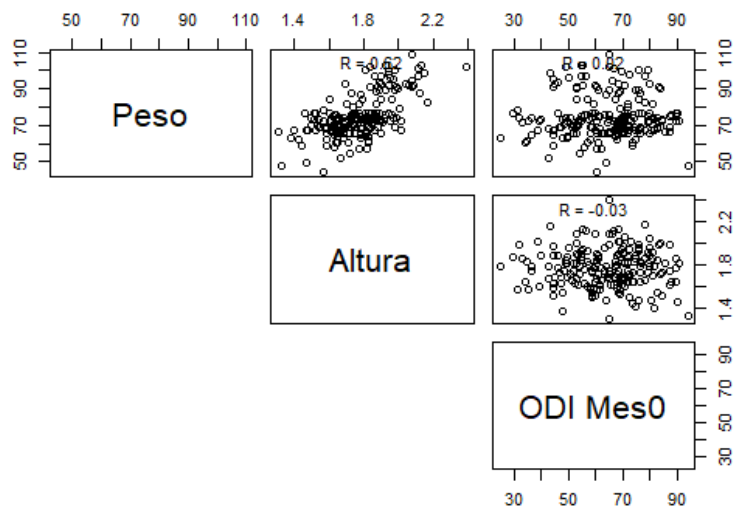
Shapiro-wilk normality test
data: dvectorPeso
W = 0.93508, p-value = 8.744e-08

> # para el altura
> dvectorAltura <- df[,2]
> shapiro.test(dvectorAltura)

Shapiro-wilk normality test
data: dvectorAltura
W = 0.99582, p-value = 0.8611

> # para el ODIMes0
> dvectorODIMes0 <- df[,3]
> shapiro.test(dvectorODIMes0)

Shapiro-wilk normality test
data: dvectorODIMes0
W = 0.98627, p-value = 0.04946
```



```

      Peso  Altura  ODI Mes0
Peso   1.00   0.62   0.02
Altura 0.62   1.00  -0.03
ODI Mes0 0.02 -0.03   1.00

n= 200
```

Paso 3 – Rellena la plantilla de contraste

Ahora calcula para el coeficiente de Pearson y Spearman:

- El Coeficiente de correlación
- La tabla de p-valores

Define la H1 o hipótesis de investigación:

Hay relación entre las medidas

Define la H0 o hipótesis nula:

No Existe una correlación entre las variables

Umbral de contraste (alpha):

5% = 0.05

Coeficiente utilizado: spearman y pearson

Spearman

Comprueba las restricciones:

Normalidad de las variables para Pearson

Altura no es normal

Peso si es normal

ODI Mes0 no es normal

Significación o p-valor:

Matriz de p-valores:

P	Peso	Altura	ODI Mes0
Peso		0.0000	0.4252
Altura	0.0000		0.9551
ODI Mes0	0.4252	0.9551	

Matriz de correlación:

	Peso	Altura	ODI Mes0
Peso	1.00	0.55	0.06
Altura	0.55	1.00	0.00
ODI Mes0	0.06	0.00	1.00
n= 200			



ANALIZANDO LA CORRELACIÓN

Respuesta:

Qué variables están relacionada y qué grado de asociación presentan

El Peso y la altura están correlacionados con un grado de asociación moderado de 0.55 directo normal.

PASO 4 - CONCLUSIÓN

Juntamos la información del test y el de la descripción de datos.

*Las variables que tienen correlación son la altura y el peso.
El P-valor del test indica que su relación es significativa.*

UNA VARIABLE ORDINAL y UNA MEDIDA

Comprueba la asociación de la variable $\text{diff_ODI} = \text{ODI}_{\text{mes0}} - \text{ODI}_{\text{mes1}}$ con el número de hernias discales (NHD)

Es decir,

- Medida = diff_ODI
- Factor = NHD

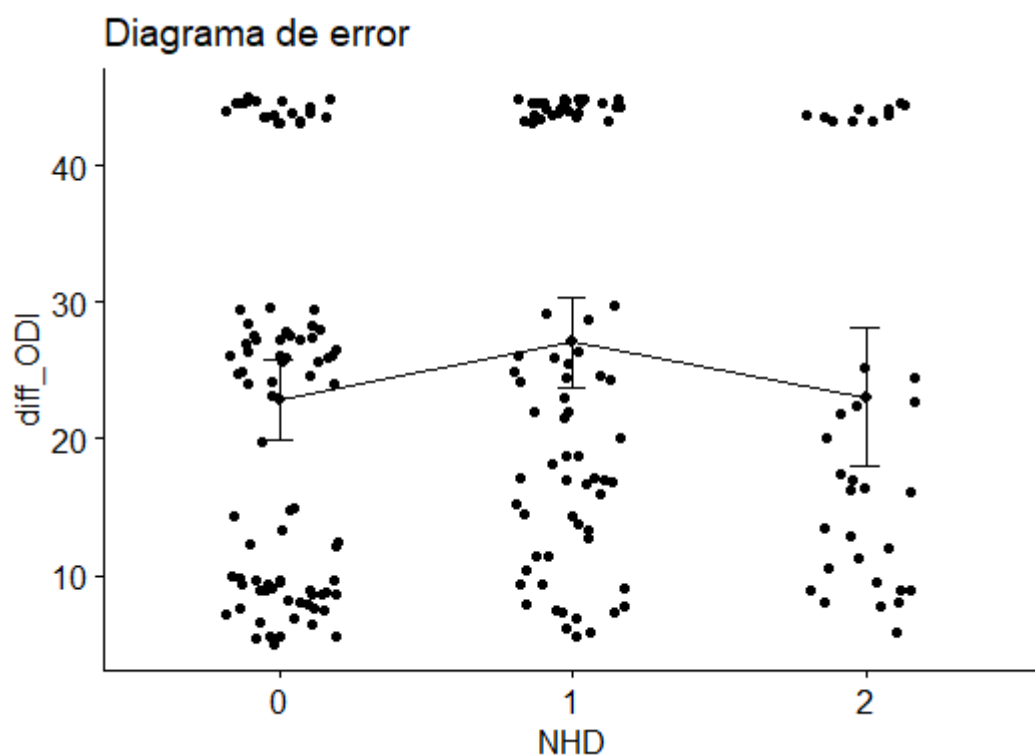
Utiliza Spearman como coeficiente de correlación

Paso 1 – Define la preguntas u objetivo

Comprobar si la mejora del paciente (diff_ODI) se asocia con el número de hernias discales (NHD)

Paso 2 – Describe la pregunta con estadística descriptiva

ANALIZANDO LA CORRELACIÓN



```
> # Coeficiente de Kendall
> m <- rcorr(as.matrix(df), type="spearman")
> m
```

	diff_ODI	NHD
diff_ODI	1.00	0.04
NHD	0.04	1.00

n= 200

P

	diff_ODI	NHD
diff_ODI		0.5287
NHD	0.5287	

Paso 3 – Rellena la plantilla de contraste

Define la H1 o hipótesis de investigación: (hipótesis de diferencias)

H1: Hay asociación entre las variables

Define la H0 o hipótesis nula: (la contraria a la H1)

H0: No existe una asociación entre las variables

ANALIZANDO LA CORRELACIÓN

Umbral de contraste (alpha):

5% = 0.05

Test Estadístico (selección del test): Spearman

Comprueba las restricciones: Una variable Ordinal + una medida

Significación o p-valor: (el resultado del test)

p-valor: 0.5287

Coeficiente de correlación: 0.04

Respuesta: (si el p-valor <5% ¿te quedas con la H1)

No hay evidencias de que exista una asociación entre las variables rechazamos la H1

Paso 4 – Conclusión

En la gráfica de barras no se observan diferencias en las medias asociadas al número de hernias.

El test de Spearman no indica que haya significancia en la asociación con un p-valor bastante alto de 0.5287 y una correlación casi cero.

Se concluye que la mejora del índice de discapacidad es indiferente al número de hernias discales que presenta el paciente.

Con esto creo que ya puedes con muchos de los problemas que te encontrarás en el sentido relacional.

¡Vamos a seguir avanzando!

¡A por ello Valiente!