

Otros coeficientes de correlación

El coeficiente de correlación de Pearson es adecuado para datos cuantitativos de intervalo o razón. En el caso de variables nominales, ordinales o dicotómicas se usan otros coeficientes, aunque algunos de ellos se reducen en sus cálculos al coeficiente de Pearson.

Coeficiente de Spearman o de rangos

Este coeficiente se popularizó en los textos de Psicología y de Ciencias de la educación, para variables de tipo ordinal (rangos). Tenía la ventaja de poseer un cálculo simple, que le daba ventaja en la época en la que aún no existían las calculadoras. Hoy en día, por su equivalencia con el de Pearson, se sustituye por este.

Su fórmula es

$$r_s = \frac{6 \sum d_i^2}{n(n^2 - 1)}$$

en la que ***d*** representa las diferencias entre los dos rangos correspondientes a un mismo dato, y ***n*** el número de esos datos.

Se usa sobre dos conjuntos que representen dos ordenaciones distintas de un mismo colectivo, en asignaturas, jueces, opiniones, etc. distintas.

Si existen empates, se asigna a cada individuo empatado el promedio de los órdenes que le hubieran correspondido.

Por ejemplo, en 1,2,3,3,3,4,5, asignaremos a los empatados el rango $(3+4+5)/3=4$

Abre el modelo ***otroscoef.ods*** y rellena la zona de datos con estas ordenaciones de 8 cantantes obtenidas en dos

curso de E.S.O. ¿Existe paralelismo entre las ordenaciones de los dos colectivos?

Curso A	1	2	3	4	5	6	7	8
Curso B	4	6	5	3	8	7	2	1

Si consultas el resultado, verás que es un coeficiente negativo y más cercano a cero que a 1. Por tanto, apenas existe paralelismo, y el que hay es negativo, porque las ordenaciones son casi inversas (con el alumnado actual esto sería casi inaudito, pues uno de los dos grupos no habría seguido las modas)

Coeficiente biserial puntual

Este coeficiente es usado cuando una de las variables es cuantitativa y la otra dicotómica. Por ejemplo, la X puede ser el número de horas de estudio (cuantitativa) y la Y el hecho de aprobar o no (dicotómica) Coincide también con el coeficiente de Pearson

Su fórmula es

$$r_{bp} = \frac{\bar{X}_1 - \bar{X}_0}{S_x} \sqrt{pq}$$

En la que las medias del numerador corresponden a los dos grupos en los que se separan los datos según la variable dicotómica, el denominador es la desviación típica de la X, y p q son las proporciones de los dos grupos respecto al total.

El siguiente ejemplo corresponde a los datos de 10 clientes de una clínica de adelgazamiento, en el que X es su índice

corporal e Y el hecho de tener o no un grado de glucemia peligroso.

X	29	28	34	29	38	40	32	41	29	27
Y	SI	NOSI	NOSI	SI	NOSI	NONO				

Busca la hoja del coeficiente biserial en el modelo **otroscoef.ods**, rellena estos datos sustituyendo el SI por 1 y el NO por 0. Consulta el resultado, que debe darte 0,74, es decir, una correlación importante entre el índice de peso corporal y el peligro de diabetes.

Coeficiente de contingencia

Este coeficiente, no muy perfecto en su diseño, es un complemento a las pruebas de bondad de ajuste y las pruebas de independencia.

Su fórmula es

$$C = \sqrt{\frac{\chi^2}{\chi^2 + N}}$$

Como ejemplo, calcula el coeficiente C de la tabla de faltas graves que hemos usado en esta sesión y te dará un valor de 0,16, de acuerdo con el resultado de la prueba de independencia, que nos indicaba que las muestras eran homogéneas.