



Análisis de Correlación Lineal

Erika Kania Kuhl
Ing. Agrónomo Dr.



Análisis de Correlación Lineal

Objetivo

Medir y evaluar el grado de asociación lineal entre dos variables CUANTITATIVAS.

Acá no interesa si una variable es la causa de otra variable, a diferencia del análisis de regresión (No existe relación Causa – Efecto)

No hay variables dependiente ni variable independiente, a diferencia del análisis de regresión



Frecuentemente estamos interesados en estudiar la manera de cómo están dos variables correlacionadas y cuantificar ese grado de asociación lineal. Por ejemplo:

- ¿Será que plantas con la parte aérea más desarrollada tienden a tener el sistema radicular más desarrollado?
- ¿Será que la materia seca de la parte aérea de las plantas está asociada con la materia seca de las raíces? ¿Crecen las dos variables en el mismo sentido?
- ¿Será que el contenido de azúcar en plantas de caña está asociado con el contenido de humedad del suelo?



Correlación: Examinar la dirección y la magnitud de la asociación entre dos variables.

- Dirección:** ¿Si al aumentar el valor de una variable, aumenta o disminuye el valor de otra variable? (Correlación positiva o negativa)
- Magnitud:** “intensidad de la asociación” entre ellas
- Significancia** de la correlación



Coefficiente de correlación lineal de Pearson

Este coeficiente varía entre -1 y 1 :

$$-1 \leq r \leq 1$$

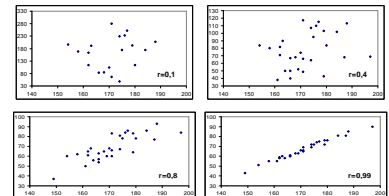
Una correlación cercana a cero indica que las dos variables no están correlacionadas

$$r = \frac{\sum xy - \sum x \sum y / n}{\text{Raíz } ((\sum x^2 - (\sum x)^2/n)(\sum y^2 - (\sum y)^2/n))}$$



Correlación lineales positivas

Una **correlación lineal cercana a 1** indica que las variables están **positivamente correlacionadas** (x1 crece, x2 crece).



Correlación lineales negativas

Una **correlación lineal cercana a -1** indica que las variables están **negativamente correlacionadas** (X1 crece, x2 disminuye).

Magnitud de la asociación

Un criterio usual (empírico) es considerar que

- $r \leq |0,20|$ es una correlación lineal baja
- $|0,20| \leq r \leq |0,60|$ es una correlación lineal media
- $|0,60| \leq r \leq |1,0|$ es una correlación lineal alta

Pruebas de hipótesis

(ρ es el coeficiente de correlación poblacional)

$H_0: \rho=0$
 $H_1: \rho \neq 0$

Archivo Proteínas

Objetivo: estudiar los alimentos que se utilizan como fuentes proteicas en las dietas de los países europeos.

Los datos corresponden a la composición proteica de dietas de habitantes de países europeos según los alimentos consumidos (Carne Vacuna, Carne de Cerdo, Huevos, Leche, Pescado, Cereal, Embutidos, Frutos Secos y Frutas y Vegetales)

El archivo contiene 9 variables correspondientes a los alimentos consumidos y un criterio de clasificación con el nombre de los países

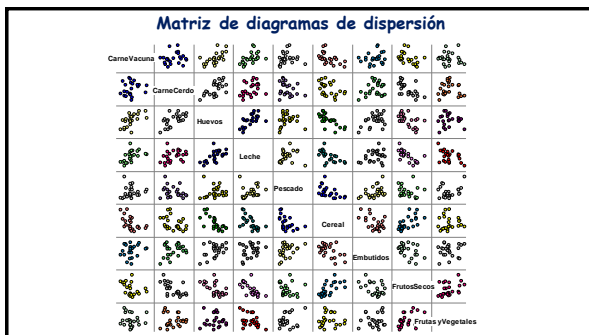


Tabla de Resultados
(Presentación de resultados en forma matricial)

Correlación de Pearson: Coeficientes y probabilidades

	CarneVacuna	CarneCerdo	Huevos	Leche	Pescado	Cereal	Embutidos	FrutosSecos	Frutas y Vegetales
CarneVacuna	1	0,4403	0,0037	0,0135	0,7541	0,0091	0,3322	0,0646	0,7095
CarneCerdo	0,1761	1	0,0015	0,1145	0,2301	0,0552	0,1995	0,0013	0,8181
Huevos	0,6068	0,6102	1	0,0013	0,793	0,0001	0,0341	0,0057	0,8632
Leche	0,497	0,3307	0,6165	1	0,4763	0,0009	0,1715	0,0002	0,0375
Pescado	0,0675	-0,2545	0,0565	0,1527	1	0,009	0,0516	0,5202	0,1993
Cereal	-0,5203	-0,3963	-0,7064	-0,6349	-0,521	1	0,0094	0,0007	0,8599
Embutidos	0,1682	0,2714	0,4339	0,2886	0,4019	-0,5187	1	0,0329	0,6203
FrutosSecos	-0,3832	-0,6161	-0,547	-0,6937	-0,1351	0,6415	-0,4367	1	0,0738
Frutas y Vegetales	-0,0802	-0,0496	-0,0372	-0,4268	0,2715	0,038	0,1065	0,3716	1

Tabla de Resultados (Presentación de resultados en forma matricial)

FACULTAD DE CIENCIAS AGRONÓMICAS UNIVERSIDAD DE CHILE

	CarneVacuna	CarneCerdo	Huevos	Leche	Pescado	Cereal	Embutidos	FrutosSecos	Frutas yVegetales
CarneVacuna	1	0,4103	0,0017	0,0135	0,7541	0,0091	0,4322	0,0646	0,7095
CarneCerdo	0,1761	1	0,0015	0,1145	0,2301	0,0552	0,1995	0,0013	0,8181
Huevos	0,6068	0,6102	1	0,0013	0,793	0,0001	0,0341	0,0057	0,8632
Leche	0,497	0,3307	0,6165	1	0,4763	0,0009	0,1715	0,0002	0,0375
Pescado	0,0675	-0,2545	0,0565	0,1527	1	0,009	0,0516	0,5292	0,1993
Cereal	-0,5203	-0,3963	-0,7064	-0,6349	-0,521	1	0,0094	0,0007	0,8599
Embutidos	0,1682	0,2714	0,4339	0,2886	0,4019	-0,5187	1	0,0329	0,6203
FrutosSecos	-0,3832	-0,6161	-0,547	-0,6937	-0,1351	0,6415	-0,4367	1	0,0738
Frutas yVegetales	-0,0802	-0,0496	-0,0372	-0,4268	0,2715	0,038	0,1065	0,3716	1

Tabla de Resultados (Presentación de resultados en forma matricial)

FACULTAD DE CIENCIAS AGRONÓMICAS UNIVERSIDAD DE CHILE

	CarneVacuna	CarneCerdo	Huevos	Leche	Pescado	Cereal	Embutidos	FrutosSecos	Frutas yVegetales
CarneVacuna	1	0,4103	0,0017	0,0135	0,7541	0,0091	0,4322	0,0646	0,7095
CarneCerdo	0,1761	1	0,0015	0,1145	0,2301	0,0552	0,1995	0,0013	0,8181
Huevos	0,6068	0,6102	1	0,0013	0,793	0,0001	0,0341	0,0057	0,8632
Leche	0,497	0,3307	0,6165	1	0,4763	0,0009	0,1715	0,0002	0,0375
Pescado	0,0675	-0,2545	0,0565	0,1527	1	0,009	0,0516	0,5292	0,1993
Cereal	-0,5203	-0,3963	-0,7064	-0,6349	-0,521	1	0,0094	0,0007	0,8599
Embutidos	0,1682	0,2714	0,4339	0,2886	0,4019	-0,5187	1	0,0329	0,6203
FrutosSecos	-0,3832	-0,6161	-0,547	-0,6937	-0,1351	0,6415	-0,4367	1	0,0738
Frutas yVegetales	-0,0802	-0,0496	-0,0372	-0,4268	0,2715	0,038	0,1065	0,3716	1

↓
Coficiente de correlación lineal de Pearson

Tabla de Resultados (Presentación de resultados en forma matricial)

FACULTAD DE CIENCIAS AGRONÓMICAS UNIVERSIDAD DE CHILE

Correlación de Pearson: Coeficientes \probabilidades

	CarneVacuna	CarneCerdo	Huevos	Leche	Pescado	Cereal	Embutidos	FrutosSecos	Frutas yVegetales
CarneVacuna	1	0,4103	0,0017	0,0135	0,7541	0,0091	0,4322	0,0646	0,7095
CarneCerdo	0,1761	1	0,0015	0,1145	0,2301	0,0552	0,1995	0,0013	0,8181
Huevos	0,6068	0,6102	1	0,0013	0,793	0,0001	0,0341	0,0057	0,8632
Leche	0,497	0,3307	0,6165	1	0,4763	0,0009	0,1715	0,0002	0,0375
Pescado	0,0675	-0,2545	0,0565	0,1527	1	0,009	0,0516	0,5292	0,1993
Cereal	-0,5203	-0,3963	-0,7064	-0,6349	-0,521	1	0,0094	0,0007	0,8599
Embutidos	0,1682	0,2714	0,4339	0,2886	0,4019	-0,5187	1	0,0329	0,6203
FrutosSecos	-0,3832	-0,6161	-0,547	-0,6937	-0,1351	0,6415	-0,4367	1	0,0738
Frutas yVegetales	-0,0802	-0,0496	-0,0372	-0,4268	0,2715	0,038	0,1065	0,3716	1

Tabla de Resultados (Presentación de resultados en forma matricial)

FACULTAD DE CIENCIAS AGRONÓMICAS UNIVERSIDAD DE CHILE

Correlación de Pearson: Coeficientes \probabilidades

	CarneVacuna	CarneCerdo	Huevos	Leche	Pescado	Cereal	Embutidos	FrutosSecos	Frutas yVegetales
CarneVacuna	1	0,4103	0,0017	0,0135	0,7541	0,0091	0,4322	0,0646	0,7095
CarneCerdo	0,1761	1	0,0015	0,1145	0,2301	0,0552	0,1995	0,0013	0,8181
Huevos	0,6068	0,6102	1	0,0013	0,793	0,0001	0,0341	0,0057	0,8632
Leche	0,497	0,3307	0,6165	1	0,4763	0,0009	0,1715	0,0002	0,0375
Pescado	0,0675	-0,2545	0,0565	0,1527	1	0,009	0,0516	0,5292	0,1993
Cereal	-0,5203	-0,3963	-0,7064	-0,6349	-0,521	1	0,0094	0,0007	0,8599
Embutidos	0,1682	0,2714	0,4339	0,2886	0,4019	-0,5187	1	0,0329	0,6203
FrutosSecos	-0,3832	-0,6161	-0,547	-0,6937	-0,1351	0,6415	-0,4367	1	0,0738
Frutas yVegetales	-0,0802	-0,0496	-0,0372	-0,4268	0,2715	0,038	0,1065	0,3716	1

↑ Probabilidades

Tabla de Resultados (Presentación de resultados en forma de lista)

FACULTAD DE CIENCIAS AGRONÓMICAS UNIVERSIDAD DE CHILE

Variable[1]	Variable[2]	n	Pearson	p-valor
CarneVacuna	CarneVacuna	24	1,0000	<0,0001
CarneVacuna	CarneCerdo	24	0,4103	
CarneVacuna	Huevos	24	0,6068	0,0017
CarneVacuna	Leche	24	0,4970	0,0135
CarneVacuna	Pescado	24	0,0675	0,7541
CarneVacuna	Cereal	24	-0,5203	0,0091
CarneVacuna	Embutidos	24	0,1682	0,4322
CarneVacuna	FrutosSecos	24	-0,3832	0,0646
CarneVacuna	Frutas yVeget	24	-0,0802	0,7095
CarneCerdo	CarneVacuna	24	0,1761	0,4103
CarneCerdo	CarneCerdo	24	1	<0,0001
CarneCerdo	Huevos	24	0,6102	0,0015
CarneCerdo	Leche	24	0,3307	0,1145
CarneCerdo	Pescado	24	-0,2545	0,2301
CarneCerdo	Cereal	24	-0,3963	0,0552
CarneCerdo	Embutidos	24	0,2714	0,1995
CarneCerdo	FrutosSecos	24	-0,6161	0,0013
CarneCerdo	Frutas yVeget	24	-0,0496	0,8181

(Nota: por espacio no se presenta la tabla completa)

Tabla de Resultados (Presentación de resultados en forma matricial)

FACULTAD DE CIENCIAS AGRONÓMICAS UNIVERSIDAD DE CHILE

Correlación de Pearson: Coeficientes \probabilidades

	CarneVacuna	CarneCerdo	Huevos	Leche	Pescado	Cereal	Embutidos	FrutosSecos	Frutas yVegetales
CarneVacuna	1	0,4103	0,0017	0,0135	0,7541	0,0091	0,4322	0,0646	0,7095
CarneCerdo	0,1761	1	0,0015	0,1145	0,2301	0,0552	0,1995	0,0013	0,8181
Huevos	0,6068	0,6102	1	0,0013	0,793	0,0001	0,0341	0,0057	0,8632
Leche	0,497	0,3307	0,6165	1	0,4763	0,0009	0,1715	0,0002	0,0375
Pescado	0,0675	-0,2545	0,0565	0,1527	1	0,009	0,0516	0,5292	0,1993
Cereal	-0,5203	-0,3963	-0,7064	-0,6349	-0,521	1	0,0094	0,0007	0,8599
Embutidos	0,1682	0,2714	0,4339	0,2886	0,4019	-0,5187	1	0,0329	0,6203
FrutosSecos	-0,3832	-0,6161	-0,547	-0,6937	-0,1351	0,6415	-0,4367	1	0,0738
Frutas yVegetales	-0,0802	-0,0496	-0,0372	-0,4268	0,2715	0,038	0,1065	0,3716	1

Tabla de Resultados (Presentación de resultados en forma de lista)

FACULTAD DE CIENCIAS AGRONÓMICAS
UNIVERSIDAD DE CHILE

$H_0: \rho=0$
 $H_1: \rho \neq 0$

Existe una correlación lineal significativa y positiva entre las variables Carne Vacuna y Huevos

Variable(1)	Variable(2)	n	Pearson	p-valor
CarneVacuna	CarneVacuna	24	1,0000	<0,0001
CarneVacuna	CarneCerdo	24	0,1761	0,4103
CarneVacuna	Huevos	24	0,6068	0,0017
CarneVacuna	Leche	24	0,4970	0,0135
CarneVacuna	Pescado	24	0,0675	0,7541
CarneVacuna	Cereal	24	-0,5203	0,0091
CarneVacuna	Embutidos	24	0,1682	0,4322
CarneVacuna	FrutosSecos	24	-0,3832	0,0646
CarneVacuna	Frutas yVeget	24	-0,0802	0,7095
CarneCerdo	CarneVacuna	24	0,1761	0,4103
CarneCerdo	CarneCerdo	24	1	<0,0001
CarneCerdo	Huevos	24	0,6102	0,0015
CarneCerdo	Leche	24	0,3307	0,1145
CarneCerdo	Pescado	24	-0,2545	0,2301
CarneCerdo	Cereal	24	-0,3963	0,0552
CarneCerdo	Embutidos	24	0,2714	0,1995
CarneCerdo	FrutosSecos	24	-0,6161	0,0013
CarneCerdo	Frutas yVeget	24	-0,0496	0,8181

(Nota: por espacio no se presenta la tabla completa)

Tabla de Resultados (Presentación de resultados en forma matricial)

FACULTAD DE CIENCIAS AGRONÓMICAS
UNIVERSIDAD DE CHILE

Correlación de Pearson: Coeficientes\probabilidades

	CarneVacuna	CarneCerdo	Huevos	Leche	Pescado	Cereal	Embutidos	FrutosSecos	Frutas yVegetales
CarneVacuna	1	0,4103	0,0017	0,0135	0,7541	0,0091	0,4322	0,0646	0,7095
CarneCerdo	0,1761	1	0,0015	0,1145	0,2301	0,0552	0,1995	0,0013	0,8181
Huevos	0,6068	0,6102	1	0,0013	0,793	0,0001	0,0341	0,0057	0,8632
Leche	0,497	0,3307	0,6165	1	0,4763	0,0009	0,1715	0,0002	0,0375
Pescado	0,0675	-0,2545	0,0565	0,1527	1	0,009	0,0516	0,5292	0,1993
Cereal	-0,5203	-0,3963	-0,7064	-0,6349	-0,521	1	0,0094	0,0007	0,8599
Embutidos	0,1682	0,2714	0,4339	0,2886	0,4019	-0,5187	1	0,0329	0,6203
FrutosSecos	-0,3832	-0,6161	-0,547	-0,6937	-0,1351	0,6415	-0,4367	1	0,0738
Frutas yVegetales	-0,0802	-0,0496	-0,0372	-0,4268	0,2715	0,038	0,1065	0,3716	1

Existe una correlación lineal significativa y negativa entre las variables Frutos Secos y Carne de Cerdo

Tabla de Resultados (Presentación de resultados en forma matricial)

FACULTAD DE CIENCIAS AGRONÓMICAS
UNIVERSIDAD DE CHILE

Correlación de Pearson: Coeficientes\probabilidades

	CarneVacuna	CarneCerdo	Huevos	Leche	Pescado	Cereal	Embutidos	FrutosSecos	Frutas yVegetales
CarneVacuna	1	0,4103	0,0017	0,0135	0,7541	0,0091	0,4322	0,0646	0,7095
CarneCerdo	0,1761	1	0,0015	0,1145	0,2301	0,0552	0,1995	0,0013	0,8181
Huevos	0,6068	0,6102	1	0,0013	0,793	0,0001	0,0341	0,0057	0,8632
Leche	0,497	0,3307	0,6165	1	0,4763	0,0009	0,1715	0,0002	0,0375
Pescado	0,0675	-0,2545	0,0565	0,1527	1	0,009	0,0516	0,5292	0,1993
Cereal	-0,5203	-0,3963	-0,7064	-0,6349	-0,521	1	0,0094	0,0007	0,8599
Embutidos	0,1682	0,2714	0,4339	0,2886	0,4019	-0,5187	1	0,0329	0,6203
FrutosSecos	-0,3832	-0,6161	-0,547	-0,6937	-0,1351	0,6415	-0,4367	1	0,0738
Frutas yVegetales	-0,0802	-0,0496	-0,0372	-0,4268	0,2715	0,038	0,1065	0,3716	1

No existe correlación lineal entre las variables Leche y Carne de Cerdo.

Análisis de Correlación Lineal

FACULTAD DE CIENCIAS AGRONÓMICAS
UNIVERSIDAD DE CHILE

Erika Kania Kuhl
Ing. Agrónomo Dr.

Archivo Añes

Se realizó un estudio con 21 variedades de añes, en donde se midieron características de: altura de planta, materia seca, sólidos solubles, número de semillas por fruto, diámetro ecuatorial, longitud del fruto, peso del fruto y número de frutos por planta.



Objetivo

Medir y evaluar el grado de asociación lineal entre las variables evaluadas

Archivo Añes

Caso	Variedad	Altura de planta	Materia Seca	Sólidos solubles	N semillas	Diámetro ecuatorial del fruto	Longitud fruto	Peso fruto	N frutos por planta
1	1	50,60	18,54	11,84	152,43	3,17	11,33	25,85	4,17
2	2	51,80	12,90	9,99	152,17	2,98	17,73	43,43	8,30
3	3	43,80	12,11	9,14	220,00	3,92	15,22	67,06	4,70
4	4	37,00	13,59	9,72	169,14	3,70	15,49	50,53	4,27
5	5	47,40	12,04	9,12	243,83	4,09	16,61	73,26	6,00
6	6	41,00	11,15	8,07	211,40	4,34	15,89	71,12	4,40
7	7	45,20	14,56	10,31	169,80	2,79	14,84	34,58	7,85
8	8	54,17	14,35	9,90	212,50	2,96	15,07	36,39	10,30
9	9	46,25	13,90	9,41	51,00	1,45	6,50	7,10	22,00
10	10	26,30	10,01	7,98	64,50	2,35	7,20	21,28	9,79
11	11	13,50	14,79	9,45	140,43	2,46	10,59	18,74	3,92
12	12	39,70	14,10	9,80	117,40	2,76	12,13	22,28	3,70
13	13	37,40	13,66	9,86	161,00	3,10	13,86	39,37	4,56
14	14	39,90	12,21	10,39	115,33	2,61	12,99	31,16	4,56
15	15	38,20	11,59	8,24	153,50	3,63	16,86	60,88	3,91
16	16	35,00	11,08	7,79	164,17	4,16	5,20	29,26	11,42
17	17	29,60	11,72	8,75	163,86	3,59	7,10	31,27	10,67
18	18	41,14	13,65	9,92	110,33	2,65	13,87	31,46	4,33
19	19	23,40	8,25	7,52	204,83	6,21	10,18	105,78	3,06
20	20	24,20	8,62	6,86	210,83	7,38	9,25	155,19	4,10
21	21	38,25	17,11	12,58	56,71	0,97	10,31	4,36	4,03



FACULTAD DE CIENCIAS
AGRONÓMICAS
UNIVERSIDAD DE CHILE

R^2 v/s r

En **Regresión lineal simple**, el Coeficiente de Determinación (R^2) coincide con el coeficiente de correlación lineal (r) al cuadrado.

$$r = \pm\sqrt{R^2}$$

