

# Asignación 5: Correlación

Dr. Marco Aurelio González Tagle

26 /08/ 2021

## Índice

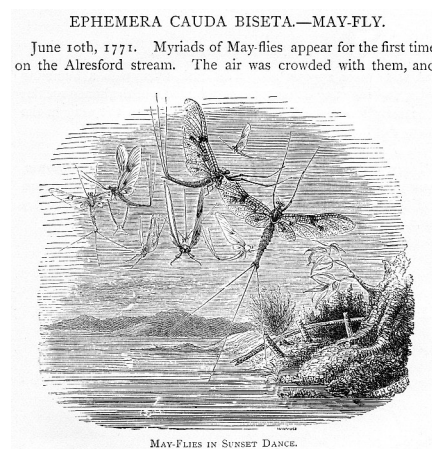
Instrucciones . . . . .	1
Ejercicio 1 . . . . .	1
Ejercicio 2 . . . . .	2

## Instrucciones

Para cada ejercicio:

- Examinar la relación que existe entre dos muestras mediante una correlación,
- Explore los datos gráficamente y explique,
- Establezca la Hipótesis nula y la Hipótesis alternativa,
- Aplique la prueba correspondiente,
- Reporte los datos (indicar valor de  $r$ , grados de libertad y probabilidad, así como la significancia de la correlación).

## Ejercicio 1



El cuadro 1 muestra los datos recopilados sobre [efímeras](#) de un arroyo y para cada muestra medimos la velocidad del arroyo y el número de efímeras.

Cuadro 1: Abundancia de las efímeras de un arroyo y la velocidad de la corriente de un arroyo.

Speed	2	3	5	9	14	24	29	34
Abundance	6	3	5	23	16	12	48	43

Si creamos un diagrama de dispersión de los datos del cuadro 1, vemos el patrón con mayor claridad (Figura 1).

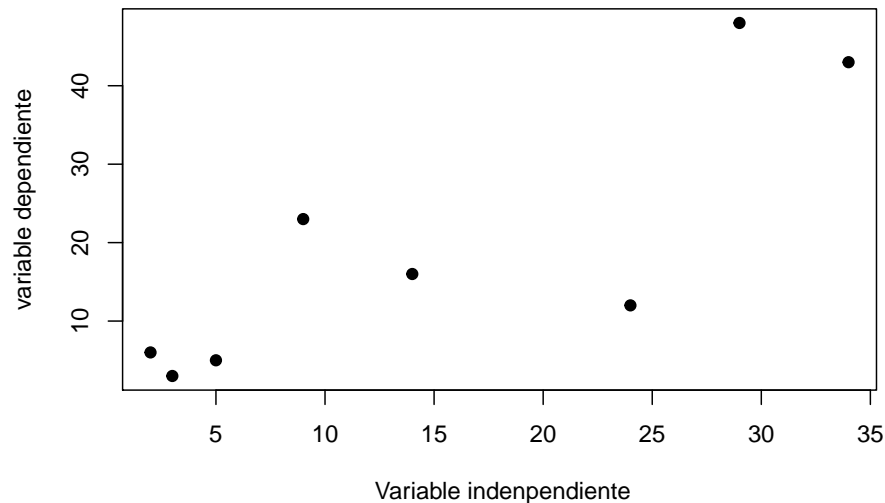


Figura 1: Diagrama de dispersión de efímeros (*Ecdyonurus dispar*) y datos de velocidad del flujo. La velocidad de la corriente es el eje independiente y los datos de la mosca de mayo el eje dependiente

Parece que puede haber una relación entre la velocidad y la abundancia, pero hay varias inconsistencias, no es una correlación perfecta. Contestar las siguientes interrogantes: + ¿Es estadísticamente significativa la correlación?

Comenzamos con una hipótesis para probar. En este caso, esperamos una mayor abundancia de efímeras a medida que aumenta la velocidad del arroyo. Escribiríamos esto formalmente (nuestra hipótesis alternativa H1) como:

“Existe una correlación positiva entre la velocidad de los arroyos y la abundancia de efímeras (*Ecdyonurus dispar*)”.

Estamos diciendo que la relación es positiva debido a la investigación previa que hemos realizado (lectura de artículos, etc.). Nuestra hipótesis nula (H0) ahora se convierte en:

“No existe una correlación entre la velocidad del arroyo y la abundancia de efímeras”.

## Ejercicio 2

Conjunto de datos: Composiciones del suelo, características físicas y químicas.

Descripción: Las características del suelo se midieron en muestras de tres áreas diferentes (cima, pendiente y depresión) y a cuatro profundidades (0-10 cm, 10-30 cm, 30-60 cm y 60-90 cm). El área se dividió en 4 bloques, en un diseño de bloques al azar (Cuadro 2).

Cuadro 2: Datos de la estructura del suelo.

Gp	Block	pH	N	Dens	P	Ca	Mg	K	Na	Conduc
T0	1	5.40	0.188	0.92	215	16.35	7.65	0.72	1.14	1.09
T0	2	5.65	0.165	1.04	208	12.25	5.15	0.71	0.94	1.35
T0	3	5.14	0.260	0.95	300	13.02	5.68	0.68	0.60	1.41
T0	4	5.14	0.169	1.10	248	11.92	7.88	1.09	1.01	1.64
T1	1	5.14	0.164	1.12	174	14.17	8.12	0.70	2.17	1.85
T1	2	5.10	0.094	1.22	129	8.55	6.92	0.81	2.67	3.18
T1	3	4.70	0.100	1.52	117	8.74	8.16	0.39	3.32	4.16

Realizar un análisis de correlación para las variables y reportar en un cuadro los valores de coeficiente de correlación y su valor de significancia (p-value):

Cuadro 3: Ejemplo de cuadro de datos con los estadísticos de interés.

Conjunto	r	valor de P
pH - N	–	–
pH - Dens		
pH - P		
pH - Ca		
pH - Mg		
pH - K		
pH - Na		

Para poder orientarte en las correlaciones, la figura 2 muestra de forma gráfica las diferentes correlaciones entre todas las variables.

Las correlaciones positivas se muestran en azul y las correlaciones negativas en rojo. La intensidad del color y el tamaño del círculo son proporcionales a los coeficientes de correlación. En el lado derecho del correlograma, el color de la leyenda muestra los coeficientes de correlación y los colores correspondientes.

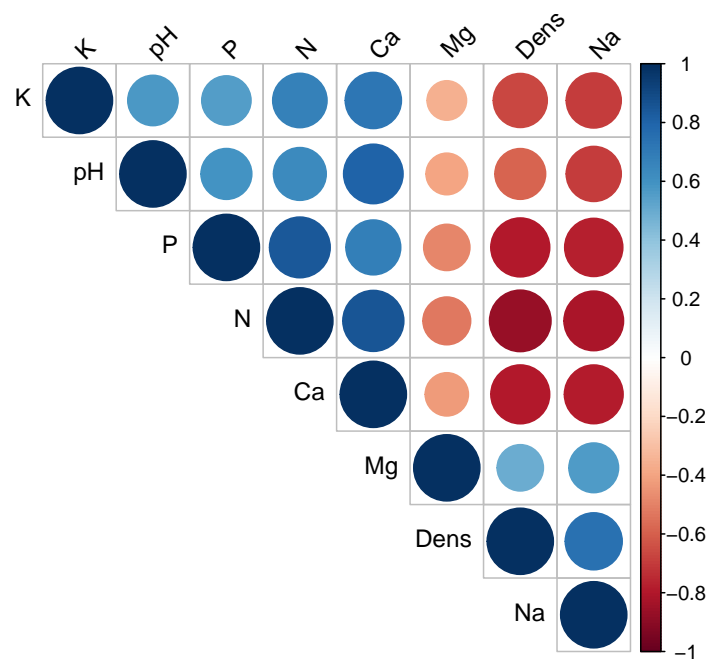


Figura 2: Gráfica de las correlaciones de la base de datos suelo