

---

# Aplicando la estadística. Correlación y regresión

---

Herramientas para practicar la endoscopia basada en la evidencia



---

Endoscopia & Mentoring

---

# Parecen lo mismo pero no son iguales

Como ya vimos en el tema 8, la relación entre variables es algo complejo porque, aunque demos que existe una asociación estadísticamente significativa, esto no significa que exista una relación causal para esto.

En estadística disponemos de 2 tipos de análisis, que la gente suele confundir, porque una cosa es la correlación y otra cosa diferente es la regresión, y por ello en este tema, vamos a tratar estos dos conceptos con detalle para que no cometamos más errores a la hora de interpretarlos.

## Correlación

En primer lugar, fijémonos en este gráfico (figura 1). Tenemos dos variables que son la variable **A** y la variable **B** para las cuales existe una nube de puntos que van a traducir la relación entre ambas variables; esta relación además queda resumida con una línea recta que no es más que el resultado de la ecuación que resume la relación de ambas.

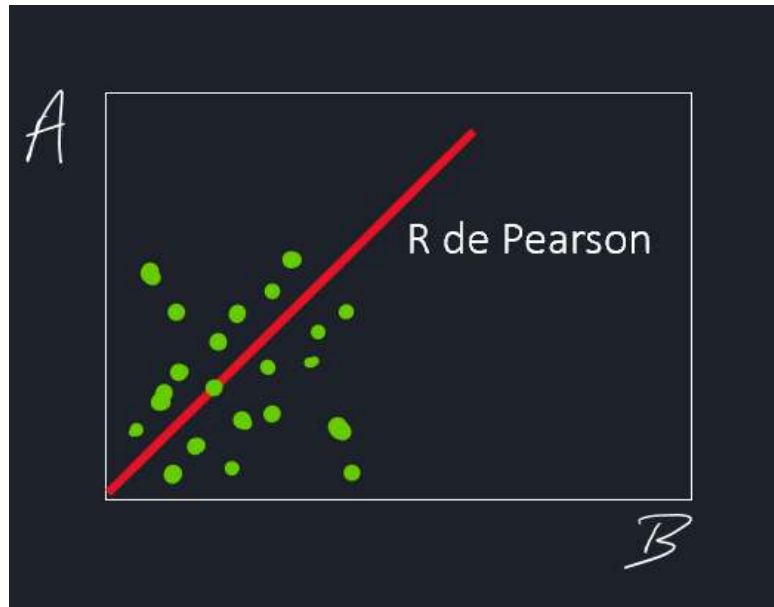
Esta relación estadística se determina mediante el cociente de correlación lineal de Pearson o también denominada R de Pearson.

### **Para la aplicación de la prueba R de Pearson, deben cumplirse estos supuestos:**

- Los datos deben tener una distribución normal, de lo contrario debemos aplicar el test para pruebas no paramétricas que sería la Rho Spearman.
- Las dos variables deben ser estructuralmente independientes, si no lo son, deben utilizarse pruebas pareadas (no paramétricas).
- Sólo deben realizarse dos mediciones por cada participante en el estudio (medición A y medición B), no más porque incurriríamos en sesgos.

---

Por supuesto cada valor **R** debe acompañarse de una significación estadística (valor P) que expresa cuál es la probabilidad de que una asociación haya surgido por el azar.



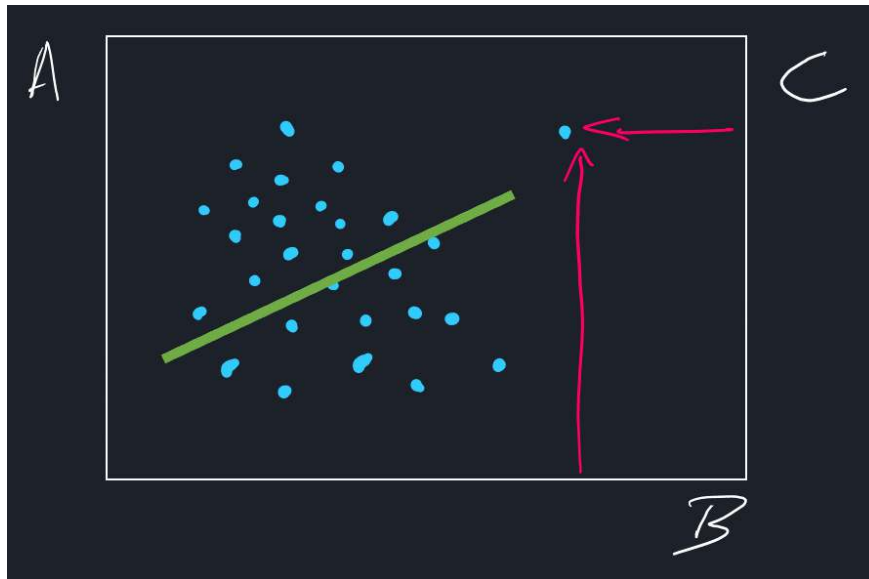
*Figura 1. Representación gráfica de la correlación lineal de Pearson*

## Regresión

El término regresión viene a definir a una ecuación matemática que permite predecir una variable a partir de otra variable diferente.

Por tanto una variable influye sobre la otra y de este modo puede predisponer a la aparición de esta. **De este modo la variable influenciada será la denominada por la variable de destino y la variable que influncia será la variable independiente.**

Cuando existen más de una variable hablamos de Covariantes, de modo que puede aparecer una tercera variable, que denominaremos C, que también influye en la ecuación. De esta forma, si realizamos una representación gráfica observaríamos como esta regresión es consecuencia de la relación de los 3 elementos, pudiendo inferir una relación en la que B y C predicen la aparición de A (Figura 2).



**Figura 2. Representación gráfica de una regresión**

Por tanto aunque correlación y regresión parezcan iguales, como hemos visto no lo son y la diferencia radica fundamentalmente en la capacidad de influencia de unas variables sobre otras; en la correlación, no se da esta circunstancia, cosa que en la regresión si ocurre.