

# 1.4 Correlación y Diseño Experimental

## 1. Estadística y Probabilidad Básica

Christian E. Galarza

**Programa New Dimensions**

# Correlación

- La correlación mide la fuerza y dirección de la relación lineal entre dos variables.
- El coeficiente de correlación de Pearson varía entre -1 y 1.
- $r = 1$ : correlación positiva perfecta;  $r = -1$ : correlación negativa perfecta;  $r = 0$ : no hay correlación.

# Caveats de la Correlación

- La correlación no implica causalidad.
- Sensible a valores atípicos.
- No puede medir relaciones no lineales.

# Transformación de Variables

- A veces, la transformación de variables puede revelar una relación lineal.
- Ejemplo: aplicar logaritmo o raíz cuadrada.

# Confounders

- Un confounder es una variable que está relacionada tanto con la variable independiente como con la dependiente.
- Puede crear o ocultar una relación aparente entre las dos variables de interés.

# ¿Qué es el Diseño de Experimentos?

- Es una técnica estadística utilizada para planificar, diseñar y analizar experimentos.
- Ayuda a maximizar la obtención de información, minimizando el uso de recursos.
- Se enfoca en identificar las causas de variación en procesos y sistemas.
- Permite estudiar los efectos de una o más variables independientes sobre una variable respuesta.

# Principios Básicos

- **Replicación:** Realizar el mismo experimento varias veces para obtener una estimación de la variabilidad.
- **Randomización:** Asignar al azar las unidades experimentales para evitar sesgos.
- **Bloqueo:** Agrupar unidades experimentales similares para reducir la variabilidad y aislar factores de interés.

Ejemplo: Si queremos evaluar un nuevo medicamento, las personas podrían ser asignadas al azar a grupos de tratamiento o control, y se podría replicar el estudio en diferentes lugares para obtener resultados más robustos.

# Tipos de Estudios

- Estudios Observacionales: el investigador observa sin intervenir.
- Estudios Experimentales: el investigador interviene y controla las variables.



# Estudios Observacionales

- Se basan en la observación y registro de eventos o características.
- No hay intervención por parte del investigador.
- Pueden ser descriptivos (simplemente observan y describen) o analíticos (buscan relaciones entre variables).
- Ejemplos: Estudios de cohortes, estudios transversales y estudios de casos y controles.

# Estudios Experimentales

- El investigador interviene y controla una o más variables.
- Se busca determinar una relación causa-efecto.
- Suelen incluir un grupo control y un grupo experimental.
- Los participantes son asignados aleatoriamente a los grupos.
- Ejemplo: Ensayos clínicos controlados, donde se evalúa la efectividad de un tratamiento.

# Estudios Longitudinales vs. Transversales

- Estudios Longitudinales: se estudian a los individuos o unidades durante un período de tiempo.
- Estudios Transversales: observan a los individuos o unidades en un solo punto en el tiempo.

# Estudios Longitudinales

- Estudian a los mismos sujetos a lo largo de un período de tiempo.
- Observan cambios o evoluciones en las características o variables de interés.
- Pueden durar desde meses hasta décadas.
- Son útiles para identificar tendencias o cambios a lo largo del tiempo.
- Ejemplo: Un estudio que sigue a una cohorte de individuos desde su infancia hasta la adultez para estudiar el desarrollo de ciertas enfermedades.

# Estudios Transversales

- Se realizan en un punto específico en el tiempo.
- Se observan diferentes grupos o categorías al mismo tiempo.
- No se sigue a los mismos sujetos a lo largo del tiempo como en los estudios longitudinales.
- Son útiles para obtener una “fotografía” instantánea de una población en un momento dado.
- Ejemplo: Una encuesta de salud que busca determinar la prevalencia de una enfermedad en diferentes grupos de edad en un año específico.

# Ejemplo: Inversión en Publicidad vs Ventas

Supongamos que una empresa quiere analizar la relación entre la inversión en publicidad (en miles de dólares) y las ventas mensuales de un producto (en miles de unidades). Se recopilan los datos de los últimos 5 meses:

- Inversión en publicidad:  $X = [10, 20, 30, 40, 50]$
- Ventas mensuales:  $Y = [50, 55, 65, 75, 90]$

El objetivo es determinar si invertir más en publicidad está correlacionado con un aumento en las ventas.

# Cálculo de la Correlación en R

Para evaluar esta relación, podemos calcular la correlación en R utilizando la función `cor()`.

```
# Datos
```

```
inversion <- c(10, 20, 30, 40, 50)
```

```
ventas <- c(50, 55, 65, 75, 90)
```

```
# Calcular la correlación
```

```
correlacion <- cor(inversion, ventas)
```