

Introducción al diseño experimental

Alejandro Zubiri

January 17, 2025

Contents

1	Introducción	2
2	Tipos de variabilidad	2
2.1	Sistemática planificada	2
2.2	Sistemática no planificada	2
3	Definiciones	2
3.1	Unidad experimental	2
3.2	Variable de interés	2
3.3	Observación experimental	2
3.4	Tamaño del experimento	2
3.5	Factores	2
3.6	Niveles	2
3.7	Tratamiento	2
4	Básicos del diseño experimental	3
4.1	Aleatorización	3
4.2	Bloqueo	3
4.3	Factorización	3
4.4	Diseño Experimental	3
5	Diseño completamente aleatorizado	3
6	Análisis de la varianza	3
6.1	Condiciones	3

1 Introducción

Cuando realizamos un experimento varias veces, se suelen producir variaciones, que se conocen como **ruido**. Las posibles fuentes de error son:

- Error experimental: factores tanto conocidos como desconocidos.
- Confusión entre correlación y causalidad.
- Complejidad de los elementos estudiados.

Cuando realizamos experimentos, queremos los métodos estadísticos que minimicen este error.

2 Tipos de variabilidad

2.1 Sistemática planificada

- Planificada por el experimentador.
- Por las condiciones del experimento.

2.2 Sistemática no planificada

Debido a causas externas desconocidas que producen resultados sesgados.

3 Definiciones

3.1 Unidad experimental

Objetos, individuos, espacio o tiempo en los que se experimenta.

3.2 Variable de interés

Lo que se desea estudiar.

3.3 Observación experimental

Cada medición que realizamos.

3.4 Tamaño del experimento

Número total de observaciones recogidas.

3.5 Factores

Variables independientes que pueden influir en los resultados.

3.6 Niveles

Cada uno de los resultados de un factor.

3.7 Tratamiento

Como se combinan los diferentes niveles de un factor.

4 Básicos del diseño experimental

4.1 Aleatorización

Tiene el objetivo de evitar el sesgo en el experimento. Se asignan al azar los tratamientos. También se evita la dependencia de las observaciones.

4.2 Bloqueo

Se dividen las unidades experimentales en grupos llamados **bloques**. Las observaciones de cada bloque deben tener condiciones experimentales similares.

4.3 Factorización

Cruzamiento de los niveles de todos los factores de tratamiento en todas las combinaciones posibles. Permite detectar la existencia de interacciones entre tratamientos. Además, es mucho más eficiente.

4.4 Diseño Experimental

Regla con la que se determina la asignación de las unidades experimentales:

- Completamente aleatorizado.
- Factor bloque.
- Factorial a dos niveles.

5 Diseño completamente aleatorizado

Los tamaños muestrales no tienen por qué ser iguales, aunque tiene ventajas:

- Asegura que cada tratamiento contribuye igual.
- Reduce problemas derivados de incumplimiento de hipótesis.
- Incrementa la potencia del test.

Realizamos la siguiente pregunta: ¿Son todas las medias iguales?

Las observaciones del tratamiento j provienen de una población común μ , un efecto del tratamiento A_j y un error aleatorio e_{ij} .

$$y_{ij} = \mu + A_i + e_{ij} \tag{1}$$

6 Análisis de la varianza

Queremos saber si cada factor tiene un efecto significativo sobre Y , o si hay dependencia entre factores.

6.1 Condiciones

- Independencia: tamaño total de la muestra $< 10\%$.
- Normalidad: La variable cuantitativa se distribuye normalmente en cada grupo.
- Homocedasticidad: varianza constante entre grupos.

Tenemos las siguientes tipos de varianza:

- Suma de cuadrados total (SCT)
- Suma de cuadrados del tratamiento ($SCTr$)
- Suma de cuadrados residual (SCR)

Se cumple que:

$$SCT = SCTr + SCR \quad (2)$$

- SCT: variabilidad total de los datos:

$$SCT = \sum_i \sum_j (y_{ij} - \bar{y})^2 \quad (3)$$

- SCTr: variabilidad de los datos asociada al efector del factor sobre la media, ponderado por el número de observaciones:

$$SCTr = \sum_i (\bar{y}_i - \bar{y})^2 n_i \quad (4)$$

- SCR: variabilidad no debida a la variable factor. Es la suma de los cuadrados dentro de los grupos:

$$SCR = \sum_i \sum_j (y_{ij} - \bar{y}_i)^2 = SCT - SCTr \quad (5)$$

Ahora, podemos analizar cada grupo Además, se cumple que:

Variabilidad	Suma de cuadrados	Grados de libertad	Cuadrados medios
Intergrupo	SCTr	$k - 1$	$CMT_r = \frac{SCT_r}{k-1}$
Intragrupos	SCR	$n - 1$	$CMR = \frac{SCR}{n-k}$
Total	SCT	$n - 1$	$CMT = \frac{SCT}{n-1}$

$$\frac{\text{intervarianza}}{\text{intravarianza}} = \frac{s_{Tr}^2}{s_R^2} \approx F_{k-1, n-k} \quad (6)$$

Siendo F la F de Fischer. Si se cumple que:

$$\frac{s_{Tr}^2}{s_r^2} > F_{k-1, n-k} \quad (7)$$

Rechazamos la igualdad de medias.