Diseño factorial general

Experimentos de tres factores con análisis de resultados.

Felipe Neira Rojas & Angel Llanos Herrera

Profesora: Carolina Marchant Fuentes
Diseño de Experimentos [IES-413]

Universidad Católica del Maule

03 de julio, 2025

Índice

Diseño factorial general: Breve historia y contexto

- Origen agronómico: necesidad de estudiar varios factores que afectan a los cultivos.
- R. A. Fisher (1926): propone el diseño factorial para controlar la variabilidad natural.
- Introduce el ANOVA como método para comparar tratamientos.
- Establece el diseño factorial completo para evaluar múltiples factores a la vez.

Diseño factorial general: Definición y ventajas

- Un diseño factorial estudia 2 factores a la vez e incluye todas las combinaciones de sus niveles; cada combinación constituye un tratamiento.
- Interacciones: revela si el efecto de un factor depende del nivel de otro, ofreciendo conclusiones más completas.
- Realismo: refleja mejor las condiciones del mundo real, donde los factores actúan juntos.
- Capacidad explicativa: al incluir más factores, se captura mayor parte de la variabilidad del sistema estudiado.

Diseño factorial general: Supuestos básicos del modelo

El análisis de DFG se realiza bajo los siguientes supuestos:

$$\varepsilon_{ijkl} \sim \mathcal{N}(0, \sigma^2), \quad \forall\, i, j, k, l$$

- **Normalidad de los Residuos:** Los errores de cada tratamiento siguen una distribución normal con media 0 y varianza σ^2 constante.
- **Homocedasticidad:** La varianza de los residuos es constante.
- **Nota:** La independencia de los errores se garantiza asignando aleatoriamente cada tratamiento a las unidades experimentales, tal como exige el diseño factorial general.

Diseño factorial con tres factores: Definición y notación

El Diseño Factorial General (DFG) se utiliza para analizar el efecto de dos o más factores sobre una variable de respuesta. A diferencia de los diseños unifactoriales, este enfoque permite evaluar no solo los efectos principales de cada factor por separado, sino también las interacciones que puedan existir entre ellos. Cada factor puede tener un número distinto de niveles, siempre que cuente con al menos dos niveles. Así, en un diseño factorial general con tres factores, se consideran: el factor A con a niveles, el factor B con b niveles y el factor C con c niveles, cumpliéndose que $r, s, t \geq 2$.

El análisis ANOVA Factorial de DFG busca:

- **Analizar los efectos principales de cada factor:** Permite determinar si cada factor, considerado de forma individual, ejerce una influencia significativa sobre la variable de respuesta.
- **Detectar interacciones entre factores:** Consiste en identificar si el efecto de un factor varía en función del nivel de otro, lo que evidencia una dependencia entre factores, sugiriendo que no actúan de manera independiente.

Diseño factorial con tres factores: Modelo estadístico

Se plantea un Modelo ANOVA factorial para la variable de respuesta y_{ijkl}, considerando 3 factores:

$$y_{ijkl} = \mu + \alpha_i + \beta_j + \gamma_k + (\alpha\beta)_{ij} + (\alpha\gamma)_{ik} + (\beta\gamma)_{jk} + (\alpha\beta\gamma)_{ijk} + \varepsilon_{ijkl}$$

- y_{ijkl}: Variable de respuesta correspondiente al nivel i del factor A, nivel j del factor B, nivel k del factor C, en la réplica l
- μ: Media global
- α_i: Efecto del nivel i del factor A, con $i = 1, 2, \ldots, r$, dónde $(a \geq 2)$
- β_j: Efecto del nivel j del factor B, con $j = 1, 2, \ldots, s$, dónde $(b \geq 2)$
- γ_k: Efecto del nivel k del factor C, con $k = 1, 2, \ldots, t$, dónde $(c \geq 2)$
- $(\alpha\beta)_{ij}$: Efecto de la interacción entre A y B
- $(\alpha\gamma)_{ik}$: Efecto de la interacción entre A y C
- $(\beta\gamma)_{jk}$: Efecto de la interacción entre B y C
- $(\alpha\beta\gamma)_{ijk}$: Efecto de la interacción entre A, B y C
- ε_{ijkl}: Error aleatorio en el nivel i del factor A, nivel j del factor B, nivel k del factor C, en la observación l.

Tipo de efecto	Condición de Estimabilidad
Efectos	$\sum_{i=1}^{a} \hat{\alpha}_i = 0$
	$\sum_{j=1}^{b} \hat{\beta}_j = 0$
	$\sum_{k=1}^{c} \hat{\gamma}_k = 0$
Interacciones dobles	$\sum_{i=1}^{a} \sum_{j=1}^{b} (\widehat{\alpha\beta})_{ij} = 0$
	$\sum_{i=1}^{a} \sum_{k=1}^{c} (\widehat{\alpha\gamma})_{ik} = 0$
	$\sum_{j=1}^{b} \sum_{k=1}^{c} (\widehat{\beta\gamma})_{jk} = 0$
Interacción triple	$\sum_{i=1}^{a} \sum_{j=1}^{b} \sum_{k=1}^{c} \widehat{(\alpha\beta\gamma)}_{ijk} = 0$

Diseño factorial con tres factores: Estimación de parámetros

La estimación de parámetros se llevará a cabo mediante el método de máxima verosimilitud considerando nuestro vector de parámetros como:

$$\boldsymbol{\theta} = (\mu, \alpha_i, \beta_j, \gamma_k, (\alpha\beta)_{ij}, (\alpha\gamma)_{ik}, (\beta\gamma)_{jk}, (\alpha\beta\gamma)_{ijk}, \sigma^2)^\top, \quad i = 1, \dots, a; \quad j = 1, \dots, b; \quad k = 1, \dots, c$$

Por lo tanto, se aplica la función de máxima verosimilitud. l = 1, . . . , n.

$$L(\boldsymbol{\theta}) = \prod_{i=1}^{a}\prod_{j=1}^{b}\prod_{k=1}^{c}\prod_{l=1}^{n} \frac{1}{\sqrt{2\pi\sigma^2}} \exp\left(-\frac{(y_{ijkl} - \mu_{ijk})^2}{2\sigma^2}\right)$$

Diseño factorial con tres factores: Estimación de parámetros

A esto aplicamos función de log-verosimilitud, quedando:

$$\ell(\boldsymbol{\theta}) = \sum_{i,j,k,l} \left[-\frac{1}{2}\log(2\pi\sigma^2) - \frac{(y_{ijkl} - \mu - \alpha_i - \beta_j - \gamma_k - (\alpha\beta)_{ij} - (\alpha\gamma)_{ik} - (\beta\gamma)_{jk} - (\alpha\beta\gamma)_{ijk})^2}{2\sigma^2} \right]$$

$$= -\frac{abcn}{2}\log(2\pi\sigma^2) - \frac{1}{2\sigma^2}\sum_{i,j,k,l}(y_{ijkl} - \mu - \alpha_i - \beta_j - \gamma_k - (\alpha\beta)_{ij} - (\alpha\gamma)_{ik} - (\beta\gamma)_{jk} - (\alpha\beta\gamma)_{ijk})^2$$

Para la estimación de los parámetros $\boldsymbol{\theta} = (\mu, \alpha_i, \beta_j, \gamma_k, (\alpha\beta)_{ij}, (\alpha\gamma)_{ik}, (\beta\gamma)_{jk}, (\alpha\beta\gamma)_{ijk})^\top$ se obtuvo la siguiente derivada:

$$\frac{1}{\sigma^2}\sum_{i=1}^{a}\sum_{j=1}^{b}\sum_{k=1}^{c}\sum_{l=1}^{n}(y_{ijkl} - \mu - \alpha_i - \beta_j - \gamma_k - (\alpha\beta)_{ij} - (\alpha\gamma)_{ik} - (\beta\gamma)_{jk} - (\alpha\beta\gamma)_{ijk}) = 0$$

Diseño factorial con tres factores: Estimación de parámetros

Para μ

$$\sum_{i,j,k,l}^{N} y_{ijkl} - \sum_{i,j,k,l}^{N} \mu = 0$$

$$\sum_{i,j,k,l}^{N} \mu = \sum_{i,j,k,l}^{N} y_{ijkl}$$

$$N\mu = \sum_{i,j,k,l}^{N} y_{ijkl}$$

$$\hat{\mu} = \frac{\sum_{i,j,k,l}^{N} y_{ijkl}}{N} = \bar{y}_{....}$$

Diseño factorial con tres factores: Estimación de parámetros

Para $(\alpha\beta)_{ij}$

$$\sum_{k,l}^{cn} y_{ijkl} - \sum_{k,l}^{cn} \mu - \sum_{k,l}^{cn} \alpha_i - \sum_{k,l}^{cn} \beta_j - \sum_{k,l}^{cn} (\alpha\beta)_{ij} = 0$$

$$cn \cdot (\alpha\beta)_{ij} = \sum_{k,l}^{cn} y_{ijkl} - cn \cdot \mu - cn \cdot \alpha_i - cn \cdot \beta_j$$

$$(\hat{\alpha\beta})_{ij} = \frac{\sum_{k,l}^{cn} y_{ijkl}}{cn} - \hat{\mu} - \hat{\alpha}_i - \hat{\beta}_j$$

$$(\hat{\alpha\beta})_{ij} = \frac{\sum_{k,l}^{cn} y_{ijkl}}{cn} - \bar{y}_{....} - (\bar{y}_{i...} - \bar{y}_{....}) - (\bar{y}_{.j..} - \bar{y}_{....})$$

$$(\hat{\alpha\beta})_{ij} = \bar{y}_{ij..} - \bar{y}_{i...} - \bar{y}_{.j..} + \bar{y}_{....}$$

Diseño factorial con tres factores: Estimación de parámetros

Para σ^2

$$\ell(\sigma^2) = -\frac{abcn}{2}\log(2\pi\sigma^2) - \frac{1}{2\sigma^2}\sum_{i,j,k,l}^{a,b,c,n}\Big(y_{ijkl}-\mu-\alpha_i-\beta_j-\gamma_k-(\alpha\beta)_{ij}-(\alpha\gamma)_{ik}-(\beta\gamma)_{jk}-(\alpha\beta\gamma)_{ijk}\Big)^2$$

$$\frac{\partial\ell}{\partial\sigma^2} = -\frac{abcn}{2\sigma^2} + \frac{1}{2\sigma^4}\sum_{i=1}^{a}\sum_{j=1}^{b}\sum_{k=1}^{c}\sum_{l=1}^{n}\Big(y_{ijkl}-\mu-\alpha_i-\beta_j-\gamma_k-(\alpha\beta)_{ij}-(\alpha\gamma)_{ik}-(\beta\gamma)_{jk}-(\alpha\beta\gamma)_{ijk}\Big)^2 = 0.$$

$$\Longrightarrow \qquad \hat{\sigma}^2 = \frac{1}{abcn}\sum_{i=1}^{a}\sum_{j=1}^{b}\sum_{k=1}^{c}\sum_{l=1}^{n}\Big(\epsilon_{ijkl}\Big)^2$$

$$\epsilon_{ijkl} = y_{ijkl} - \hat{\mu} - \hat{\alpha}_i - \hat{\beta}_j - \hat{\gamma}_k - \widehat{(\alpha\beta)}_{ij} - \widehat{(\alpha\gamma)}_{ik} - \widehat{(\beta\gamma)}_{jk} - \widehat{(\alpha\beta\gamma)}_{ijk}$$

Estimadores Insesgados para μ

$$
\begin{aligned}
\mathrm{E}[\hat{\mu}] &= \frac{1}{N} \sum_{i,j,k,l} \mathrm{E}[y_{ijkl}] \\
&= \frac{1}{N} \sum_{i,j,k,l} \Big(\mu + \alpha_i + \beta_j + \gamma_k + (\alpha\beta)_{ij} + (\alpha\gamma)_{ik} + (\beta\gamma)_{jk} + (\alpha\beta\gamma)_{ijk} \Big) \\
&= \mu + \underbrace{\frac{1}{N} \sum_{i} \alpha_i \sum_{j,k,l} 1}_{=0} + \underbrace{\frac{1}{N} \sum_{j} \beta_j \sum_{i,k,l} 1}_{=0} + \cdots \\
&= \mu.
\end{aligned}
$$

$\mathrm{E}[\hat{\mu}] = \mu \implies \hat{\mu}$ es **insesgado**

Estimadores Insesgados para σ^2

$$\mathrm{SCE} = \sum_{i=1}^{a}\sum_{j=1}^{b}\sum_{k=1}^{c}\sum_{\ell=1}^{n} e_{ijkl}^2, \qquad N = a\,b\,c\,n.$$

$$\frac{\mathrm{SCE}}{\sigma^2} = \sum_{i,j,k,\ell} \frac{e_{ijkl}^2}{\sigma^2} \sim \chi^2_{N-p}.$$

Por la propiedad de la ji-cuadrada,

$$\mathbb{E}\left[\chi^2_{N-p}\right] = N - p \quad \Longrightarrow \quad \mathbb{E}\left[\frac{\mathrm{SCE}}{\sigma^2}\right] = N - p.$$

Luego,

$$\mathbb{E}[\mathrm{SCE}] = (N-p)\,\sigma^2 \quad \Longrightarrow \quad \mathbb{E}\left[\widehat{\sigma}^2_{\mathrm{CME}}\right] = \sigma^2.$$

$$E[\hat{\sigma}^2] = \sigma^2 \implies \hat{\sigma}^2 \text{ es } \textbf{insesgado}$$

Diseño factorial con tres factores: Estimación de parámetros

$$\hat{\mu} = \bar{y}_{....}$$

$$\hat{\alpha}_i = \bar{y}_{i...} - \bar{y}_{....}, \quad \hat{\beta}_j = \bar{y}_{.j..} - \bar{y}_{....}, \quad \hat{\gamma}_k = \bar{y}_{..k.} - \bar{y}_{....}$$

$$\widehat{(\alpha\beta)}_{ij} = \bar{y}_{ij..} - \bar{y}_{i...} - \bar{y}_{.j..} + \bar{y}_{....}$$

$$\widehat{(\alpha\gamma)}_{ik} = \bar{y}_{i.k.} - \bar{y}_{i...} - \bar{y}_{..k.} + \bar{y}_{....}$$

$$\widehat{(\beta\gamma)}_{jk} = \bar{y}_{.jk.} - \bar{y}_{.j..} - \bar{y}_{..k.} + \bar{y}_{....}$$

$$\widehat{(\alpha\beta\gamma)}_{ijk} = \bar{y}_{ijk.} - \bar{y}_{ij..} - \bar{y}_{i.k.} - \bar{y}_{.jk.} + \bar{y}_{i...} + \bar{y}_{.j..} + \bar{y}_{..k.} - \bar{y}_{....}$$

Diseño factorial con tres factores: Desarrollo de las sumas de cuadrados

Suma Cuadrática por Factor

$$SC_A = \frac{1}{bcn}\sum_{i=1}^{a} y_{i\ldots}^2 - \frac{y_{\ldots\ldots}^2}{abcn}$$

$$SC_B = \frac{1}{acn}\sum_{j=1}^{b} y_{\cdot j\cdot\cdot}^2 - \frac{y_{\ldots\ldots}^2}{abcn}$$

$$SC_C = \frac{1}{abn}\sum_{k=1}^{c} y_{\cdot\cdot k\cdot}^2 - \frac{y_{\ldots\ldots}^2}{abcn}$$

Diseño factorial con tres factores: Desarrollo de las sumas de cuadrados

Suma Cuadrática por Intervención de segundo orden

$$SC_{AB} = \left(\frac{1}{cn}\sum_{i=1}^{a}\sum_{j=1}^{b} y_{ij\cdot\cdot}^2 - \frac{y_{\cdots\cdot}^2}{abcn}\right) - SC_A - SC_B$$

$$SC_{AC} = \left(\frac{1}{bn}\sum_{i=1}^{a}\sum_{k=1}^{c} y_{i\cdot k\cdot}^2 - \frac{y_{\cdots\cdot}^2}{abcn}\right) - SC_A - SC_C$$

$$SC_{BC} = \left(\frac{1}{an}\sum_{j=1}^{b}\sum_{k=1}^{c} y_{\cdot jk\cdot}^2 - \frac{y_{\cdots\cdot}^2}{abcn}\right) - SC_B - SC_C$$

Diseño factorial con tres factores: Desarrollo de las sumas de cuadrados

Suma Cuadrática por Intervención de tercer orden

$$SC_{ABC} = \left(\frac{1}{n}\sum_{i=1}^{a}\sum_{j=1}^{b}\sum_{k=1}^{c} y_{ijk\cdot}^2 - \frac{y_{\cdots\cdot}^2}{abcn}\right) - SC_A - SC_B - SC_C - SC_{AB} - SC_{AC} - SC_{BC}$$

Suma Cuadrática del Error

$$SC_E = SC_T - \left(\frac{1}{n}\sum_{i=1}^{a}\sum_{j=1}^{b}\sum_{k=1}^{c} y_{ijk\cdot}^2 - \frac{y_{\cdots\cdot}^2}{abcn}\right)$$

Suma Cuadrática Total

$$SC_T = \sum_{i=1}^{a}\sum_{j=1}^{b}\sum_{k=1}^{c}\sum_{l=1}^{n} y_{ijkl}^2 - \frac{y_{\cdots\cdot}^2}{abcn}$$

Diseño factorial con tres factores: Desarrollo de los Cuadrados Medios

Para el cálculo del Cuadrado Medio de la fuente que se requiera análizar se debe considerar el siguiente cálculo:

$$CM_{\text{fuente}} = \frac{SC_{\text{fuente}}}{G.L_{\text{fuente}}}$$

Diseño factorial con tres factores: Distribuciones y estadísticos

$$F_{calculado} = \frac{CM_{tratamiento}}{CM_{error}}$$

el cual sigue la siguiente distribución:

$$F_{calculado} \sim F_{(1-\phi),(y),(u)}$$

en dónde:
ϕ : corresponde al nivel de significancia.
y : Grados de libertad del caso a evaluar. (Si se desea evaluar un factor corresponder a sus grados de libertad, en cambio si desea evaluar la interacción corresponde a la multiplicación de los grados de libertad de los factores).
u : Grados de libertad del error.

Diseño factorial con tres factores: Formulación de hipótesis y tabla ANOVA

Por factor:

- Factor A:

$$H_0 : \alpha_1 = \alpha_2 = \cdots = \alpha_a = 0$$
$$H_1 : \text{Al menos un } \alpha_i \neq 0$$

- Factor B:

$$H_0 : \beta_1 = \beta_2 = \cdots = \beta_b = 0$$
$$H_1 : \text{Al menos un } \beta_j \neq 0$$

- Factor C:

$$H_0 : \gamma_1 = \gamma_2 = \cdots = \gamma_c = 0$$
$$H_1 : \text{Al menos un } \gamma_k \neq 0$$

Diseño factorial con tres factores: Formulación de hipótesis y tabla ANOVA

Interacciones entre 2 factores:

- Factores A y B:

$$H_0 : (\alpha\beta)_{ij} = 0 \quad \forall i, j$$
$$H_1 : \text{Existe al menos un } (\alpha\beta)_{ij} \neq 0$$

- Factores A y C:

$$H_0 : (\alpha\gamma)_{ik} = 0 \quad \forall i, k$$
$$H_1 : \text{Existe al menos un } (\alpha\gamma)_{ik} \neq 0$$

- Factores B y C:

$$H_0 : (\beta\gamma)_{jk} = 0 \quad \forall j, k$$
$$H_1 : \text{Existe al menos un } (\beta\gamma)_{jk} \neq 0$$

Diseño factorial con tres factores: Formulación de hipótesis y tabla ANOVA

Interacción entre tres factores:

$$H_0 : (\alpha\beta\gamma)_{ijk} = 0 \quad \forall i, j, k$$
$$H_1 : \text{Existe al menos un } (\alpha\beta\gamma)_{ijk} \neq 0$$

Fuente	**GL**	**(SC)**	**(CM)**	**F**
Factor A	$a-1$	SC_A	$\frac{SC_A}{a-1}$	$\frac{CM_A}{CM_E}$
Factor B	$b-1$	SC_B	$\frac{SC_B}{b-1}$	$\frac{CM_B}{CM_E}$
Factor C	$c-1$	SC_C	$\frac{SC_C}{c-1}$	$\frac{CM_C}{CM_E}$
Interacción AB	$(a-1)(b-1)$	SC_{AB}	$\frac{SC_{AB}}{(a-1)(b-1)}$	$\frac{CM_{AB}}{CM_E}$
Interacción AC	$(a-1)(c-1)$	SC_{AC}	$\frac{SC_{AC}}{(a-1)(c-1)}$	$\frac{CM_{AC}}{CM_E}$
Interacción BC	$(b-1)(c-1)$	SC_{BC}	$\frac{SC_{BC}}{(b-1)(c-1)}$	$\frac{CM_{BC}}{CM_E}$
Interacción ABC	$(a-1)(b-1)(c-1)$	SC_{ABC}	$\frac{SC_{ABC}}{(a-1)(b-1)(c-1)}$	$\frac{CM_{ABC}}{CM_E}$
Error	$abc(n-1)$	SC_E	$\frac{SC_E}{abc(n-1)}$	
Total	$abcn-1$	SC_T		

Cuadro 1 – Tabla ANOVA para un Diseño Factorial General con tres factores.

Región de Rechazo

Se rechaza H_0 bajo el criterio de la región crítica si el estadístico calculado es mayor al estadístico F a evaluar.

$$F_{calculado} > F_{(1-\phi),(y),(u)}$$

En caso de desear evaluar el valor-p, el cálculos es el siguiente

$$valor - p = 1 - P(F_{calculado} < F)$$

en dónde rechazamos si:

$$valor - p < \phi$$

Comparaciones Múltiples

El método de comparación múltiple de Duncan se aplica **solo después** de que la ANOVA indique diferencias significativas. El procedimiento esencial es:

$$S_{\bar{Y}.} = \sqrt{\frac{\mathrm{CME}}{n}}$$

Los rangos críticos de Duncan calculado corresponde a $S_{\bar{Y}.} \cdot q_{\phi(y,w)}$, dónde y representa y=[2,...,a] siendo a la mayor cantidad de niveles del factor y w siendo los grados de libertad del error.

Tratamiento Valores Faltantes

Para un diseño factorial de 3 factors A, B, C con efectos fijos, n réplicas por celda y un único valor faltante $y_{i_0 j_0 k_0 l_0}$, se utilizan las siguientes fórmulas:

$$Y_{VF} = \sum_{i=1}^{a} \sum_{j=1}^{b} \sum_{k=1}^{c} \sum_{l=1}^{n} y_{ijkl},$$

$$\alpha_{VF} = \sum_{j=1}^{b} \sum_{k=1}^{c} \sum_{\ell=1}^{n} y_{i_0 jkl}, \quad \beta_{VF} = \sum_{i=1}^{a} \sum_{k=1}^{c} \sum_{\ell=1}^{n} y_{ij_0 kl}, \quad \gamma_{VF} = \sum_{i=1}^{a} \sum_{j=1}^{b} \sum_{\ell=1}^{n} y_{ijk_0 l},$$

$$(\alpha\beta)_{VF} = \sum_{k=1}^{c} \sum_{\ell=1}^{n} y_{i_0 j_0 kl}, \quad (\alpha\gamma)_{VF} = \sum_{j=1}^{b} \sum_{\ell=1}^{n} y_{i_0 jk_0 l}, \quad (\beta\gamma)_{VF} = \sum_{i=1}^{a} \sum_{\ell=1}^{n} y_{ij_0 k_0 l}.$$

Tratamiento Valores Faltantes

La ecuación final para el valor faltante es:

$$(a-1)(b-1)(c-1)\, n\, y_{i,\hat{j},k,\ell} = Y_{VF} - a\,\alpha_{VF} - b\,\beta_{VF} - c\,\gamma_{VF} + ab\,(\alpha\beta)_{VF} + ac\,(\alpha\gamma)_{VF} + bc\,(\beta\gamma)_{VF}. \tag{1}$$

Por lo tanto, la fórmula de imputación general para n réplicas es:

$$\hat{y}_{i,j,k,\ell} = \frac{Y - a\,\alpha_{VF} - b\,\beta_{VF} - c\,\gamma_{VF} + ab\,(\alpha\beta)_{VF} + ac\,(\alpha\gamma)_{VF} + bc\,(\beta\gamma)_{VF}}{(a-1)(b-1)(c-1)\,n}. \tag{2}$$

Es importante considerar que al imputar un valor faltante, los grados de libertad se reducen a un valor menos, siendo el modelo ANOVA un análisis menos robusto en comparación a análisis de experimentos sin datos faltantes.

Ejemplo de aplicación: Crecimiento de cultivos de maravilla

La empresa .Agrícola del Maule”desea evaluar el crecimiento de un nuevo cultivo de *Maravillas*, midiendo su crecimiento mensual en milímetros como variable de respuesta. Para este estudio se han considerado tres factores experimentales:

- **Dosis de fertilizante**, con tres niveles: alta (A1), media (A2) y baja (A3).
- **Tipo de riego**, con dos niveles: goteo (B1) y aspersión (B2).
- **Cobertura de suelo**, con dos niveles: con mulch (C1) y sin mulch (C2).

Ejemplo de aplicación: Crecimiento de cultivos de maravilla

A continuación, se presentan los valores correspondientes a cada combinación de tratamientos para su análisis:

		A1						**A2**						**A3**					
		B1			**B2**			**B1**			**B2**			**B1**			**B2**		
C	**Rép**	1	2	3	1	2	3	1	2	3	1	2	3	1	2	3	1	2	3
C1	1	60	75	75	67	73	73	62	68	65	71	80	80	76	71	75	75	75	75
	2	86	70	70	67	68	68	76	65	65	72	80	80	70	68	73	75	75	77
C2	1	55	53	53	52	52	57	44	44	45	60	60	60	52	51	50	56	55	57
	2	55	55	55	52	54	54	48	48	45	67	67	65	52	48	54	59	50	55

Cuadro 2 – Datos del ejercicio factorial con 3 réplicas.

Ejemplo de aplicación: Identificación

- **Unidad Experimental:** Cultivo de Maravilla de una parcela.
- **Factores, niveles y observaciones:**

Símbolo	Factor	Niveles	Descripción de niveles
A	Dosis de fertilizante	A1, A2, A3	A1 = Alta, A2 = Media, A3 = Baja
B	Tipo de riego	B1, B2	B1 = Goteo, B2 = Aspersión
C	Cobertura de suelo	C1, C2	C1 = Con mulch, C2 = Sin mulch

Cuadro 3 – Factores y niveles del experimento.

- **Número de combinaciones de tratamiento:** $3 \times 2 \times 2 = 12$.
- **Réplicas por combinación:** $n = 6$ (independientes).
- **Observaciones totales**: $N = 12 \times 6 = 72$.

Ejemplo de aplicación: Análisis exploratorio breve

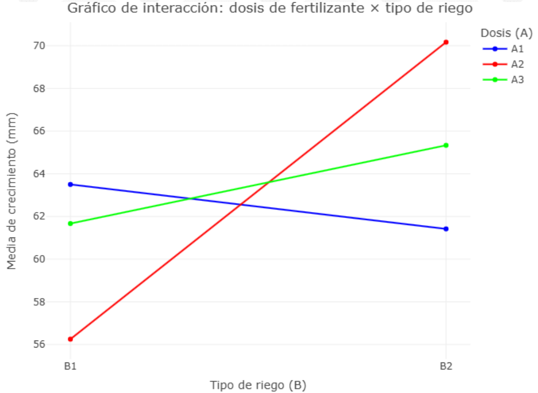

Figura 1 – Gráfico de interacción para dosis de fertilizante y tipo de riego.

- El efecto del fertilizante cambia según el riego (y cruce de promedios), señal de interacción entre ambos factores.

Ejemplo de aplicación: Análisis exploratorio breve

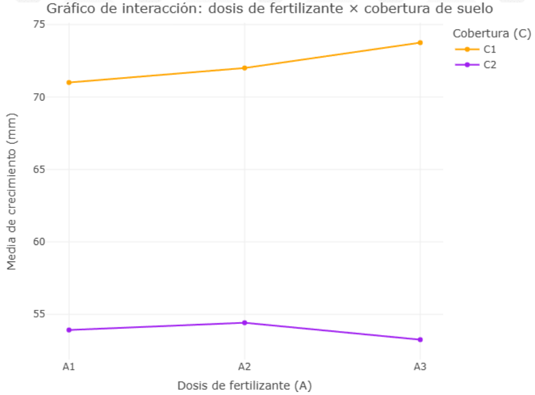

Figura 2 – Gráfico de interacción para dosis de fertilizante y cobertura de suelo.

- No se ve interacción entre dosis de fertilizante y cobertura de suelo, pero las coberturas existe posible diferencia entre coberturas de suelo en el crecimiento.

Ejemplo de aplicación: Análisis exploratorio breve

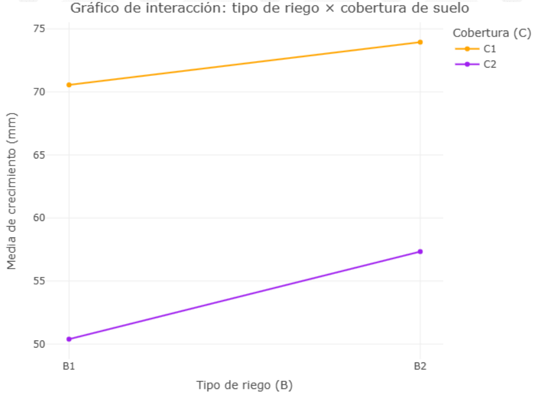

Figura 3 – Gráfico de interacción para tipo de riego y cobertura de suelo.

- No se aprecia interacción tipo de riego y cobertura de suelo, aunque las coberturas de suelo muestran diferencias en crecimiento.

Ejemplo de aplicación: Especificación de modelo a utilizar

Planteamos un modelo ANOVA de 3 factores.

$$y_{ijkl} = \mu + \alpha_i + \beta_j + \gamma_k + (\alpha\beta)_{ij} + (\alpha\gamma)_{ik} + (\beta\gamma)_{jk} + (\alpha\beta\gamma)_{ijk} + \varepsilon_{ijkl},$$

En donde,

- y_{ijkl} : Crecimiento en milímetros del l-ésimo cultivo de maravilla bajo la i-ésima dosis de fertilizante, con el j-ésimo tipo de riego, utilizando la k-ésima cobertura de suelo.
- μ: media global del crecimiento de maravillas.
- α_i: efecto de la i-ésima dosis de fertilizante.
- β_j: efecto del j-ésimo tipo de riego.
- γ_k: efecto de la k-ésima cobertura de suelo.
- $(\alpha\beta)_{ij}$: interacción dosis de fertilizante i, bajo el j-ésimo tipo de riego.
- $(\alpha\gamma)_{ik}$: interacción dosis de fertilizante i, bajo la k-ésima cobertura.
- $(\beta\gamma)_{jk}$: interacción riego j, bajo la k-ésima cobertura.
- $(\alpha\beta\gamma)_{ijk}$: interacción de la i-ésima dosis de fertilizante, bajo el tipo de riego j, con la cobertura de suelo k.
- ϵ_{ijkl} : Error experimental en milímetros del l-ésimo cultivo de maravilla bajo la i-ésima dosis de fertilizante, con el j-ésimo tipo de riego, utilizando la k-ésima cobertura de suelo.

Ejemplo de aplicación: Supuestos y condiciones de estimabilidad

$$\varepsilon_{ijkl} \sim N(0, \sigma^2) \text{ (i.i.d.)}$$

- Independencia (aleatorización)
- Normalidad en los residuos
- Varianza constante en los residuos σ^2

$$i \in \{1, 2, 3\}(A1, A2, A3),\ j \in \{1, 2\}(B1, B2),$$

$$k \in \{1, 2\}(C1, C2),\ l \in \{1, 2, 3, 4, 5, 6\}(\text{réplicas})$$

$$\sum_{i=1}^{a} \alpha_i = 0, \quad \sum_{j=1}^{b} \beta_j = 0, \quad \sum_{k=1}^{c} \gamma_k = 0,$$

$$\sum_{j=1}^{b} (\alpha\beta)_{ij} = 0\ \forall i, \quad \sum_{k=1}^{c} (\alpha\gamma)_{ik} = 0\ \forall i,$$

$$\sum_{k=1}^{c} (\beta\gamma)_{jk} = 0\ \forall j,$$

$$\sum_{i=1}^{a} \sum_{j=1}^{b} \sum_{k=1}^{c} (\alpha\beta\gamma)_{ijk} = 0.$$

Ejemplo de aplicación: Hipótesis a contrastar

Efecto	**Hipótesis nula H_0**	**Hipótesis alternativa H_1**
Dosis de fertilizante (A)	$\alpha_1 = \alpha_2 = \alpha_3 = 0$	$\exists\, i : \alpha_i \neq 0$
Tipo de riego (B)	$\beta_1 = \beta_2 = 0$	$\exists\, j : \beta_j \neq 0$
Cobertura de suelo (C)	$\gamma_1 = \gamma_2 = 0$	$\exists\, k : \gamma_k \neq 0$
Interacción $A \times B$	$(\alpha\beta)_{ij} = 0\ \forall i, j$	$\exists\, i, j : (\alpha\beta)_{ij} \neq 0$
Interacción $A \times C$	$(\alpha\gamma)_{ik} = 0\ \forall i, k$	$\exists\, i, k : (\alpha\gamma)_{ik} \neq 0$
Interacción $B \times C$	$(\beta\gamma)_{jk} = 0\ \forall j, k$	$\exists\, j, k : (\beta\gamma)_{jk} \neq 0$
Interacción triple $A \times B \times C$	$(\alpha\beta\gamma)_{ijk} = 0\ \forall i, j, k$	$\exists\, i, j, k : (\alpha\beta\gamma)_{ijk} \neq 0$

Ejemplo de aplicación: Desarrollo manual paso a paso

Consideramos la siguiente tabla que contiene la suma de observaciones por factor o combinación entre factores para facilitar los cálculos.

	A1		**A2**		**A3**		**Total**
C	**B1**	**B2**	**B1**	**B2**	**B1**	**B2**	
C1	436	416	401	463	433	452	**2 601**
C2	326	321	274	379	307	332	**1 939**
Total	**762**	**737**	**675**	**842**	**740**	**784**	**4 540**

Cuadro 4 – Suma de réplicas por combinación factorial $3 \times 2 \times 2$.

Ejemplo de aplicación: Desarrollo manual paso a paso

Es importante considerar los siguientes valores:

$a = 3, b = 2, c = 2, n = 6, N = abcn = 72$

$y_{....} = 4540$

$y^2_{....} = 20{,}611{,}600$

$\frac{y^2_{....}}{N} = 286{,}272, 22$

$$\sum_{i,j,k,l}^{N} y^2_{ijkl} = 294{,}612$$

Ejemplo de aplicación: Desarrollo manual paso a paso

$$SCT = 294{,}612 - 286{,}272,22 = 8{,}339,78$$

$$SCA = 286{,}286,0833 - 286{,}272,22 = 13,863$$

$$SCB = 286{,}752,72 - 286{,}272,22 = 480,5022$$

$$SCC = 292{,}358,944 - 286{,}272,22 = 6{,}086,7244$$

$$SCAB = 287{,}554,83 - 286{,}272,22 - 13,863 - 480,5022 = 788,25$$

$$SCAC = 292{,}413,6667 - 286{,}272,22 - 13,863 - 6{,}086,7244 = 40,86$$

$$SCBC = 292{,}896,33 - 286{,}272,22 - 480,5022 - 6{,}086,7244 = 56,89$$

$$SCABC = 293{,}770,33 - 286{,}272,22 - 13,863 - 480,5022 - 6{,}086,7244 - 788,25 - 40,86 - 56,89 = 31,02$$

$$SCE = 8{,}428,78 - 13,863 - 480,5022 - 6{,}086,7244 - 788,25 - 40,86 - 56,89 - 31,02 = 841,67$$

Ejemplo de aplicación: Tabla ANOVA completa

Por lo tanto, nos quedaría la siguiente tabla ANOVA.

Fuente de variación	**GL**	**SC**	**CM**	F	p
Factor A	2	13,863	6,932	0,494	0,613
Factor B	1	480,502	480,502	34,25	$2{,}16 \cdot 10^{-7*}$
Factor C	1	6 086,724	6 086,724	433,84	0^*
A × B	2	788,25	394,13	28,09	$2{,}45 \cdot 10^{-9*}$
A × C	2	40,86	20,43	1,46	0,241
B × C	1	56,89	56,89	4,05	$0{,}049^*$
A × B × C	2	31,02	15,51	1,11	0,338
Error	60	841,67	14,03	–	–
Total	71	8 339,78	—	—	—

Cuadro 5 – ANOVA del crecimiento de cultivos de maravillas. *Diferencias significativas a $\phi = 0{,}05$.

Ejemplo de aplicación: Regiones de rechazo

Además se comparan los estadísticos F calculados con los valores $F_{(0,95),(1),(60)} = 4,0012$ y $F_{(0,95),(2),(60)} = 3,1504$, obteniendo las siguientes comparativas:

- **Para factor A:** $F_{(0,95),(2),(60)} = 3,1504 > 0,494$ **No se rechaza** H_0
- **Para factor B:** $F_{(0,95),(1),(60)} = 4,0012 < 34,25$ **Se rechaza** H_0
- **Para factor C:** $F_{(0,95),(1),(60)} = 4,0012 < 433,84$ **Se rechaza** H_0
- **Para interacción AB:** $F_{(0,95),(2),(60)} = 3,1504 < 28,09$ **Se rechaza** H_0
- **Para interacción AC:** $F_{(0,95),(2),(60)} = 3,1504 > 1,46$ **No se rechaza** H_0
- **Para interacción BC:** $F_{(0,95),(1),(60)} = 4,0012 < 4,05$ **Se rechaza** H_0
- **Para interacción ABC:** $F_{(0,95),(2),(60)} = 3,1504 > 1,11$ **No se rechaza** H_0

Ejemplo de aplicación: Interpretación

A modo general, con un nivel de significación de $\phi = 0,05$, se concluye lo siguiente:

- Hay diferencias en el crecimiento promedio de los cultivos de maravilla por tipo de riego.
- Hay diferencias en el crecimiento promedio de los cultivos de maravilla por cobertura de suelo.
- Hay diferencias en el crecimiento promedio de los cultivos de maravilla por alguna interacción entre dosis de fertilizante y tipo de riego.
- Hay diferencias en el crecimiento promedio de los cultivos de maravilla por alguna interacción entre tipo de riego y cobertura de suelo.

Ejemplo de aplicación: Comparaciones múltiples (Duncan)

Factor B Tipos de riego (2 niveles)

Medias

$$\bar{y}_{B_1} = \frac{2177}{36} = 60{,}47, \quad \bar{y}_{B_2} = \frac{2363}{36} = 65{,}64.$$

Diferencia observada

$$D = |\bar{y}_{B_2} - \bar{y}_{B_1}| = 5{,}17.$$

Diferencia crítica Para $r = 2$, $\nu_e = 60$ y $n = 36$:

$$s_B = \sqrt{\frac{14{,}03}{36}} = 0{,}625, \quad d_2 = 2{,}828 \times 0{,}625 = 1{,}77.$$

Conclusión

$$5{,}17 > 1{,}77 \implies \bar{y}_{B_2} > \bar{y}_{B_1} \quad (\alpha = 0{,}05).$$

- Con un nivel de significación del 5 % el tipo de riego por aspersión es mayor a que por goteo, en cuanto a crecimiento promedio de los cultivos de maravilla.

Ejemplo de aplicación: Comparaciones múltiples (Duncan)

Factor C Cobertura de suelo (2 niveles)

Medias

$$\bar{y}_{C_1} = \frac{2601}{36} = 72{,}25, \quad \bar{y}_{C_2} = \frac{1939}{36} = 53{,}86.$$

Diferencia observada

$$D = 18{,}39.$$

Diferencia crítica Para $r = 2$, $\nu_e = 60$ y $n = 36$:

$$s_C = \sqrt{\frac{14{,}03}{36}} = 0{,}625, \quad d_2 = 2{,}828 \times 0{,}625 = 1{,}77.$$

Conclusión

$$18{,}39 > 1{,}77 \implies \bar{y}_{C_1} > \bar{y}_{C_2} \quad (\alpha = 0{,}05).$$

- Con un nivel de significación del 5 % la cobertura de suelo con mulch es mayor a sin mulch, en cuanto a crecimiento promedio de los cultivos de maravilla.

Ejemplo de aplicación: Comparaciones múltiples (Duncan)

Interacción dosis de fertilizante A $\times$ tipo de riego B (6 niveles)

- Disponible en el informe y realización computacional.
- En conclusión, la combinación con mayor crecimiento promedio en los cultivos de maravilla fue dosis de fertilizante media (A2) y tipo de riego por aspersión (B2).

Ejemplo de aplicación: Comparaciones múltiples (Duncan)

Interacción de tipos de riego B $\times$ coberturas de suelo C (4 niveles)
Medias

Tratamiento	B_1C_2	B_2C_2	B_1C_1	B_2C_1
Σ	907	1032	1270	1331
$\bar{y} = \Sigma/18$	50,39	57,33	70,56	73,94

(18 observaciones por célula)
Valores críticos Para $\nu_e = 60$ y $n = 18$:

$$s_{BC} = \sqrt{\frac{14{,}03}{18}} = 0{,}883,$$

r	$q_{r,60;0{,}05}$	$d_r = q_{r,60;0{,}05} \times 0{,}883$
2	2,828	2,50
3	2,980	2,63
4	3,080	2,72

Ejemplo de aplicación: Comparaciones múltiples (Duncan)

Interacción de tipos de riego B $\times$ coberturas de suelo C (4 niveles)
Comparaciones completas

$$
\begin{aligned}
r = 2: \quad & |\bar{y}_{B_2C_2} - \bar{y}_{B_1C_2}| = 6{,}94 > 2{,}50\ (*), \\
& |\bar{y}_{B_1C_1} - \bar{y}_{B_2C_2}| = 13{,}23 > 2{,}50\ (*), \\
& |\bar{y}_{B_2C_1} - \bar{y}_{B_1C_1}| = 3{,}38 > 2{,}50\ (*); \\
r = 3: \quad & |\bar{y}_{B_1C_1} - \bar{y}_{B_1C_2}| = 20{,}17 > 2{,}63\ (*), \\
& |\bar{y}_{B_2C_1} - \bar{y}_{B_2C_2}| = 16{,}61 > 2{,}63\ (*); \\
r = 4: \quad & |\bar{y}_{B_2C_1} - \bar{y}_{B_1C_2}| = 23{,}55 > 2{,}72\ (*).
\end{aligned}
$$

- La combinación que ofrece el mayor rendimiento promedio es *riego por aspersión (B2)* con *mulch (C1)*

Ejemplo de aplicación: Conclusiones e interpretaciones

Como recomendación para obtener los mejores resultados en el crecimiento de los cultivos de maravilla se recomienda el uso de,

- Tipo de riego por aspersión (B2) antes que riego por goteo (B1).
- Cobertura de suelo con mulch (C1) antes que sin mulch (C2).

Ahora, las mejores combinaciones para el crecimiento promedio de los cultivos de maravilla son,

- Tipo de riego por aspersión (B2) y cobertura de suelo con mulch (C1)
- Dosis de fertilizante media (A2) y tipo de riego por aspersión (B2)

Valores faltantes

		A1						A2						A3					
		B1			B2			B1			B2			B1			B2		
C	**Rép**	1	2	3	1	2	3	1	2	3	1	2	3	1	2	3	1	2	3
C1	1	60	75	75	67	73	73	62	68	65	71	80	80	76	71	75	75	75	75
	2	86	70	70	67	68	68	76	65	65	72	80	80	70	68	73	75	75	77
C2	1	55	53	53	52	52	57	44	44	45	60	60	60	52	51	50	56	55	57
	2	55	*	55	52	54	54	48	48	45	67	67	65	52	48	54	59	50	55

Cuadro 6 – Datos del ejercicio factorial con 3 réplicas con dato faltante.

Podemos notar el caso de un valor faltante en la observación $y_{\alpha_1 \beta_1 \gamma_2 l_5}$

Tratamiento de Valor Faltante

$$\Sigma_{A_1} = 1444, \quad \Sigma_{B_1} = 2122, \quad \Sigma_{C_2} = 1884$$
$$\Sigma_{A_1 B_1} = 707, \quad \Sigma_{B_1 C_2} = 852, \quad \Sigma_{A_1 C_2} = 592$$
$$\Sigma y = 4485$$

$$\hat{y}_{1,1,2,5} = \frac{4485 - 3 \cdot 1444 - 2 \cdot 2122 - 2 \cdot 1884 + 6 \cdot 707 + 6 \cdot 592 + 4 \cdot 852}{72} \approx 46{,}43$$

Conclusiones

- El diseño factorial general permite estimar de forma eficiente tanto los efectos principales como las interacciones de múltiples factores en un solo experimento.
- Al agrupar los factores en un mismo estudio, se reduce el número total de tratamientos y réplicas necesarias, lo que disminuye significativamente el tiempo y los costos de muestreo.
- Incluir varias fuentes de variación en el análisis incrementa la validez y el realismo de los resultados, acercando el experimento a las condiciones de la práctica real.

Referencias

Yates, F. (1937). *The design and analysis of factorial experiments* (Technical Communication No. 35). Imperial Bureau of Soil Science / Her Majesty's Stationery Office.

Montgomery, D. C. (2007). *Diseño y análisis de experimentos* (2.ª ed.). Limusa Wiley.

Figura 4 – Código QR: Repositorio GitHub.