Construcción de modelos a partir de principios generales

Cáculo de estimadores de cuadrados medios y denominadores de F

Universitat d'Alacant
Universidad de Alicante

Jose Jacobo Zubcoff

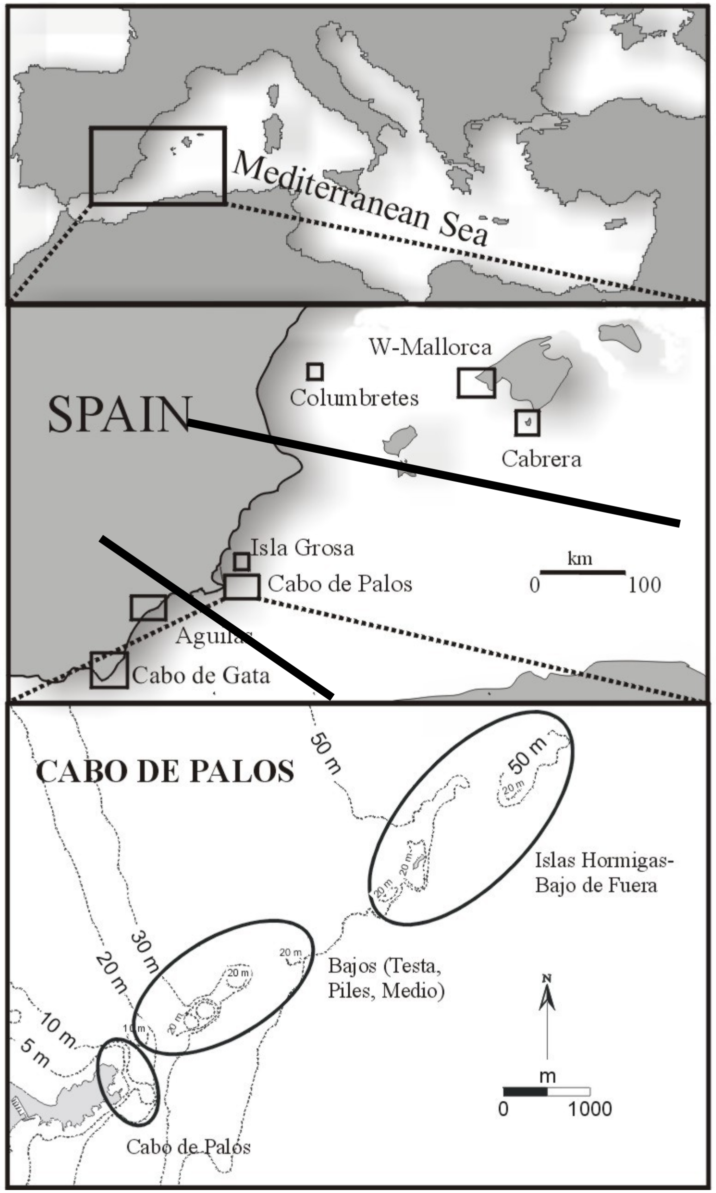
Departamento de Ciencias del Mar y Biología Aplicada

Construcción de modelos a partir de principios generales

1. Definir los modelos, hipótesis e hipótesis nulas.
2. Definir las poblaciones y variables a muestrear.
3. Determinar cuáles son los factores relevantes (a partir de las hipótesis).
4. Definir si los factores son fijos o aleatorios.
5. Definir el nivel de replicación, de modo que:
   (a) sean posibles las generalizaciones
   (b) se asegure que no hay pseudo-replicación
6. Revisar amenazas a la independencia. Organizar los controles.
7. Calcular los grados de libertad (gl), los estimadores de los Cuadrados Medios (CM) y los contrastes apropiados para el cálculo de las F.
8. Determinar la potencia del análisis y optimizar el muestreo.
9. Si es necesario, re-diseñar el experimento.

Diseños **mixtos**: parcialmente anidados

Ejemplo:



Diseños **mixtos**: parcialmente anidados

$$X_{ijkl} = \mu + P_i + R_j + P_i \times R_j \; + S_k \,(P_i \times R_j) + T_{ijklm}$$

Fuente de variación	**Tipo**	**Nº de niveles**
P	F	p = 2
R	F	r = 3
P x R		
S(P x R)	A	s = 3
Residual	A	n = 3
Total		

Cálculo de los grados de libertad (gl)

Ejemplo:

$$X_{ijkl} = \mu + P_i + R_j + P_i \times R_j + S_k\,(P_i \times R_j) + T_{ijkl}$$

Reglas:

• Cada factor ortogonal: nº de niveles – 1
• Cada factor anidado: producto de
→ nº de niveles del factor anidado – 1
→ nº de combinaciones de tratamientos en los cuales ese factor está anidado
• Cada interacción: producto del nº de gl de cada término de la interacción
• Residual: producto del nº de niveles de cada factor en el experimento por el nº de réplicas – 1
• Total: producto del nº de niveles de cada factor (incluyendo réplicas) – 1

Cálculo de los grados de libertad (gl)

Ejemplo:

$$X_{ijkl} = \mu + P_i + R_j + P_i \times R_j \; + S_k\,(P_i \times R_j) + T_{ijkl}$$

Fuente de variación	**Tipo**	**Nº de niveles**	**gl**	
P	F	p = 2		
R	F	r = 3		
P x R				
S(P x R)	A	s = 3		
Residual	A	n = 3		
Total				

Ejemplo:

$$X_{ijkl} = \mu + P_i + R_j + P_i \times R_j + S_k\,(P_i \times R_j) + T_{ijkl}$$

Fuente de variación	**Tipo**	**Nº de niveles**	**gl**	
P	F	p = 2	(p -1)	1
R	F	r = 3	(r -1)	2
P x R			(p -1)(r -1)	2
S(P x R)	A	s = 3	pr(s – 1)	12
Residual	A	n = 3	prs(n – 1)	36
Total			prsn – 1	53

Estimas de Cuadrados Medios (CM) y contraste para el cálculo de las F

Ejemplo:

$$X_{ijkl} = \mu + P_i + R_j + P_i \times R_j + S_k\,(P_i \times R_j) + T_{ijkl}$$

	f	f	a	a
Fuente de variación	***a***	***b***	***c***	***n***
P_a				
R_b				
$P \times R_{ab}$				
$S(P \times R)_{c(ab)}$				
$Residual_{n(c(ab))}$				

Estimas de Cuadrados Medios (CM) y contraste para el cálculo de las F

Ejemplo:

$$X_{ijkl} = \mu + P_i + R_j + P_i \times R_j + S_k (P_i \times R_j) + T_{ijkl}$$

Reglas:

1. Si el subíndice situado en la columna se encuentra en la fila y:

   (a) el subíndice es **FIJO**
   - **NO** incluye componentes **ANIDADOS** (no está entre paréntesis) ➔ 0
   - **SI** incluye componentes **ANIDADOS** (está entre paréntesis) ➔ 1

   (b) el subíndice es **ALEATORIO** ➔ 1

2. Si el subíndice situado en la columna no se encuentra en la fila:
   ➔ el subíndice (nº de niveles del factor)

Estimas de Cuadrados Medios (CM) y contraste para el cálculo de las F

Ejemplo:

$$X_{ijkl} = \mu + P_i + R_j + P_i \times R_j + S_k\,(P_i \times R_j) + T_{ijkl}$$

	f	f	a	a
Fuente de variación	***a***	***b***	***c***	***n***
P_a	0			
R_b	a			
$P \times R_{ab}$	0			
$S(P \times R)_{c(ab)}$	1			
$Residual_{n(c(ab))}$	1			

Estimas de Cuadrados Medios (CM) y contraste para el cálculo de las F

Ejemplo:

$$X_{ijkl} = \mu + P_i + R_j + P_i \times R_j + S_k (P_i \times R_j) + T_{ijkl}$$

	f	f	a	a
Fuente de variación	***a***	***b***	***c***	***n***
P_a	0	b		
R_b	a	0		
$P \times R_{ab}$	0	0		
$S(P \times R)_{c(ab)}$	1	1		
$Residual_{n(c(ab))}$	1	1		

Estimas de Cuadrados Medios (CM) y contraste para el cálculo de las F

Ejemplo:

$$X_{ijkl} = \mu + P_i + R_j + P_i \times R_j + S_k (P_i \times R_j) + T_{ijkl}$$

	f	f	a	a
Fuente de variación	***a***	***b***	***c***	***n***
P_a	0	b	c	
R_b	a	0	c	
$P \times R_{ab}$	0	0	c	
$S(P \times R)_{c(ab)}$	1	1	1	
$Residual_{n(c(ab))}$	1	1	1	

Estimas de Cuadrados Medios (CM) y contraste para el cálculo de las F

Ejemplo:

$$X_{ijkl} = \mu + P_i + R_j + P_i \times R_j + S_k (P_i \times R_j) + T_{ijkl}$$

	f	f	a	a
Fuente de variación	***a***	***b***	***c***	***n***
$\mathbf{P_a}$	0	b	c	n
$\mathbf{R_b}$	a	0	c	n
$\mathbf{P \times R_{ab}}$	0	0	c	n
$\mathbf{S(P \times R)_{c(ab)}}$	1	1	1	n
$\mathbf{Residual_{n(c(ab))}}$	1	1	1	1

Estimas de Cuadrados Medios (CM) y contraste para el cálculo de las F

Reglas:

1. Identificar en la fila objetivo todos los componentes de variación (σ^2_x) que pertenecen a ese CM. Estos componentes son los términos de las filas que contengan el subíndice o subíndices de la fila objetivo.

2. Cada componente identificado se ha de multiplicar por el producto de todas las entradas que aparecen en de la fila de ese componente, omitiendo las columnas que contienen el(los) subíndice(s) de la fila objetivo.

3. En la fila que contiene el residual el multiplicador es 1.

Estimas de Cuadrados Medios (CM) y contraste para el cálculo de las F

Ejemplo:

$$X_{ijkl} = \mu + P_i + R_j + P_i \times R_j + S_k (P_i \times R_j) + T_{ijkl}$$

Fuente de variación	*a*	*b*	*c*	*n*	Estimador de los CM
$\mathbf{P_a}$	0	b	c	n	$\sigma^2_e + n\sigma^2_{c(ab)} + bcn\sigma^2_a$
$\mathbf{R_b}$	a	0	c	n	
$\mathbf{P \times R_{ab}}$	0	0	c	n	
$\mathbf{S(P \times R)_{c(ab)}}$	1	1	1	n	
$\mathbf{Residual_{n(c(ab))}}$	1	1	1	1	

Estimas de Cuadrados Medios (CM) y contraste para el cálculo de las F

Ejemplo:

$$X_{ijkl} = \mu + P_i + R_j + P_i \times R_j + S_k\,(P_i \times R_j) + T_{ijkl}$$

Fuente de variación	*a*	*b*	*c*	*n*	Estimador de los CM
P_a	0	b	c	n	$\sigma^2_e + n\sigma^2_{c(ab)} + bcn\sigma^2_a$
R_b	a	0	c	n	$\sigma^2_e + n\sigma^2_{c(ab)} + acn\sigma^2_b$
$P \times R_{ab}$	0	0	c	n	
$S(P \times R)_{c(ab)}$	1	1	1	n	
$Residual_{n(c(ab))}$	1	1	1	1	

Estimas de Cuadrados Medios (CM) y contraste para el cálculo de las F

Ejemplo:

$$X_{ijkl} = \mu + P_i + R_j + P_i \times R_j + S_k (P_i \times R_j) + T_{ijkl}$$

Fuente de variación	***a***	***b***	***c***	***n***	**Estimador de los CM**
$\mathbf{P_a}$	0	b	c	n	$\sigma^2_e + n\sigma^2_{c(ab)} + bcn\sigma^2_a$
$\mathbf{R_b}$	a	0	c	n	$\sigma^2_e + n\sigma^2_{c(ab)} + acn\sigma^2_b$
$\mathbf{P \times R_{ab}}$	0	0	c	n	$\sigma^2_e + n\sigma^2_{c(ab)} + cn\sigma^2_{ab}$
$\mathbf{S(P \times R)_{c(ab)}}$	1	1	1	n	
$\mathbf{Residual_{n(c(ab))}}$	1	1	1	1	

Estimas de Cuadrados Medios (CM) y contraste para el cálculo de las F

Ejemplo:

$$X_{ijkl} = \mu + P_i + R_j + P_i \times R_j + S_k (P_i \times R_j) + T_{ijkl}$$

Fuente de variación	*a*	*b*	*c*	*n*	Estimador de los CM
P_a	0	b	c	n	$\sigma^2_e + n\sigma^2_{c(ab)} + bcn\sigma^2_a$
R_b	a	0	c	n	$\sigma^2_e + n\sigma^2_{c(ab)} + acn\sigma^2_b$
$P \times R_{ab}$	0	0	c	n	$\sigma^2_e + n\sigma^2_{c(ab)} + cn\sigma^2_{ab}$
$S(P \times R)_{c(ab)}$	1	1	1	n	$\sigma^2_e + n\sigma^2_{c(ab)}$
$Residual_{n(c(ab))}$	1	1	1	1	

Estimas de Cuadrados Medios (CM) y contraste para el cálculo de las F

Ejemplo:

$$X_{ijkl} = \mu + P_i + R_j + P_i \times R_j + S_k\,(P_i \times R_j) + T_{ijkl}$$

Fuente de variación	***a***	***b***	***c***	***n***	**Estimador de los CM**
$\mathbf{P_a}$	0	b	c	n	$\sigma^2_e + n\sigma^2_{c(ab)} + bcn\sigma^2_a$
$\mathbf{R_b}$	a	0	c	n	$\sigma^2_e + n\sigma^2_{c(ab)} + acn\sigma^2_b$
$\mathbf{P \times R_{ab}}$	0	0	c	n	$\sigma^2_e + n\sigma^2_{c(ab)} + cn\sigma^2_{ab}$
$\mathbf{S(P \times R)_{c(ab)}}$	1	1	1	n	$\sigma^2_e + n\sigma^2_{c(ab)}$
$\mathbf{Residual_{n(c(ab))}}$	1	1	1	1	σ^2_e

Estimas de Cuadrados Medios (CM) y contraste para el cálculo de las F

Ejemplo:

$$X_{ijkl} = \mu + P_i + R_j + P_i \times R_j + S_k (P_i \times R_j) + T_{ijkl}$$

Fuente de variación	*a*	*b*	*c*	*n*	Estimador de los CM	F-ratio versus
P_a	0	b	c	n	$\sigma^2_e + n\sigma^2_{c(ab)} + bcn\sigma^2_a$	
R_b	a	0	c	n	$\sigma^2_e + n\sigma^2_{c(ab)} + acn\sigma^2_b$	
$P \times R_{ab}$	0	0	c	n	$\sigma^2_e + n\sigma^2_{c(ab)} + cn\sigma^2_{ab}$	
$S(P \times R)_{c(ab)}$	1	1	1	n	$\sigma^2_e + n\sigma^2_{c(ab)}$	
$\text{Residual}_{n(c(ab))}$	1	1	1	1	σ^2_e	

Estimas de Cuadrados Medios (CM) y contraste para el cálculo de las F

Ejemplo:

$$X_{ijkl} = \mu + P_i + R_j + P_i \times R_j + S_k\,(P_i \times R_j) + T_{ijkl}$$

Fuente de variación	*a*	*b*	*c*	*n*	Estimador de los CM	F-ratio versus
P_a	0	b	c	n	$\sigma^2_e + n\sigma^2_{c(ab)} + bcn\sigma^2_a$	**$S(P \times R)_{c(ab)}$**
R_b	a	0	c	n	$\sigma^2_e + n\sigma^2_{c(ab)} + acn\sigma^2_b$	
$P \times R_{ab}$	0	0	c	n	$\sigma^2_e + n\sigma^2_{c(ab)} + cn\sigma^2_{ab}$	
$S(P \times R)_{c(ab)}$	1	1	1	n	$\sigma^2_e + n\sigma^2_{c(ab)}$	
$Residual_{n(c(ab))}$	1	1	1	1	σ^2_e	

Estimas de Cuadrados Medios (CM) y contraste para el cálculo de las F

Ejemplo:

$$X_{ijkl} = \mu + P_i + R_j + P_i \times R_j + S_k (P_i \times R_j) + T_{ijkl}$$

Fuente de variación	*a*	*b*	*c*	*n*	Estimador de los CM	F-ratio versus
P_a	0	b	c	n	$\sigma^2_e + n\sigma^2_{c(ab)} + bcn\sigma^2_a$	$S(P \times R)_{c(ab)}$
R_b	a	0	c	n	$\sigma^2_e + n\sigma^2_{c(ab)} + acn\sigma^2_b$	$S(P \times R)_{c(ab)}$
$P \times R_{ab}$	0	0	c	n	$\sigma^2_e + n\sigma^2_{c(ab)} + cn\sigma^2_{ab}$	
$S(P \times R)_{c(ab)}$	1	1	1	n	$\sigma^2_e + n\sigma^2_{c(ab)}$	
$\text{Residual}_{n(c(ab))}$	1	1	1	1	σ^2_e	

Estimas de Cuadrados Medios (CM) y contraste para el cálculo de las F

Ejemplo:

$$X_{ijkl} = \mu + P_i + R_j + P_i \times R_j + S_k (P_i \times R_j) + T_{ijkl}$$

Fuente de variación	*a*	*b*	*c*	*n*	Estimador de los CM	F-ratio versus
P_a	0	b	c	n	$\sigma^2_e + n\sigma^2_{c(ab)} + bcn\sigma^2_a$	$S(P \times R)_{c(ab)}$
R_b	a	0	c	n	$\sigma^2_e + n\sigma^2_{c(ab)} + acn\sigma^2_b$	$S(P \times R)_{c(ab)}$
$P \times R_{ab}$	0	0	c	n	$\sigma^2_e + n\sigma^2_{c(ab)} + cn\sigma^2_{ab}$	$S(P \times R)_{c(ab)}$
$S(P \times R)_{c(ab)}$	1	1	1	n	$\sigma^2_e + n\sigma^2_{c(ab)}$	
$Residual_{n(c(ab))}$	1	1	1	1	σ^2_e	

Estimas de Cuadrados Medios (CM) y contraste para el cálculo de las F

Ejemplo:

$$X_{ijkl} = \mu + P_i + R_j + P_i \times R_j + S_k\,(P_i \times R_j) + T_{ijkl}$$

Fuente de variación	*a*	*b*	*c*	*n*	Estimador de los CM	F-ratio versus
$\mathbf{P_a}$	0	b	c	n	$\sigma^2_e + n\sigma^2_{c(ab)} + bcn\sigma^2_a$	$\mathbf{S(P \times R)_{c(ab)}}$
$\mathbf{R_b}$	a	0	c	n	$\sigma^2_e + n\sigma^2_{c(ab)} + acn\sigma^2_b$	$\mathbf{S(P \times R)_{c(ab)}}$
$\mathbf{P \times R_{ab}}$	0	0	c	n	$\sigma^2_e + n\sigma^2_{c(ab)} + cn\sigma^2_{ab}$	$\mathbf{S(P \times R)_{c(ab)}}$
$\mathbf{S(P \times R)_{c(ab)}}$	1	1	1	n	$\sigma^2_e + n\sigma^2_{c(ab)}$	$\mathbf{Residual_{n(c(ab))}}$
$\mathbf{Residual_{n(c(ab))}}$	1	1	1	1	σ^2_e	

Y recuerda, no hay limite para el análisis…

El objetivo es tener el control del diseño experimental para optimizar el análisis