




Basic principles of experimental design

Principios básicos del diseño de experimentos

1



Introducción

- Los modelos de “Diseño de experimentos” son modelos estadísticos clásicos cuyo objetivo es **averiguar si factores determinados influyen en la variable de interés** y, si existe influencia de algún factor, cuantificarla.
- Ejemplos?

2

Variabilidad

- La metodología del diseño de experimentos se basa en la **experimentación**.
- Si se repite un experimento en condiciones indistinguibles, los resultados presentan *variabilidad* que puede ser grande o pequeña.

3

Error experimental

- Cuando la mayoría de las causas de variabilidad están muy controladas, el error experimental será pequeño y habrá poca variación en los resultados del experimento.

4

Objetivo del diseño experimental

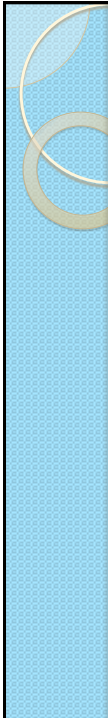
- Estudiar si utilizar un determinado tratamiento produce una mejora en el proceso o no.
- Para ello se debe experimentar utilizando el tratamiento y no utilizándolo.
- Si la variabilidad experimental es grande, sólo se detectará la influencia del uso del tratamiento cuando éste produzca grandes cambios en relación con el error de observación.

5

Un experimento se realiza por alguno de los siguientes motivos:

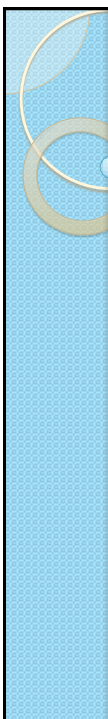
- Determinar las principales causas de variación en la respuesta
- Encontrar las condiciones experimentales con las que se consigue un valor extremo en la variable de interés o respuesta
- Comparar las respuestas a diferentes niveles de observación de variables controladas
- Obtener un modelo estadístico-matemático que permita hacer predicciones de respuestas futuras

6



- Para que la metodología de diseño de experimentos sea eficaz es fundamental que el experimento esté **bien diseñado**.

7



• **Planificación de un experimento**

8

Etapas de desarrollo de un problema de diseño experimental

1. Definir los objetivos del experimento
 - Preguntas concretas a las que debe dar respuesta el experimento
2. Identificar todas las posibles fuentes de variación, incluyendo:
 - Factores tratamiento (**cualitativos o cuantitativos**) y sus niveles,
 - Unidades experimentales,
 - Factores nuisance (molestos): factores bloque, factores ruido (variabilidad) y covariables.

9

Fuentes de variación de un experimento

Fuente	Tipo
Debida a las condiciones de interés (Factores tratamiento)	Planificada y sistemática
Debida al resto de condiciones controladas (Factores "nuisance")	Planificada y sistemática
Debida a condiciones no controladas (error de medida, material experimental, ...)	No planificada, pero ¿sistemática?

10

Etapas de desarrollo de un problema de diseño experimental

3. Elegir una regla de asignación de las unidades experimentales a las condiciones de estudio (tratamientos):
 - Diseño factorial o no
 - Anidamiento
 - Asignación al azar en determinados niveles de observación
 - Orden de asignación, etc.

11

Etapas de desarrollo de un problema de diseño experimental

4. Especificar las medidas con que se trabajará (la respuesta), el procedimiento experimental y anticiparse a las posibles dificultades.

12

Etapas de desarrollo de un problema de diseño experimental

5. Ejecutar un experimento piloto
6. Especificar el modelo
 - Modelo de efectos fijos: todos los factores son factores de efectos fijos.
 - Modelo de efectos aleatorios: todos los factores son factores de efectos aleatorios.
 - Modelo mixto: hay factores de efectos fijos y factores de efectos aleatorios

13

Etapas de desarrollo de un problema de diseño experimental

7. Esquematizar los pasos del análisis
 - Depende de:
 - Los objetivos indicados en el paso 1
 - El diseño seleccionado en el paso 3
 - El modelo asociado que se especificó en el paso 5
 - Deben incluir:
 - Estimaciones que hay que calcular
 - Contrastes a realizar
 - Intervalos de confianza que se calcularán
 - Diagnóstico y crítica del grado de ajuste del modelo a la realidad

14

Etapas de desarrollo de un problema de diseño experimental

8. Determinar el tamaño de muestras
 - Número de observaciones que se deben tomar para alcanzar los objetivos del experimento
9. Revisar las decisiones anteriores. Modificarlas si se considera necesario

15

Es fundamental tener en cuenta que:

- Ningún método de análisis estadístico, por sofisticado que sea, permite extraer conclusiones correctas en un diseño de experimentos mal planificado
 - Recíprocamente, el análisis estadístico es una etapa más que está completamente integrado en el proceso de planificación

16

Es fundamental tener en cuenta que:

- El análisis estadístico no es un segundo paso independiente de la tarea de planificación.
- Es necesario comprender la totalidad de objetivos propuestos antes de comenzar con el análisis.
- Si no se hace así, tratar que el experimento responda a otras cuestiones a posteriori puede ser (lo será casi siempre) imposible

17

Es fundamental tener en cuenta que:

- Toda persona implicada en la ejecución del experimento y en la recolección de los datos debe ser informada con precisión de la estrategia experimental diseñada

18



Principios básicos en el diseño de experimentos

19



Tres principios básicos

- Aleatorización
- Bloqueo
- La factorización del diseño
 - **Aleatorizar y bloquear** son estrategias eficientes para asignar los tratamientos a las unidades experimentales **sin preocuparse de que tratamientos considerar**
 - Por el contrario, la **factorización** del diseño define una estrategia eficiente **para elegir los tratamientos sin considerar en absoluto como asignarlos** después a las unidades experimentales.

20

Aleatorizar

- Ventajas:
 - Transforma la variabilidad sistemática no planificada en variabilidad no planificada o ruido aleatorio.
 - Evita la dependencia entre observaciones al aleatorizar los instantes de recogida muestral
 - Valida muchos de los procedimientos estadísticos más comunes

21

Bloquear

- Dividir o particionar las unidades experimentales en grupos llamados *bloques* de modo que las observaciones realizadas en cada bloque se realicen bajo condiciones experimentales lo más parecidas posibles.
 - A diferencia de lo que ocurre con los factores tratamiento, el experimentador no está interesado en investigar las posibles diferencias de la respuesta entre los niveles de los factores bloque

22

Bloquear

- La ventaja de bloquear un factor que se supone que tienen una clara influencia en la respuesta pero en el que no se está interesado, es de convertir la variabilidad sistemática no planificada en variabilidad sistemática planificada

23

La factorización del diseño

- Un diseño factorial es una estrategia experimental que consiste en cruzar los niveles de todos los factores tratamiento en todas las combinaciones posibles
- Ventajas:
 - Permite detectar la existencia de **efectos interacción** entre los diferentes factores tratamiento.
 - Es una estrategia más eficiente que la estrategia clásica de examinar la influencia de un factor manteniendo constantes el resto de los factores

24



Algunos diseños experimentales clásicos

25

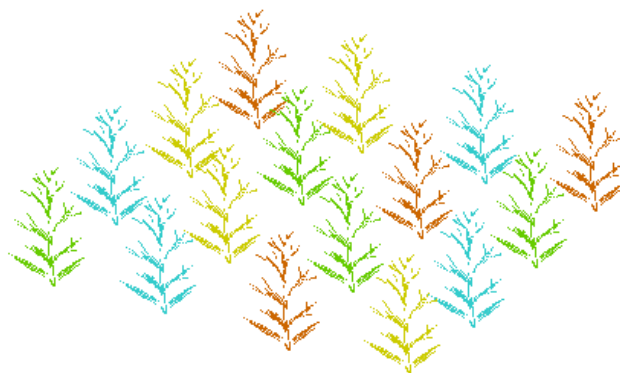


Diseño completamente aleatorizado

- El experimentador asigna las unidades experimentales a los tratamientos al azar.
- La única restricción es el número de observaciones que se toman en cada tratamiento.

26

Diseño completamente aleatorizado



A1 B1 C1 A2
 D1 A3 D2 C2
 B2 D3 C3 B3
 C4 A4 B4 D4

<http://www.tfrec.wsu.edu/>

27

Diseño completamente aleatorizado

ANOVA table format:

Source of variation	Degrees of freedom ^a	Sums of squares (SSQ)	Mean square (MS)	F
Treatments (<i>Tr</i>)	$t-1$	SSQ_{Tr}	$SSQ_{Tr}/(t-1)$	MS_{Tr}/MS_E
Error (<i>E</i>)	$t*(r-1)$	SSQ_E	$SSQ_E/(t*(r-1))$	
Total (<i>Tot</i>)	$t*r-1$	SSQ_{Tot}		

^awhere t =number of treatments and r =number of replications per treatment.

Sample ANOVA table:

Source of variation	Degrees of freedom	Sums of squares (SSQ)	Mean square (MS)	F
Treatments	3	80.92	26.97	9.68 ^a
Error	12	33.43	2.79	
Total	15	114.35		

^aF test with 3,12 degrees of freedom at $P=0.05$ is 3.49

28

Diseño en bloques o con un factor bloque

- El experimentador agrupa las unidades experimentales en bloques, a continuación determina la distribución de los tratamientos en cada bloque y, por último, asigna al azar las unidades experimentales a los tratamientos dentro de cada bloque.
- En el análisis estadístico de un diseño en bloques, éstos se tratan como los niveles de un único factor de bloqueo, aunque en realidad puedan venir definidos por la combinación de niveles de más de un factor nuisance.

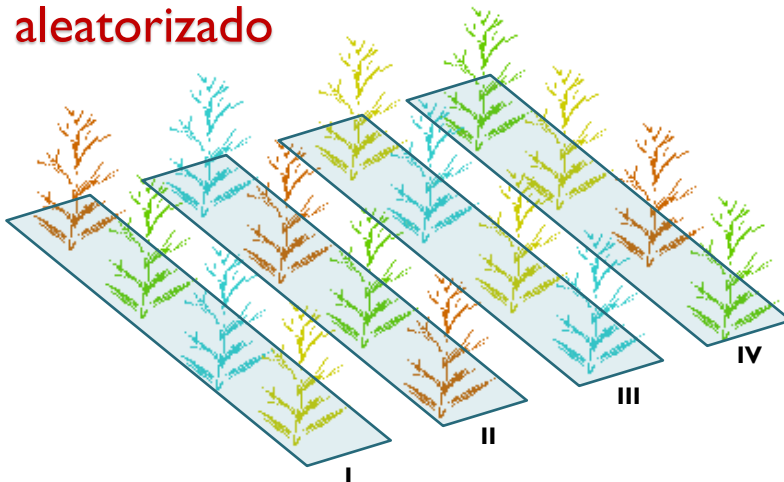
29

Diseño en bloques completamente aleatorizado

- Diseño en bloques completos con una única observación por cada tratamiento

30

Diseño en bloques completamente aleatorizado



Block I A1 B1 C1 A2
 Block II D1 A3 D2 C2
 Block III B2 D3 C3 B3
 Block IV C4 A4 B4 D4

31

Diseño en bloques completamente aleatorizado

ANOVA table format:

Source of variation	Degrees of freedom ^a	Sums of squares (SSQ)	Mean square (MS)	F
Blocks (<i>B</i>)	<i>b</i> -1	SSQ_B	$SSQ_B/(b-1)$	MS_B/MS_E
Treatments (<i>T</i>)	<i>t</i> -1	SSQ_T	$SSQ_T/(t-1)$	MS_T/MS_E
Error (<i>E</i>)	$(t-1)*(b-1)$	SSQ_E	$SSQ_E/((t-1)*(b-1))$	
Total (<i>Tot</i>)	$t*b-1$	SSQ_{Tot}		

^awhere *t*=number of treatments and *b*=number of blocks or replications.

Sample ANOVA table:

Source of variation	Degrees of freedom	Sums of squares (SSQ)	Mean square (MS)	F
Blocks	3	26.36	8.79	1.36 ^a
Treatments	3	216.23	72.08	11.19 ^a
Error	9	57.95	6.44	
Total	15	300.54		

^aF test with 3,9 degrees of freedom at $P=0.05$ is 3.86

32

Diseños con dos o más factores bloque

- En ocasiones hay dos (o más) fuentes de variación lo suficientemente importantes como para ser designadas factores de bloqueo. En tal caso, ambos factores bloque pueden ser **cruzados o anidados**.
- Los factores bloque están **cruzados** cuando existen unidades experimentales en todas las combinaciones posibles de los niveles de los factores bloques.

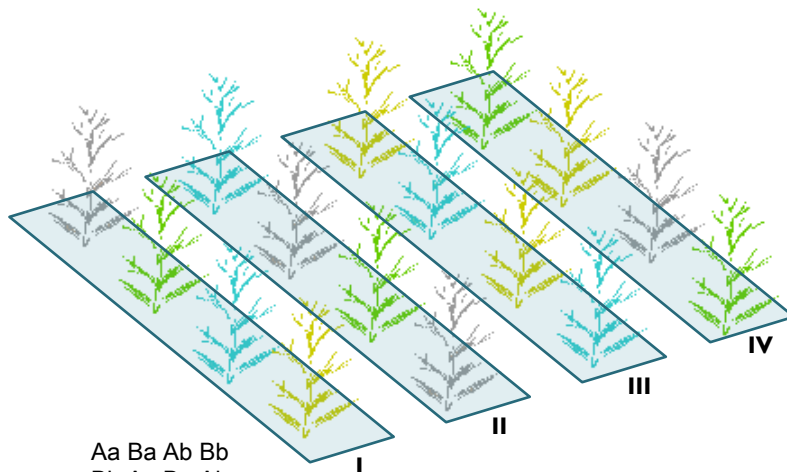
33

Diseños con dos o más factores bloque

- **Diseño con factores bloque cruzados.** También denominado diseño fila-columna, se caracteriza porque existen unidades experimentales en todas las celdas (intersecciones de fila y columna).
- Los factores bloque están **anidados** si cada nivel particular de uno de los factores bloque ocurre en un único nivel del otro factor bloque.
- **Diseño con factores bloque anidados o jerarquizados.** Dos factores bloque se dicen anidados cuando observaciones pertenecientes a dos niveles distintos de un factor bloque están automáticamente en dos niveles distintos del segundo factor bloque.

34

Diseños con dos o más factores bloque



Block IV
Block III
Block II
Block I

Aa Ba Ab Bb
Bb Aa Ba Ab
Ba Bb Ab Aa
Ab Aa Ba Bb

35

Diseños con dos o más factores bloque

ANOVA table format:

Source of variation	Degrees of freedom*	Sums of squares (SSQ)	Mean square (MS)	F
Blocks (B)	b-1	SSQ _B	SSQ _B /(b-1)	MS _B /MS _E
First factor (F1)	f-1	SSQ _{F1}	SSQ _{F1} /(f-1)	MS _{F1} /MS _E
Second factor (F2)	s-1	SSQ _{F2}	SSQ _{F2} /(s-1)	MS _{F2} /MS _E
First X Second (F1x2)	(f-1)*(s-1)	SSQ _{F1x2}	SSQ _{F1x2} /((f-1)*(s-1))	MS _{F1x2} /MS _E
Error (E)	(f*s-1)*(b-1)	SSQ _E	SSQ _E /((f*s-1)*(b-1))	
Total (Tot)	f*s*b-1	SSQ _{Tot}		

*where f=number of treatments in the first factor, s=number of treatments in the second factor and b=number of blocks or replications

Sample ANOVA table:

Source of variation	Degrees of freedom	Sums of squares (SSQ)	Mean square (MS)	F
Blocks	3	17.92	5.97	8.35 ^a
First	1	109.75	109.75	153.40 ^b
Second	1	49.64	49.64	69.38 ^b
First X Second	1	0.24	0.24	0.33 ^b
Error	9	6.44	0.72	
Total	15	183.99		

^aF test with 3,9 degrees of freedom at P=0.05 is 3.86

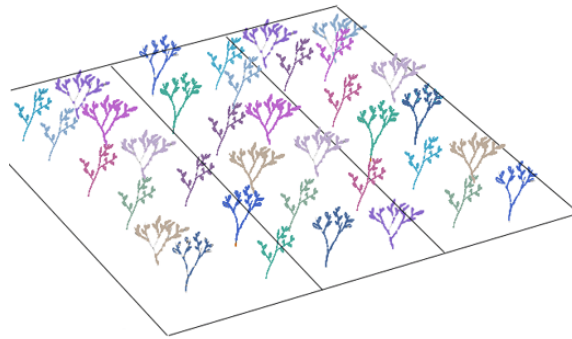
^bF test with 1,9 degrees of freedom at P=0.05 is 5.12

Diseños con dos o más factores

- En este modelo es importante estudiar la posible interacción entre dos (o más) factores. Si en cada casilla se tiene una única observación no es posible estudiar la interacción entre los dos factores, para hacerlo hay que replicar el modelo

37

Diseños con dos o más factores



Block I
C1a B2a
B1a A2a
A2b A1a
B1b B2b
A1b C2a
C2b C1b

Block II
C2a C1a
C1b B1a
B2b A2a
A1b A1a
B1b A2b
C2b B2a

Block III
B2a B1a
B2b A2a
A2b A1a
C1b C2b
C1a A1b
B1b C2a

38

Diseños con dos o más factores

Sample ANOVA table:

Source of variation	Degrees of freedom	Sums of squares (SSQ)	Mean square (MS)	F
Blocks	2	146.38	73.19	117.43 ^a
First	2	41.97	20.98	33.67 ^a
Second	1	195.70	195.70	313.99 ^b
Third	1	66.84	66.84	107.23 ^b
First X Second	2	0.04	0.02	0.03 ^a
First X Third	2	0.23	0.12	0.18 ^a
Second X Third	1	6.28	6.28	10.08 ^b
First X Second X Third	2	0.69	0.35	0.55 ^a
Error	22	13.71	0.62	
Total	35	471.85		

^aF test with 2,22 degrees of freedom at P=0.05 is 3.44
^bF test with 1,22 degrees of freedom at P=0.05 is 4.30

39

Diseños factoriales a dos niveles

- Son diseños en los que se trabaja con k factores, todos ellos con dos niveles (se suelen denotar + y -)
- Estos diseños son adecuados para tratar problemas donde hay muchos factores que pueden influir en la variable de interés descritos
- Estos diseños permiten trabajar con un número elevado de factores y son válidos para estrategias secuenciales

40

Referencias

- Juan Vilar
<http://dm.udc.es/asignaturas/estadistica2/>

41

Diseño en bloques completos

- Cada tratamiento se observa el mismo número de veces en cada bloque

42

Diseño en bloques incompletos

- Cuando el tamaño del bloque es inferior al número de tratamientos no es posible observar la totalidad de tratamientos en cada bloque

43

- En el análisis estadístico de datos históricos se pueden cometer diferentes errores, los más comunes son los siguientes:
 - Inconsistencia de los datos
 - Variables con fuerte correlación
 - El rango de las variables controladas es limitado

45