



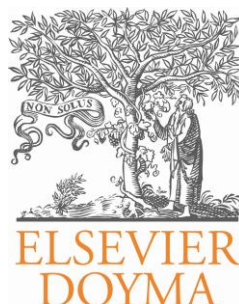
Tema 13:

# Diseños que afectan a la varianza

Clúster, análisis del cambio y cross-over

Laura Riba, Hector Rufino y Marta Vilaró  
Angel Ruiz, Erik Cobo y José Antonio González

2014



## Diseños que afectan a la varianza

Presentación .....	3
1. Asignación en grupo ( <i>cluster</i> ).....	4
1.1. Variación entre grupos .....	4
1.2. Unidad de inferencia .....	5
1.3. Correlación intraclase .....	6
1.4. Efecto del diseño .....	7
1.5. Riesgos de sesgo .....	8
1.6. Extensión de la guía consort .....	8
2. Análisis del cambio .....	8
2.1. Correlación entre Z e Y .....	10
2.2. Varianza de la variable definida como respuesta .....	10
2.3. Ancova .....	11
2.4. Diseño con determinación repetida de la respuesta .....	12
3. Diseños con intercambio (“ <i>Cross-over</i> ”) .....	13
3.1. Ventajas e inconvenientes .....	14
3.2. Diseño 2X2 .....	14
3.2.1. Inferencia para el efecto directo del tratamiento .....	15
3.2.2. Inferencia para el efecto período .....	17
3.2.3. Inferencia para el efecto residual del tratamiento anterior ( <i>carry-over</i> ). .....	18
3.3. Tamaño muestral.....	20
Soluciones a los ejercicios .....	22

### Presentación

En este tema abordamos 3 diseños específicos en los que el cálculo de la incertidumbre asociada al proceso aleatorio necesita refinarse.

En ocasiones la intervención no puede asignarse de forma individual, o bien porque se administra de forma comunitaria, global (como podría ser una intervención sobre los abastecimientos de, por ejemplo, agua) o bien porque el profesional que la administra sólo puede ser experto en una de las intervenciones en comparación. Así, la asignación en grupo a una misma intervención (*cluster*) quiebra la independencia entre unidades —de la misma forma que lo haría una contaminación del efecto, como podría ser el caso de intervenciones que afectan el nivel de transmisión de enfermedades. En ambos casos, esta correlación entre los resultados de las unidades de un mismo grupo resulta en una oscilación de los estimadores mayor que la proporcionada por el análisis clásico, que debe ser corregido para que los IC reflejen la precisión real.

Por otro lado, el análisis del cambio, ya introducido en el tema 12, y el diseño con intercambio del tratamiento o *cross-over* pretenden aumentar la precisión de la estimación, por lo que un correcto análisis conduce a IC más estrechos, que conviene aprovechar.

**Contribuciones:** Basado en transparencias de Erik Cobo y José Antonio González; Laura Riba y Hector Rufino prepararon una primera versión de análisis del cambio y crossover; y Marta Vilaró de clúster; que han sido revisadas por Ángel Ruiz y Erik Cobo.

# 1. Asignación en grupo (*cluster*)

## 1.1. Variación entre grupos



### Definición

En un EC aleatorizado en grupo, los pacientes de un mismo grupo se asignan a la misma intervención.

El diseño habitual considera al paciente como la única unidad (estadística) con variabilidad: dispone de una muestra de pacientes y desea estimar el efecto en una población objetivo que engloba a todos los que comparten aquellos criterios de elegibilidad. Y quizás, en el análisis o en el diseño, controla como influyen en esta variabilidad ciertas características, como el centro.

**Ejemplo 1.1:** Un estudio multi-céntrico aleatorizado por bloques persigue comparar los 2 tratamientos “dentro” de cada centro. Y combina los resultados en un único estimador del efecto bajo la premisa de que el efecto es el mismo en todos los centros. Luego, el análisis de subgrupos suele ‘ojea’ esta proposición. Si se sostiene que el efecto no varía entre centros, aplicarlo o “transportarlo” a otros es un reto más cualitativo que cuantitativo.

En cambio, si (1) todos los pacientes de un mismo centro se asignan a la misma intervención; y (2) el centro es variable, en el sentido de que los resultados de los pacientes pueden diferir por centros; entonces la información que aporta cada paciente no es independiente de los otros pacientes del mismo centro, lo que impide usar la metodología habitual.



### Recuerde

Un EC randomizado en grupo implica más de una unidad aleatoria

**Nota:** Otros términos similares serían: *cluster-randomized trials*, *group-randomized trials*, *multi-level trials*, *hierarquical trials*, o ensayos comunitarios.

**Ejemplo 1.2:** En una situación extrema e irreal, la variable importante sería el centro y dentro de un mismo centro todos los pacientes responderían igual. Así, la no independencia de los pacientes de un mismo centro sería extrema: conocida la respuesta de un paciente, sabríamos la de los otros de su centro.

**Ejemplo 1.3:** La otra situación extrema podría ser más real. Ahora, todos los centros serían similares entre sí, pero dentro de un mismo centro todos los pacientes responderían de forma muy distinta. Así, conocer la respuesta de un paciente no aporta ninguna información sobre los otros de su centro.

## 1.2. Unidad de inferencia

Igual que las hojas penden de las ramas, estas unidades guardan relación jerárquica entre sí.

**Ejemplo 1.4:** Sea un diseño con 3 unidades estadísticas con variabilidad: paciente, profesional y centro: los pacientes son la unidad jerárquica inferior, anidados en el profesional que, a su vez, está anidado en el centro, unidad jerárquica superior.



### Recuerde

Un diseño en grupo (clúster) contempla tantas fuentes de variación como niveles.  
En general, serán sólo 2: centro y paciente.

**Nota:** El individuo que está anidado en el grupo  $g$  anidado al tratamiento  $t$ , será un individuo que pertenece al grupo  $g$  al que se le ha asignado el tratamiento  $t$ .

Un diseño en grupo desea estimar el efecto de la intervención en las unidades inferiores, pero la intervención se asigna a una unidad de nivel superior.

**Ejemplo 1.5:** Un estudio de prevención del embarazo adolescente aleatoriza y administra la intervención a nivel escolar, pero los datos se recogen a nivel individual y se desea valorar el efecto a nivel de las micro-unidades (“inferiores”), en este caso, cada joven.



### Recuerde

Un ensayo en grupo (clúster) asigna las intervenciones al grupo, pero desea estimar el efecto en los individuos.

**Contra- ejemplo 1.6:** El estudio ‘[2on Opinion Trial](#)’ se designó para bajar las tasas de cesáreas. El objetivo de la intervención es disminuir la tasa hospitalaria de cesárea valorada a nivel del hospital. Eso implica que puede tratarse como un estudio habitual, con la particularidad de que su unidad es el centro.

En el diseño en clúster no hay independencia entre las unidades. Los miembros de un grupo pueden compartir características comunes o tener un nivel de exposición similar o interactuar entre ellos ('contaminación'). Todo ello hace más similares 2 casos del mismo grupo que 2 casos de grupos diferentes.



**Recuerde**

Un diseño *cluster* tiene grupos homogéneos internamente y heterogéneos entre sí.

**1.3. Correlación intraclase**

La correlación Intraclase (ICC= *Intraclass Correlation Coefficient*) cuantifica el nivel de relación entre las unidades de un mismo grupo.

El apartado 5.1 del capítulo 8 de Intervalos de confianza definió el coeficiente de correlación intraclase para 2 grupos, que ahora extendemos a k grupos.



**Recuerde**

Un diseño en *cluster* con 2 niveles contempla 2 variabilidades: la que diferencia los individuos (variabilidad intra-grupo,  $\sigma_I^2$ ); y la que diferencia los grupos (variabilidad entra-grupos,  $\sigma_E^2$ ).



**Definición**

CCI es la proporción de variabilidad compartida:  $CCI = \frac{\sigma_E^2}{\sigma_E^2 + \sigma_I^2}$ .



**Ejercicio 1.1**

Se desea estudiar si la comida de catering influye en la salud bucodental de los niños de entre 4 y 5 años. Se incluyen en el estudio un total de 10 colegios, a cinco de los cuales se les asigna aleatoriamente ser servidos para la empresa A o B. Al cabo de seis meses se observa una escala de salud bucodental en cada alumno. Calcule el CCI si la variabilidad entre centros es 1.7 y la entre casos 6.3. Interprete.

**Nota:** Entre las simplificaciones (o premisas) implícitas, cabe resaltar que (1) dentro de cada grupo todos los pacientes tienen la misma variabilidad; (2) en todos los grupos, los pacientes

tienen la misma variabilidad (la variabilidad de los pacientes es independiente del grupo); y (3) todos los grupos aportan la misma variabilidad.

### 1.4. Efecto del diseño

Como los pacientes de un mismo centro comparten cierta variabilidad, cada uno de ellos no aporta información “completa”, ya que parte esta repetida. Por ello, la variabilidad del estimador del efecto es mayor de lo que sería en un diseño en el que las observaciones fueran independientes. El efecto del diseño ( $DE = DEFF = Design Effect$ ) cuantifica esta pérdida de información valorando cuánto aumenta la variabilidad del estimador ( $VIF = Variance Inflation Factor$ ).



#### Fórmula

$$DE = 1 + (n - 1)CCI$$



#### Ejercicio 1.2

Calcule el efecto del diseño para el ejercicio anterior asumiendo que todos los colegios tienen 21 niños por clase. Interprete.

**Nota:** Si los grupos difieren mucho en tamaño, los resultados por centro pueden tener diferente significado y una inferencia común pierde sentido. Si difieren algo, puede estimarse el DE tomando  $n$  por el promedio del tamaño de los centros.

Una vez conoce el DE, los cálculos habituales se obtienen multiplicando por el DE a la varianza obtenida como si fuera un diseño simple. Si no se tuviera en cuenta el DE, se subestimaría la incertidumbre del estudio, resultando en inferencia demasiado optimista: los  $IC_{95\%}$  contendrían el parámetro con una confianza menor del 95%; y el riesgo alfa (declarar eficaz una intervención cuando no lo es) sería mayor del 5% deseado.



#### Fórmula

$$N = \frac{n \cdot k}{DE}$$

Dónde  $N$  es el tamaño muestral necesario de un estudio aleatorio simple para que tenga la misma potencia y precisión que un ensayo de asignación en grupo,  $n$  es el tamaño de los grupos y  $k$  es el número de grupos

**Ejemplo 1.7:** Se ha diseñado un estudio de asignación en grupo que ha incluido 23 pacientes en cada uno de los 8 centros que han participado; en total, 184. Se ha estimado un efecto del diseño de 1.34. El número de participantes que podría haber incluido un estudio de asignación aleatoria simple para obtener la misma potencia y precisión hubiera sido un total de 138 pacientes.



### Ejercicio 1.3

Se ha diseñado un estudio con asignación aleatoria simple y se ha obtenido que debe incluir 130 pacientes para obtener una potencia del 80%. Manteniendo la misma potencia en el estudio, se desea plantear un diseño de asignación en grupo. Suponiendo que habrá un total de 9 clusters y que el efecto del diseño será de 2.6, calcule el número de participantes a incluir en cada cluster.

## 1.5. Riesgos de sesgo

Los diseños aleatorizados en grupo tienen riesgos específicos de sesgo originados, en general, por la dificultad de enmascarar las intervenciones. Conocer que tal centro ha sido asignado a tal intervención puede provocar un sesgo de selección, tanto de pacientes como de profesionales: por ejemplo, ciertos pacientes pueden o cambiar de centro o dejar de acudir. Asimismo, el hecho de conocer la intervención asignada al paciente puede originar un sesgo de actuación (*performance*): por ejemplo, algunos de los profesionales pueden aumentar su uso de otras intervenciones para compensar un hipotético menor rendimiento de la intervención asignada. También podría provocarse una pérdida de seguimiento diferente entre los grupos, provocando un sesgo por atrición. Finalmente, la evaluación de los resultados podría alterarse por el conocimiento de la intervención asignada.

## 1.6. Extensión de la guía consort

CONSORT publicó una extensión para los estudios de asignación en grupo que aconseja reportar los resultados diferenciando a nivel individual y de grupos. Repásela y haga el test correspondiente.

## 2. Análisis del cambio

El análisis del cambio es posible cuando medimos la respuesta, cuantitativa, al inicio (Z: basal o pre-tratamiento) y al final (Y: post-tratamiento). En el capítulo 12 se vio como calcular el tamaño muestral para la variable diferencia o cambio entre ambas determinaciones.



**Recuerde**

El análisis del cambio se basa en la variable diferencia entre la determinación final y basal de la variable respuesta.

Vimos que este ajuste pretende bajar la varianza, y, con ello, aumentar la potencia y la precisión.

**Nota:** El curso de observacionales explica un segundo beneficio (la intención de compensar desequilibrios para disminuir sesgos potenciales) y sus retos.

**Recuerde**

El cambio o diferencia final-basal se define como:

$$D_{ij} = Y_{ij} - Z_{ij}$$

Y su varianza es:

$$\text{var } D_{ij} = V Y_{ij} - Z_{ij} = 2 \cdot \sigma_{\epsilon}^2$$

**Nota:** Sean  $V Y_{ij} = \sigma_{\alpha}^2 + \sigma_{\epsilon}^2$ ,  $V Z_{ij} = \sigma_{\alpha}^2 + \sigma_{\epsilon'}^2$  y  $\text{cov } Y_{ij}, Z_{ij} = \sigma_{\alpha}^2$

$$V D_{ij} = V Y_{ij} + V Z_{ij} - 2\text{cov } Y_{ij}, Z_{ij} = \sigma_{\alpha}^2 + \sigma_{\epsilon}^2 + \sigma_{\alpha}^2 + \sigma_{\epsilon'}^2 - 2 \cdot \sigma_{\alpha}^2 = \sigma_{\epsilon}^2 + \sigma_{\epsilon'}^2 = 2 \cdot \sigma_{\epsilon}^2$$

En el análisis del cambio, la variabilidad entre individuos desaparece y la intra-individuo se dobla.

**Ejemplo 2.1:** En la población objetivo y con el método de determinación escogido, ambas variabilidades de la presión arterial sistólica (PAS) son  $\sigma_{\alpha}^2 = 36$  y  $\sigma_{\epsilon}^2 = 12$  mmHg.

Si se considera como variable la determinación de una única medida después del tratamiento, la varianza de la respuesta sería:

$$\text{var } Y_{ij} = \sigma_{\alpha}^2 + \sigma_{\epsilon}^2 = 36 + 12 = 48$$

En cambio, si se considera como variable final la diferencia entre la respuesta final y la respuesta antes del tratamiento, se obtiene la siguiente variabilidad:

$$\text{var}(D_{ij}) = 2 \cdot \sigma_{\epsilon}^2 = 2 \cdot 12 = 24$$



**Recuerde**

En el caso que se base el análisis en la respuesta final (Y) se obtiene la siguiente varianza:

$$var(Y_{ij}) = \sigma_{\alpha}^2 + \sigma_{\varepsilon}^2$$

**Nota:** Las fórmulas anteriores descansan en las siguientes simplificaciones (premisas): (1) Independencia entre los casos (individuos); (2) Homocedasticidad entre respuesta basal y final; y (3) Homocedasticidad entre tratamientos.

**2.1. Correlación entre Z e Y**



**Recuerde**

Si asumimos que  $\varepsilon'_{ij}$ ,  $\varepsilon_{ij}$  y  $\alpha_i$  son independientes mutuamente,

$$\rho = \frac{\sigma_{\alpha}}{\sigma_Y \cdot \sigma_Z}, \text{ pero también: } \rho = \frac{\sigma_{\alpha}^2}{\sigma_{\alpha}^2 + \sigma_{\varepsilon}^2}$$

Observar la correlación entre antes y después con la ayuda del coeficiente de correlación intraclass permite ver la correlación como la proporción de la variabilidad entre casos respecto a la total.

**2.2. Varianza de la variable definida como respuesta**

Dado que un procedimiento de medida razonable requiere, por lo menos, que las diferencias observadas se correspondan más a diferencias reales entre las unidades que a oscilaciones en la unidad observada, cabe esperar que  $\sigma_{\alpha}^2 > \sigma_{\varepsilon}^2$  y, por tanto,

$$\sigma_{\alpha}^2 + \sigma_{\varepsilon}^2 > 2 \cdot \sigma_{\varepsilon}^2 \quad \rightarrow \quad var(Y_{ij}) > var(D_{ij})$$

En resumen, si el proceso de medida es razonable y no hay fuentes adicionales de variación (los pacientes permanecen estables, excepto efectos constantes como el del tratamiento), entonces la correlación  $\rho > 0.5$  y la variabilidad de la variable cambio ( $D = Y - Z$ ) será menor que la de la respuesta Y. Por tanto, será más eficiente el análisis basado en el cambio.



**Ejercicio 2.1**

La glucosa en sangre se descompone en dos variables ( $Y = Y_{\alpha} + Y_{\varepsilon}$ ) que representan dos fuentes de variabilidad independientes, entre e intra-pacientes. Sus desviaciones típicas son  $\sigma_{\alpha} = 16$  y  $\sigma_{\varepsilon} = 9$ .

- (1) Se define como variable principal de respuesta los valores de Y tras 1 mes de tratamiento, ¿qué varianza tiene la variable respuesta?
- (2) Se define como variable principal de respuesta la diferencia entre la medida final y basal, ¿qué varianza tiene la variable principal?
- (3) ¿Cuál es la correlación entre la variable basal y la variable final?
- (4) A la vista de los resultados, ¿qué planteamiento le parece más oportuno realizar?

Se ha visto que este diseño elimina la varianza entre pacientes pero duplica la intra pacientes —ya que se toman dos medidas por cada individuo. Veamos otros diseños para disminuir la varianza de la respuesta.

**2.3. Ancova**

**Nota:** El modelo lineal de regresión del curso observacionales estudia con detalle el Ancova. Aquí comentamos sus implicaciones en el diseño de EC.

El ANCOVA es más eficiente que el análisis de la variable final y que el análisis del cambio.

El efecto del tratamiento se estima por lo general como:

$$\tau_b = Y_T - Y_C - b \cdot (Z_T - Z_C)$$

Dónde:

- $Z_C$  y  $Y_C$  son las medias en el grupo control de la respuesta basal y la final, respectivamente.
- $Z_T$  y  $Y_T$  ídem para el grupo tratamiento.

Y dependiendo del valor de b, se obtiene:

- Si  $b=0$ , el análisis de la variable final:  $\tau_0 = Y_T - Y_C$
- Si  $b=1$ , el análisis del cambio:  $\tau_1 = Y_T - Y_C - Z_T - Z_C$
- Si  $b=\beta$ , el ANCOVA:  $\tau_\beta = Y_T - Y_C - \beta (Z_T - Z_C)$ , donde  $\beta$  es el coef. de regresión.

Las varianzas en cada caso son:

Variable final: $\tau_0$	Cambio post – basal: $\tau_1$	ANCOVA: $\tau_\beta$
$\sigma_y^2 = \sigma_\alpha^2 + \sigma_\epsilon^2$	$\sigma_c^2 = 2 \cdot (1 - \rho) \cdot \sigma_y^2 = 2 \cdot \sigma_\epsilon^2$	$\sigma_{Anc}^2 = (1 - \rho^2) \cdot \sigma_y^2$

Esto implica que el estimador  $\tau \beta$  es más eficiente para cualquier valor de  $\rho$  y que  $\tau 1$  es más eficiente que  $\tau 0$  para  $\rho > 0,5$ , pero menos eficiente en la situación opuesta, lo que se resume en el siguiente gráfico debido a [Stephen Senn](#).

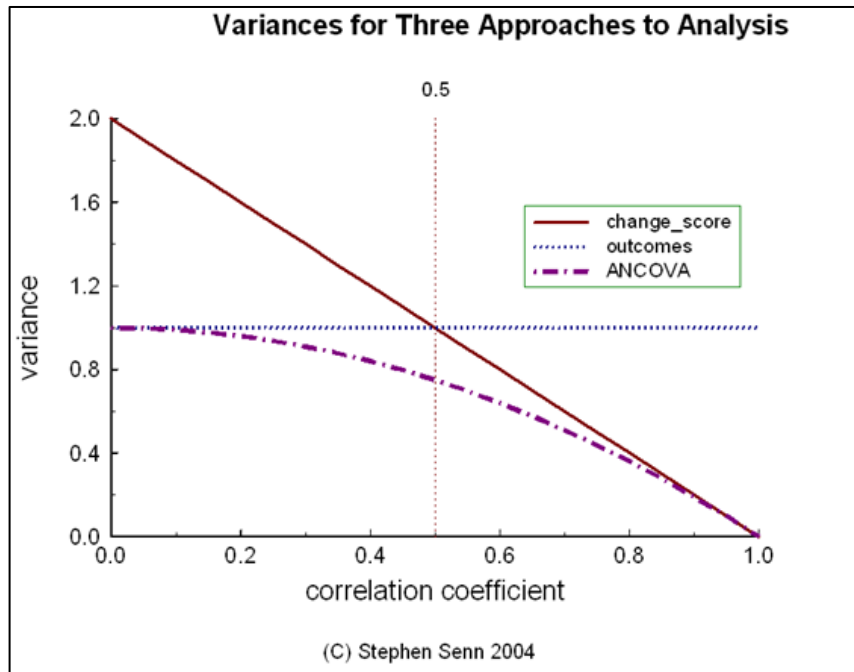


Figura 1. Variancias según el enfoque del análisis

**Ejemplo 2.2:** Siguiendo el Ejemplo 2.1, la correlación entre Z e Y es:

$$\rho = \frac{\sigma_{\alpha}^2}{\sigma_{\alpha}^2 + \sigma_{\epsilon}^2} = \frac{36}{36 + 12} = 0.75$$

Por lo tanto, si se opta por el diseño ANCOVA, la varianza de la respuesta es:

$$\sigma_{Anc}^2 = 1 - \rho^2 \cdot \sigma_y^2 = 1 - 0.75^2 \cdot 36 + 12 = 21$$

Que corresponde a la menor varianza de la respuesta en los tres diseños y conduce, por tanto, al diseño más eficiente.

#### 2.4. Diseño con determinación repetida de la respuesta

Cambio y ANCOVA pretenden reducir la varianza de la variable respuesta “atacando” a la varianza entre-pacientes,  $\sigma_{\alpha}^2$ . La intra-pacientes  $\sigma_{\epsilon}^2$  se puede controlar mediante el uso de medidas repetidas.



#### Definición

Si se obtienen  $K$  medidas repetidas independientes de la respuesta Y, una posible definición de la respuesta es la media de estas determinaciones:

$$m Y = \frac{1}{K} \sum_{k=1}^K Y_k$$

La varianza de la respuesta en este caso es:

$$\text{var } m Y = \sigma_{\alpha}^2 + \sigma_{\varepsilon}^2 / K$$

**Ejemplo 2.3:** Siguiendo el Ejemplo 2.1, si se opta por la realización de 7 medidas repetidas, la varianza de la nueva variable respuesta es:

$$\text{var } m_7 Y = \sigma_{\alpha}^2 + \sigma_{\varepsilon}^2 / 7 = 36 + 12 / 7 = 37.71$$



### Ejercicio 2.2

Con los mismos datos que en el Ejercicio 2.1, conteste:

- (1) Usando el ANCOVA (sin repeticiones), ¿qué varianza tiene la respuesta?
- (2) Si se utiliza como respuesta el promedio de la PAS repetida 5 días, ¿qué varianza tiene la respuesta?
- (3) A la vista de los nuevos resultados, ¿qué análisis usaría?

**Nota:** Estos diseños se pueden combinar, de manera que los beneficios en la reducción de la varianza de la respuesta se acumulan. Por ejemplo:

- Análisis del cambio post - basal, pero donde la respuesta basal y la final son promedios.
- ANCOVA con promedio de k medidas repetidas.

## 3. Diseños con intercambio (“*Cross-over*”)



### Definición

En un *cross-over* cada caso recibe varios tratamientos en diferente orden (o lugar) y se obtiene una respuesta bajo cada intervención.

Se requiere más de una secuencia para hacer un “*cross-over*”, es decir, se requiere más de un grupo de pacientes, cada uno con un orden (o secuencia) de administración de las intervenciones.

Los casos se asignan a secuencias de tratamientos.

**Ejemplo 3.1:** Para evaluar el efecto que dos preparados de aspirina tienen en el sangrado gástrico (evaluado por un método radioactivo), dos grupos de 8 pacientes han recibido ambos preparados en diferentes secuencias (AB o BA).

### 3.1. Ventajas e inconvenientes

Entre las ventajas de los *cross-over* se encuentran que: (1) permite observar la respuesta a A y B en los mismos sujetos y (2) es muy intuitivo para controlar el efecto individuo.

Por el contrario, mediante un buen diseño: (1) debe balancearse el período (o el lugar de administración) para controlar diferentes posibles respuestas en diferentes periodos (lugares); y (2) debe garantizarse que, al administrar un tratamiento, no permanece ningún efecto tardío (residual, arrastrado o *carry-over*) del tratamiento anterior.



#### Recuerde

*Carry-over*: persistencia de un efecto más allá de su periodo.

### 3.2. Diseño 2X2



#### Definición

El modelo que se utiliza para este tipo de estudios es:

$$Y_{itjk} = \mu + \tau_t + \pi_j + O_k + \alpha_i + \epsilon_{ij}$$

Dónde:

$\mu$  es la media general

$\tau_t$  es el efecto, fijo, inmediato o directo del tratamiento  $t=A,B$

$\pi_j$  es el efecto, fijo, del periodo

$O_k$  es el efecto, fijo, residual o tardío del tratamiento previo  $k=A,B$

$\alpha_i$  es el efecto, aleatorio, del individuo o caso  $i=1,2, \dots, n+m \rightarrow \alpha_i \sim N(0, \sigma_\alpha^2)$

$\epsilon_{ij}$  es el efecto, aleatorio, del individuo  $i$  en el periodo  $j \rightarrow \epsilon_{ij} \sim N(0, \sigma_\epsilon^2)$

**Notas:** (1) Cabe esperar que  $\sigma_\alpha^2 > \sigma_\epsilon^2$ . (Véase el punto 2.2). (2) En el primer período no hay efecto residual, al no arrastrar el efecto del tratamiento previo, dado que este no existe. (3) El efecto residual emplea el subíndice  $k$  del orden.

**Recuerde**

De existir ‘carry-over’, las diferencias en el segundo período contendrán también los efectos arrastrados de las intervenciones del primer período.

A continuación se detallan los diferentes contrastes para los efectos expresados en el modelo.

**3.2.1. Inferencia para el efecto directo del tratamiento**

La pregunta que se quiere contestar es: ¿tienen el mismo efecto los dos tratamientos?

Por tanto, se quiere realizar el siguiente contraste de hipótesis:

$$\begin{aligned} H_0: \tau_A &= \tau_B \\ H_1: \tau_A &\neq \tau_B \end{aligned}$$

Si se define  $d_i$  como la diferencia entre el primer período y el segundo:

$$\text{En el grupo 1, de orden AB: } d_{i1} = Y_{iA1k} - Y_{iB2k}$$

$$\text{En el grupo 2, de orden BA: } d_{i2} = Y_{iB1k} - Y_{iA2k}$$

Entonces, el contraste de hipótesis se puede reescribir como:

$$\begin{aligned} H_0: d_{i1} &= d_{i2} \\ H_1: d_{i1} &\neq d_{i2} \end{aligned}$$

**Nota:** Las esperanzas de las diferencias son:

$$\begin{aligned} E d_{i1} &= E Y_{iA1k} - Y_{iB2k} = E Y_{iA1k} - E Y_{iB2k} = \mu + \tau_A + \pi_1 - \mu + \tau_B + \pi_2 + O_A \\ &= \tau_A - \tau_B + \pi_1 - \pi_2 - O_A \\ E d_{i2} &= E Y_{iB1k} - Y_{iA2k} = E Y_{iB1k} - E Y_{iA2k} = \mu + \tau_B + \pi_1 - \mu + \tau_A + \pi_2 + O_B \\ &= \tau_B - \tau_A + \pi_1 - \pi_2 - O_B \end{aligned}$$

Por lo tanto,  $E(d_{i1} - d_{i2}) = E(d_{i1}) - E(d_{i2}) = 2(\tau_A - \tau_B) - (O_A - O_B)$

**Nota:** Las premisas para poder realizar este contraste son: efecto residual nulo o igual, es decir,  $O_A = O_B$  o ambos nulos; e independencia entre ambos grupos o secuencias.

Como cada caso aporta información sobre ambos tratamientos, la esperanza de la diferencia de las  $d_i$  dobla la diferencia de los efectos directos, para estimar el efecto  $\tau_A - \tau_B$ , los resultados se dividen por 2.

Nótese, como se ve en la demostración, y considerando los efectos residuales nulos:

$$(d_{i1} - d_{i2}) = 2(\tau_A - \tau_B)$$



**Recuerde**

El efecto directo se estudia con una prueba t de Student de datos independientes sobre la variable diferencia del primer menos el segundo periodo.

**Nota:** La estimación de esta prueba debe dividirse por 2.



**Ejercicio 3.1**

Diga si las siguientes afirmaciones son ciertas o no.

Para estimar el efecto (directo) del tratamiento T respecto al control C:

- (1) Calculamos la diferencia entre la respuesta a T y a C
- (2) Calculamos la diferencia de la respuesta entre primer y segundo periodo
- (3) Comparamos las medias de dichas diferencias entre ambos periodos
- (4) Comparamos las medias de dichas diferencias entre ambos tratamientos
- (5) Comparamos las medias de dichas diferencias entre ambas secuencias
- (6) Para conocer el IC95% del efecto, dividimos por 2 los resultados
- (7) Para conocer el p valor, dividimos por 2 los resultados

**Ejemplo 3.2:** Estudio cross-over del efecto de T frente a P en la enuresis en 24 adolescentes. La tabla muestra el número de noches –sobre un total de catorce- sin enuresis en los dos grupos TP y PT de 12 casos.

Grupo 1		Grupo 2	
Trat.1: T	Trat.2: P	Trat.1: P	Trat.2: T
8	5	12	11
14	10	6	8
8	0	13	9
9	7	8	8
11	6	8	9
3	5	4	8
6	0	8	14
10	6	2	4
13	12	8	13
10	2	9	7
7	5	7	10
13	13	7	6



A continuación se resuelve, con ayuda de R y de la función *t.test*, el contraste que permitirá saber si el efecto de los dos tratamientos es el mismo o no.



### Ejemplo de R

```
#Con los datos de la tabla del Ejemplo 3.2
> datos <- read.table(url("http://www-
eio.upc.es/teaching/best/datos-
ejemplos/enuresis.txt"),header=TRUE)
> d1 = datos$T_G1 - datos$P_G1
> d2 = datos$P_G2 - datos$T_G2
> t.test(d1, -d2, var.equal=TRUE)

                Two Sample t-test

data:  d1 and -d2
t = 3.7802, df = 22, p-value = 0.001029
...
95 percent confidence interval:
 2.106471 7.226863
...
```

Y dividimos por 2 el intervalo de confianza obtenido directamente del t-test, para estimar el efecto directo de  $\tau_A - \tau_B$

$$IC_{0.95} \tau_A - \tau_B = \frac{1}{2} IC_{0.95} d_{i1} - d_{i2} = \frac{1}{2} 2.11, 7.23 = 1.05, 3.61$$

Así, cambiar el tratamiento P por el T bajará el número de noches con enuresis cada 2 semanas entre 1 y 3,6, con una confianza del 95%: valores positivos que permiten establecer que T tiene efecto (relativo a C). También, como  $P=0.001$ , se concluye que existen diferencias significativas entre los dos tratamientos.

### 3.2.2. Inferencia para el efecto período

Para contestar si la respuesta Y cambia según el período en el que se recibe el tratamiento se hace el contraste de hipótesis:

$$H_0: \pi_1 = \pi_2$$

$$H_1: \pi_1 \neq \pi_2$$

Que se puede resolver mediante las diferencias  $d_i$  del anterior apartado (también dividiendo por 2 al final), cambiando el signo a los resultados de una secuencia:

$$H_0: E(d_{i1}) = E(-d_{i2})$$

$$H_1: E(d_{i1}) \neq E(-d_{i2})$$

**Ejemplo 3.3:** Siguiendo con el ejemplo anterior, podemos estudiar si el período en el que reciben el tratamiento influye en los valores de la variable respuesta.



**Ejemplo de R**

```
#Con los datos del ejemplo 3.2
> t.test(d1, d2, var.equal=TRUE)

Two Sample t-test

data: d1 and d2
t = 1.7551, df = 22, p-value = 0.09316
...
95 percent confidence interval:
 -0.3935293  4.7268627
...
```

$$IC_{0.95} \pi_A - \pi_B = \frac{1}{2} IC_{0.95} d_{i1} + d_{i2} = \frac{1}{2} -0.39, 4.73 = -0.195, 2.365$$

Por lo que no hay diferencias significativas entre periodos.

**3.2.3. Inferencia para el efecto residual del tratamiento anterior (carry-over)**

Estudiamos si son iguales los efectos tardíos mediante el contraste de hipótesis:

$$H_0: O_A = O_B$$

$$H_1: O_A \neq O_B$$

Si se define  $s_i$  como la suma de los resultados del primer período y del segundo:

En el grupo 1, de orden AB:  $s_{i1} = Y_{iA1k} + Y_{iB2k}$

En el grupo 2, de orden BA:  $s_{i2} = Y_{iB1k} + Y_{iA2k}$

Entonces, el contraste de hipótesis se puede reescribir como:

$$H_0: s_{i1} = s_{i2}$$

$$H_1: s_{i1} \neq s_{i2}$$

**Demostración:**

Las esperanzas de las diferencias son:

$$E s_{i1} = E Y_{iA1k} + Y_{iB2k} = E Y_{iA1k} + E Y_{iB2k} = \mu + \tau_A + \pi_1 + \mu + \tau_B + \pi_2 + O_A$$

$$= 2 \cdot \mu + \tau_A + \tau_B + \pi_1 + \pi_2 + O_A$$

$$E s_{i2} = E Y_{iB1k} + Y_{iA2k} = E Y_{iB1k} + E Y_{iA2k} = \mu + \tau_B + \pi_1 + \mu + \tau_A + \pi_2 + O_B = 2 \cdot \mu + \tau_B + \tau_A + \pi_1 + \pi_2 + O_B$$

Por lo tanto,

$$E(s_{i1} - s_{i2}) = E s_{i1} - E s_{i2} = O_A - O_B$$

Nótese que en este caso no hay que dividir entre 2 para estimar el efecto residual.

**Ejemplo 3.4:** Con los mismos datos que en el ejemplo anterior, nos puede interesar contrastar si los dos efectos tardíos son iguales o no.



**Ejemplo de R**

```
#Con los datos del ejemplo 3.2
> attach(datos)
> s1 <- datos$frmco_G1 + datos$pcbo_G1
> s2 = datos$pcbo_G2 + datos$frmco_G2
> t.test(s1, s2, var.equal=TRUE)
      Two Sample t-test
data:  s1 and s2
t = -0.5481, df = 22, p-value = 0.5891
...
95 percent confidence interval:
 -6.378402  3.711735
...
> detach(datos)
```

Como el intervalo de confianza incluye el 0, la ausencia de efecto tardío es compatible con los datos observados.

$$IC_{0.95} s_{i1} - s_{i2} = -6.38, 3.71$$



**Ejercicio 3.2**

La tabla siguiente contiene los datos de 2 grupos de 10 enfermos reumáticos a los que se les valoró el dolor mediante una escala convencional para comparar 2 fármacos antirreumáticos A y B, administrados cada uno a lo largo de un mes.

Grupo 1		Grupo 2	
A (mes 1)	B (mes 2)	B (mes 1)	A (mes 2)
17	17	21	10
34	41	20	24
26	26	11	32
10	3	26	26

19	-6	42	52
17	-4	28	28
8	11	3	27
16	16	3	28
13	16	16	21
11	4	-10	42

Con la ayuda de R, conteste a las siguientes preguntas:

1. Determine si el efecto (directo) del tratamiento es o no significativo.
2. Dé el intervalo de confianza del 95% del anterior efecto.
3. ¿Existen diferencias entre los dos períodos?
4. Estime el efecto período mediante un IC del 95%.
5. ¿El efecto residual es el mismo en ambos tratamientos? Justifíquelo.
6. Determine la precisión del efecto tardío mediante un IC del 95%. Compárela con la del efecto directo

### 3.3. Tamaño muestral

Como siempre, es necesario fijar de antemano: un riesgo  $\alpha$ , la potencia deseada  $(1 - \beta)$  y la magnitud  $\Delta$  del efecto que se desea establecer.



#### Recuerde

El tamaño muestral en cada grupo para la **comparación de 2 medias** era

$$n = \frac{2 \cdot \sigma^2 \cdot (Z_{\alpha/2} + Z_{\beta})^2}{\Delta^2}$$

Como en los *Cross-over* el valor de referencia se obtiene en el mismo individuo:



#### Definición

El **tamaño muestral** en un estudio cross-over es:

$$N = \frac{2 \cdot \sigma_{\epsilon}^2 \cdot (Z_{\alpha/2} + Z_{\beta})^2}{\Delta^2}$$

**Demostración:** dado que  $\tau_A - \tau_B = \frac{1}{2} d_{i1} - d_{i2}$ , se demuestra que:

$$\text{Var } \tau_A - \tau_B = \text{Var } \frac{1}{2} d_{i1} - d_{i2} = \frac{1}{4} \text{Var } d_{i1} + \text{Var } d_{i2} = \frac{1}{4} \frac{2\sigma_\epsilon^2}{\frac{N}{2}} + \frac{2\sigma_\epsilon^2}{\frac{N}{2}} = \frac{1}{4} \frac{8\sigma_\epsilon^2}{N} = \frac{2\sigma_\epsilon^2}{N}$$

Observe que la fórmula es idéntica a la de comparación de dos medias, pero (1) el error se refiere a la variabilidad *intrasujeto*  $\sigma_\epsilon^2$  y (2) 'N' ahora es la 'N' total (cada observación aporta los dos valores).

**Ejemplo 3.5:** ¿Cuántos casos totales se necesitan en un diseño paralelo y en uno con intercambio si  $\Delta=5u$ ,  $\sigma_\alpha^2=(9u)^2$ ,  $\sigma_\epsilon^2=(4u)^2$ ,  $\alpha=0.05$  bilateral y  $\beta=0.2$ .

Diseño paralelo:  $n = \frac{2 \cdot 9^2 + 4^2 \cdot 1.96 + 0.84^2}{5^2} \approx 60.84 \rightarrow 61$  casos por grupo

Con intercambio:  $N = \frac{2 \cdot 4^2 \cdot 1.96 + 0.84^2}{5^2} \approx 10.04 \rightarrow 11$  casos totales  $\rightarrow 6$  en cada orden

## Soluciones a los ejercicios

1.1.  $CCI = \frac{1.7}{6.3+1.7} = 0.2125$ . Una quinta parte de la variabilidad total entre los casos viene explicada por los centros.

Los casos de un mismo centro comparten una quinta parte de la variabilidad total. El nivel de relación o de similitud entre los casos de un mismo centro se puede cuantificar en un una quinta parte, aproximadamente. La respuesta de un niño de un centro no es independiente de la de los otros niños del mismo centro. Conocida la respuesta de un niño de un centro, ya dispongo de algo de información sobre los otros niños del mismo centro.

1.2.  $DE = 1 + 214 - 1 \cdot 0.7875 = 168.7375$

1.3.  $n = \frac{130 \cdot 2.6}{9} = 37.56 \approx 38$  pacientes en cada cluster.

2.1. A partir del enunciado, se extrae que  $\sigma_\alpha^2 = 16^2 = 256$  y  $\sigma_\epsilon^2 = 9^2 = 81$ .

a. Si se define como variable respuesta la determinación de una única medida al final del tratamiento:

$$V(Y) = V(Y_\alpha + Y_\epsilon) = V(Y_\alpha) + V(Y_\epsilon) = \sigma_\alpha^2 + \sigma_\epsilon^2 = 256 + 81 = 337$$

b. En caso de definir la respuesta como la diferencia entre la medida basal y la final, la varianza de la variable es:

$$\text{var } D_{ij} = V(Y_{ij} - Z_{ij}) = 2 \cdot \sigma_\epsilon^2 = 2 \cdot 81 = 162$$

c. La correlación entre la variable final y la basal es:

$$\rho = \frac{\sigma_\alpha^2}{\sigma_\alpha^2 + \sigma_\epsilon^2} = \frac{256}{256 + 81} = \frac{256}{337} = 0.76$$

d. Dado que la varianza en caso de utilizar la puntuación del cambio como respuesta es menor, el análisis del cambio es el diseño más eficiente de los dos propuestos.

2.2.

a. Usando el diseño ANCOVA y algunos resultados calculados en el anterior ejercicio, la varianza de la variable principal es:

$$\sigma_{Anc}^2 = 1 - \rho^2 \cdot \sigma_y^2 = 1 - 0.76^2 \cdot 337 = 142.53$$

b. La varianza de la respuesta como promedio de 5 repeticiones es:

$$\text{var } m_5 Y = \sigma_\alpha^2 + \sigma_\epsilon^2 / 5 = 256 + \frac{81}{5} = 267.57$$

c. De todos los análisis efectivos, el más eficiente es el ANCOVA, ya que es con el que se obtiene una variable respuesta con menor varianza.

3.1. (1) Falsa; (2) Cierta; (3) Falsa; (4) Falsa; (5) Cierta, se comparan las medias de las diferencias de la respuesta entre el primer y segundo periodo entre ambas secuencias mediante la t-d'Student; (6) Cierta; (7) Falsa.

3.2.

1. Se utiliza la función t.test de R:

```
> datos = read.table("clipboard", header = T)
> d1 = with(datos, A_m1 - B_m2)
> d2 = with(datos, B_m1 - A_m2)
> # Efecto directo del tratamiento
> t.test(d1, d2, var.equal=TRUE)
Two Sample t-test
data:  d1 and d2
t = 2.6792, df = 18, p-value = 0.01531
alternative hypothesis: true difference in means is not equal to 0
95 percent confidence interval:
 3.820304 31.579696
sample estimates:
```

mean of x mean of y  
4.7 -13.0

Se obtiene un p-valor inferior al 5%, por lo que se rechaza la hipótesis nula de efectos iguales. A esta misma conclusión se llega con el intervalo de confianza.

2. El intervalo de confianza para  $\tau_A - \tau_B$  es:

$$IC_{0,95} \tau_A - \tau_B = \frac{1}{2} IC_{0,95} d_{i1} - d_{i2} = \frac{1}{2} 3.82, 31.58 = 1.91, 15.79$$

3. El código de R para resolver el contraste y su solución es:

```
> # Efecto del período
> t.test(d1, -d2, var.equal=TRUE)
Two Sample t-test
data: d1 and -d2
t = -1.2563, df = 18, p-value = 0.2251
alternative hypothesis: true difference in means is not equal to 0
95 percent confidence interval:
 -22.179696  5.579696
sample estimates:
 mean of x mean of y
 4.7      13.0
```

Se obtiene un p-valor superior al 5%, por lo que se no se puede rechazar la hipótesis nula. No hay diferencias entre los períodos.

4. El intervalo de confianza para  $\pi_A - \pi_B$  es:

$$IC_{0,95} \pi_1 - \pi_2 = \frac{1}{2} IC_{0,95} d_{i1} + d_{i2} = \frac{1}{2} -22.18, 5.58 = -11.09, 2.79$$

5. El código de R así como el resultado del contraste de si hay igualdad en los dos efectos residuales se muestra a continuación:

```
> # Efecto residual del tratamiento
> s1 = with(datos, A_m1 + B_m2)
> s2 = with(datos, B_m1 + A_m2)
> t.test(s1, s2, var.equal=TRUE)
Two Sample t-test
data: s1 and s2
t = -1.737, df = 18, p-value = 0.09947
alternative hypothesis: true difference in means is not equal to 0
95 percent confidence interval:
 -34.247451  3.247451
sample estimates:
 mean of x mean of y
 29.5      45.0
```

Se obtiene un p-valor superior al 5%, por lo que se no se puede rechazar la hipótesis nula: el efecto residual es el mismo en ambos tratamientos.

6. La precisión del efecto tardío es:

$$IC_{0,95} O_A - O_B = -34.25, 3.25$$