

TEMA 5

DISEÑO EXPERIMENTAL

MÁSTER DE ESTADÍSTICA APLICADA
CON R SOFTWARE. TÉCNICAS
CLÁSICAS, ROBUSTAS, AVANZADAS Y
MULTIVARIANTES



TEMA 5. DISEÑO EXPERIMENTAL



Definición

Un *diseño experimental* es una regla que determina la asignación de las unidades experimentales a los tratamientos. El *análisis* del diseño experimental utiliza modelos estadísticos clásicos con el objetivo de estudiar el efecto de una o más variables explicativas (factores) sobre una variable de interés (respuesta), y cuantificarlo.

¿Cómo puedo obtener de los experimentos la mayor información posible y de la manera más eficiente?

A menudo los investigadores utilizan diseños inapropiados o subóptimos, que posteriormente dificultan el análisis y producen resultados dudosos. Planificar adecuadamente nuestra experimentación nos permitirá:

1. Seleccionar la estrategia experimental óptima que permita obtener la información buscada con el mínimo coste.
2. Evaluar los resultados experimentales obtenidos, garantizando la máxima fiabilidad en las conclusiones que se obtengan.

Conceptos

- :: *Unidades experimentales*: son los elementos (personas, plantas, animales u objetos) a los que se les modificarán en forma planificada ciertos factores para revisar su respuesta.
- :: *Factor*: es la variable controlada por el experimentador (Niveles del factor o tratamientos).
- :: *Variable de interés*: es la variable respuesta.
- :: *Error experimental o perturbación*: es la variación de los resultados debido a factores externos al experimento (variables no controladas por el experimentador).
- :: *Tamaño muestral*: corresponde al número total de observaciones.
- :: *Tratamiento*: es el proceso de modificación de factores de una unidad experimental cuyos efectos van a ser medidos y comparados.
- :: *Repetición o réplicas*: es el número de veces que un tratamiento se aplica al mismo número de unidades experimentales.
- :: *Bloque*: es un conjunto de unidades experimentales homogéneas.

Principios clásicos del diseño experimental

- ∴ el *control* es un grupo de unidades experimentales normalmente “no tratadas”, con las que se compararán los tratamientos experimentales; es un “nivel de referencia”.
- ∴ la *replicación* implica el uso de varias unidades experimentales que reciben el mismo tratamiento.
- ∴ la *aleatorización* es la asignación al azar de los tratamientos a las distintas unidades experimentales.
- ∴ el *bloqueo* implica dividir las unidades experimentales en grupos homogéneos llamados *bloques*. Nos permite controlar la variabilidad sistemática no planificada. A diferencia de los *factores tratamiento*, el experimentador no está interesado en investigar las diferencias entre los niveles del *factor bloque*.
- ∴ la *factorización* consiste en cruzar los niveles de todos los factores tratamientos en todas las combinaciones posibles. Utilizar diseños factoriales nos permite detectar efectos de interacción entre los distintos factores tratamiento.



Nota

Según el tamaño muestral podemos tener dos tipos de modelos:

- ∴ *Modelo equilibrado o balanceado*: mismo número de observaciones por celda.
- ∴ *Modelo no-equilibrado o desbalanceado*: distinto número de observaciones por celda.

5.1 Clasificación

Existen diseños estándar (*modelos*) que se utilizan con mucha frecuencia. Veremos a continuación algunos de estos diseños experimentales clásicos que desarrollaremos más adelante en el temario.

5.1.1 Diseños unifactoriales

5.1.1.1 Diseño completamente aleatorizado

Asigna al azar las unidades experimentales a los tratamientos. Este es el diseño más sencillo, preferentemente utilizable cuando las unidades experimentales son bastante homogéneas.

El modelo matemático de este diseño es:

Respuesta=Constante+Efecto tratamiento+Error

Ejemplo: se evalúa un conjunto de fertilizantes (tratamientos) para comprobar si éstos tienen efectos distintos sobre el rendimiento de un tipo de cereal (respuesta).

- i. **Modelo de efectos fijos:** si los tratamientos han sido seleccionados específicamente por el investigador y por ende las conclusiones solo se aplican a dichos tratamientos.
- ii. **Modelo de efectos aleatorios:** si se ha realizado una selección aleatoria de los tratamientos y por tanto las conclusiones pueden extenderse a todos los tratamientos de la población.

5.1.1.2 Diseños en bloques aleatorizados

Diseños en bloques aleatorizados

Las unidades experimentales son agrupadas en bloques (grupos homogéneos) respecto a una cierta variable de control que puede suponer una fuente adicional de variabilidad. Una vez formados los bloques se asigna al azar las unidades experimentales a los tratamientos dentro de cada bloque.

El modelo matemático de este diseño es:

$$\text{Respuesta} = \text{Constante} + \text{Efecto bloque} + \text{Efecto tratamiento} + \text{Error}$$

Cuando existe variabilidad en las unidades experimentales el investigador tiene principalmente dos opciones:

- ∴ considerar solo unidades experimentales muy homogéneas
- ∴ formar *bloques* de unidades experimentales de manera que las condiciones en cada bloque sean lo más homogéneas posibles y los bloques sean heterogéneos entre sí.

Ejemplo: las parcelas de terreno deben ser similares en calidad y superficie para eliminar del error experimental la variabilidad producida por las diferencias entre parcelas. Para ello se seleccionan parcelas con características similares (bloques) donde realizar el experimento. Los bloques no son el motivo del estudio sino que ocurren de manera natural.

- i. **Diseños en bloques completos aleatorizados:** se dice que es un diseño *completo* cada tratamiento se observa el mismo número de veces en cada bloque. Llamamos *diseños en bloques completamente aleatorizados* cuando se tiene una única observación por cada tratamiento.
- ii. **Diseños en bloques incompletos aleatorizados:** se dice que es un diseño *incompleto* cuando no es posible observar la totalidad de tratamientos en cada bloque, y aquellos que sí están representados en uno en particular se ensayan en él una sola vez.

5.1.2 Diseños factoriales

Partamos de un ejemplo del *Diseño completamente aleatorizado uni-factorial* en el que aplicamos 5 tipos de fertilizantes (niveles del tratamiento) a ciertas unidades experimentales que son similares en humedad, tipo

de suelo, pendiente, fertilidad, etc.. Tenemos un solo factor tratamiento que son los fertilizantes.

Extenderemos este ejemplo para hablar de los *diseños factoriales*, que presentan dos o más factores:

5.1.2.1 Diseños factoriales o cruzados

Los niveles de cada factor están combinados con todos los niveles de los restantes factores (*cruzados*) o, lo que es lo mismo, todas las combinaciones de los niveles se aplican a las unidades experimentales.

Ejemplo: de cuando aplicamos 2 tratamientos (2 niveles de riego y 5 niveles de fertilizante) a las unidades experimentales, que son todas similares en humedad, tipo de suelo, pendiente, fertilidad, etc. (si además es CRD).

5.1.2.2 Diseños jerárquicos o anidados

Ciertos niveles de un factor están ligados a ciertos niveles de otro factor o, lo que es lo mismo, cada nivel particular de uno de los factores bloque ocurre en un único nivel del otro factor bloque (*jerarquizados o anidados*). Entonces, todos los niveles de un factor tratamiento no ocurren en combinación con todos los niveles del otro factor tratamiento.

Ejemplo: este tipo de diseño ocurre cuando tenemos 3 niveles de fertilizantes (N) y 6 niveles de materia orgánica (OM), pero cada nivel de nitrógeno es combinado con solo 2 niveles de materia orgánica. Por tanto, tenemos solo 6 combinaciones de tratamiento: N1-OM1, N1-OM2, N2-OM3, N2-OM4, N3-OM5, N3-OM6. Si además las unidades experimentales son similares en humedad, tipo de suelo, pendiente, fertilidad, etc., entonces además es un diseño CRD.

5.1.2.3 Diseños split-plot o parcelas divididas

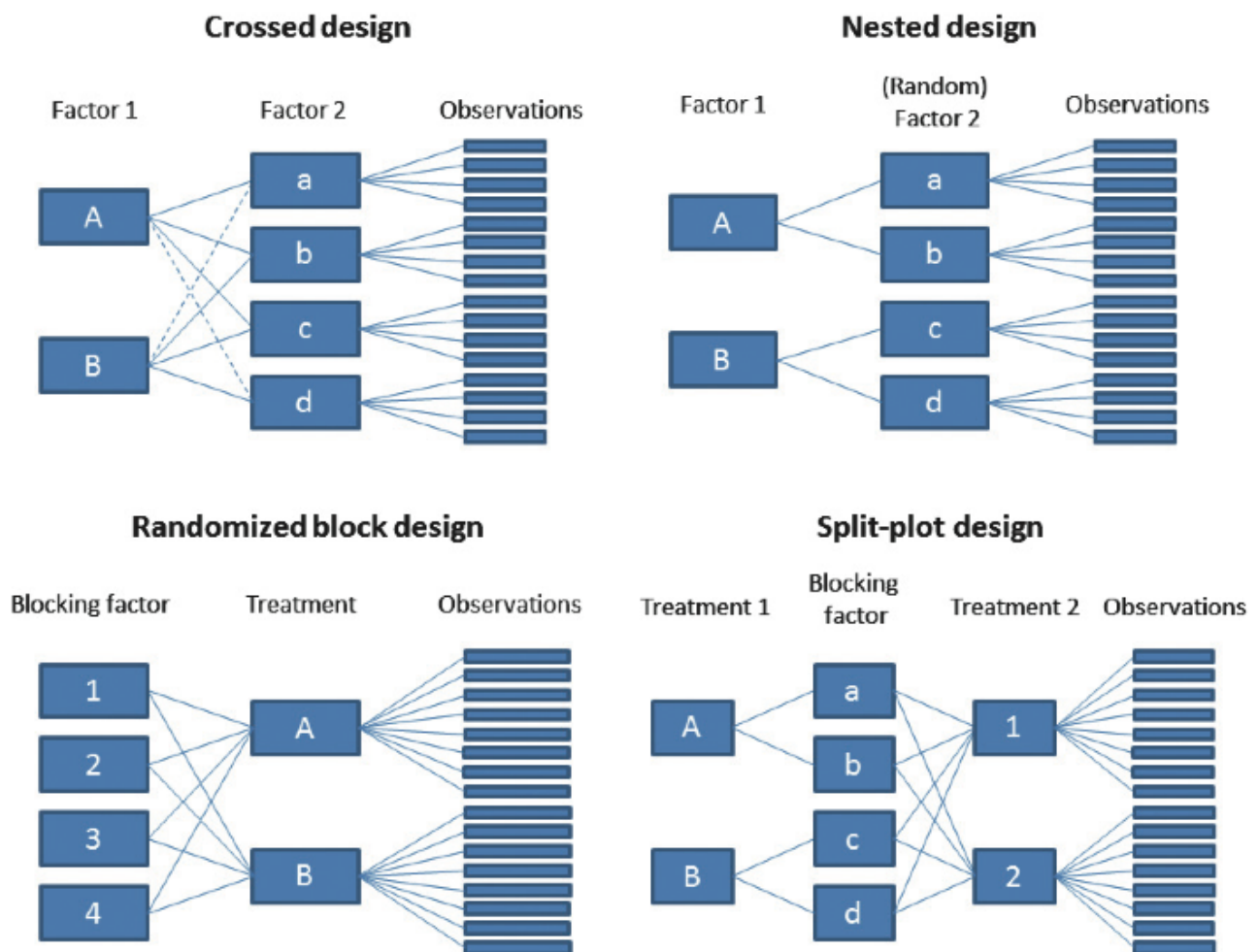
Combina factores cruzados y anidados. Ocurre cuando no es posible aleatorizar el orden de los ensayos dentro de cada bloque. Un factor se aplica a todo el grupo de unidades experimentales (a la trama entera de parcelas). Toda la trama se divide físicamente en pequeñas unidades experimentales (subparcelas) donde se asignan los niveles del segundo factor. El experimento contempla todas las combinaciones de ambos niveles.

Ejemplo: cuando aplicamos dos tratamientos, 2 niveles de riego que se aplican a todas las unidades experimentales (parcelas) y cada unidad experimental se divide en 5 subparcelas donde se aplican 5 de fertilizantes. Si además las unidades experimentales son similares en humedad, tipo de suelo, pendiente, fertilidad, etc., entonces además es un diseño CRD.

5.1.2.4 Diseños de medidas repetidas y datos longitudinales

Un factor tratamiento se aplica a una unidad experimental grande que luego es medida repetidamente en el tiempo o espacio (segundo factor de tratamiento). Cuando las medidas repetidas se toman en el tiempo, se llama estudio *longitudinal* o estudio *panel*. En este caso tenemos un factor tratamiento entre-grupos o inter-sujetos y otro factor tratamiento intra-grupo o intra-sujetos (*medidas repetidas*).

Ejemplo: tenemos 2 niveles de riego que aplicamos en 4 días. Cada unidad experimental se riega o no, a lo largo de 4 días.



En los **diseños factoriales** se desea estudiar el efecto de interacción entre los factores. Pero para más de dos factores se necesitarían demasiadas observaciones. Por ello, muchas veces se utilizan las **fracciones factoriales** que son diseños en los que se supone que muchas de las interacciones son nulas. Por ejemplo, el *cuadrado latino* supone que todas las interacciones son nulas y permite estudiar 3 factores con k niveles con k^2 observaciones en lugar de k^3 observaciones para el modelo completo.

5.1.2.5 Cuadrado latino

Se trata de un diseño para experimentos con unidades experimentales muy variables. El caso más sencillo es un cuadrado, donde las filas y columnas forman los bloques (dos factores bloque) y los tratamientos se reparten de tal manera que el número de unidades experimentales en cada fila y cada columna es igual al número de tratamientos.

Ejemplo: aplicamos 4 tratamientos de cobertura vegetal a las unidades experimentales, y estas unidades difieren en humedad y fertilidad. Tenemos 16 unidades experimentales (4 filas y 4 columnas) donde cada fila difiere en humedad y cada columna en fertilidad. El número de unidades experimentales en cada fila y cada columna es igual al número de tratamientos.

5.2 Diseño uni-factorial completamente aleatorizado

Se lo conoce en inglés como *complete random design* (CRD).

Diseño unifactorial.

Cuando se trata de 1 factor con varios niveles (tratamientos), y no se asume relación entre los niveles. Se llama estructura de tratamiento de una vía (o unifactorial).

Modelo

El modelo matemático para los datos de un diseño completamente aleatorizado es:

$$y_{ij} = \mu + \alpha_i + \varepsilon_{ij}$$

donde tenemos la observación j -ésima del tratamiento i , con el tamaño muestral n (o número de observaciones de la variable respuesta bajo el tratamiento i), μ representa la media global de y común a todos los tratamientos, α_i representa el efecto del tratamiento i (o la variación de la media del nivel i respecto a la media de la respuesta μ), y ε_{ij} es una variable aleatoria que representa el componente aleatorio de error o error experimental (independientes y normales con media cero y varianza σ^2). Los parámetros α_i se llaman **efectos**. Evaluar las diferencias entre niveles equivale a analizar las diferencias entre los parámetros α_i del modelo.

Se trata de un modelo **lineal** por lo que supone que la varianza de la respuesta es constante (homocedasticidad), que la distribución de la respuesta es normal (o de modo equivalente, que los errores son normales) y que las observaciones y_{ij} son independientes (o los errores son independientes).

Para realizar su ajuste utilizaremos la función “lm” en R-software.

Hipótesis global

Las hipótesis a plantear son:

- ∴ H_0 : las medias de los niveles de tratamiento son similares (no existe un efecto del tratamiento)
- ∴ H_1 : alguna de las medias de los niveles de tratamiento es diferente

Para realizar su contraste utilizaremos el ANOVA con la función “aov” en R-software. El ANOVA nos permite dividir la variabilidad de la respuesta en dos partes: 1) la que se debe al factor de interés (diferencias entre los niveles del tratamiento) y 2) la generada por otros factores llamados de perturbación o error experimental, conocidos o no, controlables o no.

Una medida relativa de la variabilidad explicada por el factor es el **coeficiente de determinación** R^2 .

Comparaciones múltiples post hoc

Cuando rechazamos la hipótesis nula del ANOVA, podemos afirmar que existen diferencias significativas entre al menos dos de los niveles del tratamiento. Para encontrar exactamente qué medias difieren entre sí utilizaremos las comparaciones múltiples post hoc.

Utilizaremos la función *TukeyHSD* o el *SNK.test* del paquete *agricolae* o si tenemos un grupo control utilizaremos la función *glht* del paquete *multcomp*.

Inferencia de los parámetros del modelo

Para contrastar y estimar los parámetros del modelo se consideran 2 situaciones:

- si no podemos rechazar H_0 , el modelo anterior se simplifica a $y_{ij} = \mu + \epsilon_{ij}$. Se acepta que el factor no influye ($\alpha_i = 0$) y por tanto los datos provienen de una única muestra homogénea cuyos parámetros μ y σ^2 se estimarán según las técnicas clásicas.
- si rechazamos la H_0 , el factor influye y podremos estimar los parámetros del modelo y sus intervalos de confianza.

Supuestos

- ∴ varianza homogénea entre tratamientos, o lo que es lo mismo, la varianza del error experimental σ^2 debe ser constante, para todos los niveles del factor tratamiento
- ∴ normalidad de los errores experimentales

Para verificar estos supuestos realizaremos pruebas gráficas sobre el modelo ANOVA ajustado.

5.2.1 Aplicación con R

Utilizaremos los datos de 24 tiempos de coagulación de sangre de un experimento donde 24 animales fueron aleatoriamente asignados a 4 dietas diferentes y las muestras se tomaron en orden aleatorio (Box, Hunter & Hunter, 1978).

Observamos los datos

```
library(faraway)
data(coagulation)
```



```
head(coagulation)
```

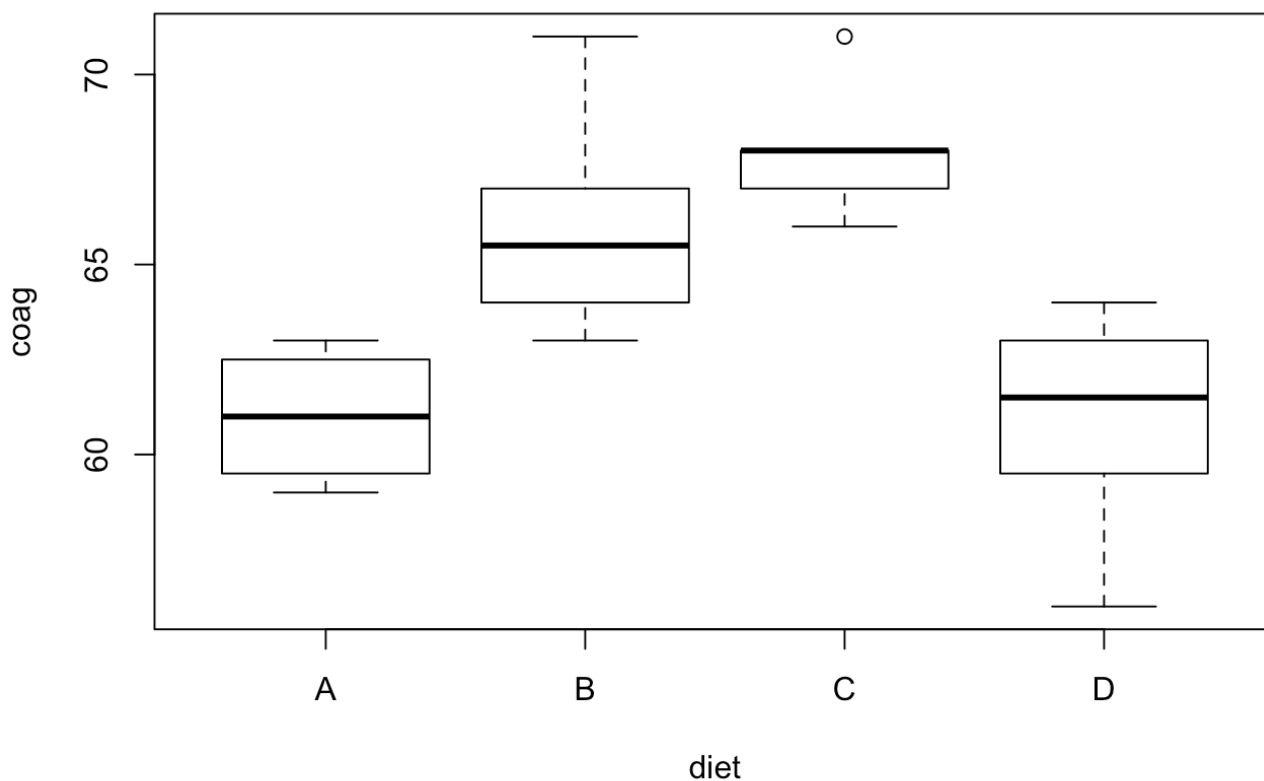
```
##   coag diet
## 1   62   A
## 2   60   A
## 3   63   A
## 4   59   A
## 5   63   B
## 6   67   B
```

```
str(coagulation)
```

```
## 'data.frame':   24 obs. of  2 variables:
##  $ coag: num   62  60  63  59  63  67  71  64  65  66 ...
##  $ diet: Factor w/ 4 levels "A","B","C","D": 1 1 1 1 2 2 2 2 2 2 ...
```

Graficamos

```
plot(coag ~ diet, data=coagulation)
```



No deberíamos detectar problemas de outliers, asimetría y varianza desigual, pero continuaremos con el ejemplo a modo de ilustración.

Modelo y ANOVA

```
fit0 <- lm(coag ~ diet, coagulation)
summary(fit0)
```

```
##
## Call:
## lm(formula = coag ~ diet, data = coagulation)
##
## Residuals:
##      Min       1Q   Median       3Q      Max
## -5.00  -1.25   0.00   1.25   5.00
##
## Coefficients:
##              Estimate Std. Error t value Pr(>|t|)
## (Intercept)  6.100e+01  1.183e+00  51.554 < 2e-16 ***
## dietB        5.000e+00  1.528e+00   3.273 0.003803 **
## dietC        7.000e+00  1.528e+00   4.583 0.000181 ***
## dietD        2.991e-15  1.449e+00   0.000 1.000000
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## Residual standard error: 2.366 on 20 degrees of freedom
## Multiple R-squared:  0.6706, Adjusted R-squared:  0.6212
## F-statistic: 13.57 on 3 and 20 DF, p-value: 4.658e-05
```

```
anova(fit0)
```

```
## Analysis of Variance Table
##
## Response: coag
##           Df Sum Sq Mean Sq F value    Pr(>F)
## diet        3     228    76.0    13.571 4.658e-05 ***
## Residuals   20     112     5.6
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
```

Existen diferencias significativas en los tiempos de coagulación de cada tipo de dieta ($F(3,20)=13.57$, $p<0.001$). El nivel de referencia corresponde al grupo A y tiene una media de 61, grupo B, C y D tienen valores 5, 7 y 0 puntos mayores que la media de A. El modelo logra un ajuste bueno ($R^2=62.12\%$).

NOTA: es equivalente realizar:

```
summary(aov(coag ~ diet, coagulation))
```

```
##           Df Sum Sq Mean Sq F value    Pr(>F)
## diet        3     228    76.0    13.57 4.66e-05 ***
## Residuals   20     112     5.6
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
```

Podemos observar directamente las medias si ajustamos un modelo sin intercepto, pero las pruebas no son útiles ya que comparan con cero (mira por ejemplo que el valor del R^2 es erróneo).

```
fit1 <- lm(coag ~ diet - 1, coagulation)
summary(fit1)
```

```
##
## Call:
## lm(formula = coag ~ diet - 1, data = coagulation)
##
## Residuals:
##      Min       1Q   Median       3Q      Max
## -5.00  -1.25   0.00   1.25   5.00
```

```
##
## Coefficients:
##      Estimate Std. Error t value Pr(>|t|)
## dietA  61.0000     1.1832   51.55  <2e-16 ***
## dietB  66.0000     0.9661   68.32  <2e-16 ***
## dietC  68.0000     0.9661   70.39  <2e-16 ***
## dietD  61.0000     0.8367   72.91  <2e-16 ***
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## Residual standard error: 2.366 on 20 degrees of freedom
## Multiple R-squared:  0.9989, Adjusted R-squared:  0.9986
## F-statistic: 4399 on 4 and 20 DF, p-value: < 2.2e-16
```

Comparaciones múltiples de medias post hoc

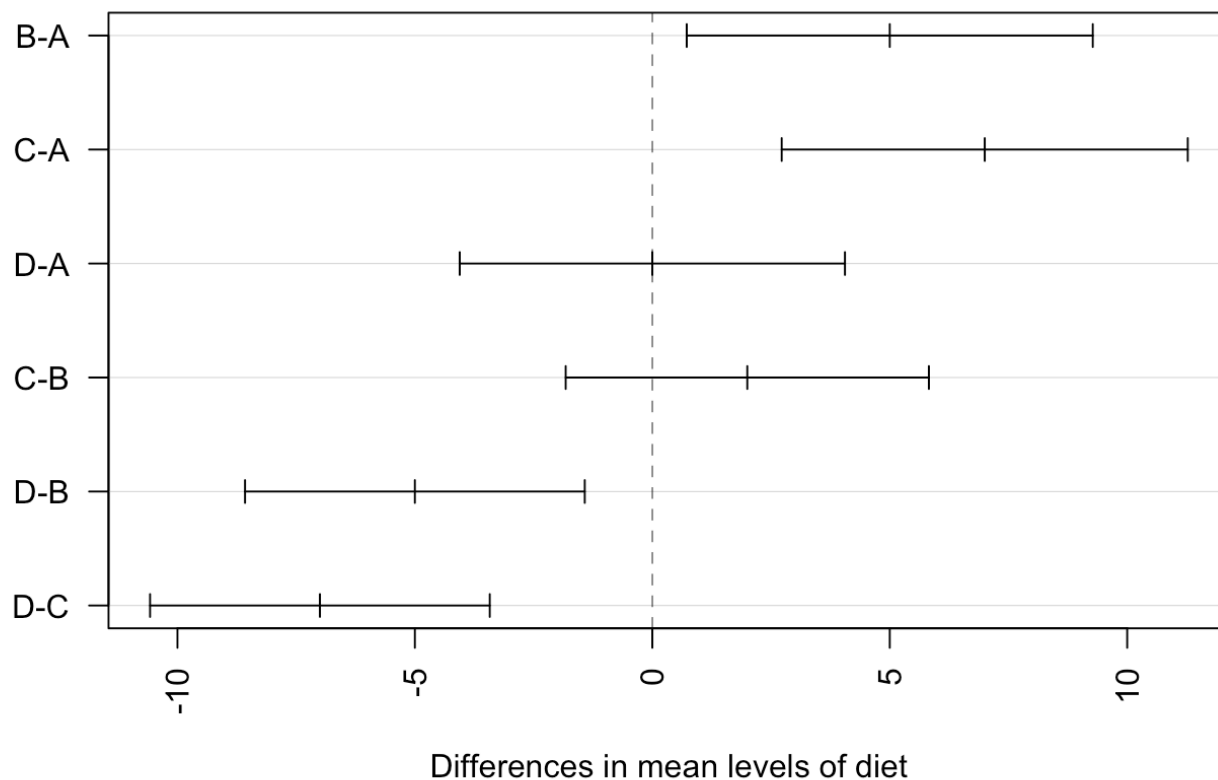
Luego de detectar diferencias significativas entre los grupos, nos interesa detectar qué niveles son diferentes.

```
(m<-TukeyHSD(aov(coag ~ diet, coagulation)))
```

```
## Tukey multiple comparisons of means
## 95% family-wise confidence level
##
## Fit: aov(formula = coag ~ diet, data = coagulation)
##
## $diet
##      diff      lwr      upr    p adj
## B-A     5  0.7245544  9.275446 0.0183283
## C-A     7  2.7245544 11.275446 0.0009577
## D-A     0 -4.0560438  4.056044 1.0000000
## C-B     2 -1.8240748  5.824075 0.4766005
## D-B    -5 -8.5770944 -1.422906 0.0044114
## D-C    -7 -10.5770944 -3.422906 0.0001268
```

```
plot(m, las=2)
```

95% family-wise confidence level



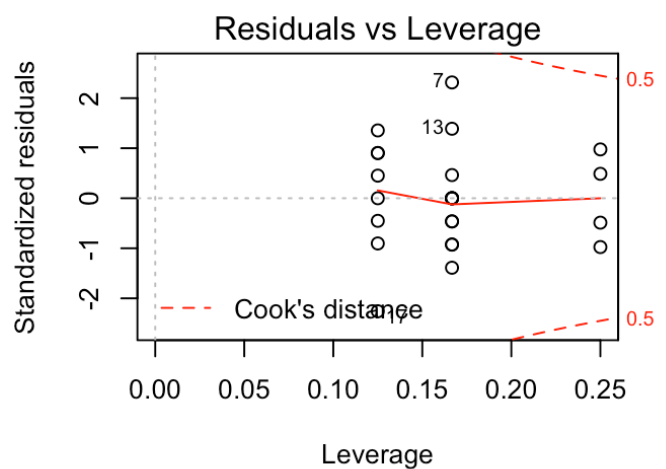
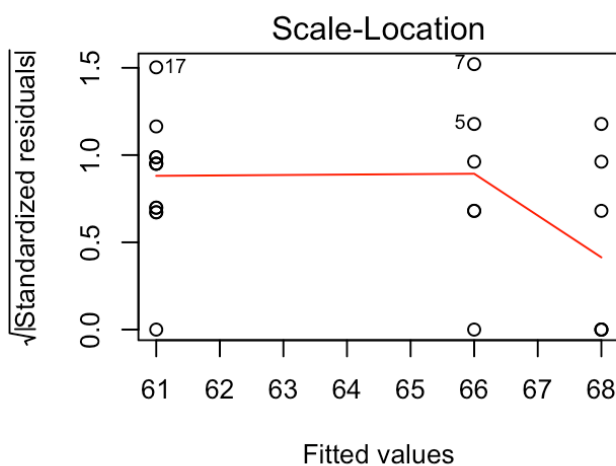
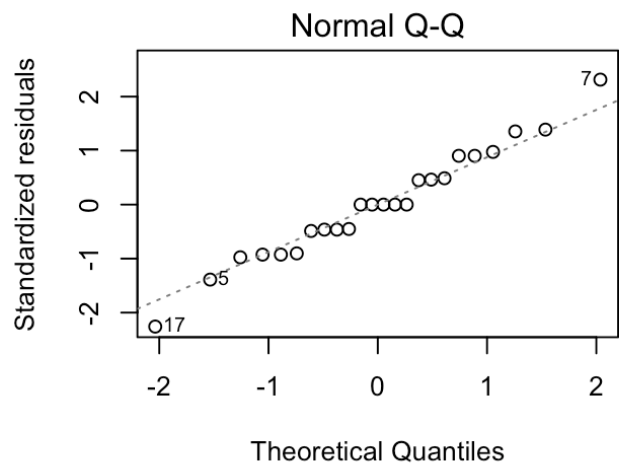
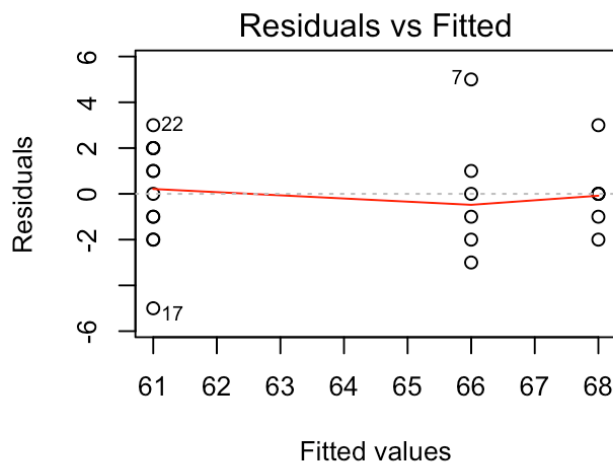
```
# obtenemos Las medias para cada grupo  
tapply(coagulation$coag,coagulation$diet, mean)
```

```
## A B C D  
## 61 66 68 61
```

Las únicas diferencias de medias que no son significativas son D-A y C-B.

Diagnóstico

```
par(mfrow=c(2,2))  
plot(fit0)
```



Los datos son aproximadamente normales y hay homogeneidad de varianza, a pesar de que los datos sean discretos.

5.3 Diseño uni-factorial en bloques aleatorizados

Se lo conoce en inglés como *Randomized complete block design* (RCBD).

Bloques

Cuando las unidades experimentales presentan un patrón de heterogeneidad definido, necesitamos formar bloques de manera tal que en cada grupo las unidades experimentales sean homogéneas entre sí. Los bloques se tratan como los niveles de un único factor de bloqueo.

Modelo

El modelo matemático para los datos de un diseño por bloques aleatorizados es:

$$y_{ij} = \mu + \alpha_i + \beta_j + \varepsilon_{ij}$$

donde tenemos que y_{ij} representa la respuesta de la observación j -ésima del tratamiento i , con el tamaño muestral n (o número de observaciones de la variable respuesta bajo el tratamiento i), μ representa la media global común a todos los tratamientos, α_i representa el efecto del tratamiento i , β_j corresponde al efecto del j -ésimo bloque ($j = 1 \dots J$) y ε_{ij} es una variable aleatoria que representa el componente aleatorio de error (independientes y normales con media cero y varianza σ^2).

El efecto de bloque (β_j) es fijo, aunque el tratamiento puede ser fijo o aleatorio.

Hipótesis

Las hipótesis a plantear son las mismas que para el modelo anterior:

- ⌘ H_0 : las medias de los niveles de tratamiento son similares (no existe un efecto del tratamiento)
- ⌘ H_1 : alguna de las medias de los niveles de tratamiento es diferente

Las hipótesis referentes a los bloques son:

- ⌘ H_0 : las medias de los bloques son similares (no existe un efecto del bloque)
- ⌘ H_1 : alguna de las medias de los bloques es diferente

En este caso lo que nos interesa es saber si bloquear el experimento resultó o no beneficioso.

Es un análisis muy similar al ANOVA de dos vías sin interacción con una observación por celda. Podemos identificar el efecto principal del tratamiento y además podemos detectar el efecto de bloque.

NOTA: Si se tiene un diseño de experimentos con dos factores tratamiento entre los que no existe interacción, el tratamiento estadístico y desarrollo que se estudia en esta sección es válido para este modelo. Sin embargo, el planteamiento y los objetivos de ambos tipos de análisis son distintos.

Supuestos

- ⌘ aditividad entre los efectos de tratamiento y bloque
- ⌘ errores experimentales normales
- ⌘ errores experimentales independientes y aleatoriamente distribuidos
- ⌘ varianzas homogéneas entre los distintos tratamientos
- ⌘ sin relación entre promedios de los tratamientos y las varianzas.
- ⌘ *se asume que no existe interacción entre el factor tratamiento y el factor bloque*

Test de Tukey

Para poner a prueba la existencia de interacción entre los efectos principales y los bloques (no aditividad de los efectos), podemos utilizar la prueba de Tukey para la no aditividad con la función `tukey.add.test` del paquete `asbio`. Si es significativa debemos verificar si su causa se debe a la presencia de *outliers* o a la heterogeneidad de varianzas. Si realmente la causa es la interacción existente entre efectos principales y bloques, esta interacción podría enmascarar los resultados.

5.3.1 Aplicación con R (bloques completos)

Utilizaremos los datos de un experimento para comparar 4 procesos (A, B, C, D) de producción de penicilina. La materia prima, licor de maíz fermentado, es bastante variable y sólo se puede hacer en mezclas suficientes para 4 ensayos. Este diseño de bloques completos al azar está sin duda determinado por la naturaleza de las unidades experimentales.

Observamos los datos

```
data(penicillin) # Library(faraway)
head(penicillin)
```

```
##   treat  blend yield
## 1    A Blend1    89
## 2    B Blend1    88
## 3    C Blend1    97
## 4    D Blend1    94
## 5    A Blend2    84
## 6    B Blend2    77
```

```
str(penicillin)
```

```
## 'data.frame':   20 obs. of  3 variables:
## $ treat: Factor w/ 4 levels "A","B","C","D": 1 2 3 4 1 2 3 4 1 2 ...
## $ blend: Factor w/ 5 levels "Blend1","Blend2",...: 1 1 1 1 2 2 2 2 3 3 ...
## $ yield: num  89 88 97 94 84 77 92 79 81 87 ...
```

```
xtabs(yield~ blend+treat,data=penicillin)
```

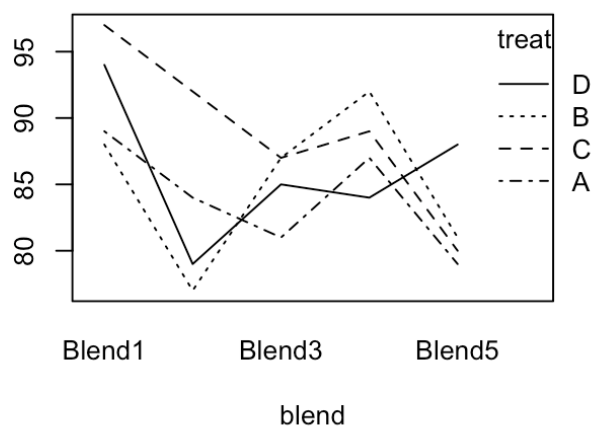
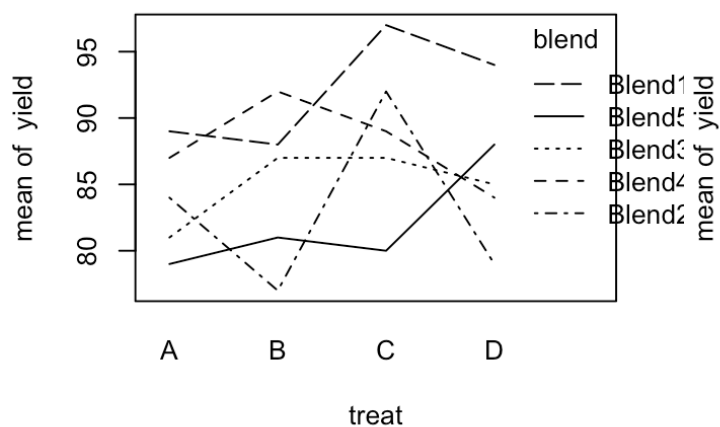
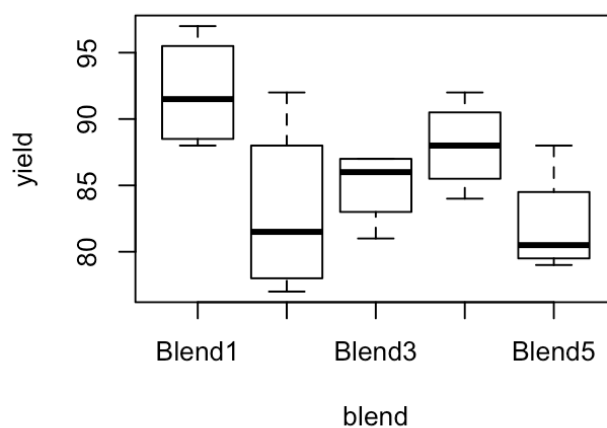
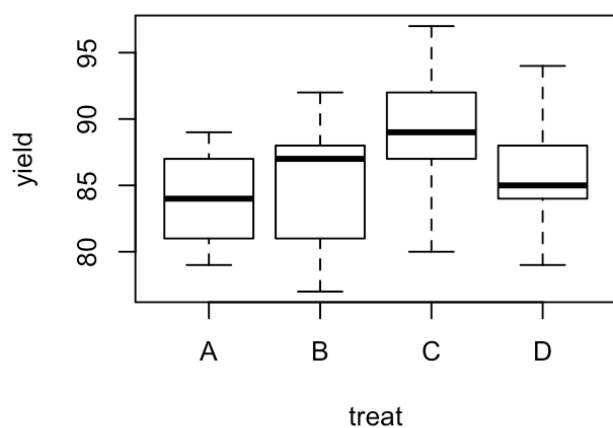
```
##           treat
## blend      A  B  C  D
## Blend1  89 88 97 94
## Blend2  84 77 92 79
## Blend3  81 87 87 85
## Blend4  87 92 89 84
## Blend5  79 81 80 88
```

Tenemos un factor de tratamiento (treat) con 4 niveles (A, B, C, D), un factor de bloque (blend) con 5 niveles (1,...,5), y una observación en cada tratamiento y bloque. Es decir, tenemos un diseño por bloques completamente aleatorizado (RCBD).

Graficamos

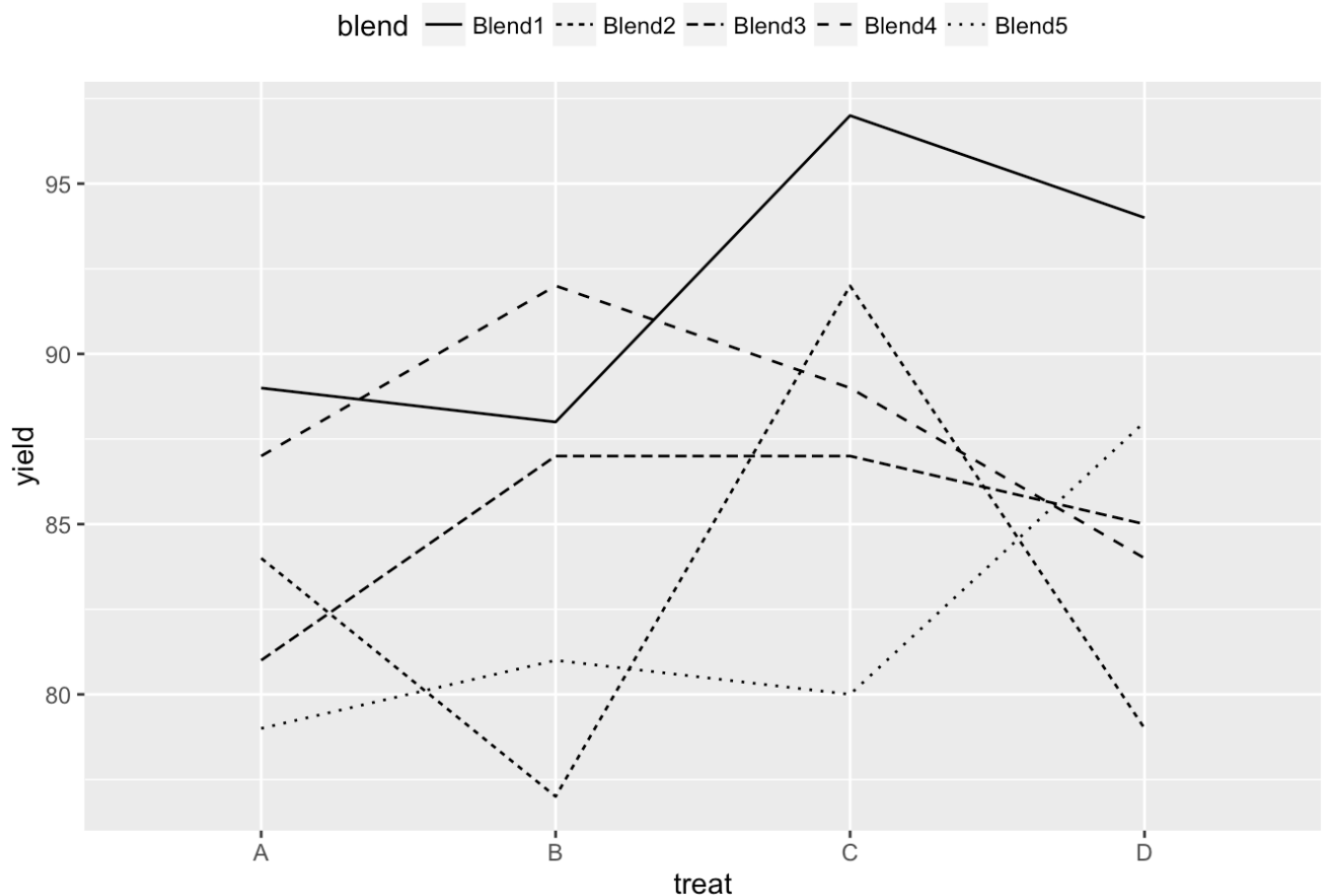
Comprobamos mediante gráficos el efecto del tratamiento (también podemos ver si existe un efecto de bloque) y el efecto de interacción entre bloque y tratamiento (luego realizaremos la prueba de hipótesis).

```
par(mfrow=c(2,2))
plot(yield ~ treat,data=penicillin)
plot(yield ~ blend,data=penicillin)
with(penicillin, interaction.plot(treat, blend, yield))
with(penicillin, interaction.plot(blend, treat, yield))
```



Si queremos un gráfico un poco más sofisticado utilizamos ggplot.

```
library(ggplot2)
ggplot(penicillin, aes(x = treat, y = yield, group = blend, linetype = blend)) +
  geom_line() +
  theme(legend.position = "top", legend.direction = "horizontal")
```

El análisis es similar al ANOVA de 2 vías con una observación por celda pero aquí asumimos que no existe interacción entre tratamiento y bloque. Probamos que no existe un efecto de interacción entre bloque y tratamiento mediante la prueba de Tukey.

```
library(asbio)
# H_0: Main effects and blocks are additive (i.e. there is no interaction effect between blocks and main factors)
# H_A: Main effects and blocks are non-additive.
with(penicillin,tukey.add.test(yield,treat,blend))
```

```
##
## Tukey's one df test for additivity
## F = 0.0982679   Denom df = 11   p-value = 0.7597822
```

No podemos rechazar la hipótesis nula, por lo que no existe efecto de interacción entre bloque y tratamiento. Podemos continuar con nuestro análisis.

Modelo y ANOVA

Ajustamos el modelo lineal a los datos.

```
g <- lm(yield ~ treat+blend,penicillin)
anova(g)
```

```
## Analysis of Variance Table
##
## Response: yield
##           Df Sum Sq Mean Sq F value Pr(>F)
## treat      3      70   23.333   1.2389 0.33866
```

```
## blend      4      264  66.000  3.5044 0.04075 *
## Residuals 12      226  18.833
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
```

No se observa un efecto significativo del tratamiento pero sí del efecto bloque. Los resultados corresponden a la prueba de modelos secuenciales: - $y \sim 1$ - $y \sim \text{treat}$ - $y \sim \text{treat} + \text{blend}$

El $p=0.339$ corresponde a la comparación de los primeros dos modelos y el $p=0.041$ a la comparación de los dos siguientes.

NOTA: el orden de los términos no importa siempre que el diseño sea ortogonal.

Comparaciones múltiples de medias post hoc

¿Qué tratamientos producen los mayores rendimientos?.

Como no hemos detectado un efecto del tratamiento no tiene sentido realizar comparaciones múltiples post hoc, sin embargo, por motivos meramente ilustrativos vamos a desarrollar el código necesario para estas pruebas.

```
# Coeficientes del modelo
summary(g)
```

```
##
## Call:
## lm(formula = yield ~ treat + blend, data = penicillin)
##
## Residuals:
##      Min       1Q   Median       3Q      Max
## -5.000  -2.250  -0.500   2.250   6.000
##
## Coefficients:
##              Estimate Std. Error t value Pr(>|t|)
## (Intercept)   90.000      2.745  32.791  4.1e-13 ***
## treatB         1.000      2.745   0.364  0.72194
## treatC         5.000      2.745   1.822  0.09351 .
## treatD         2.000      2.745   0.729  0.48018
## blendBlend2   -9.000      3.069  -2.933  0.01254 *
## blendBlend3   -7.000      3.069  -2.281  0.04159 *
## blendBlend4   -4.000      3.069  -1.304  0.21686
## blendBlend5  -10.000      3.069  -3.259  0.00684 **
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## Residual standard error: 4.34 on 12 degrees of freedom
## Multiple R-squared:  0.5964, Adjusted R-squared:  0.361
## F-statistic: 2.534 on 7 and 12 DF,  p-value: 0.07535
```

```
# media para cada celda
with(penicillin, tapply(yield, treat, mean))
```

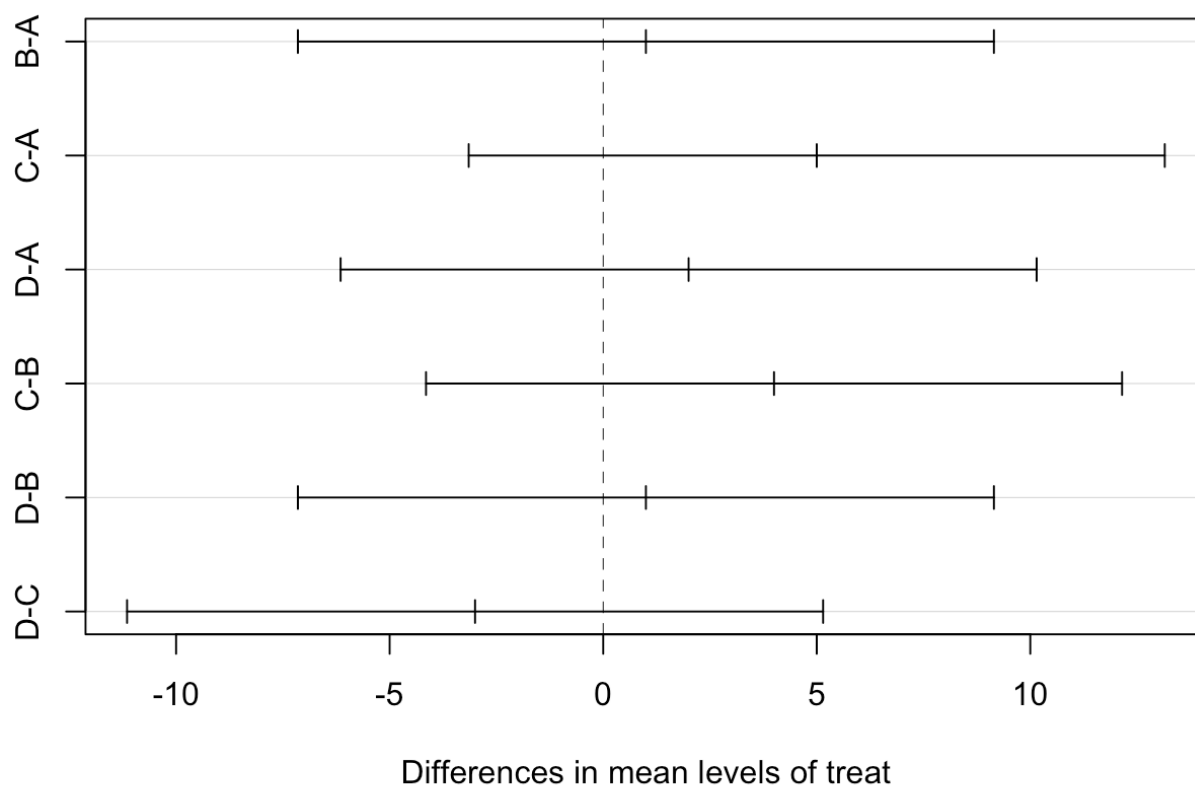
```
##  A  B  C  D
## 84 85 89 86
```

```
# pruebas de comparación múltiple, distintas funciones.
(DifT <- TukeyHSD(aov(g), "treat"))
```

```
## Tukey multiple comparisons of means
## 95% family-wise confidence level
##
## Fit: aov(formula = g)
##
## $treat
##      diff      lwr      upr    p adj
## B-A      1  -7.148719  9.148719 0.9826684
## C-A      5  -3.148719 13.148719 0.3105094
## D-A      2  -6.148719 10.148719 0.8837551
## C-B      4  -4.148719 12.148719 0.4905194
## D-B      1  -7.148719  9.148719 0.9826684
## D-C     -3 -11.148719  5.148719 0.7002271
```

```
plot(DifT)
```

95% family-wise confidence level



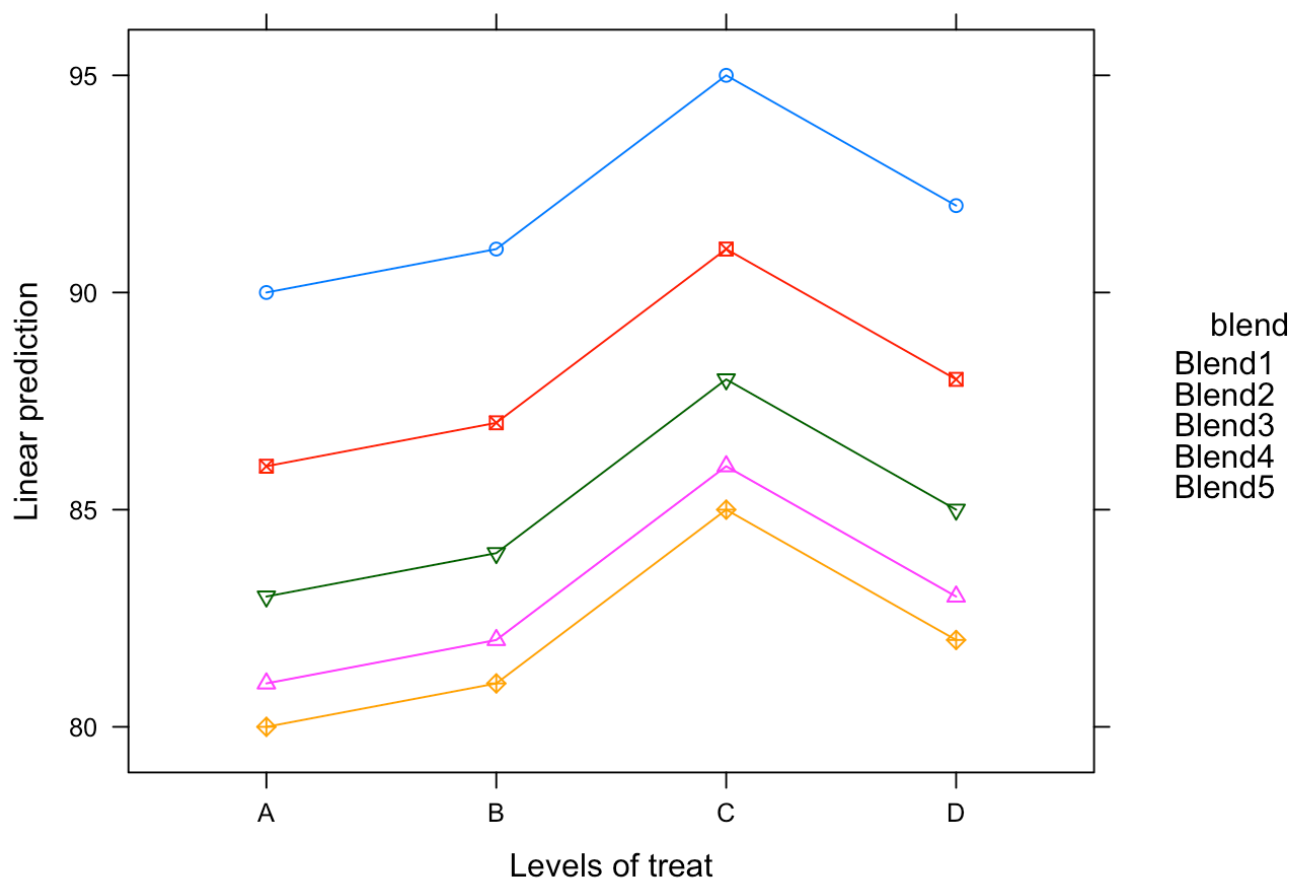
```
library(multcomp)
summary(glht(g, linfct=mcp (treat="Tukey")))
```

```
##
## Simultaneous Tests for General Linear Hypotheses
##
## Multiple Comparisons of Means: Tukey Contrasts
##
##
## Fit: lm(formula = yield ~ treat + blend, data = penicillin)
##
## Linear Hypotheses:
##      Estimate Std. Error t value Pr(>|t|)
## B - A == 0    1.000      2.745   0.364   0.983
```

```
## C - A == 0    5.000    2.745    1.822    0.311
## D - A == 0    2.000    2.745    0.729    0.884
## C - B == 0    4.000    2.745    1.457    0.491
## D - B == 0    1.000    2.745    0.364    0.983
## D - C == 0   -3.000    2.745   -1.093    0.700
## (Adjusted p values reported -- single-step method)
```

```
library(agricolae)
HSD.test(g, "treat")

library(lsmmeans) #para predicciones LS
lsmip(g, blend ~ treat)
```

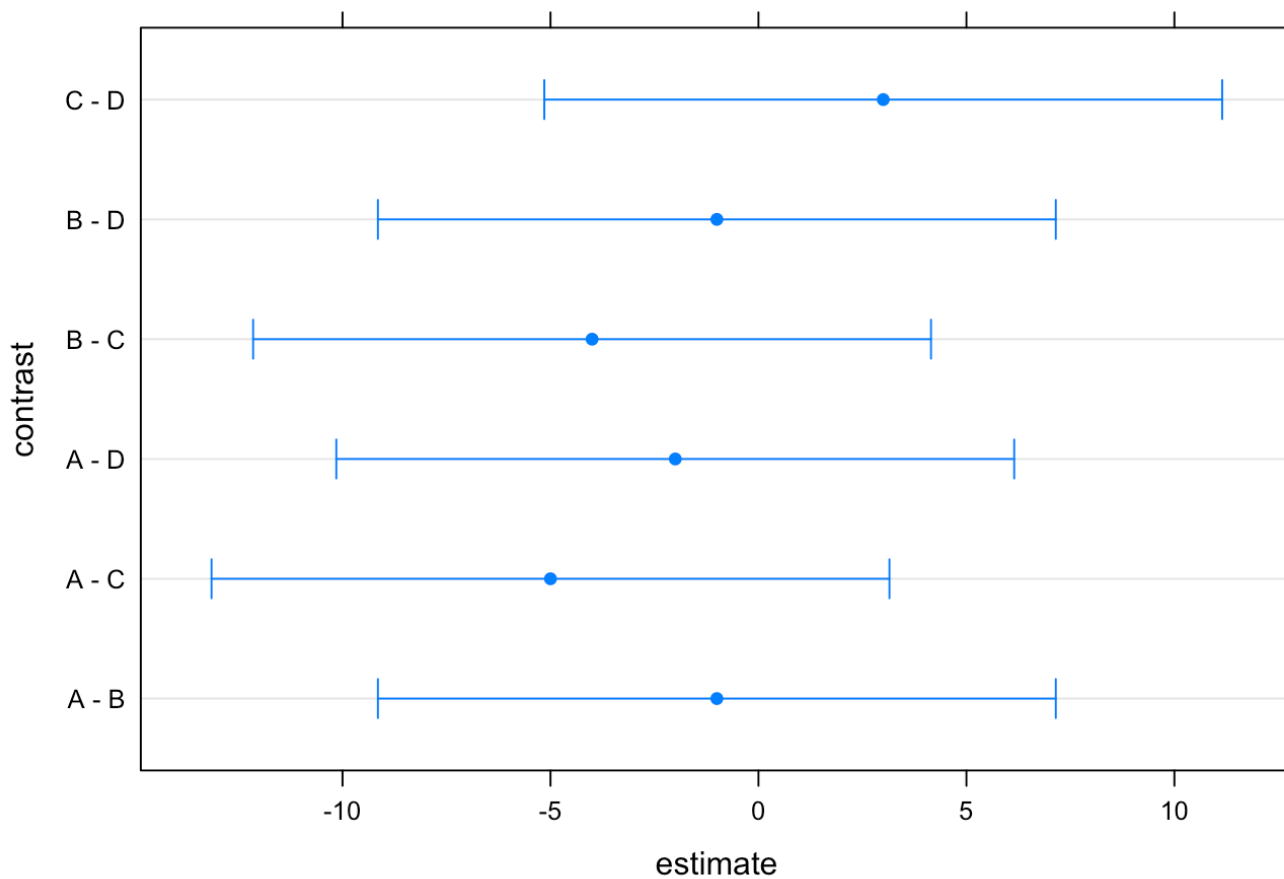


```
# lsmmeans(g, ~ treat * blend)
# lsmmeans(g, ~ treat)
(DifLS<-lsmmeans(g, pairwise ~ treat))
```

```
## $lsmmeans
## treat lsmean      SE df lower.CL upper.CL
## A      84 1.94079 12 79.77138 88.22862
## B      85 1.94079 12 80.77138 89.22862
## C      89 1.94079 12 84.77138 93.22862
## D      86 1.94079 12 81.77138 90.22862
##
## Results are averaged over the levels of: blend
## Confidence level used: 0.95
##
## $contrasts
## contrast estimate      SE df t.ratio p.value
## A - B           -1 2.744692 12  -0.364  0.9827
## A - C           -5 2.744692 12  -1.822  0.3105
```

```
## A - D      -2 2.744692 12 -0.729 0.8838
## B - C     -4 2.744692 12 -1.457 0.4905
## B - D     -1 2.744692 12 -0.364 0.9827
## C - D      3 2.744692 12  1.093 0.7002
##
## Results are averaged over the levels of: blend
## P value adjustment: tukey method for comparing a family of 4 estimates
```

```
plot(DifLS[[2]])
```

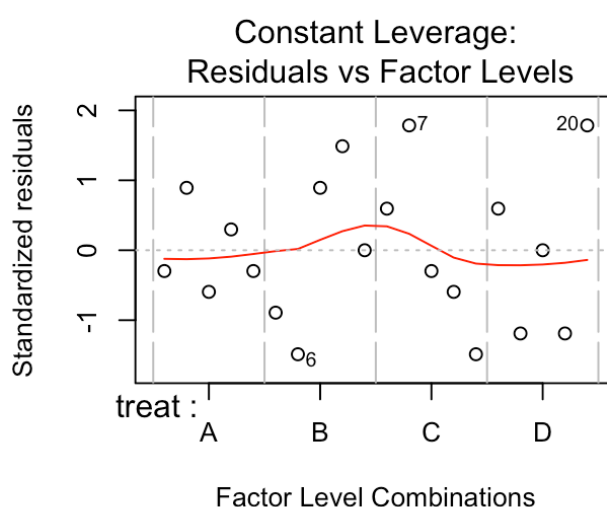
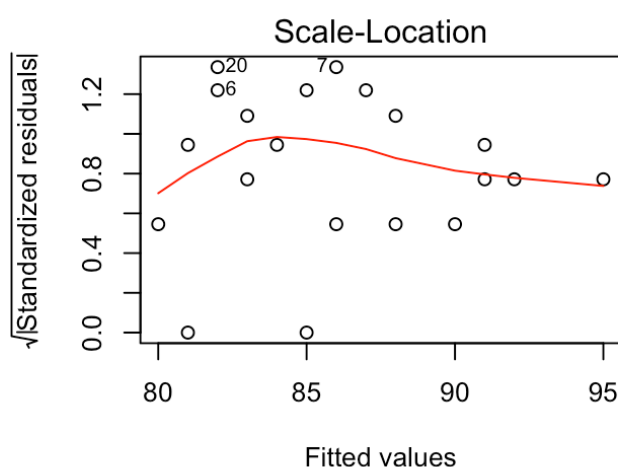
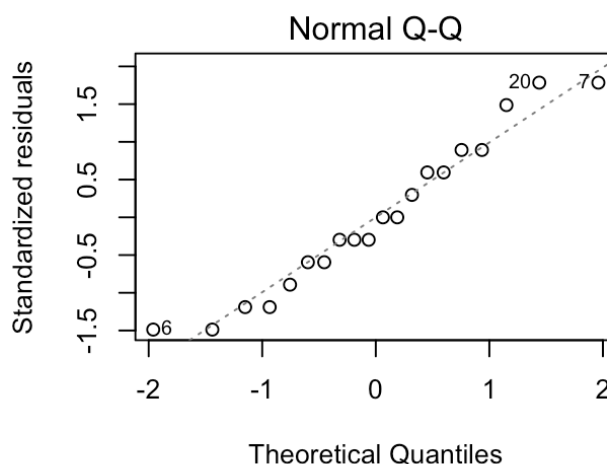
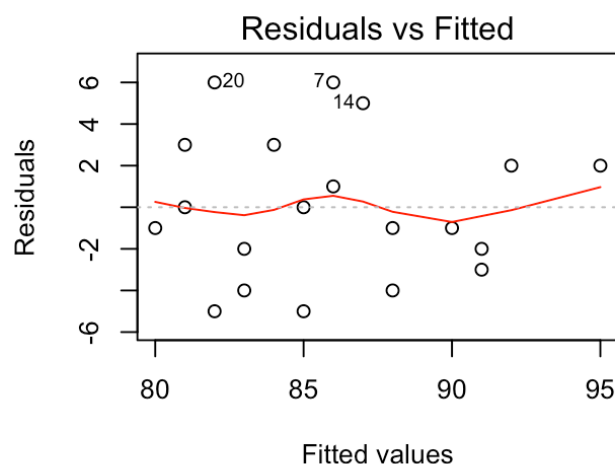


Como esperábamos, no encontramos diferencias significativas entre las medias de los tratamientos.

Diagnóstico

Comprobamos los supuestos del modelo lineal.

```
par(mfrow=c(2,2))
plot(g)
```



Debemos mirar principalmente los dos primeros gráficos. Observamos que los residuos presentan una media alrededor de 0 y que no hay grandes desviaciones, no parecen haber problemas de heterocedasticidad. En el gráfico qqplot observamos que los datos son aproximadamente normales. Sin embargo, 3 casos parecen escapar del comportamiento general, son los casos 7, 14 y 20.

5.3.2 Aplicación con R (diseño de bloques incompletos balanceado, o Balanced Incomplete Block Designs, BIBD)

Por ejemplo, si en el caso de los datos anteriores de penicilina quisiéramos comparar 6 procesos de producción con materia prima que nos permite solo 4 ensayos, deberíamos usar un diseño BIBD. En este tipo de diseños los tratamientos y bloques no son ortogonales. Los tratamientos se distribuyen en los bloques manteniendo el mismo número de réplicas por tratamiento (balanceado).

Utilizaremos un experimento de 6 tipos de dietas (A,B,C,D,E,F) y medimos el peso de una serie de conejos. Se decide que solo 3 conejos uniformes pueden seleccionarse de cada camada. Tenemos 10 camadas con

bloques de 3. Cada par de dietas aparece en el mismo bloque 2 veces.

Observamos los datos

```
data(rabbit)
head(rabbit)
```

```
##   treat gain block
## 1     f 42.2   b1
## 2     b 32.6   b1
## 3     c 35.2   b1
## 4     c 40.9   b2
## 5     a 40.1   b2
## 6     b 38.1   b2
```

```
str(rabbit)
```

```
## 'data.frame':   30 obs. of  3 variables:
## $ treat: Factor w/ 6 levels "a","b","c","d",...: 6 2 3 3 1 2 3 6 4 1 ...
## $ gain : num  42.2 32.6 35.2 40.9 40.1 38.1 34.6 34.3 37.5 44.9 ...
## $ block: Factor w/ 10 levels "b1","b10","b2",...: 1 1 1 3 3 3 4 4 4 5 ...
```

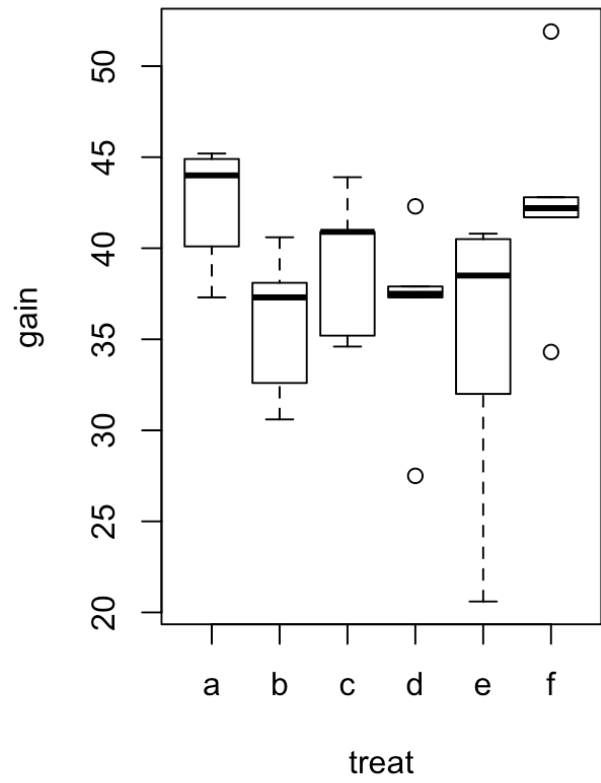
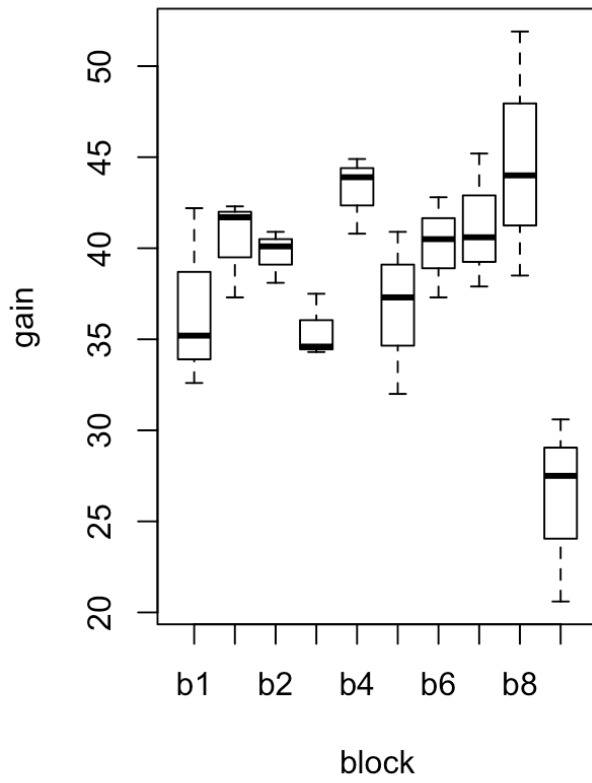
```
xtabs(gain~treat+block,data=rabbit)
```

```
##      block
## treat  b1 b10  b2  b3  b4  b5  b6  b7  b8  b9
##    a  0.0 37.3 40.1  0.0 44.9  0.0  0.0 45.2 44.0  0.0
##    b 32.6  0.0 38.1  0.0  0.0  0.0 37.3 40.6  0.0 30.6
##    c 35.2  0.0 40.9 34.6 43.9 40.9  0.0  0.0  0.0  0.0
##    d  0.0 42.3  0.0 37.5  0.0 37.3  0.0 37.9  0.0 27.5
##    e  0.0  0.0  0.0  0.0 40.8 32.0 40.5  0.0 38.5 20.6
##    f 42.2 41.7  0.0 34.3  0.0  0.0 42.8  0.0 51.9  0.0
```

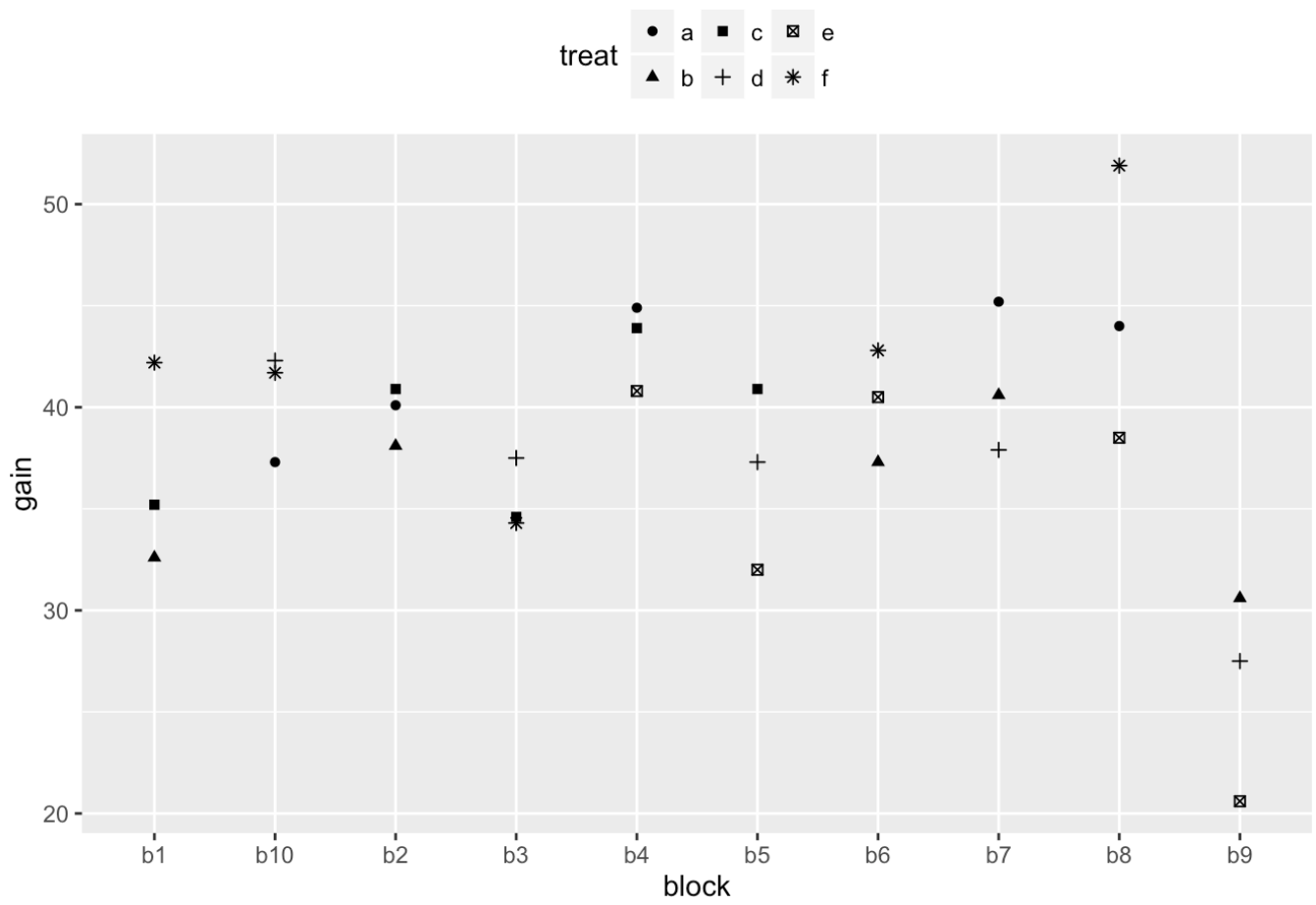
Numerosas celdas contienen ceros porque no existen datos para esa combinación de tratamiento y bloque.

Graficamos

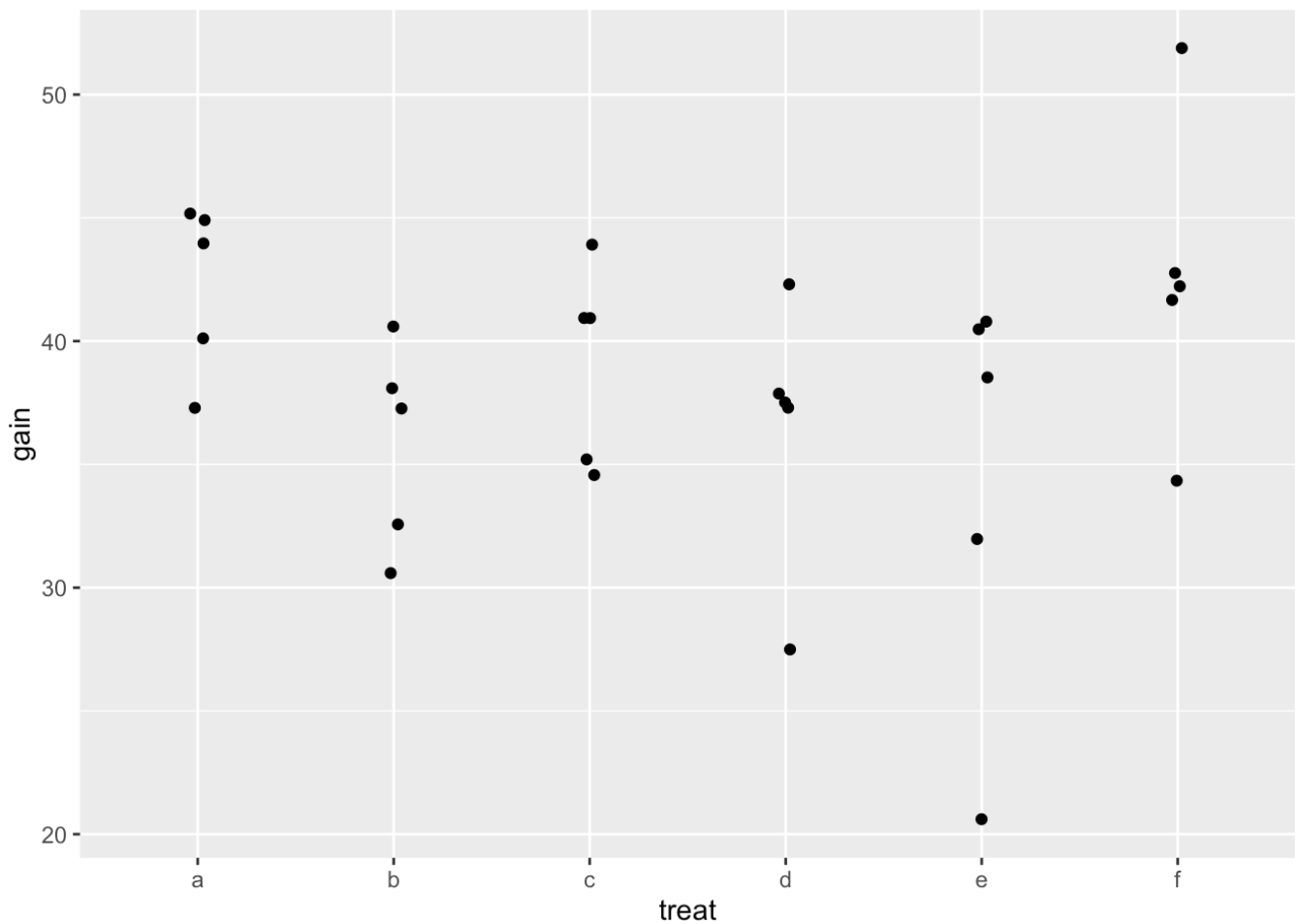
```
par(mfrow=c(1,2))
plot(gain ~ block, rabbit)
plot(gain ~ treat, rabbit)
```



```
ggplot(rabbit, aes(y = gain, x = block, shape = treat)) +
  geom_point() +
  theme(legend.position = "top", legend.direction = "horizontal")
```

```
ggplot(rabbit, aes(y = gain, x = treat)) +
  geom_point(position = position_jitter(width = 0.1))
```



Modelo y ANOVA

```
g <- lm(gain ~ block+treat,data=rabbit)
drop1(g, test = "F")
```

```
## Single term deletions
##
## Model:
## gain ~ block + treat
##      Df Sum of Sq  RSS   AIC F value    Pr(>F)
## <none>          150.77  78.437
## block   9    595.74 746.51 108.426  6.5854 0.0007602 ***
## treat   5    158.73 309.50  90.013  3.1583 0.0381655 *
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
```

```
anova(g)
```

```
## Analysis of Variance Table
##
## Response: gain
##      Df Sum Sq Mean Sq F value    Pr(>F)
## block   9  730.39   81.154   8.0738 0.0002454 ***
## treat   5  158.73   31.745   3.1583 0.0381655 *
## Residuals 15 150.77   10.052
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
```

Tanto el bloque como el tratamiento son significativos.

Comparaciones múltiples de medias post hoc

```
TukeyHSD(aov(g))$treat
```

```
##           diff      lwr      upr    p adj
## b-a -1.39333333 -7.907986  5.121320 0.97973881
## c-a  0.32000000 -6.194653  6.834653 0.99998209
## d-a  0.05333333 -6.461320  6.567986 1.00000000
## e-a -4.18000000 -10.694653  2.334653 0.34488529
## f-a  2.64000000 -3.874653  9.154653 0.77200658
## c-b  1.71333333 -4.801320  8.227986 0.95171800
## d-b  1.44666667 -5.067986  7.961320 0.97617115
## e-b -2.78666667 -9.301320  3.727986 0.73234173
## f-b  4.03333333 -2.481320 10.547986 0.38050084
## d-c -0.26666667 -6.781320  6.247986 0.99999276
## e-c -4.50000000 -11.014653  2.014653 0.27483600
## f-c  2.32000000 -4.194653  8.834653 0.84977757
## e-d -4.23333333 -10.747986  2.281320 0.33246700
## f-d  2.58666667 -3.927986  9.101320 0.78588371
## f-e  6.82000000  0.305347 13.334653 0.03766284
```

Solo la diferencia f-e fue significativa.

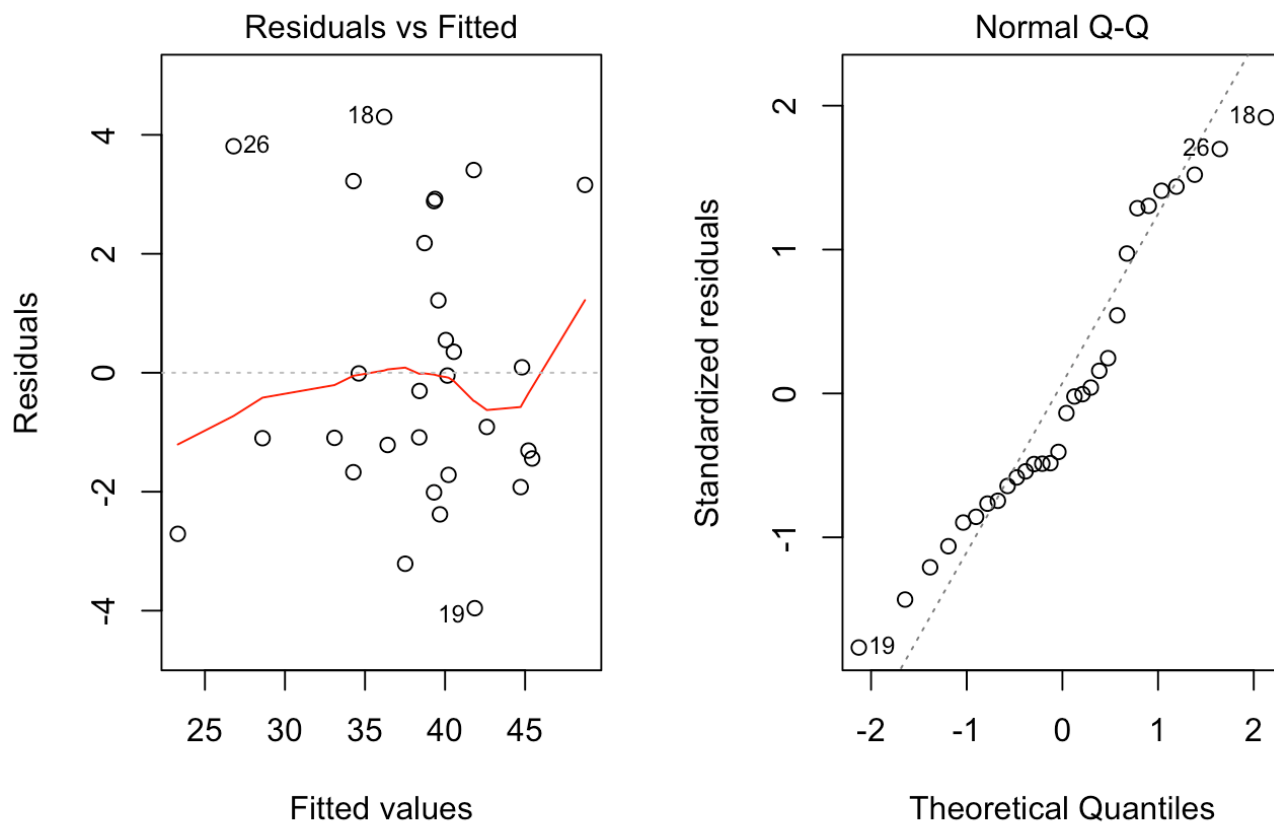
Coeficientes del modelo

```
summary(g)
```

```
##
## Call:
## lm(formula = gain ~ block + treat, data = rabbit)
##
## Residuals:
##      Min       1Q   Median       3Q      Max
## -3.9583 -1.6146 -0.6083  1.9396  4.3028
##
## Coefficients:
##              Estimate Std. Error t value Pr(>|t|)
## (Intercept)  36.01389    2.58863   13.912 5.59e-10 ***
## blockb10     3.29722    2.79604    1.179  0.25667
## blockb2      4.13333    2.69433    1.534  0.14583
## blockb3     -1.80278    2.69433   -0.669  0.51360
## blockb4      8.79444    2.79604    3.145  0.00667 **
## blockb5      2.30556    2.79604    0.825  0.42253
## blockb6      5.40833    2.69433    2.007  0.06309 .
## blockb7      5.77778    2.79604    2.066  0.05651 .
## blockb8      9.42778    2.79604    3.372  0.00419 **
## blockb9     -7.48056    2.79604   -2.675  0.01729 *
## treatb       -1.74167    2.24182   -0.777  0.44930
## treatc        0.40000    2.24182    0.178  0.86078
## treatd        0.06667    2.24182    0.030  0.97667
## treate       -5.22500    2.24182   -2.331  0.03413 *
## treatf        3.30000    2.24182    1.472  0.16169
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## Residual standard error: 3.17 on 15 degrees of freedom
## Multiple R-squared:  0.855, Adjusted R-squared:  0.7197
## F-statistic: 6.318 on 14 and 15 DF, p-value: 0.000518
```

Diagnóstico

```
par(mfrow=c(1,2))
plot(g, which = 1:2)
```



Vemos algunos casos que se escapan del comportamiento del resto y una posible falta de homocedasticidad y normalidad.

5.4 Diseño bi-factorial completamente aleatorizado

ANOVA de 2-vías (dos factores tratamiento)

Se considera un diseño completamente aleatorizado con dos factores tratamiento cuyos niveles se **cruzan**.

Modelo

Si se tienen K unidades experimentales en cada combinación de los dos factores tratamiento, con I y J niveles en cada factor (A y B), el modelo matemático del diseño es:

$$y_{ijk} = \mu + \alpha_i + \beta_j + (\alpha\beta)_{ij} + \varepsilon_{ijk}$$

donde y_{ijk} es el resultado del tratamiento i -ésimo del factor A y del tratamiento j -ésimo del factor B en la replicación k -ésima, μ representa el efecto global de y que mide el nivel medio de todos los resultados, α_i representa el efecto del tratamiento i sobre la respuesta, β_j es el efecto del tratamiento j sobre la

respuesta, $(\alpha\beta)_{ij}$ representa la interacción o el efecto “extra” sobre la respuesta debido a que se observan conjuntamente los niveles i y j de los factores 1 y 2 respectivamente, y ε_{ij} es una variable aleatoria que representa el componente aleatorio de error o error experimental o perturbación (iid, normales con media cero y varianza σ^2).

A este modelo se lo conoce también como *modelo completo de dos vías*. Se trata de un diseño **completo o equilibrado**, es decir, un diseño que se obtiene cruzando los niveles de los factores en todas las formas posibles (también llamados **diseños ortogonales**). La ortogonalidad permite separar los efectos de cada uno de los factores sobre la variable de interés.

Hipótesis

Los parámetros α_i y β_j se llaman **efectos principales**, mientras que $(\alpha\beta)_{ij}$ representa los **efectos de interacción**.

Efectos principales de cada factor

Nos permite evaluar la contribución individual de cada factor tratamiento sobre la respuesta observada. Se les denomina también *efectos marginales*. Sin embargo, cuando el efecto interacción entre ambos factores es significativa, se vuelve complejo examinar por separado cada uno de los efectos principales.

Factor A

- ⌘ H_0 Las I medias poblacionales correspondientes a los I niveles del factor A son iguales o, lo que es lo mismo, no existe efecto del primer factor.
- ⌘ H_1 el factor A influye o afecta a la respuesta (variable dependiente)

Factor B

- ⌘ H_0 Las J medias poblacionales correspondientes a los J niveles del factor B son iguales o, lo que es lo mismo, no existe efecto del segundo factor
- ⌘ H_1 el factor B influye o afecta a la respuesta (variable dependiente)

Efectos de interacción entre factores

Generalmente lo que nos interesa en estas pruebas es encontrar una interacción significativa entre los factores. Desde un punto de vista no formal, decimos que existe una interacción significativa entre dos factores cuando el efecto de uno de ellos sobre la variable dependiente no es el mismo en todos los niveles del otro factor.

- ⌘ H_0 no existe interacción
- ⌘ H_1 hay interacción

⌘ Gráfico de interacción

Nos permite observar la existencia de interacción pero hay que considerarlos con cuidado ya que no reflejan el tamaño del error experimental.

⌘ Test de Tukey

Ya lo hemos utilizado en el diseño por bloques completamente aleatorizado para contrastar si existe una interacción entre el factor tratamiento y el factor bloque. Ahora lo utilizaremos para contrastar la hipótesis de que la interacción entre tratamientos es significativa.

Se dice que *no existe interacción* entre los dos factores tratamiento cuando sus efectos sobre la respuesta son *aditivos*. Es decir, la diferencia de las respuestas medias teóricas en dos niveles cualesquiera de un factor es constante en todos los niveles del otro factor y viceversa.

5.4.1 Aplicación con R

Los datos corresponden a una investigación sobre agentes tóxicos. Se asignaron 3 tipos de veneno (I, II y III) y 4 tipos de tratamientos (A,B,C,D) a 48 ratas. Se midió el tiempo de supervivencia en decenas de horas. En este caso, tenemos un diseño entre-sujetos factorial 3x4, con los factores veneno (poison) y tratamientos (treat).

Observamos los datos

```
data(rats, package="faraway")
head(rats)
```

```
##   time poison treat
## 1 0.31      I     A
## 2 0.82      I     B
## 3 0.43      I     C
## 4 0.45      I     D
## 5 0.45      I     A
## 6 1.10      I     B
```

```
str(rats)
```

```
## 'data.frame':   48 obs. of  3 variables:
## $ time : num  0.31 0.82 0.43 0.45 0.45 1.1 0.45 0.71 0.46 0.88 ...
## $ poison: Factor w/ 3 levels "I","II","III": 1 1 1 1 1 1 1 1 1 1 ...
## $ treat : Factor w/ 4 levels "A","B","C","D": 1 2 3 4 1 2 3 4 1 2 ...
```

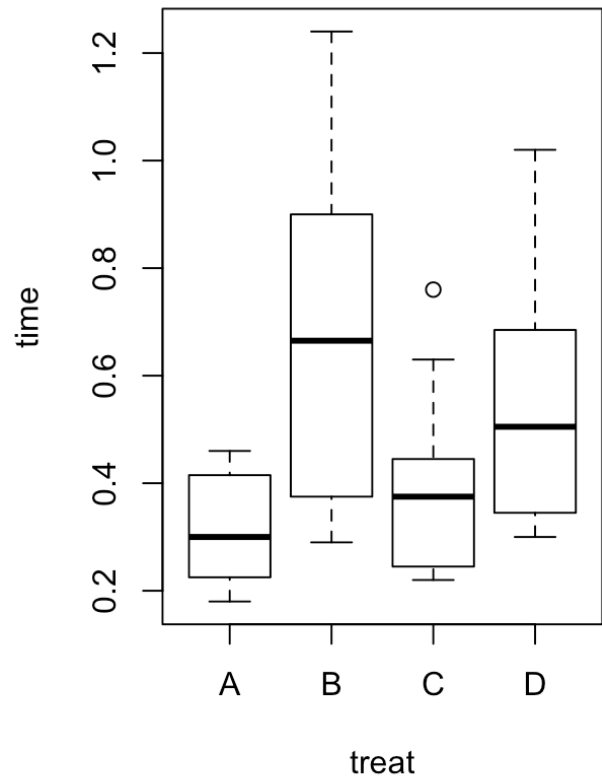
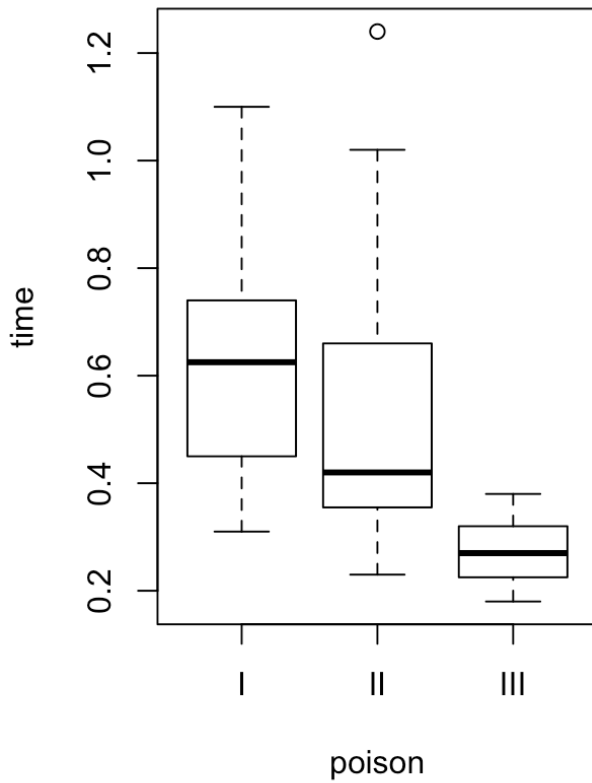
```
xtabs(time~treat+poison,data=rats)
```

```
##      poison
## treat   I   II  III
##    A 1.65 1.28 0.84
##    B 3.52 3.26 1.34
##    C 2.27 1.50 0.94
##    D 2.44 2.67 1.30
```

Encontramos todas las combinaciones de ambos factores y hay una única observación por combinación.

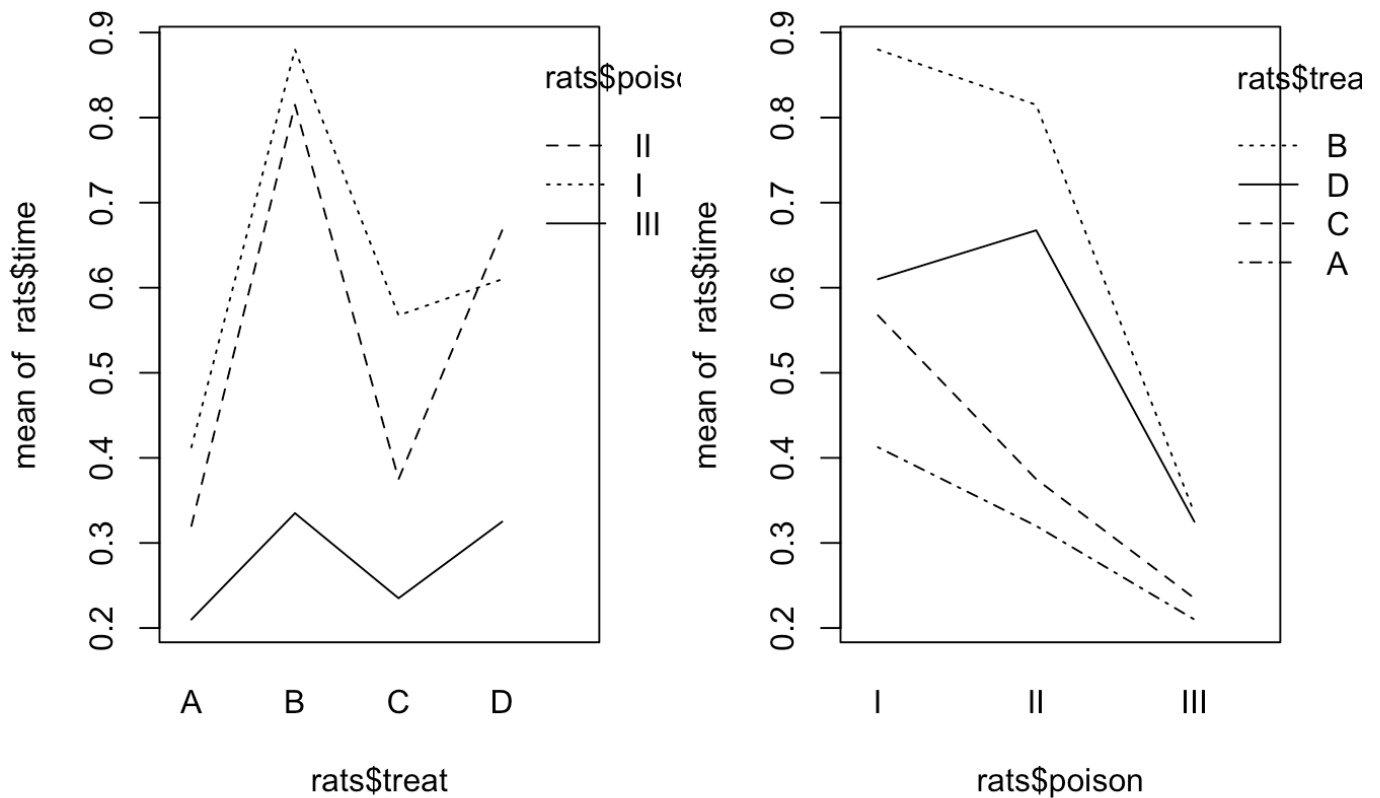
Graficamos

```
par(mfrow=c(1,2))
plot(time ~ poison, data=rats)
plot(time ~ treat, data=rats)
```



Evaluamos la **interacción de los factores**.

```
par(mfrow=c(1,2))
interaction.plot(rats$treat,rats$poison,rats$time)
interaction.plot(rats$poison,rats$treat,rats$time)
```



No observamos líneas paralelas por lo que parece existir interacción.

Modelo y ANOVA

En el siguiente ejemplo observa que las dos variables independientes están separados por un asterisco que le indica a R que queremos analizar la interacción entre ambas. Si únicamente quisiéramos analizar los dos efectos principales, tendríamos de reemplazar el asterisco con el signo de más (+).

```
g <- lm(time ~ poison*treat, rats)
anova(g)
```

```
## Analysis of Variance Table
##
## Response: time
##          Df Sum Sq Mean Sq F value    Pr(>F)
## poison     2  1.03301  0.51651  23.2217 3.331e-07 ***
## treat      3  0.92121  0.30707  13.8056 3.777e-06 ***
## poison:treat 6  0.25014  0.04169   1.8743  0.1123
## Residuals  36  0.80073  0.02224
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
```

Ambos efectos principales son significativos, pero la interacción no lo es.

NOTA: es equivalente a realizar:

```
summary(aov(time ~ poison*treat, rats))
```



```
##           Df Sum Sq Mean Sq F value    Pr(>F)
## poison      2  1.0330   0.5165   23.222 3.33e-07 ***
## treat       3  0.9212   0.3071   13.806 3.78e-06 ***
## poison:treat 6  0.2501   0.0417    1.874  0.112
## Residuals   36  0.8007   0.0222
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
```

Es posible obtener las medias marginales de poison y de treat.

```
#Vamos a ver Las medias por condición.
attach(rats)
tapply(time,list(poison,treat),mean)
```

```
##           A      B      C      D
## I    0.4125 0.880 0.5675 0.6100
## II   0.3200 0.815 0.3750 0.6675
## III  0.2100 0.335 0.2350 0.3250
```

```
# Para tener Las medias marginales de poison
# (independientemente de treat) has de emplear:

tapply(time,poison,mean)
```

```
##           I      II      III
## 0.617500 0.544375 0.276250
```

```
#Y para obtener Las medias marginales de treat, el comando es:

tapply(time, treat,mean)
```

```
##           A      B      C      D
## 0.3141667 0.6766667 0.3925000 0.5341667
```

Comparaciones múltiples de medias post hoc

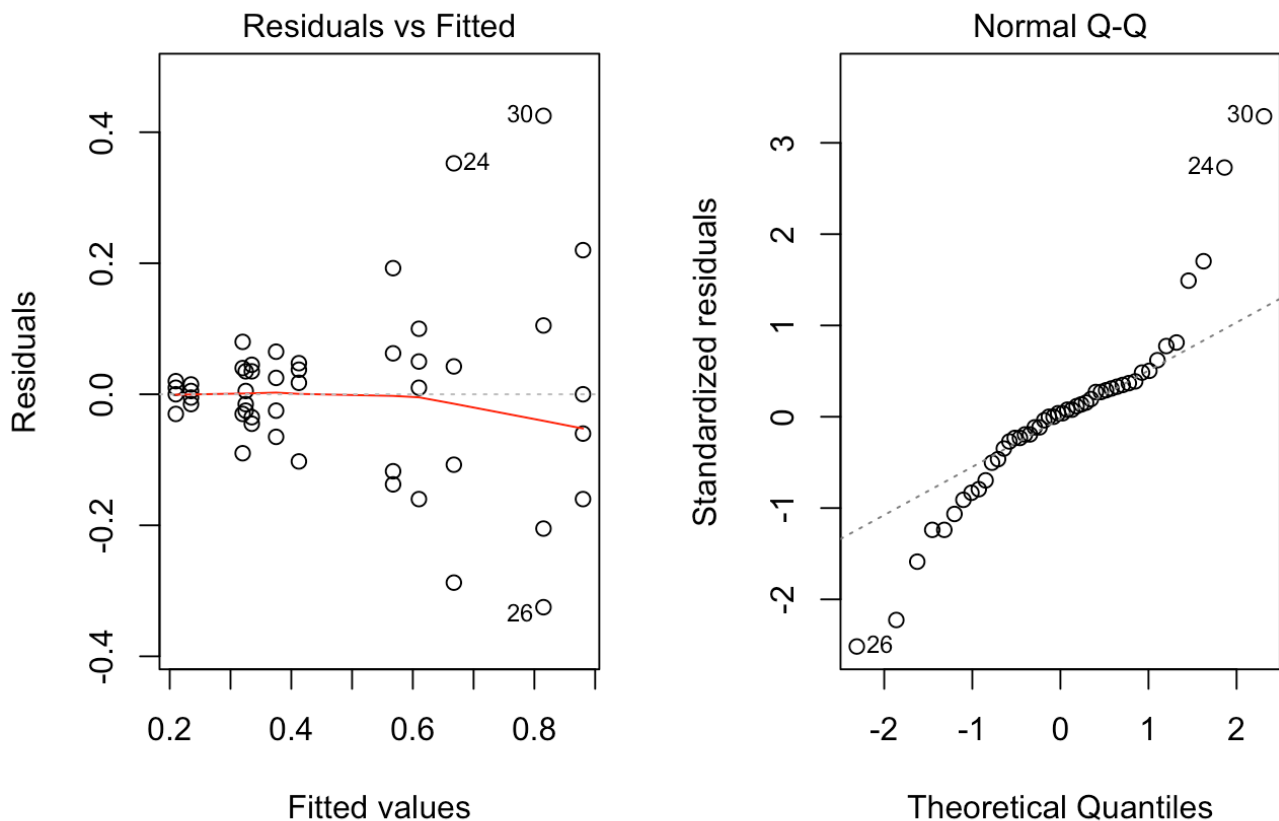
```
round(TukeyHSD(aov(g))$treat,3)
```

```
##           diff      lwr      upr p adj
## B-A    0.362    0.199    0.526 0.000
## C-A    0.078   -0.086    0.242 0.577
## D-A    0.220    0.056    0.384 0.005
## C-B   -0.284   -0.448   -0.120 0.000
## D-B   -0.142   -0.306    0.021 0.108
## D-C    0.142   -0.022    0.306 0.111
```

Las diferencias B-A, D-A y C-B fueron significativas.

Diagnóstico

```
par(mfrow=c(1,2))
plot(g, which = 1:2)
```



Los gráficos no son normales y hay heterocedasticidad, podríamos transformarlos con el logaritmo o la función recíproca.

Naturalmente, es fácil generalizar lo que hemos visto para el caso de 3 o más factores entre-grupos.

5.5 Diseño de fracciones factoriales. Cuadrados latinos

Diseño de fracciones factoriales

Los modelos de *fracciones factoriales* son útiles cuando queremos estudiar la influencia de los factores sin necesidad de realizar todas las pruebas de interacción pero manteniendo la propiedad de *ortogonalidad* de los efectos a estudiar.

En particular, el modelo *cuadrado latino* ocurre cuando tenemos dos factores de bloque y un factor tratamiento, donde cada tratamiento se asigna solo una vez a cada bloque de manera aleatoria. Todos los factores tienen K niveles, pero se necesitan K^2 observaciones en lugar de las K^3 observaciones del diseño completo de 3 factores.

El diseño de cuadrado latino se puede utilizar si se verifican las siguientes condiciones:

- ⌘ Es un diseño de experimentos con tres factores.
- ⌘ Los tres factores tienen el mismo número de niveles K .

- ⚡ No hay interacciones entre los tres factores.

Modelo

El modelo de ANOVA puede describirse de la siguiente manera:

$$y_{ij(k)} = \mu + \alpha_i + \beta_j + \phi_k + \varepsilon_{ij(k)} \text{ con } i, j, k = 1 \dots K$$

donde $y_{ij(k)}$ es el resultado del bloque i -ésimo del factor bloque 1, del bloque j -ésimo del factor bloque 2 y del nivel k -ésimo del factor tratamiento 3. Denotamos k entre paréntesis para indicar que este índice no se elige sino que viene condicionado por el par ij . Además, μ corresponde al efecto global que mide el nivel medio de todos los resultados, α_i es el efecto sobre la media global debido al bloque i del factor bloque 1, β_j es el efecto sobre la media global debido al bloque j del factor bloque 2, ϕ_k es el efecto sobre la media global debido al nivel k del factor tratamiento 3, y $\varepsilon_{ij(k)}$ es el error experimental (iid, con distribución normal de media 0 y varianza σ^2).

5.5.1 Aplicación con R

Un ingeniero quiere comparar las cualidades de las materias primas a partir de cuatro proveedores, A, B, C, D. La materia prima se utiliza para producir un componente del cual mediremos su resistencia a la rotura. Se necesita un operador un día entero para hacer uno de los componentes y hay 4 operadores y 4 días en los que se llevará a cabo el experimento. El diseño cuadrado latino es apropiado en este caso en que el operador y el día son los efectos de bloque.

Observamos los datos

```
data(breaking)
head(breaking)
```

```
##      y operator  day supplier
## 1  810      op1 day1         B
## 2 1080      op1 day2         C
## 3   700      op1 day3         A
## 4   910      op1 day4         D
## 5 1100      op2 day1         C
## 6   880      op2 day2         D
```

```
str(breaking)
```

```
## 'data.frame':  16 obs. of  4 variables:
## $ y      : num  810 1080 700 910 1100 880 780 600 840 540 ...
## $ operator: Factor w/ 4 levels "op1","op2","op3",...: 1 1 1 1 2 2 2 2 3 3 ...
## $ day      : Factor w/ 4 levels "day1","day2",...: 1 2 3 4 1 2 3 4 1 2 ...
## $ supplier: Factor w/ 4 levels "A","B","C","D": 2 3 1 4 3 4 2 1 4 1 ...
```

```
#comprobamos la estructura de cuadrado latino
xtabs(y~operator+day+supplier,data=breaking)
```

```
## , , supplier = A
##
##      day
```

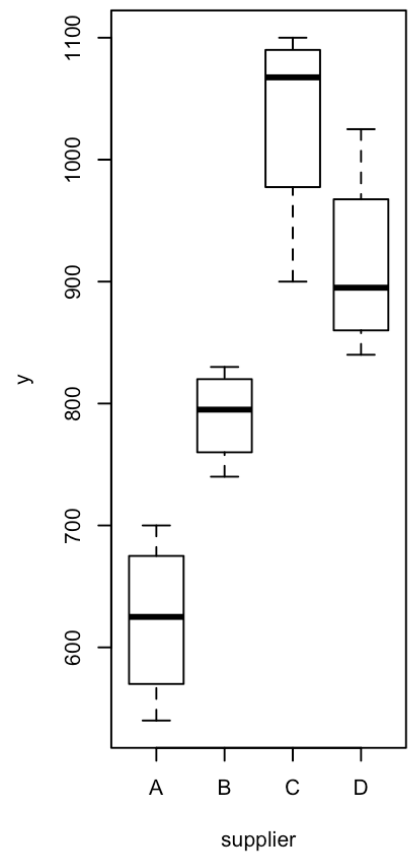
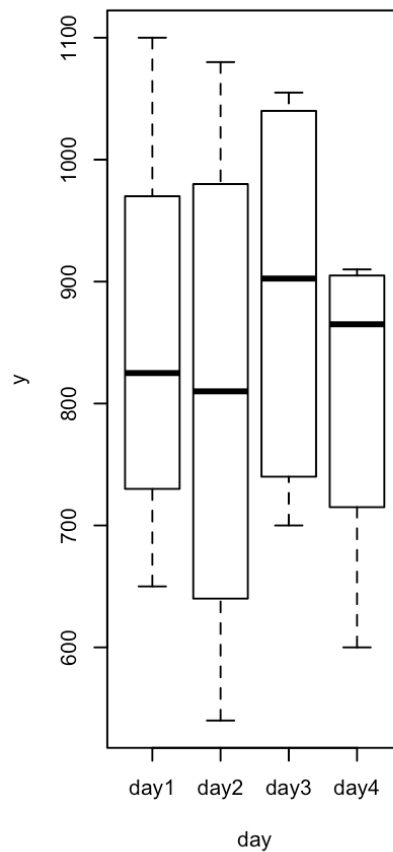
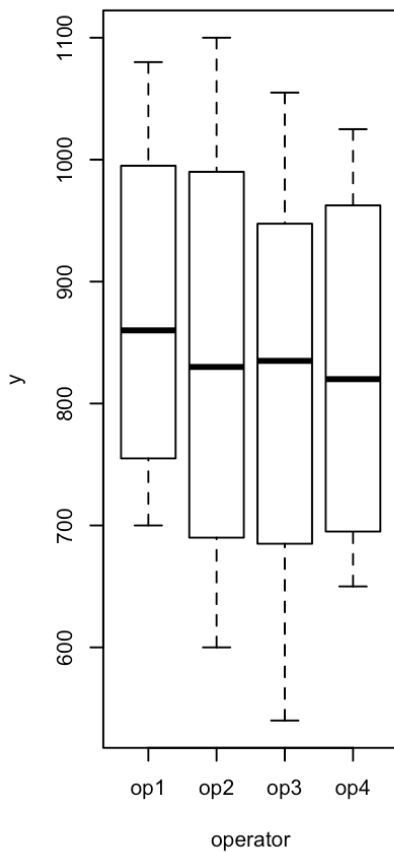
```
## operator day1 day2 day3 day4
##      op1    0    0 700    0
##      op2    0    0    0 600
##      op3    0 540    0    0
##      op4 650    0    0    0
##
## , , supplier = B
##
##      day
## operator day1 day2 day3 day4
##      op1 810    0    0    0
##      op2  0    0 780    0
##      op3  0    0    0 830
##      op4  0 740    0    0
##
## , , supplier = C
##
##      day
## operator day1 day2 day3 day4
##      op1  0 1080    0    0
##      op2 1100    0    0    0
##      op3  0    0 1055    0
##      op4  0    0    0 900
##
## , , supplier = D
##
##      day
## operator day1 day2 day3 day4
##      op1  0    0    0 910
##      op2  0 880    0    0
##      op3 840    0    0    0
##      op4  0    0 1025    0
```

```
matrix(breaking$supplier,4,4)
```

```
##      [,1] [,2] [,3] [,4]
## [1,] "B"  "C"  "D"  "A"
## [2,] "C"  "D"  "A"  "B"
## [3,] "A"  "B"  "C"  "D"
## [4,] "D"  "A"  "B"  "C"
```

Graficamos

```
par(mfrow=c(1,3))
plot(y ~ operator , breaking)
plot(y ~ day , breaking)
plot(y ~ supplier, breaking)
```



Parece haber diferencias en los proveedores pero no en las variables de bloque (operator y day). No se detectan outliers, asimetrías ni varianzas desiguales.

Modelo y ANOVA

```
g <- lm(y ~ operator + day + supplier, breaking)
anova(g)
```

```
## Analysis of Variance Table
##
## Response: y
##           Df Sum Sq Mean Sq F value    Pr(>F)
## operator   3   7663    2554   0.4114 0.750967
## day        3  17600    5867   0.9450 0.475896
## supplier   3 371138  123713 19.9268 0.001602 **
## Residuals  6   37250     6208
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
```

Existe un efecto del proveedor pero no del operador o del día.

Comparaciones múltiples de medias post hoc

Realizamos comparaciones pareadas de Tukey para la interpretación.

```
TukeyHSD(aov(g))$supplier
```

```
##           diff      lwr      upr      p adj
## B-A  167.50  -25.36942 360.36942 0.085385140
## C-A  411.25  218.38058 604.11942 0.001306761
```

```
## D-A 291.25 98.38058 484.11942 0.007838589
## C-B 243.75 50.88058 436.61942 0.018290285
## D-B 123.75 -69.11942 316.61942 0.219674691
## D-C -120.00 -312.86942 72.86942 0.237951905
```

```
#verificar
# scoefs <- c(0,g$coef[8:10])
# outer(scoefs,scoefs,"-")
```

Vemos que las diferencias B-A, D-B y D-C no son significativas.

Estimaciones de los efectos

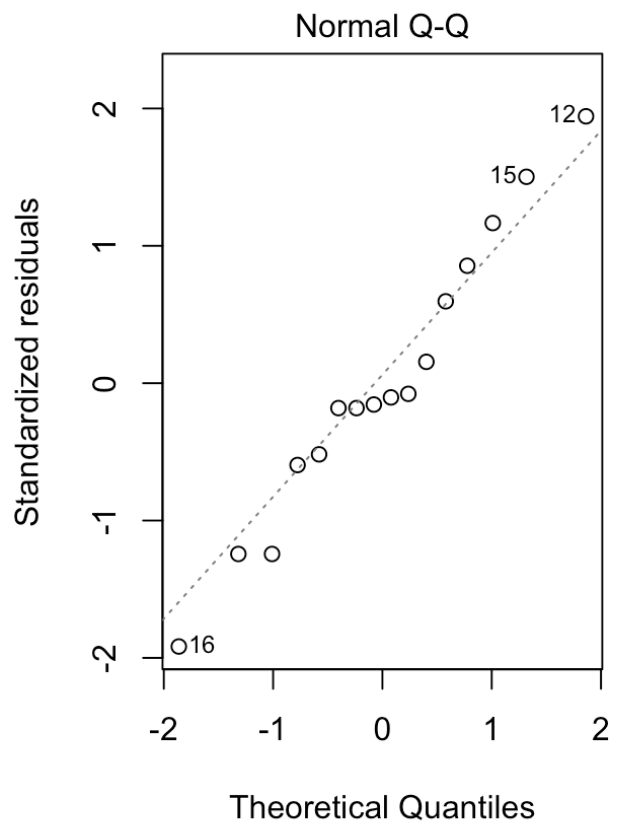
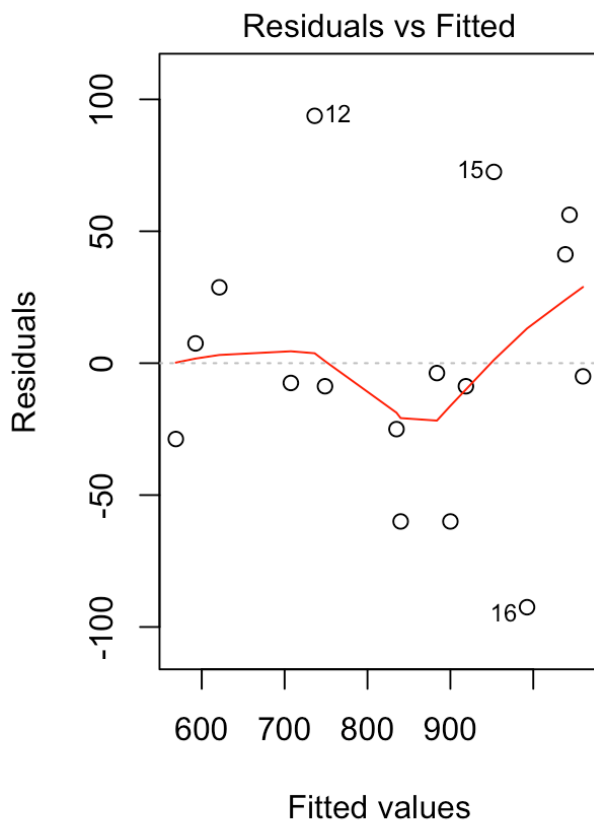
```
summary(g)
```

```
##
## Call:
## lm(formula = y ~ operator + day + supplier, data = breaking)
##
## Residuals:
##      Min       1Q   Median       3Q      Max
## -92.50 -25.94  -6.25   31.88   93.75
##
## Coefficients:
##              Estimate Std. Error t value Pr(>|t|)
## (Intercept)   667.50     62.29   10.716 3.9e-05 ***
## operatorop2   -35.00     55.72   -0.628 0.553020
## operatorop3   -58.75     55.72   -1.054 0.332266
## operatorop4   -46.25     55.72   -0.830 0.438247
## dayday2       -40.00     55.72   -0.718 0.499782
## dayday3        40.00     55.72    0.718 0.499782
## dayday4       -40.00     55.72   -0.718 0.499782
## supplierB     167.50     55.72    3.006 0.023812 *
## supplierC     411.25     55.72    7.381 0.000317 ***
## supplierD     291.25     55.72    5.227 0.001962 **
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## Residual standard error: 78.79 on 6 degrees of freedom
## Multiple R-squared:  0.9141, Adjusted R-squared:  0.7853
## F-statistic: 7.094 on 9 and 6 DF, p-value: 0.01348
```

La calidad de la materia prima es mejor para el proveedor C seguido del D.

Diagnóstico

```
par(mfrow=c(1,2))
plot(g,which=1:2)
```



Detectamos algunos outliers y falta de normalidad.

5.6 Análisis de covarianza (ANCOVA)

Todos los diseños que se han presentado hasta ahora ya sean entre-sujetos o intra-sujetos, uni-factorial o multi-factorial, aleatorizados o con restricciones por bloques, pueden incluir alguna **covariable** (cuantitativa continua) para mejorar la precisión del modelo además de las variables controlables o factores (cualitativas).

Llamaremos **ANCOVA** a aquellos diseños que incluyen una variable dependiente y alguna covariable. Combina las características del Análisis de Regresión y del Análisis de la Varianza.

Las covariables no suelen ser variables directamente sometidas a investigación sino que corresponden a variables cuyos efectos se desean controlar. La selección de la covariable se debe a su alta correlación con la variable dependiente (algunos proponen que la correlación sea mayor a 0.4). En el caso de incluir más de una covariable se mantiene el mismo criterio pero además las correlación entre covariables debe ser nula o muy escasa.

Ejemplo: si estudiamos los efectos de la herbivoría en la polinización controlando por los efectos del tamaño de la planta, o si estudiamos el efecto de distintos tipos de fertilizantes sobre el tamaño de los árboles controlado el efecto de la edad del árbol.

5.6.1 Aplicación en R

Utilizaremos los datos de un estudio sobre los efectos del abuso sexual infantil en mujeres adultas. Los datos fueron publicados por Rodríguez et al. 1997 y están formados por 45 mujeres que han sido tratadas en una clínica a las cuales se les preguntó si habían sido abusadas sexualmente en su niñez (csa) y a las que se les midió el estrés post-traumático (ptsd) y el tipo de abuso físico (cpa) en escalas estandarizadas. 31 de las mujeres no presentaron abusos sexuales en la infancia. Queremos explicar los altos valores de ptsd según los dos tipos de abuso sufridos.

Observamos los datos

```
data(sexab)
head(sexab)
```

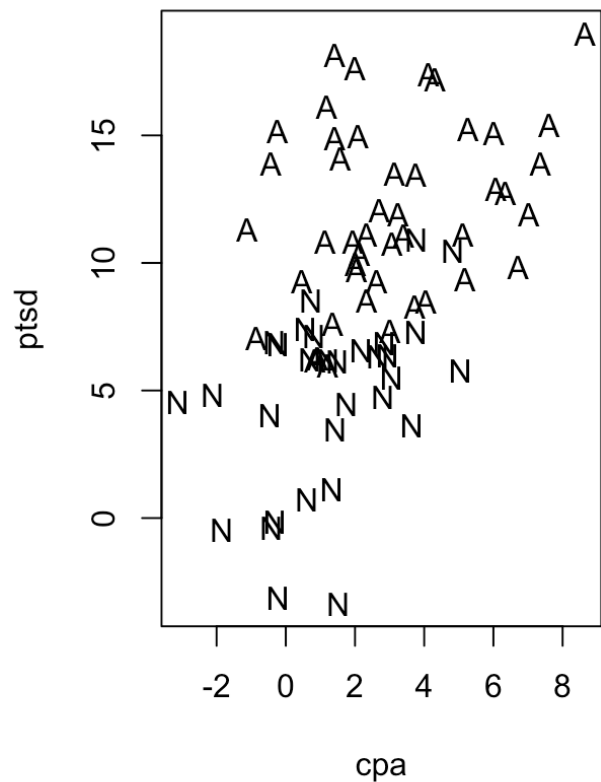
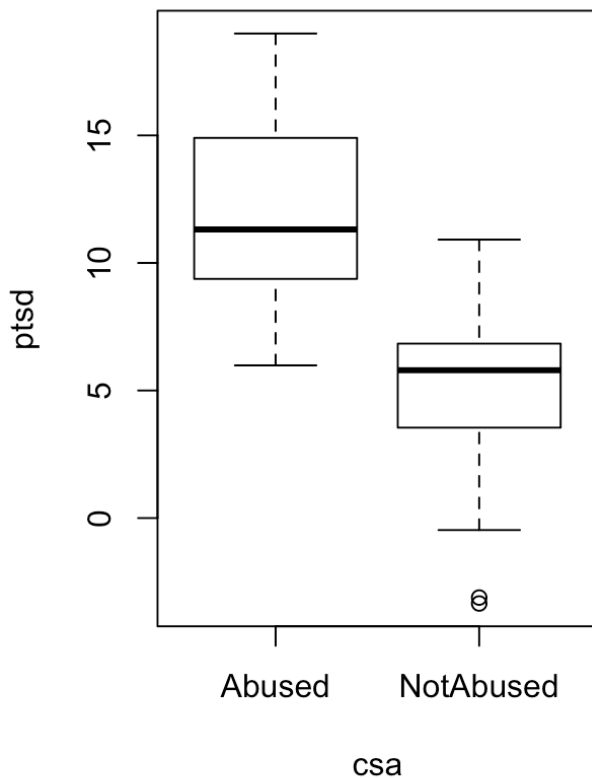
```
##      cpa    ptsd   csa
## 1  2.04786  9.71365 Abused
## 2  0.83895  6.16933 Abused
## 3 -0.24139 15.15926 Abused
## 4 -1.11461 11.31277 Abused
## 5  2.01468  9.95384 Abused
## 6  6.71131  9.83884 Abused
```

```
str(sexab)
```

```
## 'data.frame':   76 obs. of  3 variables:
## $ cpa : num  2.048 0.839 -0.241 -1.115 2.015 ...
## $ ptsd: num  9.71 6.17 15.16 11.31 9.95 ...
## $ csa : Factor w/ 2 levels "Abused","NotAbused": 1 1 1 1 1 1 1 1 1 ...
```

Graficamos

```
par(mfrow=c(1,2))
plot (ptsd~ csa, sexab)
plot (ptsd~ cpa, pch=as.character (csa), sexab)
```

Vemos que las mujeres con problemas de abuso (A) presentan altos niveles de ptsd respecto a las mujeres que no han sido abusadas sexualmente en su infancia (N). Sin embargo, también observamos que aquellas mujeres que sufrieron abuso sexual de niñas también presentaron abuso físico. Esto sugiere que el abuso físico es una explicación alternativa a los altos valores de ptsd en el grupo de abuso sexual.

El ANCOVA nos permite evaluar estas dos explicaciones.

Modelo y ANOVA

```
g <- lm(ptsd ~ cpa*csa, sexab)
summary(g)
```

```
##
## Call:
## lm(formula = ptsd ~ cpa * csa, data = sexab)
##
## Residuals:
##      Min       1Q   Median       3Q      Max
## -8.1999 -2.5313 -0.1807  2.7744  6.9748
##
## Coefficients:
##              Estimate Std. Error t value Pr(>|t|)
## (Intercept)    10.5571     0.8063   13.094 < 2e-16 ***
## cpa             0.4500     0.2085    2.159  0.0342 *
## csaNotAbused   -6.8612     1.0747   -6.384 1.48e-08 ***
## cpa:csaNotAbused  0.3140     0.3685    0.852  0.3970
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## Residual standard error: 3.279 on 72 degrees of freedom
## Multiple R-squared:  0.5828, Adjusted R-squared:  0.5654
## F-statistic: 33.53 on 3 and 72 DF, p-value: 1.133e-13
```

```
#model.matrix(g) para ver la codificación
```

Vemos que codifica “Abused=0” (nivel de referencia) y “NotAbused=1”, ya que utiliza el orden alfabético. Si quisiéramos cambiar el orden podríamos utilizar “`sexabcsa < -relevel(sexabcsa, ref="NotAbused")`”. También observamos que la interacción no es significativa (cpa: csaNotAbused) por lo que podemos simplificar el modelo:

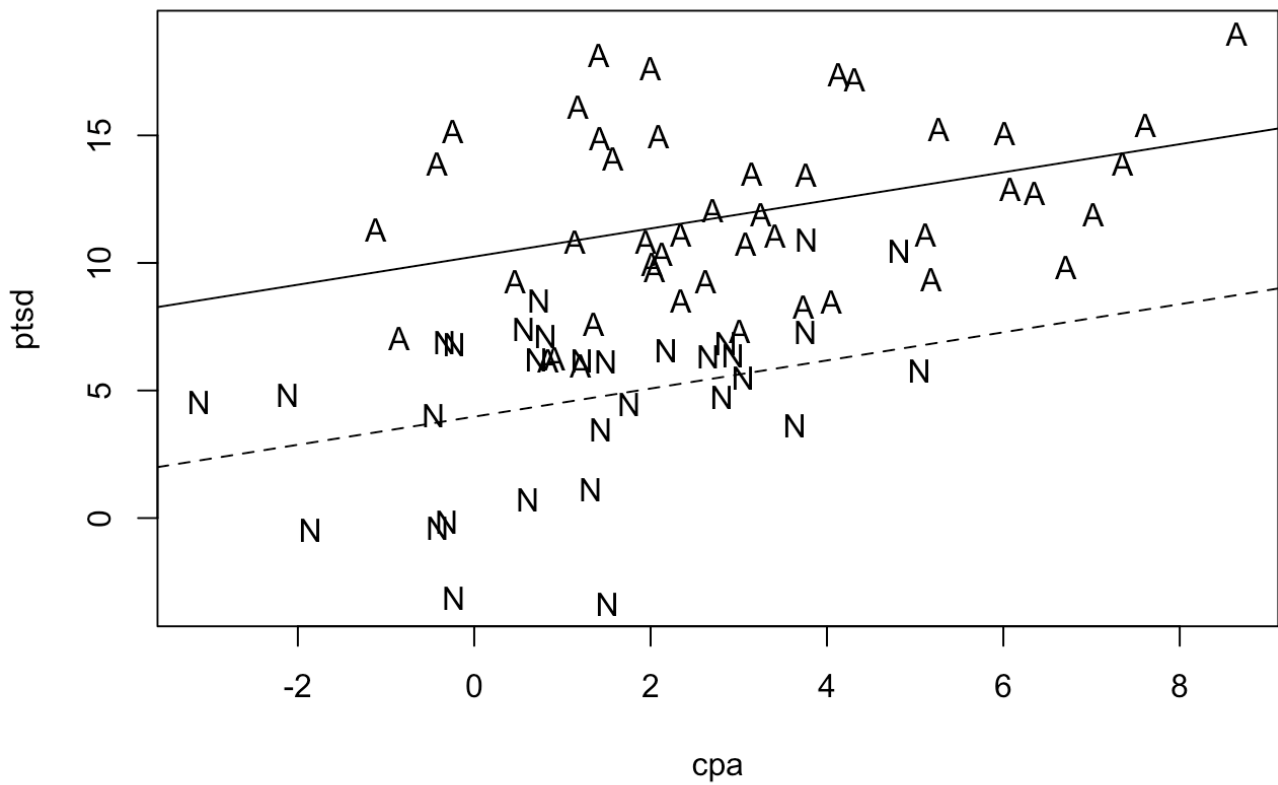
```
g <-lm(ptsd ~ cpa+csa, sexab)
summary(g)
```

```
##
## Call:
## lm(formula = ptsd ~ cpa + csa, data = sexab)
##
## Residuals:
##      Min       1Q   Median       3Q      Max
## -8.1567 -2.3643 -0.1533  2.1466  7.1417
##
## Coefficients:
##              Estimate Std. Error t value Pr(>|t|)
## (Intercept)   10.2480     0.7187  14.260 < 2e-16 ***
## cpa             0.5506     0.1716   3.209  0.00198 **
## csaNotAbused  -6.2728     0.8219  -7.632 6.91e-11 ***
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## Residual standard error: 3.273 on 73 degrees of freedom
## Multiple R-squared:  0.5786, Adjusted R-squared:  0.5671
## F-statistic: 50.12 on 2 and 73 DF,  p-value: 2.002e-14
```

Todos los términos (efectos principales) son significativos.

Para graficar el modelo debemos recordar que al no haber interacción se puede pensar en el resultado como dos rectas paralelas:

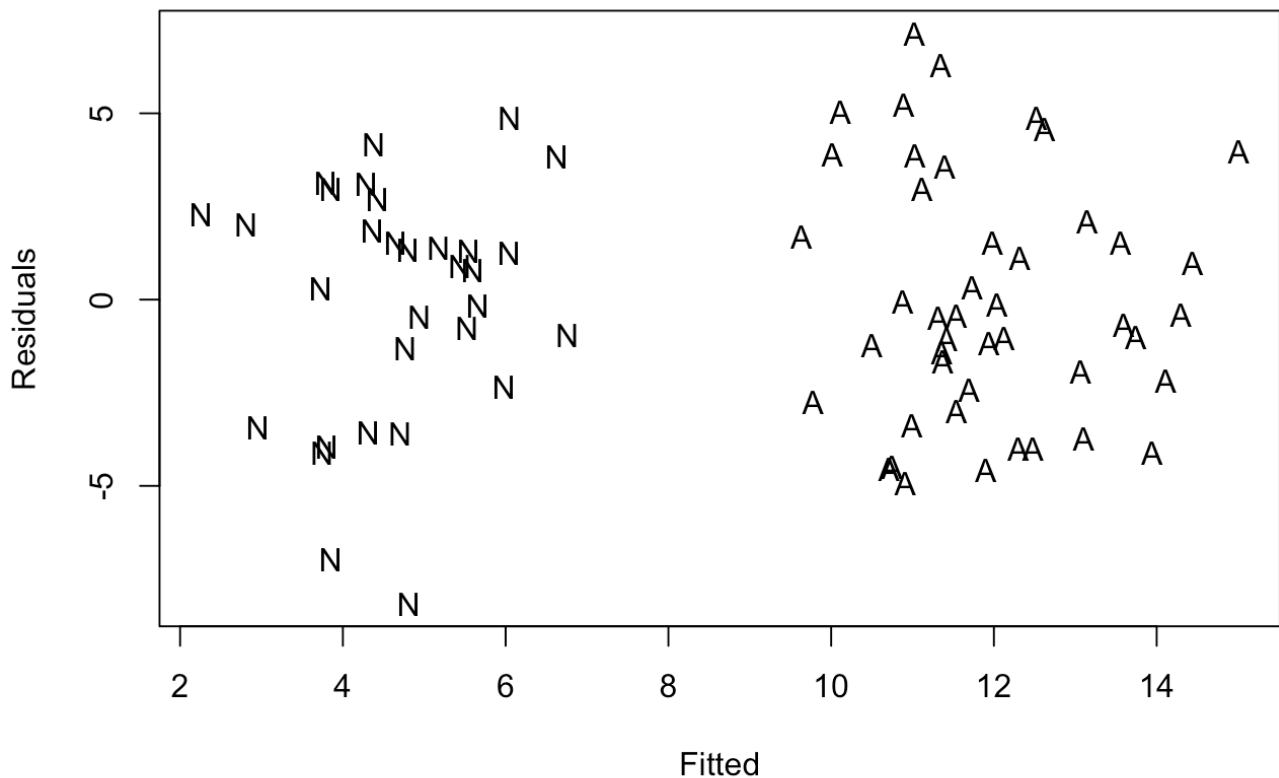
```
plot(ptsd ~ cpa, pch=as.character(csa), sexab)
abline(10.248, 0.551) # intercepto, pendiente
abline(10.248-6.273, 0.551, lty=2)
```



Esto es, luego de controlar el efecto del abuso físico infantil, el efecto estimado del abuso sexual infantil es 6.273 mayor en los sujetos “Abused” que en los “NotAbused”.

Diagnóstico

```
plot(fitted (g), residuals (g), pch=as.character (sexab$csa), xlab="Fitted", ylab="Residuals")
```



Vemos que no hay problemas de heterocedasticidad.

5.7 Resumen de funciones en R

ANOVA En los siguientes ejemplos las letras minúsculas corresponden a variables numéricas y las mayúsculas corresponden a factores (variables categóricas).

Diseño completamente aleatorizado

```
fit <- aov(y ~ A, data=mydataframe)
```

Diseño de bloques aleatorios (el factor de bloque es B)

```
fit <- aov(y ~ A + B, data=mydataframe)
```

Diseño bi-factorial o de dos vías

```
fit <- aov(y ~ A + B + A:B, data=mydataframe)
fit <- aov(y ~ A*B, data=mydataframe) # Lo mismo
```

ANCOVA

```
fit <- aov(y ~ A + x, data=mydataframe)
```

5.8 Referencias

- ⌘ ref1 (<http://www.statmethods.net/stats/anova.html>)
- ⌘ ref2 (<http://conjugateprior.org/2013/01/formulae-in-r-anova/>)
- ⌘ ref2 (<http://dawg.utk.edu/choose.htm>)

5.9 Versión robusta

A continuación veremos algunos ejemplos de ANOVA y ANCOVA robustos.

5.9.1 ANOVA robusto de dos vías

Consideremos ahora el caso del ANOVA robusto de dos vías.

La función para la comparación de medias recortadas es ***t2way***, para la comparación de medianas tenemos la función ***med2way*** y para M-estimadores más generales utilizaremos ***pbad2way***. Las pruebas de comparaciones múltiples posthoc las realizaremos con la función ***mcp2atm*** para las medias recortadas y la función ***mcp2a*** para las medianas.

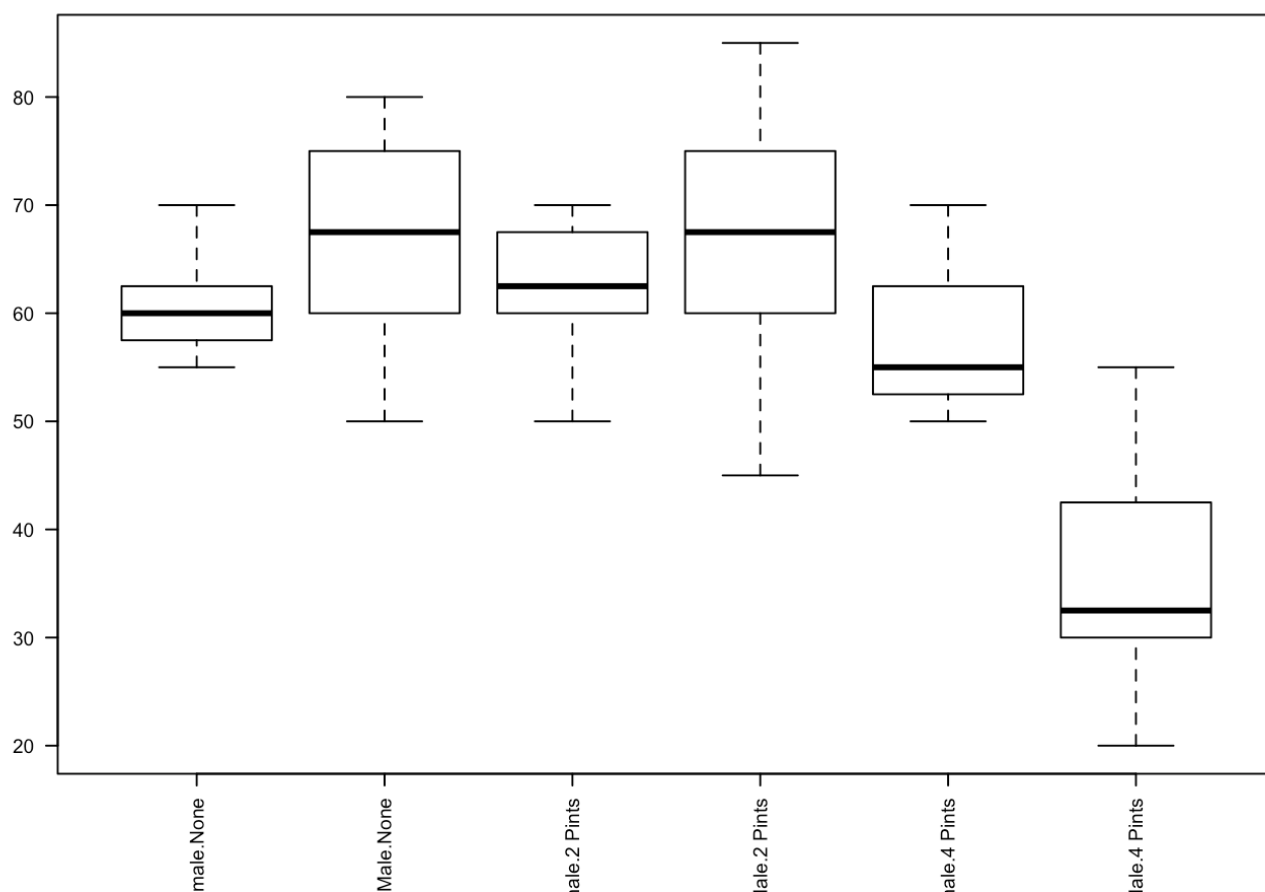
5.9.1.1 Aplicación en R

Como ejemplo utilizaremos el conjunto de datos ***goggles*** que recoge los efectos del alcohol en la selección de pareja en clubes nocturnos. Queremos averiguar si luego de consumir alcohol las percepciones subjetivas del atractivo físico se vuelven menos rigurosas (*beer goggles effect*) y si esto depende del sexo del sujeto (si ocurre con más frecuencia en los hombres respecto a las mujeres). La base de datos tiene dos factores: el sexo (24 hombres y 24 mujeres estudiantes) y el alcohol consumido (nada, 2 pintas o 4 pintas). Al final de la noche el investigador tomó una fotografía a la persona elegida por el participante. Un grupo de jueces independientes evaluó el poder de atracción de dicha persona dándole un puntaje de 0 a 100.

Observamos y graficamos los datos

Primero inspeccionamos los datos.

```
library(WRS2)
data(goggles)
par(cex=.6)
boxplot(attractiveness ~ gender*alcohol, data = goggles, las=2)
```



Observamos grandes diferencias en la distribución de los datos, lo que indica que la mejor alternativa es utilizar métodos robustos para su comparación.

Modelo y ANOVA

Realizamos la comparación de las medias recortadas, para las medianas y para el M-estimador *one-step*.

```
t2way(attractiveness ~ gender*alcohol, data = goggles)
```

```
## Call:
## t2way(formula = attractiveness ~ gender * alcohol, data = goggles)
##
##               value p.value
## gender           1.6667  0.209
## alcohol          48.2845  0.001
## gender:alcohol  26.2572  0.001
```

```
med2way(attractiveness ~ gender*alcohol, data = goggles)
```

```
## Call:
## med2way(formula = attractiveness ~ gender * alcohol, data = goggles)
##
##               value p.value
## gender           6.8444 0.0089
```

```
## alcohol          4.8207  0.0081
## gender:alcohol 12.9593  0.0015
```

```
pbad2way(attractiveness ~ gender*alcohol, data = goggles, est = "onestep")
```

```
## Call:
## pbad2way(formula = attractiveness ~ gender * alcohol, data = goggles,
##     est = "onestep")
##
##              p.value
## gender          0.1619
## alcohol          0.0000
## gender:alcohol  0.0000
```

Lo comparamos con el método clásico de ANOVA.

```
summary(aov(attractiveness ~ gender*alcohol, data = goggles))
```

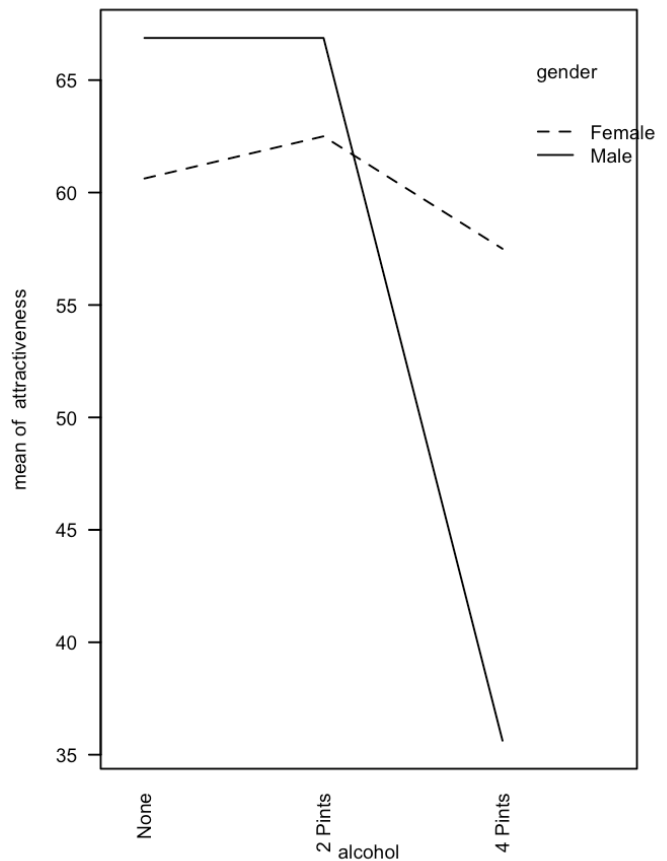
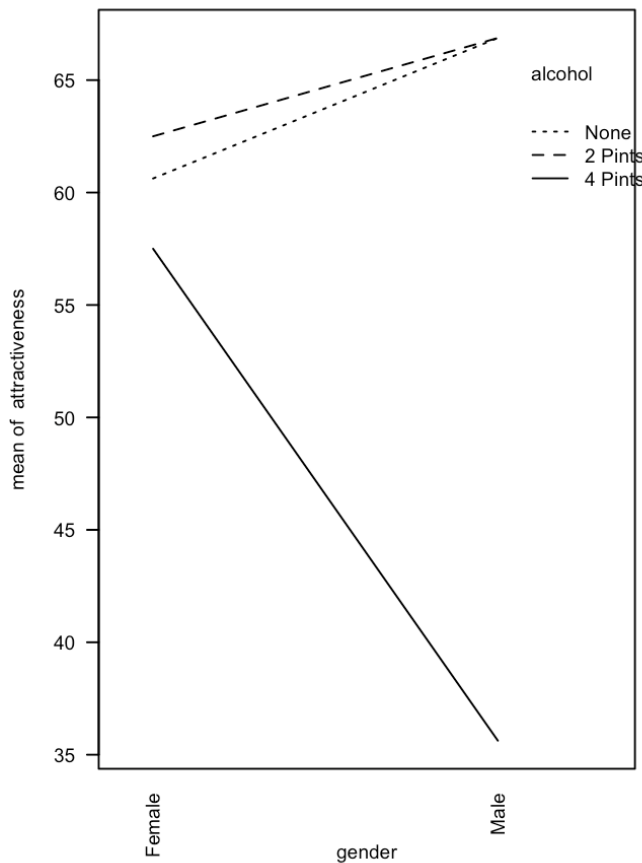
```
##              Df Sum Sq Mean Sq F value    Pr(>F)
## gender          1    169   168.7    2.032    0.161
## alcohol          2   3332  1666.1   20.065 7.65e-07 ***
## gender:alcohol  2   1978   989.1   11.911 7.99e-05 ***
## Residuals      42   3488    83.0
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
```

Obtenemos un efecto de interacción significativo.

Comparaciones múltiples de medias post hoc

Graficamos y realizamos las pruebas de comparaciones múltiples posthoc para interpretar los resultados.

```
attach(goggles)
par(mfrow=c(1,2),cex=.6)
interaction.plot(gender,alcohol,attractiveness, fixed = TRUE, las=2)
interaction.plot(alcohol,gender,attractiveness, fixed = TRUE, las=2)
```



#posthoc para medias truncadas

```
mcp2atm(attractiveness ~ gender*alcohol, data = goggles)
```

Call:

```
## mcp2atm(formula = attractiveness ~ gender * alcohol, data = goggles)
```

##

	psihat	ci.lower	ci.upper	p-value
gender1	10.00000	-6.00223	26.00223	0.20922
alcohol1	-3.33333	-20.49551	13.82885	0.61070
alcohol2	35.83333	19.32755	52.33911	0.00003
alcohol3	39.16667	22.46796	55.86537	0.00001
gender1:alcohol1	-3.33333	-20.49551	13.82885	0.61070
gender1:alcohol2	-29.16667	-45.67245	-12.66089	0.00025
gender1:alcohol3	-25.83333	-42.53204	-9.13463	0.00080

antigua versión de WRS2

```
# library(reshape)
```

```
# x=split(goggles, f=goggles[,c("gender","alcohol")])
```

```
# mcp2atm(2,3,x)
```

#posthoc para M-estimadores (mom)

```
mcp2a(attractiveness ~ gender*alcohol, data = goggles)
```

Call:

```
## mcp2a(formula = attractiveness ~ gender * alcohol, data = goggles)
```

##

	psihat	ci.lower	ci.upper	p-value
gender1	14.46429	-8.96429	28.12500	0.12855
alcohol1	-5.00000	-19.37500	13.28571	0.37229
alcohol2	35.80357	18.92857	50.62500	0.00000
alcohol3	40.80357	24.28571	55.00000	0.00000


```
## gender1:alcohol1 -5.00000 -20.37500 13.00000 0.31720
## gender1:alcohol2 -32.23214 -44.87500 -13.91667 0.00000
## gender1:alcohol3 -27.23214 -42.96429 -11.60714 0.00167
```

#posthoc para medianas

```
mcp2a(attractiveness ~ gender*alcohol, data = goggles, est = "median")
```

```
## Call:
## mcp2a(formula = attractiveness ~ gender * alcohol, data = goggles,
##       est = "median")
##
##               psihat ci.lower ci.upper p-value
## gender1          10.0      -7.5      27.5 0.17028
## alcohol1         -2.5     -20.0      15.0 0.35058
## alcohol2         40.0      20.0      52.5 0.00000
## alcohol3         42.5      22.5      55.0 0.00000
## gender1:alcohol1  -2.5     -20.0      12.5 0.23372
## gender1:alcohol2 -30.0     -45.0     -12.5 0.00000
## gender1:alcohol3 -27.5     -42.5     -10.0 0.00000
```

Observamos que las conquistas de los hombres son menos atractivas luego de 4 pintas de alcohol. El resultado más interesante es la comparación posthoc `gender1:alcohol3` ya que explica la caída en la atracción para 4 pintas de alcohol en hombres.

5.9.2 Diseños mixtos robustos

Trabajaremos con un diseño mixto, es decir, tenemos efectos dentro de los sujetos (*within-subjects*) debido a medidas repetidas y entre los sujetos (*between-subjects*) al comparar grupos.

En las pruebas paramétricas utilizamos la función **aov** o la función **ezANOVA** del paquete **ez** o con los paquetes **lme4** y **nlme**.

En las pruebas robustas utilizaremos el paquete **WRS2** con las funciones **bwtrim** para ANOVA *between-within subjects* por medias recortadas. De manera más general para los M-estimadores con bootstrap **sppba** (*between-subjects effect*), **sppbb** (*within-subjects effect*) y **sppbu** (*interaction effect*).

5.9.1.2 Aplicación en R

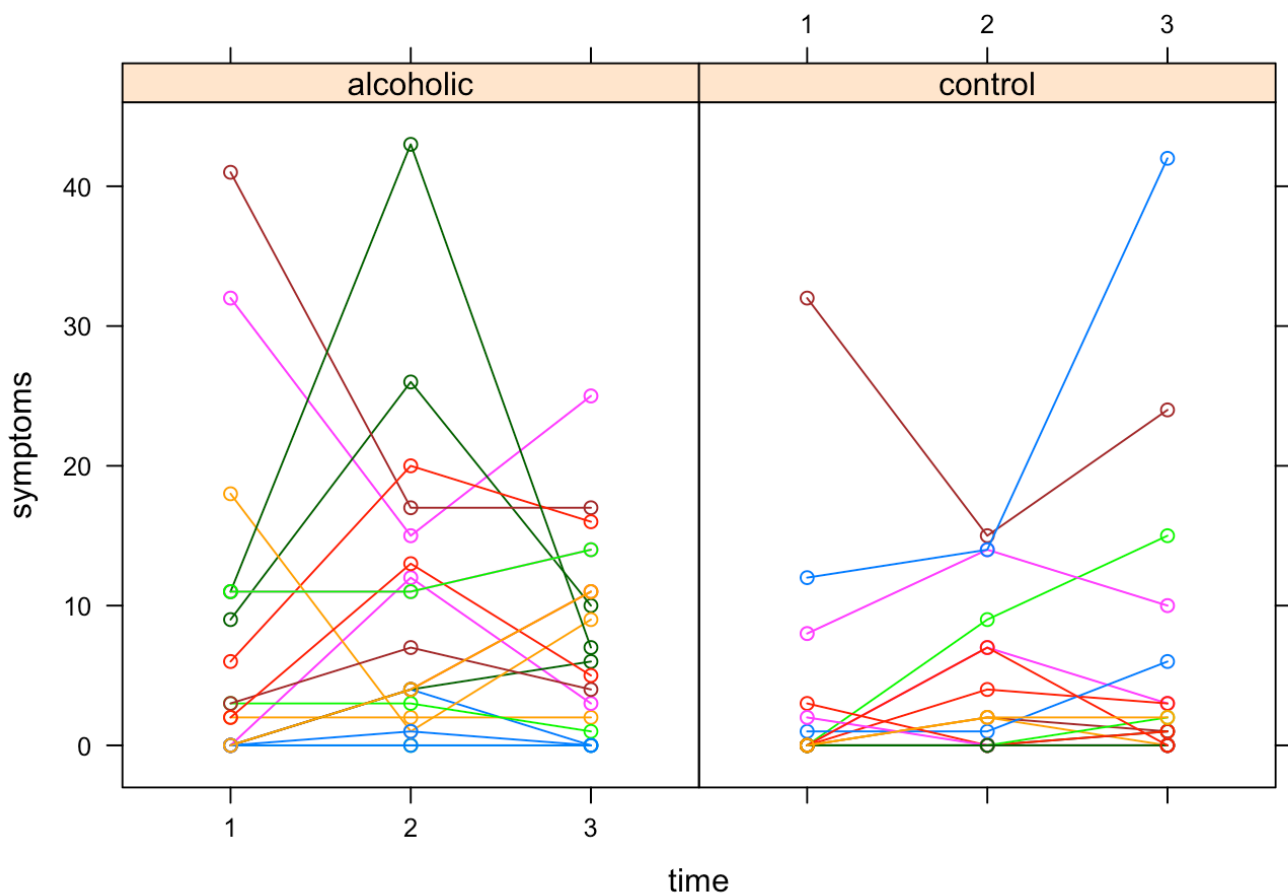
Utilizaremos el ejemplo de Wilcox (2012, p.411) sobre el estudio del efecto del consumo de alcohol. Se midieron los síntomas de la resaca para dos grupos independientes (un grupo estaba formado por hijos de alcohólicos y el otro era un grupo de control) medidos en tres ocasiones diferentes.

```
data(hangover)
head(hangover)
```

```
##   symptoms    group time id
## 1         0 alcoholic 1 1
```

```
## 2      32 alcoholic    1  2
## 3       9 alcoholic    1  3
## 4       0 alcoholic    1  4
## 5       2 alcoholic    1  5
## 6       0 alcoholic    1  6
```

```
library(lattice)
par(cex = .6)
xyplot(symptoms ~ time|group, data = hangover, groups = id,
       type = "o", panel = panel.superpose)
```



Primero realizaremos el ANOVA *between-within subjects* con las medias recortadas.

```
bwtrim(symptoms ~ group*time, id = id, data = hangover)
```

```
## Call:
## bwtrim(formula = symptoms ~ group * time, id = id, data = hangover)
##
##           value p.value
## group      6.6087 0.0218
## time       4.4931 0.0290
## group:time 0.5663 0.5790
```

Se obtiene un efecto significativo para grupo y tiempo, pero no para su interacción.

Realizamos también las comparaciones con el M-estimador de Huber aplicando las 3 funciones mencionadas, una para cada efecto.

```
sppba(symptoms ~ group*time, id, data = hangover)
```

```
## Call:
## sppba(formula = symptoms ~ group * time, id = id, data = hangover)
##
## Test statistic: 4.714
## p-value: 0.022
```

```
sppbb(symptoms ~ group*time, id, data = hangover)
```

```
## Call:
## sppbb(formula = symptoms ~ group * time, id = id, data = hangover)
##
## Test statistic: -1.8387 -0.6875 -0.1176
## p-value: 0.16
```

```
sppbi(symptoms ~ group*time, id, data = hangover)
```

```
## Call:
## sppbi(formula = symptoms ~ group * time, id = id, data = hangover)
##
## Test statistic: -0.9375 0.4157 -0.5
## p-value: 0.866
```

Esta prueba indica que existe un efecto significativo de grupo, pero no de tiempo ni de la interacción entre ambos.

Comparamos estos resultados con el ANOVA clásico seleccionando trimm=0.

```
bwtrim(symptoms ~ group*time, id = id, data = hangover, tr=0)
```

```
## Call:
## bwtrim(formula = symptoms ~ group * time, id = id, data = hangover,
##      tr = 0)
##
##              value p.value
## group      3.2770  0.0783
## time       0.8809  0.4250
## group:time 1.0508  0.3624
```

#Lo mismo con otra función

```
library(ez)
fitF <- ezANOVA(hangover, symptoms, between = group, within = time, wid = id)
fitF$ANOVA
```

##	Effect	DFn	DFd	F	p	p<.05	ges
## 2	group	1	38	3.2770015	0.07817048		0.056208518
## 3	time	2	76	0.8957333	0.41257420		0.007240111
## 4	group:time	2	76	0.9737002	0.38234407		0.007865351

No se detecta ningún efecto significativo. Aquí vemos la importancia de realizar pruebas robustas, ya que de otra manera no lograríamos visualizar ningún efecto.

5.9.3 ANCOVA robusto

El ANCOVA implica un diseño factorial y una covariable métrica que no forma parte de la manipulación experimental.

El ANCOVA asume homogeneidad en las pendiente de regresión entre los grupos cuando se relaciona la variable dependiente con la covariable. Además, asume homogeneidad de varianza en el término de error entre los grupos. Cuando no podemos asumir alguna de ellas, la mejor alternativa es utilizar el ANCOVA robusto que utiliza el suavizado por intervalos sobre las medias recortadas de cada subgrupo. Para pequeños tamaños muestrales se recomienda el suavizado de los datos.

En el paquete **WRS2** contamos con la función **ancova** para este tipo de análisis cuando tenemos 1 factor con 2 categorías y 1 covariable. Su versión bootstrap es **ancboot**.

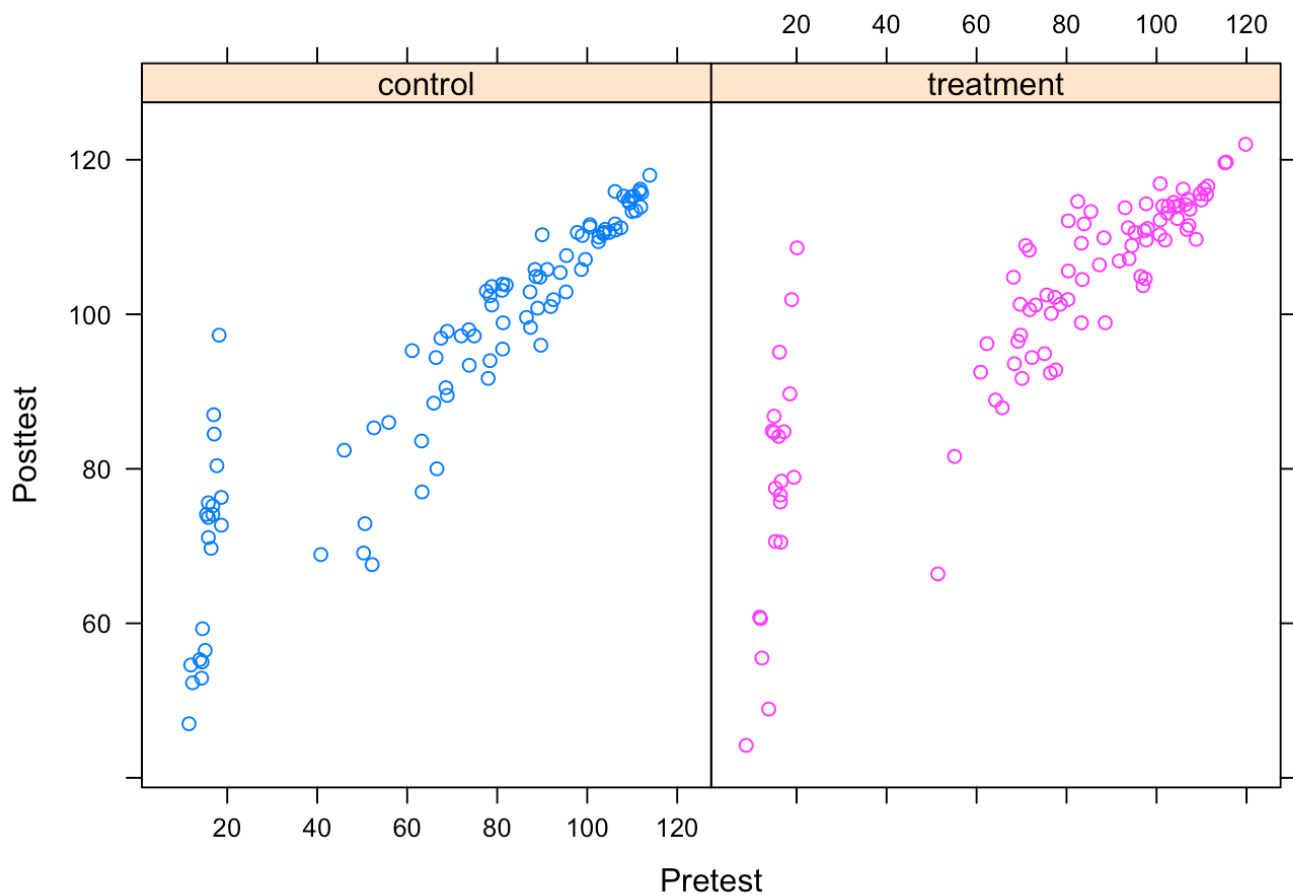
5.9.3.1 Aplicación en R

Utilizaremos los datos de un programa de televisión educativo para niños llamados “*The Electric Company*”. Se tomaron niños de 4 grados distintos y se realizó un muestreo aleatorio para colocarlos en dos grupos: control y tratamiento. Los niños del tratamiento fueron expuestos al programa de TV mientras que los niños del control no. Al comienzo y al final del año escolar, se pasó una prueba de lectura a los niños, obteniéndose puntuaciones promedios por clase. Nos interesa comparar los valores pre y posttest para lo cual utilizaremos los posttest como variable dependiente y el pretest como covariable, comparando los grupos control y tratamiento.

```
data(electric)
head(electric)
```

##	City	Grade	Pretest	Posttest	Group
## 1	F	1	13.8	48.9	treatment
## 2	F	1	16.5	70.5	treatment
## 3	F	1	18.5	89.7	treatment
## 4	F	1	8.8	44.2	treatment
## 5	F	1	15.3	77.5	treatment
## 6	F	1	15.0	84.7	treatment

```
par(cex = .6)
xyplot(Posttest ~ Pretest|Group, data = electric, group=Group, type = "p")
```



Estamos interesados en las comparaciones en 6 puntos específicos del diseño (pts).

```
comppts=c(18,70,80,90,100,110)
(fitanco<-ancova(Posttest ~ Pretest + Group, fr1 = 0.3, fr2 = 0.3,
  data = electric, pts = comppts)
)
```

```
## Call:
## ancova(formula = Posttest ~ Pretest + Group, data = electric,
##   fr1 = 0.3, fr2 = 0.3, pts = comppts)
##
##               n: control n: treatment trimmed mean diff      se lower CI
## Pretest = 18          21          20         -11.1128  4.2694 -23.3621
## Pretest = 70          20          21          -3.2186  1.9607  -8.8236
## Pretest = 80          24          23          -2.8146  1.7505  -7.7819
## Pretest = 90          24          22          -5.0670  1.3127  -8.7722
## Pretest = 100         28          30          -1.8444  0.9937  -4.6214
## Pretest = 110         24          22          -1.2491  0.8167  -3.5572
##
##               upper CI statistic p-value
## Pretest = 18      1.1364      2.6029  0.0163
## Pretest = 70      2.3864      1.6416  0.1143
## Pretest = 80      2.1528      1.6079  0.1203
## Pretest = 90     -1.3617      3.8599  0.0006
## Pretest = 100      0.9325      1.8561  0.0729
## Pretest = 110      1.0590      1.5294  0.1380
```

Graficamos el modelo resultante. Las líneas grises verticales marcan los 6 puntos del diseño. Para un nivel de significación del 0.01, solo hay una diferencia significativa, la del valor de pretest de $x = 90$.

```
plot(electric$Pretest, electric$Posttest, col = rep(1:2, each = 96),
  pch = 1, cex = 0.8,
```

```

xlab = "Pretest Score", ylab = "Posttest Score", main = "TV Show ANCOVA")
eltr <- subset(electric, subset = Group == "treatment")
elct <- subset(electric, subset = Group == "control")
ordtr <- order(eltr$Pretest)
lines(eltr$Pretest[ordtr], fitanc$fitted.values$treatment[ordtr], col = 1, lwd = 2)
abline(lm(eltr$Posttest ~ eltr$Pretest), col = 1, lty = 2)
ordct <- order(elct$Pretest)
lines(elct$Pretest[ordct], fitanc$fitted.values$control[ordct], col = 2, lwd = 2)
abline(lm(elct$Posttest ~ elct$Pretest), col = 2, lty = 2)
abline(v = comppts, lty = 2, col = "gray")
legend(30, 120, legend = c("treatment", "control"), lty = 1, col = 1:2)

```

TV Show ANCOVA

