

Linear Model: Cheese Production

For the cheese producers, the profitability is the relation between the weight of the cheese and the amount of milk used to produce it. One is interested in analyzing if the type of milk used changes the profit.

Three milk types are considered: **cow, sheep and goat**. For each milk type, one is interested in comparing the profit as a function of the **thermic milk treatment** (plane or pasteurized) and the presence of a given **additive** denoted by CaCl_2 . What is the best combination of THERMIC and CaCl_2 ?

When we have got 2 factors, we can consider:

- The main effect of the factor F_1 .
- The main effect of the factor F_2 .
- The interaction between F_1 and F_2 .

$$Y_{ijk} = \mu + \alpha_i + \beta_j + \gamma_{ij} + e_{ijk}$$

Where i is the factor F_1 , j is the factor F_2 and k is the number of repetitions. If there are no repetitions, γ_{ij} and e_{ijk} are merged (and if there are repetitions, we can divide them, so that is what we could not do in the previous example, because there were no repetitions).

Si posar clorur en llet crua o pasteurizada té efectes diferents, vol dir que hi ha interacció. I si té el mateix efecte, vol dir que la interacció no és significativa. Si hi ha interacció, no podem dir quin tractament tèrmic és millor, sinó que haurem dir quina combinació és millor (tractament i additiu CaCl_2)!

(a) Assuming that the ANOVA conditions are satisfied, for each type of milk answer the following questions:

- Which treatment is the best?
- The addition of CaCl_2 does it increase the profit?
- Does the thermic treatment increase the profit?

```
library(car)
library(emmeans)
library(tables)
library("RcmdrMisc")
```

```
dd<-read.csv2("cheese.csv")
head(dd)
```

```
##   THERMIC CaCl2    COW   GOAT  SHEEP
## 1   plane   yes 46.960 50.565 53.345
## 2   plane   yes 37.235 37.685 53.790
## 3   plane   yes 36.205 34.620 52.635
## 4   plane   yes 38.835 40.920 50.330
## 5   plane   no  61.120 44.790 73.530
## 6   plane   no  53.570 37.015 57.770
```

COW MILK

Descriptiva

```
tabular(COW*(THERMIC=1)~((CaCl2)*mean), dd)
```

		CaCl2	
		no	yes
		mean	mean
COW	THERMIC	54.55	42.9

```
tabular(COW*(THERMIC=1)~((CaCl2)*sd), dd)
```

		CaCl2	
		no	yes
		sd	sd
COW	THERMIC	5.662	5.389

```
tabular(COW*(THERMIC=1)~((CaCl2)*(n=1)), dd)
```

		CaCl2	
		no	yes
		n	n
COW	THERMIC	8	8

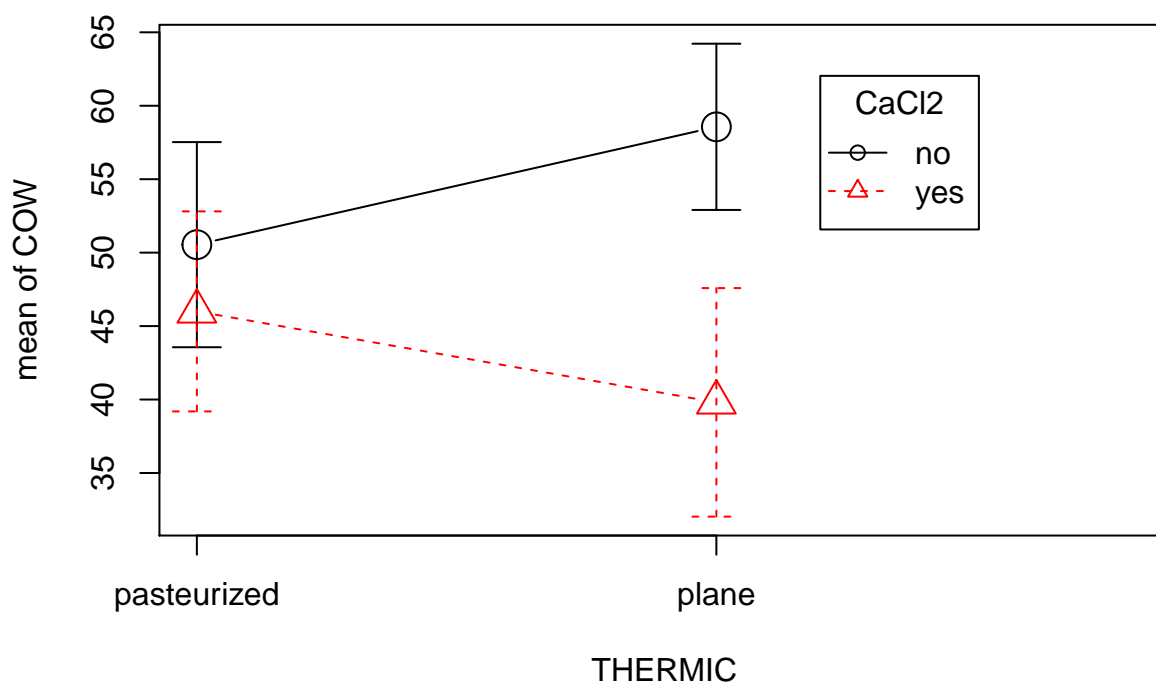
```
tabular(COW*(THERMIC+1)~((CaCl2+1)*mean), dd)
```

		CaCl2		
		no	yes	All
		mean	mean	mean
COW	pasteurized	50.54	46.00	48.27
	plane	58.56	39.81	49.19
	All	54.55	42.90	48.73

```
# plot to compare the means
```

```
with(dd, plotMeans(COW, THERMIC, CaCl2, error.bars="conf.int", level=0.95, connect=TRUE))
```

Plot of Means

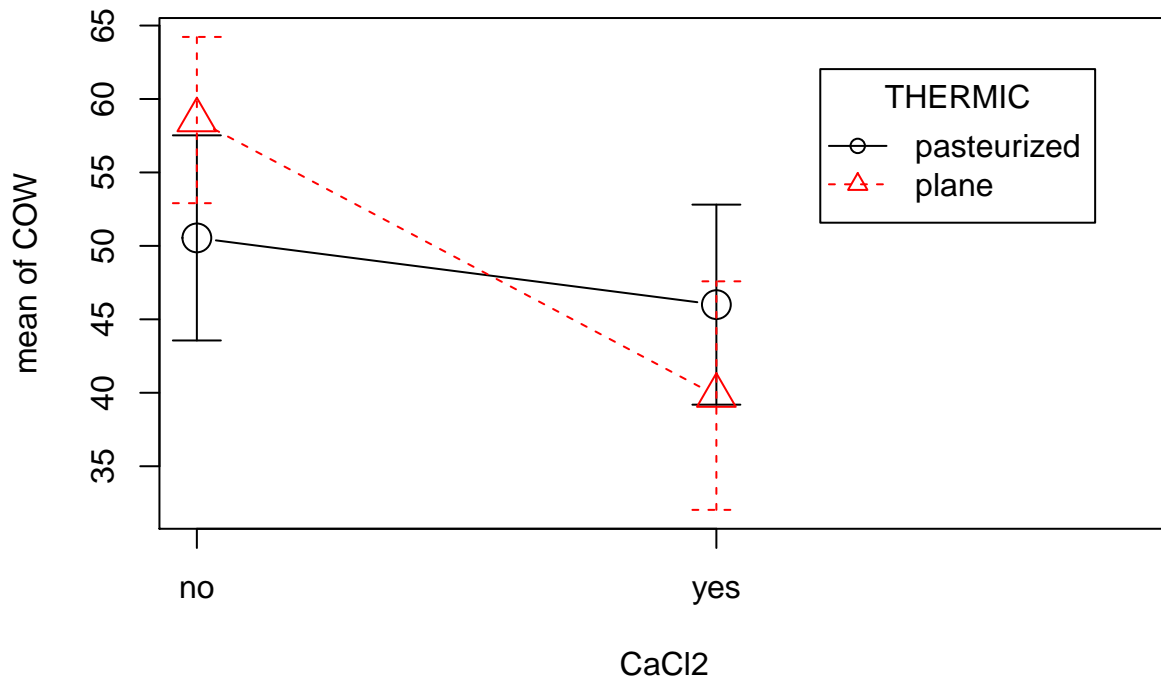


Segons les mitjanes (+ intervals de confiança), veiem que posar CaCl2 a la llet pasteuritzada gairebé no afecta, però a la crua afecta molt (posar CaCl2 disminueix molt el rendiment). És un cas clar de què hi ha interacció. Si no hi hagués interacció, aquestes rectes sortirien paral·leles. Podria ser poligonal si hi haguessin

més nivells.

```
# plot to compare the means
with(dd, plotMeans(COW, CaCl2, THERMIC, error.bars="conf.int", level=0.95, connect=TRUE))
```

Plot of Means



Modelization

$profit \sim THERMIC + CaCl_2 + THERMIC * CaCl_2 \rightarrow profit \sim THERMIC * CaCl_2$

```
# summary(mc<-lm(COW~THERMIC+CaCl2+THERMIC*CaCl2, dd))
summary(mcow<-lm(COW~THERMIC*CaCl2, dd))
```

```
##
## Call:
## lm(formula = COW ~ THERMIC * CaCl2, data = dd)
##
## Residuals:
##      Min       1Q   Median       3Q      Max
## -5.9850 -3.2138  0.1913  2.5437  7.1513
##
## Coefficients:
##              Estimate Std. Error t value Pr(>|t|)
## (Intercept)      50.543      2.152   23.484 2.12e-11 ***
## THERMICplane       8.023      3.044    2.636  0.02174 *
## CaCl2yes         -4.542      3.044   -1.492  0.16141
## THERMICplane:CaCl2yes -14.214      4.304   -3.302  0.00632 **
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## Residual standard error: 4.304 on 12 degrees of freedom
```

```
## Multiple R-squared:  0.7709, Adjusted R-squared:  0.7136
## F-statistic: 13.46 on 3 and 12 DF,  p-value: 0.0003799
```

Fixem-nos en el **test Omnibus**: la hipòtesi nul·la H_0 és “no hi ha cap efecte” ($\alpha_i = \beta_j = \gamma_{ij} = 0$), la hipòtesi alternativa H_1 : “hi ha algún efecte”. Com que el p-valor és $0.0003799 \ll 0.05$, el test és significatiu i significa que hi ha algún efecte.

Si el test sortís no significatiu, hem de plegar. Vol dir que amb aquestes dades, aquest model no serveix. O hem de completar el model, perquè hi ha algún efecte que no hem tingut en compte, o hem d'obtenir més dades, etc.

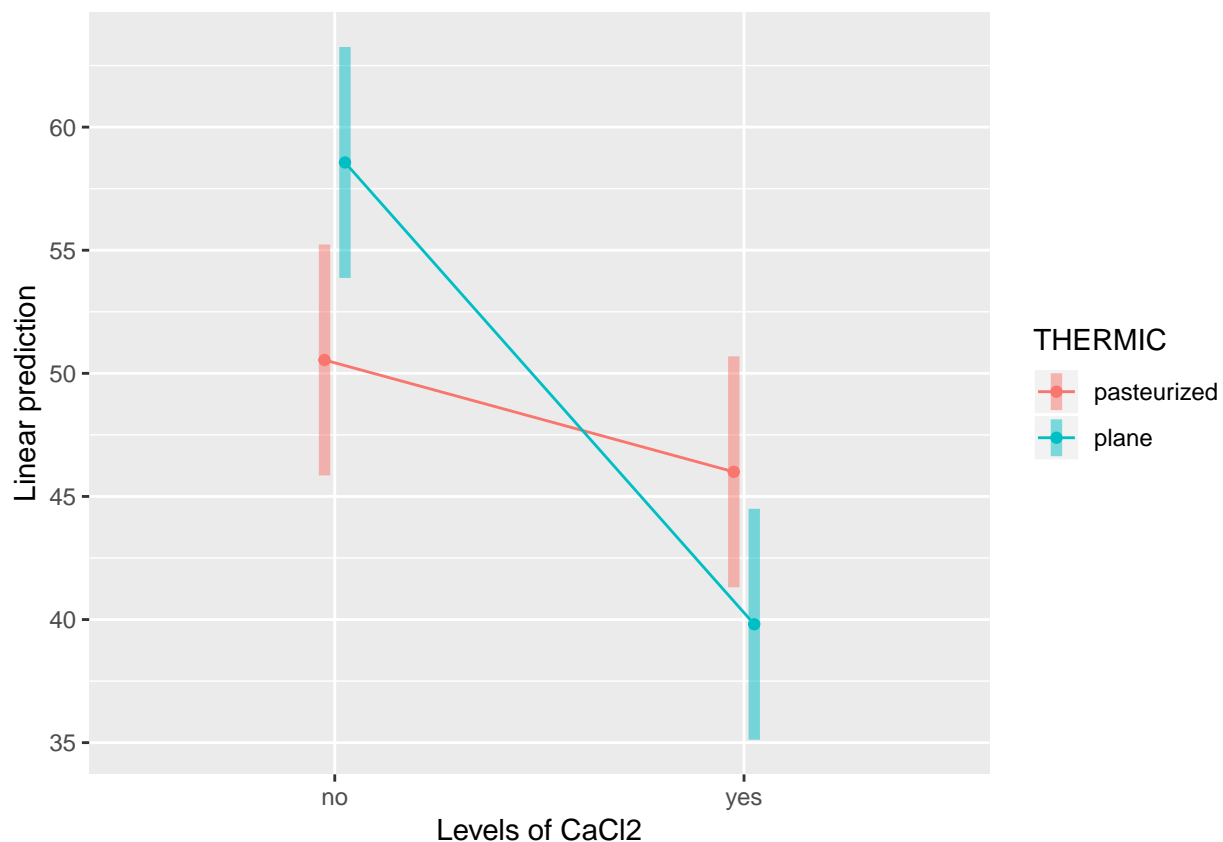
Com que en aquest cas el test ens mostra que hi ha efecte... Ens podem preguntar: **l'efecte de qui és?**

- Hi ha efecte principal tèrmic?
- Hi ha efecte principal del CaCl₂?
- Hi ha efecte interacció?

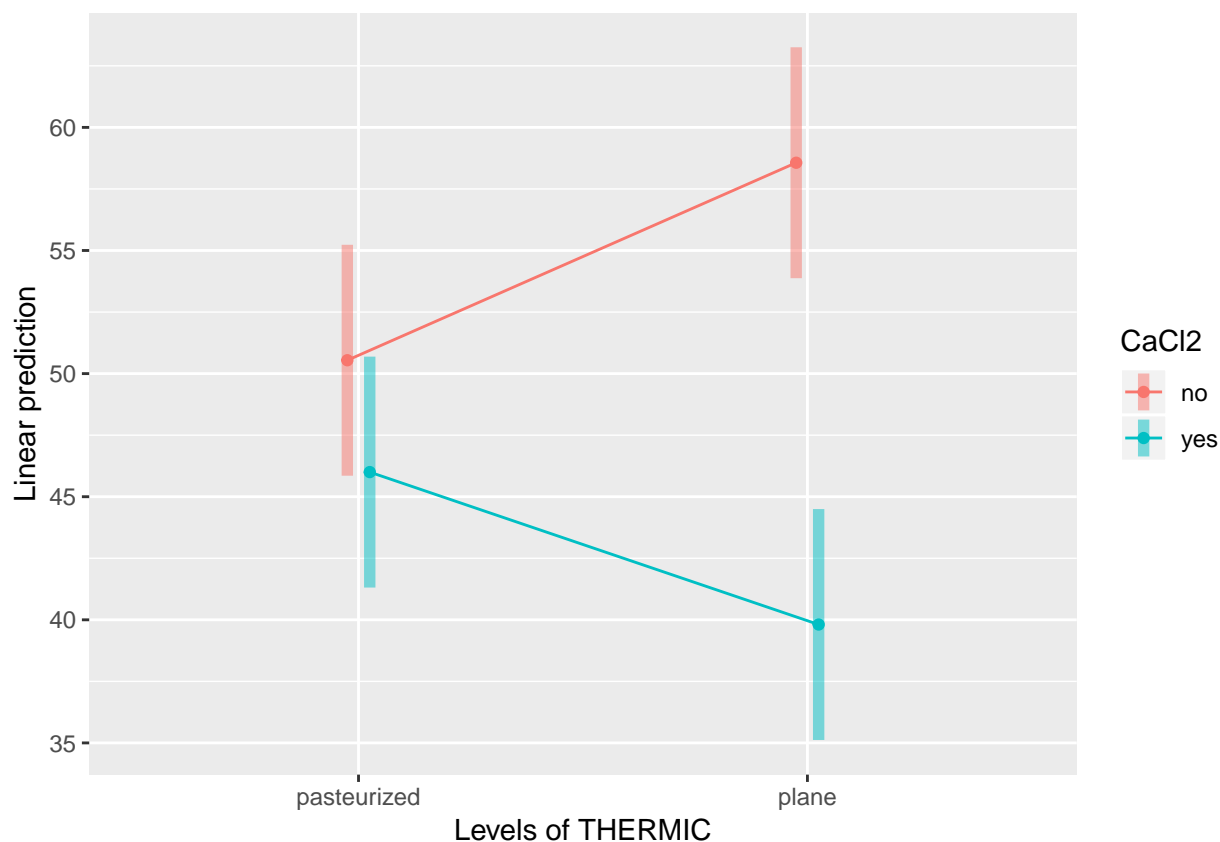
Per això fem el **test ANOVA**.

Abans, mirem-ne el *emmip*, la gràfica de la interacció, que fa un resum del que hem estimat amb el model. La diferència amb la gràfica anterior és que aquests intervals de confiança estan calculats amb la variància conjunta i en l'altre, estaven calculats amb les dades de cada grup.

```
(emmip(mcow, THERMIC~CaCl2, CIs = TRUE))
```



```
(emmip(mcow, CaCl2~THERMIC, CIs = TRUE))
```



Anova II (= anova I, en aquest cas):

```
Anova(mcow)
```

```
## Anova Table (Type II tests)
##
## Response: COW
##          Sum Sq Df F value    Pr(>F)
## THERMIC      3.35  1   0.181 0.6780528
## CaCl2       542.83  1  29.297 0.0001568 ***
## THERMIC:CaCl2 202.03  1  10.904 0.0063162 **
## Residuals    222.34 12
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
```

Observem que:

- (a) El tractament tèrmic, com a efecte principal, no afecta, però en la interacció sí.
- (b) El CaCl2 afecta tant en l'efecte principal, com en la interacció.

Quines comparacions ens interessaran? Les 4 possibles combinacions. Fem el *emmeans* de l'efecte combinació (~THERMIC*CaCl2) i fa els 4 tractaments i posa les lletres:

```
(emm<-emmeans(mcow, ~THERMIC*CaCl2))
```

```
##   THERMIC   CaCl2   emmean      SE df lower.CL upper.CL
##   pasteurized no    50.54250 2.152232 12 45.85319 55.23181
##   plane      no    58.56500 2.152232 12 53.87569 63.25431
```

```
## pasteurized yes 46.00000 2.152232 12 41.31069 50.68931
## plane yes 39.80875 2.152232 12 35.11944 44.49806
##
## Confidence level used: 0.95
```

```
CLD(emm, Letters=letters, reversed=T)
```

```
## THERMIC CaCl2 emmean SE df lower.CL upper.CL .group
## plane no 58.56500 2.152232 12 53.87569 63.25431 a
## pasteurized no 50.54250 2.152232 12 45.85319 55.23181 ab
## pasteurized yes 46.00000 2.152232 12 41.31069 50.68931 bc
## plane yes 39.80875 2.152232 12 35.11944 44.49806 c
##
## Confidence level used: 0.95
## P value adjustment: tukey method for comparing a family of 4 estimates
## significance level used: alpha = 0.05
```

Segons el resultat:

- (a) El millor tractament seria (el que té A): llet de vaca crua + sense CaCl₂ o llet de vaca pasteuritzada + sense CaCl₂.
- (b) És millor no posar CaCl₂. El rendiment més baix el té C.

En la llet crua, hi ha alguna diferència entre llet amb CaCl₂ o no? Potser és millor mirar condicionant si posem clorur o no, per cada a tractament termic per separat. És una forma més resumida:

```
(emm<-emmeans(mcow, ~THERMIC|CaCl2))
```

```
## CaCl2 = no:
## THERMIC emmean SE df lower.CL upper.CL
## pasteurized 50.54250 2.152232 12 45.85319 55.23181
## plane 58.56500 2.152232 12 53.87569 63.25431
##
## CaCl2 = yes:
## THERMIC emmean SE df lower.CL upper.CL
## pasteurized 46.00000 2.152232 12 41.31069 50.68931
## plane 39.80875 2.152232 12 35.11944 44.49806
##
## Confidence level used: 0.95
```

```
CLD(emm, Letters=letters, reversed=T)
```

```
## CaCl2 = no:
## THERMIC emmean SE df lower.CL upper.CL .group
## plane 58.56500 2.152232 12 53.87569 63.25431 a
## pasteurized 50.54250 2.152232 12 45.85319 55.23181 b
##
## CaCl2 = yes:
## THERMIC emmean SE df lower.CL upper.CL .group
## pasteurized 46.00000 2.152232 12 41.31069 50.68931 a
## plane 39.80875 2.152232 12 35.11944 44.49806 a
##
## Confidence level used: 0.95
## significance level used: alpha = 0.05
```

Observem, segons els resultats:

- (a) Quan **no** posem CaCl₂, és millor la llet crua que pasteuritzada.
- (b) Quan **sí** que hi posem, és indistingible si crua o pasteuritzada (no detectem diferències en el rendiment).

El mateix es pot fer CaCl2:

```
(emm<-emmeans(mcow,~CaCl2|THERMIC))
```

```
## THERMIC = pasteurized:
## CaCl2    emmean      SE df lower.CL upper.CL
## no      50.54250 2.152232 12 45.85319 55.23181
## yes     46.00000 2.152232 12 41.31069 50.68931
##
## THERMIC = plane:
## CaCl2    emmean      SE df lower.CL upper.CL
## no      58.56500 2.152232 12 53.87569 63.25431
## yes     39.80875 2.152232 12 35.11944 44.49806
##
## Confidence level used: 0.95
```

```
CLD(emm,Letters=letters, reversed=T)
```

```
## THERMIC = pasteurized:
## CaCl2    emmean      SE df lower.CL upper.CL .group
## no      50.54250 2.152232 12 45.85319 55.23181  a
## yes     46.00000 2.152232 12 41.31069 50.68931  a
##
## THERMIC = plane:
## CaCl2    emmean      SE df lower.CL upper.CL .group
## no      58.56500 2.152232 12 53.87569 63.25431  a
## yes     39.80875 2.152232 12 35.11944 44.49806  b
##
## Confidence level used: 0.95
## significance level used: alpha = 0.05
```

Observem, segons els resultats:

- (a) Quan la llet és **pasteuritzada**, és indistingible posar o no posar CaCl2.
- (b) Quan la llet és **crua**, és millor no posar CaCl2.

El que no podem fer, si hi ha interacció, és fer la separació de mitjanes només per CaCl2, o només per THERMIC, perquè aleshores no tenim en compte la interacció i podem tenir resultats erronis com en aquest cas (fixeu-vos, la nota ens avisa!):

```
(emm<-emmeans(mcow,~CaCl2))
```

```
## NOTE: Results may be misleading due to involvement in interactions

## CaCl2    emmean      SE df lower.CL upper.CL
## no      54.55375 1.521858 12 51.23791 57.86959
## yes     42.90438 1.521858 12 39.58853 46.22022
##
## Results are averaged over the levels of: THERMIC
## Confidence level used: 0.95
```

```
CLD(emm,Letters=letters, reversed=T)
```

```
## CaCl2    emmean      SE df lower.CL upper.CL .group
## no      54.55375 1.521858 12 51.23791 57.86959  a
## yes     42.90438 1.521858 12 39.58853 46.22022  b
##
## Results are averaged over the levels of: THERMIC
## Confidence level used: 0.95
```

```
## significance level used: alpha = 0.05
```

Diagnostics

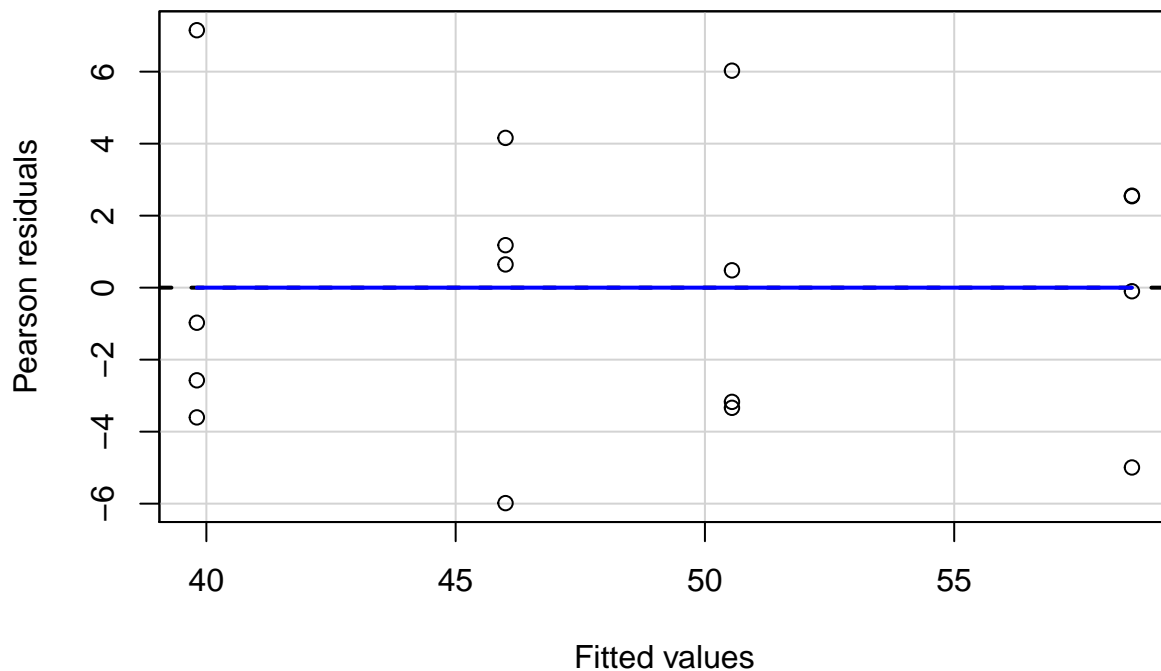
Amb 2 factors tenim:

- **Model factorial:** $Y_{ijr} = \mu + \alpha_i + \beta_j + \gamma_{ij} + e_{ijr}$
- **Model additiu:** els efectes principals dels dos factors se sumen i no hi ha interacció, $Y_{ijr} = \mu + \alpha_i + \beta_j + e_{ijr}$. És un model més senzill.
- Després tenim dos models una mica redundants, quan sabem que un dels factors no té efecte principal ni en la interacció, podem tenir: $Y_{ijr} = \mu + \alpha_i + e_{ijr}$ o $Y_{ijr} = \mu + \beta_j + e_{ijr}$.
- I finalment, el **model null:** $Y_{ijr} = \mu + e_{ijr}$.

Com podem saber que el model additiu no és l'adequat?

- Podem ajustar el model factorial i si la interacció és significativa, vol dir que entre els models factorial i additiu hi ha diferències significatives, és a dir, si la interacció és significativa, vol dir que el model factorial ajusta millor. Per tant, una manera, per veure que el model additiu no és l'adequat consisteix en ajustar un altre model - el factorial, i mirar la interacció.
- L'altra manera seria mirant els residuals. Els residuals han d'anar la voltant de zero. Si hi ha interacció, ens sortiran grups de residuals que no van al voltant de zero. Pot ser difícil veure-ho per l'efecte de l'atzar, però si veiem clarament grups de residuals que se separen del zero, vol dir que ens falta algun efecte (= la interacció).

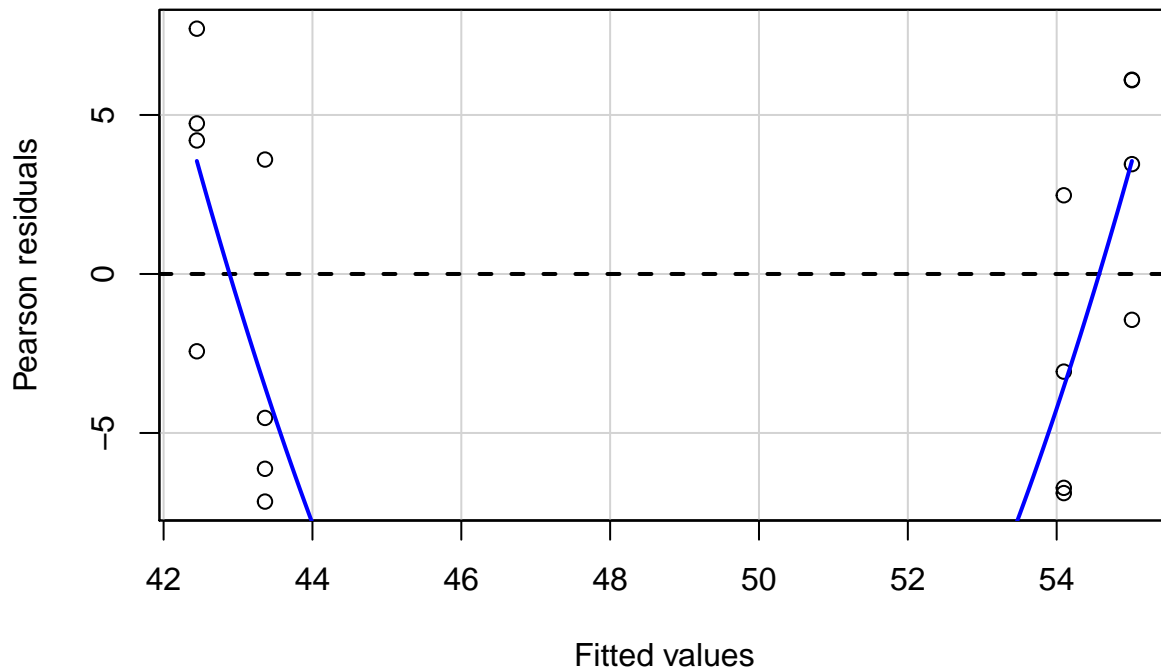
```
residualPlot(mcow)
```



Què passaria si fèssim el model sense interacció (= model additiu)? Aleshores, tota la interacció (suma de quadrats, etc) aniria a l'atzar (errors, residus) i tindriem molta més variabilitat i podria ser que els valors predits no anessin al voltant de zero (perquè falta la informació de la interacció).

L'única manera de veure, que sense interacció no funciona, és fent els diagnòstics:

```
no_mcow<-lm(COW~THERMIC+CaCl2, dd)
residualPlot(no_mcow)
```

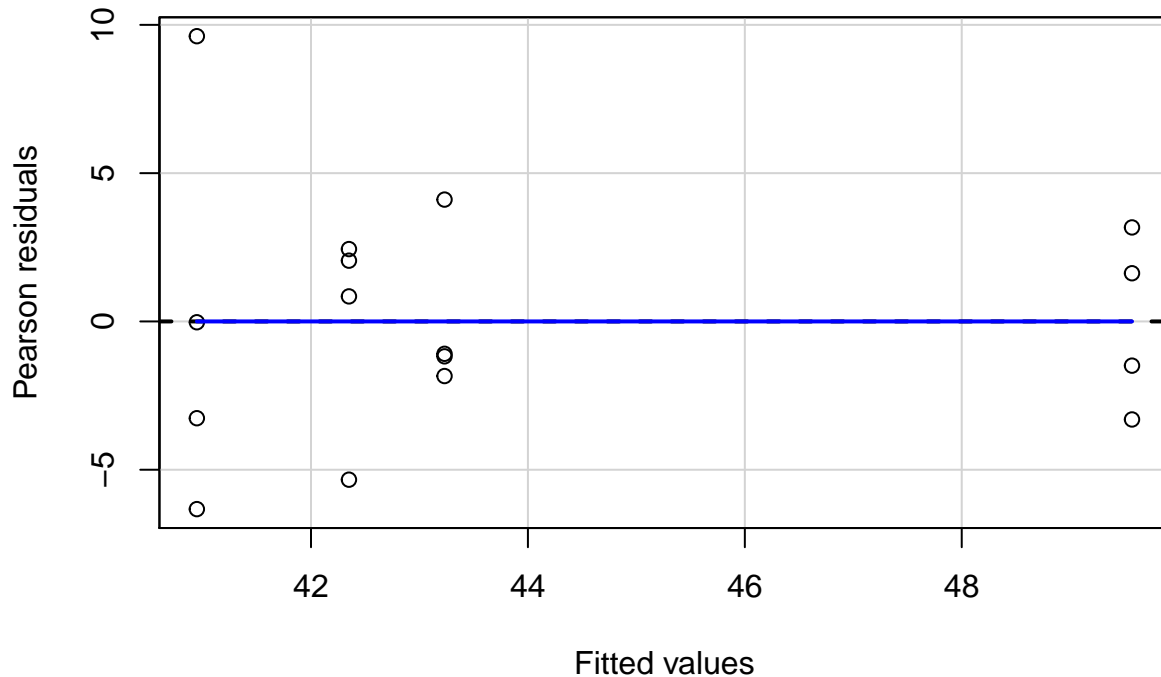
GOAT MILK

Provem el model factorial per la llet de cabra:

```
summary(mgoat1<-lm(GOAT~THERMIC*CaCl2, dd))
```

```
##
## Call:
## lm(formula = GOAT ~ THERMIC * CaCl2, data = dd)
##
## Residuals:
##      Min       1Q   Median       3Q      Max
## -6.3275 -2.1966 -0.5594  2.1475  9.6175
##
## Coefficients:
##              Estimate Std. Error t value Pr(>|t|)
## (Intercept)      49.570      2.195  22.587 3.36e-11 ***
## THERMICplane      -7.220      3.104  -2.326  0.0383 *
## CaCl2yes          -6.339      3.104  -2.042  0.0637 .
## THERMICplane:CaCl2yes  4.936      4.389   1.125  0.2827
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## Residual standard error: 4.389 on 12 degrees of freedom
## Multiple R-squared:  0.4303, Adjusted R-squared:  0.2879
## F-statistic: 3.021 on 3 and 12 DF,  p-value: 0.07157
```

```
residualPlot(mgoat1)
```

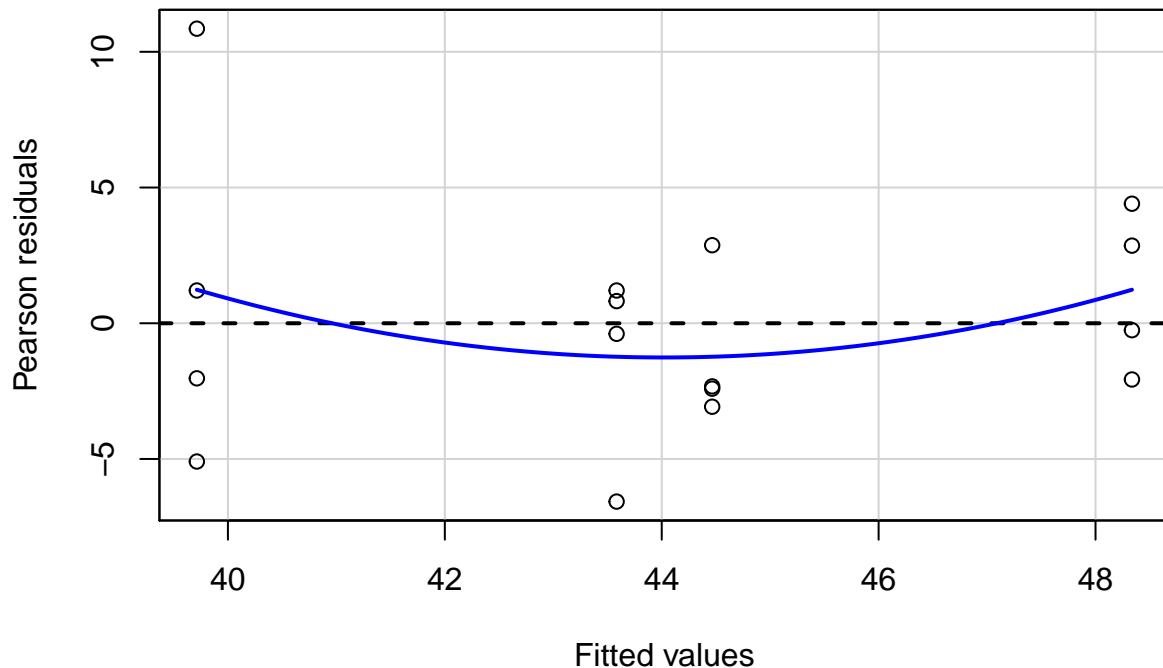


Seguidament, pel **test Omnibus** veiem que hem de parar. No ens atrevim a rebutjar la hipòtesi nul·la (no detectem que afecti res). Aquest model amb aquestes dades no va bé per modelar el rendiment de la llet de cabra. Podem provar models additius (THERMIC+CaCl2, sense interacció):

```
summary(mgoat2<-lm(GOAT~THERMIC+CaCl2, dd))
```

```
##
## Call:
## lm(formula = GOAT ~ THERMIC + CaCl2, data = dd)
##
## Residuals:
##      Min       1Q   Median       3Q      Max
## -6.5691 -2.3466 -0.3225  1.6197 10.8516
##
## Coefficients:
##              Estimate Std. Error t value Pr(>|t|)
## (Intercept)   48.336     1.920   25.177 2.05e-12 ***
## THERMICplane   -4.752     2.217   -2.143  0.0516 .
## CaCl2yes       -3.871     2.217   -1.746  0.1044
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## Residual standard error: 4.434 on 13 degrees of freedom
## Multiple R-squared:  0.3702, Adjusted R-squared:  0.2734
## F-statistic: 3.822 on 2 and 13 DF, p-value: 0.0495
```

```
residualPlot(mgoat2)
```



Arribem a una conclusió estranya:

- El **test Omnibus** ens ha sortit significatiu, el test diu que hi ha alguna diferència ($\alpha_i \neq 0$ o $\beta_i \neq 0$) però...
- Cap dels factors ens ha sortit significatiu (anova). És lògic que el test Omnibus sigui significatiu, però quan fem l'anova efecte per efecte, no detectem cap efecte? És una situació que es pot donar, el resultat és el següent: **hi ha diferències, però no sabem a on**. Això es deu a les formes de les regions d'acceptació. Cada test té la seva regió d'acceptació i és una situació quan estem en la regió d'acceptació del test Omnibus, i ens dona significatiu, però està fora de les regions d'acceptació dels dos factors. És una qüestió geomètrica que és molt difícil que es doni. Pot ser que hi hagi alguna combinació de factors que doni algun efecte.

`Anova(mgoat2)`

```
## Anova Table (Type II tests)
##
## Response: GOAT
##           Sum Sq Df F value    Pr(>F)
## THERMIC      90.321  1   4.5946 0.05156 .
## CaCl2        59.927  1   3.0485 0.10439
## Residuals  255.556 13
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
```

SHEEP MILK

El model factorial per l'ovella:

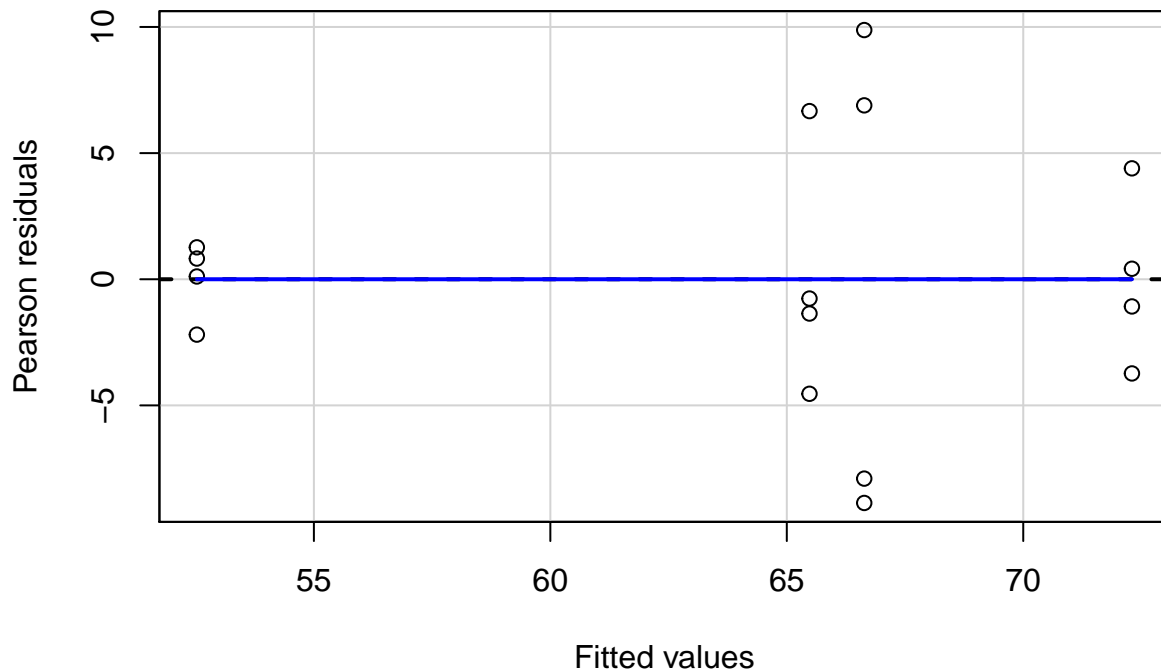
`summary(msh<-lm(SHEEP~THERMIC*CaCl2, dd))`

```
##
## Call:
## lm(formula = SHEEP ~ THERMIC * CaCl2, data = dd)
```

```
##
## Residuals:
##      Min       1Q   Median       3Q      Max
## -8.8675 -2.5797 -0.3275  2.0478  9.8775
##
## Coefficients:
##              Estimate Std. Error t value Pr(>|t|)
## (Intercept)      72.299      2.870   25.193 9.29e-12 ***
## THERMICplane     -5.661      4.059   -1.395   0.188
## CaCl2yes         -6.819      4.059   -1.680   0.119
## THERMICplane:CaCl2yes -7.294      5.740   -1.271   0.228
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## Residual standard error: 5.74 on 12 degrees of freedom
## Multiple R-squared:  0.6794, Adjusted R-squared:  0.5993
## F-statistic: 8.478 on 3 and 12 DF,  p-value: 0.002709
```

El p-valor que surt al summary és sobre el paràmetres (la parametrització), no interessa gens quan fem models amb factors (però en la regressió sí!!). Per veure si els factors són significatius o no, ho hem de mirar a l'anova. El **test Omnibus** és significatiu.

```
residualPlot(msh)
```



```
anova(msh)
```

```
## Analysis of Variance Table
##
## Response: SHEEP
##              Df Sum Sq Mean Sq F value    Pr(>F)
## THERMIC        1  346.56   346.56  10.5198 0.007042 **
## CaCl2          1  438.12   438.12  13.2988 0.003346 **
## THERMIC:CaCl2   1   53.20    53.20   1.6148 0.227899
## Residuals     12  395.33    32.94
```

```
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
```

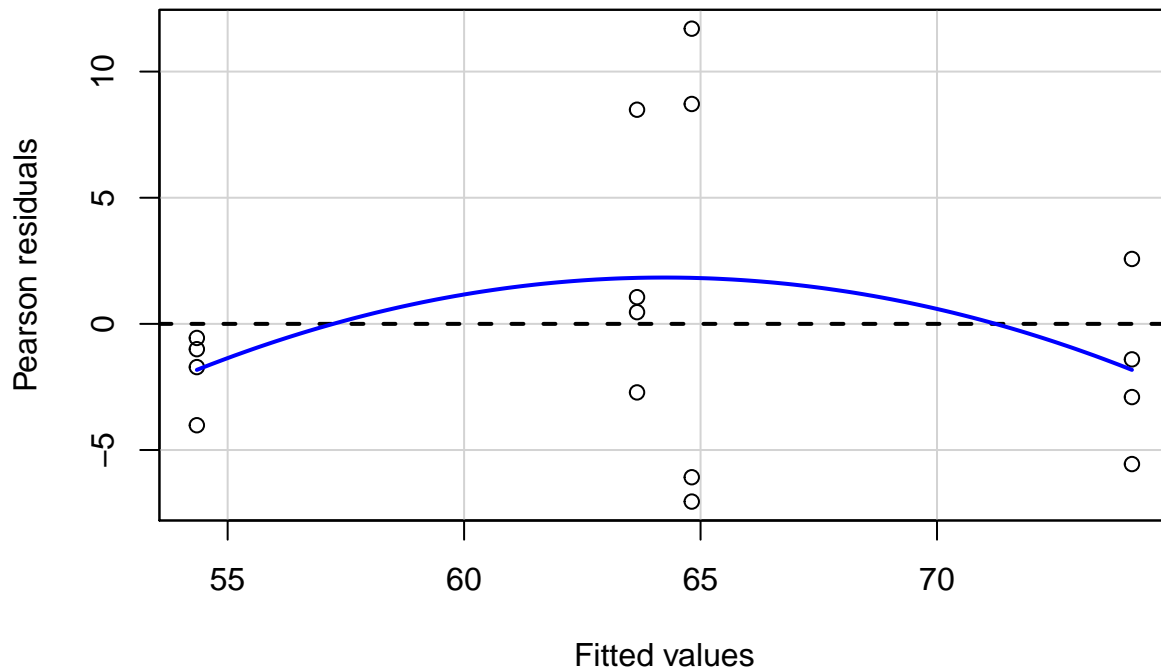
Observem que:

- (a) La interacció no és significativa.
- (b) Els dos factors són significatius.
- (c) Que no hi hagi interacció, vol dir que el model additiu pot ser vàlid. Però aquest fet no invalida el model factorial.
- (d) Com més paràmetres, menys graus de llibertat tenen els residuals (sempre!!).
- (e) Residual standard error $\hat{\sigma} = 5.74$.
- (f) Mitjana dels quadrats de l'error (anova): $\hat{\sigma}^2 = 32.94$. Que és la **variància** (= mitjana de l'error).

I el model additiu per l'ovella:

```
summary(msh2<-lm(SHEEP~THERMIC+CaCl2, dd))

##
## Call:
## lm(formula = SHEEP ~ THERMIC + CaCl2, data = dd)
##
## Residuals:
##      Min       1Q   Median       3Q      Max
## -7.044 -3.181 -1.205  1.437 11.701
##
## Coefficients:
##              Estimate Std. Error t value Pr(>|t|)
## (Intercept)    74.122     2.543   29.142 3.15e-13 ***
## THERMICplane    -9.308     2.937   -3.169  0.00739 **
## CaCl2yes       -10.466     2.937   -3.563  0.00347 **
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## Residual standard error: 5.874 on 13 degrees of freedom
## Multiple R-squared:  0.6363, Adjusted R-squared:  0.5803
## F-statistic: 11.37 on 2 and 13 DF,  p-value: 0.001396
residualPlot(msh2)
```



```
anova(msh2)
```

```
## Analysis of Variance Table
##
## Response: SHEEP
##           Df Sum Sq Mean Sq F value    Pr(>F)
## THERMIC     1 346.56   346.56   10.045 0.007391 **
## CaCl2       1 438.12   438.12   12.698 0.003466 **
## Residuals   13 448.53     34.50
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
```

Els resultats que obtenim:

- (a) Residual standard error $\hat{\sigma} = 5.874$. És una mica pitjor que en el model factorial.
- (b) El test Omnibus diu que és significatiu (= que hi ha diferències), i a l'anova veiem la mitjana de quadrats: $\frac{448.53}{13} = 34.50 = \hat{\sigma}^2$.

Resulta que els dos models són vàlids. **Quin és el millor?** Podem seguir el següent criteri:

- (a) Que la variància de l'error sigui petita: és el Residual standard error del summary al quadrat.
- (b) Que les comparacions múltiples en detectin més diferències entre l'un i l'altre.

Segons el criteri (a), el model factorial és una mica millor que l'additiu.

Anem a fer les comparacions múltiples:

```
# Model factorial
(emmsh<-emmeans(msh,~THERMIC*CaCl2))
```

```
## THERMIC CaCl2 emmean SE df lower.CL upper.CL
## pasteurized no 72.29875 2.86985 12 66.04588 78.55162
## plane no 66.63750 2.86985 12 60.38463 72.89037
## pasteurized yes 65.48000 2.86985 12 59.22713 71.73287
## plane yes 52.52500 2.86985 12 46.27213 58.77787
##
```

```
## Confidence level used: 0.95
```

```
CLD(emmsh,Letters=letters, reversed=T)
```

```
## THERMIC    CaCl2    emmean      SE df lower.CL upper.CL .group
## pasteurized no    72.29875 2.86985 12 66.04588 78.55162 a
## plane      no    66.63750 2.86985 12 60.38463 72.89037 a
## pasteurized yes   65.48000 2.86985 12 59.22713 71.73287 a
## plane      yes   52.52500 2.86985 12 46.27213 58.77787 b
##
## Confidence level used: 0.95
## P value adjustment: tukey method for comparing a family of 4 estimates
## significance level used: alpha = 0.05
```

```
# Posem la A al rendiment més alt,
# Per tant REVERSED=T
```

Ens ha sortit que entre els tres primers tractaments, no els distingim. Els tres són candidats a ser el més bo. Ens ha sortit millor pasteuritzat sense clorur càlcic, però això pot ser degut a l'atzar. Les tres tenen el rendiment més alt (*emmean*).

Anem a fer les comparacions múltiples del model additiu. Resulta que l'additiu distingeix millor (b). La millor opció és A: pasteuritzada i sense CaCl₂.

```
# Model additiu
```

```
(emmsh2<-emmeans(msh2,~THERMIC+CaCl2))
```

```
## THERMIC    CaCl2    emmean      SE df lower.CL upper.CL
## pasteurized no    74.12219 2.543455 13 68.62739 79.61699
## plane      no    64.81406 2.543455 13 59.31926 70.30886
## pasteurized yes   63.65656 2.543455 13 58.16176 69.15136
## plane      yes   54.34844 2.543455 13 48.85364 59.84324
##
## Confidence level used: 0.95
```

```
CLD(emmsh2,Letters=letters, reversed=T)
```

```
## THERMIC    CaCl2    emmean      SE df lower.CL upper.CL .group
## pasteurized no    74.12219 2.543455 13 68.62739 79.61699 a
## plane      no    64.81406 2.543455 13 59.31926 70.30886 b
## pasteurized yes   63.65656 2.543455 13 58.16176 69.15136 b
## plane      yes   54.34844 2.543455 13 48.85364 59.84324 c
##
## Confidence level used: 0.95
## P value adjustment: tukey method for comparing a family of 4 estimates
## significance level used: alpha = 0.05
```

Finalment, la decisió final... **Quin model és millor, l'additiu o el factorial?** Segons el principi de parsimònia, com més simple el model, si és adequat, millor. L'additiu és adequat i en aquest l'agafarem perquè el que perdem amb les variàncies, realment no és gaire gran, i a més distingeix millor en les comparacions. Si perdèssim molt amb les variàncies, les comparacions ja no sortirien tan bé.