Linear Model: Intelligence Quotient and Pedagogical Methods

The aim is to compare the scores (p) of two pedagogical methodologies (m) as a function of the intelligence quotient (c). Perform the parameter estimation as well as the hypothesis tests to see if the parameters are significantly different from zero or not. Moreover, perform the appropriate tests to answer the following questions:

(a) **ANCOVA**: can we consider that the two lines are not statistically different (M1 and M2)? **El mètode afecta a la puntuació? El coeficient de inteligència afecta a la puntuació? Els mètodes afecten de la mateixa manera, o de forma diferent, segons el coeficient de inteligència?**

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

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

```
##   M   C  P PP
## 1 1 110 55 55
## 2 1 121 66 66
## 3 1 108 50 50
## 4 1  95 33 33
## 5 1 107 50 50
## 6 1  89 35 35
```

```
dim(dd)
```

```
## [1] 22  4
```

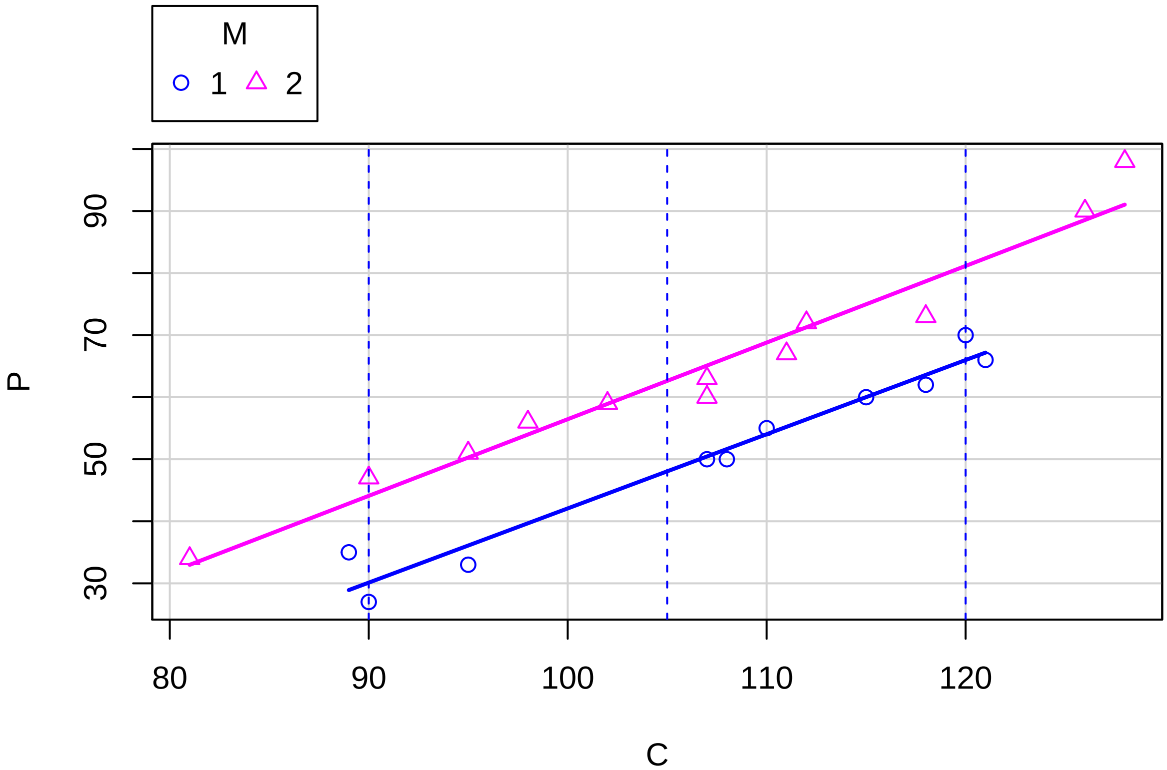
The data set contains 22 rows corresponding to individuals, and four columns. The first column corresponds to the pedagogical methodology followed by the student. The second column contains the inteligence quotient of the individual and the third and fourth columns contain scores.

We will first work with the column denoted by P. This type of analysis is called **ANCOVA** because as explanatory variables we have categorical as well as continuous variables. In this case we have one of each type. Thus, this situation corresponds to the easiest ANCOVA analysis.

Descriptive statistics

We plot the score as a function of the intelligence coefficient, using a different line for each pedagogical methodology.

```
sp(P~C|M, smooth=F, dat=dd)
abline(v=c(90,105,120),lty=2,col="blue")
```



We clearly see better results for the pedagogical method number 2. The score is always higher independetly of the intelligence coefficient. It can be seen that the lines are quite parallel ($\sim$ sembla que no hi hagi interacció). It seems that both - the intelligence quotient as well as the method - will have a significant influence upon the score.

(b) Can the two lines be considered parallel lines?
(c) Can the two intercepts be considered not statistically different?

Modelling with interaction

We start fitting a model with interaction. This means that we allow the influence of the inteligence quotient upon the score to be different in the two methods. If the interaction term is significant this would mean that the two slopes are statistically different and that the inteligence quotient has a different effect upon the scores for the two methods.

$$P_{ij} = \mu + \alpha_i + \nu C_j + \beta_i C_j + e_{ij}$$

Where: i is the method, j is the the intelligence quotient, α_i és l'efecte principal del mètode i, μ és constant per tothom, νC_j és l'efecte principal del coeficient de intel.ligència (CI) i $\beta_i C_j$ és l'efecte de la interacció *mètode-CI*.

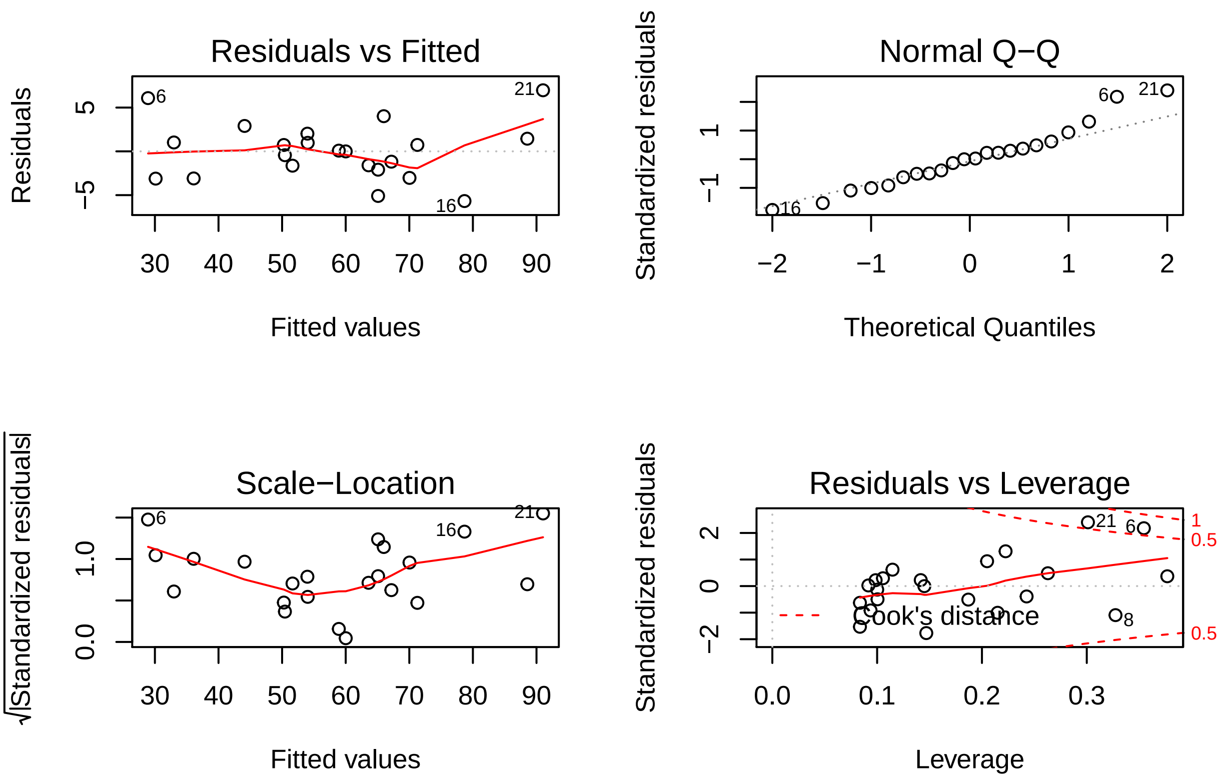
Si β_i fossin zero, tindriem dues rectes amb el mateix pendent (dues rectes paral.leles). En canvi, si $\beta_i \neq 0$ tenim dues rectes que no són paral.leles (= no hi ha interacció).

Com definim el model? Tindrem un factor que és el mètode i una covariable que és el CI. A different intercept will be interpreted as that the two methods return different results. We first define the method as a categorical variable:

```r
# We will use first the scores "P":
dd$M<-as.factor(dd$M)
summary(model<-lm(P~M*C, dd))
```

```
##
## Call:
## lm(formula = P ~ M * C, data = dd)
##
## Residuals:
##     Min      1Q  Median      3Q     Max
## -5.6767 -1.9789  0.0376  1.3367  6.9744
##
## Coefficients:
##             Estimate Std. Error t value Pr(>|t|)
## (Intercept) -77.49366   10.33232  -7.500 6.07e-07 ***
## M2           10.45276   13.05628   0.801    0.434
## C             1.19565    0.09575  12.487 2.65e-10 ***
## M2:C          0.03924    0.12133   0.323    0.750
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## Residual standard error: 3.474 on 18 degrees of freedom
## Multiple R-squared:  0.966,  Adjusted R-squared:  0.9603
## F-statistic: 170.4 on 3 and 18 DF,  p-value: 2.112e-13
```

```r
oldpar<-par(mfrow=c(2,2))
plot(model, ask=F)
```




```r
par(oldpar)
```

Omnibus test: clarament tenim diferències. We see from this first analysis that the interaction and the method 2 are not significant. Probably the method is not significant because of the interaction term. Doing the Anova type III we will see if the method is actually significant or not:

```r
Anova(model)
```

```
## Anova Table (Type II tests)
##
## Response: P
##           Sum Sq Df  F value    Pr(>F)
## M         1168.3  1  96.8287 1.144e-08 ***
## C         5192.8  1 430.3826 5.115e-14 ***
## M:C          1.3  1   0.1046    0.7501
## Residuals  217.2 18
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
```

It turns out that the method and the interaction are not significant. Let us fit the model without interaction.

Modelling without interaction

```r
summary((mP2<-lm((P~M+C), dd)))
```

```
##
## Call:
## lm(formula = (P ~ M + C), data = dd)
##
## Residuals:
##    Min     1Q Median     3Q    Max
## -5.503 -2.025 -0.088  1.529  7.296
##
## Coefficients:
##              Estimate Std. Error t value Pr(>|t|)
## (Intercept) -80.11563    6.25264  -12.81 8.50e-11 ***
## M2           14.64776    1.45307   10.08 4.62e-09 ***
## C             1.22009    0.05741   21.25 1.05e-14 ***
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## Residual standard error: 3.391 on 19 degrees of freedom
## Multiple R-squared:  0.9658, Adjusted R-squared:  0.9622
## F-statistic: 268.2 on 2 and 19 DF,  p-value: 1.186e-14
```

We see that once the interaction is suppressed, both explanatory variables are significant. Thus, we conclude that the obtained scores significatively depend on the inteligence quotient but also on the method.

The fact that the interaction is not significative means that the way in which the inteligence quotient affects upon the scores is the same for both methods. We will have thus two different lines, one for each method with the same slope but different intercepts.

Also, notice that the two variables explain 96% of the variability in the response variable.

```r
Anova(mP2)
```

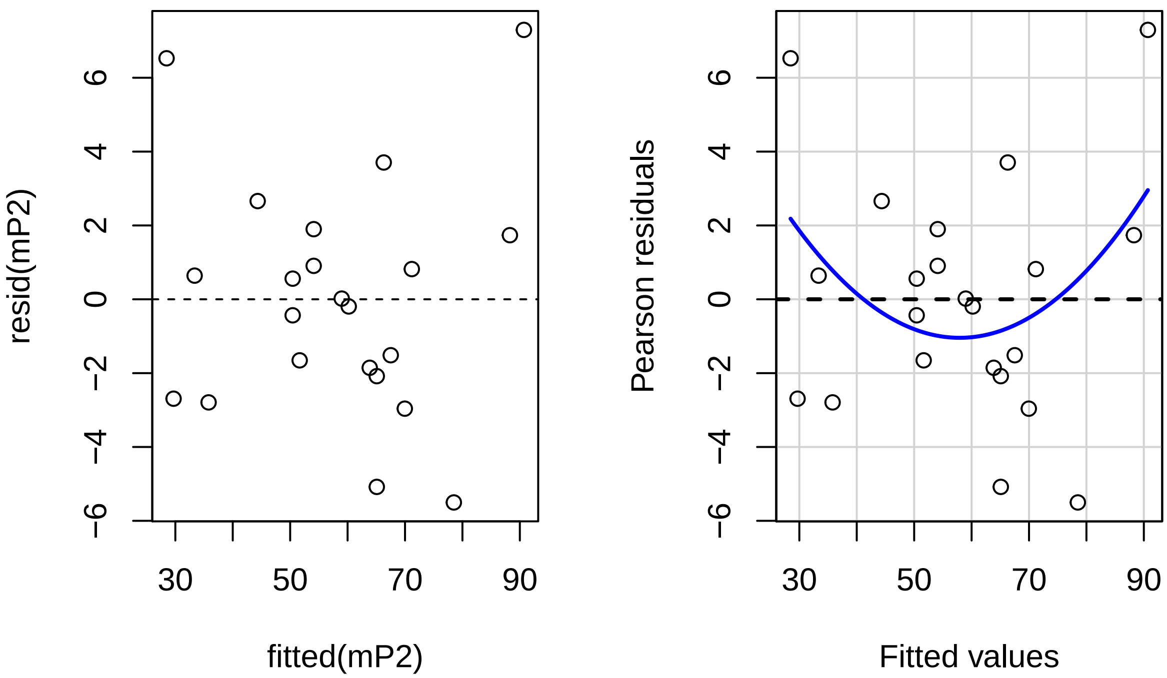
```
## Anova Table (Type II tests)
##
```

```
## Response: P
##           Sum Sq Df F value    Pr(>F)
## M         1168.3  1  101.62 4.624e-09 ***
## C         5192.8  1  451.67 1.053e-14 ***
## Residuals  218.4 19
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
```

From the Anova we also see that both explanatory variables have a significant influence upon the score. Let us see if the assumptions (hypotheses) of the linear model are satisfied.

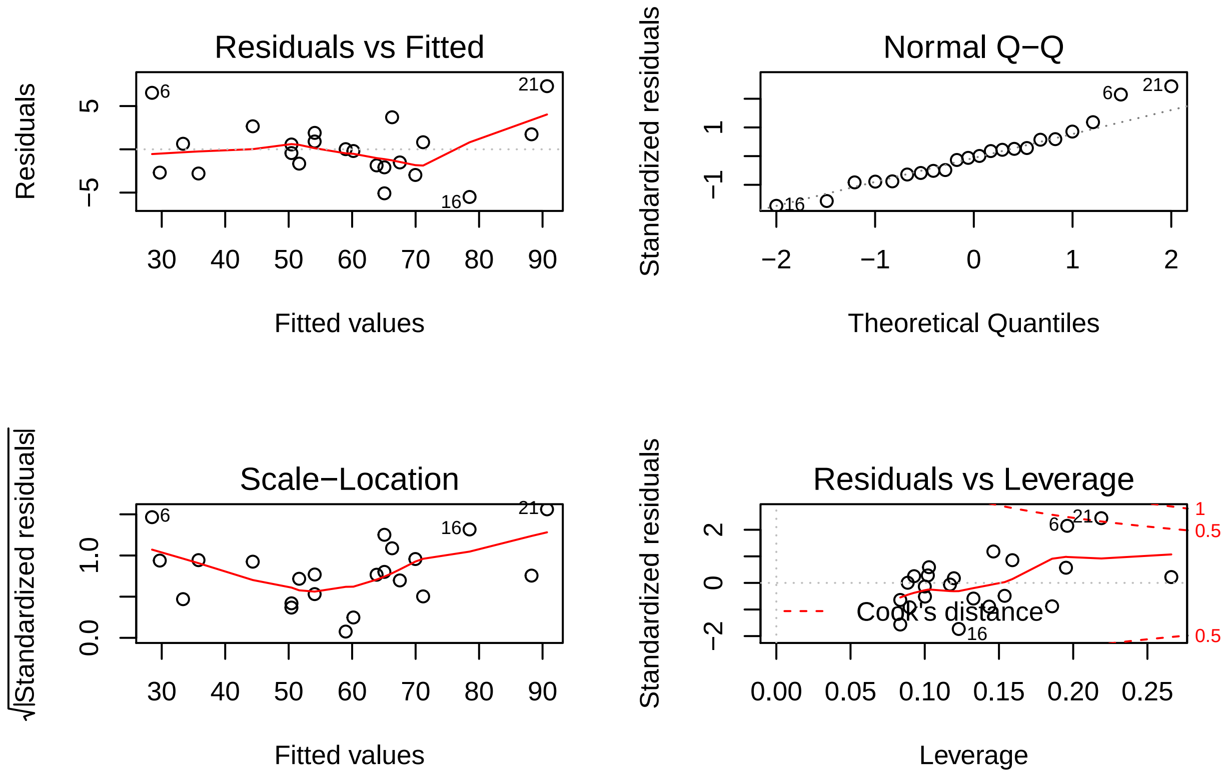
Model Diagnostics

```r
oldpar<-par(mfrow=c(1,2))
plot(fitted(mP2), resid(mP2))
abline(h=0, lty=2)
residualPlot(mP2)
```




```r
par(oldpar)
```

```r
oldpar<-par(mfrow=c(2,2))
plot(mP2, ask=F)
```




```r
par(oldpar)
```

The Normality, independence and homocedasticity assumptions may be acccepted. Which allow us to accept the additive model as a good model to explain the variability in the puctuation variable.

In what follows we compute the estimated marginal means (emmeans) for the three coefficient of inteligence levels required. Observe that the value of the method appears to be a weighted mean of one and two (the design is not balanced in this case).

(d) For each one of the following values of intelligence quotient c: 90, 105 y 120, which differences do exists in the scores for each one of the methodologies?

```r
(emm<-emmeans(mP2,~M|C, at=list(C=c(90, 105, 120))))
```

```
## C =  90:
##  M   emmean        SE df lower.CL upper.CL
##  1 29.69245 1.4615424 19 26.63340 32.75149
##  2 44.34021 1.3521782 19 41.51007 47.17035
##
## C = 105:
##  M   emmean        SE df lower.CL upper.CL
##  1 47.99379 1.0803396 19 45.73262 50.25497
##  2 62.64155 0.9814439 19 60.58737 64.69574
##
## C = 120:
##  M   emmean        SE df lower.CL upper.CL
##  1 66.29514 1.2966427 19 63.58124 69.00904
##  2 80.94290 1.2574574 19 78.31101 83.57479
##
## Confidence level used: 0.95
```

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

```
## C =  90:
##  M   emmean        SE df lower.CL upper.CL .group
##  2 44.34021 1.3521782 19 41.51007 47.17035  a
##  1 29.69245 1.4615424 19 26.63340 32.75149   b
##
## C = 105:
##  M   emmean        SE df lower.CL upper.CL .group
##  2 62.64155 0.9814439 19 60.58737 64.69574  a
##  1 47.99379 1.0803396 19 45.73262 50.25497   b
##
## C = 120:
##  M   emmean        SE df lower.CL upper.CL .group
##  2 80.94290 1.2574574 19 78.31101 83.57479  a
##  1 66.29514 1.2966427 19 63.58124 69.00904   b
##
## Confidence level used: 0.95
## significance level used: alpha = 0.05
```

```r
pairs(emm)
```

```
## C =  90:
##  contrast  estimate       SE df t.ratio p.value
##  1 - 2    -14.64776 1.453071 19 -10.081  <.0001
##
## C = 105:
##  contrast  estimate       SE df t.ratio p.value
##  1 - 2    -14.64776 1.453071 19 -10.081  <.0001
##
## C = 120:
##  contrast  estimate       SE df t.ratio p.value
##  1 - 2    -14.64776 1.453071 19 -10.081  <.0001
```

Són clarament diferents a 90, 105 i 120. Almenys en aquest rang (90-120). Podem demanar quins son els pendents?

```r
(emmp<-emtrends(mP2,~M, var="C"))
```

```
##  M C.trend         SE df lower.CL upper.CL
##  1 1.22009 0.05740927 19 1.099931 1.340249
##  2 1.22009 0.05740927 19 1.099931 1.340249
##
## Confidence level used: 0.95
```