BE 2 - Plan d'Expérience

Exercice 1

Question 1

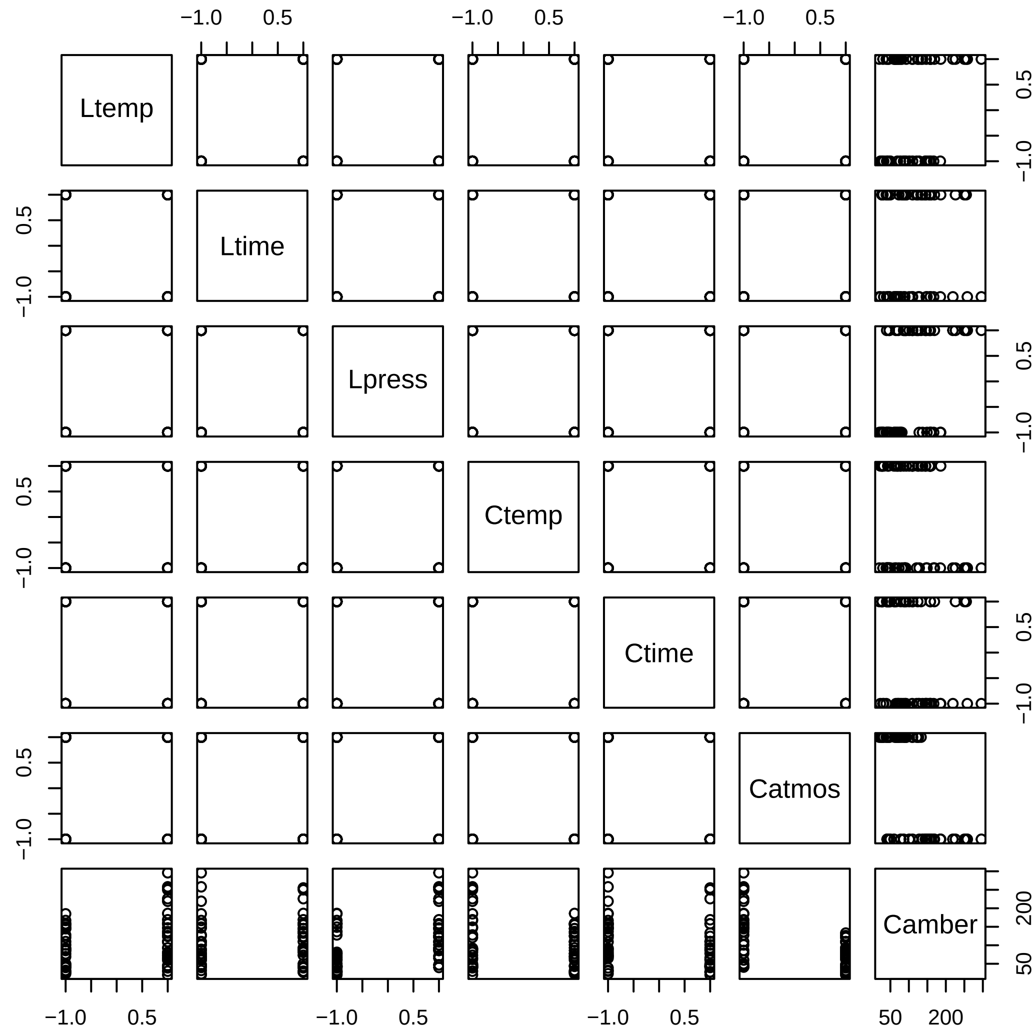
```r
silicium <- read.table("silicium.txt", header = TRUE)

head(silicium)
```

```
##   Ltemp Ltime Lpress Ctemp Ctime Catmos Camber
## 1    -1    -1     -1    -1    -1     -1    167
## 2     1    -1     -1    -1     1      1     62
## 3    -1     1     -1    -1     1     -1     41
## 4     1     1     -1    -1    -1      1     73
## 5    -1    -1      1    -1     1      1     47
## 6     1    -1      1    -1    -1     -1    219
```

```r
summary(silicium)
```

```
##      Ltemp         Ltime         Lpress        Ctemp         Ctime         Catmos
##  Min.   :-1   Min.   :-1   Min.   :-1   Min.   :-1   Min.   :-1   Min.   :-1
##  1st Qu.:-1   1st Qu.:-1   1st Qu.:-1   1st Qu.:-1   1st Qu.:-1   1st Qu.:-1
##  Median : 0   Median : 0   Median : 0   Median : 0   Median : 0   Median : 0
##  Mean   : 0   Mean   : 0   Mean   : 0   Mean   : 0   Mean   : 0   Mean   : 0
##  3rd Qu.: 1   3rd Qu.: 1   3rd Qu.: 1   3rd Qu.: 1   3rd Qu.: 1   3rd Qu.: 1
##  Max.   : 1   Max.   : 1   Max.   : 1   Max.   : 1   Max.   : 1   Max.   : 1
##      Camber
##  Min.   : 20.0
##  1st Qu.: 47.0
##  Median : 91.0
##  Mean   :107.0
##  3rd Qu.:150.5
##  Max.   :296.0
```

```r
pairs(silicium)
```




```r
X <- as.matrix(silicium[,-7])
siliciumtransposee <- t(X) %*% X
siliciumtransposee
```

```
##        Ltemp Ltime Lpress Ctemp Ctime Catmos
## Ltemp     64     0      0     0     0      0
## Ltime      0    64      0     0     0      0
## Lpress     0     0     64     0     0      0
## Ctemp      0     0      0    64     0      0
## Ctime      0     0      0     0    64      0
## Catmos     0     0      0     0     0     64
```

On remarque ici que tX.X = 64.In. Donc la matrice X est orthogonale. On dispose de 6 variables explicatives, et on peut voir que 64 expériences ont été réalisées. On pourrait donc penser à un Plan factoriel complet 2ˆp avec p = 6.

Mais il est dit dans l'énoncé que chaque expérience est répliquée quatre fois, on a donc 64/4=16 expériences soit 2^4. Le plan est donc un Plan Factoriel fractionnaire de type 2^(6-2).

```
X2 <- as.matrix(silicium[1:16,-7])
X2[,1] * X2[,2] * X2[,3] == X2[,5]
```

```
##    1    2    3    4    5    6    7    8    9   10   11   12   13   14   15   16
## TRUE TRUE TRUE TRUE TRUE TRUE TRUE TRUE TRUE TRUE TRUE TRUE TRUE TRUE TRUE TRUE
```

```
X2[,1] * X2[,3] * X2[,4] == X2[,6]
```

```
##    1    2    3    4    5    6    7    8    9   10   11   12   13   14   15   16
## TRUE TRUE TRUE TRUE TRUE TRUE TRUE TRUE TRUE TRUE TRUE TRUE TRUE TRUE TRUE TRUE
```

On retrouve bien A.B.C = E et A.C.D = F.

Les effets principaux peuvent être estimés sans confusions, en revanche les effets secondaires ne peuvent pas être estimés sans confusions. Le plan est donc de résolution IV.

Question 2

```
siliciumLm <- lm(Camber ~ Ltemp + Ltime + Lpress + Ctemp + Ctime + Catmos, data = silicium)

summary(siliciumLm)
```

```
##
## Call:
## lm(formula = Camber ~ Ltemp + Ltime + Lpress + Ctemp + Ctime +
##     Catmos, data = silicium)
##
## Residuals:
##     Min      1Q  Median      3Q     Max
## -67.672 -22.703  -3.875  28.797  81.328
##
## Coefficients:
##             Estimate Std. Error t value Pr(>|t|)
## (Intercept)  107.016      4.793  22.328  < 2e-16 ***
## Ltemp         19.453      4.793   4.059 0.000152 ***
## Ltime          2.891      4.793   0.603 0.548823
## Lpress        28.016      4.793   5.845 2.57e-07 ***
## Ctemp         -7.109      4.793  -1.483 0.143492
## Ctime        -17.234      4.793  -3.596 0.000676 ***
## Catmos       -38.734      4.793  -8.082 5.03e-11 ***
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## Residual standard error: 38.34 on 57 degrees of freedom
## Multiple R-squared:  0.6975, Adjusted R-squared:  0.6657
## F-statistic: 21.91 on 6 and 57 DF,  p-value: 3.566e-13
```

La p-value du modèle est 3.566e-13 « 0,05, ce qui signifie que ce modèle est significatif. On remarque aussi que les Variables Ltime et Ctemp ont des p-values > 0,05. On peut donc envisager de les retirer dans notre modèle.

Question 3

On sait que l'erreur standard peut s'écrire :

$$\text{StdError}_i = \sqrt{\sigma^2 \left[(X^T X)^{-1}\right]_{ii}}$$

avec σ = Residual Standard Error.
On a montré précédemment que $X^T X = 64 \cdot I$.

En remplaçant $\sigma = 38.34$:

$$\text{StdError}_i = \frac{38.34}{\sqrt{64}} = 4.793$$

Et ce résultat est valable pour toutes les variables.

Question 4

On peut supprimer les deux variables Ltime et Ctemp d'un coup, il n'est pas nécessaire de faire un algorithme backward car tX.X est diagonale, il n'y a pas de corrélation.

```r
siliciumLm_reduit <- lm(Camber ~ Ltemp + Lpress + Ctime + Catmos, data = silicium)

summary(siliciumLm_reduit)
```

```
##
## Call:
## lm(formula = Camber ~ Ltemp + Lpress + Ctime + Catmos, data = silicium)
##
## Residuals:
##     Min      1Q  Median      3Q     Max
## -74.953 -26.547  -4.016  25.844  85.547
##
## Coefficients:
##             Estimate Std. Error t value Pr(>|t|)
## (Intercept)  107.016      4.816  22.222  < 2e-16 ***
## Ltemp         19.453      4.816   4.040 0.000157 ***
## Lpress        28.016      4.816   5.818 2.59e-07 ***
## Ctime        -17.234      4.816  -3.579 0.000698 ***
## Catmos       -38.734      4.816  -8.043 4.62e-11 ***
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## Residual standard error: 38.53 on 59 degrees of freedom
## Multiple R-squared:  0.6839, Adjusted R-squared:  0.6625
## F-statistic: 31.92 on 4 and 59 DF,  p-value: 3.711e-14
```

Les coefficients sont identiques aux coefficients du modèle précédent. Cela est une nouvelle fois cohérent avec le fait qu'il n'y a pas de corrélation entre les différents paramètres.

Question 5

Les conditions expérimentales qui permettent de minimiser la courbure Camber sont celles où l'on prend la valeur minimale (-1) pour les coefficients positifs et la valeur maximale (+1) pour les coefficients négatifs. Les coefficients étant ceux que l'on peut lire dans la colonne Estimate du Summary.

On peut donc trouver un intervalle de confiance de la manière suivante :

```
frame <- data.frame(Ltemp = -1, Lpress = -1, Ctime = 1, Catmos = 1)
predict(siliciumLm_reduit, frame, interval = "confidence", level = 0.95)
```

```
##        fit      lwr      upr
## 1 3.578125 -17.9689 25.12515
```

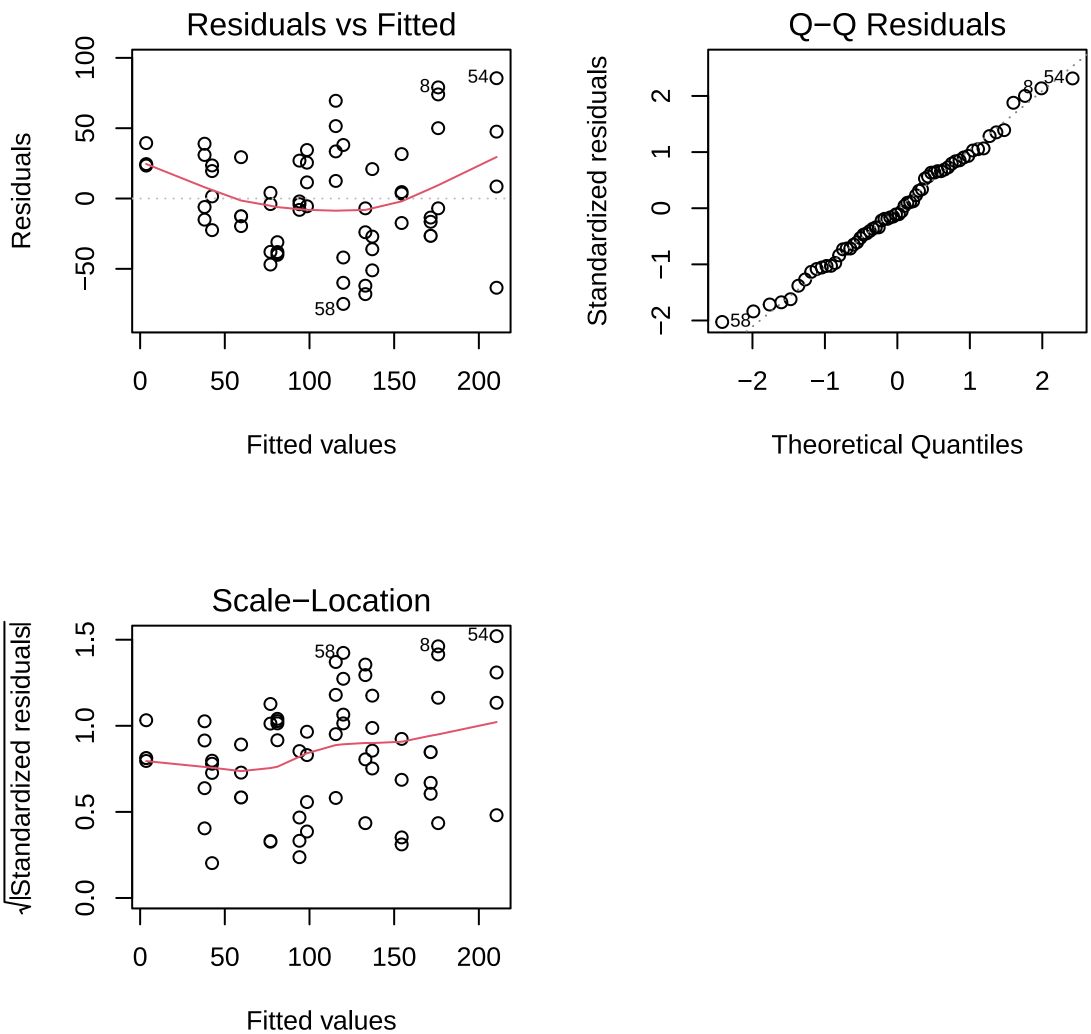
Un intervalle de confiance pour la courbure moyenne en ce point de fonctionnement optimal est donc [-17.9689 , 25.12515]. Or la courbure ne peut pas être négative, donc on prend la valeur absolue de cet intervalle : [0 , 25.12515]

La température de laminage varie entre 55 et 75 °C correspondant dans à une variation entre -1 et 1 dans le modèle. Le coefficient de cette variable est 19,453. Donc une augmentation de 5°C de la température de Laminage entraine une variation de la courbre de 2 * 19,453 / 4 = 9,7265.

Question 6

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

```
## hat values (leverages) are all = 0.078125
##  and there are no factor predictors; no plot no. 5
```



Sur le graphe Residuals vs Fitted, on observe que les résidus sont centrés sur 0, leur esperance semble bien nulle. En revanche, on observe une légère tendance en forme de U dans la répartition des résidus. Ils ne semblent donc pas réellement uniforméments répartis autour de 0.

Sur le Graphe Q-Q residuals, on remarque que les points suivent dans l'ensemble relativement bien la diagonale et sont alignés, à l'exception des extrémités.

Le modèle semble donc acceptable vis-à-vis de ses hypothèses.