Linear Statistical Modeling

Compulsory assignment 3

August Arnstad, Felix Baastad Berg, Markus Stokkenes, Ulrik Unneberg

1 Introduction

Humans' desire for coffee and tea is ever-increasing. According to a study performed in the United States [1], coffee was the most consumed beverage domestically, even moreso than tap water. When ordering a hot beverage, a short waiting time is desirable. We have therefore investigated which factors affect the waiting time significantly.

We know from personal experience that some materials isolate heat more efficiently than others, and that boiling temperatures vary across types of liquid. To better quantify the problem at hand, we have talked to an employee at Café SITO, who stated that a hot beverage should be served at 75 °C. Thus, we would like to determine factors which are crucial to the cooldown time of the beverage.

2 Factors and levels

We choose to investigate three factors: type of cup, type of beverage, and presence of a lid. We expect these to be significant for our problem. It is possible that interactions between the factors can have an effect on the problem, for instance interaction between cup and lid.
For each factor, two levels are selected. The cup is either paper or porcelain, the beverage is either tea or coffee, and the lid is either on or off. These are natural choices, as paper and porcelain are the most common cup types, and coffee and tea are the most common hot beverage types. Since the factors are discrete, it is straightforward to control their levels.

3 Response

The response variable is chosen to be time for the beverage to reach a temperature of 75 °C, starting at 85 °C. This response should be sufficient to yield a satisfying conclusion of our problem. A simple stopwatch was used to measure time, and a digital thermometer [2] was used to measure temperature. We consider the accuracy of the measurements to be quite high, as they are performed with digital instruments with easily interpretable displays. If we had used an analog thermometer, the story would perhaps be different. Therefore, we attribute any measurement errors to human vision and reaction time.

4 Choice of design

We design this study as a replicated 2^3 factorial experiment. For each of the 8 combinations of factor levels we perform 2 measurements, which in total yields 16 observations of the response. The replicates are necessary, since 8 observations might not be sufficient to show any statistical significance. Realistically, the experiment can be performed in a single location and over a short time period. It does not require any skill, and the enviromental conditions can be made more or less homogeneous. Thus, we have decided that blocking of the experiment is unnecessary.

5 Implementation

To evoke independence of the measurements, the order of the trials was randomized by simple R code. The main issue was to reset the setup before each run. Specifically, we needed to ensure that the porcelain cup was not heated from the previous run by showering it with room-temperature water. Additionally, we focused on placing the thermometer probe at roughly the same depth throughout all the trials. Marks were made on the interior of the cups to ensure an equal volume of water in each trial, regardless of the type of cup. The thermometer was mounted to a fixed position on the wall.

Each trial was then performed as follows:

- Rinse and cool the cup to room temperature
- Place the tea bag or coffee powder in the cup
- Boil water to 90° C
- Fill the cup according to the marked height
- Place the probe (through the lid if necessary)
- When the thermometer displays 85° C, start the timer
- When the thermometer displays 75° C, stop the timer and record the number of seconds

By this standard we feel that we have replicated the variability at a satisfying level and conducted each trial independently. The experiment was done at a common kitchen at NTNU, which adds some variability as people are free to use the kitchen during the experiment. Other factors that might cause some variability is the accuracy of both the thermometer and the water boiler, as previously discussed. These factors are included in the model as random noise, since they are outside of our control (irreducible error).

6 Data analysis

6.1 Calculation of effects

Once all the trials were completed, we obtained the following data:

cup	lid	beverage	time (s)
porcelain	yes	tea	412.84
paper	no	tea	233.62
porcelain	no	coffee	179.25
porcelain	yes	tea	370.23
paper	yes	tea	559.96
paper	yes	coffee	525.04
paper	no	tea	209.64
porcelain	no	tea	230.95
porcelain	yes	coffee	432.72
porcelain	yes	coffee	467.51
porcelain	no	tea	256.37
paper	yes	tea	531.09
paper	yes	coffee	557.11
porcelain	no	coffee	206.09
paper	no	coffee	266.87
paper	no	coffee	239.73

Table 1: Data in chronological order

Now we would like to analyze this data using a multiple linear regression model

$$\mathbf{Y} = X\boldsymbol{\beta} + \boldsymbol{\varepsilon},$$

where

$$\mathbf{Y} = \begin{bmatrix} Y_1 & \dots & Y_{16} \end{bmatrix}^{\mathrm{T}}$$

is the (sorted) vector of time measurements,

$$X = \begin{bmatrix} \mathbf{1} & \mathbf{x}_{\text{cup}} & \mathbf{x}_{\text{lid}} & \mathbf{x}_{\text{bev}} & \mathbf{x}_{\text{cup}}\mathbf{x}_{\text{lid}} & \mathbf{x}_{\text{cup}}\mathbf{x}_{\text{bev}} & \mathbf{x}_{\text{lid}}\mathbf{x}_{\text{bev}} & \mathbf{x}_{\text{cup}}\mathbf{x}_{\text{lid}}\mathbf{x}_{\text{bev}} \end{bmatrix}$$

is the data matrix (including interactions),

$$\boldsymbol{\beta} = \begin{bmatrix} \beta_0 & \beta_{\text{cup}} & \beta_{\text{lid}} & \beta_{\text{bev}} & \beta_{\text{cup, lid}} & \beta_{\text{cup, bev}} & \beta_{\text{lid, bev}} & \beta_{\text{cup, lid, bev}} \end{bmatrix}^{\mathrm{T}}$$

is the vector of regression coefficients and

$$\boldsymbol{\varepsilon} = \begin{bmatrix} \varepsilon_1 & \dots & \varepsilon_{16} \end{bmatrix}^{\mathrm{T}}$$

is the vector of errors, with $\boldsymbol{\varepsilon} \sim \mathcal{N}(\mathbf{0}, \sigma^2 I)$.

Coding the variables as paper/porcelain = no/yes = tea/coffee = -1/+1 and sorting the rows, we can write the linear regression as follows:

$$\underbrace{\begin{bmatrix} 209.64 \\ 233.62 \\ 230.95 \\ 256.37 \\ 531.09 \\ 559.96 \\ 370.23 \\ 412.84 \\ 239.73 \\ 266.87 \\ 179.25 \\ 206.09 \\ 525.04 \\ 557.11 \\ 432.72 \\ 467.51 \end{bmatrix}}_{\mathbf{Y}} = \underbrace{\begin{bmatrix} +1 & -1 & -1 & -1 & +1 & +1 & +1 & -1 \\ +1 & -1 & -1 & -1 & +1 & +1 & +1 & -1 \\ +1 & +1 & -1 & -1 & -1 & -1 & +1 & +1 \\ +1 & +1 & -1 & -1 & -1 & -1 & +1 & +1 \\ +1 & -1 & +1 & -1 & -1 & +1 & -1 & +1 \\ +1 & -1 & +1 & -1 & -1 & +1 & -1 & +1 \\ +1 & +1 & +1 & -1 & +1 & -1 & -1 & -1 \\ +1 & +1 & +1 & -1 & +1 & -1 & -1 & -1 \\ +1 & -1 & -1 & +1 & +1 & -1 & -1 & +1 \\ +1 & -1 & -1 & +1 & +1 & -1 & -1 & +1 \\ +1 & +1 & -1 & +1 & -1 & +1 & -1 & -1 \\ +1 & +1 & -1 & +1 & -1 & +1 & -1 & -1 \\ +1 & -1 & +1 & +1 & -1 & -1 & +1 & -1 \\ +1 & -1 & +1 & +1 & -1 & -1 & +1 & -1 \\ +1 & +1 & +1 & +1 & +1 & +1 & +1 & +1 \\ +1 & +1 & +1 & +1 & +1 & +1 & +1 & +1 \end{bmatrix}}_{X} \underbrace{\begin{bmatrix} \beta_0 \\ \beta_{\text{cup}} \\ \beta_{\text{lid}} \\ \beta_{\text{bev}} \\ \beta_{\text{cup, lid}} \\ \beta_{\text{cup, bev}} \\ \beta_{\text{lid, bev}} \\ \beta_{\text{cup, lid, bev}} \end{bmatrix}}_{\boldsymbol{\beta}} + \underbrace{\begin{bmatrix} \varepsilon_1 \\ \varepsilon_2 \\ \varepsilon_3 \\ \varepsilon_4 \\ \varepsilon_5 \\ \varepsilon_6 \\ \varepsilon_7 \\ \varepsilon_8 \\ \varepsilon_9 \\ \varepsilon_{10} \\ \varepsilon_{11} \\ \varepsilon_{12} \\ \varepsilon_{13} \\ \varepsilon_{14} \\ \varepsilon_{15} \\ \varepsilon_{16} \end{bmatrix}}_{\boldsymbol{\varepsilon}}$$

Note that sorting the measurements was not strictly necessary, but it makes for a nice structured and systematic data matrix as displayed above.

The general form of least squares estimator of $\boldsymbol{\beta}$ is

$$\hat{\boldsymbol{\beta}} = (X^{\mathrm{T}}X)^{-1}X^{\mathrm{T}}\mathbf{Y}.$$

Due to the special form of the design matrix in our case, this yields

$$\hat{\beta}_0 = \frac{1}{16}\sum_{i=1}^{16} Y_i = \bar{Y}$$

$$\hat{\beta}_j = \frac{1}{16}\sum_{x_{ij}=+1} Y_i - \frac{1}{16}\sum_{x_{ij}=-1} Y_i = \frac{1}{2}\bar{Y}^+ - \frac{1}{2}\bar{Y}^-,$$

which we readily compute as

$$\begin{aligned}\hat{\beta}_0 = \frac{1}{16}(&209.64 + 233.62 + 230.95 + 256.37 + 531.09 + 559.96 + 370.23 + 412.84 + \\ &+ 239.73 + 266.87 + 179.25 + 206.09 + 525.04 + 557.11 + 432.72 + 467.51) = 354.93875\end{aligned}$$

$$\begin{aligned}\hat{\beta}_{\text{cup}} =& \frac{1}{8}(230.95 + 256.37 + 370.23 + 412.84 + 179.25 + 206.09 + 432.72 + 467.51) - \\ &- \frac{1}{8}(209.64 + 233.62 + 531.09 + 559.96 + 239.73 + 266.87 + 525.04 + 557.11) = -35.44375\end{aligned}$$

$$\begin{aligned}\hat{\beta}_{\text{lid}} =& \frac{1}{8}(531.09 + 559.96 + 370.23 + 412.84 + 525.04 + 557.11 + 432.72 + 467.51) - \\ &- \frac{1}{8}(209.64 + 233.62 + 230.95 + 256.37 + 239.73 + 266.87 + 179.25 + 206.09) = 127.12375\end{aligned}$$

$$\begin{aligned}\hat{\beta}_{\text{bev}} =& \frac{1}{8}(239.73 + 266.87 + 179.25 + 206.09 + 525.04 + 557.11 + 432.72 + 467.51) - \\ &- \frac{1}{8}(209.64 + 233.62 + 230.95 + 256.37 + 531.09 + 559.96 + 370.23 + 412.84) = 4.35125\end{aligned}$$

$$
\begin{aligned}
\hat{\beta}_{\text{cup, lid}} =& \frac{1}{8}(209.64 + 233.62 + 370.23 + 412.84 + 239.73 + 266.87 + 432.72 + 467.51) - \\
&- \frac{1}{8}(230.95 + 256.37 + 531.09 + 559.96 + 179.25 + 206.09 + 525.04 + 557.11) = -25.79375
\end{aligned}
$$

$$
\begin{aligned}
\hat{\beta}_{\text{cup, bev}} =& \frac{1}{8}(209.64 + 233.62 + 531.09 + 559.96 + 179.25 + 206.09 + 432.72 + 467.51) - \\
&- \frac{1}{8}(230.95 + 256.37 + 370.23 + 412.84 + 239.73 + 266.87 + 525.04 + 557.11) = -2.45375
\end{aligned}
$$

$$
\begin{aligned}
\hat{\beta}_{\text{lid, bev}} =& \frac{1}{8}(209.64 + 233.62 + 230.95 + 256.37 + 525.04 + 557.11 + 432.72 + 467.51) - \\
&- \frac{1}{8}(531.09 + 559.96 + 370.23 + 412.84 + 239.73 + 266.87 + 179.25 + 206.09) = 9.18125
\end{aligned}
$$

$$
\begin{aligned}
\hat{\beta}_{\text{cup, lid, bev}} =& \frac{1}{8}(230.95 + 256.37 + 531.09 + 559.96 + 239.73 + 266.87 + 432.72 + 467.51) - \\
&- \frac{1}{8}(209.64 + 233.62 + 370.23 + 412.84 + 179.25 + 206.09 + 525.04 + 557.11) = 18.21125
\end{aligned}
$$

The main effects and interactions are simply these coefficients doubled, and are summarized here:

cup	lid	bev	cup, lid	cup, bev	lid, bev	cup, lid, bev
-70.8875	254.2475	8.7025	-51.5875	-4.9075	18.3625	36.4225

Table 2: Main effects and interactions

6.2 Statistical significance of effects

To double check our calculations and determine significant effects, we run the linear regression in R:

```
> df <- read.csv("data_sorted.csv", sep = ";")
> model  <- lm(time ~ .^3, data = df)
> summary(model)

Call:
lm(formula = time ~ .^3, data = df)

Residuals:
   Min     1Q Median     3Q    Max 
-21.30 -13.79   0.00  13.79  21.30 

Coefficients:
            Estimate Std. Error t value Pr(>|t|)    
(Intercept)  354.939      5.436  65.289 3.37e-12 ***
cup          -35.444      5.436  -6.520 0.000184 ***
lid          127.124      5.436  23.384 1.19e-08 ***
bev            4.351      5.436   0.800 0.446604    
cup:lid      -25.794      5.436  -4.745 0.001455 ** 
cup:bev       -2.454      5.436  -0.451 0.663721    
lid:bev        9.181      5.436   1.689 0.129726    
cup:lid:bev   18.211      5.436   3.350 0.010083 *  
---
Signif. codes:  0 ‘***’ 0.001 ‘**’ 0.01 ‘*’ 0.05 ‘.’ 0.1 ‘ ’ 1

Residual standard error: 21.75 on 8 degrees of freedom
Multiple R-squared:  0.9874,	Adjusted R-squared:  0.9764 
F-statistic: 89.53 on 7 and 8 DF,  p-value: 5.772e-07
```

Figure 1: Summary of the full model

The p-values in the rightmost column suggest that the significant effects are the main effects of cup and lid, the interaction between cup and lid, and the interaction between cup, lid and beverage. The beverage variable is clearly the odd one out in this model, since a 0.05-significance t-test does not provide sufficient evidence to reject the null hypothesis that β_{bev} is zero.

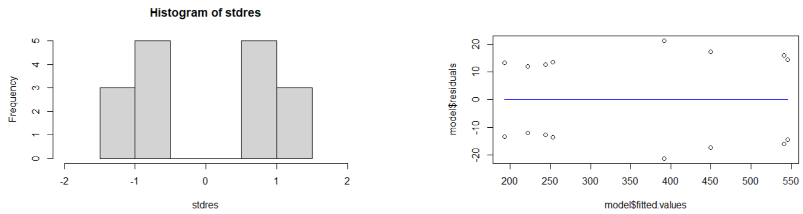
We can also observe that the null hypothesis: "all regression coefficients are identically zero" can safely be rejected by performing an F-test. The results of this test are displayed on the bottom of the regression output, and we can thus draw the conclusion that our model is significant, since the p-value is much smaller than 0.05.

6.3 Residual analysis

The model assumptions of multiple linear regression are:

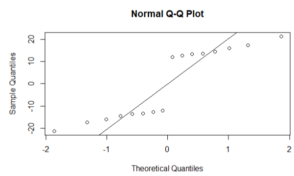
- X is of full column rank
- $\boldsymbol{\varepsilon} \sim \mathcal{N}(\mathbf{0}, \sigma^2 I)$, i.e., $\varepsilon_i \sim \mathcal{N}(0, \sigma^2)$ for all i, with ε_i and ε_j independent for $i \neq j$.

The first point is easy to verify, since the columns of the data matrix are orthogonal, so they are necessarily linearly independent and thus span the whole column space of X.
The assumptions about the error can be checked qualitatively by inspecting plots based on the residuals.



(a) (b)

Figure 2: Histogram of standardized residuals (left) and residuals plotted against fitted values (right)



(a)

Figure 3: QQ-plot

The most striking features of these plots are the clear separation and symmetry in the residuals. The symmetry is explained by the fact that we have two measurements for each configuration. Each pair of observations will be fitted to the mean of that pair, which is exactly their midpoint. This means that the measured values of each pair will be equally distant from their fitted value, but on opposite sides. The result is 8 residual pairs where the values within each pair is equal in magnitude, but with opposite signs.

To explain this mathematically, we first compute the hat matrix:

$$
\begin{aligned}
H &= X(X^{\mathrm{T}}X)^{-1}X^{\mathrm{T}} \\
&= X(16I_8)^{-1}X^{\mathrm{T}} \\
&= \frac{1}{16}XX^{\mathrm{T}} \\
&= \frac{1}{16}8 \ \mathrm{diag}(J_2, \ldots, J_2) \\
&= \frac{1}{2}\begin{bmatrix} J_2 & O_2 & \ldots & O_2 \\ O_2 & J_2 & \ldots & O_2 \\ \vdots & \vdots & \ddots & \vdots \\ O_2 & O_2 & \ldots & J_2 \end{bmatrix},
\end{aligned}
$$

which is a (16×16) block diagonal matrix where J_2 and O_2 are the 2×2 matrices consisting of ones and zeroes, respectively. The second equality follows from orthogonal columns of X, and the fourth equality from twice repeated orthogonal rows.

Thus, the residuals are given by

$$
\begin{aligned}
\hat{\boldsymbol{\varepsilon}} &= (I - H)\mathbf{Y} \\
&= \frac{1}{2}\begin{bmatrix} J_2^{\pm} & O_2 & \dots & O_2 \\ O_2 & J_2^{\pm} & \dots & O_2 \\ \vdots & \vdots & \ddots & \vdots \\ O_2 & O_2 & \dots & J_2^{\pm} \end{bmatrix} \begin{bmatrix} \mathbf{Y}_{1,2} \\ \mathbf{Y}_{3,4} \\ \vdots \\ \mathbf{Y}_{15,16} \end{bmatrix},
\end{aligned}
$$

where $J_2^{\pm} = \begin{bmatrix} 1 & -1 \\ -1 & 1 \end{bmatrix}$ and $\mathbf{Y}_{i,i+1} = \begin{bmatrix} Y_i \\ Y_{i+1} \end{bmatrix}$ for $i = 1, 3, 5, \dots, 15$. Carrying over the same notation to $\hat{\boldsymbol{\varepsilon}}$, we have the following:

$$
\begin{aligned}
\hat{\boldsymbol{\varepsilon}}_{i,i+1} &= \frac{1}{2} J_2^{\pm} Y_{i,i+1} \\
&= \frac{1}{2}\begin{bmatrix} 1 & -1 \\ -1 & 1 \end{bmatrix} \begin{bmatrix} Y_i \\ Y_{i+1} \end{bmatrix} \\
&= \frac{1}{2}\begin{bmatrix} Y_i - Y_{i+1} \\ -(Y_i - Y_{i+1}) \end{bmatrix}
\end{aligned}
$$

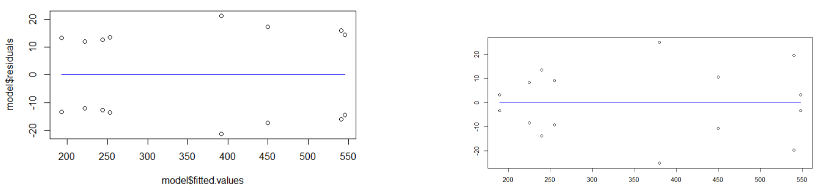
This shows that $\hat{\varepsilon}_i$ and $\hat{\varepsilon}_{i+1}$ are equal in magnitude with opposite signs for $i = 1, 3, 5, \dots, 15$.

What is more surprising, is the aforementioned separation of the residuals into two distinct clusters. After taking a closer look at our measurements, we notice a similar pattern, which is a gap within the measurement pairs. This gap is consistently around 20-30 seconds, even though each pair is two replications with identical covariate configurations. What seems strange, is that this time-gap appears to be clustered around 20-30 seconds, with smaller variations than expected.

Our theoretical model is that $Y_i = f(X_i) + \epsilon_i$ for normal and independent ϵ_i. Since $X_i = X_{i+1}$, for $i = 1, 3 \dots 15$, we would expect $\hat{\epsilon}_i$, $i = 1, 3, \dots, 15$ to be normally distributed, since

$$
\hat{\epsilon}_i = \frac{1}{2}(Y_i - Y_{i+1}) = \frac{1}{2}\Big(f(X_i) + \epsilon_i - \Big[f(\underbrace{X_{i+1}}_{=X_i}) + \epsilon_{i+1}\Big]\Big) = \frac{1}{2}(\epsilon_i - \epsilon_{i+1}),
$$

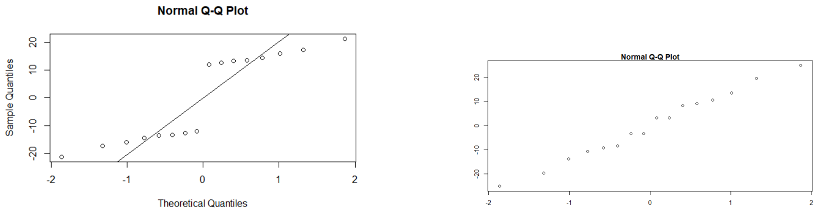
a linear combination of independent normal variables. To demonstrate an expected result, let us draw 8 normal random variables, and since $\hat{\epsilon}_i = -\hat{\epsilon}_{i+1}$, $i = 1, 3, ..., 15$, we in addition plot the negative of the obtained results, to portray the symmetry. A plot below demonstrates the difference between obtained results versus an example of a more expected result, according to the model assumptions of i.i.d. normally distributed errors.



(a) (b)

Figure 4: Obtained residual plot (a) versus sample of actual normal $\hat{\epsilon}_i$, with $\hat{\epsilon}_i = -\hat{\epsilon}_{i+1}$, $i = 1, 3, \dots, 15$, (b). We see that the variation of $|\hat{\epsilon}_i|$ in the left plot is significantly smaller than that of the right plot.

This small variance of $|\hat{\epsilon}_i|$ is also prominent in the QQ-plot, as the slopes on both sides of the theoretical zero-quantile is small. Had the variance been greater, the slopes would also increase, resulting in a more continuous-looking QQ-plot. Below is a comparison of our obtained QQ-plot and the QQ-plot of the example residual distribution presented in Figure 4b).



(a) (b)

Figure 5: QQ-plot obtained (a) versus QQ-plot of residuals shown in Figure 4b) (b)

The symmetry around the theoretical zero-quantile is simply a result of $\hat{\epsilon}_i = -\hat{\epsilon}_{i+1}, i = 1, 3, \ldots, 15$.

We unfortunately do not know where this seemingly systematic difference in the measurement pairs originates. As all the measurements were performed in random order, it is hard to imagine lack of a binary variable as a possibility. Due to the small number of observations it is possible that the time-gap's distribution simply is a matter of chance.
To summarize, the assumption of zero mean errors is fine, but the assumption that they are normally distributed is inconclusive with this small data set. We should therefore be cautious about our inference and conclusions.

7 Conclusion

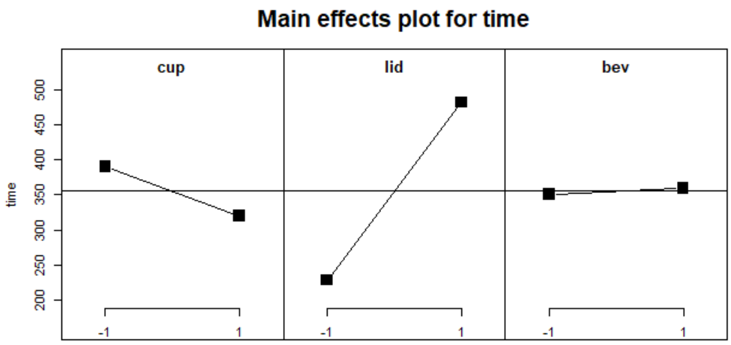


Figure 6: Main effects plot for time

A main effects plot shows the average response for when each variable is high/low. Such a plot is a good way to see if a variable has an effect on the response, as a horizontal line means that the response does not change if the variable changes from to high to low or vice versa. This supports our claim that beverage is a covariate that does not have much effect on the response. Both cup and lid clearly have a significant effect and clearly lid has the greatest impact of all three covariates.

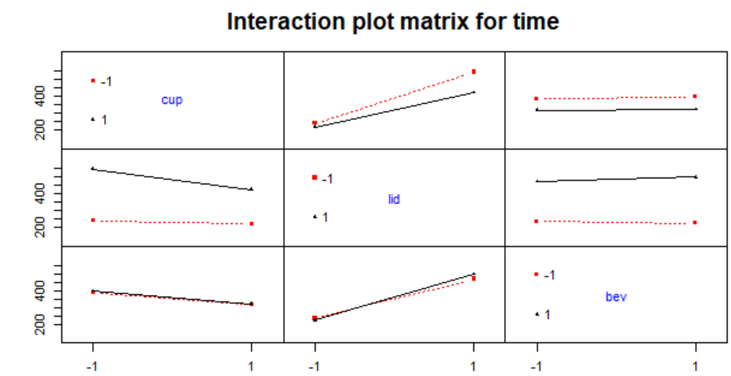


Figure 7: Interaction matrix for time

An interaction matrix is useful for choosing what interactions actually have an effect on the model. If a line is parallel between covariate A and B, then A does not impact the response, regardless if B is high or low. As seen in the plot, the only non-parallel lines are found in the interaction between "lid" and "cup". Thus, this interaction should be regarded as significant.

These results align with our intuition about hot beverages; a lid will of course keep the temperature longer, as will a material with low heat capacity, like paper. The significance of the second-order interaction can be explained by the lid blocking outflow of heat, and instead directing it towards the wall of the cup, which conducts heat based on the material. To be clear, porcelain has higher heat capacity than paper, so the porcelain cup will demand more energy from the beverage to heat up. It is also reasonable that the type of beverage should have no significant effect on the cooldown time, since both coffee and tea consist mostly of water and thus have similar heat capacities.

As we have seen through our analysis, despite only 16 observations, *lid* and *cup* are by far the most influential predictors. Further, one should note that the fact that our residuals do not behave as expected weakens the reliability of our experiment. We cannot say that our errors are independent and identically normally distributed and therefore one should be cautious when assessing the conclusion of this analysis. However, we see the evidence that *lid* and *cup* effect the cooldown time to still be significant, and therefore conclude that there is a correlation in the time it took for the beverage to cool down and these two covariates. If the aim of the experiment was prediction, the natural action would be to create a reduced model by *Best subset model selection*-analysis. This was done and the outcome was to keep only the predictors *lid* and *cup* and their interaction. By doing this the number of observations at each level doubled to four and more irrelevant variance diminished. Best subset selection then caused the model to better reflect the assumptions, e.g. normally distributed residuals seen in the QQ-plot and histogram of this model. We assumed the goal of this experiment was inference, so we chose to not include this analysis.

8 Appendix

```
```{r, eval=FALSE, echo=TRUE}
#Read file and define regression parameters
df <- read.csv("C:\\Users\\augus\\OneDrive - NTNU\\Semester 
6-LAPTOP-2AJBTS3J\\LinStat\\Compulsory 3\\data_sorted.csv", sep=";")

sample(1:16, 16)
#df <- read.csv("data_sorted.csv", sep = ";")
Y <- df[,  4]
n <- length(Y)
X <- as.matrix.data.frame(x = df[, -4])
X <- cbind(rep(1, n),
           X,
           X[,1]*X[,2], X[,1]*X[,3], X[,2]*X[,3],
           X[,1]*X[,2]*X[,3])

XTX = t(X) %*% X
XTXinv = solve(XTX)
beta <- 1/n * t(X) %*% Y
effects <- 2*beta[-1]
ybar = mean(Y)

#Fit a model
model <- lm(time ~ .^3, data = df)
```

(a)

```
#Review model
summary(model)
plot(model)
MEPlot(model)          #Main effects plots
IAPlot(model)          #Interaction plots
stdres<-rstandard(model)
par(mar=c(1, 1, 1, 1))
h <- hist(stdres, xlim=c(-2, 2), breaks=8, freq=FALSE)
xfit <- seq(-2, 2, length = 40)
yfit <- dnorm(xfit, mean = 0, sd = 1)
lines(xfit, yfit)

plot(model$residuals~model$fitted.values)
lines(lowess(model$fitted.values, model$residuals), col="blue")

#Find SSE
onesvec <- matrix(rep(1, n))
onesmat <- onesvec %*% t(onesvec)
I <- diag(rep(1, n))
H <- I - 1/n * X %*% t(X)
yvec = matrix(Y)
sse <- t(yvec) %*% H %*% yvec
```

(b)

Figure 8: All R code used, excluding output. All relevant output can be found in this report.

cup	lid	bev	time
-1	-1	-1	209.64
-1	-1	-1	233.62
1	-1	-1	230.95
1	-1	-1	256.37
-1	1	-1	531.09
-1	1	-1	559.96
1	1	-1	370.23
1	1	-1	412.84
-1	-1	1	239.73
-1	-1	1	266.87
1	-1	1	179.25
1	-1	1	206.09
-1	1	1	525.04
-1	1	1	557.11
1	1	1	432.72
1	1	1	467.51

(a)

Figure 9: CSV file used as data in the report
```

References

[1] National Coffee Association of U.S.A.(2016)
https://www.ncausa.org/Research-Trends/Economic-Impact
The economic impact of the coffee industry, 45 Broadway, Suite 1140, New York, NY 10006.

[2] BEHA-AMPROBE thermometer TPP1-C1

Så i en kilde på nett at dersom vi har PDE'en

$$u_t = -au_x$$

så har vi at $\partial_t^{(k)} = (-a)^k \partial_x^{(k)}$ når vi bruker denne operatoren på u. Dette er også brukt i utledningen til Lax-Wendroff method.

Jeg har aldri sett dette før, og lurte på om det finnes noe enkel forklaring på hvordan man kan si dette?