A2-exercise1

geoffrey

February 29, 2020

Assignment 2

Exercise 1

a. Randomized design

A randomized design with two categorical factors, with

1. the first factor having three categorical levels and
2. the second factor having two levels and
3. having three samples for each unique categorie

can be produced with the following R code:

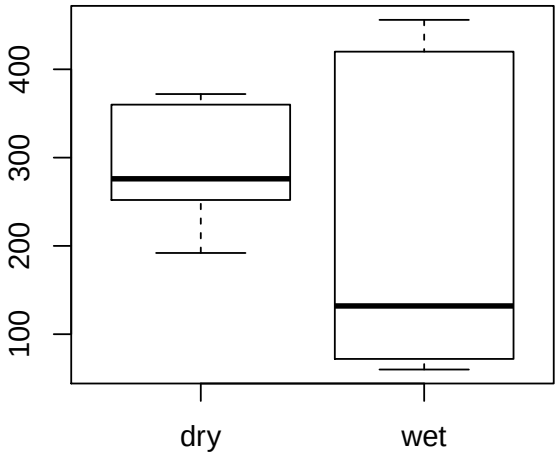
```
I=3; J=2; N=3
rbind(rep(1:I,each=N*J),rep(1:J,N*I),sample(1:(N*I*J)))
```

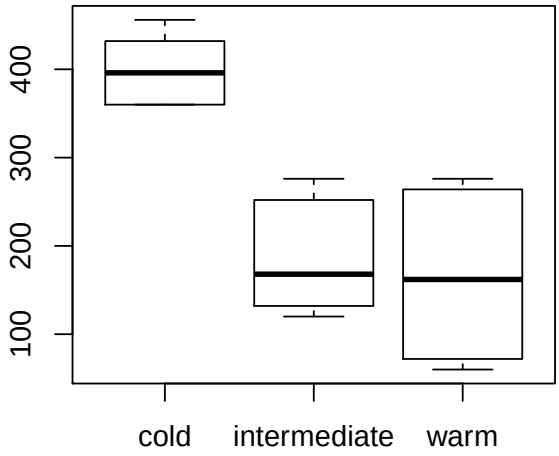
b. Plotting

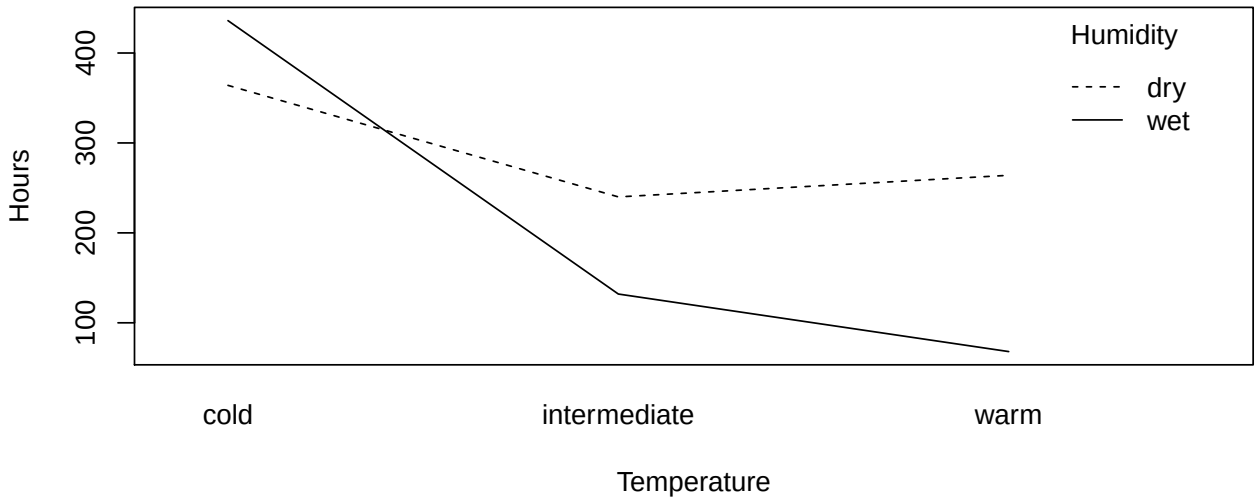
The boxplot and interaction plot below confirms our intuition:

1. A cold environment causes a much slower decay
2. Wet bread has a much wider distribution (variance)
3. On average dry bread decays slower than wet bread
4. However, wet and cold (frozen) bread has the slowest decay

From the non-parallel lines in the interaction plot and the wide distribution of the wet sample we conclude that the (wet) humidity amplifies the effect of the temperature and thus it can be explained by the strong interaction between the two factors (opposed to the errors in the measurement).







c. Two way ANOVA

From the two-way anova resulst below, we can conclude that both factors have a main effect on the decay time of bread. It can also be concluded that the factors have an interaction effect.

```
## Analysis of Variance Table
##
## Response: hours
##                       Df Sum Sq Mean Sq F value    Pr(>F)
## humidity               1  26912   26912  62.296 4.316e-06 ***
## environment            2 201904  100952 233.685 2.461e-10 ***
## humidity:environment   2  55984   27992  64.796 3.705e-07 ***
## Residuals             12   5184     432
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
```

d. Coefficients

TODO

```
##                                    Estimate Std. Error    t value
## (Intercept)                             364   12.00000  30.333333
## humiditywet                              72   16.97056   4.242641
## environmentintermediate                -124   16.97056  -7.306770
## environmentwarm                        -100   16.97056  -5.892557
## humiditywet:environmentintermediate    -180   24.00000  -7.500000
## humiditywet:environmentwarm            -268   24.00000 -11.166667
##                                       Pr(>|t|)
## (Intercept)                       1.032769e-12
## humiditywet                       1.142103e-03
## environmentintermediate           9.389760e-06
## environmentwarm                   7.336887e-05
## humiditywet:environmentintermediate 7.233671e-06
## humiditywet:environmentwarm       1.073751e-07
```

e. Diagnostics

The first requirements is that for each unique categorie, there should be atleast 2 samples, which is the case. Then the most important requirement is that the data among the factors should approximiatly have equal variances. This has been tested in b. and the conclusions was that it approximitally were the same. Another test we can do after the ANOVA, is check whether the error is normally distributed, which is to be expected of a random variable. In the following QQplot it can be seen that the residuals are approximiatly normally distributed. And it the fitted residuals plot it can be seen that the spread is approxomitelly the same among the factors, however there are 2 outliers in the middle.