

### One-way ANOVA model on pain medication treatment

In a study, 3 different treatments (A, B, C) for pain relief are investigated. The investigator randomly assigns 12 participants to the treatments. The participants are instructed to take the given medication when they feel pain and record the time (in minutes) until the pain eases off. The following table shows time to pain relief by treatments.

Treatment	Time to relief (mins)
A	14
	24
	12
	25
B	20
	14
	17
	18
C	22
	29
	36
	20

**H0:** The mean of time to pain relief is not significantly different. (mean of A=mean of B= mean of C).

**H1:** The means are significantly different, or at least one of the three means is significantly different from other two.

**Creating data frame:**

```
trmt.type = c(rep("A",4), rep("B", 4), rep("C",4))  
relf.time = c(14,24,12,25,20,14,17,18,22,29,36,20)  
pain.data <- data.frame(trmt.type,relf.time)
```

Alla Topp  
Project, One-way ANOVA

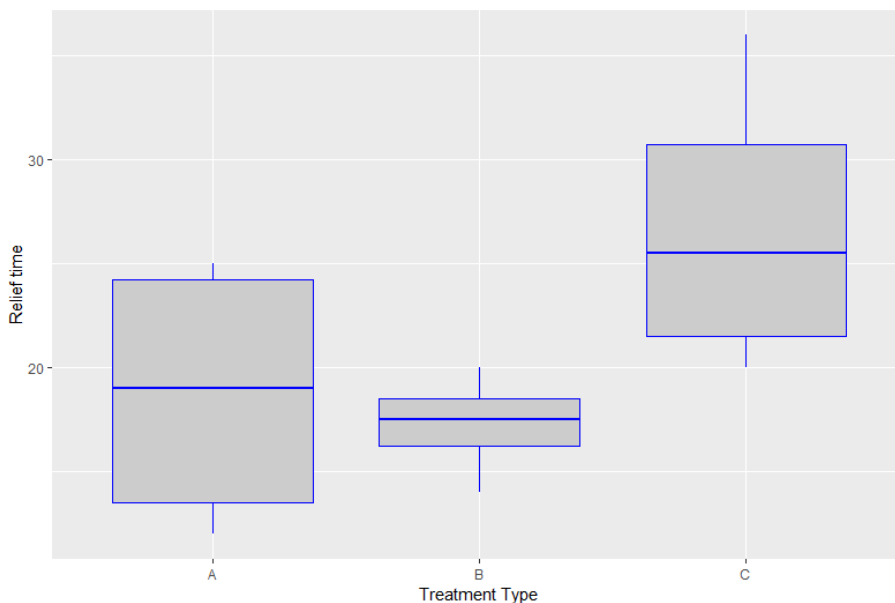
	trmt.type	relf.time
1	A	14
2	A	24
3	A	12
4	A	25
5	B	20
6	B	14
7	B	17
8	B	18
9	C	22
10	C	29
11	C	36
12	C	20

A boxplot of the distributions of the relief time for the three treatment groups is created using the ggplot package:

At this step I had to install ggplots first and then run below command to produce the plot:

```
# creating plot
ggplot(pain.data, aes(x = trmt.type, y = relf.time)) +
  geom_boxplot(fill = "grey80", color = "blue") +
  scale_x_discrete() + xlab("Treatment Type") +
  ylab("Relief time")
```

For x-axis we have treatment types and for y-axis we use relief times. Fill means the fill color of the box and color = blue means the color of the outer line of the boxes.



Looking at the plot we can see that the blue line is pretty much in the middle for all three treatment types, so we can already see that treatments are not different to the control group, but it is not that clear yet. So we proceed for further analysis. To investigate these differences, we fit the one-way ANOVA model using the **lm** function and look at the parameter estimates and standard errors for the treatment effects.

```
5 # creating 3 vectors
6 aTreat=c(14,24,12,25) # vector A treatment
7 bTreat=c(20,14,17,18) # vector B treatment
8 cTreat=c(22,29,36,20) # vector with values from C treatment
9 treatments = c(aTreat,bTreat,cTreat) # combining 3 vectors in one vector
10
11 # creating 3 levels where all the data from the vector stored
12 groups = factor(rep(letters[1:3], each = 4))
13 fit =lm(formula = treatments ~ groups) #organizing the values fitting the model
14 summary(fit) #
15 anova(fit) # produces the anova model
16 |
```

```
> summary(fit) #
```

Call:

```
lm(formula = treatments ~ groups)
```

Residuals:

Min	1Q	Median	3Q	Max
-6.750	-4.750	0.250	3.375	9.250

Coefficients:

	Estimate	Std. Error	t value	Pr(> t )
(Intercept)	18.750	2.945	6.367	0.00013 ***
groupsb	-1.500	4.165	-0.360	0.72705
groupsc	8.000	4.165	1.921	0.08695 .

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

Residual standard error: 5.89 on 9 degrees of freedom

Multiple R-squared: 0.4006, Adjusted R-squared: 0.2674

F-statistic: 3.007 on 2 and 9 DF, p-value: 0.09995

```
> anova(fit) # produces the anova model
```

Analysis of Variance Table

Response: treatments

	Df	Sum Sq	Mean Sq	F value	Pr(>F)
groups	2	208.67	104.333	3.0072	0.09995 .
Residuals	9	312.25	34.694		

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

Since p-value > 0.05, we accept the null hypothesis H0: the 3 means are statistically equal.

Since we accepted null hypothesis we don't need to run post-hoc tests, such as Tukey's Honest Significant Difference test to determine which treatment has significantly different means. On the other hand, we could verify the homoskedasticity (i.e. test for homogeneity of variances). We can do it with the Bartlett test and the Fligner-Killeen test.

```
> bartlett.test(treatments,groups)

    Bartlett test of homogeneity of variances

data:  treatments and groups
Bartlett's K-squared = 2.7008, df = 2, p-value = 0.2591

> fligner.test(treatments,groups)

    Fligner-Killeen test of homogeneity of variances

data:  treatments and groups
Fligner-Killeen:med chi-squared = 5.4992, df = 2, p-value = 0.06395
```

Looking of the last two tests p-value we can see that  $p\text{-value} > 0.05$  and we accept null hypothesis.