Linguistic Data: Quantitative Analysis and Visualisation

Ilya Schurov, Olga Lyashevskaya, George Moroz, Alla Tambovtseva

Part 1. ANOVA: Analysis of Variance

Let's load data on Icelandic we worked before:

```
phono <- read.csv("http://math-info.hse.ru/f/2018-19/ling-data/icelandic.csv")
```

When we discussed this data frame the first time, we compared the vowel duration for cases when vowels are followed by aspirated and non-aspirated consonants. So, we had two groups to compare. Now we will try to compare more groups.

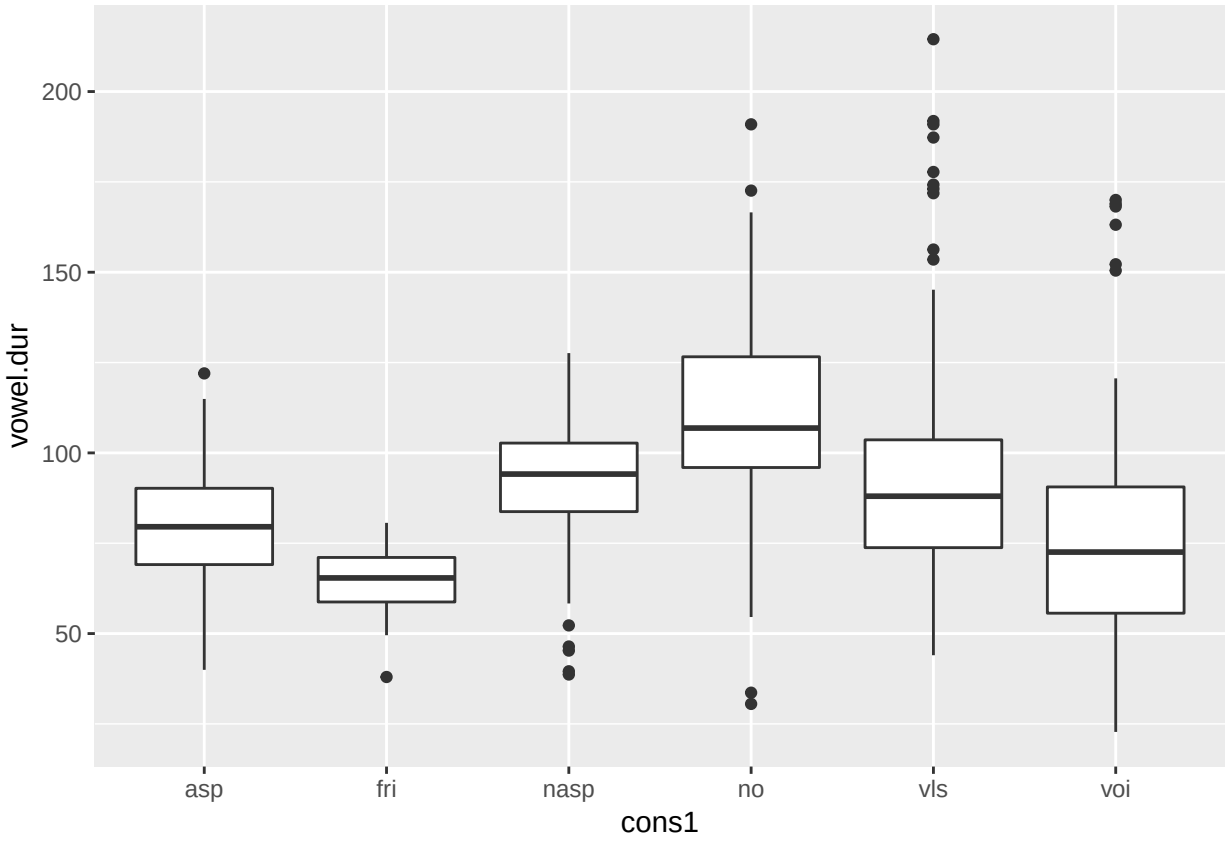
Look at all groups of consonants:

```
table(phono$cons1)
```

```
##
##  asp  fri nasp   no  vls  voi
##  304   15  133   94  142  118
```

Create a boxplot for vowel duration for each group of consonants (and revise `ggplot()` as well):

```
library(tidyverse)
ggplot(data = phono, aes(x = cons1, y = vowel.dur)) + geom_boxplot()
```



As we can see, the median values (and the shape of the distributions as well) are different for different groups of consonants.

Now let's perform ANOVA (Analysis of Variance). Formulate hypotheses:

$$H_0 : \mu_1 = \mu_2 = \cdots = \mu_k \text{ (there are no difference in population means)}$$

$$H_1 : \text{there exists at least one pair of groups with different population means}$$

Run ANOVA (the syntax is `variable of interest ~ grouping variable`):

```
res <- aov(phono$vowel.dur ~ phono$cons1)
res
```

```
## Call:
##    aov(formula = phono$vowel.dur ~ phono$cons1)
##
## Terms:
##                 phono$cons1 Residuals
## Sum of Squares      96776.3  404073.9
## Deg. of Freedom           5       800
##
## Residual standard error: 22.47426
## Estimated effects may be unbalanced
```

More informative summary:

```
summary(res)
```

```
##              Df Sum Sq Mean Sq F value Pr(>F)
## phono$cons1   5  96776   19355   38.32 <2e-16 ***
## Residuals   800 404074     505
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
```

Question: judging by the output above, can we conclude that average vowel duration differ significantly in different groups of consonants?

Answer: yes, we can, judging by p-value that is close to zero. Null hypothesis should be rejected.

Part 2: Multiple comparisons and Bonferroni correction

During the lecture we discussed a problem of multiple comparisons and concluded that it is not correct to compare groups pairwise (as is) if we have more than two groups to compare. We should either use ANOVA (or its non-parametric analogues) or perform multiple comparisons with corrections. Let's compare the vowel duration for each pair of vowel types using a t-test with no corrections:

```
# g - grouping variable
# p.adjust.method - adjustment
pairwise.t.test(phono$vowel.dur,
                g = phono$cons1,
                p.adjust.method = "none")
```

```
##
##  Pairwise comparisons using t tests with pooled SD
##
## data:  phono$vowel.dur and phono$cons1
##
```

```
##      asp     fri     nasp    no      vls
## fri  0.0085  -       -       -       -
## nasp 3.2e-07 6.8e-06 -       -       -
## no   < 2e-16 9.2e-13 8.2e-09 -       -
## vls  3.2e-10 8.9e-07 0.3556  5.1e-07 -
## voi  0.0987  0.0589  2.2e-08 < 2e-16 6.0e-11
##
## P value adjustment method: none
```

Here we have a table that contains p-values for two-sample t-tests applied for every pair of groups (vowel types). For example, p-value for the t-test comparing mean values of aspirated (`asp`) and fricative (‘fri) vowels is 0.0085.

Now let’s do the same comparisons, but with the Bonferroni correction.

```
pairwise.t.test(phono$vowel.dur, g = phono$cons1, p.adjust.method = "bonferroni")
```

```
##
##  Pairwise comparisons using t tests with pooled SD
##
## data:  phono$vowel.dur and phono$cons1
##
##      asp     fri     nasp    no      vls
## fri  0.1274  -       -       -       -
## nasp 4.9e-06 0.0001  -       -       -
## no   < 2e-16 1.4e-11 1.2e-07 -       -
## vls  4.9e-09 1.3e-05 1.0000  7.6e-06 -
## voi  1.0000  0.8838  3.3e-07 < 2e-16 9.0e-10
##
## P value adjustment method: bonferroni
```

What happened when we applied this correction? P-values became larger! What does it show? We discussed that the Bonferroni correction helps us to reduce the Type I error while perfoming multiple comparisons. Type I error is the probability of rejecting the null hypothesis when it is actually true. Here p-values are larger, so we have less chances to reject H_0 for each pair of vowel types.