Midterm tests MTH6101 2024

2024-04-18

Overall results for the tests

The first test was presented by 302 students, with an average mark of 73, while the second test was presented by 291 students, with an average mark of 74.75.

Here is the distribution of students in each classification:

```
##        0-40 40-50 50-60 60-70 70-100
## First    27    20    24    46    185
## Second   29    18    31    42    171
```

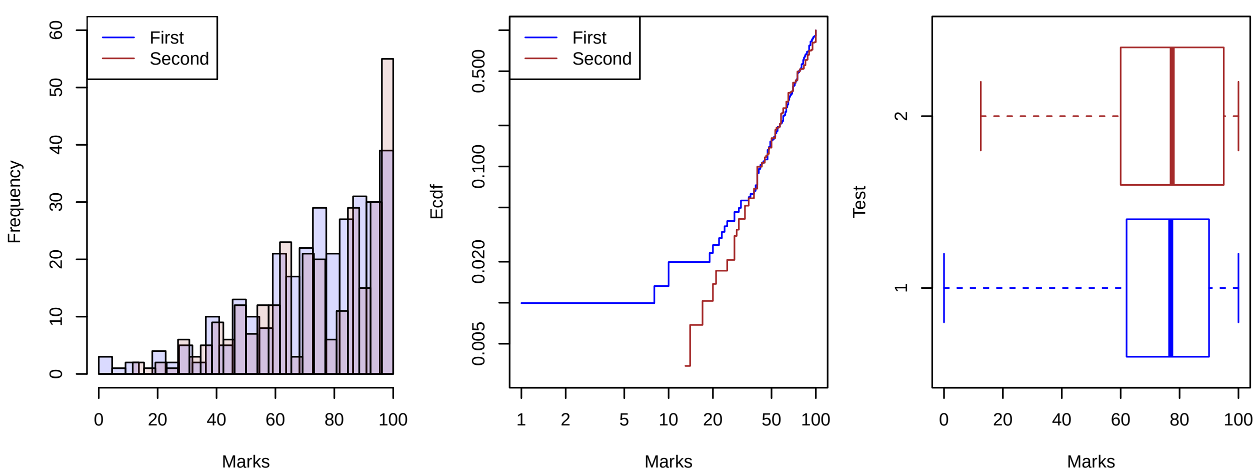
and the same distribution, written as percentages (divided by the total of students in each test):

```
##        0-40 40-50 50-60 60-70 70-100
## First  8.94  6.62  7.95 15.23  61.26
## Second 9.97  6.19 10.65 14.43  58.76
```

Were the tests results different?

If we consider the two tests as different data, we can compare them using a (KS) nonparametric test with a p-value of 0.008, which would be interpreted as a rejection indicating that the test results were different. Note that these results include all students who did a perfect exam. If we exclude these from the analysis, we have a p-value of 0.317, not rejecting equality between tests.

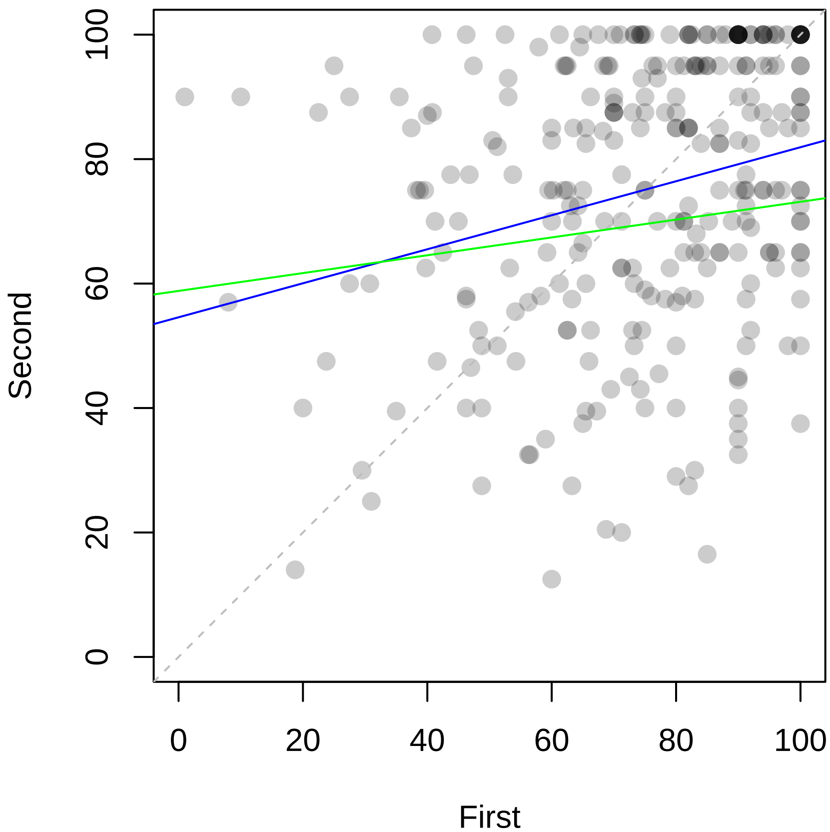
Here is a histogram of the marks obtained with the first and second tests, as well as a plot in log-log scale of the empirical cumulative densities for each test. In this latter plot a highlight is that the distributions differ in the left tail although there are only a handful of observations there. The plot also shows the big number of perfect exams in the second test, although this right tail is not as evident in the plot as the left tail. We end up with a boxplot of each test, highlighting the differences in the tails as stated.



The test results as paired data

Now we consider the data as paired data, where for each student we have two observations. The results will differ a little from the previous analysis, as we can only consider those students sitting both tests, i.e. we consider 285 students for which `complete.cases` gave value `TRUE`.

We give below a scatterplot of the test results. The shading indicates frequency of cases; the dashed grey line indicates equality between test marks; the blue line predicts the second test result using the mark of the first test, and the solid green line is the same prediction, excluding those very high exams.



An important feature in the diagram is the higher number of perfect results in the second test, shown in the top right corner. There is a weak positive correlation of $r = 0.26$ between the two tests. Concerning the regression perfomed with the tests, although both model terms are highly significant, the model explains little of the variability in the data with a very low $R^2 = 0.07$.

Paired data as a contingency table

If we consider the classification of each students' result in the test, we have the following contingency table:

```
##          Second
## First     0-40 40-50 50-60 60-70 70-100
##    0-40      5     1     3     1     11
##    40-50     3     3     3     3      6
##    50-60     4     2     3     3     10
##    60-70     5     2     6     4     28
##    70-100   12    10    13    30    114
```

We perform the independence test that the joint distribution equals to the product of (row and column) marginals. We have a p-value of 0.15 and thus we cannot reject the independence hypothesis.