

## 5.5: Answers to exercises

### Two sample tests, effect sizes

Answer to the sign test question. It is enough to write:

Here the sign test failed to find obvious differences because (like t-test and Wilcoxon test) it considers only central values.

Answer to the ozone question. To know if our data are normally distributed, we can apply the `Normality()` function:

(Here we applied `unstack()` function which segregated our data by months.)

Answer to the argon question. First, we need to check assumptions:

It is clear that in this case, nonparametric test will work better:

(We used `jitter()` to break ties. However, be careful and try to check if this random noise does not influence the p-value. Here, it does not.)

And yes, boxplots (Figure 5.2.4) told the truth: there is a statistical difference between two set of numbers.

Answer to the cashiers question. Check normality first:

Now, we can compare means:

It is likely that first cashier has generally bigger lines:

The difference is not significant.

Answer to the grades question. First, check the normality:

(Function `split()` created three new variables in accordance with the grouping factor; it is similar to `unstack()` from previous answer but can accept groups of unequal size.)

Check data (it is also possible to plot boxplots):

It is likely that the first class has results similar between exams but in the first exam, the second group might have better grades.

Since data is not normal, we will use nonparametric methods:

For the first class, we applied the paired test since grades in first and second exams belong to the same people. To see if differences between different classes exist, we used one-sided alternative hypothesis because we needed to understand not if the second class is different, but if it is *better*.

As a result, grades of the first class are not significantly different between exams, but the second class performed significantly better than first. First confidence interval includes zero (as it should be in the case of no difference), and second is not of much use.

Now effect sizes with suitable nonparametric Cliff's Delta:

Therefore, results of the second class are only *slightly better* which could even be negligible since confidence interval includes 0.

Answer to the question about ground elder leaves (Figure 5.5.1).

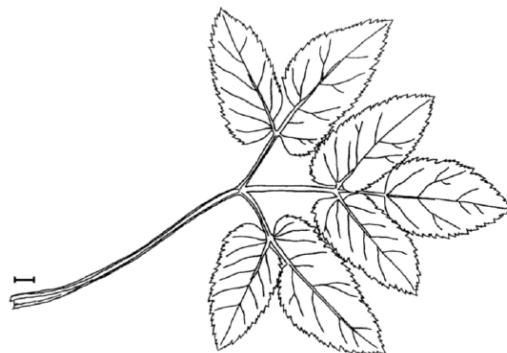


Figure 5.5.1 Leaf of *Aegopodium podagraria*., ground elder. Scale bar is approximately 10 mm.

First, check data, load it and check the object:

(We also converted `SUN` variable into factor and supplied the proper labels.)

Let us check the data for the normality and for the most different character (Figure 5.5.2):

`TERM.L` (length of the terminal leaflet, it is the rightmost one on Figure 5.5.1), is likely most different between sun and shade.

Since this character is normal, we will run more precise parametric test:

To report t-test result, one needs to provide degrees of freedom, statistic and p-value, e.g., like “in a Welch test, t statistic is 14.85 on 63.69 degrees of freedom, p-value is close to zero, thus we rejected the null hypothesis”.

Effect sizes are usually concerted with p-values but provide additional useful information about the magnitude of differences:

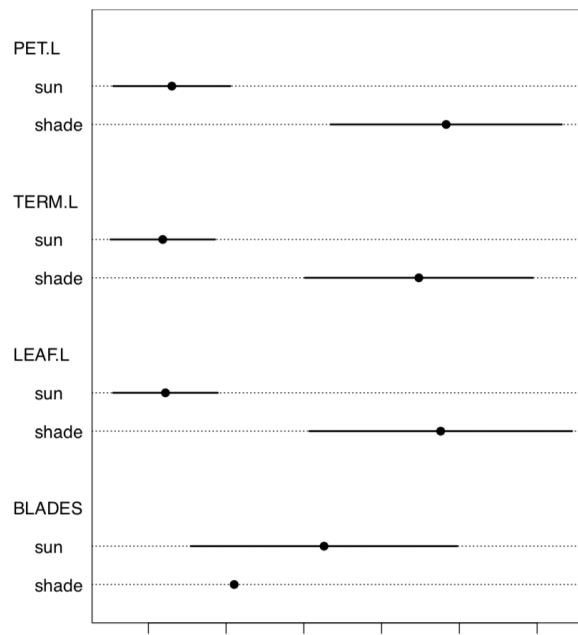


Figure 5.5.2 Medians with MADs in leaves data.

Both Cohen's d and Lyubishchev's K (coefficient of divergence) are large.

### ANOVA

Answer to the height and color questions. Yes on both questions:

There are significant differences between all three groups.

Answer to the question about differences between cow-wheats (Figure 5.5.3) from seven locations.

Load the data and check its structure:

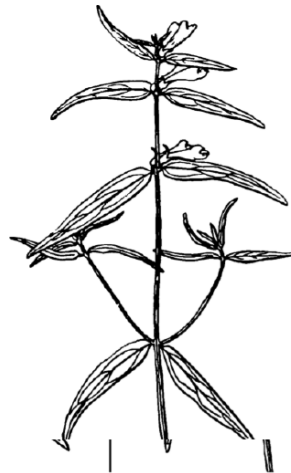


Figure 5.5.3 Top of the cow-wheat plant, *Melampyrum* sp. Size of fragment is approximately 10 cm.

Plot it first (Figure 5.5.4):

Check assumptions:

Consequently, leaf length must be analyzed with non-parametric procedure, and plant height—with parametric which does not assume homogeneity of variance (one-way test):

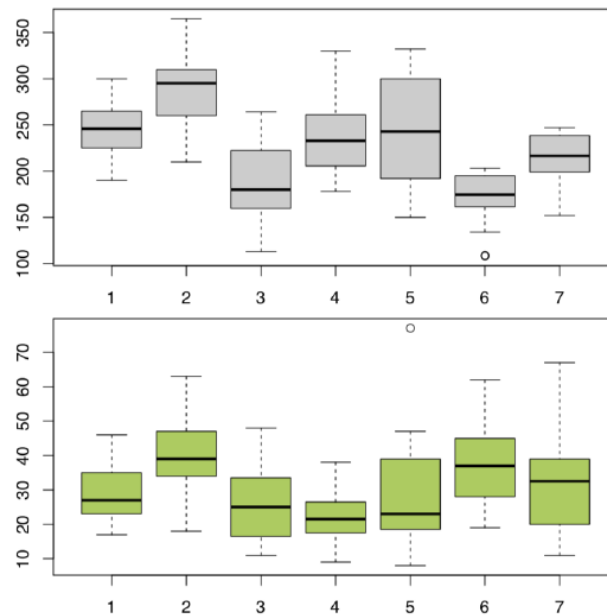


Figure 5.5.4 Cow-wheat stem heights (top) and leaf lengths (bottom) across seven different locations.

Now the leaf length:

All in all, location pairs 2–4 and 4–6 are divergent statistically in both cases. This is visible also on boxplots (Figure 5.5.5). There are more significant differences in plant heights, location #6, in particular, is quite outstanding.

### Contingency tables

Answer to the seedlings question. Load data and check its structure:

Now, what we need is to examine the table because both variables only look like numbers; in fact, they are categorical. Dotchart (Figure 5.5.5) is a good way to explore 2-dimensional table:

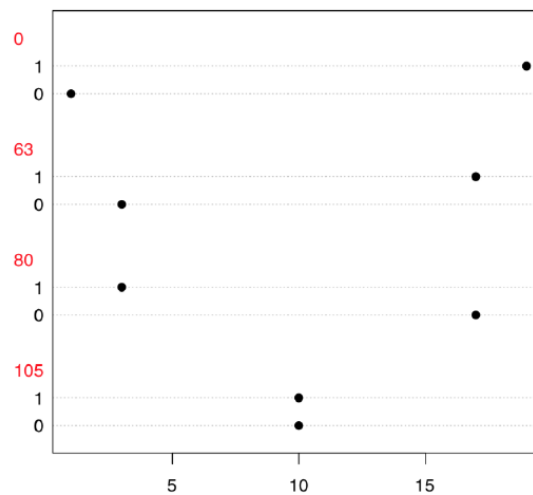


Figure 5.5.5 Dotchart to explore table made from seedlings data.

To explore possible associations visually, we employ `vcd` package:

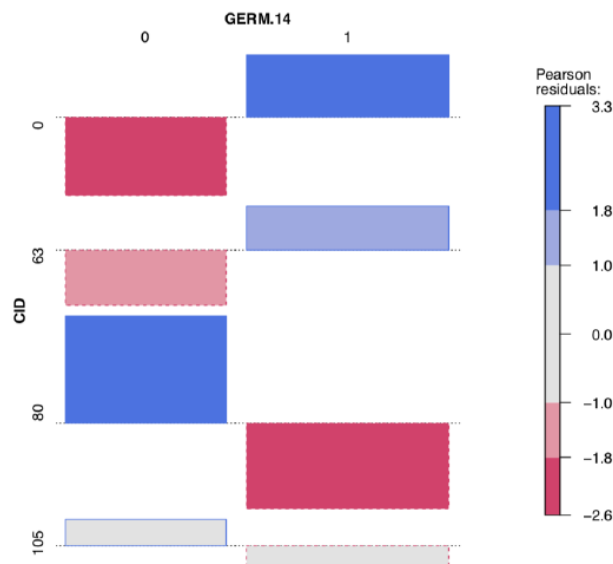


Figure 5.5.6 Advanced association plot of seedlings data.

Both table output and `vcd` association plot (Figure 5.5.6) suggest some asymmetry (especially for CID80) which is a sign of possible association. Let us check it numerically, with the chi-squared test:

Yes, there is an association between fungus (or their absence) and germination. How to know differences between particular samples? Here we need a *post hoc* test:

(Exact Fisher test was used because some counts were really small.)

It is now clear that germination patterns form two fungal infections, CID80 and CID105, are significantly different from germination in the control (CID0). Also, significant association was found in the every comparison between three infections; this means that all three germination patterns are statistically different. Finally, one fungus, CID63 produces germination pattern which is *not* statistically different from the control.

Answer to the question about multiple comparisons of toxicity. Here we will go the slightly different way. Instead of using array, we will extract p-values right from the original data, and will avoid warnings with the exact test:

(We cannot use `pairwise.Table2.test()` from the previous answer since our comparisons have different structure. But we used exact test to avoid warnings related with small numbers of counts.)

Now we can adjust p-values:

Well, now we can say that Caesar salad and tomatoes are statistically supported as culprits. But why table tests always show us two factors? This could be due to the interaction: in simple words, it means that people who took the salad, frequently took tomatoes with it.

Answer to the scurvy-grass question. Check the data file, load and check result:

(In addition, we converted `LOC` and `IS.CREEPING` to factors and provided new level labels.)

Next step is the visual analysis (Figure 5.5.7):

Some locations look different. To analyze, we need contingency table:

Now the test and effect size:

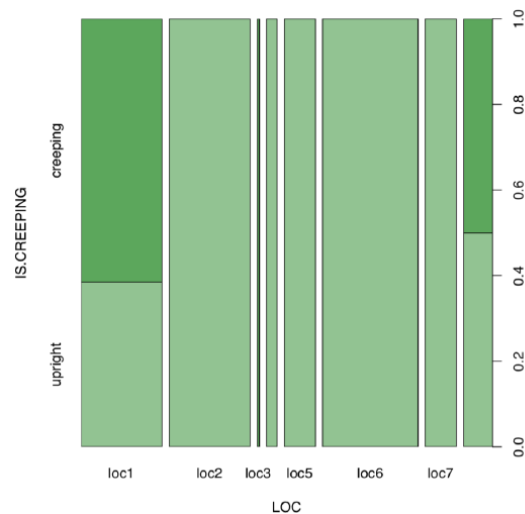


Figure 5.5.7 Spine plot: locality vs. life form of scurvy-grass.

(Run `pairwise.Table2.test(cc.lc)` yourself to understand differences in details.)

Yes, there is a large, statistically significant association between locality and life form of scurvy-grass.

Answer to the question about equality of proportions of **LOBES** character in two birch localities. First, we need to select these two localities (1 and 2) and count proportions there. The shortest way is to use the `table()` function:

Spine plot (Figure 5.5.8) helps to make differences in the table even more apparent:

(Please also note how to create two colors intermediate between black and dark green.)

The most natural choice is `prop.test()` which is applicable directly to the `table()` output:

Instead of proportion test, we can use Fisher exact:

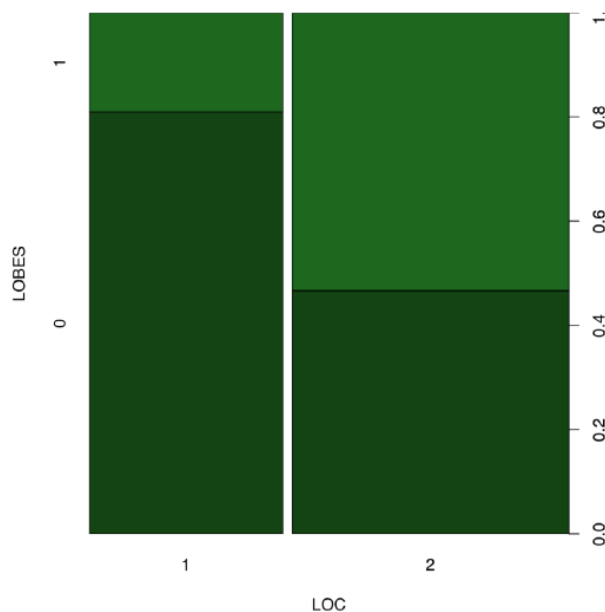


Figure 5.5.8 Spine plot of two birch characters.

... or chi-squared with simulation (note that one cell has only 4 cases), or with default Yates' correction:

All in all, yes, proportions of plants with different position of lobes are different between location 1 and 2.

And what about effect size of this association?

Answer to the question about proportion equality in the whole **betula** dataset. First, make table:

There is no apparent asymmetry. Since **betula.lw** is  $2 \times 2$  table, we can apply fourfold plot. It shows differences not only as different sizes of sectors, but also allows to check 95% confidence interval with marginal rings (Figure 5.5.9):

Also not suggestive... Finally, we need to test the association, if any. Note that samples are *related*. This is because **LOBES** and **WINGS** were measured on the *same plants*. Therefore, instead of the chi-squared or proportion test we should run McNemar's test:

We conclude that proportions of two character states in each of characters are not statistically different.

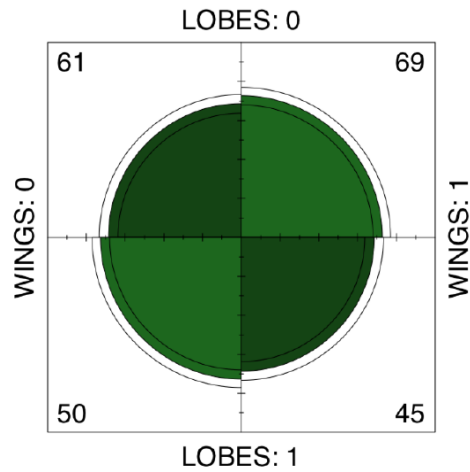


Figure 5.5.9 Fourfold plot of two birch characters.

This page titled 5.5: Answers to exercises is shared under a [Public Domain](#) license and was authored, remixed, and/or curated by Alexey Shipunov via source content that was edited to the style and standards of the LibreTexts platform.