Problem Set 2.1: Comparing multiple means

ADS2

Semester 2 2023/24

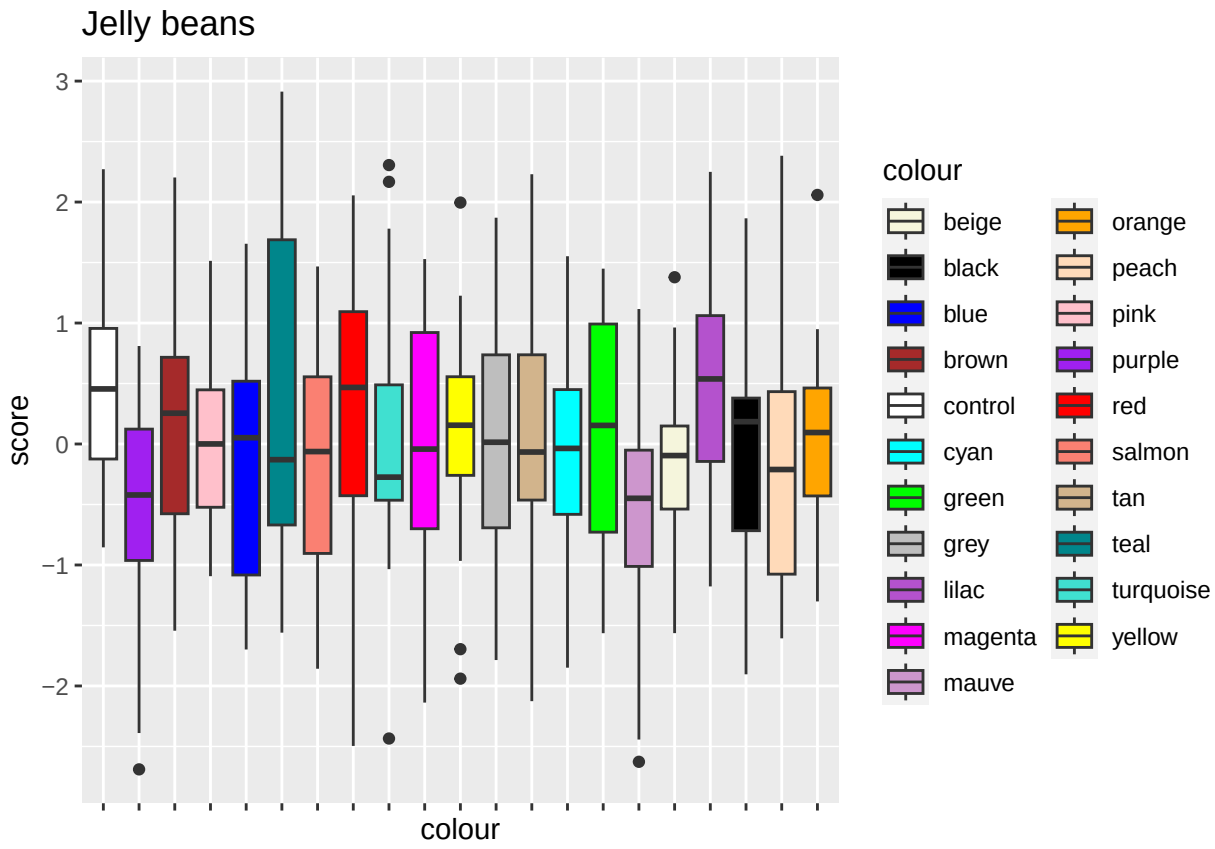
We expect this problem set to take around an hour to complete. But professors are sometimes wrong![citation missing] If this or future problem sets are too long, please let us know, so we can adjust and plan accordingly.

The problem with jelly beans

Look at this web comic: https://xkcd.com/882/

We have created a simulation of the study done in the comic. The data can be found in file *jellybeans.csv*

- For the simulation, we imagined that acne is measured by the presence of an "acne score", which is normally distributed with a mean of 0 and standard deviation of 1.
- From the comic, it looks like (after being asked to investigate individual colours), the scientists ran 20 separate experiments, one for each of 20 colours. Actually, there probably (hopefully!) was an additional "control" experiment, where participants ate no jelly beans at all (there may have been a "control" condition for each of the 20 colours, but it is very difficult to establish the exact experimental protocol for a fictional study from a web comic).
- For the sake of this exercise, we assumed that each group has 20 participants. We randomly created the datasets assuming that the Null Hypothesis is true (i.e. none of the jelly bean colours really causes acne)



Madness of multiple t-tests

- We established in lecture that multiple t-test massively increase the probability of a false positive result. Even if researchers only compared each colour to control (and not different colours to each other), how many t-tests is that? What is the false positive probability?
- Confirm that at least one of the t-tests does indeed give a false positive result. (No need to systematically compare everything - have a look at the plot first and look for a comparison that you suspect would give a significant result)

Can we look at variance instead?

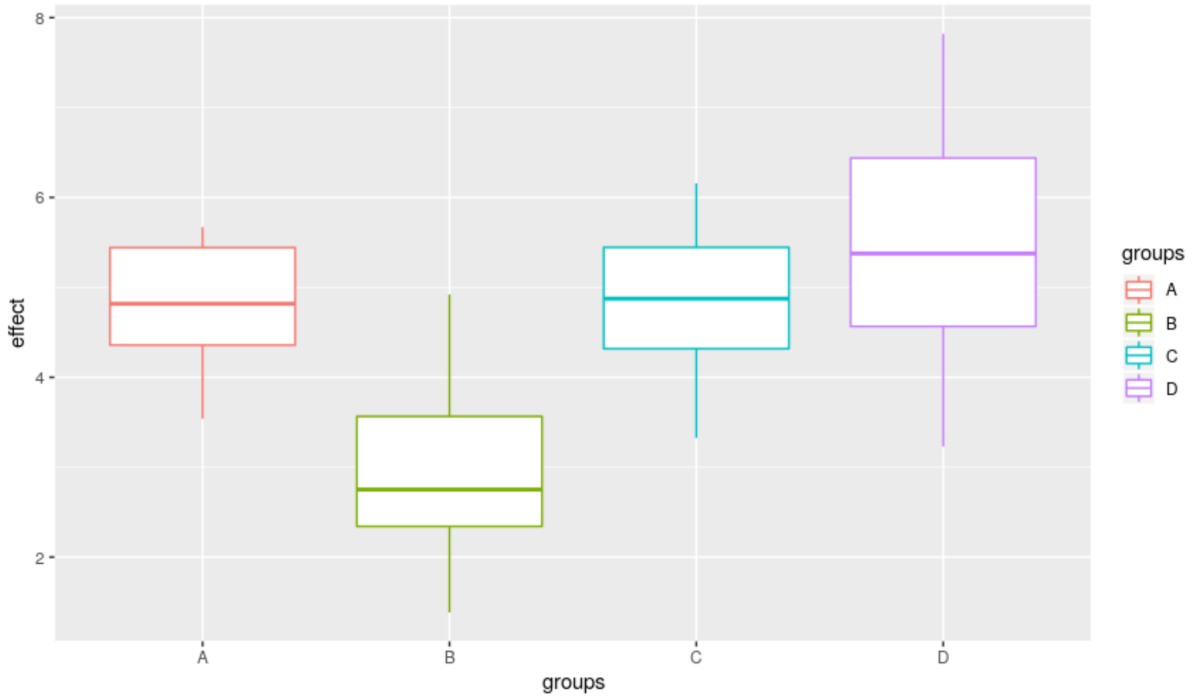
If we only want to know whether at least 2 out of the 21 conditions are different from each other, we could do so by looking at differences within and between groups again. Remember, we said in lecture that we want to find out whether the difference between two data points can be explained by them being in different groups.

You can use a similar method to what you did in the practical to repeatedly draw samples of two and take note of the (absolute) difference between them and of whether they come from the same or different groups. Note that you would probably need more iterations than you had for the practical (*why?*).

There is a difference, what now?

Let's assume we looked at within group and between group differences in some study and concluded that the four groups are *NOT* the same. But what now? How do we know what groups exactly are different from each other?

Let's go back to the simpler example we looked at in lecture: comparing just four groups.



Remember we said that running 6 individual t-tests would give us an effective false positive probability (if H0 is true) of more than 25 %.

But could we adjust the α level for each individual t-test so that the overall false positive probability remains at 0.05? If so, how?
