V.1 – STATISTICAL ASPECTS

EEG-TRAINING

OUTLINE

1. Goal
2. t-test
3. Problem
4. Correction for multiple comparisons
    1. Bonferroni
    2. Non-parametric permutation testing

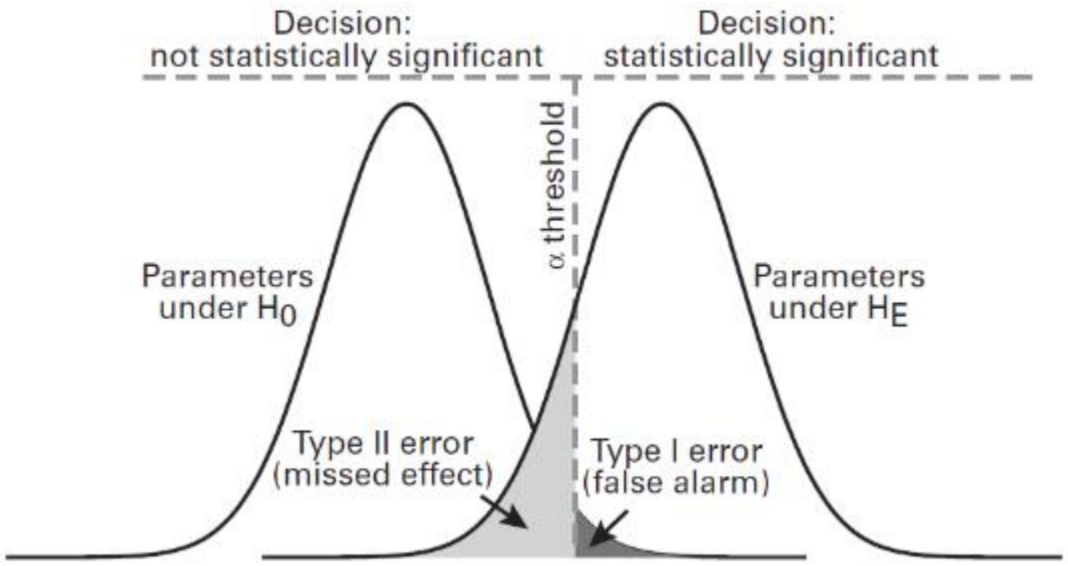
1) GOAL

- So far we have studied phenomena in time-space or time-frequence-space
- We offered qualitative observations
- In particular, we are interested in quantifying differences between two phenomena (or quantify how different a phenomenon is compared to baseline activity)

Statistical analyses

2) t –TEST

- Distributions A and B describe a phenomenon (ERP or ERSP) in two different conditions
- Question: are the two distributions identical? (this is our H_0 hypothesis, that we try to reject through t –test)
- Set a threshold for significance: α (usually set to 0.05 for one test)
- Type I error: find an effect where there is in fact none. Associate a probability p to it
- The null-hypothesis is rejected if $p < \alpha$

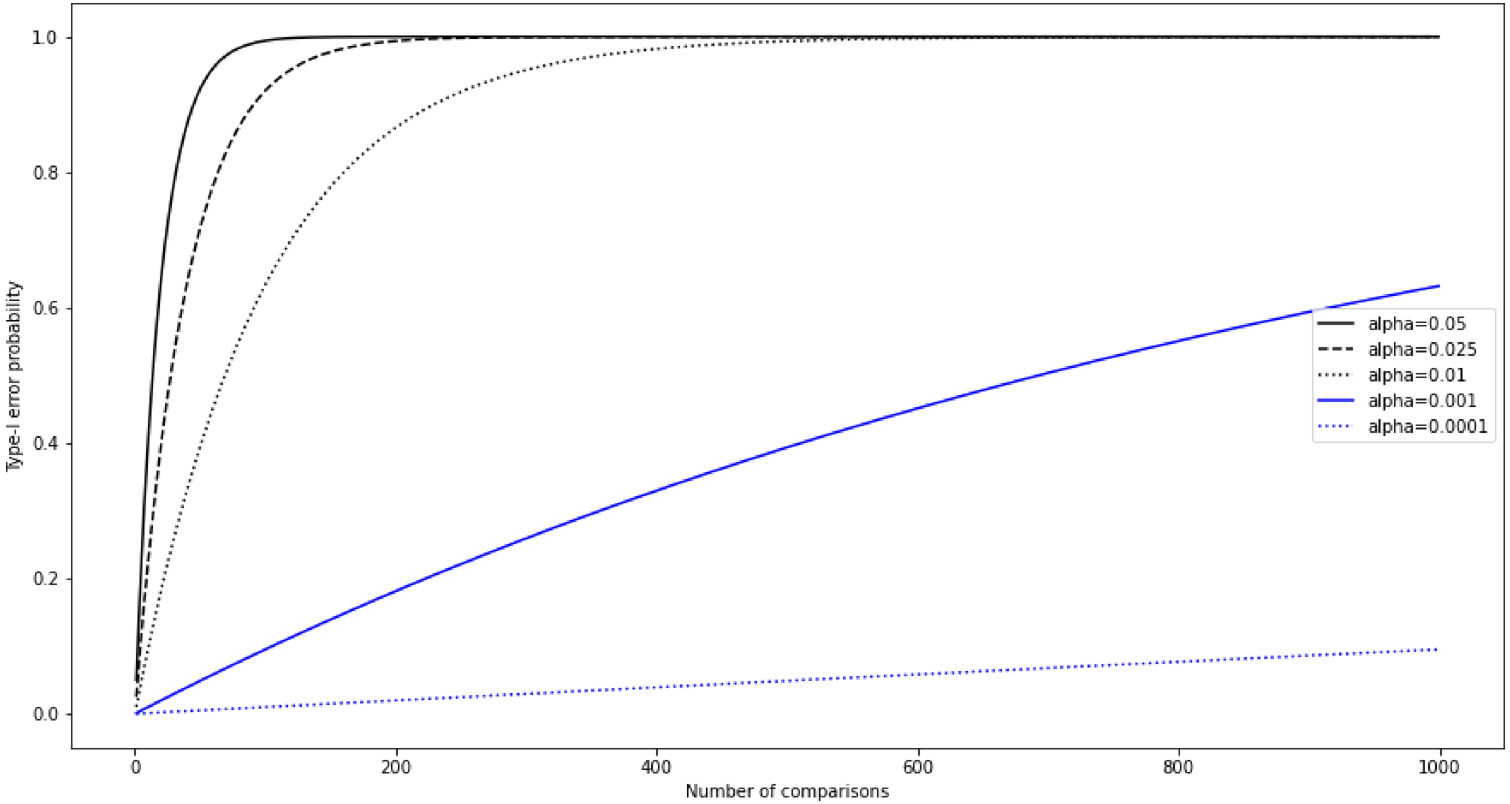


3) PROBLEM

- This approach is univariate, ie it takes into account one variable and compares it in two conditions
- In our case, it means we compare the power in one time-frequency interval at one electrode
- In reality, we have several electrodes and a lot of time-frequency pixels to test: need for several t −tests
- **Main assumption: data needs to be normally distributed**

3) PROBLEM

- Multiple tests increase the risk for Type-I errors: $\bar{\alpha} = 1 - (1 - \alpha)^m$



4) BONFERRONI CORREECTION FOR MULTIPLE COMPARISONS

- The risk with performing multiple t-tests is that we increase the chance of finding a significant effect just by chance

  $\Rightarrow$ Need to adapt the threshold to account for this risk

- Bonferroni correction: $\alpha_{multiple} = \frac{\alpha_{simple}}{n_{comparisons}}$
- But what if $n_{comparisons} = 100\ 000$ (often the case in our data)?
- What if data not normally distributed?

4) NON-PARAMETRIC PERMUTATION TESTING

- This method solves the issue of the normalization and of the multiple comparisons
- We still consider each voxel to be an independent variable
- How it's done:
  - In theory: generate a random distribution of each condition, extract a minimum and a maximum distribution, compare the observed quantities to these distributions
  - In practice: over a few iterations: shuffle the observations between the two conditions (ie: exchange the label of observations), extract the minimum and maximum of all quantities (over all voxels), and repeat. If the observed quantity is bigger than the $(1-\frac{\alpha}{2})$ quantile of the maximum distribution or smaller than the $\frac{\alpha}{2}$ quantile of the minimum distribution, then the observation is statistically significant

CONCLUSION

- Statistical testing of EEG data is delicate and it requires a good understanding of the assumption of tests
- In many cases parametric methods are not adapted because of the non-normality of the data handled and/or the number of variables taken into account
- The approaches described here all assume that the voxels are independent, which is in fact not the case. However, it is a valid assumption as long as we are aware of that and do not try to correlate findings between them
- In the next and final chapter, we describe cluster-based methods