SESSION 4
SUMMARY

- We will recap the basics of model fitting from last week
- We will recap how to use simulations from the model to validate the model (from session 2)
- We will learn how to do statistical tests (t-tests and correlations) in Matlab
- We will check whether the stress manipulation worked
- And see whether the stress manipulation affected how participants learned
- We will talk about what to include in your practical report

Part 1: TESTING YOUR RESULTS FOR STATISTICAL SIGNFICANCE
We will use the 'ttest' (paired-samples t-test), 'ttest2' (independent samples t-test) and 'corr' (Pearson or Kendall correlations), command in Matlab for the following tests:

- Are learning rates different in the stable and volatile phase (across all participants, using paired samples t-test)?
- Did the stress manipulation affect the mood ratings (independent samples t-test)?
- Is the difference in learning rates between blocks affected by the stress manipulation (independent samples t-test)? And relatedly, is this effect quantitative, i.e. do people who feel more stressed - either in response to the sounds or in general – have more problems adjusting their learning rates (correlation)?

Please note that the figures in this handout were made using fake data that I generated to illustrate how the analyses would look if certain effects did or did not exist- your real data will look different!

To get started, download the folder 'Session4' from the weblearn page. Then open the Matlab script 'Session4_StudentVersion'. In the first cell 'Load data', complete the line 'excludeParticipants' with the IDs of the participants that you decided to exclude in the last session (if any). If you have not decided yet who you want to exclude for your report, that's fine, just leave the array empty as it is. Then execute the cell to load the model fits (made using the scripts from session three) for the real data, as well as for simulated data. The variables that are defined in this cell are:

- 'InclSubs': a list of 1s and 0s that is 1 if a participant was in either group and was not excluded; the length of this variable is the number of subjects
- 'Controls': a list of 1s and 0s that is 1 if a participant was in the control group and was not excluded
- 'Stress': a list of 1s and 0s that is 1 if a participants was in the manipulation (i.e. stressful sounds) group and was not excluded
- 'Alphas': the learning rates for the stable (first column) and volatile block (second column); the size of this variable is therefore number of subjects x 2

- 'AICs': model fits for each person measured using AIC (see session 3) for the model with one learning rate (1st column) and the model with two learning rates (2nd column)
- 'Ratings': the mood ratings (either stress or happy – up to you to decide which one to use for your report) of each participant over the course of the task; the size of this array is therefore number of subjects x 8 (for the eight ratings they did over the course of the experiment)

[A.0: Interlude: Recap from last week: AIC scores]
In last week's session we have compared our models using AIC scores (Akaike information criterion). Here we are plotting these scores again, separately for the two groups. We want to check that the model from which we later extract parameters to test the two groups indeed provides the best (or at least a relatively good) fit to the data. This becomes more relevant when we construct more models. For example, for this study, we could have constructed a model that also has two learning rates, but makes decisions based on a weighted sum of reward magnitude and probability (see session 2), rather than based on their product (magnitude x probability).
Q: Why would we need to test whether the multiplicative or the additive model provides a better fit before extracting the learning rates from one model or the other for further statistical tests?

____

____

To report AIC values, there are two methods: sum up the difference in AIC scores. This tells you which model (across all subjects in general or across all subjects in one group) provides the best fit. AIC difference over 10 can be taken as very strong evidence in favour of the model with a smaller AIC score. The second way is to count how many participants are better fit by either model (i.e. have a negative or positive AIC value).
Q: Can you think of the advantages and disadvantages of the two approaches?

____

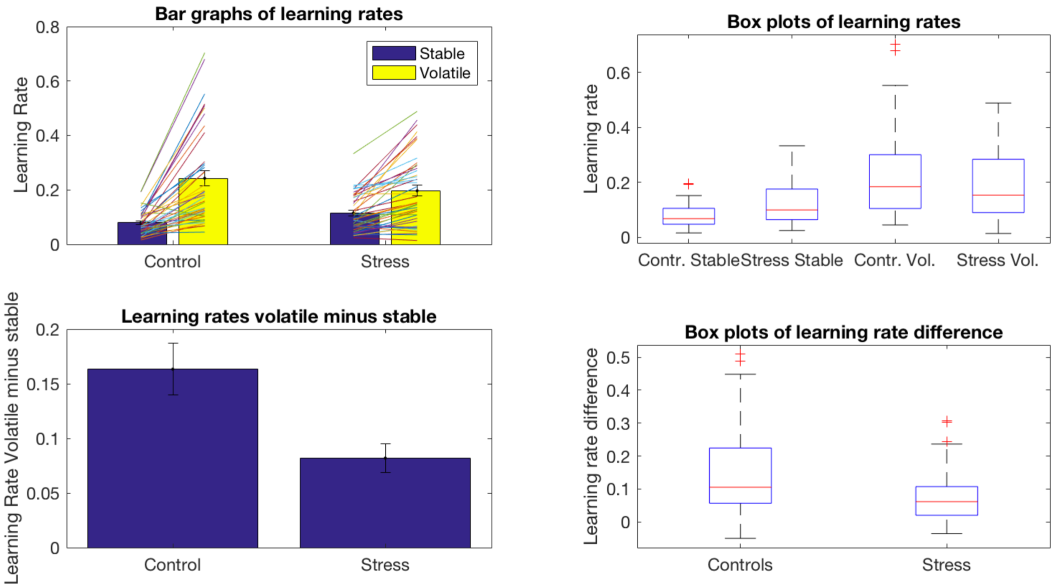
____

Make a note of what you conclude here in your practical report.

A) Paired samples t-test
We can check the hypothesis that learning is different in volatile and stable blocks using paired-samples t-tests. While t-tests can be quite robust to violations of their assumptions, it is nevertheless good practice to double-check the assumptions. Specifically, we should check that our data is approximately normally distributed and that we don't have any extreme outliers. We can do this by make:
- bar graphs and box plots
- histograms

A.1) Bar graphs and box plots
Here is an example (using fake data – this will look differently with the real data that you now have) that shows histograms of the average learning rates (with

standard errors of the mean) in the two groups could look like. On the right you can see the boxplots for the same data. In the second row you can see the same for the difference in learning rates between the volatile and the stable block (stored in the variable '**AlphaVolMinStable'** – you will need this again later).



Q: What do you notice from the bar graphs about how learning rates in the stable and the volatile block relate?

Q: Why is it useful that we also plot the data from individual participants on top of the bar graphs?

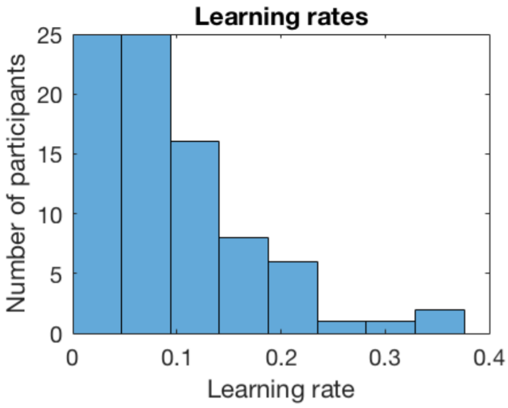
Q: What do the example box plots suggest to you? Are there any outliers?

-> Now let's use the Matlab code to have a look at the real data (execute the cell 'Bar plots, stable vs. volatile').

Q: How do the plots compare to the simulated data? How do you interpret this?

A.2) Histograms

Before we decide whether we can do t-tests with our data, let's also look at histograms. With a histogram, we can look at the distribution of learning rates and at the distribution of differences between learning rates in the stable and the volatile block. See below for an example histogram based on fake data (the histogram may look different for your real data).



Example of learning rates from a skewed distribution

Q: What does the histogram show? Does the data follow a normal distribution?

____

____

Q: Can you think of a reason why the learning rates might not follow a normal distribution?

____

____

Q: How else (other than visual inspection) could we check that our data is normally distributed? What is the advantage or disadvantage?

____

____

-> Now, run the code in the cell 'A.2 Histograms' to produce histograms of the real data.

Q: Why are we plotting histograms separately for the two groups? And why are we plotting histograms not only for the learning rates per se, but also for the difference in learning rates between the volatile and stable block?

____

____

Optional: if you prefer to use quantile-quantile plots instead of histograms to assess the distribution of your data, Matlab also has a function for this, called 'qqplot'. If you have time in the end of the practical you could explore what using this reveals.

A.3) Transforming the data

For the fake data (this might not be the case for the real data!), we can notice that the data is not normally distributed and that there are outliers. There are different ways for dealing with this problem. A simple way is to transform the data using a log transformation.

-> If you decide that your data is not normally distributed, you can try a log transformation by changing the line Alphas= '(Alphas)' in the cell 'Read data'. Write down the code for this here:
Alphas=____
Then replot all graphs.

Q: What do you observe?
____

There are alternative ways (for example if the transformation does not make the data more normally distributed) to deal with non-normal data: You can either exclude outliers or use non-parametric tests. We will look at the code for non-parametric tests below.

Q: Given what you have done in session 3 to check the data for outliers, why might you not want to additionally reject outliers at this point?
____
____

A.4) Doing paired-samples t-tests with Matlab
-> The Matlab command to do paired-samples t-tests is 'ttest', use the command 'doc ttest' to open the Matlab help to find out how you can enter the data to test (across all participants) whether the learning rate in the stable and volatile blocks is different. Then complete the code in cell A.4:
[h,p,ci,stats]= ttest(____)
Hint: don't forget to ensure that you only include subjects (across both groups) that you have not decided to reject in session 3 (i.e. use the variable 'InclSubs' to index the included participants)

Tipp: For your report, the outputs from the t-test that you will need to report are the p-value ('p'), the t-stastic (stats.tstat), the degrees of freedom (stats.df). For example, you could use the format: t(degrees of freedom)=t-statistic, p=p-value. You should also report the mean and standard error in each condition. You can find these in the cells of code that make the plots in the variables 'means' and stes'.

In addition, it is also good practice to report the effect sizes of the observed effects (e.g. Cohen's d or Hedge's g, which are quite similar). We can do this in Matlab using an additional toolbox (accessed with the command 'mes') that we have already included in the script folder for you - when using Matlab in the future, it is always worthwhile having a look on the MathWorks file exchange if someone has already made a function you are looking for. To use this, complete the code by entering the same info as you have just entered into the 'ttest' code into the two gaps for the 'mes' command (optional: if you are interested, type 'doc mes' to get more information about the syntax of the mes command):
effectSizeStats = mes(________, ________,'hedgesg','isDep',1,'nBoot',10000)
The effect size can then be found in effectSizeStats.hedgesg. To get a confidence interval on the effect size, look at effectSizeStats.hedgesgCi. For your report write-up, you may wish to consult the manual ('Documentation_MESToolbox),

found in the folder 'hhentschke-measures-of-effect-size-toolbox-b5d992f', for further information.

Q: Why would you want to report effect size in addition to the p-values? Start by thinking about what p-values and effect sizes measure.

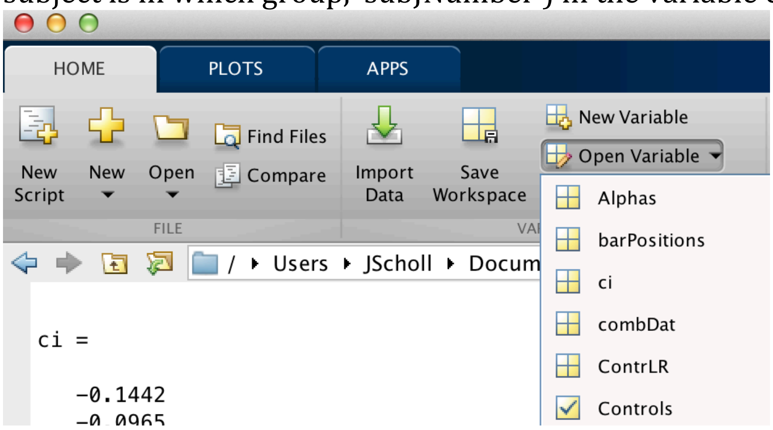
____

____

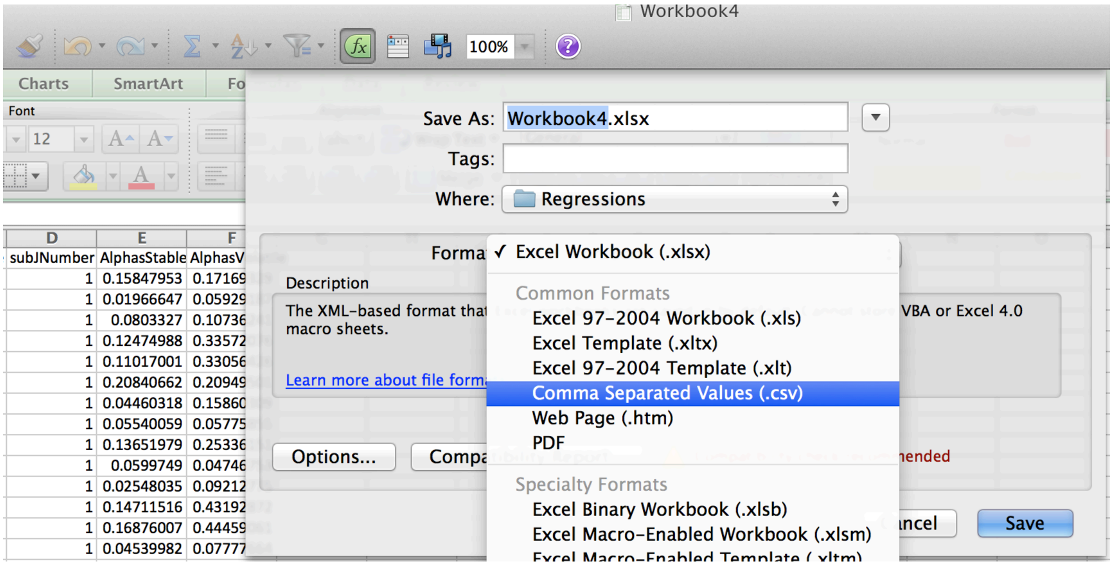
If the transformation did not solve your problem of non-normality, you can use non-parametric tests, such as the Wilcoxon signed-rank test (Matlab command: signrank); the corresponding measure from the mes toolbox is Cohen's U1 (consult the manual for an example how to write the code to get this).

Optional – Bayesian statistics to quantify evidence for the null hypothesis

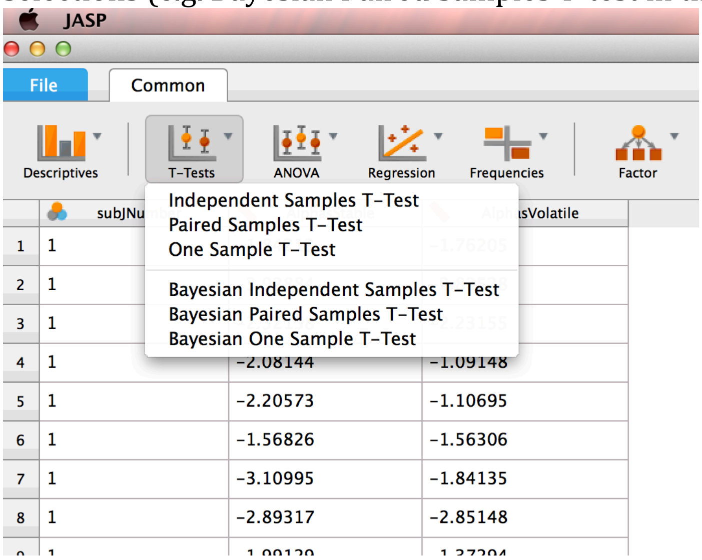
We will only get to this point if time permits: Recently, there has been a lot of discussion in Psychology about whether using p-values is problematic (especially when studies do not report effect sizes and mistakenly interpret p-values as the measure of the strength of the effect). If you are interested in the topic, Wagenmakers (2007) 'A practical solution to the pervasive problems of p values' is a good place to start. Several solutions have been suggested to this problem. One is to use Bayesian instead of frequentist statistics. In short, while a p-value means 'how unlikely is the data if the null hypothesis were true', in the Bayesian framework, we use 'Bayes Factors' that mean 'given the data, my hypothesis is X times more likely to be true than the null hypothesis'. A very practical advantage of this is that with a Bayes' Factor, we can also make a statement like 'given the data, the null hypothesis is X times more likely than my hypothesis'; rather than in a frequentist framework where we can only say 'I cannot reject the null hypothesis'. If you want to use Bayesian statistics, you can do this very easily using the freely available software JASP (https://jasp-stats.org). To get your data from Matlab to JASP, the easiest way (without involving much coding) is to open the variables (e.g. 'Alphas' and the index which subject is in which group, 'subjNumber') in the variable editor:



Then copy the variables into an Excel sheet, give them labels and save them as .csv file (make sure to delete the rows of participants that you don't want to include):

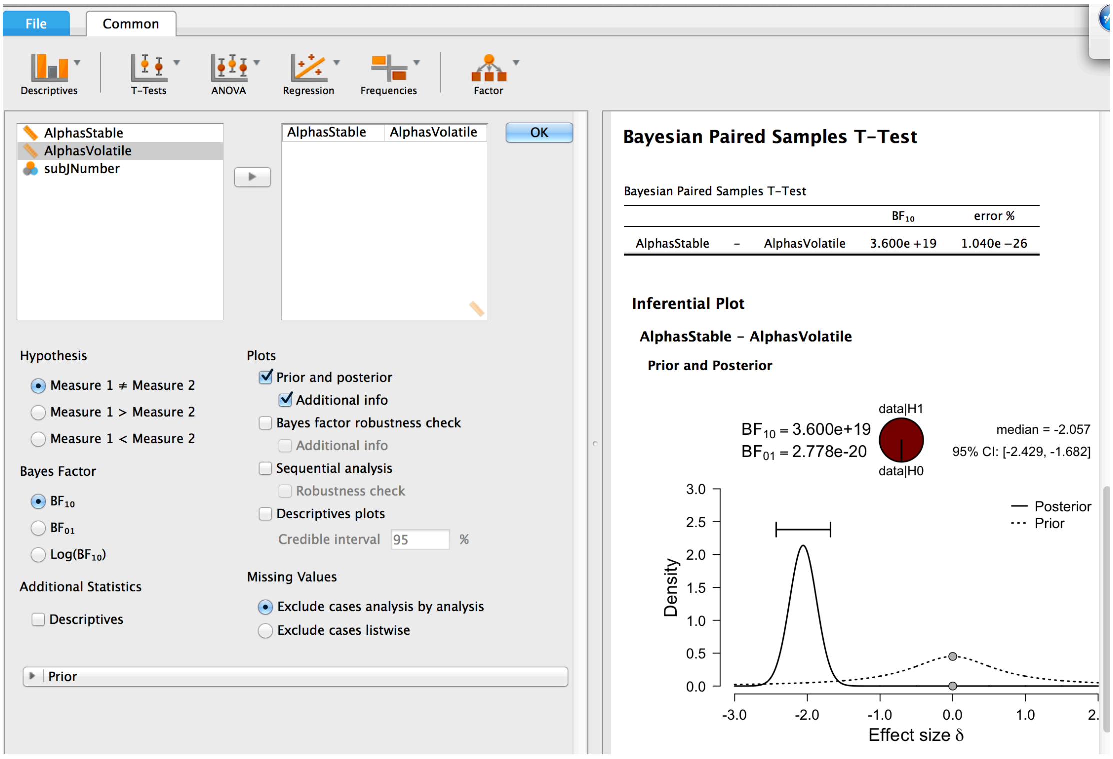


You can then open the file in JASP. To repeat the t-tests that you have done before or to run Bayesian t-tests, click on the T-Tests button and then make your selections (e.g. Bayesian Paired Samples T-test in this case).

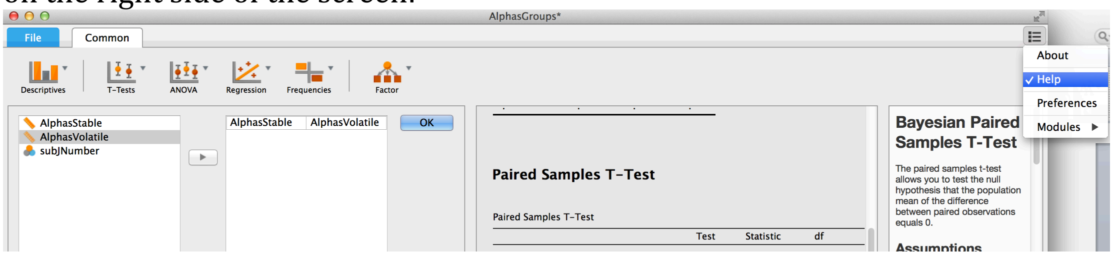


Then put the variables of interest into the variable selection window and tick the boxes for 'prior and posterior' and 'additional info' to also get estimates of effect sizes. The Bayes Factors (i.e. how much more likely your data is given your hypothesis vs. the Null hypothesis) are given in the table. JASP also gives you an estimate of the effect size (with a median and the '95% Credible Interval, CI – which means that there is a 95% probability that the true effect size is in that interval).

If you want to include Bayesian statistics in your report to measure evidence for the null hypothesis, you would probably want to include the Bayes Factor (how much more likely is the null hypothesis to be true than your experimental hypothesis) and the credible interval for the effect size. For the credible interval, you want to check whether the effect size includes 0 or not.



While JASP is very intuitive to use, if you want to get help, select the help button on the right side of the screen:



B) Independent-samples t-test

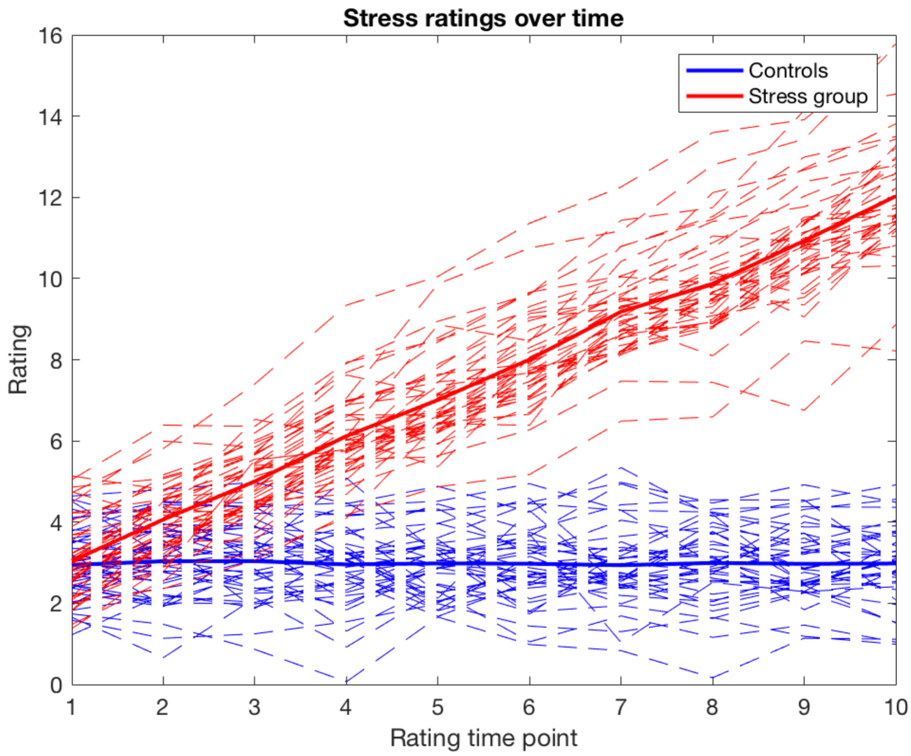
B.1) Checking the stress manipulation

For each participant, we have stored their 8 ratings of how stressed or happy they feel in the array 'Ratings'. This array is created in the first cell. You have to decide how to define it – do you want to use the stress ratings, or the happiness ratings? Or do you want to make a combination of the two? How did you decide this? (There is no right or wrong answer)

____

____

Whichever you choose, before doing any statistical tests, let's first inspect the data: In the Matlab cell 'B.1 Checking the stress manipulation' execute the code until the 'set(gca' … command (i.e. stop before the 'keyboard'). This should produce a figure similar to the one below that we've created using some simulated data:



Q: Describe what you see on this figure: How do the two groups differ?

____

____

Q: Now compare this to what you have obtained for your real data. What do you observer?

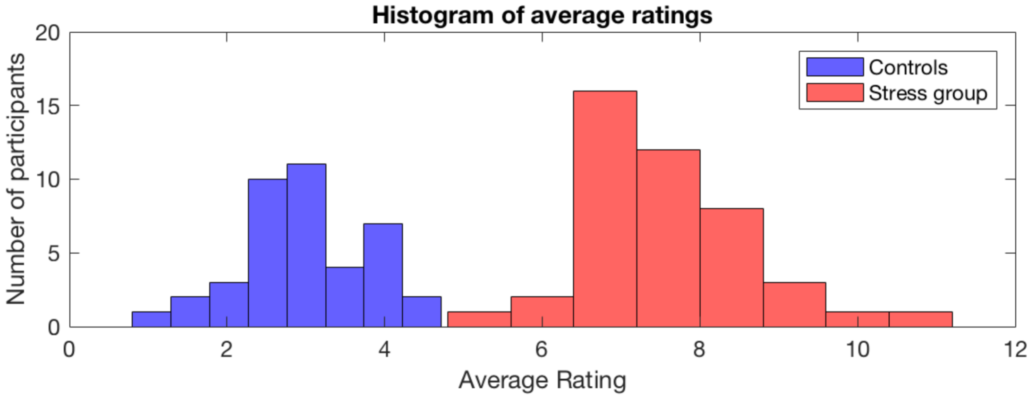
____

____

There are different ways how we could statistically test whether the manipulation has worked, i.e. whether the participants in the 'Stress' group feel more stressed than in the 'Control' group. One way is to compare the average stress rating.

-> In the cell 'B.1', first compute the average stress rating for each person. Hint: look at the Matlab help for 'mean' to find out how to make the mean of an array along the dimension that you need, i.e. make the average across the different time points for each person, rather than across the participants.
avgRating = ____

The next step, as before is to check the distributions. The plot below shows this for the example data. As you can see, the distribution is not strikingly non-normal.



Q: How does the real data look? And in case it does not look normal, how do you choose to address this?

____

____

Now you are ready to do the test:
-> Using the 'ttest2' command, do an independent samples t-test of the ratings in the two groups.
[h, p, ci, stats] = ttest2(_______________,_______________)
-> As before, compute the effect size using the 'mes' command (this time omitting "'isDep',1" because we want to know the effect size for an independent samples t-test.
effectSizeStats= mes(_______________,_______________,'hedgesg')

Q: We have checked the stress manipulation by comparing the average of the 10 ratings each participant made in the session. What could we have tested alternatively and why might this have been advantageous?

____

____

____

____

Optional: if you want to repeat the test in JASP to obtain Bayesian statistics you need to copy the variables 'Controls' and the variable 'avgRating' into 2 columns in Excel and again save it as .csv file.

Group	average Rating
1	3.01889
1	3.88321
1	2.62202
1	3.45484
1	3.98812
1	3.26749
1	4.04112
1	2.61855
1	3.19675

In JASP, then select ‘T-Test’> ‘Bayesianndependent samples t-test’ (to do a non-Bayesian t-test), assign the ratings as the dependent variable and ‘group’ as grouping variable and tick the boxes ‘Prior and posterior’ (and underneath it ‘Additional info’ in the ‘Plots’ selection’ to get both Bayes Factors and estimates of the effect size.

B.2) Does stress affect the ability to adapt learning rates to the environment?

We will use an independent samples t-test to test whether in the stress condition, participants adapt their learning rate less to the environment. In the section ‘A1’, we have already made the bar plots to show the data (variable ‘AlphaVolMinStable’, i.e. the difference between the learning rates in the two blocks, plotted separately for the two groups) and in ‘A2’ we have checked the distribution of these values is approximately normal.
-> Now all that remains to be done is run the t-test and obtain the effect size using the same commands as you have done in ‘B.1’

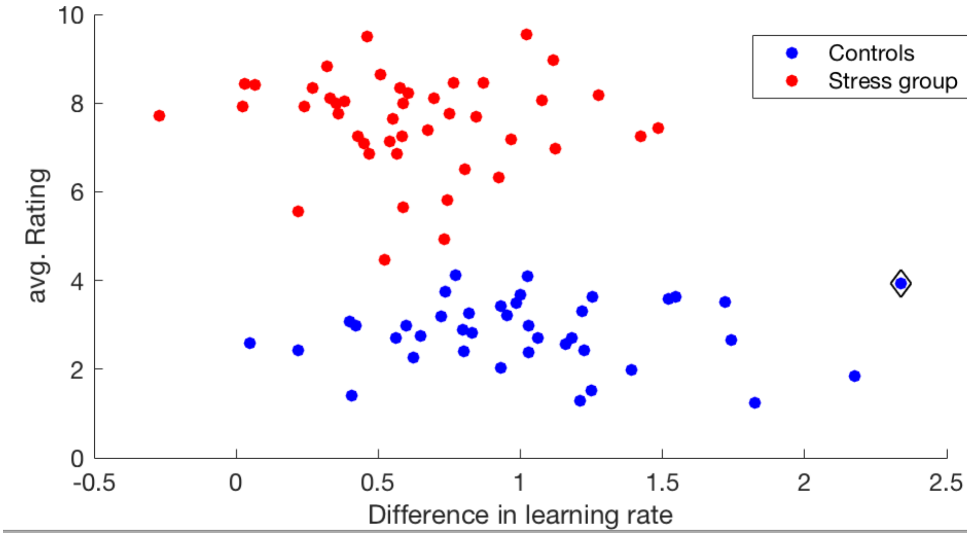
C) Correlation between average stress and adaptation of learning rates

In addition to comparing the change in learning rates between the two groups, we can also test for a correlation between the average stress level of each subject and how much they adapt their learning rates.
Q: Why might we want to test this in addition to the t-test comparing the groups? What is your hypothesis what you should find?

____

____

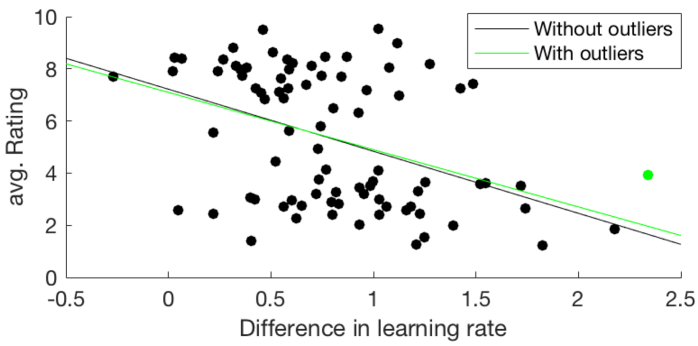
Before running a correlation, we first need to check whether the assumptions of a Pearson correlation are met. We need to check that the potential relationship that we are testing is linear (rather than e.g. following a hyperbolic curve) and that there are no significant outliers. The last point is especially important because correlations are sensitive to outliers. (In addition you also need to check that the data follows an approximately normal distribution, but you have already done this above).
-> To check these assumptions, we will do a scatter plot in Matlab (in the cell B.2). In the plots, the two groups are shown with different colours and outliers are highlighted with a diamond shape.



Q: What do you observe when you run this analyses with your real data?

Q: Have a look at the line of Matlab code that defines outliers ('outliersLRdiffs = …') Do you understand what it does? What cutoff is used to exclude outliers? And based on which data are the mean and standard deviation determined?

If you notice that you have outliers in your data, you again have several options (with again pros and cons). For example, you could exclude the outliers. However, given that you have already excluded subjects in session 3 that did not perform the task correctly, it is unlikely that the outliers that you find are the result of an experimental error. Therefore, it might be more appropriate to include the outliers. If so, you need to check whether the correlation changes if you exclude the outlier (and report this in your write what). In the Matlab code, we are plotting the least square lines (i.e. the best fitting straight line) for you with and without the outliers (following the comment '%Plot least square lines with and without outliers')



Q: In the example above, what do you think the plot tells us about how influential the outliers are?

Q: What do you think you should do if you only find a correlation if you include the outlier(s)?

____

____

An alternative to excluding outliers is to use a non-parametric correlation. Both non-parametric and parametric correlations in Matlab are done with the 'corr' command.
-> Using the Matlab documentation, write the code to do a parametric or non-parametric correlation (as appropriate):
[rho, pval] = corr(____
When you report these statistics in your write-up, make sure to report: the number of participants in the test, the correlation coefficient (e.g. Pearson's r) and the p-value. Also make a note of whether the test was parametric or non-parametric and whether outliers were excluded and what the effect of this was.

D) Some plots to help you write the report
[Optional for today during the session – but look at this for writing your report].
D.1-D.2
In these sections, Matlab is making some plots for you that might come handy when you write your report. You should already be familiar with these plots from the previous sessions. They show a schematic of the task and validations of the computational model. It is putting different elements from the previous sessions together. You will need to add labels to include the plots in your report (if you are stuck, have a look at examples for previous sessions or consult the Matlab help). And you may wish (optional) to make the figures look nicer by changing some of the figure settings (consult the Matlab help or look online for tutorials).

Reminder about validating your model. Also have a look at your notes for session 2 in which we covered the special case of validating that we have enough trials in our experiment. The aim of validating your model is to make sure that the parameters that you obtain when you fit a model to the real data reflect aspects of the participants' behaviour, or in other words the computational model that participants' brains used to solve the task (rather than e.g., aspects of the specific schedule/set of trials that you have used).

We can validate our data by simulating behaviour (i.e. choices on each trial of the experiment) based on parameters that we choose. For example, we can set the learning rate in the stable block to 0.2, the learning rate in the volatile block to 0.4 and the inverse temperature to 1; we then simulate what choices a participants with these parameter values would make on each trial. The next step is then to analyse the simulated data with the model that we would ultimately want to use to analyse the real data.

We can then compare how well simulated and fitted (or 'recovered' – called like this because we know the ground truth of the real parameters and we are trying to recover the ground truth) parameters relate. We can also check whether when we simulate data from a model that has our effect of interest (i.e. a model for which our experimental hypothesis is true; in our example this would be a lower learning rate for the stable compared to the volatile block) the recovered parameters show the effect of interest (as a statistically significant result).

Importantly, we also need to test whether when we instead use simulated data that does not have our effect of interest (e.g. the simulation was run with the same learning rate for stable and volatile blocks), the recovered parameters similarly do not show a significant test result (e.g. no significant difference in the learning rate parameters fitted to the stable and volatile block).

Q: Why do you think it is important to also simulate data from a model that does not have our effect of interest? If you are interesting in finding out more details about this, have a look at Palminteri, Wyart and Koechlin (2017, Trends in Cognitive Science)

____

____

Concerning the code that we are using here, the data is organized as follows:
SimulData = data that we have simulated for you and then fitted with the different models
SimulData.OneLearningRate: in the simulation, the same learning rate was used for the stable and the volatile block
SimulData.TwoLearningRates: in the simulation, separate learning rates were used for the stable and the volatile block (we set them to be higher in the volatile than in the stable block in agreement with our hypothesis)
SimulData.OneLearningRate.OriginalParameters: a record of the original (ground truth) parameters with which we made the simulation
SimulData.OneLearningRate.OneLearningRateModel: results from fitting the simulated data with the model with one learning rate. Subfields here: Parameters: the fitted/recovered parameters; negLogLikelihood: the error term for the model fit; AIC: the Akaishi information criterion for the model fit.
SimulData.OneLearningRate.TwoLearningRatesModel: the same simulated data, but now fitted with the model that estimates separate learning rates for the stable and the volatile block.

D.3 Excluding participants using a regression analyses

This part is optional. For your report, we only expect you to exclude participants using rules as covered in session 3. However, as you may have noticed in session 3, some participants seemed to ignore for example the magnitudes and only use the reward probabilities that they had learnt to make decisions.
Q: Can you think of why these participants pose a problem for estimating learning rates using our computational model?

____

____

It therefore makes sense to exclude participants that did not take magnitudes into account. We can quantify whether participants in fact used magnitudes to make their decisions using a logistic regression analysis (rather than using visual inspection as you did last week). In a regression, we predict participants choices (the left or the right option on each trial) based on the magnitudes and the true probabilities. Note that if we wanted to do it even more correctly, we should not use the true probabilities (do you know why?). And instead, we could use a new computational model that tells us what participants could know if they used the ideal learning rate for a given block (we will not cover this; if you are interested, have a look at O'Reilly (2013, Frontiers in Neuroscience) for an easy to read introduction).

When you run cell D.3 you will get a figure that shows the regression weights for each person for each predictor of interest. These regression weights tell you to what extend participants decision are influenced by the regressors/predictors. You can for example notice that most people have a positive regression weight for the reward probability of an option. This means that they are more likely to pick an option if it has a better reward probability.

Q: How could you use the results of this regression to exclude participants?

____

____

Part 2: REPORT GUIDELINES
Your report should be as much to the point as possible. As this block practical is concerned with data analysis using computational model, we advice you to focus most on the methods, results and discussion section, keeping the introduction to a minimum.

Abstract
Keep this very short (~120 words). Allocate between one two sentences to summarize each section of the report

Introduction:
Keep this very short: 200-400 words. Including up to 3 citations max (for example – but also feel free to pick others : Behrens et al. (2007, Nature Neuroscience; Daw and Tobler (2014, chapter 15 in Neuroeconomics (2nd edition), edited by Glimcher and Fehr; Arnsten (2009, Nature Reviews Neuroscience). The introduction should cover:

1) What each experimental question is and why we ask this question (by citing relevant previous work)
2) How exactly each experimental question will be tested in the experiment and what you predict would happen if the

Methods
This is an important section of the report. A key aspect we are looking for here is that it is clear why you do the analyses you do and how you do them. We recommend that you look back through your notes and questions-and-answers from the previous sessions (and today) to help you with this section.
This section should cover:

1) Participants: number of participants and whether you had to exclude any participants
2) Experimental task design: Look back at the hand-outs from the previous weeks (weeks one and two). These will already include descriptions of the experimental design. Feel free to re-use them. For the participants, include here whether you had to exclude any participants and how you made that decision (make this decision using code from session 3).
3) Simple analyses of the data (as covered in session 3): describe how you use analyses that do not rely on a computational model to check that a) participants have performed the task correctly, b) test whether the manipulation of stressful vs. neutral sounds has worked (as covered today).
4) Description of the computational model (with equations) used to analyse the data: Again, look at the descriptions from the hand outs (session 1-3) to complete this section. Do not forget to mention (here or in the

discussion) what the advantage of a computational model is compared to the simple analyses done above. Also include how the model has been fitted to the data (session 3)

5) Model validation: describe what is meant by model validation (covered in session two), why you do it and how exactly you do it (including what statistical tests, covered in session 4 today; code to help you make the figures is covered in the end of session 4 today)
6) Application of the computational model to the data (covered in session 3 and today): describe what measures (i.e. model parameters) that you have derived from the computational model you will use to test your experimental hypotheses. Include what statistical tests you will do and how you will check that these tests are appropriate. Also explain how you checked that your model fits the data well (see model comparison covered in session 3).

Results
This should follow the same structure as the methods section above, i.e. you should show the results of the analyses that you describe in the methods section.

For each analysis we expect you to summarize briefly what you want to test with the analysis and how the result of the analysis relates to your question, including a figure showing the data and appropriate statistical tests (see session today):

1) Figure: Each figure should have a number, a title and a short description of what is shown. You should be able to make most figures that you might need using the script from today's session. Make sure that the axes are labelled, have the correct units and that if there are several different lines on a plot, add a legend (in the figure) to describe what they mean. When reporting data across subjects, plot the average with error bars showing the standard error of the mean (and make a note of this in the figure legend).
2) Statistical tests (see session today): report all aspects of the test; e.g. for a t-test the degrees of freedom, the t-value and the p-value, as well as the effect size.

Discussion
Keep this section below 1000-1500 words. Given that this block practical is about using modelling for data analysis, the main focus of your discussion should be on interpreting the results, considering limitations to the interpretation (given the method) and discussion of additional analyses you could do in the future if you had more time. You should focus less (in terms of words) on relating these findings to previous work.

1) Start this section with a short summary (about 2-4 sentences) of the experiment and the main findings.
2) Then describe in detail how you interpret the different analyses; what do they mean? Were the results as you expected? If not, why? Hint: look at previous paper (e.g. Behrens et al. 2007 or Browning et al. 2015) and see how their models differed from ours. Does this suggest additional analyses you could do if you had more time? If you are interested (this is not a requirement for your discussion), you could also note in these

papers how they fit their data instead of fminsearch. Does this suggest directions for future analyses studies?

3) Very briefly discuss how the findings relate to the previous literature
4) Limitations to the analyses and future analyses: Describe whether any of the analyses done were not yet decisive and what analyses you could have done additionally to corroborate your findings further. For this, it might be useful to look back at your answers in the previous sessions where you were asked to reflect on additional analyses you could have done.

Summary of marking criteria

Below is a summary of the criteria that will be applied to mark your reports:

- demonstrate clear understanding of the experimental design and aims of the experiment
- show understanding of the rational for and the principles of the computational methods used to analyse the experiment
- demonstrate ability to use MATLAB to:
    - generate meaningful diagrams (with appropriate labels) to reflect results
    - use MATLAB to run statistical tests for key experimental hypotheses
    - use MATLAB outputs (figures and statistics) to generate a coherent results section
- critically interpret these results and limitations of the data set and the analyses performed, also with reference to relevant literature