Homework Project 3 - part 1

In this project, you will analyse data from one observer an experiment in which one observer was presented with 8 types of stimuli.

- Sound at an azimuthal angle of 0 or 30 degrees
- Light at an azimuthal angle of 0 or 30 degrees
- All 4 audiovisual combinations of the sound and light.

Each type of stimulus was presented 20 times. The order of the stimulus presentations was randomised.

The task of the observer was to point a laser pointer towards the sound for audio and audiovisual stimuli and towards the light for visual stimuli. The azimuthal angle of the response was recorded in 1 degree intervals but you can assume that the responses are distributed on a continuum.

The data comes in in a .csv file where each row contains the responses to one stimulus type. The correspondence between stimulus type and row number is described in the table below.

Row number	1	2	3	4	5	6	7	8
Auditory stimulus location	0	30	-	-	0	0	30	30
Visual stimulus location	-	-	0	30	0	30	0	30

Visualising the data

- Plot a probability density histogram for each type of stimulus. This height of each bar should be the proportion of responses divided by the width of the column.
- Calculate the mean and standard deviation of the responses for each type of stimulus. Plot the corresponding Gaussian probability density function.
- Based on the above, estimate whether the observer experiences the enhancement effects and the ventriloquist illusion.
- Check by visual inspection if the responses for each stimulus type follows a Gaussian distribution.
- You figure could look something like this

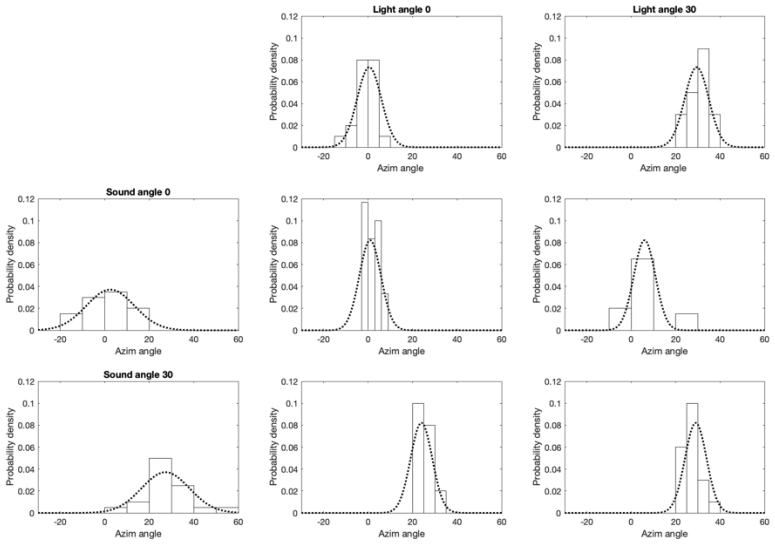

Figure 1 - Data visualisation: probability density histograms and corresponding Gaussian probability densities for each stimulus type. Rows correspond to the location of the sound and columns correspond to the location of the light.

Fit the Strong Fusion (MLE) model to the data

- Use all the data to estimate the free parameters (the mean and standard deviation for each of the two auditory stimuli and each of the two visual stimuli).
- List the values of the parameter estimates and the negative log likelihood of the fits.
- Plot the models' probability density and the corresponding histogram for each stimulus type. Evaluate the model fits by visual inspection.

Check for over-fitting using cross-validation

- Check that the results of the model fits are not due to over-fitting using 8-fold leave-one-stimulus out cross-validation. In each fold, the validation set consist of the responses to one type of stimulus. The training set is the responses to the other 7 types of stimuli.
- List the validation negative log likelihood for the two models. Compare your finding to the model fits.
- Visualise the results of the cross-validation. For each fold/stimulus type, calculate the probability density for the validation stimulus using the parameters estimated from the training set. Plot the corresponding Gaussian probability density function for each stimulus type and the corresponding histogram. Evaluate the cross-validation results by visual inspection.

Homework Project 3 - part 2

In this part of the homework project we will work with the Bayesian model of signal detection under the unequal variance assumption. Calculate the probability of a true positive response and the probability of a false positive response for an observer, which behaves according to a Bayesian model with perceptual sensitivity parameters $\mu = 1.2$ and $\sigma = 1.8$ for prior probability *P*(*s*)=1/3 and for prior probability *P*(*s*)=2/3.

Fit the unequal variance signal detection theory model to the response probabilities that you calculated above. List the parameter values for perceptual sensitivity parameters μ and σ. Compare them to the true underlying values and discuss your result.

Homework Project 3 - part 3

In this part of the homework project we will fit three Bayesian models of multisensory integration to data from an experiment on audiovisual speech perception. The three models are described in Section 3.3 of the lecture notes: The early strong fusion model, the probability matching model and the late strong fusion model. The data has been used for testing Bayesian models of audiovisual integration in the literature and is described in the lecture notes and in the paper by Andersen (JASA, 2015).

We will compare the models based on their fit measured by the negative log likelihood and also using cross-validation. As both models aim to predict how, say, an audiovisual stimulus is perceived from how the auditory and visual stimulus components are perceived we will employ leave-one-stimulus cross-validation where one stimulus is left out of the training set and used for validation. Note that the models may also predict how, say, an auditory stimulus is perceived from how an audiovisual and a visual stimulus is perceived. Therefore we will use every stimulus as a validation stimulus. As there are 35 stimuli in one experiment this means that we can cross-validate across 35 folds.

The data

The data is provided to you as five csv files. Each file contain data from one observer and contains a 7-by-5 array with response counts. The response count ranges from 0 to 24 and is the number of trials out of 24 that the observer responded that the speech stimulus was perceived as ’d’.

The first row of the array contains response counts for the auditory stimuli. The columns of the first row indicate the specific auditory stimulus ranging from very similar to a ’b’ in the first column to very similar to a ’d’ in the last column. Likewise, the second row contains response counts to the visual stimulus ranging from very similar to a ’b’ to very similar to a ’d’. Rows 3-7 contains response counts for the audiovisual stimuli. Rows indicate the visual component of the stimulus ranging from very similar to a ’b’ (row 3) to very similar to a ’d’ (row 7). Columns indicate the auditory component of the stimulus ranging from very similar to a ’b’ (column 1) to very similar to a ’d’ (column 5).

Your report

Your answer to this part of the homework assignment should contain the following

- The exact formula for the negative log likelihood function you used to fit each of the three model to the data. Specify the free parameters and data in the formula.
- A table with parameter values of the fit for each model for each subject
- A table with the negative log likelihoods for the fit of the three models models to the data from each subject. (Hint: they should lie in the range of 53.54 - 72.80) for the early strong fusion model
- A table with the validation negative log likelihoods summed across folds for both models.

- An evaluation of the three models in terms of training and validation error. Provide details such as whether the models overfit, whether one model is better for all stimuli and observers.
- A scatter plot of the response probabilities vs. the response proportions pooled across observers for each model. Make scatter plots for the response probabilities that you got from fitting the models to all the data and also for the response probabilities that you got from cross validation. Evaluate the models by visual inspection of the plots.