Tutorial on statistical analysis of single-neuron spiking activity

Joaquín Rapela*

July 1, 2021

This tutorial will provide participants with computational experience (e.g., statistics, programming, plotting) to better understand single-neuron spiking activity.

We will use spikes from one neuron recorded by Cristina Mazuski from the amygdala of a male rat while he was interacting, in different sessions, with two other female rats (i.e., `female1` and `female2`). We will try to find features of these recordings from which a downstream neuron could decode the identity of the interacting rat (`female1` or `female2`) by only looking at the spiking activity of the the analyzed neuron.

In the first part of the tutorial we will try to do this decoding using various statistical measures of the recorded data (e.g., spike counts, inter-spike intervals, autocorrelations; Section 1). In the second part we will attempt to infer response properties of the recorded neuron using statistical models and to decode the identity of the interacting rat based on these inferred properties (Section 2).

This tutorial will be very hands-on and provide practical experience in:

Statistical neuroscience

- descriptive statistics for point processes (plots of increment processes, histograms of ISIs),

- statistics for spiking activity (Fano factor, autocorrelations of increments and ISIs),

- probabilistic models for spikes (e.g., Poisson process, inverse Gaussian process),

- maximum likelihood parameter estimation,

- confidence intervals and hypothesis tests (parametric, simulation),

Python programming

- code organization: packages, modules, classes, functions,

- object-oriented programming,

- Python scientific programming libraries: numpy, scipy,

- interactive data visualization,

- exceptions, containers, command-line-arguments, . . .

---

*j.rapela@ucl.ac.uk

Open-science tools

- git, Github,

- Ancanconda,

- interactive data visualilzation (as in this document)

This tutorial will be jointly lead by a group of GCNU and SWC members (e.g., Adam Tyson, Cristina Mazuski, Lucas Simões, Joaquín Rapela). The code and data to generate all the figures in this tutorial appear here.

1 Descriptive statistics

We will display and apply statistical tests to:

1. spike times (Figure 1, doPlotSpikeTimes.py).

   [Figure 1 about here.]

2. inter-spike-intervals histograms (ISIs, Figure 2, doPlotISIsHistograms.py).

   [Figure 2 about here.]

3. binned spike increments (Figure 3, doPlotIncrements.py).

   [Figure 3 about here.]

4. autocorrelations between increments (Figures 4 and Figure 5, doPlotIncrementsAutocorrelations.py).

   [Figure 4 about here.]

   [Figure 5 about here.]

5. autocorrelations between ISIs (Figures 6 and Figure 7, doPlotISIsAutocorrelations.py).

   [Figure 6 about here.]

   [Figure 7 about here.]

2 Inferential statistics

We will fit statistical models to the ISIs from the interactions with `female1` and `female2`. We will use two types of statistical models for ISIs: exponential (Section 2.1) and inverse Gaussian (Section 2.2).

To try to decode the identity of the interaction using these models, we will take two approaches. First we will test if the estimated parameters of these models are statistically different from each other. Second, we will build a Naive Bayes Classifier to decode the identity of the interaction from calculated ISIs, we will build confusion matrices and derive statistical measures from them to assess the accuracy of these decodings.

2.1 Exponential model

Figure 8 (doLearnExpModel.py) shows histograms of ISIs and their fits by an exponential model.

2.1.1 Significant parameters differences

The title of Figure 8 shows the parameters estimated for each exponential model. The model for `female1` appears to have a larger λ parameter than that for `female2`. To test if this difference is statistical significant, we performed a bootstrap hypothesis test, which result is show in Figure 9 (doTestDiffLambdaExpModels.py) . This test indicates that the difference is not significant at the 0.05 level.

[Figure 8 about here.]

[Figure 9 about here.]

2.1.2 Decoding

Figure 10 (doDecodeExponential.bash) shows the confusion matrix corresponding to decodings from the exponential model. The title of this figure shows the corresponding precision, recall and f1-score. Decodings from the exponential model are at chance.

[Figure 10 about here.]

2.2 Inverse Gaussian model

Figure 11 (doLearnInverseGaussianModel.py) shows histograms of ISIs and their fits by an inverse Gaussian model.

[Figure 11 about here.]

2.2.1 Significant parameters differences

The title of Figure 11 shows the parameters estimated for each inverse Gaussian model. The model for `female1` appears to have a smaller μ parameter than that for `female2`. To test if this difference is statistical significant, we performed a bootstrap hypothesis test, which results are show in Figure 12 (doTestDiffParamInvGaussianModels.py). This test indicates that the difference is not significant at the 0.05 level.

[Figure 12 about here.]

From the title of Figure 11, the model for `female1` appears to have a smaller λ parameter than that for `female2`. To test if this difference is statistical significant, we performed a bootstrap hypothesis test, which results are show in Figure 13 (doTestDiffParamInvGaussianModels.py) . This test indicates that the difference is significant at the 0.05 level.

[Figure 13 about here.]

2.2.2 Decoding

Figure 14 (doDecodeInvGaussian.bash) shows the confusion matrix corresponding to decodings from the inverse Gaussian model. The title of this figure shows the corresponding precision, recall and f1-score. Decodings from the inverse Gaussian model are excellent (i.e., they achieve large precision and recall values).

[Figure 14 about here.]

It is suspicious that while we could not find significant differences between the spiking activity of `female1` and `female2` in the descriptive statistics measured in Section 1, or in the parameter values of the models estimated in this section (except the λ parameter of the inverse Gaussian model), and while the decodings using the exponential model performed at chance, we are now observing excellent decodings using the inverse Gaussian model. This suggests that decodings of the inverse Gaussian model are artifactual. To address this concern we developed the following control.

We built a control dataset, shuffling the labels `female1` and `female2` of the calculated ISIs. If decodings using the inverse Gaussian model are not artifactual we should observe chance decodings for the control dataset, as shown in Figure 15 (doDecodeInvGaussian_control.bash).

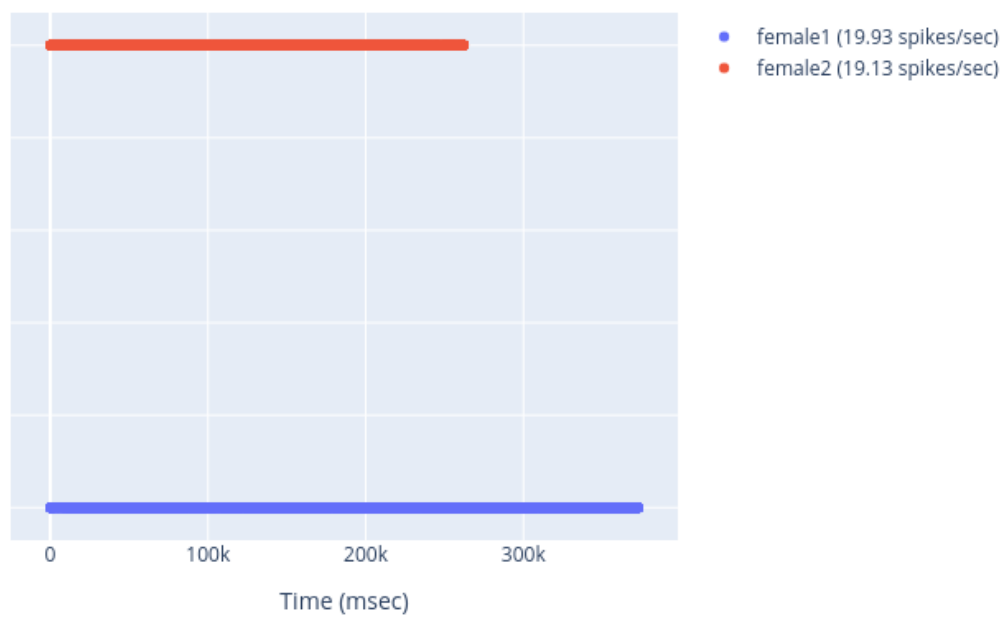[Figure 15 about here.]

4

List of Figures

Figure 1: Spikes times. Click on the figure to see its interactive version.

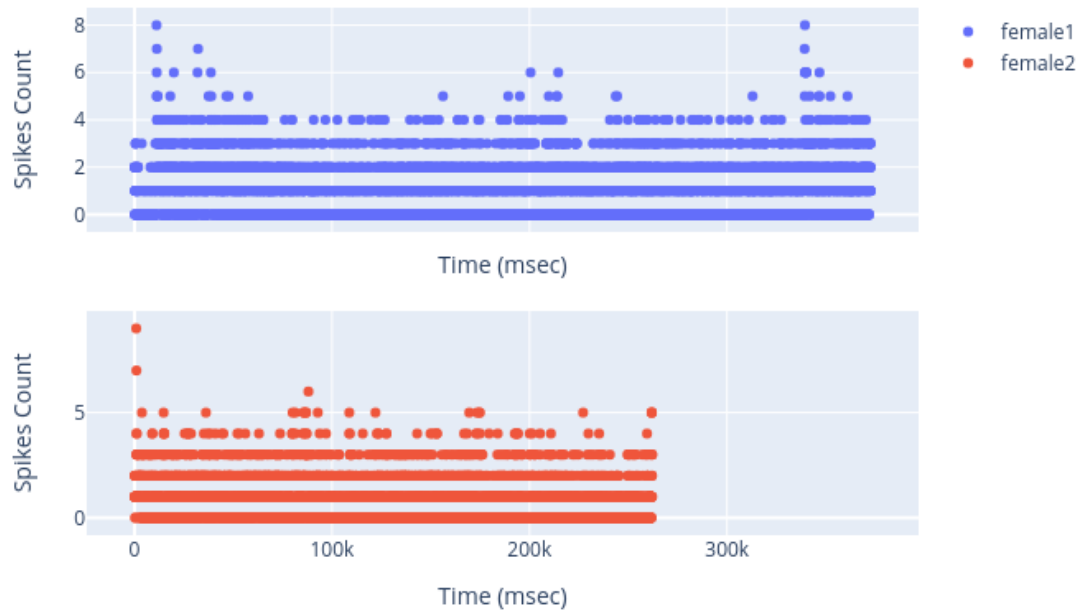Figure 2: ISIs histograms. Click on the figure to see its interactive version.

Figure 3: Binned spike increments, Fano factors and their 95% bootstrap confidence intervals. Click on the figure to see its interactive version.
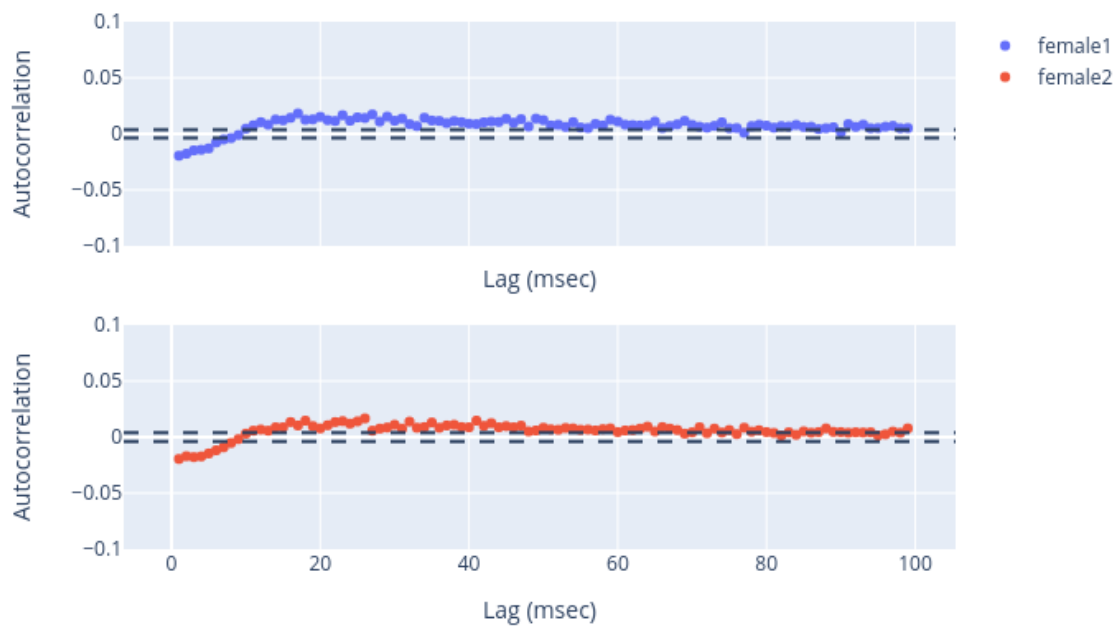
Figure 4: Binned spike increments autocorrelations, and their 95% approximate confidence intervals for lack of correlation. Click on the figure to see its interactive version.
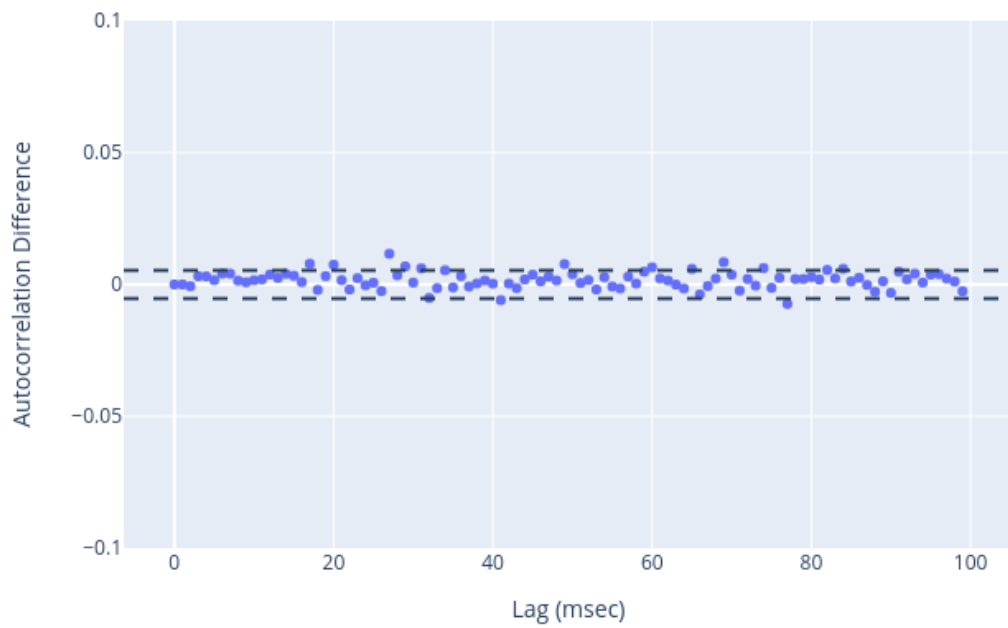
Figure 5: Difference between the increments autocorrelations of `female1` and `female2`, and their 95% approximate confidence intervals for lack of significance difference. Click on the figure to see its interactive version.
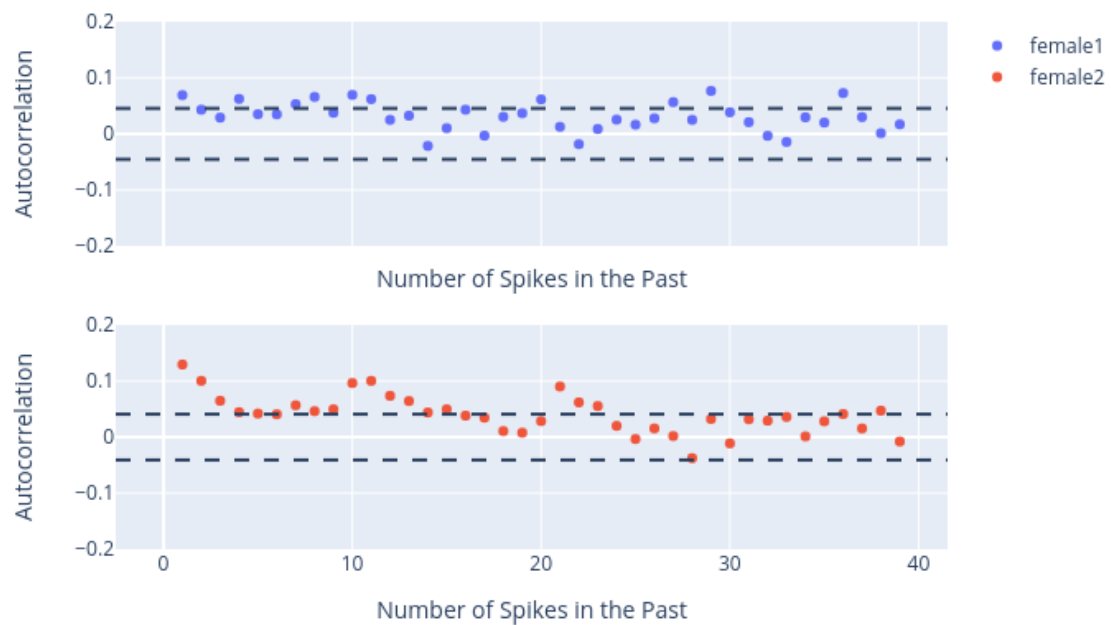
Figure 6: ISIs autocorrelations, and their 95% approximate confidence intervals for lack of correlation. Click on the figure to see its interactive version.
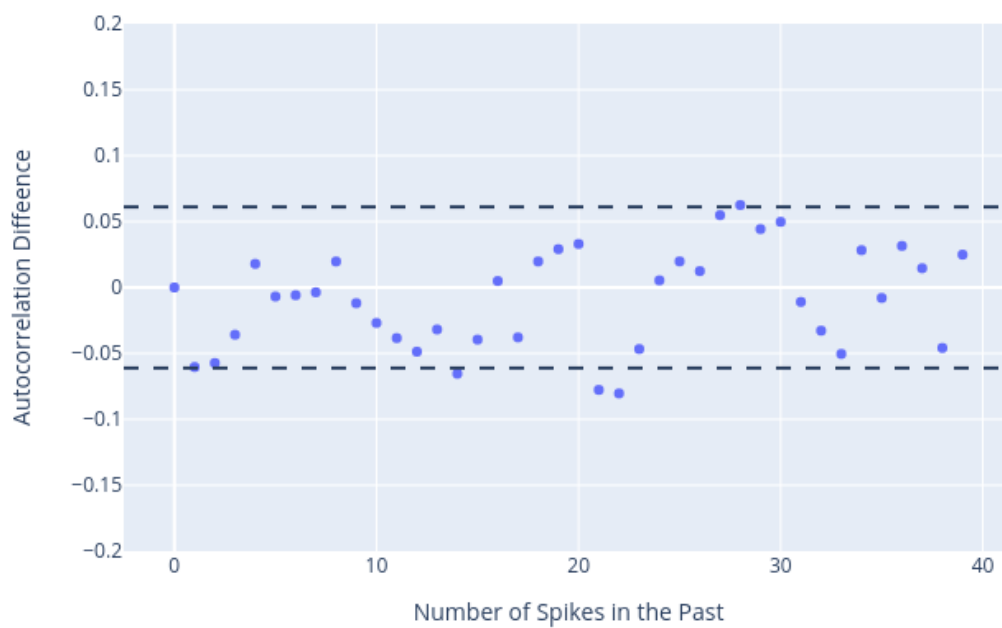
Figure 7: Difference between the ISIs autocorrelations of `female1` and `female2`, and their 95% approximate confidence intervals for lack of significance difference. Click on the figure to see its interactive version.

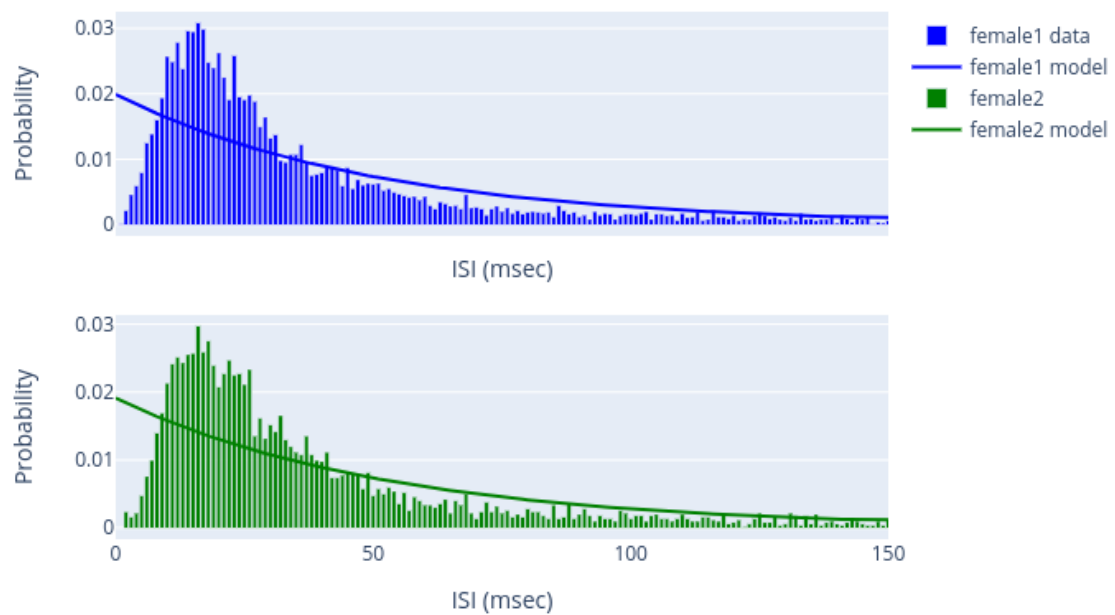Female1: $\lambda = 19.93$ Hz, Female2: $\lambda = 19.13$ Hz

Figure 8: ISIs and their fits by an exponential model. The title shows the estimated parameters for each model. Click on the figure to see its interactive version.
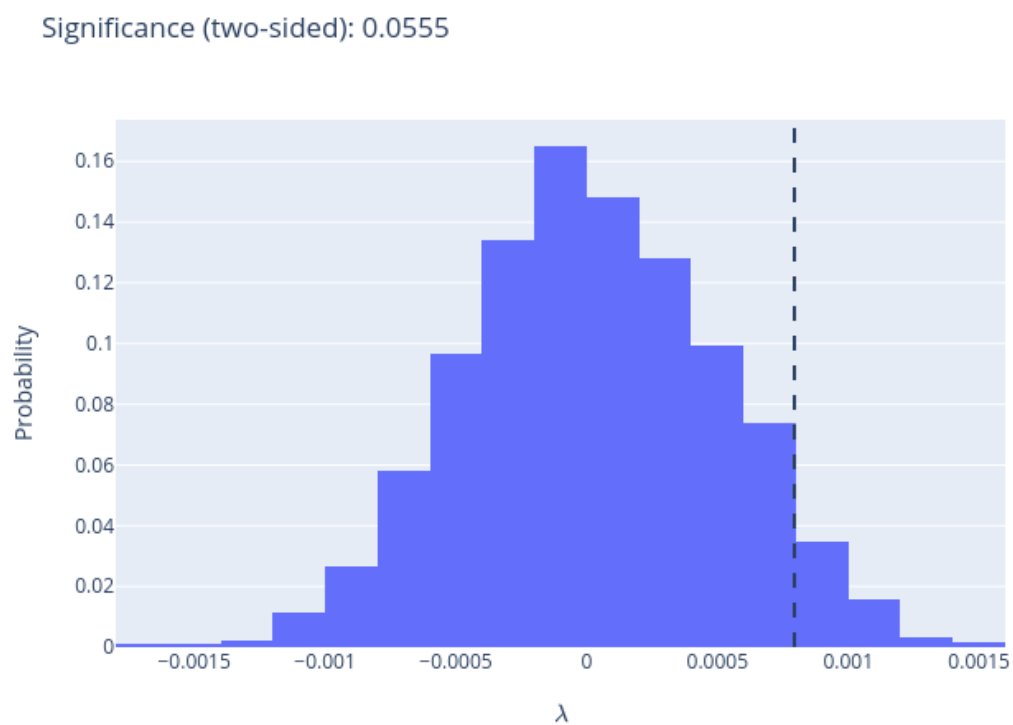
Figure 9: Results from a bootstrap hypothesis test for the significance of the difference of the λ parameters of the exponential models fitted to ISIs from `female1` and `female2`.

Precision: 0.73, Recall: 0.70, f1-score: 0.72

Figure 10: Confusion matrix corresponding to decodings using a naive Bayes classifier with the exponential model.

Female1: $\mu = 0.0502$ sec, $\lambda = 33705.90$ Hz, Female2: $\mu = 0.0523$ sec, $\lambda = 39286.30$ F
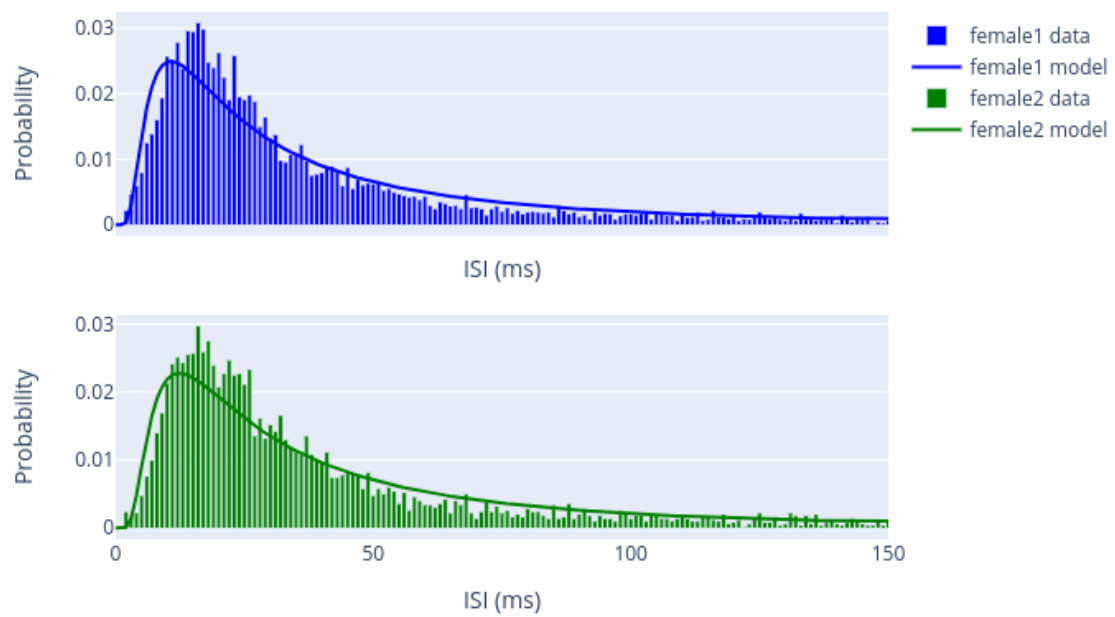
Figure 11: ISIs and their fits by an inverse Gaussian model. The title shows the estimated parameters for each model. Click on the figure to see its interactive version.
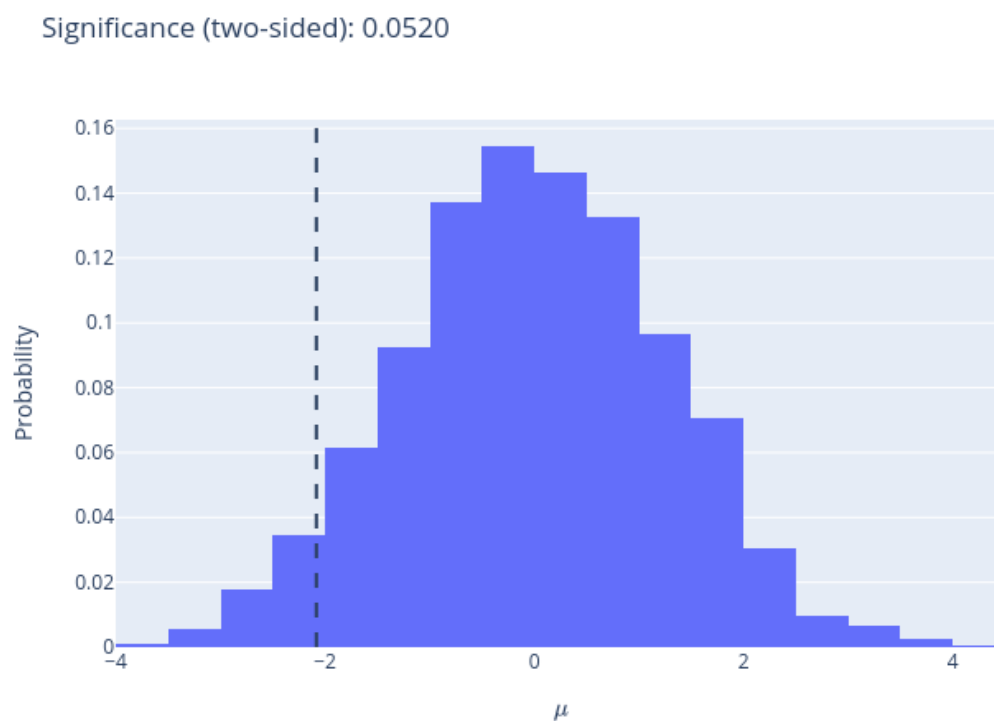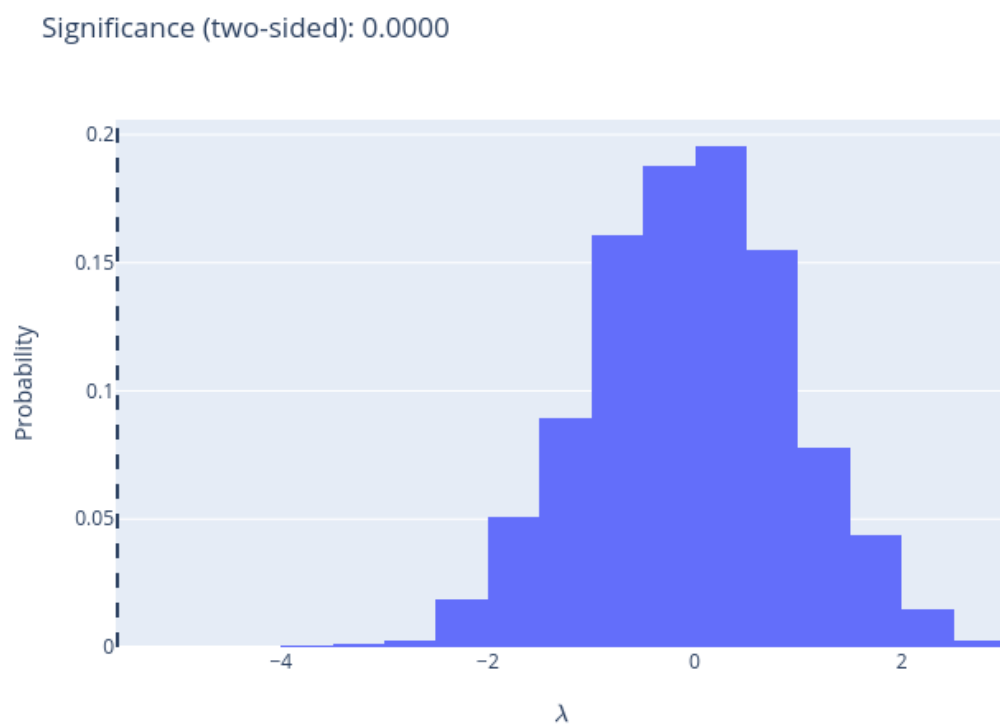
Figure 12: Results from a bootstrap hypothesis test for the significance of the difference of the μ parameters of the inverse Gaussian models fitted to ISIs from `female1` and `female2`.

Figure 13: Results from a bootstrap hypothesis test for the significance of the difference of the λ parameters of the inverse Gaussian models fitted to ISIs from `female1` and `female2`.

18

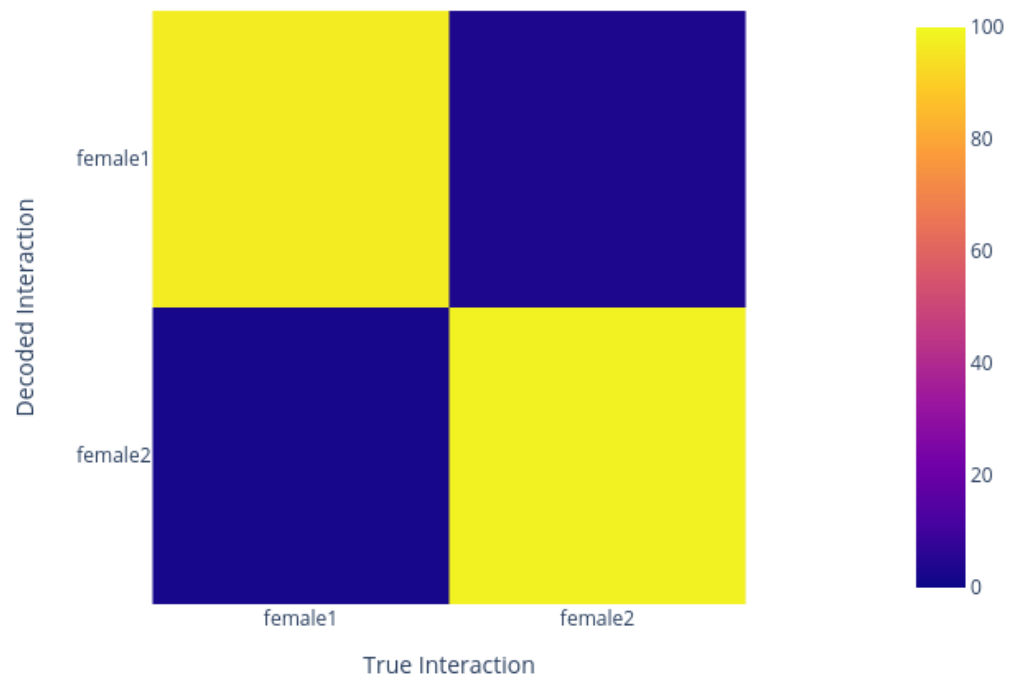Figure 14: Confusion matrix corresponding to decodings using a naive Bayes classifier with the inverse Gaussian model.
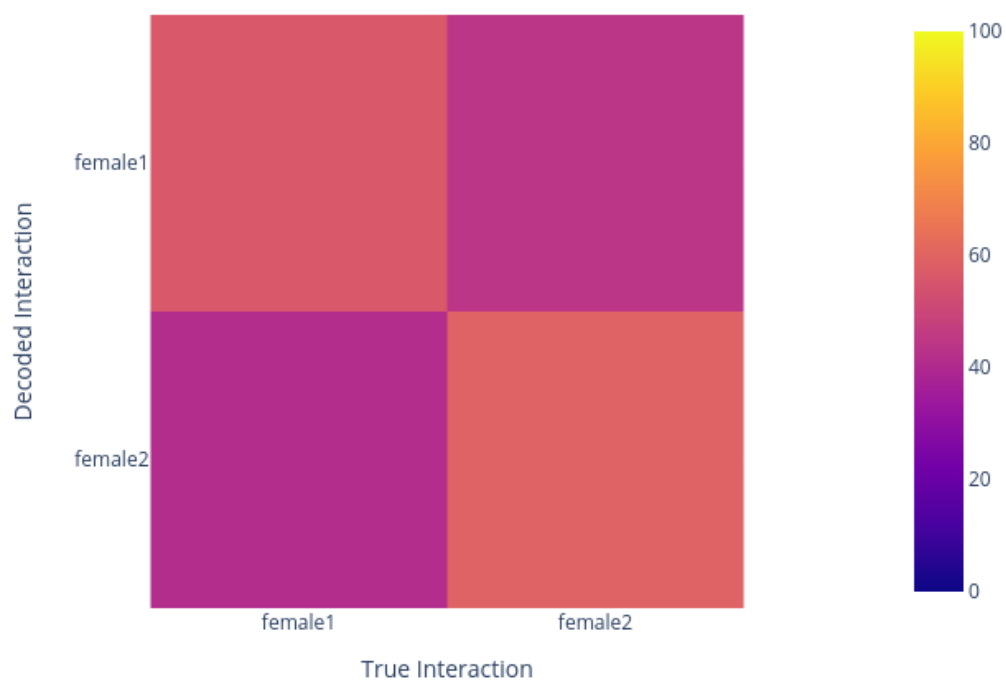
Figure 15: Confusion matrix corresponding to decodings using a naive Bayes classifier with the inverse Gaussian model and the control dataset.