**Model Details:**

**RL Model:**

**For simplicity and model accuracy, we fit the model to the concatenated session for each rat.**This method is better than fitting the model to each session since each session has few data and therefore prompts to produce a inaccurate estimate. We use MATLAB fmincon with multiple starting points to encourage global minimum of negative loglikelihood of free-choice trials. For the model, we have 2 models: model 1 assumes rewarded and unrewarded trials shares the same learning rate, thus symmetric learning, and model 2 assumes different learning rates for rewarded and unrewarded trials, thus allowing asymmetric learning such as loss aversion.

**Parameter:**

(learning rate), between 0 and 1. and **nr** for 2 rates model.

(inverse temperament), sensitivity to the difference in subjective reward values, > 0

**gamma**(memory decay rate or temporal discounting rate of action value function), between 0 and 1, 1 means no decay at all, 0 means total forget.

**b**(bias), negative means putting negative bias(averse) to sub-optimal choice, postive means giving extra value(prefer) to sub-optimal choice.

**Variables:**

**R(t)** = reward at trial t.

**Q(A)(t)** = action value function for action A at t.

**Action Value Function Update:**

**Model 1(4 parameters) - single learning rate:**

for all trial\_type:

Q(A)(t+1) =Q(A)(t) + \*(R(t) - Q(A)(t))

Q(~A)(t+1) = (1-gamma)\*Q(~A)

**Model 2(5 parameters) - 2 learning rate for rewarded and unrewarded trials:**

= rewarded trial learing rate, = unrewarded trial learing rate

for all trial\_type:

If R(t) == 1:

Q(A)(t+1) =Q(A)(t) + \*(R(t) - Q(A)(t))

Q(~A)(t+1) = (1-gamma)\*Q(~A)(t)

Else:

Q(A)(t+1) =Q(A)(t) + \*(R(t) - Q(A)(t))

Q(~A)(t+1) = (1-gamma)\*Q(~A)(t)

**For all action A1 and A2, the final output of A1 at trial t for example, is:**

P(A1) = Softmax(Q(A1)(t), Q(A2)(t),) =

Where P(A1) is the probability of choosing actions A1, given the action-value estimates Q(A1)(t) and Q(A2)(t) at time t.

Maximizing loglikelihood is same as minizing negative loglikelihood. Thus, the target function is given by minimizing mean **negative loglikelihood(NLL)** based on choice history. Forced trials are excluded **so we are only optimizing NLL with respect to free choice trials:**

**sf\_sub** = softmax prediction of informative choice probability conditioned on free choice

**choice2\_sub** = informative choice conditioned on free choice

**NLL** = -**Mean**(choice2\_sub .\* log(sf\_sub) + (1 - choice2\_sub) .\* log(1 - sf\_sub));

**Logistic Regression:**

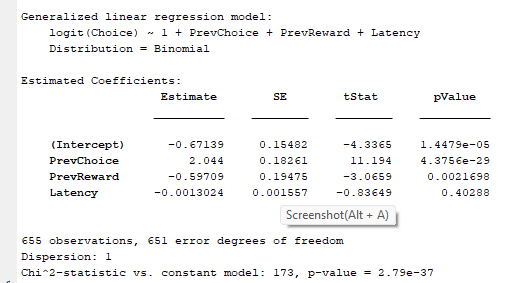
We fit logistic regression(a type of generalized linear model for binary classification) to each rat that either belongs to test(female ACC-CNO) or control conditions to predict the choice outcome, using the 3 features. We use free-choice data for the regression, and the previous choice and reward can be either free-choice or forced choice since they all contains information about a choice. For the regression, we use latency, previous choice, and previous reward as features, and extract the p-values of features to determine if a feature is a significant predictor.

Feature1: latency(continuous)

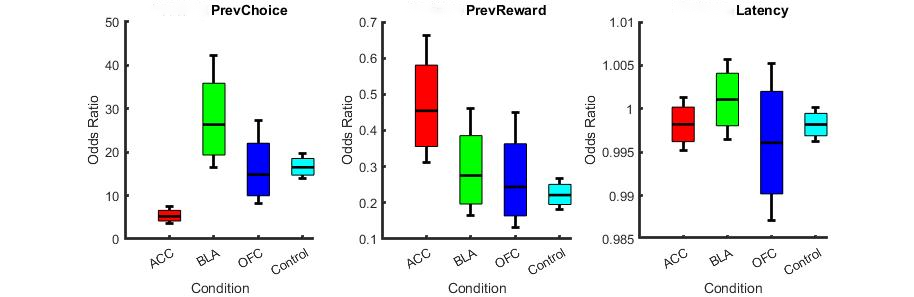
Feature2: previous choice - 0 or 1(binary)

Feature3: previous choice reward - yes or no(binary)

prediction: predicted probability of choice



Example logistic regression summary statistics of a single rat.



We compute the odds ratio and 95%,80% CI for each condition’s regression fit:

If an odds ratio is greater than 1, it represents an increase in the odds of the outcome happening given a one-unit increase in the predictor.

If the odds ratio is less than 1, it represents a decrease in the odds of the outcome happening given a one-unit increase in the predictor.

An odds ratio of exactly 1 signifies that the value of the predictor has no effect on the probability of the outcome.

The plots indicate test(ACC,OFC,BLA) and control groups respectively. The x-axis stands for each condition, and the y-axis is the odds ratio of the corresponding feature given condition. The box ,the error bar, and the red line represent 80% confidence interval, 95% confidence interval, and mean odds ratio, respectively. Mean model R^2 for control group = 0.4613, ACC group = 0.4189, OFC group = 0.4613, BLA group = 0.5181.

In conclusion, these findings reveal distinct patterns of behavioral responses across different inhibited brain areas when considering previous choices, rewards, and latency as predictors for next choice. The results suggest that previous choices have a substantial effect on subsequent behavior for all conditions, particularly under the BLA and Control conditions. Previous rewards, conversely, appear to reduce the odds of the event of interest across all conditions. Latency shows a negligible effect.

Complete odds ratio statistic:

Feature: PrevChoice

Condition: ACC

80% CI: [4.153163, 6.591847]

95% CI: [3.673560, 7.452448]

Condition: BLA

80% CI: [19.368697, 35.881787]

95% CI: [16.442696, 42.267001]

Condition: Control

80% CI: [14.718738, 18.535002]

95% CI: [13.844448, 19.705505]

Condition: OFC

80% CI: [9.973790, 22.057026]

95% CI: [8.077980, 27.233558]

Feature: PrevReward

Condition: ACC

80% CI: [0.355753, 0.580748]

95% CI: [0.312329, 0.661490]

Condition: BLA

80% CI: [0.196702, 0.385799]

95% CI: [0.164474, 0.461392]

Condition: Control

80% CI: [0.195356, 0.251202]

95% CI: [0.182735, 0.268551]

Condition: OFC

80% CI: [0.163945, 0.363321]

95% CI: [0.132709, 0.448838]

Feature: Latency

Condition: ACC

80% CI: [0.996224, 1.000198]

95% CI: [0.995171, 1.001256]

Condition: BLA

80% CI: [0.998048, 1.004093]

95% CI: [0.996449, 1.005704]

Condition: Control

80% CI: [0.996900, 0.999479]

95% CI: [0.996216, 1.000166]

Condition: OFC

80% CI: [0.990199, 1.002008]

95% CI: [0.987085, 1.005168]