Statistical analysis of AMPARO's data

Rafael Bassi Stern

Federal University of São Carlos

1. A description of the game

2. A model for learning

3. Building classifiers

A description of the game

The goalkeeper's game

The game was applied to patients with Parkinson's disease and a control group.

Steps

- Select a direction between: left, center or right.
- Algorithm determines correct direction.
- Iterate.
- Each stage uses a different algorithm.

The goalkeeper's game

Stages

- Warming-up
- Deterministic with hints
- Deterministic without hints
- ...
- Memory game

Each stage was designed to capture a specific type of information.

The goalkeeper's game

Initial goals

- Validate the information obtained from the game.
- Predict variables related to Parkinson's disease using the goalkeeper's game.

Challenge

- High dimensional features and small sample size.

A model for learning

The data

Data description

- i: the stage of the game, $i \in \{1, 2, 3, 4\}$.
- j: and id for each patient, $1 \leq j \leq 67$.
- k: a turn number.
- $X_{i,j,k}$: the indicator that patient j chose the right answer in the k-th turn of the i-th stage of the game, $X_{i,j,k} \in \{0, 1\}$.
- $Y_{i,j,k}$: the time spent by patient j in the k-th turn of the i-th stage of the game, $Y_{i,j,k} \in \mathbb{R}^+$.
- S_i: level of schooling of patient i.

The data

Challenge

- $n = 67$.
- Covariates: 8 time series per patient.
- S_i is very correlated with predicted variables.

A model for learning

Statistical model

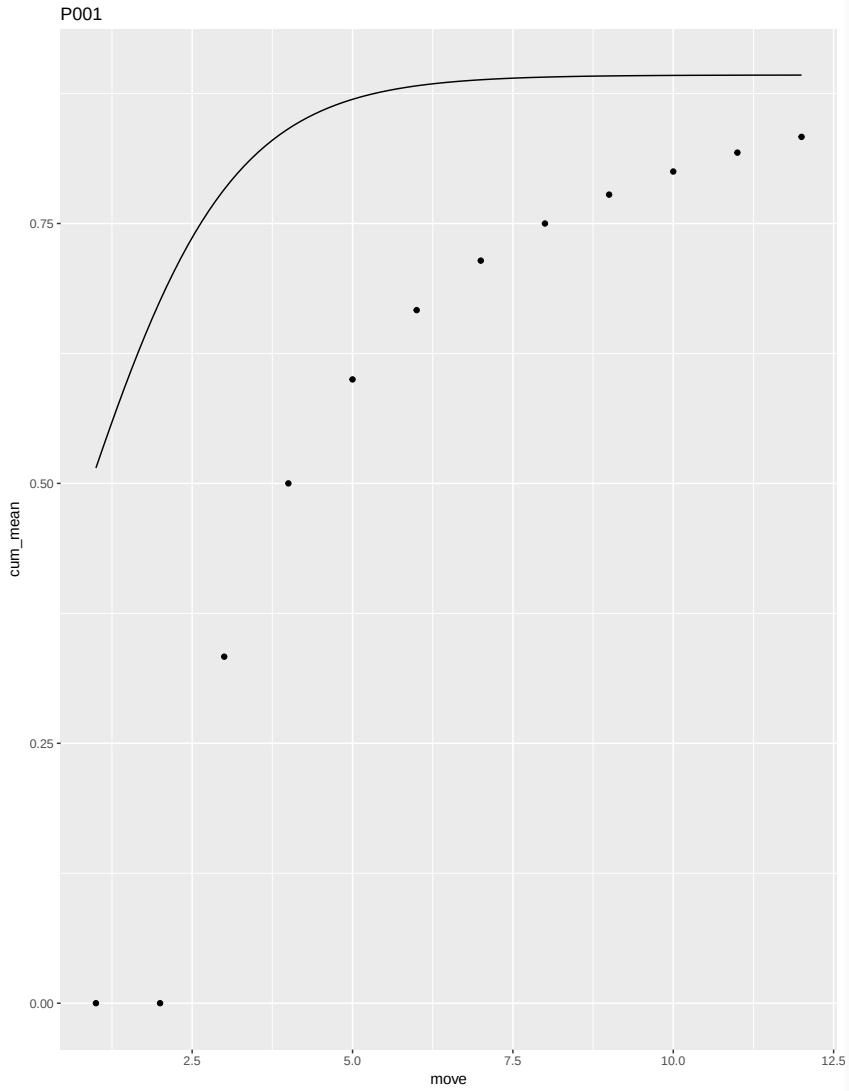
$$g(s) = \frac{\exp(s)}{1 + \exp(s)}$$

$$\mathbb{P}(X_{i,j,k} = 1) = \gamma_{i,j} \cdot g(-\log(3\gamma_{i,j}) + \alpha(S_i) + k \cdot \beta_{i,j})$$

Strategy

- Data might lie on a low dimensional manifold.
- Use estimated parameters as covariates.

A model for learning



Fitting the model

Challenge

- Obtain sparse parameter estimation.
- No analytical solution.

Solution

- Bayesian estimation with sparse priors.
- Posterior calculation via HMC (Stan).

Building classifiers

Response variables

Types

- MoCA: Montreal Cognitive Assessment.
- UPDRS III: Unified Parkinson's Disease Rating Scale.
- BEST: Balance Evaluation Scale Test.

Each variable type has several instances.

Logistic regression based on goalkeeper's game

response	baseline	accuracy	golden standard (moca)
updrs tot	0.54	0.66	0.72
updrs rig	0.52	0.65	0.73
best reat	0.52	0.66	0.72
best rest	0.5	0.66	0.72
moca evoc	0.5	0.69	0.8
best lim	0.57	0.75	0.75
best trans	0.5	0.7	0.74
moca vis	0.56	0.77	0.83
moca tot	0.52	0.72	-
best tot	0.5	0.72	0.72
best march	0.5	0.75	0.73

Referências