Modelling the feeding behaviour of a population of grey mouse lemurs via mixed hidden Markov models

Roland Langrock

Centre for Research into Ecological and Environmental Modelling

University of St Andrews

1. Data description

2. Modelling via HMMs
   - HMM for a single animal
   - HMM for multiple animals

3. Current research

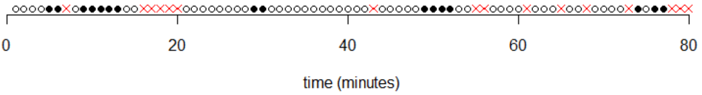
Grey mouse lemur

- largest mouse lemur (up to 70 grams)
- lives in Madagascar, mainly in high trees
- night-active
- food: fruits, leaves, insects, ...

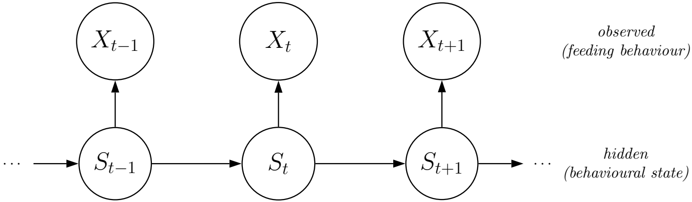
Data

- 16 female & 38 male lemurs
- feeding behaviour observed at 1-minute intervals
  - 0 (not feeding), 1 (feeding), NA (missing/out of sight)
  - time frame: 18 – 23h
  - on at most 7 different days
  - lots of missing data
- available covariates: body mass & gender

Sample:



Basic HMM structure

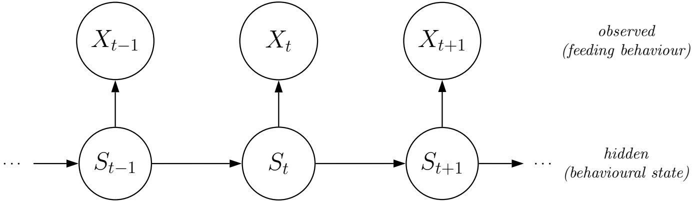


- X_t: **State-dependent process**
  $\longrightarrow$ 0 (not feeding) or 1 (feeding)

- S_t: **Markov chain**
  $\longrightarrow$ 0/"sated" or 1/"hungry"

$\longrightarrow$ *special case of a state-space model (with finite state space)*

Basic HMM structure

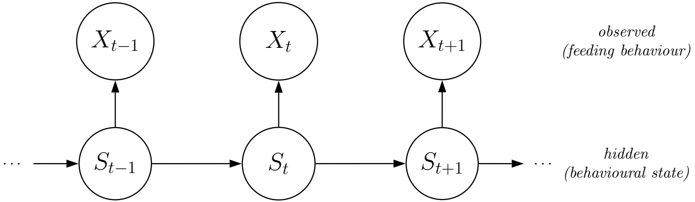


- X_t: **State-dependent process**
  $\rightarrow$ 0 (not feeding) or 1 (feeding)

- S_t: **Markov chain**
  $\rightarrow$ 0/ "sated" or 1/ "hungry"

$\rightarrow$ *special case of a state-space model (with finite state space)*

Basic HMM structure



- X_t: **State-dependent process**
  $\longrightarrow$ 0 (not feeding) or 1 (feeding)
- S_t: **Markov chain**
  $\longrightarrow$ 0/ "sated" or 1/ "hungry"

$\longrightarrow$ *special case of a state-space model (with finite state space)*

1. Data description

2. Modelling via HMMs
   - HMM for a single animal
   - HMM for multiple animals

3. Current research

Model for succession of behavioural states

Transition probabilities are modelled **time-dependently**:

$$\text{logit}(\gamma_{ii}^{(t)}) = \beta_{0,i} + \beta_{1,i} t, \quad i = 1, 2, \quad t = 0, 1, 2, \ldots$$

$$\text{where } \gamma_{ij}^{(t)} = \mathbb{P}\big(S_{t+1} = j \,|\, S_t = i\big)$$

Transition probability matrix (at time t):

$$\Gamma^{(t)} = \begin{pmatrix} \gamma_{11}^{(t)} & \gamma_{12}^{(t)} \\ \gamma_{21}^{(t)} & \gamma_{22}^{(t)} \end{pmatrix}$$

Initial distribution:

$$\delta = \big(\mathbb{P}(S_t = 1), (\mathbb{P}(S_t = 2)\big)$$

Model for succession of behavioural states

Transition probabilities are modelled **time-dependently**:

$$\text{logit}(\gamma_{ii}^{(t)}) = \beta_{0,i} + \beta_{1,i}t, \quad i = 1, 2, \quad t = 0, 1, 2, \ldots$$

$$\text{where } \gamma_{ij}^{(t)} = \mathbb{P}\big(S_{t+1} = j \,|\, S_t = i\big)$$

Transition probability matrix (at time t):

$$\mathbf{\Gamma}^{(t)} = \begin{pmatrix} \gamma_{11}^{(t)} & \gamma_{12}^{(t)} \\ \gamma_{21}^{(t)} & \gamma_{22}^{(t)} \end{pmatrix}$$

Initial distribution:

$$\boldsymbol{\delta} = \big(\mathbb{P}(S_t = 1), (\mathbb{P}(S_t = 2)\big)$$

Observation model

State-dependent (Bernoulli) distributions:

$$\mathbb{P}\big(X_t = 1 \,|\, S_t = k\big) = \pi_k$$
$$\mathbb{P}\big(X_t = 0 \,|\, S_t = k\big) = 1 - \pi_k$$

"hungry" state: π_k relatively large (typically ≈ 0.9)
"sated" state: π_k close to 0

Likelihood

$$\mathcal{L} = \boldsymbol{\delta}\mathbf{P}_{\boldsymbol{\pi}}(x_1)\boldsymbol{\Gamma}^{(1)}\mathbf{P}_{\boldsymbol{\pi}}(x_2)\boldsymbol{\Gamma}^{(2)} \cdot \ldots \cdot \mathbf{P}_{\boldsymbol{\pi}}(x_{T-1})\boldsymbol{\Gamma}^{(T-1)}\mathbf{P}_{\boldsymbol{\pi}}(x_T)\mathbf{1}^t$$

where $\mathbf{P}_{\boldsymbol{\pi}}(x_t) = \text{diag}\big(\pi_1^{x_t}(1-\pi_1)^{1-x_t}, \pi_2^{x_t}(1-\pi_2)^{1-x_t}\big)$
(for missing observations: 2×2-identity matrix)

1. Data description

2. Modelling via HMMs
   - HMM for a single animal
   - HMM for multiple animals

3. Current research

Strategies for dealing with populations in the HMM framework

- parameter set common to all subjects
  $\longrightarrow$ neglects any possible heterogeneity across subjects
- each of the parameters subject-specific
  $\longrightarrow$ enormous number of parameters
  $\longrightarrow$ impedes the possibility of comparing the resulting models
- some parameters common to all subjects, others subject-spec.
  $\longrightarrow$ special case: random effects
  $\longrightarrow$ computationally demanding
- subject-specific covariates
  $\longrightarrow$ requires that suitable covariates are available

Model for multiple animals

Transition probabilities:

$$\text{logit}(\gamma_{ii}^{(t,m)}) = \beta_{0,i} + \beta_{1,i} t + \beta_{2,i} \text{sex}^{(m)} + \beta_{3,i} \text{mass}^{(m)}$$

State-dependent feeding probabilities:

$$\pi_{i,m} \overset{iid}{\sim} \mathcal{B}eta(a_i, b_i), \qquad m = 1, \ldots, M \, (= 54)$$

Likelihood:

$$\mathcal{L} = \prod_{m=1}^{M} \int_0^1 \int_0^1 \boldsymbol{\delta} \mathbf{P}_{\boldsymbol{\pi}}(x_{1,m}) \boldsymbol{\Gamma}^{(1,m)} \mathbf{P}_{\boldsymbol{\pi}}(x_{2,m}) \boldsymbol{\Gamma}^{(2,m)} \cdot \ldots$$
$$\ldots \cdot \boldsymbol{\Gamma}^{(T-1,m)} \mathbf{P}_{\boldsymbol{\pi}}(x_{T_m,m}) \mathbf{1}^t f_1(\pi_1) f_2(\pi_2) d\pi_1 d\pi_2$$

Estimation

- Monte Carlo expectation–maximization
  $\longrightarrow$ E-step requires simulation methods
  $\longrightarrow$ M-step might require numerical optimization
- Numerical likelihood maximization
  $\longrightarrow$ integrals must be approximated numerically

Both methods are computationally demanding!

Results (part I)

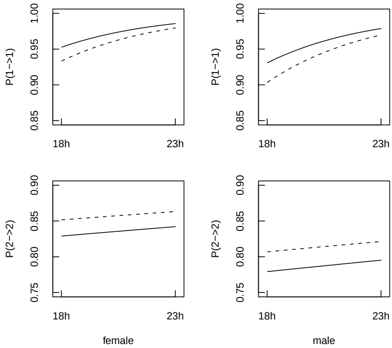


Figure: persistence in behavioural states as function of t
upper plots refer to 'sated' state, lower plots to 'hungry' state
left plots: female lemurs, right plots: male lemurs
solid lines: mass = 59 gr, dashed lines: mass = 33 gr)

Results (part II)

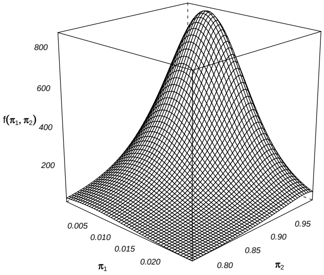


Figure: Fitted joint probability density function of π_1 and π_2.

- notice the scales $\longrightarrow$ much more heterogeneity in π_2, the probability of eating when in the "hungry" state
- doesn't take into account possible correlation between π_i's

Related research

- other types of time series
  - → continuous and/or multivariate data
  - → in particular animal movement paths