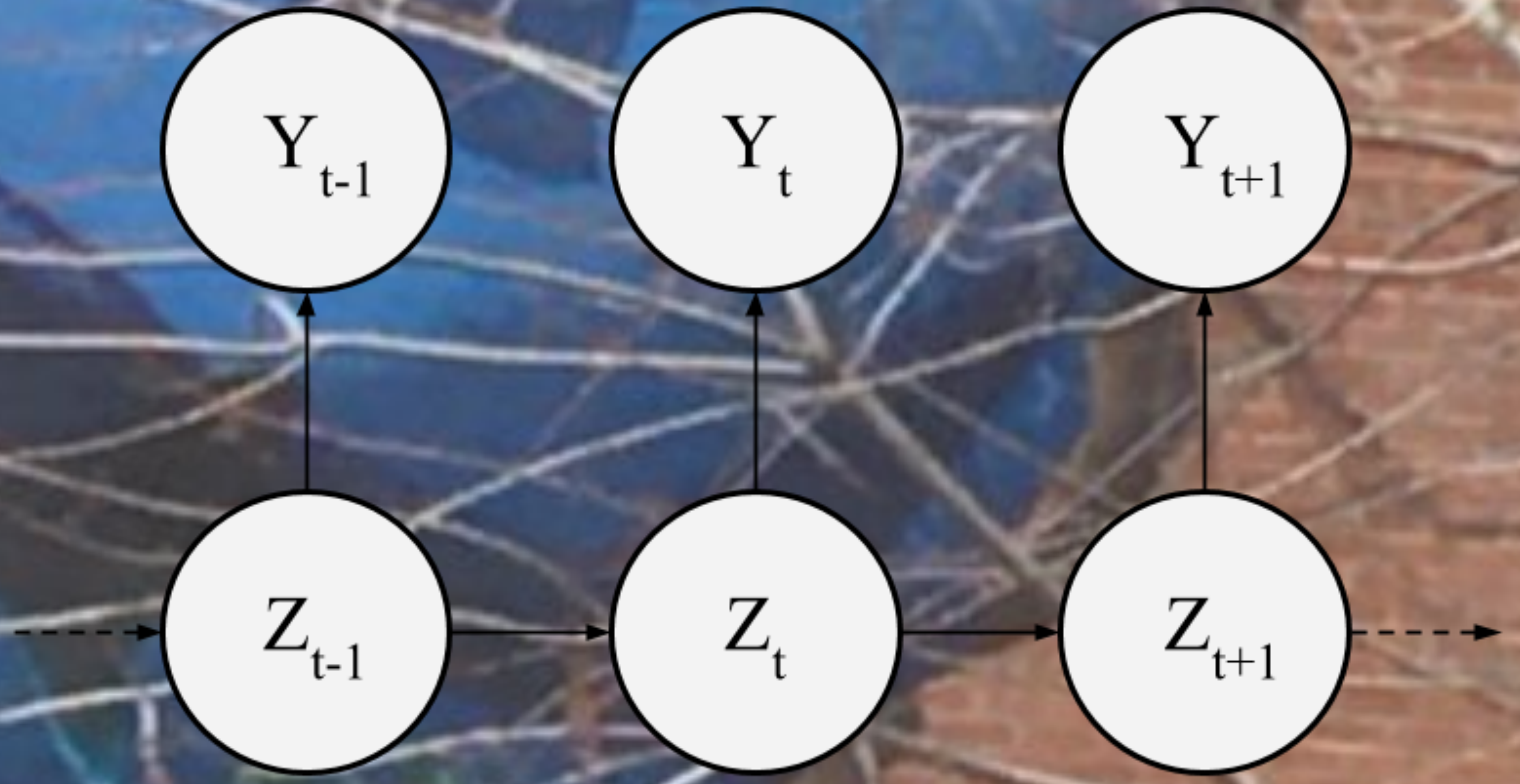
HIDDEN MARKOV MODELS IN ECOLOGY



VIANEY LEOS BARAJAS
STATISTICAL SCIENCES/SCHOOL OF THE ENVIRONMENT
UNIVERSITY OF TORONTO

INTRODUCTION

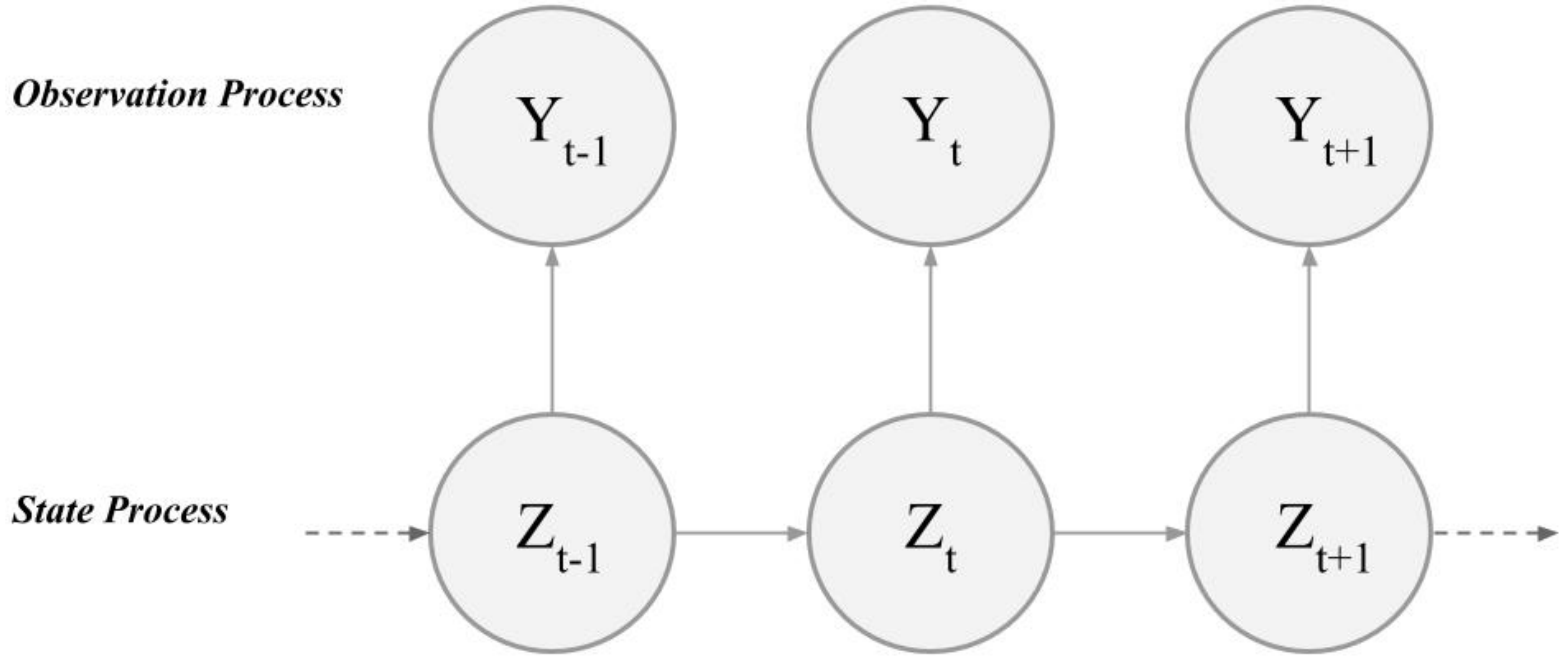
ECOLOGICAL DATA

Animal movement

Occupancy

Capture-recapture

HIDDEN MARKOV MODELS



HIDDEN MARKOV MODELS (HMMS)

- A doubly stochastic process composed of an **observation process,** $\{Y_t\}_{t=1}^{T}$**,** and an underlying **latent state process,** $\{Z_t\}_{t=1}^{T}$**, (taken to be a Markov chain)**

HIDDEN MARKOV MODELS (HMMS)

- A doubly stochastic process composed of an **observation process,** $\{Y_t\}_{t=1}^{T}$**,** and an underlying **latent state process,** $\{Z_t\}_{t=1}^{T}$**, (taken to be a Markov chain)**
- Y_t can be discrete, continuous, measured with error, multivariate, and can also be an AR(p) process, and so on…(with AR(p) we typically call them Markov-switching models)

HIDDEN MARKOV MODELS (HMMS)

- A doubly stochastic process composed of an **observation process,** $\{Y_t\}_{t=1}^{T}$**,** and an underlying **latent state process,** $\{Z_t\}_{t=1}^{T}$**, (taken to be a Markov chain)**
- Y_t can be discrete, continuous, measured with error, multivariate, and can also be an AR(p) process, and so on…(with AR(p) we typically call them Markov-switching models)
- Z_t is a discrete random variable, that takes on a value in $\{1, \ldots, N\}$ at each time t, where N is taken to represent the **'number of states'**

ESSENCE OF A (DISCRETE-TIME FINITE-STATE) HMM

MATHEMATICAL DESCRIPTION:

- Number of states, N

GENERAL INTERPRETATION:

- Number of processes of interest

ESSENCE OF A (DISCRETE-TIME FINITE-STATE) HMM

MATHEMATICAL DESCRIPTION:

- Number of states, N
- State-dependent distributions: $\{f(Y_t \mid Z_t = n\}_{n=1}^{N} = \{f_n(Y_t)\}_{n=1}^{N}$

GENERAL INTERPRETATION:

- Number of processes of interest
- How the data are generated in each state

ESSENCE OF A (DISCRETE-TIME FINITE-STATE) HMM

MATHEMATICAL DESCRIPTION:

- Number of states, N
- State-dependent distributions: $\{f(Y_t \mid Z_t = n\}_{n=1}^{N} = \{f_n(Y_t)\}_{n=1}^{N}$
- Transition probability matrix (t.p.m.), $\boldsymbol{\Gamma}$, with entries $\Gamma_{ij} = \Pr(Z_t = j \mid Z_{t-1} = i)$, for $i, j \in \{1, \ldots, N\}$

GENERAL INTERPRETATION:

- Number of processes of interest
- How the data are generated in each state
- State-switching dynamics across processes

ESSENCE OF A (DISCRETE-TIME FINITE-STATE) HMM

MATHEMATICAL DESCRIPTION:

- Number of states, N
- State-dependent distributions: $\{f(Y_t \mid Z_t = n\}_{n=1}^{N} = \{f_n(Y_t)\}_{n=1}^{N}$
- Transition probability matrix (t.p.m.), $\boldsymbol{\Gamma}$, with entries $\Gamma_{ij} = \Pr(Z_t = j \mid Z_{t-1} = i)$, for $i, j \in \{1, \ldots, N\}$
- Initial state distribution, $\boldsymbol{\delta}$, with entries $\delta_i = \Pr(Z_1 = i)$ for $i \in \{1, \ldots, N\}$

GENERAL INTERPRETATION:

- Number of processes of interest
- How the data are generated in each state
- State-switching dynamics across processes
- What is happening at the first observation — what is the first process we observe?

STATE-DECODING ALGORITHMS FOR HMMS: CONNECTING OBSERVATIONS TO LATENT STATES

- **Viterbi:** computes the most probable sequence of states that generated the data by finding the sequence of states $\mathbf{Z}^*$ that maximizes: $Pr(\mathbf{Z} \mid \mathbf{Y})$
- **Forward-backward algorithm:** computes the marginal probability of the state at time t: $Pr(Z_t \mid \mathbf{Y})$
- **Forward-filtering backward-sampling algorithm:** sample from the joint distribution of the state sequences — sample sequences that could have generated the data: $Pr(\mathbf{Z} \mid \mathbf{Y})$

SIMULATING FROM A 3-STATE HMM

PARAMETER VALUES:

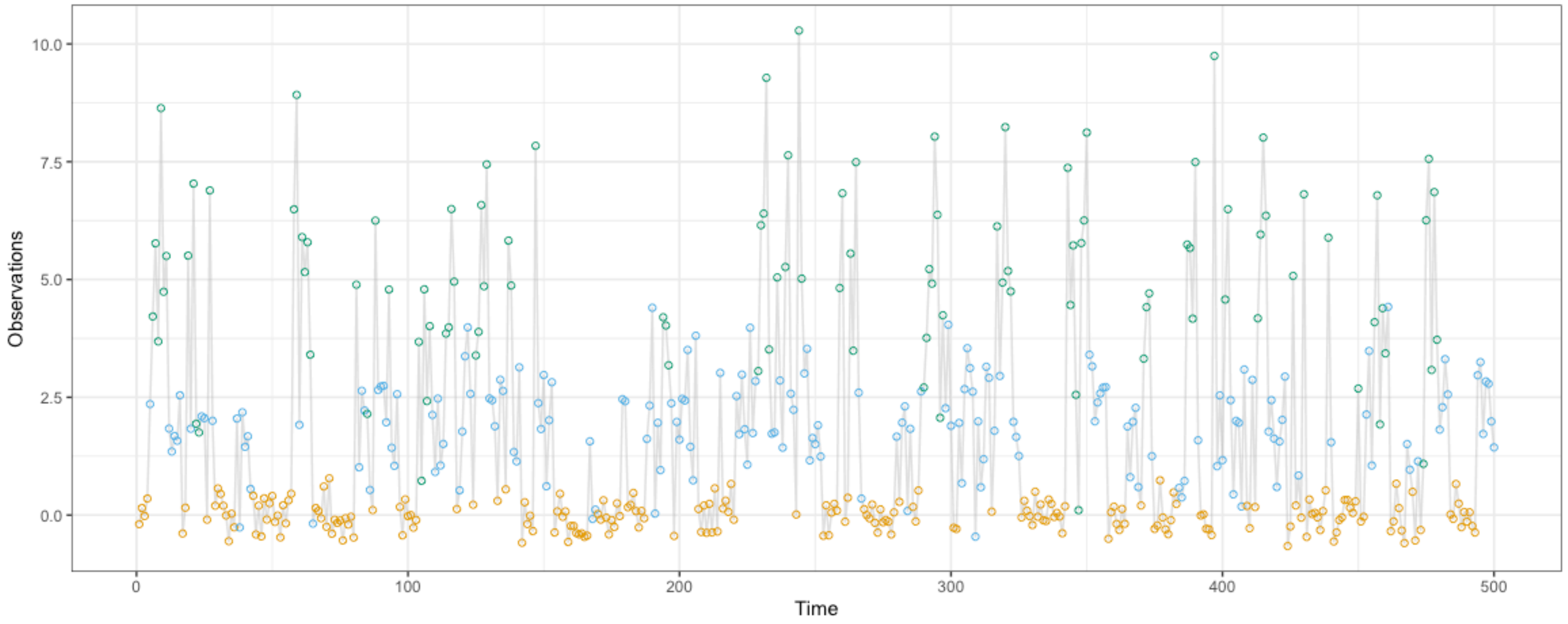
$$N = 3$$

$$f_n(Y_t) \sim N(\mu_n, \sigma_n) \quad \textbf{for} \quad n \in \{1,2,3\}$$

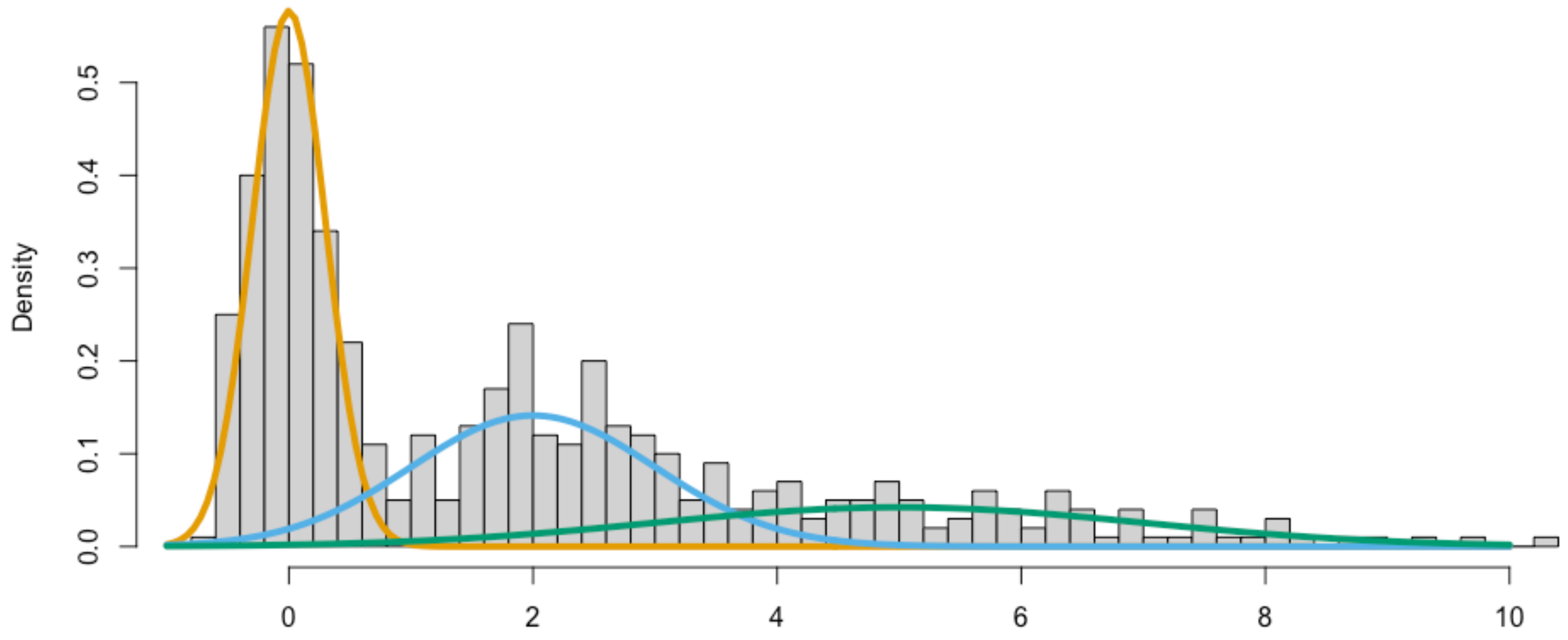
$$\mu \in \{0,2,5\} \qquad \sigma \in \{0.3,1,2\}$$

$$\boldsymbol{\delta} = [1/3,1/3,1/3]$$

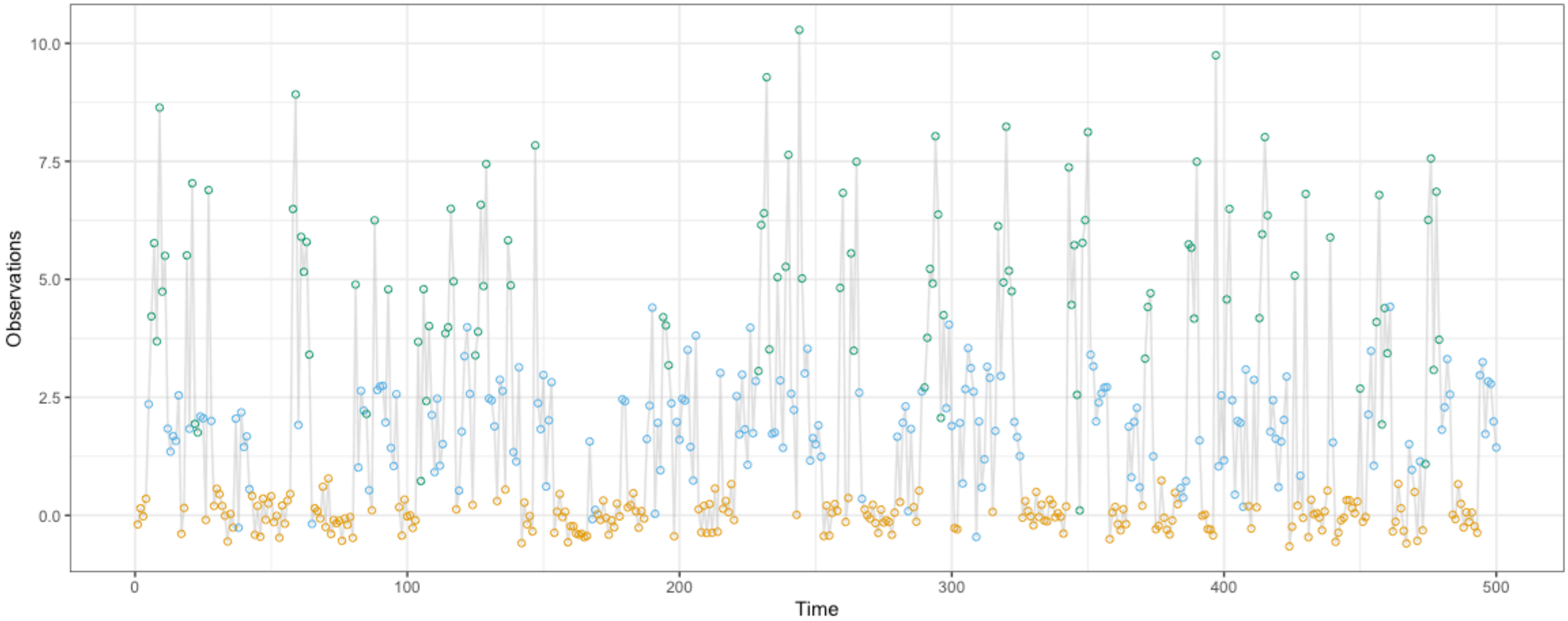
$$\boldsymbol{\Gamma} = \begin{pmatrix} 0.8 & 0.1 & 0.1 \\ 0.2 & 0.7 & 0.1 \\ 0.05 & 0.3 & 0.65 \end{pmatrix}$$



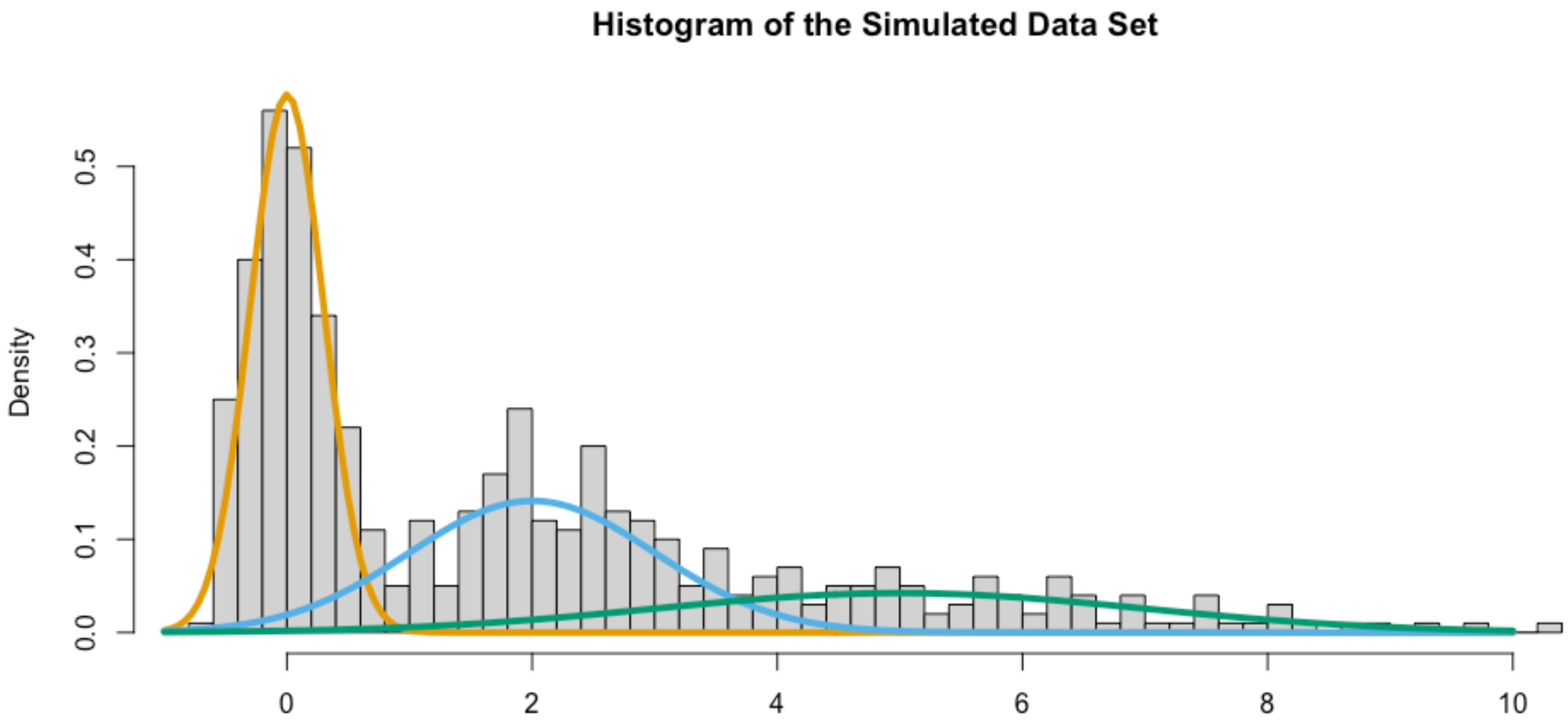
Histogram of the Simulated Data Set



SIMULATED DATA: FIRST 500 OF 2000 OBSERVATIONS



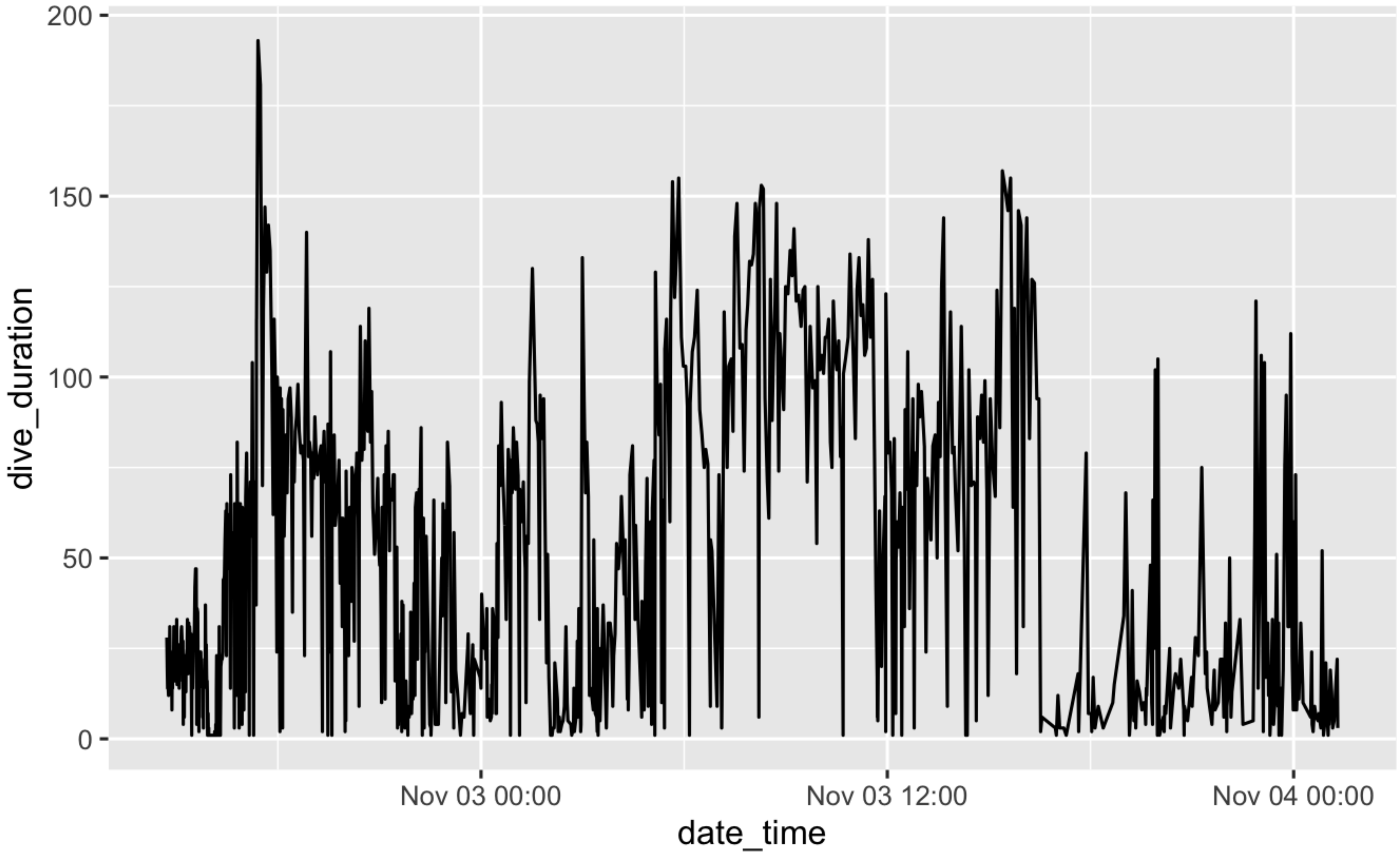
SIMULATED DATA: MARGINAL DISTRIBUTION



ANIMAL MOVEMENT

MODELING PORPOISE DIVE DATA

Data from *Multi-scale modeling of animal movement and general behavior data using hidden Markov models with hierarchical structures*, particularly from Floris M. Van Beest and Jacob Nabe-Nielsen



SETUP

What we need to define:

- **number of states,** typically starting from $N = 2$ and increasing by one in subsequent iterations/model fits
- **state-dependent distributions:**
    * for dive duration, we can see that the data is integer-valued so we can start with a Poisson distribution (but can be modified)
- **the transition probability matrix structure:** e.g. for a 2-state HMM

$$\begin{pmatrix} \Gamma_{1,1} & \Gamma_{1,2} \\ \Gamma_{2,1} & \Gamma_{2,2} \end{pmatrix}$$

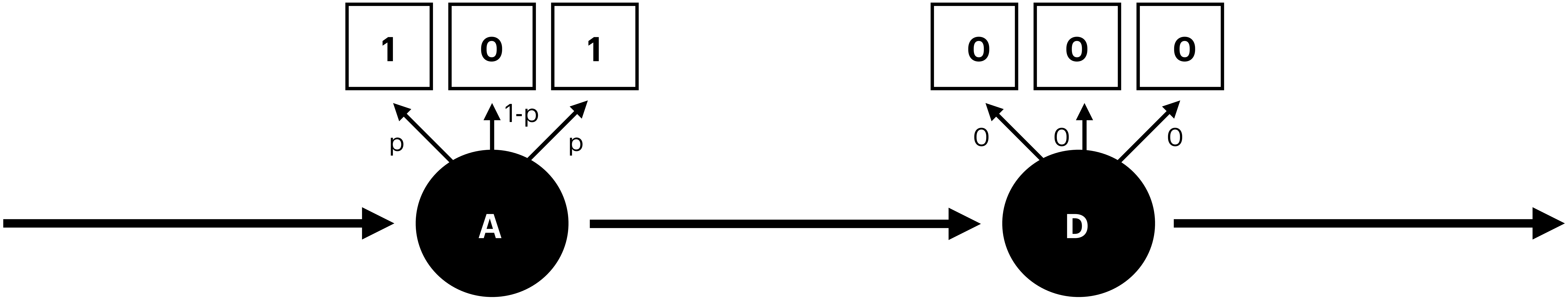
- **the initial state distribution:** e.g. for a 2-state HMM — $\boldsymbol{\delta} = (\delta_1 \quad \delta_2)$

CAPTURE-RECAPTURE

MODELING CAPTURE EVENTS

We can model capture-recapture events via an HMM framework where:

- **state process:** an animal is alive or dead
- **observation process:** seen or not seen



where p is the detection probability

SETUP

What we need to define:

- **number of states,** here we'll have two: **alive** + **dead**
- **state-dependent distributions:**

$$\begin{array}{cc} & \begin{array}{cc} 0 & 1 \end{array} \\ \begin{array}{c} A \\ D \end{array} & \begin{pmatrix} 1-p & p \\ 1 & 0 \end{pmatrix} \end{array}$$

- **the transition probability matrix structure:** switch from alive to dead to estimate survival probability, S

$$\begin{array}{cc} & \begin{array}{cc} A & D \end{array} \\ \begin{array}{c} A \\ D \end{array} & \begin{pmatrix} S & 1-S \\ 0 & 1 \end{pmatrix} \end{array}$$

- **the initial state distribution:** e.g. for a 2-state HMM — $\begin{array}{cc} A & D \\ (1 & 0) \end{array}$

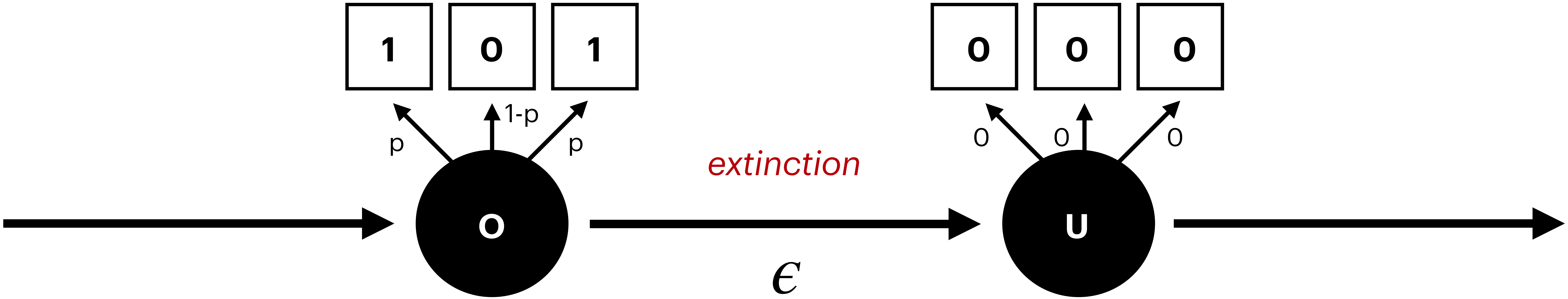
OCCUPANCY

DYNAMIC OCCUPANCY MODELS

We can model occupancy via an HMM framework where:

- **state process:** occupied vs unoccupied
- **observation process:** detected vs not detected

A key criteria is that a **species can occupy a space yet not be detected.**



SETUP

What we need to define:

- **number of states,** here we'll have two: **occupied** + **unoccupied**
- **state-dependent distributions:**

$$\begin{array}{cc} & \begin{array}{cc} 0 & 1 \end{array} \\ \begin{array}{c} U \\ O \end{array} & \begin{pmatrix} 1 & 0 \\ 1-p & p \end{pmatrix} \end{array}$$

- **the transition probability matrix structure:** switch from occupied to unoccupied termed as '**extinction**' and unoccupied to occupied as '**colonization**'

$$\begin{array}{cc} & \begin{array}{cc} U & O \end{array} \\ \begin{array}{c} U \\ O \end{array} & \begin{pmatrix} 1-\gamma & \gamma \\ \epsilon & 1-\epsilon \end{pmatrix} \end{array}$$

- **the initial state distribution:** e.g. for a 2-state HMM — $\begin{array}{cc} U & O \\ (1-\psi & \psi) \end{array}$

HMM NUANCES

DO WE HAVE STATE LABELS?

For animal movement:

- incorporation of state (behavioral) labels has not been thoroughly studied in ecology but mathematically, we can assume that some (or all) of the state sequence is observed

For capture-recapture:

- observing an animal tells us that the animal is still alive (the state is known!), but if we do not see an animal, there is uncertainty as to its true state (alive or dead)

For occupancy:

- as with capture-recapture, we have the state if the species is detected, but not if it is undetected

HOW DO WE EVALUATE THE MODEL?

'Unsupervised vs Supervised' —> can be confusing

'Do we care about explaining the data generating process OR do we care about predicting the states correctly?'

** Misspecified models might predict the states just as well as correctly specified models

** When there are no state labels available, we are generally building an HMM that captures the data generating process (assessed via residuals, model selection, simulation), with no guarantees that the states correspond to meaningful processes (if we can't validate our states — we can't assess how well the model captures those behaviors)

BAYESIAN HMMS

SETTING PRIORS

Setting priors distinguishes a Bayesian framework from a frequentist framework and is necessary to quantify the uncertainty of our parameters in a probabilistic manner.

What do we need from an HMM for it to work:

- For the **transition probability matrix** to be: **full rank** + **ergodic** *(in a process that has no specific end)*

- For the **state-dependent distributions** to be **distinct**

TIME TO FIT SOME BAYESIAN HIDDEN MARKOV MODELS

Contact: vianey.leosbarajas@utoronto.ca

REFERENCES

- Kéry, M., Guillera-Arroita, G., Lahoz-Monfort, J.J. Analysing and mapping species range dynamics using occupancy models. *Journal of Biogeography,* 1463-1474 (2013)

- Leos-Barajas, V., Gangloff, E.J., Adam, T. *et al*. Multi-scale Modeling of Animal Movement and General Behavior Data Using Hidden Markov Models with Hierarchical Structures. *JABES* 22, 232–248 (2017)

- Laake, J. L. 2013. Capture-recapture analysis with hidden Markov models. AFSC Processed Rep. 2013-04, 34 p. Alaska Fish. Sci. Cent., NOAA, Natl. Mar. Fish. Serv., 7600 Sand Point Way NE, Seattle WA 98115.

- McClintock, B. T., R. Langrock, O. Gimenez, E. Cam, D. L. Borchers, and R. Glennie. "Uncovering ecological state dynamics with hidden Markov models (vol 23, pg 1878, 2020)." *ECOLOGY LETTERS* 24, no. 5 (2021): 1117-1117.

- Zucchini, W., I. L. MacDonald, and R. Langrock. "Hidden Markov models for time series: an introduction using R." (2017).

- Stan Development Team. 2022. Stan Modeling Language Users Guide and Reference Manual, VERSION. https://mc-stan.org