Bayesian model selection for Markov chains using sparse probability vectors

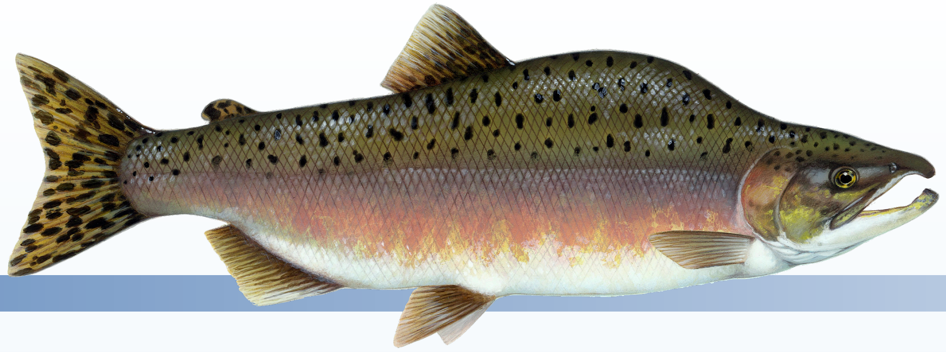
Matthew Heiner[1], Athanasios Kottas[1], and Stephan Munch[2]

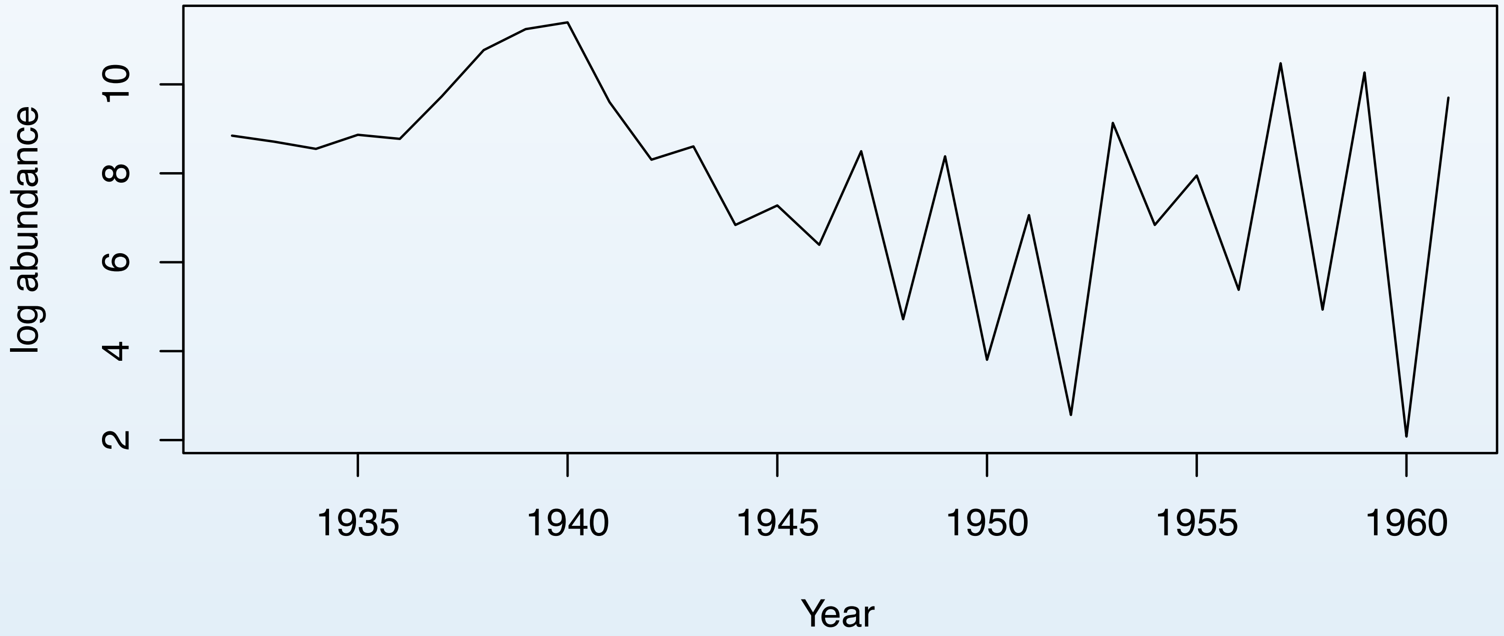
1. Department of Applied Mathematics and Statistics, University of California, Santa Cruz, California, USA
2. Fisheries Ecology Division, Southwest Fisheries Science Center, National Marine Fisheries Service, NOAA, Santa Cruz, California, USA

UCSC

Baskin Engineering UC SANTA CRUZ

Pink Salmon time series

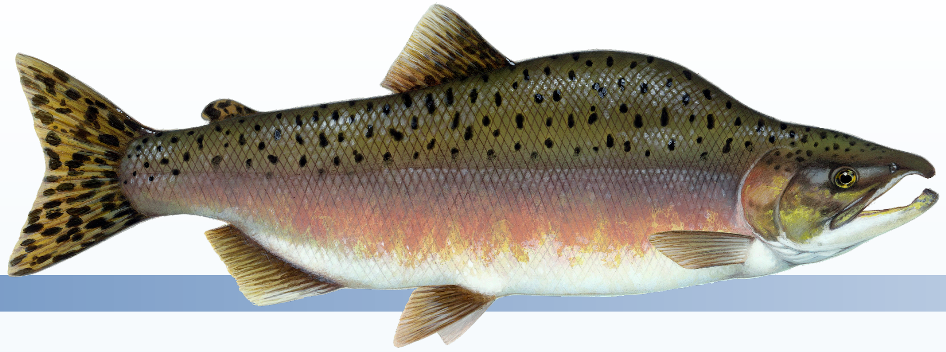


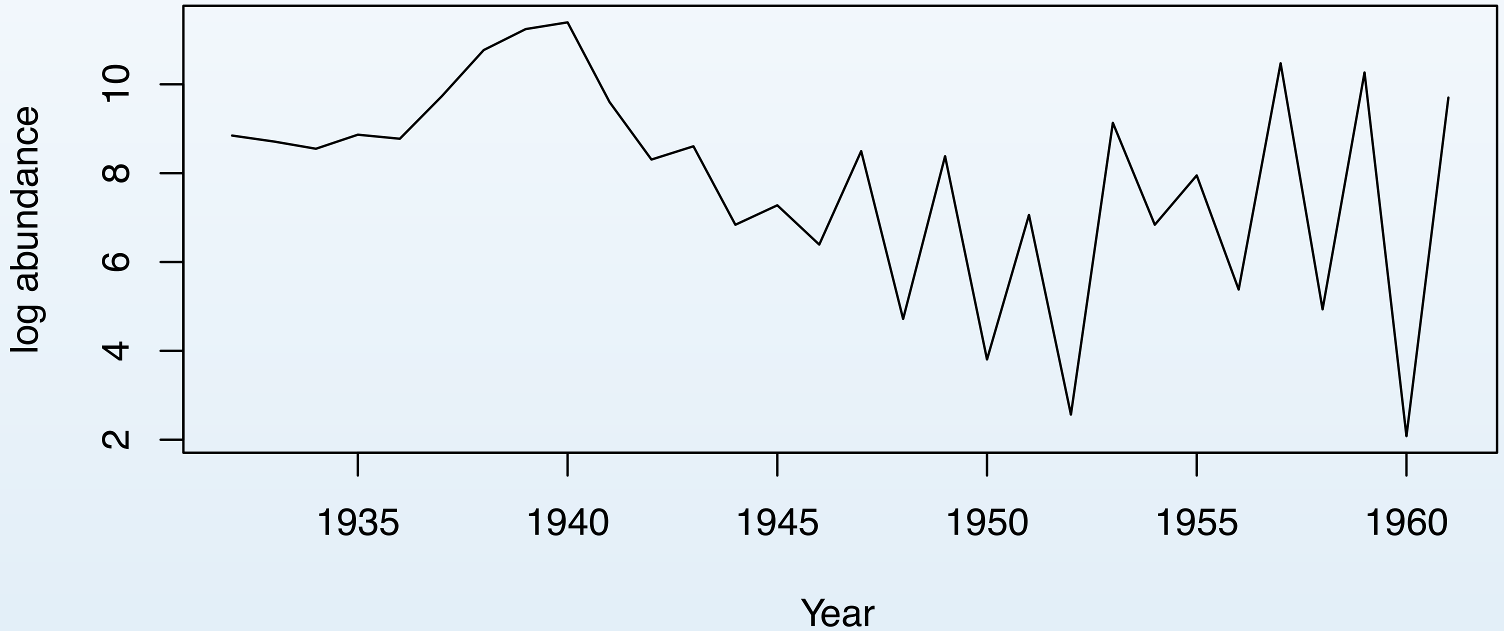


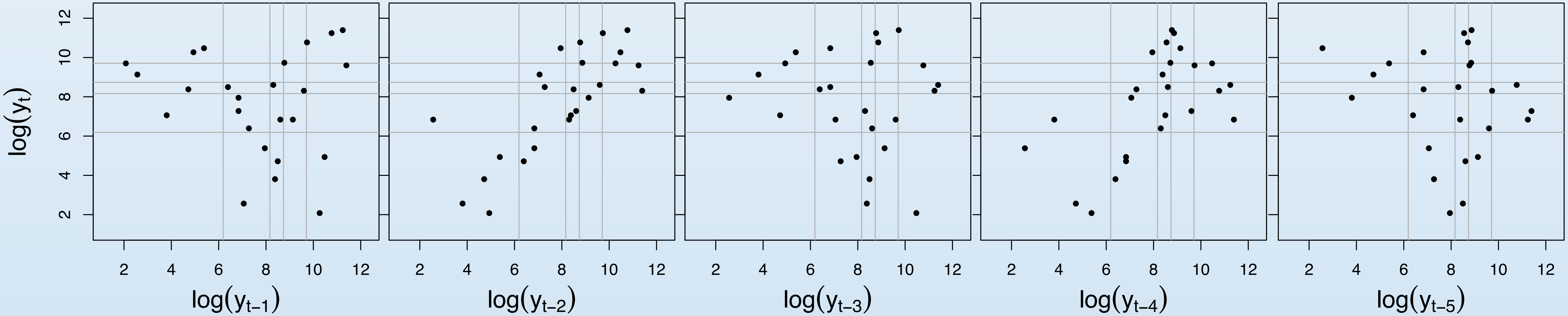
Data: Alaska Fisheries Science Center (2018), "AFSC/ABL: Pink salmon data collected at Sashin Creek Weir 1934-2002," NOAA National Centers for Environmental Information, https://inport.nmfs.noaa.gov/ inport/item/17256.

Image: U.S. Fish and Wildlife Service

Pink Salmon time series







Data: Alaska Fisheries Science Center (2018), “AFSC/ABL: Pink salmon data collected at Sashin Creek Weir 1934-2002,” NOAA National Centers for Environmental Information, https://inport.nmfs.noaa.gov/ inport/item/17256.

Image: U.S. Fish and Wildlife Service

Mixture transition distribution

Raftery (1985)

Mixture transition distribution

Raftery (1985)

Example with two states and three lags:

Mixture transition distribution

Raftery (1985)

Example with two states and three lags:

$$Q = \begin{pmatrix} q_{1,1} & q_{1,2} \\ q_{2,1} & q_{2,2} \end{pmatrix}$$

first order
transition matrix

Mixture transition distribution

Raftery (1985)

Example with two states and three lags:

$$Q = \begin{pmatrix} q_{1,1} & q_{1,2} \\ q_{2,1} & q_{2,2} \end{pmatrix}$$

first order
transition matrix

$$\Pr(s_t = 2 \mid s_{t-1} = 1, s_{t-2} = 1, s_{t-3} = 2)$$

Mixture transition distribution

Raftery (1985)

Example with two states and three lags:

$$Q = \begin{pmatrix} q_{1,1} & q_{1,2} \\ q_{2,1} & q_{2,2} \end{pmatrix}$$

first order transition matrix

$$\Pr(s_t = 2 \mid s_{t-1} = 1, s_{t-2} = 1, s_{t-3} = 2)$$

$$= \lambda_1 q_{1,2}$$

Mixture transition distribution

Raftery (1985)

Example with two states and three lags:

$$Q = \begin{pmatrix} q_{1,1} & q_{1,2} \\ q_{2,1} & q_{2,2} \end{pmatrix}$$

first order transition matrix

$$\Pr(s_t = 2 \mid s_{t-1} = 1, s_{t-2} = 1, s_{t-3} = 2)$$

$$= \lambda_1 q_{1,2} + \lambda_2 q_{1,2}$$

Mixture transition distribution

Raftery (1985)

Example with two states and three lags:

$$Q = \begin{pmatrix} q_{1,1} & q_{1,2} \\ q_{2,1} & q_{2,2} \end{pmatrix}$$

first order transition matrix

$$\Pr(s_t = 2 \mid s_{t-1} = 1, s_{t-2} = 1, s_{t-3} = 2)$$

$$= \lambda_1 q_{1,2} + \lambda_2 q_{1,2} + \lambda_3 q_{2,2}$$

Sparse probability vectors

Sparse Dirichlet mixture (SDM) prior

$$p(\boldsymbol{\theta}) \propto \mathrm{Dir}(\boldsymbol{\theta}; \boldsymbol{\alpha}) \times \sum_{k=1}^{K} \theta_k^{\beta}, \quad \beta > 1$$

Sparse probability vectors

Sparse Dirichlet mixture (SDM) prior

$$p(\boldsymbol{\theta}) \propto \mathrm{Dir}(\boldsymbol{\theta};\boldsymbol{\alpha}) \times \sum_{k=1}^{K} \theta_k^{\beta}\,, \quad \beta > 1$$

$$p_{\mathrm{SDM}}(\boldsymbol{\theta};\boldsymbol{\alpha},\beta) = \sum_{k=1}^{K} \frac{w_k}{\sum_{j=1}^{K} w_j} \mathrm{Dir}(\boldsymbol{\theta};\boldsymbol{\alpha} + \beta \mathbf{e}_k)$$

$$w_k = \prod_{j=1}^{K} \Gamma(\alpha_j + \beta 1_{(j=k)})$$

Sparse probability vectors

Sparse Dirichlet mixture (SDM) prior

$$p(\boldsymbol{\theta}) \propto \mathrm{Dir}(\boldsymbol{\theta}; \boldsymbol{\alpha}) \times \sum_{k=1}^{K} \theta_k^{\beta}\,, \quad \beta > 1$$

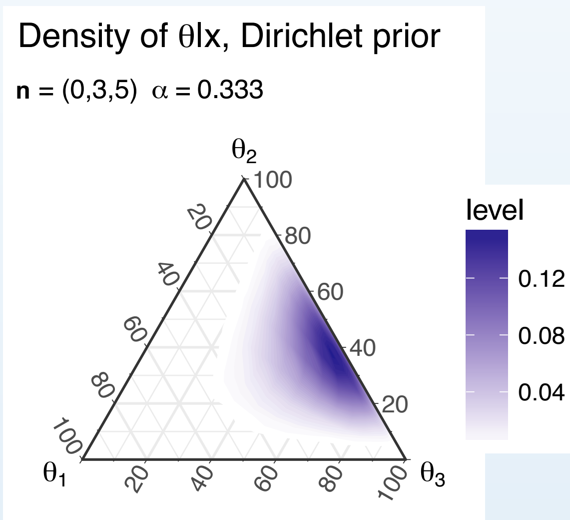
$$p_{\mathrm{SDM}}(\boldsymbol{\theta}; \boldsymbol{\alpha}, \beta) = \sum_{k=1}^{K} \frac{w_k}{\sum_{j=1}^{K} w_j} \mathrm{Dir}(\boldsymbol{\theta}; \boldsymbol{\alpha} + \beta \mathbf{e}_k)$$

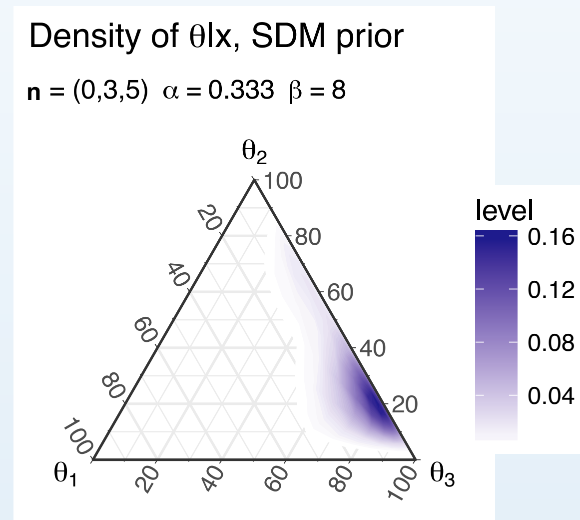
$$w_k = \prod_{j=1}^{K} \Gamma(\alpha_j + \beta 1_{(j=k)})$$

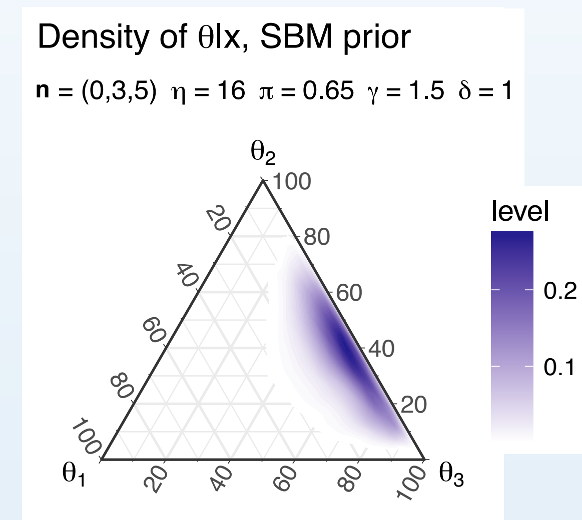
$$p(\boldsymbol{\theta} \mid \boldsymbol{n}) \quad \propto \quad \prod_{k=1}^{K} \theta_k^{n_k} \times \mathrm{Dir}(\boldsymbol{\theta}; \boldsymbol{\alpha}) \times \sum_{k=1}^{K} \theta_k^{\beta} \quad \propto \quad \mathrm{Dir}(\boldsymbol{\theta}; \boldsymbol{\alpha} + \boldsymbol{n}) \times \sum_{k=1}^{K} \theta_k^{\beta}$$

Sparse probability vectors

Multinomial data example

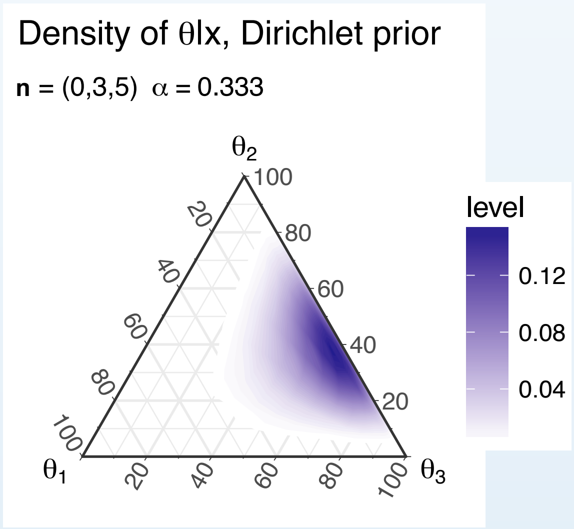


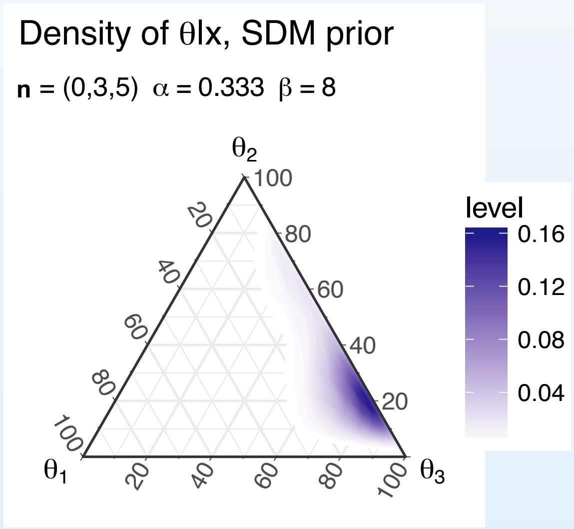


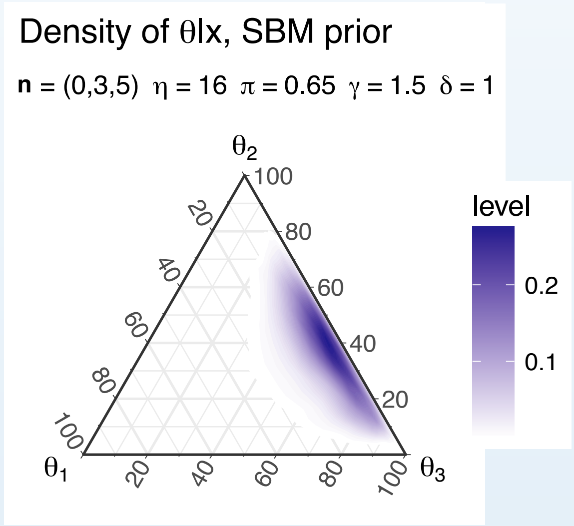


Sparse probability vectors

Multinomial data example



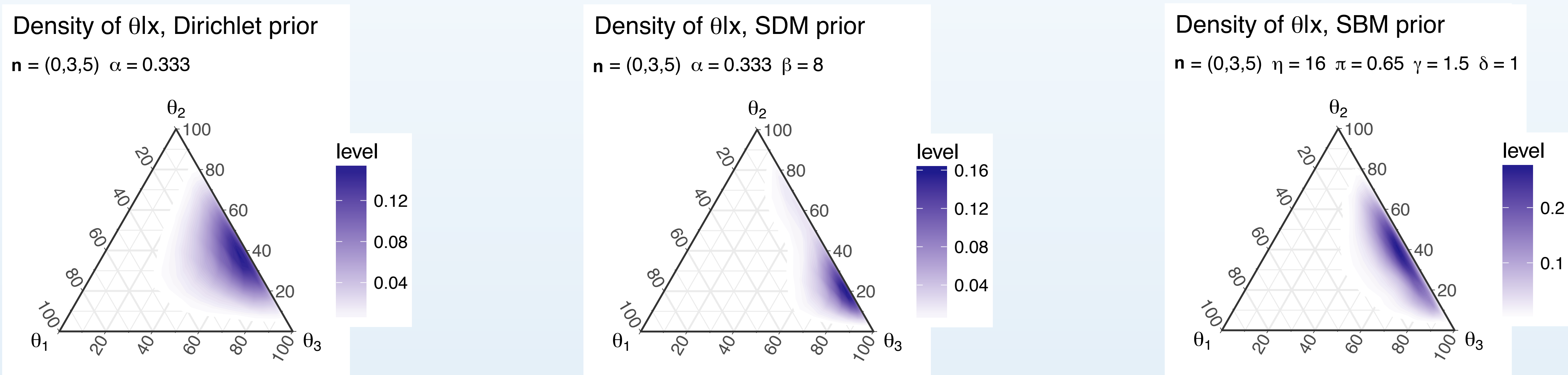




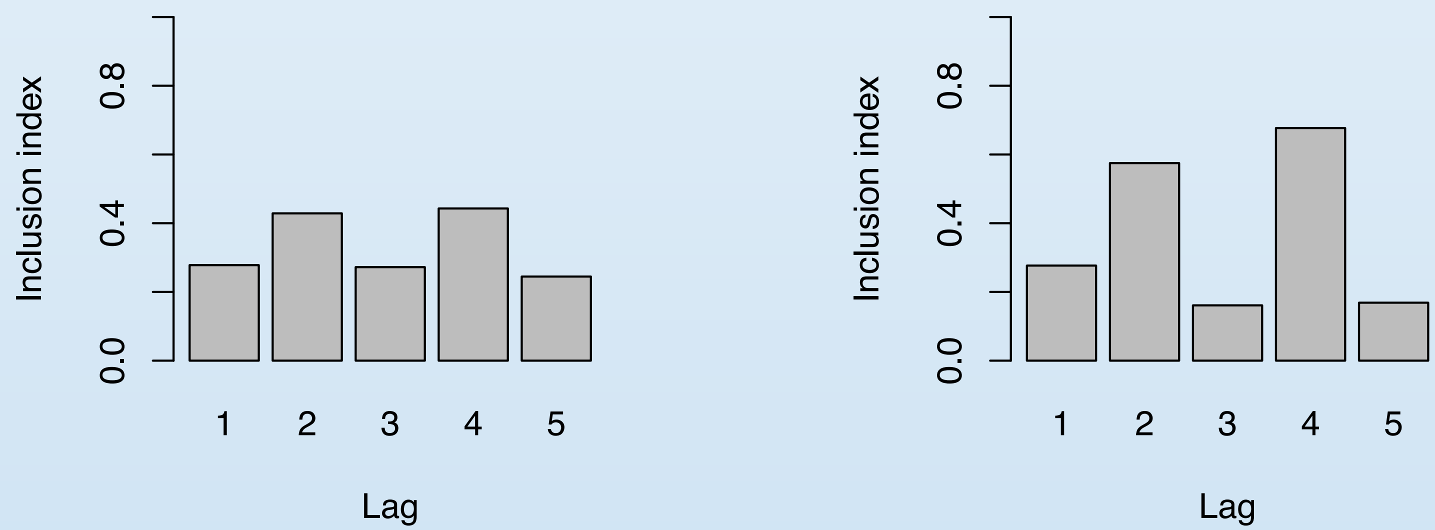
Lag inclusion inference, salmon data

Sparse probability vectors

Multinomial data example



Lag inclusion inference, salmon data



Dirichlet prior SDM prior

Acknowledgements:

San Francisco Bay Area Chapter of ASA Student Travel Award for JSM 2018.

The work of the first and the second author was supported in part by the National Science Foundation under award DMS 1310438.

References:

Berchtold, André, and Adrian E. Raftery. "The mixture transition distribution model for high-order Markov chains and non-Gaussian time series.” *Statistical Science* (2002): 328-356.

Connor, Robert J., and James E. Mosimann. "Concepts of independence for proportions with a generalization of the Dirichlet distribution." *Journal of the American Statistical Association* 64.325 (1969): 194-206.

Raftery, Adrian E. "A model for high-order Markov chains." *Journal of the Royal Statistical Society. Series B (Methodological)* (1985): 528-539.