(integrated) Step-Selection Analysis

Tal Avgar

Introduction

Habitat Selection (Probability) Function: a phenomenological model of the probability that an available discrete spatial unit will be used given its habitat type/value

- Availability varies over time and space:
  - Environments change: depletion, phenology, temperature, light, etc.
  - Cognitive processes: sensory perception, learning, and memory
  - Movement: position in geographical space dictates availability in environmental space
- Case-control design: matching used positions with a set of available positions based on spatial and/or temporal proximity

Introduction

Habitat Selection (Probability) Function: a phenomenological model of the probability that an available discrete spatial unit will be used given its habitat type/value

- Availability varies over time and space:
  - Environments change: depletion, phenology, temperature, light, etc.
  - Cognitive processes: sensory perception, learning, and memory
  - Movement: position in geographical space dictates availability in environmental space
- Case-control design: matching used positions with a set of available positions based on spatial and/or temporal proximity

Introduction

Habitat Selection (Probability) Function: a phenomenological model of the probability that an available discrete spatial unit will be used given its habitat type/value

- Availability varies over time and space:
  - Environments change: depletion, phenology, temperature, light, etc.
  - Cognitive processes: sensory perception, learning, and memory
  - Movement: position in geographical space dictates availability in environmental space
- Case-control design: matching used positions with a set of available positions based on spatial and/or temporal proximity

Step-Selection Analysis

- Match each 'observed step' (a straight line connecting two consecutive observed positions) with a random set of 'available steps' $[i = 1{:}\, n]$, sampled from the empirical distribution of observed steps
- Use conditional logistic regression to obtain maximum likelihood parameter (B) estimates:

$$L(\mathrm{B} | x_1, \ldots, x_t, \ldots, x_T) = \prod_{t=1}^{T} \frac{exp[\mathrm{B} \cdot H(x_t, t)]}{\sum_{i=1}^{n} exp\left[\mathrm{B} \cdot H\left(x_{t,i}, t\right)\right]}$$

Fortin, D., Beyer, H.L., Boyce, M.S., Smith, D.W., Duchesne, T. & Mao, J.S. (2005). Wolves influence elk movements: behavior shapes a trophic cascade in Yellowstone National Park. Ecology, 86, 1320–1330.

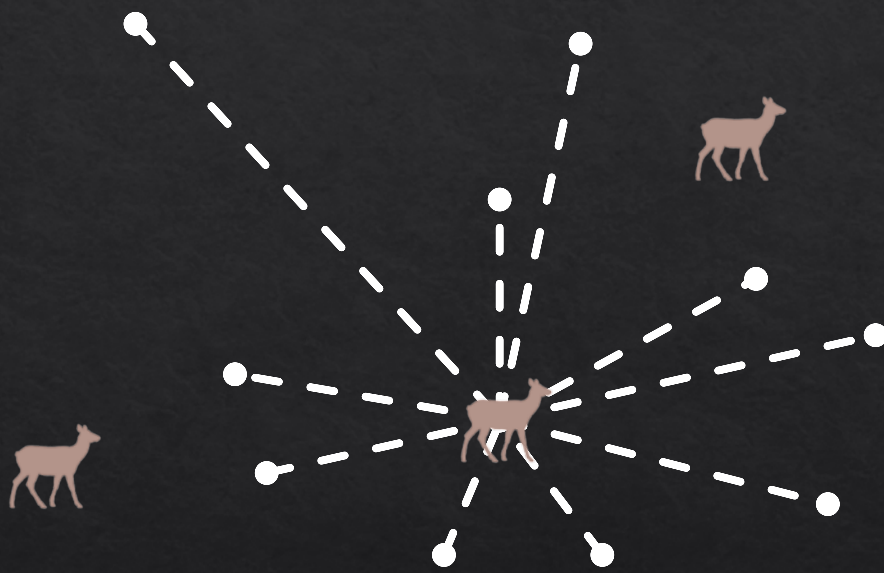
Step-Selection Analysis

- Match each ‘observed step’ (a straight line connecting two consecutive observed positions) with a random set of ‘available steps’ $[i = 1{:}\,n]$, sampled from the empirical distribution of observed steps
- Use conditional logistic regression to obtain maximum likelihood parameter (B) estimates:

$$L(\mathrm{B}|x_1, \dots, x_t, \dots, x_T) = \prod_{t=1}^{T} \frac{exp[\mathrm{B} \cdot H(x_t, t)]}{\sum_{i=1}^{n} exp\left[\mathrm{B} \cdot H\left(x_{t,i}, t\right)\right]}$$

Fortin, D., Beyer, H.L., Boyce, M.S., Smith, D.W., Duchesne, T. & Mao, J.S. (2005). Wolves influence elk movements: behavior shapes a trophic cascade in Yellowstone National Park. Ecology, 86, 1320–1330.

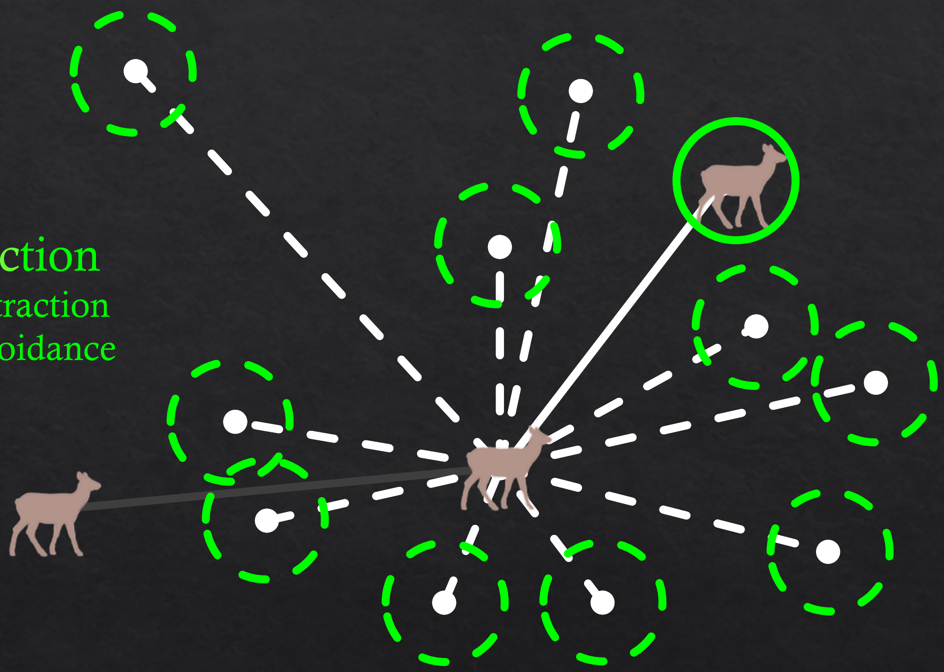
Step-Selection Analysis



graphics adapted from Prokopenko 2016

Step-Selection Analysis

selection
- attraction
- avoidance



graphics adapted from Prokopenko 2016

Step-Selection Analysis

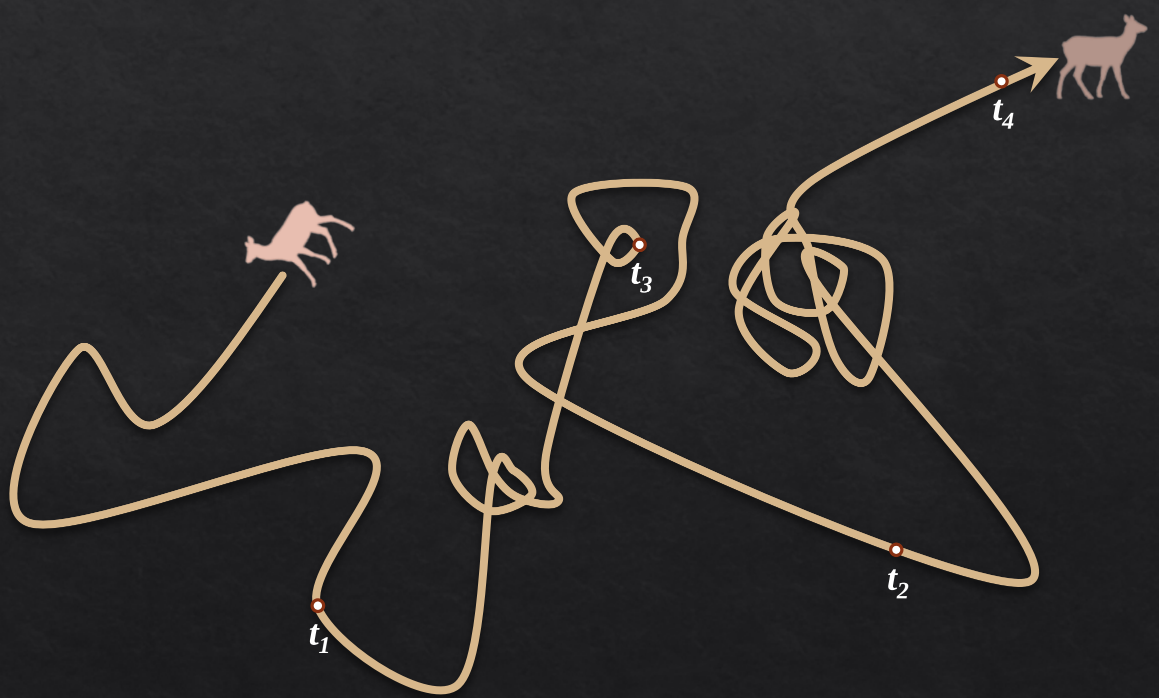
1. Obtain individual positional data
2. Define observed ('used') steps: pairs of fixes taken at a <u>constant time interval</u>
3. Match each observed steps with a random set of 'available' steps, sampled from the observed pool of steps (or pairs of steps if turn angles are also sampled)
4. Attribute all steps with environmental covariates, measured either at the step's endpoint, or 'along' the step
5. Fit a conditional logistic regression (`clogit` in R) with a binary 'used'/'available' as the response variable, environmental covariates as predictors, and the observed step's ID as the 'strata'
6. Gain inference: estimate the *Relative Selection Strength* for the various environmental covariates

Step-Selection Analysis

1. Obtain individual positional data
2. Define observed (‘used’) steps: pairs of fixes taken at a constant time interval
3. Match each observed steps with a random set of ‘available’ steps, sampled from the observed pool of steps (or pairs of steps if turn angles are also sampled)
4. Attribute all steps with environmental covariates, measured either at the step’s endpoint, or ‘along’ the step
5. Fit a conditional logistic regression (`clogit` in R) with a binary ‘used’/‘available’ as the response variable, environmental covariates as predictors, and the observed step’s ID as the ‘strata’
6. Gain inference: estimate the *Relative Selection Strength* for the various environmental covariates

Path geometry

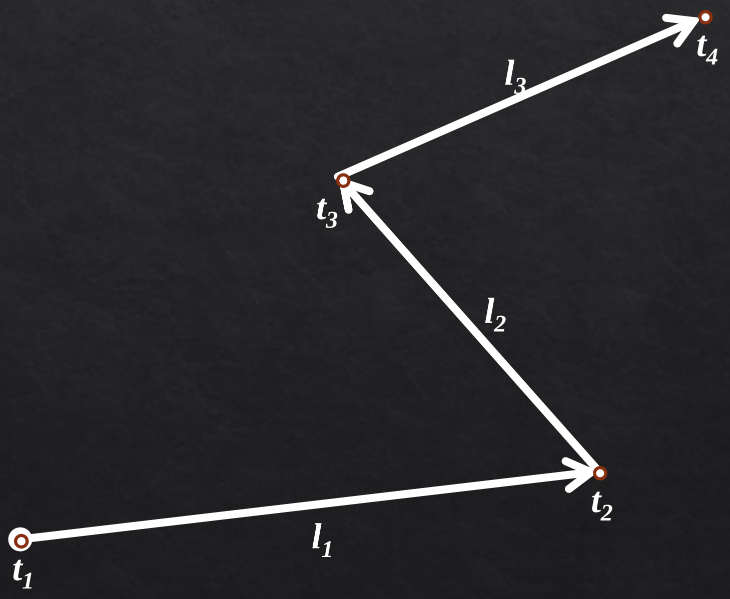
Path geometry



$$(t_4 - t_3) = (t_3 - t_2) = (t_2 - t_1)$$

Path geometry

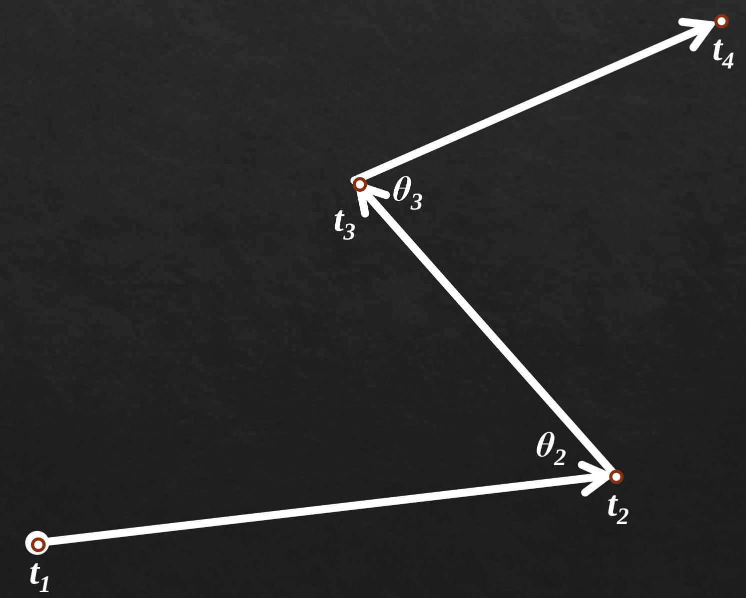
- Step length: l



$$(t_4 - t_3) = (t_3 - t_2) = (t_2 - t_1)$$

Path geometry

- Step length: l
- Directional correlation: $cos(\theta)$



$$(t_4 - t_3) = (t_3 - t_2) = (t_2 - t_1)$$

Step-Selection Analysis

1. Obtain individual positional data
2. Define observed (‘used’) steps: pairs of fixes taken at a <u>constant time interval</u>
3. Match each observed steps with a random set of ‘available’ steps, sampled from the observed pool of steps (or pairs of steps if turn angles are also sampled)
4. Attribute all steps with environmental covariates, measured either at the step’s endpoint, or ‘along’ the step
5. Fit a conditional logistic regression (`clogit` in R) with a binary ‘used’/‘available’ as the response variable, environmental covariates as predictors, and the observed step’s ID as the ‘strata’
6. Gain inference: estimate the *Relative Selection Strength* for the various environmental covariates

Step-Selection Analysis

1. Obtain individual positional data
2. Define observed (‘used’) steps: pairs of fixes taken at a constant time interval
3. Match each observed steps with a random set of ‘available’ steps, sampled from the observed pool of steps (or pairs of steps if turn angles are also sampled)
4. Attribute all steps with environmental covariates, measured either at the step’s endpoint, or ‘along’ the step
5. Fit a conditional logistic regression (`clogit` in `R`) with a binary ‘used’/’available’ as the response variable, environmental covariates as predictors, and the observed step’s ID as the ‘strata’
6. Gain inference: estimate the *Relative Selection Strength* for the various environmental covariates

Step-Selection Analysis

- In SSA, habitat-selection inference is conditional on movement, whereas movement (steps sampled from the empirical distribution) is assumed independent of habitat selection
- Include 'step length' as a covariate to control for the effect of movement and reduce bias in habitat-selection inference

Forester, J.D., Im, K.H. & Rathouz, P.J. (2009). Accounting for animal movement in estimation of resource selection functions: sampling and data analysis. *Ecology*, 90, 3554–65.

Step-Selection Analysis

- In SSA, habitat-selection inference is conditional on movement, whereas movement (steps sampled from the empirical distribution) is assumed independent of habitat selection
- Include 'step length' as a covariate to control for the effect of movement and reduce bias in habitat-selection inference

Forester, J.D., Im, K.H. & Rathouz, P.J. (2009). Accounting for animal movement in estimation of resource selection functions: sampling and data analysis. *Ecology*, 90, 3554–65.

integrated Step-Selection Analysis

- In iSSA, available step are sampled from explicitly defined theoretical distributions
- Step length and turn angles are then included as covariates, with associated ‘selection coefficients’ that correspond to the parameter vector
- Use conditional logistic regression to simultaneously obtain maximum likelihood parameter estimates for habitat selection and movement:

$$L(\mathrm{B}(\tau), \mathrm{A}(\tau) | x_1, \ldots, x_t, \ldots, x_T) = \prod_{t=1}^{T} \frac{exp[\mathrm{B}(\tau) \cdot H(x_t, \tau) + \mathrm{A}(\tau) \cdot f(x_{t-2\tau}, x_{t-\tau}, x_t)]}{\sum_{i=1} exp[\mathrm{B}(\tau) \cdot H(x_{t,i}, \tau) + \mathrm{A}(\tau) \cdot f(x_{t-2\tau}, x_{t-\tau}, x_{t,i})]}$$

Duchesne, T., Fortin, D. & Rivest, L. (2015). Equivalence between step selection functions and biased correlated random walks for statistical inference on animal movement. *PLoS ONE*, 10, e0122947.

Avgar, T., Potts, J.R., Lewis, M.A. & Boyce, M.S. (2016). Integrated step selection analysis: bridging the gap between resource selection and animal movement. *Methods Ecol. Evol.*, 7, 619–630.

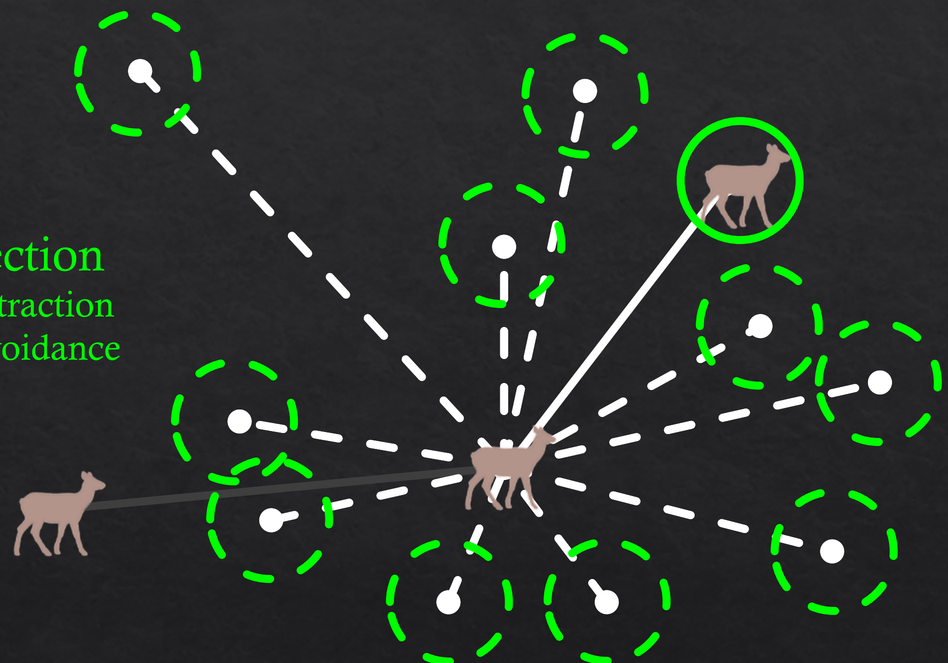
integrated Step-Selection Analysis

- In iSSA, available step are sampled from explicitly defined theoretical distributions
- Step length and turn angles are then included as covariates, with associated 'selection coefficients' that correspond to the parameter vector
- Use conditional logistic regression to simultaneously obtain maximum likelihood parameter estimates for habitat selection ($\mathrm{B}(\tau)$) and movement ($\mathrm{A}(\tau)$):

$$L(\mathrm{B}(\tau), \mathrm{A}(\tau) | x_1, \ldots, x_t, \ldots, x_T) = \prod_{t=1}^{T} \frac{exp[\mathrm{B}(\tau) \cdot H(x_t, t) + \mathrm{A}(\tau) \cdot f(x_{t-2\tau}, x_{t-\tau}, x_t)]}{\sum_{i=1}^{n} exp\left[\mathrm{B}(\tau) \cdot H\left(x_{t,i}, t\right) + \mathrm{A}(\tau) \cdot f\left(x_{t-2\tau}, x_{t-\tau}, x_{t,i}\right)\right]}$$

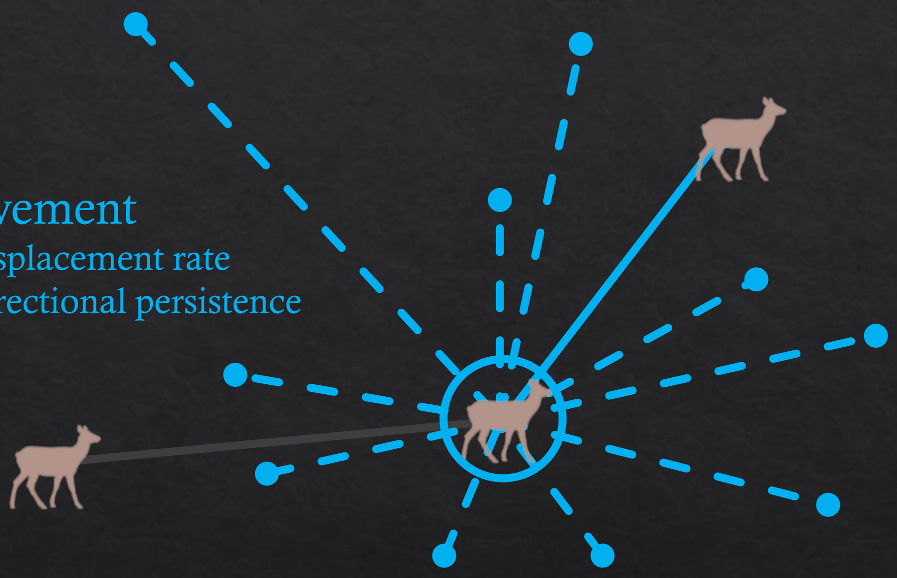
integrated Step-Selection Analysis

selection
- attraction
- avoidance



integrated Step-Selection Analysis

movement
- displacement rate
- directional persistence



integrated Step-Selection Analysis

1. Obtain individual positional data
2. Define observed ('used') steps: pairs of fixes taken at a <u>constant time interval</u>
3. Match each observed steps with a random set of 'available' steps, sampled from tentative analytical distributions of step-lengths and turn-angles
4. Attribute all steps with environmental covariates, measured either at the step's endpoint, or 'along' the step, and the step's length ($l, ln(l), l^2, ln(l)^2$) and turn angle ($cos(\theta)$)
5. Fit a conditional logistic regression (`clogit` in R) with a binary 'used'/'available' as the response variable, step attributes as the predictors, and the observed step's ID as the 'strata'
6. Gain inference: estimate the *Relative-Selection Strength* for the various environmental covariates, and the parameters of the *Selection-Free Movement Kernel*, obtained by combining the tentative parameter values with those obtained in the analysis

integrated Step-Selection Analysis

1. Obtain individual positional data
2. Define observed (‘used’) steps: pairs of fixes taken at a <u>constant time interval</u>
3. Match each observed steps with a random set of ‘available’ steps, sampled from tentative analytical distributions of step-lengths and turn-angles
4. Attribute all steps with environmental covariates, measured either at the step’s endpoint, or ‘along’ the step, and the step’s length ($l, ln(l), l^2, ln(l)^2$) and turn angle ($cos(\theta)$)
5. Fit a conditional logistic regression (`clogit` in R) with a binary ‘used’/’available’ as the response variable, step attributes as the predictors, and the observed step’s ID as the ‘strata’
6. Gain inference: estimate the *Relative-Selection Strength* for the various environmental covariates, and the parameters of the *Selection-Free Movement Kernel*, obtained by combining the tentative parameter values with those obtained in the analysis

integrated Step-Selection Analysis

Random (‘control’ or ‘available’) steps are sampled from:

step length

- Exponential: $ln\left(f_{\text{exponential}}(l|\lambda)\right) = \ln(\lambda) - \lambda l$
- Half-Normal: $ln(f_{\text{halfNormal}}(l|\sigma)) = -\ln\left(\frac{\sqrt{2}}{\sigma\sqrt{\pi}}\right) - \frac{1}{2\sigma^2} l^2$
- Log-Normal: $ln\left(f_{\text{logNormal}}(l|\sigma,\mu)\right) = \left(-ln(\sigma\sqrt{2\pi}) - \frac{\mu^2}{2\sigma^2}\right) - \left(\frac{\mu}{\sigma^2} - 1\right) ln(l) - \frac{1}{2\sigma^2} ln(l)^2$
- Gamma: $ln\left(f_{\text{gamma}}(l|k,q)\right) = -ln(\Gamma(k)q^{\kappa}) - \frac{1}{q} l + (k-1) ln(l)$

directional bias

- von Mises: $ln(f_{\text{vonMises}}(\theta|\nu,\mu = 0)) = \nu cos(\theta) - ln[2\pi I_0(\nu)]$

integrated Step-Selection Analysis

1. Obtain individual positional data
2. Define observed ('used') steps: pairs of fixes taken at a constant time interval
3. Match each observed steps with a random set of 'available' steps, sampled from tentative analytical distributions of step-lengths and turn-angles
4. Attribute all steps with environmental covariates, measured either at the step's endpoint, or 'along' the step, and the step's length ($l, ln(l), l^2, ln(l)^2$) and turn angle ($cos(\theta)$)
5. Fit a conditional logistic regression (`clogit` in R) with a binary 'used'/'available' as the response variable, step attributes as the predictors, and the observed step's ID as the 'strata'
6. Gain inference: estimate the *Relative-Selection Strength* for the various environmental covariates, and the parameters of the *Selection-Free Movement Kernel*, obtained by combining the tentative parameter values with those obtained in the analysis

integrated Step-Selection Analysis

1. Obtain individual positional data
2. Define observed ('used') steps: pairs of fixes taken at a constant time interval
3. Match each observed steps with a random set of 'available' steps, sampled from tentative analytical distributions of step-lengths and turn-angles
4. Attribute all steps with environmental covariates, measured either at the step's endpoint, or 'along' the step, and the step's length ($l, ln(l), l^2, ln(l)^2$) and turn angle ($cos(\theta)$)
5. Fit a conditional logistic regression (`clogit` in `R`) with a binary 'used'/'available' as the response variable, step attributes as the predictors, and the observed step's ID as the 'strata'
6. Gain inference: estimate the *Relative-Selection Strength* for the various environmental covariates, and the parameters of the *Selection-Free Movement Kernel*, obtained by combining the tentative parameter values with those obtained in the analysis

iSSA: calculating ‘selection-free’ movement parameters

- Exponential rate parameter: $\hat{\lambda} = \lambda_0 - \beta_l$
- half-Normal scale parameter: $\hat{\sigma} = \sigma_0 / \sqrt{1 - 2\sigma_0 \beta_{l^2}}$
- log-Normal mean: $\hat{\mu} = [\mu_0 - \sigma_0 \beta_{ln(l)}] / [1 - 2{\sigma_0}^2 \beta_{ln(l)^2}]$
- log-Normal standard deviation: $\hat{\sigma} = \sigma_0 / \sqrt{1 - 2{\sigma_0}^2 \beta_{ln(l)^2}}$
- Gamma shape parameter: $\hat{k} = k_0 + \beta_{ln(l)}$
- Gamma scale parameter: $\hat{q} = 1 / \left(\frac{1}{q_0} - \beta_l\right)$
- von Mises concentration parameter: $\hat{\nu} = \nu_0 + \beta_{cos(\theta)}$

iSSA FAQs

- Can I compare results from data sampled at different frequencies?
- Can I test effects of covariates that do not vary within a cluster?
- How many available steps do I need for each used step?
- How do I calculate the expected ‘selection-free’ speed based on my iSSA results?
- How do I interpret interactions?
- Can I estimate a Directional Bias (rather than Directional Correlation)?