Habitat Selection and Species Distribution Models

John Fieberg, Associate Professor
Department of Fisheries, Wildlife and Conservation Biology

Outline

1. Introduce Resource Selection Functions (RSFs) and Species Distribution Models (SDMs)
2. Illustrate a simple method for fitting models (logistic regression)
3. Discuss parameter interpretation

Habitat Selection versus Species Distribution Models

Habitat or resource selection functions (RSFs): models fit to observations typically collected from several individuals using tracking devices.

Species distribution models (SDMs): models fit to locations of a group of individuals (often without timestamps).

‘ISI’s Essential Science Indicators identifies species distribution modeling as the top ranked research front in ecology and the environmental sciences.’ (Renner and Warton 2013)

RSFs and SDMs

Data

- Locations of plants or animals
- Remotely sensed environmental covariates, weather (temp, precip data), habitat, etc

Objectives:

- Link species occurrence (or abundance) to resources, risks, and environmental conditions
- Predict distributions in novel environments
  - Areas not previously sampled
  - In response to climate change or habitat manipulations
Lots of modeling approaches (and jargon)

*We are modeling the spatial distribution of locations as a function of spatial covariates....*

Resources (more is better), risks (less is better), and conditions (not too much or too little)

Use-availability, presence-background

Models typically compare locations where animals are found to...

- Many ways to select points (depending on scale of inference)

‘Preference’ = used/availability depends critically on what the researcher deems is available!


Google Scholar: 3647 citations as of May 16, 2018!

Logistic Regression

Consider a prospective study:
- involving $n$ sites with camera traps
- species detections $y_i = 1$ if detected (0 otherwise)
- spatial predictors $(x_{i1}, \ldots, x_{ip})$

Model for probability of detecting a species:

$$y_i \sim \text{Bernoulli}(p_i)$$

$$\logit(p_i) = \log \left( \frac{p_i}{1 - p_i} \right) = \beta_0 + \beta_1 x_{i1} + \ldots + \beta_p x_{ip}$$

Probability (site used)

\[ p_i = \frac{\exp(\beta_0 + \beta_1 x_{i1} + \ldots + \beta_p x_{ip})}{1 + \exp(\beta_0 + \beta_1 x_{i1} + \ldots + \beta_p x_{ip})} \]

Telemetry Studies

1. Compare used locations \((y_i = 1)\) with available locations \((y_i = 0)\) that may also be used.

2. \(P(y_i = 1)\) depends on the ratio of used to available points (which is under control of the analyst). The data do not follow a Bernoulli distribution!

Lots of Historical Debate . . .

▶ Manly et al. (2002) OK if . . . availability points sampled without replacement, prior to used points being collected, no overlap between used and available points.
▶ Keating and Cherry (2004) argued strongly against
▶ Johnson et al. (2006), Lele and Keim (2006) . . . generally OK
▶ Warton & Shepherd (2010), Aarts et al. (2012), Fithian and Hastie (2013) made connections to a point process model.

Notes on probability of use

Traditionally, resource-selection functions were described as measuring “relative probabilities of use.”

▶ Probability of using a site depends on the size of the site and how long individuals are monitored.
▶ Probability of using a point in space = 0 (to ensure integration over space = 1 for continuous probability distributions).
▶ Better to think of modeling hazards (rates of use), which can be integrated over time or space to estimate utilization distributions.

Logistic Regression

For use availability designs, we focus on:

\[ w(x, \beta) = \exp(x_1 \beta_1 + x_2 \beta_2 + \ldots + x_p \beta_p) \]

Frameworks for Interpreting Resource Selection Functions:

▶ Weighted Distribution Theory
▶ Inhomogeneous Poisson Process Models
Traditional ‘Use-Availability’ likelihood

Lele and Keim (2006), weighted distribution theory:

\[
    f^u(x) = \frac{w(x, \beta)f^a(x)}{\int_{s \in A} w(s, \beta)f^a(x)ds}
\]

- \( f^u(x) \) = distribution of used habitat
- \( f^a(x) \) = distribution of available habitat

\( w(x, \beta) \) is a function that takes us from “available” to “used”.

We are modeling the spatial distribution of ‘used’ locations, as a function of covariates (through \( w(x, \beta) \)), while accounting for what is ‘available’!

IPP Model; The Grand Unifier

- **Maxent** (Aarts et al. 2012, Renner and Warton 2013, Fithian and Hastie 2013)
- **Logistic regression** (Warton & Shepherd 2010, Fithian and Hastie 2013)
  - If model is correctly specified.
  - If available points are given *arbitrarily large* weights.
- **Poisson regression** applied to grid cells (Aarts et al. 2012)
- **Weighted distribution theory** with exponential model (Lele and Keim 2006, Aarts et al. 2011)
- **Resource utilization functions** (\( \log(UD_{KDE}) \sim \text{covariates} \)) (Hooten et al. 2013)

Point process models

A whole area of spatial statistics devoted to modeling points in space.

The Inhomogeneous Poisson Process (IPP) model often serves as a starting point...

It also serves to unify many different methods for modeling use:availability or presence only data!

IPP

Model the intensity (\( \lambda \)) of a Poisson process as a log-linear function of spatial covariates: \( \log(\lambda(s)) = x(s)\beta \).

Assumptions:

- The number of events in an area \( A \) is given by a Poisson random variable with mean = \( \int_A \lambda(s)ds \).
- If \( \lambda \) is constant in \( A \), then this is equivalent to a Poisson random variable with mean that depends on the area of \( A \): \( N \sim \text{Poisson}(\lambda|A|) \).
- The number of events in disjoint areas are independent.
Likelihood: Inhomogeneous Poisson Process Model

If we condition on the total number of observed points:

\[
L(\beta; x_i) = \frac{\exp(x_i \beta)}{\int_{s \in A} \exp(x(s) \beta) \, ds}
\]

This is the same use-availability likelihood, with:

- \( w(x, \beta) = \exp(x \beta) \)
- \( f^a(x) = \) constant (i.e., uniform availability in geographic space).
- “Available points” are used to numerically evaluate the integral (Warton & Shepherd 2010, Aarts et al. 2012).

Logistic regression

Steps:

1. Observed locations \((Y_i = 1)\)
2. Sample available locations (randomly, systematically) from an area \(A\) \((Y_i = 0)\).
3. Assign weights (1 to used points, and a large number - say 1000 - to available points)
4. Fit logistic regression model using the weights (throw away the intercept).
5. Increase the number of available points until slope parameters are stable.

Logistic regression and IPP model

- Logistic regression provides unbiased estimates of \( \beta \) in the IPP model if \( n_a \) is “large enough” (Warton and Shepherd 2010).
- Fithian and Hastie (2013) showed logistic regression results in biased estimators of \( \beta \) in finite samples, unless available points are given large weights.
  - In practice, assign \( W = 1000 \) to available points, 1 to used points.

How to Create a Map

Approximate \( f^u(x) = \frac{\exp(x \beta)}{\int_{s \in A} \exp(x(s) \beta) \, ds} \) with \( f^u(x) = \frac{\exp(x \beta)}{\sum_{i=1}^{n_a} \exp(x_i \beta)} \)

---


Modeling Leroy’s Habitat Use

Leroy is a Fisher from Upstate New York, tracked as part of a larger telemetry study designed to quantify the use and importance of habitat corridors (LaPoint et al. 2013).

- Used Env-Data to merge on data layers representing population density, elevation, landcover.

³ Photo of a fisher by ForestWander Nature Photography (ForestWander.com)

Parameter Interpretation

|                        | Estimate | Std. Error | z value | Pr(|z|) |
|------------------------|----------|------------|---------|--------|
| (Intercept)            | -5.782   | 0.526      | -10.985 | 0.000  |
| elev                   | 7.707    | 0.707      | 10.902  | 0.000  |
| popD                   | 0.284    | 0.039      | 7.288   | 0.000  |
| landCgrass             | -1.503   | 0.355      | -4.237  | 0.000  |
| landCother             | 1.087    | 0.140      | 7.759   | 0.000  |
| landCshrub             | -1.394   | 0.707      | -1.971  | 0.049  |
| landCwet               | 0.267    | 0.222      | 1.206   | 0.228  |

Consider two points, $s_1$ and $s_2$, both equally accessible, in the same habitat category (let’s say “wet”), that have the same population density, and...

- we would expect the animal to select the 2nd observation with higher elevation.

Quantitative interpretation

Consider two locations, $s_1$ and $s_2$, both equally accessible, in the same habitat category (let’s say “wet”), that have the same population density, and...

- elevation at $s_2$ is 1 unit higher than at $s_1$ (it is important to know the units of elevation)

Claim: Leroy is $\exp(b_{elev}) = \exp(7.707)$ times more likely to use $s_2$ than $s_1$.

We can calculate the relative risk of an animal using $s_2$ relative to $s_1$ as:

$$\frac{f_u(x_{s_2})}{f_u(x_{s_1})} = \frac{w(x_{s_2}, \beta)f_\alpha(x_{s_2})}{w(x_{s_1}, \beta)f_\alpha(x_{s_1})}$$

where we have dropped $\int_{s \in A} w(x, \beta)f_\alpha(x)ds$ since it appears in both numerator and denominator.
Quantitative interpretation

\[
\frac{f^u(x_{s_2})}{f^u(x_{s_1})} = \frac{\exp(elev_2 \beta_{elev} + popD_2 \beta_{popD} + \beta_{wet})f^a(x_{s_2})}{\exp(elev_1 \beta_{elev} + popD_1 \beta_{popD} + \beta_{wet})f^a(x_{s_1})}
\]  
(1)

Setting:

- \(elev_2 = elev_1 + 1\)
- \(popD_2 = popD_1\)
- \(f^a(x_{s_1}) = f^a(x_{s_2})\) (assuming both locations are equally available)

\[
\Rightarrow \frac{f^u(s_1)}{f^u(s_2)} = \frac{\exp([elev_1 + 1] \beta_{elev})}{\exp([elev_1] \beta_{elev})} = \exp(\beta_{elev})
\]

For continuous variables, \(\beta\) gives the change in log-relative risk associated with increasing \(x\) by 1 unit, while:

- holding everything else constant
- and assuming equal availability

|             | Estimate | Std. Error | z value | Pr(>|z|) |
|-------------|----------|------------|---------|----------|
| (Intercept) | -5.782   | 0.526      | -10.985 | 0.000    |
| elev        | 7.707    | 0.707      | 10.902  | 0.000    |
| popD        | 0.284    | 0.039      | 7.288   | 0.000    |
| landCgrass  | -1.503   | 0.355      | -4.237  | 0.000    |
| landCother  | 1.087    | 0.140      | 7.759   | 0.000    |
| landCshrub  | -1.394   | 0.707      | -1.971  | 0.049    |
| landCwet    | 0.267    | 0.222      | 1.206   | 0.228    |

Given equal availability of all landcover classes, and holding elevation and population density constant

- this fisher would “select” locations in the “wet” class over grass, shrub, and... forest [the reference class].

But...availability is not equal!

- selection (use/available) is strongest for other, but use is highest for forest!
- the positive coefficient for wet reflects a larger use/available ratio relative to the reference category, forest.
What if we use a different reference class?

```r
FisherLeroy <- within(FisherLeroy, 
  landC <- relevel(landC, ref = "other"))
RSF.Leroy2 <- glm(case_ ~ elev + popD + landC, 
  data = FisherLeroy, 
  weight = w, 
  family = binomial)
```

|                | Estimate | Std. Error | z value | Pr(>|z|) |
|----------------|----------|------------|---------|----------|
| (Intercept)    | -4.695   | 0.561      | -8.370  | 0.000    |
| elev           | 7.707    | 0.707      | 10.902  | 0.000    |
| popD           | 0.284    | 0.039      | 7.288   | 0.000    |
| landCforest    | -1.087   | 0.140      | -7.759  | 0.000    |
| landCgrass     | -2.589   | 0.378      | -6.853  | 0.000    |
| landCshrub     | -2.480   | 0.719      | -3.449  | 0.001    |
| landCwet       | -0.819   | 0.258      | -3.174  | 0.002    |

- coefficients for `elev` and `popD` do not change

Note the coefficient for `forest` is also negative despite Leroy spending more than 90% of his time in the forest!

- coefficient for `wet` is now negative despite the fact that Leroy uses wet areas more than available... why?
  - because the ratio of used to available points is greater for the reference class (`other`) than for `wet`. 

- coefficients for `elev` and `popD` do not change
Summary

For continuous predictors:

- $\beta$ describes the change in log relative risk associated with increasing the value of the predictor by 1 unit, while holding all other predictors (and habitat availability) constant.

For categorical predictors:

- the $\beta$'s describe the log-relative risk of selecting different levels of the variable relative to a reference level, while holding all other predictor variables constant and assuming equal availability of the different levels of the categorical predictor.

\[
I^U(x) \propto w(x) \beta f^a(x)
\]

$\Rightarrow$ $\beta$ may change as we change habitat availability (functional responses in habitat selection)

Issues

- not all habitat is equally available
- when we move to new locations, typically more than 1 habitat covariate changes (so, everything else is not held constant).
- use may not increase proportionally with habitat availability

Let's consider population density and landcover:

If we compare locations in other and forest, population density is not likely to be held constant.

The importance of population density seems much more pronounced in the other and shrub categories. This effect could be modeled by including an interaction between population density and landcover class.