Survival Analyses

Trees, shrubs, squirrels, fish

Estimation of survival rates

- Survival is a critical population process in nearly every field of biology
- Many approaches to estimating survival, including capture-recapture methods discussed last week

Why survival?

- Survival provides insight into population demographics
- Interested in survival differences between "treatments" – herbicides, harvest plans, burning regimes
- Most management actions are directed at managing (increasing or decreasing) survival
- Important not only to know how many of are there, but what is happening to them

Finite survival rates

- Simplest measures of survival obtained by following a group of individuals over time
  - FINITE survival rate
  - # alive at end of t / # alive at start of o
  - Note that the time interval is the unit of time your data is measured
  - t = 1 year, this is a finite *annual* survival rate

$$\hat{S}_o = \frac{N_t}{N_o}$$

Finite rates

- Common need to convert to standard time intervals
- ts = standard time
- to = observed time

$$adjusted = observed^{\frac{t_s}{t_o}}$$

Finite survival rates

- 10 seedlings at start of experiment day 1
- 6 seedlings viable at end of 3 days

$$\hat{S}_{3days} = \frac{6}{10}$$

$$\hat{S}_o = \frac{N_t}{N_o}$$

$$\hat{S}_{3days} = 0.6$$

Finite survival rates

- 10 seedlings at start of experiment day 1
- 6 seedlings viable at end of 3 days
- What is daily survival rate?

$$\hat{S}_{3days} = \frac{6}{10}$$

$$\hat{S}_{3days} = 0.6$$

Finite survival rates

- 10 seedlings at start of experiment day 1
- 6 seedlings viable at end of 3 days
- What is daily survival rate?
- ts= standard (1 day)
- to= observed (3 day)

$$adjusted = observed^{\frac{t_s}{t_o}}$$

$$\hat{S}_{daily} = 0.6^{1/3}$$

$$\hat{S}_{daily} = 0.84$$

Finite survival rates

- 10 seedlings at start of experiment day 1
- 6 seedlings viable at end of 3 days
- What is daily survival rate?
- Daily survival = 84%

$$adjusted = observed^{\frac{t_s}{t_o}}$$

$$\hat{S}_{daily} = 0.6^{1/3}$$

$$\hat{S}_{daily} = 0.84$$

Finite survival rates

- 10 seedlings at start of experiment day 1
- 6 seedlings viable at end of 3 days
- What is weekly survival rate?

$$adjusted = observed^{\frac{t_s}{t_o}}$$

Finite survival rates

- 10 seedlings at start of experiment day 1
- 6 seedlings viable at end of 3 days
- What is weekly survival rate?

$$adjusted = observed^{\frac{t_s}{t_o}}$$

$$\hat{S}_{daily} = 0.6^{7/3}$$

$$\hat{S}_{daily} = 0.30$$

Finite survival rates

- 10 seedlings at start of experiment day 1
- 6 seedlings viable at end of 3 days
- What is weekly survival rate?
- Weekly survival = 30%

$$adjusted = observed^{\frac{t_s}{t_o}}$$

$$\hat{S}_{daily} = 0.6^{7/3}$$

$$\hat{S}_{daily} = 30\%$$

Cumulative finite rates….

- Cumulative effect of a constant finite rate is an exponential model

Instantaneous rates

If the survival fraction can be assumed to be constant, an exponential model is appropriate

Cases:

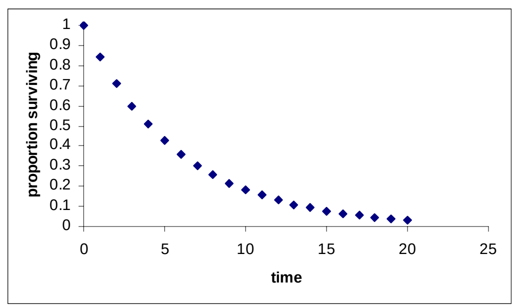
Short time periods for any analysis
Life history of organism (certain life stages)
Many uses in ecology
Instantaneous rates can sometimes be analyzed with traditional statistics such as regression, ANOVA, etc., if they meet the assumptions

Instantaneous rates

- If number of deaths in short time period proportional to the total number of individuals at that time, the rate of drop in numbers can be described by the geometric equation
  - N = number of individuals
  - z = instantaneous mortality rate (always a negative number)
  - t = time

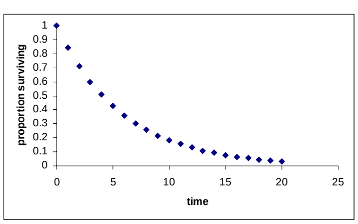
$$\frac{dN}{dt} = zN$$

Geometric population decline



Geometric population decline

$$N_t = N_o e^{-zt}$$



Geometric population decline

$$N_t = N_o e^{-zt}$$

- If t=1 and we take the log of both sides

$$\log_e \left(\frac{N_t}{N_o} \right) = -z$$

Geometric population decline

$$N_t = N_o e^{-zt}$$

- If t=1 and we take the log of both sides

$$\log_e\left(\frac{N_t}{N_o}\right) = -z$$

- What was the finite survival rate?

Geometric population decline

$$N_t = N_o e^{-zt}$$

- If t=1 and we take the log of both sides

$$\log_e\left(\frac{N_t}{N_o}\right) = -z$$

- What was the finite survival rate?

$$\hat{S}_o = \frac{N_t}{N_o}$$

Simple rules

- loge (finite survival rate) = Instantaneous mortality rate
- Instantaneous allow you to work on any scale you wish
  - Range 0 to neg. infinity
- Finite easiest to think about
  - Range 0 to 1

$$N_t = N_o e^{-zt}$$

$$S = e^{-z}$$

$$-\ln(S) = z$$

$$A = 1 - S$$

$$N_t = N_o e^{-Zt}$$
$$S = e^{-Z}$$
$$Z = -\ln(S)$$
$$A = 1 - S$$

Example:

Say the finite annual survival rate is 0.95, how many from an initial 1,000 shrubs would be alive at the end of 10 years?

$$Z = -\ln(S) = 0.0513$$

$$N_t = 1{,}000e^{-0.0513\times 10} = 599$$

$$N_t = N_o e^{-Zt}$$
$$S = e^{-Z}$$
$$Z = -\ln(S)$$
$$A = 1 - S$$

Example:

Say the finite annual survival rate is 0.5, how many from an initial 1,000 squirrels would be alive at the end of 10 years?

$$Z = -\ln(S) = 0.69$$

$$N_t = 1{,}000e^{-0.69\times 10} = 1$$

Key assumptions

- Many of these models are count based
- What is the key assumption with count data?
- If assumptions can be met, count based methods can work well

Key assumptions

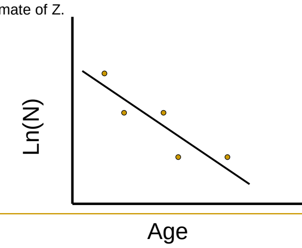
- Many of these models are count based
- What is the key assumption with count data?
  - Detection probabilities equal to 1 or are equal for different groups
- If assumptions can be met, count based methods can work well

Estimating survival from age composition

1. Changes in abundance at age

A catch curve is a regression of the ln(N) on age, which makes an exponential survival model linear.

The slope of the line is an estimate of Z.



Example:

Age	Number	ln(Number)
0	62	4.127
1	147	4.99
2	122	4.8
3	90	4.5
4	61	4.11
5	14	2.63

$$N_t = N_o e^{-Zt}$$

$$S = e^{-Z}$$

$$Z = -\ln(S)$$

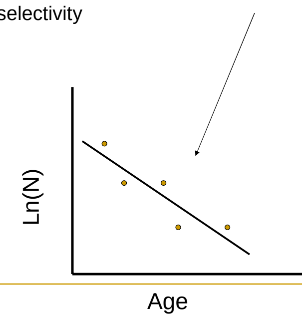
$$A = 1 - S$$

Z = -0.54
S = $e^{-0.54}$ = 0.58
A = 0.42

Assumptions of Catch Curve Analysis
1) survival is constant across ages
2) random sample of the population
3) constant recruitment each year

Estimating survival from age composition

Big assumption that this decline is a function of death!
And not emigration or selectivity



Binomial survival model

- If two mutually exclusive outcomes exist (live or dead) when assessing survival
- Used to measure nest survival, survival of radio tagged individuals, seedlings in a plot, etc.
- Things you revisit

$$n = subjects$$

$$x = \#\, survive$$

$$\#\, die = n - x$$

$$\hat{S} = \frac{x}{n}$$

$$\mathrm{var}(\hat{S}) = \frac{\hat{S}(1-\hat{S})}{n}$$

Binomial survival model

- Assumptions
  - Fates of all *n* subjects are known
    - Some subjects may not be detected
    - OR are censored (removed)
  - Fates are independent events that are identically distributed
    - No heterogeneity
    - If fates are not independent, violating binomial assumption
      - Stratify groups (different sizes)

Binomial survival

- Each time step living or dying is an independent occurrence
- Constant mortality over whole time period
- Death recorded exactly
- Example radiotelemetry survival Trent and Rongstad (1974)

$$\hat{S} = \frac{x - y}{x}$$

Daily survival
x = total number of radio days observed
y = total number of deaths observed

Binomial survival

$$\hat{S} = \frac{x - y}{x}$$

- Each time step living or dying is an independent occurrence
- Example radiotelemetry survival Trent and Rongstad (1974)
- Daily survival
- x = total number of radio days observed
- y = total number of deaths observed

$$\hat{S} = \frac{1660 - 6}{1660}$$

$$\hat{S} = 0.996$$

Binomial survival

$$\hat{S} = 0.996$$

- Convert daily rates to other time intervals
- Remember key assumption of mortality constant…realistic w/ long time intervals?
- Example 28 day survival

$$\hat{p} = \hat{S}^n$$

$$\hat{p} = .99628^{28}$$

$$\hat{p}_{28day} = 0.9036$$

Mayfield survival

- Similar approaches with nest survival (Mayfield)
- Nests are located during searches and revisited until they either fail or produce fledglings
- If all nests are found at same time, then simple binomial is fine
- Usually nests located at different times since deposition

Mayfield survival

- If S = 99% per day
- Deposition to fledging = 30 days
- Survival after 30 days is 74%
- BUT if the nest is found at the beginning of day 28
- Nest is only monitored 2 days…
- Positive bias in survival...

$$S^1 = 0.99$$

$$S^{30} = 0.74$$

$$S^2 = 0.98$$

Mayfield survival

- Mayfield (1961) original formulation to account for different lengths of time the nest are studied
- If all nests are visited each day, simple estimator works fine
- Different discovery days or unequal visits, use more recent approaches….

$$S^1 = 0.99$$

$$S^{30} = 0.74$$

$$S^2 = 0.98$$

Kaplan-Meier

- Most of the approaches we have talked about so far assume that the mortality risk can be "anticipated"
- Mortality occurs in discrete events
- Many mortality studies occur with small sample sizes, irregular "discoveries" or releases/plantings of things to monitor, or require censoring
- Kaplan-Meier method (Kaplan-Meier 1958; Pollock et al. 1989a,b)

Kaplan-Meier

- Kaplan-Meier method (Kaplan-Meier 1958; Pollock et al. 1989a,b) general approach useful in a lot of studies – very common in human health
- Provides
  - Empirical estimation of survival rates that can be graphically examined for patterns
  - Hypothesis testing based on individuals

Kaplan-Meier

- Kaplan-Meier method is based on a *hazard function* which is the instantaneous mortality rate for organisms alive at time *t*
- Allows for
  - Censoring (including temporary censoring with re-entry)
    - Really useful if with incomplete searches or temporary movement from study area
- Additions
  - New animals added to system

Kaplan-Meier Survival Estimation

Uses censored observations through conditional probabilities of survival each time period

$$S(t_j) = \prod_{t_j \le t}\left(1 - \frac{d_j}{n_j}\right)$$

Where:

$S(t_i)$ = cumulative survival probability for any particular one of the *t* time periods
n_i is the number of subjects at risk at the beginning of time period t_i, and
d_i is the number of deaths during time period t_i.

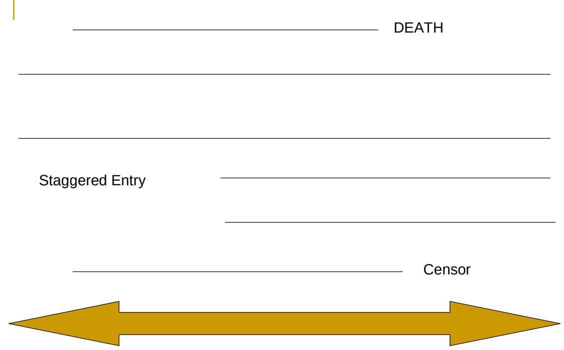
Survivorship (Cumulative Survival)

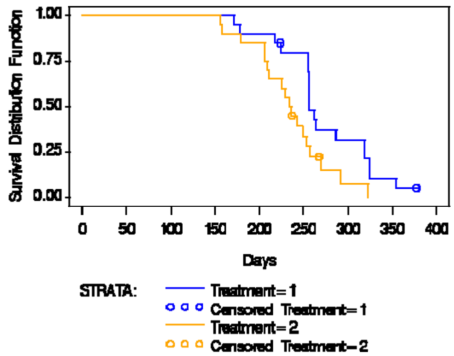
Censoring in survival data refers to observations where an event has not occurred either due to:

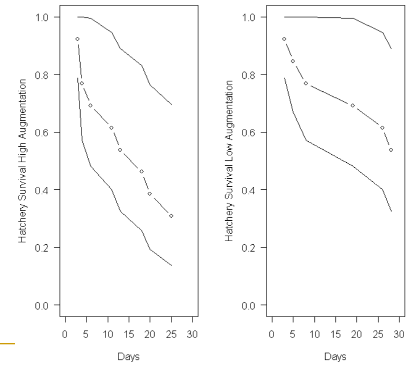
-the end of the study (only document survivors to that time)
-early removal of an observation due to a confounding issue (e.g., development of an unrelated disease)

-censoring of data must be considered in the analysis

-censoring may have some biological meaning that you wish to estimate….snail kite emigration







Kaplan-Meier

- Variety of ways to test for differences between survival distributions
- Can incorporate explanatory variables
- We will demonstrate and explore in lab using R and SAS