

Open-population capture-recapture and stopover duration models; recent advances including age-structure and heterogeneity

Tutors:

Shirley Pledger, Eleni Matechou, Murray Efford, Richard Arnold

Workshop at International Statistical Ecology Conference
NCSE, Canterbury Campus, University of Kent, U.K.
5 July 2010

Topics

- 1 Introduction
- 2 Background: the Schwarz-Arnason Model
- 3 Age-structured Models
- 4 Models with Heterogeneity
- 5 Combining Heterogeneity and Age Structure
- 6 User-defined Functions
- 7 References

1. Introduction

These notes give explanations and technical details of two suites of models for open-population capture-recapture, the age-structured models in Pledger *et al.* 2009 and the models with individual heterogeneity of capture and/or survival probabilities in Pledger *et al.* 2010.

The models may be fitted using the R code in **hetage** (to be turned into an R package).

The models are described in the main part of these notes, while more technical details of likelihoods and parameter counts are in Pledger *et al.* 2009, Pledger *et al.* 2010, and in the **hetage** package notes.

2. Background: the Schwarz-Arnason Model

Schwarz and Arnason (1996) produced a fully likelihood-based version of the Jolly-Seber model (Jolly 1965, Seber 1965) for capture-recapture in open populations.

The capture-recapture study has the simple design, with K sampling occasions widely spaced through time.
(Not the robust design, which has clusters of samples.)

Each animal has a capture history, a vector of length K with 1 for capture, 0 for non-capture, e.g.

0 0 0 0 1 0 0 1 0 1 0 0 0 0 0 0

Jolly-Seber (JS) Model

Jolly, 1965; Seber, 1965. The JS model has parameters

p_j = probability of capture for an animal present at sample j ,

ϕ_j = probability of survival to sample $j + 1$ if alive at j ,

The JS model also estimates the expected values of

N_j = number alive in the population at sample j ,

B_j = number entering the population between samples j and $j + 1$.

Schwarz-Arnason (SA) model

Schwarz and Arnason (1996) used the idea of a superpopulation N of all animals available for capture on at least one occasion (from Crosbie and Manly). They gave a fully likelihood-based version of the JS model.

The SA model has parameters

N = number in superpopulation, available for capture at least once

β_j = proportion of N arriving between samples j and $j+1$

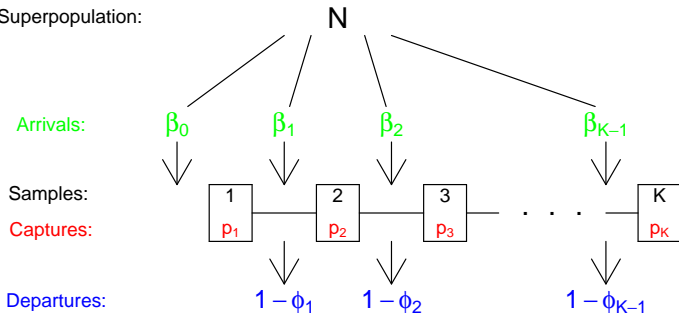
($\sum_{j=0}^{K-1} \beta_j = 1$, β_0 = proportion there at start)

ϕ_j = Prob(survive to $j+1$ given alive at j)

p_j = Prob(capture at j)

Schwarz and Arnason Parameters

Superpopulation:



Basics of SA Model

The SA Model assumes a multinomial distribution, where N animals are allocated to different capture histories.

That is, there are N trials, with allocation to different cells, where each cell is a different capture history.

Suppose we observe H different capture histories, and D distinct animals. There are y_1 animals with the first capture history, y_2 with the second, ... up to y_H with the H^{th} capture history.

$$\sum_{h=1}^H y_h = D$$

There are also $N - D$ animals unobserved, even though they were there for some of the time. They have capture history 0000...000, a vector of K zeros.

Data Matrix

Suppose there are D distinct animals observed over K samples.

The $D \times K$ capture matrix \mathbf{X} has

$x_{ij} = 1$ if animal i is seen at sample j ,

$x_{ij} = 0$ otherwise.

Row i of \mathbf{X} is the capture (encounter) history (CH) of animal i .

The data matrix \mathbf{X} may be condensed into an $H \times K$ matrix with each row being a different observed capture history. In this case, there is an associated vector \mathbf{y} of length H , where y_h is the number of animals seen with the h^{th} capture history, $h = 1, 2, \dots, H$.

Data Layout

		Matrix X								Vector y	
		Sample									
		1	2	3	4	5	6	7	8	y	
Capture History	1	0	1	0	1	1	0	0	0	6	Observed Data
	2	1	0	0	1	0	0	0	0	2	
	3	0	0	1	1	0	0	0	0	3	
	4	0	0	0	1	0	0	1	1	7	Total of D animals
	5	0	1	1	0	0	1	0	0	1	
	6	0	0	1	1	1	0	0	0	2	

The ones that got away:

0	0	0	0	0	0	0	0	0
---	---	---	---	---	---	---	---	---

$N - D$	Unobserved
---------	------------

Multinomial Likelihood for SA Model

Details of the construction of the likelihood are in Pledger *et al.* 2009.

The full likelihood of parameter vector θ (composed of N , β s, ϕ s and ps), given the data X, y , is

$$L(\theta | X, y) = \frac{N!}{y_1! y_2! \dots y_H! (N-D)!} \times \left(\prod_{h=1}^H L_h^{y_h} \right) \times (L_0)^{N-D},$$

where L_h is the likelihood of capture history h and L_0 is the likelihood of no captures.

Individual Likelihoods for SA Model

Suppose capture history h is 0 0 0 0 0 1 0 1 1 0 0 0 ($K = 12$)
Consider an arrival between samples 2 and 3, and departure
between 11 and 12.

$$L = \beta_2 \left(\prod_{j=3}^{10} \phi_j \right) (1 - \phi_{11}) \left[\prod_{j=3}^{11} p_j^{x_j} (1 - p_j)^{1-x_j} \right]$$

To find L_h , sum over all possible entry times and all possible exit times. First seen at $f_h = 6$, last seen at $\ell_h = 9$.

$$L_h = \sum_{b=1}^{f_h} \sum_{d=\ell_h}^K \beta_{b-1} \left(\prod_{j=b}^{d-1} \phi_j \right) (1 - \phi_d) \left[\prod_{j=b}^d p_j^{x_j} (1 - p_j)^{1-x_j} \right]$$

Comments on SA model

The Cormack-Jolly-Seber model (CJS) sums over possible exit times (the χ parameters).

The Schwarz-Arnason model, by using the superpopulation concept, allows summation over possible entry times as well.

Schwarz and Arnason (1996) give a simplified formula for the likelihood, as they can separate the entry parameters β from the exit parameters ϕ , and rearrange.

However, in our work on age-structured models, arrival times and departure times are not independent, so we must stay with the likelihood formulations based on individual capture histories.

Great cormorant (*Phalacrocorax carbo sinensis*)



Cormorant Data

Thanks to Thomas Bregnballe for the data, via Rachel McCrea and Ruth King.

These are monthly data from February to October 1994, of cormorants at a breeding site in Denmark.

Using all the $D = 317$ birds (file `corm_1994_BPall.csv`), we may read in the data and fit four models: the JSSA model and simpler models with ϕ and/or p constant.

Practical Session 1

- 1 Install **hetage** R files.
- 2 Have all the files in the same directory.
- 3 Open **R**.
- 4 Using the file **corm4models.R**, cut and paste the commands into **R** to do the analyses.
- 5 Interpret the output.

	MaxLL	RD	npar	AIC	relAIC	AICc	relAICc
phic.pc	-318.71	637.41	11	659.41	306.39	659.94	305.53
phic.pt	-198.58	397.15	18	433.15	80.13	434.54	80.13
phit.pc	-158.51	317.02	18	353.02	0.00	354.40	0.00
phit.pt	-153.40	306.80	24	354.80	1.79	357.27	2.86

3. Age-structured Models

Here we extend the Schwarz-Arnason capture-recapture analysis to address the question of the age of the animal, if its true age is unknown.

There are two major applications:

- 1 Settled population: entry and exit mainly by births and deaths, little or no migration.
- 2 Stopovers of migratory birds: entry and exit mainly by arrivals and departures, no births, few deaths.

New Questions

The following questions cannot be answered by the JS and SA models:

1. How can we make the survival probability depend on the age of the animal, when age is generally unknown? (How long before its first capture was the animal present?) We must address this question if we wish to detect and describe senescence.
2. How can we make departure probability depend on how long since the bird arrived, when its exact arrival time is unknown? (How long had it been there before it was first seen?) We must address this question if we want to check whether a major impetus for departure is having built up adequate body reserves of fat for the next stage of the migration. *If we want departure probability to depend on stopover duration so far, then it's somewhat of a handicap to not know when it arrived.*

Migratory Birds

Birds stop over at a site: feed, build up body condition, fly on.
They have bands for identification.

There are repeated samples, with captures (e.g. mist netting) or observation of bands.

Not all birds are observed in the sampling.

Questions

- How many stopped over?
- How long did they spend there on average?
- Is departure time related to length of stay? Body condition? Weather?
- Do birds with earlier or later arrivals differ in stopover duration? Is this related to fitness?

Notes:

- 1 The following models work best if sampling covers the whole stopover period, so most capture histories have leading and trailing zeros.
- 2 Sampling must be done at equal intervals (not necessarily every day).
- 3 If there are many samples, giving rise to a large number of parameters, a practical approach is to merge groups of samples, e.g. daily sampling could be merged into four-day blocks. At least one capture or sighting during the four days leads to a one in the new data matrix.

Assumptions for stopover models

- 1 Arrivals and departures occur, but no births or deaths.
- 2 Independence between birds (their arrivals, captures and departures).
- 3 Homogeneity of capture probabilities among birds.
- 4 Bands are read correctly, not lost, etc.

We do **not** assume independence of arrival and departure time within birds. With age-structured models, we are interested in departure time being correlated with arrival time.

Overview of how the age-structured models work

Details of the probabilities and likelihoods are in Pledger *et al.* 2009.

We use the feature of the SA models, in which there is a sum over different possible arrival times. The likelihood for bird i is a sum of likelihoods over the feasible arrival times for that bird.

Within each term of the sum (i.e. for each possible arrival time), we modify the survival parameter and/or capture parameter to take account of both the sample number (j) and the arrival time (birth cohort, b).

Modified likelihoods

Suppose bird i is first seen at sample f_i , and last seen at sample ℓ_i . Let b (d) be the unknown sample at which bird i is first (last) present and available for capture.

Schwarz-Arnason:

$$L_i = \sum_{b=1}^{f_i} \sum_{d=\ell_i}^K \beta_{b-1} \left(\prod_{j=b}^{d-1} \phi_j \right) (1 - \phi_d) \left[\prod_{j=b}^d p_j^{x_{ij}} (1 - p_j)^{1-x_{ij}} \right]$$

Age-structured models: ϕ and/or p are modified to depend on age (via the birth/arrival cohort, b) as well as on sample number j .

$$L_i = \sum_{b=1}^{f_i} \sum_{d=\ell_i}^K \beta_{b-1} \left(\prod_{j=b}^{d-1} \phi_{bj} \right) (1 - \phi_{bd}) \left[\prod_{j=b}^d p_{bj}^{x_{ij}} (1 - p_{bj})^{1-x_{ij}} \right].$$

Modified survival parameters

Schwarz-Arnason: Vector of survival probabilities,
 $\phi_j = \text{Prob}(\text{survive to sample } j+1 \mid \text{alive at sample } j)$

$$\phi_1 \quad \phi_2 \quad \phi_3 \quad \dots \quad \phi_{K-1}$$

With age structure: Matrix of survival probabilities,
 $\phi_{bj} = \text{Prob}(\text{survive to sample } j+1 \mid \text{alive at sample } j,$
 for those first present at sample $b)$

		Sample j						
		1	2	3	.	.	.	$K-1$
Arrival	1	ϕ_{11}	ϕ_{12}	ϕ_{13}	.	.	.	$\phi_{1,(K-1)}$
Time b	2		ϕ_{22}	ϕ_{23}	.	.	.	$\phi_{2,(K-1)}$
(Cohort)	3			ϕ_{33}	.	.	.	$\phi_{3,(K-1)}$

	.						.	.
	$K-1$							$\phi_{(K-1),(K-1)}$

The unseen

To complete the likelihood calculation, we also need the likelihood for an unseen animal:

$$L_0 = \sum_{b=1}^K \sum_{d=b}^K \beta_{b-1} \left(\prod_{j=b}^{d-1} \phi_{bj} \right) (1 - \phi_{bd}) \left[\prod_{j=b}^d (1 - p_{bj}) \right].$$

The full likelihood equation, which uses both L_i and L_0 , is on page 11.

For the SA model, let $\phi_{bj} = \phi_j$ and $p_{bj} = p_j$.

Connections

We have three indices, j = sample number, b = cohort (sample when first present), and a = age (or residence time).

We are now assuming equally-spaced samples, so $a = j - b + 1$.
(By convention, we say age = 1 at entry.)

If we know any two of birth time, calendar time and age, we can calculate the other one.

The previous matrix was written using cohort and sample, which is more familiar to MARK users.

However, we now switch to indexing ϕ or p by age and sample, as age is the focus of these models.

The rearrangements are done automatically in the **hetage** code.

Can't be done!

We now have a very general model with ϕ and p both varying by sample and age, ϕ_{aj} and p_{aj} .

There's only one drawback, it can't be fitted.

If we try to allow either ϕ or p to vary by both age and sample, there are altogether too many parameters.

See the next page, interactive and additive models.

Interactive and Additive Models

Interactive model

Can't estimate all ϕ_{aj} and p_{aj} . Simplify.

Additive model (logit scale)

$$\text{Let } \log \left(\frac{\phi_{aj}}{1 - \phi_{aj}} \right) = \mu + \tau_j + \alpha_a$$

where

μ = overall mean on logit scale,

τ = “time” (sample) effect, constraint (e.g.) $\sum \tau_j = 0$,

α = “age” (residence time) effect, constraint (e.g.) $\sum \alpha_a = 0$.

Similarly, model capture probabilities p_{aj} as logit-additive.

Models so far

Label models, e.g. $\phi(a)p(t)$ for age effects in survival and time effects in capture.

c = constant t = time (sample) effect

a = age effect ta = time + age effect (additive on logit scale)

		p			
		c	t	a	ta
ϕ	c	$\phi(c)p(c)^*$	$\phi(c)p(t)^*$	$\phi(c)p(a)$	$\phi(c)p(ta)$
	t	$\phi(t)p(c)^*$	$\phi(t)p(t)^*$	$\phi(t)p(a)$	$\phi(t)p(ta)$
	a	$\phi(a)p(c)$	$\phi(a)p(t)$	$\phi(a)p(a)$	$\phi(a)p(ta)$
	ta	$\phi(ta)p(c)$	$\phi(ta)p(t)$	$\phi(ta)p(a)$	$\phi(ta)p(ta)^\dagger$

* No age effects, so can have unequally spaced samples

† Also unfittable, too many parameters.

Back to the Example (at last)



Practical Session 2

Return to the cormorants, and run the analyses for all 15 models.
Use the commands file **corm.R** Interpret the best model.

	maxLL	RD	AIC	relAIC	AICc	relAICc	npar
phic.pc	-318.71	637.41	659.41	311.52	659.94	309.42	11
phic.pt	-198.57	397.15	433.15	85.26	434.54	84.02	18
phic.pa	-210.41	420.81	456.81	108.93	458.20	107.69	18
phic.pta	-177.75	355.50	405.50	57.62	408.18	57.67	25
phit.pc	-158.51	317.02	353.02	5.13	354.40	3.89	18
phit.pt	-153.40	306.80	354.80	6.92	357.27	6.75	24
phit.pa	-154.40	308.80	356.80	8.91	359.26	8.75	24
phit.pta	-147.16	294.33	356.33	8.44	360.46	9.95	31
phia.pc	-183.09	366.19	400.19	52.30	401.43	50.91	17
phia.pt	-169.34	338.68	386.68	38.80	389.15	38.63	24
phia.pa	-173.67	347.33	391.33	43.45	393.40	42.89	22
phia.pta	-168.52	337.05	397.05	49.16	400.91	50.40	30
phita.pc	-150.03	300.05	348.05	0.16	350.51	0.00	24
phita.pt	-143.94	287.89	347.89	0.00	351.75	1.24	30
phita.pa	-145.60	291.20	349.20	1.31	352.81	2.30	29
phiW.pc	-187.94	375.89	399.89	52.00	400.52	50.00	12
phiW.pt	-171.25	342.49	380.49	32.60	382.04	31.52	19
phiW.pa	-174.59	349.17	385.17	37.29	386.56	36.05	18
phiW.pta	-150.10	300.21	352.21	4.32	355.10	4.59	26

How to estimate stopover duration?

Recall the likelihood for capture history i :

$$L_i = \sum_{b=1}^{f_i} \sum_{d=\ell_i}^K \left[\beta_{b-1} \left(\prod_{j=b}^{d-1} \phi_{bj} \right) (1 - \phi_{bd}) \right] \left[\prod_{j=b}^d p_{bj}^{x_{ij}} (1 - p_{bj})^{1-x_{ij}} \right].$$

This may be seen as

$$L_i = \sum_{b=1}^{f_i} \sum_{d=\ell_i}^K [Prob(PH = (b, d))] \times [Prob(CH_i | PH = (b, d))]$$

where PH = presence history, specified by (b, d) , the arrival and departure times, and $CH | PH$ = probability of this capture history given this presence history. (*Zero probability if $f_i < b$ or $\ell_i > d$.*)

Stopover duration for one bird

For each bird, we know the capture history but not the full presence history.

Bayes' theorem gives the probability of each feasible presence history, given the capture history, $Prob(PH = (b, d) \mid CH_i)$.

Hence the expected stop-over duration for bird i is:

$$E_i(dur) = E_i(d - b + 1) = \sum_{b=1}^{f_i} \sum_{d=l_i}^K (d - b + 1) \times P(PH = (b, d) \mid CH_i)$$

Similarly the variance is

$$V_i(dur) = E_i [(d - b + 1)^2] - [E_i(d - b + 1)]^2.$$

Stopover duration over all birds

Averaging over the N birds (assuming $N = \hat{N}$), the expected stopover duration is

$$E(dur) = \frac{1}{N} \left[\sum_{i=1}^D E_i + (N - D)E_0 \right]$$

and assuming independence of birds, the variance is

$$V(dur) = \frac{1}{N^2} \left[\sum_{i=1}^D V_i + (N - D)^2 V_0 \right].$$

Time-dependent survival, $\phi(t)$

(i) We have used time as a factor: $\phi_{aj} = \phi_j$ ($j = 1, 2, \dots, K - 1$). Survival depends only on the particular sample, a calendar time effect (e.g. weather related survival).

This model for survival is well established in the capture-recapture literature, as part of the CJS and JS models.

(ii) To reduce the number of parameters, we could, for example, assume a trend over time: $\phi_{aj} = \phi_j$ with

$$\log \left(\frac{\phi_j}{1 - \phi_j} \right) = \alpha + \gamma t_j$$

where t_j = calendar time at sample j . A quadratic curve could be used, if a maximum or minimum may be present partway through the study.

Age dependent survival, $\phi(a)$

- (i) We have used age as a factor: $\phi_{aj} = \phi_a$ ($a = 1, 2, \dots, K-1$).
- (ii) To reduce the number of parameters and see broad trends, assume survival is related to age using some line or curve.

A model with

$$\log \left(\frac{\phi_a}{1 - \phi_a} \right) = \alpha + \gamma t_j$$

has a steady trend, or we could use a standard survival curve.

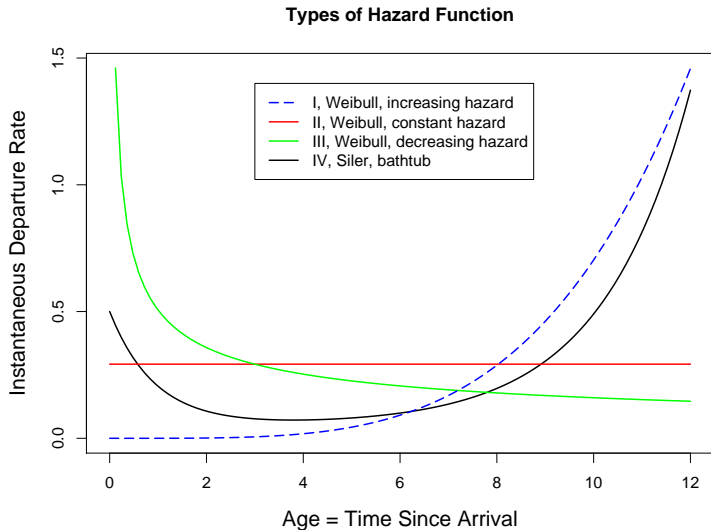
Continuous Time

We've modelled age as a factor, one parameter per sample, in discrete time.

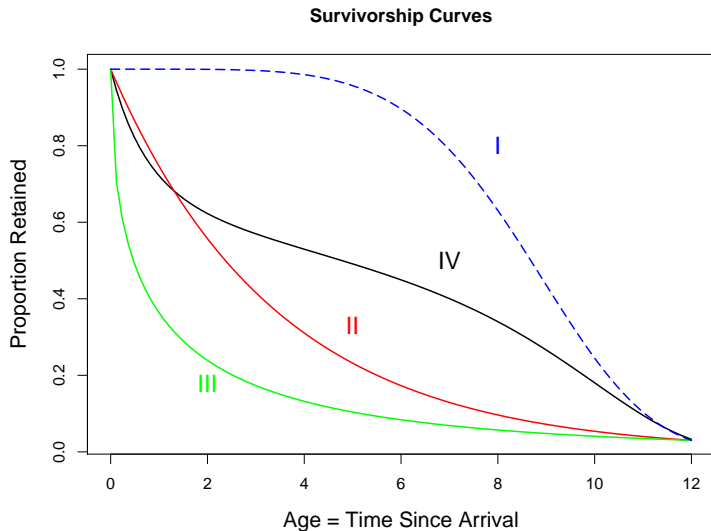
However, continuous-time survival models use fewer parameters and show any general trends or patterns over time.

There are four commonly-used shapes of survival (retention) curve, depending on four types of hazard curve (instantaneous departure rate).

Hazard Functions



Survivorship Curves



Mathematical Connections

X = random variable for lifetime (stopover duration).

Survival Curve:

$$S(x) = \text{Prob}(X \geq x) = 1 - F(x)$$

Density Function:

$$f(x) = \frac{d F(x)}{dx}$$

Hazard rate:

$$h(x) = \frac{f(x)}{S(x)}$$

Discrete survival probability from Sample j to Sample $j + 1$:

$$\phi_{aj} = \frac{S(x_{j+1})}{S(x_j)}$$

where x_j = age at sample j .

Types of Survival Curves

Type II Curve, constant survival probability

Exponential survival time, constant hazard. Special case of Weibull with shape parameter = 1.

Type I Curve, Weibull, shape parameter > 1

Ageing model - the hazard function increases with age. Useful for populations with senescence (animals tend to have similar lifetimes), or for stopovers where departure is driven by length of stay so far (birds tend to have similar stopover times).

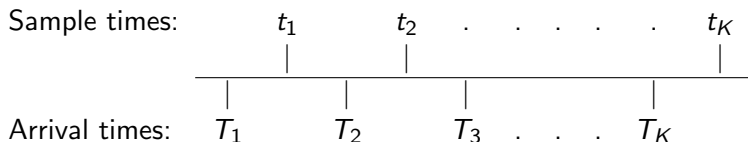
Type IV Curve, Bathtub

Higher hazard at beginning and end of life. For stopover data, may pick up difference between transients (high hazard when newly-arrived) and residents (more like J-shaped Type I hazard). (Five or six parameters, though.)

Continuous time assumptions

Assume birds in arrival cohort b arrived halfway between samples $b-1$ and b .

(Sample 0 was before any birds arrived.)



For arrival cohort b , age at time $t = t - T_b$

Weibull models for the cormorants - anything new?

Earlier table - $\phi(W)$ is Weibull survival.

	maxLL	RD	AIC	relAIC	AICc	relAICc	npars
phic.pc	-318.71	637.41	659.41	311.52	659.94	309.42	11
phic.pt	-198.57	397.15	433.15	85.26	434.54	84.02	18
phic.pa	-210.41	420.81	456.81	108.93	458.20	107.69	18
phic.pta	-177.75	355.50	405.50	57.62	408.18	57.67	25
phit.pc	-158.51	317.02	353.02	5.13	354.40	3.89	18
phit.pt	-153.40	306.80	354.80	6.92	357.27	6.75	24
phit.pa	-154.40	308.80	356.80	8.91	359.26	8.75	24
phit.pta	-147.16	294.33	356.33	8.44	360.46	9.95	31
phia.pc	-183.09	366.19	400.19	52.30	401.43	50.91	17
phia.pt	-169.34	338.68	386.68	38.80	389.15	38.63	24
phia.pa	-173.67	347.33	391.33	43.45	393.40	42.89	22
phia.pta	-168.52	337.05	397.05	49.16	400.91	50.40	30
phita.pc	-150.03	300.05	348.05	0.16	350.51	0.00	24
phita.pt	-143.94	287.89	347.89	0.00	351.75	1.24	30
phita.pa	-145.60	291.20	349.20	1.31	352.81	2.30	29
phiW.pc	-187.94	375.89	399.89	52.00	400.52	50.00	12
phiW.pt	-171.25	342.49	380.49	32.60	382.04	31.52	19
phiW.pa	-174.59	349.17	385.17	37.29	386.56	36.05	18
phiW.pta	-150.10	300.21	352.21	4.32	355.10	4.59	26

For this data set, the $\phi(a)$ models were not chosen, as there seemed to be both age and time variation in the survival rates. Hence the Weibull model, by ignoring those time effects, is also not chosen.

Estimating Mean Stopover Duration

If a fitted curve, e.g. Weibull, is chosen for the model of survival, we may use the mean and variance of this fitted distribution.

If the estimated Weibull parameters are a = shape parameter and b = scale parameter, the mean is

$$\mu = E(X) = b \Gamma \left(1 + \frac{1}{a} \right)$$

and the variance is

$$\sigma^2 = Var(X) = b^2 \Gamma \left(1 + \frac{2}{a} \right),$$

where Γ is the gamma function.

Practical Session 3

Continue with the cormorant data. Do the last plot and calculations in **corm.R**, using Model phiW.pta.

Comments on age-structured models

- It is advisable to start sampling before arrivals, and continue until after departures. We can still do the analysis using only part of the time when the birds are occupying the area, but the age-structured models carry an implicit assumption that the birds first present at sample 1 have arrived recently, and the birds still present at sample K are about to depart. This second assumption is also implicit in the CJS model, where it is assumed that $\phi(K) = 0$.
- We need more data than for JSSA analysis, because of estimating arrival times.
- We have allowed for a certain type of heterogeneity of survival by including the age effects in the models. However, we did not allow for individual heterogeneity of survival, with some animals intrinsically more likely to survive than others. Also, if there is **individual heterogeneity of capture** probability, models which fail to allow for this will underestimate N .

4. Models with Heterogeneity

See Pledger *et al.* 2010.

We now extend the JSSA model to cover the case of individual heterogeneity of capture and survival probabilities.

The heterogeneity of survival now is not assumed to be related to duration of stay, as with the age-structured models.

We now suppose both types of heterogeneity (survival and capture probabilities) are due to unknown causes, and we wish to allow for this in a general way which is not tied to any particular covariates or other extra information (e.g. age, weight of animal, location of home range in relation to trap placement).

Why model this heterogeneity?

The major reason is to reduce the bias in abundance estimates. Unmodelled heterogeneity of capture probability gives underestimates of population size, whether the superpopulation N or the JS sample-specific population sizes N_j , $j = 1, 2, \dots, K$.

There is also some bias in survival estimates caused by unmodelled heterogeneity of capture probability, but this is not a very severe problem.

Another source of bias is caused by estimating survival using only data from the animals actually caught. Much work using CJS models takes this approach. However, if there is individual heterogeneity of survival, the short-lived animals have a lower probability of at least one capture, so there is an overestimation of survival probabilities because these short-lived animals are systematically ignored. A model combining information from the uncaught animals circumvents this problem.

How is it done?

We now introduce finite mixture models, where the animals are assumed to come from different classes, although we don't know which classes they are from. These are hidden or latent classes.

There are C animal classes with membership unknown (a latent effect); each animal comes independently from class c with probability π_c ($\sum \pi_c = 1$).

An animal from class c , if present at sample j , has probability p_{jc} of capture in sample j , and probability ϕ_{jc} of survival to the next sample (with ϕ_{Kc} assumed to be zero).

Data and parameters

As before, there are D distinct animals seen, with $x_{ij} = 1$ if animal i is caught in sample j , otherwise $x_{ij} = 0$. Animal i 's capture history, CH_i , is the row vector \mathbf{x}_i .

The $N - D$ uncaught animals each have capture history $\text{CH}_0 = \mathbf{0}$, a K -vector of zeros.

There are n_h animals with capture history h , $h = 1, 2, \dots, H$. We use the data in grouped form: matrix X is $H \times K$, with rows all different, and vector y gives the frequency for each row.

Likelihoods

Writing the parameters N , β , π , p and ϕ as a parameter vector θ , the **full likelihood** is

$$L(\theta \mid \text{data}) = \frac{N!}{(N-D)! \prod_h n_h!} \times \prod_{i=1}^D L_i \times L_0^{N-D}$$

where h indexes the different observed capture histories.

The individual likelihoods are found by summing over the classes and all feasible birth and death times (b and d respectively, samples when first and last available for capture):

$$L_i = P(\text{CH}_i) = \sum_{b=1}^{f_i} \sum_{d=\ell_i}^K \sum_{c=1}^C \left[\pi_c \beta_{b-1} \left(\prod_{j=b}^{d-1} \phi_{jc} \right) (1 - \phi_{dc}) \left\{ \prod_{j=b}^d p_{jc}^{x_{ij}} (1 - p_{jc})^{1-x_{ij}} \right\} \right],$$

(where the empty product $\prod_{j=b}^{d-1} \phi_{jc} = 1$ if $b = d$).

Yet another likelihood

Similarly for an uncaught animal,

$$L_0 = P(CH_0) = \sum_{b=1}^K \sum_{d=b}^K \sum_{c=1}^C \left[\pi_c \beta_{b-1} \left(\prod_{j=b}^{d-1} \phi_{jc} \right) (1 - \phi_{dc}) \left\{ \prod_{j=b}^d (1 - p_{jc}) \right\} \right].$$

Reparameterise!

For computational effectiveness and the provision of more appropriate confidence intervals, we **reparameterise** the model, expressing the parameters ϕ_{jc} and p_{jc} on a logit scale, and N on a log scale.

Using τ for time and η for heterogeneity,

$$\text{logit}(p_{jc}) = \log\left(\frac{p_{jc}}{1 - p_{jc}}\right) = \mu + \tau_j + \eta_c + (\tau\eta)_{jc} \quad (1)$$

with constraints $\sum \tau_j = 0$, $\sum \eta_c = 0$, and each row and column of $(\tau\eta)_{jc}$ adding to 0. Similarly $\text{logit}(\phi_{jc})$ may be modelled with main effects and interaction.

Remove the interactions

Models with interactive time or heterogeneity effects may be more complicated than necessary, and have too many parameters for successful model fitting. A simpler model for p_{jc} with additive effects of time and class has $\text{logit}(p_{jc}) = \mu + \tau_j + \eta_c$ with constraints $\sum \tau_j = 0$ and $\sum \eta_c = 0$.

Similarly survival could be modelled as additive on the logit scale.

Further simplifications have capture and/or survival probabilities dependent only on time (e.g. $\text{logit}(p_{jc}) = \mu + \tau_j$), only on individual heterogeneity ($\text{logit}(p_{jc}) = \mu + \eta_c$), or constant over time and animals ($\text{logit}(p_{jc}) = \mu$), with similar simplified versions of ϕ_{jc} .

Possum Example

Murray Efford provided data from live trapping of the Australian brushtail possum (*Trichosurus vulpecula* Kerr) on a study grid in the Orongorongo Valley, near Wellington, New Zealand. There were $D = 270$ animals from $K = 9$ samples taken in February 1980-1988. Table of relative AIC from Pledger *et al.* 2010:

ϕ model	$p(.)$	$p(t)$	$p(h_2)$	$p(t+h_2)$
$\phi(.)$	39.0	34.9	13.1	4.1
$\phi(t)$	29.4	37.6	0.0	2.6
$\phi(h_2)$	43.0	30.2	15.0	5.7
$\phi(t+h_2)$	33.4	41.6	1.6	4.2

Note the dichotomy of AIC values between models with and without heterogeneity of capture.

Selection of the model $\{\phi(t), p(h_2)\}$ accords well with other knowledge of this population. There is spatially-induced heterogeneity of capture, as the traps are always set in the same location on a grid, and possums with a home range including a

Practical Session 4

Run the possum example using the file possums9.R.

Consider the different models, and their N estimates.

Find and interpret the best model.

Find the places where **hetage** is giving higher likelihoods than those found in Pledger *et al.* 2010 (where the model fitting was much slower, and laborious searches were done by hand).

Incorporating Extra Information

It is usual to allow for losses on capture (Pradel 1996, Schwarz and Arnason 1996), where some animals are known to die or are removed at sample j .

We see such “deaths” as just one of four processes which are easily incorporated into our individual-based models. These are natural or unnatural death (e.g. death on capture, or removal), and natural or unnatural arrival (e.g. translocation from another population).

No extra parameters (e.g. probability of loss on capture) are needed, as we simply modify the likelihood equation for that animal.

Is there much use for this development?

5. Combining Heterogeneity and Age Structure

There is now an obvious development, the combination of age-structure for survival probabilities with heterogeneity of capture probabilities.

This work is not yet published, but we have some working code.

The likelihoods still have summation over b , d and c (for the mixture classes or groups), but the ϕ parameters are subscripted by age and/or time, while the p parameters are subscripted by heterogeneity and/or time.

Practical Session 5

See the end of the file **possums9.R** for adding four new models, with age structure and heterogeneity combined in two of them.

	maxLL	RD	AIC	relAIC	AICc	relAICc	npar
phic.pc	-238.75	477.51	499.51	38.95	500.04	37.77	11
phic.pt	-229.73	459.47	495.47	34.91	496.85	34.59	18
phic.ph	-223.83	447.67	473.67	13.11	474.40	12.13	13
phic.pth	-212.32	424.63	464.63	4.08	466.34	4.08	20
phit.pc	-226.98	453.95	489.95	29.40	491.34	29.07	18
phit.pt	-225.06	450.12	498.12	37.57	500.58	38.32	24
phit.ph	-210.28	420.55	460.55	0.00	462.26	0.00	20
phit.pth	-205.55	411.11	463.11	2.55	466.00	3.74	26
phih.pc	-238.75	477.51	503.51	42.95	504.24	41.97	13
phih.pt	-225.40	450.79	490.79	30.24	492.50	30.24	20
phih.ph	-223.78	447.57	475.57	15.02	476.42	14.15	14
phih.pth	-212.14	424.29	466.29	5.73	468.17	5.91	21
phith.pc	-226.98	453.95	493.95	33.40	495.67	33.40	20
phith.pt	-220.90	441.80	493.80	33.25	496.69	34.43	26
phith.ph	-210.10	420.20	462.20	1.64	464.08	1.82	21
phith.pth	-205.39	410.78	464.78	4.22	467.90	5.64	27
phia.pc	-224.60	449.21	485.21	24.66	486.60	24.33	18
phia.pt	-218.98	437.96	487.96	27.40	490.63	28.37	25
phia.ph	-210.56	421.12	461.12	0.56	462.83	0.56	20
phia.pth	-205.94	411.88	463.88	3.33	466.78	4.51	26

Comments on the combined models

With the nine-year possum data set, the model (phia, ph) is almost the best. With a longer time span, which better covers all the possums' lifetimes, this becomes a better model, with $(\text{phita}, \text{ph})$ taking over with more data.

The age-structured models need to have long-term data sets (compared with the lifetime of the animal).

6. User-defined functions

The file **newts.R** has an example of a user-defined model. It goes with the newt example to follow.

It is necessary to construct the wrapping and unwrapping files for the particular model.

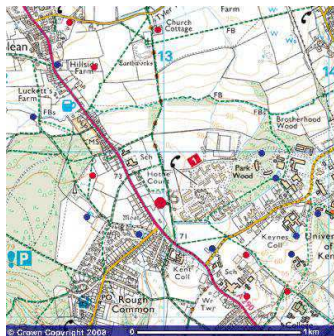
The wrapping file, “v.to.f”, takes the vector of independent parameters and wraps it up into a full list of parameters suitable for feeding to the minus likelihood function mLL.

The unwrapping file, “start”, takes a full list of parameters and turns it into the minimal set of independent parameters needed for the model.

Great crested newts (*Triturus cristatus*)

A collaboration with the *Durrell Institute of Conservation and Ecology* at the University of Kent.

Data on great crested newts (*Triturus cristatus*) have been collected every year since 2002 in a small study area in Canterbury, UK.



The area consists of 4 small ponds which were artificially created in 1998 and were left to be colonised naturally.



Each pond is 2m by 1m with the deepest end being approximately 0.7m into the ground.

The newts use the ponds during the breeding season, March-July.

Samples are taken every Friday morning using bottle traps.

The newts are individually identified by the unique patterns on their bellies.



The sample sizes every year are very small, varying between 10 and 17 newts.

However, individuals stay for a number of weeks at the ponds (about 20 samples are taken each year) and almost all of the newts in the sample are caught several times each year.

Because of the very small number of newts colonising the ponds, the biologists have given them individual names instead of individual numbers.

The biologists coordinating the study believe that they capture all of the newts that visit the ponds during the breeding season each year. Is this true?

2002 data set

Arnie	0	1	0	0	0	1	0	0	1	1	0	0	0	0	0	0	0	0	0	0	0	4
Brad	0	1	0	1	0	0	0	1	1	1	1	1	0	0	1	0	0	0	0	0	0	8
Bruce	0	0	1	0	0	0	0	0	1	0	1	0	0	0	0	0	1	0	0	0	0	4
Clint	1	1	1	1	1	1	1	1	1	1	0	0	1	1	1	1	1	0	0	0	0	15
Dustin	0	0	1	1	0	0	1	0	0	1	0	0	0	0	0	0	0	0	0	0	0	4
Mr T	0	1	0	0	1	1	1	1	0	1	0	0	0	1	0	0	1	1	0	0	0	9
Hugh	0	0	0	1	1	0	1	1	0	1	1	0	0	1	0	0	0	0	0	0	0	7
John	1	0	0	0	0	1	0	1	1	0	0	1	1	0	0	1	1	0	0	0	0	8
Leonardo	1	1	1	0	1	1	0	0	0	1	0	1	1	0	0	0	0	0	0	0	0	8
Nicholas	0	1	1	1	0	0	0	1	0	0	1	1	0	0	0	0	0	0	0	0	0	6
Sean	1	0	1	0	0	0	0	0	1	1	0	0	1	1	0	0	1	0	0	0	0	7
Vin	1	0	0	0	0	0	0	0	0	0	0	0	0	1	0	0	0	0	1	0	0	3
Gwyneth	0	0	0	1	1	0	1	1	0	1	0	1	1	0	0	1	0	0	0	0	0	8
Julia	1	1	1	1	1	0	1	1	1	1	0	1	1	1	0	1	0	0	0	0	0	13
Marilyn	0	0	0	1	0	1	0	1	1	1	1	1	0	0	1	0	0	0	0	0	0	9
Patricia	0	0	0	0	0	1	0	0	0	1	1	0	0	0	0	0	0	0	0	0	0	3
Robin	0	0	0	0	0	1	0	1	0	0	1	1	1	0	1	1	1	0	0	0	0	8

The number of captures varies between individuals. The differences could be due to the different stop-over duration of individual newts but it could also be a sign of individual heterogeneity in capture probability.

Not accounting for heterogeneity in capture probability is known to bias significantly the population size estimate.

In this case it could mean that individuals that are caught several times are more active and move around the pond more, therefore they are more likely to get caught and that there might be a number of newts that avoid the bottle traps on every sampling occasion.

The likelihood

$$L(N, \boldsymbol{\beta}, \boldsymbol{\phi}, \boldsymbol{p}, \boldsymbol{\pi} | \boldsymbol{X}) = \frac{N!}{\prod_h n_h! (N-D)!} \times$$

$$\prod_{i=1}^D \left[\sum_{g=1}^m \pi_g \left\{ \sum_{b=1}^{f_i} \sum_{d=l_i}^K \left(\beta_{b-1} \left(\prod_{j=b}^{d-1} \phi_{ja} \right) (1 - \phi_{da}) \left(\prod_{j=b}^d p_{gja}^{x_{ij}} (1 - p_{gja})^{1-x_{ij}} \right) \right) \right\} \right]$$

$$\times \left[\sum_{g=1}^m \pi_g \left\{ \sum_{b=1}^K \sum_{d=b}^K \left(\beta_{b-1} \left(\prod_{j=b}^{d-1} \phi_{ja} \right) (1 - \phi_{da}) \left(\prod_{j=b}^d (1 - p_{gja}) \right) \right) \right\} \right]^{N-D},$$

where now, π_g is the probability that an individual belongs to group g , $\sum_{g=1}^m \pi_g = 1$, and p_{gja} represents the probability of capturing an individual that belongs to group g which is alive at sample j and has been at the study area for a samples.

Constraints for the parameters

The β probabilities are generated using $\beta_{j-1} = a_{\beta}j + b_{\beta}$, $j = 1 \dots K$ and then transformed into a vector of length $K - 1$ with elements β_j^* where,

$$\beta_{j-1}^* = \frac{\beta_{j-1}}{\sum_{k=j}^K \beta_{k-1}}, j = 1 \dots K - 1,$$

in order for their sum to be equal to 1.

Correspondingly, $\phi_j = \frac{\exp(a_{\phi}j + b_{\phi})}{1 + \exp(a_{\phi}j + b_{\phi})}$, $j = 1 \dots K - 1$, to ensure that the ϕ estimates are always between 0 and 1.

Capture probability is assumed constant within each year. Only the cases of $m = 1$ and $m = 2$ are examined.

Model selection is performed using AIC within the models considered from the two different likelihoods. Because of violation of regularity conditions, comparisons between the homogeneous and heterogeneous models using traditional model-selection criteria are not meaningful.

In these cases (when comparing between models with $m = 1$ and $m = 2$), the LRT statistic will be distributed, under the null hypothesis of homogeneous capture probability, according to a 50 : 50 mixture of 0s and the χ^2_1 distribution (Self and Liang (1987)).

Because of issues of multimodality of likelihood surfaces when using finite mixtures, 100 random starting values for parameters p and π are generated for the heterogeneous model and the best fit is selected in terms of the maximised likelihood value.

AIC values

Model		λ	2002	2003	2004	2005	2006	2007
Homogeneous	$\beta_c \phi_c p_c$	3	91.828	37.885	33.452	35.410	47.218	24.415
	$\beta_c \phi_l p_c$	4	52.191	13.095	22.714	22.458	18.480	12.469
	$\beta_c \phi_{al} p_c$	4	62.452	18.786	26.880	26.388	34.192	6.453
	$\beta_l \phi_c p_c$	5	38.259	24.235	10.607	12.668	27.823	17.745
	$\beta_l \phi_l p_c$	6	0.000	1.012	0.000	0.000	0.000	6.092
	$\beta_l \phi_{al} p_c$	6	1.164	0.000	3.243	1.264	13.796	0.000
Heterogeneous	$\beta_c \phi_c p_{2g}$	5	94.453	38.839	34.049	33.457	47.218	24.415
	$\beta_c \phi_l p_{2g}$	6	54.375	14.076	23.895	20.911	18.480	12.469
	$\beta_c \phi_{al} p_{2g}$	6	64.920	19.040	27.495	25.137	34.192	6.453
	$\beta_l \phi_c p_{2g}$	7	38.860	25.178	9.937	12.265	27.823	17.745
	$\beta_l \phi_l p_{2g}$	8	0.000	1.992	0.000	0.000	0.000	6.092
	$\beta_l \phi_{al} p_{2g}$	8	0.148	0.000	1.457	2.112	13.796	0.000

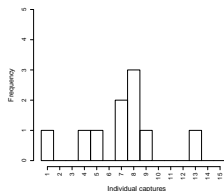
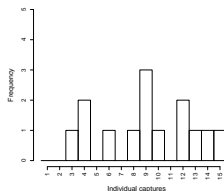
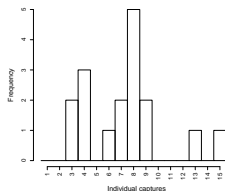
Year	Model	$\max(\log L)$	LRT	p-value
2002	$\beta_I \phi_I p_c$ $\beta_I \phi_I p_{2g}$	-169.802 -164.214	11.176	0.0004
2003	$\beta_I \phi_{aI} p_c$ $\beta_I \phi_{aI} p_{2g}$	-88.499 -88.009	0.980	0.1611
2004	$\beta_I \phi_I p_c$ $\beta_I \phi_I p_{2g}$	-166.641 -164.326	4.631	0.0157
2005	$\beta_I \phi_I p_c$ $\beta_I \phi_I p_{2g}$	-97.728 -93.839	7.777	0.0026
2006	$\beta_I \phi_I p_c$ $\beta_I \phi_I p_{2g}$	-76.256 -76.256	0	1
2007	$\beta_I \phi_{aI} p_c$ $\beta_I \phi_{aI} p_{2g}$	-111.992 -111.992	0	1

Estimation of N

Model		2002	2003	2004	2005	2006	2007
Homogeneous	$\beta_c \phi_c p_c$	18.44	11.89	14.34	10.45	13.24	10.16
	$\beta_c \phi_l p_c$	19.85	12.47	14.45	10.94	14.32	10.41
	$\beta_c \phi_{al} p_c$	17.33	11.29	14	10	13	10
	$\beta_l \phi_c p_c$	17.45	11	14	10.04	13	10
	$\beta_l \phi_l p_c$	17	11	14	10	13	10
	$\beta_l \phi_{al} p_c$	17	11	14	10	13	10
Heterogeneous	$\beta_c \phi_c p_{2g}$	18.92	11.93	14.45	10.62	13.24	10.16
	$\beta_c \phi_l p_{2g}$	20.23	12.47	14.76	11.11	14.32	10.41
	$\beta_c \phi_{al} p_{2g}$	17.48	11.34	14.16	10	13	10
	$\beta_l \phi_c p_{2g}$	17.69	11	14	10.08	13	10
	$\beta_l \phi_l p_{2g}$	17	11	14	10	13	10
	$\beta_l \phi_{al} p_{2g}$	17	11	14	10	13	10

Estimation of p

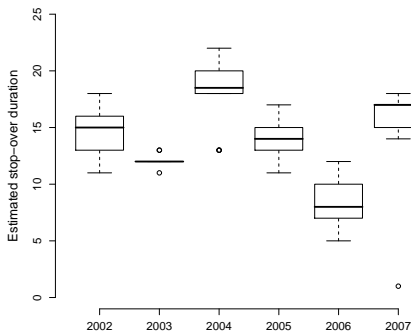
Year	2002	2003	2004	2005	2006	2007
p_1	0.835	0.434	0.604	1.000	0.718	0.455
p_2	0.392	—	0.238	0.457	—	—
π	0.118	—	0.822	0.100	—	—
$\pi \cdot p_1 + (1 - \pi) \cdot p_2$	0.444	—	0.539	0.511	—	—



Expected stop-over duration

Can calculate expected stop-over duration for newt i as:

$$\sum_{b=1}^{f_i} \sum_{d=l_i}^K (d - b + 1) \times P(\{b, d\} \mid CH_i)$$



5. References

- Jolly, G. M. 1965. Explicit estimates from capture-recapture data with both death and immigration - Stochastic model. *Biometrika* **52**, 225-247.
- Pledger, S., Efford, M.G., Pollock, K.H., Collazo, J. and Lyons, J. (2009). Stopover duration analysis with departure probability dependent on unknown time since arrival. Pages 349-363 in D.L. Thomson, E.G. Cooch & M.J. Conroy (eds), *Modeling Demographic Processes in Marked Populations*, Environmental and Ecological Statistics 3, New York: Springer.
- Pledger, S., Pollock, K.H. & Norris, J.L. (2010). Open CaptureRecapture Models with Heterogeneity: II. JollySeber Model. In press, *Biometrics*. Published Online Early: Nov 2009. DOI: 10.1111/j.1541-0420.2009.01361.x
- Seber, G. A. F. 1965. A note on the multiple-recapture census. *Biometrika* **52**, 249-259.
- Schwarz, C.J. & Arnason, A.N. (1996). A general methodology for the analysis of capture-recapture experiments in open populations. *Biometrics* **52**, 860-873.