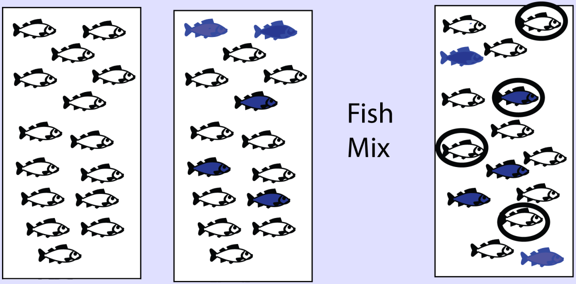
Partial Stratification in Two-Sample Capture-Recapture Experiments

Lasantha Premarathna
Carl James Schwarz

Department of Statistics and Actuarial Science,
Simon Fraser University,
Burnaby, BC.

June 24, 2014

Basic capture-recapture



N ?

Capture and Mark n_1

$n_1 = 100$

Capture n_2 of which m_2 are marked

$n_2 = 200$

$m_2 = 10$

$$\hat{N} = \frac{n_1 n_2}{m_2} = \frac{100 \times 200}{10} = 2000$$

Standard Lincoln-Petersen - Assumptions

- The population is **closed** (geographically and demographically).
- Mark status is correctly identified at each sampling occasion
- Marks are not lost between sampling occasions
- Capture and marking does not affect subsequent catchability of an animal
- Sampling is random
- **Homogeneous capture probabilities.**

$$\widehat{N} = \frac{n_1 \times n_2}{m_2} \approx \frac{Np_1 \times Np_2}{Np_1 p_2} \approx N$$

Standard Lincoln-Petersen - Need for stratification

Animals may vary in catchability, e.g. by sex.

Sex	N	p_1	p_2	n_1	n_2	m_2	$\widehat{N}$
M	1200	.10	.05	120	60	6	
F	800	.05	.10	40	80	4	
Total	2000			160	140	10	2240

Bias can go in either direction depending on correlation of catchabilities.

Standard Lincoln-Petersen - Need for stratification

Animals may vary in catchability, e.g. by sex.

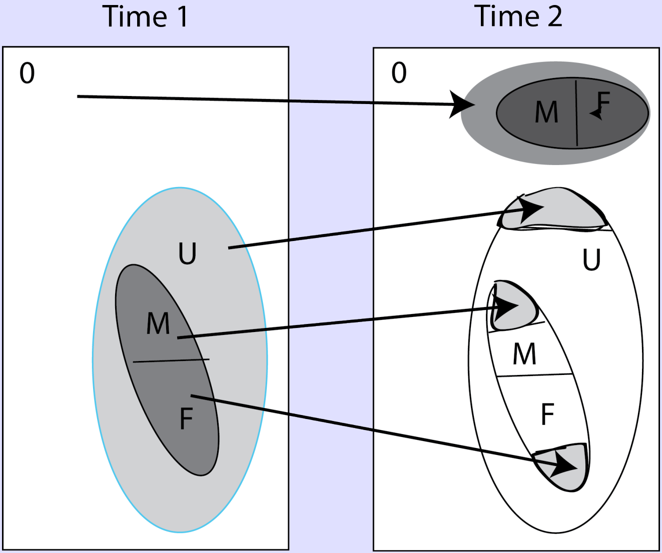
Stratify and compute separate estimates for each stratum.

Sex	N	p_1	p_2	n_1	n_2	m_2	$\widehat{N}$
M	1200	.10	.05	120	60	6	1200
F	800	.05	.10	40	80	4	800
Total	2000						2000

Assumes that you can classify animals into strata in ALL sampling occasions.

Partial Stratification

Sometimes difficult/costly to stratify when captured.
Select a sub-sample of each sample to stratify.



Partial Stratification

$3C + 4$ possible capture histories.

History	Probability
MM	$\lambda_M \; p_{1M} \; \theta_1 \; p_{2M}$
M0	$\lambda_M \; p_{1M} \; \theta_1 \; (1 - p_{2M})$
0M	$\lambda_M \; (1 - p_{1M}) \; p_{2M} \; \theta_2$
...	Similarly for other categories
UU	$\lambda_M \; p_{1M} \; (1 - \theta_1) \; p_{2M} + \lambda_F \; p_{1F} \; (1 - \theta_1) \; p_{2F}$
U0	$\lambda_M \; p_{1M} \; (1 - \theta_1) \; (1 - p_{2M}) + \lambda_F \; p_{1F} \; (1 - \theta_1) \; (1 - p_{2F})$
0U	$\lambda_M \; (1 - p_{1M}) \; p_{2M} \; (1 - \theta_2) + \lambda_F \; (1 - p_{1F}) \; p_{2F} \; (1 - \theta_2)$
00	Everything else (not observable)

Where

- p_{tC} = probability of capture of category C at time t.
- λ_C = proportion of category C in the population.
- θ_t = proportion of sample at time t that is "categorized".

Partial Stratification - Likelihood

Multinomial distribution with unknown index

$$L = \frac{N!}{n_{U0}! \quad n_{UU}! \quad n_{0U}! \ldots \quad (N-n)!} \quad \times$$

$$(P_{U0})^{n_{U0}} (P_{UU})^{n_{UU}} (P_{0U})^{n_{0U}} \quad \times$$

$$\prod_C (P_{C0})^{n_{C0}} \times \prod_C (P_{CC})^{n_{CC}} \times \prod_C (P_{0C})^{n_{0C}} \times (P_{00})^{N-n}$$

No closed form solution - standard numerical methods used.

Parameters can be constrained by using the design matrix and offset vectors

$$logit(p_{tk}) = X\beta + offset$$

Example, a model with $p_{1M} = p_{1F}$, $p_{2M} = p_{2F}$ uses design matrices and offsets

$$\begin{bmatrix} 1 & 0 \\ 1 & 0 \\ 0 & 1 \\ 0 & 1 \end{bmatrix} \text{ and } \begin{bmatrix} 0 \\ 0 \\ 0 \\ 0 \end{bmatrix}$$

Use Akaike's information criterion (AIC) (Burnham and Anderson, 2002) to rank models.

Partial Stratification - Optimal Allocation

Use linear cost function

$$C = n_1 c_1 + n_1^* c_1^* + n_2 c_2 + n_2^* c_2^* \leq C_0$$

where

- n_1 is number of fish captured at first occasion;
- n_1^* is number of fish captured at first occasion that are "categorized". $n_1 \geq n_1^*$.
- n_2 is the number of fish captured at second occasion;
- n_2^* is the number of fish captured at second occasion that are "categorized". $n_2 \geq n_2^*$.

Partial Stratification - Example

Walleye Data-Mille Lacs, MN

- Walleye are captured on the spawning grounds. Almost all the fish can be sexed in the first occasion
- All the captured fish are tagged and released and the recapture occurred 3 weeks later using gill-nets
- From a sample of fish captured at second occasion that are not tagged, a random sample is selected and sexed

Capture History	statistics
$U0$	40
UU	1
$M0$	5067
MM	40
$F0$	1551
FF	33
$0M$	41
$0F$	237
$0U$	3075

Partial Stratification - Example

Model			np	$\hat{N}$ '000s	$s.e.(\hat{N})$ '000s	$\Delta AICc$
$p(C * t)$	$\theta(t)$	$\lambda(C)$	8	205	26	0.0
$p(C * t)$	$\theta(t)$	$\lambda(.)$	7	208	24	4.7
$p(t)$	$\theta(t)$	$\lambda(C)$	6	311	36	462.2
$p(C)$	$\theta(t)$	$\lambda(C)$	6	399	54	1561.5
$p(.)$	$\theta(t)$	$\lambda(C)$	5	348	39	1572.0
$p(C)$	$\theta(t)$	$\lambda(.)$	5	399	54	11613.4
$p(.)$	$\theta(.)$	$\lambda(.)$	3	348	39	13276.8

- p = probability of capture
- λ = proportion of category in the population.
- θ = proportion of sample at time that is "categorized".
- $C * t$ = varies by category and time;
- t = varies by time but not category;
- C = varies by category but not by time
- . = no variation by time or by category.

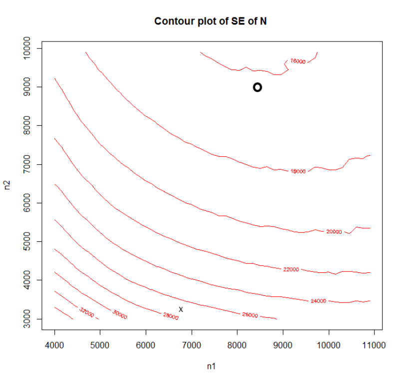
Partial Stratification - Example

Parameter	MLE	SE	SE if sex ratio known
λ_M	.33	.06	-
λ_F	.67	.06	-
θ_1	.993	.001	.001
θ_2	.082	.004	.004
N	205,000	26,000	25,000
N_M	68,000	13,000	8,000
N_F	137,000	23,000	17,000
$\hat{N}_{LP}$	311,000	36,000	

Example - optimal allocation

Let $c_1 = 2,\ c_2 = c_1/2,\ c_1^* = 4,\ c_2^* = c_1^*/2 = 4,\ C_0 = 40000$

$n_1 = 8199 \quad n_1^* = 3003 \quad n_2 = 9511 \quad n_2^* = 997$
$SE(\widehat{N}) = 16000$ a 30% reduction.
Conditional Contour plot and Graph for $SE(\widehat{N})$:



Partial Stratification - Summary

Problem

- Capture heterogeneity can cause bias in estimates in capture-recapture experiments
- Stratification may not be possible/is costly for all captured animals in each occasion

Solution

- Partial stratification
- Optimal allocation

Future work

- Bayesian solution for prior information on sex ratio
- Individual covariates such as length