



Bayesian integrated population modeling using JAGS

# Estimation of survival probabilities using capture-recapture data

# Estimation of survival



The principle to estimate survival is fairly easy:

- Follow individuals across time
- Count individuals at time  $t$  ( $C_t$ ) and assess how many of them are still alive after time  $\Delta t$  ( $L_{t+\Delta t}$ )
- Survival probability is then  $s_{\Delta t} = \frac{L_{t+\Delta t}}{C_t}$

But major problem: all individuals that are still alive, but **not seen** → classified **dead**

- Typically, we do not know  $L_{t+\Delta t}$ , but only  $p * L_{t+\Delta t}$ , where  $p$  is the detection probability
- Detection can be estimated, if we extend the experiment to at least one further occasion
- Thus, collect **capture-recapture data** (individual capture histories)

# Marking, recaptures and resightings



*Artificial marks*



*Natural marks*

*... also genetic or acoustic recapture*

# Individual capture history

ID	1982	1983	1984	1985	1986	1987	1988	1989	1990	1991	1992	1993	1994	1995	1996	1997	1998	1999	2000	2001	2002	2003	2004	2005	2006	2007
1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	0	0	0	0	0	1	1	0	0	0	0	0
2	1	1	1	1	1	0	1	1	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
3	1	1	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
4	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	0	0	0	1	1	1	0	0	0
5	0	1	1	1	1	1	1	1	1	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
6	0	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	0
7	0	1	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
8	0	0	1	1	1	1	1	1	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
9	0	0	1	1	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
10	0	0	0	1	1	1	1	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
11	0	0	0	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	0	0	0	0	0	0
12	0	0	0	0	0	1	1	1	1	1	1	1	1	1	0	0	0	1	0	0	0	0	0	0	0	0



# Individual capture history: result of two processes

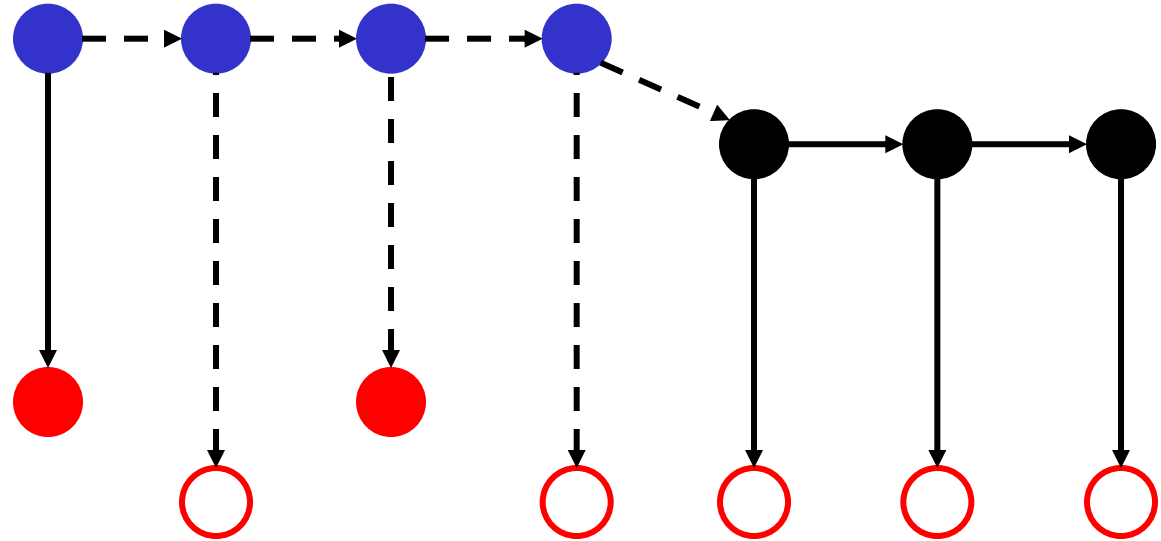
State process

Alive

Dead

Seen

Not seen



Observation process

- - - - -> Stochastic process
- > Deterministic process

# State-space likelihood to analyse CR data

## 1. Survival process

$$z_{i,first_i} = 1$$

$$z_{i,t} | z_{i,t-1} \sim \text{Bernoulli}(z_{i,t-1} \phi_{i,t-1})$$

where,

$z_{i,t}$ : matrix, indicating whether individual  $i$  is alive at time  $t$  ( $z = 1$ ), or dead ( $z = 0$ )

$\phi_{i,t}$ : apparent survival probability for individual  $i$  from time  $t$  to  $t+1$

## 2. Observation process

$$y_{i,t} | z_{i,t} \sim \text{Bernoulli}(z_{i,t} p_{i,t})$$

where,

$y_{i,t}$ : is the observed capture history for individual  $i$  at time  $t$

$p_{i,t}$ : recapture probability for individual  $i$  at time  $t$

## ***Data***

$$y_{i,t} = \begin{matrix} & 0 & 1 & 1 & 0 & 0 \\ & 1 & 0 & 0 & 0 & 0 \\ & 0 & 0 & 1 & 0 & 1 \\ & 0 & 1 & 0 & 0 & 0 \\ & \cdot & \cdot & \cdot & \cdot & \cdot \\ & \cdot & \cdot & \cdot & \cdot & \cdot \end{matrix}$$

## ***Parameters***

## ***Data***

$$y_{i,t} = \begin{matrix} & 0 & 1 & 1 & 0 & 0 \\ & 1 & 0 & 0 & 0 & 0 \\ & 0 & 0 & 1 & 0 & 1 \\ & 0 & 1 & 0 & 0 & 0 \\ & \cdot & \cdot & \cdot & \cdot & \cdot \\ & \cdot & \cdot & \cdot & \cdot & \cdot \end{matrix}$$

## ***Parameters***

$$\phi_{i,t} = \begin{matrix} & NA & \phi_{1,2} & \phi_{1,3} & \phi_{1,4} \\ \phi_{2,1} & \phi_{2,1} & \phi_{2,2} & \phi_{2,3} & \phi_{2,4} \\ NA & NA & \phi_{3,3} & \phi_{3,4} \\ NA & \phi_{4,2} & \phi_{4,3} & \phi_{4,4} \\ \cdot & \cdot & \cdot & \cdot \\ \cdot & \cdot & \cdot & \cdot \end{matrix}$$



## Data

$$y_{i,t} = \begin{matrix} & 0 & 1 & 1 & 0 & 0 \\ & 1 & 0 & 0 & 0 & 0 \\ & 0 & 0 & 1 & 0 & 1 \\ & 0 & 1 & 0 & 0 & 0 \\ & . & . & . & . & . \\ & . & . & . & . & . \end{matrix}$$

## Parameters

$$\phi_{i,t} = \begin{matrix} & NA & \phi_{1,2} & \phi_{1,3} & \phi_{1,4} \\ \phi_{2,1} & \phi_{2,1} & \phi_{2,2} & \phi_{2,3} & \phi_{2,4} \\ NA & NA & \phi_{3,3} & \phi_{3,4} \\ NA & \phi_{4,2} & \phi_{4,3} & \phi_{4,4} \\ . & . & . & . \\ . & . & . & . \end{matrix}$$

## Data

$$y_{i,t} = \begin{matrix} & 0 & \textcolor{red}{1} & 1 & 0 & 0 \\ & 1 & 0 & 0 & 0 & 0 \\ & 0 & 0 & 1 & 0 & 1 \\ & 0 & 1 & 0 & 0 & 0 \\ & \cdot & \cdot & \cdot & \cdot & \cdot \\ & \cdot & \cdot & \cdot & \cdot & \cdot \end{matrix}$$

## Parameters

$$\phi_{i,t} = \begin{matrix} NA & \phi_{1,2} & \phi_{1,3} & \phi_{1,4} \\ \phi_{2,1} & \phi_{2,2} & \phi_{2,3} & \phi_{2,4} \\ NA & NA & \phi_{3,3} & \phi_{3,4} \\ NA & \phi_{4,2} & \phi_{4,3} & \phi_{4,4} \\ \cdot & \cdot & \cdot & \cdot \\ \cdot & \cdot & \cdot & \cdot \end{matrix}$$

$$\mathbf{p}_{i,t} = \begin{matrix} \textcolor{red}{NA} & p_{1,2} & p_{1,3} & p_{1,4} \\ p_{2,1} & p_{2,2} & p_{2,3} & p_{2,4} \\ NA & NA & p_{3,3} & p_{3,4} \\ NA & p_{4,2} & p_{4,3} & p_{4,4} \\ \cdot & \cdot & \cdot & \cdot \\ \cdot & \cdot & \cdot & \cdot \end{matrix}$$

## Data

$$y_{i,t} = \begin{matrix} & 0 & 1 & 1 & 0 & 0 \\ & 1 & 0 & 0 & 0 & 0 \\ & 0 & 0 & 1 & 0 & 1 \\ & 0 & 1 & 0 & 0 & 0 \\ & \cdot & \cdot & \cdot & \cdot & \cdot \\ & \cdot & \cdot & \cdot & \cdot & \cdot \end{matrix}$$

## Parameters

$$\phi_{i,t} = \begin{matrix} NA & \phi_{1,2} & \phi_{1,3} & \phi_{1,4} \\ \phi_{2,1} & \phi_{2,2} & \phi_{2,3} & \phi_{2,4} \\ NA & NA & \phi_{3,3} & \phi_{3,4} \\ NA & \phi_{4,2} & \phi_{4,3} & \phi_{4,4} \\ \cdot & \cdot & \cdot & \cdot \\ \cdot & \cdot & \cdot & \cdot \end{matrix}$$

$$p_{i,t} = \begin{matrix} NA & p_{1,2} & p_{1,3} & p_{1,4} \\ p_{2,1} & p_{2,2} & p_{2,3} & p_{2,4} \\ NA & NA & p_{3,3} & p_{3,4} \\ NA & p_{4,2} & p_{4,3} & p_{4,4} \\ \cdot & \cdot & \cdot & \cdot \\ \cdot & \cdot & \cdot & \cdot \end{matrix}$$

# Modelling

along time axis

$$\phi_{i,t} = \begin{matrix} NA & \phi_{1,2} & \phi_{1,3} & \phi_{1,4} \\ \phi_{2,1} & \phi_{2,2} & \phi_{2,3} & \phi_{2,4} \\ NA & NA & \phi_{3,3} & \phi_{3,4} \\ NA & \phi_{4,2} & \phi_{4,3} & \phi_{4,4} \\ \cdot & \cdot & \cdot & \cdot \\ \cdot & \cdot & \cdot & \cdot \end{matrix}$$

along individual axis

Time axis:

- temporal variation (fixed, random)
- temporal covariates
- time-constant

Individual axis:

- groups (fixed, random)
- individual covariates

Interaction of time and individual axes:

- age effect (transients)
- additive effects (group + time)
- cohort effects
- trap-response

Same story for detectability

# A simple model: $\phi(\cdot)$ , $p(\cdot)$

```
# Likelihood
```

```
for (i in 1:nind){
```

```
  # Define latent state at first capture
```

```
  z[i,f[i]] <- 1
```

```
  for (t in (f[i]+1):n.occasions){
```

```
    # State process
```

```
    z[i,t] ~ dbern(phi * z[i,t-1])
```

```
    # Observation process
```

```
    y[i,t] ~ dbern(p * z[i,t])
```

```
  } #t
```

```
} #i
```

```
phi ~ dunif(0,1) # Prior for survival
```

```
p ~ dunif(0,1)   # Prior for recapture
```

$$z_{i,first_i} = 1$$

$$z_{i,t} \sim \text{Bern}(z_{i,t-1}\phi)$$

$$y_{i,t} \sim \text{Bern}(z_{i,t}p)$$

## Another simple model: $\phi(t)$ , $p(t)$

```
# Likelihood
```

```
for (i in 1:nind){
```

```
  # Define latent state at first capture
```

```
  z[i,f[i]] <- 1
```

```
  for (t in (f[i]+1):n.occasions){
```

```
    # State process
```

```
    z[i,t] ~ dbern(phi[t-1] * z[i,t-1])
```

```
    # Observation process
```

```
    y[i,t] ~ dbern(p[t-1] * z[i,t])
```

```
  } #t
```

```
} #i
```

```
phi[t] ~ dunif(0,1) # Prior for survival
```

```
p[t] ~ dunif(0,1) # Prior for recapture
```

$$z_{i,first_i} = 1$$

$$z_{i,t} \sim \text{Bern}(z_{i,t-1}\phi_{t-1})$$

$$y_{i,t} \sim \text{Bern}(z_{i,t}p_t)$$

# A general program structure

```
# Likelihood
```

```
for (i in 1:nind){
```

```
  # Define latent state at first capture
```

```
  z[i,f[i]] <- 1
```

```
  for (t in (f[i]+1):n.occasions){
```

```
    # State process
```

```
    z[i,t] ~ dbern(phi[i,t-1] * z[i,t-1])
```

```
    # Observation process
```

```
    y[i,t] ~ dbern(p[i,t-1] * z[i,t])
```

```
  } #t
```

```
} #i
```

$$z_{i,first_i} = 1$$

$$z_{i,t} \sim \text{Bern}(z_{i,t-1} \phi_{i,t-1})$$

$$y_{i,t} \sim \text{Bern}(z_{i,t} p_{i,t})$$

$$\phi_{i,t} = \begin{array}{cccc} \text{NA} & \phi_{1,2} & \phi_{1,3} & \phi_{1,4} \\ \phi_{2,1} & \phi_{2,2} & \phi_{2,3} & \phi_{2,4} \\ \text{NA} & \text{NA} & \phi_{3,3} & \phi_{3,4} \\ \text{NA} & \phi_{4,2} & \phi_{4,3} & \phi_{4,4} \\ \cdot & \cdot & \cdot & \cdot \\ \cdot & \cdot & \cdot & \cdot \end{array}$$

# A general program structure

```
# Likelihood
for (i in 1:nind){

  # Define latent state at first capture
  z[i,f[i]] <- 1

  for (t in (f[i]+1):n.occasions){
    # State process
    z[i,t] ~ dbern(phi[i,t-1] * z[i,t-1])

    # Observation process
    y[i,t] ~ dbern(p[i,t-1] * z[i,t])
  } #t
} #i
```

Example for  $\phi(\cdot)$ ,  $p(\cdot)$ :

```
# Priors and constraints
for (i in 1:nind){
  for (t in 1:(n.occasions-1)){
    phi[i,t] <- mean.phi
    p[i,t] <- mean.p
  } #t
} #i
```

```
mean.phi ~ dunif(0,1) # Prior mean
surv
mean.p ~ dunif(0,1)   # Prior mean
recap
```



# Assumptions



Survival refers to the study area: mortality and permanent emigration are confounded. Therefore we estimate **apparent survival**, but not *true* survival.

## Assumptions

- Design:
  - No mark lost
  - Identity of individuals recorded without error (no *false positives*)
  - Captured individuals are a random sample
- Model:
  - Homogeneity of survival and recapture probabilities
  - Independence between individuals

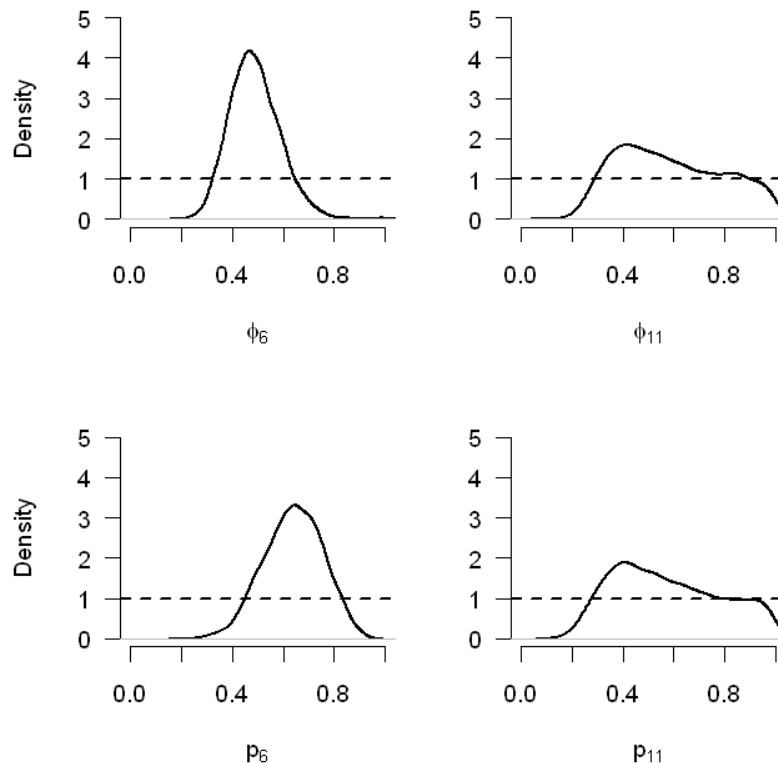
# Parameter identifiability

- **Intrinsic** non-identifiability
  - Likelihood can be expressed by a smaller number of parameters
  - Feature of the model
- **Extrinsic** non-identifiability
  - Model structure is OK, but due to lacking data a parameter cannot be estimated
  - Feature of the data

# Parameter identifiability

Slightly different in Bayesian

- Due to prior knowledge, **all parameters are defined**, but they **may not be informed by the data**
- Compare prior and posterior



Example: classic CJS ( $\phi_t, p_t$ ) with 12 years and uniform priors

# Multinomial model for capture-recapture data

From the capture-histories to the m-array data format

*Capture histories*

```
1 0 1 0
1 1 0 0
1 0 1 1
0 1 0 0
```

*m-array*

	Recapture occ.			
Release occ.	2	3	4	Never recaptured
1				
2	–			
3	–	–		

# Multinomial model for capture-recapture data

From the capture-histories to the m-array data format

*Capture histories*

**1 0 1 0**

1 1 0 0

1 0 1 1

0 1 0 0

*m-array*

	Recapture occ.			
Release occ.	2	3	4	Never recaptured
1		<b>1</b>		
2	–			
3	–	–		<b>1</b>

# Multinomial model for capture-recapture data

From the capture-histories to the m-array data format

*Capture histories*

1 0 1 0

1 1 0 0

1 0 1 1

0 1 0 0

*m-array*

Release occ.	Recapture occ.			Never recaptured
	2	3	4	
1	1	1		
2	–			1
3	–	–		1

# Multinomial model for capture-recapture data

From the capture-histories to the m-array data format

*Capture histories*

1 0 1 0

1 1 0 0

1 0 1 1

0 1 0 0

*m-array*

Release occ.	Recapture occ.			Never recaptured
	2	3	4	
1	1	1+1		
2	–			1
3	–	–	1	1

# Multinomial model for capture-recapture data

From the capture-histories to the m-array data format

*Capture histories*

1 0 1 0

1 1 0 0

1 0 1 1

0 1 0 0

*m-array*

	Recapture occ.			
Release occ.	2	3	4	Never recaptured
1	1	1+1		
2	–			1+1
3	–	–	1	1



# Multinomial model for capture-recapture data

From the capture-histories to the m-array data format

*Capture histories*

```
1 0 1 0
1 1 0 0
1 0 1 1
0 1 0 0
```

*m-array*

	Recapture occ.			
Release occ.	2	3	4	Never recaptured
1	1	2	0	0
2	–	0	0	2
3	–	–	1	1

# Multinomial model for capture-recapture data

Cell probabilities of the m-array:

	Recapture occ.			
Rel.	2	3	4	Never
1	$\phi_1 p_1$	$\phi_1 (1-p_1) \phi_2 p_2$	$\phi_1 (1-p_1) \phi_2 (1-p_2) \phi_3 p_3$	$1-\Sigma_1$
2	–	$\phi_2 p_2$	$\phi_2 (1-p_2) \phi_3 p_3$	$1-\Sigma_2$
3	–	–	$\phi_3 p_3$	$1-\Sigma_3$

Likelihood of the m-array:

$$\mathbf{m}_t \sim \text{Multinomial}(\mathbf{q}_t, R_t)$$

where

$m_t = t^{\text{th}}$  row of m-array

$q_t$  = cell probabilities of the m-array for release occasion  $t$

$R_t$  = number of released individuals at occasion  $t$

# Bayesian GOF: PPC & Bayesian P-values

- **Posterior predictive check (PPC)**: a general and very common Bayesian model-checking procedure
- In a nutshell:
  - Quantify how different the ***observed*** data is compared to what we ***expect*** from the estimated parameters [measure of **discrepancy**]
  - To assess how “bad” or “good” that difference is, compare to the same measure of discrepancy obtained for ideal data ***simulated*** from the estimated parameters
  - This double comparison is done for each sample of the MCMC chains
- Generally straightforward to implement in BUGS or in R after the analysis (from the MCMC chains)

# Bayesian GOF: PPC & Bayesian P-values

## Basic algorithm:

At each MCMC iteration  $t$ :

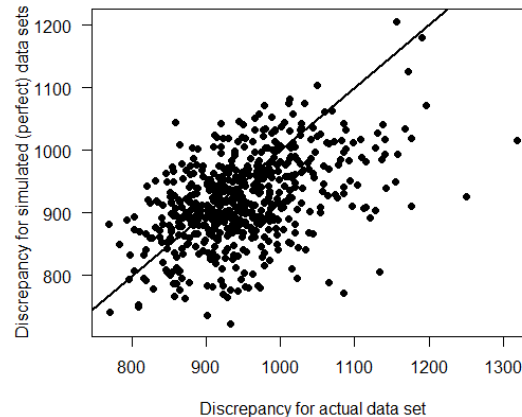
- 1) Calculate a **discrepancy statistic**  $D(x; e_t)$  to measure some “distance” between the **observed data**  $x$  and the corresponding expected values,  $e_t$  at that iteration (based on the MCMC chain values of the parameters)
- 2) Generate a new data set  $x_t$  from the fitted model (using the MCMC chain value of the parameters)
- 3) Calculate a discrepancy statistic  $D(x_t; e_t)$  to measure the “distance” between the **simulated data**  $x_t$  and the corresponding expected values,  $e_t$  at that iteration

# Bayesian GOF: PPC & Bayesian P-values

Then, use the values of  $D(x; e_t)$  and  $D(x_t; e_t)$  to check the fit...

1) ...graphically: scatterplot of  $D(x; e_t)$  and  $D(x_t; e_t)$

No evidence of lack of fit  
if  $\sim \frac{1}{2}$  points above and  
 $\sim \frac{1}{2}$  below a 1:1 line



2) ...numerically by calculating a **Bayesian p-value** = proportion of times that  $D(x; e_t) < D(x_t; e_t)$

→ close to 0.5: good fit

→ close to 0 or 1: suggests doubtful fit

# Bayesian GOF: PPC & Bayesian P-values


On what response to define *observed* and *expected* values?

- Not on binary responses (i.e. detection histories) themselves
- Use some aggregating: **m-array in CR data**


What measure of discrepancy?

- Different options, e.g.

- **Pearson chi-squared statistic**


$$D(x_t, e_t) = \sum_i \frac{(x_i - e_i)^2}{e_i}$$

- **Freeman-Tukey statistic**


$$D(x_t, e_t) = \sum_i (\sqrt{x_i} - \sqrt{e_i})^2$$

- Different measures sensible to different kinds of lack of fit; resulting p-values may differ substantially
- Normally assess some global measure of lack of fit; can also assess particular features of the model (e.g. extreme values)

# Bayesian GOF: PPC & Bayesian P-values

- Implicitly depend on priors
  - Bayesian p-values have been criticised
    - They use data twice:  
simulate data; compare them to these data
    - Not clear what values represent a good fit
- Descriptive only! (not for model selection)

## Jolly-Seber model

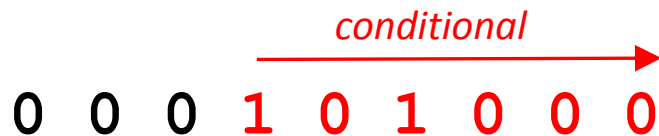
- The CJS model is *conditional* on first capture

0 0 0 1 0 1 0 0 0



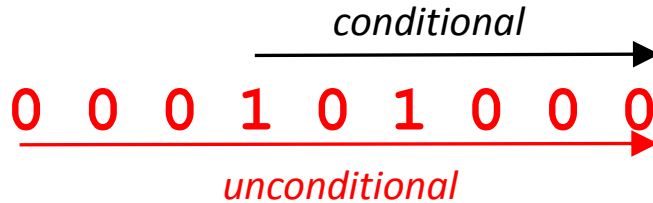
## Jolly-Seber model

- The CJS model is *conditional* on first capture



# Jolly-Seber model

- The CJS model is *conditional* on first capture



- The Jolly-Seber model is *unconditional*
- Estimation of *recruitment* and *population size* is possible
- Requires an additional assumption (capture = recapture)
- See chapter 10 in BPA book