Estimation with incomplete detection at distance zero "g(0)<1"

Chapter 6 of Advanced book (Methods for incomplete detection at distance zero by Laake and Borchers)

Borchers, D., Laake, J., Southwell, C. and Paxton, C. 2006. Accommodating unmodeled heterogeneity in double-observer distance sampling surveys. *Biometrics* **62**: 372-378

Buckland, S.T., Laake, J.L. and Borchers, D.L. 2009. Double-observer line transect methods: levels of independence. *Biometrics* **66**: 169-177

Laake, J.L., Collier, B.A., Morrison, M.L. and Wilkins, R.N. 2011. Point-based mark-recapture distance sampling. JABES **16**: 389-408

Burt, M.L., Borchers, D.L., Jenkins, K.J. and Marques, T.A.M. 2014. Using mark-recapture distance sampling methods on line transect surveys. *Methods in Ecology and Evolution* **5**: 1180-1191.

Conventional Distance sampling estimates are biased if $g(0)<1$:

$$D^* = D \times g(0)$$

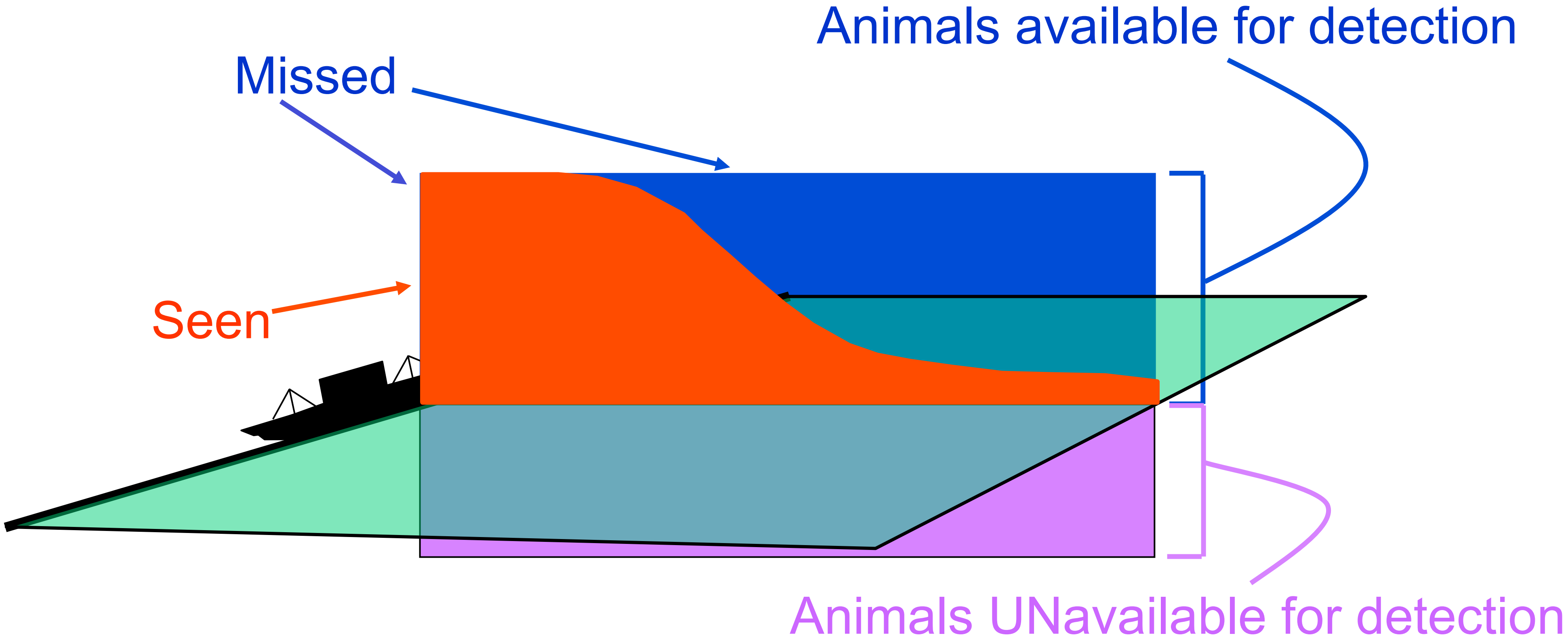
where D is the true density and D^* is the density obtained if you assume $g(0)=1$.

***g*(0)<1** when there is

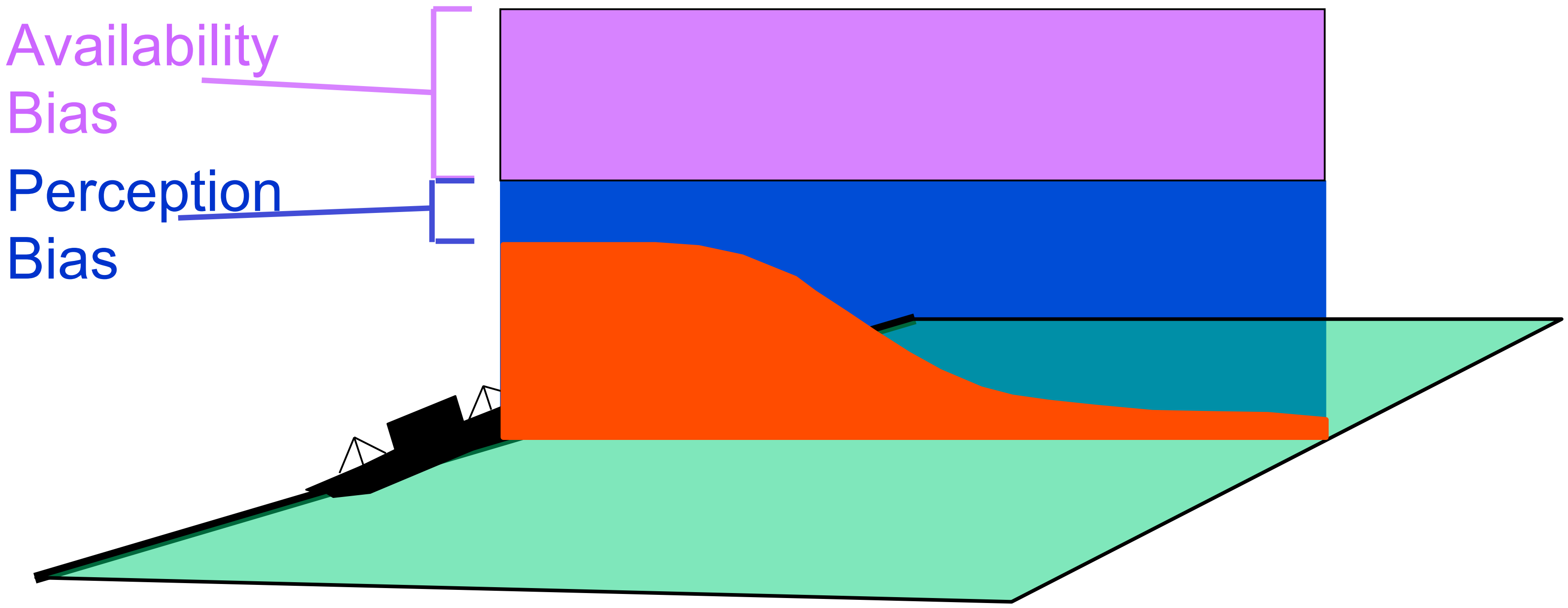
Availability Bias

Perception Bias at distance 0

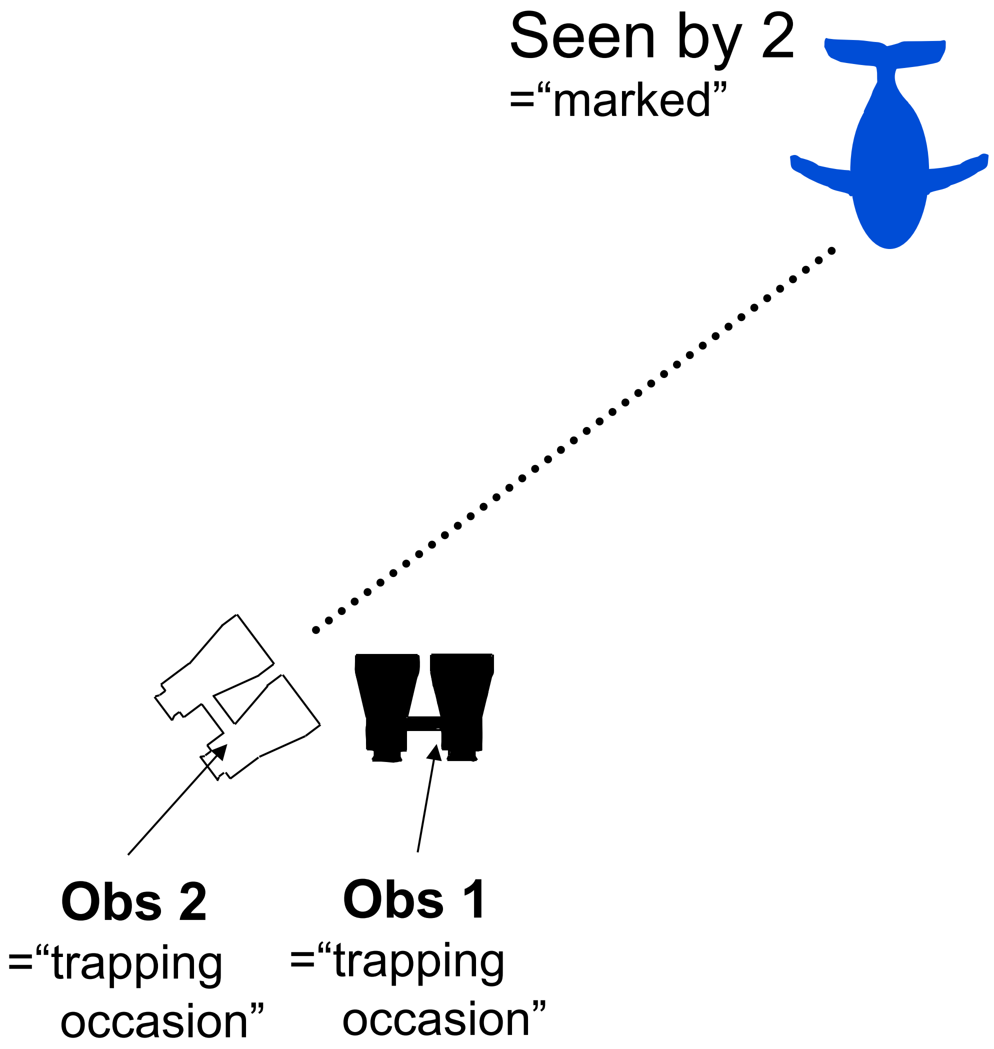
- “**Availability Bias**”: When animals are unavailable for detection.
- “**Perception Bias**”: When observers fail to detect animals at distance 0 although they are available



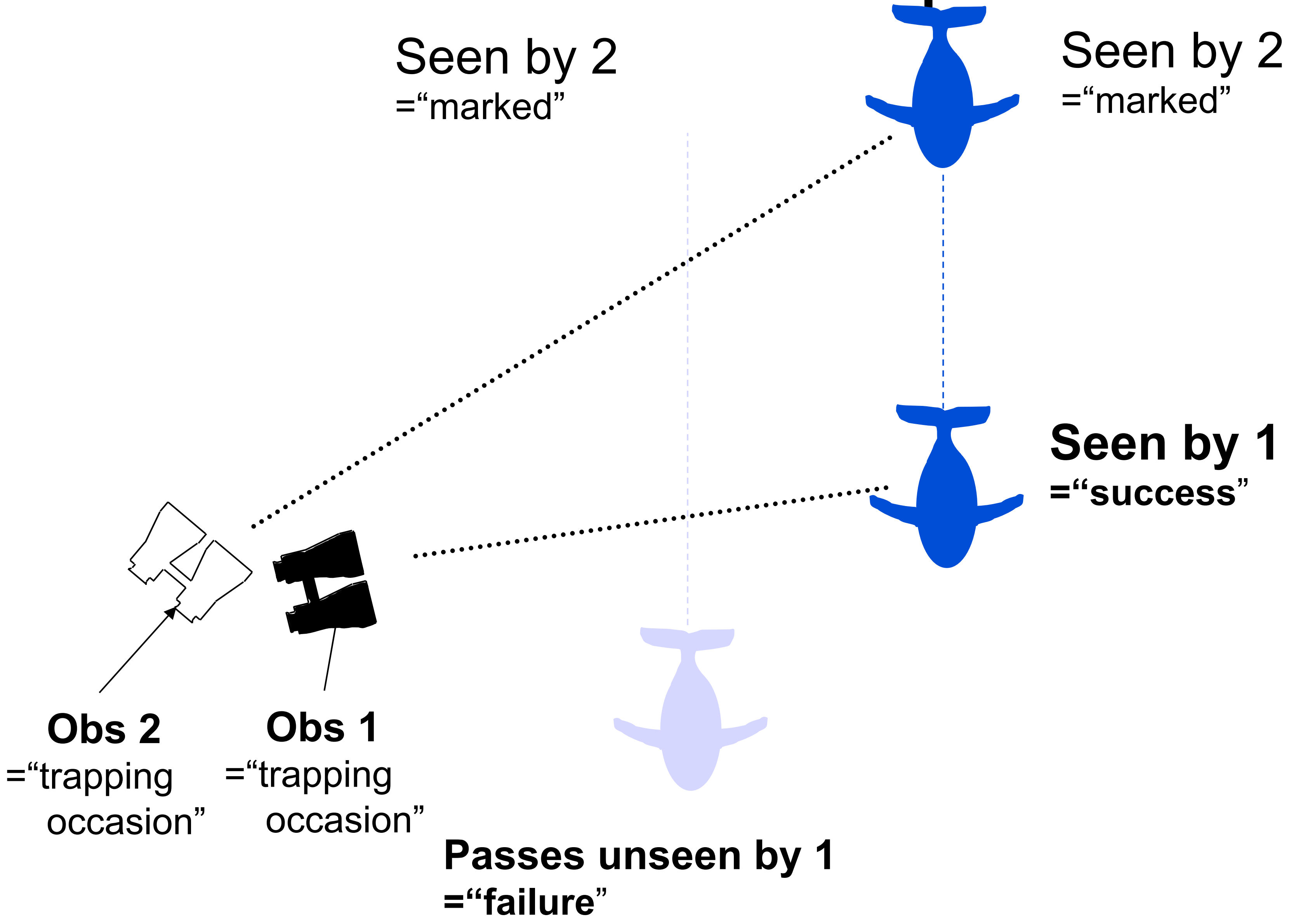
- "**Availability Bias**": When animals are unavailable for detection.
- "**Perception Bias**": When observers fail to detect animals on the transect although they are available



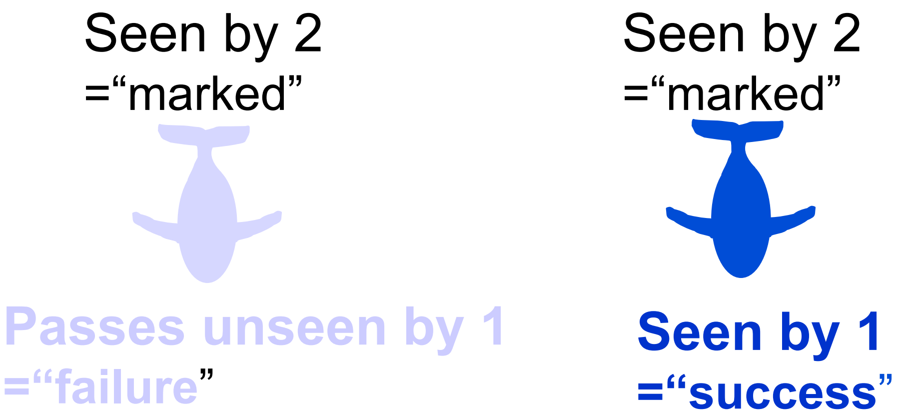
Visual Mark-Recapture



Visual Mark-Recapture



Visual Mark-Recapture



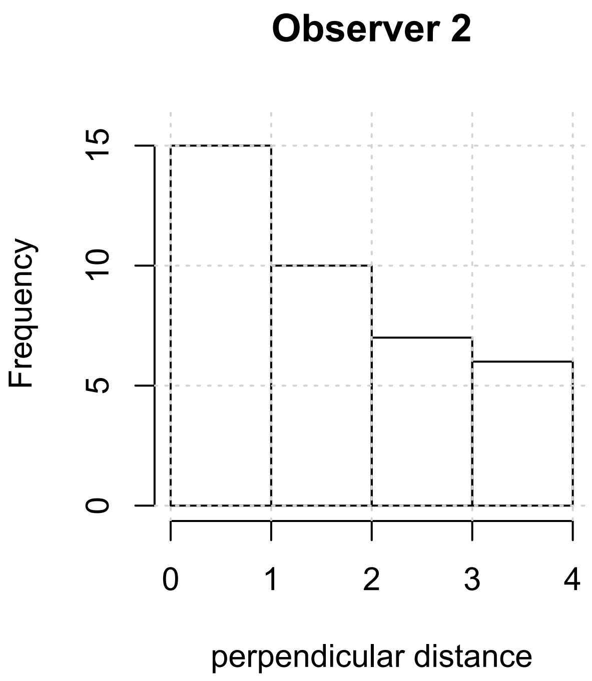
- We know 2 animals passed
  (because Obs 2 saw them)

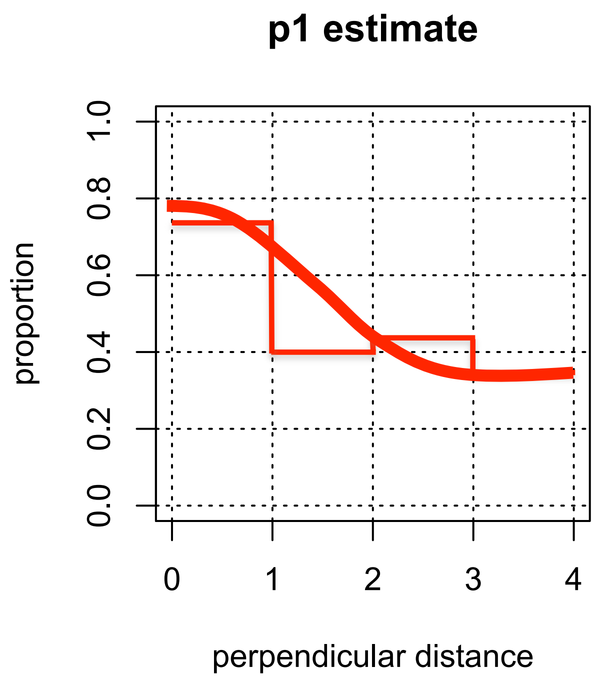
- Of these, Obs 1 saw 1

- So **estimate:**
  Pr(Obs 1 sees) = $\hat{p}_1 = \frac{1}{2} = \frac{n_{12}}{n_2}$ = $\frac{\text{number "duplicates"}}{\text{number seen by 2}}$

Note: In this section, we use *p*, not *g* for the detection function

Class Exercise



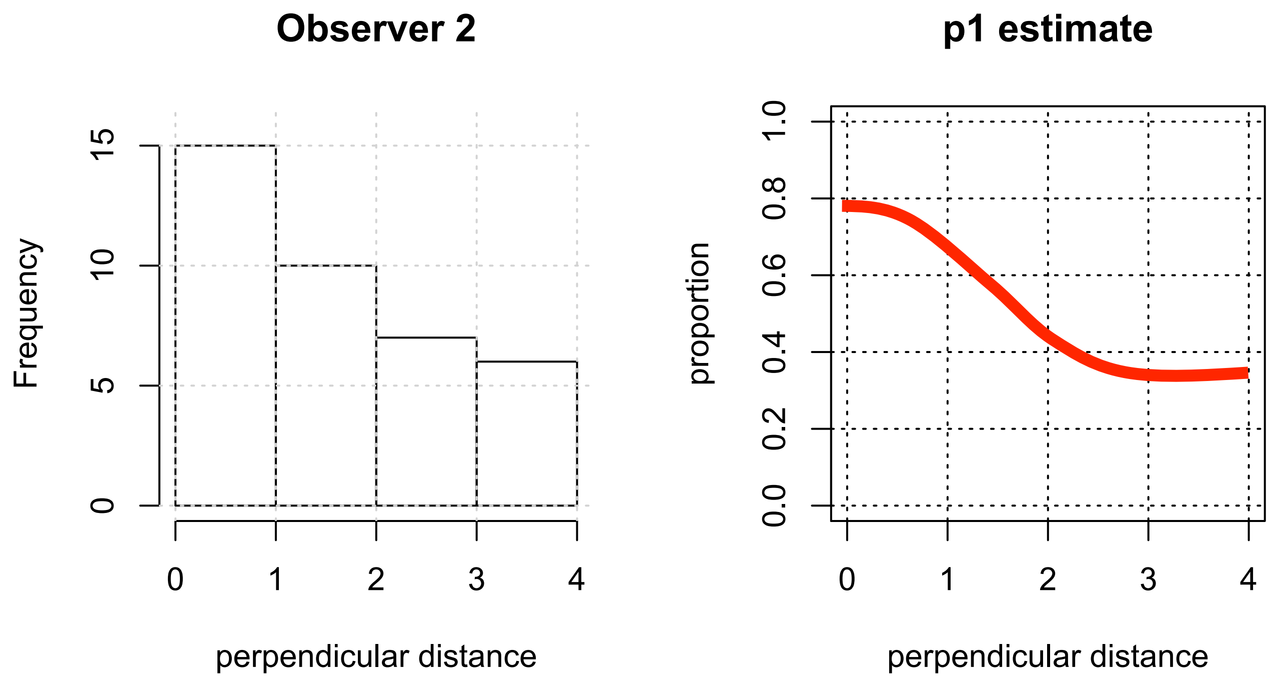


Obs 2 detections:

	n_2	$\hat{p}$	n_1	$\hat{N}_x$
100s: 101,102,103,104,105,106,107,108,111,112,114,115,116,118,134	15			
200s: 201,202,204,205,206,207,211,214,215,218	10			
300s: 301,303,304,305,307,313,314	7			
400s: 402,404,407,416,417,418	6			
$n_{dups} = 20$	38		25	

$$\hat{N}_{Petersen} = \frac{n_1}{\hat{p}_1} = \frac{25}{20/38} = 47.5$$

$$\hat{N}_{TOTAL} =$$



Fit smooth curve using Logistic Regression
(instead of grouping into distance intervals)

Duplicate Identification

Field methods

- Use a dedicated “duplicate identifier”
- Record measure of confidence in duplicate identification.
- Record positions and times as precisely as possible
- Record ancillary data
- Have at least one observer “track” animals

Duplicate Identification

Analysis methods

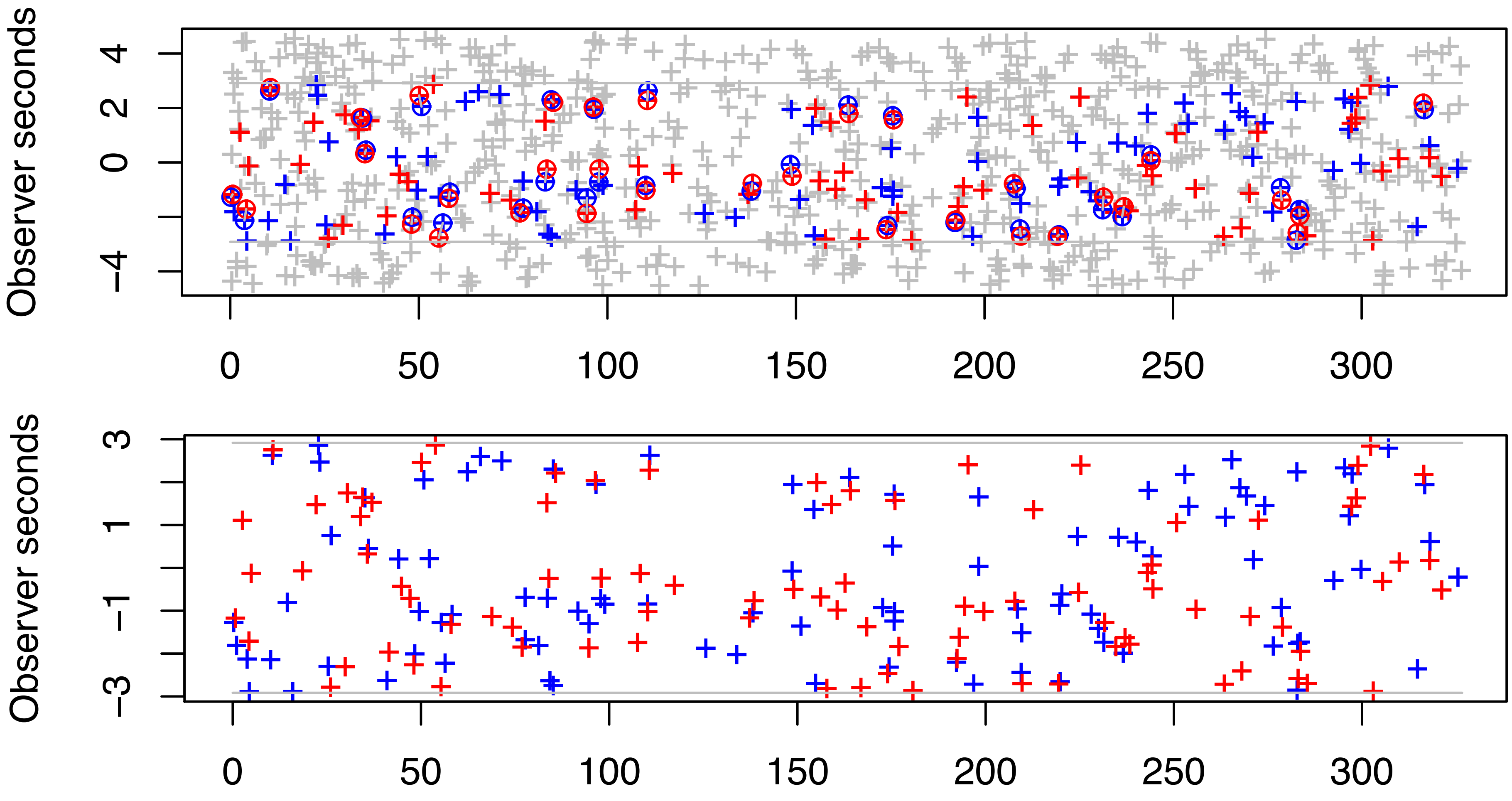
- Bracket "best" estimate by two extremes

- Rule-based duplicate identification after the survey. (e.g. Schweder et al., 1996)

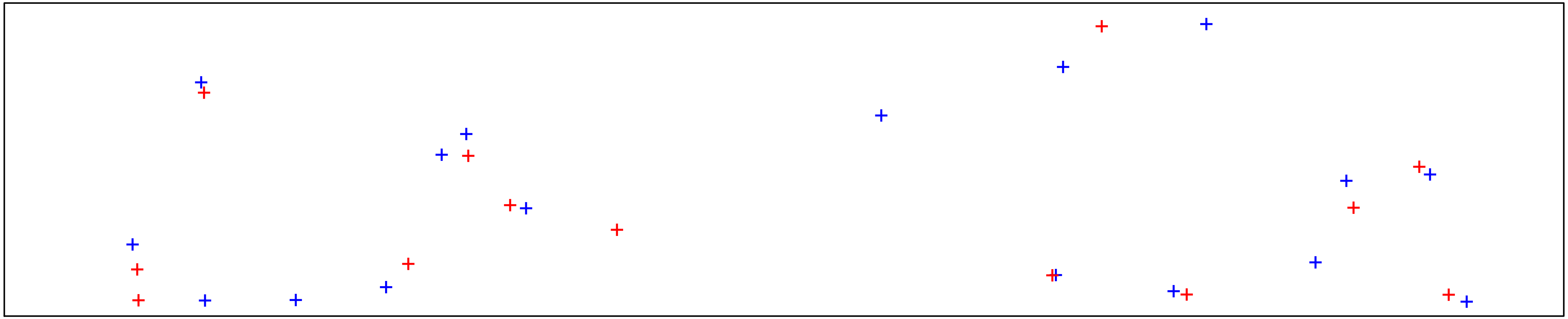
- Probabilitistic duplicate identification after the survey. (e.g. Hiby and Lovell, 1998, Borchers *et al*. *in prep*., Stevenson *et al*. *in prep*)

Schweder, T., Hagen, G., Helgeland, J. and Koppervik, I. 1996. Abundance estimation of northeastern Atlantic minke whales. *Rep. Int. Whal. Commn.* **46**: 391-405.
Hiby, A. and Lovell, P.1998. Using aircraft in tandem formation to estimate abundance of harbour porpoise. *Biometrics* **54**: 1280-1289.

Probabilistic Duplicate Identification



Probabilistic Duplicate Identification



Design to deal with availability bias

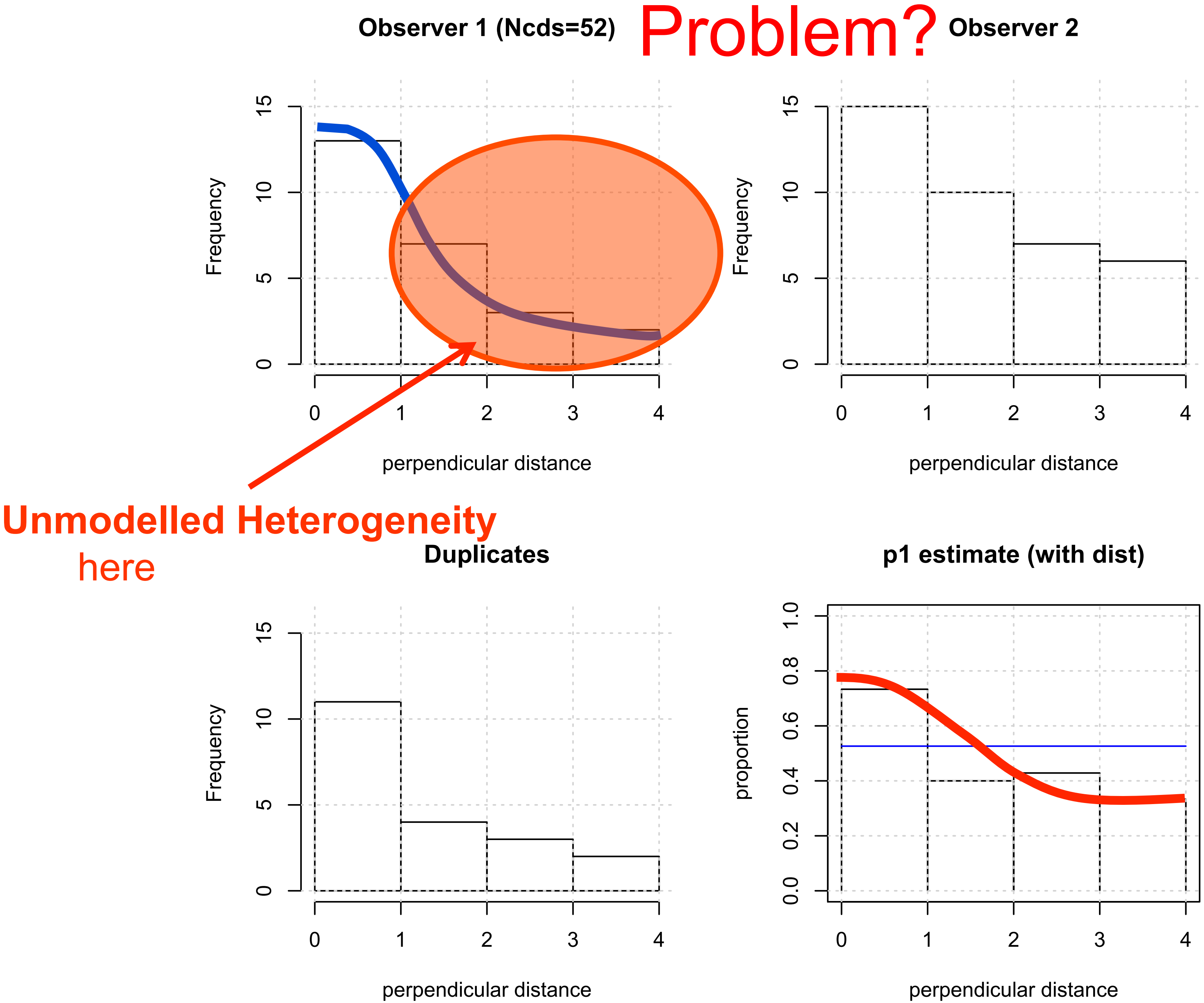
Use enough effort for certain detection at x=0: May not be possible

Use cue-based methods : Need to estimate availability process

Separate search areas of the observers (see pp 176-177 Adv. book)

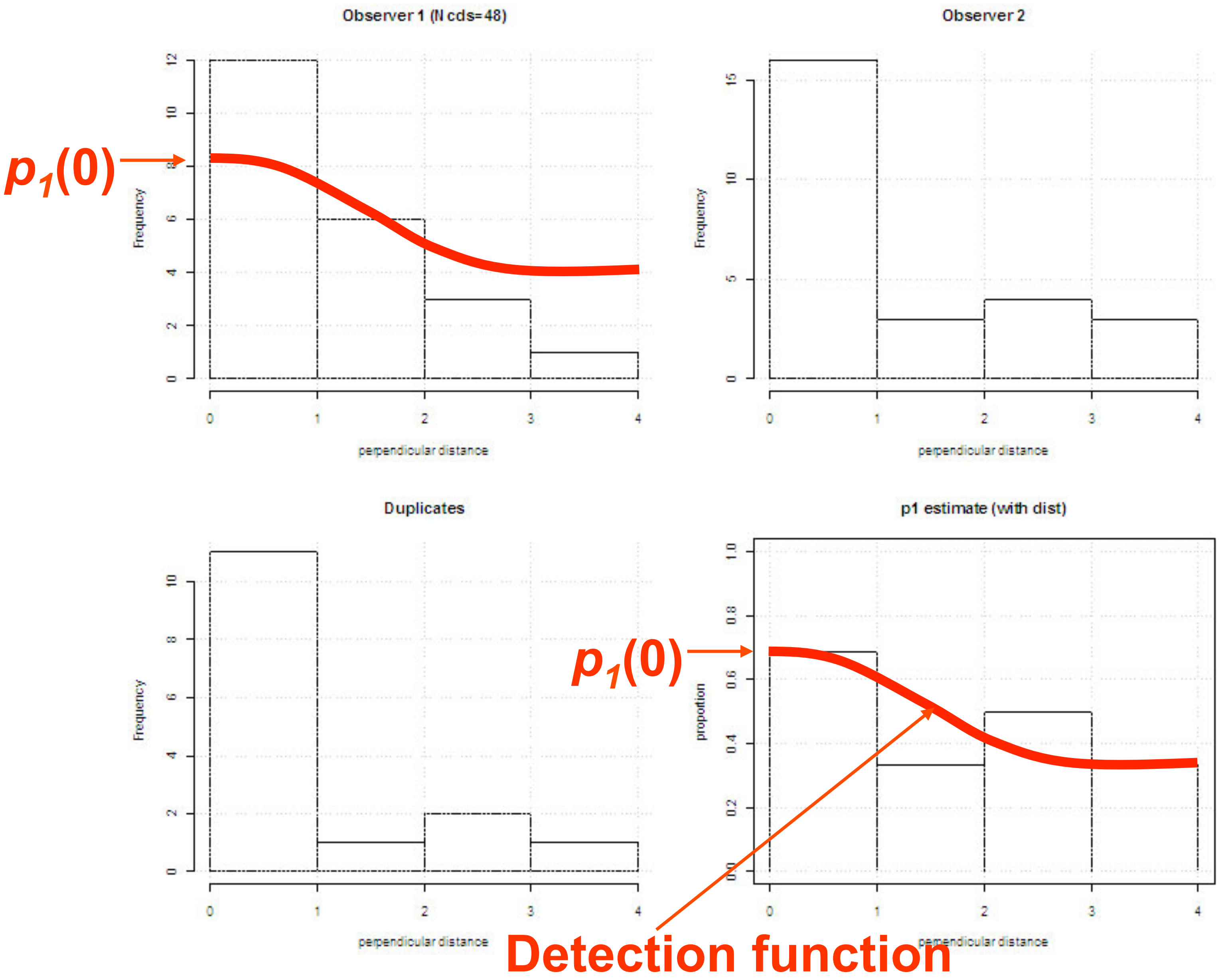
Use different types of observers (e.g. visual and acoustic; visual and radio-tag)

Availability bias correction factor: Need to be careful if animals in veiw for more than very small fraction of their availability cycle time.

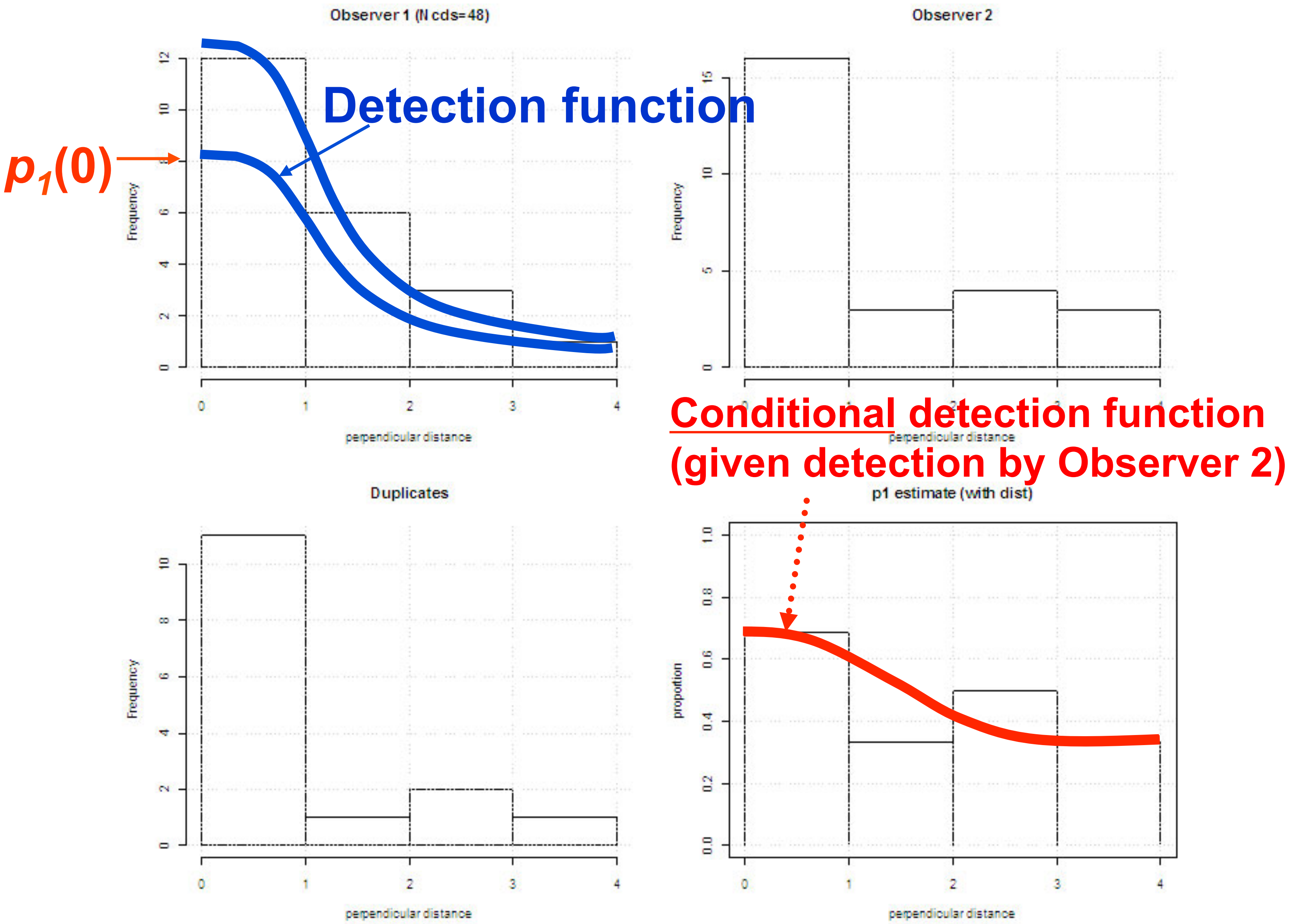


Observer 1 (Ncds=52)
Problem?
Observer 2
Frequency
15
10
5
0
0
1
2
3
4
perpendicular distance
Unmodelled Heterogeneity
here
Duplicates
p1 estimate (with dist)
proportion
1.0
0.8
0.6
0.4
0.2
0.0

Full Independence (FI) Model:



Point Independence (PI) Model:



Point vs Full Independence

Full Independence

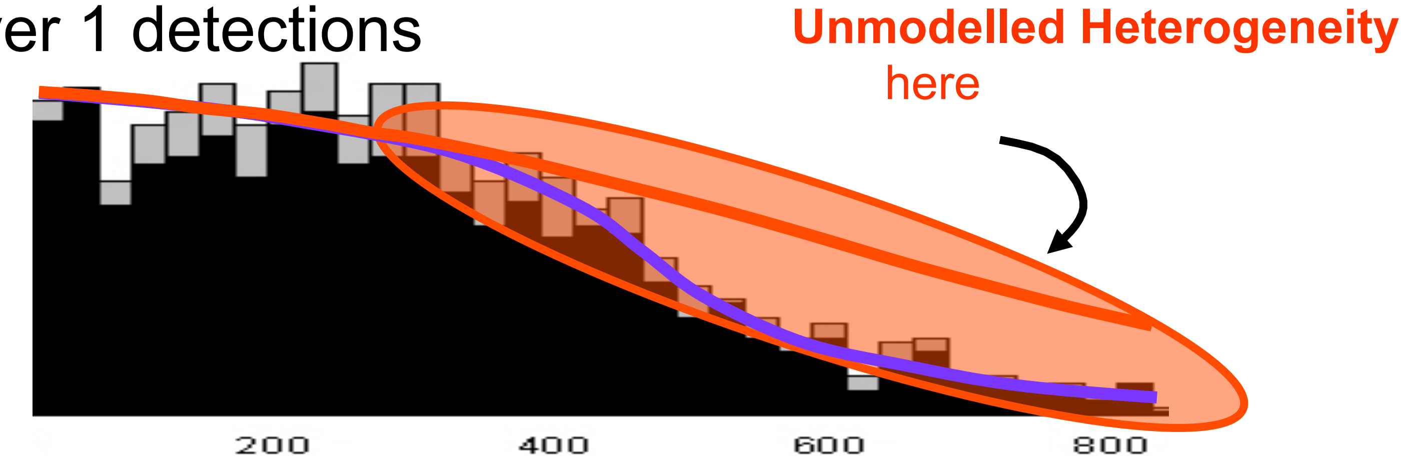
- Class e.g. Nhat= 48.
- Sensitive to unmodelled heterogeneity: negative bias.
- Assumption of uniform animal distribution not required - so useful if there is responsive movement.
- Don’t use unless you have to.

Point Independence

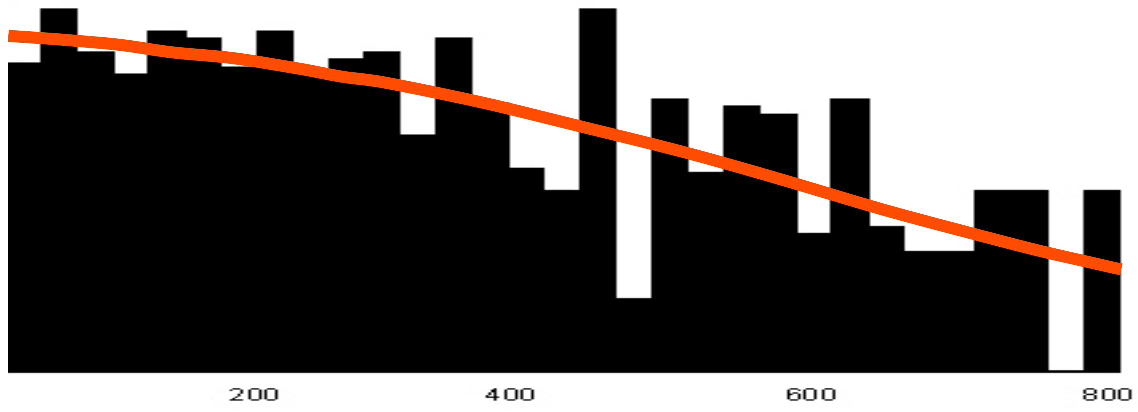
- Class e.g. Nhat= 70.
- Much less sensitive to unmodelled heterogeneity.
- Assumption of uniform animal distribution required – so no good if there is responsive movement.
- Use it unless there is responsive movement (or other non-uniform distribution).

Example: Pack-Ice Seals

Observer 1 detections

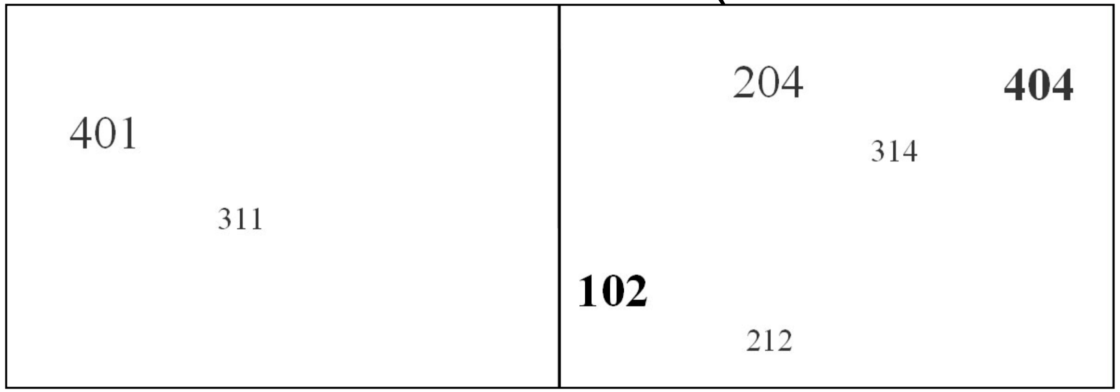


Proportion of Observer 2 detections seen by Observer 1



Sources of Heterogeneity

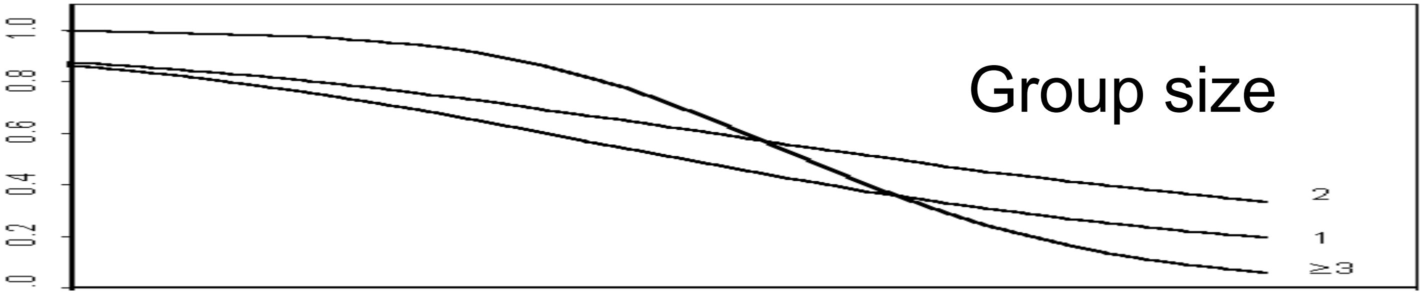
- The **animals** themselves (size, boldness)



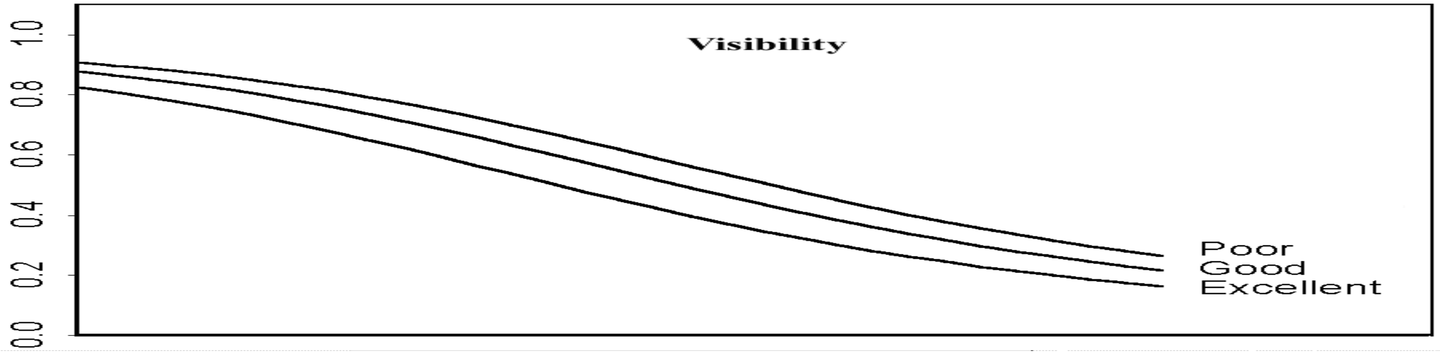
- The **environment** (clear/"misty")

Sources of Heterogeneity

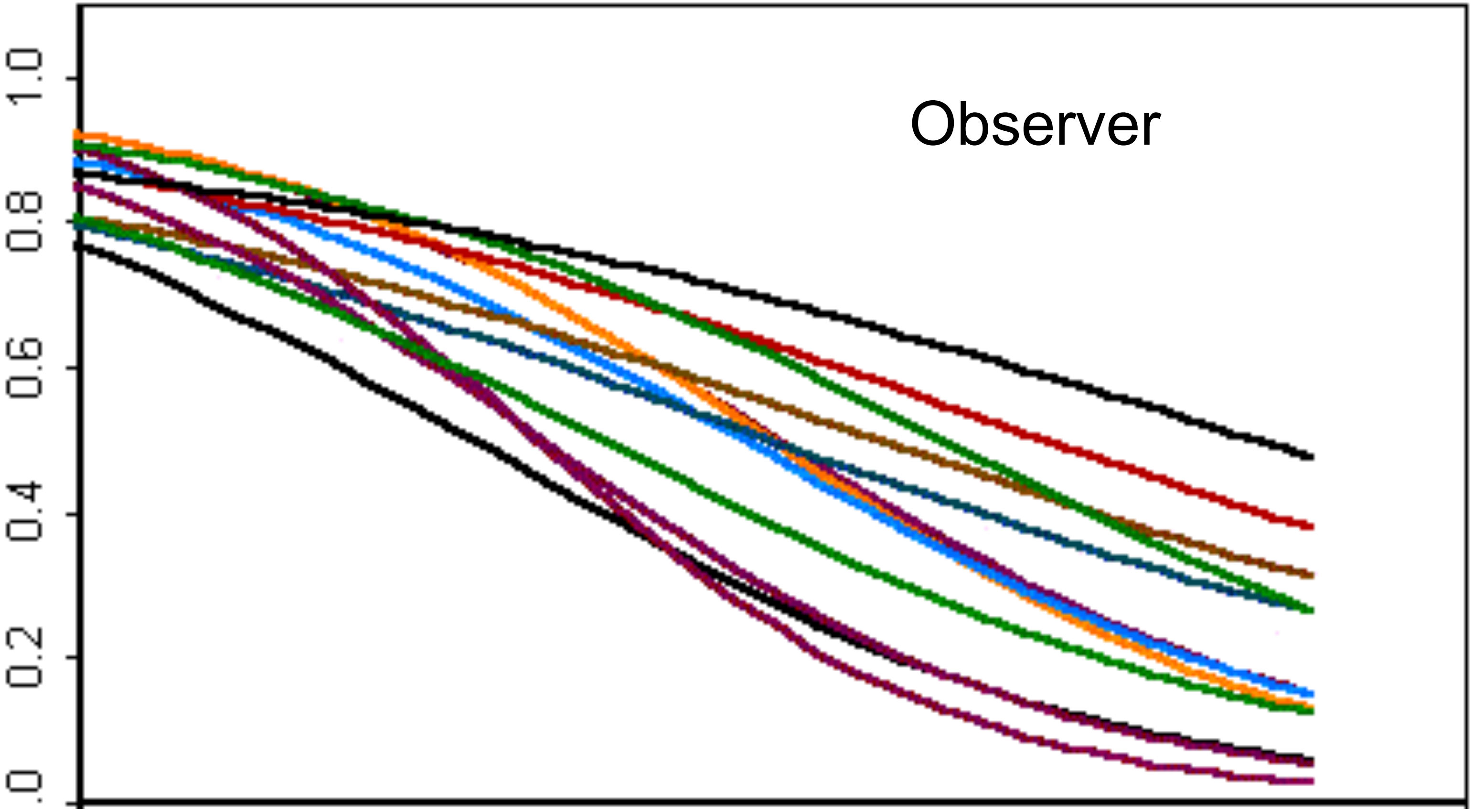
- The **animals** themselves (distance, size, availability, ...)



- The **environment** (sea state, ground cover, ...)



- The kind of **survey effort** (the observers, their platforms, ...)



Configuration: Trial-Observer

Observer 2

↓

sets up trials for

↓

Observer 1

↓

to estimate p_1

The Observer at the end of an arrow must be independent of
the Observer at the start of the arrow

Configuration: Independent Observer

to estimate p_2

Observer 2

sets up trials for

Observer 1

to estimate p_1

$$\boxed{p_{.} = p_1 + p_2 - (\, p_1 \, p_2 \,)}$$

The Observer at the end of an arrow must be independent of
the Observer at the start of the arrow

Abundance Estimation

- Trial-Observer $N = \sum_{seen\ by\ 1} 1/p(x_i, \ldots)$
- Independent Observer $N = \sum_{seen} 1/p(x_i, \ldots)$

Double-Platform Analysis Types

<u>Cue-based methods</u>:

- Cues (not animals) are units; estimate p(see cue)
- Getting adequate estimates of cue generation process can be difficult.
- Able to incorporate heterogeneity due to availability (cue-ing) process.

• <u>Animal-based methods</u>:

We focus on these; in some applications cue-based methods perform better

- Animals are units; estimate p(see animal)
- Don’t need to estimate availability/cue-ing process.
- More difficult to incorporate heterogeneity due to availability process.

Related Models not covered:

Limiting Independence

- Assume no unmodelled heterogeneity not at any point, but only as *p* approaches 1.

- See Buckland, S.T., Laake, J.L. and Borchers, D.L. 2009. Double-observer line transect methods: levels of independence. *Biometrics* **66**: 169-177

Point Transects

- Can also do full, point and limiting independence with **Point Transects**.

- See Laake, J.L., Collier, B.A., Morrison, M.L. and Wilkins, R.N. 2011. Point-based mark-recapture distance sampling. JABES **16**: 389-408

Critical Assumptions of Mark Recapture Line Transect

- Have the required independence between observers
- No unmodelled heterogeneity
- Duplicates (resightings) known (else need to include uncertainty in duplicate status in estimated variance)