Practical advice

David L Miller & Jason J Roberts

Real survey data is messy

Collecting and analysing data in the real world

- We've talked a lot about models

- We've also talked about assumptions

- Our example is relatively well-behaved

- What can we do about all the nasty real world stuff?

Some days…

🔥 🔥 🖥️ 🔥 🔥

Aims

- Here we want to cover common questions
- Not definitive answers
- Some guidance on where to look for answers

What should my sample size be?
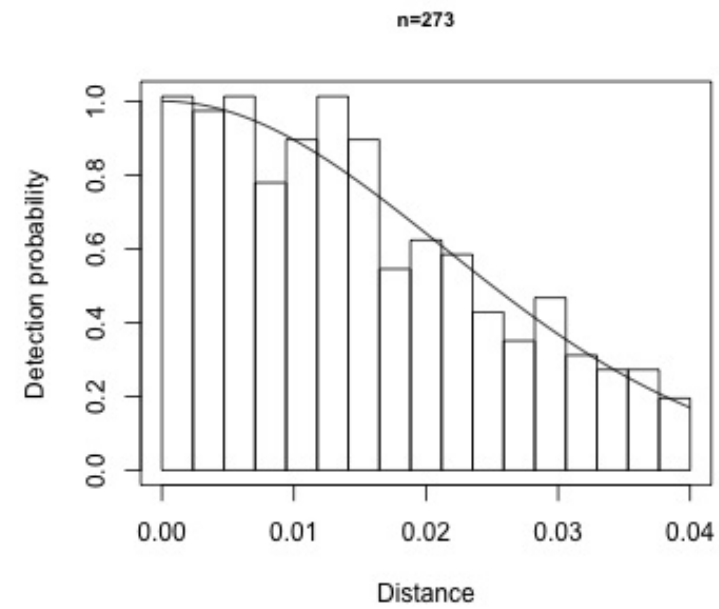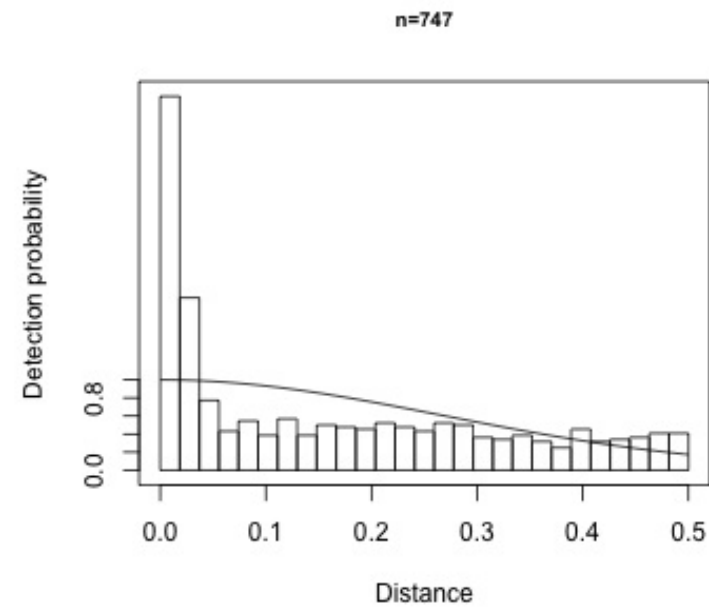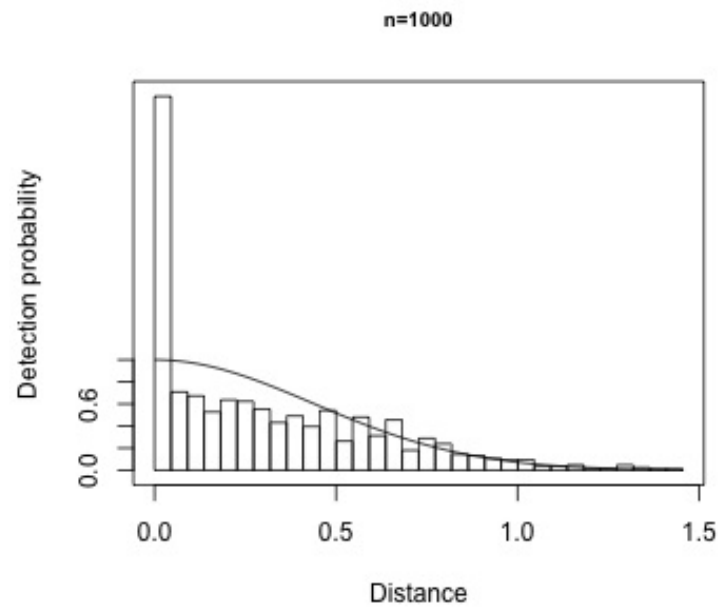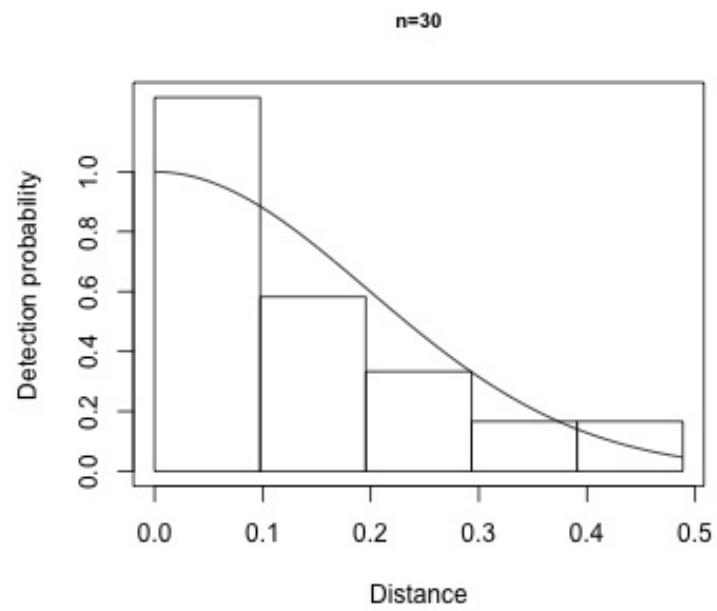
What do we mean by "sample size"?

- Number of animal (groups) recorded
  - *detection function*

- Number of segments
  - *spatial model*

- Number of segments with observations
  - *spatial model*

Re-frame

How would we know when we have enough samples?

- We don't

- Heavily context-dependent

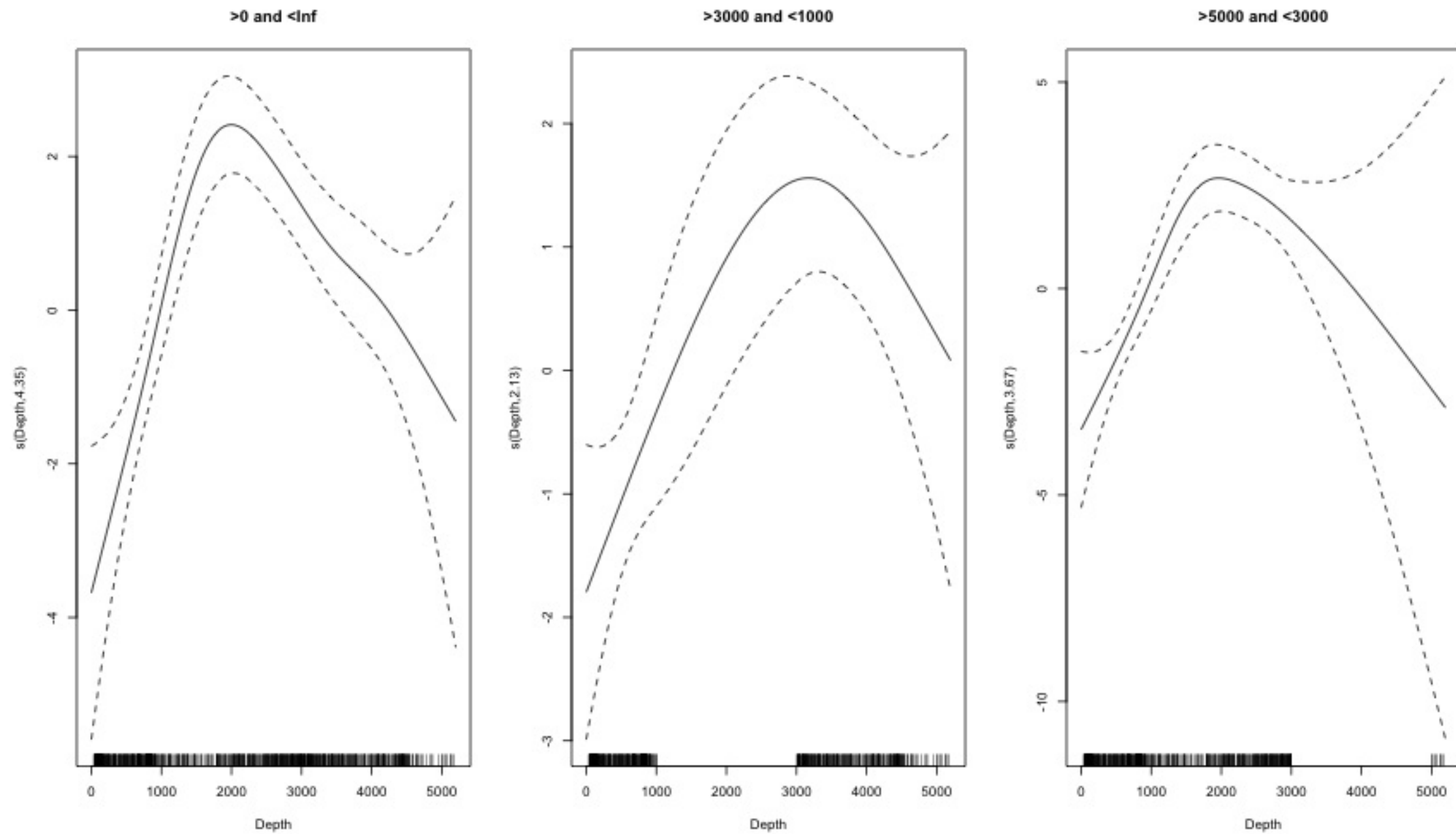- Go back to assumptions

"How much data?"

Pilot studies and "you get what you pay for"

- Designing surveys is hard

- Designing surveys is essential

- Better to fail one season than fail for 5, 10 years

- Get information early, get it cheap

  - Inform design from a pilot study

Avoiding rules of thumb

- Think about assumptions
    - Detection function
    - Spatial model
- Think about design
    - Spatial coverage
    - Covariate coverage

Coverage

Disappointment

- Sometimes you don't have enough data

- Or, enough coverage

- Or, the right covariates

Sometimes, you can't build a spatial model

"Which of X, Y, Z is correct?"

Alternatives problem

- When faced with options, try them all.

- Where does the sensitivity lie?

- What's **really** going on?

"How big should our segments be?"

Segment size

- If you think it's an issue test it
- Resolution of covariates also important
- Maybe species-/domain-dependent?

"Is our model right?"

Model validation

- Some variety of cross-validation
- Temporal replication
  - Leave out 1 year, fit to others, predict, assess
- Spatial "pseudo-jackknife"
  - Leave out every n^{th} segment, refit, …
  - (Maybe leave out 2, 3 etc…)

Getting help

Resources

- Bibliography has pointers to these topics
- Distance sampling Google Group
  - Friendly, helpful, low traffic
  - see distancesampling.org/distancelist.html