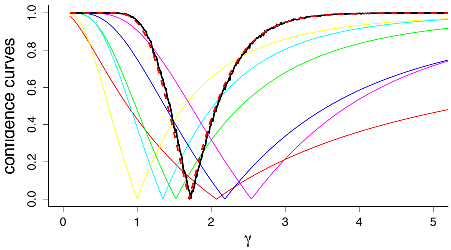
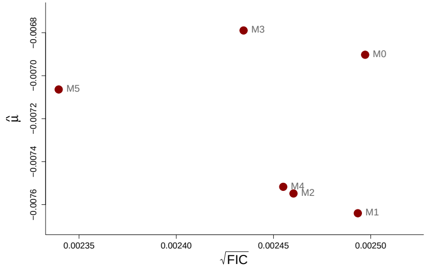
Wars and Whales: Extensions and Applications of Confidence Curves and Focused Model Selection

Céline Cunen

Department of Mathematics, University of Oslo

07/12/2018


confidence curves
0.0
0.2
0.4
0.6
0.8
1.0
0
1
2
3
4
5
γ

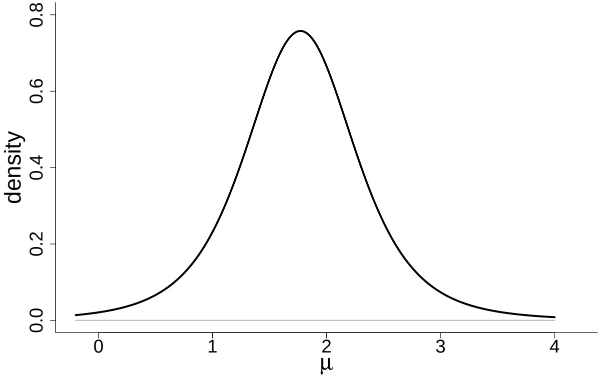

$\hat{\mu}$
−0.0068
−0.0070
−0.0072
−0.0074
−0.0076
M3
M0
M5
M4
M2
M1
0.00235
0.00240
0.00245
0.00250
$\sqrt{\text{FIC}}$

Overview

- Four papers in total, ranging from primarily methodological to primarily applied.
- Confidence curves (and confidence distributions) appear in all the papers.
- The Focused information criterion (FIC) appears in two papers.
- Change-points play a role in two (even three) papers.
- Real data problems are treated in all papers, for instance:
  - Wars;
  - Whales;
  - Tirant lo Blanch.

Confidence distributions (CDs)

- The goal of statistical inference is to obtain data-dependent statements about unknown parameters, usually with an attached uncertainty measurement.
- Statements? Point-estimates; tests; confidence intervals; densities (usually for Bayesians only);...

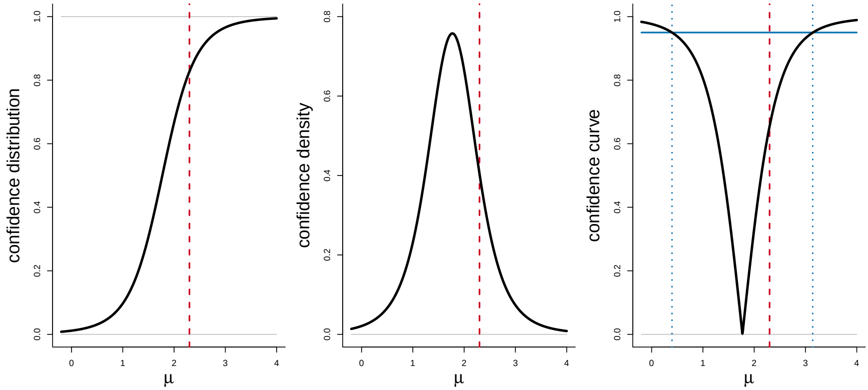


Confidence distributions (CDs)

- CD $\approx$ a posterior without having to specify a prior
  - a sample-dependent distribution function on the parameter space
  - can be used for inference (for example for constructing confidence intervals of all levels)

Confidence distributions (CDs)

- CD $\approx$ a posterior without having to specify a prior
  - a sample-dependent distribution function on the parameter space
  - can be used for inference (for example for constructing confidence intervals of all levels)



Requirements for CDs

Definition

A function $C(\theta, Y)$ is called a confidence distribution for a parameter θ if:

- $C(\theta, Y)$ is a cumulative distribution function on the parameter space
- at the true parameter value $\theta = \theta_0$, $C(\theta_0, Y)$ as a function of the random sample Y follows the uniform distribution U[0,1]

- The second requirement ensures that all confidence intervals have the correct coverage.
- We will typically construct CDs and confidence curves by exact or approximate pivots.
- Note that *any* method producing confidence intervals fulfilling these requirements can be used to make CDs (no matter the underlying paradigm).

Papers

Paper I

Cunen, Hermansen, and Hjort (2018). Confidence distributions for change-points and regime shifts. *Journal of Statistical Planning and Inference* 195, 14–34.

Paper II

Cunen, Hjort, and Nygård (2018). Statistical Sightings of Better Angels: Analysing the Distribution of Battle Deaths in Interstate Conflict over Time. *Invited to submit a revision to Journal of Peace Research*.

Paper III

Cunen, Walløe, and Hjort (2018). Focused model selection for linear mixed models, with an application to whale ecology. *Invited to resubmit to Annals of Applied Statistics*.

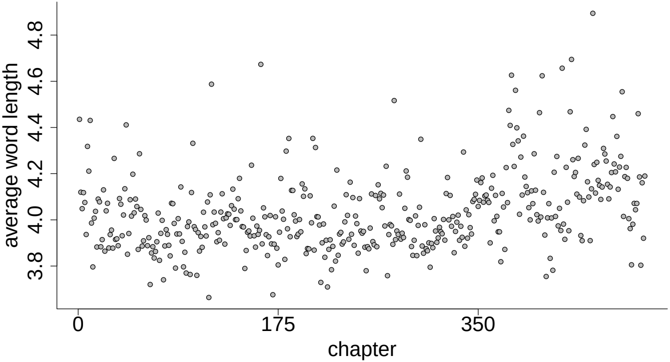
Paper IV

Cunen and Hjort (2018). Combining information across diverse sources: The II-CC-FF paradigm. *Submitted for publication*.

I: Change-point problems

We have a sequence of observations,

.

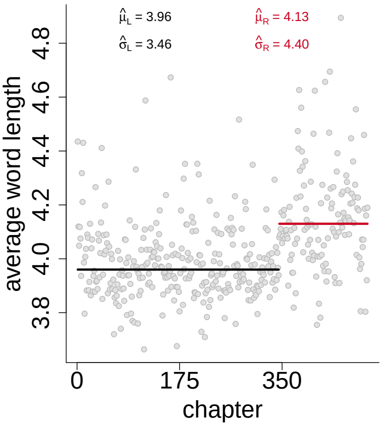


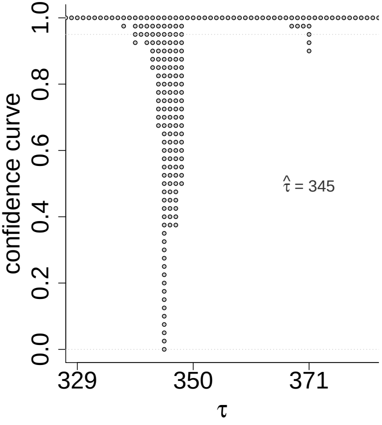
and we assume the following model,

$$y_i \sim \mathrm{N}(\mu_L, \sigma_L^2/m_i) \text{ for } i \leq \tau,$$
$$y_i \sim \mathrm{N}(\mu_R, \sigma_R^2/m_i) \text{ for } i \geq \tau + 1.$$

I: Change-point problems

We have a sequence of observations, and we would like to find the point of maximal change - and quantify the degree of change.





I: Two methods

We assume

$$y_1, \ldots, y_\tau \sim f(y, \theta_L), \qquad y_{\tau+1}, \ldots, y_n \sim f(y, \theta_R).$$

Method A relies on testing the homogeneity of the sequences $(y_1, \ldots, y_\tau)$ and $(y_{\tau+1}, \ldots, y_n)$ at each potential change-point τ. Say we have a test statistic $Z_{a,b}$ for the sequence $(y_a, \ldots, y_b)$ with null distribution $G_{a,b}$. For each level α the confidence set is given by

$$R(\tau) = \{\tau \colon Z_{1,\tau} \leq G^{-1}_{1,\tau}(\sqrt{\alpha}) \text{ and } Z_{\tau+1,n} \leq G^{-1}_{\tau+1,n}(\sqrt{\alpha})\}.$$

Flexible. Makes no a priori assumptions on the existence of a change.

I: Two methods

We assume

$$y_1, \ldots, y_\tau \sim f(y, \theta_L), \qquad y_{\tau+1}, \ldots, y_n \sim f(y, \theta_R).$$

Method B uses the log-likelihood profile

$$\begin{aligned} \ell_{n,\mathrm{prof}}(\tau) &= \max\{\ell_{1,\tau}(\theta_L) + \ell_{\tau+1,n}(\theta_R)\colon \text{all } \theta_L, \theta_R\} \\ &= \ell_{1,\tau}(\widehat{\theta}_L) + \ell_{\tau+1,n}(\widehat{\theta}_R). \end{aligned}$$

This leads (i) to the ML estimate $\widehat{\tau}$; (ii) to the deviance function $D(\tau, y) = 2\{\ell_{n,\mathrm{prof}}(\widehat{\tau}) - \ell_{n,\mathrm{prof}}(\tau)\}$. We use

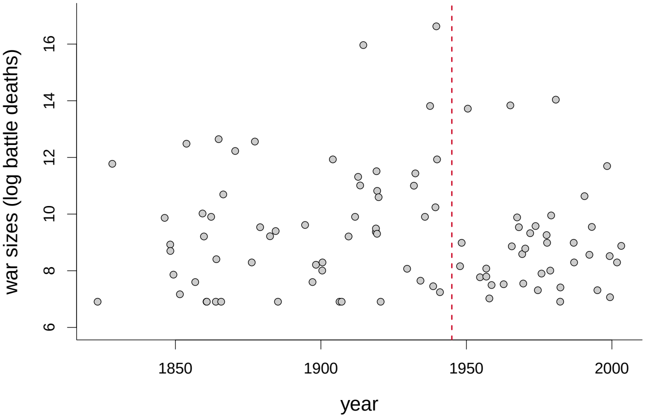
$$\mathrm{cc}(\tau, y_{\mathrm{obs}}) = \mathrm{Pr}_\tau\{D(\tau, Y) < D(\tau, y_{\mathrm{obs}})\} = K_\tau(D(\tau, y_{\mathrm{obs}})),$$

with $K_\tau(x)$ the c.d.f. of the random $D(\tau, Y)$. This requires simulations for each candidate τ. We can also compute a confidence curve for the degree of change $\rho = d(\theta_L, \theta_R)$ by similar methods.

More powerful. Explicitly assumes the existence of a change.

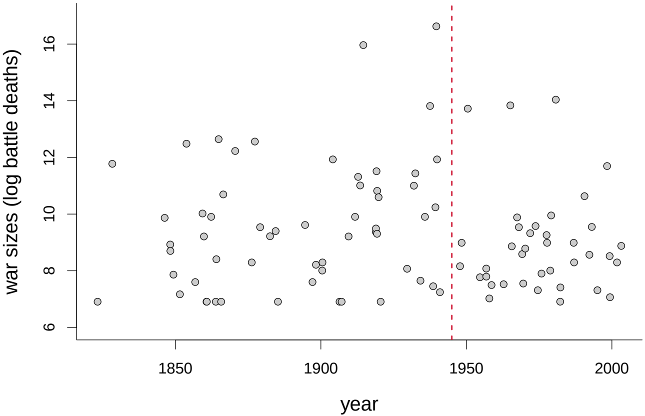
II: The long peace?

Current debate: has there been a (significant) decrease in the sizes of wars?



II: The long peace?

Current debate: has there been a (significant) decrease in the sizes of wars?



We investigate the related question: when is the point of maximal change in this sequence? And what is the magnitude of this change?

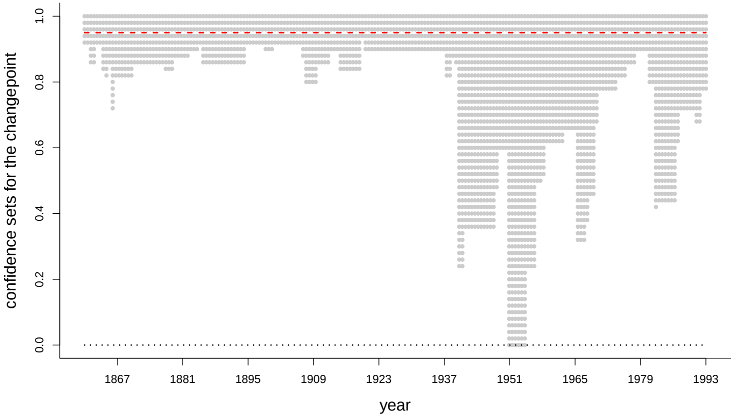
II: Change-point analysis with method B

After careful modelling of the war size distribution, and using method B from the previous paper, we find:

II: Change-point analysis with method B

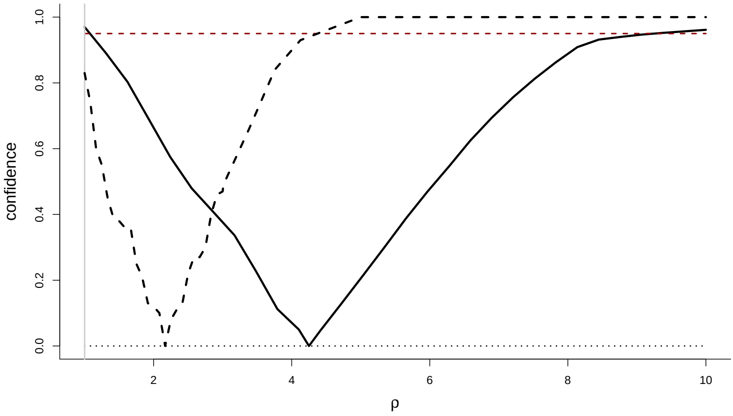
After careful modelling of the war size distribution, and using method B from the previous paper, we find:

a point estimate in $\hat{\tau} = 1950.483 \to$ the Korean War.



II: The degree of change

Let $\rho_1 = \phi_{0.50,L}/\phi_{0.50,R}$ (dashed) and $\rho_2 = \phi_{0.75,L}/\phi_{0.75,R}$ (fully drawn),



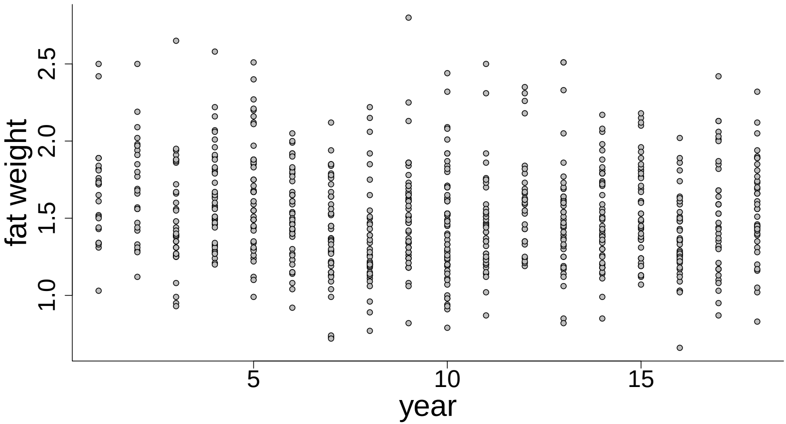
At $\rho = 1$ we are testing H_0: *the world is constant*.

II: Other efforts

- Modelling of the full sequence of war sizes
- Modelling only the largest wars
- Model selection (FIC)
- Including covariates in the change-point analysis (democracy scores)
- Analysing the timings between wars

III: Focused model selection for linear mixed models

Motivation: are Antarctic Minke whales becoming thinner?



III: Focused model selection for linear mixed models

Motivation: are Antarctic Minke whales becoming thinner?

Models:

$$
\begin{aligned}
M_0: \quad Y \sim & \mathit{Year} + \mathit{BodyLength} + \mathit{Sex} + \mathit{Diatom} + \mathit{Date} + \mathit{Date}^2 \\
& + \mathit{Latitude} + \mathit{Sex} * \mathit{FetusLength} + \mathit{Sex} * \mathit{Diatom} + \mathit{Diatom} * \mathit{Date} \\
& + \mathit{Diatom} * \mathit{Date}^2 + \mathit{Latitude} * \mathit{Date} + \mathit{Latitude} * \mathit{Date}^2 \\
& + \mathit{BodyLength} * \mathit{Date} * \mathit{Sex} + (1 + \mathit{Date} + \mathit{Date}^2 | \mathit{Year}).
\end{aligned}
$$

III: Focused model selection for linear mixed models

Motivation: are Antarctic Minke whales becoming thinner?

Models:

$$
\begin{aligned}
M_0: \quad Y \sim & \, Year + BodyLength + Sex + Diatom + Date + Date^2 \\
& + Latitude + Sex * FetusLength + Sex * Diatom + Diatom * Date \\
& + Diatom * Date^2 + Latitude * Date + Latitude * Date^2 \\
& + BodyLength * Date * Sex + (1 + Date + Date^2 | Year).
\end{aligned}
$$

$\dots$

$$M_5: \quad Y \sim Year + BodyLength + Sex + Date + (1 + Date \,|\, Year).$$

III: Focused model selection for linear mixed models

Motivation: are Antarctic Minke whales becoming thinner?

Models:

$$
\begin{aligned}
M_0: \quad Y \sim & \mathit{Year} + \mathit{BodyLength} + \mathit{Sex} + \mathit{Diatom} + \mathit{Date} + \mathit{Date}^2 \\
& + \mathit{Latitude} + \mathit{Sex} * \mathit{FetusLength} + \mathit{Sex} * \mathit{Diatom} + \mathit{Diatom} * \mathit{Date} \\
& + \mathit{Diatom} * \mathit{Date}^2 + \mathit{Latitude} * \mathit{Date} + \mathit{Latitude} * \mathit{Date}^2 \\
& + \mathit{BodyLength} * \mathit{Date} * \mathit{Sex} + (1 + \mathit{Date} + \mathit{Date}^2 | \mathit{Year}).
\end{aligned}
$$

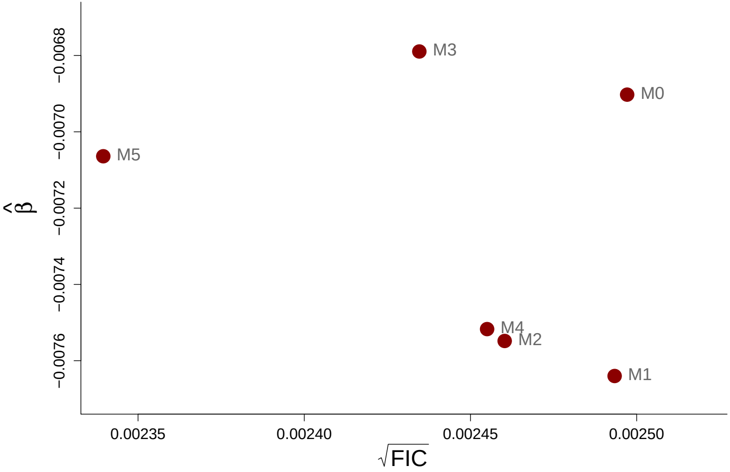
$\ldots$

$$M_5: \quad Y \sim \mathit{Year} + \mathit{BodyLength} + \mathit{Sex} + \mathit{Date} + (1 + \mathit{Date} \,|\, \mathit{Year}).$$

Which model gives the most precise estimates of β_{year}? FIC offers an anwer to this question, in particular FIC ranks models according to their estimated mean-squared error $\mathrm{FIC}(M) = \widehat{\mathrm{mse}}_M$,

$$\mathrm{mse}_M = \mathrm{E}\,(\hat\beta_{M,\mathrm{year}} - \beta_{\mathrm{year}})^2 = \mathrm{Var}\,\hat\beta_{M,\mathrm{year}} + (\mathrm{E}\,\hat\beta_{M,\mathrm{year}} - \beta_{\mathrm{year}})^2.$$

III: Whale results



Model M_5 is considered the best! (note that the scale is tons)

III: Some details

The whale models are linear mixed effect models (LMEs):

$$y_i \sim \mathrm{N}_{m_i}(X_i\beta, \sigma^2(I + Z_i D Z_i^{\mathrm{t}})),$$

with $i = 1, \ldots, n$ natural groups of m_i observations. We have a focus parameter $\mu = \mu(\beta, \sigma, D)$ and estimate it by $\widehat{\mu} = \mu(\widehat{\beta}, \widehat{\sigma}, \widehat{D})$.

III: Some details

The whale models are linear mixed effect models (LMEs):

$$y_i \sim \mathrm{N}_{m_i}(X_i\beta, \sigma^2(I + Z_i D Z_i^{\mathrm{t}})),$$

with $i = 1, \ldots, n$ natural groups of m_i observations. We have a focus parameter $\mu = \mu(\beta, \sigma, D)$ and estimate it by $\widehat{\mu} = \mu(\widehat{\beta}, \widehat{\sigma}, \widehat{D})$.

In the candidate models $y_i \sim \mathrm{N}_{m_i}(X_{M,i}\beta_M, \sigma_M^2(I + Z_{M,i} D_M Z_{M,i}^{\mathrm{t}}))$, we have $\widehat{\mu}_M = \mu(\widehat{\beta}_M, \widehat{\sigma}_M, \widehat{D}_M)$.

III: Some details

The whale models are linear mixed effect models (LMEs):

$$y_i \sim \mathrm{N}_{m_i}\big(X_i\beta, \sigma^2(I + Z_i D Z_i^{\mathrm{t}})\big),$$

with $i = 1, \ldots, n$ natural groups of m_i observations. We have a focus parameter $\mu = \mu(\beta, \sigma, D)$ and estimate it by $\widehat{\mu} = \mu(\widehat{\beta}, \widehat{\sigma}, \widehat{D})$.

In the candidate models $y_i \sim \mathrm{N}_{m_i}\big(X_{M,i}\beta_M, \sigma_M^2(I + Z_{M,i} D_M Z_{M,i}^{\mathrm{t}})\big)$, we have $\widehat{\mu}_M = \mu(\widehat{\beta}_M, \widehat{\sigma}_M, \widehat{D}_M)$.

The expectations and variances in the FIC formulas are computed wrt the wide model, using the following joint approximate distribution,

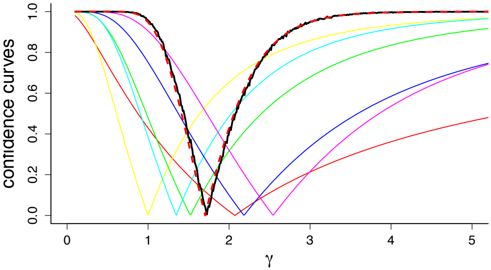
$$\begin{pmatrix} \sqrt{n}(\widehat{\mu} - \mu_{true}) \\ \sqrt{n}(\widehat{\mu}_M - \mu_{M,0}) \end{pmatrix} \approx_d \mathrm{N}_2\left(0, \begin{pmatrix} \nu_{wide} & \nu_{M,c} \\ \nu_{M,c} & \nu_M \end{pmatrix}\right),$$

with $\nu_{wide} = c^{\mathrm{t}} J_n^{-1} c$, $\nu_{M,c} = c^{\mathrm{t}} J_n^{-1} C_{M,n} J_{M,n}^{-1} c_M$, $\nu_M = c_M^{\mathrm{t}} J_{M,n}^{-1} K_{M,n} J_{M,n}^{-1} c_M$. We have explicit formulas for all the necessary quantities.

IV: Confidence curves for combination of information

We would like to combine independent sources of information $i = 1, \ldots, k$.

- They might inform on some common parameter ϕ (meta-analysis);
- or on different parameters which inform on some overall parameter $\phi = \phi(\psi_1, \ldots, \psi_k)$.
- From sources we might have the full dataset, from others we might have summaries only.
- The sources may be of different quality.



IV: II-CC-FF – a 3-step recipe

Combining information, for inference about a focus parameter $\phi = \phi(\psi_1, \ldots, \psi_k)$:
II: Independent Inspection: From data source y_i to estimates and intervals, in the form of a confidence distribution/curve:

$$y_i \Longrightarrow C_i(\psi_i)$$

CC: Confidence Conversion: From the confidence distribution to a confidence log-likelihood,

$$C_i(\psi_i) \Longrightarrow \ell_{c,i}(\psi_i)$$

FF: Focused Fusion: Use the combined confidence log-likelihood $\ell_f(\psi_1, \ldots, \psi_k) = \sum_{i=1}^k \ell_{c,i}(\psi_i)$ to construct a CD for the given focus $\phi = \phi(\psi_1, \ldots, \psi_k)$, often via profiling:

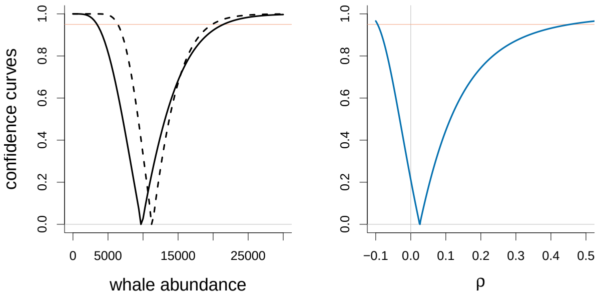
$$\ell_f(\psi_1, \ldots, \psi_k) \Longrightarrow C_{fusion}(\phi)$$

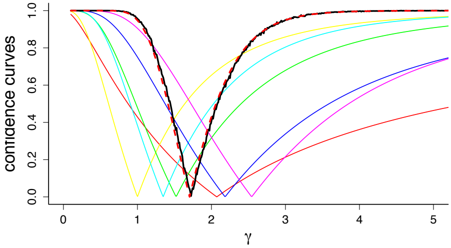
IV: Some other whales

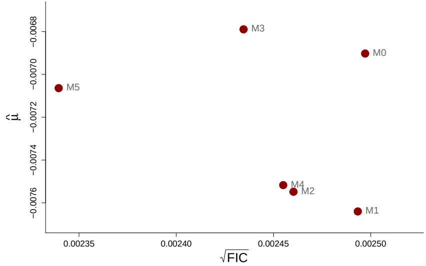
Source 1 informs on ψ_1 the North Atlantic humpback whale abundance in 1995.
Source 2 informs on ψ_2 the North Atlantic humpback whale abundance in 2001.
Both sources only report non-sufficient summaries (3 quantiles).

What can we learn about the annual growth rate?

$$\rho = (\psi_2 - \psi_1)/(6\psi_1)$$




confidence curves
0.0
0.2
0.4
0.6
0.8
1.0
0
1
2
3
4
5
γ


M3
M0
M5
M4
M2
M1
−0.0068
−0.0070
−0.0072
−0.0074
−0.0076
$\hat{\mu}$
0.00235
0.00240
0.00245
0.00250
$\sqrt{\text{FIC}}$