

一般化線形（混合）モデル

GLM & GLMM

前口上

久保拓弥（北海道大・環境科学）

kubo@ees.hokudai.ac.jp

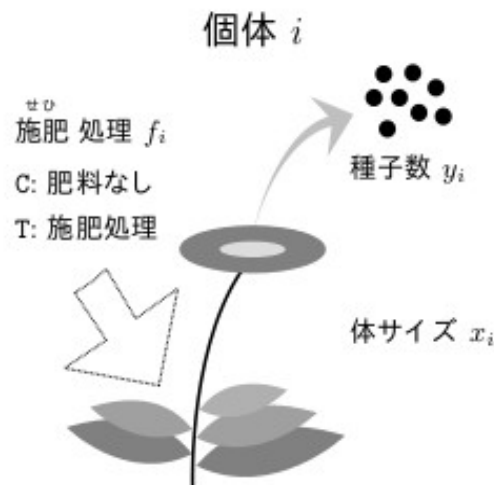
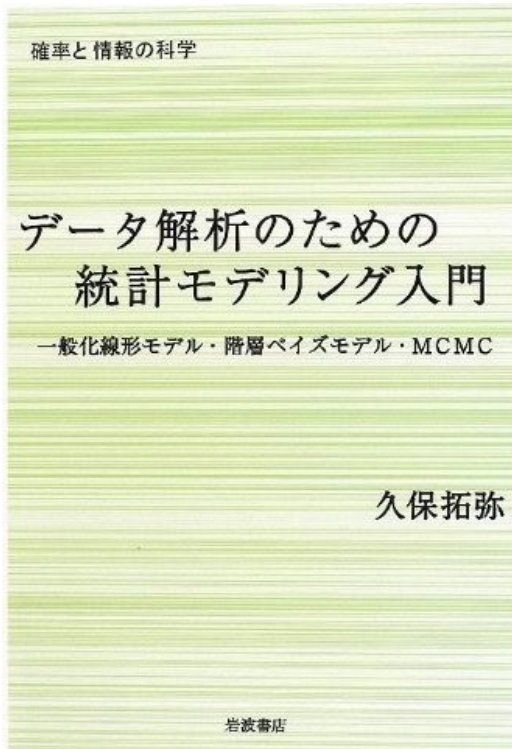


図 3.1 この例題に登場する架空植物の第 i 番目の個体。この植物の体サイズ（個体の大きさ） x_i と肥料をやる施肥処理 f_i が種子数 y_i にどう影響しているのかを知りたい。

この GLM 授業の web page

<http://goo.gl/RrHZeY>

更新: 2013-07-26 13:21:46

生態学のデータ解析 - 大阪大学「一般化線形(混合)モデル」講義 2013

- 2013年8月27日の大阪大学基礎工学部での集中講義 [データ科学特論I](#) 第2日目
- このページの短縮 URL: <http://goo.gl/RrHZeY>
 - 久保の [統計学授業](#) 一覧
- 参考文献: [統計モデリング入門](#)
 - 古い文献: [2008年の講義の一と](#) (北大図書館 HUSCAP, 無料)

(もくじ)

- 1. (10:30-12:00) 確率分布, 最尤推定, ポアソン回帰
- 2. (13:00-14:30) ロジスティック回帰と GLMM
- 3. (14:40-16:10) 階層ベイズモデルと MCMC

1. (10:30-12:00) 確率分布, 最尤推定, ポアソン回帰

- 投影資料
- ポアソン分布説明の例題データ: [data.RData](#)
 - R で読む: `load("data.RData")`
- ポアソン回帰の例題データ: [data3a.csv](#)
- CSV file をうまく読めない場合は R 上で以下の命令を実行してください:
 - `d <- read.csv("http://hosho.ees.hokudai.ac.jp/~kubo/stat/2013/Fig/poisson/data3a.csv")`

2. (13:00-14:30) ロジスティック回帰と GLMM

簡単な自己紹介

- 北大の環境科学院というところで生態学のデータ解析方法についての研究をしています
- 生態学：生物集団のあれこれを調べる生物学の一分野
- 久保は野外調査とかはやらずに他人のデータの解析ばかりやっています



統計モデルは データ解析の道具

なぜデータ解析の方法を
勉強しなければ
ならないのか？

科学のデータ解釈は統計的手法に依存

「データ→結論」のつなぎめ

- データ解析がおかしいと結論もおかしい
- データ解析を悪用して結論をねつぞうできる
- 論文を読むときにデータ解析の部分がわからないと「どうしてこのデータからこの結論が導かれたのか、妥当といえるのか」などがわからない→論文を批判的に読めない

生物学とかでは

データ解析はあまり重視されてなかった

内容がわからなくてもソフトウェアにまるなげ

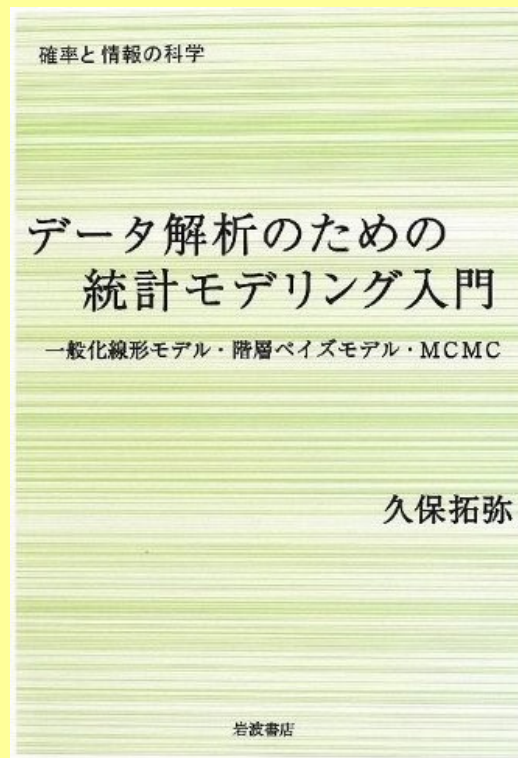
- ブラックボックス統計解析
- とにかく「ゆーい差」さえ出せばよいという発想になっている
- 大学・大学院でもあまりちゃんと教えられていない，教えられるヒトが少ない……とくに近年発達している統計モデリングについて

この授業のねらい

できるだけ内容を理解して統計ソフトウェアを使おう！

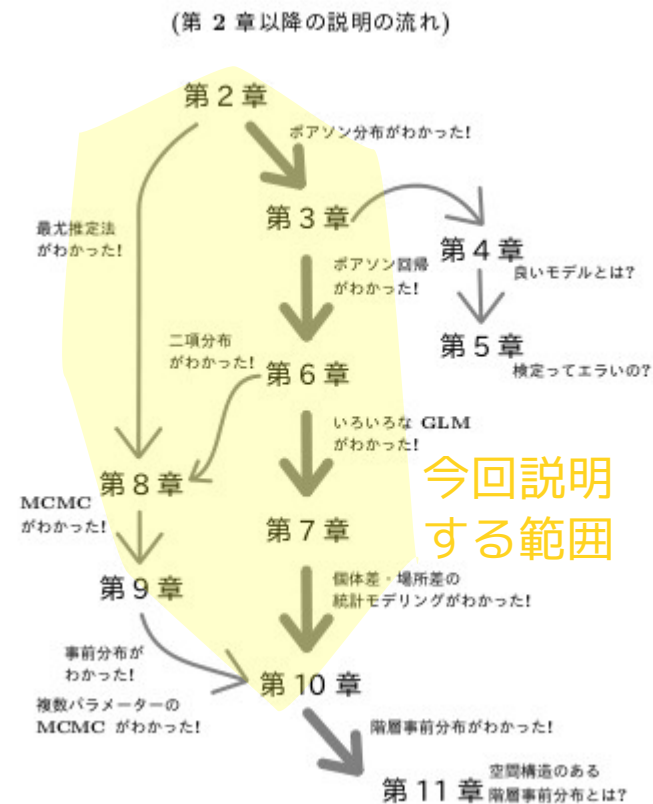
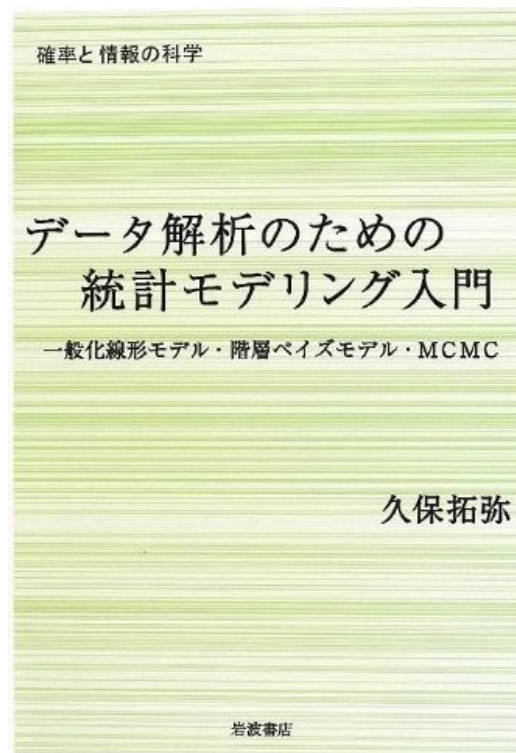
- データ解析で使われるの中でも比較的簡単な統計モデルを理解しよう
- 「ゆーい差」さえ出せばよいという発想をやめて、データと統計モデルの対応関係をよく見よう（作図重要）
- 統計ソフトウェア R を使い始めよう

教科書とソフトウェア



この授業は「統計モデリング入門」 にそった内容を説明します

著者：久保拓弥
出版社：岩波書店
2012-05-18 刊行
価格 3990 円



<http://goo.gl/Ufq2>

「統計モデリング入門」のもとになった「講義のーと」もあります



北海道大学学術成果コレクション
HUSCAP
Hokkaido University collection of Scholarly and Academic Papers
Copyright(c) 2005 Hokkaido University Library, All Rights Reserved.

北海道大学 | 附属図書館 | HUSCAP

検索 Language: 日本語

Hokkaido University Collection of Scholarly and Academic Papers >
環境科学院・地球環境科学研究所 >
雑誌発表論文等 >

フルテキスト

kubostat2008a.pdf	第1回	260.69 kB	PDF	見る/開く
kubostat2008b.pdf	第2回	156.71 kB	PDF	見る/開く
kubostat2008c.pdf	第3回	434.56 kB	PDF	見る/開く
kubostat2008f.pdf	第6回	219.43 kB	PDF	見る/開く
kubostat2008g.pdf	第7回	246.95 kB	PDF	見る/開く
kubostat2008e.pdf	第5回	238.75 kB	PDF	見る/開く
kubostat2008d.pdf	第4回	184.92 kB	PDF	見る/開く

タイトル: 講義のーと : データ解析のための統計モデリング
著者: 久保, 拓弥
キーワード: 生態学
データ解析

授業 web page に「講義のーと」へのリンクがあります! <http://goo.gl/82dgC>

統計ソフトウェア R



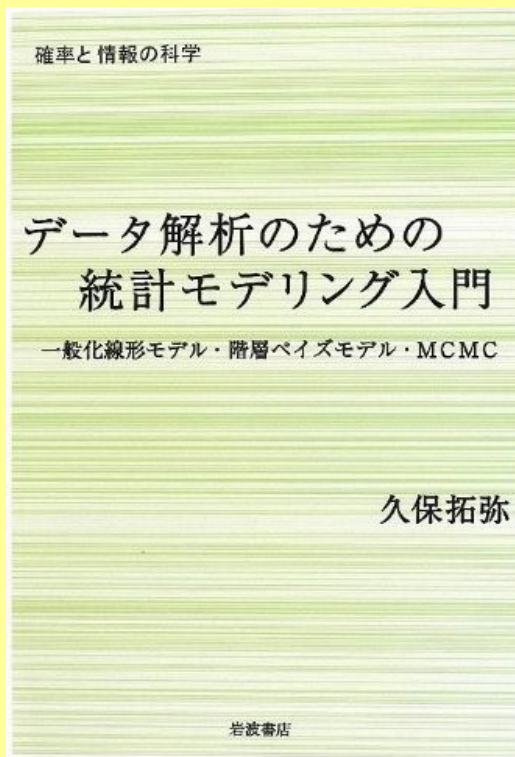
統計学の勉強には良い統計ソフトウェアが必要!

- 無料で入手できる
- 内容が完全に公開されている
- 多くの研究者が使っている
- 作図機能が強力

この教科書でも R を
使って問題を解決する
方法を説明しています



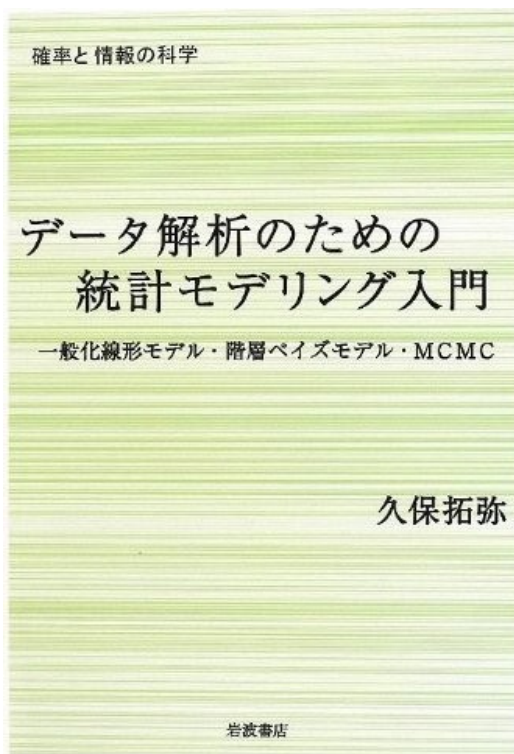
統計モデルとは何か？



「統計モデル」とは何か？

どんな統計解析においても
統計モデルが使用されている

- 観察によってデータ化された現象を説明するために作られる
- 確率分布が基本的な部品であり、これはデータにみられるばらつきを表現する手段である
- データとモデルを対応づける手つづきが準備されていて、モデルがデータにどれぐらい良くあてはまっているかを定量的に評価できる



この授業のながれ

1. 確率分布, 最尤推定, ポアソン回帰
2. ロジスティック回帰と GLMM
3. MCMC と階層ベイズモデル

