

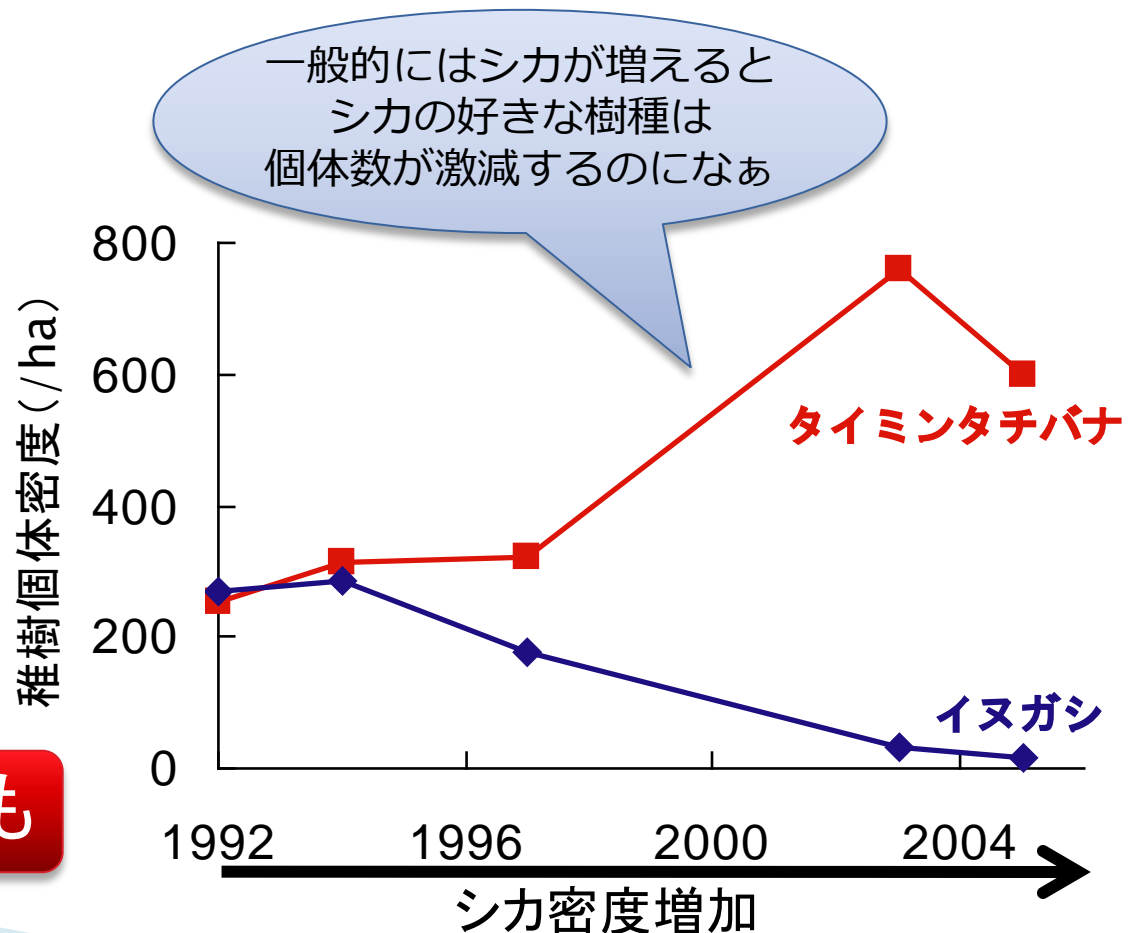
カウントデータをGLM・GLMMで解析

- ▶ 例:『シカが食べると根萌芽が増えるんじゃないか仮説』の検証



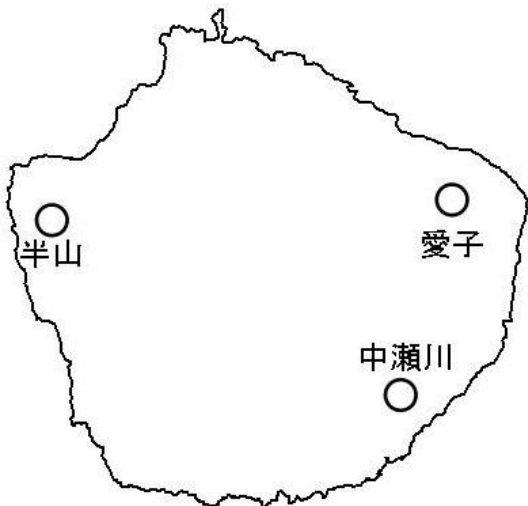
タイミンタチバナの根萌芽

根萌芽が増えたのかも



データ収集・GLMで解析

- ▶ 10×10mの調査区をたくさん作っていろいろ調べる
 - 根萌芽の数
 - シカ
 - その他影響しそうなもの：親木の数，光環境，地形



屋久島(いいところです)

根萌芽の数～シカ＋親木＋光＋地形

カウント
データ

この辺の要因は
考慮すべき
なのかしら??

ポアソン分布でGLM＋モデル選択だ！



調査地の違いは？GLMMで釈明

- ▶ シカ密度の違う3ヶ所に10数個ずつ調査区を作った
 - データが入れ子(nested)状になっている



『疑似反復 (pseudo-replication) とちゃうの？』
とかいう“いちゃもん”が付きそう…



GLMMで調査地の違いを考慮しよう！

というか
ちゃんと考慮したよ
って言い訳しよう

GLMMの関数とパッケージ



▶ いろいろある。それぞれ長所と短所がある・・・のか？

- glmmML(glmmML)
- lmer(lme4)
- negbin(aod)
- lme(nlme)
- glmm(glmm)
- glmmPQL(MASS) etc...

* 関数(パッケージ)

glmmMLとlmerしか
使ったことがないので
他は誰か解説して下さい

▶ glmmML

- 混合効果は一つだけ扱える
- ポアソン分布・二項分布が扱える
- stepAICが使える
→モデル選択が簡単
- 用途は狭いが融通が効く

▶ lmer

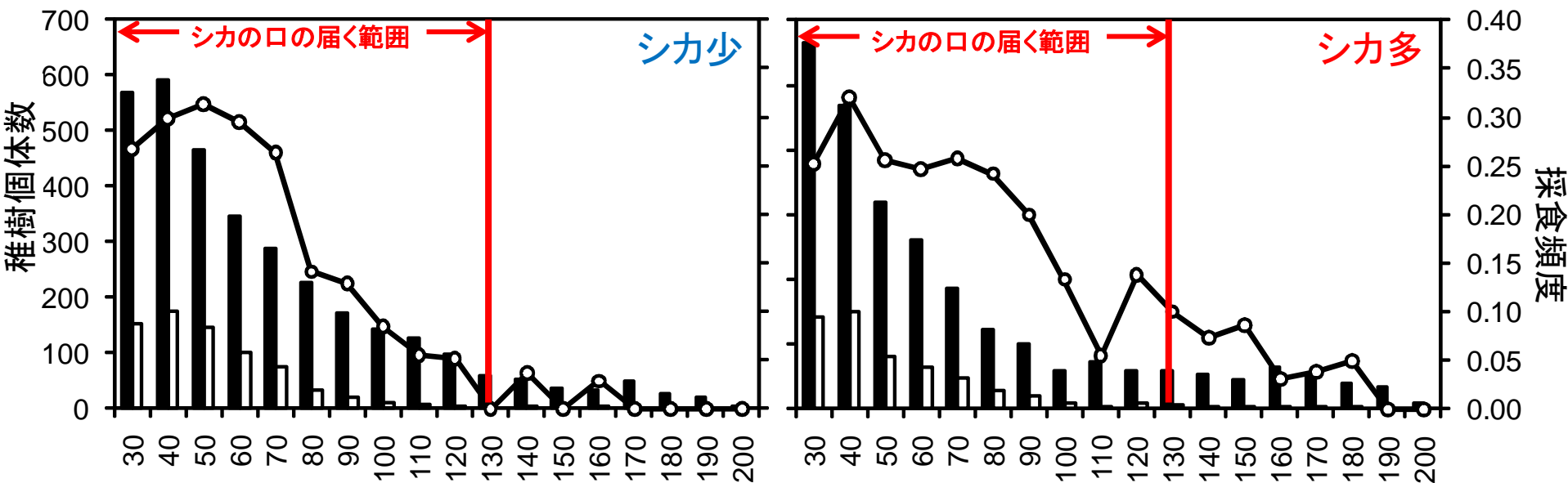
- 混合効果は複数指定できる
- “ランダム傾き”という良く分からんものも考慮することができる
- いろんな分布型が使える
- 動きが重い、無理させると落ちる
- stepAICは使えない

この2つで実践！

あり/なしデータをGLMで解析

- ▶ 例:『シカは小さい木がお好き?』仮説の検証

全体での採食頻度はほぼ同じ2か所だけど...



高さクラスごとにまとめた値での解析

個体単位での解析(データの有効活用)

二項分布を使ったGLM!



今さらながら、結果の解釈のしかた

▶ bite~height*site (siteはAIKOとHNYMの2カテゴリ)

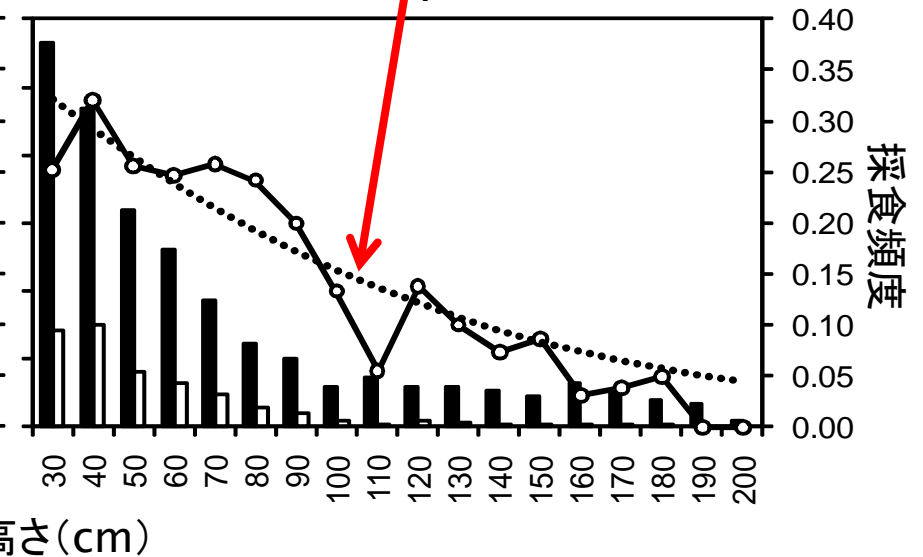
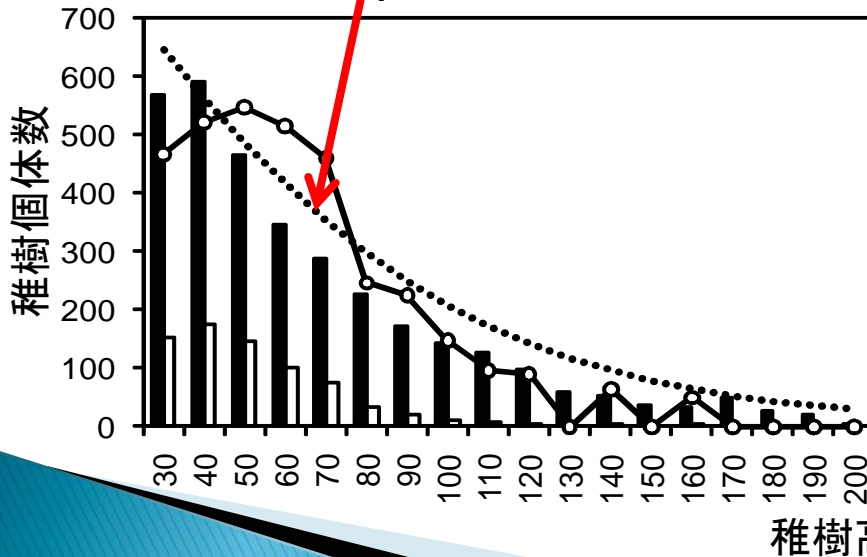
	Estimate	Pr(> z)
(Intercept)	0.090065	0.4034
height	-0.020882	< 2e-16 ***
siteHNYM	-0.429934	0.0171 *
height:siteHNYM	0.007225	0.0075 **

AIKO: bite = $-0.021 \times \text{height} + 0.09$

HNYM: bite = $-0.014 \times \text{height} - 0.34$

AIKO: $1 / (1 + \exp(0.021 \times \text{高さ} - 0.09))$

HNYM: $1 / (1 + \exp(0.014 \times \text{高さ} + 0.34))$



おまけ: offset項を使おう

▶ 割り算してたデータも有効活用しよう

- 密度, 優占度, 確率...

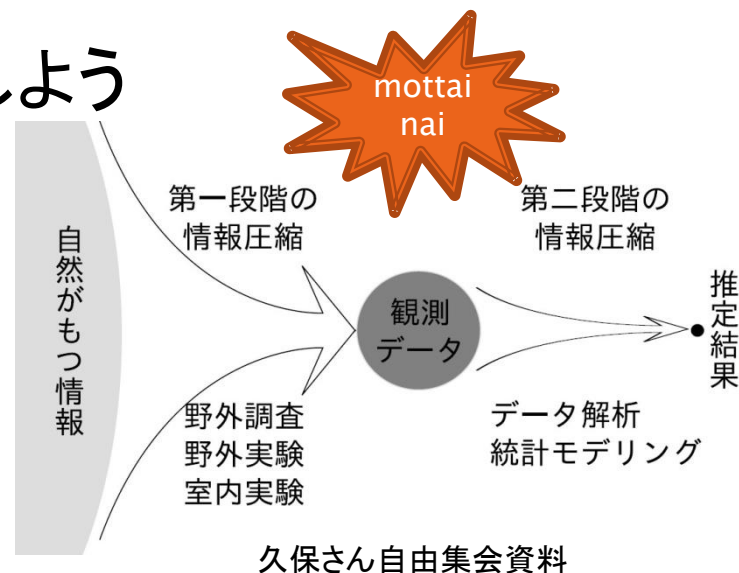
▶ 割り算を避ける方法

◦ 確率の場合(N個のうちのk個)

- 二項分布でGLM

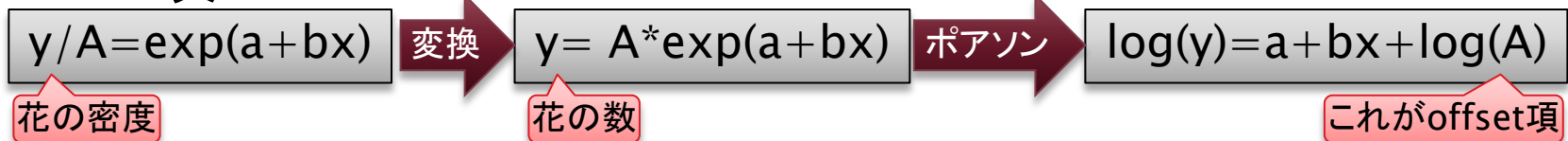
◦ 密度や優占度など(分母もデータ)

- offset項を使用+ポアソン分布でGLM



▶ offset項

例えば...面積(A)あたりの花の数(y)は光環境(x)で説明できるのか?



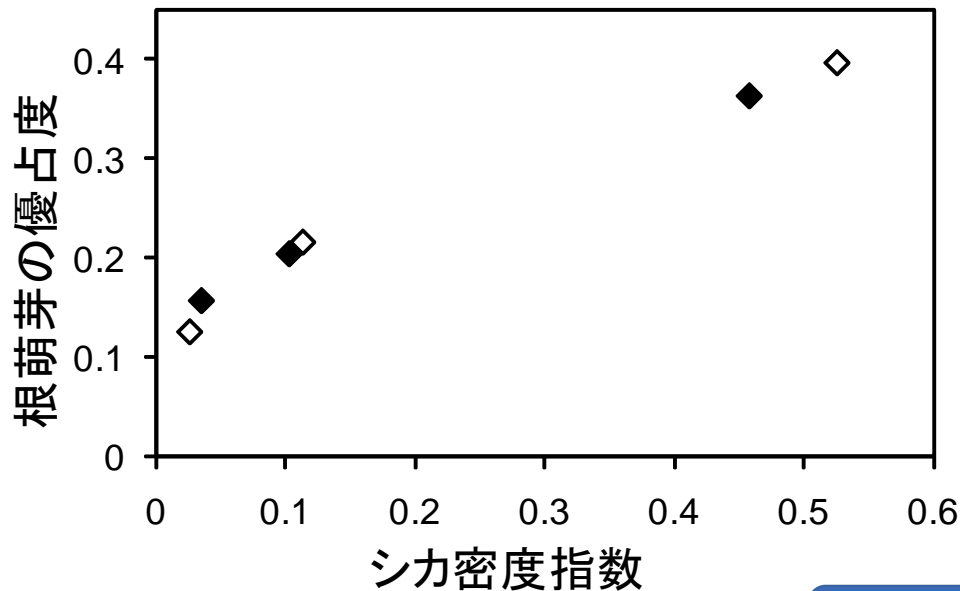
* 詳しいことは北大の久保さんの講義ノートを見てください

やってみる



(調査年の違いを含むGLMMというおまけ付き)

▶ 例: 『シカが多いところは根萌芽だらけ』仮説の検証



普通は... 優占度~シカ+(年)

根萌芽数 / 全個体数



offset項を利用

根萌芽数~シカ+offset(全個体数)+(年)

重ね重ね...

* 詳しいことは北大の久保さんの講義ノートを見てください