

数学を用いる生物学：理念・概念と実践・方法論

7月20日(水)

10:00 - 10:30 開催趣旨：「個体群行列モデルから見える数理モデルと統計モデル」島谷健一郎(統数研)

10:30 - 11:50 「マメゾウムシ-寄生蜂3種系におけるカオス動態、スイッチング捕食行動、学習規定物質」嶋田正和(東京大・産総研)

12:50 - 14:10 「生態系の情報処理能力の定量化とその利用に向けた挑戦」潮雅之(京都大白眉セ)

14:20 - 15:40 「細胞生物学研究における統計的解析の質を高める」木村暁(遺伝研)

15:50 - 17:10 「モデルとは何か～メタな視点のモデル論」松王政浩(北海道大)

17:10 - 17:50 「繁殖干渉による進化が群集構造に与える影響に関する量的遺伝学モデルの解析」森田慶一(総研大)

18:45- Free discussion session

7月21日(木)

10:00 - 10:30 ワークショップ2日目：「空間点過程に見られる数理生態学と統計的推定」島谷健一郎(統数研)

10:30 - 11:10 「個体の視点に基づく点パターンダイナミクスの導出の仕方」高須夫悟(奈良女子大)

11:10 - 11:50 「数理モデルを用いた混獲の定量的評価」井上巨人(神戸大)

12:50 - 13:30 「水産学における空間統計モデルの利用について」井嶋浩貴(水産研究・教育機構)

13:30 - 14:10 「Agrochaology - Chaos, Synchronization and Networks-」酒井憲司(農工大)

14:20 - 15:00 「粘菌ネットワーク形成過程の位相的データ解析の試み」本武陽一(統数研)

15:00 - 15:40 「湖沼堆積物を用いた、より正確な長期生態動態の推定を目指して」大竹裕里恵(東北大)

15:40 - 16:10 「アリにおける効率的な液体の輸送について」藤岡春菜(岡山大)

16:10 - 16:40 「宿主内での突然変異ウイルスの確率的絶滅・薬剤耐性出現・交差免疫回避：時間依存分枝過程によるアプローチ」林玲奈(九州大)

開催趣旨:「個体群行列モデルから見える数理モデルと統計モデル」
島谷健一郎(統数研)

生き物は数式に従って生きているわけではない。
だから、数学と生物学は相性が悪い(発言者不詳)。

定性的仮説はともかく、**定量的予測や推定で数学は不可欠**。
ところで、野外や実験研究者が数学の必要性を感じたとき

統計学か数理生物学か？ どちらの研究者と相談/共同研究？

たくさんある統計モデル、数理モデルの、どれから学習をする？

数理生物(や統計学)が専門の人は、野外や実験研究者と(将来)どう対応する？

同じ「数学を使う生物学」でも、使い方以前に使う理由・発想が異なる

数理モデル:少数の基本**原理**を数式で表現し、**演繹**的に多様な生態系の現象を説明する。

統計モデル:データ生成過程を確率分布を含む数式で表現し、**帰納**的に未知数量を**推定**し、**予測**を行う。

鳥類A

成鳥メスは春に3個の卵

ヒナは3年で成鳥、4年目から繁殖

成鳥までの生残率: 年40%の

成鳥以降の生残率: 95%

鳥類Aは絶滅? どんどん増殖? 今の個体数を保つ(安定的)?

鳥類B

成鳥のメスは春に6個の卵

ヒナは5年で成鳥、6年目から繁殖

成鳥までの生残率: 75%

成鳥の生残率: 25%

・どちらがより早く増殖あるいは絶滅に向かう?

・鳥類Aを増殖させるには,

1. 生まれたヒナや幼鳥の生残率を高める

2. 繁殖している成鳥の生残率を高める

どちらが有効だろう? 鳥類Bではどうか?

個体群行列モデルという
実例で違いを見る

統計スポットライト・シリーズ 4

編集幹事 島谷健一郎・宮岡悦良

個体群生態学と 行列モデル

統計学がつかなく野外調査と数理の世界

島谷健一郎・高田壮則 著

鳥類A

成鳥メスは春に3個の卵(メスは半分の1.5)

ヒナは3年で成鳥、4年目から繁殖

成鳥までの生残率: 年40%の

成鳥以降の生残率: 95%

鳥類Aは絶滅? どんどん増殖? 今の個体数を保つ(安定的)?

鳥類B

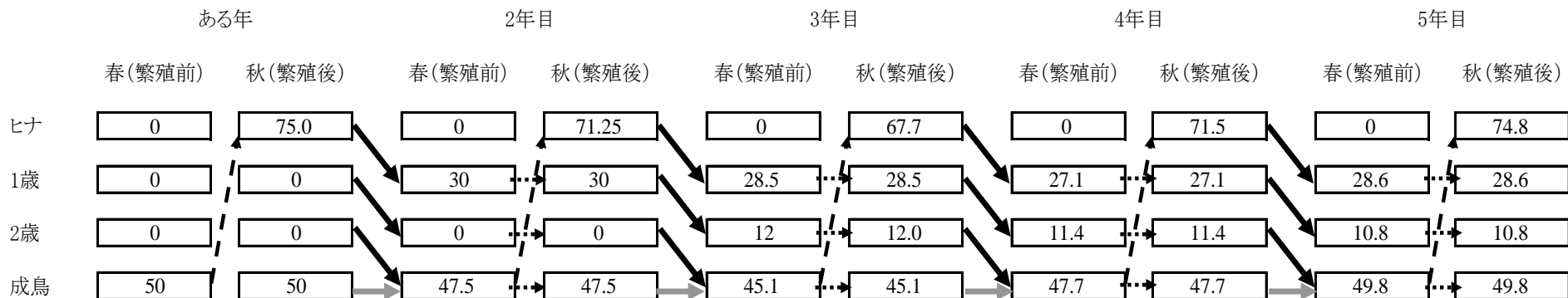
成鳥のメスは春に6個の卵(メスは3)

ヒナは5年で成鳥、6年目から繁殖

成鳥までの生残率: 75%

成鳥の生残率: 25%

鳥類A 図にしても、ややこしくてよくわからない...



ベクトルと行列で表すと、見通しがよくなる

$t + 1 \text{ 年}$
 $\begin{pmatrix} \text{ヒナの個体数} \\ \text{1歳の個体数} \\ \text{2歳の個体数} \\ \text{成鳥の個体数} \end{pmatrix}$

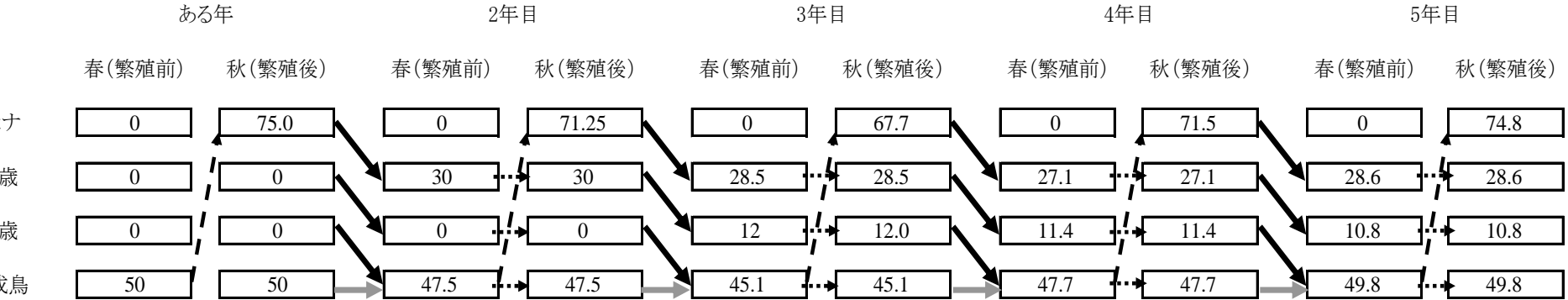
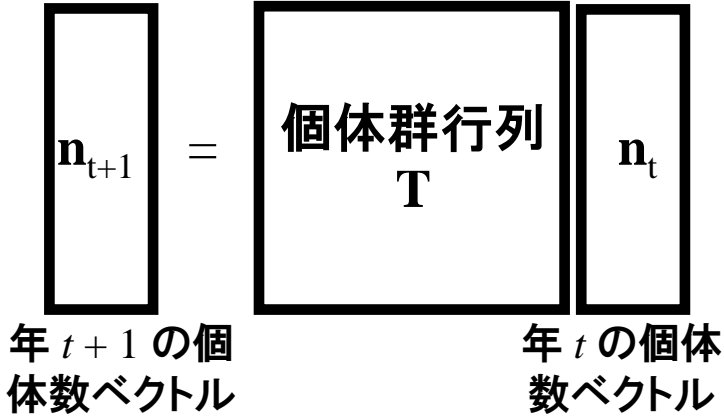
$=$

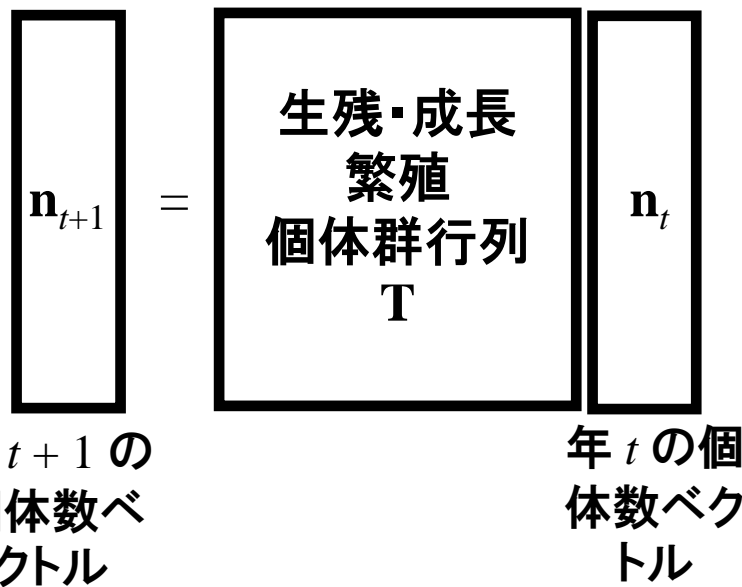
$\begin{pmatrix} 0 & 0 & 0 & 1.5 \\ 0 & 1 & 0 & 0 \\ 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 1 \end{pmatrix}$ \cdot $\begin{pmatrix} 0 & 0 & 0 & 0 \\ 0.4 & 0 & 0 & 0 \\ 0 & 0.4 & 0 & 0 \\ 0 & 0 & 0.4 & 0.95 \end{pmatrix}$

$t \text{ 年}$
 $\begin{pmatrix} \text{ヒナの個体数} \\ \text{1歳の個体数} \\ \text{2歳の個体数} \\ \text{成鳥の個体数} \end{pmatrix}$

個体群行列

$$\mathbf{T} = \begin{pmatrix} 0 & 0 & 0 & 1.5 \\ 0 & 1 & 0 & 0 \\ 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 1 \end{pmatrix} \cdot \begin{pmatrix} 0 & 0 & 0 & 0 \\ 0.4 & 0 & 0 & 0 \\ 0 & 0.4 & 0 & 0 \\ 0 & 0 & 0.4 & 0.95 \end{pmatrix} = \begin{pmatrix} 0 & 0 & 0.6 & 1.425 \\ 0.4 & 0 & 0 & 0 \\ 0 & 0.4 & 0 & 0 \\ 0 & 0 & 0.4 & 0.95 \end{pmatrix}$$





ある程度の年数を経ると
 $\mathbf{n}_{t+1} = \mathbf{T} \mathbf{n}_t \approx c \mathbf{n}_t$ が成立
 c : 行列Tの**固有値**
個体群成長率と解釈できる

鳥類A: $c = 1.039$ 増加
 鳥類B: $c = 0.990$ 減少

感度分析: 行列成分による固有値の偏微分係数:
 行列要素の(微小)変化が個体群成長率に与える影響

鳥類A

$$S = \begin{pmatrix} -- & -- & 0.011 & 0.049 \\ 0.190 & -- & -- & -- \\ -- & 0.190 & -- & -- \\ -- & -- & 0.174 & 0.781 \end{pmatrix}$$

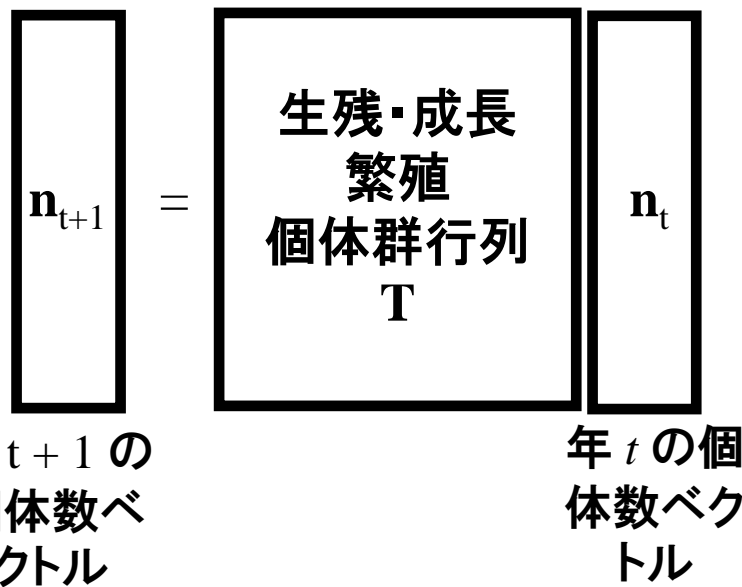
成鳥の生残率

鳥類B

$$S_B = \begin{pmatrix} -- & -- & -- & -- & 0.062 & 0.062 \\ 0.247 & -- & -- & -- & -- & -- \\ -- & 0.247 & -- & -- & -- & -- \\ -- & -- & 0.247 & -- & -- & -- \\ -- & -- & -- & 0.247 & -- & -- \\ -- & -- & -- & -- & 0.062 & 0.063 \end{pmatrix}$$

幼鳥の生残率

増殖には、鳥類Aでは成鳥の生残率を高める、鳥類Bでは幼鳥の生残率を高める、が有効



ある程度の年数を経ると

$$\mathbf{n}_{t+1} = \mathbf{T}\mathbf{n}_t \approx c\mathbf{n}_t \text{ が成立}$$

c : 行列 T の**固有値**
個体群成長率と解釈できる

鳥類A: $c = 1.039$ 増加
 鳥類B: $c = 0.990$ 減少

感度分析: 行列成分による固有値の偏微分係数:
 行列要素の(微小)変化が個体群成長率に与える影響

固有値という数学の概念と、ペロン・フロベニウスの定理
 によって、その属性に応じて生残率や繁殖が異なる個体
 が混ざる個体群でも、個体群成長率という生態学の概念
 を定量的に表現し、かつ、計算法まで示してくれる。
 その偏微分係数は保全生態学への示唆も与える。

増殖には、鳥類Aでは成鳥の生残率を高める、鳥類Bでは幼鳥の生残率を高める、が有効

鳥類A

成鳥メスは春に3個の卵

ヒナは3年で成鳥、4年目から繁殖

成鳥までの生残率: 年40%の

成鳥以降の生残率: 95%

鳥類Aは絶滅? どんどん増殖? 今の個体数を保つ(安定的)?

鳥類B

成鳥のメスは春に6個の卵

ヒナは5年で成鳥、6年目から繁殖

成鳥までの生残率: 75%

成鳥の生残率: 25%

統計スポットライト・シリーズ 4

編集幹事 島谷健一郎・宮岡悦良

個体群生態学と 行列モデル

統計学がつかなく野外調査と数理の世界

島谷健一郎・高田壮則 著

ところで、生残率はどうやったらわかる?

野外生物については、野外データから推定する。

標識調査

個体を捕獲し、標識を付けて解放、1年後、再発見調査

標識付き個体を発見	→	生きている
発見できなかった	→	死亡した？

生きていたけど発見に失敗したのかもしれない...
死んじゃったのか、発見できなかったのか、わからない...

標識調査では生残率(何%が生残するか)はわからない...?

生残率の推定には、

1. 野外調査を継続
2. 統計モデル

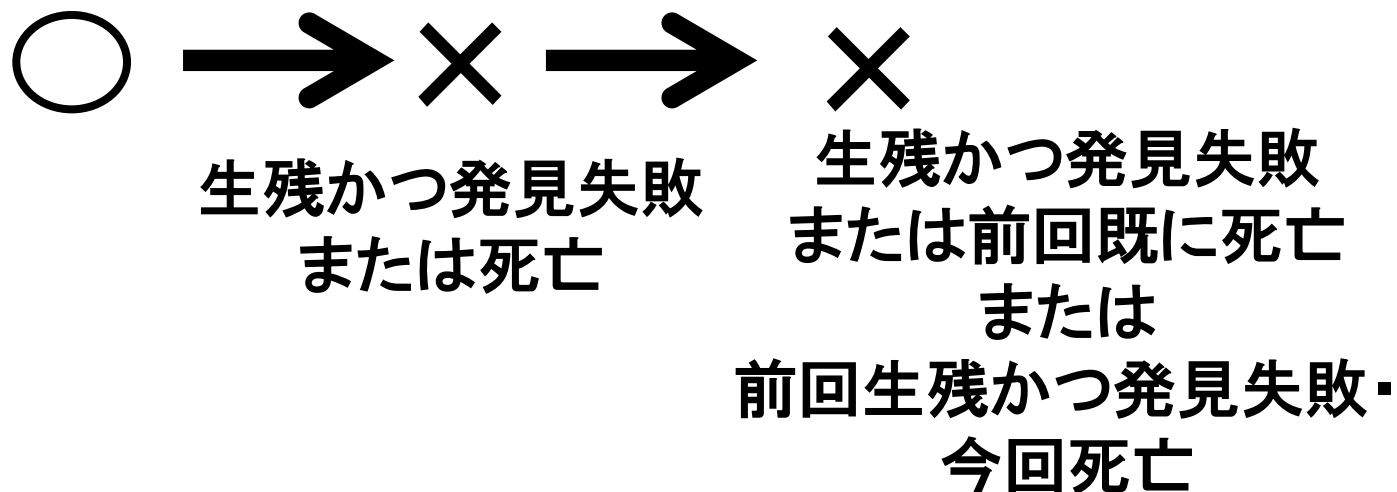
の2つが必要(どちらか一方だけではダメ)

調査を継続

確かな情報が得られる

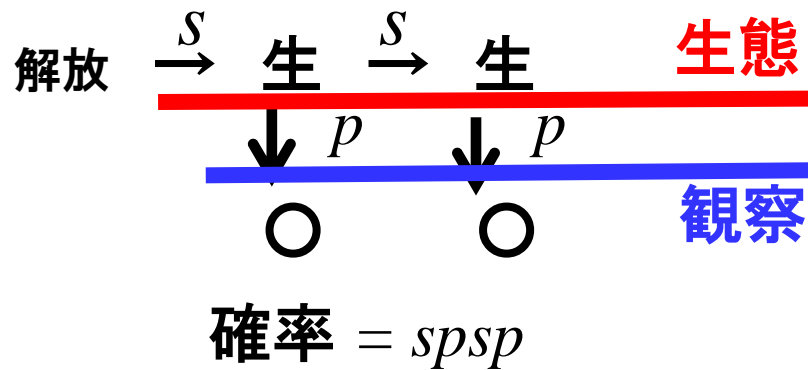


不確かな情報もある

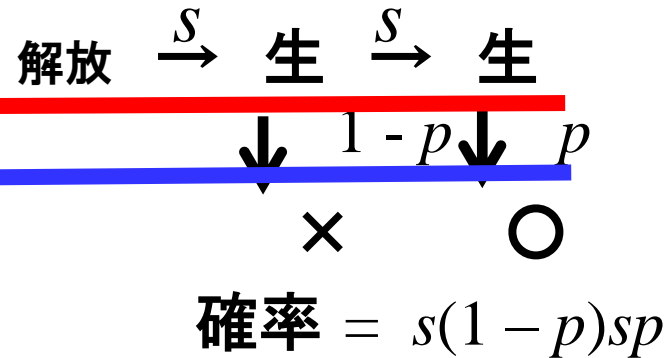


推定したい生残率と発見率を未知パラメータとする統計モデル
生死という生態的過程と、**発見という観察過程**を階層化した
状態空間モデル(のひな型)

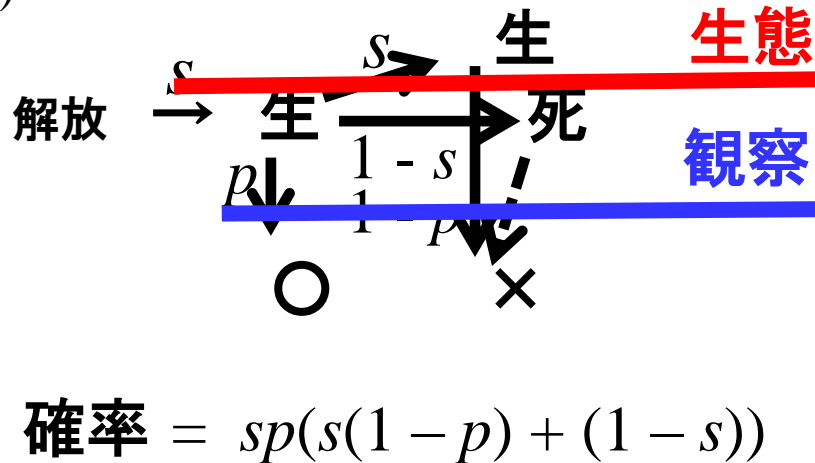
(a)



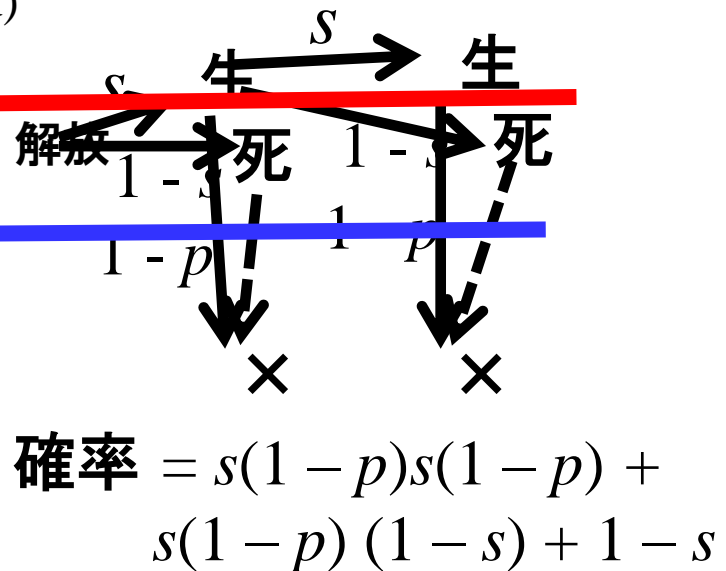
(b)



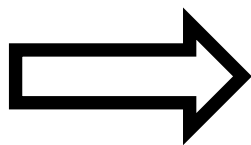
(c)



(d)



標識番号	調査1	調査2
1	○	×
2	○	○
3	○	○
4	○	×
5	○	○
6	○	×
7	○	○
8	○	○
9	×	○
10	×	×
11	○	○
12	×	○
13	×	×
14	×	○
15	○	×
⋮	⋮	⋮



パターン	データ数	確率
○○	38	$spsp$
×○	14	$s(1-p)sp$
○×	18	$sp(s(1-p) + (1-s))$
××	30	$s(1-p)s(1-p) +$ $s(1-p)(1-s) + 1-s$

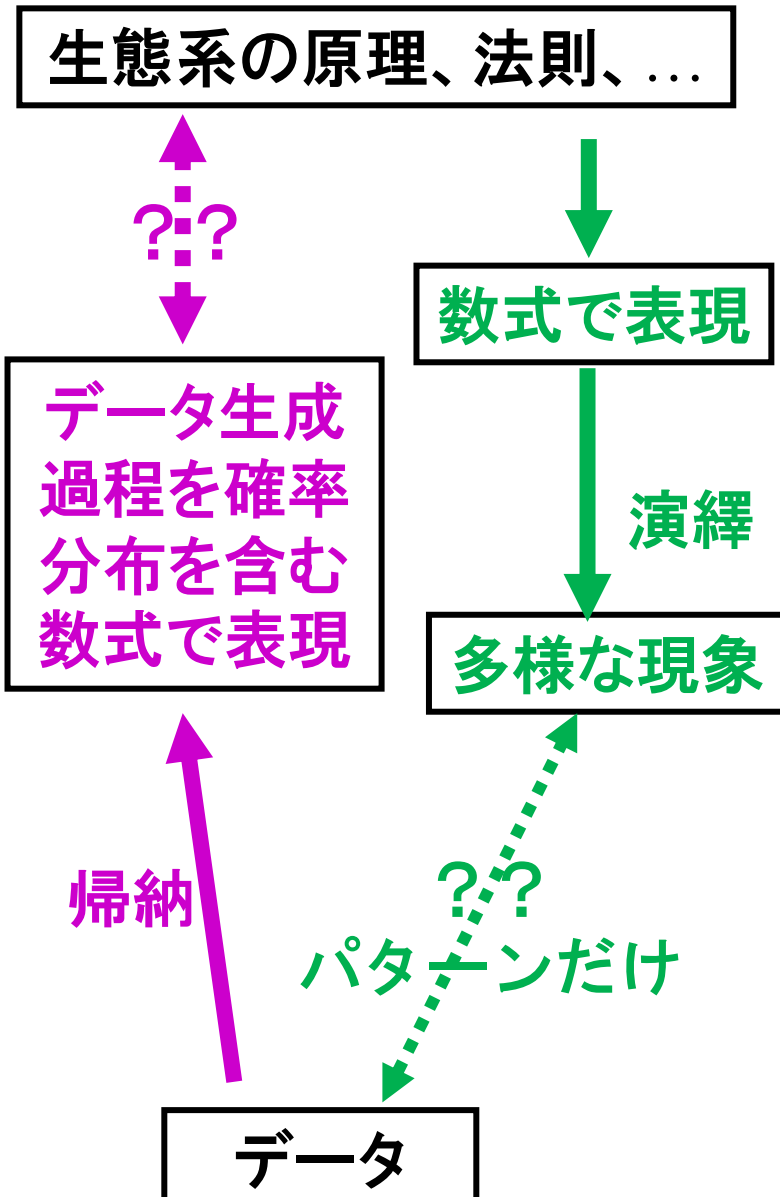
左のようなデータを集約し、データが得られる確率(尤度)

$$(spsp)^{38} \cdot (s(1-p)sp)^{14} \cdot (sp(s(1-p) + 1-s))^{18} \cdot (s(1-p)s(1-p) + s(1-p)(1-s) + 1-s)^{30}$$

が最大となる生残率 s と発見率 p を数値計算(最尤法)

数理生物学と生物統計

数学を使うときの、使い方が違う



数理モデル: 少数の基本**原理**を数式で表現し、**演繹**的に多様な生態系の現象を説明する。

統計モデル: データ生成過程を確率分布を含む数式で表現し、**帰納**的に未知数量を**推定**し、**予測**を行う。

演繹(えんえき):前提が正しければ結論は必ず正しい

AならばB
Aである

Bである

AならばB
Bでない

Aでない

検証は演繹
ではない

~~AならばB~~
~~Bである~~

~~Aである~~

科学哲学者 Karl Popper

科学とは、**反証**可能な仮説や理論の中で、これまで反証されていないものを一時的に受け入れる営み。帰納は科学の方法ではない。

帰納推論

過去のデータから予測

例

一昨日昼の学食は混んでいた
昨日の昼の学食も混んでいた
⋮

今日の昼も混むだろう

未来は過去に似ている

自然の一樣性原理

この前提がないと帰納は正しくない

Humeの問題:帰納は原理的に正当化できない

演繹の真理保存性

- ・前提が正しければ結論が必ず正しい。
- ・結論の正しさはすべて前提の正しさで決まる。
- ・結論には前提以上の内容は含まれておらず、新しい情報が得られない。

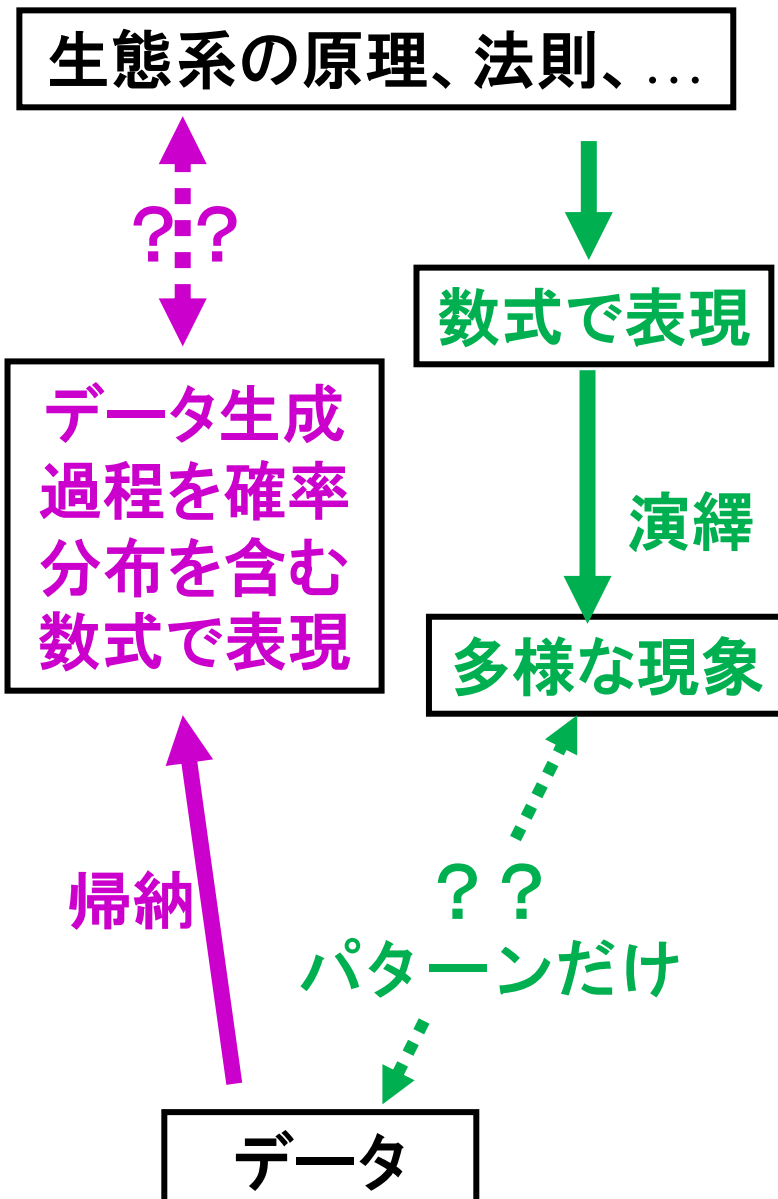
帰納の利点

- ・帰納では演繹と異なり、結論に前提以上の内容が含まれる。結論では前提にないことが言える。
- ・それはゴマカシているからではある。
- ・ゴマカシだから即ダメというわけではなく、うまくゴマカシて、豊富な内容を引き出すことができる。

- ・帰納の欠陥を補うために統計という数学(演繹)をデータに基づく帰納推論に加える。
- ・基本的に帰納推論なので、どこかにゴマカシがある。それが統計を使うときに伴う違和感やうしろめたさの原因(のひとつ)。

数理生物学と生物統計:

数学を使うときの、使い方の発想から異なる



数理モデル: 少数の基本原則を数式で表現し、**演繹**的に多様な生態系の現象を説明する。

演繹(数学の定理・証明)と、理論が予測するパターンがデータと矛盾していたら反証

推論は明瞭(仮説演繹法)

生態系の基本法則、生物多様性の根本原理を明らかにするという目標意識も伴う

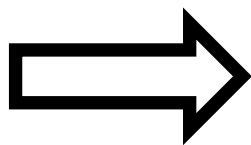
統計モデル: データ生成過程を確率分布を含む数式で表現し、**帰納**的に未知数量を推定し、予測を行う。

帰納推論の中に統計という演繹(数学)が混ざり推論の流れが不明瞭

統計が嫌われる理由

しょせん予測と推定で終わり?

標識番号	調査1	調査2
1	○	×
2	○	○
3	○	○
4	○	×
5	○	○
6	○	×
7	○	○
8	○	○
9	×	○
10	×	×
11	○	○
12	×	○
13	×	×
14	×	○
15	○	×
⋮	⋮	⋮



パターン	データ数	確率
○○	38	$spsp$
×○	14	$s(1-p)sp$
○×	18	$sp(s(1-p) + (1-s))$
××	30	$s(1-p)s(1-p) +$ $s(1-p)(1-s) + 1-s$

左のようなデータを集約し、データが得られる確率(尤度)

$$(spsp)^{38} \cdot (s(1-p)sp)^{14} \cdot (sp(s(1-p) + 1-s))^{18} \cdot (s(1-p)s(1-p) + s(1-p)(1-s) + 1-s)^{30}$$

が最大となる生残率 s と発見率 p を数値計算(最尤法)

データから統計モデルで推定して数理モデルに必要な数値を得る

統計 曼荼羅（島谷版の全体像）

2. データを観る工夫

1. 統計モデルによる予測・推定

グラフ、表、集約、
特徴量

データを自在に加
工して図示

尤度、最適化、乱数生成、
シミュレーション、...
一般化線形モデル

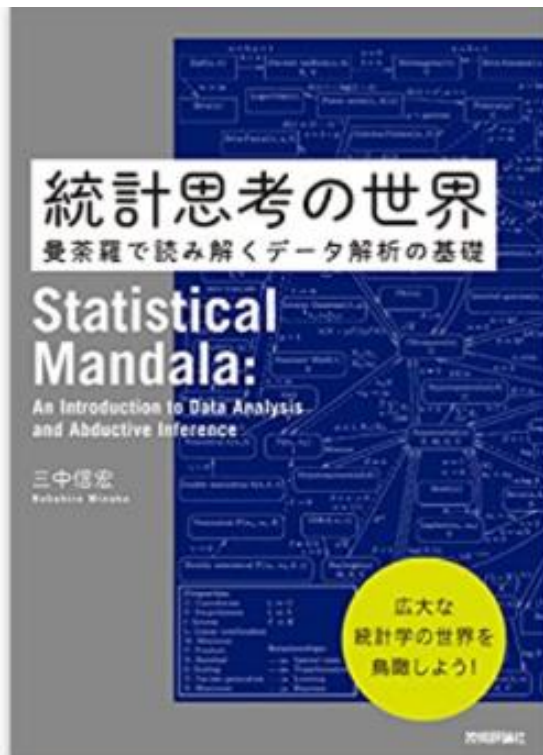
検定

統計解析の結果
(output)を観る

情報量規準
ベイズ統計
機械学習
...

いい「予測」モデルを作って満足できたのは2010年ころまで。
計算完了後に何をする？生物学的考察？統計の結果の分析？

解決法2 曼荼羅



数学を用いる生物学へのふたつめの障壁： 基礎から発展へ順に積み上げる学習法は不向き？

(少しの)基礎を終えたら

曼荼羅(描き手の主観が入った地図?)を手に旅をする…???



統計学者の世界 ～豊田國で読み解くデータ解析の基礎

統計 曼荼羅（島谷版の全体像）

2. データを観る工夫

1. 統計モデルによる予測・推定

グラフ、表、集約、
特徴量

データを自在に加
工して図示

基礎

ギャップ

先端

統計解析の結果
(output)を観る

尤度、最適化、乱数生成、
シミュレーション、...
一般化線形モデル

情報量規準
ベイズ統計
機械学習
...

いい「予測」モデルを作って満足できたのは2010年ころまで。
計算完了後に何をする？生物学的考察？統計の結果の分析？

統計 曼荼羅（島谷版の全体像）

理工学の数学（解析学）

微積分、線形代数

微分方程式
フーリエ解析
複素関数
ベクトル解析

先端理工学に必要な数学

1. 統計モデルによる予測・推定

基礎

ギャップ

先端

尤度、最適化、乱数生成、
シミュレーション、...
一般化線形モデル

情報量規準
ベイズ統計
機械学習

- ・現時点で先端統計とのギャップを埋める教育カリキュラムは未確立
- ・一般化線形モデルを(つまらない)統計学の授業で学習。卒論、修論などで**その限界を体験的に理解する**。
- ・自分のデータ・課題と相性のいい先端と、無理してつなぐ... ???

- ・数学を用いる生物学では、(本来相性の悪い?)数学と生物学をどうつなぐかが鍵
- ・今回:様々な事例を観ることから始める(で鍛える)
- ・参加希望全学生に旅費支援しようと思いました。
- ・参加教員にはご自身の研究費による来所を申し出てくださった方もいて、重ねて感謝します。
- ・参加学生の皆様も、旅費総額を予算枠に収めるべく、協力してくれました。
- ・ワークショップや若手の会などは自費(自分の研究費)参加が普通
- ・特定技能の講習(画像処理ソフト、ドローン、Rによる統計、...)には講習料を払っての参加も多い。
- ・統計(生物学と数学のつなげ方?)のような**非専門科目**に、遠隔地から自費参加する学生は(たぶん)いない
- ・それは、何万円も使う見返り(目標や収穫)が定かでないから。??
- ・必要経費(交通費+安宿の宿泊費)の支援がないと、統計(や関連諸概念、発想、...)を使う諸科学は普及しない

・数学を用いる生物学では、(本来相性の悪い?) **数学と生物学をどうつなぐか**が鍵

・今回: 様々な事例を観ることから始める(で鍛える)

概して

・基礎の段階は、習得に時間がかかる。自身の進歩を自覚する機会も少ない。強制されることも重要。専門科目では耐えられる。

・**非専門科目**では、基礎の習得がさらに苦しい

・**統計**と **academic writing** (英語論文書き)がその典型

・一つの専門を確立することが優先

・専門とは?

誰よりも多くの失敗経験と少しの成功経験(論文)を有する分野

・統計(生物学と数学のつなげ方?)のような**非専門科目**に、遠隔地から自費参加する学生は(たぶん)いない

・それは、何万円も使う見返り(目標や収穫)が定かでないから。??

・必要経費(交通費+安宿の宿泊費)の支援がないと、統計(や関連諸概念、発想、...)を使う諸科学は普及しない

数学を用いる生物学：理念・概念と実践・方法論

7月20日(水)

10:00 - 10:30 開催趣旨：「個体群行列モデルから見える数理モデルと統計モデル」島谷健一郎(統数研)

10:30 - 11:50 「マメゾウムシ寄生蜂3種系におけるカオス動態、スイッチング捕食行動、学習規定物質」嶋田正和(東京大・産総研)

12:50 - 14:10 「生態系の情報処理能力の定量化とその利用に向けた挑戦」潮雅之(京都大白眉セ)

14:20 - 15:40 「細胞生物学研究における統計的解析の質を高める」木村暁(遺伝研)

15:50 - 17:10 「モデルとは何か～メタな視点のモデル論」松王政浩(北海道大)

17:10 - 17:50 「繁殖干渉による進化が群集構造に与える影響に関する量的遺伝学モデルの解析」森田慶一(総研大)

18:45- Free discussion session