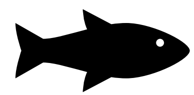
日本行動計量学会セミナー

線形モデルの理論とRを用いた分析事例

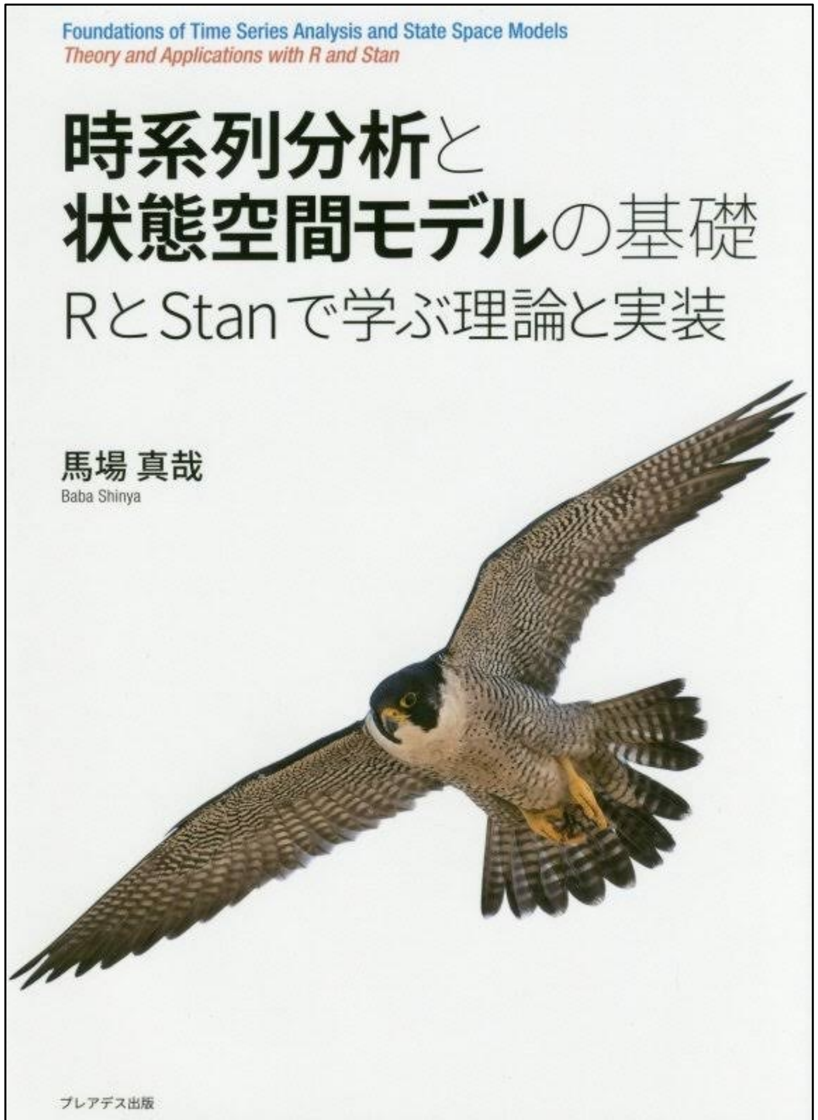
以下に、資料を配置する予定です。
https://logics-of-blue.com/bms-seminar-2025/

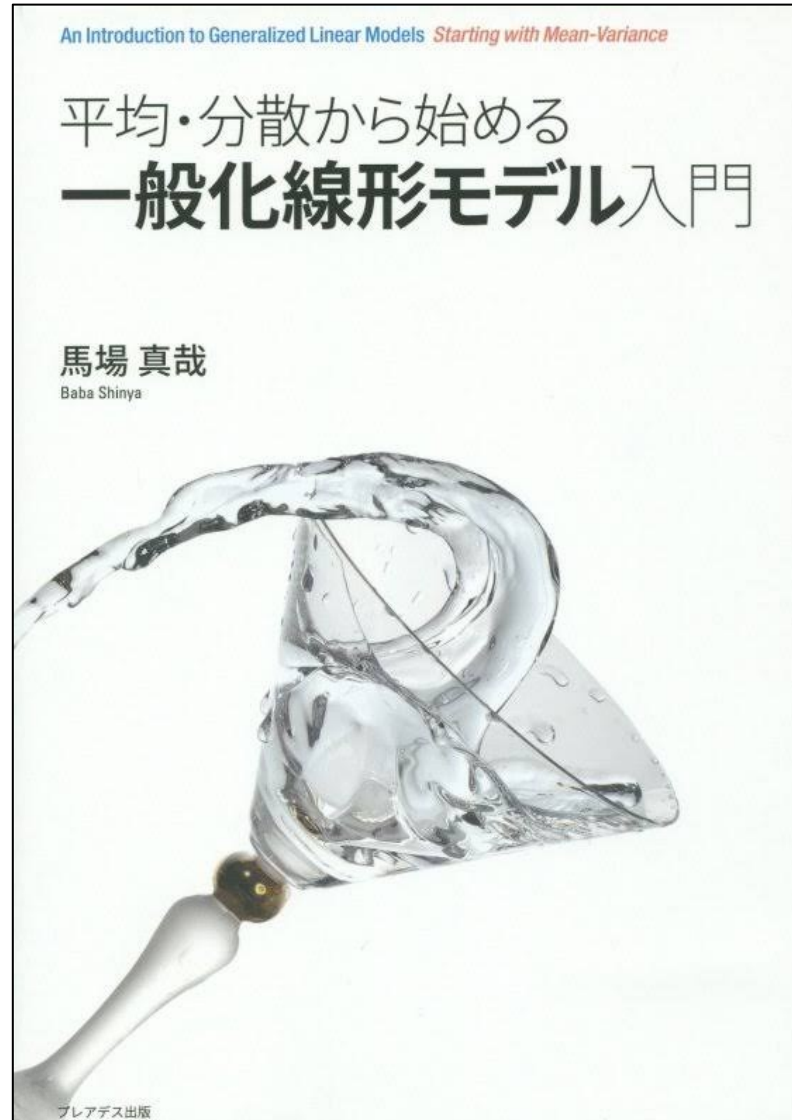
自己紹介・セミナー概要

自己紹介

名前	馬場真哉
学生の頃の専門	水産学 
前職	システムエンジニア
現職	フリーランス 東京科学大学非常勤講師 北海道大学非常勤講師 帝京大学特任講師

本も書いています


Foundations of Time Series Analysis and State Space Models
Theory and Applications with R and Stan
時系列分析と状態空間モデルの基礎
RとStanで学ぶ理論と実装
馬場 真哉
Baba Shinya
プレアデス出版


An Introduction to Generalized Linear Models Starting with Mean-Variance
平均・分散から始める
一般化線形モデル入門
馬場 真哉
Baba Shinya
プレアデス出版

セミナーのテーマ

(前半)線形回帰モデルから一般化線形モデルへ

→ ロジスティック回帰モデルの特徴を学び
分析結果を理解できるようになる

(後半)切片と傾きが時間に応じて変化する回帰モデル

→ 線形ガウス状態空間モデルの初歩を学び
簡単な分析ができるようになる

単純なモデルを発展させる
分析モデルの特徴を理解し、実装する

回帰分析の初歩

内容

1. 回帰分析とは？　回帰モデルとは？

2. 回帰モデルと正規分布

回帰分析の目的

回帰分析の目的

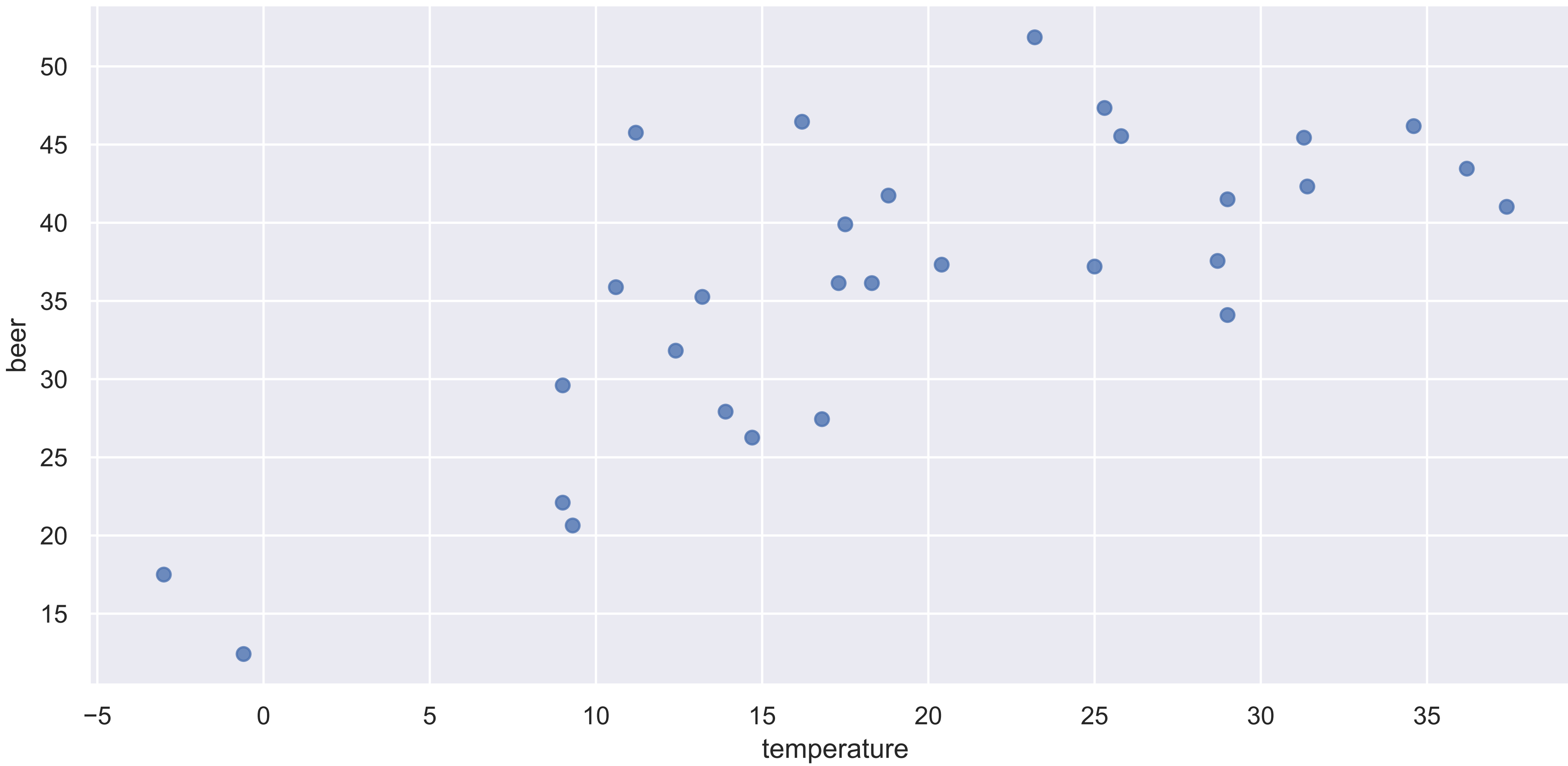
2つ以上の変数の関係性を、定量的に表す式を求める
→回帰モデル
　変数が2つの場合は「単回帰モデル」と呼ばれる

単回帰モデルによって得られた数式の例

$$ビールの売り上げ = 22.79 + 気温 \times 0.69$$

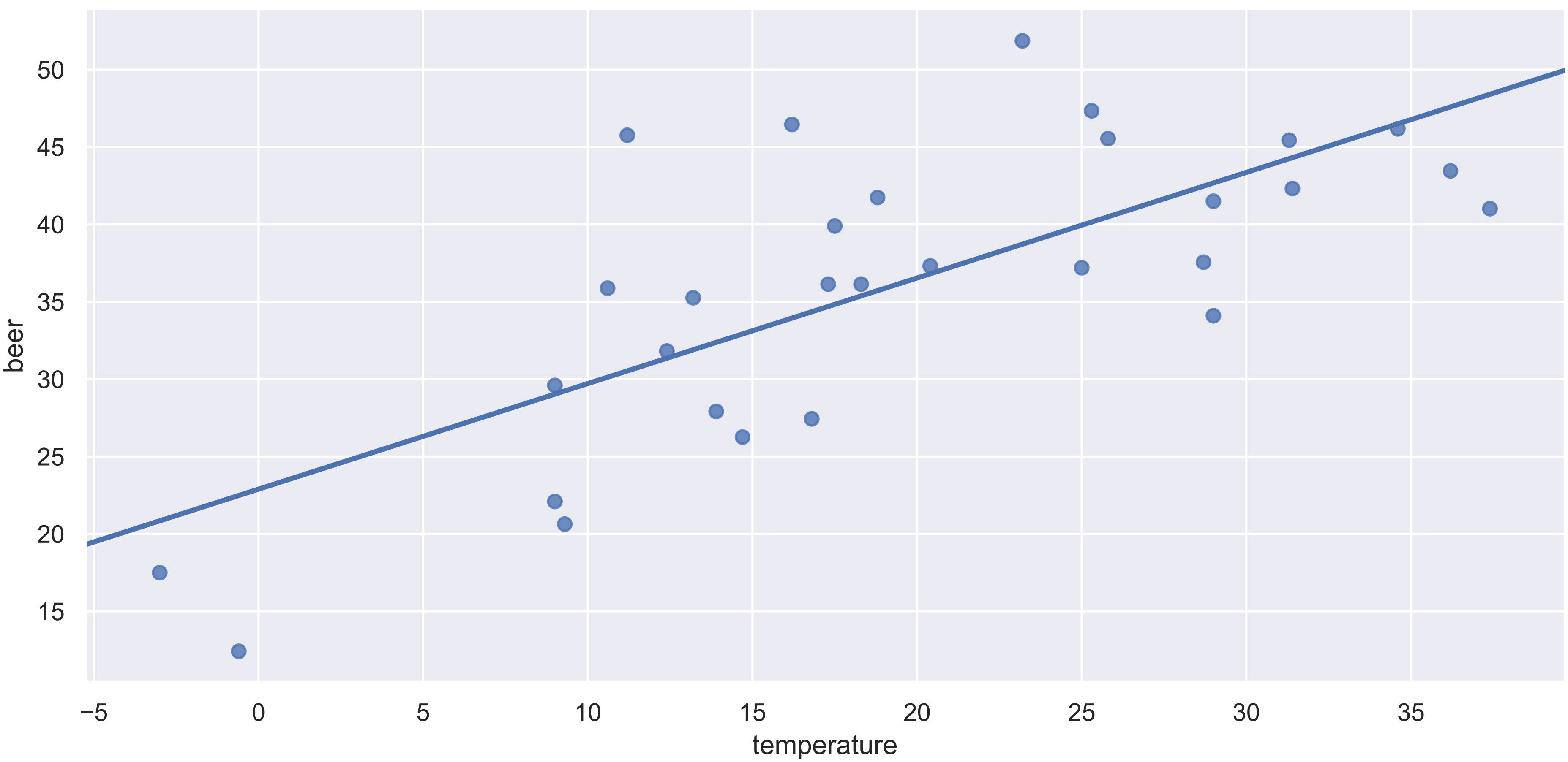
線形回帰モデル

ビールの売り上げと気温の散布図



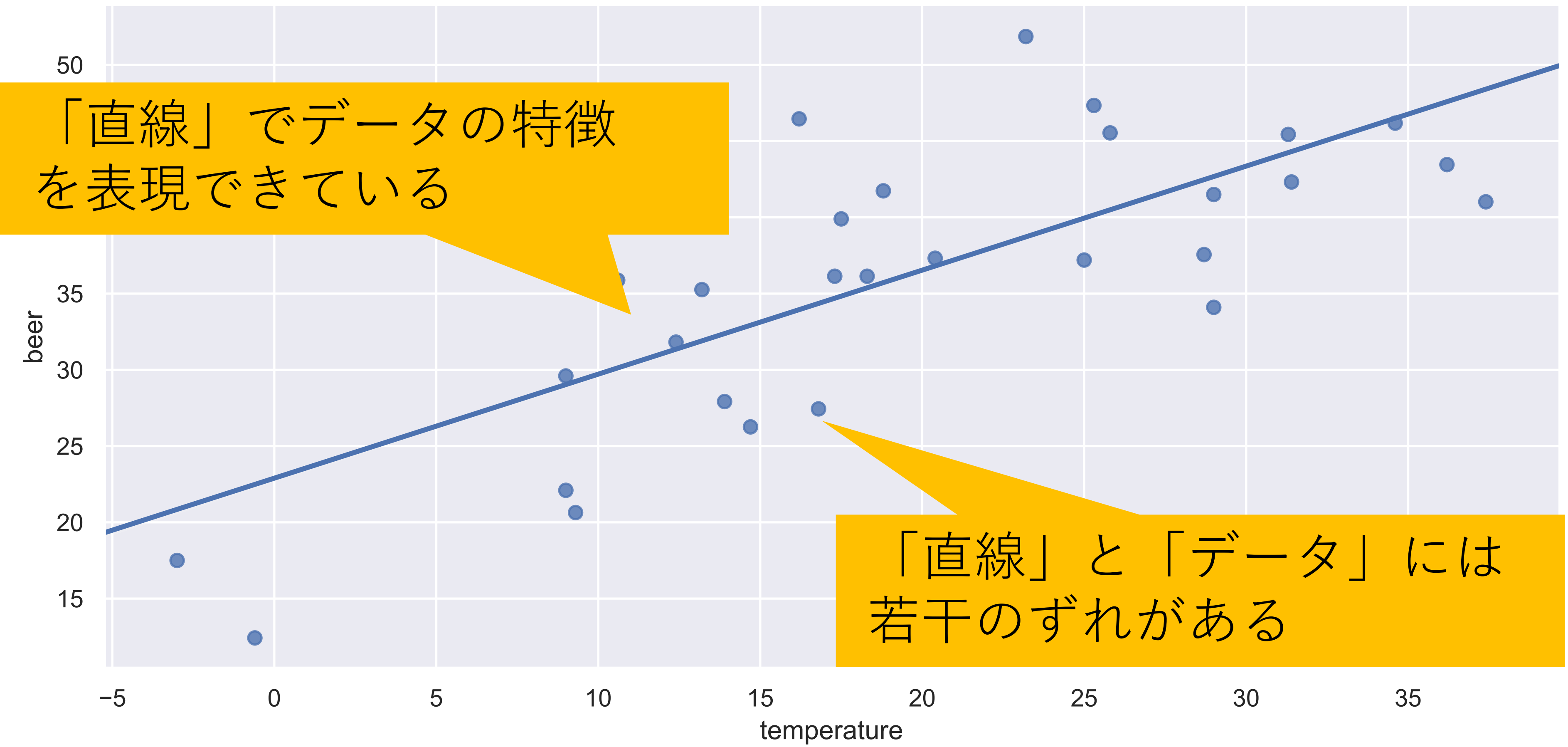
線形回帰モデル

売り上げ ＝ 22.79 ＋ 気温 × 0.69の直線を追加
（回帰直線と呼ぶ）



線形回帰モデル

売り上げ = 22.79 + 気温 × 0.69の直線を追加
（回帰直線と呼ぶ）



回帰分析の目的

単回帰モデルによって得られた数式の例

$$\text{ビールの売り上げ} = 22.79 + \text{気温} \times 0.69 + \underline{\text{誤差}}$$

誤差を入れるのを忘れない

回帰分析のメリット

解釈　売上と気温の関係性が分かる
→気温が上がると、売り上げも上がる

予測　気温をもとにして、売り上げを予測できる
→気温が10度の時の売り上げを予測する

売り上げ $= 22.79 + 10 \times 0.69 \approx 30$

およそ、30万円の売り上げになると予測

現象を数式で表現し、
現象の解釈や予測に役立たせる

線形回帰モデル

回帰分析によって得られた数式の例

ビールの売り上げ = 22.79 + 気温 × 0.69 + 誤差

切片

傾き
気温の係数

応答変数
（興味の対象）

説明変数
（興味の対象を変化させる）

一般的な記号

切片β_0、傾きβ_1、
応答変数Y_i、説明変数x_i、誤差ε_i

$$Y_i = \beta_0 + x_i \cdot \beta_1 + \varepsilon_i$$

内容

1. 回帰分析とは？　回帰モデルとは？

2. 回帰モデルと正規分布

線形回帰モデル

回帰分析によって得られた数式の例

① ビールの売り上げ = 22.79 + 気温 × 0.69 + 誤差

Y_i　　　　　　x_i

平均が0である
正規分布に従うと仮定

※　独立などの条件を満たせば、正規分布以外の分布の誤差でも、
最小二乗法を用いたパラメータ推定には支障ない
→これから説明する一般化線形モデルを理解するために、
今回の講義では最初から正規分布を仮定する

正規分布を使ってモデルを作る

線形回帰モデル

回帰分析によって得られた数式の例

① $Y_i = \boxed{22.79 + x_i \times 0.69}$ + 平均0、分散$\boxed{\sigma^2}$の$\boxed{\text{正規分布}}$に従う誤差

② $Y_i \sim \boxed{\text{Normal}}(\boxed{22.79 + x_i \cdot 0.69}, \boxed{\sigma^2})$

誤差が正規分布に従う
→ Yも正規分布に従う（正規分布の特徴）

線形回帰モデル

回帰分析によって得られた数式の例

① $Y_i = 22.79 + x_i \times 0.69 +$ 平均0、分散σ^2の正規分布に従う誤差

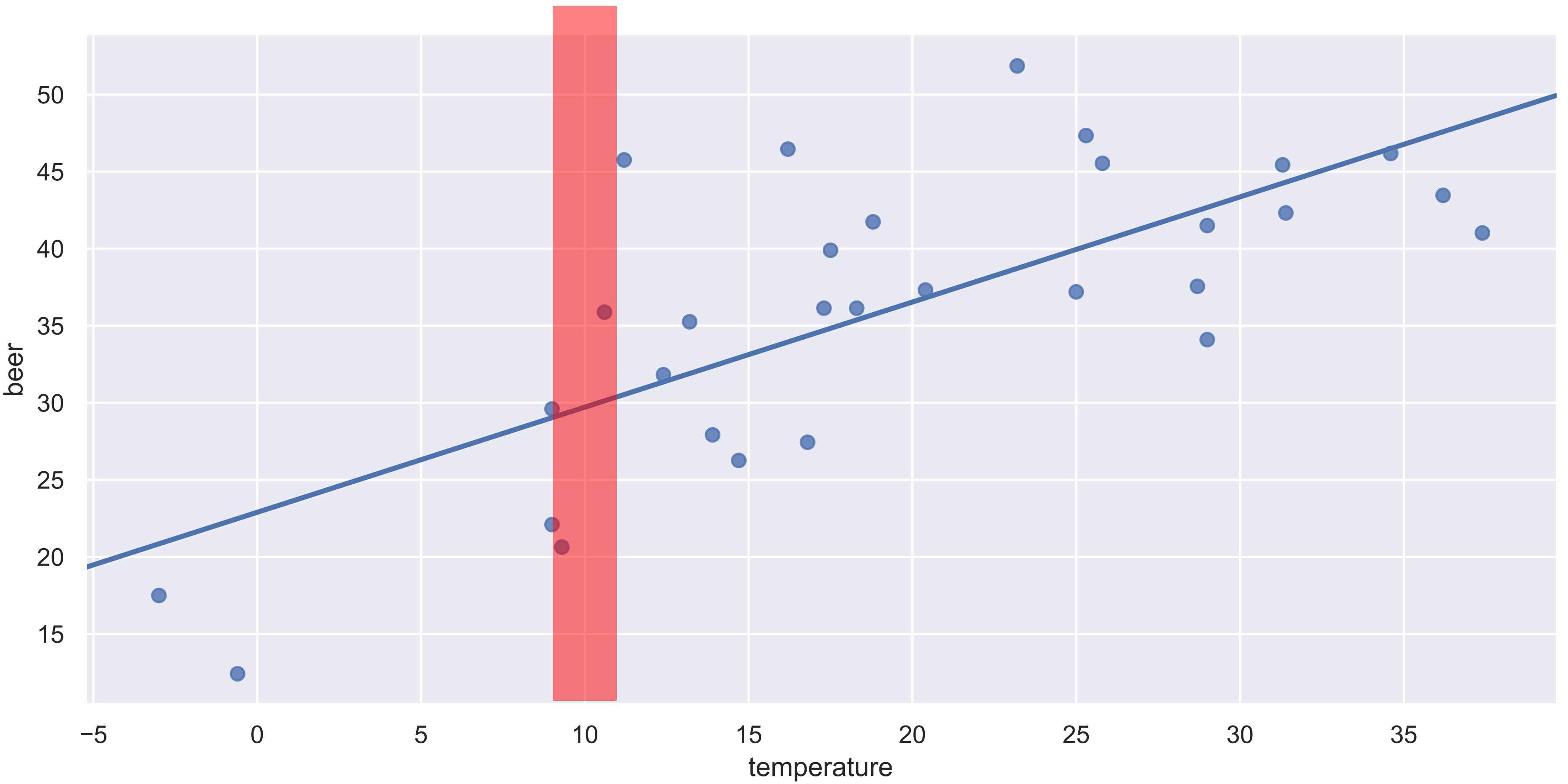
② $Y_i \sim \mathrm{Normal}(22.79 + x_i \cdot 0.69, \sigma^2)$

③ $\mu_i = 22.79 + x_i \cdot 0.69$
$Y_i \sim \mathrm{Normal}(\mu_i, \sigma^2)$

正規分布の期待値や分散を見やすくする
→式を2行に分けた。意味は全く同じ

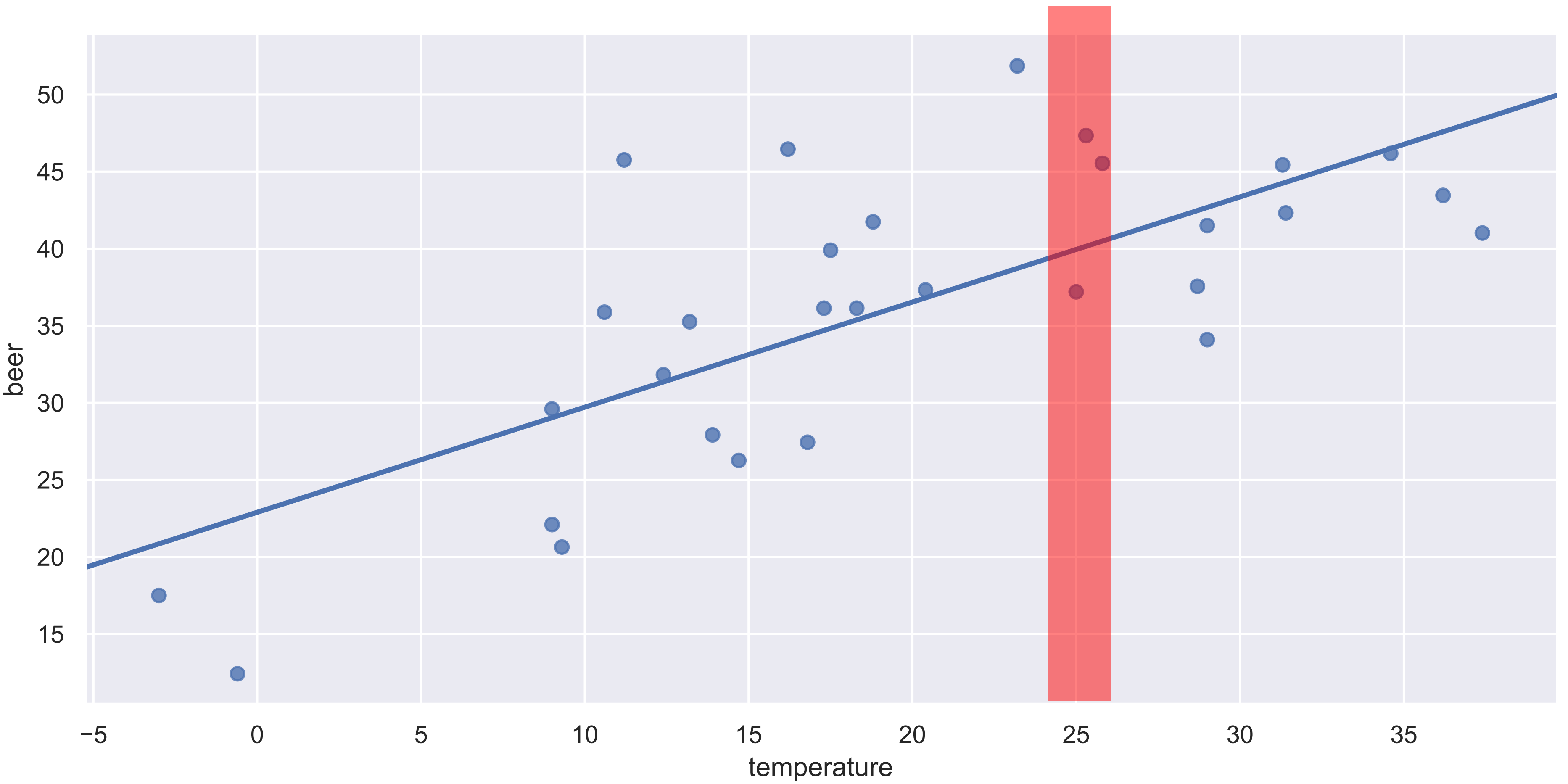
線形回帰モデル

売り上げの分布は、平均が「$22.79 + 10 \times 0.69$」の正規分布

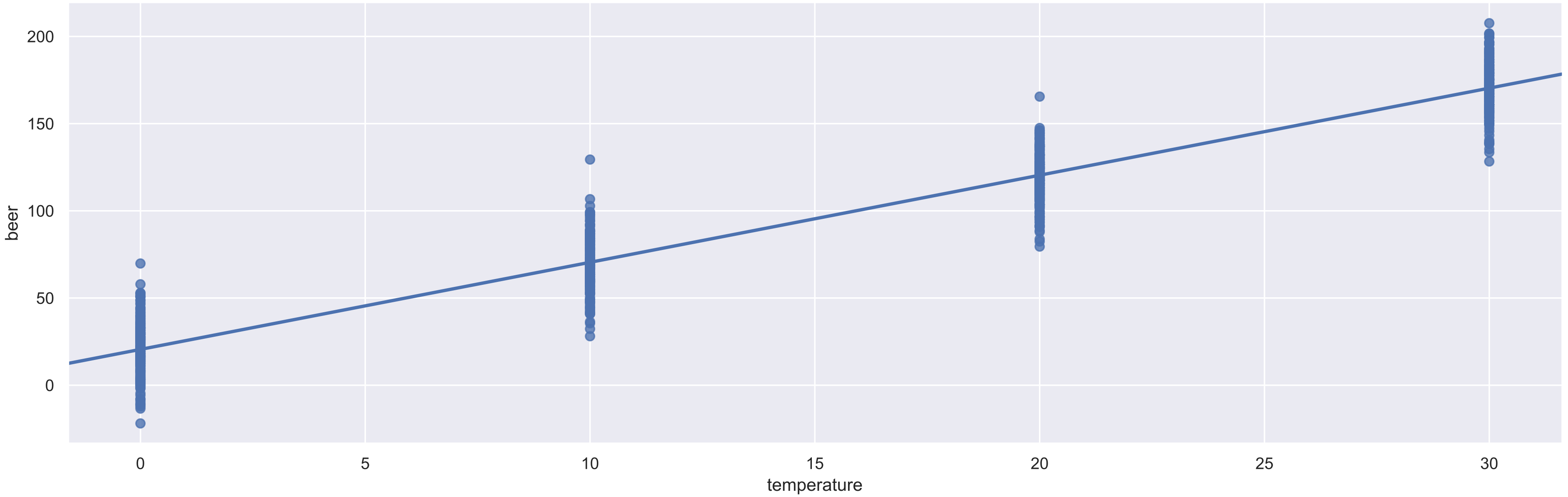


線形回帰モデル

売り上げの分布は、平均が「$22.79 + 25 \times 0.69$」の正規分布

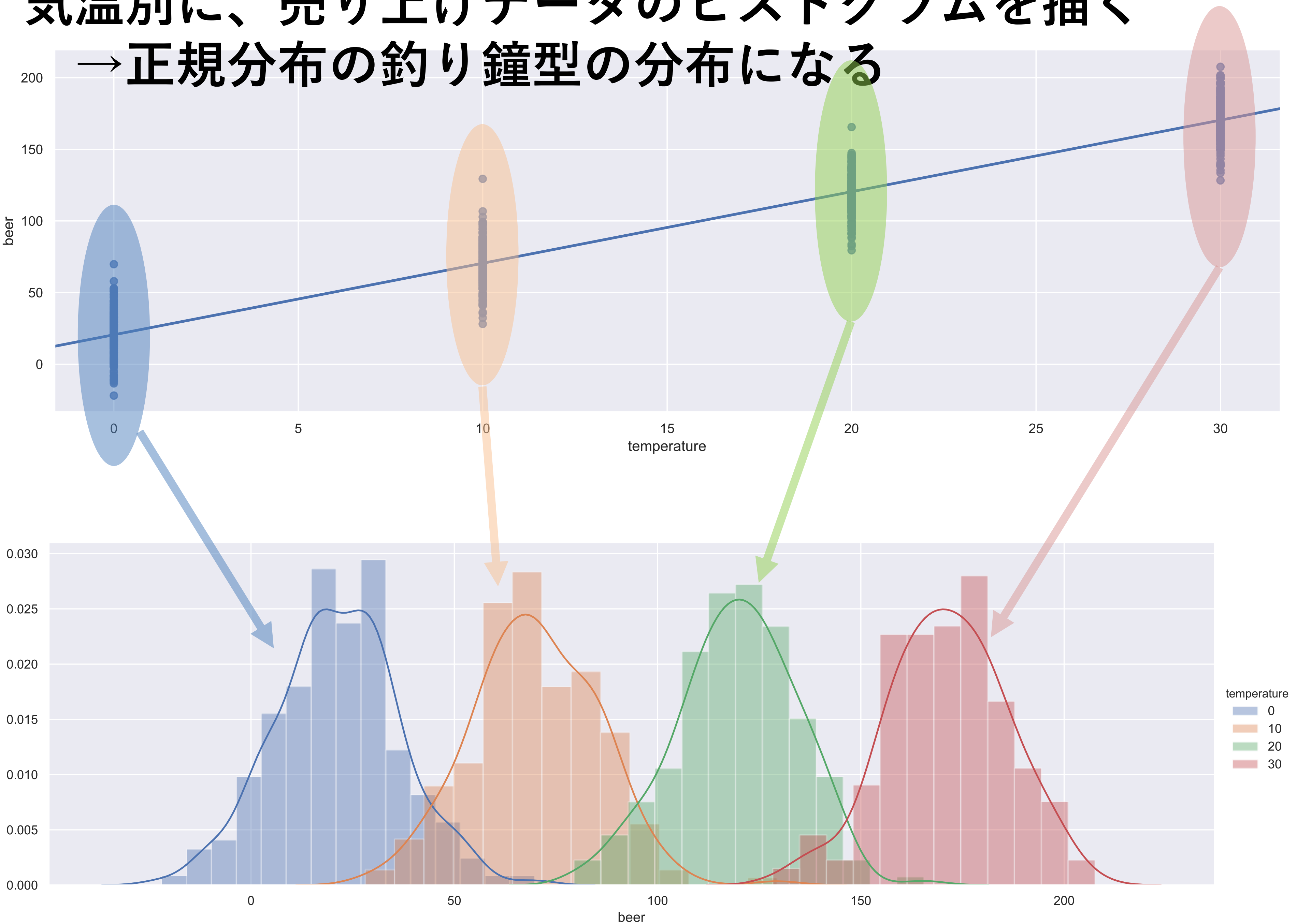


回帰モデルのイメージをデータを変えて確認



気温が「０，１０，２０，３０」の４種類のみ存在。
売り上げを応答変数として、回帰直線を描いた

気温別に、売り上げデータのヒストグラムを描く
→正規分布の釣り鐘型の分布になる



統計モデルと確率分布　再考

モデルとは

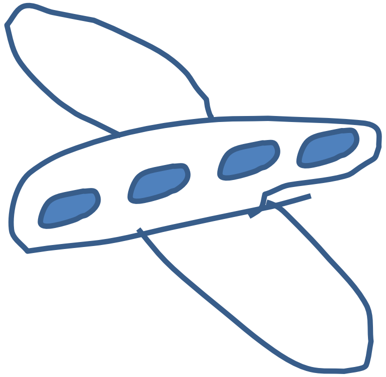
モデル

現象を単純化した「模型」のこと

プラモデルのモデル

モデルの例

飛行機のモデル（模型）を考える



ぼくの考えた「さいきょう」の
飛行機だよ！！

（馬場が描きました）

プロペラが無いので飛ばない

モデルとは

<u>モデルの使い道</u>

モデルを使えば、
実際に行動する前に、結果について議論できる

何億円ものお金をかけて本物を作る前に、
空気抵抗や推進力などいろいろなことを検討できる

データに基づいてモデルを作ることで、
将来の予測等ができるかもしれない

モデルの利用

<u>モデルと現実の対応に要注意</u>

モデルの飛行機には、エンジンもプロペラもあって
ちゃんと飛ぶことがわかりました

でもなんか、本物の飛行機は飛ばんのだが

本物の飛行機には、エンジンついてないからね〜

ダメじゃん

モデルの利用

<u>モデルと現実の対応に要注意</u>

回帰分析の結果、こういう感じになったよ！

データは線形回帰モデルが適合しそうなの？

全然違います

その分析結果を本当に使うの……？

統計学におけるモデル

モデル

模型。現実世界の模型を作る

統計学におけるモデル

観測したデータを生み出す確率的な過程を
簡潔に記述したもの。

Graham Upton, Ian Cook.（白幡慎吾 監訳）.（2010）. 統計学辞典. 共立出版

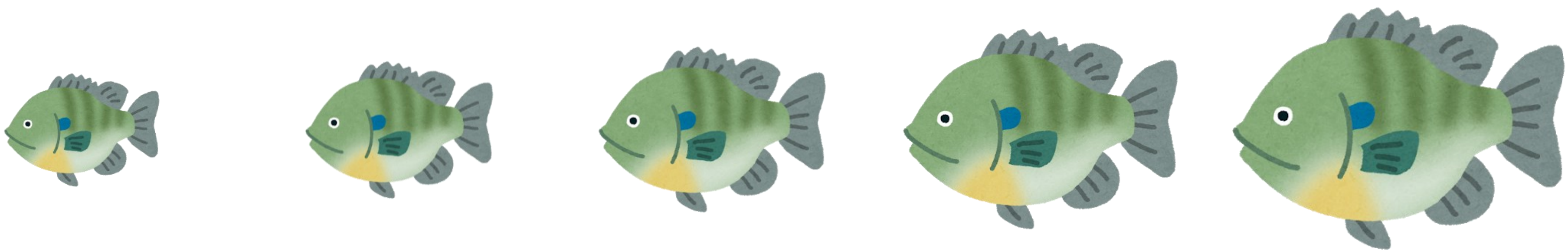
「データを生み出す確率的な過程」や
「データが得られる過程」に着目

統計学におけるモデル

今回の事例

湖からの標本抽出
→5尾の魚だけがいる湖から1尾抽出する

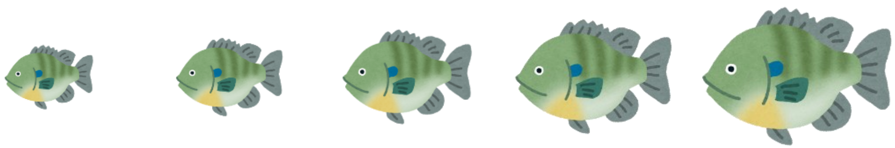
1cm **2cm** **3cm** **4cm** **5cm**



とても小さな湖から無作為に1尾を取得
このモデル(データが得られる過程)は？

統計学におけるモデル

1cm　2cm　3cm　4cm　5cm



どの体長も、0.2の確率で得られる

体長	確率
1cm	0.2
2cm	0.2
3cm	0.2
4cm	0.2
5cm	0.2

確率分布

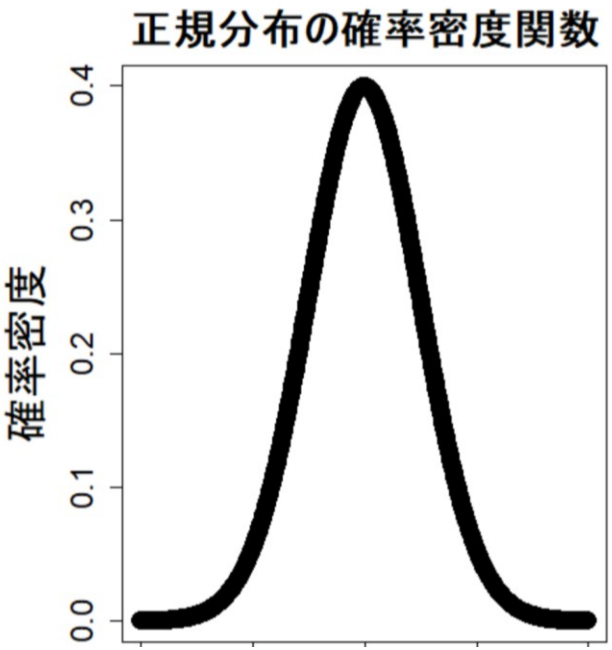
確率変数(この場合の体長)と確率の対応関係のこと
今回は右上の数表が確率分布となる

このモデルはちょっと単純すぎ？
もう少し現実味のあるモデルにしたい

正規分布を利用したモデル

正規分布の特徴

平均値(期待値)に近い値が出やすい
平均値に対して左右対称の確率分布



体長	確率
1cm	0.1
2cm	0.2
3cm	0.4
4cm	0.2
5cm	0.1

正規分布を使う方がよさそうな気がする？

正規分布を利用したモデル

ビールの売り上げは正規分布に従うか？

平均的な売り上げになりやすく、
上振れ・下振れは同等の確率で発生するとみなせそう？
→このように想定できるなら、正規分布がよさそう

どんなデータでも正規分布を使えるか？

0か1しかとらないデータは、正規分布を使えない気がする
整数の値しかとらないデータも、正規分布は不適合

正規分布以外の確率分布を使う方が現実にあうモデルになる可能性がある

ロジスティック回帰の初歩

一般化線形モデルの基本

線形回帰モデル

$$\mu_i = \beta_0 + \beta_1 \cdot x_i$$
$$Y_i \sim \mathrm{Normal}(\mu_i, \sigma^2)$$

正規分布しか扱えない

一般化線形モデル(Generalized Linear Models :GLM)

正規分布以外の分布を扱うことができる

正規分布しか使えない線形回帰モデルから様々な分布が使える一般化線形モデルへ

一般化線形モデルの基本

一般化線形モデルの構造(イメージ)

$$g(\theta_i) = \beta_0 + \beta_1 \cdot x_i$$

線形の構造(線形予測子)でパラメータが変化
リンク関数$g()$を使い、モデルの表現力を上げる(後述)

$$Y_i \sim 何らかの分布(\theta_i)$$

正規分布以外の様々な分布に対応

一般化線形モデルの基本

正規線形モデル(線形回帰モデル)

リンク関数が恒等関数であり、
正規分布を確率分布として用いる一般化線形モデル

恒等関数

$f(x) = x$となる関数のこと

線形回帰モデル(復習)

$$\mu_i = \beta_0 + \beta_1 \cdot x_i$$
$$Y_i \sim \mathrm{Normal}(\mu_i, \sigma^2)$$

一般化線形モデルの基本

ロジスティック回帰モデル

一般化線形モデルにはいろいろな種類がある
・ポアソン回帰モデル
・ガンマ回帰モデル
→用いる確率分布によって名前が変わる

今回は「ロジスティック回帰モデル」に焦点を当てる

内容

1. 二項分布を使ったモデル化
2. ロジスティック関数・ロジット関数
3. ロジスティック回帰の構造

二項分布

二項分布の確率質量関数

$$\mathrm{Bin}(X|n,\theta) = {}_n\mathrm{C}_x \cdot \theta^x \cdot (1-\theta)^{n-x}$$

確率変数Xは0以上n以下の整数を取りうる

パラメータはn, θの2つ
n：試行回数
θ：成功確率

二項分布

二項分布が適合しそうな例

表が出る確率がθであるコインをn回投げた時、
X回の表が出る確率

$$\mathrm{Bin}(X|n,\theta) = {}_n\mathrm{C}_x \cdot \theta^x \cdot (1-\theta)^{n-x}$$

期待値

$$E(X) = n\theta$$

分散

$$V(X) = n\theta(1-\theta)$$

ロジスティック回帰

線形回帰モデルとロジスティック回帰モデルの比較

モデルの構造は似ているが違いもある

確率分布の違い

線形回帰　　　　　　　：正規分布
ロジスティック回帰：二項分布
→これが大事。ここが変わると連鎖的に色々変わる

変化するパラメータの違い

線形回帰　　　　　　　：期待値μが変化する
ロジスティック回帰：成功確率θが変化する
→確率分布が違うので、パラメータも変わる
→期待値は「試行回数n」×「成功確率θ」なので注意

ロジスティック回帰

線形回帰モデルの場合

① $Y_i = \beta_0 + \beta_1 \cdot x_i +$ 正規分布に従う誤差

ロジスティック回帰モデルの場合

② $Y_i = \beta_0 + \beta_1 \cdot x_i +$ 二項分布に従う誤差・・・？

これは間違い！

誤差が加わる構造ではないので注意！

ロジスティック回帰

線形回帰モデルの場合

① $Y_i = \beta_0 + \beta_1 \cdot x_i +$ 正規分布に従う誤差

③ $\mu_i = \beta_0 + \beta_1 \cdot x_i$
$Y_i \sim \mathrm{Normal}(\underline{\mu_i}, \sigma^2)$

確率分布のパラメータが変わる！

③の構造で確率分布を変更するイメージ！
ロジスティック回帰は成功確率θが変化する

ロジスティック回帰

データの例

10人ずつの班を用意しテストを受験、 Y 人が合格
班ごとに勉強時間xを変化
→試行回数nは10回（1つの班に10人いる）
→成功確率θは、勉強時間によって変化する

データが得られる確率的な過程(仮)

$$\theta_i = \beta_0 + \beta_1 \cdot x_i$$
$$Y_i \sim \mathrm{Bin}(10, \theta_i)$$

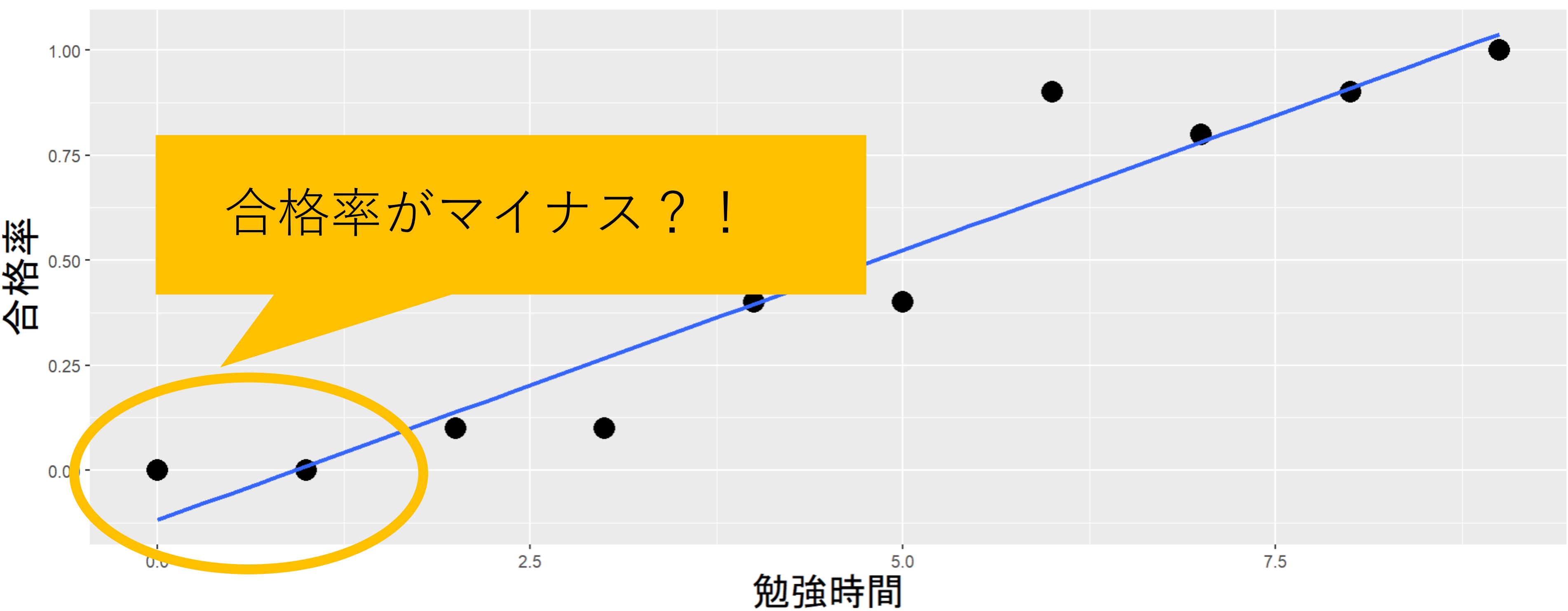
NormalをBinに変更した。実は、まだ修正が必要

ロジスティック回帰モデル

<u>**二項分布に従うデータに対して、線形構造を指定する問題点**</u>

$$\theta_i = \beta_0 + \beta_1 \cdot x_i$$

→ θ_iが0未満や1以上の値をとることがありうる



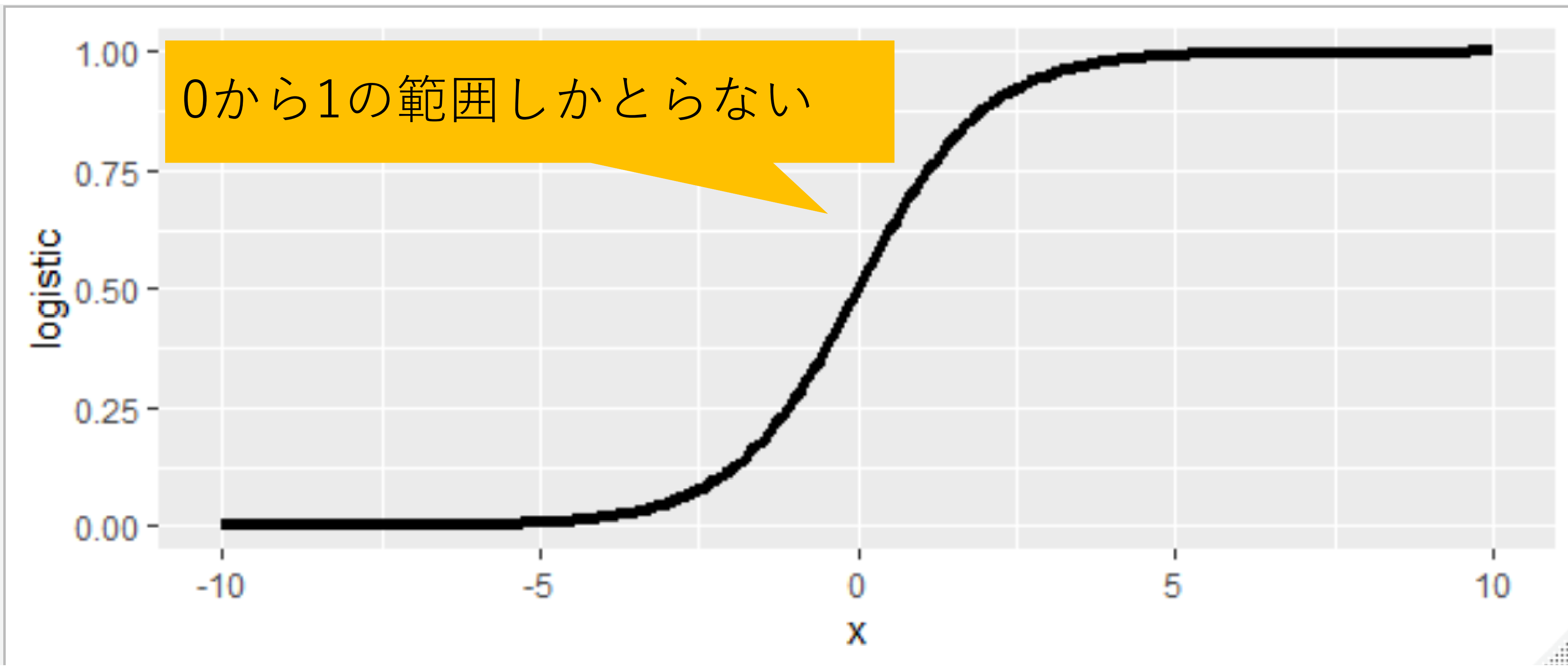
内容

1. 二項分布を使ったモデル化
2. ロジスティック関数・ロジット関数
3. ロジスティック回帰の構造

ロジスティック回帰モデル

ロジスティック関数

$$\mathrm{logistic}(x) = \frac{1}{1 + \mathrm{EXP}(-x)}$$



ロジスティック回帰モデル

ロジスティック関数は、なぜ0以上1以下しかとらないか

$$\mathrm{logistic}(x) = \frac{1}{1 + \mathrm{EXP}(-x)}$$

以下は指数関数と呼ぶ。 eはネイピア数（2.7くらいの値）

$$e^x$$

以下のように表記することもある（意味は同じ）

$$\mathrm{EXP}(x)$$

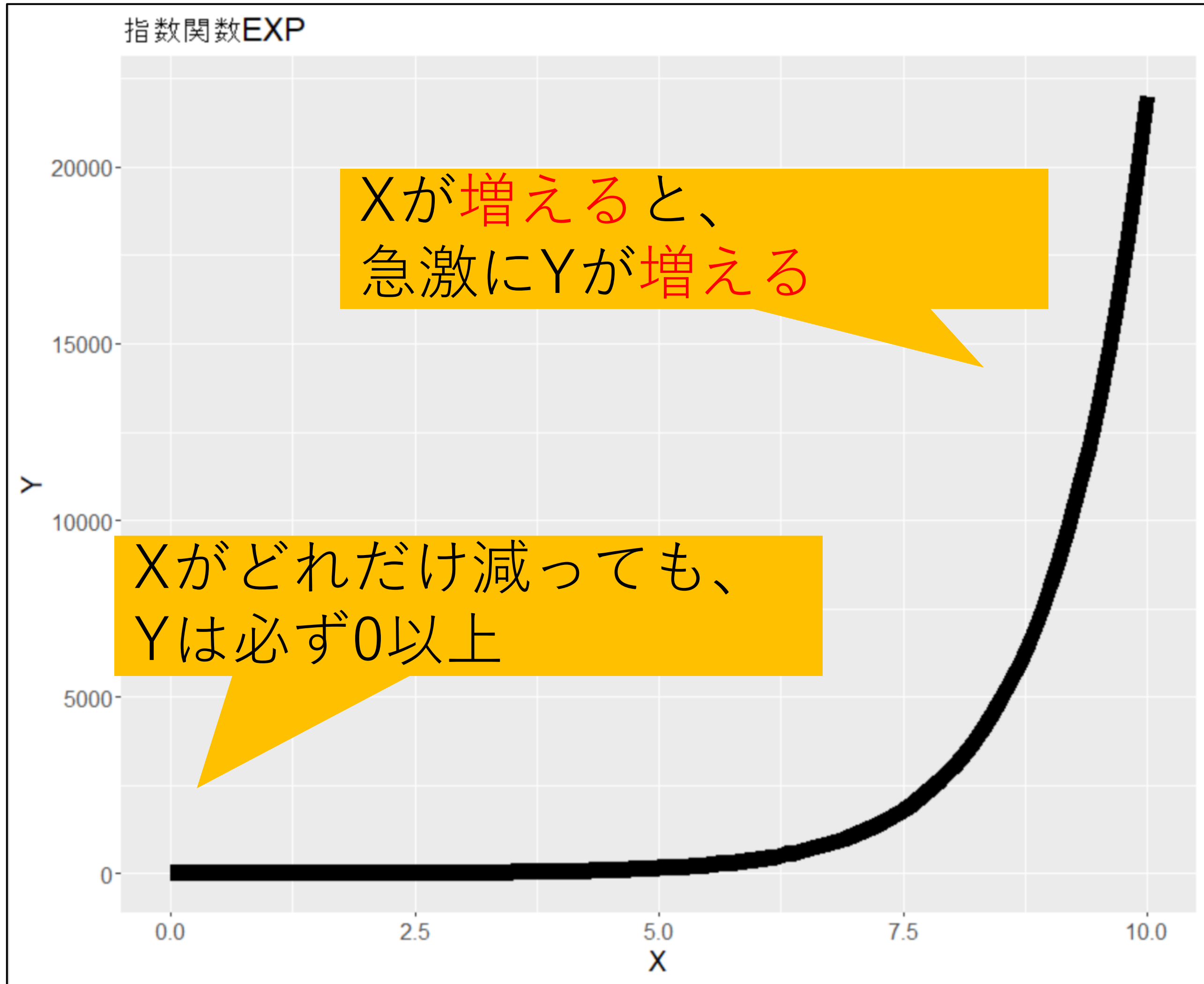
$x = 0$なら、指数関数の結果は1になる

$$e^0 = \mathrm{EXP}(0) = 1$$

ロジスティック回帰モデル

指数関数のグラフ①

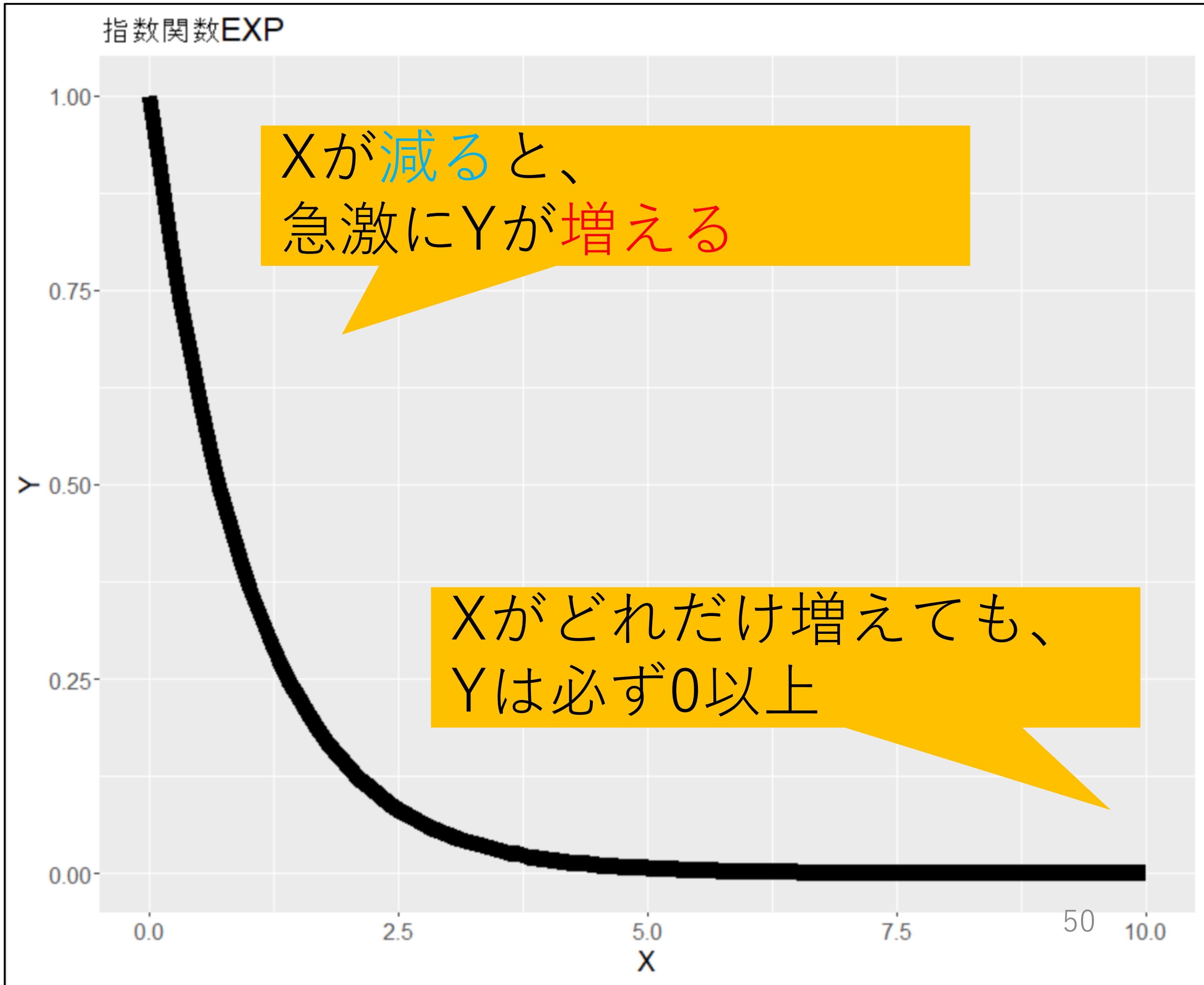
$e^{x} = \mathrm{EXP}(x)$



ロジスティック回帰モデル

指数関数のグラフ②

$e^{-x} = \mathrm{EXP}(-x)$

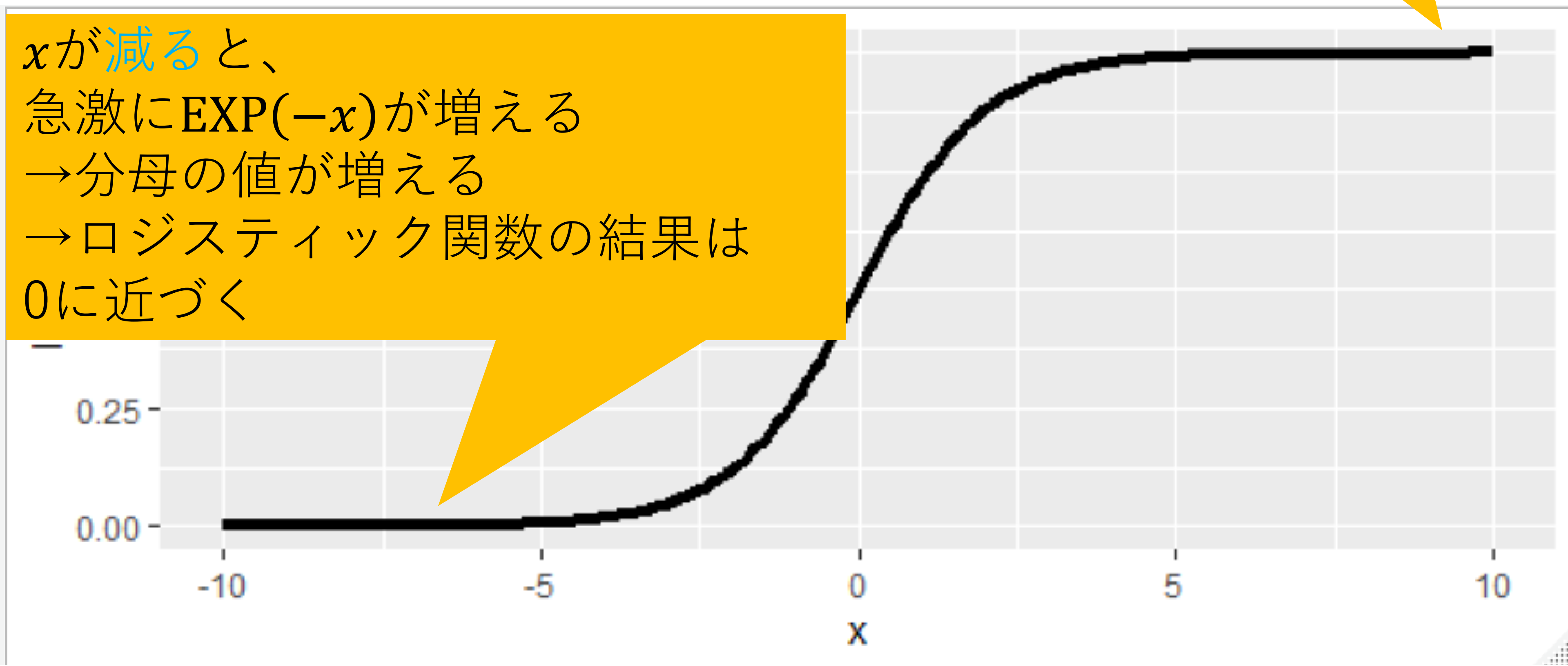


ロジスティッ

ロジスティック

xが増えると、
急激に$\mathrm{EXP}(-x)$が0に近づく
→分母の値が「1」に近づく
→ロジスティック関数の結果は1に近づく

$$\mathrm{logistic}(x) = \frac{1}{1 + \mathrm{EXP}(-x)}$$



xが減ると、
急激に$\mathrm{EXP}(-x)$が増える
→分母の値が増える
→ロジスティック関数の結果は
0に近づく

ロジスティック回帰モデル

ロジット関数

$$\mathrm{logit}(x) = \log \frac{x}{1-x}$$

ロジット関数はロジスティック関数の逆関数

$$\mathrm{logit}\big(\mathrm{logistic}(x)\big) = x$$

内容

1. 二項分布を使ったモデル化
2. ロジスティック関数・ロジット関数
3. ロジスティック回帰の構造

ロジスティック回帰モデル

データが得られる確率的な過程(修正版)

$$\mathrm{logit}(\theta_i) = \beta_0 + \beta_1 \cdot x_i$$
$$Y_i \sim \mathrm{Bin}(n, \theta_i)$$

ロジスティック回帰モデル

データが得られる確率的な過程(修正版)

$$\mathrm{logit}(\theta_i) = \beta_0 + \beta_1 \cdot x_i$$
$$Y_i \sim \mathrm{Bin}(n, \theta_i)$$

リンク関数として
ロジット関数を使う

1つ目の式の左右両辺にロジスティック関数を適用

$$\theta_i = \mathrm{logistic}(\beta_0 + \beta_1 \cdot x_i)$$
$$Y_i \sim \mathrm{Bin}(n, \theta_i)$$

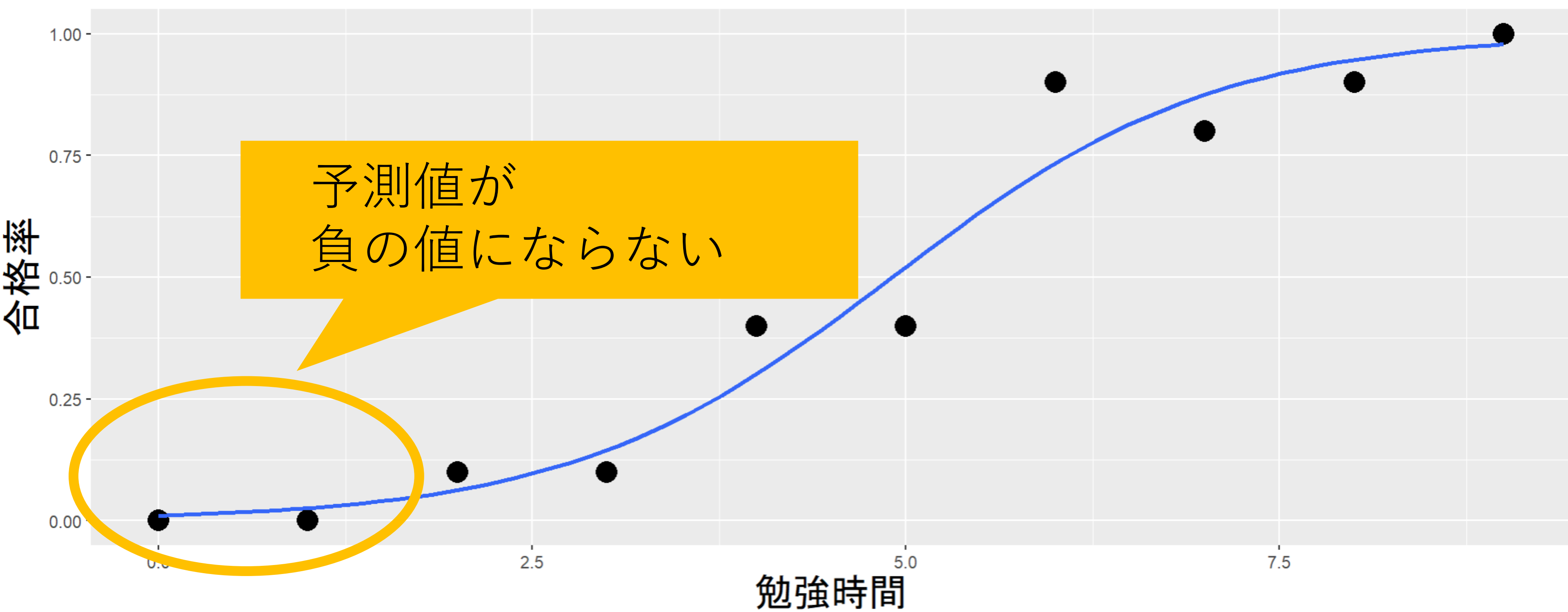
線形予測子にロジスティック関数を
適用したと解釈できる

ロジスティック回帰モデル

データが得られる確率的な過程(修正版)

$\mathrm{logit}(\theta_i) = \beta_0 + \beta_1 \cdot x_i$
$Y_i \sim \mathrm{Bin}(10, \theta_i)$

$\theta_i = \mathrm{logistic}(\beta_0 + \beta_1 \cdot x_i)$
$Y_i \sim \mathrm{Bin}(10, \theta_i)$



まとめ

$\text{logit}(\theta_i) = \beta_0 + \beta_1 \cdot x_i$

$Y_i \sim \text{Bin}(n, \theta_i)$

$\theta_i = \text{logistic}(\beta_0 + \beta_1 \cdot x_i)$

$Y_i \sim \text{Bin}(n, \theta_i)$

ロジスティック回帰モデル

リンク関数がロジット関数であり、
二項分布を確率分布として用いる一般化線形モデル

モデルの構造は、
ロジスティック関数を使って覚えるとわかりやすい

$$\theta_i = \frac{1}{1 + \text{EXP}\left(-(\beta_0 + \beta_1 \cdot x_i)\right)}$$

$$Y_i \sim \text{Bin}(n, \theta_i)$$

$$\text{logistic}(x) = \frac{1}{1 + \text{EXP}(-x)}$$

「ある」「ない」などの二値分類や
「〇個中、×個が成功」といったデータに対して適用する

尤度と最尤法

切片と傾きの計算方法

$$\mathrm{logit}(\theta_i) = \beta_0 + \beta_1 \cdot x_i$$
$$Y_i \sim \mathrm{Bin}(n, \theta_i)$$

内容

1. 尤度と最尤法の基本

2. 対数と対数尤度

3. ロジスティック回帰の尤度

尤度と最尤法

尤度

パラメータを指定したときに
手持ちのデータが得られる確率を計算することによって得る

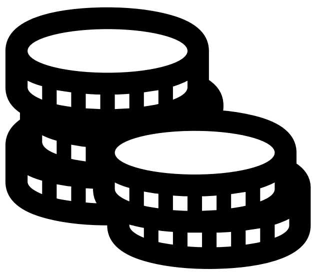
尤度関数

標本yは固定。パラメータθを入力すると尤度を出力する関数
$f(y|\theta)$と表記される

（尤度関数の合計値や積分値は1にならないので注意）
（尤度関数は確率質量関数や確率密度関数とはみなせない）

尤度と最尤法

例）コインを10枚投げた。そのうちの2枚が表だった



データ：　　表2枚

パラメータ：θ = コインを投げて表が出る確率

尤度関数：　$f(y|\theta) = \mathrm{Bin}(2|10, \theta)$

$$= {}_{10}\mathrm{C}_2 \cdot \theta^2 \cdot (1-\theta)^{10-2}$$

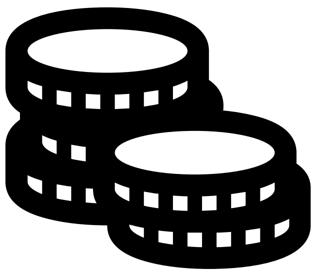
$$f(y|0.1) = {}_{10}\mathrm{C}_2 \cdot 0.1^2 \cdot (1-0.1)^{10-2} \approx 0.19$$

$$f(y|0.2) = {}_{10}\mathrm{C}_2 \cdot 0.2^2 \cdot (1-0.2)^{10-2} \approx 0.30$$

$$f(y|0.3) = {}_{10}\mathrm{C}_2 \cdot 0.3^2 \cdot (1-0.3)^{10-2} \approx 0.23$$

尤度と最尤法

例）コインを10枚投げる作業を2回行った。
1回目：2枚表　　2回目：4枚表



データ：　　表2枚・表4枚

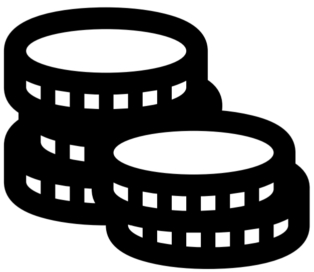
パラメータ：θ = コインを投げて表が出る確率

尤度関数： $f(y|\theta) = \mathrm{Bin}(2|10,\theta) \times \mathrm{Bin}(4|10,\theta)$

$$= {}_{10}\mathrm{C}_2 \cdot \theta^2 \cdot (1-\theta)^{10-2} \times {}_{10}\mathrm{C}_4 \cdot \theta^4 \cdot (1-\theta)^{10-4}$$

尤度と最尤法

例）コインを10枚投げる作業を2回行った。
1回目：2枚表　　2回目：4枚表



データ：　　表2枚・表4枚

パラメータ：θ = コインを投げて表が出る確率

尤度関数：　$$f(y|\theta) = {}_{10}\mathrm{C}_2 \cdot \theta^2 \cdot (1-\theta)^{10-2} \times {}_{10}\mathrm{C}_4 \cdot \theta^4 \cdot (1-\theta)^{10-4}$$

式の書き換え

データ：　$y_1 = 2, y_2 = 4$

尤度関数：$$f(y|\theta) = \prod_{i=1}^{2} {}_{10}\mathrm{C}_{y_i} \cdot \theta^{y_i} \cdot (1-\theta)^{10-y_i}$$

尤度と最尤法

<u>対数尤度</u>

尤度の対数をとったもの。計算が楽になる(後述)

<u>最尤法</u>

（対数）尤度を最大にするパラメータを選ぶという、パラメータ推定の原理のこと

尤度と最尤法

<u>例）コインを10枚投げた。そのうちの2枚が表だった</u>

データ：　表2枚

最尤法

尤度が最も大きくなる
パラメータを採用

ータ：θ = コインを投げて表が出る確率

数：　$f(y|\theta) = {}_{10}\mathrm{C}_2 \cdot \theta^2 \cdot (1-\theta)^{10-2}$

$f(y|0.1) = {}_{10}\mathrm{C}_2 \cdot 0.1^2 \cdot (1-0.1)^{10-2} \approx 0.19$

$\boxed{f(y|0.2) = {}_{10}\mathrm{C}_2 \cdot 0.2^2 \cdot (1-0.2)^{10-2} \approx 0.30}$

$f(y|0.3) = {}_{10}\mathrm{C}_2 \cdot 0.3^2 \cdot (1-0.3)^{10-2} \approx 0.23$

内容

1. 尤度と最尤法の基本
2. 対数と対数尤度
3. ロジスティック回帰の尤度

対数と対数尤度

対数と対数尤度

尤度が最大になるパラメータを採用するのが最尤法

尤度の代わりに、尤度の対数をとった「対数尤度」を使う
→なぜわざわざ対数をとるのか？

対数と対数尤度

対数

指数の計算の逆だと思えばOK

指数	対数
$2^1 = 2$	$\log_2 2 = 1$
$2^2 = 4$	$\log_2 4 = 2$
$2^3 = 8$	$\log_2 8 = 3$

対数と対数尤度

対数

指数の計算の逆だと思えばOK

指数	対数
$2^1 = 2$	$\log_2 2 = 1$
$2^2 = 4$	$\log_2 4 = 2$
$2^3 = 8$	$\log_2 8 = 3$

$\log_2 8$の「 2 」の部分を対数の底（テイ）と呼ぶ

対数と対数尤度

$$\log_2 2 = 1$$

$$\log_2 4 = 2$$

$$\log_2 8 = 3$$

対数の特徴

掛け算を足し算に変えることができる

素朴な掛け算

$$2 \times 4 = 8$$

両辺に対数をとる

$$\log_2(2 \times 4) = \log_2(8)$$

対数の中の掛け算は、対数の外の足し算になる

$$\log_2(2) + \log_2(4) = \log_2(8)$$

対数と対数尤度

$$\log_2 2 = 1$$

$$\log_2 4 = 2$$

$$\log_2 8 = 3$$

対数の特徴

掛け算を足し算に変えることができる

素朴な掛け算

$$2 \times 4 = 8$$

両辺に対数をとる

$$\log_2(2 \times 4) = \log_2(8)$$

対数の中の掛け算は、対数の外の足し算になる

$$\overset{1}{\boxed{\log_2(2)}} + \overset{2}{\boxed{\log_2(4)}} = \overset{3}{\boxed{\log_2(8)}}$$

対数と対数尤度

自然対数

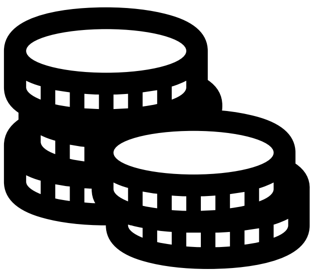
EXP（ネイピア数の指数関数）の逆は、対数の底を略する
これを自然対数と呼ぶ

対数の底は何でもよいが、
自然対数を使うことが多い

自然対数
$\log e = 1$
$\log e^2 = 2$
$\log e^3 = 3$

尤度と最尤法

例） コインを10枚投げる作業を2回行った。
1回目：2枚表　　2回目：4枚表



データ：　表2枚・表4枚

パラメータ：θ = コインを投げて表が出る確率

尤度関数：
$$f(y|\theta) = {}_{10}\mathrm{C}_2 \cdot \theta^2 \cdot (1-\theta)^{10-2} \times {}_{10}\mathrm{C}_4 \cdot \theta^4 \cdot (1-\theta)^{10-4}$$

対数尤度関数：

$$\log\bigl(f(y|\theta)\bigr) = \log\bigl({}_{10}\mathrm{C}_2 \cdot \theta^2 \cdot (1-\theta)^{10-2}\bigr) + \log\bigl({}_{10}\mathrm{C}_4 \cdot \theta^4 \cdot (1-\theta)^{10-4}\bigr)$$

尤度と最尤法

例）コインを10枚投げる作業を2回行った。
1回目：2枚表　　2回目：4枚表

データ：　$y_1 = 2, y_2 = 4$

尤度関数：$f(y|\theta) = \prod_{i=1}^{2} {}_{10}\mathrm{C}_{y_i} \cdot \theta^{y_i} \cdot (1-\theta)^{10-y_i}$

対数尤度関数：

$$\log\left(f(y|\theta)\right) = \sum_{i=1}^{2} \log\left({}_{10}\mathrm{C}_{y_i} \cdot \theta^{y_i} \cdot (1-\theta)^{10-y_i}\right)$$

対数を使うと掛け算しなくて済む！

内容

1. 尤度と最尤法の基本
2. 対数と対数尤度
3. ロジスティック回帰の尤度

尤度と最尤法

ロジスティック回帰モデルの場合

ロジスティック回帰モデルでは、成功確率が変わる

二項分布の単純なモデル

$Y \sim \mathrm{Bin}(n, \theta)$

nもθも一定

ロジスティック回帰モデル

$\mathrm{logit}(\theta_i) = \beta_0 + \beta_1 \cdot x_i$
$Y_i \sim \mathrm{Bin}(n, \theta_i)$

θがx_iの影響を受ける

尤度の計算の際、成功確率が変わる
表が出る確率が違うコインを使うイメージ

ロジスティック回帰

<u>データの例</u>

10人ずつの班を用意しテストを受験、 Y人が合格
班ごとに勉強時間xを変化
→試行回数nは10回（1つの班に10人いる）
→成功確率θは、勉強時間によって変化する

<u>対数尤度の計算の対象となるデータの例</u>

モデル(logistic表記)

$$\theta_i = \mathrm{logistic}(\beta_0 + \beta_1 \cdot x_i)$$
$$Y_i \sim \mathrm{Bin}(n, \theta_i)$$

データ

$$y_1 = 2, y_2 = 4 \qquad x_1 = 5, x_2 = 10$$

尤度と最尤法

尤度の計算例

データ：　　(10人中)2人・4人合格 / 勉強時間は5時間・10時間

パラメータ：β_0切片、 β_1傾き（勉強時間の係数）

尤度関数：

$$f(y|\theta) = \mathrm{Bin}\left(2|10, \mathrm{logistic}\,(\beta_0 + \beta_1 \cdot 5)\right) \times \mathrm{Bin}\left(4|10, \mathrm{logistic}\,(\beta_0 + \beta_1 \cdot 10)\right)$$

尤度と最尤法

尤度の計算例

データ：　(10人中)2人・4人合格 / 勉強時間は5時間・10時間

パラメータ：β_0切片、β_1傾き（勉強時

成功確率が
説明変数に応じて変化

尤度関数：

$$f(y|\theta) = \mathrm{Bin}\left(2|10, \boxed{\mathrm{logistic}\,(\beta_0 + \beta_1 \cdot 5)}\right) \times \mathrm{Bin}\left(4|10, \boxed{\mathrm{logistic}\,(\beta_0 + \beta_1 \cdot 10)}\right)$$

$$= {}_{10}\mathrm{C}_2 \cdot \mathrm{logistic}\,(\beta_0 + \beta_1 \cdot 5)^2 \cdot \left(1 - \mathrm{logistic}\,(\beta_0 + \beta_1 \cdot 5)\right)^{10-2} \times {}_{10}\mathrm{C}_4 \cdot \mathrm{logistic}\,(\beta_0 + \beta_1 \cdot 10)^4 \cdot \left(1 - \mathrm{logistic}\,(\beta_0 + \beta_1 \cdot 10)\right)^{10-4}$$

ロジスティック回帰の係数の解釈

ロジスティック回帰モデル

データが得られる確率的な過程(修正版)

$$\mathrm{logit}(\theta_i) = \beta_0 + \beta_1 \cdot x_i$$
$$Y_i \sim \mathrm{Bin}(10, \theta_i)$$

ロジスティック関数を適用

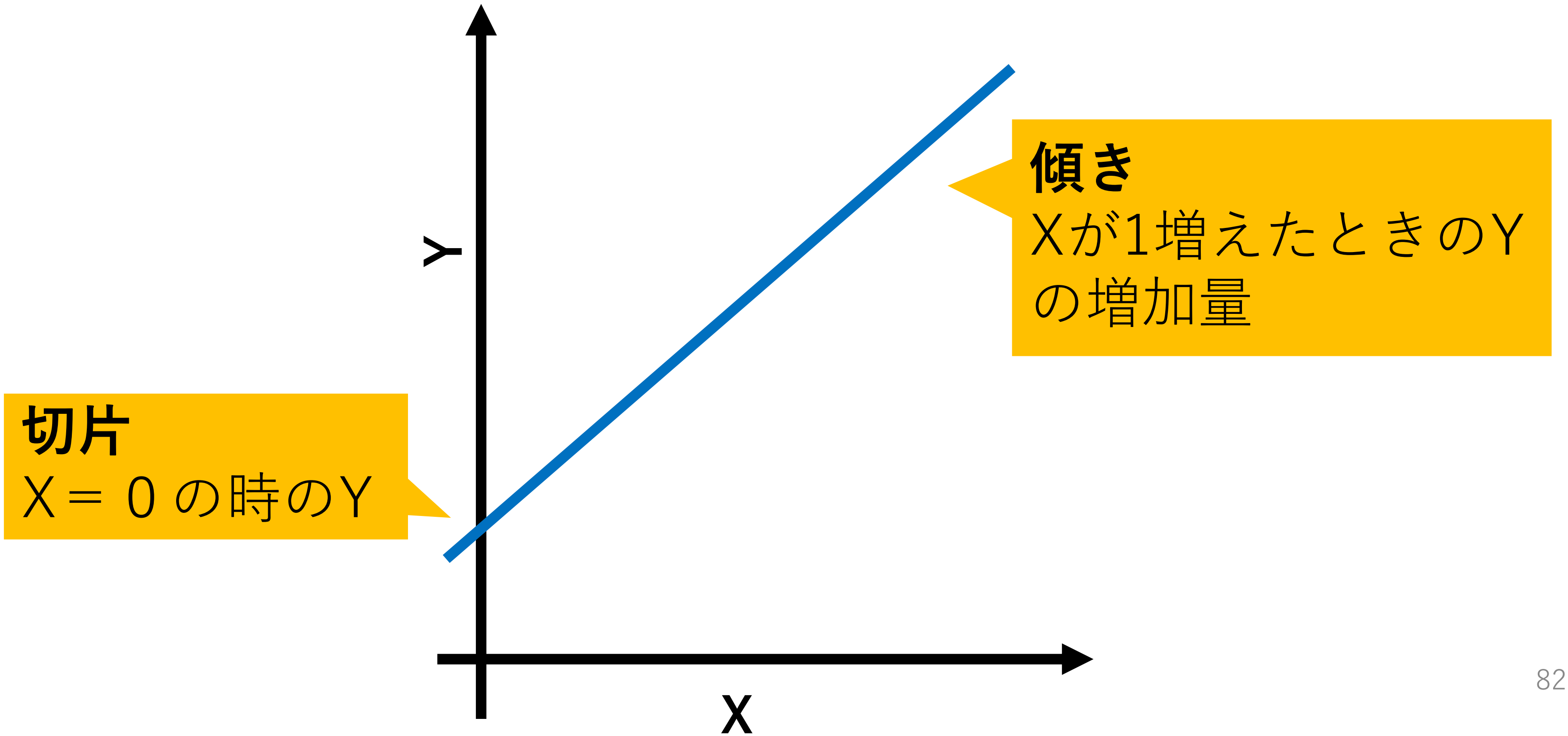
$$\theta_i = \mathrm{logistic}(\beta_0 + \beta_1 \cdot x_i)$$
$$Y_i \sim \mathrm{Bin}(n, \theta_i)$$

ロジスティック関数が適用
→係数β_0, β_1の解釈に注意

ロジスティック回帰モデル

<u>ロジスティック関数を適用すると何が変わるか</u>

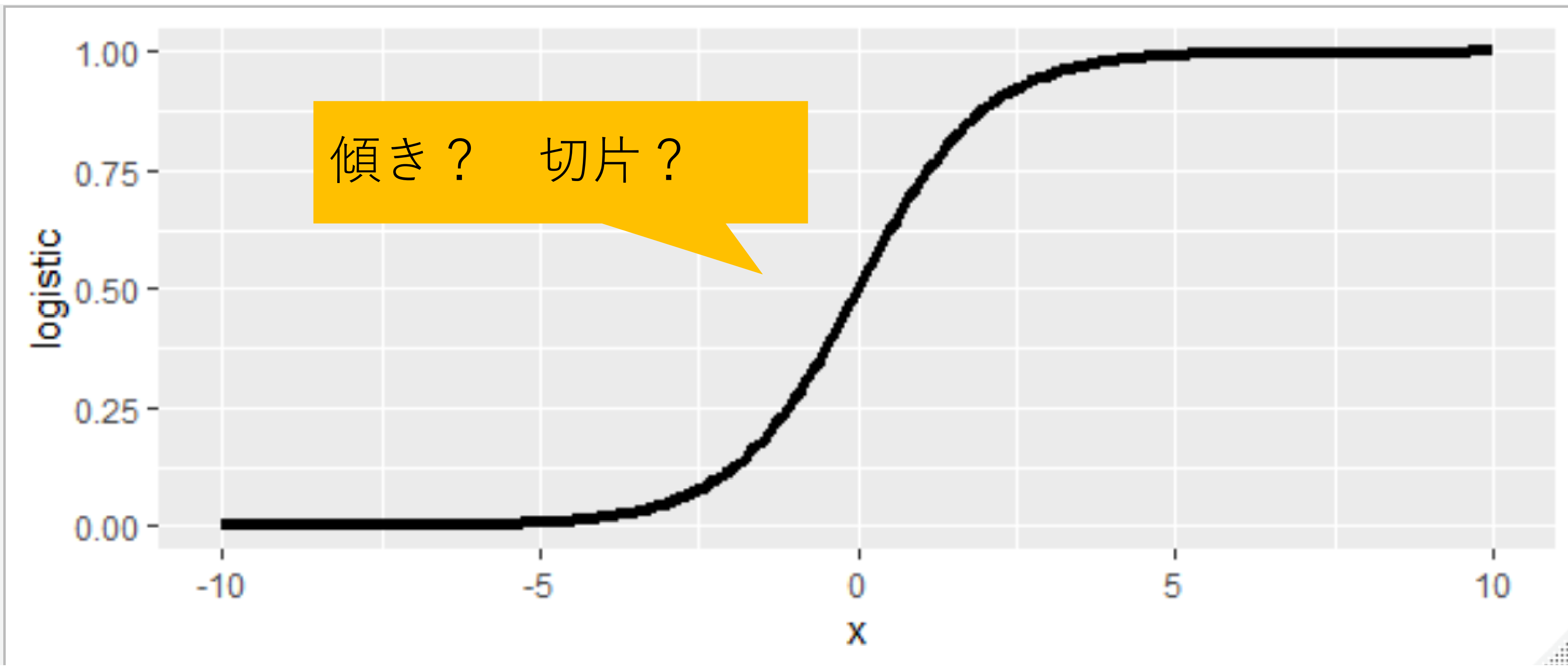
直線の関係性だと、係数の解釈は簡単



ロジスティック回帰モデル

ロジスティック回帰

$$\mathrm{logit}(p_i) = \beta_0 + \beta_1 \cdot x_i$$
$$Y_i \sim \mathrm{Bin}(n, p_i)$$



内容

1. 切片β_0の解釈
2. 傾きβ_1の解釈

ロジスティック回帰モデル

オッズ

$$\text{オッズ} = \frac{p}{1-p}$$ pは成功確率

オッズは「成功確率÷失敗確率」より
「失敗するよりも何倍成功しやすいか」を表すもの

対数オッズ

オッズの対数 $$\text{対数オッズ} = \log\frac{p}{1-p}$$

ロジスティック回帰モデル

ロジスティック回帰モデルの解釈

$$\mathrm{logit}(p_i) = \beta_0 + \beta_1 \cdot x_i$$
$$Y_i \sim \mathrm{Bin}(10, p_i)$$

β_0は、 xが0の時の
対数オッズである

ロジスティック回帰モデル

ロジット関数(参考)

$$\mathrm{logit}(p) = \log \frac{p}{1-p}$$

ロジット関数は
対数オッズの定義と同じ！

具体例

xが0の時の対数オッズ

$$\mathrm{logit}(p_0) = \beta_0 + \beta_1 \cdot 0$$

$$\log \frac{p_0}{1-p_0} = \beta_0$$

内容

1. 切片β_0の解釈

2. 傾きβ_1の解釈

ロジスティック回帰モデル

<u>オッズ比</u>

オッズの変化率　$\text{オッズ比} = \dfrac{p_2\text{のオッズ}}{p_1\text{のオッズ}} = \dfrac{p_2 / (1-p_2)}{p_1 / (1-p_1)}$

<u>対数オッズ比</u>

オッズ比の対数をとったもの

ロジスティック回帰モデル

ロジスティック回帰モデルの解釈

$$\mathrm{logit}(p_i) = \beta_0 + \beta_1 \cdot x_i$$
$$Y_i \sim \mathrm{Bin}(10, p_i)$$

$\boldsymbol{\beta_1}$は、$\boldsymbol{x}$が1単位増えた時の対数オッズ比である

ロジスティック回帰モデル

具体例

xが2から3に増えたとする

$$\mathrm{logit}(p_2) = \beta_0 + \beta_1 \cdot 2$$

$$\mathrm{logit}(p_3) = \beta_0 + \beta_1 \cdot 3$$

ロジスティック関数(参考)

$$\mathrm{logistic}(x) = \frac{1}{1 + \mathrm{EXP}(-x)}$$

両辺にロジスティック関数を適用

$\mathrm{logit}\big(\mathrm{logistic}(x)\big) = x$であることに注意

$$p_2 = \frac{1}{1 + \mathrm{EXP}\big(-(\beta_0 + \beta_1 \cdot 2)\big)}$$

$$p_3 = \frac{1}{1 + \mathrm{EXP}\big(-(\beta_0 + \beta_1 \cdot 3)\big)}$$

ロジスティック回帰モデル

具体例

$$p_2 = \frac{1}{1 + \mathrm{EXP}\big(-(\beta_0 + \beta_1 \cdot 2)\big)}$$

オッズ(参考)

$$\text{オッズ} = \frac{p}{1-p}$$

p_2のオッズをとる

$$\text{オッズ}(p_2) = \frac{p_2}{1-p_2} = \frac{\dfrac{1}{1 + \mathrm{EXP}\big(-(\beta_0 + \beta_1 \cdot 2)\big)}}{1 - \dfrac{1}{1 + \mathrm{EXP}\big(-(\beta_0 + \beta_1 \cdot 2)\big)}}$$

$$= \mathrm{EXP}(\beta_0 + \beta_1 \cdot 2)$$

ロジスティック回帰モデル

具体例

$$\text{オッズ}(p_2) = \mathrm{EXP}(\beta_0 + \beta_1 \cdot 2)$$

$$\text{オッズ}(p_3) = \mathrm{EXP}(\beta_0 + \beta_1 \cdot 3)$$

オッズ比をとる

$$\frac{\text{オッズ}(p_3)}{\text{オッズ}(p_2)} = \mathrm{EXP}(\beta_1)$$

対数をとる

$$\text{対数オッズ比} = \beta_1$$

Rを用いた実演