

11

JINSE活動報告シンポジウム  
 「論より統計！統計教育の改善に向けた活動」  
 日時：2016年2月27日(土)  
 会場：青山学院大学 本多記念国際会議場

(続)これからの統計教育  
 「統計基礎」の実際

確率変数と確率分布  
 標本分布論  
 BLUE, 検定の過誤, 検出力

12

## 「標本分布論」事始め

- ランダムな現象(確率変数)の結果を正確に予測することはできないが, 条件を整えて, 多数回その現象を観測するとヒストグラム(分布)は安定する
  - ヒストグラムの極限を(確率)分布という

$N(0,1)$  のヒストグラム  
n = 5000

$\text{Chi}^2(3)$  のヒストグラム  
n = 5000

13

### 確率的現象の「安定性」

Nが大きいとき, 確率的な現象の分布は安定する

2008年度(n=466,609)

初婚年齢(妻)

$\bar{x} = 27.69$

2011年度(n=418,223)

初婚年齢(妻)

$\bar{x} = 28.17$

初婚年齢(夫)

$\bar{y} = 29.37$

初婚年齢(夫)

$\bar{y} = 29.81$

14

### カイ二乗分布はどんな分布？

- 確率分布(密度関数の形)を予想させる

$N(0,1)$  のヒストグラム  
n = 200 Sturges公式

$\chi^2_1$  のヒストグラム  
n = 200 Sturges公式

N(0,1)

## 密度関数の形を予想

--学生に聞いてみる--

15

	H19	H20	H21	H24
0	a few	11	1	2
有限値	many	many	15	many
$+\infty$	some	12	4	3

$$Y = X^2 \text{ with } X \sim N(0, 1)$$

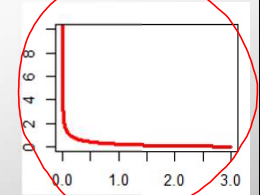
$$\begin{aligned} h(y) &= \frac{d}{dy} P(Y \leq y) = \frac{d}{dy} P(-\sqrt{y} \leq X \leq \sqrt{y}) \\ &= \frac{d}{dy} \int_{-\sqrt{y}}^{\sqrt{y}} \frac{1}{\sqrt{2\pi}} e^{-\frac{x^2}{2}} dx = \frac{d}{dy} 2 \int_0^{\sqrt{y}} \frac{1}{\sqrt{2\pi}} e^{-\frac{x^2}{2}} dx \\ &= \frac{1}{\sqrt{2\pi}} y^{-1/2} e^{-y/2} \quad (y > 0) \end{aligned}$$

ゆえに

$$h(y) = \begin{cases} \frac{1}{\sqrt{2\pi}} y^{-1/2} e^{-y/2}, & \text{if } y > 0 \\ 0, & \text{if } y \leq 0 \end{cases}$$

意外に出来ない (高校の数学)

$$\frac{d}{dy} \int_{-\infty}^{g(y)} f(x) dx = f(g(y)) g'(y)$$



16

## 学生への課題

17

- シミュレーションによってカイニ乗分布  $X^2$  を確認せよ
  - 標本サイズを色々変えてみよ
  - e.g.,  $n=200, 1000, 50000$
- 自由度1のカイニ乗分布の確率密度関数を描け
  - $dchisq(x, 1)$

## 学生のクレーム

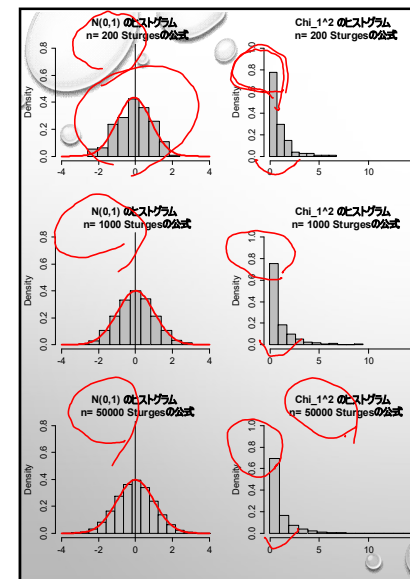
- $n$ を大きくしてヒストグラムを描いてみましたが、天から降りてきません!

どやっ!

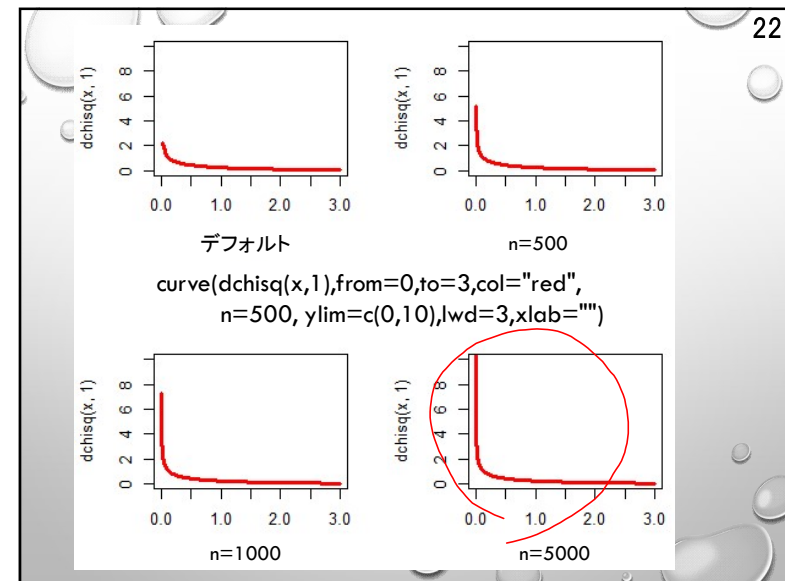
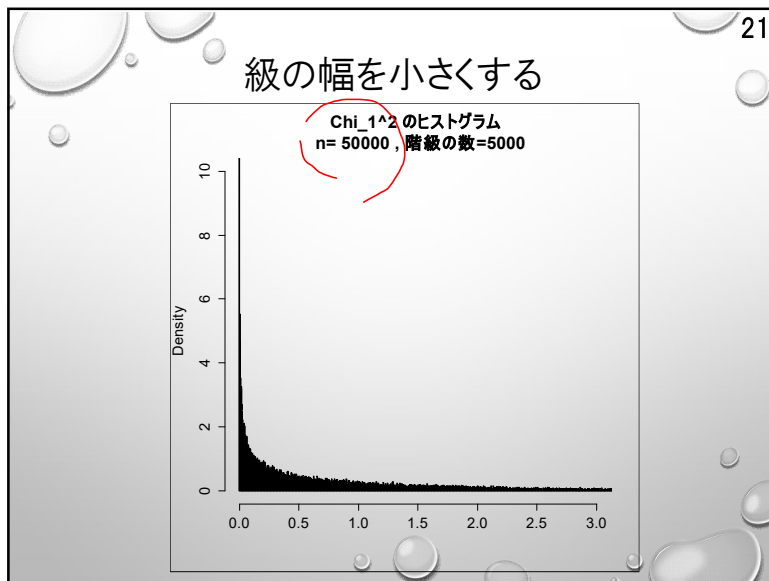
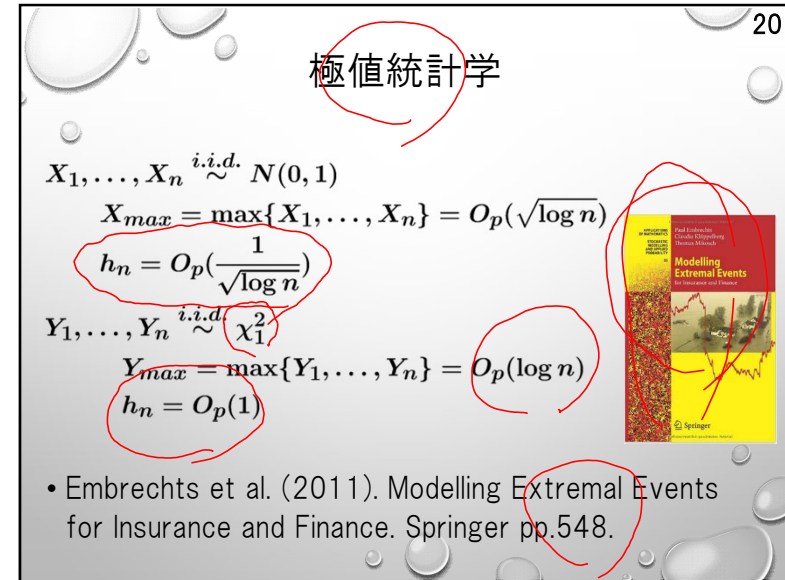
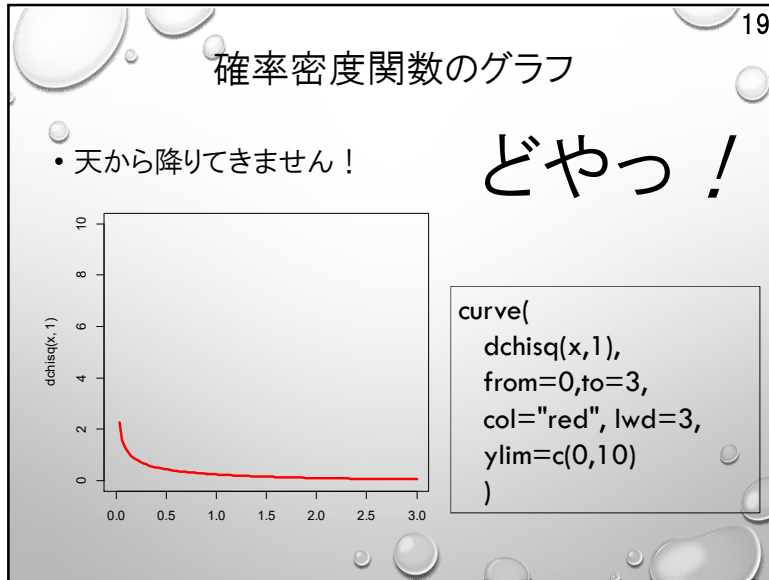
- Sturgesの公式

$$k_n = 1 + \log_2 n$$

$$h_n = (X_{max} - X_{min}) / k_n$$



18



23

## 最後に

- データ科学の4要素
  1. 統計学の概念と手法
  2. 現象の理解
  3. 数学
    - 最大値の分布
    - 収束のオーダー
    - 関数の連続の一様性
  4. 情報技術
    - グラフ描画の基礎
- Four "I's" by the Data Science Institute at Imperial College
  - Intelligence
  - Integration
  - Interdisciplinary
  - Interaction

24

## まとめ2013

おとなの学問, 統計学  
 もっと見よう数字とデータ  
 て”一たとお友達になろう  
 なんと最強統計学  
 おしまい

25

## 本日のまとめ2015

とても大事な基礎統計  
 リテラシーとしての統計学  
 プロジェクトベースの課題解決  
 ルールを守って統計解析  
 スルーできない情報技術  
 リろんと現実つなぐ統計  
 いちおく総活躍「統計学」

日本統計学会会長  
 JINSIE 運営委員

内閣総理大臣

26

## おわり

二人の子ども問題  
 封筒問題