政治学方法論 II

矢内　勇生

ベイズデータ分析の手順

1. リサーチクエスチョンに答えられるデータを特定する
2. 特定したデータを記述する統計（確率）モデルを作る
3. 推定する母数の事前分布を特定する
4. Bayes Rule に従い、情報を更新する
5. 事後予測分布がデータにフィットしているか確かめる

(Kruschke 2014: 25)

ベイズ分析における統計モデル

完全な確率モデルが必要！

分析に関連するすべての数量（観察可能なものと観察不可能なものを含む）の同時確率 (joint probability) を定式化することが必要

確率 (probability) とは？

- 不確実性 (uncertainty) を表現する手段
- 統計学における「公用語」

標本空間 (sample space) と事象 (event)

- 定義：ある試行 (experiment) で起こり得る全ての結果の集合をその試行の「標本空間 S」と呼ぶ。
- 定義：標本空間 S の部分集合を「事象 A」と呼び、試行の結果がA に含まれるとき、「A が起きた」と言う。

集合

- 和事象 (unions)
- 積事象 (intersections)
- 余事象 (complements)
- ド・モルガンの法則 (De Morgan’s laws)

長期的に得られる相対頻度としての確率

- ある事象の確率：ある試行を長期にわたって繰り返すとき、（回数n : $n \to \infty$ ）その事象が起きる割合

「信念」としての確率

- 「賭け」は公平か
  - 賭け1：サイコロを投げて偶数が出たら1万円もらえる
  - 賭け2：明日、雨が降らなかったら1万円もらえる

★どちらに掛ける？

確率の公理 (by Kolmogorov)

- 標本空間 *S*、*S* の事象（部分集合）*A*に実数を割り当てる関数 *P*(*A*) が与えられたとき、以下の性質を満たす*P*(*A*) を「確率」と呼ぶ：

1. $P(A) \geq 0, \forall A \subseteq S$

2. $P(S) = 1$

3. 事象 $A_1, A_2, \ldots, A_n$ が排反であるとき、

$$P(A_1 \cup A_2 \cup \cdots \cup A_n) = \sum_{i=1}^{n} P(A_i)$$

確率分布

- 起こり得るすべての結果（標本空間内のすべての要素）とそれに対応する確率の一覧表
    - 離散 (discrete)
    - 連続 (continuous, contiguous)

確率分布の例（離散）

コイン投げの確率分布

結果	表	裏
確率	θ	1 - θ

サイコロ投げの確率分布

結果	1	2	3	4	5	6
確率	θ_1	θ_2	θ_3	θ_4	θ_5	θ_6

確率質量関数 (PMF)

- Probability mass function: PMF
- 離散確率変数 X が「ある値 x」になる確率を与える関数

確率質量関数の例

確率 θ で表が出るコインを投げたとき、出る面をXとする。表を1、裏を0とすると、$X = x$ となる確率は、

$$p(x) = \theta^x (1 - \theta)^{1-x}$$

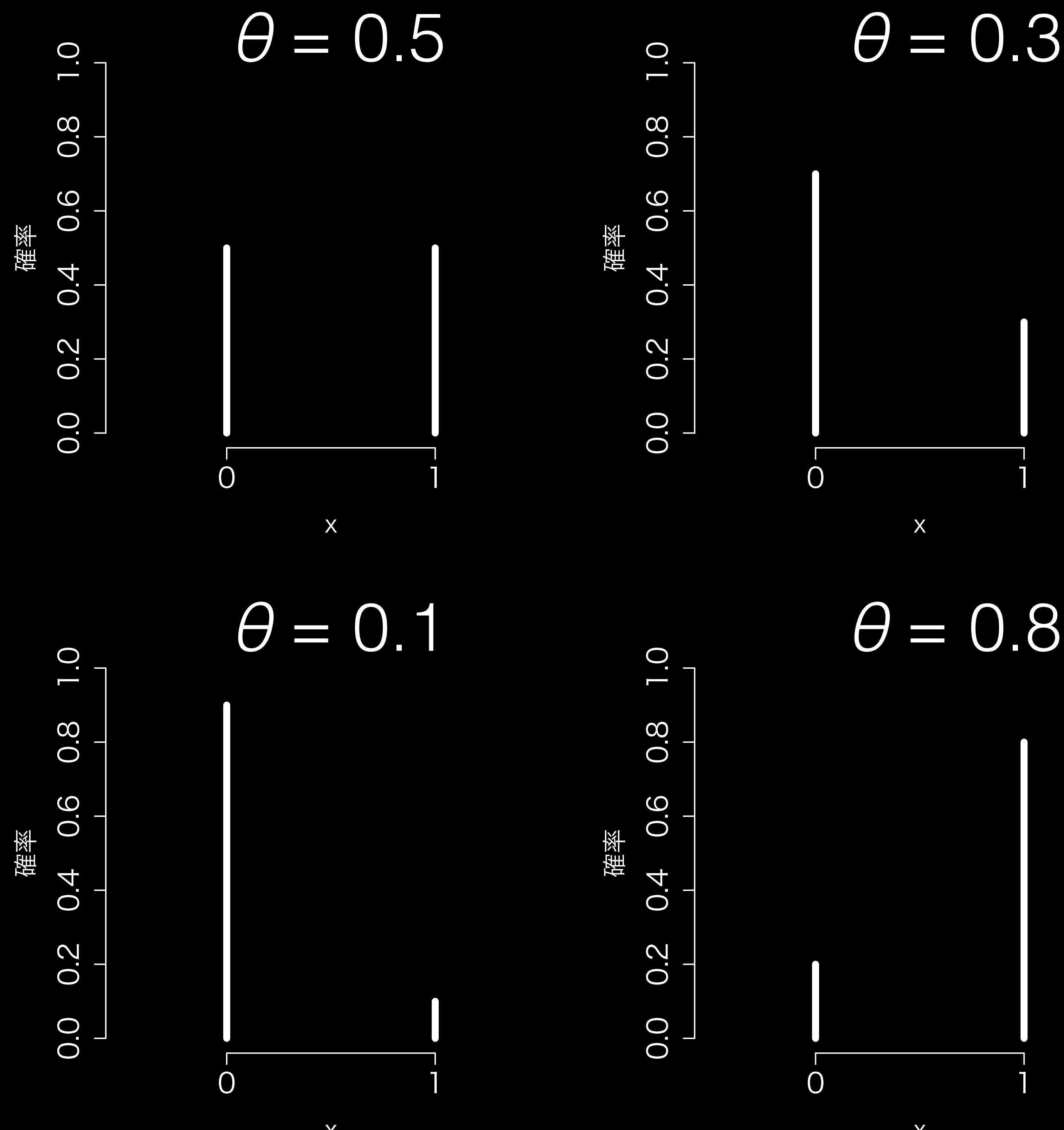
で与えられる。ただし、 $x \in \{0, 1\}$

確率質量関数の性質

$$p(x) \geq 0$$

$$\sum_x p(x) = 1$$

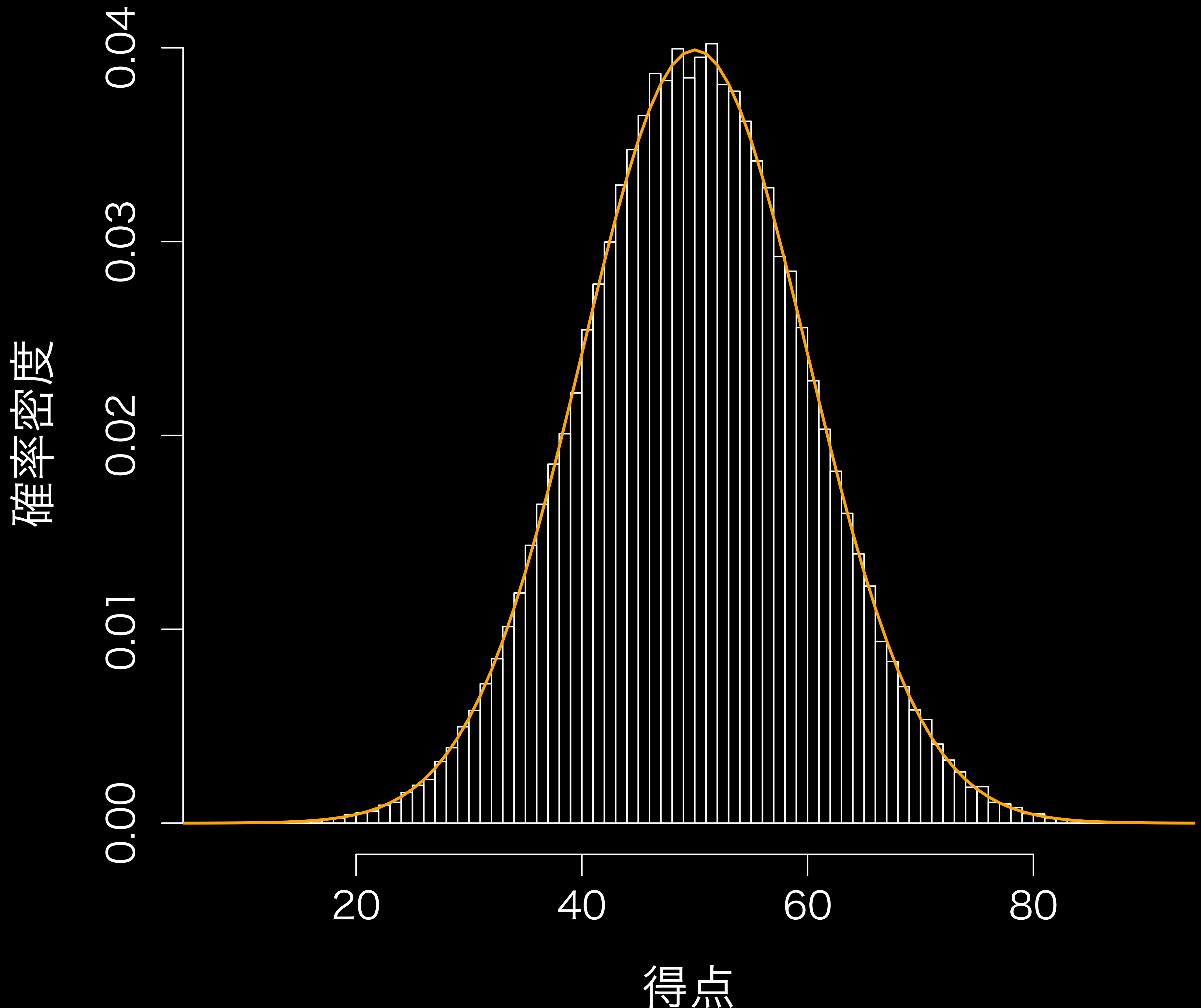
ベルヌーイ分布 Bern(θ)



確率密度関数 (PDF)

- Probability density function: PDF
- 連続確率変数Xが取り得る値での相対尤度を与える関数

ヒストグラムと確率密度関数



離散から連続へ

- ヒストグラムの階級幅を極限まで小さくすると、曲線に近づく
- その曲線を確率密度関数（pdf）と呼ぶ
- 注）確率密度 ≠ 確率

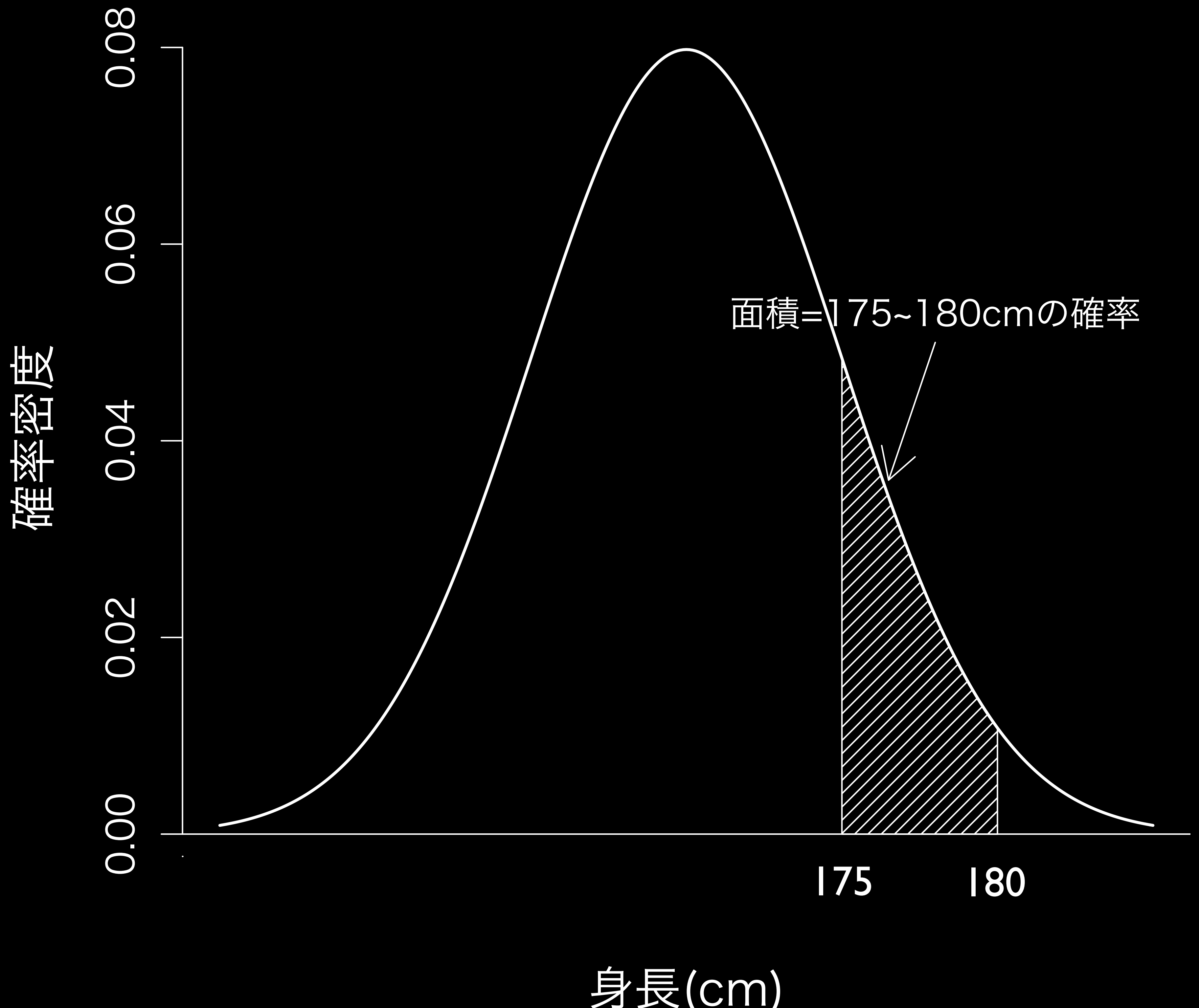
確率密度関数の特徴

- 横軸と確率密度関数の曲線に囲まれた部分の面積の合計が1（100%）になる

$$\int_x p(x)dx = 1$$

- 縦軸は、ある1点（$X = x$）の確率密度を表す
  - ある1点の面積は0 → 1点の確率は0
- 確率はある区間 $x_1 \leq X \leq x_2$ について与えられる

確率密度関数と確率



累積分布関数 (CDF)

- Cumulative distribution function (CDF)
- 確率変数Xがある値 x 以下である確率を与える関数

$$F(x) = \Pr(X \leq x)$$

$$\int_{-\infty}^{x} p(x)dx = F(x)$$

$$\lim_{x \to \infty} F(x) = \int_{-\infty}^{\infty} p(x)dx = 1$$

期待値 (mean, expected value)

- 加重平均：変数 (X) を 確率（密度）で重み付けした平均
  - 離散 $$\mathrm{E}(X) = \sum x_i p(x_i)$$
  - 連続 $$\mathrm{E}(X) = \int xp(x)dx$$

分散 (variance)

- 変数のばらつきを表す：不確実性の指標の１つ

$$\mathrm{var}(X) = \mathrm{E}[X - \mathrm{E}(X)]^2 = \mathrm{E}(X^2) - [\mathrm{E}(X)]^2$$

$$\mathrm{E}(X) = \sum x_i p(x_i)$$

$$\mathrm{E}(X^2) = \sum x_i^2 p(x_i)$$

期待値：分布の中心的傾向

- ある点Mの周りの分散を計算し、その分散を最小化する：分散を最小化するM ($\hat{M}$とする) は？

$$\min_{M} \mathrm{E}(X - M)^2$$

$$\hat{M} = \mathrm{E}(X)$$

条件付き確率 (conditional probability)

- 確率は、条件付け (conditioning) に基づいている

"*Conditioning is the soul of statistics*"
(Blitzstein and Hwang 2015: 42)

条件付き確率：定義

- AとBを事象とし、$P(B) > 0$ とすると、BにおけるAの条件付き確率 (probability of A given B) は、

$$P(A|B) = \frac{P(A, B)}{P(B)}$$

再びMonty Hall 問題

- Monty Hall 問題を条件付き確率で考える

ベイズルール（BR）

- BR: Bayes' Rule, Bayes Rule, Bayes's Rule, Bayes' Theorem, etc.

$$P(A|B) = \frac{P(A,B)}{P(B)}$$

$$P(B|A) = \frac{P(A,B)}{P(A)}$$

$$P(A|B) = \frac{P(B|A)P(A)}{P(B)}$$

合計確率の法則

- $A_1, \ldots, A_n$ が標本空間S を分割し（つまり、A_i が排反でその和事象がSである）、すべてのi について $P(A_i) \geq 0$ であるとする。このとき、

$$P(B) = \sum_{i=1}^{n} P(B|A_i)P(A_i)$$

BR + 合計確率

離散

$$p(\theta|y) = \frac{p(y|\theta)p(\theta)}{\sum_\theta p(y|\theta)p(\theta)}$$

連続

$$p(\theta|y) = \frac{p(y|\theta)p(\theta)}{\int p(y|\theta)p(\theta)d\theta}$$

条件付き確率は確率である

- 条件付き確率は、0以上1以下の値をとる
- $P(S|E) = 1, \quad P(\emptyset|E) = 0$
- $A_1, A_2, \dots$ が排反なら、$P(\bigcup_{j=1}^{\infty} A_j|E) = \sum_{j=1}^{\infty} P(A_j|E)$
- $P(A^c|E) = 1 - P(A|E)$
- $P(A \cup B|E) = P(A|E) + P(B|E) - P(A \cap B|E)$

独立 (independence)

- 事象AとBは、以下の条件を満たすとき独立であると定義される

$$P(A, B) = P(A)P(B)$$

- 独立 (independent) ≠ 排反 (disjoint)
  - 例：コイン投げにおける「表」と「裏」

条件付き独立

- 事象AとBが以下の条件を満たすとき、条件EのもとでAとBは独立である

$$P(A, B|E) = P(A|E)P(B|E)$$

- 条件付き独立は独立を意味しない！

独立 ➔ 条件付き独立？

問題

$$p(A,B) = p(A)p(B) \overset{?}{\not\Rightarrow} p(A,B|C) = p(A|C)p(B|C)$$

- 例
  - C：火災報知機の作動（作動の原因は2つ）
    - A: 火事
    - B：誤作動