

小角散乱データに基づく 試料パラメータとモデルのベイズ推論

2024/4/10 物性研究所セミナー

東京大学大学院 新領域創成科学研究科 複雑理工学専攻

岡田研究室 博士課程1年

林 悠偉

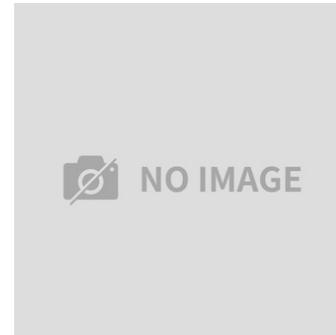
共同研究者(五十音順)



岡田真人
東京大学



片上舜
東京大学



桑本滋生
SPring-8/JASRI



永田賢二
物材研(NIMS)



水牧仁一朗
熊本大学

ベイズ推論と物性科学

2012年12月17日 物性研6階 大講義室

- 本日の招待講演の12年前にすでに物性研でベイズ推論の招待講演実施
- 講演の具体的内容
 - ベイズ推論の紹介とベイズ推論の分光学(スペクトル分解)への導入
 - 物理パラメータの事後確立分布とモデル選択
- 講演の今後の展望
 - スペクトル分解をロールモデルとし、ベイズ推論は物性科学を根本的に変革する普遍的情報数理的枠組み
- 講演のインパクトはなく、物性科学は何も変わらなかつた

ベイズ計測の物性科学における適用例

東京大学 岡田研究室

- 事後分布推定、モデル選択、ベイズ統合
 - 1. $y=ax+b$
 - 2. スペクトル分解
 - 3. X線光電子放出スペクトル(XPS)
 - 4. X線吸収スペクトル(XAS)
 - 5. メスバウア一分光
 - 6. X線小角散乱スペクトル
 - 7. NMR
 - 8. 中性子非弾性散乱スペクトル
 - 9. 比熱
 - 10. 帯磁率
 - 11. 比熱と帯磁率

NIMS

- 事後分布推定 and/or モデル選択
 - 1. XRD
 - 2. X線光電子放出スペクトル(XPS)

熊本大学 赤井研究室

- 事後分布推定 and/or モデル選択
 - 1. 可視光発光スペクトル
 - 2. 可視光吸収スペクトル
 - 3. EXAFS
 - 4. XANES
 - 5. XMCD
 - 6. 変調スペクトル
 - 7. XPS
 - 8. XRD
 - 9. コヒーレントフォノン

SPring-8/JASRI 熊本大学 水牧研究室

- 事後分布推定、モデル選択、ベイズ統合
 - 1. XRD
 - 2. XPSとXAS
 - 3. 磁気コンプトン散乱プロファイル
 - 4. ラマンスペクトル

目次

1. 自己紹介
2. ベイズ計測
 - ベイズ推論とベイズ計測
 - 物理パラメータの事後確率推定, モデル選択, 複数データの統合
3. 解析的に解けるベイズ計測 : $y = ax + b$
 - 物理パラメータの事後確率推定, モデル選択
4. 小角散乱のベイズ計測
 - 小角散乱の概説
 - 物理パラメータの事後確率推定 : 試料パラメータの推定
 - モデル選択 : 散乱強度モデルの選択
5. まとめ
 - ベイズ計測を用いた新たな物性科学のパラダイムの提案
 - ベイズ計測の社会実装に必須な条件

目次

1. 自己紹介
2. ベイズ計測
 - ベイズ推論とベイズ計測
 - 物理パラメータの事後確率推定, モデル選択, 複数データの統合
3. 解析的に解けるベイズ計測 : $y = ax + b$
 - 物理パラメータの事後確率推定, モデル選択
4. 小角散乱のベイズ計測
 - 小角散乱の概説
 - 物理パラメータの事後確率推定 : 試料パラメータの推定
 - モデル選択 : 散乱強度モデルの選択
5. まとめ
 - ベイズ計測を用いた新たな物性科学のパラダイムの提案
 - ベイズ計測の社会実装に必須な条件

自己紹介

名前：林 悠偉

東京大学大学院 新領域創成科学研究科
複雑理工学専攻 博士課程 1年

- 研究テーマ：小角散乱法のベイズ計測
- メール：hayashi@mns.k.u-tokyo.ac.jp



学術論文

- Yui Hayashi, Shun Katakami, Shigeo Kuwamoto, Kenji Nagata, Masaichiro Mizumaki and Masato Okada, Journal of the Physical Society of Japan, 92, 094002, August 2023.
arxiv : <https://arxiv.org/abs/2303.04983>
- Yui Hayashi, Shun Katakami, Shigeo Kuwamoto, Kenji Nagata, Masaichiro Mizumaki and Masato Okada, "Quantitative Selection of Sample Structures in Small-Angle Scattering Using Bayesian Methods", January 2024. (submitted to Journal of Applied Crystallography) *arxiv* : <https://arxiv.org/abs/2401.10466>

招待講演

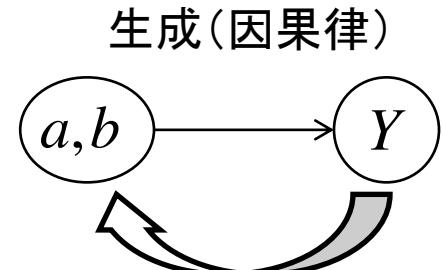
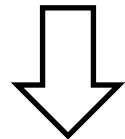
- 林悠偉, 片上舜, 桑本滋生, 永田賢二, 水牧仁一朗, 岡田真人, 「小角散乱法の計測データを用いた試料パラメータのベイズ推定」, 『第88回SPring-8先端利用技術ワークショップ/「第4回データ駆動科学によるデータ解析高度化～ベイズ計測～」』, SACLA実験研究棟大会議室 及びwebハイブリッド開催, 2023年 3月.

目次

1. 自己紹介
2. ベイズ計測
 - ベイズ推論とベイズ計測
 - 物理パラメータの事後確率推定, モデル選択, 複数データの統合
3. 解析的に解けるベイズ計測 : $y = ax + b$
 - 物理パラメータの事後確率推定, モデル選択
4. 小角散乱のベイズ計測
 - 小角散乱の概説
 - 物理パラメータの事後確率推定 : 試料パラメータの推定
 - モデル選択 : 散乱強度モデルの選択
5. まとめ
 - ベイズ計測を用いた新たな物性科学のパラダイムの提案
 - ベイズ計測の社会実装に必須な条件

自然科学的視点からのベイズ推論

$$p(Y, a, b) = \underbrace{p(Y | a, b)p(a, b)}_{\text{生成(因果律)}}$$



<ベイズの定理>

$$p(a, b | Y) = \frac{p(Y | a, b)p(a, b)}{p(Y)} \propto \exp(-nE(a, b))p(a, b)$$

$p(a, b | Y)$: 事後確率. データが与えられたもとでの, パラメータの確率.

$p(a, b)$: 事前確率. あらかじめ設定しておく必要がある.
これまで蓄積されてきた科学的知見

ベイズ計測の定義

- ・ベイズ推論には多種多様な枠組みがある.
- ・ベイズ計測は、計測科学に必要なベイズ推論の以下の3つの枠組みである.

物理パラメータの事後確率推定

従来の最小二乗法とは異なり物理パラメータを誤差評価付きで推定できる。評価関数のローカルミニマムを避けることができる。

モデル選択

実験データを説明する複数の数理モデルの候補の内、どれが適切かを数理的に選択できる。

複数データの統合

比熱と磁化率のように一つの物質を異なる実験で計測した場合のデータを数理的に統合し解析する。

ベイズ計測のご利益

■物理パラメータを確率分布として推定

大域的最適解の推定、推定結果の信頼度を評価可能

■モデル選択

- ・データに対する数理モデルの妥当性を定量化
- ・適切なモデル化によって精度改善

■複数の計測データを統合解析

解析精度の向上の可能性

ベイズ計測の物性科学における適用例

東京大学 岡田研究室

- 事後分布推定、モデル選択、ベイズ統合

1. $y=ax+b$
2. スペクトル分解
3. X線光電子放出スペクトル(XPS)
4. X線吸収スペクトル(XAS)
5. メスバウア一分光
6. X線小角散乱スペクトル
7. NMR
8. 中性子非弾性散乱スペクトル
9. 比熱
10. 帯磁率
11. 比熱と帯磁率

NIMS

- 事後分布推定 and/or モデル選択

1. XRD
2. X線光電子放出スペクトル(XPS)

熊本大学 赤井研究室

- 事後分布推定 and/or モデル選択
1. 可視光発光スペクトル
 2. 可視光吸収スペクトル
 3. EXAFS
 4. XANES
 5. XMCD
 6. 変調スペクトル
 7. XPS
 8. XRD
 9. コヒーレントフォノン

SPring-8/JASRI 熊本大学 水牧研究室

- 事後分布推定、モデル選択、ベイズ統合

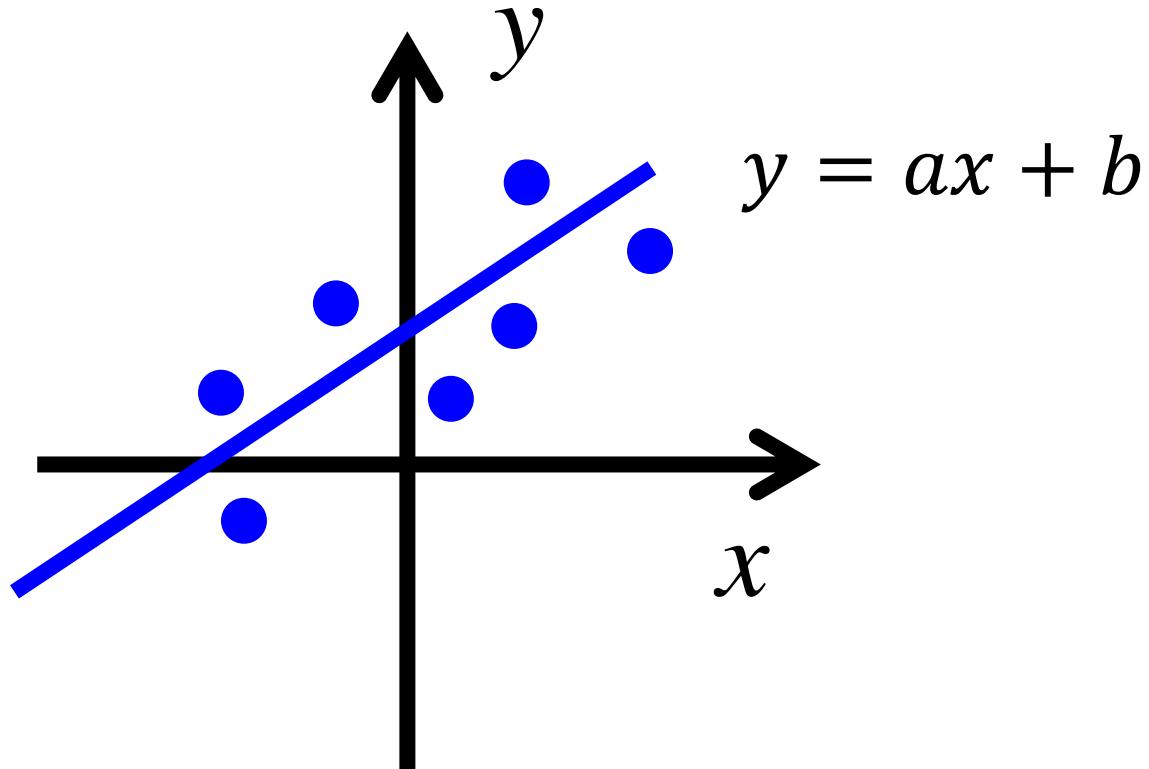
1. XRD
2. XPSとXAS
3. 磁気コンプトン散乱プロファイル
4. ラマンスペクトル

ベイズ計測の枠組みは様々な計測データに適用できる
(シミュレーションはノートPCで実行できる場合が多い)

目次

1. 自己紹介
2. ベイズ計測
 - ベイズ推論とベイズ計測
 - 物理パラメータの事後確率推定, モデル選択, 複数データの統合
3. 解析的に解けるベイズ計測 : $y = ax + b$
 - 物理パラメータの事後確率推定, モデル選択
4. 小角散乱のベイズ計測
 - 小角散乱の概説
 - 物理パラメータの事後確率推定 : 試料パラメータの推定
 - モデル選択 : 散乱強度モデルの選択
5. まとめ
 - ベイズ計測を用いた新たな物性科学のパラダイムの提案
 - ベイズ計測の社会実装に必須な条件

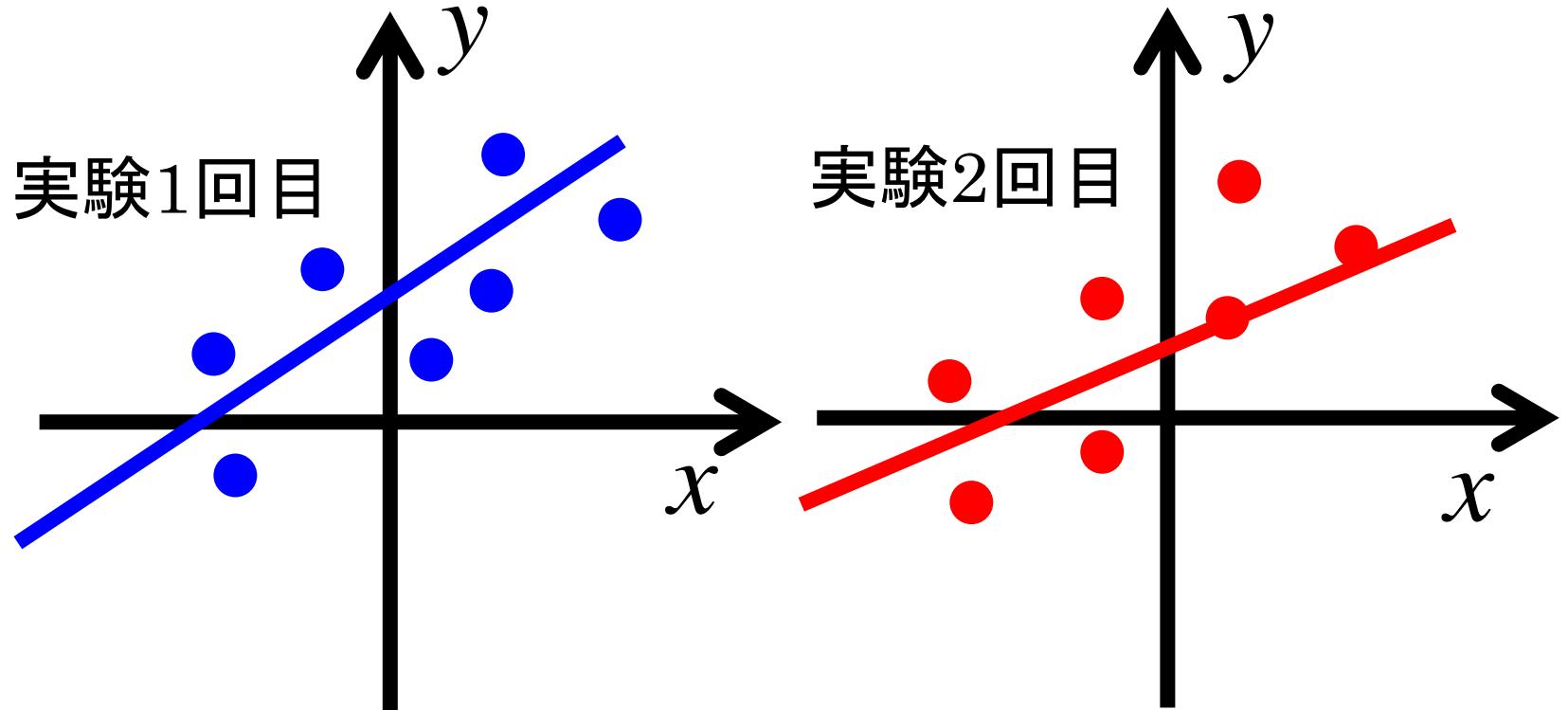
$y = ax + b$ のパラメータ推定



傾き a : バネ定数, 電気伝導度, 誘電率, 系の線形応答

モデルのパラメータ a, b を推定する → 物理量を評価

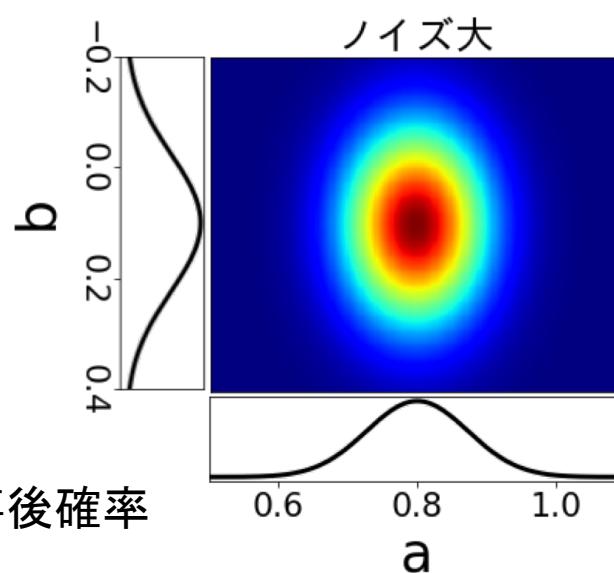
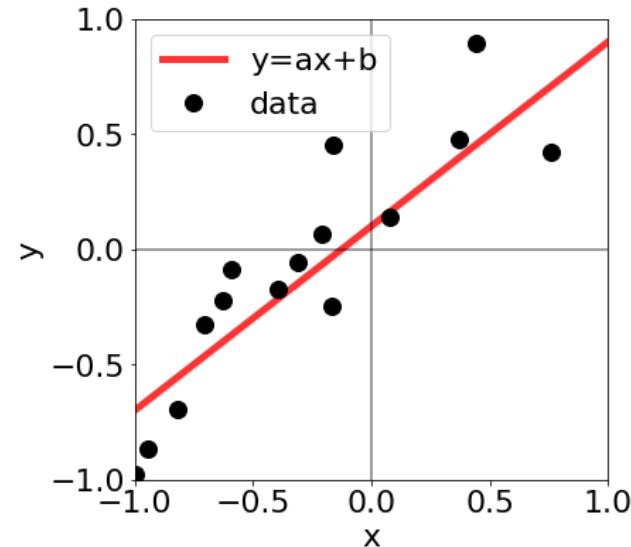
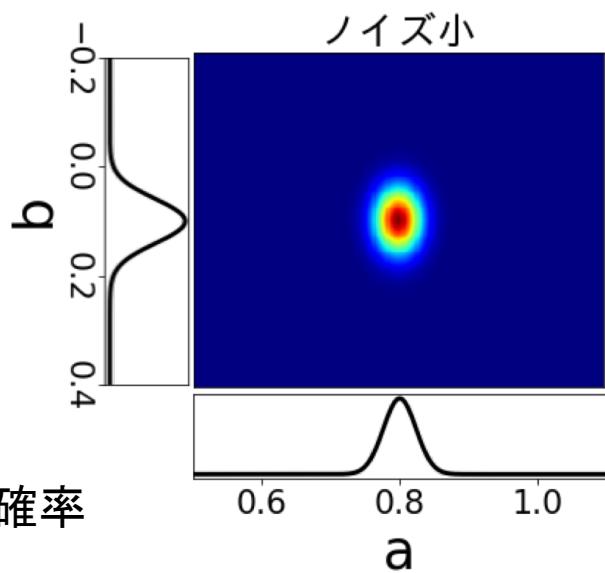
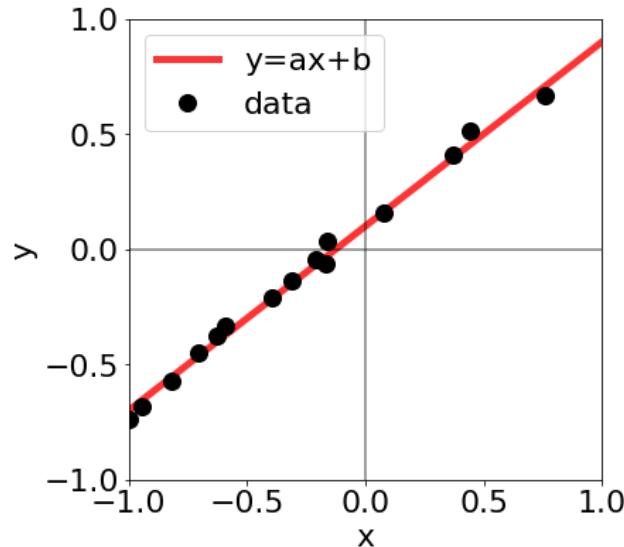
データのばらつきを考慮した物理量の評価



実験を複数回おこなって、 α のばらつきを見る

これを1回の実験データから求められないか→ ベイズ計測

まずは結論: $y = ax + b$ のベイズ計測



確率分布の幅として信頼度(不確かさ)を推定出来る

最小二乗法 (1/2)

$$E(a, b) = \frac{1}{N} \sum_{i=1}^N (y_i - (ax_i + b))^2$$

モデル

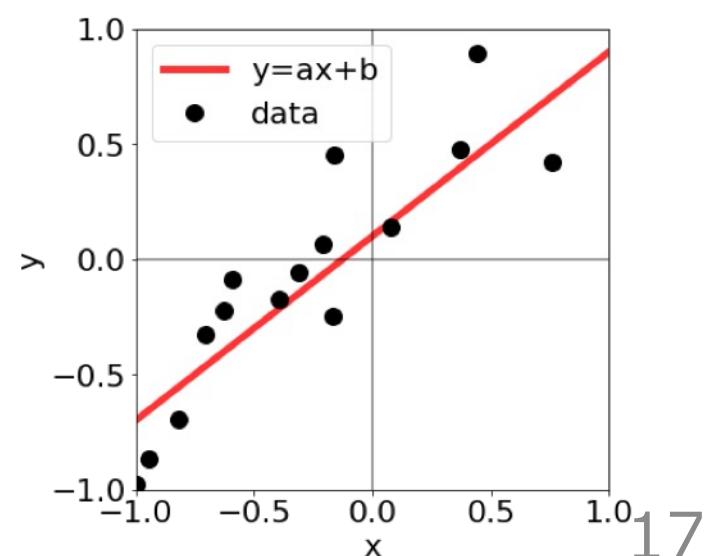
二乗誤差 $E(a, b)$ を最小にするようにパラメータをフィット(最小二乗法)

$$\bar{x} = \frac{1}{N} \sum_{i=1}^N x_i = 0 \text{ を仮定する}$$

$$E(a, b) = \frac{1}{2} \left(\overline{x^2} \left(a - \frac{\overline{xy}}{\overline{x^2}} \right)^2 + (b - \bar{y})^2 - \frac{\overline{xy}^2}{\overline{x^2}} - \bar{y}^2 + \overline{y^2} \right) \quad a_0 = \frac{\overline{xy}}{\overline{x^2}} \quad b_0 = \bar{y}$$
$$= \mathcal{E}_a(a) + \mathcal{E}_b(b) + E(a_0, b_0) \geq E(a_0, b_0)$$

平均: $\bar{y} = \frac{1}{N} \sum_{i=1}^N y_i$, 分散: $\overline{x^2} = \frac{1}{N} \sum_{i=1}^N x_i^2$

$$\overline{xy} = \frac{1}{N} \sum_{i=1}^N x_i y_i$$

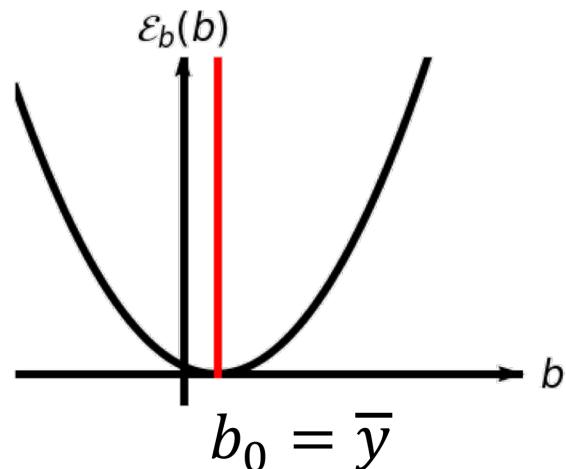
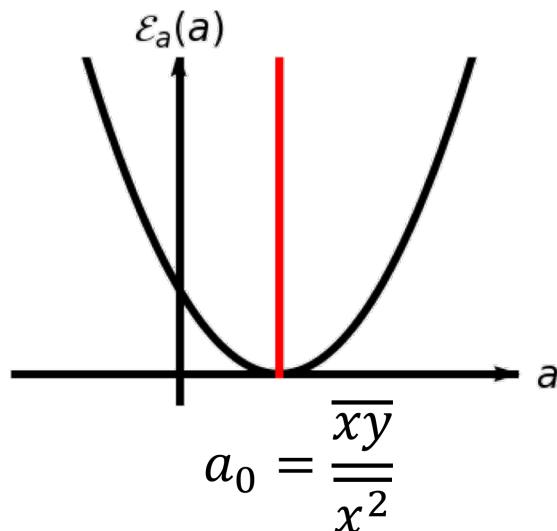


最小二乗法 (2/2)

二乗誤差 : $E(a, b) = \frac{1}{N} \sum_{i=1}^N (y_i - (ax_i + b))^2$

$$E(a, b) = \frac{1}{2} \left(\overline{x^2} \left(a - \frac{\overline{xy}}{\overline{x^2}} \right)^2 + (b - \bar{y})^2 - \frac{\overline{xy}^2}{\overline{x^2}} - \bar{y}^2 + \overline{y^2} \right) \quad a_0 = \frac{\overline{xy}}{\overline{x^2}} \quad b_0 = \bar{y}$$

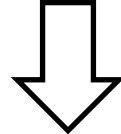
$$E(a, b) = \mathcal{E}_a(a) + \mathcal{E}_b(b) + E(a_0, b_0) \geq E(a_0, b_0)$$



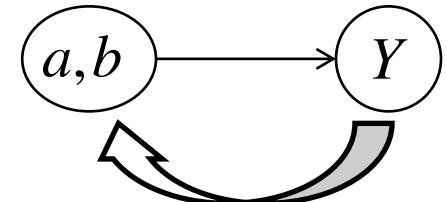
ベイズ計測：事後確率推定 自然科学的視点からのベイズの定理

Y : データ; a, b : パラメータ

$$p(Y, a, b) = \underline{p(Y|a, b)p(a, b)} = p(a, b|Y)p(Y)$$



生成(因果律)



<ベイズの定理>

$$p(a, b|Y) = \frac{p(Y|a, b)p(a, b)}{p(Y)} \propto \exp(-nE(a, b))p(a, b)$$

$p(a, b|Y)$: 事後確率. データが与えられたもとでの, パラメータの確率.

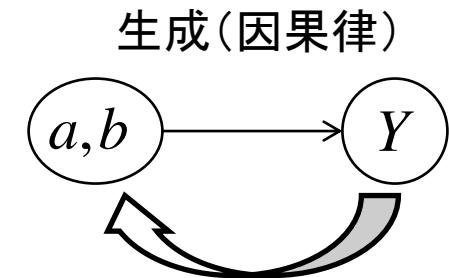
$p(a, b)$: 事前確率. あらかじめ設定しておく必要がある.
これまで蓄積されてきた科学的知見

$p(a, b|Y)$ の推定 (1/3) 最小二乗法では手で計算できる

データ点 : $y_i = ax_i + b + n_i$

ノイズの分布 : $p(n_i) = \frac{1}{\sqrt{2\pi}\sigma} \exp\left(-\frac{n_i^2}{2\sigma^2}\right)$

$$\begin{aligned} p(n_i) &= p(y_i|a, b) \\ &= \frac{1}{\sqrt{2\pi}\sigma} \exp\left(-\frac{(y_i - (ax_i + b))^2}{2\sigma^2}\right) \end{aligned}$$



データの生成確率(尤度) :

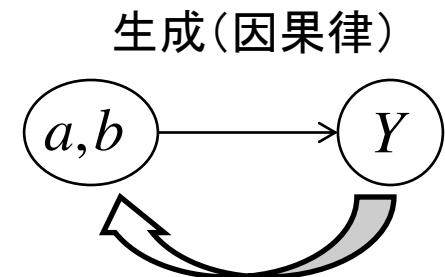
$$\begin{aligned} p(Y|a, b) &= \prod_{i=1}^N p(y_i|a, b) \\ &= \left(\frac{1}{\sqrt{2\pi}\sigma}\right)^N \exp\left(-\frac{\sum_{i=1}^N (y_i - (ax_i + b))^2}{2\sigma^2}\right) \\ &= \left(\frac{1}{\sqrt{2\pi}\sigma}\right)^N \exp\left(-\frac{N}{\sigma^2} E(a, b)\right) \end{aligned}$$

二乗誤差

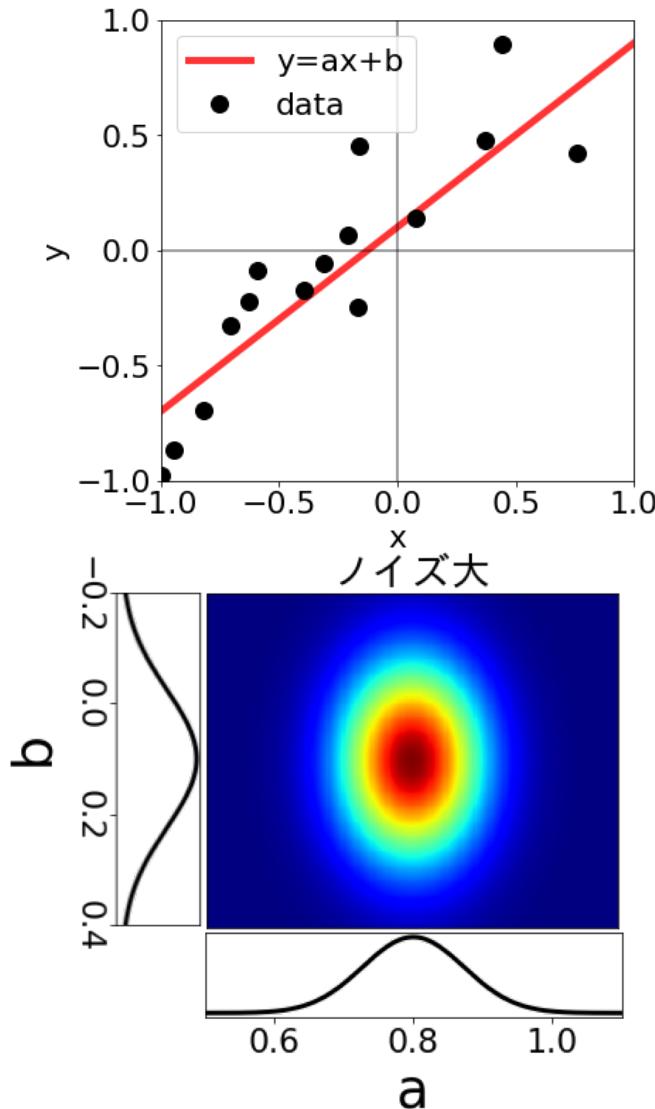
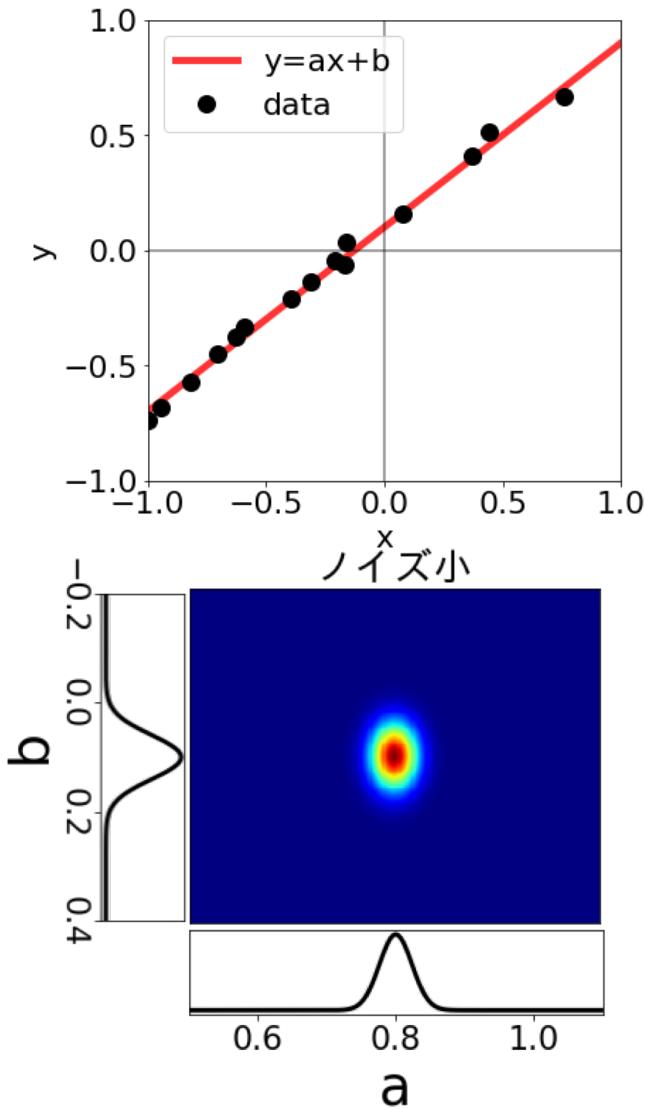
$p(a, b|Y)$ の推定 (2/3) 最小二乗法では手で計算できる

$$\begin{aligned} p(a, b|Y) &= \frac{p(Y|a, b)p(a, b)}{p(Y)} \propto p(Y|a, b) \\ &= \exp \left(-\frac{N}{\sigma^2} (\mathcal{E}_a(a) + \mathcal{E}_b(b) + E(a_0, b_0)) \right) \\ &\propto \exp \left(-\frac{N}{\sigma^2} (\mathcal{E}_a(a) + \mathcal{E}_b(b)) \right) \\ &= \exp \left(-\frac{N}{2\sigma^2} \left\{ \bar{x}^2 (a - a_0)^2 + (b - b_0)^2 \right\} \right) \end{aligned}$$

($p(a, b)$ に一様分布を仮定)



$p(a, b|Y)$ の推定 (3/3) 最小二乗法では手で計算できる

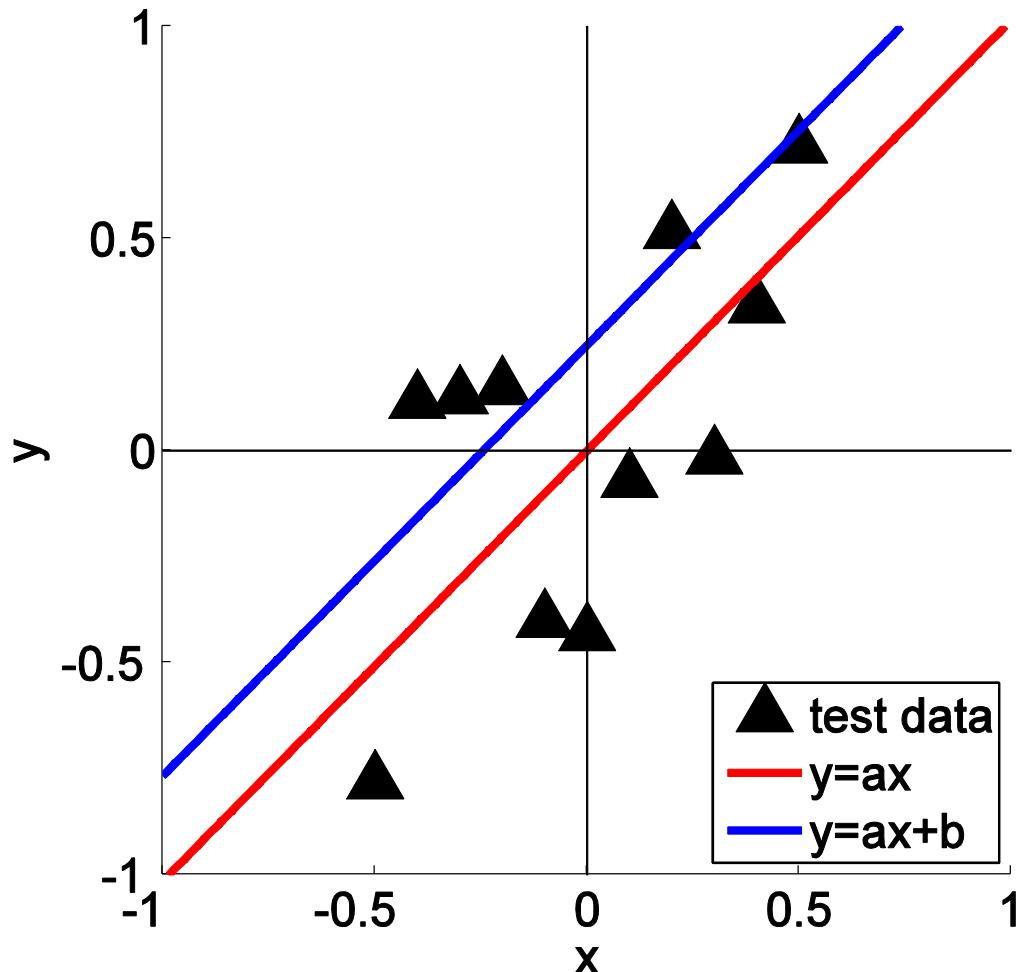


目次

1. 自己紹介
2. ベイズ計測
 - ベイズ推論とベイズ計測
 - 物理パラメータの事後確率推定, モデル選択, 複数データの統合
3. 解析的に解けるベイズ計測 : $y = ax + b$
 - 物理パラメータの事後確率推定, モデル選択
4. 小角散乱のベイズ計測
 - 小角散乱の概説
 - 物理パラメータの事後確率推定 : 試料パラメータの推定
 - モデル選択 : 散乱強度モデルの選択
5. まとめ
 - ベイズ計測を用いた新たな物性科学のパラダイムの提案
 - ベイズ計測の社会実装に必須な条件

モデル選択

$y = ax$ か $y = ax + b$ か?



	訓練誤差
$y = ax$	0.43
$y = ax + b$	0.34

	汎化誤差
$y = ax$	0.33
$y = ax + b$	0.40

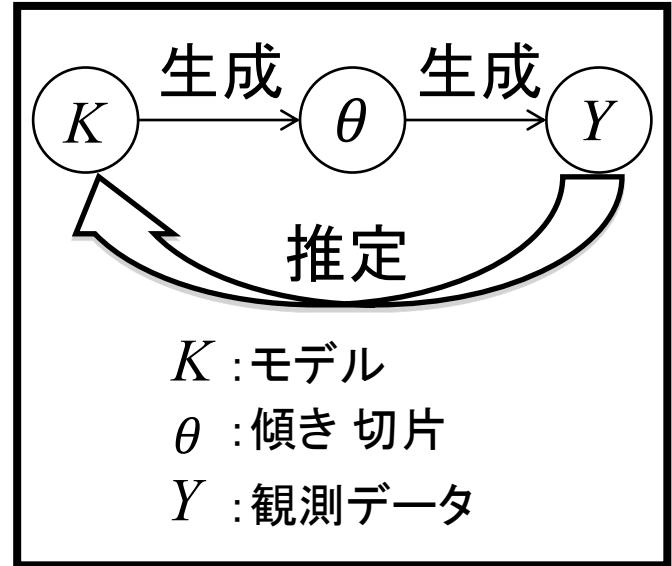
データに対する数理モデルの妥当性を理論的に評価したい

ベイズ計測：モデル選択

- 欲しいのは $p(K|Y)$
- $\theta = \{a, b\}$ がないぞ
- $p(K, \theta, Y)$ の存在を仮定

$$p(K, \theta, Y) = p(Y|\theta, K)p(K)$$

$$p(Y|\theta, K) = \prod_{i=1}^n p(y_i|\theta) \propto \exp(-nE(\theta))$$



- 無駄な自由度の系統的消去: 周辺化

$$p(K, Y) = \int p(K, \theta, Y) d\theta$$

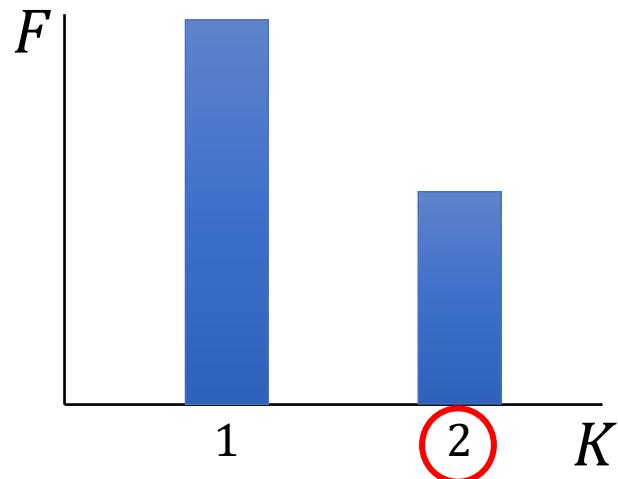
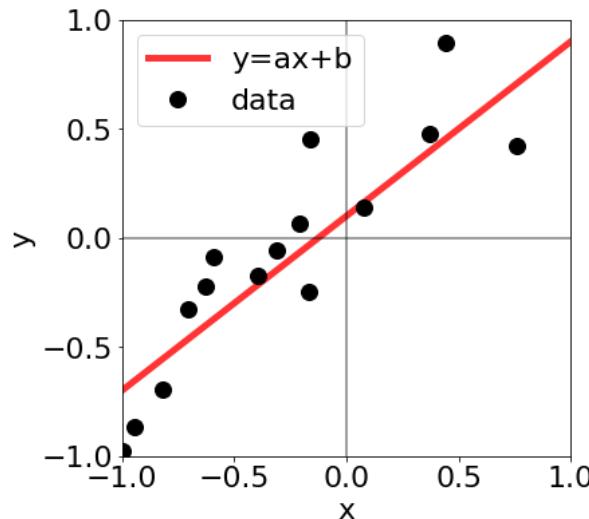
$$p(K|Y) = \frac{p(Y|K)p(K)}{p(Y)} \propto \int \exp(-nE(\theta)) p(\theta) d\theta$$

($\because p(K)$ に一様分布を仮定)

$$F(K) = -\log \int \exp(-nE(\theta)) p(\theta) d\theta$$

自由エネルギーを最小にするモデル K を求める。

モデル選択: 自由エネルギー $y = ax$ か $y = ax + b$ か?



- $K = 1 : y = ax$
- $K = 2 : y = ax + b$

$$F(K=1) = N \left\{ \frac{1}{\sigma^2} E(a_0) + \frac{\log N}{2N} \right\}$$

$$F(K=2) = N \left\{ \frac{1}{\sigma^2} E(a_0, b_0) + \frac{\log N}{N} \right\}$$

数理的にデータ生成のモデルを選択できる

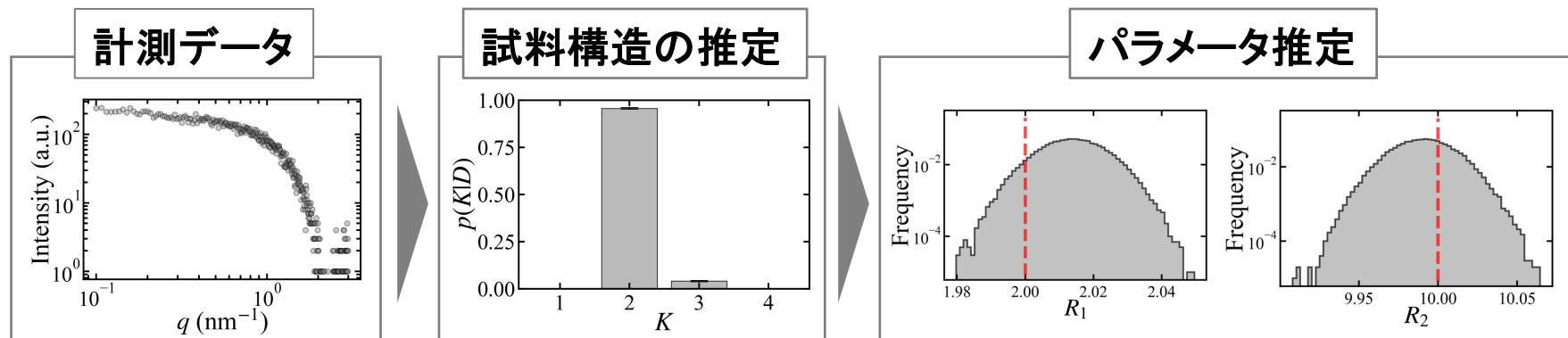
目次

1. 自己紹介
2. ベイズ計測
 - ベイズ推論とベイズ計測
 - 物理パラメータの事後確率推定, モデル選択, 複数データの統合
3. 解析的に解けるベイズ計測 : $y = ax + b$
 - 物理パラメータの事後確率推定, モデル選択
4. 小角散乱のベイズ計測
 - 小角散乱の概説
 - 物理パラメータの事後確率推定 : 試料パラメータの推定
 - モデル選択 : 散乱強度モデルの選択
5. まとめ
 - ベイズ計測を用いた新たな物性科学のパラダイムの提案
 - ベイズ計測の社会実装に必須な条件

小角散乱のベイズ計測の概要

■ 小角散乱法の計測データ解析にベイズ推論を導入する.

- 試料パラメータを確率分布として推定
→大域的最適解の推定 & 信頼度の推定
- 散乱強度の数理モデル選択
→候補の数理モデルのデータに対する妥当性を定量化し、選択する
→データから試料構造を定量的に選択できる

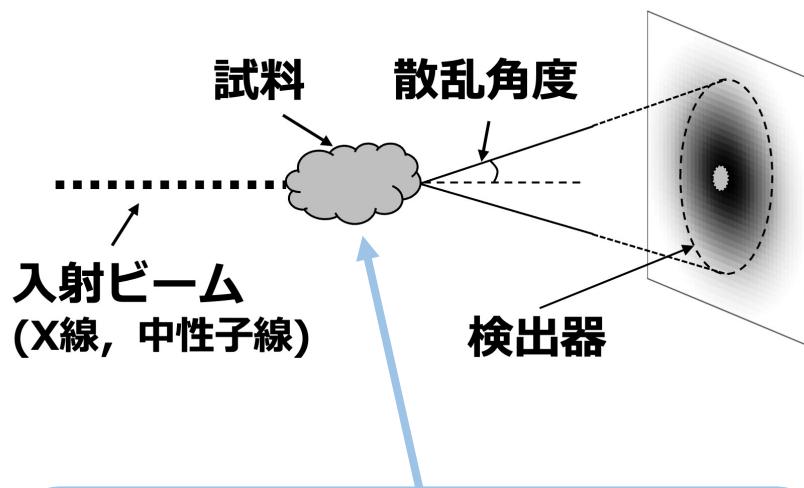


目次

1. 自己紹介
2. ベイズ計測
 - ベイズ推論とベイズ計測
 - 物理パラメータの事後確率推定, モデル選択, 複数データの統合
3. 解析的に解けるベイズ計測 : $y = ax + b$
 - 物理パラメータの事後確率推定, モデル選択
4. 小角散乱のベイズ計測
 - 小角散乱の概説
 - 物理パラメータの事後確率推定 : 試料パラメータの推定
 - モデル選択 : 散乱強度モデルの選択
5. まとめ
 - ベイズ計測を用いた新たな物性科学のパラダイムの提案
 - ベイズ計測の社会実装に必須な条件

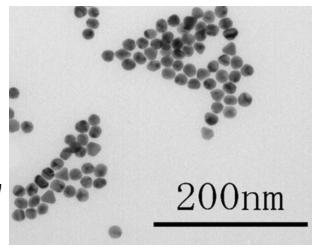
小角散乱法(SAS)

試料にビームを照射し、小さい角度の散乱強度からナノ構造解析を行う。



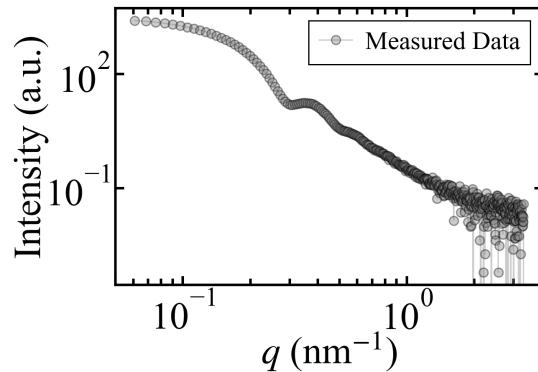
不均一系試料

- 高分子溶液
- タンパク質
- コロイド粒子
- 金属, セラミック
- 繊維 …



出典: [1]

散乱強度-散乱ベクトル

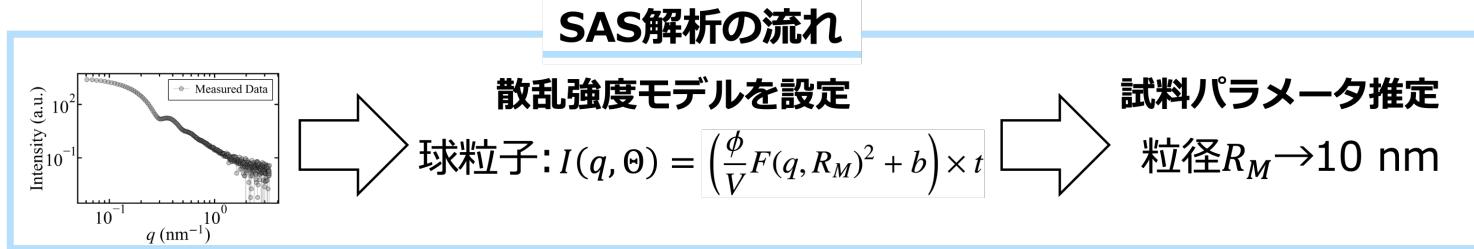


データ解析

1. 散乱強度モデルを設定
球粒子: $I(q, \Theta) = \left(\frac{\phi}{V} F(q, R_M)^2 + b \right) \times t$
2. モデルパラメータ推定
粒径 $R_M \rightarrow 10 \text{ nm}$

[1] "X線小角散乱法 (SAXS) による液体中に分散したナノ粒子の構造解析", NISSAN ARC,
URL: <https://www.nissan-arc.co.jp/services/p088/>

SAS解析の従来法とその課題



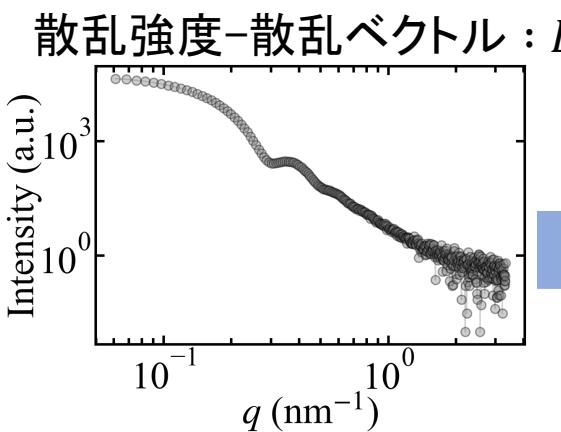
- 試料パラメータ推定に勾配法などを用いているために、局所解が得られることが多い。
→何度もパラメータの初期値を変えて対処している。
- 解析に用いる散乱強度モデルが予め分からないと解析できない。
→フィッティングの二乗誤差や経験則を基にモデル選択している。

本研究の目的

2つの従来法の課題をベイズ計測を応用し解決する。

小角散乱法におけるベイズ計測 物理パラメータ推定とモデル選択

散乱強度、計測ノイズモデルの選択や試料パラメータ推定の解析に適用出来る。



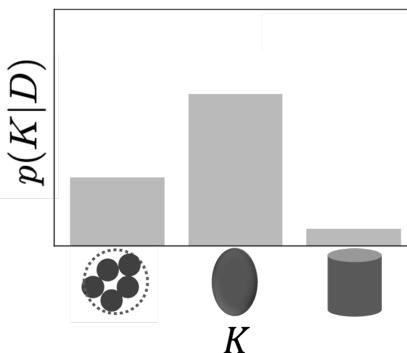
目的

データから試料構造を解析する。

モデル選択

$$p(K|\mathcal{D}) = \frac{\int p(\mathcal{D}, \Theta, K) d\Theta}{\sum_K \int p(\mathcal{D}, \Theta, K) d\Theta}$$

試料構造を表す
散乱強度モデルの選択

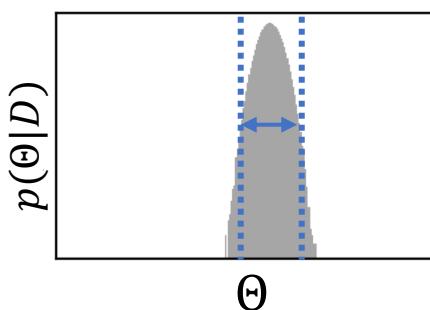


定量的な散乱強度
モデルの選択

パラメータ推定

$$p(\Theta|\mathcal{D}) = \frac{p(\mathcal{D}|\Theta)p(\Theta)}{\int p(\mathcal{D}, \Theta) d\Theta}$$

試料の大きさや密度に
関するパラメータを推定



推定結果から信頼度の
定量評価が可能

小角散乱法のベイズ計測 – 定式化

データ点 y_i の確率分布 : $p(y|q, \Theta, K) = \frac{I_K(q, \Theta)^y \exp(-I_K(q, \Theta))}{y!}$

(\because 散乱強度は、光子のカウントデータとして計測)

データ \mathcal{D} の生成確率 : $p(\mathcal{D}|\Theta, K) = \prod_{i=1}^N p(y_i|q_i, \Theta, K) = \exp(-NE(\Theta, K))$

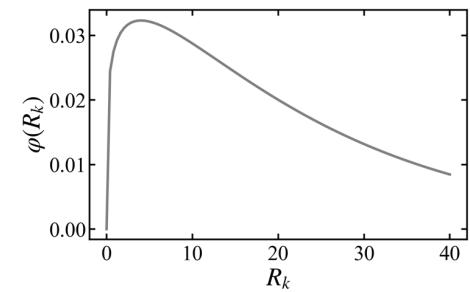
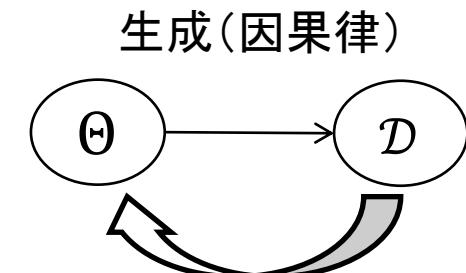
→ 誤差関数 : $E(\Theta, K) = \frac{1}{N} \sum_{i=1}^N \left\{ I_K(q_i, \Theta) - y_i \log I_K(q_i, \Theta) + \sum_{j=1}^{y_i} \log j \right\}$

パラメータ θ の事後確率 : $p(\Theta|\mathcal{D}) = \frac{p(\mathcal{D}|\Theta)\varphi(\Theta)}{\int p(\mathcal{D}, \Theta)d\Theta} \propto \exp(-NE(\Theta))\varphi(\Theta)$

モデル K の事後確率 : $p(K|\mathcal{D}) = \frac{\int p(\mathcal{D}, \Theta, K)d\Theta}{\sum_K \int p(\mathcal{D}, \Theta, K)d\Theta} = \frac{\exp(-F(K))\varphi(K)}{\sum_K \exp(-F(K))\varphi(K)}$

事前分布 $\varphi(\Theta)$: 事前知識を基に設定

散乱強度の計測値 : y_i
散乱ベクトル : q_i
試料モデル : $I(q_i; \Theta)$



目次

1. 自己紹介
2. ベイズ計測
 - ベイズ推論とベイズ計測
 - 物理パラメータの事後確率推定, モデル選択, 複数データの統合
3. 解析的に解けるベイズ計測 : $y = ax + b$
 - 物理パラメータの事後確率推定, モデル選択
4. 小角散乱のベイズ計測
 - 小角散乱の概説
 - 物理パラメータの事後確率推定 : 試料パラメータの推定
 - モデル選択 : 散乱強度モデルの選択
5. まとめ
 - ベイズ計測を用いた新たな物性科学のパラダイムの提案

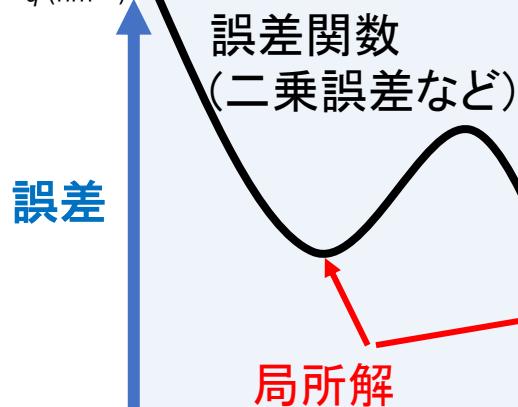
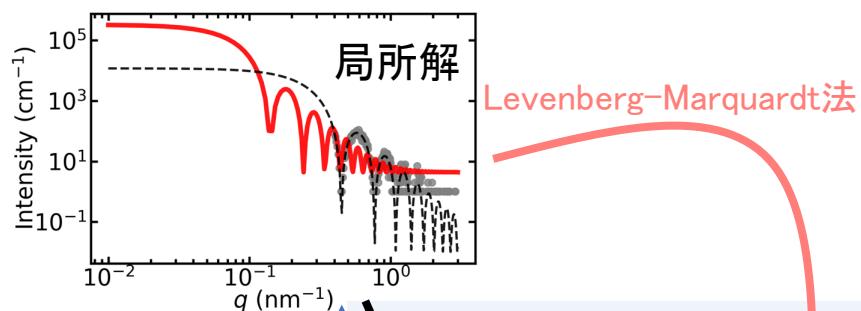
試料パラメータ推定

従来法

勾配法や準ニュートン法による
パラメータフィッティング.

課題

- 局所解にトラップされる
- 結果の信頼度が分からぬ



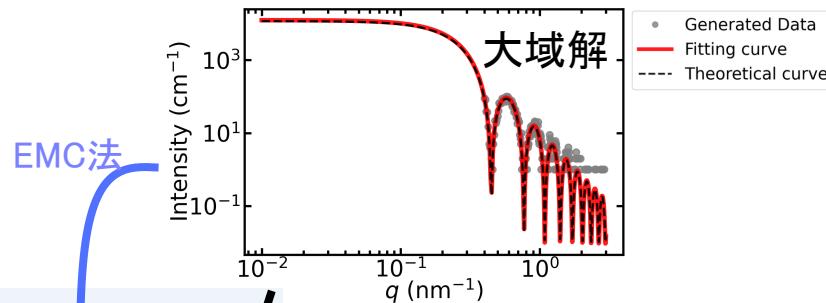
ベイズ計測

交換モンテカルロ(EMC)法[2]により事後分布をサンプリング.

課題の解決

- 大域的最適解の推定
- 事後分布から信頼度評価

Yui Hayashi et al., J. Phys. Soc. Jpn. 92, 094002 (2023).



交換モンテカルロ法

メトロポリス法

パラメータの事後確率 :

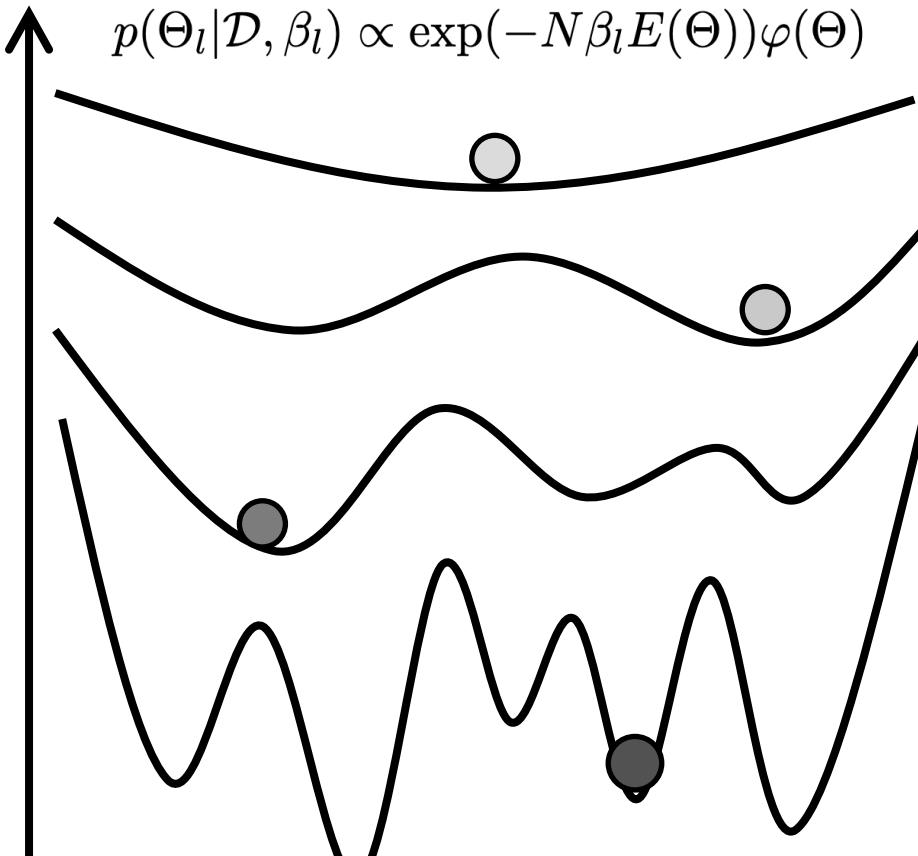
$$p(\Theta|\mathcal{D}) \propto \exp(-NE(\Theta))\varphi(\Theta)$$



交換モンテカルロ法

高温($\beta = 0$)

$$p(\Theta_l|\mathcal{D}, \beta_l) \propto \exp(-N\beta_l E(\Theta))\varphi(\Theta)$$



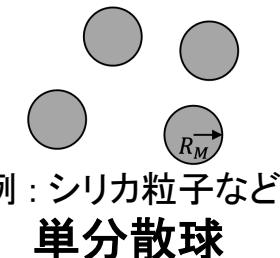
低温($\beta = 1$)

パラメータ推定の数値実験 - 設定

単分散球試料のパラメータを推定する。

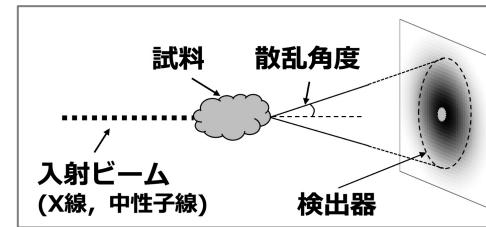
$$\text{散乱強度モデル} : I_M(q; \Theta) = \left[\frac{\phi}{V} \left\{ \frac{3\Delta\rho V [\sin(qR_M) - qR_M \cos(qR_M)]}{(qR_M)^3} \right\}^2 + b \right] \times t$$

パラメータ : $\Theta = \{R_M, b, t\}$ (粒径: R_M , バックグラウンド: b , 計測時間: t)



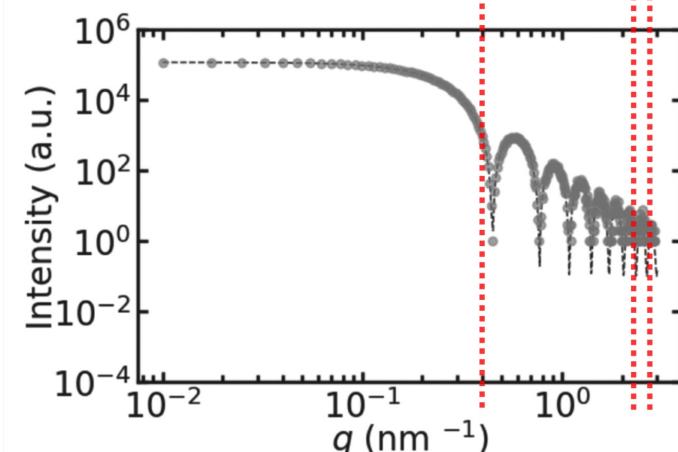
SASデータの特徴

実験機器の構造上, 低角領域のデータ欠損がしばしば起こる。



欠損度合いの異なる3種類の人工データに対して実験を行う

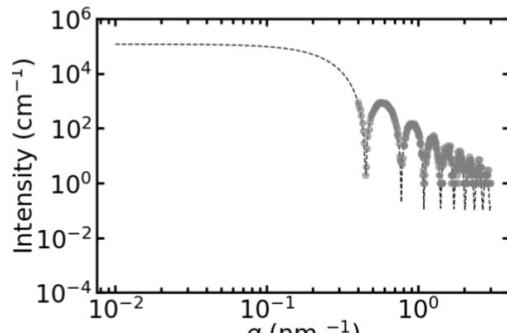
(a) (b) (c)



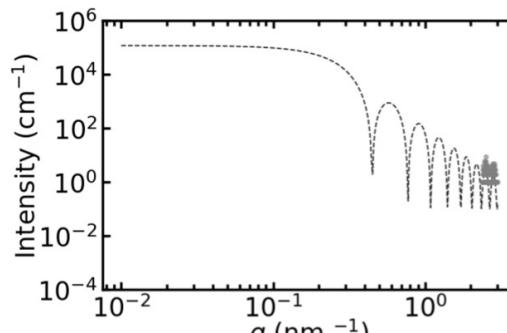
パラメータ推定の数値実験 - 結果

q_{\min} : 散乱ベクトル q の下限

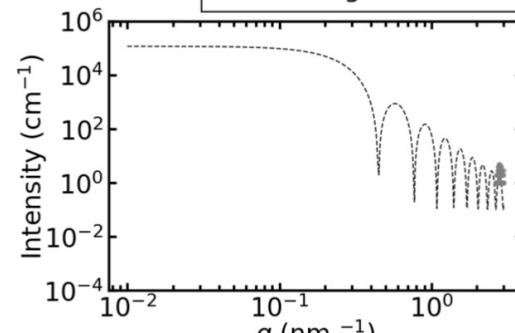
人工データ



(a) $q_{\min} = 0.4 \text{ nm}^{-1}$

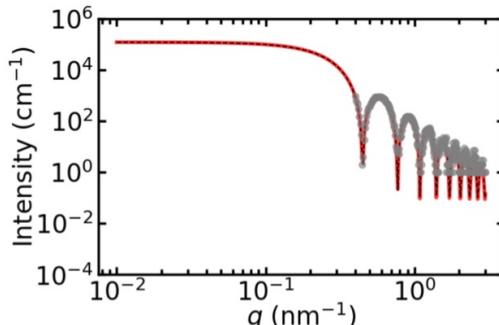


(b) $q_{\min} = 2.35 \text{ nm}^{-1}$

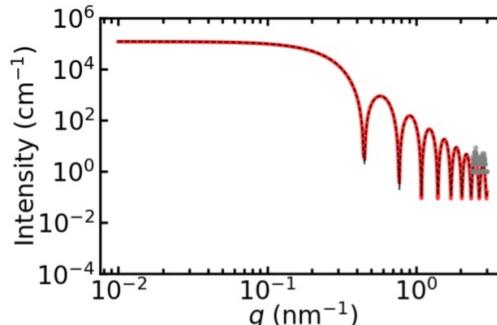


(c) $q_{\min} = 2.65 \text{ nm}^{-1}$

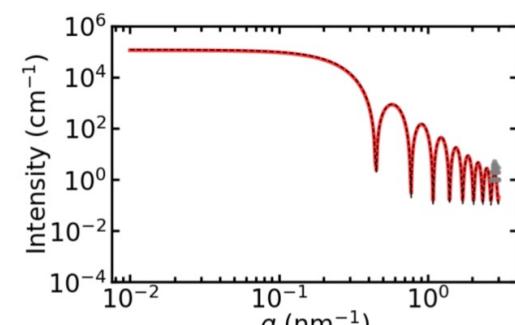
フィッティング



(a) $q_{\min} = 0.4 \text{ nm}^{-1}$



(b) $q_{\min} = 2.35 \text{ nm}^{-1}$



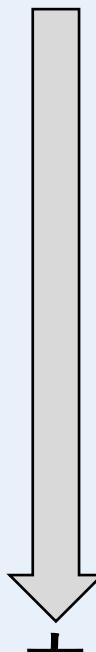
(c) $q_{\min} = 2.65 \text{ nm}^{-1}$

パラメータの事後分布

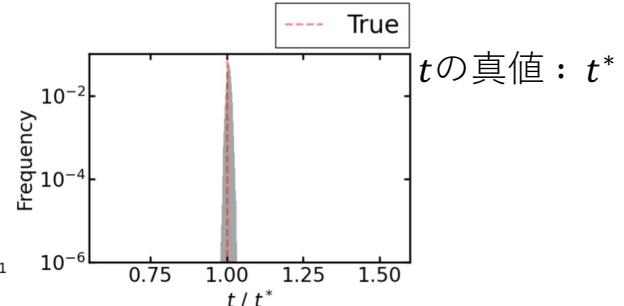
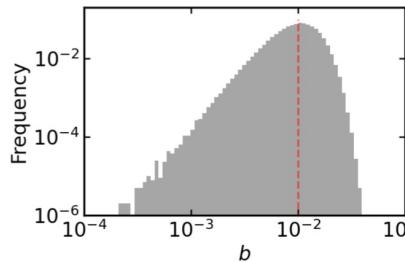
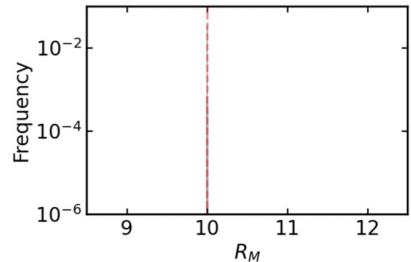
パラメータの事後分布

データ欠損

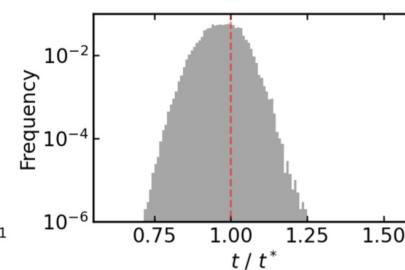
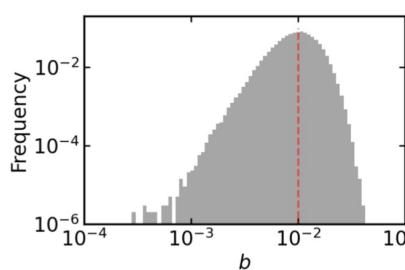
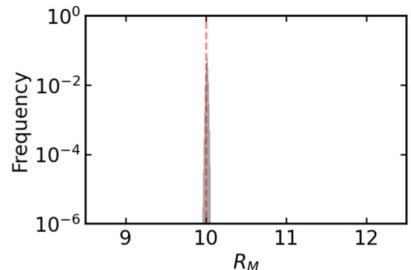
小



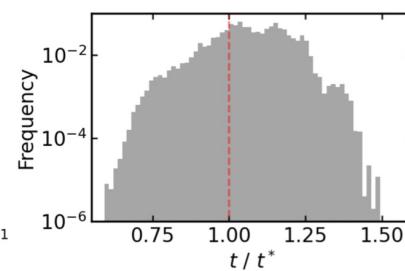
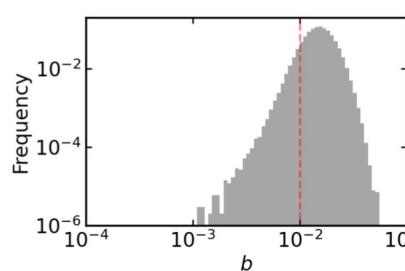
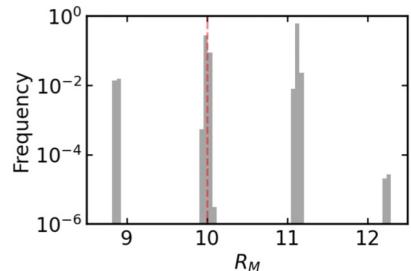
(a) $q_{\min} = 0.4$



(b) $q_{\min} = 2.35$



(c) $q_{\min} = 2.65$

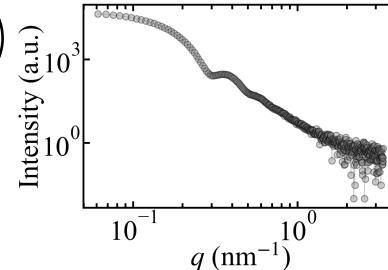
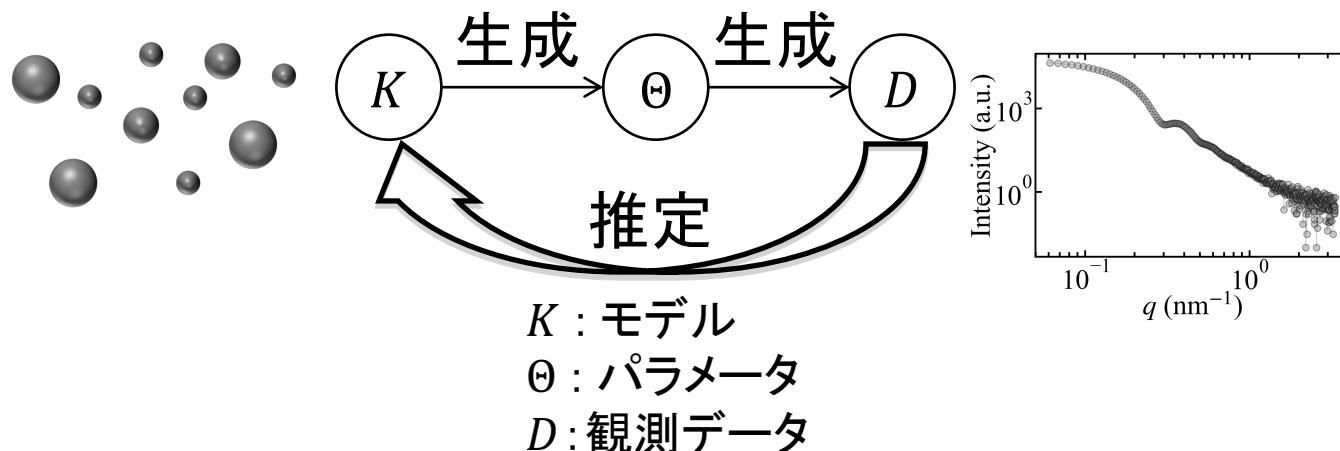


(c)のデータに対する R_M の事後分布に複数のピークが見られる。
→ $q_{\min} = 2.35 - 2.65 \text{ nm}^{-1}$ で推定限界

目次

1. 自己紹介
2. ベイズ計測
 - ベイズ推論とベイズ計測
 - 物理パラメータの事後確率推定, モデル選択, 複数データの統合
3. 解析的に解けるベイズ計測 : $y = ax + b$
 - 物理パラメータの事後確率推定, モデル選択
4. 小角散乱のベイズ計測
 - 小角散乱の概説
 - 物理パラメータの事後確率推定 : 試料パラメータの推定
 - モデル選択 : 散乱強度モデルの選択
5. まとめ
 - ベイズ計測を用いた新たな物性科学のパラダイムの提案
 - ベイズ計測の社会実装に必須な条件

数理モデルの事後確率の計算



$$\begin{aligned}\text{モデル } K \text{ の事後確率} : p(K|\mathcal{D}) &= \frac{\int p(\mathcal{D}, \Theta, K) d\Theta}{\sum_K \int p(\mathcal{D}, \Theta, K) d\Theta} \\ &= \frac{\exp(-F(K))\varphi(K)}{\sum_K \exp(-F(K))\varphi(K)}\end{aligned}$$

ベイズ自由エネルギー : $F(K) = -\log Z(K)$

$$Z(K) = \int p(\mathcal{D}, \Theta|K) d\Theta$$

$F(K)$ を交換モンテカルロ法で数値的に計算する。

自由エネルギーの数値的計算法 レプリカ交換モンテカルロ法を使う

たくさんの温度でのパラメータの事後確率 $p(\Theta|D, K, \beta)$ のサンプルを用いて計算する($\because 0 = \beta_1 < \dots < \beta_L = 1$)：

$$\begin{aligned} Z(K, \beta = 1) &= \prod_{l=1}^{L-1} \frac{Z(K, \beta_{l+1})}{Z(K, \beta_l)} \\ &= \prod_{l=1}^{L-1} \langle \exp(-N(\beta_{l+1} - \beta_l)E(\Theta, K)) \rangle_{p(\Theta_l | \mathcal{D}, K, \beta_l)} \end{aligned}$$

→ $F(K)$, $p(K|D)$ が計算できる.

ベイズ自由エネルギー : $F(K) = -\log Z(K)$.

$$\begin{aligned} \text{モデル } K \text{ の事後分布} : p(K|\mathcal{D}) &= \frac{\int p(\mathcal{D}, \Theta, K) d\Theta}{\sum_K \int p(\mathcal{D}, \Theta, K) d\Theta} \\ &= \frac{\exp(-F(K))\varphi(K)}{\sum_K \exp(-F(K))\varphi(K)} \end{aligned}$$

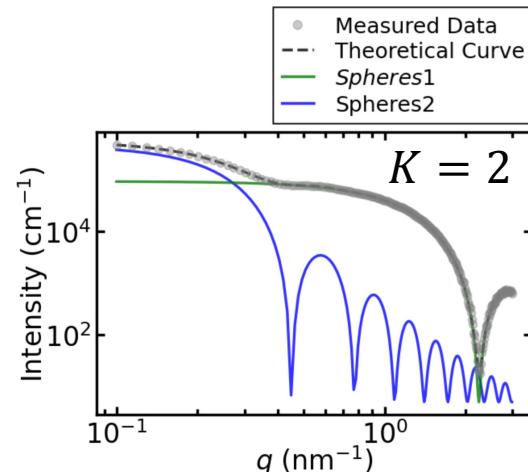
散乱強度モデル選択の数値実験 – 設定

2成分の单分散球が混合しているデータから、成分数 K を選択する。

K 成分の单分散球モデル

$$I_K(q, \Theta) = \sum_{k=1}^K S_k V_k \left(\frac{(\sin(qR_k) - qR_k \cos(qR_k))}{(qR_k)^3} \right)^2 + B$$

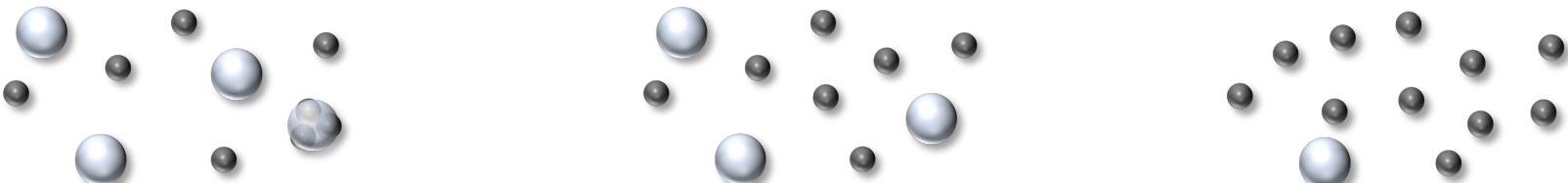
$$\Theta = \left\{ \{R_k, S_k\}_{k=1}^K, B \right\}$$



混合比 r_s を変化させて6種のデータを生成した。

$$r_s = \frac{S_2}{S_1}$$

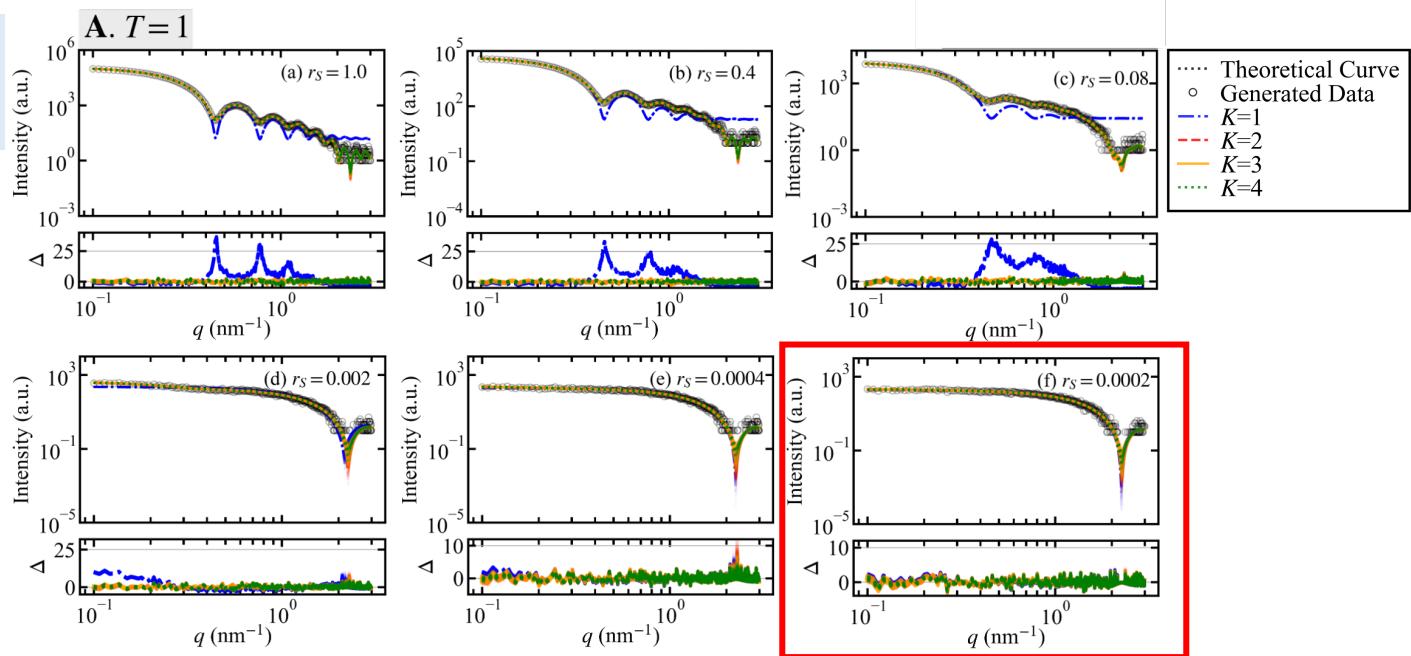
r_s 大  r_s 小



→ $K = 1, 2, 3, 4$ に対して事後確率 $p(K|D)$ を計算し比較する。

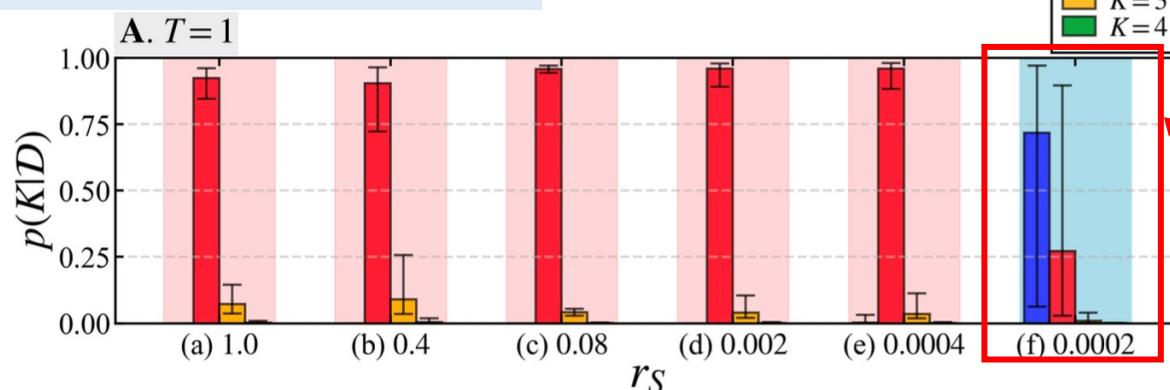
モデル選択の数値実験 - 結果(1/2)

データと フィッティング



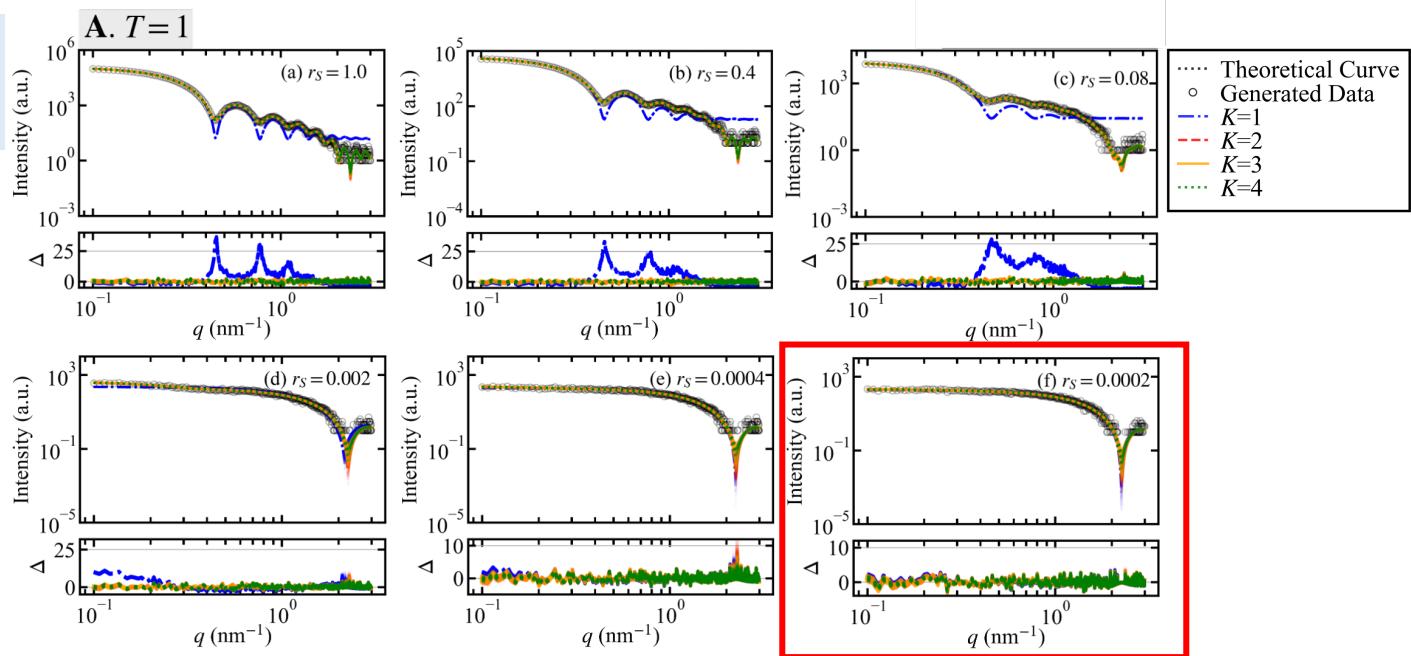
データ生成時の乱数シードを変えて10回の実験を行った結果. (真値: $K = 2$)

解析限界



モデル選択の数値実験 - 結果(2/2)

データと フィッティング



データ生成の乱数シードを変化させた場合

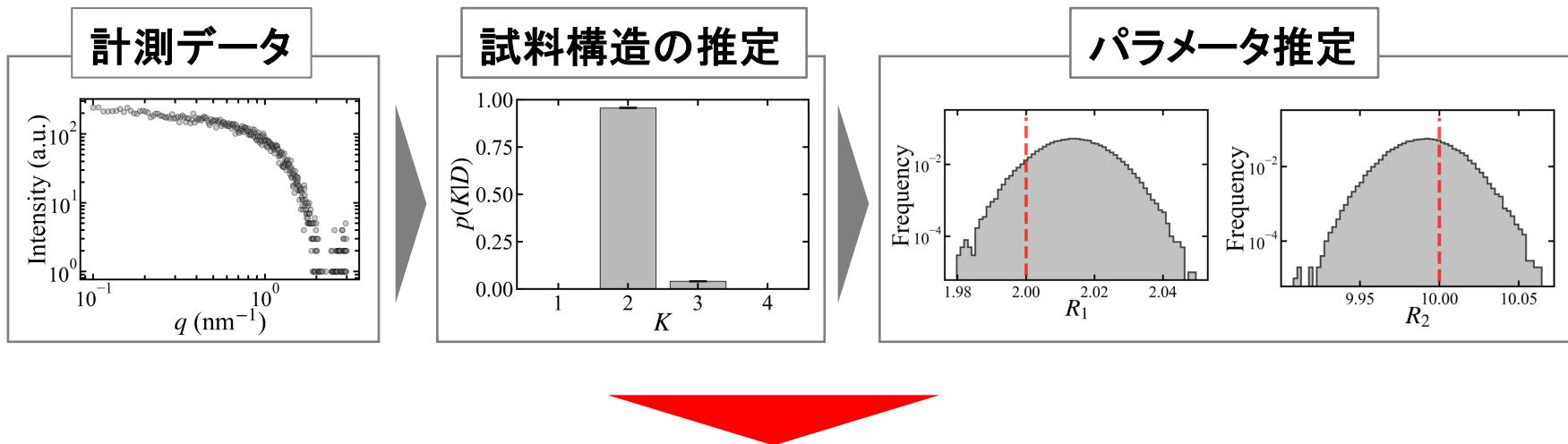
10データ中で、 $K = 1 \sim 4$ の
モデルが支持された回数

$r_s \backslash K$	1	2	3	4
(a) 1.0	0	10	0	0
(b) 0.4	0	10	0	0
(c) 0.08	0	10	0	0
(d) 0.002	0	10	0	0
(e) 0.0004	0	10	0	0
(f) 0.0002	8	2	0	0

解析限界

小角散乱のベイズ計測のまとめ

- パラメータ推定に勾配法などを用いているために、局所解が得られることが多い。
→交換モンテカルロ法を用いて事後確率分布として推定する.
→大域的最適解の推定、推定の信頼度評価ができる.
- 解析に用いる散乱強度モデルが予め分からないと解析できない.
→データドリブンで定量的な散乱強度モデル選択を可能にした.
→データから試料構造を数理的に選択できる.



今後、様々な試料のSAS実データへの適用が望まれる.

目次

1. 自己紹介
2. ベイズ計測
 - ベイズ推論とベイズ計測
 - 物理パラメータの事後確率推定, モデル選択, 複数データの統合
3. 解析的に解けるベイズ計測 : $y = ax + b$
 - 物理パラメータの事後確率推定, モデル選択
4. 小角散乱のベイズ計測
 - 小角散乱の概説
 - 物理パラメータの事後確率推定 : 試料パラメータの推定
 - モデル選択 : 散乱強度モデルの選択
5. まとめ
 - ベイズ計測を用いた新たな物性科学のパラダイムの提案
 - ベイズ計測の社会実装に必須な条件

ベイズ計測を用いた 新たな物性科学のパラダイムの提案

- ・ベイズ計測により物性科学は新たな局面を迎えると考えている.
- ・従来は、個々の分野でポスト処理としてパラメータフィッティングを行っていた.
- ・ベイズ計測の導入により物性科学全般を取り扱うことが出来る普遍的なデータ解析の数理基盤が構築できる.
- ・ベイズ計測では様々なデータを普遍的な枠組みで取り扱うことができ、データ解析技術のフラット化が実現できる.
- ・ベイズ計測によりデータ解析は実験のポスト処理ではなく、実験計画の立案に大きく寄与する.

よってベイズ計測を用いた、新たな物性科学のパラダイムが提案できると考えている。

ベイズ計測の物性科学における適用例

東京大学 岡田研究室

- 事後分布推定、モデル選択、ベイズ統合
 - 1. $y=ax+b$
 - 2. スペクトル分解
 - 3. X線光電子放出スペクトル(XPS)
 - 4. X線吸収スペクトル(XAS)
 - 5. メスバウア一分光
 - 6. X線小角散乱スペクトル
 - 7. NMR
 - 8. 中性子非弾性散乱スペクトル
 - 9. 比熱
 - 10. 帯磁率
 - 11. 比熱と帯磁率

NIMS

- 事後分布推定 and/or モデル選択
 - 1. XRD
 - 2. X線光電子放出スペクトル(XPS)

熊本大学 赤井研究室

- 事後分布推定 and/or モデル選択
 - 1. 可視光発光スペクトル
 - 2. 可視光吸収スペクトル
 - 3. EXAFS
 - 4. XANES
 - 5. XMCD
 - 6. 変調スペクトル
 - 7. XPS
 - 8. XRD
 - 9. コヒーレントフォノン

SPring-8/JASRI 熊本大学 水牧研究室

- 事後分布推定、モデル選択、ベイズ統合
 - 1. XRD
 - 2. XPSとXAS
 - 3. 磁気コンプトン散乱プロファイル
 - 4. ラマンスペクトル

目次

1. 自己紹介
2. ベイズ計測
 - ベイズ推論とベイズ計測
 - 物理パラメータの事後確率推定, モデル選択, 複数データの統合
3. 解析的に解けるベイズ計測 : $y = ax + b$
 - 物理パラメータの事後確率推定, モデル選択
4. 小角散乱のベイズ計測
 - 小角散乱の概説
 - 物理パラメータの事後確率推定 : 試料パラメータの推定
 - モデル選択 : 散乱強度モデルの選択
5. まとめ
 - ベイズ計測を用いた新たな物性科学のパラダイムの提案
 - ベイズ計測の社会実装に必須な条件

ベイズ計測の社会実装に必須な条件

- 大学、国研などのアカデミアではベイズ計測のアルゴリズム開発や新たな実験データ解析への適用が主である。
- SPring-8/JASRIなどの大規模計測施設のサービス部門は計測に関するサービスで手一杯の状態である。
- とてもベイズ計測による新たなデータ解析手法を民間企業等にサービスするための人的資源はない。
- つまり、アカデミアと民間企業の間にベイズ計測の技術移転の空洞化が生じるはずである。この空洞を埋めることが、新たなビジネスチャンスにつながると考えている。
- 日本の物質材料産業に強く依存した構造を考えると、このような形態のビジネスは単に民間企業の利益の追求だけではなく、日本の国力を下支えする重要な役割を果たすと考えている。

前述の思いの基、以下の起業をした 株式会社 a.s.ist

データの本質を解き明かし、確かな洞察で社会に変革をもたらす。

会社名	株式会社 a.s.ist
所在地	東京都足立区伊興 5-17-24
代表取締役	森口棕太
設立	2021年 6月
事業	<ul style="list-style-type: none">・ アルゴリズム提供事業・ 受託開発
Webサイト	https://www.a-s-ist.com/

- メンバー：岡田研究室の博士課程の学生
- 技術顧問：東京大学 岡田真人 教授
- ご連絡先：sales@a-s-ist.com

強み

- ・ ベイズ計測
- ・ スパースモデリング
- ・ 深層学習

事業内容

- ・ 受託データ解析
- ・ ライブライ開発
- ・ Web API開発

ベイズ計測の二日間の集中講義

新領域複雑理工 岡田真人

- 新領域基盤系のデータ駆動科学教育プログラム
- 47130-51 データ駆動科学入門 I
 - 目的と内容
 - 以下の二つでベイズ計測の俯瞰的理解
 - $y=ax+b$ のベイズ計測を手計算で解析計算
 - スペクトル分解の座学
 - 講義日程
 - 8/7(水) 2限、3限、4限、5限
 - 8/9(金) 2限、3限、4限
 - ZoomURL
 - <https://u-tokyo-ac-jp.zoom.us/j/88625776290?pwd=bQmJ4x4a1a8UDdj42YPAvg6shlmIZ0.1>
 - 東大関係者であれば聴講可能
 - ぜひ、ご聴講ください。

1. 自己紹介
2. ベイズ計測
 - ベイズ推論とベイズ計測
 - 物理パラメータの事後確率推定, モデル選択, 複数データの統合
3. 解析的に解けるベイズ計測 : $y = ax + b$
 - 物理パラメータの事後確率推定, モデル選択
4. 小角散乱のベイズ計測
 - 小角散乱の概説
 - 物理パラメータの事後確率推定 : 試料パラメータの推定
 - モデル選択 : 散乱強度モデルの選択
5. まとめ
 - ベイズ計測を用いた新たな物性科学のパラダイムの提案
 - ベイズ計測の社会実装に必須な条件

ポアソンノイズ

計測時間を T として、仮想的に n 分割した微小時間 Δt を導入し、 $T = n\Delta t$ とする。

単位時間あたりのカウント数を m とする。

ここで $n \rightarrow \infty$ を考えると、微小時間 Δt には光子が1カウントされるか、されないかのどちらかとなる。

微小時間に光子がカウントされる確率は、 $p = m\Delta t = \frac{mT}{n}$ となる。

有限時間 T の間に j 回カウントされる確率 $P(j)$ は独立な n 回の試行における j 回の実現確率となり、次式を得る。

$$P_n(j) = \frac{n!}{(n-j)! j!} \left(1 - \frac{mT}{n}\right)^{n-j} \left(\frac{mT}{n}\right)^j.$$

$n \rightarrow \infty$ を考えて、 $n \gg j$ であることを用いると、次式が得られる。

$$\begin{aligned} P(j) &= \lim_{n \rightarrow \infty} \frac{n!}{(n-j)! j!} e^{-mT} \left(1 - \frac{mT}{n}\right)^{-j} \frac{(mT)^j}{n^j} \\ &= \frac{(mT)^j}{j!} e^{-mT} \lim_{n \rightarrow \infty} \frac{n(n-1)\cdots(n-j+1)}{n^j} \\ &= \frac{(mT)^j}{j!} e^{-mT}. \end{aligned}$$

事後分布のサンプリング

交換モンテカルロ法(EMC)のアルゴリズム

1. パラメータ Θ を初期化する.
2. 逆温度 $0 = \beta_1 < \dots < \beta_l < \dots < \beta_L = 1$ を用意する.
3. L 個の各逆温度で, MCMCにより事後分布のサンプリングを行う.

$$\beta_l \text{での事後分布} : p(\Theta_l | \mathcal{D}, \beta_l) \propto \exp(-N\beta_l E(\Theta_l)) p(\Theta_l)$$

4. 任意の時点で隣り合う逆温度での状態を確率 W で交換する.

$$W(\Theta_l | \Theta_{l+1}) = \min \left[1, \frac{p(\Theta_{l+1} | \mathcal{D}, \beta_l) p(\Theta_l | \mathcal{D}, \beta_{l+1})}{p(\Theta_l | \mathcal{D}, \beta_l) p(\Theta_{l+1} | \mathcal{D}, \beta_{l+1})} \right]$$

5. 3~4の手順を繰り返し, 目的の分布 $p(\Theta_L | \mathcal{D})$ を得る.