

これからの仮説検証・モデル評価

ベイズ塾 北條大樹

東京大学 D1・日本学術振興会(DC1)





名前：北條大樹

所属：東京大学 教育学研究科 教育心理学コース D1

研究室：心理統計学研究室（指導教員：岡田 謙介）

研究関心：ベイズ統計学・心理統計学・数理心理学

HP：<https://dastatis.github.io/index.html>

OSF：osf.io/7d9zw

Github：<https://github.com/dastatis>

Twitter：@dastatis

- **心理学者におけるベイズ統計について**

- 事後分布が出てくるまで
 - 事前分布について
 - どうやってデータ生成メカニズムを考えるのか？

- **事後分布の取り扱い方**

- 事後分布の点要約
- 事後分布の区間要約

- **これからの仮説検証・モデル評価**

- 我々はどちらに興味があるのか？
- 周辺尤度 + α
- 事後予測チェックについて
- 情報量規準について

第1部 心理学者における ベイズ統計について

～ 事後分布がでてくるまで ～

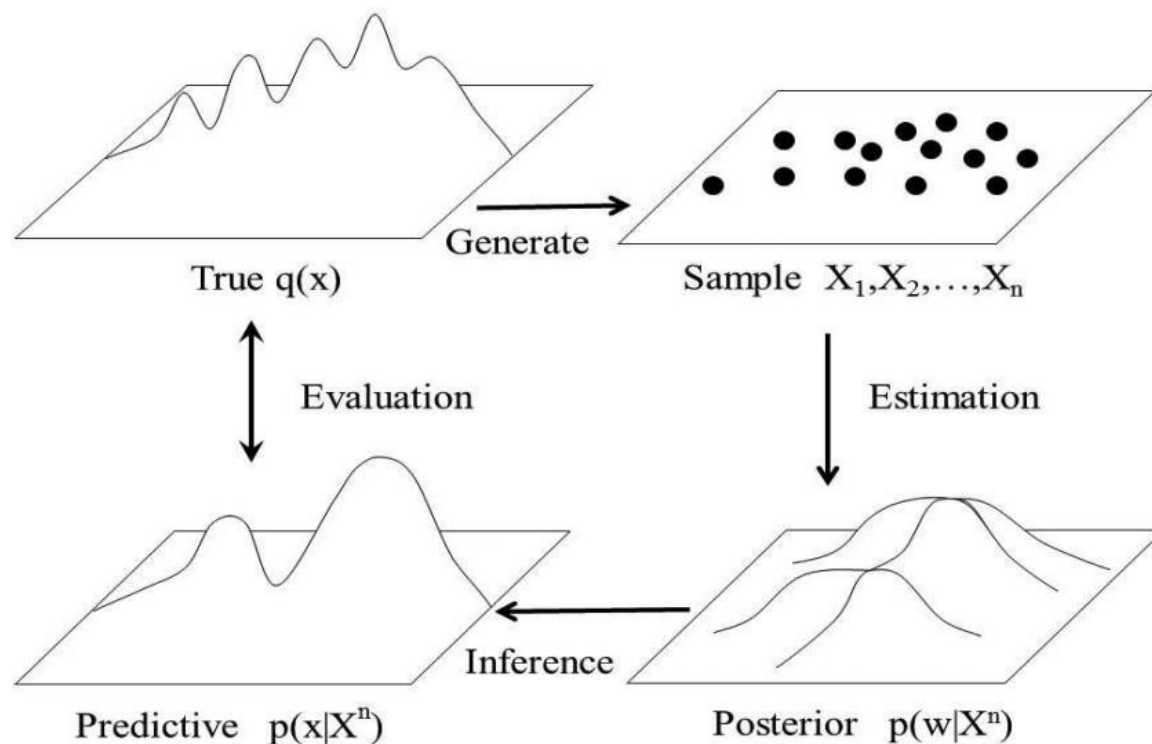


Figure 1.1: Framework of Bayesian inference. The procedure of Bayesian estimation is shown. A sample X^n is taken from unknown true distribution $q(x)$. A statistician sets a statistical model and a prior, then the posterior density $p(w|X^n)$ is obtained. The true distribution $q(x)$ is estimated by a predictive density $p(x|X^n)$, whose accuracy is evaluated by using mathematical laws.

心理学者の統計分析

• 調査データ・・・項目回答データ

ID	Q1	Q2	Q3
1	5	2	3
2	2	1	1
3	3	2	4



- カットオフを超えた？
- 平均得点は？
- 2群で差がある？

等々

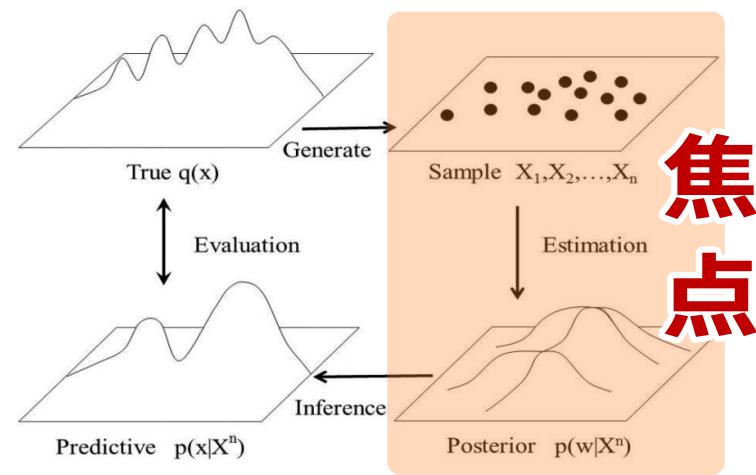
• 実験データ・・・反応時間(RT)データ

ID	条件	RT
1	1	3.25
1	2	4.17
2	1	3.36
2	2	4.01

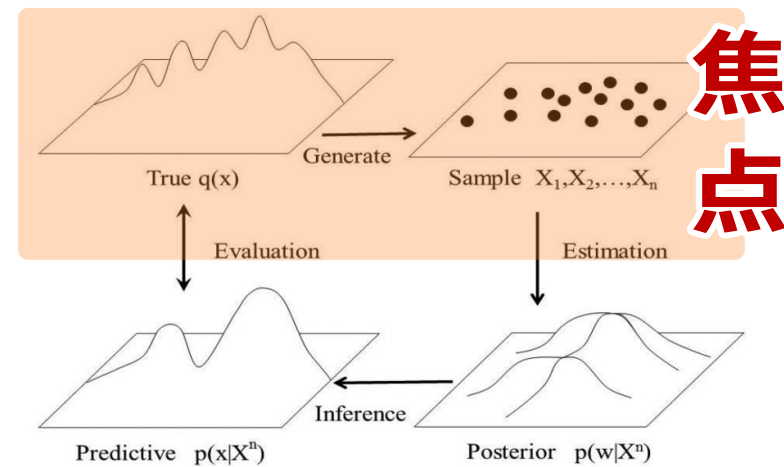


- 平均RTは？
- 2条件で差がある？

等々



これからの心理学



• 調査データ…項目回答データ

ID	Q1	Q2	Q3
1	5	2	3
2	2	1	1
3	3	2	4

- カットオフを超えた？
- 平均得点は？
- 2群で差がある？

等々



• 実験データ…反応時間(RT)データ

ID	条件	RT
1	1	3.25
1	2	4.17
2	1	3.36
2	2	4.01

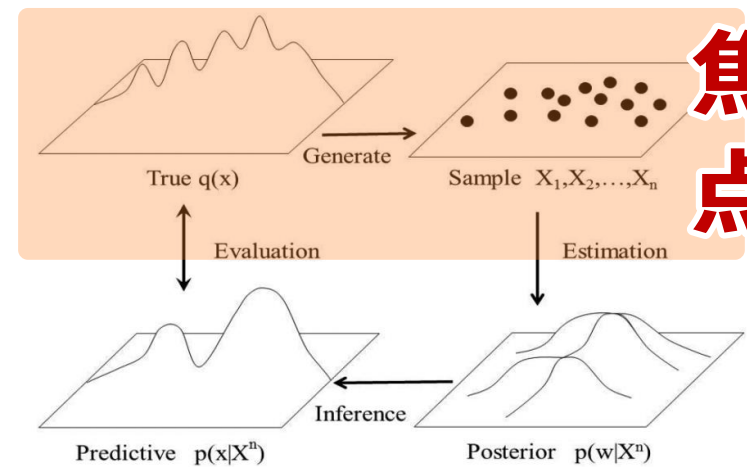
- 平均RTは？
- 2条件で差がある？

等々

これからの心理学

焦点

・ データ生成過程を考える



・ 調査データ・・・項目回答データ

ID	Q1	Q2	Q3
1	5	2	3
2	2	1	1
3	3	2	4

- ・ カットオフを超えた？
- ・ 平均得点は？
- ・ 2群で差がある？

等々



・ 実験データ・・・反応時間(RT)データ

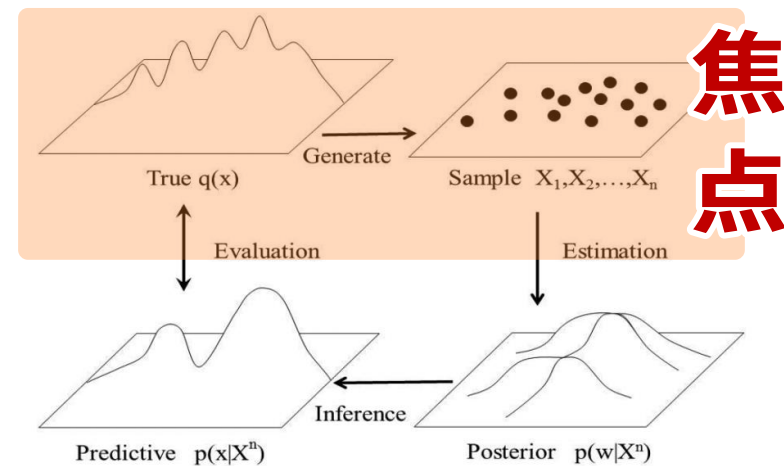
ID	条件	RT
1	1	3.25
1	2	4.17
2	1	3.36
2	2	4.01

- ・ 平均RTは？
- ・ 2条件で差がある？

等々

これからの心理学

・モデルを考える



・調査データ…項目回答データ

ID	Q1	Q2	Q3
1	5	2	3
2	2	1	1
3	3	2	4

- ・カットオフを超えた？
- ・平均得点は？
- ・2群で差がある？

等々

・実験データ…反応時間(RT)データ

ID	条件	RT
1	1	3.25
1	2	4.17
2	1	3.36
2	2	4.01

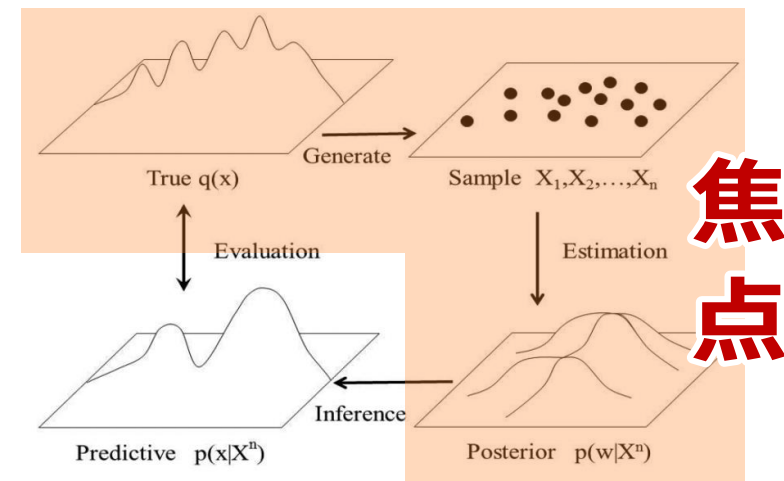
- ・平均RTは？
- ・2条件で差がある？

等々

モデルとは、尤度関数と事前分布を考えることに相当

これからの心理学

・そして、パラメータ推定へ



・調査データ…項目回答データ

ID	Q1	Q2	Q3
1	5	2	3
2	2	1	1
3	3	2	4

- ・カットオフを超えた？
- ・平均得点は？
- ・2群で差がある？

等々



・実験データ…反応時間(RT)データ

ID	条件	RT
1	1	3.25
1	2	4.17
2	1	3.36
2	2	4.01

- ・平均RTは？
- ・2条件で差がある？

等々

データ生成を表現したベイズモデル

- いままでの心理学は、**得られたデータを統計分析**
 - これからの心理学は、**データ生成メカニズムを考えたモデル**
- もちろん、ベイズ統計は、どちらにも対応できる

• 調査データ・・・項目回答データ



ID	Q1	Q2	Q3
1	5	2	3
2	2	1	1
3	3	2	4



- カットオフを超えた？
- 平均得点は？
- 2群で差がある？

等々

• 実験データ・・・反応時間(RT)データ



ID	条件	RT
1	1	3.25
1	2	4.17
2	1	3.36
2	2	4.01



- 平均RTは？
- 2条件で差がある？

等々

知らぬ間に仮定

t検定モデル・分散分析モデル

仮定を設定する必要

拡散モデル・項目反応モデル

- **パラメータ推定**の文脈で、我々の興味のあるパラメータ θ の不確かさを観測データ y から求めるとき

$$\begin{array}{ccc} \text{事後分布} & \text{尤度} & \text{事前分布} \\ p(\theta|y) = & \frac{p(y|\theta)p(\theta)}{p(y)} & \\ & \text{周辺尤度} & \end{array}$$

と表す

- しかし、実際はMCMC等で事後分布を導出するので

$$\begin{array}{ccc} \text{事後分布} & \text{尤度} & \text{事前分布} \end{array}$$

$$p(\theta|y) \propto p(y|\theta)p(\theta) \quad * \propto \dots \text{は比例を表す}$$

必要なのは、**尤度と事前分布のみ**

Next 事前分布をどうやって選ぶ? ➡

- ・事前分布の選択方法は、2種類ある(私論)

ホワイトリストベイジアン

特徴: モデルに**必要な**仮定を、事前分布に取り組む

イメージ: 良いと思う仮定を加えていくスタイル

相性: ○ **データ生成メカニズムを考えたモデル**

➡ 分析者がデータ生成メカニズムをモデルに柔軟に表現できるため

≠ 自由な事前分布

ブラックリストベイジアン

特徴: モデルに**不適切な**仮定を、事前分布から除く

イメージ: 不適切な仮定を除いていくスタイル

相性: ○ **得られたデータを統計分析**

➡ 強く制約された仮定のもとで、モデルの一部を検討するため

≠ 無情報事前分布

・ホワイトリストベイジアンの考え方

- ・得られたデータの生成メカニズムを考えるならば、 $\circ\circ$ という仮定を置くことにしよう

➡ 論文ではこれを客観的に説明する。

ex) 成人男性の身長の前分布

× $0 - \infty$ (限りなく広い正) の範囲の一般分布

○ 1.65m (成人男性の平均身長) を平均とする正規分布

- ・ **データ生成メカニズムを考えたモデル**のパラメータ推定がうまくいかないため、推定可能となるような事前分布を設定する

➡ 論文では、モデルの仮定の一つとして上記の事前分布を置いたことを説明し、モデルの挙動を正確に記述する。また、分析を再現可能な形として論文とともに提出し、査読を受ける。論文化される場合、分析コードを公開する。**現段階でモデルの最善を尽くした結果、このような仮定（事前分布）のモデルになった**

➡ これによって、別の研究者がそのコードに基づいて再現、改良を加えることができる、もしくは、計算技術の向上等で、より自然な仮定のモデルが誕生。**当該領域の研究の発展につながる**

・ブラックリストベイジアンの考え方

- ・得られたデータの生成メカニズムを考えるならば、○○という仮定は不適切だから取り除こう

➡ 誰もが納得するような事前分布をまずは設定する

ex) 成人男性の身長的事前分布

× $0 - \infty$ (限りなく広い正) の範囲の一樣分布

○ $1.2 - 2m$ の範囲の一樣分布 (身長が $2m$ 以上の人は稀だから)

- ・得られたデータを統計分析では、モデルの一部分について興味があることが多いので、細かいところまで事前分布をきっちり設定

➡ ダメダメ式で事前分布を設定

ex) 2つの群のある指標の平均得点の差を検討したい

群間の平均得点の差といったモデルの一部にしか興味がない

そのデータがどのような群(分布)から生成されたかは興味が無い

➡ この部分に関しては、既存の分析の仮定を流用する

研究目的に応じて、適切に使い分けるべき

- **事後分布**を導出するためには、**尤度**と**事前分布**が必要
 - **事前分布の設定をどうすべきか？**
 - ➡ あなたの研究目的は？何が知りたい？
 - 今回の研究では、ホワイトリストベイジアンになるのか、ブラックリストベイジアンになるのか？又はその中間で行くのか（オフホワイトw）
- 事後分布を導出するためには(再掲)

事後分布 尤度 事前分布

$$p(\theta|y) \propto p(y|\theta)p(\theta)$$

Next データ生成(尤度の部分)をどうやって考えていく？

- **得られたデータを統計分析**

➡ 既にモデルが定まっていることが多い

- **データ生成メカニズムを考えたモデル**

➡ 既存のモデルをベースを元に自分組み立てて行く

- それは、午後のセッションをお楽しみに！(・∀・)ハァハァ

13:30 発表3. ベイズモデリングの実践例

1. 認知心理学への実践 「データ生成メカニズムのベイズモデリング」
武藤拓之（大阪大学/日本学術振興会特別研究員）
2. 臨床心理学への実践 「妄想の認知モデル」
杉取恵太（専修大学）
3. 社会心理学への実践 「ベイズモデリングで見る因子構造」
難波修史（広島大学/日本学術振興会特別研究員）
4. 教育心理学への実践 「ベイズモデリングによる第2種信号検出モデルの表現」
山根高史（川崎医療福祉大学）
5. 青年心理学への実践 「AR法は心理学研究に受け入れられるのか」
ゲスト：豊田秀樹（早稲田大学）

- 熟練のベイジアンモデラーによる心理学実践例を見ることができます

- ・データ生成メカニズムを考えたモデルのために何を？

- ・データをあらゆる視点から眺める

➡ 眺めるとは、単純な可視化・記述統計・相関を算出すること
データがどのような形状(分布)なのか？(属性別・個人別にも出してみる)

ヒストグラム・散布図を描く

- ・データがどのような環境・背景・人物・時間で取れたのか
➡ 他に使えそうなデータを捨ててしまった or 取りこぼしていないか？
実験デザイン・調査デザイン全体を通したデータが得られるまでの過程
を隅々まで追求する

本来、統計分析は事前にやることがたくさん
これからの仮説検証・モデル評価には
これらが非常に重要になってくる

*ベイズか否かは、ここではあまり関係ない

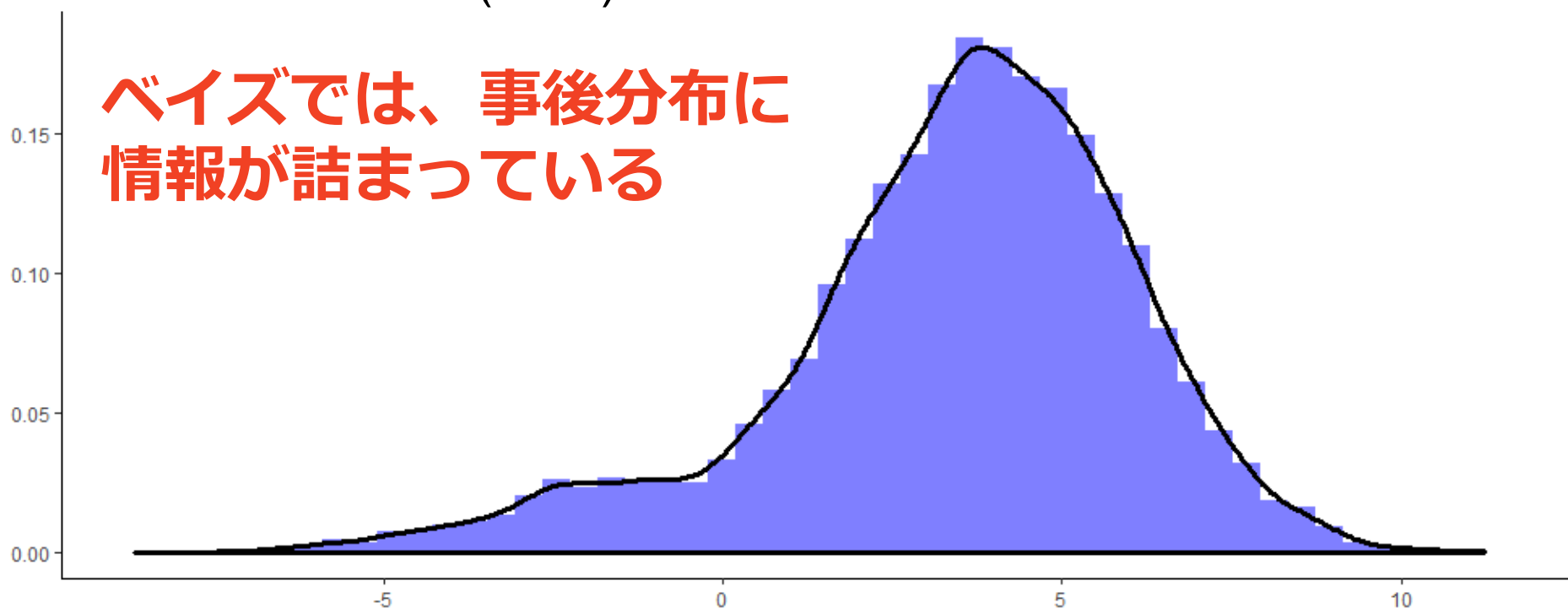
第2部 事後分布の取り扱い方

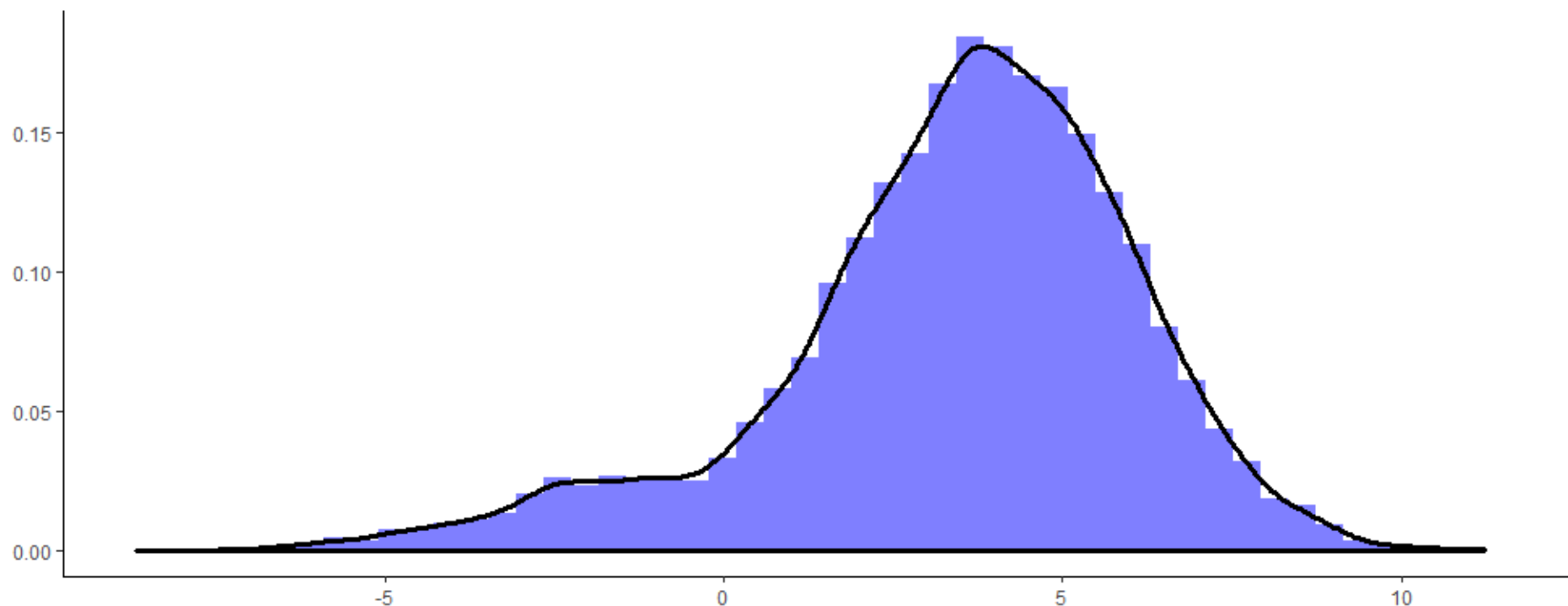
～ 事後分布がでてから ～

事後分布 尤度 事前分布

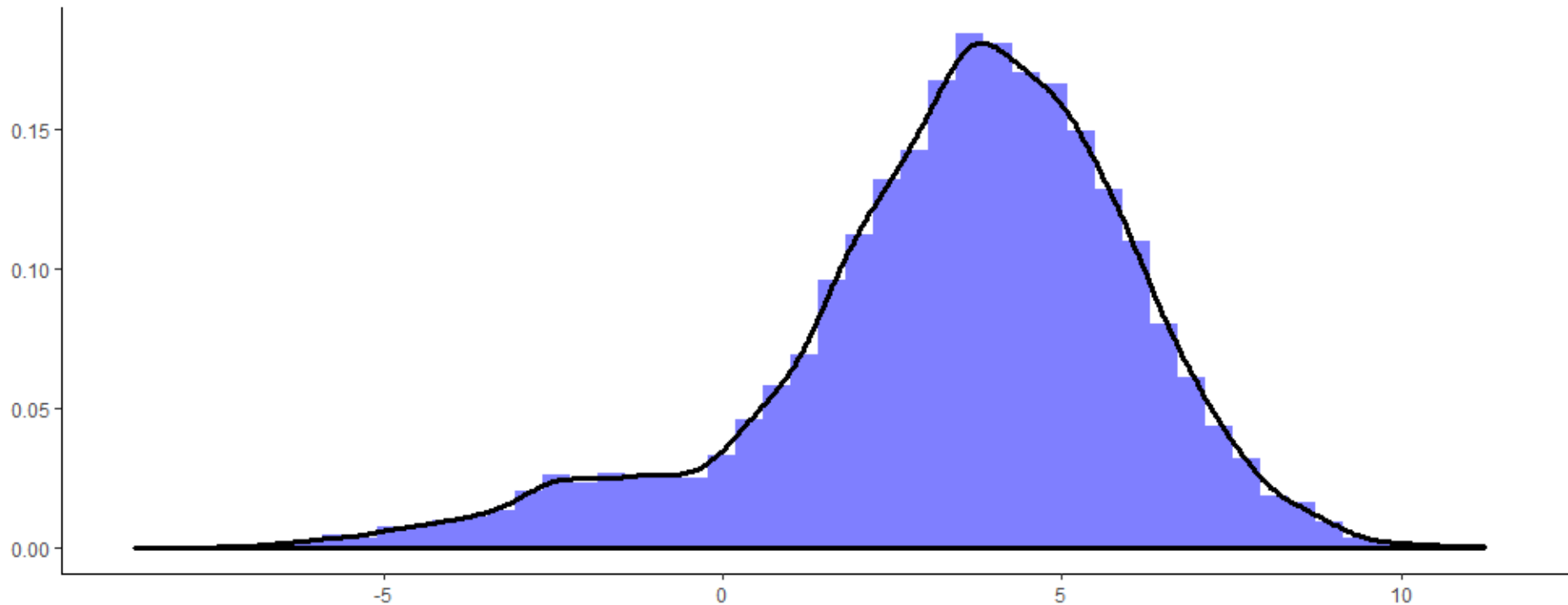
$$p(\theta|y) \propto p(y|\theta)p(\theta)$$

- 事前分布と尤度を設定し、パラメータ推定を実施
- 事後分布を得ることができた *収束していることにします
- 事後分布を可視化(**重要**)

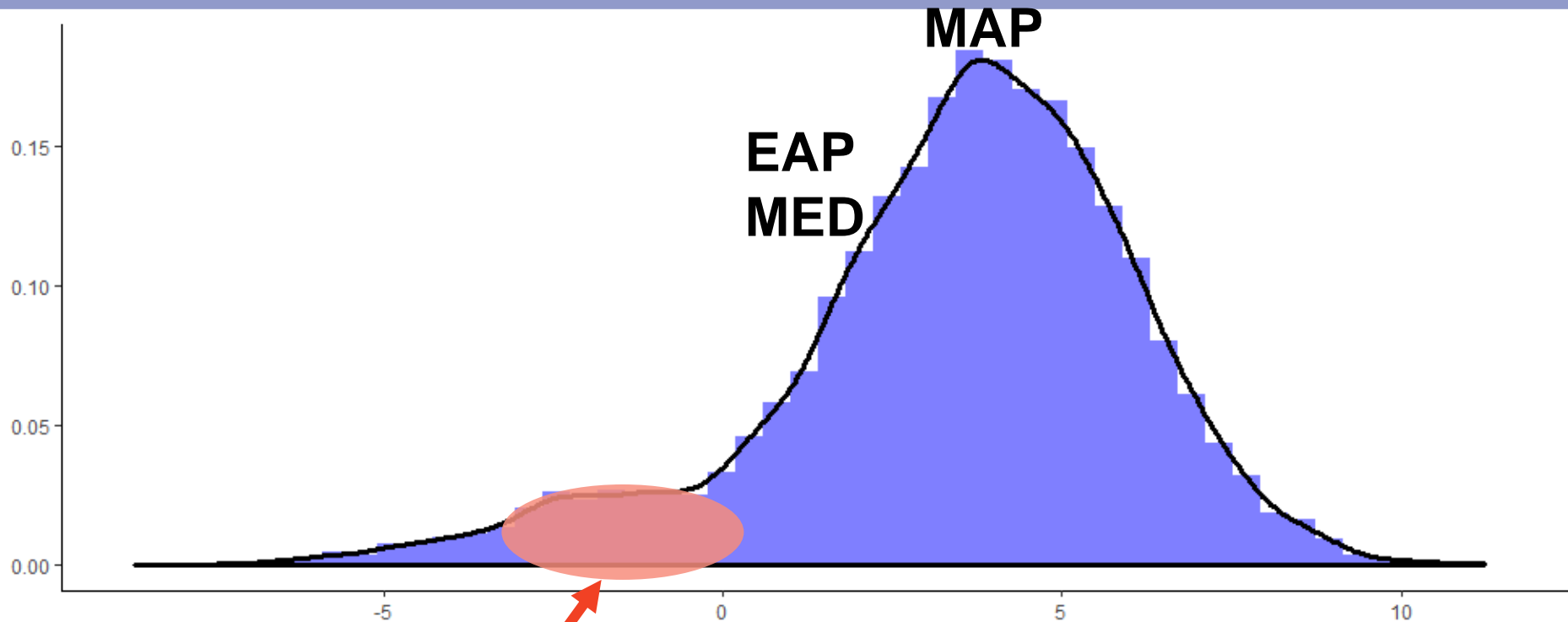




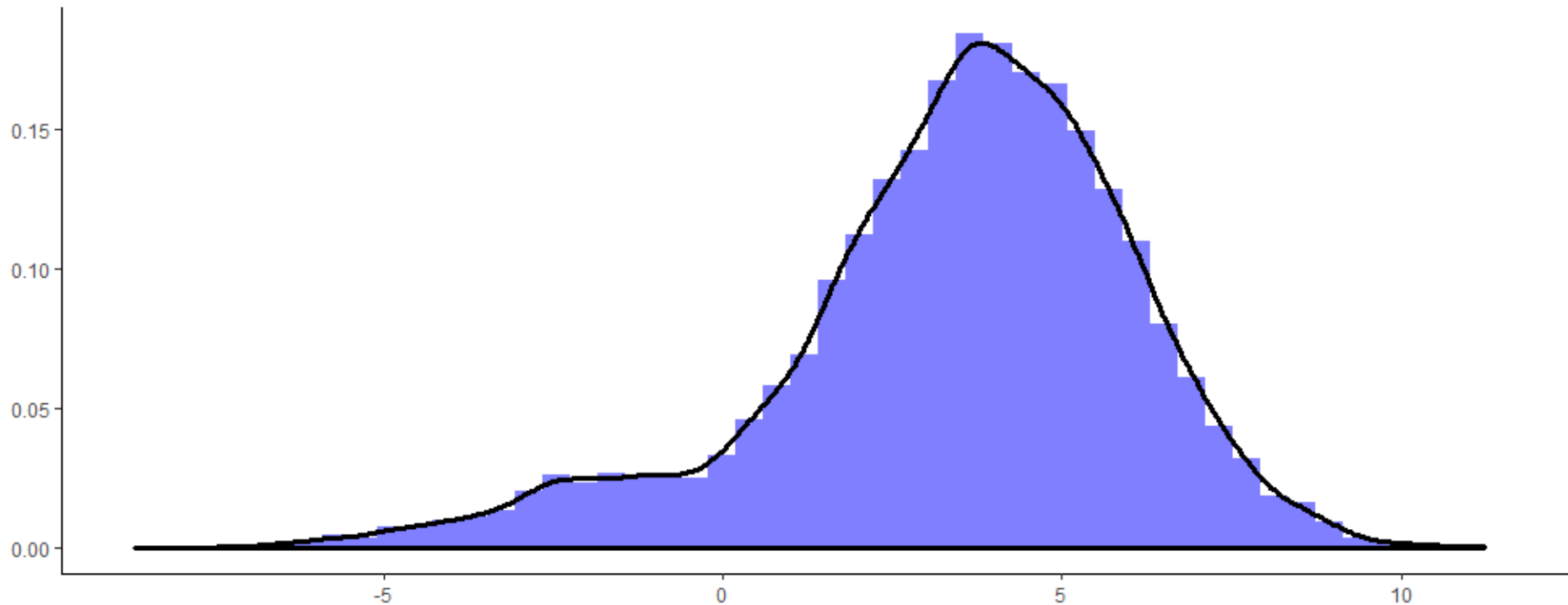
- まずは確認する
 - 左右対称？ ➡ 左右対称には見えない
 - 単峰性？ ➡ 単峰性だが、左に少し山がある
- この事後分布をどうやって、報告する？



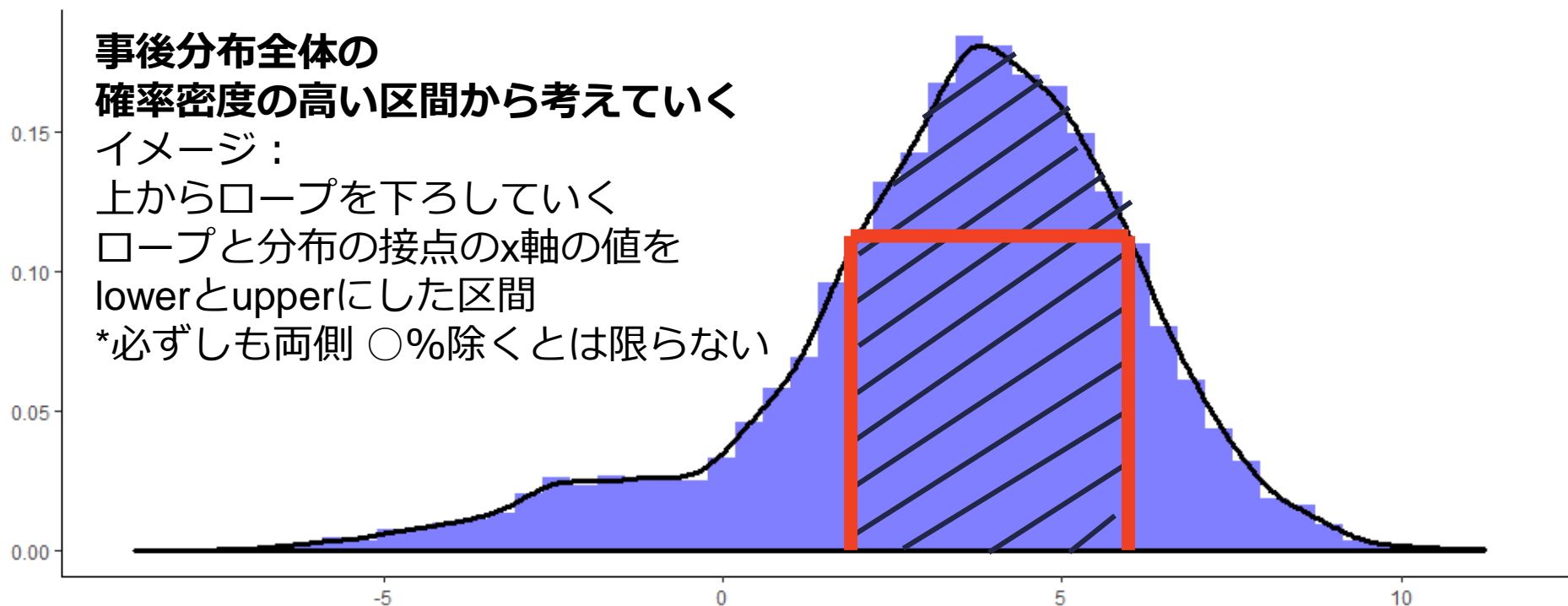
- 一つの統計量に要約する
 - **MAP推定値**・・・事後分布の確率密度が最も高い値
 - **MED推定値**・・・事後分布の中央値
 - **EAP推定値**(事後平均値)・・・事後分布の平均値
- どれを報告すればよい？ ➡ 最良のものはない



- 実際の値
 - $\text{MAP} = 3.83$ $\text{MED} = 1.28$ $\text{EAP} = 1.28$
- MEDとEAPは左に引き寄せられている
 - 分布の端の情報の影響を受けている
 - ➡ 分布の端の情報も反映していると考えることができる

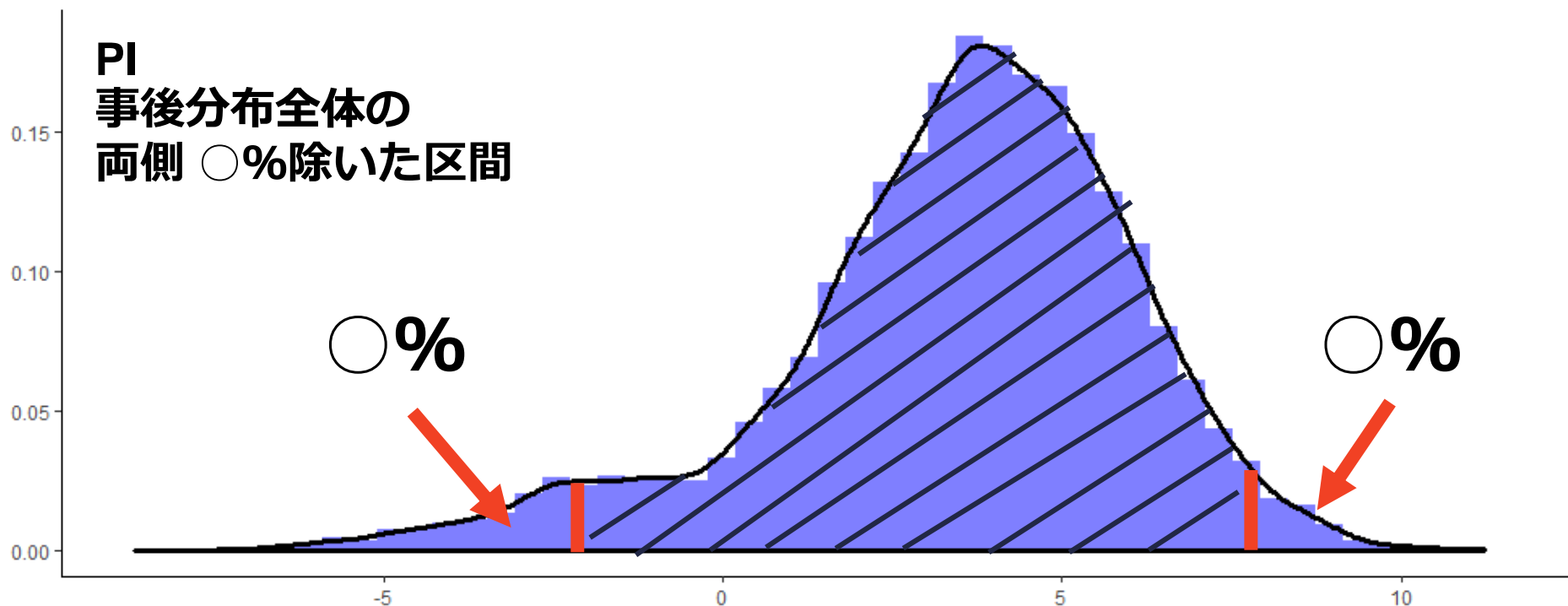


- 信用区間(Credible Interval)
 - HPDI (Highest Posterior Density Interval)
➡ 事後分布全体の確率密度の高い区間
 - PI (Percentile Interval)
➡ 事後分布の両側から同じ%分を除いた区間



- HPDI (Highest Posterior Density Interval)

95%HPDI Lower	95%HPDI Upper
-2.76	8.10



• PI 2.5%点 97.5%点

-3.17	7.79
-------	------

• HPDI 95%HPDI Lower 95%HPDI Upper

-2.76	8.10
-------	------

・事後分布を点・区間要約するには？

- ・点であれば、MAP/MED/EAP推定値というまとめ方がある
- ・MED/EAPは、分布が左右非対称の場合、分布の端の情報を反映させることができる
- ・報告の際に、余裕があるのであれば全て報告した方が良い？
- ・大体一致しているならば、その旨を書いても良いかも

・区間であれば、同じ信用区間という呼び名でも2種類（HPDIとPI）ある

- ・報告の際は、どちらの信用区間なのか記述すべき
- ・もし、二つの区間幅の違いが大きいうのであれば、そもそも区間要約は適切ではないかもしれない(Robert, 2017)

**要約は便利でわかりやすいが、
事後分布の大事な情報を削りすぎないように注意が必要
可能な限り事後分布を報告 (付録や補足資料としてOSF等にUP)**

第3部

これからの仮説検証・モデル評価

～ 複数のモデルで事後分布を算出したら ～

**突然ですが、
あなた知りたいことは現象の
当てはまり？予測？**

反応時間をとる実験 あなたの興味があるのはどちらですか？

- ・ 参加者が次の試行でも同じ反応をするか
- ・ 参加者がこの試行で反応した現象について

質問紙調査

あなたの興味があるのはどちらですか？

- ・ 参加者が次の項目でも同じ反応をするか
- ・ 参加者がこの項目に回答した理由

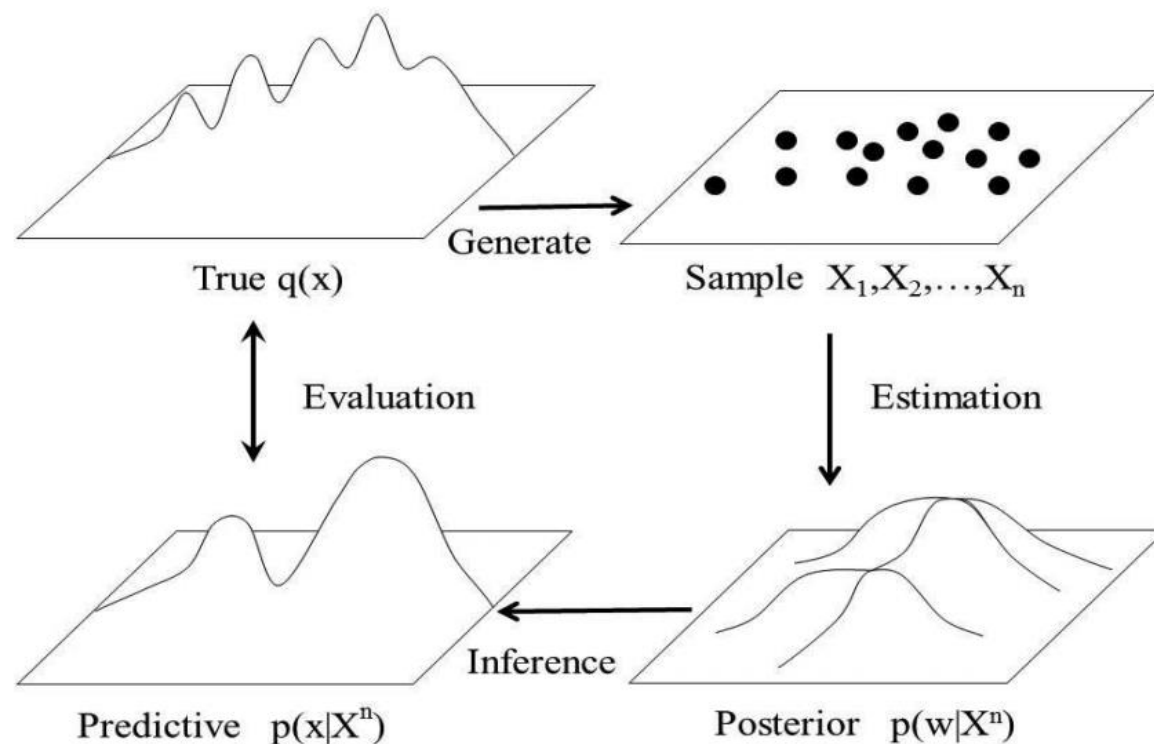


Figure 1.1: Framework of Bayesian inference. The procedure of Bayesian estimation is shown. A sample X^n is taken from unknown true distribution $q(x)$. A statistician sets a statistical model and a prior, then the posterior density $p(w|X^n)$ is obtained. The true distribution $q(x)$ is estimated by a predictive density $p(x|X^n)$, whose accuracy is evaluated by using mathematical laws.

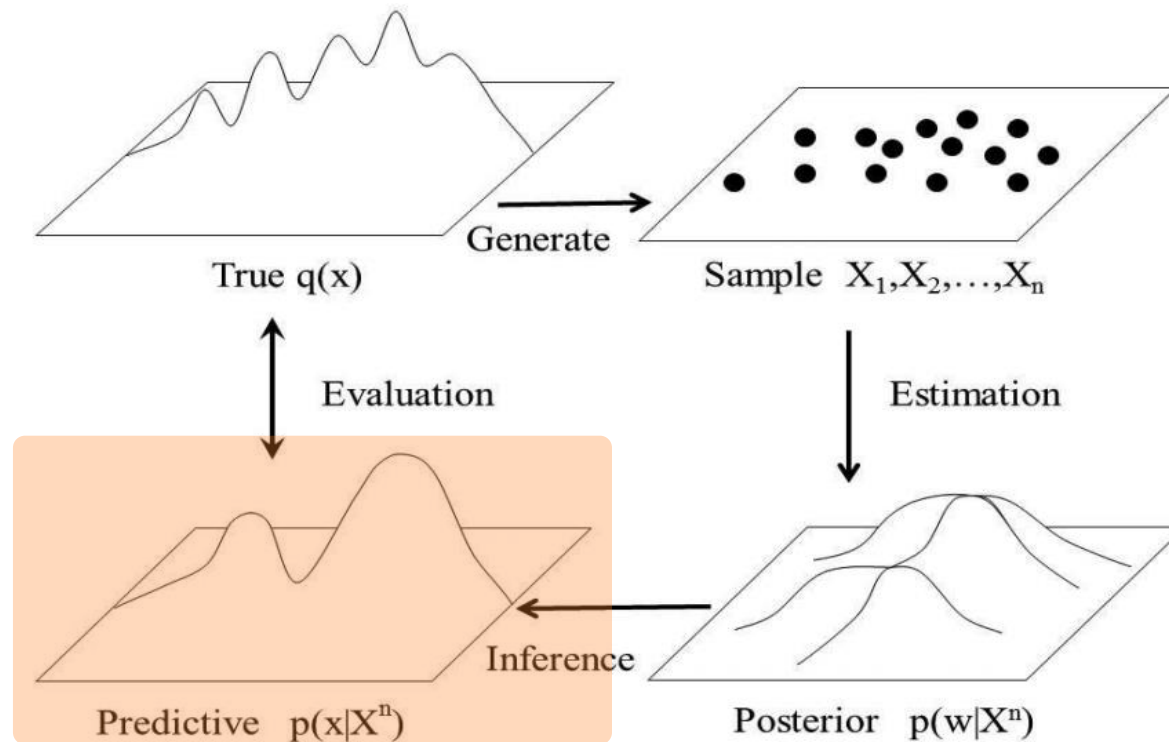
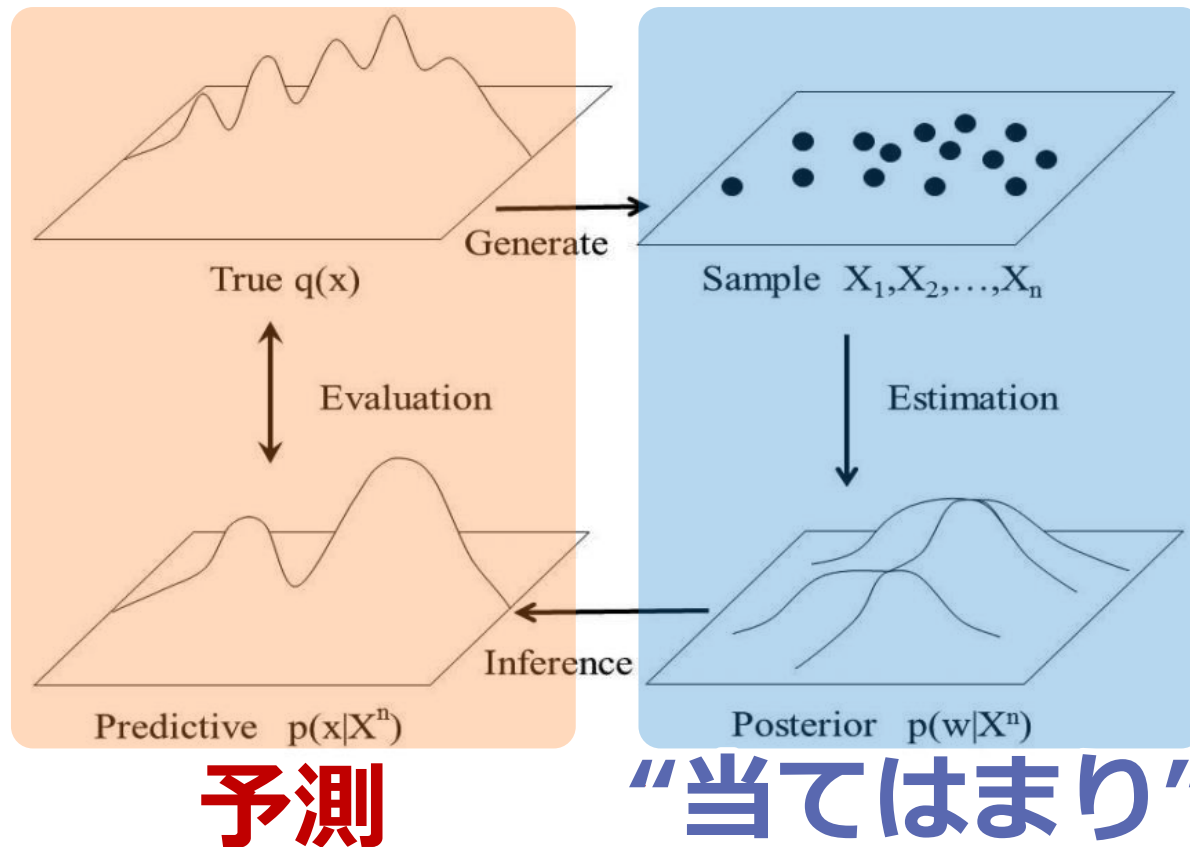


Figure 1.1: Framework of Bayesian inference. The procedure of Bayesian estimation is shown. A sample X^n is taken from unknown true distribution $q(x)$. A statistician sets a statistical model and a prior, then the posterior density $p(w|X^n)$ is obtained. The true distribution $q(x)$ is estimated by a predictive density $p(x|X^n)$, whose accuracy is evaluated by using mathematical laws.

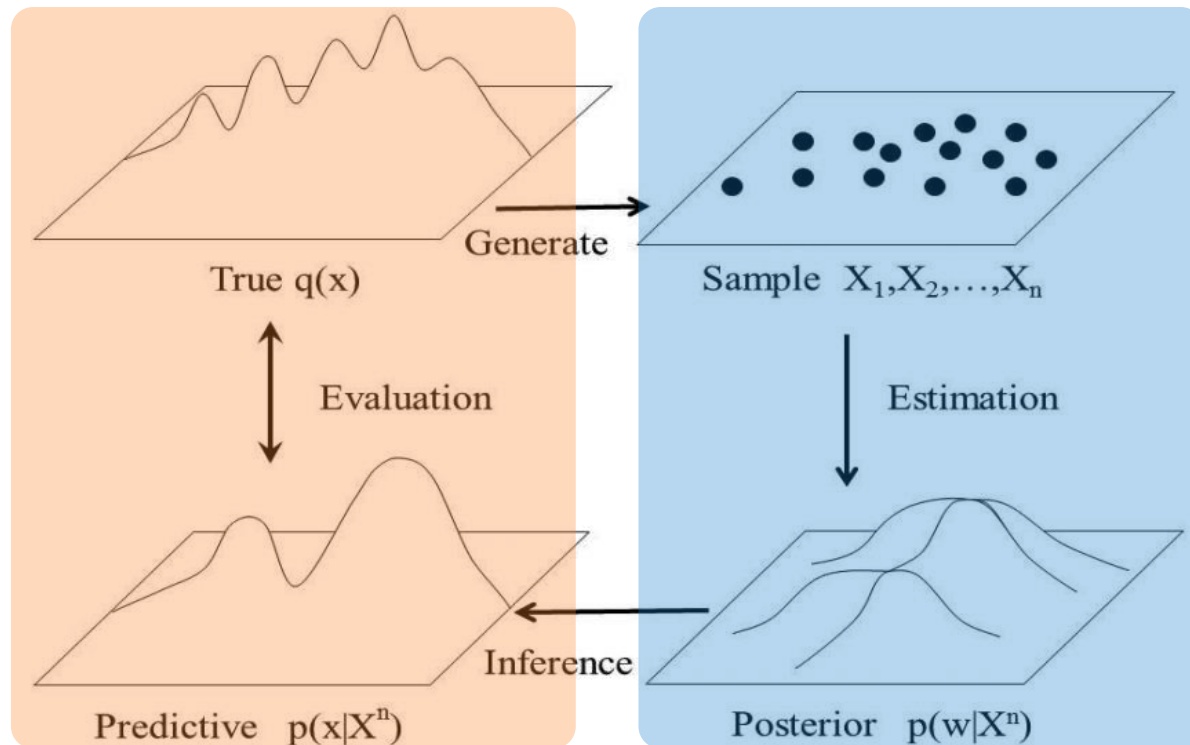


*ここでの当てはまりは
手元のデータと
モデルとの当てはまり
*当てはまりの該当箇所
はイメージです

Figure 1.1: Framework of Bayesian inference. The procedure of Bayesian estimation is shown. A sample X^n is taken from unknown true distribution $q(x)$. A statistician sets a statistical model and a prior, then the posterior density $p(w|X^n)$ is obtained. The true distribution $q(x)$ is estimated by a predictive density $p(x|X^n)$, whose accuracy is evaluated by using mathematical laws.

Watanabe (2018) P.4 Figure1.1のみ引用

(公式 : <https://www.taylorfrancis.com/books/9781482238082>)



予測



事後予測分布

“当てはまり”



周辺尤度 $+\alpha$

*ここでの当てはまりは
手元のデータと
モデルとの当てはまり
*当てはまりの該当箇所
はイメージです

• 周辺尤度って何？

$$\begin{array}{ccc} \text{事後分布} & \begin{array}{cc} \text{尤度} & \text{事前分布} \end{array} & \begin{array}{cc} \text{尤度} & \text{事前分布} \end{array} \\ p(\theta|y) = \frac{p(y|\theta)p(\theta)}{p(y)} = \frac{p(y|\theta)p(\theta)}{\int p(y|\theta)p(\theta)d\theta} \\ \text{周辺尤度} & & \text{周辺尤度} \end{array}$$

• ベイズの定理の分母の部分

- MCMC等での事後分布の導出には、尤度と事前分布があればよかったので使ってなかった

• 現在のデータとモデルの当てはまりを表す**周辺尤度+α**とは？

自由エネルギーのこと ➡ 周辺尤度から計算できる

*自由エネルギーって何？

➡これは、いろいろ端折ると**真の分布と確率モデルの近さ**を表す

• 複数のモデルにおける自由エネルギーを比較することで 現在のデータとモデルの当てはまりを評価できる

➡ **心理学現象の解釈につながる**

- 3つのモデルの自由エネルギーをそれぞれ算出した

	モデル1	モデル2	モデル3
自由エネルギー	-500	-600	-10000

- 自由エネルギーは相対指標なので、いくつ以下になるとOKみたいなものはない！
 - 上に示しているのは例であり、これより大きくなることも小さくなることも当然ある
- 論文にするならば、以下のように解釈できる
- 3つのモデルの自由エネルギーを比較するとモデル3が最も低いので、ほかの2つよりも、現在のデータから得られたモデルと真のモデルが近いと解釈できる

もっと、良いモデルがあるかもしれない?? 38

- 今回は3つのモデルを比較した

	モデル1	モデル2	モデル3
自由エネルギー	-500	-600	-10000

- 検討していない第4のモデルの自由エネルギーを算出したら
-1000000000かもしれない
- モデルは常に間違っていると前提しておく
- 今のモデルが正しいことをチェックするのではなく、観測データをどれだけ記述できていないかを考える(Robert, 2017)
 - Next 周辺尤度とベイズファクターについて ➡
 - 二つのモデルを比較するならば...

- 二つのモデル($m=2$)の比較を考える

$$p(M_1|y) = \frac{p(y|M_1)p(M_1)}{\sum_{j=1}^m p(y|M_j)p(M_j)} \quad p(M_2|y) = \frac{p(y|M_2)p(M_2)}{\sum_{j=1}^m p(y|M_j)p(M_j)}$$

← 同じ →

- 式の分母が等しいので、くっつけて整理(オッズで表す)

事後オッズ **ベイズファクター** 事前オッズ

$$\frac{p(M_1|y)}{p(M_2|y)} = \frac{p(y|M_1)}{p(y|M_2)} \times \frac{p(M_1)}{p(M_2)}$$

ベイズファクター
=周辺尤度の比

- 次のようにも

$$BF_{12} = \frac{p(y|M_1)}{p(y|M_2)} = \frac{\frac{p(M_1|y)}{p(M_2|y)} \text{ 事後オッズ}}{\frac{p(M_1)}{p(M_2)} \text{ 事前オッズ}}$$

ベイズファクターとは、データによって与えられたモデル M_2 に比してモデル M_1 を支持する程度（オッズ）の変化を表す

- モデル(仮説) 1 とモデル(仮説) 2 を用意して、比較を行った
 - すると $BF_{12} = 3$ と求まった。これはどういう意味だ
 - 現在得られたデータはモデル2より、モデル1のほうが3倍上手に表現できる (あってる?)
 - このように、2つのモデル比較を事前オッズと事後オッズ比で考えることができる
 - Nullモデルを用意して、検定っぽく比較することもあります
 - これに関連して、詳細を知りたい人はSavage-Dickey法などと併せて、検索してみてください(怖い人本等)へ
 - ただ、ベイズファクターは比なので、片方のモデルがおかしいと、とても大きな影響を受ける
- ➡事前分布等の影響を大きく受ける(後述)

- 周辺尤度を算出する
- 自由エネルギー or ベイズファクターをもとめる
- じゃあ、なんでこんなに便利なのに今まで使われなかったのか？
 - 周辺尤度を算出するのが大変だった（心理学者は特に）
 - 心理学で事前分布に関する議論が（あまり）行われなかった

- 周辺尤度を計算できないケースが多かった

どんなケースで計算が難しかったのか？

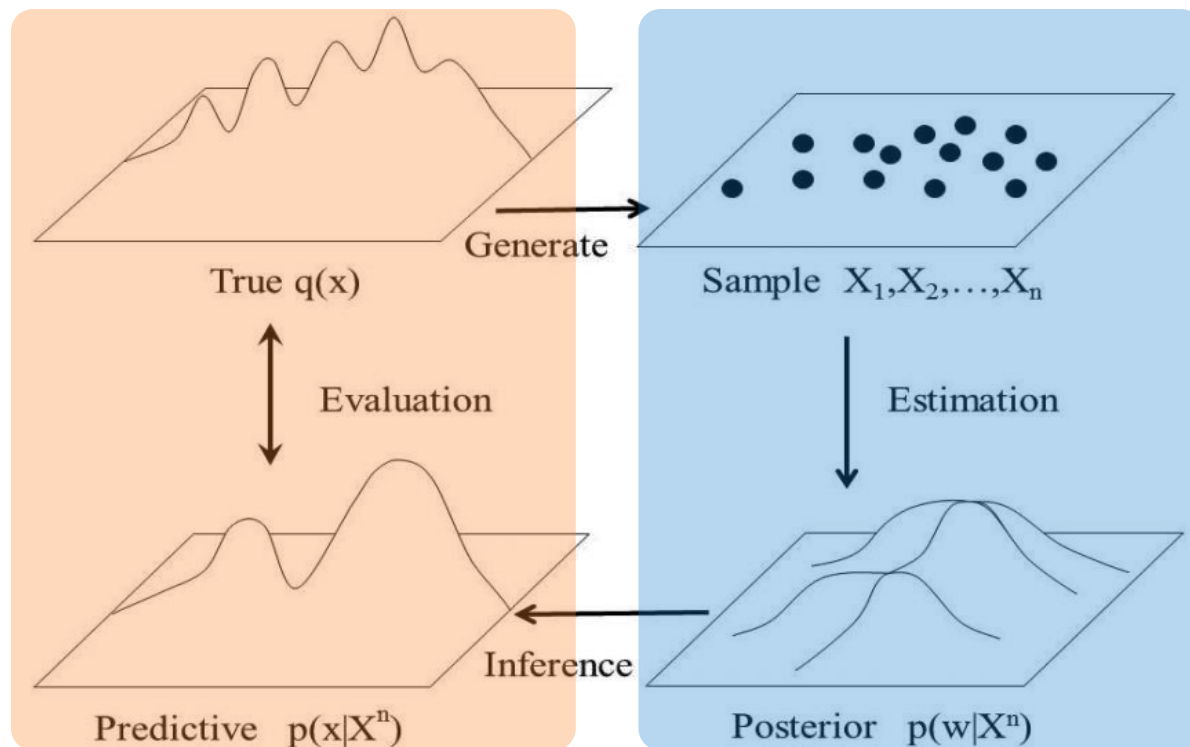
- 非線形モデル
- パラメータ数が多い場合(特に階層モデル)
- しかし、近年、周辺尤度を計算できない時代は終わりを告げようとしている
 1. Naive Monte Carlo Estimator
 2. Importance Sampling Estimator
 3. Generalized Harmonic Mean Estimator
 4. Bridge Sampling
 5. WARP-III bridge sampling or WARP-U bridge sampling

詳しい話は、[ここ](#)

$$\text{事後分布 } p(\theta|y) = \frac{\text{尤度 事前分布 } p(y|\theta)p(\theta)}{\text{周辺尤度 } p(y)} = \frac{\text{尤度 事前分布 } p(y|\theta)p(\theta)}{\int \text{周辺尤度 } p(y|\theta)p(\theta)d\theta}$$

• 周辺尤度は事前分布の影響を大きく受ける

- 事前分布の選択がとても重要になる
 - どこまでホワイトリストベイジアンで、どこまでブラックリストベイジアンでいくのか？
 - [清水\(2018\)](#)は、心理学者も真剣に事前分布を考える時だと提言
 - **Farrell & Lewandowsky (2018)**では、心理学者がベイズファクターを使用する際の事前分布をどうすべきか解説
 - ➡曖昧な事前分布はかえって、ベイズファクターに悪影響
 - ➡ベイズファクターがすべてのモデル比較を解決するわけではない
 - ➡ベイズファクターの基準を決めて判断するのはナンセンス



予測



事後予測分布

“当てはまり”



周辺尤度 $+\alpha$

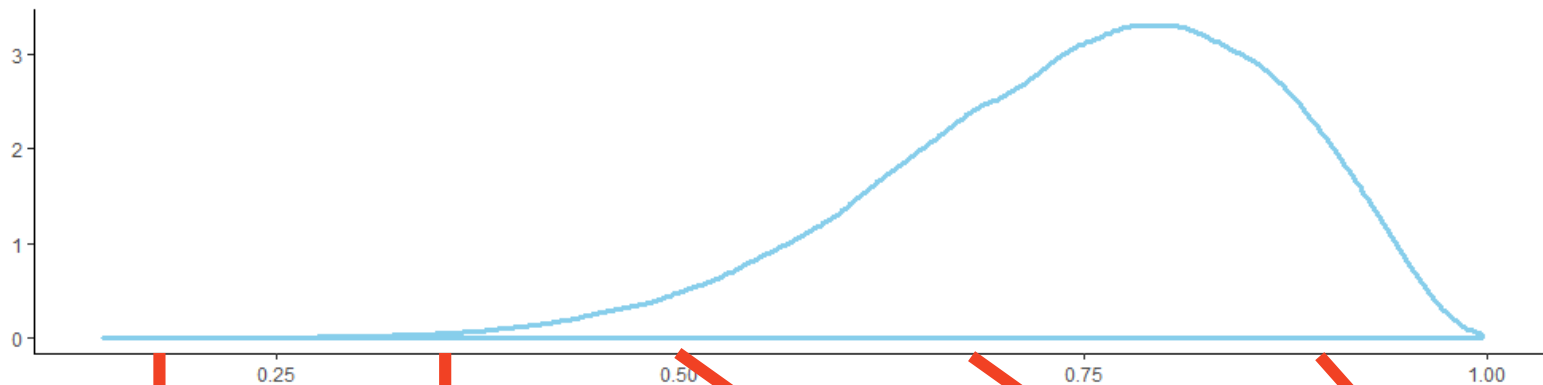
*ここでの当てはまりは
手元のデータと
モデルとの当てはまり
*当てはまりの該当箇所
はイメージです

予測の評価につながる事後予測分布

45

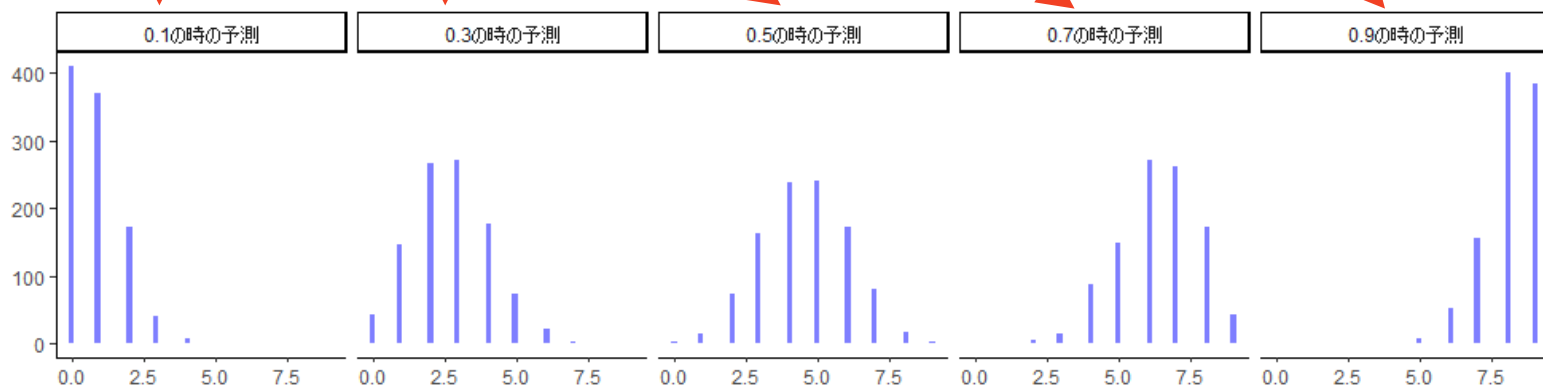
Robert (2017)を参考に作成

事後分布



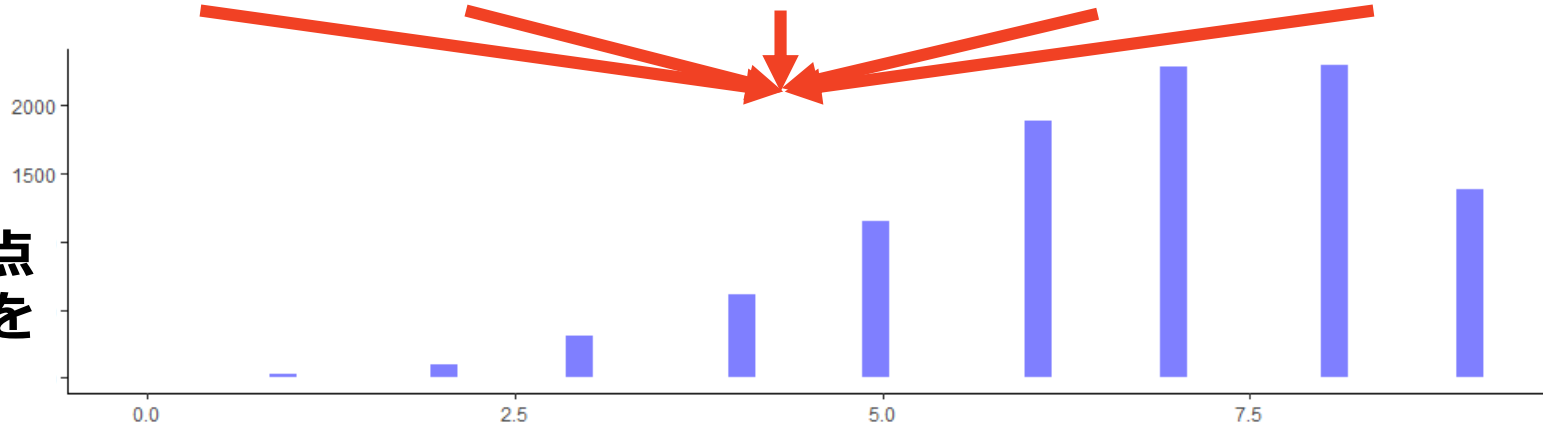
条件付き
予測分布

事後分布の
各点だけで
予測した分布



事後
予測分布

事後分布の各点
予測したものを
統合した分布



- 事後予測分布と実際のデータの分布を比較してみる
- 大きな相違がなければ、今回のベイズモデルの予測がきちんとできていることを表す

Next 事後予測分布以外の予測について検討する方法
情報量基準の話➡

- WAICやLOOという情報量基準によって、将来予測の観点からモデルを評価することができる。

• WAICって何？

- とても省略して書くと、AICの改良の改良版
- AICは、複雑すぎるモデルが”良い”モデルとされないために罰則を付加
➡しかし、**この計算には最尤推定値が用いられている**
- ベイズでいうと、MAP推定値が用いられていることを指す
- ここで、事後分布の点要約の話を出そう
- MAP推定値は事後確率密度の高いところ表現する代わりに、ゆがんだ分布だと分布の端の情報を考慮できない
➡では、**計算に事後平均値(EAP)を使ってやろう！ってのがDIC**

しかし、ベイズの本質を改めて思い出してほしい
ベイズでは、事後分布に全ての情報が込められている

➡**事後分布の各点の値を用いて、計算を行おう！ってのがWAIC**

*注：DICもWAICも
罰則項も変化している

- **LOOって何？** (LOOは、PSIS-LOO, LOO-CVとも呼ばれる)

- 正式名称は

Pareto Smooth Importance Sampling – Leave One Out – Cross Validation

- やっていることは、PSISというすごいアルゴリズムに基づいて、データを1つとって、それを残りのデータで予測できるかをすべての組み合わせで繰り返す
- WAICとLOOは近い値になる。そういう性質がある

- **予測に興味がある？当てはまりに興味がある？**
- 当てはまりなら、周辺尤度を考えることが非常に重要になってくる
 - 自由エネルギーやベイズファクターとお友達になろう
 - 適切な事前分布の選択が必要になる
- 予測なら、事後予測分布を書いてみるのが大事
 - また、別の方法として情報量規準による方法もある
 - 基本的にこれらの指標は相対的な指標である
- **どんなモデルチェックやモデル比較にも有効な手段は存在しない。状況に応じて方法を変える必要あり**
 - そのために引き出しを増やそう

**モデル比較で決めるよりも試行錯誤を繰り返して
“素晴らしい”モデルを作ろう**

心理学者が ベイズアンになるためのツール

- いままでの心理学は、**得られたデータを統計分析**
 - これからの心理学は、**データ生成メカニズムを考えたモデル**
- もちろん、ベイズ統計は、どちらにも対応できる

• 調査データ・・・項目回答データ



ID	Q1	Q2	Q3
1	5	2	3
2	2	1	1
3	3	2	4



- カットオフを超えた？
- 平均得点は？
- 2群で差がある？

等々

• 実験データ・・・反応時間(RT)データ



ID	条件	RT
1	1	3.25
1	2	4.17
2	1	3.36
2	2	4.01



- 平均RTは？
- 2条件で差がある？

等々

知らぬ間に仮定

t検定モデル・分散分析モデル

仮定を設定する必要

拡散モデル・項目反応モデル

- いままでの心理学は、**得られたデータを統計分析**

➡ こちらから始めるのが楽

- これからの心理学は、**データ生成メカニズムを考えたモデル**

➡ ベイズの雰囲気が出てきたらこっち/別に初めからでも

- どう始める？

- **得られたデータを統計分析** ➡ JASP

➡ 最後にこのお話と実践

- **データ生成メカニズムを考えたモデル** ➡ brms / JAGS / Stan

➡ こちらを学びたい人は、関連書籍や午後の話を参考にしてね

A Fresh Way to Do Statistics

 Download JASP



FREE

JASP is an open-source project supported by the University of Amsterdam.



FRIENDLY

JASP has an intuitive interface that was designed with the user in mind.



FLEXIBLE

JASP offers standard analysis procedures in both their classical and Bayesian form.

- **JASP**は、統計解析を行うための新たな”方法”(ソフト)である。
 - 完全無料
 - オープンソース
 - クロスプラットフォーム
- 見やすい画面、そして、ドラッグ&ドロップ操作で簡単に統計分析できる
- **JASP**は、**ANOVA**や**回帰分析**のような**古典的分析**だけでなく、それらの**ベイズ統計分析**も実施することができる
- ➡モデリングはできない
- まだ誕生したばかり。開発が盛ん(~~すぎて資料作りに困った~~)

- 多くの利点
 - SPSSのような商用パッケージと比較して、JASPは無料で、オープンソースで、開発が盛ん、そして、パラメータ推定法としてベイズ法を使える。
 - Rのようなフリーパッケージと比較しても、ドラッグ&ドロップで簡単に操作できる。プログラミング能力を必要としない。
 - その他の全てのパッケージと比べて、JASPは基本となる分析を古典的かつベイズ的に実行することができる。
- Windows/Mac/Linux の全てで使用することが可能である。

- JASPって何の略？
 - Jeffereys's **A** amazing **S**tatistics **P**rogram
 - Bayesian pioneer Sir Harold Jeffreysが驚愕の統計プログラム
- JASPって何のプログラミング言語で書かれているの？
 - アプリケーション自体は**C++**で書かれている。
 - 分析は、**R**と**C++**のどちらもで書かれている(R packageも導入済)。
 - その他、ソフト外観で**javascript**も使用されている。
- どのライセンスでJASPは公開されているのか？
 - [The GNU Affero General Public License, Version 3](#)

Frequentist Analyses

ANCOVA
ANOVA
Binomial Test
Contingency Tables
Correlation
Descriptive Statistics
Exploratory Factor Analysis
Hierarchical Regression
Independent Samples T-Test
Linear Regression
Logistic Regression
Log-Linear Regression
One Sample T-Test
Paired Samples T-Test
Principal Component Analysis
Repeated Measures ANOVA
Reliability Analysis
Structural Equation Modeling

Bayesian Analyses

ANCOVA
ANOVA
Binomial Test
Contingency Tables
Correlation
Independent Samples T-Test
Linear Regression
Log-Linear Regression
One Sample T-Test
Paired Samples T-Test
Repeated Measures ANOVA
Summary Stats

Functions & Modules

Data & Label Editing
Exact P-Values
How to Test Interval-Null Hypotheses in JASP
OSF support
VS-MPR
Network Module



Alcohol Attitudes*

FileCommon+

Descriptives

T-Tests

ANOVA

Regression

Frequencies

Factor

beerpos

1	1
2	43
3	15
4	40
5	8
6	17
7	30
8	34
9	34
10	26
11	1
12	7
13	22
14	30
15	40
16	15
17	20
18	9
19	14
20	15

winepos

wineneg

wine neut

waterpos

water neg

water neu

participant

Repeated Measures Factors

RM Factor 1

Level 1

Level 2

level 3

Level 4

RM Factor 2

OK

Repeated Measures Cells

beerpos

beer neg

beer neu

Level 1

Level 2

level 3

Between Subject Factors

Covariates

Bayes Factor

☒ BF_M

☐ BF_w

☐ Log(BF_M)

Order

☒ Compare to null model

☐ Compare to best model

Output

☒ Effects

☒ Across all models

☐ Across matched models

☐ Descriptives

Model

Components

RM Factor 1

Model terms

RM Factor 1

Add to null model

Bayesian Repeated Measures ANOVA

Model Comparison

Models	P(M)	P(M data)	BF _M	BF ₁₀	error %
Null model (incl. subject)	0.500	2.207e-4	2.208e-4	1.000	
RM Factor 1	0.500	1.000	4529.194	4529.194	0.920

Note: All models include subject.

Analysis of Effects

Effects	P(incl)	P(incl data)	BF _{Inclusion}
RM Factor 1	0.500	1.000	4529

Post Hoc Tests

Post Hoc Comparisons - RM Factor 1

	Prior Odds	Posterior Odds	BF _{10, U}	error %
Level 1				
Level 2	0.587	257.179	437.826	8.749e-7
level 3	0.587	36.176	61.586	1.565e-6
Level 2				
level 3	0.587	0.866	1.475	0.002

Note: The posterior odds have been corrected for multiple testing by fixing to 0.5 the prior probability that the null hypothesis holds across all comparisons (Westfall, Johnson, & Utts, 1997). Individual comparisons are based on the default t-test with a Cauchy (0, r = 1/sqrt(2)) prior. The "U" in the Bayes factor denotes that it is uncorrected.

Descriptives

Descriptives Plot

Level	Value
Level 1	24
Level 2	1
level 3	10

実際に操作をお見せします

- 操作方法についてのスライド資料



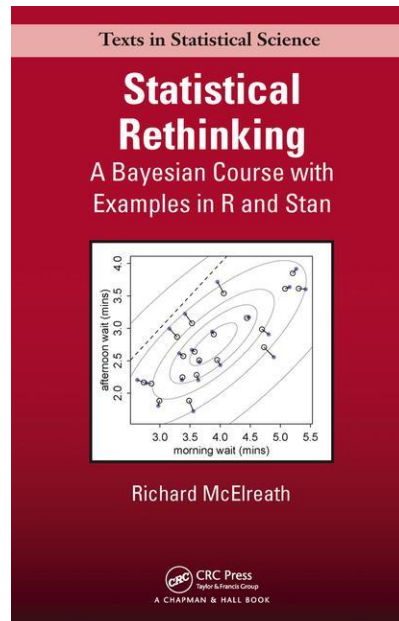
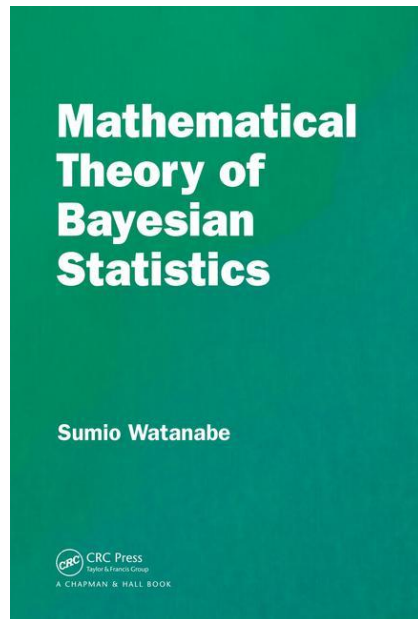
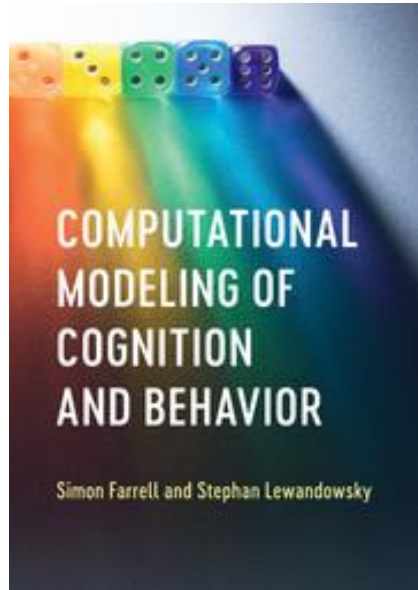
<https://www.slideshare.net/daikihojo/jasp-89875504>

- 日本心理学会 第82回大会@仙台にて
ワークショップ

「心理学者のためのJASP統計解析入門」

開催します(通ってればw)

興味があればご検討ください



- Farrell, S., & Lewandowsky, S. (2018). *Computational Modeling of Cognition and Behavior*. Cambridge University Press. [[LINK](#)]
- Robert, C. (2017). *Statistical Rethinking*. CRC Press. [[LINK](#)]
- Watanabe, S. (2018). *Mathematical theory of Bayesian statistics*. CRC Press. [[LINK](#)]

- 本スライド(主に参考書籍)は、特別研究員奨励費（18J22162）の助成を受けたものです。関係者の皆様ありがとうございます。