マーケティング領域における ベイジアンモデリングの利用

佐野　宏喜
アクセンチュア株式会社

Bayesian modeling in marketing

SANO Hiroki
Accenture Japan

要旨:

ビッグデータの拡大を背景として、マーケティング領域での統計的手法の利活用も進展してきた。

本発表ではベイジアンモデリングを用いることによる、より一層のデータ分析高度化の一端を示すとともに、SASでの実装も合わせて紹介する。

キーワード:ビッグデータ, マーケティングサイエンス, MCMC, ベイジアンモデリング, 階層ベイズ, 状態空間モデル

注）

本発表は個人の見解であり、アクセンチュアの考えを代表するものではありません。

アジェンダ

1. マーケティングにおけるベイジアンモデリングの必要性
2. ベイジアンモデリングの基礎
3. 階層ベイズを用いた「個」の理解
4. 状態空間モデルを用いた動的表現
5. まとめ

アジェンダ

1. マーケティングにおけるベイジアンモデリングの必要性
2. ベイジアンモデリングの基礎
3. 階層ベイズを用いた「個」の理解
4. 状態空間モデルを用いた動的表現
5. まとめ

マーケティング戦略/戦術のなかでの統計的手法の利用領域はどんどん増えています。

PDCAサイクル

セグメンテーション → ターゲティング → ポジショニング

	セグメンテーション	ターゲティング	ポジショニング			
			製品（Product）	価格（Price）	流通（Place）	販促（Promotion）
考慮事項	• 市場・顧客の細分化 ? 顧客はどの単位で分けるべきか ? マーケティング施策の単位を作成	• 優先セグメントを明確化 ? 収益性の高いセグメントはどれか ? 今後成長するセグメントはどれか	• 売るべき商品・サービスの明確化 ? どんな商品が求められているか	• 提供価格の決定・原価の管理 ? いくらで商品を売るべきか ? どこまで質を高めるべきか	• チャネルの選定 ? どこで商品を売るべきか	• 販促活動による売上拡大 ? 誰にコストをかけるべきか ? いくらまでコストをかけるべきか
求められる分析手法	クラスター分析		アソシエーション分析	回帰分析	階層ベイズ（LTV計測）	
	主成分分析	時系列分析	離散選択モデル		ニューラルネットワーク	項目反応理論
	因子分析	コレスポンデンス分析	MDS	回帰分析	最適化	決定木

※分析手法は様々なシーンで活用できるため、よく利用される観点で一例を示した

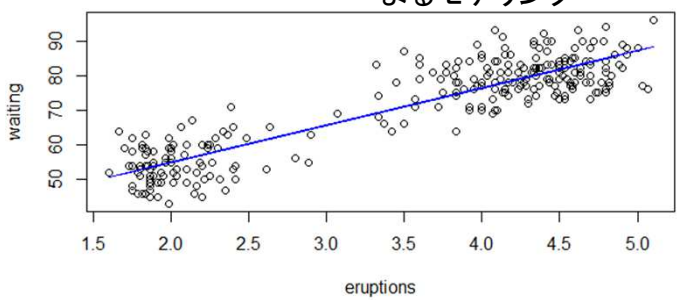
多岐に渡って、より精度の高い分析・予測が求められている

一方で、単純なモデルでは実際の利用ケースにそぐわない場合も多くあります。

単純な手法によるモデリング

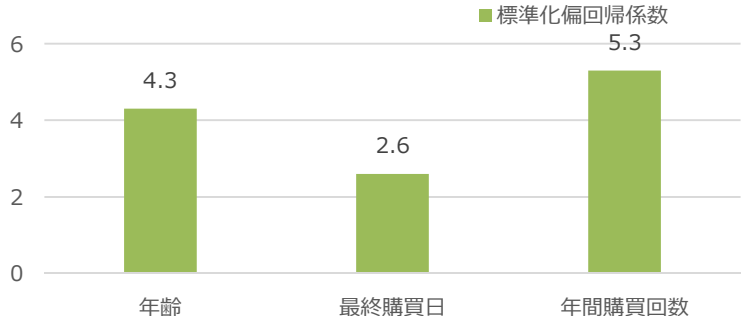
顧客	年齢	最終購買日	年間購買回数	新商品購買有無
Aさん	24歳	14/12/01	15階	有
Bさん	42歳	15/02/06	12階	無
Cさん	36歳	15/01/10	9階	有
…	…	…	…	…

線形回帰分析によるモデリング



- 線形回帰分析により特定の商品の購買確率を予測
- 投入変数として、年齢、最終購買日、年間購買回数を利用

分析結果と問題点



顧客	年齢	最終購買日	年間購買回数	予測購買確率
Xさん	2歳	14/02/27	42階	89%
Yさん	4歳	14/02/25	28階	88%
Zさん	4歳	14/02/26	29階	88%
…	…	…	…	…

予測確率高

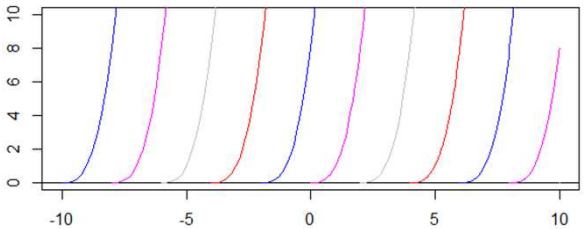
- 良いと予測された顧客は説明変数の最大(or 最小)の値を持つ顧客になってしまう

➔ 本来は優先すべきではない顧客を優先してしまう

柔軟な構造を非線形モデルで表現するには多くのパラメータが必要になります。

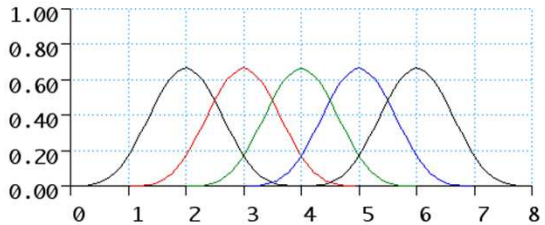
基底変換を用いたアプローチ

3次スプライン

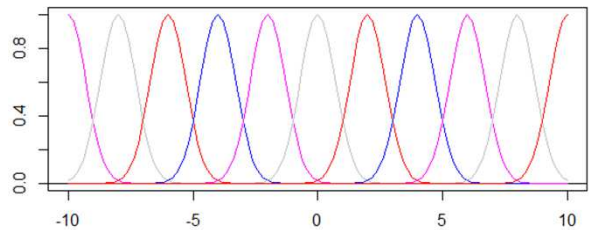


B-スプライン

※1



ガウス型基底関数



3次スプライン関数の式: $u(x;\boldsymbol{\theta}) = \beta_0 + \beta_1 x + \beta_2 x^2 + \beta_3 x^3 + \sum_{i=i}^{m} \theta_i (x - t_i)_+^3$

推定するパラメータベクトル $\boldsymbol{\theta} = (\beta_0, \beta_1, \beta_2, \theta_1, \theta_2, \ldots, \theta_m)^T$

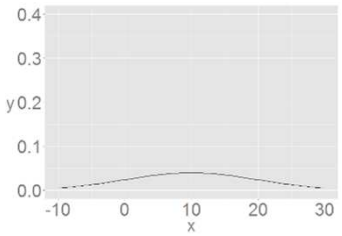
利用する説明変数が1つでも、
推定するパラメータは数十にもなる

説明変数を多い場合はかなり不安定なモデルとなってしまう

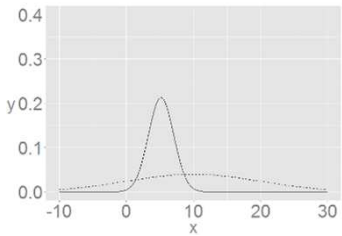
※1) http://escience.anu.edu.au/lecture/cg/Spline/printCG.en.html

ベイズ統計を用いて“常識”を教えることで、多くの事象を適切にモデル化できるようになります。

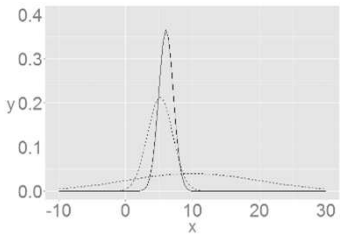
	伝統的統計学でのアプローチ	ベイズ統計を用いたアプローチ
データ数：0	• 情報は一切なし	• 事前分布を設定 • （無情報事前分布や勘と経験から）
データ数：1	• データはたまたま取れた値のみ利用できる • データの広がり(=分散)は分からない	• 事前分布を更新 • 分散が小さくなり、推定値の確からしさを分布で表現
データ数：2	• 標本分散を利用して母分散を推定可能 • 但し、2つのデータの近さからのみ推定される（=不安定）	• 事前分布を更新 • 分散が小さくなり、推定値の確からしさを分布で表現



事前分布を更新



事前分布を更新



アジェンダ

1. マーケティングにおけるベイジアンモデリングの必要性
2. ベイジアンモデリングの基礎
3. 階層ベイズを用いた「個」の理解
4. 状態空間モデルを用いた動的表現
5. まとめ

ベイズ統計はこれまで長い間、封印されてきました。

ベイズ統計の変遷

- 1740年代にトマス・ベイズ(1702-1761)がベイズの定理を発見、しかし未発表のまま没
- 1763年、リチャード・プライスが英国ロイヤルソサエティの哲学寄稿に発表
- その後、ピエール=シモン・ラプラス(1749-1827)も独自にベイズの定理を発見し、体系化
- 20世紀前半、ロナルド・フィッシャー(1890-1962)・イェジ・ネイマン(1894-1981)によって、分析手法が徹底的に手続き化される
  それとともに、徹底的にベイズ統計学が否定される
- 1980年代からベイズ統計の有用性が再評価される
- **21世紀になり、コンピューターの発展と相まって、ベイズ統計のビジネスへの利用が飛躍的に増加**

ベイズの定理を活用することで、条件付き確率を実世界のモデリングに利用できるようになります。

- ベイズの定理

Y：事象
D：データ

事象Yが起こる場合のデータDの存在確率

事象Yそのものの生起確率

$$\boxed{P(Y|D)} = \frac{\boxed{P(D|Y)}\boxed{P(Y)}}{\boxed{P(D)}}$$

データDが存在する場合の事象Yの生起確率

データDそのものの生起確率

データが与えられた場合の事象の発生確率を表現することができる

ベイズ統計の例：傘の売上から天気を予測する①

【例題】
過去2週間の天気と傘の売上データが存在する。
ある日の傘の売上が「大」の時、天気が雨の確率は何%か？

	第1週						
	月	火	水	木	金	土	日
天気							
傘売上	小	小	小	中	大	小	大

	第2週						
	月	火	水	木	金	土	日
天気							
傘売上	大	中	中	小	小	中	大

ベイズ統計の例：傘の売上から天気を予測する②

【例題】
過去2週間の天気と傘の売上データが存在する。
ある日の傘の売上が「大」の時、天気が雨の確率は何%か？

統計量の整理

	傘売上		
	小	中	大
晴れ	4日	1日	0日
曇り	2日	1日	1日
雨	0日	2日	3日

確率の計算

ベイズの定理

$$P(\text{天気雨}|\text{売上大}) = \frac{P(\text{売上大}|\text{天気雨})P(\text{天気雨})}{P(\text{売上大})}$$

- $P(\text{売上大}|\text{天気雨}) = 3/5 = 0.6$
- $P(\text{天気雨}) = 5/14 \fallingdotseq 0.36$
- $P(\text{売上大}) = 4/14 \fallingdotseq 0.29$

ベイズ統計の例：傘の売上から天気を予測する③

【例題】
過去2週間の天気と傘の売上データが存在する。
ある日の傘の売上が「大」の時、天気が雨の確率は何%か？

$$P(\text{天気雨}|\text{売上大}) = \frac{P(\text{売上大}|\text{天気雨})P(\text{天気雨})}{P(\text{売上大})}$$

$$= \frac{0.6 \times 0.36}{0.29}$$

$$= 0.75$$

⇒ 売上が「大」の時、天気が雨の確率は**75%**

ベイズの定理を統計モデルのパラメータに利用することで、データ数が少ない場合でも安定した結果を得ることができます。

- 線形ベイズモデル（切片は固定）

単回帰分析の傾きβ_1をベイズ推定する

$$\boxed{y = 20 + \boxed{\beta_1} x + \varepsilon}$$

$$\boxed{P(\beta_1|D)} = \frac{P(D|\beta_1)P(\beta_1)}{P(D)}$$

β_1はパラメータなので、確率分布を考える

事後分布　　尤度　　事前分布

$$\boxed{f(\beta_1|D)} = \frac{f(D|\beta_1)f(\beta_1)}{\boxed{f(D)}} \propto \boxed{f(D|\beta_1)}\boxed{f(\beta_1)}$$

パラメータ（β）とは無関係のため、最後に考える

事前分布と尤度を適切に指定することで、事後分布を確率密度関数として求めることができます。

- 線形ベイズモデル

尤度　事前分布　事後分布

$$f(D|\beta_1)\,f(\beta_1) \propto f(\beta_1|D)$$

	分布の仮定	式の形	事後分布
尤度	正規分布	$\prod_{i=1}^{n} \frac{1}{\sqrt{2\pi}\sigma} exp\{-\frac{1}{2\sigma^2}(y_i-\beta_1)\}$	新たな平均・分散を持つ正規分布
事前分布	正規分布	$\frac{1}{\sqrt{2\pi}\sigma_0} exp\{-\frac{(\beta_1-\mu_0)^2}{2\sigma_1^2}\}$	

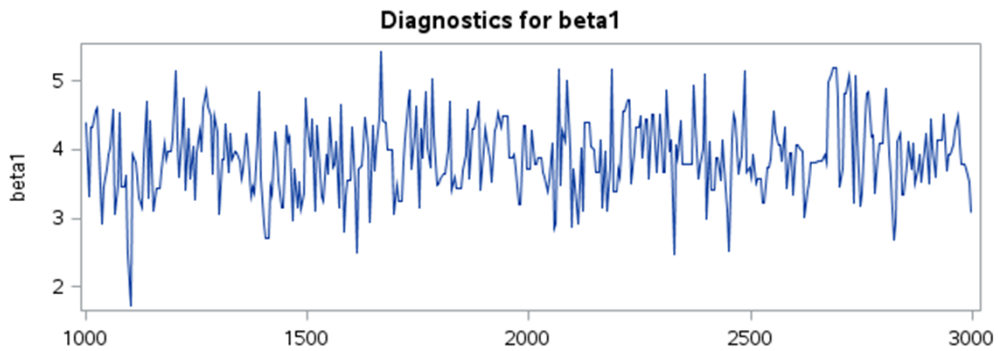
パラメータ推定

MCMC（マルコフ連鎖モンテカルロ法）を利用することで、パラメータ推定が出来ます。

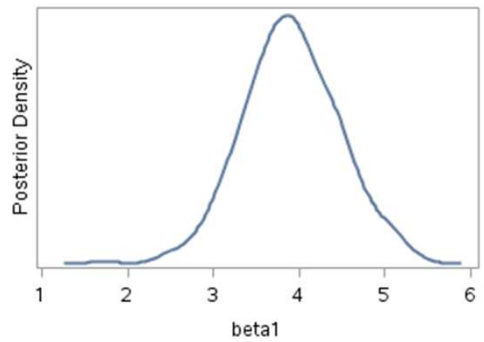
MCMC（マルコフ連鎖モンテカルロ法）とは

- 特定の確率分布のサンプリングを行うシミュレーション方法
- 様々な事前分布や尤度関数の形に対応して、事後分布が得られる
- 具体的な方法には、M-H法・ギッブスサンプリング・ハミルトニアンモンテカルロ法などが存在

事後分布のシミュレーション



事後分布の推定



線形ベイズ回帰の例①
～サンプルデータ作成

SASコード

```
data Class;
   input Name $ Height Weight @@;
   datalines;
Alfred  69.0 112.5   Alice  56.5  84.0   Barbara 65.3
98.0
Carol   62.8 102.5   Henry  63.5 102.5   James   57.3
83.0
Jane    59.8  84.5   Janet  62.5 112.5   Jeffrey 62.5  84.0
John    59.0  99.5   Joyce  51.3  50.5   Judy    64.3  90.0
Louise  56.3  77.0   Mary   66.5 112.0   Philip  72.0
150.0
Robert  64.8 128.0   Ronald 67.0 133.0   Thomas  57.5
85.0
William 66.5 112.0
;
```

作成データセット

Obs	Name	Height	Weight
1	Alfred	69.0	112.5
2	Alice	56.5	84.0
3	Barbara	65.3	98.0
4	Carol	62.8	102.5
5	Henry	63.5	102.5
6	James	57.3	83.0
7	Jane	59.8	84.5
8	Janet	62.5	112.5
9	Jeffrey	62.5	84.0
10	John	59.0	99.5
11	Joyce	51.3	50.5
12	Judy	64.3	90.0
13	Louise	56.3	77.0
14	Mary	66.5	112.0
15	Philip	72.0	150.0
16	Robert	64.8	128.0
17	Ronald	67.0	133.0
18	Thomas	57.5	85.0
19	William	66.5	112.0

線形ベイズ回帰の例②
～モデルの検討

単回帰モデルにより、身長から体重を予測

$$\boxed{Weight_i = \beta_0 + \beta_1 Height_i + \varepsilon_i}$$

目的変数の分布（≒尤度）を正規分布と考える

$$Weight_i \sim Nomral(\beta_0 + \beta_1 Height_i, \sigma^2)$$

推定パラメータ	事後分布	事前分布
β_0	正規分布	$Normal(0, 100^2)$
β_1	正規分布	$Normal(0, 100^2)$
σ^2	逆ガンマ関数	$igamma(\frac{3}{10}, \frac{10}{3})$

事前に情報がないため、平均をゼロとし、分散は幅広く取る

線形ベイズ回帰の例③
～MCMCのモデルコード

SASコード

```
ods graphics on;
proc mcmc data=class outpost=classout nmc=10000 thin=5 seed=246810;
  parms beta0 0 beta1 0;
  parms sigma2 1;
  prior beta0 beta1 ~ normal(mean = 0, var = 1e6);
  prior sigma2 ~ igamma(shape = 3/10, scale = 10/3);
  mu = beta0 + beta1*height;
  model weight ~ n(mu, var = sigma2);
run;
ods graphics off;
```

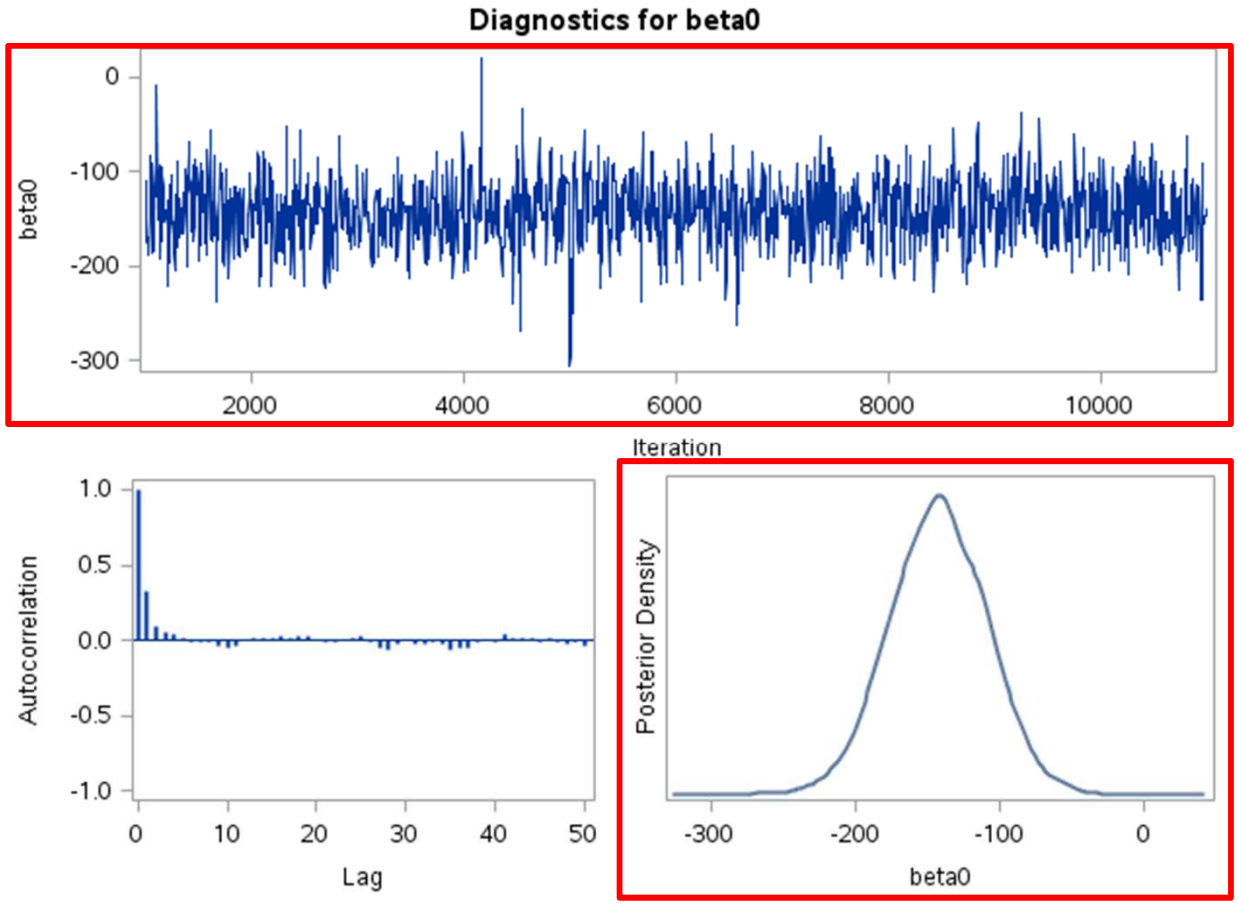
推定パラメータの指定

事前分布の指定

尤度・モデルの定義

線形ベイズ回帰の例④
～推定結果（beta0のみ）

MCMCによるパラメータ推定結果

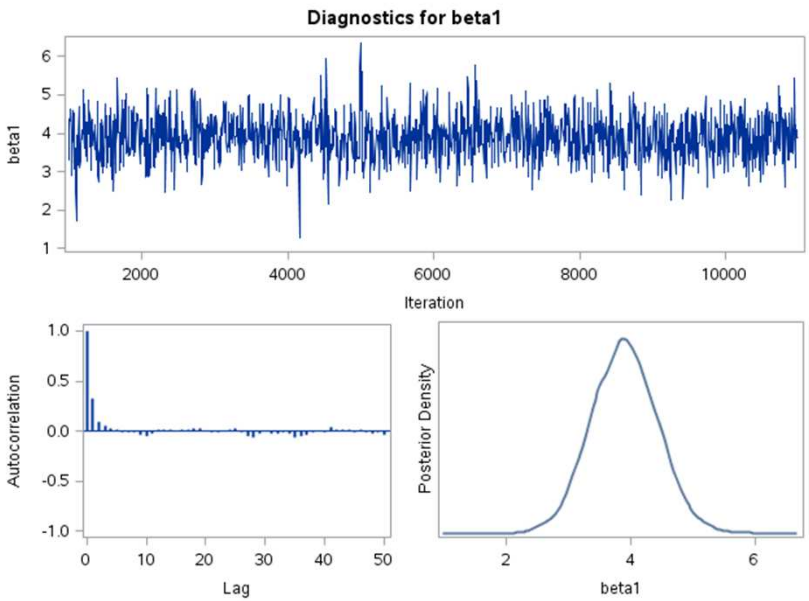


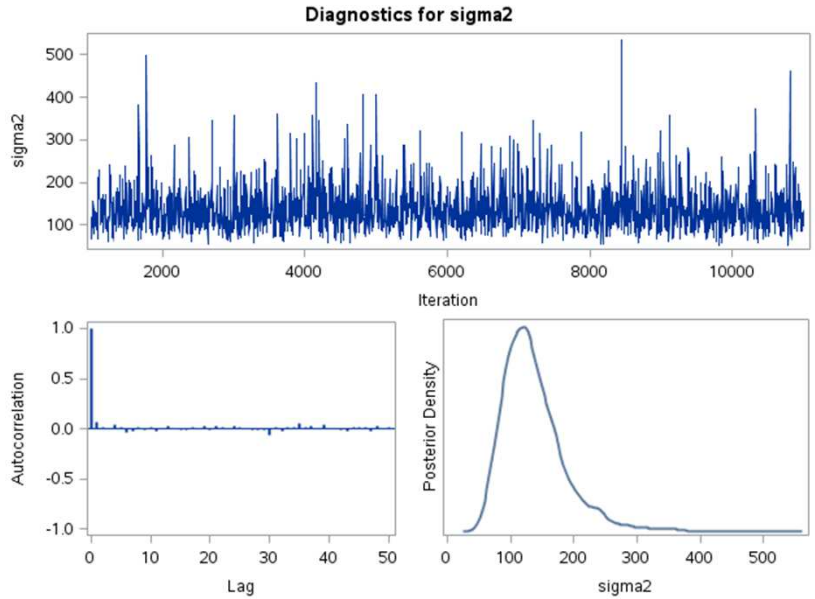
MCMCによる
サンプリング結果

推定パラメータの
事後分布

線形ベイズ回帰の例④
～推定結果（beta1, sigma2）

MCMCによるパラメータ推定結果





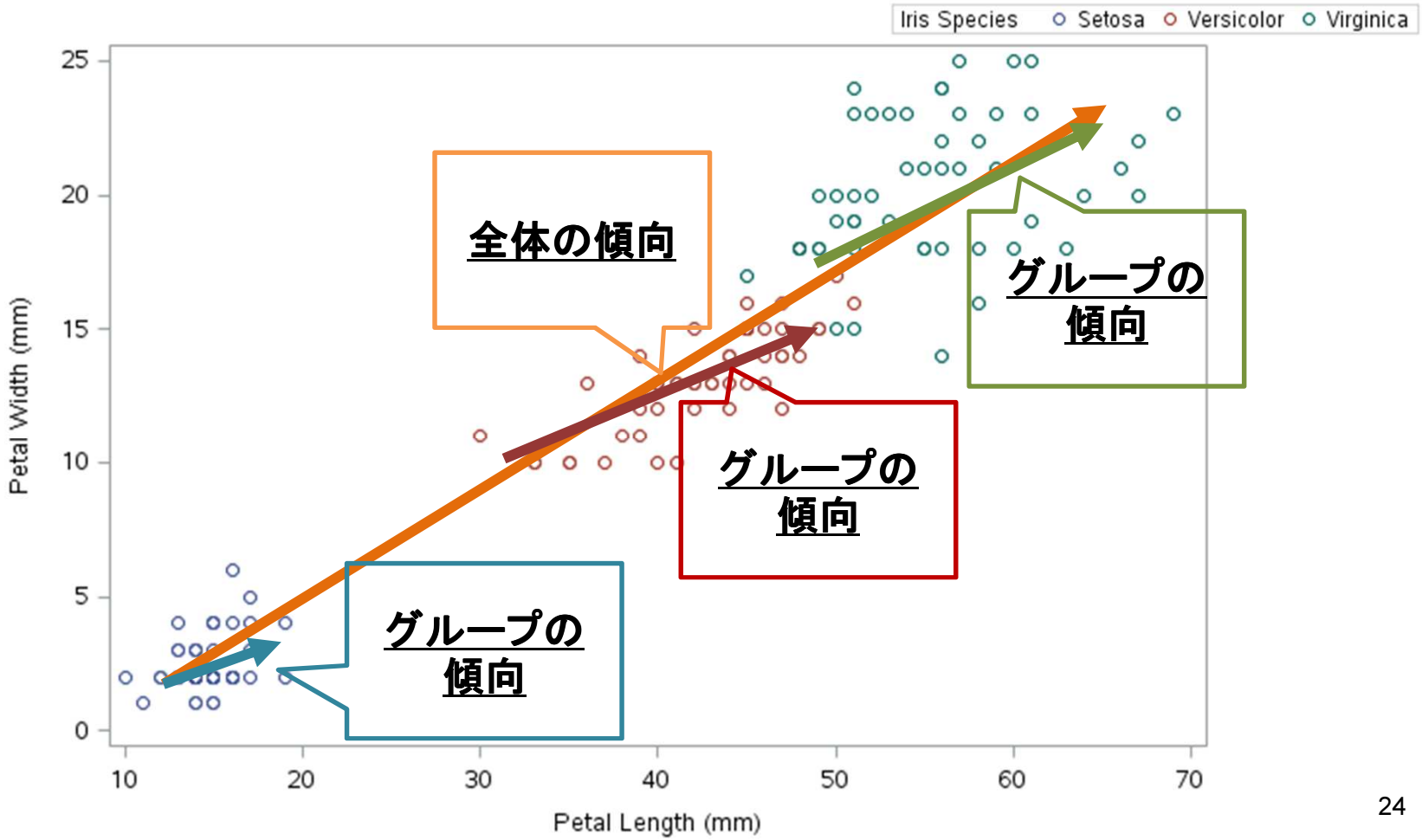
推定パラメータの分布を表現することが可能となる

アジェンダ

1. マーケティングにおけるベイジアンモデリングの必要性
2. ベイジアンモデリングの基礎
3. 階層ベイズを用いた「個」の理解
4. 状態空間モデルを用いた動的表現
5. まとめ

画一的に分析しているデータにも、グループによる違いがあります。

アヤメの花びらの幅と長さの関係



画一的に表現していたモデルにグループ差を取りこむことで、
より複雑な構造をモデリング出来ます。

i:個人
g:所属グループ

$$y_i = \beta_0 + \beta_1 x_i + e_i$$

グループ差の取りこみ

$$y_i = \beta_0 + \beta_1 x_i + \boxed{\gamma_g} + e_i$$

グループ間の違いを抽出するため、
平均ゼロの正規分布を事前分布とする

➔ 事前分布の分散はどうするか？

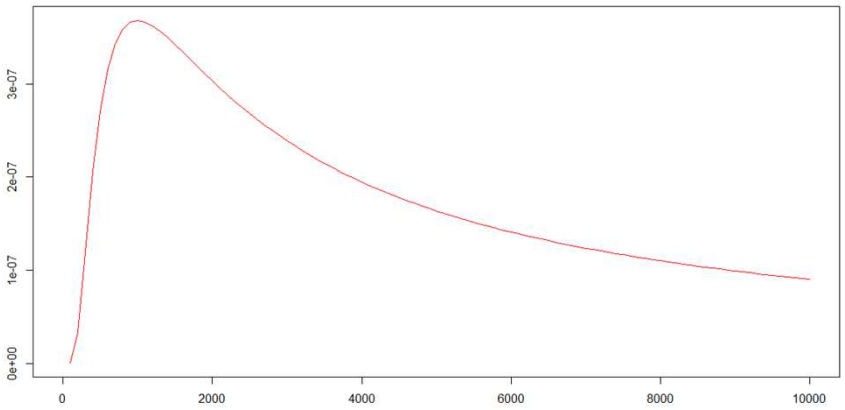
事前分布のパラメータを説明するハイパーパラメータを指定することで、客観的なモデル構造を表現出来ます。

β_0: 正規分布

β_1: 正規分布

$$y_i = \beta_0 + \beta_1 x_i + \boxed{\gamma_g} + e_i$$

σの事前分布



$$\gamma_g \sim Normal(0, \boxed{\sigma^2})$$

$$\sigma \sim igamma(shape, scale)$$

グループ差を表すγの分散に関して、更に事前分布を設定し階層化することで、何も情報がない事実をモデルで表現する

階層ベイズの例①
～サンプルデータ作成

SASコード

```
data iris;
   set sashelp.iris;
   if Species = 'Setosa' THEN g = 1;
   else if Species = 'Versicolor' THEN g = 2;
   else g = 3;
run;
```

作成データセット

Obs	Species	SepalLength	SepalWidth	PetalLength	PetalWidth	g
1	Setosa	50	33	14	2	1
2	Setosa	46	34	14	3	1
3	Setosa	46	36	10	2	1
4	Setosa	51	33	17	5	1
5	Setosa	55	35	13	2	1
6	Setosa	48	31	16	2	1
7	Setosa	52	34	14	2	1
8	Setosa	49	36	14	1	1
9	Setosa	44	32	13	2	1
10	Setosa	50	35	16	6	1
11	Setosa	44	30	13	2	1
12	Setosa	47	32	16	2	1
13	Setosa	48	30	14	3	1
14	Setosa	51	38	16	2	1
15	Setosa	48	34	19	2	1
16	Setosa	50	30	16	2	1
17	Setosa	50	32	12	2	1
18	Setosa	43	30	11	1	1
19	Setosa	58	40	12	2	1
20	Setosa	51	38	19	4	1
21	Setosa	49	30	14	2	1
22	Setosa	51	35	14	2	1
23	Setosa	50	34	16	4	1
24	Setosa	46	32	14	2	1
25	Setosa	57	44	15	4	1

…

階層ベイズの例②
～モデルの検討

階層ベイズモデルにより、花びらの長さから花びらの幅を予測

$$\boxed{petalwidth_i = \beta_0 + \beta_1 petallength_i + \gamma_g + e_i}$$

目的変数の分布（≒尤度）を正規分布と考える

$$petalwidth_i \sim Nomral(\beta_0 + \beta_1 petallength_i + \gamma_g + e_i, s^2)$$

推定パラメータ	事後分布	事前分布
β_0	正規分布	$Normal(0, 100^2)$
β_1	正規分布	$Normal(0, 100^2)$
γ_g	正規分布	$Normal(0, \sigma^2)$
σ^2	逆ガンマ関数	$igamma(0.001, 1000)$
s^2	逆ガンマ関数	$igamma(0.001, 1000)$

階層構造

階層ベイズの例③
～MCMCのモデルコード

SASコード

```
ods graphics on;
proc mcmc data=iris outpost=postout thin=10 nmc=50000 seed=1234;
  array gamma[3];

  parms b0 0 b1 0 gamma: 0;
  parms s2 1 ;  parms s2g 1;

  prior b: ~ normal(0, var = 10000);
  prior gamma: ~ normal(0, var = s2g);
  prior s2: ~ igamma(0.001, scale = 1000);

  mu = b0 + b1 * petallength + gamma[g];
  model petalwidth ~ normal(mu, var = s2);
run;
ods graphics off;
```

推定パラメータの指定

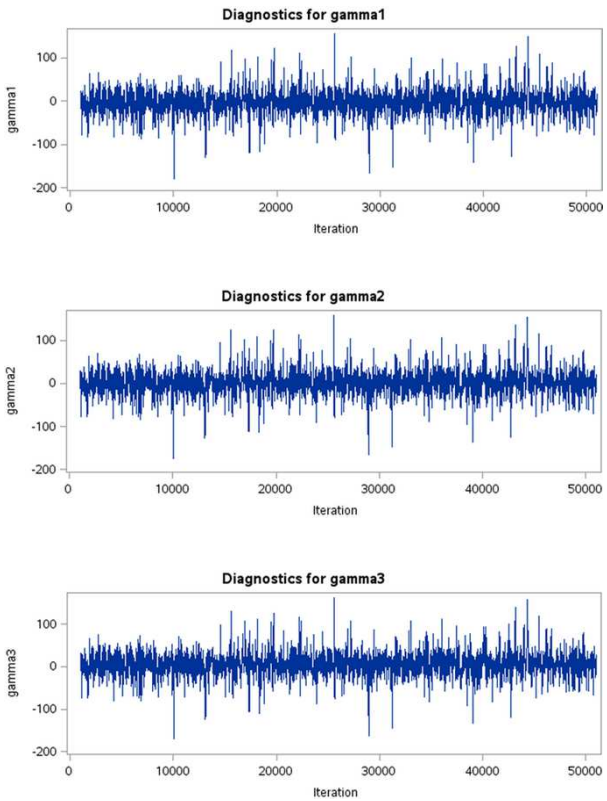
事前分布の指定

尤度・モデルの定義

階層ベイズの例④
～推定結果

MCMCによるパラメータ推定結果

Posterior Summaries and Intervals					
Parameter	N	Mean	Standard Deviation	95% HPD Interval	
b0	5000	2.7771	27.5225	-51.7588	60.9259
b1	5000	0.2317	0.0778	0.0796	0.3831
gamma1	5000	-3.6991	27.4220	-61.6121	50.6615
gamma2	5000	0.6268	27.4341	-53.7505	58.9006
gamma3	5000	4.6134	27.4684	-50.7204	62.2599
s2	5000	17.1615	2.0115	13.4979	21.2488
s2g	5000	4661.4	31927.1	121.2	12970.1



グループ差が推計出来ている

階層ベイズの例⑤
～線形回帰との結果の比較

	線形回帰	階層ベイズ
モデル	$y_i = \beta_0 + \beta_1 x_i + e_i$	$y_i = \beta_0 + \beta_1 x_i + \gamma_g + e_i$
推定方法	最小二乗法	MCMC
決定係数	0.927	0.946

階層ベイズにより、推定精度が向上している

アジェンダ

1. マーケティングにおけるベイジアンモデリングの必要性
2. ベイジアンモデリングの基礎
3. 階層ベイズを用いた「個」の理解
4. 状態空間モデルを用いた動的表現
5. まとめ

状態空間モデルを用いることで、時点の変化に関わる影響と特定の時点で受ける要因の影響をモデル化出来ます。

時点の変化に関わる影響

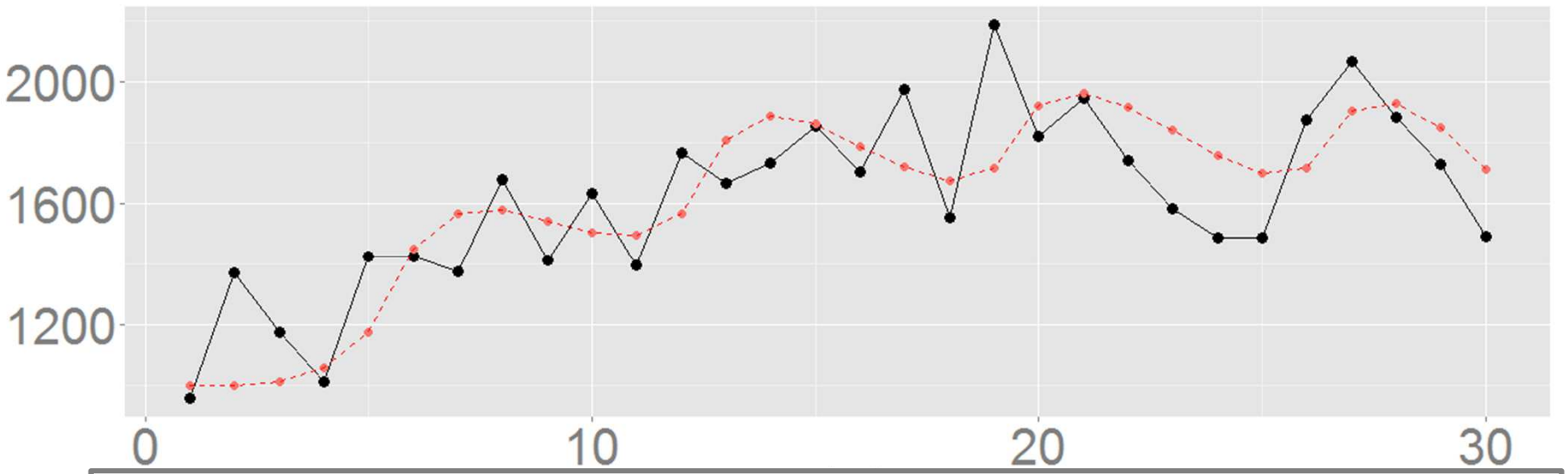
- ✓ 週末になったので、来客数が増加した
- ✓ 夏になったので、来客数が増加した

特定時点での要因による影響

- ✓ 割引を実施したので、来客数が増加した
- ✓ たまたま雨が降ったので、来客数が増加した

観測値から特定時点での要因を排除した“トレンド”を抽出することで時系列の分析を正確に行うことが可能になります

あるレストランの日別来店者数



観測値：実際に取得されたデータ

トレンド：観測値から特定時点の要因を排除した潜在的な値

- トレンドの変化を抽出することで時点の変化に関わる影響を分析
- トレンドと観測値を用いて、特定時点での要因による影響を分析

t期前のデータをモデルに組み込むことで、時点の変化に関わる影響をモデルに組み込むことが出来ます。

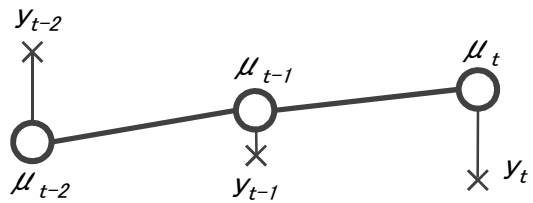
- 以下の3つの条件を満たす形で変化するデータについて考える
  1. 時点の変化に伴う変化量は、1時点前の変化量と同様となる傾向がある
  2. 時点の変化に伴って一定のノイズが発生する
  3. 実際の観測に伴って一定のノイズが発生する

時点の変化に関わる影響

- t時点における変動している変数をμ_tと置く
- （仮定1から）μ_tは、$\mu_t - \mu_{t-1} = \mu_{t-1} - \mu_{t-2}$を満たす⇒$\mu_t = 2\mu_{t-1} - \mu_{t-2}$
- （仮定2から）ノイズをvtのN(0, σ_1^2)に従う正規分布とする（更新ノイズ）

特定時点での要因による影響

- t時点における観測値をy_tと置く
- （仮定3から）ノイズをw_tのN(0, σ_2^2)に従う正規分布とする（観測ノイズ）



$$\boxed{\mu_t} = 2\mu_{t-1} - \mu_{t-2} + v_t$$

トレンド

$$\boxed{y_t = \mu_t + w_t}$$

y_tはy_{t-1}からの影響は受けず、μ_tからの影響のみ受ける

ベクトル・行列を用いることで、時系列データを状態空間モデルとして表現することが可能になります。

時点の変化に関わる影響	特定時点での要因による影響

$$\mu_t = 2\mu_{t-1} - \mu_{t-2} + v_t \qquad y_t = \mu_t + w_t$$

ベクトル・行列で表現

$$\begin{pmatrix} \mu_t \\ \mu_{t-1} \end{pmatrix} = \begin{pmatrix} 2 & -1 \\ 1 & 0 \end{pmatrix} \begin{pmatrix} \mu_{t-1} \\ \mu_{t-2} \end{pmatrix} + \begin{pmatrix} 1 \\ 0 \end{pmatrix} \cdot v_t \qquad y_t = \begin{pmatrix} 1 & 0 \end{pmatrix} \begin{pmatrix} \mu_t \\ \mu_{t-1} \end{pmatrix} + w_t$$

$$\boxed{\boldsymbol{x}_t = \begin{pmatrix} \mu_t \\ \mu_{t-1} \end{pmatrix}} \qquad \boldsymbol{F}_t = \begin{pmatrix} 2 & -1 \\ 1 & 0 \end{pmatrix} \qquad \boldsymbol{G}_t = \begin{pmatrix} 1 \\ 0 \end{pmatrix} \qquad \boldsymbol{H}_t = \begin{pmatrix} 1 & 0 \end{pmatrix}$$

状態ベクトル

一般化

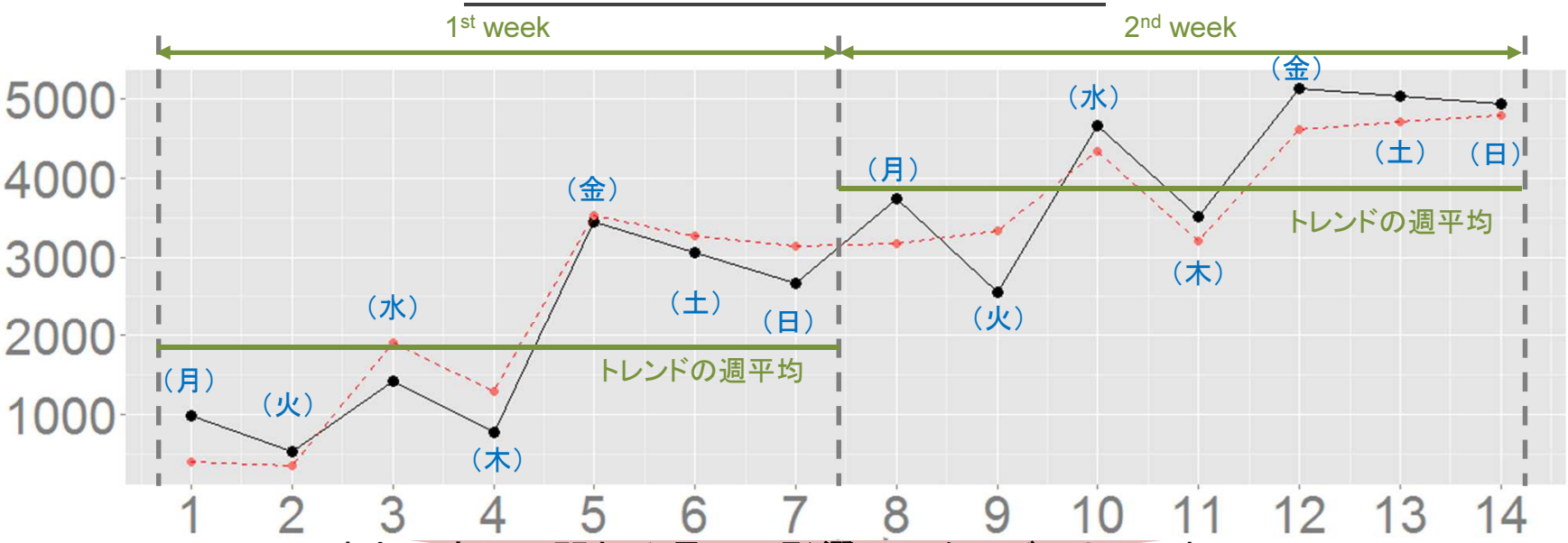
線形ガウス状態空間モデル

$$\boldsymbol{x}_t = \boldsymbol{F}_t \boldsymbol{x}_{t-1} + \boldsymbol{G}_t \boldsymbol{v}_t \qquad \boldsymbol{y}_t = \boldsymbol{H}_t \boldsymbol{x}_t + \boldsymbol{w}_t$$

システムモデル　　観測モデル

曜日による影響などの周期変動は周期内での和がゼロとなる成分を考えます。

あるレストランの日別来店者数



時点の変化に関わる**曜日の影響のみ**をモデルとして表現

$$s_t = -\sum_{i=1}^{6} s_{t-i} + v_t \qquad v_t \sim N(0, \alpha_s^2 \sigma^2)$$

時点の変化の関係をシステムモデルに組み込んでいく

周期成分をシステムモデルに組み込むことで、状態空間モデルに取り込むことが可能になります。

時点の変化に関わる影響

$$\boldsymbol{x}_t = \boldsymbol{F}_t \boldsymbol{x}_{t-1} + \boldsymbol{G}_t \boldsymbol{v}_t$$

$$\begin{pmatrix} \mu_t \\ \mu_{t-1} \end{pmatrix} = \begin{pmatrix} 2 & -1 \\ 1 & 0 \end{pmatrix} \begin{pmatrix} \mu_{t-1} \\ \mu_{t-2} \end{pmatrix} + \begin{pmatrix} 1 \\ 0 \end{pmatrix} \cdot v_t$$

$$s_t = -\sum_{i=1}^{6} s_{t-i} + v_t$$

曜日効果の取込

$$\begin{pmatrix} \mu_t \\ \mu_{t-1} \\ \boxed{s_t} \\ s_{t-1} \\ s_{t-2} \\ s_{t-3} \\ s_{t-4} \\ s_{t-5} \end{pmatrix} = \begin{pmatrix} 2 & -1 & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & -1 & -1 & -1 & -1 & -1 & -1 \\ 0 & 0 & -1 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & -1 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & -1 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & -1 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & -1 & 0 \end{pmatrix} \begin{pmatrix} \mu_t \\ \mu_{t-1} \\ s_{t-1} \\ s_{t-2} \\ s_{t-3} \\ s_{t-4} \\ s_{t-5} \\ s_{t-6} \end{pmatrix}$$

特定時点での要因による影響

$$\boldsymbol{y}_t = \boldsymbol{H}_t \boldsymbol{x}_t + \boldsymbol{w}_t$$

$$y_t = \begin{pmatrix} 1 & 0 \end{pmatrix} \begin{pmatrix} \mu_t \\ \mu_{t-1} \end{pmatrix} + w_t$$

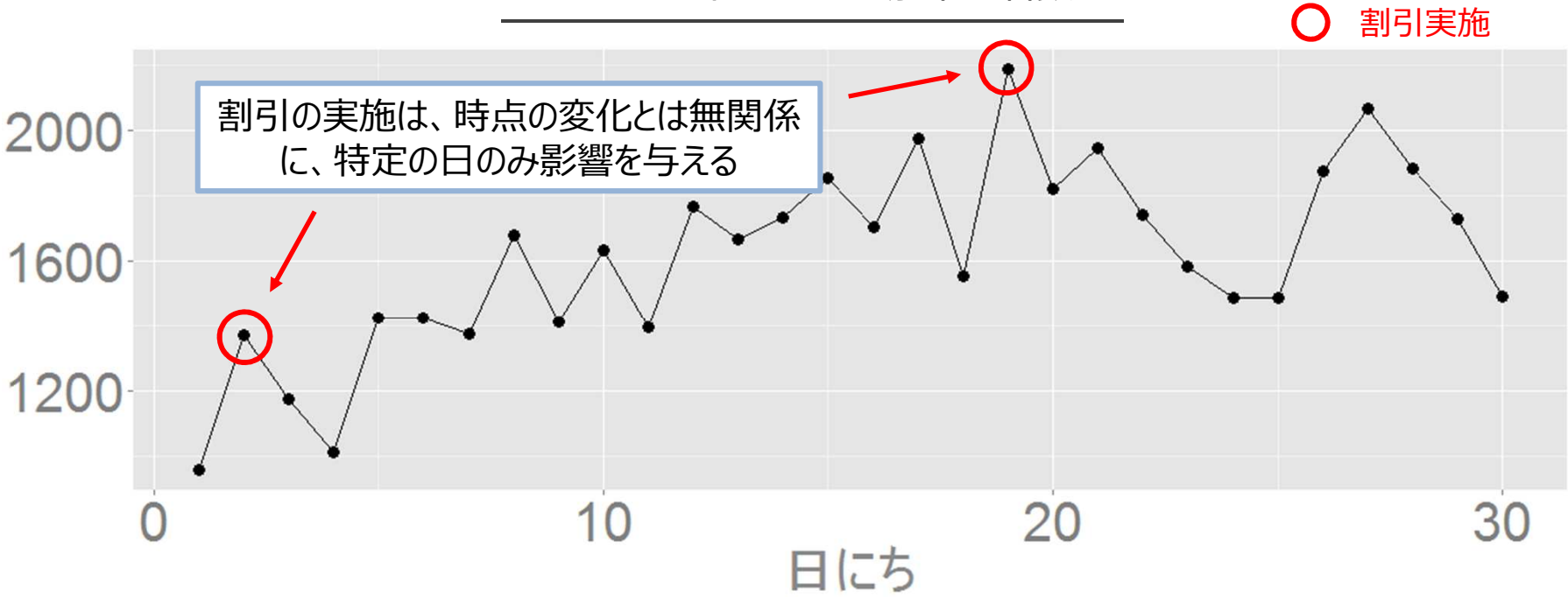
$$y_t = \begin{pmatrix} 1 & 0 & \boxed{1} & 0 & 0 & 0 & 0 & 0 \end{pmatrix} \begin{pmatrix} \mu_t \\ \mu_{t-1} \\ \boxed{s_t} \\ s_{t-1} \\ s_{t-2} \\ s_{t-3} \\ s_{t-4} \\ s_{t-5} \end{pmatrix}$$

- 観測モデルはその時点の成分以外は排除する

状態ベクトルを拡張することで、曜日効果も表現できる

割引などのイベント効果は、特定時点にのみ影響を与えるものだと考えられます。

あるレストランの日別来店者数



時点の変化に関わらず一定の値を持つ成分を表現

$$\gamma_t = \gamma_{t-1} + v_t \qquad v_t \sim N(0, \alpha_\gamma^2 \sigma^2)$$

一定の値を持つ成分をシステムモデルに組み込み、割引効果を測定

一定の値を持つ成分をシステムモデルに組み込むことで、特定時点の要因を状態空間モデルに取り込むことが出来ます。

時点の変化に関わる影響

$$\boldsymbol{x}_t = \boldsymbol{F}_t \boldsymbol{x}_{t-1} + \boldsymbol{G}_t \boldsymbol{v}_t$$

$$\begin{pmatrix} \mu_t \\ \mu_{t-1} \\ s_t \\ s_{t-1} \\ s_{t-2} \\ s_{t-3} \\ s_{t-4} \\ s_{t-5} \end{pmatrix} = \begin{pmatrix} 2 & -1 & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & -1 & -1 & -1 & -1 & -1 & -1 \\ 0 & 0 & -1 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & -1 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & -1 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & -1 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & -1 & 0 \end{pmatrix} \begin{pmatrix} \mu_t \\ \mu_{t-1} \\ s_{t-1} \\ s_{t-2} \\ s_{t-3} \\ s_{t-4} \\ s_{t-5} \\ s_{t-6} \end{pmatrix}$$

$$\gamma_t = \gamma_{t-1} + v_t$$

時点によらない項の組込

$$\begin{pmatrix} \mu_t \\ \mu_{t-1} \\ s_t \\ s_{t-1} \\ s_{t-2} \\ s_{t-3} \\ s_{t-4} \\ s_{t-5} \\ \boxed{\gamma_t} \end{pmatrix} = \begin{pmatrix} 2 & -1 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & -1 & -1 & -1 & -1 & -1 & -1 & 0 \\ 0 & 0 & -1 & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & -1 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & -1 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & -1 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & -1 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & \boxed{1} \end{pmatrix} \begin{pmatrix} \mu_t \\ \mu_{t-1} \\ s_{t-1} \\ s_{t-2} \\ s_{t-3} \\ s_{t-4} \\ s_{t-5} \\ s_{t-6} \\ \boxed{\gamma_{t-1}} \end{pmatrix}$$

特定時点での要因による影響

$$\boldsymbol{y}_t = \boldsymbol{H}_t \boldsymbol{x}_t + \boldsymbol{w}_t$$

$$y_t = \begin{pmatrix} 1 & 0 & 1 & 0 & 0 & 0 & 0 & 0 \end{pmatrix} \begin{pmatrix} \mu_t \\ \mu_{t-1} \\ s_t \\ s_{t-1} \\ s_{t-2} \\ s_{t-3} \\ s_{t-4} \\ s_{t-5} \end{pmatrix}$$

f_tはt時点の値引きの有無を表すフラグ

観測モデルへの反映

$$y_t = \begin{pmatrix} 1 & 0 & 1 & 0 & 0 & 0 & 0 & 0 & \boxed{f_t} \end{pmatrix} \begin{pmatrix} \mu_t \\ \mu_{t-1} \\ s_t \\ s_{t-1} \\ s_{t-2} \\ s_{t-3} \\ s_{t-4} \\ s_{t-5} \\ \boxed{\gamma_t} \end{pmatrix}$$

時点によらない成分を組込むことで、値引き効果も表現できる

SASを用いた状態空間モデルは、STATESPACEプロシジャ・SSMプロシジャ・MCMCプロシジャによって実装出来ます。

	STATESPACE プロシジャ	SSMプロシジャ	MCMCプロシジャ
推定方法	カルマン フィルタ	カルマン フィルタ	MCMC
パネルデータ	×	○	○
一般化状態空間モデル	×	×	○
実行速度	中	中	低

アジェンダ

1. マーケティングにおけるベイジアンモデリングの必要性
2. ベイジアンモデリングの基礎
3. 階層ベイズを用いた「個」の理解
4. 状態空間モデルを用いた動的表現
5. まとめ

まとめ

- ベイジアンモデリングを活用することで、柔軟なモデリングが可能になります。
- ベイジアンモデリングのパラメータ推定の方法として、MCMCが利用されます。
- 状態空間モデルを用いることで、時系列データもベイジアンモデリングの1つとして表現できます。
- 状態空間モデルというダイナミクスの中で、階層ベイズ等を用いることで、自由なモデリングが実装出来ます。
- ベイジアンモデリングを用いて、需要予測・キャンペーン効果測定・消費者行動分析等を行いマイクロマーケティングを実行していくことで、今後の更なるビジネス発展が見込めます。

参考文献

- 照井 伸彦 (2008)「ベイズモデリングによるマーケティング分析 」東京電機大学出版局
- 樋口 知之 (2011)「予測にいかす統計モデリングの基本」講談社
- 佐藤 忠, 樋口 知之 (2013)「ビッグデータ時代のマーケティング」講談社
- 久保 拓弥 (2012)「データ解析のための統計モデリング入門」講談社
- 豊田 秀樹 (2015)「基礎からのベイズ統計」朝倉書店
- MatsuuraKentaro (ブログ)「Small Data Scientist Memorandum」