

ベイズ統計学の基礎と 官能評価

予稿ダイジェスト版

日本女子大学
岡本安晴

1

官能評価データ： ランダムな変動（ゆらぎ）

↓

確率モデルによる表現

↓

ベイズ統計学（分析）

ベイズ分析例

1. 回帰分析（中心化、標準得点）
2. 匂いの順応曲線（一般線形モデル、リンク関数）
3. 温度と甘味度（ダミー変数）
4. チョコレートの選択頻度（階層モデル、ロジット変換）
5. 信号検出理論（階層モデル）
6. 感覚の弁別
7. 感覚の強さ
8. 3点識別試験ベイズ分析（ d' の回帰分析）

2

関連ウェブサイト

1. 回帰分析（中心化、標準得点）
<http://y-okamoto-psy1949.la.coocan.jp/Python/tips/SimpleRegression/>
<http://y-okamoto-psy1949.la.coocan.jp/Python/tips/MultipleRegression/>
2. 匂いの順応曲線（一般線形モデル、リンク関数）
<http://y-okamoto-psy1949.la.coocan.jp/Python/tips/RegCurveOdor/>
3. 温度と甘味度（ダミー変数）
<http://y-okamoto-psy1949.la.coocan.jp/Python/tips/SweetnessTemperature/>
4. チョコレートの選択頻度（階層モデル、ロジット変換）
<http://y-okamoto-psy1949.la.coocan.jp/Python/tips/ChocolatePreference/>
<http://y-okamoto-psy1949.la.coocan.jp/Python/tips/ChocolatePersistence/>
5. 信号検出理論（階層モデル）
http://y-okamoto-psy1949.la.coocan.jp/Python/tips/SDT_hyper/
6. 感覚の弁別
http://y-okamoto-psy1949.la.coocan.jp/Python/tips/PSE_JND_Bysn/
7. 感覚の強さ
<http://y-okamoto-psy1949.la.coocan.jp/Python/tips/CategoricalScaling/>
8. 3点識別試験ベイズ分析（ d' の回帰分析）
<http://y-okamoto-psy1949.la.coocan.jp/Python/misc/TriTestDRgrssn/>

3

最尤法とベイズ分析

$P(X|\theta)$ と $P(\theta, X) = P(\theta)P(X|\theta)$ ：「木を見る」と「森を見る」

製品の品質	検査	
	合格	不合格
問題なし： $1-p$	1	0
問題あり： p	$1-s$	s

最尤法：

合格 $P(\text{合格}|\text{問題なし}) = 1 > 1-s = P(\text{合格}|\text{問題あり})$
➡ 問題なし

不合格 $P(\text{不合格}|\text{問題なし}) = 0 < s = P(\text{不合格}|\text{問題あり})$
➡ 問題あり

4

ベイズ分析

製品の品質	検査	
	合格	不合格
問題なし: $1-p$	1	0
問題あり: p	$1-s$	s

不合格のとき

$$P(\text{問題あり}|\text{不合格}) = \frac{P(\text{不合格, 問題あり})}{P(\text{不合格})} = \frac{ps}{ps} = 1$$

$$P(\text{問題なし}|\text{不合格}) = \frac{P(\text{不合格, 問題なし})}{P(\text{不合格})} = \frac{0}{ps} = 0$$

5

$P(*|\text{合格})$ のグラフ

$p = P(\text{問題あり})$

ベイズ分析では、製品における「問題あり」の全般的な比率 p を考慮した「森を見る」判断ができる

6

ベイズ分析の基礎

パラメータ θ とデータ D の同時分布 (joint probability distribution)

$$P(\theta, D) = P(\theta)P(D; \theta)$$

$P(\theta)$ 事前分布 (prior distribution)

$P(D; \theta)$ データ分布 (data distribution)
尤度関数 (Dを固定して θ の関数と見る)

$$P(D|\theta) = \frac{P(D, \theta)}{P(\theta)} = \frac{P(\theta)P(D; \theta)}{P(\theta)} = P(D; \theta)$$

Gelman, A., Carlin, J. B., Stern, H. S., Dunson, D. B., Vehtari, A., & Rubin, D. B. (2014). *Bayesian Data Analysis, third edition*. CRC Press. pp. 6-8.

岡本安晴 (2018) 「心理学における量的研究法・質的研究法の統合」
日本情報教育学会誌 Vol.1、pp.70-77. https://doi.org/10.24711/jite.1.1_70

7

事後分布の分析

1. 解析的 (数学的) 分析
共役事前分布
2. シミュレーション (乱数) による分析
Markov chain Monte Carlo (MCMC) サンプリング

Markov chain

$X_t \rightarrow P(X_{t+1}|X_t) \rightarrow X_{t+1}$

8

事後分布をシミュレーションで求めるソフトStan

<https://mc-stan.org/users/documentation/>

岡本安晴「いまさら聞けないPythonでデータ分析」丸善出版、第10～12章

サンプリング法としてHMCが使える

HMC: Hamiltonian Monte Carlo
Hybrid Monte Carlo

NUTS: No-U-turn sampler

HMC can be more effective than the various samplers in JAGS and BUGS

Kruschke, J. K. (2015) Doing Bayesian Data Analysis, 2nd ed.
Academic Press.

9

Stanのインストール

下記ウェブサイトに解説を用意している

<http://y-okamoto-psy1949.la.coocan.jp/bayesian/pystan/>

<http://y-okamoto-psy1949.la.coocan.jp/Python/tips/UsingPython/>

Windowsの場合の注意事項

Samplingの実行において、パラメータn_jobs=1を設定する。

現在(2022.07.29)、実行時にVSCコンパイラが選ばれてしまうという不具合がある。これは、AnacondaにおいてPython 3.7の仮想環境を用意し、pystan 2.17.1.0をインストールすれば回避できる。

WindowsにubuntuをインストールしてStanを使う例

http://y-okamoto-psy1949.la.coocan.jp/Python/tips/ubuntu_wsl/

10

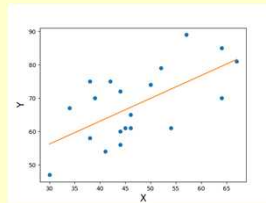
単回帰モデル

表 人工データ (X_i, Y_i)

(41, 54)	(45, 61)	(44, 56)	(34, 67)	(46, 65)
(39, 70)	(30, 47)	(64, 85)	(54, 61)	(44, 72)
(38, 58)	(46, 61)	(64, 70)	(42, 75)	(50, 74)
(44, 60)	(67, 81)	(52, 79)	(57, 89)	(38, 75)

$$Y_i = a + bX_i + e_i$$

$$e_i \sim \text{Normal}(0, \sigma^2)$$



```
data {
  int N;
  real X[N];
  real Y[N];
}
parameters {
  real a;
  real b;
  real<lower = 0> sgm;
}
model {
  a ~ normal(0, 100);
  b ~ normal(0, 100);
  sgm ~ exponential(0.001);
  for (i in 1:N)
    Y[i] ~ normal(a + b*X[i], sgm);
}
```

11

重回帰モデル

表 人工データ (Y, X_1, X_2)

(924,202,60)	(711,114,46)	(943,173,65)	(906,256,60)	(844,250,48)
(611,116,43)	(805,165,57)	(743,220,43)	(893,217,59)	(905,201,55)
(958,182,64)	(708,215,33)	(694,168,43)	(1109,282,73)	(911,220,58)
(825,158,62)	(753,183,45)	(804,235,52)	(853,217,50)	(707,268,24)

$$Y = b_0 + b_1X_1 + b_2X_2 + \varepsilon,$$

$$\varepsilon \sim \text{Normal}(0, \sigma^2)$$

$$b_0 \sim \text{Normal}(0, 1000^2), \quad b_1 \sim \text{Normal}(0, 1000^2)$$

$$b_2 \sim \text{Normal}(0, 1000^2), \quad \sigma \sim \text{Exponential}(0.01)$$

12

```

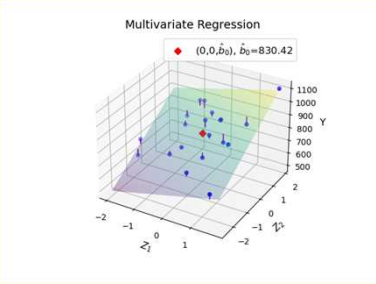
data {
  int N;
  real Y[N];
  real X1[N];
  real X2[N];
}
parameters {
  real b0;
  real b1;
  real b2;
  real<lower = 0.0> sgm;
}
model {
  b0 ~ normal(0, 1000);
  b1 ~ normal(0, 1000);
  b2 ~ normal(0, 1000);
  sgm ~ exponential(0.01);
  for (i in 1:N)
    Y[i] ~ normal(b0 + b1 * X1[i] + b2 * X2[i], sgm);
}
    
```

13

標準化後の変量をZで表すと、次式で与えられる。

$$Z = \frac{X - \bar{X}}{std(X)}$$

\bar{X} は平均値、 $std(X)$ は標準偏差

$$Y = b_0 + b_1 Z_1 + b_2 Z_2 + \varepsilon$$


14

匂いの順応曲線

(重) 回帰モデルを、従属変数と独立変数の線形結合との一般的な関数関係 (リンク関数: link function) に拡張したモデルは一般線形モデル (generalized linear model) と呼ばれている (Gelman et al., 2014)。リンク関数を g で表すと、一般線形モデルは次の形のものである。

従属変数 = g^{-1} (独立変数の線形結合) + 残差

時間 t 後の匂いの強さ y_t を、次のモデルで分析してみる。

$$y_t = 100 \times \exp(a \times t) + e_t, \quad a < 0$$

$$e_t \sim \text{Normal}(0, (sa + sb \times t)^2),$$

$$sa > 0, sb > 0$$

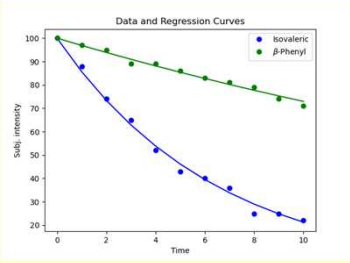
Gelman, A. G., Carlin, J. B., Stern, H. S., Dunson, D. B., Vehtari, A., & Rubin, D. B. (2014). *Bayesian data analysis, third edition*. CRC Press.

15

表のデータは、高島 (1997) の Fig. 3 から読み取ったものである。

高島靖弘 (1997). 香りと官能評価. 日本官能評価学会誌, 1, 10-17.

time	B-phenyl	Isovaleric
0	100	100
1	97	88
2	95	74
3	89	65
4	89	52
5	86	43
6	83	40
7	81	36
8	79	25
9	74	25
10	71	22



16

温度と甘味度：ダミー変数の利用

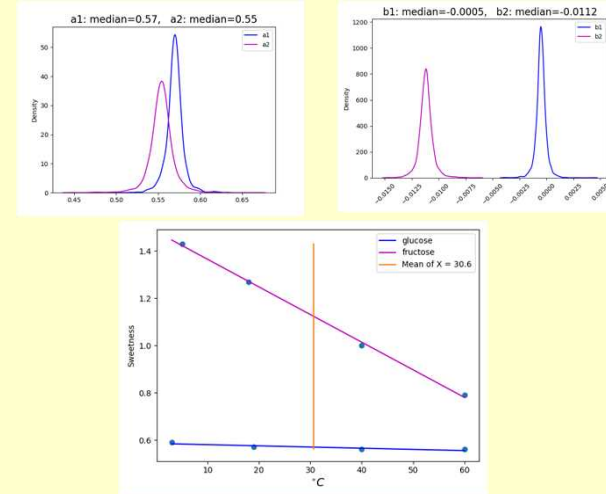
前橋（2011）の第2図からの読み取り

表 温度と甘味度			
フラクトース		グルコース	
温度°C	甘味度	温度°C	甘味度
5	1.43	3	0.59
18	1.27	19	0.57
40	1	40	0.56
60	0.79	60	0.56

$$\text{ダミー変数 } G = \begin{cases} 0 & \text{グルコース} \\ 1 & \text{フラクトース} \end{cases}$$

$$Y = a_1 + a_2G + (b_1 + b_2G)X + e$$

前橋健二（2011）. 甘味の基礎知識. 日本醸造協会誌、106、818-825.



製品の選択頻度の分析（チョコレートの選択）

3種類のチョコレートの嗜好を、ノーズクリップの脱着によりレトロネーザルアロマの影響をコントロールして調べたデータが星野ら（2020）に報告されている。3種のチョコレート、ガーナ、ベネズエラ、エクアドルからの選択を、ノーズクリップなしの条件とノーズクリップあり（レトロネーザルアロマの関与の除外）の2条件で行われた結果が示されている星野ら（2020）のTable 4から、各条件ごとのチョコレートの選択頻度を抜き出して下表にまとめた。

表 チョコレートの選択			
	ガーナ	ベネズエラ	エクアドル
ノーズクリップなし	55	24	16
ノーズクリップあり	38	27	30

星野亜由美、石川華衣、岡田綾、飯田文子（2020）. カカオ豆の香がチョコレート嗜好に及ぼす影響. 日本官能評価学会誌、24、29-42.

3種のチョコレートから1つを選ぶという実験において、それぞれのチョコレートが選ばれた頻度を f_1, f_2, f_3 とおき、それぞれのチョコレートが選ばれる確率を p_1, p_2, p_3 とおけば、多項分布により次式で表される。

$$P(k_1, k_2, k_3) \propto p_1^{k_1} p_2^{k_2} p_3^{k_3}$$

ここで、 $p_1 + p_2 + p_3 = 1$ (1) である。

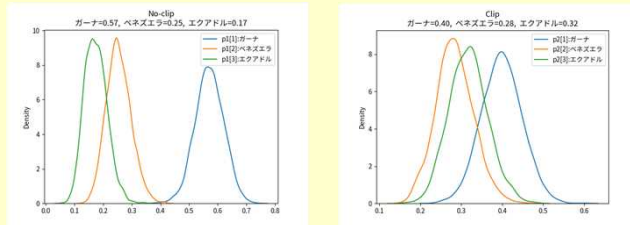
式(1)の条件を満たすパラメータ (p_1, p_2, p_3) は、事前分布としてDirichlet分布を用いると得ることができる。Dirichlet分布は3変量 (p_1, p_2, p_3) の場合、次式で表される。

$$\text{Dirichlet}(p_1, p_2, p_3; \alpha_1, \alpha_2, \alpha_3) \propto p_1^{\alpha_1-1} p_2^{\alpha_2-1} p_3^{\alpha_3-1}$$

ここで、 $p_1 \geq 0, p_2 \geq 0, p_3 \geq 0, p_1 + p_2 + p_3 = 1$
 $\alpha_1 > 0, \alpha_2 > 0, \alpha_3 > 0$

である。

上の表のデータに対して多項分布のパラメータを求めると以下ようになる。



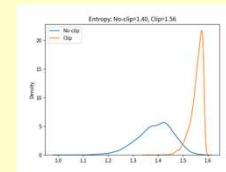
左側の「ノーズクリップなし」ではガーナの事後分布が右側の「ノーズクリップあり」での事後分布に比べて他の2つのチョコレートより右側に移動している様子が見て取れる。

この事後分布の位置の違いをエントロピーで見してみる。

選択確率(p_1, p_2, p_3)に対するエントロピーは次式で与えられる。

$$\text{Entropy} = -p_1 \log_2 p_1 - p_2 \log_2 p_2 - p_3 \log_2 p_3$$

「ノーズクリップなし」条件と「ノーズクリップあり」条件でのエントロピーを求めると、下図のような事後分布になる。「ノーズクリップあり」条件でのエントロピーの方が大きいことが分かる。



「ノーズクリップなし」条件のエントロピーを E_1 で、「ノーズクリップあり」条件のエントロピーを E_2 で表すとき、事後分布において E_1 の方が E_2 より小さい確率を次式

$$P(E_1 < E_2) = \int P(E_2 > E_1 | E_1) dP(E_1)$$

により算出する。

$$P(E_1 < E_2) \approx 0.99$$

星野ら (2020) Table 4によれば、「ノーズクリップなし」条件でチョコレートの選択が「ノーズクリップあり」条件でも同じである対象者は下表に示す人数である。

MN.2	ノーズクリップなし	ノーズクリップありでも同じチョコレートを選択
ガーナ	55	31
ベネズエラ	24	8
エクアドル	16	9

「ノーズクリップなし」条件であるチョコレートが N 人に選ばれ、その N 人の内 k 人が同じチョコレートを「ノーズクリップあり」条件で選んでいるとき、2項分布で

$$P(k) = \binom{N}{k} p_c^k (1 - p_c)^{N-k}$$

と表すことができる。

ここで、 p_c はチョコレート c に対して「ノーズクリップなし」で選ばれたときに「ノーズクリップあり」でも選ばれる確率である。

階層モデル

3種のチョコレートの選択確率は実験で用いられたチョコレートという共通点があり、何らかのまとまりがあると考えられる。いま、選択確率 p_c のロジット変換した値が共通の正規分布に従っていると考える (hierarchical model; Lambert, 2018)。

ロジット変換するのは、確率の変域が0から1の値であり、これらをロジット変換すると値の変域は $-\infty$ から $+\infty$ となり、正規分布が使えるからである。

$$\text{logit}(p_c) \sim \text{Normal}(\mu_h, \sigma_h^2)$$

ここで、

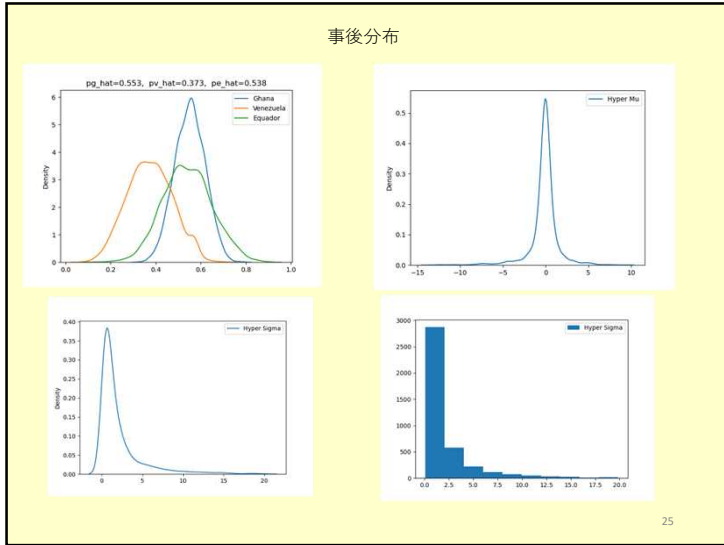
$$\text{logit}(p) = \log\left(\frac{p}{1-p}\right)$$

である。

右辺の正規分布のパラメータ μ_h と σ_h はhyperparameterと呼ばれている (Gelman et al., 2014, p. 108)。これらのhyperparameterの事前分布を平らな事前分布として次式で与える。

$$\mu_h \sim \text{Normal}(0, 100^2), \quad \sigma_h \sim \text{Exponential}(0.01)$$

これらhyperparameterの事前分布はhyperprior分布と呼ばれている (Gelman et al., 2014, p. 108)。



信号検出理論

信号検出理論の実験を4条件、A、B、C、D、で行い、表の結果を得たとする。

	条件A		条件B		条件C		条件D	
	No	Yes	No	Yes	No	yes	No	Yes
ノイズ刺激	7 7	2 3	5 0	5 0	7 7	2 3	6 0	4 0
シグナル刺激	2 3	7 7	2 3	7 7	1 0	9 0	1 1	8 9

各条件においてノイズ刺激が100回、シグナル刺激が100回、ランダムに提示されたときのNo反応（「ノイズ刺激である」という判断）とYes反応（「シグナル刺激である」という判断）の頻度がまとめられている。条件Aを基準にして考えると、条件B]ではシグナル刺激に対する反応の比率は条件Aと同じであるが、ノイズ刺激に対するYes反応が2 3回から5 0回に増えている。条件Cでは、ノイズ刺激に対する判断の比率は条件Aと同じであるが、シグナル刺激に対するYes反応が7 7回から9 0回に増えている。条件Dでは、条件Aと比べて、ノイズ刺激とシグナル刺激いずれに対してもYes反応が増えている。

26

条件A、B、C、Dにおける d' を da 、 db 、 dc 、 dd で表すとき、これらは共通の分布 $Normal(dmu_h, dsigm_h)$ に従うものとする。

条件A、B、C、Dにおける C も、共通の分布 $Normal(Cmu_h, Csgm_h)$ に従うものとする。

Hyperparameter dmu_h と Cmu_h は、hyperprior 分布 $Normal(0, 5^2)$ に従うものとし、 $dsigm_h$ と $Csgm_h$ は、 $Uniform(0, 5)$ に従うものとする。

Stan スクリプトを以下のように作成した。

```

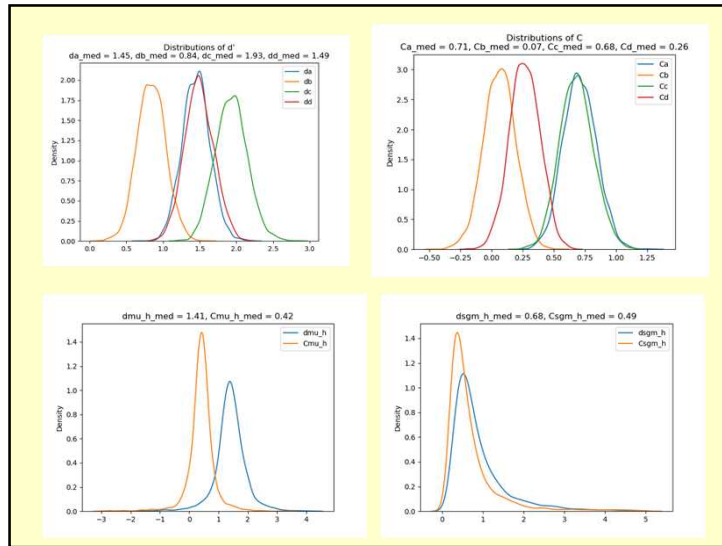
data {
  int NFAa; int NCRa; int NHa; int NMa;
  int NFAb; int NCRb; int NHb; int NMb;
  int NFAC; int NCRc; int NHc; int NMc;
  int NFAd; int NCRd; int NHd; int NMd;
}
parameters {
  real dmu_h;
  real Cmu_h;
  real<lower = 0> dsigm_h;
  real<lower = 0> Csgm_h;
  real da; real Ca; real db; real Cb;
  real dc; real Cc; real dd; real Cd;
}
    
```

27

```

model {
  dmu_h ~ normal(0, 5); Cmu_h ~ normal(0, 5);
  dsigm_h ~ uniform(0, 5); Csgm_h ~ uniform(0, 5);
  da ~ normal(dmu_h, dsigm_h); db ~ normal(dmu_h, dsigm_h);
  dc ~ normal(dmu_h, dsigm_h); dd ~ normal(dmu_h, dsigm_h);
  Ca ~ normal(Cmu_h, Csgm_h); Cb ~ normal(Cmu_h, Csgm_h);
  Cc ~ normal(Cmu_h, Csgm_h); Cd ~ normal(Cmu_h, Csgm_h);
  NFAa ~ binomial(NFAa+NCRa, 1.0 - normal_cdf(Ca, 0.0, 1.0));
  NHa ~ binomial(NHa+NMa, 1.0 - normal_cdf(Ca, da, 1.0));
  NFAb ~ binomial(NFAb+NCRb, 1.0 - normal_cdf(Cb, 0.0, 1.0));
  NHb ~ binomial(NHb+NMb, 1.0 - normal_cdf(Cb, db, 1.0));
  NFAC ~ binomial(NFAC+NCRc, 1.0 - normal_cdf(Cc, 0.0, 1.0));
  NHc ~ binomial(NHc+NMc, 1.0 - normal_cdf(Cc, dc, 1.0));
  NFAd ~ binomial(NFAd+NCRd, 1.0 - normal_cdf(Cd, 0.0, 1.0));
  NHd ~ binomial(NHd+NMd, 1.0 - normal_cdf(Cd, dd, 1.0));
}
    
```

28



恒常法／上下法（階段法）

標準刺激 S_{sd} 比較刺激 $S_{cmp(i)}$

判断：3件法：「強い」「判らない／同じ」「弱い」

3件法は、2件法より精度が良い（RMSEが小さい）
 （岡本、2017、2019）

恒常法：比較刺激の組：実験前に決められている
 ランダムに選んで提示

上下法：観察者の判断に応じて
 次の提示刺激値を決める。

$$R_t^{(i)} = + \text{のとき, } X_{t+1}^{(i)} = X_t^{(i)} - \Delta_t^{(i)}$$

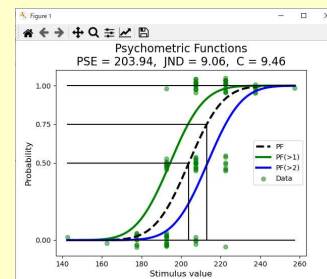
$$R_t^{(i)} = ? \text{のとき, } X_{t+1}^{(i)} = \begin{cases} X_t^{(i)} - \Delta_t^{(i)}, & \text{if } X_t^{(i)} < X_{t-1}^{(i)} \\ X_t^{(i)} + \Delta_t^{(i)}, & \text{if } X_t^{(i)} > X_{t-1}^{(i)} \end{cases}$$

$$R_t^{(i)} = - \text{のとき, } X_{t+1}^{(i)} = X_t^{(i)} + \Delta_t^{(i)}$$

岡本安晴（2017）基礎心理学研究、36、85-91.
 岡本安晴（2019）基礎心理学研究、38、90-104.

3件法上下法データの分析例

t	s	r	
2	0	257.5	3
3	1	237.5	3
4	2	222.5	3
5	3	207.5	3
6	4	192.5	1
7	5	207.5	2
8	6	222.5	2
9	7	237.5	3
10	8	222.5	3
11	9	207.5	2
12	10	192.5	1
13	11	207.5	1
14	12	222.5	3
15	13	207.5	3
16	14	192.5	2
17	15	177.5	1
18	16	192.5	1
19	17	207.5	2
20	18	222.5	2



感覚の強さ

カテゴリ尺度法（method of category scale）

観察者は、刺激の感覚の強さを順序カテゴリで答え、それに基づいて尺度を構成する方法は、カテゴリ尺度法（Gescheider, 1997, p. 214）、評定尺度法（難波・桑野、1998, pp. 72-86）、あるいは継次カテゴリ法／系列カテゴリ法（method of successive categories）、継次間隔法（method of successive intervals）（中谷、1969, p. 162）と呼ばれる。カテゴリ尺度法では、刺激の感覚量とカテゴリ境界の関係にThurstoneの判断の基本モデルを適用したカテゴリ判断の法則（The law of categorical judgment; Torgerson, 1958p. 205）を用いて分析することができる。

Gescheider, G. A. (1997). Psychophysics: The fundamentals, Third edition. Mahwah: Lawrence Erlbaum Associates, Publishers.

中谷和夫（1969）. 第5章 尺度構成法. 八木晃（監修）/ 田中良久（編）講座心理学2 計量心理学. 東京大学出版会.

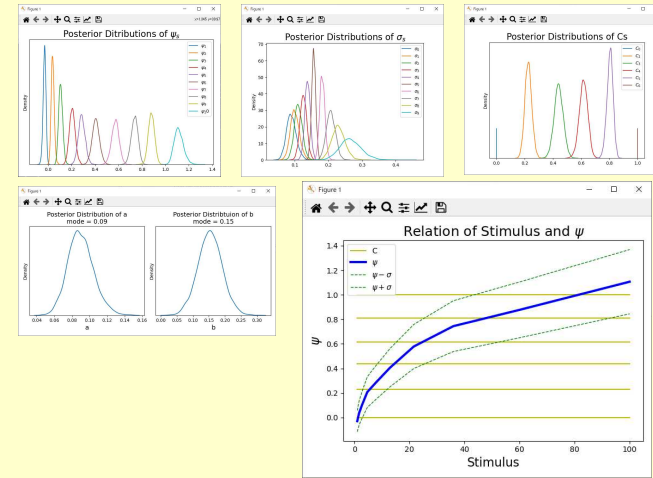
難波精一郎・桑野園子（1998）. 音の評価のための心理学的測定法. コロナ社.

Torgerson, W. S. (1958). Theory and methods of scaling. New York: John Wiley & Sons, Inc.

順序カテゴリ尺度法サンプルデータ (シミュレーション)

<http://y-okamoto-psy1949.la.coocan.jp/Python/tips/CategoricalScaling/>

St	R1	R2	R3	R4	R5	R6	R7
1							
2	1	63	37	0	0	0	0
3	1.668	33	65	2	0	0	0
4	2.783	18	67	15	0	0	0
5	4.642	4	53	40	3	0	0
6	7.743	3	28	56	13	0	0
7	12.915	1	13	45	30	11	0
8	21.544	0	3	21	34	33	8
9	35.938	0	0	8	20	36	25
10	59.948	0	0	1	15	21	32
11	100	0	0	0	2	7	26
12							
13							



3点識別試験ベイズ分析
d'の回帰分析例

3つの刺激の内、2つが同じ刺激Aで1つが異なる刺激Bであるとき、3つの中で他の2つと異なるものとして正しくBが選ばれる(あるいは、Bが他の2つと最も異なるものとして選ばれる)確率を調べる実験(3点識別試験: triangular test, 3AFC Oddity)において、実験協力者の特性(年齢など)によりAとBの感覚の差d'がどのように変化しているかを調べる場合を想定する。

信号検出理論に基づいて、刺激Aおよび刺激Bの感覚が、それぞれ独立に平均 μ_A および平均 μ_B の分散 σ^2 の正規分布(ガウス分布)に従うものとする(サーストンモデル)。2つの刺激Aの感覚を X_A および X'_A で表し、刺激Bの感覚を X_B で表すと

$$X_A \sim N(\mu_A, \sigma^2), \quad X'_A \sim N(\mu_A, \sigma^2), \quad X_B \sim N(\mu_B, \sigma^2)$$

である。

3点識別試験法において正しくBが選ばれるのは

$$|X_A - X'_A| < |X_A - X_B| \quad \text{かつ} \quad |X_A - X'_A| < |X'_A - X_B|$$

が成り立つときである (Frijters et al., 1980)。

Frijters, J. E. R., Kooistra, A., & Vereijken, P. F. G. (1980). Tables od d' for the triangular method and the 3-AFC signal detection procedure. *Perception & Psychophysics*, 1980, 27, 176-178.

いま、信号検出理論の慣例に倣って、

$$d' = \frac{\mu_B - \mu_A}{\sigma}$$

とおく。このとき、正答確率 (Bが選ばれる確率) P_c と感覚の差 d' の関係は次式で表される (Frijters et al., 1980)。

$$(1) \quad P_c = 2 \int_0^{\infty} \left[\Phi(-u\sqrt{3} + d'\sqrt{2/3}) + \Phi(-u\sqrt{3} - d'\sqrt{2/3}) \right] e^{-1/2 u^2} / \sqrt{2\pi} du$$

3点識別試験データが、実験参加者の判断の正誤と参加者の特性のペアで与えられているとする。例えば、右のデータ(仮想データ)では、判断の正誤が変数Rの値として1(正答)・0(誤答)で与えられ、判断者の年齢が変数Aの値として記入されている。変数CaseIDはデータの識別値で右のデータでは通し番号が記入されている。CaseIDが1の判断者の反応は正答(1)であり、年齢は33歳である。

CaseID	R	A
1	1	33
2	0	20
3	0	42
4	1	20
5	0	44
6	1	23
7	0	54
8	1	24
9	0	59
10	0	40
11	0	40

感覚の差 d' と年齢の関係を調べるが、 d' には値が負でないという制約があるので、リンク関数を用いて d' の範囲を $-\infty$ から $+\infty$ に変換して年齢（変数 X で表す）の1次式として表す（単回帰モデル）。すなわち、

$$\log(d') = a + bX$$

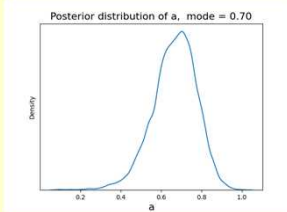
とおく。
このとき、

$$(2) \quad d' = \exp(a + bX)$$

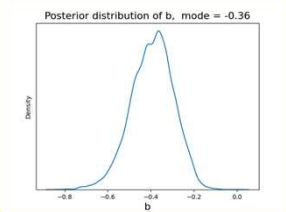
である。

式（1）と式（2）に基づいてベイズ分析を行うと以下の結果を得る。

パラメータ a の分布とモード パラメータ b の分布とモード



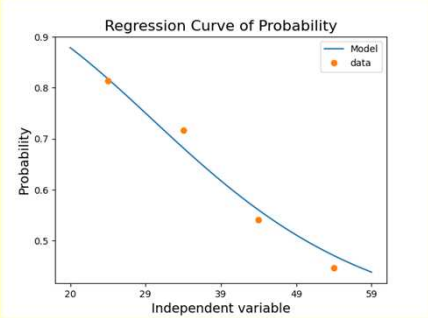
Posterior distribution of a, mode = 0.70



Posterior distribution of b, mode = -0.36

37

X と Pc の関係



Regression Curve of Probability

$$(1) \quad Pc = 2 \int_0^{\infty} \left[\Phi(-u\sqrt{3} + d'\sqrt{2/3}) + \Phi(-u\sqrt{3} - d'\sqrt{2/3}) \right] e^{-1/2 \cdot u^2} / \sqrt{2\pi} du$$

$$(2) \quad d' = \exp(a + bX)$$

38

まとめ

1. 官能評価（感覚）データには「ゆらぎ」があり、確率モデルで表される。
2. 確率モデルの分析法として、ベイズ分析（統計学）がある。
3. ベイズ分析を用いると、様々な回帰モデルの分析ができる。
4. 複数の条件のSDTデータの分析が、1つの枠内で行える。
5. 感覚の弁別特性を表す心理特性関数（PF）をベイズ分析で求めた。
6. 感覚の強さを、サーストンのカテゴリ判断の法則に基づいてベイズ分析で求めた。
7. d' を一般化回帰モデルによる分析を行った。

39