

# MPTモデルによる認知過程の分析 ーソースモニタリングデータによるチュートリアルー

PD 井関龍太

# MPTモデルとは

---

- MPTモデル (multinomial processing tree model) …カテゴリーカルデータ（頻度データ）を潜在的な認知過程を表すモデルに基づいて分析する手法

- 仮定した認知過程の測度（確率推定値, パラメータ）を得られる

- 複数のモデルを比較することによって, 統計的仮説検定を行える

- 今回の目的：ソースモニタリング実験における適用を具体例として, MPTモデルによる分析を紹介する

# 発表の構成

---

- ソースモニタリング実験のパラダイムと分析上の問題
- ソースモニタリングデータのためのMPTモデルの構成
- ソフトウェアを用いた適合度の検定とパラメータ推定
- 複数のモデルによる仮説検定

# ソースモニタリング実験の手続き

---

## ■ 典型的な手続き：

■ **学習段階**：刺激（単語など）をA（e.g.,男声）とB（e.g.,女声）の2つの文脈のいずれかで提示

## ■ **テスト段階**：

- ① **再認判断**：学習段階で提示されたものか，提示されなかったものか判断する
- ② **ソース判断**：①で“提示された”と判断した場合のみ，その項目が文脈A（男声）で提示されたか，文脈B（女声）で提示されたかを判断する

# ソースモニタリング実験のデータ構造

■ **項目のタイプ** : ソース A 項目, ソース B 項目, 新項目

■ **反応のタイプ** : ソース A 項目, ソース B 項目, 新項目

→ 3 × 3 のクロス表形式

	男声 A 反応	女声 B 反応	新反応
男声 A 項目	612	151	77
女声 B 項目	123	643	74
新項目	19	18	383

(Dodson et al., 1998, Table 2Aより) 4

# なぜMPTモデルか

---

■ ソースモニタリングデータの伝統的な分析法：比率の計算による

■ **再認率**：正しく再認できた旧項目の数／旧項目の総数

■ **ソース同定率**：ソースを正しく判断できた数／正しく再認できた旧項目の数

．．． など

# 伝統的な分析法の問題点 (Batchelder & Riefer, 1990)

---

- 別々のソースを通してデータを結合する
  - ソースごとの検出可能性（再認率）とソース弁別可能性の違いを無視している
- ヒットとフォルスアラームを別々に扱っている
  - 旧新再認の部分は、信号検出理論にしたがった方がよい
- ソース同定率は、異なる認知過程の組み合わせ効果を反映している可能性がある
  - ソース同定率 = 旧項目を正しく再認できた場合の正ソース同定の条件つき確率であり、実験の中での全体的な再認のレベルに依存する

# モデルの構築

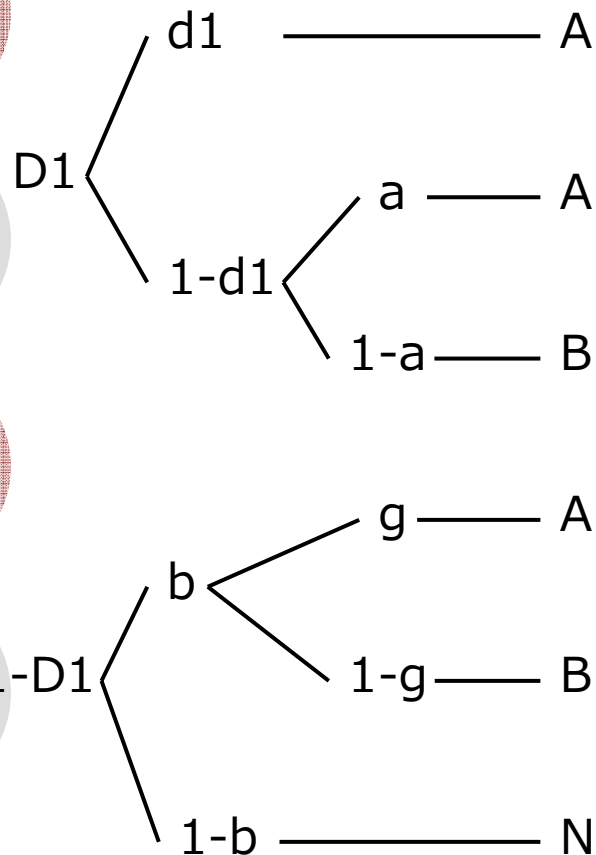
---

- データの構造と研究者の仮定する理論に基づいて、**分析ごとに**適切なモデルを構築する（≡構造方程式モデリング：SEM）
- ただし、ソースモニタリングデータの分析では、よく使用される**定型的なモデル**がある（Batchelder & Riefer, 1990）
  - 研究を通してデータ構造と研究者が仮定するであろう理論が似ている
  - アレンジして使うことも可能



# ソースモニタリングモデル

## ■ 基本的なモデル



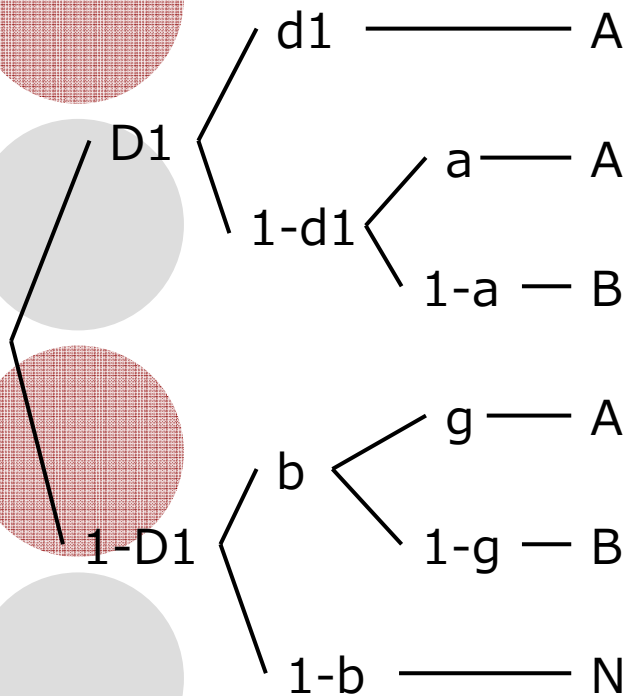
※A, B, Nは, それぞれ, ソースA, ソースB, 新の反応頻度

## ■ パラメータの意味

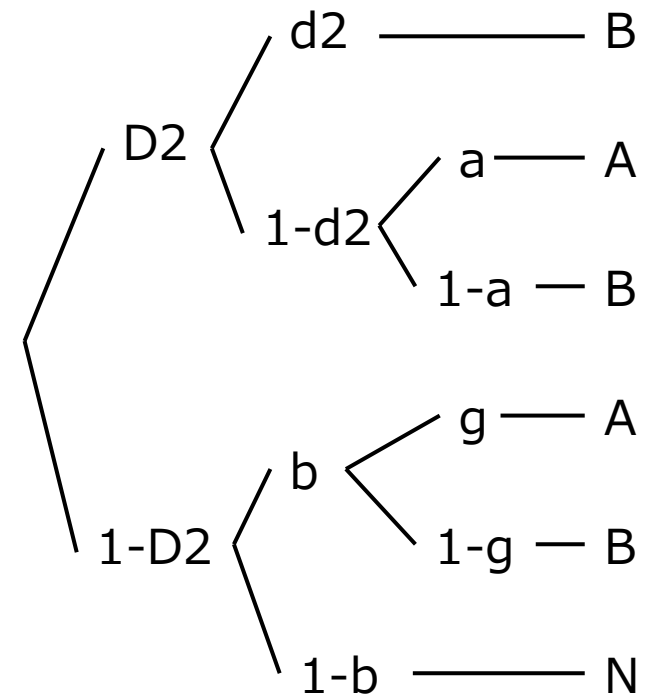
- **D = 検出可能性** ; 旧項目を正しく「旧項目」とであると判断する確率
- **d = ソース弁別可能性** ; 項目のソースを正しく判断する確率
- **b = 反応バイアス** ; 正しく検出できなかった項目を「旧項目」とであると判断する確率 (フォールスアラーム)
- **a = 推測率1** ; 正しく検出した項目を推測に基づいて「ソースA項目」とであると判断する確率
- **g = 推測率2** ; 正しく検出できなかった項目を推測に基づいて「ソースA項目」とであると判断する確率

# モデル全体の様子

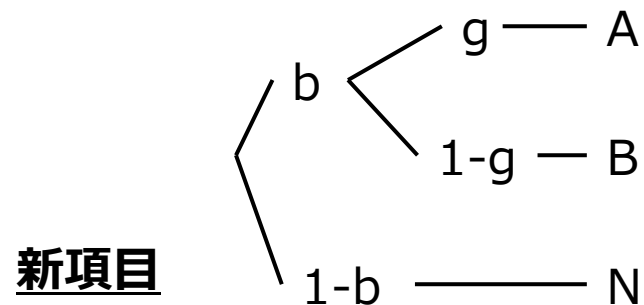
## ■ 項目のタイプごとにツリーを作る



ソースA項目



ソースB項目



新項目

# モデル式による表現

## ■ ソース A 項目 :

- $p_{11} = D_1 d_1 + D_1(1-d_1)a + (1-D_1)bg$

- $p_{12} = D_1(1-d_1)(1-a) + (1-D_1)b(1-g)$

- $p_{13} = (1-D_1)(1-b)$

## ■ ソース B 項目 :

- $p_{21} = D_2(1-d_2)a + (1-D_2)bg$

- $p_{22} = D_2 d_2 + D_2(1-d_2)(1-a) + (1-D_2)b(1-g)$

- $p_{23} = (1-D_2)(1-b)$

## ■ 新項目 :

- $p_{31} = bg$

- $p_{32} = b(1-g)$

- $p_{33} = (1-b)$

# パラメータの制約

■ **識別問題**：MPTモデルでは、モデル式に対してパラメータが多すぎると連立方程式を一意に解くことができず、**分析不能**に陥る（SEMと同じ）

■ **ソースモニタリングモデル**：

■ 自由度 = 6 ;  $= 3 \times 3 - 3$

■ パラメータ = 7 ; D1, D2, d1, d2, a, b, g  
そのままでは識別不能

→ **サブモデルの選択**：パラメータ間に**等値制約**を置くことで推定する個数を減らす

# ソースモニタリングモデルにおける 典型的なパラメータ制約

---

- 以下の 3 つの制約を組み合わせて用いる（全部で 7 つのサブモデル）
  - $D1 = D2$ 
    - 2 つのソースの間で検出可能性（再認率）は等しい
  - $d1 = d2$ 
    - 2 つのソースの間でソース弁別可能性は等しい
  - $a = g$ 
    - 旧項目の検出に成功した場合と失敗した場合で、ソース推測率は等しい
- 今回は、すべての制約を置いたモデルを使用
  - Dodson et al. (1998) にしたがう

# ソフトウェア

---

## ■ multiTree ←今回は、これを使います

- <http://psycho3.uni-mannheim.de/index.php?n=Main.MultiTree>

- Windows 版, Mac版, Linux版あり ; 二股分岐のモデルのみ扱える

## ■ HMMTree

- <http://www.psychologie.uni-freiburg.de/Members/stahl/HMMTree>

- Windows XP/Vista/7 ; 二股分岐のモデルのみ扱える

## ■ GPT.EXE: Analysis for General Processing Tree Models

- <http://www.xiangenhu.info/course/view.php?id=53>

- Windows版, Mac OS X版の両方がzipフォルダに含まれる

## ■ Excelのソルバー

- <http://socrates.berkeley.edu/~wprinz/>

- Windows用とMac用の両方のシートがzipフォルダに含まれる

# multiTree : モデルの定義

## ■ モデル式のファイルを作る

■ 1 行目 . . . すべてのツリーの方程式の合計数を記述する

■ 2 行目以降

■ 1 列目 . . . ツリーの違いを表すラベル

■ 2 列目 . . . 反応カテゴリーの違いを表すラベル

■ 3 列目 . . . 方程式

■ 拡張子を“\*eqn”にすること

15

1 1 D1\*d1

1 1 D1\*(1-d1)\*a

1 2 D1\*(1-d1)\*(1-a)

1 1 (1-D1)\*b\*g

1 2 (1-D1)\*b\*(1-g)

1 3 (1-D1)\*(1-b)

2 5 D2\*d2

2 4 D2\*(1-d2)\*a

2 5 D2\*(1-d2)\*(1-a)

2 4 (1-D2)\*b\*g

2 5 (1-D2)\*b\*(1-g)

2 6 (1-D2)\*(1-b)

3 7 b\*g

3 8 b\*(1-g)

3 9 (1-b)

ソースモニタリング  
モデルの場合

列はスペース  
で区切ります

# multiTree : データセットの入力

## ■ データファイルを作る

■ 1行目・・・タイトル  
(任意の名前)

■ 2行目以降

■ 1列目・・・反応カテゴリーの違いを表すラベル  
(eqnファイルのラベルと一致させること)

■ 2列目・・・反応頻度

■ 最終行・・・"==="を入力する

■ 拡張子を"\*mdt"にすること

voice\_change

1 612

2 151

3 77

4 123

5 643

6 74

7 19

8 18

9 383

===

Dodson et al. (1998)  
のデータの場合

列はスペースで  
区切っています



# multiTree : ファイルの読み込み

---

- 作成したeqnファイルとmdtファイルをmultiTreeに読み込む
  - メニューバーの“File”から“Open multiTree File”を選択
  - ダイアログから目的のファイルを選択
    - eqnファイルとmdtファイルの両方について、この読み込み作業を行う
    - それぞれ, “Equations”タブと“Data”タブに反映される (→確認)

# multiTree : パラメータの制約

## ■ “Parameters”タブを開く

- モデルで指定した潜在変数が並んでいる
- デフォルトでは, どのパラメータも“free”

## ■ パラメータを制約する

- “free”になっているボタンをクリック  
→展開されたメニューからオプションを選択
- 例 : “D1 = D2”としたい
  - “D2”の隣のボタンをクリック, オプションから  
“= D1”を選択→ボタンの表示が“= D1”に !
  - “D1”の側は “free”にしておくこと
- “constant”を選ぶと定数を指定できる
  - ボタン右側のウィンドウの中に固定したい数値を入力する

# multiTree : 分析実行

---

- 後は分析ボタンを押すだけ
  - メニューバーの下にある緑の矢印をクリックする
  - または、メニューバーの “Analysis” から “Run analysis” を選択  
→ 結果が出力される
    - 分析実行前にモデル式を保存するか尋ねられたら、YESかNOを選択する

# モデル適合の検討

## ■ “Model Fit” セクション

- $PD^\lambda$  : 一般に “ $G^2$ ” として知られる統計量と同じ ( $\lambda = 0$  のとき) ; カイ二乗の近似値

- $df$  : モデルの自由度

- $p$  : カイ二乗検定の有意確率

- 有意 = モデルとデータが合致しない

- 有意でない = モデルを容認できる

- $AIC, BIC$  : 情報量基準 ; モデルの相対比較に用いる

- $\Delta AIC, \Delta BIC$  : 飽和帰無仮説モデルとの比較の指標 (正 = モデルを棄却, 負 = モデルを支持)

# パラメータの検討

---

## ■ “Parameter Estimates”セクション

■ パラメータ：MPTモデルにより推定された確率

■ () 内：“asymptotic variances”（たぶん）

■ [] 内：信頼区間（たぶん）

■ MPTモデルでは、パラメータはツリーの各分岐での確率を表す

→値が大きいほど分岐の一方の方向が選ばれやすいことを示していると思われる

# 仮説検定

- Dodson et al. (1998) のデータ例：
  - 単語を聴覚提示する際の声の性別を操作
  - ソース判断のとき、声が同性か異性かということも判断の手がかりになるのではないか

Dodson et al. (1998). Table 2A, 2Bのデー

タ	男声 A反	女声 B 反応	新反応		男声 A 反応	男声 C 反応	新反応
男声 A 項目	612	151	77	男声 A 項目	521	201	109
女声 B 項目	123	643	74	男声 C 項目	172	566	102
新項目	19	18	383	新項目	28	32	360

# 統合したデータファイルを作る

■ 2つの条件のデータをまとめて1ファイルにする

■ 類似条件と非類似条件で反応カテゴリーのレベルを変えておく

voice\_similarity

1 612

2 151

3 77

4 123

5 643

6 74

7 19

8 18

9 383

10 521

11 210

12 109

13 172

14 566

15 102

16 28

17 32

18 360

===

非類似条件のデータ

類似条件のデータ

# 統合したモデル式を作る

30

1 1 LD1\*Ld1

1 1 LD1\*(1-Ld1)\*La

1 2 LD1\*(1-Ld1)\*(1-La)

1 1 (1-LD1)\*Lb\*Lg

1 2 (1-LD1)\*Lb\*(1-Lg)

1 3 (1-LD1)\*(1-Lb)

2 5 LD2\*Ld2

2 4 LD2\*(1-Ld2)\*La

2 5 LD2\*(1-Ld2)\*(1-La)

2 4 (1-LD2)\*Lb\*Lg

2 5 (1-LD2)\*Lb\*(1-Lg)

2 6 (1-LD2)\*(1-Lb)

3 7 Lb\*Lg

3 8 Lb\*(1-Lg)

3 9 (1-Lb)

4 10 HD1\*Hd1

4 10 HD1\*(1-Hd1)\*Ha

4 11 HD1\*(1-Hd1)\*(1-Ha)

4 10 (1-HD1)\*Hb\*Hg

4 11 (1-HD1)\*Hb\*(1-Hg)

4 12 (1-HD1)\*(1-Hb)

5 14 HD2\*Hd2

5 13 HD2\*(1-Hd2)\*Ha

5 14 HD2\*(1-Hd2)\*(1-Ha)

5 13 (1-HD2)\*Hb\*Hg

5 14 (1-HD2)\*Hb\*(1-Hg)

5 15 (1-HD2)\*(1-Hb)

6 16 Hb\*Hg

6 17 Hb\*(1-Hg)

6 18 (1-Hb)

■ 類似条件と非類似条件のデータに同じ構造の別のツリーを当てはめている

→ 6本のツリー

■ 非類似条件のパラメータはL, 類似条件のパラメータはHを頭につけて示す

※表示しきれないので, 2段に分けています



# 仮説検定の手続き（1）

- 2つの条件を統合したモデルで推定を行う
  - 個別に分析した場合と同じパラメータが得られる
  - $G^2$ は2つのモデルの $G^2$ を足した値に一致（ $G^2 = .98$ ）
  - この結果を等値制約を加えたモデルと比較
- 2つの条件でソース弁別可能性は異なるか
  - 類似条件 :  $Hd1 = .49$
  - 非類似条件 :  $Ld1 = .65$
  - 非類似条件でよりソース弁別できているように見える（このことを検定する）

## 仮説検定の手続き（２）

### ■ “ $Hd1 = Ld1$ ” と制約を置く（帰無仮説）

■ ソース弁別可能性に差がないという仮定

■ multiTreeの“Parameters”タブを開いて、この制約を適用する

→ 分析：  $G^2 = 26.79$ ,  $df = 5$

### ■ カイ二乗検定

■ 制約を置く前のモデルと置いたモデルの  $G^2$  と  $df$  を引き算する

■  $G^2 = 26.79 - 0.98 = 25.81$

■  $df = 5 - 4 = 1$

■ 検定：  $G^2(1) = 25.81$  のとき,  $p < .001$

# 検定結果の解釈

■ 2つのモデルの適合度に**有意な差**が見られた

■ **示唆**：“ $Hd1 = Ld1$ ”の制約を置いたモデルで大きく適合度が下がっているので、この制約を置かない方がよい

■ **結論**：“ $Hd1 = Ld1$ ”でないと仮定した方がよい、つまりは、類似条件と非類似条件のソース弁別可能性は**有意に異なる**

■ 同様のやり方で、他のパラメータについても条件間で差があるか検定できる

# MPTモデルの応用可能性

---

- 対連合学習による干渉パラダイム
- 自由再生における項目のクラスタリング
- 過程分離手続きのデータ
- 目撃証言
- 後知恵バイアス
- 4枚カード問題
- 項目反応理論
- レビューとして, Batchelder & Riefer (1999), Erdfelder et al. (2009) など

# 文献

- Batchelder, W. H., & Riefer, D. M. (1990). Multinomial processing models of source monitoring. *Psychological Review*, **97**, 548-564.
- Batchelder, W. H., & Riefer, D. M. (1999). Theoretical and empirical review of multinomial process tree modeling. *Psychonomic Bulletin & Review*, **6**, 57-86.
- Dodson, C. S., Prinzmetal, W., & Shimamura, A. P. (1998). Using Excel to estimate parameters from observed data: An example from source memory data. *Behavior Research Methods, Instruments, & Computers*, **30**, 517-526.
- Erdfelder, E., Auer, T.-S., Hilbig, B. E., Aßfalg, A., Moshagen, M., & Nadarevic, L. (2009). Multinomial processing tree models: A review of the literature. *Zeitschrift für Psychologie/Journal of Psychology*, **217**, 108-124.
- Hu, X., & Phillips, G. A. (1999). GPT.EXE: A powerful tool for the visualization and analysis of general processing tree models. *Behavior Research Methods, Instruments, & Computers*, **31**, 220-234.
- Moshagen, M. (2010). multiTree: A computer program for the analysis of multinomial processing tree models. *Behavior Research Methods*, **42**, 42-54.
- Stahl, C., & Klauer, K. C. (2007). HMMTree: A computer program for latent-class hierarchical multinomial processing tree models. *Behavior Research Methods*, **39**, 267-273.