

# 丁寧語の通時的構文交替： 状態空間モデルを用いた ベイズ時系列分析



The Diachronic Constructional Alternation  
of Japanese addressee-honorific markers:  
A Bayesian Time Series Analysis  
using State-Space Models

山田彬堯 Akitaka Yamada

言語文化共同研究プロジェクト

テキストマイニングとデジタルヒューマニティーズ2021

日時: 2022年2月4日 (金) 17:00-

---

# はじめに

## 言語統計学A・B

大学院での授業が一年間終わったので、そのご報告と、授業で扱った内容と今日の発表の関係について、冒頭で簡単にお伝えします。

### 言語統計学A

### 言語統計学B

変数：観測変数

従属変数：正規分布

独立変数：

	質	量
固定効果	1つ t 検定 2つ以上 分散分析	単回帰 重回帰

混合効果	分散分析 (対応あり)	一般線形 混合効果 モデル
------	----------------	---------------------

正規分布以外

固定効果	一般化 線形 モデル
------	------------------

混合効果	一般化 線形 混合効果 モデル
------	--------------------------

潜在変数

- ・ 隠れマルコフモデル
- ・ 順序ロジスティック回帰
- ・ トピックモデル
- ・ 共分散構造方程式モデリング
- ・ 状態空間モデル (時系列分析)

# はじめに

## 時系列分析

観測される値

23.1

27.1

21.1

1936

1937

1938

## 時系列データ

これは、時点ごとの確率変数の測定値を、その時間順に並べた系列のこと。

# はじめに

## 時系列分析

観測される値

0

1936

1

1937

1

1938

## 時系列データ

これは、時点ごとの確率変数の測定値を、その時間順に並べた系列のこと。

# はじめに

## 時系列分析

観測される値

to v adv

1936

to adv v

1937

to adv v

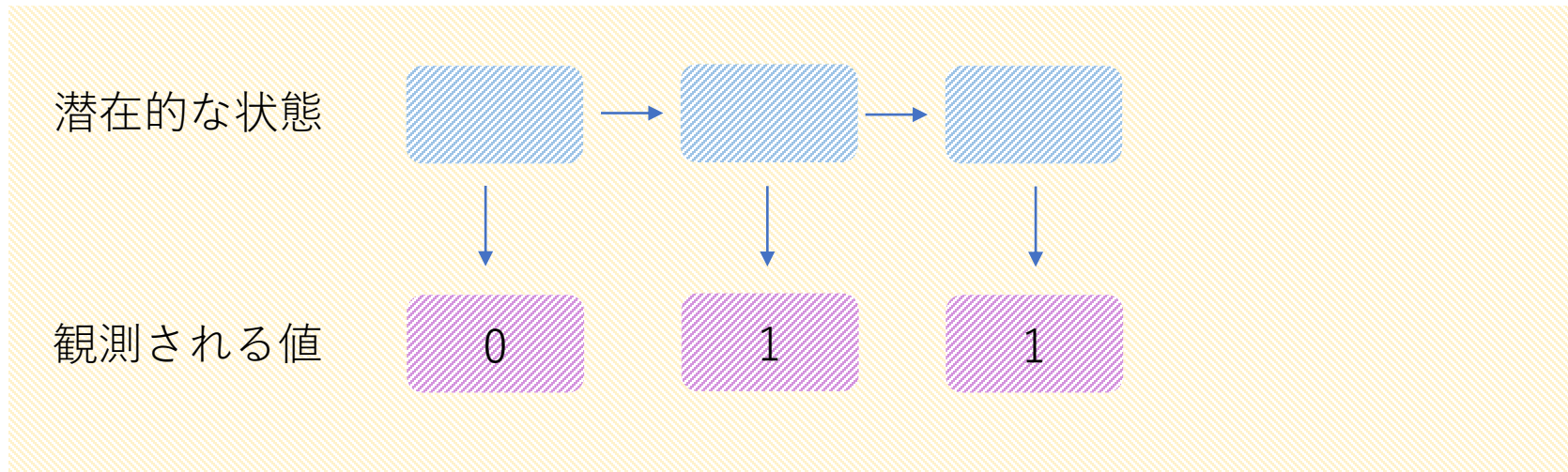
1938

### 時系列データ

これは、いつ採取されたのかという時点が併せて観察されている確率変数の測定値の系列のこと。

# はじめに

## 状態空間モデル

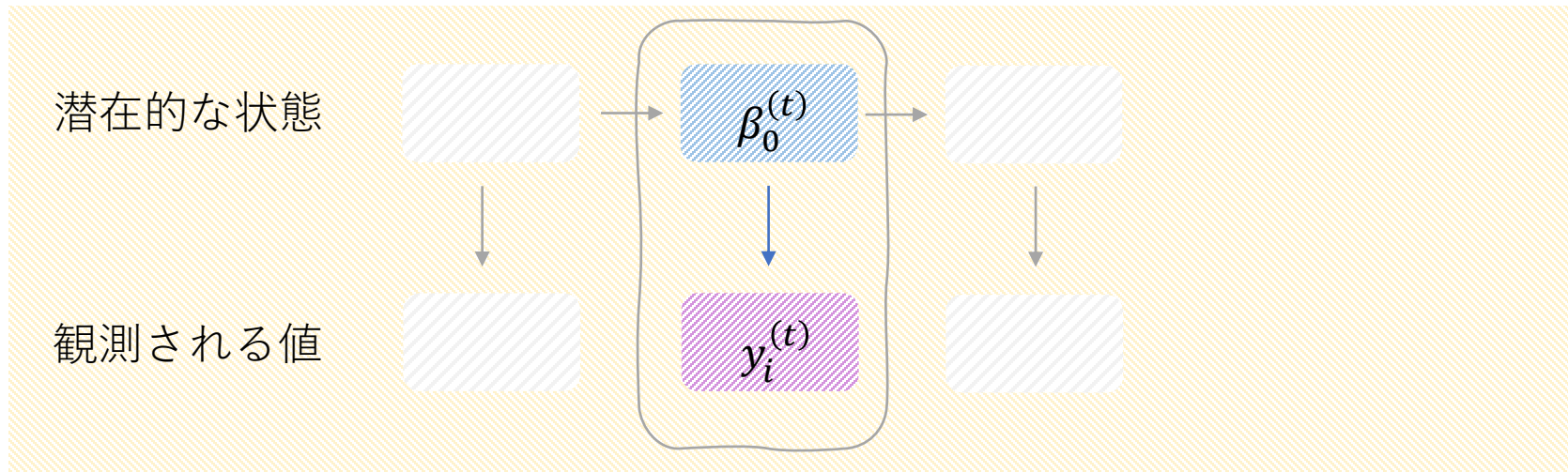


### 状態空間モデル State Space Model (SSM)

これは、観測されている変数の背後に、その観測値を生み出す目に見えない状態の存在を仮定し、この潜在的な状態が時刻ごとに変化するというメカニズムを想定する統計モデル。

# はじめに

## 状態空間モデル



### 観測モデル Observation Model

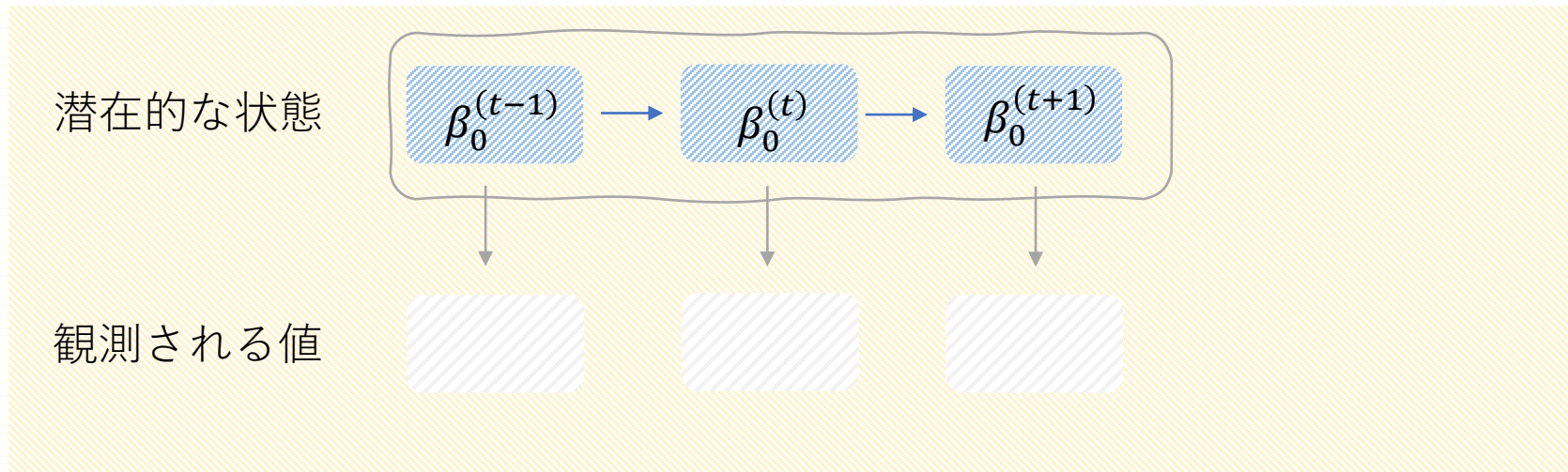
これは、潜在的な状態をパラメータに、どのように観測値が得られるかを表す統計モデル。

$$\text{(例)} \quad y_i^{(t)} \sim \text{Cat} \left( F^{-1} \left( \beta_0^{(t)} \right) \right)$$

⇒共時的な一点に注目すれば  
なじみのあるカテゴリカル分布

# はじめに

## 状態空間モデル



### 状態モデル State Model

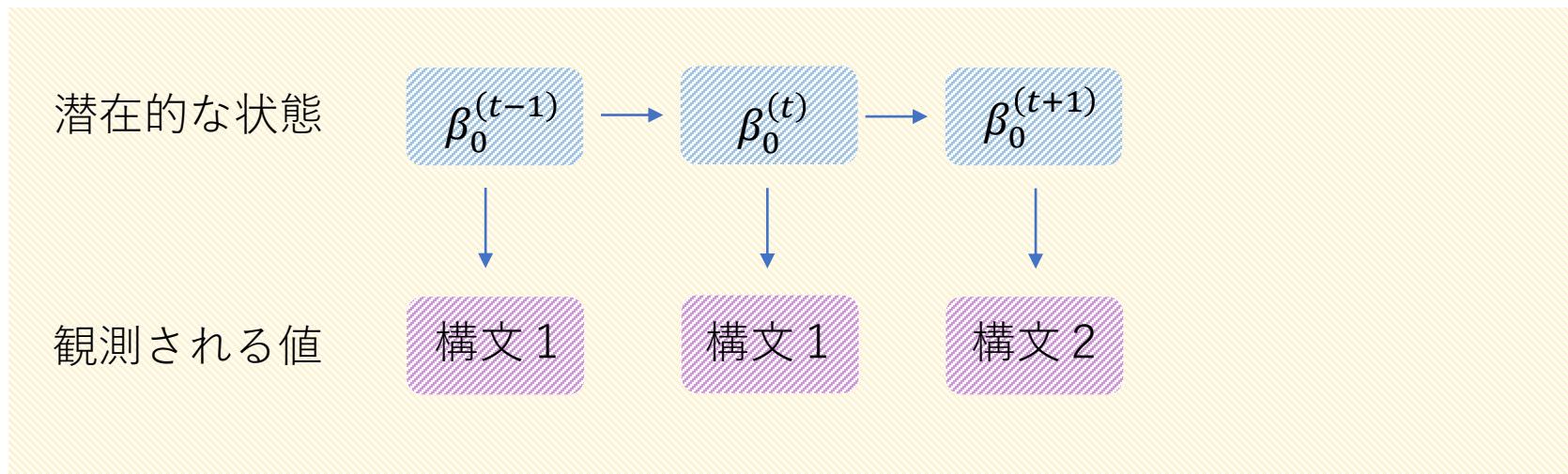
これは、ある時点の潜在的な状態がそれ以前の状態にどのように依存しているのかを表す統計モデル。

(例)  $\beta_0^{(t)} \sim N(\beta_0^{(t-1)}, \sigma_w^2)$



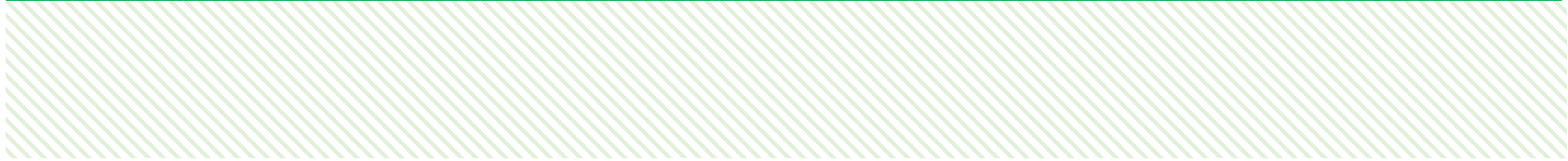
# はじめに

## 今日の発表のあらすじ

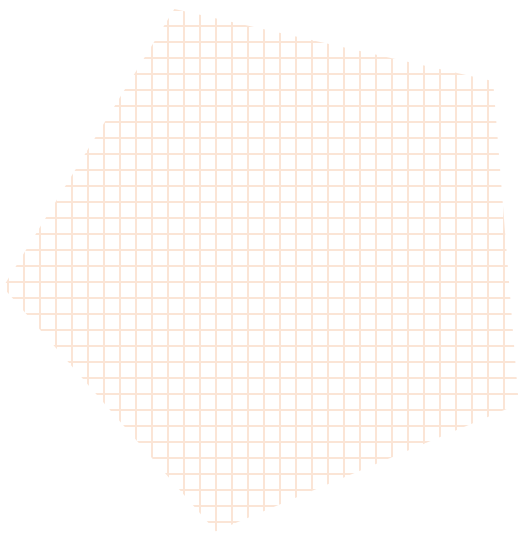
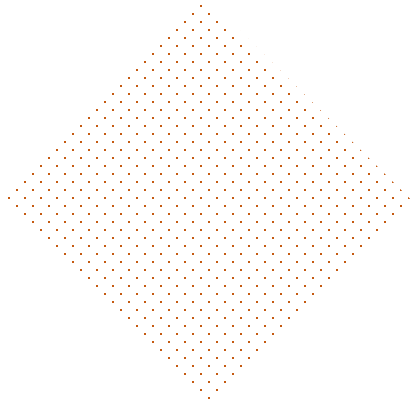


## 目的

- ・ 状態空間モデルを用いて、丁寧語に起こった「言語変化」をモデル化する。
- ・ ネイティブの意見が聞けず、コーパスに頼ることしかできない昔の言語の動態について統計解析を施すことで、定性的な分析だけでは明示的に示しにくい変化の素描が具体的／客観的に分析できることを示す。



# 先行研究



# 先行研究：日本語の丁寧語の“ふしぎさ”

## 日本語の丁寧語に特徴的な点(1): 否定丁寧形

### 普通形

#### 現在形

- |                 |                      |
|-----------------|----------------------|
| (1) 走る          | (3) 走らない             |
| <i>hasir-u.</i> | <i>hasira-nai.</i>   |
| run-PRS         | run-NEG              |
| ‘(He) runs.’    | ‘(He) does not run.’ |

#### 過去形

- |                  |                          |
|------------------|--------------------------|
| (2) 走った          | (4) 走らなかった               |
| <i>hasit-ta.</i> | <i>hasira-nak at-ta.</i> |
| run-PST          | run-NEG COP-PST          |
| ‘(He) ran.’      | ‘(He) did not run.’      |

### 丁寧形

#### 現在形

- |                      |                       |
|----------------------|-----------------------|
| (5) 走ります             | (7) 走りません             |
| <i>hasiri-mas-u.</i> | <i>hasiri-mas-en.</i> |
| run-AH-PRS           | run-AH-NEG            |
| ‘(He) runs.’         | ‘(He) does not run.’  |

#### 過去形

- |                        |                               |
|------------------------|-------------------------------|
| (6) 走りました              | (8) 走りませんでした                  |
| <i>hasiri-masi-ta.</i> | <i>hasiri-mas-en desi-ta.</i> |
| run-AH-PST             | run-AH-NEG COP.AH-PST         |
| ‘(He) ran.’            | ‘(He) did not run.’           |

## 興味深い点

否定丁寧形では二回「丁寧語」が発音されているのに敬意の度合いは他と同じ。

# 先行研究：日本語の丁寧語の“ふしぎさ”

## 日本語の丁寧語に特徴的な点(1): 否定丁寧形

### 普通形

#### 現在形

- |                 |                      |
|-----------------|----------------------|
| (1) 走る          | (3) 走らない             |
| <i>hasir-u.</i> | <i>hasira-nai.</i>   |
| run-PRS         | run-NEG              |
| ‘(He) runs.’    | ‘(He) does not run.’ |

#### 過去形

- |                  |                          |
|------------------|--------------------------|
| (2) 走った          | (4) 走らなかった               |
| <i>hasit-ta.</i> | <i>hasira-nak at-ta.</i> |
| run-PST          | run-NEG COP-PST          |
| ‘(He) ran.’      | ‘(He) did not run.’      |

### 丁寧形

#### 現在形

- |                      |                       |
|----------------------|-----------------------|
| (5) 走ります             | (7) 走りません             |
| <i>hasiri-mas-u.</i> | <i>hasiri-mas-en.</i> |
| run-AH-PRS           | run-AH-NEG            |
| ‘(He) runs.’         | ‘(He) does not run.’  |

#### 過去形

- |                        |                               |
|------------------------|-------------------------------|
| (6) 走りました              | (8) 走りませんでした                  |
| <i>hasiri-masi-ta.</i> | <i>hasiri-mas-en desi-ta.</i> |
| run-AH-PST             | run-AH-NEG COP.AH-PST         |
| ‘(He) ran.’            | ‘(He) did not run.’           |

(9) 中世日本語（東洋文庫弁官補任裏文書；鎌倉遺文 #3078；1223年; Yamada 2019）

*pitobito motii-sooraw-azu sooraw-iki*  
people use-HON<sub>K</sub>-NEG HON<sub>K</sub>-PST  
‘(i) people did not use (it);’

語源的に全く関係のない「候ふ」でも doubling が後期になると生まれている  
⇒繰り返す丁寧語の歴史

# 先行研究：日本語の丁寧語の“ふしぎさ”

## 日本語の丁寧語に特徴的な点(2):マーキングされる場所

### 普通形

#### 現在形

- |                 |                      |
|-----------------|----------------------|
| (1) 走る          | (3) 走らない             |
| <i>hasir-u.</i> | <i>hasira-nai.</i>   |
| run-PRS         | run-NEG              |
| ‘(He) runs.’    | ‘(He) does not run.’ |

#### 過去形

- |                  |                          |
|------------------|--------------------------|
| (2) 走った          | (4) 走らなかった               |
| <i>hasit-ta.</i> | <i>hasira-nak at-ta.</i> |
| run-PST          | run-NEG COP-PST          |
| ‘(He) ran.’      | ‘(He) did not run.’      |

### 丁寧形

#### 現在形

- |                      |                       |
|----------------------|-----------------------|
| (5) 走ります             | (7) 走りません             |
| <i>hasiri-mas-u.</i> | <i>hasiri-mas-en.</i> |
| run-AH-PRS           | run-AH-NEG            |
| ‘(He) runs.’         | ‘(He) does not run.’  |

#### 過去形

- |                        |                               |
|------------------------|-------------------------------|
| (6) 走りました              | (8) 走りませんでした                  |
| <i>hasiri-masi-ta.</i> | <i>hasiri-mas-en desi-ta.</i> |
| run-AH-PST             | run-AH-NEG COP.AH-PST         |
| ‘(He) ran.’            | ‘(He) did not run.’           |

### 興味深い点

- ・聞き手を目指す表現は「文の末尾」に出ると予想されている (Speas and Tenny 2003)
- cf. 間主観化

# 先行研究：日本語の丁寧語の“ふしぎさ”

## 日本語の丁寧語に特徴的な点(2):マーキングされる場所

### 普通形

#### 現在形

- |                 |                      |
|-----------------|----------------------|
| (1) 走る          | (3) 走らない             |
| <i>hasir-u.</i> | <i>hasira-nai.</i>   |
| run-PRS         | run-NEG              |
| ‘(He) runs.’    | ‘(He) does not run.’ |

#### 過去形

- |                  |                          |
|------------------|--------------------------|
| (2) 走った          | (4) 走らなかった               |
| <i>hasit-ta.</i> | <i>hasira-nak at-ta.</i> |
| run-PST          | run-NEG COP-PST          |
| ‘(He) ran.’      | ‘(He) did not run.’      |

### 丁寧形

#### 現在形

- |                      |                       |
|----------------------|-----------------------|
| (5) 走ります             | (7) 走りません             |
| <i>hasiri-mas-u.</i> | <i>hasiri-mas-en.</i> |
| run-AH-PRS           | run-AH-NEG            |
| ‘(He) runs.’         | ‘(He) does not run.’  |

#### 過去形

- |                        |                               |
|------------------------|-------------------------------|
| (6) 走りました              | (8) 走りませんでした                  |
| <i>hasiri-masi-ta.</i> | <i>hasiri-mas-en desi-ta.</i> |
| run-AH-PST             | run-AH-NEG COP.AH-PST         |
| ‘(He) ran.’            | ‘(He) did not run.’           |

## (10) 韓国語

Inho-ka choysen-ul ta ha-ess-supnita.  
Inho-NOM best-ACC all do-PST-AH.DECL.FORM  
‘Inho did his best’ (Portner et al. 2019: 6)

文末の丁寧語は  
韓国語やタイ語が有名

# 先行研究：日本語の丁寧語の“ふしぎさ”

## 日本語の丁寧語に特徴的な点(3):イ形容詞との相性

### ナ形容詞（形容動詞）

- (11) 温厚**です**
- (12) 温厚**でした**
- (13) 温厚ではあり**ません**
- (14) 温厚ではあり**ませんでした**

### イ形容詞（形容詞）

- (15) (?)美し(い)**です**
- (16) \*美し(い)**でした**
- (17) 美しくはあり**ません**
- (18) 美しくはあり**ませんでした**

### 興味深い点

イ形容詞は「丁寧語」が使えなかった！

⇒ 20世紀前半の規範文法書は「**美しゅうございます**」を推奨。

# 先行研究：日本語の丁寧語の“ふしぎさ”

## 日本語の丁寧語に特徴的な点(4):新規形式の登場

### 旧形式 (規範形式)

#### 現在形

- (5) 走ります (7) 走りません  
*hasiri-mas-u.*      *hasiri-mas-en.*  
run-AH-PRS      run-AH-NEG  
'(He) runs.'      '(He) does not run.'

#### 過去形

- (6) 走りました (8) 走りませんでした  
*hasiri-masi-ta.*      *hasiri-mas-en desi-ta.*  
run-AH-PST      run-AH-NEG      COP.AH-PST  
'(He) ran.'      '(He) did not run.'

### 新規形式

#### 現在形

- (5') \*走ります (7') 走りません  
*hasir-u des-u*      *hasira-nai des-u*  
run-PRS AH-*u*      run-AH-NEG      COP.AH-*u*

#### 過去形

- (6') 走りました (8') 走りませんでした  
*hasit-ta des-u*      *hasira-nak at-ta des-u.*  
run-PST AH-*u*      run-NEG      COP-PST AH-*u*  
'(He) ran.'      '(He) did not run.'



# リサーチクエスチョン

## 日本語の丁寧語に特徴的な点:まとめ

- (1) 否定丁寧形
- (2) マーキング位置
- (3) イ形容詞との相性
- (4) 新規形式の登場

### 規範文法の扱い

- 国語調査委員会（編）（1916）『口語法』
- ・「美しゅうございます」（規範形）
  - ・「美しいです」（記載なし）

# リサーチクエスチョン

## 日本語の丁寧語に特徴的な点:まとめ

- (1) 否定丁寧形
- (2) マーキング位置
- (3) イ形容詞との相性
- (4) 新規形式の登場

### 規範文法の扱い

国語調査委員会（編）（1916）『口語法』

- ・「美しゅうございます」（規範形）
- ・「美しいです」（記載なし）

三尾（1942）

「ちかごろは、[...]元来叙述機能のない語へつくはずの「です」を、形容詞の基本形につけて用いるやうになつてきました。まだ語感にしつくりするところまでは、用法が熟し切つてゐるとはいへないかもしれませんが、[...]」

# 先行研究：日本語の丁寧語の“ふしぎさ”

## 日本語の丁寧語に特徴的な点：まとめ

- (1) 否定丁寧形
- (2) マーキング位置
- (3) イ形容詞との相性
- (4) 新規形式の登場

### 国定教科書

川口 (2014:89)

表3. 国定国語教科書における動詞・イ形容詞丁寧形の規範

	第一期 (1904)	第二期 (1910)	第三期 (1918)	第四期 (1933)	第五期 (1941)	第六期 (1947)
おもしろいです						△
おもしろうございます	○	○	○	○	○	○
おもしろく(は)ありません	○	○	○	○	○	○
おもしろくないです						
行きます	○	○	○	○	○	○
行きません	○	○	○	○	○	○
行かないです						

第一期：明治 37 (1904) 年～『尋常小學讀本』卷一～八

第二期：明治 43 (1910) 年～『尋常小學讀本』卷一～十二

第三期：大正 7 (1918) 年～『尋常小學國語讀本』卷一～十二

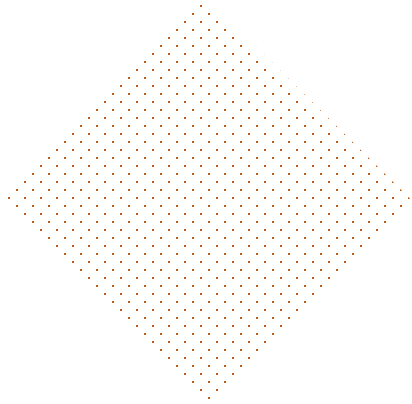
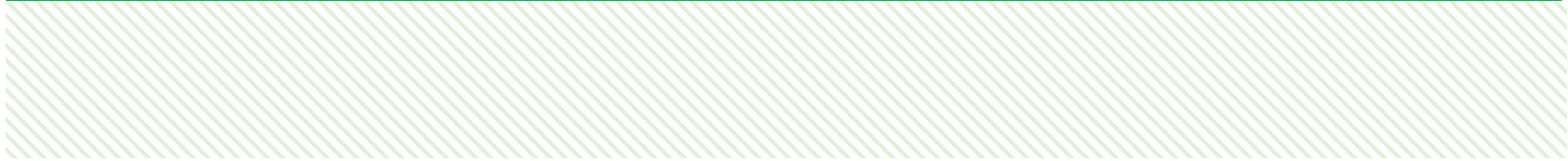
# 先行研究：日本語の丁寧語の“ふしぎさ”

## 日本語の丁寧語に特徴的な点：まとめ

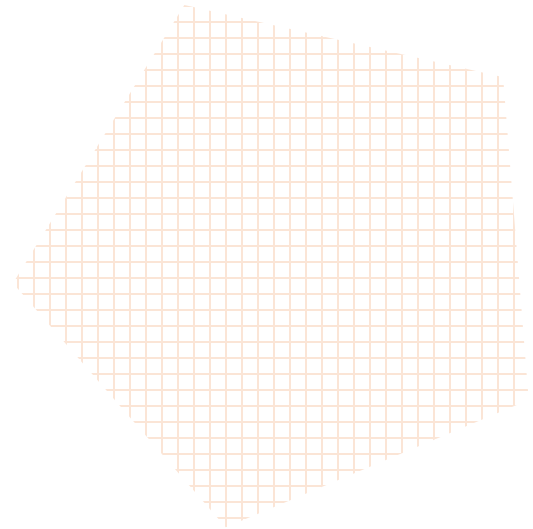
- (1) 否定丁寧形
- (2) マーキング位置
- (3) イ形容詞との相性
- (4) 新規形式の登場

## リサーチクエスチョン

- (19) 先行研究において取り残されているリサーチクエスチョン
  - a. 新形式はいつごろから用いられてきたのか。
  - b. どのような社会言語学的環境で新形式はその使用を拡大させていったのか。
  - c. どのような理論言語学的要因で新形式はその使用を拡大させていったのか。
  - d. イ形容詞の間に新形式を用いる選好度合いに違いがあったのか。



# データ



# データ

## コーパス

まとめて検索（試験公開版）

# KOTONOHA

書字形出現形で検索

○ 書字形出現形で検索 ● 語彙素で検索

この検索窓では、現在コーパス開発センターが作っている『まとめて検索 KOTONOHA』を使い、あなたが利用申請している（【検索対象】が灰色でない）コーパスをまとめて一括検索できます。  
検索は短単位の手書き出現形と語彙素の2種類が選択できます。検索ボタンをクリックすると、KOTONOHA 試用版に移動し、そこに検索結果が表示されます。

【検索対象】 [現代日本語書き言葉均衡コーパス](#) [国語研日本語ウェブコーパス](#) [日本語話し言葉コーパス](#) [日本語日常会話コーパス](#) [昭和話し言葉コーパス](#)  
[名大会話コーパス](#) [現日研・職場談話コーパス](#) [日本語歴史コーパス](#) [日本語談話方言コーパス](#) [多言語母語の日本語学習者横断コーパス](#)

■「中納言」「KOTONOHA」セキュリティ診断のお知らせ■  
「中納言」「KOTONOHA」のセキュリティ診断日が1/18(火)と2/7(月)-2/9(水)に決まりました。  
セキュリティ診断中もサービスはご利用になれますが、つながりにくくなることもございます。

■国立国語研究所ネットワーク工事のお知らせ■  
2022/02/11(祝・金)～02/14(月)13:00に国立国語研究所のネットワーク工事を予定しており、「中納言」「KOTONOHA」を含めた国語研の各種ウェブサービスが利用できなくなります。

【個別検索】ご利用になりたいコーパス名をクリックしてください。

コーパス名	略称	個別検索	包括的検索	備考
<a href="#">書き言葉</a> 現代日本語書き言葉均衡コーパス 中納言版	BCCWJ	✓	✓	従来より利用いただいている BCCWJ のデータです（コーパスの紹介ページ）。こちらのページから BCCWJ アノテーションデータをダウンロードできます。
<a href="#">書き言葉</a> 国語研日本語ウェブコーパス 中納言版	NWJC	準備中	✓	
<a href="#">話し言葉</a> 日本語話し言葉コーパス	CSJ	✓	✓	<a href="#">コーパスの紹介ページ</a>

# データ

## コーパス

The screenshot shows the '中納言' (Chunagon) corpus search application. The header includes the title '日本語歴史コーパス CHJ' and navigation links for 'コーパス選択画面' and 'ログアウト'. The main navigation bar has four tabs: '短単位検索' (Short Unit Search), '長単位検索' (Long Unit Search), '文字列検索' (Text Search), and '位置検索' (Position Search). The '短単位検索' tab is active, showing a search form with a search box, a '検索' button, and various filters. The search form includes a 'キー' field with a dropdown menu, a '語' dropdown, and a checkbox for 'キーの条件を指定しない'. Below the search box is a 'ダウンロードオプション' section with a '設定を表示する' button. The interface is clean and professional, with a light green and white color scheme.

日本語歴史コーパス CHJ

中納言 2.5.2 データバージョン 2021.03

日本語歴史コーパス マニュアル 語数について

現在のサーバ負荷状況: [Progress Bar]

現在 30 人ログイン中

短単位検索 長単位検索 文字列検索 位置検索

短単位検索

検索フォームで検索 検索条件式で検索 履歴で検索

前方共起条件の追加

キー [ ] 語 [ ] キーの条件を指定しない

書字形出現形 [ ] が [ ] 短単位の条件の追加

後方共起条件の追加

検索対象 設定を隠す

検索対象を選択 検索対象をクリア

全て

検索動作 設定を隠す

文脈中の区切り記号 [ ] 文脈中の文区切り記号 [ # ]

前後文脈の語数 [ 20 ] 副本文 [ 副本文を検索対象に含まない ]

共起条件の範囲 [ 文境界をまたがない ]

ダウンロードオプション 設定を表示する

検索 検索結果をダウンロード 条件クリア キャンセル

# データ

## コーパス検索

- ・日本語歴史コーパス (CHJ)  
バージョン2021.3 (中納言2.5.2) 最終アクセス：2021年11月24日

### 検索式

- ・「美しゅうございます」型  
キー: 品詞 LIKE "形容詞%"  
AND 後方共起: 語彙素="御座る" ON 1 WORDS FROM キー  
AND 後方共起: 語彙素="ます" ON 2 WORDS FROM キー

- ・「美しいです」型  
キー: 品詞 LIKE "形容詞%"  
AND 後方共起: 語彙素="です" ON 1 WORDS FROM キー  
キー: 品詞 LIKE "形容詞%"  
AND 後方共起: 語彙素="た" ON 1 WORDS FROM キー  
AND 後方共起: 語彙素="です" ON 2 WORDS FROM キー

### 検索期間

- ・明治から昭和



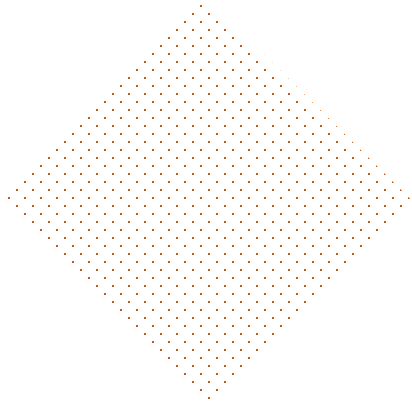
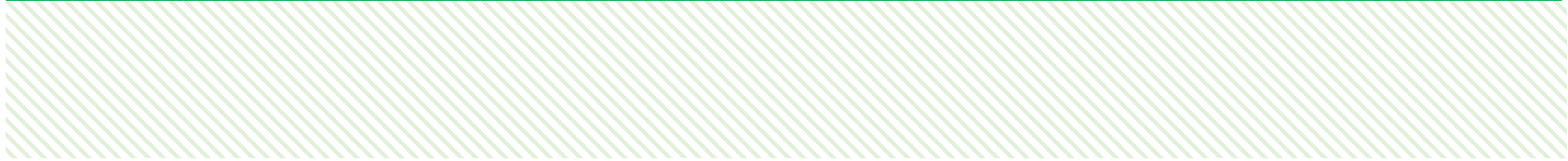
# 統計モデル

## 従属変数

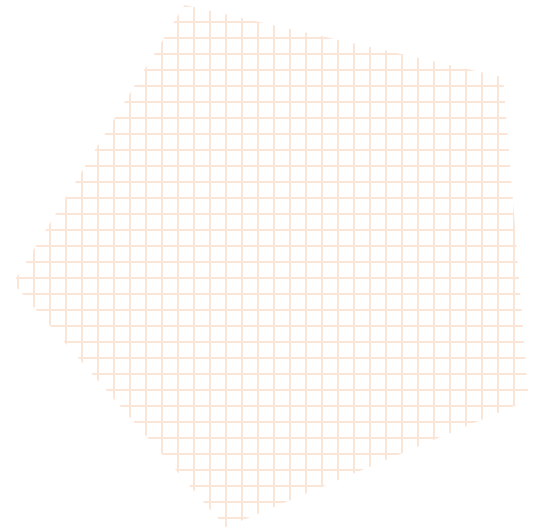
### 離散的選択

- ・ **用いられた変異** $y$  : 二値の値を取る変数。  
新規形式なら1。（「美しいです」型）  
旧形式なら0。（「美しゅうございます」型）

⇒ロジスティック回帰



# 統計モデル



# 統計モデル

## 独立変数

### 固定効果

#### (社会言語学的な要因)

- ・ **教科書** $x_1$  : ダミー変数。**教科書**からのサンプルなら1 (それ以外0)。
- ・ **文学** $x_2$  : ダミー変数。**文学系**のジャンルから採られたら1 (それ以外0)。

#### (言語学的な要因)

- ・ **終助詞** $x_3$  : ダミー変数。**終助詞**と共起していたら1 (それ以外0)。
- ・ **認識的モーダル** $x_4$  : ダミー変数。認識的モーダル「**よう**」-(y)ooと共起したら1 (それ以外0)。
- ・ **過去時制** $x_5$  : ダミー変数。過去形なら1 (それ以外0)。

# 統計モデル

## 独立変数

### 変量効果

(言語学的な要因)

- ・ **形容詞** $u_{0j}$ :  $j$ 番目の形容詞の独自性。 $N(0, \tau^2)$ に従うと仮定。

※変量効果を適切に推定するため、サンプル数が5以上のもの限定し、ここでは、32個のイ形容詞（1031トークン）に限定して考察を進める。

⇒ロジスティック回帰  
＋混合効果

# 統計モデル

## 統計モデル

ここまでの情報を統合すると…

Model

(Version 1 out of 2)

$$y_{ij} \sim \text{Bern}(\pi_{ij})$$

$$\pi_{ij} \sim \text{logistic}(\eta_{ij})$$

$$\eta_{ij} = \beta_0 + \beta_1 x_{1i} + \beta_2 x_{2i} + \dots + \beta_5 x_{5i} + u_{0j}$$

$$u_{0j} \sim N(0, \tau^2)$$

⇒ロジスティック回帰  
+混合効果

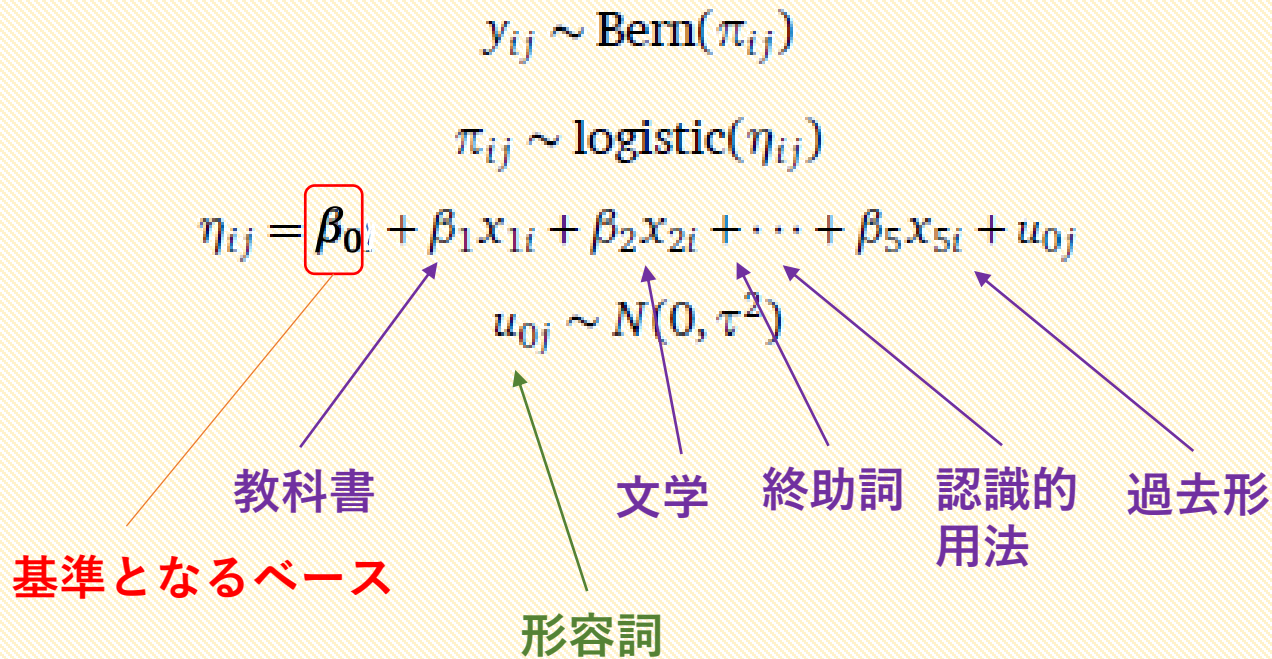
# 統計モデル

## 統計モデル

ここまでの情報を統合すると…

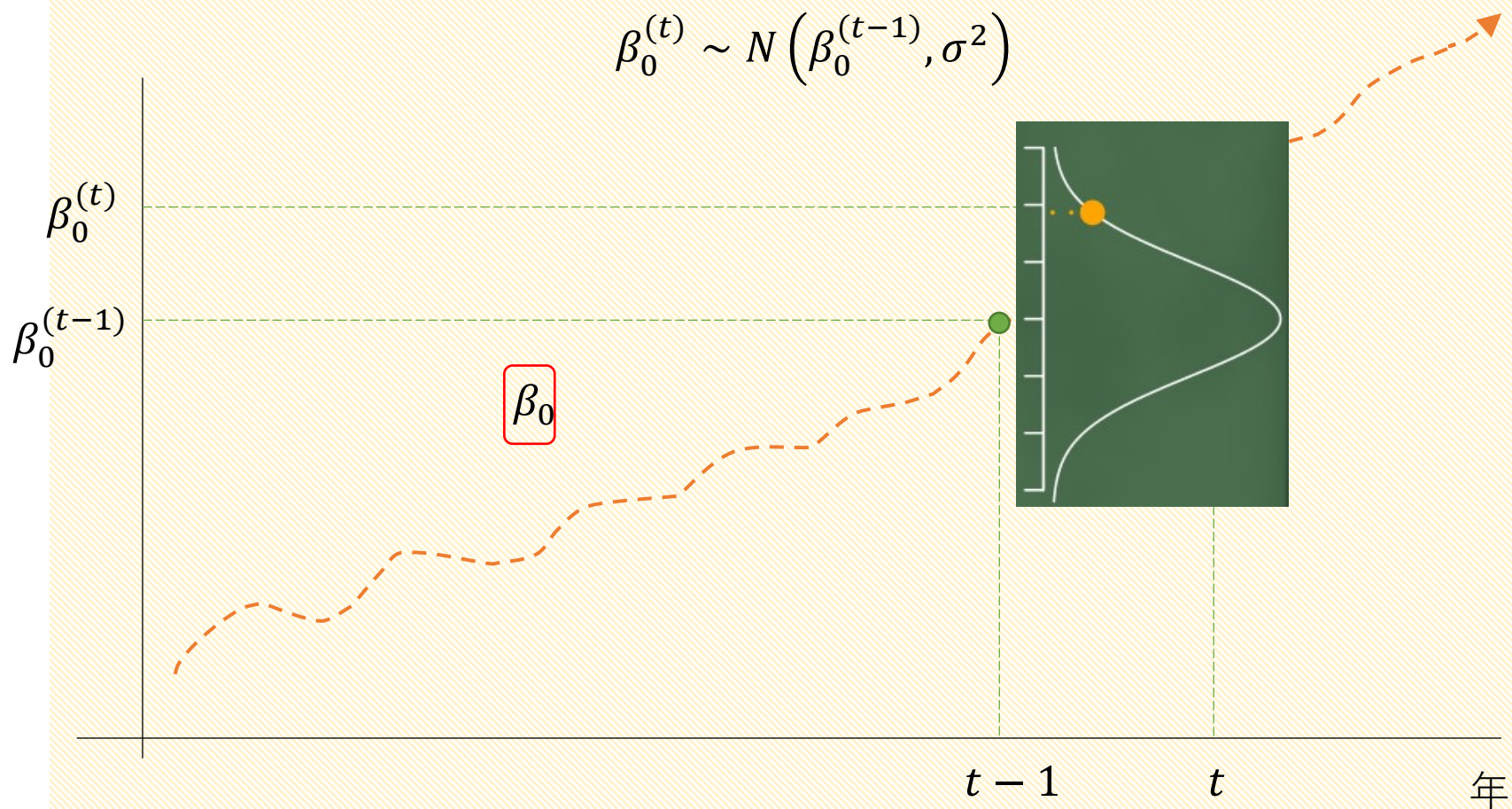
Model

(Version 1 out of 2)



# 統計モデル

## 統計モデル



# 統計モデル

## 統計モデル

ここまでの情報を統合すると…

Model

(Version 2 out of 2)

$$y_{ij}^{(t)} \sim \text{Bern}(\pi_{ij}^{(t)})$$

$$\pi_{ij}^{(t)} \sim \text{logistic}(\eta_{ij}^{(t)})$$

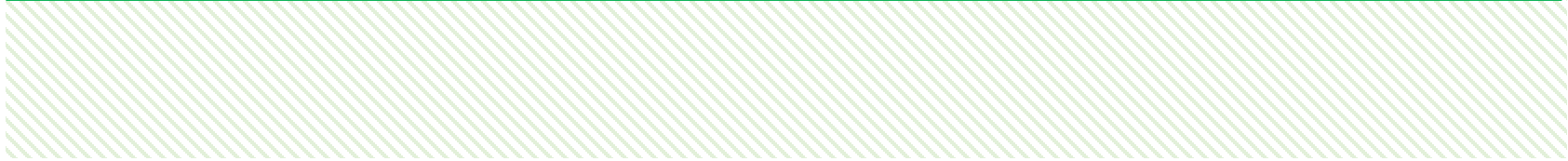
$$\eta_{ij}^{(t)} = \beta_0^{(t)} + \beta_1 x_{1i}^{(t)} + \beta_2 x_{2i}^{(t)} + \beta_3 x_{3i}^{(t)} + \beta_4 x_{4i}^{(t)} + u_{0j}$$

$$u_{0j} \sim N(0, \tau^2)$$

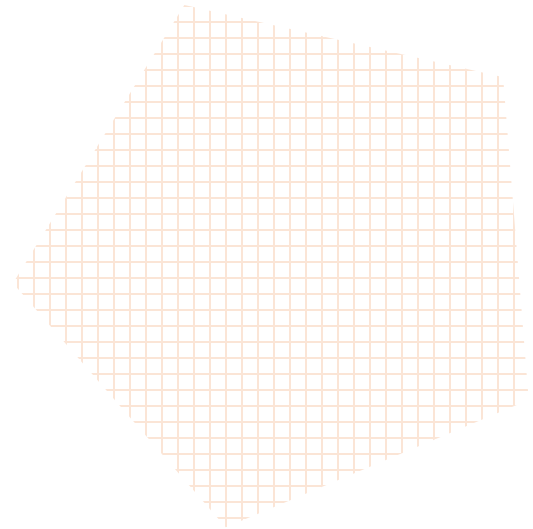
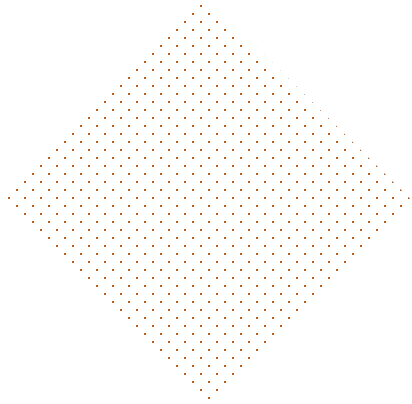
$$\beta_0^{(t)} \sim N(\beta_0^{(t-1)}, \sigma_\zeta^2)$$

⇒ロジスティック回帰  
+ 混合効果  
+ ランダムウォーク





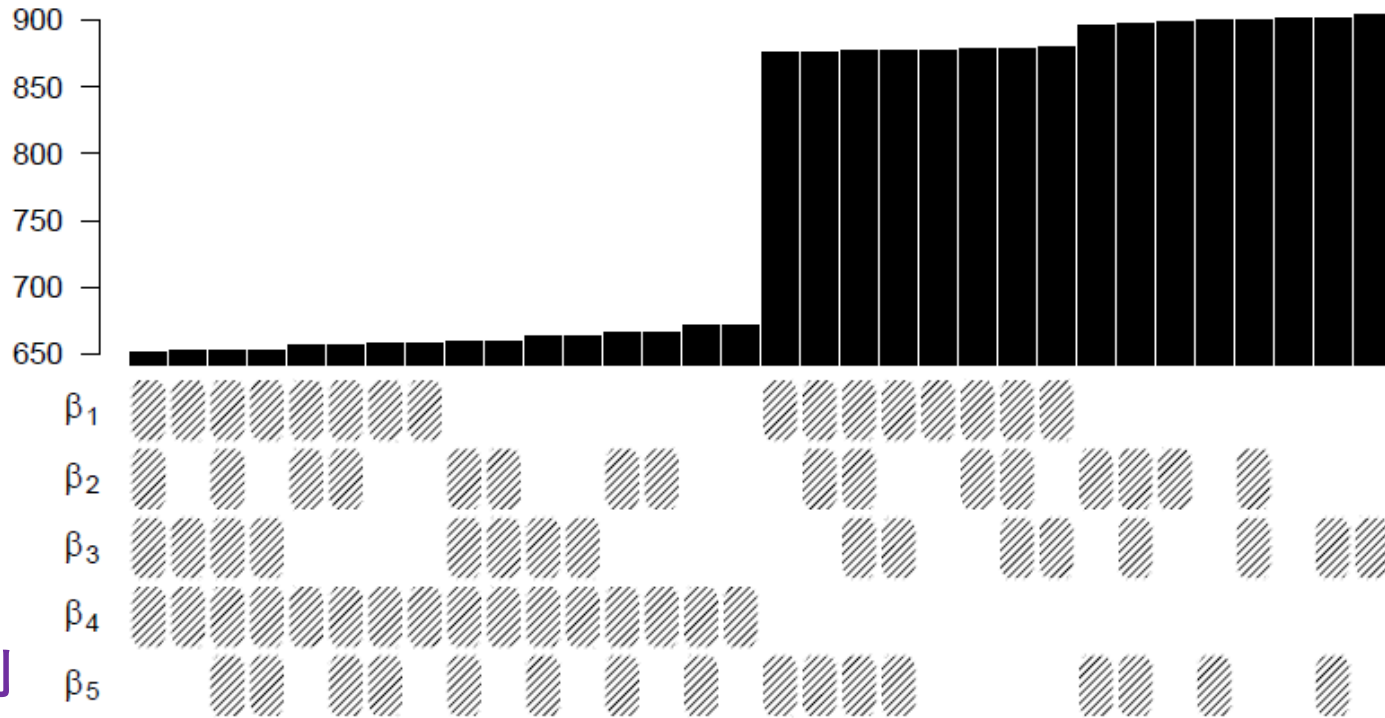
# 結果



# 分析結果

## モデル比較

**Figure 4:** LOOIC: The LOOIC of each model is shown as a barplot in the upper panel; the models are ordered in the increasing order. The barcode plot in the lower panel represents the variables included in the given model. For example, the leftmost model contains the variables from  $x_1$  to  $x_4$  but lacks the last variable.

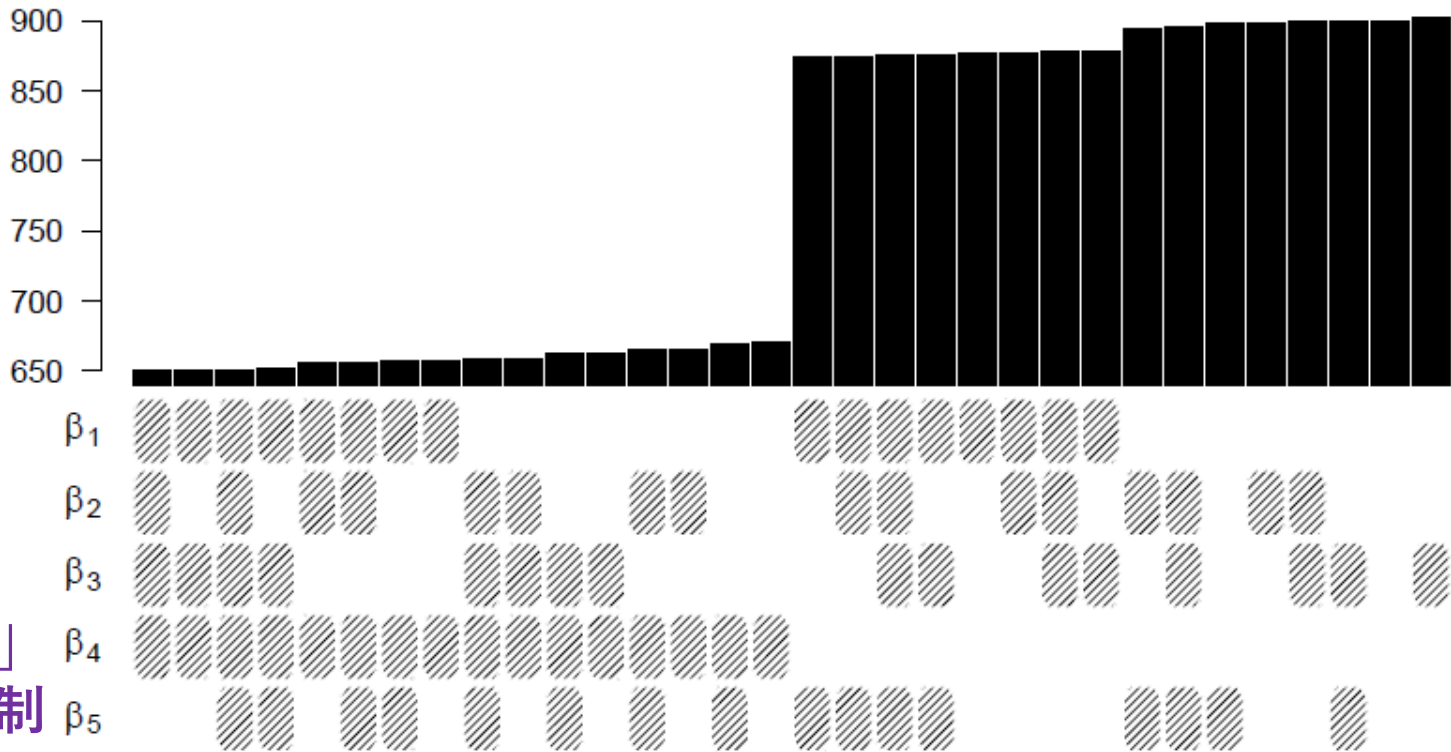


- 教科書
- 文学
- 終助詞
- 「よう」
- 過去時制

# 分析結果

## モデル比較

**Figure 5:** WAIC: The WAIC of each model is shown as a barplot in the upper panel; the models are ordered in the increasing order. The overall tendency is the same as the results in Figure 4



- 教科書
- 文学
- 終助詞
- 「よう」
- 過去時制

# 分析結果

## 統計モデル

Model

(Version 2 out of 2)

$$y_{ij}^{(t)} \sim \text{Bern}(\pi_{ij}^{(t)})$$

$$\pi_{ij}^{(t)} \sim \text{logistic}(\eta_{ij}^{(t)})$$

$$\eta_{ij}^{(t)} = \beta_0^{(t)} + \beta_1 x_{1i}^{(t)} + \beta_2 x_{2i}^{(t)} + \beta_3 x_{3i}^{(t)} + \beta_4 x_{4i}^{(t)} + u_{0j}$$

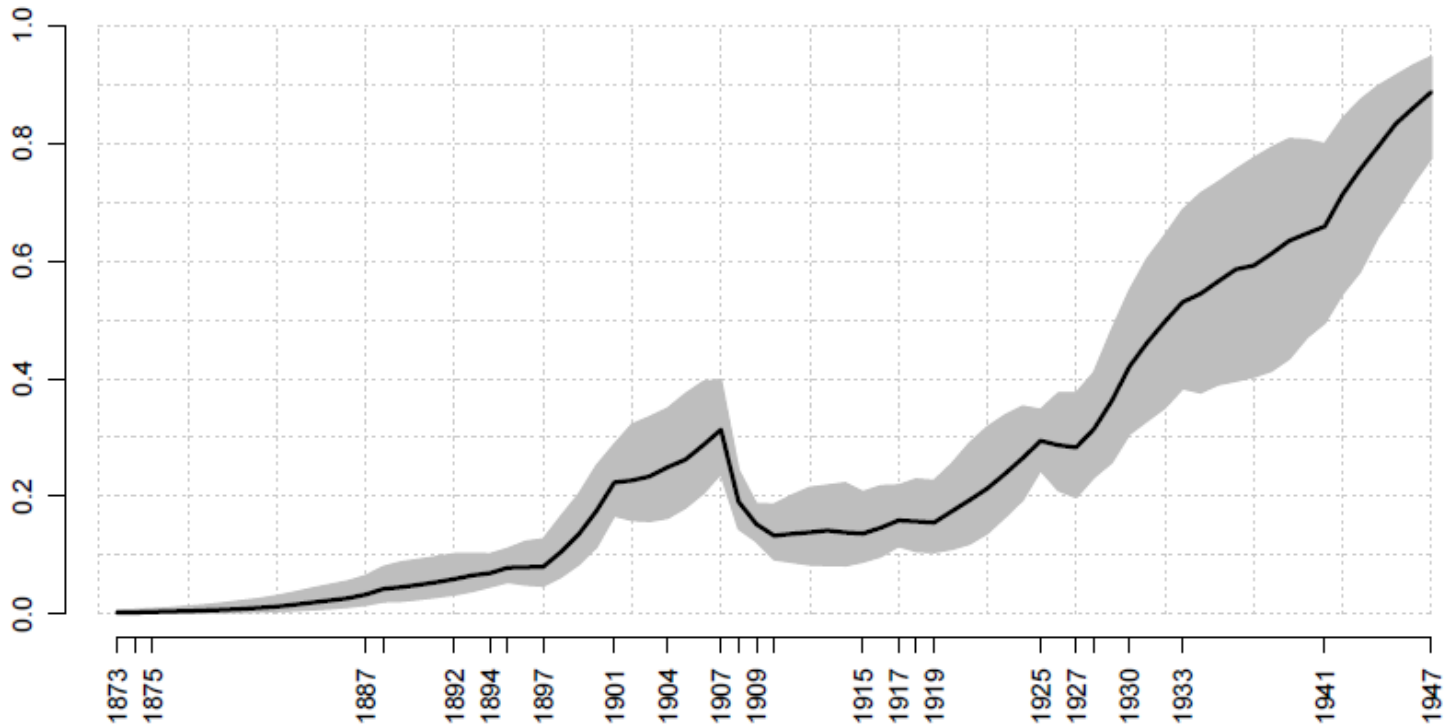
$$u_{0j} \sim N(0, \tau^2)$$

$$\beta_0^{(t)} \sim N(\beta_0^{(t-1)}, \sigma_\zeta^2)$$

- 教科書
- 文学
- 終助詞
- 「よう」

# 分析結果

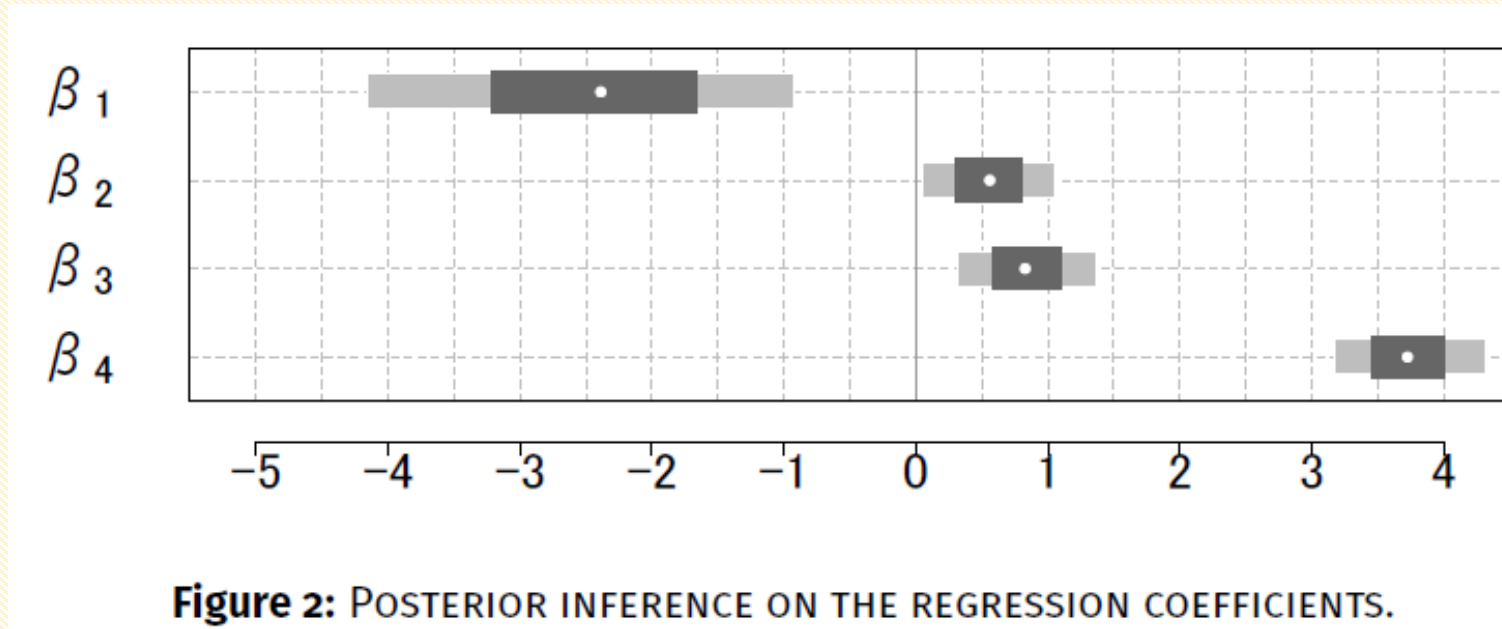
## 統計モデル



**Figure 1:** POSTERIOR INFERENCE ON  $\{\beta_0^{(t)} : t \in \{1873, 1874, \dots, 1947\}\}$ . For each year, the posterior median of  $\beta_0^{(t)}$  (the solid line) and its 95% credible intervals (the shaded gray area) are drawn.

# 分析結果

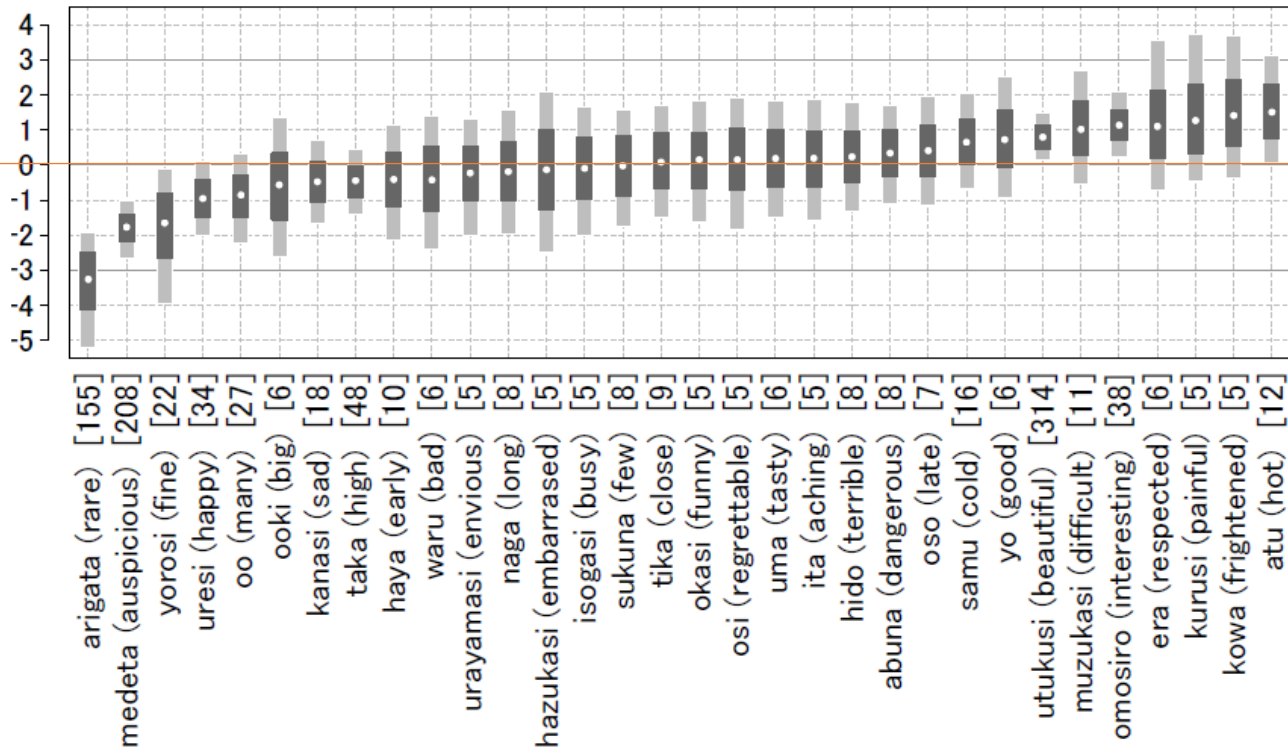
## 統計モデル



- 教科書
- 文学
- 終助詞
- 「よう」

# 分析結果

## 統計モデル



**Figure 3:** POSTERIOR INFERENCE ON THE RANDOM EFFECTS. Each bar represents the 95% and the 66% credible intervals of the random-effects term  $u_{0j}$ . The white circles are the posterior medians. The number in the square brackets shows how many times the adjective is used with an AH form in the CHJ.



---

# まとめと今後の課題



# まとめ

## リサーチクエスチョン

(19)

- a. 新形式はいつごろから用いられてきたのか。
- b. どのような社会言語学的環境で新形式はその使用を拡大させていったのか。
- c. どのような理論言語学的要因で新形式はその使用を拡大させていったのか。
- d. イ形容詞の間に新形式を用いる選好度合いに違いがあったのか。

# まとめ

## 主張

(20)

- a. 19世紀末から新形式はコーパス上に散見され始めるが、大きくその使用を拡大させるのは1930年代前後である。
- b. その表現が生起するジャンルにより構文の選好度合いが変わる。国定教科書では旧形式への指向が高く、文芸では新形式への指向が高くなる。
- c. 認識的モーダル「よう」や終助詞が共起することで新形式への指向が強まる。
- d. 「有難い」「めでたい」という一部のイ形容詞は旧形式でイディオムの意味を持ち、旧形式への依存が高い一方、「良い」（「面白い」）は新形式への指向が他の形容詞に比べると少し強い傾向が見て取れる。

# 今後の課題

## 定量分析から言語理論へ

- (21) a. なぜ「よう」において新規形式への指向が見られたのか？  
b. なぜ「終助詞」において新規形式への指向が見られたのか？

### スキーマの形成

- (5) 走ります (7) 走りません  
*hasiri-mas-u.* *hasiri-mas-en.*  
run-AH-PRS run-AH-NEG  
'(He) runs.' '(He) does not run.'

### 脱カテゴリー化（脱コピュラ化）

suffixの前に「です」を挿入

- (6) 走りました (8) 走りませんでした  
*hasiri-masi-ta.* *hasiri-mas-en* *desi-ta.*  
run-AH-PST run-AH-NEG COP.AH-PST  
'(He) ran.' '(He) did not run.'

- (22) 美しい ですね  
*utukusii* *des-ne.*  
beautiful AH-SFP  
'... is beautiful.'

# 参考文献

---

- Durbin, James and Siem Jan Koopman (2001) *Time Series Analysis by State Space Methods*. Oxford: Oxford University Press.
- Hagiwara, Junichiro (2021) *Time Series Analysis for the State-Space Model with R/Stan*. Singapore: Springer.
- 川口良 (2014) 『丁寧体否定形のバリエーションに関する研究』くろしお出版.
- 國語調査委員会 (編) (1916) 『口語法』 國定教科書共同販賣所. DOI: 10.11501/1870063
- 三上章 (1953) 『現代語法序説—シンタクスの試み—』 刀江書院
- 三尾砂 (1958) 『話しことばの文法』 法政大学出版会.
- R Core Team (2020) R: A Language and Environment for Statistical Computing. R Foundation for Statistical Computing Vienna, Austria. <https://www.R-project.org/>.
- Stan Development Team (2020) Stan Modeling Language Users Guide and Reference Manual. version. 2.21.0, <https://mc-stan.org>.
- Shumway, Robert H. and David S. Stoffer (2017) *Time Series Analysis and its Applications: with R examples*. New York: Springer.
- Vehtari, Aki, Andrew Gelman and Jonah Gabry (2016) *loo: Efficient leave one-out cross-validation and WAIC for Bayesian models*. R package version 0.1.6. <https://github.com/stan-dev/loo>.

# 参考文献

---

Vehtari, Aki, Andrew Gelman and Jonah Gabry (2017) Practical Bayesian model evaluation using leave-one-out cross-validation and WAIC. *Statistics and Computing* 27. 1413–1432.

Vehtari, Aki, Andrew Gelman, Daniel Simpson, Bob Carpenter and Paul-Christian Bürkner (2021) Rank-normalization, folding, and localization: An improved  $\hat{R}$  for assessing convergence of MCMC (with discussion). *Bayesian Analysis* 16(2). 667–718.

Watanabe, Sumio (2010) Asymptotic equivalence of Bayes cross validation and Widely Applicable Information Criterion in singular learning theory. *Journal of Machine Learning Research* 11. 3571–3594.

湯澤幸吉郎 (1953) 『口語法精説』 明治書院.