



時系列データ解析に おけるモデル説明(2)

蘆田 誠人



ARIMAモデル

- 自己回帰モデルと移動平均モデルの合成モデル
- ビッグデータ解析に有用性を増す傾向に



ARIMAモデル

- ARIMAモデルの活用はあまり進んでいない
 - ✓ パラメータ推計に時間がかかる
 - ✓ モデルの構築に人的負担が大きい



実用に向けての改良が進んでいる



定常性とは

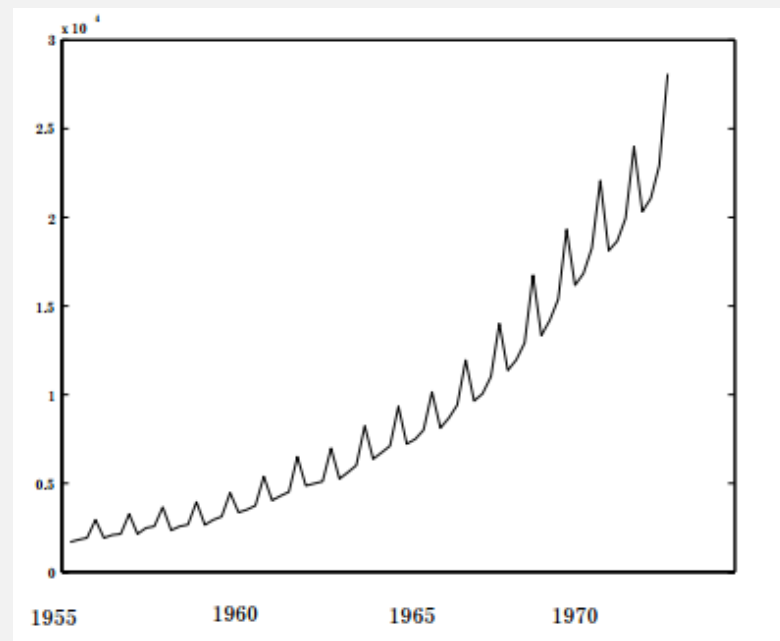
- 時系列データの確率分布は常に同じであり、時間や位置によってその確率過程の性質
- 定常性
サイコロの出目
- 非定常性
1年間の平均気温の推移

非定常データの定常化

データ処理
(修正, 異常値除去)

階差除去

季節階差除去



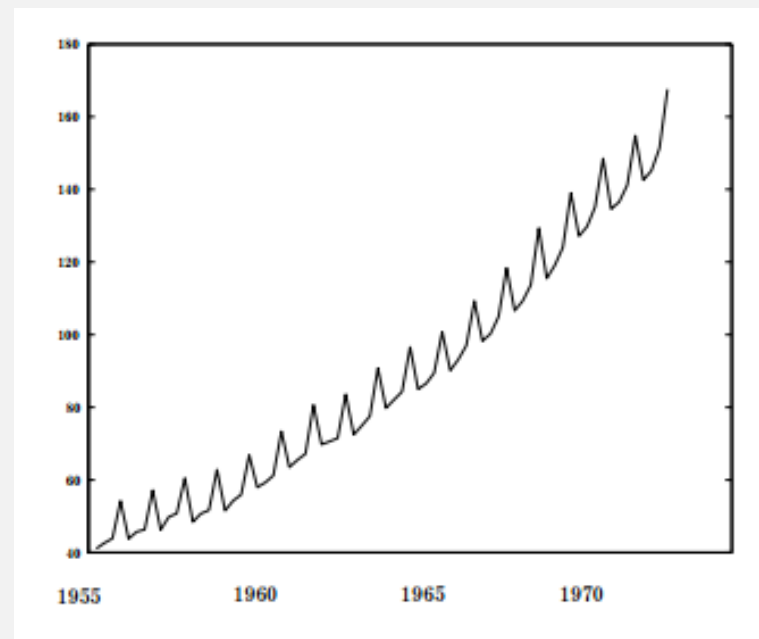
GDPのデータ

非定常データの定常化

データ処理
(修正, 異常値除去)

階差除去

季節階差除去



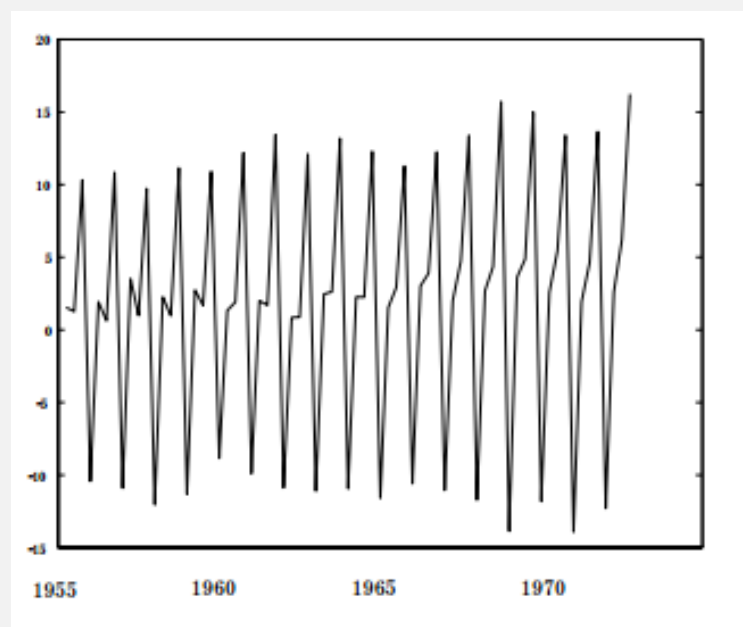
平方根変換後のデータ

非定常データの定常化

データ処理
(修正, 異常値除去)

階差除去

季節階差除去



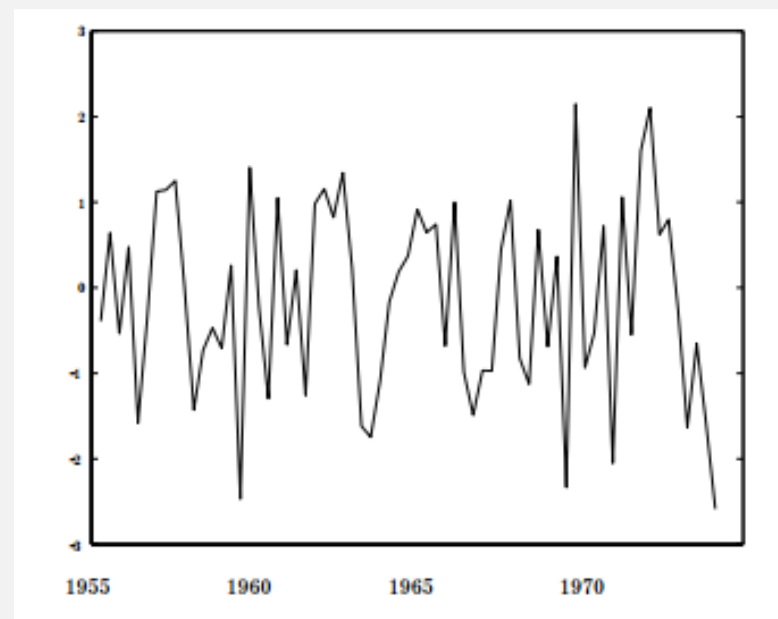
階差除去後のデータ

非定常データの定常化

データ処理
(修正, 異常値除去)

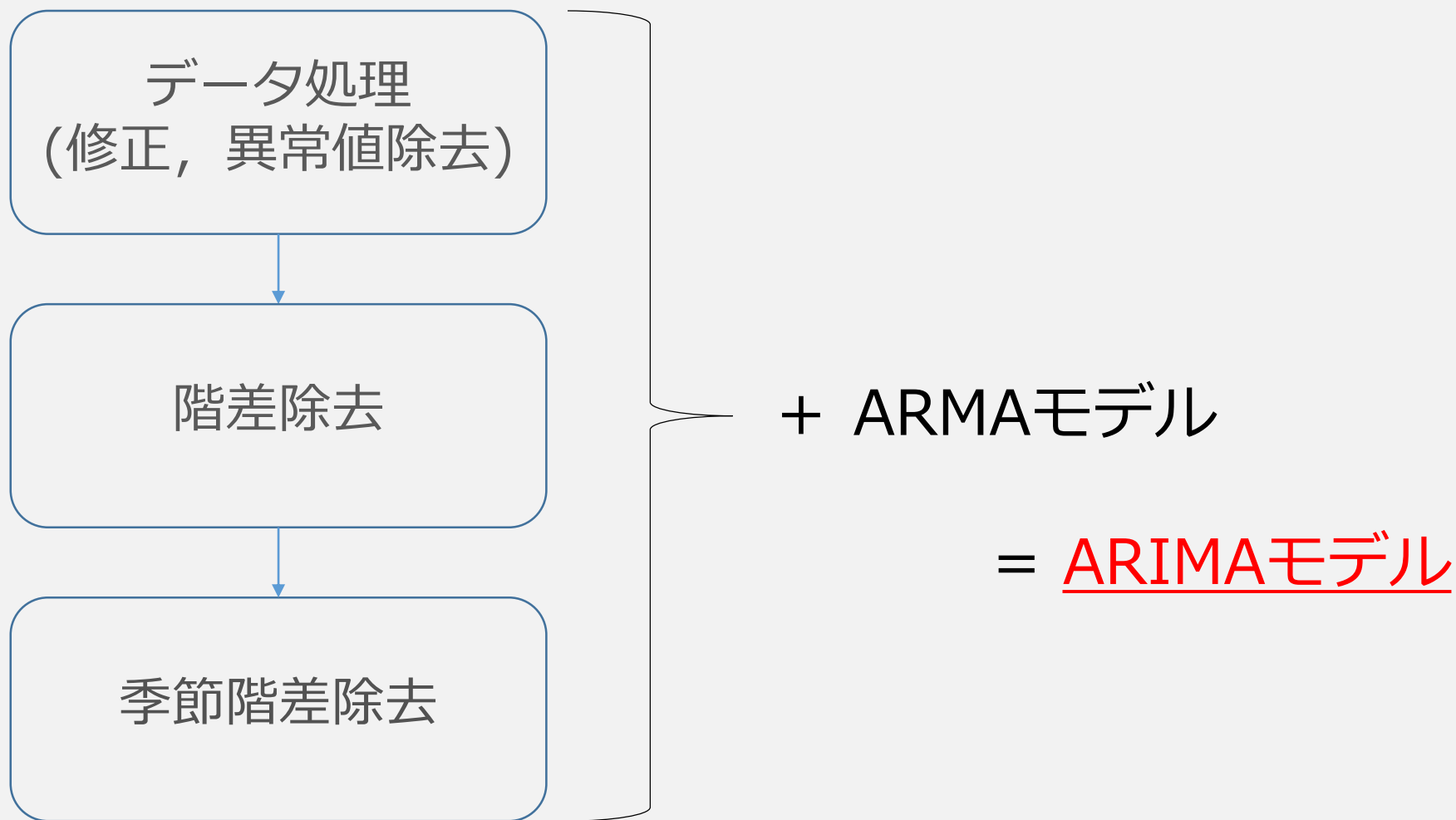
階差除去

季節階差除去



季節階差除去後のデータ

非定常データの定常化





VARモデル

- ARモデル

1変数の時系列データを利用

- VAR(Vector AutoRegressive)モデル

多変量の時系列データを利用

VARモデル

PCとスマホのサイト訪問数

➤ARモデルの場合

明日の訪問数 = a * 今日の訪問数 + 定数 c

- ✓ $a = 0.5$, $c = 1000$ の場合,
- 今日 = 3000 → 明日 = 2500
 - 今日 = 1000 → 明日 = 1500

VARモデル

PCとスマホのサイト訪問数

➤ VARモデルの場合

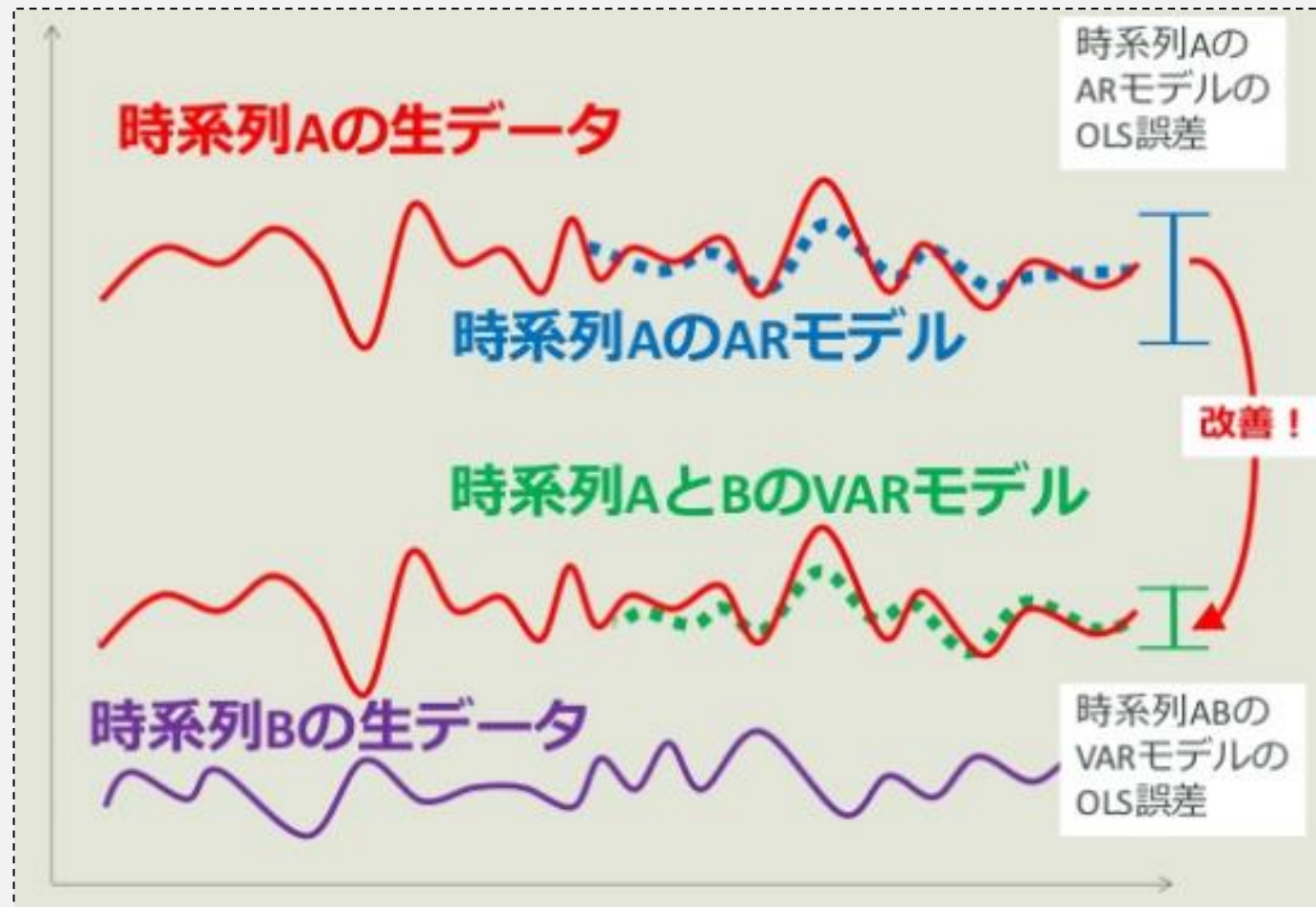
明日のPC訪問数 =

$$a1 * \text{PC訪問数} + b1 * \text{SP訪問数} + \text{定数}c1$$

明日のSP訪問数 =

$$a2 * \text{PC訪問数} + b2 * \text{SP訪問数} + \text{定数}c2$$

グレンジャー因果性



BからAへのグレンジャー因果性が存在する



グレンジャー因果性をRで

キャベツ, たけのこ, 豚肉の因果関係

- ✓ キャベツ, たけのこ, 豚肉の支出データ
- ✓ 階差除去済み
- ✓ Rで実行

たけのことと豚肉からキャベツが予測できるか？

グレンジャー因果性をRで

```
> causality(VAR(x, p=VARselect(x)$selection[1]),cause="キャベツ")  
$Granger
```

```
Granger causality H0: キャベツ do not Granger-cause たけのこ 豚肉
```

```
data: VAR object VAR(x, p = VARselect(x)$selection[1])
```

```
F-Test = 1.2936, df1 = 12, df2 = 105, p-value = 0.2332
```

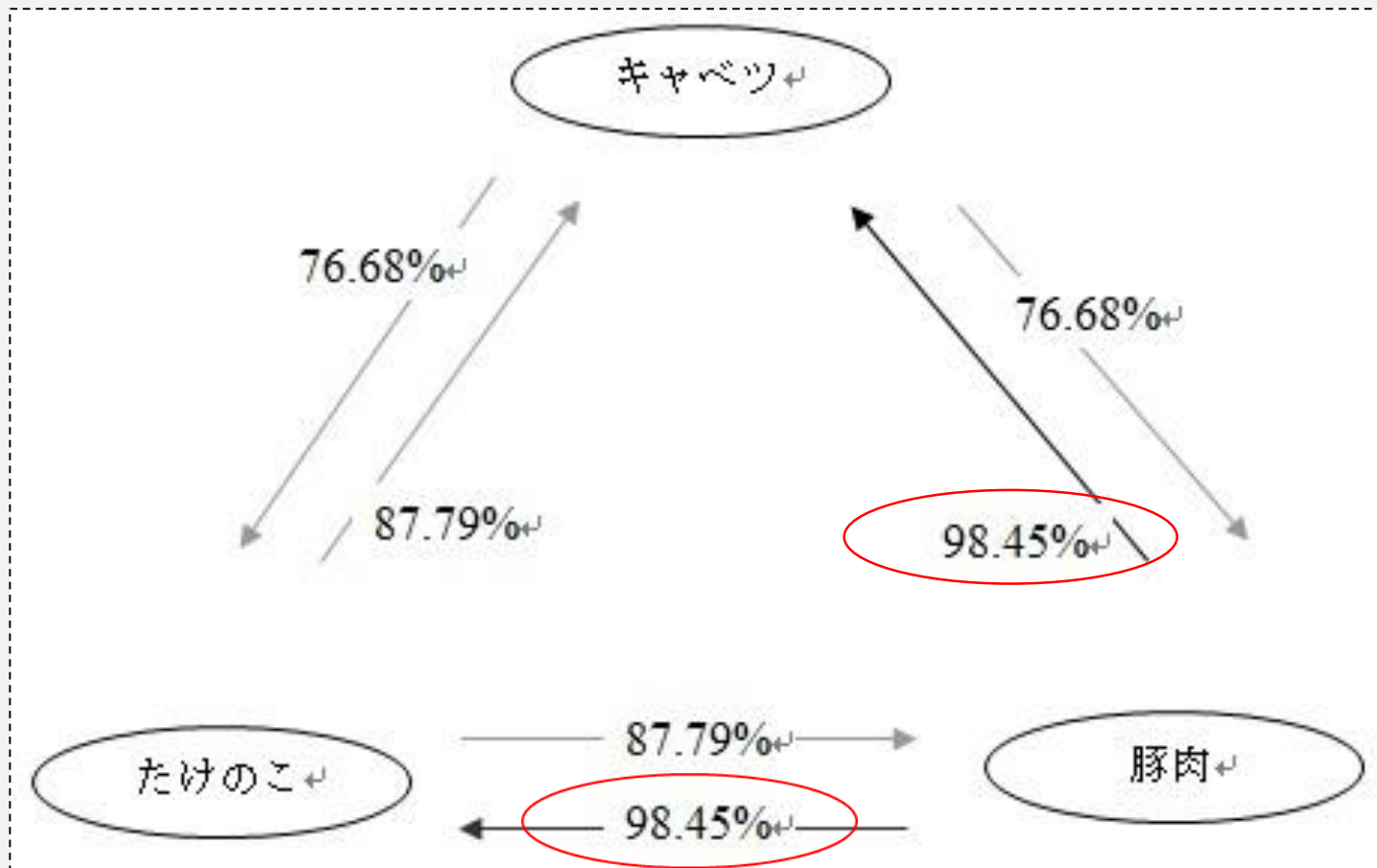
有意水準95%で有意でない

たけのこと豚肉の支出はキャベツに対して
グレンジャー因果を持たない



たけのこと豚肉の支出からキャベツの支出は
予測できない

グレンジャー因果性をRで



豚肉からたけのこととキャベツの予測は可能



まとめ

- ARIMAモデルの改良が進んでいる
- 非定常なデータは定常化
- VARモデルはARモデルの多変量版
- ARよりもVARが誤差が改善されれば, グレンジャー因果性が存在する.