

経済産業省生産動態統計調査 欠測値補完法の妥当性に関する分析結果報告書

令和元年6月11日
総務省 統計研究研修所

目次

	ページ
1. 要旨	2
2. 緒言	3 – 4
2.1 背景	
2.2 目的	
3. 分析方法	5 – 16
3.1 使用データ	
3.2 ジャックナイフ法を応用したシミュレーション	
3.3 計算方法等	
4. 結果及び評価	17 – 20
4.1 データ基礎属性	
4.2 長期に渡る「過去の回答結果を使用した単一補完」の影響評価	
4.3 改善案の検討	
5. 結語	21

1. 要旨

分析課題の整理

・統計委員会による指摘

明確な使用期限なく「過去の回答結果を使用した単一補完」（以下、「LOCF※」。）が行われており、統計精度の悪化に繋がる懸念

↓↓
ポリシーの策定が必要

・欠測値検査の実施

シミュレーションを実施し

- ・補完データの使用期限
- ・より適切な対応方法を報告

※Last Observation Carried Forward

「自動車」に係る調査票を用いて検証

長期LOCFの影響評価

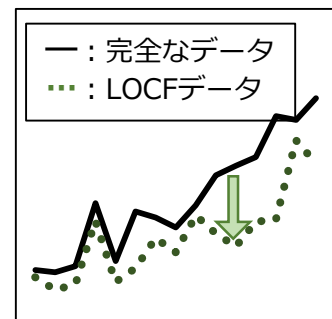


図1.1 LOCFによる乖離のイメージ

完全なデータとの乖離を分析

- ・シミュレーションを行い、完全なデータと、欠測を発生させLOCFしたデータとを比較した
- ・品目によるが、数年であれば顕著な精度の悪化は起きない

精度改善案の提案

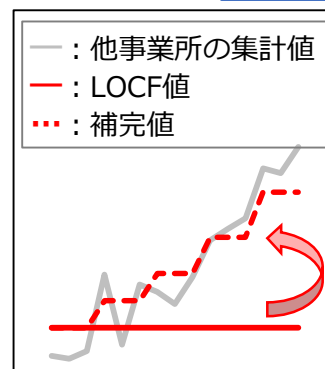


図1.2 伸び率によるLOCF値の補正イメージ

「伸び率」による補正

- ・直近の回答値をベースに他の事業所の伸び率でLOCF値を補正
- ・偏りの抑制・安定化に効果
- ・急激な変動や、事業所の追加又は削除による精度悪化のリスクが存在

2.1. 緒言：背景

◆統計委員会での検討と指摘

- ・平成28年度統計法施行状況に関する審議結果報告書（統計精度検査編）
における指摘事項

経済産業省生産動態統計調査など、一部の統計調査で「明確な使用期限を定めることなく過去の回答結果を使用した単一補完が行われており、長期にわたってデータを使用し続けた場合、統計精度の悪化につながる懸念がある。」旨が指摘

経済産業省所管の調査については「回収率90%以上を保持していることから、致命的な対応とは判断されないが、平成30年度から総務省の支援を得て、計画的にシミュレーションを行いデータの使用期限も含めた補完方法の検証を行うことが必要である。」



- ・平成30年度オプション検査の実施

総務省統計研究研修所が、上記シミュレーションを実施するとともにより適切な欠測値補完方法を提案

2.2. 緒言：目的

◆主目的：LOCFの継続による結果精度への影響を明らかにする

1. LOCFを長期継続した場合の影響評価

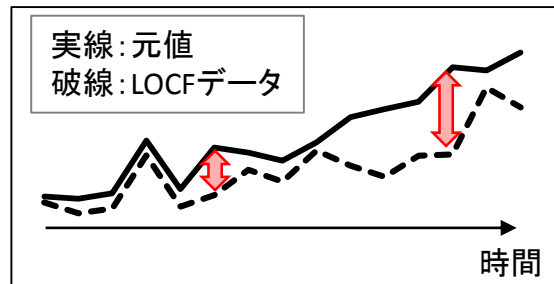


図2.2 LOCFによる乖離のイメージ

有効な回答が得られている事業所の一部に欠測を発生させて、現行の集計方法と同様、LOCFによるシミュレーションを実施した。

図2.2の矢印で示したように、LOCFによる補完値を用いた合計値と、欠測値補完を行っていない本来の値※（以下、「元値」。）との乖離幅を、LOCFを継続した場合の影響として評価する。

※シミュレーションのための審査を行い、データクリーニングを行ったため、公表結果とは必ずしも一致しない。

2. 改善案の検討

上記1. の検討の結果、各品目又は産業ごとの販売金額の水準変化により、LOCFによる補完値と元値との間に継続的な乖離が発生することが予想される。

本研究では、結果精度を改善する方法についていくつかシミュレーションを行い、その効果を比較検討する。

3.1. 分析方法：使用データ

◆調査票及び品目の選定

経済産業省生産動態統計調査には生產品目グループ別に109種類の調査票がある。このうち、以下の理由から「機械器具月報(その40)自動車」を選んだ。

- ア 販売金額が高額
- イ 当該品目を生産又は販売している事業所数が少ない
- ウ 販売金額別の事業所数の分布に偏りがみられる

LOCFの影響を
受けやすいと考えられる

このうち、分析対象期間の時系列データを得ることが出来る、次ページ表3.1に掲げる13調査品目を分析対象とした。

◆項目

主要な指標である「販売金額」をシミュレーションの対象とした。

◆分析対象期間

リーマンショック前後を含む2003年1月分～2017年12月分までの15年分（180か月）とした。

表3.1 分析に使用した品目一覧

品目番号	名称	資料編における略称
0101	軽乗用車	軽自動車
0102	小型乗用車	小型自動車
0103	普通乗用車	普通自動車
0104	バスシャシー（完成車を含む） - 小型バス	小型バス
0105	バスシャシー（完成車を含む） - 大型バス	大型バス
0106	トラックシャシー（完成車を含む） - 軽自動車	トラック（軽自動車）
0107	トラックシャシー（完成車を含む） - 小型自動車 - ガソリン車	トラック（小型・ガソリン）
0108	トラックシャシー（完成車を含む） - 小型自動車 - ディーゼル車	トラック（小型・ディーゼル）
0109	トラックシャシー（完成車を含む） - 普通自動車 - ガソリン車	トラック（普通・ガソリン）
0110	トラックシャシー（完成車を含む） - 普通自動車 - ディーゼル車	トラック（普通・ディーゼル）
0111	トラックシャシー（完成車を含む） - けん引車	トラック（けん引）
0112	特殊自動車	特殊自動車
0113	トレーラ	トレーラ

3.2.1 分析方法：ジャックナイフ法を 応用したシミュレーション

• ジャックナイフ法とは

ジャックナイフ法は、標本の性質を分析するシミュレーション手法である。

図3.2はビン数（除外事業所数）が1つでの概要を示している。この方法は、あるデータセットに対して、以下の操作を行い、母数の推定又は交差検証を行うものである。

- ①ローテーションで1事業所を全期間除外した複数(N個)のデータセットを作成
- ②各データセットの平均や総和、分散などの統計量 (θ) を計算
- ③上記②で求めたN個の統計量から母数($\hat{\theta}$)や標本誤差(S_{θ})等を計算

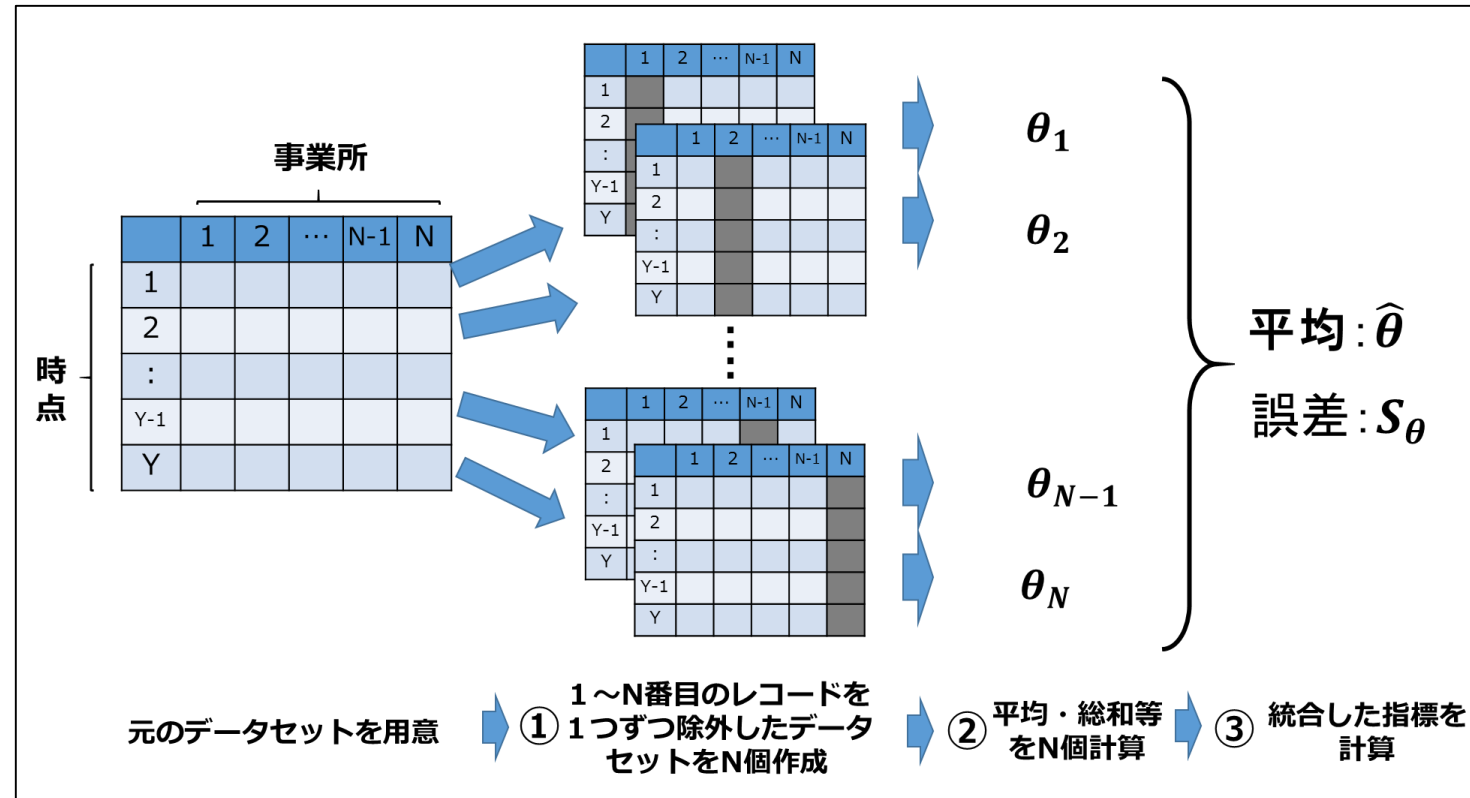


図3.2 ジャックナイフ法概要

◆本分析での応用

本分析では、このジャックナイフ手法を応用し、1事業所を全期間欠測させて、その補完を行った。

シミュレーションには、対象期間（2003年1月～2017年12月の180か月）に販売金額（元値）があり、1年以上のLOCFを行っていない事業所のデータを用いた。13品目の有効回答があった事業所の平均は10.8事業所であった。

1. LOCFを長期継続した場合の影響評価

現行の補完手法であるLOCFを長期間継続した場合の影響を評価した。

(1) 2003年1月の値のLOCFの継続（現行と同様の手法）

2. 改善案の提案

補完法の改善案として、以下に示した平均値代入法及び3通りのLOCFの補正によりシミュレーションを行い、結果を上記1.(1)と比較した。

(2) 平均値代入法※

(3) 年間の伸び率によるLOCFの補正

(4) 3年ごとの伸び率によるLOCFの補正

(5) 前年同月比を使用したLOCFの補正

※現行のLOCFと同様に簡易な手法として、比較対象とした。

3.3.1 分析方法：集計結果の計算方法

・販売金額の集計

以下の式により、評価対象として事業所、年、月別の「販売金額」($S_{i \cdot y \cdot m}$)の年間合計値(X_y)を算出した。ここで求めた X_y は、「元値」に相当する。

$$X_y = \sum_{m=1}^{12} \sum_{i=1}^N S_{i \cdot y \cdot m} \quad \dots \textcircled{1}$$

y ：調査年 m ：調査月, i ：事業所番号 (通番), N ：事業所数
 X_y ：年 y における販売金額の年間合計値, $S_{i \cdot y \cdot m}$ ：第 i 番目の事業所の y 年 m 月における販売金額

・欠測値補完を行ったデータセットの販売金額の集計


以下の式により、欠測値補完を行ったデータセットの販売金額の年間合計値 ($x_{j \cdot y}$) 及びその平均 (\bar{x}_y) を算出した。

$$x_{j \cdot y} = \sum_{m=1}^{12} \left(\sum_{i \neq j} S_{i \cdot y \cdot m} + Z_{j \cdot y \cdot m} (q) \right) \quad \dots \textcircled{2}, \quad \bar{x}_y = \frac{1}{N} \sum_{j=1}^N x_{j \cdot y} \quad \dots \textcircled{3}$$

j ：欠測を発生させ、補完を行う事業所番号, $x_{j \cdot y}$ ：第 j 事業所を欠測値補完した年 y における販売金額の年間合計値,
 $Z_{j \cdot y \cdot m}$ ：第 j 番目の事業所の y 年 m 月における方法 q による補完値

3.3.2 分析方法：補完値の計算方法

欠測とみなした事業所



年/月	1	2	...	M-1	M
2003/01					
2003/02					
:					
2017/11					
2017/12					

図3.3.1 欠測値の生成イメージ

まず、ローテーションで1事業所のデータを全期間欠測させたものを第1～N番目の事業所までN個作成する（図3.3.1）。次に、(1)～(5)の方法により欠測値を補完して、販売額の年間合計値を計算し、N個の年間合計値の平均乖離率を確認する。

以下に掲げるイメージ図（図3.3.2.1～3.3.2.5）では、第2番目の事業所を欠測させた事例について示す。

(1) 2003年1月の値のLOCFの継続

以下の式により、LOCFによる補完値（ $z_{j \cdot y \cdot m(LOCF)}$ ）を算出した。

$$z_{j \cdot y \cdot m(LOCF)} \leftarrow S_{j \cdot 2003 \cdot 01} \quad \dots \quad \textcircled{4}$$

j ：欠測を発生させ、補完を行う事業所番号、

$z_{j \cdot y \cdot m(LOCF)}$ ：第 j 番目の事業所に対する y 年 m 月のLOCFによる補完値

年/月	1	2	...	N-1	N
2003/01		100			
2003/02		100			
:					
2017/11		100			
2017/12		100			




図3.3.2.1 LOCFによる補完イメージ

(2) 平均値代入法

以下の式により、平均値代入による補完値 ($z_{j \cdot y \cdot m(MVS)}$) を算出した。

$$z_{j \cdot y \cdot m(MVS)} \leftarrow \tilde{x}_{j \cdot y \cdot m} = \frac{\sum_{i \neq j} S_{i \cdot y \cdot m}}{N - 1} \dots \textcircled{5}$$

$\tilde{x}_{j \cdot y \cdot m}$: 第 j 番目を除いた事業所の y 年 m 月における販売金額の平均値

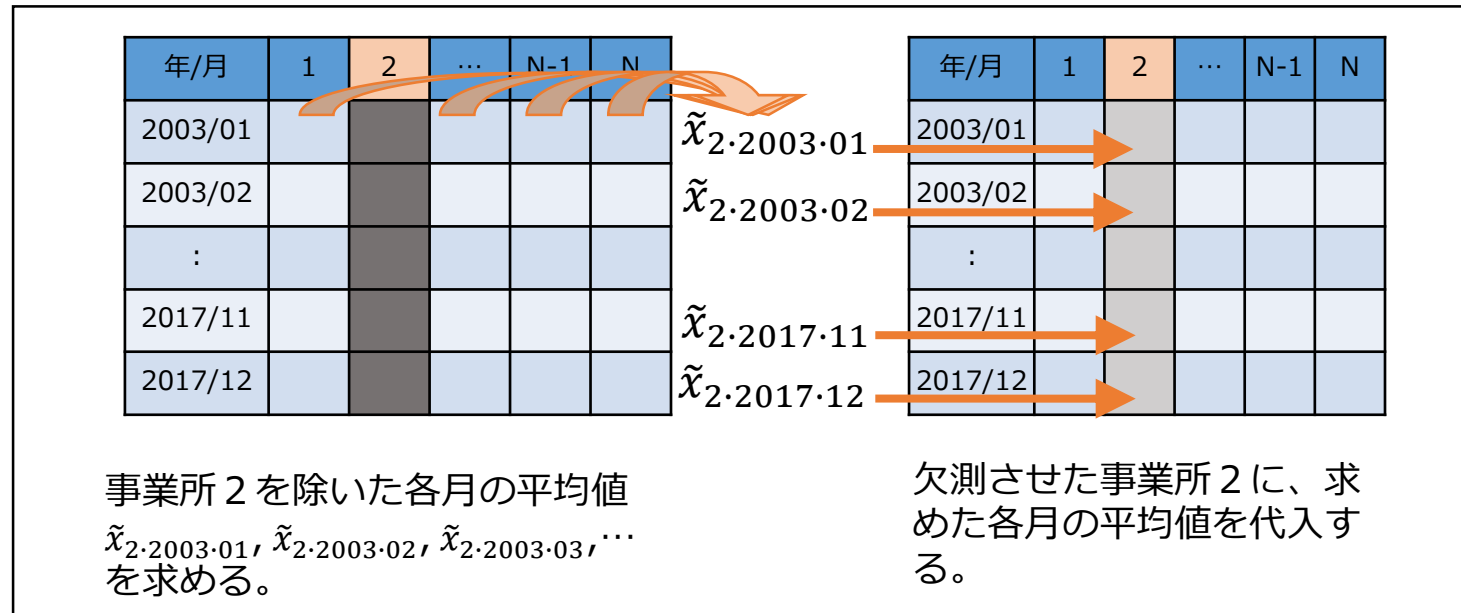


図3.3.2.2 平均値代入法による補完イメージ

(3) 年間の伸び率による補正補完値

以下の式により、伸び率によるLOCFの補正を行った補完値 ($z_{j \cdot y \cdot m(GR1)}$) を算出した。

$$g_{j \cdot y(GR1)} = \frac{\sum_{i \neq j} S_{i \cdot y}}{\sum_{i \neq j} S_{i \cdot y-1}} \quad (\text{ただし } g_{j \cdot 2003(GR1)} = 1) \quad \dots \textcircled{6}$$

$$z_{j \cdot y \cdot m(GR1)} = z_{j \cdot y \cdot m(LOCF)} \times \prod_{k=2003}^y g_{j \cdot k(GR1)} \quad \dots \textcircled{7}$$

$g_{j \cdot y(GR1)}$: 年 y における第 j 番目の事業所を除いて求めた前年比 (伸び率) , $S_{i \cdot y}$: 第 i 番目の事業所の y 年における販売金額合計

年	1	2	...	N-1	N	合計金額	伸び率 $g_{2 \cdot y \cdot (GR1)}$
2003						$\sum_{i \neq 2} S_{i \cdot 2003}$	1
2004						$\sum_{i \neq 2} S_{i \cdot 2004}$	$\frac{\sum_{i \neq 2} S_{i \cdot 2004}}{\sum_{i \neq 2} S_{i \cdot 2003}}$
:							
2016						$\sum_{i \neq 2} S_{i \cdot 2016}$	$\frac{\sum_{i \neq 2} S_{i \cdot 2016}}{\sum_{i \neq 2} S_{i \cdot 2015}}$
2017						$\sum_{i \neq 2} S_{i \cdot 2017}$	$\frac{\sum_{i \neq 2} S_{i \cdot 2017}}{\sum_{i \neq 2} S_{i \cdot 2016}}$

年	1	2	...	N-1	N
2003		100			
2004		$100 \times g_{2 \cdot 2004 \cdot (GR1)}$			
:					
2016		$100 \times g_{2 \cdot 2004 \cdot (GR1)} \times g_{2 \cdot 2005 \cdot (GR1)} \times \dots \times g_{2 \cdot 2016 \cdot (GR1)}$			
2017		$100 \times g_{2 \cdot 2004 \cdot (GR1)} \times \dots \times g_{2 \cdot 2017 \cdot (GR1)}$			

事業所 2 を除いた各年の合計金額 $\sum_{m \neq 2} S_{m \cdot t}$ から伸び率 (前年比) $g_{2 \cdot t \cdot (1)}$ を計算する。

2003年1月値のLOCF値 (例では100) に当該年の年間の伸び率を乗じた値で事業所2を補完する。2005年以降は、前年の補完値に、更に当該年の伸び率を乗じて補完値を算出する。

図3.3.2.3 年間の伸び率によるLOCFの補正イメージ

(4) 3年ごとの伸び率による補正補完値

以下の式により、伸び率によるLOCFの補正を行った補完値 ($Z_{j \cdot y \cdot m(GR3)}$) を算出した。

$$g_{j \cdot y(GR3)} = \frac{\sum_{i \neq j} S_{i \cdot y}}{\sum_{i \neq j} S_{i \cdot y-3}} \quad (\text{ただし } \{g_{j \cdot y(GR3)} \mid y = \left\{ \begin{array}{l} 2004, 2005, 2007, 2008, \\ 2010, 2011, 2013, 2014, \\ 2016, 2017 \end{array} \right\} = 1) \quad \dots \textcircled{8}$$

$$Z_{j \cdot y \cdot m(GR3)} = Z_{j \cdot y \cdot m(LOCF)} \times \prod_{k=2003}^y g_{j \cdot k(GR3)} \quad \dots \textcircled{9}$$

$g_{j \cdot y(GR3)}$: 年 y における第 j 番目の事業所を除いて求めた3年前比 (伸び率)

年	1	2	...	N-1	N	合計金額	伸び率 $g_{2 \cdot y(GR3)}$
2003						$\sum_{i \neq 2} S_{i \cdot 2003}$	1
⋮							
2006						$\sum_{i \neq 2} S_{i \cdot 2003}$	$\frac{\sum_{i \neq 2} S_{i \cdot 2006}}{\sum_{i \neq 2} S_{i \cdot 2003}}$
⋮							
2009						$\sum_{i \neq 2} S_{i \cdot 2003}$	$\frac{\sum_{i \neq 2} S_{i \cdot 2009}}{\sum_{i \neq 2} S_{i \cdot 2003}}$
⋮							

年	1	2	...	N-1	N
2003		100			
2004		$100 \times g_{2 \cdot 2004(GR3)} (=1)$			
2005		$100 \times 1 \times g_{2 \cdot 2005(GR3)} (=1)$			
2006		$100 \times 1 \times 1$ $\times g_{2 \cdot 2006(GR3)}$			
⋮					

2003年を基点とし、(3)同様3年ごとに伸び率を計算す

補完値は3年間同じ値を使用する。

図3.3.2.4 3年ごとの伸び率によるLOCFの補正イメージ

(5) 前年同月比によるLOCFの補正補完値

以下の式により、前年同期比によるLOCF値の補正を行った補完値（ $Z_{j \cdot y \cdot m(RPV)}$ ）を算出した。

$$r_{j \cdot y \cdot m} = \frac{\sum_{i \neq j} S_{i \cdot y \cdot m}}{\sum_{i \neq j} S_{i \cdot y-1 \cdot m}} \quad (\text{ただし } r_{j \cdot 2003 \cdot m} = 1) \dots \textcircled{10} \quad , \quad Z_{j \cdot y \cdot m(RPV)} = Z_{j \cdot y \cdot m(LOCF)} \times \prod_{k=2003}^y r_{j \cdot k \cdot m} \dots \textcircled{11}$$

$r_{j \cdot y \cdot m}$: 第 j 番目の事業所を除いた y 年 m 月における前年同月比

年/月	1	2	...	N-1	N	合計金額	前年同月比 $r_{2 \cdot t \cdot q}$
2003/01						$\sum_{i \neq 2} S_{i \cdot 2003 \cdot 01}$	1
2003/02						$\sum_{i \neq 2} S_{i \cdot 2003 \cdot 02}$	1
:							
2004/01						$\sum_{i \neq 2} S_{i \cdot 2004 \cdot 01}$	$\frac{\sum_{i \neq 2} S_{i \cdot 2004 \cdot 01}}{\sum_{i \neq 2} S_{i \cdot 2003 \cdot 01}}$
2004/02						$\sum_{i \neq 2} S_{i \cdot 2004 \cdot 02}$	$\frac{\sum_{i \neq 2} S_{i \cdot 2004 \cdot 02}}{\sum_{i \neq 2} S_{i \cdot 2003 \cdot 02}}$
:							
2005-01						$\sum_{i \neq 2} S_{i \cdot 2005 \cdot 01}$	$\frac{\sum_{i \neq 2} S_{i \cdot 2005 \cdot 01}}{\sum_{i \neq 2} S_{i \cdot 2004 \cdot 01}}$
2005/02						$\sum_{i \neq 2} S_{i \cdot 2005 \cdot 02}$	$\frac{\sum_{i \neq 2} S_{i \cdot 2005 \cdot 02}}{\sum_{i \neq 2} S_{i \cdot 2004 \cdot 02}}$
:							

年/月	1	2	...	N-1	N
2003/01		100 x 1			
2003/02		100 x 1			
:					
2004/01		100 x $r_{2 \cdot 2004 \cdot 01}$			
2004/02		100 x $r_{2 \cdot 2004 \cdot 02}$			
:					
2005/01		100 x $r_{2 \cdot 2004 \cdot 01}$ x $r_{2 \cdot 2005 \cdot 01}$			
2005/02		100 x $r_{2 \cdot 2004 \cdot 02}$ x $r_{2 \cdot 2005 \cdot 02}$			
:					
2017/12		100 x $r_{2 \cdot 2004 \cdot 12}$ x ... x $r_{2 \cdot 2017 \cdot 12}$			

事業所 2 を除いた合計金額から前年同月比を計算する。

LOCF値（例では100）に前年同月比を乗じた値で事業所 2 を補完する。
2005年以降は、補完値に更に当該年の前年同月比を乗じて補完値を算出する。

3.3.3 分析方法：評価指標の計算方法

以下の式により、 $j = 1 \sim N$ までの N 個全ての事業所を順に欠測させ補完して計算した販売金額の年間合計値 ($x_{j \cdot y}$) ごとに、元値 (X_y) との乖離率 ($e_{j \cdot y}$) を求めた。乖離率は、偏りを評価するための「符号付」(式 ⑫ - a) と、ばらつきを評価するための「絶対値」(式 ⑫ - b) の2通り計算した。

そこから、 N 個のデータセットから求めた乖離率を平均し、平均乖離率 (\bar{e}_y) を求め、評価指標とした。

$$e_{j \cdot y} = \frac{x_{j \cdot y} - X_y}{X_y} \quad \dots \quad \text{⑫ - a} \quad , \quad e_{j \cdot y} = \frac{|x_{j \cdot y} - X_y|}{X_y} \quad \dots \quad \text{⑫ - b}$$

$$\bar{e}_y = \frac{1}{N} \sum_{j=1}^N e_{j \cdot y} \quad \dots \quad \text{⑬}$$

y : 調査年, N : 事業所数, j : 欠測を発生させ、補完を行う事業所番号,
 $x_{j \cdot y}$: 第 j 事業所を欠測値補完した年 y における販売金額の年間合計値, X_y : 年 y における販売金額の年間合計値,
 $e_{j \cdot y}$: 第 j 事業所を欠測値補完した年 y における乖離率, \bar{e}_y : 年 y における N 個のデータセットの乖離率の平均

3.3.4 分析方法：ウインドウ分析

全期間を用いたシミュレーションは、2008年のリーマンショックなど、急激な水準変化による影響を受ける可能性がある。そこで、LOCFの長期継続の影響評価を補強するため、「ウインドウ分析」を行った。

ウインドウ分析では、以下の式に示したとおり、LOCFの開始年 $a = \{2003, 2004, 2005, 2006, 2007, 2008\}$ を1年ずつずらした10年間のLOCFのシミュレーションを、2003年開始から2008年開始まで6セット行い、LOCFの開始後の経過年数 b ごとの平均乖離率の平均 (\tilde{e}_b) を求めた。

$$Z_{j \cdot y \cdot m \cdot a(\text{LOCF})} = S_{j \cdot a \cdot 01} \dots \textcircled{14}, \quad x_{j \cdot y \cdot a} = \sum_{m=1}^{12} \left(\sum_{i \neq j} S_{i \cdot y \cdot m} + Z_{j \cdot y \cdot m \cdot a(\text{LOCF})} \right) \dots \textcircled{15}$$

$$e_{j \cdot y \cdot a} = \frac{x_{j \cdot y \cdot a} - X_y}{X_y} \quad \left(\text{又は } e_{j \cdot y \cdot a} = \frac{|x_{j \cdot y \cdot a} - X_y|}{X_y} \right) \dots \textcircled{16}, \quad \bar{e}_{y \cdot a} = \frac{1}{N} \sum_{j=1}^N e_{j \cdot y \cdot a} \dots \textcircled{17}, \quad \tilde{e}_b = \frac{1}{6} \sum_{a=2003}^{2008} \bar{e}_{(a+b) \cdot a} \dots \textcircled{18}$$

- j : 欠測を発生させ、補完を行う事業所番号, a : LOCFの開始年,
- $S_{i \cdot y \cdot m}$: 第 i 番目の事業所の y 年 m 月における販売金額
- $Z_{j \cdot y \cdot m \cdot a(\text{LOCF})}$: 第 j 番目の事業所の a 年から開始したLOCFによる y 年 m 月の補完値
- $x_{j \cdot y \cdot a}$: 第 j 事業所を a 年から開始したLOCFにより欠測値補完したときの年 y における販売金額の年間合計値,
- $e_{j \cdot y \cdot a}$: 第 j 事業所を a 年から開始したLOCFにより欠測値補完したときの年 y における元値との乖離率,
- $\bar{e}_{y \cdot a}$: 年 y における N 個のデータセットの $e_{j \cdot y \cdot a}$ の平均
- \tilde{e}_b : 年 a から b 年経過した時の $a = 2003 \sim 2006$ の6個の平均乖離率の平均値

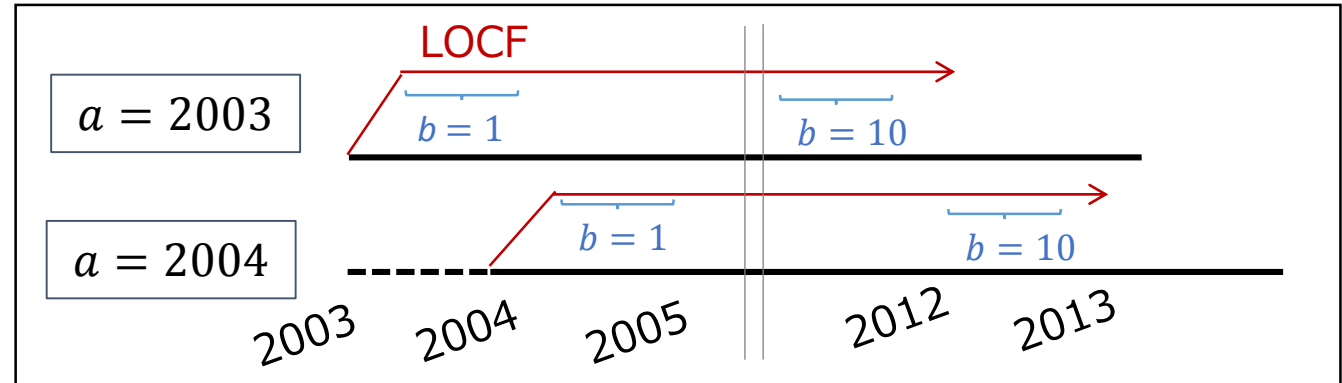


図3.3.3 ウインドウ分析の模式図

4.1. 結果及び評価：データの基礎属性

「機械器具月報(その40)自動車」について、以下の特徴を確認した。

- ア 有効回答があった事業所数は、8から17事業所（平均10.8、標準偏差3.1）だった。
- イ 事業所間の販売金額の分布に偏りがみられた（図4.1.1）
- ウ 今回LOCFに用いた1月の季節指数は、0.86である。（図4.1.2及び資料編2章）



今回分析した品目は、少サンプルかつ販売金額の偏りが大きい。また、LOCFに用いる2003年1月は12か月で最も小さい値である。ここから、本分析で行うシミュレーションは、乖離率が比較的大きい値をとりやすい（結果が振れやすい）といった条件の悪いケースであるといえる。

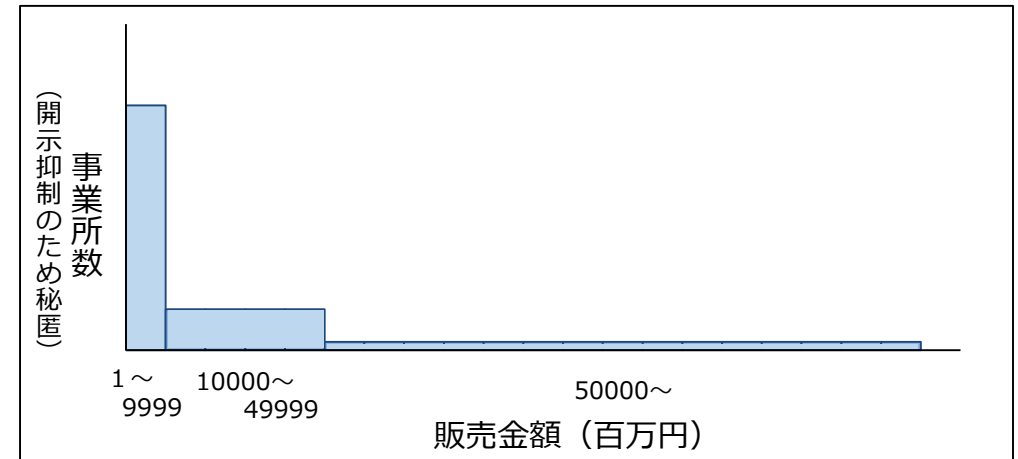


図4.1.1 トレーラ（例）の販売金額階級別事業所数

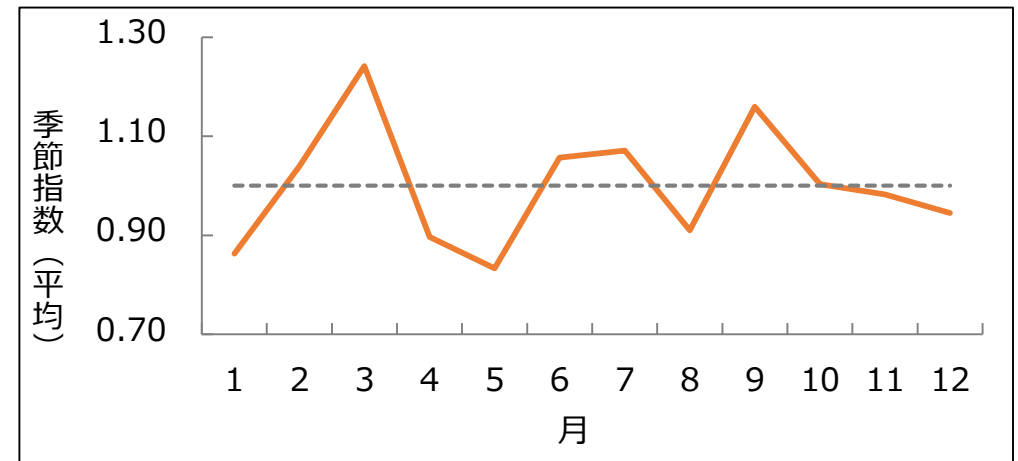


図4.1.2 販売金額の季節指数（2003-2017年の月別平均の全品目平均）

※季節指数はX-12-ARIMAにより算出した。条件は資料編2章の参考を参照

4.2.結果及び評価：長期LOCFの影響評価

現行手法と同様にLOCFを長期間継続した場合のシミュレーション結果の特徴は以下のとおり（資料編1、4、5章）。

ア 平均乖離率（符号付）は、販売金額にほぼ反比例する（全ての品目でPearsonの相関係数が-0.90より小さい）。（図4.2.1）

イ 平均乖離率（符号付）は大きく変動せず（図4.2.1）、平均乖離率（絶対値）も5年程度のLOCFであれば5%程度に収まった。（図4.2.2）

LOCFの開始時点の金額から、品目全体の金額の水準が離れるほど乖離が大きくなる。
数年程度のLOCFでは顕著に結果が悪化することはない※。

※ 「小型バス」、「トラック（けん引）」や「特殊自動車」のように、全体の水準が急激に変化する場合、顕著な乖離が発生しやすい。詳しくは資料編の個別品目を参照。

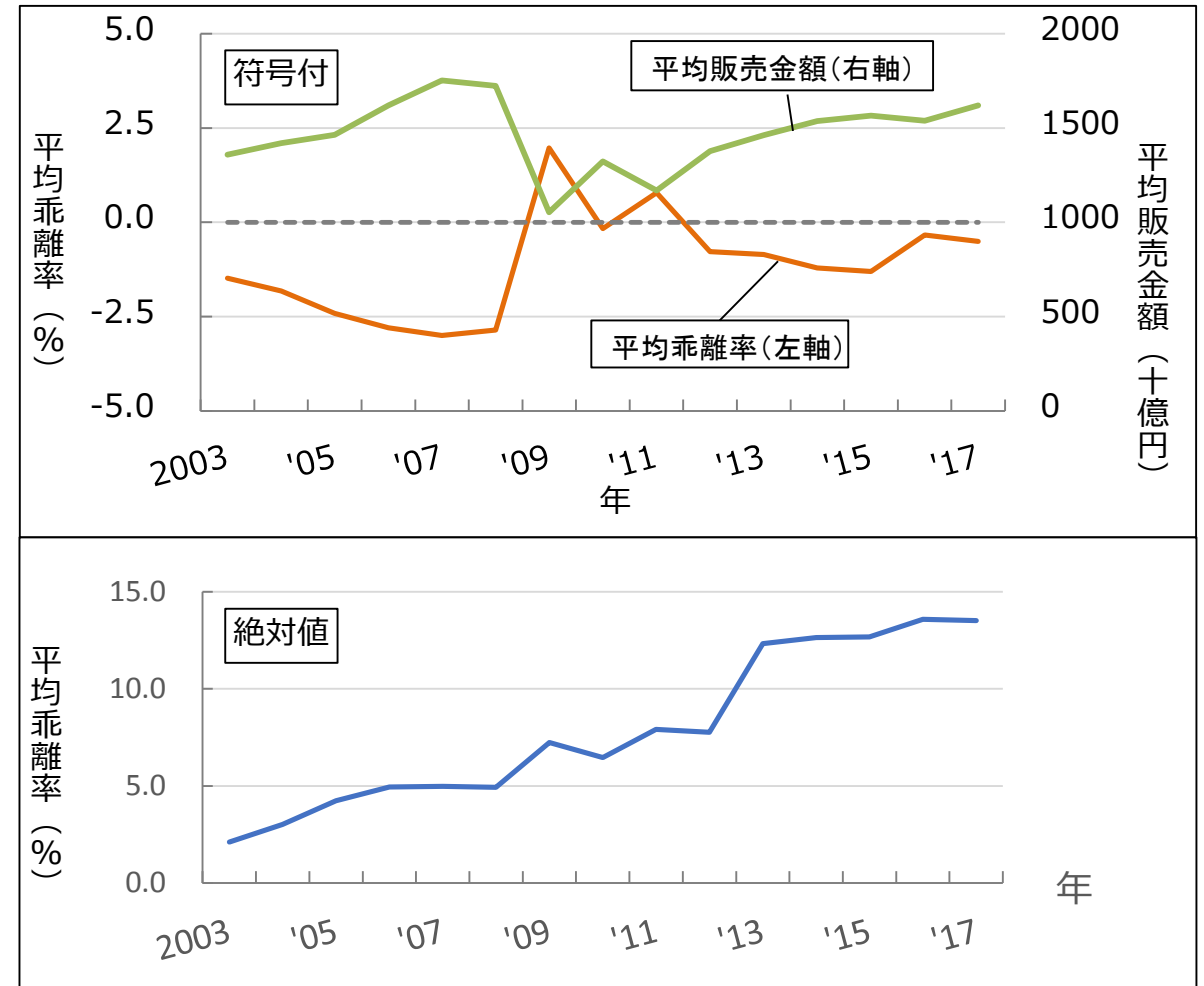


図4.2.1 LOCFの継続による販売金額及び平均乖離率の全品目平均の推移

4.2. 長期LOCFの影響評価（つづき）

更に、ウィンドウ分析によって得られた結果は以下のとおり。（図4.2.2及び資料編6、7章）

- ア 継続期間が10年程度であれば、平均乖離率（符号付）は5%以内の水準に収まった（「特殊自動車」を除く）。
- イ 平均乖離率（絶対値）は、5年以内でも5品目で5%を上回った。

平均乖離率（符号付）の結果から、ほぼすべての品目で、LOCFにより発生する偏りは結果精度に重大な影響を及ぼすことはないと考えられる。

しかし、平均乖離率（絶対値）の結果から、LOCFを長期継続した場合、上記結果がばらつく可能性があり、注意が必要である。特に、「特殊自動車」のように、金額の水準に極端な違いが生じる場合、LOCFによる補完には限界がある。

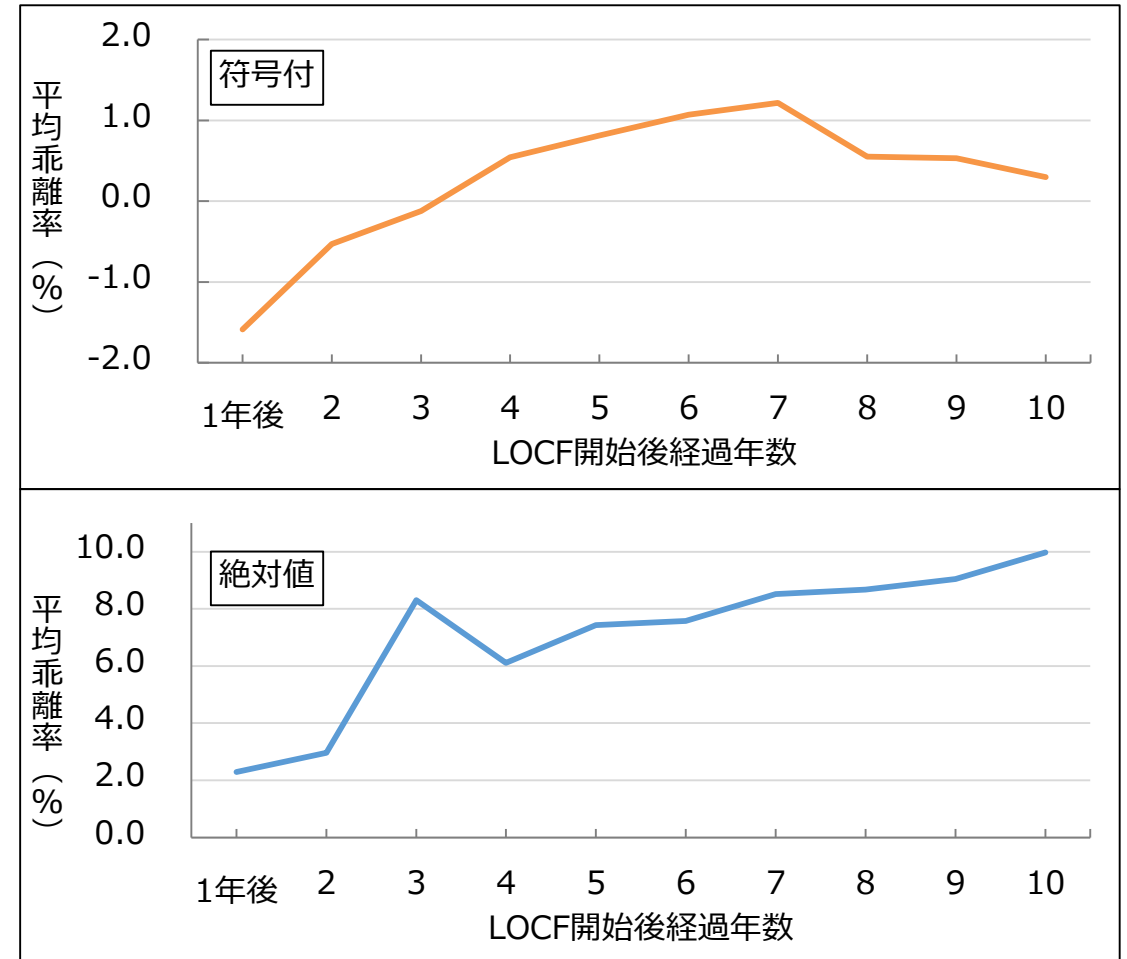


図4.2.2 ウィンドウ分析によるLOCF開始後経過年数による平均乖離率の全品目平均の推移

4.3 結果及び評価：改善案の検討

各種改善法を用いたシミュレーションにより得られた示唆は以下のとおり。（図4.3及び資料編4, 5章）。

ア 平均乖離率（符号付）は、3種類の補正方法を用いることでLOCF開始時の水準が維持された。（平均値代入法では計算の性質から0となった。）

イ 平均乖離率（絶対値）は、平均値代入法で高水準となり、他の3種類の補正方法ではLOCFと差が見られなかった。

ウ いずれの指標でも、一部の品目で2013年付近に断層が発生し、その後乖離率の推移が平坦となる事例が見られた。（データ上事業所交代の影響とみられることを確認）



平均値代入法は、事業所間の偏った販売金額の分布から、極端な値を持つ事業所の数値に影響されやすく、欠測値補完法としては不適切。

伸び率等による補正は劇的な改善効果は見られないが、「特殊自動車」のように、金額の水準の極端な変化から大きな乖離が発生する場合におけるLOCFの補正に効果的と考えられる。

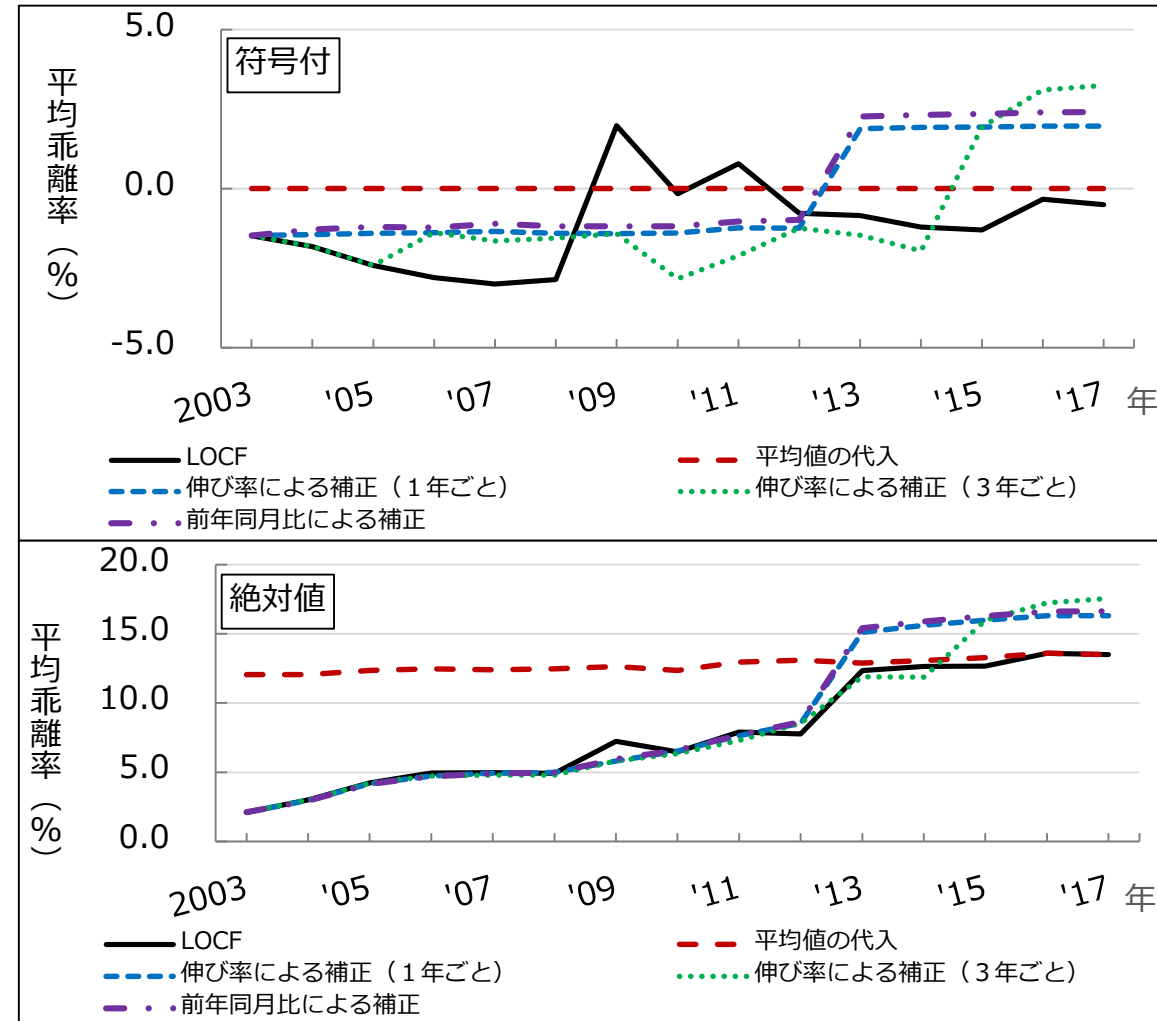


図4.3 各種補完法による平均乖離率の全品目平均

5. 結語

(1) LOCFの長期継続について

- ア 長期に渡り継続しても必ずしも精度が下がる訳ではない。
- イ 販売金額の水準が大きく変化する場合、補完値は元値から大きく乖離する。
- ウ 数年程度LOCFを継続した場合、販売金額の水準の変化を確認するとともに、補正の要否を検討することが望ましい。

(2) 改善案の提案について

伸び率等で、金額の水準の変化に合わせてLOCF値の補正を行うことで、真値からの乖離を一定程度に抑えることができることが分かった。特に、販売金額の水準が大きく変化した場合、この補正は効果的である。