

## ESRI統計より：景気統計

## 機械受注統計調査における欠測値補完方法の検討について

総務省統計局統計調査部国勢統計課労働力人口統計室<sup>†</sup>

高野 正博

近年、統計調査を取り巻く環境が総じて厳しさを増す中で、統計精度を向上させる観点から、欠測値処理について関心が高まっている。一般的な欠測値の補完方法としては、調査客体の過去の回答値を利用する方法（横置き補完）や、類似した属性を持つ他の客体の回答値を加工して利用する方法などが挙げられる。

本稿では、機械受注統計調査における欠測値処理のあり方について検討する。同調査は、機械製造業者が製造する設備用機械類の毎月の受注状況に関する調査であり、設備投資の動向を早期に把握し、経済動向分析の基礎資料を得ることを目的としている。調査対象は機械等を製造する企業のうち、調査対象機種の大分類ごとにカバレッジが80%程度となるよう選定された280社をベースにしている。調査票回収率は現在ほぼ100%であるが、未回収企業が発生した場合には、当該企業の過去の回答値を利用する方法（横置き補完）を採用している。

## 1. 欠測企業のモデル分析

最も適切な補完方法を検討するには、どのようなメカニズムで欠測が発生しているかを調べる必要がある。このため、2005年4月から2011年9月のデータを使用して、欠測状況と企業属性との関係をロジット・モデルによって分析し<sup>1</sup>、過去に欠測となった企業の

特徴を把握する。ここでは企業属性として、上記期間内における需要者別及び機種別の受注状況を設定した。その結果、需要者別には統計的に有意な傾向は見られなかったものの、機種別には運搬機械、産業用ロボットを製造する企業がそれぞれ5%、1%の水準でプラスに有意となり、欠測しやすい傾向のあることが示唆された<sup>2</sup>。

## 2. 補完方法の違いによる精度の比較

次に、近年の経済環境の変化を考慮して、現行の横置き補完とは異なる方法である伸び率補完を検討した。伸び率補完とは、調査票が提出された企業の回答値を基に、前期、または欠測企業が最後に回答した時点から現在までの伸び率を算出し、欠測した企業の過去の回答値にその伸び率を乗じた値を補完する方法である。本稿では横置き補完と伸び率補完の精度を比較<sup>3</sup>するため、2011年1月から12月の回答が得られたデータに対して、4種類の補完方法（Ⅰ：前年同月値で補完、Ⅱ：前年同月値×伸び率<sup>4</sup>で補完、Ⅲ：前月値で補完、Ⅳ：前月値×伸び率で補完）を適用し、実際の回答値と補完値との誤差率を機種別に計測した。具体的な手順は以下のとおりである。

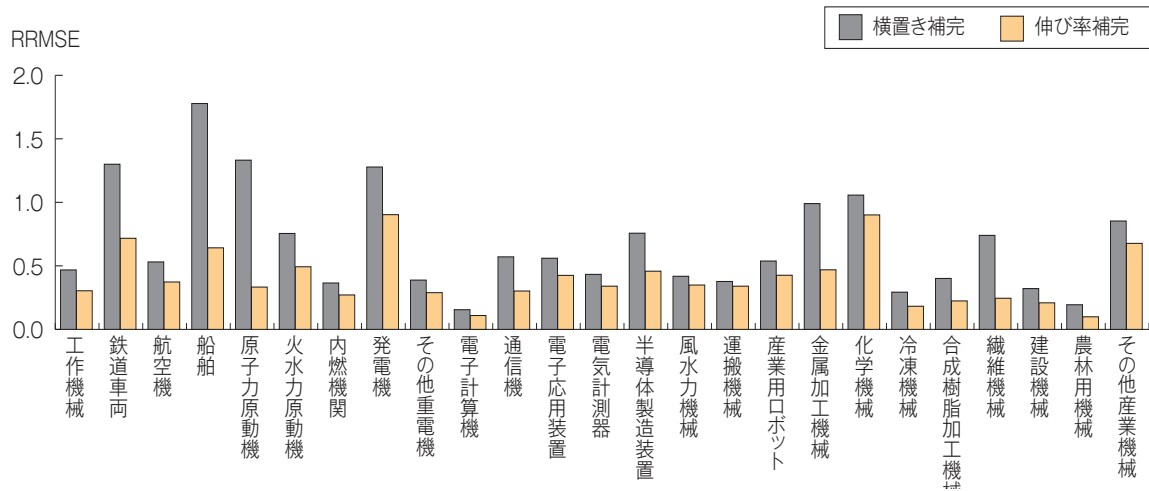
- ①データの抽出 : 対象となるデータを機種別に抽出
- ②欠測値の作成 : 回答値を仮想的に欠測化<sup>5</sup>
- ③補完方法の適用 : 欠測値に上記の補完方法（Ⅰ～Ⅳ）を適用
- ④精度の比較 : 回答値と補完値の誤差率（RRMSE<sup>6</sup>）を評価

図表1は前年同月値による横置き補完（Ⅰ）と伸び率補完（Ⅱ）の誤差率を比較した結果である。機種によって誤差率の減少幅に違いはあるものの、全ての機種で横置き補完よりも伸び率補完を適用した方が、誤

<sup>†</sup> 平成25年3月まで経済社会総合研究所景気統計部研究専門職に在職。

1 目的変数を「当月の欠測状況」（欠測＝1、非欠測＝0）、説明変数を「需要者分類（36区分）または機種分類（27区分）別の受注状況」、「前月の欠測状況」、「前々月の欠測状況」としてモデルを構築した。受注状況は、期間内に受注があった場合を1、受注がなかった場合を0に設定した。  
 2 係数の推定値は、運搬機械が1.52（標準誤差0.73）、産業用ロボットが2.40（標準誤差0.92）、前月の欠測状況が2.23（標準誤差0.25）、前々月の欠測状況が1.59（標準誤差0.27）、定数項が-7.22（標準誤差0.77）となった。  
 3 一般に横置き補完では、補完が容易であるという長所の反面、外部環境変化が反映されないなどの短所が指摘されている。伸び率補完では、外部環境変化を反映できる反面、欠測企業の属性変化が反映されないなどの短所がある。  
 4 前年同月および前月からの伸び率は、回答が得られた企業を、調査年月、機種分類、企業規模（機種別年間受注額の四分位に基づいた区分）で層に分類した上で算出した。  
 5 欠測化の処理に際しては、各企業が同一の確率で欠測になると仮定した。

図表1 前年同月値による横置き補完（Ⅰ）と伸び率補完（Ⅱ）の誤差率比較<sup>7</sup>（機種別）



差率が減少する結果を示した。また、前月値を使用した横置き補完（Ⅲ）と伸び率補完（Ⅳ）を比較した場合についても同様の結果が得られた。なお、前年同月値と前月値との比較では、総じてみると受注額に明確なトレンドがある機種では前年同月値の誤差率の方が小さかった。

参考文献

宇都宮浄人・園田桂子『『全国企業短期経済観測調査』における欠測値補完の検討』『日本銀行ワーキングペーパーシリーズ』、2001、Working Paper 01-11  
 平川貴大・鳩貝淳一郎「ビジネスサーベイにおける欠測値補完の検討—全国企業短期経済観測調査（短観）のケース—」『日本銀行ワーキングペーパーシリーズ』、2012、No.12-J-8

高野 正博（たかの まさひろ）

コラム

景気基準日付について（第14回景気動向指数研究会の結果概要）

内閣府では、各景気循環における経済活動の比較などのため、主要な経済指標の中心的な転換点である景気基準日付（景気の山、谷）を決定しており、これまで戦後14回の景気循環を設定してきた。

今般、第15循環の景気の山の暫定設定に関して、平成25（2013）年8月21日に第14回景気動向指数研究会を開催した。同研究会では、①景気の波及度を示すヒストリカルDIは平成24（2012）年5月から同年11月まで50%を下回っており（過半の指標が下降）、同年4月が山の候補となる、②これについては、平成24（2012）年4月以降のCI一致指数の下降率は過去の後退局面と比して量的変化が十分であるとともに、平成21（2009）年3月の谷からの拡張期間も37ヶ月と過去平均（36.2ヶ月）を上回っており、景気の山とみなすに十分である、③平成23（2011）年3月の東日本大震災に伴う経済活動の急速な低下は、経済活動の収縮が経済の大半の部門に持続的に波及する景気後退局面とはみなされない等に関する検証結果を基にご議論いただいた。

この結果、研究会での結論に従い、第15循環の景気の山を平成24（2012）年4月に暫定的に設定することとした。

6 RRMSE (Relative Root Mean Squared Error) : 回答値をx、補完値をx̂、回答値の平均値をE(・)、標準偏差をσで表したとき、次の式で計算される。

$$RRMSE = \frac{\sqrt{E\{(x - \hat{x})^2\}}}{\sigma}$$

7 前年同月および前月からの伸び率は、回答が得られた企業を、調査年月、機種分類、企業規模（機種別年間受注額の四分位に基づいた区分）で層に分類した上で算出した。