

中国 GDP の 3 つの代理変数に関する研究

～重回帰分析の観点から～

Study on three proxy variables of China's GDP: From the perspective of multiple regression analysis

薛 晓燕[†], 岡崎 一浩^{††}
Xiaoyan Xue[†], Kazuhiro Okazaki^{††}

Abstract

The aim of this study is to examine the usefulness of multiple regression analysis by conducting an analysis using Chinese GDP as an objective variable and the Li Keqiang index, namely, electricity usage by the industrial sector, railway freight quantity, and growth of bank loans as explanatory variables.

Firstly, we used correlation analysis to determine the correlation between our objective variable and explanatory variables. We found the explanatory variables electricity usage by the industrial sector and growth of bank loans to have a strong correlation with objective variable GDP.

However, this correlation analysis found only weak correlation between railway freight quantity and GDP. From this study, we learned that if we add road freight quantity to railway freight quantity, using the resulting freight quantity as an explanatory variable, there was a strong correlation with GDP. We also found that when a revised version of the Li Keqiang index is used, there is a very strong correlation between these three explanatory variables and GDP. Additionally, we considered multicollinearity while conducting this study.

The fact that these three new variables—specific economic activities—each showed strong correlation with China's GDP suggests that GDP statistics accurately reflect economic conditions. Therefore, Chinese GDP statistics are useful. We also believe that these three new variables are proxy variables for GDP.

1. はじめに

本研究では、重回帰分析から説明変数と目的変数の相関を検討するため、中国の GDP よりも信頼できると言われる経済指標「工業電力消費量、鉄道貨物輸送量、銀行融資残高の伸び」という、いわゆる「李克強 3 指数」¹⁾を分析した²⁾。まず、重回帰分析を行う有

用性を検討するため、相関分析によって各説明変数と

[†] 愛知工業大学大学院 経営情報研究科 (名古屋市)
^{††} 愛知工業大学 経営情報研究科 (名古屋市)

¹⁾ 李克強が 2007 年に米国大使に「道貨物輸送量、銀行融資残高、電力消費の推移を見ている」と語ったとされることから来ている。シティバンク等の国際金融機関も可用性を認めている。

<https://ja.wikipedia.org/wiki/%E6%9D%8E%E5%85%8B%E5%B0%B7%E6%8C%87%E6%95%B0>

²⁾ 三尾幸吉郎によれば中国の GDP に影響される説明変数は工業生産、製造業 PMI、非製造業 PMI、小売売上高、工程資産投資、輸出、電力消費量、貨物運送量、工業生産社出荷価格、通貨供給量であるとされる。
<http://www.nliresearch.co.jp/report/detail/id=53981&pno=4?site=nli>

目的変数の相関を検討し、説明変数としての工業電力消費量と銀行融資残高の伸びは目的変数である GDP との相関が強いことを確認した。

しかし、同じ相関分析から鉄道貨物輸送量と GDP の相関が弱いことがわかり、本研究では鉄道貨物輸送量と道路貨物輸送量を加えて「貨物輸送量」として GDP の説明変数とすれば相関が高いことが分かった。修正を施した新 3 指数を用いれば、これら 3 つの説明変数と GDP とには非常に強い相関が認められることがわかった。さらに本研究では多重共線性にも考察を加えた。

また、中国の GDP は新 3 変数具体的な経済活動とそれぞれ非常に強い相関が認められることから、GDP 統計自体も正しく経済実態が表れているとも考えられる。つまり、これら新 3 変数の動きは中国 GDP の代理変数との相関が高いと考えられる。従って、中国の GDP 統計自体にも信頼性があると考えられる。これら新 3 変数は GDP の代理変数としての有用性があると認められる。

本研究では統計的重回帰分析の手法でよりよい説明変数と目的変数の相関が簡単に調べることができることを示している。そして、中国の旺盛な消費などを考慮にいれると、GDP 統計を全面的に否定することには無理があると考えられ、中国の GDP 統計が虚偽に満ちて、有用性がないとするいくつかの説を支持しない。

2. 先行研究と研究目的

谷口 (2015) [9], 谷口 (2015) [10] では中国の GDP と鉄道貨物輸送量及び電力消費量の関係、鉄道輸送量と電力消費量の中国全国と地区の動向などから、「李克強 3 指数」の信頼性が疑わしいと論じている。高橋 (2015) [11] では中国の GDP 統計は全て正しくないと指摘している³⁾。Peter Cai⁴⁾ (2016) では「李

克強 3 指数」は現在中国の経済には意味がないと論じている。

本研究では 2006 年から 2016 年まで中国の GDP と「李克強 3 指数」のデータを用いて重回帰分析でそれらの相関を再検討する。そこで、中国 GDP 統計自体の信頼性及び 3 指数は GDP の代理変数としての有用性について検討を加えた。

3. 研究フロー

3・1 データ収集

表 1 で使用している工業電力消費量データは中国全国工業電力統計快報各年⁵⁾から収集し、鉄道貨物輸送量、高速道路輸送量のデータは中国統計年鑑各年版⁶⁾から収集した。銀行融資残高は百度文庫銀行融資残高の伸びデータ各年⁷⁾から収集した。

表 1 2006 から 2016 年まで元データ

	説明変数			目的変数
	鉄道貨物輸送量 (百万トン)	銀行融資残高の伸び (千億元)	工業電力消費 (千万 KWH)	
2006 年	2,882	31,800	211,540	31,654
2007 年	3,142	48,000	245,660	36,149
2008 年	3,304	49,100	258,640	39,620
2009 年	3,333	95,900	266,650	43,265
2010 年	3,643	75,000	308,860	47,853
2011 年	3,933	74,700	346,330	52,400
2012 年	3,904	85,000	360,610	56,539
2013 年	3,967	88,900	384,710	60,949
2015 年	3,358	112,000	403,300	69,911
2016 年	3,332	126,500	413,830	74,595

3・2 「李克強 3 指数」の妥当性の再検討

単相関分析によって重回帰式を求める意義があるか否かを検討する。目的変数に相関が強いと思われる説明変数が複数ある場合、重回帰分析とする。

単相関分析 (銀行融資残高の伸びと GDP)

³⁾ 高橋洋一:『中国 GDP の大嘘』に以下の説明がある。「中国の統計システムは、社会主義国家の「先輩」であるソビエト連に学んだ。…」 「ソ連国内で五〇年間も使用され続け、デタラメな統計を生み出して来た。…」 「官僚主義の問題と偽造統計システムの手法は、それをそっくり導入した社会主義国家としての「後輩」である中国にも引き継がれている。…」 中国当局が発表する統計データや経済指標はおしなべて信用できないと論じている。…」 (PP28-56) 高橋は『李克強指数が中国の経済失速を示しているという事実が多分にあるからに違いない。…」 「中国の地域ごと産業構造が違うなど状況が異なっている。…」 「李克強指数は中国経済の動きを正確なデータとして捉えるのにはふさわしくないとしている。…」 (PP81-82)。

⁴⁾ Peter Cai 「The Li Keqiang Index Is Dead, Long Live the New Li Keqiang Index」 (2016)
<http://nationalinterest.org/blog/the-buzz/the-li-keqiang-index-dead-long-live-the-new-li-keqiang-index-15469>

⁵⁾ www.docin.com/p-1481195583.html

⁶⁾ <http://money.163.com/17/0120/20/CB8H80LU002580S6.html>
www.askci.com/news/data/2011/08/0215134045152.shtml

⁷⁾ <https://ameblo.jp/kisaragisearchina/entry-12157843185.html>

www.phaj.or.jp/distribution/data/2013china.pdf

図1により、2006年から2016年まで銀行融資残高の伸びとGDPの単相関には強い正相関が見られる⁸⁾

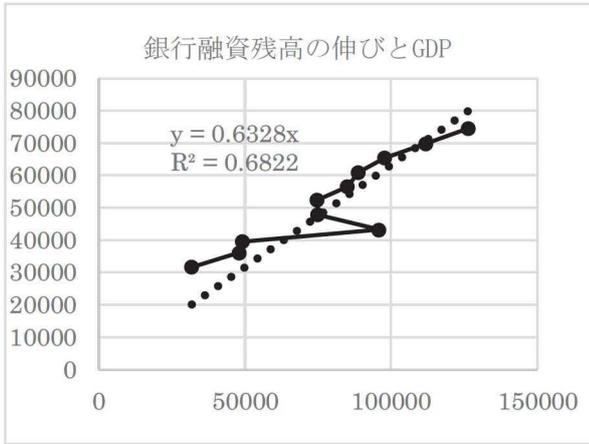


図1 銀行融資残高の伸びとGDPの単相関図 筆者作成

単相関分析 (工業電力消費量の伸びとGDP)

図2により、2006年から2016年まで工業電力消費量もGDPとともに直線的に上がっており、この2個の変数には強い正相関が見られる。

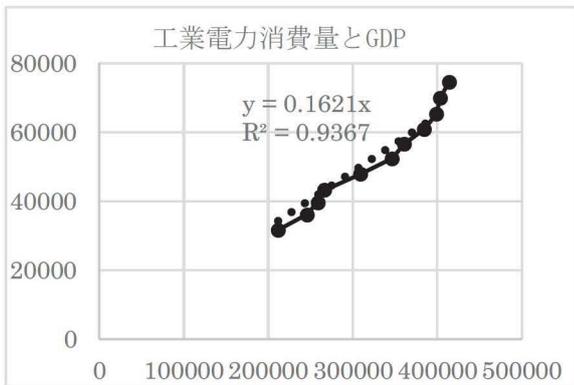


図2 工業電消費量とGDPの単相関図 筆者作成

単相関分析 (鉄道貨物運送量とGDP)

鉄道貨物運送量のみの場合

図3により2006年から2011年までは鉄道貨物運送量とGDPとは強い順相関が見られる2011年から2016年までは負の相関が混在している。つまり、2011年から2016年までの期間は中国高速道路の整備が進んでおり、鉄道貨物運送量のみがGDPに寄与するのではないと考えられる。

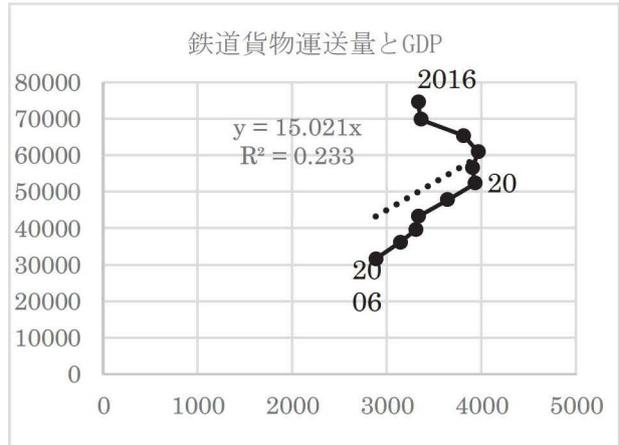


図3 鉄道貨物運送量とGDPとの単相関図 筆者作成

単相関分析 (貨物運送量とGDP)

鉄道貨物運送量と高速貨物運送量の場合

図3より、本研究では中国のGDPに影響する貨物運送量は鉄道のみではなく、高速道路の運送量も加味して分析するのが有効であると考えられる。鉄道貨物運送量と高速貨物運送量を加えた貨物運送量とGDPとの単相関図は下の通りである。

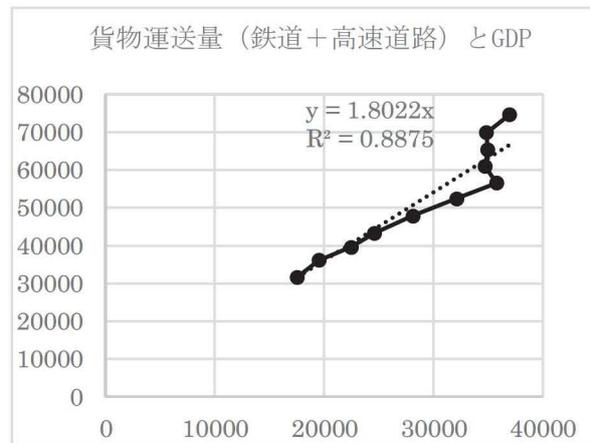


図4 貨物運送量とGDPとの単相関図 筆者作成

上記の表1では鉄道貨物運送量であるが、高速道路運送量も含めたデータは表2ようになる。

表2 GDPと新3指数データ

年度(12月)	説明変数			目的変数
	貨物運送量(鉄道と高速道路(百万トン))	銀行融資残高の伸び(千億元)	工業電力消費(千万KWH)	GDP(10億元)
2006年	17,545	31,800	211,540	31,654
2007年	19,536	48,000	245,660	36,149

⁸⁾ R²はVIの3.1で参照する。

2008 年	22,472	49,100	258,640	39,620
2009 年	24,611	95,900	266,650	43,265
2010 年	28,124	75,000	308,860	47,853
2011 年	32,134	74,700	346,330	52,400
2012 年	35,789	85,000	360,610	56,539
2013 年	34,737	88,900	384,710	60,949
2014 年	34,943	97,800	398,980	65,398
2015 年	34,858	112,000	403,300	69,911
2016 年	36,962	12,6500	413,830	74,595

経済の発展と 3 つの指標と GDP との相関

表 2 のデータで作られたグラフが図 5 である。図 5 により 2006 年から 2016 年まで中国の経済の型 (タイプ) が 2008 年を境に、プロダクト型経済からファイナンス型経済 [13] 武田 (2008) に移行した様子がわかる。中国の銀行融資残高の伸びはリーマンショックには影響されておらず、GDP と共に右肩上がりをしていいる。工業電力消費量と貨物運送量も GDP と同じ動きで登っている。

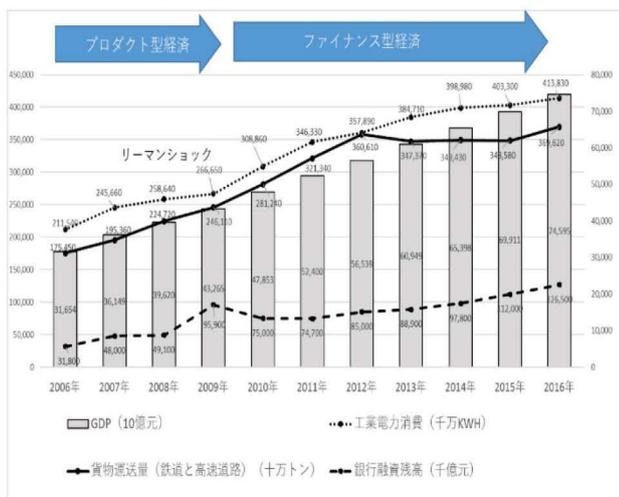


図 5 2006 年から 2016 年まで中国経済状況

注：単位調整済み

GDP と 3 変数の相関行列

説明変数と目的変数の相関を把握するために相関行列を作り、結果は表 3 の通りとなった⁹⁾。図 3 により貨物運送量と工業電力消費量の相関は 0.9773、貨物運送量と銀行融資残高の伸びとの相関は 0.8459、工業電力消費量と銀行融資残高の伸びは 0.8592 である。GDP、貨物運送量および銀行融資残高の伸びと GDP との間の相関係数はそれぞれ 0.9431、0.9113、0.9814 である。

⁹⁾ エクセルあるいは R 言語その他の統計パッケージソフトで作成可能である。

表 3 各項目の相関行列表

	貨物運送量 (鉄道と高速道路) (百万トン)	銀行融資残高の伸び (千億元)	工業電力消費 (千万 KWH)	GDP (10 億元)
貨物運送量 (鉄道と高速道路) (百万トン)	1	0.8459	0.9773	0.9431
銀行融資残高の伸び (千億元)	0.8459	1	0.8592	0.9113
工業電力消費 (千万 KWH)	0.9773	0.8592	1	0.9814
GDP (10 億元)	0.9431	0.9113	0.9814	1

表 3 を図にすれば下の通りである (図 6)。ここで太い線は相関係数が 0.9 以上の強い相関を示す。

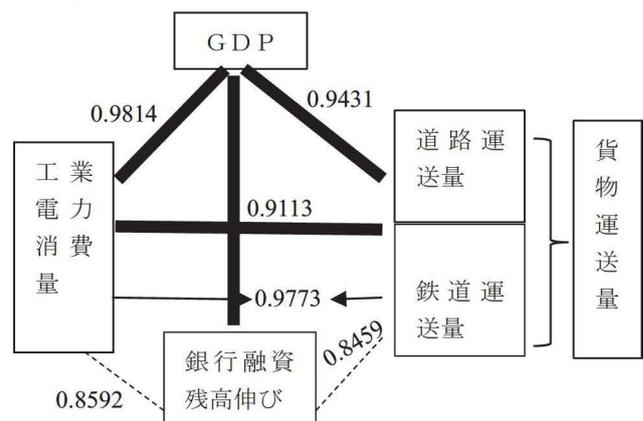


図 6 GDP と新 3 指数の相関図 筆者作成

図 6 により、本研究では、中国の GDP に影響する経済指標は「工業電量消費量、貨物運送量、銀行融資残高の伸び」のデータを使って、重回帰分析でこの 3 指数の妥当性を検討する。

しかし重回帰分析では多重共線性¹⁰⁾の問題がある。

¹⁰⁾ 多重共線性の問題があるかどうかは、VIF やトレランスと呼ばれる指標の値で確認できる。EXCEL には行列式の値をもとめる「MDETERM」という関数が存在するので、それで行列式の値を確認する手法もある。

相関係数と VIF 統計量との関係は次の通りである。『VIF 統計量は、1 つの説明変数 x_i を目的変数とした他の説明変数による重回帰分析での重相関係数 r_i^2 を用いて以下のように定義される。

$$VIF_i = 1 / (1 - r_i^2)$$

これによると、VIF の値が 10 程度というのは、重相関係数が約 0.95 という事になる。』福井正康

(<http://www.heisei-u.ac.jp/ba/fukui/CATips.html>)。

ここに多重共線性とは、説明変数の中に、相関係数が高い相関の組み合わせがあることをいう。分析結果の精度が高くなるため、説明変数の相関係数が極めて高い場合、説明変数として両方を使わない方が精度が向上するとされる。そこで多重共線性を回避するため、表3の相関行列表により貨物運送量と工業電力消費量の相関が0.9以上であるので、この2つ説明変数は多重共線性となるおそれがある。本研究ではR言語を用いて客観的に多重共線性を検討した。その分析結果は以下通りとなる。VIF統計量の値は10より大きい場合には多重共線性の疑いがあるとされ、VIF統計量が10より小さい場合には多重共線性の疑いがないとされる。本研究ではR言語により貨物運送量と工業電力消費量の多重共線性の値はそれぞれ22.386及び24.329である、銀行融資残高の伸びの多重共線性の値は3.831である。従って貨物運送量と電力消費量の多重共線性があるということが分かった。貨物運送量と工業電力消費量どちらが一つ削除をすることになる。しかも表4により貨物運送量の有意確率(3.2で説明する)は5%を上回っているため、その係数が統計的に意味を持っていないと判断できるので本研究では貨物運送量を削除する。

表4 GDPと新3指数の多重共線性の確認結果 (R言語)

	偏回帰係数	標準誤差	t値	有意確率
切片	-8345.00	3117.00	-2.678	0.0317
貨物運送量 (鉄道と高速道路)(百万トン)	-0.79	0.39	-2.022	0.0829
銀行融資残高の伸び(千億元)	0.13	0.04	3.322	0.0127
工業電力消費(千万KWH)	0.22	0.04	5.564	0.0008

表4により図は下の通りになる。

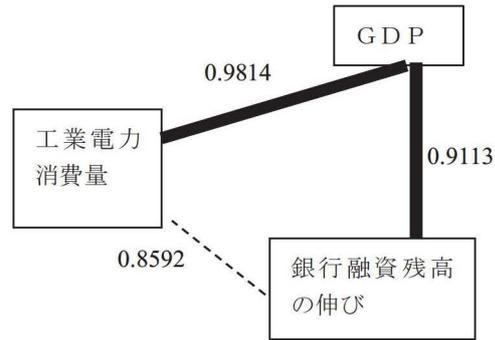


図7 GDPと工業電力消費量,銀行融資残高の伸びの相関図 筆者作成

削除したデータは下の通りである。

表5 削除したデータ

	説明変数		目的変数
	銀行融資残高の伸び(千億元)	工業電力消費(億KWH)	GDP(10億元)
2006年	31,800	21,154	31,654
2007年	48,000	24,566	36,149
2008年	49,100	25,864	39,620
2009年	95,900	26,665	43,265
2010年	75,000	30,886	47,853
2011年	74,700	34,633	52,400
2012年	85,000	36,061	56,539
2013年	88,900	38,471	60,949
2014年	97,800	39,898	65,398
2015年	112,000	40,330	69,911
2016年	126,500	41,383	74,595

表5から相関行列表は表6の通りになる。変数GDPと説明変数工業電力消費量,銀行融資残高の伸びの相関はそれぞれ0.9113及び0.8592である。つまり、目的変数と説明変数の相関があり、説明変数の多重共線の可能性もないため、本研究では次の章で表5のデータを使って重回帰分析に入る。

表6 削除したデータの相関係数表

	銀行融資残高の伸び(千億元)	工業電力消費(億KWH)	GDP(10億元)
銀行融資残高の伸び(千億元)	1	0.8592	0.9113

工業電力消費 (億 KWH)	0.8592	1	0.9814
GDP (10 億円)	0.9113	0.9814	1

4. 重回帰分析

本研究では重回帰分析を予測ではなく、目的変数に対する各説明変数の影響度を調べるために利用する¹¹⁾。

4・1 重回帰式の導出

表 5 のデータを R で重回帰分析の最小二乗法で偏回帰係数 (表 7) を求めた。目的変数 GDP を Y とすれば、切片は -6693、回帰係数銀行融資残高の伸び (千億円) と工業電力消費量 (億 KWH) はそれぞれ 0.13 と 1.49 である。偏回帰係数の値の大きさに注目すると「工業電力消費量」のものが最大である。つまり、工業電力消費量が GDP に最も影響を及ぼしている。重回帰式は下式 1 の通りである。

$$y = 1.49 \times \text{工業電力消費量} + 0.13 \times \text{銀行融資残高の伸び} - 6693$$

……………式 1

表 7 R 言語による重回帰分析 (最小二乗法)

	偏回帰係数	標準誤差	t 値	有意確率
切片	-6693.00	3541.00	-1.89	0.09540
銀行融資残高の 伸び (千億円)	0.13	0.05	2.728	0.02590
工業電力消費 (億 KWH)	1.49	0.19	7.948	0.00005

4・2 P 値と t 値

重回帰分析において t 値 (t-value) は各説明変数が目的変数に与える影響の大きさを表している。絶対値が大きいほど影響が強いことを意味する。t 値は絶対値 2 以下である場合は説明変数が目的変数には全く影響されないとされる¹²⁾。本研究では t 値 (表) の絶対値はどれも 2 以上になっており、つまり、全ての説明変数が目的変数に影響している。

$\text{Pr}(>|t|)$ (有意確率) は説明変数の係数や定数項が偶然その値である確率を示している。通常は 5% 未満で、その係数が統計的に意味を持っていることが判断でき

る¹³⁾。今回の分析結果により有意確率 (表 6) はどれも 5% を下回っている。本研究では説明変数と目的変数の相関があるということは偶然ではないと判断できた。

ちなみに切片 (Intercept) は重回帰分析においては全ての独立変数の値がゼロの場合の従属変数の予測値である。

4・3 重回帰分析の精度

重回帰式の精度を確認

重回帰式の精度を確認するは主に 2 つ重要なものがある。

- ① 変数の説明力
- ② 変数の信頼性

4・3・1 寄与率と自由度調整済み寄与率で回帰式説明力を確認

寄与率 (Multiple R-squared)、寄与率は決定係数とも言い、全体の説明力を意味し重回帰分析には最も重要な指標である。つまり、説明変数はどの程度目的変数を説明できるかを示している指標である¹⁴⁾。しかし、寄与率は説明変数が多いほど寄与率の値が大きくなる欠点があり、見かけ上の精度が高くなるため、通常では自由度調整済み寄与率 (Adjusted R-squared) を使用する。本研究の自由度調整済み寄与率は 0.976 であり、説明変数が目的変数の 98% 状況が説明できる。

寄与率と P 値 (R 言語)

標準誤差 (Residual standard error) : 2193 on 8 degrees of freedom

寄与率 (Multiple R-squared) : 0.981

調整済み寄与率 (Adjusted R-squared) : 0.9762

F 値 (F-statistic) : 206 on 2 and 8 DF

P 値 (p-value) : 0.0000001316

4・3・2 寄与率の P 値で回帰式の信頼性を確認

今回の分析では寄与率の P 値が「0.0000001316」とほぼ 0 になり、説明変数は 99% 以上で信頼できると言える。

¹¹⁾https://www.albert2005.co.jp/knowledge/statistics_analysis/multivariate_analysis/multivariate_method

¹²⁾「7 つの統計用語を知りましょう | マーケティングと重回帰分析」(2017)

<http://xica.net/magellan/marketing-idea/stats/statistics-words/>

¹³⁾「t 値と p 値の違いを教えて | マーケティングと重回帰分析」(2017)

<http://xica.net/magellan/marketing-idea/stats/tvalue-and-pvalue/>

¹⁴⁾<http://www.geisya.or.jp/~mwm48961/statistics/coef.html>

5. 結論

李克強の3指数はGDPと強い相関がある。2011年以降について鉄道貨物輸送量のみならず高速道路貨物輸送量を加味した貨物輸送量、銀行融資残高の伸び及び工業電力消費量の新3指数は、中国GDPの動向を示す代理変数として機能していることが判明した。

中国のGDPは、新3変数で示される実体的な経済活動とそれぞれ非常に強い相関が認められることから、GDP統計自体も中国の経済実態を表しているとも考えられる。従って、中国のGDP統計自体にも信頼性があると考えられる。

参考文献

- [1]Wang, BY, Zhang, YH: "A study on city motor vehicle emission factors by tunnel test", 科学出版社(2001)
- [2]Jeremy Yuen Chun Teoh, Peter Ka Fung Chiu: "Risk of new-onset diabetes after androgen deprivation therapy for prostate cancer in the Asian population", Volume 7, Issue 5 September 2015 Pages 672-680
- [3]君山由良, 重回帰分析の利用法, データ分析研究所, (2004)
- [4]孟健軍, 張紅咏: 中国の鉄道物流構造変化に関する実証分析 2015年5月, 経済産業研究所(2015)
- [5]中川徹, 小柳義夫: 最小二乗法による実験データ解析, 東京大学出版会, (1982)
- [6]仁科健, 永田靖: 重回帰分析における偏回帰係数の解釈について, 東京: 日本品質管理学会(2002)
- [7]大村平: 多変量解析のはなし: 複雑さから本質を探る, 日科技連出版社(2006)
- [8]瀬口清之: 李克強指数で中国経済を判断すると間違える理由, JBpress(2015)
- [9]谷口洋志: 克強指数はGDP統計よりも中国経済の実態を正確に反映するか(2): 鉄道貨物輸送量のケース, IERCUDiscussion Paper (No.253), 2015-06
- [10]谷口洋志: 克強指数はGDP統計よりも中国経済の実態を正確に反映するか(2): 電力消費量のケース, IERCUDiscussion Paper (No.252), 2015-05
- [11]高橋洋一: 中国GDPの大嘘, 講談社(2015)
- [12]豊田秀樹: 「もう一つの重回帰分析: 予測変数を直行化する方法」, 東京図書(2017)
- [13]武田隆二: 最新財務諸表論, 中央経済社(2008年)
- [14]山本真理子, 松井豊, 山成由紀子: 認知された自己の諸側面の構造, J-STAGE 30巻(1982)1号 p. 64-68
- [15]泰地秀信, 神崎仁: 突発性難聴の予後の回帰分析, J-STAGE 91巻(1988)7号 p. 1023-1030
<https://www.jstage.jst.go.jp/article/jibiinkoka1947/9>

1/7/91_7_1023/_pdf

[16]堤盛人, 清水英範, 井出裕史: 誤差項の仮定からの違背が空間データを用いた回帰分析の結果に及ぼす影響, Theory and Applications of GIS, 1999, Vol. 7 No. 1, pp. 19-26

[17]シャーロット・ガオ, 中国経済の「復活」のIMFがお墨付き, Newsweek ニューズウィーク日本版, (2017), pp. 20

(受理 平成30年3月10日)