

過去のデータから将来をどう予測するか

時系列回帰分析を用いた予測手法

はじめに

損害保険料率の算出を主たる業務としている本機構や損害保険会社においては、過去のデータを用いて将来を予測することがよく行われる。しかし、これらの方法は、保険業界に限ったことではなく、私企業・公企業を問わず、おおよそ事業を営むあらゆる業態において、需要予測に基づき、生産・サービス供給活動を行っている以上、多かれ少なかれ将来予測は行われている。最もよく知られているのは、政府系や民間の経済研究機関等が毎年行っている経済成長率予測であろう。

ところで、日本人の人口の年次推移や65歳以上の人口割合の年次推移などについては、比較的年度間で大きく変動することはなく、長期的なスパンにおいて概ねトレンド的な動きをすることが知られているので、単純な時系列回帰モデルを用いて予測したとしてもそれほど実態と大きく乖離することはないであろう。一方で、外部環境の急激な変化がその後の推移にダイレクトな変化を及ぼすような場合、例えば、出国日本人の年次別推移に基づき近未来の予測を行うとなると話は違ってくる。なぜなら、一昨年のニューヨークにおける同時多発テロの勃発というような、当初だれも想定していなかった事件が突然発生し、その日を境にして日本人の海外旅行者が激減したからである。

また、交通事故死者数の推移等についても、飲酒運転や、危険運転の罰則を強化した改正道交法に加え、これまで継続してきた安全対策の効果等により、近年、交通事故者数の減少傾向が見受けられる。これらの事例をみると、いずれも人為的な要因により大きく変動しているので、より精緻な予測を行うためには、月別推移を見るなど、よりきめ細かな分析が求められる。

本論文では、月次データが得られるような場合に、年次別トレンドおよび月別のサイクリックな動向を加味した推移予測をどう行えばよいのか、事例を用いながら、数量化理論類という回帰分析を用いた予測手法の一例について以下に紹介する¹。

¹複雑な社会現象を説明する時系列経済モデルとしては、トレンド、サイクル、季節変動、不規則変動の4つの要素から組み立てられた決定論的因果関係モデルが採用されている。また、最近では、これらの4要素に分解することなく同時に体系的に扱うARIMA (Autoregressive Integrated Moving Average) モデルが用いられるのが一般であるが、本論文では単に確率的因果関係に基づく分析を中心に扱っている。

1. 年次データを使用した時系列回帰による予測方法

1.1 時系列回帰による予測

1.1.1 直線回帰による予測

表1.1.1 は、わが国の65歳以上の人口割合の推移である。このデータから2050年の65歳以上の人口割合を予測するために、少々乱暴ではあるが、以下のような単純な時系列回帰（直線回帰）モデルを用いることから始めたい。

表1.1.1 65歳以上の人口割合の推移

年次	65歳以上 人口の割合
1930	4.75
1940	4.80
1950	4.94
1960	5.73
1970	7.07
1980	9.10
1990	12.05
2000	17.24
2010	?
2020	?
2030	?
2040	?
2050	?

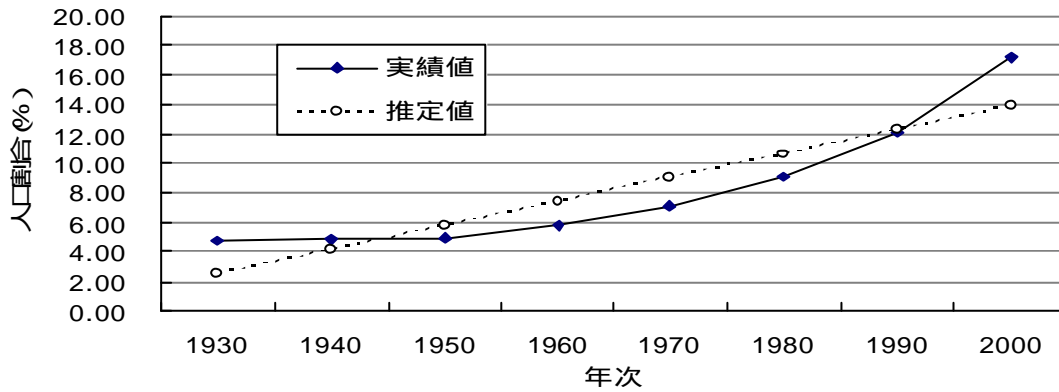


図1.1.1 65歳以上の人口割合の年次推移

x を年次、y を65歳以上の人口割合とすれば、回帰直線(最小二乗直線)
 $y = a x + b$ から、a および b は以下のとおり求められる。

$$a = \frac{S(xy)}{S(xx)} = \frac{\sum (x_i - \bar{x})(y_i - \bar{y})}{\sum (x_i - \bar{x})^2} \dots\dots\dots$$

$$b = \bar{y} - a\bar{x} \dots\dots\dots$$

65歳以上				
年次 x	人口の割合 y	$[x_i - E(x)]^2$	$[y_i - E(y)]^2$	$[x_i - E(x)] * [y_i - E(y)]$
1930	4.75	1,225	11.97	121.1
1940	4.80	625	11.63	85.25
1950	4.94	225	10.69	49.05
1960	5.73	25	6.15	12.4
1970	7.07	25	1.30	-5.7
1980	9.10	225	0.79	13.35
1990	12.05	625	14.75	96
2000	17.24	1,225	81.54	316.05
1965	8.21	4,200	138.82	687.5000
E(x)	E(y)	S(xx)	S(yy)	S(xy)
サンプルサイズ	n			8
定数項	$a = S(xy)/S(xx) =$			0.1637
回帰係数	$b = E(y) - aE(x) =$			-313.442

以上の結果から、2050年における65歳以上の人口割合を推定すると、 $y = a x + b$ から、 $y = 0.1637 \times 2050 - 313.442 = 22.12$ となる。

ところで、国立社会保障・人口問題研究所の予測によれば、2010年以降10年ごと2050年までの65歳以上の人口割合は、以下のとおりとなっている。

表1.1.2 65歳以上の人口割合の年次推移（予測）

年次	65歳以上 人口の割合	予測式 $y = ax + b$
1930	4.75	2.48
1940	4.80	4.12
1950	4.94	5.75
1960	5.73	7.39
1970	7.07	9.03
1980	9.10	10.67
1990	12.05	12.30
2000	17.24	13.94
2010	(22.04)	15.58 *
2020	(26.85)	17.21 *
2030	(27.97)	18.85 *
2040	(30.95)	20.49 *
2050	(32.29)	22.12 *

()内は人口問題研究所の予測値 *印は予測値

国立社会保障・人口問題研究所による推定は、単なるトレンド予測手法ではなく、既存人口については将来生命表を用いて各年齢別の将来人口を予測するとともに、新たに生まれる人口については、将来の出生率を用いて生存数を予測するコーホート要因法と呼ばれる手法を採用している。したがって、コーホート要因法によって将来人口を推計するためには、基準人口、将来生残率、将来の出生率、将来の出生性比、将来の国際人口移動数（率）の5つのデータが必要とされる。

上記のとおり、2050年の65歳以上の人口割合については、人口問題研究所の予測値は32.29%であるのに対して、単純に回帰直線モデルに当てはめると、22.12%であり、かなり下回っていることが分かる。

1.2.1 指数回帰による予測

次に、指数回帰モデルによる推定を行ってみた。推定式は、 $y = b e^{ax}$ すなわち、 $\log y = ax + \log b$ とおいて、以下のとおり a 、 $\log b$ を求める。

65歳以上					
年次 x	人口の割合 y	z = log y	$[x_i - E(x)]^2$	$[z_i - E(z)]^2$	$[x_i - E(x)] * [z_i - E(z)]$
1930	4.75	1.55814	1,225	0.19	68720.4649
1940	4.80	1.56862	625	0.18	49085.7846
1950	4.94	1.59737	225	0.16	29451.0395
1960	5.73	1.74572	25	0.06	9816.2714
1970	7.07	1.95586	25	0.00	-9815.2207
1980	9.10	2.20827	225	0.04	-29441.8759
1990	12.05	2.48906	625	0.24	-49062.7734
2000	17.24	2.84723	1,225	0.72	-68675.3469
1965	8.21	1.99628	4,200	1.61	78.3436
E(x)		E(z)	S(xx)	S(zz)	S(xz)
サンプルサイズ	n			8	
定数項		$a = S(xz)/S(xx) =$		0.01865	
回帰係数		$\log b = E(z) - aE(x) =$		-34.657	
		$b = e^{-34.657} =$		8.882E-16	

同様にして、2010年以降10年ごと2050年までの65歳以上の人口割合は、表1.2.1のとおりとなっている。

表1.2.1 65歳以上の人口割合の年次推移（推定）

年次 x	65歳以上 人口の割合 y	予測式 $y = b e^{ax}$
1930	4.75	3.83
1940	4.80	4.62
1950	4.94	5.56
1960	5.73	6.71
1970	7.07	8.08
1980	9.10	9.74
1990	12.05	11.74
2000	17.24	14.14
2010	(22.04)	17.04 *
2020	(26.85)	20.54 *
2030	(27.97)	24.75 *
2040	(30.95)	29.82 *
2050	(32.29)	35.94 *

()内は人口問題研究所の予測値 *印は予測値

以上のとおり、指数回帰モデルであると、2040年あたりは、人口問題研究所の予測値とかなり近い値が得られているが、2050年になると、大分上回ってしまうことが分かる。

日本の将来人口推計については、コーホート要因法という精緻な要因データに基づき行われており、その予測手法により得られた予測値に対して、単純な回帰モデルとの比較は的確さを欠いているとの批判は免れないと思われる。それでも、上記のような65歳以上の人口割合といった、比較的マクロ的推移を把握するような場合には、単純な回帰モデルであっても、よく適合する回帰曲線が選択できさえすれば、予測値は実績値とそれほど大きく乖離することはないであろう。

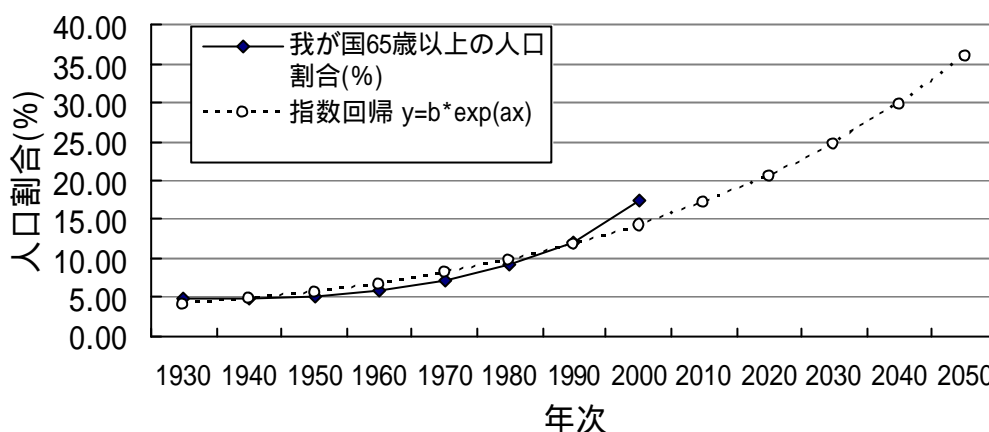


図1.2.1 65歳以上の人口割合の年次推移

しかしながら、これらのマクロ的な予測方法は、企業が行う日々の生産活動やサービス供給活動においては、ほとんど意味をなさない。例えば、メーカーなどは需要予測に基づいて製品の生産計画を立て、それに基づき生産、在庫、出荷、販売を行っている。需要予測を過大に評価してしまうと、在庫コストがかさばり、最終的には死蔵在庫の山となって、大きな営業損失を招くことになる。そのため、製品によっては、月単位あるいは週単位の販売予測、さらには鮮度が求められる生鮮食料品等にあって時間単位に基づく、生産・出荷計画が行われるのが一般的であるからである。例えば、スーパーなどがその日の商品別入荷量を決定するに当たっては、日々の天候・気温・湿度などの気象データをはじめとして、競合店舗の有無や販売戦術、運動会等の地域行事の有無などのコーザルデータ（売上に影響を及ぼす商品価格以外のデータ）を総合的に活用して、予測すると言われている。

2. 月次データを使用した時系列回帰による予測方法

政府や民間調査機関等が公表している各種統計データの中には、年次別データの他に、月次別データが入手できるものが案外多い。そこで、月次データが得られるような場合に、年次別トレンドおよび月次のサイクリックな動向を加味した推移予測を行うにはどうすればよいのか、具体的な事例を用いながら、以下に紹介する。

2.1 出国日本人の予測

法務省出入国管理統計によると、出国日本人数の年次別推移は、表2.1.1のとおりである。海外出国者の大多数が海外旅行傷害保険に加入するという実態からすると、損害保険会社にとっても、予算計画を立てる上で出国日本人数の正確な予測を行うことは、極めて関心が高いテーマである。

表2.1.1 出国日本人の年次別推移

	Total
1996年	16,694,769
1997年	16,802,750
1998年	15,806,218
1999年	16,357,572
2000年	17,818,590
2001年	16,215,657

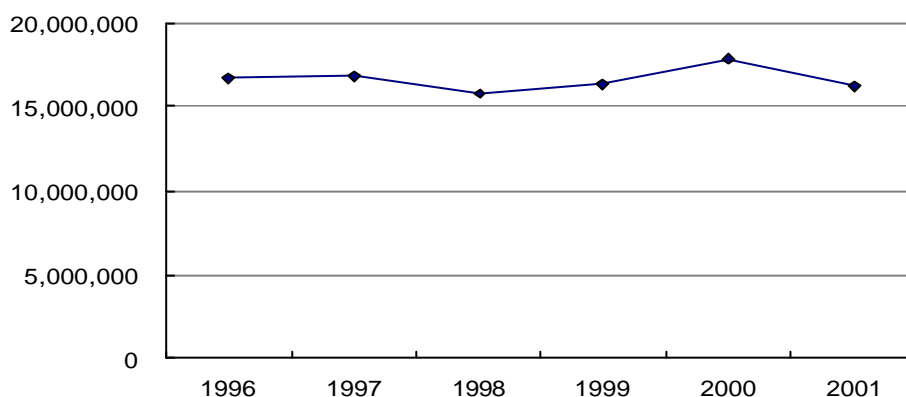


図2.1.1 出国日本人の年次別推移（折れ線グラフ）

図2.1.1 をみると、1996年から毎年出国者数が上昇傾向であったが、2001年に急激に落ち込んだことが分かる。その理由を探るため、次に月別の出国日本人の推移を表したのが表2.1.2および図2.1.2である。

表2.1.2 出国日本人の月別推移

	1月	2月	3月	4月	5月	6月
1996年	1,301,035	1,310,529	1,503,189	1,177,272	1,215,418	1,339,160
1997年	1,346,750	1,367,897	1,549,137	1,233,412	1,264,947	1,340,096
1998年	1,195,570	1,266,536	1,365,750	1,132,109	1,220,230	1,248,827
1999年	1,245,306	1,258,616	1,471,483	1,184,801	1,235,395	1,280,099
2000年	1,228,599	1,414,251	1,573,517	1,305,417	1,369,655	1,421,924
2001年	1,361,711	1,501,532	1,612,008	1,370,049	1,366,727	1,460,542

	7月	8月	9月	10月	11月	12月	Total
1996年	1,470,207	1,660,720	1,551,960	1,363,782	1,316,778	1,484,719	16,694,769
1997年	1,512,137	1,671,698	1,545,922	1,344,536	1,311,081	1,315,137	16,802,750
1998年	1,459,841	1,628,575	1,501,426	1,233,945	1,264,895	1,288,514	15,806,218
1999年	1,473,633	1,686,134	1,572,340	1,384,130	1,358,036	1,207,599	16,357,572
2000年	1,583,129	1,759,090	1,677,031	1,522,313	1,531,695	1,431,969	17,818,590
2001年	1,596,737	1,791,166	1,331,411	925,142	860,698	1,037,934	16,215,657

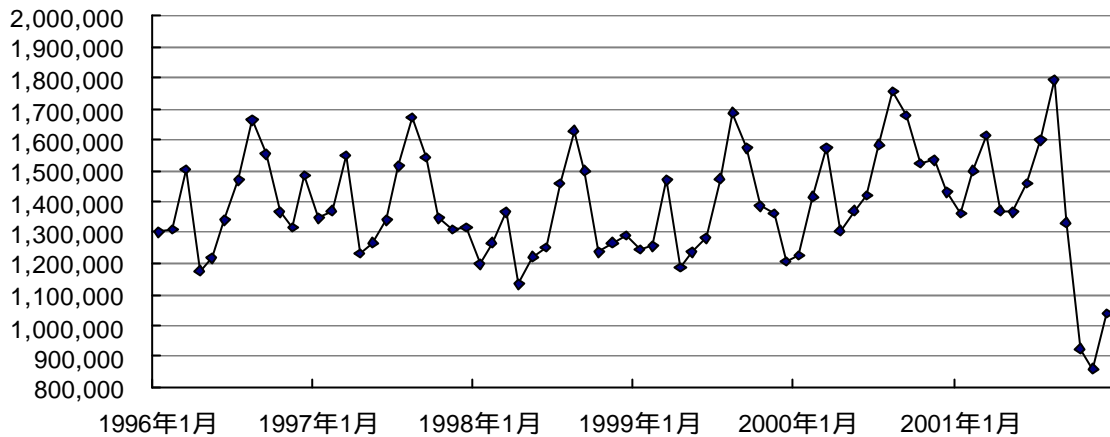


図2.1.2 出国日本人の月別推移(折れ線グラフ)

図2.1.2をみると、1996年～2000年までの5年間の毎月の出国日本人の推移は、観光旅行者が大半を占める中で、3月の春休みに学生等の卒業旅行などで多くなり、4月に落ち込み、7月、8月のサマーバケーション時期に向けて右肩上がりに上昇し、9月以降は、急激に下降線を辿り、11月～翌年1月が底となり、また3月に向けて上昇するという、毎年同じようなサイクルで推移していることが分かる。しかしながら、2001年をみると、8月までは同じように月別推移をしていたのにもかかわらず、9月になると激減しており、10月、11月にはさらに落ち込んでいることが分かる。この理由としては、2001年9月11日の米国ニューヨーク市を中心とした同時多発テロの後、テロを警戒して日本人による海外旅行や海外出張等の自粛や取り止めが大きく影響している。このように、年次別の出国日本人の推移をみている限りでは分からなかった構造変化が、月別推移を辿ってみることによって初めて鮮明に浮かび上がってくることがある。

そこで、もし、あのような痛ましい事件が発生しなければ、出国日本人の月別推移はどうであっただろうかを推定するため、以下に回帰分析による推定を行ってみることとする。まず、回帰分析が可能となるように以下のようなデータを作成することからはじめる。

表2.1.3 回帰分析が可能なデータ

年次	出国 日本人数	96 年	97 年	98 年	99 年	2 月	3 月	4 月	5 月	6 月	7 月	8 月	9 月	10 月	11 月	12 月
1996年1月	1,301,035	1	0	0	0	0	0	0	0	0	0	0	0	0	0	0
2月	1,310,529	1	0	0	0	1	0	0	0	0	0	0	0	0	0	0
3月	1,503,189	1	0	0	0	0	1	0	0	0	0	0	0	0	0	0
4月	1,177,272	1	0	0	0	0	0	1	0	0	0	0	0	0	0	0
5月	1,215,418	1	0	0	0	0	0	0	1	0	0	0	0	0	0	0
6月	1,339,160	1	0	0	0	0	0	0	0	1	0	0	0	0	0	0
7月	1,470,207	1	0	0	0	0	0	0	0	0	1	0	0	0	0	0
8月	1,660,720	1	0	0	0	0	0	0	0	0	0	1	0	0	0	0
9月	1,551,960	1	0	0	0	0	0	0	0	0	0	0	1	0	0	0
10月	1,363,782	1	0	0	0	0	0	0	0	0	0	0	0	1	0	0
11月	1,316,778	1	0	0	0	0	0	0	0	0	0	0	0	0	1	0
12月	1,484,719	1	0	0	0	0	0	0	0	0	0	0	0	0	0	1

2000年1月	1,228,599	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
2月	1,414,251	0	0	0	0	1	0	0	0	0	0	0	0	0	0	0
3月	1,573,517	0	0	0	0	0	1	0	0	0	0	0	0	0	0	0
4月	1,305,417	0	0	0	0	0	0	1	0	0	0	0	0	0	0	0
5月	1,369,655	0	0	0	0	0	0	0	1	0	0	0	0	0	0	0
6月	1,421,924	0	0	0	0	0	0	0	0	1	0	0	0	0	0	0
7月	1,583,129	0	0	0	0	0	0	0	0	0	1	0	0	0	0	0
8月	1,759,090	0	0	0	0	0	0	0	0	0	0	1	0	0	0	0
9月	1,677,031	0	0	0	0	0	0	0	0	0	0	0	1	0	0	0
10月	1,522,313	0	0	0	0	0	0	0	0	0	0	0	0	1	0	0
11月	1,531,695	0	0	0	0	0	0	0	0	0	0	0	0	0	1	0
12月	1,431,969	0	0	0	0	0	0	0	0	0	0	0	0	0	0	1

次に、上記のデータを活用して、Excelの「分析ツール」を選び、「回帰分析」を選択すると、以下のような回帰分析結果が得られる。

表2.1.4 回帰分析結果

概要

回帰統計	
重相関 R	0.962665125
重決定 R2	0.926724143
補正 R2	0.901743737
標準誤差	47796.92075
観測数	60

分散分析表

	自由度	変動	分散	観測された分散比	有意 F
回帰	15	1.27128E+12	84752169347	37.09804178	5.18393E-20
残差	44	1.0052E+11	2284545633		
合計	59	1.3718E+12			

	係数	標準誤差	t	P-値	下限 95%	上限 95%
切片	1357002.85	24682.22374	54.97895425	3.38831E-42	1307259.098	1406746.602
96年	-93651.75	19513.01119	-4.799451459	1.86442E-05	-132977.6386	-54325.86144
97年	-84653.33333	19513.01119	-4.338301891	8.26563E-05	-123979.2219	-45327.44477
98年	-167697.6667	19513.01119	-8.594145982	5.75063E-11	-207023.5552	-128371.7781
99年	-121751.5	19513.01119	-6.239503419	1.49965E-07	-161077.3886	-82425.61144
2月	60113.8	30229.42694	1.988585497	0.052988816	-809.6045892	121037.2046
3月	229163.2	30229.42694	7.580798685	1.6252E-09	168239.7954	290086.6046
4月	-56849.8	30229.42694	-1.880611237	0.066653844	-117773.2046	4073.604589
5月	-2323	30229.42694	-0.076845651	0.939094708	-63246.40459	58600.40459
6月	62569.2	30229.42694	2.069810986	0.044372153	1645.795411	123492.6046
7月	236337.4	30229.42694	7.818123726	7.37116E-10	175413.9954	297260.8046
8月	417791.4	30229.42694	13.82068541	1.29603E-17	356867.9954	478714.8046
9月	306283.8	30229.42694	10.13197506	4.4519E-13	245360.3954	367207.2046
10月	106289.2	30229.42694	3.516083854	0.001028937	45365.79541	167212.6046
11月	93045	30229.42694	3.077961093	0.00358131	32121.59541	153968.4046
12月	82135.6	30229.42694	2.717074331	0.009381917	21212.19541	143059.0046

1996年～2000年の5カ年間データを用いて、2001年における出国日本人の月別推移を予測値と実測値を比較したのが、表2.1.7および図2.1.4である。

表2.1.7 出国日本人の月別推移の比較

月	実測値	予測値	残差	相対誤差
1月	1,361,711	1,308,507	53,204	3.91%
2月	1,501,532	1,368,621	132,911	8.85%
3月	1,612,008	1,537,670	74,338	4.61%
4月	1,370,049	1,251,657	118,392	8.64%
5月	1,366,727	1,306,184	60,543	4.43%
6月	1,460,542	1,371,076	89,466	6.13%
7月	1,596,737	1,544,844	51,893	3.25%
8月	1,791,166	1,726,298	64,868	3.62%
9月	1,331,411	1,614,791	-283,380	-21.28%
10月	925,142	1,414,796	-489,654	-52.93%
11月	860,698	1,401,552	-540,854	-62.84%
12月	1,037,934	1,390,643	-352,709	-33.98%
相対誤差(絶対値)の平均値				17.87%

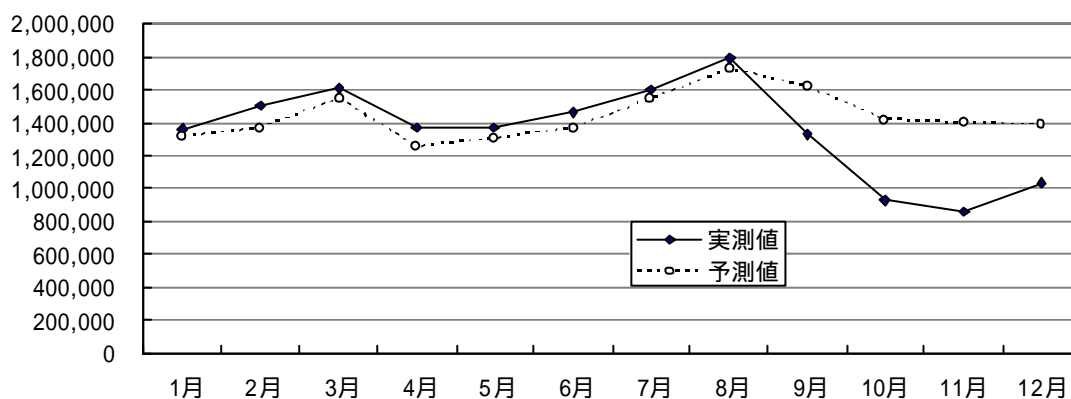


図2.1.4 出国日本人の月別推移の比較

図2.1.4をみて分るとおり、2001年の1月～8月までは実測値と予測値はかなり似通った推移をしているが、9月以降は月を追って実測値と乖離してきていることが分かる。

逆に言えば、この乖離こそが、もし、あのような痛ましい事件がなければ推移したであろう出国日本人の推移であると想定することができる。

2.2 交通事故者数の予測

「2.1 出国日本人の予測手法」と同様な方法によって、以下に交通事故死者数の予測について検討してみよう。

(財)交通事故総合センターの「交通事故データ」によると、交通事故死者数の年次別推移は、表2.2.1のとおりである。

表2.2.1 交通事故死者数の年次別推移

	Total
1996年	9,942
1997年	9,640
1998年	9,211
1999年	9,006
2000年	9,066
2001年	8,747

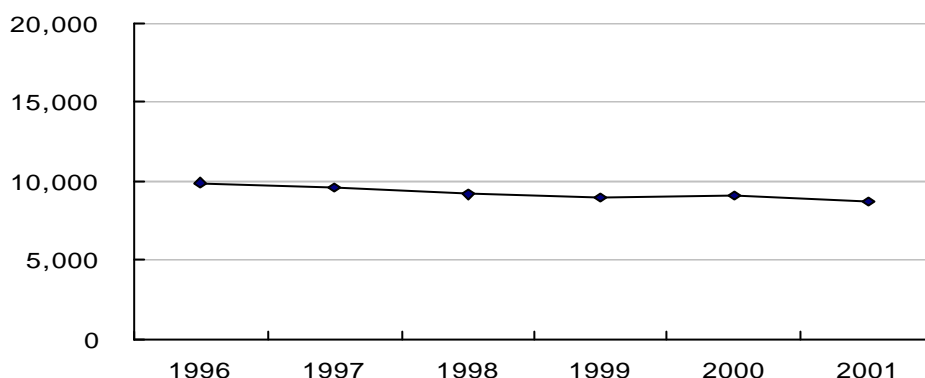


図2.2.1 交通事故死者数の年次別推移 (折れ線グラフ)

1995年～2001年までの交通事故死者数の年次別推移をみると、2000年を例外として、毎年下降傾向で推移していることが分かる。次に、表2.2.2および図2.2.2に示すとおり、交通事故死者数の月別推移を1996年～2001年の6カ年間ながめてみると、毎年とも、1月が月平均な死者数を記録した後、2月に減少し、いったん3月でピークを迎える。これは、2月が他の月に比べて日数が短いということが大きな理由と考えられる。また、4月～8月までは、大の月、小の月によるジグザグはあるものの、やや右肩上がり傾向で推移し、8月をピークに9月は大きく落ち込み、10月からは急激な右肩上がり傾向で12月に最高ピークを迎えるのがほぼ毎年の傾向である。しかしながら、年次別のトレンドを眺めてみると、近年になるにしたがい、全体的にみて下降ぎみで推移していることが分かる。これらの理由として、飲酒運転や、危険運転の罰則を強化した改正道交法に加え、これまで続けてきた安全対策の効果等が指摘されているが、一方で、警察庁統計では、発生から24時間以内の死亡しかカウントされないことから、救急医療体制の充実が交通事故死者数の見かけ上の減少に寄与している部分もあるのではないかと推測も完全には否定できない。

表2.2.2 交通事故死者数の月別推移

	1月	2月	3月	4月	5月	6月
1996年	792	688	785	787	787	760
1997年	813	722	801	740	796	743
1998年	704	621	783	679	764	676
1999年	715	639	699	679	742	717
2000年	728	666	780	694	695	697
2001年	619	637	763	665	662	659

	7月	8月	9月	10月	11月	12月	Total
1996年	829	859	814	942	900	999	9,942
1997年	786	804	760	845	851	979	9,640
1998年	734	772	749	838	931	960	9,211
1999年	758	770	695	805	853	934	9,006
2000年	747	806	685	835	866	867	9,066
2001年	743	745	726	823	833	872	8,747

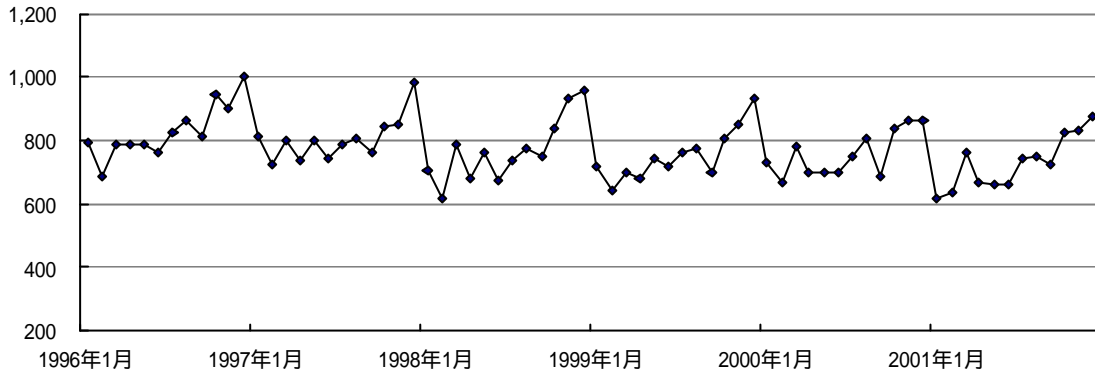


図2.2.2 交通事故死者数の年次別推移

表2.2.3 回帰分析結果

概要

回帰統計	
重相関 R	0.956939105
重決定 R ²	0.915732451
補正 R ²	0.887004878
標準誤差	28.67228627
観測数	60

分散分析表

	自由度	変動	分散	観測された分散比	有意 F
回帰	15	393084.1833	26205.61222	31.87642893	1.04227E-18
残差	44	36172.4	822.1		
合計	59	429256.5833			

	係数	標準誤差	t	P-値	下限 95%	上限 95%
切片	724.8166667	14.80630496	48.95324448	5.16728E-40	694.9765209	754.6568125
96年	73	11.70541185	6.236431568	1.51534E-07	49.40929338	96.59070662
97年	47.83333333	11.70541185	4.086428904	0.000182738	24.24262671	71.42403995
98年	12.08333333	11.70541185	1.032286047	0.30757903	-11.50737329	35.67403995
99年	-5	11.70541185	-0.427152847	0.671351096	-28.59070662	18.59070662
2月	-83.2	18.13394607	-4.588080261	3.70887E-05	-119.7465655	-46.65343446
3月	19.2	18.13394607	1.058787752	0.29547455	-17.34656554	55.74656554
4月	-34.6	18.13394607	-1.908023762	0.062926866	-71.14656554	1.946565544
5月	6.4	18.13394607	0.352929251	0.725826689	-30.14656554	42.94656554
6月	-31.8	18.13394607	-1.753617215	0.086460202	-68.34656554	4.746565544
7月	20.4	18.13394607	1.124961987	0.266704582	-16.14656554	56.94656554
8月	51.8	18.13394607	2.856521124	0.00651533	15.25343446	88.34656554
9月	-9.8	18.13394607	-0.540422915	0.591630665	-46.34656554	26.74656554
10月	102.6	18.13394607	5.657897052	1.07139E-06	66.05343446	139.1465655
11月	129.8	18.13394607	7.157846368	6.71496E-09	93.25343446	166.3465655
12月	197.4	18.13394607	10.88566158	4.57297E-14	160.8534345	233.9465655

表2.2.4 年次別の回帰分析実行結果

概要

回帰統計	
重相関 R	0.935699
重決定 R2	0.875533
補正 R2	0.834044
標準誤差	13.68735
観測数	5

分散分析表

	自由度	変動	分散	観測された分散比	有意 F
回帰	1	3953.469	3953.469	21.102779	0.0193832
残差	3	562.0306	187.3435		
合計	4	4515.5			

	係数	標準誤差	t	P-値	下限 95%	上限 95%
切片	39752.48	8647.986	4.596733	0.0193495	12230.707	67274.258
年次	-19.8833	4.32832	-4.59378	0.0193832	-33.65799	-6.108674

$$Y(\text{年毎の回帰係数}) = 39,752.48 - 19.8833 X(\text{年次})$$

X = 2001を代入すると、 Y = -34.0033
 という2001年の回帰係数の予測値が求まる。

表2.2.3および表2.2.4を用いて、交通事故死者数を求める式を算出したのが、表2.2.5である。

表2.2.5 交通事故死者数を求める式

$$\begin{aligned}
 \text{死者数} = & 724.8167 + 0.0 \text{ (1月)} \\
 & -83.2 \text{ (2月)} \\
 & 19.2 \text{ (3月)} \\
 & 73 \text{ (1996年)} -34.6 \text{ (4月)} \\
 & 47.83333 \text{ (1997年)} 6.4 \text{ (5月)} \\
 & + 12.08333 \text{ (1998年)} + -31.8 \text{ (6月)} \\
 & -5 \text{ (1999年)} 20.4 \text{ (7月)} \\
 & 0.0 \text{ (2000年)} 51.8 \text{ (8月)} \\
 & -34.0033 \text{ (2001年)} -9.8 \text{ (9月)} \\
 & 102.6 \text{ (10月)} \\
 & 129.8 \text{ (11月)} \\
 & 197.4 \text{ (12月)}
 \end{aligned}$$

表2.2.4に基づき、交通事故死者数の月別推移を推定したのが、図2.2.3である。

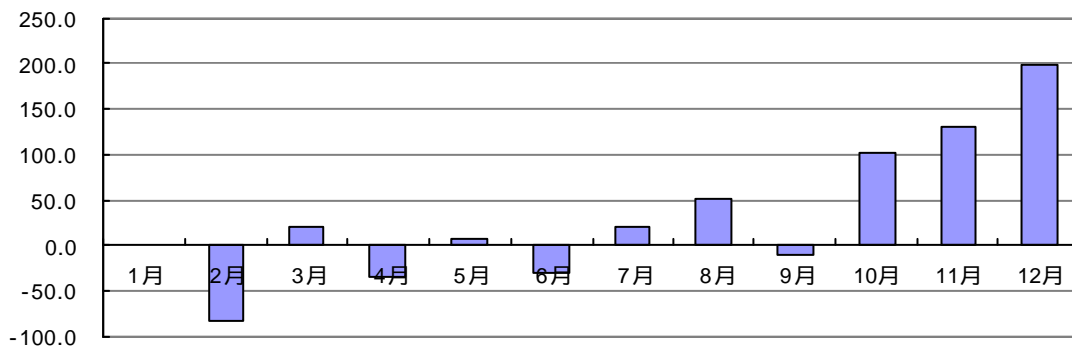


図2.2.3 交通事故死者数の月別推移 (棒グラフ)

1996年～2000年の5カ年間データを用いて、2001年における交通事故死者数の月別推移について予測値と実測値を比較したのが、表2.2.6および図2.2.4である。

表2.2.6 交通事故死者数の月別推移の比較（2001年）

月	実測値	予測値	残差	相対誤差
1月	619	690.8	-71.8	-11.60%
2月	637	607.6	29.4	4.61%
3月	763	710.0	53.0	6.94%
4月	665	656.2	8.8	1.32%
5月	662	697.2	-35.2	-5.32%
6月	659	659.0	0.0	0.00%
7月	743	711.2	31.8	4.28%
8月	745	742.6	2.4	0.32%
9月	726	681.0	45.0	6.20%
10月	823	793.4	29.6	3.59%
11月	833	820.6	12.4	1.49%
12月	872	888.2	-16.2	-1.86%
相対誤差（絶対値）の平均値				3.96%

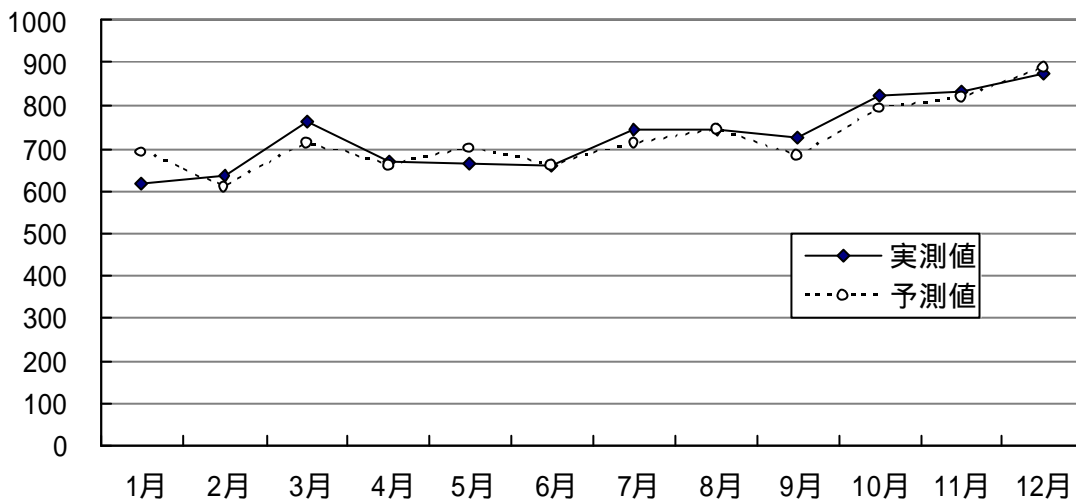


図2.2.4 交通事故死者数の月別推移の比較（2001年）

以上のとおり、この回帰手法を用いると、年次別推移としてみた場合の交通事故死者数の下降傾向が取り込まれるとともに、月次別の推移も織り込まれることにより、図2.2.4をみれば分るとおり、月次別の予測値が実測値に実によく適合している。

つまり、この回帰分析手法を用いれば、出国日本人数の場合のような、ニューヨーク同時多発テロといった、その前日まで誰も予測し得ない突発的な事象の織り込みは無理にしても、少なくとも今回取り上げた観察期間における交通事故死者数の予測といった、時系列的にみるとトレンドがみられるような場合には、その傾向をも包含された形での予測が可能であることが分かった。

3．おわりに

本論文では、具体的な事例を用いて、数量化理論 類と呼ばれる多変量解析手法を用いた時系列回帰を試みた。もちろん、一般的に行われている需要予測においては、単に年次や月次などを用いた時系列回帰に止まらず、いろいろな先行指標を用いた重回帰分析手法等を用いて需要予測を行うことが多い。これらの手法のうち、どちらがより優れているかという問題ではなく、まず、ブラックス・モデルとして、過去データを活用した時系列回帰を行うことによって、大よそのトレンドを把握し、併せて先行指標を用いた重回帰分析手法も行ってみるというように、双方からアプローチするという方法がとられる場合もある。それらの手法のさらにきめ細かい活用事例等については、別の機会に紹介したいと思っている。

【参考文献】

- 1) 国立社会保障・人口問題研究所：日本の将来推計人口
- 2) 法務省：出入国管理統計 2001.
- 3) 同友館：企業診断2002.10、vol.49
- 4) 交通事故総合センター：交通事故死者数の月別推移

(研究部)