

不具合発生率と整備不良事故発生率の回帰モデルの検証について

国土交通省より、不具合発生率（X:%）と整備不良事故発生率（Y:百万台当たりの事故発生件数）の相関式の推計に用いたデータ（別紙）をご提供いただいた（3月11日（金）夕刻）ので、標記検証を行った。

1. 回帰係数が有意な車種は3車種のみ

提供されたデータを用いた単回帰モデル（ $Y = \quad + \quad \times X$ ）の推計結果は、表1のとおり。

・回帰係数のt-検定（有意水準5%）において、有意なものは、自家用乗用車（普通・小型）、8トン以上自家用貨物車、8トン以上事業用貨物車の3車種であり、他の3車種は有意でなかった。また、回帰式全体についてのF-検定（有意水準5%）では、8トン以上自家用貨物車も有意でなく、両検定にパスしたものは、自家用乗用車（普通・小型）、8トン以上事業用貨物車の2車種であった。

・提供されたデータで推計したが、事業用乗用車、8トン以上自家用貨物車、乗合車の推計結果については、国土交通省よりの回答（平成17年2月17日付 別紙の項目17）に示された推計結果（ただし、回帰係数のみの回答で他の統計量は提示されていない）と異なるものとなった。

・なお、二輪車については、データが提供されていない。

2. 外れ値（車齢1年のデータ）を除くと有意な車種は自家用乗用車（普通・小型）のみ

・整備不良事故発生率（Y:百万台当たりの事故発生件数）は、車齢1年の値が他の車齢の値と比べて極端に小さい、いわゆる「外れ値」になっている（グラフ1）。車齢1年データが外れ値になっていることに合理的な理由があるか検証が必要である。

・外れ値を入れて回帰モデルの推計をすると、外れ値の影響で「見かけの相関」が高くなる。この影響をみるために、車齢1年のデータを除いて単回帰モデルを推計した（表2）。その結果、自家用乗用車（普通・小型）を除く5車種で極端に決定係数が低下しており、表1の推計結果が車齢1年のデータに引きずられていることが明瞭である。

・自家用乗用車（普通・小型）のみは、逆に、決定係数が高まり、回帰の標準誤差が大きく減少するなど、推計結果が改善する。この際に、回帰係数の推定値は、表1の2.43から1.37へとほぼ半減している（つまり、不具合率の上昇による事故件数の増加は表1の場合のほぼ半分になる）。

3. 区間推定でみた予測値の信頼度は低い

国土交通省の今回の調査では、上記(表1)の回帰モデルを用い、次式から、車検年延長による不具合率増加(ΔX)を原因とする整備不良事故発生率(Y :百万台当たりの事故発生件数)の増加(ΔY)を予測(点推定)していると考えられる。

$$\Delta \hat{Y} = \hat{\beta} \times \Delta X$$

しかし、外れ値を推計に入れた場合には、見かけの相関は高くなる一方、回帰の標準誤差が大きくなるために、予測値の分散が大きくなり、その信頼性は小さくなる。この点を見るために、表1の推計結果で回帰式が有意であった、自家用乗用車(普通・小型)、8トン以上事業用貨物車の2車種について、上記予測値の95%の信頼区間を計算した(ΔX の関数となる)。

結果は、グラフ2のとおりであり、自家用乗用車(普通・小型)について、不具合率10%ポイントの上昇があった場合、点推定では、百万台当たりの事故発生件数は24.34件増加するという結果であるが、95%の信頼区間は、-52.30~100.98件とマイナスの領域を含むものとなり、不具合率の増加が事故発生件数の増加をもたらすことの統計的信頼度は低い(予測値=0という帰無仮説を有意水準5%のt-検定で棄却できない)。

なお、回帰モデルの説明変数 X (不具合率)の非車検年のデータは、(1)「走行劣化+経年劣化」効果の推計および(2)製品劣化曲線の推計に基づく推計値であり、これらの推計誤差の影響も無視できないと思われるが、(1)(2)ともに自由度ゼロ(=標準誤差無限大)の推計(注)であり誤差の影響の統計的評価ができない。

(注)例えば、製品劣化曲線は、原点通過制約のある指数関数 $1 - \exp(- (tm / \quad))$ の2つのパラメータ(m 、 \quad)を2点のみのデータで推計しており、自由度ゼロ(=標準誤差無限大)である。

以上

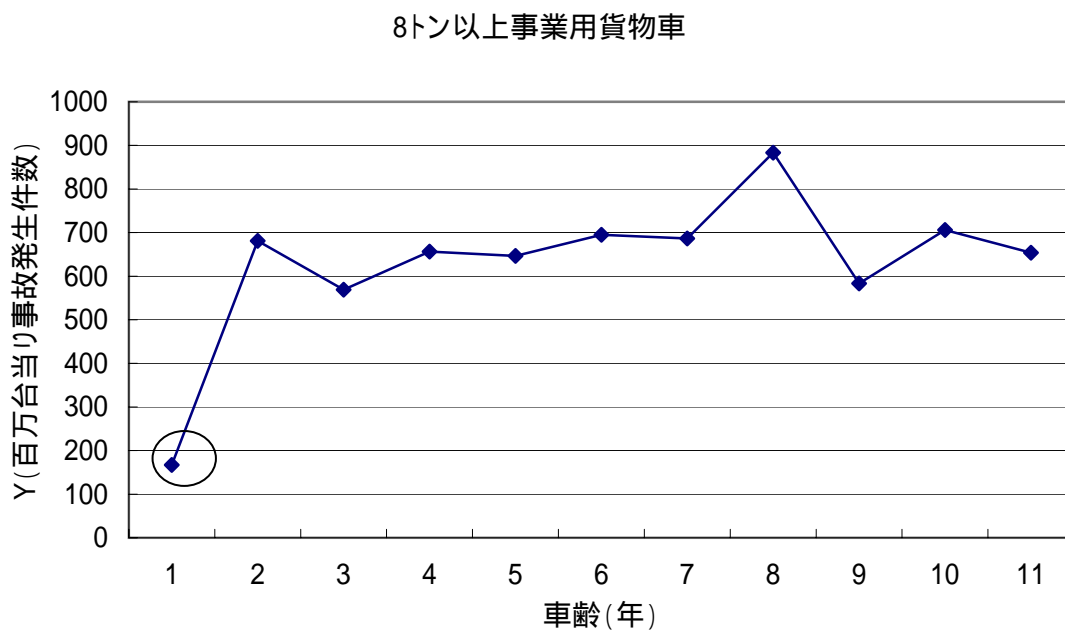
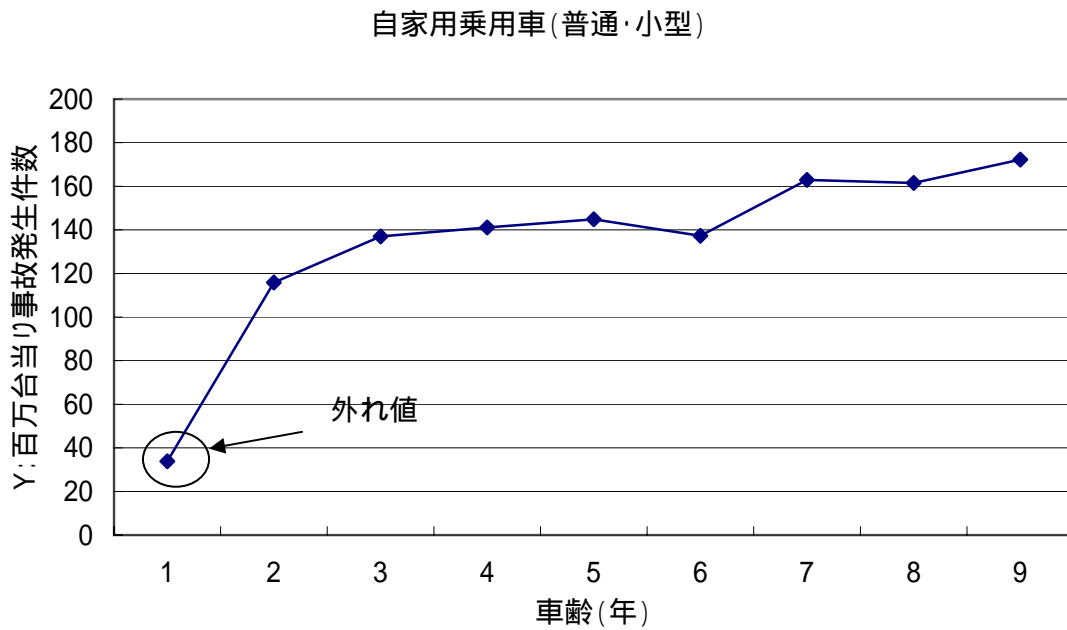
表1 回帰モデルの推計結果

車種	回帰係数(下段はt-値)		サンプル数 n	決定係数 R ²	回帰の標準誤差 S.E.	F-値	のt-検定 有意水準 5%	回帰式の F-検定 有意水準 5%	備考 *印は国土交通省 の推計結果と大きく 異なるもの
	(切片)	(傾き)							
自家用乗用車(普通・小型)	49.10 2.38	2.43 4.43	9	0.74	22.58	19.67			
自家用乗用車(軽)	51.18 4.56	0.44 1.63	9	0.27	15.11	2.64	×	×	
事業用乗用車	-268.48 -0.95	13.18 2.02	7	0.45	93.41	4.08	×	×	*
8トン以上自家用貨物車	32.83 0.27	4.49 2.26	9	0.42	58.99	5.10		×	*
8トン以上事業用貨物車	-260.08 -1.14	13.63 3.94	11	0.59	110.91	15.53			
乗合車	-39.73 -0.22	6.12 1.18	8	0.19	105.65	1.39	×	×	*

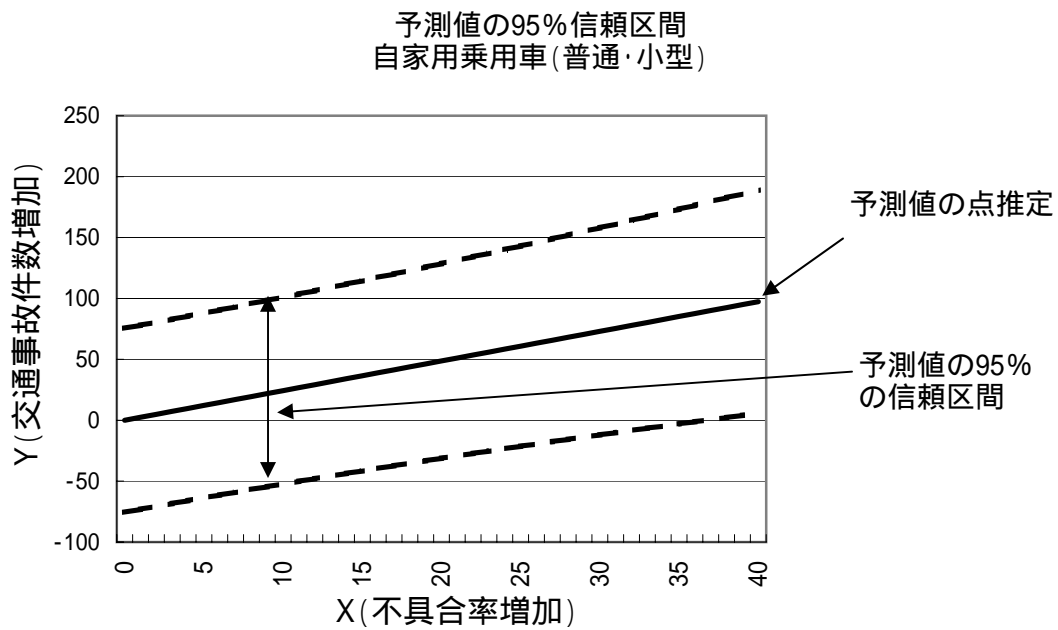
表2 車齢1年のデータを除いた場合の推計結果

車種	回帰係数(下段はt-値)		サンプル数 n	決定係数 R ²	回帰の標準誤差 S.E.	F-値	のt-検定 有意水準 5%	回帰式の F-検定 有意水準 5%	
	(切片)	(傾き)							
自家用乗用車(普通・小型)	94.63 11.49	1.37 6.60	8	0.88	6.81	43.56			
自家用乗用車(軽)	72.01	0.01	8	0.00	8.87	0.00	×	×	
事業用乗用車	158.81 0.37	3.87 0.41	6	0.04	87.77	0.17	×	×	
8トン以上自家用貨物車	341.60 2.44	-0.35 -0.16	8	0.00	42.09	0.02	×	×	
8トン以上事業用貨物車	370.32 1.26	4.52 1.05	10	0.12	85.33	1.10	×	×	
乗合車	196.04 0.77	-0.16 -0.02	7	0.00	100.81	0.00	×	×	

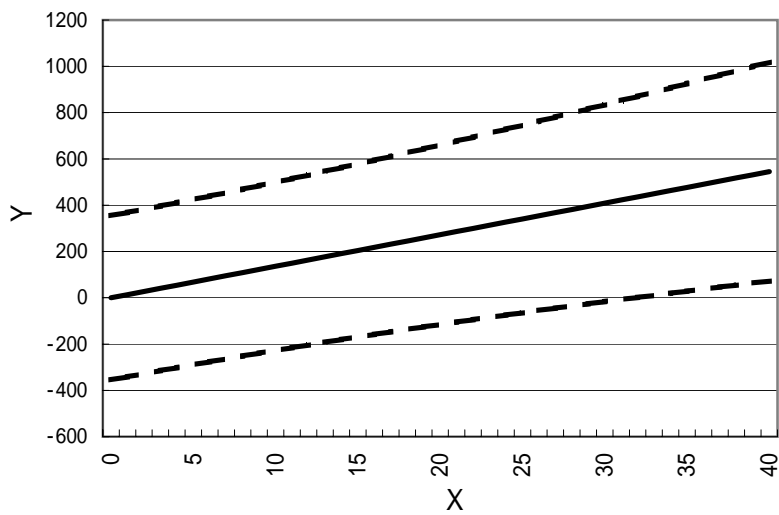
グラフ1 車齢別の整備不良事故発生率(Y:百万台当り事故発生件数)



グラフ2 予測値の区間推定



同
8トン以上事業用貨物車



車齡別交通事故発生率及び車齡別不具合発生率

車齡	自家用乗用車(普通・小型)		自家用乗用車(軽)		事業用乗用車		8才以上自家用貨物車		8才以上事業用貨物車		乗合車	
	不具合率	交通事故発生率	不具合率	交通事故発生率	不具合率	交通事故発生率	不具合率	交通事故発生率	不具合率	交通事故発生率	不具合率	交通事故発生率
1	11.4284	33.8685	6.3354	28.2705	32.9009	97.4327	37.6289	131.9488	41.2338	167.0922	20.1882	0.0000
2	20.1942	115.8953	17.6620	67.1615	38.3074	303.5971	50.6787	326.8265	53.5907	681.0110	26.2761	191.9982
3	26.4646	137.0354	28.9885	81.0905	41.7324	357.4413	53.0488	335.4253	59.3706	568.9185	30.4855	227.2496
4	30.0958	141.1170	28.6704	84.3016	44.2727	355.9783	59.7536	261.3374	65.1990	656.2140	33.7745	86.7537
5	40.0640	144.8769	51.6416	64.9485	46.3020	185.6184	61.9469	379.1559	67.5834	646.4458	36.4985	250.8213
6	36.4296	137.2960	33.9360	60.8603	47.9956	415.2215	65.3137	309.8097	70.5797	694.8736	38.8343	287.8122
7	49.9234	162.0900	61.0847	72.0844	49.4508	373.4692	67.7108	333.5269	70.7294	686.6280	40.8842	40.8110
8	41.7287	161.5602	36.5161	70.3926			67.4653	296.3851	73.0297	883.3727	42.7136	
9	57.9335	172.2509	65.3994	78.2677			70.0629	344.4627	70.7147	583.2349		246.9483
10									71.9651	706.0528		
11									74.3012	653.7585		