

都道府県別データから見る 交通事故発生要因の研究

東海大学 情報通信学部 経営システム工学科
須田泰珠

目次

1. 研究背景・事前調査
 - 先行研究
 - 1-1.近年の交通事故件数推移
 - 1-2.高齢者人口と交通事故
 - 1-3.飲酒運転と交通事故
 2. 分析方法
 - 2-1.分析の方針
 - 2-2.使用データ
 - 2-3.使用データから用いる変数
 - 2-4.分析の流れ
 3. 分析
 - 3-1.データの整理
 - 3-2.主成分分析
 - 3-3.クラスター分析
 4. 仮説・検証
 5. まとめと課題
- 参考文献
- appendix

先行研究

- **交通統計(2017年度版)^[12]**

公益財団法人 交通事故分析センターが毎年発表している、年別の交通事故件数推移や、都道府県別の交通事故発生件数をまとめた統計データ

- **高齢者の道路横断中の交通事故^[13]**

高齢者の道路横断中における交通事故において、考えられる要因やある条件での死亡率などを分析・解説している。



高齢者運転の交通事故への関連や、都道府県ごとの交通事故要因の分析はされていないようであった

1-1. 近年の交通事故件数推移

年度(和暦)	事故件数
12	931,950
13	947,253
14	931,950
15	946,253
16	952,720
⋮	⋮
22	725,924
23	692,084
24	665,157
25	629,033
26	573,842

増加

過去最多

減少

交通事故件数
H16年度まで増加傾向
以降、徐々に減少傾向

予想される
原因

- ・高齢者人口の増加
- ・飲酒運転に対する意識変化
- ・電気自動車の普及

表1-1. 年度ごとの交通事故件数推移^[1]

1-2.高齡者増加と交通事故

分析に使用

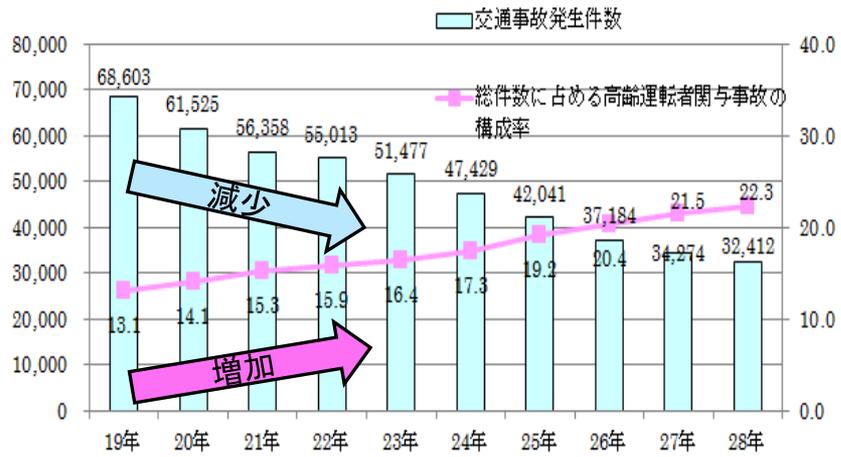


図1-1. 交通事故発生件数と高齢運転手^[2]

表1-2. 交通事故発生件数と高齢運転手

年度 (和暦)	高齢者関与と事故割合	交通事故発生件数	高齢者関与事故件数
19	13.1	68,603	8987
20	14.1	61,525	8675
21	15.3	56,358	8623
22	15.9	66,013	10496
23	16.4	51,477	8442
24	17.3	47,429	8205
25	19.2	42,041	8072
26	20.4	37,184	7586

事故件数に占める高齢者ドライバー関与の構成率

- **増加**
- 交通事故発生件数は減少傾向
- 高齢者人口(高齢の運転免許保有者人口)と事故数は関連している?
- 交通事故発生件数と高齢者関与交通事故件数には相関があった
- 高齢者の関与割合は増加傾向にあるが、実際の高齢者事故件数は減少しているといえる

表1-3. 交通事故発生件数と高齢者関与の相関

		交通事故発生件数	高齢者関与事故件数
交通事故発生件数	Pearson の相関係数	1	.811*
	有意確率 (両側)		.015
	度数	8	8
高齢者関与事故件数	Pearson の相関係数	.811*	1
	有意確率 (両側)	.015	
	度数	8	8

1-3. 飲酒運転と0交通事故

分析に使用

表1-4. 総事故件数に占める飲酒運転事故の割合

年度(和暦)	飲酒運転 事故件数	交通事故 発生件数	発生件数に占め る飲酒運転率
7	7,561	68,603	11.02
8	6,219	61,525	10.11
9	5,726	56,358	10.16
10	5,556	66,013	8.42
11	5,030	51,477	9.77
12	4,603	47,429	9.71
13	4,335	42,041	10.31
14	4,155	37,184	11.17



図1-2. 飲酒運転の関わる交通事故発生件数推移^[3]

交通事故件数に占める飲酒運転 関与の事故率

- **大きな変化なし**
- 飲酒運転と交通事故との関連は低いのか
- 相関分析の結果、相関があった
- 飲酒運転事故件数の減少とともに交通事故発生件数も減少する

表1-5. 総事故件数と飲酒運転事故件数の相関

		飲酒運転 事故件数	交通事故 発生件数
飲酒運転 事故件数	Pearson の 相関係数	1	.897**
	有意確率 (両側)		.003
	度数	8	8
交通事故 発生件数	Pearson の 相関係数	.897**	1
	有意確率 (両側)	.003	
	度数	8	8

2-1.分析の方針

目的

研究背景から判別できない、交通事故発生に関わる要因を発見する

都道府県ごとのデータを用いて事故件数と関わる要因を把握する

分析手順

各都道府県別の交通事故件数、事故発生に関わると思われるデータから分析
(本分析では2014年度のものを使用)



2-2.使用データ

- 警視庁 道路の交通に関する統計 交通事故の発生状況
平成28年度における交通事故の発生状況 状況(P2)^[1]
- 保険の窓口 インズウェブ 交通事故発生ランキング(都道府県別)
都道府県別道路交通事故(平成26年)^[2]
- 都道府県データランキング 高齢者ドライバー^[5]
- 都道府県データランキング 自家用車普及台数^[6]
- 都道府県データランキング 自家用車保有台数推移 【2014年】^[7]
- 都道府県データランキング 人口・面積・人口密度 【2014年】^[8]
- 交通事故発生ランキング(都道府県別)
- 年平均気温の都道府県ランキング 年平均気温ランキング^[9]
- 年間快晴日数の都道府県ランキング 年間快晴日数ランキング^[10]
- 年間快降水日数の都道府県ランキング 年間降水日数ランキング^[11]

2-3.使用データから用いる変数

1. 交通事故件数
2. 運転免許保有者全体
3. 免許保有者あたりの事故率
4. 運転免許保有者65歳以上
5. 自家用乗用車台数
6. 世帯あたり台数(台/世帯)
7. 乗用車密度(台/km²)
8. 1台あたり人口(人/台)
9. 1人あたり台(台/人)
10. 増減台数
11. 増減率
13. 鉄道駅数
14. 人口(人)
15. 面積(km²)
16. 人口密度(人/km²)
17. 標高(m)
18. 平均気温
19. 年間快晴日数
20. 年間降水日数
21. 65歳以上人口 15-64歳人口
22. 15歳以下人口

2-4.分析の流れ

1. 分析データの整理

使用データから選択したデータから、目的変数に関連の強いものを削除

2. 残ったデータ間の関連を探る

主成分分析を用いて整理後のデータ間で関連の強いものをまとめる

3. 都道府県をグループ分け

主成分分析を用いてクラスタ分析を行い、都道府県をグループ分け・グループの特徴を探る

4. 仮説立てと検証

推定されたグループの特徴から交通事故に影響を及ぼす要因を推定、実際の関わりを検証する

3-1.データの整理

相関係数 < 0.80

表3-1.データクリーニング後使用データ

初期データ:22変数

相関分析

目的半数との相関
が強い変数を削除
(0.80以上)

データ整理後:9変数

(内訳-目的変数:

- ・交通事故件数
- ・免許保有者当たりの事故率)

説明変数:

- ・世帯当たり台数
- ・鉄道駅数
- ・人口密度
- ・標高
- ・平均気温
- ・年間快晴日数
- ・年間降水日数

	交通事故件数	免許保有者あたりの事故率	世帯あたり台数	鉄道駅数	人口密度	標高	平均気温	年間快晴日数	年間降水日数
交通事故件数	1	.380*	-.542**	.756**	.708**	-.125	.232	.428**	-.411**
Pearsonの相関係数									
有意確率(両側)		.010	.000	.000	.000	.412	.120	.003	.005
度数	45	45	45	45	45	45	45	45	45
免許保有者あたりの事故率	.380*	1	.017	-.070	.004	-.043	.413**	.330*	-.430**
Pearsonの相関係数									
有意確率(両側)	.010		.910	.646	.978	.781	.005	.027	.003
度数	45	45	45	45	45	45	45	45	45
世帯あたり台数	-.542**	.017	1	-.638**	-.694**	.301*	-.291	-.057	.308*
Pearsonの相関係数									
有意確率(両側)	.000	.910		.000	.000	.045	.053	.709	.040
度数	45	45	45	45	45	45	45	45	45
増減台数	.766**	.202	-.208	.536**	.308*	-.058	.067	.415**	-.309*
Pearsonの相関係数									
有意確率(両側)	.000	.184	.171	.000	.040	.706	.662	.005	.039
度数	45	45	45	45	45	45	45	45	45
増減率	-.287	.149	.352*	-.510**	-.407**	.028	.328*	-.162	.043
Pearsonの相関係数									
有意確率(両側)	.056	.328	.018	.000	.006	.853	.028	.287	.781
度数	45	45	45	45	45	45	45	45	45
鉄道駅数	.756**	-.070	-.638**	1	.740**	-.052	-.108	.211	-.187
Pearsonの相関係数									
有意確率(両側)	.000	.646	.000		.000	.736	.480	.164	.220
度数	45	45	45	45	45	45	45	45	45
面積.km2	-.075	-.251	-.043	.324*	-.194	.047	-.525**	-.058	.206
Pearsonの相関係数									
有意確率(両側)	.625	.097	.777	.030	.201	.758	.000	.705	.176
度数	45	45	45	45	45	45	45	45	45
人口密度.人km2	.708**	.004	-.694**	.740**	1	-.110	.216	.263	-.268
Pearsonの相関係数									
有意確率(両側)	.000	.978	.000	.000		.471	.154	.081	.076
度数	45	45	45	45	45	45	45	45	45
標高m	-.125	-.043	.301*	-.052	-.110	1	-.389**	-.023	-.210
Pearsonの相関係数									
有意確率(両側)	.412	.781	.045	.736	.471		.008	.883	.167
度数	45	45	45	45	45	45	45	45	45
平均気温	.232	.413**	-.291	-.108	.216	-.389**	1	.126	-.323*
Pearsonの相関係数									
有意確率(両側)	.126	.005	.053	.480	.154	.008		.409	.030
度数	45	45	45	45	45	45	45	45	45
年間快晴日数	.428**	.330*	-.057	.211	.263	-.023	.126	1	-.494**
Pearsonの相関係数									
有意確率(両側)	.003	.027	.709	.164	.081	.883	.409		.001
度数	45	45	45	45	45	45	45	45	45
年間降水日数	-.411**	-.430**	.308*	-.187	-.268	-.210	-.323*	-.494**	1
Pearsonの相関係数									
有意確率(両側)	.005	.003	.040	.220	.076	.167	.030	.001	
度数	45	45	45	45	45	45	45	45	45

3-2.主成分分析

分析を用いる理由

データの整理でまとめたデータ間において、関連性の強いものをまとめ、クラスター分析に使用する成分を作成する

結果

S-plusによる主成分分析の結果から、カイザー・ガットマン基準※1を用いて使用する成分を決める
主成分1,2,3が**固有値1.0以上**である。
→第3主成分を採用し、クラスター分析に使用する

※1 カイザー・ガットマン基準：
因子数の基準となる固有値の最小値を1とする

Relative Importance of Principal Components

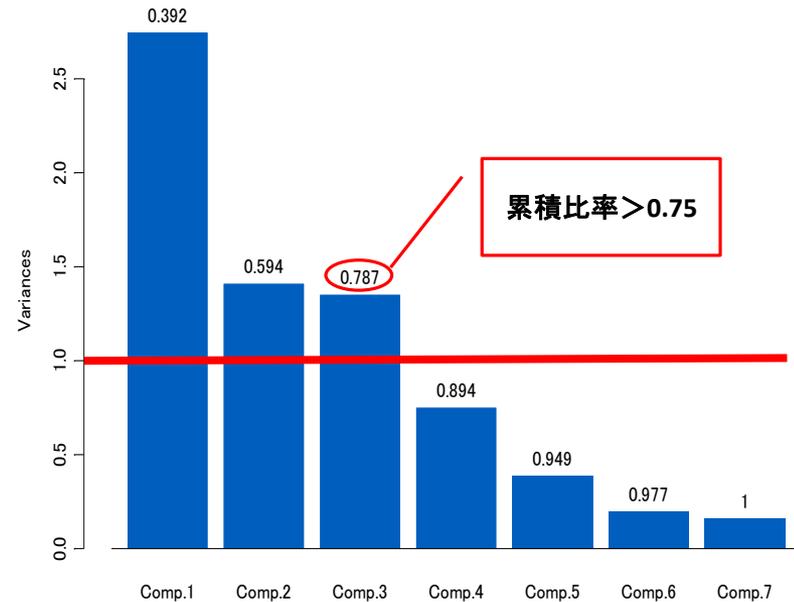


図3-1.生成要因の累積比率

表3-2.各成分に対する変数の固有値

	成分		
	1	2	3
鉄道駅数	0.922	0.07	0.16
人口密度.人km ²	0.876	0.216	-0.109
世帯あたり台数台世帯	-0.829	-0.092	0.334
年間降水日数	-0.159	-0.885	0.033
年間快晴日数	0.115	0.772	-0.001
平均気温	-0.01	0.354	-0.832
標高m	-0.158	0.273	0.811
因子抽出法: 主成分分析			

3-3. クラスタ分析

分析を用いる理由

主成分分析の結果をもとに、都道府県をいくつかのグループに分類する
分類したグループごとの特徴を推定する

結果

- クラスタ①: 東北地方中心
- クラスタ②: 人口数が比較的少ない
人口密度低い傾向
- クラスタ③: 首都圏周辺
- クラスタ④: その他
- クラスタ⑤: 主要都市
人口密度Top3

人口密度が
関係？

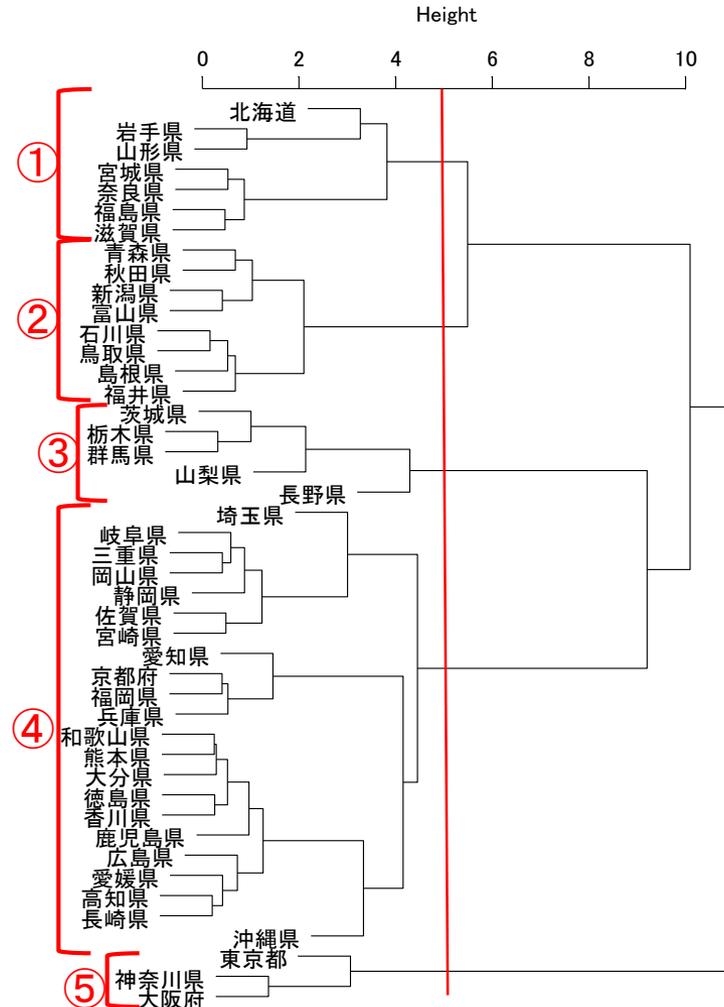


図3-2. クラスタ分析で得られたデンドログラム

4.仮説・検証

主成分分析・クラスター分析の結果をもとに仮説を立て、検証

仮説

人口密度の大きさと交通事故発生件数にも関連があるのではないかと

検証方法

・都道府県別データのうち、人口密度と交通事故発生件数の相関を調べる

表4-1.交通事故件数と人口密度の相関

		交通事故件数	人口密度 (人/km ²)
交通事故 件数	Pearson の 相関係数	1	.708**
	有意確率 (両側)		.000
	度数	45	45
人口密度 (人/km ²)	Pearson の 相関係数	.708**	1
	有意確率 (両側)	.000	
	度数	45	45

結果

交通事故発生件数と人口密度の間に強い相関が見られた
人口密度が多い都道府県は交通事故が多いといえる

5.まとめと課題

表5-1.交通事故件数(降順)と人口密度(Top6)

都道府県	交通事故件数	人口密度
愛知県	46131	1439.28
埼玉県	30821	1905.8
静岡県	33499	475.36
大阪府	42729	4646.01
東京都	37184	6106.43
福岡県	41168	1021.28

表5-2. 人口密度(降順)と交通事故件数(Top6)

都道府県	交通事故件数	人口密度
東京都	37184	6106.43
大阪府	42729	4646.01
神奈川県	30434	3766.43
埼玉県	30821	1905.8
愛知県	46131	1439.28
福岡県	41168	1021.28

まとめ

事前調査から、飲酒運転の減少とともに交通事故発生
件数も減少すると推定できた

分析では高齢者人口と事故件数の相関が高すぎたこと
からデータクリーニングを行い、都道府県ごとの人口密度
と交通事故発生件数には強い相関があると分かった

「人口密度が高いほど交通事故が起こりやすく、低いほど
交通事故が起こりにくい」と言えるが、表5-1・5-2より、
主要都市よりも人口密度が低い、地方都市のほうが交通
事故が起きやすいと分かる

5.まとめと課題

考察

- ・北海道・福島県・山形県などは前後と比べて人口密度が低いが、交通事故件数は多く現れている
- ・京都府・沖縄県などは人口密度が高いにも関わらず交通事故件数がそれほど多くない

→各都道府県の面積や、観光地としての整備が交通事故減少につながっているのかもしれない

今後の課題

本研究で扱えなかった電気自動車の普及や、乗用車の安全性向上による交通事故件数への影響を考慮したい
また、考察で例外と読み取った都道府県について、なぜ予測からずれたのかを調査する

表5-3.交通事故件数(降順)と人口密度

順位	都道府県	交通事故件数	人口密度
11	広島県	12479	334.18
12	北海道	12274	68.86
13	岡山県	12271	270.51
14	京都府	10185	566.67
15	宮崎県	9759	144.01
16	長野県	9283	155.47
17	宮城県	9142	319.69
18	香川県	8942	522.68
19	佐賀県	8870	342.13
20	鹿児島県	8425	181.57
21	岐阜県	8250	192.23
22	三重県	8100	315.27
23	福島県	7710	140.5
24	熊本県	7584	242.21
25	滋賀県	6598	352.59
26	長崎県	6465	335.29
27	山形県	6426	121.27
28	栃木県	6413	309.13
29	新潟県	6317	183.87
30	沖縄県	6242	623.64
31	奈良県	5868	372.93
32	愛媛県	5745	245.87

参考文献

[1]統計データ検索結果 政府統計の総合窓口 GL02020101

<http://www.e-stat.go.jp/SG1/estat/GL02020101.do?method=extendTclass&refTarget=toukeihyo&listFormat=hierarchy&statCode=00130002&statCode=000001027457&tclass1=&tclass2=&tclass3=&tclass4=&tclass5=> (最終閲覧日:2017/10/13)

[2]交通事故発生ランキング(都道府県別) | 自動車事故徹底ガイド | 保険の窓口インズウェブ

<https://www.insweb.co.jp/jidousya-jiko/jiko-ranking.html> (最終閲覧日:2017/10/13)

[3]防ごう!高齢者の交通事故! 警視庁

<http://www.keishicho.metro.tokyo.jp/kotsu/jikoboshi/koreisha/koreijiko.html> (最終閲覧日:2017/10/13)

[4]飲酒運転の実態 | 危険な飲酒運転 | お酒と健康 | キリン

<http://www.kirin.co.jp/csv/arp/driving/actualcondition.html> (最終閲覧日:2017/10/13)

[5]高齢者ドライバー

<http://uub.jp/pdr/t/sd.html> (最終閲覧日:2017/10/13)

[6]自家用車普及台数

<http://uub.jp/pdr/t/cr.html> (最終閲覧日:2017/10/13)

[7]自家用車保有台数推移

http://uub.jp/pdr/t/ct_2a.html (最終閲覧日:2017/10/13)

[8]人口・面積・人口密度

http://uub.jp/pdr/j/j_4.html (最終閲覧日:2017/10/13)

[9]年平均気温の都道府県ランキング - 都道府県格付研究所

<http://grading.jpn.org/SRB02101.html> (最終閲覧日:2017/10/13)

[10]年間快晴日数の都道府県ランキング - 都道府県格付研究所

<http://grading.jpn.org/SRB02301.html> (最終閲覧日:2017/10/13)

[11]年間降水日数の都道府県ランキング - 都道府県格付研究所

<http://grading.jpn.org/SRB02303.html> (最終閲覧日:2017/10/13)

[12]交通統計-交通事故総合分析センター 平成27年度版

<http://www.itarda.or.jp/materials/publications.php?page=4> (最終閲覧日:2017/10/13)

[13]第19回交通事故-調査分析研究発表会-交通事故総合分析センター

https://www.itarda.or.jp/ws/index_19.php (最終閲覧日:2017/10/13)

Appendix

主成分分析

表5-1.各主成分の固有ベクトル値

	Com p.1	Com p.2	Com p.3	Com p.4	Com p.5	Com p.6	Com p.7
世帯あたり台数/世帯	0.509	0.218	0.143	-0.257	0.399	-0.545	-0.388
鉄道駅数	-0.465		-0.455	-0.136		-0.683	0.294
人口密度/人km ²	-0.526		-0.218		0.532	0.236	-0.579
標高m	0.152	0.669	-0.216	0.482	0.347	0.145	0.33
平均気温	-0.233	-0.315	0.626	0.326	0.444	-0.231	0.313
年間快晴日数	-0.262	0.427	0.349	-0.695	0.118	0.24	0.27
年間降水日数	0.319	-0.468	-0.405	-0.304	0.475	0.222	0.385

第一主成分得点 = $0.509 * \text{世帯あたり台数} + (-0.465) * \text{鉄道駅数} + (-0.526) * \text{人口密度} + 0.152 * \text{標高}$
 $+ (-0.233) * \text{平均気温} + (-0.262) * \text{年間快晴日数} + 0.319 * \text{年間降水日数}$

第二主成分得点 = $0.218 * \text{世帯あたり台数} + 0.669 * \text{標高} + (-0.315) * \text{平均気温} + 0.427 * \text{年間快晴日数}$
 $+ (-0.468) * \text{年間降水日数}$

第三主成分得点 = $0.143 * \text{世帯あたり台数} + (-0.455) * \text{鉄道駅数} + (-0.218) * \text{人口密度} + (-0.216) * \text{標高}$
 $+ 0.626 * \text{平均気温} + 0.349 * \text{年間快晴日数} + (-0.405) * \text{年間降水日数}$

主成分分析

表3-2.各成分に対する変数の固有値(再掲)

	成分		
	1	2	3
鉄道駅数	0.922	0.07	0.16
人口密度.人km ²	0.876	0.216	-0.109
世帯あたり台数台世帯	-0.829	-0.092	0.334
年間降水日数	-0.159	-0.885	0.033
年間快晴日数	0.115	0.772	-0.001
平均気温	-0.01	0.354	-0.832
標高m	-0.158	0.273	0.811
因子抽出法:主成分分析			

抽出された3主成分の固有値から、それぞれの関連をまとめる

- ・主成分1: 鉄道駅数、人口密度、世帯あたり台数への固有値が高い
→ 地域ごとの発展度合いに関連しているのではないか
- ・主成分2: 年間降水日数、年間快晴日数への固有値が高い
→ 天候の変化に関連していると予測
季節ごとの交通事故件数を調べてみることも検討
- ・主成分3: 平均気温、標高への固有値が高い
→ 地形や地方の違いに関連していると予測できる

クラスター分析

・クラスター内平方和

$$E = \sum_{h=1}^k \sum_{i=1}^{n_h} \sum_{j=1}^p \left(x_{ji}^{(h)} - \bar{x}_j^{(h)} \right)^2$$

- ・ウオード法(Ward Method): クラスター内平方和が最小となるように D_{mh} を求める

$$D_{mh} = \frac{1}{n_m + n_h} \left\{ (n_f + n_h) D_{fh} + (n_g + n_h) D_{gh} - n_h D_{fg} \right\}$$

・クラスター分析の詳細

Call:

```
agnes(x = menuModelFrame(data = `最終`, variables = "Comp.1,Comp.2,Comp.3",
  subset = NULL, na.rm = TRUE), diss = FALSE, metric = "euclidean",
  stand = TRUE, method = "ward", save.x = TRUE, save.diss = TRUE)
```