

# 品質・技術力向上に繋げる SQCと機械学習の よりよい使い方について

---

トヨタ自動車(株) カスタマーファースト推進本部

監査改良領域 業務品質改善部

渡邊克彦

# トヨタグローバルビジョン

笑顔のために、期待を超えて

**人々を安全・安心に運び、心までも動かす。**

そして、世界中の生活を、社会を、豊かにしていく。  
それが、未来のモビリティ社会をリードする、  
私たちの想いです。

**一人ひとりが高い品質を造りこむこと。**

**常に時代の一步先のイノベーションを追い求めること。**

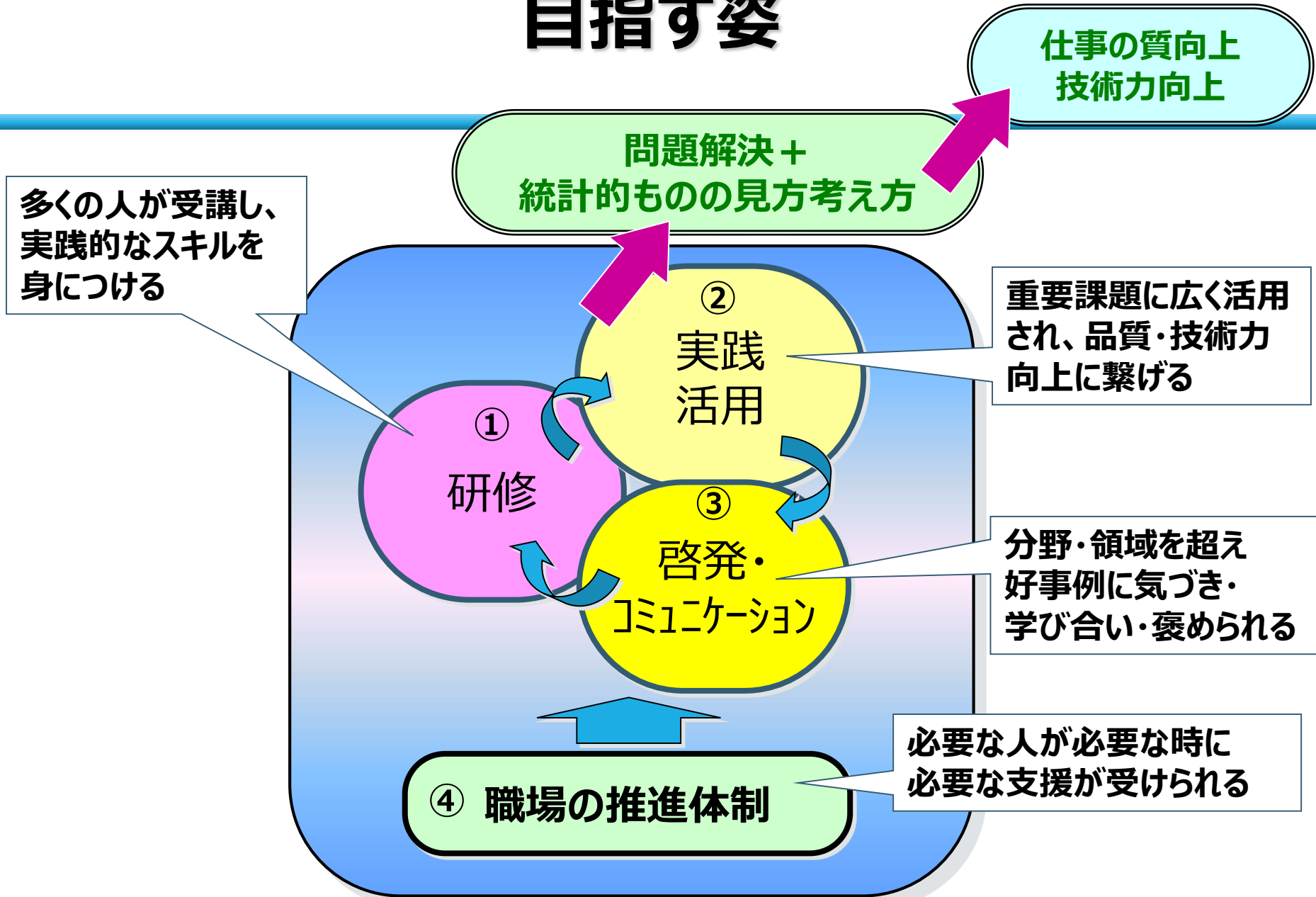
地球環境に寄り添う意識を持ち続けること。

その先に、期待を常に超え、  
お客様そして地域の笑顔と幸せに  
つながるトヨタがあると信じています。

**「今よりもっとよい方法がある」**その改善の精神とともに、

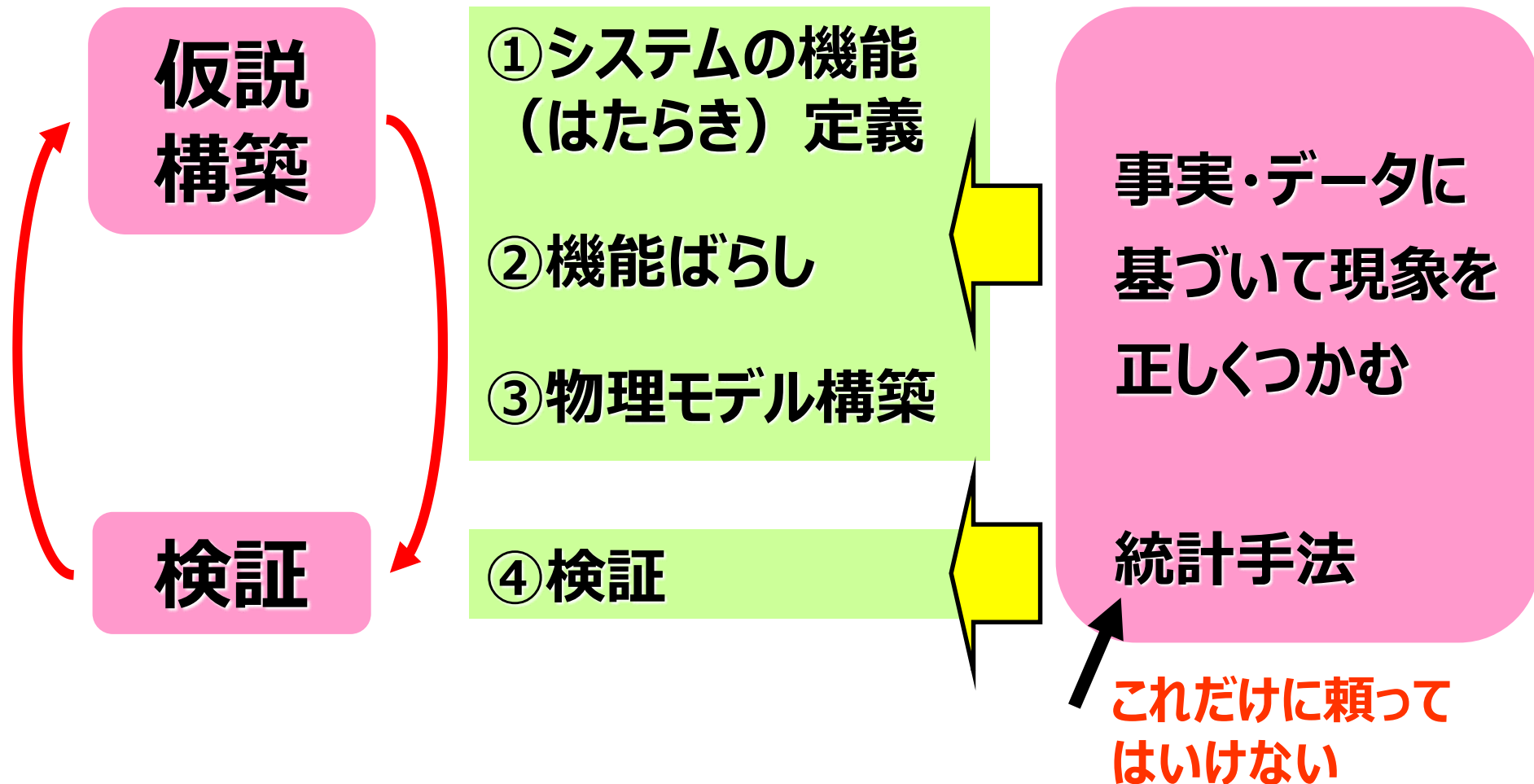
トヨタを支えてくださる皆様の声に真摯に耳を傾け、  
常に自らを改革しながら、高い目標を実現していきます。

# 目指す姿



TQMとは品質・仕事の質向上と、人と組織の活性化により、企業体質を強化する活動

# 原理原則に基づく あるべきサイクル



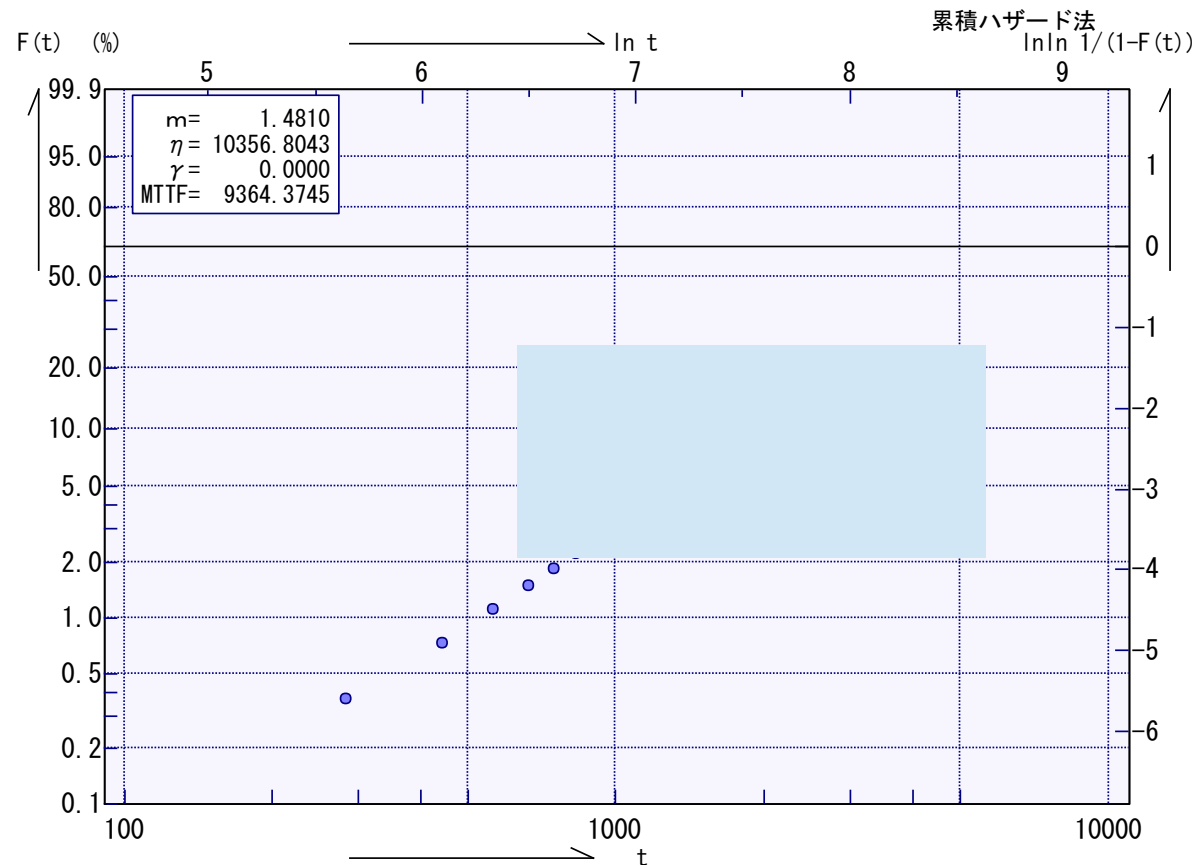
# 信頼性データ解析シンポジウム（2012年）より

# ③ ワイブル解析と S-S モデルの対応

2012年信頼性データ解析シンポジウム

市場故障データをワイブルプロット

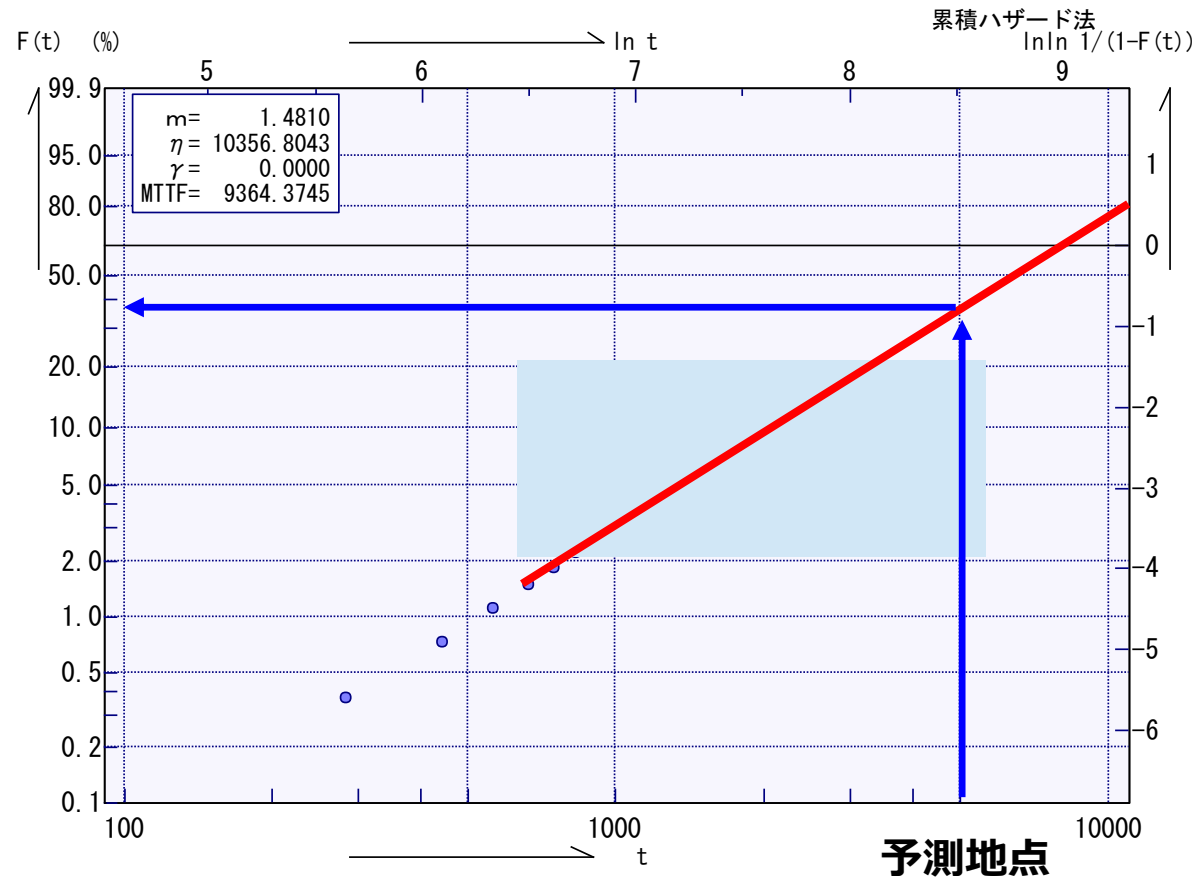
今後の故障数を予測する



# ③ワイブル解析とS-Sモデルの対応

2012年信頼性データ解析シンポジウム

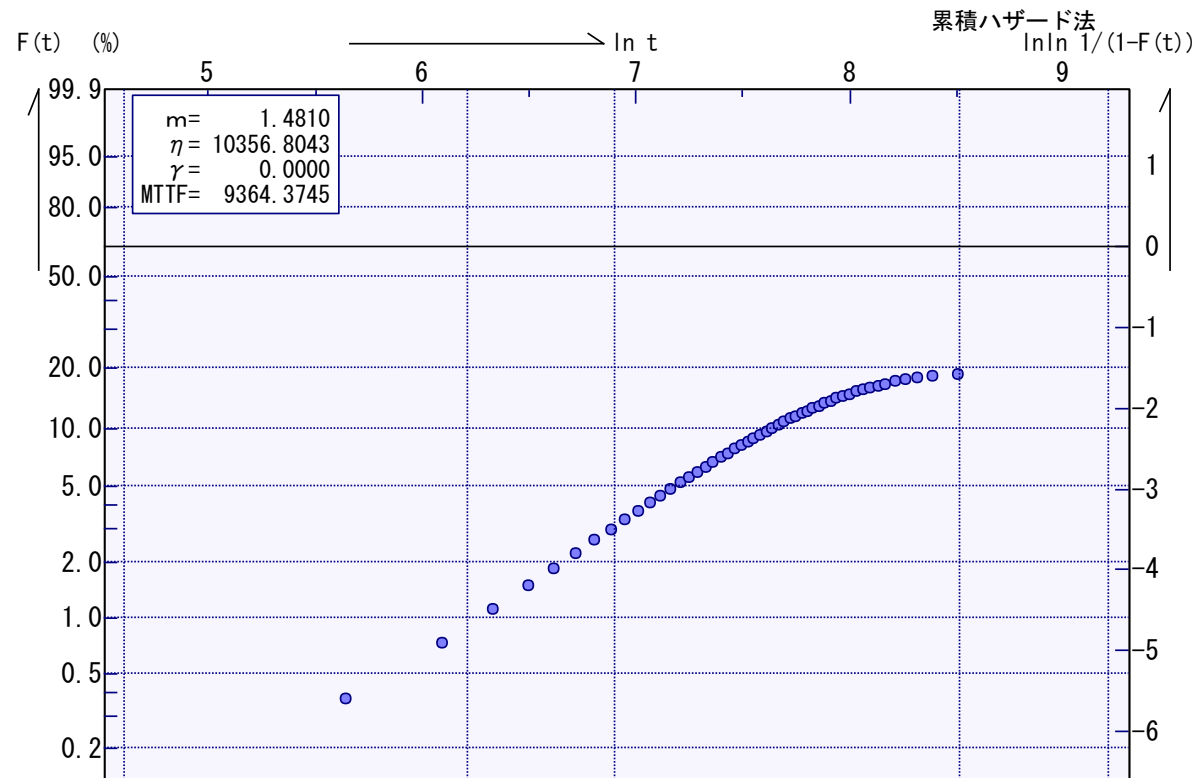
○年後に○%と予測



# ③ ワイブル解析と S-S モデルの対応

2012年信頼性データ解析シンポジウム

ところが実際は20%で飽和する



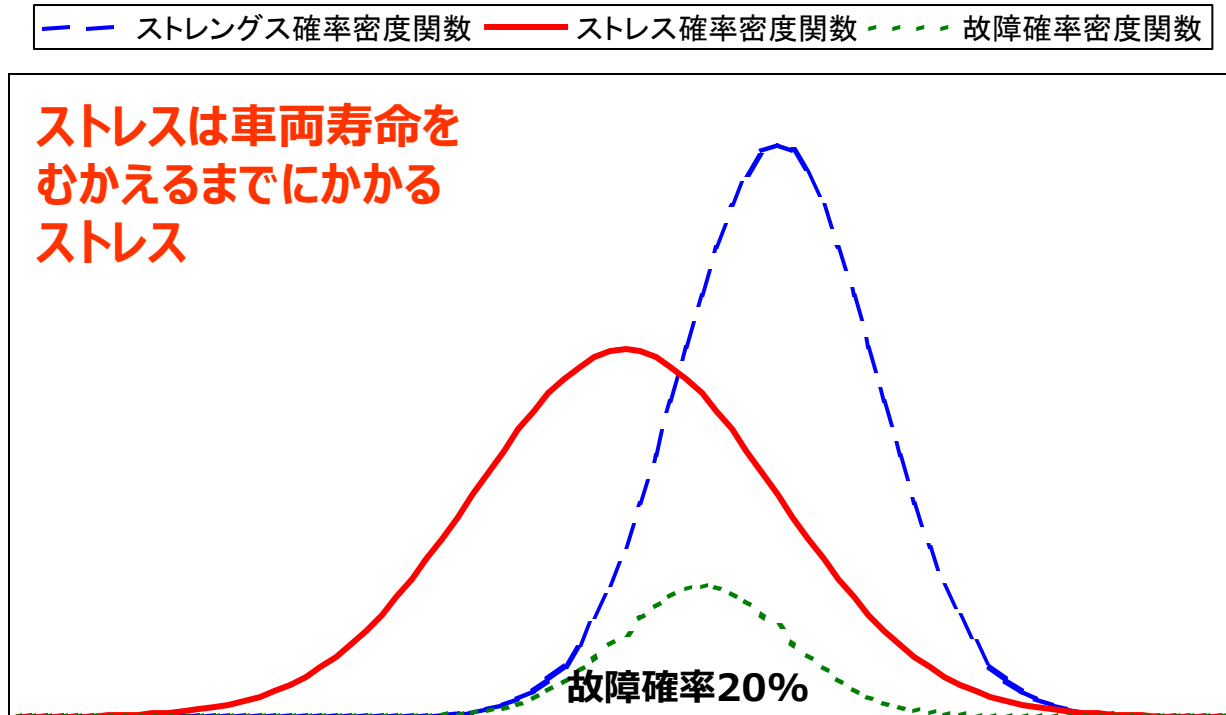
ワイブル解析に頼りすぎると物事の本質を見極められないことも考えられる



# ③ワイブル解析とS-Sモデルの対応

2012年信頼性データ解析シンポジウム

S-Sモデルを考察すると

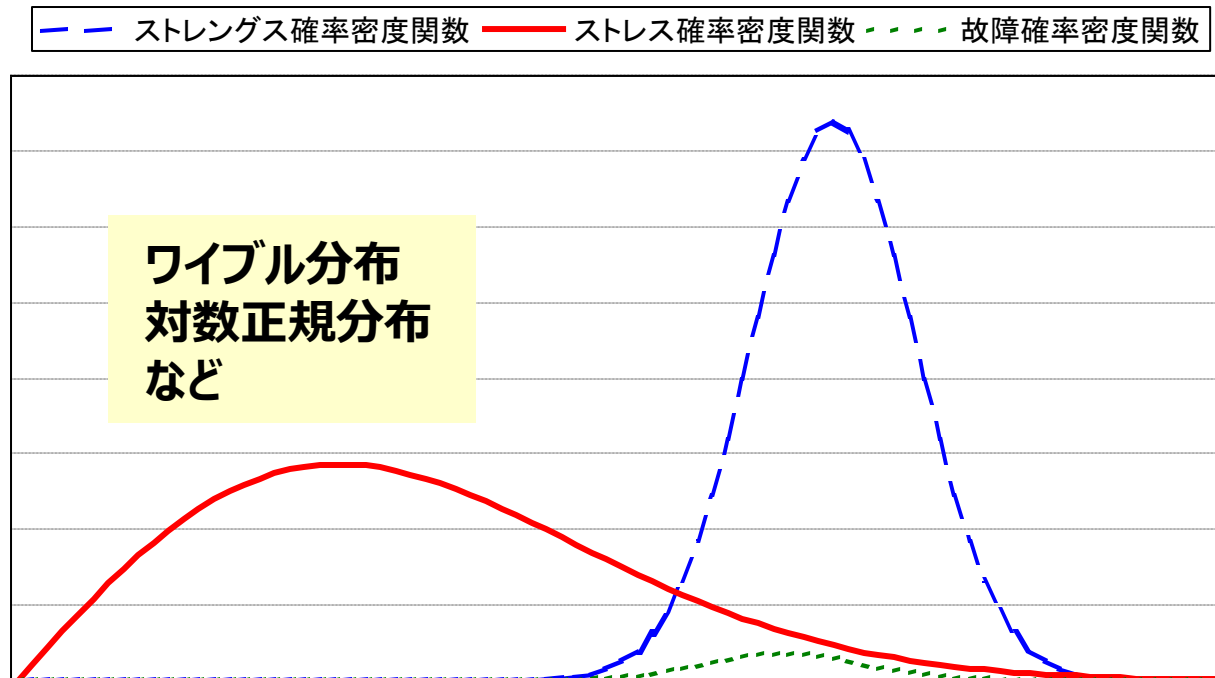


原理原則で20%以上は故障しない

# ③ ワイブル解析とS-Sモデルの対応

2012年信頼性データ解析シンポジウム

[補足]現実には正規分布以外の分布が多い

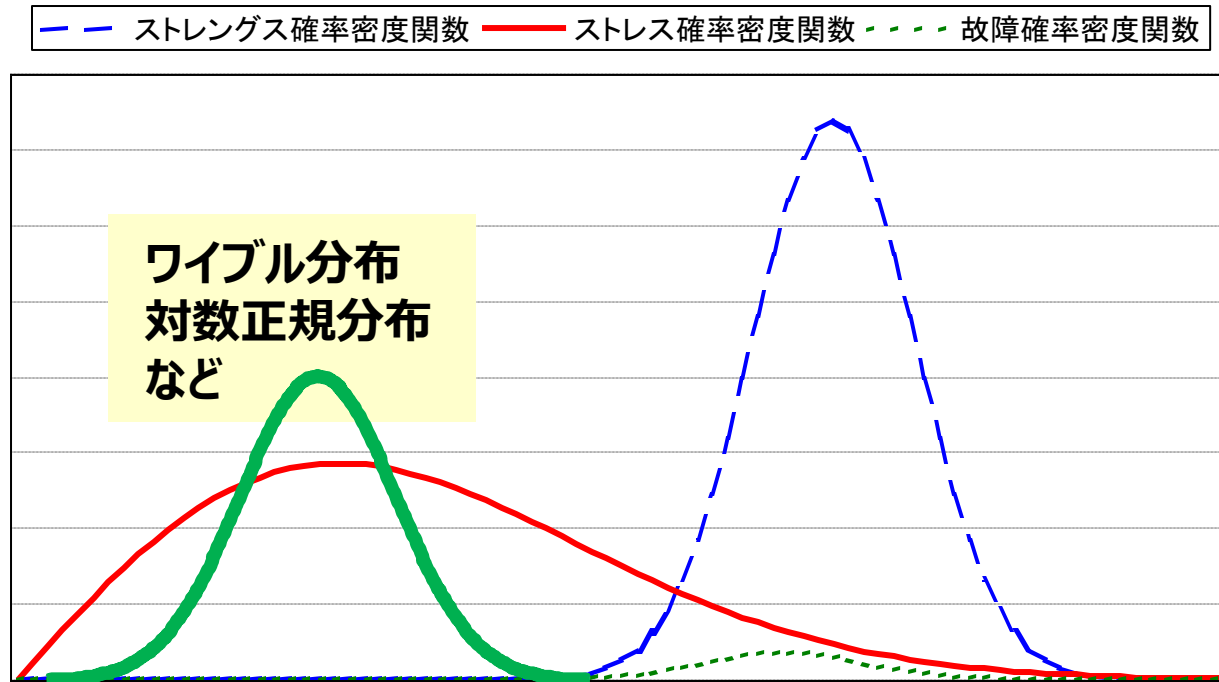


JUSE-StatWorks V5に新機能として追加

# ③ワイブル解析とS-Sモデルの対応

2012年信頼性データ解析シンポジウム

[補足]現実には正規分布以外の分布が多い



JUSE-StatWorks V5に新機能として追加

# ③ ワイブル解析とS-Sモデルの対応

2012年信頼性データ解析シンポジウム

## ③のまとめ

- ワイブル解析は100%故障が前提の解析である
- ストレス-ストレングスモデルなど原理原則と照らし合わせる  
ことが重要

正しい分布で計算することで技術者の  
誤った判断の防止に繋がる

# 対象とする手法、データサイズ

## JUSE-StatWorks/V5

解析の目的		● SQC (総合編プレミアムなど2011年～)	● 機械学習 (機械学習編 2019年～)
①データ可視化		多変量連関図、モニタリング	濃淡散布図、密度プロット、等高線図
②層別		階層的クラスター分析 非階層的クラスター分析(k-means法)	混合ガウス分布
③情報の要約		主成分分析	カーネル主成分分析
④予測 (回帰)		重回帰分析	正則化回帰分析 (リッジ回帰、lasso回帰、Elastic Net)
⑤分類	⑤- 1	判別分析	サポートベクターマシン (SVM)
	⑤- 2	AID、CAID	ランダムフォレスト
⑥外れ値検出		多変量管理図、MT法	1クラスSVM
⑦因果分析		グラフィカルモデリング	glasso

ミドルデータの  
領域が対象

データサイズ: 1,000変数 × 100,000サンプル

# トヨタでの解析ソフト

1980年代 TPOS (Toyota Promotional Original SQC Soft) を開発

TPOS32-PM 多変量解析ソフト

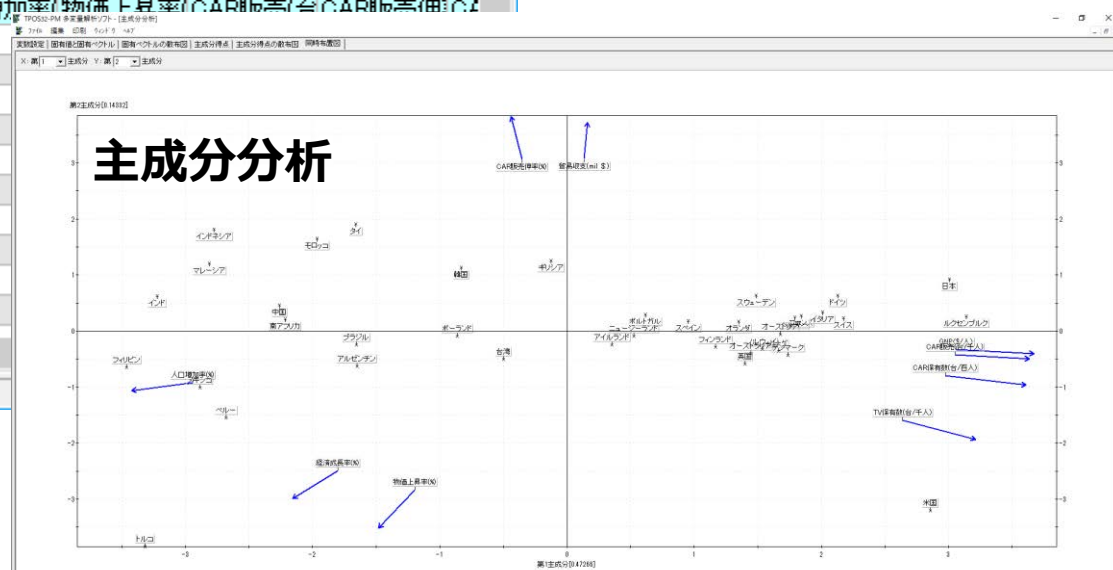
ファイル 編集 印刷 ウィンドウ ヘルプ

データ操作 - C:\Users\1333012\Documents\TQMでのデータ式\TPOSデータ\多変量07\_TPOS配布...

多変量連関図 基本統計量 重回帰分析 判別分析 主成分分析  
 因子分析 クラスター分析 数量化理論I類 数量化理論II類 数量化理論III類

変数の数:9 ケースの数:38

			X1	X2	X3	X4	X5	X6	X7
			<input checked="" type="checkbox"/>	<input checked="" type="checkbox"/>	<input checked="" type="checkbox"/>	<input checked="" type="checkbox"/>	<input checked="" type="checkbox"/>	<input checked="" type="checkbox"/>	<input checked="" type="checkbox"/>
			数値	数値	数値	数値	数値	数値	数値
			GNP(\$/人)	経済成長率	貿易収支(m)	人口増加率	物価上昇率	CAR販売台数	CAR販売価
C1	<input checked="" type="checkbox"/>	米国	29.08	3.8	-210.3	1.0			
C2	<input checked="" type="checkbox"/>	カナダ	19.64	3.8	13.5	1.3			
C3	<input checked="" type="checkbox"/>	メキシコ	3.70	7.0	-11.2	3.0			
C4	<input checked="" type="checkbox"/>	ブラジル	4.79	3.0	-12.0	1.4			
C5	<input checked="" type="checkbox"/>	アルゼンチン	8.95	8.4	-4.8	1.3			
C6	<input checked="" type="checkbox"/>	ペルー	2.61	7.2	-3.4	1.8			
C7	<input checked="" type="checkbox"/>	ドイツ	28.28	2.0	66.9	0.5			
C8	<input checked="" type="checkbox"/>	フランス	26.30	2.3	20.1	0.5			
C9	<input checked="" type="checkbox"/>	イタリア	20.17	1.5	30.0	0.0			
C10	<input checked="" type="checkbox"/>	英国	20.87	3.4	-25.5	0.4			



# 対象とする手法、データサイズ

## JUSE-StatWorks/V5

解析の目的	● SQC (総合編プレミアムなど2011年～)	● 機械学習 (機械学習編 2019年～)
①データ可視化	多変量連関図、モニタリング	濃淡散布図、密度プロット、等高線図
②層別	階層的クラスター分析 非階層的クラスター分析(k-means法)	混合ガウス分布
③情報の要約	主成分分析	カーネル主成分分析
④予測 (回帰)	重回帰分析	正則化回帰分析 (リッジ回帰、lasso回帰、Elastic Net)
⑤分類	⑤- 1 判別分析	サポートベクターマシン (SVM)
	⑤- 2 AID、CAID	ランダムフォレスト
⑥外れ値検出	多変量管理図、MT法	1クラスSVM
⑦因果分析	グラフィカルモデリング	glasso

データサイズ: 1,000変数 × 100,000サンプル

ワークシート 手法選択 解析 装飾

開く 保存 印刷 変数一括編集 変数属性 変数情報 挿入 削除 変換 編集 表示 ソート 範囲 全データ 欠測 解析対象 検索 ジャンプ データ 多変量 統計量 連関図 グラフ 基本解析 システム 接続

	● S1	● N2	● N3	● N4	● N5	● N6	● N7	● N8	● N9	● N10
	用紙銘柄	y1	y2	y3	y4	y5	y6	y7	y8	y9
● 1	base-paper1	118.6	87.8	19.14	0.95	3.52	2.22	11.00	1.27	
● 2	base-paper2	99.3	124.6	23.92	1.24	0.83	-10.24	10.91	1.43	
● 3	base-paper3	72.3	71.6	17.00	0.69	2.62	-3.87	11.69	1.24	
● 4	base-paper4	50.0	77.5	20.30	0.64	1.03	-10.91	10.33	1.60	
● 5	base-paper5	131.3	137.6	21.97	1.37	2.27	-5.64	10.92	1.50	
● 6	base-paper6	98.6	123.9	24.53	1.40	1.26	-10.30	10.74	1.56	
● 7	base-paper7	98.3	86.3	20.65	1.06	3.02	-4.07	11.09	1.39	
● 8	base-paper8	139.0	154.0	26.66	1.56	2.18	-6.45	11.39	1.41	
● 9	base-paper9	60.9	97.1	18.67	0.85	1.40	-13.17	10.28	1.34	
● 10	base-paper10	98.4	90.3	19.90	1.01	2.73	-6.44	10.85	1.27	
● 11	base-paper11	131.8	137.8	21.13	1.39	2.00	-9.19	10.94	1.39	
● 12	base-paper12	64.1	71.3	21.03	0.70	2.02	-5.85	10.60	1.48	
● 13	base-paper13	46.7	84.8	20.88	0.80	1.10	-10.72	10.40	1.36	
● 14	base-paper14	94.5	103.5	24.42	0.99	2.21	-8.18	11.10	1.67	
● 15	base-paper15	107.0	105.1	23.90	1.14	2.63	-4.28	11.11	1.56	
● 16	base-paper16	107.3	99.6	21.02	1.11	3.38	-7.00	11.45	1.50	
● 17	base-paper17	66.1	67.6	15.30	0.80	2.63	-6.29	11.08	1.33	
● 18	base-paper18	146.2	126.7	23.19	1.32	3.62	-4.35	10.97	1.35	
● 19	base-paper19	79.1	82.7	20.87	0.89	2.89	-5.00	11.06	1.63	
● 20	base-paper20	116.7	150.2	30.03	1.63	1.08	-11.43	11.00	1.64	
● 21	base-paper21	152.5	159.4	26.74	1.82	2.33	-7.38	10.23	1.46	
● 22	base-paper22	77.4	70.5	21.49	0.85	2.87	-6.97	11.62	1.62	
● 23	base-paper23	152.8	100.2	21.21	1.23	4.84	-1.02	11.14	1.28	
● 24	base-paper24	104.9	134.6	27.71	1.31	1.53	-9.71	10.63	1.58	
● 25	base-paper25	58.2	85.7	22.57	0.89	0.75	-12.75	10.23	1.60	
					1.41	1.36	-8.73	10.71	1.47	

StatWorks サンプルデータ ML\_M03\_02.SW5



# ① データ可視化

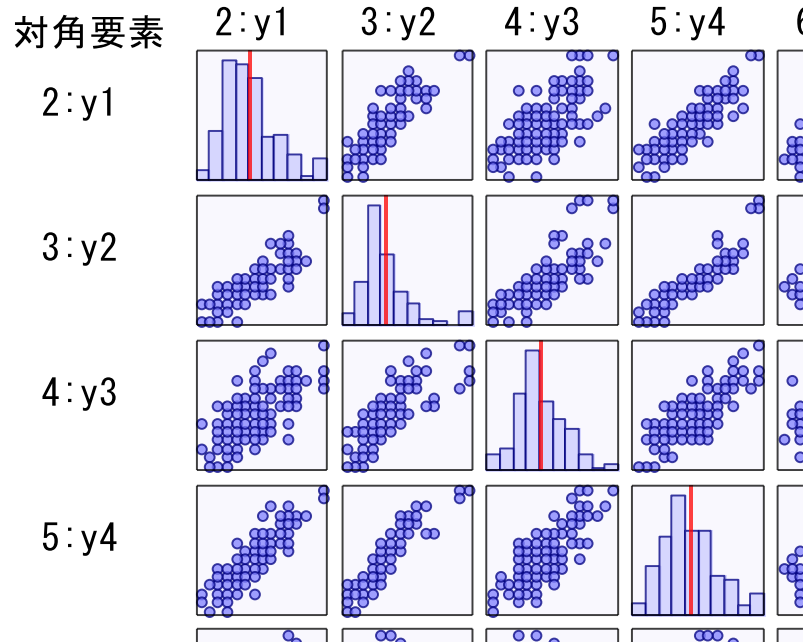
● SQC

● 機械学習

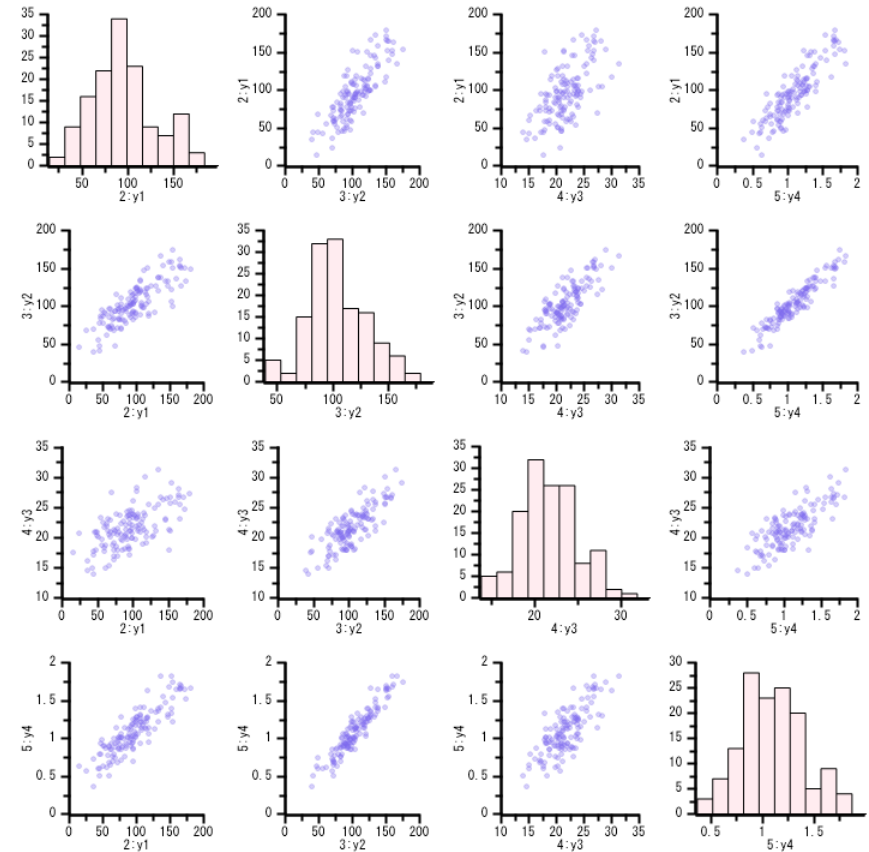
多変量連関図、モニタリング

濃淡散布図、密度プロット、等高線図

## ● 多変量連関図



## ● 濃淡散布図



# ① データ可視化

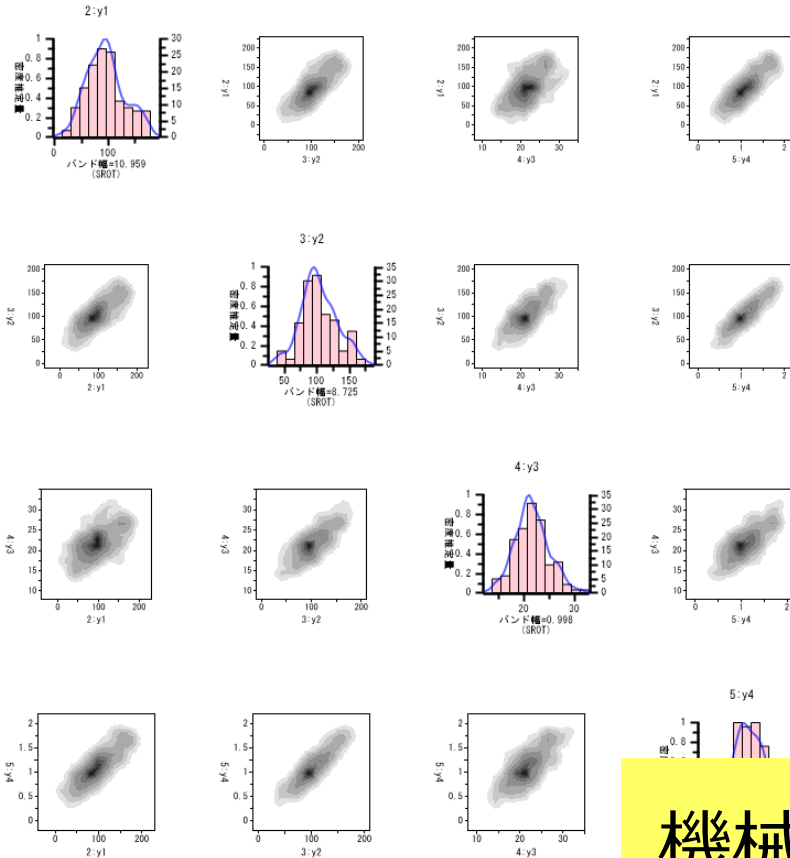
● SQC

● 機械学習

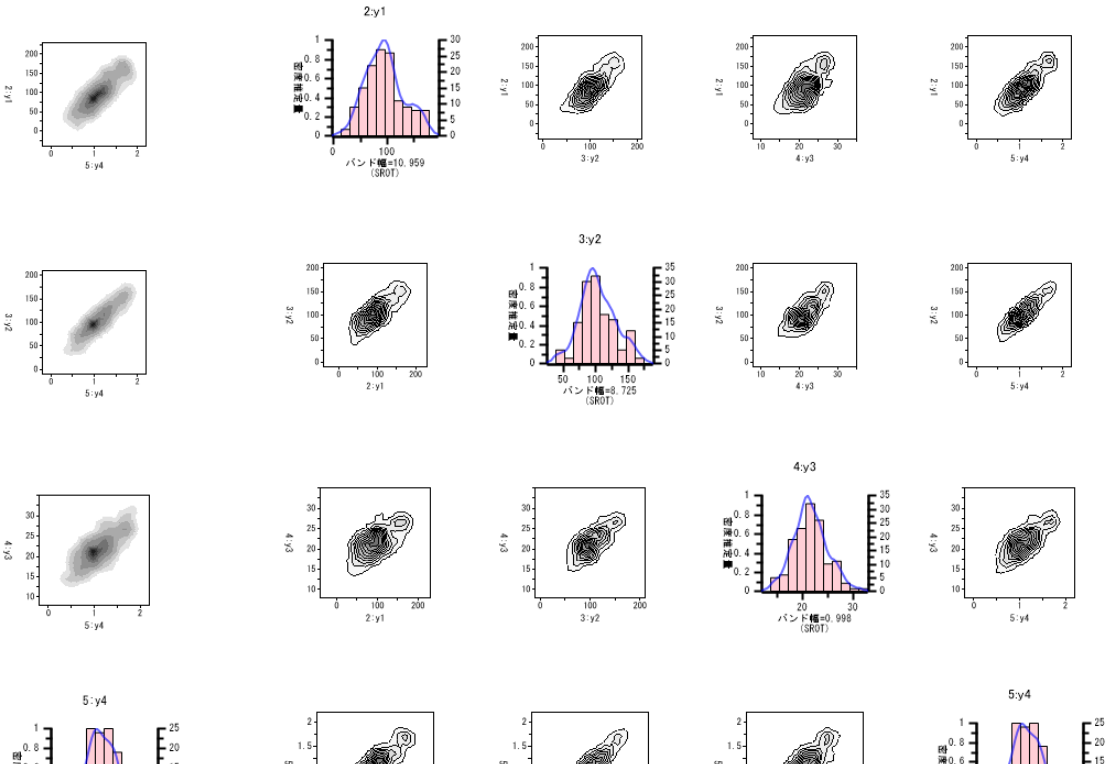
多変量連関図、モニタリング

濃淡散布図、密度プロット、等高線図

## ● 密度プロット



## ● 等高線図



機械学習編ではサンプルが多い場合のデータ可視化が充実

# ① データ可視化

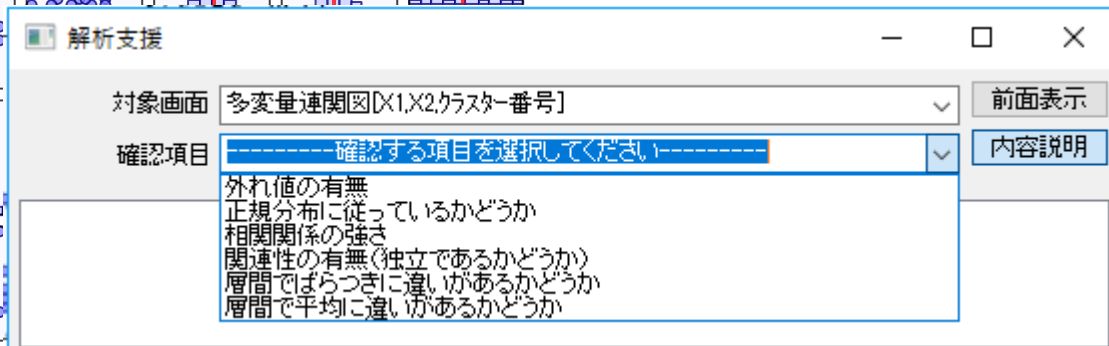
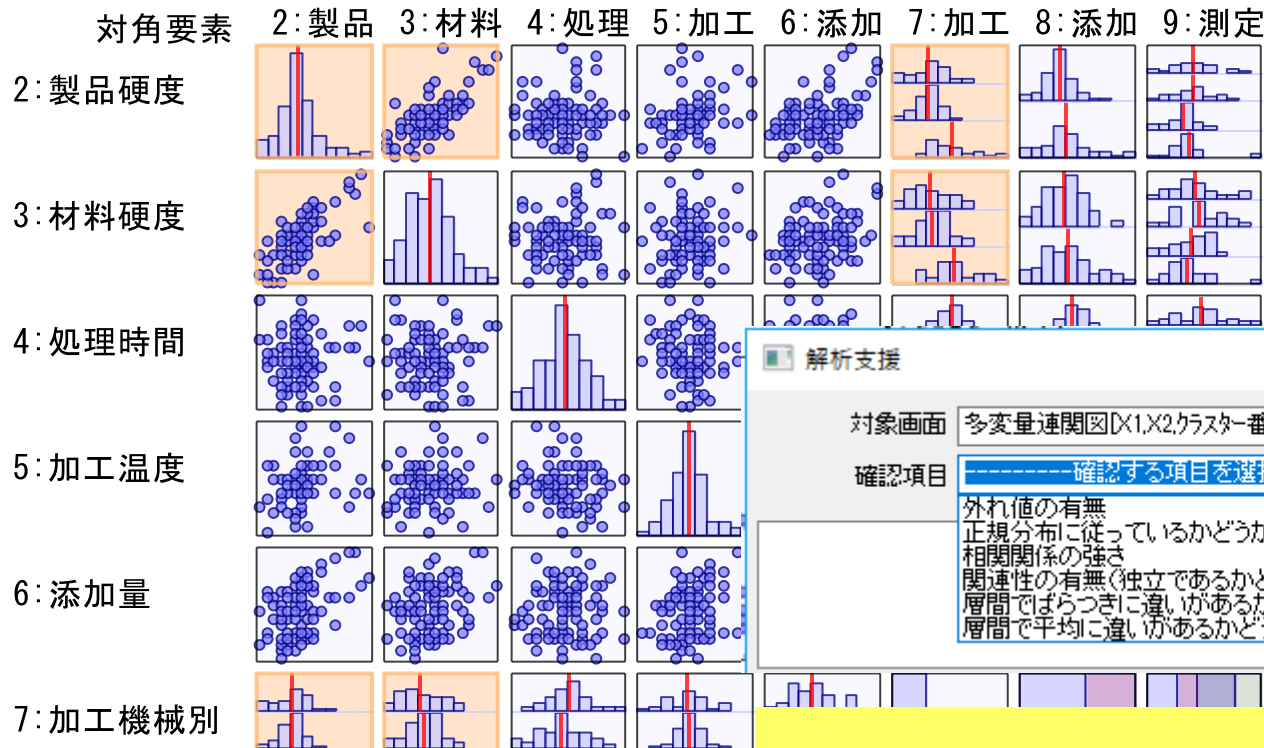
● SQC

● 機械学習

多変量連関図、モニタリング

濃淡散布図、密度プロット、等高線図

## ● 多変量連関図



多変量連関図は質的変数も扱え、  
各種検定を有している

# ① データ可視化

● SQC	● 機械学習
多変量連関図、モニタリング	濃淡散布図、密度プロット、等高線図

- SQCの多変量連関図はn数が多くなると、重なりを表現することが難しい。  
ただし、質的変数を扱え、検定も可能
- 機械学習の各手法は、データが多い場合の可視化機能が充実。  
一方、質的変数は扱えない

両者をうまく使い分け、データの傾向や特徴を掴むこと

# (参考) 基本動作の重要性

1. 重回帰分析とは

## 1-2. 重回帰分析のイメージ

### ◆ 重回帰分析から得られる結果

《重回帰式》

$$Y = a_1X_1 + a_2X_2 + \dots + a_iX_i + b + \varepsilon$$

Y: 目的変数 (結果)

$a_i$ : 偏回帰係数

$X_i$ : 説明変数 (原因)

b: 定数項

$\varepsilon$ : 誤差 (Yの変動のうち、 $X_i$ で説明できない部分)

この式を重回帰式 (予測式、モデル式) という。

例えば、

- 野球の勝率 =  $a_1$ (チーム打率) +  $a_2$ (防御率) +  $a_3$ (本塁打数)
- 車速(最高速) =  $a_1$ (エンジンの出力) +  $a_2$ (車重) +  $a_3$ (空気抵抗)
- 肺がん発生率 =  $a_1$ (タバコの本数/日) +  $a_2$ (環境) +  $a_3$ (年齢)

1. 重回帰分析とは

## 1-3. 重回帰分析の目的

### ◆ 重回帰分析の特徴

#### ① 変動要因解析 (要因の絞込み)

各要因 (説明変数) が、特性 Y (目的変数) に対してどれだけの影響があるかをみる

例) 引き続き実験、調査を行うための予備的解析として使用

#### ② 予測 (現象のモデル化)

各要因 (説明変数) の値を使い、特性 Y (目的変数) の値を予測する

例) 製品の強度 (Y) を、材料強度、形状、板厚など ( $X_i$ ) から予測

$$Y = a_1X_1 + a_2X_2 + a_3X_3 + \dots + b \text{ の予測式を作成}$$

# (参考) 回帰分析の要因解析

## 早稲田大学 永田靖先生ご講演より

- 第19回 JUSEシンポジウム (2009.11.30)

【特別講演】

**回帰分析の用途・実験計画法の意義  
グラフィカルモデリングの活用**

- 2019年度デミング賞本賞 受賞報告講演要旨

# 対象とする手法、データサイズ

## JUSE-StatWorks/V5

解析の目的		● SQC (総合編プレミアムなど2011年～)	● 機械学習 (機械学習編 2019年～)
①データ可視化		多変量連関図、モニタリング	濃淡散布図、密度プロット、等高線図
②層別		階層的クラスター分析 非階層的クラスター分析(k-means法)	混合ガウス分布
③情報の要約		主成分分析	カーネル主成分分析
④予測 (回帰)		重回帰分析	正則化回帰分析 (リッジ回帰、lasso回帰、Elastic Net)
⑤分類	⑤- 1	判別分析	サポートベクターマシン (SVM)
	⑤- 2	AID、CAID	ランダムフォレスト
⑥外れ値検出		多変量管理図、MT法	1クラスSVM
⑦因果分析		グラフィカルモデリング	glasso

## ②層別

● SQC

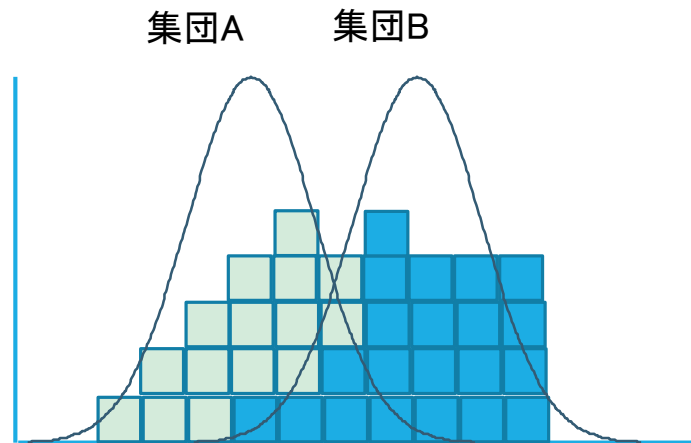
● 機械学習

階層的クラスター分析  
非階層的クラスター分析

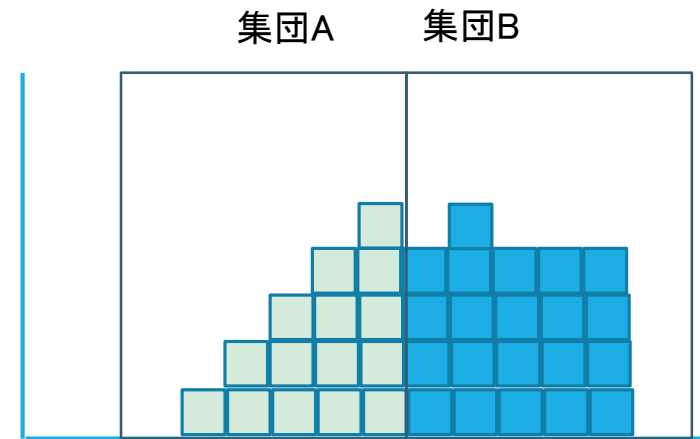
混合ガウス分布

### 1変数の例

● 混合ガウス分布



● クラスタ分析



原理原則・固有技術で考えて、  
同一座標で重なるか重ならないかで選択



# 対象とする手法、データサイズ

## JUSE-StatWorks/V5

解析の目的	● SQC (総合編プレミアムなど2011年～)	● 機械学習 (機械学習編 2019年～)
①データ可視化	多変量連関図、モニタリング	濃淡散布図、密度プロット、等高線図
②層別	階層的クラスター分析 非階層的クラスター分析(k-means法)	混合ガウス分布
③情報の要約	主成分分析	カーネル主成分分析
④予測 (回帰)	重回帰分析	正則化回帰分析 (リッジ回帰、lasso回帰、Elastic Net)
⑤分類	⑤- 1 判別分析	サポートベクターマシン (SVM)
	⑤- 2 AID、CAID	ランダムフォレスト
⑥外れ値検出	多変量管理図、MT法	1クラスSVM
⑦因果分析	グラフィカルモデリング	glasso

JUSE Package Software - [1:シート1 [ML\_M03\_03\_カーネル法による分類.sw5]]

ワークシート 手法選択 解析 装飾

開く 保存 印刷 コピー 貼付 変数一括編集 変数属性 編集 表示 変数情報  
 入力 出力 クリップボード

データ 基本解析 システム ウィンドウ ヘルプ  
 ソート

	S1	C2	N3	N4	N5	N6
	サンプル名	class	X1	X2	変数5	変数6
1	s1		2	2.508	1.479	
2	s2		2	-2.589	1.348	
3	s3		1	0.421	-0.850	
4	s4		1	-0.872	-0.627	
5	s5		2	-1.627	-2.790	
6	s6		1	-0.048	1.630	
7	s7		1	0.183	-1.818	
8	s8		2	1.829	2.375	
9	s9		2	0.259	2.959	
10	s10		1	1.192	0.740	
11	s11		2	2.173	2.161	
12	s12		2	2.867	0.383	
13	s13		2	-3.177	-1.131	
14	s14		1	-0.621	0.486	
15	s15		2	1.124	-2.843	
16	s16		2	2.557	1.209	
17	s17		2	-2.430	1.414	
18	s18		1	-0.935	0.155	
19	s19		2	-0.868	-2.857	
20	s20		2	2.047	-2.178	
21	s21		2	2.554	-0.007	
22	s22		2	-0.186	2.939	
23	s23		1	0.018	0.633	

新規ブック  
シート1

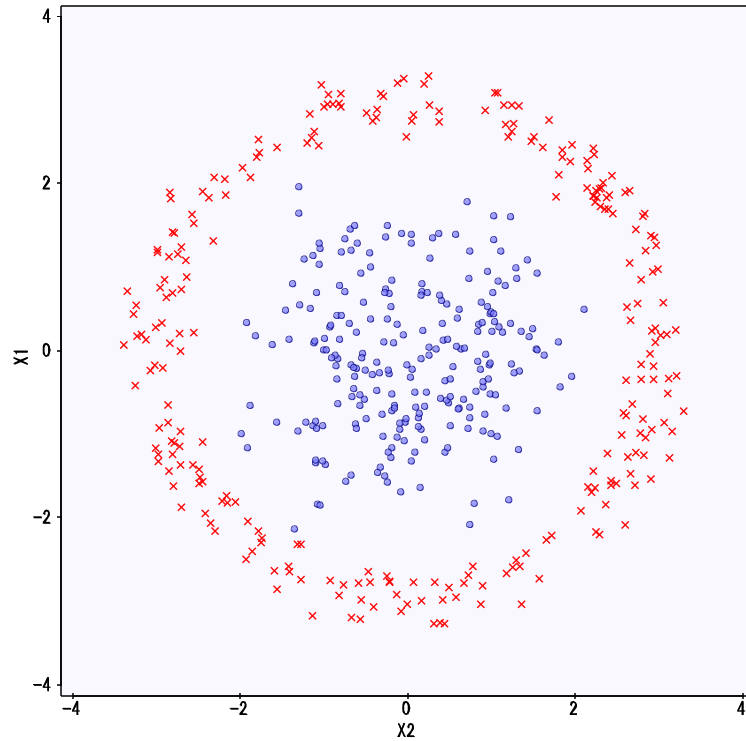
StatWorks サンプルデータ ML\_M03\_03.SW5

データ

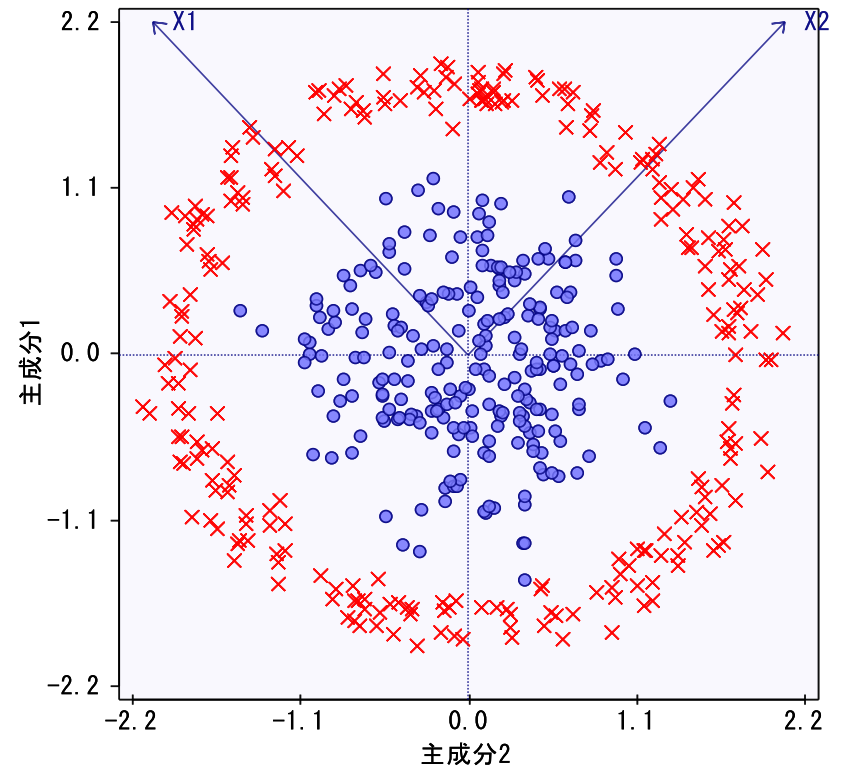
# ③情報の要約

● SQC	● 機械学習
主成分分析	カーネル主成分分析

## 多変量連関関 (散布図)



## ● 主成分分析



# ③情報の要約

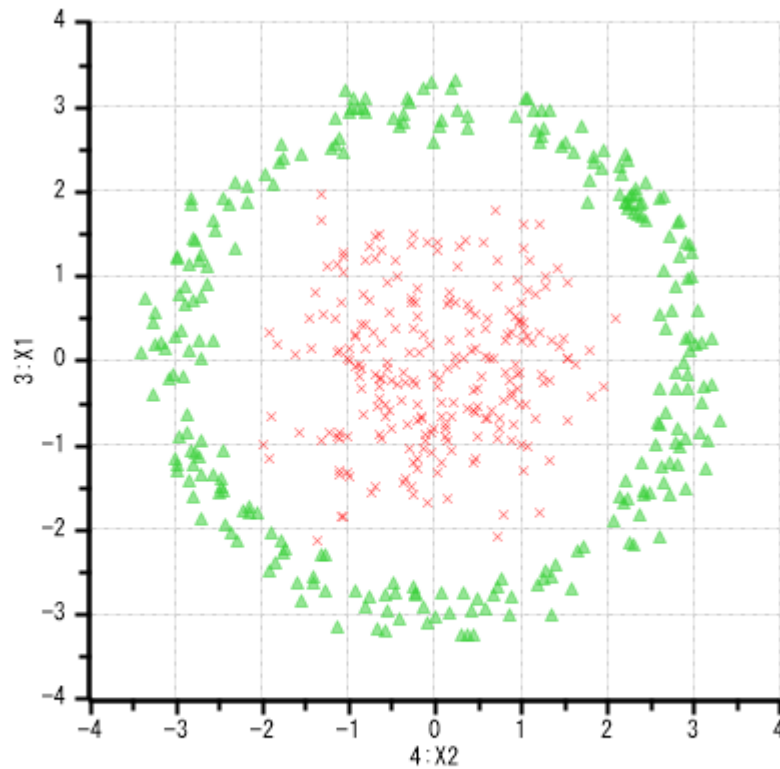
● SQC

主成分分析

● 機械学習

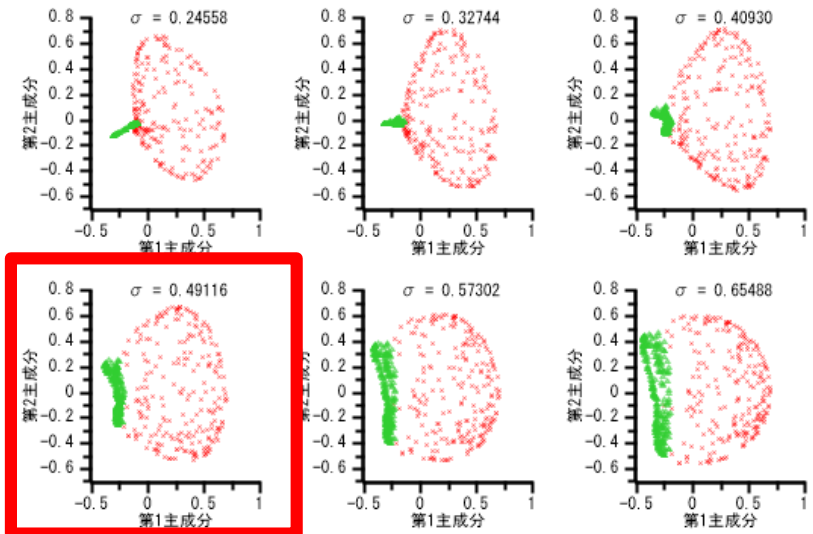
カーネル主成分分析

## ●カーネル主成分分析



層別:class

StatWorks サンプルデータ ML\_M03\_03.SW5



ワークシート 手法選択 **解析** 装飾

保存 印刷 コピー 出力

表示

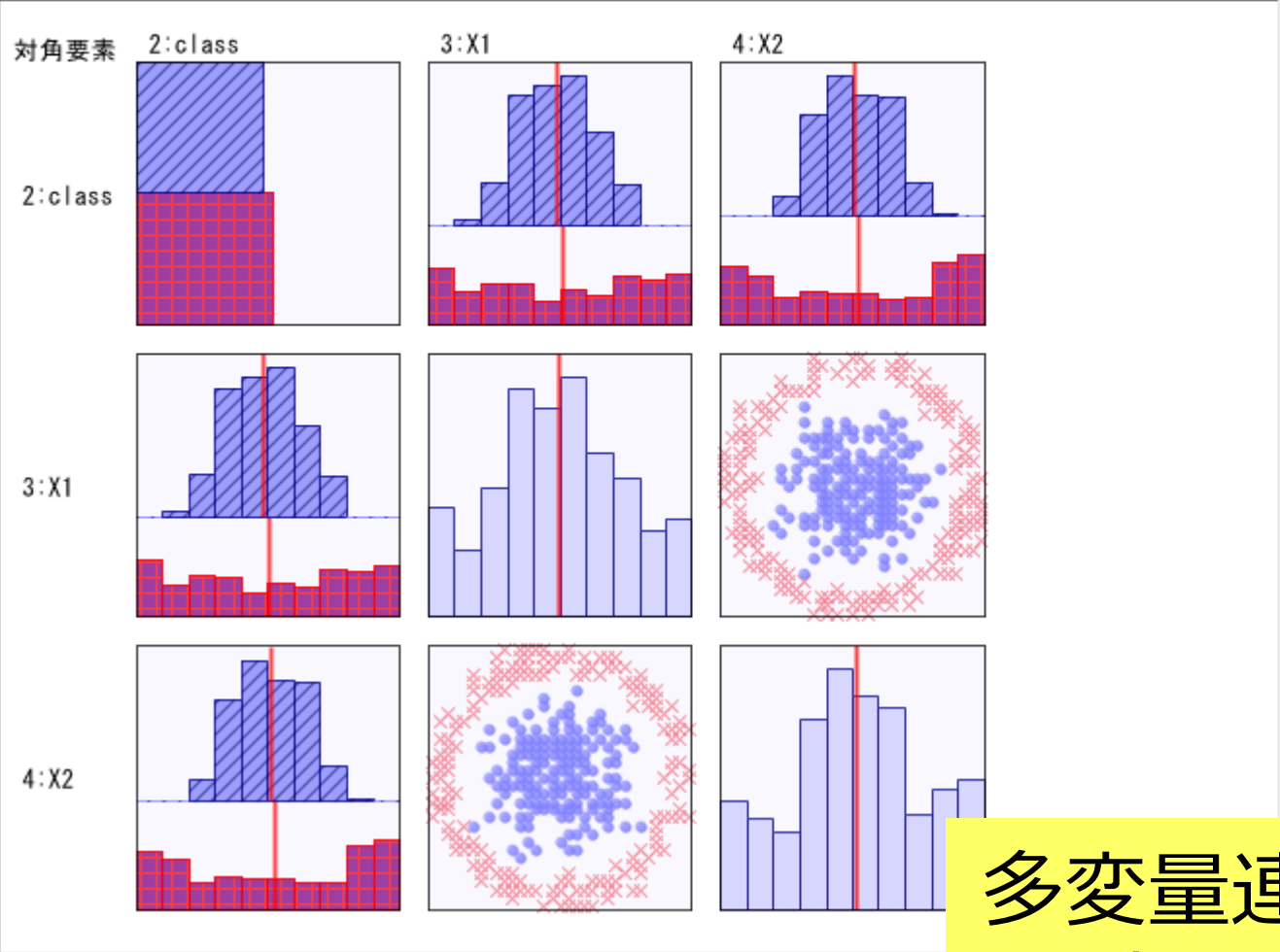
ソート

解析アドバイザー  
変数再指定  
解析支援

解析操作

ウインドウ

ヘルプ



新規ブック

シート1

多変量連関図[X1,X2,class]

多変量連関図を描けば  
層別はできる

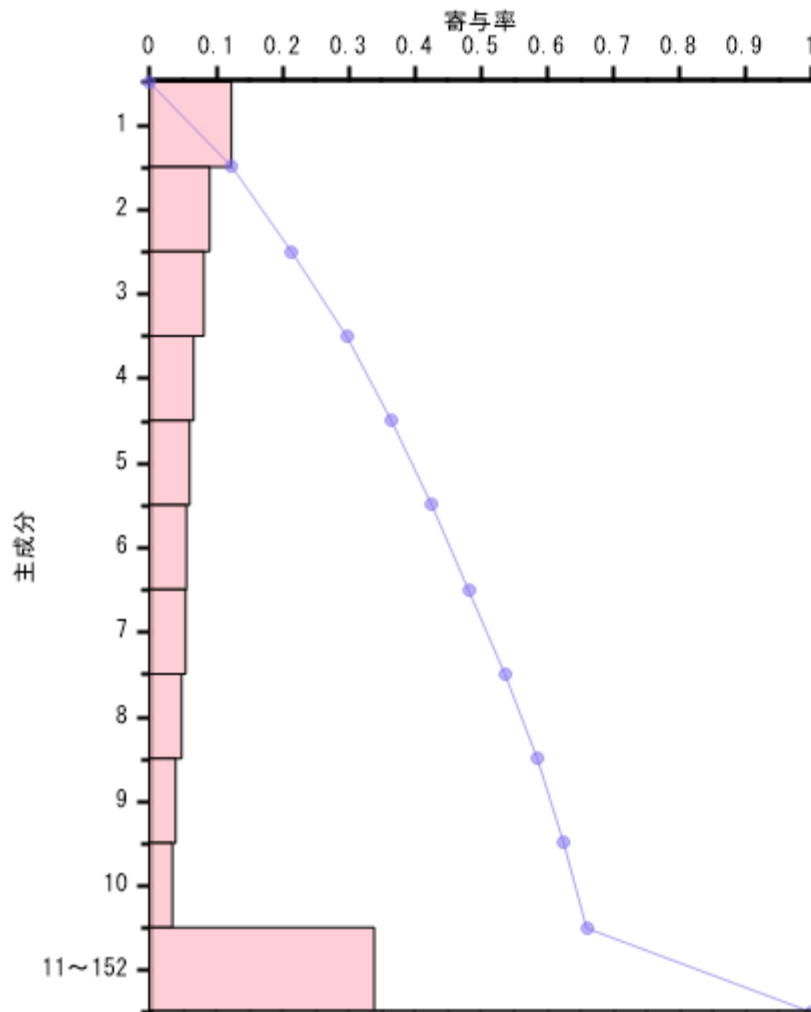
### ③情報の要約

● SQC	● 機械学習
主成分分析	カーネル主成分分析

- ドーナツ型の内側/外側の層別に対して、カーネル主成分分析は有効である。ただし、軸の解釈が難しい
- 一方、主成分分析はデータサイズによらず計算可能のため、まずは主成分分析で解析するのがよい。軸の解釈をしやすいメリットもある
- 主成分分析で解析後、それでは見つからなかった異常値をカーネル主成分分析で見つけることに使えそう

手順としては  
主成分分析→カーネル主成分分析

# (参考) カーネル主成分分析での固有値と累積寄与率



	固有値	寄与率	累積寄与率
1	55.551	0.124	0.124
2	40.712	0.091	0.215
3	36.901	0.083	0.298
4	30.037	0.067	0.365
5	27.190	0.061	0.426
6	25.081	0.056	0.482
7	24.549	0.055	0.537
8	21.792	0.049	0.586
9	17.762	0.040	0.625
10	15.579	0.035	0.660
11	15.150	0.034	0.694
12	13.894	0.031	0.725
13	13.195	0.030	0.755
14	11.875	0.027	0.781
15	11.040	0.025	0.806
16	9.604	0.021	0.827
17	7.078	0.016	0.843
18	6.606	0.015	0.858
19	6.500	0.015	0.872
20	5.570	0.012	0.885
21	5.046	0.011	0.896
22	4.670	0.010	0.907
23	4.163	0.009	0.916
24	4.004	0.009	0.925
25	2.778	0.006	0.931
26	2.576	0.006	0.937
27	2.524	0.006	0.943
28	2.229	0.005	0.948
29	2.067	0.005	0.952
30	1.888	0.004	0.956
31	1.794	0.004	0.960
32	1.465	0.003	0.964

# 対象とする手法、データサイズ

## JUSE-StatWorks/V5

解析の目的		● SQC (総合編プレミアムなど2011年～)	● 機械学習 (機械学習編 2019年～)
①データ可視化		多変量連関図、モニタリング	濃淡散布図、密度プロット、等高線図
②層別		階層的クラスター分析 非階層的クラスター分析(k-means法)	混合ガウス分布
③情報の要約		主成分分析	カーネル主成分分析
④予測 (回帰)		重回帰分析	正則化回帰分析 (リッジ回帰、lasso回帰、Elastic Net)
⑤分類	⑤- 1	判別分析	サポートベクターマシン (SVM)
	⑤- 2	AID、CAID	ランダムフォレスト
⑥外れ値検出		多変量管理図、MT法	1クラスSVM
⑦因果分析		グラフィカルモデリング	glasso



## ④予測（回帰）

● SQC

● 機械学習

重回帰分析

正則化回帰分析（リッジ回帰、  
lasso回帰、Elastic Net）

- まずは単回帰分析。ビックデータになって変数が膨大なら目的変数と相関が高い説明変数が見つかる可能性あり（相関係数が0.9以上表示できる機能があるとよい）

$$y = ax + b$$

- 単回帰分析で説明力が不十分ならば、重回帰分析

$$y = a_1x_1 + a_2x_2 + b$$

- $p > n$  ( $p$ :説明変数の数、 $n$ :サンプル数)の状況で正則化回帰分析が登場

## ④予測（回帰）

● SQC	● 機械学習
重回帰分析	正則化回帰分析（リッジ回帰、lasso回帰、Elastic Net）

- 正則化回帰分析は予期せぬ線形制約にも回帰係数を求めることができる
- ただし、以下の注意点が必要
  - ◆ 回帰係数が「縮小」される
  - ◆ 意のままに変数選択できない

手順としては  
単回帰分析→重回帰分析→正則化回帰分析

# 対象とする手法、データサイズ

## JUSE-StatWorks/V5

解析の目的	● SQC (総合編プレミアムなど2011年～)	● 機械学習 (機械学習編 2019年～)
①データ可視化	多変量連関図、モニタリング	濃淡散布図、密度プロット、等高線図
②層別	階層的クラスター分析 非階層的クラスター分析(k-means法)	混合ガウス分布
③情報の要約	主成分分析	カーネル主成分分析
④予測 (回帰)	重回帰分析	正則化回帰分析 (リッジ回帰、lasso回帰、Elastic Net)
⑤分類	⑤- 1 判別分析	サポートベクターマシン (SVM)
	⑤- 2 AID、CAID	ランダムフォレスト
⑥外れ値検出	多変量管理図、MT法	1クラスSVM
⑦因果分析	グラフィカルモデリング	glasso

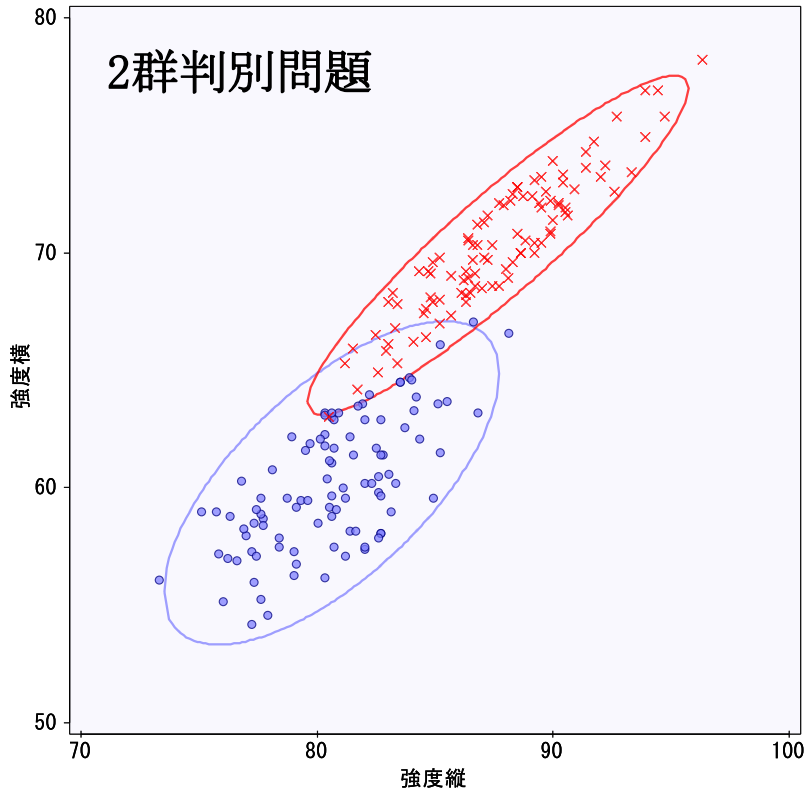
# ⑤-1 分類

● SQC

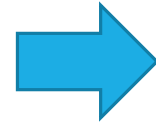
● 機械学習

判別分析

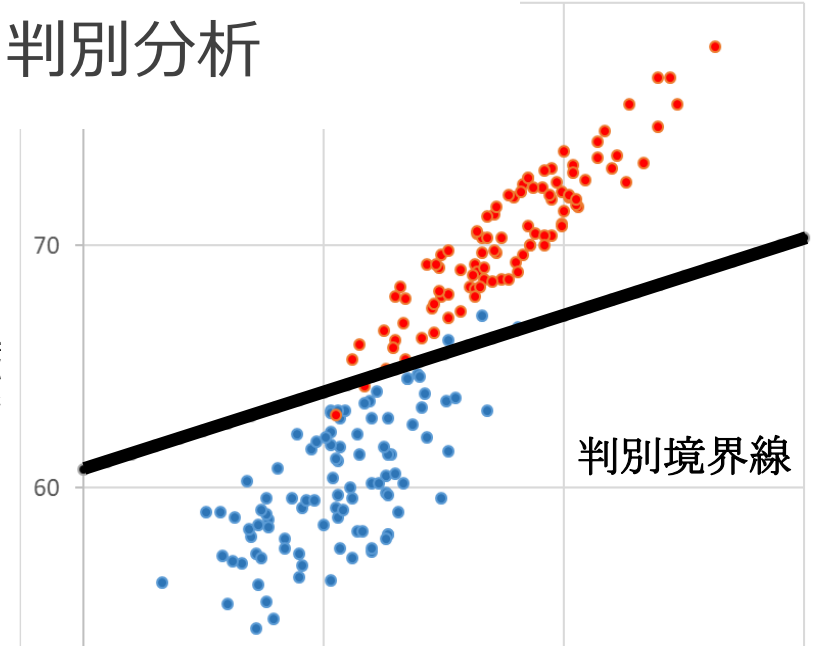
サポートベクターマシン (SVM)



判別分析



● 判別分析



目的変数		判別効率 $D^2$	$D^{*2}$	$D^{**2}$		群	重心 $\Delta$
機械	マハラノビス距離	13.697	13.519	13.347		1: A	6.848
	誤判別率(%)	3.212	3.300	3.387		2: B	-6.848
		判別効率 $D^2$	変化量	誤判別率	F値	P値(上側)	判別係数
vNo.	定数項						66.983
IN 3	強度縦	12.7	-1.0	3.7	11.6	0.0	0.6
IN 4	強度横	5.3	-8.4	12.5	178.4	0.0	-1.7

	定数項	強度縦	強度横
判別関数	F1= +66.983	+0.6 X3	-1.7 X4

# ⑤-1 分類

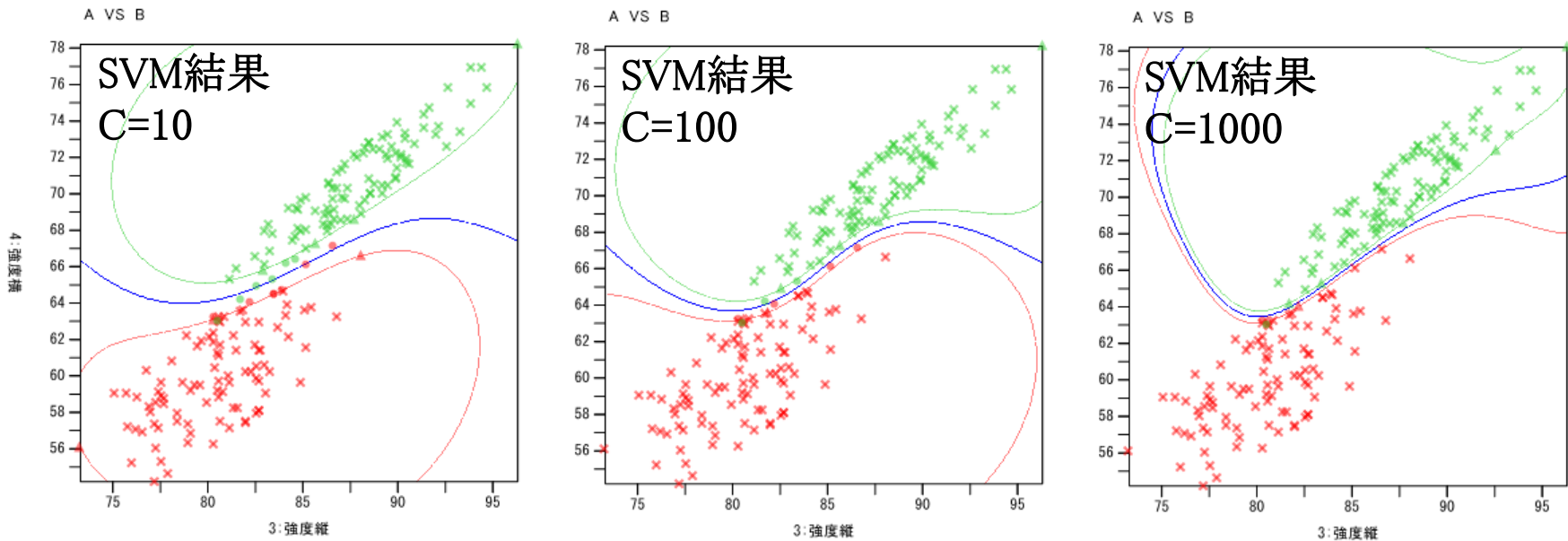
● SQC

判別分析

● 機械学習

サポートベクターマシン (SVM)

## ● SVM



群が正規分布にみなせれば、判別分析を推奨。  
 それで不十分ならSVM。  
 ただし、判別力が強力だが判別境界の技術的な解釈は難しい

# 対象とする手法、データサイズ

## JUSE-StatWorks/V5

解析の目的		SQC (総合編プレミアムなど2011年～)	機械学習 (機械学習編 2019年～)
①データ可視化		多変量連関図、モニタリング	濃淡散布図、密度プロット、等高線図
②層別		階層的クラスター分析 非階層的クラスター分析(k-means法)	混合ガウス分布
③情報の要約		主成分分析	カーネル主成分分析
④予測 (回帰)		重回帰分析	正則化回帰分析 (リッジ回帰、lasso回帰、Elastic Net)
⑤分類	⑤- 1	判別分析	サポートベクターマシン (SVM)
	⑤- 2	AID、CAID	ランダムフォレスト
⑥外れ値検出		多変量管理図、MT法	1クラスSVM
⑦因果分析		グラフィカルモデリング	glasso

JUSE Package Software - [1:勤続年数等によるローン審査(CAID) [G2\_0204\_勤続年数等によるローン審査(CAID).SW5]]

ワークシート 手法選択 解析 装飾

開く 保存 印刷 出力  
コピー 貼付 切取 クリップボード  
変数一括編集 変数属性 サンプル名 変数情報  
編集 表示 ソート  
データ 基本解析 システム ウィンドウ ヘルプ

	● S1	● C2	● C3	● C4	● C5	● C6	● N7
	サンプル名	職業	勤続年数	年収	信用情報	ローン審査	変数7
● 1	1	会社員	3年以上	400万円以上	優	安全	
● 2	2	公務員	1年以上	400万円以上	優	危険	
● 3	3	自営業	1年未満	400万円未満	良	危険	
● 4	4	会社員	1年未満	400万円未満	可	危険	
● 5	5	公務員	1年未満	400万円以上	優	安全	
● 6	6	会社員	3年以上	400万円以上	優	危険	
● 7	7	その他	3年以上	400万円未満	良	安全	
● 8	8	その他	3年以上	400万円以上	優	安全	
● 9	9	自営業	1年以上	400万円以上	優	安全	
● 10	10	会社員	3年以上	400万円未満	優	安全	
● 11	11	会社員	3年以上	400万円以上	優	安全	
● 12	12	会社員	3年以上	400万円未満	優	安全	
● 13	13	会社員	3年以上	400万円未満	可	安全	
● 14	14	会社員	1年以上	400万円未満	優	安全	
● 15	15	その他	3年以上	400万円以上	良	危険	
● 16	16	自営業	1年未満	400万円以上	良	安全	
● 17	17	会社員	3年以上	400万円未満	良	危険	
● 18	18	自営業	3年以上	400万円未満	優	安全	
● 19	19	会社員	3年以上	400万円以上	優	安全	
● 20	20	会社員	1年以上	400万円以上	優	危険	
● 21	21	会社員	3年以上	400万円以上	可	安全	

勤続年数等によるローン審査(CAID)  
勤続年数等によるローン審査(CAID)

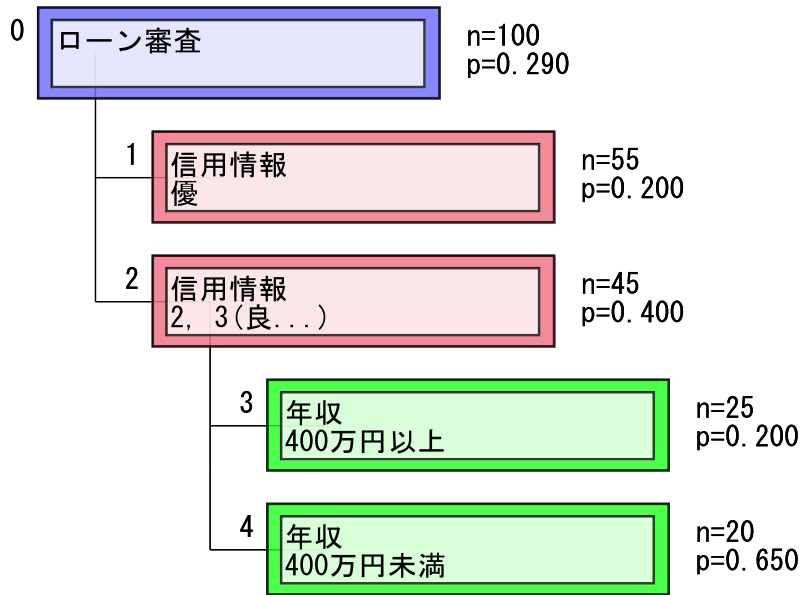
StatWorks サンプルデータ G2\_0204.SW5

データ

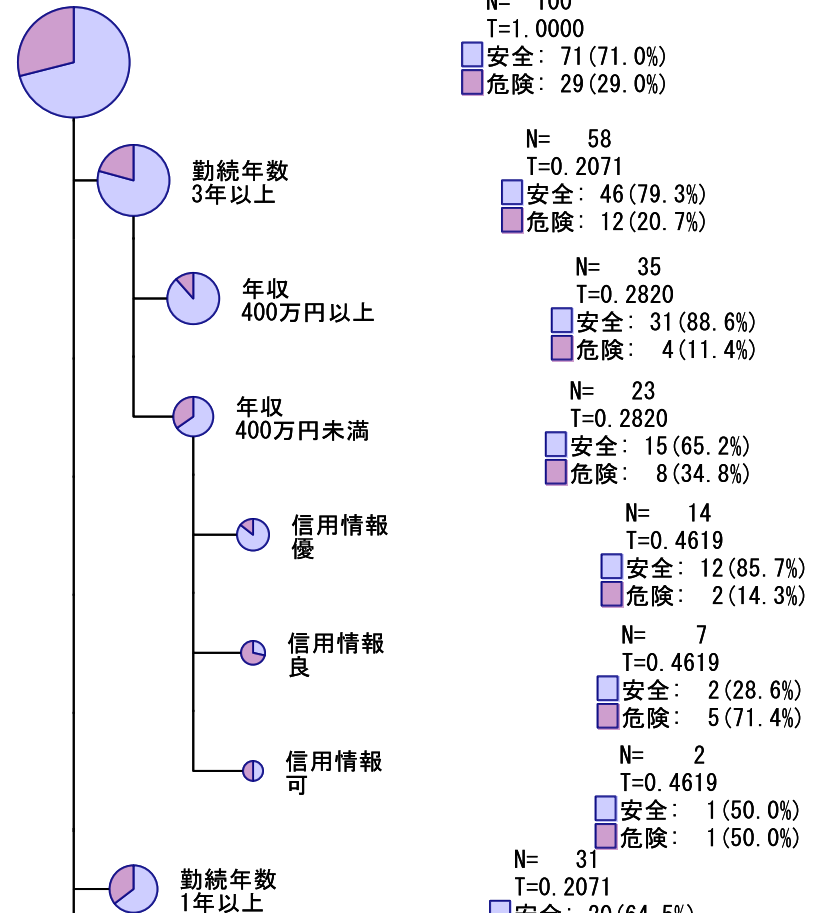
# ⑤-2 分類

● SQC	● 機械学習
AID、CAID	ランダムフォレスト

## ● AID



## ● CAID



StatWorks サンプルデータ G2\_0204.SW5

AID、CAIDは分岐条件がわかりやすい

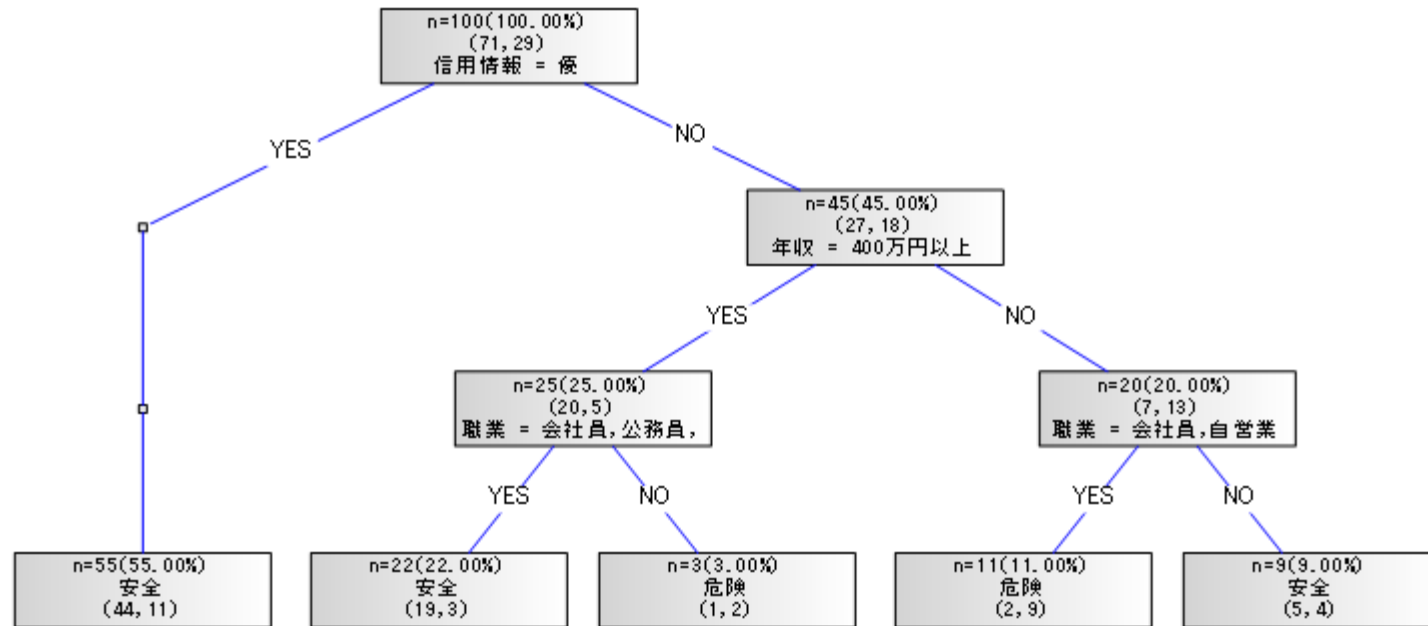
安全: 5 (45.5%)



## ⑤-2 分類

● SQC	● 機械学習
AID、CAID	ランダムフォレスト

### ● 決定木 (ランダムフォレスト内)



ランダムフォレストはAID,CAIDを包含している。  
ただし、3分岐以上が必要ならCAIDが候補

# 対象とする手法、データサイズ

## JUSE-StatWorks/V5

解析の目的		● SQC (総合編プレミアムなど2011年～)	● 機械学習 (機械学習編 2019年～)
①データ可視化		多変量連関図、モニタリング	濃淡散布図、密度プロット、等高線図
②層別		階層的クラスター分析 非階層的クラスター分析(k-means法)	混合ガウス分布
③情報の要約		主成分分析	カーネル主成分分析
④予測 (回帰)		重回帰分析	正則化回帰分析 (リッジ回帰、lasso回帰、Elastic Net)
⑤分類	⑤- 1	判別分析	サポートベクターマシン (SVM)
	⑤- 2	AID、CAID	ランダムフォレスト
⑥外れ値検出		多変量管理図、MT法	1クラスSVM
⑦因果分析		グラフィカルモデリング	glasso

# ⑥ 外れ値検出

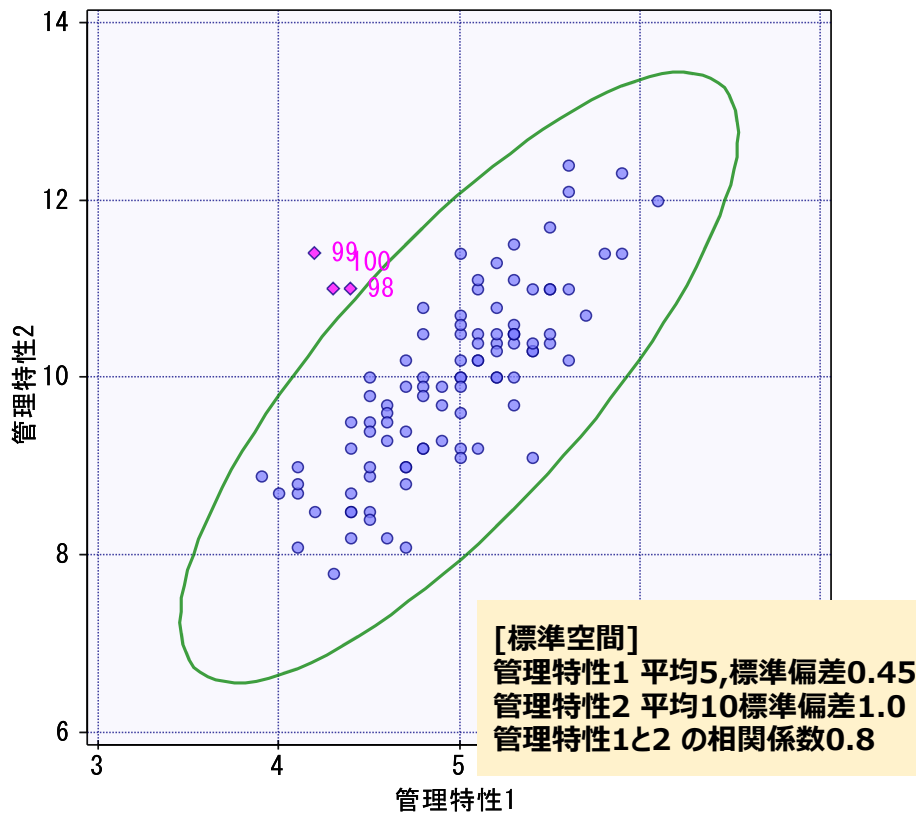
● SQC

● 機械学習

多変量管理図、MT法

1クラスSVM

## ● 多変量管理図



### オプション指定

管理限界値 補助線・装飾 境界値 規格値

#### 管理限界

$\chi^2$ 分布の上側0.27%点 (1変数の $3\sigma$ に相当)

指定されたパーセント点に対する $\chi^2$ 分布の上側パーセント点

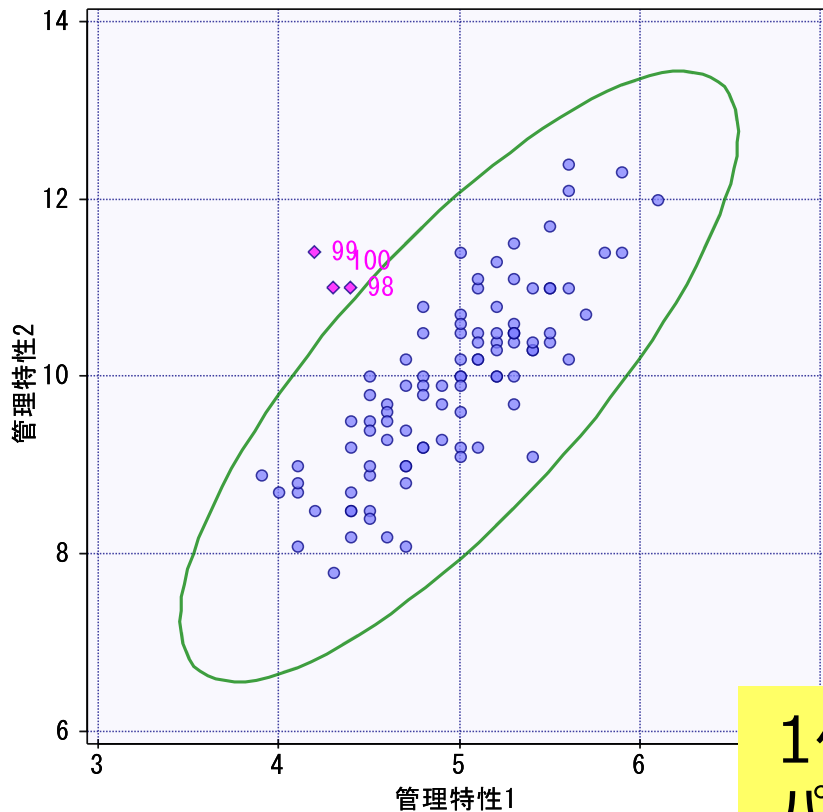
パーセント (0~100) :

89	3/17AM	1.235	0.667	0.000
90	3/17PM	2.299	-1.333	-1.500
91	3/18AM	4.250	-2.000	-1.900
92	3/18PM	1.049	-0.667	-1.000
93	3/19AM	0.051	0.222	0.200
94	3/19PM	3.040	0.667	1.500
95	3/20AM	2.250	0.000	-0.900
96	3/20PM	4.744	1.333	2.100
97	3/21AM	4.444	-2.000	-1.200
98	3/21PM	13.642	-1.333	1.000
99	3/22AM	25.285	-1.778	1.400
100	3/22PM	16.413	-1.556	1.000

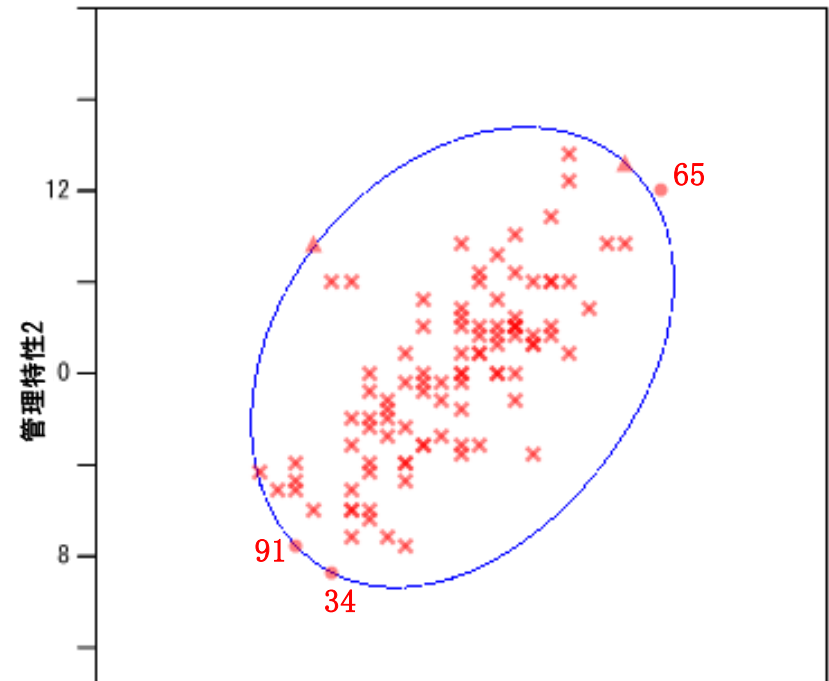
## ⑥ 外れ値検出

● SQC	● 機械学習
多変量管理図、MT法	1クラスSVM

### ● 多変量管理図



### ● 1クラスSVM



1クラスSVMでは異なる点が発見される。  
パラメータによっても検出点が変わるため、  
技術的な考察が必要

# 対象とする手法、データサイズ

## JUSE-StatWorks/V5

解析の目的		● SQC (総合編プレミアムなど2011年～)	● 機械学習 (機械学習編 2019年～)
①データ可視化		多変量連関図、モニタリング	濃淡散布図、密度プロット、等高線図
②層別		階層的クラスター分析 非階層的クラスター分析(k-means法)	混合ガウス分布
③情報の要約		主成分分析	カーネル主成分分析
④予測 (回帰)		重回帰分析	正則化回帰分析 (リッジ回帰、lasso回帰、Elastic Net)
⑤分類	⑤- 1	判別分析	サポートベクターマシン (SVM)
	⑤- 2	AID、CAID	ランダムフォレスト
⑥外れ値検出		多変量管理図、MT法	1クラスSVM
⑦因果分析		グラフィカルモデリング	glasso

JUSE Package Software - [1:中古住宅価格の予測 [ML\_M03\_05\_中古住宅価格の予測 (正規化回帰・glasso) .sw5]]

ワークシート

手法選択

解析

装飾

開く 保存 印刷 入力 出力

コピー 貼付 切取 クリップボード

変数一括編集 変数属性 変数情報

編集

表示

ソート

範囲 欠測 検索

解析対象データ 解析対象

データ

多変量 統計量 連関図 グラフ

基本解析

プロパティ 接続 システム

閉 全開 並び ウインドウ

	S1	N2	N3	N4	N5	N6	N7	N8	N9	N
	No.	中古住宅価格	土地面積 (m)	建物面積 (m)	築後年数	各停電車時間	バス時間	徒歩時間	新宿通勤時間	変数
1	1	17800	138.6	300.0	8	5	0	9	14	
2	2	3980	60.4	70.1	31	4	0	4	8	
3	3	3980	60.4	70.1	31	4	0	4	8	
4	4	5180	98.5	139.1	14	5	0	14	19	
5	5	6480	69.6	113.0	2	5	0	10	15	
6	6	8980	99.9	146.6	16	5	0	10	15	
7	7	17800	138.6	300.1	8	5	0	9	14	
8	8	4080	53.6	83.8	15	7	0	15	22	
9	9	4080	53.6	83.8	15	7	0	15	22	
10	10	4080	53.6	83.8	15	7	0	15	22	
11	11	4080	53.6	83.8	15	7	0	15	22	
12	12	4080	53.6	83.8	15	7	0	15	22	
13	13	4080	53.6	83.8	15	7	0	15	22	
14	14	5880	78.6	163.6	23	7	0	13	20	
15	15	6170	63.5	88.4	18	7	0	4	11	
16	16	6580	89.1	103.1	76	7	0	12	19	
17	17	6980	137.2	225.3	22	7	0	17	24	
18	18	7600	62.6	121.7	7	7	0	4	11	
19	19	8800	70.3	97.5	4	7	0	7	14	
20	20	4530	48.8	77.4	15	8	0	8	16	
21	21	5400	100.0	100.0	34	8	0	17	25	
22	22	3400	101.3	74.4	54	11	0	17	28	

中古住宅  
中古1

StatWorks サンプルデータ ML\_M03\_05.SW5

# ⑦因果分析

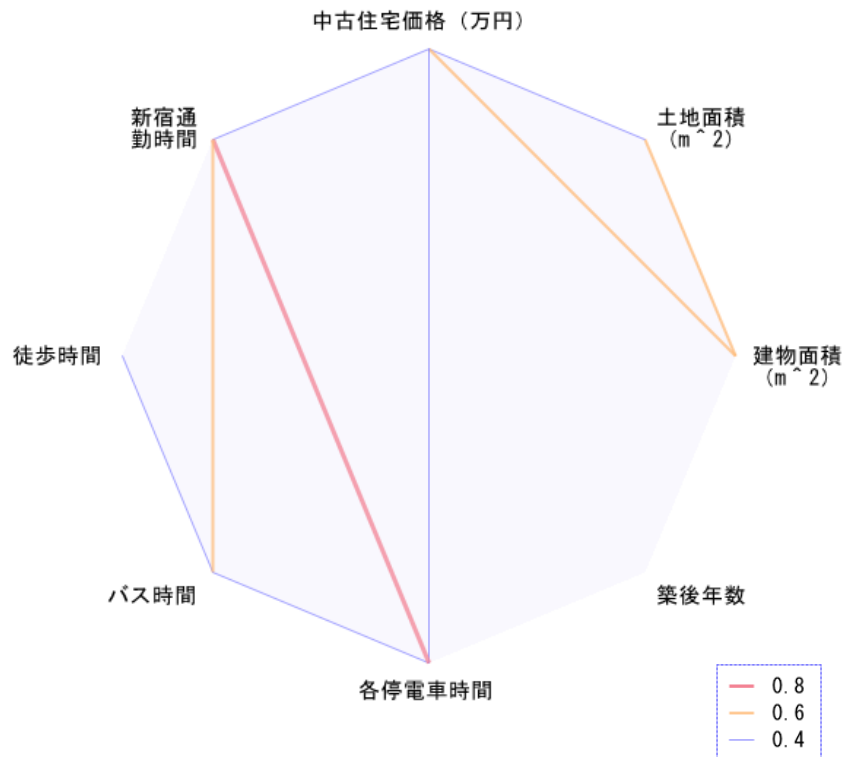
● SQC

● 機械学習

グラフィカルモデリング

glasso

## 相関係数の可視化グラフ



## ● グラフィカルモデリング

	中古住宅	土地面積	建物面積	築後年数	各停電車	バス時間	徒歩時間	新宿通勤
1	17800.0	138.580	300.020	8.0000	5.0000	0.0000	9.0000	14.0000
2	3980.00	60.3500	70.0900	31.0000	4.0000	0.0000	4.0000	8.0000
3	3980.00	60.3500	70.0900	31.0000	4.0000	0.0000	4.0000	8.0000
4	5180.00	98.4800	139.100	14.0000	5.0000	0.0000	14.0000	19.0000
5	6480.00	69.5700	113.020	2.0000	5.0000	0.0000	10.0000	15.0000
6	8980.00	99.9000	146.550	16.0000	5.0000	0.0000	10.0000	15.0000
7	17800.0	138.580	300.060	8.0000	5.0000	0.0000	9.0000	14.0000
8	4080.00	53.6000	83.8300	15.0000	7.0000	0.0000	15.0000	22.0000
9	4080.00	53.6000	83.8300	15.0000	7.0000	0.0000	15.0000	22.0000
10	4080.00	53.6000	83.8300	15.0000	7.0000	0.0000	15.0000	22.0000
11	4080.00	53.6000	83.8300	15.0000	7.0000	0.0000	15.0000	22.0000
12	4080.00	53.6000	83.8300	15.0000	7.0000	0.0000	15.0000	22.0000
13	4080.00	53.6000	83.8300	15.0000	7.0000	0.0000	15.0000	22.0000
14	5880.00	78.6000	102.770	34.0000	8.0000	0.0000	17.0000	25.0000
15	6170.00	63.5000	97.5500	4.0000	7.0000	0.0000	7.0000	14.0000
16	6580.00	89.0000	102.770	34.0000	8.0000	0.0000	17.0000	25.0000
17	6980.00	137.0000	102.770	34.0000	8.0000	0.0000	17.0000	25.0000
18	7600.00	62.6000	97.5200	46.0000	11.0000	0.0000	21.0000	32.0000
19	8800.00	70.3300	97.2000	11.0000	11.0000	0.0000	15.0000	26.0000
20	4530.00	48.8400	77.3800	15.0000	8.0000	0.0000	8.0000	16.0000
21	5490.00	120.890	102.770	34.0000	8.0000	0.0000	17.0000	25.0000
22	5490.00	120.890	102.770	34.0000	8.0000	0.0000	17.0000	25.0000
23	3400.00	101.280	74.3800	54.0000	11.0000	0.0000	17.0000	28.0000
24	4200.00	92.5600	97.5200	46.0000	11.0000	0.0000	21.0000	32.0000
25	4280.00	77.6500	74.0700	25.0000	11.0000	0.0000	15.0000	26.0000
26	4800.00	66.6300	97.2000	11.0000	11.0000	0.0000	15.0000	26.0000

Juse SEM 因果分析編

偏相関係数行列を求めることができません。

OK

StatWorks サンプルデータ ML\_M03\_05.SW5

この場合、線形制約があり  
偏相関係数が求まらない

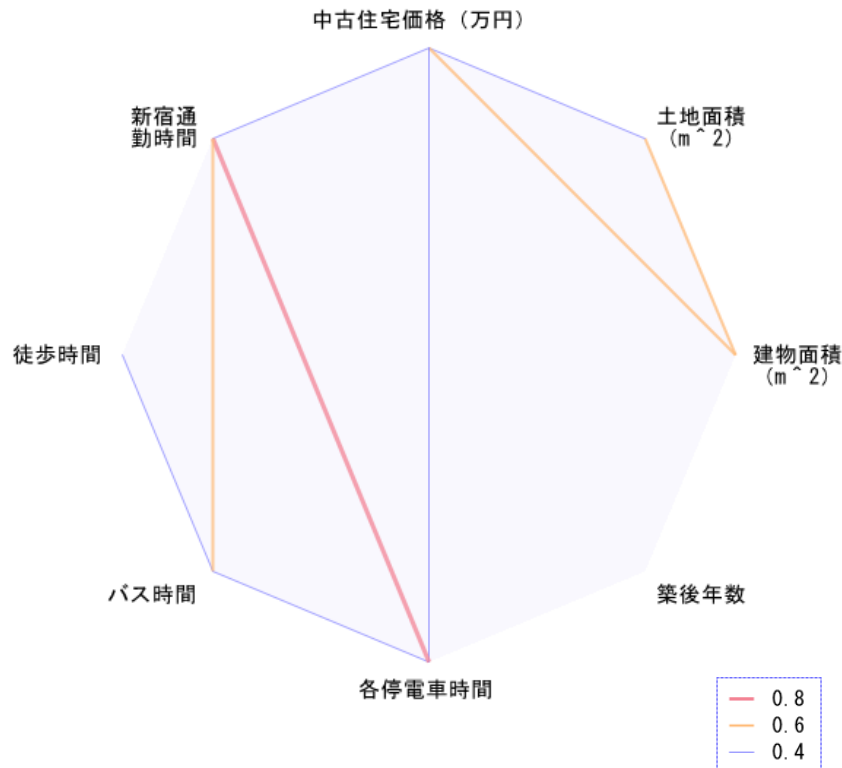
# ⑦因果分析

● SQC

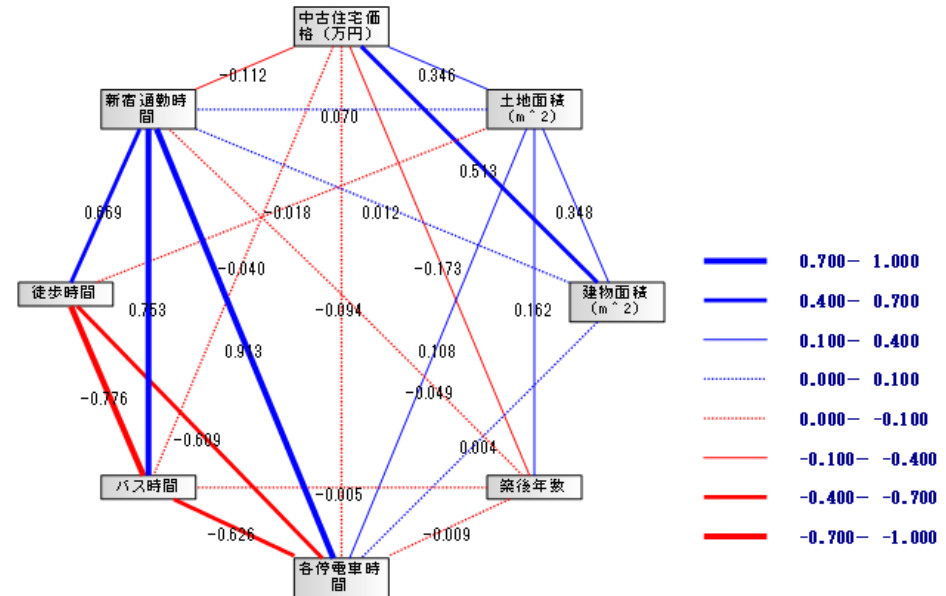
● 機械学習

グラフィカルモデリング

glasso



## ● glasso



glassoは解くことができる  
(ただし、係数の縮小あり)



# ⑦因果分析

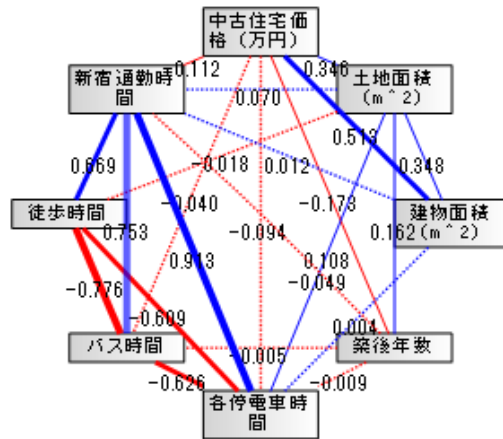
● SQC

● 機械学習

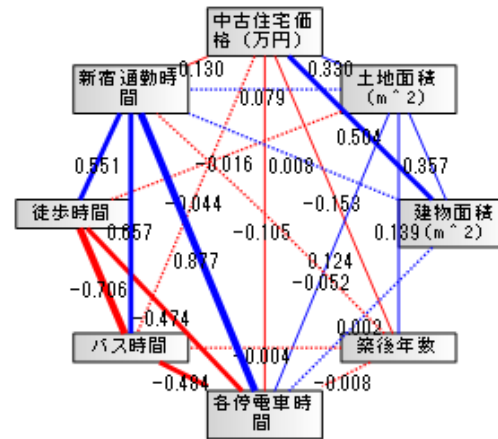
グラフィカルモデリング

glasso

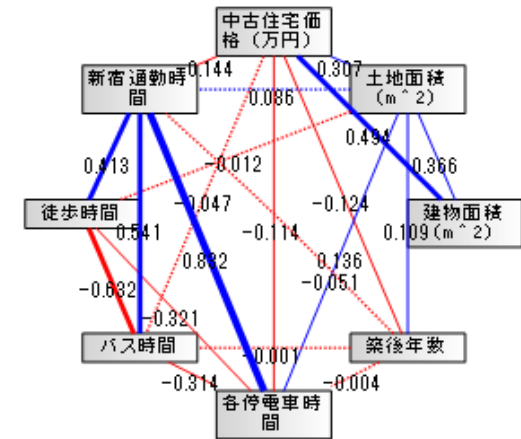
$\rho : 0.01000$  線の数: 22 BIC: 1215.217



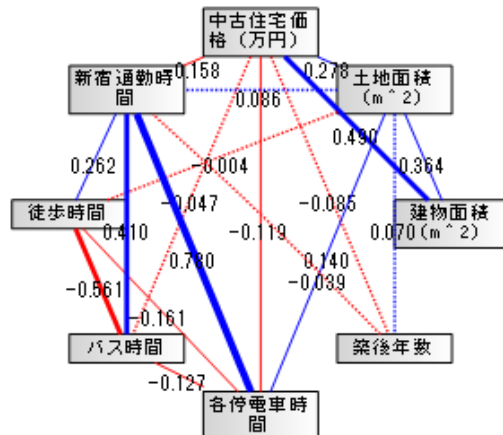
$\rho : 0.01550$  線の数: 22 BIC: 1507.318



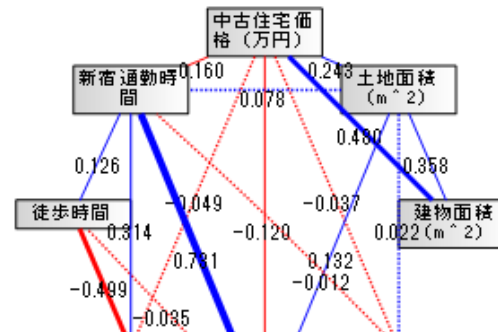
$\rho : 0.02402$  線の数: 20 BIC: 1798.962



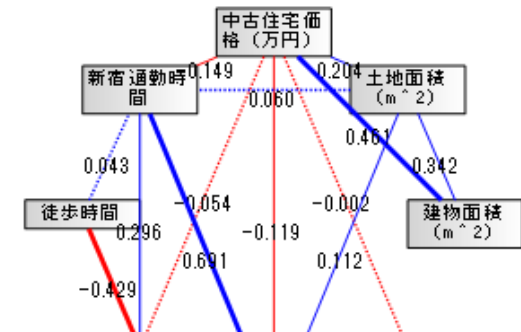
$\rho : 0.03723$  線の数: 18 BIC: 2109.357



$\rho : 0.05771$  線の数: 16 BIC: 2419.642



$\rho : 0.08944$  線の数: 13 BIC: 2729.029



正則化パラメータを変化し、線の数、  
偏相関係数（縮小）を確認できる

# ⑦因果分析

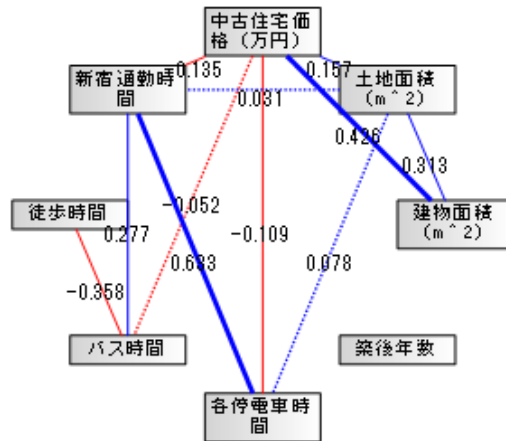
● SQC

● 機械学習

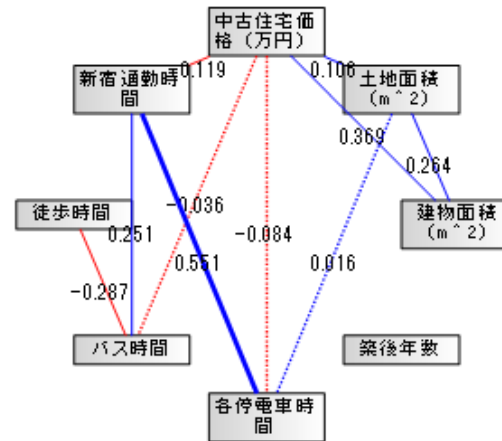
グラフィカルモデリング

glasso

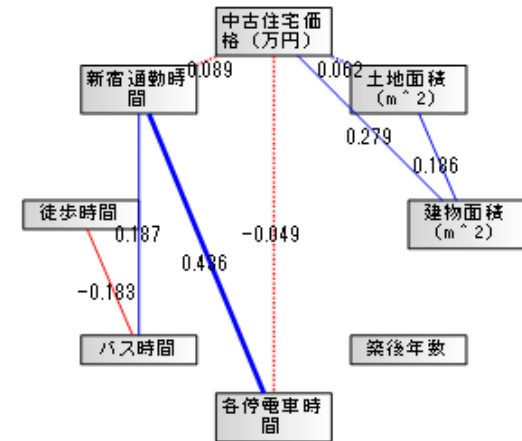
$\rho : 0.13863$  線の数: 11 BIC: 3137.210



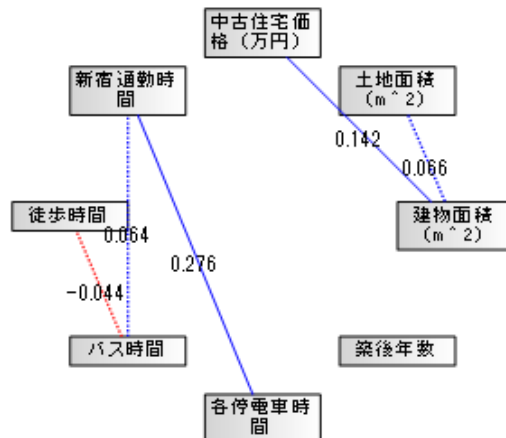
$\rho : 0.21486$  線の数: 10 BIC: 3673.389



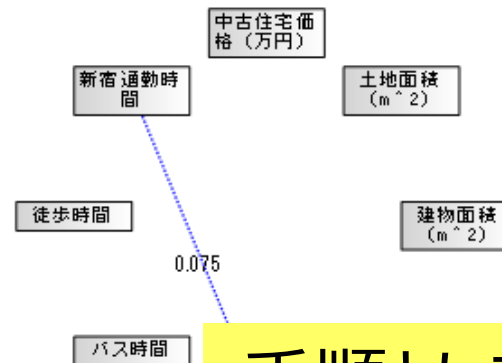
$\rho : 0.33302$  線の数: 8 BIC: 4325.300



$\rho : 0.51616$  線の数: 5 BIC: 5124.499



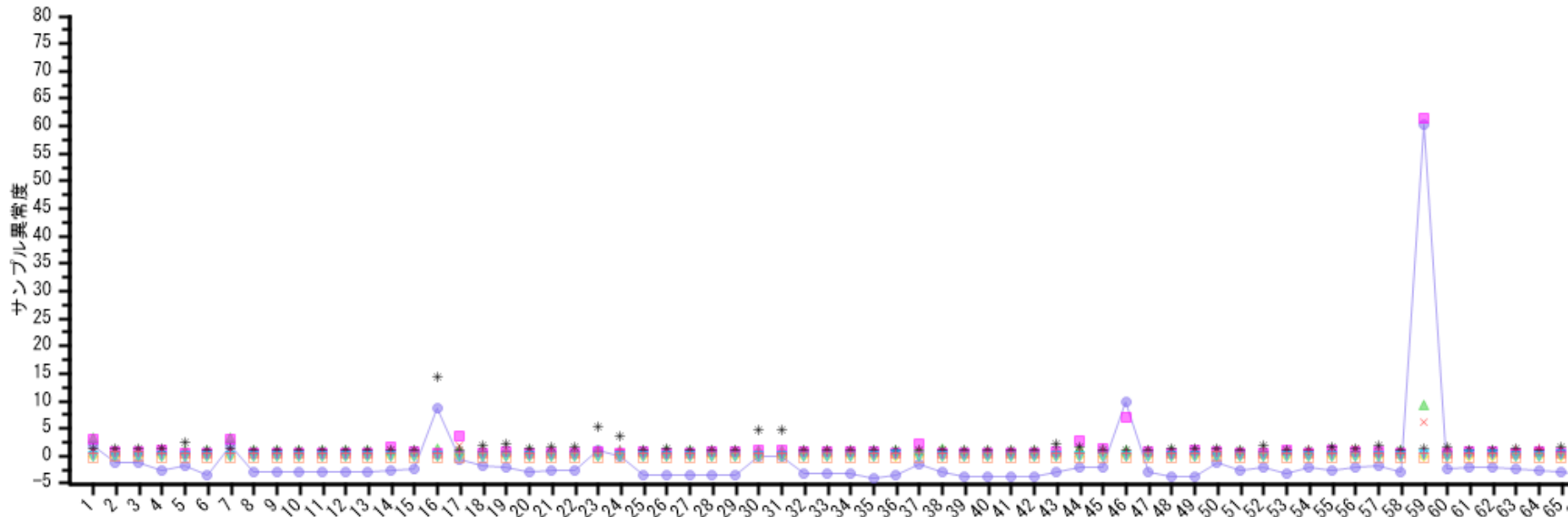
$\rho : 0.80000$  線の数: 1 BIC: 5793.259



手順としては  
グラフィカルモデリング → glasso

# (参考) その他のglassoのうれしさ

## サンプル異常度の表示

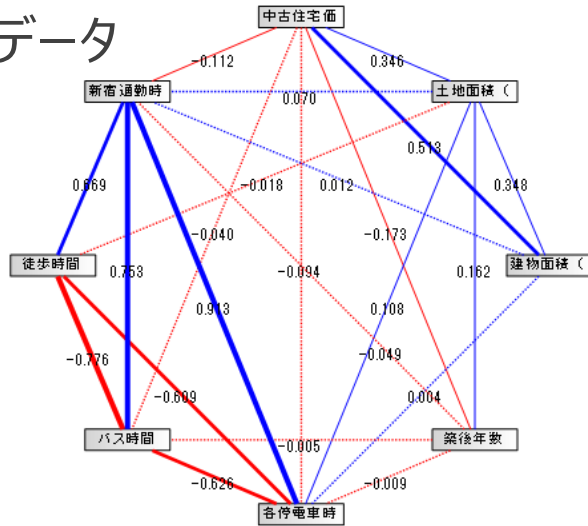


マハラビス距離が求まらない状況でも、  
正則化にてサンプルの異常度を求められる

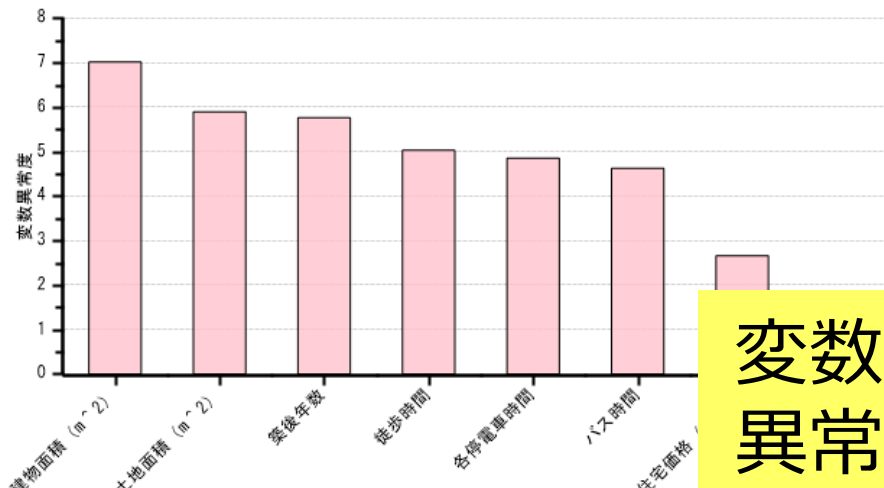
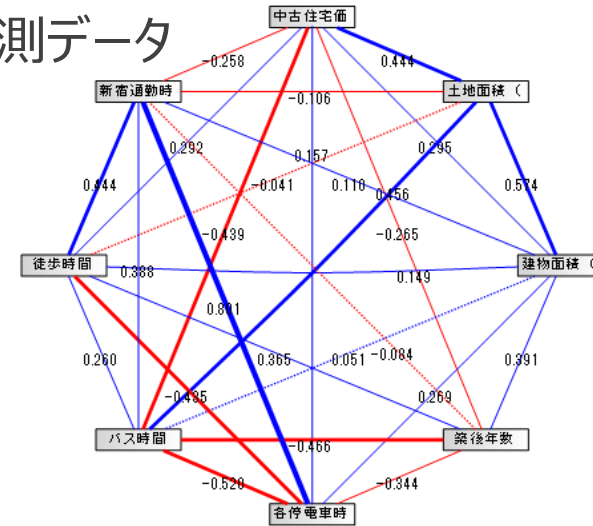
# (参考) その他のglassoのうれしさ

## 変数異常度の表示

分析データ



予測データ



変数の関係性に対する  
異常も検出できる

# まとめ

- ものづくりにおける「SQC」と「機械学習」のより良い使い方について提案した（報文の表2）
- 「機械学習はどんな場合でも万能であり、SQCは不要」ということは決してない。
- 問題解決のツールとして両者をうまく使い分け、品質・技術力の向上を目指すことが重要である。

---

**ご清聴ありがとうございました**

---

本著作物は原著作者の許可を得て、株式会社日本科学技術研修所（以下弊社）が掲載しています。本著作物の著作権については、制作した原著作者に帰属します。

原著作者および弊社の許可なく営利・非営利・イントラネットを問わず、本著作物の複製・転用・販売等を禁止します。

所属および役職等は、公開当時のものです。

■公開資料ページ

弊社ウェブページで各種資料をご覧ください <http://www.i-juse.co.jp/statistics/jirei/>

■お問い合わせ先

(株)日科技研 数理事業部 パッケージサポート係 <http://www.i-juse.co.jp/statistics/support/contact.html>