

情報処理応用B 第12回

藤田 一寿

■ QC 7 つ道具

- QCとはQuality Controlの略で品質管理のこと.
- QC 7 つ道具とは品質管理で有用な 7 つの手法のこと.
- QC手法は工業製品の品質管理の手法であったが, 現在では仕事の問題解決に用いられている.

■ QC7つ道具

- グラフ
- パレート図
- 管理図
- 散布図
- 特性要因図
- チェックシート
- 層別
- ヒストグラム

グラフ

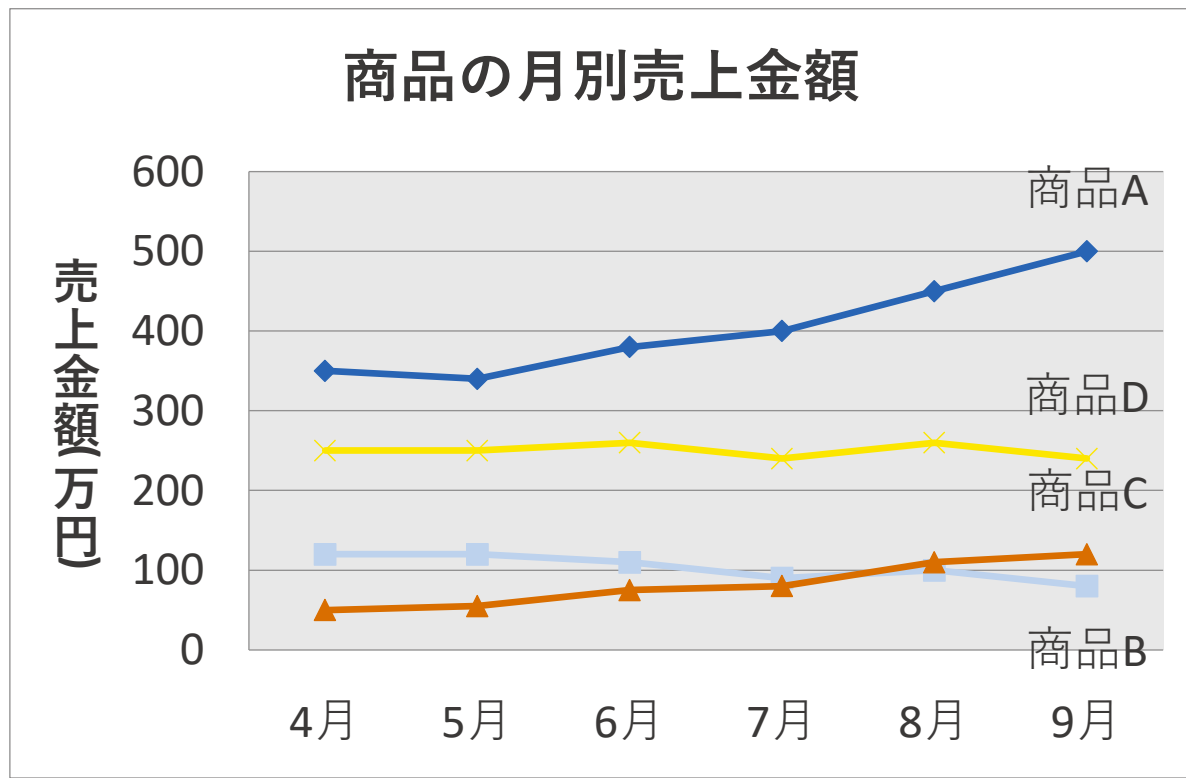
■ グラフ

- データの可視化
- グラフは目的に応じて使い分ける
- 折れ線グラフ、棒グラフ、円グラフ、帯グラフ、など

商品	4月	5月	6月	7月	8月	9月
商品A	350	340	380	400	450	500
商品B	120	120	110	90	100	80
商品C	50	55	75	80	110	120
商品D	250	250	260	240	260	240

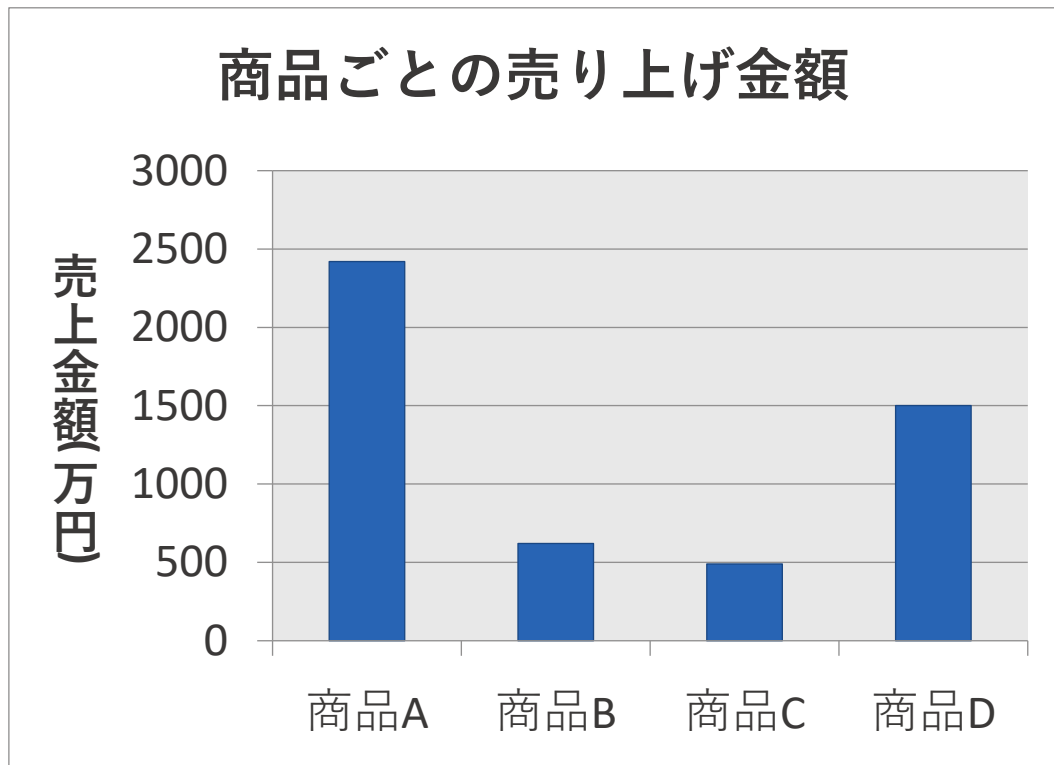
■ 折れ線グラフ

- 数量の時間的変化（もしくはある数値に対する依存性）を見るために用いる。



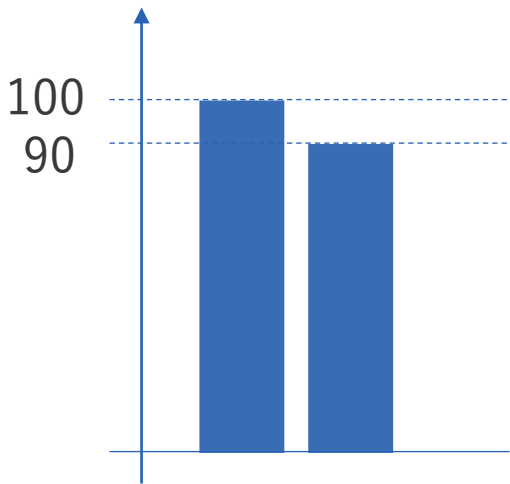
■ 棒グラフ

- 数量の大小関係を見るために用いる。

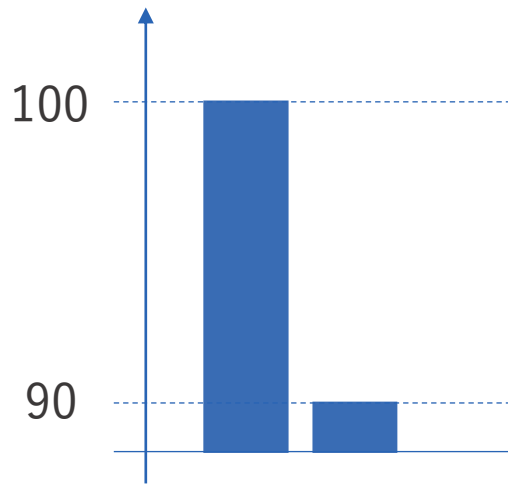


■ 折れ線グラフと棒グラフを見る上での注意

- 軸の範囲を変えて印象操作することが出来る.



元のグラフ

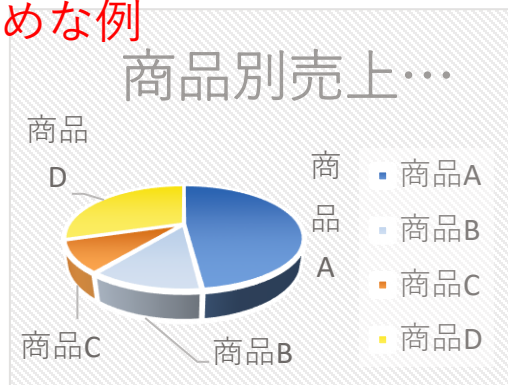


軸の範囲を変えたグラフ
元のグラフと同じ値なのに
差が広がって見える.

■ 円グラフ

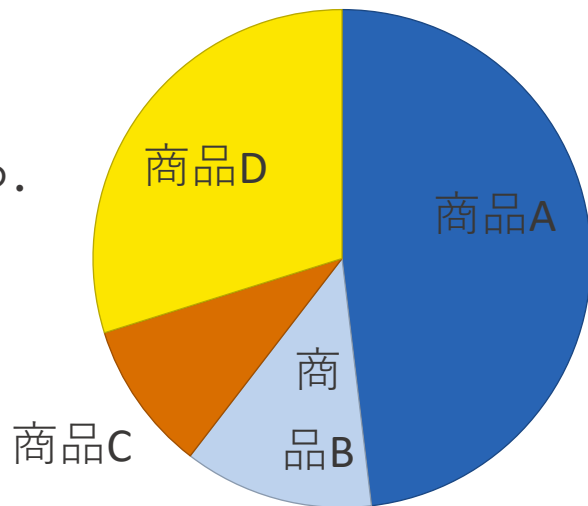
- 割合(占有率)を見る.
- 1周100%になるようにする.
- 割合に応じた扇の角度にする.
 - 10%しか無いのに、60度の扇にすると捏造になる.
- 3Dにしない.
 - 斜めから見ると割合が分からなくなる.

だめな例



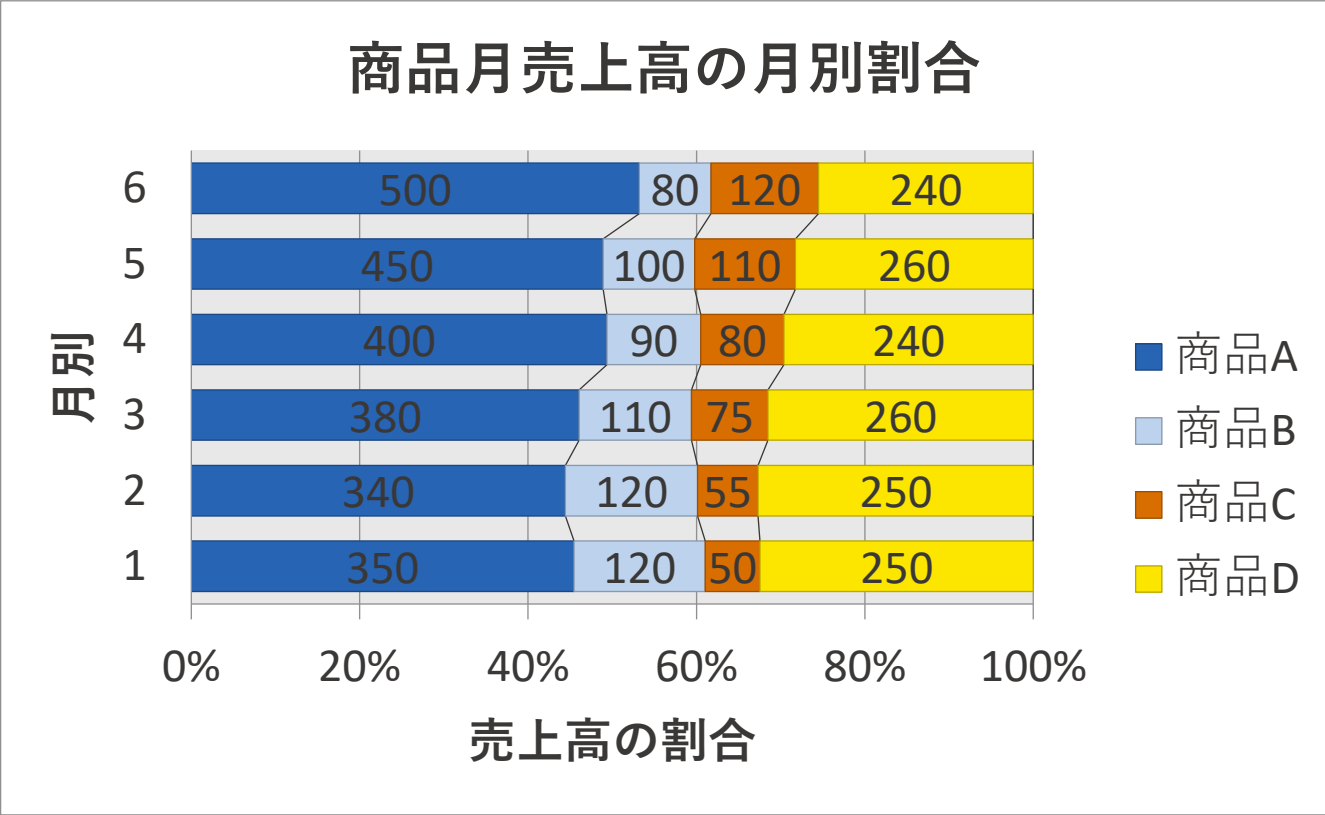
実際より商品Bの割合が大きく見える気がする.

商品別売上金額の割合



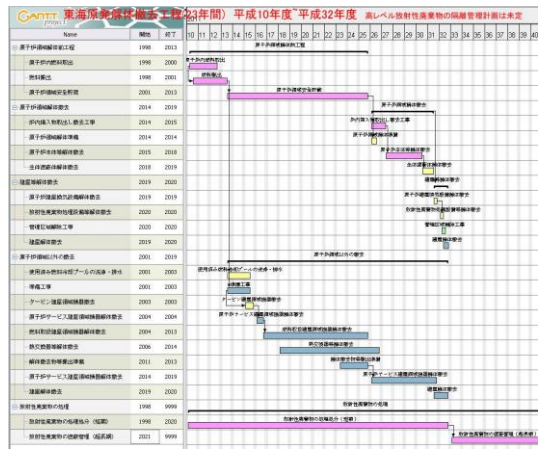
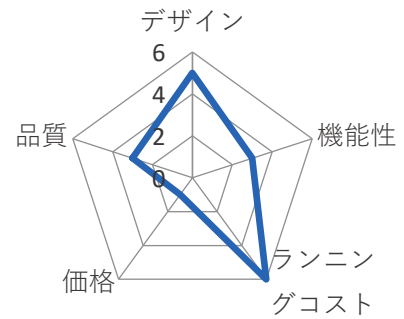
■ 帯グラフ

- 総量、割合の変化をみる場合に使う。



■ その他

- レーダーチャート
 - 項目ごとの評価の把握
- ガントチャート
 - 作業の流れと進捗状況の確認



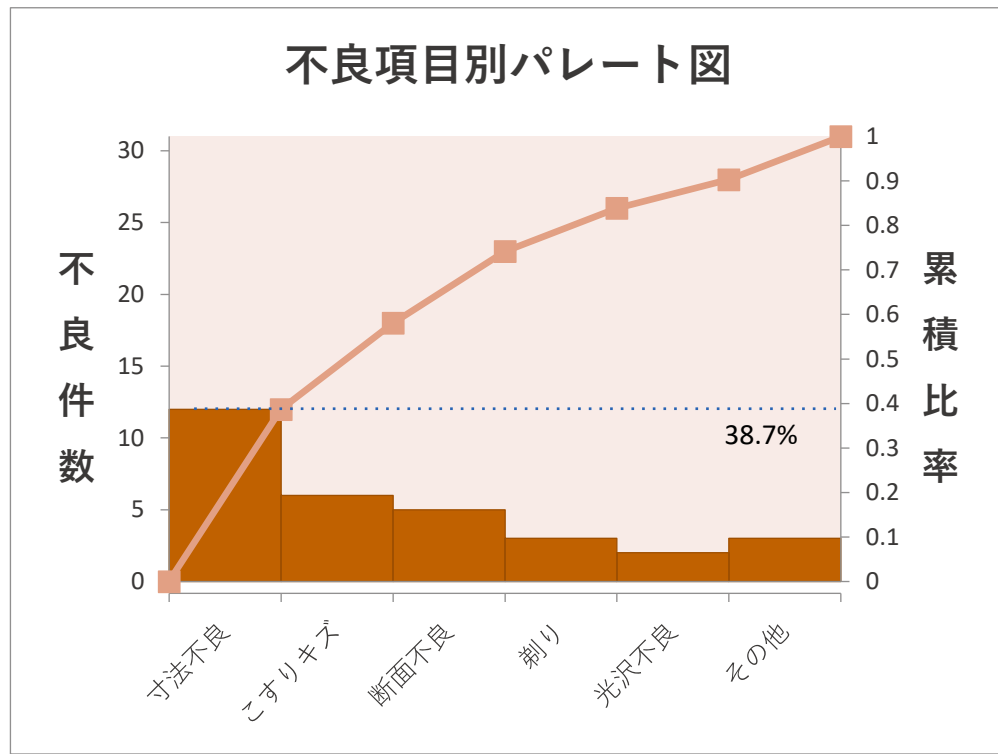
(wikipedia)

パレート図

■ パレート図

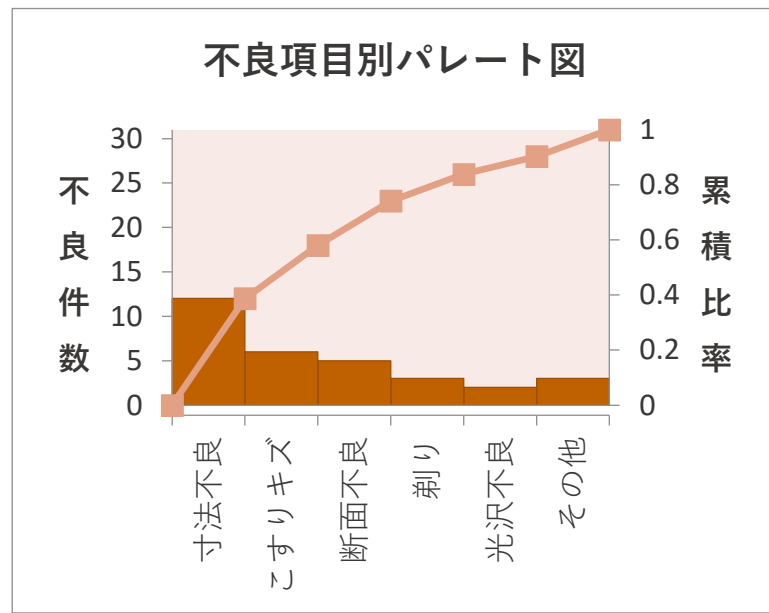
- 件数順に項目を並べ、グラフにしたもの
- 最も問題(重要)な項目を探すための手法

No.	不良項目	件数	累積件数	累積比率
1	寸法不良	12	12	38.7%
2	こすりキズ	6	18	58.1%
3	断面不良	5	23	74.2%
4	剃り	3	26	83.9%
5	光沢不良	2	28	90.3%
6	その他	3	31	100.0%
	合計	31		

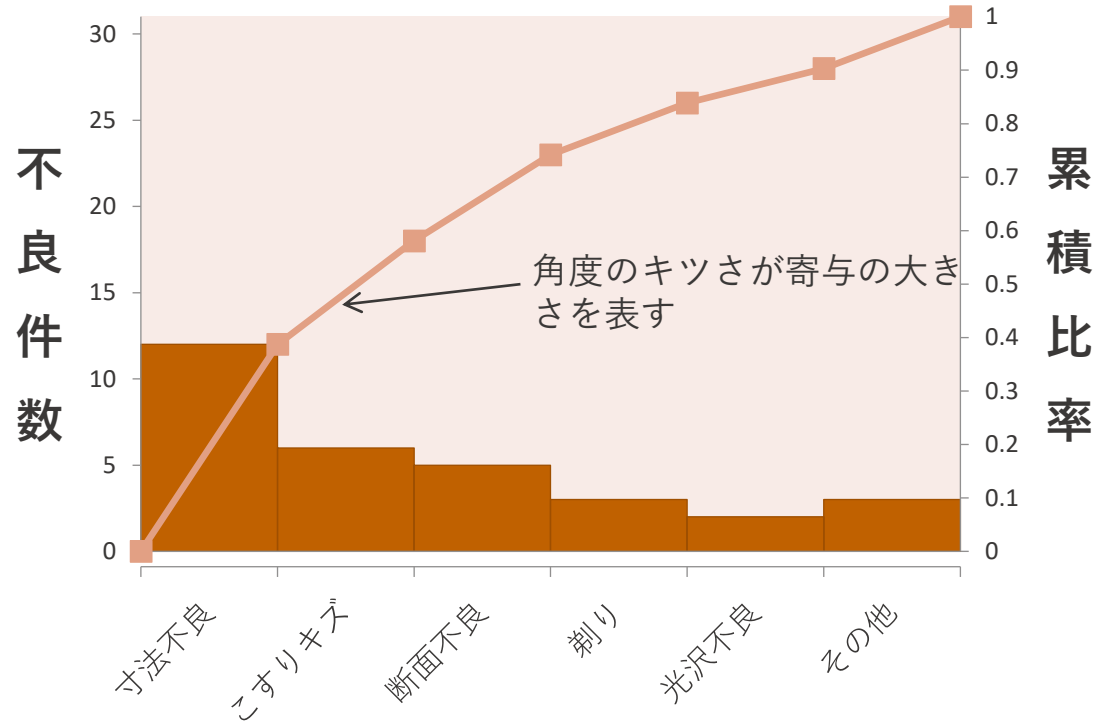


■ パレート図の役割

- 重要な項目を見つける
 - パレートの法則（8対2の法則）
 - 事象の8割は2割の要因から生じる。
 - 2割の人が全体の8割の富をしめる、など
 - 主要な項目に対し、改善をした方が効果的
- 視覚的に不具合の割合をみる



不良項目別パレート図



寸法不良とこすりキズの不良が全体の6割を占めることがわかる。

管理図

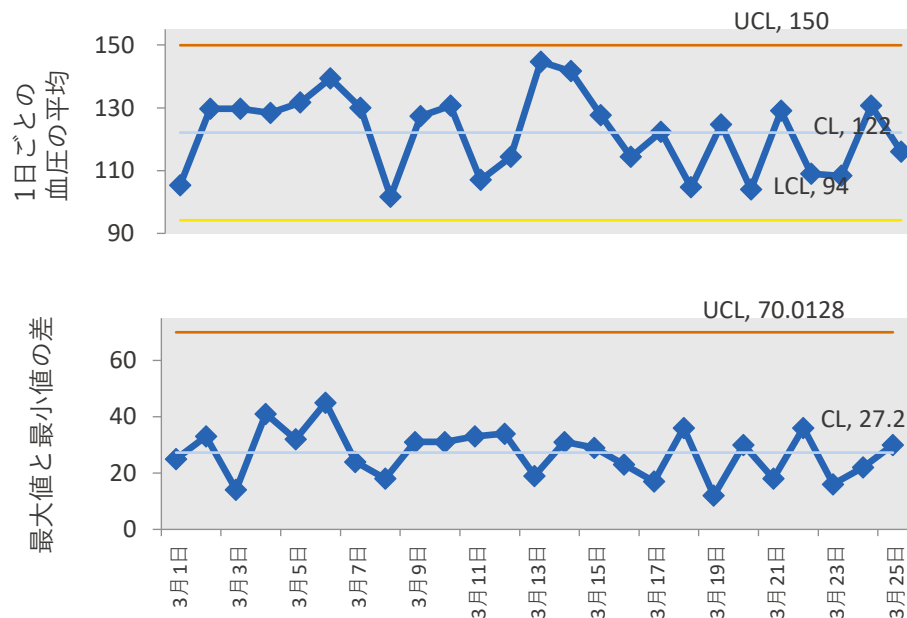
■ 管理図

- 工程の状態を時間推移により把握
- よい状態の維持と管理，異常を見つける

表：血圧の値

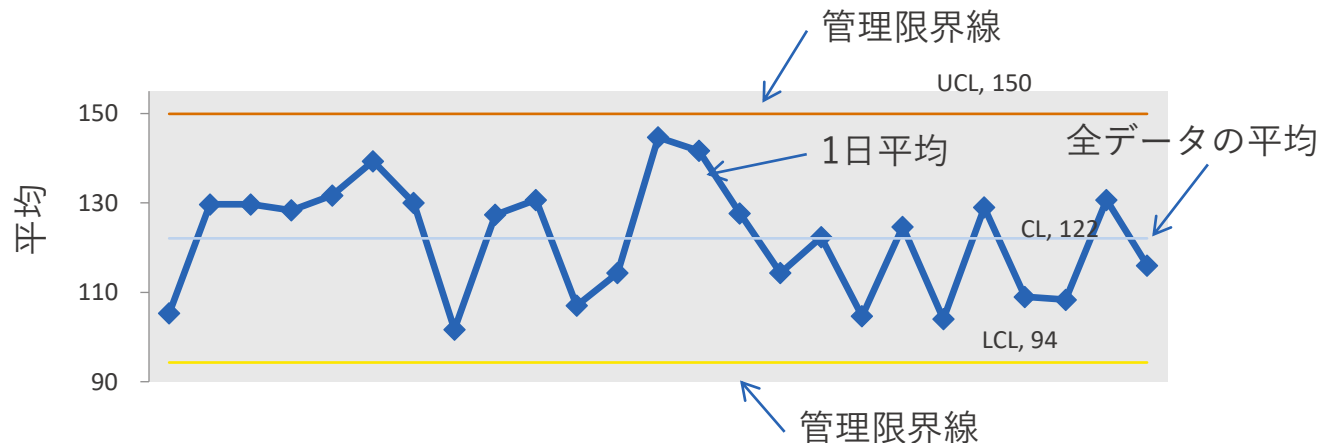
月日	x1	x2	x3
3月1日	95	120	101
3月2日	150	117	122
3月3日	137	129	123
3月4日	143	140	102
3月5日	143	111	141
3月6日	141	116	161
3月7日	128	143	119
3月8日	93	111	101
3月9日	131	110	141
3月10日	116	129	147
3月11日	90	123	108
3月12日	129	95	119
3月13日	153	147	134
3月14日	162	132	131
3月15日	117	120	146
3月16日	128	105	110
3月17日	131	114	122
3月18日	116	117	81
3月19日	128	129	117
3月20日	93	123	96
3月21日	120	129	138
3月22日	117	123	87
3月23日	107	117	101
3月24日	141	132	119
3月25日	105	135	108

X-R管理図



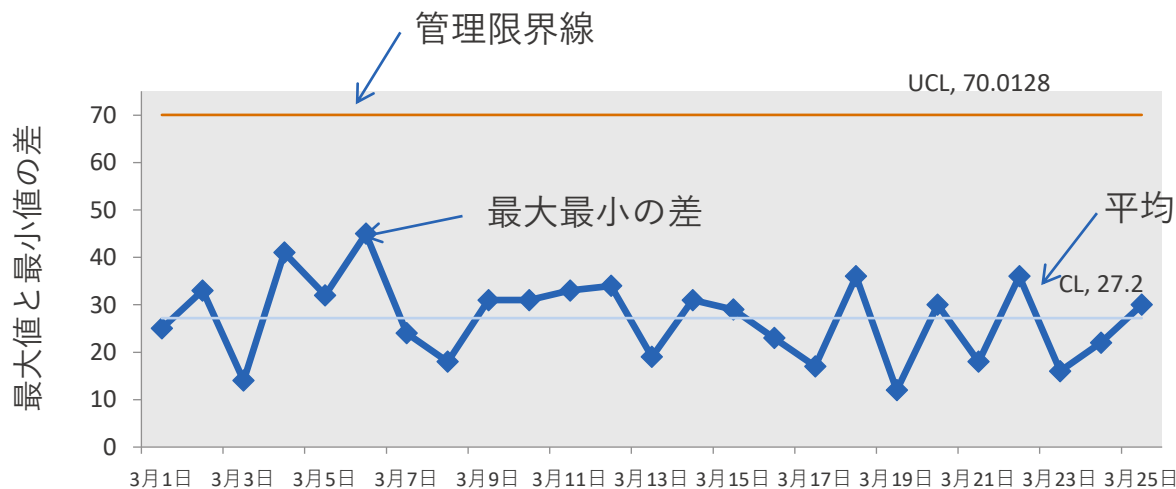
■ 平均の推移

- この図は1日ごと（ある期間ごと）の平均の推移を表す。
- 1日ごとの平均は全体の平均の周りに分布する。
- 全体の平均から1日の平均が大きく外れた場合、異常が発生した可能性がある。
- 管理限界線を超えたら大きく外れたと考える。

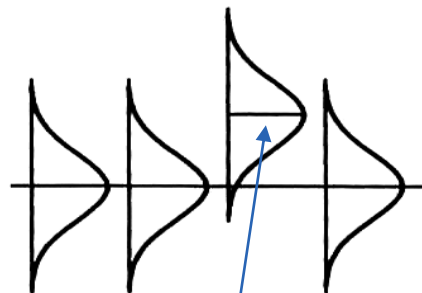
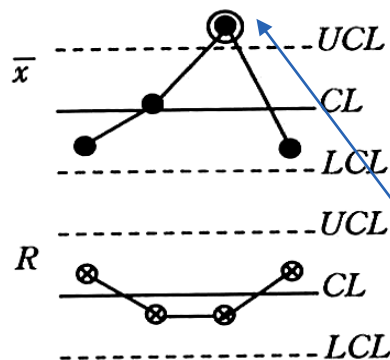


■ 最大と最初の差の推移

- この図は最大と最小の差は1日（ある期間）のデータの分布の広がりを表している。
- 異常が発生しなければ，分布の広がりも変わらなと考えられる。

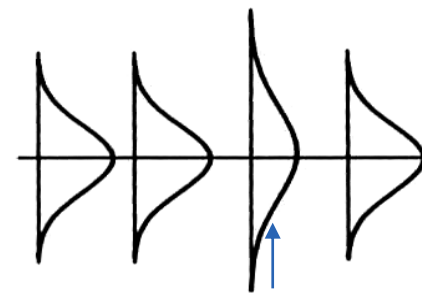
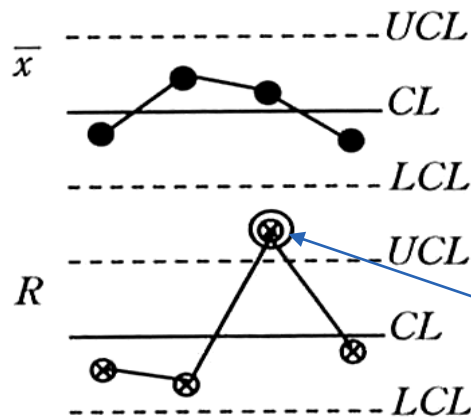


■ 管理図と分布の関係



1日平均が上にずれているので、分布も上にずれる。

1日平均（最大と最小の差）が全体の平均から大きく外れた場合、その日のデータの分布は他の日と異なると考えられる。



1日平均が上にずれているので、分布も上にずれる。

■ 管理図に必要な統計量

- 平均
- 平均の平均
- 最大値、最小値
- 最大値最小値の差 R
- 最大値最小値の差の平均
- 管理限界線(UCL, LCL)

■ 管理限界の計算式

- X管理図

- 上方管理限界
- 下方管理限界

$$UCL = \bar{x} + A_2 \bar{R}$$

$$LCL = \bar{x} - A_2 \bar{R}$$

- R管理図

- 上方管理限界
- 下方管理限界

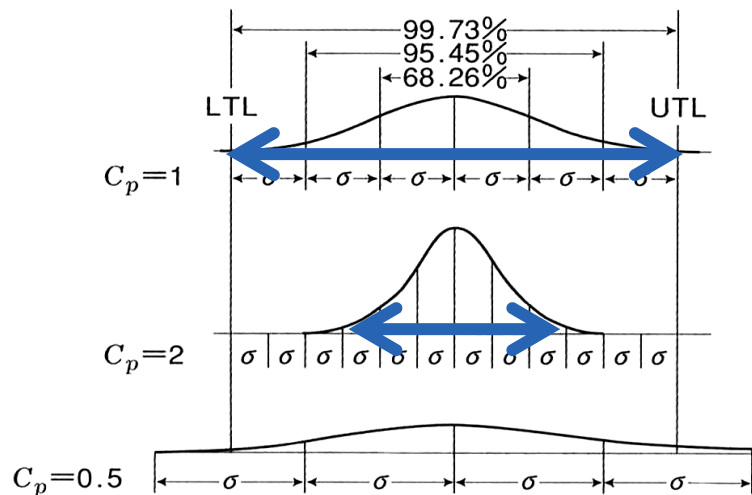
$$UCL = D_4 \bar{R}$$

$$LCL = D_3 \bar{R}$$

サンプルの 大きさ	A2	D3	D4
2	1.88		3.267
3	1.023		2.754
4	0.729		2.282
5	0.577		2.114
6	0.483		2.004
7	0.419	0.076	1.924
8	0.373	0.136	1.864
9	0.337	0.184	1.816
10	0.308	0.223	1.777

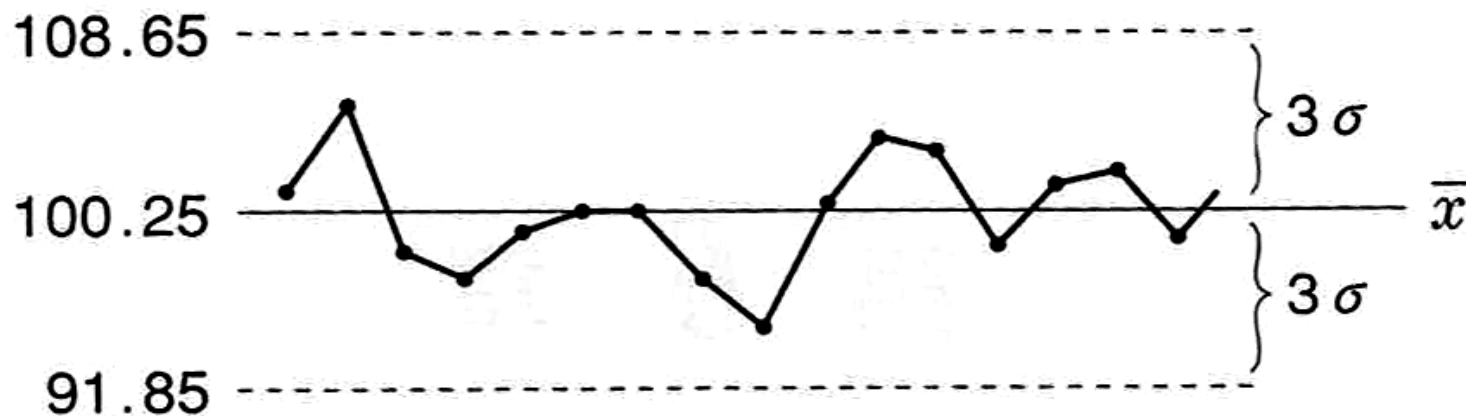
■ 管理限界線の意味

- 管理限界線は平均 ± 3 * 標準偏差を示す.
- 管理限界線を超える可能性0.3%しかないと考えられる.
- つまり, 管理限界線を超えたばあい0.3%の確率でしか起こらない珍しいことが起こった (異常が起こった可能性が高い) といえる.



統計的管理状態

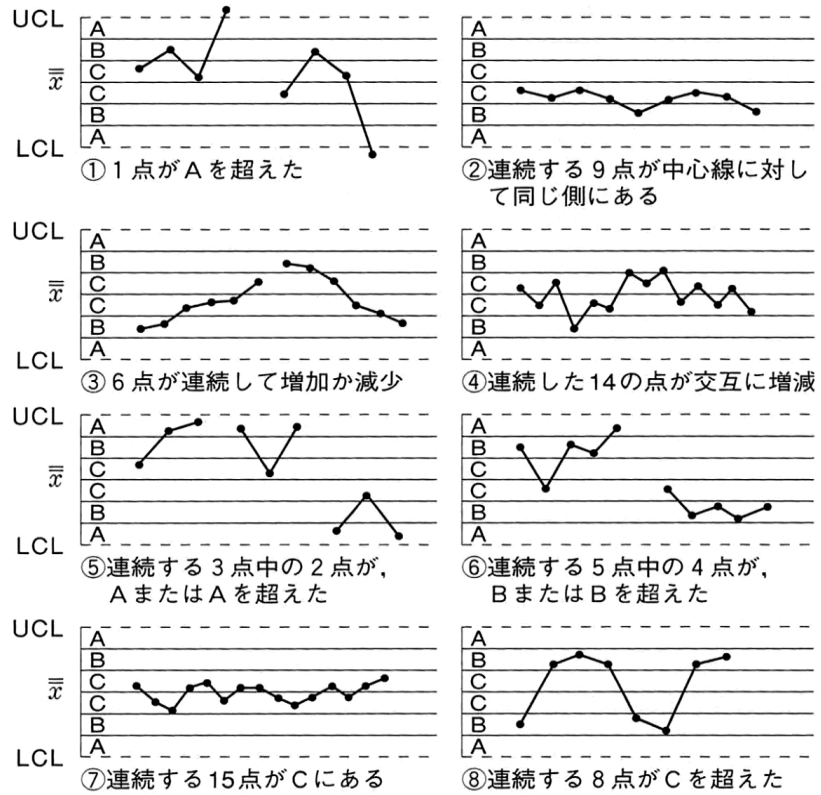
- 基本的に管理限界線の間で状態が推移している状態を統計的管理状態という。



(QC数学の話 大村平より)

■ 管理図から異常を判断

- 統計的にみて通常あり得ない状態を見つける

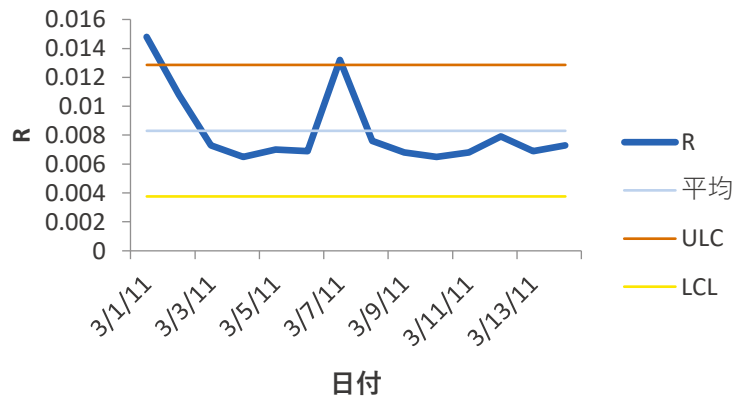
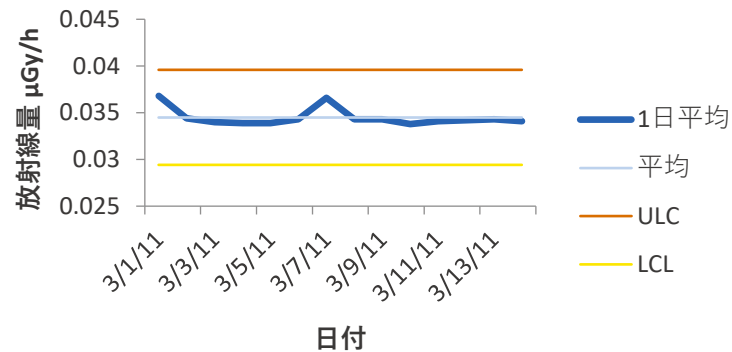


区間	確率
超A	0.00135
A	0.02134
B	0.1360
C	0.3413
C	0.3413
B	0.1360
A	0.02134
超A	0.00135

(QC数学の話 大村平より)

• 新宿の放射線量

X-R管理図

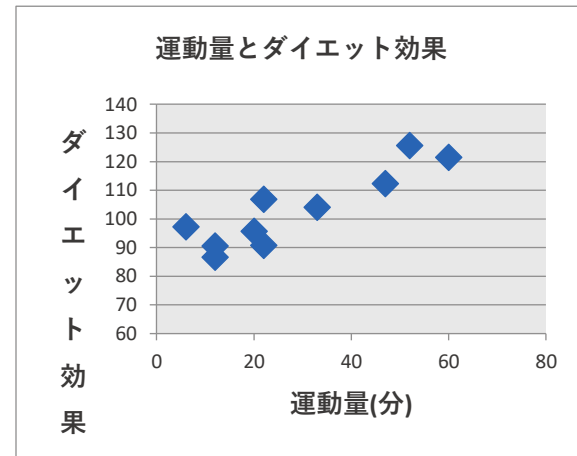
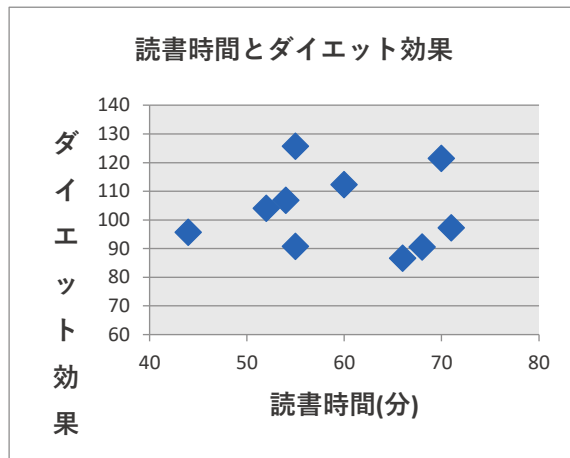
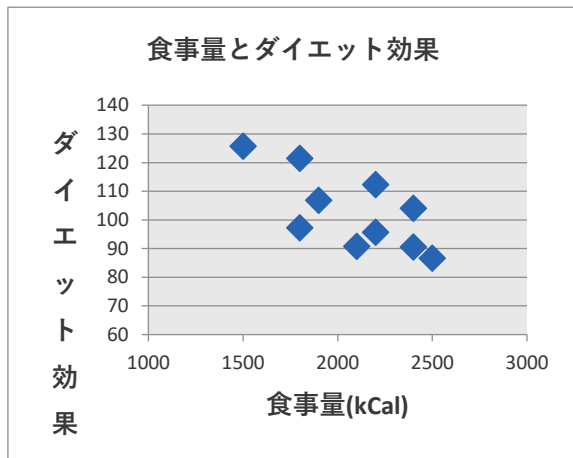


散布図

■ 散布図

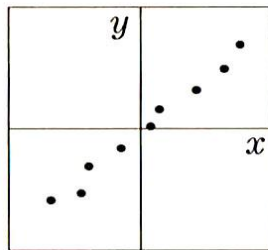
- 要素の関係性をみる.
- 2つの変数間に直線関係に近い傾向がある場合, 相関があるという.

メンバー	食事量 (kCal)	読書時間(分)	運動量(分)	ダイエット効果
スタッフA	1800	70	60	121.5
スタッフB	2200	44	20	95.7
スタッフC	2100	55	22	90.8
スタッフD	2500	66	12	86.7
スタッフE	2400	68	12	90.6
スタッフF	1900	54	22	106.9
スタッフG	1500	55	52	125.7
スタッフH	2200	60	47	112.4
スタッフI	2400	52	33	104.1
スタッフJ	1800	71	6	97.3

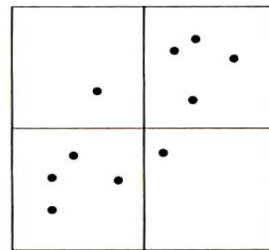


■ 相関

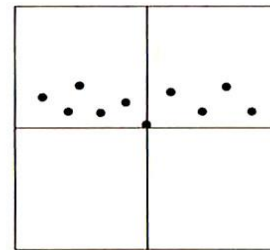
- 一方の変数が増加し、他方の変数も増加する場合を正の相関があるという。
- 逆に、一方の変数が増加し、他方の変数が減少する場合を負の相関があるという。



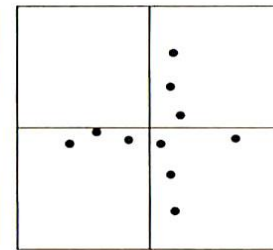
(a) 強い正の相関



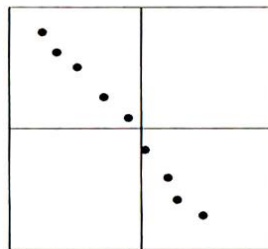
(c) 弱い正の相関



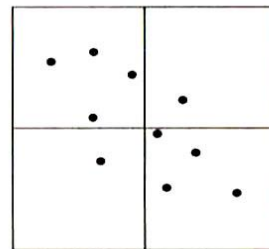
(e) 相関なし



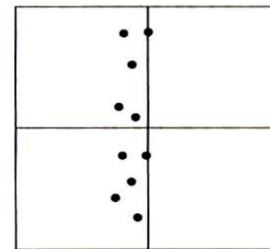
(g) 相関なし



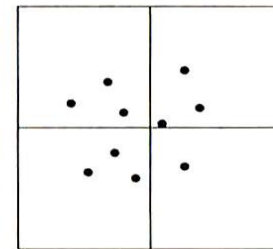
(b) 強い負の相関



(d) 弱い負の相関



(f) 相関なし



(h) 相関なし

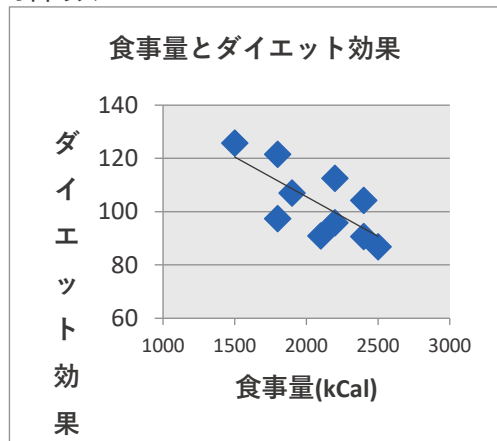
(QC数学の話 大村平より)

■ 相関係数 r

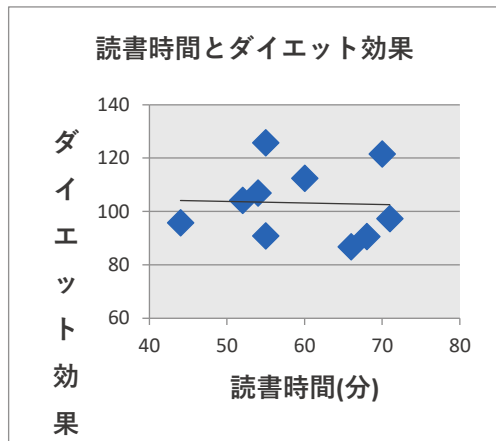
- 2つのデータがどれくらい関係を持っているのかを表す統計量
- -1から1までの数値

$$r = \frac{\sum_{i=1}^n (x_i - \bar{x})(y_i - \bar{y})}{\sqrt{\sum_{i=1}^n (x_i - \bar{x})^2} \sqrt{\sum_{i=1}^n (y_i - \bar{y})^2}}$$

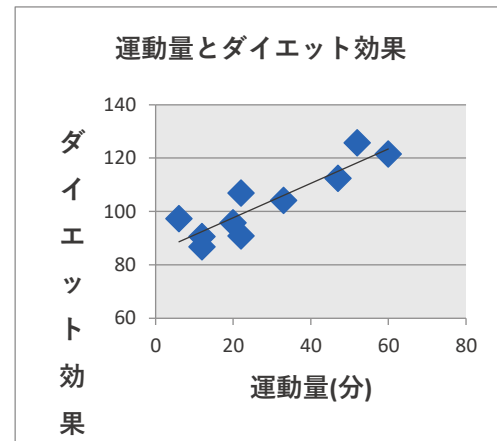
相関係数 $r = -0.72$



$r = -0.038$



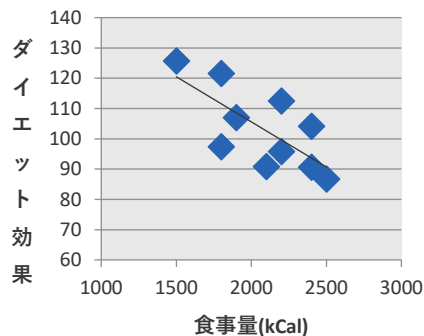
$r = 0.9$



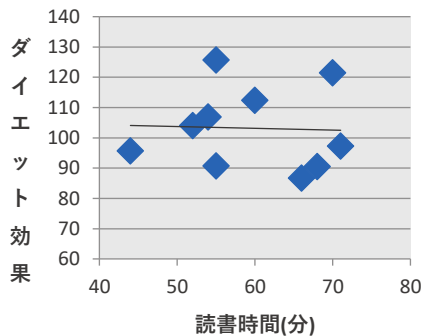
■ 回帰直線

- 2つの変数の関係を $y=bx+a$ に当てはめて得られた直線を回帰直線という。

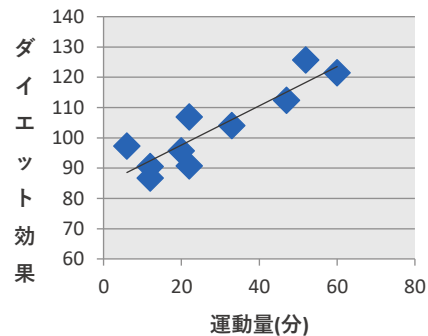
食事量とダイエット効果



読書時間とダイエット効果



運動量とダイエット効果



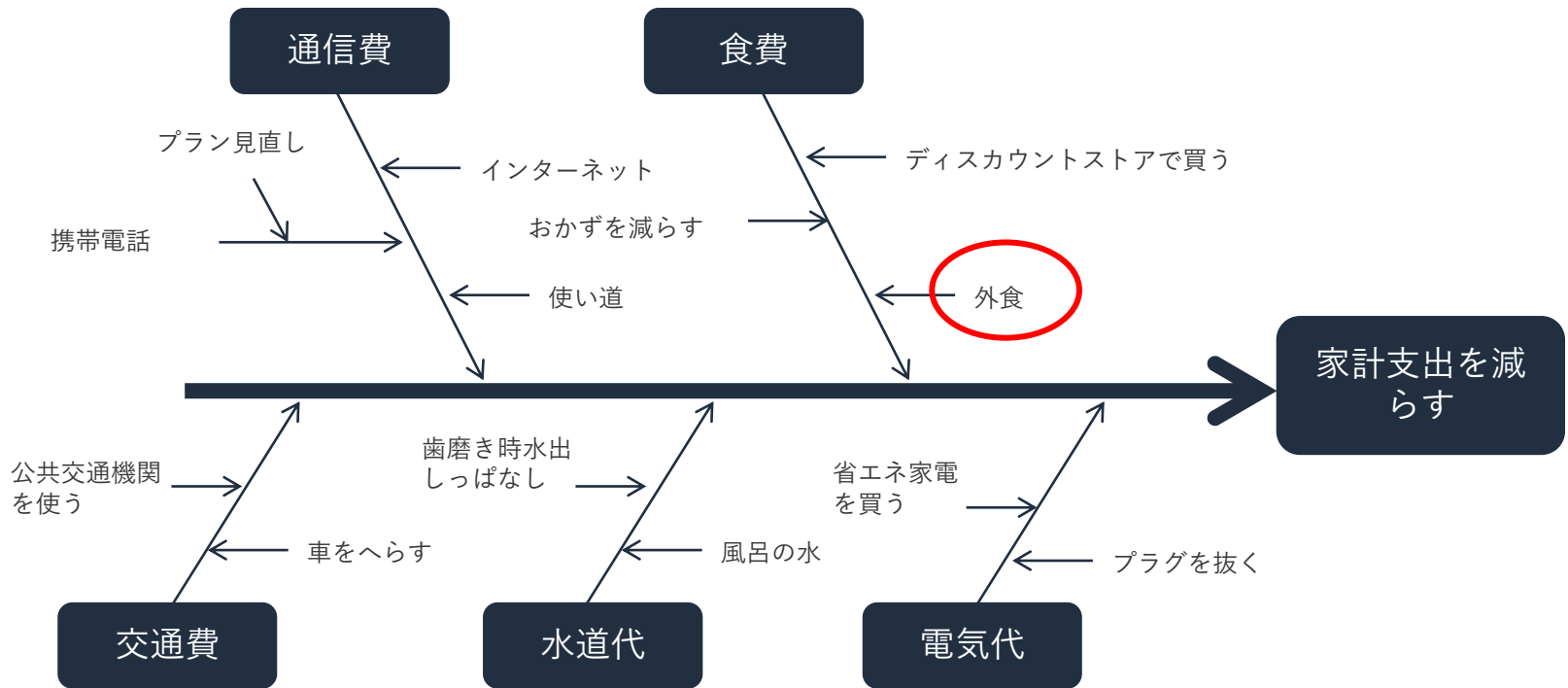
■ おまけ

- 相関があるからと言って因果関係があるわけではない。
 - 例
 - 部屋の明るさと視力の悪さを調べたら、部屋が明るいほど視力が悪かった（負の相関があった）。
 - この結果から、部屋が明るいと視力が悪くなると考えてよいか？
 - 実は、家族の視力が悪いため部屋を明るくしていただけだった。
- 因果関係があるからと言って相関があるわけではない。
 - 例
 - 車のアクセルは踏めば踏むほど速度が出る（因果関係がある）。
 - 因果関係があるからといって、アクセルを踏む量と速度に相関があるわけではない。
- 研究者や技術者でも間違えて使うので注意すること。

特性要因図

■ 特性要因図

- 結果と原因との関係を1つの図に整理してわかりやすくしたもの



チェックリスト

■ チェックリスト

- あらかじめ確認すべき項目を列挙しておいたシートを使って、確認結果を記入していく。

海外旅行持ち物リスト	
持ち物	チェック
パスポート	
海外旅行保険保険証	
航空券控え	
日本円	
現地通貨	
スマホ充電器	
歯ブラシ	
ティッシュ	
パジャマ	
下着	
靴下	
上着	

演習

■ 演習

- クラスの学生の8科目の成績をそれぞれ5段階で評価した。クラス
の平均点と学生の成績の比較や，科目間の成績のバランスを
評価するために用いるグラフとして，最も適切なものはどれか。
(ITパスポート平成25年秋期)

1. 円グラフ
2. 散布図
3. パレート図
4. レーダチャート

■ 演習

- クラスの学生の8科目の成績をそれぞれ5段階で評価した。クラス
の平均点と学生の成績の比較や，科目間の成績のバランスを
評価するために用いるグラフとして，最も適切なものはどれか。
(ITパスポート平成25年秋期)

1. 円グラフ
2. 散布図
3. パレート図
4. レーダチャート

- パレート図を説明したものはどれか。（基本情報平成24年春期）
1. 原因と結果の関連を魚の骨のような形態に整理して体系的にまとめ、結果に対してどのような原因が関連しているかを明確にする。
 2. 時系列的に発生するデータのばらつきを折れ線グラフで表し、管理限界線を利用して客観的に管理する。
 3. 収集したデータを幾つかの区間に分類し、各区間に属するデータの個数を棒グラフとして描き、品質のばらつきを捉える。
 4. データを幾つかの項目に分類し、出現頻度の大きさの順に棒グラフとして並べ、累積和を折れ線グラフで描き、問題点を絞り込む。

- パレート図を説明したものはどれか。（基本情報平成24年春期）
1. 原因と結果の関連を魚の骨のような形態に整理して体系的にまとめ、結果に対してどのような原因が関連しているかを明確にする。
 2. 時系列的に発生するデータのばらつきを折れ線グラフで表し、管理限界線を利用して客観的に管理する。
 3. 収集したデータを幾つかの区間に分類し、各区間に属するデータの個数を棒グラフとして描き、品質のばらつきを捉える。
 4. データを幾つかの項目に分類し、出現頻度の大きさの順に棒グラフとして並べ、累積和を折れ線グラフで描き、問題点を絞り込む。

■ 演習

- 特性要因図の説明として、適切なものはどれか。（基本情報平成17年春期）
1. 原因と結果の関連を魚の骨のような形態に整理して体系的にまとめ、結果に対してどのような原因が関連しているかを明確にする。
 2. 時系列データのばらつきを折れ線グラフで表し、管理限界線を利用して客観的に管理する。
 3. 収集したデータを幾つかの区間に分類し、各区間に属するデータの個数を棒グラフとして描き、品質のばらつきをとらえる。
 4. データを幾つかの項目に分類し、横軸方向に大きさの順に棒グラフとして並べ、累積値を折れ線グラフで描き、問題点を整理する。

■ 演習

- 特性要因図の説明として、適切なものはどれか。（基本情報平成17年春期）
1. 原因と結果の関連を魚の骨のような形態に整理して体系的にまとめ、結果に対してどのような原因が関連しているかを明確にする。
 2. 時系列データのばらつきを折れ線グラフで表し、管理限界線を利用して客観的に管理する。
 3. 収集したデータを幾つかの区間に分類し、各区間に属するデータの個数を棒グラフとして描き、品質のばらつきをとらえる。
 4. データを幾つかの項目に分類し、横軸方向に大きさの順に棒グラフとして並べ、累積値を折れ線グラフで描き、問題点を整理する。