

2009.08.26

「地域データ分析活用講座  
～データをもとに地域と観光を考えてみよう～」  
データ分析実習

# マーケティングにつながる データ分析手法

奈良県立大学 地域創造学部  
遠藤 英樹

# 地域創造のステップ

地域の現状把握



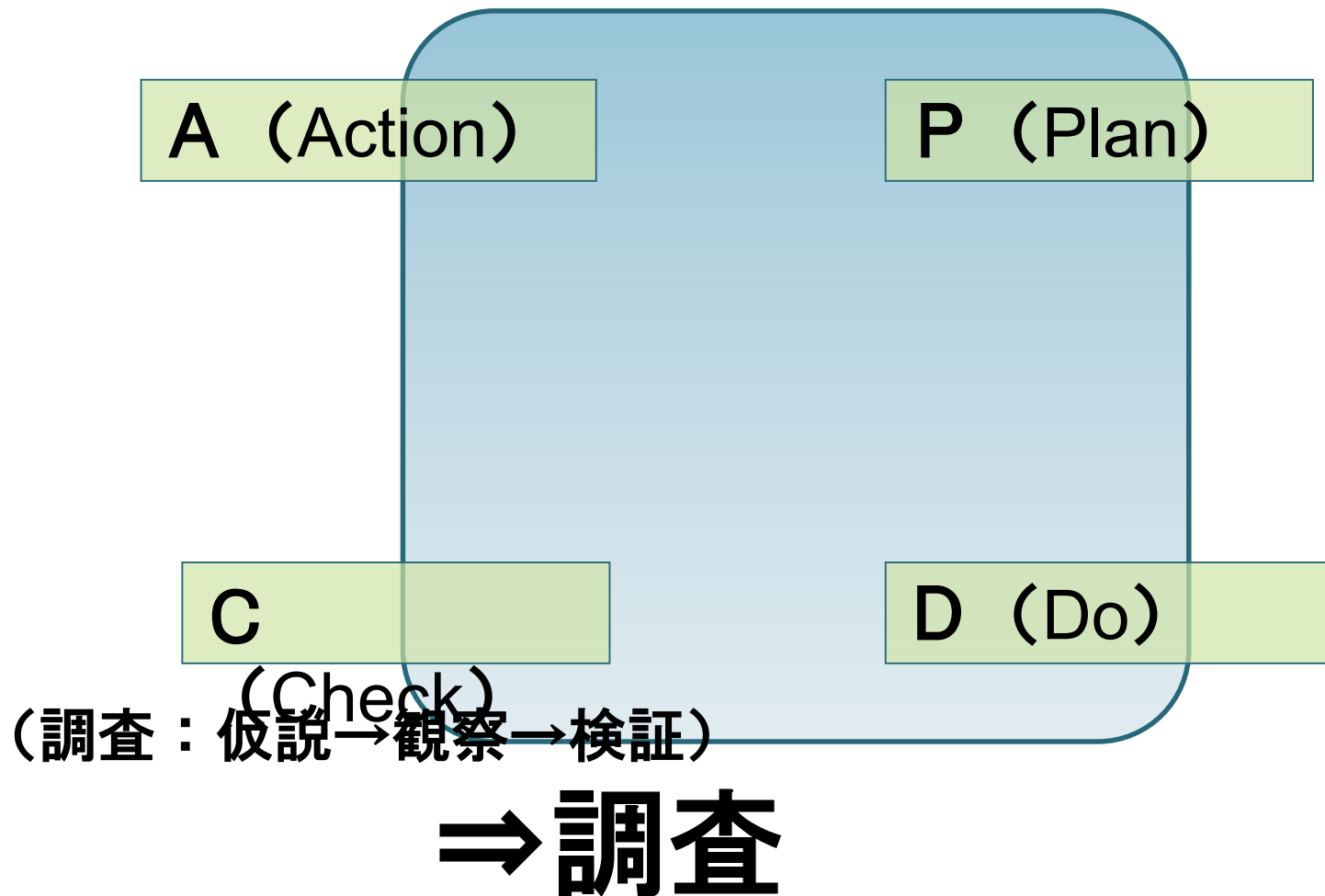
地域づくり

# 地域の現状把握

- 地域や観光が現在どのようなものなのか。
- 地域の中にどのような資源が発見できるのか。
- 地域や観光が現在かかえている問題点は何か。
- そもそも、地域や観光とは何なのか？それを良くするとはどういうことなのか？

● ⇒ 調査

# 地域づくりのサイクル



# 「PDCAサイクル」とは？

- Plan（計画）：従来の実績をはじめとするさまざまなデータ等をもとにして、業務計画を作成する段階
- Do（実施）：計画に沿って業務を行う段階
- Check（評価）：業務を実施した効果を測定し、業務全体を評価する段階
- Action（改善）：評価にもとづいて、改善を行う段階

# 数字で語ることの意味

- **エビデンス・ベーストな地域づくり、観光マーケティング**
- **=数量的データをふまえた、カンだけに頼らない地域づくり、観光マーケティング**

# 数字データにおける2つの種類

- 質問紙調査データ
- →マーケティングにおいて頻繁に用いられる調査データ。
  
- 品質管理データ、経済調査データ、人口調査データ

# 質問紙調査データ

- 人びとがどのように感じているのか、考えているのかを知るための調査。
- これを知るための「ものさし」
- =とても「あやふや」。自己申告のデータであり、回答者の気分によって変化する。
- →これをふまえたうえで、そのデータを分析する。



## 品質管理データ、経済調査データ、人口調査データなど

- それに対して、「ものづくり」で頻繁に用いられる品質管理データ、さらには経済調査データや人口調査データなどは—
- →「あやふや」では困る。非常に精確な数字を出すことができる「測り方」「ものさし」が必要となる。

# 測定する「尺度」の次元

- 1) 名義尺度
- 2) 順序尺度
- 3) 間隔尺度
- 4) 比例尺度

# 尺度の妥当性と信頼性

- 「信頼性」：調査を繰り返しても、同じようなデータが得られること
- 「妥当性」：測定したい事柄を正しく測定できているかどうか。

# 質問紙調査の手順

- 1) 調査企画段階
- 2) 調査設計段階
  - 調査方法を考える：母集団、サンプリングの方法、配布回収方法
  - 質問項目を決定する：概念の操作的定義、変数、仮説
  - 調査票を作成する：ワーディング、選択肢、質問の順番、プリテスト
- 3) 実査段階
  - 調査票印刷、調査の手引き作成など
  - サンプリング作業、郵送作業、調査の実施
- 4) 調査票のデータ化段階
  - 回答のチェック、データ入力、データクリーニング
- 5) データ分析、公表段階

# 調査票におけるいくつかの作成ポイント

- 1) あいまいな言葉は使わない
- 2) 難しい言葉は使わない
- 3) ステレオタイプの言葉は使わない
- 4) 誘導的な質問をしない
- 5) キャリーオーバー効果に気をつける
- 6) ダブルバーレルに気をつける

# 1) あいまいな言葉は使わない

- 「音楽番組がよく放送されますが、あなたはこのような傾向に賛成ですか、反対ですか」
- →どこが「あいまいな言葉」か？ = 「音楽番組」
- →「音楽番組」とは？

## 2) 難しい言葉は使わない

- 「あなたは、 $\text{No}_x$ 規制に賛成ですか、反対ですか」
- → 「 $\text{No}_x$  (窒素酸化物) 規制」は誰でも知っている用語か？

### 3) ステレオタイプの言葉は使わない

- 「官僚が民間企業や政府の外郭団体などに天下りすることについて、どう思いますか」
- → 「官僚」「天下り」
- 「上級国家公務員が民間企業や政府の外郭団体などに再就職することについて、どう思いますか」



## 4) 誘導的な質問をしない

- 前に誘導的な説明をつけない
- 「最近、米国で同時多発テロが起っています。もし、あなたの身内の人や親しい人が、米国に留学などで長期滞在したいと言ったら、どう思いますか」

## 5) キャリーオーバー効果に気をつける

- 後の質問を回答する際に、それに影響を与えるような内容を前に聞いていないか？
- 「イスラム急進派によって、アメリカのニューヨークで起きたテロ行為を知っているか」
- ↓
- 「あなたは、アフガニスタンに対する最近のアメリカの軍事行動について賛成ですか、反対ですか」

## 6) ダブルバーレルに気をつける

- 1つの質問に2つ以上の論点を含めていないか？
- 「あなたは、若い女性がたばこを吸ったり、お酒を飲んだりすることをどう思いますか」

# それでもデータはウソをつく

- 1) 願望によるウソ
- 2) 社会的規範にしばられたウソ
- 3) 世間体や見栄によるウソ
- 4) 質問者を喜ばせるためのウソ
- 5) 「坊主憎けりゃ袈裟まで憎い」ウソ

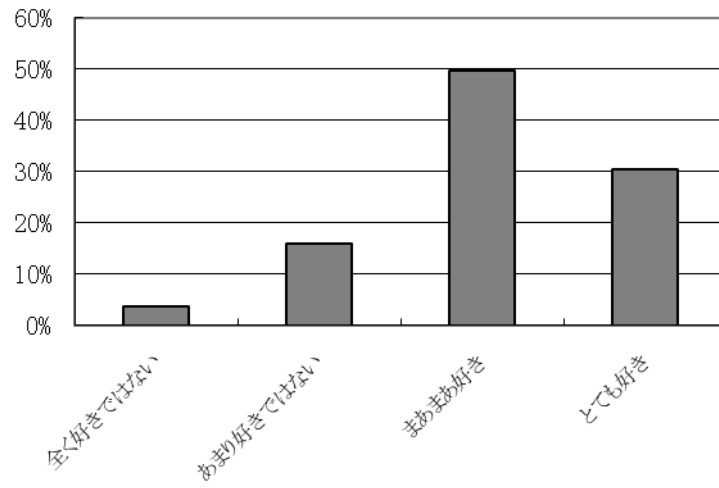
■ ■ ■ ■ ■



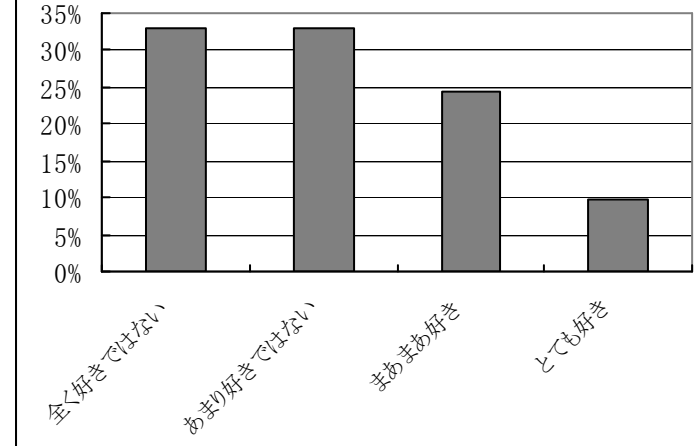
# 基礎統計

- 代表値とは何か？
- 散らばりとは何か？

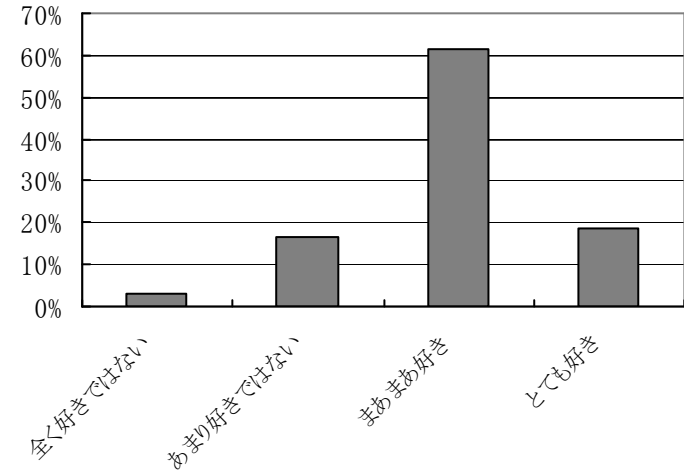
コブクロ



松田聖子



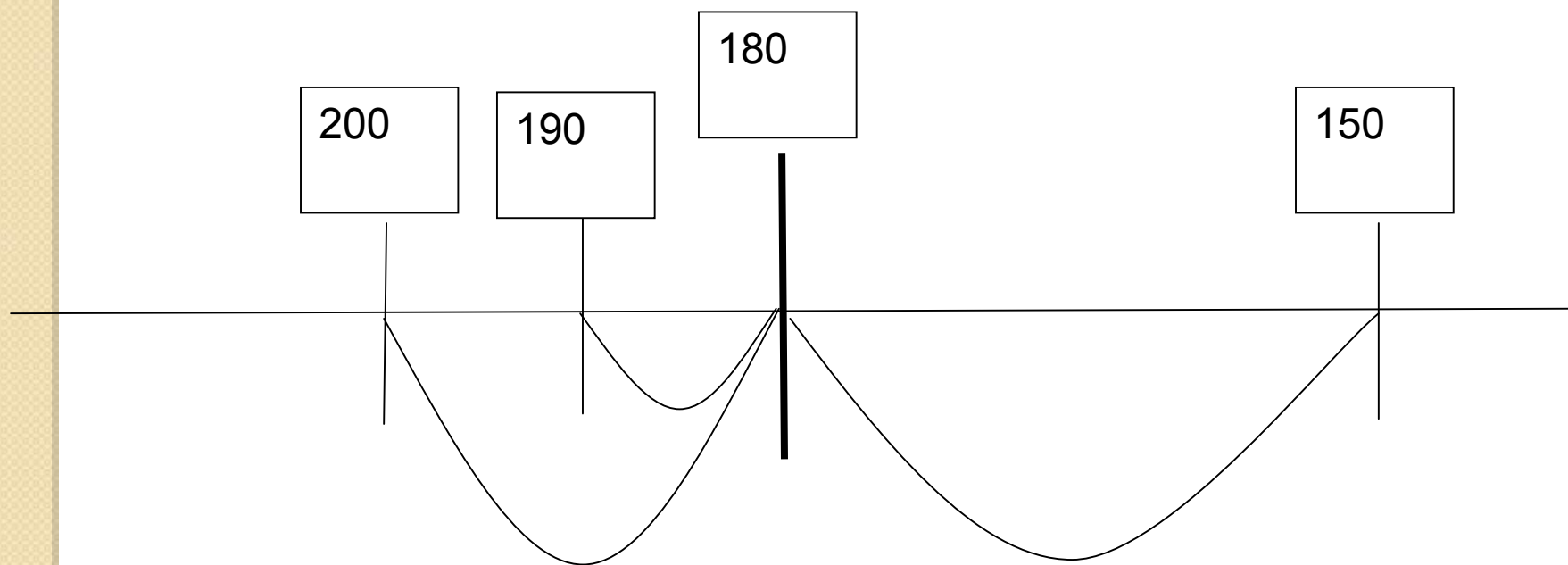
スマップ



# 代表値とは

- 1) 平均値
- 2) 中央値
- 3) 最頻値

# 散らばりとは





# 散らばりとは

- 1) 偏差平方和

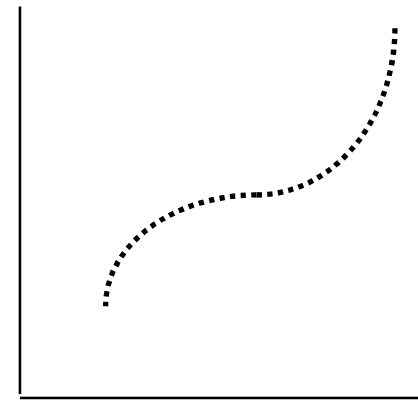
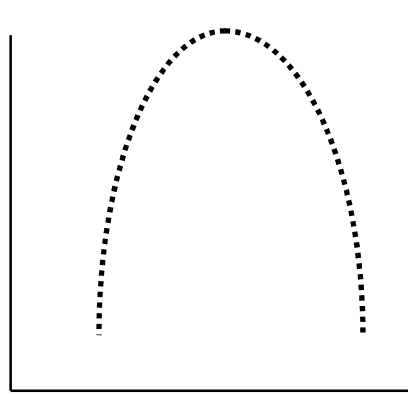
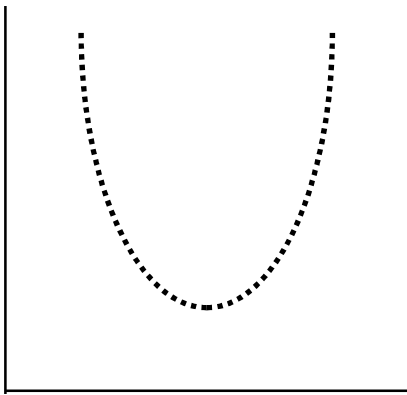
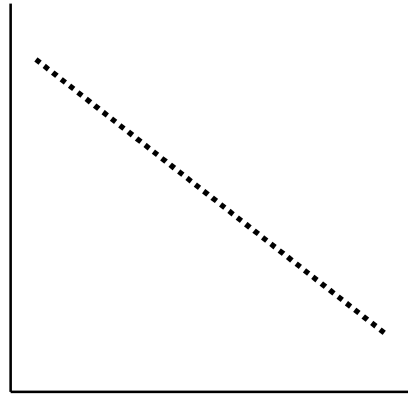
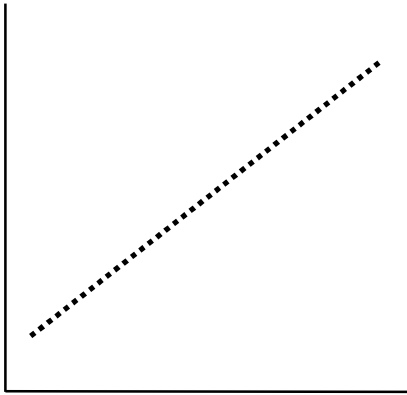
- 2) 分散

偏差平方和 ÷ 個 (人) 数

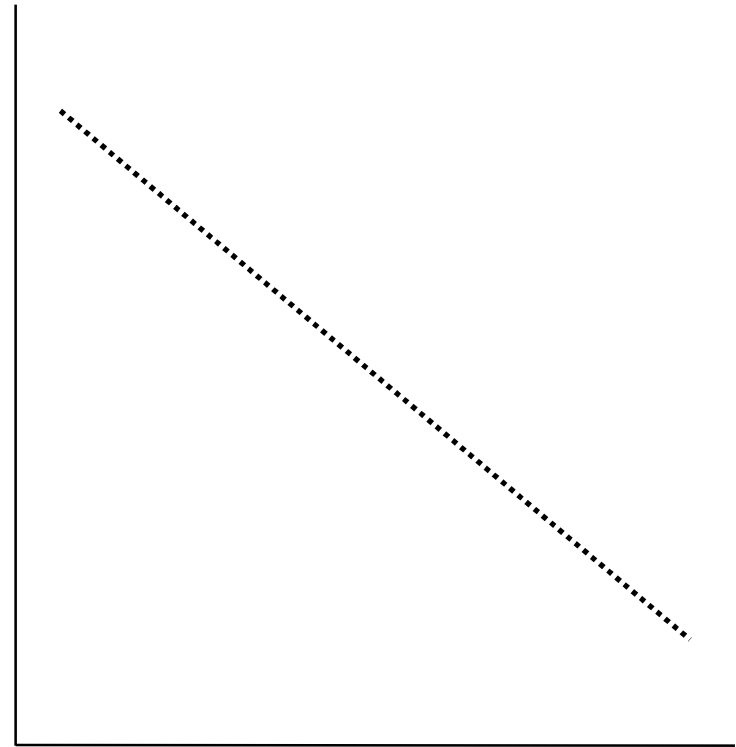
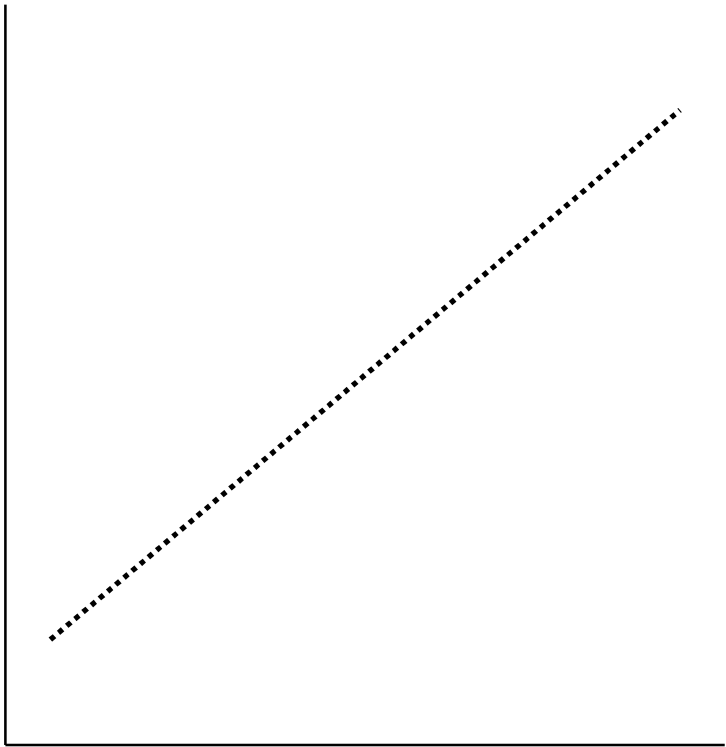
- 3) 標準偏差

$\sqrt{\text{分散}}$

# データの世界のいろいろな関係



# 相関係数でわかる関係とは

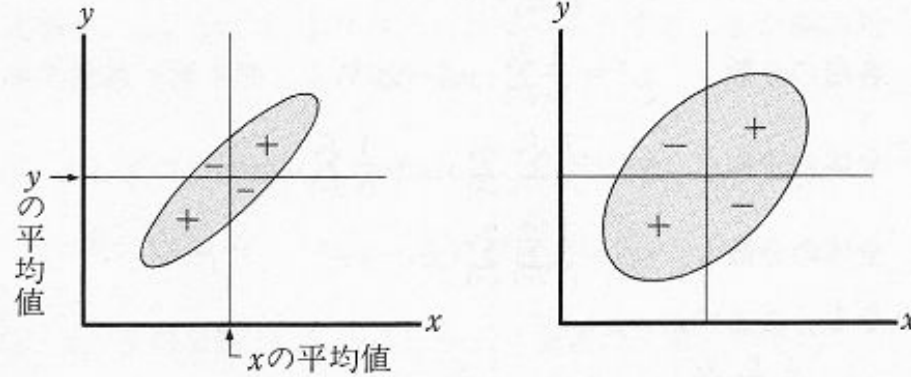


# 相関係数の数値の見方

- $-0.2 \sim 0.2$  : ほとんど相関がない
- $0.2 \sim 0.4$  ( $-0.2 \sim -0.4$ ) : やや相関がある
- $0.4 \sim 0.7$  ( $-0.4 \sim -0.7$ ) : かなり相関がある
- $0.7 \sim 1.0$  ( $-0.7 \sim -1.0$ ) : 強い相関がある

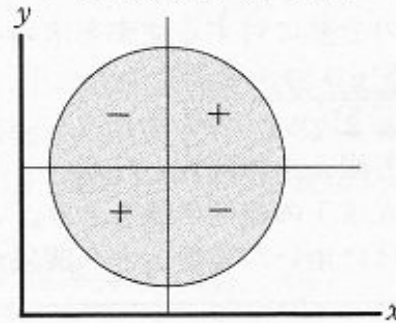
図表 1-7

a 強い正の共変関係がある場合    b 弱い正の共変関係がある場合

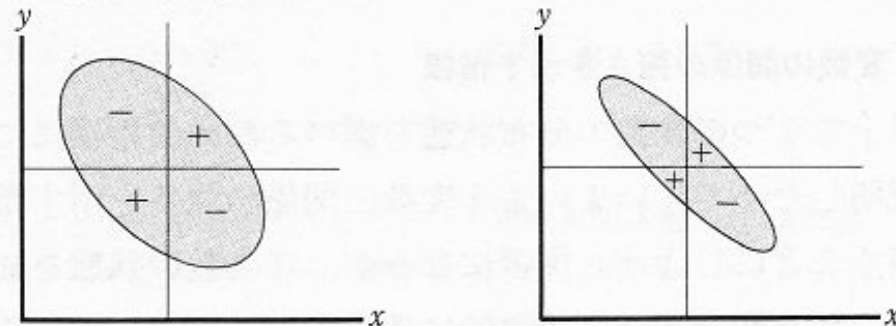


c 共変関係がない場合

+ : 偏差の積が  
正の領域  
- : 偏差の積が  
負の領域



d 弱い負の共変関係がある場合    e 強い負の共変関係がある場合



# 相関係数の計算方法

(xのデータ - xの平均値) × (yのデータ - yの平均値)

／xの偏差平方和のルート × yの偏差平方和のルート