

大学生のための 統計学入門

—テーマ4. 2変量の関連—

京都大学大学院医学研究科
臨床統計学/臨床統計家育成コース 田中司朗



テーマ4. 2変量の関連

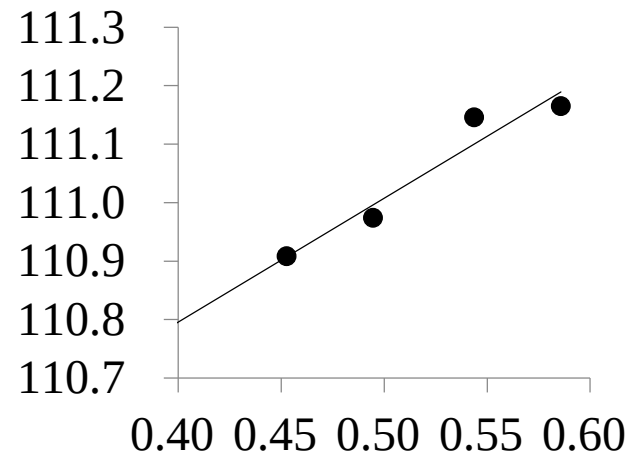
- 地球の大きさを測る
 - 相関
 - 回帰分析
 - 最小二乗法
- **相関と因果**
- ランダム化臨床試験と医療の進歩
 - ランダム化による因果関係の証明



直線関係はほんとうに正しいのか

- 数学者Gaussは測地データに直線を当てはめた
- このケースでは直線関係がある程度正しいことは, 地球が楕円体であることから明らか
- 一般には, 回帰分析を行うときに直線関係が正しいという保証はない

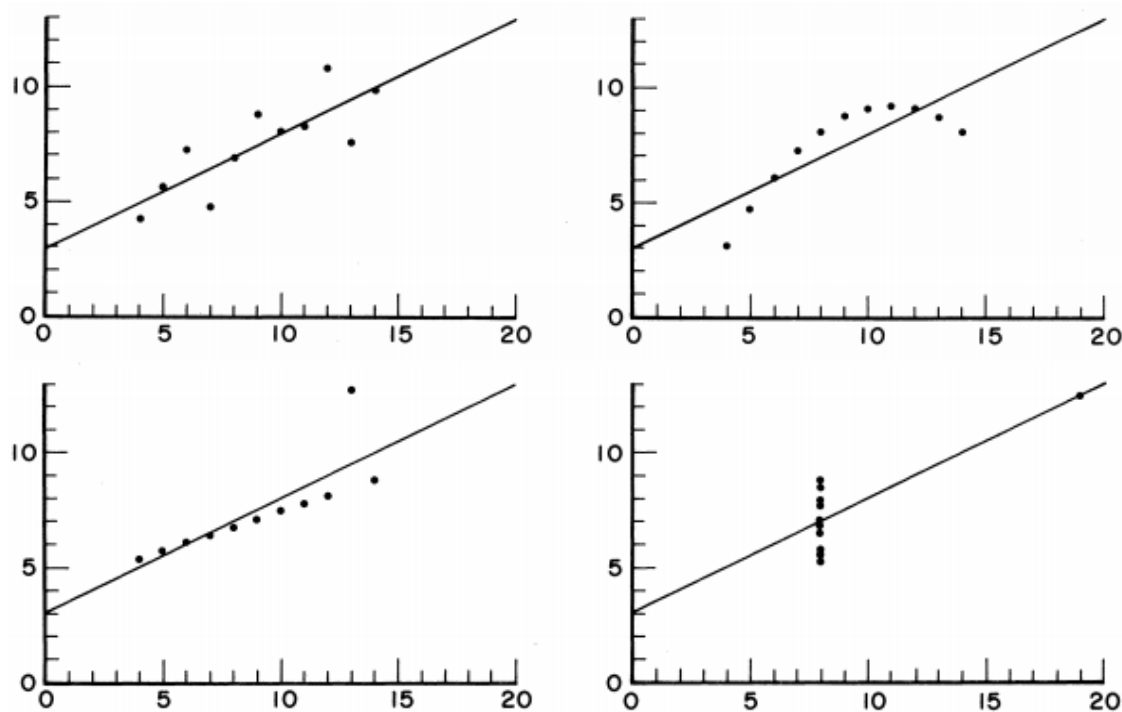
Y=距離/角度



X=sin²(緯度)

回帰直線の誤特定

- 四つの散布図で, YとXの関係はまったく異なる
- 最小二乗法で回帰直線を引くと同じになる

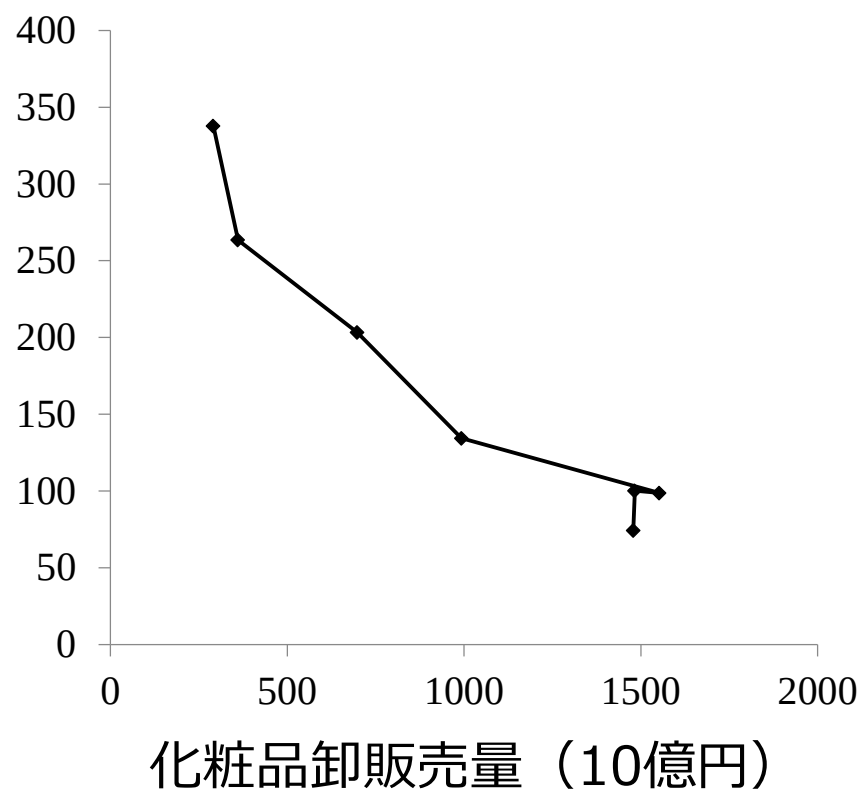


相関関係と因果関係

- 相関
 - 二つの確率変数 X と Y が独立でないこと
- 因果
 - 原因と結果の関係があること
- $Y = \alpha + \beta X + \varepsilon$
 - 2変量の関連を調べるとき, どちらが結果変数でどちらが説明変数なのかは, 解析するものが指定

化粧品普及は脳卒中死亡の原因?

年齢調整脳卒中死亡率



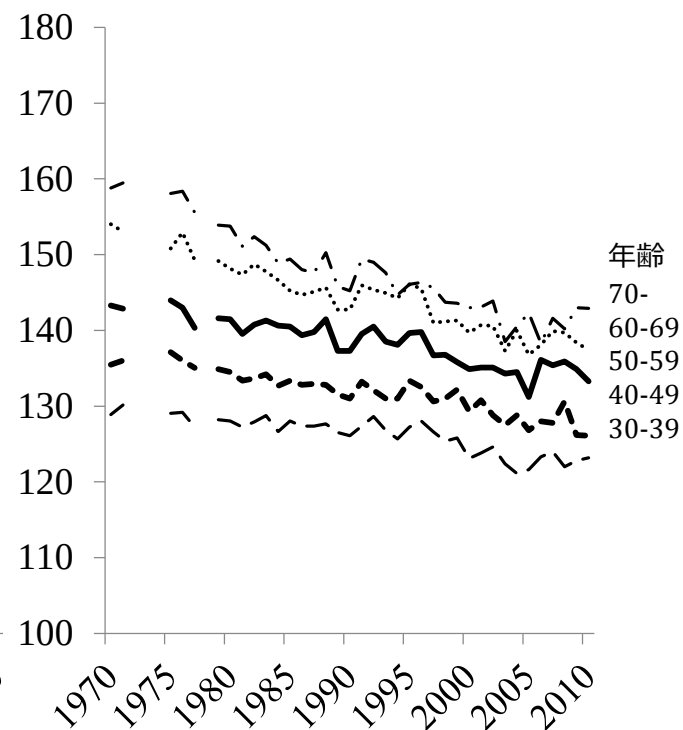
商業統計（産業細分類別） (<http://www.meti.go.jp/statistics/tyo/syougyo/index.html>)
人口動態調査 (<http://www.mhlw.go.jp/toukei/list/81-1.html>)

食塩摂取→血圧→脳卒中死亡？

食塩摂取量 (g/day)



収縮期血圧 (mmHg)



年齢調整死亡率

