



データ解析実習 (3 回目)

2011.10.24 月曜、4・5限

担当教員：高木英至、田端章明

■本日の授業

- ▶ 基本統計量とSPSS
- ▶ 前半(4限)：基本統計量
 - ▶ 代表値、分散・標準偏差、共分散
 - ▶ 相関係数
- ▶ 後半(5限)
 - ▶ SPSS導入



▶ 2

■基本統計量

- ▶ 基礎統計量
 - ▶ 『心理統計学の基礎』、第1～3章
 - ▶ 『社会心理学研究入門』のpp.169-176
- ▶ (1) 代表値(average)
 - ▶ 平均値 (mean)
 - ▶ メディアン
 - ▶ モード
- ▶ (2) 散布度
 - ▶ 分散・標準偏差
 - ▶ 四分位
 - ▶ 共分散
- ▶ (3) 相関係数
 - ▶ ピアソンの積率相関係数
 - ▶ 順位相関係数 (スピアマン、ケンドール...)

▶

▶ 平均
$$\bar{X} = \frac{\sum_{i=1}^n X_i}{n}$$

▶ 分散
$$S_{xx} = \frac{\sum_{i=1}^n (x_i - \bar{x})^2}{n-1} \quad S_{yy} = \frac{\sum_{i=1}^n (y_i - \bar{y})^2}{n-1}$$

▶ 共分散
$$S_{xy} = \frac{\sum_{i=1}^n (x_i - \bar{x})(y_i - \bar{y})}{n-1}$$

▶ 相関係数
$$r_{xy} = \frac{\sum_{i=1}^n (x_i - \bar{x})(y_i - \bar{y})}{\sqrt{\sum_{i=1}^n (x_i - \bar{x})^2 \cdot \sum_{i=1}^n (y_i - \bar{y})^2}} = \frac{S_{xy}}{\sqrt{S_{xx} S_{yy}}}$$

▶

共分散

$$S_{xy} = \frac{\sum_{i=1}^n (x_i - \bar{x})(y_i - \bar{y})}{n-1}$$

- ▶ xが大きいほどyは大きい → 共分散は正の値
- ▶ xが大きいほどyは小さい → 共分散は負の値
- ▶ しかし共分散の値はx,yの尺度の取り方に依存する

▶ 5

相関係数

$$r_{xy} = \frac{\sum_{i=1}^n (x_i - \bar{x})(y_i - \bar{y})}{\sqrt{\sum_{i=1}^n (x_i - \bar{x})^2 \sum_{i=1}^n (y_i - \bar{y})^2}} = \frac{S_{xy}}{\sqrt{S_{xx}S_{yy}}}$$

- ▶ 正確には「ピアソンの積率相関係数」
- ▶ 値の正負については共分散と同じ
 - ▶ xが大きいほどyは大きい → 相関係数は正の値
 - ▶ xが大きいほどyは小さい → 相関係数は負の値
 - ▶ $r_{xy}=0$ のとき、x,y は無相関（関連なし）
- ▶ しかし
 - ▶ 常に $-1 \leq r_{xy} \leq 1$
 - ▶ 相関係数の値は変数の正1次変換に対して不変

▶ 6

線形変換（テキスト p.38）

$$x' = cx + d \quad (2.20) \quad c, d \text{ は定数}$$

$$\bar{x}' = c\bar{x} + d$$

$$s_{x'} = |c|s_x$$

$c > 0$ のとき、正1次変換

▶ 7

標準化：一種の線形変換

$$z = \frac{x - \bar{x}}{s_x} \quad \begin{array}{l} x \text{ が素点 raw score、} z \text{ が標準得点} \\ \text{standard score} \end{array}$$

$$z = \frac{1}{s_x} x - \frac{\bar{x}}{s_x} \quad \bar{z} = 0 \quad s_z = 1$$

$$x = \bar{x} + z s_x$$

x_i, y_i の標準得点 z_{xi}, z_{yi} をかけあわせて総和をとると、 x, y の相関係数になる。

$$\sum_{i=1}^n z_{xi} z_{yi} = \sum_{i=1}^n \left(\frac{x_i - \bar{x}}{s_x} \cdot \frac{y_i - \bar{y}}{s_y} \right) = \frac{\sum_{i=1}^n \frac{(x_i - \bar{x})}{n} \cdot \frac{(y_i - \bar{y})}{n}}{\sqrt{\sum_{i=1}^n \frac{(x_i - \bar{x})^2}{n} \sum_{i=1}^n \frac{(y_i - \bar{y})^2}{n}}} = r_{xy}$$

▶ 8

Q 1

▶ Q 1. x と y を線形変換して x' と y' を求めたとき、 x と y の相関係数と、 x' と y' の相関係数は等しくなるのはなぜか？ 説明せよ

▶ $x' = cx + d, y' = ky + l$

▶ $\rightarrow |r_{xy}| = |r_{x'y'}|$

▶ $c < 0$ の線形変換

▶ 逆転項目

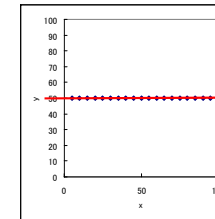
賛成	1	$\rightarrow$	5
やや賛成	2	$\rightarrow$	4
どちらでもない	3	$\rightarrow$	3
やや反対	4	$\rightarrow$	2
反対	5	$\rightarrow$	1

$$x' = 6 - x = -1 \cdot x + 6$$

▶

Q 2

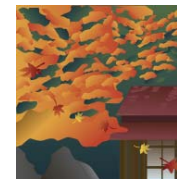
▶ 次のような直線上にデータがあると相関係数 (r_{xy}) はどうかの？



▶

▶ Q 3. 変数 x と y があるとして、 x の値が高ければ y の値も高くなる傾向があるとき、 x と y の相関係数が正になる理由を簡単に説明しなさい。

▶



4限は おしまい

▶ 12