

## 第7回 5月13日の授業内容

- ◆ § 2.4 分布の量的記述 (続)
  - § 2.4.3 その他の代表値
  - § 2.4.4 記述例
- ◆ § 2.5 2変量データの記述
  - 散布図
  - 共分散
  - 相関係数

5/13/03

1

---

---

---

---

---

---

---

---

## § 2.4.3 その他の代表値

- ◆ 変動係数 coefficient of variation
  - 相対的な散らばりの大きさを測る指標  
変動係数 = 標準偏差 / 平均
- ◆ 歪度 skewness
  - 分布の歪みを測る指標
  - 左右対称の形に近ければ、0近辺の値をとる。
- ◆ 尖度 kurtosis
  - 分布の尖り、裾の厚さを測る指標

5/13/03

2

---

---

---

---

---

---

---

---

## § 2.4.3 その他の代表値

- ◆ Note: 偏差値
  - 平均が50、標準偏差が10になるように変換したもの

$$H_i = \frac{(X_i - \bar{X})}{\sigma} \times 10 + 50$$

- ◆ Note: 標準化 (基準化)
  - 平均が0、標準偏差が1になるように変換すること

5/13/03

3

---

---

---

---

---

---

---

---

### § 2.4.4 記述例(1)

- ◆データ: ある属性をもつ標本の通学時間
- ◆標本数  $n=15$ , 単位(分)
- ◆90, 30, 120, 75, 70, 100, 20, 110, 100, 45, 50, 115, 100, 110, 80
- ◆データの処理の簡単化のために昇順に並び替える
  - 20, 30, 45, 50, 70, 75, 80, 90, 100, 100, 100, 110, 110, 115, 120

5/13/03

4

---

---

---

---

---

---

---

---

---

---

---

---

### § 2.4.4 記述例(2) 計算過程

No.(j)	X(j)	X(j)-Xbar	square	
1	20	-61	3721	
2	30	-51	2601	
3	45	-36	1296	
4	50	-31	961	
5	70	-11	121	
6	75	-6	36	
7	80	-1	1	
8	90	9	81	
9	100	19	361	
10	100	19	361	
11	100	19	361	
12	110	29	841	
13	110	29	841	
14	115	34	1156	
15	120	39	1521	
sum	1215	0	14260	
mean	81		950.6667	

5/13/03

5

---

---

---

---

---

---

---

---

---

---

---

---

### § 2.4.4 記述例(3)

- ◆分布の中心を示す代表値
  - メディアン: 90(分)
  - 算術平均: 81(分)
  - 切り落とし平均(大小2個を落とす): 84.5(分)
- ◆分布の広がりを示す代表値
  - レンジ:  $120-20=100$
  - 分散: 950.6667
  - 標準偏差: 30.83(分)

5/13/03

6

---

---

---

---

---

---

---

---

---

---

---

---

### § 2.5 2変量データの記述(1)

- ◆ペア(x, y)になったデータの記述  
関係性の記述
- ◆散布図: 視覚からの把握
- ◆共分散: xとyが共に動く大きさを示す指標  
(共変動)
- ◆相関係数: xとyの関係が線形に近いかを  
示す指標

5/13/03

7

---

---

---

---

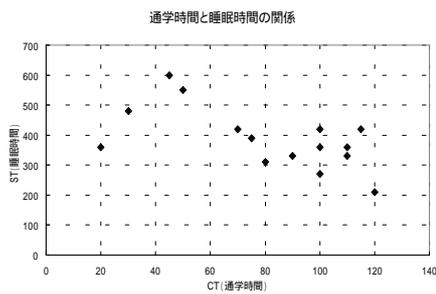
---

---

---

---

### § 2.5 2変量データの記述(2) 散布図



5/13/03

8

---

---

---

---

---

---

---

---

### § 2.5 2変量データの記述(3)

- ◆共分散: 2変量(ペアになった変数)の共変動の大きさをあらわす代表値

◆定義

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

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

5/13/03

9

---

---

---

---

---

---

---

---

§ 2.5 2変量データの記述(4)

◆相関係数: 2変量間の線形関係の強弱を表す指標  $-1 \leq r \leq 1$

◆定義

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

5/13/03

10

---

---

---

---

---

---

---

---

§ 2.5 2変量データの記述(5) 計算過程

No.	CT	ST	CT-Ctbar	ST-Stbar	CT*STb
1	90	330	9	-57.33	-516
2	30	480	-51	52.67	-4726
3	120	210	39	-177.33	-6916
4	75	390	-6	2.67	-16
5	70	420	-11	32.67	-359.333
6	100	360	19	-27.33	-519.333
7	20	360	-61	-27.33	1667.333
8	110	330	29	-57.33	-1662.67
9	100	270	19	-117.33	-2229.33
10	45	600	-36	212.67	-7656
11	50	550	-31	162.67	-5042.67
12	115	420	34	32.67	1110.667
13	100	420	19	32.67	620.667
14	110	360	29	-27.33	-792.667
15	80	310	-1	-77.33	77.33333
sum	1215	5810	0	0.00	-26960
mean	81	387.33			-1797.33

5/13/03

11

---

---

---

---

---

---

---

---

§ 2.5 2変量データの記述(6) 実際例

◆通学時間と平均睡眠時間の関係

- 共分散:  $-1797.33$
- 標準偏差: 通学時間  $30.83$   
平均睡眠時間  $97.81$
- 相関係数:  $r = -0.596$   
 $= -1797.33 / (30.83 \times 97.81)$

5/13/03

12

---

---

---

---

---

---

---

---