

第5回 5月6日の授業内容

- ◆ § 2. 記述統計
- ◆ § 2.3 データの分布
- ◆ § 2.4 分布の量的記述

5/6/03

1

§ 2.3 データの分布

- ◆ データのまとまりを理解しやすくする方法
 - 度数分布表(頻度分布表)
 - ヒストグラム
 - 幹葉表示(stem and leaf display)

5/6/03

2

データ例

- ◆ O大学下宿生の通学時間(分)
標本数19
40, 15, 60, 15, 45, 10, 15, 15, 18, 20, 47,
10, 5, 15, 20, 60, 15, 25, 10

5/6/03

3

度数分布表 (頻度分布表)

frequency distribution

データ(数字・記号)の羅列を表の形で整理したもの。変数のタイプによって作成方法は異なる。

◆質的データの場合

- 質的データはとる値(ラベル・順序など)によって分類することができる。そして分類した各値に属する個体数を数えればよい。

5/6/03

4

度数分布 (頻度分布) (2)

◆量的・離散データの場合

- 質的データと同様に度数分布を作成。但し、注意すべき点は、とる値の順番に意味がある(昇順に並べる)。また、ある値以下の数値をとる度数もデータの記述に意味を持つ。

累積度数、累積相対度数

5/6/03

5

度数分布 (頻度分布) (3)

◆量的・連続データの度数分布

- 連続データは実数値をとるので、離散データとは異なり、同じ値が複数観測されることは稀。観測値を等間隔に幾つかのグループ(区間・階級)に区切り、その区間に入る個体の数をカウントする。

5/6/03

6

度数分布(頻度分布)(4)

- ◆連続データの度数分布について
 - グループ化されていることに注意。グループ化の方法によって、見え方が変わる。
- ◆区間の取り方
 - 区間の個数 多すぎても少なすぎてもダメ。
 - Note . スタージェス(Starjes)の公式

$$\text{区間数} = 1 + 3.3 \times \log(\text{標本数})$$
 - 区間の始点と終点

5/6/03

7

データ例(度数分布)

階級(区間)	階級値	度数	累積度数	相対度数	累積相対度数
0より大 ~ 10以下	5	4	4	0.21	0.21
10 ~ 20	15	9	13	0.47	0.68
20 ~ 30	25	1	14	0.05	0.74
30 ~ 40	35	1	15	0.05	0.79
40 ~ 50	45	2	17	0.11	0.89
50 ~ 60	55	2	19	0.11	1.00
計		19		1.00	

5/6/03

8

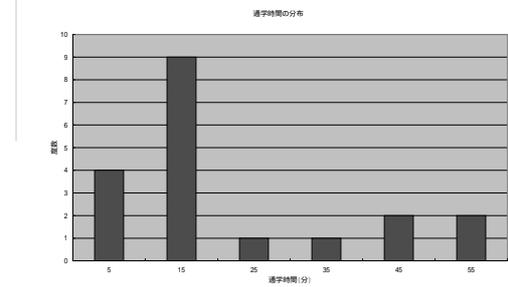
ヒストグラム histogram

- ◆度数分布表、相対度数分布表を棒グラフ(bar chart)で表したもの

5/6/03

9

データ例(ヒストグラム)



5/6/03

10

幹葉表示(stem and leaf display)

◆ヒストグラムでは、区間内のデータの情報が失われてしまう。そこで、区間内のデータの情報も表示できるように工夫したヒストグラム。

5/6/03

11

データ例(幹葉表示)

```

0 | 5000
1 | 555555800
2 | 5
3 | 0
4 | 57
5 | 00

```

5/6/03

12
