Exploratory Statistical Data Analysis With R Software (ESDAR)

Swayam Prabha

Lecture 21

Mode

Shalabh

Department of Mathematics and Statistics

Indian Institute of Technology Kanpur

Slides can be downloaded from

http://home.iitk.ac.in/~shalab/sp

Mode

Examples:

- **A fruit juice shop owner wants to know that which of the fruit juice is more preferred.**
- **A clothing shop owner wants to know which size of shirt and/ or trouser is highest in demand.**

Mode of *n* observations $x_1, x_2, \ldots, x_n$ is the value which occurs the most, compared with all other values.

Mode

Mode is the value which occurs most frequently in a set of observations.

Mode is not at all affected by extreme observations.

Distributions having only one mode are called unimodal and the one with two modes is called bimodal.

Mode for Ungrouped Data

For discrete variables, the mode of a variable is the value of the variable having the highest frequency in a unimodal distribution.

Mode for Ungrouped Data

R command

Step 1: Create a table of given data vector or matrix `data`, say `modetab` as follows:

```
modetab = table(as.vector(data))
```

The first row of `modetab` is a sorted list of all unique values in `data`.

Step 2: Following returns the names of the values having the highest count in second row of `modetab` which is the mode.

```
names(modetab)[modetab == max(modetab)]
```

Mode for Ungrouped Data

R command

```
> data = c(10,10,10,10,2,2,3,4,5,6)
```

Create a table of given data vector `data`

```
> modetab = table(as.vector(data))
> modetab
 2  3  4  5  6 10
 2  1  1  1  1  4
```

sorted list of all unique values in `data`

names of values having highest count

```
> names(modetab)[modetab == max(modetab)]
[1] "10"
```

Mode for Ungrouped Data

R command

R Console

```
> data = c(10,10,10,10,2,2,3,4,5,6)
>
> modetab = table(as.vector(data))
> modetab

 2  3  4  5  6 10
 2  1  1  1  1  4
>
> names(modetab)[modetab == max(modetab)]
[1] "10"
```

Mode for Ungrouped Data

R command

```
> data = matrix(nrow= 3, ncol=3, data=c(1,2,2,
3,3,4,5,6,6))


> data
     [,1] [,2] [,3]
[1,]    1    3    5
[2,]    2    3    6
[3,]    2    4    6
```

Mode for Ungrouped Data

R command

Create a table of given data matrix `data`

```
> modetab = table(as.vector(data))


> modetab
1 2 3 4 5 6
1 2 2 1 1 2


> names(modetab)[modetab == max(modetab)]
[1] "2" "3" "6"
```

Mode for Ungrouped Data

R command

R Console

```
> data = matrix(nrow=3, ncol=3, data=c(1,2,2,3,3,4,5,6,6))
> data
     [,1] [,2] [,3]
[1,]    1    3    5
[2,]    2    3    6
[3,]    2    4    6
> modetab = table(as.vector(data))
> modetab

1 2 3 4 5 6
1 2 2 1 1 2
> names(modetab)[modetab == max(modetab)]
[1] "2" "3" "6"
>
```

Note:

R command `mode` does not provide mode.

It provides the mode of the data vector.

Mode for Grouped Data

For continuous variable, the mode is the value of the variable with the highest frequency density corresponding to the ideal distribution which would be obtained if the total frequency were increased indefinitely and if, at the same time, the width of the class intervals were decreased indefinitely.

Mode for Grouped Data

Class intervals	Mid point (m_i)	Absolute frequency (f_i)
e_1- e_2	$m_1 = (e_1 + e_2)/2$	f_1
e_2- e_3	$m_2 = (e_2 + e_3)/2$	f_2
...	...	...
e_{K-1}- e_K	$m_K = (e_{K-1} + e_K)/2$	f_K

$$\sum_{i=1}^{K} f_i = n$$

Modal class: Class corresponding to the maximum frequency.

Mode for Grouped Data

$$\overline{x}_{mo} = e_l + \frac{f_0 + f_{-1}}{(f_0 - f_1) + (f_0 - f_{-1})} d_l$$

e_l : lower limit of modal class

d_l : class width

f_0 : frequency of modal class

f_{-1} : frequency of the class just before the modal class

f_1 : frequency of the class just after the modal class

Mode for Grouped Data

Example

The time (in minutes) taken by a customer to arrive in a shop in a month on different days are recorded as follows:

Day	1	2	3	4	5	6	7	8	9	10	
No. of minutes	30	31	30	30	29	29	29	29	29	28	
Day	11	12	13	14	15	16	17	18	19	20	
No. of minutes	28	28	27	27	27	26	26	26	26	25	
Day	21	22	23	24	25	26	27	28	29	30	31
No. of minutes	25	25	25	25	25	24	24	23	22	21	21

Mode for Grouped Data

Example: Considering the data as grouped data, we can present the data as

Class intervals	Mid point (x_i)	Absolute frequency (f_i)
15 – 20	17.5	$f_1 = 0$
20 – 25	22.5	$f_2 = 12$
25 – 30	27.5	$f_3 = 18$
30 – 35	32.5	$f_4 = 1$
35 – 40	37.5	$f_5 = 0$

Modal class: Class corresponding to the maximum frequency.

l = 3 : 25 – 30

Mode for Grouped Data

Example:

e_l = 25 : lower limit of modal class

d_l = 5 : class width

f_0 = 18 : frequency of modal class

f_{-1} = 12 frequency of the class just before the modal class

f_1 = 1: frequency of the class just after the modal class

$$\bar{x}_{mo} = e_l + \frac{f_0 + f_{-1}}{(f_0 - f_1) + (f_0 - f_{-1})} d_l$$

$$= 25 + \frac{18 + 12}{(18 - 1) + (18 - 12)} \times 5 \approx 31.52$$