Exploratory Statistical Data Analysis With R Software (ESDAR)
Swayam Prabha

Lecture 11

Frequency Distribution and Cumulative Distribution Functions

Shalabh

Department of Mathematics and Statistics

Indian Institute of Technology Kanpur

Slides can be downloaded from

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

1

Frequency Distribution

- **Arrangement of ungrouped data in the form of group is called <u>frequency distribution of data</u>.**

- **Classify the data into different classes by dividing the entire range of the values of variables into suitable number of groups called <u>class</u>.**

Frequency Distribution

- Lower and upper boundary figures of a class are called the

  <u>lower limit and upper limit</u> respectively.

- Difference between the limits is called the <u>width</u> of the class or

  class interval.

- The value of variate lies in the middle of lower and upper limits.

Frequency Distribution

- The number of observations in a particular class is called **absolute frequency** or frequency.

- The number of observations in a particular class divided by total frequency is called **relative frequency**.

Frequency Distribution

- The <u>**cumulative frequency**</u> corresponding to any variate value is the number of observations less than or equal to that value.

- The <u>**cumulative frequency**</u> corresponding to a class is the total number of observations less than or equal to the upper limit of the class.

Frequency Distribution

Example:

Following are the time taken (in seconds) by 20 participants in a race.

32, 35, 45, 83, 74, 55, 68, 38, 35, 55, 66, 65, 42, 68, 72, 84, 67, 36, 42, 58.

The data is summarized in class intervals

31-40,   41-50,   51-60,   61-70,   71-80 and   81-90

Frequency Distribution

Example (contd.):

Class intervals	Mid point	Absolute frequency (or frequency)	Relative Frequency	Cumulative Frequency
31 – 40	35.5	5	5/20 = 0.25	5
41 – 50	45.5	3	3/20 = 0.15	5+3 = 8
51 – 60	55.5	3	3/20 = 0.15	5+3+3 = 11
61 – 70	65.5	5	5/20 = 0.25	5+3+3+5 = 16
71 – 80	75.5	2	2/20 = 0.01	5+3+3+5+2 = 18
81 - 90	85.5	2	2/20 = 0.01	5+3+3+5+2+2 = 20
	Total	20	1	

Frequency Distribution

General, if there are k class intervals, n observations are divided into k class intervals $a_1, a_2, ..., a_k$ containing $n_1, n_2, ..., n_k$ observations respectively.

Relative frequency of jth class : $f_j = n_j/n$

Frequency distribution:

Class interval (a_j)	a_1	a_2	...	a_k
Absolute frequency (n_j)	n_1	n_2	...	n_k
Relative frequency (f_j)	f_1	f_2	...	f_k

Frequency Distribution

Class interval (a_j)	a_1	a_2	...	a_k
Absolute frequency (n_j)	n_1	n_2	...	n_k
Relative frequency (f_j)	f_1	f_2	...	f_k
Cumulative frequency	n_1	$n_1 + n_2$	...	$n_1 + n_2 + ... + n_k$
Cumulative frequency	f_1	$f_1 + f_2$	...	$f_1 + f_2 + ... + f_k = 1$

Frequency Distribution

Example (contd.):

Class intervals	Interpretation	Cumulative Frequency
31 – 40	Less than 41	5
41 – 50	Less than 51	5+3 = 8
51 – 60	Less than 61	5+3+3 = 11
61 – 70	Less than 71	5+3+3+5 = 16
71 – 80	Less than 81	5+3+3+5+2 = 18
81 - 90	Less than 91	5+3+3+5+2+2 = 20

Frequency Distribution

Example (contd.):

Interpretation:

Class intervals	Interpretation	Cumulative Frequency
31 – 40	Less than 41	5
41 – 50	Less than 51	5+3 = 8
51 – 60	Less than 61	5+3+3 = 11
61 – 70	Less than 71	5+3+3+5 = 16
71 – 80	Less than 81	5+3+3+5+2 = 18
81 - 90	Less than 91	5+3+3+5+2+2 = 20