

Statistics

Lecture 14: | Histogram for continuous Data

Prof. Dr. Mostafa Elhosseini

<https://youtube.com/drmelhosseini>

| Histogram for Continuous Data

- **subdividing the measurement axis into a suitable number of class intervals or classes, such that each observation is contained in exactly one class**
- One potential **difficulty** is that occasionally an **observation lies on a class boundary** so therefore does not fall in exactly one interval
- an observation on a boundary is placed in the interval to the right of the boundary
- 10 falls in the class $10- < 20$ rather than in the class $0- < 10$

| Histogram for Continuous Data

- Determine the **frequency** and **relative frequency** for each class.
- **Mark the class boundaries on a horizontal measurement axis.**
- **Above each class interval, draw a rectangle** whose height is the corresponding **relative frequency** (or **frequency**).

| Example

- 1, 3, 27, 32, 5, 63, 26, 25, 18, 16, 4, 45, 29, 19, 22, 51, 58, 9, 42, 6

Bucket	Frequency
0 – 9	6
10 – 19	3
20 – 29	5
30 – 39	1
40 - 49	2
50 - 59	2
60 - 69	1

| Example

Bucket	Frequency
0 – 9	6
10 – 19	3
20 – 29	5
30 – 39	1
40 - 49	2
50 - 59	2
60 - 69	1

Bucket	Frequency
0 – 9 0-<10	6
10 – 19 10-<20	3
20 – 29 20-<30	5
30 – 39 30-<40	1
40 – 49 40-<50	2
50 – 59 50-<60	2
60 – 69 60-<70	1

| Histogram for Discrete Data

- In general, it is good to have more intervals rather than fewer, but it is also good to have large numbers of sample points in the intervals. Striking the **proper balance** is a matter of judgment and of **trial and error**
- There is no hard-and-fast rule as to how to choose the endpoints of the **class intervals**
- **Density** presents the **relative frequency divided by the class width**
- When the **classes are of equal width**, the frequencies, relative frequencies, and densities are **proportional to one another**

| Histogram for Discrete Data

- In some cases, **histograms are drawn** with class intervals of **differing widths**. This may be done when it is desired for the histogram to have a **smoother appearance**, or when the **data come** in the form of a frequency table in which the classes have **unequal widths**
- When the **class intervals are of unequal widths**, the heights of the rectangles must be set equal to the **densities**. The **areas** of the rectangles will then be the **relative frequencies**.
- Examples power forecast Devore page 16 BTU
- How to implement with Python