

Statistics

Lecture 18: | Unequal Bucket Width

Prof. Dr. Mostafa Elhosseini

<https://youtube.com/drmelhosseini>

| Unequal Bucket Width

- 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

| Unequal Bucket Width

Class Interval (g/gal)	Frequency	Relative Frequency	Density
1-< 3	12	0.1935	0.0968
3-< 5	11	0.1774	0.0887
5-< 7	18	0.2903	0.1452
7-< 9	9	0.1452	0.0726
9-< 11	5	0.0806	0.0403
11-< 13	1	0.0161	0.0081
13-< 15	2	0.0323	0.0161
15-< 17	0	0.0000	0.0000
17-< 19	2	0.0323	0.0161
19-< 21	1	0.0161	0.0081
21-< 23	0	0.0000	0.0000
23-< 25	1	0.0161	0.0081

Class Interval (g/gal)	Frequency	Relative Frequency	Density
1-< 3	12	0.1935	0.0968
3-< 5	11	0.1774	0.0887
5-< 7	18	0.2903	0.1452
7-< 9	9	0.1452	0.0726
9-< 11	5	0.0806	0.0403
11-< 15	3	0.0484	0.0121
15-< 25	4	0.0645	0.0065

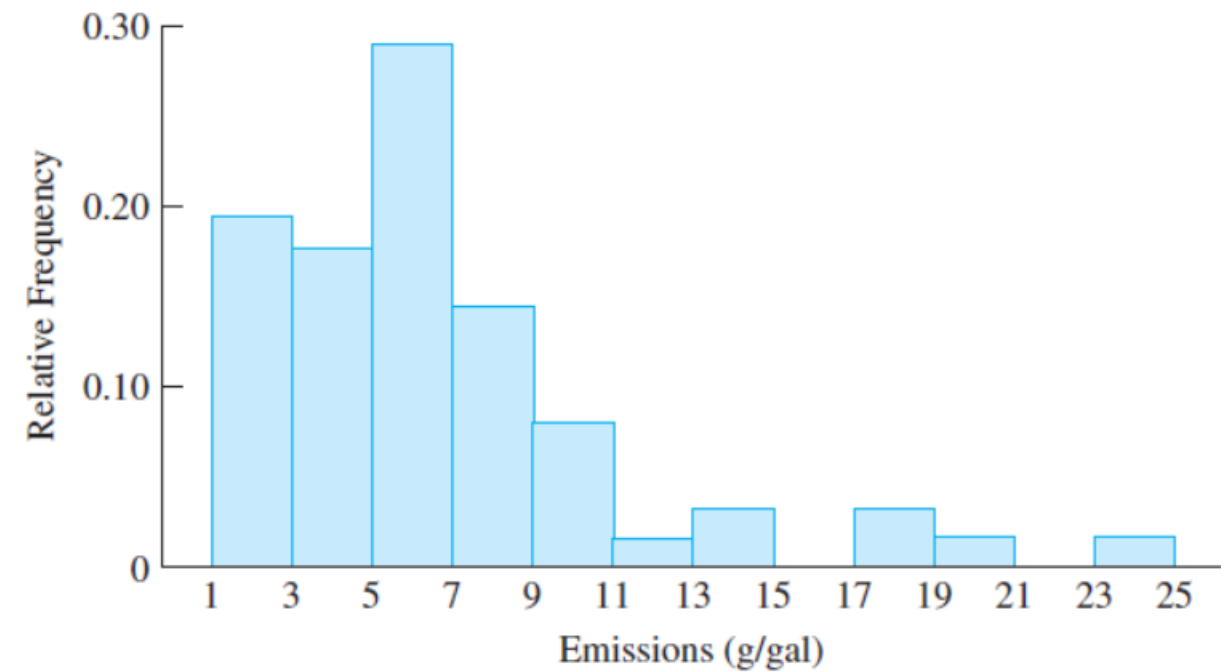
| Unequal Bucket Width

- It is important to note that **because the class widths vary in size, the densities are no longer proportional to the relative frequencies.**
- Instead, the **densities adjust the relative frequency** for the width of the class
- **wider classes tend to contain more sample items** than the narrower classes, and thus tend to have larger relative frequencies
- **Dividing the relative frequency by the class width** to obtain the density adjusts for this tendency

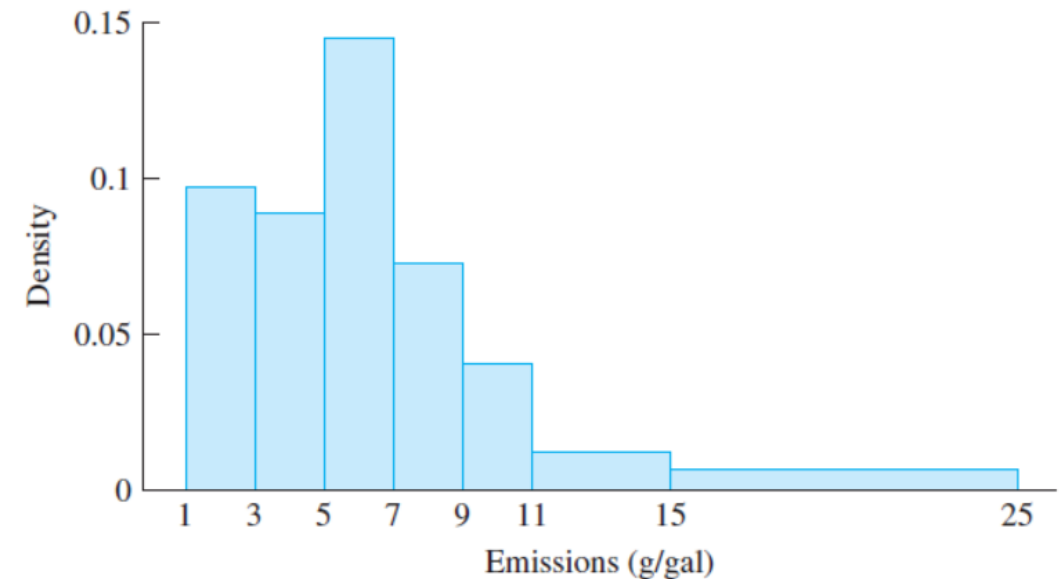
| Unequal Bucket Width

- **When the classes have unequal widths, the heights of the rectangles must be set equal to the densities. The areas of the rectangles then represent the relative frequencies.**

Class Interval (g/gal)	Frequency	Relative Frequency	Density
1-< 3	12	0.1935	0.0968
3-< 5	11	0.1774	0.0887
5-< 7	18	0.2903	0.1452
7-< 9	9	0.1452	0.0726
9-< 11	5	0.0806	0.0403
11-< 13	1	0.0161	0.0081
13-< 15	2	0.0323	0.0161
15-< 17	0	0.0000	0.0000
17-< 19	2	0.0323	0.0161
19-< 21	1	0.0161	0.0081
21-< 23	0	0.0000	0.0000
23-< 25	1	0.0161	0.0081

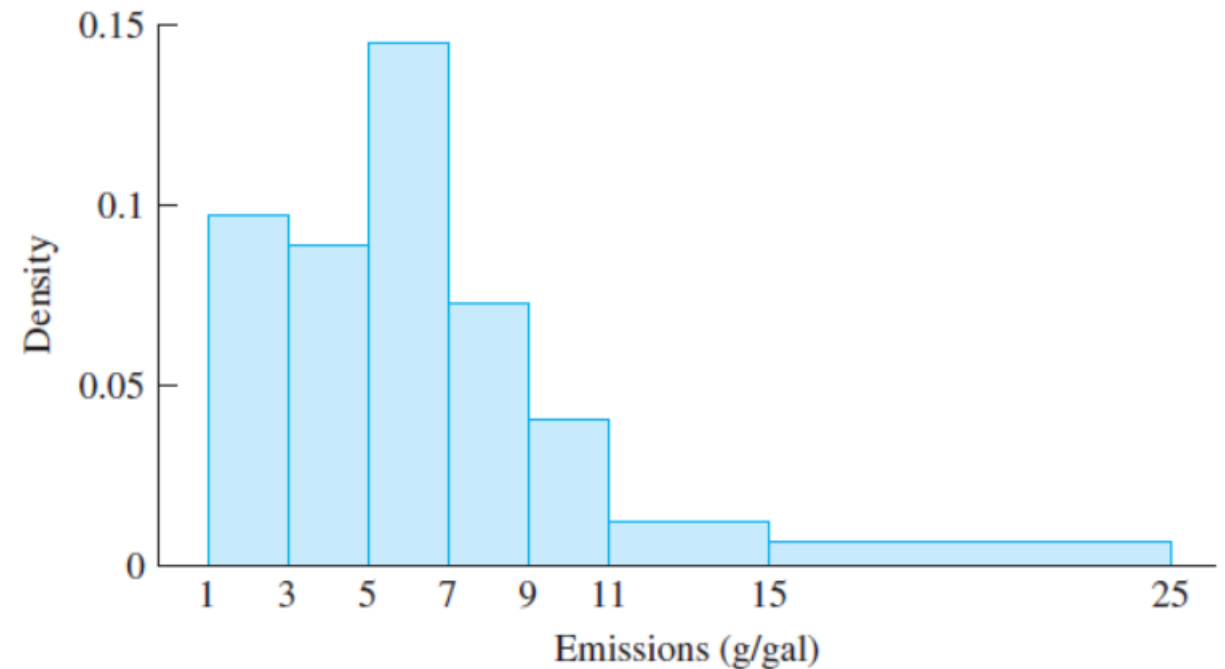


Class Interval (g/gal)	Frequency	Relative Frequency	Density
1-< 3	12	0.1935	0.0968
3-< 5	11	0.1774	0.0887
5-< 7	18	0.2903	0.1452
7-< 9	9	0.1452	0.0726
9-< 11	5	0.0806	0.0403
11-< 15	3	0.0484	0.0121
15-< 25	4	0.0645	0.0065



Use the next histogram to determine the proportion of the vehicles in the sample with emissions between 9 and 15 g/gal.

Class Interval (g/gal)	Frequency	Relative Frequency	Density
1–< 3	12	0.1935	0.0968
3–< 5	11	0.1774	0.0887
5–< 7	18	0.2903	0.1452
7–< 9	9	0.1452	0.0726
9–< 11	5	0.0806	0.0403
11–< 15	3	0.0484	0.0121
15–< 25	4	0.0645	0.0065



Summary

To construct a histogram:

- Choose boundary points for the class intervals.
- Compute the frequency and relative frequency for each class. (Relative frequency is optional if the classes all have the same width.)
- Compute the density for each class, according to the formula

$$\text{Density} = \frac{\text{Relative Frequency}}{\text{Class Width}}$$

(This step is optional if the classes all have the same width.)

- Draw a rectangle for each class. If the classes all have the same width, the heights of the rectangles may be set equal to the frequencies, the relative frequencies, or the densities. If the classes do not all have the same width, the heights of the rectangles must be set equal to the densities.