Today, we will:

- Do a couple review example problems – basic statistics
- Review the pdf module: **Histograms** and do some examples
- Review the pdf module: **Probability Density Functions (PDFs)** and do some examples

Example: Basic statistics

Given: The true (exact) speed in a wind tunnel is 44.63 m/s. Ten velocity readings are taken. The sample mean is 44.580 m/s:

Original data	Data in increasing order
44.54	44.53
44.62	44.54
44.57	44.56
44.56	44.57
44.61	44.57
44.53	44.58
44.57	44.59
44.59	44.61
44.63	44.62
44.58	44.63

mode

median → 44.57, 44.57 } median = avg. of these two

44.57
+ 44.58
89.15 / 2 = 44.575

To do: Calculate the following, giving your answers to the appropriate number of significant digits:

(a) The systematic (or bias) error of the instrument, based on these readings.
(b) The sample median
(c) The sample mode

Solution: (a) $\bar{x}$ = sample mean = 445.80 / 10 = 44.580 [also = average value]

Systematic error = bias error = average value − true value
= 44.580 − 44.63 = −0.050

Systematic error = −0.05

(b) sample median = half-way point. I re-ordered the data in increasing order. See above ↑

- Since there are 10 values, the median is the avg. of the two middle values
- median = 44.575 ≃ 44.8 after rounding

(either answer is acceptable here) → median = 44.575 or 44.58

(c) Sample mode → 44.57 occurs twice. All others are not repeated.

So, mode = 44.57

[If no repeats, mode = undefined]
[If more than one repeat, take value closest to median]

[See also Excel spreadsheet on the website for this same problem]

Example: Basic statistics

Given: Ten houses are sold in the State College area during a particular time period. The selling prices are listed (in increasing order), rounded to the nearest 500 dollars.

1	$242,000
2	$253,000
3	$264,000
4	$267,500
5	$269,000
6	$278,000
7	$286,500
8	$299,000
9	$327,000
10	$748,000

This one house is a mansion! It skews the mean.

median = avg. of these two

Note: This decimal place is a partial significant digit, either 0 or 5

To do: Calculate the mean, mode, and median, and discuss.

Solution:

Mean = (sum of all values)/(number of values) = $3,234,000/10

= $323,400

Mean = $323,400

Median:

$$\begin{array}{r} 269,000 \\ +\ 278,000 \\ \hline 547,000 \end{array}$$

median = $547,000/2

median = $273,500

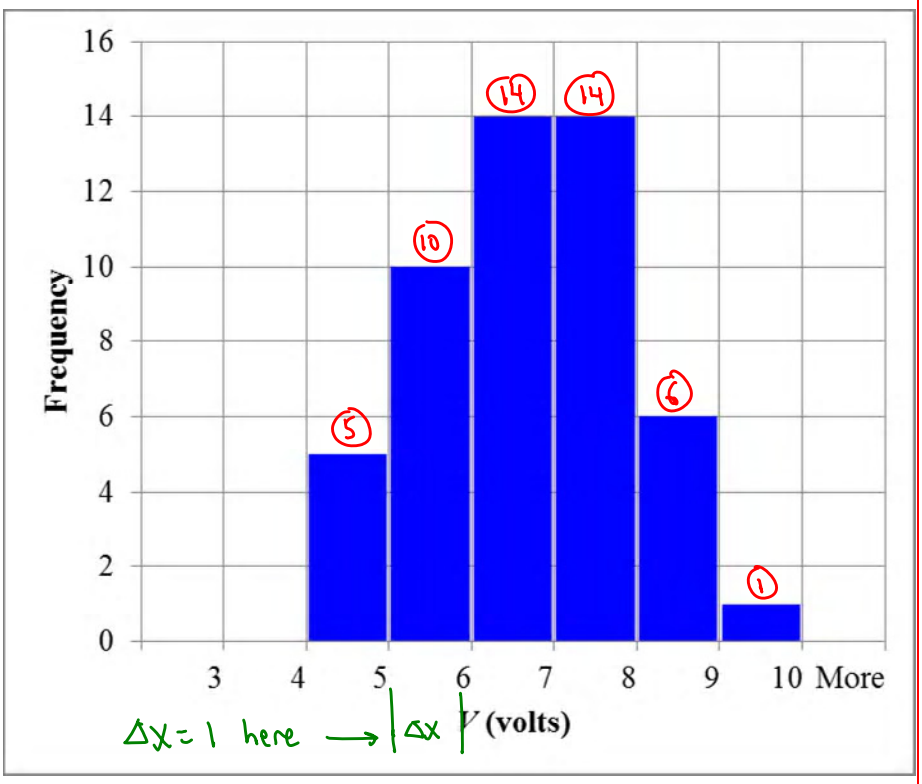
NOTE: MEDIAN IS A BETTER INDICATOR THAN MEAN, SINCE HALF THE HOMES ARE LESS EXPENSIVE & HALF ARE MORE EXPENSIVE THAN THE MEDIAN

[NOTE: ONLY 2 HOUSES ARE MORE EXPENSIVE THAN THE MEAN!]

Mode: No repeats, so Mode = Undefined

Example: Histograms

Given: The histogram shown here (not normalized), produced in Excel from 50 voltage measurements.



To do:

(*a*) How many data points have a voltage less than or equal to 6 V?

(*b*) How many data points have a voltage that lies between 6 V and 8 V?

(*c*) How many data points have a voltage that lies between 3 V and 10 V?

(*d*) What is the *probability* (in percent) that a given reading lies between 5 V and 6 V?

(*e*) When we transform the vertical axis from frequency (number of data points) to $f(x)$ (vertically normalized histogram), what is the value of $f(x)$ for the bin between 5 and 6 on the horizontal axis?

(*f*) The sample mean is 6.76 and the sample standard deviation is 1.31. When we transform the vertically normalized histogram $f(x)$ into a PDF, and then into a normalized PDF, what is the value of $f(z)$ for $x = 8.5$?

Solution:

(a) 5 pts + 10 pts = **15 pts** between $-\infty$ and 6 (first two bins)

(b) 14 pts + 14 pts = **28 pts** between 6 and 8 (middle two bins)

(c) All of them! **50 pts** between 3 & 10

(d) For $5 < x \le 6$, frequency = 10 (second bin). Probability = $\frac{\text{Freq}}{n} = \frac{10}{50} = 0.20$

So, probability that x lies between 5 & 6 is 20% → $P(5 < x \le 6) = 20\%$

(e) $f(x) = \frac{\text{Prob}}{\Delta x} = \frac{\text{Frequency}}{n \cdot \Delta x} = \frac{10}{50(1)} = 0.20$. Thus, $f(x) = 0.20$ @ $x = 5.5$ V

mid value of the bin

(f) @ $x = 8.5$, $f(x) = \frac{\text{Freq.}}{n \cdot \Delta x} = \frac{6}{50 \cdot (1)} = 0.120$. Transform: $z = \frac{x-\mu}{\sigma} \approx \frac{x-\bar{x}}{s} = \frac{8.5-6.76}{1.31} = 1.33$

So, at $z = 1.33$, $f(z) = s \cdot f(x) \approx s \cdot f(x) = (1.31)(0.120) = 0.1572$ → $f(z) = 0.157$