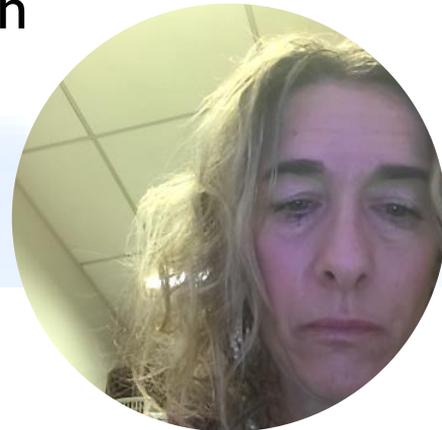


V4 - Descriptive Statistics

Part 3

Course: Statistical Testing & Regression
Dr. Renee Clark
Swanson School of Engineering
Industrial Engineering
University of Pittsburgh



Descriptive Statistics – Part 3

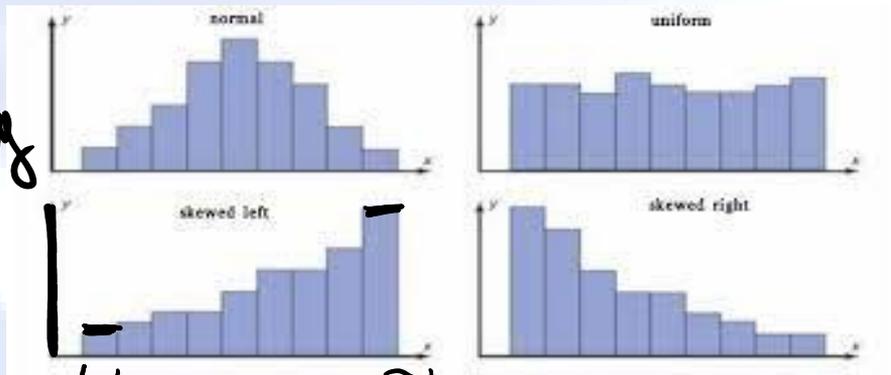
- ❑ Distribution
- ❑ Symmetry
- ❑ Skewness
- ❑ Kurtosis
 - ❑ Descriptive statistics



What is a Data Distribution?

❖ Shape of graph when all possible values of your variable (e.g., height) are plotted on X axis.

❖ How often each value occurs shown on Y axis.



frequency

6' x: height 7'



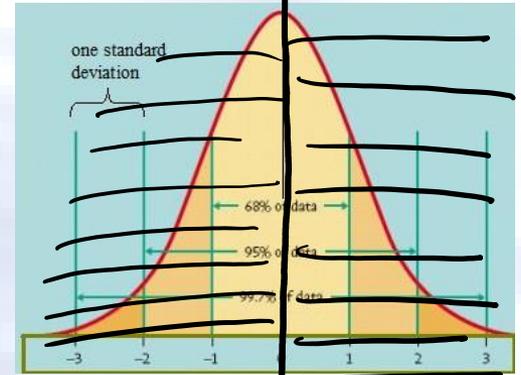
Symmetry

- A distribution is **symmetric** if...

Vertical line through center divides it into 2 halves that are mirror images of each other

- normal distribution is symmetric

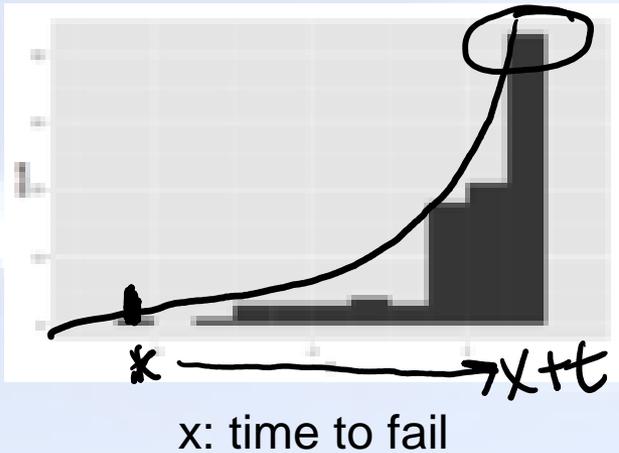
- Half of data is to left of the center line & other half is to the right of this line.



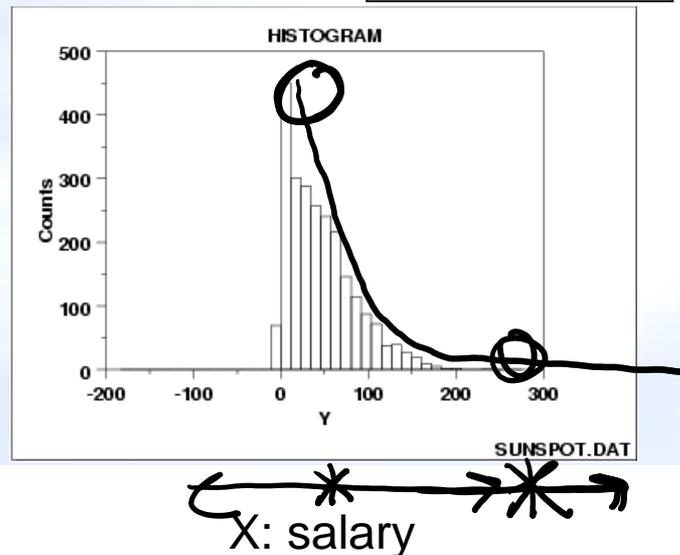
Opposite of Symmetric: Skew

- If not symmetric \rightarrow skewed
 - Skewed to right – long right tail
 - Skewed to left – long left tail

Skewed left

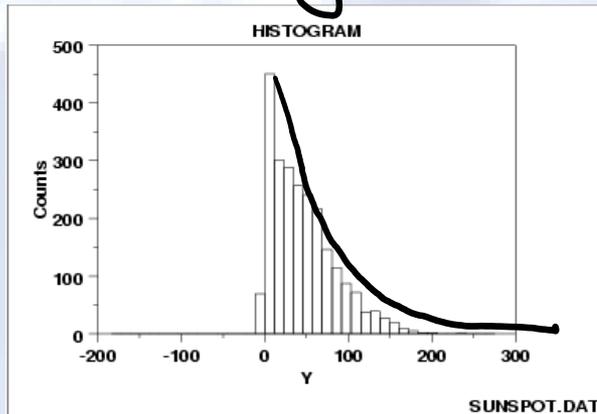


Skewed right



Descriptive Statistics for Shape

Skewness: number that measures lack of symmetry in distribution



Skewed
right
(positive
skew)

- As symmetry increases, skewness \rightarrow 0
- If skewness = 0 then perfectly symmetrical!



Descriptive Statistics for Shape

- Allow software to calculate skewness value
- We interpret using rules of thumb
- **Rules of thumb** (Bulmer, 1979):
 - **Highly skewed:** skew $> +1$ or < -1
 - **Moderately skew:** skew between $(+1/2$ to $+1)$ or $(-1$ to $-1/2)$
 - **Fairly symmetric:** skew between $-1/2$ and $+1/2$

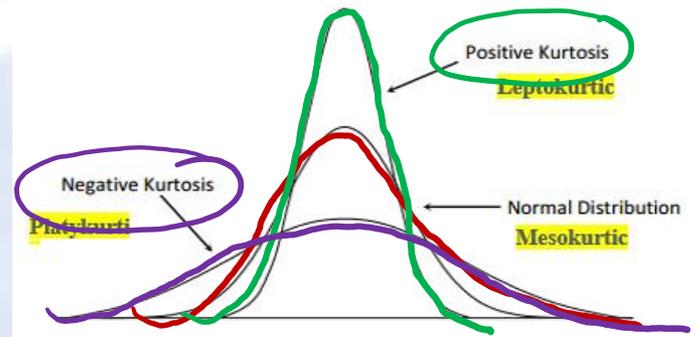


Descriptive Statistics for Shape

Kurtosis: measures "pointy-ness" or "flatness" of distribution

Negative → relatively flat (vs. normal distribution)

Positive → taller/higher peak (vs. normal distribution)



Statistical methods require approx. normal data

❖ For normal data, skewness & kurtosis both 0. Unlikely to occur.

Approximate normality:

kurtosis of +/-1 → very good ✓

+/-2 also acceptable ✓

MTB or EXCEL will calculate ✓





Acknowledgement

This material is based upon work partially supported by the National Science Foundation under Grant# 2335802. Any opinions, findings, and conclusions, or recommendations expressed in this material are those of the author(s) and do not necessarily reflect the views of the National Science Foundation.





You are free

- to **Share** – to copy, distribute, display and perform the work
- to **Remix** – to make derivative works

Under the following conditions

- **Attribution** — You must attribute the work in the manner specified by the author or licensor (but not in any way that suggests that they endorse you or your use of the work).
- **Noncommercial** — You may not use this work for commercial purposes.
- **Share Alike** — If you alter, transform, or build upon work, you may distribute the resulting work only under the same or similar license to this one.





THE END

