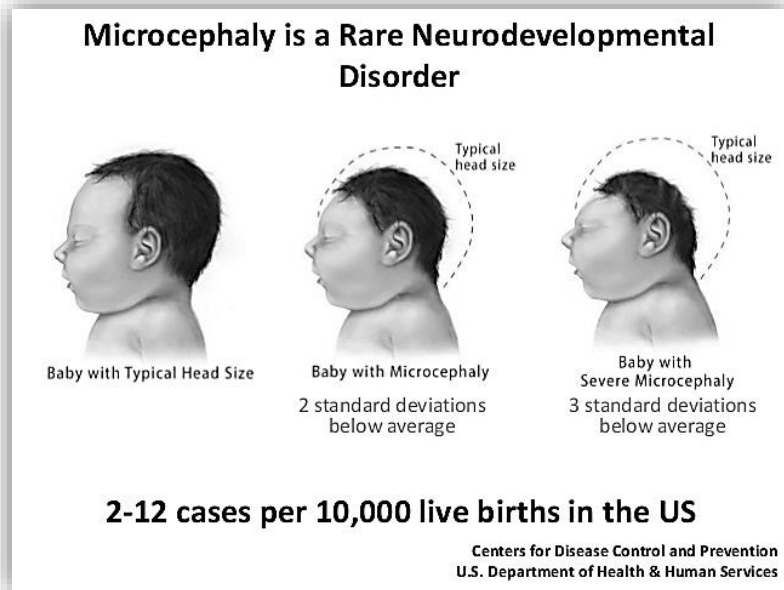
Math 247: What's Unusual? The Empirical Rule and z-scores (Section 3.2)

Physical characteristics that can be measured (height, length of feet, arms, legs, etc.) within populations **are often normally distributed**. This can help, say, a pediatrician monitor a child's growth to see whether the child has unusual measurements that could signal a problem.

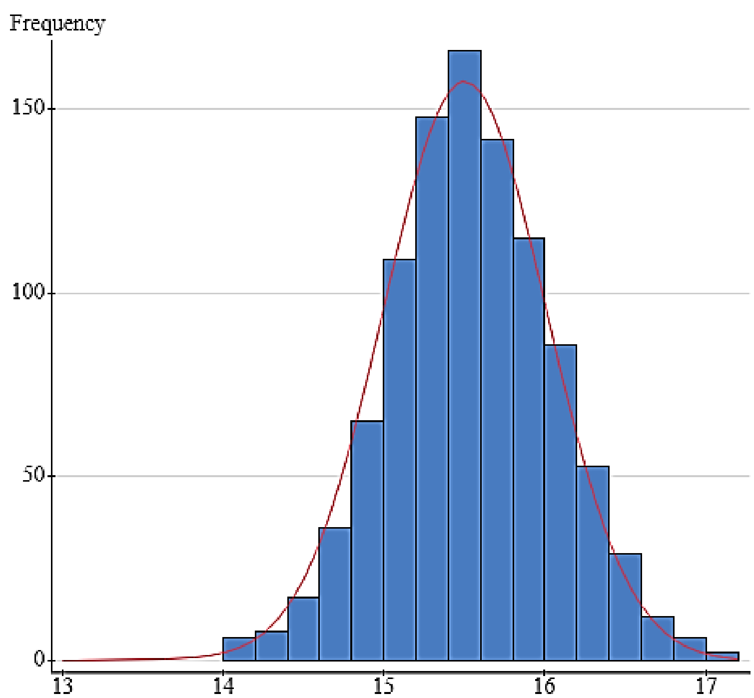
One such problem is "microcephaly" which is a condition where a child's head is much smaller than what is usual. The Zika virus, carried by mosquitoes, can infect a pregnant woman and cause this often devastating birth defect.

How does a physician tell whether a child's head size is unusually small?



Example. Suppose the head size data for 2-month-old baby girls was all gathered together and graphed in a histogram as shown.

What would you guess that the mean and median would be for this distribution?



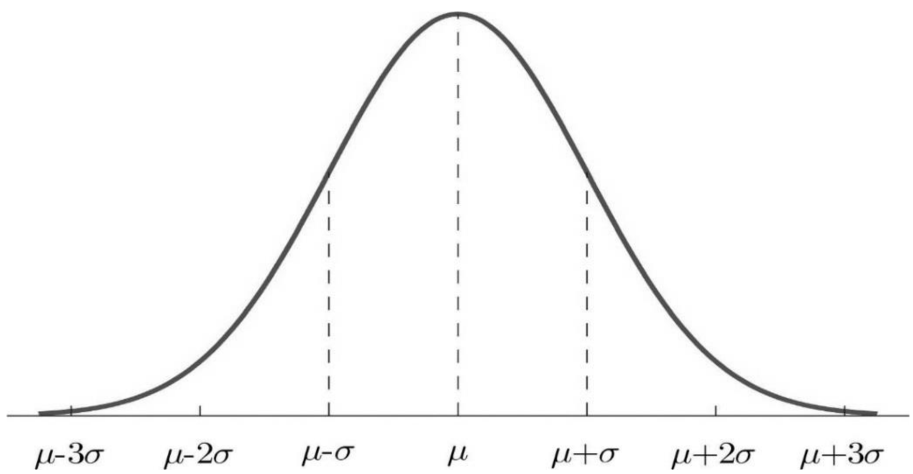
What are common head sizes for these baby girls?

The summary statistics for these data show that the standard deviation is 0.5.

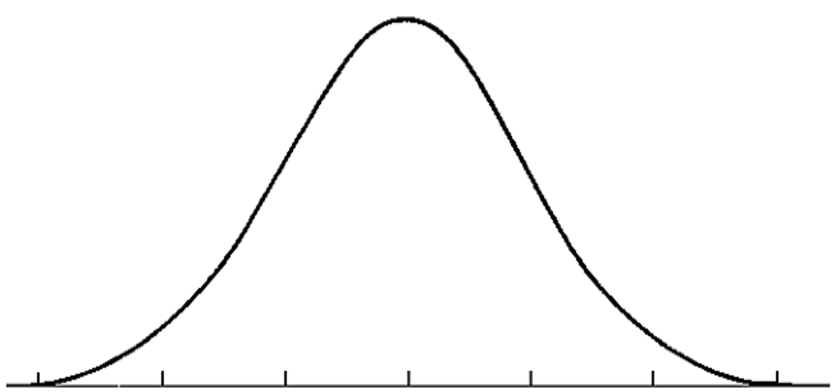
How does this number relate to your answers above; i.e., how could you tell **what's a common or uncommon head size based on just the mean and standard deviation?**

The Empirical Rule: When a population has a distribution that is roughly symmetric and bell-shaped ("normal"), with a population mean of μ and a population standard deviation of σ, then

- approximately 68% of the data will be within 1 standard deviation of the mean.
- approximately 95% of the data will be within 2 standard deviations of the mean.
- 99.7% (all or almost all) of the data will be within 3 standard deviations of the mean.



Example: Assume the distribution of head size (circumference) of 2-month-old baby girls is symmetric (approximately normal) with a mean of 15.5 inches and standard deviation of 0.5 inches. Sketch a well-labeled normal distribution and use it to answer the following questions:



(a) Between what two values would 68% of the head circumferences fall?

(b) Between what two values would 95% of the head circumferences fall?

(c) What head size is the cut-off for the top 2.5% of these baby girls? The bottom 2.5%?

Standardizing measurement: The **z-score** of a data value, x, measures how many standard deviations the data value is from the mean.

z-score Formula: $$z = \frac{\text{observed value - center}}{\text{standard deviation}} = \frac{x - \mu}{\sigma}$$

:

(c) Using the same information as above ($\bar{x}$ = 15.5 inches; s = 0.5 inches), find the z-scores for the following head sizes. Then draw a z-axis under the x-axis and fill in the z-scores.

16 in

16.5 in

15.5 in

15 in

14.5 in

(d) Would a head circumference of 14.9 inches be unusual? Explain.
Plot the x-value and the z-score on the graph.

(e) Would a head size of 13.8 inches be unusual? Explain.
Plot the x-value and the z-score on the graph.

(f) Would a head size with a z-score of 2.63 be above the mean or below the mean? ________________
Would this be an unusually large or small head?

Calculate the head size for this z-score.

Why we standardize measurements…comparing apples to apples.

The CDC commonly uses z-scores since they allow a researcher to compare measurements that come from symmetric distributions with very different means and standard deviations.

> **For instance, as babies grow, the mean head size of the age group will change as will the standard deviation, but the <u>z-score will consistently track</u> how the baby compares to the overall age group.**

Example: Here are two distributions with very different means and standard deviations; nevertheless, we can still compare data values, using z-scores.

Two tests used for admission to Medical School and to an MBA program are the MCAT and GMAT, respectively. The scores on these exams are adjusted so they are normally distributed.

Suppose Nanda, a pre-med major, and Adi, a pre-MBA student, took these exams and were comparing their scores.

The mean MCAT score is 25, with a standard deviation of 5.4. Nanda scored a 36.

The mean GMAT score is 150, with standard deviation of 12.4. Adi scored a 180.

Who did better on their respective test?

Nanda:

Adi:

INTERPRET: