

Please go to Update page and Course page to see information about class and to keep up to date. The links are in main page of Canvas and also on <http://nature-lover.net/math>. You can see old notes from spring 22 etc at this website. It will help you prepare for class.

Quiz 7 Wednesday: Mean, Mode, Median, Std Deviation, Percentile, Normal distribution

Today: More about normal distribution

Review:

MEASURES OF POSITION: PERCENTILES

Arrange data in ascending order, find position, divide by total number, multiply by 100.

Suppose that we take the 7 predicted temperatures,

73 63 58 60 61 69 69

Arranging in ascending order: 58 60 61 63 69 69 73

Percentile score = percentage of data at or below given data point.

63 is in what percentile? $4/7$ or 57.1% at or below, so it is in the 57th percentile.

69 appears twice, so are there 5 or 6 values at or below 69?

Take average of 5 and 6 then we get 5.5

So 69 is in $5.5/7$ or 78.57% or 78th percentile

What are in the 25th percentile?

25th percentile : $25/100 = 1/4 = 0.25$.

Say $x/7 = 0.25$ means $x = 7 \times 0.25 = 1.75$ So only the first number is in 25th percentile. The 2nd number is in $2/7$ or 0.286 or 28.6% from bottom or 28th percentile. If two numbers were both in first position, then their position is 1.5 (average of 1 and 2) and it would also be in 25th percentile.

Cancer distribution:

Article about cancer prevalence in a neighborhood of New Orleans:

(also posted on update page)

<https://www.washingtonpost.com/climate-environment/2022/04/01/new-orleans-gordon-plaza-epa/>

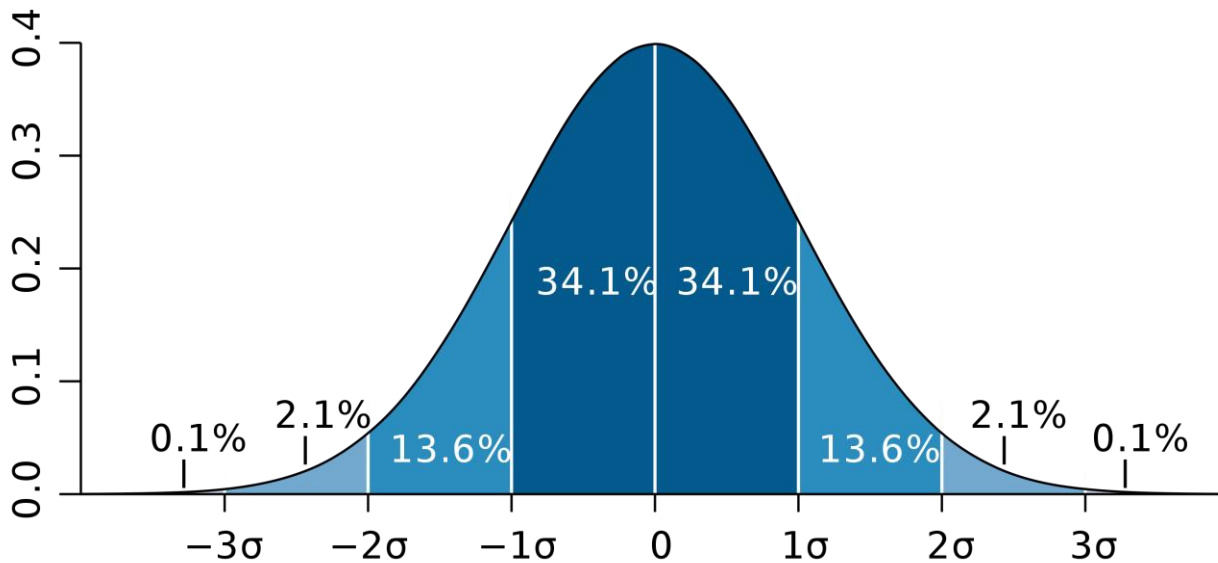
Statistics has been used to check if distribution of cancer in an area is abnormal. If it is distributed according to normal distribution, we can check that.

Description of behavior of normal distribution

The picture below is showing how data is distributed in a normal distribution. The values in the data are along x-axis and they go from negative infinity to positive infinity. It is the graph of $y = \frac{1}{\sqrt{2\pi}} e^{-x^2/2}$

What is the mode? 0 is the mode because its frequency is the highest.

You can also see that it is the median. But it is also the mean!



The area under the curve between two x values tells you “how many” data points are in that range. σ (sigma) stands for standard deviation. Here $\sigma = 1$.

68% within one standard deviation of the average.

95% within two standard deviations.

99.7% within three standard deviations.

Average: 69, std dev 3

Question: What percentage are between 63 and 75 inches?

If heights are between 63 and 75, then 95% would be in that range, because they are $69-6$ and $69+6$ respectively (within two sigma)

What to do if the mean (average) is not zero?

Z scores come to our help.

Remember, the main point is, how many std deviations are you from the mean?

Z-score of the value x is given by $Z(x) = \frac{x-\mu}{\sigma}$ where μ (mu) is the average and σ (sigma) stands for standard deviation.

So for example, 75 is $(75-69)/3 = 2$ standard deviations.

Segue to probability:

Probability is about using statistics to predict the chance of something happening.

So for example, the chance that a random US male is of height between 63 and 75 inches (5'6" and 6'3") is 95%.

Z-scores can be used to find area between any two values, not just whole multiples of std deviations.

Example: Average 30, std dev 5

Z-score of 37 equals $(37-30)/5 = 7/5 = 1.4$. This means there are 1.4 standard deviations separating 37 from the average. Using tables for Z-scores we can find the percentage of the population within 1.4 standard Deviations.

PRACTICE QUESTIONS FOR TODAY

1. If average attendance of a group of students is 90% and the standard deviation is 2% find the percentage of students scoring under 94%. Use the information in these notes.
2. Find the Z-score of a student with attendance 86% in problem 1. What is his percentile rank in the group? In other words, what percentage of students had attendance at or below 86%? Use the graph of normal distribution in the notes above to find the answer.