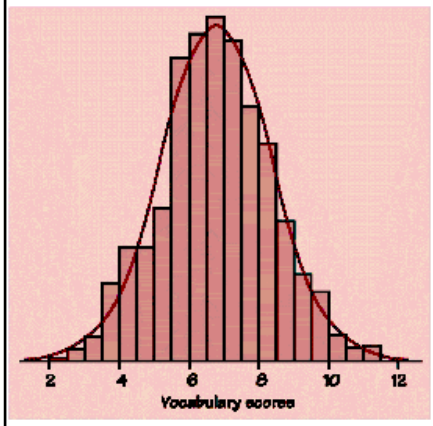
So far:
1. Plot/graph data
2. SOCS CG
3. Numerical Analysis: (5 number summary, IQR, mean, s).

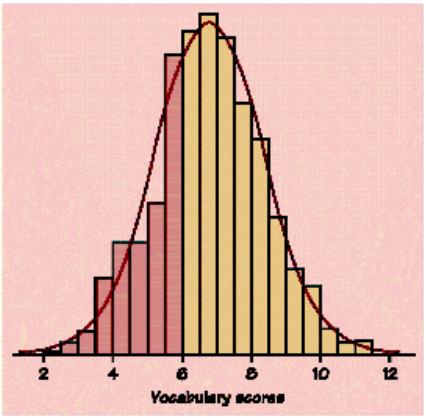
Next: realize that the overall pattern of many observations might so regular that we can describe it by a smooth curve.

Let's explore the idea of a smooth curve fitting a distribution.

http://www.shodor.org/interactivate/activities/normaldist



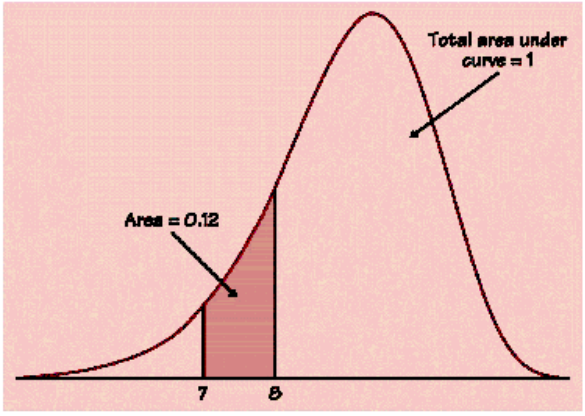
- smooth curve
- mathematical model
- idealized description
- general picture
- ignores minor irregularities
- ignores outliers



- adjust scale
- **total area under curve = 1**
- call this a **DENSITY CURVE**.

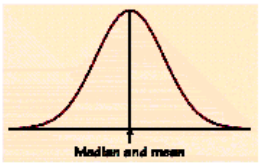
Shaded portion corresponds to % of students who had a vocabulary score from 0 to 6. Actual proportion (287/947 = 0.303) ≈ shaded "area under the curve" (0.293).

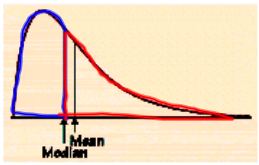
Density Curve:
1. Always on or above the horizontal axis
2. Area under it = 1

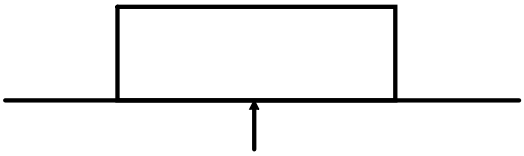


- describes the overall pattern of a distribution
- area under curve and above a range of values = proportion of all observations in that range.

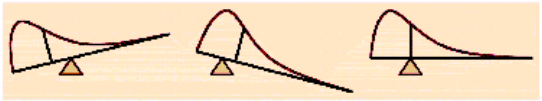
Density curves (like distributions) may be symmetric, skewed, or other shapes.







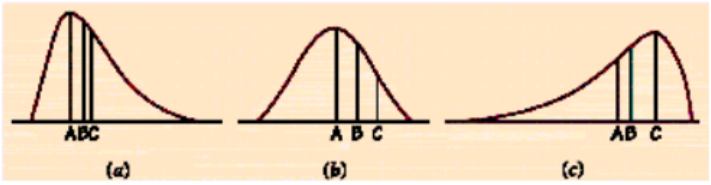
median: equal-areas point
half the area under the curve to its left and half to the right.



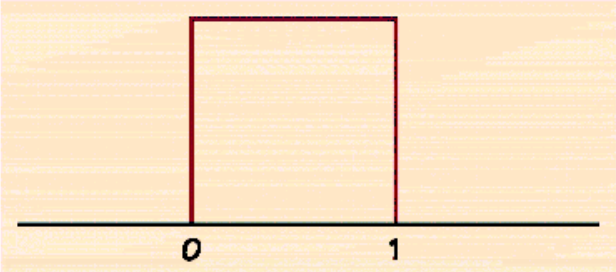
mean: balance point
if made of solid material, the curve would balance

symmetric: median ≈ mean
skewed: mean pulled toward the long tail. "It's mean to pull the dog's tail."

See if you can identify where the mean, median, and mode are located for each of these density curves.



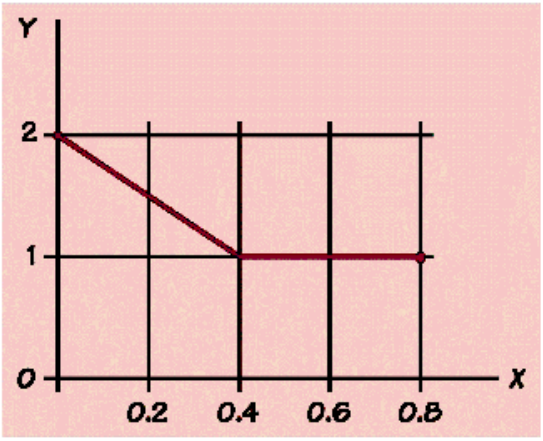
Answer	Answer	Answer



Is this square a density curve? Why or why not?

Verify: 1. On or above x-axis? (yes)
2. Total Area = 1? (yes)

So this is a density curve, a special type called a Uniform Distribution



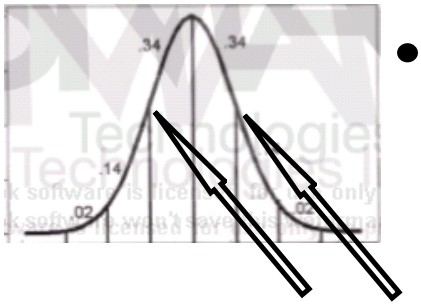
Is this a density curve? Why or why not?

If it is a density curve, explain why the median is a point between X = 0.2 and X = 0.4

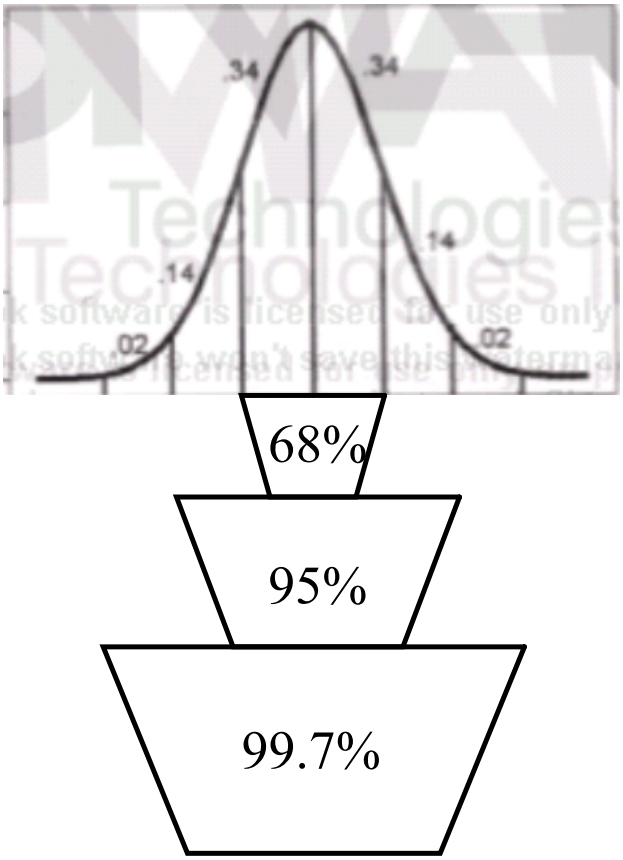
extension

special density curve: the **normal curve**

- symmetric
- single-peaked
- bell-shaped



These points where the normal curve changes from concave up to concave down or from concave down to concave up are called **inflection points** and are one standard deviation on each side of the mean.



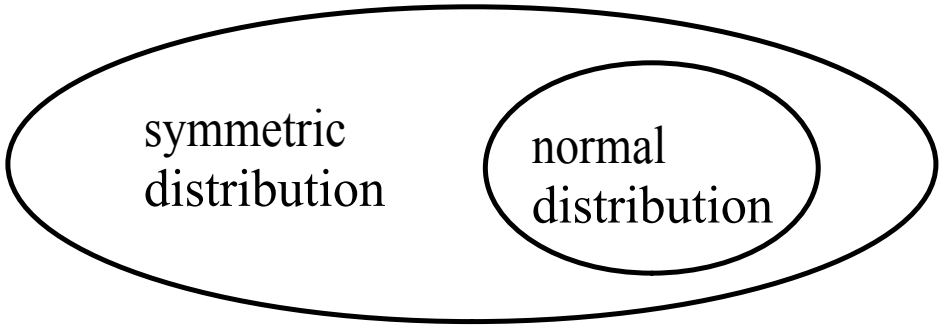
The "**<u>68-95-99.7 rule</u>**" or "**<u>empirical rule</u>**"
In the normal distribution with mean μ and standard deviation σ:

68% of the observations fall within σ of the mean μ.

95% of the observations lie within 2σ of μ.

99.7% of the observations lie within 3σ of μ.

Where are the quartiles?



normal automatically symmetric
symmetric not necessarily normal