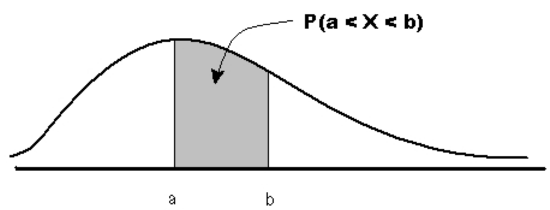
Class #21 - Monday, May 2

In Ch5, we studied *discrete* random variables, with a probability distribution given as a table listing the discrete possible values of X and the probability $P(X = x_i)$ of each value. Such probability distributions were represented graphically as histograms.

Section 6.2: Continuous Random Variables: random variables who set of possible values is an interval of real numbers

- Continuous random variables are typically used when measuring "continuous" quantities such as time, weight, height, length or distance
- It is not possible to list all the possible values of a continuous random variable X, and so we can't make a table for the probability distribution.
- Instead, the probability distribution of a continuous random variable is represented with a curve (i.e., a function), called a **probability density function**
- The probability of an observation of X falling in a given interval (a, b) is the area under the curve over that interval:



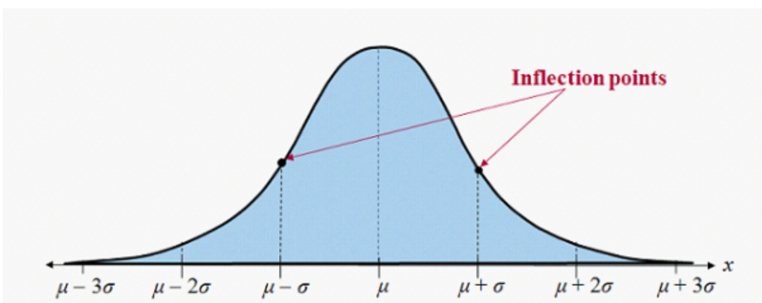
- If you have taken MAT1575 (2nd semester calculus), recall that area under the curve is given by the definite integral, so:

$$P(a < X < b) = \int_a^b f(x)dx$$

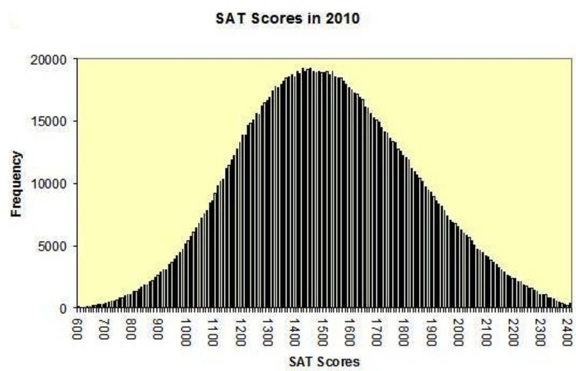
Section 6.3: Normal Random Variables: the most important continuous random variables, which arise naturally in many situations

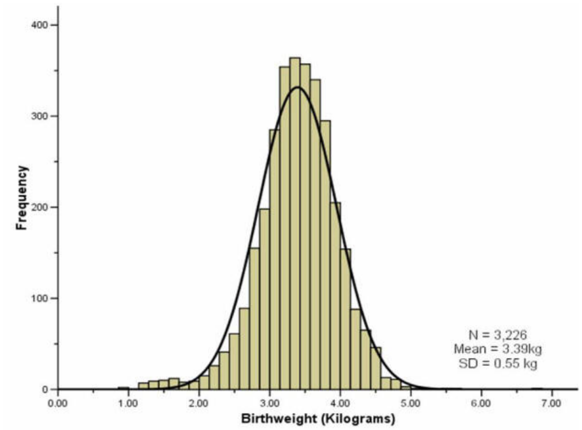
Definition: A **normal (or Gaussian) random variable** is a continuous random variable whose probability density function is determined by its expected value $\mu = E[X]$ and standard deviation $\sigma = \text{SD}(X)$, and which has the following properties:

- It forms a symmetric bell-shaped "normal" curve which is centered about its mean μ
- The total area under the curve is 1 (true for any probability density function)
- The curve extends indefinitely to left and right, i.e., the curve gets closer and closer to, but never touches, the x-axis (we say the distribution has "long tails")
- Between $\mu - \sigma$ & $\mu + \sigma$ (in the center of the curve, "within one standard deviation of the mean") the graph curves downward, and contains approximately 0.68 of the area
- The graph curves upward to the left of $\mu - \sigma$ and to the right of $\mu + \sigma$. These points on the curve are called *inflection points*. Finding the inflection points can help you see how large the standard deviation σ is



Example 1: Many naturally occurring random variables can be approximated (or "modeled") as normal random variables (even though the data usually cannot "extend indefinitely in both directions"; why not?)

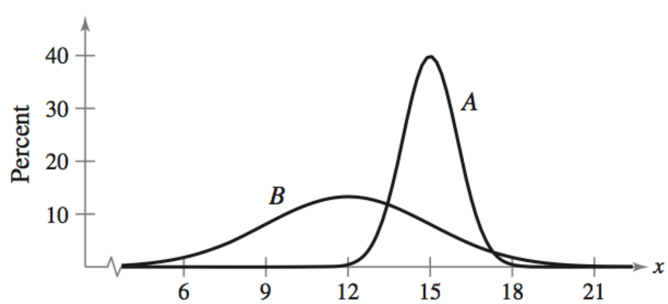




A normal distribution can have any mean and any positive standard deviation. The mean is center of the distribution, and the standard deviation is the distance from the mean to the inflection points (and thus measures how "spread out" the distribution is)

Example 2:

- Which normal distribution has a greater mean? Find μ_A and μ_B.
- Which normal distribution has a greater standard deviation ? Estimate σ_A and σ_B.



Example 3: Estimate the mean and standard deviation of the following normal distributions: