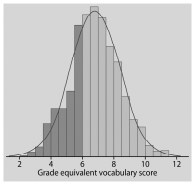
Section 1.3

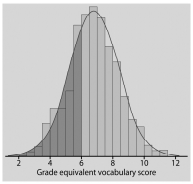
Density Curves

- Mathematical model for the distribution of a variable
- Alternative to graphical and numerical summaries
- Idealized description - histogram of the data is not exactly the same as the curve
- Compact representation of a distribution

Relative Frequency of Scores Less Than or Equal to 6.0



0.303 (using the histogram)



0.293 (using the density curve)

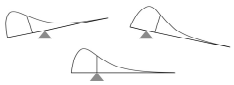
Properties of Density Curves

- A density curve is always on or above the horizontal axis.
- The area underneath a density curve is exactly 1.
- The area under a density curve and above any range of values is the relative frequency of all observations that fall in that range.

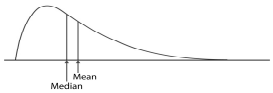
Density curves, like data distributions, can come in many shapes - symmetric, right-skewed, left-skewed.

Observations that are outliers are not described by the density curve.

- The **mode** of a density curve is a peak point of the curve or a location where the curve is highest.
- The **median** of a density curve is the point that divides the area under the curve in half.
- The **mean** of a density curve is the point at which the curve would balance if made out of solid material.



- For a symmetric density curve, the median = mean.
- For a right-skewed density curve, the mean > median.
- For a left-skewed density curve, the mean < median.



We need to distinguish the mean and standard deviation of a density curve (characteristics of the mathematical model) from the mean and standard deviation of the data.

Notation:

Data distribution: mean = sample average ($\bar{x}$), standard deviation (s)

Density curve: mean (μ), standard deviation (σ)