Statistical Moments (and the Shape of Distributions)

The mean and the variance provide information on the location and variability (spread, dispersion) of a set of numbers, and by doing so, provide some information on the appearance of the distribution (for example, as shown by the histogram) of the numbers. The mean and variance are the first two *statistical moments*, and the third and fourth moments also provide information on the shape of the distribution.

For comparison, the

$$\text{First moment} = \sum_{i=1}^{n} (x_i - \overline{X})^1,$$

is by definition is equal to zero. (One might think of the mean as being that value of *x* that makes the above statement true, and consequently indicates where the individual numbers generally lie.

The second moment is recognized as the numerator of the variance

$$\text{Second moment} = \sum_{i=1}^{n} (x_i - \overline{X})^2,$$

which gives information on the spread or scale of the distribution of numbers.

The third moment, $\sum_{i=1}^{n} (x_i - \overline{X})^3$, is used to define the *skewness* of a distribution:

$$\text{Skewness} = \frac{\sum_{i=1}^{n} (x_i - \overline{X})^3}{ns^3}.$$

Skewness is a measure of the symmetry of the shape of a distribution. If a distribution is symmetric, the skewness will be zero. If there is a long tail in the positive directin, skewness will be positive, while if there is a long tail in the negative direction, skewness will be negative.

The fourth moment, $\sum_{i=1}^{n} (x_i - \overline{X})^4$, is used to define the kurtosis of a distribution:

$$Kurtosis = \frac{\sum_{i=1}^{n}(x_i - \bar{X})^4}{ns^4}.$$

Kurtosis is a measure of the flatness or peakedness of a distribution. Flat-looking distributions are referred to as "platykurtic," while peaked distributions are referred to as "leptokurtic."