Variance, Distribution and Standard Deviation[1]
Instructional Primer[2]

The correct specification of the ***variance*** formula reads as follows: $\frac{\sum(E_i - \bar{E})^2}{n}$. That is, the sum of the terms for each of the observations in the sample (population), with the terms being the square of the earnings less the mean earnings for the sample, divided by the number in the sample – this is the variance. If there are 100 observations in the sample (an observation being a household in this case), then this term must be calculated for each of the 100 and then summed, with the summation being divided by the number in the sample. This is the ***variance*** and it is a representation of the absolute value of the breadth of distribution within a sample.

The ***variance*** is simply a statistical term that leads us to the ***standard deviation*** and the ***coefficient of variation***, but it also tells us something of the breadth of the distribution of the sample. The higher the ***variance***, the greater the distribution and the greater the earnings inequality within the sample.

The ***coefficient of variation*** is a representation of the relative breadth of the distribution within a sample and might be preferable the simply using the ***variance***. The ***coefficient of variation*** can be found by taking the square root of the ***variance*** (also referred to as the ***standard deviation***: $\sqrt{variation}$ = *standard deviation*) divided by the mean earnings of the sample: $\frac{\sqrt{variation}}{\bar{E}} = \frac{Standard\ Deviation}{\bar{E}}$ = ***coefficient of variation***. As with the variance, the higher the ***coefficient of variance***, the greater the distribution and the greater the earnings inequality within the sample.

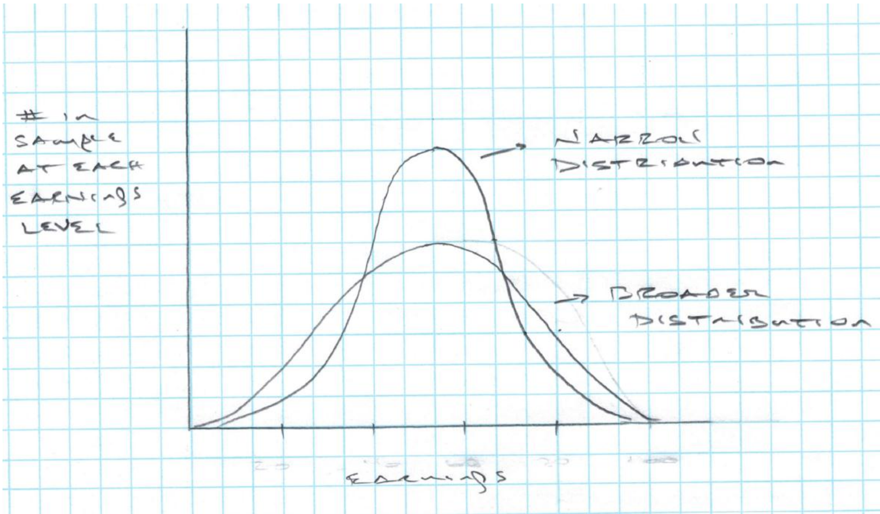
Also, Ryan and I stayed after for a bit and discussed something that might have been helpful to all of us, that is a further representation of normal distributions, their relative shapes and how these might look compared to a distribution representative of earnings inequality.

I represented the following normal distributions and suggested that a narrower distribution would have less dispersion of income than a broader distribution – which is correct.

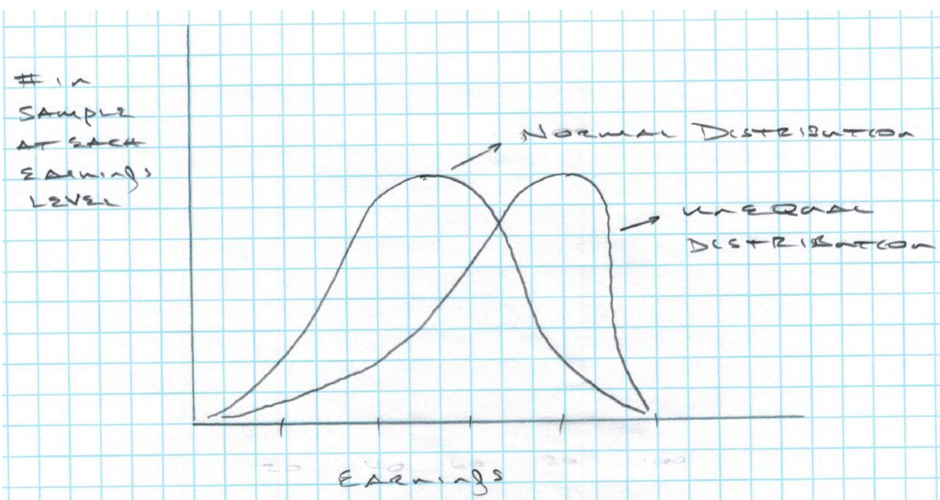
---

[1] This primer is intended to present an abbreviated discussion of the included economic concepts and is not intended to be a full or complete representation of them or the underlying economic foundations from which they are built.

[2] This primer was developed by Rick Haskell (rick.haskell@utah.edu), Ph.D. Student, Department of Economics, College of Social and Behavioral Sciences, The University of Utah, Salt Lake City, Utah (2013)



What I didn't represent as well as might have been helpful to you is what an unequal distribution (one representative of income inequality) might look like. Remember that a normal distribution is equally distributed and has no obvious inequality but an unequal distribution shows to which portion(s) of the distribution the inequality is weighted. The following represents a sample for which the lower quintile has a lesser portion of the aggregate income while the upper quintiles have significantly more (inequality).



Okay, now a little more about ***standard deviations***.

In statistics and probability theory, ***standard deviation*** shows how much variation or dispersion exists from the mean (average). A low ***standard deviation*** indicates that the data in the sample tends to be very close to the mean – it has a lesser dispersion and a more narrow distribution; a high ***standard deviation*** indicates that the data is spread out over a larger range – it has a greater dispersion or a broader distribution.