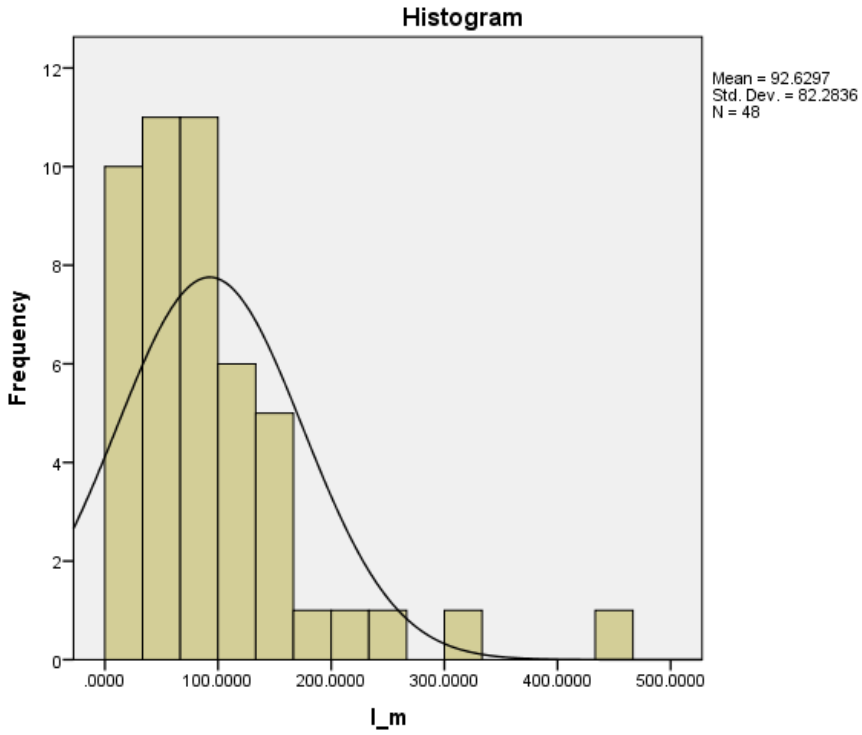
Kernel Density Plots

These plots produce a smooth curve estimating the probability density function of a continuous variable from a set of scores that likely contain some error. See Howell (2013). ***Statistical methods for psychology*** (8th ed.). Belmont, CA: Cengage Wadsworth, pages 22-24, for a simple explanation of these plots. To illustrate the production of these plots, I shall use some of the data from my dissertation. The variable used here is the latency to enter a house-mouse scented tunnel by an adult house mouse when there was also a rat-scented tunnel available.

Descriptives, Frequencies, in SPSS will produce a smooth plot of another sort. It computes the mean and standard deviation of the variable and then superimposes, over a histogram of the scores, what the probability density plot would look like were the variable normally distributed. Here is such a plot for the latency data:

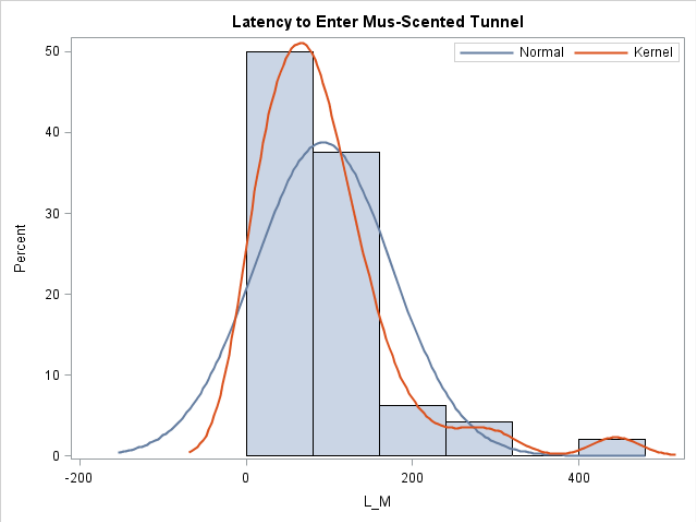


I find plots like this a convenient way to visualize the extent to which the distribution of a variable differs from a normal distribution. This variable differs from normal by a lot, $g_1 = 2.19$.

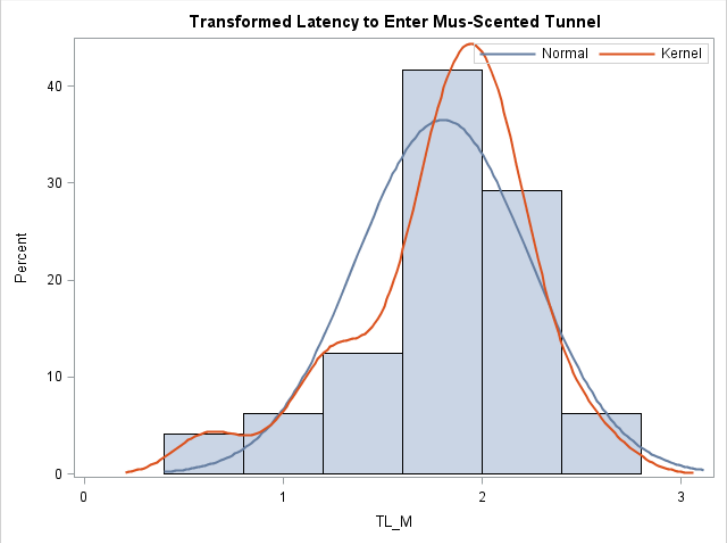
A kernel plot produces a smooth curved without assuming normality. Here is SAS code for producing a histogram with a normal-assumed plot and a kernel plot superimposed.

```
proc sgplot;
  title "Latency to Enter Mus-Scented Tunnel";
  histogram l_m;
density L_m;
  density l_m / type=kernel;
 keylegend / location=inside position=topright;
run;
```

Here is the output:



The kernel plot makes it clear that the distribution is distinctly skewed, not normal. I applied a monotonic but nonlinear transformation to these data to reduce the skewness prior to further analysis. Here is the plot of the transformed data, which had g_1 = -.878 (still skewed, but much less).



Karl L. Wuensch, August, 2016.

- Return to my Stats Lesson Page.
- Do it with R or SPSS