

# Today: 1-D Continuous

Sam Ventura  
36-315

Today: Boxplots, Histograms, and Density Estimates  
Conditional Distributions for Continuous Variables

Department of Statistics  
Carnegie Mellon University

February 13, 2017

# 1-D Continuous Data

Structure:

Summary:

In R:

# 1-D Continuous Distributions

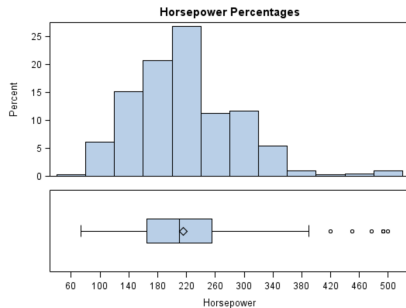
How do we describe continuous distributions?

Do the data appear to fit some common/known distribution:

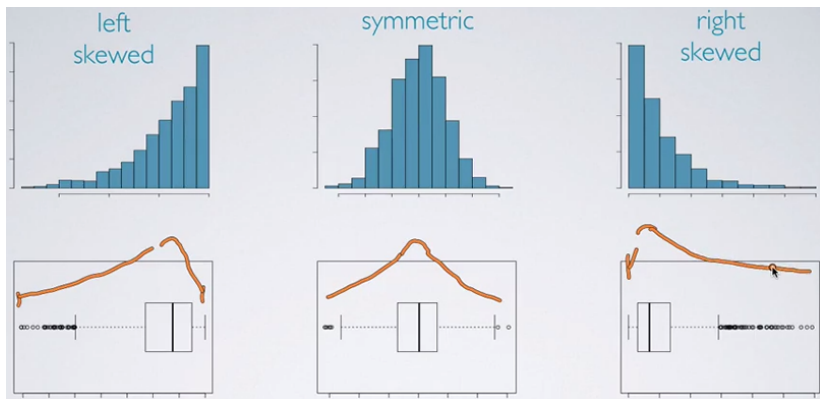
# Visualizing 1-D Continuous Distributions

How do we visualize continuous distributions?

# Box Plots vs. Histograms: Advantages and Disadvantages



# Determining Skew Visually



# How Can We Improve Boxplots?

Standard Boxplot:

Improved Boxplot?

# Continuous Densities

Theoretical:

What if we don't know the underlying distribution?

How can we estimate the empirical distribution?



# 1-D Kernel Density Estimation



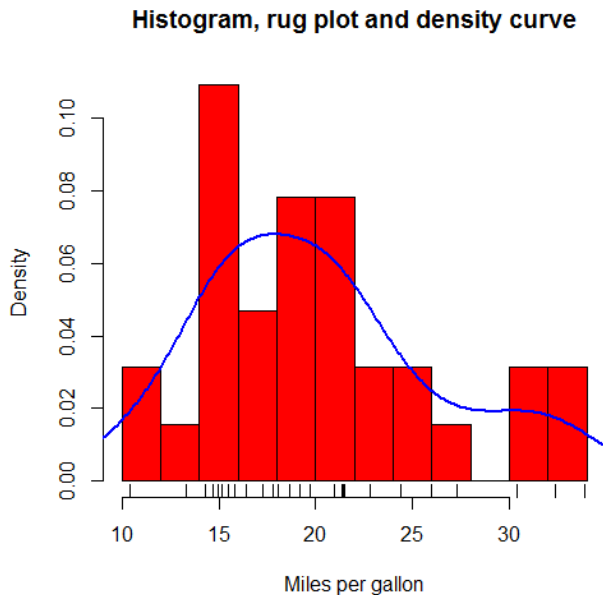
# Comparing 1-D Continuous Densities

**Kolmogorov-Smirnoff (KS) Test:**  
nonparametric test of the equality of  
multiple 1-D continuous distributions

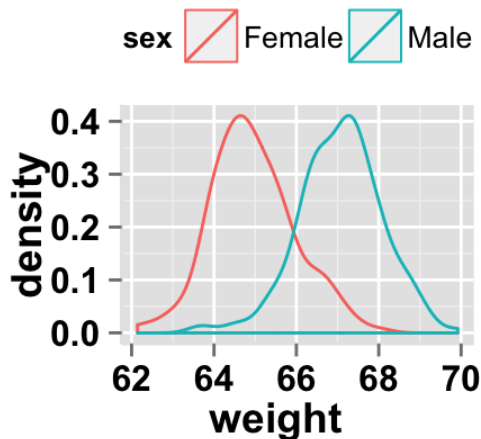
In R: **ks.test**

**Kullback-Liebler Divergence:**  
Measure of the difference between two probability distributions  
(discrete/categorical or continuous)

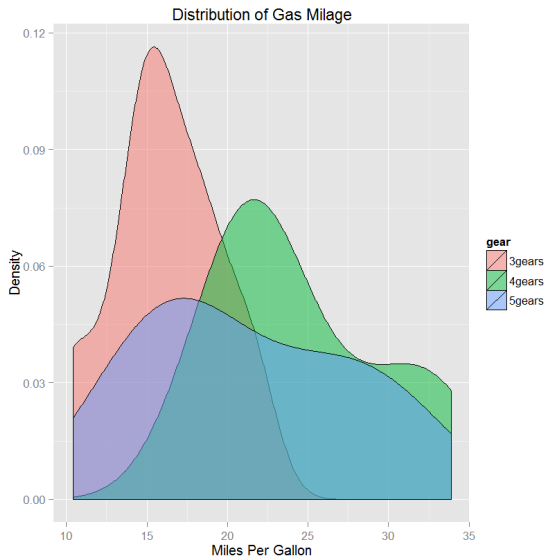
# Histogram with Density Curve and Rug



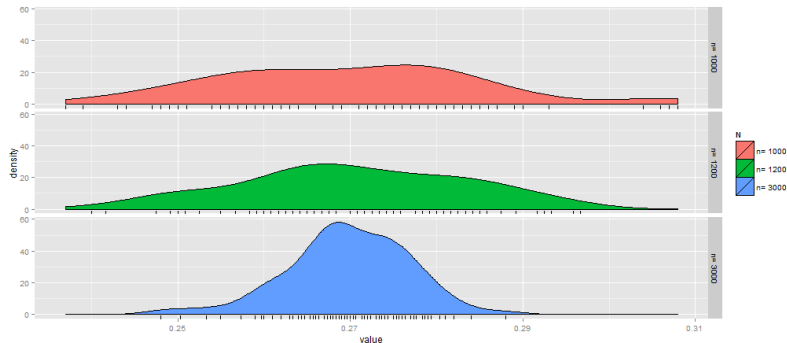
# Density Estimates and Conditional Distributions



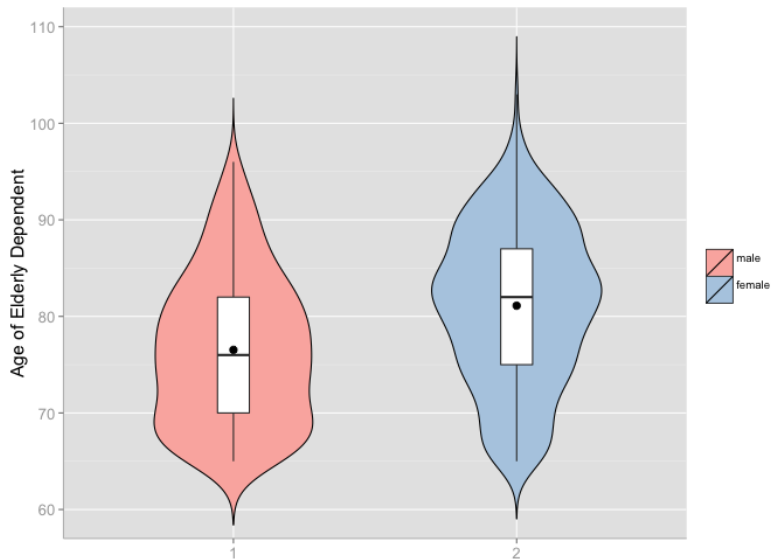
# Density Estimates and Conditional Distributions



# Density Estimates with “Rugs”

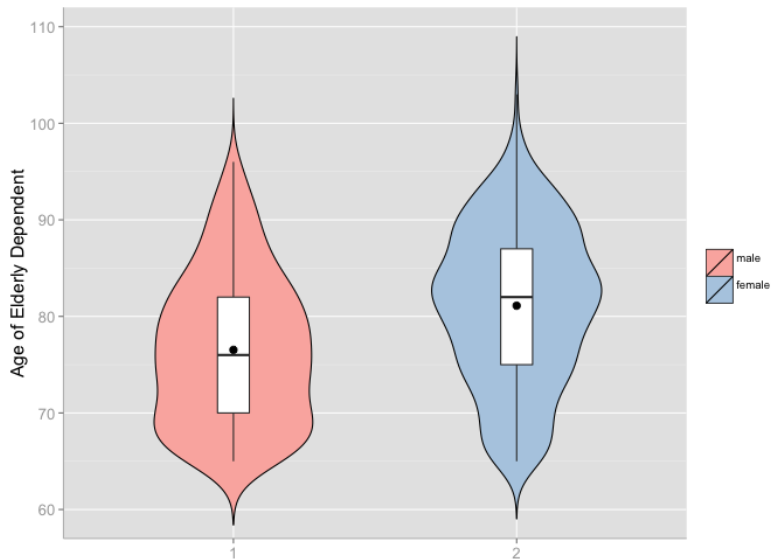


# Violin Plots vs. Boxplots

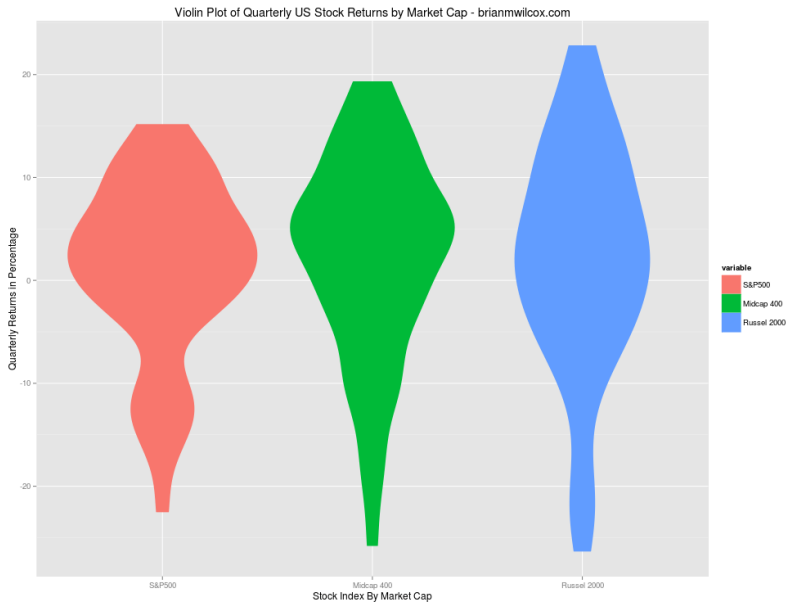




# Issues with Violin Plots?



# Issues with Violin Plots?



# Improved Violin Plot?

