

Math 502. 2nd Homework. Due Friday 25, February, 2005.

Name:

When you do graphs, like histograms, you have to a graph which show the shape of the data. Besides handing me your homework you have to e-mail me your Splus programs.

1. The San Andreas Fault is a geological fault that runs a length of roughly 800 miles (1,300 km) through western and southern California. Most of earthquakes happening in California are in this fault (like the 1905 San Francisco earthquake). For the Quakes Data Set take the variables: "longitude of quake location". and "latitude of quake location". Draw the scaterplot of the two variables. Describe where the bulk of the data is.
2. For the Quakes Data Set take the variable: "magnitude: the size of the quake".
 - (a) Find the five number summary of the data (minimum, first quartile, median, 3rd quartile, maximum). Find the mean of this variable.
 - (b) Draw a relative frequency histogram that you obtain when you divide the data using the cutpoints: 2.5, 3, 3.5, 4, 4.5, 5, 5.5 and 6. Describe the shape of the distribution. Is it symmetric? Is it skewed?
 - (c) Find the boxplot. Find how many observations are outliers according with the rule that says that an observation is an outlier, if it does not fall in the interval

$$(Q_1 - 1.5(Q_3 - Q_1), Q_3 + 1.5(Q_3 - Q_1)).$$