**Computing and Data Analysis in Geosciences**

**Week 3: Error and Uncertainty Analysis.** Do the whole exercise in R. Use “sections” (as Nuss demonstrated last week) to make your script neat and easier to follow. Comment your code appropriately. Use the print() and paste0() functions to display answers. As usual, there is an Excel sheet if you need help!

**Note:** You do not need specific knowledge of volcanology, geochronology, or seismology to answer any of these questions. **You just have to figure out whether your equations are propagating errors in sums, differences, products, or quotients.**

1) The distance to a building is estimated from a map to be about 2550 25 m. With a theodolite, a student determines the angle between the horizontal plane and the building to be 1°21’ 1’. What is the height of the building and the fractional uncertainty in this estimate?

**Hints:** Recall basic trigonometry!

You will need to convert the DMS (degree, minutes, seconds) degrees to decimal degrees, and then to radians. Both the value and the uncertainty needs to be converted.

You should write a function that converts DMS degrees to radians, and use the function to convert both the value and the uncertainty. Pi is a built-in variable in R.

2) From compositional data you infer that two volcanic rock samples represent the initial and final pulse of volcanic activity. Your two samples have been radiometrically dated at 25.53 0.1 Ma and 29.66 0.2 Ma respectively. What is the likely duration of volcanic activity (including the error?)

3) Load the datatable ex3\_eqscals.txt. This is a table of real measurements of earthquake size, at multiple stations, from a single earthquake in California. There are 3 columns,

*X* (km) = the distance of the station from the earthquake

*r* (m) = the estimated fault radius

*Mo* (Nm) = the seismic moment, which is related to earthquake magnitude.

Parameters *r* and *Mo* are estimates made from the seismogram recorded at each station, and they vary.

Note that the csv file has NO HEADERS, and the separator is a tab. So you need to specify these arguments! Tab separators can be indicated by sep = ‘’

a) What is the mean, median, standard deviation, and median absolute deviation (MAD) for *r* and *Mo*?

b) Make one descriptive plot each, of *Mo* and *r* (eg. histograms, boxplots, scatter, logged etc). Are there any obvious outliers that can cause problems?

c) Are there any outliers apparent, either from looking at the numbers or plotting? A good criterion is eliminating points exceeding 3\*MAD from the median. Eliminate them to make a TRIMMED set, and recalculate the mean, median and standard deviation. What is your “best” estimate, and uncertainty, in *Mo*?

d) The “moment” magnitude of an earthquake is calculated from the equation

Mw =

Use your “best” estimate and uncertainty in *Mo* from (c) , to calculate *Mw* and its uncertainty.