Task 2: Descriptive Statistics in Python

Haris Themistocleous

February 2, 2017

1 Task 1

1. What are the main measures that are used to determine the shape of a distribution.
2. What is the difference between measures of central tendency and measures of dispersion.
3. How do you calculate the variance, and the standard deviation, write the formulae.

2 Task 2

The table fricatives contains measurements of fricatives spectrum and there are estimations here of the probability distribution of the spectrum of each fricative in terms of spectral moments. More information about this study exist in the attached paper. The values in the table may not correspond to that of the paper as there were some modifications in the table.

The table of contains the following information:

id an index of the table

duration the duration of consonants

intensity

cog the center of gravity which basically the mean of the spectrum

sdev Spectral SD

skew Spectral Skewness

kurt Spectral Kurtosis

Segment (see task 2)

Vowel two vowels: /i/ and /a/

Variety two varieties: CG and SMG

Stress two stresses: Stressed and Unstressed

Voice the consonant is voiced or voiceless

Position

1. Open the table in Python or R.
2. Find how many unique segments (levels of the variable Segment) exist in the column Segment of the table.
3. Plot histograms of all columns (only those that have numbers inside: duration intensity cog sdev skew kurt).
4. Create boxplots for duration and center of gravity (cog).
5. Find the mean and standard deviation of all columns.
6. Estimate for every segment the mean and standard deviation of the columns: duration intensity cog sdev skew kurt
7. Estimate the mean and standard deviation of duration intensity cog sdev skew kurt for every level i.e., stressed and unstressed of the variable stress.
8. Estimate the mean, mode, range, standard deviation, variance, sum for every level of the variable Segment for intensity.
9. Conduct a t-test and find out whether duration in different in the two levels of the variable Dialect.