Measures of Dispersion for Grouped Data

A. Recall Measures of Dispersion for *Ungrouped* Data

Measures of dispersion or spread for *ungrouped* data describe how far each individual data value has *strayed* from the **mean** (also described as how closely the individual data values *cluster* around the mean). We use the formulas learned in the last lesson when we have all of the **raw data** and can compare each data point to the mean.

But what if we don't have access to the raw data? What if we ONLY have *grouped* data (i.e. a frequency table with intervals and/or midpoints)? The best we can do is ESTIMATE the measures of dispersion by treating the midpoints as though they are the data points. This requires slightly different formulas!

B. Measures of Dispersion for Grouped Data

We will now consider the same measures of spread for grouped data (i.e. data that has been organized into intervals and frequency tables). The formulas below are for standard deviation. To find the variance, simply square the standard deviation!

Population Standard Deviation (Grouped)	**Sample Standard Deviation (Grouped)**
$\sigma \approx \sqrt{\frac{\sum f_i(m_i - \mu)^2}{N}}$	$s \approx \sqrt{\frac{\sum f_i(m_i - \bar{x})^2}{n-1}}$

** Note the denominator of n – 1 for the sample standard deviation. This compensates for the fact that a sample from a population tends to underestimate the deviations in the population.*

Recall that *f* is the frequency for a given interval and *m* is the midpoint of the interval. It should be noted that calculating standard deviations from raw, ungrouped data will give more accurate results, and that measures of dispersion from grouped data are only estimates.

Ex. 1: Use Excel to calculate the standard deviation for the salaries listed below:

Midpoint Salary, ($1000)	28	30	32	43	36	38
Frequency	4	6	7	4	2	1

To find the standard deviation for grouped data:

i) Find the grouped mean using the appropriate formula from Lesson 2.

ii) Find the midpoint deviations.

iii) Square the midpoint deviations.

iv) Multiply the squared midpoint deviations by their frequency, so that the values are weighted appropriately.

v) Find the sum of the weighted midpoint deviations and divide by N or *n* – 1.

vi) Square root the result!

CW: p. 148 #7a, 9a