

When possible, I'll post the lecture notes that I work from, like the notes below, so that you'll have a list of some of the topics covered during the lecture. These may not be complete, and will be very brief, but if you find them useful then feel free to use them to supplement your own notes.

chapter 2. measures of central tendency

- how to summarize and describe data? (descriptive statistics)
 - one goal: measure of central tendency
 - = a single number that reports the centre of a distribution
- mean (average)
 - population mean: parameter, μ
 - sample mean: statistic, \bar{x}
 - add up scores and divide by number of scores
 - write formulas using sigma notation
- median (50th percentile, P_{50})
 - if odd number of data points, use middle data point
 - if even, use average of middle two data points
 - useful when data is skewed, or when there are outliers
 - to find middle data points
 - if n is odd, $n/2$ rounded up
 - if n is even, $(n/2)$ and $(n/2)+1$
- mode (most common value)
- mean, median, and mode are all measures of central tendency
- choice between measures depends on purpose, e.g., mean income (tax collector) vs. median income (sociologist)
- skewed distributions
 - population or sample with many data points at one end of distribution and few at the other (opposite of: symmetric)
 - positively skewed (= right-skewed): tail longer at higher end
 - negatively skewed (= left-skewed): tail longer at lower end
 - for a skewed distribution the mean, median, and mode are usually different
 - mean is pulled up or down, in direction of skew

chapter 3: measures of variability

- mean, median, and mode give information about central tendency
 - but no information about how widely the scores are distributed
- measures of variability, or dispersion
 - = differences between scores in a distribution
 - range = maximum - minimum
 - not robust; i.e., it is sensitive to outliers
 - interquartile range (IQR): 75th percentile - 25th percentile
 - percentile position: $i = (P/100)n$ ("index method")
 - if integer, use average of data points in position i and $i+1$
 - if not integer, round up
 - first quartile = 25th percentile
 - second quartile = 50th percentile = median
 - third quartile = 75th percentile
 - variance: see formulas.pdf
 - population variance (σ^2) and sample variance (s^2)
 - definitional formula and computational formula
 - standard deviation: square root of variance