Stat 145, Fri 10-Sep-2021 -- Fri 10-Sep-2021
Biostatistics
Spring 2021

Friday, September 10th 2021

Wk 2, Fr
Topic:: Standard deviation
Read:: Lock5 2.3
HW:: PS03 due Wed.
HW:: PS04 due Sat.

Administrative:
- settings in RStudio
- escape key use
- reset console to blank screen : Use Ctrl-L

Boxplots revisited
- other names
- 1.5 x IQR criterion for outliers

Measures of spread
- topic only for quantitative data (looking at one variable)
- "spread" can be thought of amount of variability expressed in values
- available measures: range, IQR, standard deviation
- these measures do what they're advertised as doing

```
gf_histogram(~ haircut | sex, data=ssurv)  # note visual difference
favstats(~ haircut | sex, data=ssurv)

Or, work with "data" which has only one value
  roll a die 50 times that with all six faces the same (three dots, perhaps)
    rolls <- rep(3, 50)
    gf_dotplot(~rolls, binwidth=1, dotsize=.02)
    favstats(~rolls)
```

- focus on standard deviation and variance
  formula
  example calculation

[redacted]

Not yet covered

- parameter labeled sigma, s is used for sample
- same units as that of quantitative variable

- blurring a histogram as number of data points increases
  the "binning" done in constructing a histogram builds in choppiness
  gf_density() command
    can be used on any quantitative variable
    smooths out choppiness (natural to the eye?)
    compare
      gf_histogram(~ TotChol, data=NHANES, color="black") with
      gf_dhistogram(~ TotChol, data=NHANES, color="black") with
      gf_density(~ TotChol, data=NHANES)
- bell-shaped distributions (i.e., normal or Gaussian)
  idealized curve, but encountered often, at least approximately
    favstats(~ TotChol, data=NHANES)
    gf_density(~ TotChol, data=NHANES) %>%
        gf_dist("norm", params=c(mean=4.88, sd=1.08))
  sigma as a "unit" of measure
    visualizing
    standardizing a score
      Z = ((unstandardized score) - mean) / (standard deviation)

Not yet discussed

Q4: Who performed better?
  Millie with score of 1410 on the SAT (mean = 1026, sd = 209), or
  Michal with score of 27 on the ACT (mean = 20.8, sd = 4.8), or