Stat 145, Mon 5-Apr-2021 -- Mon 5-Apr-2021
Biostatistics
Spring 2021

Monday, April 5th 2021

Wk 10, Mo
Topic:: Goodess-of-fit
Read:: Lock5 7.1

Example:
In 90 of 400 randomly-selected questions from AP exams where
- Answers were multiple choice
- Five options were offered, A through E

90 of the questions had correct answer B. Is this higher than we would expect from an equally-likely model?

Let p = proportion in population of AP questions that B is correct

Framed as a question we have encountered in the past
- write null, alternative hypotheses

$H_0: p = 0.2, 1-p = 0.8$ $\quad H_a: p \neq 0.2$

- assess normal model appropriateness

$n\hat{p} = (400)(\frac{90}{400}) = 90,$ $\quad n(1-\hat{p}) = 400(1-\frac{90}{400}) = 310$

- analyze within Statkey using randomization

Note:
- More complete look at the data:
  In these 400 questions, the frequency table for correct answers:

observed counts

A:	85
B:	90 — 10 times more often
C:	79
D:	78
E:	68 — 12 times less often

$$\chi^2 = \underbrace{\frac{(85-80)^2}{80}}_{0.313} + \underbrace{\frac{(90-80)^2}{80}}_{1.25} + \frac{(79-80)^2}{80} + \frac{(78-80)^2}{80} + \frac{(68-80)^2}{80} = 3.426$$

Q1: What frequency might we expect? Under equally-likely model, 80

There is another answer, E, that is even farther away from expected count

Q2: Is there a way to
assess whether this sort of freq table arises in equally-likely world
all, not just one, of the discrepancies?

Goodness-of-fit test:

- generalizes 1-proportion hypothesis test to non-binary categorical variables
  does not have a corresponding confidence interval construction

- null hypothesis asserts a null proportion for all values

H_0: $p_A = 0.2$, $p_B = 0.2$, $p_C = 0.2$, $p_D = 0.2$, $p_E = 0.2$

- alternative hypothesis

H_a: At least one of these proportions is incorrect.

- the chi-square statistic
  blame the name on Karl Pearson, inventor around 1900

$$\chi^2 = \sum \frac{[(\text{observed count}) - (\text{expected count})]^2}{\text{expected count}}$$

- bootstrapping
  using a bag
  using a bag
  P-value assesses only the right tail

- theoretical chi-square distribution
  app on class webpage

Another example:

M&M's ~~Hersheys~~ candies, original packs, have been claimed to have
these proportions of colors:

brown:	0.13
red:	0.13
yellow:	0.14
green:	0.16

H_0: $p_{Br} = 0.13$, $p_{Re} = 0.13$, $p_{Ye} = 0.14$, $p_{Gr} = 0.16$,

$p_{Bl} = 0.24$, $p_{Or} = 0.2$

H_a = at least one is different than hypothesized

```
  blue:   0.24
  orange: 0.20

State null/alternative hypotheses



Does the following observed counts taken from a bag with 233 candies
offer significant evidence, at 5% level, against this null hypothesis?
  brown:  35
  red:    27
  yellow: 33
  green:  40
  blue:   47
  orange: 51
```