Stat 145, Wed 20-Oct-2021 -- Wed 20-Oct-2021

Biostatistics

Spring 2021

Wednesday, October 20th 2021

Due:: Moodle Quiz Ch. 4 at 11 pm

Wednesday, October 20th 2021

Wk 8, We

Topic:: Inference for one proportion wrapup

Read:: Lock5 6.2-6.3

~~HW((WW ch06Part1 due Sat.~~

~~HW((PS10 Ellenberg~~

Some finer points concerning inference on one proportion

Confidence intervals for p:

$$SE_{\hat{p}} = \sqrt{\frac{p(1-p)}{n}}$$

- Centered interval approach

critical value

$$\widehat{p} \pm (z^*)(\text{SE}_{\widehat{p}})$$

This approach makes sense if it's reasonable to assume $\widehat{p}$ has a normal distribution, so check the rules of thumb:

$$n\widehat{p} \geq 10 \qquad \text{and} \qquad n(1-\widehat{p}) \geq 10.$$

$\left.\begin{matrix} np \\ n(1-p) \end{matrix}\right\}$ both at least 10

(In CI construction, you have $\widehat{p}$, but not p.)

- Formula $\text{SE}_{\widehat{p}} = \dfrac{\sqrt{p(1-p)}}{\sqrt{n}}$ is modified to $\text{SE}_{\widehat{p}} = \dfrac{\sqrt{\widehat{p}(1-\widehat{p})}}{\sqrt{n}}$ when building CI for p.

Margin of error $= z^*(SE)$

Note that $\text{SE}_{\widehat{p}}$ decreases as n grows. Specifically,

  - can cut $\text{SE}_{\widehat{p}}$ in half by quadrupling ($4\times$) sample size
  - can cut $\text{SE}_{\widehat{p}}$ to one third its size at n by increasing sample size by a factor of 9
  - other similar statements?

Want $\frac{1}{5}$ the size of SE, make n $25\times$ larger

- Can "tailor" a margin of error ME

$$ME = (z^*)(SE_{\hat{p}}) = z^* \sqrt{\frac{p(1-p)}{n}}$$

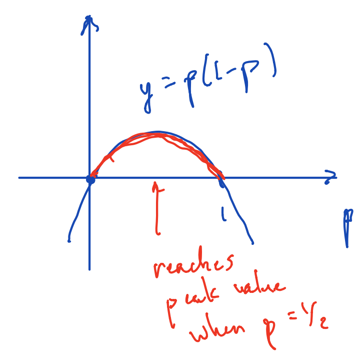
Square both sides

$$(ME)^2 = (z^*)^2 \, \frac{p(1-p)}{n}$$

Solve for n:

$$n \approx \left(\frac{z^*}{ME}\right)^2 \underbrace{p(1-p)}_{\text{worst case at } p = 1/2}$$

problem? Don't know p



So, if my goal is a ME no bigger than M, should be able to achieve that by taking sample size

$$n \geq \left(\frac{z^*}{M}\right)^2 \left(\frac{1}{2}\right)\left(1 - \frac{1}{2}\right) = \left(\frac{z^*}{2M}\right)^2$$

Hypothesis tests involving null hypothesis $\mathbf{H}_0 : p = p_0$: fixed, called "null value"

- several **test statistics** one might use from the sample
  1. the sample proportion $\widehat{p}$
     - natural to use, as an unbiased estimator of p
     - approximating normal distribution: $\mathsf{Norm}(p_0, \overbrace{\sqrt{p_0(1-p_0)/n}}^{\text{SE}})$

       null value at center
  2. the count of successes X
     - natural to use, as an unbiased estimator of p
     - approximating normal distribution: $\mathsf{Norm}(np_0, \sqrt{np_0(1-p_0)})$
  3. the Z-score of either of the previous, calculated either as

$$z = \frac{\widehat{p} - p_0}{\sqrt{p_0(1-p_0)/n}}, \qquad \text{or} \qquad z = \frac{X - np_0}{\sqrt{np_0(1-p_0)}}.$$

     - called the **standardized test statistic**
     - approximating normal distribution: $\mathsf{Norm}(0, 1)$
     - one advantage: immediate comparison with critical z^* value

Practice:

1. Obtain the critical z* value for 96% confidence.

qnorm(0.98), or $z^* = 2.054$

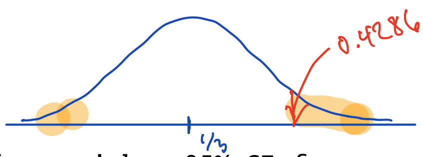
2. One might expect the 3 options in the "rock, paper, scissors" game are equally likely. If a certain player shows "rock" in 51 out of 119 tries, test that rock appears one-third of the time vs. a 2-sided alternative. Use significance level alpha = 0.04.

p = proportion of "rock"

$H_0: p = 1/3$ vs. $H_a: p \neq 1/3$

$\hat{p} = 51/119 = 0.4286$ Estimated SE: $\sqrt{\frac{p(1-p)}{n}} = \sqrt{\frac{(1/3)(2/3)}{119}} = 0.04321$

Can (Option 1) consult Norm(1/3, 0.04321) or (Option 2) standardize the test statistic



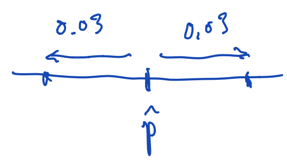
$$z = \frac{0.4286 - 1/3}{0.04321} = 2.205$$

and consult Normal(0, 1)

3. If you wish a 95% CI for p to have a margin of error around 0.03, how large a sample should you have?

$$n \geq \left(\frac{z^*}{2M}\right)^2$$

$M = 0.03$, $z^* = 1.96$



$$n \geq \left[\frac{1.96}{2(0.03)}\right]^2 = 1067.1$$

Take n at least 1068.

4. What might be the minimum sample size n in the last question if you have the additional information that p cannot be larger than 0.25?

$$n \geq \left(\frac{z^*}{M}\right)^2 p(1-p) = \left(\frac{1.96}{0.03}\right)^2 (0.25)(1-0.25) =$$