STOR 155 Introductory Statistics

Lecture 18: Inference for a single proportion

Section 8.1
Normal Approximation for Counts and Proportions

- Let $X \sim B(n, p)$ and $\hat{p} = X / n$.
- If $n$ is large, then

$$\hat{p} \text{ is approx. } N(p, \sqrt{p(1-p)/n}).$$

- Rule of Thumb: $np \geq 10, \ n(1 - p) \geq 10.$
Confidence Interval for $p$

- **Expression:** $[ \hat{p} - m, \hat{p} + m ]$

  where the margin of error $m = z^* \sqrt{\hat{p}(1 - \hat{p})/n}$

- **Assumption:** $n$ is large
- **Confidence level** $C$ determines $z^*$
Hypothesis Testing for $p$

• For a hypothesized value $p_0$, we want to test $H_0 : \ p = p_0$ versus some alternative (1-sided or 2-sided). Recall the 4 steps …

• Step 1: only need to specify $H_a$

• Step 2: Test statistic $z = (\hat{p} - p_0)/\sigma_0$

where $\sigma_0 = \sqrt{p_0(1 - p_0)/n}$
Hypothesis Testing for $p$ (continued)

- **Step 3:** The $P$-value will be equal to
  
  $P(Z > z)$ for 1-sided (upper tail) \( H_a : p > p_0 \)
  
  $P(Z < z)$ for 1-sided (lower tail) \( H_a : p < p_0 \)
  
  $2 \ P(Z > |z|)$ for 2-sided \( H_a : p \neq p_0 \)

- **Step 4:** Compare the $P$-value with the significance level $\alpha$ and draw your conclusion.
“Biased” one-Euro Coin?

- A group of statistics students spun the Belgian one-Euro coin 250 times, and heads came up 140 times.
- \( p = P(H) \) in each spin
- **Claim**: the coin is biased (more specifically, \( p \) is greater than 0.5)
“Biased” one-Euro Coin? (continued)

- Sample: 140 heads among 250 spins of a Belgian one-Euro coin (a hint)
- $p = P(H)$ in each spin

- $H_0$: $p = 0.5$ vs $H_a$: $p > 0.5$ (one-sided upper)

$$z = (140/250 - 0.5) \div \sqrt{0.5(1 - 0.5)/250} = 1.897$$

- $P$-value = $P(Z > 1.897) = 1 - 0.9713 = 0.0287$
- … Conclude based on a given $\alpha$
“Biased” one-Euro Coin? (continued)

What about a 95% CI for $p$?

Note: Margin of error

$$m = z^* \sqrt{\hat{p}(1 - \hat{p})/n} = 1.96 \sqrt{0.56(1 - 0.56)/250} \approx 0.06$$

95% CI = $[0.56 - 0.06, \ 0.56 + 0.06] = [0.5, 0.62]$
Take Home Message

- CI for a single proportion $p$
- Margin of error $m$ in CI
- Hypothesis testing for $p$: 4 steps
- Assumption: large $n$

How large?

$$np_0 \geq 10 \quad \text{and} \quad n(1 - p_0) \geq 10$$