20/05/2020

Date..............

Hypothesis Testing

→ Why do we do Hypothesis Tests?

Could the value of the parameter be __ ?

with that question

use data to help support the claim

One-proportion

Example:

C.S. Mott Children's Hospital Poll

→ C.S. Mott Children's Hospital conducted a national poll on an issue in children's health, sleep habits. We will be looking at an example about lack of sleep in teens.

→ Research Question

In previous years 52% of parents believed that electronics & social media was the cause of their teenager's lack of sleep. Do more parents today believe that their teenager's lack of sleep is caused due to electronics & social media?

→ Population - Parents with a teenager (13-18)

Parameters of interest - p

→ Main goal

Test for a significant increase in the proportion of parents with a teenager who believe that electronics & social media is the cause for lack of sleep.

Date..............

Survey result

→ Hypotheses:

$H_0 : p = 0.52$

$H_1 : p > 0.52$

because "significant increase"

$p = \hat{p} = 0.56$

where $p \to$ population proportion of parents with a teenager who believe that electronics & social media is the cause of their teenager's lack of sleep.

→ $\alpha = 0.05$

significance level.

→ Survey Results

A random sample of 1018 parents with a teenager was taken and 56% said they believe electronics & social media was the cause of their teenager's lack of sleep.

→ Assumptions

↳ We need a random sample of parents

↳ We also need a large enough sample size to ensure our distribution of sample proportions is normal.

That is: $n.p$ be atleast $10 \to n.p_0$

$n.(1-p)$ be at least $10 \to n.p_0$

Date...............

→ Checking Assumptions

Random Sample ✓

$n \cdot p_0 = 1018 \cdot (0.52) = 529$ ✓

$n \cdot (1-p_0) = 1018(1-0.52) = 489$ ✓

Testing a one-population proportion

→ Test statistic

Z - test statistic

$$\frac{\text{Best estimate} - \text{Hypothesized estimate}}{\text{standard error of estimate}}$$

$$= \frac{\hat{p} - p_0}{s.e.}$$

$$s.e.(\hat{p}) = \sqrt{\frac{p \cdot (1-p)}{n}} \Rightarrow \underset{\text{Null}}{\uparrow} s.e.(\hat{p}) = \sqrt{\frac{p_0 \cdot (1-p_0)}{n}}$$

$$Z = \frac{0.56 - 0.52}{0.0157}$$

$$Z = 2.555$$

Interpretation: That means that our observed sample proportion is 2.555 null standard errors above our hypothesized population proportion.

Scanned with CamScanner

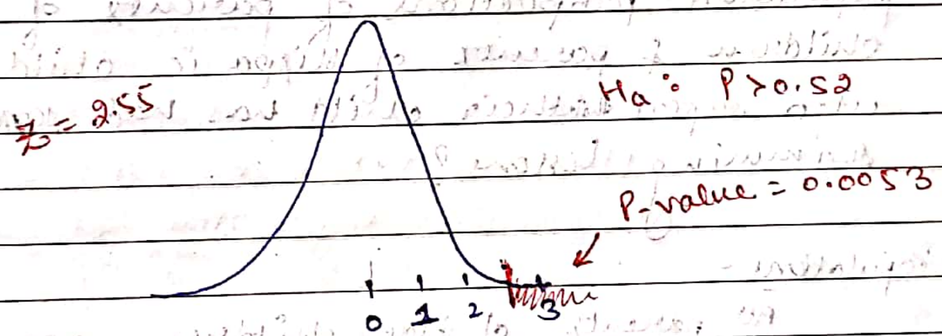
Date..............

- A z-test statistic is another random variable! It has a distribution.
- The z test statistic will always follow a $N(0,1)$
  ↓
  Normal

- This is due to us centring & scaling our original data.

$$\frac{\overbrace{\hat{p} - p_0}^{\text{Centres data}}}{\underbrace{s.e.(\hat{p})}_{\text{Scales data}}}$$

→ The P-value



p-value = 0.0053 < α = 0.05
↓
Reject the null Hypothesis ($H_0 : p = 0.52$)
↓
There is sufficient evidence to conclude that the population proportion of parents with a teenager who believe that electronics & social media is the cause for lack of sleep is greater than 52%.

Two - proportion

Date..............

1. Setting up a Test of difference in Population proportions

→ C.S. Mott Children's Hospital Poll

C.S. Mott Children's Hospital conducted a national poll on an issue in children's health, water safety. We will be looking at an example about swimming lessons.

→ Research Question

Is there a significant difference b/w the population proportions of parents of black children & parents of Hispanic children who report their child has had some swimming lessons?

→ Populations -

All parents of black children age 6-18 & all parents of Hispanic children age 6-18

→ Parameter of interest → $p_1 - p_2$

let, group 1 = black

group 2 = Hispanic

Date..............

→ Main goal

Test for a significance difference in the population proportions of parents reporting that their child has had swimming lessons at the 10% significance level.

→ Hypotheses

$H_0 : p_1 - p_2 = 0$

$H_1 : p_1 - p_2 \neq 0$

→ $\alpha = 0.10$

→ Survey Results

- A sample of 247 parents of black childrens age 6-18 was taken with 91 saying that their child has had some swimming lessons.
- A sample of 308 parents of Hispanic children age 6-18 was taken with 120 saying that their child has had some swimming lessons.

→ Assumptions

- We need to assume that we have two independent random samples.
- We also need large enough sample sizes to assume that the distribution of our estimate is normal. That is, we need $n_1\hat{p}$, $n_1(1-\hat{p})$, $n_2\hat{p}$ & $n_2(1-\hat{p})$ to all be atleast 10.

- Calculating $\hat{p}$

That is, we need to estimate the common proportion and then make sure that we would expect atleast 10 yes's. & 10 no's in each sample.

→ $\hat{p} = \frac{(91+120)}{(247+308)} = \frac{211}{555} = 0.38$

→ Checking Assumptions

- $247\,(0.38) = 94$
- $247\,(1-0.38) = 153$
- $308\,(0.38) = 117$
- $308\,(1-0.38) = 191$

★ If this assumption is not met, we can perform different tests that bypass this assumption.

→ Best Estimate of the Parameter

$\hat{p}_1 = 91/247 = 0.37$ $\boxed{1 = \text{black}}$

$\hat{p}_2 = 120/308 = 0.39$ $\boxed{2 = \text{hispanic}}$

$\hat{p}_1 - \hat{p}_2 = 0.37 - 0.39 = \boxed{-0.02}$

This -ve means that the sample proportion for black children is less than sample proportion of Hispanic children.

Date..............

→ <u>Test Statistic</u>

- $= \dfrac{\text{Best Estimate} - \text{Hypothesized estimate}}{\text{Standard Errors of estimate}}$

$$= \frac{(\hat{p}_1 - \hat{p}_2) - 0}{se(\hat{p})}$$

where, $se(\hat{p}) = \sqrt{\hat{p}(1-\hat{p})\left(\frac{1}{n_1} + \frac{1}{n_2}\right)}$

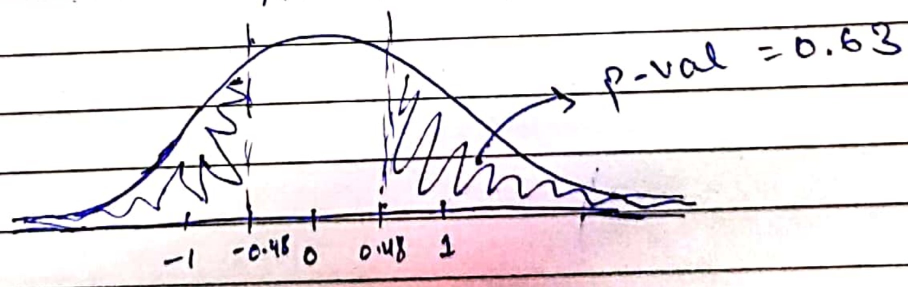
$$Z_0 = \frac{-0.02}{0.041}$$

$$Z_0 = -0.48$$

- Interpretation

That means that our observed differences in sample proportion is 0.48 estimated standard errors below our hypothesized mean of equal population proportions.

~~Normal (0,1)~~



→ <u>Decision & Calculation</u>

p-val = 0.63 > 0.10 → fail to reject null hypothesis.

α

Date...............

- don't have evidence against equal proportions.

→ Formally, based on our sample and our p-value, we fail to reject the null hypotheses. We conclude that there is no significant difference b/w the population proportion of parents of black & Hispanic children who report their child has had swimming lessons.