One-Sample Correlation Case

Course: Statistics 1

Lecturer: Dr. Courtney Pindling

Introduction

Is the correlation coefficient significantly different from ***0*** or some reference value, *a*?

Test whether the linear relationship between x and y is significant by testing hypothesis about the population correlation coefficient, ρ_{xy}:

Case 1: $H_0: \rho_{xy} = 0$ ***or***

Case 2: $H_0: \rho_{xy} = a$

Note: Will only examine Case 1 in this lecture

Approach 1

Critical Value from *t*-distribution

- Given: a (0.05 or 0.01) and *df* (n – 2)
- Sample: r_{xy} and *n*

Case 1 (H_0: $\rho = 0$): Hypothesis

- Null Hypothesis:
  - H_0: $\rho = 0$
- Alternative Hypothesis:
  - H_0: $\rho \neq 0$

Example: A sample with $\boldsymbol{n}$ **= 10** (x and y pairs) produced a correlation coefficient of $\boldsymbol{r_{xy}}$ **= 0.4501**. Is this correlation different from zero?

Case 1 (H_0: $\rho = 0$): Rejection Criteria

- We use hypothesized $\rho = 0$
  - Underlying t-distribution
  - Large ***n*** approaches standard normal distribution
  - ***df*** **=** $n - 2$ or *8*
  - Critical Value, CV:

    for two-tailed t with $df = 8$ and $\boldsymbol{a = 0.05}$ is

    $t_{cv} = \boldsymbol{2.306}$

Case 1 (H_0: ρ = 0): Test Statistics

$$t = r_{xy}\sqrt{\frac{n-2}{1-r_{xy}^2}}, \text{ where } r_{xy} \text{ is sample } r$$

$$t_{stat} = r_{xy}\sqrt{\frac{n-2}{1-r_{xy}^2}} = 0.4501\sqrt{\frac{10-2}{1-0.4501^2}} = 1.4256$$

Case 1 (H_0: $\rho = 0$): Decision

- Results:
  - Given: a = 0.05, *df* = 8, t_{cv} = 2.306
  - t_{stat} = 1.4256
- Decision: (do **not** reject H_0):
  - Since $t_{stat} < t_{cv}$ or 1.4256 < 2.306
- Conclusion:
  - The *r* = 0.4501 is not statistically different from 0, so there is no linear relationship between *x* and *y*

Approach 2

Correlation Critical Value Table

- Given: a (0.05 or 0.01) and *df* (n – 2)
- Sample: r_{xy} and n

Case 1 (H_0: $\rho = 0$): CV Table

Level of Significance (a) for a Two-Tailed Test

df (n-2)	0.1	0.05	0.02	0.01
1	0.988	0.997	0.9995	0.9999
2	0.9	0.95	0.98	0.99
3	0.805	0.878	0.934	0.959
4	0.729	0.811	0.882	0.917
5	0.669	0.754	0.833	0.874
6	0.622	0.707	0.789	0.834
7	0.582	0.666	0.75	0.798
8	0.549	0.632	0.716	0.765

Case 1 (H_0: $\rho = 0$): CV Criteria

- *The value found in the CV Table (a and df) is minimum **r** needed to be 95% (a = 0.05) confident that a relationship exists*

- *If the absolute value of **r** is above **0.632**, reject H_0 (there is no relationship) and accept the H_a. There is a statistically significant relationship between x and y ($p < 0.05$)*

- *If the absolute value of **r** is below **0.632**, do not reject H_0. There is not a statistically significant relationship between x and y ($p < 0.05$)*

- *A **p =< 0.05** means that **r** exceeds the critical value found in the CV Table and you are 95% confident that a relationship exists. A **p > 0.05** means that **r** was less than the critical value in the CV Table and you cannot be 95% confident that a relationship exists*

Case 1 (H_0: ρ = 0): CV Decision

- Results:
  - Given: a = 0.05, *df* = 8, r_{cv} = 0.632
- Decision: (do **not** reject H_0):
  - Since $r_{sample} < r_{cv}$ or 0.4501 < 0.632
- Conclusion:
  - The *r* = 0.4501 is not statistically different from 0, so there is no linear relationship between *x* and *y*
  - *Same conclusion as when used the t-critical value*