

Inference and simulation

Effect size



Effect size

- We would like to know how significant a result is, not just whether or not we can reject the null hypothesis.
- The **effect size** is how big of a difference there is between two distributions.
- There are different ways to compute an effect size.

Phi coefficient

- Categorical response variable
- Categorical explanatory variable

Effect size	Phi
Small	0.10
Medium	0.30
Large	0.50

Cohen's *d*

- Numerical response variable
- Categorical explanatory variable

Effect size	<i>d</i>
Very small	0.01
Small	0.20
Medium	0.50
Large	0.80
Very large	1.20
Huge	2.00

Effect size of gender discrimination experiment

The formula for computing the phi coefficient is as follows,

$$\varphi = \sqrt{\frac{\chi^2}{N}}$$

χ^2 is the Chi-Square statistic and N is the number of observations in the dataset.

Effect size of gender discrimination experiment

The formula for computing the phi coefficient is as follows,

$$\varphi = \sqrt{\frac{\chi^2}{N}}$$

χ^2 is the Chi-Square statistic and N is the number of observations in the dataset.

The Chi-Square statistic can be easily computed using `infer`, and the rest of the computation can be handled with `mutate()`,

```
num_samples <- count(applicants_data)
applicants_data %>%
  specify(outcome ~ sex, success = "Promoted") %>%
  calculate(stat = "Chisq", order = combine("Male", "Female")) %>%
  mutate(effect_size = sqrt(stat / pull(num_samples, n))) %>%
  pull(effect_size)
```

```
## [1] 0.2812843
```

Effect size of gender discrimination experiment

The formula for computing the phi coefficient is as follows,

$$\varphi = \sqrt{\frac{\chi^2}{N}}$$

χ^2 is the Chi-Square statistic and N is the number of observations in the dataset.

The Chi-Square statistic can be easily computed using `infer`, and the rest of the computation can be handled with `mutate()`,

```
num_samples <- count(applicants_data)
applicants_data %>%
  specify(outcome ~ sex, success = "Promoted") %>%
  calculate(stat = "Chisq", order = combine("Male", "Female")) %>%
  mutate(effect_size = sqrt(stat / pull(num_samples, n))) %>%
  pull(effect_size)
```

```
## [1] 0.2812843
```

The effect size for the observed result is *medium*.

Credits

License

Creative Commons Attribution-NonCommerical-ShareAlike 4.0 International