Page
Date / /

5. A/B Testing for Website Conversion Rate Optimization

Dataset

| Group | Converted | Count |
| --- | --- | --- |
| A | 0 | 4695 |
| | 1 | 305 |
| B | 0 | 4607 |
| | 1 | 393 |

| Group | Converted | COUNTA of Use |
|---|---|---|
| A | 0 | 4695 |
| | 1 | 305 |
| A Total | | 5000 |
| B | 0 | 4607 |
| | 1 | 393 |
| B Total | | 5000 |
| **Grand Total** | | **10000** |

| Converted | COUNTA of Use |
|---|---|
| 0 | 9302 |
| 1 | 698 |
| **Grand Total** | **10000** |

| Group | COUNTA of Use |
|---|---|
| A | 5000 |
| B | 5000 |
| **Grand Total** | **10000** |

Page
Date / /

Group (A) (Control):

- no. of users: $n_A = 5000$
- no. of conversions: $x_A = 305$

Group (B) (Treatment):

- no. of users: $n_B = 5000$
- no. of conversions: $x_B = 393$

sample conversion rates:

- $p_A = \frac{x_A}{n_A} = \frac{305}{5000} = 0.0610 = 6.10\%$
- $p_B = \frac{x_B}{n_B} = \frac{393}{5000} = 0.0786 = 7.86\%$

Compute (SE) for each proportion:

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

Page

Date / /

$SE_A = 0.003386$

$SE_B = 0.003806$

Construct the 95% Confidence Intervals:

A 95% confidence interval for a proportion is given by:

$$\hat{p} \pm Z_{\alpha/2} \times SE(\hat{p})$$

for 95% CI, $Z_{\alpha=0.5/2} = Z_{0.025} = 1.96$

Group A:

- Lower Bound:

$$L_A = p_A - 1.96 \times SE_A = 0.0610 - (1.96 \times 0.003380)$$

$$= 0.0610 - 0.006635$$

$$= 0.05437$$

Upper Bound:

$$U_A = p_A + 1.96 \times SE_A = 0.06764$$

Page
Date / /

- 95% CI for Group A is (0.0544, 0.0676)
or (5.44%, 6.76%).

Group B:

Lower Bound:

$L_B = p_B - 1.96 \times SE_B = 0.0786 - 1.96 \times 0.003806$

$= 0.07114$

Upper Bound:

$U_B = p_B + 1.96 \times SE_B = 0.08606$

- 95% CI for Group B is (0.0711, 0.0861)
or (7.11%, 8.61%)

Interpretation

(A) Conversion rate is 6.10% with a 95% CI of approx.
(5.44%, 6.76%)

(B) C.R. is 7.86% with a 95% CI (7.11%, 8.16%)

The treatment group shows a higher conversions rate than the control group and C.I. suggest that the difference might be statistically significant.

Define Hypothesis -

- Null Hypothesis ($H_0$):

The conversion rates are equal between the two groups.

$$H_0 : p_A = p_B$$

- Alternative hypothesis ($H_1$):

The conversion rates differ b/w the two groups:

$$H_1 : p_A \neq p_B$$

* Data and sample proportions -

Page
Date / /

Under null hypothesis, we assume conviouiarivate

$$\hat{p} = \frac{x_A + x_B}{n_A + n_B} = \frac{698}{10000} = 0.0698$$

· Compute the (SE)·

$$SE = \sqrt{\hat{p}(1-\hat{p})\left(\frac{1}{n_A} + \frac{1}{n_B}\right)}$$

$$S.E. = \sqrt{0.00002596} \approx 0.00510$$

· Z-statistic -

$$z = \frac{p_B - p_A}{SE} = \frac{0.0786 - 0.0610}{0.00510} \approx \frac{0.0176}{0.00510} = 3.45$$

· Determine the p-value -

- $Z = 3.45$, one tailed prob. is very small (around 0.0003)
- for two-tailed test:-

$$p\text{-value} = 2 \times P(Z > 3.45)$$
$$= 2 \times 0.0003$$
$$= 0.0006$$

Page
Date / /

p-value << 0.05, reject $H_0$.

Interpretation -

- the two-proportion Z-test ⇒ Z statistic value of 3.45 with p-value of 0.0006.

- this provides strong evidence against the null hypothesis, indicating that new checkout page has significantly higher conversions rate than the existing page.

Potential risks of Type I (false positive) and Type II (false negative) errors.

Type I (False Positive):

- Rejecting the null hypothesis when it is actually true.

- Concluding that the new checkout page (treatment) improves conversion rates when, in reality, there is no difference.

Page
Date / /

- could lead to implementing a change that doesn't truly benefit the company.

x————————x

- Failing to reject the null hypothesis when the alternative hypothesis is actually true.

- Concluding that there is no difference between the checkout pages when new page does in fact improve conversion rates.

- this would mean missing an opportunity to increase conversions

x————————x