

Analysis of Form Variations and Their Effects on Conversion Rate

Analysis methodology

Experiment results were analyzed using a typical A/B Testing analysis approach, which entails calculating a sample z-value and comparing it to a critical z-value, which in turn is determined based on the desired confidence interval for acceptance/rejection of the null hypothesis. Z-testing was used since sample size was greater than 30. Since we are interested only in the probability of the 'conversion' (quote sent to consumer by service provider) rate of the experimental form(s) being higher/lower than the baseline form, we use a right-hand z-test.

Calculations were performed using Excel (technically google sheets). Online research was utilized to refresh conceptual statistical knowledge. Documentation of primary online sources can be found in the References section at the end of the report.

Assumptions

- Normal distribution of the data
- No special events affecting data (holidays, fiscal period pushes, world events, etc)
 - *(In other words, that the data for the experimental period is representative of a typical conversion/sales period for this product and industry)*
- A 95% confidence interval is desired by those that will make the decisions for form changes based on these experimental results

Hypotheses

Null hypothesis:

$$H_0 : p - p_c \leq 0$$

Where

p = conversion rate of Variation 3

p_c = conversion rate of Baseline (control)

*The conversion rate of Variation 3 is **no greater than** the conversion rate of the Baseline form*

Alternate hypothesis:

$$H_0 : p - p_c > 0$$

*The conversion rate of Variation 3 is **greater than** the conversion rate of the baseline form*

Quantitative Results

Bucket	Quotes	Views	Views Quoted	pc	Nc	p	N	z	Rel. Uplift	Result vs. Baseline
Baseline	32	595	5.38%	0.054	595	-	-	-	-	-
Variation 1	30	599	5.01%	-	-	0.050	599	0.288	-6.88%	Statistically insignificant difference
Variation 2	18	622	2.89%	-	-	0.029	622	2.173	-46.19%	Statistically worse performance
Variation 3	51	606	8.42%	-	-	0.084	606	2.083	56.48%	Statistically better performance
Variation 4	38	578	6.57%	-	-	0.066	578	0.864	22.24%	Statistically insignificant difference

Right-hand test, 95% confidence interval → Critical z-value for significance is 1.645

Calculations

For the sake of completeness, all calculations – even simple ones – will be sampled.

Variables will be defined only the first time they introduced.

Views Quoted Percentage (aka Conversion Percentage)

$$p = C/N$$

where:

- C = number of service providers that paid to send quotes to consumers (ie number of conversions)

- N = sample size of experimental treatment (ie total number of viewers of Form Var 3)

Z-Score

$$z = \frac{p - p_c}{\sqrt{\frac{p(1-p)}{N} + \frac{p_c(1-p_c)}{N_c}}}$$

where:

- z = the z-score, or the number of standard deviations from the mean
- N_c = sample size of control treatment (ie total number of viewers of the Baseline Form)

Relative Uplift

$$RU = \frac{p - p_c}{p_c}$$

where:

- RU = the relative uplift

Conclusions

Each of the four form variations had a different outcome for the statistical significance of the difference in their performance vs. the baseline form. Form 1, which had a 6.88% lower number of conversions vs. baseline, did not have a statistically significant difference in performance, since its calculated z-score of -0.288 did not exceed the (-)1.645 critical z-score for the 95% confidence interval. Form 2, which had a 46.19% lower number of conversions vs. baseline, did have a statistically significant difference in performance, since its calculated z-score of -2.173 did exceed the (-)1.645 critical z-score for the 95% confidence interval.

The second two forms both showed a greater number of conversions vs. baseline, but only one was determined to be significant. Form 3, which had a +56.48% higher number of conversions vs. baseline, did have a statistically significant difference in performance, with its calculated z-score of 2.083 safely exceeding the 95% confidence interval critical z-score of 1.645. Form 4, however, which had a +22.24% higher number of conversions vs. baseline, was – counter-intuitively – **not** found to have a statistically

significant difference in performance. Although 22.24% seems like a big difference in performance, with this sample size and 95% confidence interval, its calculated z-score of 0.864 was well below the significance threshold of 1.645.

Form Variation 3 is the obvious choice of forms for the Acme Corporation to use going forward. This change from the Baseline form to Variation 3, which is intuitive from the calculated conversion percentages, can be enacted with a confidence in excess of 95% that it will have a greater conversion rate than the baseline.

In further statistical language, based on the above analysis we can reject our null hypothesis for Variation 3, and accept our alternate hypothesis.

Follow-up Questions and Potential for Further Study

What is Acme's goal conversion rate?

We proved in this analysis that Form Variation 3 is statistically significantly better than the Baseline Form within the assumed 95% confidence interval. However, what we have not addressed is if this performance meets the business and performance goals of the organization.

Were there any special events in the timing of the experiment that could have affected the results?

Any external factors that are not typically present in regular operations can influence the data set and potentially invalidate the results. Before making any changes, Acme should confirm that there were no special events during the experimental period.

What if any effect did segmentation have on the results?

Insurance providers often target a specific segment of the market, which could potentially have impacted results. An example that comes to mind is insurance providers that target individuals and families in the military; if Variation 3 had a disproportionate number of military consumers, and/or a military-focused insurance provider had been engaged in a marketing push, this could have affected the results. Before making any changes, Acme should do further investigation to confirm that such effects were not reflected in the experiment.

A final comment on the setup of the experiment itself. This was analyzed as an A/B test, but technically it can be more accurately described as a A/B/C/D/E test, which, for a total sample size of 3000, is not optimal for obtaining solid, statistically significant data sets. In my opinion, I would consider this to be an initial exploratory test, and recommend further testing – perhaps with Variations 3, 4, and the Baseline. The Variations that have been shown to negatively affect conversions (in the case of Variation

2) , or do not look promising to show a positive effect (Variation 1) can be eliminated, and further experimentation with larger sample size can be done without any considerable risk of lowering conversion rates (ie dropping revenues) during the experiment.