

question3

August 27, 2018

1 DSI Technical Exercise, Question 3

1.1 Business Case Study for Experiment and Analysis

- A consumer posts a **request** for a service needed. Every request is in some **category** (e.g., Catering, Personal Training, Interior Design) and some **location** (e.g., New York, San Francisco).
- We match the request up with appropriate service providers and send each of those providers an **invite** to quote on the request.
- Providers view the invite and some choose to send a **quote** to the consumer expressing interest.

1.2 Split Test (or A/B/n) Experimental Design and Test Results

I've just concluded a test of our *quote form*. After receiving an invite, service providers come to the quote form to view the consumer request and choose whether or not to pay to send a quote. My goal was to determine if certain changes to the design of the form would cause more providers to send a quote after coming to the page.

Over the course of a week, I divided invites from about 3000 requests among four new variations of the quote form as well as the baseline form we've been using for the last year. Here are my results:

- Baseline: 32 quotes out of 595 viewers
- Variation 1: 30 quotes out of 599 viewers
- Variation 2: 18 quotes out of 622 viewers
- Variation 3: 51 quotes out of 606 viewers
- Variation 4: 38 quotes out of 578 viewers

1.3 Split Test (or A/B/n) Analysis

Analytical Assumptions As population inferences will be made using the sample statistics generated as part of the test execution these assumptions must be made: the sample statistic distribution is approximately normal, the samples are independent, and the sample size is significantly large.

The significance level, or α , will be 5% for this analysis.

Experimental Design The summary of the experimental design overview indicates human involvement with the assignment of service providers and a particular summary page. Further discussion is needed to ensure that the assignment process is sufficiently randomize.

```
In [34]: # import standard libraries for analytical work
import pandas as pd
import statsmodels.api as sms
```

```
In [35]: # load csv test results into a dataframe
df = pd.read_csv("./test_data/acme_corp.csv")
# take a look at the dataframe
df.head()
```

```
Out[35]:
```

| | Bucket | Quotes | Views |
|---|-------------|--------|-------|
| 0 | Baseline | 32 | 595 |
| 1 | Variation 1 | 30 | 599 |
| 2 | Variation 2 | 18 | 622 |
| 3 | Variation 3 | 51 | 606 |
| 4 | Variation 4 | 38 | 578 |

The Hypotheses The A/B/n (or Split Test) testing analysis is a controlled experiment leveraging hypothesis statements. Let's go ahead and document all four null/alternative statements.

Null and alternative hypothesis for the controlled experiment of the variation 1 of the submittal page versus the baseline or existing submittal page.

$$H_0 : \pi_{\text{variation1}} - \pi_{\text{baseline}} = 0$$

$$H_a : \pi_{\text{variation1}} - \pi_{\text{baseline}} > 0$$

Null and alternative hypothesis for the controlled experiment of the variation 2 of the submittal page versus the baseline or existing submittal page.

$$H_0 : \pi_{\text{variation2}} - \pi_{\text{baseline}} = 0$$

$$H_a : \pi_{\text{variation2}} - \pi_{\text{baseline}} > 0$$

Null and alternative hypothesis for the controlled experiment of the variation 3 of the submittal page versus the baseline or existing submittal page.

$$H_0 : \pi_{\text{variation3}} - \pi_{\text{baseline}} = 0$$

$$H_a : \pi_{\text{variation3}} - \pi_{\text{baseline}} > 0$$

Null and alternative hypothesis for the controlled experiment of the variation 4 of the submittal page versus the baseline or existing submittal page.

$$H_0 : \pi_{\text{variation4}} - \pi_{\text{baseline}} = 0$$

$$H_a : \pi_{\text{variation4}} - \pi_{\text{baseline}} > 0$$

Each of these null statements assert that no difference exists between the population proportion, where proportion is represented as π .

All 4 of the alternative statements assert that the population proportion is greater than the existing population proportion, where proportion is represented as π .

Built-In Function "proportions_ztest" The A/B/n testing analysis will utilize a built-in statsmodel function called `proportions_ztest`. Looking at the help page for this function several arguments are required for this calculation, which will be performed in the next cell.

```

In [36]: # baseline count and nob
baseline_df = df.query("Bucket == 'Baseline'")
baseline_count = baseline_df.Quotes.sum()
baseline_nobs = baseline_df.Views.sum()
print(f"Baseline count and nob is: {baseline_count} and {baseline_nobs}")

# variation 1 count and nob
variation1_df = df.query("Bucket == 'Variation 1'")
variation1_count = variation1_df.Quotes.sum()
variation1_nobs = variation1_df.Views.sum()
print(f"Variation 1 count and nob is: {variation1_count} and {variation1_nobs}")

# variation 2 count and nob
variation2_df = df.query("Bucket == 'Variation 2'")
variation2_count = variation2_df.Quotes.sum()
variation2_nobs = variation2_df.Views.sum()
print(f"Variation 2 count and nob is: {variation2_count} and {variation2_nobs}")

# variation 3 count and nob
variation3_df = df.query("Bucket == 'Variation 3'")
variation3_count = variation3_df.Quotes.sum()
variation3_nobs = variation3_df.Views.sum()
print(f"Variation 3 count and nob is: {variation3_count} and {variation3_nobs}")

# variation 3 count and nob
variation4_df = df.query("Bucket == 'Variation 4'")
variation4_count = variation4_df.Quotes.sum()
variation4_nobs = variation4_df.Views.sum()
print(f"Variation 4 count and nob is: {variation4_count} and {variation4_nobs}")

```

```

Baseline count and nob is: 32 and 595
Variation 1 count and nob is: 30 and 599
Variation 2 count and nob is: 18 and 622
Variation 3 count and nob is: 51 and 606
Variation 4 count and nob is: 38 and 578

```

```

In [37]: # calculate the z-score and p-value for variation1 and baseline experiment
z_score_variation1, p_value_variation1 = sms.stats.proportions_ztest([variation1_count, baseline_count])
print(f"The p-value for the Variation 1 and Baseline experiment is {p_value_variation1}")

# calculate the z-score and p-value for variation2 and baseline experiment
z_score_variation2, p_value_variation2 = sms.stats.proportions_ztest([variation2_count, baseline_count])
print(f"The p-value for the Variation 2 and Baseline experiment is {p_value_variation2}")

# calculate the z-score and p-value for variation3 and baseline experiment
z_score_variation3, p_value_variation3 = sms.stats.proportions_ztest([variation3_count, baseline_count])
print(f"The p-value for the Variation 3 and Baseline experiment is {p_value_variation3}")

```

```
# calculate the z-score and p-value for variation4 and baseline experiment
z_score_variation4, p_value_variation4 = sms.stats.proportions_ztest([variation4_count, baseline_count])
print(f"The p-value for the Variation 4 and Baseline experiment is {p_value_variation4}")
```

The p-value for the Variation 1 and Baseline experiment is 0.6133099143530945

The p-value for the Variation 2 and Baseline experiment is 0.9854676438929962

The p-value for the Variation 3 and Baseline experiment is 0.018986239169411456

The p-value for the Variation 4 and Baseline experiment is 0.19360793468873894

Conclusions The p-value for Variation 1, 2, and 4 experiment is greater than the significance level of 5% and the null hypothesis will not be rejected. The p-value for experiment is less than the significance level of 5% for the third variation of the submittal page, make that experiment the only variation allowing us to reject the null hypothesis.

Recommending that Variation 3 of the submittal page be scheduled for installation into the production domain for customer use.

References

1. "A Refresher on A/B Testing". Retrieved from <https://hbr.org/2017/06/a-refresher-on-ab-testing#>
2. "proportions_ztest". Retrieved from http://knowledgetack.com/python/statsmodels/proportions_ztest/