# P(A>B) Based Analysis - Unbalanced Data

Rain

5/27/2021

This analysis is based on Bayesian result according to P(A>B) > 95% and unbalance ratio = 0.25 (treatmet:control = 4:1).

Recall that

| Case ID | Application | Distribution | Effect |
|---|---|---|---|
| 1 | Payer Conversion | Bernoulli | False |
| 2 | Payer Conversion | Bernoulli | True |
| 3 | Total Moves | Poisson | False |
| 4 | Total Moves | Poisson | True |

## False/True Positive Rate

All the percentages in the following tables are proportion of the simulated experiments that led to the conclusion that the treatment is chosen.

In Case 1 and 3 where the underlying treatment and control have no difference, the percentages mean false positive rates. Therefore, the lower the percentage, the better.

In Case 2 and 4 where the treatment is better than control, the percentages mean true positive rates. Therefore, the higher the percentage, the better.

## Areas of Analysis

### 1. Effect of peeking

We calculate peek_multiplier - how many times more likely we would choose treatment if monitor daily and stop the experiment earlier when we see a result that's positive enough. Note that only sample size of 500 is used because when the sample size is large enough, % of accepting treatment goes to 100% very quickly and would skew the peek_multiplier.

Result: Bayesian suffers from peeking as well.

| case | avg_peek_multiplier_freq | avg_peek_multiplier_bayes |
|---|---|---|
| 1 | 4.8 | 3.4 |
| 3 | 4.5 | 5.0 |

## 2. Effect of sample size

Conclusion: When treatment is not better than control, the false positive rate is controled at 5%.

| case | sample_size_per_day | avg_freq_treat | avg_freq_treat_peek | avg_bayes_treat | avg_bayes_treat_peek |
|---|---|---|---|---|---|
| 1 | 500 | 4.15% | 19.75% | 3.6% | 12.65% |
| 1 | 5000 | 4.45% | 24.55% | 4.45% | 19.6% |
| 1 | 50000 | 5.05% | 28.3% | 3.9% | 23.35% |
| 3 | 500 | 6.33% | 28.2% | 6.8% | 33.73% |
| 3 | 5000 | 5.4% | 28.33% | 5.87% | 28.87% |
| 3 | 50000 | 5.8% | 27% | 5.27% | 24% |

Conclusion: When treatment is better than control, Bayesian has slightly more power.

| case | sample_size_per_day | avg_freq_treat | avg_freq_treat_peek | avg_bayes_treat | avg_bayes_treat_peek |
|---|---|---|---|---|---|
| 2 | 500 | 10.5% | 26.6% | 15.07% | 29.2% |
| 2 | 5000 | 52.2% | 64.65% | 57.35% | 67.7% |
| 2 | 50000 | 82.35% | 89.2% | 87.33% | 93.4% |
| 4 | 500 | 9.03% | 32.03% | 14.93% | 45.93% |
| 4 | 5000 | 38.87% | 55.9% | 44.27% | 62.07% |
| 4 | 50000 | 56.03% | 67.67% | 59.6% | 72.07% |

## 3. Effect of prior parameter selections

We compare average false/true positive rate by directional, confident, and wrong priors.

Conclusion: Bayesian's prior parameters do not matter much for the sample sizes we have.

| case | prior | avg_freq_treat | avg_freq_treat_peek | avg_bayes_treat | avg_bayes_treat_peek |
|---|---|---|---|---|---|
| 1 | confident | 4.33% | 23.8% | 3.67% | 17.53% |
| 1 | directional | 4.93% | 24.07% | 4.53% | 20.67% |
| 1 | neutral | 4.2% | 25.07% | 4.2% | 20.53% |
| 1 | wrong | 4.73% | 23.87% | 3.53% | 15.4% |
| 2 | confident | 48.9% | 60.07% | 53.1% | 61.7% |
| 2 | directional | 48.47% | 60.23% | 53.43% | 63.23% |
| 2 | neutral | 48.37% | 60.4% | 54.17% | 66.47% |
| 2 | wrong | 47.67% | 59.9% | 52.3% | 62.33% |
| 3 | confident | 6.2% | 28% | 6.67% | 32.13% |
| 3 | directional | 6.27% | 28.07% | 6.6% | 27.8% |
| 3 | wrong | 5.07% | 27.47% | 4.67% | 26.67% |
| 4 | confident | 34.87% | 51.9% | 40.4% | 62.97% |
| 4 | directional | 33.6% | 50.97% | 38.03% | 58.17% |
| 4 | wrong | 35.47% | 52.73% | 40.37% | 58.93% |