Meets Specifications

SHARE YOUR ACCOMPLISHMENT

Dear student,

I've reviewed your very first submission, well done successfully addressing every previous issue, you made it for an excellent project. I hope this hard work helped you gather a practical knowledge and understanding of the many concepts involved in A/B testing and deal with the complex and challenging decision nodes involved in designing experiments and in interpreting results.

Congratulations on passing your exam!

Metric Choice

✓ **A good set of metrics have been selected for the experiment, without missing any necessary or valuable metrics.**

✓ **Each metric has a clear and well-reasoned explanation of why it was or was not chosen as an invariant metric and as an evaluation metric.**

Please note that the click through probability is not a good evaluation metric but it could be a very good invariant metric, it would make it for a better invariant metric compared to the number of clicks, as it normalizes to the size of the control and experiment group.

✓ **The report clearly states what results we look for in order to launch the experiment and the stated results are aligned with the experiment goals.**

Variability

✓ **The standard deviations for all evaluation metrics have been correctly calculated.**

✓ **Each evaluation metric has a clear and correct explanation of whether the analytic variability is likely to match the empirical variability.**

Sizing

- ✓ The number of pageviews given is correct given the students choice of whether to use the Bonferroni correction.
- ✓ A well-reasoned argument about how risky the experiment will be is made and a fraction of traffic to divert is chosen accordingly.
- ✓ The duration of the experiment is correctly calculated given the fraction of traffic to divert that was chosen.

Sanity Checks

- ✓ The sanity checks have been correctly calculated for all chosen invariant metrics.
- ✓ The passing or failure of all sanity checks have been evaluated. If sanity checks failed, analysis has been performed to discover why the sanity checks may have failed and the experiment has not been continued.

Effect Size Tests

- ✓ Correctly calculated confidence intervals have been reported for the difference in all evaluation metrics.
- ✓ Statistical and practical significance have been correctly reported for all evaluation metrics.

Sign Tests

- ✓ P-value and statistical significance have been correctly reported for all evaluation metrics.

Results Summary

✓ **The report provides good justification for the choice of whether to use the Bonferroni correction.**

Correct: When we are considering multiple metrics at the same time, and we need all of them to inform our decision (in our case, we are looking at both gross and net conversion to decide whether to launch or not), the risk is of a type II error. That is not what the Bonferroni correction is designed for. The Bonferroni correction is designed to reduce the risk that one metric is deemed significant by mistake. If we were in the situation where we need just one metric to meet expectations in order to launch an experiment, then we would need Bonferroni. In our case we would need multiple metrics to match our expectations to launch the experiment therefore Bonferroni is neither necessary nor helpful.

✓ **A well-reasoned and plausible explanation for each discrepancy between the effect size tests and the sign tests has been provided.**

Recommendation

✓ **A recommendation is made that is well-reasoned and supported by the data.**

Follow-Up Experiment

✓ **A plausible experiment that would be worth testing has been made. A hypothesis for results of the experiment is clearly stated.**

✓ **The metrics chosen in the report will be sufficient to evaluate the hypothesis of the experiment, would be possible to measure under most infrastructures, and are well-supported by reasoning in the report.**

Well done providing a coherent set of metrics considering that you're dealing with enrolled students.

✓ **The report describes a reasonable unit of diversion and gives good support for this choice.**

DOWNLOAD PROJECT