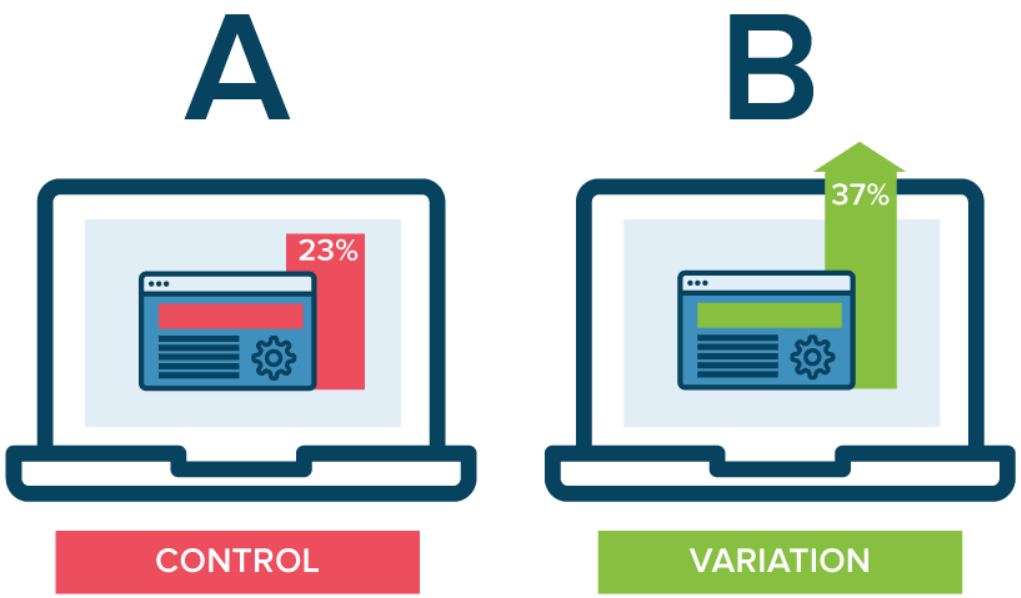
A/B Testing on a Mobile Game Data

- Sameer Kumar Mohapatra
(M13969443)

What is A/B Testing?



- It is a process of checking variation in a product without taking account of the changes that were caused by outside factors
- Two groups (control and test variants)are tested for the difference in the metrics.
- Use Cases: Website Redesign, Setting up difficulty levels in a game, Call to action button design, Advertisements etc.

Experiment design

1

Formulate the Hypothesis

H0: New feature in the test hasn’t changed anything (No change in metrics)

H1: New feature has changed the metrics

2

Decide the metrics to be checked

Categorical metrics (Yes/No, Churned/Returned etc.)

Continuous metrics (Average value, rate of increase etc.)

3

Estimate the Sample size

Confidence intervals for a proportion

Test of difference between sample means

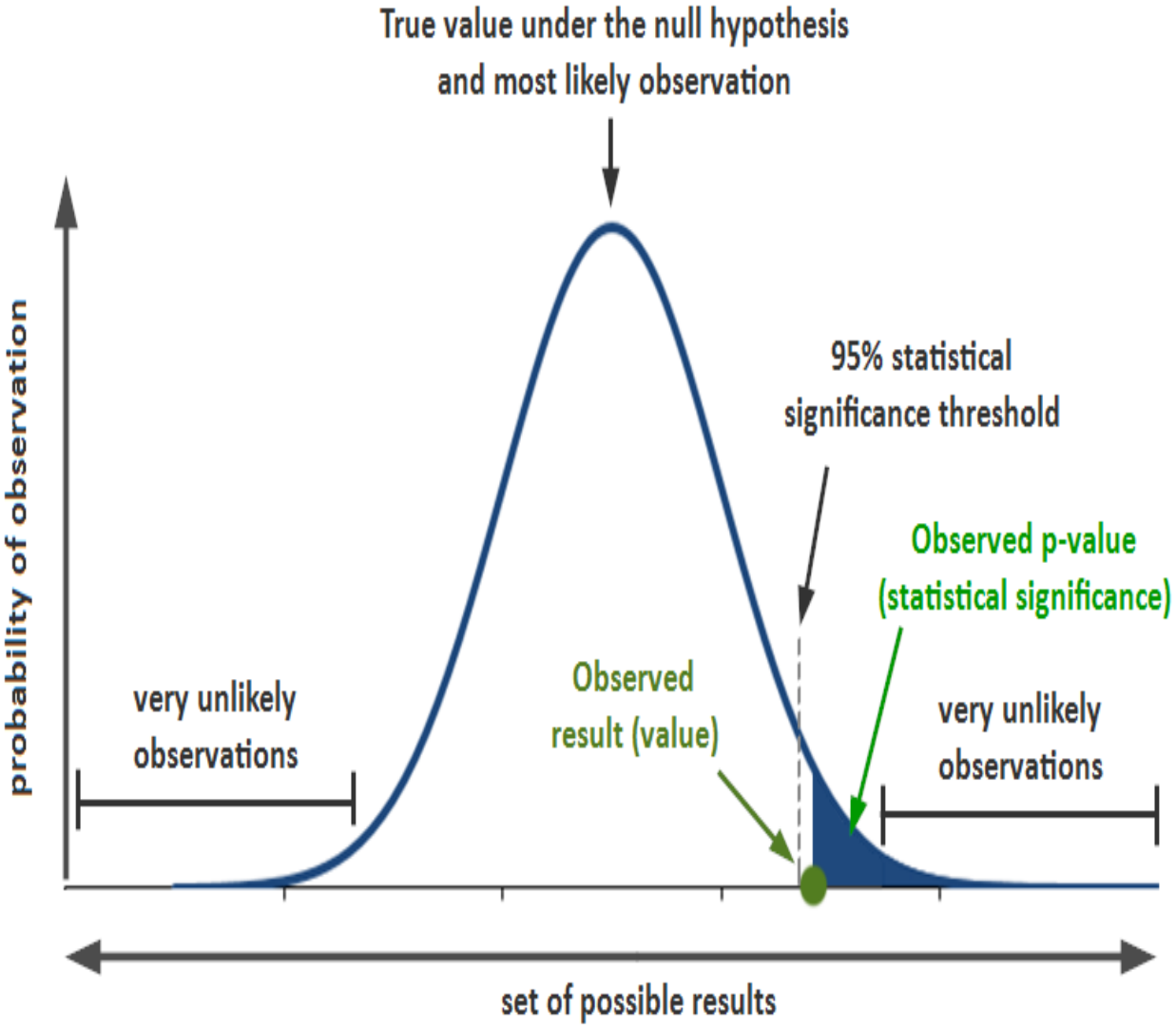
Testing

- **Parametric Testing:**

It is used when the assumed distribution is close to normal. The most widely used tests are t-test and ANOVA test.

- **Non-Parametric testing:**

It is used when the data is discrete or the continuous data is not normally distributed. Tests like Chi-squared test, Mann-Whitney U-test etc. Bootstrapping is also used to compare the means of control and variant samples

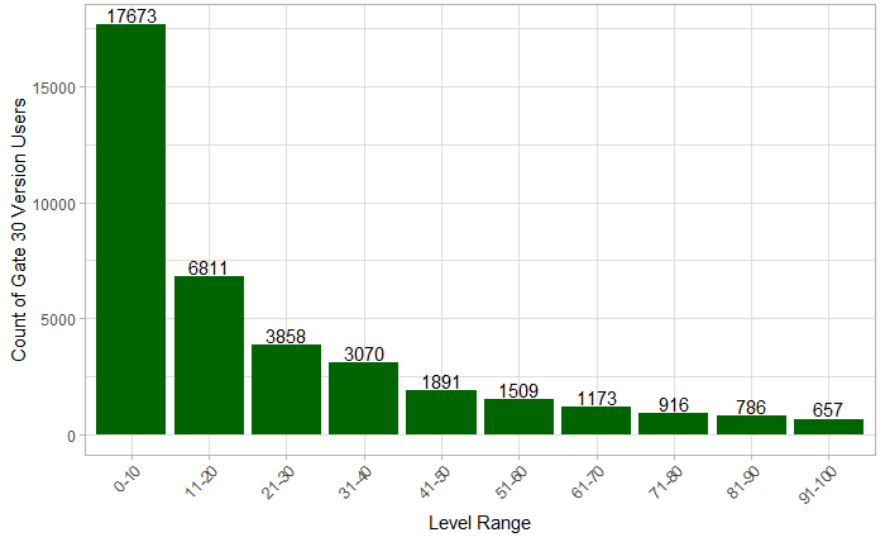


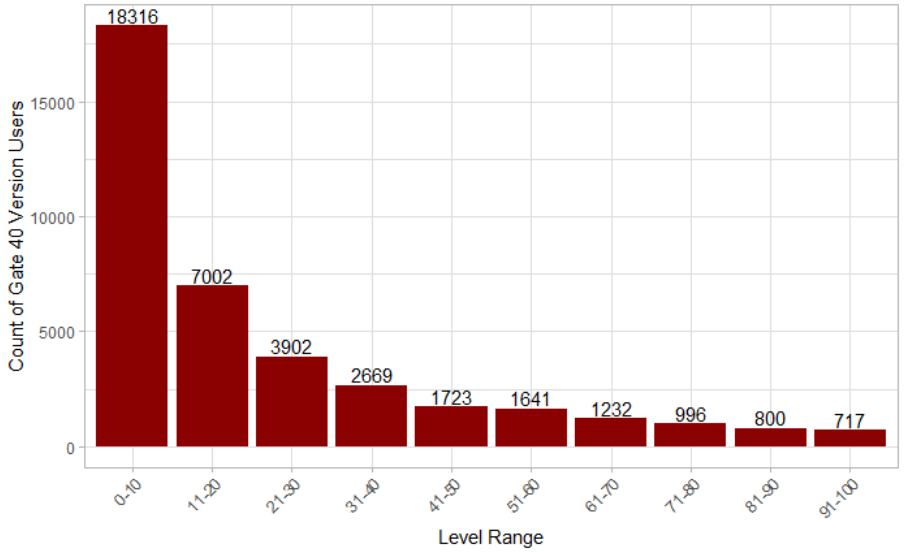
Problem Statement

- **About the game:** Cookie Cats is a popular mobile puzzle game developed by Tactile Entertainments. It's a connect the dots style puzzle game where the player must connect tiles of the same food item to clear the board and win the level by feeding the cats.
- **A/B testing problem:** The game has multiple levels and as the players progress, they will encounter gates that force them to collect keys before moving to next level or do in-app purchase to progress instantly. So, we have the gate at level 30, and we will test moving the gate at level 40. We will analyze the player retention and take action accordingly.

Game Data Description

- There are total of 90189 players tested for the analysis
- Total of 44700 players are of Gate 30 game mode and 45489 players are of Gate 40 game mode
- Count of users of different game modes in top 100 level distributions is shown in respective bar chart





Retention Analysis

1-Day Retention:

Overall, 44.52% 1-day retention is there

Version	Total Count	Retention Count	Retention Ratio
Gate_30	44700	20034	0.4481879
Gate_40	45489	20119	0.4422827

7-Day Retention:

Overall, 18.6% 7-day retention is there

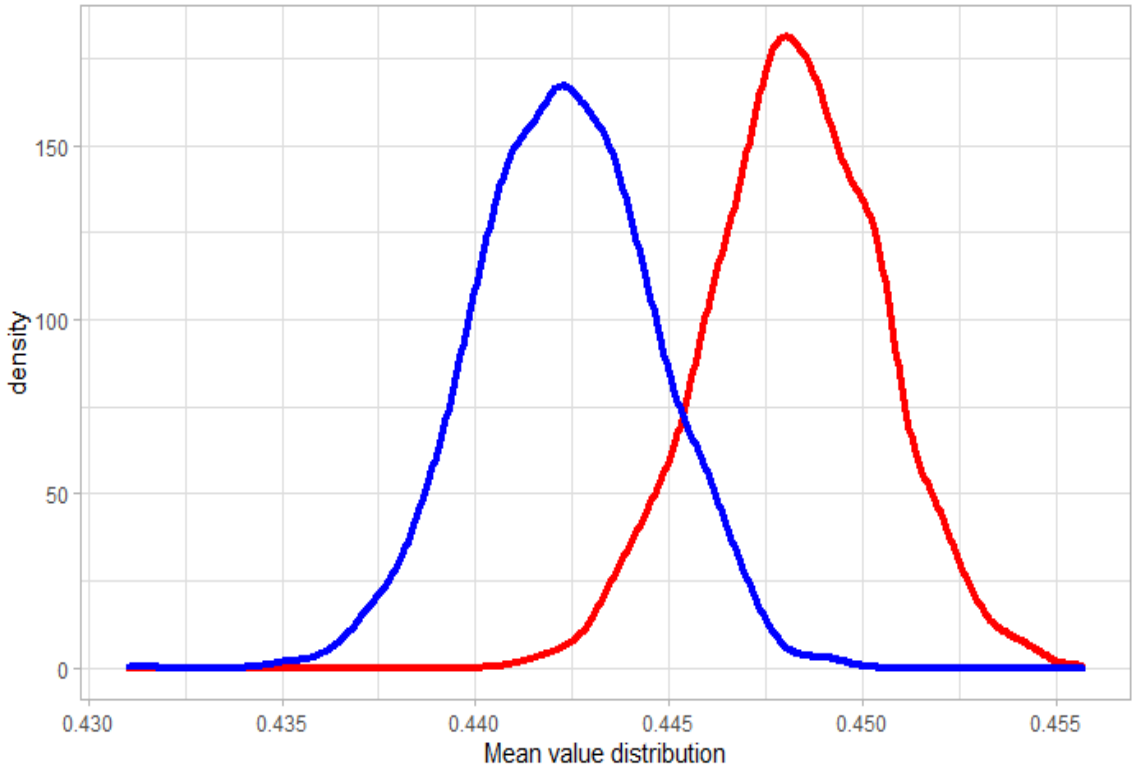
Version	Total Count	Retention Count	Retention Ratio
Gate_30	44700	8502	0.1902013
Gate_40	45489	8279	0.1820000

1-Day Retention Testing

- **Lift:** 1.32% (Lift in retention for Gate_30 with respect to Gate_40)
- **Mean Value Distribution:** Using bootstrap method, the obtained 95% CI are:

Version	2.5%	97.5%
Gate_30	0.4436683	0.4525951
Gate_40	0.4377762	0.4467893

- **Test of Proportions:** Using Chi-squared test to compare the two means, the p-value obtained is 0.0755 which is >0.05 and hence, the null hypothesis that the two proportions are same cannot be rejected

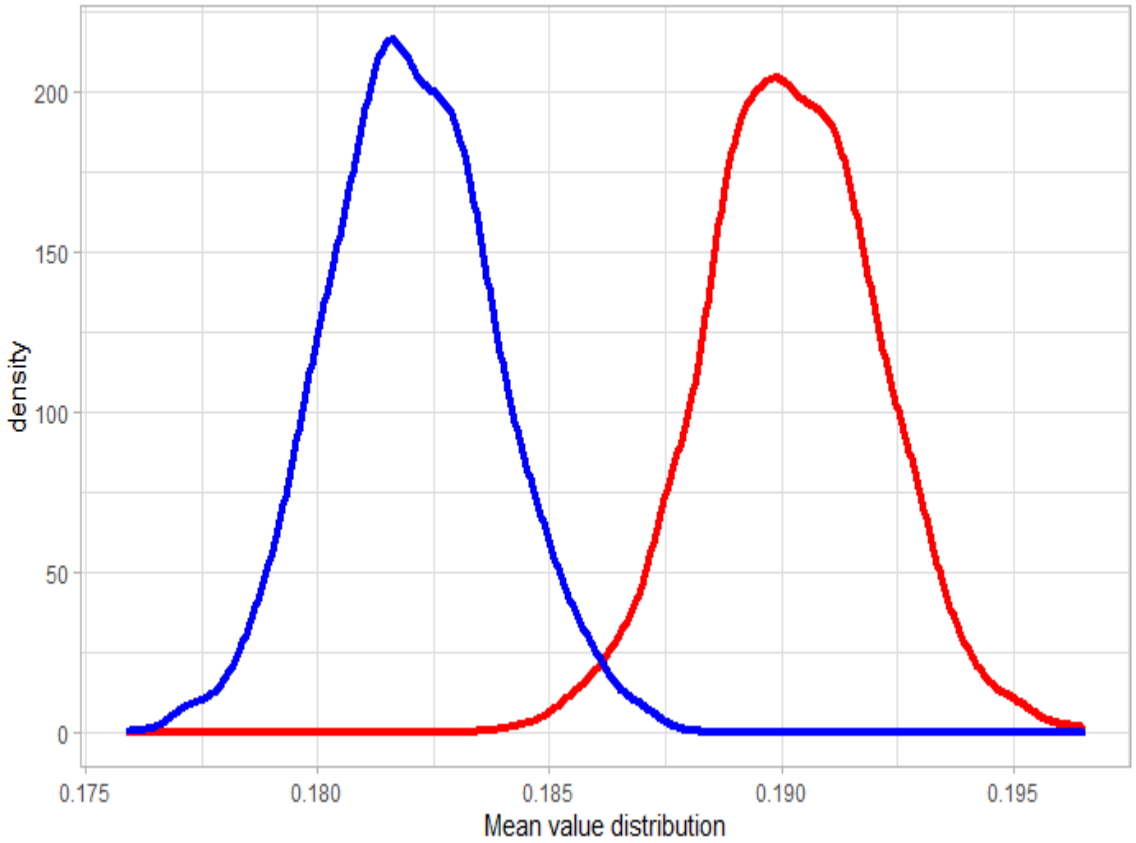


7-Day Retention Testing

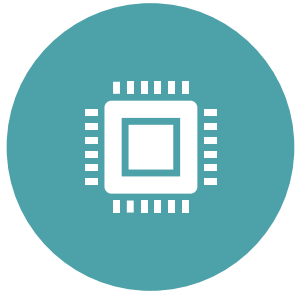
- **Lift:** 4.5% (Lift in retention for Gate_30 with respect to Gate_40)
- **Mean Value Distribution:** Using bootstrap method, the obtained 95% CI are:

Version	2.5%	97.5%
Gate_30	0.1864871	0.1939161
Gate_40	0.1785706	0.1857812

- **Test of Proportions:** Using Chi-squared test to compare the two means, the p-value obtained is 0.0016 which is <0.05 and hence, the null hypothesis that the two proportions are same can be rejected and we can conclude that the two proportions are significantly different.



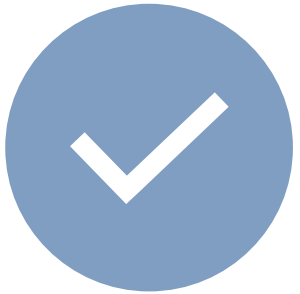
Conclusion



The A/B testing of the two versions of the game suggests that placing of Gate at level 40 does not result in any improvement in retention. Rather, the retention rates are slightly higher for gate at level 30

The reason could be that players may find it more interesting when they get challenges at level 30 as compared to level 40 which can keep them more engaged with the game

The 1-day retention is similar for keeping gates at both the levels whereas, the 7-day retention is 4.5% more in case of having gates at level 30 than at level 40



Thus, we can conclude that it is better to not move the gates to level 40 and keep it at level 30

References

https://medium.com/swlh/the-ultimate-guide-to-a-b-testing-part-1-experiment-design-8315a2470c63

https://www.analyticsvidhya.com/blog/2020/10/ab-testing-data-science/

https://blog.analytics-toolkit.com/2017/statistical-significance-ab-testing-complete-guide/2017-09-11-statistical-significance-p-value-2/

http://www.sthda.com/english/wiki/two-proportions-z-test-in-r