Part 2. The Gotham-Metropolis Bridge

Setup

The neighboring cities of Gotham and Metropolis have complementary circadian rhythms: on weekdays, Ultimate Gotham is most active at night, and Ultimate Metropolis is most active during the day. On weekends, there is reasonable activity in both cities. However, a toll bridge, with a two way toll, between the two cities causes driver partners to tend to be exclusive to each city. The Ultimate managers of city operations for the two cities have proposed an experiment to encourage driver partners to be available in both cities, by reimbursing all toll costs.

Question 1

What would you choose as the key measure of success of this experiment in encouraging driver partners to serve both cities, and why would you choose this metric?

I would use a metric that tracks the city of origin of a ride for each driver, and compare the ratio of Gotham to Metropolis ride origins *on the weekends* from before and after the change.

Question 2

Describe a practical experiment you would design to compare the effectiveness of the proposed change in relation to the key measure of success. Please provide details on:

a.how you will implement the experiment
b.what statistical test(s) you will conduct to verify the significance of the observation
c.how you would interpret the results and provide recommendations to the city operations team along with any caveats.

I would perform an A/B test style experiment to compare the effectiveness of the toll reimbursement. The drivers who are a part of the experiment should be ones who drive frequently on the weekend. We will do a semi-random split to ensure that there are similar metrics between groups in terms of:

- Split between which city they frequent
- Even distribution of hours driven, trips made, etc
- Even distribution of times of day that the driver operates

Half the group would be told they will get the toll reimbursements for the weekend, while the second half would operate as usual. This test would probably need to be performed over several weekends, and each weekend can have its own groups.

A straightforward way to test the significance of any differences we observe would be to do a permutation test. Once we have the trip data, we could randomly sort the weekend trip city ratios and see if we still observe a similar difference (meaning the difference likely occurred by chance) or if the observed difference is so unlikely to have happened by chance that it is almost certainly a real change in driver behavior.

If the difference proves to be significant, I would recommend that city operations move forward with the change. However, I would also recommend that after a certain trial period that the results be re-evaluated to make sure that the program remains effective among the larger driver population, and that more drivers having access to the toll reimbursement doesn't in some way devalue the fact, causing drivers to go back to their usual patterns.