Case 1 : Digital Advertising

Talha Amin, Ziyue Wang, Yeyi Su, Evan Sun

2025-10-06

Executive Summary

Oz Collectibles recently conducted an experiment running three Instagram advertisements—A, B, and C—to determine if these ads generated higher conversions and revenue compared to a control group with no ads. The experiment included 100,000 participants, evenly split into four groups of 25,000 each. Our goal was to identify which ad should be continued based on performance.

We calculated the proportion of conversions for each group and conducted two-sample proportion tests, using a two-sided approach to assess whether ads performed better or worse than the control. Applying Bonferroni's correction, we multiplied p-values by 3. All ads showed significantly higher conversion rates than the control group: **A (5.0%), B (3.1%), and C (4.4%)** versus **2.6%** for control ($p < 0.001$).

We also conducted t-tests comparing revenue among converters. Ads A and C ($p < 0.01$) generated significantly higher revenue than control, while Ad B did not ($p = 2.55$). After calculating cost per incremental conversion, profit per incremental conversion, and ROI, we concluded that Ads **A (ROI = 47.9%)** and **C (ROI = 40.8%)** should be continued, while **B (ROI = −69.2%)** should be discontinued.

Glinda also explored whether a multi-armed bandit approach could have identified top ads earlier. This method reallocates impressions toward better-performing ads over time.Calculating a baseline conversion rate of 2.6% and a minimum detectable effect of 20%, Ad B was identified as the weakest, while A and C performed significantly better. Using a bandit approach, the experiment could have stopped around **60,000 total impressions (15,000 per condition)**, reallocating 40% of sample, avoiding **166** lost conversions, and gaining approximately **$1,961** in additional profit through faster learning and optimized allocation.

Analysis

```
load('IGAds.Rdata')

str(IGAds)
```

```
## 'data.frame':    100000 obs. of  4 variables:
##  $ Observation: int  1 2 3 4 5 6 7 8 9 10 ...
##  $ Condition  : chr  "A" "B" "C" "Control" ...
##  $ Conversion : int  0 0 0 0 0 0 0 0 0 0 ...
##  $ Revenue    : num  0 0 0 0 0 0 0 0 0 0 ...
```

```
head(IGAds)
```

```
##   Observation Condition Conversion Revenue
## 1           1         A          0       0
```

```
## 2           2           B           0           0
## 3           3           C           0           0
## 4           4     Control           0           0
## 5           5           A           0           0
## 6           6           B           0           0
```

Question 1

Treatment Effect by each Advertisement

We first run different proportions tests on the conversions through each advertisement and control and calculate the treatment effect for each ad.

```r
nbin = 25000

A <- IGAds[IGAds$Condition == 'A',]
B <- IGAds[IGAds$Condition == 'B',]
C <- IGAds[IGAds$Condition == 'C',]
Control <- IGAds[IGAds$Condition == 'Control',]

countA <- sum(A$Conversion)
countB <- sum(B$Conversion)
countC <- sum(C$Conversion)
countCountrol <- sum(Control$Conversion)

### Two-sample proportions test (one-sided) for conversion: Is ad > control?
a_ab <- prop.test(x=c(countA,countCountrol), n=c(nbin,nbin), alternative = 'greater')
b_ab <- prop.test(x=c(countB,countCountrol), n=c(nbin,nbin), alternative = 'greater')
c_ab <- prop.test(x=c(countC,countCountrol), n=c(nbin,nbin), alternative = 'greater')

# Treatment effect

a_ab$estimate[1] - a_ab$estimate[2]
```

```
## prop 1
## 0.0248
```

```r
b_ab$estimate[1] - b_ab$estimate[2]
```

```
##  prop 1
## 0.00532
```

```r
c_ab$estimate[1] - c_ab$estimate[2]
```

```
##  prop 1
## 0.01904
```

```r
# Test output
a_ab
```

```
## 
##  2-sample test for equality of proportions with continuity correction
## 
## data:  c(countA, countCountrol) out of c(nbin, nbin)
## X-squared = 209.21, df = 1, p-value < 2.2e-16
## alternative hypothesis: greater
## 95 percent confidence interval:
##  0.02195019 1.00000000
## sample estimates:
##  prop 1  prop 2
## 0.05048 0.02568
```

b_ab

```
## 
##  2-sample test for equality of proportions with continuity correction
## 
## data:  c(countB, countCountrol) out of c(nbin, nbin)
## X-squared = 12.655, df = 1, p-value = 0.0001873
## alternative hypothesis: greater
## 95 percent confidence interval:
##  0.002838968 1.000000000
## sample estimates:
##  prop 1  prop 2
## 0.03100 0.02568
```

c_ab

```
## 
##  2-sample test for equality of proportions with continuity correction
## 
## data:  c(countC, countCountrol) out of c(nbin, nbin)
## X-squared = 132.87, df = 1, p-value < 2.2e-16
## alternative hypothesis: greater
## 95 percent confidence interval:
##  0.01629242 1.00000000
## sample estimates:
##  prop 1  prop 2
## 0.04472 0.02568
```

All the ads have significantly higher conversion rate than Control Group ($p < 0.05/3 = 0.0167$).

Two-Sample t-tests for Revenue

Now, we conduct different t-tests to compare the revenues generated by each ad as compared to the control.

```
A_rev <- A$Revenue[A$Conversion == 1]
B_rev <- B$Revenue[B$Conversion == 1]
C_rev <- C$Revenue[C$Conversion == 1]
Contorl_rev <- Control$Revenue[Control$Conversion == 1]

a_t <- t.test(A_rev, Contorl_rev, alternative = 'two.sided', var.equal = FALSE)
```

```
b_t <- t.test(B_rev, Contorl_rev, alternative = 'two.sided', var.equal = FALSE)
c_t <- t.test(C_rev, Contorl_rev, alternative = 'two.sided', var.equal = FALSE)

a_t$p.value*3
```

```
## [1] 0.002235682
```

```
b_t$p.value*3
```

```
## [1] 2.545441
```

```
c_t$p.value*3
```

```
## [1] 2.11795e-161
```

Among converted customers, Ads A and C have significantly higher average revenue than Control Group ($p < 0.05/3 = 0.0167$).

ROI Through each Advertisement

```
## Computing ROI for each advertisement
c = 0.20
CPIC_a <- c / (a_ab$estimate[1] - a_ab$estimate[2])
CPIC_b <- c / (b_ab$estimate[1] - b_ab$estimate[2])
CPIC_c <- c / (c_ab$estimate[1] - c_ab$estimate[2])

piPIC_a <- mean(A_rev)
piPIC_b <- mean(B_rev)
piPIC_c <- mean(C_rev)

ROI_a <- (piPIC_a - CPIC_a) / CPIC_a
ROI_b <- (piPIC_b - CPIC_b) / CPIC_b
ROI_c <- (piPIC_c - CPIC_c) / CPIC_c

ROI_a
```

```
##    prop 1
## 0.4791569
```

```
ROI_b
```

```
##     prop 1
## -0.6921265
```

```
ROI_c
```

```
##    prop 1
## 0.4079041
```

Advertisement A and C have positive ROIs, therefore we should continue to use A and C.

Question 2

https://www.optimizely.com/sample-size-calculator/#/?conversion=2.568&effect=20&significance=95

In order to detect minimum detectable effect, we decide to stop the experiment at 60,000 people (15,000 for each Ad and control group) through the above website. Assume we found that Ad B performs worst.

Therefore, for the rest of 30,000 people (excluding 10,000 control group), we decide to exclude Ad B and evenly distribute it among Ad A and C.

```r
con_a <- countA / nbin
con_b <- countB / nbin
con_c <- countC / nbin

con_ABC <- (con_a*10000 + con_b*10000 + con_c*10000)
con_excludeB <- (con_a*15000 + con_c*15000)

average_piPIC <- mean(piPIC_a,piPIC_b,piPIC_c)

lost_conversion <- con_excludeB - con_ABC
lost_conversion
```

```
## [1] 166
```

```r
additional_profit <- (con_excludeB - con_ABC) * average_piPIC
additional_profit
```

```
## [1] 1980.162
```

Therefore, for the remaining 30,000 people, if we had eliminated the advertisement B earlier by applying multi-armed bandit, we could have earned an additional 166 coneversions and profit of approximately $1,961.