BA860 Assignment 2

Ying Wu, Yixuan Wang, Ruchika Venkateswaran

06/21/2021

```r
# load dataset
df <- fread('/Users/wangyixuan/Downloads/PS1-4003988.csv')
```

1. What is the standard deviation of sales (not past_sales) in the treatment group among exposed users who saw the campaign in "PS1-4003988.csv" (from Assignment #1)? Use this number for your calculations below. (5 points)

```r
sd1 <- sd(df[Treatment == 1 & saw_ads==1, sales])
round(sd1,2)
```

```
## [1] 6.27
```

The standard deviation of sales in the treatment group among exposed users who saw the campaign is 6.27.

2. Suppose you have the budget to spend $0.01 per person on advertising. As your 'reasonable signal' or benchmark ad lift, you assume that the campaign will break even (0 profit).

a. If your profitability margin on sales is 50%, what is the 'reasonable signal' for ad lift that you are assuming? (2 points)

```r
camp_lift <- (500000*0.01)/0.5
camp_lift_per_person <- camp_lift/500000
round(camp_lift_per_person,2)
```

```
## [1] 0.02
```

The 'reasonable signal' for ad lift that we are assuming is 0.02.

b. For a 95% confidence interval, calculate the statistical power to successfully detect the signal for the following three experimental designs (9 points):

i. 20% of users are assigned to the control group.

```r
power1 <- pwr.t2n.test(d=camp_lift_per_person/sd1, n1=0.2*500000, n2=0.8*500000)
power1
```

```
##
##      t test power calculation
##
##              n1 = 1e+05
##              n2 = 4e+05
##               d = 0.003187707
##       sig.level = 0.05
##           power = 0.1470565
##     alternative = two.sided
```

The statistical power is 14.71% for 20% of users are assigned to the control group.

ii. 30% of users are assigned to the control group.

```
power2 <- pwr.t2n.test(d=camp_lift_per_person/sd1, n1=0.3*500000, n2=0.7*500000)
power2
```

```
##
##      t test power calculation
##
##              n1 = 150000
##              n2 = 350000
##               d = 0.003187707
##       sig.level = 0.05
##           power = 0.178337
##     alternative = two.sided
```

The statistical power is 17.83% for 30% of users are assigned to the control group.

iii. 80% of users are assigned to the control group.

```
power3 <- pwr.t2n.test(d=camp_lift_per_person/sd1, n1=0.8*500000, n2=0.2*500000)
power3
```

```
##
##      t test power calculation
##
##              n1 = 4e+05
##              n2 = 1e+05
##               d = 0.003187707
##       sig.level = 0.05
##           power = 0.1470565
##     alternative = two.sided
```

The statistical power is 14.71% for 80% of users are assigned to the control group.

Which design is best according to this criterion? (1 point)

The design with 30% of the users assigned to the control group is the best because its statistical power is higher than the other 2 designs with 20% and 80% of the users assigned to the control group respectively (the statistical power for the design with 30% of the users assigned to the control group is 3.12% higher than that of the other two designs). That's why the one with 30% of users assigned to the control group is the best according to the statistical power.

c. Design i – iii may not be the 'optimal' design. What treatment assignment (%assigned to the control group) maximizes statistical power? (5 points)

```
power4 <- pwr.t2n.test(d=camp_lift_per_person/sd1, n1=0.5*500000, n2=0.5*500000)
power4
```

```
##
##      t test power calculation
##
##              n1 = 250000
##              n2 = 250000
##               d = 0.003187707
##       sig.level = 0.05
##           power = 0.2034498
##     alternative = two.sided
```

The optimal design that we recommend is allocating a 50-50 split between the assignment of people. This is because when the treatment & control have equal variances, a 50-50 split gives you the maximum power for a given sample size. The statistical power is around 20.34%.

3. Now suppose instead that you have a fixed ad budget of $2,000 to spend on your own ads. You ignore the cost of control ads because your partner publisher is providing them for free.

a. Calculate the average ad spend per person for the three experimental designs above. Report your answer to 4 decimals. (6 points)

```
e1 <- round((2000/(0.8*500000)),4)
e2 <- round((2000/(0.7*500000)),4)
e3 <- round((2000/(0.2*500000)),4)

cat('The average ad spend per person for the 1st experiment with 80% of the users
    assigned to the treatment group is: $', round(e1,4),'\n',
    'The average ad spend per person for the 2nd experiment with 70% of the users
    assigned to the treatment group is: $', round(e2,4),'\n',
    'The average ad spend per person for the 3rd experiment with 20% of the users
    assigned to the treatment group is: $', round(e3,4),'\n')
```

```
## The average ad spend per person for the 1st experiment with 80% of the users
##     assigned to the treatment group is: $ 0.005
##  The average ad spend per person for the 2nd experiment with 70% of the users
##     assigned to the treatment group is: $ 0.0057
##  The average ad spend per person for the 3rd experiment with 20% of the users
##     assigned to the treatment group is: $ 0.02
```

b. Maintaining your assumption that the campaign will break-even for each experimental design, what is the 'reasonable signal' for each of the three designs above? Report your answer to 4 decimals. (6 points)

```r
lift1 <- 2000/0.5

incre1 <- lift1/(0.8*500000)
incre2 <- lift1/(0.7*500000)
incre3 <- lift1/(0.2*500000)

cat('The ‘reasonable signal’ for the 1st experiment with 80% of the users
    assigned to the treatment group is: ', round(incre1,4),'\n',
    'The ‘reasonable signal’ for the 2nd experiment with 70% of the users
    assigned to the treatment group is: ', round(incre2,4),'\n',
    'The ‘reasonable signal’ for the 3rd experiment with 20% of the users
    assigned to the treatment group is: ', round(incre3,4),'\n')
```

```
## The ‘reasonable signal’ for the 1st experiment with 80% of the users
##     assigned to the treatment group is:  0.01
##  The ‘reasonable signal’ for the 2nd experiment with 70% of the users
##     assigned to the treatment group is:  0.0114
##  The ‘reasonable signal’ for the 3rd experiment with 20% of the users
##     assigned to the treatment group is:  0.04
```

c. For a 95% confidence interval, what is the statistical power now for each of the three possible experimental designs? Which is highest now? (9 points)

```r
power_e1 <- pwr.t2n.test(d=incre1/sd1, n1=0.2*500000, n2=0.8*500000)
power_e2 <- pwr.t2n.test(d=incre2/sd1, n1=0.3*500000, n2=0.7*500000)
power_e3 <- pwr.t2n.test(d=incre3/sd1, n1=0.8*500000, n2=0.2*500000)

power_e1
```

```
##
##      t test power calculation
##
##              n1 = 1e+05
##              n2 = 4e+05
##               d = 0.001593854
##       sig.level = 0.05
##           power = 0.07358897
##     alternative = two.sided
```

```r
power_e2
```

```
##
##      t test power calculation
##
##              n1 = 150000
##              n2 = 350000
##               d = 0.001821547
##       sig.level = 0.05
##           power = 0.09077059
##     alternative = two.sided
```

```
power_e3
```

```
##
##      t test power calculation
##
##              n1 = 4e+05
##              n2 = 1e+05
##               d = 0.006375414
##       sig.level = 0.05
##           power = 0.4378135
##     alternative = two.sided
```

```
cat('The statistical power for the 1st experiment with 80% of the users
    assigned to the treatment group is: ', percent(power_e1$power, accuracy=0.01),'\n',
    'The statistical power for the 2nd experiment with 70% of the users
    assigned to the treatment group is: ', percent(power_e2$power, accuracy=0.01),'\n',
    'The statistical power for the 3rd experiment with 20% of the users
    assigned to the treatment group is: ', percent(power_e3$power, accuracy=0.01),'\n')
```

```
## The statistical power for the 1st experiment with 80% of the users
##     assigned to the treatment group is:  7.36%
##  The statistical power for the 2nd experiment with 70% of the users
##     assigned to the treatment group is:  9.08%
##  The statistical power for the 3rd experiment with 20% of the users
##     assigned to the treatment group is:  43.78%
```

The 3rd experiment with 20% of the users assigned to the treatment group has the highest statistical power (43.78%).

d. What changes your answer between 2(b) and 3(c)? Which is the best of the six options in terms of statistical power? (5 points)

The statistical power decreases for the first two experimental designs but increases for the third experiment. It decreases by 7.35% for the first experiment, decreases by 8.76% for the second experiment but increases by 29.07% for the third experiment. The experiment with 20% of the users in the treatment group, given that a fixed ad budget of $2,000 is spent on ads is the most successful experiment with the highest statistical power at 43.78%.

4. You have been thinking about the possibility your ads may wear out so that their effectiveness decreases as you increase the average ad spend. You revisit your previous assumption and instead assume the following for the 3 possible experimental designs:

a. What is the 'reasonable' signal now? Fill out the above table. Report your answer to 4 decimals. (6 points)

```
e1_new <- e1*2
e2_new <- e2*1.9
e3_new <- e3*1.2
```

```r
incre1_new <- round((e1_new/0.5),4)
incre2_new <- round((e2_new/0.5),4)
incre3_new <- round((e3_new/0.5),4)

tab <- matrix(c('From 3(a)', '2x cost=0.02', 'From 3(a)', '1.9x cost=0.0217',
                'From 3(a)', '1.2x cost=0.048'), ncol=2, byrow=TRUE)
colnames(tab) <- c('Cost Per Person','Reasonable signal')
rownames(tab) <- c('20% control','30% control','80% control')
tab <- as.table(tab)
tab
```

```
##             Cost Per Person Reasonable signal
## 20% control From 3(a)       2x cost=0.02
## 30% control From 3(a)       1.9x cost=0.0217
## 80% control From 3(a)       1.2x cost=0.048
```

```r
cat('The ‘reasonable signal’ for the 1st experiment with 80% of the users
    assigned to the treatment group is: ', incre1_new,'\n',
    'The ‘reasonable signal’ for the 2nd experiment with 70% of the users
    assigned to the treatment group is: ', incre2_new,'\n',
    'The ‘reasonable signal’ for the 3rd experiment with 20% of the users
    assigned to the treatment group is: ', incre3_new,'\n')
```

```
## The ‘reasonable signal’ for the 1st experiment with 80% of the users
##     assigned to the treatment group is:  0.02
##  The ‘reasonable signal’ for the 2nd experiment with 70% of the users
##     assigned to the treatment group is:  0.0217
##  The ‘reasonable signal’ for the 3rd experiment with 20% of the users
##     assigned to the treatment group is:  0.048
```

b. Under your revised assumption, what is the statistical power of each option? (9 points) Which is best now? (1 point)

```r
power_e1_new <- pwr.t2n.test(d=incre1_new/sd1, n1=0.2*500000, n2=0.8*500000)
power_e2_new <- pwr.t2n.test(d=incre2_new/sd1, n1=0.3*500000, n2=0.7*500000)
power_e3_new <- pwr.t2n.test(d=incre3_new/sd1, n1=0.8*500000, n2=0.2*500000)

power_e1_new
```

```
##
##      t test power calculation
##
##              n1 = 1e+05
##              n2 = 4e+05
##               d = 0.003187707
##       sig.level = 0.05
##           power = 0.1470565
##     alternative = two.sided
```

```
power_e2_new
```

```
##
##      t test power calculation
##
##              n1 = 150000
##              n2 = 350000
##               d = 0.003458662
##       sig.level = 0.05
##           power = 0.2017023
##     alternative = two.sided
```

```
power_e3_new
```

```
##
##      t test power calculation
##
##              n1 = 4e+05
##              n2 = 1e+05
##               d = 0.007650497
##       sig.level = 0.05
##           power = 0.5808103
##     alternative = two.sided
```

```
cat('The statistical power for the 1st experiment with 80% of the users assigned to
    the treatment group is: ', percent(power_e1_new$power, accuracy =0.01),'\n',
    'The statistical power for the 2nd experiment with 70% of the users assigned to
    the treatment group is: ', percent(power_e2_new$power, accuracy =0.01),'\n',
    'The statistical power for the 3rd experiment with 20% of the users assigned to
    the treatment group is: ', percent(power_e3_new$power, accuracy =0.01),'\n')
```

```
## The statistical power for the 1st experiment with 80% of the users assigned to
##     the treatment group is:  14.71%
##  The statistical power for the 2nd experiment with 70% of the users assigned to
##     the treatment group is:  20.17%
##  The statistical power for the 3rd experiment with 20% of the users assigned to
##     the treatment group is:  58.08%
```

The experiment with 20% of the users in the treatment group has the highest statistical power (58.08%) among the three experiments.

5.a. You budget $40,000 weekly for the paid search campaign. For your power calculation, you expect the search ads to generate a lift equivalent to 3X the ad cost. What is the expected ad lift for the average store (the 'reasonable' signal)? (2 points)

```
lift <- 40000*3
incre <- lift/30
cat('The expected ad lift for average store (`reasonable' signal) is: $', incre,'\n')
```

```
## The expected ad lift for average store (‘reasonable’ signal) is: $ 4000
```

5.b. Across the 60 locations, the stores have the same average weekly sales of $200,000 with standard deviation $30,000. Given this and using a 90% confidence interval as your standard, calculate the statistical power for this experiment if you run it for a single week. (6 points)

```
power1 <- pwr.t2n.test(d=incre/30000, n1=30, n2=30, sig.level = 0.1)
power1
```

```
##
##      t test power calculation
##
##             n1 = 30
##             n2 = 30
##              d = 0.1333333
##      sig.level = 0.1
##          power = 0.143871
##    alternative = two.sided
```

```
cat('The statistical power for this experiment is: ',
    percent(power1$power,accuracy =0.01),'\n')
```

```
## The statistical power for this experiment is:  14.39%
```

5.c. At current spending levels, how many weeks would you need to run this experiment before the statistical power of the experiment surpasses 50%? (8 points)

```
power2 <- pwr.t2n.test(n1=30, n2=30, power=0.5, sig.level = 0.1)
week <- (power2$d)/power1$d
week
```

```
## [1] 3.220359
```

```
cat('It would take about', round(week,0),'weeks to run this experiment before
    the statistical power of the experiment surpasses 50%', '\n')
```

```
## It would take about 3 weeks to run this experiment before
##     the statistical power of the experiment surpasses 50%
```

6.a. Given the response rates, how many surveys do you expect to collect in each of the treatment and control groups? (2 points)

```
survey_treat <- 300000*0.2/100
survey_control <- 100000*0.2/100
cat('There are', survey_treat,'surveys that we expect to collect in the treatment group', '\n',
    'There are', survey_control,'surveys that we expect to collect in the control group', '\n')
```

```
## There are 600 surveys that we expect to collect in the treatment group
##  There are 200 surveys that we expect to collect in the control group
```

6.b. DynamicLogic survey proposal is expensive and you can only pay them for this one ad campaign. What is the likelihood that the 95% confidence intervals on your ad effect estimates exclude 0 for each survey measure?

i. Brand favorability. (3 points)

```
power1 <- pwr.t2n.test(d=4.1*0.02/0.68, n1=survey_treat, n2=survey_control)
power1
```

```
##
##      t test power calculation
##
##               n1 = 600
##               n2 = 200
##                d = 0.1205882
##        sig.level = 0.05
##            power = 0.31419
##      alternative = two.sided
```

```
cat('The likelihood for brand favorability is: ',
        percent(power1$power, accuracy = 0.01),'\n')
```

```
## The likelihood for brand favorability is:  31.42%
```

ii. Purchase intent (3 points)

```
power2 <- pwr.t2n.test(d=1.6*0.04/1.23, n1=survey_treat, n2=survey_control)
power2
```

```
##
##      t test power calculation
##
##               n1 = 600
##               n2 = 200
##                d = 0.05203252
##        sig.level = 0.05
##            power = 0.09754985
##      alternative = two.sided
```

```
cat('The likelihood for intention to purchase is: ',
        percent(power2$power, accuracy = 0.01),'\n')
```

```
## The likelihood for intention to purchase is:  9.75%
```

6.c. In a one-paragraph summary for a coworker, explain your decision on whether to use DynamicLogic's services. Be sure to discuss the strengths and limitations of survey analysis and explain implications of your power calculation for the experiment. (6 points)

The final recommendation is that IAMS should proceed with DynamicLogic's services. Since 90% of the company's sales occurs in stores and only a small percentage of sales occurs through online channels, the key objective as a CMO of IAMS is to increase the company's brand awareness online. The power analysis also shows that the brand favorability is significantly higher than that of the intention to purchase (the statistical power for brand favorability is 21.67% higher than that of the purchase intent). The results of our power analysis supports the company's objectives in increasing brand awareness and at this juncture, we recommend proceeding with DynamicLogic's services. The main limitation with this approach however is that the likelihood for purchase intent is low, but this can be improved by executing different campaigns with targeted promotions for customers to increase their purchase to intent.

6.d. Suppose that 0.5% of users who see the treatment ad fill out the survey, but 0.1% of users who see the control ad fill the survey. How would this affect your interpretation of the experimental results? (6 points)

```
survey_treat2 <- 300000*0.5/100
survey_control2 <- 100000*0.1/100
cat('There are', survey_treat2,'surveys that collected in the treatment group', '\n',
    'There are', survey_control2,'surveys that collected in the control group', '\n')
```

```
## There are 1500 surveys that collected in the treatment group
##  There are 100 surveys that collected in the control group
```

```
power3 <- pwr.t2n.test(d=4.1*0.02/0.68, n1=survey_treat2, n2=survey_control2)
power4 <- pwr.t2n.test(d=1.6*0.04/1.23, n1=survey_treat2, n2=survey_control2)
power3
```

```
##
##      t test power calculation
##
##              n1 = 1500
##              n2 = 100
##               d = 0.1205882
##       sig.level = 0.05
##           power = 0.2147505
##     alternative = two.sided
```

```
power4
```

```
##
##      t test power calculation
##
##              n1 = 1500
##              n2 = 100
##               d = 0.05203252
##       sig.level = 0.05
##           power = 0.07951229
##     alternative = two.sided
```

```
cat('The statistical power for brand favorability is: ',
        percent(power3$power, accuracy = 0.01),'\n',
    'The statistical power for intention to purchase is: ',
        percent(power4$power, accuracy = 0.01),'\n')
```

```
## The statistical power for brand favorability is:  21.48%
##  The statistical power for intention to purchase is:  7.95%
```

```
brand_fav_change <- (power3$power-power1$power)/power1$power
purchase_inten_change <- (power4$power-power2$power)/power2$power
cat('The change for brand favorability is: ',
        percent(brand_fav_change, accuracy = 0.01),'\n',
    'The change for intention to purchase is: ',
        percent(purchase_inten_change, accuracy = 0.01),'\n')
```

```
## The change for brand favorability is:  -31.65%
##  The change for intention to purchase is:  -18.49%
```

The sample size increases when 0.5% of the users who see the treatment ad fill out the survey and 0.1% of the users who see the control ad fill the survey. When the sample size increases, the statistical power is also expected to increases, but the calculations above display that the statistical power decreases for both brand favorability and purchase intent. The brand favorability observed a 31.65% reduction in statistical power and the purchase intent observed a dip in statistical power by 18.49%. Since the statistical power for both brand favorability and purchase intent decreases even when the sample size increases, IAMS should not proceed with DynamicLogic's services and should reevaluate the experiment design.