Fast Food Marketing Analysis

Yipeng Wei (leader)

Zibo Hong, Zhicheng Ma, Qinyuan Xing, Haocheng Yu

Abstract

This project is aimed at determining better promotion for fast food chains based on the Response-Adaptive design and Covariate-Adaptive design. First, we discretize continuous covariate(Age of store) and response(Sales in thousand) for our analysis. Next, after discretization, we define the success promotion rate based on the summary of the sales data. Then, we re-design the experiment by using the response-adaptive design and covariate-adaptive design to compare the two promotions. For response-adaptive design, DBCD with $\gamma = 2$ and target allocation $\rho = \frac{\sqrt{P_1}}{\sqrt{P_1}+\sqrt{P_2}}$ is recommended since it has the largest power, and more restaurants are assigned to better promotion 1. For covariates-adaptive design, Pocock and Simon's design performs well in marginal imbalance. STR-PB performs well in within-stratum imbalance. HuHu procedure maintains a good balance from all overall imbalances, marginal imbalance, and within-stratum imbalance. For statistical inference, type I error is conservative for STR-PB, P-S, and HuHu except for the full linear model. When the linear model is fixed, CR, STR-PB, P-S, and HuHu have similar power.

1. Introduction

The fast-food marketing campaign is always seeking a higher rate of promotion success for new products. The target people, the way of promotion, and even the time they send the promotion can be critical to the efficiency of the final promotion results. The data contain three promotions and the corresponding results. For simplicity, we only use promotions 1 and 2 in our analysis.

First, we discretize the continuous covariate and continuous response into categorical variables. Then, after discretization, we re-design the experiment by using the response-adaptive design and covariate-adaptive design to compare the two promotions. We use complete randomization, randomized play-the-winner rule, DBCD with different parameters and ERADE in the response-adaptive design. The estimated success rate is from the original data and is used in the response-adaptive design. The number of restaurants assigned to promotion 1 and the power of the tests are evaluated based on 1000 replications. Next, we use complete randomization, Stratified permuted block design, Pocock and Simon's procedures, and HuHu's procedure in the covariates-adaptive design. The overall imbalance, overall imbalance, and with-in stratum imbalance are evaluated. The type I error and power based on the linear model are also assessed based on 1000 replications for these 4 procedures.

2. Explore Data Analysis

2.1 Data Information

In the dataset, each observation contains the following variables:

- **Market size**: the size of the market area by sales
- **Age of store**: age of store in years
- **Promotion**: one of three promotions that were tested (Only promotions 1 and 2 are used in our analysis)
- **Week**: one of four weeks when the promotions happened
- **Sales in thousands**: sales amount for a specific promotion and week

Among those variables, Age of stores is a continuous covariate. Market size, Week are discrete covariates. Promotion is the treatment effect, and Sales in thousands is the response variable.

2.2 Data preprocessing and summary statistics

First, we discretize Age of Stores (covariate), and the Sales in thousands (response) for analysis.

For Age of Stores, according to the summary statistics, we categorize it into four strata based on the 1^{st} quarter, median, and 3^{rd} quarter. The discretization is as follows:

Table 1 Discretization of Age of Stores

Original	[0,4]	(4,7]	(7,12]	(12,∞)
Strata	0	1	2	3

To examine whether a promotion is successful, we also categorize the Sales in thousands (response variable) into successful promotion group and unsuccessful promotion group. Based on the summary statistics of the Sales in thousands, 50.2 is the median of all observations. For the promotion that finally has sales of more than 50.2, we regard it as a successful promotion, otherwise, it will be an unsuccessful promotion. The discretization for the Sales in thousands is as follows:

Table 2 Discretization of Sales in thousands

Original	[0,50.2]	(50.2,∞)
Indicator	0	1

The following table shows summary statistics for promotions 1 and 2.

Table 3 Summary Statistics for promotions

Promotion	sample size	Sales (μ)	Sales (p)
1	172	58.10	0.65
2	188	47.33	0.32

From table 3 we can intuitively see the significant difference between promotion 1 and promotion 2 in terms of their average sales amount and successful promotion rate with similar sample sizes.

Next, we will re-design the experiment by using the response-adaptive design and covariate-adaptive design to compare the two promotions.

3. Response-Adaptive Design

3.1 Outline of the design

An adaptable design is described as a design that permits changes to the trial's statistical processes after the trial has begun without jeopardizing the validity and integrity of the trial. To place more customers into a better promotion plan and ensure the validity of the experiment, we use the response-adaptive design with the following requirements. $\hat{p}_1 = 0.65$ and $\hat{p}_2 = 0.32$ from the estimation of the original data are implemented as the success rate for the adaptive-response design:

- Complete randomization with $p = 0.5$
- Randomized play-the-winner rule starts with (1,1)
- DBCD with $\gamma = 0$ and target allocation $\rho = \frac{q_2}{q_1+q_2}$
- DBCD with $\gamma = 2$ and target allocation $\rho = \frac{q_2}{q_1+q_2}$
- ERADE with $\alpha = 0.5$
- DBCD with $\gamma = 2$ and target allocation $\rho = \frac{\sqrt{P_1}}{\sqrt{P_1}+\sqrt{P_2}}$

3.2 Results for response-adaptive design

We would like to evaluate the number of restaurants assigned to promotion 1 and the power of the tests based on 1000 replications. As there is a big difference of the successful promotion rate between the two promotions in the original data. The power of the tests is based on the null hypothesis $p_1 - p_2 < 0.2$. The test is one-sided with 95% confidence level. The results are as follows, note that N_1 represents the number of restaurants assigned to promotion 1.

Table 4 Response-Adaptive Design Results

	CR	RPW	$DBCD_1$	$DBCD_2$	ERADE	$DBCD_3$
power	0.8010	0.7680	0.770	0.7700	0.734	0.8050
mean of N_1	179.9760	234.7340	233.979	235.9830	235.458	211.6170
sd of N_1	9.6024	14.5083	14.185	10.1338	8.656	7.5336

4. Covariate-adaptive Design

4.1 Outline of the procedures

Covariate-adaptive design is a randomization procedure that incorporates covariates and balances treatment allocation over covariates. It has two typical features:

- Achieving statistical efficiency by preserving type I errors while increasing power.
- Balancing treatment allocation for influential covariates.

For this design, we used 4 different randomization procedures discussed in our lectures and homework. Here are some of our settings and notes:

- **Complete randomization (CR):** $p = 0.5$
- **Stratified permuted block design (STR-PB)**: The block size = 4.
- **Pocock-Simon procedure (P-S)**: The weights of the margins =

(0.35,0.35,0.3), biased coin probability $p = 0.75$.

- **Hu and Hu's procedure (HuHu)**: The weight of overall imbalance = 0.35, the weight of within-stratum imbalance = 0.35, the weights of margins = (0.1,0.1,0.1), biased coin probability $p = 0.75$.

The reason why we set different marginal weights for the P-S procedure is that we think the first two covariates in this data, "market size" and "age of store" are much more related to the response than the third one "week". For the HuHu procedure, it doesn't matter because the overall and within-stratum imbalance dominate.

Next, the table 5 shows the distribution of the samples (restaurants) in the strata:

Table 5 distribution of the restaurants

# of restaurants within stratum	0	1	2	3	4 and more
# of strata	0	4	4	8	32

From the table above, there is no empty stratum. Only 4 strata have 1 sample in each of them. Most strata, 32 of 48 strata to be exact, have 4 or more samples, which is good for our stratified randomization.

4.2 Imbalance properties of the procedures

In this part, we measured and compared the imbalance properties of the 4 procedures based on 1000 replications.

Firstly, the overall imbalance:

Table 6 Mean absolute overall imbalance

	Complete randomization	STR–PBR	Pocock and Simon	Hu and Hu's
mean	15.88	4.96	1.886	1.206
median	14.00	4.00	2.000	2.000
95% quantile	38.00	12.00	4.000	4.000

Table 6 shows that CR has the highest overall imbalance and then the STR-PB. HuHu procedure has the lowest overall balance, which means the best performance among the 4 procedures.

Secondly, the marginal imbalances:

Table 7 Absolute marginal imbalances

	Complete randomization	STR–PBR	Pocock and Simon's	Hu and Hu's
market=small imbalance	4.830	3.250	1.416	1.738
market=medium imbalance	11.234	3.312	1.320	1.876
market=large imbalance	8.972	1.714	1.502	2.006
age=0 imbalance	9.232	2.270	1.474	2.236
age=1 imbalance	6.916	2.216	1.414	2.182
age=2 imbalance	7.098	2.434	1.428	2.232
age=3 imbalance	7.002	2.924	1.430	2.250
week=1 imbalance	7.760	2.554	1.550	2.244
week=2 imbalance	7.516	2.396	1.448	2.180
week=3 imbalance	7.368	2.454	1.540	2.206
week=4 imbalance	7.396	2.548	1.448	2.424

Table 7 above gives the mean absolute marginal imbalances. For the "market size", "age of store" and "week", the table explicitly lists the mean values on these 10 margins. P-S procedure has the best performance; HuHu procedure is slightly worse, but still acceptable. CR also has the worst performance in terms of marginal imbalances.

Finally, the within-stratum imbalances:

Table 8 Absolute with-in stratum imbalances

# of restaurants within stratum	D(n)	Complete randomization	STR-PBR	Pocock and Simon's	Hu and Hu's
2	prob(=2)	0.525	0.3325	0.47	0.41
	prob(=0)	0.475	0.6675	0.53	0.59
	mean	1.0500	0.665	0.74	0.8200
3	prob(=3)	0.27125	0	0.21125	0.11625
	prob(=1)	0.72875	1	0.78875	0.88375
	mean	1.5425	1.000	1.4225	1.2325

Table 8 displays the distribution and absolute mean of within-stratum imbalances for strata with 2 or 3 restaurants. According to this criterion, STR-PB has the lowest mean, HuHu has a slightly larger value, CR is still the procedure with the highest imbalance.

It is worth mentioning that, for strata containing 3 patients, since the block size is 4 for STR-PB, it is impossible to get an absolute value of 3. Hence, the mean absolute imbalance is 1, the minimum among the three methods.

In summary, HuHu procedure maintains a good balance from all three perspectives and should be favored when we consider the imbalance properties.

4.3 Statistical Inference of Covariate-adaptive Design

4.3.1 Outline of the inference

According to the feature we mentioned before, the design achieves statistical efficiency by preserving type I errors while increasing power. Therefore, we do a hypothesis test based on linear models to measure and compare type I errors and powers of 4 procedures. We consider promotion 1 as group A and promotion 2 as group B.

Null hypothesis: $\mu_A - \mu_B = 0$

Two-sided test with 95% confidence level

The linear model we fit based on the original data:

y_0: sales in thousands, t_0: treatment, z_1: market size, z_2: age of store

```
> summary(fit0)

Call:
lm(formula = y0 ~ t0 + z1 + z2)

Residuals:
    Min      1Q  Median      3Q     Max 
-26.897 -10.160  -3.294   9.303  32.523 

Coefficients:
            Estimate Std. Error t value Pr(>|t|)    
(Intercept)  19.7181     3.0641   6.435 3.98e-10 ***
t0           11.3461     1.4641   7.750 9.73e-14 ***
z1           12.3860     1.1951  10.364  < 2e-16 ***
z2           -0.2531     0.6251  -0.405    0.686    
---
Signif. codes:  0 ‘***’ 0.001 ‘**’ 0.01 ‘*’ 0.05 ‘.’ 0.1 ‘ ’ 1

Residual standard error: 13.87 on 356 degrees of freedom
Multiple R-squared:  0.3153,	Adjusted R-squared:  0.3096 
F-statistic: 54.65 on 3 and 356 DF,  p-value: < 2.2e-16
```

Then we used the coefficients b_0, b_1 and b_2 to generate response y with different treatment effects (diff)

$$y = b_0 + diff * t + b_1 z_1 + b_2 z_2 + \varepsilon \qquad (1)$$

The models we use for hypothesis test:

$$y \sim t$$

$$y \sim t + z_1$$

$$y \sim t + z_2$$

$$y \sim t + z_1 + z_2$$

4.3.1 Type I error

Let diff = 0, for 1000 simulations, we generated response y using (1) and fitted models. The type I errors results of the 4 procedures and 4 models:

Table 9 Type I error

	CR	STR–PB	P-S	HuHu
fit1 = lm(y~t)	0.050	0.027	0.036	0.028
fit2 = lm(y~t+z1)	0.047	0.045	0.053	0.047
fit3 = lm(y~t+z2)	0.048	0.028	0.036	0.028
fit4 = lm(y~t+z1+z2)	0.046	0.044	0.054	0.046

From the table we can see the type I errors are less than 0.05 which is aligned with our settings. All 4 procedures and 4 models have acceptable type I error. Also, we note that the Type I error for STR-PB, P-S and HuHu is conservative when z_1 is not included. This is reasonable because the coefficient for z_2 in the original linear model fit_0 is not significant. Namely, the market size does not affect sales. Theoretically, the Type I error for STR-PB, P-S and HuHu should be conservative whenever the linear model is not fully specified.

Then we used "carat" package to do the corrected t-test with 3 procedures except for CR. The corrected type I error results:

Table 10 Corrected Type I error

	STR–PB	P-S	HuHu
Corrected t-test	0.053	0.052	0.054

Compared with the former results, the type I errors of corrected t-test are less conservative and close to 0.05. And the three procedures have similar performance.

4.3.2 Power analysis

For the power analysis, we used different "diff" (coefficient of t) in (1) to generate response y. We set the absolute value of the mean difference between promotion 1 (group A) and promotion 2 (group B) of response y as our reference value (i.e., $|\mu_A - \mu_B|$). And the coefficients:

$$\text{diff} = k * |\mu_A - \mu_B| \qquad (k = 0.2, 0.5, 1, 1.5).$$

And for 1000 simulations, the power results are:

Table 11 Power based on multiple difference

diff = $0.2 * \|\mu_A - \mu_B\|$	CR	STR–PB	P-S	HuHu
fit1 = lm(y~t)	0.1510	0.1305	0.1350	0.1155
fit2 = lm(y~t+z1)	0.1855	0.1785	0.1940	0.1775
fit3 = lm(y~t+z2)	0.1490	0.1300	0.1370	0.1180
fit4 = lm(y~t+z1+z2)	0.1855	0.1800	0.1955	0.1760

diff = $0.5 * \|\mu_A - \mu_B\|$	CR	STR–PB	P-S	HuHu
fit1 = lm(y~t)	0.3933	0.3917	0.3963	0.3843
fit2 = lm(y~t+z1)	0.4383	0.4373	0.4473	0.4370
fit3 = lm(y~t+z2)	0.3900	0.3910	0.3977	0.3867
fit4 = lm(y~t+z1+z2)	0.4377	0.4380	0.4483	0.4353

diff = $\|\mu_A - \mu_B\|$	CR	STR–PB	P-S	HuHu
fit1 = lm(y~t)	0.5450	0.5438	0.5473	0.5383
fit2 = lm(y~t+z1)	0.5788	0.5780	0.5855	0.5778
fit3 = lm(y~t+z2)	0.5425	0.5433	0.5483	0.5400
fit4 = lm(y~t+z1+z2)	0.5783	0.5785	0.5863	0.5765

diff = $1.5 * \|\mu_A - \mu_B\|$	CR	STR–PB	P-S	HuHu
fit1 = lm(y~t)	0.6360	0.6350	0.6378	0.6306
fit2 = lm(y~t+z1)	0.6630	0.6624	0.6684	0.6622
fit3 = lm(y~t+z2)	0.6340	0.6346	0.6386	0.6320
fit4 = lm(y~t+z1+z2)	0.6626	0.6628	0.6690	0.6612

From the power results, the bigger "diff" we set, the larger power we get. When the model is fixed, the powers of the 4 procedures are similar. When the design is fixed, fit_2 and fit_4 linear models have larger power compared with the others.

5. Conclusion

In this project, we realized the response-adaptive design and the covariate adaptive design. And we made some comparisons for the different procedures. Here are the conclusions:

For response-adaptive design:

1. Promotion 1 is better than promotion 2.
2. DBCD with $\gamma = 2$ and target allocation $\rho = \frac{\sqrt{P_1}}{\sqrt{P_1}+\sqrt{P_2}}$ has the largest power and lowest standard deviation of restaurants assigned to promotion 1. Besides, more restaurants are assigned to better promotion 1. This method is recommended.
3. The mean of restaurants assigned to promotion 1 converges to the target allocation ρ desired.

For covariate-adaptive design:

1. Promotion 1 and promotion 2 have different effects on sales.

2. P-S performs well in marginal imbalance. STR-PB performs well in within-stratum imbalance. HuHu procedure maintains a good balance from all overall imbalance, marginal imbalance and within-stratum imbalance.
3. Except for the full model (fit_4), the type I error is conservative for STR-PB, P-S, and HuHu as it is less than 0.05. After correction, the Type I error is not conservative anymore.
4. When the design is fixed, the full model has the largest power. When the model is fixed, CR, STR-PB, P-S, and HuHu have similar power. And the power increases as the treatment effect increases.

References

1. Hu, F., & Rosenberger, W. F. (2003). Optimality, Variability, Power. *Journal of the American Statistical Association*, *98*(463), 671–678. https://doi.org/10.1198/016214503000000576
2. Hu, F., Zhang, L.-X., & He, X. (2009). Efficient randomized-adaptive designs. *The Annals of Statistics*, *37*(5A). https://doi.org/10.1214/08-aos655
3. Hu, Y., & Hu, F. (2012). Asymptotic properties of covariate-adaptive randomization. *The Annals of Statistics*, *40*(3), 1794–1815. https://doi.org/10.1214/12-aos983
4. Ma, W., Qin, Y., Li, Y., & Hu, F. (2019). Statistical Inference for Covariate-Adaptive Randomization Procedures. *Journal of the American Statistical Association*, *115*(531), 1488–1497. https://doi.org/10.1080/01621459.2019.1635483