Construction of Mixed-Level Orthogonal Arrays for Testing in Digital Marketing

Vladimir Brayman

Webtrends

October 19, 2012

Advantages of Conducting Designed Experiments in Digital Marketing

- Availability of Data
- Ease of Creating Tests
- Automation of the Analysis

Challenges of Conducting Designed Experiments in Digital Marketing

- Wide Range of Factor-Level Combinations, Including Mixed-Level Designs
- Binary and continuous response variable
- Designs Must Be Small
- Designs Must Be Robust
- Must Isolate Effects
- Must Produce Results Fast
- Unsophisticated Users Robustness

Outline

- Motivation
- Fractional Factorial Design
- Clear Effects
- Minimum Aberration Criterion
- Orthogonal Array Construction
- Conclusion

Fractional Factorial Design

Motivation: for economic reasons, full factorial designs are seldom used in practice for large k ($k \geq 7$).

Fractional Factorial Design: a subset or fraction of full factorial designs.

- "Optimal" fractions: are chosen according to the **resolution** or **minimum aberration** criteria.
- **Aliasing** of effects: a price one must pay for choosing a smaller design.

Design r^{k-p}, where

- r: level of the factors.
- k: number of the factors.
- p: number of design generators.
- $n = r^{k-p}$: run size.

An Example

No.	A	B	C	D	E
1	$-$	$+$	$+$	$-$	$-$
2	$+$	$+$	$+$	$+$	$-$
3	$-$	$-$	$+$	$+$	$-$
4	$+$	$-$	$+$	$-$	$-$
5	$-$	$+$	$-$	$+$	$-$
6	$+$	$+$	$-$	$-$	$-$
7	$-$	$-$	$-$	$-$	$-$
8	$+$	$-$	$-$	$+$	$-$
9	$-$	$+$	$+$	$-$	$+$
10	$+$	$+$	$+$	$+$	$+$
11	$-$	$-$	$+$	$+$	$+$
12	$+$	$-$	$+$	$-$	$+$
13	$-$	$+$	$-$	$+$	$+$
14	$+$	$+$	$-$	$-$	$+$
15	$-$	$-$	$-$	$-$	$+$
16	$+$	$-$	$-$	$+$	$+$

Balance and Orthogonality

Two key properties of the designs: **balance** and **orthogonality**.

- Balance: Each factor level appears in the same number of runs.
- Orthogonality: Two factors are called orthogonal if all their level combinations appear in the same number of runs. A design is called orthogonal if all pairs of its factors are orthogonal.

Design Generators

- 2^{5-1} design: 16 runs, which is a $\frac{1}{2}$ fraction of a 2^5 full factorial design.
- Aliasing: D and ABC, i.e., main effect of D is aliased with the $A \times B \times C$ interaction.
- The aliasing is denoted by the **design generator** $D = ABC$, $x_4 = x_1 + x_2 + x_3 \pmod 2$.
- Since $2x_4 = x_1 + x_2 + x_3 + x_4 = 0 \pmod 2$, we can get the **defining relation** $I = ABCD$ ($I = 1234$).

	Number	Factors
Main effects	5	A,B,C,D,E
Two-factor	10	AB,AC,AD,AE,BC,...,DE
Three-factor	10	ABC,ABD,ABE,BCD,...,CDE
Four-factor	5	ABCD,ABCE,ABDE,ACDE,BCDE
Five-factor	1	ABCDE

Outline

- Motivation
- Fractional Factorial Design
- Clear Effects
- Minimum Aberration Criterion
- Orthogonal Array Construction
- Conclusion

Clear Main Effects and Two-factor Interaction Effects

Clear effect: a main effect or two-factor interaction is clear if none of its aliases are main effects or two-factor interactions.

	Number	Factors
Main effects	5	A,B,C,D,E
Two-factor	4	AE,BE,CE,DE

From $x_1 + x_2 + x_3 + x_4 = 0 \pmod 2$, we can get:

- $A = BCD, B = ACD, C = ABD$, so all the main effects are clear.
- $AB = CD, AC = BD, AD = BC, ..., AE = BCDE, BE = ACDE, CE = ABDE, DE = ABCE$, so only the two-factor interactions including E are clear, all the others aliased with other two-factor interactions.

More Than One Design Generators

Consider the 2^{6-2} design with design generators:
$E = AB$, $F = ACD$.

- We get the **defining contrast subgroups**:
  $I = ABE = ACDF = BCDEF$.
- A_i: the number of words of length i in its defining contrast subgroup, **wordlength pattern** $W = (A_3, A_4, ..., A_k)$.
- **Resolution**: the smallest r such that $A_r \geq 1$, i.e., the length of the shortest word in the defining contrast subgroup.
- The above design, resolution $R = 3$ and $W = (1, 1, 1, 0, 0, ...)$.
- **Maximum Resolution Criterion**: Box and Hunter (1961).
- Resolution III design, some main effects are not clear.
- Resolution IV design, main effects are clear, those with the largest number of clear two-factor interactions are the best.
- Resolution V design, two-factor interactions are clear.

Outline

- Motivation
- Fractional Factorial Design
- Clear Effects
- Minimum Aberration Criterion
- Orthogonal Array
- Conclusion

Minimum Aberration Criterion

- Question: for the same r^{k-p} designs d_1 and d_2 with different design generators, which one is better?
- Consider the following two 2^{7-2} designs:
  d_1: $I = 4567 = 12346 = 12357$,
  d_2: $I = 1236 = 1457 = 234567$.
- Fries and Hunter (1980): For any two 2^{k-p} designs d_1 and d_2, let r be the smallest integer such that $A_r(d_1) \neq A_r(d_2)$. Then d_1 is said to have **less aberration** than d_2 if $A_r(d_1) < A_r(d_2)$. If there is no design with less aberration than d_1, then d_1 has **minimum aberration**.
- For the above d_1 and d_2, we have wordlength patterns:
  $W(d_1) = (0, 1, 2, 0, 0)$,
  $W(d_2) = (0, 2, 0, 1, 0)$,
  so d_1 is better than d_2.

Maximum Number of Clear Effects Criterion

- Consider the following two 2^{9-4} designs:
  d_1: $6 = 123, 7 = 124, 8 = 125, 9 = 1345$,
  d_2: $6 = 123, 7 = 124, 8 = 134, 9 = 2345$.
  d_1: $I = 1236 = 1247 = 1258 = 3467 = 3568 = 4578$,
  d_2: $I = 1236 = 1247 = 1348 = 3467 = 2468 = 2378 = 1678$.
- For the above d_1 and d_2, we have:
  $A_3(d_1) = A_3(d_2) = 0$,
  $A_4(d_1) = 6 < A_4(d_2) = 7$,
  so d_1 is better than d_2 from minimum aberration criterion.
- While all the 9 main effects in d_1 and d_2 are clear, d_2 has 15 clear two-factor interactions but d_1 has only 8, so one would judge that d_2 is better than d_1.

Experiments at Mixed Levels

- When $r = 3$, $A \times B$: AB, AB^2, $A \times B \times C$: $ABC, ABC^2, AB^2C, AB^2C^2$.
- Consider a $2^{3-1} \times 3^{3-1}$ (asymmetric) **product** design:
  d_1: $C = AB$ for the two-level factors A, B, C; $I = ABC$.
  d_2: $D = EF$ for the three-level factors D, E, F; $I = DEF^2$.
- Type 1: find 3 aliasing relations A_1, A_2, A_3 of the two-level factors A, B, C, from $C = AB$:
  A_1: $A = BC$
  A_2: $B = AC$
  A_3: $C = AB$
- Type 2: find 4 aliasing relations B_1, B_2, B_3, B_4 of the three-level factors D, E, F, from $D = EF$:
  B_1: $D = DE^2F = EF^2$
  B_2: $E = DF^2 = DE^2F^2$
  B_3: $F = DE = DEF$
  B_4: $DE^2 = DF = EF$.

Experiments at Mixed Levels (Continued)

- Type 3: find 12 aliasing relations C_1 to C_{12} from Type 1 and Type 2 aliasing relations:
  C_1 (from A_1 and B_1):
  $AD = ADE^2F = AEF^2 = BCD = BCDE^2 = BCDEF^2$.
  C_2 (from A_1 and B_2):
  $AE = ADF^2 = ADE^2F^2 = BCE = BCDF^2 = BCDE^2F^2$.
  C_3: $AF = ADE = ADEF = BCF = BCDE = BCDEF$.
  C_4: $ADE^2 = ADF = AEF = BCDE^2 = BCDF = BCEF$.
  C_5: $BD = BDE^2F = BEF^2 = ACD = ACDE^2F = ACEF^2$.
  C_6: $BE = BDF^2 = BDE^2F^2 = ACE = ACDF^2 = ACDE^2F^2$.
  C_7: $BF = BDE = BDEF = ACF = ACDE = ACDEF$.
  C_8: $BDE^2 = BDF = BEF = ACDE^2 = ACDF = ACEF$.
  C_9: $CD = CDE^2F = CEF^2 = ABD = ABDE^2F = ABEF^2$.
  C_{10}:
  $CE = CDF^2 = CDE^2F^2 = ABE = ABDF^2F = ABDE^2F^2$.
  C_{11}: $CF = CDE = CDEF = ABF = ABDE = ABDEF$.
  C_{12}: $CDE^2 = CDF = CEF = ABDE^2 = ABDF = ABEF$.

Outline

- Motivation
- Fractional Factorial Design
- Clear Effects
- Minimum Aberration Criterion
- Orthogonal Array Construction
- Conclusion

Orthogonal Array Construction Problem

Problem: given factors vector, R, how to construct the orthogonal array (OA) (design matrix) that has the minimum possible run size n?

- R=3 or R=4 for symmetric design.
- For asymmetric design, we define $R = min(R_1, R_2, ..., R_m)$. Sometimes have to be $R = 1$.
- level=$(r_1, .., r_1, r_2, ..., r_2, ..., r_k, ..., r_m)$. For example, level=(2,2,2,3,3,3,3) or level=(2,2,2,2,3,3,3).
- minimum possible run size $n \Longrightarrow$ minimum possible run size vector $n = (n_1, n_2, ..., n_m) \Longleftrightarrow$ maximum possible $p = (p_1, p_2, ..., p_m)$.
- Get $(OA_1, OA_2, ..., OA_m)$, then cross them together to get the (asymmetric) product design OA.

Maximum Possible p Table

Problem: in each symmetric group, given r, k, R, how to get the maximum possible p?

For r=2:

k	3	4	4	5	5	5	6	6	6	6
R	3	3(1)	4	3	4(1)	5	3	4	5	6
max p	1	1^*	1	2	1^*	1	3	2	2	1

For r=3:

k	3	4	4	5	5	5	6	6	6	6
R	3	3	4	3	4	5	3	4	5	6
max p	1	2	1	2	2	1	3	3	2	1

Notice: For $r = 2$, it might be not compatible for some given R. FYI, for (2,2,2,2) and $R = 3$, we can only assign 1 design generator $3 = 12$, then factor 4 will be the extra factor.

An Example

No.	A	B	C	D
1	$-$	$+$	$-$	$+$
2	$+$	$+$	$+$	$+$
3	$-$	$-$	$+$	$+$
4	$+$	$-$	$-$	$+$
5	$-$	$+$	$-$	$-$
6	$+$	$+$	$-$	$-$
7	$-$	$-$	$-$	$-$
8	$+$	$-$	$-$	$-$

- For factors A, B, C, it is a 2^{3-1} design with design generator $C = AB$, and D is the extra factor. A, B, C, D makes a
- $2^{3-1} \times 2$ product design.
- It is a design with $R = min(R_1, R_2) = min(3, 1) = 1$.

Orthogonal Array Construction Algorithm

Inputs and outputs of the function codes:

- Input: level vector, R.
- Output: OA ($OA_1, OA_2, ..., OA_m$ are intermediate outputs).

Algorithm:
(1) From R, generator all the possible resolution combination vector $(R_1, R_2, ..., R_m)$.
(2) In each symmetric group (given r, k, R_i), check the compatibility of the given level and resolution R_i.

- If not, stop.
- If yes, continue to step (2).

(3) In each symmetric group (given r, k, R_i), find the maximum possible p.
(4) In each symmetric design (given r, k, p_i), get all the possible design generators $(d.g)$.

Orthogonal Array Construction Algorithm (Continued)

(5) In each symmetric design (given r, k, $d.g$), from all the possible design generators, get the one which can achieves the minimum aberration.

- For each possible design generators, get the wordlength.
- Rank all the wordlengths through minimum aberration criterion.
- Pick up the best wordlength, find its corresponding design generators ($d.g_o$).

(6) In each symmetric design (given r, k, $d.g_o$), generate the OA_i.
(7) Cross all the OA_is to get the product design OA.

Outline

- Motivation
- Fractional Factorial Design
- Clear Effects
- Minimum Aberration Criterion
- Orthogonal Array Construction
- Conclusion

Conclusions

- We introduce the basic ideas of fractional factorial design, design generators and minimum aberration criterion.
- We generalize all the ideas from symmetric design to asymmetric (mixed-level) design.
- We provide an algorithm to generate the orthogonal array based on the minimum aberration criterion.

References

- Box, G. E. P., Hunter, W.G, and Hunter,J.S. (1978), "Statistics for experimenters," New York: John Wiley & Sons.
- Fries, A., and Hunter, W. G. (1980), "Minimum aberration 2^{k-p} designs," *Technometrics*, 22, 601-608.
- Wu, C. F. Jeff. and Hamada, Micheal. S. (2009), "Experiments: planning,analysis, and optimization", (2nd edition), Wiley.