Plackett-Burman Tables

fF2 vs PB Tables, low n

Donald Palahnuk

2024.11.10
version 1.0

InsightHatch WhitePaper: fF2 vs PB Tables For (low n)
IH-fF2 PB Tables-v1.0

Contents

Revision	**Date**	**Author(s)**	**Description**
0.0	2023.10.10	DP	created!
0.1	2023.10.10	DP	compared fF2 vs PB calls, added kable tables
0.2	2023.10.11	DP	add resolution, and some math
1.0	2024.10.27	DP	update for github issue
1.0	2024.11.10	DP	plan extensions - subject area application

1 Fractional Factorial and Plackett-Burman DOE

Design of Experiments (DOE) is a methodological framework for systemically planning experiments, collecting data, and analyzing results. 2 Level Fractional Factorial (fF2) and Plackett-Burman (PB) designs are two common techniques within DOE that aim to reduce the number of experiments you need to run when investigating multiple factors.

1.1 Difference between Fractional Factorial and Plackett-Burman Design

1.1.1 Fractional Factorial (fF2) 2 Level Design

In a full factorial (FF) design, you study all possible combinations of the factors. However, this can become impractical with a large number of factors. In a fractional factorial design, only a “fraction” of the full set of experiments is carried out. The fraction is chosen carefully so that certain effects (main effects or interactions) can still be estimated. Note in this write-up we use the conventions: $n = 2^{(f-r)}$, where f=factors, r=reductions).

Variable Resolution: Fractional factorial designs can be of various resolutions (III, IV, V, etc.), depending on how you fractionate the full factorial design.

Resolution III: Main effects are confounded with two-factor interactions but not with each other.

Resolution IV: Main effects are clear of all two-factor interactions. Two-factor interactions are confounded with each other but not with main effects.

Resolution V and above: Higher resolutions allow you to separate higher-order interactions from lower-order effects.

1.1.2 Plackett-Burman Design

The Plackett-Burman design is another approach to reducing the number of experimental runs. It’s commonly used for screening experiments where you have many factors and you want to identify which ones have significant effects on the outcome. These designs are specialized for estimating main effects and **assume that interaction effects are negligible.**

Resolution III: Plackett-Burman designs are generally resolution III designs. This means that main effects are confounded with two-factor interactions but not with each other. In other words, you cannot distinguish between a main effect and a two-factor interaction that it is confounded with. However, main effects are clear of each other.

1.2 *Pros* and *Cons*

1.2.1 Fractional Factorial Design

Pros: More flexible in terms of defining the "fraction" of experiments. Allows for the estimation of interaction effects. Can be optimal or near-optimal for a given effect size and error.

Cons: More complex to set up and analyze, especially for higher-order fractions. Not efficient for screening a large number of factors, especially when most are not influential.

1.2.2 Plackett-Burman Design

Pros: Simple to set up and analyze. Efficient for screening a large number of factors to identify the few that are most significant.

Cons: Assumes that interaction effects are negligible, which may not be true. Less flexible in design choices compared to fractional factorial.

In summary, fractional factorial designs offer more flexibility and the ability to study interactions but can be complex and often require more runs than Plackett-Burman designs. Plackett-Burman designs are more straightforward and are generally used for initial screening when you have a large number of factors and you want to identify the most significant ones quickly.

1.3 Discussion of increasing n in the context of fF2 & PB designs.

The quality of multivariate residuals $(y\hat{-}y)$ in a Design of Experiments (DOE) largely depends on the underlying structure of the problem, the number and nature of factors, and the presence of interactions and higher-order effects. Here's a breakdown based on the two designs you've mentioned:

Lets discuss a fF2 of $n = 12$ vs PB of $n = 8$ runs to further understand how these concepts manifest.

fF2 with 12 Runs: A 12-run fF2 design will generally allow you to estimate more effects (both main effects and interactions) than an 8-run PB design. This would generally make the model fit better to the actual observed data, potentially yielding better residuals.

PB with 8 Runs: An 8-run PB design is generally intended for screening a large number of factors to find the few that are most significant. It is usually not efficient for estimating interaction effects, so if your process has significant interactions, a Plackett-Burman design might not capture them well, leading to poorer residuals. However, it allows for a reduced number of experiments which can help to save in experimental execution costs.

Key Points to Consider:

Interactions: If interactions between factors are significant, a fractional factorial design would generally yield better residuals as it allows for the estimation of interactions.

Number of Factors: The Plackett-Burman design is useful for screening a large number of factors. If most factors are not influential, then a Plackett-Burman design may perform reasonably well even with fewer runs.

Assumptions: Plackett-Burman assumes negligible interactions. If this assumption holds true, then it may give good residuals even with fewer runs.

Model Complexity: Fractional Factorial designs are generally more flexible and can fit more complex models than Plackett-Burman designs.

In summary, if you believe your system is better represented by a model that includes interaction terms, or if you want to estimate these terms, a 12-run fractional factorial design is more likely to give better multivariate residuals. On the other hand, if you are primarily interested in identifying the main effects quickly and cheaply, and are willing to assume that interactions are negligible, then an 8-run Plackett-Burman design could suffice.

2 fF2 vs PB in terms of base design n experiments and f factors

Comparison of n between Plackett-Burman and 2^{f-r}_{III}

Number of factors	PB	fF2 III
$4 \leq f \leq 7$	8	8
$8 \leq f \leq 11$	12	16
$12 \leq f \leq 15$	16	16
$16 \leq f \leq 19$	20	32
$20 \leq f \leq 23$	24	32
$24 \leq f \leq 27$	28	32
$28 \leq f \leq 31$	32	32
$32 \leq f \leq 35$	36	64
$\vdots$	$\vdots$	$\vdots$
$64 \leq f \leq 67$	68	128

3 fF2 & PB Tables Set for 4 and 5 factors

Series of low n Placket-Burman and fF2 (reduced factorial) Design Tables

- Motivation: Develop reference of commonly used fF2 & PB Tables for typical BioProcess PD work
- Compile Interesting Examples - for useful consideration

3.1 Comments

3.1.1 PB Designs

PB designs are done in increments of 4 … n = 4, 8, 12, etc.

3.1.2 fF2 Designs

fF2 designs are based on powers of 2 ($2^{(f-r)}$, where f=factors, r=reductions) … n = 4, 8, 16, etc. (or if not a power of 2 -> design must be truncated)

4 fF2 & PB for 4 Factors

4.1 PB vs fF2 8 runs, 4 factors

```r
library("FrF2")
design <- pb(8, 4)
knitr::kable(design, align = "rrrr", caption = "PB, n=8, f=4")
```

Table 1: PB, n=8, f=4

A	B	C	D
-1	-1	-1	-1
-1	-1	1	1
-1	1	1	-1
1	-1	1	-1
1	-1	-1	1
1	1	-1	-1
-1	1	-1	1
1	1	1	1

```r
library("FrF2")
design <- FrF2(8, 4)
knitr::kable(design, align = "rrrr", caption = "fF, n=8, f=4")
```

Table 2: fF, n=8, f=4

A	B	C	D
-1	-1	-1	-1
-1	1	1	-1
1	1	-1	-1
-1	-1	1	1
1	-1	-1	1
1	1	1	1
1	-1	1	-1
-1	1	-1	1

4.2 PB vs fF2 12/16 runs, 4 factors

```
library("FrF2")
design <- pb(12, 4)
knitr::kable(design, align = "rrrr", caption = "PB, n=12, f=4")
```

Table 3: PB, n=12, f=4

A	B	C	D
1	1	-1	1
-1	-1	1	-1
-1	1	-1	1
-1	-1	-1	1
1	-1	1	1
-1	1	1	-1
-1	1	1	1
1	1	1	-1
-1	-1	-1	-1
1	-1	1	1
1	1	-1	-1
1	-1	-1	-1

```
library("FrF2")
design <- FrF2(16, 4)
knitr::kable(design, align = "rrrr", caption = "fF, n=12, f=4")
```

Table 4: fF, n=12, f=4

A	B	C	D
-1	-1	1	-1
1	1	1	-1
-1	-1	-1	-1
1	-1	-1	1
-1	1	-1	-1
-1	1	-1	1
-1	-1	1	1
1	-1	-1	-1
-1	-1	-1	1

A	B	C	D
-1	1	1	1
1	1	-1	-1
1	1	1	1
1	-1	1	1
1	1	-1	1
-1	1	1	-1
1	-1	1	-1

5 fF2 & PB for 5 Factors

5.1 PB vs fF2 8 runs, 5 factors

```
library("FrF2")
design <- pb(8, 5)
knitr::kable(design, align = "rrrrr", caption = "PB, n=8, f=5")
```

Table 5: PB, n=8, f=5

A	B	C	D	E
-1	-1	-1	-1	-1
1	1	-1	1	-1
-1	1	1	1	-1
1	-1	1	-1	-1
-1	-1	1	1	1
-1	1	-1	-1	1
1	1	1	-1	1
1	-1	-1	1	1

```
library("FrF2")
design <- FrF2(8, 5)
knitr::kable(design, align = "rrrrr", caption = "fF, n=8, f=5")
```

Table 6: fF, n=8, f=5

A	B	C	D	E
1	-1	1	-1	1
-1	-1	-1	1	1
-1	1	1	-1	-1
-1	1	-1	-1	1
1	-1	-1	-1	-1
1	1	1	1	1
1	1	-1	1	-1
-1	-1	1	1	-1

5.2 PB vs fF2 12 runs, 5 factors

```
library("FrF2")
design <- pb(12, 5)
knitr::kable(design, align = "rrrr", caption = "PB, n=12, f=5")
```

Table 7: PB, n=12, f=5

A	B	C	D	E
1	1	-1	-1	-1
-1	1	1	-1	1
-1	1	-1	1	1
-1	-1	-1	-1	-1
-1	-1	-1	1	-1
1	-1	1	1	-1
-1	-1	1	-1	1
1	1	1	-1	-1
-1	1	1	1	-1
1	1	-1	1	1
1	-1	-1	-1	1
1	-1	1	1	1

```
library("FrF2")
design <- FrF2(16, 5)
knitr::kable(design, align = "rrrr", caption = "fF, n=12, f=5")
```

Table 8: fF, n=12, f=5

A	B	C	D	E
1	-1	1	1	-1
-1	1	1	1	-1
-1	-1	1	1	1
-1	-1	-1	1	-1
1	1	1	-1	-1
1	1	-1	1	-1
1	1	-1	-1	1
1	-1	-1	1	1
1	-1	-1	-1	-1

A	B	C	D	E
-1	-1	1	-1	-1
-1	-1	-1	-1	1
-1	1	1	-1	1
-1	1	-1	-1	-1
1	-1	1	-1	1
1	1	1	1	1
-1	1	-1	1	1

6 Hadamard Matrix

In mathematics, a Hadamard matrix, named after the French mathematician Jacques Hadamard, is a square matrix whose entries are either +1 or -1 and whose rows are mutually orthogonal. In geometric terms, this means that each pair of rows in a Hadamard matrix represents two perpendicular vectors, while in combinatorial terms, it means that each pair of rows has matching entries in exactly half of their columns and mismatched entries in the remaining columns. It is a consequence of this definition that the corresponding properties hold for columns as well as rows.

The n-dimensional parallelotope spanned by the rows of an $n \times n$ Hadamard matrix has the maximum possible n dimensional volume among parallelotopes spanned by vectors whose entries are bounded in absolute value by 1. Equivalently, a Hadamard matrix has maximal determinant among matrices with entries of absolute value less than or equal to 1 and so is an extremal solution of Hadamard's maximal determinant problem.

Let H be a Hadamard matrix of order n. The transpose of H is closely related to its inverse. In fact:

$$HH^\top = nI_n$$

where I_n is the $n \times n$ identity matrix and $H^\top$ is the transpose of H. To see that this is true, notice that the rows of H are all orthogonal vectors over the field of real numbers and each have length $\sqrt{n}$. Dividing H through by this length gives an orthogonal matrix whose transpose is thus its inverse. Multiplying by the length again gives the equality above. As a result,

$$\det(H) = \pm n^{n/2},$$

where $\det(H)$ is the determinant of H. Suppose that M is a complex matrix of order n, whose entries are bounded by $|M_{ij}| \leq 1$, for each i, j between 1 and n. Then Hadamard's determinant bound states that

$$|\det(M)| \leq n^{n/2}.$$

Equality in this bound is attained for a real matrix M if and only if M is a Hadamard matrix. The order of a Hadamard matrix must be 1,2 , or a multiple of 4[1]

6.1 Example of Hadamard for n=8

```
library(pracma)
H <- hadamard(8)
design <- H
knitr::kable(design, align = "rrrrrrrr", caption = "Hadamard, n=8")
```

Table 9: Hadamard, n=8

1	1	1	1	1	1	1	1
1	-1	1	-1	1	-1	1	-1
1	1	-1	-1	1	1	-1	-1
1	-1	-1	1	1	-1	-1	1
1	1	1	1	-1	-1	-1	-1
1	-1	1	-1	-1	1	-1	1
1	1	-1	-1	-1	-1	1	1
1	-1	-1	1	-1	1	1	-1

6.2 Example of Hadamard for n=12

```
library(pracma)
H <- hadamard(12)
design <- H
knitr::kable(design, align = "rrrrrrrrrrrr", caption = "Hadamard, n=12")
```

Table 10: Hadamard, n=12

1	1	1	1	1	1	1	1	1	1	1	1
1	-1	1	-1	1	1	1	-1	-1	-1	1	-1
1	-1	-1	1	-1	1	1	1	-1	-1	-1	1
1	1	-1	-1	1	-1	1	1	1	-1	-1	-1
1	-1	1	-1	-1	1	-1	1	1	1	-1	-1
1	-1	-1	1	-1	-1	1	-1	1	1	1	-1
1	-1	-1	-1	1	-1	-1	1	-1	1	1	1
1	1	-1	-1	-1	1	-1	-1	1	-1	1	1
1	1	1	-1	-1	-1	1	-1	-1	1	-1	1
1	1	1	1	-1	-1	-1	1	-1	-1	1	-1
1	-1	1	1	1	-1	-1	-1	1	-1	-1	1
1	1	-1	1	1	1	-1	-1	-1	1	-1	-1

6.3 Calculate Hadamard Property for n=8, $HH^\top = nI_n$

```
library(pracma)
H <- hadamard(8)
H_T <- t(H)
design <- H %*% H_T
knitr::kable(design, align = "rrrrrrrr", caption = "Hadamard, n=8")
```

Table 11: Hadamard, n=8

8	0	0	0	0	0	0	0
0	8	0	0	0	0	0	0
0	0	8	0	0	0	0	0
0	0	0	8	0	0	0	0
0	0	0	0	8	0	0	0
0	0	0	0	0	8	0	0
0	0	0	0	0	0	8	0
0	0	0	0	0	0	0	8

References

Montgomery, Douglas C. 2020. *Design and Analysis of Experiments.* Tenth edition. Hoboken, NJ: Wiley.