

Das Risiko für falsche Informationen aus Sieb-Versuchsplänen

Dipl.-Ing. Michael Radeck, TEQ® Training & Consulting GmbH

Für die Untersuchung von Ursache-Wirk-Beziehungen werden u.a. Versuchspläne eingesetzt. In der folgenden Grafik ist grob angedeutet, dass die Anwendung der Versuchsplanung in mehrere, aufeinander aufbauende Phasen unterteilt ist. In diesem Beitrag soll die Phase des Identifizierens von wichtigen Einflussgrößen beleuchtet werden, das Haupteinsatzgebiet der Sieb-Versuchspläne.

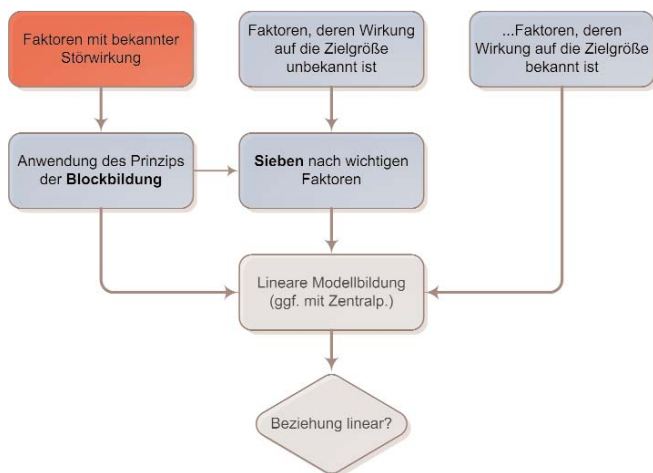


Abbildung 1: Das Identifizieren wichtiger Faktoren (Einflussgrößen) mit Sieb-Versuchsplänen in der Phase Sieben

Mit einem Sieb-Versuchsplan kann man zwischen schwach und stark wirksamen Faktoren unterscheiden und somit die stark wirksamen als wichtige Faktoren auszusieben. Für diesen Zweck werden oft Plackett-Burman Pläne, teilfaktorielle Pläne und die Lateinischen Quadrate nach Taguchi verwendet. Der wesentliche Reiz an diesen Plänen ist der geringe Umfang an Versuchen. Jedoch gibt es auch Nachteile, von denen zwei in diesem Artikel an einem Mini-Beispiel betrachtet werden:

1. Mit Sieb-Versuchsplänen lassen sich nicht alle Faktorwechselwirkungen schätzen
2. Die ermittelten Koeffizienten der Faktorwirkungen weisen einen systematischen Fehler auf

Beispiel: Zweistufiger Versuchsplan mit acht Faktoren

Würde man für das Sieben von acht Faktoren einen vollfaktoriellen Plan verwenden, so hätte man insgesamt $N = 2^8 = 256$ Versuchskombinationen zu bewältigen, was praktisch unmöglich ist. Aus diesem Grund wurden spezielle Sieb-Versuchspläne entwickelt, mit denen die Anzahl der Versuche z.T. dramatisch reduziert werden kann. Im Folgenden werden zwei Klassiker der Sieb-Versuchspläne betrachtet: Ein Plackett-Burman Plan mit 12 Versuchskombinationen und ein teilfaktorieller 2^{8-4} -Versuchsplan.

Sieb-Versuchsplan Nr. 1: Plackett-Burman Plan mit 12 Versuchskombinationen

Das Regressionsmodell des Plackett-Burman Planes für acht Faktoren lautet:

$$y = b_0 + b_1 \cdot x_1 + b_2 \cdot x_2 + b_3 \cdot x_3 + b_4 \cdot x_4 + b_5 \cdot x_5 + b_6 \cdot x_6 + b_7 \cdot x_7 + b_8 \cdot x_8$$

Aus der Darstellung wird deutlich, dass hier nur die Koeffizienten der Hauptfaktorwirkungen ermittelt werden. Die zu diesem Modellansatz gehörige Modellmatrix X ist in der Tabelle 1 abgebildet:

Konstante	A	B	C	D	E	F	G	H
1	1	1	-1	1	1	1	-1	-1
1	-1	1	1	-1	1	1	1	-1
1	1	-1	1	1	-1	1	1	1
1	-1	1	-1	1	1	-1	1	1
1	-1	-1	1	-1	1	1	-1	1
1	-1	-1	-1	1	-1	1	1	-1
1	1	-1	-1	-1	1	-1	1	1
1	1	1	-1	-1	-1	1	-1	1
1	1	1	1	-1	-1	-1	1	-1
1	-1	1	1	1	-1	-1	-1	1
1	1	-1	1	1	1	-1	-1	-1
1	-1	-1	-1	-1	-1	-1	-1	-1

Tabelle 1: Modellmatrix X des Plackett-Burman Planes für 8 Faktoren

Als Referenzmodell für den Vergleich wird der vollfaktorielle 2^8 -Plan gewählt. Dieser Plan würde die Bestimmung sämtlicher Koeffizienten der Faktor- und Wechselwirkungsterme ermöglichen. Das sind bei einem 2^8 -Plan immerhin 255 Koeffizienten für die Faktor- und Wechselwirkungskoeffizienten - die Konstante b_0 nicht mitgezählt. Wirft man nun wieder einen Blick auf die Tabelle 1, so erkennt man: Durch den Einsatz des Plackett-Burman Planes bleiben die Koeffizienten der 247 Faktorwechselwirkungen unbekannt.

Das allein wäre deshalb nicht so schlimm, da ja das Ziel eines Sieb-Versuches nicht das Modell ist, sondern nur das Reduzieren der Faktorenanzahl. Der eigentliche Nachteil steckt in der Tatsache, dass die mit dem Plackett-Burman Plan ermittelten Koeffizienten systematisch verfälscht sind.

Der systematische Modellfehler beim Plackett-Burman Plan

Die angedeutete systematische Verfälschung der Koeffizienten wird wie folgt ermittelt:

Schritt 1: Zunächst bleibt festzustellen, dass man mit dem Regressionsansatz des vollfaktoriellen Planes 255 Koeffizienten und die Konstante bestimmen kann. Mit dem Plackett-Burman Plan erhält man jedoch nur die

acht Koeffizienten der Faktoren A bis H und die Konstante. Somit lassen sich mit dem Plackett-Burman Plan alle 247 Koeffizienten der Faktorwechselwirkungen nicht ermitteln. Diese im Plackett-Burman Modellansatz fehlenden Koeffizienten sind nun im Fokus von Schritt 2.

Schritt 2: Man bildet die Differenzmatrix D. Sie besteht aus den Vorzeichenspalten für die 247 fehlenden Koeffizienten aus Schritt 1. Um den Aufbau der Differenzmatrix D nachvollziehen zu können, sind in der Tabelle 2 die ersten und letzten drei Spalten der Differenzmatrix D dargestellt. Die Vorzeichen in der Spalte AB von Tabelle 2 sind dadurch entstanden, dass man von der Tabelle 1 die Vorzeichen der Spalte A mit den Vorzeichen der Spalte B multipliziert hat. Analog sind die Vorzeichen in der Spalte AC durch die Multiplikation der Vorzeichen der Spalten A und C aus Tabelle 1 entstanden, usw.

AB	AC	AD	...	ACDEFGH	BCDEFGH	ABCDEFGH
1	-1	1	...	-1	-1	-1
-1	-1	1	...	-1	1	-1
-1	1	1	...	-1	1	1
-1	1	-1	...	-1	1	-1
1	-1	1	...	-1	-1	1
1	1	-1	...	1	1	-1
-1	-1	-1	...	-1	1	1
1	-1	-1	...	1	1	1
1	1	-1	...	1	1	1
-1	-1	-1	...	1	-1	1
-1	1	1	...	-1	1	1
1	1	1	...	-1	-1	1

Tabelle 2: Auszug der Differenzmatrix D für den Plackett-Burman Plan

Schritt 3: Mit der Aliasmatrix A wird die Vermengungsinformation berechnet. Das heißt, diese Matrix enthält die Information, welche der mit dem Plackett-Burman Plan berechneten Koeffizienten eine systematische Verfälschung aufweisen.

$$A = (X^T X)^{-1} X^T D$$

Da die Aliasmatrix A insgesamt 247 Spalten und neun Zeilen umfasst, bleibt die Darstellung auf eine Zeile beschränkt. Das ist die Vermengungsinformation für den Koeffizienten des Faktors A:

$$[A] = A - \frac{1}{3}BC - \frac{1}{3}BD - \frac{1}{3}BE + \frac{1}{3}BF - \frac{1}{3}BG - \frac{1}{3}BH + \frac{1}{3}CD - \frac{1}{3}CE - \frac{1}{3}CF + \frac{1}{3}CG - \frac{1}{3}CH + \frac{1}{3}DE + \frac{1}{3}DF - \frac{1}{3}DG - \frac{1}{3}DH - \frac{1}{3}EF - \frac{1}{3}EG - \frac{1}{3}EH - \frac{1}{3}FC + \frac{1}{3}FH + \frac{1}{3}GH + 126 \text{ weitere Summanden}$$

(Koeffizienten der Wechselwirkungen Ordnung ≥ 3).

Der Buchstabe A in eckigen Klammern ist der Wert des Koeffizienten, den man für den Faktor A mit dem Plackett-Burman Plan berechnen wird. Dieser Wert ist in Wahrheit überlagert mit der Summe der rechts vom Gleichheitszeichen stehenden 147 Wechselwirkungskoeffizienten. Ein wenig tröstend ist allein, dass die Koeffizienten der Wechselwirkungen nicht mit ihrem vollen Betrag eingehen.

Der mit dem Plackett-Burman Plan berechnete Koeffizient von Faktor A ist im Beispiel insgesamt mit 147 Wechselwirkungskoeffizienten vermengt.

Darunter sind 21 Koeffizienten, die zu Zweifaktorwechselwirkungen gehören. Tröstend ist allein, dass die Wechselwirkungsterme nicht mit ihrem vollen Betrag an der systematischen Verfälschung beteiligt sind.

Sieb-Versuchsplan Nr. 2: Der teilfaktorielle 2⁸⁻⁴Plan

Das Regressionsmodell des teilfaktoriellen 2⁸⁻⁴-Planes lautet:

$$y = b_0 + b_1 \cdot x_1 + b_2 \cdot x_2 + b_3 \cdot x_3 + b_4 \cdot x_4 + b_5 \cdot x_5 + b_6 \cdot x_6 + b_7 \cdot x_7 + b_8 \cdot x_8$$

Die zugehörige Modellmatrix X ist in der Tabelle 3 dargestellt:

Konstante	A	B	C	D	E	F	G	H
1	1	1	-1	1	1	1	-1	-1
1	-1	1	1	-1	1	1	1	-1
1	1	-1	1	1	-1	1	1	1
1	-1	1	-1	1	1	-1	1	1
1	-1	-1	1	-1	1	1	-1	1
1	-1	-1	-1	1	-1	1	1	-1
1	1	-1	-1	-1	1	-1	1	1
1	1	1	-1	-1	-1	1	-1	1
1	1	1	1	-1	-1	-1	1	-1
1	-1	1	1	1	-1	-1	-1	1
1	1	-1	1	1	1	-1	-1	-1
1	-1	-1	-1	-1	-1	-1	-1	-1

Tabelle 3: Modellmatrix X des teilfaktoriellen 2⁸⁻⁴-Planes

Sowohl die Differenzmatrix D als auch die Aliasmatrix A werden nach dem gleichen Prinzip ermittelt, wie es für den Plackett-Burman Plan beschrieben wurde. Auch wird wieder auf die Darstellung der Aliasmatrix verzichtet und allein die Vermengungsinformation von dem Koeffizienten des Faktors A gezeigt. Es bleibt zu beachten, dass die Vermengungsinformation dieses Mal vollständig ist:

$$[A] = A + BCG + BDH + BEF + CDF + CEH + DEG + FGH + ABCDE + ABCFH + ABDFG + ABEGH + ACDGH + ACEFG + ADEFH + BCDEFGH$$

Man erkennt, dass nur Koeffizienten von Wechselwirkungen der Ordnung drei oder höher an der Vermengung beteiligt sind. Oft sind die Werte der Koeffizienten von Wechselwirkungen dieser Größenordnung vernachlässigbar klein. Aus diesem Grund wird das Risiko der Fehldeutung der Faktorwirkungen für den teilfaktoriellen 2⁸⁻⁴-Plan geringer eingestuft als für den Plackett-Burman Plan mit 12 Versuchen.

Fazit: Screening-Pläne benötigen deutlich weniger Versuchskombinationen als vollfaktorielle Pläne. Jedoch erkaufte man sich diesen Vorteil mit dem Nachteil der Vermengung. Eine Verfälschung der berechneten Koeffizienten der Faktorwirkungen ist bei Screening-Plänen also immer gegeben. Daher ist vor der Auswahl eines Screening-Planes auch das Risiko der Fehlinterpretation zu beachten. Diese Information erhält man aus der Aliasmatrix.