

RANDOMISATION von VERSUCHSANLAGEN mit Hilfe der Prozedur PLAN von SAS®.

Hanspeter Thöni
Institut für Angewandte Mathematik und Statistik
Fachgebiet **BIOMETRIE**
Universität Hohenheim
D-70593 Stuttgart
thoeni@uni-hohenheim.de

Einleitung.

Die *Notwendigkeit* einer sauberen und einwandfreien *Randomisation* eines Experimentes und die *qualitätssichernden Eigenschaften* hiervon sind seit langem bekannt und in der einschlägigen Literatur breit dokumentiert. Das Prinzip der Randomisation lässt sich zurückverfolgen bis in die Anfänge der biometrischen Literatur und wurde bereits von R.A.FISHER in das Konzept des wissenschaftlichen Denkens eingeführt ¹.

Randomisation dient im Wesentlichen drei Aspekten eines Experiments ^{2,3,4} :

1. Vermeidung von systematischen Verzerrungen der Prüfglied-Effekte
2. Erzeugung von unkorrelierten stochastischen Residuen
3. Validierung der Wahrscheinlichkeitsaussagen im Zusammenhang mit statistischen Tests und Angabe von Vertrauensintervallen bzw. -bereichen.

Grundprinzip aller Randomisationsvorgänge ist :

***Alle möglichen Anordnungen der Prüfglieder (Behandlungen) über die Versuchseinheiten („experimental units“) sind gleichwahrscheinlich.
Aus der Menge aller möglichen gleichwahrscheinlichen Anordnungen wird eine Anordnung zufällig ausgewählt.***

Dieser Vorgang kann nicht „auf's Geratewohl“ („haphazardly“) sozusagen mit der linken Hand vollzogen werden. Es lässt sich nachweisen, dass dabei die Gleichwahrscheinlichkeit aller möglichen Anordnungen nicht erreicht wird, sondern gewisse (individuell unterschiedliche !) Regelmässigkeiten auftreten. Auch genügt es nicht, scheinbar „randomisierte“ Anlagemuster aus Lehrbüchern und/oder älteren Vorarbeiten zu übernehmen. Um das Postulat der zufälligen Auswahl aus der Grundmenge der Möglichkeiten zu erfüllen, muss bei jedem Randomisationsvorgang eine erneute Zufallsauswahl aus allem möglichen Anordnungen erfolgen.

Die nachstehend vorgestellten Programme sollen diesen Vorgang erleichtern. Für eine Reihe von häufig verwendeten Versuchsplänen werden Programmbeispiele aufgeführt :

Versuche mitvollständig zufälliger Zuordnung	CRD
Versuche in randomisierten vollständigen Blöcken	RCBD
Versuche in lateinischen Quadraten	LQ
Versuche in lateinischen Quadraten mit Wiederholungen des Grundplanes	

Versuche in ausgewogenen unvollständigen Blöcken	BIBD
Versuche in YOUDEN'schen Rechtecken	YR
Versuche in ausgewogenen und teilweise ausgewogenen quadratischen Gittern	LATTICE
Zwei-, Dreisatz-Gitter	

Zweifaktorielle Versuchsanlagen in randomisierten vollständigen Blöcken
 Zweifaktorielle Spaltanlagen, geblockt
 Zweifaktorielle Streifenanlagen, geblockt

PROC PLAN.

Die Prozedur PLAN ermöglicht die Erzeugung von Zufallspermutationen einer vorgegebenen Zahlenfolge. Der im Programm eingebaute Zufallszahlengenerator startet mit einer Startzahl, welche im Programm vorgegeben werden kann. Wird keine Startzahl vorgegeben, wählt das Programm die interne Uhrzeit als Startzahl. Um die Randomisationsvorgänge dokumentierbar und reproduzierbar zu machen und dabei die Zufallseigenschaften zu erhalten, ist in den Programmen die Möglichkeit vorgesehen, eine Zufallszahl als Startzahl einzugeben.

PROC PLAN SEED=xxxxxxx ;

Wählt man für xxxxxx eine 5-7-stellige^a ungerade Zahl, z.B. entnommen einer Tafel von Zufallszahlen, so werden die erzeugten Randomisationspläne reproduzierbar und damit wird der Randomisationsvorgang dokumentierbar. Setzt man für jeden Randomisationsvorgang eine neue, zufällig einer Tafel entnommene Ziffernkombination ein, erhält man voneinander unabhängige, gleichwahrscheinliche Randomisationspläne.

^a SAS Institute Inc., SAS/STAT™ User's Guide, Release 6.03 Edition, Cary, NC, 1988. p.737.

Programme.

Die Programme sind so gestaltet, dass die *Parameter der Versuchspläne* als sog. *Makro-Variable* eingegeben werden. So können z.B. Blockgrößen und Prüfglied-Anzahlen unmittelbar in die PROC PLAN übernommen werden.

Prüfglied-Bezeichnungen und *standardisierte Grundpläne*^{5,6,7} (z.B. für Lateinische Quadrate, balanciert unvollständige Block- und Gitterpläne) werden in eine Datei („data-step“) geschrieben.

Für die Angabe der Quelle der Grundpläne ist eine *footnote* vorgesehen.

Die Felder, in welchen Angaben zu den Plänen zu machen sind, sind in den Programm-beschreibungen **grau hinterlegt**.

Versuchspläne.

1. Versuche mit **vollständig zufälliger Zuteilung**.

Es können k Prüfglieder mit jeweiligen n_i Wiederholungen vollständig zufällig auf die N Versuchseinheiten verteilt werden.

Die Eingabe erfordert die Angabe der Gesamtzahl der Versuchseinheiten sowie der Prüf-

glied-Bezeichnungen und die zugehörige Anzahl der Wiederholungen.

Es wird eine geordnete Liste der Versuchseinheiten (Plots) mit dem zugehörigen Prüfglied erstellt.

Das Programmbeispiel basiert auf einem Versuch, im welchem neun Testsorten mit zwei Standardsorten verglichen werden sollen; um eine optimale^b Aufteilung der Zahl der Wiederholungen zum Vergleich zwischen Standard- und Testsorten zu erreichen, werden die Wiederholungszahlen dem Verhältnis

$$\sqrt{9/2}$$

möglichst nahe gebracht.

^b H.THÖNI, Optimale Aufteilung des Stichprobenumfanges zum Vergleich von s Standard- und t Testbehandlungen. *EDV in Medizin und Biologie* **14**, 95-97; 1983.

2. Versuche in *randomisierten vollständigen Blöcken*.

Die Eingabe erfordert die Angabe der Anzahl der Blöcke b und der Zahl der Prüfglieder k sowie die Eingabe der Prüfglied-Bezeichnungen.

Das Programm erstellt eine Liste der b Blöcke mit in jedem Block separat randomisiert angeordneten Prüfgliedern.

3. Versuche in *Lateinischen Quadraten*.

Die Eingabe erfordert die Angabe der Anzahl der Prüfglieder, die Eingabe der Prüfglied-Bezeichnungen, sowie die Eingabe eines Grundplanes (Lateinisches Quadrat in Standardform).

Das Programm erstellt eine randomisierte Form des Quadrates durch zufälliges Permutieren der Zeilen und der Spalten, und ordnet die Prüfgliedern den 'Buchstaben' des Quadrates zufällig zu.

Erstellt wird eine nach Zeilen und Spalten geordnete Liste des randomisierten Quadrats mit Angabe der 'Buchstaben' und der Prüfglied-Bezeichnungen.

4. Versuche in *ausgewogenen unvollständigen Blöcken*.

Die Eingabe erfordert die Angabe der Anzahl der Prüfglieder t , der Anzahl der Wiederholungen r , der Anzahl der Blöcke b , und der Blockgrösse k , die Eingabe der Prüfglied-Bezeichnungen, sowie die Eingabe eines Grundplanes.

Das Programm erstellt eine randomisierte Form des Blockplanes durch zufälliges Permutieren der Blöcke und der Prüfglieder innerhalb der Blöcke, und ordnet die Prüfglieder den 'Buchstaben' des Grundplanes zufällig zu.

Erstellt wird eine nach Blöcken und Versuchseinheiten geordnete Liste des randomisierten Grundplanes mit Angabe der 'Buchstaben' und der Prüfglied-Bezeichnungen.

5. **Versuche in einem unvollständigen lateinischen Quadrat („YOUDEN'sches Rechteck“).**

Die Eingabe erfordert die Angabe der Anzahl der Zeilen r und der Spalten c des unvollständigen lateinischen Quadrates, die Eingabe der Prüfglied-Bezeichnungen, sowie die Eingabe eines Grundplanes.

Das Programm erstellt eine randomisierte Form des Blockplanes durch zufälliges Permutieren der Zeilen und Spalten des Grundplanes, und ordnet die Prüfglieder den 'Buchstaben' des Grundplanes zufällig zu.

Erstellt wird eine nach Zeilen und Spalten geordnete Liste des randomisierten Grundplanes mit Angabe der 'Buchstaben' und der Prüfglied-Bezeichnungen.

6. **Versuche in ausgewogenen und teilweise ausgewogenen quadratischen Gittern.**

Es müssen die Anzahl der Zeilen und Spalten des Gitterquadrats k , die Anzahl der Prüfglieder $t = k^2$ und die Zahl der Wiederholungen $r \leq k+1$, die Prüfglied-Bezeichnungen und der Grundplan eingegeben werden.

Das Programm erstellt eine randomisierte Form durch zufälliges Permutieren der Zeilen und Spalten in jeder Wiederholung, Randomisation der Wiederholungen, und zufällige Zuordnung der Prüfglieder zu den 'Buchstaben'.

Erstellt wird eine nach Wiederholungen, Zeilen und Spalten geordnete Liste des randomisierten Grundplanes mit Angabe der 'Buchstaben' und der Prüfglied-Bezeichnungen.

7. **Zwei-, Dreisatz-Gitter.**

Es müssen die Anzahl der Zeilen und Spalten des Gitterquadrats k , die Anzahl der Prüfglieder $t = k^2$ und die Zahl der Wiederholungen r , die Prüfglied-Bezeichnungen und der Grundplan eingegeben werden.

Das Programm erstellt eine randomisierte Form durch Randomisation der Prüfglieder in jeder Zeile, zufälliges Permutieren der Zeilen in jeder Wiederholung, Randomisation der Wiederholungen, und zufällige Zuordnung der Prüfglieder zu den 'Buchstaben'.

Erstellt wird eine nach Wiederholungen, Zeilen und Spalten geordnete Liste des randomisierten Grundplanes mit Angabe der 'Buchstaben' und der Prüfglied-Bezeichnungen.

8. **Zweifaktorielle Versuche in vollständigen Blöcken.**

8.1 **Zweifaktorielle Versuche in randomisierten vollständigen Blöcken.**

Es werden die Anzahl der Wiederholungen (=Blöcke) r , die Anzahl a der Stufen des Faktors A und die Anzahl b der Stufen des Faktors B sowie die Anzahl $k = ab$ der Faktorenkombinationen (=Blockgröße) und Bezeichnungen für die ab Faktorenkombinationen eingegeben.

Das Programm erstellt eine Liste mit den in jedem Block separat randomisiert angeordneten Faktorenkombinationen.

8.2 Zweifaktorielle Versuche als *Spalt-Anlage* („Split-Plot-Design“).

Es müssen die Anzahl der Wiederholungen (=Blöcke) r , die Anzahl a der Stufen des Faktors A und die Anzahl b der Stufen des Faktors B eingegeben werden.

Das Programm erzeugt eine Liste mit in jedem Block randomisiert angeordneten Stufen des Grossparzellenfaktors A und in jeder Grossparzelle randomisiert angeordneten Stufen des Kleinparzellenfaktors B.

8.3 Zweifaktorielle Versuche als *Streifen-Anlage* („Split-Block-Design“).

Es müssen die Anzahl der Wiederholungen (=Blöcke) r , die Anzahl a der Stufen des Faktors A und die Anzahl b der Stufen des Faktors B eingegeben werden.

Das Programm erzeugt für die r Blöcke je eine Liste mit der randomisierten Anordnung für die beiden Faktoren A und B.

Literatur

¹ FISHER.,R.A. *The Design of Experiments*. Oliver and Boyd, Edinburgh. 8th edition, 1966 [pp.17-21, 41-44].

² HINKELMANN, K., and O.KEMPTHORNE, *Design and Analysis of Experiments. Volume I: Introduction to Experimental Design*. John Wiley & Sons, Inc., New York, 1994 [pp.35,41,128-144].

³ JOHN, P.W.M., *Statistical Design and Analysis of Experiments*. The Macmillan Company, New York, 1971 [p.4].

⁴ JOHN, J.A., and M.H. QUENOUILLE, *Experiments : Design and Analysis*. Charles Griffin & Company Ltd., London, 2nd edition 1977 [p.13, 18-19]

Tafelwerke mit Grundplänen

⁵ W.H.BEYER (Ed.) : *Handbook of Tables for Probability and Statistics*. pp. 64-65 , 84-105. The Chemical Rubber Company, Ohio, 2nd edition 1968.

⁶ R.A.FISHER and F.YATES : *Statistical Tables for Biological, Agricultural and Medical Research*. pp.86-89 Oliver and Boyd, Edinburgh, 6th edition 1963/1974.

⁷ W.G.COCHRAN and Gertrude M.COX : *Experimental Designs*. pp 428-437, 469-482, 497-506, 520-544. J.Wiley & Sons, Inc., New York-London-Sydney, 2nd edition 1957.