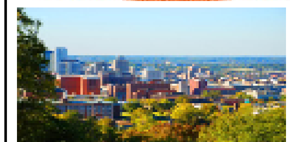
Youden Square to George Box to Latin Hypercube: The Evolution of Experimental Design

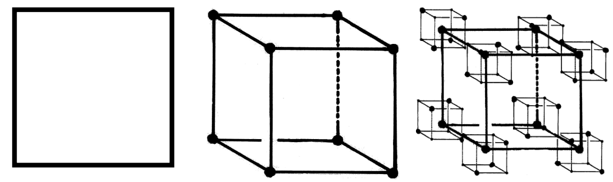
Dennis Lin
University Distinguished Professor
The Pennsylvania State University

Fall Technical Conference
Birmingham, Alabama
07 November, 2010

54th Annual
Fall Technical Conference
Quality & Statistics: The Engines of Success

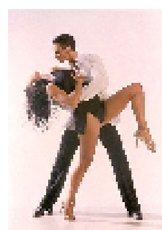
October 6-8, 2010
The Wynfrey Hotel
Birmingham, AL

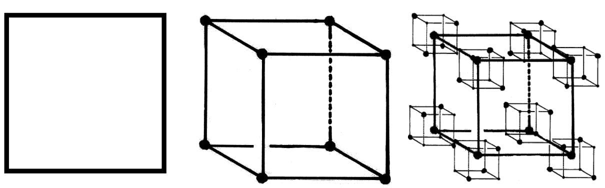




Youden **George** **Latin**



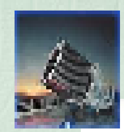




Youden Square, George Box & Latin Hypercube

The Evolution of Science

- **Observational Science**
  - Scientist gathers data by direct observation
  - Scientist analyzes data
- **Analytical Science**
  - Scientist builds analytical model
  - Makes predictions.
- **Computational Science**
  - Simulate analytical model
  - Validate model and makes predictions
- **Data Exploration Science**
  - **Data-driven science**
    Data captured by instruments
    or data generated by simulation
  - Processed by software
  - Placed in a database / files
  - Scientist(s) analyze(s) database / files
  - *Access crucial*



Things are similar,
this makes science *possible*.

Things are different,
this makes science *necessary*.

Statistics based upon data

EDA: Exploratory Data Analysis

(Tukey)

- Learning from Data
- What can the data tell you?

- Show me your problems?
- Show me your data?
- Tell me how were the data collected?

Statistics based upon Probability

- Probability Theory
- Likelihood Function
- Maximal Likelihood Principle
  - Point Estimate
  - Hypothesis Testing
  - Inference
- The role of the data.

Likelihood Principle

- Full Likelihood
- Partial Likelihood
- Empirical Likelihood
- Pseudo Likelihood
- Quasi Likelihood
- Penalized Likelihood
- Posterior Likelihood
- Composite Likelihood
- Profile Likelihood

Randomization in *Theory* vs. Randomization in *Reality*

R.A. Fisher (1920)

How Should the Data be collected?

Randomly

or

Systematically

What type of Jobs?

- *1990—1950+* *Agriculture*
- *1950—1980+* *Industry (Manufacturing)*
- *1980—present* *Service Industry*
  Informatics
- 2010—2050 What's Next?

Design of Experiment-I

- Treatment Comparison
- Complete Randomized Block Design
- Greco-Latin Square
- Youden Square
- Balanced Incomplete Block Design
- Split-Plot Design
- Nested Designs
- Row-Column Design
- etc

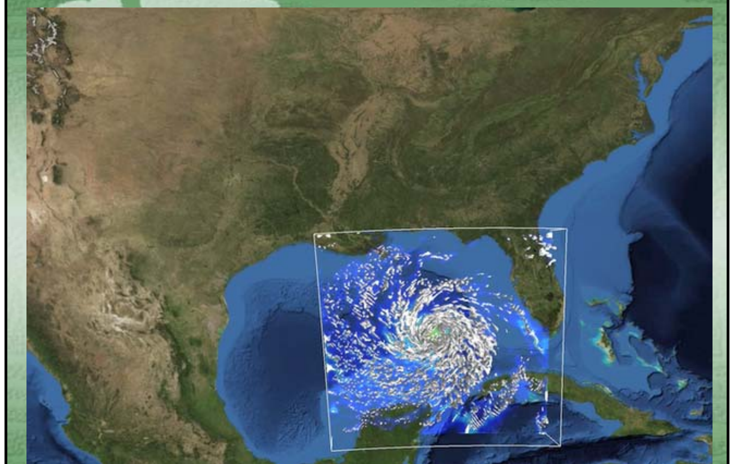
Designing **Industrial** Experiments
is very different from
Designing **Agricultural** Experiments

Lin (1991)

Design of Experiment-II

- Response Surface Design
  - Central Composite Design
  - Box-Behenken Design
- Robust Design
- Fractional Factorial Design
- Orthogonal Arrays
- Supersaturated Design
- Design for Six-Sigma
- Multi-Stage Design

Visualizing High-resolution Hurricane Models



Cloud water in the 1.5-km nested grid WRF/EnKF forecast of Ike plotted by TACC

Designing **Computer** Experiments
is very different from
Designing **Industrial** Experiments

Lin (1998)

- What type of Jobs?
- *Agriculture*
- *Industry (Manufacturing)*
- *Service Industry*
  - *Informatics*
- What's Next?

Youden

Box

Lin

- What type of Designs?
- BIBD/Split/Treatment …
- RSM, Factorials, OA Robust Design, …
- Design for Six Sigma etc
- Computer Experiment

Classroom *vs.* Reality

Classroom
- **Theoretical**
- **Pure Knowledge**
- **Orderly**
- **Pristine**
- **Controlled**
- **General**
- **Covering Laws**
- **Predictions**
- **Certain**

Reality
- **Practical**
- **Applied to problems**
- **Disorderly**
- **Contaminated**
- **Chaotic**
- **Specific**
- **Approximations**
- **Conjectures**
- **Uncertain**

Will Computer Model be another version of Ivory Tower?

Latin Square

α	β	γ	δ
γ	δ	α	β
δ	γ	β	α
β	α	δ	γ

Greco-Latin Square

A	B	C	D
C	D	A	B
D	C	B	A
B	A	D	C

1	2	3	4
4	3	2	1
2	1	4	3
3	4	1	2

1A	2B	3C	4D
4C	3D	2A	1B
2D	1C	4B	3A
3B	4A	1D	2C

A 4x4 Greco-Latin Square

♥ A	♠ K	♣ Q	♦ J
♣ J	♦ Q	♥ K	♠ A
♦ K	♣ A	♠ J	♥ Q
♠ Q	♥ J	♦ A	♣ K

Latin Square & Youden Square

(doubly balanced incompleted block design)

BIBD with
The number of block
is equal to
The number of treatment

ELSEVIER
Computational Statistics & Data Analysis 40 (2002) 329–338
COMPUTATIONAL STATISTICS & DATA ANALYSIS
www.elsevier.com/locate/csda

Missing observations in Youden square designs

Ralph Mansson, Philip Prescott*
Department of Mathematics, University of Southampton, Highfield, Southampton, SO17 1BJ, UK

Received 1 June 2001; received in revised form 1 November 2001; accepted 1 November 2001

Missing Value(s) Problems in Latin Square & Youden Square

Missing one or two,
Systematic "Missing"

Systematic
Missing Values
in
Latin Square

Sudoku

Sudoku

9			1					5
		5		9		2		1
8				4				
				8				
			7					
				2	6			9
2			3					6
			2			9		
		1	9		4	5	7	

Design &Analysis of Sudoku Design

(Fractional Latin/Youden Square)

Computer Experiment

What is Computer Simulation?

What for?

And How?

What to Simulate???

$$y = f(x, \theta) + \varepsilon$$

You Could

- Simulate y
- Simulate f
- Simulate x
- Simulate θ
- Simulate ε

What to Simulate??? More

$$y = f(x, \theta) + \varepsilon$$

You could also

- Simulate $y \mid x$, $x \mid y$, …
- *Simulate* $\theta \mid x$, …
- Simulate $\{u_1, u_2, \dots, u_m\}$
  - Take them all,
  - or use reject-accept strategy;
- Simulate $u_t \mid u_{t-1}$, … etc

All simulations look alike

Did you use the correct simulation???

$$y = f(x, \theta) + \varepsilon$$

Statistics vs. Engineering Models

$$y = f(x, \theta) + \varepsilon$$

Statistical Model, f

$$y = \beta_0 + \Sigma\beta_i x_i + \Sigma\beta_{ij} x_i x_j + \varepsilon$$

A Typical Engineering Model f (page 1 of 3, in Liao and Wang, 1995)

$$
\begin{aligned}
&\rho_b A_b \frac{\partial^2 w}{\partial t^2} + E_b I_b \frac{\partial^4 w}{\partial x^4} \\
&+\{(\rho_s A_s + \rho_c A_c)\frac{\partial^2 w}{\partial t^2} + \rho_s A_s\left(\frac{t_b+t_s}{2}\right)\left(\frac{\partial^3 u_b}{\partial x \partial t^2} - \frac{t_b+t_s}{2}\frac{\partial^4 w}{\partial x^2 \partial t^2} + \frac{t_s}{2}\frac{\partial^3 \beta}{\partial x \partial t^2}\right) \\
&+\rho_c A_c a\left(\frac{\partial^3 u_b}{\partial x \partial t^2} - a\frac{\partial^4 w}{\partial x^2 \partial t^2} + t_s \frac{\partial^3 \beta}{\partial x \partial t^2}\right) + C_{11}^D I_c \frac{\partial^4 w}{\partial x^4} - E_c A_c a\left(\frac{\partial^3 u_b}{\partial x^3} - a\frac{\partial^4 w}{\partial x^4} + t_s \frac{\partial^3 \beta}{\partial x^3}\right)\}[H(x-x_1) - H(x-x_2)] \quad (1) \\
&+\{\rho_s A_s\left(\frac{t_b+t_s}{2}\right)\left(\frac{\partial^2 u_b}{\partial t^2} - \frac{t_b+t_s}{2}\frac{\partial^3 w}{\partial x \partial t^2} + \frac{t_s}{2}\frac{\partial^2 \beta}{\partial t^2}\right) + \rho_c A_c a\left(\frac{\partial^2 u_b}{\partial t^2} - a\frac{\partial^3 w}{\partial x \partial t^2} + t_s \frac{\partial^2 \beta}{\partial t^2}\right) \\
&+2C_{11}^D I_c \frac{\partial^3 w}{\partial x^3} - 2E_c A_c a\left(\frac{\partial^2 u_b}{\partial x^2} - a\frac{\partial^3 w}{\partial x^3} + t_s \frac{\partial^2 \beta}{\partial x^2}\right)\}[\delta(x-x_1) - \delta(x-x_2)] \\
&+\{C_{11}^D I_c \frac{\partial^2 w}{\partial x^2} - E_c A_c a\left(\frac{\partial u_b}{\partial x} - a\frac{\partial^2 w}{\partial x^2} - t_s \frac{\partial \beta}{\partial x}\right) + b d_{31} E_c a V(t)\}[\delta'(x-x_1) - \delta'(x-x_2)] = f(x,t)
\end{aligned}
$$

$$
\begin{aligned}
&\rho_b A_b \frac{\partial^2 u_b}{\partial t^2} - E_b A_b \frac{\partial^2 u_b}{\partial x^2} \\
&+\{\rho_s A_s\left(\frac{\partial^2 u_b}{\partial t^2} - \frac{t_b+t_s}{2}\frac{\partial^3 w}{\partial x \partial t^2} + \frac{t_s}{2}\frac{\partial^2 \beta}{\partial t^2}\right) + \rho_c A_c\left(\frac{\partial^2 u_b}{\partial t^2} - a\frac{\partial^3 w}{\partial x \partial t^2} + t_s \frac{\partial^2 \beta}{\partial t^2}\right) \\
&-E_c A_c\left(\frac{\partial^2 u_b}{\partial x^2} - a\frac{\partial^3 w}{\partial x^3} + t_s \frac{\partial^2 \beta}{\partial x^2}\right)\}[H(x-x_1) - H(x-x_2)] \quad (2) \\
&+\{-E_c A_c\left(\frac{\partial u_b}{\partial x} - a\frac{\partial^2 w}{\partial x^2} - t_s \frac{\partial \beta}{\partial x}\right) + b d_{31} E_c V(t)\}[\delta(x-x_1) - \delta(x-x_2)] = 0
\end{aligned}
$$

$$
\begin{aligned}
&[\rho_s A_s \frac{t_s}{2}\left(\frac{\partial^2 u_b}{\partial t^2} - \frac{t_b+t_s}{2}\frac{\partial^3 w}{\partial x \partial t^2} + \frac{t_s}{2}\frac{\partial^2 \beta}{\partial t^2}\right) + \rho_c A_c t_s\left(\frac{\partial^2 u_b}{\partial t^2} - a\frac{\partial^3 w}{\partial x \partial t^2} + t_s \frac{\partial^2 \beta}{\partial t^2}\right) \\
&+A_s(G*\beta) - E_c A_c t_s\left(\frac{\partial^2 u_b}{\partial x^2} - a\frac{\partial^3 w}{\partial x^3} + t_s \frac{\partial^2 \beta}{\partial x^2}\right)][H(x-x_1) - H(x-x_2)] \quad (3)
\end{aligned}
$$

"Statistical" Simulation Research

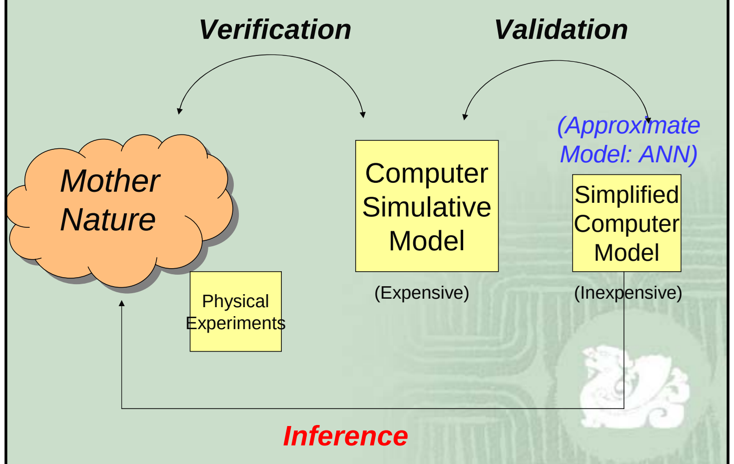
- Random Number Generators
  - Deng and Lin (1997, 2001, 2007)
- Robustness of transformation (Sensitivity Analysis)
  - From Uniform random numbers to other distributions

"Engineering" Computer Experiments

Mostly deterministic
Many input variables
Time consuming
Grid Search is not feasible

Analysis of Computer Experiments

- Complicate mean model, with relatively simple error structure
  - Polynomial model for mean model
  - $\varepsilon \sim N(0,\sigma^2)$ for error
- Simple mean model, with relatively complicated error structure
  - Gaussian Process Model
    - Intercept model for mean
    - Matern Covariance for error
- Comparisons on pros & cons: Theoretically and Empirically.



Irrelevant Issues

- Replicates
- Blocking
- Randomization

Question: How can a computer experiment be run in an efficient manner?

Lin (1997)

Space Filling Design

- Original Problem Setup
- Uniform Design
  - Fang and Wang (1982)
  - Fang, Lin, Winker & Yang (*Technometrics*, 1999)
  - Fang and Lin (Handbook of Statistics, Vol 22, 2003)
- Latin Hypercube Design
  - McKay, Beckman & Conover (1979)
- Orthogonal Latin Hypercube
  - Beattie and Lin (1997)
  - Steinberg and Lin (2006, *Biometrika*)
  - Sun, Liu and Lin (2009, *Biometrika*)

Uniform Design

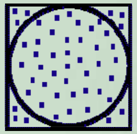
A uniform design provides uniformly scatter design points in the experimental domain.

http://www.math.hkbu.edu.hk/UniformDesign

How to estimate π ?

- Randomly (uniformly) drop n points into the square, suppose that there are a points fell in the circle. Then…

2r



$$\frac{\pi}{4} = \frac{\pi r^2}{4r^2} = \frac{a}{n}$$

$$\pi = \frac{4a}{n}$$

Now, suppose I do know π (=3.14159…),
how could I know how uniform are these points?

L_p-star Discrepancy

$$D_p(P) = \left[\int_{C^s} \left| \frac{|P \cap [0,\mathbf{x})|}{n} - \mathrm{Vol}([0,\mathbf{x})) \right|^p dx \right]^{1/p}$$

where

- $[0,\mathbf{x}) = [0,x_1) \times [0,x_2) \times \cdots \times [0,x_s)$;
- $|P \cap [0,\mathbf{x})|$: **the number of points of** P **falling in** $[0,\mathbf{x})$;
- $d_p([0,\mathbf{x})) = \left| \frac{|P \cap [0,\mathbf{x})|}{n} - \mathbf{Vol}([0,\mathbf{x})) \right|$ is called the discrepancy of P over the rectangular [0,x);

$D_p(P)$ is called the L_p -star discrepancy of the set P.

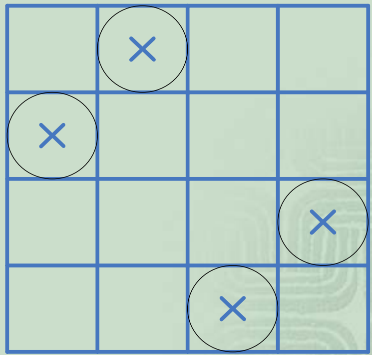
Uniform Design: Summary

- Uniformity
- Model Robustness
- Flexibility in experimental runs
- Flexibility in the number of levels

References

- Fang and Lin (2003)
  Handbook of Statistics, Statistics in Industry (Vol.22).
- Fang, Lin, Winker and Zhang
  (Technometrics, 2000)
- Website
  www.math.hkbu.edu.hk/UniformDesign

What is a Latin Hypercube?



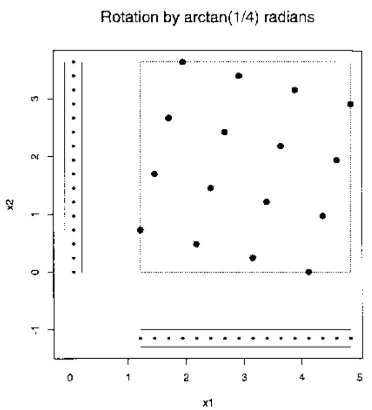
Bayesian Designs

- Maximin Distance Designs, Johnson, Moore,and Ylvisaker (1990)
- Maximizes the Minimum Interpoint Distance (MID)
- Moves design points as far apart as possible in design space $MID = \min_{x_1,x_2 \in D} d(x_1,x_2)$
- D* is a Maximin Distance Design if

$$MID = \min_{x_1,x_2 \in D^*} d(x_1,x_2) = \max_{D} \min_{x_1,x_2 \in D} d(x_1,x_2)$$

Rotated Factorial Designs

Beattie and Lin (1997)



- Rotation Theorem
- Orthogonality Theorem

Theorem 1

Rotation Theorem *For nontrivial rotations between 0 and 45 degrees, a rotated standard* p^2 *factorial design will produce equally-spaced projections to each dimension if and only if the rotation angle is* $\tan^{-1}(1/k)$ *where* $k \in \{1,\ldots,p\}$. *These equally-spaced projections will be unique if and only if the rotation angle is* $\tan^{-1}(1/p)$.

Theorem 2

Orthogonality Theorem *Any rotated standard factorial design, regardless of the rotation angle, has uncorrelated regression effect estimates (that is, orthogonal design matrix columns).*

Orthogonal Latin Hypercube Designs

$$D = X \cdot V$$

- D: desirable design
- X: factorial design
- V: rotated matrix

$$X = \begin{bmatrix} 1 & 1 \\ 2 & 1 \\ \vdots & \vdots \\ p & 1 \\ 1 & 2 \\ 2 & 2 \\ \vdots & \vdots \\ p & 2 \\ \vdots & \vdots \\ 1 & p \\ 2 & p \\ \vdots & \vdots \\ p & p \end{bmatrix}_{p^2 \times d} \qquad V = \begin{bmatrix} \upsilon_1 & \upsilon_3 \\ \upsilon_2 & \upsilon_4 \end{bmatrix}_{d \times d}$$

Beattie & Lin (1998): Rotating Full Factorials

$$D = X \cdot V$$

- D: desirable design
- X: Two-level fractional factorial design
- V: rotated matrix

$$V = \begin{bmatrix} S_1 & & & & \\ & S_2 & & & \\ & & \ddots & & \\ & & & S_i & \\ & & & & \ddots \\ & & & & & S_m \end{bmatrix}_{n \times n}$$

Bursztyn & Steinberg (2002): Rotating in Groups

Now,
Put these two ideas together!

- Grouping all design columns into groups,
- each forms a full factorial design,
- then rotate each group (in block).

LHD's as Rotated Factorial Designs

Steinberg and Lin:

$$DR = [D_1 | \cdots | D_t] \begin{bmatrix} R & 0 & 0 \\ 0 & \ddots & 0 \\ 0 & 0 & R \end{bmatrix}$$

$$= [D_1 R | \cdots | D_t R]$$

Bursztyn & Steinberg

Lin & Beattie

The resulting design is an orthogonal Latin hypercube.

Grouping Example (16 runs)

- Full (2^4=16) Factorial Design
  - Basic factors: *a, b, c, d*
- Fractional Factorial
  - Basic Factors: *a, b, c, d*
  - Generators:
    ab, ac, ad, bc, bd, cd, abc, abd, acd, bcd, abcd
- Grouping into three: each form a full factorial
  - (*a,b,c,d*),
  - (*ab,ac,ad,abcd*), and
  - (*abc, abd, acd, bcd*)

Steinberg & Lin (2006, *Biometrika*)

Table 1: *Example 1. An orthogonal Latin hypercube design for 12 input factors in 16 runs. The numbers in the table should be divided by 15 to scale the design to the unit hypercube*

1	2	3	4	5	6	7	8	9	10	11	12
−15	5	9	−3	7	11	−11	7	−9	3	−15	5
−13	1	1	13	−7	−11	11	−7	−1	−13	−13	1
−11	7	−7	−11	13	−1	−1	−13	9	−3	15	−5
−9	3	−15	5	−13	1	1	13	1	13	13	−1
−7	−11	11	−7	11	−7	7	11	5	15	−3	−9
−5	−15	3	9	−11	7	−7	−11	13	−1	−1	−13
−3	−9	−5	−15	1	13	13	−1	−5	−15	3	9
−1	−13	−13	1	−1	−13	−13	1	−13	1	1	13
1	13	13	−1	−9	3	−15	5	11	−7	7	11
3	9	5	15	9	−3	15	−5	3	9	5	15
5	15	−3	−9	−3	−9	−5	−15	−11	7	−7	−11
7	11	−11	7	3	9	5	15	−3	−9	−5	−15
9	−3	15	−5	−5	−15	3	9	−7	−11	11	−7
11	−7	7	11	5	15	−3	−9	−15	5	9	−3
13	−1	−1	−13	−15	5	9	−3	7	11	−11	7
15	−5	−9	3	15	−5	−9	3	15	−5	−9	3

Steinberg and Lin (2006, *Biometrika*)

Biometrika (2006), 93, 2, *pp.* 279–288
© 2006 Biometrika Trust
Printed in Great Britain

A construction method for orthogonal Latin hypercube designs

By DAVID M. STEINBERG
Department of Statistics and Operations Research, Tel-Aviv University, Tel-Aviv 69978, Israel
dms@post.tau.ac.il

AND DENNIS K. J. LIN
Department of Supply Chain & Information Systems, The Pennsylvania State University, University Park, Pennsylvania 16802, U.S.A.
dkl5@psu.edu

Pang, Liu and Lin (*Statistica Sinica*, 2009)

Statistica Sinica 19 (2009), 1721-1728

A CONSTRUCTION METHOD FOR ORTHOGONAL LATIN HYPERCUBE DESIGNS WITH PRIME POWER LEVELS

Fang Pang[1], Min-Qian Liu[1] and Dennis K. J. Lin[2]

[1] *Nankai University and* [2] *The Pennsylvania State University*

Abstract: Latin hypercube design (LHD) is popularly used in designing computer experiments. This paper explores how to construct LHDs with p^d ($d = 2^c$) runs and up to $(p^d - 1)/(p-1)$ factors in which all main effects are orthogonal. This is accomplished by rotating groups of factors in a p^d-run regular saturated factorial design. These rotated factorial designs are easy to construct and preserve many attractive properties of standard factorial designs. The proposed method covers the one by Steinberg and Lin (2006) as a special case and is able to generate more orthogonal LHDs with attractive properties. Theoretical properties as well as the construction algorithm are discussed, with an example for illustration.

Key words and phrases: Computer experiment, factorial design, galois field, rotation.

$$D = X \cdot V$$

desirable design (D)

p-level fractional factorial design (X)

rotated matrix (V)

$$\begin{bmatrix} S_1 & & & & & \\ & S_2 & & & & \\ & & \ddots & & & \\ & & & S_i & & \\ & & & & \ddots & \\ & & & & & S_m \end{bmatrix}_{n \times n}$$

Pang, Liu & Lin (2009)
Rotating in Groups for
P-level (not 2-level)

Ye (1998, JASA)

Table 1. A 5 × 2 Orthogonal Latin Hypercube

1	−2
2	1
0	0
−1	2
−2	−1

Table 2. A 9 × 4 Orthogonal Latin Hypercube

1	−2	4	3
2	1	3	−4
3	−4	−2	−1
4	3	−1	2
0	0	0	0
−4	−3	1	−2
−3	4	2	1
−2	−1	−3	4
−1	2	−4	−3

Second-Order Orthogonality

(a) All main effects are orthogonal, and

(b) All main effects are orthogonal to all quadratic & two factor interactions.

Biometrika (2009), **96**, 4, *pp.* 971–974
doi: 10.1093/biomet/asp058
Advance Access publication 26 October 2009
Printed in Great Britain

Miscellanea

Construction of orthogonal Latin hypercube designs

By FASHENG SUN, MIN-QIAN LIU

The Key Laboratory of Pure Mathematics and Combinatorics, School of Mathematical Sciences, Nankai University, Tianjin 300071, China

sfxsfx2001@mail.nankai.edu.cn mqliu@nankai.edu.cn

AND DENNIS K. J. LIN

Department of Statistics, The Pennsylvania State University, University Park, Pennsylvania 16802, U.S.A.

dkl5@psu.edu

Second-Order Orthogonality

Sun, Liu & Lin (2009, Biometrika)

THEOREM 1. (i) *The T_c in* (2) *consists of rows and columns of permutations of the 2^c elements $1, \ldots, 2^c$, up to sign changes.*

(ii) *The L_c in* (3) *is a Latin hypercube design $L(2^{c+1}+1, 2^c)$ with properties* (a) *and* (b).

(a) each column is orthogonal to the others in the design;

(b) the elementwise square of each column and the elementwise product of every two columns are orthogonal to all columns in the design.

THEOREM 3. *If $L(n,k) = (l_{ij})$ is a centered Latin hypercube design with properties* (a) *and* (b), *then $k \le \lfloor n/2 \rfloor$, where $\lfloor x \rfloor$ is the integer part of x.*

Orthogonal Latin Hypercube (n=2^c +1 or 2^c)

Design	Ye (1998) *JASA*	C&L (2007) *Technometrics*	S&L (2006) *Biometrika*	PLL (2009) *Sinica*	SLL (2009) *Biometrika*
No. of Factor	2(c-1)	c(c-1)/2+1	c[(n-1)/c]	c[(n-1)/c/(q-1)]	2^{c-1}
c=4 c=8 c large	6 14	7 29	12 -	12 -	8 256
Main Orthog	Yes	Yes	Yes	Yes	Yes
Second-Order Othog	**Yes**	yes	No	No	**Yes**

Sun, Liu and Lin (2009, *Biometrika*)

1. Let $M_1 = \begin{pmatrix} 1 & 2 \\ 2 & 1 \end{pmatrix}$, $\quad S_1 = \begin{pmatrix} 1 & 1 \\ 1 & -1 \end{pmatrix}$.

2. Let $M_c = \begin{pmatrix} M_{c-1} & M_{c-1} + 2^{c-1} J_{2^{c-1}} \\ M_{c-1} + 2^{c-1} J_{2^{c-1}} & M_{c-1} \end{pmatrix}$,

$$S_c = \begin{pmatrix} S_{c-1} & -S^*_{c-1} \\ S_{c-1} & S^*_{c-1} \end{pmatrix}.$$

$$M_2 = \left(\begin{array}{cc|cc} 1 & 2 & 3 & 4 \\ 2 & 1 & 4 & 3 \\ \hline 3 & 4 & 1 & 2 \\ 4 & 3 & 2 & 1 \end{array}\right), \quad S_2 = \left(\begin{array}{cc|cc} 1 & 1 & 1 & 1 \\ 1 & -1 & -1 & 1 \\ \hline 1 & 1 & -1 & -1 \\ 1 & -1 & 1 & -1 \end{array}\right)$$

Sun, Liu and Lin (2009, *Biometrika*)

- Let $T_c = M_c \odot S_c$, i.e., $(T_c)_{ij} = (M_c)_{ij} \cdot (S_c)_{ij}$.

$$T_2 = M_2 \odot S_2 = \begin{pmatrix} 1 & 2 & 3 & 4 \\ 2 & -1 & -4 & 3 \\ 3 & 4 & -1 & -2 \\ 4 & -3 & 2 & -1 \end{pmatrix}, \ldots$$

- $L_c = \begin{pmatrix} T_c \\ \mathbf{0}'_{2^c} \\ -T_c \end{pmatrix}$ is an OLHD with $n = 2^{c+1} + 1$ runs and $m = 2^c$ factors.

$L_2 = \begin{pmatrix} T_2 \\ \mathbf{0}'_4 \\ -T_2 \end{pmatrix}$ is an OLHD with $n = 9$ runs and $m = 4$ factors.

$$M_2 = \left(\begin{array}{cc|cc} 1 & 2 & 3 & 4 \\ 2 & 1 & 4 & 3 \\ \hline 3 & 4 & 1 & 2 \\ 4 & 3 & 2 & 1 \end{array}\right), \quad S_2 = \left(\begin{array}{cc|cc} 1 & 1 & 1 & 1 \\ 1 & -1 & -1 & 1 \\ \hline 1 & 1 & -1 & -1 \\ 1 & -1 & 1 & -1 \end{array}\right)$$

$$T_2 = M_2 \odot S_2 = \begin{pmatrix} 1 & 2 & 3 & 4 \\ 2 & -1 & -4 & 3 \\ 3 & 4 & -1 & -2 \\ 4 & -3 & 2 & -1 \end{pmatrix} \qquad L_2 = \begin{pmatrix} T_2 \\ \mathbf{0}'_4 \\ -T_2 \end{pmatrix}$$

$n \times k$ matrix $L(n, k)$: n runs, k factors, each factor includes n uniformly spaced levels.

$$L_2 = \begin{pmatrix} T_2 \\ \mathbf{0}'_4 \\ -T_2 \end{pmatrix} \quad L(9,4) = \begin{pmatrix} 1 & 2 & 3 & 4 \\ 2 & -1 & -4 & 3 \\ 3 & 4 & -1 & -2 \\ 4 & -3 & 2 & -1 \\ 0 & 0 & 0 & 0 \\ -1 & -2 & -3 & -4 \\ -2 & 1 & 4 & -3 \\ -3 & -4 & 1 & 2 \\ -4 & 3 & -2 & 1 \end{pmatrix}$$

Beyond
Orthogonal Latin Hypercube

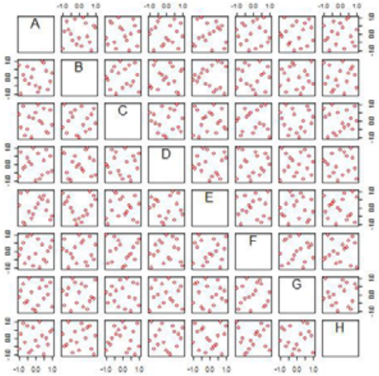
Near-Orthogonal Latin Hypercube

&

Orthogonal near-Latin Hypercube

(Nguyen, Steinberg and Lin, 2010)

8 factors in 17 runs
—Nguyen and Lin (2010)



After all, simulation means “not real”

Good for “description,”

But

Not necessary good for a solid proof!

There are many types of simulations, they must be used with care!

Looking Ahead:

Design and Analysis

All-in-One

- Telephone
- Television
- Radio
- Computer
- Recorder
- Game (Compute
- Camera
- Internet
- etc

ipod





Future Design & Analysis

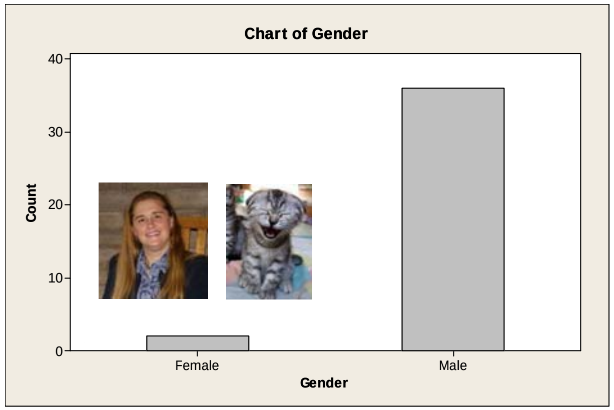
Number
Text
Image
Voice
Film
Web data: Facebook. Google
Video (YouTube)
etc

All-In-One Data

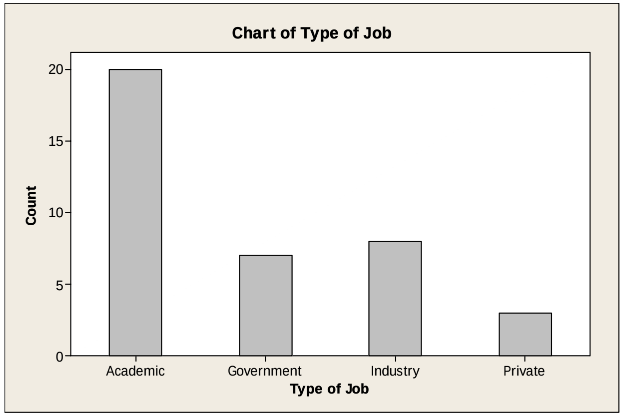
Example: Youden Speakers

1973	**Box**		1986	**Schilling**		1999	**Bacon**
1974	**Daniel**		1987	**Hahn**		2000	**Vining**
1975	**Eisenhart**		1988	**Godfrey**		2001	**Myers**
1976	**Smith**		1989	**Marquardt**		2002	**Meeker**
1977	**Hunter**		1990	**Golomski**		2003	**Hahn**
1978	**Mandel**		1991	**Lawton**		2004	**Woodall**
1979	**Freund**		1992	**Cornell**		2005	**Bisgaard**
1980	**Nelson**		1993	**Hare**		2006	**Cornell**
1981	**Smith**		1994	**Hoerl**		2007	**Nair**
1982	**Nicholson**		1995	**MacGregor**		2008	**Anderson-Cook**
1983	**Hunter**		1996	**Lucus**		2009	**Piepel**
1984	**Joiner**		1997	**Ranney**		2010	**Lin**
1985	**Snee**		1998	**Montgomery**			

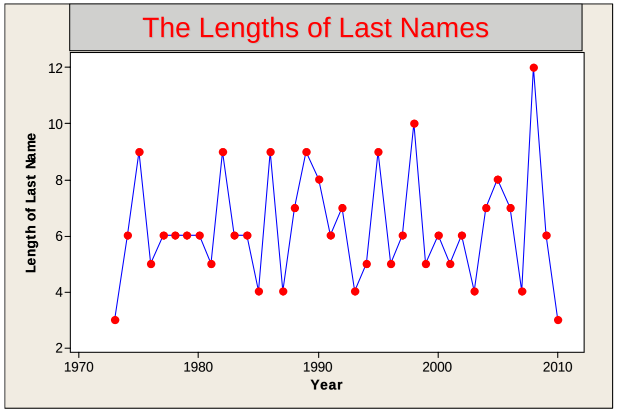
Gender



Affiliations



What is this?



Youden Squear, George Box and Latin Hypercube: Conclusion

- Youden Square is not a Square
  - it's in fact a rectangle!
- George Box is not a box
  - AND he did not invent Box plot!
- Latin Hypercube is not a Latin
  - Although it is indeed a hypercube!
- New design (data collection) concept is needed for the informatic/computational age.

Remembering Youden (1900—1971)

- PhD in Chemistry (1924, Columbia Univ)—An Excellent Chemist!
- First (formal) Statistical Course, through Hotelling (Columbia University) in 1932.
- Youden Square, termed by Fisher, in 1938.
- Youden Diagram in 1959.
- Youden will never be abel to receive COPPS Award (40-)
- Spent his life improving the way measurements are taken.

Remembering Box (1910—)

- ***All Modes are wrong,***
  some are useful.

- ***Statistician, like artist,***
  has the bad habit of
  easily falling in love with his model.

Remember Today (Lin)

- 1997—Rotated Full Factorials (to form a new class of Latin Hypercube)
- Orthogonal Latin Hypercube
- Second-Order Orthogonal Latin Hypercube
  - Steinberger and Lin (*Biometrika, 2006*)
  - Pang, Liu and Lin (*Statistica Sinica, 2009*)
  - Sun, Liu and Lin (Biometrika, 2009)
- Nearly Orthogonal Latin Hypercube (for flexible run sizes)
- New design (data collection) concept is needed for the informatic/computational age.

STILL QUESTION?

Send $500 to

- Dennis Lin
  University Distinguished Professor
  317 Thomas Building
  Department of Statistics
  Penn State University

- +1 814 865-0377 (phone)
- +1 814 863-7114 (fax)
- DKL5@psu.edu

(Customer Satisfaction or your money back!)