Recent Advances on Computer Experiment

Dennis Lin
The Pennsylvania State University
DKL5@psu.edu

12 February, 2008
OR Seminar
Penn State

All Chinese Look Alike? Why?

- (US) criteria for people classification (as in your driver license):
  - Height — *Short*
  - Weight — *Light*
  - Hair Color — *Black*
  - Eye Color — *Black*

You must simulate under the "correct" (right subject/model).

Where have all the Data gone?

- No need for data (Theoretical Development)
- Survey Sampling and Design of Experiment (Physical data collection)
- Computer Simulation (Experiment)
  - Statistical Simulation (Random Number generation)
  - Engineering Simulation
- Data from Internet
  - On-line auction
  - Search Engine

Computer Experiment

What is Computer Simulation?
What for?
And How?

Simulation

A device that enables the operator to reproduce under test conditions phenomena likely to occur in actual performance —Webster

Under H_o is true

Computer Experiment

- Stochastic
  Deterministic
- Expensive (really expensive)
  Inexpensive (really cheap)

Model function

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

Response
Output
Dependent Variable

Input
Predicted Variable
Independent Variable

Parameter

Error
Noise

Parameter

Constant
c_1, c_2, …

What to Simulate???

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

You Could
- Simulate *y*
- Simulate *f*
- Simulate *x*
- Simulate θ
- Simulate ε

What do you mean by simulation??

- A sequence of points that follows
  - a specific desirable distribution π
    - Simple π
    - Complicated π
  - geometrical properties
    - Equal spacing (Uniform)
    - orthogonal
- Independent?
- Conditional Independent?

What to Simulate??? More

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

You could also

Simulate *y |x,*

Simulate θ *|x, …*

Simulate $\{u_1, u_2, \ldots, u_m\}$

Take them all,

or use reject-accept strategy;

Simulate $u_t|u_{t-1}$, … etc

Did you use the correct simulation???

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

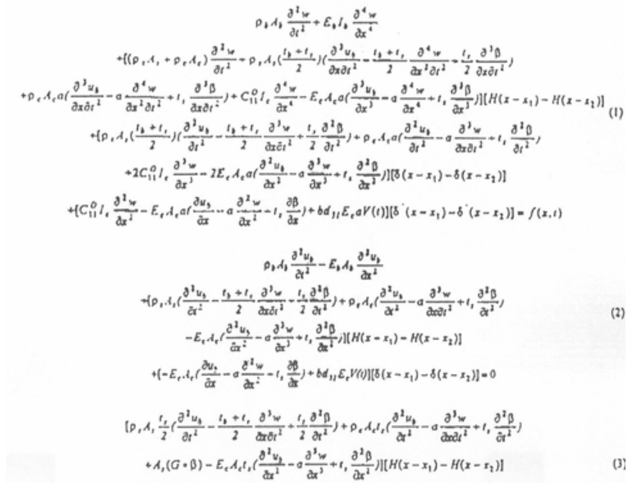
Statistics vs. Engineering Models

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

Statistical Model, f

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

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



In this Talk

- Distribution Theory: Coin Example
- Distribution Theory: R^2 Story
- Distribution Theory: Significance
- Bootstrapping
- MCMC (Markov Chain Monte Carlo)
- Random Number Generation
- Orthogonal Latin Hypercube Design and more
- Uniform Design
- Strategy for Computer Experiments
- Special Topics

If time permits…

- Bootstrapping (Re-sampling)
  - Sampling from the “samples”
    - Treat “Samples” as “Population”
    - Alternatively, delete some data from the “samples”
  - Evaluate “statistic” from the current sample—obtain one value
  - Repeat this many times
  - Average these “statistics” as the estimate
- MCMC (Markov Chain Monte Carlo)
  - OR-seminar by Dr. Murali Haran (Nov 2007)

Distribution Theory

Coin Example

R^2 Story

Significance in Regression Model

Statistical Hypothesis Testing

H_0: Null Hypothesis

vs

H_1: Alternative Hypothesis

Statistical Hypothesis Testing

Does Not Prove anything, but

Could be powerful to

Disapprove “something” (H_0)

Statistical Hypothesis Testing

1. Under the Null Hypothesis H_o (typically implies the white noise) , what will the test statistics behave?
   - what is “typical” and what is “abnormal”?
2. Now, compare your “observed” test statistic to the distribution in (1)
   - If it is “typical”→ accept Null Hypothesis (H_0)
   - If it is “abnormal” → Reject H_o

Coin Example

Throw a fair coin 200 times and the results were recorded: One sequence is real and the other one is fake.

(A)
1111000000000001000001011000000100000101001111001100
01111110010110110101101001111001100011011101100000
10001001111110100100001011001011101101110001010010
01100111111100011100101000101001110011100010100111

(B)
01110010010010100010011110010100010011010111001110
01111011010111101101001000111001101011010101101001
00101001110110100100001110101101101001110101100110
01110011110110001110011010111001110011110010100111

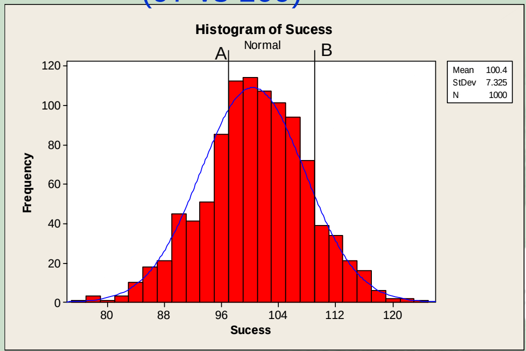
Potential Indexes/Statistics

- Number of Heads (1's)
- Number of Sign Changes
- Maximal Length

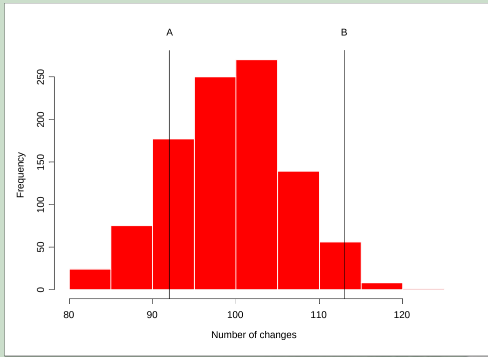
Computer Simulation

- Assumption: For a fair coin (50-50)
- Do the followings
  - Toss the coin 200 times
  - From this 200 "0-1" sequence, evaluate its statistic (say, number of 1's)
  - Now, repeat this process for 1000 times (say), and you will have 1000 statistics.
  - Put these statistics in a histogram (this is the distribution of such a statistic)

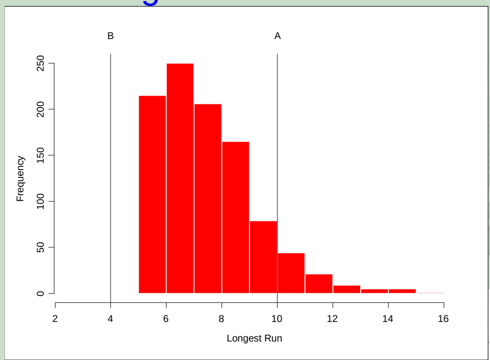
Number of 1's in 1000 Simulations
(97 vs 109)



Numbers of Change in 1000 Simulations



Max-Length in 1000 Simulations



Regression Model: JMP Demo

Significance

R^2

“Statistical” Simulation Research

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

Goodness of Random Number Generators

- Period Length
- Efficiency
- Portability
- Theoretical Justification:
  - Uniformity
  - Independence
- Empirical Performance
  - Small & Big Crash Tests

LCG: Linear Congruential Generator
Classical Random Number Generators

- $X_t = (B\,X_{t-1} + A) \mod m$

  Length=m

  Lehmer (1951); Knuth (1981)
- With proper choice of A & B

  Length=m=2^{31}-1=2147483647 (=2.1×10^9)

Random Number Generation for the New Century

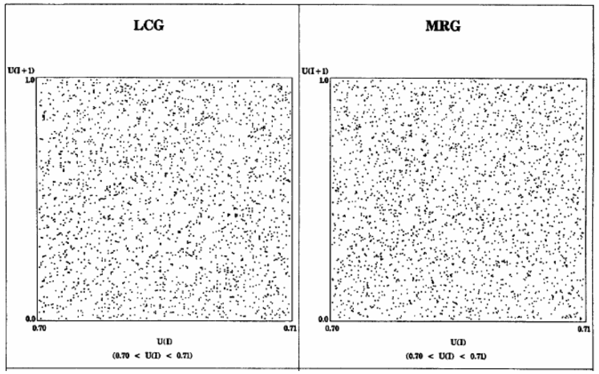
Lih-Yuan DENG and Dennis K. J. LIN

Use of empirical studies based on computer-generated random numbers has become a common practice in the development of statistical methods, particularly when the analytical study of a statistical procedure becomes intractable. The quality of any simulation study depends heavily on the quality of the random number generators. Classical uniform random number generators have some major defects—such as the (relatively) short period length and the lack of higher-dimension uniformity. Two recent uniform pseudo-random number generators (MRG and MCG) are reviewed. They are compared with the classical generator LCG. It is shown that MRG/MCG are much better random number generators than the popular LCG. Special forms of MRG/MCG are introduced and recommended as the random number generators for the new century. A step-by-step procedure for constructing such random number generators is also provided.

KEY WORDS: Linear congruential generator (LCG); Matrix congruential generator (MCG); Multiple recursive generator (MRG); Portable and efficient generator.

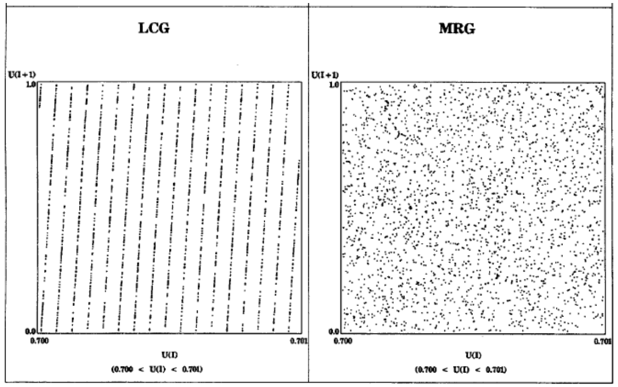
Deng & Lin (2000)
The American Statistician

Dependence: y_t vs y_{t-1}



Range=(0.70, 0.71)

Dependence: y_t vs y_{t-1}



Range=(0.700, 0.701)

Briefings & Update

- We have found a system of random number generators breaking the current world record. (Recall $p=2^{31}-1$ is about 10^9)

  Old world record:
  - MT19937 (1998)
  - Period length $2^{19937}-1=10^{6001.6}$

  New record with $p=2^{31}-1$:
  - DX-1597 [Deng, 2005]
  - Period length: $10^{14903.1}$
- Longest Period found so far:
  - Deng and Lin (2007)—A Penn State Patent
  - Period= 10^{69980}.
  - Survived from all (Small & Big Crash) Tests

Normal Random Numbers: Examples

- Central Limit Theorem
  - X_i~iid U(0,1) → $Z=\Sigma X_i-6$
- Box-Muller Transformation
  - X_i~ ind U(0,1), i=1 & 2 →

    $Z_1=\sqrt{-2\ln x_1}\cos(2\pi x_2)$

    $Z_2=\sqrt{-2\ln x_1}\sin(2\pi x_2)$.
- Rejection Polar Method

Other Approaches

- Kinderman and Ramage (1976)
- Triangular Acceptance/Rejection Method
- Trapezoidal Method
  - (Ahrens, 1977)
- Ratio of Uniform
  - (Kinderman & Monahan, 1976)
- Rectangle/Wedge/Tail Method
  - (Marsaglia, Maclaren & Bray, 1964)

Box-Muller Transformation

$$z_1 = \sqrt{-2\ln x_1}\cos(2\pi x_2) \qquad z_2 = \sqrt{-2\ln x_1}\sin(2\pi x_2).$$

$$x_1 = e^{-(z_1^2+z_2^2)/2} \qquad x_2 = \frac{1}{2\pi}\tan^{-1}\left(\frac{z_2}{z_1}\right).$$

$$\frac{\partial(x_1, x_2)}{\partial(z_1, z_2)} = \begin{vmatrix} \frac{\partial x_1}{\partial z_1} & \frac{\partial x_1}{\partial z_2} \\ \frac{\partial x_2}{\partial z_1} & \frac{\partial x_2}{\partial z_2} \end{vmatrix} = -\left[\frac{1}{\sqrt{2\pi}} e^{-z_1^2/2}\right]\left[\frac{1}{\sqrt{2\pi}} e^{-z_2^2/2}\right].$$

"Engineering" Computer Experiments

Mostly deterministic

Many input variables

Time consuming

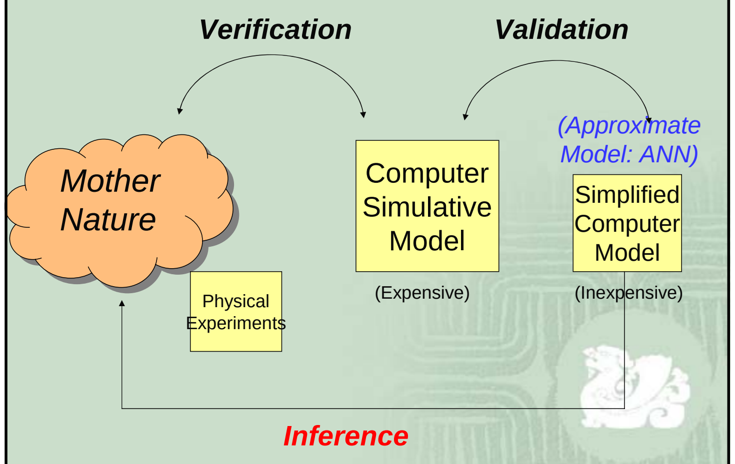
Goals—Computer Experiment

- Confirmation
- Sensitivity Analysis
- Empirical Model Building
- Optimization
- Model Validation
- High Dimension Integration

Space Filling Design

How to (optimally) put n points in d dimensional space?

Optimal=cover as much space as possible



A Structured Roadmap for Verification and Validation-- Highlighting the Critical Role of Experiment Design

James J. Filliben

National Institute of Standards and Technology
Information Technology Division
Statistical Engineering Division

2004 Workshop on Verification & Validation of Computer Models of High-Consequence Engineering Systems
NIST Administration Building
Lecture Room D
3:10-3:25, November 8, 2004

Computer Experiment

- Expensive simulation
- When Monte Carlo study is infeasible, how to run simulation?
- Latin Hypercube

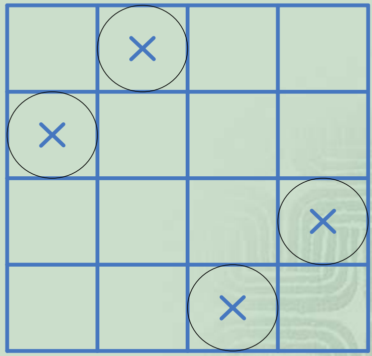
Irrelevant Issues

- Replicates
- Blocking
- Randomization

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

Lin (1997)

What is a Latin Hypercube?



Why Latin Hypercube Designs?

- Replication is worthless in CEs
- Factor levels are easily changed in CEs (not so in PEs)
- Suppose certain terms have little influence
  - Factorial designs produce replication when terms dropped
  - Can estimate high-order terms for other factors
- Provides pseudo-randomness since CEs are deterministic
- Smaller variance than random sampling or stratified random sampling (McKay, Beckman, and Conover (1979)

A special class of LHC

x_1	x_2
1	τ_1
2	τ_2
3	τ_3
4	τ_4
.	.
.	.
.	.
16	τ_{16}

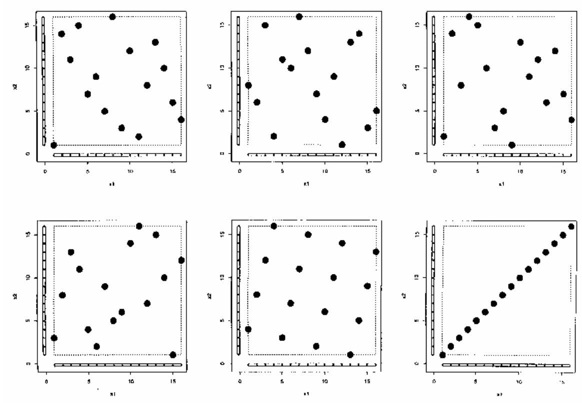
τ_i: permutation of {1, …, 16}

16!

n! for size *n* &

$(n!)^{d-1}$ for *d*-dim

Some Latin Hypercube Designs

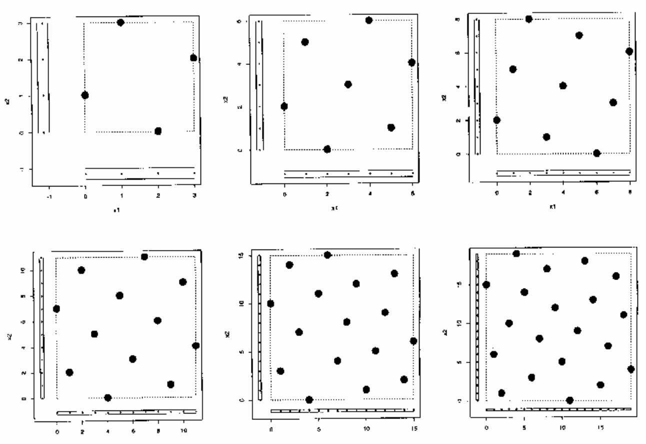


Bayesian Designs

- 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

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

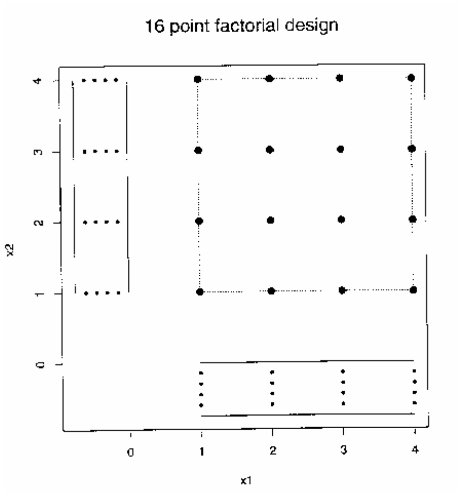
Maximin Latin Hypercube Designs

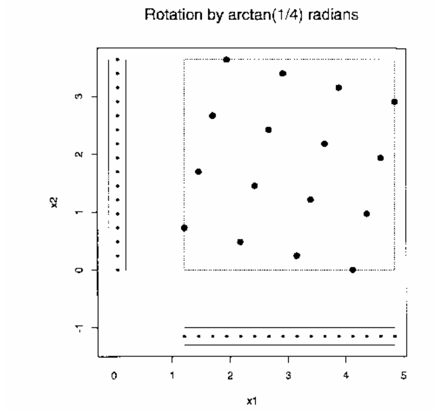


Rotated Factorial Designs

- Computer experiments are gaining in popularity
  - One main research area of the next 10 years
- Rotated factorial designs
  - good factorial design properties (orthogonality and structure)
  - good Latin hypercube properties (unique and equally-spaced projections)
  - easy to construct
  - comparable by Bayesian criteria
  - very suitable for computer experiments

Lin (1997)





- Rotation Theorem
- Orthogonality Theorem

$$D = X \cdot V$$

desirable design — factorial design — rotated matrix

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

For d = 2

$$V_1 = [v_1 \quad v_2] = \begin{bmatrix} +1 & +p \\ +p & -1 \end{bmatrix}$$

For d = 2^c

$$V_c = \begin{bmatrix} V_{c-1} & -(p^{2^{c-1}} V_{c-1})^* \\ p^{2^{c-1}} V_{c-1} & (V_{c-1})^* \end{bmatrix}$$

where the operator (•) works on any matrix with an even number of rows by multiplying the entries in the top half of the matrix by -1 and leaving those in the bottom half unchanged.*

Rotated Factorial with other n ($\neq p^k$) Points

Rotation Theorem for Mixed Level Design

d = 2

$$R=\begin{bmatrix} 1 & \sqrt{pq} \\ -\sqrt{pq} & 1 \end{bmatrix}$$

d = 4

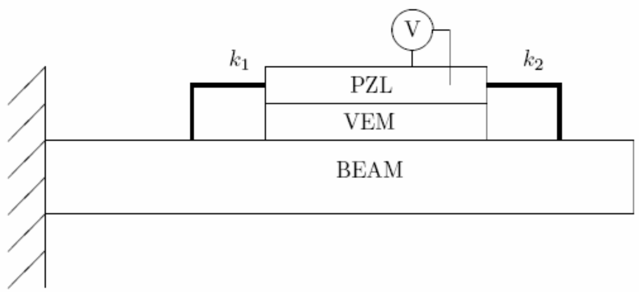
$$R=\begin{bmatrix}
\frac{1}{\sqrt{1+pq}\sqrt{1+pqrs}} & \frac{\sqrt{pq}}{\sqrt{1+pq}\sqrt{1+q^2r^2}} & \frac{\sqrt{pqrs}}{\sqrt{1+pq}\sqrt{1+pqrs}} & \frac{qr\sqrt{pq}}{\sqrt{1+pq}\sqrt{1+q^2r^2}} \\
\frac{-\sqrt{pq}}{\sqrt{1+pq}\sqrt{1+pqrs}} & \frac{1}{\sqrt{1+pq}\sqrt{1+q^2r^2}} & \frac{-\sqrt{pqrs}\sqrt{pq}}{\sqrt{1+pq}\sqrt{1+pqrs}} & \frac{qr}{\sqrt{1+pq}\sqrt{1+q^2r^2}} \\
\frac{-\sqrt{pqrs}}{\sqrt{1+rs}\sqrt{1+pqrs}} & \frac{-qr\sqrt{rs}}{\sqrt{1+rs}\sqrt{1+q^2r^2}} & \frac{1}{\sqrt{1+rs}\sqrt{1+pqrs}} & \frac{\sqrt{rs}}{\sqrt{1+rs}\sqrt{1+q^2r^2}} \\
\frac{\sqrt{pqrs}\sqrt{rs}}{\sqrt{1+rs}\sqrt{1+pqrs}} & \frac{-qr}{\sqrt{1+rs}\sqrt{1+q^2r^2}} & \frac{-\sqrt{rs}}{\sqrt{1+rs}\sqrt{1+pqrs}} & \frac{1}{\sqrt{1+rs}\sqrt{1+q^2r^2}}
\end{bmatrix}$$

d = 8

d = 2^c

Beattie & Lin (2004)

Beam Example



Some Comments

- Computer experiments are gaining in popularity
  - main research area of the next 10 years
- Rotated factorial designs
  - good factorial design properties
  - (orthogonality and structure)
  - good Latin hypercube properties
  - (unique and equally-spaced projections)
  - easy to construct
  - comparable by Bayesian criteria
  - very suitable for computer experiments
- Extensions
  - Type U and Type E designs
    - extension to sizes other than p^2
  - higher dimensional extension promising

Basic Idea-1

$$D = X \cdot V$$

desirable design ← D; factorial design ← X; rotated matrix ← V

factorial design:

$$\begin{bmatrix} 1 & 1 \\ 2 & 1 \\ \vdots & \vdots \\ p & 1 \\ 1 & 2 \\ 2 & 2 \\ \vdots & \vdots \\ p & 2 \\ \vdots & \vdots \\ 1 & p \\ 2 & p \\ \vdots & \vdots \\ p & p \end{bmatrix}_{p^2 \times d}$$

rotated matrix:

$$\begin{bmatrix} \upsilon_1 & \upsilon_3 \\ \upsilon_2 & \upsilon_4 \end{bmatrix}_{d \times d}$$

Beattie & Lin (1998): Rotating Full Factorials

Basic Idea-2

$$D = X \cdot V$$

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

rotated matrix:

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

Bursztyn & Steinberg (2002): Rotating in Groups

Basic Idea-3

Now,
Put these two ideas together!

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

Steinberg and Lin (2006)

Biometrika (2006), 93, 2, pp. 279–288

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

A Construction Method for Orthogonal Latin Hypercube Designs (with p-level)

Pang, Liu and Lin

(2007)

Uniform Designs

Fang, Lin, Winker & Yang
(*Technometrics*, 1999)
Fang and Lin
(Handbook of Statistics, Vol 22, 2003)

Uniform Design

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

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

Uniform Design

$\hat{F}_n(x)$ = **Empirical** Cumulative Distribution Function

$F(x)$ = Uniform Cumulative Distribution Function

Find $x = (x_1, x_2, ..., x_n)$

such that $\hat{F}_n(x)$ is closest to $F(x)$.

Discrepancy

$$D = \left[\int_{\Omega} \left\| \hat{F}_n(x) - F(x) \right\|^p dx \right]^{1/p}$$

- *Wang & Fang (1980)*

The centered L_p-discrepancy is invariant under exchanging coordinates from x to 1-x. Especially, the centered L_2-discrepancy, denoted by CL_2, has the following computation formula:

$$(CL_2(\mathrm{P}))^2$$

$$= \left(\frac{13}{12}\right)^s - \frac{2}{n}\sum_{k=1}^{n}\prod_{i=1}^{s}\left(1 + \frac{1}{2}|x_{ki} - \frac{1}{2}| - \frac{1}{2}|x_{ki} - \frac{1}{2}|^2\right)$$

$$+ \frac{1}{n^2}\sum_{k=1}^{n}\sum_{j=1}^{n}\prod_{i=1}^{s}\left[1 + \frac{1}{2}|x_{ki} - \frac{1}{2}| + \frac{1}{2}|x_{ji} - \frac{1}{2}| - \frac{1}{2}\left|x_{ki} - x_{ji}\right|\right].$$

Sampling Strategies for Computer Experiments: Design and Analysis

Simpson, T.W., Lin, D.K.J., and Chen, W. (2001)

Recent Research on

- Obtaining information which are not possible, without modern technology
  - Censor
  - RFID
  - Simulation
- How to (optimally) design these devices?
- How to analyze the outcomes (data)?

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!

This talk is based on

- Deng and Lin (2007) Patent on Random Number Generation.
- Steinberg and Lin (2006) “A Construction Method for Orthogonal Latin Hypercube Designs,” *Biometrika*, **93**, 279-288.
- Fang and Lin (2007) “Uniform Design in Computer and Physical Experiments,” The Grammar of Technology Development, ed. Shu Yamada, pp.99-119.

<http://www.personal.psu.edu/users/j/x/jxz203/lin/Lin_pub/>

STILL QUESTION?

Send $500 to

- Dennis Lin

  University Distinguished Professor

  483 Business Building

  Department of Supply Chain & Information Systems

  Penn State University

- +1 814 865-0377 (phone)
- +1 814 863-7076 (fax)
- DKL5@psu.edu

(Customer Satisfaction or your money back!)