

Forty Years of Statistical Design and Analysis of Simulation Experiments (DASE)

Jack P.C. Kleijnen Tilburg University & Wageningen University

"Titans of Simulation" seminar Winter Simulation Conference 2005 Orlando, December 4-7, 2005

Overview

- Source: 200 publications; see LPAA speech at INFORMS Simulation Society, 6 Dec. (5:15 PM)
- Discrete-event & deterministic simulations
- Classic designs (e.g., fractional factorials):
 - Few factors & few levels per factor
 - Low-order polynomial regression (meta)models
- *Modern* designs (e.g., space-filling: LHS):
 - Many factors & many levels per factor
 - Kriging (meta)models
- Screening designs (e.g., Sequential Bifurcation)
- Goals: Validation, sensitivity analysis, optimization, risk analysis

Black boxes & metamodels

- DOE treats simulation model as a black box Example: Single-server queue simulation $w = f(\lambda, \mu, r_0)$ with average simulated waiting time w, etc. Implicit function f, random via PRN seed r_0
- Metamodels with 'track record':
- Example: First-order polynomial in two 'factors' (λ , μ) $y = \beta_0 + \beta_1 \lambda + \beta_2 \mu + e$ with predictor y, and *noise e* caused by PRN & lack of fit
- First-order polynomials using transformations
 - Basic factor, traffic rate $y = \beta_0 + \beta_1 (\lambda / \mu) + e$
 - Log of factors (elasticity coefficients β) In y = $\beta_0 + \beta_1 \ln \lambda + \beta_2 \ln \mu + e$

Note: Remains *linear* regression analysis!

• Kriging metamodels; see (end of) next slide

Linear regression & DOE

Estimate metamodel parameters: Experiment with simulation (generate I/O data)

Example 1:
$$y = \beta_0 + \beta_1 (\lambda / \mu) + e = \beta_0 + \beta_1 x + e$$

- BLUE estimator: Two x-values only; as far apart as 'possible' Assumption: Valid metamodel Holds locally (Taylor series; see optimization through RSM)
- Example 2: $y = \beta_0 + \beta_1 x + \beta_2 x^2 + e$ Popular (but, not optimal) designs: 3^k design, CCD (5 values)

General: BLUE b if OLS used and IID noise e with zero mean

Global metamodel (0 < λ / μ < 1): Kriging metamodel (Kleijnen & Van Beers) 11/29/2005 WSC 2005, Titans of Simulation

Multiple factors: One-at-a-time design

This design is popular, but *inferior*:

- 1. Interactions not estimable (singular matrix of independent variables; see next slide)
- 2. Higher *variances* of (OLS) estimated main effects Minimum variance: Orthogonal design (see next slides)
- 3. Often three values (two is optimal; see preceding slide)

2^k design for k factors

Example: Three factors; - means –1; empty means +1

scenario	X 0	X ₁	x ₂	X 3	<i>x</i> ₁ <i>x</i> ₂	<i>x</i> ₁ <i>x</i> ₃	X ₂ X ₃	$\boldsymbol{X}_1 \boldsymbol{X}_2 \boldsymbol{X}_3$
1		-	-	-				-
2			-	-	-	-		
3		-		-	-		-	
4				-		-	-	-
5		-	-			-	-	
6			-		-		-	-
7		-			-	-		-
8 (= 2 ³)								

11/29/2005

WSC 2005, Titans of Simulation

2^k design continued

Properties:

- Balanced: each column has four plusses; four minuses (except x₀: only plusses)
- 2. Orthogonal: zero innerproduct for two different columns

2³ design assumes following metamodel:

$$y = \beta_0 + \beta_1 x_1 + \beta_2 x_2 + \beta_3 x_3 +$$
(main effects)
+ $\beta_{12} x_1 x_2 + \beta_{13} x_1 x_3 + \beta_{23} x_2 x_3 +$
+ $\beta_{123} x_1 x_2 x_3 + e$ (three-factor interaction)

q = 8 effects β ; n = 8 (= 2³) scenarios (factor combinations) Problem: No degrees of freedom left

Solutions:

- 1. Effects are zero (see later slide)
- 2. Scenarios are replicated m > 1 times (N = nm runs)

11/29/2005

WSC 2005, Titans of Simulation

Regresion analysis of 2^k design

OLS estimator: Orthogonal X implies 'simpler' b = (X' w) / n Balanced X implies subtract two specific averages

Covariance matrix of b: Simple iff white noise Practical problems:

- Variance heterogeneity (e.g., queuing simulation)
- Common Random Numbers (CRN): Correlated outputs Solutions (Kleijnen '92, '98):
- a. OLS *point* estimate & more complicated *cov(b)*
- b. GLS with cov(w) estimated from m (> n) replicates
 t test for significance???

Significance vs. importance: Scale factors between –1 and +1 Factor with highest absolute value is most important

2^{k-p} design & other incomplete designs Assumptions:

- *High*-order interactions (e.g., β_{123}) are hard to *interpret*
- *High*-order interactions are *unimportant* ($\beta_{123} = 0$)

Consequence: *q* decreases, so *n* may decrease too

Example: 2³⁻¹ design: Delete rows with – sign in column 1.2.3

scenario	x ₀	X ₁	X ₂	X 3	X ₁ X ₂	X ₁ X ₃	X ₂ X ₃	X ₁ X ₂ X ₃
2			-	-	-	-		
3		-		-	-		-	
5		-	-			-	-	
8								

11/29/2005

Bias in 2^{k-p} designs

2³⁻¹ design: *Identical* 0 and 1.2.3 columns (namely +1) Consequence: $E(b_0) = \beta_0 + \beta_{123}$ (*bias*, unless $\beta_{123} = 0$), etc. Application area: *Optimization* through *RSM*

First-order polynomial (Taylor series); see Angün et al.

Example 2: A 27-	⁴ design	(balanced	&	orthogonal)
------------------	---------------------	-----------	---	-------------

Scen.	1	2	3	4 (= 1.2)	5 (= 1.3)	6 (= 2.3)	7(=1.2.3)
1	-	-	-				-
2		-	-	-	-		
3	-		-	-		-	
4			-		-	-	-
5	-	-			-	-	
6		-		-		-	-
7	-			-	-		-
8 (= 2 ⁷⁻⁴)							

11/29/2005

Designs of resolution 3, 4, and 5

Resolution-3 (R-3) designs: Assumes *first*-order polynomial Subclass 1: 2^{k-p} design with p high enough (e.g., 2^{7-4}) Subclass 2: k = 11 and n = 12 (not power of 2) Tabulated by Plackett-Burman ('46), up to n = 100

Resolution-4 designs: Unbiased by two-factor interactions Construction: Add *mirror* design to R-3 design (multiply -1) Consequence: Doubles the number of scenarios simulated

Resolution-5: Unbiased estimators of all two-factor interaction: Subclass 1: 2^{k-p} design with p low enough (e.g. 2^{5-1} , 2^{8-2} , 2^{11-4}) Subclass 2: Saturated (n = q) Rechtschaffner ('67) designs

Second-order polynomial metamodel

Central Composite Designs (CCD):

- Resolution-5 design for all effects except *pure quadratic* effects
- One-factor-at-a-time ('star') design (add 2k scenarios)
- Base scenario



- Five values/factor; non-orthogonal
- CCD for *k* < 121; *n* < 33,011: Sanchez & Sanchez (2005)
- Saturated designs: See Kleijnen ('87)

Kriging

Track record in *deterministic* simulation Examples: CAE for airplanes, t.v. monitors; see Sacks et al. ('89), Simpson et al. ('01)

New: *Random* simulation; see Kleijnen & Van Beers ('04) **Basics: Interpolation method; see Cressie ('93) (geostatistics)** Weighted linear combination of old outputs, to predict new Weights depend on *distance* between old and new inputs Assumption: *Closer* inputs generate *more* correlated outputs Method: Correlogram (variogram) Assumption: Stationary covariance process Notes: *Exact* interpolator in *deterministic* simulation Weights vary with location of new input (unlike regression b) *Estimated* optimal weights: *Non*-linear predictor Solution: Bootstrap; see Den Hertog et al. ('06)

Designs for Kriging

Popular: Latin Hypercube Sampling (LHS); see @Risk, etc. Origin: McKay et al. ('79) for risk analysis of deterministic simulation (No Kriging analysis!)

Flexible, simple, and space filling (non-extreme values) designs; example: k = 2, n = 4



Future research

- Robust optimization of simulation models Optimization: Glover's OptQuest, Angün et al.'s RSM, etc. Robust: Environment changes unexpectedly Approach: Taguchi ('87) concepts & modern DASE
- 2. Sequential designs
 Screening: Sequential Bifurcation (low-order polynomial)
 Kriging: Kleijnen & Van Beers ('04), Jin et al. ('02)
- 3. *Replication* versus *new* scenarios
- 4. *Multiple* outputs (standard in simulation practice)
- 5. *Kriging* in *random* (discrete-event) simulation: Software
- 6. Comparisons of different metamodels & designs: Tilburg test bed
- References: Kleijnen (2005), Invited review, *EJOR,* 287-300 Kleijnen (2006?): *DASE*, Kluwer

Appendix 1: Background

Classic *discrete-event simulation* textbooks: Major attention: *Tactical* issues (steady state, VRT) Minor attention: *Classic* Design Of Experiments (DOE)

This talk:

Strategic issues:

Which scenarios to simulate?

How to *analyze* the resulting simulation Input/Output (I/O) data?

Origin of DOE:

Agriculture ('30s), engineering ('50s):

Few factors & few levels

DOE in simulation:

Kleijnen ('75, '87, '98), Kleijnen et al. (2005)

Appendix 2: Screening

- Practical simulation models: Hundreds of factors Example: Ericsson supply chain with 92 factors Solution? Plackett-Burman design: 96 scenarios
- Problem: Too much computer time?
- Solution? Focus on a few factors, selected *intuitively*
- Solution? *Supersaturated* designs: *n* < *k*
- Subclass: Aggregate factors into groups
- **Problem: Cancellation of effects**
- Solution: Define -1 and +1 such that *each* effect is *positive* Practice: Users do know *direction* of (nearly) each effect!
- Group-screening design types: Saltelli et al. ('04) Kleijnen et al. ('04): 11 of the 92 factors identified as important, after only 21 simulated scenarios