Comparing Systems via Simulation (Part 5)

Dr.Çağatay ÜNDEĞER

Öğretim Görevlisi
Bilkent Üniversitesi Bilgisayar Mühendisliği Bölümü

e-mail : cagatay@undeger.com
cagatay@cs.bilkent.edu.tr

Comparing Systems via Simulation (Outline)

- Introduction
- Comparison Problems
  - Comparing Two Systems
  - Screening Problems
  - Selecting the Best
  - Comparison with a Fixed Performance

Introduction

- Simulation experiments are usually performed to compare some alternative solutions/designs.
- Method that is appropriate depends on the type of comparison (problem) and output data.

Introduction (Statistical Methods)

- Statistical methods are applicable to computer simulations
- Since the important assumptions of statistics can usually be satisfied approximately.

Introduction (Assumptions)

- Normally distributed data can be secured by batching large number of outputs.
- Independence can be obtained by controlling random-number assignments.
- Multiple-stage sampling is feasible because a subsequent stage can be initialized simply;
  - By retaining the final random number seed from the preceding stage or
  - By regenerating the entire sample.

Comparison Problems

- In this lesson, we will examine 4 classes of problems.
- Objective is to select the correct answer to the problem with a high probability.

Comparison Problems

- Comparing Two Systems:
  - To compute the difference in expected performance of two systems.
- Screening problems:
  - To compare substantial number of designs in order to eliminate clearly inferior (not successful) performers.
- Selecting the best:
  - To find the system with the largest or smallest performance measure.
- Comparsion with a fixed performance:
  - To find the best system, provided that its performance exceeds a known, fixed performance standard.

Comparison Problems (Background)

- Compare *k* different system (design points) via simulation.
- Let *Y* be a random variable that represents the output.
- Let Y_{ij} represents the j^{th} simulation output (replication or batch) of i^{th} system.
- Let n_i be the number of replications/batches from i^{th} system.
- Simulations of design points are either independent or using common random numbers.

Comparison Problems (Background)

- Let $\overline{Y}_i$ *be the sample mean of i^{th}* system.

$$\overline{Y}_i = \frac{1}{n_i} \sum_{j=1}^{n_i} Y_{ij}$$

- Let S^2_i *be the sample variance of i^{th}* system.

$$S^2_i = \frac{1}{n_i - 1} \sum_{j=1}^{n_i} (Y_{ij} - \overline{Y}_i)^2$$

- Let S^2 *be the pooled sample variance of all systems*.

$$S^2 = \frac{1}{k} \sum_{i=1}^{k} S^2_i$$

CS-503 9

Comparing Two Systems

- The goal is to compute the difference in expected performance of two systems
  - In order to determine whether one is better or they are practically equivalent.

CS-503 10

Comparing Two Systems

- If systems 1 & 2 are simulated with *n* replications/batches using independent random number streams then
  - The difference between system 1 and 2 with (1-α) confidence interval is:

$$\mu_1 - \mu_1 \text{ range limits} = (\bar{Y}_1 - \bar{Y}_2) \pm t^* \sqrt{\frac{S_1^2 + S_2^2}{n}}$$

$t^* = t_{2n-2,1-\alpha/2} = 1-\alpha/2$ probability value for t-distribution with n-1 degrees of freedom

Comparing Two Systems

- With (1-α) confidence, If the range limits are;
  - Both positive then (e.g. [3,9])
    - Performance metric of system 1 is greater than system 2,
  - Both negative then (e.g. [-9,-2])
    - Performance metric of system 1 is smaller than system 2,
  - In different sides of zero then (e.g. [-1,3])
    - Systems are equivalent.

Screening Problems

- The goal is to compare substantial number of system designs in order to;
  - Group those with similar performance, and
  - Eliminate clearly inferior performers for examining high performers in more details.
- For instance,
  - 20 potential system designs for a company is produced.
  - Response time is the performance measure.
  - You would like to reduce the number of potential designs using a plot study before a more detailed study is performed.

Screening Problems (Techniques)

- Multiple Comparison Approach
- Subset Selection Approach

Screening Problems (Multiple Comparison Approach)

- Approaches the screening problem by forming simultaneous confidence intervals on parameters $\mu_i - \mu_l$ for all $i \neq l$.
- $k(k-1)/2$ confidence intervals will be formed.
- Indicate magnitude & direction of the difference between each pair of alternatives.

Screening Problems (Multiple Comparison Approach)

- Simulate systems;
  - With independent random number streams,
- Compute;
  - Sample means ($\bar{Y}_i$), and
  - Pooled sample variance (S^2).

Screening Problems (Multiple Comparison Approach)

- Simultaneous confidence intervals of $\mu_i - \mu_l$ for all $i \neq l$ (Tukey's procedure) :

$$(\bar{Y}_i - \bar{Y}_l) \pm \frac{Q^{(\alpha)}_{k,v}}{\sqrt{2}} S \sqrt{\frac{1}{n_i} + \frac{1}{n_l}}$$

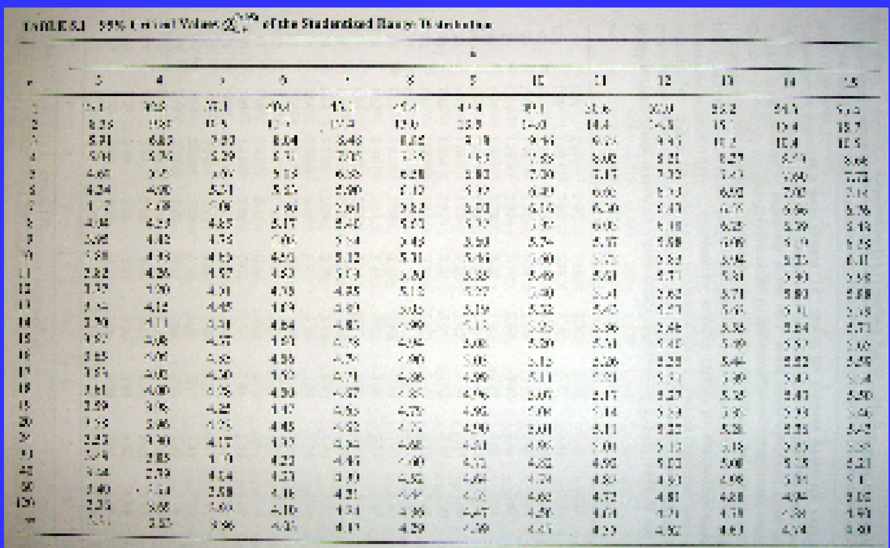
$1\text{-}\alpha$ quantile of Studentized range distribution with parameter k and v (degrees of freedom)

$$v = \sum_{i=1}^{k} (n_i - 1)$$

difference between sample means

pooled standard deviation

Studentized Range Distribution Table



Screening Problems (Multiple Comparison Approach)

- Suppose that:
  - k = 4 system architectures.
  - n = 6 replications are obtained for each.

$\bar{Y}_1 = 72$

$\bar{Y}_2 = 85$

$\bar{Y}_3 = 76$

$\bar{Y}_4 = 62$

$S^2 = 100.9$

$$\bar{Y}_i = \frac{1}{n_i}\sum_{j=1}^{n_i} Y_{ij}$$

$$S^2_i = \frac{1}{n_i-1}\sum_{j=1}^{n_i} (Y_{ij} - \bar{Y}_i)^2$$

$$S^2 = \frac{1}{k}\sum_{i=1}^{k} S^2_i$$

Response time is performance metric (smaller is better)

Screening Problems (Multiple Comparison Approach)

- Suppose that:
  - Objective is to eliminate system architectures with low performance (high response time) with 0.95 confidence.
  - So compare pairs:
    - 1 and 2
    - 1 and 3
    - 1 and 4
    - 2 and 3
    - 2 and 4
    - 3 and 4

Screening Problems (Multiple Comparison Approach)

- Suppose that:
  - We compare system 2 and 4 :

$$v = \sum_{i=1}^{4} (n_i - 1) = (6-1) + (6-1) + (6-1) + (6-1) = 20$$

$$(\bar{Y}_i - \bar{Y}_l) \pm \frac{Q^{(0.05)}_{4,20}}{\sqrt{2}} \, S\sqrt{\frac{1}{n_i} + \frac{1}{n_l}}$$

$$(85 - 62) \pm \frac{3.96}{\sqrt{2}} \, 10.04\sqrt{\frac{1}{6} + \frac{1}{6}} = 23 \pm 16 = [\,7{,}39\,]$$

With 95% probability, system 2 is different from system 4 in range 7 to 39. Since range of system 2 is greater than 4, and range does not contain 0, system 2 can be eliminated (because shorter response time is preferrable).

Screening Problems (Subset Selection Approach)

- Approaches the screening problem by producing a subset of designs that contains the best system with a probability *1-α*.
- Applicable in cases when data from competing designs are;
  - Independent (different random numbers),
  - Balanced ($n_1=n_2=...=n_k=n$), and
  - Normally distributed with a common variance.

Screening Problems (Subset Selection Approach)

- Simulate systems;
  - With independent random number streams,
  - With equal number of replications/batches.
- Compute;
  - Sample means ($\overline{Y}_i$), and
  - Pooled sample variance (S^2).

CS-503 23

Screening Problems (Subset Selection Approach)

- Include i^th design in the subset if;

$$\overline{Y}_i \le \min_{1 \le j \le k} (\overline{Y}_j) + g\, S \sqrt{\frac{2}{n}}$$ If smaller performance metric is better

$$\overline{Y}_i \ge \max_{1 \le j \le k} (\overline{Y}_j) - g\, S \sqrt{\frac{2}{n}}$$ If greater performance metric is better

$g = T^{(\alpha)}_{k-1,k(n-1)}$ is a critical value from a multivariate t-distribution.

CS-503 24

Multivariate t-Distribution Table

TABLE 8.2 95% Critical Values $t_{p,\nu}^{(1/2)}$ of the Multivariate t-Distribution with Common Correlation 1/2

ν	p = 1	2	3	4	5	6	7	8	9
1	6.31	8.51	11.38	13.10	14.27	15.23	16.04	16.75	[illegible]
2	2.92	3.80	4.24	4.51	5.00	5.24	5.41	5.50	5.75
3	2.35	2.94	3.28	3.52	3.70	3.82	3.93	[illegible]	4.12
4	2.13	2.61	2.88	3.08	3.22	3.34	3.44	3.52	3.59
5	2.02	2.44	2.68	2.85	2.98	3.08	3.16	3.24	3.30
6	1.94	2.34	2.56	2.71	2.83	2.92	3.00	3.06	3.12
7	1.89	2.27	2.48	2.62	2.73	2.82	2.89	2.95	3.00
8	1.86	2.22	2.42	2.55	2.66	2.74	2.81	2.87	2.92
9	1.83	2.18	2.37	2.50	2.60	2.68	2.75	2.81	2.86
10	1.81	2.15	2.34	2.47	2.56	2.64	2.70	2.76	2.81
11	1.80	2.13	2.31	2.43	2.53	2.60	2.67	2.72	2.77
12	1.78	2.11	2.29	2.41	2.50	2.58	2.64	2.69	2.74
13	1.77	2.09	2.27	2.39	2.48	2.55	2.61	2.66	2.71
14	1.76	2.08	2.25	2.37	2.46	2.53	2.59	2.64	2.69
15	1.75	2.07	2.24	2.36	2.44	2.51	2.57	2.62	2.67
16	1.75	2.06	2.23	2.34	2.43	2.50	2.56	2.61	2.65
17	1.74	2.05	2.22	2.33	2.42	2.49	2.54	2.59	2.64
18	1.73	2.04	2.21	2.32	2.41	2.48	2.53	2.58	2.62
19	1.73	2.03	2.20	2.31	2.40	2.47	2.52	2.57	2.61
20	1.72	2.03	2.19	2.30	2.39	2.46	2.51	2.56	2.60
25	1.71	2.00	2.17	2.27	2.36	2.42	2.48	2.52	2.56
30	1.70	1.99	2.15	2.25	2.34	2.40	2.45	2.50	2.54
35	1.69	1.98	2.14	2.24	2.32	2.38	2.44	2.48	2.52
40	1.68	1.97	2.13	2.23	2.31	2.37	2.42	2.47	2.51
45	1.68	1.96	2.12	2.22	2.30	2.36	2.41	2.46	2.50
50	1.68	1.96	2.11	2.22	2.29	2.36	2.41	2.45	2.49
55	1.67	1.95	2.11	2.21	2.29	2.35	2.40	2.44	2.48
60	1.67	1.95	2.10	2.21	2.28	2.35	2.40	2.44	2.48
120	1.66	1.93	2.08	2.18	2.26	2.32	2.37	2.41	2.45
∞	1.65	1.92	2.06	2.16	2.23	2.29	2.34	2.38	2.42

Screening Problems (Subset Selection Approach)

- Suppose that:
  - k = 4 system architectures.
  - n = 6 replications are obtained for each.

$\bar{Y}_1 = 72$

$\bar{Y}_2 = 85$

$\bar{Y}_3 = 76$

$\bar{Y}_4 = 62$

$S^2 = 100.9$

$$\bar{Y}_i = \frac{1}{n}\sum_{j=1}^{n} Y_{ij}$$

$$S^2_i = \frac{1}{n-1}\sum_{j=1}^{n} (Y_{ij} - \bar{Y}_i)^2$$

$$S^2 = \frac{1}{k}\sum_{i=1}^{k} S^2_i$$

Response time is performance metric (smaller is better)

Screening Problems (Subset Selection Approach)

- Suppose that:
  - Objective is to produce a subset including the best system with 0.95 confidence.
  - Smaller response time is preferred so we use:

$$\bar{Y}_i \le \min_{1 \le j \le k} (\bar{Y}_j) + g\, S \sqrt{\frac{2}{n}}$$

$$\bar{Y}_i \le \min(72,85,76,62) + T^{(\alpha)}_{k-1,\, k(n-1)}\ 10.04 \sqrt{\frac{2}{6}}$$

$$\bar{Y}_i \le 62 + T^{(0.05)}_{3,20}\ 5.8 = 62 + 2.19\ 5.8 = 74.7$$

  - Select system 1 and 4 ($72 \le 74.7$ and $62 \le 74.7$)

Selecting the Best

- The goal is to find the system with the largest or smallest performance measure.
- For instance,
  - 20 potential system designs for a company is produced.
  - Number of jobs processed per hour is the performance measure.
  - Differences of less than about 5 jobs are considered practically equivalent.
  - You reduced the number of alternatives to 4 with a plot study.
  - Now you would like to determine the best one among 4 with a more detailed study.

Selecting the Best (Techniques)

- Multiple Comparison Approach
  - Procedure Rinott + MCB
  - Procedure NM + MCB
  - Procedure Bonferroni + MCB
- Multinomial Selection Approach
  - Procedure BEM
  - Procedure BG

Selecting the Best (Multiple Comparison Approach)

- In stochastic simulations, correct selection can never be guaranteed with certainty.
- A solution offered by Multiple Comparison integrated with Indifference-zone selection is;
  - To guarantee to select the best system with high probability,
    - Whenever it is at least a user-specified amount better than others.
- This practically significant difference is called Indifference-zone (δ) (e.g. δ = 5 jobs).

Selecting the Best (Multiple Comparison Approach)

- If some system happens to be within δ of the best,
  - Then these are considered to be practically equivalent,
  - And probability of selecting one of the good systems is at least 1-α,
  - So any of them can be chosen considering other important metrics (e.g. cost).

Selecting the Best (Multiple Comparison Approach)

- Uses Indifference-zone.
- Approaches the screening problem by forming simultaneous confidence intervals on parameters;
  - $\mu_i - max_{l \neq i} \mu_l$ (greater is better) or
  - $\mu_i - min_{l \neq i} \mu_l$ (smaller is better) for all i.
- Bound the difference between expected performance of each system and the best of the others, with probability *1-α.*

Selecting the Best (Multiple Comparison Approach)

- Performs multiple comparisons with the best (MCB).
- Combines indifference-zone selection and MCB.
- Provides information about how close each of the inferior systems is to the best,
  - Which is useful if secondary criteria such as cost, ease of installation are not reflected to the performance measure.
  - e.g. δ = 5 jobs difference is equivalent.

CS-503 33

Selecting the Best
(Proc. Rinott+MCB (Independent Sampling))

- Takes observations in two stages:
  - First stage uses $n_0 \geq 2$ (10 recomended) independent observations from each system
    - To estimate marginal variance.
  - Second stage uses marginal variance
    - To compute additional number of observations required to meet the indifference-zone probability.

CS-503 34

Selecting the Best

(Proc. Rinott+MCB (Independent Sampling))

- Specify;
  - δ (indifference-zone)
  - α (confidence interval probability)
  - n_0 (first-stage sample size).
- Take n_0 independent replications/batches from each of the k systems.

CS-503 35

Selecting the Best

(Proc. Rinott+MCB (Independent Sampling))

- Compute marginal sample variance for all i:

$$S^2_i = \frac{1}{n_0-1} \sum_{j=1}^{n_0} (Y_{ij} - \overline{Y}_i)^2$$

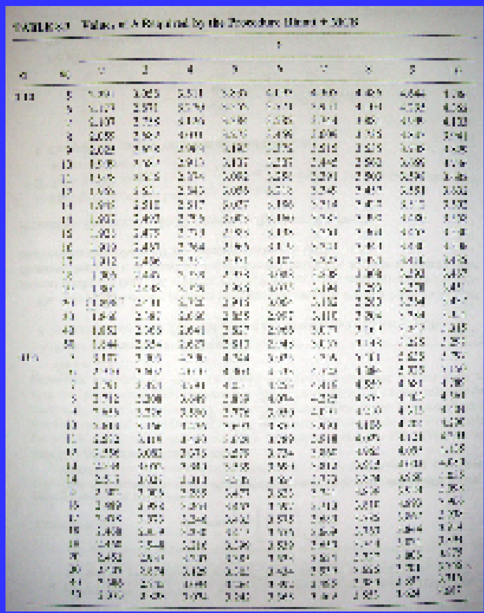
- Compute final sample size for each system:

$$N_i = \max \left(n_0, \left\lceil \left(\frac{h_{\alpha,n_0} S_i}{\delta} \right)^2 \right\rceil \right)$$

Find h_{α,n_0} from Procedure Rinott+MCB table.

CS-503 36

Procedure Rinott+MCB Table



CS-503 37

Selecting the Best

(Proc. Rinott+MCB (Independent Sampling))

- For each system i,
  - Take $N_i - n_0$ additional (or restart all) observations independently of the first-stage.
  - Compute overall sample mean:

$$\bar{\bar{Y}}_i = \frac{1}{N_i} \sum_{j=1}^{N_i} Y_{ij}$$

- Select system with;
  - Largest $\bar{\bar{Y}}_i$ (greater is better) or
  - Smallest $\bar{\bar{Y}}_i$ (smaller is better) as the best.

CS-503 38

Selecting the Best
(Proc. Rinott+MCB (Independent Sampling))

- Simultanously form MCB confidence interval for each system i:
  - If greater is better:

    $\min(0, Y_i - \max_{i \neq l} Y_l - \delta), \max(0, Y_i - \max_{i \neq l} Y_l + \delta)$

  - If smaller is better:

    $\min(0, Y_i - \min_{i \neq l} Y_l - \delta), \max(0, Y_i - \min_{i \neq l} Y_l + \delta)$

- If range of a system *i* is not on one side of the zero, it is an equivalent of best (within δ).

CS-503 39

Selecting the Best
(Proc. NM+MCB (Common Random Numbers))

- In Rinott + MCB,
  - Systems are simulated independently.
- However, under fairly conditions, assigning common random numbers (CRN) to simulation of each system decreases variances of estimates.
- NM+MCB uses CRN.

CS-503 40

Selecting the Best

(Proc. NM+MCB (Common Random Numbers))

- Similarly use δ, α and n_0.
- Take n_0 replications/batches from each of the k systems using CRN across systems.

Selecting the Best

(Proc. NM+MCB (Common Random Numbers))

- Compute approximate sample variance:

$$S^2 = \frac{2 \sum_{i=1}^{k} \sum_{j=1}^{n_0} (Y_{ij} - \overline{Y}_{i\cdot} - \overline{Y}_{\cdot j} + \overline{Y}_{\cdot\cdot})^2}{(k-1)(n_0-1)}$$

"." means average of all i

".." means average of all i and j

- Compute final sample size for all systems:

$$N = \max \left(n_0, \left\lceil \left(\frac{g\,S}{\delta} \right)^2 \right\rceil \right)$$

Find $g = T^{(\alpha)}_{k-1,(k-1)(n_0-1)}$ from Multivariate t-distribution table.

Selecting the Best
(Proc. NM+MCB (Common Random Numbers))

- For each system i,
  - Take $N - n_0$ additional (or restart all) observations using CRN across systems.
  - Compute overall sample mean:

$$\bar{\bar{Y}}_i = \frac{1}{N} \sum_{j=1}^{N} Y_{ij}$$

- Select system with;
  - Largest $\bar{\bar{Y}}_i$ (greater is better) or
  - Smallest $\bar{\bar{Y}}_i$ (smaller is better) as the best.

Selecting the Best
(Proc. NM+MCB (Common Random Numbers))

- Simultaneously from MCB confidence intervals as in Rinott + MCB.

Selecting the Best
(Proc. Bonferroni+MCB (CRN))

- NM + MCB works under a complex set of conditions.
- Conferroni + MCB works under more general conditions.
- But, tends to require more observations, especially when k is large.

Selecting the Best
(Proc. Bonferroni+MCB (CRN))

- Similarly use δ, α and n_0.
- Take n_0 replications/batches from each of the k systems using CRN across systems.

Selecting the Best
(Proc. Bonferroni+MCB (CRN))

- Compute sample variances of differences for all i≠l :

$$S^2_{il} = \frac{1}{n_0-1} \sum_{j=1}^{n_0} [\,(Y_{ij} - Y_{lj}) - (\bar{Y}_i - \bar{Y}_l)\,]^2$$

- Compute final sample size for all systems:

$$N = \max\left(n_0, \left\lceil \max_{l \neq i} \left(\frac{t\, S_{il}}{\delta} \right)^2 \right\rceil \right)$$

Find $t = t_{n_0-1,\,1-\alpha/(k-1)}$, from t-distribution table.

Selecting the Best
(Proc. Bonferroni+MCB (CRN))

- For each system i,
  - Take $N - n_0$ additional (or restart all) observations using CRN across systems.
  - Compute overall sample mean:

$$\bar{\bar{Y}}_i = \frac{1}{N} \sum_{j=1}^{N} Y_{ij}$$

- Select system with
  - Largest $\bar{\bar{Y}}_i$ (greater is better) or
  - Smallest $\bar{\bar{Y}}_i$ (smaller is better) as the best.

Selecting the Best
(Proc. Bonferroni+MCB (CRN))

- Simultaneously from MCB confidence intervals as in Rinott + MCB.

Selecting the Best (Techniques)

- Multiple Comparison Approach
  - Procedure Rinott + MCB
  - Procedure NM + MCB
  - Procedure Bonferroni + MCB
- Multinomial Selection Approach
  - Procedure BEM
  - Procedure BG

Selecting the Best (Multinomial Selection Approach)

- Solution offered by Multinomial Selection is;
  - To select the best system with probability 1-α
    - Whenever the ratio of selecting the best to the second best p_i is greater than a user-specified constant.
- This practically significant smallest ratio worth detecting is called Indifference-constant (θ>1) (e.g. θ = 1.2).

$$\theta = \min \frac{p_{best}}{p_{second\ best}} \text{ required}$$

CS-503 51

Selecting the Best (Procedure BEM)

- Specify;
  - *θ* (indifference-constant)
  - *α* (confidence interval probability)
- Take a random sample of *n* independent multinomial observations from each of the *k* systems,
  - Where *n* is found from Multinomial procedure table using *α*, *θ* and *k*.

CS-503 52

Multinomial Procedure Table

TABLE 8.4 Sample Size n for Multinomial Procedure BEM, and Truncation Numbers n_T for Procedure BG

		$k = 2$		$k = 3$		$k = 4$		$k = 5$	
α	θ	n	n_T	n	n_T	n	n_T	n	n_T
0.25	3.0	1	1	5	5	8	9	11	12
	2.0	5	5	12	13	20	24	29	34
	1.8	5	7	17	18	29	35	41	50
	1.6	9	9	26	32	46	57	68	86
	1.4	17	19	52	71	92	124	137	184
	1.2	55	67	181	285	326	495	486	730
0.10	3.0	7	10	11	12	16	19	21	24
	2.0	15	15	29	34	43	53	58	71
	1.8	19	27	40	50	61	75	83	104
	1.6	31	41	64	83	98	126	134	172
	1.4	59	79	126	170	196	274	271	374
	1.2	199	267	437	670	692	1050	964	1460
0.05	3.0	9	11	17	20	23	26	29	34
	2.0	23	27	42	52	61	74	81	98
	1.8	33	35	59	71	87	106	115	142
	1.6	49	59	94	123	139	180	165	240
	1.4	97	151	186	266	278	380	374	510
	1.2	327	455	645	960	979	1500	1331	2000

Selecting the Best (Procedure BEM)

- A random sample is taken from each of the *k* systems for each replication.
- So we have a matrix;

$$\begin{bmatrix} Y_{11}, Y_{21}, Y_{31}, \dots, Y_{k1} \\ Y_{12}, Y_{22}, Y_{32}, \dots, Y_{k2} \\ Y_{12}, Y_{22}, Y_{32}, \dots, Y_{k2} \\ \dots \\ Y_{1n}, Y_{2n}, Y_{3n}, \dots, Y_{kn} \end{bmatrix}$$

replications (↓)

system designs (→)

Selecting the Best (Procedure BEM)

- From *Y*, determine multinomial observations *X*.

$$\begin{bmatrix} Y_{11}, Y_{21}, Y_{31}, \ldots, Y_{k1} \\ Y_{11}, Y_{21}, Y_{31}, \ldots, Y_{k1} \\ Y_{12}, Y_{22}, Y_{32}, \ldots, Y_{k2} \\ \ldots. \\ Y_{1n}, Y_{2n}, Y_{3n}, \ldots, Y_{kn} \end{bmatrix} \longrightarrow X = \begin{bmatrix} X_{11}, X_{21}, X_{31}, \ldots, X_{k1} \\ X_{11}, X_{21}, X_{31}, \ldots, X_{k1} \\ X_{12}, X_{22}, X_{32}, \ldots, X_{k2} \\ \ldots. \\ X_{1n}, X_{2n}, X_{3n}, \ldots, X_{kn} \end{bmatrix} \downarrow \text{replications}$$

$$X_{ij} = \begin{cases} 1 & \text{if } Y_{ij} > \max_{l \neq i} Y_{lj} \\ 0 & \text{otherwise} \end{cases}$$

On j^{th} replication, if system *i* is best, set $X_{ij} = 1$ else $X_{ij} = 0$

Thus there will be only a single 1 on each row.

Selecting the Best (Procedure BEM)

- From *X*, determine W.

$$\begin{bmatrix} X_{11}, X_{21}, X_{31}, \ldots, X_{k1} \\ X_{11}, X_{21}, X_{31}, \ldots, X_{k1} \\ X_{12}, X_{22}, X_{32}, \ldots, X_{k2} \\ \ldots. \\ X_{1n}, X_{2n}, X_{3n}, \ldots, X_{kn} \end{bmatrix} \longrightarrow W = \begin{bmatrix} W_1, W_2, W_3, \ldots, W_k \end{bmatrix}$$

sum up column i

$$W_i = \sum_{j=1}^{n} X_{ij}$$

The number of times system i was the best

- Select the design having largest W_i as the best (one with highest probability).
- If there are equal W_is, pick any of them

Selecting the Best (Procedure BG)

- Procedure BEM, uses a fixed number of replications to select the best.
- This may sometimes be inefficient.
- Procedure BG;
  - Uses a more efficient but complex procedure, and
  - Stops when one design is sufficiently ahead of the others.

Selecting the Best (Procedure BG)

- Specify;
  - θ (indifference-constant)
  - α (confidence interval probability)
- To select the best system among *k* systems.
- Uses incremental number of observations.
- Upper limit of observations, the truncation number (n_T), is found from Multinomial procedure table using α, θ and *k*.

Selecting the Best (Procedure BG)

- The method advances stage by stage.
- A stage contains one observation from each system totally making *k* observations.

$$\begin{bmatrix} Y_{11}, Y_{21}, Y_{31}, \dots, Y_{k1} \\ Y_{11}, Y_{21}, Y_{31}, \dots, Y_{k1} \\ Y_{12}, Y_{22}, Y_{32}, \dots, Y_{k2} \\ \dots \\ Y_{1..}, Y_{2..}, Y_{3..}, \dots, Y_{k..} \end{bmatrix}$$

← a stage

advance stage by stage
until stoping criteria is satisfied.

Selecting the Best (Procedure BG)

- At the m^{th} stage of observations,
  – We determine X_m from Y:

$$\begin{bmatrix} Y_{1m}, Y_{2m}, Y_{3m}, \dots, Y_{km} \end{bmatrix} \longrightarrow X_m = \begin{bmatrix} X_{1m}, X_{2m}, X_{3m}, \dots, X_{km} \end{bmatrix}$$

$$X_{ij} = \begin{cases} 1 & \text{if } Y_{ij} > \max_{l \neq i} Y_{lj} \\ 0 & \text{otherwise} \end{cases}$$

Selecting the Best (Procedure BG)

- Then compute W_m from X:

$$\begin{bmatrix} X_{11}, X_{21}, X_{31}, \ldots, X_{k1} \\ X_{11}, X_{21}, X_{31}, \ldots, X_{k1} \\ X_{12}, X_{22}, X_{32}, \ldots, X_{k2} \\ \ldots. \\ X_{1m}, X_{2m}, X_{3m}, \ldots, X_{km} \end{bmatrix} \longrightarrow W_m = \begin{bmatrix} W_{1m}, W_{2m}, W_{3m}, \ldots, W_{km} \end{bmatrix}$$

sum up column i

$$W_{im} = \sum_{j=1}^{n} X_{ij}$$

- And sort W_m ascending:

$$W_{[1]m} \le W_{[2]m} \le W_{[3]m} \le \ldots \le W_{[k]m}$$

[i] is the index of i^{th} smallest W_{jm}

Selecting the Best (Procedure BG)

- Compute Z_m:

$$Z_m = \sum_{i=1}^{k-1} (1/\theta)^{W_{[k]}-W_{[i]m}}$$

- Stop sampling when any of the following occur:
  $Z_m \le \alpha/(1-\alpha)$
  $m = n_T$
  $(W_{[k]m} - W_{[k-1]m}) \ge (n_T - m)$

- Take system i with the largest W_i as the best.

Comparison With a Fixed Performance

- To find the best system, provided that its performance exceeds a known, fixed performance standard (μ_0).
- For instance,
  - There are 4 potential risky investment strategies for a company.
  - But it is also possible to get a known, fixed return if the money is deposited in a bank.
  - Therefore, we would like to chose none of the strategies unless its expected return is larger than the fixed return.

Comparison With a Fixed Performance (Procedure BT)

- Takes observations in two stages:
  - First stage uses $n_0 \geq 2$ (10 recomended) independent observations from each system
    - To estimate marginal variance.
  - Second stage uses marginal variance
    - To compute number of observations required to meet the probability requirement.

Comparison With a Fixed Performance (Procedure BT)

- Specify;
  - μ_0 *(fixed performance standard)*
  - δ (indifference-zone)
  - α (confidence interval probability)
  - n_0 (first-stage sample size).
- Take n_0 independent replications/batches from each of the k systems.

Comparison With a Fixed Performance (Procedure BT)

- Compute estimated sample variance:

$$S^2 = \frac{\sum_{i=1}^{k} \sum_{j=1}^{n_0} (Y_{ij} - \overline{Y}_i)^2}{k(n_0-1)}$$

- Compute final sample size for all systems:

$$N = \max \left(n_0, \left\lceil \left(\frac{g\,S}{\delta} \right)^2 \right\rceil \right)$$

Find *g*, *h* from Procedure BT table.

Procedure BT Table

TABLE 8.5 Values of g (Top Entries) and h (Bottom Entries) for Procedure BT for Comparing Normal Treatments with a Standard with 1 [illegible] $\nu = n_0 - 1$ [illegible]

n_0	$1-\alpha$	$k=2$		$k=3$		$k=4$		$k=5$	
		0.90	0.95	0.90	0.95	0.90	0.95	0.90	0.95
2		4.775	7.176	4.416	6.011	4.269	5.382	4.197	5.242
		2.743	4.075	2.648	3.551	2.620	3.340	2.615	3.234
3		3.679	4.920	3.688	4.701	3.714	4.619	3.743	4.584
		2.072	2.722	2.172	2.696	2.243	2.704	2.300	2.721
4		3.401	4.406	3.487	4.357	3.554	4.356	3.608	4.369
		1.963	2.417	2.042	2.479	2.136	2.531	2.207	2.576
5		3.276	4.181	3.394	4.201	3.478	4.239	3.544	4.267
		1.829	2.285	1.982	2.381	2.085	2.451	2.162	2.507
6		3.204	4.055	3.340	4.112	3.434	4.163	3.506	4.208
		1.787	2.211	1.947	2.325	2.055	2.405	2.136	2.468
8		3.126	3.920	3.281	4.014	3.384	4.085	3.463	4.142
		1.740	2.132	1.909	2.263	2.022	2.354	2.107	2.424
10		3.084	3.848	3.248	3.951	3.358	4.043	3.440	4.107
		1.715	2.090	1.888	2.230	2.004	2.327	2.091	2.400
12		3.058	3.804	3.227	3.923	3.341	4.016	3.426	4.084
		1.699	2.064	1.875	2.210	1.993	2.309	2.081	2.385
∞		2.945	3.615	3.139	3.786	3.266	3.900	3.361	3.985
		1.632	1.955	1.818	2.121	1.943	2.234	2.037	2.319

Comparison With a Fixed Performance (Procedure BT)

- For each system i,
  - Take $N - n_0$ additional (or restart all) independent observations.
  - Compute overall sample mean:

$$\bar{\bar{Y}}_i = \frac{1}{N} \sum_{j=1}^{N} Y_{ij}$$

Comparison With a Fixed Performance (Procedure BT)

- If greatest performance measure is better:

  If $_{max}\bar{\bar{Y}}_i > \mu_0 + h\delta/g$ then

  Select the strategy i as best,

  Otherwise retain the standard as best.

- If smallest performance measure is better:

  If $_{min}\bar{\bar{Y}}_i < \mu_0 - h\delta/g$ then

  Select the strategy i as best,

  Otherwise retain the standard as best.