Blocking Multiple Sources of Error in Small Analytic Studies

Ramón V. León and Robert W. Mee

Department of Statistics

University of Tennessee

Knoxville Tennessee

2001 Quality and Productivity Research Conference

Austin, Texas

May 25, 2001

Abstract

Blocking is most commonly used to reduce error variation. But in small analytic studies blocking is sometimes used instead to broaden the scope of inference. In this note we show an example of a Graeco-Latin square design where blocking is being used for this second reason and where a traditional analysis that assumes the first reason for blocking leads to misleading conclusions. The ideas discussed apply in general to the use of blocking in small analytic studies

References

- R. V. León and R. W. Mee (June 2000), Blocking Multiple Sources of Error in Small Analytic Studies *Quality Engineering*. **12**, No. 4, June 2000
- Montgomery, D. C. (1997), *Design and Analysis of Experiments, Fourth Edition*, New York: John Wiley.
- Hahn, G. J. and Meeker, W. Q. (1993), “Assumptions for Statistical Inference,” *The American Statistician*, **47**, 1-11.
- Deming, W. E. (1975), “On Probability as a Basis for Action,” *The American Statistician* **29**, 146-152
- Moen, R. D., Nolan, T. W., and Provost, L. P. (1991) *Improving Quality Through Planned Experimentation*, New York: McGraw-Hill.

Why Block?

- To improve the precision of comparisons, that is, to reduce error variation
- To make the inferences as broad as possible

Blocking Adage:

- Minimize the differences within blocks
- Maximize the differences between blocks

Large Differences Between Blocks

"For an analytic study, the investigation is made as broad as possible so as to reduce the almost inevitable gap between the sampled process and the process of interest."

Hahn and Meeker (1993)

<u>Large differences between blocks provides:</u>

- Breadth
- Opportunity to investigate the consistency of any treatment effect across the varied conditions represented by the blocks

Model Used When Blocking is Used to Increase Precision of Comparisons

- Model with additive block effect
- Block-by-treatment interactions are assumed not to be present
- Analysis includes blocks only as main effects

Model Used When Blocking is Used to Broaden the Scope of Inference

Block-by-treatment interactions are of primary interest, since they indicate inconsistency of the treatment effect over the variety of conditions included in the experiment

With such a profound difference in analysis, it is prudent not to confuse the two purposes of blocking

Example
(Montgomery 1997, Problem 5-23)

- An engineer is interested in comparing whether four different methods of assembling color television components produce systematic differences in assembly time in a production line.

- Operators and workplace are considered as possible sources of variation. It is also possible that there is an order effect when an operator uses the assembly methods sequentially

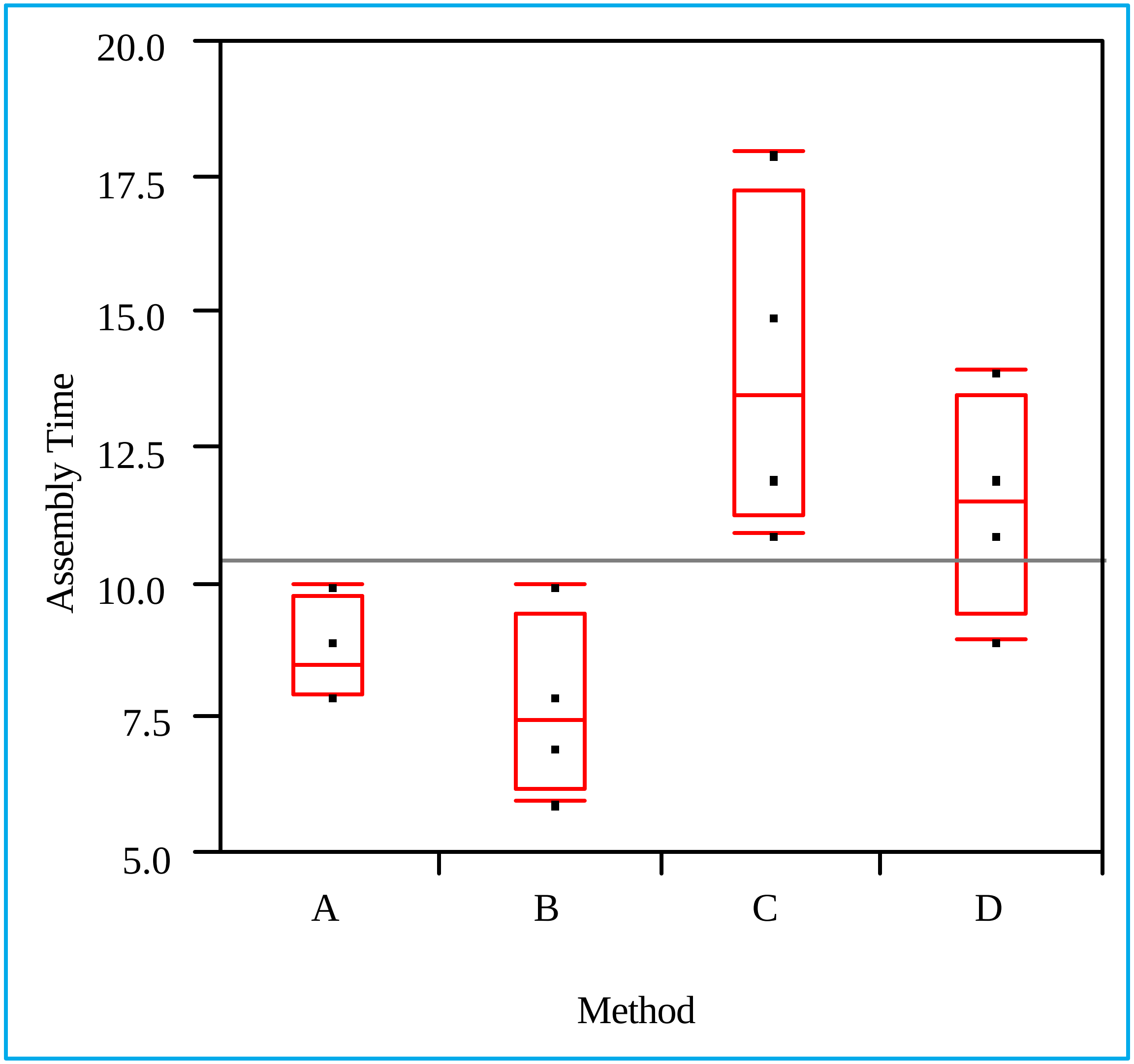
Graeco-Latin Square Design

The engineer decides to conduct a quick exploratory study to see if one can rule out some of the methods for production

Order of Assembly	**Operator**			
	1	**2**	**3**	**4**
1	Cβ=11	Bγ=10	Dδ=14	Aα=8
2	Bα=8	Cδ=12	Aγ=10	Dβ=12
3	Aδ=9	Dα=11	Bβ=7	Cγ=15
4	Dγ=9	Aβ=8	Cα=18	Bδ=6

A, B, C, D are assembly methods and α, β, γ, δ are workplaces

Graphical Analysis



Is the engineer justified in ruling out methods C and D for not being as quick as methods A and B?

A Reasoning About the Box Plots

- Each assembly method is used once by each of four operators, once in each of four workplaces, and once in each of four orders of assembly
- This systematic balance in the data ensures *fairer* comparisons than if the four observations per box plot had been chosen at random from among the operators, workplaces, and orders of assembly
- Since each box plot incorporates a wider range of the sources of variation than would probably be included under random sampling the experiment result is more likely to be generalizable to production conditions

Traditional Greaco-Latin Square Analysis

Source	DF	Sum of Squares	Mean Square	F Ratio	Prob>F
Operator	3	19.0	6.3333	0.6909	0.6157
Order	3	0.5	0.1667	0.0182	0.9960
Workplace	3	7.5	2.5000	0.2727	0.8429
Method	3	95.5	31.8333	3.4727	0.1669
Error	3	27.5	9.1667		
Total	15	150.0			

- Do we conclude that there is no difference between assembly methods?
- Is this analysis "correct" in the sense that it corresponds to the engineer's rationale for the experiment?

Goal of Traditional Anova Analysis for Graeco-Latin Square

To detect a difference in the assembly methods in the presence of *only the variation within* operators, workplaces, and orders.
Thus, this test is of limited relevance given that the engineer's goal is to see if there is a difference of *practical* significance in the assembly methods in production where operators, workplaces, and other circumstances would be all changing.

Traditional Anova Assumptions

- Any difference in method are consistent from one operator to the next, and are also the same for any combination or order and workplace, that is, **it assumes no any treatment-by-block interaction**

- We suspect that any engineer who planned such a study with the intention of generalizing to production conditions would find the assumptions of no treatment-by-block interactions unreasonable

- The question of generalizability is answered by considering method*block interactions, not by ignoring them!

A Possible Way Out of the Paradox?

- Assume only one source of extraneous variation – say, operator (i.e., that we have a randomized block design rather than a Graeco-Latin square)
- Treat possible operator effect and method*operator interaction as random
- Test for method would be

  $$MS_{method}/MS_{method*operator} = (95.5/3)/(35.5/9) = 8.07$$

  which is above the 99th percentile of the $F_{3,9}$ distribution
- <u>Conclusion:</u> Differences exist in method means for the pool of workers from where we sampled
- Computation ignores two sources of variation – order and workplace. Yet their possible presence does not weaken the conclusion that differences exists between the method means since an order or workplace effect would have inflated the denominator above

Remarks

- The analysis above is not motivated by "a pool into error insignificant sums of squares" argument
- A similar argument can be made using workplace rather than operator
- A similar justification cannot be used with order since orders 1 to 4 cannot be viewed as a random sample or even a representative sample from some conceptual population. Nevertheless, the fact that the method*order mean square is small compared to the method mean square is suggestive.

Remarks

- The method mean square is significantly larger than each of the three method*block mean squares obtained by ignoring two blocking sources at the time
- Thus the lost of power of the tests for method resulting from the presence of the ignored sources of variation does not affect our conclusion in this case of significant difference among methods across operators, workplaces, and order of assembly
- If significant treatment differences had not been found then the lost of power would have been more problematic

Some Quotes

"All analytic studies are conducted under judgment samples"
Deming (1975)

"… it is rare in analytic studies that a random selection of conditions is preferable to a judgment selection."
Moen, Nolan, and Provost (1991)

Final Comments

- Need to understand that analytic studies are best done with judgment sample
  - Better to select operators and workplaces as different from each others as possible rather than at random

How I Would do the Analysis

- Treat the data as a completely randomized design with method as the treatment:

  $F=(95.5/3)/(54.5/12)=7.01$ with 3,12 d.f.

- This has a p-value of 0.00559
- This analysis would be valid if there were no operator, workplace, or order effect
- If these effects are present they would inflate the F denominator.
- Since we are rejecting this denominator inflation does not reduce our confidence that there is a method effect that shows up above the operator, workplace, and order variation