SAND99-0064J

RECEIVED
APR 07 1999

Statistical Analyses of Scatterplots to Identify Important Factors in Large-Scale Simulations, 2: Robustness of Techniques

Jack P.C. Kleijnen, Department of Information Systems and Auditing, Tilburg University, 5000 LE Tilburg, The Netherlands
Jon C. Helton, Department of Mathematics, Arizona State University, Tempe, AZ 85287-1804, USA

Abstract

The robustness of procedures for identifying patterns in scatterplots generated in Monte Carlo sensitivity analyses is investigated. These procedures are based on attempts to detect increasingly complex patterns in the scatterplots under consideration and involve the identification of (i) linear relationships with correlation coefficients, (ii) monotonic relationships with rank correlation coefficients, (iii) trends in central tendency as defined by means, medians and the Kruskal-Wallis statistic, (iv) trends in variability as defined by variances and interquartile ranges, and (v) deviations from randomness as defined by the chi-square statistic. The following two topics related to the robustness of these procedures are considered for a sequence of example analyses with a large model for two-phase fluid flow: the presence of Type I and Type II errors, and the stability of results obtained with independent Latin hypercube samples. Observations from analysis include: (i) Type I errors are unavoidable, (ii) Type II errors can occur when inappropriate analysis procedures are used, (iii) physical explanations should always be sought for why statistical procedures identify variables as being important, and (iv) the identification of important variables tends to be stable for independent Latin hypercube samples.

Key Words: Chi-square, correlation coefficient, epistemic uncertainty, interquartile range, Kruskal-Wallis, Latin hypercube sampling, mean, median, Monte Carlo, partial correlation coefficient, rank transform, scatterplot, sensitivity analysis, standardized regression coefficient, subjective uncertainty, variance.

Please send page proof to:

Jon C. Helton
Department 6848, MS 0779
Sandia National Laboratories
Albuquerque, NM 87185-0779, USA
Phone: 505-284-4808
Fax: 505-844-2348

DISCLAIMER

DISCLAIMER

1. Introduction

Procedures for identifying patterns in scatterplots generated in Monte Carlo sensitivity analyses are described and illustrated in the preceding article.[1] These procedures are based on attempts to recognize increasingly complex patterns in the scatterplots under consideration and involve the identification of (i) linear relationships with correlation coefficients, (ii) monotonic relationships with rank correlation coefficients, (iii) trends in central tendency as defined by means, medians and the Kruskal-Wallis statistic, (iv) trends in variability as defined by variances and interquartile ranges, and (v) deviations from randomness as defined by the chi-square statistic. The robustness of these procedures is now considered. In particular, the presence of Type I and II errors is considered (Sects. 2, 3), and the stability of results obtained with independent Latin hypercube samples (LHSs)[2] is examined (Sect. 4).

2. Type I and II Errors

The sensitivity analysis techniques under discussion use p-values to indicate if a relationship appears to exist between an uncertain analysis input and a predicted analysis outcome (Sect. 8, Ref. 1). Clearly, it is desirable that the techniques identify the inputs that actually affect analysis outcomes (i.e., to avoid Type II errors, which correspond to the failure to identify important variables). As shown by the example analyses in Sect. 10 of Ref. 1, Type II errors can occur when the test for variable importance is inappropriate for the relationships that exist between analysis inputs and analysis outcomes (e.g., see the analyses for *E2:WAS_PRES* in Sect 10.4, Ref. 1). Thus, a good analysis strategy is to use several tests for variable importance and thus reduce the likelihood of overlooking an important variable (i.e., committing a Type II error).

In addition, it is also important that the techniques not identify inputs as having effects that are not actually present (i.e., to avoid Type I errors, which correspond to the indication of nonexistent effects for unimportant variables). Unfortunately, the "price" of using multiple tests for variable importance is an increase in the number of Type I errors (i.e., in "false alarms"); however, it is the responsibility of the subject-area experts to explain why individual variables are identified as being important. Ultimately, if such explanations cannot be developed, then the analysis is suspect and the observed results may be due to errors in the implementation of the analysis.

If a variable has no effect on a particular analysis outcome and the assumptions of the statistical test in use are satisfied, then the corresponding p-values generated from repeated random sampling should have a uniform distribution on the interval (0, 1). Specifically, $prob\ (\hat{p} < p) = prob\ (\hat{t} > t_p) = p$, and thus $\hat{p}$ has a uniform distribution on (0, 1), where $0 \le p \le 1$, *prob* denotes probability, and t_p and $\hat{t}$ are values of the statistic with p-values of p and $\hat{p}$, respectively. Similarly, if multiple unimportant variables are involved, their p-values from a single sampling should be uniformly distributed on (0, 1). Thus, for a specified p-value (i.e., p) and n unimportant variables, the likelihood $prob\ (\text{Type I} \mid p, n)$ of committing a Type I error (actually, one or more Type I errors) is given by

$$prob\ (\text{Type I} \mid p, n) = 1 - (1 - p)^n, \tag{1}$$

with $prob$ (Type I | p, n) increasing as each of p and n increases (Fig. 1). Thus, Type I errors cannot be avoided, and their likelihood of occurrence is defined by Eq. (1) provided that the p-values for unimportant variables follow a uniform distribution.

The LHSs indicated in Eqs. (8)-(10) of Ref. 1, and on which the examples in Sect. 10 of Ref. 1 are based, involved 75 variables (Table 2, Ref. 1). However, 49 of these variables were not used in the calculation of the model results *E0:WAS_PRES* and *E0:BRAALIC*; and 48 of these variables were not used in the calculation of the model results *E2:WAS_SATB* and *E2:WAS_PRES* (Table 1, Ref. 1). Thus, the p-values associated with these 49/48 variables should have uniform distributions on the interval (0, 1). The Kolmogorov-Smirnov test[3] can be used to indicate if the distributions of p-values for these variables do indeed have uniform distributions on (0, 1). In particular, the 0.9 and 0.99 two-sided Kolmogorov-Smirnov bounds around the cumulative distribution function (CDF) for the true distribution (i.e., uniform on (0, 1)) are given by $1.22/(n+\sqrt{n/10})^{1/2}$ and $1.63/(n+\sqrt{n/10})^{1/2}$, respectively, where n is the sample size (Table A14, Ref. 3). For $n = 48, 49$, the corresponding 0.9 and 0.99 bounds are 0.17 and 0.23, respectively.

As 4 variables (i.e., *E0:WAS_PRES, E0:BRAALIC, E2:WAS_SATB, E2:WAS_PRES)* and 8 tests (i.e., CC, RCC, CMN, CL, CMD, CV, CIQ, SI) are under consideration (see Sect. 10, Ref. 1), 32 distributions of p-values result (Fig. 2). The p-values that give rise to these 32 distributions were calculated with the analytic rather than the Monte Carlo procedures described in Ref. 1. Of these 32 distributions, 24 are within the 0.9 bounds. Further, 6 of the 9 distributions that are outside the bounds are for the variable/test pairs (*E0:WAS_PRES*, CC), (*E0:BRAALIC*, CC), (*E2:WAS_SATB*, CC), (*E0:WAS_PRES*, RCC), (*E0:BRAALIC*, RCC), and (*E2:WAS_SATB*, RCC). As results obtained with CCs and RCCs are not independent, the indicated deviations of (*E0:WAS_PRES*, CC) (*E0:BRAALIC*, CC) and (*E2:WAS_SATB*, CC) from a uniform distribution on (0, 1) are not independent of the indicated deviations for (*E0:WAS_PRES*, RCC), (*E0:BRAALIC*, RCC), and (*E2:WAS_SATB*, RCC). The most notable deviation occurs for the pair (*E0:BRAALIC*, CV), with no p-values exceeding 0.7. There is something associated with *E0:BRAALIC* that is causing an underrepresentation of large p-values for unimportant variables. This underrepresentation probably derives from the fact that *E0:BRAALIC* has a few large values and many very small values (Fig. 2b, Ref. 1). Fortunately, the shape of the individual CDFs in Fig. 2 does not suggest any tendency for the tests to produce unusual numbers of very small p-values; thus, there does not appear to be a tendency to produce excessive numbers of Type I errors in the examples under consideration. However, the results in Fig. 2 do suggest that the p-values for unimportant variables may not have a uniform distribution on (0, 1). Because of this behavior, additional simulations were carried out as described in the next section.

3. Type I and Type II Errors: Additional Simulations

An additional set of simulations was carried out to provide a check on the reasonableness of the distributions of p-values in Fig. 2. In particular, 10 independent LHSs of size 300 were generated with the Iman and Conover restricted pairing technique[4] from 50 independent variables with uniform distributions on the interval [0, 1]. These LHSs were then associated with the calculated values for *E0:WAS_PRES*, *E0:BRAALIC*, *E2:WAS_SATB* and *E2:WAS_PRES* obtained with the original LHS of size 300 discussed in Sect. 2 of Ref. 1, and the corresponding distributions of p-values were calculated for the preceding four output variables, each of the eight tests under consideration, and each of the 10 independent LHSs. Again, the p-values were calculated with the analytic procedures described in Ref. 1. The outcome is 10 CDFs for each of the 32 test/output variable pairs.

If the assumptions of the tests are met and the calculations are implemented correctly, then the CDFs for each test/dependent variable pair should approximate a uniform distribution on [0, 1]. This generally appears to be the case. For example, the original CDFs for *E0:WAS_PRES* and tests based on CCs and RCCs move across the 0.99 Kolmogorov-Smirnov boundary (Figs. 2a, b). In contrast, the current exercise with 10 independently-generated LHSs produces CDFs of p-values that generally stay within the 0.9 Kolomogorov-Smirnov bounds (Fig. 3).

Twenty-nine of the remaining 30 test/output variable pairs produced distributions of p-value CDFs that were similar to the two CDF distributions in Fig. 3. The exception to this consistency occurred for *E0:BRAALIC* and the CVs test (Fig. 4). For this test/output variable pair, the p-values remain below approximately 0.7, which was also the case in Fig. 2f. The variable *E0:BRAALIC* has a large number of values that are effectively zero (Figs. 2, 4, Ref. 1). As a result, the estimated variances t_{ql} in Eq. (50) of Ref. 1 used to define the F statistic for the CVs test do not have a normal distribution for the individual independent variables, and so the associated p-values do not have a uniform distribution on [0, 1] even though the independent variables have no effect on *E0:BRAALIC*.

4. Robustness with Respect to Repeated Independent Samples

The examples in Sect. 10 of Ref. 1 use a sample of size 300 obtained by pooling the three samples of size 100 each indicated in Eqs. (8)-(10) of Ref. 1. The availability of these three independent samples provides a way to examine the robustness of the techniques under consideration. In particular, the analyses in Sect. 10 of Ref. 1 with each of the 8 techniques can be repeated with the individual samples of size 100. The extent to which the individual samples agree in the identification of important variables then provides an indication of how robust the techniques are with respect to repeated independent samples and also reductions in sample size (Table 1).

When comparing the variable selections in Table 1, it is important to keep in mind that the likelihood of a Type I error increases rapidly as p-values increase (Fig. 1), with 25 variables and a p-value of 0.01 producing a probability of 0.22 of a Type I error as indicated in Eq. (1). Further, the p-values for unimportant variables may not be random on (0, 1) due to patterns that are imposed on the data by the effects of other variables (Fig. 4). Thus. the probabilities in Fig. 1 are, at best, only an indication of the likelihood of a Type I error. As a result, the comparison of sets of important variables obtained with different replicates is probably valid only for variables with fairly low p-values. As p-values increase (e.g., > 0.01), such comparisons become less and less meaningful.

The overall pattern that emerges from the results in Table 1 is that the most important variables identified with the pooled sample of size 300 are also identified as being important with the three individual samples of size 100. In particular, the two most important variables as defined by the size of their p-values are typically the same across all four samples for the individual tests (i.e., CCs, RCCs, CMs, CLs, CMDs, CVs, CIQ, SI), although it should be recognized that the results obtained with the pooled sample are not independent of the results obtained with the individual samples. Hence, the use of a sample size of 300 or 100 made little difference with respect to the variables identified as being most important, although the larger sample size did tend to indicate likely effects for more variables than was the case for the smaller sample size. Similar robustness has been observed in several other studies involving Latin hypercube sampling.[5-7]

The most notable deviations from this consistency occur for the CVs test for *E0:BRAALIC* and *E2:WAS_PRES* and the CIQ test for *E0:BRAALIC*. The variable *E0:BRAALIC* is significantly affected by both *WMICDFLG* and *ANHPRM* (Fig. 4, Ref. 1). However, as *WMICDFLG* is being missed by the CVs test, it is perhaps not surprising that the individual samples are not producing consistent results. A logarithmic transformation improved the results obtained with the CVs test for *WMICDFLG* with the pooled sample (Table 16, Ref. 1) and also produced somewhat better results for the individual samples (Table 2). The variable *E2:WAS_PRES* is almost completely dominated by *BHPRM* (Fig. 6, Ref. 1), with this effect being missed by the CVs test for replicate R3; further, although *BHPRM* is identified by the CVs test as the most important variables affecting *E2:WAS_PRES* for replicate R2, the p-value is high (i.e., 0.0633). The CIQ test misses the effect of *ANHPRM* on *E0:BRAALIC* for replicates R1 and R2, with this behavior probably resulting from the large number of zero and near-zero values associated with *E0:BRAALIC* (Fig. 4, Ref. 1). The CVs and CIQ tests attempt to detect important variables on the basis of variable spread rather than variable location as is the case for the CMNs, CLs and CMDs tests. For the output variables under consideration, the tests based on location appear to be more effective in identifying important variables than tests based on spread.

An important point that emerges from the individual replicates is that consistency across independent analyses does not necessarily imply that these analyses are properly identifying the dominant variables. For example, all four analyses with both CCs and RCCs identify *HALPRM* and *ANHPRM* as being the two most important variables with respect to *E2:WAS_PRES* (Table 1) and completely fail to identify the dominant role played by *BHPRM* (Fig. 6, Ref. 1). For *E2:WAS_PRES*, the three replicates are producing similar patterns, which in turn are producing similar outcomes when analyzed with CCs and RCCs.

5. Discussion

Two aspects of statistical analyses of scatterplots to identify important factors in large-scale simulations have been examined: the occurrence of Type I and Type II errors, and the stability of results obtained with independent LHSs.

The occurrence of Type I errors is unavoidable in sampling-based sensitivity analyses (Fig. 1), with the likelihood of such errors increasing as the number of independent variables under consideration increases and also as more tests are applied to a given dependent variable. Although the possibility of Type I errors exists, this is not viewed as a serious problem for two reasons. First, the really important variables typically display a sufficiently strong effect that there is little likelihood that this effect could have originated from chance alone. Second, a variable should never be assumed to be important solely on the basis of a statistical test. Rather, an explanation for its indicated importance should be developed on the basis of the properties of the model under consideration. If such an explanation cannot be developed, then the effect may be spurious or, as occurs with disconcerting frequency, there may be an error in the implementation of the model.

The occurrence of Type II errors is a real possibility when statistical tests are used that are inappropriate for the patterns that occur in the analysis results under consideration. In a large analysis, there may be hundreds of dependent variables that are investigated in sensitivity analyses in a rote manner (i.e., the same test or tests are applied to each dependent variable rather than a unique sequence of tests being developed for each dependent variable). A good analysis strategy is to apply a sequence of tests to each dependent variable. Then, there is a high likelihood that at least one of these tests will be appropriate for a given dependent variable and correctly identify the factors affecting this variable. A possible sequence of tests is correlation coefficients (CCs), rank correlation coefficients (RCCs), common locations (CLs) or common medians (CMDs), and statistical independence (SI) (Sect. 11, Ref. 1).

Sample size is often an important consideration in sensitivity analyses for long-running models. In particular, the computational cost of evaluating the model may be a significant limitation on the number of model evaluations that can be carried out, with Latin hypercube sampling having been developed to make efficient use of a limited number of model evaluations.[2] Given the need to limit sample size, the stability of results obtained with independent, relatively small samples is a concern. In the empirical investigations reported here, individual LHSs of size 100 typically identified the same dominant variables as identified with a sample of size 300 obtained by pooling the three individual samples. Thus, relatively-small samples led to the identification of the important variables provided an appropriate statistical test was used. An inappropriate test will fail regardless of sample size. However, success at identifying less important variables rather unsurprising goes up as the sample size increases. The preceding suggests that a small sample size will lead to an identification of the most important variables, with an increased sample size leading to greater resolution of the effects associated with less important variables. The authors' experience is that the uncertainty in individual model predictions tends to be dominated by a small number of variables even though the model itself may have a large number of uncertain inputs.

Acknowledgment

Work was performed in part while both authors were visitors at Sandia National Laboratories (SNL), which is a multiprogram laboratory operated by Sandia Corporation, a Lockheed Martin Company, for the United States Department of Energy under contract DE-AC04-94AL85000. At SNL, programming support was provided by M.J. Shortencarier, and review was provided by M. Chavez, C. Crawford, and J.W. Garner. Editorial support was provided by L. Harrison, T. Allen, and H. Olmstead of Tech Reps, Inc.

6. References

1. Kleijnen, J.P.C., and J.C. Helton. 1998. "Statistical Analyses of Scatterplots to Identify Important Factors in Large-Scale Simulations, 1: Review and Comparison of Techniques." *Reliability Engineering and System Safety* (to appear).

2. McKay, M.D., R.J. Beckman, and W.J. Conover. 1979. "A Comparison of Three Methods for Selecting Values of Input Variables in the Analysis of Output from a Computer Code." *Technometrics* **21:** 239-245.

3. Conover, W.J. 1980. *Practical Non-parametric Statistics: Second Edition*. New York: John Wiley & Sons, Inc.

4. Iman, R.L., and W.J. Conover. 1982. "A Distribution-Free Approach to Inducing Rank Correlation Among Input Variables." *Communications in Statistics: Part B—Simulation and Computation* **B11:** 311-334.

5. Iman, R.L., and J.C. Helton. 1988. "An Investigation of Uncertainty and Sensitivity Analysis Techniques for Computer Models." *Risk Analysis* **8:** 71-90.

6. Iman, R.L., and J. C. Helton. 1991. "The Repeatability of Uncertainty and Sensitivity Analyses for Complex Probabilistic Risk Assessments." *Risk Analysis* **11:** 591-606.

7. Helton, J.C., J.D. Johnson, M.D. McKay, A.W. Shiver, and J.L. Sprung. 1995. "Robustness of an Uncertainty and Sensitivity Analysis of Early Exposure Results with the MACCS Reactor Accident Consequence Model." *Reliability Engineering and System Safety* **48:** 129-148.

Figure Captions

Fig. 1. Contour plots for probability of a Type I error, *prob* (Type I | p, n), as a function of p-value, p, and number of unimportant variables, n (See Eq. (1)).

Fig. 2. Distribution of p-values and associated Kolmogorov-Smirnov bounds for individual tests and variables in LHS that do not affect *E0:WAS_PRES*, *E0:BRAALIC*, *E2:WAS_SATB* and *E2:WAS_PRES*.

Fig. 3. Distributions of p-values for 10 independently-generated LHSs: (3a) CCs for *E0:WAS_PRES* and (3b) RCCs for *E0:WAS_PRES*.

Fig. 4. Distribution of p-values for 10 independently-generated LHSs for CVs test and *E0:BRAALIC*.

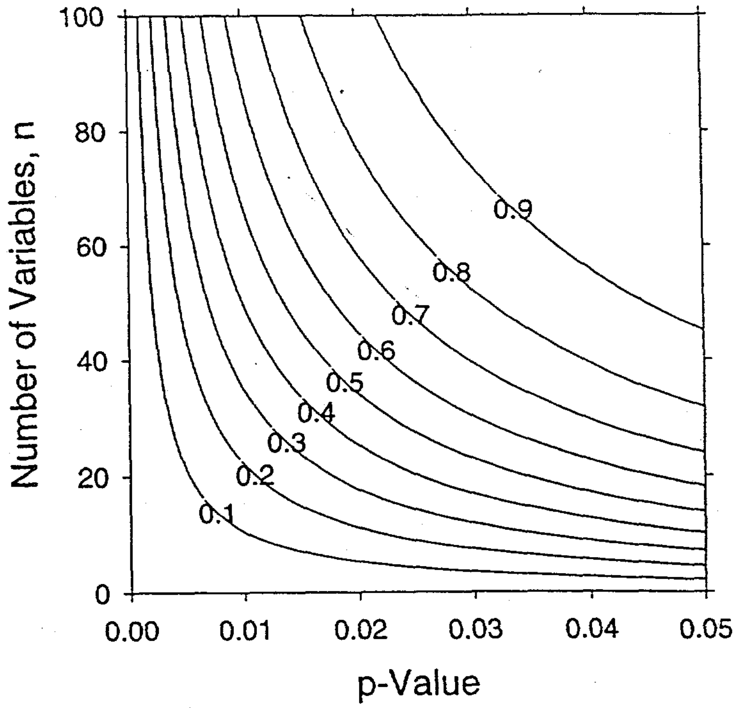

Fig. 1. Contour plots for probability of a Type I error, *prob* (Type I | p, n), as a function of p-value, p, and number of unimportant variables, n (See Eq. (1)).

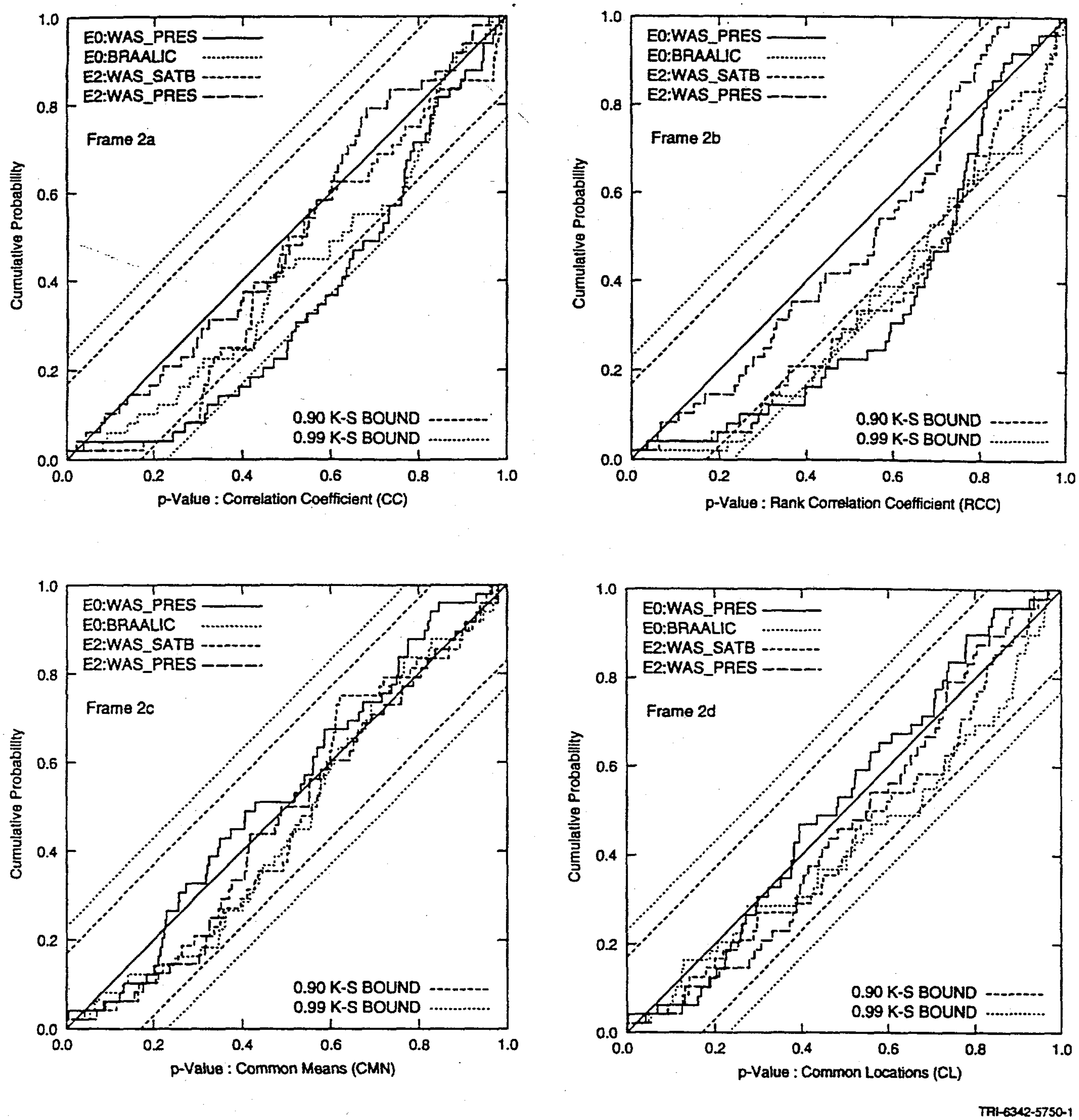

Fig. 2. Distribution of p-values and associated Kolmogorov-Smirnov bounds for individual tests and variables in LHS that do not affect *E0:WAS_PRES*, *E0:BRAALIC*, *E2:WAS_SATB* and *E2:WAS_PRES.*

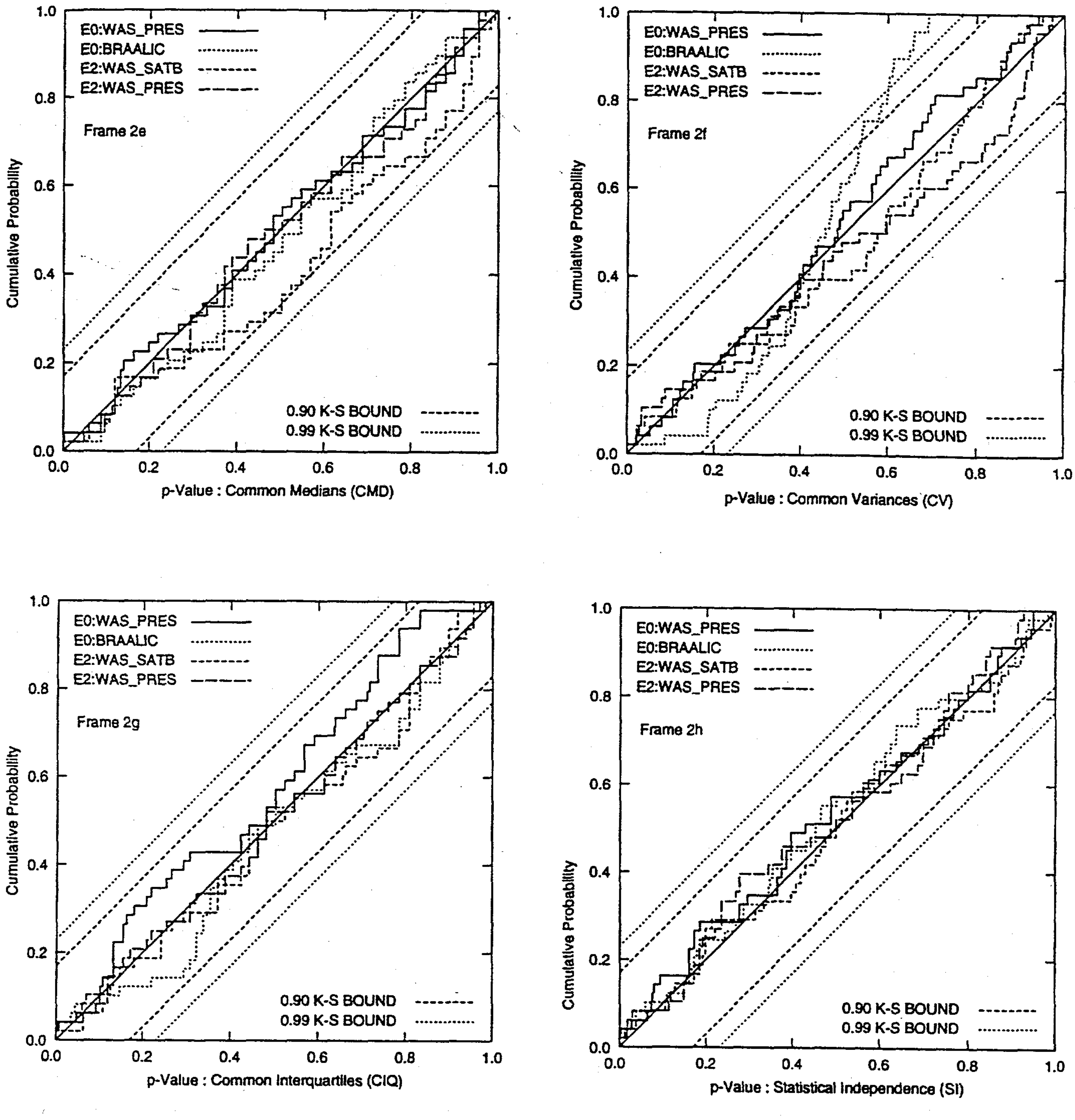

Fig. 2. Distribution of *p*-values and associated Kolmogorov-Smirnov bounds for individual tests and variables in LHS that do not affect *E0:WAS_PRES*, *E0:BRAALIC*, *E2:WAS_SATB* and *E2:WAS_PRES* (continued).

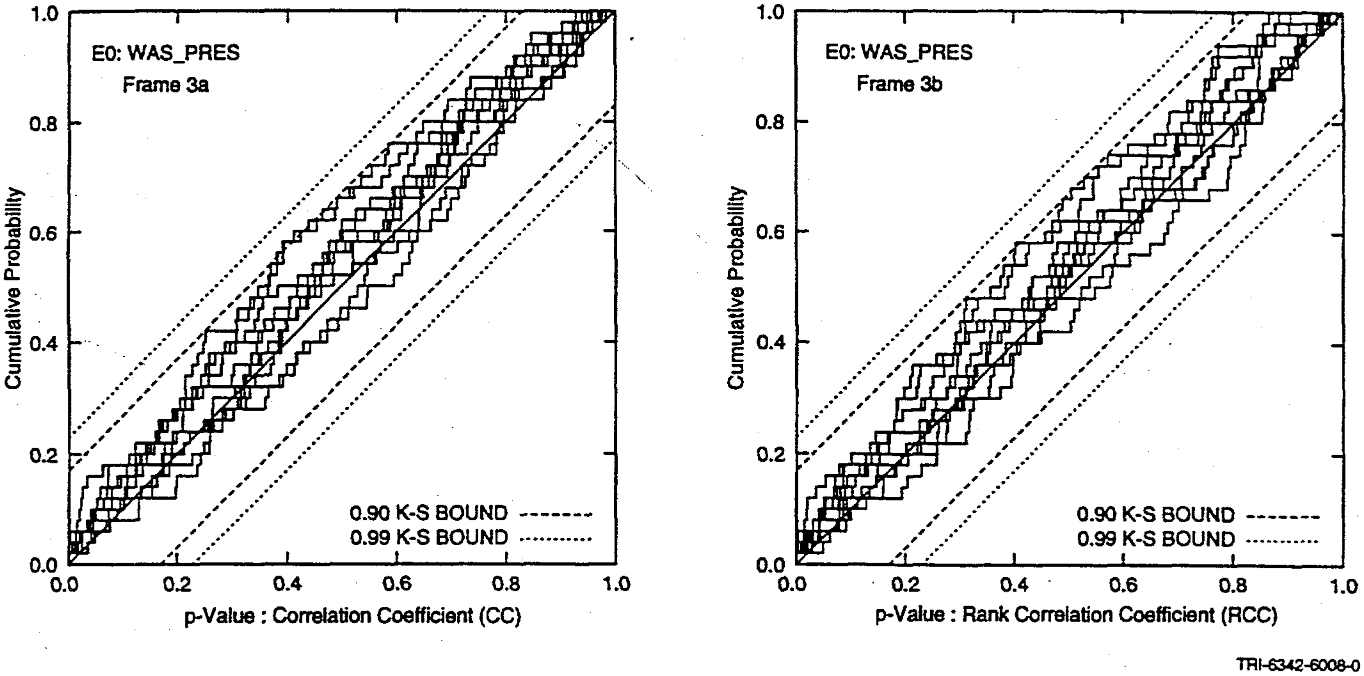

Fig. 3. Distributions of *p*-values for 10 independently-generated LHSs: (3a) CCs for *E0:WAS_PRES* and (3b) RCCs for *E0:WAS_PRES*.

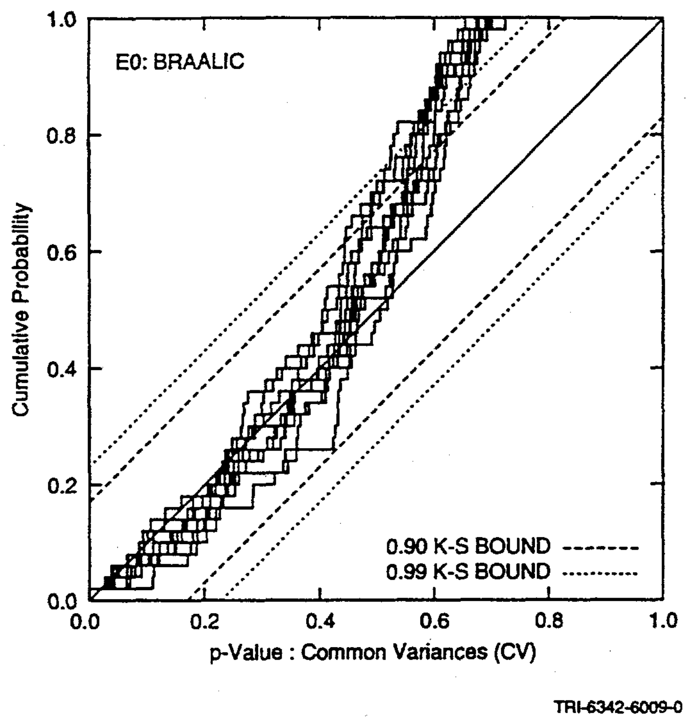

Fig. 4. Distribution of *p*-values for 10 independently-generated LHSs for CVs test and *E0:BRAALIC*.

Table 1. Comparison of Variable Rankings Obtained with Different Analysis Procedures[a] for Three Independent Samples of Size 100 (Columns AP:R1, AP:R2, AP:R3, where AP ~ CC, RCC, CMN, CL, CMD, CV, CIQ, SI as appropriate), Pooled Sample of Size 300 (Column AP:All), and a Maximum of Five Classes of Values for Each Variable (i.e., nX=5)[b]

Variable Name	CC: All Rank	CC: All *p*-Val	CC: R1 Rank	CC: R1 *p*-Val	CC: R2 Rank	CC: R2 *p*-Val	CC: R3 Rank	CC: R3 *p*-Val
Correlation Coefficients (CCs) for *E0:WAS_PRES*								
WMICDFLG	1.0	0.0000	1.0	0.0000	1.0	0.0000	1.0	0.0000
HALPOR	2.0	0.0000	2.0	0.0000	2.0	0.0000	2.0	0.0001
WGRCOR	3.0	0.0000	3.0	0.0180	3.0	0.0051	3.0	0.0018
ANHPRM	4.0	0.0241	9.0	0.3947	4.0	0.2371	4.0	0.0598
SALPRES	5.0	0.0855	4.0	0.0822	18.0	0.8602	7.0	0.2824
Correlation Coefficients (CCs) for *E2:WAS_SATB*								
BHPRM	1.0	0.0000	1.0	0.0000	1.0	0.0000	1.0	0.0000
ANHPRM	2.0	0.0000	2.0	0.0003	2.0	0.0281	2.0	0.0033
HALPOR	3.0	0.0006	3.0	0.0884	4.0	0.0706	3.0	0.0159
WGRCOR	4.0	0.0017	6.0	0.1241	3.0	0.0547	5.0	0.0473
WRGSSAT	5.0	0.0081	8.0	0.1367	11.0	0.5224	4.0	0.0175

Variable Name	CC: All Rank	CC: All *p*-Val	CC: R1 Rank	CC: R1 *p*-Val	CC: R2 Rank	CC: R2 *p*-Val	CC: R3 Rank	CC: R3 *p*-Val
Correlation Coefficients (CCs) for *E0:BRAALIC*								
ANHPRM	1.0	0.0000	1.0	0.0000	1.0	0.0000	1.0	0.0000
WMICDFLG	2.0	0.0000	2.0	0.0016	2.0	0.0060	2.0	0.0000
WASTWICK	3.0	0.0045	4.0	0.0584	9.0	0.2948	4.0	0.0333
WGRCOR	4.0	0.0048	5.0	0.0957	3.0	0.0318	9.0	0.3018
ANHBCEXP	5.0	0.0095	3.0	0.0420	6.0	0.1474	12.0	0.4274
Correlation Coefficients (CCs) for *E2:WAS_PRES*								
HALPRM	1.0	0.0000	1.0	0.0013	1.0	0.0000	1.0	0.0001
ANHPRM	2.0	0.0000	2.0	0.0020	2.0	0.0303	2.0	0.0267
HALPOR	3.0	0.0090	4.0	0.1417	3.0	0.0680	5.0	0.2188
ANHBCVGP	4.0	0.1123	8.0	0.3286	5.0	0.1492	20.0	0.7457
SHPRMASP	5.0	0.1606	10.0	0.3784	16.0	0.5907	6.0	0.3115

Variable Name	RCC: All Rank	RCC: All *p*-Val	RCC: R1 Rank	RCC: R1 *p*-Val	RCC: R2 Rank	RCC: R2 *p*-Val	RCC: R3 Rank	RCC: R3 *p*-Val
Rank Correlation Coefficients (RCCs) for *E0:WAS_PRES*								
WMICDFLG	1.0	0.0000	1.0	0.0000	1.0	0.0000	1.0	0.0000
HALPOR	2.0	0.0000	2.0	0.0000	2.0	0.0000	2.0	0.0001
WGRCOR	3.0	0.0000	3.0	0.0286	3.0	0.0041	3.0	0.0051
ANHPRM	4.0	0.0268	9.0	0.4366	4.0	0.1070	5.0	0.1268
SALPRES	5.0	0.0664	4.0	0.1111	16.0	0.7611	4.0	0.0957
Rank Correlation Coefficients (RCCs) for *E2:WAS_SATB*								
BHPRM	1.0	0.0000	1.0	0.0000	1.0	0.0000	1.0	0.0000
WRGSSAT	2.0	0.0000	2.0	0.0000	2.0	0.0048	2.0	0.0000
ANHPRM	3.0	0.0001	3.0	0.0013	3.0	0.1182	3.0	0.0335
SHPRMHAL	4.0	0.0225	4.0	0.1842	5.0	0.1243	9.0	0.2595
HALPOR	5.0	0.0269	8.5	0.4570	4.0	0.1236	4.0	0.1398

Variable Name	RCC: All Rank	RCC: All *p*-Val	RCC: R1 Rank	RCC: R1 *p*-Val	RCC: R2 Rank	RCC: R2 *p*-Val	RCC: R3 Rank	RCC: R3 *p*-Val
Rank Correlation Coefficients (RCCs) for *E0:BRAALIC*								
WMICDFLG	1.0	0.0000	1.0	0.0000	1.0	0.0000	1.0	0.0000
ANHPRM	2.0	0.0000	2.0	0.0000	2.0	0.0000	2.0	0.0000
HALPRM	3.0	0.0014	5.0	0.1867	5.0	0.0998	3.0	0.0140
WGRCOR	4.0	0.0057	4.0	0.1772	6.0	0.1383	4.0	0.0570
HALPOR	5.0	0.0087	3.0	0.0980	3.0	0.0396	7.0	0.3723
Rank Correlation Coefficients (RCCs) for *E2:WAS_PRES*								
HALPRM	1.0	0.0000	2.0	0.0110	1.0	0.0000	1.0	0.0001
ANHPRM	2.0	0.0000	1.0	0.0036	2.0	0.0820	2.0	0.0086
HALPOR	3.0	0.0184	4.0	0.1194	6.0	0.2015	5.0	0.2157
ANHBCVGP	4.0	0.1099	11.0	0.2611	4.0	0.1347	18.0	0.8795
WGRMICI	5.0	0.1477	5.0	0.1275	8.0	0.3344	23.0	0.9673

Variable Name	CMN: All,1×5 Rank	CMN: All,1×5 *p*-Val	CMN: R1,1×5 Rank	CMN: R1,1×5 *p*-Val	CMN: R2,1×5 Rank	CMN: R2,1×5 *p*-Val	CMN: R3,1×5 Rank	CMN: R3,1×5 *p*-Val
Common Means (CMNs) for *E0:WAS_PRES*								
WMICDFLG	1.0	0.0000	1.0	0.0000	1.0	0.0000	1.0	0.0000
HALPOR	2.0	0.0000	2.0	0.0002	2.0	0.0000	2.0	0.0010
WGRCOR	3.0	0.0000	3.0	0.0051	3.0	0.0093	3.0	0.0107
ANHPRM	4.0	0.0195	10.0	0.4751	6.0	0.2920	7.0	0.2881
SHPRMASP	5.0	0.1439	21.0	0.8597	5.0	0.1824	12.0	0.5410
Common Means (CMNs) for *E2:WAS_SATB*								
BHPRM	1.0	0.0000	1.0	0.0000	1.0	0.0000	1.0	0.0000
ANHPRM	2.0	0.0000	2.0	0.0031	2.0	0.0020	3.0	0.0544
WGRMICH	3.0	0.0021	4.0	0.0416	3.0	0.0471	2.0	0.0070
HALPOR	4.0	0.0124	10.0	0.2900	8.0	0.2345	10.0	0.3619
WRGSSAT	4.0	0.0143	5.0	0.0429	20.0	0.5575	6.0	0.1363

Variable Name	CMN: All,1×5 Rank	CMN: All,1×5 *p*-Val	CMN: R1,1×5 Rank	CMN: R1,1×5 *p*-Val	CMN: R2,1×5 Rank	CMN: R2,1×5 *p*-Val	CMN: R3,1×5 Rank	CMN: R3,1×5 *p*-Val
Common Means (CMNs) for *E0:BRAALIC*								
ANHPRM	1.0	0.0000	1.0	0.0014	1.0	0.0000	1.0	0.0000
WMICDFLG	2.0	0.0000	2.0	0.0040	2.0	0.0069	2.0	0.0001
SHPRMCON	3.0	0.0057	12.0	0.3818	8.0	0.3098	7.0	0.1531
WGRCOR	4.0	0.0636	5.0	0.1989	11.0	0.3914	20.0	0.5713
WFBETCEL	5.0	0.0732	10.0	0.3274	18.0	0.6060	12.0	0.2874
Common Means (CMNs) for *E2:WAS_PRES*								
BHPRM	1.0	0.0000	1.0	0.0000	1.0	0.0000	1.0	0.0000
HALPRM	2.0	0.0000	3.0	0.0288	2.0	0.0016	2.0	0.0027
ANHPRM	3.0	0.0002	2.0	0.0286	6.0	0.1137	5.0	0.1184
ANHBCEXP	4.0	0.0405	6.0	0.1860	5.0	0.1137	4.0	0.0230
HALPOR	5.0	0.0415	16.0	0.5971	4.0	0.0956	15.0	0.6365

Table 1. Comparison of Variable Rankings Obtained with Different Analysis Procedures[a] for Three Independent Samples of Size 100 (Columns AP:R1, AP:R2, AP:R3, where AP ~ CC, RCC, CMN, CL, CMD, CV, CIQ, SI as appropriate), Pooled Sample of Size 300 (Column AP:All), and a Maximum of Five Classes of Values for Each Variable (i.e., nX=5)[b] (continued)

Variable Name	CL: All,1×5 Rank	CL: All,1×5 *p*-Val	CL: R1,1×5 Rank	CL: R1,1×5 *p*-Val	CL: R2,1×5 Rank	CL: R2,1×5 *p*-Val	CL: R3,1×5 Rank	CL: R3,1×5 *p*-Val
	Common Locations (CLs) for *E0:WAS_PRES*							
WMICDFLG	1.0	0.0000	1.0	0.0000	1.0	0.0000	1.0	0.0000
HALPOR	2.0	0.0000	2.0	0.0003	2.0	0.0000	2.0	0.0023
WGRCOR	3.0	0.0000	3.0	0.0112	3.0	0.0093	3.0	0.0179
ANHPRM	4.0	0.0187	6.0	0.3792	6.0	0.2595	8.0	0.3770
SHPRMASP	5.0	0.1237	19.0	0.7696	5.0	0.1901	11.0	0.4537
	Common Locations (CLs) for *E2:WAS_SATB*							
BHPRM	1.0	0.0000	1.0	0.0000	1.0	0.0000	1.0	0.0000
WRGSSAT	2.0	0.0000	2.0	0.0000	3.0	0.0450	2.0	0.0001
ANHPRM	3.0	0.0001	3.0	0.0102	2.0	0.0184	7.0	0.2010
WGRMICH	4.0	0.0059	9.0	0.1714	5.0	0.0979	3.0	0.0206
SHPRMCON	5.0	0.0202	13.0	0.4691	10.0	0.2278	10.0	0.3785

Variable Name	CL: All, 1×5 Rank	CL: All, 1×5 *p*-Val	CL: R1,1×5 Rank	CL: R1,1×5 *p*-Val	CL: R2,1×5 Rank	CL: R2,1×5 *p*-Val	CL: R3,1×5 Rank	CL: R3,1×5 *p*-Val
	Common Locations (CLs) for *E0:BRAALIC*							
WMICDFLG	1.0	0.0000	1.0	0.0000	1.0	0.0000	1.0	0.0000
ANHPRM	2.0	0.0000	2.0	0.0000	2.0	0.0000	2.0	0.0000
HALPRM	3.0	0.0019	4.0	0.2667	6.0	0.2321	3.0	0.0125
WGRCOR	4.0	0.0427	6.0	0.3340	10.0	0.3212	13.0	0.4371
SHPRMDRZ	5.0	0.1060	5.0	0.2785	15.0	0.5898	9.0	0.2393
	Common Locations (CLs) for *E2:WAS_PRES*							
BHPRM	1.0	0.0000	1.0	0.0000	1.0	0.0000	1.0	0.0000
HALPRM	2.0	0.0000	4.0	0.1176	2.0	0.0025	2.0	0.0028
ANHPRM	3.0	0.0000	2.0	0.0154	3.0	0.0523	4.0	0.0419
ANHBCEXP	4.0	0.0602	7.0	0.2213	6.0	0.1191	5.0	0.0438
HALPOR	5.0	0.0940	18.0	0.5620	9.0	0.2452	11.0	0.5243

Variable Name	CMD: All, 2×5 Rank	CMD: All, 2×5 *p*-Val	CMD: R1,2×5 Rank	CMD: R1,2×5 *p*-Val	CMD: R2,2×5 Rank	CMD: R2,2×5 *p*-Val	CMD: R3,2×5 Rank	CMD: R3,2×5 *p*-Val
	Common Medians (CMDs) for *E0:WAS_PRES*							
WMICDFLG	1.0	0.0000	1.0	0.0000	1.0	0.0000	1.0	0.0000
HALPOR	2.0	0.0000	2.0	0.0001	2.0	0.0000	2.0	0.0123
WGRCOR	3.0	0.0025	5.0	0.1712	3.0	0.0663	4.0	0.1712
ANHPRM	4.0	0.0663	16.5	0.7358	16.0	0.6626	5.0	0.1991
SHPRMASP	5.0	0.2427	16.5	0.7358	7.0	0.2674	6.0	0.2674
	Common Medians (CMDs) for *E2:WAS_SATB*							
BHPRM	1.0	0.0000	2.0	0.0000	1.0	0.0000	2.0	0.0001
WRGSSAT	2.0	0.0000	1.0	0.0000	2.0	0.0015	1.0	0.0000
ANHPRM	3.0	0.0003	3.0	0.0073	13.0	0.2674	8.0	0.2674
WGRMICH	4.0	0.0130	9.0	0.1712	6.5	0.0477	4.0	0.0916
SHPRMCON	5.0	0.0206	16.0	0.7358	3.0	0.0244	6.5	0.1991

Variable Name	CMD: All,2×5 Rank	CMD: All,2×5 *p*-Val	CMD: R1,2×5 Rank	CMD: R1,2×5 *p*-Val	CMD: R2,2×5 Rank	CMD: R2,2×5 *p*-Val	CMD: R3,2×5 Rank	CMD: R3,2×5 *p*-Val
	Common Medians (CMDs) for *E0:BRAALIC*							
WMICDFLG	1.0	0.0000	1.0	0.0000	1.0	0.0000	1.0	0.0000
ANHPRM	2.0	0.0000	2.0	0.0009	2.0	0.0001	2.0	0.0003
HALPRM	3.0	0.0050	17.5	0.7358	13.0	0.4060	3.0	0.0021
HALPOR	4.0	0.0155	8.5	0.3084	5.0	0.0563	11.0	0.4060
WGRCOR	5.0	0.0231	3.0	0.1257	3.0	0.0244	16.0	0.5918
	Common Medians (CMDs) for *E2:WAS_PRES*							
BHPRM	1.0	0.0000	1.0	0.0000	1.0	0.0000	1.0	0.0000
HALPRM	2.0	0.0000	5.0	0.0663	2.0	0.0113	2.5	0.0289
ANHPRM	3.0	0.0007	3.0	0.0477	4.0	0.0780	5.0	0.1074
ANHBCEXP	4.0	0.0595	2.0	0.0289	3.0	0.0663	2.5	0.0289
HALPOR	5.0	0.0700	15.0	0.5918	5.5	0.1468	21.5	0.8781

Variable Name	CV: All,1×5 Rank	CV: All,1×5 *p*-Val	CV: R1,1×5 Rank	CV: R1,1×5 *p*-Val	CV: R2,1×5 Rank	CV: R2,1×5 *p*-Val	CV: R3,1×5 Rank	CV: R3,1×5 *p*-Val
	Common Variances (CVs) for *E0:WAS_PRES*							
WMICDFLG	1.0	0.0000	1.0	0.0016	1.0	0.0058	1.0	0.0004
ANHPRM	2.0	0.0042	17.0	0.8561	2.0	0.0171	2.0	0.0671
HALPRM	3.0	0.1184	14.0	0.6965	10.0	0.4818	9.0	0.2016
WGRCOR	4.0	0.1244	16.0	0.8147	7.0	0.3521	3.0	0.0943
SHPRMCON	5.0	0.1287	22.0	0.9555	3.0	0.0381	10.0	0.2053
	Common Variances (CVs) for *E2:WAS_SATB*							
BHPRM	1.0	0.0000	1.0	0.0000	1.0	0.0001	1.0	0.0000
ANHPRM	2.0	0.0000	2.0	0.0030	2.0	0.0018	2.0	0.0134
HALPOR	3.0	0.0011	5.0	0.0243	6.0	0.1228	12.0	0.3601
WGRMICH	4.0	0.0050	3.0	0.0156	7.0	0.1353	3.0	0.1013
WGRCOR	5.0	0.0067	4.0	0.0210	3.0	0.0080	10.0	0.3387

Variable Name	CV: All,1×5 Rank	CV: All,1×5 *p*-Val	CV: R1,1×5 Rank	CV: R1,1×5 *p*-Val	CV: R2,1×5 Rank	CV: R2,1×5 *p*-Val	CV: R3,1×5 Rank	CV: R3,1×5 *p*-Val
	Common Variances (CVs) for *E0:BRAALIC*							
ANHPRM	1.0	0.0078	1.0	0.2779	1.0	0.0576	1.0	0.0005
SHPRMCON	2.0	0.0426	13.0	0.4026	6.0	0.0938	22.0	0.6452
SHBCEXP	3.0	0.1463	17.0	0.4412	21.0	0.5811	23.0	0.6840
ANRBRSAT	4.0	0.1994	6.0	0.3463	19.0	0.5557	9.0	0.1909
WGRCOR	5.0	0.2125	11.0	0.3969	9.0	0.4550	15.0	0.4175
	Common Variances (CVs) for *E2:WAS_PRES*							
BHPRM	1.0	0.0000	1.0	0.0082	1.0	0.0633	11.0	0.2843
HALPRM	2.0	0.0014	4.0	0.1392	4.0	0.1719	9.0	0.2669
WGRCOR	3.0	0.0296	2.0	0.0329	8.0	0.3272	5.0	0.0741
SHPRMDRZ	4.0	0.0298	3.0	0.0713	5.0	0.2094	13.0	0.3795
ANHBCVGP	5.0	0.1173	20.0	0.7401	9.0	0.4415	1.0	0.0178

Table 1. Comparison of Variable Rankings Obtained with Different Analysis Procedures[a] for Three Independent Samples of Size 100 (Columns AP:R1, AP:R2, AP:R3, where AP ~ CC, RCC, CMN, CL, CMD, CV, CIQ, SI as appropriate), Pooled Sample of Size 300 (Column AP:All), and a Maximum of Five Classes of Values for Each Variable (i.e., nX=5)[b] (continued)

Variable Name	CIQ: All, 2×5 Rank	CIQ: All, 2×5 *p*-Val	CIQ: R1,2×5 Rank	CIQ: R1,2×5 *p*-Val	CIQ: R2,2×5 Rank	CIQ: R2,2×5 *p*-Val	CIQ: R3,2×5 Rank	CIQ: R3,2×5 *p*-Val
	Common Interquartile (CIQ) for *E0:WAS_PRES*							
WMICDFLG	1.0	0.0000	1.0	0.0057	1.0	0.0012	1.0	0.0001
HALPOR	2.0	0.0000	2.0	0.1257	2.0	0.0206	2.0	0.0061
ANHPRM	3.0	0.0007	8.5	0.4628	3.0	0.0342	5.0	0.0780
SHPRMCON	4.0	0.0244	17.0	0.6626	16.0	0.4628	10.0	0.3546
WGRMICI	5.0	0.0595	13.5	0.5918	6.5	0.1468	17.0	0.8088
	Common Interquartile (CIQ) for *E2:WAS_SATB*							
WRGSSAT	1.0	0.0000	1.0	0.0001	1.0	0.0001	1.0	0.0000
WGRCOR	2.0	0.0019	9.0	0.2311	4.0	0.0563	6.0	0.1468
BHPRM	3.0	0.0054	2.0	0.0206	13.5	0.3546	2.0	0.0289
ANHPRM	4.0	0.0628	13.0	0.4628	6.0	0.1074	5.0	0.1074
SHRBRSAT	5.0	0.1257	11.0	0.2674	5.0	0.0663	20.0	0.6626

Variable Name	CIQ: All,2×5 Rank	CIQ: All,2×5 *p*-Val	CIQ: R1,2×5 Rank	CIQ: R1,2×5 *p*-Val	CIQ: R2,2×5 Rank	CIQ: R2,2×5 *p*-Val	CIQ: R3,2×5 Rank	CIQ: R3,2×5 *p*-Val
	Common Interquartile (CIQ) for *E0:BRAALIC*							
ANHPRM	1.0	0.0000	17.0	0.7358	6.5	0.1468	1.0	0.0001
WMICDFLG	2.0	0.0000	2.0	0.0392	1.0	0.0321	2.0	0.0057
SHRGSSAT	3.0	0.0628	13.5	0.5918	5.0	0.1074	5.5	0.2674
WGRMICI	4.0	0.0780	5.0	0.2311	3.0	0.0663	7.0	0.3084
SHRBRSAT	5.0	0.1395	7.0	0.3084	20.0	0.8088	24.0	0.8781
	Common Interquartile (CIQ) for *E2:WAS_PRES*							
BHPRM	1.0	0.0000	1.0	0.0002	1.0	0.0021	2.0	0.0061
WGRCOR	2.0	0.0130	2.0	0.1074	7.0	0.2674	7.0	0.1468
SHRGSSAT	3.0	0.0289	11.0	0.4628	2.0	0.1074	11.0	0.3546
ANRBRSAT	4.0	0.0739	22.5	0.8781	14.5	0.5918	10.0	0.3084
SHRBRSAT	5.0	0.2093	18.5	0.7358	8.0	0.4060	1.0	0.0051

Variable Name	SI: All, 5×5 Rank	SI: All, 5×5 *p*-Val	SI: R1,5×5 Rank	SI: R1,5×5 *p*-Val	SI: R2,5×5 Rank	SI: R2,5×5 *p*-Val	SI: R3,5×5 Rank	SI: R3,5×5 *p*-Val
	Statistical Independence (SI) for *E0:WAS_PRES*							
WMICDFLG	1.0	0.0000	1.0	0.0000	1.0	0.0000	1.0	0.0000
HALPOR	2.0	0.0000	2.0	0.0034	2.0	0.0000	2.0	0.0000
WGRCOR	3.0	0.0003	13.5	0.4884	4.0	0.0316	9.5	0.2687
ANHPRM	4.0	0.0049	4.5	0.1785	8.5	0.2202	3.0	0.1010
ANHBCVGP	5.0	0.0194	3.0	0.1712	13.0	0.3546	18.0	0.7358
	Statistical Independence (SI) for *E2:WAS_SATB*							
WRGSSAT	1.0	0.0000	1.0	0.0000	1.0	0.0000	1.0	0.0000
BHPRM	2.0	0.0000	2.0	0.0000	2.0	0.0000	2.0	0.0001
ANHPRM	3.0	0.0002	3.0	0.0316	7.5	0.0415	7.0	0.2954
ANRBRSAT	4.0	0.0495	14.0	0.4530	5.0	0.0275	23.5	0.9134
WGRMICH	5.0	0.0564	11.0	0.3856	9.0	0.0895	8.5	0.3540

Variable Name	SI: All, 5×5 Rank	SI: All, 5×5 *p*-Val	SI: R1,5×5 Rank	SI: R1,5×5 *p*-Val	SI: R2,5×5 Rank	SI: R2,5×5 *p*-Val	SI: R3,5×5 Rank	SI: R3,5×5 *p*-Val
	Statistical Independence (SI) for *E0:BRAALIC*							
WMICDFLG	1.0	0.0000	1.0	0.0000	1.0	0.0000	1.0	0.0000
ANHPRM	2.0	0.0000	2.0	0.0003	2.0	0.0001	2.0	0.0000
HALPRM	3.0	0.0517	7.5	0.3540	21.5	0.7776	5.0	0.1137
HALPOR	4.0	0.0698	17.5	0.7089	8.0	0.2202	17.0	0.7440
SHRBRSAT	5.0	0.1917	23.0	0.8392	10.0	0.2687	3.0	0.0540
	Statistical Independence (SI) for *E2:WAS_PRES*							
BHPRM	1.0	0.0000	1.0	0.0000	1.0	0.0000	1.0	0.0000
HALPRM	2.0	0.0002	10.0	0.2954	3.0	0.1137	7.0	0.2202
WGRCOR	3.0	0.0002	2.0	0.0100	18.0	0.7089	2.0	0.0362
ANHPRM	4.0	0.0049	14.0	0.3856	9.0	0.3856	5.0	0.1785
SHRGSSAT	5.0	0.0698	17.5	0.5615	16.0	0.6359	11.0	0.3856

a Twenty-four (24) variables included in analysis for *E0:WAS_PRES* and *E0:BRAALIC* (see Footnote b to Table 4, Ref. 1); twenty-five (25) variables included in analysis for *E2:WAS_SATB* and *E2:WAS_PRES* (see Footnote b to Table 17, Ref. 1); for each test and dependent variable, top five variables based on their ordering with *p*-values obtained from pooled sample of size 300 are included in table.

b See Footnote c, Table 4, Ref. 1.

Table 2. Comparison of Variable Rankings Obtained with Common Variances (CVs) Test with Use of Logarithms[a] for Three Independent Samples of Size 100 (Column CV:R1, CV:R2, CV:R3) and Pooled Sample of Size 300 (Column CV:All) for *y=E0:BRAALIC*[b]

Variable	CV: All.1×5		CV: R1.1×5		CV: R2.1×5		CV: R3.1×5	
Name	Rank	*p*-Val	Rank	*p*-Val	Rank	*p*-Val	Rank	*p*-Val
ANHPRM	1.0	0.0000	1.0	0.0000	1.0	0.0000	2.0	0.0000
WMICDFLG	2.0	0.0002	10.0	0.0251	7.0	0.0035	1.0	0.0000
SHPRMCON	3.0	0.0019	11.0	0.0257	5.0	0.0022	22.0	0.7184
SHBCEXP	4.0	0.0130	15.0	0.0528	19.0	0.2129	21.0	0.6442
WASTWICK	5.0	0.0144	13.0	0.0387	4.0	0.0002	17.0	0.3413

a See Footnote a, Table 10, Ref. 1, for description of test.

b See Footnote a, Table 1.