New Jersey Public Employment Relations Commission

Interest Arbitrator By-Lot Selection
Program:
Random Number Generation
Testing for Re-Certification

By

Francis A. Steffero, PhD, CISA

December 27, 2011
I. INTRODUCTION

In September, 1991, the New Jersey Public Employment Relations Commission (PERC) implemented a computer-assisted system to create interest arbitration panels. The system was designed to assign interest arbitrators to panels in a random manner. The system used a computer-based random number generator supplied by the equipment manufacturer, Wang Laboratories, Inc.

PERC commissioned a study to certify that the computer system performed in a random manner consistent with requirements set forth in N.J.S.A. 34:13A-16 and N.J.A.C. 19:16-5.6. The study (Steffero, 1991) used statistical techniques recommended by Knuth (1981) and confirmed the system performed as expected. The system was modified in 1996 to comply with a revision in N.J.S.A. 34:13A-16e(2) which changed the selection of interest arbitrators from a panel selection process to a direct by-lot appointment process. PERC commissioned a second study (Steffero, 1996) which certified that the system assigned interest arbitrators in an unbiased manner.

In 2005, the Wang Laboratories, Inc., hardware and software used to create and operate the computer-assisted system reached the end of its life cycle. PERC selected Specialty Systems, Inc. (SSI) to develop a new system based on the original requirements. SSI used Lotus Notes, an IBM product, and Microsoft’s Windows 2003 Server as the hardware and software platform. Lotus Script is the programming language for Lotus Notes and was used to program the new system. SSI used the random number generator provided by IBM in the Lotus Script programming language as the source of random numbers used in the algorithm to select interest arbitrators.

The Lotus Notes system was tested in 2005 and 2009 (Steffero, 2005, 2009) to confirm that the new computer assisted system assigned interest arbitrators in a random manner. The methodology of the study applied a statistical test described by Donald E. Knuth (1981, 1998), professor emeritus from Stanford University. The results of the study confirmed that the random number generator provided by IBM in Lotus Script generated random numbers. The results of the study also confirmed that the programming provided by SSI selected interest arbitrators in a random manner (Steffero, 2005, 2009).

In 2011, the Lotus Notes system was retested to confirm that the computer assisted system complies with the interest arbitrator appointment procedures amended by L. 2010 c. 105 effective January 1, 2011; assigned interest arbitrators in a random manner; and followed the methodology from the past studies (Steffero, 2005, 2009). This PRNG (Pseudo Random Number Generator) test was not repeated because there had been no changes to the IBM random number generator between 2009 and 2011. The Completed Application Test was performed on November 15, December 22 and December 23, 2011. The results of the November 15 and December 22 testing identified the need to improve the interest arbitrator selection process for very small arbitrator pool sizes. SSI installed a programming update on December 23 and subsequent testing confirmed that the system performed successfully with very small pool sizes to select interest arbitrators in a random manner.
II. BACKGROUND INFORMATION

In this study, the term random is defined as “…a process of selection in which each item of a set has an equal probability of being chosen” (Flexner, 1987). Therefore, if each item of a set has an equal chance of being selected, then the selection process is free from bias. In this study, if every eligible interest arbitrator has an equal probability of being selected, then the selection process behaves in a random manner.


Knuth (1998) explained that true randomness comes from natural phenomenon. He pointed out that digital computers are deterministic which means that they use algorithms, or formulae, to create random numbers. He used the term pseudo-random number to describe a random number generated by a digital computer and he called the computer programs that create them “pseudo-random number generators,” or PRNGs. Knuth (1998) also described testing methods for PRNGs in detail. He called the Chi-square test “…perhaps the best known of all statistical tests, and it is a basic method that is used in connection with many other tests” (p. 42).

The Chi-square test compares the observed results of the PRNG with the expected results, and then determines the probability that the results are random or not random. For example, if one tosses an unbiased coin 100 times, one would expect the perfect result to be ‘heads” 50 times, and tails “50” times. To determine if the method of tossing the coin is biased or unbiased, the coin must be tossed many times and the results examined. If the method of tossing the coin is unbiased, then the observed results will approach the expected results as the test is repeated over and over again. If the coin toss method is biased, then the observed results will not match the expected results.

The Chi-square test is also known as a “Goodness of Fit” test (Siegel, 1956) and means that the goal of the test is to measure how well the coin toss results will “fit” the expected distribution. Since the purpose of this study was to compare the observed results of the computer-assisted system with the expected results of a random selection process, the Chi-square goodness of fit test was selected.

The PRNG in Lotus Script is called the “Rnd” function. A critical component of a PRNG is the method it uses to obtain a “seed” value. The “seed” can directly determine the random value a PRNG will produce. If the same seed value is used each time a PRNG is executed, then the same pseudo-random value will be produced. In the present study, the computer-assisted system required that a unique pseudo-random value was generated each time the PRNG was executed.
The method in Lotus Script which ensures that a unique “seed” is provided to the "Rnd" function by the use of two subordinate functions, "Randomize" and "Timer." The "Randomize" function obtains the "seed" value from the "Timer" function. The "seed" value in the "Timer" function is the number of seconds elapsed since midnight expressed in hundredths of a second. Therefore, the combination of "Rnd," "Randomize," and "Timer" ensures that a unique "seed" value is obtained each time the PRNG function is executed.

Knuth (1998, p. 184) confirms that system clock functions are a common source for obtaining initial values to "seed" computer based random number generators. The method implemented by IBM in Lotus Script appears consistent with good practices. The study author conducted a computer “code” review with SSI and PERC staff and verified that the PRNG developed by SSI using Lotus Script is consistent with implementation guidelines recommended in the IBM Lotus Script documentation.
III. METHODOLOGY

The present study examined two possible sources for bias, or non-random behavior, in the PERC computer-assisted system arbitrator selection process. The first source of possible bias is performance of the IBM Lotus Script “Rnd” function supplied by the manufacturer, IBM and used by Specialty Systems, Inc., in a function called "getrandoms." The purpose of the PRNG test is to confirm that the basic function by itself is behaving in a random manner.

Even if the basic random function performs as designed, it is still possible that its use in the full information system could introduce bias. Therefore, the second test focuses on the selection process using the complete application. This was called the Completed Application Test.

Production Server Environment

All certification testing was performed on the production environment at PERC. The major components of the environment at PERC were the server hardware, operating system and Lotus Notes Server. The production server hardware was a Hewlett-Packard ProLiant, DL380 G4 server with dual 3.6 gigahertz processors, 4 gigabytes of random access memory (RAM) and a high performance, SCSI disk subsystem. The production server operating system was Windows 2003 Server, Standard Edition, Version 5.2, and Service Pack 1, by Microsoft Corporation. The Lotus software version was Lotus Domino Server, Release 7.1 for Windows, January 17, 2006. The server hardware, operating system, and Lotus Notes software used for the PERC system were consistent with generally accepted standards for high performance, production server environments at the time of this study.

There were no changes in the production environment between the 2009 and 2011 re-certification testing. Therefore, the PRNG Test was not performed. A description of the test is included to keep all certification report descriptions of methodology consistent and repeatable for future certifications.

PRNG Test (Steffero, 2009)

To perform the PRNG test, the Lotus Script “Rnd” function was executed 1,000 times in the production environment using a script requested by the author and written by SSI for this study. The script used the “Rnd” function to generate 1,000 pseudo-random numbers between 0 and 1, and then rounded each number to produce a test value between 1 and 10.

If one were to select the number 1 through 10 at random 1,000 times, one would expect to obtain the value “1” 100 times, the value “2” 100 times, and so on through the value “10.” To test the randomness of the actual computed values, the study compared the actual outcome with the expected outcome. If the actual outcome matched the expected outcome, then the outcome is random. The Chi-square test was selected to measure the
goodness of fit. The level of precision, or significance, was set at the .01 level. This means that if the test was repeated an infinite number of times, the probability that the results would be the same is 99%.

*Completed Application Test*

The Completed Application Test examined the actual arbitrator selection functionality of the system. To determine if the procedure of selecting one arbitrator from a pool of eight arbitrators behaved in a random manner, an automated test was executed 300 times and the results were recorded, analyzed and presented in Table 2, Test 1, on December 23, 2011. The automated test script was executed two more times to produce Test 2 and Test 3, respectively, on that date to comply with Knuth's (1998, p. 47) recommendation to perform the test 3 times.

If there was no bias in the selection of arbitrators reported in Table 2, then one would expect to select the first arbitrator 37.5 times (300/8 = 37.5), the second arbitrator 37.5 times, and so on until all arbitrators were selected. If the computer-generated results match the expected random results and pass the Chi-square test, then the outcome is random. The level of precision, or significance, was set at the .01 level. This means that if the tests were repeated an infinite number of times, the probability that the results would be the same is 99%.

Results appear in the next section.
IV. RESULTS

The results are divided into two sections: PRNG Test (results from Steffero, 2009) and Completed Application Test for Interest Arbitrator Selection.

**PRNG Test (from Steffero, 2009)**

The results of the PRNG Test are presented below in the Table 1 below. The Chi-square test accepted the null hypothesis that there was no significant difference between the observed and expected results at the .01 level of significance. Therefore, there is a 99% probability that the pseudo-random number generator is behaving in a random manner, as designed by the manufacturer.

**Table 1.** Results of the PRNG Test in 2009
(n = 1,000)

<table>
<thead>
<tr>
<th>CHOICE</th>
<th>TEST</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>97</td>
</tr>
<tr>
<td>2</td>
<td>99</td>
</tr>
<tr>
<td>3</td>
<td>80</td>
</tr>
<tr>
<td>4</td>
<td>89</td>
</tr>
<tr>
<td>5</td>
<td>114</td>
</tr>
<tr>
<td>6</td>
<td>112</td>
</tr>
<tr>
<td>7</td>
<td>97</td>
</tr>
<tr>
<td>8</td>
<td>92</td>
</tr>
<tr>
<td>9</td>
<td>114</td>
</tr>
<tr>
<td>10</td>
<td>106</td>
</tr>
</tbody>
</table>

At the .01 Level of Significance with df = 9, Chi-square must be less than 21.67. The test indicates that the results do not differ from a random distribution.
Completed Application Test for Interest Arbitrator By-Lot Selection

The results of the Completed Application Test for Interest Arbitrator By-Lot Selection are presented in Table 2 below. The Chi-square test accepted the null hypothesis that there was no significant difference between the observed and expected results at the .01 level of significance. Therefore, there is a 99% level of confidence that the selection of arbitrators from a pool of 8 interest arbitrators is behaving in a random manner.

Table 2. Results of Completed Application Test: Interest Arbitrator Selection (n=300)

<table>
<thead>
<tr>
<th>Actual Arbitrator</th>
<th>Test 1</th>
<th>Test 2</th>
<th>Test 3</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>33</td>
<td>43</td>
<td>39</td>
</tr>
<tr>
<td>2</td>
<td>33</td>
<td>43</td>
<td>35</td>
</tr>
<tr>
<td>3</td>
<td>36</td>
<td>30</td>
<td>38</td>
</tr>
<tr>
<td>4</td>
<td>37</td>
<td>38</td>
<td>33</td>
</tr>
<tr>
<td>5</td>
<td>39</td>
<td>31</td>
<td>30</td>
</tr>
<tr>
<td>6</td>
<td>46</td>
<td>36</td>
<td>46</td>
</tr>
<tr>
<td>7</td>
<td>35</td>
<td>48</td>
<td>27</td>
</tr>
<tr>
<td>8</td>
<td>41</td>
<td>31</td>
<td>52</td>
</tr>
<tr>
<td>k=8</td>
<td>300</td>
<td>300</td>
<td>300</td>
</tr>
<tr>
<td>Chi-Square</td>
<td>3.63</td>
<td>8.37</td>
<td>12.75</td>
</tr>
</tbody>
</table>

At the .01 Level of Significance with df = 7, Chi-square must be less than 18.48. The test indicates that the results do not differ from a random distribution.
V. CONCLUSION

The study confirmed that the random behavior of the computer-assisted method is consistent with the requirements set forth in N.J.S.A. 34:13A-16e and N.J.A.C. 19:16-5.6. The pseudo-random number generator provided by IBM/Lotus behaved in a random manner based on prior testing (Steffero, 2009). The computer-assisted processes developed by Specialty Systems, Inc. and modified on December 23, 2011, for selecting interest arbitrators by-lot behaved in a random manner.
BIBLIOGRAHY


Signature Page

I hereby certify to the authenticity of the report entitled:

New Jersey Public Employment Relations Commission

Interest Arbitrator By-Lot Selection
Program:

Random Number Generation
Testing for Re-Certification

Francis A. Steffero, PhD, CISA

December 27, 2011

Date