Evaluating TreeAge's Pseudo Random Number Generator and Implementation of Statistical Distributions

Bernhard Klingenberg
Department of Mathematics & Statistics, Williams College, Williamstown, MA

Wojciech (Al) Chrosny
TreeAge Software, Inc., Williamstown, MA

1.) Validity of the uniform number generator used in TreeAge Pro

To establish if the uniform number generator implemented in the TreeAge software delivers the expected outcome we investigated the probabilistic and statistical properties of pseudo uniform random numbers generated with TreeAge in several ways. We computed the first four empirical central and non-central moments of a uniform sequence generated with TreeAge and compared them to their theoretical counterparts[1]. We also conducted a global goodness of fit test known as the Kolmogorov-Smirnov test (K-S test) [2] and looked at a Quantile-Quantile plot (Q-Q plot)[3] to inspect deviation from the expected outcome graphically. Finally, we employed a whole battery of tests designed to detect deviations from the expected outcome, known as the Die Hard battery of tests [4]. The collection of 17 different tests we used checks whether the generated uniform sequence from TreeAge has the probabilistic and statistical properties expected of such a sequence.

We used a sequences of length $1.5 \cdot 10^8$, each generated from the uniform distribution on [0,1] in TreeAge. Each sequence was read into the statistical software program R [5], which was used to

- i.) compute the central and non-central moments of the generated sequence and compare them to the theoretical ones,
- ii.) run the K-S test and obtain a P-value for that test,
- iii.) construct the Q-Q plot and
- iv.) convert the stream of TreeAge uniform numbers into 32-bit integers, which is the input required for running the battery of Die Hard tests.

---

[1] Weisstein, Eric W. "Uniform Distribution." From MathWorld--A Wolfram Web Resource. http://mathworld.wolfram.com/UniformDistribution.html

[2] Refer to the Wikipedia page http://en.wikipedia.org/wiki/Kolmogorov%E2%80%93Smirnov_test for a good description and references.

[3] Refer to the Wikipedia page http://en.wikipedia.org/wiki/Q-Q_plot for a good description and references.

[4] George Marsaglia. The Marsaglia random number CDROM including the diehard battery of tests of randomness. Also at http://stat.fsu.edu/pub/diehard., 1996. See also the Wikipedia entry at http://en.wikipedia.org/wiki/Diehard_tests.

[5] R Development Core Team. R: A Language and Environment for Statistical Computing. R Foundation for Statistical Computing, Vienna, Austria, 2013. URL http://www.R-project.org.

For running the Die Hard battery of tests, we used software provided by Robert Brown[6]. For each of these four points we deemed the sequence as passing if no unusual outcomes were observed. For point i.), this means all central and non-central moments were reasonably close to their theoretical counterparts, for point ii.) this means a P-value larger than 5% for the K-S test, for point iii.) this means no visual deviations from a 45 degree line for the Q-Q plot and for point iv.) this means all 17 tests that part of the Die Hard battery of tests indicate a passing result as defined in Brown[6]. The results (whether or not the TreeAge sequence fails or passes any of the four points mentioned above) are summarized in Table 1.

From running all these checks on the sequence of $1.5 \cdot 10^8$ random numbers and the results in Table 1, we infer that a sequence of pseudo uniform numbers as generated with TreeAge has the probabilistic and statistical properties expected from a sequence of uniform random numbers on the interval [0,1]. **We conclude that using the uniform random number generator in TreeAge is valid and delivers the expected results.**

Table 1: Evaluating TreeAge Pro's Uniform Random Number Generator, using an input sequence of $1.5 \cdot 10^8$ pseudo uniform random numbers generated with TreeAge.

Distribution: Uniform on the interval [0,1]

Method to Evaluate TreeAge Sequence	*Result*
Moments	Passed
QQ plot	Passed
K-S test	Passed
Die Hard battery	Passed

2.) Validity of implementations of *continuous distributions* in TreeAge Pro

We used a similar technique to evaluate distributions other than the uniform that are available in TreeAge Pro. In particular, we tested the validity of the implementation to generate random numbers from the Exponential, Weibull, Normal, Gamma, Rayleigh and Erlang distributions. (Results for the Gamma distributions are preliminary.) For each distribution, a sequence of $1.5 \cdot 10^8$ pseudo random numbers was generated with TreeAge Pro and read into the statistical software program R. Each sequence was then subjected to the following checks: i.) Central, non-central or standardized empirical moments were compared to their theoretical counterparts to see if they agreed up to numerical accuracy , ii.) the K-S test was carried out comparing the empirical cumulative distribution function (cdf) of the TreeAge sequence to the theoretical cdf of the corresponding distribution and iii.) the Die Hard battery of tests was applied to a transformed sequence. To prepare the TreeAge sequence for the Die Hard battery of tests, which needs input of uniform 32-bit integers, we computed the cumulative distribution function (cdf) at each element in the TreeAge generated sequence using R. Results from

[6] Robert G. Brown. dieharder: A Random Number Test Suite, 2007. URL: http://www.phy.duke.edu/~rgb/General/dieharder.php. C program archive dieharder, version *3.31.1*.

probability theory imply that the resulting cdf-transformed sequence must be independent and identically distributed (i.i.d.) following the uniform distribution[7], and this was tested with 17 tests from the Die Hard battery for each distribution.
As each distribution depends on one or two parameters, the above scenario was repeated for a set of representative parameter values that are likely to come up in practical use of TreeAge Pro.

We also investigated the sensitivity of our checks. For this purpose, we generated sequences using slightly different parameter values than specified. E.g., for the normal distribution, while theoretical moments, Q-Q plots, the K-S test or the battery of Die Hard tests were implemented using $\sigma^2 = 1$, we actually generated the TreeAge sequence using $\sigma^2 = 1.02^2$ or mixing together two sequences with slightly different variances σ^2. The goal was to see how sensitive our results (whether or not the generated sequence passes or fails the various checks) are for small departures from the specified parameter values. We found that in all instances, at least one of our checks immediately signaled a failing result, even under the slightest changes of the parameter values. This gives us confidence that our checks are sensitive in detecting unexpected TreeAge output.

a. The Exponential Distribution

From the help file, TreeAge parameterizes the density (pdf) of the exponential distribution as

$$f(x;\lambda) = \lambda e^{-\lambda x}, \qquad x > 0, \qquad \lambda > 0,$$

so that $E[X] = 1/\lambda$ and $Var[X] = 1/\lambda^2$. To check if a sequence from the exponential distribution generated with TreeAge has all the probabilistic and statistical properties of a true exponential sequence, a sequence of length $1.5 \cdot 10^8$ was generated in TreeAge and the above mentioned checks performed, for a variety of values for the parameter λ. Results are listed in Tables 2a to 2g. **We conclude that the implementation of the exponential distribution in TreeAge is valid and produces the expected outcome.**

Table 2: Evaluating TreeAge Pro's Exponential Random Number Generator, using an input sequence of $1.5 \cdot 10^8$ pseudo exponential random numbers generated with TreeAge.

a. Distribution: Exponential with $\lambda = 1$

Method to Evaluate TreeAge Sequence	*Result*
Moments	Passed
K-S test	Passed
Die Hard battery	Passed

[7] This is known as the Probability Integral Transform, see e.g., the Wikipedia entry at http://en.wikipedia.org/wiki/Probability_integral_transform or Theorem 2.1.10 on page 54 in Casella & Berger, Statistical Inference, 2nd ed. Pacific Grove, CA:Wadsworth, 2001.

b. Distribution: Exponential with $\lambda = 5$

Method to Evaluate TreeAge Sequence	*Result*
Moments	Passed
K-S test	Passed
Die Hard battery	Passed*

** Two of the seventeen Die Hard tests indicated a "weak" agreement between the theoretical and provided sample. This is not alarming, for two reasons.*[8]

c. Distribution: Exponential with $\lambda = 30$

Method to Evaluate TreeAge Sequence	*Result*
Moments	Passed
K-S test	Passed
Die Hard battery	Passed

d. Distribution: Exponential with $\lambda = 2500$

Method to Evaluate TreeAge Sequence	*Result*
Moments	Passed
K-S test	Passed
Die Hard battery	Passed

e. Distribution: Exponential with $\lambda = 0.005$

Method to Evaluate TreeAge Sequence	*Result*
Moments	Passed
K-S test	Passed
Die Hard battery	Passed

[8] *Firstly, the Die Harder software declares a test result as "weak" when its P-value is smaller than 0.005 or larger than 0.995. While a P-value less than 0.005 indicates significant departure of the TreeAge distribution from the theoretical one, a P-value larger than 0.995 refers to an agreement between the supplied sample and the theoretical distribution that is almost too good to be true. For instance, the P-value of 0.9985 for one of the "weak" results implies that such a good agreement would be observed in less than 15 out of 10,000 cases if indeed the two agree. Since this is unlikely to happen, Die Harder flags the result. For our purpose, results flagged as "weak" that are associated with a **large** P-value have no consequence. They indicate almost perfect agreement between TreeAge Pro and theoretical distributions, so perfect, that there may be suspicion that someone has forged the data to agree perfectly. Since this is not the case, we do not need to worry about a "weak" (or even "failed") result when the corresponding P-value is large.*
Secondly, as with any statistical decision, there is the possibility of a Type 1 error (false positive). Since a test is deemed weak at the 0.5% significance level, there is a small chance (i.e. 0.5% chance) that this is actually the wrong decision. Since a total of 17 Die Hard tests are run, there is the possibility of committing this error several times. (This is called multiplicity in statistics.) So, it is not surprising that we may get a few weak results when running several of such tests.

f. Distribution: Exponential with $\boldsymbol{\lambda = 0.0007}$

Method to Evaluate TreeAge Sequence	*Result*
Moments	Passed
K-S test	Passed
Die Hard battery	Passed

g. Distribution: Exponential with $\boldsymbol{\lambda = 10^{-8}}$

Method to Evaluate TreeAge Sequence	*Result*
Moments	Passed
K-S test	Passed
Die Hard battery	Passed

b. The Weibull Distribution

From the help file, TreeAge parameterizes the density (pdf) of the Weibull distribution as

$$f(x; \lambda, k) = \lambda k x^{k-1} e^{-\lambda x^k}, \qquad x > 0, \qquad \lambda, k > 0,$$

with expected value equal to $\mathrm{E[X]} = \lambda^{-1/\mathrm{k}}\Gamma(1 + 1/\mathrm{k})$. To check if a sequence from the Weibull distribution generated with TreeAge has all the probabilistic and statistical properties of a true Weibull sequence, a sequence of length $1.5 \cdot 10^8$ was generated in TreeAge and the above mentioned checks performed, for a variety of values for the parameters λ and k. To test a relatively broad parameter spectrum, we generate TreeAge sequences with $\lambda = 0.001, 1, 50$ and 1000 and $k = 0.1, 1,$ and 20, in all possible combinations. Results are summarized in Tables 3a to 3p. **We conclude that the implementation of the Weibull distribution in TreeAge is valid and delivers the expected output, but care needs to be given to output when using extreme parameter settings. These extreme settings are, however, unlikely to be used in practice.**

Table 3: Evaluating TreeAge Pro's Weibull Random Number Generator, using an input sequence of $1.5\cdot 10^8$ pseudo Weibull random numbers generated with TreeAge.

a. Distribution: Weibull with $\lambda = 0.001, k = 0.1$

Method to Evaluate TreeAge Sequence	*Result*
Moments	Failed*
K-S test	Passed
Die Hard battery	Passed**

** At these values of λ and k the resulting values are not likely to be in a practical range of TreeAge Applications (mean =* $3.6\cdot 10^{36}$ *and standard deviation =* $1.6\cdot 10^{39}$*) The resulting empirical mean was less than 0.5% off from theoretical mean but the standard deviation and higher moments were off by more than 42%. This is due to numerical limitations for representing large numbers in computers.*
***One out of seventeen Die Hard tests indicated a "weak" result, with a P-value for the K-S test larger than 0.9999. Again, this indicates that the agreement is almost too good, see footnote 8.*

b. Distribution: Weibull with $\lambda = 0.001, k = 1$

Method to Evaluate TreeAge Sequence	*Result*
Moments	Passed
K-S test	Passed
Die Hard battery	Passed

c. Distribution: Weibull with $\lambda = 0.001, k = 20$

Method to Evaluate TreeAge Sequence	*Result*
Moments	Passed
K-S test	Passed
Die Hard battery	Passed

d. Distribution: Weibull with $\lambda = 0.001, k = 100$

Method to Evaluate TreeAge Sequence	*Result*
Moments	Passed
K-S test	Passed
Die Hard battery	Passed

e. Distribution: Weibull with $\lambda = 1, k = 0.1$

Method to Evaluate TreeAge Sequence	*Result*
Moments	Failed*
K-S test	Passed
Die Hard battery	Passed

** For values of k ~ 0.1 the empirical mean was less than 0.2% off from the theoretical mean, but the standard deviation, skewness and kurtosis were off by more than 32%. See note in table 3a (Calculated mean =* $3.6\cdot 10^{6}$ *and standard deviation =* $1.6\cdot 10^{9}$*).*

f. Distribution: Weibull with $\lambda = 1, k = 1$

Method to Evaluate TreeAge Sequence	*Result*
Moments	Passed
K-S test	Passed
Die Hard battery	Passed

g. Distribution: Weibull with $\lambda = 1, k = 20$

Method to Evaluate TreeAge Sequence	*Result*
Moments	Passed
K-S test	Passed
Die Hard battery	Passed*

**Two out of seventeen Die Hard test indicated a weak result. While one is associated with a large P-value, the other is associated with a small P-value and therefore raises a flag and we will investigate this further. This also could be a Type I error. Since all other Die Harder tests did not signal a problem, we preliminary declared an overall passing grade.*

h. Distribution: Weibull with $\lambda = 1, k = 100$

Method to Evaluate TreeAge Sequence	*Result*
Moments	Passed
K-S test	Passed
Die Hard battery	Passed

i. Distribution: Weibull with $\lambda = 50, k = 0.1$

Method to Evaluate TreeAge Sequence	*Result*
Moments	Failed*
K-S test	Passed
Die Hard battery	Passed**

** The empirical mean was less than 3% off from the theoretical mean, but standard deviation was off by 37%, skewness by 96% and kurtosis by 100%. See note in table 3a (Calculated mean =* $3.7 \cdot 10^{-11}$ *and standard deviation =* $1.6 \cdot 10^{-8}$*).*

***One out of seventeen Die Hard test indicated Weak result. Since it is associated with a small P-value, it raises a flag and we will investigate this further. Since all other Die Harder tests did not signal a problem, we preliminary declared an overall passing grade.*

j. Distribution: Weibull with $\lambda = 50, k = 1$

Method to Evaluate TreeAge Sequence	*Result*
Moments	Passed
K-S test	Passed
Die Hard battery	Passed

k. Distribution: Weibull with $\lambda = 50, k = 20$

Method to Evaluate TreeAge Sequence	*Result*
Moments	Passed
K-S test	Passed
Die Hard battery	Passed*

**One out of seventeen Die Hard test indicated Weak result. Since it is associated with a small P-value, it raises a flag and we will investigate this further. Since all other Die Harder tests did not signal a problem, we preliminary declared an overall passing grade.*

l. Distribution: Weibull with $\lambda = 50, k = 100$

Method to Evaluate TreeAge Sequence	*Result*
Moments	Passed
K-S test	Passed
Die Hard battery	Passed*

**One out of seventeen Die Hard test indicated Weak result. Since it is associated with a small P-value, it raises a flag and we will investigate this further. Since all other Die Harder tests did not signal a problem, we preliminary declared an overall passing grade.*

m. Distribution: Weibull with $\lambda = 1000, k = 0.1$

Method to Evaluate TreeAge Sequence	*Result*
Moments	Failed*
K-S test	Passed
Die Hard battery	Passed

** The empirical mean was less than 3.1% off from the theoretical mean, but the standard deviation was off by 34%, skewness by 95% and kurtosis by 100%. See note in table 3a (Calculated mean =* $3.6 \cdot 10^{-24}$ *and standard deviation =* $1.6 \cdot 10^{-21}$*).*

n. Distribution: Weibull with $\lambda = 1000, k = 1$

Method to Evaluate TreeAge Sequence	*Result*
Moments	Passed
K-S test	Passed
Die Hard battery	Passed

o. Distribution: Weibull with $\lambda = 1000, k = 20$

Method to Evaluate TreeAge Sequence	*Result*
Moments	Passed
K-S test	Passed
Die Hard battery	Passed

p. Distribution: Weibull with $\lambda = 1000, k = 100$

Method to Evaluate TreeAge Sequence	*Result*
Moments	Passed
K-S test	Passed
Die Hard battery	Passed

c. The Normal Distribution

The normal distribution as documented in TreeAge follows the standard notation. We ran the same checks as with the exponential and Weibull distribution. However, the cdf of the normal distribution is not available in closed form, but the numerical approximation implemented in the statistical software package R is very accurate, so we used it to turn the TreeAge sequence into a sequence of independent and identically distributed (i.i.d.) uniform variables for the Die Hard battery of tests. Alternatively, we made use of the fact that if X_1 and X_2 are i.i.d. Normal with mean μ and standard deviation σ, then

$$exp\left[-\frac{1}{2}\left(\frac{(X_1-\mu)^2}{\sigma^2}+\frac{(X_2-\mu)^2}{\sigma^2}\right)\right]$$

follows a uniform distribution on [0,1]. Hence, we can turn the sequence $X_1, X_2, X_3, \ldots$ of normal pseudo random variables from TreeAge into a sample (of half the length) of i.i.d uniform random numbers on [0,1] with the above transformation applied to successive pairs X_i and $X_{i+1.}$ This avoids having to compute the cdf of a normal random variable. For the results displayed in Tables 4a to 4d below, we used the cdf method, but we also got the same results when using the transformation above.

We deemed it sufficient to test the TreeAge output of pseudo normal random variables with mean $\mu = 0$ and variance $\sigma^2 = 1$, (the standard normal distribution), but just to make sure we also tested cases with $\mu = -4000, \sigma^2 = 200$, $\mu = 200, \sigma^2 = 80$, and $\mu = 100{,}000, \sigma^2 = 1000$. **We conclude that the implementation of the normal distribution in TreeAge is valid and delivers the expected output.**

Table 4: Evaluating TreeAge Pro's Normal Random Number Generator, using an input sequence of $1.5 \cdot 10^8$ pseudo normal random numbers generated with TreeAge.

a. Distribution: Normal with $\mu = 0, \sigma^2 = 1$

Method to Evaluate TreeAge Sequence	*Result*
Moments	Passed
K-S test	Passed
Die Hard battery	Passed

b. Distribution: Normal with $\mu = -4000, \sigma^2 = 200$

Method to Evaluate TreeAge Sequence	*Result*
Moments	Passed
K-S test	Passed
Die Hard battery	Passed

c. Distribution: Normal with $\mu = 200, \sigma^2 = 80$

Method to Evaluate TreeAge Sequence	*Result*
Moments	Passed
K-S test	Passed
Die Hard battery	Passed

d. Distribution: Normal with $\mu = 100,000, \sigma^2 = 1000$

Method to Evaluate TreeAge Sequence	*Result*
Moments	Passed
K-S test	Passed
Die Hard battery	Passed

d. The Gamma Distribution

From the help file, TreeAge parameterizes the density (pdf) of the Gamma distribution as

$$f(x;\lambda,\alpha) = \frac{\lambda^\alpha x^{(\alpha-1)}}{\Gamma(\alpha)} e^{-\lambda x}, \qquad x > 0, \qquad \lambda, \alpha > 0,$$

with expected value equal to $\mathrm{E[X]} = \frac{\alpha}{\lambda}$. To check if a sequence from the Gamma distribution generated with TreeAge has all the probabilistic and statistical properties of a true Gamma sequence, a sequence of length $1.5 \cdot 10^8$ was generated in TreeAge and the Moments and K-S checks performed, for a variety of values for the parameters λ and α. **We conclude that the implementation of the gamma distribution in TreeAge is valid and delivers the expected output**. However, since we yet did not finish subjecting the TreeAge output to the Die Harder battery of tests, these results are preliminary.

Table 5: Evaluating TreeAge Pro's Gamma Random Number Generator, using an input sequence of $1.5 \cdot 10^8$ pseudo gamma random numbers generated with TreeAge.

a. Distribution: Gamma with $\alpha = 0.01, \lambda = 0.01$

Method to Evaluate TreeAge Sequence	*Result*
Moments	Passed
K-S test	Failed*

** For small values of α such as 0.01, the resulting values are not likely to be in a practical range of TreeAge Applications (mean =* $9.3 \cdot 10^{155}$ *and standard deviation =* **not able to calculate***). For example for $\alpha = 0.015, \lambda = 0.01$ (mean =* $8.7 \cdot 10^{91}$ *and standard deviation =* $2.8 \cdot 10^{111}$*). Please see results further in this table for α values 0.02 etc.*

b. Distribution: Gamma with $\alpha = 1.5, \lambda = 0.01$

Method to Evaluate TreeAge Sequence	*Result*
Moments	Passed
K-S test	Passed

c. Distribution: Gamma with $\alpha = 20.5, \lambda = 0.01$

Method to Evaluate TreeAge Sequence	*Result*
Moments	Passed
K-S test	Passed

d. Distribution: Gamma with $\alpha = 0.01, \lambda = 1$

Method to Evaluate TreeAge Sequence	*Result*
Moments	Passed
K-S test	Failed*

** For small values of α such as 0.01, numerical inaccuracies of representing numbers in a computer may have caused problems. See note above in table 5a.*

e. Distribution: Gamma with $\alpha = 1.5, \lambda = 1$

Method to Evaluate TreeAge Sequence	*Result*
Moments	Passed
K-S test	Passed

f. Distribution: Gamma with $\alpha = 20.5, \lambda = 1$

Method to Evaluate TreeAge Sequence	*Result*
Moments	Passed
K-S test	Passed

g. Distribution: Gamma with $\alpha = 0.01.5, \lambda = 50$

Method to Evaluate TreeAge Sequence	*Result*
Moments	Passed
K-S test	Failed*

** For small values of α such as 0.01, numerical inaccuracies of representing numbers in a computer may have caused problems. See note above in table 5a.*

h. Distribution: Gamma with $\alpha = 1.5, \lambda = 50$

Method to Evaluate TreeAge Sequence	*Result*
Moments	Passed
K-S test	Passed

i. Distribution: Gamma with $\alpha = 20.5, \lambda = 50$

Method to Evaluate TreeAge Sequence	*Result*
Moments	Passed
K-S test	Passed

j. Distribution: Gamma with $\alpha = 0.01, \lambda = 1000$

Method to Evaluate TreeAge Sequence	*Result*
Moments	Passed
K-S test	Failed*

** For small values of α such as 0.01, numerical inaccuracies of representing numbers in a computer may have caused problems. See note above in table 5a.*

k. Distribution: Gamma with $\alpha = 1.5, \lambda = 1000$

Method to Evaluate TreeAge Sequence	*Result*
Moments	Passed
K-S test	Passed

l. Distribution: Gamma with $\alpha = 20.5, \lambda = 1000$

Method to Evaluate TreeAge Sequence	*Result*
Moments	Passed
K-S test	Passed

m. Distribution: Gamma with $\alpha = 0.02, \lambda = 0.01$

Method to Evaluate TreeAge Sequence	*Result*
Moments	Passed
K-S test	Passed

n. Distribution: Gamma with $\alpha = 0.02, \lambda = 1$

Method to Evaluate TreeAge Sequence	*Result*
Moments	Passed
K-S test	Passed

o. Distribution: Gamma with $\boldsymbol{\alpha = 0.02, \lambda = 50}$

Method to Evaluate TreeAge Sequence	*Result*
Moments	Passed
K-S test	Passed

p. Distribution: Gamma with $\boldsymbol{\alpha = 0.02, \lambda = 1000}$

Method to Evaluate TreeAge Sequence	*Result*
Moments	Passed
K-S test	Passed

p. Distribution: Gamma with $\boldsymbol{\alpha = 1, \lambda = 0.01}$

Method to Evaluate TreeAge Sequence	*Result*
Moments	Passed
K-S test	Passed

q. Distribution: Gamma with $\boldsymbol{\alpha = 10, \lambda = 0.1}$

Method to Evaluate TreeAge Sequence	*Result*
Moments	Passed
K-S test	Passed

r. Distribution: Gamma with $\boldsymbol{\alpha = 1000, \lambda = 10}$

Method to Evaluate TreeAge Sequence	*Result*
Moments	Passed
K-S test	Passed

e. The Rayleigh Distribution

From the help file, TreeAge parameterizes the density (pdf) of the Rayleigh distribution as

$$f(x;\alpha) = \frac{x}{\alpha^2} e^{\left(\frac{-x^2}{2\alpha^2}\right)}, \qquad x, \alpha > 0$$

To check if a sequence from the Rayleigh distribution generated with TreeAge has all the probabilistic and statistical properties of a true Rayleigh sequence, a sequence of length $1.5 \cdot 10^8$ was generated in TreeAge and the Moments, K-S and Die Hard checks performed, for different of values for the parameter α. **We conclude that the implementation of the Rayleigh distribution in TreeAge is valid and delivers the expected output.**

Table 6: Evaluating TreeAge Pro's Normal Random Number Generator, using an input sequence of $1.5 \cdot 10^8$ pseudo normal random numbers generated with TreeAge.

a. Distribution: Rayleigh with $\alpha = 0.1$

Method to Evaluate TreeAge Sequence	*Result*
K-S test	Passed
Moments	Passed
Die Hard	Passed

b. Distribution: Rayleigh with $\alpha = 0.09$

Method to Evaluate TreeAge Sequence	*Result*
K-S test	Passed
Moments	Passed
Die Hard	Passed

c. Distribution: Rayleigh with $\alpha = 0.08$

Method to Evaluate TreeAge Sequence	*Result*
K-S test	Passed
Moments	Passed
Die Hard	Passed

d. Distribution: Rayleigh with $\alpha = 0.05$

Method to Evaluate TreeAge Sequence	*Result*
K-S test	Passed
Moments	Passed
Die Hard	Passed

e. Distribution: Rayleigh with $\alpha = 0.01$

Method to Evaluate TreeAge Sequence	*Result*
K-S test	Passed
Moments	Passed
Die Hard	Passed

f. Distribution: Rayleigh with $\alpha = 1$

Method to Evaluate TreeAge Sequence	*Result*
K-S test	Passed
Moments	Passed
Die Hard	Passed*

**Two out of seventeen Die Hard tests indicated Weak result. Both are associated with large P-values and therefore not worrisome.*

g. Distribution: Rayleigh with $\alpha = 13$

Method to Evaluate TreeAge Sequence	*Result*
K-S test	Passed
Moments	Passed
Die Hard	Passed

h. Distribution: Rayleigh with $\alpha = 50$

Method to Evaluate TreeAge Sequence	*Result*
K-S test	Passed
Moments	Passed
Die Hard	Passed

h. Distribution: Rayleigh with $\alpha = 1000$

Method to Evaluate TreeAge Sequence	*Result*
K-S test	Passed
Moments	Passed
Die Hard	Passed*

**One out of seventeen Die Hard tests indicated Weak result. Since it is associated with a large P-value it is not worrisome.*

f. The Erlang Distribution

From the help file, TreeAge parameterizes the density (pdf) of the Erlang distribution as

$$f(x;\ \lambda, k) = \frac{(k\lambda)^k x^{k-1}}{(k-1)!} e^{-k\lambda x}, \qquad x, \alpha > 0, k = 1,2, \dots n\ integer$$

To check if a sequence from the Erlang distribution generated with TreeAge has all the probabilistic and statistical properties of a true Erlang sequence, a sequence of length $1.5 \cdot 10^8$ was generated in TreeAge and the above mentioned checks performed, for a variety of values for the parameters λ and α. **We conclude that the implementation of the Erlang distribution in TreeAge is valid and delivers the expected output.**

Table 7: Evaluating TreeAge Pro's Erlang Random Number Generator, using an input sequence of $1.5 \cdot 10^8$ pseudo Erlang random numbers generated with TreeAge.

a. Distribution: Erlang with $\lambda = 0.01\ and\ k = 1$

Method to Evaluate TreeAge Sequence	*Result*
K-S test	Passed
Moments	Passed
Die Hard	Passed

b. Distribution: Erlang with $\lambda = 0.01$ and $k = 2$

Method to Evaluate TreeAge Sequence	*Result*
K-S test	Passed
Moments	Passed
Die Hard	Passed*

**Two out of seventeen Die Hard tests indicated Weak result. Both are associated with large P-values and therefore not worrisome.*

c. Distribution: Erlang with $\lambda = 0.01$ and $k = 10$

Method to Evaluate TreeAge Sequence	*Result*
K-S test	Passed
Moments	Passed
Die Hard	Passed

d. Distribution: Erlang with $\lambda = 1$ and $k = 1$

Method to Evaluate TreeAge Sequence	*Result*
K-S test	Passed
Moments	Passed
Die Hard	Passed*

**One out of seventeen Die Hard tests indicated Weak result. Since it is associated with a large P-value it is not worrisome.*

e. Distribution: Erlang with $\lambda = 1$ and $k = 3$

Method to Evaluate TreeAge Sequence	*Result*
K-S test	Passed
Moments	Passed
Die Hard	Passed

f. Distribution: Erlang with $\lambda = 1$ and $k = 20$

Method to Evaluate TreeAge Sequence	*Result*
K-S test	Passed
Moments	Passed
Die Hard	Passed

g. Distribution: Erlang with $\lambda = 50$ and $k = 1$

Method to Evaluate TreeAge Sequence	*Result*
K-S test	Passed
Moments	Passed
Die Hard	Passed

h. Distribution: Erlang with $\lambda = 50$ and $k = 4$

Method to Evaluate TreeAge Sequence	*Result*
K-S test	Passed
Moments	Passed
Die Hard	Passed

i. Distribution: Erlang with $\lambda = 50$ and $k = 4$

Method to Evaluate TreeAge Sequence	*Result*
K-S test	Passed
Moments	Passed
Die Hard	Passed*

**One out of seventeen Die Hard tests indicated Weak result. Since it is associated with a large P-value it is not worrisome.*

i. Distribution: Erlang with $\lambda = 50$ and $k = 20$

Method to Evaluate TreeAge Sequence	*Result*
K-S test	Passed
Moments	Passed
Die Hard	Passed

j. Distribution: Erlang with $\lambda = 1000$ and $k = 1$

Method to Evaluate TreeAge Sequence	*Result*
K-S test	Passed
Moments	Passed
Die Hard	Passed*

**One out of seventeen Die Hard tests indicated Weak result. Since it is associated with a large P-value it is not worrisome.*

k. Distribution: Erlang with $\lambda = 1000$ and $k = 5$

Method to Evaluate TreeAge Sequence	*Result*
K-S test	Passed
Moments	Passed
Die Hard	Passed

l. Distribution: Erlang with $\lambda = 1000$ and $k = 20$

Method to Evaluate TreeAge Sequence	*Result*
K-S test	Passed
Moments	Passed
Die Hard	Passed

3.) Validity of implementations of *discrete distributions* in TreeAge Pro

For checking the validity of discrete distributions, modifications are needed because of the discrete nature. While we can still use the moments test, modifications have to be made for carrying out the K-S test. Also, a different strategy has to be selected to use the Die Hard battery of tests, because the probability integral transform which was used to create the Die Hard input files does not hold for discrete random variables. We will treat validating TreeAge's generation of binary samples first, followed by samples from the Binomial, Poisson and discrete uniform distribution.

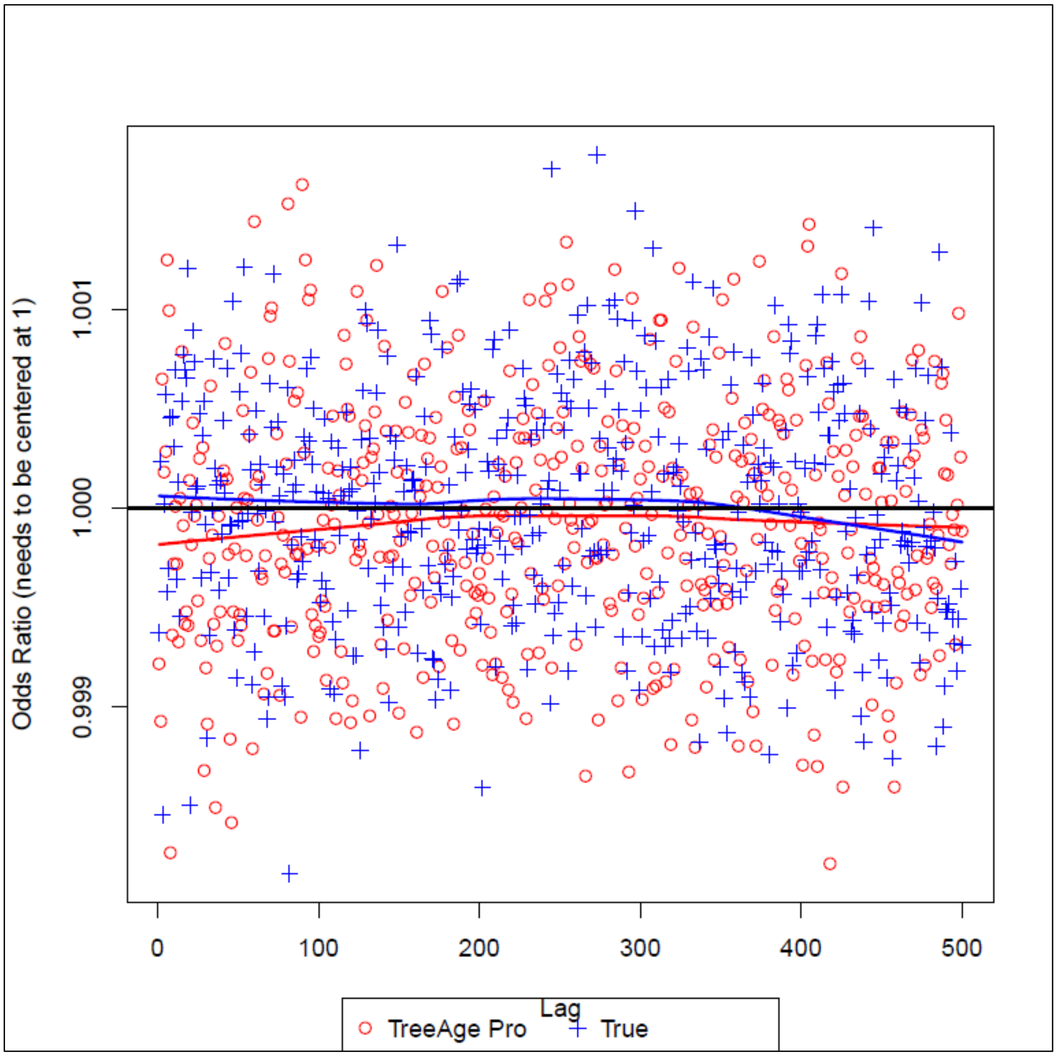
a. The Bernoulli Distribution

TreeAge Pro generates Bernoulli sequence using Binomial distribution with parameter N=1 and desired probability *p*. To test if a Bernoulli sequence generated in TreeAge with success probability *p* actually follows this distribution, we generate, in the statistical program R, a complementary Bernoulli sequence with success probability *1-p*. We then mix the two sequences to create a new sequence, so that each element of the new sequence has 50% chance of being from the original or the complementary sequence. In theory, this new sequence must follow a Bernoulli sequence with success probability 50%. We then transform every consecutive subsequence of length 32 into a 32-bit integer. If the original TreeAge sequence is indeed Bernoulli with success probability p, the derived sequence of 32-bit integers must be i.i.d. uniform on the numbers from 0 to 2^{32} - 1. This can be tested with an overall K-S test and the Die Hard battery of tests.

We also developed a plot that takes the binary sequence generated in TreeAge and computes the empirical odds ratio at the first 1000 lags. The odds ratio (OR) at lag *h* for a binary sequence $X_1, X_2, X_3, \dots$ is defined as

$$OR_h = \frac{\text{Prob}(X_i = 1, X_{i+h} = 1) \, * \, \text{Prob}(X_i = 0, X_{i+h} = 0)}{\text{Prob}(X_i = 1, X_{i+h} = 0) * \, \text{Prob}(X_i = 0, X_{i+h} = 1)}$$

If the TreeAge sequence truly is i.i.d. Bernoulli with success probability *p*, the odds ratio at each lag *h* should be equal to 1. In the plot, we plot the empirical odds ratio versus the theoretical ones. If the plot in any way systematically deviates from the horizontal line at 1, this raises a flag. Additionally, we overlay the plot with the OR at lag *h* computed from a sequence of equal size, generated with the same probability *p* in the statistical software program R. If there is any structural break (such as a non-constant success probability) in the TreeAge sequence, it should deviate from the one generated with R and a flag is raised. In theory, both the OR plot for the TreeAge sequence and the one from the one generated with R should have a similar shape and spread.



Odds Ratio Plot for *p=0.5*

This OR plots shows agreement between the TreeAge generated binary sequence and the theoretical properties of such a sequence.

Summing up, we employ the following four checks on a Bernoulli sequence generated with TreeAge:

i.) Compare the first four empirical central and non-central moments of the original TreeAge sequence with probability *p* to the theoretical moments of the Bernoulli distribution with probability *p*,

ii.) run a K-S test on the derived sequence, which must be i.i.d uniform on the numbers 0 to 2^{32} - 1,

iii.) visually inspect the OR plot for deviations from 1 and agreement with an equal sequence generated with R and

iv.) run the Die Hard battery of tests on the derived sequence.

These checks were performed on Bernoulli sequences of length $32 \cdot 150 \cdot 10^6$ (so that the Bernoulli sequence results in 150 million uniform numbers) generated in TreeAge with success probability p = 0.001, 0.03, 0.1, 0.2, 0.3, 0.5, 0.6 and 0.9999.
We again validated that if the distribution in TreeAge is generated with slightly different parameter values, or if a TreeAge sequence does not have constant probability of success, this would immediately be flagged in at least one of the checks. We also found that the OR plot may not be sensitive to certain departures from an i.i.d Bernoulli sequence, but that these departures are flagged in at least one of the other tests.
Results for applying the checks are given in Table 8a to 8h. **We conclude that the implementation of the Bernoulli distribution in TreeAge is valid and delivers the expected output.**

Table 8: Evaluating TreeAge Pro's Bernoulli Random Number Generator, using an input sequence of $32 \cdot 150 \cdot 10^6$ pseudo Bernoulli random numbers generated with TreeAge.

a. Distribution: Bernoulli with $p = 0.001$

Method to Evaluate TreeAge Sequence	*Result*
Moments	Passed
K-S test	Passed
OR plot	Passed
Die Hard battery	Passed

b. Distribution: Bernoulli with $p = 0.03$

Method to Evaluate TreeAge Sequence	*Result*
Moments	Passed
K-S test	Passed
OR plot	Passed
Die Hard battery	Passed

c. Distribution: Bernoulli with $p = 0.1$

Method to Evaluate TreeAge Sequence	*Result*
Moments	Passed
K-S test	Passed
OR plot	Passed
Die Hard battery	Passed

d. Distribution: Bernoulli with $p = 0.2$

Method to Evaluate TreeAge Sequence	*Result*
Moments	Passed
K-S test	Passed
OR plot	Passed
Die Hard battery	Passed

e. Distribution: Bernoulli with $p = 0.3$

Method to Evaluate TreeAge Sequence	*Result*
Moments	Passed
K-S test	Passed
OR plot	Passed
Die Hard battery	Passed

f. Distribution: Bernoulli with $p = 0.5$

Method to Evaluate TreeAge Sequence	*Result*
Moments	Passed
K-S test	Passed
OR plot	Passed
Die Hard battery	Passed

g. Distribution: Bernoulli with $p = 0.6$

Method to Evaluate TreeAge Sequence	*Result*
Moments	Passed
K-S test	Passed
OR plot	Passed
Die Hard battery	Passed

h. Distribution: Bernoulli with $p = 0.9999$

Method to Evaluate TreeAge Sequence	*Result*
Moments	Passed
K-S test	Passed
OR plot	Passed
Die Hard battery	Passed

b. The Binomial Distribution

To evaluate the binomial distribution with size *n* and success probability *p* we again compare empirical and theoretical central and non-central moments. We also convert each generated binomial number x into a random sequence of x successes (i.e. ones) and n-x failures (i.e. zeros). This creates a new binary sequence with success probability p. We can then proceed along the lines of testing the validity of the binary sequence. In particular, we can create the complementary sequence, mix the two sequences together and transform them to 32-bit integers. If the TreeAge sequence is indeed binomial with success probability p, the derived sequence of 32-bit integers needs to be uniform on the numbers 0 to $2^{32} - 1$. This is again tested using the Die Hard battery of tests.

A limitation in the previous approach is that very long sequences may arise when n is moderately large, and this can lead to memory problems. So in addition to the above test based on Die Harder, we devised a test not based on it. For this new test, we split the input sequence from TreeAge into distinct blocks of length 10,000 (this can be varied). For each block, we obtain the frequency table of the numbers 0, 1, 2, … ,n (i.e., we tabulate how often each of these numbers occur in a block of length 10,000). We then compare these frequencies to the theoretical binomial frequencies, using four different measures:

i.) The Pearson goodness of fit statistic
ii.) The Likelihood Ratio statistic
iii.) The Freeman-Tukey statistic
iv.) The Power Divergence statistic with $\lambda = 2$

Each of these statistics is a special case of the Power Divergence Statistic [9] and provides a different measure of how the observed frequencies deviate from the ones expected under the theoretical binomial distribution in each block. To judge the significance of each of the four statistics, we used a permutation approach which provides a P-value for each statistic in each block. If the pseudo random sample generated with the Binomial distribution in TreeAge is valid, for a given statistic, the distribution of these p-values across the blocks should be uniform on the interval [0,1], which we test with a final K-S test. If the P-value of that final K-S test is larger than 0.0125 for each of the four measures, we declared that the sequence passed the test, if not we declared it failed.

Results for applying the above Die Hard battery of tests and the Power Divergence tests (for larger n, only the Power Divergence test were carried out due to memory issues) to sequences of binomial pseudo random variables generated with TreeAge are given in Tables 9a to 9p. We considered several possible parameter combinations of $n = 5, 10, 50 \text{ and } 100$ and $p = 0.01, 0.1, 0.2, 0.5, 0.9 \text{ and } 0.999$.

Table 9: Evaluating TreeAge Pro's Binomial Random Number Generator, using an input sequence of $2.5 \cdot 10^6$ (for the Power Divergence tests) pseudo binomial random numbers generated with TreeAge.

a. Distribution: Binomial with $\boldsymbol{n = 5, p = 0.01}$

Method to Evaluate TreeAge Sequence	*Result*
Moments	Passed
Power Divergence	Passed
K-S Test	Passed
Die Hard	Passed
Q-Q Plot	Passed

b. Distribution: Binomial with $\boldsymbol{n = 5, p = 0.2}$

Method to Evaluate TreeAge Sequence	*Result*
Moments	Passed
Power Divergence	Passed
K-S Test	Passed
Die Hard	Passed
Q-Q Plot	Passed

9 Timothy Read and Noel Cressie, "Goodness-of-Fit Statistics for Discrete Multivariate Data", Springer, New-York, 1988.

c. Distribution: Binomial with $n = 5, p = 0.5$

Method to Evaluate TreeAge Sequence	*Result*
Moments	Passed
Power Divergence	Passed
K-S Test	Passed
Die Hard	Passed
Q-Q Plot	Passed

d. Distribution: Binomial with $n = 5, p = 0.9$

Method to Evaluate TreeAge Sequence	*Result*
Moments	Passed
Power Divergence	Passed
K-S Test	Passed
Die Hard	Passed
Q-Q Plot	Passed

e. Distribution: Binomial with $n = 10, p = 0.01$

Method to Evaluate TreeAge Sequence	*Result*
Moments	Passed
Power Divergence	Passed
K-S Test	Passed
Q-Q Plot	Passed

f. Distribution: Binomial with $n = 10, p = 0.1$

Method to Evaluate TreeAge Sequence	*Result*
Moments	Passed
Power Divergence	Passed
K-S Test	Passed
Q-Q Plot	Passed

g. Distribution: Binomial with $n = 10, p = 0.3$

Method to Evaluate TreeAge Sequence	*Result*
Moments	Passed
Power Divergence	Passed
K-S Test	Passed
Q-Q Plot	Passed

h. Distribution: Binomial with $n = 10, p = 0.5$

Method to Evaluate TreeAge Sequence	*Result*
Moments	Passed
Power Divergence	Passed
K-S Test	Passed
Q-Q Plot	Passed

i. Distribution: Binomial with $n = 10, p = 0.8$

Method to Evaluate TreeAge Sequence	*Result*
Moments	Passed
Power Divergence	Passed
K-S Test	Passed
Q-Q Plot	Passed

j. Distribution: Binomial with $n = 10, p = 0.999$

Method to Evaluate TreeAge Sequence	*Result*
Moments	Passed
Power Divergence	Passed
K-S Test	Passed
Q-Q Plot	Passed

k. Distribution: Binomial with $n = 10, p = 0.999$

Method to Evaluate TreeAge Sequence	*Result*
Moments	Passed
Power Divergence	Passed
K-S Test	Passed
Q-Q Plot	Passed

l. Distribution: Binomial with $n = 100, p = 0.001$

Method to Evaluate TreeAge Sequence	*Result*
Moments	Passed
Power Divergence	Passed
K-S Test	Passed
Q-Q Plot	Passed

m. Distribution: Binomial with $n = 100, p = 0.01$

Method to Evaluate TreeAge Sequence	*Result*
Moments	Passed
Power Divergence	Passed
K-S Test	Passed
Q-Q Plot	Passed

n. Distribution: Binomial with $n = 100, p = 0.1$

Method to Evaluate TreeAge Sequence	*Result*
Moments	Passed
Power Divergence	Passed
K-S Test	Passed
Q-Q Plot	Passed

o. Distribution: Binomial with $n = 100, p = 0.3$

Method to Evaluate TreeAge Sequence	*Result*
Moments	Passed
Power Divergence	Passed
K-S Test	Passed
Q-Q Plot	Passed

p. Distribution: Binomial with $n = 100, p = 0.5$

Method to Evaluate TreeAge Sequence	*Result*
Moments	Passed
Power Divergence	Passed
K-S Test	Passed
Q-Q Plot	Passed

c. The Poisson Distribution

For the Poisson distribution with parameter λ, we again compared empirical and theoretical moments. We also use a similar approach to the Power Divergence tests for the Binomial sequence: The Poisson TreeAge input sequence was split into blocks of length 10,000 and for each block the observed frequencies for the integers from *l* to *u* were obtained. Here, *l* and *u* are the lowest and highest number observed in the entire input sequence. Within each block of 10,000 numbers, we then computed the 4 different Power Divergence statistics which are different measures comparing the observed frequencies to the ones expected under a Poisson distribution in each block. Using a permutation approach, we obtain a P-value for each of the four Power Divergence statistics (i.e. judging their significance), for each block. If the input sequence is truly Poisson with parameter λ, for a given Power Divergence statistic, the P-values across the blocks should follow a uniform distribution on the interval [0,1], which we test with a final K-S test. If the P-value of this K-S test is larger than 0.0125 for all four Power Divergence statistics, we declare the sequence as passed, otherwise the sequence fails.

Results from running the above tests on TreeAge input files generated with a Poisson distribution of $\lambda = 0.5, 1, 3, 5, 10, 50$ and 100 is shown in Table 10a to 10h. **We conclude that the implementation of the Poisson distribution in TreeAge is valid and delivers the expected output.**

Table 10: Evaluating TreeAge Pro's Poisson Random Number Generator, using an input sequence of $1.5 \cdot 10^7$ pseudo Poisson random numbers generated with TreeAge.

a. Distribution: Poisson with $\boldsymbol{\lambda = 0.2}$

Method to Evaluate TreeAge Sequence	*Result*
Moments	Passed
Power Divergence	Passed

b. Distribution: Poisson with $\boldsymbol{\lambda = 0.5}$

Method to Evaluate TreeAge Sequence	*Result*
Moments	Passed
Power Divergence	Passed

c. Distribution: Poisson with $\boldsymbol{\lambda = 1}$

Method to Evaluate TreeAge Sequence	*Result*
Moments	Passed
Power Divergence	Passed

d. Distribution: Poisson with $\boldsymbol{\lambda = 3}$

Method to Evaluate TreeAge Sequence	*Result*
Moments	Passed
Power Divergence	Passed

e. Distribution: Poisson with $\boldsymbol{\lambda = 5}$

Method to Evaluate TreeAge Sequence	*Result*
Moments	Passed
Power Divergence	Passed

f. Distribution: Poisson with $\boldsymbol{\lambda = 10}$

Method to Evaluate TreeAge Sequence	*Result*
Moments	Passed
Power Divergence	Passed

g. Distribution: Poisson with $\lambda = 50$

Method to Evaluate TreeAge Sequence	*Result*
Moments	Passed
Power Divergence	Passed

h. Distribution: Poisson with $\lambda = 100$

Method to Evaluate TreeAge Sequence	*Result*
Moments	Passed
Power Divergence	Passed

d. The Discrete Uniform Distribution

For the discrete uniform distribution with parameters *low* and *high*, a similar strategy was selected as for the Poisson distribution. We first compared empirical and theoretical moments. In addition, we split the input sequence from TreeAge into distinct blocks of length 10,000 and compute, for each block, the observed frequencies for the integers from *low* to *high.* Within each block we then computed the 4 different Power Divergence statistics which are different measures comparing the observed frequencies to the ones expected under the discrete uniform distribution. Using a permutation approach, we obtain a P-value for each of the four Power Divergence statistics, in each block. If the input sequence is truly discrete uniform with parameters *low* and *high*, for a given Power Divergence statistic, the P-values across the blocks should follow a uniform distribution on the interval [0,1], which we test with a final K-S test. If the P-value of this K-S test is larger than 0.0125 for all four Power Divergence statistics, we declare the sequence as passed, otherwise the sequence fails.

Results from running the above tests on TreeAge input files generated with a discrete uniform distribution with (*low, high*) = (*1, 6*), (*1, 20*), (*10, 20*), *(-10, 10*), *(-23, 15*), (*500, 10,000*) and (*-20,000, 80,000*) is shown in Table 11a to 11g. **We conclude that the implementation of the discrete uniform distribution in TreeAge is valid and delivers the expected output.**

Table 11: Evaluating TreeAge Pro's Discrete Uniform Random Number Generator, using an input sequence of $2.5 \cdot 10^7$ pseudo discrete uniform random numbers generated with TreeAge.

a. Distribution: Discrete Uniform with (*low, high*) = (*1, 6*)

Method to Evaluate TreeAge Sequence	*Result*
Moments	Passed
Power Divergence	Passed

b. Distribution: Discrete Uniform with (*low, high*) = (*1, 20*)

Method to Evaluate TreeAge Sequence	*Result*
Moments	Passed
Power Divergence	Passed

c. Distribution: Discrete Uniform with (*low, high*) = (*10, 20*)

Method to Evaluate TreeAge Sequence	*Result*
Moments	Passed
Power Divergence	Passed

d. Distribution: Discrete Uniform with (*low, high*) = (-*10, 10*)

Method to Evaluate TreeAge Sequence	*Result*
Moments	Passed
Power Divergence	Passed

e. Distribution: Discrete Uniform with (*low, high*) = (-23*, 15*)

Method to Evaluate TreeAge Sequence	*Result*
Moments	Passed
Power Divergence	Passed

f. Distribution: Discrete Uniform with (*low, high*) = (500, 10,000)

Method to Evaluate TreeAge Sequence	*Result*
Moments	Passed
Power Divergence	Passed

g. Distribution: Discrete Uniform with (*low, high*) = (-20,000, 80,000)

Method to Evaluate TreeAge Sequence	*Result*
Moments	Passed
Power Divergence	Passed