Soft versus Hard: A comparison of random number generators between R, GSL and a non-deterministic generator

Dirk Eddelbuettel
dirk@eddelbuettel.com

Submitted to *Directions in Statistical Computing (DSC) 2007*
University of Auckland, February 15-16, 2007

Random number generators are critically important for simulation-based estimation and inference used throughout statistical computing. 'Good' random numbers are therefore a crucial aspect of a statistical, or quantitative, computing environment.

Extending work with the **random** package (Eddelbuettel, 2006) which provides functions access a non-deterministic random number generator (NDRNG) based on a physical source of randomness, we compare this NDRNG to the ones implemented in GNU R itself, as well as several from the GNU GSL, a well-known general-purpose scientific computing library.

Recent versions GNU R provide six different random number generators, and GNU GSL provides more than fourty. The overlap of RNGs allows for a direct comparison of implementations between R and GSL, and thus a dual-benchmark for the NDRNG. For these tests, we use the *dieharder* test suite by Brown (2006) which extends the well-known *diehard* test suite by Marsaglia.

Initial results, presented in table 1 below, show that the Mersenne-Twister, the default generator in R, performs well across a variety of tests. For comparison, the non-deterministic generator is seen as competitive with most of the deterministic (i.e. "software") generators. However, it appears to be slightly weaker than the Mersenne-Twister.

Additional tests shown in table 2 compares the non-deterministic RNG to the GSL generators also used in R. These are three different implementations of the Mersenne-Twister as well as two of the Knuth 'TAOCP' algorithm. We see that the non-deterministic RNG outperforms the Knuth algorithm (of the which the second version is seen to be surprisingly slow). A direct comparison of Mersenne-Twister implementations between R and the GSL (not shown here) suggests that further improvement may be avilable.

Open issues: possible comparison the *hotbits* NDRNG, integration of *dieharder* test suite into R, easier access of GSL RNGs from R.

References

Robert G. Brown. ***dieharder**: A Random Number Test Suite*, 2006. URL `http://www.phy.duke.edu/~rgb/General/dieharder.php`. C program archive **dieharder**, version 1.4.24.

Dirk Eddelbuettel. **random***: True random numbers using random.org*, 2006. URL `http://cran.r-project.org/src/contrib/Descriptions/random.html`. R package **random**, version 0.1.0.

Test		GNU R					
	random org	Wichmann-Hill	Marsaglia MultiCarry	Super Duper	Mersenne Twister	Knuth TAOCP	Knuth TAOCP2
RGB							
Timing (10^6 per second)	8.60	5.91	14.47	14.97	13.66	10.51	10.84
Bit Persistence	√	√	√	√	√	√	√
Bit Distribution	√	√	√	√	√	√	√
Diehard							
Birthdays test (mod.)	√	√	√	√	√	√	√
Overlapping 5-Permutations	¬	¬	¬	¬	¬	¬	¬
32x32 Binary Rank Test	¬	√	√	√	√	¬	¬
6x8 Binary Rank Test	√	√	∼	√	√	√	√
Bitstream Test	¬	∼	√	¬	∼	¬	¬
Overlapping Pairs (OPSO)	√	√	√	√	√	¬	¬
Overlapping Quadruples (OQSO)	√	√	∼	¬	√	¬	¬
DNA Test	√	√	√	√	√	¬	¬
Count the 1s (stream) (mod.)	√	√	√	√	√	¬	¬
Count the 1s (byte) (mod.)	√	√	√	√	√	¬	¬
Parking Lot Test (mod.)	√	√	√	√	∼	√	√
2d Circle Minimum Distance	√	√	√	√	√	√	√
3d Sphere Minimum Distance	√	√	¬	√	√	√	√
Squeeze Test	≈	√	√	√	√	√	√
Sums Test	∼	√	√	√	√	≈	∼
Runs Test (up)	√	√	√	√	√	√	√
Runs Test (down)	√	√	√	√	√	√	√
Craps Test (mean)	√	√	√	√	√	√	√
Craps Test (freq)	√	√	√	√	√	∼	√
Other							
Marsaglia/Tsang GCD	¬	√	√	√	√	√	√
Marsaglia/Tsang Gorilla (preli.)	√	√	√	√	√	√	√
STS Monobit Test	√	√	√	√	√	√	√
STS Runs Test	√	√	√	√	√	√	√
User Example Lagged Sums	√	√	√	√	√	√	√

Note: Version 1.4.24 of Brown's `dieharder` was used.
The √ symbol denotes a 'pass', i.e. a p-value above the 5% level.
The ≈ symbol denotes a 'weak' result as is assigned to p-value between 1% and 5%.
The ∼ symbol denotes a 'poor' result below 1%, but above 0.01% level.
The ¬ symbol denotes a test failure with a p-value below 0.01%.
All tests pass the RGB Bit Persitence for tuples sized $n = 1$ to $n = 5$, and fail for $n = 6$
with the exception of Knuth Ran2 with also passes $n = 6$ but fails $n = 6$.

Table 1: Results of **dieharder** for `random.org` and R

Test		GNU GSL				
	random org	Mersenne Twister	Mersenne Tw. 1999	Mersenne Tw. 1998	Knuth Ran	Knuth Ran2
RGB						
Timing (10^6 per second)	8.60	33.88	33.43	33.20	36.71	2.08
Bit Persistence	√	√	√	√	√	√
Bit Distribution	√	√	√	√	√	√
Diehard						
Birthdays test (mod.)	√	√	√	√	√	√
Overlapping 5-Permutations	¬	¬	¬	¬	¬	¬
32x32 Binary Rank Test	¬	√	√	√	¬	¬
6x8 Binary Rank Test	√	√	√	√	√	√
Bitstream Test	¬	≈	∼	√	¬	¬
Overlapping Pairs (OPSO)	√	√	√	√	¬	¬
Overlapping Quadruples (OQSO)	√	√	√	√	¬	¬
DNA Test	√	√	√	√	¬	¬
Count the 1s (stream) (mod.)	√	√	√	√	¬	¬
Count the 1s (byte) (mod.)	√	√	√	∼	¬	¬
Parking Lot Test (mod.)	√	√	√	∼	∼	√
2d Circle Minimum Distance	√	√	√	√	√	√
3d Sphere Minimum Distance	√	√	√	√	√	√
Squeeze Test	≈	√	√	√	√	√
Sums Test	∼	√	√	√	√	√
Runs Test (up)	√	√	√	√	√	√
Runs Test (down)	√	√	√	√	√	√
Craps Test (mean)	√	√	√	√	√	√
Craps Test (freq)	√	√	√	√	√	∼
Other						
Marsaglia/Tsang GCD	¬	√	√	√	√	√
Marsaglia/Tsang Gorilla (preli.)	√	√	√	√	√	√
STS Monobit Test	√	√	√	√	√	√
STS Runs Test	√	√	√	√	√	√
User Example Lagged Sums	√	√	√	√	√	√

Note: Version 1.4.24 of Brown's `dieharder` was used.
The √ symbol denotes a 'pass', i.e. a p-value above the 5% level.
The ≈ symbol denotes a 'weak' result as is assigned to p-value between 1% and 5%.
The ∼ symbol denotes a 'poor' result below 1%, but above 0.01% level.
The ¬ symbol denotes a test failure with a p-value below 0.01%.
All tests pass the RGB Bit Persitence for tuples sized $n = 1$ to $n = 5$, and fail for $n = 6$
with the exception of Knuth Ran2 with also passes $n = 6$ but fails $n = 6$.

Table 2: Results of **dieharder** for `random.org` and `GSL`