StatSoft®

Business White Paper

Random Number Generation in *STATISTICA*

U.S. Headquarters: StatSoft, Inc. • 2300 E. 14th St. • Tulsa, OK 74104 • USA • (918) 749-1119 • Fax: (918) 749-2217 • info@statsoft.com • www.statsoft.com

Australia: StatSoft Pacific Pty Ltd.
Brazil: StatSoft South America
Czech Republic: StatSoft Czech Rep. s.r.o.
France: StatSoft France
Germany: StatSoft GmbH
Hungary: StatSoft Hungary Ltd.
Israel: StatSoft Israel Ltd.
Italy: StatSoft Italia srl
Japan: StatSoft Japan Inc.
Korea: StatSoft Korea
Netherlands: StatSoft Benelux BV
Poland: StatSoft Polska Sp. z o. o.
Portugal: StatSoft Iberica Ltda.
Russia: StatSoft Russia
Singapore: StatSoft Singapore
S. Africa: StatSoft S. Africa (Pty) Ltd.
Spain: StatSoft Espana
Sweden: StatSoft Scandinavia AB
Taiwan: StatSoft Taiwan
UK: StatSoft Ltd.

Table of Contents

Overview

Many areas of statistical analysis, research, and simulation rely on high quality random number generators. Most programs for statistical data analysis contain a function for generating uniform random numbers. A two-part article (McCullough, 1998, 1999) which appeared in The American Statistician tested the random number generators of several programs using the so-called DIEHART suite of tests (Marsaglia, 1998). DIEHARD applies various methods of assembling and combining uniform random numbers, and then performs statistical tests that are expected to be non-significant; this suite of tests has become a standard method of evaluating the quality of uniform random number generator routines. Our Statistical Development department recently applied the DIEHARD suite of tests to the random number generator available in *STATISTICA* (Version 5.5 BASIC, and Version 6 *STATISTICA* Visual Basic), and we are happy to report that - as expected - our program passed all tests.

It should be pointed out that the DIEHARD suite of tests applies various approaches in order to detect non-randomness in the stream of (supposedly) uniform random numbers, and success (passing all tests) is by no means guaranteed. For example, in the review by McCullough (1999) SAS and SPSS failed one of the tests ("Count the Ones Test"), and S-Plus failed multiple tests (see Table 2, p. 156, of McCullough, 1999). Also, in a recent review (Altman, 2000), the SAS JMP program was found to fail 3 of the DIEHARD tests, and Altman (2000) concluded that "This argues against its [JMP's] use for any serious simulation."

So we believe that this is a very convincing, independent demonstration and documentation of the quality of *STATISTICA's* random number generator, which makes the *STATISTICA* programming environment suitable, for example, for demanding Monte Carlo simulations or analyses using advanced Bayesian methods.

The actual results of the DIEHARD tests follows.

References

Altman, M. Software review: JMP, V. 4.02. At http://lilt.ilstu.edu/gmklass/ITPnews/fall00/MAltmanf00.htm

McCullough, B. D., 1998. Assessing the reliabilty of statistical software: Part I. The American Statistician, 52, 358-366.

McCullough, B. D., 1999. Assessing the reliability of statistical software: Part II. The American Statistician, 53, 149-159.

Marsaglia, G., 1996. DIEHARD: A battery of tests of randomness. At http://stat.fsu.edu/~geo/diehard.html.

Diehard Test Results

```
NOTE: Most of the tests in DIEHARD return a p-value, which
should be uniform on [0,1) if the input file contains truly
independent random bits.   Those p-values are obtained by
p=F(X), where F is the assumed distribution of the sample
random variable X---often normal. But that assumed F is just
an asymptotic approximation, for which the fit will be worst
in the tails. Thus you should not be surprised with
occasional p-values near 0 or 1, such as .0012 or .9983.
When a bit stream really FAILS BIG, you will get p's of 0 or
1 to six or more places.  By all means, do not, as a
Statistician might, think that a p < .025 or p> .975 means
that the RNG has "failed the test at the .05 level".  Such
p's happen among the hundreds that DIEHARD produces, even
with good RNG's.  So keep in mind that " p happens".
::::::::::::::::::::::::::::::::::::::::::::::::::::::::::::::::::::
::            This is the BIRTHDAY SPACINGS TEST                 ::
:: Choose m birthdays in a year of n days.  List the spacings    ::
:: between the birthdays.  If j is the number of values that     ::
:: occur more than once in that list, then j is asymptotically   ::
:: Poisson distributed with mean m^3/(4n).  Experience shows n   ::
:: must be quite large, say n>=2^18, for comparing the results   ::
:: to the Poisson distribution with that mean.  This test uses   ::
:: n=2^24 and m=2^9,  so that the underlying distribution for j  ::
:: is taken to be Poisson with lambda=2^27/(2^26)=2.  A sample   ::
:: of 500 j's is taken, and a chi-square goodness of fit test    ::
:: provides a p value.  The first test uses bits 1-24 (counting  ::
:: from the left) from integers in the specified file.           ::
::   Then the file is closed and reopened. Next, bits 2-25 are   ::
:: used to provide birthdays, then 3-26 and so on to bits 9-32.  ::
:: Each set of bits provides a p-value, and the nine p-values    ::
:: provide a sample for a KSTEST.                                ::
::::::::::::::::::::::::::::::::::::::::::::::::::::::::::::::::::::
 BIRTHDAY SPACINGS TEST, M= 512 N=2**24 LAMBDA=  2.0000
           Results for bin2.txt
                   For a sample of size 500:     mean
           bin2.txt        using bits  1 to 24   1.954
  duplicate       number       number
  spacings      observed     expected
        0          67.       67.668
        1         144.      135.335
        2         139.      135.335
```

```
       3          86.       90.224
       4          39.       45.112
       5          15.       18.045
  6 to INF        10.        8.282
 Chisquare with  6 d.o.f. =     2.56 p-value=  .137910
 ::::::::::::::::::::::::::::::::::::::::
              For a sample of size 500:     mean
          bin2.txt        using bits  2 to 25   1.894
  duplicate       number       number
  spacings       observed     expected
       0          73.       67.668
       1         145.       135.335
       2         136.       135.335
       3          88.       90.224
       4          36.       45.112
       5          14.       18.045
  6 to INF         8.        8.282
 Chisquare with  6 d.o.f. =     3.93 p-value=  .313184
 ::::::::::::::::::::::::::::::::::::::::
              For a sample of size 500:     mean
          bin2.txt        using bits  3 to 26   1.984
  duplicate       number       number
  spacings       observed     expected
       0          74.       67.668
       1         139.       135.335
       2         124.       135.335
       3          88.       90.224
       4          49.       45.112
       5          16.       18.045
  6 to INF        10.        8.282
 Chisquare with  6 d.o.f. =     2.62 p-value=  .145118
 ::::::::::::::::::::::::::::::::::::::::
              For a sample of size 500:     mean
          bin2.txt        using bits  4 to 27   2.062
  duplicate       number       number
  spacings       observed     expected
       0          53.       67.668
       1         136.       135.335
       2         154.       135.335
       3          84.       90.224
       4          40.       45.112
       5          23.       18.045
  6 to INF        10.        8.282
 Chisquare with  6 d.o.f. =     8.48 p-value=  .795161
```

```
::::::::::::::::::::::::::::::::::::::::
                  For a sample of size 500:     mean
         bin2.txt        using bits  5 to 28   2.084
 duplicate       number       number
 spacings       observed     expected
       0          64.       67.668
       1         117.       135.335
       2         147.       135.335
       3          96.       90.224
       4          47.       45.112
       5          19.       18.045
 6 to INF         10.        8.282
Chisquare with  6 d.o.f. =     4.54 p-value=  .396545
::::::::::::::::::::::::::::::::::::::::
                  For a sample of size 500:     mean
         bin2.txt        using bits  6 to 29   1.954
 duplicate       number       number
 spacings       observed     expected
       0          72.       67.668
       1         140.       135.335
       2         133.       135.335
       3          85.       90.224
       4          43.       45.112
       5          19.       18.045
 6 to INF          8.       8.282
Chisquare with  6 d.o.f. =      .94 p-value=  .012211
::::::::::::::::::::::::::::::::::::::::
                  For a sample of size 500:     mean
         bin2.txt        using bits  7 to 30   1.858
 duplicate       number       number
 spacings       observed     expected
       0          70.       67.668
       1         150.       135.335
       2         136.       135.335
       3          95.       90.224
       4          29.       45.112
       5          15.       18.045
 6 to INF          5.       8.282
Chisquare with  6 d.o.f. =     9.49 p-value=  .852364
::::::::::::::::::::::::::::::::::::::::
                  For a sample of size 500:     mean
         bin2.txt        using bits  8 to 31   1.970
 duplicate       number       number
 spacings       observed     expected
```

```
     0        61.      67.668
     1       147.     135.335
     2       131.     135.335
     3        95.      90.224
     4        44.      45.112
     5        17.      18.045
 6 to INF          5.       8.282
Chisquare with  6 d.o.f. =     3.44 p-value=  .248396
 :::::::::::::::::::::::::::::::::::::::
                For a sample of size 500:     mean
         bin2.txt        using bits  9 to 32   2.000
 duplicate       number       number
 spacings       observed     expected
     0        64.      67.668
     1       140.     135.335
     2       128.     135.335
     3        99.      90.224
     4        47.      45.112
     5        15.      18.045
 6 to INF          7.       8.282
Chisquare with  6 d.o.f. =     2.40 p-value=  .120735
 :::::::::::::::::::::::::::::::::::::::
  The 9 p-values were
       .137910   .313184   .145118   .795161   .396545
       .012211   .852364   .248396   .120735
 A KSTEST for the 9 p-values yields  .891928

$$$$$$$$$$$$$$$$$$$$$$$$$$$$$$$$$$$$$$$$$$$$$$$$$$$$$$$$$$$$$$$

    ::::::::::::::::::::::::::::::::::::::::::::::::::::::::::::::::
    ::            THE OVERLAPPING 5-PERMUTATION TEST                 ::
    :: This is the OPERM5 test.  It looks at a sequence of one mill- ::
    :: ion 32-bit random integers.  Each set of five consecutive     ::
    :: integers can be in one of 120 states, for the 5! possible or- ::
    :: derings of five numbers.  Thus the 5th, 6th, 7th,...numbers   ::
    :: each provide a state. As many thousands of state transitions  ::
    :: are observed,  cumulative counts are made of the number of    ::
    :: occurences of each state.  Then the quadratic form in the     ::
    :: weak inverse of the 120x120 covariance matrix yields a test   ::
    :: equivalent to the likelihood ratio test that the 120 cell     ::
    :: counts came from the specified (asymptotically) normal dis-   ::
    :: tribution with the specified 120x120 covariance matrix (with  ::
    :: rank 99).  This version uses 1,000,000 integers, twice.       ::
    ::::::::::::::::::::::::::::::::::::::::::::::::::::::::::::::::
```

```
          OPERM5 test for file bin2.txt
    For a sample of 1,000,000 consecutive 5-tuples,
chisquare for 99 degrees of freedom= 87.771; p-value= .216861
          OPERM5 test for file bin2.txt
    For a sample of 1,000,000 consecutive 5-tuples,
chisquare for 99 degrees of freedom=106.111; p-value= .705785
    ::::::::::::::::::::::::::::::::::::::::::::::::::::::::::::::
    :: This is the BINARY RANK TEST for 31x31 matrices. The leftmost ::
    :: 31 bits of 31 random integers from the test sequence are used ::
    :: to form a 31x31 binary matrix over the field {0,1}. The rank  ::
    :: is determined. That rank can be from 0 to 31, but ranks< 28   ::
    :: are rare, and their counts are pooled with those for rank 28. ::
    :: Ranks are found for 40,000 such random matrices and a chisqua-::
    :: re test is performed on counts for ranks 31,30,29 and <=28.   ::
    ::::::::::::::::::::::::::::::::::::::::::::::::::::::::::::::
   Binary rank test for bin2.txt
        Rank test for 31x31 binary matrices:
       rows from leftmost 31 bits of each 32-bit integer
      rank   observed  expected (o-e)^2/e  sum
        28       235     211.4  2.630382    2.630
        29      5131    5134.0   .001765    2.632
        30     23205   23103.0   .449916    3.082
        31     11429   11551.5  1.299589    4.382
  chisquare= 4.382 for 3 d. of f.; p-value= .795071
--------------------------------------------------------------
    ::::::::::::::::::::::::::::::::::::::::::::::::::::::::::::::
    :: This is the BINARY RANK TEST for 32x32 matrices. A random 32x ::
    :: 32 binary matrix is formed, each row a 32-bit random integer. ::
    :: The rank is determined. That rank can be from 0 to 32, ranks  ::
    :: less than 29 are rare, and their counts are pooled with those ::
    :: for rank 29.  Ranks are found for 40,000 such random matrices ::
    :: and a chisquare test is performed on counts for ranks  32,31, ::
    :: 30 and <=29.                                                  ::
    ::::::::::::::::::::::::::::::::::::::::::::::::::::::::::::::
   Binary rank test for bin2.txt
        Rank test for 32x32 binary matrices:
       rows from leftmost 32 bits of each 32-bit integer
      rank   observed  expected (o-e)^2/e  sum
        29       227     211.4  1.148427    1.148
        30      5213    5134.0  1.215303    2.364
        31     23066   23103.0   .059406    2.423
        32     11494   11551.5   .286461    2.710
  chisquare= 2.710 for 3 d. of f.; p-value= .611275
--------------------------------------------------------------
```

```
$$$$$$$$$$$$$$$$$$$$$$$$$$$$$$$$$$$$$$$$$$$$$$$$$$$$$$$$$$$$$$$$$$$$$$$$$$$$

    ::::::::::::::::::::::::::::::::::::::::::::::::::::::::::::::::::::
    :: This is the BINARY RANK TEST for 6x8 matrices.  From each of  ::
    :: six random 32-bit integers from the generator under test, a   ::
    :: specified byte is chosen, and the resulting six bytes form a  ::
    :: 6x8 binary matrix whose rank is determined.  That rank can be ::
    :: from 0 to 6, but ranks 0,1,2,3 are rare; their counts are     ::
    :: pooled with those for rank 4. Ranks are found for 100,000     ::
    :: random matrices, and a chi-square test is performed on        ::
    :: counts for ranks 6,5 and <=4.                                 ::
    ::::::::::::::::::::::::::::::::::::::::::::::::::::::::::::::::::::
       Binary Rank Test for bin2.txt
      Rank of a 6x8 binary matrix,
    rows formed from eight bits of the RNG bin2.txt
    b-rank test for bits  1 to  8
               OBSERVED   EXPECTED     (O-E)^2/E      SUM
       r<=4         986       944.3       1.841       1.841
       r =5       21692     21743.9        .124       1.965
       r =6       77322     77311.8        .001       1.967
                  p=1-exp(-SUM/2)= .62592
      Rank of a 6x8 binary matrix,
    rows formed from eight bits of the RNG bin2.txt
    b-rank test for bits  2 to  9
               OBSERVED   EXPECTED     (O-E)^2/E      SUM
       r<=4         962       944.3        .332        .332
       r =5       21666     21743.9        .279        .611
       r =6       77372     77311.8        .047        .658
                  p=1-exp(-SUM/2)= .28024
      Rank of a 6x8 binary matrix,
    rows formed from eight bits of the RNG bin2.txt
    b-rank test for bits  3 to 10
               OBSERVED   EXPECTED     (O-E)^2/E      SUM
       r<=4         919       944.3        .678        .678
       r =5       21945     21743.9       1.860       2.538
       r =6       77136     77311.8        .400       2.938
                  p=1-exp(-SUM/2)= .76979
      Rank of a 6x8 binary matrix,
    rows formed from eight bits of the RNG bin2.txt
    b-rank test for bits  4 to 11
               OBSERVED   EXPECTED     (O-E)^2/E      SUM
       r<=4         923       944.3        .481        .481
       r =5       21762     21743.9        .015        .496
```

```
    r =6       77315    77311.8       .000       .496
                p=1-exp(-SUM/2)= .21953
  Rank of a 6x8 binary matrix,
rows formed from eight bits of the RNG bin2.txt
b-rank test for bits  5 to 12
              OBSERVED  EXPECTED    (O-E)^2/E     SUM
    r<=4         980       944.3      1.350      1.350
    r =5       21961    21743.9       2.168       3.517
    r =6       77059    77311.8        .827       4.344
                p=1-exp(-SUM/2)= .88604
  Rank of a 6x8 binary matrix,
rows formed from eight bits of the RNG bin2.txt
b-rank test for bits  6 to 13
              OBSERVED  EXPECTED    (O-E)^2/E     SUM
    r<=4         947       944.3       .008       .008
    r =5       21808    21743.9        .189       .197
    r =6       77245    77311.8        .058       .254
                p=1-exp(-SUM/2)= .11944
  Rank of a 6x8 binary matrix,
rows formed from eight bits of the RNG bin2.txt
b-rank test for bits  7 to 14
              OBSERVED  EXPECTED    (O-E)^2/E     SUM
    r<=4         869       944.3      6.005      6.005
    r =5       21874    21743.9        .778       6.783
    r =6       77257    77311.8        .039       6.822
                p=1-exp(-SUM/2)= .96699
  Rank of a 6x8 binary matrix,
rows formed from eight bits of the RNG bin2.txt
b-rank test for bits  8 to 15
              OBSERVED  EXPECTED    (O-E)^2/E     SUM
    r<=4         937       944.3       .056       .056
    r =5       21710    21743.9        .053       .109
    r =6       77353    77311.8        .022       .131
                p=1-exp(-SUM/2)= .06352
  Rank of a 6x8 binary matrix,
rows formed from eight bits of the RNG bin2.txt
b-rank test for bits  9 to 16
              OBSERVED  EXPECTED    (O-E)^2/E     SUM
    r<=4         975       944.3       .998       .998
    r =5       21615    21743.9        .764       1.762
    r =6       77410    77311.8        .125       1.887
                p=1-exp(-SUM/2)= .61071
  Rank of a 6x8 binary matrix,
rows formed from eight bits of the RNG bin2.txt
```

b-rank test for bits 10 to 17

	OBSERVED	EXPECTED	(O-E)^2/E	SUM
r<=4	930	944.3	.217	.217
r =5	21622	21743.9	.683	.900
r =6	77448	77311.8	.240	1.140

p=1-exp(-SUM/2)= .43445

Rank of a 6x8 binary matrix,
rows formed from eight bits of the RNG bin2.txt
b-rank test for bits 11 to 18

	OBSERVED	EXPECTED	(O-E)^2/E	SUM
r<=4	975	944.3	.998	.998
r =5	21634	21743.9	.555	1.553
r =6	77391	77311.8	.081	1.635

p=1-exp(-SUM/2)= .55838

Rank of a 6x8 binary matrix,
rows formed from eight bits of the RNG bin2.txt
b-rank test for bits 12 to 19

	OBSERVED	EXPECTED	(O-E)^2/E	SUM
r<=4	988	944.3	2.022	2.022
r =5	21534	21743.9	2.026	4.048
r =6	77478	77311.8	.357	4.406

p=1-exp(-SUM/2)= .88951

Rank of a 6x8 binary matrix,
rows formed from eight bits of the RNG bin2.txt
b-rank test for bits 13 to 20

	OBSERVED	EXPECTED	(O-E)^2/E	SUM
r<=4	916	944.3	.848	.848
r =5	21884	21743.9	.903	1.751
r =6	77200	77311.8	.162	1.913

p=1-exp(-SUM/2)= .61568

Rank of a 6x8 binary matrix,
rows formed from eight bits of the RNG bin2.txt
b-rank test for bits 14 to 21

	OBSERVED	EXPECTED	(O-E)^2/E	SUM
r<=4	1007	944.3	4.163	4.163
r =5	21988	21743.9	2.740	6.903
r =6	77005	77311.8	1.218	8.121

p=1-exp(-SUM/2)= .98276

Rank of a 6x8 binary matrix,
rows formed from eight bits of the RNG bin2.txt
b-rank test for bits 15 to 22

	OBSERVED	EXPECTED	(O-E)^2/E	SUM
r<=4	967	944.3	.546	.546
r =5	21962	21743.9	2.188	2.733

r =6	77071	77311.8	.750	3.483

p=1-exp(-SUM/2)= .82477

Rank of a 6x8 binary matrix,
rows formed from eight bits of the RNG bin2.txt
b-rank test for bits 16 to 23

	OBSERVED	EXPECTED	(O-E)^2/E	SUM
r<=4	947	944.3	.008	.008
r =5	21852	21743.9	.537	.545
r =6	77201	77311.8	.159	.704

p=1-exp(-SUM/2)= .29670

Rank of a 6x8 binary matrix,
rows formed from eight bits of the RNG bin2.txt
b-rank test for bits 17 to 24

	OBSERVED	EXPECTED	(O-E)^2/E	SUM
r<=4	968	944.3	.595	.595
r =5	21830	21743.9	.341	.936
r =6	77202	77311.8	.156	1.092

p=1-exp(-SUM/2)= .42063

Rank of a 6x8 binary matrix,
rows formed from eight bits of the RNG bin2.txt
b-rank test for bits 18 to 25

	OBSERVED	EXPECTED	(O-E)^2/E	SUM
r<=4	935	944.3	.092	.092
r =5	21982	21743.9	2.607	2.699
r =6	77083	77311.8	.677	3.376

p=1-exp(-SUM/2)= .81511

Rank of a 6x8 binary matrix,
rows formed from eight bits of the RNG bin2.txt
b-rank test for bits 19 to 26

	OBSERVED	EXPECTED	(O-E)^2/E	SUM
r<=4	912	944.3	1.105	1.105
r =5	21668	21743.9	.265	1.370
r =6	77420	77311.8	.151	1.521

p=1-exp(-SUM/2)= .53263

Rank of a 6x8 binary matrix,
rows formed from eight bits of the RNG bin2.txt
b-rank test for bits 20 to 27

	OBSERVED	EXPECTED	(O-E)^2/E	SUM
r<=4	922	944.3	.527	.527
r =5	21530	21743.9	2.104	2.631
r =6	77548	77311.8	.722	3.352

p=1-exp(-SUM/2)= .81292

Rank of a 6x8 binary matrix,
rows formed from eight bits of the RNG bin2.txt

b-rank test for bits 21 to 28

	OBSERVED	EXPECTED	(O-E)^2/E	SUM
r<=4	925	944.3	.395	.395
r =5	21713	21743.9	.044	.438
r =6	77362	77311.8	.033	.471

p=1-exp(-SUM/2)= .20983

Rank of a 6x8 binary matrix,
rows formed from eight bits of the RNG bin2.txt
b-rank test for bits 22 to 29

	OBSERVED	EXPECTED	(O-E)^2/E	SUM
r<=4	949	944.3	.023	.023
r =5	21611	21743.9	.812	.836
r =6	77440	77311.8	.213	1.048

p=1-exp(-SUM/2)= .40793

Rank of a 6x8 binary matrix,
rows formed from eight bits of the RNG bin2.txt
b-rank test for bits 23 to 30

	OBSERVED	EXPECTED	(O-E)^2/E	SUM
r<=4	956	944.3	.145	.145
r =5	21665	21743.9	.286	.431
r =6	77379	77311.8	.058	.490

p=1-exp(-SUM/2)= .21715

Rank of a 6x8 binary matrix,
rows formed from eight bits of the RNG bin2.txt
b-rank test for bits 24 to 31

	OBSERVED	EXPECTED	(O-E)^2/E	SUM
r<=4	941	944.3	.012	.012
r =5	21738	21743.9	.002	.013
r =6	77321	77311.8	.001	.014

p=1-exp(-SUM/2)= .00709

Rank of a 6x8 binary matrix,
rows formed from eight bits of the RNG bin2.txt
b-rank test for bits 25 to 32

	OBSERVED	EXPECTED	(O-E)^2/E	SUM
r<=4	933	944.3	.135	.135
r =5	21741	21743.9	.000	.136
r =6	77326	77311.8	.003	.138

p=1-exp(-SUM/2)= .06679

TEST SUMMARY, 25 tests on 100,000 random 6x8 matrices
These should be 25 uniform [0,1] random variables:

.625920	.280241	.769795	.219525	.886041
.119442	.966992	.063521	.610709	.434450
.558376	.889513	.615684	.982758	.824768
.296697	.420634	.815111	.532631	.812924

```
   .209830     .407926     .217153     .007093     .066788
 brank test summary for bin2.txt
     The KS test for those 25 supposed UNI's yields
                    KS p-value= .012692

$$$$$$$$$$$$$$$$$$$$$$$$$$$$$$$$$$$$$$$$$$$$$$$$$$$$$$$$$$$$$$$$$

   :::::::::::::::::::::::::::::::::::::::::::::::::::::::::::::::::
   ::                  THE BITSTREAM TEST                           ::
   :: The file under test is viewed as a stream of bits. Call them  ::
   :: b1,b2,... .  Consider an alphabet with two "letters", 0 and 1 ::
   :: and think of the stream of bits as a succession of 20-letter  ::
   :: "words", overlapping.  Thus the first word is b1b2...b20, the ::
   :: second is b2b3...b21, and so on.  The bitstream test counts   ::
   :: the number of missing 20-letter (20-bit) words in a string of ::
   :: 2^21 overlapping 20-letter words.  There are 2^20 possible 20 ::
   :: letter words.  For a truly random string of 2^21+19 bits, the ::
   :: number of missing words j should be (very close to) normally  ::
   :: distributed with mean 141,909 and sigma 428.  Thus            ::
   ::  (j-141909)/428 should be a standard normal variate (z score) ::
   :: that leads to a uniform [0,1) p value.  The test is repeated  ::
   :: twenty times.                                                 ::
   :::::::::::::::::::::::::::::::::::::::::::::::::::::::::::::::::
 THE OVERLAPPING 20-tuples BITSTREAM  TEST, 20 BITS PER WORD, N words
   This test uses N=2^21 and samples the bitstream 20 times.
  No. missing words should average  141909. with sigma=428.
---------------------------------------------------------
 tst no  1:  141023 missing words,   -2.07 sigmas from mean, p-value= .01919
 tst no  2:  141873 missing words,    -.08 sigmas from mean, p-value= .46618
 tst no  3:  141525 missing words,    -.90 sigmas from mean, p-value= .18460
 tst no  4:  141676 missing words,    -.55 sigmas from mean, p-value= .29282
 tst no  5:  142017 missing words,     .25 sigmas from mean, p-value= .59931
 tst no  6:  141931 missing words,     .05 sigmas from mean, p-value= .52019
 tst no  7:  142477 missing words,    1.33 sigmas from mean, p-value= .90764
 tst no  8:  141820 missing words,    -.21 sigmas from mean, p-value= .41734
 tst no  9:  142136 missing words,     .53 sigmas from mean, p-value= .70181
 tst no 10:  141667 missing words,    -.57 sigmas from mean, p-value= .28563
 tst no 11:  141974 missing words,     .15 sigmas from mean, p-value= .56005
 tst no 12:  141761 missing words,    -.35 sigmas from mean, p-value= .36446
 tst no 13:  141773 missing words,    -.32 sigmas from mean, p-value= .37504
 tst no 14:  141828 missing words,    -.19 sigmas from mean, p-value= .42465
 tst no 15:  141846 missing words,    -.15 sigmas from mean, p-value= .44119
 tst no 16:  141545 missing words,    -.85 sigmas from mean, p-value= .19732
 tst no 17:  142610 missing words,    1.64 sigmas from mean, p-value= .94919
```

```
tst no 18:  141442 missing words,   -1.09 sigmas from mean, p-value= .13744
tst no 19:  141578 missing words,    -.77 sigmas from mean, p-value= .21943
tst no 20:  141299 missing words,   -1.43 sigmas from mean, p-value= .07693

$$$$$$$$$$$$$$$$$$$$$$$$$$$$$$$$$$$$$$$$$$$$$$$$$$$$$$$$$$$$

    ::::::::::::::::::::::::::::::::::::::::::::::::::::::::::::::::::
    ::            The tests OPSO, OQSO and DNA                    ::
    ::        OPSO means Overlapping-Pairs-Sparse-Occupancy        ::
    :: The OPSO test considers 2-letter words from an alphabet of    ::
    :: 1024 letters.  Each letter is determined by a specified ten   ::
    :: bits from a 32-bit integer in the sequence to be tested. OPSO ::
    :: generates  2^21 (overlapping) 2-letter words  (from 2^21+1    ::
    :: "keystrokes")  and counts the number of missing words---that  ::
    :: is 2-letter words which do not appear in the entire sequence. ::
    :: That count should be very close to normally distributed with  ::
    :: mean 141,909, sigma 290. Thus (missingwrds-141909)/290 should ::
    :: be a standard normal variable. The OPSO test takes 32 bits at ::
    :: a time from the test file and uses a designated set of ten    ::
    :: consecutive bits. It then restarts the file for the next de-  ::
    :: signated 10 bits, and so on.                                 ::
    ::                                                             ::
    ::    OQSO means Overlapping-Quadruples-Sparse-Occupancy        ::
    ::   The test OQSO is similar, except that it considers 4-letter ::
    :: words from an alphabet of 32 letters, each letter determined  ::
    :: by a designated string of 5 consecutive bits from the test    ::
    :: file, elements of which are assumed 32-bit random integers.   ::
    :: The mean number of missing words in a sequence of 2^21 four-  ::
    :: letter words,  (2^21+3 "keystrokes"), is again 141909, with   ::
    :: sigma = 295.  The mean is based on theory; sigma comes from   ::
    :: extensive simulation.                                        ::
    ::                                                             ::
    ::    The DNA test considers an alphabet of 4 letters::  C,G,A,T,::
    :: determined by two designated bits in the sequence of random   ::
    :: integers being tested.  It considers 10-letter words, so that ::
    :: as in OPSO and OQSO, there are 2^20 possible words, and the   ::
    :: mean number of missing words from a string of 2^21  (over-    ::
    :: lapping)  10-letter  words (2^21+9 "keystrokes") is 141909.   ::
    :: The standard deviation sigma=339 was determined as for OQSO   ::
    :: by simulation.  (Sigma for OPSO, 290, is the true value (to   ::
    :: three places), not determined by simulation.                 ::
    ::::::::::::::::::::::::::::::::::::::::::::::::::::::::::::::::::
OPSO test for generator bin2.txt
 Output: No. missing words (mw), equiv normal variate (z), p-value (p)
```

		mw	z	p
OPSO for bin2.txt	using bits 23 to 32	142614	2.430	.9924
OPSO for bin2.txt	using bits 22 to 31	141787	-.422	.3366
OPSO for bin2.txt	using bits 21 to 30	141739	-.587	.2785
OPSO for bin2.txt	using bits 20 to 29	141865	-.153	.4393
OPSO for bin2.txt	using bits 19 to 28	142254	1.189	.8827
OPSO for bin2.txt	using bits 18 to 27	142214	1.051	.8533
OPSO for bin2.txt	using bits 17 to 26	142219	1.068	.8572
OPSO for bin2.txt	using bits 16 to 25	142144	.809	.7908
OPSO for bin2.txt	using bits 15 to 24	141837	-.249	.4015
OPSO for bin2.txt	using bits 14 to 23	142199	.999	.8411
OPSO for bin2.txt	using bits 13 to 22	142124	.740	.7704
OPSO for bin2.txt	using bits 12 to 21	141454	-1.570	.0582
OPSO for bin2.txt	using bits 11 to 20	141940	.106	.5421
OPSO for bin2.txt	using bits 10 to 19	142167	.889	.8129
OPSO for bin2.txt	using bits 9 to 18	142065	.537	.7043
OPSO for bin2.txt	using bits 8 to 17	141922	.044	.5174
OPSO for bin2.txt	using bits 7 to 16	141925	.054	.5215
OPSO for bin2.txt	using bits 6 to 15	141867	-.146	.4420
OPSO for bin2.txt	using bits 5 to 14	141560	-1.205	.1142
OPSO for bin2.txt	using bits 4 to 13	141716	-.667	.2525
OPSO for bin2.txt	using bits 3 to 12	141926	.057	.5229
OPSO for bin2.txt	using bits 2 to 11	142025	.399	.6550
OPSO for bin2.txt	using bits 1 to 10	141464	-1.536	.0623

OQSO test for generator bin2.txt

Output: No. missing words (mw), equiv normal variate (z), p-value (p)

		mw	z	p
OQSO for bin2.txt	using bits 28 to 32	141962	.179	.5709
OQSO for bin2.txt	using bits 27 to 31	141568	-1.157	.1236
OQSO for bin2.txt	using bits 26 to 30	142119	.711	.7614
OQSO for bin2.txt	using bits 25 to 29	142002	.314	.6233
OQSO for bin2.txt	using bits 24 to 28	141941	.107	.5427
OQSO for bin2.txt	using bits 23 to 27	142288	1.284	.9004
OQSO for bin2.txt	using bits 22 to 26	141959	.168	.5669
OQSO for bin2.txt	using bits 21 to 25	141152	-2.567	.0051
OQSO for bin2.txt	using bits 20 to 24	141497	-1.398	.0811
OQSO for bin2.txt	using bits 19 to 23	142223	1.063	.8562
OQSO for bin2.txt	using bits 18 to 22	141870	-.133	.4470
OQSO for bin2.txt	using bits 17 to 21	142104	.660	.7453
OQSO for bin2.txt	using bits 16 to 20	141609	-1.018	.1543
OQSO for bin2.txt	using bits 15 to 19	142298	1.318	.9062
OQSO for bin2.txt	using bits 14 to 18	142314	1.372	.9149
OQSO for bin2.txt	using bits 13 to 17	141808	-.343	.3656
OQSO for bin2.txt	using bits 12 to 16	141656	-.859	.1952

```
  OQSO for bin2.txt      using bits 11 to 15       141981   .243  .5960
  OQSO for bin2.txt      using bits 10 to 14       141801  -.367  .3567
  OQSO for bin2.txt      using bits  9 to 13       142085   .595  .7242
  OQSO for bin2.txt      using bits  8 to 12       142145   .799  .7878
  OQSO for bin2.txt      using bits  7 to 11       141509 -1.357  .0874
  OQSO for bin2.txt      using bits  6 to 10       142913  3.402  .9997
  OQSO for bin2.txt      using bits  5 to  9       142289  1.287  .9010
  OQSO for bin2.txt      using bits  4 to  8       142100   .646  .7410
  OQSO for bin2.txt      using bits  3 to  7       141711  -.672  .2507
  OQSO for bin2.txt      using bits  2 to  6       141819  -.306  .3797
  OQSO for bin2.txt      using bits  1 to  5       141351 -1.893  .0292
```

DNA test for generator bin2.txt

Output: No. missing words (mw), equiv normal variate (z), p-value (p)

```
                                                   mw      z      p
  DNA for bin2.txt       using bits 31 to 32       141695  -.632  .2636
  DNA for bin2.txt       using bits 30 to 31       141603  -.904  .1831
  DNA for bin2.txt       using bits 29 to 30       141940   .090  .5360
  DNA for bin2.txt       using bits 28 to 29       141583  -.963  .1679
  DNA for bin2.txt       using bits 27 to 28       142538  1.854  .9682
  DNA for bin2.txt       using bits 26 to 27       142105   .577  .7181
  DNA for bin2.txt       using bits 25 to 26       142143   .689  .7547
  DNA for bin2.txt       using bits 24 to 25       141115 -2.343  .0096
  DNA for bin2.txt       using bits 23 to 24       141557 -1.039  .1493
  DNA for bin2.txt       using bits 22 to 23       141425 -1.429  .0765
  DNA for bin2.txt       using bits 21 to 22       141338 -1.685  .0460
  DNA for bin2.txt       using bits 20 to 21       141694  -.635  .2627
  DNA for bin2.txt       using bits 19 to 20       141809  -.296  .3836
  DNA for bin2.txt       using bits 18 to 19       141667  -.715  .2374
  DNA for bin2.txt       using bits 17 to 18       141940   .090  .5360
  DNA for bin2.txt       using bits 16 to 17       141635  -.809  .2092
  DNA for bin2.txt       using bits 15 to 16       142054   .427  .6652
  DNA for bin2.txt       using bits 14 to 15       141861  -.143  .4433
  DNA for bin2.txt       using bits 13 to 14       141224 -2.022  .0216
  DNA for bin2.txt       using bits 12 to 13       141928   .055  .5220
  DNA for bin2.txt       using bits 11 to 12       142015   .312  .6224
  DNA for bin2.txt       using bits 10 to 11       142376  1.377  .9157
  DNA for bin2.txt       using bits  9 to 10       141542 -1.084  .1393
  DNA for bin2.txt       using bits  8 to  9       142089   .530  .7019
  DNA for bin2.txt       using bits  7 to  8       142450  1.595  .9446
  DNA for bin2.txt       using bits  6 to  7       142020   .326  .6280
  DNA for bin2.txt       using bits  5 to  6       141920   .031  .5126
  DNA for bin2.txt       using bits  4 to  5       142257  1.026  .8475
  DNA for bin2.txt       using bits  3 to  4       141615  -.868  .1926
  DNA for bin2.txt       using bits  2 to  3       142641  2.158  .9845
```

```
    DNA for bin2.txt       using bits  1 to  2       141660  -.735  .2310

$$$$$$$$$$$$$$$$$$$$$$$$$$$$$$$$$$$$$$$$$$$$$$$$$$$$$$$$$$$$$$

    ::::::::::::::::::::::::::::::::::::::::::::::::::::::::::::::::
    ::     This is the COUNT-THE-1's TEST on a stream of bytes.      ::
    :: Consider the file under test as a stream of bytes (four per   ::
    :: 32 bit integer).  Each byte can contain from 0 to 8 1's,      ::
    :: with probabilities 1,8,28,56,70,56,28,8,1 over 256.  Now let  ::
    :: the stream of bytes provide a string of overlapping  5-letter ::
    :: words, each "letter" taking values A,B,C,D,E. The letters are ::
    :: determined by the number of 1's in a byte::  0,1,or 2 yield A,::
    :: 3 yields B, 4 yields C, 5 yields D and 6,7 or 8 yield E. Thus ::
    :: we have a monkey at a typewriter hitting five keys with vari- ::
    :: ous probabilities (37,56,70,56,37 over 256).  There are 5^5   ::
    :: possible 5-letter words, and from a string of 256,000 (over-  ::
    :: lapping) 5-letter words, counts are made on the frequencies   ::
    :: for each word.   The quadratic form in the weak inverse of    ::
    :: the covariance matrix of the cell counts provides a chisquare ::
    :: test::  Q5-Q4, the difference of the naive Pearson sums of    ::
    :: (OBS-EXP)^2/EXP on counts for 5- and 4-letter cell counts.    ::
    ::::::::::::::::::::::::::::::::::::::::::::::::::::::::::::::::
   Test results for bin2.txt
 Chi-square with 5^5-5^4=2500 d.of f. for sample size:2560000
                          chisquare  equiv normal  p-value
  Results fo COUNT-THE-1's in successive bytes:
 byte stream for bin2.txt        2777.65      3.927      .999957
 byte stream for bin2.txt        3106.10      8.572     1.000000

$$$$$$$$$$$$$$$$$$$$$$$$$$$$$$$$$$$$$$$$$$$$$$$$$$$$$$$$$$$$$$

    ::::::::::::::::::::::::::::::::::::::::::::::::::::::::::::::::
    ::     This is the COUNT-THE-1's TEST for specific bytes.        ::
    :: Consider the file under test as a stream of 32-bit integers.  ::
    :: From each integer, a specific byte is chosen , say the left-  ::
    :: most::  bits 1 to 8. Each byte can contain from 0 to 8 1's,   ::
    :: with probabilitie 1,8,28,56,70,56,28,8,1 over 256.  Now let   ::
    :: the specified bytes from successive integers provide a string ::
    :: of (overlapping) 5-letter words, each "letter" taking values  ::
    :: A,B,C,D,E. The letters are determined  by the number of 1's,  ::
    :: in that byte::  0,1,or 2 ---> A, 3 ---> B, 4 ---> C, 5 ---> D,::
    :: and  6,7 or 8 ---> E.  Thus we have a monkey at a typewriter  ::
    :: hitting five keys with with various probabilities::  37,56,70,::
    :: 56,37 over 256. There are 5^5 possible 5-letter words, and    ::
```

```
    :: from a string of 256,000 (overlapping) 5-letter words, counts ::
    :: are made on the frequencies for each word. The quadratic form ::
    :: in the weak inverse of the covariance matrix of the cell      ::
    :: counts provides a chisquare test::  Q5-Q4, the difference of  ::
    :: the naive Pearson  sums of (OBS-EXP)^2/EXP on counts for 5-   ::
    :: and 4-letter cell counts.                                     ::
    :::::::::::::::::::::::::::::::::::::::::::::::::::::::::::::::::::
 Chi-square with 5^5-5^4=2500 d.of f. for sample size: 256000
                chisquare  equiv normal  p value
  Results for COUNT-THE-1's in specified bytes:
           bits  1 to  8  2447.83      -.738     .230332
           bits  2 to  9  2582.49      1.167     .878318
           bits  3 to 10  2509.70       .137     .554561
           bits  4 to 11  2565.68       .929     .823534
           bits  5 to 12  2579.26      1.121     .868825
           bits  6 to 13  2454.29      -.646     .258985
           bits  7 to 14  2505.89       .083     .533182
           bits  8 to 15  2471.59      -.402     .343948
           bits  9 to 16  2510.24       .145     .557584
           bits 10 to 17  2503.48       .049     .519625
           bits 11 to 18  2461.67      -.542     .293874
           bits 12 to 19  2407.35     -1.310     .095050
           bits 13 to 20  2528.11       .398     .654514
           bits 14 to 21  2537.19       .526     .700553
           bits 15 to 22  2485.19      -.209     .417030
           bits 16 to 23  2579.94      1.131     .870877
           bits 17 to 24  2728.11      3.226     .999372
           bits 18 to 25  2462.59      -.529     .298370
           bits 19 to 26  2447.84      -.738     .230370
           bits 20 to 27  2574.59      1.055     .854244
           bits 21 to 28  2480.23      -.280     .389913
           bits 22 to 29  2417.68     -1.164     .122189
           bits 23 to 30  2546.97       .664     .746720
           bits 24 to 31  2427.09     -1.031     .151250
           bits 25 to 32  2640.33      1.985     .976402

$$$$$$$$$$$$$$$$$$$$$$$$$$$$$$$$$$$$$$$$$$$$$$$$$$$$$$$$$$$$$$$$$

    :::::::::::::::::::::::::::::::::::::::::::::::::::::::::::::::::::
    ::              THIS IS A PARKING LOT TEST                       ::
    :: In a square of side 100, randomly "park" a car---a circle of  ::
    :: radius 1.   Then try to park a 2nd, a 3rd, and so on, each    ::
    :: time parking "by ear".  That is, if an attempt to park a car  ::
    :: causes a crash with one already parked, try again at a new    ::
```

```
:: random location. (To avoid path problems, consider parking    ::
:: helicopters rather than cars.)   Each attempt leads to either ::
:: a crash or a success, the latter followed by an increment to  ::
:: the list of cars already parked. If we plot n:  the number of ::
:: attempts, versus k::  the number successfully parked, we get a::
:: curve that should be similar to those provided by a perfect   ::
:: random number generator.  Theory for the behavior of such a   ::
:: random curve seems beyond reach, and as graphics displays are ::
:: not available for this battery of tests, a simple characteriz ::
:: ation of the random experiment is used: k, the number of cars ::
:: successfully parked after n=12,000 attempts. Simulation shows ::
:: that k should average 3523 with sigma 21.9 and is very close  ::
:: to normally distributed.  Thus (k-3523)/21.9 should be a st-  ::
:: andard normal variable, which, converted to a uniform varia-  ::
:: ble, provides input to a KSTEST based on a sample of 10.      ::
:::::::::::::::::::::::::::::::::::::::::::::::::::::::::::::::::::
     CDPARK: result of ten tests on file bin2.txt
      Of 12,000 tries, the average no. of successes
          should be 3523 with sigma=21.9
      Successes: 3504    z-score:  -.868 p-value: .192812
      Successes: 3486    z-score: -1.689 p-value: .045562
      Successes: 3530    z-score:   .320 p-value: .625377
      Successes: 3506    z-score:  -.776 p-value: .218799
      Successes: 3506    z-score:  -.776 p-value: .218799
      Successes: 3493    z-score: -1.370 p-value: .085365
      Successes: 3529    z-score:   .274 p-value: .607947
      Successes: 3533    z-score:   .457 p-value: .676028
      Successes: 3522    z-score:  -.046 p-value: .481790
      Successes: 3525    z-score:   .091 p-value: .536382

     square size   avg. no.  parked   sample sigma
       100.           3513.400       15.711
      KSTEST for the above 10: p=  .787029

$$$$$$$$$$$$$$$$$$$$$$$$$$$$$$$$$$$$$$$$$$$$$$$$$$$$$$$$$$$$$$$$$

:::::::::::::::::::::::::::::::::::::::::::::::::::::::::::::::::::
::               THE MINIMUM DISTANCE TEST                       ::
:: It does this 100 times::   choose n=8000 random points in a   ::
:: square of side 10000.  Find d, the minimum distance between   ::
:: the (n^2-n)/2 pairs of points.  If the points are truly inde- ::
:: pendent uniform, then d^2, the square of the minimum distance ::
:: should be (very close to) exponentially distributed with mean ::
:: .995 .  Thus 1-exp(-d^2/.995) should be uniform on [0,1) and  ::
```

```
:: a KSTEST on the resulting 100 values serves as a test of uni- ::
:: formity for random points in the square. Test numbers=0 mod 5 ::
:: are printed but the KSTEST is based on the full set of 100    ::
:: random choices of 8000 points in the 10000x10000 square.      ::
:::::::::::::::::::::::::::::::::::::::::::::::::::::::::::::::::::
          This is the MINIMUM DISTANCE test
         for random integers in the file bin2.txt
Sample no.    d^2     avg     equiv uni
      5     .9525   1.2528    .616061
     10     .2523   1.1304    .223983
     15    1.2425   1.1435    .713130
     20     .9199   1.1138    .603270
     25    1.5022   1.1907    .779041
     30    2.2629   1.1487    .897124
     35     .1656   1.1107    .153347
     40     .1797   1.0272    .165202
     45     .0064    .9392    .006461
     50     .2021    .9189    .183835
     55    2.5716    .9766    .924567
     60    1.3520    .9941    .743031
     65     .8384    .9872    .569434
     70     .7880    .9745    .547062
     75    2.6903    .9803    .933053
     80     .2643    .9576    .233251
     85    1.1764    .9345    .693417
     90     .0812    .9569    .078349
     95     .0128    .9594    .012796
    100     .5111    .9616    .401706
MINIMUM DISTANCE TEST for bin2.txt
     Result of KS test on 20 transformed mindist^2's:
                        p-value= .157869

$$$$$$$$$$$$$$$$$$$$$$$$$$$$$$$$$$$$$$$$$$$$$$$$$$$$$$$$$$$$$$

:::::::::::::::::::::::::::::::::::::::::::::::::::::::::::::::::::
::             THE 3DSPHERES TEST                                ::
:: Choose  4000 random points in a cube of edge 1000.  At each   ::
:: point, center a sphere large enough to reach the next closest ::
:: point. Then the volume of the smallest such sphere is (very   ::
:: close to) exponentially distributed with mean 120pi/3.  Thus  ::
:: the radius cubed is exponential with mean 30. (The mean is    ::
:: obtained by extensive simulation).  The 3DSPHERES test gener- ::
:: ates 4000 such spheres 20 times.  Each min radius cubed leads ::
:: to a uniform variable by means of 1-exp(-r^3/30.), then a     ::
```

```
    ::  KSTEST is done on the 20 p-values.                ::
    :::::::::::::::::::::::::::::::::::::::::::::::::::::::::
           The 3DSPHERES test for file bin2.txt
 sample no:  1     r^3=  27.580     p-value= .60122
 sample no:  2     r^3=  69.627     p-value= .90181
 sample no:  3     r^3=  31.036     p-value= .64460
 sample no:  4     r^3=   3.202     p-value= .10124
 sample no:  5     r^3=   8.543     p-value= .24780
 sample no:  6     r^3=  60.743     p-value= .86798
 sample no:  7     r^3=  92.601     p-value= .95435
 sample no:  8     r^3=  33.032     p-value= .66749
 sample no:  9     r^3=  35.169     p-value= .69034
 sample no: 10     r^3=  22.806     p-value= .53242
 sample no: 11     r^3=  35.717     p-value= .69595
 sample no: 12     r^3=  17.904     p-value= .44943
 sample no: 13     r^3=  41.284     p-value= .74744
 sample no: 14     r^3=   2.973     p-value= .09435
 sample no: 15     r^3=  27.739     p-value= .60332
 sample no: 16     r^3=   5.186     p-value= .15876
 sample no: 17     r^3=   7.059     p-value= .20967
 sample no: 18     r^3=  13.432     p-value= .36093
 sample no: 19     r^3=  16.929     p-value= .43125
 sample no: 20     r^3=  18.512     p-value= .46047
  A KS test is applied to those 20 p-values.
---------------------------------------------------------
      3DSPHERES test for file bin2.txt           p-value= .141646
$$$$$$$$$$$$$$$$$$$$$$$$$$$$$$$$$$$$$$$$$$$$$$$$$$$$$$$$$$$$$$$$$

    :::::::::::::::::::::::::::::::::::::::::::::::::::::::::
    ::      This is the SQEEZE test                         ::
    ::  Random integers are floated to get uniforms on [0,1). Start- ::
    ::  ing with k=2^31=2147483647, the test finds j, the number of  ::
    ::  iterations necessary to reduce k to 1, using the reduction   ::
    ::  k=ceiling(k*U), with U provided by floating integers from    ::
    ::  the file being tested.  Such j's are found 100,000 times,    ::
    ::  then counts for the number of times j was <=6,7,...,47,>=48  ::
    ::  are used to provide a chi-square test for cell frequencies. ::
    :::::::::::::::::::::::::::::::::::::::::::::::::::::::::
         RESULTS OF SQUEEZE TEST FOR bin2.txt
       Table of standardized frequency counts
    ( (obs-exp)/sqrt(exp) )^2
       for j taking values <=6,7,8,...,47,>=48:
    -.1     1.3     -.6     .2     .0     -.2
    -.9     -.4     1.9     -.8     .2     .5
```

```
  -.7    1.4    1.1    -.3    -.6    -.4
   .3    -.6     .0     .1     .3    -.7
  -.7     .4    -.4    -.9    1.3     .8
 -1.3     .6    1.4   -1.0    -.9     .6
  -.5     .5    -.8     .4    -.6   -1.0
  -.1
      Chi-square with 42 degrees of freedom: 25.755
        z-score= -1.772  p-value= .023009
```

$$

```
    :::::::::::::::::::::::::::::::::::::::::::::::::::::::::::::::
    ::            The  OVERLAPPING SUMS test                     ::
    :: Integers are floated to get a sequence U(1),U(2),... of uni-  ::
    :: form [0,1) variables.  Then overlapping sums,               ::
    ::   S(1)=U(1)+...+U(100), S2=U(2)+...+U(101),... are formed.    ::
    :: The S's are virtually normal with a certain covariance mat-   ::
    :: rix.  A linear transformation of the S's converts them to a   ::
    :: sequence of independent standard normals, which are converted ::
    :: to uniform variables for a KSTEST. The  p-values from ten     ::
    :: KSTESTs are given still another KSTEST.                     ::
    :::::::::::::::::::::::::::::::::::::::::::::::::::::::::::::::
             Test no.  1      p-value  .409725
             Test no.  2      p-value  .586482
             Test no.  3      p-value  .186793
             Test no.  4      p-value  .990710
             Test no.  5      p-value  .001333
             Test no.  6      p-value  .372953
             Test no.  7      p-value  .725250
             Test no.  8      p-value  .467133
             Test no.  9      p-value  .272107
             Test no. 10      p-value  .771743
  Results of the OSUM test for bin2.txt
      KSTEST on the above 10 p-values:  .293007
```

$$

```
    :::::::::::::::::::::::::::::::::::::::::::::::::::::::::::::::
    ::     This is the RUNS test.  It counts runs up, and runs down, ::
    :: in a sequence of uniform [0,1) variables, obtained by float-  ::
    :: ing the 32-bit integers in the specified file. This example   ::
    :: shows how runs are counted:  .123,.357,.789,.425,.224,.416,.95::
    :: contains an up-run of length 3, a down-run of length 2 and an ::
```

:: up-run of (at least) 2, depending on the next values. The ::
:: covariance matrices for the runs-up and runs-down are well ::
:: known, leading to chisquare tests for quadratic forms in the ::
:: weak inverses of the covariance matrices. Runs are counted ::
:: for sequences of length 10,000. This is done ten times. Then ::
:: repeated. ::
:::
The RUNS test for file bin2.txt
Up and down runs in a sample of 10000

Run test for bin2.txt :
runs up; ks test for 10 p's: .497116
runs down; ks test for 10 p's: .522927
Run test for bin2.txt :
runs up; ks test for 10 p's: .913122
runs down; ks test for 10 p's: .322519

$$$

:::
:: This is the CRAPS TEST. It plays 200,000 games of craps, finds::
:: the number of wins and the number of throws necessary to end ::
:: each game. The number of wins should be (very close to) a ::
:: normal with mean 200000p and variance 200000p(1-p), with ::
:: p=244/495. Throws necessary to complete the game can vary ::
:: from 1 to infinity, but counts for all>21 are lumped with 21. ::
:: A chi-square test is made on the no.-of-throws cell counts. ::
:: Each 32-bit integer from the test file provides the value for ::
:: the throw of a die, by floating to [0,1), multiplying by 6 ::
:: and taking 1 plus the integer part of the result. ::
:::
Results of craps test for bin2.txt
No. of wins: Observed Expected
98580 98585.86
98580= No. of wins, z-score= -.026 pvalue= .48955
Analysis of Throws-per-Game:
Chisq= 30.94 for 20 degrees of freedom, p= .94399

Throws	Observed	Expected	Chisq	Sum
1	66912	66666.7	.903	.903
2	37501	37654.3	.624	1.527
3	26859	26954.7	.340	1.867
4	19392	19313.5	.319	2.186
5	13746	13851.4	.802	2.989
6	9757	9943.5	3.500	6.488

7	7258	7145.0	1.786	8.275
8	5191	5139.1	.525	8.799
9	3695	3699.9	.006	8.806
10	2694	2666.3	.288	9.094
11	1869	1923.3	1.535	10.628
12	1451	1388.7	2.791	13.420
13	1004	1003.7	.000	13.420
14	681	726.1	2.806	16.226
15	549	525.8	1.020	17.246
16	410	381.2	2.184	19.430
17	264	276.5	.569	19.998
18	170	200.8	4.733	24.731
19	163	146.0	1.983	26.714
20	116	106.2	.901	27.616
21	318	287.1	3.322	30.938

SUMMARY FOR bin2.txt
p-value for no. of wins: .489546
p-value for throws/game: .943986

$$

Results of DIEHARD battery of tests sent to file die.res