Given a pseudorandom number generator, we want to measure how well numbers generated imitate the true behavior of random numbers. There is no end to the number of tests that can be designed. In general , if a generator passes half a dozen statistical tests, testing different aspects of the sequence, then we consider the generator sufficiently random.

1 Two Main Procedures

First we will discuss two main procedures (goodness-of-fit tests) that are used in many other statistical tests.

- χ^2 goodness-of-fit test
- Kolmogorov-Smirnov goodness-of-fit test

1.1 The χ^2 goodness-of-fit test

This test was developed by Pearson in a paper in 1900. Before this paper, people would simply plot experimental results graphically and claim they were correct. This paper is regarded as one of the foundations of modern statistics.

Let a random experiment have k mutually exclusive and exhaustive outcomes, say $A_1, ..., A_k$. Put $p_i = P(A_i)$ and $p_1 + ... + p_k = 1$. The experiment is repeated n times, independently, and we let Y_i the number of times the experiment results in A_i. For example, if the experiment is rolling a pair of die and if the outcome is the sum of die, then

A_i	2	3	4	5	6	7	8	9	10	11	12
p_i	1/36	1/18	1/12	1/9	5/36	1/6	5/36	1/9	1/12	1/18	1/36

and the results of an actual roll of the dice $n = 144$ times are

Y_i	2	4	10	12	22	29	21	15	14	9	6

Observe that np_i is the expected number of times the experiment would result in A_i. In this example,

np_i	4	8	12	16	20	24	20	16	12	8	4

Pearson showed that the expression

$$Q_{k-1} = \sum_{i=1}^{k} \frac{(Y_i - np_i)^2}{np_i}$$

has an approximate χ^2 distribution with $k-1$ degrees of freedom (Notation: $\chi^2(k-1)$)

Remark 1 *Recall the definition of the chi-square distribution: X has a chi-square distribution with r degrees of freedom if*

$$f(x) = \frac{1}{\Gamma(r/2)2^{r/2}} x^{r/2-1} e^{-x/2}, \ 0 \leq x < \infty$$

where $\Gamma(t) = \int_0^\infty y^{t-1} e^{-y} dy, \ t > 0.$ *The mean, variance, and moment generating function of* $\chi^2(r)$ *are:*

$$E[X] = r, \ Var(X) = 2r, \ \textit{and} \ M(t) = (1-2t)^{-r/2}, t < 1/2.$$

We use this fact to design a test as follows: Suppose we are trying to test a probability model that assigns the probability p_i to outcome A_i. We generate n independent samples to test this model, and compute Q_{k-1}, which we will call the χ^2- test statistic. If Q_{k-1} is too small or too large, the model will be suspect. For example, let X be a random variable with the $\chi^2(k-1)$ distribution, and let Q_{k-1} be the outcome of the above calculation for a specific random sample. Then, if $P(X > Q_{k-1}) < 0.01$ or $P(X < Q_{k-1}) < 0.01$, we would reject the null hypothesis H_0 that assigns the probability p_i to outcome A_i, at the 1% significance level.

Question Q_{k-1} is approximately $\chi^2(k-1)$ as n gets large. In a practical test, how large should n be?

Rule of thumb Choose n such that $np_i \geq 5$ for all i. However, the approximation is pretty good even if $np_i < 5$. In such cases, guard against a particular np_i becoming so small that the corresponding term $(Y_i - np_i)^2/np_i$ dominates the others.

1.1.1 How to use the χ^2 test with random numbers

Take a block of numbers of size n, $u_{m+1}, ..., u_{m+n}$, from the random number generator: Divide $(0,1)$ into k subintervals of the form $I_i = [\frac{i-1}{k}, \frac{i}{k}), i = 1, ..., k$. Let p_i be the probability that a number belongs to I_i. When the numbers are from a (uniform) generator, and the subintervals are of equal length as we have chosen, the probability p_i is simply $1/k$. However, we can easily consider non-uniform generators, and intervals of different length if we want to. Finally, let Y_i be the number of $u's$ that are in I_i. Now we can compute Q_{k-1} and apply the test as discussed earlier.

In general, we would apply the test to different blocks of numbers (varying m) from the sequence, and try different values for k. Knuth [1] recommends testing at lest 3 blocks, and try 4 or 5 different k values.

Rule of thumb (Knuth) If 3 different blocks are tested and if 2 blocks fail at the 5% level, reject the generator. If one block fails at the 1% level, reject the generator.

We will learn a better way of deciding whether to reject a generator based on the χ^2 test in the next section.

1.2 Kolmogorov-Smirnov Test

Consider a sample $X_1, ..., X_N$ from a continuous distribution $F(x)$. The sample cumulative distribution function (c.d.f) is defined as

$$F_N(x) = \frac{1}{N}(\text{number of } X_i \leq x) = \frac{1}{N}\sum_{i=1}^{N} 1_{(-\infty,x]}(X_i). \tag{1}$$

Observe that $1_{(-\infty,x]}(X_i)$ is a Bernoulli random variable with probability of success $p = F(x)$. Therefore $\sum_{i=1}^{N} 1_{(-\infty,x]}(X_i)$ is a binomial random variable with parameters N (number of trials) and $F(x)$ (probability of success). This can be written as $\sum_{i=1}^{N} 1_{(-\infty,x]}(X_i) \sim bin(N, F(x))$, using the standard notations from probability. Since $\sum_{i=1}^{N} 1_{(-\infty,x]}(X_i) = NF_N(x)$, we have

$$\begin{aligned} E[NF_N(x)] &= NF(x) \Rightarrow E[F_N(x)] = F(x) \\ Var(NF_N(x)) &= NF(x)(1-F(x)) \Rightarrow Var(F_N(x)) = \frac{F(x)(1-F(x))}{N} \end{aligned}$$

and

$$P\{F_N(x) = k/N\} = \binom{N}{k} F(x)^k (1-F(x))^{N-k}.$$

Conclusion 1 *For a given $x, F_N(x)$ is an unbiased estimator for $F(x)$ and $Var(F_N(x)) \to 0$ as $N \to \infty$. From Central Limit Theorem and equation 1, we also know that the distribution of $F_N(x)$ is approximately normal. Therefore*

$$F_N(x) \sim \mathcal{N}\left(F(x), F(x)(1-F(x))/N\right).$$

The previous results are pointwise results; they hold for a fixed value of x. The next theorem, often stated as the Fundamental Theorem of Statistics, tells us more.

Theorem 1 (Glivenko-Cantelli) *For any $\epsilon > 0$,*

$$\lim_{N\to\infty} P\{\sup_{-\infty<x<\infty} |F_N(x) - F(x)| > \epsilon\} = 0.$$

The Kolmogorov-Smirnov statistic is defined as $D_N = \sup_{-\infty<x<\infty} |F_N(x) - F(x)|$. We have

Theorem 2

$$\lim_{N\to\infty} P(\sqrt{N}D_N \leq x) = \underbrace{1 - 2\sum_{j=1}^{\infty}(-1)^{j-1}e^{-2j^2x^2}}_{H(x)}$$

Remark 2 *1. Numerical values for $H(x)$ are tabulated and the approximation is sufficient when $N > 35$.*

2. $H(x)$ is independent of the distribution $F(x)$!

Designing the test: Designing the Kolmogorov-Smirnov test consists of three steps:

1. Obtaining a sample: $X_1, ..., X_N$
2. State H_0, the null hypothesis: "The sample is from $F(x)$"
3. Test H_0 : Compute D_N. Reject H_0 if it is too large.

We will now discuss the computation of D_N in detail.

1.2.1 Computing D_N

We start with a sample $X_1, ..., X_N$. Next sort the sample so that $X_1 \leq X_2 \leq ... \leq X_N$ (here there is an abuse of notation since I use the same variables for both the original sample, and the sorted sample.) We have

$$D_N = \sup_{-\infty < x < \infty} |F_N(x) - F(x)|$$

where

$$F_N(x) = \begin{cases} 0 \text{ if } x < X_1 \\ k/N \text{ if } X_k \leq x < X_{k+1} \\ 1 \text{ if } x > X_N \end{cases}$$

Therefore

$$D_N = \max \left[\sup_{x < X_1} |-F(x)|, \sup_{X_1 \leq x < X_2} \left| \frac{1}{N} - F(x) \right|, ..., \sup_{x > X_N} |1 - F(x)| \right]$$

The first and last terms are

$$\begin{aligned} \sup_{x < X_1} |-F(x)| &= F(X_1) \\ \sup_{x > X_N} |1 - F(x)| &= 1 - F(X_N) \end{aligned}$$

For the other terms, observe that

$$\sup_{X_k \leq x < X_{k+1}} \left| \frac{k}{N} - F(x) \right| = \max \left(F(x_{k+1}) - \frac{k}{N}, \frac{k}{N} - F(x_k) \right); k = 1, ..., N-1$$

and thus

$$\begin{aligned} D_N &= \max \left[F(X_1), \max_{k=1,...,N-1} \left(F(x_{k+1}) - \frac{k}{N}, \frac{k}{N} - F(x_k) \right), 1 - F(X_N) \right] \quad (2) \\ &= \max \left[\underbrace{\max_{k=1,...,N} \left(\frac{k}{N} - F(x_k) \right)}_{D^+}, \underbrace{\max_{k=1,...,N} \left(F(x_k) - \frac{k-1}{N} \right)}_{D^-} \right] \\ &= \max(D^+, D^-) \end{aligned}$$

If we are testing for the uniform distribution $U(0,1)$, then the above formulas simplify by noting that $F(x) = x$.

Remark 3 *1. Kolmogorov-Smirnov (KS)-test is used when the distribution F is continuous - no jumps! The χ^2 test, on the other hand, is essentially for distributions consisting of jumps . However, we can apply the χ^2 test to continuous distributions by grouping data as we have done in Section 1.1.1.*

2. *The χ^2 test may miss local non-random behavior as a result of grouping. The KS-test makes more use of the available data.*

3. *The KS-test can be used in conjunction with the χ^2-test for a better procedure than the rule of thumb used in Section 1.1.1 for deciding whether we should reject a generator based on several χ^2-tests. Consider M blocks of size N numbers from a random number generator. Fix k, the number of groups, i.e., intervals. Compute the χ^2- statistic Q_{k-1} for each block. This gives us the values $c_1, ..., c_M$, where c_i corresponds to the χ^2- statistic computed for the ith block. Now let F_M be the sample c.d.f of the data $c_1, ..., c_M$. The theoretical distribution of the data, under the null hypothesis, is the $\chi^2(k-1)$ distribution: apply the KS-test to check this.*

2 Generalizations of KS-statistic

KS-statistic computes, $D_N = \sup_{-\infty<x<\infty} |F_N(x) - F(x)|$, which measures, in a way, the distance between functions F_N and F. There are certainly other ways to measure distance!

Cramer-von Mises Family A general statistic is defined as

$$Q = N \int_{-\infty}^{\infty} (F_N(x) - F(x))^2 \psi(x) dF(x)$$

where $\psi(x)$ is a suitable function which gives "weights" to the squared differences.

Special Cases 1. $\psi(x) \equiv 1$ gives the classical Cramer-von Mises statistic

$$\begin{aligned} W^2 &= N \int_{-\infty}^{\infty} (F_N(x) - F(x))^2 dF(x) \\ &= \sum_{i=1}^{N} \left(Z_{(i)} - \frac{2i-1}{2N} \right)^2 + \frac{1}{12N} \end{aligned}$$

where $Z_{(i)}$ are the ordered values $(Z_{(1)} \leq Z_{(2)} \leq ... \leq Z_{(N)})$ of the sequence $Z_i = F(X_i)$ and $X_1, ..., X_N$ is the sample.

2. $\psi(x) = \dfrac{1}{F(x)(1-F(x))}$ gives the popular Anderson-Darling statistic A^2.

$$A^2 = -N - \frac{1}{N}\sum_{i=1}^{N}(2i-1)(\log Z_{(i)} + \log(1 - Z_{(N+1-i)}))$$

Another equivalent expression is

$$A^2 = -N - \frac{1}{N}\sum_{i=1}^{N}\big((2i-1)\log Z_{(i)} + (2N+1-2i)\log(1-Z_{(i)})\big)$$

The values for the statistics W^2 and A^2 are tabulated. A test based on these statistics reject the null hypothesis (with significance level α) that the sample is distributed according to $F(x)$, if W^2 (or, A^2) is greater than the critical value in the upper tail given at level α. In other words, for a given α (1%, 5%, etc.), let the real number Y_α be the number such that the probability that W^2 will exceed Y_α is α. Then, if an observed value of W^2 is larger than Y_α, then the null hypothesis is rejected at significance level α.

A good reference for goodness-of-fit tests, including Anderson-Darling and others, see D'Agostino and Stephens [2].

3 Empirical Tests

In this part we will discuss several statistical tests for random number generators. You will see that each test uses one of the two goodness-of-fit tests we discussed previously.

Given $u_1, u_2, ...$, a sequence of real numbers between 0 and 1, we want to test the claim that these numbers are independent realizations of a $U(0,1)$ random variable. There are also some tests designed for integer valued sequences only. In that case, define $y_n = Floor(du_n)$, where d is a positive integer. Then $y_1, y_2, ...$, is a sequence of integers between 0 and $d-1$.If u_n's are uniformly distributed on $(0,1)$, then y_n's are uniformly distributed as integers.

Knowledge of computer arithmetic can make this transformation from u_n to y_n very efficient. On a binary computer, choose d to be 64 or 128, then y_n represents 6 or 7 most significant bits of the binary representation of u_n. Indeed, if

$$u_n = (0.a_1a_2...a_m)_2 = \frac{a_1}{2} + \frac{a_2}{2^2} + ... + \frac{a_m}{2^m}$$

then

$$y_n = Floor(64u_n) = Floor(2^5a_1 + 2^4a_2 + ... + a_6 + \frac{a_7}{2} + ...) = (a_1a_2...a_6)_2$$

4 Equidistribution (frequency) test

This is a test we have already talked about: it checks whether the empirical distribution of the numbers is indeed the uniform distribution $U(0, 1)$.

1. Apply the KS-test to the sequence u_n with $F(x) = x$ for $0 < x < 1$.
2. Apply the χ^2- test to the sequence y_n, taking k (number of categories) to be d and $p_i = 1/d$ for each category.

5 Serial test

This tests whether pairs of successive numbers are uniformly distributed. Consider the integer sequence $y_n = Floor(du_n), n = 1, 2, ...$ There are d^2 distinct values (d_1, d_2) that can be attained by the pair $(y_{2n}, y_{2n+1}), n = 1, 2, ...$ Each value should be taken by probability $1/d^2$ if uniform distribution hypothesis is true.

Apply the χ^2- test taking k (number of categories) to be d^2 and $p_i = 1/d^2$. Count the number of pairs $(y_{2n}, y_{2n+1}), n = 1, 2, ...$ that take the value (d_1, d_2) for each of the $1/d^2$ values. (This gives the number of times the experiment results in a specific outcome (d_1, d_2), which was denoted by Y in Section 1.1. Note that n should be large compared to d^2 (see the rule of thumb for χ^2 approximation), so you will probably need a larger n value than you would use in the equidistribution test (or, a smaller d value).

6 Gap test

Consider a subinterval of $(0, 1)$, say, $J = (0, 1/2)$ and consider the numbers 0.2, 0.3, 0.6, 0.7, 0.4, 0.9, 0.8, 0.7, 0.1. I will explain the meaning of "gap of length 1 (or, 2, 3, etc.)" by using this finite sequence of numbers and J as an example. Later I will give a general definition. Study the diagram below:

$$\underset{\text{gap of length 0}}{\underset{\downarrow\in J}{0.2}}\quad,\quad\underset{\text{gap of length 0}}{\underset{\downarrow\in J}{0.3}}\quad,\ \underbrace{0.6, 0.7, \underset{\downarrow\in J}{0.4}}_{\text{gap of length 2}},\underbrace{0.9, 0.8, 0.7, \underset{\downarrow\in J}{0.1}}_{\text{gap of length 3}}$$

The probability that a number falls in J is $1/2$, but for a general discussion, let's call this number p. From the diagram above, observe that the probability of having a gap of length 0 is p, the probability of having a gap of length 2 is $(1-p)^2p$, and the probability of having a gap of length 3 is $(1-p)^3p$.

In general, consider a subinterval $J = (\alpha, \beta)$. We say a gap of length r occurs whenever a subsequence

$$u_j, u_{j+1}, ..., u_{j+r}$$

such that $u_j, u_{j+1}, ..., u_{j+r-1} \notin J$ and $u_{j+r} \in J$ is observed. The probability that a gap of length i occurs is $P_i = (1-p)^ip, i = 0, 1, ..., t-1$, where $p = \beta - \alpha$. The probability that a gap of length t or more occurs is $P_t = (1-p)^t$.

Apply a χ^2- test as follows: First, pick a t value. Then, pick a value for n such that $nP_i \geq 5$, verifying the rule of thumb for the χ^2- test. The outcomes of the test are: "gap of length 0", "gap of length 1",...,"gap of length $t-1$”, and "gap of length t or more". We know the probability of each outcome occurring. Now, given a sequence of numbers $u_1, ...,$ start at the beginning of the sequence, x_1, and start going through the sequence to identify subsequences as "gap of length i”, for some i. (See the example above.) Stop when there are a total of n gaps found (observe that this algorithm may not terminate if the numbers are sufficiently nonrandom.) Count the number of gaps of length $i, 0 \leq i \leq t-1$, and the number of gaps of length t or more. These counts are the Y_i values you need in the χ^2- test (see Section 1.1).

Example 1 *Study the Mathematica file Gap Test Examples.*

7 The Maximum Test

Consider $U_1, U_2, ...,$ i.i.d. random variables from the $U(0,1)$ distribution. Let $V_1 = \max(U_1, U_2, ..., U_t), V_2 = \max(U_{t+1}, U_{t+2}, ..., U_{2t})$, etc. In other words, V_i is the maximum of a block of t uniform random variables, for some fixed t. The following theorem gives us the distribution function (c.d.f.) of the maximum values, V_i. Note that $V_1, V_2, ...,$ are i.i.d. random variables, so we write

$$P(V \leq x) = P(\max(U_1, U_2, ..., U_t) \leq x)$$

where V has the same distribution as V_i. Since the maximum of a set of numbers is less than or equal to x iff each number in the set is less than or equal to x, we get

$$P(V \leq x) = P(U_1 \leq x, ..., U_t \leq x) = P(U \leq x)^t$$

where $P(U \leq x)$ is the c.d.f. of the $U(0,1)$ random variable, i.e., $P(U \leq x) = x$. Then the above expression simplifies to

$$P(V \leq x) = x^t.$$

We can design a test based on this fact: Apply the Kolmogorov-Smirnov test to $V_1, V_2, .., V_n$ where the distribution function is $F(x) = x^t$.

8 Run Test

Given a sequence of unequal numbers, I will explain what a run-up and a run-down of length p is by giving an example: Consider the sequence

$$2,\ 7,\ 8,\ 1,\ 9,\ 6,\ 4,\ 0,\ 3,\ 11,\ 10,\ 17$$

Finding run-ups: Put a vertical line at the beginning and end of the sequence. Also put a vertical line between x_j and x_{j+1} if $x_j > x_{j+1}$. We get

$$|2\ 7\ 8|\ 1\ 9|\ 6|\ 4|\ 0\ 3\ 11|\ 10\ 17|$$

Count the numbers in each block. We have a run-up of length 3, 2, 1, 1, 3, 2.

Finding run-downs: Put a vertical line at the beginning and end of the sequence. Also put a vertical line between x_j and x_{j+1} if $x_j < x_{j+1}$. We get

$$|2\ |7|\ 8\ 1|\ 9\ 6\ 4\ 0|\ 3|\ 11\ 10|\ 17$$

We have run-downs of length 1, 1, 2, 4, 1, 2, 1.

Let $u(i)$ be the number of runs-up of length i for $i = 1, 2, ..., 5$ and let $u(6)$ be the number of runs-up of length 6 or more. Similarly define $d(i)$, the number of runs-down of length i. Then

$$u = \frac{1}{n}\sum_{i=1}^{6}\sum_{j=1}^{6}(u(i) - nb_i)(u(j) - nb_j)a_{ij}$$

has the χ^2- distribution with 6 degrees of freedom when n is large (rule of thumb: $n \geq 4000$). A similar result applies for runs-down.

A detailed discussion of this test is given by Knuth [1]. A Fortran code for the test and some numerical results is given by Grafton [3] (this paper is posted on Blackboard).

9 Collision test

This is a generalization of the serial test to higher dimensions. Consider M urns, and imagine tossing (randomly) a small number, N, of balls into M urns. We toss balls one after another. Observe that a ball falls into a specific urn with probability $1/M$. We say there is a collision if a ball falls into an urn that already has a ball.

Note that the maximum number of collisions is $N-1$; this happens when all N balls fall into one urn. Consider an example where 3 urns are occupied when you toss 5 balls. This means each urn has at least one ball, and the remaining two balls fall into one or two urns that already contained one ball. So there are 2 collisions. In general, if j urns are occupied in n tosses ($j \leq n$) then there are $n-j$ collisions. The following theorem gives a recursion for the probability that j urns are occupied in n tosses.

Theorem 3

$$\begin{aligned} P\{j \textit{ urns occupied in } n \textit{ tosses}\} &= \frac{j}{M}P\{j \textit{ urns occupied in } n-1 \textit{ tosses}\} \\ &+\frac{M-(j-1)}{M}P\{j-1 \textit{ urns occupied in } n-1 \textit{ tosses}\} \end{aligned} \quad (3)$$

Proof. *If j urns are occupied at the end of n tosses, then either j urns, or $j-1$ urns must be occupied at the end of $n-1$ tosses. If the former event occurs, i.e., j urns occupied in $n-1$ tosses, then the nth toss has to fall into one of the already occupied j urns, which occurs with probability j/M. This observation* *gives us the first probability on the right hand side of equation 3. If $j-1$ urns are occupied in $n-1$ tosses, then the nth toss has to into one of the empty urns, which occurs with probability $(M-(j-1))/M$.(The numerator is the number of empty urns, and the denominator is the total number of urns.) This observation yields the second probability on the right hand side of equation 3. The final result that shows the probability on the left equals the probability on the right follows from a conditioning argument.* ■

Let's define $p_{jn} = P\{j \text{ urns occupied in } n \text{ tosses}\}$. The above recursion with this new notation becomes

$$p_{jn} = \frac{j}{M} p_{j(n-1)} + \frac{M-(j-1)}{M} p_{(j-1)(n-1)}; j = 1, 2, ..., n; n = 1, 2, ..., N$$

with initial conditions

$$\begin{aligned} p_{11} &= 1 \text{ and } p_{j1} = 0 \text{ for all } j > 1 \\ p_{1n} &= M^{-n+1}, n > 1 \end{aligned}$$

Also note that

$$p_{jn} = P\{j \text{ urns occupied in } n \text{ tosses}\} = P\{\text{there are } n-j \text{ collisions}\} \tag{4}$$

as explained just before the theorem.

We can design a test based on this theorem. For an example, consider 100 random numbers, $u_1, ..., u_{100}$, from a uniform generator. Consider the first 3 digits of these numbers in their base 10 expansion:

$$u_i = 0.a_{i1}a_{i2}a_{i3}$$

There are 10^3 possible values for the digits (a_1, a_2, a_3) - think of these values as your urns (imagine a three dimensional integer grid, where each axes go from 0 to 9). Think of the u_i as the balls. Each one will fall into one of these urns, based on the first three digits of its decimal expansion in base 10. The idea is to count the number of collisions, and reject the generator if this number is too small or too large.

To decide if a given number of collisions is too small or too large, we need to find the percentage points of the distribution of the number of collisions. Let's abbreviate the event "there are j collisions" as "j collisions". Then we want to compute the probabilities

$$P\{j \text{ collisions}\}, j = 0, 1, ..., 10^3 - 1$$

and then compute cumulative probabilities in order to obtain some percentage points. In the example we consider, we have $M = 1000$ urns, and $N = 100$ balls. We toss all these balls, and compute the number of collisions. The probabilities we want to compute are

$$\begin{aligned} P\{0 \text{ collision}\} &= P\{N \text{ urns occupied in } N \text{ tosses}\} = p_{NN} \\ P\{1 \text{ collision}\} &= P\{N-1 \text{ urns occupied in } N \text{ tosses}\} = p_{(N-1)N} \\ &\cdots \\ P\{N-1 \text{ collisions}\} &= P\{1 \text{ urn occupied in } N \text{ tosses}\} = p_{1N} = M^{-N+1} \end{aligned}$$

where we make use of the equation (4).

Once we compute the probabilities $P\{0 \text{ collision}\}, ..., P\{N-1 \text{ collisions}\}$, we can obtain the cumulative probabilities and percentage points for its distribution. Using the percentage points, we can decide whether the number of actual collisions observed in a random number generator is too small or too big.

Example 2 *Study the Mathematica Collision examples file.*

Remark 4 *Knuth [1] labels the run and collision tests as strong tests, meaning several generators that were thought to be satisfactory in the past failed these tests. In contrast, the frequency test is the weakest; almost all known generators pass it.*

References

[1] Donald E. Knuth. The Art of Computer Programming, Vol 2. Addison Wesley.

[2] R. B. D'Agostino and M. A. Stephens. Goodness-of-Fit Techniques. Marcel Dekker.

[3] R. G. T. Grafton. "Algorithm AS 157. The Runs-Up and Runs-Down Tests", Applied Statistics, Vol 30, No 1, 1981, pp 81-85.