

Testing Random Numbers: Theory and Practice

Prof. Michael Mascagni

Applied and Computational Mathematics Division, Information Technology Laboratory
National Institute of Standards and Technology, Gaithersburg, MD 20899-8910 **USA**

AND

Department of Computer Science

Department of Mathematics

Department of Scientific Computing

Graduate Program in Molecular Biophysics

Florida State University, Tallahassee, FL 32306 **USA**

E-mail: mascagni@fsu.edu

URL: <http://www.cs.fsu.edu/~mascagni>

Overview

Chi-Square Test

The Kolmogorov-Smirnov (K-S) Test

Empirical Tests

- Equidistribution Test (Frequency Test)

- Serial Test

- Gap Test

- Poker Test

- Coupon Collector's Test

- Permutation Test

- Runs Test

- Maximum of t Test

- Collision Test

- Serial Correlation Test

The Spectral Test

Chi-Square Test

Eg. "Throwing 2 dice"

s : Value of the sum of the 2 dice.

p_s : Probability.

s	2	3	4	5	6	7	8	9	10	11	12
p_s	$\frac{1}{36}$	$\frac{1}{18}$	$\frac{1}{12}$	$\frac{1}{9}$	$\frac{5}{36}$	$\frac{1}{6}$	$\frac{5}{36}$	$\frac{1}{9}$	$\frac{1}{12}$	$\frac{1}{18}$	$\frac{1}{36}$

If we throw dice 144 times:

s	2	3	4	5	6	7	8	9	10	11	12
Observed: Y_s	2	4	10	12	22	29	21	15	14	9	6
Expected: np_s	4	8	12	16	20	24	20	16	12	8	4

Chi-Square Test

Is a pair of dice loaded?

We can't make a definite yes/no statement, but we can give a probabilistic answer.

We can form the Chi-Square Statistic.

$$\begin{aligned}\chi^2 &= \sum_{1 \leq s \leq k} \frac{(Y_s - np_s)^2}{np_s} \\ &= \frac{1}{n} \sum_{1 \leq s \leq k} \left(\frac{Y_s^2}{p_s} \right) - n\end{aligned}$$

$\chi^2 = k - 1$: degrees of freedom

k : Number of categories

n : Number of observances

Table of Chi-Square Distribution

Entry in row χ^2 under column p is x , which means

“The quantity χ^2 will be less than or equal to x , with approximate probability p , if n is large enough.”

Example:

Value of s	2	3	4	5	6	7	8	9	10	11	12
Experiment 1, Y_s	4	10	10	13	20	18	18	11	13	14	13
Experiment 2, Y_s	3	7	11	15	19	24	21	17	13	9	5

$$\chi_1^2 = 29 \frac{59}{120}$$

$$\chi_2^2 = 1 \frac{11}{120}$$

Table of Chi-Square Distribution

$$\chi_1^2 = 29 \frac{59}{120}$$

$$\chi_2^2 = 1 \frac{11}{120}$$

Discussion:

χ_1^2 is too high, χ^2 0.1% of the time.

χ_2^2 is too low, χ^2 0.01% of the time.

Both represent x with a significant departure from randomness.

To use Chi-Square distribution table, n should be large.

How large should n be?

Rule of thumb:

n should be large enough to make each np_s be 5 or greater.

Chi-Square Test

1. Large number n of independent observations.
2. Count the number of observations on k categories.
3. Compute χ^2 .
4. Look up Chi-Square distribution table.

$\chi^2 < 1\%$ or $\chi^2 > 99\%$	reject
$1\% < \chi^2 < 5\%$ or $95\% < \chi^2 < 99\%$	suspect
$5\% < \chi^2 < 10\%$ or $90\% < \chi^2 < 95\%$	almost suspect
otherwise	accept

The Kolmogorov-Smirnov (K-S) Test

χ^2 Test : for discrete random data

K-S Test : for continuous random data

Def: $F(x) = P[(X \leq x)]$, cumulative distribution function (CDF) for r.v. X

n independent observations of X : X_1, X_2, \dots, X_n

Def: Empirical CDF $F_n(x)$ based on the X_i 's

$$F_n(x) = \frac{1}{n} \sum_{i=1}^n \mathcal{I}_{[0,x)}(x_i)$$

where $\mathcal{I}_{[0,x)}$ is the characteristic function of the interval $[0, x)$

The Kolmogorov-Smirnov Test

The K-S Test is based on $F(x) - F_n(x)$

$$K_n^+ = \sqrt{n} \max_{-\infty < x < +\infty} (F_n(x) - F(x))$$

maximum deviation when F_n is greater than F .

$$K_n^- = \sqrt{n} \max_{-\infty < x < +\infty} (F(x) - F_n(x))$$

maximum deviation when F_n is less than F .

$K_n = \max(K_n^+, K_n^-)$, table like Chi-Square to find the percentile, but unlike χ^2 , the table fits any size of n .

The Kolmogorov - Smirnov Test

Simple procedure to obtain K_n^+ , K_n^- , to test hypothesis that $X_j \sim F$

1. Obtain observations X_1, X_2, \dots, X_n .
2. Rearrange (sort) into ascending order (with renumbering).

$$X_1 \leq X_2 \leq \dots \leq X_n$$

3. Calculate K_n^+ , K_n^-

$$K_n^+ = \sqrt{n} \max_{1 \leq j \leq n} \left(\frac{j}{n} - F(X_j) \right), \quad K_n^- = \sqrt{n} \max_{1 \leq j \leq n} \left(F(X_j) - \frac{j-1}{n} \right)$$

The Kolmogorov-Smirnov Test

Dilemma: We need a large n to differentiate F_n and F .
Large n will average out local random behavior.

Compromise: Consider a moderate size for n , say 1000.

Make a fairly large number of K_{1000}^+ on different parts of the random sequence $K_{1000}^+(1), K_{1000}^+(2), \dots, K_{1000}^+(r)$.

Apply another KS Test. The distribution of K_n^+ is approximated.

$$F_\infty(x) = 1 - e^{-2x^2}$$

Significance: Detects both local and global random behavior.

Empirical Tests

Empirical Tests: 10 tests

Test of real number sequence

$$\langle \mathcal{U}_n \rangle = \mathcal{U}_0, \mathcal{U}_1, \mathcal{U}_2 \dots$$

Test of integer number sequence

$$\langle \mathcal{Y}_n \rangle = \mathcal{Y}_0, \mathcal{Y}_1, \mathcal{Y}_2 \dots$$

$$\mathcal{Y}_n = \lfloor d\mathcal{U}_n \rfloor$$

$$\mathcal{Y}_n : \text{integers}[0, d - 1]$$

A. Equidistribution Test (Frequency Test)

Two ways:

1. Use χ^2 test



d intervals

Count the number of sequence $\langle \mathcal{Y}_n \rangle = \mathcal{Y}_0, \mathcal{Y}_1, \mathcal{Y}_2, \dots$ falling into each interval

$k=d$

$$p_s = \frac{1}{d}$$

2. Use KS Test

Test $\langle \mathcal{U}_n \rangle = \mathcal{U}_0, \mathcal{U}_1, \mathcal{U}_2, \dots$

$$F(x) = x \text{ for } 0 \leq x \leq 1$$

C. Gap Test

Examine length of “gaps” between occurrences of $U_j \in \mathcal{I} = (\alpha, \beta)$, where $0 \leq \alpha < \beta \leq 1$, and $p = \beta - \alpha$. A gap is the length r where Length of $U_j, U_{j+1}, \dots, U_{j+r}$ have $U_j, U_{j+r} \in \mathcal{I}$ and all the other are not. Algorithm:

1. Initialize: $j \leftarrow -1, s \leftarrow 0$
2. $r \leftarrow 0$
3. if $(\alpha \leq U_j \leq \beta)$, $j \leftarrow j+1$
else goto 5.
4. $r \leftarrow r+1$, goto 3.
5. record gap length.
if $r \geq t$, $COUNT[t] \leftarrow COUNT[t]+1$
else $COUNT[r] \leftarrow COUNT[r]+1$
6. Repeat until n gaps are found.

C. Gap Test

$COUNT[0], COUNT[1], \dots, COUNT[t]$ should have the following probability:

- ▶ $p_0 = p, p_1 = p(1 - p), p_2 = p(1 - p)^2, \dots, p_{t-1} = p(1 - p)^{t-1}, p_t = p(1 - p)^t$

Now, we can apply the χ^2 test.

Special cases:

- ▶ $(\alpha, \beta) = (0, \frac{1}{2}) \leftarrow$ runs above the mean
- ▶ $(\alpha, \beta) = (\frac{1}{2}, 1) \leftarrow$ runs below the mean

D. Poker Test

Consider 5 successive integers $(\mathcal{Y}_{sj}, \mathcal{Y}_{sj+1}, \mathcal{Y}_{sj+2}, \mathcal{Y}_{sj+3}, \mathcal{Y}_{sj+4})$

Pattern	Example	Pattern	Example
All different	abcde	Full house	aaabb
One Pair	aabcd	Four of a kind	aaaab
Two Pairs	aabbc	Five of a kind	aaaaa
Three of a kind	aaabc		

Simplify:

5 different	all different
4 different	one pair
3 different	two pairs or three of a kind
2 different	full house or four of a kind
5 same numbers	five of a kind

D. Poker Test

Generalized:

n groups of k successive numbers (k - tuples) with r different values.

$$pr = \frac{d(d-1)\dots(d-r+1)}{d^k} \begin{Bmatrix} k \\ r \end{Bmatrix}$$

$d = \text{number of categories}$

Then, the χ^2 test can be applied.

Stirling Numbers of the Second Kind

- ▶ Notation: $S(n, k)$ or $\left\{ \begin{matrix} n \\ k \end{matrix} \right\}$
- ▶ Definition: counts the number of ways to partition a set of n labelled objects into k nonempty unlabelled subsets **or**
- ▶ Also counts the number of different equivalence relations with precisely k equivalence classes that can be defined on an n element set
- ▶ Obviously, $\left\{ \begin{matrix} n \\ n \end{matrix} \right\} = 1$ and for $n \geq 1$, $\left\{ \begin{matrix} n \\ 1 \end{matrix} \right\} = 1$: the only way to partition an "n"-element set into "n" parts is to put each element of the set into its own part, and the only way to partition a nonempty set into one part is to put all of the elements in the same part.
- ▶ They can be calculated using the following explicit formula:

$$\left\{ \begin{matrix} n \\ k \end{matrix} \right\} = \frac{1}{k!} \sum_{j=0}^k (-1)^{k-j} \binom{k}{j} j^n$$

E. Coupon Collector's Test

In the sequence $\mathcal{Y}_0, \mathcal{Y}_1, \dots$, the lengths of the segments $\mathcal{Y}_{j+1}, \mathcal{Y}_{j+2}, \dots, \mathcal{Y}_{j+r}$ are collected to get a complete set of integers from 0 to $d-1$.

Algorithm:

1. Initialize $j \leftarrow -1, s \leftarrow 0, COUNT[r] \leftarrow 0$ for $d \leq r < t$.
2. $q \leftarrow r \leftarrow 0, OCCURS[k] \leftarrow 0$ for $0 \leq k < d$.
3. $r \leftarrow r+1, j \leftarrow j+1$
4. Complete Set? $OCCURS[\mathcal{Y}_j] \leftarrow 1$ and $q \leftarrow q+1$
if $q=d$, a complete set
 $q < d$, goto 3.
5. Record the length.
if $r \geq t, COUNT[t] \leftarrow COUNT[t]+1$
else $COUNT[r] \leftarrow COUNT[r]+1$
6. Repeat until n values are found.

E. Coupon Collector's Test

Chi-Square Test can be applied to $COUNT[d]$, $COUNT[d+1]$, ..., $COUNT[t]$

$$p_r = \frac{d!}{d^r} \binom{r-1}{d-1}, d \leq r < t$$

$$p_t = 1 - \frac{d!}{d^{t-1}} \binom{t-1}{d}$$

F. Permutation Test

A t -tuple $(U_{jt}, U_{jt+1}, \dots, U_{jt+t-1})$ can have $t!$ possible relative orderings.

For Example: $t=3$

There should be $3! = 6$ categories

1 < 2 < 3	2 < 1 < 3	2 < 3 < 1
1 < 3 < 2	3 < 1 < 2	3 < 2 < 1

$$k = t! \qquad \rho_s = \frac{1}{t!}$$

We can apply χ^2 test now.

G. Run Test

Examine the length of monotone subsequences.

“Runs up”: increasing

“Runs down”: decreasing

For Run i , the length of the run is $COUNT[i]$.

$$| \underbrace{1 \ 2 \ 9}_3 | \underbrace{8}_1 | \underbrace{5}_1 | \underbrace{3 \ 6 \ 7}_3 | \underbrace{0 \ 4}_2 |$$

Note: χ^2 test cannot be directly applied because of lack of independence (each segment depends on previous segment).

Then, we need to calculate

$$\chi^2 = \frac{1}{n} \sum_{1 \leq i, j \leq 6} (COUNT[i] - nb_i) (COUNT[j] - nb_j) a_{ij}$$

G. Runs Test

$$\begin{bmatrix} a_{11} & a_{12} & a_{13} & a_{14} & a_{15} & a_{16} \\ a_{21} & a_{22} & a_{23} & a_{24} & a_{25} & a_{26} \\ a_{31} & a_{32} & a_{33} & a_{34} & a_{35} & a_{36} \\ a_{41} & a_{42} & a_{43} & a_{44} & a_{45} & a_{46} \\ a_{51} & a_{52} & a_{53} & a_{54} & a_{55} & a_{56} \\ a_{61} & a_{62} & a_{63} & a_{64} & a_{65} & a_{66} \end{bmatrix} = \begin{bmatrix} 4529.4 & 9044.9 & 13568 & 18091 & 22615 & 27892 \\ 9044.9 & 18097 & 27139 & 36187 & 45234 & 55789 \\ 13568 & 27139 & 40721 & 54281 & 67852 & 83685 \\ 18091 & 36187 & 54281 & 72414 & 90470 & 111580 \\ 22615 & 45234 & 67852 & 90470 & 113262 & 139476 \\ 27892 & 55789 & 83685 & 111580 & 139476 & 172860 \end{bmatrix}$$

$$(b_1 \ b_2 \ b_3 \ b_4 \ b_5 \ b_6) = \left(\frac{1}{6} \quad \frac{5}{24} \quad \frac{11}{120} \quad \frac{19}{720} \quad \frac{29}{5040} \quad \frac{1}{890} \right)$$

Then, χ^2 should have the same χ^2 distribution with degree 6.

H. Maximum-of- t Test

Examine the maximum value of t uniform random variables.

Let $\chi^2_j = \max(\mathcal{U}_{tj}, \mathcal{U}_{tj+1}, \dots, \mathcal{U}_{tj+t-1})$.

The distribution is $F(x) = X^t$

Then, we can apply the Kolmogorov - Smirnov Test here.

I. Collision Test

Suppose we have m urns and n balls, $m \ll n$.

Most of the balls will fall in an empty urn.

If a ball falls in an urn that already has a ball, we call it a “collision”.

A generator passes the collision test only if it doesn't induce too many or too few collisions.

Probability of c collisions occurring:

$$\frac{m(m-1)\dots(m-n+c+1)}{m^n} \binom{n}{n-c}$$

J. Serial Correlation Test

Consider the observations $(\mathcal{U}_0, \mathcal{U}_1, \dots, \mathcal{U}_{n-1})$ and $(\mathcal{U}_1, \dots, \mathcal{U}_{n-1}, \mathcal{U}_0)$
 Test the correlation between these two tuples.

We compute:

$$C = \frac{n(\mathcal{U}_0\mathcal{U}_1 + \mathcal{U}_1\mathcal{U}_2 + \dots + \mathcal{U}_{n-2}\mathcal{U}_{n-1} + \mathcal{U}_{n-1}\mathcal{U}_0) - (\mathcal{U}_0 + \mathcal{U}_1 + \dots + \mathcal{U}_{n-1})^2}{n(\mathcal{U}_0^2 + \mathcal{U}_1^2 + \dots + \mathcal{U}_{n-1}^2) - (\mathcal{U}_0 + \mathcal{U}_1 + \dots + \mathcal{U}_{n-1})^2}$$

A “good” C should be between $\mu_n - 2\delta_n$ and $\mu_n + 2\delta_n$.

$$\mu_n = \frac{-1}{n-1} \quad , \quad \delta_n = \frac{1}{n-1} \sqrt{\frac{n(n-3)}{n+1}} \quad , \quad n > 2$$

The Spectral Test

Idea underlying the test: *Congruential Generators generate random numbers in grids!*

In t -dimensional space, $\{(\mathcal{U}_n, \mathcal{U}_{n+1}, \dots, \mathcal{U}_{n+t-1})\}$

Compute the distance between lines (2D), planes (3D), parallel hyperplanes (>3D).

- $1/\nu_2$: Maximum distance between lines.
Two dimensional accuracy.
- $1/\nu_3$: Maximum distance between planes.
Three dimensional accuracy.
- $1/\nu_t$: Maximum distance between hyperplanes.
 t - dimensional accuracy.

The Spectral Test

Differentiate between truly random sequences and periodic sequences.

Truly random sequences: accuracy remains same in all dimensions

Periodic sequences: accuracy decreases as t increases

Spectral Test is by far the most powerful test.

- ▶ All “good” generators pass it.
- ▶ All known “bad” generators fail it.

Summary

1. Basic idea of empirical tests:
The combination of random numbers is expected to conform to a specific distribution.
 - 1.1 Build the combination.
 - 1.2 Use χ^2 or KS test to test the deviation from the expected distribution.
2. We can perform an infinite number of tests.
3. We might be able to construct a test to “kill” a specific generator.

Other resources for RNG Testing

1. FFT, Metropolis, Wolfgang Tests (spectrum).
2. Diehard (<http://www.stat.fsu.edu/pub/diehard>)
3. SPRNG (implements most of the empirical tests and spectrum tests).
<http://sprng.cs.fsu.edu>
4. TestU01

Questions?

© Michael Mascagni, 2026