Tests for Random Numbers

When a random number generator is devised, one needs to test its property. The two properties we are concerned most are uniformity and independence. A list of tests will be discussed. The first one tests for uniformity and the second to fifth ones test independence.

Frequency test

Runs test

Autocorrelation test

Gap test

Poker test

The algorithms of testing a random number generator are based on some statistics theory, i.e. testing the hypotheses. The basic ideas are the following, using testing of uniformity as an example.

We have two hypotheses, one says the random number generator is indeed uniformly distributed. We call this , known in statistics as null hypothesis. The other hypothesis says the random number generator is not uniformly distributed. We call this , known in statistics as alternative hypothesis.

We are interested in testing result of , reject it, or fail to reject it.

To see why we don't say accept H null, let's ask this question: what does it mean if we had said accepting H null? That would have meant the distribution is truely uniform. But this is impossible to state, without exhaustive test of a real random generator with infinite number of cases. So we can only say failure to reject H null, which means no evidence of non-uniformity has been detected on the basis of the test. This can be described by the saying ``so far so good''.

On the other hand, if we have found evidence that the random number generator is not uniform, we can simply say reject H null.

It is always possible that the is true, but we rejected it because a sample landed in the region, leading us to reject . This is known as Type I error. Similarily if is false, but we didn't reject it, this also results in an error, known as Type II error.

With these information, how do we state the result of a test? (How to perform the test will be the subject of next a few sections)

A level of statistical significance has to be given. The level is the probability of rejecting the H null while the H null is true (thus, Type I error).

We want the probability as little as possible. Typical values are 0.01 (one percent) or 0.05 (five percent).

Decreasing the probability of Type I error will increase the probability of Type II error. We should try to strike a balance.

For a given set of random numbers produced by a random number generator, the more tests are, the more accurate the results will be.

Gap test

The gap test is used to determine the significance of the interval between recurrence of the same digit.

A gap of length x occurs between the recurrence of some digit.

See the example on page 313 where the digit 3 is underlined. There are a total of eighteen 3's in the list. Thus only 17 gaps can occur.

The probability of a particular gap length can be determined by a Bernoulli trail.

If we are only concerned with digits between 0 and 9, then

The theoretical frequency distribution for randomly ordered digits is given by

Steps involved in the test.

Step 1.

Specify the cdf for the theoretical frequency distribution given by Equation (8.14) based on the selected class interval width (See Table 8.6 for an example).

Step 2.

Arrange the observed sample of gaps in a cumulative distribution with these same classes.

Step 3.

Find D, the maximum deviation between F(x) and as in Equation 8.3 (on page 299).

Step 4.

Determine the critical value, , from Table A.8 for the specified value of and the sample size N.

Step 5.

If the calculated value of D is greater than the tabulated value of , the null hypothesis of independence is rejected.

Poker test

The poker test for independence is based on the frequency in which certain digits are repeated in a series of numbers.

For example 0.255, 0.577, 0.331, 0.414, 0.828, 0.909, 0.303, 0.001... In each case, a pair of like digits appears in the number.

In a three digit number, there are only three possibilities.

The individual digits can be all different. Case 1.

The individual digits can all be the same. Case 2.

There can be one pair of like digits. Case 3.

P(case 1) = P(second differ from the first) \* P(third differ from the first and second) = 0.9 \* 0.8 = 0.72

P(case 2) = P(second the same as the first) \* P(third same as the first) = 0.1 \* 0.1 = 0.01 P(case 3) = 1 - 0.72 - 0.01 = 0.27

[www.thegreatmartinicompany.com](http://www.thegreatmartinicompany.com)

[www.math.leidenuniv.nl/statscience](http://www.math.leidenuniv.nl/statscience)

Frequency test

The frequency test is a test of uniformity.

Two different methods available, Kolmogorov-Smirnov test and the chi-square test. Both tests measure the agreement between the distribution of a sample of generated random numbers and the theoretical uniform distribution.

Both tests are based on the null hypothesis of no significant difference between the sample distribution and the theoretical distribution.

The Kolmogorov-Smirnov test

This test compares the cdf of uniform distribution F(x) to the empirical cdf of the sample of N observations.

As N becomes larger, should be close to F(x)

Kolmogorov-Smirnov test is based on the statistic

that is the absolute value of the differences.

Here D is a random variable, its sampling distribution is tabulated in Table A.8.

If the calcualted D value is greater than the ones listed in the Table, the hypothesis (no disagreement between the samples and the theoretical value) should be rejected; otherwise, we don't have enough information to reject it.

Following steps are taken to perform the test.

Rank the data from smallest to largest

Compute

Compute

Determine the critical value, , from Table A.8 for the specified significance level and the given sample size N.

If the sample statistic D is greater than the critical value , the null hypothsis that the sample data is from a uniform distribution is rejected; if , then there is no evidence to reject it.

Example 8.6 on page 300.

Chi-Square test

The chi-square test looks at the issue from the same angle but uses different method. Instead of measure the difference of each point between the samples and the true distribution, chi-square checks the ``deviation'' from the ``expected'' value.

where n is the number of classes (e.g. intervals), is the number of samples obseved in the interval, is expected number of samples in the interval. If the sample size is N, in a uniform distribution,