FIPS 140-2(Change Notice 1) Random Number Tests

—Distribution Functions and Observed values of RPG100—

Ananda Vithanage† Takakuni Shimizu‡
2/10/2003
FDK CORPORATION RPG Business Promotion Dept.
1 Kamanomae, Kamiyunagaya-machi, Jyoban, Iwaki-shi, Fukushima 972-8322, Japan
E-mail: † kalin@fdk.co.jp, ‡ tshimizu@fdk.co.jp

Introduction

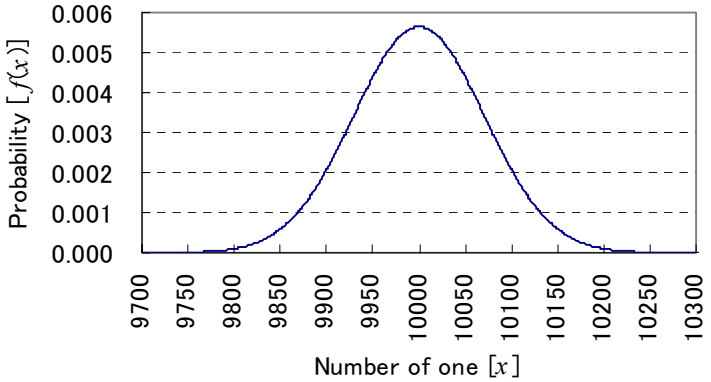
Federal Information Processing Standards (FIPS) 140-2 publication for cryptographic modules specifies four statistical tests for randomness. Instead of making the user select appropriate significance levels for these tests, explicit bounds are provided that the computed value of a statistic must satisfy. A single bit stream of 20000 bits, output from a generator, is subjected to each of the four tests. If any of the tests fail, then the generator fails the FIPS 104-2 statistical test for randomness. The distribution functions are derived and significance level are calculated for each of the four statistical tests. The observed statistical values of RPG100 are also shown with the calculated theoretical values.

1. The Monobit Test

The number of ones is counted in a bit stream of 20,000 bits.
Statistic x : denotes the number of ones in the bit stream
Acceptance Region: $9{,}725 < x < 10{,}275$

Distribution Function:
Considering a bit stream of 20,000 bits in which all bit occurrences can be considered as Bernoulli trials with success probability of 1/2. Let x denotes the number of ones occurred in n bits, then the number of distinct patterns are ${}_nC_x = n!/[(n-x)!x!]$. The probability distribution function $f(x)$ describes the probability of x number of ones in n bits.

$$f(x) = nCx(1/2)^x(1-1/2)^{n-x} = nCx(1/2)^n \qquad (1)$$



Significance Level α *:* $\alpha = 1 - \sum_{x=9726}^{10274} f(x) \cong 0.0001$

2. The Poker Test

A bit stream of 20,000 bits are divided into 5,000 non-overlapping consecutive 4 bits segments. The total number of patterns of 4 bits segments is $2^4 = 16$. The number of occurrences of each of the 16 possible patterns is counted. Let g_i $(i = 0 \sim 15)$ be the number of occurrences of each pattern.

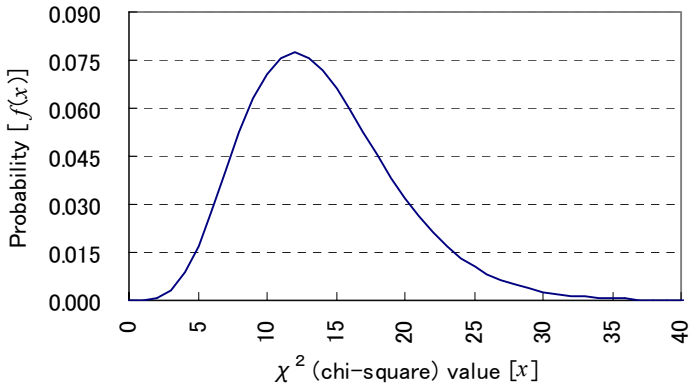
Statistic x : $x = (16/5000) \times \sum_{i=0}^{15} g_i^{\,2} - 5000$

Acceptance Region : $2.16 < x < 46.17$

Distribution Function :

The x statistic approximately follows a χ^2 distribution with 15 degrees of freedom. The total number of patterns considered is 16, then the probability of occurrence of each pattern is 1/16, and the degree of freedom of the distribution is 15 because if the frequency of 15 patterns are fixed then the remaining pattern will automatically be fixed. The probability density function $f(x)$ describes the probability of x statistic.

$$f(x) = \frac{1}{2^{15/2}\Gamma(15/2)} x^{15/2-1} e^{-x/2} \qquad \text{where,} \quad \Gamma(y) \equiv \int_0^\infty z^{y-1} e^{-z} dz \tag{2}$$



Significance Level α *:* $\alpha = 1 - \int_{2.16}^{46.17} f(x)dx \cong 0.0001$

3. The Runs Test

A run is defined as a sequence of consecutive values of one or zero. In this test, the number of runs in a bit stream of 20,000 bits is counted. The 12 tests are individually performed for each run length of 1, 2, 3, 4, 5 and 6 or greater 6 of either one or zero.

Statistic x : denotes the number of runs of each length appear in the bit stream

Acceptance Regions :

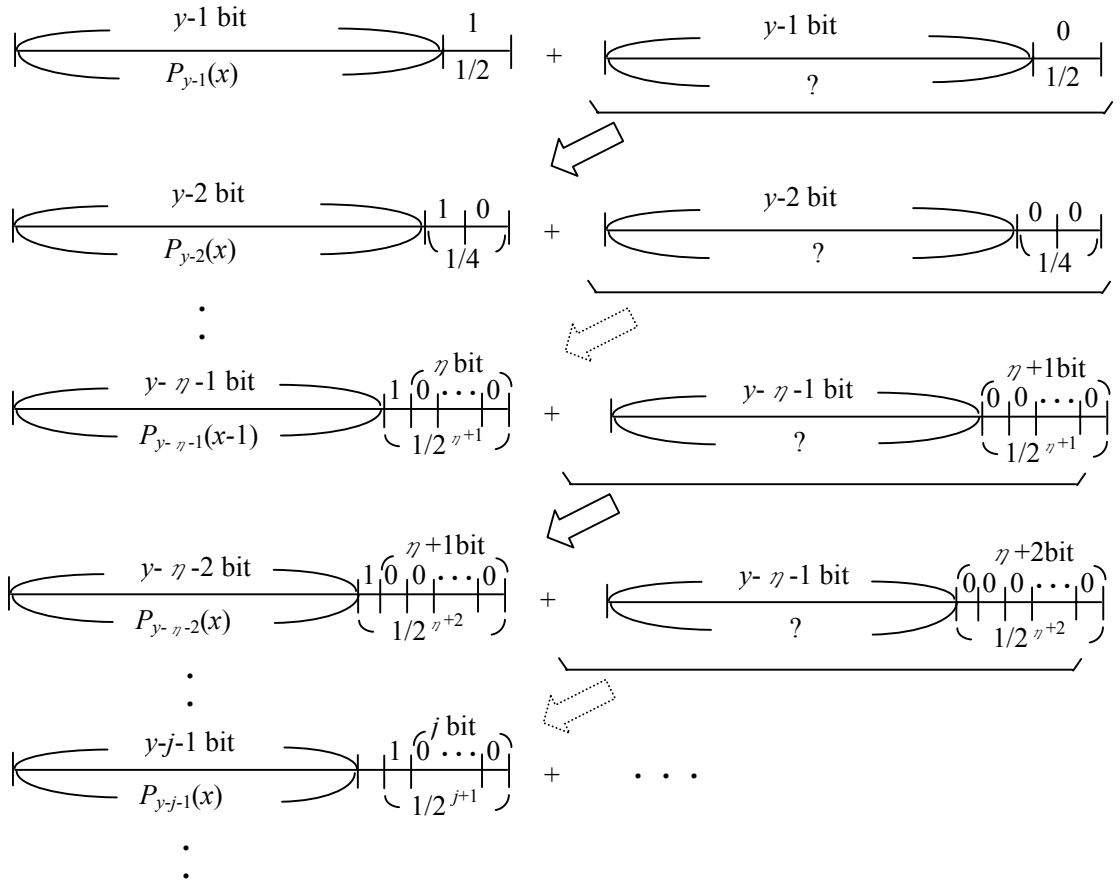
Length of runs	Acceptance Region
1	2,315 ≦ x ≦ 2,685
2	1,114 ≦ x ≦ 1,386
3	527 ≦ x ≦ 723
4	240 ≦ x ≦ 384
5	103 ≦ x ≦ 209
6+	103 ≦ x ≦ 209

Distribution Function:

For a fixed bit length the distribution function of runs is not simple. When use combinations in calculation it consumes too much time due to calculation of factorials and summations and it is not appropriate to use approximations in this case. The calculation of probability for each length of run, the recursive approach is straightforward.

Probability of Runs of Length η :

Let $P_y(x)$ be the probability of x ($x \geqq 2$) number of runs of length η of zeros appears in a bit stream of y bits. It can be considered as two reduced simpler forms either y's LSB fixed to one or zero. Then the probability $P_y(x)$ is the addition of the probabilities of the two simpler forms. However, the probability of x number of runs when y's LSB is fixed to one can easily be written as $P_{y-1}(x)/2$, but when it is fixed to zero it cannot easily be obtained. The reason is that the rightmost run of length η might be contained that zero bit. The calculations are proceed by fixing the bit next to the y's LSB either zero or one and reduce the problem to another simpler form. As mentioned before the calculation is easy if the newly fixed bit is one and the probability becomes $P_{y-2}(x)/4$. The following figure explains the stapes and reduced forms of calculation of original $P_y(x)$.



If this procedure is continued until the bit to be fixed is $(y - 1)^{th}$ then $P_y(x)$ can be obtained by adding all calculated probabilities $P_{y-1}(x)/2$, $P_{y-2}(x)/4$, ... However, when the length of zero of fixed bits is η, the number of remaining zero η length runs should be $x-1$ because the zero bits already fixed are η, and the probability becomes $P_{y-\eta-1}(x-1)/2^{\eta+1}$. If the fixed zero bit length is greater than η then for further calculation steps consider $P_{y-\eta-2}(x-1)$ as $P_{y-\eta-2}(x)$. When number of runs x is 0 or 1, the equations for these two cases have some differences, but the explanation is omitted because the basic idea remains unchanged. The recurrence relations for

each of the three different cases are given below.

Probability of no runs of length η in k bits

$$P_k(0) = 1 \qquad (\ k = 1 \sim \eta - 1\ )$$

$$P_k(0) = 1 - 2^{-\eta} \qquad (\ k = \eta\ )$$

$$P_k(0) = \sum_{i=1}^{\eta} 2^{-i} P_{\eta+1-i}(0) + 2^{-(\eta+1)} \qquad (\ k = \eta + 1\ )$$

$$P_k(0) = \sum_{i=1}^{\eta} 2^{-i} P_{k-i}(0) + \sum_{i=\eta+2}^{k-1} 2^{-i} P_{k-i}(0) + 2^{-(k-1)} \qquad (\ k \geqq \eta + 2\ ) \qquad (3)$$

Probability of one run of length η in k bits

$$P_k(1) = 0 \qquad (\ k = 1 \sim \eta - 1\ )$$

$$P_k(1) = 2^{-\eta} \qquad (\ k = \eta\ ,\ \eta + 1\ )$$

$$P_k(1) = \sum_{i=1}^{\eta} 2^{-i} P_{k-i}(1) + \sum_{i=\eta+2}^{k-\eta} 2^{-i} P_{k-i}(1) + 2^{-(\eta+1)} P_{k-\eta-1}(0) \qquad (\ k \geqq \eta + 2\ ) \qquad (4)$$

Probability of more than one run of length η in k bits

$$P_k(x) = 0 \qquad (\ k = 1 \sim (\eta+1)x - 2\ )$$

$$P_k(x) = 2^{-((\eta+1)x-1)} \qquad (\ k = (\eta+1)x - 1\ )$$

$$P_k(x) = \sum_{i=1}^{\eta} 2^{-i} P_{k-i}(x) + \sum_{i=\eta+2}^{k-(\eta+1)x+1} 2^{-i} P_{k-i}(x) + 2^{-(\eta+1)} P_{k-\eta-1}(x-1) \qquad (\ k \geqq (\eta+1)x\ ) \qquad (5)$$

Probability of Runs Longer Than η :

The probability of runs longer than η can be led by same idea as the above-mentioned. The difference is $P(x\text{-}1)$ is used throughout the calculation without changing to $P(x)$ even after the number of fixed zero bits greater than η.

Probability of no runs of length longer than η in k bits

$$P_k(0) = 1 \qquad (\ k = 1 \sim \eta - 1\ )$$

$$P_k(0) = 1 - 2^{-\eta} \qquad (\ k = \eta\ )$$

$$P_k(0) = \sum_{i=1}^{\eta} 2^{-i} P_{k-i}(0) \qquad (\ k \geqq \eta + 1\ ) \qquad (6)$$

Probability of one run of length longer than η in k bits

$$P_k(1) = 0 \qquad (\ k = 1 \sim \eta - 1\ )$$

$$P_k(1) = 2^{-\eta} \qquad (\ k = \eta\ )$$

$$P_k(1) = \sum_{i=1}^{\eta} 2^{-i} P_{k-i}(1) + \sum_{i=\eta+1}^{k-1} 2^{-i} P_{k-i}(0) + 2^{k-1} \qquad (\ k \geqq \eta + 1\ ) \qquad (7)$$

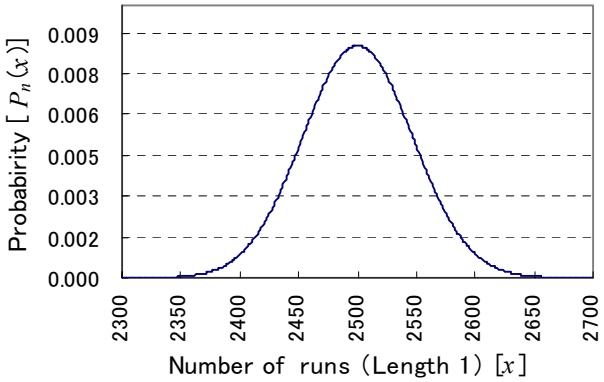
Probability of more than one run of length longer η in k bits

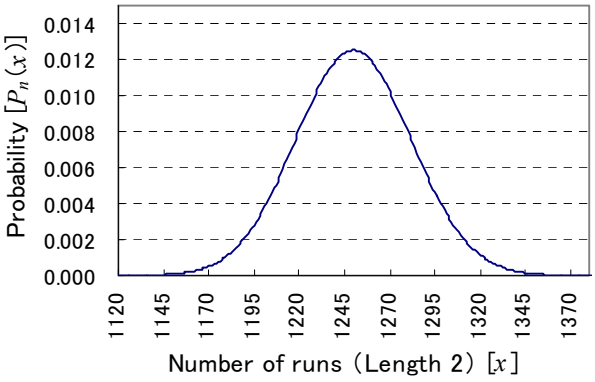
$$P_k(x) = 0 \qquad (\ k = 1 \sim (\eta+1)x - 2\ )$$

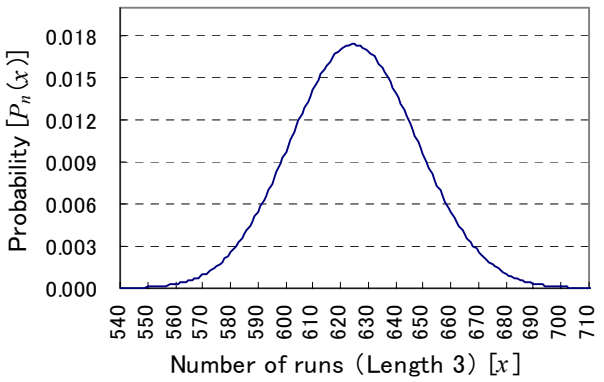
$$P_k(x) = 2^{-((\eta+1)x-1)} \qquad (\ k = (\eta+1)x - 1\ )$$

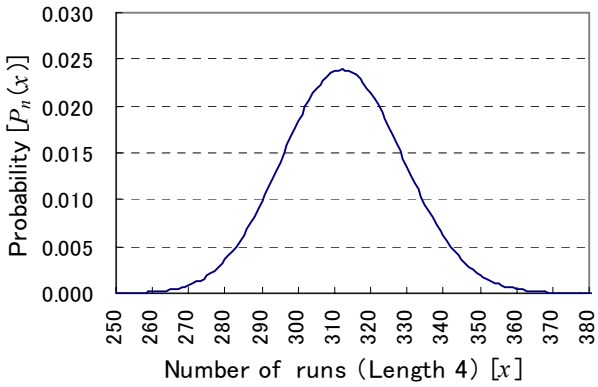
$$P_k(x) = \sum_{i=1}^{\eta} 2^{-i} P_{k-i}(x) + \sum_{i=\eta+1}^{k-(\eta+1)(x-1)+1} 2^{-i} P_{k-i}(x-1) \qquad (\ k \geqq (\eta+1)x\ ) \qquad (8)$$

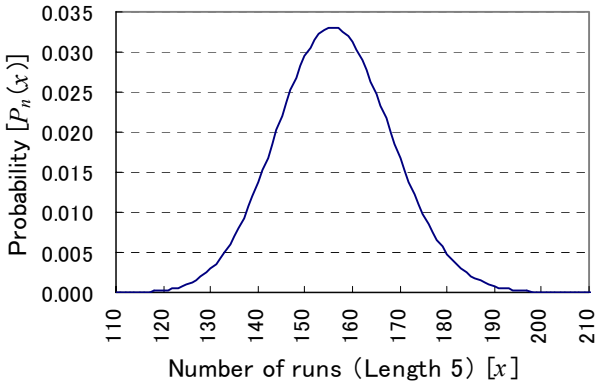
In these recursions, the amount of calculations can be reduced by omitting higher-order powers of 2. However, in cases of no run of and one run of length η is recommended to keep accuracy as much as possible. These value of $P_n(x)$, n = 20000 gives distribution of statistics of Runs Test in FIPS140-2.

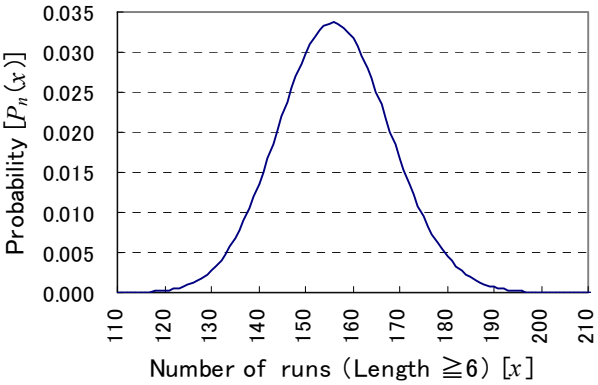












Significance Level α : $\alpha_i = 1 - \sum_{x=\min(i)}^{\max(i)} P_n(X)$

Where i is the length of run, $max(i)$ and $min(i)$ are the limits of the acceptance region for each i and n = 20000.

	Length 1 α_1	Length 2 α_2	Length 3 α_3	Length 4 α_4	Length 5 α_5	Length 6 α_{6+}
S L α_i	0.00007355	0.00001875	0.00001871	0.00001675	0.00001299	0.00000902

4. The Long Run Test

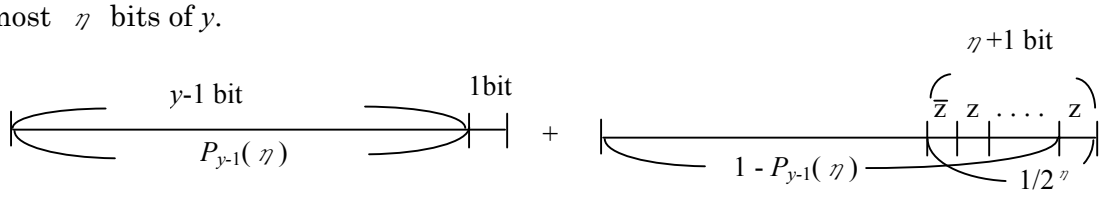
The length of the longest run is examined which appear in 20,000bit.

Statistic x: The length of the longest run in bit stream of 20,000bit (both of one and zero)

Acceptance Region: $x < 26$

Distribution Function:

The method above-mentioned in Runs Test can also be used to calculate of probability for this test, but there is another more efficient method of calculation of long runs probability. Let $P_y(\eta)$ be the probability of a run (one and zero) of length longer than η appears in a bit stream of y bits. This can be expressed by addition of two mutually exclusive event probabilities. One of them is the probability that the run longer than η appears in $y-1$ bits reducing a LSB of y and the other is the probability that the no run longer than η appears in $y-1$ bits, but appears in right most η bits of y.

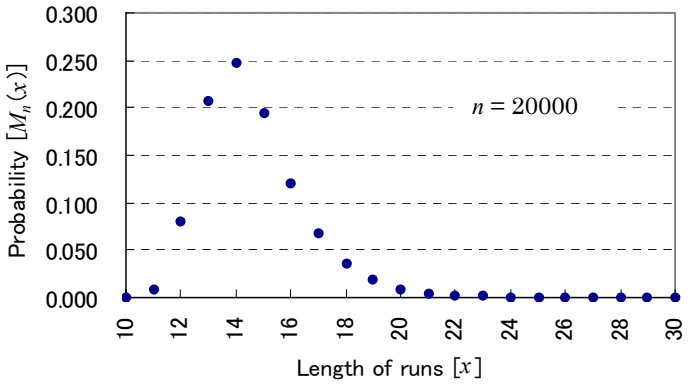


The probability that a run longer than η appears in $y-1$ bits is $P_{y-1}(\eta)$, the probability that no run longer than η in $y-1$ bits is $1 - P_{y-1}(\eta)$, and the probability that a run of length η appears in right most η bits of y bits is $1/2^{\eta}$. Considering the above probabilities and brought them into together the following recurrence relation can be derived.

$$\begin{aligned} P_k(\eta) &= 0 && (\ k = 1 \sim \eta - 1\ ) \\ P_k(\eta) &= 2^{-(\eta-1)} && (\ k = \eta\ ) \\ P_k(\eta) &= P_{k-1}(\eta) + 2^{-\eta}(1 - P_{k-1}(\eta)) && (\ k \geqq \eta + 1\ ) \end{aligned} \tag{9}$$

The following expression can be used to calculate the probability of the longest run of length η appears in n bits.

$$M_n(\eta) = P_n(\eta) - P_n(\eta + 1) \tag{10}$$



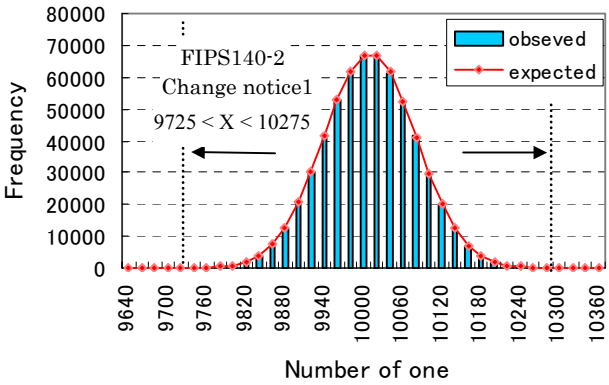
Significance Level α *:* $\alpha = P_n \cong 0.000298$ Where $n = 20000$, $\eta = 26$.

5 Measured Test Statistics of RPG100

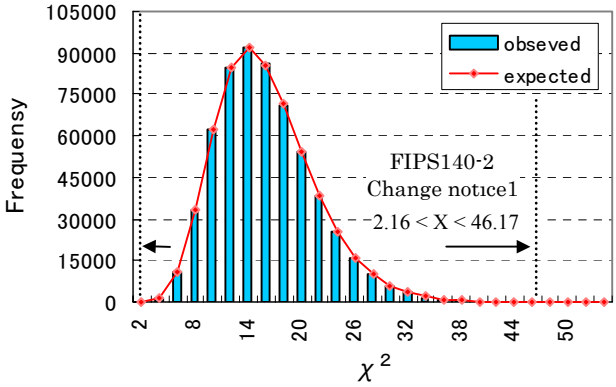
The FIPS140-2 random numbers test is performed for 600,000 different bit samples of RPG100. The theoretical values obtained from above calculations are compared with the data that is obtained form various measurements.

The random digits are collected form a single sample of RPG100 with 3.3 Volts and 250KHz at 25 Degrees of Celsius.

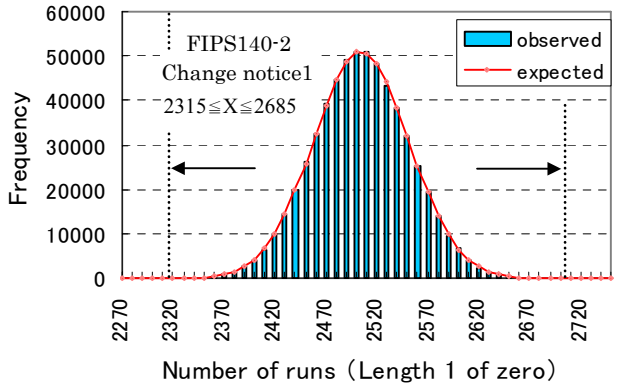
The monobit test

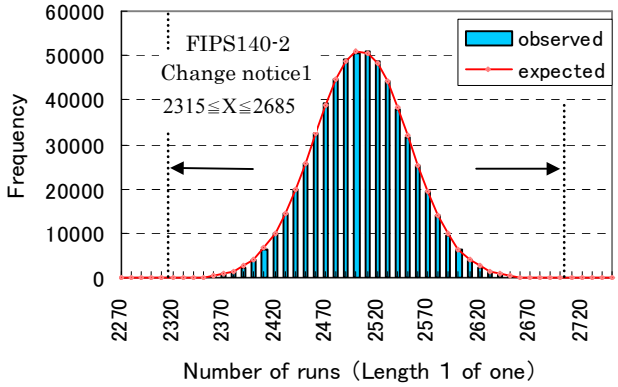


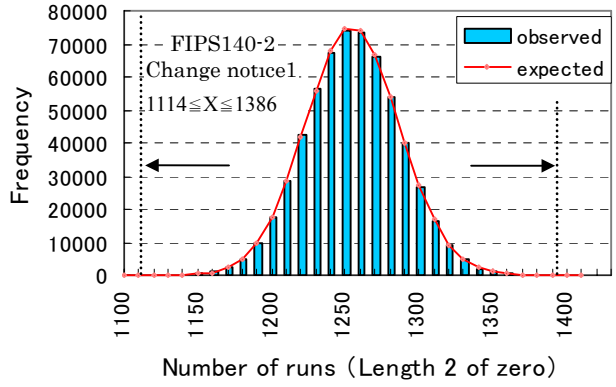
The poker test

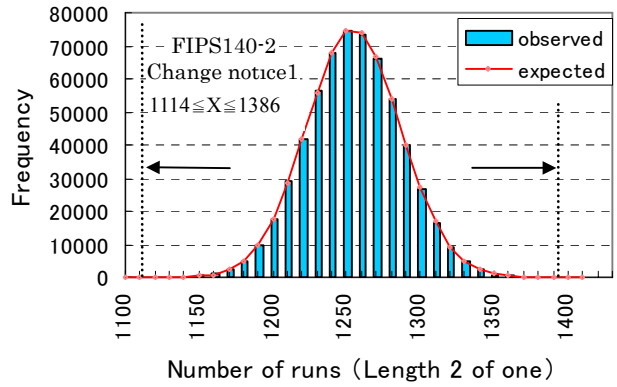


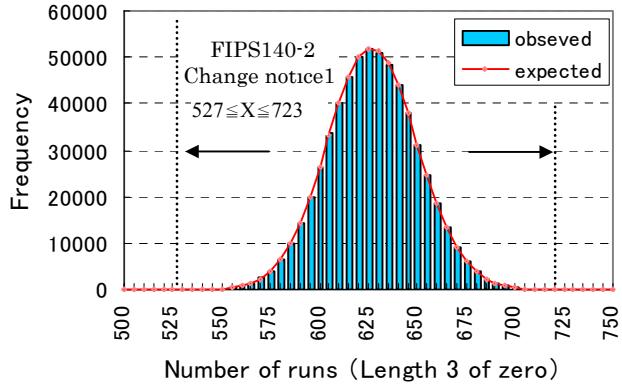
The runs test

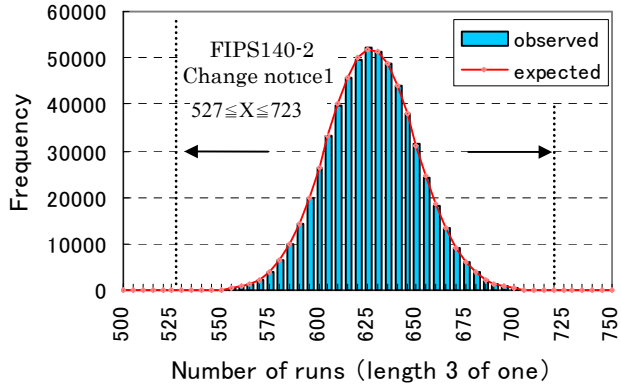


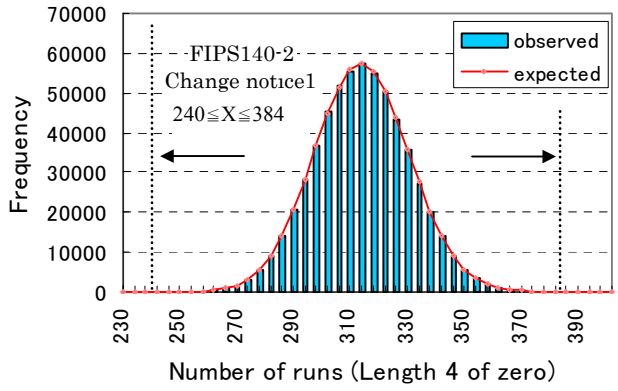


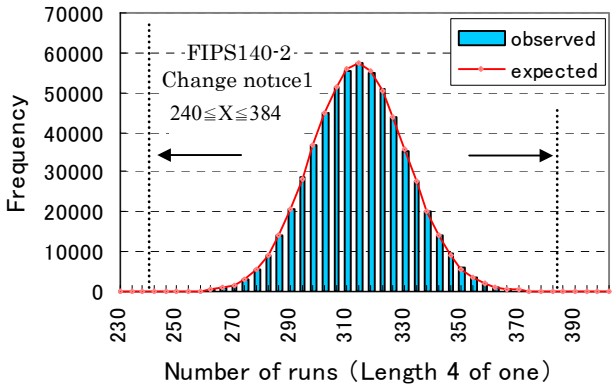


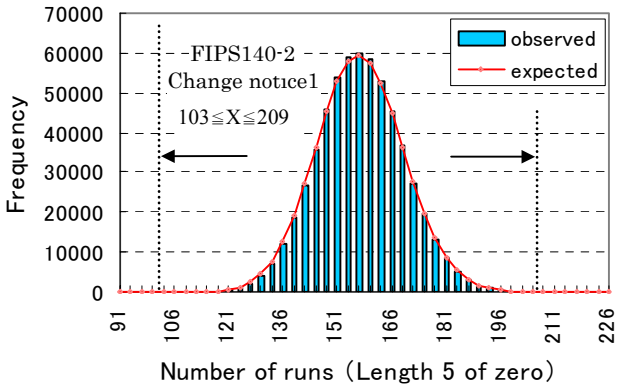


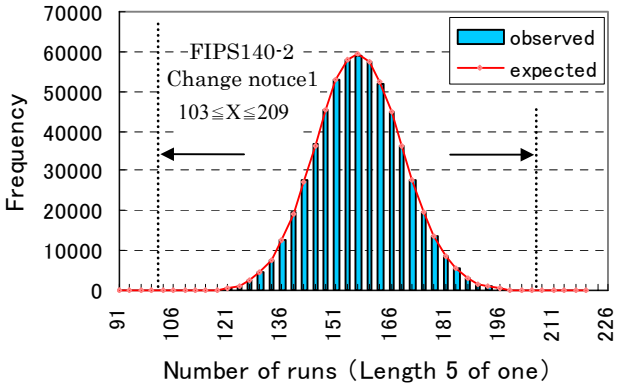


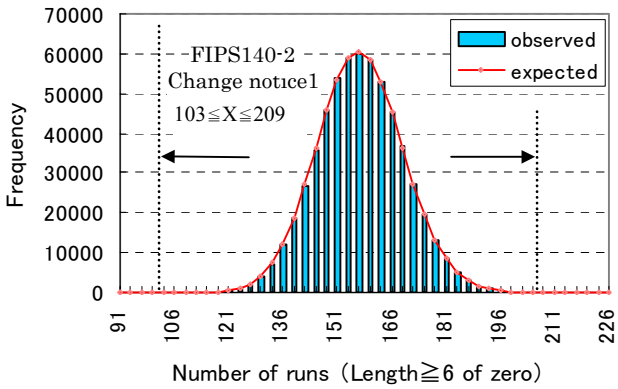


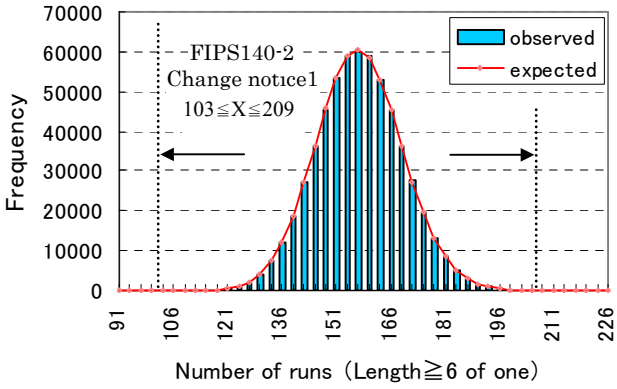




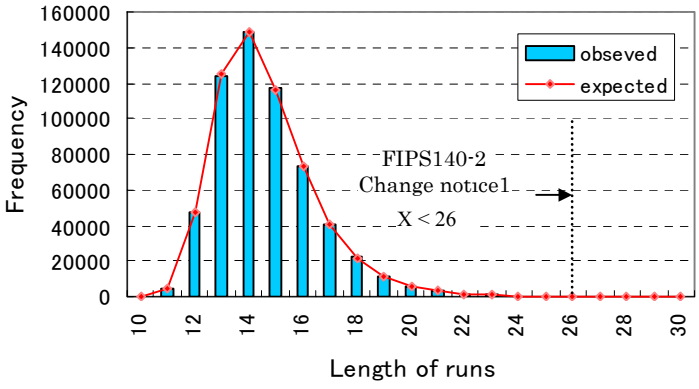








The longruns test



6 Distribution of Number of Failures Based on FIPS140-2 Test

The number of test failures is counted by performing FIPS140-2 statistical test to 4000 different bit samples. The following graphs explain the number of test failures in each 4000 tests (total sample size: 1500 x 4000).

The random digits are collected form a single sample of RPG100 with 3.3 Volts and 250KHz at 25 Degrees of Celsius.