

**Корректировки двухуровневого подхода
к тестированию генераторов случайных чисел,
обеспечивающие его надежность,
на примере частотного теста из набора тестов
NIST SP800-22**

Александр Серов

Академия криптографии Российской Федерации

Великий Новгород

3 июня 2025 г.

Генераторы случайных чисел (ГСЧ) являются основными компонентами разнообразных *прикладных компьютерных программ*, являются источниками *энтропии* и важными элементами *криптографических систем*, безопасность которых основана на *совершенной случайности* выходных последовательностей ГСЧ.

ГСЧ делятся на два типа: *истинные* и *псевдослучайные*.

Проверка качества порождаемых ГСЧ последовательностей выполняется с помощью статистических критериев проверки нулевой гипотезы $\mathcal{H}_0$.

Если гипотеза $\mathcal{H}_0$ верна, то тестируемый ГСЧ, называют *идеальным ГСЧ*.

Аналогичным статистическим критериям проверки $\mathcal{H}_0$ подвергаются *блочные шифры* или *хэш-функции*.

Традиционным способом контроля выходных последовательностей ГСЧ является порождение им N непересекающихся подпоследовательностей длины n .

Пусть $S: X^n \rightarrow E$ — произвольная измеримая функция (статистика), где X^n — выборочное пространство, E — множество значений статистики S .

Определение. p -значением $P_v(S)$ статистики S называется вероятность получения значения s статистики S , равного или “более экстремального”, чем наблюдаемое значение статистики $s_{obs} = S(\omega)$, $\omega \in X^n$, в предположении, что нулевая гипотеза $\mathcal{H}_0$ верна,

$$P_v(S) = \mathbf{P} \{S \geq s_{obs} \mid \mathcal{H}_0\}.$$

Тогда при заданном уровне значимости α , последовательности, порождаемые ГСЧ, можно считать неотклоняющими гипотезу $\mathcal{H}_0$, если, например, все критерии из некоторого их набора дают p -значения не меньше α .

Разумной альтернативой традиционному способу является **двухуровневый подход к тестированию** ГСЧ, который, учитывая близость распределения $P_v(S)$ к равномерному на $[0, 1]$, заключается в контроле принадлежности долей отрезков последовательности, прошедших тест, заданным интервалам и проверке гипотезы о принадлежности выборки из N наблюдаемых p -значений равномерному распределению на $[0, 1]$ с помощью критерия согласия Пирсона.

Последовательность длины $n \cdot N$ битов разбивается на N отрезков. Выбранный статистический тест применяется к каждому отрезку последовательности в отдельности, и эмпирическое распределение N полученных p -значений оценивается следующим образом:

- (а) определяется область допустимых значений доли ζ отрезков посл-ти, для которых p -значения оказались больше α , в предположении, что распределение ζ может быть аппроксимировано нормальной случайной величиной $\mathcal{N}(\mu, \sigma)$, где $\mu = 1 - \alpha$, $\sigma = \sqrt{\alpha(1 - \alpha)/N}$:

$$\left[1 - \alpha - 3\sqrt{\frac{\alpha(1 - \alpha)}{N}}, 1 - \alpha + 3\sqrt{\frac{\alpha(1 - \alpha)}{N}} \right];$$

- (б) распределение N p -значений проверяется на равномерность, путём разбиения интервала $[0, 1]$ на k непересекающихся подинтервалов, определении чисел p -значений, оказавшихся в каждом подинтервале, и вычислении p -значения статистики критерия согласия (хи-квадрат) χ_{obs}^2 от чисел p -значений в подинтервалах

$$p_{II} = \frac{1}{\Gamma((k-1)/2)} \int_{\chi_{obs}^2/2}^{\infty} t^{(k-1)/2-1} e^{-t} dt, \text{ где } \chi_{obs}^2 = \sum_{i=1}^k \frac{(\pi_i - N/k)^2}{N/k},$$

π_i — число p -значений в i -м подинтервале, $i \in \{1, 2, \dots, k\}$, $\Gamma(\cdot)$ — гамма-функция; для заданного α_{II} , гипотеза $\mathcal{H}_0$ отклоняется, если $p_{II} \leq \alpha_{II}$, иначе — $\mathcal{H}_0$ принимается.

Rukhin A., Soto J., Nechvatal J., Smid M., Barker E., Leigh S., Levenson M., Vangel M., Banks D., Heckert A., Dray J., Vo S. "A statistical test suite for the validation of random number generators and pseudo random number generators for cryptographic applications", NIST Special Publication 800-22 Revision 1a, 27 April 2010.

Pareschi F., Rovatti R., Setti G. "Second-level NIST randomness test for improving test reliability", ISCAS 2007 (New Orleans, USA, May 27–30, 2007), 1437–1440.

Pareschi F., Rovatti R., Setti G. "On statistical tests for randomness included in the NIST SP800-22 test suite and based on the binomial distribution", IEEE Trans. Inf. For. Sec. **7**:2 (2012), 491–505.

$$F_p(x) = x + \left(\inf_{p \in \mathcal{P}, p > x} p - x \right) + \left(F_{|| \cdot ||}^{\infty}(\xi) - F_{|| \cdot ||}(\xi) \right),$$

где ξ неявно определяется из $1 - F_{|| \cdot ||}^{\infty}(\xi) = \inf_{p \in \mathcal{P}, p > x} p$.

"The output of these tests is the so-called p -values which is a random variable whose **distribution converges to the uniform distribution in the interval** $[0, 1]$ when testing an **increasing number of samples** from an ideal generator."

Результаты двухуровневого тестирования AES-ГСЧ

#	Test Name	χ^2 test, N sequences			
		$N = 10^3$	$N = 10^4$	$N = 10^5$	$N = 2^{20}$
1	Frequency	0.616305	0.290806	0.588411	0.000803
2	Block Frequency	0.187581	0.773212	0.374097	0.000125
3	Cumulative Sums	0.401199	0.124765	0.959543	0.000009
4	Runs	0.150340	0.885418	0.910568	0.107966
5	Longest Run	0.610070	0.239883	0.000355	0.000000
6	Binary Matrix Rank	0.878618	0.341017	0.000000	0.000000
7	Discrete Fourier Transform	0.371941	0.014836	0.000000	0.000000
8	Overlapping Templ. Match.	0.071177	0.202268	0.000000	0.000000
9	Universal statistical test	0.574903	0.108534	0.000000	0.000000
10	Approximate Entropy	0.246750	0.078038	0.219501	0.000000
11	Serial	0.942198	0.174057	0.213964	0.572679
12	Linear Complexity	0.839507	0.279152	0.299852	0.117305
13	Non-overlap. Templ. Match.	0.092041	0.372782	0.121382	0.275416
14	Random Excursion	0.914727	0.663838	0.346173	0.000028
15	Random Excursion Variant	0.238264	0.133576	0.000080	0.000000

Из определения $P_v(S)$ статистики S следует, что его можно интерпретировать в качестве случайной величины со значениями в интервале $[0, 1]$.

$$\begin{aligned}\mathbf{P}\{P_v(S) \leq x \mid \mathcal{H}_0\} &= \mathbf{P}\{1 - x \leq F(S) \mid \mathcal{H}_0\} = \mathbf{P}\{F^{-1}(1 - x) \leq S \mid \mathcal{H}_0\} \\ &= 1 - F(F^{-1}(1 - x)) = x.\end{aligned}$$

Определим функцию распределения F_n статистики S_n следующим (удобным для дальнейшего изложения) образом

$$F_n(x) = \mathbf{P}\{S_n < x \mid \mathcal{H}_0\}.$$

Пусть S_n — статистика, дискретная функция распределения которой при условии $\mathcal{H}_0$ известна, множество всех r возможных значений статистики $E = \{e_1, \dots, e_r\}$ упорядочено по возрастанию, $e_i < e_{i+1}$, $i \in \{1, \dots, r-1\}$, $\mathbf{P}\{S_n = e_i \mid \mathcal{H}_0\} = p_i$, $i = 1, \dots, r$.

$$\mathbf{P}\{P_v(S_n) = x \mid \mathcal{H}_0\} = \begin{cases} 0, & \text{если } x \notin \{1 - F_n(e_2), \dots, 1 - F_n(e_r)\}, \\ p_1, & \text{если } x = 1 - F_n(e_1) = \sum_{i=1}^r p_i = 1, \\ p_2, & \text{если } x = 1 - F_n(e_2) = \sum_{i=2}^r p_i, \\ \dots, & \dots \quad \dots \quad \dots \\ p_r, & \text{если } x = 1 - F_n(e_r) = p_r. \end{cases}$$

Тогда множество значений функции распределения p -значения $P_v(S_n)$ статистики S_n совпадает с множеством значений её аргумента в $(r+1)$ точках, а именно,

$$\mathbf{P}\{P_v(S_n) \leq x \mid \mathcal{H}_0\} = x$$

при $x \in \{0, 1 - F_n(e_r), 1 - F_n(e_{r-1}), \dots, 1 - F_n(e_2), 1 - F_n(e_1)\}$.

Пусть $\varepsilon = \varepsilon_1, \varepsilon_2, \dots, \varepsilon_n$, $\varepsilon_i \in \{0, 1\}$, $i \in \{1, \dots, n\}$, длины n .

Статистикой Частотного теста является

$$|S_n| = |X_1 + \dots + X_n| = |2(\varepsilon_1 + \dots + \varepsilon_n) - n|, X_i = 2\varepsilon_i - 1,$$

и $\mathcal{H}_0$ состоит в том, что последовательность ε состоит из независимых одинаково распределенных бернуллиевских случайных величин.

Очевидно, что $|S_n| \in \{0, 2, 4, \dots, n\}$ при чётном n и

$$\mathbf{P}\{|S_n| = 2r \mid \mathcal{H}_0\} = \begin{cases} \frac{1}{2^n} \binom{n}{n/2}, & \text{если } r = 0, \\ \frac{1}{2^{n-1}} \binom{n}{n/2+r}, & \text{если } r \in \{1, \dots, n/2\}. \end{cases}$$

Для любого $x \geq 0$ функция распределения $P_v(|S_n|)$ имеет вид

$$\mathbf{P}\{P_v(|S_n|) \leq x \mid \mathcal{H}_0\}$$

$$= \begin{cases} 0, & \text{если } 0 \leq x < 1 - F_n(n), \\ \frac{1}{2^{n-1}} \binom{n}{n/2}, & \text{если } 1 - F_n(n) = \frac{1}{2^{n-1}} \leq x < 1 - F_n(n-2), \\ 1 - F_n(n-2), & \text{если } 1 - F_n(n-2) \leq x < 1 - F_n(n-4), \\ \dots, & \dots \quad \dots \quad \dots \\ 1 - F_n(2), & \text{если } 1 - F_n(2) \leq x < 1 - F_n(0), \\ 1, & \text{если } 1 - F_n(0) = 1 \leq x. \end{cases}$$

Если $\mathcal{H}_0$ верна, то согласно ЦПТ для случайного блуждания S_n

$$\lim_{n \rightarrow \infty} F_{S_n}(z) = \Phi(z) = \frac{1}{\sqrt{2\pi}} \int_{-\infty}^z e^{-u^2/2} du,$$

где $F_{S_n}(z) = \mathbf{P} \left\{ \frac{S_n}{\sqrt{n}} < z \right\}$, значит, предельное распределение $|S_n|$ —

полуноормальное, т. е. $F_{|S_n|}(z) = 2F_{S_n}(z) - 1$, $F_{|S_n|}(z) = \mathbf{P} \left\{ \frac{|S_n|}{\sqrt{n}} \leq z \right\}$ при

$z \geq 0$, и

$$\lim_{n \rightarrow \infty} F_{|S_n|}(z) = 2\Phi(z) - 1.$$

Наблюдаемому значению $|S_n(\varepsilon)| = |X_1(\varepsilon_1) + \dots + X_n(\varepsilon_n)|$ статистики $|S_n|$ соответствует предельное p -значение

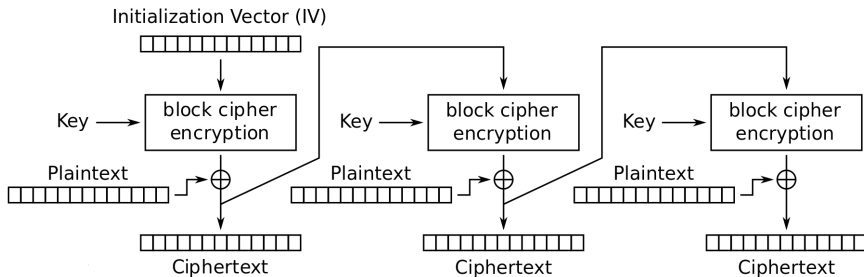
$$p_I(\varepsilon) = \lim_{n \rightarrow \infty} \left(1 - F_{|S_n|} \left(\frac{|S_n(\varepsilon)|}{\sqrt{n}} \right) \right) = 2 \left(1 - \Phi \left(\frac{|S_n(\varepsilon)|}{\sqrt{n}} \right) \right) = \operatorname{erfc} \left(\frac{|S_n(\varepsilon)|}{\sqrt{2n}} \right),$$

где $\operatorname{erfc}(x) = \frac{2}{\sqrt{\pi}} \int_x^{\infty} e^{-t^2}$

Поскольку распределение $2\Phi(z) - 1$, используемое в тесте в качестве приближения точного распределения статистики $|S_n|$, является непрерывным, постольку равномерное на $[0, 1]$ распределение является аппроксимирующим для распределения $P_v(|S_n|)$.

Тесты батареи NIST SP 800-22 были применены к последовательностям, сгенерированным с помощью AES.

Для генерации последовательностей использовался AES в режиме CFB (Cipher Feedback Block) с нулевыми вектором инициализации (IV) и блоком открытого текста (Plaintext); ключ для каждой последовательности выбирался случайным образом из множества 128-битных двоичных последовательностей, длина в битах всех последовательностей была выбрана одинаковой, равной $n = 2^{20} = 1\,048\,576$.



Значение α_{II} было выбрано равным $\alpha = \alpha_I = 0,01$.

Распределение чисел p -значений по $k = 10$ подинтервалам $C_1 = [0, \frac{1}{k})$, $C_2 = [\frac{1}{k}, \frac{2}{k})$, ..., $C_k = [\frac{k-1}{k}, 1)$ отрезка $[0, 1]$, полученным для значений статистики $|S_n|$ по формулам выше для предельных p -значений, оказалось следующим:

C_1	C_2	C_3	C_4	C_5	C_6	C_7	C_8	C_9	C_{10}
104280	104904	105253	105403	103993	105003	104480	105722	104261	105277

и этим подинтервалам соответствовали следующие значения статистики $|S_n|$:

C_1	1686, 1688, ..., 1 048 576,	C_6	538, 540, ..., 690,
C_2	1314, 1316 ..., 1684,	C_7	396, 398, ..., 536,
C_3	1062, 1064, ..., 1312,	C_8	260, 262, ..., 394,
C_4	862, 864, ..., 1060,	C_9	130, 132 ..., 258,
C_5	692, 694, ..., 860,	C_{10}	0, 2, ..., 128.

Значение статистики критерия согласия $\chi_{obs}^2 = \sum_{i=1}^k \frac{(\pi_i - N/k)^2}{N/k}$, где $N = 2^{20}$, $k = 10$, π_i — число p -значений в C_i , $i = 1, 2, \dots, k$, равно 28.42, p -значение p_{II} , соответствующее этому значению статистики χ_{obs}^2 ,

$$p_{II} \approx 0.000812,$$

предполагающее, при $\alpha_{II} = 0.01$, отклонение гипотезы $\mathcal{H}_0$.

Обозначим через $\lfloor x \rfloor = \max\{m \in \mathbb{Z} \mid m \leq x\}$, $\lceil x \rceil = \min\{n \in \mathbb{Z} \mid n \geq x\}$

Тогда по определению квантиля $x_\alpha(n, p)$ уровня α биномиальной случайной величины $X_{n,p}$

$$x_\alpha(n, p) = \min \{k: \mathbf{P}\{X_{n,p} \leq k\} \geq \alpha\} \in \{0, 1, \dots, n\},$$

$$x_0(n, p) = 0 \text{ и } x_1(n, p) = n, \quad x_{1/2}(n, 1/2) = \lfloor n/2 \rfloor.$$

Следствие. Если $0 < \alpha < 1/2$, то справедливы следующие оценки

$$\left\lceil \frac{n}{2} + \frac{\sqrt{n}}{2} \Phi^{-1}(\alpha) \right\rceil - 1 \leq x_\alpha(n, 1/2) \leq \left\lceil \frac{\frac{7n}{4} + \frac{3}{8} \Phi^{-1}(\alpha)^2 + \frac{1}{4} \Phi^{-1}(\alpha) \sqrt{42n + (7/2)^2 \Phi^{-1}(\alpha)^2}}{\frac{7}{2} + \Phi^{-1}(\alpha)^2/n} \right\rceil,$$

если $1/2 < \alpha < 1$, то

$$\left\lceil \frac{\frac{7n}{4} + \frac{5}{8} \Phi^{-1}(\alpha)^2 + \frac{1}{4} \Phi^{-1}(\alpha) \sqrt{42n + (7/2)^2 \Phi^{-1}(\alpha)^2}}{\frac{7}{2} + \Phi^{-1}(\alpha)^2/n} \right\rceil - 1 \leq x_\alpha(n, 1/2) \leq \left\lceil \frac{n}{2} + \frac{\sqrt{n}}{2} \Phi^{-1}(\alpha) \right\rceil,$$

если $\alpha = 1/2$, то

$$x_{1/2}(n, 1/2) = \begin{cases} (n-1)/2, & \text{если } n \text{ — нечётно,} \\ n/2, & \text{если } n \text{ — чётно.} \end{cases}$$

Оценки $|x_\alpha(n, 1/2) - n/2|$

α	нижняя оценка	верхняя оценка
0.45	59	65
0.4	120	130
0.35	182	198
0.3	248	269
0.25	319	346
0.2	398	431
0.15	491	531
0.1	607	657
0.05	779	843

α	нижняя оценка	верхняя оценка
0.55	59	65
0.6	120	130
0.65	182	198
0.7	248	269
0.75	319	346
0.8	398	431
0.85	491	531
0.9	607	657
0.95	779	843

Точное соответствие между подинтервалами C_i и значениями статистики $|S_n|$

C_1	1686, 1688, ..., 1 048 576,
C_2	1314, 1316 ..., 1684,
C_3	1064, 1066, ..., 1312,
C_4	864, 866, ..., 1062,
C_5	692, 694, ..., 862,

C_6	538, 540, ..., 690,
C_7	396, 398, ..., 536,
C_8	262, 264, ..., 394,
C_9	130, 132 ..., 260,
C_{10}	0, 2, ..., 128.

Соответствие между подинтервалами C_i и предельными p -значениями статистики $|S_n|$

C_1	1686, 1688, ..., 1 048 576,
C_2	1314, 1316 ..., 1684,
C_3	1062, 1064, ..., 1312,
C_4	862, 864, ..., 1060,
C_5	692, 694, ..., 860,

C_6	538, 540, ..., 690,
C_7	396, 398, ..., 536,
C_8	260, 262, ..., 394,
C_9	130, 132 ..., 258,
C_{10}	0, 2, ..., 128.

Прежнее распределение p -значений по подинтервалам C_i , $i = 1, 2, \dots, 10$,

C_1	C_2	C_3	C_4	C_5	C_6	C_7	C_8	C_9	C_{10}
104280	104904	105253	105403	103993	105003	104480	105722	104261	105277

было скорректировано согласно точному соответствию между C_i и значениями статистики $|S_n|$

C_1	C_2	C_3	C_4	C_5	C_6	C_7	C_8	C_9	C_{10}
104280	104904	104326	105160	105163	105003	104480	104122	105861	105277

Числа p -значений в 3-м, 4-м и 8-м подинтервалах были уменьшены на 927, $243=(1170-927)$ и 1600 соответственно, а в 5-м и 9-м — увеличены на 1170 и 1600 соответственно.

Значение $\chi^2_{obs} = \sum_{i=1}^k \frac{(\pi_i - N/k)^2}{N/k}$, где $N = 2^{20}$, $k = 10$, отвечающее новым числам p -значений π_i в C_i , приблизительно равно 25.66 и новое p -значение p_{II} —

$$p_{II} \approx 0.00232,$$

которое оказывается больше прежнего значения 0.000812, хотя $\alpha_{II} = 0.01$ новое p -значение снова предполагает отклонение гипотезы $\mathcal{H}_0$.

Проведённые эксперименты опровергают утверждение о ненадёжности двухуровневого подхода к тестированию вследствие систематической погрешности аппроксимации

Корректировки только частот π_i при определении значения статистики χ_{obs}^2 недостаточно для надёжности двухуровневого подхода. Так как из сходимости функции распределения p -значения $P_v(S_n)$ к равномерному распределению (по длине n тестируемых последовательностей)

$$\lim_{n \rightarrow \infty} \mathbf{P} \{P_v(S_n) \leq x \mid \mathcal{H}_0\} = x$$

не следует (при фиксированном n) сходимость этого распределения по числу N тестируемых отрезков, то утверждать связь ненадёжности двухуровневого подхода с погрешностью аппроксимации ошибочно и с теоретической точки зрения.

Полного устранения несоответствия между теоретическим α и эмпирическим уровнями значимости тестов при двухуровневом подходе к тестированию можно добиться путём следующей корректировки

$$\chi_{obs,cor}^2 = \sum_{i=1}^k \frac{(\pi_i - \mu_i)^2}{\mu_i},$$

где $\mu_i = N(F_n(r_i) - F_n(r_{i-1}))$, r_i — минимальное значение $|S_n|$, отвечающее подинтервалу C_i , π_i — числа точных p -значений в C_i , $i = 1, 2, \dots, k$, $k = 10$.

Значения μ_i/N при $n = N = 2^{20}$, вычисленные по формулам для $F_n(\cdot)$

μ_1/N	μ_2/N	μ_3/N	μ_4/N	μ_5/N	μ_6/N	μ_7/N	μ_8/N	μ_9/N	μ_{10}/N
0.0998651	0.0998981	0.0994669	0.100125	0.100445	0.100191	0.099697	0.099126	0.100937	0.10025

Значение скорректированной статистики

$$\chi_{obs,cor}^2 = \sum_{i=1}^{10} \frac{(\pi_i - \mu_i)^2}{\mu_i},$$

где $\mu_i = N(F_n(r_i) - F_n(r_{i-1}))$, r_i — минимальное значение $|S_n|$, отвечающее подинтервалу C_i , $i = 1, 2, \dots, 10$, значения π_i взяты из

C_1	C_2	C_3	C_4	C_5	C_6	C_7	C_8	C_9	C_{10}
104280	104904	104326	105160	105163	105003	104480	104122	105861	105277

приблизительно равно 3.19, а соответствующее ему p -значение

$$p_{II} \approx 0.956221,$$

которое при $\alpha_{II} = 0.01$ говорит о необходимости принятия гипотезы $\mathcal{H}_0$.