Тесты NIST (National Institute of Standards and Technology)

Статистические критерии NIST – Statistical Testing Suite of Random Number Generators или NIST STS – программный пакет, разработанный лабораторией информационных технологий (ITL NIST). Всего в составе пакета 15 тестов, предназначенных для проверки бинарных последовательностей. Базовыми идеями этих тестов являются тесты, описанные Д. Кнутом, а также тесты, разработанные Дж. Марсальей в пакете DIEHARD.

В основе тестов лежит понятие нулевой гипотезы. Нулевая гипотеза – принимаемое по умолчанию предположение о том, что не существует связи между наблюдаемыми событиями. Таким образом, нулевая гипотеза считается верной, пока нельзя доказать обратное.

Существует также альтернативная гипотеза, которая опровергает нулевую гипотезу, то есть она утверждает, что между явлениями взаимосвязь существует.

В терминах случайных чисел за нулевую гипотезу принимается предположение, что последовательность является истинно случайной (знаки которой появляются равновероятно и независимо друг от друга). Следовательно, если нулевая гипотеза верна, то рассматриваемый генератор производит достаточно «хорошие» случайные числа.

Итак, с одной стороны, статистика, подсчитанная на основе фактически собранных данных (то есть рассматриваемая последовательность). С другой стороны, есть эталонная статистика, получаемая математическими методами (теоретически вычисленная), которую бы имела бы истинно случайная последовательность. Очевидно, что собранная статистика не может сравняться с эталонной. Поэтому вводят некую погрешность, например 5%. То есть, если собранная статистика отклоняется от эталонной больше, чем на 5%, то делается вывод о том, что нулевая гипотеза не верна с большой надежностью.

Существует 4 варианта:

1) Сделан вывод о том, что последовательность случайна, и это верный вывод.

2) Сделан вывод о том, что последовательность не случайна, хотя она была на самом деле случайна. Такие ошибки называют ошибками первого рода.

3) Последовательность признана случайной, хотя на самом деле таковой не является. Такие ошибки называют ошибками второго рода.

4) Последовательность справедливо отбракована.

В каждом тесте вычисляется P_{value} ($value$ – значение). Это вероятность того, что последовательность не хуже, чем гипотетически истинная. Если $P_{value} = 1$, то тестируемая последовательность считается абсолютно случайной. Если $P_{value} = 0$, то тестируемая последовательность считается абсолютно неслучайной (полностью предсказуема).

Для теста выбирается уровень значимости α. Если $P_{value} \geq \alpha$, то нулевая гипотеза принимается, то есть последовательность признается случайной. Если $P_{value} < \alpha$, то нулевая гипотеза отвергается, то есть последовательность признается неслучайной и отбрасывается.

Обычно $\alpha \in [0{,}001; 0{,}01]$. То есть значение $\alpha = 0{,}01$ означает, что одна последовательность из 100 будет отвергнута тестом, при том, что эта последовательность действительно случайна.

В тестах обычно берется $\alpha = 0{,}01$. Тогда:

- Если $P_{value} \geq 0{,}01$, то последовательность признается случайной с уровень доверия (с уверенностью) 99%.
- Если$P_{value} < 0{,}01$, то последовательность признается неслучайной с уровнем доверия 99%.