Байесовский подход оценки случайности на основе гипотез

Александр Розенкевич[1]

улица Адам, дом 4 квартира 3, Иерусалим, Израиль

Предложен байесовский метод для динамической оценки случайности последовательности экспериментальных данных на основе гипотез. Даны примеры тестирования генераторов псевдослучайных чисел.

Ключевые слова: байесовский метод.

Априорная вероятность выпадения орла или решки равна 0.5. Поэтому при многократном бросании число выпадений орла или решки должно быть примерно одинаковое. Если в пятикратном бросании пять раз выпадал орел, в следующем броске интуитивно ожидается решка. Бросание монеты физический процесс, все влияющие факторы учесть невозможно. Ранее считалось, что идеальным генератором случайных чисел является подбрасывание «честной» монеты, но исследователи [1] в 2023 году (!) в результате 350 757 испытаний доказали, что как правило, монеты падают той же стороной, с которой они начинали.

В данной статье рассматривается возможность применения байесовского метода, основанного на экспериментальных данных и наблюдениях, для простого и надёжного выявления и оценки расхождений с теорией, что, в свою очередь, может направить поиск физических причин таких расхождений.

Наблюдения за бросками монет (событие B) можно учесть при расчете апостериорной вероятности, подтверждающей или не подтверждающей гипотезу выпадения той или иной стороны монеты при следующем N+1 бросании. В байесовском методе апостериорная вероятность равна:

$$P(R/B)=\frac{p(R)\cdot p(B/R)}{p(R)\cdot p(B/R)+p(O)\cdot p(B/O)} \qquad (1)$$

$$P(O/B)=\frac{p(O)\cdot p(B/O)}{p(R)\cdot p(B/R)+p(O)\cdot p(B/O)} \qquad (2)$$

Здесь: $P(R/B)$ и $P(O/B)$ условные (апостериорные) вероятности подтверждения гипотез, а именно: при следующем N+1 бросании монеты выпадет соответственно решка или орел, при условии, что результаты N наблюдений (событие B) истинны;

– $p(R)$ и $p(O)$ – априорные вероятности выпадения соответственно решке и орла;

– $p(B/R)$ и $p(B/O)$ – условные вероятности, что произойдут события соответственно R и O.

В общем случае для многогранника с S сторонами, апостериорная вероятность гипотезы выпадения грани S_k при N+1 наблюдении равна:

$$P(S_k/B_k) = \frac{p(S_k)\cdot p(B_k/S_k)}{\sum_{1}^{S} p(S_k)\cdot p(B_k/S_k)} \qquad (3)$$

Здесь: $p(S_k)$ - априорная вероятность грани S_k,, при правильном многограннике - 1/ S ;

$p(B_k/S_k)$ -влияние результатов наблюдений на подтверждения гипотезы :

$$p(B_k/S_k) = \sum_{i=1}^{N} \frac{1}{(S_{ki}+1)} \qquad (4)$$

где N – число наблюдений грани S_k при N испытаниях.

S_{ki} =1 когда грань наблюдается и S_{ki} = 0, когда нет.

Обратная частота в формуле (4) предполагает, что данная грань не должна появляться в следующем наблюдении, а единица в знаменателе подтверждает гипотезу через рекурсию.

На рисунке 1 приведены графики результатов тестирования на C++ популярных генераторов псевдослучайных чисел (имитация подбрасывания «честной» монеты) : Xorosshiro128+, Linear_Congruential, Squares_RNG и Mersenne_Twister. По оси X отложено число наблюдений, всего их 100, а по оси У отношения натуральных логарифмов вероятности гипотез для двух сторон — Ln(P1/P2), исчисленных по формулам (3) и (4). Априорная вероятность для обеих граней принята равной 0.5. ,

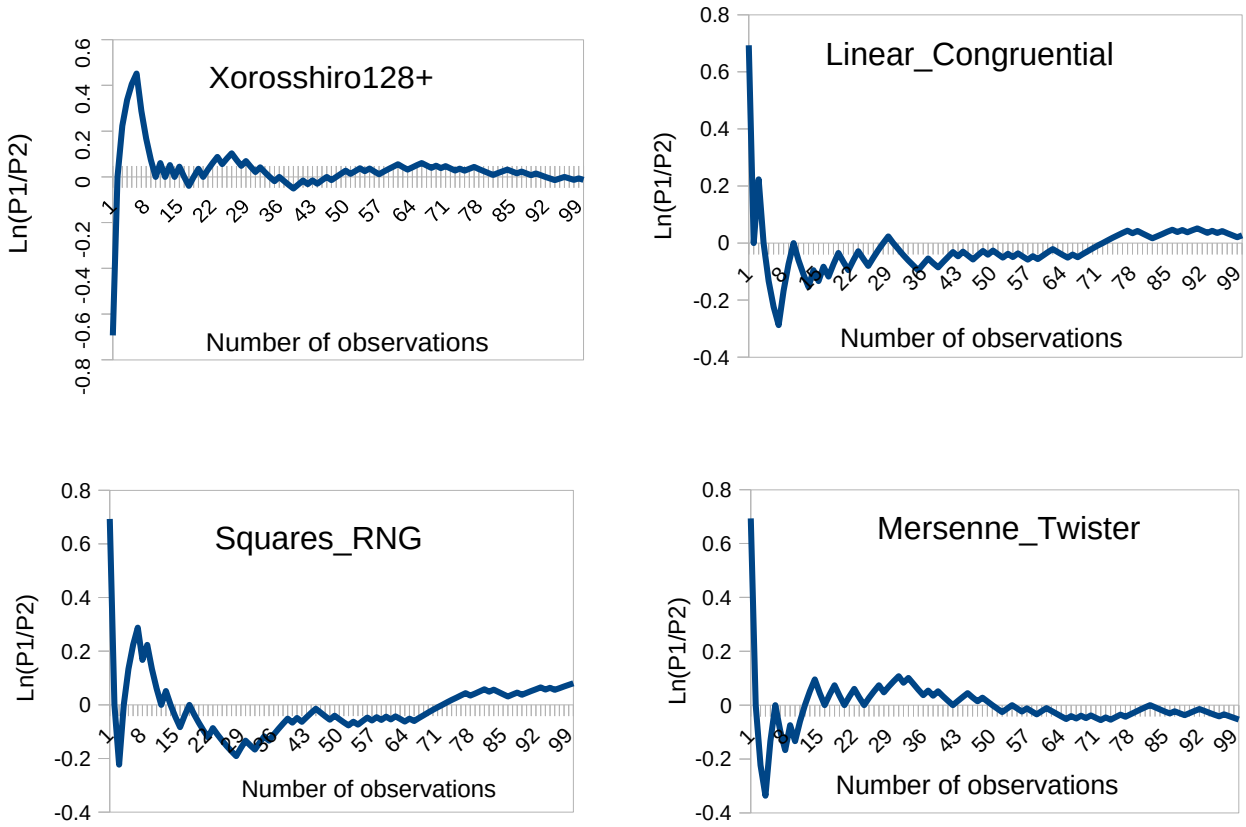

Рис .1 Зависимость логарифма отношения вероятностей гипотез от числа наблюдений

На рисунке 2 представлены коррелограммы , графики функции зависимости коэффициентов автокорреляции (ACF) логарифмов вероятности гипотез от лага.

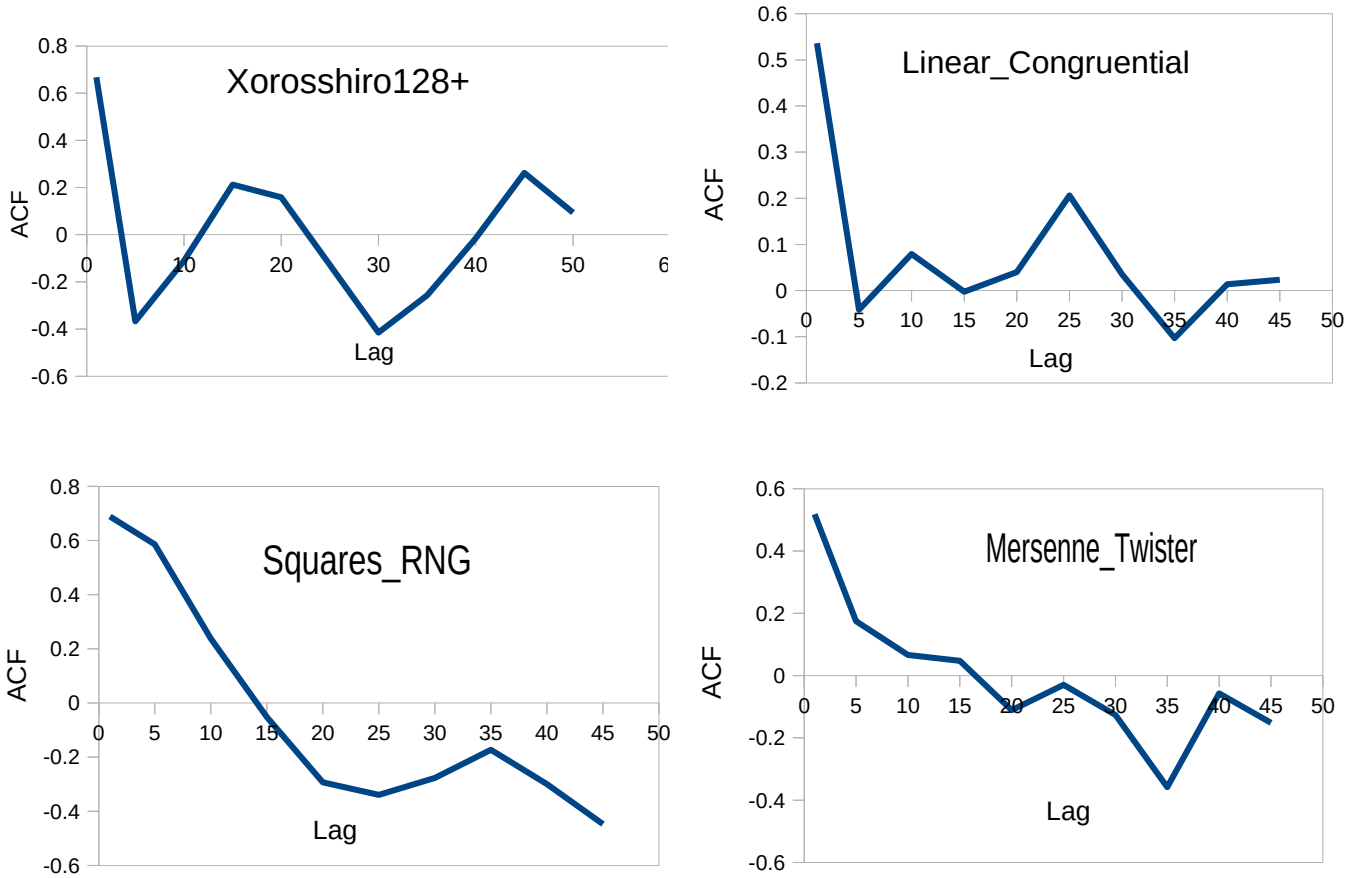

Рис.2 Коррелограммы, автокорреляции отношений логарифмов гипотез (ACF).

В таблицу сведены результаты проверки гипотез на ACF , выполнения закона повторного логарифма (закон LIL) [5], частоту смены знака и относительного *псевдопериода случайности (локальный период эргодичности)*. Закон повторного логарифма проверялся по формуле:

$$\frac{\left|\sum \ln(P1/P2)\right|}{\sigma\sqrt{2n\ln\ln(n)}}$$

здесь: σ^{2-} дисперсия, n- число наблюдения, в нашем случае — 100.

Генераторы	*ACF*	σ^2	*Закон LIL*	*Частота смены знака*	Период эргодичности
Xoroshiro128+	≠ 0	*0.012*	1.6363	*26*	*1*
Linear Congruential	≠ 0	*0.093*	1.3050	*12*	*0.46*
Mersenne_Twister	≠ 0	*0.089*	1.6507	*20*	*0.77*
Squares _RNG	≠ 0	*0.013*	0.5728	*24*	*0.92*

Согласно таблицы, проверка автокорреляции (ACF) указывает на отсутствие случайности в последовательности логарифмов гипотез, а несоответствие закону повторного логарифма (LIL) свидетельствует о связи внутренних параметров двух взаимоисключающих гипотез . Отсутствия случайных параметров в гипотезах, их закономерность, позволяют, на наш взгляд, оценивать эффективность и качество работы генераторов по стабильности осцилляций логарифмов гипотез (рис. 2).

Перечислим признаки наличия структур в последовательности случайных чисел:

- чем ниже дисперсия блуждающих логарифмов (закон LIL), то есть чем меньше отклонения от 1 , тем больше предсказуемости в числах.

- низкая частота смены знака при больших амплитудах указывает , что есть скрытая структура; и наоборот, высокая частота смены знака — признак высокого псевдопериода случайности, или другими словами больше локальный период эргодичности.

- если осцилляции долго остаются выше или ниже нуля, значит, одна гипотеза «доминирует» над другой, что говорит о несбалансированности генератора.

- в правильной случайной последовательности осцилляции логарифмов гипотез быстро пересекают ноль и не задерживаются в одной области, это говорит об отсутствии структуры, то есть об истинной случайности; и наоборот, длинные волны осцилляций указывают на тенденцию или тренд, характерный для скрытых или явных структур.

Основываясь на перечисленных выше критериях заключаем , что все тестируемые генераторы отвечают требованиям датчиков псевдослучайных чисел, но качество работы у них разное. Наилучшие результаты у Xoroshiro128+ и Squares _RNG.

Предложенный простой метод динамического анализа осцилляций логарифма гипотез , основанный на данных экспериментов и наблюдений, неограничен тестированием генераторов.

Список литературы

[1]. Fair coins tend to land on the same side they started: Evidence from 350,757 flips arXiv:2310.04153v3

[2]. "Bayes Theorem - Formula, Statement, Proof | Bayes Rule". *Cuemath*. Retrieved 2023- 10-20

[3]. *"Bayes' Theorem: Introduction"*. *Trinity University*. Archived from the original on 21 August 2004. Retrieved 5 August 2014.

[4]. ttp://csrc.nist.gov/groups/ST/toolkit/rng/stats_tests.html

[5]. https://en.wikipedia.org/wiki/Law_of_the_iterated_logarithm

[6]. Вентцель Е.С. Теория вероятностей. М.:Высш.шк. 2001

[1] Email: alexroz2008@gmail.com