Теорема о наличии памяти у случайных последовательностей

Автор Reshetov вкл Thursday, 3 December 2015
Предположим, что имеется некий генератор случайных временных рядов, об алгоритме которого нам известно лишь то, что вероятность сгенерированного любого его последующего значения независима от любого предыдущего сгенерированного значения. Предположим, что данный генератор с интервалом в 1 единицу времени, выдаёт некую последовательность чисел:

$X_1, X_2, ...X_{t-1}, X_t$

Где:

X_i - некое значение случайного числа, сгенерированное генератором на i-й единице времени с момента его запуска.

t - текущее время в единицах времени с момента запуска генератора

Предположим, что нам известны два произвольных значения, сгенерированные в прошлом: X_{t-k} и X_{t-m}

При условии:

$0 \leq k \leq m \leq t$

Предположим, что имеется некий букмекер, который предлагает сделать ставку в момент t на разницу между значениями чисел $\delta = X_{t-m} - X_{t+n}$, где n - натуральное число, при условиях:

- Если мы делаем ставку на понижение, то размер нашего выигрыша будет равен δ в денежных единицах
- Если мы делаем ставку на повышение, то размер нашего выигрыша будет равен -δ в денежных единицах

Возможно ли получить математическое ожидание выше значения 0 денежных единиц при ставках у вышеуказанного букмекера, если нам заведомо известны значения двух произвольных предыдущих чисел ряда: X_{t-k} и X_{t-m}? Ведь очевидно, что значение числа ряда X_{t+n} нам пока неизвестно, т.к. генератор выдаст его в будущем только через n единиц времени.

Для начала попробуем доказать, что знание только лишь одного значения числа X_{t-m} явно недостаточно для получения положительного математического ожидания по правилам вышеприведённой игры. Хотя вроде бы и очевидно по вышеуказанным условиям, вероятности сгенерированных значений попарно независимы друг от друга. Но это пока ещё не доказано для математического ожидания.

Рассмотрим случай 1.1. при условии: $X_{t-m} = X_{t+n}$

Поскольку в этом случае $\delta = |X_{t-m} - X_{t+n}| = 0$, то очевидно, что математическое ожидание всегда тоже будет равно M = 0.

Рассмотрим случай 1.2. при условии: $X_{t-m} \neq X_{t+n}$

Найдём минимум от значений X_{t-m} и X_{t+n}

$V_{min} = \min(X_{t-m}, X_{t+n})$

Найдём максимум от значений X_{t-m} и X_{t+n}

$V_{max} = \max(X_{t-m}, X_{t+n})$

Вычислим между ними разницу:

$dV = V_{max} - V_{min}$

В таком случае равновозможны всего лишь два несовместных варианта, приведенные в нижеуказанной таблице по строкам:

Номер варианта	Значение X_{t-m}	Значение X_{t+n}	Разница: $X_{t-m} - X_{t+n}$
1	V_{min}	V_{max}	-dV
2	V_{max}	V_{min}	dV

Поскольку вышеуказанные варианты 1 и 2 равновозможны, т.е. p(1) = p(2) = ½, то математическое ожидание:

$M = (dV - dV) / 2 = 0$

Вывод: Знание значения единственного произвольной случайной величины случайного временного ряда не даёт никакого преимущества в математическом ожидании разницы между этим значением и любым заведомо неизвестным числом этого же ряда при условии попарной независимости значений всех чисел в этом же ряду.

С двумя числами случайногов временного ряда, значение одного из которых нам заведомо известно, а второго неизвестно, мы разобрались. Теперь перейдём к трём числам, значения двух из которых нам заведомо известны, а третьего неизвестно.

Рассмотрим 2.1. случай, когда $X_{t-m} = X_{t-k} = X_{t+n}$

Очевидно, что всегда δ = 0, а следовательно и M = 0

Рассмотрим 2.2. случай, когда два значения равны между собой, а третье не равно ни одному из остальных двух значений

Т.е. может иметь место только один из трёх нижеприведённых равновозможных вариантов:

1. $X_{t-m} = X_{t-k}$ и $X_{t+n} \neq X_{t-m}$ и $X_{t+n} \neq X_{t-k}$
2. $X_{t-m} \neq X_{t-k}$ и $X_{t+n} = X_{t-m}$ и $X_{t+n} \neq X_{t-k}$
3. $X_{t-m} \neq X_{t-k}$ и $X_{t+n} \neq X_{t-m}$ и $X_{t+n} = X_{t-k}$

Найдём минимум от значений X_{t-m}, X_{t-k} и X_{t+n}

$V_{min} = \min(X_{t-m}, X_{t-k}, X_{t+n})$

Найдём максимум от значений X_{t-m}, X_{t-k} и X_{t+n}

$V_{max} = \max(X_{t-m}, X_{t-k}, X_{t+n})$

Вычислим между ними разницу:

$dV = V_{max} - V_{min}$

У нас могут быть всего шесть равновозможных и попарно несовместных вариантов:

Номер варианта	Значение X_{t-k}	Значение X_{t-m}	Значение X_{t+n}	Разница:
1	V_{max}	V_{max}	V_{min}	dV

Номер варианта	Значение $X_{t - k}$	Значение $X_{t - m}$	Значение $X_{t + n}$	Разница:
2	V_{max}	V_{min}	V_{max}	-dV
3	V_{min}	V_{max}	V_{max}	0
4	V_{max}	V_{min}	V_{min}	0
5	V_{min}	V_{max}	V_{min}	dV
6	V_{min}	V_{min}	V_{max}	-dV

И тут мы видим две интересные закономерности:

1. Если $X_{t - k} > X_{t - m}$, то $X_{t - m} \le X_{t + n}$
2. Если $X_{t - k} < X_{t - m}$, то $X_{t - m} \ge X_{t + n}$

Поскольку значения $X_{t - k}$ и $X_{t - m}$ нам заведомо известны, то в данном случае мы можем без всякого риска сделать ставку на знак разницы: $X_{t - m} - X_{t + n}$.

Что касается условия $X_{t - k} = X_{t - m}$, то в силу того, что нам заведомо неизвестно являются ли значения этих случайных величин максимальными или минимальными из трёх возможных, то математическое ожидание значения разницы $|X_{t - m} - X_{t + n}|$ всегда будет нулевым, т.к. варианты с номерами 1 и 6 равновозможны.

Вычислим матожидание для вариантов, когда $X_{t - k} \ge X_{t - m}$, т.е. для вариантов с номерами 2 и 4:

$M(X_{t - k} > X_{t - m}) = (-dV + 0) / 3 = -dV / 3 < 0$

Вычислим матожидание для вариантов, когда $X_{t - k} < X_{t - m}$, т.е. для вариантов с номерами 3 и 5:

$M(X_{t - k} > X_{t - m}) = (dV + 0) / 3 = dV / 3 > 0$

Ну и конечно же:

$M(X_{t - k} = X_{t - m}) = 0$

Рассмотрим 2.3. случай, когда все известные и неизвестное значения случайных величин попарно не равны друг другу

Найдём минимум от значений $X_{t - m}$, $X_{t - k}$ и $X_{t + n}$

$V_{min} = min(X_{t - m}, X_{t - k}, X_{t + n})$

Найдём максимум от значений $X_{t - m}$, $X_{t - k}$ и $X_{t + n}$

$V_{max} = max(X_{t - m}, X_{t - k}, X_{t + n})$

Найдём среднее из значений $X_{t - m}$, $X_{t - k}$ и $X_{t + n}$

$V_{mid} = mid(X_{t - m}, X_{t - k}, X_{t + n})$

Вычислим между ними разницу:

$dV_a = V_{mid} - V_{min}$

$dV_b = V_{max} - V_{mid}$

$dV_c = V_{max} - V_{min} = dV_a + dV_b$

У нас могут быть всего шесть равновозможных и попарно несовместных вариантов:

Номер варианта	Значение $X_{t - k}$	Значение $X_{t - m}$	Значение $X_{t + n}$	Разница:
1	V_{max}	V_{mid}	V_{min}	dV_a
2	V_{max}	V_{min}	V_{mid}	$-dV_a$
3	V_{mid}	V_{max}	V_{min}	$dV_a + dV_b$
4	V_{mid}	V_{min}	V_{max}	$-dV_a - dV_b$
5	V_{min}	V_{max}	V_{mid}	dV_b
6	V_{min}	V_{mid}	V_{max}	$-dV_b$

Вычислим матожидание для вариантов, когда $X_{t - k} > X_{t - m}$, т.е. для вариантов с номерами 1, 2 и 4:

$M(X_{t - k} > X_{t - m}) = (\cancel{dV_a} - \cancel{dV_a} - dV_a - dV_b) / 3 = - (dV_a + dV_b) / 3 < 0$

Вычислим матожидание для вариантов, когда $X_{t - k} < X_{t - m}$, т.е. для вариантов с номерами 3, 5 и 6:

$M(X_{t - k} > X_{t - m}) = (dV_a - dV_b + \cancel{dV_a} - \cancel{dV_a}) / 3 = (dV_a + dV_b) / 3 > 0$

Обобщив закономерности для случаев 2.1., 2.2. и 2.3. мы получаем закономерность:

$\mathbf{M(X_{t - n} - X_{t + m}) = p(X_{t - m} \neq X_{t - k}) \cdot signum(X_{t - m} - X_{t - k}) \cdot (V_{max} - V_{min}) \geq p(X_{t - m} \neq X_{t - k}) \cdot signum(X_{t - m} - X_{t - k}) \cdot |X_{t - k} - X_{t - m}|}$

Соответственно, если заведомо зная значения не менее двух предыдущих величин случайной временного ряда можно однозначно вычислить матожидания для разницы между последним известным значением числа этого ряда и пока ещё неизвестным значением того же самого ряда, то можно сделать окончательный вывод о том, что:

Случайные временные ряды обладают памятью на глубину не менее двух заведомо известных значений

P. S. Если бы было известно значение для $p(X_{t - m} \neq X_{t - k})$ то можно было бы вычислить и абсолютное значение этого матожидания, а оно нам неизвестно, т.к. мы не обладаем никакими сведениями об алгоритме генератора.

Юрий Решетов

Tags: Математика
Теория вероятностей

Source URL: http://yury-reshetov.com/node/36