

Лекция 12. Случайные процессы. Мартингалы

12 мая 2022 г.

В остатке курса мы будем изучать случайные процессы. Вообще говоря, случайным процессом называется любое не менее, чем счетное, упорядоченное множество случайных величин $\{X_t\}$. При этом с.в. в этом множестве могут зависеть или не зависеть друг от друга. Индекс t чаще всего интерпретируется как время, и в большинстве случаев принимает либо дискретные значения из \mathbb{N} , либо непрерывные из \mathbb{R}_+ .

Заметьте, что в случае дискретного $t \in \mathbb{N}$ множество $\{X_t\}$ является последовательностью. Мы начнем знакомство со случайными процессами с мартингалов. Очень грубо говоря, это такие процессы, матожидание изменения которых в единицу времени равно нулю. Но сначала нам надо ввести несколько дополнительных понятий.

1 Условное матожидание

У нас уже фигурировало матожидание, условное на событии или на другой с.в. В обоих случаях это было обычное матожидание, но с учетом какой-то дополнительной информации, которая у нас есть. Иногда эту информацию “кодируют” в σ -алгебре, и обуславливаются на ней.

Пусть у нас есть вероятностное пространство $(\Omega, \mathcal{F}_0, \Pr)^1$. Пусть также есть какая-то с.в. X (измеримая и интегрируемая по \mathcal{F}_0) и есть какая-то под-алгебра $\mathcal{F} \subset \mathcal{F}_0$. Тогда под $E[X \mid \mathcal{F}]$ мы понимаем любую с.в. Y , такую что

1. Y измерима по \mathcal{F}
2. Для любого множества $A \in \mathcal{F}$ выполняется равенство

$$\int_A X d\Pr = \int_A Y d\Pr$$

Оно существует и единственно почти наверное. Единственность почти наверное значит, что если какие-то две с.в. Y_1 и Y_2 удовлетворяют обоим условиям, то $\Pr(Y_1 \neq$

¹Заметьте, что мы обозначили сигма-алгебру буквой \mathcal{F} , а не Σ , но в этом нет особого скрытого смысла. Просто литература, по которой готовилась лекция, использовала это обозначение.

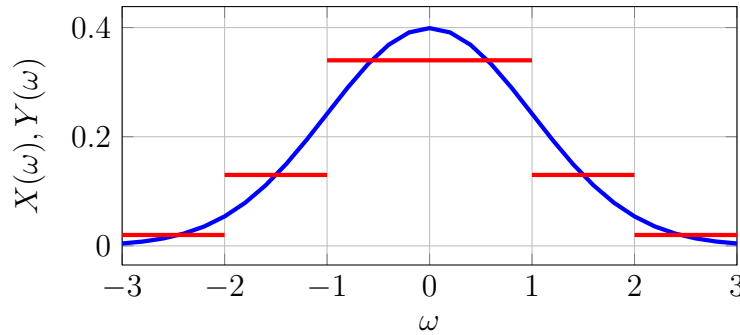
$Y_2) = 0$. Мы не будем доказывать существование и почти единственность, дабы не углубляться в теорию меры.

Как и в других условных вероятностях и матожиданиях, \mathcal{F} в условии стоит воспринимать как имеющуюся у нас информацию. А именно стоит трактовать это так: мы знаем интеграл с.в. на всех множествах из \mathcal{F} .

1.1 Примеры

Пример 1.

Для того, чтобы понять, как это работает, рассмотрим следующий пример. Пусть у нас дано вероятностное пространство $(\Omega, \mathcal{F}_0, \Pr)$, где $\Omega = \mathbb{R}$, а \mathcal{F}_0 — стандартная σ -алгебра, построенная на множестве всех полуинтервалов. Пусть \mathcal{F} — σ -алгебра, построенная на множестве полуинтервалов с концами в целых числах. Значит, если на исходном пространстве задана какая-то с.в. X , то когда мы говорим про $E[X | \mathcal{F}]$, мы говорим о какой-то с.в., которая может менять свое значение только в целых числах (иначе она будет не измерима по \mathcal{F}). Во всех остальных точках пространства она должна быть константой. На графике изображен пример, когда X — какая-то непрерывная с.в. (график нормального распределения взят просто для удобства, не путайте, это не плотность вероятности!), а Y , обозначенный красным, как раз является $E[X | \mathcal{F}]$.



Пример 2.

Пусть у нас случайная величина X измерима по \mathcal{F} . Это значит, что мы знаем ее значение на любом возможном событии, то есть у нас есть вся информация о ней. Поэтому $E[X | \mathcal{F}] = X$.

Пример 3.

Пусть с.в. X независима от \mathcal{F} . По определению независимости это значит, что \mathcal{F} не дает никакой информации об X . Формально это можно записать так: для любого события $A \in \mathcal{F}_0$ и $B \in \mathcal{F} \subset \mathcal{F}_0$ верно, что

$$\Pr(\{\omega : X(\omega) \in A\} \cap B) = \Pr(\{\omega : X(\omega) \in A\}) \Pr(B).$$

А так как мы нигде ничего не знаем о с.в., то наше лучшее предположение о с.в. — это ее среднее значение, то есть $E[X | \mathcal{F}] = E[X]$.

1.2 Свойства условного матожидания

Свойства все те же, что и у обычного. Мы просто перешли в мир, где нам известна какая-то новая информация.

- Линейность: $E[\alpha X + \beta Y \mid \mathcal{F}] = \alpha E[X \mid \mathcal{F}] + \beta E[Y \mid \mathcal{F}]$.
- Монотонность: $\forall \omega \ X(\omega) \leq Y(\omega) \Rightarrow E[X \mid \mathcal{F}] \leq E[Y \mid \mathcal{F}]$.
- Неравенство Йенсена: $\phi(x)$ — вогнута, тогда $\phi(E[X \mid \mathcal{F}]) \leq E[\phi(X) \mid \mathcal{F}]$.

И пара неочевидных, хотя ничего нового для нас тут нет

- Пусть $\mathcal{F}_1 \subset \mathcal{F}_2$. Тогда $E[E(X \mid \mathcal{F}_1) \mid \mathcal{F}_2] = E[E(X \mid \mathcal{F}_2) \mid \mathcal{F}_1] = E[X \mid \mathcal{F}_1]$.
- Если X измерима по \mathcal{F} , то $E[XY \mid \mathcal{F}] = X E[Y \mid \mathcal{F}]$.

2 Фильтрации

Фильтрацией называется неубывающая последовательность σ -алгебр в вероятностном пространстве. То есть у нас есть $(\Omega, \mathcal{F}, \text{Pr})$ и последовательность $\mathcal{F}_1 \subset \mathcal{F}_2 \subset \dots$, причем для любого $n \in \mathbb{N}$ выполняется $\mathcal{F}_n \subset \mathcal{F}$.

Фильтрацию стоит воспринимать как процесс набора информации, количество которой не убывает со временем. Если дана последовательность с.в. $\{X_n\}_{n \in \mathbb{N}}$, то *естественной фильтрацией* обычно называют

$$\mathcal{F}_n = \sigma(X_1, \dots, X_n),$$

то есть объединение σ -алгебр, задаваемых первыми n с.в. последовательности. В этом смысле запись $E[X_{n+1} \mid \mathcal{F}_n]$ означает с.в., равную X_{n+1} , когда нам уже все известно про X_1, \dots, X_n .

3 Мартингалы

Случайный процесс (последовательность с.в. $\{X_n\}_{n \in \mathbb{N}}$) называется *Мартингалом* относительно фильтрации $\{\mathcal{F}_n\}_{n \in \mathbb{N}}$, если выполнены три условия:

- Все X_n абсолютно интегрируемы, то есть $E[|X_n|] < \infty$
- Для всех n X_n измерима по \mathcal{F}_n
- Для всех n выполняется $E[X_{n+1} \mid \mathcal{F}_n] = X_n$

Последнее условие иногда записывается так:

$$E[X_{n+1} - X_n \mid \mathcal{F}_n] = 0$$

Случайный процесс называется *субмартингалом*, если в последнем условии стоит \geq . То есть следующая с.в. ожидаемо не меньше предыдущей. Случайный процесс называется *супермартингалом*, если в последнем условии стоит \leq . Да, казалось бы, что названия должны быть наоборот, но на самом деле “суб” относится к текущему наблюдению, которое не превышает ожидание следующего наблюдения $E[X_{t+1} \mid \mathcal{F}_n]$.

Чаще всего используется естественная фильтрация, которая означает, что мы знаем первые n шагов процесса. В таком случае вместо второго и третьего условия пишут просто:

$$E[X_{n+1} \mid X_1, \dots, X_n] = X_n,$$

что есть равенство с.в. То есть случайный процесс зависит только от своего прошлого.

Также возможен случай, когда один случайный процесс X_n является мартингалом относительно другого случайного процесса Y_n . Это происходит, когда вместо второго и третьего условия (в определении через фильтрацию) выполняется

$$E[X_{n+1} \mid Y_1, \dots, Y_n] = X_n,$$

Примеры мартингалов Самый простой пример: пусть есть последовательность $\{Y_n\}_{n \in \mathbb{N}}$ независимых с.в. с нулевыми матожиданиями, и есть X_n , который задается так:

$$X_n = \begin{cases} x, & \text{if } n = 1, \\ X_{n-1} + Y_{n-1}, & \text{otherwise.} \end{cases}$$

Возьмем фильтрацию $\mathcal{F}_n = \sigma(Y_1, \dots, Y_{n-1})$. По свойствам условного матожидания легко проверить, что это мартингал:

$$E[X_{n+1} \mid \mathcal{F}_n] = E[X_n \mid \mathcal{F}_n] + E[Y_n \mid \mathcal{F}_n] = X_n + E[Y_n] = X_n.$$

Второй пример. Пусть есть последовательность $\{Y_n\}_{n \in \mathbb{N}}$ независимых с.в. с единичными матожиданиями, и есть X_n , который задается так:

$$X_n = \prod_{i=1}^n Y_i,$$

он является мартингалом относительно $\mathcal{F}_n = \sigma(Y_1, \dots, Y_n)$:

$$E[X_{n+1} \mid \mathcal{F}_n] = E[X_n Y_{n+1} \mid \mathcal{F}_n] = X_n E[Y_{n+1} \mid \mathcal{F}_n] = X_n E[Y_{n+1}] = X_n.$$

Пара важных свойств (суб-/супер-)мартингалов:

Лемма 1. Если X_n — субмартингал относительно \mathcal{F}_n , то для любых $m > n$ выполняется $E[X_m | \mathcal{F}_n] \geq X_n$

Если X_n — супермартингал относительно \mathcal{F}_n , то для любых $m > n$ выполняется $E[X_m | \mathcal{F}_n] \leq X_n$

Если X_n — мартингал относительно \mathcal{F}_n , то для любых $m > n$ выполняется $E[X_m | \mathcal{F}_n] = X_n$

Доказательство. Первое доказывается по индукции. Пусть $m = n + k$. Для $k = 1$ утверждение верно по определению. Для больших k воспользуемся свойствами условного матожидания:

$$E[X_{n+k} | \mathcal{F}_n] = E[E[X_{n+k} | \mathcal{F}_{n+k-1}] | \mathcal{F}_n] \geq E[X_{n+k-1} | \mathcal{F}_n].$$

Для доказательства для супермартингала достаточно заметить, что $-X_n$ является субмартингалом, а для доказательства для мартингала достаточно заметить, что он является и суб-, и супер-мартингалом. \square

Лемма 2. Если X_n — мартингал относительно \mathcal{F}_n , а ϕ — вогнутая функция, то $\phi(X_n)$ — субмартингал. Если ϕ — неубывающая, то достаточно, чтобы X_n был субмартингалом.

Доказательство. По неравенству Йенсена в первом случае получаем:

$$E[\phi(X_{n+1}) | \mathcal{F}_n] \geq \phi(E[X_{n+1} | \mathcal{F}_n]) = \phi(X_n)$$

Во втором случае последнее равенство заменяется на неравенство (пользуемся неубыванием ϕ):

$$E[\phi(X_{n+1}) | \mathcal{F}_n] \geq \phi(E[X_{n+1} | \mathcal{F}_n]) \geq \phi(X_n)$$

\square

3.1 Предсказуемые с.в.

Последовательность с.в. $\{H_n\}_{n \in \mathbb{N}}$ называется предсказуемой относительно фильтрации $\{\mathcal{F}_n\}_{n \in \mathbb{N}}$, если для любого n H_n измерима относительно \mathcal{F}_{n-1} .

Классический пример — ставки игрока в казино. Если нам известна его стратегия, то когда мы знаем результат его первых n ставок, то мы обычно можем сказать, какой будет его $n + 1$ -ая ставка. Например, известная стратегия — удваивать ставку после проигрыша и прекращать игру после первой победы. Такая тактика (при бесконечном имеющимся капитале) гарантирует вам прибыль в размере первой ставки.

Для как раз таких стратегий часто используется следующая последовательность с.в.:

$$(H \cdot X)_n = \sum_{i=1}^n H_i(X_i - X_{i-1}),$$

Которая по сути равна балансу игрока, делающего ставки по стратегии H_i , где X_i — баланс игрока, который всегда делает единичную ставку.

Теорема 1. Если X_n — супремартингал, а $H_n \geq 0$ — предсказуемая, ограниченная последовательность с.в. Тогда $(H \cdot X)_n$ — супремартингал.

Доказательство. Так как $(H \cdot X)_n$ измеримо по \mathcal{F}_n , а H_n также измеримо по \mathcal{F}_{n-1} ,

$$\begin{aligned} E[(X \cdot H)_{n+1} | \mathcal{F}_n] &= (H \cdot X)_n + E[H_{n+1}(X_{n+1} - X_n) | \mathcal{F}_n] \\ &= (H \cdot X)_n + H_{n+1}E[(X_{n+1} - X_n) | \mathcal{F}_n] = (H \cdot X)_n. \end{aligned}$$

□

3.2 Ограничение прибыли

В данном разделе мы рассмотрим upcrossing inequality (ограничение на пересечения вверх/ограничение на прибыль), которое играет важную роль, например, в биржевых торгах.

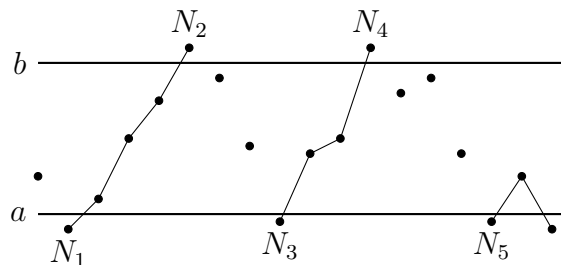
Для начала введем определение *времени останова*. Натуральная с.в. N называется временем останова относительно фильтрации $\{\mathcal{F}_n\}$, если для всех n событие $\{N = n\}$ измеримо относительно \mathcal{F}_n . Если мы введем индикаторную величину $H_n = [N \geq n]$, то она будет измерима относительно \mathcal{F}_{n-1} (так как событие $\{N \geq n\} = \{N \leq n-1\}^c \in \mathcal{F}_{n-1}$), а значит, она будет предсказуема.

Таким образом, если X_n — супермартингал, то $(H \cdot X)_n = X_{n \cap N} - X_0$ — супермартингал. Мы тут обозначили $X_{\min(n, N)} = X_{n \cap N}$ для более короткой записи. Таким образом, и $X_{n \cap N}$ — супермартингал (как сумма двух супермартингалов).

Введем теперь понятие upcrossing (пересечение вверх). Пусть у нас есть случайный процесс X_n , для простоты представим, что это стоимость акций в момент времени n . И у нас следующая стратегия: мы покупаем акции, когда их цена становится ниже какого-то a и продаем, когда их цена становится выше какого-то $b > a$. Тогда за время n наша суммарная прибыль будет $(b-a)$ умножить на число случаев, когда цена акций возрастала от a до b (что есть число пересечений полосы $[a, b]$ снизу вверх до времени n). Введем также времена, когда мы покупаем и продаем акции:

$$\begin{aligned} N_{2k-1} &= \inf\{m > N_{2k-2} : X_m \leq a\} \text{ (время покупки)} \\ N_{2k} &= \inf\{m > N_{2k-1} : X_m \geq b\} \text{ (время продажи)} \end{aligned}$$

и положим $N_0 = -1$. В таком случае upcrossing — это время от покупки до продажи (на рисунке они обозначены линиями).



Введем H_n — индикаторную величину, показывающую, на руках ли у нас акции в момент времени n , или другими словами,

$$H_n = \begin{cases} 1, & \text{если } n \text{ лежит между } N_{2k-1} \text{ и } N_{2k}, \\ 0, & \text{иначе.} \end{cases}$$

Нагонец, введем с.в. $U_n = \sup\{k : N_{2k} \leq n\}$ — число upcrossing'ов, завершенных к моменту времени n . Для него известно следующее утверждение

Теорема 2 (Upcrossing inequality). *Если X_m — субмартингал, тогда*

$$(b - a)E[U_n] \leq E[(X_n - a)^+] - E[(X_0 - a)^+]$$

Доказательство. Пусть $Y_m = a + (X_m - a)^+$. Это субмартингал по Лемме 2. Он пересекает полосу $[a, b]$ столько же раз, сколько X_m (единственное отличие — он не уходит ниже a). Тогда наша прибыль описывается как $(H \cdot Y)_n \geq (b - a)U_n$, так как в случае с Y последний начатый upcrossing не приносит убытка.

Пусть теперь $K_m = 1 - H_m$. Тогда $Y_n - Y_0 = (H \cdot Y)_n + (K \cdot Y)_n$. Так как $(K \cdot Y)_n$ — субмартингал, $(K \cdot Y)_n \geq (K \cdot Y)_0 = 0$. То есть $E[(H \cdot Y)_n] \leq E[Y_n - Y_0]$. Остается все подставить и воспользоваться линейностью матожидания. \square

Из этого результата следует другой, куда более общий результат о сходимости мартингалов.

Теорема 3. *Если X_n — субмартингал, причем $\sup E[X_n^+] < +\infty$, тогда X_n сходится почти наверное к какому-то X с $E[|X|] < \infty$*

Мы опустим доказательство из-за его излишней техничности, но в целом оно основано на том, что число upcrossing'ов любого отрезка $[a, b]$ конечно, откуда следует сходимость верхнего и нижнего пределов X_n .

3.3 Декомпозиция Дуба

Важным моментом, касающимся предсказуемых с.в. является декомпозиция Дуба.

Теорема 4 (Декомпозиция Дуба). *Пусть X_n — субмартингал. Тогда он может быть представлен как $X_n = M_n + H_n$, где M_n — мартингал, а H_n — предсказуемая неубывающая последовательность.*

Доказательство. Возьмем $A_0 = 0$ и каждый следующий A_n такой, что

$$A_{n+1} - A_n = E[X_{n+1} | \mathcal{F}_n] - X_n$$

Правая часть равенства измерима по \mathcal{F}_n , значит, по индукции можно доказать, что A_{n+1} измерима по \mathcal{F}_n , то есть является предсказуемой. Также заметим, что правая

часть неотрицательна (так как X_n — субмартингал), а значит, A_n — возрастающая последовательность.

Остается показать, что M_n — мартингал, то есть

$$\begin{aligned} E[M_n | \mathcal{F}_{n-1}] &= E[X_n - A_n | \mathcal{F}_{n-1}] \\ &= E[X_n | \mathcal{F}_{n-1}] - A_n = X_{n-1} - A_{n-1} = M_{n-1}. \end{aligned}$$

□

4 Ветвящиеся процессы

Ветвящиеся процессы в общем случае это процессы, описывающие бесполое размножение каких-то организмов с помощью генеалогических деревьев. Одним из простейших примеров является следующий процесс, называемый процессом Гальтона-Ватсона.

Пусть у нас есть двумерная последовательность независимых, одинаково распределенных с.в. $X_i^n \in \mathbb{N}$. И также есть процесс $\{Z_n\}_{n \in \mathbb{N}}$, описываемый следующим образом: $Z_0 = 1$ и

$$Z_{n+1} = \begin{cases} \sum_{i=1}^{Z_n} X_i^{n+1}, & \text{if } Z_n > 0, \\ 0, & \text{else.} \end{cases}$$

В данном случае X_i^n — это число потомков, которых породит i -ая особь из поколения n , если она вообще будет существовать. А сам Z_n отображает размер поколения.

Лемма 3. Пусть $\mathcal{F}_n = \sigma(X_i^m : i \geq 1, 1 \leq m \leq n)$ (фильтрация, знающая все про первые n поколений). И пусть $\mu = EX_i^n$. Тогда $\frac{Z_n}{\mu^n}$ мартингал относительно фильтрации \mathcal{F}_n .

Доказательство.

$$E[Z_{n+1} | \mathcal{F}_n] = E[X_1^{n+1} + \dots + X_k^{n+1} | \mathcal{F}_n] = k\mu = \mu Z_n.$$

Поэтому $E[\frac{Z_{n+1}}{\mu^{n+1}} | \mathcal{F}_n] = \frac{Z_n}{\mu^n}$.

□

Отсюда понятно, что Z_n ведет себя примерно как μ^n , откуда следуют следующие наблюдения.

Лемма 4. Если $\mu < 1$, то $Z_n = 0$ начиная с какого-то n почти наверное, то есть $\frac{Z_n}{\mu^n} \rightarrow 0$.

Доказательство.

$$\Pr[Z_n > 0] \leq E[Z_n] = \mu^n \rightarrow 0.$$

□

Лемма 5. Если $\mu = 1$ и $\Pr[X_i^n = 1] < 1$, тогда $Z_n = 0$ начиная с какого-то n почти наверное.

Доказательство. Z_n является мартингалом, значит, по теореме о сходимости он сходится к какой-то с.в. Так как Z_n целочисленна, то и предел целочисленный.

Пусть этот предел не ноль. Тогда

$$\Pr[Z_n = k \text{ for all } n \geq N] = 0,$$

поэтому единственный вариант предела есть ноль. \square

Для следующей леммы мы введем понятие производящей функции. Пусть $\phi(s) = \sum_{k \geq 0} p_k s^k$, где $p_k = \Pr[X_i^n = k]$

Лемма 6. Если $\mu > 1$ и $Z_0 = 1$, тогда $\Pr[Z_n = 0 \text{ начиная с какого-то } N] = \rho$, где ρ — единственное решение уравнения $\phi(\rho) = \rho \in [0, 1)$.

Оставим без доказательства (слишком сложно).

5 Полезные неравенства

5.1 Неравенство Азумы

Теорема 5. Пусть $\{X_n\}_{n \in \mathbb{N}}$ — супермартингал. Пусть также его изменение ограничено, то есть существует такая последовательность $\{c_n\}_{n \in \mathbb{N}}$, что для любого n верно, что $|X_{n+1} - X_n| \leq c_n$. Тогда для любой $\delta > 0$ верно, что

$$\Pr[X_n - X_0 \geq \delta] \leq \exp \left(-\frac{\delta^2}{2 \sum_{i=1}^{n-1} c_i} \right).$$

Очень легко переделать это в неравенство для субмартингала и в двустороннее неравенство для мартингала:

Теорема 6. Пусть $\{X_n\}_{n \in \mathbb{N}}$ — субмартингал. Пусть также его изменение ограничено, то есть существует такая последовательность $\{c_n\}_{n \in \mathbb{N}}$, что для любого n верно, что $|X_{n+1} - X_n| \leq c_n$. Тогда для любой $\delta > 0$ верно, что

$$\Pr[X_n - X_0 \leq -\delta] \leq \exp \left(-\frac{\delta^2}{2 \sum_{i=1}^{n-1} c_i} \right).$$

Теорема 7. Пусть $\{X_n\}_{n \in \mathbb{N}}$ — мартингал. Пусть также его изменение ограничено, то есть существует такая последовательность $\{c_n\}_{n \in \mathbb{N}}$, что для любого n верно, что $|X_{n+1} - X_n| \leq c_n$. Тогда для любой $\delta > 0$ верно, что

$$\Pr[|X_n - X_0| \geq \delta] \leq 2 \exp \left(-\frac{\delta^2}{2 \sum_{i=1}^{n-1} c_i} \right).$$

Данное неравенство доказывается через общие границы Чернова, примененные к $X_n - X_0 = \sum_{i=1}^{n-1} (X_{i+1} - X_i)$, затем выделению каждого члена суммы в отдельный множитель, который ограничивается через лемму Хёффдинга.

5.2 Неравенство МакДиармида

Теорема 8. Пусть X_1, \dots, X_n — независимые с.в., определенные на $\Omega_1, \dots, \Omega_n$ соответственно. Пусть $\Omega = \Omega_1 \times \dots \times \Omega_n$. Пусть также есть функция $f : \Omega \rightarrow \mathbb{R}$, для которой существует такой набор чисел c_1, \dots, c_n , что для любых двух векторов элементарных исходов $\omega, \nu \in \Omega$, отличающихся только в i -ом компоненте верно, что $|f(\omega) - f(\nu)| \leq c_i$. Пусть также $X = f(X_1, \dots, X_n)$. Тогда

$$\Pr[X - E[X] \geq \lambda] \leq \exp \left(-\frac{2\delta^2}{\sum_{i=1}^{n-1} c_i} \right),$$
$$\Pr[X - E[X] \leq -\lambda] \leq \exp \left(-\frac{2\delta^2}{\sum_{i=1}^{n-1} c_i} \right).$$