Случайные процессы

Конспект теории

Содержание

Источники:

1. семинары Черноусовой Е.О. в группе 776;
2. лекции Гасникова А.В. по случайным процессам на ФУПМе;
3. Гасников А.В. Лекции по случайным процессам;
4. Миллер Б.М., Панков А.Р. Теория случайных процессов в примерах и задачах.

Используемые обозначения:

- $\mathbb{E}\xi$ — матожидание случайной величины ξ (expected value);
- $\mathbb{V}\xi$ — дисперсия случайной величины ξ (variance);
- $\mathcal{N}(m, \sigma^2)$ — нормальное распределение со средним m и дисперсией σ^2;
- $\mathcal{U}[a, b]$ — равномерное распределение на отрезке $[a, b]$;
- $\mathrm{Exp}(\lambda)$ — экспоненциальное (показательное) распределение с параметром λ;
- $\mathrm{Poiss}(\lambda)$ — распределение Пуассона с параметром λ;
- $m_X(t) = \mathbb{E}X(t)$ — матожидание случайного процесса $X(t)$;
- $R_X(t, s) = \mathrm{cov}\big(X(t), X(s)\big)$ — корреляционная (ковариационная) функция случайного процесса $X(t)$;

1 Элементы теории вероятностей

1.1 Нормальное распределение

Опр. Случайный вектор ξ имеет *нормальное распределение* с параметрами $(\mathbf{m}, \Sigma)$, если его характеристическая функция имеет вид

$$\varphi_\xi(\mathbf{t}) = \exp\left(i\mathbf{t}^T\mathbf{m} - \frac{1}{2}\mathbf{t}^T\Sigma\mathbf{t}\right)$$

В случае, когда $\Sigma \succ 0$, существует функция плотности:

$$f(\mathbf{x}) = \frac{1}{\sqrt{(2\pi)^n \det\Sigma}} \exp\left(-\frac{1}{2}(\mathbf{x}-\mathbf{m})^T\Sigma^{-1}(\mathbf{x}-\mathbf{m})\right), \qquad \mathbb{E}\xi = \mathbf{m}, \qquad \mathbb{V}\xi = \Sigma$$

Свойства нормального случайного вектора:

1. Если X — нормальный случайный вектор, то $Y = AX + b$ — тоже. При этом
$$Y = AX + b, \quad X \sim \mathcal{N}(\mathbf{m}, \Sigma) \qquad \Longrightarrow \qquad Y \sim \mathcal{N}(A\mathbf{m} + b, A\Sigma A^T)$$
2. Пусть $X_i \sim \mathcal{N}(m_i, \sigma_i^2),\ i = \overline{1, n}$ — независимые нормальные случайные величины. Тогда
$$X_1 + \ldots + X_n \sim \mathcal{N}\big(m_1^2 + \ldots + m_n^2,\ \sigma_1^2 + \ldots + \sigma_n^2\big)$$
3. Пусть X, Y — подвекторы нормального случайного вектора:
$$\begin{bmatrix} X \\ Y \end{bmatrix} \sim \mathcal{N}\left(\begin{bmatrix} \mathbf{m}_X \\ \mathbf{m}_Y \end{bmatrix}, \begin{bmatrix} \Sigma_{XX} & \Sigma_{XY} \\ \Sigma_{YX} & \Sigma_{YY} \end{bmatrix}\right)$$
Тогда маргинальные распределения X и Y являются нормальными:
$$X \sim \mathcal{N}(\mathbf{m}_X, \Sigma_{XX}), \qquad Y \sim \mathcal{N}(\mathbf{m}_Y, \Sigma_{YY}),$$
и условное распределение вида $Y|X = x$ тоже является нормальным с параметрами $(\mathbf{m}, \Sigma)$:
$$\mathbf{m} = \mathbb{E}[Y \mid X = x] = \mathbf{m}_Y + \Sigma_{YX}\Sigma_{XX}^{-1}(x - \mathbf{m}_X), \qquad \Sigma = \mathbb{V}[Y \mid X = x] = \Sigma_{YY} - \Sigma_{YX}\Sigma_{XX}^{-1}\Sigma_{XY}$$
Здесь $\mathbb{E}[Y \mid X = x]$ и $\mathbb{V}[Y \mid X = x]$ — условные матожидания и дисперсия соответственно.
4. Для компонент $X_1, \ldots, X_n$ некоторого нормального случайного вектора:
$$X_1, \ldots, X_n \text{ независимы в совокупности} \qquad \Longleftrightarrow \qquad X_1, \ldots, X_n \text{ не коррелируют}$$
Для произвольных нормальных случайных величин данное свойство неверно.
5. Если $\xi \sim \mathcal{N}(0, \sigma^2)$, то
$$\mathbb{E}\xi^{2k+1} = 0, \qquad \mathbb{E}\xi^{2k} = (2k-1)!!\,\sigma^{2k}$$
6. ξ — нормальный случайный вектор $\iff$ $\forall c \in \mathbb{R}^n$ случайная величина $c^T\xi$ является нормальной (или константой).
7. $X_1, \ldots, X_n$ — нормальные случайные величины $\not\Longrightarrow$ $(X_1, \ldots, X_n)$ — нормальный случайный вектор.
Пример: пусть $X \sim \mathcal{N}(0,1),\ \xi = \pm 1$ равновероятно, ξ и X независимы, $Y = X\xi$. Тогда легко показать, что $Y \sim \mathcal{N}(0,1)$. Допустим, (X, Y) — нормальный случайный вектор, значит, по свойству 6, $X + Y$ — нормальная случайная величина, что не так (можно найти характеристическую функцию).

Теорема Вика

Пусть $(X_1, \ldots, X_n) \sim \mathcal{N}(0, R)$, где R — произвольная симметричная матрица. Тогда

- если n нечетно, то
$$\mathbb{E}\big[X_1 \cdots X_n\big] = 0$$
- если n четно, то
$$\mathbb{E}\big[X_1 \cdots X_n\big] = \sum R_{p_1 q_1} \cdots R_{p_n q_n},$$

где сумма берется по всем неупорядоченным разбиения $\{p_i\}$ и $\{q_i\}$ множества $\{1, \ldots, n\}$ на $\frac{n}{2}$ пар.

Пример: $(X_1, X_2, X_3, X_4) \sim \mathcal{N}(0, R) \implies \mathbb{E}X_1X_2X_3X_4 = R_{12}R_{34} + R_{13}R_{24} + R_{14}R_{23}$.

1.2 Экспоненциальное (показательное) распределение

Опр. Случайная величина ξ имеет *экспоненциальное распределение с параметром* λ ($\xi \sim \mathrm{Exp}(\lambda)$), если она имеет функцию плотности

$$f_\xi(x) = \begin{cases} \lambda e^{-\lambda x} & , \quad x \geq 0, \\ 0 & , \quad x < 0 \end{cases}, \qquad F_\xi(x) = 1 - e^{-\lambda x}$$

Опр. Случайная величина η имеет *гамма-распределение* с параметром λ и числом степеней свободы k, если она является суммой независимых экспоненциальных случайных величин:

$$\eta = \xi_1^2 + \ldots + \xi_k^2 \sim \mathrm{Gamma}(k, \lambda), \qquad \text{если } \ \xi_1, \ldots, \xi_k \sim \mathrm{Exp}(\lambda)$$

$$f_\eta(x) = \begin{cases} \dfrac{\lambda(\lambda x)^{n-1}}{(n-1)!} e^{-\lambda x} & , \quad x \geq 0, \\ 0 & , \quad x < 0. \end{cases}$$

Свойства экспоненциально распределенной случайной величины ($\xi \sim \mathrm{Exp}(\lambda)$):

1. $\mathbb{E}\xi = \dfrac{1}{\lambda}, \qquad \mathbb{V}\xi = \dfrac{1}{\lambda^2}$.
2. Характеристическая функция: $\varphi_\xi(t) = \left(1 - \dfrac{it}{\lambda}\right)^{-1}$
3. Если $\eta \sim \mathrm{Exp}(1)$, то $\xi = \dfrac{\eta}{\lambda} \sim \mathrm{Exp}(\lambda)$.
4. Если $w \in \mathcal{U}[0, 1]$, то $\xi = -\ln w \sim \mathrm{Exp}(\lambda)$.
   Таким образом можно генерировать экспоненциальные случайные величины.
5. *Отсутствие памяти:*
   $$\mathbb{P}\{\xi > x + y \mid \xi > y\} = \mathbb{P}\{\xi > x\}, \qquad \forall x, y > 0$$
6. Пусть $\xi_i \sim \mathrm{Exp}(\lambda_i)$, $i = \overline{1, n}$ — независимые. Тогда
   $$\eta = \min_{i=1,\ldots,n} \xi_i \sim \mathrm{Exp}(\lambda_1 + \ldots + \lambda_n)$$
   $$I = \arg\min_{i=1,\ldots,n} \xi_i \qquad \implies \qquad \mathbb{P}\{I = k\} = \frac{\lambda_k}{\lambda_1 + \ldots + \lambda_k}$$
   Кроме того, случайные величины η и I независимы.
7. Экспоненциальная случайная величина — единственная случайная величина в классе непрерывных случайных величин, обладающая свойством отсутствия памяти.
   В классе дискретных случайных величин таковой является геометрическая случайная величина.

Опр. *Гамма-функцией* называется функция $\Gamma(z)$:

$$\Gamma(z) = \int\limits_0^\infty t^{z-1} e^{-t} dt, \qquad \mathrm{Re}(z) > 0$$

Свойства гамма-функции:

$$\Gamma(z+1) = z\Gamma(z), \qquad \Gamma(n+1) = n!, \qquad \Gamma(z)\Gamma(1-z) = \frac{\pi}{\sin \pi z}$$

1.3 Сходимости случайных величин

Опр. Пусть $(\Omega, \mathcal{F}, \mathbb{P})$ — вероятностное пространство, $\{\xi_n\}_{n=1}^{\infty}$ — последовательность случайных величин на нем.

- $\{\xi_n\}_{n=1}^{\infty}$ *сходится по вероятности* к ξ, если

$$\xi_n \xrightarrow{\mathbb{P}} \xi \qquad \Longleftrightarrow \qquad \mathbb{P}\big\{|\xi_n - \xi| > \varepsilon\big\} \longrightarrow 0 \qquad \forall \varepsilon > 0$$

- $\{\xi_n\}_{n=1}^{\infty}$ *сходится почти наверное* к ξ, если

$$\xi_n \xrightarrow{\text{п.н.}} \xi \qquad \Longleftrightarrow \qquad \mathbb{P}\big\{w \in \Omega \;\big|\; \xi_n(w) \to \xi(w)\big\} = 1$$

- $\{\xi_n\}_{n=1}^{\infty}$ *сходится в среднем p-го порядка* к ξ, если

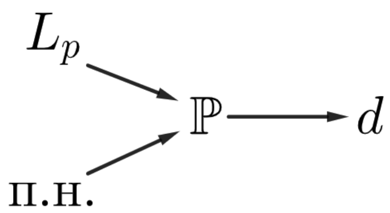
$$\xi_n \xrightarrow{L_p} \xi \qquad \Longleftrightarrow \qquad \mathbb{E}\big[|\xi_n - \xi|^p\big] \longrightarrow 0$$

  При $p = 1$ — *сходимость в среднем*, при $p = 2$ — *сходимость в среднем квадратичном.*

- $\{\xi_n\}_{n=1}^{\infty}$ *сходится по распределению* к ξ, если

$$\xi_n \xrightarrow{d} \xi \qquad \Longleftrightarrow \qquad F_{\xi_n}(t) \longrightarrow F_{\xi}(t) \;\; \text{в точках } t\text{, в которых } F_{\xi} \text{ непрерывна}$$

Связь между различными типами сходимости:



Неравенство Маркова. Пусть ξ — неотрицательная с.в., $\mathbb{E}\xi < +\infty$. Тогда

$$\mathbb{P}\{\xi \geq t\} \leq \frac{\mathbb{E}\xi}{t} \qquad \forall t > 0$$

Неравество Чебышева. Пусть ξ — с.в., $\mathbb{V}\xi < +\infty$. Тогда

$$\mathbb{P}\big\{|\xi - \mathbb{E}\xi| \geq t\big\} \leq \frac{\mathbb{V}\xi}{t^2} \qquad \forall t > 0$$

Опр. Говорят, что для последовательности случайных величин $\{\xi_n\}_{n=1}^{\infty}$ выполняется

- *(слабый) закон больших чисел (ЗБЧ)*, если

$$\frac{1}{n}\sum_{k=1}^{n} \xi_k - \frac{1}{n}\sum_{k=1}^{n} \mathbb{E}\xi_k \xrightarrow{\mathbb{P}} 0$$

- *усиленный закон больших чисел (УЗБЧ)*, если

$$\frac{1}{n}\sum_{k=1}^{n} \xi_k - \frac{1}{n}\sum_{k=1}^{n} \mathbb{E}\xi_k \xrightarrow{\text{п.н.}} 0$$

Достаточные условия выполнения ЗБЧ (для последовательности $\{\xi_n\}_{n=1}^{\infty}$)

1. **ЗБЧ в форме Чебышева:**

$$\left.\begin{aligned} &\exists C > 0 : \quad \mathbb{V}\xi_n \leq C \quad \forall n \in \mathbb{N} \\ &\text{все } \xi_n \text{ попарно не коррелируют} \end{aligned}\right\} \qquad \Longrightarrow \qquad \text{выполняется ЗБЧ}$$

2. **ЗБЧ в форме Маркова:**

$$\mathbb{V}\left[\frac{1}{n}\sum_{k=1}^{n}\xi_k\right] \longrightarrow 0 \qquad \Longrightarrow \qquad \text{выполняется ЗБЧ}$$

3. **УЗБЧ в форме Колмогорова:**

$$\left.\begin{array}{c}\{\xi_n\}_{n=1}^{\infty} - \text{i.i.d.} \\ \mathbb{E}|\xi_1| < +\infty\end{array}\right\} \qquad \Longrightarrow \qquad \text{выполняется УЗБЧ}$$

Теорема Леви о непрерывности

Последовательность случайных величин $\{\xi_n\}_{n=1}^{\infty}$ сходится по распределению к случайной величине ξ тогда и только тогда, когда их характеристические функции сходятся поточечно:

$$\xi_n \xrightarrow{d} \xi \qquad \Longleftrightarrow \qquad \varphi_{\xi_n}(t) \longrightarrow \varphi_{\xi}(t) \qquad \forall t \in \mathbb{R}$$

Теорема Пуассона

Пусть $X_n \sim \mathrm{Binom}(n, p_n)$, причем $p_n \to 0,\ np_n \to \lambda > 0$ при $n \to \infty$.

Тогда $X_n \xrightarrow{d} X$, где $X \sim \mathrm{Poiss}(\lambda)$.

Центральная предельная теорема

Пусть $\{\xi_n\}_{n=1}^{\infty}$ — i.i.d., $\mathbb{E}\xi_1 = m,\ \mathbb{V}\xi_1 = \sigma^2$. Тогда если $\eta_n = \sum\limits_{k=1}^{n}\xi_k$, то

$$\frac{\eta_n - \mathbb{E}\eta_n}{\sqrt{\mathbb{V}\eta_n}} = \frac{\sum\limits_{k=1}^{n}\xi_k - mn}{\sigma\sqrt{n}} \xrightarrow{d} \mathcal{N}(0,1)$$

Неравенство Берри-Эссеена

Пусть $\{\xi_n\}_{n=1}^{\infty}$ — i.i.d., $\mathbb{E}\xi_1 = 0,\ \mathbb{V}\xi_1 = \sigma^2,\ \mathbb{E}|\xi_1|^3 \le \rho$. Тогда если $\zeta_n = \dfrac{1}{\sigma\sqrt{n}}\sum\limits_{k=1}^{n}\xi_k$, то

$$\sup_{x\in\mathbb{R}}\left|F_{\zeta_n}(x) - \Phi(x)\right| \le \frac{C\rho}{\sigma^{3/2}\sqrt{n}},$$

где $C < 0.48$ — константа (точное значение неизвестно), а $\Phi(x)$ — функция распределения $\mathcal{N}(0,1)$.

1.4 Полезные факты из теории вероятностей

1. Пусть $\varphi_\xi(t), \varphi_\eta(t)$ — характеристические функции случайных величин ξ и η соответственно. Тогда ξ и η независимы тогда и только тогда, когда

$$\varphi_{(\xi,\eta)}(t_1, t_2) = \varphi_\xi(t_1)\cdot\varphi_\eta(t_2),$$

то есть характеристическая функция совместного распределения есть произведение характеристических функций.

Данное свойство удобно использовать для того, чтобы показать, что какой-то набор случайных величин является независимым в совокупности.

2. Если $c \in \mathbb{R}$ — константа, то $\xi_n \xrightarrow{\mathbb{P}} c \iff \xi_n \xrightarrow{d} c$.

Данное свойство удобно использовать, когда нужно проверить сходимость по вероятности, то есть когда есть предположение о том, что может быть пределом:

$$\xi_n \xrightarrow{\mathbb{P}} \xi \qquad \Longleftrightarrow \qquad \xi_n - \xi \xrightarrow{\mathbb{P}} 0 \qquad \Longleftrightarrow \qquad \xi_n - \xi \xrightarrow{d} 0 \quad \overset{\text{Теорема Леви}}{\Longleftrightarrow} \quad \varphi_{\xi_n - \xi}(t) \longrightarrow 1,$$

где $\varphi_0(t) \equiv 1$ — характеристическая функция тождественного нуля.

2 Основные понятия

2.1 Определение случайного процесса

Опр. *Случайным процессом* называется семейство случайных величин

$$\{X(t) = X(w,t) \mid t \in T\},$$

заданных на одном вероятностном пространстве $(\Omega, \mathcal{F}, \mathbb{P})$ и зависящих от одного параметра $t \in T \subset \mathbb{R}$.

Опр. При фиксированном $t = t_0$ случайная величина $X(t_0)$ называется *сечением процесса* $X(t)$.

Опр. При фиксированном исходе $w = w_0$ неслучайная функция времени $X(w_0, t)$ называется *траекторией* или *реализацией* процесса $X(t)$.

Опр. Пусть $n \in \mathbb{N}$, а $X(t_1), \ldots, X(t_n)$ — сечения случайного процесса $X(t)$. Функция распределения случайного вектора $\big(X(t_1), \ldots, X(t_n)\big)$

$$F\big(x_1, \ldots, x_n; t_1, \ldots, t_n\big) = \mathbb{P}\big\{X(t_1) < x_1, \ldots, X(t_n) < x_n\big\}$$

называется *n-мерной функцией распределения* случайного процесса $X(t)$.

Опр. Совокупность n-мерных функций распределения для всех $n \in \mathbb{N}$ и всех наборов $t_1, \ldots, t_n \in T$ называется *семейством конечномерных распределений* случайного процесса $X(t)$:

$$F = \Big\{F\big(x_1, \ldots, x_n; t_1, \ldots, t_n\big) \;\Big|\; n \in \mathbb{N},\ t_1, \ldots, t_n \in T\Big\}$$

Найти семейство конечномерных распределений процесса $X(t)$ — любым образом задать распределения случайного вектора $\big(X(t_1), \ldots, X(t_n)\big)$. Можно найти его функцию распределения, можно задать вероятности для каждого набора значений (если это дискретный случайный вектор), можно задать плотность распределения (если это непрерывный случайный вектор).

Свойства n-мерной функции распределения случайного процесса $X(t)$:

(a) $0 \leq F\big(x_1, \ldots, x_n; t_1, \ldots, t_n\big) \leq 1$;

(b) Функция $F\big(x_1, \ldots, x_n; t_1, \ldots, t_n\big)$ непрерывна слева по каждой переменной x_i;

(c) Если хотя бы одна переменная $x_i \to -\infty$, то

$$F\big(x_1, \ldots, x_n; t_1, \ldots, t_n\big) \longrightarrow 0$$

Если все переменные $x_i \to +\infty$, то

$$F\big(x_1, \ldots, x_n; t_1, \ldots, t_n\big) \longrightarrow 1;$$

(d) Функция $F\big(x_1, \ldots, x_n; t_1, \ldots, t_n\big)$ монотонна в следующем смысле:

$$\Delta_1 \ldots \Delta_n F\big(x_1, \ldots, x_n; t_1, \ldots, t_n\big) \geq 0,$$

где Δ_i — оператор конечной разности по переменной x_i:

$$\Delta_i F = F\big(x_1, \ldots, x_i + h_i, \ldots, x_n; t_1, \ldots, t_n\big) - F\big(x_1, \ldots, x_i, \ldots, x_n; t_1, \ldots, t_n\big),$$

а $h_i \geq 0$ — произвольны.

Данное свойство по сути означает, что вероятность случайного вектора попасть в любой прямоугольник вида $[x_1, x_1 + h_1) \times \ldots \times [x_n, x_n + h_n)$ неотрицательна (это следует из формулы включений-исключений).

(e) Свойство симметричности: для любой перестановки $\{k_1, \ldots, k_n\}$ индексов $\{1, \ldots, n\}$

$$F\big(x_1, \ldots, x_n; t_1, \ldots, t_n\big) = F\big(x_{k_1}, \ldots, x_{k_n}; t_{k_1}, \ldots, t_{k_n}\big)$$

(f) Свойство согласования: для любых $1 \le k < n$ и $x_1, \ldots, x_k \in \mathbb{R}$

$$F\big(x_1, \ldots, x_k; t_1, \ldots, t_k\big) = F\big(x_1, \ldots, x_k, +\infty, \ldots, +\infty; t_1, \ldots, t_n\big).$$

Это свойство указывает на связь k-мерной и n-мерной функций распределения.

Теорема Колмогорова (о семействе конечномерных распределений)

> Пусть задано некоторое семейство функций
>
> $$F = \Big\{F\big(x_1, \ldots, x_n; t_1, \ldots, t_n\big) \;\Big|\; n \in \mathbb{N},\ t_1, \ldots, t_n \in T\Big\},$$
>
> удовлетворяющих свойствам $(a) - (f)$. Тогда существуют вероятностное пространство $(\Omega, \mathcal{F}, \mathbb{P})$ и случайный процесс $\{\xi(t), t \in T\}$ на нем, такие, что семейство конечномерных распределений F_ξ случайного процесса $\xi(t)$ совпадает с семейством F.

Опр. 1 Два случайных процесса $\{\xi(t), t \in T\}$ и $\{\eta(t), t \in T\}$, определенных на одном вероятностном пространстве, называются *(стохастически) эквивалентными*, если

$$\mathbb{P}\big\{\xi(t) = \eta(t)\big\} = 1, \qquad \forall t \in T$$

При этом процесс $\eta(t)$ называется *модификацией* процесса $\xi(t)$, и наоборот.

Утв. 1 Семейства конечномерных распределений эквивалентных процессов совпадают.
Обратное неверно.

2.2 Непрерывность траекторий

Опр. Процесс $\xi(t)$ называется *процессом с непрерывными траекториями*, если почти все его траектории непрерывны, т.е. если

$$\mathbb{P}\big\{\xi(t) \text{ — непрерывная функция}\big\} = 1$$

Опр. Процесс $\eta(t)$ называется *непрерывной модификацией* процесса $\xi(t)$, если он является его модификацией и имеет непрерывные траектории.

Допустим, мы хотим вычислить вероятность

$$\mathbb{P}\left\{\sup_{t \in [0,1]} \xi(t) \le x\right\} = \mathbb{P}\left\{\bigcap_{t \in [0,1]} \{\xi(t) \le x\}\right\}$$

События $\{\xi(t) \le x\}$ лежат в сигма-алгебре $\mathcal{F}$ на вероятностном пространстве Ω, а их несчетное пересечения, вообще говоря, там не лежит. Поэтому такая вероятность не определена.

Пусть теперь $\xi(t)$ имеет непрерывные траектории. Тогда искомое событие можно записать в виде

$$\bigcap_{t \in [0,1]} \{\xi(t) \le x\} = \bigcap_{t \in [0,1] \cap \mathbb{Q}} \{\xi(t) \le x\},$$

то есть можно брать пересечение только по рациональным точкам. Теперь это счетное пересечение, и оно лежит в сигма-алгебре, поэтому вероятность определена.

Такая же вероятность определена для процессов, у которых есть непрерывная модификация. В связи с этим часто возникает вопрос проверки, есть ли у процесса непрерывная модификация.

Теорема Колмогорова (о существовании непрерывной модификации)

> Пусть для случайного процесса $\big\{\xi(t),\ t \in [a,b]\big\}$ существуют такие $\alpha > 0,\ \beta > 0,\ c < \infty$, что
>
> $$\mathbb{E}\big|\xi(t+h) - \xi(t)\big|^\alpha \le c|h|^{1+\beta}, \qquad \forall t, t+h \in [a,b]$$
>
> Тогда процесс $\xi(t)$ имеет непрерывную модификацию.

2.3 Моментные характеристики

Опр. *Математическим ожиданием* случайного процесса $X(t)$ называется функция $m_X : T \to \mathbb{R}$:

$$m_X(t) = \mathbb{E}X(t)$$

Опр. *Корреляционной функцией* случайного процесса $X(t)$ называется функция $R_X : T \times T \to \mathbb{R}$:

$$R_X(t,s) = \mathrm{cov}\big(X(t), X(s)\big) = \mathbb{E}\big[\overset{\circ}{X}(t)\overset{\circ}{X}(s)\big] = \mathbb{E}\Big[\big(X(t) - \mathbb{E}X(t)\big)\big(X(s) - \mathbb{E}X(s)\big)\Big]$$

В учебнике Гасникова А.В. "Лекции по случайным процессам" эта функция называется *корреляционной функцией*, а в учебнике Миллера, Панкова "Теория случайных процессов в примерах и задачах" — *ковариационной функцией*.

Опр. *Ковариационной функцией* случайного процесса $X(t)$ называется функция $K_X : T \times T \to \mathbb{R}$:

$$K_X(t,s) = \mathbb{E}[X(t)X(s)] = R_X(t,s) + m_X(t)m_X(s)$$

Опр. *Дисперсией* случайного процесса $X(t)$ называется функция $D_X : T \to \mathbb{R}$:

$$D_X(t) = \mathbb{V}X(t) = \mathbb{E}\big[\overset{\circ}{X}{}^2(t)\big]$$

Видно, что $D_X(t) = R_X(t,t)$.

Опр. *Характеристической функцией* случайного процесса $X(t)$ называется функция

$$\varphi_{X(t)}(s) = \mathbb{E}\big[e^{isX(t)}\big], \qquad s \in \mathbb{R}$$

Опр. Два случайных процесса $X(t)$ и $Y(t)$ называются *независимыми*, если для любых $n \in \mathbb{N}$ и любых $t_1, \ldots, t_n \in T$ независимы случайные векторы

$$\xi = \big(X(t_1), \ldots, X(t_n)\big), \qquad \eta = \big(Y(t_1), \ldots, Y(t_n)\big).$$

3 Процессы Леви

3.1 Винеровский процесс

Опр. Случайный процесс $\{X(t),\ t \geq 0\}$ называется *гауссовским* или *нормальным*, если любое его конечномерное распределение является нормальным.

Утв. Гауссовский процесс $X(t)$ однозначно определяется своим матожиданием $m_X(t)$ и корреляционной функцией $R_X(t,s)$.

Опр. Процесс $\{X(t),\ t \geq 0\}$ называется *процессом с независимыми приращениями*, если для любых $n \in \mathbb{N}$ и для любых $0 \leq t_1 \leq t_2 \leq \ldots \leq t_n$ случайные величины

$$X(0), \quad X(t_1) - X(0), \quad X(t_2) - X(t_1), \quad \ldots, \quad X(t_n) - X(t_{n-1})$$

независимы в совокупности.

Два эквивалентных определения винеровского процесса:

Опр. 1 *Винеровским процессом* с параметром $\sigma > 0$ называется случайный процесс $\{W(t),\ t \geq 0\}$, удовлетворяющий свойствам

(a) $W(0) = 0$ почти наверное;

(b) $W(t)$ — процесс с независимыми приращениями;

(c) Для любых $t, s \geq 0$ выполнено $W(t) - W(s) \sim \mathcal{N}\big(0, \sigma^2|t-s|\big)$

По умолчанию обычно подразумевается $\sigma = 1$. Такой процесс называется *стандартным винеровским процессом.*

Опр. 2 *Винеровским процессом* называется процесс $\{W(t),\ t \geq 0\}$, удовлетворяющий свойствам

(a) W(t) — гауссовский процесс;

(b) $m_W(t) = 0$ для любого $t \geq 0$;

(c) $R_W(t,s) = \min(t,s)$ для любых $t, s \geq 0$.

Из определения винеровского процесса несложно найти семейство его конечномерных плотностей. Заметим, что

$$\begin{pmatrix} W(t_1) \\ W(t_2) \\ \vdots \\ W(t_n) \end{pmatrix} = \begin{bmatrix} 1 & 0 & \ldots & 0 \\ 1 & 1 & \ddots & \vdots \\ \vdots & \ddots & \ddots & 0 \\ 1 & \ldots & 1 & 1 \end{bmatrix} \begin{pmatrix} W(t_1) \\ W(t_2) - W(t_1) \\ \vdots \\ W(t_n) - W(t_{n-1}) \end{pmatrix}$$

При линейном преобразовании $Y = AX$ плотность распределения меняется по правилу

$$f_Y(x) = \frac{1}{\det A} \cdot f_X\big(A^{-1}x\big),$$

поэтому, учитывая, что случайный вектор в правой части состоит из независимых нормальных компонент, находим n-мерную плотность распределения винеровского процесса:

$$f_Y(x) = \prod_{k=1}^{n} \frac{1}{\sqrt{2\pi(t_k - t_{k-1})}} \exp\left(-\frac{(x_k - x_{k-1})^2}{2(t_k - t_{k-1})}\right), \qquad x_0 = 0, \qquad t_0 = 0$$

Свойства винеровского процесса $W(t)$:

1. Матожидание и корреляционная функция:

$$m_W(t) = \mathbb{E}W(t) = 0, \qquad R_W(t,s) = \min(t,s)$$

2. Следующие процессы тоже являются винеровскими:

$$X_1(t) = -W(t), \qquad X_2(t) = \frac{1}{\sqrt{c}} \cdot W(ct), \qquad X_3(t) = t \cdot W\left(\frac{1}{t}\right)$$

3. Траектории винеровского процесса лежат между графиками функций $y = \pm 3\sqrt{t}$ с вероятностью 99.7% (согласно правилу трех сигм).

4. Для винеровского процесса существует непрерывная модификация (в теореме Колмогорова достаточно взять $\alpha = 4,\ \beta = 1,\ c = 3$).

3.2 Пуассоновский процесс

Следующие определения пуассоновского процесса эквивалентны.

Опр. 1 Пусть $\xi_1, \xi_2, \ldots \sim \mathrm{Exp}(\lambda)$ независимы в совокупности. Случайный процесс

$$K(t) = \max\left\{n \;\middle|\; S_n < t\right\}, \qquad S_0 = 0, \qquad S_n = \xi_1 + \ldots + \xi_n \sim \mathrm{Gamma}(n, \lambda)$$

называется *пуассоновским процессом интенсивности* λ и обозначается $K(t) \sim \text{ПП}(\lambda)$.

Опр. 2 *Пуассоновским процессом интенсивности* λ называется процесс $\{K(t),\ t \geq 0\}$, удовлетворяющий свойствам:

(a) $K(0) = 0$ почти наверное;

(b) $K(t)$ — процесс с независимыми приращениями;

(c) $\forall t > s \geq 0$ выполнено $K(t) - K(s) \sim \mathrm{Poiss}(\lambda(t-s))$.

Из второго определения можно получить первое, определив случайные величины

$$S_0 = 0, \qquad S_k = \inf\left\{t \;\middle|\; K(t) \geq k\right\}$$

Тогда случайные величины $\xi_k = S_k - S_{k-1} \sim \mathrm{Exp}(\lambda)$.

Опр. *Простейшим потоком событий* называется последовательность событий, происходящих во времени по следующим правилам:

(a) *Стационарность:* распределение числа событий, происшедших на интервале $[t, s]$ есть функция только длины интервала $s - t$.

(b) *Отсутствие последействия:* числа событий, происшедших на непересекающихся временных интервалах, независимы в совокупности.

(c) *Ординарность:* вероятность того, что на интервале $[t, t + \Delta t]$ произойдет хотя бы одно событие, равна $\lambda(t)\Delta t + o(\Delta t),\ \Delta t \to 0$, где $\lambda(t) \equiv \lambda$ для стационарного процесса (не зависит от t).

Через $\nu(t, s)$ обозначим число событий в простейшем потоке на интервале $[t, s]$.

Опр. 3 Случайный процесс $K(t) = \nu(0, t)$, равный числу событий на интервале $[0, t]$ в простейшем потоке событий, называется *пуассоновским процессом интенсивности* λ.

Свойства пуассоновского процесса $K(t)$:

1. *Отсутствие памяти:* $\forall t, s > 0$ и $\forall u \in [0, s]$ случайные величины $K(t+s) - K(s)$ и $K(u)$ независимы, и имеет место равенство распределений случайных величин

$$K(t+s) - K(s) \overset{d}{=} K(t) \sim \mathrm{Poiss}(\lambda t)$$

То есть если для фиксированного $s > 0$ определить процесс $X(t) = K(t+s) - K(s)$, то он тоже будет пуассоновским с интенсивностью λ и не будет зависеть от всей предыстории.

2. Матожидание и корреляционная функция:

$$m_K(t) = \lambda t, \qquad R_K(t,s) = \lambda \min(t,s)$$

3. Конечномерные функции вероятности: $(t_1 < \ldots < t_n,\ k_1 \le \ldots \le k_n)$

$$\mathbb{P}\big\{K(t_1) = k_1, \ldots, K(t_n) = k_n\big\} = \prod_{j=1}^{n} \frac{\big[\lambda(t_j - t_{j-1})\big]^{k_j - k_{j-1}}}{(k_j - k_{j-1})!} e^{-\lambda(t_j - t_{j-1})}$$

4. Пусть $S_1, \ldots, S_n$ — с.в. из определения 1 процесса $K(t)$. Пусть $K(t) = n$. Тогда условное распределение

$$\big(S_1, \ldots, S_n\big)\ \Big|_{K(t)=n} \overset{d}{=} \big(u_{(1)}, \ldots, u_{(n)}\big),$$

где $u_1, \ldots, u_n \sim \mathcal{U}[0,t]$ — i.i.d., а случайные величины $u_{(k)}$ — их порядковые статистики, т.е.

$$u_{(1)} = \min_{1 \le i \le n} u_{(i)}, \quad \ldots, \quad u_{(n)} = \max_{1 \le i \le n} u_{(i)}$$

5. Два способа сгенерировать пуассоновский процесс $K(t) \sim \Pi\Pi(\lambda)$:

   - Процесс до значения n: сгенерировать последовательность $\xi_1, \ldots, \xi_n \sim \mathrm{Exp}(\lambda)$ и построить процесс $K(t)$ согласно первому определению:

   $$K(t) = \max\big\{n \bigm| S_n < t\big\}, \qquad S_0 = 0, \qquad S_n = \xi_1 + \ldots + \xi_n$$

   - Процесс до времени T: сгенерировать случайную величину $K(T) \sim \mathrm{Poiss}(\lambda T)$, потом сгенерировать $n = K(T)$ случайных величин $S_1, \ldots, S_n \sim \mathcal{U}[0,T]$ и упорядочить их. Процесс $K(t)$ построить аналогично.

3.3 Сложный (обобщенный) пуассоновский процесс

Опр. Случайный процесс

$$Q(t) = \sum_{i=1}^{K(t)} V_i, \qquad t \ge 0$$

где $K(t)$ — пуассоновский процесс, а $\{V_i\}_{i=1}^{\infty}$ — i.i.d. и не зависят от $K(t)$, называется *сложным (обобщенным) пуассоновским процессом.*

Свойства сложного пуассоновского процесса $Q(t)$:

1. $Q(t)$ — процесс с независимыми приращениями.
2. Матожидание и корреляционная функция: (если $\mathbb{E}V_1^2 < +\infty$)

$$m_Q(t) = \lambda t \mathbb{E}V_1, \qquad R_Q(t,s) = \lambda \min(t,s)\, \mathbb{E}V_1^2$$

3. Характеристическая функция:

$$\varphi_{Q(t)}(s) = \exp\big\{\lambda t\big(\varphi_{V_1}(s) - 1\big)\big\}$$

4. Если $V_1 \sim \mathrm{Be}(p)$, то $Q(t)$ — пуассоновский процесс с интенсивностью λp.
   Данное свойство пуассоновского процесса называется *устойчивостью относительно случайного прореживания.*

3.4 Безгранично делимые случайные величины

Опр. Случайная величина X называется *безгранично делимой*, если для любого $n \in \mathbb{N}$ существует набор случайный величин $\{X_{kn}\}_{k=1}^{n}$ — i.i.d., такой, что

$$X \overset{d}{=} \sum_{k=1}^{n} X_{kn}, \qquad \forall n \in \mathbb{N}$$

Теорема Леви-Хинчина

> Случайная величина Y является безгранично делимой $\Longleftrightarrow$ логарифм ее характеристической функции $\varphi_Y(s)$ имеет вид
>
> $$\ln \varphi_Y(s) = ibs - \frac{\sigma^2 s^2}{2} + \int\limits_{-\infty}^{+\infty} \Big(e^{isx} - 1 - isx\mathbb{I}\{|x| < 1\}\Big) d\nu(x),$$
>
> где $\mathbb{I}$ — индикаторная функция, $b \in \mathbb{R}$, $\sigma^2 \geq 0$ — некоторые числа, $\nu(x)$ — некоторая мера на $\mathbb{R}$ со свойствами
>
> $$\nu\big(\{0\}\big) = 0, \qquad \int_{\mathbb{R}} \min(1, x^2) d\nu(x) < +\infty,$$
>
> а интеграл является интегралом Лебега по мере ν.
> При этом тройка (b, σ^2, ν) для каждой безгранично делимой случайной величины определяется единственным образом.

Теорема 2

> Пусть для любого $n \in \mathbb{N}$ набор с.в. $\{X_{kn}\}_{k=1}^n$ — i.i.d., и выполнено
>
> $$\sum_{k=1}^{n} X_{kn} \xrightarrow{d} X, \qquad \forall n \in \mathbb{N}$$
>
> Тогда X — безгранично делимая случайная величина.

Примеры безгранично делимых случайных величин:

1. $X \equiv m$ — тождественная случайная величина $\big(b = m,\ \sigma^2 = 0,\ \nu(x) = 0\big)$.
2. $X \sim \mathcal{N}(\mu, s^2)$ $\big(b = \mu,\ \sigma^2 = s^2,\ \nu(x) = 0\big)$.
3. $X \sim \mathrm{Poiss}(\lambda)$ $\big(b = 0,\ \sigma^2 = 0,\ \nu(x) = \lambda\delta(x-1)\big)$, где $\delta(x)$ — дельта-функция.

   В этом случае интеграл вырождается в значение подынтегральной функции в $x = 1$ со множителем λ:

   $$\ln \varphi_X(s) = \lambda\big(e^{is} - 1\big)$$

3.5 Процессы Леви

Опр. *Процессом Леви* называется случайный процесс $\big\{X(t),\ t \geq 0\big\}$, удовлетворяющий свойствам

(a) $X(0) = 0$ почти наверное;

(b) $X(t)$ — процесс с независимыми приращениями;

(c) $\forall t, s \geq 0$ случайная величина $X(t+s) - X(t)$ имеет распределение, не зависящее от t (стационарность приращений), т.е.

$$X(t+s) - X(t) \stackrel{d}{=} X(s)$$

(d) $X(t)$ — стохастически непрерывный процесс, т.е.

$$\forall t \geq 0: \quad X(t+\varepsilon) \xrightarrow{\mathbb{P}} X(t) \ \text{ при } \ \varepsilon \to 0$$

Процессами Леви являются винеровский процесс, пуссоновский и сложный пуассоновский процессы.

Свойства процесса Леви $X(t)$:

1. Любое сечение $X(t)$ — безгранично делимая случайная величина.
2. Если F — безгранично делимое распределение, то существует процесс Леви $X(t)$, такой, что с.в. $X(1)$ имеет распределение F.

3. Если $X(t)$ и $Y(t)$ — процессы Леви, и $X(1) \stackrel{d}{=} Y(1)$, то конечномерные распределение процессов $X(t)$ и $Y(t)$ совпадают (т.е. они эквивалентны).

4. Характеристическая функция, матожидание и дисперсия обладают свойствами:

$$\varphi_{X(t)}(s) = \left[\varphi_{X(1)}(s)\right]^t, \qquad \mathbb{E}X(t) = t \cdot \mathbb{E}X(1), \qquad \mathbb{V}X(t) = t \cdot \mathbb{V}X(1)$$

5. Характеристическая функция $\varphi_{X(t)}(s)$ имеет вид, согласованный с *теоремой Леви-Хинчина* (§3.4):

$$\ln \varphi_{X(t+\tau)-X(t)}(s) = \tau \cdot \left[ibs - \frac{\sigma^2 s^2}{2} + \int\limits_{-\infty}^{+\infty} \Big(e^{isx} - 1 - isx\mathbb{I}\big\{|x| < 1\big\} \Big) d\nu(x) \right]$$

4 Введение в стохастический анализ

4.1 Пространство L_2

Опр. Пространством L_2 *случайных величин второго порядка* называется множество случайных величин с конечным вторым моментом со скалярным произведением:

$$L_2 = \{X \mid \mathbb{E}X^2 < +\infty\}, \qquad \langle X, Y\rangle = \mathbb{E}[XY]$$

Порожденная норма:

$$\|X\|_2 = \sqrt{\mathbb{E}X^2}$$

Сходимость по норме обозначается следующим образом:

$$X_n \xrightarrow{L_2} X \qquad \Longleftrightarrow \qquad \|X_n - X\|_2 \longrightarrow 0 \qquad \Longrightarrow \qquad X = \underset{n\to\infty}{\text{l.i.m.}}\, X_n$$

Свойства пространства L_2.

1. *неравенство Коши-Буняковского*: $\langle X, Y\rangle^2 \le \langle X, X\rangle\langle Y, Y\rangle$;
2. *непрерывность скалярного произведения*:

$$X_n \xrightarrow{L_2} X, \quad Y_m \xrightarrow{L_2} Y \qquad \Longrightarrow \qquad \langle X_n, Y_m\rangle \xrightarrow{n,m\to\infty} \langle X, Y\rangle$$

   Отсюда следует непрерывность матожидания, дисперсии и ковариации.

3. пространство L_2 полно;
4. *критерий сходимости в* L_2:

$$X_n \xrightarrow{L_2} X \qquad \Longleftrightarrow \qquad \exists c \in \mathbb{R}: \text{ для любых подпоследовательностей } \{X_{n_k}\}_{k=1}^{\infty},\ \{X_{n_m}\}_{m=1}^{\infty}$$

$$\langle X_{n_k}, X_{n_m}\rangle \xrightarrow{k,m\to\infty} c$$

Опр. Будем говорить, что случайный процесс $X(t) \in L_2$, если каждое сечение $X(t)$ лежит в L_2.

4.2 Непрерывность, дифференцируемость и интегрируемость в среднем квадратичном

Опр. Случайный процесс $X(t) \in L_2$ называется *непрерывным в среднем квадратичном* или *с.к.-непрерывным*, если

$$\forall t \ge 0: \qquad X(t+\varepsilon) \xrightarrow[\varepsilon\to 0]{L_2} X(t)$$

Критерий с.к.-непрерывности

> Следующие утверждения эквивалентны:
>
> (a) Процесс $X(t) \in L_2$ является с.к.-непрерывным.
>
> (b) Ковариационная функция $K_X(t_1, t_2) = \mathbb{E}[X(t_1)X(t_2)]$ непрерывна на множестве $[0, +\infty)^2$.
>
> (c) Корреляционная функция $R_X(t_1, t_2)$ непрерывна на $[0, +\infty)^2$, а матожидание $m_X(t)$ непрерывно на $[0, +\infty)$.

Для следствия $(b) \Rightarrow (a)$ достаточно непрерывности $K_X(t_1, t_2)$ только на множестве $\{(t_1, t_2) \mid t_1 = t_2\}$.

Опр. Случайный процесс $X(t) \in L_2$ называется *дифференцируемым в среднем квадратичном* или *с.к.-дифференцируемым*, если существует процесс $Y(t) \in L_2$, такой, что

$$\forall t \ge 0: \qquad \frac{X(t+\varepsilon) - X(t)}{\varepsilon} \xrightarrow[\varepsilon\to 0]{L_2} Y(t)$$

При этом процесс $Y(t)$ называется *с.к.-производной* процесса $X(t)$ и обозначается $Y(t) = X'(t)$.

При фиксированном исходе $\omega \in \Omega$ траектория $X'(t)$ является обычной производной реализации $X(t)$ (если эта производная существует).

Критерий с.к.-дифференцируемости

> Следующие утверждения эквивалентны:
>
> (a) Процесс $X(t) \in L_2$ является с.к.-дифференцируемым.
>
> (b) Для любых $t_1, t_2 \geq 0$ существует конечный предел
>
> $$\lim_{\varepsilon_1,\varepsilon_2\to 0} \frac{1}{\varepsilon_1\varepsilon_2}\big[K_X(t_1+\varepsilon_1, t_2+\varepsilon_2) - K_X(t_1+\varepsilon_1, t_2) - K_X(t_1, t_2+\varepsilon_2) + K_X(t_1,t_2)\big] < \infty$$
>
> (c) Для любого $t \geq 0$ функция $m_X(t)$ дифференцируема и существует конечный предел
>
> $$\lim_{\varepsilon_1,\varepsilon_2\to 0} \frac{1}{\varepsilon_1\varepsilon_2}\big[R_X(t+\varepsilon_1, t+\varepsilon_2) - R_X(t+\varepsilon_1, t) - R_X(t, t+\varepsilon_2) + R_X(t,t)\big] < \infty$$

Для следствия $(b) \Rightarrow (a)$ достаточно существование данного предела только на множестве $\{(t_1,t_2) \mid t_1 = t_2\}$, что уже учтено в случае (c).

Предел в данной теореме называется *обобщенной смешанной производной*. Из его существования следует существование обычных смешанных производных второго порядка, а обратное неверно.

Достаточным условием с.к.-дифференцируемости является дважды непрерывная дифференцируемость ковариационной функции $K_X(t,s)$.

Опр. Пусть процесс $X(t) \in L_2$ определен от отрезке $[a,b]$. Пусть дано разбиение:

$$a = t_0 < t_1 < \ldots < t_n = b, \qquad \Delta t_i = t_i - t_{i-1}, \qquad \tau_i \in [t_{i-1}, t_i),$$

где τ_i — произвольные точки. Если существует предел

$$\sum_{i=1}^{n} X(\tau_i)\Delta t_i \xrightarrow[\substack{\max \Delta t_i \to 0 \\ n\to\infty}]{L_2} Y,$$

не зависящий от разбиения $\{t_i\}$ и выбора точек $\{\tau_i\}$, то процесс $X(t)$ называется *с.к.-интегрируемым* на $[a,b]$, а случайная величина Y называется ее *с.к.-интегралом* или *стохастическим интегралом Римана* на отрезке $[a,b]$ и обозначается

$$Y = \int_a^b X(t)dt$$

Критерий с.к.-интегрируемости

> Следующие утверждения эквивалентны:
>
> (a) Процесс $X(t) \in L_2$ является с.к.-интегрируемым.
>
> (b) Существует интеграл Римана
>
> $$\int_a^b\int_a^b K_X(t_1,t_2)dt_1dt_2 < \infty$$
>
> (c) Существуют интегралы Римана
>
> $$\int_a^b m_X(t)dt < \infty, \qquad \int_a^b\int_a^b R_X(t_1,t_2)dt_1dt_2 < \infty$$

Связь между данными понятиями:

$$\text{с.к.-дифференцируемость} \implies \text{с.к.-непрерывность} \implies \text{с.к.-интегрируемость}$$

Полезные формулы: (все процессы считаются достаточно с.к.-гладкими)

$$\mathbb{E}X'(t) = \frac{d}{dt}\mathbb{E}X(t), \qquad \mathbb{E}\int_a^b X(t)dt = \int_a^b \mathbb{E}X(t)dt, \qquad R_{X'}(t,s) = \frac{\partial^2 R_X(t,s)}{\partial t\, \partial s}$$

4.3 Непрерывность и дифференцируемость по вероятности

Опр. Случайный процесс $X(t)$ называется *непрерывным по вероятности* или *стохастически непрерывным*, если

$$\forall t \geq 0: \qquad X(t+\varepsilon) \xrightarrow[\varepsilon\to 0]{\mathbb{P}} X(t)$$

$$\text{с.к.-непрерывность} \implies \text{стохастическая непрерывность}$$

Опр. Случайный процесс $X(t)$ называется *дифференцируемым по вероятности* или *стохастически дифференцируемым*, если существует процесс $Y(t)$, такой, что

$$\forall t \geq 0: \qquad \frac{X(t+\varepsilon) - X(t)}{\varepsilon} \xrightarrow[\varepsilon\to 0]{\mathbb{P}} Y(t)$$

$$\text{с.к.-дифференцируемость} \implies \text{стохастическая дифференцируемость}$$

Аналогично можно ввести понятия непрерывности и дифференцируемости по распределению.

Пусть

- $W(t)$ — винеровский процесс;
- $K(t)$ — пуассоновский процесс;
- $Q(t)$ — сложный пуассоновский процесс.

	непрерывность		дифференцируемость	
	в среднем квадратичном	по вероятности	в среднем квадратичном	по вероятности
$W(t)$	✓	✓	✗	✗
$K(t)$	✓	✓	✗	✓
$Q(t)$	✓, если $Q(t) \in L_2$	✓	✗	✓

4.4 Интегралы от случайных процессов и формула Ито

Опр. Пусть процесс $X(t) \in L_2$ определен на отрезке $[a,b]$, а $g(t)$ — неслучайная непрерывная функция. Предел в среднем квадратичном Y

$$\sum_{k=1}^{n} g(\tau_k)X(\tau_k)(t_k - t_{k-1}) \quad \xrightarrow[\substack{a=t_0<\ldots<t_n=b \\ \max_k \Delta t_k \to 0 \\ \tau_k\in[t_{k-1},t_k)}]{L_2} \quad Y,$$

если он существует, называется *с.к.-интегралом процесса $X(t)$ с непрерывной функцией $g(t)$*

$$Y = \int_a^b g(t)X(t)\,dt,$$

а сам процесс $X(t)$ называется *с.к.-интегрируемым с непрерывной функцией* $g(t)$.

Этот интеграл еще иногда называется *стохастическим интегралом Римана по матожиданию процесса* $X(t)$ *с непрерывной функцией* $g(t)$.

Критерий с.к.-интегрируемости (с непрерывной функцией $g(t)$)

> Следующие утверждения эквивалентны:
>
> (a) Процесс $X(t) \in L_2$ является с.к.-интегрируемым с непрерывной функцией $g(t)$.
>
> (b) Существует интеграл Римана
>
> $$\int\limits_a^b \int\limits_a^b g(t_1)g(t_2)K_X(t_1,t_2)\,dt_1dt_2 < +\infty$$
>
> (c) Существуют интегралы Римана
>
> $$\int\limits_a^b g(t)m_X(t)\,dt < +\infty, \qquad \int\limits_a^b \int\limits_a^b g(t_1)g(t_2)R_X(t_1,t_2)\,dt_1dt_2 < +\infty$$

Из с.к.-непрерывности следует с.к. интегрируемость с любой непрерывной функцией $g(t)$.

Опр. Пусть процесс $X(t) \in L_2$ определен на отрезке $[a,b]$, а $g(t)$ — неслучайная непрерывная функция. Предел в среднем квадратичном Y

$$\sum_{k=1}^{n} g(\tau_k)\big(X(t_k) - X(t_{k-1})\big) \xrightarrow[\substack{a=t_0<\ldots<t_n=b \\ \max_k \Delta t_k \to 0 \\ \tau_k \in [t_{k-1},t_k)}]{L_2} Y,$$

если он существует, называется *интегралом Римана-Стилтьеса процесса* $X(t)$ *с непрерывной функцией* $g(t)$

$$Y = \int\limits_a^b g(t)\,dX(t)$$

Критерий существования интеграла Римана-Стилтьеса (с непрерывной функцией $g(t)$)

> Следующие утверждения эквивалентны:
>
> (a) Существует интеграл Римана-Стилтьеса процесса $X(t)$ с непрерывной функцией $g(t)$.
>
> (b) Существует интеграл Римана-Стилтьеса
>
> $$\int\limits_a^b \int\limits_a^b g(t_1)g(t_2)\,dK_X(t_1,t_2) < +\infty$$
>
> Определяется подобно интегралу Римана-Стилтьеса в $\mathbb{R}$. Мера клетки вводится как
>
> $$\mu\big([a_1,a_2]\times[b_1,b_2]\big) = K_X(a_2,b_2) - K_X(a_1,b_2) - K_X(a_2,b_1) + K_X(a_1,b_1)$$

Далее вводится понятие стохастического интеграла процесса $X(t)$ по винеровскому процессу $W(t)$.

Опр. *Сигма-алгеброй*, порожденной процессом $\{X(t),\ t \in T\}$ называется минимальная сигма-алгебра, относительно которой измеримы все сечения $X(t)$ на множестве $t \in T$:

$$\sigma\{X(t), t \in T\} \overset{\text{def}}{=} \sigma\left\{\big(X(t)\big)^{-1}(B) \;\middle|\; B \in \mathcal{B}(\mathbb{R}),\ t \in T\right\}$$

Так как все $X(t)$ измеримы относительно основной сигма-алгебры $\mathcal{F}$ на общем вероятностном пространстве, то $\sigma\{X(t), t \in T\} \subset \mathcal{F}$.

Опр. Случайный процесс $\{X(t), t \in [a,b]\}$ называется *неупреждающим* относительно случайного процесса $\{W(t), t \in [a,b]\}$, если

$$\sigma\{X(t)\} \subset \sigma\{W(s), s \le t\}, \qquad \forall t \in [a,b]$$

Тут $\sigma\{X(t)\}$ — сигма-алгебра, порожденная случайной величиной $X(t)$. Грубо говоря, данное включение означает, что события для процесса $X(t)$ в момент t связаны с событиями для $W(s)$ только до момента t и не связаны с будущим $W(s)$ (при $s > t$).

Опр. Пусть с.к.-непрерывный процесс $X(t) \in L_2,\ t \in [a,b]$ — неупреждающий относительно винеровского процесса $W(t)$. Предел в среднем квадратичном $I^\theta(X)$

$$\sum_{k=1}^{\infty} X_n(\tau_k^\theta)\big(W(t_{k+1}) - W(t_k)\big) \xrightarrow[\substack{a=t_0<\ldots<t_n=b \\ \max_k \Delta t_k \to 0}]{L_2} I^\theta(X), \qquad \tau_k^\theta = (1-\theta)t_k + \theta t_{k-1},$$

если он существует, называется *стохастическим θ-интегралом процесса $X(t)$ по винеровскому процессу $W(t)$*

$$I^\theta(X) = \int_a^b X(t)\, dW(t)$$

Значение стохастического θ-интеграла зависит от выбора θ.

Опр. Стохастический интеграл при $\theta = 0$ называется *стохастическим интегралом Ито.*

Опр. Стохастический интеграл при $\theta = 1/2$ называется *стохастическим интегралом Стратоновича.*

Исходя из определения можно вычислить следующие стохастические интегралы винеровского процесса:

$$I^\theta(W) = \int_0^T W(t)\, dW(t) = \begin{cases} \dfrac{W^2(T) - T}{2} & , \ \theta = 0 \text{ (стохастический интеграл Ито)} \\ \dfrac{W^2(T)}{2} & , \ \theta = 1/2 \text{ (стохастический интеграл Стратоновича)} \\ \dfrac{W^T(T) + T}{2} & , \ \theta = 1 \end{cases}$$

Опр. Пусть $f(t)$ и $g(t)$ — неупреждающие относительно $W(t)$ случайные процессы ($t \in [0,T]$), такие, что почти все (т.е. с вероятностью 1) их траектории квадратично интегрируемы. Тогда случайный процесс

$$X(t) = X(0) + \underbrace{\int_0^t f(\omega, s)\, ds}_{\text{потраекторный интеграл}} + \underbrace{\int_0^t g(s)\, dW(s)}_{\text{стохастический интеграл Ито}}, \qquad t \in [0,T]$$

называется *процессом Ито.*

Опр. Процесс $X(t)$ *имеет стохастический дифференциал*

$$dX(t) = f(\omega,t)dt + g(\omega,t)dW(t),$$

если $X(t)$ является процессом Ито с соответствующими процессами $f(t)$ и $g(t)$.

То есть запись стохастического дифференциала — просто краткая запись определения процесса Ито.

Теорема (формула Ито)

Пусть

- неслучайная функция $F(t,x)$ непрерывно дифференцируема по $t \ge 0$ и дважды непрерывно дифференцируема по $x \in \mathbb{R}$,

- процесс $X(t)$ имеет стохастический дифференциал

$$dX(t) = f(\omega,t)dt + g(\omega,t)dW(t),$$

Тогда процесс $F(t, X(t)),\ t \geq 0$ имеет стохастический дифференциал, определяемый *формулой Ито*:

$$dF(t, X(t)) = \left[\frac{\partial F}{\partial t} + f(t)\frac{\partial F}{\partial x} + \frac{1}{2}g^2(t)\frac{\partial^2 F}{\partial x^2}\right] dt + \frac{\partial F}{\partial x} g(t)\, dW(t)$$

Формулу Ито можно запомнить в виде

$$dF = \frac{\partial F}{\partial t}dt + \frac{\partial F}{\partial x}dX(t) + \frac{1}{2}\frac{\partial^2 F}{\partial x^2}\big(dX(t)\big)^2, \qquad dt\, dt = 0, \quad dW(t)\, dt = 0, \quad \big(dW(t)\big)^2 = dt$$

Рассмотрим пример использования этой формулы. Вычислим стохастический интеграл Ито $Y(t) = \int_0^t W(s)\, dW(s)$.

Заметим, что $Y(t)$ — процесс Ито при $Y(0) \overset{\text{п.н.}}{=} 0,\ f_Y(t) \equiv 0,\ g_Y(t) = W(t)$. Его стохастический дифференциал:

$$dY(t) = W(t)\, dW(t)$$

Возьмем в формуле Ито $X(t) = W(t)$, тогда $f(t) \equiv 0,\ g(t) \equiv 1$. Найдем такую функцию $F(t, x)$, что $dF(t, X(t))$ как-то похоже на $dY(t)$. Возьмем $F(t, x) = x^2$, тогда по формуле Ито

$$d\big(W^2(t)\big) = dt + 2\underbrace{W(t)\, dW(t)}_{dY(t)}$$

По определению процесса Ито, это запись эквивалентна

$$W^2(t) = W^2(0) + \int_0^t ds + 2\int_0^t W(s)\, dW(s) = t + 2Y(t)$$

Отсюда находим

$$Y(t) = \frac{W^2(t) - t}{2}$$

5 Стационарность и эргодичность случайных процессов

5.1 Стационарность в широком и узком смыслах

Опр. Пусть $X(t),\ Y(t),\ t \in T$ — случайные процессы на вероятностном пространстве $(\Omega, \mathcal{F}, \mathbb{P})$. Тогда

$$Z(\omega, t) = X(\omega, t) + iY(\omega, t), \qquad \omega \in \Omega, \qquad t \in T$$

называется *комплекснозначным случайным процессом.*

Матожидание и корреляционная функция определяются как

$$m_Z(t) = \mathbb{E}Z(t) = \mathbb{E}X(t) + i\mathbb{E}Y(t), \qquad R_Z(t, s) = \mathbb{E}\overset{\circ}{Z}(t)\overline{\overset{\circ}{Z}(s)}$$

Опр. Процесс $Z(t)$ называется *комплекснозначным процессом второго порядка*, если

$$\mathbb{E}\big|Z(t)\big|^2 < +\infty, \qquad \forall t \in T$$

Множество всех таких процессов будем обозначать CL_2.

Опр. Процесс $X(t)$ называется *стационарным в узком смысле* или *сильно стационарным*, если $\forall n \in \mathbb{N}$, $\forall t_1, \dots, t_n, \ \ \forall h > 0$

$$\begin{bmatrix} X(t_1) \\ \vdots \\ X(t_n) \end{bmatrix} \overset{d}{=} \begin{bmatrix} X(t_1 + h) \\ \vdots \\ X(t_n + h) \end{bmatrix}$$

Данное определение применимо не только для процессов класса CL_2.

Опр. Процесс $X(t) \in CL_2$ называется *стационарным в широком смысле* или *слабо стационарным*, если $m_X(t) \equiv \text{const}$ и $R_X(t_1, t_2)$ зависит только от разности $t_2 - t_1$.

Процесс $X(t) \equiv \xi \in \text{Cauchy}(0, 1)$ не лежит в CL_2, но он стационарен в узком смысле. Определение стационарности в широком смысле к нему не применимо.

Свойства стационарных процессов:

1. $X(t) \in CL_2$ стационарен в узком смысле $\implies$ $m_X(t) \equiv \text{const}$, $R_X(t_1, t_2) = R(t_2 - t_1)$, где $R(\tau)$ — некоторая функция одного аргумента.

2. $X(t) \in CL_2$ стационарен в узком смысле $\implies$ $X(t)$ стационарен в широком смысле.

3. $R_X(t_1, t_2) = R(t_2 - t_1) \implies D_X(t) \equiv \text{const}$.

4. Процесс вида $X(t) = \xi f(t)$, где $\mathbb{E}\xi = 0$, $\mathbb{E}|\xi|^2 < +\infty$ и неслучайная функция $f(t)$ непрерывна, стационарен в широком смысле $\iff$

$$X(t) = \xi e^{i(\omega t + \theta)}, \qquad \omega, \theta \in \mathbb{R}$$

5. Процесс вида

$$X(t) = \sum_{k=1}^{n} \xi_k e^{i\omega_k t}, \qquad \omega_k \neq \omega_m, \qquad \mathbb{E}\xi_k = 0$$

стационарен в широком смысле $\iff$ $\mathbb{E}\big[\xi_k \overline{\xi_m}\big] = 0$, то есть с.в. ξ_k попарно некоррелированы. При этом корреляционная функция имеет вид

$$R_X(t, s) = R(\tau) = \sum_{k=1}^{n} \mathbb{E}|\xi_k|^2 e^{i\omega_k \tau}, \qquad \tau = t - s$$

Теорема. Пусть $X(t)$ — гауссовский процесс. Тогда

$$X(t) \text{ стационарен в узком смысле} \qquad \iff \qquad X(t) \text{ стационарен в широком смысле}$$

Опр. Пусть процесс $X(t)$ стационарен в широком смысле. Функция

$$R_X(t) \stackrel{\text{def}}{=} R_X(t,0)$$

называется *корреляционной функцией стационарного процесса в широком смысле.*

Свойства функции $R_X(t)$: (для стационарного в широком смысле процесса $X(t)$)

1. $R_X(0) = \mathbb{V}X(0) \equiv \mathbb{V}X(t) \geq 0$
2. $R_X(t)$ — *эрмитова функция*, то есть обладает свойством
$$R_X(-t) = \overline{R_X(t)}, \qquad \forall t \in \mathbb{R}$$
3. $\left|R_X(t)\right| \leq R_X(0), \quad \forall t \in \mathbb{R}$

   Если $X(t)$ — действительнозначный процесс и существуют производные R'_X и R''_X, то из условия максимума
$$R'_X(0) = 0, \qquad R''_X(0) \leq 0$$
4. $R_X(t)$ — неотрицательно определенная функция (одного переменного), то есть
$$\sum_{i=1}^{n}\sum_{j=1}^{n} z_i\overline{z_j}R_X(t_i - t_j) \geq 0, \qquad \forall z_1,\ldots,z_n \in \mathbb{C}, \qquad \forall t_1,\ldots,t_n, \qquad \forall n \in \mathbb{N}$$

   Для любого (необязательно стационарного) процесса корреляционная функция $R_X(t,s)$ — неотрицательно определенная функция двух переменных, то есть
$$\sum_{i=1}^{n}\sum_{j=1}^{n} z_i\overline{z_j}R_X(t_i, t_j) \geq 0, \qquad \forall z_1,\ldots,z_n \in \mathbb{C}, \qquad \forall t_1,\ldots,t_n, \qquad \forall n \in \mathbb{N}$$
5. Если $R_X(t)$ непрерывна в $t = 0$, то она непрерывна всюду на $\mathbb{R}$.
6. Следующие классы функций вида $R(t,s)$ совпадают:
   - класс неотрицательно определенных комплекснозначных функций двух переменных;
   - класс корреляционных функций комплекснозначных процессов второго порядка;
   - класс корреляционных функций комплекснозначных гауссовских процессов.

   Следующие классы функций вида $R(t)$ совпадают:
   - класс неотрицательно определенных комплекснозначных функций одной переменной;
   - класс корреляционных функций стационарных в широком смысле комплекснозначных процессов второго порядка;
   - класс корреляционных функций комплекснозначных стационарных в широком смысле гауссовских процессов.

   Те же утверждения верны и для действительных процессов и корреляционных функций.
7. Выполнено равенство
$$\frac{1}{T^2}\int\limits_0^T\int\limits_0^T R_X(t_1,t_2)\,dt_1dt_2 = \frac{2}{T}\int\limits_0^T \left(1 - \frac{1}{T}\right) R_X(\tau)d\tau$$

   Этой теоремой удобно пользоваться для проверки критерия эргодичности для стационарных процессов.

5.2 Спектральное представление процессов

Теорема Бохнера-Хинчина

Функция $R(\tau)$ является корреляционной функцией некоторого стационарного в широком смысле и с.к.-непрерывного случайного процесса $\iff$ $R(\tau)$ представима в виде

$$R(\tau) = \int\limits_{-\infty}^{+\infty} e^{i\nu\tau} dS(\nu),$$

где $S(\nu) = \mathbb{E}|\xi|^2 \cdot \mathbb{P}\{\Omega < \nu\}$ для некоторых случайных величин ξ и Ω.

Интеграл понимается как интеграл Лебега-Стилтьеса по функции $S(\nu)$, которая не убывает, так как пропорциональна функции распределения случайной величины Ω.

Опр. Функция $S(\nu)$ из теоремы Бохнера-Хинчина называется *спектральной функцией*.
Для процессов в дискретном времени ($t \in \mathbb{Z}_+$) есть аналог этой теоремы, называемый теоремой Герглотца:

$$R(n) = \int\limits_{-\pi}^{\pi} e^{i\nu n} dS(\nu), \qquad n \in \mathbb{Z}$$

Опр. Если функция $S(\nu)$ абсолютно непрерывна, то есть существует неотрицательная функция $\rho(\nu)$, такая, что

$$S(\nu) = \int\limits_{-\infty}^{\nu} \rho(\tau) d\tau, \qquad \forall \nu \in \mathbb{R}$$

то функция $\rho(\nu)$ называется *спектральной плотностью*.

При этом корреляционная функция стационарного процесса представима в виде (случаи непрерывного и дискретного времени)

$$R_X(\tau) = \int\limits_{-\infty}^{+\infty} e^{i\nu\tau} \rho(\nu) d\nu, \qquad R_X(n) = \int\limits_{-\pi}^{\pi} e^{i\nu n} \rho(\nu) d\nu$$

В случае наличия спектральной плотности интегралы уже понимаются как интегралы Римана.

Дисперсия выражается через спектральную плотность как

$$\mathbb{V}X(t) = R_X(0) = \int\limits_{-\infty}^{+\infty} \rho(\nu) d\nu$$

Достаточное условие существования спектральной плотности (наличие интеграла Фурье)

Пусть корреляционная функция $R_X(\tau)$ с.к.-непрерывного процесса абсолютно интегрируема, т.е.

$$\int\limits_{-\infty}^{+\infty} \big|R_X(\tau)\big| d\tau < +\infty$$

Тогда спектральная функция процесса $X(t)$ обладает непрерывной и ограниченной спектральной плотностью $\rho(\nu)$, причем

$$\rho(\nu) = \frac{1}{2\pi} \int\limits_{-\infty}^{+\infty} e^{-i\nu\tau} R_X(\tau) d\tau$$

Если корреляционная функция $R_X(\tau)$ вещественная (а, значит, и четная), то выражение для спектральной плотности упрощается:

$$\rho(\nu) = \frac{1}{\pi} \int\limits_{0}^{+\infty} \cos(\nu\tau) R_X(\tau) d\tau$$

Опр. Случайный процесс $V(t) \in CL_2$ называется *процессом с ортогональными приращениями*, если

$$\mathbb{E}\big(V(t_4) - V(t_3)\big)\overline{\big(V(t_2) - V(t_1)\big)} = 0, \qquad \forall t_1 < t_2 \leq t_3 < t_4$$

Теорема Крамера (спектральное разложение стационарного процесса)

> Пусть процесс $X(t) \in CL_2$ с.к.-непрерывен и стационарен в широком смысле. Тогда существует случайный процесс $V(\nu) \in CL_2$, заданный на том же вероятностном пространстве, такой, что
>
> - $V(\nu)$ — процесс с ортогональными приращениями,
> - почти наверное выполнено
>
> $$X(t) = m_X + \int\limits_{-\infty}^{+\infty} e^{i\nu t}\, dV(\nu), \qquad \forall t \in \mathbb{R},$$
>
> где интеграл понимает в смысле Римана-Стилтьеса.

Для процессов в дискретном времени есть аналогичное разложение, только интеграл берется по отрезку $[-\pi, \pi]$.

Процесс $V(\nu)$ иногда называется *случайной ортогональной мерой*. Он связан со спектральной функцией $S(\nu)$ соотношением

$$dS(\nu) = \mathbb{E}|dV(\nu)|^2$$

Свойства спектральных функций и разложений

1. Пусть $X(t), Y(t) \in CL_2$ — центрированные стационарные процессы со спектральными плотностями $\rho_X(\nu)$ и $\rho_Y(\nu)$. Пусть для них справедливы спектральные разложения

$$X(t) = \int\limits_{-\infty}^{+\infty} e^{i\nu t} dV_X(\nu), \qquad Y(t) = \int\limits_{-\infty}^{+\infty} e^{i\nu t}\Phi(\nu) dV_X(\nu),$$

где $\Phi(\nu)$ — некоторая квадратично интегрируемая функция по мере $S_X(\nu)$. Тогда

$$\rho_Y(\nu) = \big|\Phi(\nu)\big|^2 \rho_X(\nu)$$

2. Производную можно вносить под знак стохастического интеграла:

$$X(t) = \int\limits_{-\infty}^{+\infty} e^{i\nu t} dV_X(\nu) \qquad \Longrightarrow \qquad X^{(k)}(t) = \int\limits_{-\infty}^{+\infty} e^{i\nu t} (i\nu)^k dV_X(\nu),$$

если выполнено условие $\int_{\mathbb{R}} \nu^{2k} dS_X(\nu) < \infty$.

3. *Закон больших чисел для стационарных в широком смысле процессов:* (для дискретного и непрерывного времени)

$$\frac{1}{N}\sum_{t=1}^{N} X_t \xrightarrow[N\to\infty]{L_2} m_X + V(\{0\}),$$

$$\frac{1}{T}\int\limits_0^T X(t)dt \xrightarrow[T\to\infty]{L_2} m_X + V(\{0\}),$$

где интеграл понимается в среднем квадратичном, а случайная величина $V(\{0\}) = V(0) - V(0-0)$ есть случайная мера нуля.

Данное свойство можно неформально вывести из спектрального разложения стационарного процесса:

$$\frac{1}{N}\sum_{t=1}^{N} X_t = m_X + \int\limits_{-\pi}^{\pi} \left(\frac{1}{N}\sum_{t=1}^{N} e^{i\nu t}\right) dV(\nu) \xrightarrow[N\to\infty]{} m_X + \int\limits_{-\pi}^{\pi} \left\{ \begin{array}{l} 1,\ \nu = 0 \\ 0,\ \nu \neq 0 \end{array} \right\} dV(\nu) = m_X + V(\{0\})$$

5.3 Эргодические случайные процессы

Опр. Вещественный случайный процесс $X(t) \in L_2,\ t \geq 0$ называется *эргодичным по матожиданию*, если его матожидание постоянно $m_X(t) \equiv m_X \in \mathbb{R}$ и

$$\langle X \rangle_T \overset{\text{def}}{=} \frac{1}{T} \int\limits_0^T X(t)\,dt \xrightarrow[T\to\infty]{L_2} m_X$$

Здесь процесс $X(t)$ подразумевается с.к.-интегрируемым на любом отрезке вида $[0, T],\ T > 0$, а интеграл понимается в смысле с.к.-интеграла.

Понятие эргодичности не связано с понятием стационарности, то есть из наличия одного свойства не следует наличие другого и наоборот.

Критерий эргодичности

> Процесс $X(t) \in L_2$ с постоянным матожиданием $m_X(t) \equiv m_X$ является эргодичным по матожиданию тогда и только тогда, когда
>
> $$\frac{1}{T^2} \int\limits_0^T \int\limits_0^T R_X(t_1, t_2) dt_1 dt_2 \xrightarrow[T\to\infty]{} 0$$

Достаточное условие эргодичности

> Пусть $X(t) \in L_2$ — процесс с постоянным матожиданием $m_X(t) \equiv m_X$ и выполнено
>
> $$\lim_{|t_1 - t_2| \to \infty} R_X(t_1, t_2) = 0$$
>
> Тогда $X(t)$ является эргодичным по матожиданию.

Это достаточное условие удобно применять для стационарных процессов.

Опр. Случайный процесс $X(t) \in L_2$ называется *эргодичным по дисперсии*, если случайный процесс

$$Y(t) = \big(X(t) - m_X(t)\big)^2$$

эргодичен по матожиданию.

Опр. Случайный процесс $X(t) \in L_2$ называется *эргодичным по корреляционной функции*, если для любого $\tau \geq 0$ случайный процесс

$$Y_\tau(t) = \overset{\circ}{X}(t)\overset{\circ}{X}(t+\tau)$$

эргодичен по матожиданию.

Опишем практический смысл эргодических случайных процессов. Допустим, у нас есть одна реализация на отрезке $[0, T]$ процесса $X(t)$ с постоянным матожиданием $m_X(t) \equiv m$. Как нам оценить это матожидание?

Рассмотрим случайную величину

$$\widehat{m} = \frac{1}{T} \int\limits_0^T X(t)\,dt,$$

где интеграл понимается в среднем квадратичном. При фиксированном исходе ω — это просто интеграл Римана от данной нам реализации $X(t)$. Покажем, что это будет хорошей оценкой.

$$\mathbb{E}\widehat{m} = \frac{1}{T} \int\limits_0^T m_X(t)\,dt = m$$

Но нам дана только одна реализация. Как нам гарантировать, значение $\widehat{m}$ на ней будет близко к матожиданию? Вычислим для этого дисперсию:

$$\mathbb{V}\widehat{m} = \frac{1}{T^2} \int\limits_0^T \int\limits_0^T R_X(t, s)\,dt\,ds$$

Если процесс $X(t)$ эргодический, то по критерию эргодичности, $\mathbb{V}\widehat{m} \xrightarrow{T\to\infty} 0$, что означает, что если данная нам реализация достаточно длинная, то оценка $\widehat{m}$ будет сколь угодно близка к истинному матожиданию m.

5.4 Эргодические динамические системы и эргодическая теорема

Опр. Совокупность $M = (\Omega, \mathcal{F}, \mathbb{P}, T)$, где $(\Omega, \mathcal{F}, \mathbb{P})$ — вероятностное пространство, а отображение $T : \Omega \to \Omega$ обладает свойством:

$$\mathbb{P}\big\{T^{-1}(B)\big\} = \mathbb{P}\{B\}, \qquad \forall B \in \mathcal{F}$$

называется *динамической системой.*

То есть вероятностная мера $\mathbb{P}$ инвариантна относительно преобразования T.

Опр. Будем говорить, что события $A, B \in \mathcal{F}$ *равны почти наверное*, если мера их симметрической разности равна 0:

$$A \overset{\text{п.н.}}{=} B \qquad \Longleftrightarrow \qquad \mathbb{P}\big\{(A \cup B) \setminus (A \cap B)\big\} = 0$$

Опр. Динамическая система T называется *эргодической динамической системой*, если имеет место следствие

$$T^{-1}(B) \overset{\text{п.н.}}{=} B \qquad \Longrightarrow \qquad B \overset{\text{п.н.}}{=} \varnothing \;\text{ или }\; B \overset{\text{п.н.}}{=} \Omega$$

Пусть $(\Omega, \mathcal{F}, \mathbb{P}, T)$ — динамическая система. Пространство случайных величин на вероятностном пространстве $(\Omega, \mathcal{F}, \mathbb{P})$ будем обозначать $L_2(\Omega, \mathcal{F}, \mathbb{P})$.

Теорема фон Неймана (эргодическая теорема для L_2-сходимости)

> Пусть $(\Omega, \mathcal{F}, \mathbb{P}, T)$ — эргодическая динамическая система. Тогда для любой случайной величины $\xi \in L_2(\Omega, \mathcal{F}, \mathbb{P})$ выполняется
>
> $$\frac{1}{N}\sum_{k=1}^{N} \xi(T^k\omega) \xrightarrow[N\to\infty]{L_2} \mathbb{E}\xi$$

Теорема Биркгофа-Хинчина (эргодическая теорема для сходимости почти наверное)

> Пусть $(\Omega, \mathcal{F}, \mathbb{P}, T)$ — эргодическая динамическая система. Тогда для любой случайной величины $\xi \in L_2(\Omega, \mathcal{F}, \mathbb{P})$ выполняется
>
> $$\frac{1}{N}\sum_{k=1}^{N} \xi(T^k\omega) \xrightarrow[N\to\infty]{\text{п.н.}} \mathbb{E}\xi$$

Теорема фон Неймана фактически утверждает, что случайная последовательность $X_k = \xi(T^k\omega) \in L_2$ является эргодической по матожиданию.

Действительно, матожидание постоянно:

$$\mathbb{E}X_k = \int\limits_{\Omega} \xi(T^k\omega)\, d\mathbb{P}(\omega) = \Big/ \begin{array}{c}\text{замена } \omega = T^{-k}\omega' \\ \text{и инвариантность меры } \mathbb{P}\end{array} \Big/ = \int\limits_{\Omega} \xi(\omega')\, d\mathbb{P}(\omega') = \mathbb{E}\xi$$

и усреднение стремится к $\mathbb{E}\xi$ в среднем квадратичном по теореме фон Неймана. Это определение эргодичности.

Данные теоремы можно переформулировать на языке теории меры. В частности:

Теорема Биркгофа-Хинчина (на языке теории меры)

> Пусть $(X, \mathcal{F}, \mu, T)$ — эргодическая динамическая система. Тогда для любой квадратично интегрируемой по мере μ функции $f \in L_2(X, \mathcal{F}, \mu)$ имеет место сходимость μ-почти всюду
>
> $$\frac{1}{N}\sum_{k=1}^{N} f(T^k x) \xrightarrow[N\to\infty]{\mu-\text{п.в.}} \int\limits_{X} f(t)\, d\mu(t), \qquad x \in X$$

6 Дискретные цепи Маркова

6.1 Определение дискретной марковской цепи

Опр. Случайная последовательность $\{X_k\}_{k=0}^{\infty}$ называется *дискретной цепью*, если множество значений E случайных величин X_k не более чем счетно.

Обычно $E = \{0, 1, 2, \ldots\},\ |E| \le \infty$.

Опр. Элементы множества E называются *состояниями* дискретной цепи $\{X_k\}_{k=0}^{\infty}$.
Слово "дискретная" в этом определении указывает на дискретность времени, а "цепь" — на дискретность множества состояний.

Опр. Дискретная цепь $\{X_k\}_{k=0}^{\infty}$ называется *дискретной марковской цепью (ДМЦ)*, если она обладает *марковским свойством*:

$$\mathbb{P}\big\{X_n = i_n \bigm| X_{n-1} = i_{n-1}, \ldots, X_0 = i_0\big\} = \mathbb{P}\big\{X_n = i_n \bigm| X_{n-1} = i_{n-1}\big\}, \qquad \forall n \ge 1,\ i_k \in E$$

Опр. Число $p_{i,j} = \mathbb{P}\{X_n = j \mid X_{n-1} = i\}$ называется *вероятностью перехода из состояния i в состояние j за 1 шаг в момент времени n.*

Опр. *Переходной матрицей* (в момент времени n) называется матрица

$$P(n) = \begin{bmatrix} p_{0,0} & p_{0,1} & \cdots & p_{0,N} \\ p_{1,0} & p_{1,1} & & \vdots \\ \vdots & & \ddots & \vdots \\ p_{N,0} & \cdots & \cdots & p_{N,N} \end{bmatrix}, \qquad N = |E| - 1 \le \infty$$

Переходная матрица $P(n)$ может быть бесконечной. Сумма элементов по строкам равна 1.

Опр. ДМЦ называется *однородной дискретной марковской цепью (ОДМЦ)*, если переходная матрица $P(n) \equiv P$ не зависит от момента времени n.

Опр. Вероятность $p_i(n) = \mathbb{P}\{X_n = i\}$ называется *вероятностью состояния i в момент $n \ge 0$.*

Опр. Вектор $p(n) = \big[p_0(n), p_1(n), \ldots\big]^T$ называется *распределением вероятностей состояний в момент $n \ge 0$.*

Задать ОДМЦ $\equiv$ задать начальное распределение состояний $p(0)$ и переходную матрицу P.

Опр. Для ОДМЦ число $p_{i,j}(m) = \mathbb{P}\big\{X_m = j \bigm| X_0 = i\big\}$ называется *вероятностью перехода из состояния i в состояние j за m шагов.*

Свойства переходной матрицы $P(n)$:

1. Распределения вероятностей состояний в моменты $n-1$ и n связаны соотношением

$$p(n) = P(n)^T\, p(n-1)$$

   Для ОДМЦ выполнены равенства:

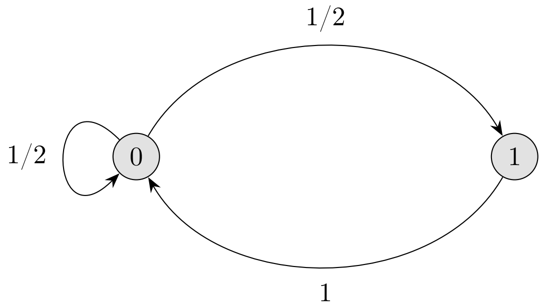
$$p(m) = (P^T)^m\, p(0), \qquad p_{i,j}(m) = \big(P^m\big)_{ij}$$

2. *Уравнение Колмогорова-Чепмена:* для ОДМЦ для любых $m, n \ge 1$ выполнено

$$p_{i,j}(m+n) = \sum_{k \in E} p_{i,k}(m)\, p_{k,j}(n)$$

3. Матрицы P и P^T всегда имеют собственное значение $\lambda = 1$. Остальные собственные значения по модулю не превосходят единицы.

ОДМЦ удобно представлять в виде *стохастического графа*, в вершинах которого находятся состояние цепи, а веса ребер равны вероятностям перехода за один шаг. Например, матрице $P = \begin{bmatrix} 1/2 & 1/2 \\ 1 & 0 \end{bmatrix}$ соответствует граф



6.2 Классификация состояний ОДМЦ

Есть два способа классификации состояний: по алгебраическим и по асимптотическим свойствам.

Введем следующие обозначения:

- Вероятность того, что за n шагов мы впервые вернулись в i-ое состояние:

$$f_i(n) = \mathbb{P}\big\{X_n = i,\ X_k \neq i,\ k = \overline{1, n-1} \ \big|\ X_0 = i\big\}, \qquad f_i(0) \overset{\text{def}}{=} 0$$

- Производящая функция последовательности $\{p_{i,i}(n)\}_{n=0}^{\infty},\ p_{i,i}(0) = 1$:

$$P_i(z) = \sum_{n=0}^{\infty} p_{i,i}(n)\, z^n$$

- Производящая функция последовательности $\{f_i(n)\}_{n=0}^{\infty}$:

$$F_i(z) = \sum_{n=0}^{\infty} f_i(n)\, z^n$$

  Справедливо равенство (если ряды сходятся):

$$P_i(z) - 1 = F_i(z)\, P_i(z)$$

- Вероятность вернуться в i-ое состояние за конечное число шагов

$$F_i = F_i(1) = \sum_{n=1}^{\infty} f_i(n)$$

а) Классификация на основе алгебраических свойств

Опр. Состояние i называется *несущественным*, если существует состояние $j \in E$, такое, что

$$\exists m \geq 1 : \quad p_{i,j}(m) > 0, \qquad \forall n \geq 1 : \quad p_{j,i}(n) = 0$$

Иначе состояние i называется *существенным*.

Состояние i несущественно $\iff$ есть такая вершина j в стохастическом графе, куда есть путь из i, но обратного пути из j в i нет.

Опр. Состояния i и j называются *сообщающимися*, если

$$\exists m \geq 1 : \quad p_{i,j}(m) > 0, \qquad \exists n \geq 1 : \quad p_{j,i}(n) > 0$$

Иначе состояния i и j называются *несообщающимися.*

Состояния i и j сообщаются $\iff$ в стохастическом графе есть путь между от вершины i к j и обратно.

Отношение сообщаемости является отношением эквивалентности, поэтому все состояния цепи разбиваются на классы эквивалентности из сообщающихся состояний. В одном классе могут быть либо только существенные состояния, либо только несущественные состояния.

Опр. Класс эквивалентности называется *открытым*, если он состоит из несущественных состояний.

Опр. Класс эквивалентности называется *замкнутым*, если он состоит из существенных состояний.

Опр. Если замкнутый класс состоит из одного состояния, то это состояние называется *поглощающим.*

Опр. Если в цепи все состояния сообщаются, то есть имеется только 1 класс эквивалентности, то такая цепь называется *неразложимой.* Иначе цепь называется *разложимой.*

Наличие классов эквивалентности приводит к блочной структуре матрицы P (если нужным образом занумеровать состояния), потому что из замкнутого класса нельзя попасть в другой замкнутый класс или в открытый класс.

Опр. Пусть $\exists n_0 \geq 1: \ p_{i,i}(n_0) > 0$. Пусть

$$d_i = \text{НОД}\big\{n \geq 1 \ \big| \ p_{i,i}(n) > 0\big\}$$

Если $d_i > 1$, то состояние i называется *периодическим с периодом* d_i. Если $d_i = 1$, то состояние i называется *апериодическим.*

Если одно состояние в классе эквивалентности E_* имеет период d, то все состояния из E_* имеют период d. При этом $E_* = \bigcup_{j=0}^{d-1} C_j$ — класс разбивается на d *циклических подклассов.* Из подкласса C_j можно попасть только в подкласс C_{j+1}. Это тоже приводит к блочной структуре подматрицы класса E_*.

б) Классификация на основе асимптотических свойств

Опр. Состояние i называется *нулевым*, если существует предел

$$\lim_{n\to\infty} p_{i,i}(n) = 0$$

Если этот предел не равен нулю или вообще не существует, то состоянии i называется *ненулевым* или *положительным.*

Опр. Состояние i называется *невозвратным*, если выполнено одно из эквивалентных свойств:

(a) $F_i < 1$ (вероятность вернуться в i за конечное число шагов).

(b) Цепь возвращается в i-ое состояние бесконечно часто с нулевой вероятностью:

$$\mathbb{P}\left\{\bigcap_{n=1}^{\infty}\bigcap_{k=n}^{\infty}\{X_k = i\} \ \middle| \ X_0 = i\right\} = 0$$

(c) $\displaystyle\sum_{n=0}^{\infty} p_{i,i}(n) < +\infty.$

Иначе состояние i называется *возвратным.*

В задачах удобнее всего пользоваться определением (c), оно часто называется *критерием возвратности.*

Эквивалентность этих определений частично следует из леммы Бореля-Кантелли и равенства $P_i(z) - 1 = F_i(z)P_i(z)$.

Из определения следует, что невозвратное состояние всегда нулевое.

Опр. Нулевое возвратное состояние называется *нуль возвратным.*

Опр. Ненулевое возвратное состояние называется *положительно возвратным.*

Лемма Бореля-Кантелли

(1) Если $\sum_{n=1}^{\infty} \mathbb{P}\{A_n\} < +\infty$, то события A_n *происходят бесконечно часто* с нулевой вероятностью:

$$\mathbb{P}\left\{\bigcap_{n=1}^{\infty} \bigcup_{k=n}^{\infty} A_k\right\} = 0$$

(2) Если события $\{A_n\}$ независимы и $\sum_{n=1}^{\infty} \mathbb{P}\{A_n\} = +\infty$, то

$$\mathbb{P}\left\{\bigcap_{n=1}^{\infty} \bigcup_{k=n}^{\infty} A_k\right\} = 1$$

Теорема солидарности

Для неразложимой ОДМЦ (конечной или счетной) справедливо, что

(1) Если хотя бы одно состояние возвратное, то все состояния возвратные.

(2) Если хотя бы одно состояние нулевое, то все состояния нулевые.

(3) Если хотя бы одно состояние имеет период d, то все состояния имеют период d.

Если хотя бы одно состояние апериодично, то все состояния апериодичны.

Данная теорема утверждает, что внутри одного класса эквивалентности (открытого или замкнутого) все состояния классифицируются одинаково. Состояния открытых классов обычно не классифицируются, потому что там все состояния невозвратные и нулевые.

Опр. Распределение вероятностей $\pi = (\pi_0, \pi_1, \ldots)$, $\sum_i \pi_i = 1$, называется *стационарным* или *инвариантным*, если $P^T \pi = \pi$.

Если состояние i возвратно, то $F_i = \sum_{n=1}^{\infty} f_i(n) = 1$, значит, последовательность $\{f_i(n)\}$ образует распределение вероятностей случайной величины σ_i — времени первого возвращения в i-ое состояние.

Если i-ое состояние нуль возвратно, то $\mu_i \stackrel{\text{def}}{=} \mathbb{E}\sigma_i = \sum_{n=1}^{\infty} n\, f_i(n) = \infty$. Если i-ое состояние положительно возвратно, то $\mu_i = \frac{1}{\pi_i} > 0$, где π — стационарное распределение.

в) Связи между различными свойствами

1. В конечной цепи всегда есть существенное состояние (а, значит, и замкнутый класс эквивалентности). В счетной цепи существенного состояния может не быть.

2. Для счетных цепей и классов эквивалентности:

$$\text{несущественное состояние} \Longrightarrow \text{невозвратное состояние} \Longrightarrow \text{нулевое состояние}$$

$$\text{существенное состояние} \Longleftarrow \text{возвратное состояние} \Longleftarrow \text{ненулевое состояние}$$

3. Для конечных цепей и классов эквивалентности: (верно в обе стороны)

$$\text{несущественное состояние} \Longleftrightarrow \text{невозвратное состояние} \Longleftrightarrow \text{нулевое состояние}$$

$$\text{существенное состояние} \Longleftrightarrow \text{возвратное состояние} \Longleftrightarrow \text{ненулевое состояние}$$

4. Периодичность состояния не связана с его существенностью и возвратностью.

6.3 Стационарные распределения в ОДМЦ

Опр. Распределение вероятностей $\pi = (\pi_0, \pi_1, \ldots)$, $\sum_i \pi_i = 1$, называется *стационарным* или *инвариантным*, если $P^T\pi = \pi$.

То есть π — неотрицательный и нормированный на единицу собственный вектор матрицы P^T, соответствующий собственному значению $\lambda = 1$.

Свойства матрицы P и стационарных распределений:

1. ОДМЦ (конечная или счетная) — стационарный в узком смысле случайный процесс $\iff$ начальное распределение $p(0)$ стационарно.

2. Матрицы P и P^T всегда имеют собственное значение $\lambda = 1$. Остальные собственные значения по модулю не превосходят единицы.

   Из теоремы Перрона-Фробениуса следует, что существует собственный вектор, соответствующий $\lambda = 1$, с неотрицательными компонентами. В случае конечной цепи его можно отнормировать и получить стационарное распределение.

3. Алгебраическая кратность $\lambda = 1$ матрицы P (или P^T) равна $m \iff$ ОДМЦ (конечная или счетная) содержит m замкнутых классов эквивалентности.

   В частности, если замкнутый класс единственен, то и стационарное распределение (если имеется) единственно.

4. ОДМЦ содержит замкнутый класс эквивалентности с периодом $d > 1 \implies$ существуют d собственных чисел матрицы P (или P^T), равных по модулю 1:

$$\lambda_k = \exp\left(i\,\frac{2\pi k}{d}\right), \qquad k = \overline{0, d-1}$$

5. ОДМЦ конечна $\implies$ существует стационарное распределение.

   У счетных цепей не всегда имеется стационарное распределение. Для них есть свое достаточное условие (см. пункт 8).

6. $$\left.\begin{array}{c}\pi \text{ — стационарное распределение ОДМЦ (конечной или счетной)} \\ \text{i — несущественное состояние}\end{array}\right\} \implies \pi_i = 0.$$

7. Пусть ОДМЦ содержит $m \le \infty$ замкнутых классов $E_1, \ldots, E_m$, а E_0 — все состояния открытый классов. Каждый замкнутый класс E_k, как самостоятельная подцепь, имеет единственное стационарное распределение $\pi^{(k)}$.

   Тогда ОДМЦ имеет стационарное распределение

$$\pi = \Big[\underbrace{0, \ldots, 0}_{|E_0|}, \alpha_1\pi^{(1)}, \ldots, \alpha_m\pi^{(m)}\Big], \qquad \sum_{k=1}^{m} \alpha_k = 1, \qquad \alpha_k \ge 0$$

   для любых α_k, удовлетворяющих этим условиям.

8. ОДМЦ содержит положительно возвратный класс эквивалентности $\implies$ существует стационарное распределение.

   Если такой класс единственен, то и стационарное распределение единственное. Если не единственен, то, как и в пункте 7, возникает смесь стационарных распределений.

   Пример: случайное блуждание на $\mathbb{Z}_+$. При $p > \frac{1}{2}$ оно невозвратно и нет стационарного распределения, при $p = \frac{1}{2}$ оно нуль возвратно и нет стационарного распределения, при $p = \frac{1}{2}$ оно положительно возвратно и есть стационарное распределение.

9. Для неразложимой ОДМЦ (счетной или конечной):

$$\begin{array}{c}\text{цепь положительно}\\ \text{возвратна}\end{array} \iff \begin{array}{c}\text{существует}\\ \text{стационарное распределение}\end{array}$$

   И при этом это стационарное распределение единственно.

Итак, **основной результат**:

> Для конечных ОДМЦ:
>
> существует единственный замкнутый класс эквивалентности $\iff$ существует единственное стационарное распределение
>
> Для *неразложимых* счетных ОДМЦ:
>
> цепь положительно возвратна $\iff$ существует единственное стационарное распределение

Как найти стационарное распределение?

(a) Решить систему

$$P^T\pi = \pi, \qquad \sum_{i\in E}\pi_i = 1, \qquad \pi_i \geq 0$$

Часто ее решение сводится к решению рекуррентного уравнения.

(b) Если матрица P — *дважды стохастическая* (суммы по строкам и по столбцам равны 1), то равномерное распределение является стационарным:

$$\pi_i = \frac{1}{|E|}, \qquad i \in E$$

При этом оно необязательно единственное.

(c) Если ОДМЦ описывает случайные блуждания по неориентированному связному графу с n ребрами, то единственное стационарное распределение:

$$\pi_i = \frac{\deg(i)}{2n}, \qquad \deg(i) = \#\text{ смежных с } i \text{ вершин}$$

(d) Решение *уравнения детального баланса*

$$\pi_i p_{ij} = \pi_j p_{ji}, \qquad \forall i, j \in E$$

является стационарным распределением. Обратное неверно.

Уравнение детального баланса в некотором смысле является законом сохранения массы.

6.4 Эргодические ОДМЦ

Опр. Конечная ОДМЦ называется *эргодической*, если стационарное распределение единственно.

Для конечных ОДМЦ:

цепь эргодична $\iff$ существует единственный замкнутый класс эквивалентности $\iff$ существует единственное стационарное распределение

Для счетных цепей не удается получить похожие результаты, поэтому вводится более узкое понятие сильной эргодичности.

Обычно в задачах под эргодичностью понимают именно сильную эргодичность.

Опр. Конечная или счетная ОДМЦ называется *сильно эргодической*, если для любого $j \in E$ существует не зависящий от $i \in E$ строго положительный предел

$$\lim_{n\to\infty} p_{i,j}(n) = p_j^* > 0$$

Так как $p_{i,j}(n) = (P^n)_{ij}$, то данное условие означает, что $P^n \xrightarrow{n\to\infty} \Pi$, где матрица Π имеет одинаковые строки, и все ее элементы положительны.

Для конечных цепей из сильной эргодичности следует эргодичность. При этом $p_j^* = \pi_j > 0$, где π — единственное стационарное распределение.

Эргодическая теорема (для ОДМЦ)

> Пусть ОДЦМ $\{X_n\}_{n=1}^\infty$ сильно эргодична, а π — ее стационарное распределение. Тогда для произвольной функции, заданной на состояниях цепи
>
> $$f:\ E \to \mathbb{R}, \qquad \sum_{k\in E} \pi_k |f(k)| < +\infty$$
>
> выполнено
>
> $$\frac{1}{N}\sum_{k=1}^{N} f(X_k) \xrightarrow[n\to\infty]{\text{п.н.}} \mathbb{E}f(X) = \sum_{k\in E} \pi_k\, f(k),$$
>
> где случайная величина X имеет распределение π, то есть $\mathbb{P}\{X = k\} = \pi_k$.

В некотором смысле данная теорема похожа на усиленный закон больших чисел или теорему Биркгофа-Хинчина. Со временем цепь забывает свое начальное распределение, поэтому при больших n распределения сечений практически одинаковы.

Критерий сильной эргодичности

> $$\text{ОДМЦ сильно эргодична} \quad \Longleftrightarrow \quad \begin{cases} \text{цепь неразложима} \\ \text{цепь апериодична} \\ \text{цепь положительно возвратна} \end{cases}$$

Для конечных цепей этот критерий можно записать в виде

$$\begin{matrix}\text{конечная ОДМЦ} \\ \text{сильно эргодична}^*\end{matrix} \quad \Longleftrightarrow \quad \begin{cases} \exists \text{ единственный замкнутый класс} \\ \text{цепь апериодична} \end{cases}$$

$*$ Важно отметить, что в для конечных цепей в определении сильной эргодичности можно разрешить наличие нулевых состояний, то есть допустить, что пределы $p_j^* \geq 0$ — могут равняться нулю.

В таком случае, для конечных цепей сильная эргодичность отличается от эргодичности только требованием апериодичности.

Требование апериодичности, вообще говоря, тоже можно снять, если заменить предел в определении сильной эргодичности на предел по Чезаро. Тогда эргодичность в таком смысле будет эквивалентна единственности замкнутого класса, что, в свою очередь, эквивалентно единственности стационарного распределения. Это и есть самое первое определение эргодичности для конечных цепей.

Для счетных цепей этот критерий можно записать в виде

$$\begin{matrix}\text{счетная ОДМЦ} \\ \text{сильно эргодична}\end{matrix} \quad \Longleftrightarrow \quad \begin{cases} \text{цепь неразложима} \\ \text{цепь апериодична} \\ \exists \text{ стационарное распределение} \end{cases}$$

так как для неразложимых цепей положительная возвратность эквивалентна существованию стационарного распределения.

Понятие сильной эргодичности для счетных цепей можно ослабить, разрешив наличие нулевых состояний, т.е. допустив, что пределы $p_j^* \geq 0$ — могут равняться нулю. Тогда выполнено (только в одну сторону)

$$\begin{matrix}\text{счетная ОДМЦ} \\ \text{сильно эргодична}^* \\ (p_j^* \geq 0)\end{matrix} \quad \Longleftarrow \quad \begin{cases} \exists \text{ и ! положительно возвратный апериодичный класс} \\ \text{число невозвратных состояний конечно} \end{cases}$$

Можно ослабить еще больше, заменив пределы на пределы по Чезаро, и тогда в первом условие можно убрать апериодичность. В таком случае первое условие будет эквивалентно единственности стационарного распределения.

В задачах обычно под эргодичностью понимается именно *сильная* эргодичность. Для выполнения эргодической теоремы достаточно, чтобы конечная цепь была *эргодична*, а счетная цепь была *сильно эргодична*.

Свойства эргодических ОДМЦ:

1. Если конечная или счетная ОДМЦ сильно эргодична, то стационарное распределение единственно. При этом $p_j^* = \pi_j > 0$, где π — единственное стационарное распределение.

2. Если ОДМЦ сильно эргодична, то матрица $P^n \xrightarrow{n\to\infty} \Pi$, где матрица Π имеет одинаковые строки, совпадающие со стационарным распределением π.

3. Если ОДМЦ сильно эргодична, то для любого начального распределения $p(0)$:
$$p(n) = (P^n)^T p(0) \xrightarrow{n\to\infty} \Pi\, p(0) = \pi,$$
то есть цепь со временем забывает свое начальное распределение.

4. Если ОДМЦ эргодична (стационарное распределение π единственно), то вне зависимости от начального распределения
$$\frac{1}{N}\sum_{n=1}^{N} \mathbb{I}\{X_n = k\} \xrightarrow[n\to\infty]{\text{п.н.}} \pi_k, \qquad \frac{1}{n}\sum_{n=1}^{N} \mathbb{E}\big[\mathbb{I}\{X_n = k\}\big] \xrightarrow[n\to\infty]{} \pi_k$$
где левую часть можно интерпретировать как долю времени, проведенную в k-ом состоянии.
Эту теорему можно использовать для оценки различных параметров эргодичных цепей.

5. Если стационарное распределение единственно, то
$$\frac{1}{\mu_j} = \pi_j, \qquad \forall j \in E,$$
где $\mu_j = \mathbb{E}\sigma_j = \sum_{n=1}^{\infty} n f_j(n)$ — среднее время до первого возвращения в j-ое состояние.
Верно даже при $\mu_j = \infty$.

6. Конечная ОДМЦ сильно эргодична $\iff \exists n_0 \geq 1 : \quad \forall i,j \in E \quad p_{i,j}(n_0) > 0$.

7. У конечных сильно эргодичных ОДМЦ имеется ровно один замкнутый класс эквивалентности, и он апериодичен, поэтому собственное значение $\lambda = 1$ матрицы P имеет кратность 1, а все остальные собственные числа по модулю меньше 1.

   Пусть $\rho = \max\big\{|\lambda| \bigm| \lambda \neq 1\big\} < 1$. Тогда скорость сходимости $p(n) \to \pi$ геометрическая (экспоненциальная) с параметром ρ.

   Если собственные значения найти сложно, то параметр ρ можно оценить следующим образом. Возьмем любое такое n, что все элементы P^n отличны от нуля (это возможно, см. свойство 6). Тогда
$$\rho \geq \max_{i,j\in E} p_{i,j}(n)$$

7 Непрерывные цепи Маркова

7.1 Определение непрерывной марковской цепи

Все определения являются обобщением определений со случая дискретных марковских цепей.

Опр. Случайный процесс $\{X_t\}_{t\geq 0}$ называется *непрерывной цепью*, если его множество состояний E не более чем счетно.

Слово "непрерывная" в этом определении указывает на непрерывность времени.

Опр. Непрерывная цепь $\{X_t\}_{t\geq 0}$ называется *непрерывной марковской цепью (НМЦ)*, если она обладает *марковским свойством*:

$$\mathbb{P}\big\{X_{t_n} = i_n \bigm| X_{t_{n-1}} = i_{n-1}, \dots, X_{t_1} = i_1\big\} = \mathbb{P}\big\{X_{t_n} = i_n \bigm| X_{t_{n-1}} = i_{n-1}\big\}$$

для любых $0 \leq t_1 \leq \ldots \leq t_n$ и любых состояний $i_k \in E$.

Опр. *Вероятностью перехода* из состояния i в состояние j за время от s до t называется функция двух аргументов

$$p_{i,j}(s,t) = \mathbb{P}\big\{X_t = j \bigm| X_s = i\big\}, \qquad 0 \leq s \leq t$$

Опр. *Переходной матрицей* (между моментами времени s и t) называется матрица

$$P(s,t) = \begin{bmatrix} p_{0,0} & p_{0,1} & \cdots & p_{0,N} \\ p_{1,0} & p_{1,1} & & \vdots \\ \vdots & & \ddots & \vdots \\ p_{N,0} & \cdots & \cdots & p_{N,N} \end{bmatrix}, \qquad N = |E| - 1 \leq \infty$$

Опр. НМЦ называется *однородной непрерывной марковской цепью (ОНМЦ)*, если для любых $s, t \geq 0$:

$$P(s, s+t) = P(0,t) \equiv P(t)$$

Далее рассматриваются только однородные НМЦ, если не оговорено противного.

Опр. Вероятность $p_i(t) = \mathbb{P}\{X_t = i\}$ называется вероятностью i-го состояния в момент $t \geq 0$.

Опр. Вектор $p(t) = \big[p_0(t), p_1(t), \dots\big]^T$ называется *распределением вероятностей состояний в момент времени* $t \geq 0$.

Опр. Вероятности $p_{i,j}(t)$ называются вероятностями перейти из i-го состояния в j-ое за время t.

Свойства переходной матрицы $P(t)$ для ОНМЦ:

1. Для любого $t \geq 0$ сумма элементов в каждой строке $P(t)$ равна 1:

$$\sum_{j\in E} p_{i,j}(t) = 1, \qquad \forall t \geq 0, \qquad \forall i \in E;$$

2. $p_{i,j}(0) = \delta_{ij} = \begin{cases} 1, & i = j \\ 0, & i \neq j \end{cases};$

3. *Уравнение Колмогорова-Чепмена*: для ОНМЦ для любых $s, t \geq 0$ выполнено

$$P(s+t) = P(s)P(t)$$

Уравнение Колмогорова-Чепмена является следствием формулы полной вероятности.

4. Для ОНМЦ для любых $s, t \geq 0$ выполнено

$$p(s+t) = \big[P(t)\big]^T p(s)$$

Следствие формулы полной вероятности и марковского свойства.

Для получения содержательных результатов про ОНМЦ, нужно сделать *базовые предположения*:

(a) Если $X_t = i$, то вероятность покинуть это состояние в течение времени Δt равна $\Lambda_i \Delta t + o(\Delta t)$ при $\Delta t \to 0$. Числа Λ_i не зависят от времени t.

(b) Вероятность того, что следующим после i-го состояния будет состояние j равна $\widehat{p}_{i,j}$. Эти вероятности не зависят от того, когда цепь придет и выйдет из i-го состояния.

По определению, $\widehat{p}_{i,i} = 0$.

Поведение ОНМЦ во времени: пусть в момент t цепь находится в состоянии i. Она будет находиться в нем некоторое время $\xi \sim \mathrm{Exp}(\Lambda_i)$, потом перейдет в некоторое состояние $j \neq i$ с вероятностью $\widehat{p}_{i,j}$. Далее ситуация повторяется.

Опр. *Генератором* или *инфинитезимальной матрицей* ОНМЦ называется матрица

$$\Lambda = \left\|\Lambda_{i,j}\right\|_{i,j\in E}, \qquad \Lambda_{i,j} = \begin{cases} \Lambda_i \widehat{p}_{i,j} & , \ i \neq j \\ -\Lambda_i & , \ i = j \end{cases}$$

Опр. Числа $\Lambda_{i,j}$ $(i \neq j)$ называются *интенсивностями перехода* из состояния i в состояние j.

Теорема (уравнения Колмогорова-Феллера)

Для ОНМЦ в базовых предположениях справедливы

- *прямое уравнение Колмогорова-Феллера*:

$$\frac{dP(t)}{dt} = P(t)\Lambda, \qquad P(0) = I$$

- *обратное уравнение Колмогорова-Феллера*:

$$\frac{dP(t)}{dt} = \Lambda P(t), \qquad P(0) = I$$

- $\dfrac{dp(t)}{dt} = \Lambda^T p(t).$

При этом решением уравнений Колмогорова-Феллера является матричная экспонента

$$P(t) = e^{t\Lambda} = \sum_{k=0}^{\infty} \frac{t^k \Lambda^k}{k!}$$

Матричная экспонента сходится, если $\sup\limits_{i\in E} \sum\limits_{j\in E} |\Lambda_{i,j}| \leq C < +\infty$. В случае конечной цепи это выполнено всегда.

Задать ОНМЦ $\equiv$ задать генератор Λ и начальное распределение $p(0)$.

Свойства матрицы-генератора Λ ОНМЦ:

1. $\Lambda_{i,j} \geq 0$ при $i \neq j$.

2. Сумма по строкам матрицы Λ равна 0:

$$\sum_{j\in E} \Lambda_{i,j} = 0, \qquad \forall i \in E$$

3. Матрицы Λ и Λ^T всегда имеют собственной значение $\lambda = 0$.

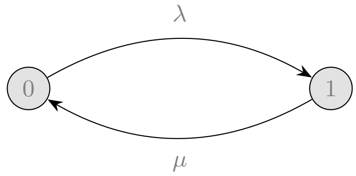
4. *Уравнения Колмогорова-Феллера*:

$$\dot{P}(t) = P(t)\Lambda = \Lambda P(t), \qquad \dot{p}(t) = \Lambda^T p(t)$$

5. Для любого $t \geq 0$ матрица $P(t) = e^{t\Lambda}$ является стохастической матрицей (сумма по строкам $= 1$).

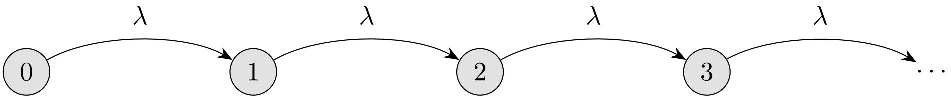
ОНМЦ удобно представлять в виде ориентированного графа, вершинами которого являются состояния, в веса ребер — интенсивности перехода между состояниями. Важно отметить, что в таком графе нет петель.

Например, граф



соответствует генератору $\Lambda = \begin{bmatrix} -\lambda & \lambda \\ \mu & -\mu \end{bmatrix}$.

Примером ОНМЦ является пуассоновский процесс с интенсивностью λ. Ему соответствует граф



7.2 Классификация и эргодичность состояний ОНМЦ

Все определения аналогичны случаю дискретных цепей.

Опр. Состояние i называется *несущественным*, если существует состояние $j \in E$, такое, что

$$\exists t > 0 : \ \ p_{i,j}(t) > 0, \qquad \forall s > 0 : \ \ p_{j,i}(s) = 0$$

Иначе состояние i называется *существенным*.

Опр. Состояния i и j называются *сообщающимися*, если

$$\exists t > 0 : \ \ p_{i,j}(t) > 0, \qquad \exists s > 0 : \ \ p_{j,i}(s) > 0$$

Иначе состояния i и j называются *несообщающимися*.

Определения неразложимой цепи, открытых и замкнутых классов эквивалентности такие же, как и в дискретном случае. Для непрерывных цепей понятие периода состояния не вводится.

Опр. Состояние i называется *невозвратным*, если выполнено одно из эквивалентных свойств:

(a) Цепь приходит в i-ое состояние бесконечно часто с нулевой вероятностью:

$$\mathbb{P}\big\{\text{множество } \{t \mid X_t = i\} \text{ неограничено}\big\} = 0$$

Событие под знаком $\mathbb{P}$ принадлежит остаточной сигма-алгебре, поэтому его вероятность может быть равна только 0 или 1. В случае дискретной цепи было аналогично.

(b) *Критерий возвратности*: $\displaystyle\int\limits_0^{+\infty} p_{i,i}(t)\,dt = +\infty.$

Опр. Распределение π называется *стационарным* или *инвариантным*, если

$$\big[P(t)\big]^T \pi = \pi, \qquad \forall t \geq 0$$

Данное требование эквивалентно $\dot{\pi}(t) = 0 \ \ \Leftrightarrow \ \ \Lambda^T \pi = 0$. Последнее уравнение называется *алгебраическим уравнением Колмогорова.*

Стационарное распределение π — собственный вектор Λ^T для $\lambda = 0$. Такой собственный вектор есть всегда, но в случае счетной цепи он может не удовлетворять условию нормировки $\sum_{k \in E} \pi_k = 1$.

Опр. ОНМЦ называется *сильно эргодической*, если для любого $j \in E$ существует не зависящий от $i \in E$ строго положительный предел
$$\lim_{t\to\infty} p_{i,j}(t) = p_j^* > 0$$

В ОНМЦ из-за отсутствия явления периодичности всегда существуют пределы $\lim\limits_{t\to\infty} p_{i,j}(t)$.

Если ОНМЦ сильно эргодична, то стационарное распределение π единственно, причем $p^* = \pi$.

Критерий сильной эргодичности (для конечных ОНМЦ):

> Следующие утверждения эквивалентны:
>
> (a) конечная ОНМЦ сильно эргодична;
>
> (b) существует единственное стационарное распределение, компоненты которого положительны.
>
> (c) конечная ОНМЦ неразложима;

Для счетных цепей надо исследовать на эродичность *цепь скачков* (см. следующий раздел).

Свойства эргодических ОНМЦ:

1. Если ОНМЦ сильно эргодична, то стационарное распределение π единственно, причем $p^* = \pi$.
2. Если ОНМЦ сильно эргодична, то матрица $P(t) \xrightarrow{t\to\infty} \Pi$, где матрица Π имеет одинаковые строки, совпадающие со стационарным распределением π.
3. Если ОНМЦ сильно эргодична, то для любого начального распределения $p(0)$:
$$p(t) = \big[P(t)\big]^T p(0) \xrightarrow{t\to\infty} \Pi\, p(0) = \pi,$$
то есть цепь со временем забывает свое начальное распределение.
4. Если ОНМЦ сильно эргодична, то вне зависимости от начального распределения
$$\frac{1}{T}\int\limits_0^T \mathbb{I}\big\{X_t = k\big\}\, dt \xrightarrow[T\to\infty]{\text{п.н.}} \pi_k, \qquad \forall k \in E,$$
где интеграл понимается как интеграл по траекториям.
5. Если стационарное распределение единственно, то
$$\frac{1}{\mu_j} = \Lambda_j \pi_j, \qquad \forall j \in E,$$
где $\mu_j = \mathbb{E}\sigma_j$ — среднее время до первого возвращения в j-ое состояние.
Верно даже при $\mu_j = \infty$.
6. Конечная ОНМЦ сильно эргодична $\iff$ $\exists\, t_0 > 0 : \quad \forall i, j \in E \quad p_{i,j}(t_0) > 0$.

7.3 Цепь скачков

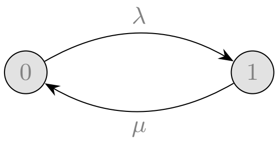
На практике проводить классификацию состояний и проверять эргодичность удобно с использованием *цепи скачков*.

Опр. ОДМЦ $\{Y_n\}_{n=0}^{\infty}$ называется *цепью скачков* для ОНМЦ $\{X_t\}_{t\geq 0}$, если если ее матрица переходов $\widehat{P}$:
$$\widehat{P} = \big\|\widehat{p}_{i,j}\big\|_{i,j\in E}, \qquad \widehat{p}_{i,j} = \begin{cases} -\dfrac{\Lambda_{i,j}}{\Lambda_{i,i}} & , \ i \neq j \\ \ 0 & , \ i = j \end{cases},$$

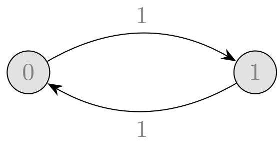
где Λ — генератор ОНМЦ $\{X_t\}_{t\geq 0}$.

Вероятности $\widehat{p}_{i,j}$ — те же самые, что и в базовых предположениях. Цепь скачков не имеет петель.

Например, цепи с генератором $\Lambda = \begin{bmatrix} -\lambda & \lambda \\ \mu & -\mu \end{bmatrix}$ соответствует следующая цепь скачков:



ОНМЦ $\{X_t\}_{t\geq 0}$



цепь скачков $\{Y_n\}_{n=0}^{\infty}$

Свойства цепи скачков:

1. Классификация состояний в ОНМЦ $\{X_t\}_{t\geq 0}$ эквивалентна классификации состояний в цепи скачков $\{Y_n\}_{n=0}^{\infty}$.
2. ОНМЦ $\{X_t\}_{t\geq 0}$ сильно эргодична $\iff$ цепь скачков $\{Y_n\}_{n=0}^{\infty}$ сильно эргодична.
3. ОНМЦ $\{X_t\}_{t\geq 0}$ имеет стационарное распределение π $\iff$ цепь скачков $\{Y_n\}_{n=0}^{\infty}$ имеет стационарное распределение $\widehat{\pi}$, причем
$$\widehat{\pi}_k = \frac{\Lambda_k \pi_k}{\sum\limits_{i\in E} \Lambda_i \pi_i}, \qquad k \in E$$