Алтайский Государственный Университет
Математический факультет
Кафедра математического анализа

Конспект лекций
по теории случайных процессов

Барнаул, 2004

1. Введение

По мере нашего продвижения по курсу теории вероятностей перед нами возникают все более и более сложные случайные объекты. Сначала это случайные события, которые можно отождествить с их индикаторами, принимающими значения 0 и 1. Затем (действительнозначные) случайные величины, которые могут своими значениями охватывать всю действительную прямую. Следом появляются комплекснозначные случайные элементы, конечномерные случайные векторы. Наконец, при изучении основных конструкций математической статистики, мы доходим до рассмотрения случайных элементов, принимающих значения в бесконечномерном пространстве (выборки бесконечного объема, как случайные векторы). Естественно, что все эти объекты возникают в результате каких-то случайных экспериментов. Нетрудно представить себе такой эксперимент, результатом которого будет определенная кривая линия или собственно функция. Так мы можем получить случайный элемент, принимающий значения в некотором функциональном пространстве, или случайную функцию. В этом случае говорят о том, что перед нами случайный процесс. Встречаются в литературе также названия вероятностный процесс, стохастический процесс, а иногда и просто процесс, если заранее ясно, о чем речь. Дадим теперь определение.

Пусть T – некоторое множество. $A(T)$ – пространство действительнозначных функций, определенных на T. Рассмотрим также вероятностное пространство $< \Omega, \mathcal{F}, \mathbf{P} >$. Отображение $\xi : T \times \Omega \to R$ назовем *случайной функцией* если $\forall t \in T \quad \xi(t) = \xi(t, \cdot)$ – случайная величина. Точка вместо второго аргумента означает здесь и далее, что мы рассматриваем $\xi(t)$ как функцию $\omega \in \Omega$ в этом контексте. Ясно также, что $\xi(\cdot, \omega) \in A(T)$. Если $T \subset R$ и параметр t интерпретируется как время, то случайную функцию называют *случайным процессом*. При записи формул, связанных со случайным процессом, как и при записи случайных величин в теории вероятностей, случайный аргумент ω обычно опускается. Если T представляет собой класс целых чисел $\mathbf{Z}$ или натуральных чисел $\mathbf{N}$, то говорят о *случайной последовательности*. Отметим, что последовательность случайных величин – объект для нас относительно знакомый, поэтому мы будем часто привлекать его в качестве примера.

Если мы фиксируем $\omega \in \Omega$, то полученная неслучайная функция $\xi(\omega, \cdot)$ называется *реализацией* случайного процесса. Наряду с этим термином употребляются также названия *траектория, выборочная функция.* Функция

$$K(t,s) \;=\; \mathbf{cov}(\xi(t),\, \xi(s)) \;=\; \mathbf{M}\xi(t)\xi(s) \;-\; \mathbf{M}\xi(t)\,\mathbf{M}\xi(s)$$

называется *ковариационной функцией* случайного процесса ξ.

Лемма 1. *Ковариационная функция любого случайного процесса обладает свойством неотрицательной определенности:*

$$(\forall k)(\forall\ c_1, ..., c_k \in R)\ (\forall t_1, ..., t_k \in T)\ \ \sum_{i=1}^{k}\sum_{j=1}^{k} c_i c_j K(t_i, t_j)\ \geq\ 0. \qquad (1)$$

Доказательство. Пусть $m(t) = \mathbf{M}\xi(t)$. Тогда

$$\sum_{i=1}^k \sum_{j=1}^k c_i c_j K(t_i, t_j)\ =\ \mathbf{M}\sum_{i,j} c_i c_j (\xi(t_i) - m(t_i))(\xi(t_j) - m(t_j))\ =$$
$$= \mathbf{M}\left(\sum_{j=1}^k c_j(\xi(t_j) - m(t_j))\right)^2\ =\ \mathbf{D}\left(\sum_{j=1}^k c_j \xi(t_j)\right)\ \geq\ 0.$$

Условимся через $\bar{t} = \bar{t}(n)$ обозначать конечные подмножества T, например, $\bar{t} = \{t_1, t_2, ..., t_n\} \subset T$. Если подмножество $\bar{t}(n)$ зафиксировано, то через $R^{\bar{t}}$ условимся обозначать n-мерное евклидово пространство. Договоримся предполагать $R^{\bar{s}} \subset R^{\bar{t}}$ при $\bar{s} \subset \bar{t}$. Будем также считать все такие пространства подпространствами $R^T\ =\ \{f \mid f : T \to R\}$. Такое включение станет сразу понятно, если считать, что

$$R^{\bar{t}}\ =\ \{(f(t_1), ..., f(t_n)) \mid f \in R^T\}.$$

Введем также обозначение $\xi_{\bar{t}}\ =\ (\xi(t_1), ..., \xi(t_n))$ для заданного $\bar{t}$ и произвольного отображения $\xi : T \to R$.

Пусть $\xi(t)$ – случайный процесс. Распределения векторов $\xi_{\bar{t}}$, когда $\bar{t}$ пробегает все конечные подмножества множества T, называют *конечномерными распределениями* случайного процесса ξ. Как вы помните из курса теории вероятностей, эти распределения задаются формулами

$$\left(\forall B \in \mathcal{B}(R^{\bar{t}})\right)\ \mathbf{P}_{\bar{t}}(B)\ =\ \mathbf{P}(\xi_{\bar{t}} \in B).$$

Пусть $\bar{t} \supset \bar{s}$, $\pi_{\bar{s},\bar{t}}$ - естественная проекция $R^{\bar{s}}$ на $R^{\bar{t}}$. Тогда во введенных обозначениях справедливо

$$(\forall A \in \mathcal{B}(R^{\bar{s}}))\ \mathbf{P}_{\bar{s}}(A)\ =\ \mathbf{P}_{\bar{t}}(\pi^{-1}_{\bar{t},\bar{s}}(A)). \qquad (2)$$

Это условие называется условием согласованности семейства распределений $\mathbf{P}_{\bar{t}}$, когда $\bar{t}$ пробегает класс всех конечных подмножеств T.

Теорема 1 (А.Н.Колмогоров). *Пусть семейство распределений удовлетворяет условию согласованности (2). Тогда найдется некоторое вероятностное пространство* $< \Omega, \mathcal{F}, \mathbf{P} >$ *и случайный процесс* $\xi(t)$ *на нем, такой, что* $\mathbf{P}_{\bar{t}}$ *есть распределение* $\xi_{\bar{t}}$ *для произвольного конечного* $\bar{t} \subset T$.

Другими словами, если семейство распределений удовлетворяет условию согласованности, то оно является семейством конечномерных распределений некоторого случайного процесса.

Изложим здесь идею доказательства. Полное доказательство можно найти в книге [1, приложение 2].

Положим $\Omega = R^T$, $\xi(t,\omega) = \omega(t)$. Множество $B \subset R^T$ назовем цилиндрическим, если оно имеет вид $\{\omega|(\omega(t_1), \ldots, \omega(t_n)) \in A\}$ для некоторого набора $t_1, \ldots, t_n$ точек T и какого-нибудь $A \in \mathcal{B}(R^n)$. Пусть $\mathcal{C}$ - класс всех цилиндрических подмножеств R^T, $\mathcal{F} = \sigma(\mathcal{C})$. Для $B \in \mathcal{C}$ зададим

$$\mathbf{P}(B) \;=\; \mathbf{P}_{\bar{t}}(A).$$

Для завершения доказательства осталось проверить условия теоремы Каратеодори о продолжении вероятностной меры. Тогда, продолжая построенную вероятность на σ-алгебру борелевских подмножеств R^T, мы получим, что на вероятностном пространстве $< \Omega, \mathcal{F}, \mathbf{P} >$ процесс $\xi(t)$ обладает нужным нам семейством конечномерных распределений.

Иногда полезным является вариант условия согласованности (2) в терминах характеристических функций. Пусть, как и раньше, $\bar{t}$ - конечное подмножество T. Определим

$$< \lambda, x >_{\bar{t}} = \sum_{\{k: t_k \in \bar{t}\}} \lambda(t_k) x(t_k), \quad \lambda, x \in R^T.$$

Тогда характеристическая функция распределения $\mathbf{P}_{\bar{t}}$ – это

$$\varphi_{\bar{t}}(\lambda) \;=\; \int_{R^T} \exp\{i < \lambda, x >_{\bar{t}}\} d\mathbf{P}_{\bar{t}}.$$

Условие согласованности распределений во введенных обозначениях эквивалентно при этом

$$(\forall \lambda \in R^T)\,(\bar{s} \subset \bar{t}) \;\Rightarrow\; (\varphi_{\bar{s}}(\lambda) = \varphi_{\bar{t}}(\pi_{\bar{t},\bar{s}}(\lambda)))\,. \tag{3}$$

Здесь $(\pi_{\bar{t},\bar{s}}(\lambda))(u) = \lambda(u)$ при $u \in \bar{s}$ и равно нулю при $u \in \bar{t} \setminus \bar{s}$.

2. Гауссовские случайные процессы

Случайный процесс $\xi(t)$ называется *гауссовским*, если все его конечномерные распределения нормальны. Таким образом, характеристическая функция любого конечномерного распределения гауссовского процесса имеет вид

$$\varphi(\vec{\lambda}) \;=\; \exp\{i < \vec{\lambda}, \vec{a} > -\frac{1}{2} < B\vec{\lambda}, \vec{\lambda} >\},$$

где $\vec{a}$ - вектор математических ожиданий, а B - ковариационная матрица координат соответствующего вектора. Как известно из курса теории вероятностей (и следует из приведенной выше формулы), для задания нормального распределения в конечномерном случае достаточно задать $\vec{a}$ и матрицу ковариаций B. Оказывается, для гауссовских случайных процессов (соответствующих нормальным распределениям в функциональных пространствах) имеет место аналогичный результат.

Теорема 2. *Для произвольно заданной функции $a(t)$ и любой функции $K(t,s)$, удовлетворяющей условию неотрицательной определенности (1), существует гауссовский случайный процесс с математическим ожиданием $a(t)$ и ковариационной функцией $K(t,s)$.*

Доказательство. Заметим, что достаточно построить процесс с тождественно нулевым математическим ожиданием, а затем прибавить к нему неслучайную функцию $a(t)$. Зафиксируем $\bar{t} = \{t_1, ..., t_n\}$, и пусть $\mathbf{P}_{\bar{t}}$ - нормальное распределение в R^n с нулевым средним и ковариационной матрицей $B = B_{\bar{t}}$, элементы которой вычислены по правилу $B_{i,j} = K(t_i, t_j)$. Получившаяся матрица будет неотрицательна определена, что следует из леммы 1. Таким образом, описанное построение возможно. Чтобы завершить доказательство, достаточно заметить, что

$$\varphi_{\bar{t}}(\vec{\lambda}) = \exp\{-\frac{1}{2} < B\vec{\lambda}, \vec{\lambda} >\} -$$

характеристическая функция построенного распределения и проверить условия согласованности получившихся распределений (например, в форме (3)).

Действительно, поскольку $\pi_{\bar{t}\bar{s}}(\lambda)(t_j) = 0$ при $j \notin \bar{s}$, то

$$\begin{aligned}\varphi_{\bar{t}}(\pi_{\bar{t}\bar{s}}(\lambda)) = \exp\left\{-\tfrac{1}{2}\textstyle\sum_{i,j\in\bar{t}} B_{i,j}\pi_{\bar{t}\bar{s}}(\lambda)(t_i)\pi_{\bar{t}\bar{s}}(\lambda)(t_j)\right\} = \\ = \exp\left\{-\tfrac{1}{2}\textstyle\sum_{i,j\in\bar{s}} B_{i,j}\lambda(t_i)\lambda(t_j)\right\} = \varphi_{\bar{s}}(\lambda),\end{aligned}$$

и теорема доказана.

Очень важным примером для нас является *винеровский процесс*, который служит моделью броуновского движения. Винеровским процессом $w(t,\omega)$ мы будем называть гауссовский процесс с независимыми приращениями такой, что $w(0) = 0$, $w(t) - w(s)$ имеет нормальное распределение $N(0, t-s)$ при $t > s$

Выпишем конечномерные распределения для винеровского процесса. Для этого зафиксируем $t_1, ..., t_n$ так, что $0 < t_1 < ... < t_n$ и рассмотрим вектор $\overrightarrow{w} = (w_1, ...w_n)$, где $w_j = w(t_j)$, $j = 1, 2...n$. Заметим, что плотность распределения вектора $X = (w_1, w_2 - w_1, ..., w_n - w_{n-1})$ легко вычислить, т.к. его координаты – независимые нормально распределенные случайные величины:

$$\mathbf{P}_X(x_1, ..., x_n) = \prod_{i=1}^{n} \left([2\pi(t_i - t_{i-1})]^{-1/2} \exp\left\{ -\frac{x_i^2}{2(t_i - t_{i-1})} \right\} \right),$$

где $t_0 = 0$. При этом $\overrightarrow{w} = AX$, где

$$A = \begin{pmatrix} 1 & 0 & 0 & ... & 0 \\ 1 & 1 & 0 & ... & 0 \\ ... & ... & ... & ... & ... \\ 1 & 1 & 1 & ... & 1 \end{pmatrix}.$$

Как известно, в этом случае

$$\mathbf{P}_{\overrightarrow{w}}(X) = |\det A|^{-1} \mathbf{P}_X(A^{-1}X).$$

Легко проверить, что $\det A = 1$,

$$A^{-1} = \begin{pmatrix} 1 & 0 & 0 & ... & 0 & 0 \\ -1 & 1 & 0 & ... & 0 & 0 \\ 0 & -1 & 1 & ... & 0 & 0 \\ ... & ... & ... & ... & ... & ... \\ 0 & 0 & 0 & ... & -1 & 1 \end{pmatrix},$$

откуда $A^{-1}X = (x_1, x_2 - x_1, ..., x_n - x_{n-1})$, и

$$\mathbf{P}_{\overrightarrow{w}}(x_1, ..., x_n) = \prod_{i=1}^{n} ([2\pi(t_i - t_{i-1})]^{-1/2} \exp\{-(x_i - x_{i-1})^2 / 2(t_i - t_{i-1})\}),$$

где $x_0 = 0$. Итак, конечномерные распределения винеровского процесса являются нормальными (гауссовскими).

Винеровский процесс (броуновское движение) является гауссовским процессом, выходящим из 0. В приложениях иногда встречается так называемый *броуновский мост* - процесс, задаваемый равенством

$$w^0(t) \ = \ w(t) - tw(1), \quad t \in [0, 1].$$

Очевидно, что $w^0(0) = w^0(1) = 0$ с вероятностью 1. Вычислим ковариационную функцию броуновского моста. Пусть $t \geq s$.

$$\begin{aligned} K(t, s) \ &= \ \mathbf{M}(w(t) - tw(1))(w(s) - sw(1)) \ = \ \mathbf{M}w(t)w(s) - \\ &- \ t\mathbf{M}w(1)w(s) - s\mathbf{M}w(1)w(t) + st\mathbf{M}w^2(1). \end{aligned}$$

Заметим, что для произвольного a $\mathbf{M}w^2(a) = a$, а при $h > u$

$$\mathbf{M}w(h)w(u) \ = \ \mathbf{M}(w(h) - w(u))w(u) \ + \ \mathbf{M}w^2(u) \ = \ u,$$

откуда

$$K(t, s) \ = \ s - ts - st + st \ = \ s(1 - t), \ t \geq s.$$

3. Процессы с независимыми приращениями

Пусть $T = [a, b]$. Процесс $\xi(t)$, $t \in T$ называется *процессом с независимыми приращениями*, если для любого k и любых $t_1 < ... < t_k \in T$) приращения $\xi(t_1) - \xi(a)$, $\xi(t_2) - \xi(t_1), ...$, $\xi(b) - \xi(t_k)$ являются независимыми случайными величинами.

Обозначим через $\mathbf{P}^a$ распределение $\xi(a)$, через $\mathbf{P}_{s,t}$ распределение $\xi(t) - \xi(s)$, $t > s$. Зная указанные распределения, мы можем восстановить все конечномерные распределения процесса:

$$\xi(t_j) \ = \ \sum_{i=1}^{j}(\xi(t_i) \ - \ \xi(t_{i-1})), \quad (t_0 = a),$$

а значит, распределение $\xi_{\bar{t}}$ - совместное распределение случайных величин $\xi_0 + \xi_1$, $\xi_0 + \xi_1 + \xi_2, ..., \xi_0 + ... + \xi_n$, где $\xi_0 = \xi(a)$, $\xi_i = \xi(t_i) - \xi(t_{i-1})$ при $i \geq 1$. Но оказывается, при некоторых дополнительных предположениях для задания конечномерных распределений процесса с независимыми приращениями достаточно знать гораздо меньшее количество распределений.

Говорят, что $\xi(t)$ – *стационарный процесс*, если для произвольных s, t, h распределения $\xi(t) - \xi(s)$ и $\xi(t + h) - \xi(s + h)$ совпадают. Процесс $\xi(t)$ называется *стохастически непрерывным*, если при $s \to t$ справедливо $\xi(s) \xrightarrow{\mathbf{P}} \xi(t)$.

Теорема 3. *Пусть случайный процесс $\xi(t)$ стохастически непрерывен, стационарен и имеет независимые приращения. Тогда для задания его конечномерных распределений достаточно, кроме $\mathbf{P}^a$, задать только одно распределение. Этим распределением может быть любое из распределений $\mathbf{P}_{t,s}$.*

Доказательство. Заметим, что в силу стационарности можно считать, что $a = 0$. Пусть $\varphi_t(\lambda)$ – характеристическая функция распределения $\mathbf{P}_{t,0}$. В силу представления

$$\xi(t) - \xi(0) \ = \ \sum_{i=1}^{n}\left(\xi\left(\frac{it}{n}\right) - \xi\left(\frac{(i-1)t}{n}\right)\right)$$

и стационарности случайного процесса имеет место соотношение

$$(\forall t, n, \lambda) \ \varphi_t(\lambda) \ = \ \varphi^n_{\frac{t}{n}}(\lambda). \qquad (4)$$

Покажем далее, что ни при одном λ $\varphi_t(\lambda)$ не обращается в 0. Действительно, при $n \to \infty$ справедливо

$$\xi\left(\frac{t}{n}\right) - \xi(0) \xrightarrow{\mathbf{P}} 0,$$

откуда, по основной теореме о характеристических функциях, $\varphi_{t/n}(\lambda)$ при $n \to \infty$ стремится к тождественной единице, характеристической функции нулевой константы. Теперь требуемое следует из (4).

Таким образом, существует функция $\Psi_t(\lambda)$ с условием

$$\varphi_t(\lambda) = \exp\left\{\Psi_t(\lambda)\right\}.$$

Следующей нашей целью является доказать, что

$$(\forall t, \lambda)\ \Psi_t(\lambda) = t\Psi_1(\lambda) \equiv t\Psi(\lambda). \tag{5}$$

Пусть сначала $t = \frac{1}{n}$. Выберем в (4) $t = 1$, откуда

$$\varphi_1(\lambda) = \varphi^n_{1/n}(\lambda) \implies \exp\Psi(\lambda) = \exp\{n\Psi_{1/n}(\lambda)\},$$

что означает справедливость (5) для $t = \frac{1}{n}$. Заметим далее, что

$$\xi\left(\frac{m}{n}\right) - \xi(0) = \sum_{j=1}^{m}\left(\xi\left(\frac{j}{n}\right) - \xi\left(\frac{j-1}{n}\right)\right),$$

а значит, $\varphi_{m/n}(\lambda) = \varphi^m_{1/n}(\lambda)$, что влечет справедливость равенства

$$\Psi_{\frac{m}{n}}(\lambda) = m\Psi_{\frac{1}{n}}(\lambda) = \frac{m}{n}\Psi(\lambda).$$

Если, наконец, $t > 0$ – произвольное действительное число, рассмотрим последовательность x_n рациональных чисел, сходящуюся к t. В силу стохастической непрерывности

$$\xi(x_n) - \xi(0) \xrightarrow{\mathbf{P}} \xi(t) - \xi(0),$$

откуда $\varphi_{x_n}(\lambda) \to \varphi_t(\lambda)$, и $x_n\Psi(\lambda) \to \Psi_t(\lambda)$. С другой стороны, $x_n\Psi(\lambda) \to t\Psi(\lambda)$, и доказываемое соотношение (5) немедленно следует из единственности предела.

Итак, пусть мы знаем $\mathbf{P}_{0,1}$. Тогда мы знаем его характеристическую функцию, а значит и $\Psi(\lambda)$. По формуле (5) и t восстанавливаем Ψ_t, затем φ_t, а значит, и произвольное $\mathbf{P}_{0,t}$. Теорема доказана.

В заключение приведем формулу для характеристической функции распределения вектора $\xi_{\bar{t}} = (\xi(t_1), ..., \xi(t_n))$, вывод которой оставляем читателю в качестве упражнения, полезного для восстановления навыков в использовании свойств характеристических функций.

$$\varphi_{\bar t}(\vec\lambda) \;=\; \varphi_a(\Lambda_n)\cdot\exp\left\{\sum_{j=1}^{n}(t_j-t_{j-1})\Psi(\Lambda_n-\Lambda_{j-1})\right\},$$

где

$$\Lambda_k \;=\; \sum_{j=1}^{k}\lambda_j,\quad \Lambda_0=0,$$

φ_a – характеристическая функция $\mathbf{P}^a$.

Одним из часто применяемых обобщений процессов с независимыми приращениями является понятие марковского процесса. Здесь мы дадим только определение, за формулировками и доказательствами свойств марковских процессов и их частного случая – дискретных однородных цепей Маркова отсылаем читателя к книге [1].

Пусть $\xi(t)$ – случайный процесс, обозначим

$$\mathcal{F}_{\le t}=\sigma(\{\xi(s),\ s\le t\}),\ \mathcal{F}_{=t}=\sigma(\xi(t)),\ \mathcal{F}_{\ge t}=\sigma(\{\xi(s),\ s\ge t\}).$$

Процесс $\xi(t)$ называют *марковским*, если

$$(\forall t)(\forall B\in\mathcal{F}_{\ge t})\ \mathbf{P}(B/\mathcal{F}_{\le t}) \;=\; \mathbf{P}(B/\mathcal{F}_{=t}).$$

4. Линейная теория случайных процессов

4.1. Гильбертово пространство случайных величин с комплексными значениями

Пусть $<\Omega,\mathcal{F},\mathbf{P}>$ – вероятностное пространство, $\xi:\Omega\to\mathbf{C}$ – комплекснозначный случайный элемент, т.е. $\Re e\xi$, $\Im m\xi$ – случайные величины. Введем

$$\mathbf{L}_2 \;=\; \mathbf{L}_2(\Omega,\mathcal{F},\mathbf{P}) \;=\; \{\xi\mid \mathbf{M}|\xi|^2<\infty\}.$$

На этом пространстве определим скалярное произведение и норму:

$$<\xi,\eta>=\ \mathbf{M}\xi\bar\eta,\quad \|\xi\|^2_{\mathbf{L_2}}=<\xi,\xi>=\ \mathbf{M}|\xi|^2.$$

Обозначим $m(t)=\mathbf{M}\xi(t)$, $\xi_*(t) \;=\; \xi(t)-m(t)$. Очевидно, что ковариационная функция процесса должна быть в этой ситуации определена следующим образом

$$K(t,s) \;\equiv\; \mathbf{M}\xi_*(t)\bar\xi_*(s),$$

а свойство неотрицательной определенности (1) примет вид

$$(\forall k)(\forall\ c_1,...,c_k \in \mathbf{C})\ (\forall t_1,...,t_k \in T)\ \ \sum_{i=1}^{k}\sum_{j=1}^{k} c_i\bar{c}_j K(t_i,t_j)\ \geq\ 0. \qquad (6)$$

Заметим также, что из определения следует

$$(\forall s,t)\ K(t,s)\ =\ \overline{K(s,t)}.$$

Полнота построенного пространства во введенной норме проверяется теми же методами, что доказывается полнота пространства интегрируемых с квадратом функций в функциональном анализе. Нам понадобится следующий довольно простой признак существования предела:

Лемма 2. *В пространстве* $\mathbf{L}_2$ *существует предел процесса* $\xi(t)$ *при* $t \to t_0$ *тогда и только тогда, когда* $(\exists m)\ m(t) \to m$ *при* $t \to t_0$, *и*

$$(\exists K < \infty)\ \lim_{s,t\to t_0} K(t,s)\ =\ K.$$

Доказательство. Необходимость. Пусть $\xi(t) \xrightarrow{\mathbf{L}_2} \xi$. Тогда, выбрав $m = \mathbf{M}\xi,\ K = \mathbf{M}\xi_*\bar{\xi}_*$, где $\xi_* = \xi - m$, получим

$$|m(t)-m|\ \leq\ \mathbf{M}|\xi(t)-\xi|\ \leq\ \sqrt{\mathbf{M}|\xi(t)-\xi|^2}\ \to\ 0,$$

а также

$$\begin{aligned}|K(t,s)-K|\ &\leq |\mathbf{M}\xi_*(t)\bar{\xi}_*(s)-\mathbf{M}\xi_*(t)\bar{\xi}_*|\ +\ |\mathbf{M}\xi_*(t)\bar{\xi}_*-\mathbf{M}\xi_*\bar{\xi}_*| \leq \\ &\leq \sqrt{\mathbf{M}|\xi_*(t)|^2\mathbf{M}|\xi_*(s)-\xi_*|^2}\ +\sqrt{\mathbf{M}|\xi_*|^2\mathbf{M}|\xi_*(t)-\xi_*|^2}\ \leq \\ &+c(\|\xi_*(s)-\xi_*\|\ +\ \|\xi_*(t)-\xi_*\|),\end{aligned}$$

что стремится к 0 при $s,t \to t_0$.

Достаточность. Заметим, что

$$\|\xi(t)-\xi(s)\|^2\ =\ |m(t)-m(s)|^2+\|\xi_*(t)-\xi_*(s)\|^2.$$

Но

$$\|\xi_*(t)-\xi_*(s)\|^2_{\mathbf{L}_2}\ =\ K(t,t)+K(s,s)-2\Re e K(t,s),$$

откуда при выполнении условий леммы следует, что

$$\|\xi(t)-\xi(s)\|^2_{\mathbf{L}_2}\ \longrightarrow 0$$

при $(t-s) \to 0$, что означает существование нужного предела в силу полноты рассматриваемого пространства. Лемма доказана.

Отметим, что в случае сходимости $\xi(t)$ к ξ в смысле пространства $\mathbf{L}_2$ имеет место также сходимость по вероятности и по распределению. Это немедленно следует из результатов курса теории вероятностей и неравенства П.Л.Чебышёва:

$$\mathbf{P}(|\xi(t)-\xi| \geq \varepsilon) \ \leq \ \frac{\mathbf{M}|\xi(t)-\xi|}{\varepsilon} \leq \ \frac{\|\xi(t)-\xi\|_{\mathbf{L}_2}}{\varepsilon}$$

для произвольного $\varepsilon > 0$.

4.2. Дифференцирование и интегрирование в среднем квадратическом

В дальнейшем, если $\xi(t) \xrightarrow{\mathbf{L}_2} \xi$, при $t \to a$, условимся писать

$$\mathrm{l.i.m.}_{t\to a}\xi(t) \ = \ \xi$$

и называвать этот предел *пределом в среднем квадратическом.* Положим при $h, a \in R$

$$\Delta_h\xi(a) \ = \ \frac{\xi(a+h)-\xi(a)}{h}.$$

Если

$$(\exists z \in \mathbf{L}_2) \ z = \mathrm{l.i.m.}_{h\to 0}\Delta_h\xi(a),$$

то $\xi(t)$ называют *дифференцируемым в среднем квадратическом* в точке a, а z – *производной процесса* в этой точке. Обозначение – $z = \frac{d\xi}{dt}(a)$. Из леммы 2 следует, что справедливы следующие необходимые и достаточные условия дифференцируемости в среднем квадратическом:

Лемма 3. *Процесс $\xi(t)$ будет дифференцируемым в среднем квадратическом в точке a тогда и только тогда, когда выполнены два условия:*

- *в этой точке существует производная $m(t)$ в обычном смысле;*
- *ковариационная функция процесса имеет в этой точке обобщенную вторую производную, а именно*

$$D_sD_tK(t,s)|_{t=s=a} \ = \ \lim_{h,q\to 0}\frac{K(a+h,a+q)-K(a,a+q)-K(a+h,a)+K(a,a)}{hq} < \infty.$$

Теорема 4. *Пусть для произвольного $a \in [A,B]$ определена обобщенная вторая производная $D_sD_tK(t,s)|_{t=s=a}$, тогда случайный процесс $\xi_*(t)$ дифференцируем в среднем квадратическом на $[A,B]$, существуют обычные частные производные $\frac{\partial K}{\partial t}$, $\frac{\partial K}{\partial s}$ и $\frac{\partial^2 K}{\partial t\partial s}$, причем для характеристик производной процесса имеют место соотношения*

$$\mathbf{M}\xi_*(t)\overline{\frac{d\xi_*}{dt}(s)} \ = \ \frac{\partial K}{\partial s}(t,s), \quad \mathbf{M}\frac{d\xi_*}{dt}(t)\overline{\frac{d\xi_*}{dt}(s)} \ = \ \frac{\partial^2 K}{\partial t\partial s}(t,s). \qquad (7)$$

Из условия видно, что существование обобщенной второй производной означает, вообще говоря, нечто большее, чем существование второй смешанной частной производной. Рассуждая так же, как в курсе математического анализа доказывалась теорема о независимости вида производной от порядка дифференцирования, выводим

Следствие. *Если ковариационная функция $K(t,s)$ дважды непрерывно дифференцируема, то процесс имеет производную и выполнены* (7).

Доказательство теоремы. Прежде всего заметим, что если $\mathbf{M}\xi^2 < \infty$ и $\exists \text{l.i.m.}_{h\to a}\eta(h) \ = \ \eta$, то существует и конечный предел

$$\lim_{h\to a} \mathbf{M}\xi\overline{\eta(h)} \ = \ \mathbf{M}\xi\overline{\eta}.$$

Действительно,

$$|\mathbf{M}\xi\overline{\eta(h)} - \mathbf{M}\xi\overline{\eta}|^2 \ \le \ \mathbf{M}\xi^2 \cdot \|\eta(h)-\eta\|^2_{\mathbf{L_2}} \ \to 0.$$

Поскольку $\mathbf{M}\xi_* = 0$, то первое условие леммы 3 проверять не нужно, а значит определен предел $\Delta_h\xi_*(s)$ при $h\to 0$. Таким образом, определен и

$$\lim_{h\to 0} \mathbf{M}\xi_*(s)\overline{\Delta_h\xi_*(t)} \ = \ \lim_{h\to 0}\frac{K(s,\, t+h)-K(s,\, t)}{h},$$

который равен $\frac{\partial K}{\partial t}$. Но, с другой стороны, вновь привлекая только что сделанное замечание, он же равен и $\mathbf{M}\xi_*(t)\overline{\frac{d\xi_*}{dt}(s)}$. Существование другой частной производной первого порядка очевидно из соображений симметрии. Осталось лишь заметить, что

$$\begin{gathered}\mathbf{M}\tfrac{d\xi_*}{dt}(s)\overline{\tfrac{d\xi_*}{dt}(t)} = \lim\nolimits_{h\to 0}\mathbf{M}\Delta_h\xi_*(s)\overline{\tfrac{d\xi_*}{dt}(t)} = \\ = \lim\nolimits_{h\to 0}\tfrac{1}{h}\left(\mathbf{M}\xi_*(s+h)\overline{\tfrac{d\xi_*}{dt}(t)} \ - \ \mathbf{M}\xi_*(s)\overline{\tfrac{d\xi_*}{dt}(t)}\right) \ = \\ = \lim\nolimits_{h\to 0}\tfrac{K'_t(s+h,t)-K'_t(s,t)}{h} \ = \ \tfrac{\partial^2 K}{\partial s\partial t}(s,t).\end{gathered}$$

Теорема доказана.

Пусть случайный процесс $\xi(t)$ задан на $[A,B]$, $A=t_0<t_1 \ < \ \ldots \ < t_m = B$, $\Delta t_j \ = \ t_j - \ t_{j-1}$. Рассмотрим

$$S_m \ = \ \sum_{j=1}^{m}\xi(\theta_j)\Delta t_j, \quad \theta_j\in[t_{j-1},t_j],\ j=1,...,m.$$

Говорят, что процесс $\xi(t)$ *интегрируем в среднем квадратическом*, если найдется элемент $I\in\mathbf{L}_2$ такой, что независимо от выбора точек t_j,θ_j справедливо

$$I \ = \ \text{l.i.m.}_{\Delta\to 0}S_m, \quad \Delta \ = \ \max_j \Delta t_j.$$

Обозначение, принятое для интеграла I от процесса в среднем квадратическом совпадает с обозначением обычного интеграла:

$$I = \int_A^B \xi(t)\, dt.$$

При этом, конечно же, не следует забывать, что I является интегрируемой с квадратом случайной величиной. Полностью аналогично предыдущей доказывается следующая

Теорема 5. *Пусть $m(t)$ интегрируема на $[A, B]$ и определен двойной интеграл $\int_A^B \int_A^B K(t,s)dtds$, тогда процесс $\xi(t)$ интегрируем в среднем квадратическом, причем*

$$\mathbf{M}\xi_*(t)\overline{\int_A^B \xi_*(t)dt} = \int_A^B K(t,s)ds, \ \mathbf{M}\int_A^B \xi_*(t)dt\overline{\int_A^B \xi_*(t)dt} = \int_A^B\int_A^B K(t,s)dtds.$$

4.3. Стохастический интеграл от неслучайной функции

Пусть U – множество, $\mathcal{A}$ – некоторый подкласс $\mathcal{P}(U)$ (системы подмножеств множества U), $\hat{m}$ – конечная неотрицательная мера на $\sigma(\mathcal{A})$. $\mu : \mathcal{A} \to \mathbf{L}_2(<\Omega, \mathcal{F}, \mathbf{P}>)$ называется *элементарной ортогональной стохастической мерой со структурной функцией* $\hat{m}$, если:

1) $\mu(A_1 \cup A_2) = \mu(A_1) + \mu(A_2)$ при $A_1 \cap A_2 = \emptyset$;

2) $(\forall A)\ \mathbf{M}\mu(A) = 0,\ \mathbf{M}|\mu(A)|^2 = \hat{m}(A)$;

3) $\mathbf{M}\mu(A)\mu(B) = 0$ при $AB = \emptyset$ (условие ортогональности).

Рассмотрим пространство

$$\mathbf{L}_2(\hat{m}) = \left\{\psi : U \to \mathbf{C} \mid \int_U |\psi(t)|^2 d\hat{m} < \infty\right\}.$$

Пусть $A_1, ..., A_k \in \mathcal{A}$ образуют конечное разбиение U. Тогда

$$\varphi(t) = \sum_{j=1}^k c_j \mathbf{1}_{A_j}(t) \in \mathbf{L}_2(\hat{m})-$$

простая функция для произвольного набора констант $c_1, ..., c_k$ (здесь $\mathbf{1}_A$ – индикатор множества A.) По определению положим

$$\int\limits_U \varphi(t) d\mu \;=\; \sum_{j=1}^{k} c_j \mu(A_j).$$

Для произвольной неотрицательной функции $\varphi \in \mathbf{L}_2(\hat{m})$ можно обычными способами построить последовательность простых функций, сходящихся к φ в поточечном смысле. Для такой последовательности полагаем

$$\int\limits_U \varphi(t) d\mu \;=\; \text{l.i.m.} \int\limits_U \varphi_n(t) d\mu.$$

В общем случае определение завершается разбиением неслучайной функции из $\mathbf{L}_2(\hat{m})$ на положительную и отрицательную части.

Пусть $\varphi(t),\ \psi(t)$ – простые функции. Без ограничения общности можно считать, что

$$\varphi(t) \;=\; \sum_j f_j \mathbf{1}_{A_j}, \quad \psi(t) \;=\; \sum_j g_j \mathbf{1}_{A_j}$$

для одного набора $A_j,\ j = 1, ..., n$. Тогда

$$\begin{aligned} &< \textstyle\int_U \varphi(t) d\mu,\ \int_U \psi(t) d\mu >_{\mathbf{L}_2} \;=\; \mathbf{M} \int_U \varphi(t) d\mu \cdot \int_U \overline{\psi(t)} d\mu \;= \\ &=\; \textstyle\sum_{j,k} f_j \bar{g}_k \mathbf{M} \mu(A_j) \overline{\mu(A_k)} \;=\; \sum_k f_k \bar{g}_k \hat{m}(A_k) \;= < \varphi, \psi >_{\mathbf{L}_2(\hat{m})} . \end{aligned}$$

Если функции $\varphi, \psi \in \mathbf{L}_2(\hat{m})$ не были простыми, то при помощи очевидным образом организованного предельного перехода может быть получено соотношение

$$< \int\limits_U \varphi(t) d\mu,\ \int\limits_U \psi(t) d\mu >_{\mathbf{L}_2(\Omega, \mathcal{F}, \mathbf{P})} = < \varphi, \psi >_{\mathbf{L}_2(\hat{m})} .$$

Таким образом, доказана

Теорема 6. *Стохастический интеграл осуществляет изометрическое соответствие между* $\mathbf{L}_2(< \Omega, \mathcal{F}, \mathbf{P} >)$ *и* $\mathbf{L}_2(\hat{m})$.

5. Спектральные представления и задача прогнозирования случайного процесса

5.1. Спектральная функция

Пусть $\xi(t)$ – стационарный случайный процесс, заданный на вещественной прямой, или на каком-то ее отрезке $[A, B]$, $\mathbf{M}\xi(t) \;=\; 0$. Ковариационная функция в этой ситуации имеет вид

$$K(t, s) \;=\; \mathbf{M}\xi(t)\overline{\xi(s)} \;=\; \mathbf{M}\xi(t - s + A)\overline{\xi(A)} \;=\; K(t - s),$$

то есть может быть заменена функцией одного аргумента. Условие неотрицательной определенности (6) перепишется теперь в виде

$$(\forall k)(\forall\ c_1, ..., c_k \in \mathbf{C})\ (\forall t_1, ..., t_k \in [A, B])\ \ \sum_{i=1}^{k} \sum_{j=1}^{k} c_i \bar{c}_j K(t_i - t_j)\ \geq\ 0. \quad (8)$$

Лемма 4. *Пусть выполнено (8). Тогда*

1) $K(0)$ *– неотрицательное действительное число;*

2) $(\forall t)\ K(t)\ =\ \overline{K(-t)}$;

3) $(\forall t)\ |K(t)| \leq K(0)$.

Доказательство. 1). Взяв в (8) $k = 1,\ c_1 = 1,\ t_1 = 0$, получим $K(0) \geq 0$.

2). Полагая в условии неотрицательной определенности $k = 2,\ t_1 = t,\ t_2 = 0$, видим, что для произвольных комплексных констант c_1, c_2 выполнено

$$(|c_1|^2 + |c_2|^2)K(0)\ +\ c_1\bar{c}_2 K(t)\ +\ \bar{c}_1 c_2 K(-t)\ \geq 0. \quad (9)$$

В силу произвольности констант и пункта 1 леммы отсюда следует, что

$$(\forall a \in \mathbf{C})\ aK(t) + \bar{a}K(-t)\ \in\ \mathbf{R}.$$

Пусть $a = a_1 + ia_2,\ K(t) = K_1 + iK_2,\ K(-t) = K_3 + iK_4$. Тогда, подставляя эти выражения в формулу

$$\Im m(aK(t) + \bar{a}K(-t))\ =\ 0,$$

получаем, что для произвольных действительных чисел a_1, a_2 должно выполняться

$$a_1(K_2 + K_4)\ +\ a_2(K_1 - K_3)\ =\ 0,$$

откуда $K_2 = -K_4,\ K_1 = K_3$, что и требовалось доказать.

3). Выберем в (8) $k = 2,\ t_1 = t,\ t_2 = 0, c_1 = -K(0),\ c_2 = K(t)$. Тогда из (9) видим, что

$$(K^2(0) + |K(t)|^2)K(0)\ -\ K(0)|K(t)|^2\ -\ K(0)|K(t)|^2\ \geq 0,$$

что после раскрытия скобок превращается в неравенство

$$K^2(0)\ -\ |K(t)|^2\ \geq 0.$$

Лемма доказана.

Теорема 7 (Бохнер-Хинчин). *Комплекснозначная функция действительного аргумента $K(t)$ удовлетворяет условию неотрицательной неопределенности (8) тогда и только тогда, когда найдется функция распределения $F(\lambda)$ такая, что*

$$(\forall t)\ \frac{K(t)}{K(0)} \ = \int\limits_{-\infty}^{\infty} e^{it\lambda} dF(\lambda). \tag{10}$$

Функция, существование которой утверждается в теореме, носит название *спектральной функции*, а представление (10) называют *спектральным представлением* ковариационной функции. Доказательство теоремы Бохнера - Хинчина можно найти в книге [5, параграф 37]. Здесь мы сформулируем и докажем подобную теорему для случайных последовательностей.

Теорема 8. $K(n)$, $n \in \mathbf{Z}$ *– неотрицательно определенная комплекснозначная последовательность в смысле (8) тогда и только тогда, когда найдется такая функция распределения $F(\lambda)$, сосредоточенного на интервале $[-\pi, \pi]$, что для произвольного целого n справедливо*

$$K(n) \ = \ K(0) \int\limits_{-\pi}^{\pi} e^{in\lambda} dF(\lambda),$$

причем эта функция распределения определена однозначно.

Доказательство. Можно считать, что $K(0) = 1$ (иначе перейдем к рассмотрению $\tilde{\xi}(n) = \xi(n)/\sqrt{K(0)}$). Если спектральное представление имеется, то

$$\sum_{j,k} c_j \bar{c}_k K(n_j - n_k) = \int\limits_{-\pi}^{\pi} \sum_{j,k} c_j \bar{c}_k e^{i\lambda(n_j - n_k)} dF(\lambda) \ = \ \int\limits_{-\pi}^{\pi} \left| \sum_j c_j e^{i\lambda n_j} \right|^2 dF(\lambda) \geq 0,$$

что означает завершение доказательства необходимости.

Обратно. Пусть справедливо (8). Возьмем $\rho \in (0,1)$. При $u \in [-\pi, \pi]$ рассмотрим

$$f_\rho(u) \ = \ \frac{1}{2\pi} \sum_{t \in \mathbf{Z}} K(t) \rho^{|t|} e^{-iut}.$$

Так как $|K(t)| \leq K(0) = 1$, то по признаку Вейерштрасса выписанный ряд сходится абсолютно и равномерно относительно u. При этом из обычных формул для преобразований Фурье имеем

$$K(t)\rho^{|t|} \ = \int\limits_{-\pi}^{\pi} e^{it\lambda} f_\rho(\lambda) d\lambda, \ 1 \ = \int\limits_{-\pi}^{\pi} f_\rho(\lambda) d\lambda.$$

Из второй части выписанной формулы следует, что для проверки того, что функция f_ρ – плотность вероятностного распределения, достаточно проверить ее неотрицательность. Подставим полученные представления в (8), выбрав $c_n = \rho^n e^{iun}$, $t_n = -n$.

$$0 \ \le \ \sum_{m=0}^{\infty}\sum_{n=0}^{\infty} \rho^n e^{iun} \rho^m e^{-ium} K(m-n).$$

Обозначим $m-n$ за z. Тогда $m = n+z \ge 0$, откуда $n \ge -z$. Это означает, что при $z \le 0$ n может принимать любые неотрицательные значения, а при $z > 0$ обязательно $n \ge -z$. Итак, $I_1 + I_2 \ge 0$, где

$$I_1 = \sum_{z\le 0}\sum_{n=-z}^{\infty} \rho^{m+n} e^{-iuz} K(z) = \sum_{z\le 0}\sum_{m=0}^{\infty} \rho^{|z|+2m} e^{-iuz} K(z),$$

$$I_2 = \sum_{z=1}^{\infty}\sum_{n=0}^{\infty} \rho^{z+2n} e^{-iuz} K(z).$$

Собирая две последние формулы (с очевидным переобозначением в формуле для I_1), получаем

$$0 \le \sum_{z\in\mathbf{Z}}\sum_{n=0}^{\infty} \rho^{|z|+2n} e^{-iuz} K(z) = 2\pi f(\rho) \sum_{n=0}^{\infty} \rho^{2n} = \frac{2\pi f(\rho)}{1-\rho^2},$$

откуда и следует, что f_ρ – плотность. Введем

$$F_\rho(\lambda) \ = \ \int\limits_{-\pi}^{\lambda} f_\rho(u)du.$$

Понятно, что эта функция распределения имеет характеристическую функцию $\rho^{|t|}K(t)$. При $\rho \to 1$ эта функция сходится к $K(t)$. По основной теореме о характеристических функциях, $F_\rho \to F$, где F – некоторая функция распределения с характеристической функцией $K(t)$, т.е.

$$K(t) \ = \ \int\limits_{-\pi}^{\pi} e^{it\lambda} dF(\lambda).$$

Тем самым теорема полностью доказана.

Будем до конца раздела, если не оговорено иное, считать, что $\mathcal{A}$ – алгебра отрезков в $[-\pi,\pi]$, если рассматривается случайная последовательность и в $\mathbf{R}$ в непрерывном случае. Обозначим через $\mathbf{H}_\xi$ замыкание в $\mathbf{L_2}(<\Omega,\mathcal{F},\mathbf{P}>)$ множества конечных линейных комбинаций вида $\sum_j \alpha_j \xi(t_j)$, и пусть $\mathbf{P}_\xi$ – мера, порожденная спектральной функцией, т.е. спектральное распределение.

Теорема 9. *Существует элементарная ортогональная стохастическая мера $\mu : \mathcal{A} \to \mathbf{H}_\xi$ со структурной функцией $m(A) = K(0)\mathbf{P}_\xi(A)$ такая, что*
$$(\forall t)\xi(t) = \int e^{itu} d\mu(u), \tag{11}$$
причем интеграл берется по интервалу $[-\pi, \pi]$ в случае стационарной случайной последовательности и по $(-\infty, \infty)$ в непрерывном случае.

Доказательство. Построим изометрическое отображение $I : \mathbf{L}_2(m) \to \mathbf{H}_\xi$. Для начала положим
$$I\left(\sum_{j=1}^{k} c_j e^{it_j u}\right) = \sum_{j=1}^{k} c_j \xi(t_j).$$

Ясно, что для функций рассматриваемого вида выполняется свойство изометричности:
$$< I(e^{itu}),\ I(e^{isu}) >_{\mathbf{L}_2(<\Omega,\mathcal{F},\mathbf{P}>)} = \mathbf{M}\xi(t)\bar{\xi}(s) = K(t-s),$$
$$< e^{itu},\ e^{isu} >_{\mathbf{L}_2(m)} = \int e^{itu}\overline{e^{isu}} dm(u) = K(t-s).$$
Отметим теперь, что конечные линейные комбинации функций e^{itu} всюду плотны в $\mathbf{L}_2(m)$. Действительно, в этом пространстве всюду плотно множество ограниченных функций, в нем – дважды дифференцируемых, в последнем – периодических, в котором всюду плотно множество частичных сумм сходящихся рядов Фурье, т.е. именно таких линейных комбинаций.

Продолжим I на $L_2(m)$ с сохранением изометричности. В силу единственности такого продолжения на всем $L_2(m)$ будет справедливо
$$I(f) = \int f(u)\, dm.$$

Пусть
$$\mu(A) = I(\mathbf{1}_A), \quad A \in \mathcal{A},$$
тогда, поскольку I – линейная функция, а для непересекающихся множеств A, B справедливо $\mathbf{1}_{A\cup B} = \mathbf{1}_A + \mathbf{1}_B$, то
$$\mu(A \cup B) = \mu(A) + \mu(B).$$

Очевидно также, что
$$(\forall A \in \mathcal{A})\ \mathbf{M}|\mu(A)|^2 = \int \mathbf{1}_A^2 dm = m(A).$$

Наконец, для непересекающихся множеств A, B справедливо

$$\begin{aligned}\mathbf{M}\mu(A)\overline{\mu(B)} &= <\mu(A),\ \mu(B)>_{\mathbf{L}_2(<\Omega,\mathcal{F},\mathbf{P}>)} = \\ = <\mathbf{1}_A, \mathbf{1}_B>_{\mathbf{L}_2(m)} &= \int \mathbf{1}_A \mathbf{1}_B dm \ = \ m(A \cap B) = 0.\end{aligned}$$

Тем самым, μ – элементарная ортогональная стохастическая мера со структурной функцией m. Согласно самому определению I, для произвольного t выполнено

$$\xi(t) \ = \ I(e^{itu}) \ = \ \int e^{itu} d\mu(u).$$

Теорема доказана.

Следствие. *Стохастический интеграл $\int f(u)d\mu(u)$ осуществляет изометрическое соответствие между пространствами $\mathbf{L}_2(m)$ и $\mathbf{H}_\xi$, причем*

$$e^{iut} \ \mapsto \xi(t).$$

Отметим, что вместо стохастической меры в последнем представлении можно использовать случайный процесс

$$Z(u) \ = \ \mu((-\infty, u)).$$

При этом соотношение между введенным ранее спектральным представлением случайного процесса (11) и записью

$$\xi(t) \ = \ \int e^{itu} dZ(u)$$

такое же, как соотношение между обычной записью интеграла Лебега - Стильтьеса и записью этого интеграла как интеграла по вероятностной мере на $(-\infty, \infty)$. Процесс $Z(u)$ в последней записи обладает *ортогональными приращениями*, т.е. при произвольном выборе чисел $a < b < c < d$ справедливо

$$\mathbf{M}(Z(b) - Z(a))(Z(d) - Z(c)) \ = \ 0.$$

Пример. Пусть A, η – неотрицательные случайные величины, φ имеет равномерное распределение на $[-\pi, \pi]$ и не зависит от A, η. Рассмотрим случайный процесс

$$\xi(t) \ = \ A\cos(\eta t + \varphi) = \frac{1}{2}Ae^{i\eta t}e^{i\varphi} + \frac{1}{2}Ae^{-i\eta t}e^{-i\varphi}.$$

Последние два слагаемых и представляют собой спектральное представление этого процесса. Соответствующий процесс с ортогональными приращениями постоянен на каждом из интервалов $(-\infty, -\eta), (-\eta, \eta), (\eta, \infty)$. В точке $-\eta$ он совершает скачок величины $\frac{1}{2}Ae^{-i\varphi}$, а в точке η – $\frac{1}{2}Ae^{i\varphi}$. Спектральная мера μ рассматриваемого процесса $\xi(t)$ сосредоточена в двух точках – $\pm\eta$.

Упражнение. Докажите, что случайный процесс, рассмотренный в примере, стационарен.

5.2. Формула Котельникова - Шеннона

В качестве первого примера применения спектральных представлений получим формулу, находящую свои приложения при организации систем связи. Пусть $\xi(t)$, $t \in \mathbf{Z}$ – стационарная случайная последовательность. Предположим, что при некотором натуральном a носитель спектральной меры m этой последовательности сосредоточен в $[-\frac{\pi}{a}, \frac{\pi}{a}]$. Тогда на этом носителе равномерно (и абсолютно) сходится ряд

$$\sum_{z \in \mathbf{Z}} \alpha_z e^{izua} = e^{iut},$$

где

$$\alpha_z = \frac{\sin\left(\frac{\pi t}{a} + \pi z\right)}{t + \pi z} \sqrt{\frac{2a}{\pi}} -$$

коэффициент Фурье. Из равномерной сходимости следует сходимость в $\mathbf{L}_2$, а значит имеет место представление

$$\xi(t) = \int_{-\pi}^{\pi} e^{iut} d\mu(u) = \sum_{z \in \mathbf{Z}} \alpha_z \int_{-\pi}^{\pi} e^{izau} d\mu(u),$$

откуда следует, что

$$\xi(t) = \sum_{z \in \mathbf{Z}} \alpha_z \xi(az).$$

Эта формула и носит название формулы Котельникова - Шеннона. Она, в частности, означает, что в этой ситуации для восстановления любого значения последовательности ξ достаточно знать, какие значения она принимала в моменты времени, кратные a.

5.3. Линейные преобразования стационарных процессов

Здесь мы рассмотрим одно из применений спектральной теории стационарных случайных процессов, что позволит нам изучать интегралы и производные от процессов в среднеквадратическом смысле. Пусть A – линейный оператор, являющийся комбинацией линейных дифференциальных операторов и интегральных операторов с ядрами, зависящими лишь от разности аргументов. Применение таких операторов к стационарным случайным процессам в результате снова дает стационарный процесс. Получим формулу для вычисления спектральной плотности результирующего процесса.

Сначала рассмотрим оператор дифференцирования. Пусть

$$\xi(t) = \int e^{it\lambda} d\mu(\lambda),$$

где интеграл без пределов по заключенному выше соглашению берется либо по отрезку $[-\pi,\pi]$, либо по всей числовой прямой, тогда

$$\frac{d\xi}{dt}(t) \;=\; \mathrm{l.i.m.}_{h\to 0}\frac{\xi(t+h)-\xi(t)}{h} \;=\; \mathrm{l.i.m.}_{h\to 0}\int \frac{e^{i(t+h)\lambda}-e^{it\lambda}}{h}d\mu(\lambda).$$

Последний интеграл перепишем в виде

$$\int \lim_{h\to 0}\frac{e^{i(t+h)\lambda}-e^{it\lambda}}{h}d\mu(\lambda) \;=\; \int i\lambda e^{it\lambda}d\mu(\lambda).$$

Эта формула показывает, что для получения спектрального представления производной стационарного процесса достаточно формально произвести дифференцирование спектрального представления первоначального процесса под знаком интеграла по t.

Теперь предположим, что

$$\eta(t) \;\equiv\; A\xi(t) \;=\; \int\limits_{-\infty}^{\infty} B(t-s)\xi(s)ds.$$

Тогда

$$\eta(t) \;=\; \int\limits_{-\infty}^{\infty}\int B(t-s)e^{is\lambda}d\mu(\lambda)ds \;=\; \int g(\lambda)e^{i\lambda t}d\mu(\lambda),$$

где

$$g(\lambda) \;=\; \int\limits_{-\infty}^{\infty} B(-u)e^{i\lambda u}du.$$

Таким образом, нами фактически доказана следующая

Лемма 5. *Спектральное представление $\eta \;=\; A\xi(t)$, где A – оператор описанного выше вида, задается формулой*

$$\eta(t) \;=\; \int g(\lambda)e^{i\lambda t}d\mu(\lambda),$$

где функция $g(\lambda)$ строится следующим образом: применим оператор A к функции $e^{it\lambda}$ и подставим туда $t=0$:

$$g(\lambda) \;=\; Ae^{it\lambda}\,|_{t=0}$$

Продемонстрируем, как эта лемма может быть использована для решения стационарных стохастических дифференциальных уравнений.

Пусть P – многочлен. Рассмотрим дифференциальное уравнение, формально записываемое в виде

$$P\left(\frac{d}{dt}\right)\eta(t) \;=\; \xi(t).$$

Тогда из леммы вытекает, что решение его имеет вид

$$\eta(t) \ = \ \int \frac{e^{i\lambda t}}{P(i\lambda)} d\mu(\lambda).$$

При этом должно выполняться

$$\int \frac{dm(\lambda)}{|P(i\lambda)|^2} \ < \ \infty$$

(m – структурная функция меры μ). Если последнее условие нарушается, то ни одного стационарного решения рассматриваемого уравнения не существует. Это следует из леммы 2. Действительно, если учесть легко проверяемое индукцией по степени P соотношение

$$P\left(\frac{d}{dt}\right) e^{it\lambda}|_{t=0} \ = \ P(i\lambda),$$

то

$$P\left(\frac{d}{dt}\right)\eta(t) = \int P\left(\frac{d}{dt}\right) e^{i\lambda t}|_{t=0} e^{i\lambda t} \frac{d\mu(\lambda)}{P(i\lambda)} \ = \ \xi(t).$$

В частности, стационарное решение уравнения

$$\frac{d\eta}{dt}(t) \ = \ \xi(t)$$

существует в том и только том случае, когда

$$\int \frac{dm(\lambda)}{\lambda^2} \ < \ \infty,$$

при этом оно задается формулой

$$\eta(t) \ = \ \int \frac{e^{it\lambda}}{i\lambda} d\mu(\lambda) \ + \ \zeta,$$

где ζ– произвольная случайная величина, не коррелированная с $\mu(C)$ (или, что то же самое, с $Z(\lambda)$).

Если, например, $\xi(t) = A\cos(\eta t + \varphi)$ – процесс из последнего примера, то

$$\eta(t) = \frac{1}{2} A \left(e^{-i\varphi} \frac{e^{-i\eta t}}{-i\eta} + e^{i\varphi} \frac{e^{i\eta t}}{i\eta} \right) + \zeta \ = \ A\sin(\eta t + \varphi) + \zeta.$$

5.4. Задача прогноза стационарного случайного процесса – сингулярный случай

Начнем с общей постановки задачи прогнозирования. Пусть $\xi(t)$ – случайный процесс, а случайная величина $\eta \in \mathbf{L}_2(<\Omega, \mathcal{F}, \mathbf{P}>)$. Требуется, наблюдая значения процесса $\xi(t)$, предсказать значение η, недоступной непосредственному наблюдению. Обозначим через $\hat{\eta}_{\leq t}$ прогноз η по значениям $\xi(s)$, $s \leq t$ наилучший в том смысле, что

$$M(t) \;=\; \mathbf{M}\left|\hat{\eta}_{\leq t} \;-\; \eta\right|^2$$

минимально среди всех допустимых интегрируемых с квадратом прогнозов.

Процесс $\xi(t)$ называется *регулярным слева*, если

$$(\forall \eta \in \mathbf{L}_2(<\Omega, \mathcal{F}, \mathbf{P}>)) \; \lim_{t \to -\infty} M(t) \;=\; \mathbf{D}\eta,$$

и *сингулярным слева*, если этот предел равен нулю.

Смысл этих определений состоит в том, что в сингулярном случае значения любой случайной величины предсказываются абсолютно точно по сколь угодно удаленным по времени в прошлое результатам наблюдений, а в регулярном случае невозможно указать значение η с точностью более высокой, чем ее дисперсия.

Подобным же образом даются определения процессов, регулярных и сингулярных справа (в этом случае предел в определении вычисляется при $t \to +\infty$), а также процессов, *линейно регулярных слева* и *линейно сингулярных слева.* вся разница в том, что при определении наилучшего прогноза минимизация $M(t)$ происходит среди всех линейных прогнозов $\hat{\eta}_{\leq t}$. Возможны, конечно же, и такие понятия, как линейная регулярность (или сингулярность) справа.

Пусть теперь $\xi(t)$, $t \in \mathbf{Z}$ – стационарная случайная последовательность. Задачу линейного прогноза $\eta \;=\; \xi(m)$ по значениям $\xi(t)$, $t \leq 0$ изучим подробнее. Эту задачу можно назвать *задачей прогноза на m шагов вперед.* Пусть $\tilde{m}$ – спектральная мера последовательности $\xi(t)$. Поскольку $\mathbf{L}_2(\tilde{m})$ изометрично $\mathbf{H}_\xi$ – изометрия осуществляется отображением

$$I(f) \;=\; \int_{-\pi}^{\pi} f\, d\mu \;-$$

то для осуществления наилучшего линейного прогноза надо найти такое g в $\mathcal{L}$ – линейной оболочке функций $e^{iz\lambda}$, $z \leq 0$ так, чтобы

$$\left(e^{im\lambda} \;-\; g\right) \;\perp\; \mathcal{L},$$

т.е.

$$\int\limits_{-\pi}^{\pi} \left(e^{im\lambda} - g(\lambda)\right) e^{-in\lambda} d\tilde{m}(\lambda) \;=\; 0, \quad n \le 0. \tag{12}$$

Линейная сингулярность слева означает возможность аппроксимировать $e^{im\lambda}$ при помощи линейных комбинаций $e^{in\lambda}$, $n \le 0$ сколь угодно точно. При этом точность понимается в смысле формулы (12).

Теорема 10. *Замкнем интервал $[-\pi, \pi)$ в окружность. Если носитель спектральной меры $\xi(t)$ сосредоточен на дуге, занимающей менее половины этой окружности, то последовательность $\xi(t)$ линейно сингулярна слева.*

Доказательство. Обозначим через z число $e^{-i\lambda}$. Пусть z_0 – середина той дуги окружности, на которой сосредоточен носитель $\tilde{m}$. Ясно, что можно считать, что $z_0 \neq 0$. Тогда (это можно легко доказать, сделав соответствующий рисунок) найдется такое N, что для любого элемента z носителя спектральной меры справедливо $|Nz_0 - z| \le N$. Запишем

$$e^{i\lambda} \;=\; \frac{1}{Nz_0 + (z - Nz_0)} \;=\; \sum_{n=0}^{\infty} \frac{(Nz_0 - e^{-i\lambda})^n}{(Nz_0)^{n+1}},$$

а значит, применяя бином Ньютона, извлекаем отсюда требуемое разложение по отрицательным степеням e:

$$e^{i\lambda} \;=\; \sum_{n=0}^{\infty} \sum_{k=0}^{n} (-1)^k \frac{C_n^k}{(Nz_0)^{n+1}} (Nz_0)^{n-k} e^{-ik\lambda}.$$

Теорема доказана.

Отметим, что в сингулярном случае мы заодно получили формулы наилучшего линейного прогноза. В частности, формула для прогноза на один шаг выглядит следующим образом:

$$\hat{\xi}(1)_{\le 0} \;=\; \sum_{n=0}^{\infty} \frac{1}{(Nz_0)^{n+1}} \sum_{k=0}^{n} (-1)^k C_n^k (Nz_0)^{n-k} \xi(-k).$$

5.5. Задача прогноза стационарного случайного процесса – регулярный случай

Построение прогностических формул

Сейчас рассмотрим случай, когда $\xi(t)$, $t \in \mathbf{Z}$ – регулярная стационарная случайная последовательность. Ограничимся ситуацией, в которой она имеет ограниченную спектральную плотность. Это означает, что найдутся такие положительные постоянные c_1, c_2, что для произвольного $\lambda \in [-\pi, \pi]$ имеет место неравенство $c_1 < f(\lambda) < c_2$. Тем самым

$$\mathbf{L}_2(\tilde{m}) \ = \ \mathbf{L}_2[-\pi,\ \pi].$$

Пусть $\mathbf{H}_{\leq 0}$ – замыкание $\mathcal{L}$ в $\mathbf{L}_2[-\pi,\ \pi]$, $\mathbf{H}_{>0}$ – замыкание линейной оболочки $e^{in\lambda}$, $n > 0$ в том же пространстве. Из существования и единственности разложения функций в ряд Фурье следует, что

$$(\forall g \in \mathbf{L}_2[-\pi,\ \pi])\, (\exists! h_1 \in \mathbf{H}_{\leq 0}, h_2 \in \mathbf{H}_{>0})\ g = h_1 + h_2.$$

Условие принадлежности пространству $\mathbf{H}_{\leq 0}$ очевидно, состоит в том, что коэффициенты Фурье с положительными индексами все равны 0.

Перепишем формулу (12) в следующем виде

$$\int_{-\pi}^{\pi} \left(e^{im\lambda} - g(\lambda)\right) e^{-in\lambda} f(\lambda) d\lambda \ = \ 0, \quad n \leq 0.$$

Тем самым, необходимо найти такую функцию $g \in \mathbf{H}_{\leq 0}$, чтобы

$$\left(e^{im\lambda} \ - \ g(\lambda)\right) f(\lambda) \ \in \mathbf{H}_{>0}. \tag{13}$$

Введем еще несколько обозначений. Через $\mathbf{C}_{\geq 0}$ обозначим замыкание линейной оболочки $e^{in\lambda}$, $n \geq 0$ в смысле равномерной сходимости, через $\mathbf{C}_{\leq 0}$ – аналогичное замыкание $e^{in\lambda}$, $n \leq 0$. Нетрудно понять, что введенные классы функций являются алгебрами функций, т.е. замкнуты относительно суммирований и умножений. Более того, ясно, что они замкнуты и относительно взятия экспонент, т.е. если, например, $g \in \mathbf{C}_{\leq 0}$, то $e^g \in \mathbf{C}_{\leq 0}$. Читателю в качестве несложного упражнения предлагается доказать следующее утверждение.

Лемма 6. *Справедливы следующие утверждения:*

$$(g_1 \in \mathbf{H}_{\leq 0},\ g_2 \in \mathbf{C}_{\leq 0}) \ \Longrightarrow \ g_1 g_2 \in \mathbf{H}_{\leq 0},$$
$$(g_1 \in \mathbf{H}_{>0},\ g_2 \in \mathbf{C}_{\geq 0}) \ \Longrightarrow \ g_1 g_2 \in \mathbf{H}_{>0}.$$

Предположим, нам удалось найти такие функции $f_1 \in \ \mathbf{C}_{\leq 0}$, $f_2 \in \mathbf{C}_{\geq 0}$, что $\frac{1}{f_1} \in \mathbf{C}_{\leq 0}$, $\frac{1}{f_2} \in \mathbf{C}_{\geq 0}$ и

$$(\forall \lambda)\ f(\lambda) \ = \ f_1(\lambda) f_2(\lambda).$$

Такое представление мы условимся называть *факторизацией спектральной плотности.* В силу единственности разложения в ряд Фурье и того факта, что $f(\lambda)$ – действительное число, можно (возможно, лишь позаботившись о константе – см. примеры ниже) считать, что $f_1 = \overline{f_2}$. Ниже мы докажем, что факторизация в рассматриваемом случае возможна и укажем некоторые алгоритмы ее получения. Домножая обе части (13) на $1/f_2$, получаем эквивалентную форму этого условия:

$$\left(e^{im\lambda} \;-\; g(\lambda)\right) f_1(\lambda) \;\equiv\; h(\lambda) \in \mathbf{H}_{>0},$$

откуда

$$e^{im\lambda} f_1(\lambda) \;=\; g(\lambda) f_1(\lambda) \;+\; h(\lambda).$$

Прибегнем к разложению Фурье. Пусть

$$f_1(\lambda) \;=\; c_0 + \sum_{j=1}^{\infty} c_{-j} e^{-ij\lambda}.$$

Отсюда

$$e^{im\lambda} f_1(\lambda) \;=\; \sum_{k=0}^{m-1} c_{-k} e^{i(m+k)\lambda} \;+\; c_{-m} \;+\; \sum_{j=1}^{\infty} c_{-m-j} e^{-ij\lambda}.$$

Учитывая, что первая группа слагаемых здесь имеет неотрицательные коэффициенты при мнимой единице в показателях экспонент: а вторая – отрицательные, можно выбрать

$$g(\lambda) \;=\; \frac{c_{-m} + \sum_{j=1}^{\infty} c_{-m-j} e^{-ij\lambda}}{f_1(\lambda)}.$$

Предположим, построенная только что функция g разложена в ряд Фурье

$$g(\lambda) \;=\; b_0 \;+\; \sum_{j=1}^{\infty} b_{-j} e^{-ij\lambda},$$

тогда наилучший линейный прогноз на m шагов задается следующим образом:

$$\hat{\xi_{\le 0}} \;=\; \sum_{j=0}^{\infty} b_{-j} \xi(-j).$$

Вычислим среднеквадратическую ошибку прогноза за m шагов, воспользовавшись изометричностью пространств $\mathbf{L}_2(\Omega)$ и $\mathbf{L}_2(\bar{m})$.

$$\sigma^2(m) \;=\; \int_{-\pi}^{\pi} \left| e^{im\lambda} - g(\lambda) \right|^2 f(\lambda) d\lambda.$$

В силу сделанных выше допущений $f(\lambda) = f_1(\lambda)\overline{f_1(\lambda)}$, а значит,

$$\sigma^2(m) \;=\; \int_{-\pi}^{\pi} \left| \left(e^{im\lambda} - g(\lambda)\right) f_1(\lambda) \right|^2 d\lambda \;=\; \int_{-\pi}^{\pi} |h(\lambda)|^2 d\lambda.$$

Отсюда, с учетом ортонормированности в $\mathbf{L}_2[-\pi,\pi]$ базиса $\frac{1}{\sqrt{2\pi}}e^{in\lambda},\ n\in\mathbf{Z}$, получаем

$$\sigma^2(m) \ = \ \int\limits_{-\pi}^{\pi}\left|c_0e^{im\lambda}+c_1e^{i(m-1)\lambda}\ +...+c_{-m+1}e^{i\lambda}\right|^2d\lambda \ = \ 2\pi\sum_{j=1}^{m}|c_{-m+j}|^2.$$

В силу свойства стационарности рассмотрение прогноза на бесконечно большое число шагов совпадает с задачей прогнозирования по "очень удаленному прошлому". При $m\to\infty$ видим, что

$$\sigma^2(m)\to 2\pi\sum_{j=1}^{\infty}|c_{-j}|^2 \ = \ \int\limits_{-\pi}^{\pi}|f_1(\lambda)|^2d\lambda \ = \ K(0),$$

т.е. дисперсии случайной величины $\xi(m)$, что еще раз подтверждает, что в этой ситуации имеется регулярный случай.

Как осуществить факторизацию

Сначала – общие рекомендации. Пусть спектральная плотность f достаточно гладкая. Рассмотрим

$$\ln f(\lambda) \ = \ \sum_{z\in\mathbf{Z}}a_ze^{iz\lambda}.$$

При всех значениях аргумента этот логарифм является действительным числом, т.к. $f(\lambda)$ принимает только положительные значения. Отсюда

$$a_0>0,\quad (\forall z)\ a_{-z}=\overline{a_z}.$$

Выберем

$$R_1(\lambda) \ = \ \frac{a_0}{2}+\sum_{z<0}a_ze^{iz\lambda},\ \ R_2(\lambda) \ = \ \frac{a_0}{2}+\sum_{z>0}a_ze^{iz\lambda}.$$

Тогда $R_1\in\mathbf{C}_{\le 0},\ R_2\in\mathbf{C}_{\ge 0},\ \overline{R_1}=R_2$. Положим

$$f_1(\lambda) \ = \ \exp R_1(\lambda),\quad f_2(\lambda) \ = \ \exp R_2(\lambda).$$

Нетрудно понять, что эти функции и осуществляют нужную факторизацию. Используя разложение экспоненты в степенной ряд, можно определить коэффициенты разложения Фурье функции f_1 через соответствующие коэффициенты R_1:

$$f_1(\lambda) \ = \ 1+\sum_{j=1}^{\infty}\frac{R_1^j(\lambda)}{j!} \ = \ 1+\sum_{j=1}^{\infty}\frac{1}{j!}\left(\frac{a_0}{2}+\sum_{z<0}a_ze^{iz\lambda}\right)^j.$$

В частности, отсюда

$$c_0 \;=\; 1 + \sum_{j=1}^{\infty} \frac{1}{j!} \left(\frac{a_0}{2}\right)^j \;=\; e^{a_0/2},$$

а следовательно, среднеквадратическая ошибка прогноза за один шаг

$$\sigma^2(1) \;=\; 2\pi |c_0|^2 \;=\; 2\pi e^{a_0},$$

или

$$\sigma^2(1) \;=\; 2\pi \exp\left\{ \frac{1}{2\pi} \int\limits_{-\pi}^{\pi} \ln f(\lambda) d\lambda \right\}.$$

В конце раздела рассмотрим частный случай, когда спектральная плотность представляется в виде отношения двух многочленов, т.е.

$$f(\lambda) \;=\; \frac{P(e^{i\lambda})}{Q(e^{i\lambda})},$$

P, Q – многочлены. Если комплексное число z лежит на единичной окружности с центром в начале координат, то из нашего представления следует, что $P(z)/Q(z)$ – действительное число, а значит каждому корню z_i одного из многочленов, отличному от 0, соответствует корень $(\overline{z_i})^{-1}$, симметричный z_i относительно единичной окружности. Тогда наша дробь может быть записана в виде

$$\frac{P(z)}{Q(z)} \;=\; c' z^k \frac{(z - z_1)(z - \bar{z}_1^{-1})...(z - z_n)(z - \bar{z}_n^{-1})}{(z - w_1)(z - \bar{w}_1^{-1})...(z - w_m)(z - \bar{w}_m^{-1})},$$

где c' – комплексная постоянная, k – разность кратностей корня 0 у многочленов P и Q. Условимся считать, что корни z_i, w_j по модулю больше 1, а сопряженные их обратным, соответственно, меньше 1. Заметим, что

$$(z - z_i)(z - \bar{z}_i^{-1}) \;=\; \frac{(z_i - z)(\bar{z}_i - z^{-1})}{-z^{-1}\bar{z}_i}.$$

В результате получили

$$\frac{P(z)}{Q(z)} \;=\; c z^{k-n+m} \frac{(z_1 - z)(\bar{z}_1 - z^{-1})...(z_n - z)(\bar{z}_n - z^{-1})}{(w_1 - z)(\bar{w}_1 - z^{-1})...(w_m - z)(\bar{w}_m - z^{-1})}.$$

Если $z = e^{i\lambda}$, то $z^{-1} = \bar{z}$ и

$$(z_j - z)(\bar{z}_j - z^{-1}) \;=\; |z_j - z|^2 \;\geq\; 0.$$

Следовательно, c – действительная постоянная, $k + n - m = 0$ и можно выбрать

$$f_1(\lambda) \;=\; \sqrt{c}\frac{(\bar{z}_1 - e^{-i\lambda})...(\bar{z}_n - e^{-i\lambda})}{(\bar{w}_1 - e^{-i\lambda})...(\bar{w}_m - e^{-i\lambda})}. \tag{14}$$

Надо лишь убедиться, что $f_1, f_1^{-1} \in \mathbf{C}_{\leq 0}$. Для этого достаточно показать, что для произвольного комплексного a, по модулю большего 1, справедливо $\left(a - e^{i\lambda}\right)^{-1} \;\in\; \mathbf{C}_{\leq 0}$. Запишем

$$\frac{1}{a - e^{-i\lambda}} \;=\; \frac{\frac{1}{a}}{1 - \frac{e^{-i\lambda}}{a}} \;=\; \frac{1}{a}\left(1 + \sum_{j=1}^{\infty}\left(\frac{e^{-i\lambda}}{a}\right)^j\right).$$

Последняя сумма очевидно лежит в $\mathbf{C}_{\leq 0}$. При выводе мы воспользовались формулой Тейлора и тем фактом, что в силу выбора корней

$$\left|\frac{e^{-i\lambda}}{a}\right| \;<\; 1.$$

Подведем итог. Для получения факторизации спектральной плотности в частном случае достаточно разложить числитель и знаменатель дроби на линейные множители, выбрать корни, по модулю большие 1, и задать f_1 по формуле (14), $f_2 = f/f_1$.

Несколько несложных примеров

Пусть случайная величина ξ такова, что $\mathbf{M}\xi \;=\; 0$, $\mathbf{M}\xi^2 \;=\; 1$, зафиксируем $s \in \mathbf{R}$. Как будет ясно из дальнейшего, можно без ограничения общности считать, что $s \in [-\pi, \pi]$. Пусть $g \in \mathbf{C}$ и

$$\xi_t \;=\; g\xi e^{ist}, \qquad t \in \mathbf{Z}.$$

Нетрудно проверить, что мы задали стационарную случайную последовательность, причем $K(n) = |g|^2 e^{isn}$, $n \in \mathbf{Z}$. Здесь ясно, что

$$K(n) \;=\; \int_{-\pi}^{\pi} e^{i\lambda n} d\mu(\lambda),$$

где стохастическая мера μ сосредоточена в точке s и равна в этой точке g. Итак, мы имеем дело с сингулярным случаем, и в формулах прогноза раздела 5.4 можно взять $z_0 = e^{is}$, $N = 1$.

Слегка обобщим предыдущий пример. Пусть комплекснозначные случайные элементы $z_1, ..., z_N$ таковы, что

$$\mathbf{M}z_i \;=\; 0,\; \mathbf{M}|z_i|^2 = \sigma_i^2,\; \mathbf{M}z_i\overline{z_k} = 0$$

при всех i и $k \neq i$. Фиксируем числа $s_k,\ k = 1, ..., N$ в интервале $[-\pi, \pi]$ и определим стационарную случайную последовательность

$$\xi_n \ = \ \sum_{k=1}^{N} z_k e^{is_k n}, \quad n \in \mathbf{Z}.$$

Понятно, что для соответствующей ковариационной функции

$$K(n) \ = \ \sum_{k=1}^{N} \sigma_k^2 e^{is_k n}$$

спектральная мера сосредоточена в точках $s_1, ..., s_N$ и принимает в них значения $\sigma_1^2, ..., \sigma_N^2$. Вопрос о наличии или отсутствии сингулярности здесь решается на основе расположения этих точек в интервале $[-\pi, \pi]$.

Два предыдущих примера дают нам основание говорить о том, что соответствующие ковариационные функции являются комбинацией небольшого числа гармоник. Следующий пример показывает ситуацию, когда в равной степени задействованы все гармоники. Такой процесс принято называть белым шумом по аналогии с белым цветом, в котором равным образом участвуют все цвета спектра.

Пусть случайные элементы $\xi_k,\ k \in \mathbf{Z}$ независимы, одинаково распределены, имеют нулевые математические ожидания, и $\mathbf{M}|\xi_k|^2 \ = \ 1,\ k \in \mathbf{Z}$. Тогда белым шумом мы назовем стационарную случайную последовательность $\xi_n,\ n \in \mathbf{Z}$. Ее ковариационная функция $K(n)$, равная 1 при $n = 0$ и нулю при остальных n, представима в виде

$$K(n) \ = \ \int_{-\pi}^{\pi} e^{i\lambda n} dF(\lambda),$$

где F – функция равномерного распределения на $[-\pi, \pi]$. В обозначениях предыдущего раздела,

$$f_1(\lambda) \ = \ \frac{1}{\sqrt{2\pi}},$$

в силу чего мы имеем дело с регулярным случаем, но наилучший прогноз здесь равен 0, и величина ошибки даже за один шаг совпадает с дисперсией.

Естественно, что более содержательные примеры регулярного случая связаны с более сложно устроенными спектральными плотностями. Пусть спектральная плотность имеет вид

$$f(\lambda) \ = \ 5 + 4\cos\lambda.$$

Тогда

$$f(\lambda) \ = \ 5 + 2e^{i\lambda} + 2e^{-i\lambda}.$$

Заметим, что ковариационная функция соответствующей стационарной последовательности имеет вид

$$K(t) \;=\; \int\limits_{-\pi}^{\pi} e^{i\lambda t} f(\lambda) d\lambda \;=\; \frac{10\sin\pi t}{t} + \frac{4\sin\pi(t+1)}{t+1} + \frac{4\sin\pi(t-1)}{t-1}.$$

При всех t отличных от 0 и ± 1 это выражение рано 0, $K(0) = 10\pi$, $K(\pm 1) = 4\pi$. Запишем

$$f(\lambda) \;=\; \frac{2e^{2i\lambda} + 5e^{i\lambda} + 2}{e^{i\lambda}} \;=\; \frac{P(e^{i\lambda})}{Q(e^{i\lambda})}.$$

Разлагая числитель на множители, получим

$$f(\lambda) \;=\; 2\frac{\left(e^{i\lambda} + \frac{1}{2}\right)\left(e^{i\lambda} + 2\right)}{e^{i\lambda}} \;=\; \left(2 + e^{-i\lambda}\right)\left(e^{i\lambda} + 2\right).$$

Это означает, что можно выбрать

$$f_1(\lambda) = 2 + e^{-i\lambda}, \quad f_2(\lambda) = e^{i\lambda} + 2,$$

и, в обозначениях предыдущего подраздела, $c_0 = 2$, $c_{-1} = 1$, $c_{-k} = 0$ при $k \geq 4$.

Выпишем формулу прогноза на 1 шаг ($m = 1$).

$$g(\lambda) \;=\; c_{-1} f_1^{-1}(\lambda) \;=\; \frac{1}{2}\frac{1}{1 + \frac{1}{2}e^{-i\lambda}},$$

следовательно, применяя формулу суммы бесконечно убывающей геометрической прогрессии, получим:

$$g(\lambda) \;=\; \frac{1}{2}\left(1 + \sum_{j=1}^{\infty} \frac{(-1)^j e^{-ij\lambda}}{2^j}\right).$$

Окончательно, наилучший линейный прогноз за один шаг задан формулой

$$\hat{\xi}(1)_{\leq 0} \;=\; \frac{1}{2}\xi(0) \;+\; \sum_{j=1}^{\infty} (-1)^j \frac{\xi(-j)}{2^{j+1}}.$$

Вычислим величину ошибки прогноза:

$$\sigma^2(1) \;=\; 2\pi c_0^2 \;=\; 8\pi,$$

что указывает на низкую надежность прогноза. Если бы мы взялись прогнозировать на 2 шага, то получили бы, что

$$\sigma^2(2) \;=\; 2\pi(c_0^2 + c_{-1}^2) \;=\; 10\pi = \mathbf{D}\xi(0),$$

что подтверждает регулярность рассматриваемой последовательности – невозможно дать прогноз точнее, чем с точностью до дисперсии. Таким образом, при прогнозе на два и более шагов наилучшим прогнозом следует признать математическое ожидание $\xi(t)$, или его оценку

$$m^* \;=\; \frac{\xi(-N)+...+\xi(0)}{N+1}.$$

Задача статистического оценивания спектральной плотности

Из предыдущих подразделов становится ясно, что знание спектральной плотности наблюдаемого случайного процесса является крайне важным для решения задач прогноза. С помощью спектральной плотности решаются также многие другие интересные задачи (например, задача решения стационарных дифференциальных уравнений из числа рассмотренных нами). Поэтому дадим рекомендации по ее оценке по данным наблюдений над случайным процессом.

Как обычно, ограничимся случаем стационарной случайной последовательности. Пусть $\mathbf{M}\xi_n = m$,

$$K(n) \;=\; \int\limits_{-\pi}^{\pi} e^{i\lambda n} dF(\lambda).$$

В нашем распоряжении имеются N наблюдений элементов случайной последовательности $x_0, x_1, ..., x_{N-1}$. Естественной оценкой m является, конечно же, их среднее арифметическое m_N, а

$$\hat{K}_N(n) \;=\; \begin{cases} \frac{1}{N-n}\sum_{k=0}^{N-n-1} x_{k+n}\overline{x}_k, & n \ge 0 \\ \frac{1}{N+n}\sum_{k=-n}^{N} x_{k+n}\overline{x}_k, & n < 0. \end{cases}$$

В любом случае число тех пар (i,j), для которых $i-j=m$, равно $N-|m|$. Нетрудно проверить, что $\hat{K}_N(n)$ – несмещенная оценка $K(n)$.

Введем в рассмотрение

$$f_N(\lambda) = \frac{1}{2\pi N}\sum_{s=0}^{N-1}\sum_{t=0}^{N-1} K(s-t)e^{-i\lambda(s-t)} \;\ge\; 0.$$

Ясно, что

$$f_N(\lambda) = \frac{1}{2\pi N}\sum_{|m|<N}\sum_{t=1}^{N-|m|} K(m)e^{-i\lambda m} = \frac{1}{2\pi}\sum_{|m|<N} K(m)\left(1-\frac{|m|}{N}\right)e^{-i\lambda m}.$$

Если $F(\lambda)$ – функция распределения с плотностью f_N, то

$$\int\limits_{-\pi}^{\pi} e^{i\lambda n} dF_N(\lambda) \;=\; \begin{cases} \left(1-\frac{|m|}{N}\right)K(n), & |n|<N, \\ 0, & |n| \ge N \end{cases}$$

в силу ортонормированности системы функций $\frac{1}{\sqrt{2\pi}}e^{-i\lambda n}$, $n \in \mathbf{Z}$. Пусть

$$F(\lambda) \;=\; \lim_{N\to\infty} F_N(\lambda)$$

в смысле слабой сходимости. Тогда из выписанных формул следует, что F – спектральная функция, а $f(\lambda) = \lim_{N\to\infty} f_N(\lambda)$ – спектральная плотность нашей последовательности.

Статистика

$$\hat{f}_N(\lambda) \;=\; \frac{1}{2\pi} \sum_{|M|<N} \hat{K}_N(m) \left(1 - \frac{|m|}{N}\right) e^{-i\lambda m},$$

являющаяся оценкой $f_N(\lambda)$, называется *периодограммой*. Оказывается,

$$\hat{f}_N(\lambda) \;=\; \frac{1}{2\pi N} \left| \sum_{s=0}^{N-1} x_s e^{-i\lambda s} \right|. \tag{15}$$

Действительно, правую часть (15) распишем в виде

$$\frac{1}{2\pi N} \sum_{s=0}^{N-1} \sum_{t=0}^{N-1} x_s \overline{x}_t e^{-i\lambda(s-t)} = \frac{1}{2\pi N} \sum_{|m|<N} \sum_{t} x_{t+m} \overline{x}_t e^{-i\lambda m},$$

и, поскольку во внутренней сумме $N - |m|$ слагаемых, то можно продолжить выписываемое равенство до

$$\frac{1}{2\pi N} \sum_{|m|<N} (N - |m|) \hat{K}(m) e^{-i\lambda m} \;=\; \hat{f}_N(\lambda).$$

Наша цель – заметить, что можно использовать (15) в качестве оценки спектральной плотности. Т.к. $\mathbf{M}\hat{K}(n) = K(n)$, то $\hat{f}(\lambda)$ является несмещенной оценкой $f_N(\lambda)$. Заметим далее, что

$$f_N(\lambda) \;=\; \frac{1}{2\pi N} \sum_{s=0}^{N-1} \sum_{t=0}^{N-1} K(s-t) e^{-i\lambda(s-t)} \;=$$

$$= \frac{1}{2\pi N} \sum_{s=0}^{N-1} \sum_{t=0}^{N-1} \int_{-\pi}^{\pi} e^{i\mu(s-t)} f(\mu) d\mu \; e^{-i\lambda(s-t)} \;=\; \frac{1}{2\pi N} \int_{-\pi}^{\pi} \left| \sum_{k=0}^{N-1} e^{i(\mu-\lambda)k} \right|^2 f(\mu) d\mu.$$

Функция

$$\Phi_N(\lambda) = \frac{1}{2\pi N} \left| \sum_{k=0}^{N-1} e^{i\lambda k} \right|^2 = \frac{1}{2\pi N} \left| \frac{1 - e^{i\lambda N}}{1 - e^{i\lambda}} \right|^2 = \frac{1}{2\pi N} \left| \frac{\sin(\lambda N/2)}{\sin(\lambda/2)} \right|^2$$

известна как *ядро Фейера*. Из курса функционального анализа известна следующая сходимость, связанная с этим ядром:

$$\mathbf{M}\hat{f}_N(\lambda) = f_N(\lambda) = \int_{-\pi}^{\pi} \Phi(\mu - \lambda) f(\mu)\, d\mu \ \to \ f(\lambda)$$

при любом λ и $N \to \infty$. Итак, $\hat{f}_N(\lambda)$ является асимптотически несмещенной оценкой для $f(\lambda)$.

Но в практических задачах ошибки двух в целом неплохих приближений обычно накапливаются, что приводит к тому, что оценка (15) не может использоваться в сколько-нибудь удовлетворительном смысле. Пусть, например, ξ_n – белый шум, составленный из стандартных нормальных величин. Тогда, как мы уже знаем, $f(\lambda) \ \equiv \ \frac{1}{2\pi}$, а

$$\hat{f}_N(\lambda) \ = \ \frac{1}{2\pi} \left| \frac{1}{\sqrt{N}} \sum_{k=0}^{N-1} x_k e^{i\lambda k} \right|^2 .$$

Даже при $\lambda = 0$ у нас $\hat{f}_N(0) \sim \frac{1}{2\pi}\xi^2$, где ξ имеет стандартное нормальное распределение, и

$$\mathbf{M}|\hat{f}_N(0) - f(0)|^2 \ = \ \frac{1}{4\pi^2}\mathbf{M}|\xi^2 - 1|^2 \not\longrightarrow 0.$$

Обычно оценку (15) улучшают за счет введения специальной весовой функции W_N, называемой *спектральным окном*, и полагают

$$f_N^W(\lambda) \ = \ \int_{-\pi}^{\pi} W_N(\lambda - \nu)\hat{f}_N(\nu) d\nu.$$

От спектрального окна при этом требуют

1) $W_N(0)$ – максимальное значение W_N;

2) $\int_{-\pi}^{\pi} W_N(\lambda) d\lambda = 1$;

3) $(\forall\lambda)\ \lim_{N\to\infty} \mathbf{M}|\hat{f}_N^W(\lambda) - f(\lambda)|^2 = 0$.

Можно предложить много способов построения спектральных окон. Вот наиболее распространенные. Пусть $a_N = o(N) \to \infty$ – произвольная числовая последовательность,

$$W_N(\lambda) \ = \ a_N B(a_N \lambda).$$

Если

$$B(\mu) \ = \ \frac{1}{2\pi} \left| \frac{\sin(\mu/2)}{\mu/2} \right|^2 ,$$

то мы получаем *окно Бартлетта.* Если в тех же обозначениях выбрать

$$B(\mu) \;=\; \frac{3}{8\pi}\left|\frac{\sin(\mu/4)}{\mu/4}\right|^4,$$

то получаем *окно Парзена.* Наконец, последовательность a_N можно брать не произвольной, а построить определенным образом, взять любое $\alpha \in (0,2]$ и

$$B(\mu) \;=\; \begin{cases} \frac{\alpha+1}{2\alpha}\,(1-|\mu|^\alpha), & |\mu| \le 1,\\ 0, & |\mu| > 1, \end{cases}$$

то мы получим *окно Журбенко.* Каждый из этих подходов имеет свои достоинства и недостатки, которые здесь мы обсуждать не будем.

6. Процессы размножения и гибели

6.1. Основные предположения

Пусть в некоторой области имеются частицы. Будем считать, что их количество в этой области случайно и обозначим $p_n(t)$ вероятность того, что в момент времени t имелось ровно n частиц. Условимся считать, что количество частиц в области образует цепь Маркова с состояниями 0,1, ... а вероятности перехода удовлетворяют следующим условиям

1) $P_{n,n+1}(\Delta t) = \lambda_n \Delta t + o(\Delta t)$;

2) $P_{n,n-1}(\Delta t) = \mu_n \Delta t + o(\Delta t)$;

3) $P_{n,n}(\Delta t) = 1 - (\lambda_n + \mu_n)\Delta t + o(\Delta t)$;

4) $P_{i,j}(0) \;=\; \delta_{i,j}$, где $\delta_{i,j}$– символ Кронекера, равный нулю при несовпадении индексов и единице при совпадении.

Известную последовательность чисел λ_n естественно назвать интенсивностями размножения, а последовательность μ_n – интенсивностями гибели. Если $(\forall n)$ $(\mu_n = 0)$, то говорят о чистом размножении, если же $(\forall n)$ $(\lambda_n = 0)$, то о чистой гибели. Заметим, что из сформулированных допущений следует, что вероятность рождения или гибели за время Δt одновременно двух или более частиц есть $o(\Delta t)$.

Укажем один набор достаточных условий для того, чтобы были справедливы все формулы, выписанные выше. Пусть частицы размножаются и гибнут независимо друг от друга, причем каждая из них за промежуток Δt порождает новую с вероятностью b или гибнет с вероятностью d $(b + d \le 1)$. Предположим, что

$$b = \lambda\Delta t + o(\Delta t), \quad d = \mu\Delta t + o(\Delta t).$$

При этом очевидно, что вероятность за время Δt родиться или погибнуть двум или большему количеству частиц будет равна $o(\Delta t)$ независимо от того, сколько в области имелось к этому моменту частиц. Теперь, обозначая $b_n(\Delta t)$, $d_n(\Delta t)$ числа родившихся и погибших частиц, получим

$$\mathbf{P}(b_n(\Delta t) = 1) \ = \ n\lambda\Delta t + o(\Delta t), \ \mathbf{P}(d_n(\Delta t) = 1) \ = \ n\mu\Delta t + o(\Delta t),$$

а значит,

$$P_{n,n+1}(\Delta t) \ = \ \mathbf{P}(b_n(\Delta t) = 1, d_n(\Delta t) = 0) + o(\Delta t) \ =$$
$$= \ (n\lambda\Delta t + o(\Delta t))(1 - n\mu\Delta t + o(\Delta t)) + o(\Delta t) \ = \ n\lambda\Delta t + o(\Delta t);$$
$$P_{n,n-1}(\Delta t) \ = \ n\mu\Delta t \ + \ o(\Delta t);$$
$$P_{n,n}(\Delta t) = \mathbf{P}(b_n(\Delta t) = 0, d_n(\Delta t) = 0) + \mathbf{P}(b_n(\Delta t) = 1, d_n(\Delta t) = 1) +$$
$$+ o(\Delta t) = = \ 1 - (\lambda + \mu)\Delta t \ + \ o(\Delta t).$$

Тем самым, в этом случае все предположения оказываются выполненными.

6.2. Одномерные распределения

Получим дифференциальные уравнения, описывающие процесс размножения и гибели. По формуле полной вероятности

$$p_n(t + \Delta t) = p_n(t)(1 - n\lambda\Delta t + o(\Delta t))(1 - n\mu\Delta t + o(\Delta t)) +$$
$$+ p_{n-1}(t)((n-1)\lambda\Delta t + o(\Delta t))(1 - (n-1)\mu\Delta t + o(\Delta t)) +$$
$$+ p_{n+1}(t)((n+1)\mu\Delta t + o(\Delta t))(1 - (n+1)\lambda\Delta t + o(\Delta t)).$$

Отсюда нетрудно получить

$$\frac{p_n(t + \Delta t) - p_n(t)}{\Delta t} = -n(\lambda + \mu)p_n(t) + (n-1)\lambda p_(n-1)(t) +$$
$$+ (n+1)\mu p_{n+1}(t) + o(1).$$

Переходя к пределу, имеем

$$p_n'(t) = -n(\lambda + \mu)p_n(t) + (n-1)\lambda p_{n-1}(t) + (n+1)\mu p_{n+1}(t).$$

Здесь предполагалось, что $n \geq 1$. Если же $n = 0$, то аналогичными рассуждениями выводим

$$p_0'(t) = \mu p_1(t).$$

Умножим n-ное уравнение полученной системы на x^n и, складывая, получим

$$\sum_{n=0}^{\infty} p_n'(t)x^n = \mu \sum_{n=1}^{\infty} p_n(t)x^{n-1}n + \lambda x^2 \sum_{n=1}^{\infty} p_n(t)x^{n-1}n -$$

$$-(\lambda+\mu)\sum_{n=0}^{\infty} p_n(t)x^n n.$$

Если мы теперь обозначим

$$F(x,t) = \sum_{n=0}^{\infty} p_n(t)x^n \, ,$$

то полученное выше уравнение можно переписать в виде

$$\frac{\partial F}{\partial t} = (\mu - (\lambda+\mu)x + \lambda x^2)\frac{\partial F}{\partial x}.$$

Предположим для простоты, что $p_1(0) = 1$, т.е. в начальный момент времени в области находилась одна частица. Отсюда, в частности, следует, что $p_n(0) = 0$ при $n \neq 1$. Чтобы решить уравнение в частных производных, составим обыкновенное дифференциальное уравнение.

$$dt \ = \ \frac{-dx}{\mu - (\lambda+\mu)x + \lambda x^2} \, ,$$

откуда

$$t - c = \int \frac{-dx}{\mu - (\lambda+\mu)x + \lambda x^2} \ = \ \frac{1}{\mu-\lambda}\int (\frac{1}{1-x} - \frac{\lambda}{\mu - \lambda x})dx \ =$$

$$= \ \frac{1}{\mu-\lambda}\ln\left|\frac{\mu-\lambda x}{1-x}\right|.$$

Общее решение уравнения в частных производных, т.о., будет

$$F(t,x) \ = \ R\left(\frac{1}{\mu-\lambda}\ln\left|\frac{\mu-\lambda x}{1-x}\right| + t\right) \, ,$$

где R - произвольная дифференцируемая функция. При $t = 0 \quad F(x,t) = x$, или

$$R\left(\frac{1}{\mu-\lambda}\ln\left|\frac{\mu-\lambda x}{1-x}\right|\right) = x,$$

и R есть обратная функция для $u \ = \ \frac{1}{\mu-\lambda}\ln\left|\frac{\mu-\lambda x}{1-x}\right|$. Следовательно,

$$R(u) = \frac{\mu - e^{(\mu-\lambda)u}}{\lambda - e^{(\mu-\lambda)u}}.$$

Подставляя в это выражение вместо u аргумент R, имеем

$$F(x,t) = \frac{\mu(1-e^{(\mu-\lambda)t}) + x(\lambda e^{(\mu-\lambda)t} - \mu)}{\lambda - \mu e^{(\mu-\lambda)t} + \lambda x(e^{(\mu-\lambda)t} - 1)}.$$

Разлагая F в ряд по степеням x, нетрудно получить, что

$$p_n(t) = \frac{(\mu-\lambda)^2 e^{(\mu-\lambda)t}\left(1-e^{(\mu-\lambda)t}\right)^{n-1}\lambda^{n-1}}{(\lambda-\mu e^{(\mu-\lambda)t})^{n+1}} \quad, \ n \geq 1,$$

$$p_0(t) = \frac{\mu\left(1-e^{(\mu-\lambda)t}\right)}{\lambda-\mu e^{(\mu-\lambda)t}}, \ p_{\geq 1}(\lambda) = \frac{\lambda-\mu}{\lambda-\mu e^{(\mu-\lambda)t}}.$$

Здесь $p_{\geq 1}(t) = \sum_{k=1}^{\infty} p_k(t)$.

Пусть $\mu < \lambda$. Тогда при $t \to \infty$

$$p_0(t) \longrightarrow \frac{\mu}{\lambda}, \quad p_{\geq 1}(t) \longrightarrow 1 - \frac{\mu}{\lambda},$$

а если, наоборот, $\mu > \lambda$, то

$$p_0(t) \to 1, \quad p_{\geq 1}(t) \to 0.$$

Выводы сделайте сами!

6.3. Время пребывания и уравнения Колмогорова

Обозначим через T_n время, в течение которого общее число частиц в области остается неизменным и равным n (время пребывания процесса в состоянии n), $F_n(t) = \mathbf{P}(T_N \geq t)$. Тогда можно с учетом сделанных ранее предположений считать, что

$$F_n(t+\Delta t) = F_n(t)P_{n,n}(\Delta t) + o(\Delta t) =$$

$$= F_n(t) - F_n(t)(\lambda_n + \mu_n)\Delta t + o(\Delta t),$$

откуда

$$\frac{F_n(t+\Delta t) - F_n(t)}{\Delta t} = (\lambda_n + \mu_n)F_n(t) + o(1).$$

Перейдем к пределу при $\Delta t \to 0$ и решим получившееся дифференциальное уравнение. С учетом условия $F_n(0) = 1$ получим

$$F_n(t) = \exp\{-(\lambda_n + \mu_n)t\},$$

что означает, что время пребывания имеет экспоненциальное распределение с параметром $\lambda_n + \mu_n$:

$$\mathbf{P}(T_n < t) = 1 - e^{-((\lambda_n+\mu_n)t}, \quad t \geq 0.$$

Выведем также и уравнения, связывающие вероятности перехода процесса из одного состояния в другое. По формуле полной вероятности

$$\begin{aligned} P_{n,k}(t+h) &= P_{n,n-1}(h)P_{n-1,k}(t) + P_{n,n}(h)P_{n,k}(t) + \\ &\quad + P_{n,n+1}(h)P_{n+1,k}(t) + o(h), \qquad\qquad n > 0, \\ P_{0,k}(t+h) &= P_{0,0}(h)P_{0,k}(t) + P_{0,1}(h)P_{1,k}(t) + o(h). \end{aligned}$$

Деля на h и устремляя его к 0, получим систему дифференциальных уравнений:

$$\begin{aligned} P'_{0,k}(t) &= \lambda_0 P_{0,k}(t), \\ P'_{n,k}(t) = \mu_n P_{n-1,k}(t) &- (\lambda_n + \mu_n)P_{n,k}(t) + \lambda_n P_{n+1,k}(t), \quad n > 0. \end{aligned}$$

Эта система носит название *системы обратных дифференциальных уравнений Колмогорова*. Аналогичным образом можно получить *систему прямых дифференциальных уравнений Колмогорова*:

$$\begin{aligned} P'_{n,0}(t) &= -\lambda_0 P_{n,0}(t) + \mu_1 P_{n,1}(t), \\ P'_{n,k}(t) = \mu_{k+1} P_{n,k+1}(t) &- (\lambda_k + \mu_k)P_{n,k}(t) + \lambda_{k-1} P_{n,k-1}(t), \quad n > 0. \end{aligned}$$

Разница в названиях и способах получения этих систем объясняется тем, что для обратной системы у нас изучается как бы "предыстория" попадания в состояние k, а для прямой системы нас интересуют "последствия" выхода системы из состояния n.

Если, независимо от выбора i для каждого j определены

$$p_j = \lim_{t\to\infty} P_{i,j}(t),$$

причем $\sum_j p_j = 1$, то говорят, что числа $p_0, p_1, \ldots$ образуют *стационарное распределение* соответствующего процесса размножения и гибели. Эти числа, очевидно, можно интерпретировать, как вероятности обнаружить в области j частиц через достаточно продолжительное время.

Если стационарное распределение существует, то из системы прямых уравнений Колмогорова, переходя в них к пределу при $t \to \infty$, получим систему для их определения:

$$\begin{aligned} -\lambda_0 p_0 + \mu_1 p_1 &= 0, \\ \lambda_{k-1}p_{k-1} - (\lambda_k + \mu_k)p_k &+ \mu_{k+1}p_{k+1} = 0, \quad k > 0. \end{aligned} \tag{16}$$

Пусть

$$\pi_0 = 1, \quad \pi_j = \frac{\lambda_0 \ldots \lambda_{j-1}}{\mu_1 \ldots \mu_j}, \; j \geq 1.$$

Покажем при помощи математической индукции, что решение системы (16) имеют вид
$$p_j \;=\; \frac{\pi_j}{\sum_k \pi_k},\;\; j \geq 0. \tag{17}$$
Действительно, из первого уравнения в (16) следует
$$p_1 \;=\; \frac{\lambda_0}{\mu_1} p_0 \;=\; \pi_1 p_0.$$
Пусть при $k \leq j$ доказана формула $p_k = \pi_k p_0$. Тогда, вновь привлекая систему, получаем
$$p_{j+1} \;=\; \frac{\lambda_j \pi_j}{\mu_{j+1}} p_0,$$
откуда сразу же имеем $p_{j+1} = \pi_{j+1} p_0$. Сложим все получившиеся соотношения и учтем, что $\sum_k p_k = 1$. Тогда
$$1 \;=\; p_0 \sum_k \pi_k \;\Rightarrow\; p_0 = \frac{\pi_0}{\sum_k \pi_k}.$$
Отсюда немедленно следует формула (17).

6.4. Примеры

Линейный рост с иммиграцией

Пусть в модели процесса размножения и гибели $\lambda_n \;=\; \lambda n + a,\; \mu_n \;=\; \mu n$, что можно интерпретировать как наличие постоянной интенсивности роста и гибели при наличии неизменной относительно общего числа частиц составляющей роста a. Эту составляющую удобно интерпретировать как иммиграцию, не связанную с количеством частиц в области. Привлекая систему прямых уравнений Колмогорова, получим
$$\begin{aligned} P'_{n,0}(t) \;&=\; -aP_{n,0} + \mu P_{n,1}(t), \\ P'_{n,j}(t) \;&=\; (a + \lambda(j-1))P_{n,j-1}(t) - \\ -((\lambda+\mu)j + a)&P_{n,j}(t) + \mu(j+1)P_{n,j+1}(t), \quad j \geq 1. \end{aligned}$$
Умножим j-е уравнение на j и суммируем. Обозначим $M(t)$ математическое ожидание процесса и допустим, что $\xi(0) = i$ с вероятностью 1. Тогда получим
$$M'(t) \;=\; a + (\lambda - \mu)M(t).$$
Если $\lambda \neq \mu$, то, проинтегрировав это уравнение, получим
$$M(t) \;=\; Ce^{(\lambda-\mu)t} \;-\; \frac{a}{\lambda - \mu}.$$

Подставляя сюда $t = 0$, находим

$$c \;=\; i \;+\; \frac{a}{\lambda - \mu}.$$

Окончательно,

$$M(t) \;=\; \frac{a}{\lambda - \mu}\left(e^{(\lambda-\mu)t} - 1\right) \;+\; ie^{(\lambda-\mu)t}.$$

Отметим, что при $\lambda > \mu$ $M(t)$ неограниченно увеличивается с течением времени, если же $\lambda < \mu$, то, монотонно убывая, стремится к $\lambda/(\mu - \lambda)$.

Если же $\lambda = \mu$, то

$$M(t) \;=\; at + i, \qquad M(t) \longrightarrow \infty.$$

Образование очереди в системе массового обслуживания

Рассмотрим обслуживание клиентов некоторой системой – например, выбивание чеков кассой или стрижка в парикмахерской. Пусть $\xi(t)$ – число клиентов в зоне обслуживания (считая тех, кто ожидает своей очереди и тех, кто уже обслуживается) в момент времени t. Будем считать, что, независимо от числа уже имеющихся клиентов, число прибывающих клиентов имеет одну и ту же интенсивность, интенсивность обслуживания клиентов также постоянна. Если канал обслуживания только один (т.е. одновременно не может обслуживаться более одного клиента), то в рамках принятой модели отсюда следует, что $(\forall n)\,\lambda_n = \lambda,\ \mu_n = \mu$. Обозначим через τ_j время обслуживания j-го клиента. Ранее было показано, что случайные величины τ_j имеют экспоненциальные распределения, более точно

$$(\forall t \geq 0)\ \mathbf{P}(\tau_j \geq t) \;=\; e^{-\mu t}.$$

Если всего имеется n каналов обслуживания (касс, парикмахерских или стоматологических кресел и т.п.), то одно из них освободится через время τ – минимум из величин $\tau_1, ..., \tau_n$. Нетрудно видеть, что

$$\mathbf{P}(\tau \geq t) \;=\; \prod_{j=1}^{n} \mathbf{P}(\tau_j \geq t) \;=\; e^{-\mu n t},$$

что означает, что время ожидания первого освободившегося канала обслуживания также имеет экспоненциальное распределение. Таким образом, если имеется n каналов обслуживания, то $\mu_j = j\mu$ при $j \leq n$ и $\mu_j \;=\; n\mu$ при $j > n$.

Вернемся к системе с одним каналом обслуживания. Во введенных выше обозначениях (см (17)):

$$\pi_j \;=\; \left(\frac{\lambda}{\mu}\right)^j, \quad \sum_{j=1}^{\infty} \pi_j \;=\; \frac{\mu}{\mu - \lambda}$$

(конечно же, только для $\mu > \lambda$, иначе очередь растет, и ни о каких стационарных распределениях говорить не имеет смысла). Привлекая (17), видим, что

$$p_n \;=\; \frac{\mu-\lambda}{\mu}\,\left(\frac{\lambda}{\mu}\right)^n, \quad n \ge 0.$$

Эти числа равны вероятностям того, что через продолжительное время с начала обслуживания длина очереди будет равна $n-1$ (один клиент обслуживается).

Рассмотрим другой крайний случай – наличие бесконечного числа каналов обслуживания. Это означает, что каждый появляющийся в зоне обслуживания клиент немедленно начинает обслуживаться. Такую модель системы массового обслуживания иногда называют моделью телефонного узла. Здесь

$$\lambda_n = \lambda, \quad \mu_n = n\mu,$$

а следовательно,

$$\pi_j \;=\; \frac{\lambda^j}{j!\mu^j}, \quad \sum_{j=1}^{\infty}\pi_j \;=\; e^{\lambda/\mu}$$

и

$$p_n \;=\; \frac{1}{n!}\left(\frac{\lambda}{\mu}\right)^n e^{-\frac{\lambda}{\mu}}, \quad n \ge 0.$$

Это – вероятности того, что одновременно через телефонный узел ведутся n разговоров. Как мы видим, число одновременно ведущихся разговоров имеет пуассоновское распределение.

7. Непрерывность реализаций

В этой главе будем считать, что все рассматриваемые случайные процессы заданы при $t \in [a,b]$. Начнем со следующего определения. Два случайных процесса $\xi(t), \eta(t)$ называются *стохастически эквивалентными*, если

$$(\forall t)\ \mathbf{P}(\xi(t) = \eta(t)) \;=\; 1.$$

Любой процесс, стохастически эквивалентный $\xi(t)$ будем называть *модификацией* $\xi(t)$.

Существование модификаций процесса, обладающих непрерывными модификациями, весьма удобно. Для таких процессов мы имеем возможность исследовать, например, наибольшие и наименьшие значения и их распределения как случайных величин. Сформулируем две теоремы, в которых даны условия для существования непрерывных реализаций.

Теорема 11. *Пусть при некоторых* $r, \alpha > 0$

$$(\forall s,t)\ \mathbf{M}|x(t)-x(s)|^r \ \le\ c|t-s|^{1+\alpha},$$

тогда найдется модификация $\xi(t)$ *случайного процесса* $x(t)$*, обладающая непрерывными реализациями.*

Теорема 12. *Пусть нашлись такие монотонно возрастающие при положительных значениях аргументов функции* g, q*, что*

$$(\forall t, h>0)\ \mathbf{P}(|x(t+h)-x(t)| \ge g(h)) \ \le\ q(h),$$

и, вдобавок,

$$\sum_{n=0}^{\infty} 2^n q\left(\frac{b-a}{2^n}\right) \ <\ \infty, \quad \sum_{n=0}^{\infty} g\left(\frac{b-a}{2^n}\right) \ <\ \infty.$$

Тогда найдется модификация $\xi(t)$ *случайного процесса* $x(t)$*, обладающая непрерывными реализациями.*

Докажем, что теорема 11 следует из теоремы 12 . Действительно, из условий первой теоремы следует, что можно найти такое $\gamma > 0$, что $\alpha - r\gamma > 0$. Положим

$$g(h) \ =\ h^{\gamma}, \quad q(h) \ =\ ch^{1+\alpha-r\gamma},$$

где постоянная c также из условия теоремы 11 . Из неравенства Чебышева с r-м моментом следует, что

$$\mathbf{P}(|x(t+h)-x(t)| \ge g(h)) \ \le\ \frac{ch^{1+\alpha}}{g^r(h)} \ =\ q(h).$$

Осталось проверить сходимость рядов в условии теоремы 12 , что сделать совсем несложно.

Для доказательства теоремы 12 нам понадобится следующий вспомогательный результат, оказывающийся полезным и во многих других задачах теории вероятностей, особенно связанных со сходимостью почти наверное.

Лемма 7 (Борель - Кантелли). *Пусть для событий* $A_1, A_2, \ldots$ *справедливо*

$$\sum_{n=1}^{\infty} \mathbf{P}(A_n) \ <\ \infty.$$

Тогда с вероятностью 1 одновременно может произойти лишь конечное число из этих событий.

Доказательство. Используя интерпретацию кванторов существования и всеобщности через теоретико-множественные операции, заметим, что если произошло бесконечное число из событий, то произошло

$$A \;=\; \cap_{k=1}^{\infty} \cup_{n=k}^{\infty} A_n.$$

Введем в рассмотрение монотонно убывающую последовательность событий $B_k \;=\; \cup_{n\geq k} A_n$. При этом

$$\mathbf{P}(B_k) \;\leq\; \sum_{n=k}^{\infty} \mathbf{P}(A_n) \;\longrightarrow\; 0,\;\; k\to\infty.$$

Применяя аксиому непрерывности вероятности, получаем, что

$$\mathbf{P}(A) \;=\; \lim_{k\to\infty} \mathbf{P}(B_k) \;=\; 0.$$

Лемма доказана.

Доказательство теоремы 12. Для упрощения формул будем считать, что процессы заданы на [0,1]. Рассмотрим случайный процесс

$$X^{(n)}(t) \;=\; 2^n \left(x(t_{n,k})(t_{n,k+1}-t) + x(t_{n,k+1})(t-t_{n,k})\right),\;\; t\in[t_{n,k},\; t_{n,k+1}],$$

где $t_{n,k}=\frac{k}{2^n},\;\; k=0,...,2^n$. Заметим, что

$$(\forall n,k)\; X^{(n)}(t_{n,k}) \;=\; x(t_{n,k})$$

по построению. Максимальное значение модуля разности $|X^{(n)}(t) - X^{(n+1)}(t)|$ достигается в одной из точек вида $\frac{2k+1}{2^{n+1}}$. Привлекая тот факт, что вероятность объединения событий не превосходит суммы их вероятностей, видим, что

$$P_n \equiv \mathbf{P}\left(\max|X^{(n)}-X^{(n+1)}| \geq g(2^{-(n+1)})\right) \;\leq\; 2^{n+1}q(2^{-(n+1)}).$$

Из этого неравенства и условий теоремы следует, что ряд $\sum_{n=1}^{\infty} P_n$ сходится. Итак, привлекая лемму, видим, что произошло лишь конечное число событий вида $\{\max|X^{(n)}-X^{(n+1)}| \geq g(2^{-(n+1)})\}$, т.е., начиная с некоторого, конечного с вероятностью 1, N, эти события не происходят, что, в силу сходимости ряда $\sum_{n=1}^{\infty} g(2^{-(n+1)})$ означает равномерную сходимость ряда $\sum_{n=1}^{\infty}\left(X^{(n)}(t)-X^{(n+1)}(t)\right)$. Из соотношения

$$X^{(n)}(t) \;=\; X^{(1)}(t) - \sum_{j=1}^{n-1}\left(X^{(n)}(t)-X^{(n+1)}(t)\right)$$

следует, что существует непрерывный

$$\hat{X}(t) \;=\; \lim_{n\to\infty} X^{(n)}(t).$$

Для завершения доказательства осталось положить

$$\xi(t) \;=\; \begin{cases}\hat{X}(t), & N<\infty\;, \\ 0, & N=\infty.\end{cases}$$

8. Условные математические ожидания

Изложение этой темы можно найти во многих учебниках, но с целью большей полноты собранного в этом конспекте материала мы приведем некоторые свойства и определения здесь. Напомним, что условной вероятностью события A при условии B, если $\mathbf{P}(B) \neq 0$ называется

$$\mathbf{P}(A/B) \;=\; \frac{\mathbf{P}(AB)}{\mathbf{P}(B)}.$$

Определим для таких B и случайных величин ξ

$$\mathbf{M}(\xi/B) \;=\; \int_{\Omega} \xi d\mathbf{P}(\cdot/B).$$

Тогда по определению интеграла по вероятности для произвольных событий A, B и случайной величины ξ

$$\int_B \mathbf{M}(\xi/B) d\mathbf{P} \;=\; \mathbf{M}(\xi/B)\cdot\mathbf{P}(B) \;=\; \int_B \xi d\mathbf{P}.$$

Пусть $< \Omega, \mathcal{F}, \mathbf{P} >$ – вероятностное пространство, $\mathcal{U} \;=\; \{B_1, B_2, ...\}$ – разбиение Ω, т.е.

$$\cup_j B_j = \Omega,\; i \neq j \Rightarrow B_i \cap B_j = \emptyset,\; (\forall j)\, B_j \in \mathcal{F},\; \mathbf{P}(B_j) \neq 0.$$

Случайная величина

$$\hat{\xi} \;=\; \sum_j \mathbf{M}(\xi/B_j) 1_{B_j}$$

называется условным математическим ожиданием ξ при условии $\mathcal{U}$ и обозначается $\mathbf{M}(\xi/\mathcal{U})$.

Пусть $B \in \sigma(\mathcal{U})$. Тогда по определению сигма-алгебры можно выделить такую подцепочку из разбиения, что $B \;=\; \cup_k B_{i_k} \;=\; \cup_k C_k$. Отметим, что

$$\int_B \xi d\mathbf{P} \;=\; \sum_k \int_{C_k} \xi d\mathbf{P},$$

иначе говоря, для произвольного $B \in \sigma(\mathcal{U})$ справедливо

$$\int_B \xi d\mathbf{P} \;=\; \int_B \mathbf{M}(\xi/\mathcal{U}) d\mathbf{P}$$

(на B все "лишние" индикаторы обращаются в 0). Отметим, что если η – $\sigma(\mathcal{U})$-измеримая случайная величина и

$$(\forall B \in \sigma(\mathcal{U}))\; \int_B \eta d\mathbf{P} \;=\; \int_b \mathbf{M}(\xi/\mathcal{U}) d\mathbf{P},$$

то $\eta = \mathbf{M}(\xi/\mathcal{U})\,(\mathrm{mod}\mathbf{P})$. Дадим теперь определение.

Пусть $\mathcal{G}$ – σ-алгебра. $\mathcal{G}$-измеримая случайная величина $\mathbf{M}(\xi/\mathcal{G})$ называется *условным математическим ожиданием* случайной величины ξ относительно этой σ-алгебры, если для произвольного $B \in \mathcal{G}$ справедливо

$$\int_B \xi d\mathbf{P} = \int_B \mathbf{M}(\xi/\mathcal{G}) d\mathbf{P}.$$

Равенство из определения с учетом определения скалярного произведения в L^2 можно переписать в виде

$$(\forall B \in \mathcal{G}) \ <\xi, 1_B>_{L^2(<\Omega,\mathcal{F},\mathbf{P}>)} = <\mathbf{M}(\xi/\mathcal{G}), 1_B>_{L^2(<\Omega,\mathcal{F},\mathbf{P}>)},$$

что можно интерпретировать, как равенство углов, образуемых случайной величиной и ее условным математическим ожиданием с порождающими $\mathcal{G}$. Существование условного математического математического ожидания (в этом контексте называемого иногда производной сигма-аддитивной функции по мере) следует из теоремы функционального анализа, формулируемой ниже.

Теорема 13 (Радон - Никодим). *Если $g(B)$ – σ-аддитивная функция на $\mathcal{F}$, μ – мера на $<\Omega,\mathcal{F}>$, то найдется единственная $mod\,\mathbf{P}$ неотрицательная μ-измеримая функция f такая, что*

$$(\forall B \in \mathcal{F}) \ g(B) = \int_B f(x) d\mu.$$

Перейдем теперь к формулировкам и доказательствам свойств условных математических ожиданий.

1) *Если ξ - $\mathcal{F}$-измерима, то* $\mathbf{M}(\xi/\mathcal{F}) = \xi$.

   Действительно,

$$(\forall B \in \mathcal{F}) \ \int_B \xi d\mathbf{P} = \int_B \mathbf{M}(\xi/\mathcal{F}) d\mathbf{P}$$

   по определению условного математического ожидания.

2) $\mathbf{M}\,\mathbf{M}(\xi/\mathcal{F}) = \mathbf{M}\xi$.

   Для доказательства этого свойства достаточно выбрать в определении $B = \Omega$:

$$\mathbf{M}\,\mathbf{M}(\xi/\mathcal{F}) = \int_\Omega \mathbf{M}(\xi/\mathcal{F}) d\mathbf{P} = \int_\Omega \xi d\mathbf{P} = \mathbf{M}\xi.$$

3) $\mathbf{M}(\alpha\xi + \beta\eta/\mathcal{F}) = \alpha\mathbf{M}(\xi/\mathcal{F}) + \beta\mathbf{M}(\eta/\mathcal{F})$.

Зафиксируем $B \in \mathcal{F}$. Заметим, что правая часть доказываемого равенства $\mathcal{F}$-измерима как комбинация измеримых величин. Отсюда, привлекая свойство линейности интеграла по вероятности, имеем

$$\begin{gathered} \textstyle\int_B (\alpha\mathbf{M}(\xi/\mathcal{F}) + \beta\mathbf{M}(\eta/\mathcal{F}))\, d\mathbf{P} \ = \\ = \ \alpha \textstyle\int_B \xi d\mathbf{P} + \beta \textstyle\int_B \eta d\mathbf{P} \ = \ \textstyle\int_B (\alpha\xi + \beta\eta) d\mathbf{P}, \end{gathered}$$

что и требовалось доказать.

4) *Если ξ не зависит от $\mathcal{F}$, то* $\mathbf{M}(\xi/\mathcal{F}) \ = \ \mathbf{M}\xi$.

Понятно, что $(\forall B \in \mathcal{F})(\forall x)\ \mathbf{P}(\xi < x,\ B) \ = \ \mathbf{P}(\xi < x)\mathbf{P}(B)$, т.е. ξ и 1_B независимы. Отсюда

$$\int_B \xi d\mathbf{P} \ = \ \mathbf{M}\xi 1_B \ = \ \mathbf{M}\xi\mathbf{M}1_B \ = \ \mathbf{M}\xi \int_B d\mathbf{P} = \int_B \mathbf{M}\xi d\mathbf{P},$$

что завершает доказательство.

5) *Если η - $\mathcal{F}$-измерима, то для произвольной ξ* $\mathbf{M}(\xi\eta/\mathcal{F}) \ = \ \eta\mathbf{M}(\xi/\mathcal{F})$.

Пусть $\eta = 1_C,\ C \in \mathcal{F}$. Тогда для $B \in \mathcal{F}$

$$\int_B \eta\mathbf{M}(\xi/\mathcal{F}) d\mathbf{P} \ = \ \int_{B\cap C} \mathbf{M}(\xi/\mathcal{F}) d\mathbf{P} \ = \ \int_{B\cap C} \xi d\mathbf{P} \ = \ \int_B \xi\eta d\mathbf{P},$$

т.е. в этой ситуации все доказано. Случай линейных комбинаций индикаторов очевиден, общий случай получается предельным переходом.

6) *Пусть $\mathcal{F} \subset \mathcal{F}_1$ – сигма-алгебры. Тогда* $\mathbf{M}\left(\mathbf{M}(\xi/\mathcal{F}_1)/\mathcal{F}\right) \ = \ \mathbf{M}(\xi/\mathcal{F})$.

Действительно, при $B \in \mathcal{F}$ справедливо

$$\int_B \mathbf{M}\left(\mathbf{M}(\xi/\mathcal{F}_1)/\mathcal{F}\right) d\mathbf{P} = \int_B \mathbf{M}(\xi/\mathcal{F}_1) d\mathbf{P} = \int_B \xi d\mathbf{P} = \int_B \mathbf{M}(\xi/\mathcal{F}) d\mathbf{P},$$

что и заканчивает доказательство.

7) *Если $\mathcal{F} \subset \mathcal{F}_1,\ \eta$ – $\mathcal{F}_1$-измерима, то* $\mathbf{M}(\xi\eta/\mathcal{F}) \ = \ \mathbf{M}\left(\eta\mathbf{M}(\xi/\mathcal{F}_1)/\mathcal{F}\right)$.

Доказательство следует из свойств 6 и 5.

9. Мартингалы

9.1. Определения

В последнем разделе курса рассмотрим процессы, устроенные более сложно, чем процессы с независимыми приращениями. Будем считать, что все рассматриваемые ниже процессы заданы на полуинтервале $T = [a,b)$, его замыкании или каком-либо дискретном подмножестве этого полуинтервала.

Будем говорить, что семейство $\sigma-$алгебр $\mathcal{F}_t, t \in T$ образует *поток*, если при $s < t$ справедливо $\mathcal{F}_s \subset \mathcal{F}_t$. Поток будем называть непрерывным справа в точке a, если $\sigma-$алгебры $\mathcal{F}_a$ и $\mathcal{F}_{a+} = \cap\{\mathcal{F}_t, t > a\}$ совпадают. Заметим, что $\sigma-$алгебры $\mathcal{F}_{t+},\ t \in T$ образуют поток, непрерывный справа в любой точке T.

Условимся тот факт, что при каждом t случайная величина $\xi(t)$ $\mathcal{F}_t-$измерима, называть *согласованностью случайного процесса* $\xi(t)$ *с потоком.* Очевидно, что любой случайный процесс $\xi(t)$ согласован с потоком $\mathcal{F}_t^\xi = \sigma(\xi(s),\ s \le t)$.

Случайный процесс называется *мартингалом* относительно потока сигма-алгебр, если он согласован с этим потоком, имеет конечное математическое ожидание, и

$$(\forall t, s \in T)\ (s < t)\ \Rightarrow\ \mathbf{M}\left(\xi(t)/\mathcal{F}_s\right) = \xi(s)$$

с вероятностью 1. Если $\mathbf{M}\left(\xi(t)/\mathcal{F}_s\right) \ge \xi(s)$ для произвольных пар аргументов с условием $s < t$, то процесс называется *субмартингалом*, если же в последнем определении знак $\ge$ изменен на $\le$, то процесс $\xi(t)$ называют *супермартингалом.* Обобщающим понятием служит *полумартингал* – это или субмартингал, или супермартингал.

Приведем здесь три примера мартингалов.

1) Пусть $\xi(t)$ – процесс с независимыми приращениями и постоянным математическим ожиданием. Тогда это – мартингал относительно $\mathcal{F}_t^\xi$ $t \in T$. Действительно, при $s < t$

$$\mathbf{M}(\xi(t)/\mathcal{F}_s^\xi) \ = \ \mathbf{M}((\xi(t) - \xi(s))/\mathcal{F}_s^\xi) \ + \ \mathbf{M}(\xi(s)/\mathcal{F}_t^\xi),$$

   но, согласно свойствам измеримости, первое слагаемое здесь равно $\mathbf{M}(\xi(t) - \xi(s)) \ = \ 0$, а второе – $\xi(s)$.

2) Пусть ξ – случайная величина, и задан поток $\sigma-$алгебр $\mathcal{F}_t,\ t \in T$. Тогда $X(t) \ = \ \mathbf{M}(\xi/\mathcal{F}_t)$ – мартингал. Это следует из следующей цепочки равенств $(s < t)$:

$$\mathbf{M}(X(t)/\mathcal{F}_s) \ = \ \mathbf{M}(\mathbf{M}(\xi/\mathcal{F}_t)/\mathcal{F}_s) \ = \ \mathbf{M}(\xi/\mathcal{F}_s) \ = \ X(s).$$

3) Последовательность сумм независимых одинаково распределенных случайных величин также образует мартингал (с дискретным временем). Это немедленно получается из рассуждений первого примера и уже отмечавшегося ранее того факта, что такие суммы образуют процесс с независимыми приращениями.

Очевидно, что определение субмартингала можно переписать в следующем виде

$$(\forall s,t)(\forall B \in \mathcal{F}_s)\ (s < t) \Rightarrow \int_B \xi(s)d\mathbf{P} \le \int_B \xi(t)d\mathbf{P}. \qquad (18)$$

Теорема 14. *Пусть $\xi(t)$ – мартингал, $f(\cdot)$ – измеримая функция, выпуклая вниз и такая, что $\mathbf{M}f(\xi(t)) < \infty$. Тогда $f(\xi(t))$ – субмартингал.*

Доказательство. Согласованность с потоком $\mathcal{F}_t, t \in T$ очевидна. Проверим условие (18) для $f(\xi(t))$. Из выпуклости вниз следует существование в каждой точке графика функции опорной прямой, т.е.

$$(\forall x_0)(\exists C)(\forall x)\ f(x)\ \ge\ f(x_0) + C(x_0)\cdot(x - x_0).$$

Фиксируем $N \in \mathbf{N}, B \in \mathcal{F}_s$ и введем $B_N\ =\ B \cap \{|C(\xi(s))| \le N\}$. Тогда

$$\int_{B_N} f(\xi(t))d\mathbf{P}\ \ge \int_{B_N} f(\xi(s))d\mathbf{P} + \int_{B_N} C(\xi(s))(\xi(t) - \xi(s))d\mathbf{P}. \qquad (19)$$

Заметим, что $B_N \in \mathcal{F}_s$ и

$$\mathbf{M}(C(\xi(s)\cdot(\xi(t) - \xi(s))/\mathcal{F}_s)\ =\ C(\xi(s))\mathbf{M}((\xi(t) - \xi(s))/\mathcal{F}_s)\ =\ 0,$$

откуда следует, что второй интеграл в (19) равен 0. Таким образом,

$$\int_{B_N} f(\xi(t))d\mathbf{P}\ \ge\ \int_{B_N} f(\xi(s))d\mathbf{P}.$$

Доказательство теоремы немедленно следует из этого неравенства предельным переходом при $N \to \infty$.

Простым следствием теоремы является возможность строить многочисленные примеры субмартингалов. Например, если $\xi(t)$ – мартингал, то $|\xi(t)|$, $\xi^2(t)$, $e^{\xi(t)}$ – субмартингалы. Сколько угодно супермартингалов можно построить, если иметь ввиду вариант доказанной теоремы с функцией, выпуклой вверх.

Некоторые полезные факты о полумартингалах можно найти в книгах Вентцеля и Гихмана - Скорохода (см. список литературы). Там, в частности, доказано, что почти все реализации полумартингалов не имеют разрывов второго рода. Нам потребуются следующие два неравенства (частный случай теоремы Дуба для субмартингалов). Пусть

$$\eta_b\ =\ \sup_{a \le t \le b} \xi(t).$$

Теорема 15. *Непрерывный справа неотрицательный субмартингал $\xi(t)$ удовлетворяет неравенствам*

$$\mathbf{P}(\eta_b \ge b)\ \le\ \frac{\mathbf{M}\xi(b)}{x},$$

$$\mathbf{M}\eta_b^p\ \le\ \left(\frac{p}{p-1}\right)^p \mathbf{M}\xi^p(b),\ p > 1.$$

9.2. Полнота гильбертова пространства мартингалов

Пусть $< \Omega, \mathcal{F}, \mathbf{P} >$ – вероятностное пространство, $\mathcal{F}_t$, $t \in T$ – непрерывный в каждой точке справа поток $\sigma-$алгебр. Предположим также, что все $\sigma-$алгебры, входящие в поток, полны относительно $\mathbf{P}$, т.е. содержат все возможные подмножества своих событий нулевой вероятности. Введем в рассмотрение класс $\mathcal{M}(T)$ всех мартингалов относительно этого потока, у которых реализации с вероятностью 1 не имеют разрывов второго рода и

$$\mathbf{M}\xi^2(t) \ < \ \infty, \ t \in T.$$

Пусть $\mathcal{M}_C(T)$ – подкласс мартингалов из $\mathcal{M}(T)$, обладающих непрерывными реализациями с вероятностью 1.

Все значения мартингала восстанавливаются по $\xi(b)$. Поэтому естественно ввести в $\mathcal{M}(T)$ скалярное произведение по формуле

$$< \xi, \ \eta >_T \ = \ \mathbf{M}\xi(b)\eta(b).$$

Теорема 16. *$\mathcal{M}(T)$ с введенным скалярным произведением является гильбертовым пространством, а $\mathcal{M}_C(T)$ – замкнутое подпространство этого пространства.*

Доказательство. Пусть $\xi_n(t)$ – фундаментальная последовательность в $\mathcal{M}(T)$. Тогда последовательность $\xi_n(b)$ в пространстве $L^2(< \Omega, \mathcal{F}, \mathbf{P} >)$ сходится к некоторой случайной величине μ. Положим

$$\mu(t) \ = \ \mathbf{M}(\mu/\mathcal{F}_t).$$

Этот случайный процесс является мартингалом, ограничения, наложенные на поток $\mathcal{F}_t$, $t \in T$ в начале главы позволяют выбрать модификацию этого случайного процесса, не имеющую разрывов второго рода. Это доказывает первое утверждение теоремы.

Как мы знаем, $|\xi_n(t) - \mu(t)|$ – субмартингал. Из неравенств теоремы 15 следует

$$\mathbf{M}\left(\sup_t |\xi_n(t) - \mu(t)|\right)^2 \ \leq \ 4\mathbf{M}(\xi_n(b) - \mu(b))^2 \ \to \ 0.$$

Можно так построить подпоследовательность $n(k)$, что

$$\sum_{k=1}^{\infty} \mathbf{M} \sup |\xi_{n(k)}(t) - \mu(t)| \ = \ \mathbf{M} \sum_{k=1}^{\infty} \sup |\xi_{n(k)} - \mu(t)| \ < \ \infty.$$

Используя критерий сходимости почти наверное, видим, что с вероятностью 1 реализации $\xi_{n(k)}$ сходятся к реализациям $\mu(t)$. Следовательно, если первоначально выбирались мартингалы с непрерывными реализациями, то и почти все реализации μ непрерывны. Если заменить все разрывные реализации предельного мартингала непрерывными функциями (а это не нарушает его согласованности с потоком, так как все $\sigma-$алгебры потока полны), то получается непрерывная модификация процесса $\mu(t)$. Теорема доказана.

9.3. Стохастический интеграл по мартингалу

Пусть, как и раньше, у нас имеется поток $\sigma-$алгебр и случайный процесс. Процесс $\xi(t) : [a,b] \to \mathbf{R}$ называем *прогрессивно измеримым* относительно потока $\mathcal{F}_t$, $t \in [a,b]$, если при любом t сужение его на множество $[a,t]$ измеримо относительно $\sigma(\mathcal{F}_t \times \mathcal{B}[a,t])$.

Прогрессивно измеримый процесс обладает тем свойством, что для любого t случайная величина $\xi(t)$ измерима относительно $\mathcal{F}_t$. Докажем утверждение, являющееся в некотором смысле обратным к этому.

Теорема 17. *Пусть $(\forall t)$ $\xi(t)$ – случайная величина, измеримая относительно $\mathcal{F}_t$. Если процесс $\xi(t)$ имеет непрерывные справа (или слева) реализации, то он прогрессивно измерим относительно потока $\mathcal{F}_t$, $t \in [a,b]$.*

Доказательство. Докажем теорему для случая, когда реализации непрерывны справа. Выберем последовательность разбиений нашего отрезка

$$a = t_0 \ < \ t_1(n) \ < \ldots < t_n(n) \ = \ b$$

так, что максимальная длина отрезка разбиения стремится к нулю при $n \to \infty$. Пусть

$$\xi_n(t) \ = \ \sum_{j=1}^{n} \xi(t_j(n)) 1_{\{t_{j-1}(n) \le t < t_j(n)\}}.$$

Тогда при $n \to \infty$ и любых t, ω выполнено $\xi_n(t) \to \xi(t)$. Поскольку все функции $\xi_n(t,\omega)$ измеримы относительно $\sigma(\mathcal{F}_t \times \mathcal{B}[a,t])$, то это же относится и к их пределу. Теорема доказана.

Пусть $\mu(t)$, $t \in [0,b]$ – квадратично интегрируемый мартингал. Будем говорить, что от имеет *абсолютно непрерывную характеристику* A, если существует прогрессивно измеримый неотрицательный процесс $a(t)$, такой, что $\mu^2(t) - A(t)$ – мартингал относительно потока $\mathcal{F}_t$, $t \in T$, где

$$A(t) \ = \ \int_0^t a(s)\, ds.$$

Будем рассматривать квадратично интегрируемый мартингал $\mu(t)$ с абсолютно непрерывной характеристикой относительно непрерывного справа потока $\sigma-$алгебр, составляющие которого полны. Дополнительно потребуем, чтобы почти все реализации этого мартингала были бы непрерывны. Легко видеть, что при $t > s$

$$\mathbf{M}((\mu(t)-\mu(s))^2/\mathcal{F}_s) \; = \; \mathbf{M}(\mu^2(t)/\mathcal{F}_s)-\mu^2(s) \; = \; \mathbf{M}((A(t)-A(s))/\mathcal{F}_s).$$

Сначала определим интеграл по мартингалу для простых случайных процессов. Разобьем наш отрезок неслучайными точками

$$0 = t(0) \; < \; t(1) \; < ... < t(N) = b$$

и рассмотрим простой случайный процесс

$$\eta(t) \; = \; \sum_{j=1}^{N} \eta_j 1_{\{t(j-1)<t\le t(j)\}},$$

где случайные величины η_j измеримы относительно $\mathcal{F}_{t(j)},\; j = 1,...,N$ и ограничены с вероятностью 1. Тогда, по определению,

$$\eta\circ\mu(t) \; \equiv \; \int_0^t \eta(t)d\mu(t) \; =$$
$$= \; \sum_{k\le m(t)} \eta_k(\mu(t(k+1))-\mu(t(k))) \; + \; \eta_{m(t)}(\mu(t)-\mu(t(m(t))),$$

где $m(t) = \max\{k|\; t(k) \le t\}$.

Случайный процесс $\eta\circ\mu(t)$ имеет непрерывные реализации и является квадратично интегрируемым мартингалом того же типа, что и μ. Действительно, если $t(k) \ge s$, то

$$\mathbf{M}(\eta_k(\mu(t(k+1))-\mu(t(k)))/\mathcal{F}_s) \; =$$
$$= \; \mathbf{M}(\mathbf{M}(\eta_k(\mu(t(k+1))-\mu(t(k)))/\mathcal{F}_{t(k)})/\mathcal{F}_s) \; =$$
$$= \; \mathbf{M}(\eta_k\mathbf{M}((\mu(t(k+1))-\mu(t(k)))/\mathcal{F}_{t(k)})/\mathcal{F}_s) \; = \; 0.$$

Аналогичным образом можно проверить, что при $t(k) < s < t$

$$\mathbf{M}(\eta_k(\mu(t(k+1))-\mu(t(k)))/\mathcal{F}_s) \; = \; \eta_k(\mu(s)-\mu(t(k)).$$

Поэтому из определения интеграла от простого процесса следует равенство

$$\mathbf{M}(\eta\circ\mu(t)/\mathcal{F}_s) \; = \; \eta\circ\mu(s),\; s < t.$$

Пусть $s < t < u < v$, η измерима относительно $\mathcal{F}_s$, $\bar\eta$ – относительно $\mathcal{F}_u$, тогда

$$\mathbf{M}(\eta(\mu(t)-\mu(s))\bar\eta(\mu(v)-\mu(u))/\mathcal{F}_s) \; =$$
$$= \; \mathbf{M}(\mathbf{M}(\eta(\mu(t)-\mu(s))\bar\eta(\mu(v)-\mu(u))/\mathcal{F}_u)/\mathcal{F}_s) \; =$$
$$\mathbf{M}(\eta(\mu(t)-\mu(s))\bar\eta\mathbf{M}((\mu(v)-\mu(u))/\mathcal{F}_u)/\mathcal{F}_s) \; = \; 0.$$

Из этой формулы следует, что

$$\mathbf{M}((\eta \circ \mu(t))^2/\mathcal{F}_s) - (\eta \circ \mu(s))^2 \;=\; \mathbf{M}\left(\int_s^t \eta^2(u) a(u)\, du/\mathcal{F}_s\right),$$

т.е., что мартингал $\eta \circ \mu$ также обладает абсолютно непрерывной характеристикой.

Если реализации процесса $\eta(t)$ непрерывны слева и ограничены неслучайной постоянной, то интеграл от него можно определить с помощью предельного перехода. Точнее, пусть

$$\eta_n(t) \;=\; \sum_{k=0}^{n-1} \eta(\frac{bk}{n}) 1_{\{bk/n \;<\; t \;<\; b(k+1)/n\}},$$

тогда, по определению,

$$\eta \circ \mu(t) \;=\; \lim_{n\to\infty} \eta_n \circ \mu(t).$$

Теорема 16 позволяет понимать этот предельный переход как равномерную по t сходимость почти всех реализаций $\eta_n \circ \mu(t)$. Действительно,

$$\begin{aligned} \mathbf{M}(\eta_n \circ \mu(b) - \eta_m \circ \mu(b))^2 \;&=\; \mathbf{M}((\eta_n - \eta_m) \circ \mu(b))^2 \;= \\ =\; \mathbf{M} \textstyle\int_0^b (\eta_n(m) - \eta_m(u))^2 a(u)\, du \;&\longrightarrow\; 0. \end{aligned}$$

Аналогичным образом, с помощью предельного перехода, определяется стохастический интеграл по мартингалу для всех процессов $\eta(t)$, которые являются пределами последовательностей непрерывных слева ограниченных прогрессивно измеримых процессов в том смысле, что

$$\mathbf{M}\int_0^b (\eta_n(m) - \eta_m(u))^2 a(u)\, du \;\longrightarrow\; 0.$$

Класс таких процессов совпадает с классом прогрессивно измеримых процессов, для которых

$$\mathbf{M}\int_0^b \eta^2(t) a(t)\, dt \;<\; \infty.$$

При этом полагают

$$\eta \circ \mu(t) \;=\; \lim_{n\to\infty} \lim_{u\to\infty} \eta_{n,u} \circ \mu(t),$$

где

$$\eta_{n,u}(t) \;=\; \int_0^t \max\{\eta(s), u\} \exp\{-(s-t)n\} n\, ds.$$

10. Стохастические дифференциалы и интегралы Ито

10.1. Стохастические интегралы Ито

Пусть выполнены все предположения предыдущего раздела относительно потока σ-алгебр и прогрессивной измеримости процесса $\xi(t)$. Но теперь вместо произвольного мартингала $\mu(t)$ рассмотрим винеровский процесс $w(t)$. Для него соответствующая элементарная стохастическая мера

$$\mu([a,b]) \ = \ w(b) - w(a)$$

обладает тем свойством, что для непересекающихся отрезков Δ_1 и Δ_2 случайные величины $\mu(\Delta_1)$ и $\mu(\Delta_2)$ независимы, а $\xi(a)$ не зависит от $\mu([a,b])$ в силу сделанных предположений. Интеграл

$$\int\limits_a^b \xi(t)\, dw(t)$$

называют *стохастическим интегралом Ито*. Очевидно, что для непрерывного в среднем квадратическом процесса $\xi(t)$ справедливо

$$\int\limits_a^t \xi(s) dw(s) \ = \ \text{l.i.m.}_{n\to\infty} \sum_{k=1}^n \xi(t_{k-1})(w(t_k) - w(t_{k-1})), \qquad (20)$$

где $t_k, k = 0, ..., n$ – произвольное разбиение отрезка $[a,t]$, выбранное так, что

$$\Delta = \max_k (t_k - t_{k-1}) \ \to 0, \ n \to \infty.$$

Нетрудно также проверить, что при $\xi(t) \neq 0$ и достаточно малых h в среднеквадратическом смысле

$$\int\limits_t^{t+h} \xi(s) dw(s) = \xi(t)(w(t+h) - w(t)) + o(\sqrt{h}).$$

Вычислим, например, $\int_0^t w(s) dw(s)$ по формуле (20). Используем равенство

$$2w(t_{k-1})(w(t_k) - w(t_{k-1})) = (w^2(t_k) - w^2(t_{k-1})) - (w(t_k) - w(t_{k-1}))^2.$$

Тогда

$$\sum_{k=1}^n w(t_{k-1})(w(t_k) - w(t_{k-1})) \ = \ \frac{1}{2} w^2(t) - \frac{1}{2} \sum_{k=1}^n (w(t_k) - w(t_{k-1}))^2. \qquad (21)$$

Докажем, что

$$\text{l.i.m.}_{n\to\infty}\sum_{k=1}^{n}(w(t_k)-w(t_{k-1}))^2=t. \qquad (22)$$

Обозначим

$$\delta_k=(w(t_k)-w(t_{k-1}))^2-(t_k-t_{k-1}),\ k=1,...,n.$$

Тогда

$$\sum_{k=1}^{n}(w(t_k)-w(t_{k-1}))^2-t\ =\ \sum_{k=1}^{n}\delta_k,$$

причем случайные величины δ_k независимы, $\mathbf{M}\delta_k=0$, и

$$\mathbf{D}\delta_k=\mathbf{M}(w(t_k)-w(t_{k-1}))^4-(t_k-t_{k-1})^2=2(t_k-t_{k-1})^2,$$

что следует из того, что для случайной величины ξ, имеющей нормальное распределение с параметрами 0 и σ^2 справедливо $\mathbf{M}\xi^4=3\sigma^4$. Отсюда

$$\mathbf{M}(\sum_{k=1}^{n}\delta_k)^2=2\sum_{k=1}^{n}(t_k-t_{k-1})^2\leq 2\Delta\sum_{k=1}^{n}(t_k-t_{k-1})=2\Delta t\to 0,$$

т.е. (22) доказано. Окончательно, из (21) и (22),

$$\int_0^t w(s)dw(s)=\frac{1}{2}(w^2(t)-t^2). \qquad (23)$$

10.2. Стохастический дифференциал

Говорят, что случайный процесс $\xi(t)$ имеет *стохастический дифференциал*

$$d\xi(t)\ =\ a(t)dt+b(t)dw(t), \qquad (24)$$

где $a(t)=a(t,\xi(t))$, $b(t)=b(t,\xi(t))$ – прогрессивно измеримые случайные процессы, если

$$\xi(t)=\xi(t_0)+\int_{t_0}^{t}a(s)ds+\int_{t_0}^{t}b(s)dw(s),\quad t\geq t_0, \qquad (25)$$

а выписанные интегралы существуют.

Поскольку

$$\int_s^{s+h}\sqrt{\mathbf{M}a^2(t)}\,dt\to 0,\quad \int_s^{s+h}\sqrt{\mathbf{M}b^2(t)}\,dt\to 0$$

при $h \to 0$ и произвольном s, то нетрудно проверить, что процесс, имеющий стохастический дифференциал, непрерывен в среднем квадратическом.

Оказывается, процессы $a(t, \xi(t))$ и $b(t, \xi(t))$, фигурирующие в (24), могут быть интерпретированы следующим образом. Если представить себе, что $\xi(t),\ t \geq t_0$ описывает движение некоторой частицы, то при условии, что $\xi(s) = x$, за малое время h математическое ожидание смещения этой частицы будет равно $a(s,x)h + o(h)$, а само это смещение с точностью до $o(\sqrt{h})$ будет случайной величиной, имеющей нормальное распределение с дисперсией $b(s,x)$. Поэтому $a(t)$ иногда называют *коэффициентом сноса*, а $b(t)$ – *коэффициентом диффузии.* Процесс, имеющий стохастический дифференциал вида (24), называют *диффузионным.* Уравнение (24) относительно $\xi(t)$ называют *стохастическим дифференциальным уравнением.* Простое изложение основных моментов теории стохастических дифференциальных уравнений можно найти в [8].

В заключение покажем, что

$$dw^2(t) \ = \ dt + 2w(t)dt.$$

Действительно, из (23),

$$w^2(t) = w(0) + \frac{1}{2}t^2 + 2\int\limits_0^t w(s)dw(s),$$

что совпадает с (25).

Литература

[1] Боровков А.А. Теория вероятностей. М.:Наука, 1986

[2] Вентцель А.Д. Курс теории случайных процессов. М.: Наука, 1975

[3] Гихман И.И., Скороход А.В. Введение в теорию случайных процессов. М.: Наука, 1977

[4] Гихман И.И., Скороход А.В. Теория случайных процессов в 3 томах.

[5] Гнеденко Б.В. Курс теории вероятностей. - М. Наука, 1969 и более поздние издания

[6] Карлин С. Основы теории случайных процессов. М.: Мир, 1971

[7] Основы теории случайных процессов - интернет-методичка и список задач. На сайте МФ АГУ www.math.dcn-asu.ru

[8] Розанов Ю.А. Введение в теорию случайных процессов. М.: Наука, 1982