Задания по курсу «Случайные процессы»

1 Выявление сезонной компоненты случайного процесса по выборочной траектории методом разложения в ряд Фурье

Описание и пояснения

В траекториях случайных процессов может существовать периодическая повторяемость отрезков возрастания и убывания процесса синусоидального вида. На практике такую повторяемость можно встретить, например, при анализе температуры воздуха за несколько лет. Температура воздуха каждый год в среднем выше летом и ниже зимой, и характер изменения этой средней температуры можно приблизительно описать графиком синуса с периодом в 1 год. Одновременно с этим есть и периодичность изменения температуры воздуха в течение суток. Получается, что с некоторой погрешностью температуру воздуха можно описывать как сумму синусоид с периодами в год и в 1 день, а также случайной компоненты, обладающей меньшей амплитудой изменений, чем исходный процесс.

Для выявления периодических компонент по выборочной траектории случайного процесса в качестве базового метода используют разложение выборочной траектории (как функции от дискретного аргумента) в ряд Фурье и построение периодограммы для наглядного определения слагаемых-гармоник ряда, обладающих существенно большими амплитудами.

Пусть $x_1, \ldots, x_n$ — выборочные значения траектории процесса $X(t)$ на равномерной сетке моментов времени $t_1, \ldots, t_n$. Разложение в ряд Фурье функции $x(i) = x_i,\ i \in \{1, \ldots, n\}$ в общем случае имеет вид:

$$x(i) = x_i = \sum_{j=1}^{n} \left(\alpha_j \cos\left(2\pi \frac{j}{n} i\right) + \beta_j \sin\left(2\pi \frac{j}{n} i\right) \right).$$

Однако, для вещественной функции $x(i)$ у гармоник j и $n-j$ будут одинаковые амплитуды. В силу этого при поиске сезонных компонент обычно откидывают гармоники с $j > n/2$ (у этих гармоник, кроме того, периоды <2) и используют представление:

$$x(i) = x_i \approx \sum_{j=1}^{[n/2]} \left(\alpha_j \cos\left(2\pi \frac{j}{n} i\right) + \beta_j \sin\left(2\pi \frac{j}{n} i\right) \right), \tag{1}$$

где $[\cdot]$ обозначает операцию взятия целой части числа, а

$$\alpha_j = \frac{2}{n} \sum_{i=1}^{n} x_i \cos\left(2\pi \frac{j}{n} i\right), \quad \beta_j = \frac{2}{n} \sum_{i=1}^{n} x_i \sin\left(2\pi \frac{j}{n} i\right).$$

Заметим, что n/j равняется периоду j-го слагаемого/гармоники.

Далее строится график периодограммы, задаваемой как:

$$P(j) = \alpha_j^2 + \beta_j^2.$$

Значение $P(j)$ показывает амплитуду гармоники с периодом n/j. По графику / значениям $P(j)$ нужно выявить одну или несколько гармоник, составляющих большую часть суммарной амплитуды. Индексы выбранных гармоник обозначим через J. В дальнейшем в качестве сезонной компоненты $Seas(t)$ берётся представление (1), в котором выбираются только слагаемые с наибольшими амплитудами согласно графику периодограммы:

$$Seas(t) = \sum_{j \in J} \left(\alpha_j \cos\left(2\pi \frac{j}{n} t\right) + \beta_j \sin\left(2\pi \frac{j}{n} t\right) \right).$$

Здесь $t \in T$ можно уже воспринимать как непрерывный аргумент.

Задание

Для данной выборочной траектории:

1. Постройте периодограмму и определите по ней две самые существенные сезонные компоненты
2. Предоставьте характеристики найденных сезонных компонент: период и доля её амплитуды к сумме амплитуд всех (кроме $j = n$) гармоник
3. Нарисуйте сравнительный график исходной выборочной траектории и найденной сезонной компоненты

2 Занятие 1. Общая теория

Примеры

$\langle \mathbf{1} \rangle$ Пусть случайный процесс $\xi(t) = (t-\gamma)^2,\ t \in T = [0;1]$, где $\gamma \sim \mathbb{U}([0;1])$. Задания:

1. Опишите множество всех траекторий
2. Вычислите $\mathbf{P}(\max_t \xi(t) \geqslant 1/2)$ и $\mathbf{P}(\max_t \xi(t) \geqslant 1/2,\ \xi(1) < 0.9)$
3. Вычислите $\xi(1/2),\ \mathbf{E}\,\xi(1/2),\ \mathrm{cov}(\xi(0), \xi(1))$
4. Опишите семейство конечномерных функций распределения
5. $\tau = \arg\min_t \xi(t)$

$\langle \mathbf{2} \rangle$ Пусть случайный процесс описывается условиями:

1. $T = [0, \infty)$
2. $\xi(0) = 0$
3. $\xi(t_2) - \xi(t_1) = \sum_{i=1}^{3} \mathbb{I}(z_i \in [t_1; t_2])\gamma_i$, где $\gamma_1, \gamma_2, \gamma_3 \sim Bern(1/2)$ и независимы, а $z_1 = 1$, $z_2 = 2,\ z_3 = 4$

Опишите траектории процесса.

$\langle \mathbf{3} \rangle$ Пусть $\xi(t)$ — случайный процесс с дискретным временем $t \in T = \{1, 2, 3\}$ и дискретными возможными значениями $\{1, 2, 3\}$. Пусть распределение $\xi(t)$ задаётся набором вероятностей:

- $\mathbf{P}(\xi(0) = j) = A_{1j}$, где матрица $A = \begin{pmatrix} 1/2 & 0 & 1/2 \end{pmatrix}$
- $\mathbf{P}(\xi(1) = j \mid \xi(0) = i) = B_{ij}$, где матрица $B = \begin{pmatrix} 1/3 & 2/3 & 0 \\ - & - & - \\ 2/3 & 1/3 & 0 \end{pmatrix}$; прочерки в матрице добавлены, чтобы символизировать, что $\mathbf{P}(\xi(1) = j \mid \xi(0) = 2)$ не существуют, т.к. $\mathbf{P}(\xi(0) = 2) = 0$
- $\mathbf{P}(\xi(2) = j \mid \xi(1) = i,\ \xi(0) = k) = C_{ij}$, где $k \in \{1, 2, 3\}$, матрица $C = \begin{pmatrix} 0 & 1 & 0 \\ 0 & 0 & 1 \\ 1 & 0 & 0 \end{pmatrix}$

Задания:

1. Опишите траектории процесса.
2. Выпишите все множества, составляющие базисные множества порождённой процессом $\xi(t)$ минимальной цилиндрической алгебры. Например, выпишите все множества траекторий, соответствующие цилиндрическим множествам вида $\{\xi(t_1) \leqslant x_1\}$, вида $\{\xi(t_1) \leqslant x_1,\ \xi(t_2) \leqslant x_2\}$ и вида $\{\xi(1) \leqslant x_1,\ \xi(2) \leqslant x_2,\ \xi(3) \leqslant x_3\}$.

Задания

2.1 Пусть $\gamma \sim \mathbb{U}([0;1])$. Про случайный процесс $\xi(t),\ t \in T = [0,1]$ известно, что $\xi(0) = 0$, $\xi(t_2) - \xi(t_1) = t_2 - t_1$ для $\gamma \leqslant t_1 < t_2$; и $\xi(t_2) - \xi(t_1) = -(t_2 - t_1)$ для $t_1 < t_2 \leqslant \gamma$. Опишите класс всех траекторий $\mathcal{X}$ процесса $\xi(t)$. Укажите, как связаны $\min_{t \in T} \xi(t)$ и γ.

2.2 Пусть $\gamma \sim \mathbb{U}([0;1])$. Про случайный процесс $\xi(t),\ t \in T = [0,1]$ известно, что $\xi(0) = 0$, $\xi(t_2) - \xi(t_1) = \gamma$ для $\gamma \in (t_1; t_2]$; и $\xi(t_2) - \xi(t_1) = 0$ для $\gamma \notin (t_1; t_2]$. Опишите класс всех траекторий $\mathcal{X}$ процесса $\xi(t)$. Укажите, как связаны $\max_{t \in T} \xi(t)$ и γ.

2.3 Пусть $\gamma_1, \gamma_2 \sim Bern(1/2)$. Зададим случайный процесс как $\xi(t) = 1 + \sum_{i=1}^{t} \gamma_i$, $t \in T = \{0, 1, 2\}$. Выпишите все множества, составляющие базисные множества порождённой процессом $\xi(t)$ минимальной цилиндрической алгебры. Например, выпишите все множества траекторий, соответствующие цилиндрическим множествам вида $\{\xi(t_1) \leqslant x_1\}$, вида $\{\xi(t_1) \leqslant x_1,\ \xi(t_2) \leqslant x_2\}$ и вида $\{\xi(0) \leqslant x_1,\ \xi(1) \leqslant x_2,\ \xi(2) \leqslant x_3\}$.

2.4 Пусть распределение случайного процесса $\xi(t),\ t \in T = \{0, 1, 2\}$ задаётся набором вероятностей:

- $\mathbf{P}(\xi(0) = j) = A_{1j}$, где матрица $A = \begin{pmatrix} 1/2 & 1/2 & 0 \end{pmatrix}$
- $\mathbf{P}(\xi(1) = j \mid \xi(0) = i) = B_{ij}$, где матрица $B = \begin{pmatrix} 1/2 & 1/2 & 0 \\ 0 & 1/2 & 1/2 \\ - & - & - \end{pmatrix}$; прочерки в матрице добавлены, чтобы символизировать, что $\mathbf{P}(\xi(1) = j \mid \xi(0) = 3)$ не существуют, т.к. $\mathbf{P}(\xi(0) = 3) = 0$
- $\mathbf{P}(\xi(2) = j \mid \xi(1) = i,\ \xi(0) = k) = C_{ij}$, где $k \in \{1, 2, 3\}$, матрица $C = \begin{pmatrix} 1/2 & 0 & 1/2 \\ 1/2 & 1/2 & 0 \\ 0 & 1/2 & 1/2 \end{pmatrix}$

Выпишите или нарисуйте все траектории процесса. Вычислите $\mathbf{E}\,\xi(t_0)$, $\operatorname{cov}(\xi(t_1), \xi(t_2))$, где $t_0, t_1, t_2 \in T$.

2.5 Пусть $\gamma \sim \mathbb{U}([0;1])$. Пусть случайный процесс $\xi(t) = \gamma t,\ t \in T = [0, \infty)$. Опишите семейство конечномерных функций распределений случайного процесса $\xi(t)$.

2.6 Пусть случайный процесс $\xi(t) = \gamma + \zeta t,\ t \in T = [0, \infty)$, где $\gamma, \zeta \sim \mathbb{U}([0;1])$ и независимы. Найдите вероятность $\mathbf{P}(\xi(1) \geqslant 1,\ \xi(2) \leqslant 2)$.

2.7 Пусть $\gamma \sim \mathbb{U}([0;1])$ и случайный процесс $\xi(t) = \mathbb{I}(\gamma > t),\ t \in T = [0;1]$. Вычислите $\mathbf{E}\,\xi(t_0)$, $\operatorname{med}\xi(t_0)$, $\operatorname{cov}(\xi(t_1), \xi(t_2))$, где $t_0, t_1, t_2 \in T$.

2.8 Пусть случайный процесс $\xi(t) = \gamma t,\ t \in T = [0, \infty)$, где $\gamma \sim \mathbb{U}([0;1])$. Опишите распределение случайной величины $\tau = \inf\{t \colon \xi(t) \geqslant 1\}$.

2.9 Пусть случайный процесс

$$\xi(t) = \Big(\frac{t}{1+\gamma} - 1\Big)^2, \quad t \in T = [0, \infty),$$

где $\gamma \sim \mathbb{U}([0;1])$. Опишите распределение случайной величины $\tau = \arg\min_{t \in T} \xi(t)$.

3 Занятие 2. Типы непрерывности траекторий случайного процесса

Примеры

$\langle$**1**$\rangle$ Пусть $\gamma \sim Bern(1/2)$ и случайный процесс $\xi(t) = \mathbb{I}(t \leqslant 1)\gamma + \mathbb{I}(t > 1)(1-\gamma)$, $t \in T = [0;2]$. Покажите, что все траектории разрывны в точке $t = 1$, процесс не является стохастически непрерывным и не являетсянепрерывным в среднеквадратичном в точке $t = 1$.

$\langle$**2**$\rangle$ Пусть $\gamma \sim Bern(1/2)$ и $\zeta \sim \mathbb{U}([0;2])$ — независимы. Пусть случайный процесс $\xi(t) = \mathbb{I}(t \leqslant \zeta)\gamma + \mathbb{I}(t > \zeta)(1-\gamma)$, $t \in T = [0;2]$. Покажите, что все траектории разрывны, процесс является стохастически непрерывным и непрерывным в среднеквадратичном.

$\langle$**3**$\rangle$ Пусть $\xi(t)$, $t \in T = [0;+\infty)$ — случайный процесс с независимыми приращениями $\xi(t_2) - \xi(t_1) \sim \mathcal{N}(0, (t_2 - t_1))$, $t_1 \leqslant t_2$. Покажите, что у процесса $\xi(t)$ существует непрерывная модификация.

Задания

3.1 Пусть случайный процесс $\xi(t) = [t + \zeta]$, $t \in T = [0;\infty)$, где $\zeta \sim \mathbb{U}([0;1])$, а $[\cdot]$ обозначает операцию взятия целой части числа. С какой вероятностью у $\xi(t)$ будут непрерывные траектории? Докажите, что процесс $\xi(t)$ является стохастически непрерывным и непрерывным в среднеквадратичном. Будет ли процесс $\xi(t)$ эквивалентен процессу $\gamma(t) = t$?

3.2 Пусть про случайный процесс $\xi(t)$, $t \in T = [0,\infty)$ известно, что $\xi(0) = 0$ и для $t_1 \leqslant t_2$ выполняется $\xi(t_2) - \xi(t_1) \sim \mathbb{U}([0;(t_2 - t_1)])$. Покажите, что для процесса $\xi(t)$ существует непрерывная модификация (т.е. эквивалентный ему процесс, все траектории которого непрерывны).

3.3 Пусть случайный процесс $\xi(t) = t/\gamma$, $t \in T = [0;1]$, $\gamma \sim \mathbb{U}([0;1])$. Покажите, что у процесса $\xi(t)$ почти наверное непрерывные траектории, он стохастически непрерывен, но не является непрерывным в среднеквадратическом смысле.

3.4 Пусть $\gamma \sim \mathbb{U}([0;1])$. Для $T = [0;1]$ определим случайные процессы $\xi(t) = \mathbb{I}(t = \gamma)\gamma$ и $\zeta(t) = \mathbb{I}(t = 1/2)\gamma$. Покажите, что $\xi(t)$ стохастически непрерывен, а $\zeta(t)$ — нет. Покажите, что процессы $\xi(t)$ и $\zeta(t)$ не эквивалентны.

3.5 Пусть случайный процесс $\xi(t) = (\gamma - t + a)^{-1}$, $t \in T = [0;1]$, где $a \in \mathbb{R}$, $\gamma \sim \mathbb{U}([0;1])$. При каких a у процесса $\xi(t)$ почти наверное будут непрерывные траектории?

4 Занятие 3. Стационарность и независимость приращений

Примеры

$\langle \mathbf{1} \rangle$ Пусть $T = [0, \infty)$ и $\xi(t) = \{t + \gamma\}$, $t \in T$, где $\gamma \sim \mathbb{U}([0;1])$, а $\{\cdot\}$ обозначает операцию взятия дробной части числа. Покажите, что $\xi(t)$ стационарен в широком и узком смыслах. Покажите, что он обладает марковским свойством.

$\langle \mathbf{2} \rangle$ Пусть $\xi(t)$ — процесс с независимыми приращениями и $T = [0, \infty)$. Пусть $\xi(0) = 0$, $\xi(t_2) - \xi(t_1) \sim Poisson((t_2 - t_1)\lambda)$. Вычислите $\mathbf{P}(\xi(1) \leqslant 3, \xi(2) > 5)$ и $\mathbf{E}\,\xi(1)\xi(2)$.

Задания

4.1 Пусть $\xi(t) = \sum_{i=0}^{t} \gamma_i$, $t \in T = \mathbb{Z}_+$, где $\gamma_i \sim Bern(1/2)$ — последовательность независимых случайных величин. Покажите, что $\xi(t)$ не является стационарным процессом (ни в широком, ни в узком смыслах).

4.2 Пусть $\xi(t) = \max_{t \leqslant i \leqslant t+2} \gamma_i$, $t \in T = \mathbb{Z}_+$, где $\gamma_i \sim \mathcal{N}(0,1)$ — последовательность независимых случайных величин. Покажите, что $\xi(t)$ является стационарным процессом (в узком и в широком смыслах).

4.3 Пусть случайный процесс $\xi(t)$ с $T = [0, \infty)$ и с независимыми приращениями такой, что $\xi(0) = 0$, $\xi(t_2) - \xi(t_1) \sim \mathcal{F}_{[t_1;t_2]}$, где $t_1 < t_2 \in T$ и $\mathcal{F}([t_1;t_2])$ — непрерывное распределение с функцией распределения $F_{[t_1;t_2]}(x)$, функцией плотности $f_{[t_1;t_2]}(x)$ и носителем $[0, \infty)$. Выразите $\mathbf{P}(\xi(1) < 1, \xi(3) > 2)$, $\operatorname{cov}(\xi(t_2), \xi(t_1))$ через распределение приращений.

4.4 Пусть случайный процесс $\xi(0) = 0$, $\xi(t) = \xi(t-1) + \varepsilon_t$ с $T = \mathbb{Z}_+$, где $\varepsilon_t \sim \mathcal{N}(0,1)$ — последовательность независимых случайных величин. Покажите, что для $\xi(t)$ выполняется марковское свойство.

4.5 Пусть случайный процесс $\xi(t) = \max_{i \in \{1,2\}} (\eta_i \, \mathbb{I}(\tau_i \leqslant t))$, $t \in T = [0;1]$, где $\tau_i \sim \mathbb{U}(T)$, $\eta_i \sim \mathbb{U}(\{0,1,2\})$ и независимы в совокупности. Покажите, что для $\xi(t)$ выполняется марковское свойство.

5 Занятие 4. Ковариационная функция

Примеры

$\langle \mathbf{1} \rangle$ Пусть $T = [0; \infty)$, $\xi(t) = t^{\gamma}$, где $\gamma \sim \mathbb{U}([0;1])$. Найдите $\mu_\xi(t)$, $R_\xi(t_1, t_2)$ и покажите исходя из их свойств, что $\xi(t)$ — непрерывный в среднеквадратическом процесс.

Задания

5.1 Пусть $T = [0; \infty)$, $\xi(0) = 0$, $\xi(t_0 + h) - \xi(t_0) = h\eta_k$ при $[t_0; t_0 + h] \subseteq [k-1; k]$, $k \in \mathbb{N}$, где $\eta_k \sim \mathbb{U}([0;1])$ и независимы. Найдите $R_\xi(t_0, t_1)$, $t_0, t_1 \in T$. Исходя из свойств $\mu_\xi(t)$ и $R_\xi(t_0, t_1)$, покажите, что $\xi(t)$ — непрерывный в среднеквадратическом.

5.2 Пусть $T = [0; \infty)$, $\xi(t) = \eta t$, $\eta \sim \mathbb{U}([0;1])$. Найдите $R_\xi(t_0, t_1)$, $t_0, t_1 \in T$. Исходя из свойств $\mu_\xi(t)$ и $R_\xi(t_0, t_1)$, покажите, что $\xi(t)$ — непрерывный в среднеквадратическом.

5.3 Пусть $T = [0; 2]$, $\xi(t) = \eta t - 2\eta \, \mathbb{I}(t \geqslant 1)$, где $\eta \sim \mathbb{U}([-1;1])$. Найдите $R_\xi(t_0, t_1)$, $t_0, t_1 \in T$. Исходя из свойств $\mu_\xi(t)$ и $R_\xi(t_0, t_1)$, покажите, что $\xi(t)$ не является непрерывным в среднеквадратическом.

5.4 **(технически сложное)** Пусть случайный процесс $\xi(t)$, $t \in T = [0;1]$ задан семейством конечномерных распределений: $F(x_1, \ldots, x_k; t_1, \ldots, t_k) = \min_i \left\{ \frac{x_i}{t_i}, 1 \right\} \mathbb{I}(x_i \geqslant 0 \forall i)$. Найдите $R_\xi(t_0, t_1)$, $t_0, t_1 \in T$. Исходя из свойств $\mu_\xi(t)$ и $R_\xi(t_0, t_1)$, покажите, что $\xi(t)$ является непрерывным в среднеквадратическом.

5.5 Пусть $T = \mathbb{Z}_+$, $\xi(0) = 0$, $\xi(t) = \xi(t-1) + \varepsilon_t$, где $\varepsilon_t \sim \mathcal{N}(\mu, \sigma^2)$ и $\{\varepsilon_t\}_t$ независимы в совокупности. Вычислите $R_\xi(t_0, t_1)$.

6 Занятие 5. Спектральная плотность и ковариационная функция стационарного процесса с дискретным временем

В общем виде прямое и обратное преобразования Фурье с дискретным временем для некой (например, абсолютно интегрируемой по считающей мере) функции $f(t)$, $t \in \mathbb{Z}$ имеют вид:

1. прямое преобразование: $\hat{f}(\theta) = \displaystyle\sum_{t=-\infty}^{\infty} f(t) \mathrm{e}^{-\mathrm{i}\theta t}, \quad \theta \in [-\pi; \pi]$
2. обратное преобразование: $f(t) = \displaystyle\frac{1}{2\pi} \int_{-\pi}^{\pi} \mathrm{e}^{\mathrm{i}\theta t} \hat{f}(\theta) \, \mathrm{d}\theta$

При преобразовании функции ковариации стационарного процесса $R(n) = R(t, t + n)$ можно учесть, что $R(n)$ – чётная и вещественная функция (для рассматриваемых нами процессов). Поэтому преобразование Фурье для неё будет тоже чётной и вещественной, а формулы могут быть представлены в виде:

1. прямое преобразование: $g(\theta) = \displaystyle\frac{R(0)}{2\pi} + \frac{1}{\pi} \sum_{n=1}^{\infty} R(n) \cos(\theta n), \quad \theta \in [-\pi; \pi]$
2. обратное преобразование: $R(n) = 2 \displaystyle\int_0^{\pi} \cos(\theta t) g(\theta) \, \mathrm{d}\theta$

Заметим, что здесь множитель $1/(2\pi)$ занёс в выражение для спектральной плотности $g(\theta)$.

Период гармоники $\mathrm{e}^{\mathrm{i}\theta t}$ равен $2\pi/\theta$, а частота — $\theta/(2\pi)$.

Примеры

$\langle \mathbf{1} \rangle$ Пусть $T = \mathbb{N}$ и случайный процесс $\xi(t) = \gamma_{t-1} + \gamma_t$, где $\gamma_t \sim Bern(1/2)$ и независимы. Найдите спектральную плотность процесса $g(\theta)$. Исходя из значений/графика $g(\theta)$ опишите, гармоники с каким периодом вносят наибольший вклад в дисперсию процесса $\xi(t)$. Произведите обратное преобразование — выразите $R(n)$ из $g(\theta)$.

Задания

6.1 Пусть $T = \mathbb{N}$ и $\xi(t) = \gamma_t + \gamma_{t+1}$, где $\gamma_t \sim \mathbb{U}([0;1])$ и независимы. Найдите спектральную плотность процесса $\xi(t)$, нарисуйте её график и опишите, гармоники с какими периодами наиболее характерны для процесса.

6.2 Пусть $T = \mathbb{N}$ и $\xi(t) = \gamma_t + \gamma_{t+3}$, где $\gamma_t \sim \mathcal{N}(0,1)$ и независимы. Найдите спектральную плотность процесса $\xi(t)$, нарисуйте её график и опишите, гармоники с какими периодами наиболее характерны для процесса.

6.3 Пусть $T = \mathbb{Z}$ и известно, что у стационарного процесса $\xi(t)$ спектральная плотность задана формулой $g(\theta) = 3/\pi + 2\cos(4\theta)/\pi$. Найдите значения ковариационной функции процесса $R(n)$.

6.4 **(технически сложное)** Пусть $T = \mathbb{Z}$, а про распределение процесса $\xi(t)$ известно, что он принимает значения из отрезка $[0;1]$ и двумерное распределение процесса $F(x,y;\, t,t+1) = x\,y\min\{x,y\}$, $F(x,y;\, t,t+1+k) = x^2y^2$, где $k \in \mathbb{N}$. Найдите спектральную плотность процесса $\xi(t)$, нарисуйте её график и опишите, гармоники с какими периодами наиболее характерны для процесса.

6.5 Пусть $T = \mathbb{Z}$, $\xi(t) = \sum_{i=-\infty}^{\infty} 2^{-|i-t|}\gamma_i$, где $\gamma_i \sim \mathcal{N}(0,1)$ и независимы. Найдите выражение для спектральной плотности.

7 Занятие 6. Процесс Пуассона

Примеры

$\langle\mathbf{1}\rangle$ Пусть $\xi(t) \sim Poisson(\lambda)$. Известно, что $\mathbf{E}\,\xi(10) = 15$. Найдите $\mathbf{P}(\xi(1) = 1 \mid \xi(3) = 3)$.

Задания

7.1 При производстве деталей на заводе каждые 30 дней в среднем 397 деталей оказываются бракованными. Опишите явление возникновения бракованных деталей как пуассоновский процесс. Опишите распределение вероятностей количества бракованных деталей за 7 дней.

7.2 Предположим, что возникновение голов в футбольном матче описывается пуассоновским процессом, и что за один матч (90 минут) ожидается 2.6 голов. Найдите вероятность того, что в игре и в первом, и во втором тайме (каждый тайм — 45 минут) произойдёт не больше одного гола.

7.3 В типографию приходит в среднем 4 заказа за рабочий день (8 часов). Предположим, что возникновение заказов описывается пуассоновским процессом. Найдите вероятность того, что по итогу рабочего дня типография получит 5 заказов, если за первый час рабочего дня к ней уже поступил один заказ.

7.4 Пусть $\xi(t) \sim Poisson(\lambda)$. Опишите условное распределение случайной величины $\xi(2)$ при условии $\xi(4) = 5$, например, найдя $\mathbf{P}(\xi(2) = x \mid \xi(4) = 5) \quad \forall x \in \mathbb{Z}_+$.

7.5 Пусть $\xi(t) \sim Poisson(\lambda)$. Напишите выражение для наибольшего значения константы $C \in \mathbb{Z}_+$ такого, что $\sup_{\lambda_0 \geqslant 4} \mathbf{P}(\xi(2) \leqslant C;\, \lambda = \lambda_0) \leqslant \alpha$.

8 Занятие 7. Промежутки времени между пуассоновскими событиями

Примеры

$\langle\mathbf{1}\rangle$ Пусть случайный процесс $\xi(t) \sim Poisson(\lambda)$. Найдите распределение τ_2 — момента 2-го пуассоновского события. Найдите ожидаемое время до второго пуассоновского события, если известно, что $\xi(2) = 1$.

Задания

8.1 Пусть случайный процесс $\xi(t) \sim Poisson(\lambda)$. Найдите ожидаемое время наступления 3-го пуассоновского события.

8.2 Пусть случайный процесс $\xi(t) \sim Poisson(\lambda)$, τ_k — момент k-го пуассоновского события в процессе $\xi(t)$. Вычислите $\operatorname{cov}(\tau_1, \tau_2)$.

8.3 Пусть случайный процесс $\xi(t) \sim Poisson(\lambda)$, τ_k — момент k-го пуассоновского события в процессе $\xi(t)$. Положим $\gamma_i \sim \mathcal{N}(0,1)$, а $\zeta(t) = \sum_{i=1}^{\infty} \mathbb{I}(\tau_i \leqslant t)\gamma_i$. Найдите $\mathbf{P}(\zeta(2) \leqslant 2)$.

8.4 Пусть случайный процесс $\xi(t) \sim Poisson(\lambda)$, τ_k — момент k-го пуассоновского события в процессе $\xi(t)$. Найдите $\mathbf{P}(\tau_3 \leqslant 3 \mid \xi(1) = 1)$

8.5 Пусть случайный процесс $\xi(t) \sim Poisson(\lambda)$, τ_k — момент k-го пуассоновского события в процессе $\xi(t)$. Упростите, насколько возможно, выражение для поиска наименьшей константы C такой, чтобы выполнялось неравенство $\sup_{\lambda_0 \geqslant 4} \mathbf{P}(\tau_1 \geqslant C) \leqslant \alpha$.

9 Оценка вероятности переполнения системы массового обслуживания с пуассоновским потоком заявок методом Монте-Карло

Описание и пояснения

Типичная задача, в которой используются пуассоновские процессы — модели систем массового обслуживания. Модель используется для прогнозирования нагрузки системы, планирования количества обслуживающих элементов или схемы обслуживания для минимизации среднего размера очереди, уменьшения среднего времени ожидания и т.п. Пуассоновский процесс и его вариации используются как модель появления новых заявок на обслуживание.

Мы рассмотрим следующую упрощённую модель системы массового обслуживания. Пусть $T = [0; 8]$ — 8 часов рабочего дня. Положим, что $\xi(t) \sim Poisson(\lambda)$ обозначает общее количество поступивших заявок к t-му моменту времени. В этой модели каждое возрастание процесса $\xi(t)$ интерпретируется как появление новой заявки. Обозначим через τ_k время прихода k-ой заявки.

Предположим, что каждая заявка требует для своей обработки случайное количество времени $\gamma_k \sim \mathcal{G}$, где $\mathcal{G}$ — некоторое распределение с неотрицательными значениями. Положим, что γ_k не зависят ни от друг друга, ни от $\xi(t)$.

Нашей задача — по заданным параметрам модели определить минимальное достаточное количество обслуживающих элементов s, при котором вероятность того, что при возникновении новой заявки все обслуживающие элементы будут заняты обслуживанием предыдущих заявок, было не больше заданного ограничения α. Т.е. выполнялось условие:

$$\mathbf{P}\left(\sup_{t\in[0;8]}\sum_{k=1}^{\infty}\mathbb{I}(t\in[\tau_k;\tau_k+\gamma_k])>s\right)\leqslant\alpha,$$

где α — заданное ограничение на вероятность.

Для численной оценки этой вероятности будем применять метод Монте-Карло (симуляция случайного процесса).

Применение метода Монте-Карло

Идея метода: если мы хотим найти $\mathbf{P}(\xi\in A)$ для некоторого процесса $\xi(t)$, то можно сгенерировать несколько траекторий процесса x_i, $i\in\{1,\ldots,N\}$ и оценить $\mathbf{P}(\xi\in A)\approx N^{-1}\sum_{i=1}^{N}\mathbb{I}(x_i\in A)$. Здесь N указывает на количество итераций метода Монте-Карло.

Конкретизируем, как это выглядит в нашей задаче.

1. Обозначим i-ую итерацию метода Монте-Карло как $\mathcal{M}_i$.
2. На каждой итерации метода Монте-Карло нам понадобится сгенерировать последовательность τ_k и γ_k. Для этого:
   (a) Генерируем последовательность независимых случайных величин $\eta_k\sim$ $\sim Exponential(\lambda)$ — времена ожиданий до следующих пуассоновских событий в текущей реализации процесса
   (b) Полагаем $\tau_k=\sum_{i=1}^{k}\eta_i$
   (c) Случайные величины η_k и τ_k нужно генерировать до индекса n, для которого $\tau_n\geqslant 8$
   (d) Генерируем последовательность независимых $\gamma_1,\ldots,\gamma_n\sim\mathcal{G}$ — время обработки каждой из заявок
3. Пусть $n=\max\{i\colon\tau_i\leqslant 8\}$. Подсчитываем пиковое количество активных заявок для текущей итерации $\mathcal{M}_i$:

$$Q_i=\sup_{j\in\{1,\ldots,n\}}\sum_{l=1}^{j}\mathbb{I}(\tau_j\leqslant\tau_l+\gamma_l).$$

4. Искомое s выбираем как

$$s=\min\left\{s\colon\frac{1}{N}\sum_{i=1}^{N}\mathbb{I}(Q_i>s)\leqslant\alpha\right\},$$

где N — количество итераций метода Монте-Карло, α — заданное ограничение на вероятность превышения максимального количества активных заявок над количеством обслуживающих элементов s.

Задание

Каждому студенту представлен набор параметров:

1. α (`alpha`) — ограничение на вероятность превышения количества заявок
2. λ (`lambda`) — интенсивность появления заявок
3. $\mathcal{G}$ (`G`) — распределение продолжительности обработки одной заявки

Исходя из $N = 1\,000$ итераций метода Монте-Карло определите минимальное количество обслуживающих элементов s, достаточных для того, чтобы вероятность превышения пикового количества активных заявок над s не превосходила α.

В качестве ответа приведите:

1. значение s
2. график оценок вероятностей превышения пикового количества заявок для выбранного и всех меньших значений s

Примечения

У экспоненциального распределения существует два варианта параметризации:

- с параметром λ, называемым rate, которому соответствует функция плотности $f(x;\lambda) = \lambda \mathrm{e}^{-\lambda x}$ и математическим ожиданием λ^{-1}
- с параметром θ, называемым scale, которому соответствует функция плотности $f(x;\theta) = \theta^{-1}\mathrm{e}^{-\theta^{-1}x}$ и математическим ожиданием θ

В этом курсе мы используем первый вариант (rate / λ), однако в статистических пакетах встречаются и тот, и другой варианты. Например

- функция `numpy.random.exponential` библиотеки `numpy` языка `Python` реализует вариант с параметризацией через scale / θ
- функция `rexp` языка `R` реализует вариант с параметризацией через rate / λ

Чтобы убедиться, какой вариант реализован в используемом вами варианте, можно, например, посчитать среднее арифметическое сгенерированной выборки и сравнить его с математическими ожиданиями того и другого варианта параметризации.