Лекция 12

ПРОЦЕССЫ ВОССТАНОВЛЕНИЯ

Процесс восстановления — это стохастическая модель для описания физического явления, в котором некоторое событие происходит в случайные моменты времени.

Пример 1 (*обслуживание заявок*). Оператор в колл центре отвечает на звонки, которые возникают в непредсказуемые моменты времени. Длительность ответов заранее предсказать невозможно. Важной характеристикой работы оператора является количество звонков, на которые он смог ответить в течение определенного времени. Эту характеристику можно использовать не только для оценки работы оператора, но и для оптимизации работы всего колл центра.

Пример 2 (*работа страховой компании*). Страховая компания время от времени получает запросы от своих клиентов на возмещение убытков в случае возникновения обстоятельств, предусмотренных страховыми полисами. Запросы возникают в случайные моменты времени; суммы на возмещение убытков заранее предсказать невозможно. Количество запросов, поступивших в течение определенного времени, является одной из важнейших для работы страховой компании. Еще более важной характеристикой является сумма компенсаций, которую компания должна выплатить держателям страховых полисов.

1. Определение процесса восстановления

Обозначим через $0 \le T_1 \le T_2 \le \dots$ случайные моменты, в которые происходит некоторое событие. Тогда

$$N(t) = \sum_{n=1}^{\infty} \mathbb{1}(T_n \le t), \qquad t \ge 0, \tag{1}$$

— это количество членов последовательности $\{T_n\}$, принадлежащих интервалу $[0,t]$.

Лемма 1. *$N(t) < \infty$ для любого $t \ge 0$ тогда и только тогда, когда $T_n \to \infty$, $n \to \infty$.*

Определение 1. $N(t)$, определенный формулой (1), называется *процессом восстановления*, если

(a) $0 < T_1 < T_2 < \dots$ и $T_n \to \infty$, $n \to \infty$;
(b) случайные величины $\xi_n = T_n - T_{n-1}$, $n \ge 1$, являются независимыми и одинаково распределенными, где $T_0 = 0$.

Случайные величины T_n называются *моментами восстановления*, а ξ_n — временами между моментами восстановления. Сам процесс $N(t)$ — это количество восстановлений на интервале $(0,t]$.

[0]Printed from the file [TSP-00-14.tex] on 19.11.2014

Пример 3 (*замена лампочек*). Электрическая лампочка освещает помещение. Как только она перегорает, она немедленно заменяется на идентичную, то есть освещение "*восстанавливается*". Если обозначить через ξ_n время в течение которого горела n-ая лампочка, то $T_n = \xi_1 + \cdots + \xi_n$ — это момент, когда "*восстановление*" произошло в n-ый раз, а $N(t)$ — это количество "*восстановлений*" на интервале $(0, t]$.

Пример 4 (*поиск совпадений*). Сервер обслуживает поток данных. Одна из комбинаций символов представляет интерес. Момент прохождения такой комбинации через сервер — это и есть момент "*восстановления*". Количество комбинаций, прошедших через сервер, — это и есть процесс восстановления.

Пример 5 (*комбинации в ДНК*). Заменить в примере 4 слово "*сервер*" на "*ДНК*", а выражение "*комбинация символов*" — на "*специфическая последовательность молекул*".

2. Распределения процесса восстановления

Чтобы определить вероятностные характеристики процесса восстановления, необходимо задать функцию распределения времен ξ_n. Мы будем обозначать эту функцию распределения через F. Будем также предполагать, что $F(0) = 0$.

Лемма 2. *Функция распределения случайной величины T_n вычисляется по формуле свертки*

$$\mathsf{P}(T_n < t) = F \underbrace{* \cdots *}_{n-1 \text{ раз}} F(t) = F^{n*}(t), \qquad t \in \mathbf{R}.$$

Напомним, что свертка $F * G$ двух функций распределения F и G равна

$$F * G(x) = \int_{-\infty}^{\infty} F(x - z)\, dG(z), \qquad x \in \mathbf{R}.$$

$\boxed{\text{почему } F * F(0) = 0, \text{ если } F(0) = 0?}$

Если функция G дифференцируема, то есть имеет плотность распределения $g(z) = G'(z)$, то Напомним, что свертка $F * G$ двух функций распределения F и G равна

$$F * G(x) = \int_{-\infty}^{\infty} F(x - z) g(z)\, dz, \qquad x \in \mathbf{R}.$$

Если же и функция F дифференцируема, то есть имеет плотность распределения $f(z) = F'(z)$, то и $F * F$ имеет плотность, причем

$$f_{F*G}(x) = \int_{-\infty}^{\infty} f(x - z) g(z)\, dz, \qquad x \in \mathbf{R}. \tag{2}$$

2.1. Процесс Пуассона. Предположим, что время между последовательными восстановлениями имеет показательное распределение с параметром $\lambda > 0$, то есть

$$F(x) = \begin{cases} 0, & x \le 0, \\ 1 - e^{-\lambda x}, & x > 0, \end{cases} \quad \Longleftrightarrow \quad f(x) = \begin{cases} 0, & x \le 0, \\ \lambda e^{-\lambda x}, & x > 0. \end{cases}$$

Лемма 3. *Если F является экспоненциальной, то свертка F^{n*} имеет плотность f_n, причем $f_n(x) = 0$, $x \le 0$, и*

$$f_n(x) = \lambda^n x^{n-1} \frac{e^{-\lambda x}}{(n-1)!}, \qquad x > 0.$$

Доказательство. Мы докажем эту формулу по индукции. При $n = 1$ она справедлива, так как f_1 — это плотность показательного распределения. Предположив справедливость формулы для $n - 1 \ge 1$, докажем ее для n. Для этого воспользуемся равенством (2): для $x > 0$ имеем

$$\begin{aligned} f_n(x) &= f_{F*F^{(n-1)*}}(x) = \int_{-\infty}^{\infty} f(x-z) f_{n-1}(z)\,dz = \int_0^{\infty} f(x-z) f_{n-1}(z)\,dz \\ &= \int_0^x f(x-z) f_{n-1}(z)\,dz = \int_0^x \lambda e^{-\lambda(x-z)} \cdot \lambda^{n-1} z^{n-2} \frac{e^{-\lambda z}}{(n-2)!}\,dz \\ &= \lambda^n \frac{e^{-\lambda x}}{(n-2)!} \int_0^x z^{n-2}\,dz = \lambda^n x^{n-1} \frac{e^{-\lambda x}}{(n-1)!}. \end{aligned}$$

□

Поскольку F^{n*} — это функция распределения случайной величины T_n, то

$$\mathsf{P}(T_n < x) = \int_0^x \lambda^n z^{n-1} \frac{e^{-\lambda z}}{(n-1)!}\,dz.$$

Лемма 4.

$$\mathsf{P}(T_n < x) = 1 - \sum_{k=0}^{n-1} e^{-\lambda x} \frac{(\lambda x)^k}{k!}.$$

Воспользуемся соотношением *двойственности*, которое докажем позднее:

$$T_n \le t \quad \Longleftrightarrow \quad N(t) \ge n. \tag{3}$$

Из (3) вытекает, что

$$\mathsf{P}(N(t) \ge n) = 1 - \sum_{k=0}^{n-1} e^{-\lambda t} \frac{(\lambda t)^k}{k!}.$$

Поэтому

$$\mathsf{P}(N(t) = n) = \mathsf{P}(N(t) \ge n) - \mathsf{P}(N(t) \ge n+1) = e^{-\lambda t} \frac{(\lambda t)^n}{n!}, \qquad n \ge 0.$$

Теорема 1. *Если F — это экспоненциальная функция распределения, то $N(t)$, $t \ge 0$, имеет распределение Пуассона с параметром λt.*

Определение 2. Процесс востановления, построенный по экспоненциальным временам между последовательными восстановлениями, называется процессом Пуассона.

3. Дуальность и другие элементарные свойства

3.1. Доказательство свойства дуальности (3). Если $T_n \le t$ для некоторого n, то $T_k \le t$ для всех $k \le n$. Это означает, что не менее n слагаемых в сумме (1) равны 1, то есть $N(t) \ge n$.

Если же наоборот $N(t) \ge n$, то не менее n слагаемых в сумме (1) равны 1. Так как последовательность $\{T_k\}$ возрастает, то $T_n \le t$. Эти два коротких рассуждения доказывают (3).

3.2. Другие свойства дуальности. Аналогичным образом доказываются следующие свойства, связывающие последовательность $\{T_n\}$ и процесс $N(t)$:

$$T_{N(t)} \le t < T_{N(t)+1}, \tag{4}$$

$$N(T_n) = n, \tag{5}$$

$$N(t) = \max\{n : T_n \le t\} = \min\{n : T_n > t\} - 1. \tag{6}$$

доказать (4)–(6)

4. Примеры моделей с процессами восстановления

4.1. Обслуживание. Техническое обслуживание автомобиля необходимо проводить после пробега L километров или прошествия M дней после последнего обслуживания в зависимости от того, какое из этих условий выполнится первым. Обозначим через τ время, необходимое для пробега L километров, а через $N(t)$ — количество технических осмотров до момента t. Если считать, что пробеги между осмотрами являются независимыми и одинаково распределенными, то $N(t)$ — это процесс восстановления, причем функция распределения времени между восстановления равна $F(t) = \mathsf{P}(\tau \wedge M < t)$.

Модели с такой функцией F являются довольно распространенными. Например, в теории надежности модель старения предусматривает замену компонента системы, если от отказывает или если достигает определенного возраста.

см. задачу 19, p. 160,Serfozo

4.2. Циклический процесс восстановления. Пусть случайный процесс $X(t)$ оказывается в состояниях $S_0, S_1, \ldots, S_{K-1}$ (для краткости будем писать $0, 1, \ldots, K-1$) по кругу. Последовательность состояний процесса изображена ниже

$$0 \longrightarrow 1 \longrightarrow \ldots \longrightarrow K-1 \longrightarrow 0 \longrightarrow \ldots$$

Считаем, что времена, которое процесс проводит в состояниях $0, 1, \ldots, K-1$, независимы. Обозначим через F_j функцию распределения времени пребывания в состоянии j. Если $F_j(0) = 0$, то вероятность того, что $X(t)$ "*проскочит*" состояние j равна 0. почему?

Время полного цикла (время возвращения в 0) имеет функцию распределения $F = F_0 * F_1 * \cdots * F_{K-1}$. доказать! С помощью времен последовательного возвращения в 0 $0 = T_0 \le T_1 \le \ldots$ можно построить процесс восстановления, который называется *циклическим процессом восстановления*.

Если предположить, что $F_0(0) > 0$, то можно построить другой процесс восстановления, используя моменты времени, когда $X(t)$ "*проскакивает*" состояние 0 и оказывается в состоянии 1 сразу из $K-1$.

см. задачу 7, p. 157,Serfozo

Альтернирующий процесс восстановления. Рассмотрим случай $K = 2$. В этом случае процесс $X(t)$ “*колеблется*” между двумя альтернативными состояниями, например, “исправен” и “неисправен”, “доступен” и “недоступен” или “готов” и “не готов”.

Если предположить, что функция распределения F_0 времени “*сидения*” в 0 является вырожденной, то есть $F_0(x) = \mathbb{1}_{(0,\infty)}(x)$, то время пребывания в 0 равно 0 с вероятностью 1. Эта ситуация отвечает классической модели, когда работоспособность системы восстанавливается немедленно (см. пример 3).

Трихотомический процесс восстановления. Рассмотрим случай $K = 3$. Одна из интерпретаций этого случая возникает в экономической теории, в которой различают такие состояния экономики: “бум”, “застой”, “рецессия”.

4.3. Процесс восстановления с премиями. Предположим, что в каждый момент T_n возникает обязательство заплатить определенную сумму Y_n (называемую *премией*). Например, страховая компания при возникновении каждого страхового случая обязана заплатить по обязательствам согласно страховому полису (см. пример 2). Тогда общая сумма обязательств на интервале $(0, t]$ равна

$$Z(t) = \sum_{n=1}^{\infty} Y_n \mathbb{1}(T_n \le t) = \sum_{n=1}^{N(t)} Y_n, \qquad t > 0.$$

Такой случайный процесс называется *процессом восстановления с премиями.*

4.4. Процесс восстановления с дисконтными премиями. В экономической теории принято считать, что деньги теряют свою ценность с течением времени. Этот процесс называется *инфляцией.* Зависимость ценности денег от времени выражается с помощью экспоненциальной функции; сумма y “*сегодня*” будет иметь ценность $ye^{-\alpha t}$ через время t.

Этот факт можно пояснить следующим образом, используя понятие сложных процентов для непрерывного времени: депозит в банке величины y “*превратится*” в сумму $ye^{\alpha t}$, где α — это процентная ставка. Таким образом, сумма депозита увеличится в $e^{\alpha t}$ раз. Если в момент t мы будем обладать суммой y, то это означает, что “*сегодня*” мы имеем сумму $ye^{-\alpha t}$.

С учетом инфляции, процесс восстановления с премиями надо корректировать так

$$Z(t) = \sum_{n=1}^{N(t)} Y_n e^{-\alpha T_n}, \qquad t > 0.$$

Такой случайный процесс называется *процессом восстановления с дисконтными премиями.*