

Ю. Н. Каштанов

**Статистическое моделирование
систем массового обслуживания**

Санкт-Петербург
2014

УДК 519.245

Печатается по решению кафедры статистического моделирования
Санкт-Петербургского государственного университета

В настоящем издании рассмотрены методы статистического моделирования для численного решения задач теории массового обслуживания. Оно может быть использовано в качестве учебного пособия для студентов кафедры статистического моделирования СПбГУ.

Подготовлено при финансовой поддержке РФФИ (грант 14-01-00271).

ISBN

© Ю.Н.Каштанов 2014

Содержание

1	Модели систем массового обслуживания	2
1.1	Марковские процессы	2
1.2	Полумарковские процессы	11
2	Статистическое моделирование	15
2.1	Ответвление траекторий.	17
2.2	Вложенная марковская цепь	25
2.3	Ветвящийся процесс	31
2.4	Численный пример	33

В первом разделе приводятся модели и примеры систем массового обслуживания, а также утверждения об эргодичности описывающих их процессов. Во втором разделе рассматривается задача эффективно-го вычисления методом Монте-Карло некоторых характеристик систем массового обслуживания. Одна из сложностей таких вычислений состоит в оценке вероятностей редких событий, т.к. при этом требуется моделировать траектории большой длины. Для того, чтобы уменьшить вычислительные затраты были предложены различные методы, которые и рассматриваются в параграфах 2.1-2.3. В параграфе 2.4 приведено численное сравнение рассмотренных методов на примере простой системы массового обслуживания.

1 Модели систем массового обслуживания

1.1 Марковские процессы

Пусть $(X, \mathcal{B})$ - измеримое пространство и дана случайная последовательность $x_k \in X$, $k = 0, 1, \dots$; положим $\mathcal{F}_k = \sigma(x_0, \dots, x_k)$. Последовательность x_k называется марковской (цепью), если существуют вероятностные меры $p_0(\cdot)$, $p_k(x, \cdot)$, $k \geq 1$, на $\mathcal{B}$, такие что

$$\mathbf{P}(x_0 \in A_0, \dots, x_k \in A_k) = \int_{A_0} p(dy_0) \int_{A_1} p_1(y_0, dy_1) \dots p_k(y_{k-1}, A_k).$$

Мера $p_0(\cdot)$ определяет начальное распределение x_0 , а $p_k(x, dy)$ - вероятности перехода на k -ом шаге. Если $p_0(\cdot)$ сосредоточена в точке x , то будем говорить, что последовательность начинается в этой точке и для вероятностей применять обозначение $\mathbf{P}_x$. Если $p_k(x, A)$ не зависит от k , то марковская цепь называется однородной; в дальнейшем будем рассматривать только однородные марковские цепи.

Вероятность перехода за n шагов обозначим $p^{(n)}(x, A) = \mathbf{P}_x(x_n \in A)$. Из определения марковской цепи следует уравнение Чепмена - Колмогорова:

$$p^{(n+m)}(x, A) = \int_X p^{(n)}(x, dy) p^{(m)}(y, A).$$

Введем оператор P на пространстве ограниченных функций, связанный с ядром $p(x, dy)$: $P_x f = \int_X p(x, dy) f(y)$. Очевидно, что имеют место равенства $\mathbf{E}_x f(x_k) = \int_X p^{(n)}(x, y) f(y) = P_x^n f$.

1.1.1. Стационарное распределение. Предположим, что для некоторого n имеет место неравенство $p^{(n)}(x, A) \geq \varepsilon \mu(A)$, где μ - вероятностная мера, тогда

$$\begin{aligned} p^{(n)}(x, A) - p^{(n)}(y, A) &= 1 - p^{(n)}(x, X \setminus A) - p^{(n)}(y, A) \\ &\leq 1 - \varepsilon \mu(X \setminus A) - \varepsilon \mu(A) = 1 - \varepsilon; \end{aligned}$$

откуда следует, что $\sup_{x, y} |p^{(n)}(x, A) - p^{(n)}(y, A)| \leq 1 - \varepsilon$.

Полный заряд меры $Q(\cdot) = p^{(n)}(x, \cdot) - p^{(n)}(y, \cdot)$ равен нулю, поэтому $Q(X^+) = -Q(X^-)$, где X^+ , X^- - множества положительности и отрицательности меры Q . Отсюда следует, что

$$\begin{aligned} p^{(m)}(x, A) - p^{(m)}(y, A) &= \\ &= \int_{X^+} [p^{(n)}(x, dz) - p^{(n)}(y, dz)] p^{(m-n)}(z, A) + \\ &+ \int_{X^-} [p^{(n)}(x, dz) - p^{(n)}(y, dz)] p^{(m-n)}(z, A) \leq \\ &\leq Q(X^+) [\sup_z p^{(m-n)}(z, A) - \inf_z p^{(m-n)}(z, A)] \leq \\ &\leq (1 - \varepsilon) \sup_{x_1, y_1} [p^{(m-n)}(x_1, A) - p^{(m-n)}(y_1, A)]. \end{aligned}$$

Итерируя полученное неравенство, имеем $|p^{(m)}(x, A) - p^{(m)}(y, A)| \leq C \delta^m$, где $\delta = (1 - \varepsilon)^{1/n}$ и, значит, существует предельное распределение

$$\pi(A) = \lim_k p^{(k)}(x, A), \quad (1)$$

причем

$$|\pi(A) - p^{(k)}(x, A)| \leq C \delta^k. \quad (2)$$

Из (1) следует также, что

$$\int_X \pi(dx) p(x, A) = \pi(A), \quad (3)$$

поэтому π называется стационарным распределением.

1.1.2. Закон больших чисел. Пусть $h(x)$ - ограниченная функция, $J = \int_X \pi(dx)h(x)$, покажем, что

$$\bar{J}_n = \frac{1}{n} \sum_{i=0}^{n-1} h(x_i) \xrightarrow{\mathbf{P}} J.$$

Обозначим $f(x) = h(x) - J$. Из неравенства Чебышева

$$\begin{aligned} \mathbf{P}(|\bar{J}_n - J| > \varepsilon) &\leq \frac{1}{\varepsilon^2 n^2} \mathbf{E} \left(\sum_{i=0}^{n-1} f(x_i) \right)^2 \\ &= \frac{1}{\varepsilon^2 n^2} \left[\sum_{i=0}^{n-1} \mathbf{E} f^2(x_i) + 2 \sum_{i=0}^{n-1} \mathbf{E} f(x_i) \sum_{j=i+1}^{n-1} \mathbf{E}_{x_i} f(x_j) \right]. \end{aligned}$$

Используя (2), получаем оценку

$$|\mathbf{E}_{x_i} f(x_j)| = \left| \int_X [p^{(j-i)}(x_i, dy) - \pi(dy)] h(y) \right| \leq C \gamma^{j-i}.$$

Таким образом, $\mathbf{P}(|\bar{J}_n - J| > \varepsilon) \leq C/n \rightarrow 0$.

1.1.3. Замечание. Имеет место и более сильное утверждение ([2] Т.1): Если существует стационарное распределение π , то предел для $\bar{J}_n$ существует п.н. для п.в. x по мере π . Если дополнительно выполнено $p^{(k)}(x, A) \rightarrow \pi(A)$ и $\int_X \pi(dx) |h(x)| < \infty$, тогда $\lim_n \bar{J}_n = J$ п.н.

1.1.4. Скачкообразный марковский процесс с непрерывным временем конструктивно можно описать следующим образом. Процесс x_t , находится в состоянии x время τ , которое распределено экспоненциально с параметром $\Lambda(x)$, затем осуществляется переход в новое состояние y в соответствии с переходными вероятностями $p(x, dy)$, после чего процесс ведет себя аналогично. Траектории процесса будем считать непрерывными справа. Пусть t_n - моменты скачков и $x_n = x(t_n)$ - вложенная

марковская цепь, будем обозначать $\tau_{n+1} = t_{n+1} - t_n$, т.е. x_t находится в состоянии x_n в течение времени τ_{n+1} .

1.1.5. Пуассоновский процесс. Пусть τ_i - последовательность независимых экспоненциально распределенных величин с параметром μ :

$$\mathbf{P}(\tau_i > t) = e^{-\mu t}.$$

Положим $t_n = \sum_{i=1}^n \tau_i$, $n = 0, 1, \dots$; будем считать, что t_n это моменты прихода заявок, для которых определен некоторый порядок обслуживания. Распределение t_n называется распределением Эрланга; найдем плотность распределения $p_n(t)$. Очевидно, $p_0(t) = \delta_0(t)$; покажем, что для $n \geq 1$

$$p_n(t) = \mu \frac{(\mu t)^{n-1}}{(n-1)!} e^{-\mu t}. \quad (4)$$

Формулу (5) докажем индукцией: для $n = 1$ формула верна; пусть она верна для n , тогда

$$\mathbf{P}(t_{n+1} < t) = \int_0^t ds p_1(s) \mathbf{P}(t_n < t - s).$$

Дифференцируя, получим

$$\begin{aligned} p_{n+1}(t) &= \int_0^t ds \mu e^{-\mu s} p_n(t-s) = \int_0^t du \mu e^{-\mu(t-u)} \mu \frac{(\mu u)^{n-1}}{(n-1)!} \\ &= e^{-\mu t} \frac{\mu}{(n-1)!} \int_0^{\mu t} dv v^{n-1} = \mu e^{-\mu t} \frac{(\mu t)^n}{n!}. \end{aligned} \quad (5)$$

Обозначим $N_t = \max\{n : t_n \leq t\}$; процесс N_t называется пуассоновским. Заметим, что для данного процесса $X = 0, 1, \dots$, $\Lambda(x) = \mu$, $p(x, y) = 1$ при $y = x + 1$ и $p(x, y) = 0$ в остальных случаях.

Зафиксируем t и найдем распределение случайной величины N_t ; из (5) следует, что для $k \geq 0$

$$\mathbf{P}(N_t = k) = \int_0^t ds p_k(s) e^{-\mu(t-s)} = \frac{(\mu t)^k}{k!} e^{-\mu t} = P_k(t).$$

Покажем, что случайные величины N_t и $N_s - N_t$ независимы. Для этого надо доказать $\mathbf{P}(N_t = k, N_s - N_t = m) = P_k(t)P_m(s - t)$. При $m = 0$

$$\begin{aligned}\mathbf{P}(N_t = k, N_s - N_t = 0) &= \int_0^t du p_k(u) e^{-\mu(t-u)} = \\ &= \int_0^t du p_k(u) e^{-\mu(t-u)} e^{-\mu(s-t)} = P_k(t)P_0(s - t).\end{aligned}$$

При $m \geq 1$

$$\begin{aligned}\mathbf{P}(N_t = k, N_s - N_t = m) &= \\ &= \int_0^t du p_k(u) \int_{t-u}^{s-u} dv p_1(v) P_{m-1}(s - u - v) = \\ &= \int_0^t du p_k(u) \int_0^{s-t} dz \mu e^{-\mu(t-u+z)} P_{m-1}(s - t - z) = \\ &= \int_0^t p_k(u) e^{-\mu(t-u)} \int_0^{s-t} dz \mu e^{-\mu z} P_{m-1}(s - t - z) = \\ &= P_k(t)P_m(s - t).\end{aligned}$$

1.1.6. Для марковских процессов с непрерывным временем имеют место утверждения, аналогичные 1.1.1 – 1.1.3. При этом стационарное распределение этого процесса $\Pi(A)$ может быть выражено как

$$\Pi(A) = \lim_{T \rightarrow \infty} \frac{1}{T} \int_0^T dt 1_A(x_t) \quad \text{п.н.} \quad (6)$$

Укажем выражение для этого предела в терминах стационарного распределения $\pi(dx)$ вложенной марковской цепи и функции $\Lambda(x)$. Пусть для переходных вероятностей $p(x, dy)$ выполнено $p^{(k)}(x, A) \rightarrow \pi(A)$; определим $h_0(x) = 1_A(x)/\Lambda(x)$, $h_1(x) = 1/\Lambda(x)$ и предположим, что $\int_X \pi(dx) h_1(x) < \infty$; покажем, что тогда

$$\Pi(A) = \frac{\int_X \pi(dx) h_0(x)}{\int_X \pi(dx) h_1(x)} = \frac{J_0}{J_1}. \quad (7)$$

Рассмотрим последовательность $\bar{x}_n = (x_n, \tau_{n+1}) \in \bar{X} = X \times R_+$. Эта последовательность представляет из себя марковскую цепь с переходными вероятностями из $\bar{x} = (x, t)$ в $d\bar{y} = dy \times ds$

$$\bar{p}(\bar{x}, d\bar{y}) = p(x, dy)\Lambda(y)e^{-\Lambda(y)s}ds,$$

причем

$$\bar{p}^{(k)}(\bar{x}, d\bar{y}) = p^{(k)}(x, dy)\Lambda(y)e^{-\Lambda(y)s}ds \rightarrow \pi(dy)\Lambda(y)e^{-\Lambda(y)s}ds = \bar{\pi}(d\bar{y}).$$

В выражении (6) подставим t_n на место T , тогда

$$\frac{1}{t_n} \int_0^{t_n} dt 1_A(x_t) = \frac{\sum_{i < n} 1_A(x_i) \tau_{i+1}}{\sum_{i < n} \tau_{i+1}}$$

В соответствии с п. 1.1.3

$$\frac{1}{n} \sum_{i < n} 1_A(x_i) \tau_{i+1} \rightarrow \int_{\bar{X}} \pi(dx) \int_0^\infty dt \Lambda(x) e^{-\Lambda(x)t} 1_A(x) t = J_0 \quad \text{п.н.}$$

Для доказательства (7) теперь достаточно показать что существует предел в (6). Это доказательство будет проведено в следующем параграфе для более общего случая полумарковских процессов.

Пример 1. Рассмотрим следующую систему массового обслуживания:

- входной поток заявок – пуассоновский с параметром μ ;
- одно обслуживающее устройство; время обслуживания экспоненциальное с параметром λ ;
- очереди нет; если устройство занято, то заявка уходит без обслуживания.

Данная система часто носит обозначение $(M/M/1/0)$ [1], в котором первый параметр обозначает распределение времени между приходом заявок (D - детерминированное, M - экспоненциальное, G - произвольное и др.), второй - распределение времени обслуживания, третий - число обслуживающих устройств, четвертый - число мест в очереди.

Введем процесс $x_t \in X = \{0, 1\}$, равный числу заявок в системе в момент времени t . Этот процесс марковский и функция $\Lambda(x)$ из п.1.1.4 принимает значения $\Lambda = \{\mu, \mu + \lambda\}$. Найдем уравнения для $P_i(t) = \mathbf{P}(x_t = i)$. Обозначим A_k событие, заключающееся в том, что на промежутке $(t, t + \Delta t)$ пришло k заявок. Поскольку A_k не зависит на от того, сколько заявок пришло до момента t , ни от того как они были обслужены, то

$$\mathbf{P}(A_k | x_t = i) = \frac{(\mu \Delta t)^k}{k!} e^{-\mu \Delta t}.$$

Получим уравнения для $P_i(t)$, $i = 0, 1$:

$$\begin{aligned} P_0(t + \Delta t) &= \sum_{k=0}^2 \mathbf{P}(x_{t+\Delta t} = 0, A_k) + O((\Delta t)^2) = \\ &= \sum_{k=0}^2 \sum_{i=0}^2 \mathbf{P}(x_{t+\Delta t} = 0, A_k | x_t = i) P_i(t) + O((\Delta t)^2) \end{aligned} \quad (8)$$

Легко убедиться, что

$$\begin{aligned} \mathbf{P}(x_{t+\Delta t} = 0, A_0 | x_t = 0) &= e^{-\mu \Delta t}, \\ \mathbf{P}(x_{t+\Delta t} = 0, A_0 | x_t = 1) &= 1 - e^{-\lambda \Delta t}, \\ \mathbf{P}(x_{t+\Delta t} = 0, A_1 | x_t = 0) &= \int_t^{t+\Delta t} ds \mu e^{-\mu s} (1 - e^{-\lambda(\Delta t - s)}) \leq C(\Delta t)^2, \\ \mathbf{P}(x_{t+\Delta t} = 0, A_1 | x_t = 1) &\leq (1 - e^{-\lambda \Delta t})(1 - e^{-\mu \Delta t}) \leq C(\Delta t)^2 \end{aligned}$$

Используя полученные выражения в (8), получим

$$P_0(t + \Delta t) = (1 - \mu \Delta t) P_0(t) + \lambda \Delta t P_1(t) + O((\Delta t)^2).$$

Переходя к пределу при $\Delta t \rightarrow 0$, имеем

$$P'_0(t) = -\mu P_0(t) + \lambda P_1(t). \quad (9)$$

Добавляя условие нормировки $P_0(t) + P_1(t) = 1$, при $P_0(0) = 1$ получаем

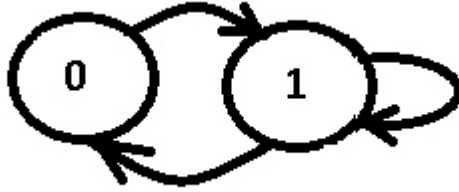
$$P_0(t) = \frac{\lambda}{\lambda + \mu} + \frac{\mu}{\lambda + \mu} e^{-(\lambda + \mu)t}.$$

Отсюда

$$\lim_{t \rightarrow \infty} P_0(t) = \frac{\lambda}{\lambda + \mu} = \Pi_0.$$

Можно проверить, что Π_0 не зависит от $P_0(0)$.

Теперь вычислим Π_0 другим способом – применяя формулу (7). Рассмотрим последовательность $x_n = x(t_n)$. Граф переходов имеет вид



Вероятности переходов $\{p(x, y)\}$, $x, y \in X$, равны $p = \begin{pmatrix} 0 & 1 \\ \frac{\lambda}{\lambda + \mu} & \frac{\mu}{\lambda + \mu} \end{pmatrix}$.

Стационарное распределение π находим из уравнений (3), и уравнения нормировки, которые в данном случае имеют вид

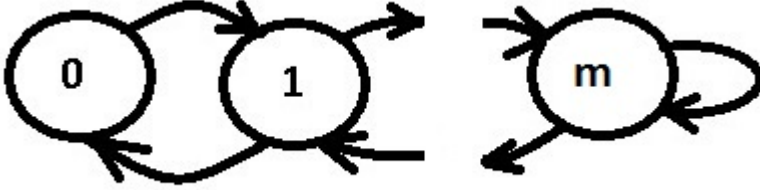
$$\pi_0 = \frac{\lambda}{\lambda + \mu} \pi_1, \quad \pi_0 + \pi_1 = 1.$$

Отсюда $\pi = \left(\frac{\lambda}{2\lambda + \mu}, \frac{\lambda + \mu}{2\lambda + \mu} \right)$. Далее вычисляем функционалы J_0 и J_1 от стационарного распределения

$$J_0 = \sum_x \pi(x) h_0(x) = \frac{\lambda/\mu}{2\lambda + \mu}, \quad J_1 = \sum_x \pi(x) h_1(x) = \frac{\lambda/\mu + 1}{2\lambda + \mu}.$$

Применяем формулу (7): $\Pi_0 = J_0/J_1 = \lambda/(\lambda + \mu)$.

Пример 2. Рассмотрим систему $(M/M/m/0)$, которая отличается от предыдущей наличием m обслуживающих устройств; множество состояний $X = (0, 1, \dots, m)$. Пусть $x_n = x(t_n)$, граф переходов выглядит следующим образом



Вероятности переходов

$$p_{ij} = \begin{cases} \frac{\lambda i}{\lambda i + \mu}, & j = \max(i - 1, 0), \\ \frac{\mu}{\lambda i + \mu}, & j = \min(i + 1, m), \\ 0, & \text{в остальных случаях.} \end{cases}$$

Из уравнений (3) находим соотношение для вероятностей стационарного распределения

$$\pi_i \frac{\lambda i}{\lambda i + \mu} = \pi_{i-1} \frac{\mu}{\lambda(i-1) + \mu}.$$

Итерируя, выражаем π_i через π_0 :

$$\pi_i = \frac{1}{i!} \left(\frac{\mu}{\lambda} \right)^i (\lambda i + \mu) \pi_0.$$

Применяя формулу (7), после нормировки имеем

$$\Pi_i = \frac{1}{i!} \left(\frac{\mu}{\lambda} \right)^i / \sum_{k=0}^n \frac{1}{k!} \left(\frac{\mu}{\lambda} \right)^k.$$

Пример 3. Если обслуживающие устройства имеют различные интенсивности обслуживания λ_i , то множество состояний определяются вектором

$$x = (\delta_1, \dots, \delta_m), \quad \delta_i = 0, 1.$$

При этом количество состояний 2^m ; уже при $m = 10$ число состояний (уравнений) больше 1000, что делает невозможным получение аналитического решения. Численные результаты для этого примера будут приведены в следующем разделе.

1.2 Полумарковские процессы

Не все системы массового обслуживания могут быть описаны скачкообразными марковскими процессами. Укажем некоторое расширение марковской модели, для которого можно сформулировать эргодическую теорему, аналогичную п. 1.1.3, а именно определим конструктивно полумарковский процесс ([2], т.2).

Пусть X пространство состояний и $x_n \in X$, $n = 0, 1, \dots$ - марковская цепь с переходными вероятностями $p(x, dy)$; предполагается, что $p^{(k)}(x, A) \rightarrow \pi(A)$. Будем считать, что для любых $x, y \in X$ определена функция $\varphi_{x,y} : (0, 1) \rightarrow R_+$; пусть α равномерно распределена на $(0, 1)$, будем обозначать $F_{x,y}(t) = \mathbf{P}(\varphi_{x,y}(\alpha) > t)$; заметим, что $\int_0^1 \varphi_{xy}(u) du = \int_0^\infty F_{x,y}(t) dt$. Пусть α_n , $n = 1, 2, \dots$ - независимые в совокупности и от x_n равномерно-распределенные на $(0, 1)$ случайные величины, $\tau_n = \varphi_{x_{n-1}, x_n}(\alpha_n)$; $t_n = \sum_{i=1}^n \tau_i$, предполагаем, что $t_n \uparrow \infty$ п.н. Положим $x_t = x_n$ при $t_n \leq t < t_{n+1}$. Определенный таким образом процесс x_t называется полумарковским процессом.

Рассмотрим, например, систему (M/G/1), т.е. систему, у которой входной поток пуассоновский с параметром μ ; одно обслуживающее устройство; про время обслуживания η известно, что $\mathbf{P}(\eta > t) = H(t)$; очередь неограниченна. Пусть t_n - моменты прихода или ухода заявок. Построим полумарковский процесс, описывающий данную систему. Пусть $x_n = (z_n, \zeta_n)$, где z_n - количество заявок в момент времени t_n , а ζ_n - время, прошедшее с момента начала обслуживания текущей заявки или 0, если заявок нет. Опишем переход к следующему состоянию. Очевидно, что из состояния $x_n = (0, 0)$ цепь переходит в состояние $x_{n+1} = (1, 0)$, при этом время пребывания в x_n - экспоненциальное с показателем μ . Предположим, что $z_n > 0$ и пусть η_n - время обслуживания текущей заявки, s_n - время оставшееся до окончания обслуживания при условии $\eta_n > \zeta_n$, тогда $\mathbf{P}(s_n > s | \eta_n > \zeta_n) = H(s + \zeta_n)/H(\zeta_n)$; время γ_n от момента t_n до прихода следующей заявки распределено экспоненциально с параметром μ . Таким образом, мы можем промоделировать s_n и γ_n . Если $\gamma_n < s_n$, то $x_{n+1} = (z_n + 1, \zeta_n + \gamma_n)$; если $\gamma_n > s_n$, то $x_{n+1} = (z_n - 1, 0)$. Время пребывания $\tau_n = \min(\gamma_n, s_n)$. При таком опи-

сании переход осуществляется с помощью некоторого распределения $q_x(dy \times d\tau)$ и формально не соответствует конструкции полумарковского процесса, однако мы можем определить $p(x, dy) = q_x(dy \times R_+)$ и $F_{xy}(s) = q_x(dy \times (s, \infty))/p(x, dy)$. Таким образом, процесс $x_t = x_n$ при $t_n \leq t < t_{n+1}$ является полумарковским для системы $(M/G/1)$.

Вернемся к общему полумарковскому процессу. Будем предполагать, что $\max_{x,y}[1 - F_{xy}(s)] \rightarrow 0$ при $s \rightarrow 0$ и

$$J_1 = \int_{X^2} \pi(dx) p(x, dy) \int_0^\infty dt F_{x,y}(t) < \infty.$$

Покажем, что существует

$$\lim_T \frac{1}{T} \int_0^T dt 1_A(x_t). \quad (10)$$

Определим процесс $\bar{x}_t = (x_{n-1}, x_n, t - t_{n-1}) \in \bar{X} = X^2 \times R_+$, при $t_{n-1} \leq t < t_n$; иногда для n будем использовать обозначение n_t .

Покажем, что $\bar{x}_t$ является однородным марковским процессом. Действительно, σ -алгебры $\mathcal{F}_t$ определяются значениями x_i при $i \leq n_t$ и моментами скачков t_i при $i < n_t$. Для того, чтобы промоделировать значение процесса в момент времени $T > t$ при условии, что известна его траектория до момента t , надо промоделировать последовательности $x_{n+1}, \dots$ и $t_n, t_{n+1}, \dots$ до шага m , когда выполнено $t_{m-1} \leq T < t_m$, в этом случае $\bar{x}_T = (x_{m-1}, x_m, T - t_m)$. Распределение величины x_{n+1} равно $p(x_n, \cdot) = p(\bar{x}_{t,2}, \cdot)$ (здесь $\bar{x}_{t,2}$ обозначает вторую компоненту вектора $\bar{x}_t$), т.е. зависит только от $\bar{x}_t$. Величина $\tau_n - t$ при том, что известна траектория процесса $\bar{x}_s$, $s \leq t$, имеет функцию распределения $1 - Q_{\bar{x}_t}(s)$, где $Q_{\bar{x}}(s) = F_{x,y}(s + t)/F_{x,y}(t)$ при $\bar{x} = (x, y, t)$.

Покажем теперь, что стационарное распределение $\Pi(d\bar{x})$ для процесса $\bar{x}_t$ имеет вид

$$\Pi(d\bar{x}) = \Pi(dx \times dy \times dt) = c_1 \pi(dx) p(x, dy) F_{x,y}(t) dt. \quad (11)$$

где $c_1 = 1/J_1$ - константа нормировки.

Для этого достаточно показать, что для любой ограниченной функции $f(\bar{x})$ выполнено $\int_{\bar{X}} \Pi(d\bar{x}) \mathbf{E}_{\bar{x}} f(\bar{x}_s) = \int_{\bar{X}} \Pi(d\bar{x}) f(\bar{x})$. Обозначим левую часть $G(s) = \int_{\bar{X}} \Pi(d\bar{x}) \mathbf{E}_{\bar{x}} f(\bar{x}_s)$. Легко видеть, что при $\bar{x} = (x, y, t)$

$$G(s) = f(x, y, t + s) Q_{\bar{x}}(s) - \int_X p(y, dz) \int_0^s du Q_{\bar{x}}(u) \mathbf{E}_{y,z,0} f(\bar{x}_{s-u}). \quad (12)$$

Заметим, что $F_{x,y}(t) Q_{\bar{x}}(s) = F_{x,y}(t + s)$, поэтому, интегрируя первое слагаемое, получаем

$$c_1 \int_{\bar{X}} \pi(dx) p(x, dy) F_{x,y}(t + s) dt f(x, y, t + s) = c_1 \int_{X^2} \int_s^\infty \Pi(d\bar{x}) f(\bar{x}). \quad (13)$$

Далее, $\int_0^\infty dt du F_{x,y}(t + u) = -F_{x,y}(u) du$, поэтому интегрируя второе слагаемое, получим

$$\begin{aligned} & c_1 \int_{X^3} \pi(dx) p(x, dy) p(y, dz) \int_0^s du F_{x,y}(u) \mathbf{E}_{y,z,0} f(\bar{x}_{s-u}) = \\ & = c_1 \int_{X^3} \int_0^s du \pi(dx) p(x, dy) p(y, dz) \mathbf{E}_{y,z,0} f(\bar{x}_{s-u}) + \\ & + c_1 \int_{X^3} \int_0^s du \pi(dx) p(x, dy) p(y, dz) [F_{x,y}(u) - 1] \mathbf{E}_{y,z,0} f(\bar{x}_{s-u}) \end{aligned} \quad (14)$$

Второе слагаемое в (14) не превосходит по абсолютной величине

$$\Delta_s = c_1 \|f\| \int_0^s du \int_{X^2} \pi(dx) p(x, dy) [1 - F_{xy}(u)] = o(s) \quad \text{при } s \rightarrow 0.$$

В первом слагаемом в (14) можно проинтегрировать по x , тогда получим $\int_X \pi(dx) p(x, dy) = \pi(dy)$. Кроме того, в силу (12) имеем оценку

$$|\mathbf{E}_{y,z,0} f(\bar{x}_{s-u}) - f(y, z, s - u) F_{yz}(s - u)| \leq \|f\| [1 - Q_{y,z,0}(s - u)].$$

Итак, первое слагаемое в (14) может быть представлено в виде

$$\int_{X^2} \pi(dy) p(y, dz) \int_0^s du F_{yz}(s - u) f(y, z, s - u) + o(s). \quad (15)$$

Складывая (13) и (15), получаем $G(s) = \int_{\bar{X}} \Pi(d\bar{x})f(\bar{x}) + o(s)$. Отсюда следует, $\lim_{s \rightarrow 0} G(s) = \int_{\bar{X}} \Pi(d\bar{x})f(\bar{x})$. С другой стороны, подставляя в $G(s)$ вместо $f(\bar{x})$ функцию $\mathbf{E}_{\bar{x}}f(\bar{x}_T)$ имеем $G(T+s) - G(T) = o(s)$. Устремляя s к нулю, получим $G'(T) = 0$.

Таким образом, формула (11) определяет стационарное распределение и можно построить меру $\mathbf{P}_{\Pi}(\cdot) = \int_{\bar{X}} \mathbf{P}_{\bar{x}}(\cdot)$, относительно которой процесс $\bar{x}_t$ будет стационарным. Из п.1.1.3 следует, что п.н. существует предел

$$\lim_T \frac{1}{T} \int_0^T dt f(\bar{x}_t).$$

Если положить $f(\bar{x}) = 1_A(\bar{x}_1)$, то мы получим существование предела в (10). Теперь найдем выражение для этого предела. Как и в предыдущем параграфе введем марковскую цепь $\bar{x}_n = (x_n, x_{n+1}, \tau_{n+1})$, переходные вероятности которой из точки $\bar{z} = (z_1, z_2, t)$ в $d\bar{y} = dy_1 \times dy_2 \times ds$ равны

$$\bar{p}(\bar{z}, d\bar{y}) = \delta_{z_2}(dy_1)p(z_2, dy_2)d_s(1 - F_{y_1, y_2}(s)).$$

Переходные вероятности за k шагов имеют вид

$$\bar{p}^{(k)}(\bar{z}, d\bar{y}) = p^{(k-1)}(z_2, dy_1)p(y_1, dy_2)d_s(1 - F_{y_1, y_2}(s)),$$

то есть $\bar{p}^{(k)}(\bar{z}, d\bar{y})$ стремится к стационарному распределению

$$\bar{\pi}(d\bar{y}) = \pi(dy_1)p(y_1, dy_2)d_s(1 - F_{y_1, y_2}(s)).$$

Из утверждения 1.1.3 следует, что

$$\frac{1}{n} \sum_{i < n} f(\bar{x}_i) \rightarrow \int_{\bar{X}} \bar{\pi}(d\bar{x})f(\bar{x}) = J(f) \quad \text{п.н.} \quad (16)$$

при условии $\int_{\bar{X}} \bar{\pi}(d\bar{x})|f(\bar{x})| < \infty$.

Представим в виде суммы интеграл $\int_0^{t_n} dt 1_A(x_t) = \sum_{i < n} 1_A(x_i)\tau_{i+1}$ и в (16) положим $f(\bar{x}) = 1_A(\bar{x}_1)$, тогда предел в (10) равен

$$\lim_n \frac{1}{t_n} \int_0^{t_n} 1_A(x_t)dt = \frac{J(1_A)}{J(1)}.$$

2 Статистическое моделирование

В предыдущем разделе было показано, что важные для практики характеристики системы массового обслуживания, задаваемые в виде

$$\Pi(A) = \lim_T \frac{1}{T} \int_0^T dt \, 1_A(x_t), \quad (17)$$

могут быть представлены как отношение функционалов от некоторого стационарного распределения марковской цепи:

$$\Pi(A) = \frac{\int_X \pi(dx) h_0(x)}{\int_X \pi(dx) h_1(x)} = \frac{J_0}{J_1}. \quad (18)$$

Поскольку аналитическое решение возможно лишь в простейших случаях, то практически единственным методом, позволяющим вычислять характеристики сложных СМО, является метод Монте-Карло. Основой применения метода этого метода при вычислении функционалов вида $J = \int_X \pi(dx) h(x)$ является утверждение 1.1.3.

Рассмотрим вопрос об оценке погрешности. Предположим, что мы можем разбить траекторию $\{x_n\}$ на независимые одинаково распределенные циклы $\{\bar{x}_k\}$, где $\bar{x}_k = (x_{n_{k-1}+1}, \dots, x_{n_k})$. Далее обозначим $\xi_k = \sum_{x_i \in \bar{x}_k} h_0(x_i)$, $\eta_k = \sum_{x_i \in \bar{x}_k} h_1(x_i)$, и пусть a , b - математические ожидания ξ_k и η_k соответственно. Считая что на траектории имеется N циклов, имеем

$$D = \mathbf{E} \left(\frac{\sum_{i=1}^n h_0(x_i)}{\sum_{i=1}^n h_1(x_i)} - \frac{J_0}{J_1} \right)^2 = \mathbf{E} \left(\frac{\bar{\xi}}{\bar{\eta}} - \frac{a}{b} \right)^2,$$

где $\bar{\xi} = \frac{1}{N} \sum_k \xi_k$, $\bar{\eta} = \frac{1}{N} \sum_k \eta_k$. Раскладываем $\bar{\xi}/\bar{\eta}$ в ряд Тейлора вокруг точки a/b , получим

$$\frac{\bar{\xi}}{\bar{\eta}} \approx \frac{a}{b} + \frac{\bar{\xi} - a}{b} - \frac{a(\bar{\eta} - b)}{b^2},$$

откуда, предполагая конечность соответствующих моментов, имеем

$$D \approx \frac{1}{N} \left[\frac{\mathbf{E}(\xi - a)^2}{b^2} + \frac{a^2 \mathbf{E}(\eta - b)^2}{b^4} - \frac{2a \mathbf{E}(\xi - a)(\eta - b)}{b^3} \right]. \quad (19)$$

Повышение эффективности. Определим относительную погрешность как $q_n = \mathbf{E}^{1/2}(\bar{J}_n - J)^2/J$. Если J мало, то относительная погрешность может быть велика; действительно, пусть, например, x_i - независимые величины, $h(x) = 1_A(x)$, тогда $q_n = \sqrt{1 - \bar{J}}/\sqrt{n\bar{J}}$ и для достижения нужной точности надо брать n достаточно большим.

Чтобы повысить эффективность моделирования, можно применить метод существенной выборки. Однако, простейшая оценка, построенная по этому методу, оказывается несостоятельной [5]; покажем это. Пусть $r(x, dy)$ - переходные вероятности, такие что существует стационарное распределение $\sigma(dy)$: $\int_X \sigma(dx) r(x, dy) = \sigma(dy)$; $p(x, \cdot) \ll r(x, \cdot)$ и $\rho(x, y) = p(x, dy)/r(x, dy)$. Пусть ξ_i - марковская цепь с переходными вероятностями $r(x, dy)$, определим оценку

$$\bar{J}_n(r) = \frac{1}{n} \sum_{k=1}^n \rho(\xi_1, \xi_2) \dots \rho(\xi_{k-1}, \xi_k) h(\xi_k).$$

Очевидно, что имеет место несмещенность по отношению к оценке $\bar{J}_n$, т.е. $\mathbf{E} \bar{J}_n(r) = \mathbf{E} \bar{J}_n$. Рассмотрим вопрос о состоятельности $\bar{J}_n(r)$. Представим оценку в виде

$$\bar{J}_n(r) = \frac{1}{n} \sum_{k=1}^n h(\xi_k) e^{(k-1)\bar{I}_{k-1}} \quad (20)$$

где $\bar{I}_k = \frac{1}{k} \sum_{i=1}^k \ln \rho(\xi_i, \xi_{i+1})$ Обозначим $z_i = (\xi_i, \xi_{i+1})$ - марковская цепь в $X^2 = \{(x, y)\}$. Легко проверить, что она имеет стационарное распределение $\bar{\sigma}(dz) = \sigma(dx) r(x, dy)$. Обозначим $f(z) = \ln \rho(x, y)$. По закону больших чисел для марковских процессов имеем

$$\bar{I}_k \rightarrow \int_{X^2} \bar{\sigma}(dz) f(z) = a.$$

Применяя неравенство Йенсена, получаем

$$a = \int_{X^2} \bar{\sigma}(dz) f(z) \leq \ln \int_{X^2} \sigma(dx) r(x, dy) \rho(x, y) = \ln 1 = 0,$$

причем, если $\rho(x, y) \neq \text{const}$, то $a < 0$. Таким образом, если $r \neq p$, то, как следует из (20), $\bar{J}_n(r) \rightarrow 0$ п.н.

2.1 Ответвление траекторий.

Метод ответвления траекторий был предложен в [5] и затем развит в [6]. Суть метода состоит в том, что моделируется основная траектория с исходными переходными вероятностями $p(x, dy)$, а также ответвления с произвольными, вообще говоря, вероятностями перехода $r_i(x, dy)$. Будем предполагать, что $r_i(x, X) \leq 1$, т.е. возможен обрыв траектории; вероятность обрыва на i -ом шаге равна $g_i(x) = 1 - r_i(x, X)$. Обозначим через $\bar{X}$ - множество траекторий, конечных и бесконечных; для обрывающейся траектории $\bar{x} = (x_0, \dots, x_n)$ положим $\tau_0(\bar{x}) = n + 1$ - момент обрыва; если траектория $\bar{x}$ бесконечная, положим $\tau_0(\bar{x}) = \infty$. Пусть $\mathbf{R}_{x_0}(\cdot)$ - мера на $\bar{X}$, порожденная переходными вероятностями $r_i(x, dy)$, $\mathbf{M}_{x_0}$ - математическое ожидание по мере $\mathbf{R}_{x_0}$; меру и математическое ожидание, порожденную исходными переходными вероятностями будем обозначать $\mathbf{P}_{x_0}$ и $\mathbf{E}_{x_0}$. Далее построим марковскую цепь со значениями в $\bar{X}$; для этого определим переходные вероятности $\bar{P}(\bar{x}, \cdot) = \int_X p(x_0, dy_0) \mathbf{R}_{y_0}(\cdot)$. Очевидно, что стационарное распределение для $\bar{P}(\bar{x}, \cdot)$ равно $\Pi(\cdot) = \int_X \pi(dx) \mathbf{R}_x(\cdot)$.

Плотности исходных переходных вероятностей по моделируемым переходным вероятностям на i -ом шаге равны $\rho_i(x, y) = p(x, dy)/r_i(x, dy)$, а плотность после i шагов

$$\rho_i(\bar{x}) = 1_{\{i < \tau_0\}} \prod_{j=1}^i \rho_j(x_{j-1}, x_j). \quad (21)$$

Введем дополнительно величины $\alpha_i(x)$ - вес i -го шага в точке x , предполагая, что $\sum_i \alpha_i(x) = 1$, а также величины $w_i(\bar{x})$, которые задают оценки так называемого «единичного класса» [4], определяемого

равенством $1_{\{i < \tau_0\}} \mathbf{M}_x(w_i | \mathcal{F}_i) = 1_{\{i < \tau_0\}}$. Наиболее известными оценками данного класса являются оценки «по столкновениям», для которых $w_i(\bar{x}) = 1$, и оценки «по поглощениям», где $w_i(\bar{x}) = 1_{\{i+1=\tau_0\}}/g_i(x_i)$. Покажем, что последние действительно являются оценками «единичного класса»:

$$\begin{aligned} 1_{\{i < \tau_0\}} \mathbf{M}(w_i | \mathcal{F}_i) &= 1_{\{i < \tau_0\}} \mathbf{M}_x \left(1_{\{\tau_0=i+1\}}/g(x_i) | \mathcal{F}_i \right) = \\ &= 1_{\{i < \tau_0\}}/g(x_i) \mathbf{R}_x \left(\tau_0 = i+1 | \mathcal{F}_i \right) = \\ &= 1_{\{i < \tau_0\}}/g(x_i) [1 - r_i(x_i, X)] = 1_{\{i < \tau_0\}}. \end{aligned}$$

Теперь определим оценку на траектории $\bar{x}$: $H(\bar{x}) = \sum_i \rho_i w_i \alpha_i(x_i) h(x_i)$. и покажем, что $\int \Pi(d\bar{x}) H(\bar{x}) = J$. Сначала вычислим $\varphi(x) = \mathbf{M}_x H(\bar{x})$:

$$\begin{aligned} \varphi(x) &= \mathbf{M}_x \sum_i \rho_i w_i \alpha_i(x_i) h(x_i) = \sum_i \mathbf{M}_x \rho_i \alpha_i(x_i) h(x_i) \mathbf{M}_x \left(w_i | \mathcal{F}_i \right) = \\ &= \sum_i \mathbf{M}_x \rho_i \alpha_i(x_i) h(x_i) = \sum_i \mathbf{E}_x 1_{\{i < \tau_0\}} \alpha_i(x_i) h(x_i) = \\ &= \sum_i \mathbf{E}_x \alpha_i(x_i) h(x_i) = \sum_i \int_X p^{(i)}(x, dy) \alpha_i(y) h(y). \end{aligned}$$

Поскольку $\int_{\bar{X}} \Pi(d\bar{x}) H(\bar{x}) = \int_X \pi(dx) \varphi(x)$, то, учитывая, что $\pi(dx)$ является стационарным распределением для $p(x, dy)$, имеем

$$\begin{aligned} \int \Pi(d\bar{x}) H(\bar{x}) &= \sum_i \int_{X^2} \pi(dx) p^{(i)}(x, dy) h(y) \alpha_i(y) = \\ &= \int_X \pi(dy) h(y) \sum_i \alpha_i(x) = \int_X \pi(dy) h(y) = J. \end{aligned}$$

Пусть имеется $\bar{x}_k$ - марковская цепь с переходными вероятностями $\bar{P}(\bar{x}, \cdot)$, тогда оценка $\bar{J}_n = \frac{1}{n} \sum_{k=1}^n H(\bar{x}_k)$ является состоятельной для J и, кроме того, можно выбрать переходные вероятности $r(x, dy)$

так, чтобы уменьшить дисперсию этой оценки. Мы будем рассматривать ограниченные функции $h(x)$, тогда можно считать, что $h(x)$ строго положительна. Действительно, если $h(x) > -c$, то можно взять $h_c(x) = h(x) + c > 0$ и построить оценку $\bar{J}_n(h_c) - c$, которая является состоятельной для J и имеет ту же дисперсию, что и оценка $\bar{J}_n(h_c)$.

Вычислим дисперсию оценки. Пусть процесс $\bar{x}_k$ находится в стационарном режиме, тогда

$$\begin{aligned} \mathbf{D}\bar{J}_n &= \mathbf{M} \left[\frac{1}{n} \sum_{k=1}^n H(\bar{x}_k) - J \right]^2 = \\ &= \frac{1}{n} \mathbf{M}[H(\bar{x}) - J]^2 + \frac{2}{n^2} \sum_{k < m} \mathbf{M}[H(\bar{x}_k) - J][H(\bar{x}_m) - J] = \\ &= \frac{1}{n} \int_X \pi(dx) [\mathbf{M}_x H^2 - J^2] + \\ &+ \frac{2}{n^2} \sum_{k < m} \int_X \pi(dx) \mathbf{M}_x [H - J] \cdot p^{(m-k)}(x, dy) \mathbf{M}_y [H - J] \end{aligned}$$

Обозначим $\psi(x) = \mathbf{M}_x H^2$, тогда $\mathbf{D}\bar{J}_n$ можно представить в виде

$$\begin{aligned} \mathbf{D}\bar{J}_n &= \frac{1}{n} \int_X \pi(dx) [\psi(x) - J^2] + \\ &+ \frac{2}{n^2} \sum_{k=1}^{n-1} \int_X \pi(dx) [\varphi(x) - J][P_x^k \varphi - J](n - k). \end{aligned} \tag{22}$$

Заметим, что $|P_x^k \varphi - J| \leq C \sup_A |p^{(k)}(x, A) - \pi(A)| \leq C \delta^k$, поэтому второе слагаемое в (22) равно

$$\frac{2}{n} \sum_{k=1}^{n-1} \int_X \pi(dx) [\varphi(x) - J][P_x^k \varphi - J] + O\left(\frac{1}{n^2}\right).$$

Таким образом, для уменьшения дисперсии нужно уменьшить $\psi(x)$. Рассмотрим различные случаи выбора $\alpha_i(x)$. Заметим, что для доказательства оптимальности выбора переходных вероятностей $r_i(x, dy)$,

достаточно показать, что $H(\bar{x}) = \varphi(x_0)$ п.н., так как в этом случае $\psi(x) = \varphi^2(x)$.

2.1.1. Пусть заданы целые числа $L \leq M$, положим $\alpha_i = \frac{1}{N} 1_{[L, M]}(i)$, где $N = M - L + 1$, тогда $\varphi(x) = \sum_{i=L}^M P_x^i h / N$. Рассмотрим оценки по столкновениям: $w_i = 1$; покажем, что оптимальные переходные плотности в этом случае равны

$$r_k^*(x, dy) = c_k(x) p(x, dy) \sum_{i=L \wedge k}^M P_y^{i-k} h, \quad r_{M+1}^* = 0, \quad (23)$$

где $c_k(x)$ – коэффициент нормировки.

Действительно, при таком выборе r_k траектории имеют длину $M+1$ и на этих траекториях $H^*(\bar{x}) = \sum_{i=L}^M \rho_i^* h(x_i) / N$. Сначала покажем, что для $k \geq L$ имеет место равенство

$$\sum_{i=k}^M \rho_i^* h(x_i) = \rho_k^* \sum_{i=k}^M P_{x_k}^{i-k} h. \quad (24)$$

При $k = M$ представление (24) верно, поскольку левая и правая части равны $\rho_M^* h(x_M)$. Сделаем индукционное предположение, что (24) верно для $L < k \leq M$, тогда имеем

$$\begin{aligned} \sum_{i=k-1}^M \rho_i^* h(x_i) &= \rho_{k-1}^* h(x_{k-1}) + \rho_k^* \sum_{i=k}^M P_{x_k}^{i-k} h = \\ &= \rho_{k-1}^* \left[h(x_{k-1}) + \frac{p(x_{k-1}, dx_k)}{r_k^*(x_{k-1}, dx_k)} \sum_{i=k}^M P_{x_k}^{i-k} h \right]. \end{aligned}$$

Подставляя сюда выражение (23), получаем

$$\sum_{i=k-1}^M \rho_i^* h(x_i) = \rho_{k-1}^* \sum_{i=k-1}^M P_{x_{k-1}}^{i-(k-1)} h.$$

Таким образом, (24) верно при любом $k \geq L$, откуда следует, что

$$H^*(\bar{x}) = \rho_L^* \sum_{i=L}^M P_{x_L}^{i-L} h.$$

Непосредственно проверяется, что $\rho_L^* = \varphi(x_0) / \sum_{i=L}^M P_{x_L}^{i-L} h$ и, значит, $H^*(\bar{x}) = \varphi(x_0)$, что доказывает оптимальность r_n^* .

2.1.2. Рассмотрим случай $\alpha_i = (1 - \gamma)\gamma^i$, $\gamma < 1$, тогда

$$\varphi(x) = \sum_{i=0}^{\infty} \gamma^i P_x^i h (1 - \gamma) = \sum_{i=0}^{\infty} K_x^i f,$$

где $k(x, dy) = \gamma p(x, dy)$, $f(x) = (1 - \gamma)h(x)$, т. е. $\varphi(x)$ является решением уравнения $\varphi(x) = K_x \varphi + f(x)$.

2.1.2.a. Рассмотрим оценку по поглощениям. В этом случае

$$H(\bar{x}) = \rho_1(x_0, x_1) \dots \rho_n(x_{n-1}, x_n) f(x_n) / g(x_n),$$

где $\rho_i(x, y) = k(x, dy) / r_i(x, dy)$. Покажем, что оптимальные переходные вероятности равны $r_i^*(x, dy) = k(x, dy) \varphi(y) / \varphi(x)$. Действительно, для вероятности поглощения имеем выражение

$$g_i^*(x) = 1 - \frac{1}{\varphi(x)} \int_X k(x, dy) \varphi(y) = \frac{f(x)}{\varphi(x)}.$$

Подставляя его в $H(\bar{x})$, где $\bar{x} = (x_0, \dots, x_n)$, получаем

$$H^*(\bar{x}) = \frac{k(x_0, dx_1)}{r^*(x_0, dx_1)} \dots \frac{k(x_{n-1}, dx_n)}{r^*(x_{n-1}, dx_n)} \frac{f(x_n)}{g^*(x_n)} = \varphi(x_0),$$

откуда и следует оптимальность.

2.1.2.b. Рассмотрим теперь оценку по столкновениям. Чтобы исключить случай бесконечных траекторий, наложим условие

$$r(x, X) \leq c < 1. \quad (25)$$

Сначала найдем представление для $\psi(x)$. Для этого определим функцию $s(x) = L_x \varphi^2 - (K_x \varphi)^2$ и заметим, что $s(x) \geq 0$; далее положим $\sigma(x) = \sum_{i=0}^{\infty} L_x^i s \geq 0$. Будем предполагать, что $r(x, dy)$ таковы, что $\sigma(x) < \infty$ для любого x . Заметим, что $H = \sum_{i=0}^{\infty} \rho_i f(x_i)$, где ρ_i определяются формулой (21). Этот ряд можно также представить в виде $H = \varphi(x_0) + \sum_{i=1}^{\infty} \xi_i$, где $\xi_i = \rho_{i-1} [1_{\{i < \tau\}} \rho(x_{i-1}, x_i) \varphi(x_i) - K_{x_{i-1}} \varphi]$. Нетрудно видеть, что $\mathbf{E}(\xi_i | \mathcal{F}_{i-1}) = 0$, и $\mathbf{E}_x \xi_i^2 = L_x^i s$. Из этих равенств и сходимости ряда $\sigma(x)$ следует ([2], т. 1, с.89), что ряд H сходится в $\mathbf{L}_2$ и $\psi(x) = \varphi^2(x) + \sigma^2(x)$. Заметим, что следствием этого представления является уравнение

$$\psi(x) = L_x \psi + a(x). \quad (26)$$

где $a(x) = \varphi^2(x) - (K_x \varphi)^2$. Пусть c таково, что

$$\gamma^2 < c. \quad (27)$$

Покажем, что в этом случае оптимальные переходные вероятности имеют вид

$$r^*(x, dy) = ck(x, dy)b(y)/K_x b, \quad (28)$$

где $b(x)$ - неотрицательное решение уравнения

$$b^2(x) = \frac{1}{c} (K_x b)^2 + a(x). \quad (29)$$

Сначала покажем, что уравнение (29) имеет единственное неотрицательное решение. Определим оператор A в пространстве ограниченных функций выражением

$$A_x b = \left[\frac{1}{c} K_x b + a(x) \right]^{1/2}, \quad (30)$$

тогда из неравенства

$$|(\beta_1 + \alpha)^{1/2} - (\beta_2 + \alpha)^{1/2}| \leq |\beta_1 - \beta_2| \quad (31)$$

верного для $\alpha \geq 0$, следует, что

$$|A_x b_1 - A_x b_2| \leq \frac{1}{\sqrt{c}} K_x |b_1 - b_2| \leq \frac{\gamma}{\sqrt{c}} \|b_1 - b_2\|,$$

где $\|b\| = \sup_x |b(x)|$. Отсюда при $\gamma^* = \gamma/\sqrt{c} < 1$ выполнено

$$\|Ab_1 - Ab_2\| < \gamma^* \|b_1 - b_2\|,$$

т.е. A - сжимающий оператор в полном пространстве и, значит, уравнение $b(x) = A_x b$ имеет единственное решение. Покажем, что $b^2(x)$ является оценкой снизу для $\psi(x)$. В силу неравенства Коши-Буняковского

$$\begin{aligned} b^2(x) &= \frac{1}{c} \left(\int_X k(x, dy) \sqrt{\rho(x, y)} \frac{b(y)}{\sqrt{\rho(x, y)}} \right)^2 + a(x) \leq \\ &\leq \frac{1}{c} \int_X k(x, dy) \rho(x, y) b^2(y) \int_X \frac{k(x, dy)}{\rho(x, y)} + a(x) = \\ &= \frac{1}{c} \int_X l(x, dy) b^2(y) \int_X r(x, dy) + a(x) = \\ &\leq L_x b^2 + a(x) \leq \dots \leq \sum_{i=0}^{\infty} L_x^i a = \psi(x). \end{aligned}$$

Теперь покажем, что $\psi^*(x) \leq b^2(x)$. Обозначим $L^* = L(r^*)$, тогда из (29) следует, что $(L^*)_x^n a \leq C(\gamma^*)^n$, таким образом, ряд $\sum_n (L^*)_x^n a$ сходится и $\psi^*(x)$ является единственным решением уравнения

$$\psi^*(x) = L_x^* \psi^* + a(x) = \frac{1}{c} K_x b K_x \psi^* / b + a(x),$$

т.е. $\psi^*(x) = b^2(x)$.

2.1.3. Рассмотрим еще одну оценку, которая формально может быть отнесена к рассматриваемой схеме.

Пусть $\sigma(x)$ - произвольная функция, $H(x) = h(x) + \sigma(x) - P_x \sigma$, тогда $\int \pi(dx) H(x) = J$, т.е. в оценке $\bar{J}_n$ мы можем заменить $h(x)$ на $H(x)$, используя $\sigma(x)$ для повышения эффективности.

Из (2) следует, что $|P_x^k h - J| \leq C\delta^k$, и, таким образом, сходится ряд $\sum_{k=0}^{\infty} P_x^k (J - h) = \sigma^*(x)$. Нетрудно видеть, что $\sigma^*(x)$ удовлетворяет уравнению $\sigma(x) = P_x \sigma + J - h(x)$ и для $H^*(x)$ верно равенство $H^*(x) = J$, что доказывает оптимальность выбора σ .

Заметим, что в отличие от предыдущих оценок уменьшение дисперсии достигается не за счет моделирования переходных вероятностей, отличных от исходных, а за счет добавления некоторой функции, интеграл от которой известен (в данном случае равен нулю); поэтому данная оценка может считаться аналогом метода выделения главной части, применяемого при вычислении интегралов [4].

2.2 Вложенная марковская цепь

Рассмотрим еще одну схему [7]. Пусть на траекториях $\bar{x} = (x_0, x_1, \dots)$ задан марковский момент $\bar{\tau}(\bar{x})$; на основе этого момента построим рекуррентно последовательность марковских моментов $\bar{\tau}_n$: $\bar{\tau}_1(\bar{x}) = \bar{\tau}(\bar{x})$, $\bar{\tau}_n(\bar{x}) = \bar{\tau}_{n-1}(\bar{x}) + \bar{\tau}(\theta_{\bar{\tau}_{n-1}} \bar{x})$, где $\theta_k \bar{x} = (x_k, x_{k+1}, \dots)$ - траектория, получающаяся сдвигом траектории $\bar{x}$ на k шагов.

Обозначим $\xi_n = x_{\tau_n}$ и предположим, что у ξ_n имеется стационарное распределение $\bar{\pi}(dx)$; и $\int_X \bar{\pi}(dx) \mathbf{E}_x \bar{\tau} < \infty$. Покажем, что имеет место равенство

$$J = \int_X \pi(dx) h(x) = \frac{\int_X \bar{\pi}(dx) \mathbf{E}_x \sum_{k=0}^{\bar{\tau}-1} h(x_k)}{\int_X \bar{\pi}(dx) \mathbf{E}_x \bar{\tau}}. \quad (32)$$

Обозначим $I(n) = \sum_{k=0}^{n-1} h(x_k)$ и для $s > 0$ положим

$$F_x(s) = \mathbf{E}_x \sum_{n=0}^{\infty} e^{s(iI(n)-n)} = \mathbf{E}_x \sum_{n=0}^{\bar{\tau}-1} e^{s(iI(n)-n)} + \mathbf{E}_x e^{s(iI(\bar{\tau})-\bar{\tau})} F_{x_{\bar{\tau}}}(s)$$

Поскольку $\bar{\pi}(dx)$ является стационарным распределением для переходных вероятностей $\bar{p}(x, A) = \mathbf{P}_x(x_{\bar{\tau}} \in A)$, то проинтегрировав предыдущее равенство по $\bar{\pi}(dx)$, получим

$$\int_X \bar{\pi}(dx) \mathbf{E}_x \sum_{n=0}^{\bar{\tau}-1} e^{s(iI(n)-n)} = \int_X \bar{\pi}(dx) \mathbf{E}_x [1 - e^{s(iI(\bar{\tau})-\bar{\tau})}] F_{x_{\bar{\tau}}}(s). \quad (33)$$

Устремим s к нулю; поскольку $|e^{s(iI(\bar{\tau})-\bar{\tau})} - 1| \leq 2$ и $\int_X \bar{\pi}(dx) \mathbf{E}_x \bar{\tau} < \infty$, то можно перейти к пределу под знаком интеграла в левой части (33) и мы получим $\int_X \bar{\pi}(dx) \mathbf{E}_x \bar{\tau}$. Подынтегральное выражение в правой части представим в виде произведения $f_1(s)f_2(s)$, где $f_1(s) = (1 - e^{s(iI(\bar{\tau})-\bar{\tau})})/s$, $f_2(s) = sF_{x_{\bar{\tau}}}(s)$. Поскольку $|f_1(s)| \leq C\bar{\tau}$ и $|f_2(s)| \leq s/(1 - e^{-s}) = O(1)$, то можно перейти к пределу под знаком интеграла и в правой части (33), при этом $\lim_s f_1(s) = \tau - iI(\tau)$. Покажем, что $\lim_s f_2(s) = 1/(1 - iJ)$.

Обозначим $\Delta(s, x) = |sF_x(s) - (1 - iJ)^{-1}|$ и воспользуемся равенством $\sum_{n=0}^{\infty} e^{ns(iJ-1)} = (1 - e^{s(iJ-1)})^{-1}$, тогда

$$\Delta(s, x) \leq \left| \frac{1}{1 - iJ} - \frac{s}{1 - e^{s(iJ-1)}} \right| + s \sum_{n=0}^{\infty} \mathbf{E}_x |e^{s(iI(n)-n)} - e^{sn(iJ-1)}|.$$

Первое слагаемое стремится к нулю при $s \rightarrow 0$; второе слагаемое при произвольных N и δ не превосходит

$$\begin{aligned} & s \sum_{n=0}^{N-1} \mathbf{E}_x |e^{s(iI(n)-n)} - e^{sn(iJ-1)}| + s \sum_{n=N}^{\infty} e^{-ns} \mathbf{E}_x |e^{sni(I(n)/n-J)} - 1| \leq \\ & \leq 2sN + 2s \sum_{n=N}^{\infty} e^{-ns} \mathbf{P}_x(|I(n)/n - J| > \delta) + 2s \sum_{n=N}^{\infty} e^{-ns} sn\delta = \sum_{i=1}^3 \Delta_i. \end{aligned}$$

В последнем выражении $\Delta_3 \leq Cs^2(1 - e^{-s})^{-2} \leq C\delta$ при $s \leq 1$ и, значит, при заданном ε можно выбрать δ таким, чтобы $\Delta_3 \leq \varepsilon/2$. Поскольку $\mathbf{P}_x(|I(n)/n - J| > \delta) \rightarrow 0$ при $n \rightarrow \infty$, то можно выбрать N , так чтобы $\Delta_2 \leq \varepsilon/2$. Устремляя s к нулю, получим, что $\lim_s \Delta(s, x) \leq \varepsilon$, а поскольку ε произвольно, то $\lim_s sF_x(s) = (1 - iJ)^{-1}$. Переходя к пределу в (33), получим $(1 - iJ)\mathbf{E}_x \bar{\tau} = \mathbf{E}_x(\bar{\tau} - iI(\bar{\tau}))$, откуда следует (32).

Рассмотрим случай, когда $\bar{\tau}$ - момент достижения некоторого множества D и $\int_X \pi(dx) \mathbf{E}_x \bar{\tau} < \infty$; покажем, что в этом случае $\pi(D) > 0$, $\bar{\pi}(B) = \pi(B|D)$ и $\int_X \bar{\pi}(dx) \mathbf{E}_x \bar{\tau} = 1/\pi(D)$. Действительно, обозначим $G = X \setminus D$, $k(x, y) = p(x, y)1_G(y)$ и для произвольного множества B положим $p_B(x) = P(x, B \cap D)$, тогда $\mathbf{P}_x(x_{\bar{\tau}} \in B) = \sum_{n=0}^{\infty} K_x^n p_B$. Поскольку при $n > 0$

$$\int_D \pi(dx) K_x^n p_B = \int_G \pi(dx) K_x^{n-1} p_B - \int_G \pi(dx) K_x^n p_B, \quad (34)$$

то $\int_D \pi(dx) \mathbf{P}_x(x_{\bar{\tau}} \in B) = \pi(B \cap D)$. Из (34) следует также, что

$$\int_D \pi(dx) \mathbf{E}_x \bar{\tau} = \int_D \pi(dx) \sum_{n=1}^{\infty} n K_x^{n-1} p_D = \pi(D) + \sum_{n=0}^{\infty} \int_G \pi(dx) K_x^n p_D = 1,$$

откуда и следуют утверждения.

Укажем пример, который может быть сведен к моменту достижения. Предположим, что относительно некоторой вероятностной меры ν существуют плотности $p(x, dy)/\nu(dy) = p(x, y)$, тогда существует плотность $\pi(dx)/\nu(dx)$, которую будем обозначать $\pi(x)$. Предположим, что $p(x, y) \geq \gamma(x)$, причем $\int_X \pi(dx) \gamma(x) > 0$. Наряду с последовательностью x_i зададим случайную последовательность величин α_i , равномерно распределенных на $[0, p(x_{i-1}, x_i)]$. Последовательность $v_i = (x_{i-1}, x_i, \alpha_i)$ образует марковскую цепь со значениями в $V = \{(x, y, \alpha)\}$; эта последовательность имеет стационарное распределение, которое мы также будем обозначать $\pi(dv)$, с плотностью $\pi(v) = \pi(x) 1_{[0, p(x, y)]}(\alpha)$ относительно $\nu(dx)\nu(dy)d\alpha$. Положим $D = \{v : \alpha \leq \gamma(x)\}$, $\bar{\tau}$ - момент достижения множества D . Заметим, что $\pi(D) = \int_X \pi(dx) \gamma(x)$ и, согласно предыдущим утверждениям, стационарная плотность соответствующей вложенной цепи $\bar{\pi}(v) = 1_D(v) \pi(v) / \pi(D)$. Если положить $h(v) = h(y)$, то получим, что $\int_V \pi(dv) h(v) = \int_X \pi(dx) h(y)$ и, значит, для J можно использовать представление (32): $J = \int_V \bar{\pi}(dv) \mathbf{E}_v \sum_{k=0}^{\bar{\tau}-1} h(v_k) / \int_V \bar{\pi}(dv) \mathbf{E}_v \bar{\tau}$. Рассмотрим вопрос о моделировании распределения $\bar{\pi}(dv)$ или, другими словами, о моделировании v_0 . Поскольку $h(v_0)$ зависит только от y_0 , то достаточно промоделировать только y_0 , плотность распределения которой относительно меры ν равномерное, т.к. $\int_X \pi(dx) \int_0^{\gamma(x)} d\alpha / \pi(D) = 1$.

Из (32) получаем выражение для (18):

$$\frac{J_0}{J_1} = \frac{\int_X \bar{\pi}(dx) \mathbf{E}_x \sum_{i=0}^{\bar{\tau}-1} h_0(x_i)}{\int_X \bar{\pi}(dx) \mathbf{E}_x \sum_{i=0}^{\bar{\tau}-1} h_1(x_i)} = \frac{\int_X \bar{\pi}(dx) \varphi_0(x)}{\int_X \bar{\pi}(dx) \varphi_1(x)},$$

где $\varphi_k(x) = \mathbf{E}_x \sum_{i=0}^{\bar{\tau}-1} h_k(x_i)$, $k = 0, 1$.

Далее будем предполагать, что $\bar{\tau}$ - момент достижения множества D и $G = X \setminus D$. Определим оператор K с ядром $k(x, dy) = p(x, dy) 1_G(y)$ и предположим, что

$$\mathbf{P}_x(\bar{\tau} > n) = k^{(n)}(x, X) \leq C \gamma^n. \quad (35)$$

Пусть заданы переходные вероятности $r_k(x, dy)$, которые определяют на траекториях вероятности $\mathbf{P}_x^{(r_k)}$ и математическое ожидание $\mathbf{E}_x^{(r_k)}$. Введем оценки $H_k(\bar{x}) = \sum_{i < \tau} \rho_k(x_0, x_1) \dots \rho_k(x_{i-1}, x_i) w_i h_k(x_i)$, где плотности $\rho_k(x, y)$ равны $\rho_k(x, y) = p(x, dy)/r_k(x, dy)$. Имеем $\mathbf{E}_x^{(r_k)} H = \varphi_k(x)$ и, следовательно,

$$\frac{J_0}{J_1} = \frac{\int_X \bar{\pi}(dx) \mathbf{E}_x^{(r_0)} H_0}{\int_X \bar{\pi}(dx) \mathbf{E}_x^{(r_1)} H_1}.$$

2.2.1. Оценки по различным траекториям. Часто для знаменателя при $r_1 = p$ получаются оценки с приемлемой дисперсией. Рассмотрим выбор переходных вероятностей для числителя. функция φ_0 равна

$$\varphi_0(x) = \mathbf{E}_x \sum_{i=0}^{\bar{\tau}-1} h_0(x_i) = \sum_{i=0}^{\infty} K_x^i h_0.$$

Для этой функции можно построить аналогично пп. 2.1.2.a – 2.1.2.b:

2.2.1.a. оценку по поглощениям;

2.2.1.b. оценку по столкновениям.

Оптимальные переходные вероятности оценок определяются соответствующими формулами из п. 2.1.2.

2.2.2. Оценки по тем же траекториям. Здесь $r_0 = r_1$ и оценка для J_0/J_1 имеет вид

$$\frac{\frac{1}{N} \sum_{k=1}^N H_0(\bar{x}_k)}{\frac{1}{N} \sum_{k=1}^N H_1(\bar{x}_k)} = \frac{\bar{H}_0}{\bar{H}_1}. \quad (36)$$

Рассмотрим вопрос о погрешности оценки. Положим $H = H_0 - H_1 J_0/J_1$ и определим функции $\varphi(x) = \mathbf{E}_x^{(r)} H$ и $\psi(x) = \mathbf{E}_x^{(r)} H^2$. Интеграл от ψ по распределению $\bar{\pi}$ обозначим d ; соответствующий интеграл от φ равен

нулю, поэтому

$$\begin{aligned} & \mathbf{P} \left(\left| \frac{\bar{H}_0}{\bar{H}_1} - \frac{J_0}{J_1} \right| < \sqrt{\frac{d}{N}} \frac{\varepsilon}{\bar{H}_1} \right) = \\ & = \mathbf{P} \left(-\varepsilon < \frac{(\bar{H}_0 - \bar{H}_1 J_0/J_1) \sqrt{N}}{\sqrt{d}} < \varepsilon \right) \rightarrow \Phi(\varepsilon) - \Phi(-\varepsilon). \end{aligned}$$

Поскольку при больших N величина $\bar{H}_1$ близка к $\int \bar{\pi}(dx) \varphi_1(x)$ и, значит, не зависит от r , то для уменьшения длины доверительного интервала остается уменьшить d , т.е. $\psi(x)$. Заметим, что $\varphi(x) = \sum_{i=0}^{\infty} K_x^i f$, где функция f равна $f(x) = h_0(x) - h_1(x) J_0/J_1$.

2.2.2.а. В случае оценки по поглощениям $\psi(x) = \sum_{i=0}^{\infty} L_x^i a(x)$, где $a(x) = f^2(x)/g(x)$. Определим $\bar{\varphi}(x) = \sum_{i=0}^{\infty} K_x^i |f|$ и будем предполагать, что $\bar{\varphi}(x) > 0$. Покажем, что оптимальные переходные вероятности равны $r^*(x, dy) = k(x, dy) \bar{\varphi}(y) / \bar{\varphi}(x)$. Пусть для переходных вероятностей $r(x, dy)$ существуют плотности $\rho(x, y) = k(x, dy) / r(x, dy)$; положим $m(x) = r(x, X)$, $g(x) = 1 - m(x)$. Дважды используя неравенство Гельдера, имеем

$$\begin{aligned} \bar{\varphi}^2(x) &= \left(\sqrt{m(x)} \frac{K_x \bar{\varphi}}{\sqrt{m(x)}} + \sqrt{g(x)} \frac{|f(x)|}{\sqrt{g(x)}} \right)^2 \leq \\ &\leq \left(\frac{(K_x \bar{\varphi})^2}{m(x)} + \frac{f^2(x)}{g(x)} \right) (m(x) + g(x)) = \\ &= \frac{1}{m(x)} \left(\int_X \sqrt{\rho(x, y)} \frac{k(x, dy)}{\sqrt{\rho(x, y)}} \bar{\varphi}(y) \right)^2 + a(x) \leq \\ &\leq \frac{1}{m(x)} \int_X \frac{k(x, dy)}{\rho(x, y)} \int_X \rho(x, y) k(x, dy) \bar{\varphi}^2(y) + a(x). \end{aligned}$$

Поскольку $\int_X k(x, dy) / \rho(x, y) = \int_X r(x, dy) = m(x)$, то получаем неравенство $\bar{\varphi}^2(x) \leq \int_X l(x, dy) \bar{\varphi}^2(y) + a(x)$, итерируя которое, имеем оценку снизу для $\psi(x)$: $\psi(x) \geq \bar{\varphi}^2(x)$. С другой стороны, при $r = r^*$ получаем $l^*(x, dy) = \bar{\varphi}(x) k(x, dy) / \bar{\varphi}(y)$, $a^*(x) = |f(x)| \bar{\varphi}(x)$, откуда следует, что $\psi^*(x) = \bar{\varphi}(x) \sum_i K_x^i |f| = \bar{\varphi}^2(x)$, т.е. оптимальность.

2.2.2.b. В случае оценки по столкновениям будем предполагать выполненными условия (25) и (27) из п.2.1.2.b. Докажем, что оптимальные переходные вероятности выражаются формулами (28)-(29). Приведенное в п. 2.1.2.b доказательство в данном случае неприменимо, т.к. функция $a(x)$ является знакопеременной.

Докажем, что уравнение (29) имеет решение. Определим последовательность $b_0 = |\varphi|$, $b_n = Ab_{n-1}$, где A – оператор, задаваемый формулой (30). Легко проверить, что эта последовательность неубывающая. Чтобы доказать, что существует конечный предел, рассмотрим также последовательность $\bar{b}_0 = |\varphi|$, $\bar{b}_n = \bar{A}\bar{b}_{n-1}$, где оператор $\bar{A}$ получается из оператора A заменой функции a на φ^2 . С одной стороны, $\bar{b}_n$ – неубывающая последовательность, а с другой стороны, для $\bar{A}$ выполнено условие (31), откуда следует неравенство $\bar{b}_{n+1} - \bar{b}_n \leq K(\bar{b}_n - \bar{b}_{n-1})/\sqrt{c}$, итерируя которое, в силу (35), получаем оценку $\|\bar{b}_{n+1} - \bar{b}_n\| \leq C(\gamma^*)^n$; из этого неравенства следует $\|\bar{b}_n\| \leq C$, а поскольку $b_n(x) \leq \bar{b}_n(x)$, то и $\|b_n\| \leq C$. Таким образом, существует предел последовательности b_n , который обозначим b , удовлетворяющий уравнению (29).

Теперь покажем, что $b^2(x) \leq \psi(x)$. Перепишем уравнение (26) в виде $\psi(x) = (K_x\sqrt{\psi})^2/c + \tilde{a}(x)$, где $\tilde{a}(x) = a(x) + L_x\psi - (K_x\sqrt{\psi})^2/c \geq a(x)$. Пусть оператор $\tilde{A}$ получается из A заменой a на $\tilde{a}$. Определим последовательность $\tilde{b}_0 = |\varphi|$, $\tilde{b}_n = \tilde{A}\tilde{b}_{n-1}$, тогда $b_n^2(x) \leq \tilde{b}_n^2(x) \leq \psi(x)$, откуда следует, что $b^2(x) \leq \psi(x)$.

Наконец, покажем, что $\psi^* \leq b^2$. Рассмотрим последовательность $\sigma_0 = 0$, $\sigma_n = L^*\sigma_{n-1} + s^*$. Индукцией доказывается, что $\sigma_n \leq b^2 - \varphi^2$; действительно, если это неравенство верно при n , то

$$\sigma_{n+1} \leq L^*(b^2 - \varphi^2) + s^* = \frac{1}{c}(Kb)^2 - (K\varphi)^2 = b^2 - \varphi^2.$$

Отсюда следует, что $\sigma^* \leq b^2 - \varphi^2$ и, значит, $\psi^* = \varphi^2 + \sigma^* \leq b^2$.

2.3 Ветвящийся процесс

В данном методе [8] повышение эффективности достигается не за счет выбора моделируемой плотности, как в методе существенной выборки, и не за счет изменения функционала, как в методе выделения главной части, а за счет увеличения числа траекторий на некоторых множествах. Пусть задано субстохастическое ядро $k(x, dy)$ и ряд

$$Q(x, dy) = \sum_{i=0}^{\infty} k^{(i)}(x, dy)$$

определяет конечную меру; например, в случае схемы вложенной цепи $k(x, dy) = p(x, dy)1_G(y)$.

Пусть также задана функция $b(x)$, такая что $0 < b_0 \leq b(x) \leq b_1 < \infty$ и целочисленная случайная величина $\beta(x, y)$, для которой выполнено $\mathbf{E}\beta(x, y) = b(y)/b(x)$.

Построим ветвящийся процесс: переход из точки x в точку y осуществляется в соответствии с субстохастическими вероятностями $k(x, dy)$; число частиц, рождающихся в точке y равно $\beta(x, y)$. Если траектория $\bar{x}$ начинается в точке x , то будем обозначать это выражением $\bar{x}(0) = x$, а саму траекторию представим в виде $\bar{x} = \left(x, \{\bar{y}_i\}_{i=1}^{\beta}\right)$ где $\bar{y}_i(0) = y$. Число частиц на траектории равно $\eta(\bar{x}) = 1 + \sum_{i=1}^{\beta} \eta(\bar{y}_i)$; вычислим математическое ожидание:

$$n(x) = \mathbf{E}_x \eta = 1 + \int_X k(x, dy) \frac{b(y)}{b(x)} n(y).$$

Итерацией получаем, что

$$n(x) = \frac{1}{b(x)} \int_X Q(x, dy) b(y) \leq \frac{b_1}{b_0} Q(x, X). \quad (37)$$

Положим $\varphi(x) = \int_X Q(x, dy) h(y)$ и определим на траектории $\bar{x}$ функционал

$$\check{\varphi}(\bar{x}) = h(x) + \frac{b(x)}{b(y)} \sum_{i=1}^{\beta} \check{\varphi}(\bar{y}_i).$$

Случайная величина $\check{\varphi}$ является несмещенной оценкой для $\varphi(x)$, т.к.

$$\mathbf{E}_x \check{\varphi} = h(x) + \int_X k(x, dy) \frac{b(x)}{b(y)} \mathbf{E} \beta(x, y) \mathbf{E}_y \check{\varphi} = \sum_{i=0}^{\infty} K_x^i h = \varphi(x).$$

Второй момент $\psi(x) = \mathbf{E}_x \check{\varphi}^2$ равен

$$\psi(x) = h^2(x) + 2h(x)K_x\varphi + \int_X k(x, dy) \frac{b^2(x)}{b^2(y)} \mathbf{E}_{xy} \left(\sum_{i=1}^{\beta} \check{\varphi}_i \right)^2.$$

Вычислим математическое ожидание в последнем слагаемом

$$\begin{aligned} \mathbf{E}_{xy} \left(\sum_{i=1}^{\beta} \check{\varphi}_i \right)^2 &= \varphi^2(y) \mathbf{E}_{xy} (\beta^2 - \beta) + \psi(y) \mathbf{E}_{xy} \beta \\ &= [\psi(y) - \varphi^2(y)] \frac{b(y)}{b(x)} + \varphi^2(y) \left[\frac{b^2(y)}{b^2(x)} + \mathbf{D}_{xy} \beta \right] \end{aligned}$$

Ясно, что минимум $\mathbf{D}_{xy} \beta$ достигается при

$$\beta = \begin{cases} \lfloor d \rfloor, & \text{с вероятностью } \lceil d \rceil - d, \\ \lceil d \rceil, & \text{с вероятностью } d - \lfloor d \rfloor, \end{cases}$$

где $d = d(x, y) = b(y)/b(x)$, при этом $\mathbf{D}_{xy} \beta = (\lceil d \rceil - d)(d - \lfloor d \rfloor) \leq 0.25$.

Таким образом, $\psi(x)$ удовлетворяет уравнению

$$\psi(x) = \int_X \frac{b(x)k(x, dy)\psi(y)}{b(y)} + a(x) + c(x),$$

где

$$\begin{aligned} a(x) &= \varphi^2(x) - (K_x \varphi)^2 + K_x \varphi^2 \\ c(x) &= \int_X k(x, dy) \frac{\varphi^2(y)}{d^2(x, y)} [\mathbf{D}_{xy} \beta - d(x, y)]. \end{aligned}$$

Решение этого уравнения представляется в виде $\psi(x) = b(x)Q_x(a+c)/b$. Заметим, что при $d(x, y) \geq 1$ выполнено $\mathbf{D}_{xy} \beta \leq d(x, y)$ и, следовательно, $c(x) \leq 0$; если же $d(x, y) \leq 1$, то $\mathbf{D}_{xy} \beta = [1 - d(x, y)]d(x, y) \leq d(x, y)$, откуда также следует $c(x) \leq 0$. Значит, $\psi(x) \leq b(x)Q_x a/b = \Psi(x)$.

Пусть зафиксировано время моделирования M и необходимо найти функцию $b(x)$, которая минимизировала бы длину доверительного интервала для оценки $\bar{\varphi} = \frac{1}{N} \sum_{i=1}^N \check{\varphi}_i$ при фиксированном уровне доверия $1 - \alpha$. Пусть σ^2 - дисперсия случайной величины $\check{\varphi}$; в силу центральной предельной теоремы $\mathbf{P}(|\bar{\varphi} - \varphi(x)| < \varepsilon) = 1 - \alpha$, где $\varepsilon = c_\alpha \sigma / \sqrt{N}$, $c_\alpha = -\Phi^{-1}(\alpha/2)$.

Время моделирования M пропорционально среднему числу переходов на траектории n ; можно считать, что $M = Nn$, поэтому задача минимизации ε сводится к минимизации $n\sigma^2$. Величина $n(x)$ определяется формулой (37), а $\sigma^2(x)$ оценивается сверху $\Psi(x)$. Нетрудно видеть, что минимум произведения $n(x)\Psi(x) = Q_x b Q_x a/b$ достигается при $b = \sqrt{a}$.

2.4 Численный пример

Пусть в примере 3 на стр.10 требуется оценить вероятность того, что в стационарном режиме будет занято не менее K устройств. Обозначим $|x| = \sum_{i=1}^m \delta_i$ - число занятых устройств в состоянии x , множество A в (17) имеет вид $A = \{x : |x| > K\}$, а функция $h_1(x) = 1/(\mu + \sum_i \lambda_i \delta_i)$.

Чтобы приближенно оценить величины, участвующие в оптимальных оценках, построим некоторое приближение к заданной системе, а именно, пусть $\tilde{X} = \{0, 1, \dots, m\}$ - множество состояний; состояние $\tilde{x} \in \tilde{X}$ обозначает количество занятых устройств; переходные вероятности приблизим следующим образом: вычислим среднюю интенсивность обработки $\tilde{\lambda} = \frac{1}{m} \sum_{i=1}^m \lambda_i$ и положим

$$\tilde{p}_{ij} = \begin{cases} \frac{i\tilde{\lambda}}{i\tilde{\lambda}+\mu}, & j = \max(0, i-1), \\ \frac{\mu}{i\tilde{\lambda}+\mu}, & j = \min(0, i+1), \\ 0, & \text{в остальных случаях.} \end{cases}$$

Поскольку размерность матрицы $\tilde{P} = \{\tilde{p}_{ij}\}$ небольшая, то можно легко вычислять итерации $\tilde{P}^k$, а значит приближенно оценивать функционал J , функции $\varphi(x)$, $\sigma(x)$ и т.д.

Ниже приведены результаты моделирования по различным схемам при $m = 10$, $K = 9$, $\mu = 2$, $\lambda_i = 1 + i/M$; время счета для всех схем одинаково. В качестве меры эффективности рассматривается отношение дисперсии имитационного моделирования к дисперсии рассматриваемой оценки, полученных по формуле (19). Это отношение показывает во сколько надо увеличить время имитационного моделирования для достижения той же точности, какую показывает рассматриваемая схема.

Метод	Параметры	Эфф – сть	Метод	Эфф–сть
2.1.1	$L = 0, M = 10$	2.3	2.2.1.a	450
2.1.1	$L = 10, M = 10$	85	2.2.1.b	300
2.1.2.a	$\gamma = 0.9$	3.0	2.2.2.a	400
2.1.2.b	$\gamma = 0.9, c = 0.9$	1.7	2.2.2.b	100
2.1.3		800	2.3	50

Список литературы

- [1] А. Кофман, Р. Крюон. Массовое обслуживание. Теория и приложения. Москва, Мир, 1965.
- [2] И. И. Гихман, А. В. Скороход. Теория случайных процессов. Москва, Наука, Т.1 - 1971, Т.2 - 1974.
- [3] Л. Кляйнрок. Теория массового обслуживания. Москва, Машиностроение, 1975.
- [4] С. М. Ермаков. Метод Монте-Карло и смежные вопросы. Москва, Наука, 1975.
- [5] И. Ю. Линник. Об улучшении сходимости метода Монте-Карло при вычислении параметров систем массового обслуживания. В сб.: Вероятностные методы решения задач математической физики. Новосибирск, 1972, с. 156–164.

- [6] Ю. Н. Каштанов. Улучшение сходимости при вычислении методом Монте-Карло функционалов от стационарного распределения марковской цепи. Вестник Ленингр. ун-та, 1980, № 13, с. 34–40.
- [7] Ю. Н. Каштанов. Некоторые способы уменьшения дисперсии при оценивании методом Монте-Карло функционалов от стационарного распределения марковской цепи. Вестник Ленингр. ун-та, 1981, № 7, с. 42–49.
- [8] С.М.Ермаков, В.Б.Мелас. Математический эксперимент с моделями сложных систем. СПб, Из-во С.-Петербургского университета, 1993.

Научное издание

Кацтанов Юрий Николаевич

**Статистическое моделирование
систем массового обслуживания**

Подписано в печать XX.XX.14 Формат 60x84 1/16. Бумага офсетная.

Печать ризографическая. Усл. печ. л.

Тираж 50 экз. Заказ .

Отпечатано в отделе оперативной полиграфии НИИХ СПбГУ.
198504, Санкт-Петербург, Старый Петергоф, Университетский пр. 26