Марковские случайные процессы

При исследовании различных операций с точки зрения выбора оптимального решения часто возникают ситуации, когда обстановка приведения операции характеризуется случайными неконтролируемыми факторами.

В этом случае операция развивается по схеме случайного процесса, протекание которого зависит от сопровождающих операцию случайных факторов.

В случае стохастических задач исследования операций построение математической модели является достаточно сложным. Исключение составляет особый случай, когда исследуемая операция представляет собой так называемый **марковский процесс**.

Количественно случайный процесс описывается случайной функцией времени t, которая может принимать различные значения с заданным распределением вероятностей. Т.о. для любого t=ti значение

$$\xi_i = \xi(t_i)$$

является случайной величиной.

Случайный процесс определяется совокупностью функций времени и законами, характеризующими свойства этой совокупности. Каждая из функций этой совокупности называется реализацией случайного процесса. Реализация обозначается

$$\xi^{(q)}(t), \text{ где } q = 1, 2, \ldots .$$

В зависимости от того, принадлежат ли возможные значения времени t и реализации $\xi(t)$, дискретному множеству чисел или интервалу

действительных чисел, различают четыре типа случайных процессов:

1. **Случайный процесс общего типа:**
t и $\xi(t)$ могут принимать любые значения.
2. **Дискретный случайный процесс:**
t-непрерывно, а значения $\xi(t)$ дискретны.
3. **Случайная последовательность общего типа:**
t-дискретно, а $\xi(t)$ принимает любые значения.
4. **Дискретная случайная последовательность:**
t и $\xi(t)$ дискретны.

Для описания случайного процесса используют функции распределения: одномерная интегральная функция распределения вероятностей случайного процесса

$$F_1(x_1,t_1) = P(\xi(t_1) \le x_1)$$

$$\partial F_1(x_1,t_1) / \partial x_1 = f(x_1,t_1)$$

Определим n-мерную функцию распределения вероятностей случайного процесса

$$F_n(x_1,t_1,\ldots,x_n,t_n) = P\{\xi(t_1) \le x_1, \xi(t_2) \le x_2, \ldots, \xi(t_n) \le x_n\}$$

$$\partial^n F^I{}_n(x_1,t_1,\ldots,x_n,t_n) / \partial x_1 * x_n = f_n(x_1,t_1,\ldots,x_n,t_n)$$

Случайный процесс будет марковским, если выполняется условие

$$f\left(x_n, t_n \mid x_{n-1}, t_{n-1} .. x_1, t_1\right) = f\left(x_n, t_n \mid x_{n-1}, t_{n-1}\right)$$

$$f\left(x_1, t_1, \ldots., x_n, t_n\right) = f\left(x_1, t_1\right) f\left(x_2, t_2 \mid x_1, t_1\right) f\left(x_3, t_3 \mid x_2, t_2\right) \ldots.$$

$f(x_i, t_i \mid x_{i-1}, t_{i-1})$ называется плотностью вероятности перехода.

Если плотность вероятности перехода зависит от разности $t_i - t_{i-1}$

$$f(x_i, t_i \mid x_{i-1}, t_{i-1}) = f(x_i, t_1 - t_{i-1} \mid x_{i-1})$$

то такой процесс называется однородным.

Пример системы с дискретными состояниями и непрерывным временем

техническое устройство Q состоит из двух узлов каждый из которых в случайный момент времени может выйти из строя, после чего мгновенно начинается его ремонт, который продолжается случайное время.

- Возможные состояния системы:
- **Q0** - оба узла исправны.
- **Q1** - первый узел ремонтируется, а второй исправен.
- **Q2** - второй узел ремонтируется, а первый исправен.
- **Q3** - оба узла ремонтируются. Переход системы Q из состояния в состояние происходит практически мгновенно в случайные моменты времени выхода какого-то узла из строя или состояния его ремонта.

Потоки событий

Для рассмотрения случайных процессов, протекающих в системах с дискретными состояниями и непрерывным временем определим понятие ***"поток событий"****.*

Потоком событий называется последовательность однородных событий, следующих одно за другим в случайные моменты времени (*поток автобусов на данной остановке, поток отказов какой-то системы и т. п.*) Поток событий будем изображать последовательностью точек на оси времени

будем рассматривать потоки событий, обладающие свойствами: *стационарность, отсутствие последействия, ординарность*.

Поток событий называется простейшим, если он стационарен, однороден и не имеет последействия.

Для простейшего потока интервал t между соседними событиями имеет показательное распределение $f(t)=\lambda e^{-\lambda t}$

Поток событий называется рекуррентным или потоком "Пальма", если он стационарен, ординарен, а интервалы времени между событиями представляют собой независимые случайные величины с одинаковым произвольным распределением

Уравнения Колмогорова для вероятностей состояний

Формула Пуассона

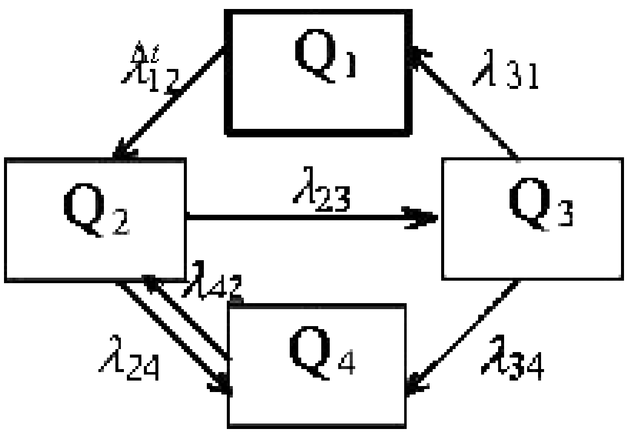
$$P(m, \Delta t) = ((\lambda \Delta t)^m / m!) e^{-\lambda \Delta t}$$

С учетом ординарности

$$P(0, \Delta t) = (\lambda \Delta t)^0 \cdot \exp(-\lambda \Delta t) / 0! \approx 1, \Delta t \to 0$$

$$P(1, \Delta t) = (\lambda \Delta t)^1 \cdot \exp(-\lambda \Delta t) / 1! \approx \lambda \Delta t, \Delta t \to 0$$

$$P(0, \Delta t) + P(1, \Delta t) \approx 1, \text{ а } P(1, \Delta t) \approx \lambda \Delta t.$$

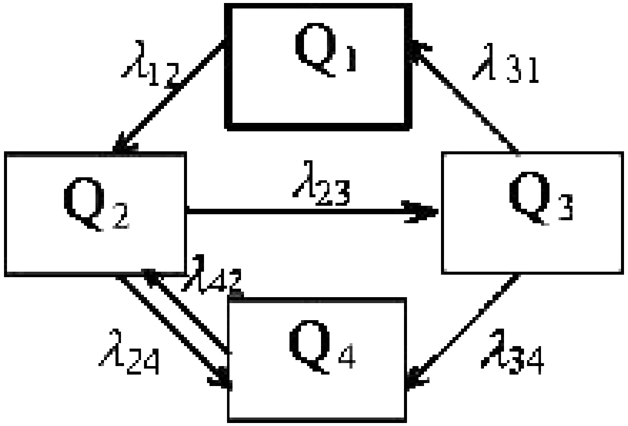


Найдем вероятность $p_1(t)$, что в момент t система будет находиться в состоянии Q_1.

Придадим t приращение Δt и найдем вероятность того, что в момент $t + \Delta t$ система будет находиться в состоянии Q_1

Это событие может осуществиться двумя способами:

1. В момент t система была в состоянии Q_1 и за время Δt из него не вышла;
2. В момент t система была в состоянии Q_3 и за Δt перешла в Q_1.



1. В момент t система была в состоянии Q1 и за время Δt из него не вышла;

Вероятность равна произведению $p_1(t)$ на условную вероятность того, что за Δt не произойдет перехода $Q_1 \rightarrow Q_2$

Условная вероятность, что не произойдет переход равна вероятности события обратному возникновению перехода $1-\lambda_{12}\Delta t$

В целом вероятность события 1 равна $p_1(t)(1-\lambda_{12}\Delta t)$

2. В момент t система была в состоянии Q3 и за Δt перешла в Q1 .
$p_3(t)(\lambda_{31}\Delta t)$

$$p_1(t+\Delta t) = p_1(t)(1-\lambda_{12}\Delta t) + p_3(t)\lambda_{31}\Delta t$$

Как получаются диф уравнения

$$p_1(t+\Delta t) = p_1(t)(1-\lambda_{12}\Delta t) + p_3(t)\lambda_{31}\Delta t$$

$$p_1(t+\Delta t) = p_1(t) - p_1(t)\lambda_{12}\Delta t + p_3(t)\lambda_{31}\Delta t$$

$$p_1(t+\Delta t) - p_1(t) = -p_1(t)\lambda_{12}\Delta t + p_3(t)\lambda_{31}\Delta t$$

Делим на Δt

$$\frac{p_1(t+\Delta t) - p_1(t)}{\Delta t} = -p_1(t)\lambda_{12} + p_3(t)\lambda_{31}, \Delta t \to 0$$

$$\frac{dp_1(t)}{dt} = -p_1(t)\lambda_{12} + p_3(t)\lambda_{31}$$

Переходим к
к пределу при
Δt -> 0

Аналогично можно найти еще три уравнения вероятностей состояний (уравнения Колмогорова)

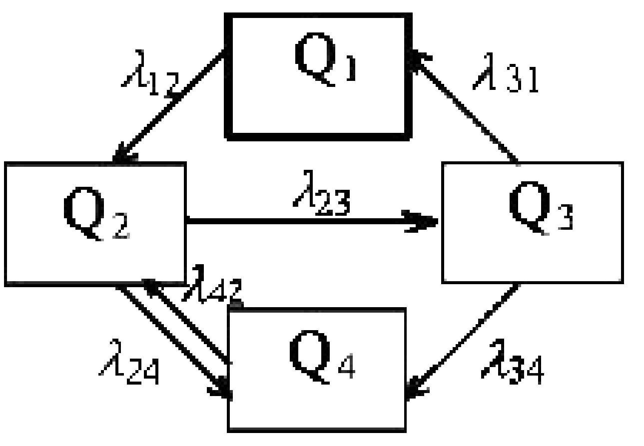
$$\frac{dp_1(t)}{dt} = -\lambda_{12} p_1(t) + \lambda_{31} p_3(t)$$

$$\frac{dp_2(t)}{dt} = -\lambda_{23} p_2 - \lambda_{24} p_2 + \lambda_{12} p_1 + \lambda_{42} p_4$$

$$\frac{dp_3(t)}{dt} = -\lambda_{31} p_3 - \lambda_{34} p_3 + \lambda_{23} p_2$$

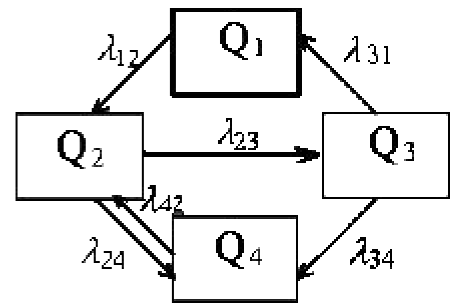
$$\frac{dp_4(t)}{dt} = -\lambda_{42} p_4 + \lambda_{24} p_2 + \lambda_{34} p_3$$

- Интегрируя эту систему уравнений, можно найти вероятности состояний, как функции времени. Для этого необходимо задать начальные условия при t=0.
  Например $p_1=p_3=p_4=0$, $p_2=1$ - это означает, что при t=0 система находится в состоянии Q2.



Правило составления дифференциальных уравнений

- *В* левой части каждого уравнения стоит производная вероятности состояния, а правая содержит столько членов, сколько стрелок связано с данным состоянием. Если стрелка направлена из состояния, соответствующий член имеет знак "-", если в состояние знак "+". Каждый член равен произведению плотности вероятности перехода, соответствующему данной стрелке, умноженной на вероятность состояния, из которого исходит стрелка.



$$\frac{dp_1(t)}{dt} = -\lambda_{12}p_1(t) + \lambda_{31}p_3(t)$$

$$\frac{dp_2(t)}{dt} = -\lambda_{23}p_2 - \lambda_{24}p_2 + \lambda_{12}p_1 + \lambda_{42}p_4$$

$$\frac{dp_3(t)}{dt} = -\lambda_{31}p_3 - \lambda_{34}p_3 + \lambda_{23}p_2$$

$$\frac{dp_4(t)}{dt} = -\lambda_{42}p_4 + \lambda_{24}p_2 + \lambda_{34}p_3$$

Предельные вероятности состояний

предельные состояния системы $p_1, p_2, \dots, p_n$ при $t \to \infty$

Предельные или финальные вероятности характеризуют установившийся стационарный режим, для которого $dp_i / dt = 0$

$$\sum p_i = 1$$

$$\lambda_{31} p_3 - \lambda_{12} p_1 = 0$$

$$-\lambda_{32} p_2 - \lambda_{24} p_2 + \lambda_{12} p_1 + \lambda_{42} p_4 = 0$$

$$-\lambda_{31} p_3 - \lambda_{34} p_3 + \lambda_{23} p_2 = 0$$

$$-\lambda_{42} p_4 + \lambda_{24} p_2 + \lambda_{34} p_3 = 0$$

Использование предельных вероятностей состояний

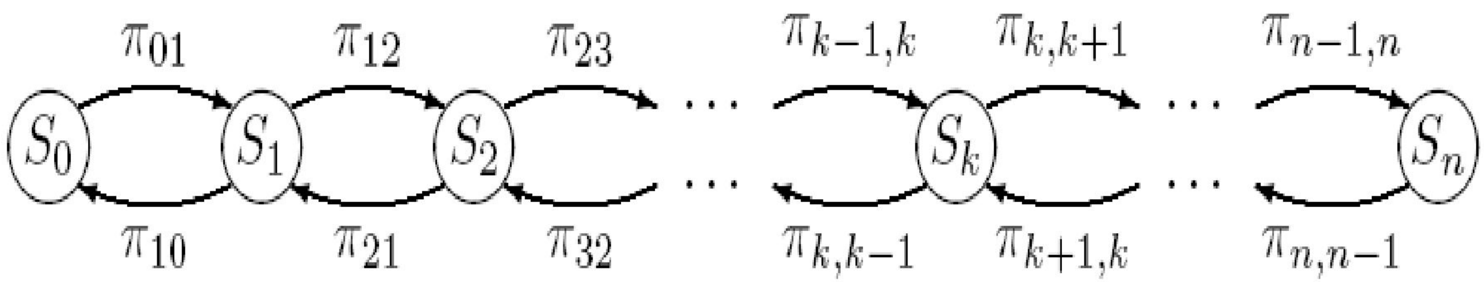
- Знание финальных вероятностей можно использовать, например, для оценки эффективности работы всей системы. Если предположить, что, находясь в состоянии Q_i, система приносит доход w_i, тогда средний доход в стационарном режиме равен

$$W = \sum_i w_i p_i$$

Можно ставить и решать задачу оптимизации системы.

Схема гибели и размножения

Термин «схема гибели и размножения» в биологии описывает изменение численности популяции. Схема гибели и размножения очень часто встречается и в задачах теории массового обслуживания, поэтому мы и начинаем с ее рассмотрения.



Граф состояний для схемы гибели и размножения

S_k – состояние означает: численность популяции равна k

Обозначим через $p_k(t)$ вероятность того, что в момент времени t система находится в состоянии S_k. Для достаточно малого $\Delta t > 0$ в момент времени $t + \Delta t$ система окажется в состоянии S_k $(1 < k < n)$

- с вероятностью $\pi_{k-1,k}\Delta t$, если в момент t она была в состоянии S_{k-1};
- с вероятностью $1 - (\pi_{k,k-1} + \pi_{k,k+1})\Delta t$, если в момент t она была в состоянии S_k;
- с вероятностью $\pi_{k+1,k}\Delta t$, если в момент t она была в состоянии S_{k+1}.

Поэтому справедливо равенство

$$\begin{aligned} p_k(t+\Delta t) &= p_{k+1}(t)\pi_{k+1,k}\Delta t + p_{k-1}(t)\pi_{k-1,k}\Delta t \\ &\quad + p_k(t)\left(1 - (\pi_{k,k+1} + \pi_{k,k-1})\Delta t\right). \end{aligned}$$

Разделив обе части равенства на Δt, получим

$$\frac{p_k(t+\Delta t)-p_k(t)}{\Delta t}=\pi_{k+1,k}p_{k+1}(t)+\pi_{k-1,k}p_{k-1}(t)-(\pi_{k,k+1}+\pi_{k,k-1})p_k(t).$$

Переходя к пределу при $\Delta t\to 0$, получим

$$\frac{dp_k(t)}{dt}=\pi_{k+1,k}p_{k+1}(t)+\pi_{k-1,k}p_{k-1}(t)$$
$$-(\pi_{k,k+1}+\pi_{k,k-1})p_k(t),\quad k=1,\ldots,n-1.$$

Аналогично, можно получить уравнения для $k=0$ и $k=n$:

$$\frac{dp_0(t)}{dt}=\pi_{10}p_1(t)-\pi_{01}p_0(t),$$
$$\frac{dp_n(t)}{dt}=\pi_{n-1,n}p_{n-1}(t)-\pi_{n,n-1}p_n(t).$$

Если в системе установился стационарный режим, то все вероятности $p_k(t) \stackrel{\text{def}}{=} p_k$ постоянны (независят от времени). Мы можем вычислить *финальные вероятности* $p_0, p_1, \ldots, p_n$[12] состояний системы, решая систему с учетом того, что $\frac{dp_k(t)}{dt} = 0$ для $k = 0, 1, \ldots, n$.

Для состояния S_0 справедливо равенство:

$$\pi_{01} p_0 = \pi_{10} p_1.$$

Для состояния S_1 имеем:

$$(\pi_{10} + \pi_{12}) p_1 = \pi_{01} p_0 + \pi_{21} p_2.$$

последнее равенство приводится к виду

$$\pi_{12} p_1 = \pi_{21} p_2.$$

Далее, совершенно аналогично получаем равенство

$$\pi_{23}p_2 = \pi_{32}p_3$$

и для любого $k = 1, \ldots, n$ имеем:

$$\pi_{k-1,k}p_{k-1} = \pi_{k,k-1}p_k.$$

Итак, финальные вероятности $p_0, p_1, \ldots, p_n$ удовлетворяют системе

$$\pi_{01}p_0 = \pi_{10}p_1,$$

$$\pi_{12}p_1 = \pi_{21}p_2,$$

$$\ldots$$

$$\pi_{k-1,k}p_{k-1} = \pi_{k,k-1}p_k,$$

$$\ldots$$

$$\pi_{n-1,n}p_{n-1} = \pi_{n,n-1}p_n.$$

$$\pi_{01}p_0 = \pi_{10}p_1,$$ выразим p_1 через p_0: $$p_1 = \frac{\pi_{01}}{\pi_{10}}p_0.$$

$$\pi_{12}p_1 = \pi_{21}p_2, \qquad p_2 = \frac{\pi_{12}}{\pi_{21}}p_1 = \frac{\pi_{01}\pi_{12}}{\pi_{21}\pi_{10}}p_0.$$

Аналогично выражаем p_3

$$p_3 = \frac{\pi_{23}}{\pi_{32}}p_2 = \frac{\pi_{01}\pi_{12}\pi_{23}}{\pi_{32}\pi_{21}\pi_{10}}p_0.$$

В общем, для любого $k = 1, \ldots, n$ имеем:

$$p_k = \frac{\pi_{01}\pi_{12}\ldots\pi_{k-1,k}}{\pi_{k,k-1}\ldots\pi_{21}\pi_{10}}p_0.$$

$$p_k = \frac{\pi_{01}\pi_{12}\cdots\pi_{k-1,k}}{\pi_{k,k-1}\cdots\pi_{21}\pi_{10}} p_0.$$

Заметим, что в формуле числитель есть произведение всех интенсивностей, стоящих у дуг, ведущих слева направо от состояния S_0 до состояния S_k, а знаменатель есть произведение всех интенсивностей, стоящих у дуг, ведущих справа налево от состояния S_k до состояния S_0. Таким образом, все вероятности состояний $p_1, \ldots, p_n$ выражаются через состояние p_0. Подставив эти выражения в нормировочное равенство

$$p_0 + p_1 + p_2 + \cdots + p_n = 1,$$

найдем

$$p_0 = \left(1 + \frac{\pi_{01}}{\pi_{10}} + \frac{\pi_{01}\pi_{12}}{\pi_{21}\pi_{10}} + \cdots + \frac{\pi_{01}\pi_{12}\cdots\pi_{n-1,n}}{\pi_{n,n-1}\cdots\pi_{21}\pi_{10}}\right)^{-1}$$

Теория массового обслуживания

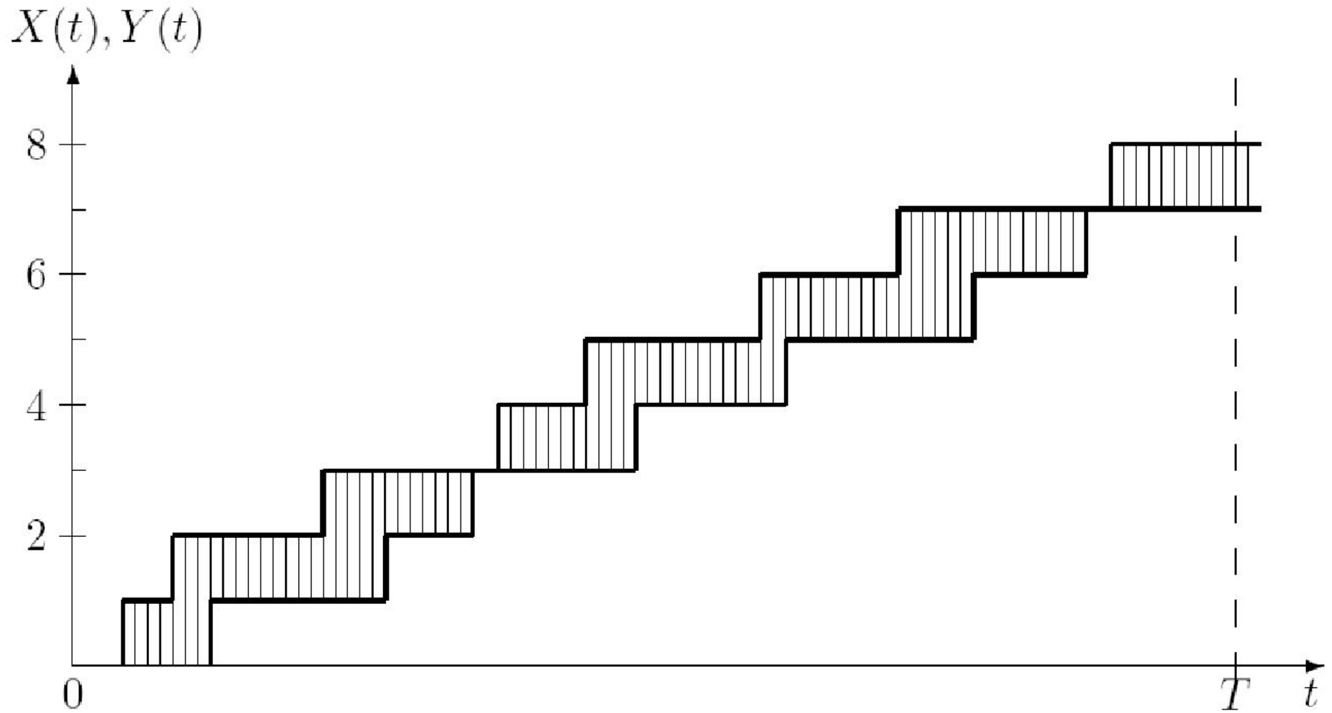
- Каждая система массового обслуживания (**СМО**) состоит из одного или нескольких «приборов», которые называются каналами обслуживания. Каналами могут быть: линии связи, билетные кассы, лифты, такси, вебсерверы, серверы баз данных и др. СМО могут быть одноканальными и многоканальными.
- Всякая СМО предназначена для обслуживания некоторого потока заявок (или «требований»), которые поступают в случайные моменты времени. Обслуживание заявки продолжается какое-то время (в общем случае продолжительность обслуживания заявки есть случайная величина), после чего канал освобождается и готов к обслуживанию следующей заявки. Случайный характер потока заявок и продолжительности их обслуживания приводит к тому, что в некоторые периоды времени на входе СМО может скапливаться излишне большое число заявок (они либо становятся в очередь, либо покидают СМО необслуженными); в другие же периоды отдельные каналы СМО могут простаивать.
- Процесс работы СМО представляет собой случайный процесс с дискретными состояниями и непрерывным временем; состояние СМО меняется скачком в моменты, когда появляется новая заявка, или завершается обслуживание некоторой заявки, или заявка, которой надоело ждать в очереди, покидает очередь.
- В дальнейшем, будем предполагать, что все потоки заявок и обслуживаний являются пуассоновскими.

Поскольку средняя продолжительность интервала между последовательными событиями $E(T_j) = 1/\lambda$, то параметр λ можно рассматривать как *интенсивность потока*, которая равна среднему количеству событий, происходящих в единицу времени.

Формулы Литтла

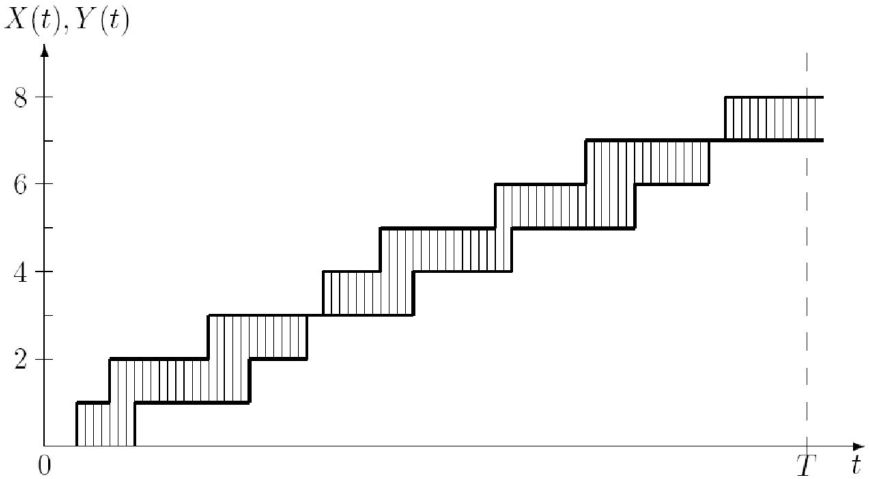
Выведем важную формулу, связывающую (для предельного стационарного режима) среднее число заявок Lсист , находящихся в СМО (т. е. обслуживаемых или стоящих в очереди), и среднее время пребывания заявки в системе Wсист .

Рассмотрим любую СМО (одноканальную или многоканальную, марковскую или немарковскую, с неограниченной или с ограниченной оче редью, и т. д.) и связанные с нею два потока событий: поток заявок, поступающих в СМО, и поток заявок, покидающих СМО. Если в системе установился предельный стационарный режим, то среднее число заявок, поступаюших в СМО, равно среднему числу заявок, покидающих СМО, т. е. оба потока имеют одну и ту же интенсивность λ.



- Обозначим через X(t) число заявок, поступивших в СМО до момента времени t, а через Y (t) число заявок, покинувших СМО до момента t. И та и другая функции являются случайными, X(t) увеличиваются на единицу в момент поступления новой заявки, а Y (t) увеличивается на единицу в момент, когда некоторая заявка покидает систему. Поведение функций X(t) и Y (t) проиллюстрировано на рисунке. Для любого мо мента t разность Z(t) = X(t) − Y (t) есть число заявок, находящихся в СМО. Когда Z(t) = 0, в системе нет заявок.

Рассмотрим очень большой промежуток времени T и вычислим для него среднее число заявок, находящихся в СМО. Оно будет равно



$$\frac{1}{T}\int_0^T Z(t)dt.$$

Этот интеграл равен площади фигуры, заштрихованной на рисунке Фигура состоит из прямоугольников, k-й из которых имеет высоту, равную единице, и основание, равное времени t_k пребывания в системе заявки, поступившей k-й по счету. Отметим, что в конце промежутка T некоторые прямоугольники войдут в заштрихованную фигуру не полностью, а частично, но при достаточно больших T. k(T) обозначает количество заявок, поступивших в систему за время T .

$$\int_0^T Z(t)dt \approx \sum_{k=1}^{k(T)} t_k,$$

Отсюда получаем

$$L_{\text{сист}} = \lim_{T\to\infty}\frac{1}{T}\int_0^T Z(t)dt = \lim_{T\to\infty}\frac{1}{T}\sum_{k=1}^{k(T)} t_k = \lambda\lim_{T\to\infty}\frac{1}{T\lambda}\sum_{k=1}^{k(T)} t_k.$$

Но величина $T\lambda$ есть среднее число заявок, поступивших за время T. Поэтому

$$\lim_{T\to\infty}\frac{1}{T\lambda}\sum_{k=1}^{k(T)} t_k$$

есть среднее время пребывания заявки в системе $W_{\text{сист}}$. Итак $L_{\text{сист}} = \lambda W_{\text{сист}}$, или

$$W_{\text{сист}} = \frac{1}{\lambda}L_{\text{сист}}.$$

Первая формула Литтла

$$W_{\text{сист}} = \frac{1}{\lambda} L_{\text{сист}}.$$

Для любой СМО, при любом характере потока заявок, при любом распределении времени обслуживания, при любой дисциплине обслуживания среднее время пребывания заявки в системе равно среднему числу заявок в системе, деленному на интенсивность потока заявок.

Вторая формула Литтла, связывающая среднее время пребывания заявки в очереди Wоч и среднее
число заявок в очереди Lоч

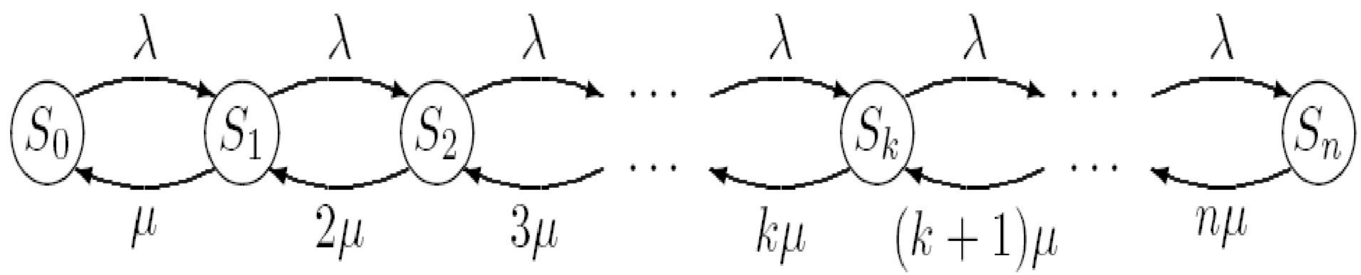
$$W_{\text{оч}} = \frac{1}{\lambda} L_{\text{оч}}.$$

Для вывода формулы достаточно заменить функцию Y на функцию U , где U (t) есть количество заявок, покинувших очередь до момента t (если заявка, поступая в систему, обслуживается сразу, не становясь в очередь, то можно считать, что она пробыла в очереди нулевое время).

Многоканальная СМО с отказами

- Одна из первых по времени «классических» задач теории массового обслуживания. Эта задача возникла из практических нужд телефонии и была решена в начале 19-го века датским математиком Эрлангом.
- Имеется n каналов (линий связи), на которые поступает поток заявок с интенсивностью λ. Поток обслуживаний имеет интенсивность μ.
- Нужно найти финальные вероятности состояний СМО, а также характеристики ее эффективности:
- • A — абсолютную пропускную способность, т. е. среднее число заявок, обслуживаемых в единицу времени;
- • Q — относительную пропускную способность, т. е. среднею долю пришедших заявок, обслуженных системой;
- • Ротк — вероятность отказа, т. е. того, что заявка покинет СМО необслуженной;
- • $\bar{k}$ — среднее число занятых каналов.
- Состояние данной СМО определяется числом заявок в системе (в данном случае оно совпадает с числом занятых каналов): S_k — в СМО находится k заявок (k = 1, . . . , n). Граф состояний СМО соответствует схеме гибели и размножения

Многоканальная СМО с отказами



$$p_0 = \left(1 + \frac{\pi_{01}}{\pi_{10}} + \frac{\pi_{01}\pi_{12}}{\pi_{21}\pi_{10}} + \cdots + \frac{\pi_{01}\pi_{12}\cdots\pi_{n-1,n}}{\pi_{n,n-1}\cdots\pi_{21}\pi_{10}}\right)^{-1} \qquad p_k = \frac{\pi_{01}\pi_{12}\cdots\pi_{k-1,k}}{\pi_{k,k-1}\cdots\pi_{21}\pi_{10}} p_0.$$

По формулам предельных вероятностей состояний находим

$$p_0 = \left(1 + \frac{\lambda}{\mu} + \frac{\lambda^2}{2\mu^2} + \frac{\lambda^3}{3!\mu^3} + \cdots + \frac{\lambda^k}{k!\mu^k} + \ldots \frac{\lambda^n}{n!\mu^n}\right)^{-1}. \qquad p_k = \frac{\lambda^k}{k!\mu^k} p_0, \quad k = 1, \ldots, n.$$

Члены разложения являются коэффициентами при p0 в выражениях для p1 , . . . , pn

Формулы Эрланга

- Обозначим отношение λ/μ через ρ и назовем его «приведенной интенсивностью потока заявок». Заметим, что ρ есть среднее число заявок, приходящее за среднее время обслуживания одной заявки. Пользуясь этим обозначением, перепишем формулы следующим образом:

$$p_0 = \left(\sum_{k=0}^{n} \frac{\rho^k}{k!} \right)^{-1},$$

$$p_k = \frac{\rho^k}{k!} p_0, \quad k = 1, \ldots, k.$$

По формулам Эрланга определяем характеристики СМО

Вероятность того, что пришедшая заявка получит отказ (не будет обслужена) равна

$$P_{\text{отк}} = p_n = \frac{\rho^n}{n!} p_0.$$

Далее находим относительную пропускную способность — вероятность того, что пришедшая заявка будет обслужена:

$$Q = 1 - P_{\text{отк}} = 1 - \frac{\rho^n}{n!} p_0.$$

Абсолютную пропускную способность получим, умножая интенсивность потока заявок λ на Q:

$$A = \lambda Q = \lambda \left(1 - \frac{\rho^n}{n!} p_0 \right).$$

Среднее число занятых каналов $\bar{k}$

$$\bar{k} = 0 \cdot p_0 + 1 \cdot p_1 + 2 \cdot p_2 + \cdots + n \cdot p_n,$$

$$\bar{k} = A/\mu$$

А – средняя интенсивность обслуживания СМО (абсолютная пропускная способность, число обслуженных заявок в единицу времени)
μ – средняя интенсивность обслуживания одним каналом

$$\bar{k} = \rho \left(1 - \frac{\rho^n}{n!} p_0\right).$$

Пример

- Станция связи имеет три канала (n = 3), интенсивность потока заявок 1.5 заявки в минуту, среднее время обслуживания одной заявки 2 минуты. Найти финальные вероятности состояний и характеристики эффективности СМО: A, Q, Ротк , k. Сколько нужно каналов, чтобы удовлетворять не менее 80 % заявок? Какая доля каналов при этом будет простаивать?
- Решение. Здесь λ = 3/2, μ = 1/2 и ρ = (3/2)/(1/2) = 3. По формуле вычислим

$$p_0 = \left(1 + \frac{\lambda}{\mu} + \frac{\lambda^2}{2\mu^2} + \frac{\lambda^3}{3!\mu^3} + \cdots + \frac{\lambda^k}{k!\mu^k} + \ldots \frac{\lambda^n}{n!\mu^n}\right)^{-1}$$

$$p_0 = \frac{1}{1+\rho+\rho^2/2+\rho^3/6} = \frac{1}{1+3+9/2+27/6} = 1/13.$$

Теперь мы можем вычислить вероятность отказа $P_{\text{отк}} = p_3 = (3^3/6) \cdot (1/13) = 9/26$, относительную пропускную способность системы $Q = 1 - P_{\text{отк}} = 1 - 9/26 = 15/26$, абсолютную пропускную способность системы $A = \lambda Q = (3/2) \cdot (15/26) = 45/52$, и среднее число занятых каналов $\bar{k} = A/\mu = (45/52)/(1/2) = 45/26$. □