

Лекция 1

по Теории Массового Обслуживания

от Татаринова Никиты Алексеевича, БПИ193

2022.09.01

Литература

- D.Cross, C.M.Harris "Fundamentals of Queuing Theory"

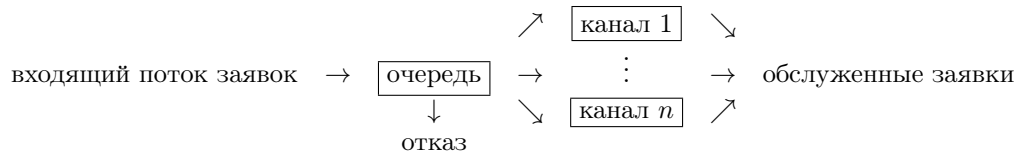
Теория массового обслуживания

Родилась из нужд телекоммуникационных компаний – организовать телекоммуникации оптимальным образом. Совмещает теорию вероятностей и исследование операций. 1909г – датский инженер А.К.Эрланг описал модель появления событий во времени – "Theory of Probability and Telephone Conversations".

Что интересует?

- частота звонков (или среднее время между звонками) – интенсивность входящего потока звонков
- среднее время на обслуживание звонка – интенсивность обслуживания
- число телефонисток

Модель



Нотация Кендалла

Последовательность из нескольких символов и цифр:

$$A/B/X/Y/Z$$

, который характеризуют некоторые особенности системы массового обслуживания.

- A – закон распределения времени между поступлением заявок
 - M – экспоненциальное (беспамятное распределение)
 - E_k/Er_k – распределение Эрланга порядка k
 - D – детерминированное время
 - G – произвольное распределение (отсутствие ограничений)
- B – закон распределения времени обслуживания (те же обозначения)
- X – число каналов обслуживания (может быть бесконечным, но нетипично)
- Y – ёмкость системы – максимальное число заявок в системе одновременно (может быть бесконечным, но нетипично)

- Z – дисциплина обслуживания – показывает, в каком порядке обслуживаются заявки
 - $FCFS/FIFO$ – First Come First Served
 - $LCFS/LIFO$ – Last Come First Served
 - $RSS/SIRO$ – Random Selection for Service
 - PS – processor sharing (например, 2 канала обслуживают одновременно, но в 2 раза медленнее)
 - PRI – приоритетное обслуживание (заявки разбиваются на разные типы с разными приоритетами)
 - GD – general discipline

Основные случайные величины, связанные с процессом обслуживания

Для числа заявок:

$$\left\{ \begin{array}{l} N = N_q + N_s \\ \text{число заявок в системе} = \text{число заявок в очереди} + \text{число заявок на обслуживании} \\ L = E[N] \quad - \text{среднее число заявок в системе} \\ L_q = E[N_q] \quad - \text{среднее число заявок в очереди} \\ L_s = U = E[N_s] \quad - \text{коэффициент загрузки мощностей} \end{array} \right.$$

Для времени:

$$\left\{ \begin{array}{l} T = T_q + T_s \\ \text{время пребывания заявки в системе} = \text{время ожидания в очереди} + \text{время обслуживания} \\ W = E[T] \quad - \text{среднее время в системе} \\ W_q = E[T_q] \quad - \text{среднее время ожидания в очереди} \\ \mu = \frac{1}{E[N_s]} \quad - \text{интенсивность обслуживания} \end{array} \right.$$

Кроме того:

$$\left\{ \begin{array}{l} P_{loss} \quad - \text{вероятность потери} \\ \lambda \quad - \text{интенсивность входящего потока заявок} \\ ? \quad - \text{пропускная способность} \end{array} \right.$$

Пример

Заявки поступают в моменты 0, 20, 40, Время обслуживания равномерное на отрезке [14; 24] минут. Один канал обслуживания. Если заявка поступает, а канал занят, то она получает отказ. Найти вероятность $P_l(k)$ потери заявки с номером k .

$$P_l(1) = 0$$

$$P_l(2) = \frac{24-20}{24-14} = 0.4$$

$$P_l(3) = (1 - P_l(2)) \cdot 0.4 = 0.24$$

$$P_l(4) = (1 - P_l(3)) \cdot 0.4 = 0.304$$

$$P_l(k) = (1 - P_l(k-1)) \cdot 0.4$$

График представляет затухающие колебания вокруг некоего значения. Первые этапы называются переходным периодом, дальнейшие - стационарным периодом.

Геометрическая интерпретация математического ожидания

$$G_T(t) = \mathbf{P}\{T > t\} = 1 - F_T(t) \quad - \text{дополнительная функция распределения} \\ \text{(complementary cdf)}$$

Тогда:

$$E[T] = \int_0^{+\infty} G_T(t) \cdot dt - \int_{-\infty}^0 F_T(t) \cdot dt$$

Для неотрицательной СВ T :

$$E[T] = \int_0^{+\infty} G_T(t) \cdot dt$$

Пример 1

Пусть $T \sim \text{Exp}(\lambda)$, т.е.

$$G_T(t) = 1 - F_T(t) = \begin{cases} 1 & t < 0 \\ e^{-\lambda t} & t \geq 0 \end{cases}$$

Тогда:

$$E[T] = \int_0^{+\infty} e^{-\lambda t} \cdot dt = -\frac{1}{\lambda} \cdot e^{-\lambda t} \Big|_0^{+\infty} = 0 - \left(-\frac{1}{\lambda}\right) = \frac{1}{\lambda}$$

Пример 2

Пусть:

$$G_T(t) = 0.7 \cdot e^{-t} \quad t \geq 0$$

Найти:

- а) долю заявок, которым приходится ждать в очереди
- б) среднее время ожидания
- в) вероятность, что заявке придётся ждать более 2 минут при условии, что ей придётся ждать

Решение

а)

$$\mathbf{P}\{T > 0\} = G_T(0) = 0.7$$

б)

$$E[T] = \int_0^{+\infty} 0.7 \cdot e^{-t} \cdot dt = -0.7 \cdot e^{-t} \Big|_0^{+\infty} = 0 - (-0.7) = 0.7$$

в)

$$\mathbf{P}\left(\{T > 2\} \mid \{T > 0\}\right) = \frac{\mathbf{P}\left(\{T > 2\} \cap \{T > 0\}\right)}{\mathbf{P}\{T > 0\}} = \frac{\mathbf{P}\{T > 2\}}{\mathbf{P}\{T > 0\}} = \frac{G_T(2)}{G_T(0)} = e^{-2}$$