

Санкт-Петербургский государственный политехнический университет

Лабораторная работа № 2

по курсу «Стохастические модели»

«Определение параметров распределения потока заявок
по наблюдениям нагрузки системы»

Студент:	Руцкий В. В.
Группа:	5057/2
Преподаватель:	Иванков А. А.

Санкт-Петербург 2011

1 Постановка задачи

Данной работе производится анализ лога загруженности процессора сервера при поступающих заявках на обработку информации.

В отсутствие заявок величина загруженности процессора представляет собой сумму некоторой постоянной величины загрузки m и случайных отклонений:

$$B(t) = m + \sigma \mathcal{W}(t),$$

где $\mathcal{W}(t)$ — это винеровский процесс.

Интенсивность поступления заявок подчиняются закону распределения Пуассона $\mathcal{P}(\lambda)$.

При поступлении одной заявки нагрузка на процессор мгновенно возрастает, а затем экспоненциально снижается до прежнего уровня. Увеличение загрузки процессора от одной заявки, поступившей в момент времени t_c выражается следующим образом:

$$K_{t_c}(t) = \mathcal{N}(m_c, \sigma_c^2) \cdot \mathbf{I}(t - t_c) \cdot e^{-\lambda_c(t - t_c)},$$

где $\mathbf{I}(x)$ — функция Хевисайда.¹

В логе загруженности процессора наблюдается общая загрузка процессора:

$$X(t) = B(t) + \sum_{t_c \in T_c} K_{t_c}(t),$$

где T_c — это моменты времени поступления заявок.

Необходимо по дискретным наблюдениям $X(t_i)$, $i = 1, \dots, N$

1. оценить моменты времени поступления заявок T_c ,
2. оценить параметры модели $m, \sigma, \lambda, m_c, \sigma_c^2, \lambda_c$.

Наблюдения производятся через равные промежутки времени $\Delta t = t_{i+1} - t_i$.

Для упрощения решения λ_c принимается равным величине близкой к нулю, т.е. каждая пришедшая заявка увеличивает загрузку процессора на некоторую фиксированную величину.

2 Решение

2.1 Оценка моментов времени поступления заявок

Предположим, что в отрезке времени $[t_k, t_{k+n+1}]$ не пришло ни одной заявки. Тогда наблюдения $X(t_k), \dots, X(t_{k+n+1})$ представляют собой наблюдения $B(t)$. Оценим по этим наблюдениям параметры $B(t)$.

Рассмотрим разности соседних наблюдений — они представляют собой наблюдения нормально распределённой случайной величины:

$$B(t_{i+1}) - B(t_i) = \sigma \mathcal{W}(t_{i+1}) - \sigma \mathcal{W}(t_i) = \sigma \mathcal{N}(0, \Delta t) = \mathcal{N}(0, \sigma^2 \Delta t).$$

Построим точечную оценку $\hat{\sigma}$ методом максимального правдоподобия:²

$$\hat{\sigma}^2 = \frac{1}{\Delta t} \cdot \frac{1}{n-1} \sum_{i=1}^n ((X(t_{k+i+1}) - X(t_{k+i})) - 0)^2.$$

Обозначим гипотезу о том, что в промежутке времени $[t_{k+n+1}, t_{k+n+2}]$ не пришло ни одной заявки, как H_0 . Тогда

$$\mathcal{L}(X(t_{k+n+2}) - X(t_{k+n+1}) | H_0) = \mathcal{N}(0, \sigma^2 \Delta t).$$

В качестве статистики для отвержения гипотезы H_0 возьмём вероятность наблюдения $X(t_{k+n+2})$:

$$T(X_{k+n+2}) = \mathbf{P}(B(t_{k+n+2}) = X(t_{k+n+2})) = \mathbf{P}(\mathcal{N}(0, \hat{\sigma}^2 \Delta t) = X(t_{k+n+2}) - X(t_{k+n+1})),$$

¹Функция Хевисайда: $\mathbf{I}(x) = \begin{cases} 0, & x < 0 \\ 1, & x \geq 0 \end{cases}$.

²§ 3.5 пункт 1 в [?].

а критерием отвержения гипотезы H_0 с уровнем значимости α будет служить следующее выражение:

$$H_0 \text{ rejected} \iff T(X_{k+n+2}) < \alpha.$$

Алгоритм нахождения моментов времени поступления заявок T_c состоит в следующем:

1. В предположении, что в первые $n + 1$ наблюдений не пришло ни одной заявки, оценим $\hat{\sigma}$ и построим критерий для отвержения H_0 .
2. Будем добавлять к первым $n + 1$ наблюдениям по одному наблюдению и проверять гипотезу H_0 . Если H_0 не отвергается, то $\hat{\sigma}$ и критерий для отвержения H_0 пересчитываются для добавленного наблюдения.
3. Как только встретится наблюдение $n + 1 + l$, для которого гипотеза H_0 отвергается, то $t_{n+1+l} \in T_c$. Все наблюдения до $t_{n+1+l+1}$ отбрасываются и алгоритм начинается с шага 1 для поиска следующего момента времени прихода заявки.

2.2 Оценка интенсивности поступления заявок λ

Зная время прибытия заявок T_c интенсивность поступления заявок можно оценить методом максимального правдоподобия:³

$$\hat{\lambda} = \frac{1}{|T_c|} \sum_{i=0}^{|T_c|-1} (t_{c_{i+1}} - t_{c_i}).$$

2.3 Оценка параметров распределения величины нагрузки поступающих заявок

Рассмотрим ненормированный разностный аналог производной $X(t)$:

$$dX(t) = X(t) - X(t - \Delta t).$$

$dX(t)$ в момент времени прихода заявки t_c выражается следующим образом:

$$\begin{aligned} dX(t_c) &= X(t_c) - X(t_c - \Delta t) = B(t_c) + K_{t_c}(t_c) - B(t_c - \Delta t) = \\ &= \sigma \mathcal{W}(t_c) - \sigma \mathcal{W}(t_c - \Delta t) + \mathcal{N}(m_c, \sigma_c^2) = \\ &= \sigma \mathcal{N}(0, \Delta t) + \mathcal{N}(m_c, \sigma_c^2) = \mathcal{N}(m_c, \sigma^2 \Delta t + \sigma_c^2), \end{aligned}$$

предполагая, что в момент времени $t_c - \Delta t$ заявки не было.

Во время отсутствия заявок $dX(t)$ выражается как:

$$dX(t) = X(t) - X(t - \Delta t) = B(t) - B(t - \Delta t) = \mathcal{N}(0, \sigma^2 \Delta t).$$

Значит $dX(t)$ представляет собой наблюдение смеси двух нормально распределённых случайных величин. Оценим их параметры ЕМ-алгоритмом.

Введём скрытые случайные величины Z_i , $i = 1, \dots, N$, принимающие значения 1 или 2, в зависимости от того, пришла ли заявка в момент времени t_i или нет соответственно.

$$X(t_i) | (Z_i = 1)$$

3 Результаты работы

³http://en.wikipedia.org/wiki/Poisson_distribution#Maximum_likelihood или в общем случае в §3.5 пункт 1 в [?].