# ВТОРЖЕНИЕ КАК РАЗЛАДКА

Баранов В.А.

vb0407@gmail.com

*ГОУ «Санкт-Петербургский государственный политехнический университет»*

# Представление системы

ВХОД $\rightarrow$ $\Phi = (a_1, a_2, \ldots, a_N)$ $\rightarrow$ ВЫХОД

$$A_1, A_2, \ldots, A_t\colon\ A_j \in \Phi,\ j = 1, \ldots, t$$

$$A_1, A_2, \ldots, A_t \longrightarrow \xi_1, \xi_2, \ldots, \xi_t$$

# «Разладка»

Стабильная работа системы:

$$P(\xi_i = a_j) = p_j, j = 1, \dots, N, i = 1, \dots, t.$$

$$\xi_1, \xi_2, \dots, \xi_t \longrightarrow \xi_1, \xi_2, \dots, \xi_{\theta-1}, \xi_\theta, \dots, \xi_t,$$

$\theta$ – момент вторжения или сбоя.

$$\xi_1, \dots, \xi_{\theta-1}: P(\xi_i = a_j) = p_j, j = 1, \dots, N, i = 1, \dots, \theta\text{-}1,$$
$$\xi_\theta, \dots, \xi_n: P(\xi_i = a_j) = \pi_j, \quad j = 1, \dots, N, \quad i = \theta, \dots, t$$

Частотные распределения :

$$P = (p_1, \dots, p_N),$$
$$\Pi = (\pi_1, \dots, \pi_N),$$
$$P \neq \Pi$$

# Метод обнаружения разладки

$$\hat{\theta} = \arg\max_{1 \leq k \leq t} f(k)\,,$$

$k$ – соответствует точке разбиения последовательности наблюдений длины $t$;
$\hat{\theta}$ – предполагаемый момент разладки;
$f(k)$ – статистическая функция;
$\mu_i(k)$ – набор частот $i$-ого исхода на отрезке наблюдений $1, \dots, k$;
$\nu_i(k)$ – набор частот $i$-ого исхода на отрезке наблюдений $k+1, \dots, t$.

# Статистика $Cr(k, t-k, \lambda)$

$$Cr(k, t-k, \lambda) =$$

$$\left(c_{\lambda+1}^{2}\right)^{-1} \sum_{i=1}^{N} \left(\left(\frac{\mu_i(k)t}{k\left(\mu_i(k)+\nu_i(k)\right)}\right)^{\lambda} - 1\right)\mu_i(k) +$$

$$+\left(\left(\frac{\nu_i(k)t}{(t-k)\left(\mu_i(k)+\nu_i(k)\right)}\right)^{\lambda} - 1\right)\nu_i(k)$$

$\lambda$ – некоторый действительный параметр, определяющий конкретный вид статистики

# Статистика $L(k, t-k)$

$$L(k, t-k) =$$

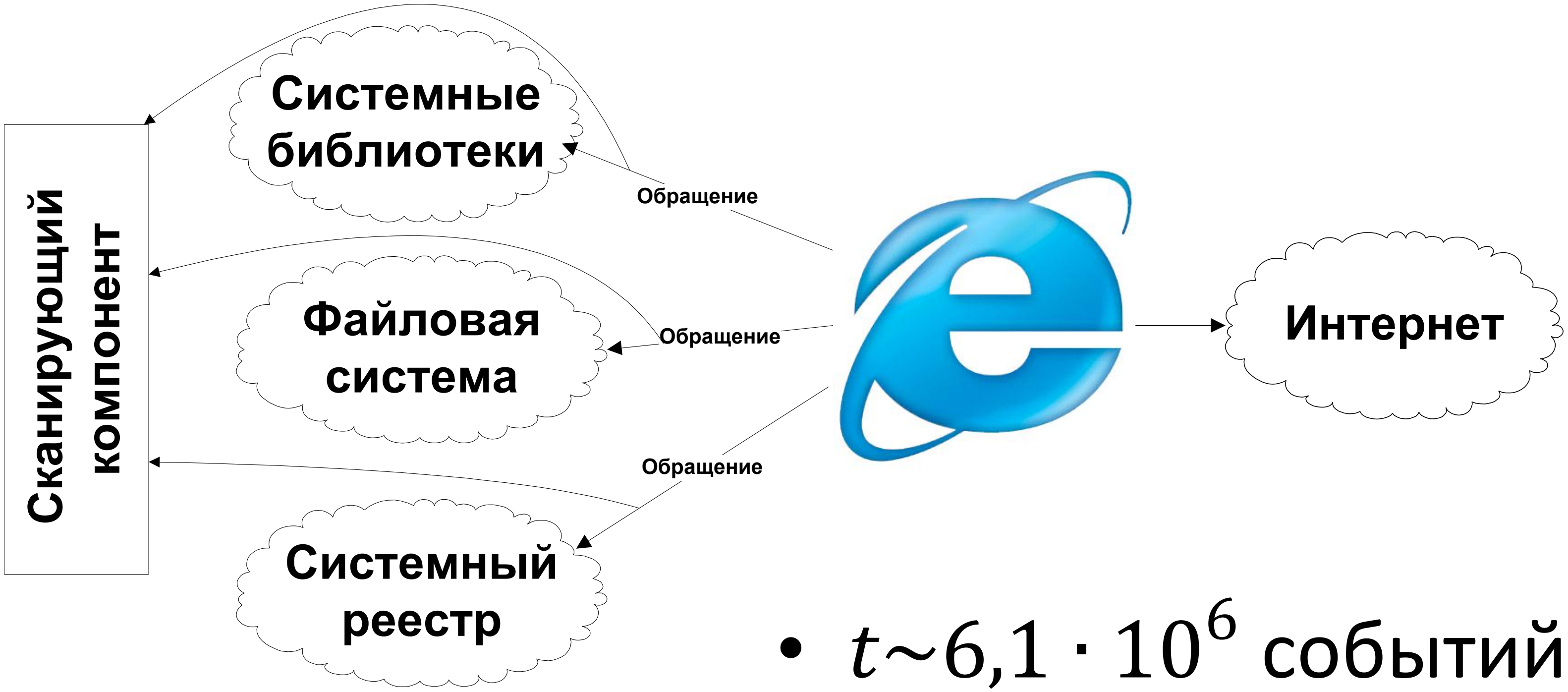
$$\sum_{i=1}^{N} \mu_i(k) \ln \frac{t\mu_i(k)}{k\big(\mu_i(k) + \nu_i(k)\big)} +$$

$$+\nu_i(k) \ln \frac{t\nu_i(k)}{(t-k)(\mu_i(k) + \nu_i(k))}$$

# Статистика $Y_t(k)$

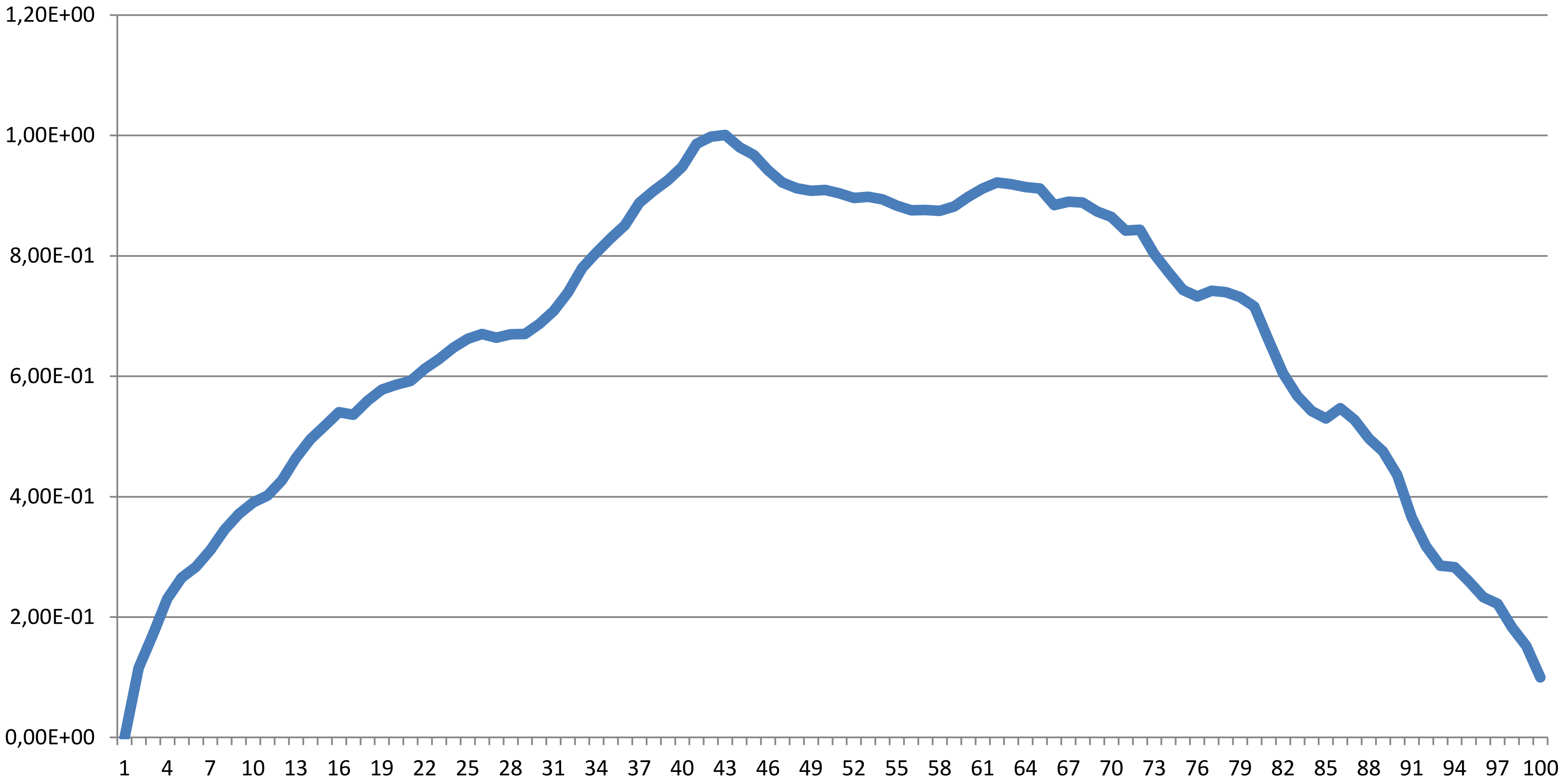
$$Y_t(k) = \sum_{i=1}^{N} \left( \frac{\mu_i(k)}{k} - \frac{\nu_i(k)}{t-k} \right)^2$$

# Экспериментальное исследование



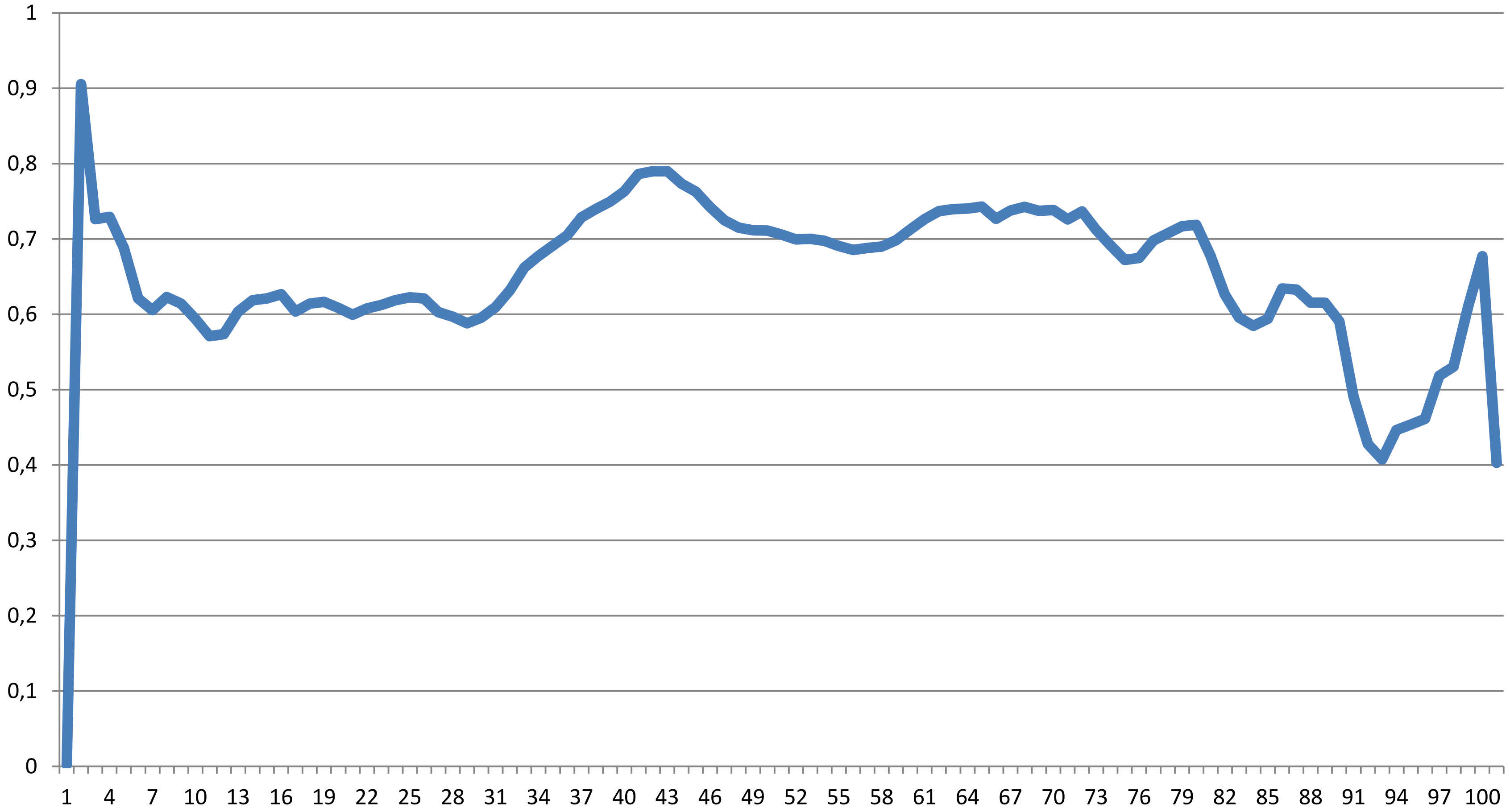

- $t \sim 6{,}1 \cdot 10^6$ событий
- $\theta \sim 2{,}5 \cdot 10^6$
- $N \sim 2{,}2 \cdot 10^4$

# Среднее для $Cr(k, t-k, \lambda)$



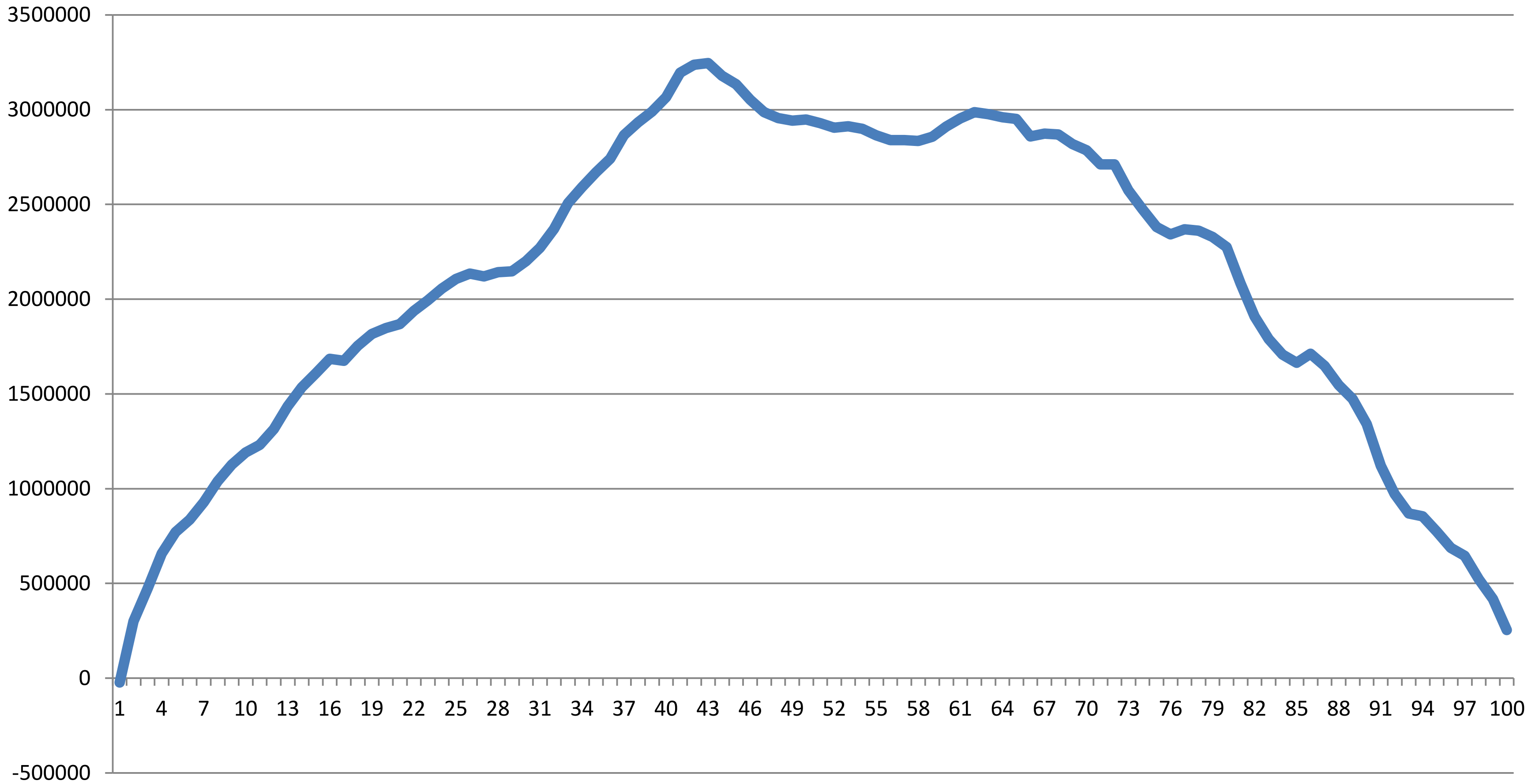

$$\lambda = 0{,}1, \hat{\theta} = 43, \theta \sim 41$$

# Среднее для $Cr(k, t-k, \lambda)$



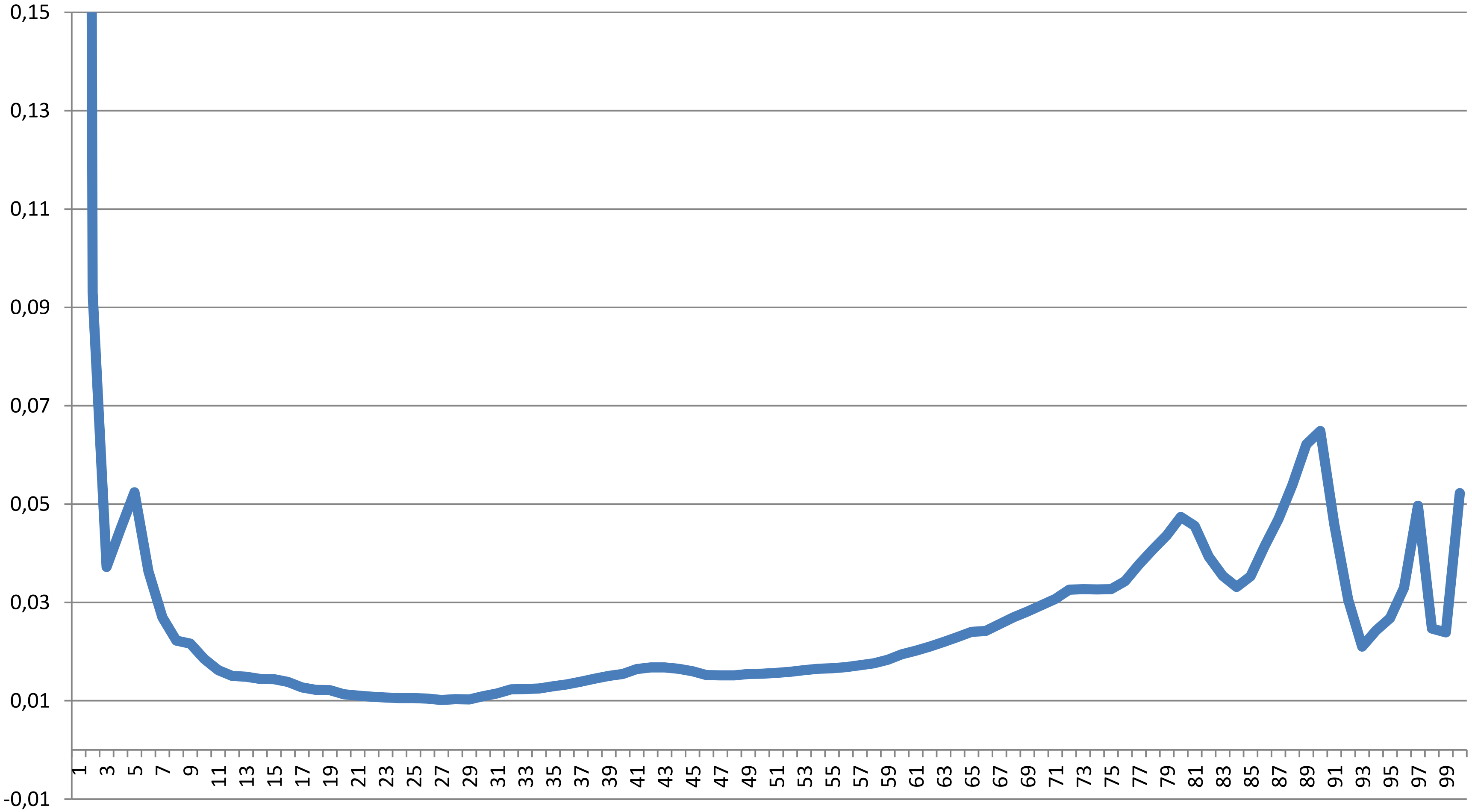

$\lambda = 1, \hat{\theta} = 2, \theta \sim 41$

# Среднее для $L(k, t-k)$




$$\hat{\theta} = 43, \theta \sim 41$$

# Среднее для $Y_t(k)$




$$\hat{\theta} = 1, \theta \sim 41$$

# Выводы

- При помощи статистики $Y_t(k)$ не удалось определить момент разладки;
- Статистика $Cr(k, t-k, \lambda)$, при $\lambda \to 0$ эффективна для определения момента разладки;
- При использовании статистики $Cr(k, t-k, \lambda)$, при увеличении значения параметр $\lambda$, момент разладки «размывается», что делает статистику неприменимой;
- Статистика $L(k, t-k)$ и $Cr(k, t-k, \lambda)$ при $\lambda = 0{,}1$ ожидаемо имеют практически идентичный вид. $L(k, t-k)$ – эффективна для определения момента разладки;
- Статистики $Y_t(k)$, $Cr(k, t-k, \lambda)$ и $L(k, t-k)$ теоретически ожидались одинаково работоспособны для данных полностью соответствующих моделям.

# Спасибо за внимание!