

---

# Индекс нестационарности временного ряда и примеры его применения

- 1. Теоретические аспекты:** согласованный уровень стационарности и индекс нестационарности временного ряда; оптимальный объем выборки и горизонт прогноза; индикатор разладки
- 2. Достоверность прогнозов по землетрясениям**
- 3. Анализ энцефалограмм:** прогноз приступа эпилепсии

---

I

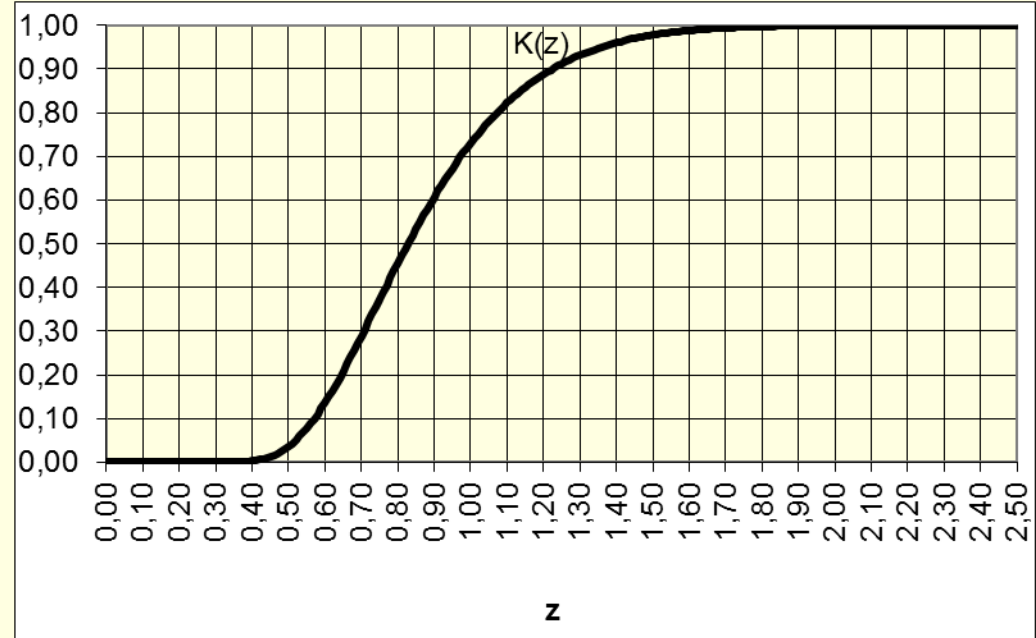
Согласованный уровень  
стационарности  
и индекс нестационарности

# Непараметрическое сравнение выборок

Пусть случайные величины имеют непрерывное стационарное распределение и являются независимыми. Тогда вероятность того, что две выборки объема  $N$  различаются между собой в норме  $S$  менее, чем на  $\varepsilon$ , равна

$$K(\varepsilon\sqrt{N/2})$$

$$S_N = \sup_x |F_{1,N}(x) - F_{2,N}(x)|$$



$$\lim_{N \rightarrow \infty} P \left\{ 0 < \sqrt{\frac{N}{2}} S_N < z \right\} = K(z)$$

# Согласованный уровень значимости

На каком расстоянии следует отцепить «хвост» распределения расстояний между распределениями, чтобы оставшийся квантиль был бы равен эмпирически наблюдаемому уровню доверия в задаче распознавания «своих» выборок?

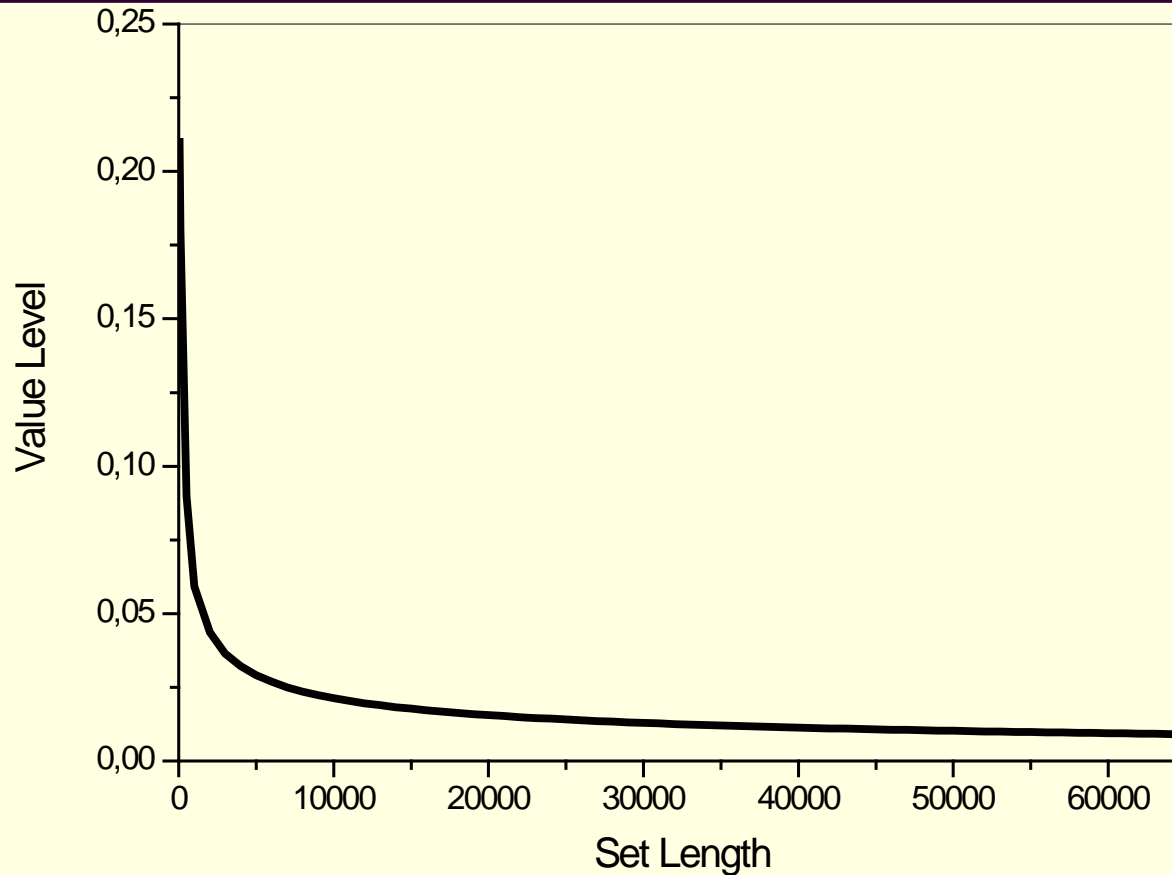
Известно, что если СВ  $r$  (это расстояние между выборками) имеет ФР  $F(x)$ , то СВ  $u=F(r)$  (в нашей задаче это уровень значимости) имеет равномерное распределение на  $[0;1]$ . Следовательно, согласованный с экспериментом уровень значимости как квантиль равномерно распределенной СВ есть функция, линейно зависящая от расстояния между выборками, т.е.  $\alpha = \varepsilon$ .



В норме  $S$  две выборки длины  $N$ , расстояние между которыми  $\varepsilon$ , различны на уровне значимости  $\alpha$ , если

$$1 - K\left(\sqrt{\frac{N}{2}}\varepsilon\right) < \alpha$$

# Уровень стационарного шума в зависимости от длины $N$ выборки



$$1 - \varepsilon = K \left( \sqrt{\frac{N}{2}} \varepsilon \right)$$

$$\varepsilon = \varepsilon^* (N)$$

Согласованный уровень значимости: доля расстояний, превышающих его, равна уровню значимости критерия

# Расстояние между нестационарными ВПФР

1. Ряд расстояний между встык-выборками переменной длины:

$$\rho_N(t) = \|F_N(x, t + N) - F_N(x, t)\|_C$$

2. Функция распределения расстояний между распределениями:

$$G_N(r) = \text{Prob}\{\rho_N \leq r\}$$

3. Численно решается уравнение по нахождению СУС:

$$G_N(r) = 1 - r \quad \Rightarrow \quad r = r^*(N)$$

4. Индекс нестационарности:

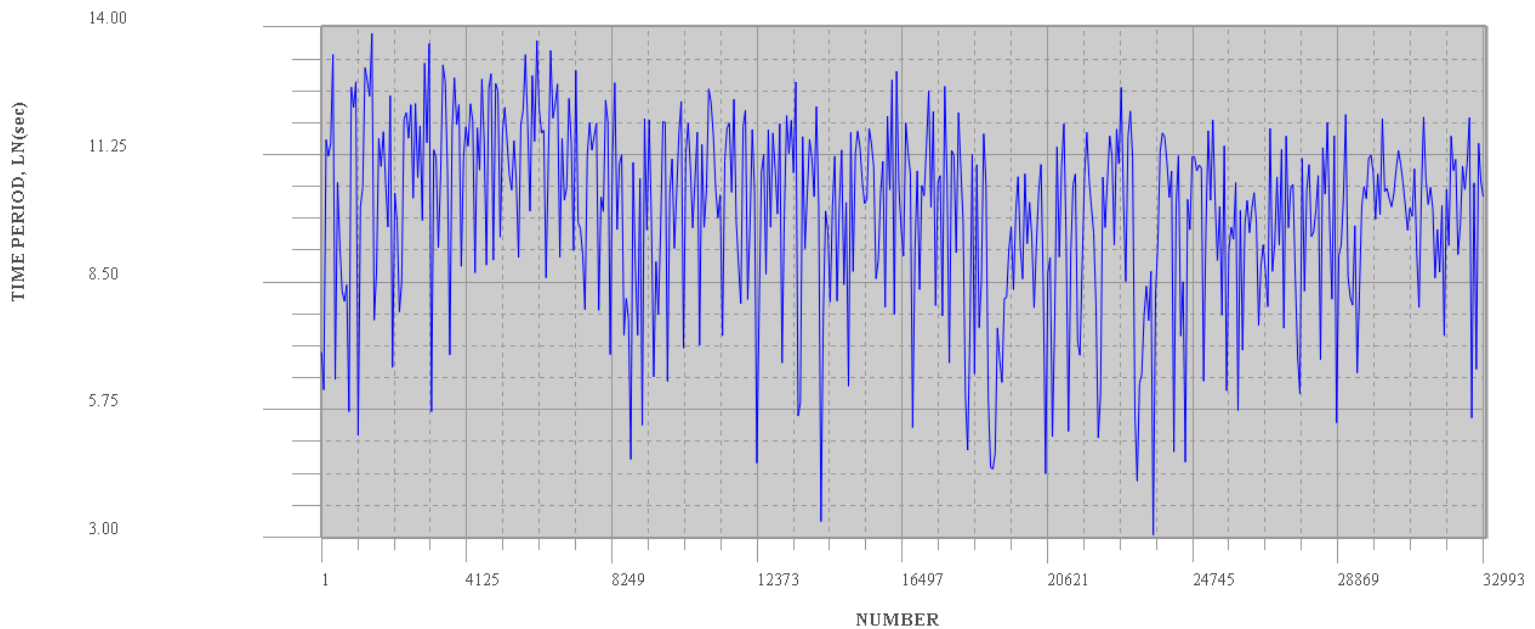
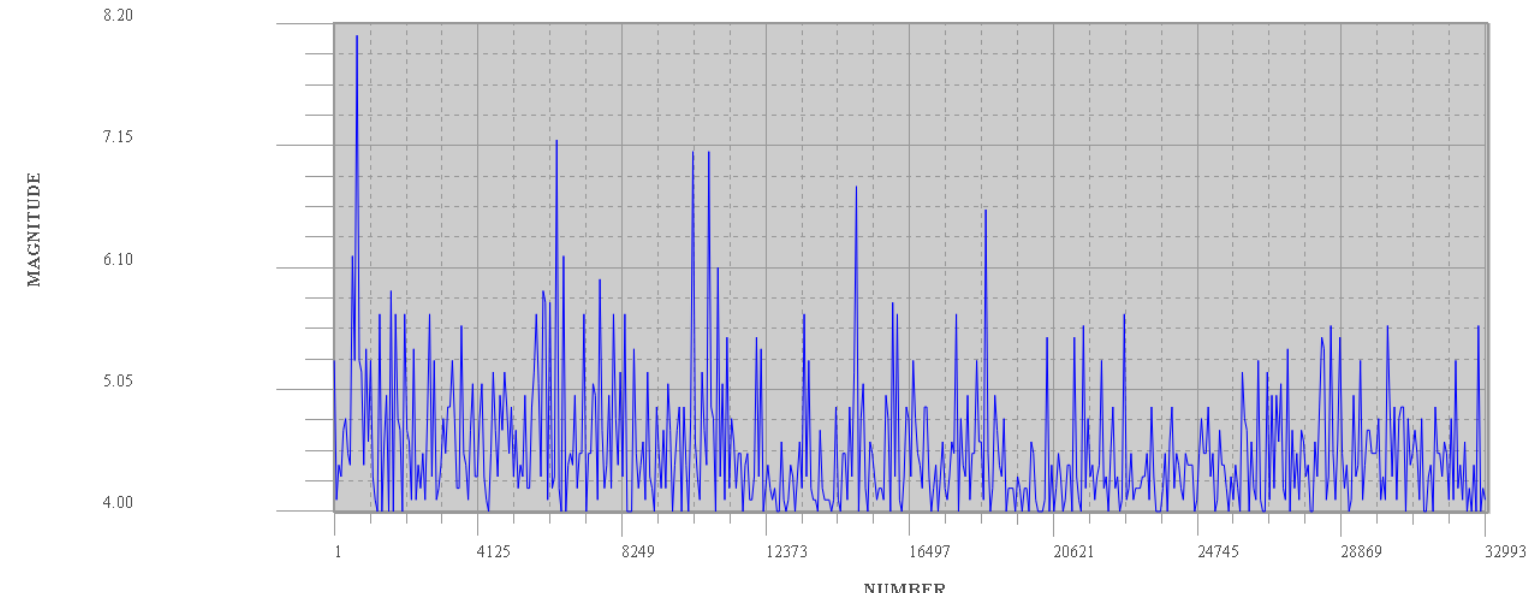
$$J(N) = \frac{r^*(N)}{\varepsilon^*(N)}$$

---

## II

# Точность прогнозирования землетрясений (анализ мирового атласа по Японии с 1900 г.)

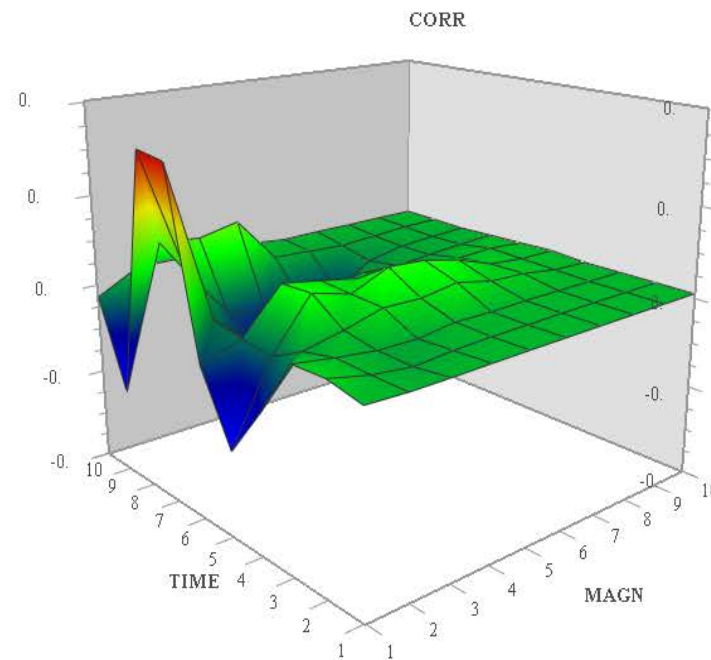
# Результаты наблюдений за землетрясениями





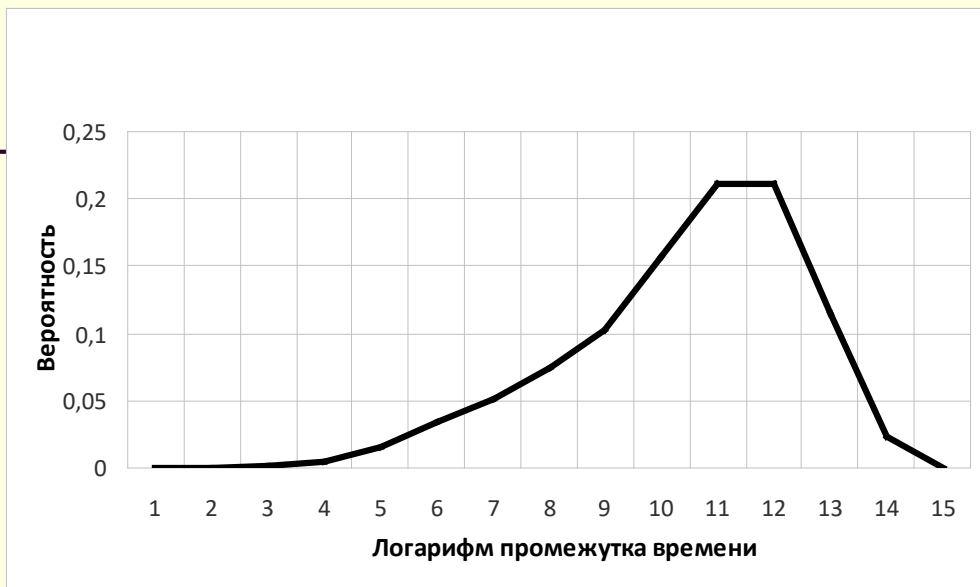
# Корреляционная функция «магнитуда-время»

$$R(x,t) = F(x,t) - f(x)\varphi(t)$$

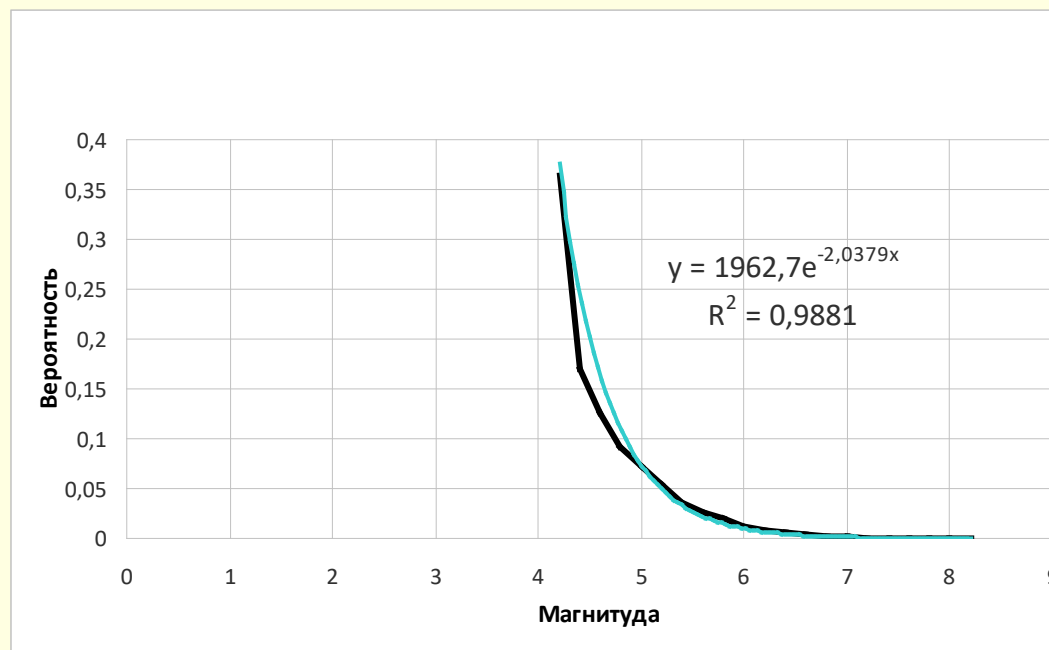


Магнитуда и промежутки времени между событиями независимы: максимум модуля корреляционной функции 0,01 при максимуме совместного распределения 0,20.

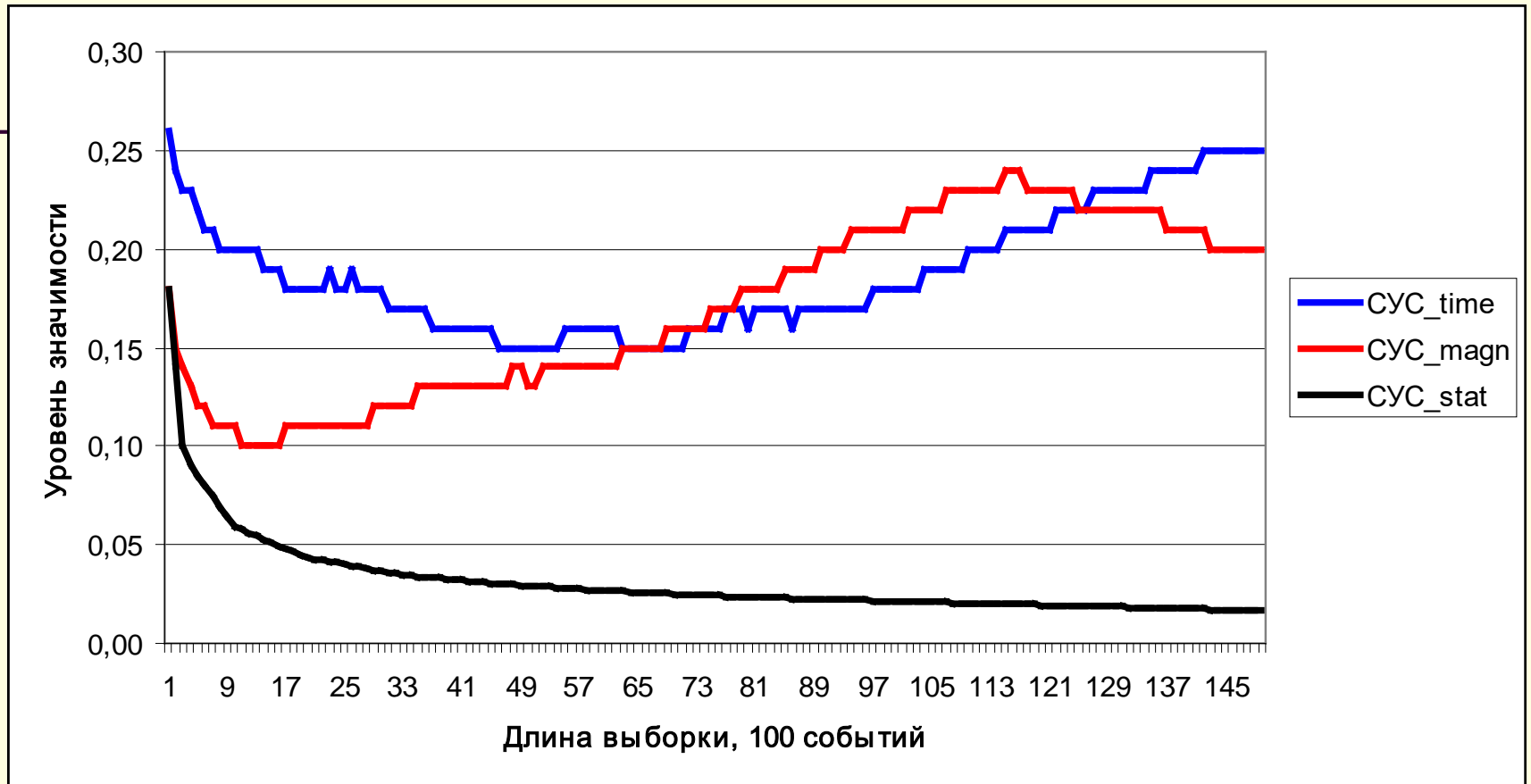
# Плотности распределений



Плотности  
распределений магнитуд  
и промежутков времени  
за 116 лет  
(34 тыс. событий)

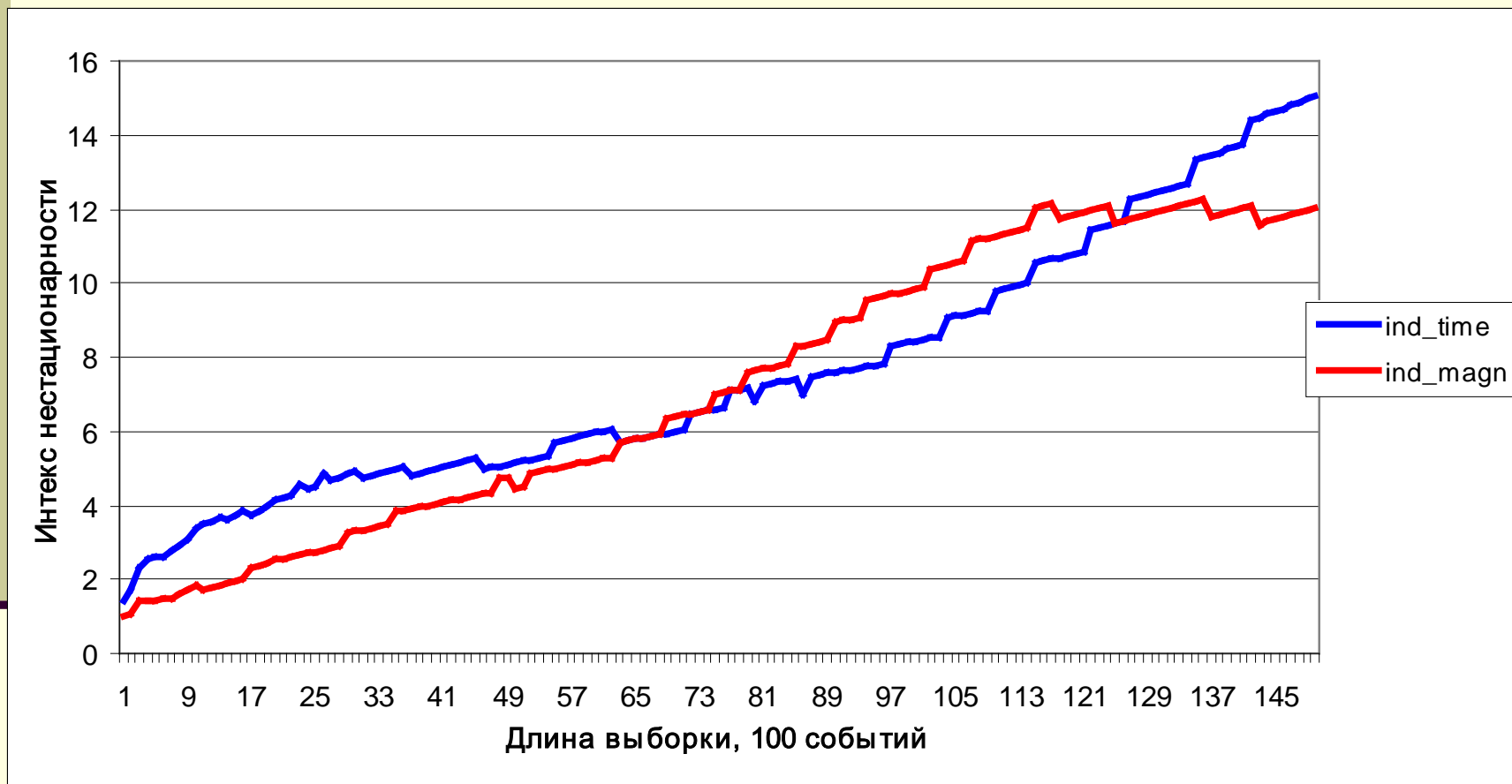


# Согласованный уровень стационарности



Наилучшая значимость по магнитуде 0,10 по выборке 1000-1200 данных (3-4 года), а по промежуткам времени 0,15 по выборке 4500-7000 данных (13-21 лет)

# Индекс нестационарности



# III

Анализ энцефалограмм:  
индикация разрядки перед  
приступом эпилепсии

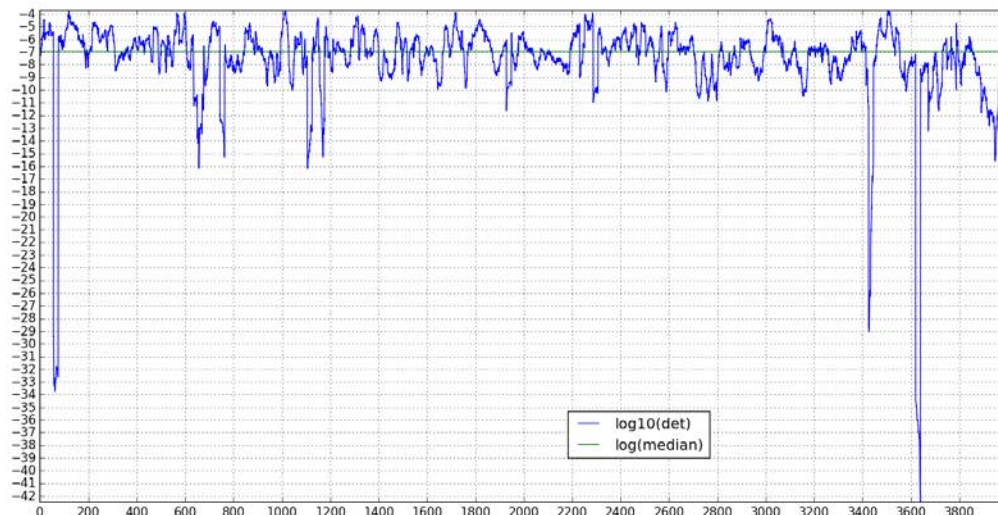
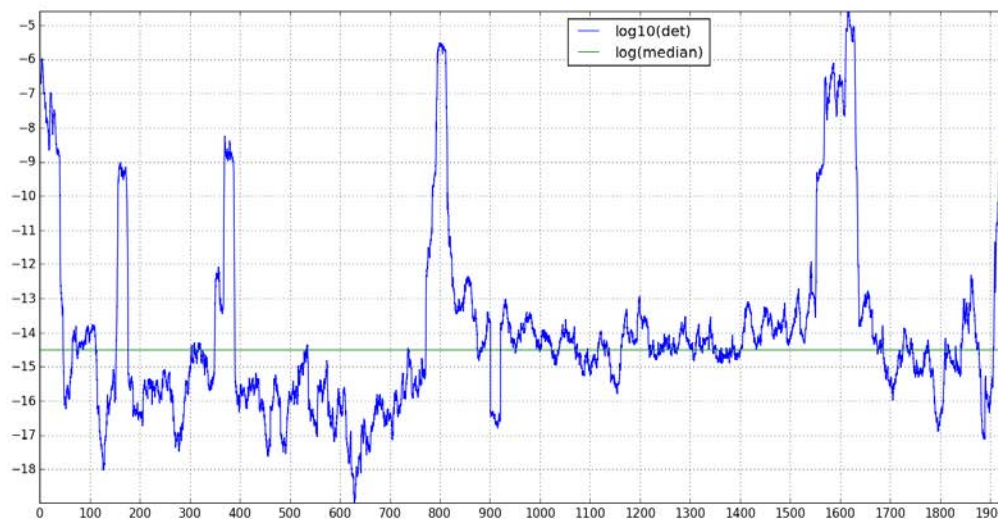
# Определитель корреляционной матрицы

Исходные  
данные:

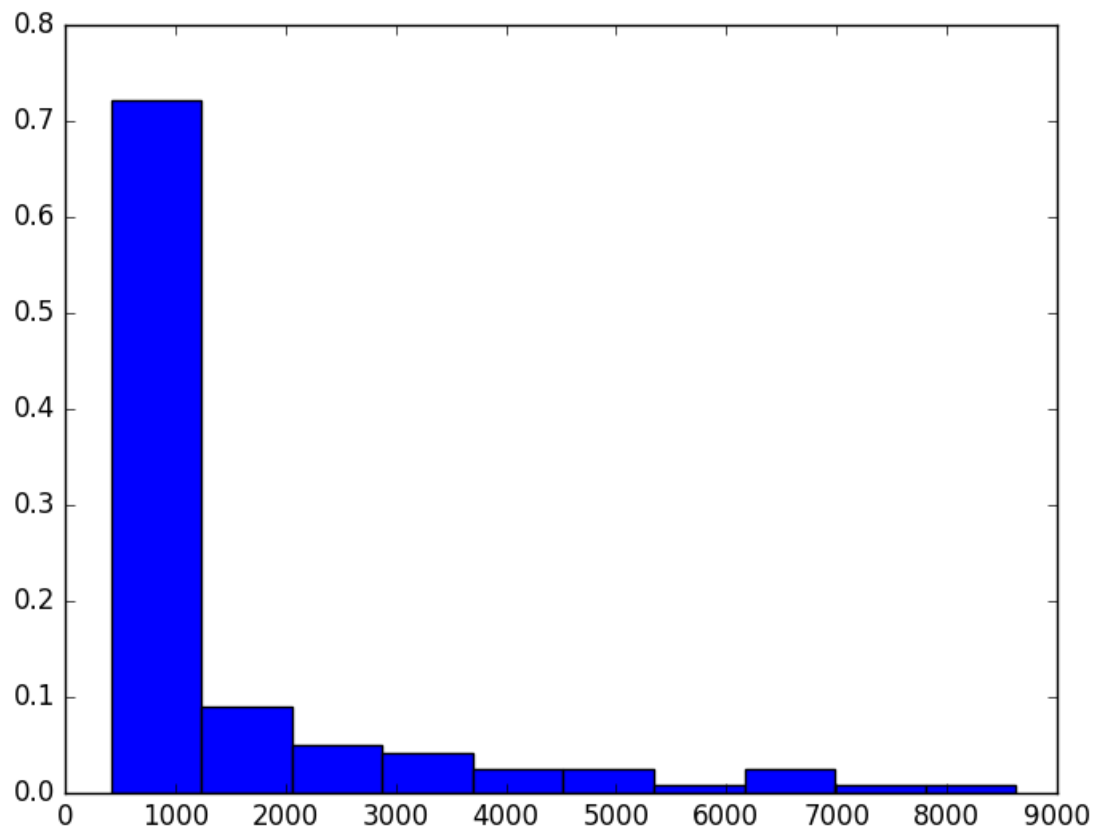
22 ряда длиной  
2 млн данных  
(частота 512 Гц);

Определитель  
считался в окне  
30 тыс. (1 мин.);

По-видимому,  
не является  
надежным  
индикатором

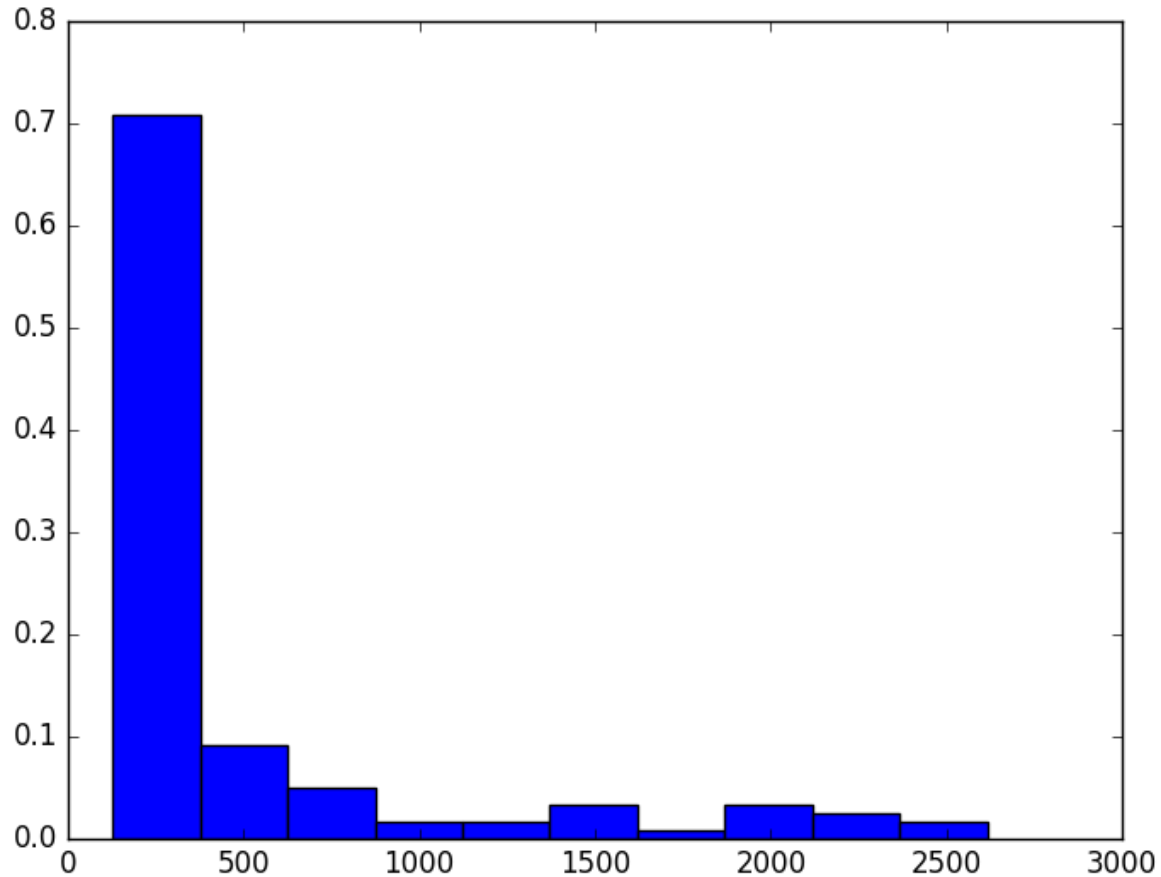


# Распределение промежутков max-max



Для стационарного ряда это распределение геометрического типа; здесь же «хвост» убывает медленнее, распределение нестационарно.

# Распределение промежутков max-min



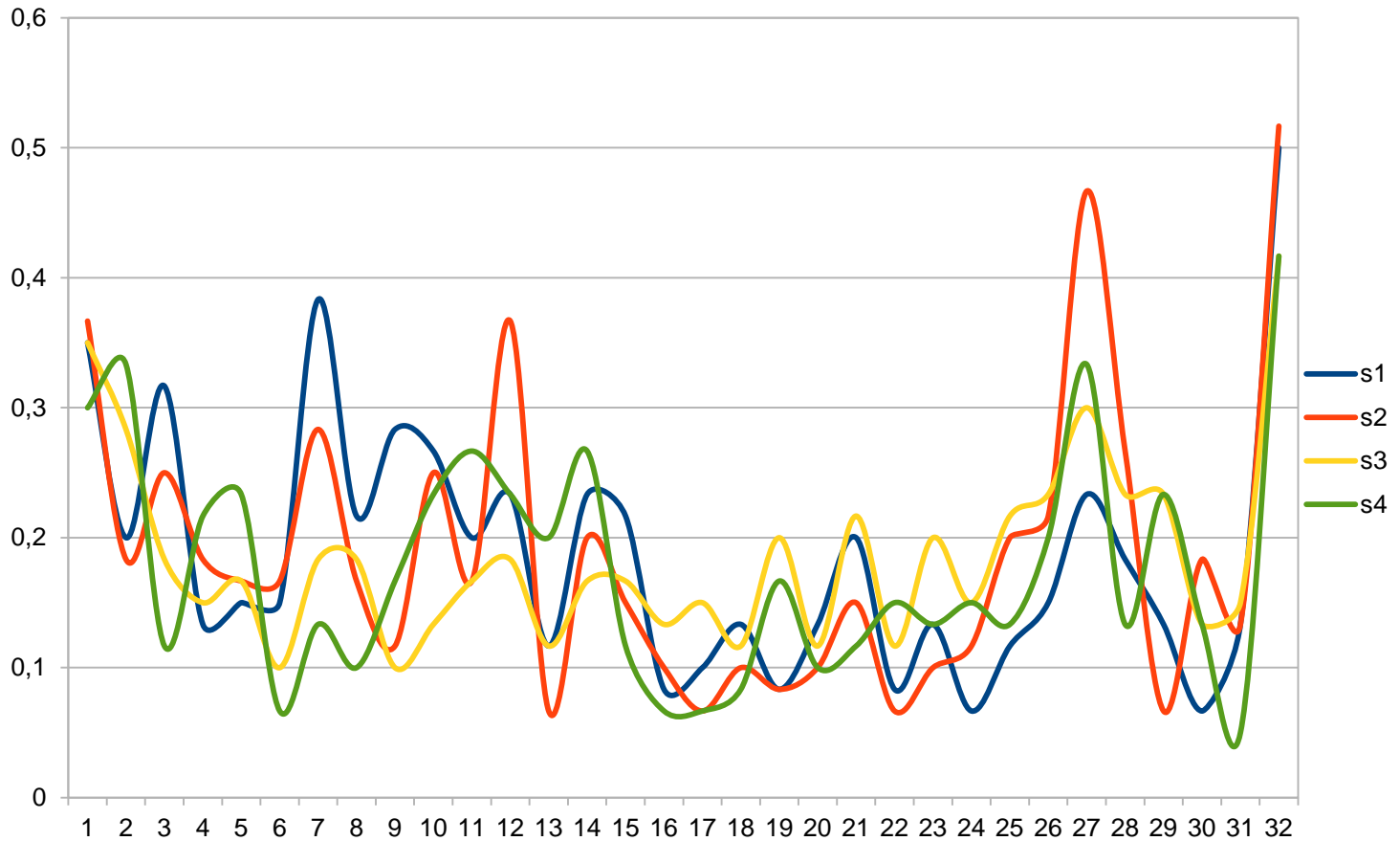


# Согласованный уровень стационарности рядов для выборки длины 5 тыс. (10 сек.)

---

СУС вычислялся по всему набору данных (2 млн) в скользящем окне 5 тыс. с шагом 1 тыс. для каждого ряда (22) для трех пациентов

# Динамика СУС в окне 60 тыс. (2 мин), первый образец

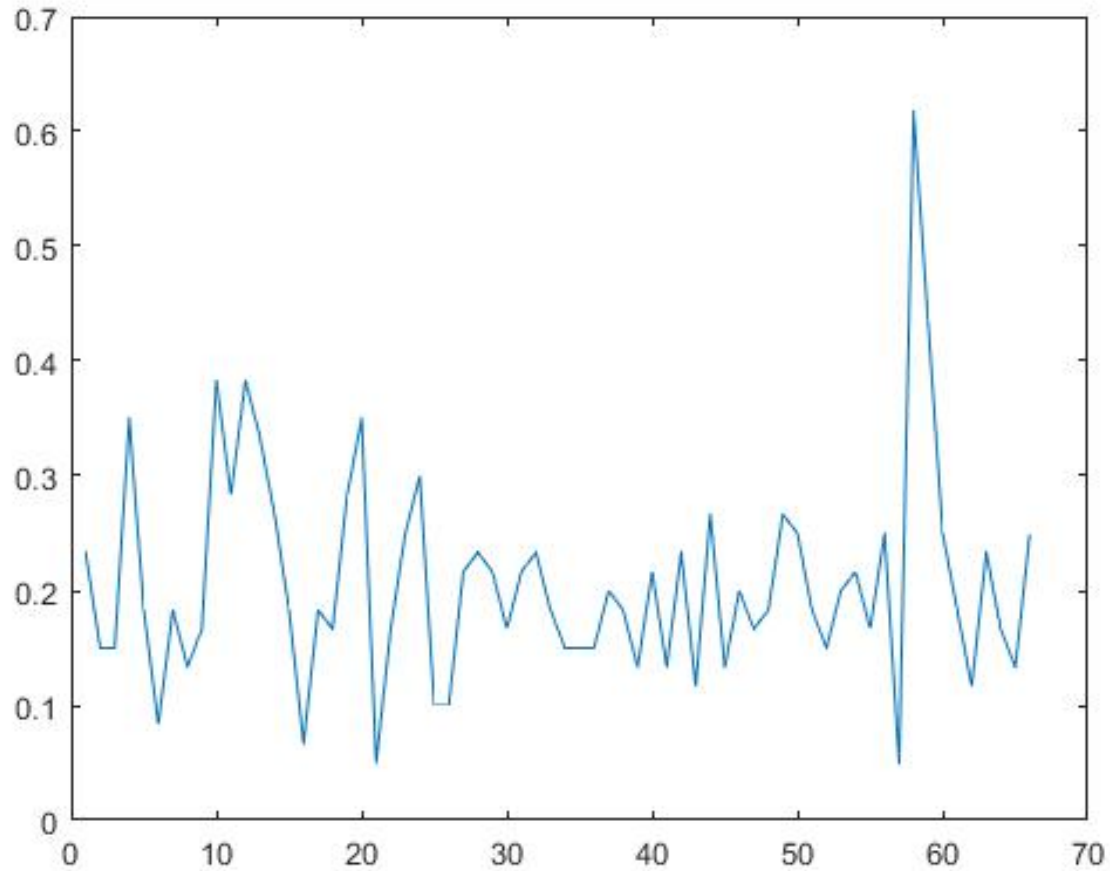


# Динамика СУС в окне 60 тыс. (2 мин), второй образец

---

В последние две минуты наблюдается значительное превышение СУС (в 2 раза) по более чем половине рядов.

# Усредненное движение СУС для выборки длины 5 тыс. в окне 60 тыс. (2 мин)



# ВЫВОДЫ

---

1. Во многих практических задачах (анализ биржевых рядов, геофизических и медицинских данных) статистика расстояний между распределениями и индикатор СУС оказываются весьма эффективными для получения количественных оценок
2. В отсутствие надежных физико-математических моделей нестационарные временные ряды следует анализировать кинетическим методом, т.е. изучать эволюцию их выборочных распределений
3. Применительно к анализу разладки для энцефалограмм необходимо собрать более репрезентативную статистику, а также изучить возможность выделения характерных паттернов функций распределения

**СПАСИБО ЗА ВНИМАНИЕ**