

И.П. Симаков

**Презентация лекции  
на тему:**

**Алгоритмическое и программное обеспечение  
для решения задач обработки статистической  
информации о наработках и отказах объектов при  
эксплуатации**

# ЦЕЛИ И ЗАДАЧИ РАБОТЫ

**Цель работы** - развитие вычислительных процедур и алгоритмов и разработка программного обеспечения статистической обработки информации, получаемой из сферы эксплуатации, для решения задач объективной оценки характеристик и показателей надежности оборудования с использованием методов моментов и максимального правдоподобия для полных и цензурированных выборок без обращения к огромному числу таблиц.

## **Задачи работы:**

1. Анализ и практическое освоение известных классических методов и алгоритмов обработки статистических данных для информационно-аналитических систем различного назначения.
2. Разработка эффективных алгоритмов и программ реализации методов моментов и максимального правдоподобия для обработки статистической информации по полным и цензурированным выборкам.
3. Отработка алгоритмов и программ статистической проверки гипотез о теоретическом законе распределения с применением критерия согласия А. Н. Колмогорова и системы неравенств, устанавливающих принадлежность функции распределения к классу функций с возрастающей интенсивностью «опасности» (или интенсивностью отказов для технических систем).
4. Решение конкретных задач обработки экспериментальных (наблюдаемых) данных с «распознаванием» теоретической функции распределения и оценкой ее параметров.

## Расчет оценок начальных и центральных моментов

по выборке независимых наблюдений  $t_1, \dots, t_N$  над случайной величиной

### 1. Оценки первых четырех начальных выборочных моментов

$$\hat{m}_k = \frac{1}{N} \sum_{i=1}^N t_i^k \quad k = 1, 2, 3, 4, \dots$$

### 2. Оценки первых четырех центральных моментов $\mu_1, \mu_2, \mu_3, \mu_4$ по известным связующим соотношениям между ними

$$\mu_1 \equiv \bar{T} = \hat{m}_1 \quad - \text{математическое ожидание случайной величины}$$

$$\mu_2 = \hat{m}_2 - \hat{m}_1^2 \quad - \text{дисперсия}$$

$$\mu_3 = \hat{m}_3 - 3\hat{m}_1\hat{m}_2 + 2\hat{m}_1^3 \quad - \text{характеристика асимметрии распределения}$$

$$\mu_4 = \hat{m}_4 - 4\hat{m}_1\hat{m}_3 + 6\hat{m}_1^2\hat{m}_2 - 3\hat{m}_1^4 \quad - \text{характеристика островершинности}$$

### 3. Оценки центральных моментов по выборке

$$\hat{\mu}_2 = \hat{D} = \frac{1}{N} \sum_i (t_i - \hat{T})^2; \quad \hat{\mu}_3 = \frac{1}{N} \sum_i (t_i - \hat{T})^3; \quad \hat{\mu}_4 = \frac{1}{N} \sum_i (t_i - \hat{T})^4; \quad S\hat{k} = \frac{\hat{\mu}_3}{\hat{\sigma}^3}; \quad E\hat{x} = \frac{\hat{\mu}_4}{\hat{\sigma}^4} - 3.$$

### 4. Расчет несмещенных центральных моментов

$$\hat{T} = \frac{1}{N} \sum_{i=1}^N t_i; \quad \hat{\mu}_{2H} = \frac{1}{N-1} \sum_{i=1}^N (t_i - \hat{T})^2; \quad \hat{\sigma}_H = \sqrt{\hat{\mu}_{2H}}$$

$$\hat{\mu}_{3H} = \frac{N^2}{(N-1)(N-2)} \hat{\mu}_3; \quad \hat{\mu}_{4H} = \frac{N(N^2 - 2N + 3)\hat{\mu}_4 - 3N(2N-3)\hat{\mu}_2^2}{(N-1)(N-2)(N-3)}; \quad S\hat{k}_H = \frac{\sqrt{N(N-1)}}{N-2} S\hat{k}; \quad E\hat{x}_H = \frac{N-1}{(N-2)(N-3)} [(N+1)E\hat{x} + 6].$$

# Эмпирическая (выборочная) функция распределения

1. Выборка независимых наблюдений,  $t_N$  над случайной величиной  $\xi$

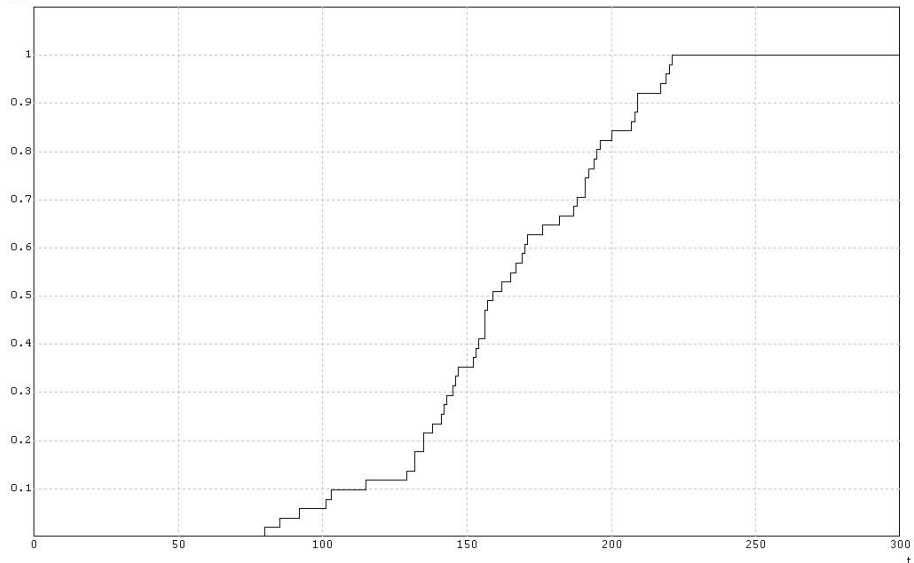
2. Вариационный ряд  $t_1^* \leq t_2^* \leq \dots \leq t_N^*$

3. Эмпирическая функция распределения

$$F_N(t)$$

$$F_N(t) = \begin{cases} 0, & t \leq t_1^* \\ \frac{k}{N}, & t_k < t \leq t_{k+1}^* \\ 1, & x > x_N^* \end{cases}$$

График эмпирической функции распределения  $F_N(t)$



4. Выборочные несмещенные центральные моменты

$$\hat{T} = \frac{1}{N} \sum_{i=1}^N t_i; \hat{\mu}_{2H} \equiv \hat{D} = \frac{1}{N-1} \sum_{i=1}^N (t_i - \hat{T})^2; \hat{\sigma}_H = \sqrt{\hat{\mu}_{2H}}$$

$$\hat{\mu}_{3H} = \frac{N^2}{(N-1)(N-2)} \hat{\mu}_3; \hat{\mu}_{4H} = \frac{N(N^2 - 2N + 3)\hat{\mu}_4 - 3N(2N-3)\hat{\mu}_2^2}{(N-1)(N-2)(N-3)}; S\hat{k}_H = \frac{\sqrt{N(N-1)}}{N-2} S\hat{k}; E\hat{x}_H = \frac{N-1}{(N-2)(N-3)} [(N+1)E\hat{x} + 6]$$

На основе полученных численных значений оценок центральных моментов могут выдвигаться гипотезы о предполагаемых теоретических законах распределения.

## Метод моментов

для точечной оценки неизвестных параметров заданного распределения



Метод моментов предложен в 1894 г..  
**Карлом Пирсоном** (1857 – 1936) -  
английский математик-статистик,  
биолог, философ, основоположник  
знаменитого журнала «Биометрика».

**Идея метода** проста - приравнивание теоретических моментов соответствующим эмпирическим моментам того же порядка. Если распределение определяется двумя параметрами, то приравнивают два теоретических момента двум соответствующим эмпирическим моментам того же порядка. Заметим – метод не использует информацию о третьем и четвертом выборочных моментах.

## Метод моментов

для точечной оценки неизвестных параметров заданного распределения

1. Для нахождения двух неизвестных параметров  $a$  и  $b$  распределения

**Вейбулла**

$$F(t) = 1 - \exp\left(-\left(\frac{t}{b}\right)^a\right), \quad f(t) = \frac{a}{b}\left(\frac{t}{b}\right)^{a-1} \cdot \exp\left(-\left(\frac{t}{b}\right)^a\right), \quad \lambda(t) = \frac{a}{b} \cdot \left(\frac{t}{b}\right)^{a-1},$$

достаточно использовать два первых

соотношения:

$$T = b \cdot \Gamma\left(1 + \frac{1}{a}\right), \quad D = b^2 \cdot \left\{ \Gamma\left(1 + \frac{2}{a}\right) - \left[ \Gamma\left(1 + \frac{1}{a}\right) \right]^2 \right\},$$

заменив теоретические значения моментов их выборочным несмещенным оценкам.

Выражая  $b$  из соотношения для первого момента и подставляя во второе соотношение, получаем сложное алгебраическое уравнение для нахождения параметра  $a$

$$\hat{D} = \left( \frac{\hat{T}}{a} \right)^2 \cdot \left\{ \Gamma\left(1 + \frac{2}{a}\right) - \left[ \Gamma\left(1 + \frac{1}{a}\right) \right]^2 \right\},$$

где  $\Gamma(x)$  — гамма-функция или Эйлеров интеграл второго рода

$$\Gamma(x) = \int_0^{\infty} t^{x-1} \cdot e^{-t} \cdot dt,$$

численные значения которого заданы **таблично в интервале**  $1 \leq x \leq 2$ .

Значком «тильдой» означают выборочные моменты.

## Метод моментов

для точечной оценки неизвестных параметров заданного распределения

2. Для нахождения двух неизвестных параметров  $\lambda$  и  $a$  Гамма - распределения с плотностью вероятности наработки до отказа

$$f(t) = \frac{\lambda^a}{\Gamma(a)} \cdot t^{a-1} \cdot e^{-\lambda t}; \quad t > 0,$$

где  $\lambda$  — параметр масштаба ( $\lambda > 0$ ),  $a$  — параметр формы ( $a > 0$ ),

$\Gamma(a)$  — Эйлера интеграл второго рода  $\Gamma(a) = \int_0^{\infty} t^{a-1} \cdot e^{-t} \cdot dt$

решение задачи таково  $a = \frac{\bar{T}^2}{D}$ ,  $\lambda = \frac{\bar{T}}{D}$ ,

где в правые части подставляются соответственно выборочные оценки математического ожидания и дисперсии

$$\hat{T} = \frac{1}{N} \sum_{i=1}^N t_i; \quad \hat{D} = \frac{1}{N-1} \sum_{i=1}^N (t_i - \hat{T})^2.$$

**(Заметим – метод не использует информацию о третьем и четвертом выборочных моментах)**

# Метод максимального правдоподобия

для точечной оценки неизвестных параметров заданного распределения



Предложен в 1912 г. Рональдом Айлмер Фишером (1890 -1962) - английский генетик, математик-статистик внес огромный вклад в теорию вероятностей и математическую статистику, является родоначальником дисперсионного анализа и вместе с К. Пирсоном заложил основы теории проверки статистических гипотез.

Идея метода такова.

1. Получена из генеральной совокупности полная выборка  $\{t_1, t_2, \dots, t_N\}$  - совокупность возможных значений независимых, одинаково распределенных случайных величин.

2. Принимается гипотеза о виде функции распределения или плотности вероятности  $f(t; a, b)$ , где  $a$  и  $b$  - параметры, подлежащие определению.

3. Вероятность получения выборки  $\{t_1, t_2, \dots, t_N\}$  равна  $f(t_1; a, b) \cdot f(t_2; a, b) \dots f(t_N; a, b) dt_1 dt_2 \dots dt_N$

4. Вводится функция правдоподобия  $L(t_1, \dots, t_N; a, b) = \prod_{i=1}^N f(t_i; a, b)$

которая должна быть максимизирована по параметрам  $(a, b)$ .

Удобно использовать функцию  $\ln L(t_1, \dots, t_N; \theta)$  и

имеющую максимум в той же точке, что

$$\left. \begin{aligned} \frac{\partial}{\partial a} (\ln L(t_1, \dots, t_N; \theta)) &= 0 \\ \frac{\partial}{\partial b} (\ln L(t_1, \dots, t_N; \theta)) &= 0 \end{aligned} \right\}$$

Необходимые условия оптимальности имеют вид:

$$(\hat{a}, \hat{b}).$$

При выполнении достаточно общих условий эти оценки являются состоятельными и асимптотически эффективными. Решение системы алгебраических уравнений дает оптимальные  $(\hat{a}, \hat{b})$

В общем случае оценки являются смещенными (см. на с. 544 в кн. Крамер Г. «Математические методы статистики». – М.: Мир, 1975. – 648 с.).



## Метод максимального правдоподобия

для точечной оценки неизвестных параметров заданного распределения

1. Для **экспоненциального** распределения с плотностью вероятности  $f(x) = \lambda e^{-\lambda x}$

функция правдоподобия имеет вид  $L(t_1, \dots, t_N; \lambda) = \lambda^N e^{-\lambda \sum_{i=1}^N t_i}$  или  $L(t_1, \dots, t_N; \lambda) = N \ln \lambda - \lambda \sum_{i=1}^N t_i$

Необходимое условие экстремума при этом имеет вид  $\frac{\partial \ln L(t_1, \dots, t_N; \lambda)}{\partial \lambda} = \frac{N}{\lambda} - \sum_{i=1}^N t_i = 0$ ,

из которого находим решение  $\hat{\lambda} = \frac{1}{\frac{1}{N} \sum_{i=1}^N t_i} = \frac{1}{\bar{T}}$ .

2. Для **нормального** распределения с плотностью  $f(x; \bar{\mu}, \sigma) = \frac{1}{\sigma \sqrt{2\pi}} e^{-\frac{1}{2} \left( \frac{x - \bar{\mu}}{\sigma} \right)^2}$

функция правдоподобия  $L(t_1, \dots, t_N; \bar{T}, \sigma) = \left( \frac{1}{\sigma \sqrt{2\pi}} \right)^N e^{-\frac{1}{2} \frac{\sum_{i=1}^N (t_i - \bar{T})^2}{\sigma^2}}$  или

$$\ln L(t_1, \dots, t_N; \bar{T}, \sigma) = -N \ln \sigma + \ln \left( \frac{1}{(\sqrt{2\pi})^N} \right) - \frac{1}{2} \cdot \frac{\sum_{i=1}^N (t_i - \bar{T})^2}{\sigma^2}.$$

Необходимые условия экстремума

$$\frac{\partial \ln(L(t_1, \dots, t_N; \bar{T}, \sigma))}{\partial \bar{T}} \equiv \frac{\sum_{i=1}^N t_i - N\bar{T}}{\sigma^2} = 0, \quad \frac{\partial \ln(L(t_1, \dots, t_N; \bar{T}, \sigma))}{\partial \sigma} \equiv -\frac{1}{2} \cdot \frac{N}{\sigma^2} + \frac{\sum_{i=1}^N (t_i - \bar{T})^2}{2\sigma^4} = 0$$

Решение этой системы алгебраической уравнений

$$\hat{\bar{T}} = \frac{1}{N} \sum_{i=1}^N t_i \quad \text{и} \quad \hat{\sigma}^2 = \frac{1}{N} \sum_{i=1}^N (t_i - \hat{\bar{T}})^2.$$

## Метод максимального правдоподобия

для точечной оценки неизвестных параметров заданного распределения

Для **Гамма** – распределения с плотностью вероятности

$$f(t; \lambda) = \frac{\lambda \cdot (\lambda t)^{a-1}}{\Gamma(a)} \cdot e^{-\lambda t}; (t \geq 0; a > 0, \lambda > 0)$$

Функция правдоподобия  $L(t_1, \dots, t_N; \lambda, a) = \left(\frac{\lambda}{\Gamma(a)}\right)^N \cdot \prod_{i=1}^N [(\lambda t_i)^{a-1} \cdot e^{-\lambda t_i}]$  или эквивалентная ей

$$\ln(L(t_1, \dots, t_N; \lambda, a)) = N \cdot a \cdot \ln(\lambda) - N \cdot \ln(\Gamma(a)) + (a-1) \cdot \sum_{i=1}^N \ln(t_i) - \lambda \sum_{i=1}^N t_i.$$

Необходимое условие экстремума:

$$\left\{ \begin{array}{l} \frac{\partial \ln(L(t_1, \dots, t_N; \lambda, a))}{\partial \lambda} = 0 \\ \frac{\partial \ln(L(t_1, \dots, t_N; \lambda, a))}{\partial a} = 0 \end{array} \right\} \Rightarrow \left\{ \begin{array}{l} N \cdot \frac{a}{\lambda} - \sum_{i=1}^N t_i = 0, \\ N \cdot \ln(\lambda) + \sum_{i=1}^N \ln(t_i) - N \cdot \frac{d\Gamma(a)}{da} \equiv 0 \end{array} \right.$$

Из первого уравнения следует  $\lambda = \frac{a}{\frac{1}{N} \cdot \sum_{i=1}^N t_i} \equiv \frac{a}{\hat{T}}$ . После подстановки во второе

уравнение получим уравнение для получения оценки параметра  $a$   $\frac{1}{N} \cdot \sum_{i=1}^N \ln\left(\frac{t_i}{\hat{T}}\right) = \frac{d(\ln(\Gamma(a)))}{da} - \ln(a) \equiv G(a)$ .

Или в развернутом виде:

$$\frac{1}{N} \cdot \sum_{i=1}^N \ln\left(\frac{t_i}{\hat{T}}\right) = \frac{d\Gamma(a)}{\Gamma(a) da} - \ln(a) \equiv G(a).$$

Получив оценку параметра  $\hat{a}$ , вычисляем и оценку параметра  $\hat{\lambda}$  по формуле  $\hat{\lambda} = \frac{\hat{a}}{\hat{T}}$ .

## Метод максимального правдоподобия

для точечной оценки неизвестных параметров заданного распределения

Для распределения **Вейбулла** с неизвестных параметров  $a$  и  $b$  для **полной выборки**

$$F(t) = 1 - \exp\left(-\left(\frac{t}{b}\right)^a\right), \quad f(t) = \frac{a}{b}\left(\frac{t}{b}\right)^{a-1} \cdot \exp\left(-\left(\frac{t}{b}\right)^a\right), \quad \lambda(t) = \frac{a}{b} \cdot \left(\frac{t}{b}\right)^{a-1},$$

Функция правдоподобия  $L(t_1, \dots, t_N; a, b) = \left(\frac{a}{b}\right)^N \cdot \prod_{i=1}^N \left\{ \left(\frac{t_i}{b}\right)^{a-1} \cdot e^{-\left(\frac{t_i}{b}\right)^a} \right\}$ .

$$\ln(L(t_1, \dots, t_N; a, b)) = N \cdot \ln(a) - N \cdot a \cdot \ln(b) + (a-1) \sum_{i=1}^N \ln(t_i) - b^{-a} \sum_{i=1}^N t_i^a.$$

Необходимое условие экстремума:

$$\begin{cases} \frac{\partial \ln(L(a, b))}{\partial b} \equiv -N \cdot \frac{a}{b} + \frac{a}{b^{a+1}} \cdot \sum_{i=1}^N t_i^a = 0, \\ \frac{\partial \ln(L(a, b))}{\partial a} \equiv \frac{N}{a} - N \cdot \ln(b) + \sum_{i=1}^N \ln(t_i) + b^{-a} \cdot \ln(b) \cdot \sum_{i=1}^N t_i^a - b^{-a} \sum_{i=1}^N [t_i^a \cdot \ln(t_i)] = 0. \end{cases}$$

Из первого уравнения находим выражение для  $\left(\frac{\sum_{i=1}^N t_i^a}{N}\right)^{\frac{1}{a}}$

и, подставляя его во второе уравнение, приходим к уравнению:

$$\frac{N}{a} + \sum_{i=1}^N \ln(t_i) - N \cdot \frac{\sum_{i=1}^N t_i^a \cdot \ln(t_i)}{\sum_{i=1}^N t_i^a} = 0,$$

которое надо разрешить относительно параметра  $a$ . После чего остается вычислить  $\hat{b}$ .

## Метод максимального правдоподобия

для точечной оценки неизвестных параметров заданного распределения

Расчет параметров  $a$  и  $b$  распределения **Вейбулла** для **цензурированной выборки**

Имеем усеченную выборку объемом  $N = m + s$ , содержащую:

- ряд наработок с отказами  $\tau_1, \tau_2, \dots, \tau_i, \dots, \tau_m$
- ряд безотказных наработок  $t_1, t_2, \dots, t_j, \dots, t_s$

Функция правдоподобия  $\ln L = m \ln a - ma \ln b + (a-1) \sum_{i=1}^m \ln(\tau_i) - b^{-a} \left( \sum_{i=1}^m \tau_i^a + \sum_{j=1}^s t_j^a \right)$

Необходимое условие экстремума:

$$\begin{cases} \frac{\partial L}{\partial b} = -m \frac{a}{b} + \frac{a}{b^{a+1}} \left( \sum_{i=1}^m \tau_i^a + \sum_{j=1}^s t_j^a \right) = 0 \\ \frac{\partial L}{\partial a} = \frac{m}{a} - m \ln b + \sum_{i=1}^m \ln(\tau_i) + b^{-a} \ln b \left( \sum_{i=1}^m \tau_i^a + \sum_{j=1}^s t_j^a \right) - b^{-a} \left( \sum_{i=1}^m \tau_i^a \ln \tau_i + \sum_{j=1}^s t_j^a \ln t_j \right) = 0 \end{cases}$$

Из первого уравнения находим выражение для  $\hat{b} = \frac{\left( \sum_{i=1}^m \tau_i^a + \sum_{j=1}^s t_j^a \right)^{1/a}}{m}$

и, подставляя его во второе уравнение, приходим к уравнению:

$$\frac{m}{a} - \sum_{(i)} \ln \tau_i - m \frac{\sum_{i=1}^m \tau_i^a \ln \tau_i + \sum_{j=1}^s t_j^a \ln t_j}{\sum_{i=1}^m \tau_i^a + \sum_{j=1}^s t_j^a} = 0$$

которое надо разрешить относительно параметра  $a$ . После чего остается вычислить  $\hat{b}$ .

## Вычислительные трудности решения сложных алгебраических уравнений при применении ММ и ММП

1. Алгебраическое уравнение для нахождения параметра  $a$  распределения Вейбулла при применении ММ

$$\frac{\bar{D}}{D} - \left( \frac{\bar{T}}{T} \right) / \left( 1 + \frac{1}{a} \right)^2 \cdot \left\{ \Gamma \left( 1 + \frac{2}{a} \right) - \left[ \Gamma \left( 1 + \frac{1}{a} \right) \right]^2 \right\} = 0,$$

2. Алгебраическое уравнение для нахождения параметра  $\hat{a}$  Гамма-распределения методом МП

$$\frac{1}{N} \cdot \sum_{i=1}^N \ln \left( \frac{t_i}{T} \right) - \frac{\frac{d\Gamma(a)}{da}}{\Gamma(a)} + \ln(a) = 0.$$

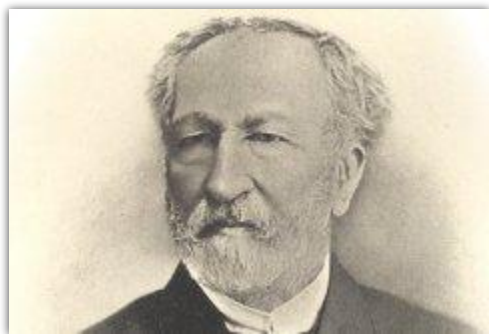
$$\text{Введем обозначение } \frac{\frac{d\Gamma(a)}{da}}{\Gamma(a)} - \ln(a) \equiv G(a).$$

3. Алгебраическое уравнение для нахождения параметра  $\hat{a}$  распределения Вейбулла

$$\frac{N}{a} + \sum_{i=1}^N \ln(t_i) - N \cdot \frac{\sum_{i=1}^N t_i^a \cdot \ln(t_i)}{\sum_{i=1}^N t_i^a} = 0.$$

## Применение разложения Стирлинга

для решения сложных алгебраических уравнений

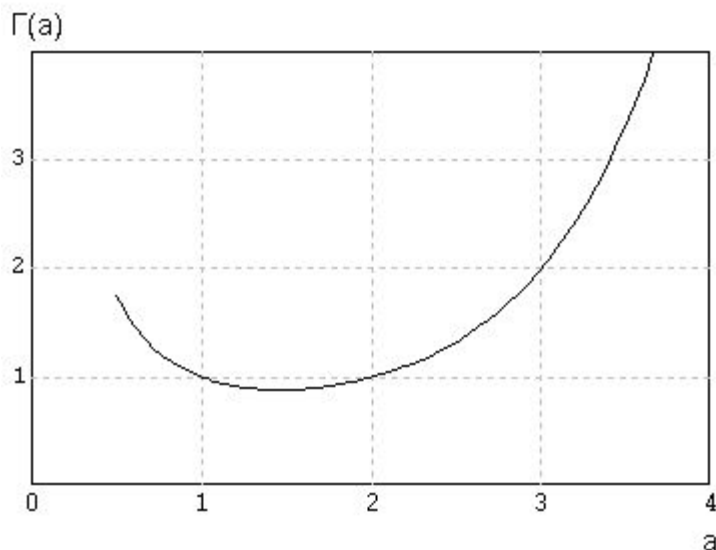


Джеймс Стирлинг (Stirling) (1692 – 1770) - шотландский математик, член Лондонского королевского общества. Впервые дал асимптотическое разложение для гамма-функции и логарифма от нее.

### Разложения Джеймса Стирлинга для Эйлера интеграла 2-го рода

$$\Gamma(a) = e^{-a} \cdot a^{a-\frac{1}{2}} \cdot \sqrt{2\pi} \cdot \left[ 1 + \frac{1}{12a} + \frac{1}{288a^2} - \frac{139}{51840a^3} - \frac{571}{2488320a^4} + O(a^{-5}) \right].$$

График  $\Gamma(a)$



Показана приемлемость формулы Стирлинга в широком диапазоне изменения параметра  $a$

	$a = \frac{1}{2} \quad \Gamma(\frac{1}{2}) = \sqrt{\pi}$	$a = \frac{3}{2} \quad \Gamma(\frac{3}{2}) = \sqrt{\pi} / 2$
Таблица	1,77245	0,88623
Стирлинг	1,75666	0,886155

## Решение нелинейного уравнения для параметра $a$ распределения Вейбулла при использовании метода моментов

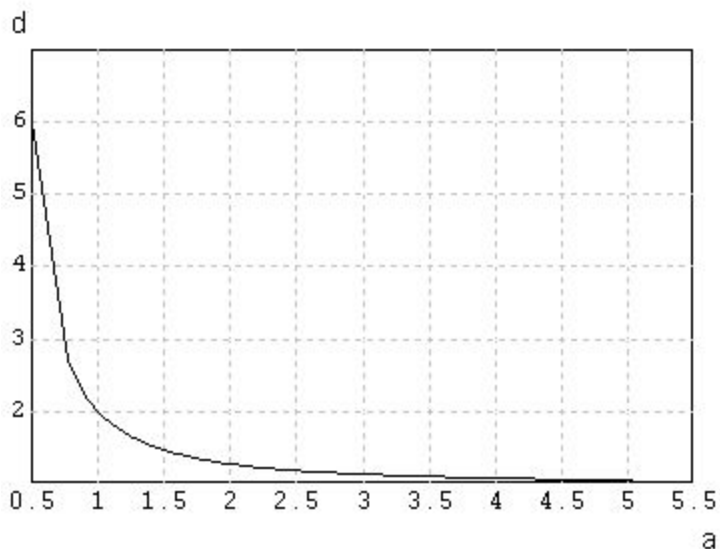
$$\frac{\hat{D}}{\hat{T}^2} + 1 - \frac{\Gamma(1 + \frac{2}{a})}{\Gamma^2(1 + \frac{1}{a})} = 0.$$

Учитывая свойство строгой вогнутости функции  $\frac{\Gamma(1 + \frac{2}{a})}{\Gamma^2(1 + \frac{1}{a})} = 0$

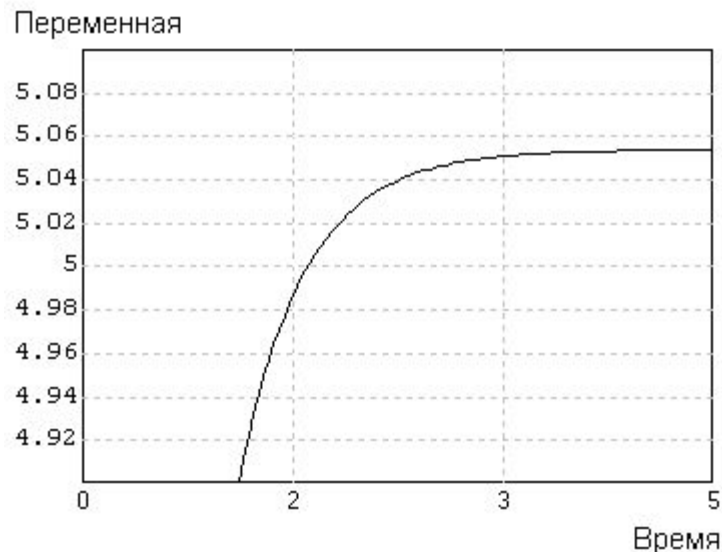
а следовательно единственности решения будем искать решение путем численного интегрирования уравнения (введением «отрицательной обратной

связи»)  $\frac{da}{dt} = \frac{\hat{D}}{\hat{T}^2} + 1 - \frac{\Gamma(1 + \frac{2}{a})}{\Gamma^2(1 + \frac{1}{a})}$  при любом начальном условии.

График  $\Gamma^2/\Gamma^2$  от "а"



Расчет параметра "а"



## Решение нелинейного уравнения для параметра $a$ Гамма-распределения при использовании метода максимального правдоподобия

$$\frac{1}{N} \cdot \sum_{i=1}^N \ln\left(\frac{t_i}{T}\right) - \frac{d\Gamma(a)}{\Gamma(a) da} + \ln(a) = 0. \quad \text{Обозначим } \frac{d\Gamma(a)}{\Gamma(a) da} - \ln(a) \equiv G(a), \quad \text{и}$$

найдем аналитическое выражение для  $G(a)$

$$\Gamma(a) = e^{-a} \cdot a^{a-0.5} \cdot \sqrt{2\pi} \cdot Q(a); \quad Q(a) = \left[ 1 + \frac{1}{12a} + \frac{1}{288a^2} - \frac{139}{51840a^3} - \frac{571}{2488320a^4} + O(a^{-5}) \right].$$

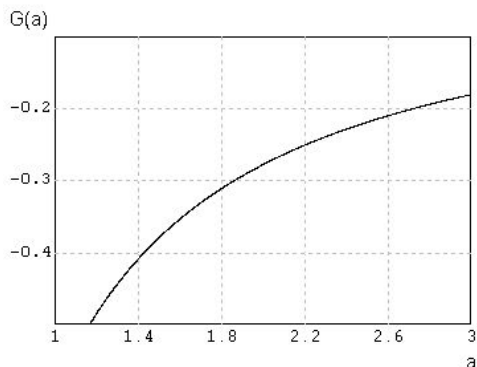
$$\ln(\Gamma(a)) = \ln(e^{-a}) + \ln(a^{a-0.5}) + \ln(\sqrt{2\pi}) + \ln(Q(a)). \quad \ln(a^{a-0.5}) = (a-0.5) \cdot \ln(a).$$

$$\frac{d(\ln(\Gamma(a)))}{da} = -1 + \ln(a) + (a-0.5)/a + (-1/(12 \cdot a^2) - 2/(288 \cdot a^3) + 139 \cdot 3/(51840 \cdot a^4) + 571 \cdot 4/(2488320 \cdot a^5))/Q(a).$$

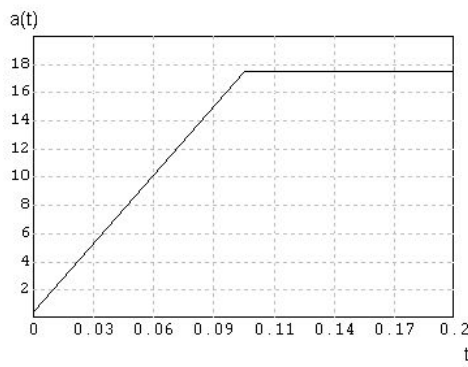
В итоге получим окончательно аналитическое выражение

$$G(a) = -1 + (a-0.5)/a + (-1/(12 \cdot a^2) - 2/(288 \cdot a^3) + 139 \cdot 3/(51840 \cdot a^4) + 571 \cdot 4/(2488320 \cdot a^5))/Q(a).$$

График  $G(a)$



Расчет параметра  $a(t)$  при  $t \rightarrow \infty$



Тогда алгоритм решения уравнения может быть принят в виде дифференциального уравнения  $a' = -K \cdot (0.0288041 + G(a))$  или  $a' = -K \cdot \text{sign}(0.0288041 + G(a))$  при произвольном начальном значении  $a = 0.5$ , решение которого асимптотически приведет к искомому результату.



## Решение нелинейного уравнения для параметра $a$ распределения Вейбулла при использовании метода максимального правдоподобия

Для решения уравнения предлагается использовать ту же идею для нахождения  $a$  - решать следующее дифференциальное уравнение вида

$$\frac{N}{a} + \sum_{i=1}^N \ln(t_i) - N \cdot \frac{\sum_{i=1}^N t_i^a \cdot \ln(t_i)}{\sum_{i=1}^N t_i^a} = 0.$$

$$\frac{da}{dt} = + \left[ \frac{N}{a} + \sum_{i=1}^N \ln(t_i) - N \cdot \frac{\sum_{i=1}^N t_i^a \cdot \ln(t_i)}{\sum_{i=1}^N t_i^a} \right],$$

с различными начальными условиями, подтверждающих единственность «корня»  $a(t)$ . Тогда даёт искомое значение оценки

параметра  $\hat{a}$ . После чего остается вычислить  $\hat{b}$  по формуле  $\left( \frac{\sum_{i=1}^N t_i^{\hat{a}}}{N} \right)^{\frac{1}{\hat{a}}}$ .

График "а"

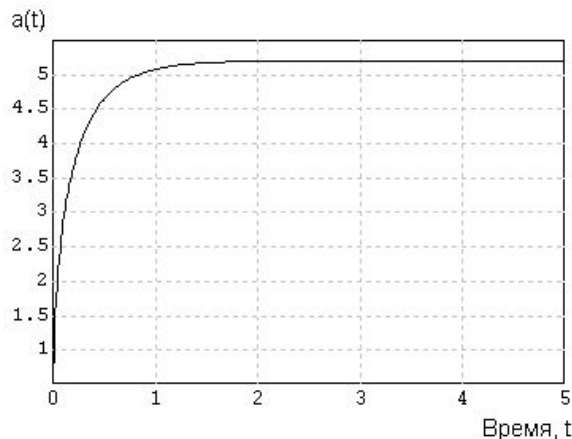
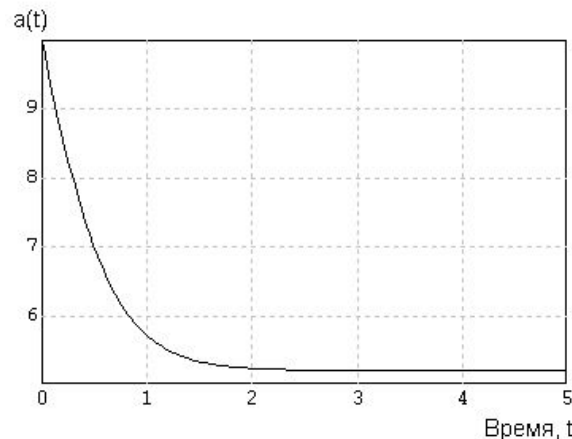
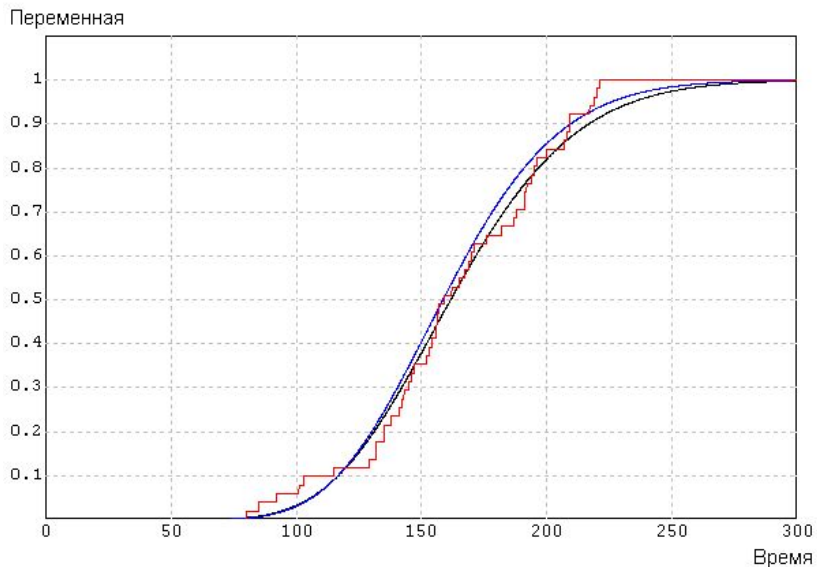


График "а"



# Решение тестовой задачи

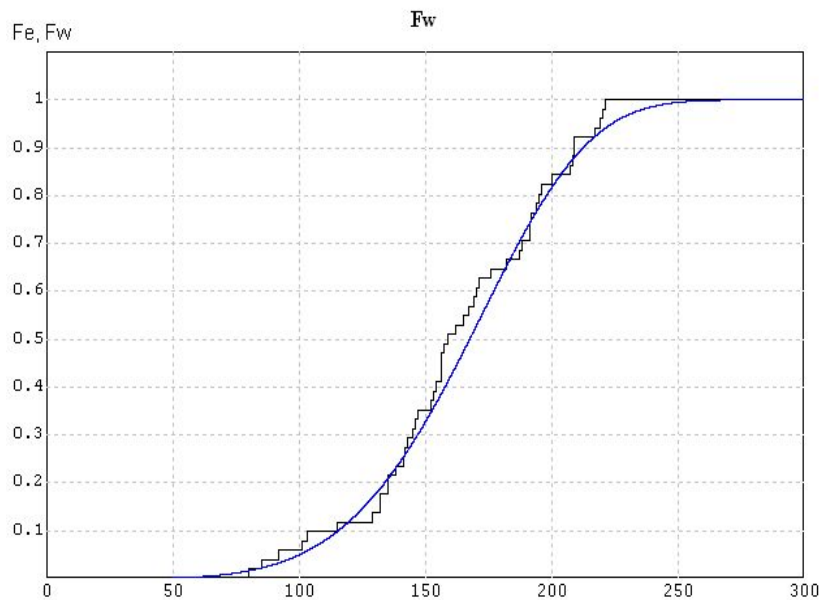
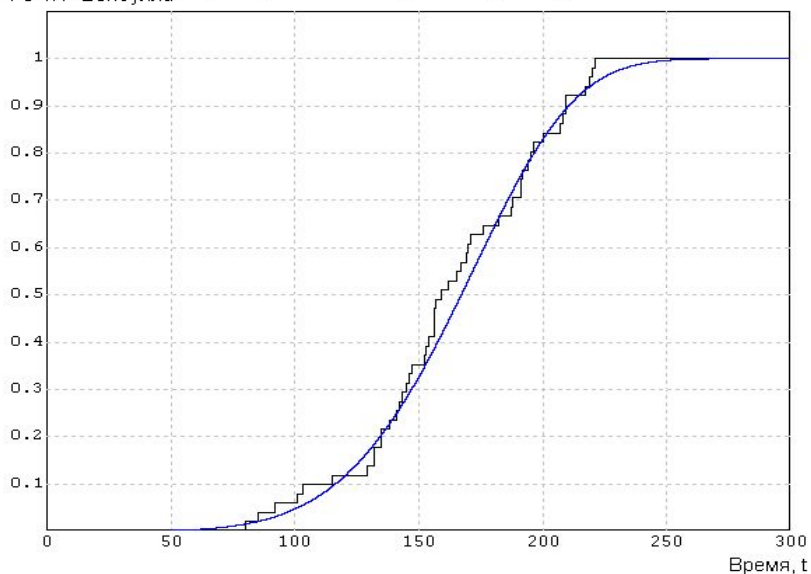
Графики  $F_e$ ,  $F_{MMП}$  - черный,  $F_{MM}$  - синий для Гамма-распределения



Распределение Вейбулла		Гамма-распределение	
MM	MMП	MM	MMП
$a = 5.054$	$a = 5.2061$	$a = 19,294$	$a = 17,52$
$b = 179.308$	$b = 179.232$	$\lambda = 0,121433$	$\lambda = 0.106401$

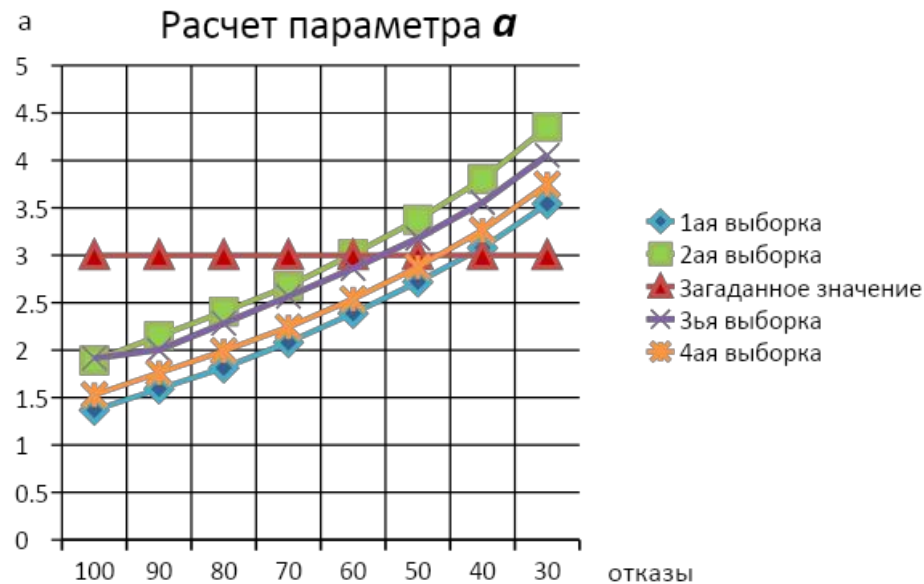
Графики эмпирического  $F_e$  распределения и аппроксимирующего его распределения Вейбулла

Графики эмпирического и аппроксимирующего распределения Вейбулла с параметрами по методу ПРАВДОПОДОБИЯ  $F_e$  и  $F$ -Вейбулла



Решение нелинейного уравнения  
для параметра  $a$  распределения Вейбулла  
для **цензурированной** выборки  
при использовании метода **максимального правдоподобия**

Имеем четыре выборки по 100 значений.



Можно утверждать, что примерно 50% отказов уже дают представление о том, что выборка принадлежит распределению Вейбулла.

## Результаты работы

1. Разработано алгоритмическое и программное обеспечение для получения точечных оценок параметров двух практически важных распределений - распределения Вейбулла и гамма-распределения методом моментов, методом максимального правдоподобия и методом вероятностных сеток («вероятностной бумаги»). Разработан и программно реализован также алгоритм проверки выполнимости критерия А.Н. Колмогорова – критерия согласия аппроксимирующего распределения эмпирическому (ступенчатому) распределению.
2. Системы расчетов полностью реализованы в среде отечественного Программного комплекса «МВТУ 3.5».
3. Показано, что применение предложенных алгоритмов позволило исключить необходимость в «ручном» проведении расчетов показателей надежности и в обращении к многочисленным числовым статистическим таблицам. Все необходимые данные получаются расчетным путем с использованием соответствующих моделей и формул. По информации о первых выборочных четырех начальных и центральных моментах и об эмпирической функции распределения расчетным путем автоматически находятся параметры практически всех типовых распределений, а также значения гамма-функции (эйлерова интеграла второго рода).

## Вывод

В работе сделан научно-технический задел для решения на ЭВМ перспективных задач обработки информации с малыми или ограниченными по объему выборками, в том числе с различными типами цензурирования, то есть усеченными выборками, методами моментов и максимального правдоподобия, как с точечными, так и интервальными оценками параметров предполагаемых распределений.

# Критерии согласия

- Критерий согласия

Колмогорова

$$\lambda = D \cdot \sqrt{N} = \max_t |F_{эм}(t) - F(t)| \cdot \sqrt{N}$$

$$\text{Вер}\{D\sqrt{N} > x_\alpha\} = 1 - \sum_{k=-\infty}^{k=+\infty} (-1)^k \cdot \exp(-2k^2 x_\alpha^2) = \alpha$$

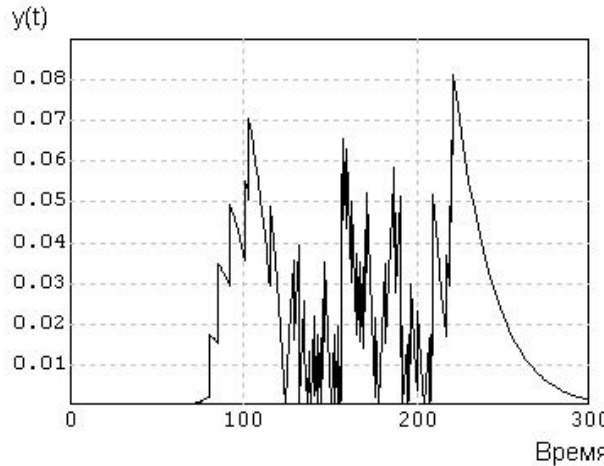
- Критерий “Стареющего” распределения

$$m_k = \frac{1}{N} \sum_{i=1}^{i=N} t_i^k \quad k = 1, 2, 3, 4.$$

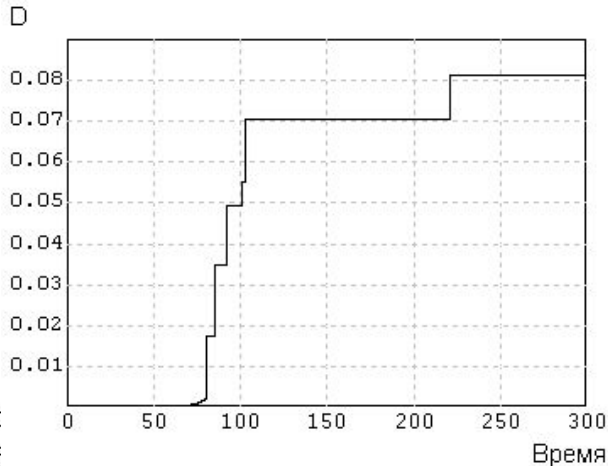
$$M_{k+1} \cdot M_{k-1} \leq M_k^2 \quad M_k = \frac{m_k}{k!}$$

## Критерий согласия Колмогорова

График  $y(t)=\text{abs}(F_e-F)$



Расчет статистики Колмогорова  $D = \sup(y(t))$  по t



$$\lambda = D \cdot \sqrt{N} = \max_t |F_{\text{эм}}(t) - F(t)| \cdot \sqrt{N}$$

$$\text{Вер} \{ D \sqrt{N} > x_\alpha \} = 1 - \sum_{k=-\infty}^{k=+\infty} (-1)^k \cdot \exp(-2k^2 x_\alpha^2) = \alpha$$

a	0,20	0,10	0,05	0,02	0,01	0,001
x	1,073	1,224	1,358	1,520	1,625	1,950

*Очень спасибо  
за внимание !*

График плотности вероятности гамма-распределения с параметрами, найденными методом моментов

