

Применение метода минимизации функции риска с учетом усеченных слева и цензурированных справа данных для оценки параметров распределения Вейбулла

Using the Bayes estimator for Weibull parameters estimation taking into account left-truncated and right-censored data

Николаев Д.А.^{1*}, Антонов А.В.², Чепурко В.А.³

Dmitry A. Nikolaev^{1*}, Alexander V. Antonov², Valery A. Chepurko³

¹АО «Росатом Автоматизированные системы управления», Москва Российская Федерация, ²АНО ДПО «Техническая академия Росатома», Обнинск, Российская Федерация, ³АО «Росатом Автоматизированные системы управления», Москва, Российская Федерация

¹Rusatom Automated Control Systems, Moscow, Russian Federation, ²Rosatom Technical Academy, Obninsk, Russian Federation, ³Rusatom Automated Control Systems, Moscow, Russian Federation

*dafanday@gmail.com



Николаев Д.А.



Антонов А.В.



Чепурко В.А.

Резюме. В настоящее время при решении вопросов организации эксплуатации промышленных установок предъявляются повышенные требования к безопасности, надежности и эффективности их функционирования. В основе методов и процедур, используемых для решения задач повышения безопасности и надежности работы промышленных установок, лежит информация о надежности комплектующих элементов, систем и оборудования. Для определения объективных характеристик надежности указанных объектов организуют наблюдения за их поведением в процессе эксплуатации. В ходе наблюдения за функционированием объектов отмечают периоды времени непрерывного исправного функционирования, времена простоев, причины простоев, отказы, дефекты и неисправности объектов, периодичность проведения и глубину профилактического обслуживания элементов и систем, и прочую информацию. Следует заметить, что элементы и системы современных промышленных объектов, таких как атомные электростанции, нефтехимические комплексы и т.п., относятся к категории высоконадежного оборудования. Отказы такого оборудования – события редкие. Количество однотипных объектов крайне мало.

Цель. В связи с этим возникает задача разработки методов достоверного оценивания характеристик надежности объектов на основании статистической информации ограниченного объема. **Метод.** В статье рассматривается метод расчета показателей надежности объектов на основании статистической информации, полученной на этапе эксплуатации – метод минимизации функции риска с учетом усеченных слева и цензурированных справа данных для оценки параметров распределения Вейбулла. **Выводы.** В качестве примера приведен метод оценки показателей надежности на основании полных, цензурированных справа и усеченных слева наработок, т.к. на практике такая комбинация встречается довольно часто. Приведен вид функций правдоподобия для распределения Вейбулла. Рассмотрен тестовый пример, в котором с помощью метода минимизации функции риска находятся оценки параметров распределения Вейбулла тестовой выборки, содержащей полные, усеченные слева и цензурированные справа данные. Исследовано изменение значений оценок параметров распределения Вейбулла и их точности в зависимости от доли усеченных и от доли цензурированных данных.

Abstract. Today, when operation of commercial plants is organised, they are expected to comply to constantly increasing requirements for safety, dependability and efficiency of operation. The methods and procedures that are employed for the purpose of improving the safety and dependability of commercial plants are based on the information on the dependability of components, systems and equipment. In order to identify the objective dependability characteristics of such facilities, their behaviour in operation is monitored. In the course of facility operation monitoring, periods of continuous fault-free operation, periods of downtime, causes of downtime, failures, defects and malfunctions of items, frequency and depth of preventive maintenance of elements and systems, as well as other information are recorded. It should be noted that elements and systems of today's industrial facilities, such as nuclear power plants, petrochemical complexes, etc., are classified as highly dependable equipment. Failures of such equipment are rare. The number of same-type facilities is extremely small.

Aim. Given the above, the problem arises of developing methods for reliable estimation if item dependability characteristics on the basis of limited statistical information. **Method.** The paper

examines a method for calculating facility dependability indicators on the basis of statistical information obtained in operation, i.e. a method for minimising the risk function while taking into account left-truncated and right-censored data for the purpose of Weibull distribution parameter estimation. **Conclusions.** By way of example, the authors refer to a method for evaluating dependability indicators based on complete, right-censored and left-truncated operation times, as, in practice, such combination is quite common. The form of likelihood functions for the Weibull distribution is given. A test case is examined, whereas, using the risk function minimisation, estimates of the Weibull distribution parameters are obtained for a sample that contains full, left-truncated and right-censored data. The authors examined the variation in the Weibull distribution values and their accuracy depending on the proportion of truncated and censored data.

Ключевые слова: функция правдоподобия, метод минимизации функции риска, моделирование данных, распределение Вейбулла, полные данные, цензурированные данные, усеченные данные.

Keywords: maximum likelihood, Bayes estimator, data modelling, Weibull distribution, complete data, censored data, truncated data.

Для цитирования: Николаев Д.А., Антонов А.В., Чепурко В.А. Применение метода минимизации функции риска с учетом усеченных слева и цензурированных справа данных для оценки параметров распределения Вейбулла // Надежность. 2022. №3. С. 53-61. <https://doi.org/10.21683/1729-2646-2022-22-3-53-61>

For citation: Nikolaev D.A., Antonov A.V., Chepurko V.A. Using the Bayes estimator for Weibull parameters estimation taking into account left-truncated and right-censored data. Dependability 2022;3: 53-61. <https://doi.org/10.21683/1729-2646-2022-22-3-53-61>

Поступила 01.04.2022 г. / **После доработки** 14.08.2022 / **К печати** 19.09.2022 г.

Received on: 01.04.2022 / **Revised on:** 14.08.2022 / **For printing:** 19.09.2022.

Введение

Эксплуатация оборудования представляет собой наиболее информативный вид испытаний на надежность, так как при этом оборудование подвергается такому широкому спектру внешних воздействий, режимов работы и обслуживания, который невозможно воссоздать в условиях искусственного эксперимента. Негативной стороной эксплуатационных наблюдений является малый объем статистических данных и необходимость наблюдения за функционированием объектов длительное время.

В последнее время развиваются методы оценки характеристик надежности, которые используют не только информацию о наработках объектов до отказа, но и весь объем имеющейся информации о функционировании объектов. К таким данным относится информация о периодах работы объектов, не завершившихся отказами, информация о периодах с различного рода неопределенностями.

Постановка задачи

Рассмотрим более подробно виды информации, на основании которой выполняют оценку характеристик надежности. Итак, будем полагать, что под наблюдением находится N однотипных объектов. Если в ходе наблюдения за функционированием i -того объекта наблюдается отказ с известной наработкой, то будем иметь наработку объекта до отказа ξ_i . Для вычисления наработки до отка-

за объекта необходимо знать время начала функционирования объекта (установки в систему) и время его отказа. Следующий тип наработки, который будем учитывать при расчетах – цензурированные справа наработки. Цензурированные справа наработки характеризуются тем, что функционирование объектов наблюдения успешно продолжается на временном интервале после завершения наблюдений. В этом случае исследователь имеет информацию о начале эксплуатации объекта (время установки объекта в систему) и длительность периода наблюдения. За время наблюдения объект не отказал и продолжает успешно функционировать. Цензурированную справа наработку обозначим Y_j .

И, наконец, третий тип данных, которые имеют место при сборе информации о функционировании объектов, – пропущенные данные. Механизм возникновения пропущенных данных характеризуется следующими обстоятельствами. Например, известно, что систему ввели в эксплуатацию в некоторый момент времени, а наблюдение и фиксирование информации о функционировании оборудования началось позже. Данные о работе объектов анализа за период с момента ввода объекта в эксплуатацию до момента начала фиксирования информации отсутствуют, известно лишь то, что оборудование эксплуатировалось. Механизм формирования данных с пропусками называется усечением.

Задача, таким образом, состоит в разработке метода оценки надежности объектов с учетом наработок до отказа, цензурированных и усеченных данных.

Представление полных, цензурированных справа и усеченных слева наработок

В некоторый момент времени в эксплуатацию вводится система, в составе которой работают элементы, являющиеся объектами исследования. Момент начала работы каждого элемента известен. Обозначим его μ_i . Наблюдения за работой элементов и фиксация отказов началась не с нулевого момента времени, а с некоторого более позднего момента T_0 и продолжается до некоторого момента времени T_k (рис. 1).

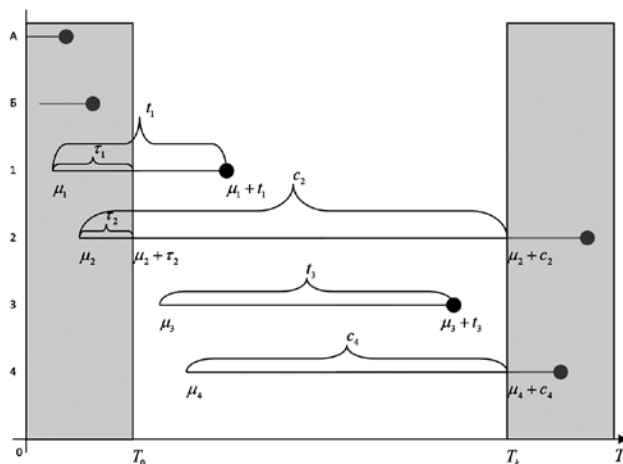


Рис. 1. Возможные виды наработок до отказа

T_k – это момент времени завершения наблюдений за функционированием объектов.

Рассмотрим возможные ситуации поведения объектов во время их функционирования. Ситуации отражены на рис. 1.

А. Объект введен в эксплуатацию вместе с системой и отказал до начала организации наблюдений за функционированием.

Б. Объект введен в эксплуатацию в некоторый момент времени после начала функционирования системы, и отказал до начала организации наблюдений за функционированием.

Для этих двух случаев времена работы объектов являются «потерянными». Они должны были бы вносить определенный вклад в функцию правдоподобия, однако, их не наблюдали, поскольку фиксация истории функционирования объектов началась после их выхода из строя. Часть относительно малых наработок выпала из статистической выборки и этот факт необходимо учесть при построении функции правдоподобия. Игнорирование факта потери части данных может привести к некорректной, смещенной оценке параметров закона распределения.

1. Объект введен в эксплуатацию в некоторый момент времени после начала функционирования системы – μ_1 (момент начала его эксплуатации), но до начала организации наблюдений за функционированием. Зная время установки объекта в систему можно найти промежуток времени, который система проработала до начала фиксации отказов t_1 . В

момент времени $\mu_1 + t_1$ система отказала. В данном случае имеем усеченную слева наработку до отказа t_1 .

Такое определение отличает наработки данного типа от остальных и вот почему. С одной стороны, казалось бы, это полная наработка, но она реализовалась с условием, что эта полная наработка объекта «пережила» время t_1 . То есть имеется дополнительное условие $t > t_1$, которое необходимо учесть в функции правдоподобия. Учесть это возможно путем введения дополнительного множителя – условной плотностью распределения при условии $t > t_1$: $\frac{f(\theta; t)}{1 - F(\theta; t_1)}$. Такой знаменатель частично компенсирует потерянные данные, упомянутые выше.

2. Объект введен в эксплуатацию в некоторый момент времени после начала организации наблюдений за функционированием – μ_2 , но до начала организации наблюдений за функционированием. На момент завершения наблюдений объект продолжает функционировать, то есть, имеем цензурированную справа наработку. Из соотношения $\mu_2 + \tau_2 = T_0$ можно найти τ_2 – промежуток времени, который система проработала к моменту начала наблюдений. Поскольку система продолжает работать, то c_2 является усеченной слева цензурированной справа наработкой.

Учет такого рода наработок приведет к появлению в функции правдоподобия множителей вида: $\frac{1 - F(\theta; c)}{1 - F(\theta; t)}$ – условной вероятности того, что наработка больше времени c , при условии, что она больше времени t . В совокупности с упомянутыми выше усеченными слева полными наработками этот тип наработок полностью компенсирует потерю части данных при получении оценки.

3. Объект введен в эксплуатацию в некоторый момент времени после начала организации наблюдений за функционированием – μ_3 и в момент времени $\mu_3 + t_3$ объект отказал. Таким образом, имеем полную наработку до отказа t_3 .

4. Объект введен в эксплуатацию в некоторый момент времени после начала организации наблюдений за функционированием – μ_4 и проработал без отказа до момента завершения наблюдений T_k . На момент завершения наблюдений объект продолжает функционировать, то есть, имеем цензурированную справа наработку. Нарботка объекта равна c_4 – цензурированная справа наработка.

В функции правдоподобия учет полных и цензурированных справа наработок приведет к появлению классических множителей $f(\theta; t)$ и $[1 - F(\theta; c)]$ соответственно.

Применение метода минимизации функции риска для определения параметров по зафиксированной информации

Введем ряд предположений.

Пусть под наблюдением находится n объектов. Для каждого объекта имеются данные о его времени работы

такого типа, как это показано на рис. 1. Т.е. зафиксирована выборка $t_1, \dots, t_n$, где t_i – наработки до отказа, $c_1, \dots, c_n$, где c_i – цензурированные справа данные, и $\tau_1^L, \dots, \tau_n^L$, где τ_i^L – усеченные слева данные. В этом случае совместная плотность будет выражаться согласно формуле [3, 4, 5]:

$$L(\theta | DATA) = \prod_{i=1}^n f(\theta; t_i)^{\delta_i v_i} \left[\frac{f(\theta; t_i)}{1 - F(\theta; \tau_i^L)} \right]^{\delta_i (1 - v_i)} \times \\ \times [1 - F(\theta; c_i)]^{(1 - \delta_i) v_i} \left[\frac{1 - F(\theta; c_i)}{1 - F(\theta; \tau_i^L)} \right]^{(1 - \delta_i) (1 - v_i)} \quad (1)$$

где θ – вектор параметров распределения; v_i – индикатор пропущенных данных, который принимает значение 0, если для данного i пропуск имел место, и 1 в противном случае; δ_i – индикатор цензурированных данных, равен 0, если i -е наблюдение цензурировано, и 1 в противном случае, т.е. в i -м наблюдении зафиксирован отказ.

Параметр $\theta \in \Theta$ – параметр или вектор параметров распределения, в общем случае имеющий априорную плотность распределения $h(\theta)$.

Оценка d параметра θ задана на множестве возможных решений D .

Функция потерь $u(\theta, d)$ определена на (Θ, D) и выражает потери, обусловленные ошибочным решением.

В общем виде функция потерь выглядит следующим образом:

$$u(\theta, d) = \lambda(\theta) W(|d - \theta|),$$

где $W(0) = 0$; $W(x)$ – монотонно возрастающая функция, $x > 0$; $\lambda(\theta)$ – положительно определенная конечная функция.

Известна плотность распределения наблюдаемой величины, которая имеет вид $f(\theta | DATA)$, где θ – вектор параметров.

Классическим методом решения задачи определения вектора параметров закона распределения является метод максимального правдоподобия. При решении задачи данным методом оптимизируется функция правдоподобия (1).

В данном изложении будем решать задачу оценки параметров закона распределения методом минимизации функции риска. Для этой цели рассмотрим формулу Байеса [1, 2]. Запишем выражение для апостериорной плотности распределения параметра при условии, что в результате проведения опыта реализовались данные о наработках в виде:

$$h_{аност}(\theta | \{DATA\}) = \frac{h(\theta) L(\theta | DATA)}{\int_0^\infty h(\theta) L(\theta | DATA) d\theta},$$

где $h(\theta)$ – априорная плотность распределения оцениваемого параметра. В случае отсутствия априорной информации, в качестве априорной плотности распределения выбирают несобственную плотность распределения из класса сопряженных к функции распределения наработок плотностей. Тогда апостериорная плотность распределения преобразуется к виду

$$h_{аност}(\theta | \{DATA\}) = \frac{L(\theta | DATA)}{\int_0^\infty L(\theta | DATA) d\theta}.$$

Для определения оценки d параметра θ введем функцию риска:

$$R(\theta, d) = \int_\Theta \lambda(\theta) W(|d - \theta|) h(\theta | \{DATA\}) d\theta.$$

При квадратичной функции потерь $W(|d - \theta|) = (d - \theta)^2$ и $\lambda(\theta) = 1$ функция риска имеет вид:

$$R(\theta, d) = \int_\Theta (d - \theta)^2 h(\theta | \{DATA\}) d\theta. \quad (2)$$

Минимизируем данную функцию риска, определяем оценку параметра θ . Возьмем производную от функции (2) и приравняем ее нулю:

$$d(DATA) \int_\Theta h(\theta | \{DATA\}) d\theta - \int_\Theta \theta h(\theta | \{DATA\}) d\theta = 0.$$

Значение $\int_\Theta h(\theta | \{DATA\}) d\theta = 1$, т.к. представляет собой интеграл от плотности распределения по всей области определения параметра θ . Получим оценку параметра

$$d(DATA) = \int_\Theta \theta h(\theta | \{DATA\}) d\theta = \frac{\int_0^\infty \theta L(\theta | DATA) d\theta}{\int_0^\infty L(\theta | DATA) d\theta}.$$

Выражение для дисперсии оцениваемого параметра имеет вид:

$$\sigma_0^2 = \int_\Theta \theta^2 h(\theta | \{DATA\}) d\theta - d^2(DATA) = \\ = \frac{\int_0^\infty \theta^2 L(\theta | DATA) d\theta}{\int_0^\infty L(\theta | DATA) d\theta} - d^2(DATA).$$

Применим изложенный подход к задаче оценки параметров закона распределения Вейбулла, с соответствующей функцией распределения:

$$F(t; \theta_1; \theta_2) = 1 - \exp\left(-\left(\frac{t}{\theta_1}\right)^{\theta_2}\right), \quad \theta_1 = \alpha, \quad \theta_2 = \eta.$$

Запишем для распределения Вейбулла функцию правдоподобия:

$$L((\alpha; \eta) | DATA) = \left(\frac{\eta}{\alpha}\right)^{\sum_{i=1}^n \delta_i} \times \\ \times \prod_{i=1}^n \left(\frac{t_i}{\alpha}\right)^{(\eta-1)\delta_i} \left(e^{-\left(\frac{t_i}{\alpha}\right)^\eta}\right)^{\delta_i v_i} \left(e^{-\left(\frac{t_i}{\alpha}\right)^\eta + \left(\frac{\tau_i^L}{\alpha}\right)^\eta}\right)^{\delta_i (1-v_i)} \times \\ \times \left(e^{-\left(\frac{c_i}{\alpha}\right)^\eta}\right)^{(1-\delta_i)v_i} \left(e^{-\left(\frac{c_i}{\alpha}\right)^\eta + \left(\frac{\tau_i^L}{\alpha}\right)^\eta}\right)^{(1-\delta_i)(1-v_i)}.$$

Оценка параметра α распределения Вейбулла:

$$d(DATA) = \frac{\int_0^\infty \alpha L((\alpha; \eta) | DATA) d\alpha}{\int_0^\infty L((\alpha; \eta) | DATA) d\alpha}. \quad (3)$$

Дисперсии оцениваемого параметра α имеет вид:

$$\sigma_{\alpha}^2 = \frac{\int_0^{\infty} \alpha^2 L((\alpha; \eta) | DATA) d\alpha}{\int_0^{\infty} L((\alpha; \eta) | DATA) d\alpha} - d^2(DATA). \quad (4)$$

Для второго параметра η распределения Вейбулла оценка и дисперсия могут быть найдены аналогичным образом. В данном исследовании параметр η предполагается известным и оцениванию не подлежит.

Тестовый пример

Пусть набор данных, используемый в исследовании, содержит 400 наблюдений. Эти данные накладываются на модель, описанную в разделе «Представление полных, цензурированных справа и усеченных слева наработок».

Для моделирования описанных ранее ситуаций выбирался регулярный массив моментов начала работы i -го объекта $\mu_i = 10 \cdot i$, $i = 1, 2, \dots, n$; $n = 400$ (см. рис. 1). Время жизни каждого объекта моделировалось с помощью распределения Вейбулла с параметрами $\alpha = 2000$ и $\eta = 2$. В этом случае математическое ожидание соответствующей наработки будет равно

$$\mu_i + \alpha \Gamma(3/2) = 10 \cdot i + \alpha \sqrt{\pi} / 2.$$

Таким образом, можно полагать, что отказ всех объектов произойдет в среднем в момент времени равный

$$\frac{1}{n} \sum_{i=1}^n (\mu_i + \alpha \Gamma(3/2)) = 5(n+1) + \alpha \sqrt{\pi} / 2 \approx 3777 \text{ ч.}$$

Для получения различных видов наработок до отказа положим момент начала наблюдения равным $T_0 = 0$ ч, 200 ч, 400 ч, 600 ч, 800 ч, 1000 ч, 2000 ч, а момент окончания $T_k = 100\,000$ ч (данное число выбрано экспертно, т.к. с большой вероятностью при таком значении все наблюдаемые объекты уже откажут), 5000 ч, 3000 ч, 2500 ч.

Следовательно, чем больше T_0 , тем больше доля усеченных слева данных, а чем меньше T_k , тем больше доля цензурированных справа данных.

Оценка параметра α распределения Вейбулла и ее дисперсия находятся путем решения уравнений (3) и (4) соответственно. Решение данных уравнений выполняется с помощью методов численного интегрирования [1]. Параметр формы η не оценивается, он фиксирован и равняется 2.

Далее в работе для удобства построения графиков рассматривается обратная величина параметра масштаба $\lambda = \frac{1}{\alpha} = 5 \cdot 10^{-4} \frac{1}{\text{ч}}$, называемая интенсивностью отказов.

Графики получены с помощью ядерной оценки плотности распределения [1]. Ядерная оценка плотности — это непараметрический, т.е. свободный от предположений о законе распределения, способ оценки плотности распределения случайной величины. Данная оценка позволяет визуализировать распределение данных на непрерывном интервале или периоде времени. Ядерная

оценка плотности, по сути, является сглаживающей оценкой, когда делается заключение о плотности распределения случайной величины на основе выборки конечного объема. Ядерная оценка плотности тесно связана с гистограммой, но, при этом, она является гладкой и непрерывной. Это осуществляется за счет выбора соответствующей функции ядра. Полученные в работе графики представляют собой разновидность оценки плотности распределения, в которой в качестве ядра применялась Гауссовская плотность распределения. За счет этой функции обеспечивается непрерывность оценки.

В табл. 1–4 приведены результаты исследований, полученные методом максимального правдоподобия (МП) и методом минимизации функции риска (МФР), в зависимости от количества усеченных данных в выборке для определенного момента времени T_k (цензурированные данные справа).

Каждая таблица имеет 8 столбцов:

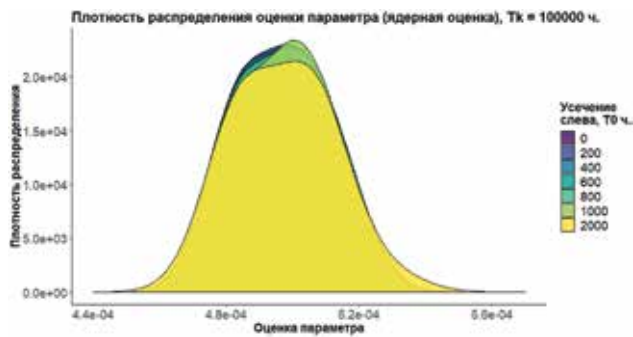
1. момент начала наблюдения;
2. оценка параметра (интенсивность отказа) полученная методом максимального правдоподобия, среднее значение для текущей выборки;
3. дисперсия оцениваемого параметра полученная методом максимального правдоподобия, среднее значение для текущей выборки;
4. оценка параметра (интенсивность отказа) полученная методом минимизации функции риска, среднее значение для текущей выборки;
5. дисперсия оцениваемого параметра полученная методом минимизации функции риска, среднее значение для текущей выборки;
6. объем полной выборки, для которого производилось моделирование и вычисления;
7. количество объектов, которое не попало в анализ по причине усечения (см. рис. 1, случаи А и Б);
8. количество объектов, которое не попало в анализ по причине цензурирования (т.е. данные объекты еще не были введены в эксплуатацию).

Рассмотрим случай, когда наблюдение за объектами начинается в $T_0 = 600$ ч., а заканчивают в момент времени $T_k = 3000$ ч. (пример из табл. 3). Получаем, что в данный временной интервал у нас под наблюдение попадают 295 объектов (столбец 6). Получается, что наблюдаемых объектов стало меньше, т.к. два объекта не попало в анализ по причине усечения (столбец 7), а 103 — по причине цензурирования (столбец 8). В столбцах с 2 по 3 для метода МП и с 4 по 5 для МФР для выборки из 295 объектов показаны результаты оценки параметра и ее дисперсии.

На рис. 2–9 более наглядно проиллюстрированы результаты, приведенные в табл. 1 – 4, а именно, приведены значения оценки и точность оценок обратного параметра масштаба распределения Вейбулла методом минимизации функции риска в зависимости от момента начала наблюдения T_0 для конкретного времени прекращения наблюдения T_k .

Табл. 1. Результаты оценки параметра интенсивности отказа распределения Вейбулла в зависимости от момента начала наблюдения T_0 . $T_k = 100000$ ч.

T_0 , ч.	МП		МФР		Объем полной выборки	Полностью отсутствуют по причине усеечения, шт.	Полностью отсутствуют по причине цензурирования, шт.
	$\lambda \cdot 10^{-4}$	$\sigma \cdot 10^{-4}$	$\lambda \cdot 10^{-4}$	$\sigma \cdot 10^{-4}$			
1	2	3	4	5	6	7	8
0	4,972214	0,124305	4,967547	0,124221	400	0	0
200	5,071529	0,126801	5,066767	0,126717	400	0	0
400	5,072779	0,126919	5,068009	0,126837	399	1	0
600	5,073737	0,127107	5,068959	0,127023	398	2	0
800	5,074037	0,127468	5,069237	0,127376	396	4	0
1000	5,075058	0,128051	5,070214	0,127951	393	7	0
2000	5,078850	0,135530	5,073414	0,135485	350	50	0

Рис. 2. Результаты оценки параметра интенсивности отказа распределения Вейбулла методом МФР в зависимости от момента начала наблюдения T_0 . $T_k = 100000$ ч.Рис. 3. Результаты оценки точности параметра интенсивности отказа распределения Вейбулла методом МФР в зависимости от момента начала наблюдения T_0 . $T_k = 100000$ ч.Табл. 2. Результаты оценки параметра интенсивности отказа распределения Вейбулла в зависимости от момента начала наблюдения T_0 . $T_k = 5000$ ч.

T_0 , ч.	МП		МФР		Объем полной выборки	Полностью отсутствуют по причине усеечения, шт.	Полностью отсутствуют по причине цензурирования, шт.
	$\lambda \cdot 10^{-4}$	$\sigma \cdot 10^{-4}$	$\lambda \cdot 10^{-4}$	$\sigma \cdot 10^{-4}$			
1	2	3	4	5	6	7	8
0	5,053371	0,142987	5,047325	0,142952	399	1	0
200	4,973095	0,143150	4,967035	0,143115	400	0	0
400	4,982129	0,143252	4,976060	0,143217	399	1	0
600	5,073287	0,143521	5,067196	0,143485	398	2	0
800	5,043619	0,144039	5,037483	0,144004	396	4	0
1000	5,064900	0,144882	5,058694	0,144847	393	7	0
2000	4,959545	0,155993	4,952352	0,155955	350	50	0

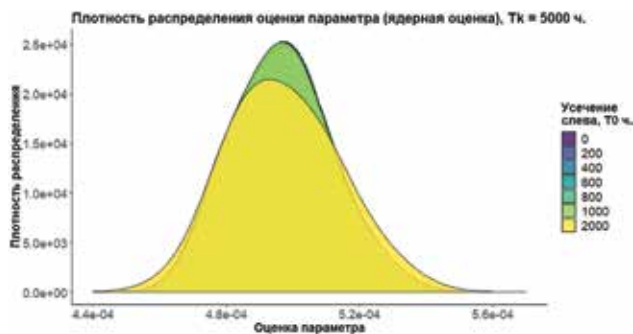
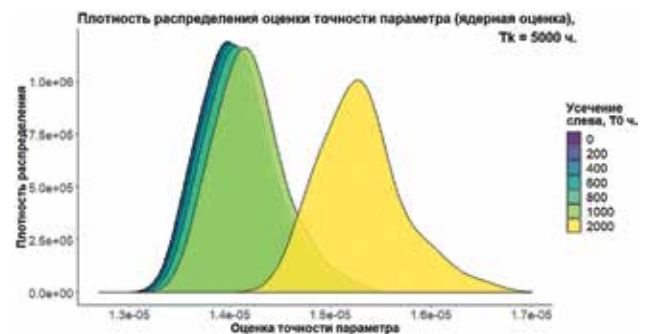
Рис. 4. Результаты оценки параметра интенсивности отказа распределения Вейбулла методом МФР в зависимости от момента начала наблюдения T_0 . $T_k = 5000$ ч.Рис. 5. Результаты оценки точности параметра интенсивности отказа распределения Вейбулла методом МФР в зависимости от момента начала наблюдения T_0 . $T_k = 5000$ ч.

Табл. 3. Результаты оценки параметра интенсивности отказа распределения Вейбулла в зависимости от момента начала наблюдения T_0 . $T_k = 3000$ ч.

T_0 , ч.	МП		МФР		Объем полной выборки	Полностью отсутствуют по причине усечения, шт.	Полностью отсутствуют по причине цензурирования, шт.
	$\lambda \cdot 10^{-4}$	$\sigma \cdot 10^{-4}$	$\lambda \cdot 10^{-4}$	$\sigma \cdot 10^{-4}$			
1	2	3	4	5	6	7	8
0	4,991958	0,222956	4,977257	0,222848	298	1	101
200	4,991782	0,224347	4,976923	0,224238	297	0	103
400	4,989463	0,224755	4,974544	0,224645	296	1	103
600	5,012335	0,225798	5,017284	0,225686	295	2	103
800	5,023641	0,227834	5,028321	0,227719	293	4	103
1000	4,997294	0,231213	4,981527	0,231093	290	7	103
2000	4,981366	0,286938	4,977115	0,286708	247	50	103



Рис. 6. – Результаты оценки параметра интенсивности отказа распределения Вейбулла методом МФР в зависимости от момента начала наблюдения T_0 . $T_k = 3000$ ч.

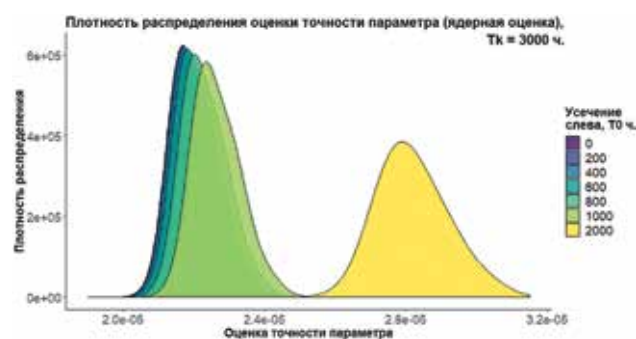


Рис. 7. Результаты оценки точности параметра интенсивности отказа распределения Вейбулла методом МФР в зависимости от момента начала наблюдения T_0 . $T_k = 3000$ ч.

Табл. 4. Результаты оценки параметра интенсивности отказа Вейбулла в зависимости от момента начала наблюдения T_0 . $T_k = 2500$ ч.

T_0 , ч.	МП		МФР		Объем полной выборки	Полностью отсутствуют по причине усечения, шт.	Полностью отсутствуют по причине цензурирования, шт.
	$\lambda \cdot 10^{-4}$	$\sigma \cdot 10^{-4}$	$\lambda \cdot 10^{-4}$	$\sigma \cdot 10^{-4}$			
1	2	3	4	5	6	7	8
0	4,959233	0,274137	4,936879	0,273933	246	1	153
200	4,977356	0,273976	4,965023	0,273772	246	0	154
400	4,973461	0,274665	4,960996	0,274459	245	1	154
600	4,977281	0,276573	4,964518	0,276363	244	2	154
800	4,978054	0,280340	4,964666	0,280121	242	4	154
1000	4,992010	0,286706	4,987563	0,286473	239	7	154
2000	5,060327	0,423020	5,006371	0,422233	196	50	154



Рис. 8. Результаты оценки параметра интенсивности отказа распределения Вейбулла методом МФР в зависимости от момента начала наблюдения T_0 . $T_k = 2500$ ч.



Рис. 9. Результаты оценки точности параметра интенсивности отказа распределения Вейбулла методом МФР в зависимости от момента начала наблюдения T_0 . $T_k = 2500$ ч.

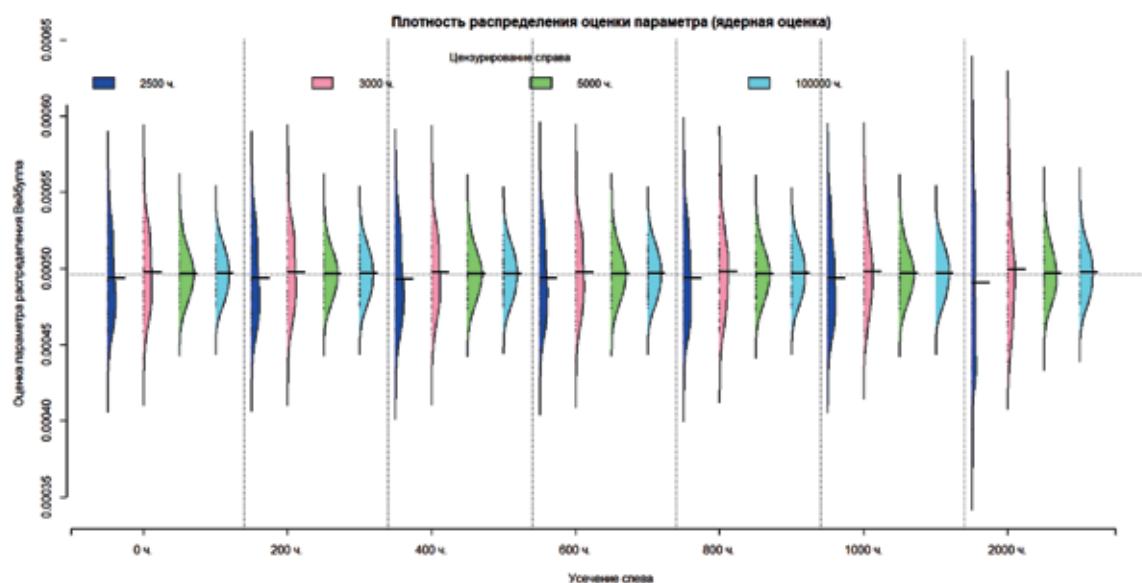


Рис. 10. Результаты оценки параметра интенсивности отказа распределения Вейбулла методом МФР в зависимости от момента начала наблюдения

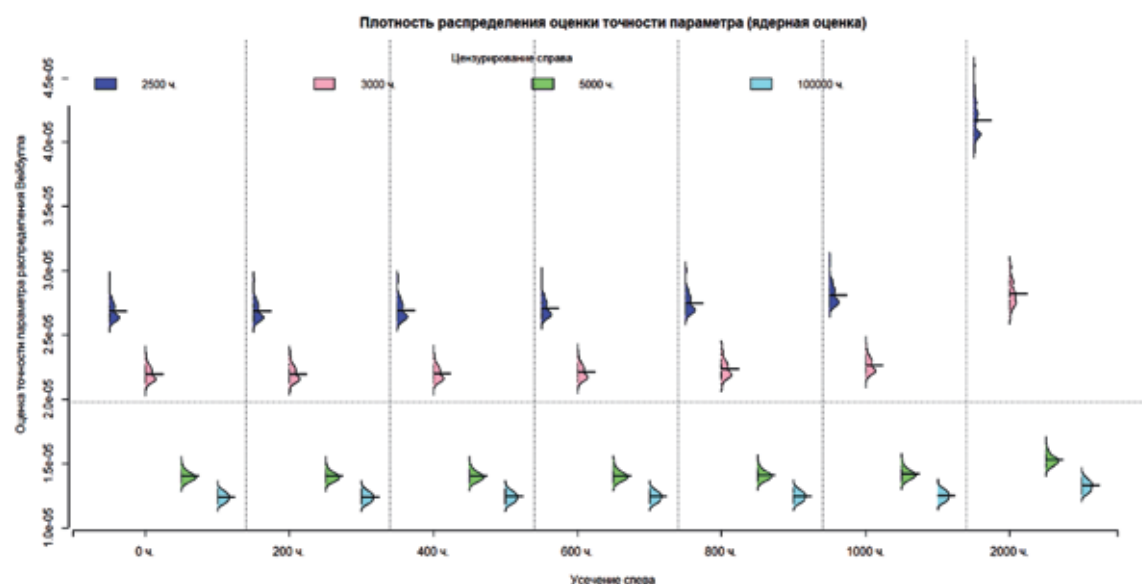


Рис. 11. Результаты оценки точности параметра интенсивности отказа распределения Вейбулла методом МФР в зависимости от момента начала наблюдения

На рис. 10–11 представлена обобщенная картина изменения оценок и точности оценок от доли цензурированных и усеченных данных.

Анализируя полученные рисунки можно прийти к следующим выводам:

- корректный учет специфики имеющихся данных в методике оценивания, а именно усечение слева и цензура справа, приводит к оценкам, не имеющим смещения;
- глубина цензурирования и усечения влияют на точность оценки, т.к. они в той или иной степени вносят неопределенность в исходные данные;
- увеличение цензуры справа приводит к более точным оценкам интенсивности отказа;
- увеличение усечения слева, при таком выборе моделирования, практически сказывается на точности оценки

в ситуации, когда точка усечения слева приближается к точке цензуры справа.

Результаты

По результатам моделирования можно сделать вполне ожидаемый вывод.

Чем позже мы начали наблюдать за объектами (больше усечение слева), тем больше дисперсия оценки. Чем позже мы прервали наблюдения (меньше глубина цензурирования), тем меньше дисперсия.

В результате показана возможность получения оценок характеристик надежности с учетом не только полных наработок объектов, но и наработок, цензурированных справа и усеченных слева. Выполнены исследования изменения значений оценок параметра

масштаба распределения Вейбулла и точности в его оценке в зависимости от объема усеченных данных и глубины цензурирования методом минимизации функции риска.

Приведенный пример иллюстрирует тот факт, что учет цензурированной и усеченной информации приводит к существенному усложнению процедуры вычисления параметров рассматриваемого закона распределения. Следует также отметить, что сложность решения задачи оценивания компенсируется точностью оценок, которую в итоге удается достичь за счет учета цензурированной и усеченной информации.

Показано, что оба метода (метод максимального правдоподобия и метод минимизации функции риска) дают одинаковые результаты при совершенно разных подходах к оцениванию. Однако, к положительной стороне метода минимизации функции риска можно отнести возможность учета априорной информации.

Библиографический список

1. Антонов А.В. Системный анализ: учебник. 4-е изд., перераб. и доп. М.: ИНФРА-М, 2017. 366 с.
2. Антонов А.В., Ершов А.Н., Чепурко В.А. Оценивание параметров закона распределения при неполной информации об отказах // Диагностика и прогнозирование состояния сложных систем, Сб. науч. тр. № 17 каф. АСУ. – Обнинск: ИАТЭ, 2007. с. 16-21.
3. Кокс Д.Р., Оукс Д. Анализ данных типа времени жизни. М.: Финансы и статистика, 1988. 191 с.
4. Balakrishnan N., Debanjan M. Likelihood inference for left truncated and right censored data // Computational Statistics and Data Analysis. September 2011. P. 58.
5. Hong Y.Q., Meeker W.Q., McCalley J.D. Prediction of remaining life of power transformers based on left truncated and right censored lifetime data // The Annals of Applied Statistics. 2009. Vol. 3. № 2. Pp. 857-879.

References

1. Antonov A.V. [System Analysis: A Study Guide. 4-th edition, revised and extended]. Moscow: INFRA-M; 2017. (in Russ.)
2. Antonov A.V., Yershov A.N., Chepurko V.A. [Estimating the parameters of the distribution law in cases of incomplete failure information]. In: [Diagnostics and prediction of the state of complex systems. Proceedings no. 17 of the ACS Department]. Obninsk: OINPE; 2007:16-21. (in Russ.)
3. Cox D.R., Oakes D. Analysis of Survival Data. Moscow: Finansy i statistika; 1988.

4. Balakrishnan N., Debanjan M. Likelihood inference for left truncated and right censored data. *Computational Statistics and Data Analysis*; 2011. P. 58.

5. Hong Y.Q., Meeker W.Q., McCalley J.D. Prediction of remaining life of power transformers based on left truncated and right censored lifetime data. *The Annals of Applied Statistics* 2009;3(2):857-879.

Сведения об авторах

Николаев Дмитрий Александрович – ведущий инженер отдела расчетных обоснований проектных решений АО «РАСУ», Москва, Российская Федерация, e-mail: dafanday@gmail.com

Антонов Александр Владимирович – доктор технических наук, профессор, главный эксперт АНО ДПО «Техническая Академия Росатома», Обнинск, Российская Федерация, e-mail: AVAntonov@rosatomtech.ru

Чепурко Валерий Анатольевич – кандидат физико-математических наук, доцент, главный специалист отдела расчетных обоснований проектных решений АО «РАСУ», Москва, Российская Федерация, e-mail: VAChepurko@rasu.ru

About the authors

Dmitry A. Nikolaev, Lead Engineer, Division for Computational Substantiation of Design Solutions, RASU, Moscow, Russian Federation, e-mail: dafanday@gmail.com.

Alexander V. Antonov, Doctor of Engineering, Professor, Chief Expert, Rosatom Technical Academy, Obninsk, Russian Federation, e-mail: AVAntonov@rosatomtech.ru.

Valery A. Chepurko, Candidate of Physics and Mathematics, Associate Professor, Chief Specialist, Division for Computational Substantiation of Design Solutions, RASU, Moscow, Russian Federation, e-mail: VAChepurko@rasu.ru.

Вклад авторов в статью

Николаев Д.А. выполнил обзор литературы, провел анализ пропущенных данных, разработал подход для применения метода минимизации функции риска для учета усеченных слева и цензурированных справа данных для оценки параметров распределения Вейбулла, применил подход на тестовом примере.

Антонов А.В. выполнил обзор литературы, постановку задачи и анализ исходных данных.

Чепурко В.А. провел анализ возможных видов работ до отказа.

Конфликт интересов

Авторы заявляют об отсутствии конфликта интересов.