

УДК 004.5; 62-192

М.С. Абрамович¹, М.Н. Мицкевич¹, Н.Н. Пыжик²

АЛГОРИТМИЧЕСКОЕ И ПРОГРАММНОЕ ОБЕСПЕЧЕНИЕ СТАТИСТИЧЕСКОГО ОЦЕНИВАНИЯ ПОКАЗАТЕЛЕЙ НАДЕЖНОСТИ АВТОТРАНСПОРТНЫХ СРЕДСТВ ПО ЦЕНЗУРИРОВАННЫМ ВЫБОРКАМ

Рассматривается подход к статистическому оцениванию показателей надежности автотранспортных средств по выборкам ограниченного объема. Приводится описание программного комплекса оценивания показателей надежности.

Введение

Важное место в комплексе мероприятий по обеспечению надежности автомобильных транспортных средств (АТС): автомобилей, прицепного состава, их сборочных единиц и деталей (далее объектов) – занимают статистические методы [1]. Их можно классифицировать на две группы. К первой группе относятся методы, которые основаны на использовании теоретических законов распределения случайных величин и предназначены для вычисления показателей долговечности, безотказности, сохраняемости на основе параметров законов распределения, ко второй – методы, которые основаны на аппроксимации эмпирических зависимостей уравнениями регрессии и предназначены для вычисления комплексных показателей надежности, ремонтпригодности, технико-экономических и эксплуатационных показателей использования объектов в зависимости от их наработки с начала эксплуатации.

Как показывает практика обработки данных о надежности АТС, в качестве теоретических законов распределения, как правило, используются нормальный, экспоненциальный, трехпараметрические логнормальный и Вейбулла [1, 2].

Применяемые при этом выборки данных имеют две основные характерные особенности:

– часть объектов за период эксплуатации или испытаний (или на момент обработки данных) имеет отказы (наработки на момент отказа известны), а часть – не имеет (известны наработки до цензурирования – наработки работоспособных объектов). Это обусловлено двумя причинами: разной долговечностью деталей АТС, заменой некоторых деталей в работоспособном состоянии;

– не всегда предоставляется возможность сформировать выборки объектов достаточно большого объема, но, тем не менее, уже по выборкам ограниченного объема требуется оценить надежность объектов до начала серийного выпуска, чтобы по возможности предотвратить тиражирование отказов.

Оценка показателей надежности для конкретных моделей АТС требует больших вычислительных и информационных ресурсов, связанных с созданием и поддержкой баз данных большой номенклатуры и объема, выбором и обработкой информации из баз данных, процедурами вычисления показателей надежности, формированием выходных форм отчетов. Все перечисленные функции реализованы в программном комплексе, описанном в работе.

1. Оценивание показателей надежности на основе функций распределения

На основе функций распределения определяются следующие показатели надежности: средний и гамма-процентный ресурс, среднее квадратическое отклонение ресурса, вероятность безотказной работы, интенсивность отказов, которые вычисляются на основании оценок параметров законов распределения.

Процедура оценивания показателей надежности включает оценку параметров законов распределения, выбор теоретического закона распределения, соответствующего эмпирическим данным с использованием критериев согласия, и оценку показателей надежности на основе параметров выбранного закона распределения.

1.1. Оценивание параметров распределений по цензурированным выборкам

Предположим, что сведения о надежности изделий представлены наработками работоспособных и отказавших объектов:

$$t_1, t_2, \dots, t_r; \tau_1, \tau_2, \dots, \tau_n, \quad (1)$$

где t_i ($i=1, \dots, r$) – наработки изделий до отказа; τ_j ($j=1, \dots, n$) – наработки изделий до цензурирования; объем рассматриваемой выборки $N = r + n$.

По результатам эксплуатационных испытаний вида (1) требуется оценить показатели надежности объектов. В выборке (1) в общем случае значения наработок до отказа и наработок до цензурирования не равны между собой и могут чередоваться случайным образом. Такая модель выборки называется многократно цензурированной случайной выборкой [1].

Пусть многократно цензурированная случайная выборка описывается некоторым законом распределения с плотностью $f(t, \Theta_1, \Theta_2, \dots, \Theta_l)$ и функцией распределения $F(t, \Theta_1, \Theta_2, \dots, \Theta_l)$, где $\Theta = (\Theta_1, \Theta_2, \dots, \Theta_l)$ – вектор неизвестных параметров. Рассмотрим подход для вычисления оценок параметров распределения по цензурированной выборке с использованием метода максимального правдоподобия. Как и в [1], будем предполагать, что в (1) отказы и цензурирования независимы и закон распределения отказов и цензурирований один и тот же. Тогда функция правдоподобия для этой выборки может быть записана в виде

$$L = \prod_{i=1}^r f(t_i) \times \prod_{j=1}^n [1 - F(\tau_j)].$$

Оценки максимального правдоподобия параметров распределения (МП-оценки) получаются решением уравнений правдоподобия:

$$\frac{\partial L}{\partial \Theta_i} = 0, \quad i = 1, \dots, l.$$

Для нормального и экспоненциального законов распределения МП-оценки параметров распределений по цензурированным выборкам подробно описаны в [1]. Рассмотрим алгоритм построения МП-оценок для трехпараметрического распределения Вейбулла с плотностью

$$f(t, a, b, c) = \frac{b}{a} \left(\frac{t-c}{a} \right)^{b-1} e^{-((t-c)/a)^b}, \quad t > c.$$

Функция правдоподобия для нахождения оценок параметров распределения Вейбулла запишется в виде

$$L = \prod_{i=1}^r \frac{b}{a^b} (t_i - c)^{b-1} e^{-\left(\frac{t_i-c}{a}\right)^b} \times \prod_{j=1}^n \left[e^{-\left(\frac{\tau_j-c}{a}\right)^b} \right],$$

а ее логарифм – в виде

$$\ln L = r \ln b + (b-1) \sum_{i=1}^r \ln(t_i - c) - br \ln a - \sum_{i=1}^r \left(\frac{t_i - c}{a} \right)^b - \sum_{j=1}^n \left(\frac{\tau_j - c}{a} \right)^b. \quad (2)$$

Зафиксировав параметр сдвига c , используем необходимое условие экстремума и на основании (2) получим следующую систему уравнений для нахождения оценок параметров a и b :

$$\frac{1}{\hat{b}} = \frac{\sum_{i=1}^r (t_i - c)^{\hat{b}} \ln(t_i - c) + \sum_{j=1}^n (\tau_j - c)^{\hat{b}} \ln(\tau_j - c) - \sum_{i=1}^r \ln(t_i - c)}{\sum_{i=1}^r (t_i - c)^{\hat{b}} + \sum_{j=1}^n (\tau_j - c)^{\hat{b}} - r}; \quad (3)$$

$$\hat{a} = \left(\frac{\sum_{i=1}^r (t_i - c)^{\hat{b}} + \sum_{j=1}^n (\tau_j - c)^{\hat{b}}}{r} \right)^{1/\hat{b}}. \quad (4)$$

Уравнение (3) решается методом последовательных приближений. Начальное приближение \hat{b}_0 выбирается следующим образом [1]:

$$\hat{b}_0 = \frac{r+1}{\prod_{i=1}^r t_i \ln \frac{i-1}{t_{(1)}} (0,23r+3,75)}, \quad (5)$$

где $t_{(1)}$ – минимальное значение наработки до отказа. Затем k -е приближение определяется по формуле

$$\hat{b}_k = \left(\frac{\sum_{i=1}^r (t_i - c)^{\hat{b}_{k-1}} \ln(t_i - c) + \sum_{j=1}^n (\tau_j - c)^{\hat{b}_{k-1}} \ln(\tau_j - c) - \sum_{i=1}^r \ln(t_i - c)}{\sum_{i=1}^r (t_i - c)^{\hat{b}_{k-1}} + \sum_{j=1}^n (\tau_j - c)^{\hat{b}_{k-1}} - r} \right)^{-1}. \quad (6)$$

Итеративный процесс продолжается до тех пор, пока не будет достигнута требуемая относительная погрешность ε :

$$\left| \frac{\hat{b}_k - \hat{b}_{k-1}}{\hat{b}_k} \right| \leq \varepsilon. \quad (7)$$

В качестве оценки параметра b принимается $\hat{b} = \hat{b}_k$. Оценка параметра a определяется в соответствии с (4). Для этих значений оценок параметров проверяется достаточное условие максимума логарифма функции правдоподобия (2):

$$r \left(\sum_{i=1}^r \left(\frac{t_i - c}{a} \right)^b \ln^2 \left(\frac{t_i - c}{a} \right) + \sum_{j=1}^n \left(\frac{\tau_j - c}{a} \right)^b \ln^2 \left(\frac{\tau_j - c}{a} \right) \right) - 2r \sum_{i=1}^r \ln \left(\frac{t_i - c}{a} \right) - \left(\sum_{i=1}^r \ln \left(\frac{t_i - c}{a} \right) \right)^2 > 0.$$

Для оценки параметра сдвига c трехпараметрического распределения Вейбулла выполняется следующая процедура:

1. Выбирается некоторое натуральное число K . Пусть $\Delta = x_{(2)} - x_{(1)}$, где $x_{(i)}$ – i -я порядковая статистика выборки. Отметим, что в качестве правой границы изменения параметра c можно взять минимальный элемент выборки $x_{(1)}$, а в качестве левой – 0. Тогда $K = \left\lceil \frac{x_{(1)}}{\Delta} \right\rceil$.

2. В цикле меняется значение параметра c : $c_m = x_{\min} - m\Delta$, $m = 1, \dots, K$. В качестве текущей оценки для параметра c принимается значение c_m и на основании (6), (7) определяются оценки \hat{a}_m, \hat{b}_m и $\ln L_m$.

3. Определяется номер шага $m^* = \arg \max_m (\ln L_m)$. В качестве оценок параметров распределения Вейбулла принимаются значения $\hat{a} = \hat{a}_{m^*}$, $\hat{b} = \hat{b}_{m^*}$, $\hat{c} = c_{m^*}$.

1.2. Построение точечных и интервальных оценок показателей надежности

Как отмечалось выше, оценки показателей надежности вычисляются на основании оценок параметров распределения. Так, например, если выборка (1) подчиняется закону распределения Вейбулла, оценка среднего ресурса \hat{L} вычисляется следующим образом [1]:

$$\hat{L} = \hat{a} \Gamma \left(1 + \frac{1}{\hat{b}} \right) + \hat{c}.$$

Известно [3], что из-за потерь информации при цензурировании выборки уменьшается точность оценивания и имеет место смещение традиционных оценок параметров. При этом величина смещения зависит от степени цензурирования и от объема выборки.

Количество информации Фишера по цензурированной выборке определяется следующим соотношением [4]:

$$J_c(\theta) = \frac{1}{P_1(\theta)} \left[\frac{\partial P_1(\theta)}{\partial \theta} \right] \left[\frac{\partial P_1(\theta)}{\partial \theta} \theta \right]^T + \int_{x_1}^{x_2} \left[\frac{\partial \ln f(x, \theta)}{\partial \theta} \right] \left[\frac{\partial \ln f(x, \theta)}{\partial \theta} \right]^T f(x, \theta) dx + \\ + \frac{1}{P_2(\theta)} \left[\frac{\partial P_2(\theta)}{\partial \theta} \right] \left[\frac{\partial P_2(\theta)}{\partial \theta} \right]^T,$$

где $P_1(\theta)$ и $P_2(\theta)$ – вероятности попадания в область цензурирования слева и справа, а наблюдаемая область лежит в пределах от x_1 до x_2 .

В работе [4] методом статистического моделирования исследована эффективность оценивания параметров законов распределения по цензурированной выборке по отношению к оцениванию по полной выборке. Соотношение (5) позволяет судить о потерях информации при оценивании параметров распределения в зависимости от степени цензурирования.

Эффективность оценивания параметров определяется соотношением

$$S = \frac{\det J_c(\theta)}{\det J(\theta)},$$

где $J(\theta)$ – информационная матрица Фишера, построенная по полной выборке; $J_c(\theta)$ – информационная матрица Фишера, построенная по цензурированной выборке; \det – определитель матрицы.

Результаты проведенного в [4] исследования показали, что при умеренной доле цензурирования (до 40 % от объема полной выборки) имеет место высокая эффективность оценивания параметров законов распределения, т. е. сохраняющаяся в цензурированной выборке информация является достаточной для оценивания параметров распределения с требуемой точностью. Таким образом, основной проблемой при оценке показателей надежности при умеренном цензурировании является ограниченность объема анализируемой выборки.

Как отмечается в [4], оценки показателей надежности, построенные по выборкам ограниченного объема, являются смещенными.

Для построения более точных оценок рассмотрим процедуру, основанную на методе бутстрепа [5] и состоящую из следующих шагов:

1. По цензурированной выборке объема N вычисляются МП-оценки $\hat{\Theta}_1, \hat{\Theta}_2, \dots, \hat{\Theta}_l$ параметров закона распределения $F(t, \Theta_1, \Theta_2, \dots, \Theta_l)$.

2. Генерируются M независимых выборок объема N распределения $F(t, \hat{\Theta}_1, \hat{\Theta}_2, \dots, \hat{\Theta}_l)$. В качестве параметров для моделирования используются оценки параметров, полученные на шаге 1.

3. Для каждой из M выборок вычисляются МП-оценки параметров закона распределения и на основании их оценка показателя надежности. Формируется выборка оценок показателей надежности $R = \{\hat{R}_i, i = 1, \dots, M\}$.

4. В качестве искомого значения оценки показателя надежности принимается среднее значение выборки R .

Для нахождения непараметрических доверительных интервалов показателей надежности по полученной выборке оценок показателей надежности $R = \{\hat{R}_i, i = 1, \dots, M\}$ формируется вариационный ряд этих оценок $V = \{\hat{R}_{(i)}, i = 1, \dots, M\}$. Доверительные границы для показателя надежности R определяются значениями вариационного ряда $[\hat{R}_{(l)}, \hat{R}_{(u)}]$, где $l = \left[\frac{1-\nu}{2} M \right]$, $u = \left[\frac{1+\nu}{2} M \right]$, $[\cdot]$ – целая часть числа, ν – доверительная вероятность [5].

2. Оценивание показателей надежности и эффективности на основе уравнений регрессии

Показатели надежности и эффективности АТС включают: комплексные показатели надежности (коэффициенты готовности, оперативной готовности, технического использования); технико-экономические показатели использования АТС (затраты на запасные части, оплату труда ремонтных рабочих, компенсацию простоя); технико-эксплуатационные показатели использования АТС (среднюю эксплуатационную и техническую скорость, коэффициенты использования пробега и грузоподъемности, коэффициент выпуска, транспортную работу, объем перевозок и др.) [6].

Оценивание данных показателей осуществляется за отдельные отчетные периоды или с накоплением от начала эксплуатации (в зависимости от физического смысла и цели исследований). Аппроксимация эмпирических значений показателей проводится по следующим типам кривых:

$$kx + d, dx^k, de^{kx}, d + \frac{k}{x}, d + k \ln x, \frac{1}{d + kx}, de^{k/x}, \frac{1}{d + ke^{-x}}, (d + k \ln x)x,$$

где x – значение пробега или продолжительности эксплуатации АТС.

Для вычисления МНК-оценок параметров проводится линеаризация кривых. Для проверки значимости уравнения регрессии применяется F -критерий.

Алгоритм выбора кривой, наиболее точно аппроксимирующей эмпирические данные, состоит из следующих шагов:

1. Из представленного набора аппроксимирующих кривых исключаются те, различие которых с эмпирическими данными являлось значимым по F -критерию.

2. Из оставшихся исключаются такие кривые, экстраполяция которых некорректна (значения в отрицательной области, появление точек разрыва) и характер изменения которых противоречит физическому смыслу величины (например, коэффициент готовности будет монотонно убывать с ростом значения пробега, а затраты на запасные части – возрастать).

3. В качестве кривой, наиболее адекватно описывающей эмпирические данные, выбирается та, для которой среднеквадратическая ошибка прогноза минимальна.

Построенные уравнения зависимости показателей надежности и эффективности от пробега используются для прогнозирования этих показателей в краткосрочном периоде. На рис. 1 показаны эмпирические и теоретические значения удельной стоимости запасных частей, применяемых для ремонта АТС, в зависимости от пробега. Для вычисления теоретических значений в приведенном примере была использована линейная зависимость $y = kx + b$.

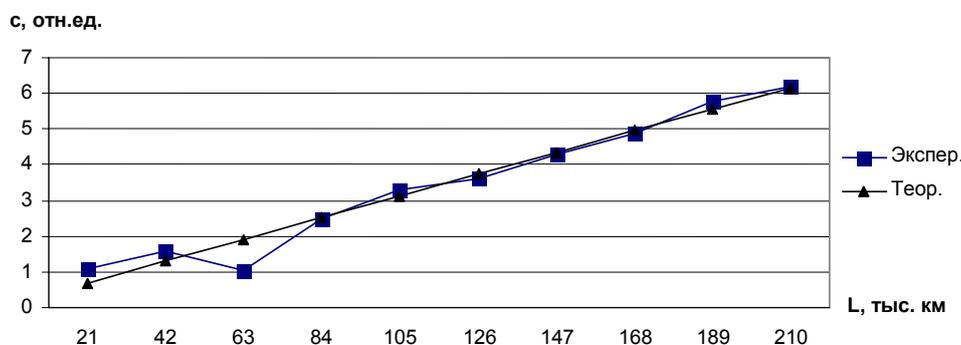


Рис. 1. Зависимость удельной стоимости запасных частей от пробега АТС

3. Структура программного комплекса оценки показателей надежности

Программный комплекс оценки показателей надежности (рис. 2) предназначен для сбора и обработки информации о надежности, вычисления показателей надежности и эффективности АТС и формирования соответствующих отчетов.

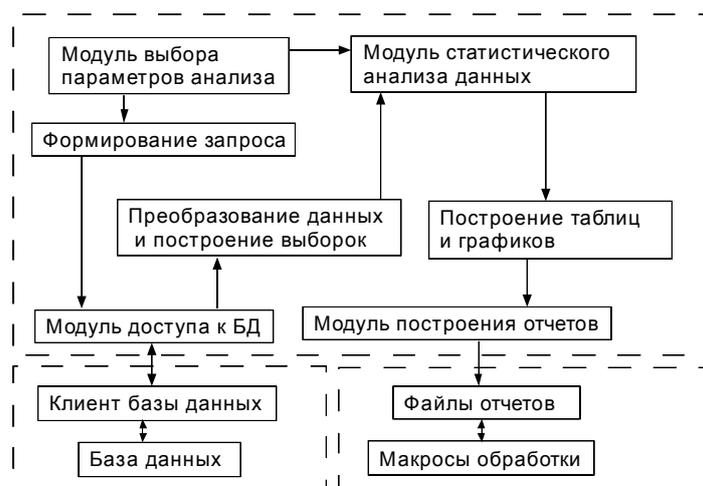


Рис. 2. Структура программного комплекса

Модуль выбора параметров анализа включает в себя выбор объекта статистического анализа (детали, сборочной единицы, АТС в целом), выбор модели, контрольной группы АТС, расчетного периода и вычисляемого показателя надежности.

База данных об отказах автомобилей МАЗ, разработанная в среде Oracle, содержит справочную информацию, а также отчетные данные по отказам деталей за определенный период – с начала эксплуатации на дату наличия данных (информационную карту о надежности). Для работы с информационными картами разработан модуль автоматизированного рабочего места сбора и обработки информации о надежности.

Модуль доступа к базе данных осуществляет:

- редактирование информации базы данных (добавление, исправление, удаление) по всей исходной накопленной информации;
- редактирование справочников и классификаторов объектов (АТС, деталей, сборочных единиц), неисправностей (по характеру, причинам, влиянию на безопасность использования, сложности и способам устранения);

– получение по запросу исходных данных или результатов расчета.

Создание и корректировка базы данных, классификаторов и справочников объектов и неисправностей проводится по мере возникновения неисправностей объектов.

В модуле статистического анализа данных реализованы алгоритмы вычисления показателей надежности и эффективности, предусмотрены расчет и вывод промежуточных результатов для обеспечения возможности оперативного анализа.

Модуль построения отчетов позволяет выполнять:

- подготовку таблиц по заданным показателям, объектам, классификационным группам объектов и признаков классификации показателей;
- построение графиков изменения показателей надежности;
- передачу данных в файл Excel и запуск макросов обработки отчетов;
- формирование отчетов.

Заключение

В настоящей работе рассмотрен подход к построению оценок параметров законов распределения по многократно цензурированным выборкам объектов с использованием метода максимального правдоподобия. Для построения точечных и интервальных оценок показателей надежности в случае выборок ограниченного объема предложен алгоритм, основанный на методе бутстрепа [5].

Разработанный программный комплекс используется в управлении главного конструктора ОАО «МАЗ» для оценки показателей надежности серийных моделей автомобилей семейства МАЗ и прогнозирования показателей надежности проектируемых моделей.

Работа выполнена в рамках Государственной комплексной программы научных исследований «Инфотех» (задание «Инфотех 51»).

Список литературы

1. Аронов, И.З. Оценка надежности по результатам сокращенных испытаний / И.З. Аронов, Е.И. Бурдасов – М. : Изд-во стандартов, 1987. – 184 с.
2. Анализ надежности технических систем по цензурированным выборкам / В.М. Скрипник [и др.]. – М. : Радио и связь, 1988. – 184 с.
3. Лемешко, Б.Ю. Об оценивании параметров распределений и проверке гипотез по цензурированным выборкам / Б.Ю. Лемешко // Методы менеджмента качества. – 2001. – № 4. – С. 32–38.
4. Лемешко, Б.Ю. К оцениванию параметров надежности по цензурированным выборкам / Б.Ю. Лемешко, С.Я. Гильдебрант, С.Н. Постовалов // Заводская лаборатория. Диагностика материалов. – 2001. – Т. 67, № 1. – С. 52–64.
5. Diccio, T.J. Bootstrap Confidence Intervals / T.J. Diccio, B. Efron // Statistical Science. – 1996. – Vol. 11, № 3. – P. 189–228.
6. Алексеева, И.М. Статистика автомобильного транспорта / И.М. Алексеева, О.И. Ганченко, Е.В. Петрова. – М. : Экзамен, 2005. – 352 с.

Поступила 19.03.2010

¹Научно-исследовательский институт
прикладных проблем математики и информатики,
Минск, пр. Независимости, 4
e-mail: abramovichms@bsu.by

²ОАО «Минский автомобильный завод»,
Минск, Социалистическая, 2

M.S. Abramovich, M.M. Mitskevich, N.N. Pygk

**ALGORITHMS AND SOFTWARE
FOR THE STATISTICAL ESTIMATION OF RELIABILITY INDICATORS
OF VEHICLES ON THE CENSORED SAMPLES**

An approach to the statistical estimation of reliability indicators for vehicles with the limited size censored samples is considered. The description of the software for the estimation of reliability indicators is given.