ДОКЛАДЫ АН ВШ РФ
2015 январь–март № 1 (26)
ТЕХНИЧЕСКИЕ НАУКИ

УДК 519.2

ПРОВЕРКА АДЕКВАТНОСТИ ПАРАМЕТРИЧЕСКИХ РЕГРЕССИОННЫХ МОДЕЛЕЙ НАДЕЖНОСТИ ПО УСЕЧЕННЫМ СЛЕВА И ЦЕНЗУРИРОВАННЫМ СПРАВА ДАННЫМ

Е.В. Чимитова, М.А. Семёнова
Новосибирский государственный технический университет

В настоящей работе рассмотрены основные проблемы построения вероятностных моделей надежности с учетом объясняющих переменных по усеченным слева и цензурированным справа выборкам. Проведено исследование точности оценок максимального правдоподобия параметров распределения Вейбулла по усеченным слева выборкам. Показано, что при одновременном оценивании двух параметров определитель ковариационной матрицы для усеченной выборки значительно больше, чем для полной выборки отказов. С ростом глубины усечения определитель ковариационной матрицы для усеченной выборки увеличивается. Предложен универсальный подход к проверке гипотезы о согласии с вероятностной моделью надежности по усеченным слева и цензурированным справа выборкам. Предлагаемый метод заключается в использовании модифицированных критериев типа Колмогорова, Крамера–Мизеса–Смирнова и Андерсона–Дарлинга для проверки гипотезы о равномерном распределении выборок остатков, полученных в соответствии с проверяемой вероятностной моделью. Сформулированный алгоритм статистического моделирования неизвестных распределений статистик критериев согласия позволяет их корректное применение для усеченных и цензурированных выборок. В результате исследования мощности рассматриваемых критериев показано, что метод проверки согласия на основе выборок остатков позволяет проверять как предположения относительно вида регрессионной зависимости, так и предположения о виде закона распределения отказов.

Ключевые слова: цензурированная выборка, усеченная выборка, модель надежности, модель пропорциональных интенсивностей, модель ускоренных испытаний, проверка адекватности, критерии согласия.

DOI: 10.17212/1727-2769-2015-1-104-120

Введение

Одной из важнейших задач в теории надежности является построение модели зависимости функции надежности от объясняющих переменных, в качестве которых обычно выступают воздействия различного типа, такие как температура, давление, напряжение, механические нагрузки и другие. Для описания подобного рода зависимостей наиболее широко используются такие модели, как модель ускоренных испытаний [1] и модель пропорциональных интенсивностей Кокса, предложенная в [2]. Если зависимость функции надежности от объясняющих переменных не удается описать данными моделями, в [3] предложены их обобщения.

В зависимости от характера привлекаемой априорной информации различают три типа вероятностных моделей надежности: непараметрические модели, полупараметрические модели, требующие спецификации вида регрессионной зависи-

Исследование выполнено при поддержке Министерства образования и науки Российской Федерации в рамках проектной части государственного задания в сфере научной деятельности № 2.541.2014К от 17.07.2014.

мости, и параметрические модели, для построения которых необходима информация как о виде регрессионной зависимости, так и о законе распределения отказов. Один из методов построения непараметрической регрессионной модели надежности был предложен в [4] и получил дальнейшее развитие в [5]. Методы построения полупараметрических моделей пропорциональных интенсивностей и ускоренных испытаний рассматривались, например, в [6–8]. В данной работе рассматриваются вопросы построения параметрических моделей надежности, на основе которых рассчитываются интересующие исследователя показатели надежности при заданных значениях объясняющих переменных, например, время безотказной работы с заданной вероятностью, средняя наработка на отказ или интенсивность отказов.

В задачах анализа данных типа времени жизни полученные в результате проведенного эксперимента выборки, как правило, являются цензурированными. Цензурированные наблюдения в отличие от полных наблюдений не содержат информации о точном времени наступления отказа, указывая лишь на время, в течение которого отказа не произошло. Более того, на практике часто возникают ситуации, когда в выборку попадают только те объекты, время безотказной работы которых не меньше некоторой величины. Такие наблюдения называют усеченными, а соответствующие выборки отказов называются усеченными слева.

Неизвестные параметры моделей, как правило, оценивают методом максимального правдоподобия, так как данный метод является универсальным относительно формы регистрации наблюдений и позволяет получать состоятельные и асимптотически эффективные оценки. Для вычисления оценок максимального правдоподобия (ОМП) по усеченным слева и цензурированным справа данным в [9–11] применяется *EM*-алгоритм.

Основной проблемой, возникающей при построении вероятностных моделей надежности, является проверка их адекватности, т. е. проверка справедливости предположения о виде регрессионной зависимости и распределении отказов, для чего используются статистические критерии согласия. Проблема заключается в том, что проверка гипотезы о согласии осуществляется на основе цензурированной и усеченной выборки наблюдений, которые имеют различные распределения и получены при разных значениях объясняющих переменных, вследствие чего использование классических критериев согласия невозможно. В случае усеченных слева и цензурированных справа данных в [12] предложен параметрический критерий проверки гипотезы о согласии с моделью Кокса против обобщенной модели пропорциональных интенсивностей. В настоящей работе предлагается универсальный метод проверки гипотезы о согласии с вероятностной моделью надежности по усеченным слева и цензурированным справа выборкам. Предлагаемый метод заключается в использовании модифицированных критериев типа Колмогорова, Крамера–Мизеса–Смирнова и Андерсона–Дарлинга для проверки сложной гипотезы о виде распределения выборок остатков, полученных в соответствии с проверяемой вероятностной моделью.

1. Данные типа времени жизни

Пусть T_x – неотрицательная случайная величина, определяющая время безотказной работы, которое зависит от вектора объясняющих переменных (воздействий) $x = (x_1, x_2, ..., x_m)^{\mathrm{T}}$, область значений которых определяется условиями проведения эксперимента и представляет собой отрезок числовой прямой.

В зависимости от условий проведения эксперимента и способа регистрации отказов можно выделить следующие виды выборок отказов: полные, цензурированные справа и усеченные слева. Рассмотрим каждую из них подробнее.

Выборка отказов называется полной, если для каждого объекта в выборке определены значения времен отказов:

$$\mathrm{X}_n = \left\{ \left(T_1, x^1\right), \left(T_2, x^2\right), ..., \left(T_n, x^n\right) \right\}, \qquad (1)$$

где n – объем выборки, x^i – значение вектора объясняющих переменных для i-го объекта.

Цензурированной справа называют выборку вида

$$\mathrm{X}_n = \left\{ \left(X_1, x^1, \delta_1\right), \left(X_2, x^2, \delta_2\right), ..., \left(X_n, x^n, \delta_n\right) \right\}, \qquad (2)$$

где $X_i = \min\{T_i, C_i\}$ – время наступления отказа T_i или момент цензурирования C_i (время завершения наблюдения за i-м объектом), δ_i – индикатор цензурирования, который содержит информацию о причине прекращения наблюдения, $i = \overline{1,n}$. Если в ходе эксперимента было зафиксировано время отказа, то $X_i = T_i$, $\delta_i = 1$, и данное наблюдение называется полным. Если же T_i неизвестно по причине окончания наблюдения в момент $C_i \le T_i$, то $X_i = C_i$, $\delta_i = 0$, наблюдение называется цензурированным справа.

Цензурированные справа выборки можно разделить на три основных типа и их комбинации. Если время эксперимента ограничено, т. е. наблюдение за объектами ведется до заранее определенного момента времени c, тогда $\forall \delta_i = 0 : C_i = c$, и полученная в результате выборка называется цензурированной I типа. В случае, если эксперимент продолжается до наступления определенного количества отказов r, и наблюдение за остальными объектами прекращается в момент наступления r-го отказа, то полученная в результате выборка наблюдений называется цензурированной II типа, и $\forall \delta_i = 0 : C_i = T_{(r)}$, где $T_{(r)}$ – время наступления последнего отказа. Если C_i, $i = \overline{1,n}$, – случайные величины, то выборка наблюдений (2) называется цензурированной III типа или случайно цензурированной. Данный тип цензурирования, в свою очередь, делится на неинформативное цензурирование, когда случайные величины T_i и C_i, $i = \overline{1,n}$, являются независимыми, и информативное цензурирование, при котором параметры распределения случайных величин C_i зависят от распределения отказов T_i, $i = \overline{1,n}$. В данной работе при исследовании свойств оценок параметров и распределений статистик критериев по случайно цензурированным выборкам рассматривается неинформативное цензурирование.

Рассмотрим понятие усеченной слева выборки. Пусть имеется генеральная совокупность объектов, для которых необходимо построить вероятностную модель надежности. При этом часть объектов была введена в эксплуатацию до начала эксперимента. Обозначим через $F(t)$ функцию распределения случайной величины T – времени с начала эксплуатации до отказа. Предположим, что в момент времени t_0 началась регистрация отказов исследуемых объектов с целью анализа их надежности. При этом объекты из рассматриваемой генеральной совокупности, которые отказали до момента времени t_0, не включены в выборку. Необходимо отметить, что если рассматривать полученную выборку времен отказов как обычную выборку из распределения $F(t)$, то полученный результат будет слишком оптимистичным, поскольку чем больше продолжительность безотказной работы объекта, тем больше у него шансов попасть в выборку, в то время как объекты с ранними отказами в выборку не попадают. Обозначим через D_i время с начала эксплуатации *i*-го объекта до начала регистрации отказов t_0 (время усечения). Тогда распределение наблюдаемой случайной величины при условии, что время отказа не меньше времени усечения, имеет вид

$$F_{LT}(t) = \frac{F(t) - F(D_i)}{1 - F(D_i)}, \ t > D_i,$$

аналогично, плотность распределения определяется соотношением

$$f_{LT}(t) = \frac{f(t)}{1 - F(D_i)}, \ t > D_i.$$

В выборку могут быть включены также объекты, эксплуатация которых началась уже после момента времени t_0. В этом случае соответствующий элемент выборки представляет собой неусеченное наблюдение, для которого $D_i = 0$.

Усеченную слева выборку можно представить в следующем виде:

$$\mathbf{X}_n = \left\{ \left(T_1, D_1, x^1\right), \left(T_2, D_2, x^2\right), ..., \left(T_n, D_n, x^n\right) \right\}, \qquad (3)$$

где T_i – время наступления отказа i-го объекта, D_i – время усечения, x^i – значение вектора объясняющих переменных для *i*-го объекта, $i = \overline{1, n}$.

В общем виде рассматриваемые в работе выборки можно записать следующим образом:

$$\mathbf{X}_n = \left\{ \left(X_1, D_1, x^1, \delta_1\right), \left(X_2, D_2, x^2, \delta_2\right), ..., \left(X_n, D_n, x^n, \delta_n\right) \right\}. \qquad (4)$$

Если $\forall i \ \ D_i = 0$ и $\delta_i = 1$, то выборка (4) является полной выборкой (1); если $\forall i \ \ D_i = 0$, то выборка (4) является цензурированной выборкой вида (2); если же $\forall i \ \ \delta_i = 1$, то выборка (4) является усеченной слева выборкой (3). В противном случае выборка (4) представляет собой усеченную слева и цензурированную справа выборку отказов.

2. Вероятностные модели надежности

Закон распределения случайной величины T_x определяется одной из следующих функций: функция плотности распределения $f_x(t)$, функция распределения $F_x(t) = \int_0^t f_x(u)du$, функция надежности $S_x(t) = P\left(T_x \geq t\right) = 1 - F_x(t)$, функция интенсивности $\lambda_x(t) = \lim_{h \to 0} \left(P\left\{ t \leq T_x < t + h \mid T_x > t \right\} / h \right)$ и кумулятивная функция риска $\Lambda_x(t) = \int_0^t \lambda_x(u)du = -\ln\left(S_x(t)\right)$.

Нелинейную регрессионную модель зависимости времени безотказной работы от объясняющих переменных в общем виде можно записать следующим образом:

$$T_x = g(x; \varepsilon), \qquad (5)$$

где $g(\cdot)$ – неотрицательная функция, случайная ошибка ε имеет положительно определенное распределение $F_0(t)$ с конечной дисперсией, которое совпадает с законом распределения случайной величины T_x при нулевых значениях объясняющих переменных: $F_{x=0}(t) = F_0(t)$. Функция $F_0(t)$ называется базовой функцией распределения.

Поскольку в теории надежности объектом моделирования является не столько время безотказной работы, сколько вероятность безотказной работы за наработку t, т.е. функция надежности $S_x(t)$, поэтому регрессионную модель (5) принято записывать через функцию надежности

$$S_x(t) = g_1\left(x; S_0(t)\right), \tag{6}$$

или кумулятивную функцию риска

$$\Lambda_x(t) = g_2\left(x; \Lambda_0(t)\right),$$

где $g_2(x;s) = g_1\left(x; \exp(-s)\right)$, функция $g_1(\cdot)$ определяет изменение базовой функции надежности для различных значений объясняющих переменных.

При построении параметрических моделей надежности базовая функция надежности $S_0(t)$ и функция риска $\Lambda_0(t)$ соответствуют некоторому параметрическому семейству распределений $F_0(t;\,\theta)$. В качестве базовых законов часто используются распределения, функции плотности которых представлены в табл. 1.

Таблица 1 / Table 1

Функции плотности базовых распределений
Density functions of baseline distributions

Распределение / Distribution	**Функция плотности** $f_0(t;\,\theta)$ / **Density function** $f_0(t;\,\theta)$
Экспоненциальное / Exponential	$\frac{1}{\theta_0}\exp\left(-\frac{t}{\theta_0}\right)$
Логнормальное / Lognormal	$\frac{1}{t\cdot\theta_1\sqrt{2\pi}}\exp\left(-\frac{1}{2}\left(\frac{1}{\theta_1}\ln\left(\frac{t}{\theta_0}\right)\right)^2\right)$
Вейбулла / Weibull	$\frac{\theta_1}{t}\left(\frac{t}{\theta_0}\right)^{\theta_1}\exp\left(-\left(\frac{t}{\theta_0}\right)^{\theta_1}\right)$
Гамма / Gamma	$\frac{1}{\theta_0\cdot\Gamma(\theta_1)}\left(\frac{t}{\theta_0}\right)^{\theta_1-1}\exp\left(-\frac{t}{\theta_0}\right)$

Простейшим примером модели (5) является **модель ускоренных испытаний**

$$T_x = r(x;\,\beta)\cdot\varepsilon,$$

которую можно переписать в виде вероятностной модели надежности вида (6) следующим образом:

$$S_x(t;\,\beta,\,\theta) = P\left\{T_x \ge t\right\} = P\left\{r(x;\,\beta)\cdot\varepsilon \ge t\right\} = S_0\left(\frac{t}{r(x;\,\beta)};\theta\right),$$

где $r(x;\beta)$ – положительная функция от воздействий, которая обычно задается в виде

$$r(x;\beta)=\exp\left(\beta_1\varphi_1(x)+\ldots+\beta_m\varphi_m(x)\right),$$

$\varphi_i(x)$ – некоторые функции от компонентов вектора объясняющих переменных (регрессоры), $i=\overline{1,m}$.

Естественным обобщением модели ускоренных испытаний является **модель изменения масштаба и формы** (Changing Scale and Shape model) или CHSS-модель [3]

$$S_x(t;\beta,\gamma,\theta)=S_0\left(\left(\frac{t}{r(x;\beta)}\right)^{\nu(x;\gamma)};\theta\right),$$

где $\nu(x;\gamma)$ и $r(x;\beta)$ являются положительными функциями от объясняющих переменных.

Модель пропорциональных интенсивностей Кокса обычно определяется в терминах кумулятивной функции риска следующим образом [2]:

$$\Lambda_x(t;\beta,\theta)=\exp\left(\beta^{\mathrm{T}}x\right)\cdot\Lambda_0(t;\theta), \tag{7}$$

где β – m-мерный вектор параметров, $\Lambda_0(t;\theta)$ – базовая кумулятивная функция риска. Несмотря на широкую популярность данной модели, обусловленную в том числе наличием возможности ее построения в известных статистических пакетах, таких как SPSS и Statistica, ее применение возможно только при выполнении предположения о пропорциональности интенсивностей:

$$\frac{\lambda_{x=a}(t)}{\lambda_{x=b}(t)}=\frac{\exp\left(\beta^{\mathrm{T}}a\right)}{\exp\left(\beta^{\mathrm{T}}b\right)}=\mathrm{const}.$$

В случае базового распределения Вейбулла модель Кокса совпадает с моделью ускоренных испытаний.

Модель Ксая, полученная возведением базовой функции риска модели пропорциональных интенсивностей (7) в степень $\exp\left(\gamma^{\mathrm{T}}x\right)$, позволяет описать непостоянное отношение функций интенсивности и имеет следующий вид [13]:

$$\Lambda_x(t;\beta,\gamma,\theta)=\exp\left(\beta^{\mathrm{T}}x\right)\left\{\Lambda_0(t;\theta)\right\}^{\exp\left(\gamma^{\mathrm{T}}x\right)}. \tag{8}$$

Параметры β и обобщающие параметры γ являются m-мерными. При $\gamma=\overline{0}$ данная модель представляет собой модель пропорциональных интенсивностей, тогда как при $\gamma\neq\overline{0}$ функции интенсивности при разных значениях ковариат пересекаются [13].

Достаточно подробный обзор различных вероятностных моделей надежности можно найти в [1, 3].

3. Метод максимального правдоподобия

Вектор неизвестных параметров $\eta=\left(\theta^{\mathrm{T}},\beta^{\mathrm{T}},\gamma^{\mathrm{T}}\right)^{\mathrm{T}}$, где θ – вектор параметров базового закона распределения, β – вектор параметров функции от воздействий и

γ – вектор обобщающих параметров обобщенных моделей, по имеющейся выборке отказов можно оценить с помощью метода максимального правдоподобия. Для этого максимизируется логарифм функции правдоподобия:

$$\ln L(\mathbf{X}_n;\eta)=\sum_{i=1}^{n}\left(\delta_i \ln\left(\frac{f_{x^i}(X_i;\eta)}{S_{x^i}(D_i;\eta)}\right)+(1-\delta_i)\ln\left(\frac{S_{x^i}(X_i;\eta)}{S_{x^i}(D_i;\eta)}\right)\right). \quad (9)$$

Если функция правдоподобия дифференцируема по η, то для нахождения ОМП параметра $\eta=\left(\theta^{\mathrm{T}},\beta^{\mathrm{T}},\gamma^{\mathrm{T}}\right)^{\mathrm{T}}$ решают систему уравнений правдоподобия:

$$\frac{\partial \ln L(\mathbf{X}_n;\eta)}{\partial \eta_i}=0,\ i=\overline{1,s}.$$

Оценки, построенные методом максимального правдоподобия, широко применяются на практике, поскольку при достаточно больших объемах выборок для них характерны следующие свойства.

Асимптотическая несмещенность: $M\hat{\eta}\to\eta$, $n\to\infty$.

Состоятельность: $\hat{\eta}\xrightarrow{P}\eta$, $n\to\infty$.

При выполнении условий регулярности модели ОМП обладают свойствами асимптотической эффективности и нормальности:

$$\sqrt{n}\left(\hat{\eta}-\eta\right)\xrightarrow{P}\xi\succ N\left(0,J_n^{-1}(\eta)\right),\ n\to\infty,$$

где $J_n(\eta)$ – информационное количество Фишера о параметре η, содержащееся в выборке $\mathbf{X}_n$.

При построении вероятностных моделей надежности по усеченным и цензурированным выборкам особый интерес представляет исследование свойств ОМП параметров базового закона распределения. Об относительной эффективности оценивания параметров по усеченным и цензурированным выборкам по сравнению с оцениванием по полным выборкам можно судить по величине $J_n(\theta)/J_n^*(\theta)$, где $J_n(\theta)$ – количество информации Фишера в усеченной и/или цензурированной выборке, $J_n^*(\theta)$ – количество информации Фишера в полной выборке. Для цензурированных I или II типа выборок в [14] для ряда законов распределения найдены значения $J_n(\theta)/J_n^*(\theta)$, либо значения отношения определителей соответствующих информационных матриц в случае оценивания векторного параметра.

Информационное количество Фишера, содержащееся в выборке усеченных слева наблюдений, имеет вид

$$J_n(\theta)=n\int_D^\infty\left[\frac{\partial \ln f_{LT}(t;\theta)}{\partial\theta}\right]^2 f_{LT}(t;\theta)\,dt. \quad (10)$$

В табл. 2 для базового закона распределения Вейбулла найдены значения $J_n(\theta)/J_n^*(\theta)$ и значения отношения определителей соответствующих информационных матриц для случая векторного параметра в зависимости от величины усечения. При этом для удобства сравнения время усечения D выбиралось как квантиль базового распределения отказов $F_0^{-1}(d)$, $0 < d < 1$, где $F_0^{-1}(\cdot)$ – функция, обратная к $F_0(t)$. Величину d назовем глубиной усечения. В табл. 2 представлены также значения аналогичных отношений для цензурированных I и II типа выборок в зависимости от степени цензурирования.

Таблица 2 / Table 2

Отношение количества информации Фишера в усеченной слева или цензурированной справа выборке к количеству информации в полной выборке
Ratio of the Fisher information in truncated sample or in censored sample to the Fisher information in complete sample

Глубина усечения / Truncation depth	**О параметре θ_0 / About parameter θ_0**	**О параметре θ_1 / About parameter θ_1**	**О двух параметрах / About both parameters**
0,1	1,0000	0,7327	0,4948
0,2	1,0000	0,7733	0,3455
0,3	1,0000	0,8774	0,2574
0,4	1,0000	1,0204	0,1970
0,5	1,0000	1,1979	0,1521
0,6	1,0000	1,4144	0,1170
0,7	1,0000	1,6836	0,0882
0,8	1,0000	2,0400	0,0635
Степень цензурирования / Censoring degree	**О параметре θ_0 / About parameter θ_0**	**О параметре θ_1 / About parameter θ_1**	**О двух параметрах / About both parameters**
10 %	0,9000	0,7410	0,7133
20 %	0,8000	0,5920	0,5240
30 %	0,7000	0,4949	0,3791
40 %	0,6000	0,4343	0,2657
50 %	0,5000	0,4010	0,1771
60 %	0,4000	0,3878	0,1093
70 %	0,3000	0,3859	0,0595
80 %	0,2000	0,3814	0,0257

Как видно из табл. 2, в случае оценивания только параметра масштаба θ_0 распределения Вейбулла при известном значении параметра формы θ_1, точность ОМП не зависит от глубины усечения. Что интересно, для случая оценивания только параметра формы θ_1 при фиксированном параметре масштаба и глубине усечения $d > 0,4$ информационное количество Фишера $J_n(\theta_1)$ больше, чем $J_n^*(\theta_1)$, и с ростом d увеличивается. При одновременном оценивании двух параметров распределения Вейбулла определитель информационной матрицы для усеченного наблюдения, напротив, значительно меньше, чем для полного наблюдения, и уменьшается с ростом глубины усечения.

В отличие от усечения цензурирование выборки всегда приводит к потере информации Фишера. Однако в зависимости от степени цензурирования можно оценить минимально необходимый объем выборки, при котором дисперсия $J_n^{-1}(\theta)$ должна быть не выше заданной. Или, наоборот, по объему выборки можно оценить наибольшую степень цензурирования, при которой обеспечивается требуемая точность оценивания.

4. Проверка гипотез о согласии на основе выборок остатков

После оценивания неизвестных параметров предполагаемой вероятностной модели необходимо проверить адекватность полученной модели. Как правило, эту задачу сводят к проверке сложной гипотезы о виде вероятностной модели, которая может быть решена с использованием одного из критериев согласия. Для проверки такого рода гипотез по выборкам независимых одинаково распределенных случайных величин может использоваться целый ряд критериев согласия, например, критерии типа хи-квадрат, непараметрические критерии типа Колмогорова, Крамера–Мизеса–Смирнова, Андерсона–Дарлинга. Однако к выборкам вида (4) классические критерии согласия неприменимы, поскольку элементы данной выборки не являются одинаково распределенными.

Необходимым этапом решения задачи является формирование выборки остатков по усеченным слева и цензурированным справа выборкам. В частности, элементы выборки остатков для модели (6) могут быть получены в результате следующего преобразования:

$$R_i = \frac{S_{x^i}\left(X_i;\hat{\eta}\right)}{S_{x^i}\left(D_i;\hat{\eta}\right)},\ i=\overline{1,n}. \tag{11}$$

Если модель верна, то полученная цензурированная выборка остатков $\mathbf{R}_n = \{(R_1,\delta_1),\dots,(R_n,\delta_n)\}$ принадлежит равномерному на интервале (0, 1) распределению.

Для проверки гипотезы о принадлежности цензурированной выборки остатков равномерному распределению можно использовать модифицированные критерии согласия типа Колмогорова, Крамера–Мизеса–Смирнова и Андерсона–Дарлинга, в выражениях статистик которых вместо эмпирической функции распределения $F_n(t)$ используется оценка Каплана–Мейера [15] для функции распределения.

Областью определения случайных величин R_i, $i=\overline{1,n}$, является интервал $(0, 1)$. Обозначим через $0 < a_1 < a_2 < \dots < a_k = \tau < 1$, $k \le n$, неповторяющиеся значения полных наблюдений $(R_i, \delta_i = 1)$ в выборке остатков. Тогда оценку Каплана–Мейера можно вычислить по формуле

$$\hat{F}_n(t) = 1 - \prod_{a_i \le t}\left(1 - \frac{d_i}{r_i}\right), \tag{12}$$

где $d_i = \sum_{R_j = a_i} \delta_j$, r_i – количество наблюдений, для которых $R_j \ge a_i$, $j=\overline{1,n}$.

В модифицированном критерии Колмогорова для цензурированных выборок в качестве расстояния между оценкой Каплана–Мейера и равномерным законом распределения используется величина

$$D_n = \sup_{0 \le t \le \tau}\left|\hat{F}_n(t) - t\right|.$$

В модифицированном критерии Колмогорова будем использовать статистику с поправкой Большева

$$S_K^C = \frac{6nD_n + 1}{6\sqrt{n}}, \tag{13}$$

где $D_n = \max\left\{D_n^+, D_n^-\right\}$, $D_n^+ = \max\limits_{1\le i\le k}\left\{\hat{F}_n(a_i) - a_i\right\}$, $D_n^- = \max\limits_{1\le i\le k}\left\{a_i - \hat{F}_n\left(a_{i-1}\right)\right\}$.

В модифицированном критерии Крамера–Мизеса–Смирнова в качестве расстояния между распределениями используется величина

$$\omega^2 = \int_0^\tau \left(\hat{F}_n(t) - t\right)^2 dt .$$

Статистика модифицированного критерия Крамера–Мизеса–Смирнова с оценкой Каплана–Мейера принимает вид

$$S_\omega^C = \frac{n\cdot a_1}{3} + n\sum_{j=1}^{k-1}\left[\hat{F}_n^2(a_j)(a_{j+1}-a_j) - \hat{F}_n(a_j)\left(a_{j+1}^2 - a_j^2\right) + \frac{\left(a_{j+1}^3 - a_j^3\right)}{3}\right]. \quad (14)$$

В модифицированном критерии Андерсона-Дарлинга в качестве меры рассматривается величина

$$\Omega^2 = \int_0^\tau \left(\hat{F}_n(t) - t\right)^2 \frac{dt}{t(1-t)} .$$

Соответственно, статистика модифицированного критерия Андерсона–Дарлинга имеет вид

$$S_\Omega^C = n\left\{-a_1 + \sum_{j=1}^{k-1}\left[a_j - a_{j+1} + \hat{F}_n^2\left(a_j\right)\left(\ln a_{j+1} - \ln a_j\right) - \right.\right.$$

$$\left.\left. -\left(1-\hat{F}_n\left(a_j\right)\right)^2\left(\ln\left(1-a_{j+1}\right) - \ln\left(1-a_j\right)\right)\right] - \ln\left(1-a_1\right)\right\}. \quad (15)$$

Проверяемая гипотеза о согласии отвергается при больших значениях статистик. Необходимо отметить, что проверяемая гипотеза о равномерности остатков является сложной, поскольку при вычислении остатков по формуле (11) используются ОМП параметров модели. Аналитические выражения для распределений статистик рассматриваемых критериев в этом случае неизвестны. Более того, они зависят от множества факторов, определяющих «сложность» гипотезы, включая факторы, связанные с зависимостью от свойств оценок параметров, от типа и степени цензурирования, от наличия усеченных наблюдений, что полностью исключает возможность аналитического построения распределений статистик критериев. Поэтому вычисление критических значений статистик (или достигнутых уровней значимости), необходимых при проверке гипотезы о согласии с использованием данных критериев, возможно только с опорой на численные оценки распределений статистик, получаемые с использованием статистического моделирования.

5. Распределения статистик и мощность критериев согласия

Как сказано выше, проверяемая гипотеза о согласии с построенной вероятностной моделью является сложной, поскольку параметры модели оцениваются по той же выборке, по которой проверяется гипотеза. В этом случае на распределения статистик критериев типа Колмогорова, Крамера–Мизеса–Смирнова и Андерсона–Дарлинга влияют как свойства непараметрической оценки Каплана–Мейера, так и свойства ОМП параметров модели.

Если исходная выборка $\mathbf{X}_n$ является цензурированной, то и выборка остатков $\mathbf{R}_n$ также будет цензурированной, при этом одинаковые времена цензурирования в исходной выборке (при I или II типах цензурирования) после преобразования (11) будут различными. В [16] было показано, что свойства оценки Каплана–Мейера зависят от типа и степени цензурирования, а также от распределения моментов цензурирования в случае III типа.

Оценка Каплана–Мейера вычисляется по выборке остатков, тогда как ОМП параметров модели вычисляются по исходной выборке $\mathbf{X}_n$, содержащей в общем случае как полные, так и усеченные слева и цензурированные справа наблюдения. Как показано в разделе 3, на свойства ОМП параметров вероятностных моделей существенное влияние оказывают вид базового закона распределения отказов, тип и количество оцениваемых параметров базового закона, а также степень цензурирования и глубина усечения выборки. Кроме этого для некоторых законов распределения (например, для гамма-распределения) свойства ОМП зависят и от конкретного значения параметра формы [17].

Таким образом, алгоритм моделирования распределений статистик, учитывающий влияние указанных факторов, можно сформулировать следующим образом.

1. На основе построенной вероятностной модели надежности, с которой проверяется гипотеза о согласии, сгенерировать выборку $\mathbf{Y}_n$, аналогичную исходной выборке $\mathbf{X}_n$, по заданному плану эксперимента (при заданных значениях вектора объясняющих переменных x и количествах объектов в группах, имеющих разные значения x).

1.1. Если $\mathbf{X}_n$ представляет собой полную выборку вида (1), то элементы выборки $\mathbf{Y}_n$ моделируются в соответствии с выражением $T_i = F_{x^i}^{-1}(\xi_i)$, где $\xi_i \succ \text{Uniform}(0,1)$, $i = 1,...,n$.

1.2. Если $\mathbf{X}_n$ представляет собой цензурированную выборку вида (2), то сначала моделируется полная выборка (см. п. 1.1), а затем полученная выборка преобразуется в соответствии с типом цензурирования выборки $\mathbf{X}_n$.

1.3. Если $\mathbf{X}_n$ представляет собой усеченную слева выборку вида (3), то для формирования i-го элемента выборки $\mathbf{Y}_n$, $i = \overline{1,n}$, генерируется последовательность величин $\xi_j \succ \text{Uniform}(0,1)$, $j = 1,2,...$, пока не выполнится условие $T_i = F_{x^i}^{-1}(\xi_j) \ge D_i$, где D_i – время усечения i-го наблюдения выборки $\mathbf{X}_n$.

1.4. Если $\mathbf{X}_n$ представляет собой выборку вида (4), усеченную слева и цензурированную справа, то сначала моделируется усеченная выборка в соответствии с п. 1.3, а затем полученная выборка преобразуется в цензурированную (в соответствии с типом цензурирования выборки $\mathbf{X}_n$).

2. По полученной выборке $\mathbf{Y}_n$ оценить параметры модели методом максимального правдоподобия.

3. Сформировать выборку остатков $\mathbf{R}_n$.

4. По выборке $\mathbf{R}_n$ вычислить значение статистики соответствующего критерия (статистики S_K^C, S_ω^C или S_Ω^C).

5. Повторить пункты 1–4 N раз. По полученной выборке статистик объема N построить эмпирическую функцию распределения статистики критерия $G_N(S \mid H_0)$.

По построенному в соответствии с данным алгоритмом распределению статистики $G_N(S \mid H_0)$ вычисляется оценка достигнутого уровня значимости $\alpha_n = 1 - G_N(S_n \mid H_0)$, где S_n – значение соответствующей статистики, полученное по исходной выборке $\mathbf{X}_n$. Если α_n не превышает заданного уровня значимости α, то гипотеза о согласии отвергается.

В качестве примера исследуем мощность рассматриваемых модификаций критериев согласия относительно следующей пары конкурирующих гипотез:

H_0: модель пропорциональных интенсивностей Кокса (7) с базовым распределением Вейбулла;

H_1: модель Ксая (8) с тем же базовым распределением.

Конкурирующая гипотеза H_1 представляет собой обобщение модели пропорциональных интенсивностей Кокса, при котором функции надежности для разных значений воздействий пересекаются. Таким образом, данное исследование позволит выявить чувствительность рассматриваемых критериев согласия к изменению вида регрессионной зависимости.

Распределения статистик критериев согласия при справедливости гипотезы H_0 моделировались на основе модели Кокса со скалярной объясняющей переменной $x \in \{0, 1\}$ и параметрами $\theta_0 = 22,0$, $\theta_1 = 1,1$, $\beta = 0,2$. Количество наблюдений с одинаковым значением ковариаты – $n_i = 100$ для $x = 0$ и $x = 1$, $i = \{1, 2\}$. При справедливой конкурирующей гипотезе H_1 выборки моделировались в соответствии с моделью Ксая с обобщающим параметром $\gamma = 0,5$ при тех же прочих условиях.

В табл. 3 представлены оценки мощности рассматриваемых критериев, полученные по выборкам вида (2) с цензурированием II типа при $\alpha = 0,1$, объем моделирования $N = 10000$.

Таблица 3 / Table 3

Оценки мощности критериев согласия для первой пары гипотез
Estimates of the power of goodness-of-fit tests for the first pair of hypotheses

Степень цензурирования / Censoring degree	Критерий согласия / Goodness-of-fit test		
	Колмогорова / Kolmogorov	Крамера–Мизеса-Смирнова / Cramer–von Mises–Smirnov	Андерсона–Дарлинга / Anderson–Darling
0 %	0,19	0,22	0,23
10 %	0,16	0,19	0,21
30 %	0,13	0,14	0,16
50 %	0,13	0,13	0,14

Из табл. 3 видно, что непараметрические критерии согласия типа Колмогорова, Крамера–Мизеса–Смирнова и Андерсона–Дарлинга способны различать близкие гипотезы, соответствующие моделям с разными видами регрессионной зависимости. Наибольшей мощностью для всех рассмотренных степеней цензурирования обладает критерий типа Андерсона–Дарлинга.

Для того чтобы исследовать чувствительность критериев к изменению вида базового закона распределения, сравним мощности критериев согласия для следующей пары близких конкурирующих гипотез:

H_0: модель пропорциональных интенсивностей Кокса (7) с базовым распределением Вейбулла;

H_1: модель пропорциональных интенсивностей Кокса (7) с базовым гамма-распределением.

В данном случае при справедливости гипотезы H_0 распределения статистик критериев согласия моделировались на основе модели Кокса со скалярной объясняющей переменной $x \in \{0,0;\ 0,5;\ 1,0;\ 1,5;\ 2,0\}$ и значениями параметров $\theta_0 = 2$, $\theta_1 = 2$, $\beta = 2$. Количество наблюдений с одинаковым значением ковариаты – $n_i = 40$, $i = \overline{1,5}$. При справедливости конкурирующей гипотезы H_1 выборки моделировались в соответствии с базовым гамма-распределением при значениях параметров $\theta_0 = 0,5577$, $\theta_1 = 3,1215$, $\beta = 2$ по тому же плану эксперимента.

В табл. 4 представлены оценки мощности рассматриваемых критериев согласия в случае усеченных слева и цензурированных справа выборок вида (4) при различной величине глубины усечения d и степени цензурирования. При этом времена усечения $D_i = F^{-1}_{x^i}(d)$.

Таблица 4 / Table 4

Оценки мощности критериев согласия для второй пары гипотез
Estimates of the power of goodness-of-fit tests for the second pair of hypotheses

Глубина усечения / Truncation depth	Степень цензурирования / Censoring degree	Критерий согласия / Goodness-of-fit test		
		Колмогорова / Kolmogorov	Крамера–Мизеса–Смирнова / Cramer–von Mises–Smirnov	Андерсона–Дарлинга / Anderson–Darling
0,0	0 %	0,41	0,49	0,56
0,0	10 %	0,19	0,21	0,29
0,0	30 %	0,20	0,18	0,18
0,0	50 %	0,14	0,14	0,13
0,1	0 %	0,25	0,31	0,34
0,1	10 %	0,19	0,22	0,25
0,1	30 %	0,17	0,17	0,18
0,1	50 %	0,15	0,15	0,14
0,3	0 %	0,20	0,23	0,25
0,3	10 %	0,16	0,18	0,19
0,3	30 %	0,15	0,15	0,15
0,3	50 %	0,12	0,12	0,11
0,5	0 %	0,16	0,18	0,19
0,5	10 %	0,14	0,15	0,17
0,5	30 %	0,13	0,13	0,12
0,5	50 %	0,11	0,10	0,10

Как видно из табл. 4, с увеличением глубины усечения и ростом степени цензурирования мощность критериев согласия падает. Однако потери в мощности критериев для усеченных слева выборок оказались меньше, чем аналогичные потери для цензурированных выборок на данной паре конкурирующих гипотез. Также стоит отметить, что при малых степенях цензурирования большей мощностью обладает критерий типа Андерсона–Дарлинга, тогда как с ростом степени цензурирования мощность всех рассматриваемых критериев приближается к заданному уровню значимости $\alpha = 0,1$.

Заключение

В настоящей работе рассмотрена проблема проверки адекватности вероятностных моделей надежности с учетом объясняющих переменных по усеченным слева и цензурированным справа выборкам. Проведено исследование точности ОМП параметров распределения Вейбулла по усеченным слева выборкам. Показано, что в случае оценивания параметра масштаба при известном значении параметра формы точность ОМП не зависит от глубины усечения; в случае оценивания параметра формы при фиксированном параметре масштаба и глубине усечения $d > 0,4$ дисперсия ОМП по усеченной выборке меньше, чем по полной выборке отказов; при одновременном оценивании двух параметров определитель ковариационной матрицы для усеченной выборки, напротив, значительно больше, чем для полной, и увеличивается с ростом глубины усечения.

Предложен универсальный подход к проверке адекватности построенных вероятностных моделей надежности, основанный на применении модификаций критериев согласия типа Колмогорова, Крамера–Мизеса–Смирнова и Андерсона-Дарлинга для проверки сложных гипотез о принадлежности выборок остатков, вычисленных в соответствии с проверяемой моделью, равномерному на $(0, 1)$ распределению. Предложен и реализован алгоритм статистического моделирования неизвестных распределений статистик критериев при справедливости проверяемой гипотезы, позволяющий в итоге вычислять достигнутый уровень значимости и принимать корректное решение о результатах проверки. Показано, что предложенный подход позволяет проверять как предположения относительно вида регрессионной зависимости, так и предположения о виде базового закона распределения отказов.

ЛИТЕРАТУРА

1. **Meeker W.Q., Escobar L.A.** Statistical methods for reliability data. – New York: John Wiley & Sons, 1998. – 680 p.
2. **Cox D.R.** Regression models and life tables // Journal of the Royal Statistical Society. Series B (Methodological). – 1972. – Vol. 34, iss. 2. – P. 187–220.
3. **Bagdonavicius V., Nikulin M.** Accelerated life models: modeling and statistical analysis. – Boca Raton, Florida: Chapman & Hall/CRC, 2002. – 334 p.
4. **Beran R.** Nonparametic regression with randomly censored survival data: technical report. – Berkeley: University of California, Department of Statistics, 1981.
5. **Демин В.А., Чимитова Е.В.** Выбор оптимального параметра сглаживания для непараметрической оценки регрессионной модели надежности // Вестник Томского государственного университета. Управление, вычислительная техника и информатика. – 2013. – № 1 (22) . – С. 59–65.
6. **Breslow N.E.** Analysis of survival data under the proportional hazards model // International Statistical Review. – 1975. – Vol. 43, N 1. – P. 45–57.
7. **Lawless J.F.** Statistical models and methods for lifetime data. – Hoboken, New Jersey: John Wiley & Sons, 2003. – xx, 630 p.
8. **Чимитова Е.В., Ведерникова М.А., Галанова Н.С.** Непараметрические критерии согласия в задачах проверки адекватности моделей надежности по цензурированным данным // Вестник Томского государственного университета. Управление, вычислительная техника и информатика. – 2013. – № 4 (25). – С. 115–124.
9. **Balakrishnan N., Mitra D.** Left truncated and right censored Weibull data and likelihood inference with an illustration // Computational Statistics and Data Analysis. – 2012. – Vol. 56, iss. 12. – P. 4011–4025. – doi: 10.1016/j.csda.2012.05.004.
10. **Balakrishnan N., Mitra D.** Likelihood inference for lognormal data with left truncation and right censoring with an illustration // Journal of Statistical Planning and Inference. – 2011. – Vol. 141, iss. 11. – P. 3536–3553. – doi: 10.1016/j.jspi.2011.05.007.
11. **Balakrishnan N., Mitra D.** Likelihood inference based on left truncated and right censored data from a gamma distribution // IEEE Transactions on Reliability. – 2013. – Vol. 62, iss. 3. – P. 679–688. – doi: 10.1109/TR.2013.2273039.

12. **Bagdonavičius V., Levuliené R., Nikulin M.S.** Goodness-of-fit criteria for the Cox model from left truncated and right censored data // Journal of Mathematical Sciences. – 2010. – Vol. 167, iss. 4. – P. 436–443. – doi: 10.1007/s10958-010-9929-6.
13. **Hsieh F.** On heteroscedastic hazards regression models: theory and application // Journal of the Royal Statistical Society. Series B (Statistical Methodology). – 2001. – Vol. 63, iss. 1. – P. 63–79. – doi: 10.1111/1467-9868.00276.
14. **Лемешко Б.Ю., Гильдебрант С.Я., Постовалов С.Н.** К оцениванию параметров надежности по цензурированным выборкам // Заводская лаборатория. Диагностика материалов. – 2001. – Т. 67, № 1. – С. 52–64.
15. Testing goodness-of-fit of parametric AFT and PH models with residuals / N. Balakrishnan, E.V. Chimitova, N.S. Galanova, M.A. Vedernikova // Communications in Statistics – Simulation and Computation. – 2013. – Vol. 42, iss. 6. – P. 1352–1367. – doi: 10.1080/03610918. 2012.659824.
16. **Лемешко Б.Ю., Чимитова Е.В., Ведерникова М.А.** Модифицированные критерии согласия Колмогорова, Крамера–Мизеса–Смирнова и Андерсона–Дарлинга для случайно цензурированных выборок. Ч. 2 // Научный вестник НГТУ. – 2013. – № 1 (50). – С. 3–16.
17. Статистический анализ данных, моделирование и исследование вероятностных закономерностей. Компьютерный подход / Б.Ю. Лемешко, С.Б. Лемешко, С.Н. Постовалов, Е.В. Чимитова. – Новосибирск: Изд-во НГТУ, 2011. – 887 с. – (Монографии НГТУ).

TESTING GOODNESS-OF-FIT OF PARAMETRIC RELIABILITY REGRESSION MODELS WITH LEFT TRUNCATED AND RIGHT CENSORED DATA

E.V. Chimitova, M.A Semenova.
Novosibirsk State Technical University, Novosibirsk, Russian Federation

The paper addresses basic issues of constructing probabilistic reliability models with explanatory variables for left truncated and right censored samples. The study of the accuracy of the maximum likelihood estimates of the Weibull distribution parameters has been carried out for left truncated samples. It has been shown that the determinant of the covariance matrix for the truncated sample is significantly larger than the determinant for a complete sample. Moreover, the determinant of the covariance matrix for the truncated sample increases as the depth of truncation grows. A universal approach to testing the goodness-of-fit hypothesis for the probabilistic reliability model by left truncated and right censored samples is proposed. This approach is based on using modified goodness-of-fit criteria for testing the hypothesis of the uniformity of residual samples corresponding to the model tested. In the paper, the Kolmogorov type, Cramer–von Mises–Smirnov type and Anderson–Darling type goodness-of-fit tests are considered. The proposed algorithm of statistical simulation of unknown distributions of test statistics makes it possible to use these tests in the case of left truncated and right censored samples correctly. The study of the tests power has shown that the approach based on residual samples can be applied for testing the assumption of the regression dependence kind as well as for testing the choice of baseline distribution.

Keywords: censored sample, truncated sample, reliability model, proportional hazards model, accelerated failure time model, goodness-of-fit tests.

DOI: 10.17212/1727-2769-2015-1-104-120

REFERENCES

1. Meeker W.Q., Escobar L.A. *Statistical methods for reliability data.* New York, John Wiley & Sons, 1998. 680 p.
2. Cox D.R. Regression models and life tables. *Journal of the Royal Statistical Society. Series B (Methodological)*, 1972, vol. 34, iss. 2, pp. 187–220.
3. Bagdonavicius V., Nikulin M. *Accelerated life models: modeling and statistical analysis*. Boca Raton, Florida, Chapman & Hall/CRC, 2002. 334 p.
4. Beran R. *Nonparametric regression with randomly censored survival data*. Technical report. Berkeley, University of California, Department of Statistics, 1981.
5. Demin V.A., Chimitova E.V. Vybor optimal'nogo parametra sglazhivaniya dlya neparametricheskoi otsenki regressionnoi modeli nadezhnosti [Choice of optimal smoothing parame-

ter for nonparametric estimation of regression reliability model]. *Vestnik Tomskogo gosudarstvennogo universiteta. Upravlenie, vychislitel'naya tekhnika i informatika – Tomsk State University Journal of Control and Computer Science,* 2013, no. 1 (22), pp. 59–65.
6. Breslow N.E. Analysis of survival data under the proportional hazards model. *International Statistical Review*, 1975, vol. 43, no. 1, pp. 45–57.
7. Lawless J.F. *Statistical models and methods for lifetime data*. Hoboken, New Jersey, John Wiley & Sons, 2003. xx, 630 p.
8. Chimitova E.V., Vedernikova M.A., Galanova N.S. Neparametricheskie kriterii soglasiya v zadachakh proverki adekvatnosti modelei nadezhnosti po tsenzurirovannym dannym [Nonparametric goodness-of-fit tests in testing adequacy of reliability models for right censored data]. *Vestnik Tomskogo gosudarstvennogo universiteta. Upravlenie, vychislitel'naya tekhnika i informatika – Tomsk State University Journal of Control and Computer Science*, 2013, no. 4 (25), pp. 115–124.
9. Balakrishnan N., Mitra D. Left truncated and right censored Weibull data and likelihood inference with an illustration. *Computational Statistics and Data Analysis*, 2012, vol. 56, iss. 12, pp. 4011–4025. doi: 10.1016/j.csda.2012.05.004
10. Balakrishnan N., Mitra D. Likelihood inference for lognormal data with left truncation and right censoring with an illustration. *Journal of Statistical Planning and Inference*, 2011, vol. 141, iss. 11, pp. 3536–3553. doi: 10.1016/j.jspi.2011.05.007
11. Balakrishnan N., Mitra D. Likelihood inference based on left truncated and right censored data from a gamma distribution. *IEEE Transactions on Reliability*, 2013, vol. 62, iss. 3, pp. 679–688. doi: 10.1109/TR.2013.2273039
12. Bagdonavičius V., Levuliené R., Nikulin M.S. Goodness-of-fit criteria for the Cox model from left truncated and right censored data. *Journal of Mathematical Sciences*, 2010, vol. 167, iss. 4, pp. 436–443. doi: 10.1007/s10958-010-9929-6
13. Hsieh F. On heteroscedastic hazards regression models: theory and application. *Journal of the Royal Statistical Society. Series B (Statistical Methodology)*, 2001, vol. 63, iss. 1, pp. 63–79. doi: 10.1111/1467-9868.00276
14. Lemeshko B.Yu., Gil'debrant S.Ya., Postovalov S.N. K otsenivaniyu parametrov nadezhnosti po tsenzurirovannym vyborkam [About the estimation of reliability parameters by censored samples]. *Zavodskaya laboratoriya. Diagnostika materialov – Industrial laboratory. Materials diagnostics*, 2001, vol. 67, no. 1, pp. 52–64.
15. Balakrishnan N., Chimitova E., Galanova N., Vedernikova M. Testing goodness-of-fit of parametric AFT and PH models with residuals. *Communications in Statistics – Simulation and Computation*, 2013, vol. 42, iss. 6, pp. 1352–1367. doi: 10.1080/03610918.2012.659824
16. Lemeshko B.Yu., Chimitova E.V., Vedernikova M.A. Modifitsirovannye kriterii sogla-siya Kolmogorova, Kramera–Mizesa–Smirnova i Andersona–Darlinga dlya sluchaino tsenzurirovannykh vyborok. Ch. 2 [Modified goodness-of-fit tests of Kolmogorov, Cramer–von Mises–Smirnov and Anderson–Darling for randomly censored samples]. *Nauchnyi vestnik NGTU – Science bulletin of Novosibirsk state technical university*, 2013, no. 1 (50), pp. 3–16.
17. Lemeshko B.Yu., Lemeshko S.B., Postovalov S.N., Chimitova E.V. *Statisticheskii analiz dannykh, modelirovanie i issledovanie veroyatnostnykh zakono-mernostei. Komp'yuternyi podkhod* [Statistical data analysis, simulation and study of probability regularities. Computer approach]. Novosibirsk, NSTU Publ., 2011. 887 p.

СВЕДЕНИЯ ОБ АВТОРАХ

Чимитова Екатерина Владимировна – родилась в 1977 году, канд. техн. наук, доцент, доцент кафедры теоретической и прикладной информатики Новосибирского государственного технического университета. Область научных интересов: статистические методы анализа данных типа времени жизни. Опубликовано 80 научных работ. (Адрес: 630073, Россия, г. Новосибирск, пр. Карла Маркса, 20. E-mail: chimitova@corp.nstu.ru).

Chimitova Ekaterina Vladimirovna (b. 1977) – PhD (Eng.), associate professor at the Theoretical and Applied Informatics Department of the Novosibirsk State Technical University. Her research interests are currently focused on statistical methods of lifetime data analysis. She is the author of 80 scientific papers (Address: 20, Karl Marx Av., Novosibirsk, 630073, Russian Federation. E-mail: chimitova@corp.nstu.ru).

Семёнова Мария Александровна – родилась в 1989 году, аспирант кафедры теоретической и прикладной информатики Новосибирского государственного технического университета. Область научных интересов: проверка статистических гипотез при построении моделей выживаемости по цензурированным выборкам. Опубликовано 10 научных работ. (Адрес: 630073, Россия, г. Новосибирск, пр. Карла Маркса, 20. E-mail: vedernikova.m.a@gmail.com).

Semenova Maria Alexandrovna (b. 1989) – a postgraduate student at the Theoretical and Applied Informatics Department in the Novosibirsk State Technical University. Her research interests are currently focused on statistical tests for survival models by censored samples. She is the author of 10 scientific papers. (Address: 20, Karl Marx Av., Novosibirsk, 630073, Russian Federation. E-mail: vedernikova.m.a@gmail.com)

Статья поступила 10 декабря 2014 г.
Received on December 10, 2014

To Reference:

Chimitova E.V., Semenova M.A. Proverka adekvatnosti parametricheskikh regressionnykh modelei nadezhnosti po usechennym sleva i tsenzurirovannym sprava dannym [Testing goodness-of-fit of parametric reliability regression models with left truncated and right censored data]. *Doklady Akademii nauk vysshei shkoly Rossiiskoi Federatsii – Proceedings of the Russian higher school Academy of sciences*, 2015, no. 1 (26), pp. 104–120. doi: 10.17212/1727-2769-2015-1-104-120