

Особенности применения регрессии Кокса в различных инструментальных средах

Крамаренко Инна Владимировна

Канд. экон. наук, доц. каф. математических методов в экономике и управлении
ORCID: 0000-0003-0861-8772, e-mail: i.v.kramarenko@kiber-guu.ru

Константинова Любовь Андреевна

Канд. экон. наук, доц. каф. математических методов в экономике и управлении
ORCID: 0000-0003-0228-4011, e-mail: konstantinova11@mail.ru

Государственный университет управления, г. Москва, Россия

Аннотация

Наличие больших объемов данных в информационно-аналитических системах порождает необходимость их изучения с использованием методов машинного обучения и искусственного интеллекта. Эти модели требуют определения настроечных параметров, связанных с особенностями предметной области. В статье представлена модель регрессии Кокса для решения проблемы оттока клиентов. Регрессия Кокса признана моделью с высокой точностью предсказаний в здравоохранении. Поэтому интересным представляется применение модели в иных отраслях. В работе приведены результаты и сравнительный анализ расчетов на модели Кокса с использованием трех инструментов: статистического пакета Statistical Package for the Social Sciences, языка программирования R и российского программного обеспечения – аналитической платформы Logipom. Отличительной особенностью разработанной вероятностной модели является определение риска наступления события в условиях неполноты данных, а также выделение показателей, оказывающих существенное влияние на степень его проявления.

Ключевые слова

Регрессия Кокса, риск, управление лояльностью клиентов, аналитическая платформа, статистический пакет

Благодарности. Авторы выражают благодарность и глубокую признательность генеральному директору ЗАО «Аналитические технологии» Алексею Ивановичу Арустамову за предоставление программного обеспечения в рамках сотрудничества с вузами, Николаю Борисовичу Паклину за рекомендации и наставления при разработке компонентов, рецензентам за советы и ценные замечания в процессе работы над статьей.

Для цитирования: Крамаренко И.В., Константинова Л.А. Особенности применения регрессии Кокса в различных инструментальных средах // Вестник университета. 2022. № 10. С. 80–88.



Features of using Cox regression in various instrumental environments

Inna V. Kramarenko

Cand. Sci. (Econ.), Assoc. Prof. at the Mathematical Methods in Economics and Management Department
ORCID: 0000-0003-0861-8772, e-mail: i.v.kramarenko@kiber-guu.ru

Lyubov A. Konstantinova

Cand. Sci. (Econ.), Assoc. Prof. at the Mathematical Methods in Economics and Management Department
ORCID: 0000-0003-0228-4011, e-mail: konstantinova11@mail.ru

State University of Management, Moscow, Russia

Abstract

The presence of large amounts of data in information and analytical systems makes it necessary to study them using machine learning and artificial intelligence methods. These models require the definition of tuning parameters related to the specifics of the subject area. The article presents a Cox regression model to solve the problem of customer churn. Cox regression is recognized as a model with high accuracy of predictions in health-care. Therefore, it is interesting to use the model in other industries. The paper presents the results and comparative analysis of calculations on the Cox model using three tools: Statistical Package for the Social Sciences, programming language R and Russian software – analytical platform Loginom. A distinctive feature of the developed probabilistic model is the determination of the risk of event occurrence in conditions of incomplete data, as well as the identification of indicators that have a significant impact on the degree of its manifestation.

Keywords

Cox regression, risk, customer loyalty management, analytical platform, statistical package

Acknowledgements. The authors express gratitude and deep appreciation to Alexey Arustamov, General Director of CJSC Analytical Technologies, for providing software in the course of cooperation with universities, to Nikolay Paklin for recommendations and instructions in the development of components and reviewers for advice and valuable comments when working on the article.

For citation: Kramarenko I.V., Konstantinova L.A. (2022) Features of using Cox regression in various instrumental environments. *Vestnik universiteta*, no. 10, pp. 80–88.

ВВЕДЕНИЕ

Существует исторически сложившееся мнение, что регрессионная модель Кокса применима исключительно в биологических и медицинских испытаниях [1]. Однако современные исследования доказывают, что модель пропорциональных рисков может успешно применяться в банковской сфере [2], в страховании [3], в управлении производством и других сферах экономики и управления [4].

Теоретическая и методологическая основа исследования базируется на научных трудах российских и зарубежных авторов, посвященных особенностям построения и реализации регрессии Кокса в различных сферах жизнедеятельности общества, моделированию оттока клиентов с помощью различных методов, а также методологии проведения анализа выживаемости в различных программных средах [5–7]. В качестве базовой модели исследования используется классическая модель пропорциональных рисков Кокса, представленная в работе под названием «Регрессионные модели Дэвида Кокса и таблицы смертности»

© Kramarenko I.V., Konstantinova L.A., 2022.

This is an open access article under the CC BY 4.0 license (<http://creativecommons.org/licenses/by/4.0/>).



(англ. “Cox David Regression models and life-tables”) в 1972 г., и ее модификации, которые получили широкое распространение после 1986 г. Исследование также опирается на более поздние работы Д.Г. Кляйнбаума, посвященные анализу выживаемости.

Анализ выживаемости – класс статистических моделей, характеризующий взаимоотношения между конечными точками времени до наступления события и независимыми переменными (ковариатами). Последующее изучение этой проблемы привело к формированию нового статистического метода, названного моделью пропорциональных рисков Кокса (англ. proportional hazard model) или регрессией Кокса (Cox regression) [1; 8]. Им были сформулированы четыре основных стимула, на основании которых произошло формирование регрессии Кокса как самостоятельного инструмента анализа и прогнозирования [9].

Новая функция правдоподобия соответствует четко определенной, хотя и довольно сложной серии условных операций, названной частичным правдоподобием (англ. partial likelihood). Д. Кокс считал, что ее доказательство привело бы к теореме с запутанными условиями регулярности, не представляющими научного интереса, однако последующие работы других специалистов достигли поразительно элегантных результатов. Удивительно быстро многие производители статистических пакетов включили метод в свои библиотеки, и именно с этого момента он начал активно использоваться в исследованиях [6; 10–12].

На сегодняшний день модель пропорциональных рисков Кокса является одной из самых известных публикаций по статистике. По данным Google Scholar (Google Академия), по состоянию на март 2022 г. она получила более 2 400 цитирований.

Существует множество универсальных инструментов для реализации модели Кокса, однако их тонкая настройка невозможна единообразном подходе [5–7; 13; 14].

ТЕОРИЯ И МЕТОДЫ

Цель исследования состоит в разработке и программной реализации модели пропорциональных рисков (регрессии Кокса) с использованием различных инструментальных сред. Выдвигается гипотеза о том, что не существует универсального инструмента для тонкой настройки модели. Для реализации модели, связанной с конкретной предметной областью, необходимо иметь инструменты, которые позволяют доработать готовые универсальные решения.

Выделим две основные задачи исследования.

1. Проведение сравнительного анализа моделей и методов для моделирования и прогнозирования оттока клиентов.

2. Построение регрессии Кокса с помощью статистического пакета Statistical Package for the Social Sciences (далее – SPSS), языка программирования R и реализации компонента «Регрессия Кокса» в аналитической платформе Logiном.

В таблице 1 представлена сравнительная характеристика моделей и методов, с помощью которых возможен анализ оттока клиентов.

Таблица 1

Сравнительный анализ методов, моделирующих отток клиентов

Метод	Достоинства	Недостатки
Регрессия Кокса	<ul style="list-style-type: none"> – учитывает время жизни; – способна работать в условиях неполноты данных и наличия категориальных переменных; – возможность графической интерпретации 	<ul style="list-style-type: none"> – несостоятельность результатов в связи с наличием цензурированных данных
Логистическая (бинарная) регрессия	<ul style="list-style-type: none"> – имеет широкую сферу применения; – проста построения; – простота интерпретации параметров; – высокая точность; – функционал содержит оценку качества 	<ul style="list-style-type: none"> – предсказываемый параметр, является непрерывным числом; – требует большого количества данных

Метод	Достоинства	Недостатки
Методы кластеризации (K-средних)	<ul style="list-style-type: none"> – прост в построении; – хорошо интерпретирует результаты 	<ul style="list-style-type: none"> – плохо работает с дискретными данными; – результат зависит от выбора исходных центров кластеров; – высокая сложность одного прогноза
Деревья решений	<ul style="list-style-type: none"> – эффективны при работе с данными с большим количеством факторов и классов; – работают с дискретными и непрерывными признаками; – нечувствительны к монотонным преобразованиям 	<ul style="list-style-type: none"> – большая размерность моделей; – склонны к переобучению
Нейронные сети	<ul style="list-style-type: none"> – возможность использования при неидентифицируемых закономерностях; – устойчивость к шумам; – самообучаемость; – отказоустойчивость 	<ul style="list-style-type: none"> – неинтерпретируемость результатов в связи с отсутствием логики причин принятого решения; – отсутствие гарантий получения однозначных повторяемых результатов

Составлено авторами по материалам источника [9]

Каждый из рассмотренных методов имеет свои достоинства и недостатки. Отметим, что нейронная сеть, несмотря на ряд достоинств, таких как адаптация к изменениям среды и устойчивость к шумам, не всегда дает качественные результаты для прогнозирования оттока клиентов. Деревья решений эффективны в задачах классификации и кластеризации, что важно при моделировании оттока клиентов, однако модели имеют большую размерность данных [2].

Для построения регрессии Кокса, моделирующей отток клиентов, воспользуемся выборкой из набора данных клиентской базы от анонимного оператора, предоставленной платформой IBM Developer [15]. Этот массив данных содержит информацию за 3 года в общей сложности о 4 042 клиентах и 11 атрибутах, основанных на личных характеристиках пользователей, предоставленных услугах и деталях контрактов. В качестве индикатора для этой оценки лояльности будет использоваться переменная Churn [12].

Количество активных клиентов (2 351) превышает количество абонентов (1 511), отказавшихся от услуг компании, в 1,7 раз.

В качестве рассматриваемых переменных модели Кокса для телекоммуникационной отрасли могут выступать [12]:

- *CustomerID* – идентификационный номер клиента;
- *Gender* – пол (male – мужской, female – женский);
- *SeniorCitizen* – индикатор принадлежности клиента к группе пенсионного возраста (yes – да, no – нет);
- *Partner* – индикатор семейного положения (yes – есть партнер, no – нет партнера);
- *Dependents* – индикатор наличия иждивенцев (yes – да, no – нет);
- *Tenure* – время жизни клиента (в месяцах);
- *PhoneService* – использование абонентом услуг телефонной связи (yes – да, no – нет);
- *InternetService* – предоставление пользователю доступа к сети «Интернет» (yes – да, no – нет);
- *Contract* – срок действия контракта с пользователем (month – месяц, one year – один год, two years – два года);
- *MonthlyCharges* – ежемесячная сумма, взимаемая с клиента (в долларах США);
- *Churn* – индикатор отказа клиента от услуг компании (yes – да, no – нет).

Поясним трактовку нескольких переменных. В исследовании под событием рассматривается уход клиента из телекоммуникационной компании (Churn). Период наблюдения (Tenure) – количество месяцев от момента, когда клиент начал пользоваться услугами представленной фирмы. Группа риска – активные клиенты компании. Независимые переменные, влияющие на отток, – демографические показатели (пол (Gender), возраст (SeniorCitizen), семейное положение (Partner), наличие иждивенцев (Dependents), услуги, используемые клиентами (телефонная связь (PhoneService), сеть «Интернет» (InternetService), информация об аккаунте пользователя (срок действия контракта (Contract), ежемесячная оплата (MonthlyCharges).

Построим классическую модель Кокса, представленную формулой:

$$h(T) = h_0(T) e^{\sum_{i=1}^p x_i \beta_i}, \quad (1)$$

где x_1, x_2, \dots, x_p – независимые переменные (ковариаты); $\beta_1, \beta_2, \dots, \beta_p$ – коэффициенты регрессии, которые необходимо оценить; T – рассматриваемый период времени; $h_0(T)$ – базовый уровень риска, когда все ковариаты равны нулю.

В таблице 2 представлен основной отчет SPSS [16] по построенной регрессии Кокса. Это базовые характеристики для ковариат при их добавлении в модель на каждом шаге.

Таблица 2

**Характеристика переменных в регрессии Кокса в статистическом пакете
Statistical Package for the Social Sciences**

Этапы	Переменные модели Кокса для телекоммуникационной отрасли	Значение коэффициента регрессии B	Стандартная ошибка	Критерий Вальда	Степени свободы	Значимость	Спрогнозированное значение риска при изменении независимой переменной на единицу $Exp(B)$	Доверительный интервал 95 % для $Exp(B)$	
								Нижняя	Верхняя
Шаг 1	<i>InternetService</i>	-1,325	0,101	172,894	1	0,000	0,266	0,218	0,324
Шаг 2	<i>InternetService</i>	-1,056	0,101	109,474	1	0,000	0,348	0,286	0,424
	<i>Contact</i>	3,864	0,708	29,740	1	0,000	47,633	11,881	190,962
Шаг 3	<i>Partner</i>	0,237	0,28	70,277	1	0,000	1,267	1,199	1,339
	<i>InternetService</i>	-1,066	0,101	111,667	1	0,000	0,344	0,283	0,420
	<i>Contact</i>	3,804	0,708	28,834	1	0,000	44,881	11,196	179,919
Шаг 4	<i>Partner</i>	0,183	0,031	35,737	1	0,000	1,201	1,131	1,276
	<i>Dependents</i>	0,312	0,076	16,907	1	0,000	1,367	1,178	1,586
	<i>InternetService</i>	-1,033	0,101	104,163	1	0,000	0,356	0,292	0,434
	<i>Contact</i>	3,743	0,709	27,896	1	0,000	42,206	10,525	169,252

Составлено авторами по материалам исследования

Таблица 2 показывает, что основные характеристики предикторов содержат значение коэффициента регрессии для каждой вводимой переменной, стандартную ошибку, критерий Вальда (аналог статистики для модели пропорциональных рисков) и его значимость, а также спрогнозированное значение риска при изменении независимой переменной на единицу. Так как значимость всех ковариат, включенных в модель, меньше принятого уровня значимости, то эти факторы являются значимыми и могут обуславливать риск ухода клиента.

На основании показателя можно сделать следующие выводы.

1. Риск ухода клиентов, не имеющих партнера, в 1,201 раз выше, чем для тех, кто состоит в отношениях.
2. Клиенты, имеющие на попечении иждивенцев, имеют склонность к уходу в 1,367 ниже, чем абоненты, у которых их нет.
3. Для клиентов, не пользующихся интернет-услугами, риск ухода выше в 0,356 раз, чем для пользователей, имеющих доступ к сети «Интернет».
4. Риск отказа от услуг компании абонентов, которые подписывают контракт на срок менее одного года, в 42 раза выше, чем для тех, кто имеет соглашение на более длительный период.

Приведем результаты построения модели Кокса с использованием языка программирования R. Отчет по модели пропорциональных рисков, написанный в среде Rstudio, также выводит все основные характеристики модели, однако в отличие от алгоритма, реализованного в SPSS, для оценки значимости коэффициентов используется не критерий Вальда, а Z-статистика [14].

Итоговый отчет модели представлен на рисунке 2. Проверка на пропорциональность рисков представлена в нижней части рисунка 2.

```
>res.cox1
Call:
Coxph(formula=Surv0bj ~ Contact+ Depends, data=DT, ties="breslow")
```

	coef	Exp(coef)	Se(coef)	z	p
Contact	4.14305	62.99485	0.70790	5.853	4.84e-09
Dependents	0.55931	1.74947	0.06978	8.016	1.09e-15

	chisq	df	p
Contact	3.36	1	0.067
Dependents	1.71	1	0.190
GLOBAL	4.95	1	0.084

Составлено авторами по материалам исследования

Рис. 2. Сводная статистика по регрессии Кокса, реализованная в среде Rstudio с проверкой на пропорциональность рисков

Отметим, что в SPSS оценка модели Кокса осуществлялась с помощью приближения Бреслоу, так как только данный тип реализован в статистическом пакете. С помощью языка программирования R можно задать иной метод приближения (Эфрона) для оценки параметров и на основании информационных критериев качества определить наилучшую модель.

Все более нарастающая популярность модели Кокса отражается и в разнообразии инструментальных средств для ее реализации. В качестве примера можно привести запатентованный программный продукт [7] для реализации оценки Каплана–Мейера и модели пропорциональных рисков с изменяющимися во времени переменными.

Третий инструмент реализации – это российская аналитическая платформа Loginom [5]. На ней был разработан сценарий, содержащий шесть узлов «Python» [14], реализующих 6 подмоделей. Базовый отчет подмодели «Выход регрессии» представлен на рисунке 3, «Коэффициенты регрессии» – на рисунке 4.

Регрессия Кокса • Выход регрессии • Быстрый просмотр данных									
#	ab customerID	12 Dependents	12 time	12 Contract	0 ₁ indicator	9.0 Условный риск	9.0 Базовый риск	9.0 Предсказанный риск	
1	7590-VHVEG	1	1	1	false	1,50	0,07	0,11	
2	8779-QRDMV	1	1	1	true	1,50	0,07	0,11	
3	1066-JKSGK	1	1	1	true	1,50	0,07	0,11	
4	8665-UTDZH	0	1	1	true	0,85	0,07	0,06	
5	7310-EGVHZ	1	1	1	false	1,50	0,07	0,11	
6	3413-BMNZE	1	1	1	false	1,50	0,07	0,11	
7	2273-QCKXA	1	1	1	false	1,50	0,07	0,11	
8	5919-TMRGD	0	1	1	true	0,85	0,07	0,06	
9	2424-WVHPL	1	1	1	false	1,50	0,07	0,11	
10	6380-ARCEH	1	1	1	false	1,50	0,07	0,11	
11	3679-XASPY	0	1	1	false	0,85	0,07	0,06	
12	3930-ZGWWE	1	1	1	false	1,50	0,07	0,11	
13	3091-FYHKI	1	1	1	true	1,50	0,07	0,11	
14	0390-DCFDQ	1	1	1	true	1,50	0,07	0,11	
15	2135-RXIHG	1	1	1	true	1,50	0,07	0,11	

Составлено авторами по материалам исследования

Рис. 3. Результат работы компонента «Выход регрессии» сценария Loginom

№	Метка	Значение
1	ab Ковариаты	Contract
2	9.0 B	4,16
3	9.0 exp(B)	64,29
4	9.0 Стандартная ошибка	0,71
5	9.0 Нижняя граница ДИ для B	2,78
6	9.0 Верхняя граница ДИ для B	5,55
7	9.0 Нижняя граница ДИ для exp(B)	16,06
8	9.0 Верхняя граница ДИ для exp(B)	257,41
9	9.0 z-статистика	5,88
10	9.0 p-значение	4,05e-9
11	9.0 Формула Хартли	27,88

Составлено авторами по материалам исследования

Рис. 4. Результат работы компонента «Коэффициенты регрессии» сценария Loginom

ОСНОВНЫЕ РЕЗУЛЬТАТЫ

Полученная регрессия Кокса с помощью метода Эфрона имеет следующий вид и может быть определена по формуле:

$$h(T) = h_0(T)e^{4,16361 \cdot \text{Contract} + 0,57015 \cdot \text{Dependents}} \quad (2)$$

Качество модели оценивалось с использованием индекса конкордации (C-индекс) и информационных критериев Акаике (англ. Akaike information criterion, AIC) и Шварца (англ. Bayesian information criterion, BIC). Результаты сравнения моделей представлены в таблице 3. Жирным выделен метод, дающий лучший результат вычислений.

Таблица 3

Итоговые результаты оценки качества моделей Кокса

Метод оценивания	C-индекс	Информационный критерий Акаике (AIC)	Информационный критерий Шварца (BIC)
Метод Бреслоу	0,586	22 782,4	22 793,04
Метод Эфрона	0,586	22 706,08	22 716,72

Составлено авторами по материалам исследования

На основании результатов, представленных в таблице 3, можно сделать следующие выводы.

1. Чем меньше значение показателей AIC и BIC, тем модель лучше, а значит, следует предпочесть модель, оцениваемую методом Эфрона.

2. Индекс конкордации (C-индекс) для обоих приближений близок к 0,5, а значит, прогнозы оценки риска можно признать удовлетворительными [4; 9].

Сравнительная характеристика трех инструментальных сред представлена в таблице 4.

Таблица 4

Сравнительная характеристика программных сред для реализации модели пропорциональных рисков

Критерий	Статистический пакет Statistical Package for the Social Sciences	Язык программирования R	Аналитическая платформа Loginom
Спецификация модели	+	+	+
Оценка параметров модели с помощью различных приближений	-	+	+/-
Оценка качества модели и ее параметров	+	+	+
Возможность графической интерпретации функций выживаемости и риска	+	+	+
Проверка на пропорциональность рисков	-	+	+
Вывод остатков	+	+	+
Расчет характеристик качества модели	-	+	+
Удобство интерфейса	+	-	+

Составлено авторами по материалам исследования

Исходя из данных таблицы 4, разработанный компонент «Регрессия Кокса» на платформе Loginom может стать хорошей альтернативой для построения полной и функциональной модели пропорциональных рисков.

ЗАКЛЮЧЕНИЕ

Таким образом, существенное влияние на отток клиентов оказывает срок заключения контракта с фирмой, а также наличие иждивенцев. Однако средний уровень индекса конкордации указывает на то, что прогностические свойства модели не являются достаточно надежными для определения ухода клиента.

Основываясь на логике и характеристиках качества полученной модели, можно утверждать, что модель Кокса применима для решения широкого круга задач в области экономики и управления. Однако для доказательства этого утверждения следует провести более глубокое исследование.

В ходе работы также установлено, что статистический пакет SPSS обладает удобным интерфейсом, но не включает в себя тестирование на пропорциональность рисков и вычисление характеристик качества, а язык программирования R, способный провести грамотную реализацию модели, не подходит для аналитики рядовому пользователю, который не умеет программировать.

Сценарий, разработанный при помощи аналитической платформы Loginom, допускающей использование узлов, реализованных с применением языка программирования Python, способен дополнить функционал различных инструментальных сред. Таким образом, любой пользователь может получить всю необходимую информацию о параметрах модели и ее качестве, а также провести визуальный анализ на основе графиков кривых выживаемости и риска и остатков. Помимо этого, аналитическая платформа позволяет дополнить функционал построения регрессии Кокса такими элементами, как: настройка пользователем метода оценивания регрессии Кокса (приближения Бреслоу, Эфрона и других); настройка пользователем метода оценивания остатков; проведение перекрестной проверки для оценки качества модели; автоматическая интерпретация графиков выживаемости и риска.

В современной цифровой среде существует множество технологий, которые могут решить перечень стандартных задач. Однако любая предметная область требует тонкой настройки. Особенности предметной области не позволяют использовать готовые инструменты в силу множества ограничений. Современные аналитические платформы позволяют устранить эти недостатки и становятся гибким элементом, который позволяет донастроить модель под особенности любой предметной области. Разработанный программный инструментариум дает возможность использовать гибкую настройку параметров модели Кокса для прогнозирования вероятности оттока клиентов с использованием платформы Loginom.

Библиографический список

1. Шарашова Е.Е., Холматова К.К., Горбатова М.А., Гржибовский А.М. Применение анализа выживаемости в здравоохранении с использованием пакета статистических программ SPSS. *Наука и здравоохранение*. 2017;5:5–28.
2. Ширококова М.А., Пивкин К.С. Интерпретация моделей случайного леса выживаемости в применении к расчету банковского кредитного риска. В сб.: *Статистические методы анализа экономики и общества: Труды 11-й Международной научно-практической конференции студентов и аспирантов, Москва, 12–15 мая 2020 г.* М.: Национальный исследовательский университет «Высшая школа экономики»; 2020. С. 163–164. <https://doi.org/10.17323/978-5-7598-2305-6>
3. Красоткина О.В., Попов В.А., Нгуен Т.Ч., Моттль В.В. Байесовский подход к оцениванию факторов риска в анализе продолжительности жизни. *Известия ТулГУ. Технические науки*. 2013;2:168–176.
4. Изергина А.Р. Обзор статистических методов оценки надежности. В сб.: Новиков Д.А. (ред.) *Математические модели современных экономических процессов, методы анализа и синтеза экономических механизмов. Актуальные проблемы и перспективы менеджмента организаций в России: Сборник статей XII Всероссийской научно-практической конференции, Самара, 28–31 мая 2018 г.* Самара: Самарский научный центр РАН; 2018. С. 45–50.
5. Аналитическая платформа Loginom. *О платформе*. <https://loginom.ru/platform> (дата обращения: 01.08.2022).
6. Шарашова Е.Е., Холматова К.К., Горбатова М.А., Гржибовский А.М. Применение регрессии Кокса в здравоохранении с использованием пакета статистических программ SPSS. *Наука и здравоохранение*. 2017;6:5–27.
7. Рей А.И., Андропова Е.С. *Анализ выживаемости фирм отрасли: оценка Каплана–Мейера и модель пропорциональных рисков с изменяющимися во времени переменными. Свидетельство о государственной регистрации программы для ЭВМ № 2020664382 Российская Федерация. RU2020663158 (Патент) 2020.*
8. HR портал. *Анализ выживаемости*. <https://hr-portal.ru/statistica/gl14/gl14.php> (дата обращения: 01.08.2022).

9. Грищенко Д.А., Катаев А.В. Анализ методов моделирования и прогнозирования оттока клиентов. *Вестник науки и образования*. 2018;2(5-41):21–23.
10. Паклин Н.Б., Кацко И.А., Кремьянская Е.В. Моделирование LTV подписчика сайта с использованием low-code платформы Loginom. *САЕС*. 2021;3:94–106.
11. Румянцева Е.В., Фурманов К.К. Использование вневыборочных остатков Кокса–Снелл при прогнозировании наступления событий. *Бизнес-информатика*. 2021;1(15):7–18. <https://doi.org/10.17323/2587-814x.2021.1.7.18>
12. Vafeiadis T., Diamantaras K.I., Sarigiannidis G., Chatzisavvas K. Customer churn prediction in telecommunications. Simulation Modelling: Practice and Theory. *Journal of Applied Technology and Innovation*. 2018;1(2):7–14.
13. Яблонцева А.Д. Реализация модели пропорциональных рисков Кокса средствами языка программирования R. *Modern Science*. 2021;7:380–384.
14. Аналитическая платформа Loginom. *Помощник по языку программирования Python в Loginom*. <https://help.loginom.ru/userguide/processors/programming/python/index.html/> (дата обращения: 20.08.2022).
15. Kaggle. *Telco Customer Churn*. <https://www.kaggle.com/blastchar/telco-customer-churn> (accessed 10.08.2022).
16. R Documentation. *Cox Regression*. <https://www.rdocumentation.org/packages/survival/versions/3.2-10/topics/coxph/> (accessed 09.08.2022).

References

1. Sharashova E.E., Kholmatova K.K., Gorbatova M.A., Grzhibovsky A.M. Application of survival analysis in healthcare using the SPSS statistical software package. *Science and healthcare*. 2017;5:5–28.
2. Shirobokova M.A., Pivkin K.S. Interpretation of models of random forest survival in application to the calculation of bank credit risk. In: *Statistical methods of analysis of economics and society: Proceedings of the 11th International Scientific and Practical Conference of Students and Postgraduates, Moscow, 12–15 May, 2020*. Moscow: National Research University “Higher School of Economics”; 2020. P. 163–164. <https://doi.org/10.17323/978-5-7598-2305-6>
3. Krasotkina O.V., Popov V.A., Nguyen T.Ch., Mottl V.V. Bayesian approach to risk factors assessment in life expectancy analysis. *News of TulaSU. Technical Sciences*. 2013;2:168–176.
4. Izergina A.R. Review of statistical methods for assessing reliability. In: Novikov D.A. *Mathematical models of modern economic processes, methods of analysis and synthesis of economic mechanisms. Actual problems and prospects of management of organizations in Russia: Proceedings of the 12th All-Russian Scientific and Practical Conference, Samara, 28–31 May 2018*. Samara: Samara Scientific Center of the Russian Academy of Sciences; 2018. P. 45–50.
5. Loginom analytical platform. *About the platform*. <https://loginom.ru/platform>. (accessed 01.08.2022).
6. Sharashova E.E., Kholmatova K.K., Gorbatova M.A., Grzhibovsky A.M. Application of Cox regression in healthcare using the SPSS statistical software package. *Science and healthcare*. 2017;6:5–27.
7. Rey A.I., Andronova E.S. *Analysis of the survival of firms in the industry: Kaplan–Meyer assessment and a model of proportional risks with time-varying variable*. Certificate of state registration of the computer program No. 2020664382 Russian Federation. RU2020663158: application 28.10.2020 (Patent) 2020.
8. HR portal. *Survival analysis*. <https://hr-portal.ru/statistica/gl14/gl14.php> (01.08.2022).
9. Grishchenko D.A., Kataev A.V. Analysis of methods for modeling and forecasting customer outflow. *Bulletin of Science and Education*. 2018;2(5-41):21–23.
10. Paklin N.B., Katsko I.A., Kremyanskaya E.V. LTV modeling of a site subscriber using the low-code Loginom platform. *САЕС*. 2021;3:94–106.
11. Rumyantseva E.V., Furmanov K.K. The use of non-selective Cox-Snell residues in predicting the occurrence of events. *Business informatics*. 2021;1(15):7–18. <https://doi.org/10.17323/2587-814x.2021.1.7.18>
12. Vafeiadis T., Diamantaras K.I., Sarigiannidis G., Chatzisavvas K. Customer churn prediction in telecommunications. Simulation Modelling: Practice and Theory. *Journal of Applied Technology and Innovation*. 2018;1(2):7–14.
13. Yablontseva A.D. Implementation of the cox proportional risks model by means of the programming language R. *Modern Science*. 2021;7:380–384.
14. Loginom analytical platform. *Python programming language Assistant in Loginom*. <https://help.loginom.ru/userguide/processors/programming/python/index.html/> (accessed 20.08.2022).
15. Kaggle. *Telco Customer Churn*. <https://www.kaggle.com/blastchar/telco-customer-churn> (accessed 10.08.2022).
16. R Documentation. *Cox Regression*. <https://www.rdocumentation.org/packages/survival/versions/3.2-10/topics/coxph/> (accessed 09.08.2022).