

ИТОГИ НАУКИ И ТЕХНИКИ

СЕРИЯ

ОБЩИЕ ПРОБЛЕМЫ БИОЛОГИИ

Том 6

ПОПУЛЯЦИОННАЯ ГЕРОНТОЛОГИЯ

Под редакцией проф. Е. Б. Бурлаковой
и канд. биол. наук Л. А. Гаврилова

Серия издается с 1962 г.



МОСКВА 1987

ПЕРВИЧНАЯ СТАТИСТИЧЕСКАЯ ОБРАБОТКА ДАННЫХ ПО ВЫЖИВАЕМОСТИ ОРГАНИЗМОВ

С.П. Ермаков, Н.С. Гаврилова

ВВЕДЕНИЕ

В последние годы раздел статистики, посвященный анализу данных по выживаемости, пополнился большим количеством новых методов. Именно в это время появились специальные монографии, посвященные анализу выживания в медико-биологических экспериментах [28, 33, 39]. Толчком к такому развитию в значительной степени послужила необходимость обработки результатов большого количества клинических и экспериментальных исследований в онкологии, радиобиологии и эпидемиологии инфекционных заболеваний. Поэтому неудивительно, что именно в этих областях науки новейшие статистические методы анализа получили наибольшее распространение [7, 8, 26, 51]. Исходно большинство статистических методов анализа выживаемости было разработано в математической теории надежности для задач обработки результатов испытания и эксплуатации промышленных изделий [1, 2]. Однако эти подходы оказались пригодными и для анализа выживаемости биологических объектов [7, 8].

К сожалению, в современной экспериментальной геронтологии эти методы практически не используются [43], хотя именно здесь анализу продолжительности жизни и смертности должно придаваться первостепенное значение. Как правило, весь статистический анализ полученных данных в экспериментальной геронтологии ограничивается расчетом средней и медианной продолжительностей жизни с соответствующими стандартными ошибками и сравнением двух значений средней продолжительности жизни по критерию Стьюдента [12, 50]. Между тем, как показали специальные исследования [3, 9, 10] нормальный закон распределения продолжительности жизни не выполняется ни в популяциях лабораторных животных [9, 10], ни в популяциях человека [3]. Поэтому использование критерия Стьюдента при сравнении средних значений продолжитель-

ности жизни для небольших по численности выборок (менее 50 организмов) может привести к ошибочным выводам. Кроме того, публикация результатов опыта в виде значений средней и медианной продолжительностей жизни приводит к существенной потере информации об изучаемых объектах, так как сравниваемые выборки могут отличаться и по другим параметрам выживания [43], а публикация значений максимальной продолжительности жизни, встречающаяся в статьях геронтологов, оказывается абсолютно бесполезной при статистической обработке результатов эксперимента, поскольку это понятие лишено смысла с точки зрения математической статистики [3].

Цель данного обзора состоит в том, чтобы познакомить читателей с современными статистическими методами анализа выживаемости.

Особенности данных по выживаемости. Данные по выживаемости обладают некоторыми особенностями, которые требуют разработки для них специальных методов статистического анализа.

Первая особенность данных по выживанию заключается в том, что распределение времени жизни не подчиняется нормальному закону, и конкретный вид распределения продолжительности жизни, как правило, неизвестен. Это обстоятельство заставляет использовать методы непараметрической статистики, которые по сравнению с обычными параметрическими методами являются более новыми и менее известными (в основном они были разработаны после 1940 г.). Их достоинство в том, что они применимы для любого вида распределения анализируемых данных.

Второй особенностью данных по выживанию является то, что наблюдаемые выборки, состоящие из набора времен жизни, не всегда бывают полными. Что это означает? Дело в том, что в ряде случаев объект эксперимента может уйти из-под наблюдения до момента своей гибели. Например, часть лабораторных животных может регулярно забиваться для проведения различных физиологических и биохимических анализов. Аналогичная проблема возникает при клинических испытаниях, когда больные по тем или иным причинам не могут участвовать в опыте до его завершения и уходят из-под наблюдения. Оказывается, информацию об объектах, ушедших из-под наблюдения, также можно использовать в статистическом анализе. В данном случае для статистики важно то, что продолжительность жизни таких объектов больше, чем длительность наблю-

дения за ними [38]. Такое явление ухода из-под наблюдения получило название цензурирования, а моменты ухода из-под наблюдения были названы временами цензурирования [7,38]. Выборки, содержащие цензурированные наблюдения, называются цензурированными выборками. В опыте и клинике наиболее часто может возникать случай так называемого однократного цензурирования справа [7], что означает прекращение наблюдения до того, как умрут все организмы в изучаемой группе. Тогда все выжившие к моменту прекращения эксперимента будут иметь одно и то же время цензурирования, совпадающее с концом наблюдения.

Цензурирование значительно осложняет статистический анализ данных, и усилия статистиков, в последние годы были направлены главным образом на то, чтобы распространить имеющиеся методы на анализ неполных данных. В этом направлении уже удалось добиться определенных успехов, хотя не все разработанные методы анализа хорошо изучены для небольших выборок [38]. Поэтому если в ходе эксперимента нельзя обойтись без цензурированных данных (например, если часть животных необходимо забивать для проведения анализов), численность животных в таком эксперименте должна быть достаточно большой, а цензурированные наблюдения — составлять не более 10% всех наблюдений.

К счастью, в экспериментальной геронтологии цензурированные наблюдения встречаются не столь часто, чего нельзя сказать о клинических испытаниях (например, в гериатрической клинике). В последнем случае, однако, возможно использование специальных статистических методов, разработанных для анализа цензурированных выборок и позволяющих исследователю сделать необходимые и обоснованные выводы.

Предварительные замечания о планировании длительных экспериментов по выживаемости. Прежде чем начинать длительный и трудоемкий эксперимент по изучению продолжительности жизни, необходимо тщательно его продумать. В первую очередь это касается определения необходимого размера выборки, который может сильно влиять на статистические выводы. Например, если в эксперименте ожидают найти достоверные различия в смертности при использовании какого-либо препарата, предположительно обладающего геропротекторным действием, то, взяв для эксперимента слишком малое количество животных, можно таких различий не обнаружить, хотя в действительности они существуют. При этом чем меньшие различия в выживаемости требуется обнаружить, тем большее ко-

личество животных следует брать для эксперимента. Если в опытах на лабораторных животных наблюдение обычно ведется до тех пор, пока не умрет последняя особь, то при клинических испытаниях препаратов складывается иная ситуация. Поэтому при планировании клинических экспериментов следует иметь в виду, что статистические выводы зависят не от того, сколько испытуемых взято под наблюдение, а от того, сколько из них умерло в ходе наблюдения. Исследование ста больных, из которых 50 умерло за период наблюдения, примерно столь же чувствительно, как и наблюдение за тысячей больных, из которых умерло 50 [46]. Эти и другие вопросы, связанные с планированием клинических испытаний, подробно освещены в работе большого коллектива английских статистиков [46, 47].

При планировании размера выборки необходимо стараться по возможности увеличить ее численность. При прочих равных условиях лучше провести один эксперимент на большой выборке (или одно сравнение двух больших выборок), чем много экспериментов на малых выборках. Это пожелание связано с тем обстоятельством, что почти все статистические методы теоретически выведены для асимптотического случая бесконечно больших выборок. В работе [46] неоднократно подчеркивается, что даже если в опытах на малых выборках обнаружены достоверные различия в выживаемости, эти результаты не следует безоговорочно принимать на веру. Какую же выборку следует считать достаточно большой? В работе [46] выборка считается большой, если количество организмов в опыте и контроле превышает 250. В других работах по статистике выживаемости выборку обычно считают достаточно большой, если ее численность превышает 100. По-видимому, именно эту цифру следует взять за основу при определении размера выборки, если опыты только начинаются и никакой дополнительной информации о выживаемости еще нет. Существуют и точные методы определения размера выборки, но они основаны на определенных предположениях относительно вида распределения продолжительности жизни. Например, в работе [26] приведен метод определения размера выборки, основанный на модели экспоненциального распределения.

Другое замечание относительно планирования опытов по изучению продолжительности жизни состоит в том, что не всегда следует сравнивать две группы, равные по численности [46]. Наоборот, целесообразно увеличить опытную группу, получающую новое, более перспективное лечение или геропр-

тектор, с тем, чтобы приблизительно уравнивать число умерших в опыте и контроле, если эксперимент будет проводиться не до полного вымирания всех изучаемых объектов (предполагается, что смертность в опытной группе будет меньше, чем в контроле).

И, наконец, последнее замечание касается выбора контрольной группы. Некоторые исследователи в своих работах используют так называемый ретроспективный контроль, то есть данные о выживаемости в ранее проведенных экспериментах. При этом исходят из предположения, что условия содержания животных или условия пребывания больных в клинике не претерпели заметных изменений. Использование ретроспективного контроля обычно оправдывают тем, что при этом можно примерно вдвое увеличить численность опытной группы за счет предполагаемой контрольной. Авторы работы [46] предостерегают от подобного выбора контрольной группы. Действительно, как правило, с течением времени происходят систематические изменения в условиях содержания лабораторных животных и в методах поддерживающей терапии в клинике, которые могут быть незаметны исследователю. В этом случае использование ретроспективного контроля приведет к систематическому искажению результатов.

1. ОСНОВНЫЕ ФУНКЦИИ, ХАРАКТЕРИЗУЮЩИЕ РАСПРЕДЕЛЕНИЕ ПРОДОЛЖИТЕЛЬНОСТИ ЖИЗНИ

1.1. Функция дожития и связанные с ней характеристики выживания

Данные о выживаемости обычно представлены набором соответствующих времен жизни

$$t_1 < t_2 \dots t_i < \dots < t_N,$$

рассматриваемых как случайные величины, которым соответствует функция распределения

$$F(t) = P\{t_i < t\}, \quad (1)$$

называемая также кумулятивной функцией распределения вероятности смерти. Эта функция отражает вероятность того, что

организм проживет время, меньшее, чем t . Однако в анализе продолжительности жизни более часто рассматривается вероятность противоположного события, называемая функцией дожития.

$$S(t) = 1 - F(t) = P\{t_i > t\}, \quad (2)$$

то есть вероятность того, что организм проживет время, большее, чем t . Поведение функции дожития отражает известная геронтологам кривая выживания.

Другой функцией, описывающей выживаемость, является функция плотности вероятности смерти — предел отношения вероятности смерти организма в интервале времени от t до $t + \Delta t$ к длине интервала Δt , стремящейся к нулю:

$$f(t) = \frac{dF(t)}{dt} = -\frac{dS(t)}{dt}. \quad (3)$$

Данная функция описывает известную геронтологам кривую распределения организмов по срокам их жизни. В демографии эта кривая называется кривой числа умерших (см. [3]).

Важной характеристикой выживаемости является функция интенсивности смертности, называемая также силой смертности или функцией риска:

$$\mu(t) = -\frac{d \ln S(t)}{dt}. \quad (4)$$

Эта функция характеризует удельную скорость гибели и связана с функциями $f(t)$ и $S(t)$ соотношениями:

$$\mu(t) = f(t)/S(t) \quad (5)$$

$$S(t) = \exp\left(-\int_0^t \mu(t) dt\right). \quad (6)$$

Наряду с функцией интенсивности смертности в анализе выживаемости используется кумулятивная интенсивность смертности, называемая в теории надежности функцией ресурса [2]:

$$\Lambda(t) = \ln(1/S(t)) = \int_0^t \mu(t) dt \quad (7)$$

Перечислив функции, характеризующие выживаемость, перейдем к методам получения их эмпирических оценок.

1.2. Получение эмпирических оценок для функции дожития

Для получения эмпирических оценок функции дожития используют два основных метода: метод множительных оценок Каплана-Мейера и метод таблиц смертности. Первый метод используется для выборок небольших размеров (менее 50), когда точно известны моменты смерти каждого члена группы. Для больших выборок применение метода Каплана-Мейера неудобно, так как требует громоздких расчетов, а представление результатов становится трудным для восприятия. Оценки, полученные методом таблиц смертности, сводятся к оценкам Каплана-Мейера при уменьшении временных интервалов в таблице смертности до тех пор, пока в каждом интервале окажется не более одного наблюдения.

1.2.1. Оценивание функции дожития для малых выборок: множительная оценка Каплана-Мейера

Для начала рассмотрим метод расчета эмпирических оценок функции дожития на примере полных данных о выживании (в отсутствие цензурирования). При этом исходные данные представляют собой набор времен жизни: $t_1 < t_2, \dots, < t_i, \dots, < t_N$.

Эмпирическую оценку кумулятивной функции распределения вероятности смерти можно получить из формулы [2]:

$$\hat{F}(t_i) = D(t_i)/N, \quad (8)$$

где $D(t_i) = \sum_{j=1}^i d_j$ - суммарное число умерших к моменту

t_i включительно, d_j - число умерших в момент t_j ; N - исходная численность выборки, равная в нашем примере 41 (табл. 1).

Функция дожития, в отличие от функции распределения вероятности смерти, определяет вероятность противоположного события, и ее оценка соответственно рассчитывается по формуле:

$$\hat{S}(t_i) = 1 - \hat{F}(t_i) = 1 - \frac{D(t_i)}{N}. \quad (9)$$

Таблица 1

Представление результатов опыта по изучению выживаемости лабораторных крыс. Исходная численность $N = 41$

Время жизни (дни) t_i	Суммарное число умерших, $D(t_i)$	Функция дожития, $\hat{S}(t_i)$	Стандартная ошибка функции дожития, $\hat{\sigma}_S(t_i)$	Функция ресурса, $\hat{A}(t_i)$
1	2	3	4	5
2	1	0,9756	0,0244	0,0244
31	2	0,9512	0,0341	0,0494
45	3	0,9268	0,0412	0,0750
88	4	0,9024	0,0469	0,1013
115	5	0,8780	0,0517	0,1284
122	6	0,8537	0,0559	0,1562
128	7	0,8293	0,0595	0,1847
142	8	0,8049	0,0627	0,2141
213	9	0,7805	0,0654	0,2444
243	10	0,7561	0,0679	0,2757
251	11	0,7317	0,0701	0,3079
256	12	0,7073	0,0719	0,3413
262	13	0,6829	0,0736	0,3758
269	14	0,6585	0,0750	0,4115
283	15	0,6341	0,0762	0,4485
288	16	0,6098	0,0771	0,4870
291	17	0,5854	0,0779	0,5270
297	18	0,5610	0,0785	0,5686
301	19	0,5366	0,0788	0,6121
304	20	0,5122	0,0790	0,6576
309	21	0,4878	0,0790	0,7052
313	22	0,4634	0,0788	0,7552

Продолжение табл. 1

1	2	3	4	5
320	23	0,4390	0,0785	0,8078
327	24	0,4146	0,0779	0,8634
340	25	0,3902	0,0771	0,9222
354	26	0,3659	0,0762	0,9847
360	27	0,3415	0,0750	1,0514
367	28	0,3171	0,0736	1,1228
375	29	0,2927	0,0719	1,1997
381	30	0,2683	0,0701	1,2831
395	31	0,2439	0,0679	1,3740
400	32	0,2195	0,0654	1,4740
409	33	0,1951	0,0627	1,5851
423	34	0,1707	0,0595	1,7101
426	35	0,1463	0,0559	1,8529
437	36	0,1220	0,0517	2,0196
451	37	0,0976	0,0469	2,2196
459	38	0,0732	0,0412	2,4696
465	39	0,0488	0,0341	2,8029
479	40	0,0244	0,0244	3,3029
496	41	0,0000		

Примечание. Рассчитано на основании данных С.И. Потапова

Пример расчета $\hat{S}(t_i)$ приведен в табл. 1 (колонка 3). График функции $\hat{S}(t)$ представляет собой ступенчатую функцию с высотой ступеньки, равной $d(t_i)/N$.

Приведенные оценки $\hat{F}(t)$ и $\hat{S}(t)$ являются несмещенными, то есть их математические ожидания совпадают с оцениваемыми величинами [2]. Дисперсии этих оценок равны между собой [2]:

$$[\text{var } \hat{F}(t)] = [\text{var } \hat{S}(t)] = \frac{\hat{F}(t) \cdot \hat{S}(t)}{N} \quad (10)$$

Несмещенной оценкой дисперсии функции дожития является величина, равная [2]:

$$[\text{var } \hat{S}(t_i)] = \frac{D(t_i)[N - D(t_i)]}{N^2(N-1)} \quad (11)$$

Стандартную ошибку для функции дожития рассчитывают, извлекая квадратный корень из дисперсии:

$$\hat{\sigma}_{S(t_i)} = \frac{1}{N} \sqrt{D(t_i)[N - D(t_i)] / (N-1)} \quad (12)$$

Теперь рассмотрим метод построения эмпирических оценок функции дожития для цензурированной выборки. В этом случае исходные данные представляют собой упорядоченный набор как времен жизни (t_i), так и времен цензурирования (t_j^*):

$$t_1 < t_2 < t_3^* < \dots < t_l < \dots < t_j^* < \dots < t_n$$

В 1958 г. Каплан и Мейер предложили непараметрическую оценку функции дожития по неполным данным, названную Product limit estimate, то есть оценку в виде произведения с конечным числом сомножителей [34]. В отечественной литературе эту оценку называют множительной оценкой [2]. Множительную оценку функции дожития для i -го момента смерти рассчитывают по формуле:

$$\hat{S}(t_i) = \left(1 - \frac{\delta_1}{N-1+1}\right) \left(1 - \frac{\delta_2}{N-2+1}\right) \dots \left(1 - \frac{\delta_i}{N-i+1}\right) \quad (13)$$

$$= \prod_{j=1}^i \left(1 - \frac{\delta_j}{N-j+1}\right),$$

где $\delta_j = \begin{cases} 1, & \text{если } t_j - \text{ момент смерти} \\ 0, & \text{если } t_j - \text{ момент цензурирования.} \end{cases}$

Здесь произведение берется для всех моментов времени, накопленных к i -му моменту смерти.

Проще,

$$\hat{S}(t_i) = \prod_{j=1}^i \left(1 - \frac{1}{r_j}\right), \quad (14)$$

где r_j обозначает число организмов непосредственно перед j -м моментом смерти.

Предыдущие выражения можно записать следующим образом:

$$\hat{S}(t_i) = \hat{S}(t_j) \left(1 - \frac{1}{r_i}\right), \quad (15)$$

где t_j - момент смерти предшествующий t_i .

В ряде случаев возможны ситуации, когда момент смерти совпадает с моментом цензурирования. В этом случае предполагают, что смерть наступает несколько раньше, чем цензурирование. Если в один момент времени происходит несколько смертей, то для оценки функции дожития можно воспользоваться формулой [17]:

$$\hat{S}(t_i) = \prod_{j=1}^i \left(1 - \frac{d_j}{r_j}\right), \quad (16)$$

где d_j - число умерших в момент t_j , а r_j - численность организмов непосредственно перед моментом смерти t_j .

Расчет стандартной ошибки для функции дожития проводится по формуле, предложенной Гринаудом [25, 26, 52]:

$$\hat{\sigma}_{S(t_i)} = \hat{S}(t_i) \sqrt{\sum_{j=1}^i \frac{d_j}{(N-j)(N-j+1)}} \quad (17)$$

или

$$\hat{\sigma}_{S(t_i)} = \hat{S}(t_i) \sqrt{\sum_{j=1}^i \frac{d_j}{r_j(r_j - d_j)}}.$$

Для оценки функции ресурса (кумулятивной интенсивности смертности) можно воспользоваться формулой [52]:

$$\hat{\lambda}(t_i) = \sum_{j=1}^i \frac{d_j}{r_j} \quad (18)$$

или формулой [22]:

Таблица 2
Оформление результатов исследования длительности ремиссий у больных лейкозом, получавших препарат 6-меркаптопурин. Исходная численность $N = 21$. Обработка данных [23]

Длительность ремиссии (мес)	t_i	Число больных перед моментом t_i, r_i	Число рецидивов в момент t_i, d_i	Функция дожития $\hat{S}(t_i)$	Σ_i	Стандартная ошибка функции дожития, $\hat{\sigma}_{S_i}$	Функция ресурса, $\hat{\lambda}(t_i)$
6	21		3	0,857	0,008	0,006	0,143
7	17		1	0,807	0,012	0,008	0,202
9	16		0	0,807	0,012	0,008	0,202
10	15		1	0,753	0,016	0,009	0,268
11	13		0	0,753	0,016	0,009	0,268
13	12		1	0,690	0,024	0,011	0,352
16	11		1	0,627	0,033	0,013	0,442
17	10		0	0,627	0,033	0,013	0,442
19	9		0	0,627	0,033	0,013	0,442
20	8		0	0,627	0,033	0,013	0,442
22	7		1	0,538	0,057	0,016	0,585
23	6		1	0,448	0,090	0,018	0,752
25	5		0	0,448	0,090	0,018	0,752
32	4		0	0,448	0,090	0,018	0,752
34	2		0	0,448	0,090	0,018	0,752
35	1		0	0,448	0,090	0,018	0,752

$$\hat{\Lambda}(t_i) = \sum_{j=1}^i \sum_{k=0}^{d_j-1} (r_j-k)^{-1} \quad (19)$$

Если в каждый момент времени происходит не более одной смерти, то оценки (18) и (19) идентичны. Другой оценкой функции ресурса является:

$$\hat{\Lambda}(t_i) = - \ln \hat{S}(t_i).$$

Последней оценкой предпочтительнее пользоваться в конце периода наблюдения, когда в живых остается менее 10% от исходной численности.

В качестве примера выборки, состоящей из неполных данных, рассмотрим клинические данные по длительности ремиссий (аналог длительности жизни) у больных лейкозом, получавших препарат 6-меркаптопурин (взято из работы [23]). Длительности ремиссий у этих больных составляли (в месяцах): 6*, 6, 6, 6, 7, 9*, 10*, 10, 11*, 13, 16, 17*, 19*, 20*, 22, 23, 25*, 32*, 32*. Звездочками отмечены моменты цензурирования. В табл. 2 для этих данных приведены значения функции дожития с соответствующей стандартной ошибкой и значения функции ресурса. Через \sum_i обозначены накопленные к i -му моменту времени суммы:

$$\sum_{j=1}^i \frac{d_j}{r_j(r_j-d_j)}$$

Оценка Каплана-Мейера обладает тем достоинством, что для ее получения нет необходимости группировать данные по возрастным интервалам, то есть она не зависит от выбора интервала разбиения. С другой стороны, при больших численностях выборок (>100) ее расчеты становятся слишком громоздкими и без большой потери информации могут быть заменены оценкой таблиц смертности.

1.2.2. Оценка функции дожития для больших выборок: метод таблиц смертности

Метод таблиц смертности, называемый также актуариальным методом, был впервые подробно изложен в работе S. Cutler и F. Edeger в 1958 г. [18]. Суть его состоит в группировке моментов гибели по интервалам времени (не обязательно равным). Весь период наблюдения разбивают на интервалы и для

каждого интервала определяют число организмов в начале интервала (n_i), число организмов, умерших в течение данного интервала (d_i) и число организмов, ушедших из-под наблюдения в интервале (c_i). Следующий этап анализа состоит в расчете числа организмов, подверженных риску умереть в течение рассматриваемого интервала:

$$r_i = n_i - \frac{1}{2} c_i.$$

В данном случае предполагают, что цензурирование в рассматриваемом интервале происходит случайно (моменты цензурирования равномерно распределены на интервале). Следовательно, организмы, ушедшие из-под наблюдения, подвержены риску смерти примерно в течение половины интервала времени.

Исходя из этих данных, можно оценить условную вероятность смерти в i -м интервале:

$$\hat{q}_i = \frac{d_i}{r_i}, \quad (19)$$

где d_i - число организмов, умерших в i -м интервале. Величина \hat{q}_i является оценкой вероятности того, что организм умрет в течение i -го интервала времени при условии, что он дожил до начала данного интервала. При расчете \hat{q}_i предполагают, что риск гибели на всем интервале постоянен [26]. Условную вероятность дожития \hat{p}_i рассчитывают как вероятность противоположного события:

$$\hat{p}_i = 1 - \hat{q}_i \quad (20)$$

В табл. 3 в качестве примера приведены данные по выживаемости больных лейкозом, взятые из работы [26].

Оценка функции дожития методом таблиц смертности или кумулятивная функция дожития определяется как кумулятивная доля доживающих до начала i -го интервала:

$$\hat{S}_i = \hat{p}_{i-1} \cdot \hat{S}_{i-1}, \quad i > 1, \quad (21)$$

где $\hat{S}_1 = 1$. Основой для получения этой оценки является соображение, что для дожития до начала i -го интервала требуется дожить до начала $(i-1)$ -го интервала и не умереть в течение этого интервала. Приближенная оценка стандартной

Предварительные данные, необходимые для построения таблицы смертности для больных лейкозом. Взято из [26]

Начало интервала в мес	Середина интервала в мес	Длина интервала в мес h_i	Число живых в начале интервала n_i	Число выживших в интервале c_i	Число подверженных риску в интервале r_i	Число умерших в интервале d_i	Вероятность смерти \hat{q}_i	Вероятность дожития \hat{p}_i
0,0	2,50	5,0	235	0	235,0	68	0,2894	0,7106
5,0	7,50	5,0	167	0	167,0	33	0,1976	0,8024
10,0	12,50	5,0	134	0	134,0	14	0,1045	0,8955
15,0	17,50	5,0	120	0	120,0	13	0,1083	0,8917
20,0	22,50	5,0	107	1	106,5	10	0,0939	0,9061
25,0	27,50	5,0	96	0	96,0	8	0,0833	0,9167
30,0	32,50	5,0	88	2	87,0	1	0,0115	0,9885
35,0	37,50	5,0	85	0	85,0	4	0,0471	0,9529

ошибки для табличной функции дожития была выведена Гринвудом еще в 1926 г. [13, 25, 26]:

$$\hat{\sigma}_{S_i} = \hat{S}_i \sqrt{\frac{1}{\sum_{j=1}^i (\hat{q}_j / r_j \hat{p}_j)}} \quad (22)$$

Как выяснилось впоследствии, данная оценка может сильно занижать истинное значение стандартной ошибки при большом количестве цензурированных наблюдений [28].

Функция интенсивности смертности оценивается в середине каждого интервала [18, 25, 26]:

$$\hat{\mu}_i = \frac{1}{h_i} \frac{d_i}{(r_i - d_i/2)} = \frac{2\hat{q}_i}{h_i(1 + \hat{p}_i)} \quad (23)$$

где h_i - длина интервала. Приближенная стандартная ошибка интенсивности смертности может быть рассчитана по формуле [25, 26]:

$$\hat{\sigma}_{\mu_i} = \sqrt{\frac{\lambda^2}{r_i \hat{q}_i} \left(1 - \left(\frac{\hat{\mu}_i h_i}{2}\right)^2\right)} \quad (24)$$

Эта оценка также может существенно занижать истинную стандартную ошибку при высокой степени цензурирования [28].

Существуют и другие способы оценки функции интенсивности смертности. Так, например, известный радиобиолог и геронтолог G.Sacher еще в 1956 г. [48] предложил следующую формулу для оценки интенсивности смертности:

$$\hat{\mu}_i = \frac{1}{h_i} [-\ln(1 - d_i/r_i)] \quad (25)$$

Исследование, выполненное методом Монте Карло, показало, что оценка (25) дает большее смещение, чем оценка (23) [27]. С другой стороны, оценка интенсивности смертности, даваемая формулой (23), не может превышать значение, равное $2/h_i$, в то время как оценка G.Sacher не накладывает никаких ограничений на рост интенсивности смертности. По-видимому, оценку (23) следует использовать при небольших значениях интенсивности смертности, а в тех случаях, когда она велика (превышает 1,0), лучше пользоваться оценкой G. Sacher.

Таблица смертности больных лейкозом. Взято из [26]

Начало интервала в мес	Середина интервала в мес	Длина интервала в мес	Функция дожития S_i	Функция плотности распределения f_i	Интенсивность смертности μ_i	Стандартная ошибка S_i	Стандартная ошибка f_i	Стандартная ошибка μ_i
0,0	2,50	5,0	1,0000	0,0579	0,0677	0,0000	0,0059	0,0081
5,0	7,50	5,0	0,7106	0,0281	0,0439	0,0296	0,0045	0,0076
10,0	12,50	5,0	0,5702	0,0119	0,0220	0,0323	0,0031	0,0059
15,0	17,50	5,0	0,5106	0,0111	0,0229	0,0326	0,0030	0,0063
20,0	22,50	5,0	0,4553	0,0086	0,0197	0,0325	0,0026	0,0062
25,0	27,50	5,0	0,4126	0,0069	0,0174	0,0321	0,0024	0,0061
30,0	32,50	5,0	0,3782	0,0009	0,0023	0,0317	0,0009	0,0023
35,0	37,50	5,0	0,3738	0,0035	0,0096	0,0316	0,0017	0,0048

Функция плотности вероятности смерти оценивается в середине интервала по формуле [25, 26]:

$$\hat{f}_i = \frac{\hat{S}_i - \hat{S}_{i+1}}{h_i} = \frac{\hat{p}_i \hat{q}_i}{h_i},$$

а ее приближенная стандартная ошибка по формуле [25,26]:

$$\delta_{f_i} = \frac{\hat{S}_i \hat{q}_i}{h_i} \sqrt{\sum_{j=1}^{i-1} \frac{\hat{q}_j}{r_j \hat{p}_j} + \frac{\hat{p}_i}{r_i \hat{q}_i}},$$

В заключение приведем в качестве примера таблицу смертности, построенную на основании данных табл. 3 по выживаемости больных лейкозом (табл. 4).

2. ДОВЕРИТЕЛЬНЫЕ ИНТЕРВАЛЫ И ДОВЕРИТЕЛЬНАЯ ПОЛОСА ДЛЯ ФУНКЦИИ ДОЖИТИЯ

2.1. Методы расчета доверительных интервалов для функции дожития

Полученные в предыдущем разделе точечные оценки функции дожития не позволяют судить о точности и надежности полученных оценок. Сделать это позволяют интервальные оценки функции дожития или доверительные интервалы.

Для начала рассмотрим полную выборку без цензурированных наблюдений (см. табл. 1). Мы имеем набор эмпирических оценок функции дожития - $\hat{S}(t)$. Используя центральную предельную теорему теории вероятностей, можно показать, что в пределе при неограниченном увеличении размера выборки (при $N \rightarrow \infty$) распределение разности между истинной и эмпирической функцией дожития будет стремиться к нормальному [2]. На этом факте и основан метод построения доверительных интервалов для функции дожития. Можно показать, что при $N \rightarrow \infty$ вероятность выполнения неравенств:

$$\hat{S}(t_i) - z_\alpha \sqrt{\frac{D(t_i)[N-D(t_i)]}{N^2(N-1)}} \leq S(t_i) \leq \hat{S}(t_i) + z_\alpha \sqrt{\frac{D(t_i)[N-D(t_i)]}{N^2(N-1)}} \quad (26)$$

будет стремиться к $1-\alpha$. z_α обозначает квантиль уровня α стандартного нормального распределения:

$$1-\alpha = 2\Phi(z_\alpha).$$

Значения z_α находятся из таблиц нормального распределения. Для $\alpha = 0,05$ z_α составляет 1,96. Выражение (26) можно также переписать в виде:

$$\hat{S}(t_i) - z_\alpha \hat{\sigma}_{S_i} \leq S(t_i) \leq \hat{S}(t_i) + z_\alpha \hat{\sigma}_{S_i}. \quad (27)$$

Этот прием получения доверительных интервалов для больших выборок в случае асимптотически нормальных оценок широко используется в математической статистике.

Для неполных (цензурированных) выборок метод построения доверительных интервалов практически не отличается от описанного выше. В качестве эмпирической оценки функции дожития берется множительная оценка Каплана-Мейера, и доверительные интервалы функции дожития соответствуют величинам: $\hat{S}(t_i) \pm z_\alpha \hat{\sigma}_{S_i}$, где $\hat{\sigma}_{S_i}$ обозначает стандартную ошибку функции дожития в момент t_i . Таким образом, доверительный интервал уровня α для функции дожития можно получить из выражения [53]:

$$S(t_i) = \hat{S}(t_i) \pm z_\alpha \cdot \hat{S}(t_i) \sqrt{\sum_{j=1}^i \frac{d_j}{r_j(r_j - d_j)}}, \quad (28)$$

где для оценки стандартной ошибки функции дожития используется формула Гринвуда (17) или из следующего выражения [51]:

$$S(t_i) = \hat{S}(t_i) \pm z_\alpha \cdot \hat{S}(t_i) \sqrt{\frac{1 - \hat{S}(t_i)}{r_i}}, \quad (29)$$

где стандартная ошибка функции дожития рассчитывается по формуле, предложенной в работе [47]. В последнем случае оценка стандартной ошибки дает завышенные значения для больших выборок и при большом числе цензурированных наблюдений [52].

В нашем примере (см. табл. 2) 95%-ный доверительный

интервал для функции дожития составляет для $t = 10$ месяцев: $0,753 \pm 1,96 \cdot 0,009$, то есть: $0,735 \leq S(10) \leq 0,771$.

2.2. Расчет доверительной полосы для функции дожития

Доверительные интервалы определяют точность и надежность оценок функции дожития лишь в те моменты времени, для которых они определены. Чтобы оценить точность определения функции дожития на всем интервале наблюдения в любой момент времени, необходимо построить для нее доверительную полосу. Считается, что две функции $S_1(t)$ и $S_2(t)$ определяют на интервале $(0, t_N) \gamma$ -доверительную полосу для вероятности дожития $S(t)$, если вероятность того, что истинная функция дожития лежит в полосе, определяемой функциями $S_1(t)$ и $S_2(t)$, не меньше, чем γ :

$$P\{S_1(t) \leq S(t) \leq S_2(t)\} \geq \gamma \quad (30)$$

для всех t в интервале $(0, t_N)$. Обычно γ выражают в процентах (95 или 99%).

Существующие методы нахождения доверительной полосы для функции дожития также разработаны для асимптотического случая, когда $N \rightarrow \infty$. При этом γ -доверительная полоса для функции $S(t)$ в отсутствие цензурирования имеет вид [2]:

$$\hat{S}(t) - \lambda_\gamma / \sqrt{N} \leq S(t) \leq \hat{S}(t) + \lambda_\gamma / \sqrt{N}, \quad (31)$$

где λ_γ - квантиль уровня γ распределения Колмогорова-Смирнова (находится из таблиц). Для $\gamma = 0,95$ (или 95%) его значение равно 1,36, а для $\gamma = 0,99$ $\lambda = 1,63$. Например, для данных из табл. 1 95%-ную доверительную полосу составляют функции:

$$S_1(t) = \hat{S}(t) + 0,2124 \quad \text{и} \quad S_2(t) = \hat{S}(t) - 0,2124.$$

Метод построения доверительной полосы для функции дожития по неполным данным был предложен сравнительно недавно [29]. Чтобы построить доверительную полосу для функции дожития согласно этому методу, необходимо рассчитать следующие промежуточные величины:

2843

$$C_N(t_i) = N \sum_{j=1}^i \frac{\delta_j}{(N-j)(N-j+1)} \quad (32)$$

$$K_N(t_i) = 1/(1 + C_N(t_i))$$

$$D_N(t_i) = \sqrt{N} \cdot \hat{S}(t_i)/K_N(t_i)$$

Тогда при $N \rightarrow \infty$ вероятность выполнения неравенства

$$\hat{S}(t) - \lambda_y^* D_N(t) \leq S(t) \leq \hat{S}(t) + \lambda_y^* D_N(t) \quad (33)$$

на всем интервале наблюдения $0 \leq t \leq t_N$ будет стремиться к γ . В отсутствие цензурированных наблюдений λ_y^* равно квантилю уровня γ распределения Колмогорова-Смирнова. Если имеются цензурированные данные, то для нахождения λ_y^* необходимо оценить величину a , характеризующую степень цензурирования [29]:

$$\hat{a}_i = \frac{1}{1 + C_N(t_i)} \quad (34)$$

Ниже приводится таблица значений λ_y^* для разных уровней a и γ , взятая из работы [29] (табл. 5).

Таблица 5
Значения λ_y^* для разных a и γ

γ	$a = 0,50$	$a = 0,60$	$a = 0,75$	$a = 0,90$	$a = 1,00$
0,99	1,552	1,600	1,626	1,628	1,628
0,95	1,273	1,321	1,354	1,358	1,358
0,90	1,133	1,181	1,217	1,224	1,224

Более подробная таблица значений λ_y^* приведена в работе [29]. Доверительная полоса для функции дожития неполной выборки отличается тем, что ее ширина неодинакова в разных участках кривой. Например, по данным табл. 2, при $t = 10$ мес. можно рассчитать, что $C_N(10) = 0,336$, $K_N(10) = 0,749$, $D_N(10) = 0,219$. Величина $\hat{a} = 1/1,336 = 0,749$. Таким образом, из табл. 5 находим, что величина λ_y^* для 95%-ной доверительной полосы составляет 1,354. Следовательно, ширина доверительной полосы в 10 мес равна $2 \times 0,297$. В то же

время можно рассчитать, что ширина 95%-ной доверительной полосы в 16 месяцев составляет $2 \times 0,305$.

3. ОСНОВНЫЕ ХАРАКТЕРИСТИКИ РАСПРЕДЕЛЕНИЯ ПРОДОЛЖИТЕЛЬНОСТИ ЖИЗНИ

Наряду с оценкой функции дожития, существуют методы расчета обобщенных характеристик распределения продолжительности жизни. К таким характеристикам относятся средняя продолжительность жизни, медианная продолжительность жизни, асимметрия и эксцесс распределения.

3.1. Средняя продолжительность жизни

Средняя продолжительность жизни для набора времен жизни в отсутствие цензурированных наблюдений рассчитывается как обычное среднее арифметическое для полной выборки:

$$\bar{t} = \frac{1}{N} \sum_{i=1}^n t_i \cdot d_i \quad (35)$$

Выборочная дисперсия распределения продолжительности жизни оценивается по формуле [6,11]:

$$\hat{\sigma}_t^2 = \frac{1}{N-1} \sum_{i=1}^n (t_i - \bar{t})^2 \cdot d_i \quad (36)$$

Эту оценку дисперсии распределения продолжительности жизни не следует путать с оценкой дисперсии для функции дожития, которая характеризует степень рассеивания данных "по вертикали" (то есть по оси $S(t)$), в то время как дисперсия распределения продолжительности жизни отражает степень рассеивания данных "по горизонтали" (то есть по оси времени). Среднее квадратичное отклонение (стандартная ошибка) распределения продолжительности жизни определяется как корень квадратный из дисперсии:

$$\hat{\sigma}_t = \sqrt{\hat{\sigma}_t^2} \quad (37)$$

Оценка дисперсии для средней продолжительности жизни в N раз меньше выборочной дисперсии распределения продолжительности жизни:

$$\frac{\hat{\sigma}_{\bar{t}}^2}{\hat{\sigma}_t^2} = \frac{1}{N} \frac{\hat{\sigma}_t^2}{t} = \frac{1}{N^2} \sum_{i=1}^n (t_i - \bar{t})^2 d_i. \quad (38)$$

Отсюда можно определить приближенный доверительный интервал для теоретической средней продолжительности жизни в случае больших выборок:

$$\bar{t} - z_\alpha \hat{\sigma}_{\bar{t}} / \sqrt{N} \leq t_{cp} \leq \bar{t} + z_\alpha \hat{\sigma}_{\bar{t}} / \sqrt{N}, \quad (39)$$

где z_α - квантиль уровня α стандартного нормального распределения, равный 1,96 для $\alpha = 0,05$ и 2,57 для $\alpha = 0,01$.

Для цензурированной выборки среднюю продолжительность жизни можно определить из формулы [34]:

$$\bar{t} = \sum_{i=1}^n \hat{S}(t_{i-1})(t_i - t_{i-1}), \quad (40)$$

где суммирование проводится по всем моментам смерти. Если наблюдение прекращают до того, как умрут все организмы, составляющие группу, то среднюю продолжительность жизни однозначно определить нельзя. Ирвин предлагает в этом случае рассчитывать "среднюю продолжительность жизни, ограниченную временем L " (см. [34]). В качестве времени L обычно выбирают последний момент цензурирования, подставляя его в формулу (40) вместо последнего момента смерти. При этом оценка средней продолжительности жизни будет всегда заниженной по сравнению с истинным значением.

Среднюю продолжительность жизни для данных, сгруппированных в виде таблицы смертности, можно рассчитать, используя формулу [34]:

$$\bar{t} = \frac{1}{2} \sum_{i=1}^n (\hat{S}_{i-1} + \hat{S}_i)(t_i - t_{i-1}). \quad (41)$$

Дисперсия средней продолжительности жизни, полученной из формулы (40), может быть рассчитана следующим образом [34]:

$$\text{var}(\bar{t}) = \sum_{i=1}^n \frac{A_i^2}{r_i(r_i - 1)}, \quad (42)$$

где A_i определяется последовательно, начиная с наибольшего момента смерти, по формулам: $A_n = \hat{S}(t_{n-1})(t_n - t_{n-1})$;

$$A_i = A_{i+1} + \hat{S}(t_{i-1})(t_i - t_{i-1}), \quad i = n-1, n-2, \dots, 1.$$

Е.Карлан и Р.Меер предлагают домножать оценку (42) на величину $\hat{D}/(D-1)$, где D - суммарное число умерших за весь период наблюдения [34].

В нашем примере (табл. 2) средняя продолжительность ремиссий больных лейкозом, ограниченная 35 месяцами, составляет:

$$\bar{t}_{[35]} = 23,28 \text{ мес}$$

Для расчета дисперсии средней продолжительности ремиссий использовалась формула:

$$\text{var}(\bar{t}) = \sum_{i=1}^n \frac{A_i^2 \cdot d_i}{r_i(r_i - d_i)},$$

так как в отдельные моменты времени происходило более одного рецидива. Рассчитанная таким образом оценка дисперсии составляет 7,99. Умножив ее на отношение $D/(D-1)$, равное в данном случае 9/8, получаем, что дисперсия средней продолжительности ремиссий составляет 8,99, а ее стандартная ошибка - 3,00. Таким образом, в нашем примере $\bar{t}_{[35]} = 23,28 \pm 3,00$ мес.

Метод расчета дисперсии средней продолжительности предстоящей жизни, определяемой для таблиц смертности, подробно описан в работе [16].

3.2. Медианная продолжительность жизни и построение непараметрических доверительных интервалов для медианы

Средняя продолжительность жизни является не самой лучшей характеристикой центра распределения. Это связано с тем, что средняя продолжительность жизни очень чувствительна к случайным выбросам данных на концах распределения. Например, если в экспериментальной выборке какой-нибудь организм проживет значительно дольше остальных, то это приведет к сильному сдвигу средней продолжительности жизни в область больших возрастов. Поэтому полученная величина средней про-

должительности жизни не будет отражать характерных значений продолжительности жизни для всей выборки.

Этого недостатка лишена медианная продолжительность жизни, которая соответствует времени, до которого доживает 50% исходной популяции. Если данные не сгруппированы в таблицу смертности, то медианной продолжительностью жизни является момент смерти с номером $(N+1)/2$ для выборок, содержащих нечетное число наблюдений. Для выборок с четным числом наблюдений медианная продолжительность жизни определяется как полусумма времен с номерами $N/2$ и $N/2+1$.

В нашем примере (табл. 1) медианная продолжительность жизни равна 309 дням.

Для цензурированной выборки медианная продолжительность жизни определяется как наименьший момент времени, для которого выполняется неравенство: $\hat{S}(t) \leq 1/2$.

Если данные представлены в виде таблицы смертности, то медианную продолжительность жизни можно рассчитать по формуле [6]:

$$Me = t_{Me} + h_i \left(\frac{0,5N - D_{i-1}}{d_i} \right), \quad (43)$$

где t_{Me} - начало интервала, в который входит медиана, h_i - длина интервала, N - объем выборки, D_{i-1} - суммарное число умерших к началу интервала, в котором находится медиана, d_i - число умерших в интервале, где находится медиана. В нашем примере (см. табл. 4) медианная продолжительность жизни равна:

$$Me = 15,0 + 5,0 \cdot \frac{(0,5 \cdot 235 - 115)}{13} = 15,96 \text{ мес}$$

Несмотря на преимущество показателя медианной продолжительности жизни перед средней продолжительностью жизни, его использование до последнего времени было ограничено отсутствием хороших методов оценки доверительных интервалов для медианы. В то же время этот показатель характеризуется значительным разбросом, особенно в случаях, когда кривая дожития мало меняется вблизи медианы [46].

Однако в последние годы появилось много работ, в которых были предложены различные методы оценки непараметрических доверительных интервалов для медианной продолжительности жизни [15, 20, 51, 52]. Е. Slud с соавторами предла-

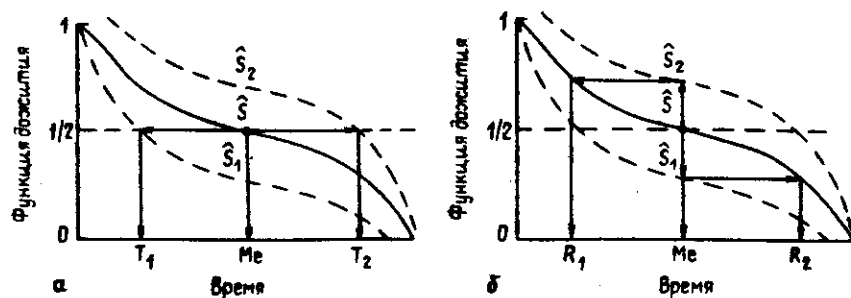


Рис. 1. Схематическое представление критерияльных (test-based) и отраженных (reflected) интервалов для медианной продолжительности жизни. \hat{S} - эмпирическая оценка функции дожития; \hat{S}_1 и \hat{S}_2 - доверительные границы для функции дожития уровня α . а) T_1 и T_2 - границы для критерияльного интервала (прямая оценка разброса по горизонтали). б) R_1 и R_2 - границы для отраженного интервала (оценка разброса по горизонтали путем отражения вертикального разброса)

гают делить существующие методы построения доверительных интервалов на так называемые "критерияльные" и "отраженные" [52]. Рис. 1 поясняет различия в принципе построения "отраженных" и "критерияльных" доверительных интервалов для медианы.

Рассмотрим четыре метода построения доверительных интервалов для медианной продолжительности жизни. В основе трех из них лежит асимптотическое неравенство [52]:

$$\left(\hat{S}(t) - \frac{1}{2} \right)^2 \leq \chi_{1,\alpha}^2 \cdot \hat{\sigma}^2 S(t), \quad (43)$$

где $\hat{S}(t)$ - оценка функции дожития, $\hat{\sigma}^2 S(t)$ - оценка дисперсии функции дожития, $\chi_{1,\alpha}^2$ - значение статистики хи-квадрат с одной степенью свободы для уровня значимости α .

Метод построения критерияльных доверительных интервалов для медианы, предложенный R. Brookmeyer и J. Crowley [15], использует для расчета дисперсии функции дожития формулу Гринвуда:

$$\hat{\sigma}_G^2(t_j) = [\hat{S}(t_j)]^2 \sum_{j=1}^l \frac{d_j}{r_j(r_j - d_j)} \quad (44)$$

Таким образом, подставив выражение (44) в неравенство (43) и задавшись уровнем значимости $\alpha = 0,05$, находим, что в 95%-ный доверительный интервал для медианной продолжительности жизни входят те моменты времени, для которых выполняется неравенство:

$$\frac{(\hat{S}(t) - \frac{1}{2})^2}{\hat{\sigma}_G^2(t)} \leq 3,84 \quad (45)$$

В нашем примере 95%-ный доверительный интервал для медианы составляет 2837354 дня (см. табл. 1).

R. Simon и J. Lee [51] рассчитывают критериальные доверительные интервалы для медианной продолжительности жизни, основываясь на оценке дисперсии функции дожития, предложенной в работе [47]:

$$\hat{\sigma}_P^2(t_i) = [\hat{S}(t_i)]^2 [1 - \hat{S}(t_i)] / r_i \quad (46)$$

где r_i - число организмов непосредственно перед i -м моментом смерти. В остальном метод Саймона и Ли аналогичен методу Брукмеера и Кроули. Следует отметить, что оценка $\hat{\sigma}_P(t_i)$ завышает истинное значение дисперсии при большом объеме выборки и большом числе цензурированных наблюдений [52].

Построение отраженных доверительных интервалов для медианы отличается от построения критериальных тем, что расчет дисперсии функции дожития проводится только для момента времени, равного медианной продолжительности жизни. Таким образом, мы приходим к "простым отраженным доверительным интервалам" для медианной продолжительности жизни, которые были предложены В. Efron [20]:

$$(\hat{S}(t) - \frac{1}{2})^2 \leq \chi_{1,\alpha}^2 \cdot \hat{\sigma}_G^2(Me), \quad (47)$$

где Me - медианная продолжительность жизни.

E. Slud и соавторы [52] предложили другой способ построения отраженных доверительных интервалов для медианной продолжительности жизни. Он основан на расчете кумулятивной интенсивности смертности или функции ресурса (см. раздел 1.2). Авторы называют эти доверительные интервалы "трансформированными отраженными доверительными интерва-

лами" [52]. Эти интервалы находятся из неравенства:

$$|\hat{\Lambda}(t) - \hat{\Lambda}(Me)|^2 \leq 4\chi_{1,\alpha}^2 \cdot \hat{\sigma}_G^2(Me), \quad (48)$$

где $\hat{\Lambda}(t)$ - функция ресурса, а Me - медианная продолжительность жизни.

Общая схема расчета доверительных интервалов для медианной продолжительности жизни по четырем рассмотренным методам приведена в табл. 6. В ней используются данные о длительности ремиссий у больных лейкозом, приведенные в табл. 2. Z_1^2, Z_2^2, Z_3^2 и Z_4^2 обозначают наблюдаемые значения величины $\chi_{1,\alpha}^2$, которые сравниваются с табличными критическими (3,84 для 95%-ного доверительного интервала и 6,61 для 99%-ного), соответственно для интервалов Эфрона, интервалов Слуда с соавторами, интервалов Брукмеера-Кроули и интервалов Саймона-Ли. Через Σ_i обозначены накопленные к i -му моменту времени суммы:

$$\sum_{j=1}^i \frac{d_j}{r_j(r_j - d_j)}$$

Расчет доверительных интервалов для медианной продолжительности жизни проводится следующим образом. Обозначим через $C_N(t)$ величины, приведенные в колонке с номером N табл. 6. Тогда наблюдаемые значения статистики хи-квадрат составляют:

$$Z_1^2(t) = C_7(t) - 4(C_2(t) - 1/2)^2 / C_4(Me) \quad \text{для метода Эфрона;}$$

$$Z_2^2(t) = C_8(t) - (C_3(t) - C_3(Me))^2 / C_4(Me) \quad \text{для метода Слуда и соавторов;}$$

$$Z_3^2(t) = C_9(t) - (C_2(t) - 1/2)^2 / C_6(t) \quad \text{для метода Брукмеера-Кроули;}$$

$$Z_4^2(t) = C_{10}(t) - (C_2(t) - 1/2)^2 / C_5(t) \quad \text{для метода Саймона-Ли}$$

Какой же метод построения доверительных интервалов для медианной продолжительности жизни наиболее предпочтителен? В работе [52] содержатся следующие рекомендации по использованию приведенных выше методов.

Схема расчета доверительных интервалов для медианной продолжительности жизни на примере данных по длительности ремеслей у больших ткацком. Взято из [52]

t_i	$\hat{S}(t_i)$	$\hat{\Lambda}(t_i)$	Σ_i	$\hat{\sigma}_P^2(t_i)$	$\hat{\sigma}_G^2(t_i)$	$Z_1^2(t_i)$	$Z_2^2(t_i)$	$Z_3^2(t_i)$	$Z_4^2(t_i)$
6	0,857	0,143	0,008	0,011	0,006	5,66	4,12	4,12	12,0
7	0,807	0,202	0,012	0,012	0,008	4,17	3,36	3,36	7,81
9*	0,807	0,202	0,012	0,013	0,008	4,17	3,36	3,36	7,46
10	0,753	0,268	0,016	0,013	0,009	2,84	2,60	2,60	5,01
11*	0,753	0,268	0,016	0,014	0,009	2,84	2,60	2,60	4,42
13	0,690	0,352	0,024	0,015	0,011	1,60	1,78	1,78	2,46
16	0,627	0,442	0,033	0,015	0,013	0,72	1,06	1,06	1,11
17*	0,627	0,442	0,033	0,016	0,013	0,72	1,06	1,06	1,04
18*	0,627	0,442	0,033	0,017	0,013	0,72	1,06	1,06	0,93
20*	0,627	0,442	0,033	0,020	0,013	0,72	1,06	1,06	0,83
22	0,538	0,585	0,057	0,020	0,016	0,06	0,31	0,31	0,07
23	0,448	0,752	0,090	0,019	0,018	0,12	0,00	0,00	0,14
25*	0,448	0,752	0,090	0,022	0,018	0,12	0,00	0,00	0,12
32*	0,448	0,752	0,090	0,028	0,018	0,12	0,00	0,00	0,10
34*	0,448	0,752	0,090	0,056	0,018	0,12	0,00	0,00	0,05
35*	0,448	0,752	0,090	0,112	0,018	0,12	0,00	0,00	0,02

Примечание. Звездочками отмечены моменты цензурирования; значение медианной продолжительности жизни подчеркнуто.

1. Метод Саймона-Ли не следует использовать в случаях, когда имеются цензурированные наблюдения.

2. Метод Брукмеера-Кроули можно использовать для построения доверительных интервалов в случае больших выборок, но не следует применять, когда выборки невелики и имеются цензурированные наблюдения.

3. Оба метода построения отраженных доверительных интервалов рекомендуется использовать в случае малых выборок при наличии цензурирования. Благодаря большей простоте расчета, эти методы (особенно метод Эфрона) являются, по мнению авторов работы [52], наиболее корректными и удобными методами расчета доверительных интервалов для медианной продолжительности жизни.

3.3. Асимметрия и эксцесс распределения продолжительности жизни

После того как мы определили величину средней и медианной продолжительности жизни, полезно оценить и такие характеристики распределения, как асимметрия и эксцесс.

Асимметрия, или скошенность распределения, может быть оценена с помощью показателя асимметрии, определяемого по формуле [6]:

$$As = \frac{\sum_{i=1}^n (t_i - \bar{t})^3 \cdot d_i}{N \cdot \hat{\sigma}_t^3}, \quad (49)$$

где As - показатель асимметрии, \bar{t} - средняя продолжительность жизни, t_i - время жизни (для таблиц смертности - середина соответствующего интервала времени), N - численность выборки, d_i - число умерших в данный момент времени (для таблиц смертности - в данном интервале времени), $\hat{\sigma}_t$ - среднее квадратичное отклонение распределения продолжительности жизни (рассчитывается по формуле (37)).

Асимметрию можно оценить также по степени отклонения средней продолжительности жизни от модальной [6]:

$$K_{As} = \frac{\bar{t} - Mo}{\hat{\sigma}_t},$$

где K_{As} - коэффициент асимметрии, \bar{t} - средняя продолжи-

тельность жизни, M_0 – модальная продолжительность жизни, $\hat{\sigma}_t$ – среднее квадратичное отклонение.

Модальная продолжительность жизни или возраст, на который приходится максимальное количество смертей, является еще одной характеристикой распределения и находится из формулы [6]:

$$M_0 = t_M + h_M \left(\frac{d_M - d_{M-1}}{2d_M - d_{M-1} - d_{M+1}} \right), \quad (50)$$

где t_M – начало интервала, в котором находится мода, h_M – длина интервала, в котором находится мода, d_{M-1} – число умерших в интервале, предшествующем модальному, d_M – число умерших в модальном интервале, d_{M+1} – число умерших в интервале, следующем за модальным. Рассчитать модальную продолжительность жизни можно только для таблиц смертности.

Показатели асимметрии могут принимать как положительные, так и отрицательные значения. Если распределение подчиняется нормальному закону, то показатель асимметрии равен нулю. На этом свойстве основан один из методов проверки нормальности исследуемого распределения [6]. Для распределения продолжительности жизни человека и других биологических видов характерна отрицательная асимметрия [3, 9].

Формула (49) дает смещенные оценки показателя асимметрии, поэтому для малых выборок рекомендуется пользоваться другой оценкой асимметрии распределения [11]:

$$\bar{A}_s = \frac{\sqrt{N(N-1)}}{N-2} A_s. \quad (51)$$

Стандартная ошибка этой оценки показателя асимметрии рассчитывается по формуле:

$$\hat{\sigma}_{\bar{A}_s} = \sqrt{\frac{6N(N-1)}{(N-2)(N+1)(N+3)}} \quad (52)$$

Для больших выборок можно пользоваться приближенной формулой:

$$\hat{\sigma}_{\bar{A}_s} = \sqrt{\frac{6}{N}}. \quad (53)$$

Эмпирические распределения продолжительности жизни могут быть как островершинными, так и плосковершинными. Величиной, характеризующей крутость распределения в области вершины, является эксцесс распределения. Рассчитать его можно по формуле:

$$E = \frac{\sum_{i=1}^n (t_i - \bar{t})^4 d_i}{N \hat{\sigma}_t^4} - 3, \quad (54)$$

где E – показатель эксцесса, \bar{t} – средняя продолжительность жизни, t_i – время жизни (или середина интервала времени для таблиц смертности), d_i – число умерших в момент t_i (для таблиц смертности – в данном интервале), N – численность выборки, $\hat{\sigma}_t$ – среднее квадратичное отклонение распределения продолжительности жизни.

Приведенная оценка эксцесса является смещенной, поэтому для малых выборок рекомендуется пользоваться несмещенной оценкой показателя эксцесса [11]:

$$\bar{E} = \frac{N^2 - 1}{(N-2)(N-3)} \left[E + \frac{6}{N+1} \right] \quad (55)$$

с соответствующей стандартной ошибкой:

$$\hat{\sigma}_{\bar{E}} = \sqrt{\frac{24N(N-1)^2}{(N-3)(N-2)(N+3)(N+5)}} \quad (56)$$

Для больших выборок можно пользоваться приближенной формулой [6]:

$$\hat{\sigma}_{\bar{E}} = \sqrt{\frac{24}{N}} = 2\sqrt{\frac{6}{N}} = 2\hat{\sigma}_{A_s}$$

Показатель эксцесса не может принимать значения меньше, чем -2 . В положительной области эксцесс может принимать сколь угодно большие значения. Показатель эксцесса нормального распределения равен нулю. На этом свойстве основан один из методов проверки нормальности распределения [6]. Для распределения продолжительности жизни человека характерен положительный эксцесс.

4. КРИТЕРИИ СРАВНЕНИЯ ВЫЖИВАЕМОСТИ В ДВУХ ВЫБОРКАХ

При проведении экспериментов в лаборатории и клинике часто бывает необходимо проверить, достоверно ли отличается выживаемость в группе, получавшей какой-либо препарат, от выживаемости в контрольной группе. На этот вопрос позволяют ответить статистические критерии, разработанные для анализа данных по выживаемости.

При сравнении двух выборок с помощью статистических критериев проверяется гипотеза о равенстве двух распределений продолжительности жизни, называемая нулевой гипотезой:

$$H_0: F_1(t) = F_2(t).$$

Эта гипотеза проверяется против альтернативной гипотезы о том, что функция распределения 1-й выборки больше, чем 2-й (так называемая односторонняя альтернатива):

$$H_1: F_1(t) > F_2(t),$$

либо о том, что функция распределения 1-й выборки не равна функции распределения 2-й выборки (двусторонняя альтернатива, так как $F_1(t)$ может быть либо больше, либо меньше, чем $F_2(t)$):

$$H_1: F_1(t) \neq F_2(t).$$

Оценку достоверности различий между двумя выборками с помощью статистического критерия начинают с выбора уровня значимости критерия, обозначаемого α . Уровень значимости определяет вероятность ошибочного принятия нулевой гипотезы, то есть принятия гипотезы об отсутствии существенных различий, когда они на самом деле имеются. Обычно задаются величиной $\alpha = 0,05$ или $\alpha = 0,01$.

Вероятность отвергнуть нулевую гипотезу, когда она неверна, характеризует так называемую мощность критерия.

Поскольку вид распределения продолжительности жизни в сравниваемых выборках, как правило, бывает неизвестен, ниже приводятся только непараметрические критерии.

4.1. Критерий Манталя-Кокса и логранговый критерий

Критерий Манталя-Кокса и логранговый критерий благодаря простоте расчета наиболее часто используются при анализе выживаемости. Следует, однако, иметь в виду, что эти критерии лучше всего применять в тех случаях, когда выполняется модель пропорциональных рисков, то есть отношение интенсивностей смертности в сравниваемых группах остается постоянным в течение всего периода наблюдения [22]:

$$\mu_1(t)/\mu_2(t) = k = \text{const.}$$

При выполнении модели пропорциональных рисков мощность этих критериев максимальна [22, 26].

Чтобы проверить, выполняется ли модель пропорциональных рисков в данном конкретном случае, необходимо для каждой выборки построить график зависимости $\ln[-\ln S(t)]$ от времени. Если построенные графики различаются на постоянную величину, то это значит, что выполняется модель пропорциональных рисков [54].

Расчет критерия N.Mantel, называемого также критерием Mantel-Haenszel, основан на построении для каждого i -го момента времени (или интервала времени) таблицы сопряженности признаков 2×2 следующего вида [41, 42, 44]:

	Умершие	Выжившие	Сумма
Группа 1	d_{1i}	$N_{1i} - d_{1i}$	N_{1i}
Группа 2	d_{2i}	$N_{2i} - d_{2i}$	N_{2i}
Сумма	d_i	$N_i - d_i$	N_i

где N_{1i} - число индивидов, подверженных риску в i -й момент времени (или в i -м интервале), d_{1i} - число умерших в i -й момент времени (или в i -м интервале) для группы 1. Для группы 2 обозначения аналогичны. Если моменты смерти и цензурирования совпадают, то предполагается, что смерть происходит несколько раньше цензурирования. Если данные сгруппированы в таблицу смертности и часть организмов уходит из-под наблюдения в i -м интервале времени, то предполагается, что этот уход происходит в конце интервала. Данное допущение отличается от предположения, используемого при построении таблиц смертности, где полагалось, что только

половина организмов, ушедших в интервале из-под наблюдения, подвержена риску смерти [18] (см. также раздел 1.2.2).

Далее, как и в случае таблиц сопряженности, рассчитывается величина критерия хи-квадрат, основанная на разности между теоретически ожидаемым и наблюдаемым числом смертей в 1-й группе.

Рассмотрим сначала пример, когда сравниваются две кривые дожития, построенные для малых выборок. В этом случае нам известны точные моменты смерти в каждой группе. Расположим их в одну колонку в порядке возрастания. В табл. 7 приведено сравнение выживаемости крыс контрольной группы (группа 1) с опытной, получавшей препарат центрофеноксин (группа 2). В две соседние колонки запишем число умерших в каждый момент времени соответственно в 1-й и 2-й группе. Для каждого i -го момента времени рассчитаем ожидаемое число умерших в 1-й группе:

$$Ed_{1i} = N_{1i}/(N_{1i}+N_{2i}), \quad (57)$$

где N_{1i} и N_{2i} - число организмов в 1-й и 2-й группе непосредственно перед i -м моментом смерти. Значения Ed_{1i} затем суммируют по всем моментам смерти. Далее, для каждого i -го момента смерти рассчитывают дисперсию числа умерших по формуле:

$$\text{var}(d_{1i}) = N_{1i}N_{2i}/(N_{1i}+N_{2i})^2. \quad (58)$$

После чего суммируют полученные для каждого момента времени значения дисперсий:

$$\sum_{i=1}^k \text{var}(d_{1i}),$$

где k - суммарное число моментов смерти.

Затем рассчитывают величину критерия N.Mantel по формуле:

$$T_M = \frac{(|\sum_{i=1}^k d_{1i} - \sum_{i=1}^k Ed_{1i}| - 1/2)^2}{\sum_{i=1}^k \text{var} d_{1i}}. \quad (59)$$

Эта величина асимптотически имеет распределение хи-квадрат с одной степенью свободы. Как известно, табличное критическое

Таблица 7

Сравнение выживаемости крыс контрольной группы (группа 1) и крыс, получавших препарат центрофеноксин (группа 2)

Время жизни, дни	d_{1i}	d_{2i}	N_{1i}	N_{2i}	N_i	Ed_{1i}	Ed_{2i}
1	2	3	4	5	6	7	8
2	1	0	41	29	70	0,586	0,414
31	1	0	40	29	69	0,580	0,420
45	1	1	39	29	68	0,574	0,426
88	1	0	38	28	66	0,576	0,424
115	1	0	37	28	65	0,431	0,431
122	1	0	36	28	64	0,563	0,438
128	1	0	35	28	63	0,556	0,444
142	1	1	34	28	62	0,548	0,452
213	1	1	33	27	60	0,550	0,450
243	1	0	32	26	58	0,552	0,448
251	1	0	31	26	57	0,544	0,456
256	1	1	30	26	56	0,536	0,464
262	1	0	29	25	54	0,537	0,463
269	1	0	28	25	53	0,528	0,472
283	1	1	27	25	52	0,519	0,481
288	1	0	26	24	50	0,520	0,480
291	1	0	25	24	49	0,510	0,490
297	1	1	24	24	48	0,500	0,500
301	1	0	23	23	46	0,500	0,500
304	1	0	22	23	45	0,489	0,511
309	1	0	21	23	44	0,477	0,523
313	1	1	20	23	43	0,465	0,535

2843

Продолжение табл. 7

1	2	3	4	5	6	7	8
320	1	0	19	22	41	0,463	0,537
327	1	1	18	22	40	0,450	0,550
335	0	1	17	21	38	0,447	0,553
340	1	1	17	20	37	0,459	0,541
354	1	1	16	19	35	0,457	0,543
360	1	0	15	18	33	0,455	0,545
362	0	1	14	18	32	0,438	0,563
367	1	1	14	17	31	0,452	0,548
375	1	0	13	16	29	0,448	0,552
381	1	0	12	16	28	0,429	0,571
395	1	1	11	16	27	0,407	0,593
400	1	1	10	15	25	0,400	0,600
409	1	1	9	14	23	0,391	0,609
418	0	1	8	13	21	0,381	0,619
423	1	1	8	12	20	0,400	0,600
426	1	0	7	11	18	0,389	0,611
437	1	1	6	11	17	0,353	0,647
451	1	1	5	10	15	0,333	0,667
459	1	0	4	9	13	0,308	0,692
465	1	1	3	9	12	0,250	0,750
470	0	1	2	8	10	0,200	0,800
479	1	1	2	7	9	0,222	0,778
492	0	1	1	6	7	0,143	0,857
496	1	0	1	5	6	0,167	0,833
500	0	1	0	5	5	0,000	1,000
506	0	1	0	4	4	0,000	1,000

Продолжение табл. 7

1	2	3	4	5	6	7	8
515	0	1	0	3	3	0,000	1,000
521	0	1	0	2	2	0,000	1,000
550	0	1	0	1	1	0,000	1,000
Сумма:	41	29				20,483	30,381

Примечание: Рассчитано на основании данных С.И. Потапова

кое значение хи-квадрат для уровня значимости $\alpha = 0,05$ равно 3,84. Таким образом, если рассчитанное значение хи-квадрат превышает критическое (например, 3,84 для $\alpha = 0,05$), то можно утверждать, что различия в выживаемости сравниваемых групп являются значимыми. В нашем примере (табл. 7)

$\sum E d_{1i} = 20,483$, $\sum \text{var } d_{1i} = 10,800$. Следовательно,

$$T_M = \frac{(|41 - 20,483| - 1/2)^2}{10,800} = 37,10$$

Полученная величина T_M , равная 37,10, говорит о высокой достоверности наблюдаемых различий в выживании ($\alpha \ll 0,001$).

Если данные сгруппированы в виде таблицы смертности (с учетом приведенного выше замечания относительно цензурированных наблюдений), то величины $E d_{1i}$ и $\text{var } d_{1i}$ рассчитываются несколько иначе. В этом случае ожидаемое число умерших в i -м интервале для 1-й группы рассчитывается по формуле:

$$E d_{1i} = N_{1i}(d_{1i} + d_{2i}) / (N_{1i} + N_{2i}), \quad (60)$$

а дисперсия числа умерших в 1-й группе определяется следующим образом:

$$\text{var } d_{1i} = N_{1i} N_{2i} (d_{1i} + d_{2i}) (N_{1i} - d_{1i} - d_{2i}) / N_i^2 (N_i - 1). \quad (61)$$

Расчет критерия, предложенного D. Cox [17], в принципе не отличается от описанного выше. Сначала по формуле (57) рассчитывают ожидаемое число умерших в 1-й группе Ed_{1i} . Затем определяют величину статистики $U(0)$:

$$U(0) = N_1 - \sum_i^k d_i Ed_{1i} \quad (62)$$

где N_1 — исходная численность 1-й группы, а d_i — суммарное число умерших в обеих группах в i -й момент времени, и величину дисперсии $I(0)$:

$$I(0) = \sum_i^k \frac{d_i(N_i - d_i)Ed_{1i}(1 - Ed_{1i})}{(N_i - 1)} \quad (63)$$

Отношение, равное $U(0)/\sqrt{I(0)}$, асимптотически имеет стандартное нормальное распределение, и критические значения для него находятся из таблиц нормального распределения. Например, для уровня значимости $\alpha = 0,05$ критическое значение равно 1,96.

Наиболее часто на практике пользуются аппроксимацией критерия Мантеля, которая получила название логрангового критерия (Log rank test) [36, 44]. Если сравниваются две группы, то этот критерий можно записать следующим образом:

$$T_{LP} = \frac{(\sum_i^k d_{1i} - \sum_i^k Ed_{1i})^2}{\sum_i^k Ed_{1i}} + \frac{(\sum_i^k d_{2i} - \sum_i^k Ed_{2i})^2}{\sum_i^k Ed_{2i}} \quad (64)$$

где Ed_{1i} и Ed_{2i} рассчитываются по формуле (60) соответственно для 1-й и 2-й групп. Величина T_{LP} асимптотически имеет распределение хи-квадрат с одной степенью свободы. В нашем примере (табл. 7) логранговый критерий рассчитывается следующим образом:

$$T_{LP} = \frac{(41 - 20,483)^2}{20,483} + \frac{(29 - 30,381)^2}{30,381} = 20,61.$$

Таким образом, и в этом случае подтверждается высокая достоверность различий в выживании между контролем и опытом ($\alpha \ll 0,001$).

В работе [45] было показано, что всегда $T_{LP} \leq T_M$. Таким образом, если логранговый критерий не выявляет различий в выживаемости, следует применять критерий Мантеля.

Иногда при сравнении двух групп исследователь может столкнуться с ситуацией, когда наблюдаемое число смертей столь мало, что встает вопрос о правомерности использования асимптотического критерия. В этих случаях стандартный логранговый критерий часто заменяется на логранговый перестановочный критерий [31]. Рассчитывается он следующим образом. Сначала вычисляют величину статистики $U_p(0)$:

$$U_p(0) = \sum_{i=1}^k (d_{1i} - Ed_{1i}), \quad (65)$$

где Ed_{1i} определяется по формуле (57). Расчет дисперсии этой статистики производится иначе, чем в предыдущих случаях:

$$I_p(0) = \frac{N_1 N_2}{N(N-1)} (k - \sum_{i=1}^k N_i^{-1}), \quad (66)$$

где k — суммарное число моментов времени t_i , в которые произошла смерть. Отношение $U_p(0)/\sqrt{I_p(0)}$ асимптотически имеет стандартное нормальное распределение. Полученное отношение $U_p(0)/\sqrt{I_p(0)}$ сравнивается затем с табличными критическими значениями из таблиц нормального распределения.

Рассмотренные выше варианты критерия Мантеля-Кокса рекомендуется применять в тех случаях, когда выполняется модель пропорциональных рисков, а также тогда, когда различия между кривыми выживания наиболее выражены в конце периода наблюдения [22].

4.2. Критерий Гехана-Бреслоу

Другой критерий, который часто используется в анализе выживаемости, — это обобщенный критерий Вилкоксона, впервые предложенный в 1965 г. E. Gehan [24]. Расчет критерия Гехана для сравнения полных выборок (в отсутствие цензурированных наблюдений) проводится следующим образом. Сначала вычисляют статистику W :

$$W = \sum_{i=1}^k (N_i d_{1i} - N_{1i} d_i) \quad (67)$$

и ее дисперсию [24]:

$$\text{var}(W|H_0) = \frac{1}{3} N_1 N_2 (N_1 + N_2 + 1). \quad (68)$$

Здесь используются те же обозначения, что и при расчете критерия Мантеля. Величина T_G , равная $W^2/\text{var}(W|H_0)$, асимптотически имеет распределение хи-квадрат с одной степенью свободы [40]. Если рассчитанное значение T_G превышает критическое, взятое из таблиц распределения хи-квадрат, то гипотезу о равенстве двух кривых выживания отбрасывают.

В нашем примере (см. табл. 7) величина статистики W равна 472, а $\text{var}(W|H_0) = 28139,7$. Таким образом, $T_G = (472)^2/28139,7 = 7,92$, что говорит о достоверности наблюдаемых различий ($\alpha < 0,005$).

Первоначальный вариант критерия Гехана для сравнения цензурированных выборок [24] требовал довольно громоздких расчетов и впоследствии был существенно упрощен. N. Breslow [14] удалось модифицировать этот критерий так, что он стал независим относительно распределений моментов цензурирования в сравниваемых выборках (в то время как критерий Мантеля предполагает равенство этих распределений).

Расчет критерия Гехана-Бреслоу для цензурированных выборок можно представить следующим образом [36]. Для каждой из сравниваемых групп определяют величины W и G по формулам:

$$\begin{aligned} W_1 &= \sum_{i=1}^k (N_i d_{1i} - N_{1i} d_i) \\ W_2 &= -W_1 = \sum_{i=1}^k (N_i d_{2i} - N_{2i} d_i) \\ G_1 &= \sum_{i=2}^k \{ N_i d_i (N_i - d_i) N_{1i} / N_{i-1} \} \\ G_2 &= \sum_{i=2}^k \{ N_i d_i (N_i - d_i) N_{2i} / N_{i-1} \}. \end{aligned} \quad (69)$$

Критерий Гехана-Бреслоу строится как сумма:

$$T_{GB} = \frac{W_1^2}{G_1} + \frac{W_2^2}{G_2}, \quad (70)$$

которая асимптотически имеет распределение хи-квадрат с одной степенью свободы. Рассчитанное значение критерия затем сравнивается с табличным для данного уровня значимости.

Для малых выборок, когда в каждый момент времени происходит не более одной смерти ($d_{1i} + d_{2i} = 1$ для всех $i = 1, k$), расчет критерия Гехана-Бреслоу упрощается [40]:

$$W = \sum_{i=1}^k (N_i d_{1i} - N_{1i}) \quad (71)$$

$$\text{var}(W|H_0) = \sum_{i=1}^k N_{1i} \cdot N_{2i} \quad (72)$$

Отношение $W^2/\text{var}(W|H_0)$ асимптотически имеет распределение хи-квадрат с одной степенью свободы и сравнивается с табличным критическим значением.

Критерий Гехана и критерий Гехана-Бреслоу рекомендуется применять в тех случаях, когда различия в кривых дожития наиболее выражены в начальный период наблюдения, а также в случаях, когда не выполняется модель пропорциональных рисков [22].

4.3. Критерий Колмогорова-Смирнова

Критерий Колмогорова-Смирнова, в отличие от предыдущих критериев, чувствителен к любым различиям в форме сравниваемых распределений, а не только к различиям в средних тенденциях [5]. В одной из недавних работ [43] критерий Колмогорова-Смирнова рекомендовался для сравнения результатов геронтологических экспериментов.

Расчет критерия Колмогорова-Смирнова для сравнения полных выборок довольно прост. Сначала рассчитывают величину m :

$$m = \frac{N_1 N_2}{N_1 + N_2}, \quad (73)$$

где N_1 и N_2 - соответственно численности 1-й и 2-й сравниваемых групп. Затем рассчитывают оценки функции дожития в 1-й и 2-й группе для каждого i -го момента времени по формуле (9) или для каждого i -го интервала в случае таблиц смертности по формуле (21): $\hat{S}_1(t)$ и $\hat{S}_2(t)$. На основании полученных оценок для каждого момента времени (или для

каждого интервала) определяют разность между функциями дожития в 1-й и 2-й группе. Из полученных разностей выбирают максимальную по абсолютной величине:

$$V = \max_{t_i} |\hat{S}_1(t) - \hat{S}_2(t)|. \quad (74)$$

После чего рассчитывают значение критерия Колмогорова-Смирнова:

$$T_{КС} = \sqrt{m} \cdot V. \quad (75)$$

Если $T_{КС}$ превышает критическое значение, взятое из таблицы распределения Колмогорова-Смирнова для данного уровня значимости, то гипотезу о равенстве кривых дожития отвергают. Ниже приводятся критические значения $T_{КС}$ для разных уровней значимости α :

α	0,2	0,1	0,05	0,02	0,01
$T_{КС}$	1,07	1,22	1,36	1,52	1,63

Приведенные критические значения $T_{КС}$ справедливы для выборок, размер которых превышает 20 [43].

В нашем примере (табл. 7) расчет разностей между $\hat{S}_1(t)$ и $\hat{S}_2(t)$ дает максимальную величину, равную 0,3197. Величина \sqrt{m} при этом равна: $\sqrt{41 \cdot 29 / 70} = \sqrt{16,99} = 4,1214$. Таким образом, величина критерия равна: $4,1214 \cdot 0,3197 = 1,3176$. Сравнивая это значение с табличным, делаем вывод, что различия между опытом и контролем являются значимыми на уровне, меньшем 0,1, но не на уровне, меньшем 0,05.

Полученный результат отличается от результата применения к этим данным критерия Мантеля, где были обнаружены высокозначимые различия, и отражает одно из свойств критерия Колмогорова-Смирнова - его сравнительно низкую мощность. Это означает, что вероятность принятия неверной нулевой гипотезы о равенстве кривых дожития у этого критерия выше, чем у предыдущих.

С целью повышения мощности критерия Колмогорова-Смирнова и распространения его на анализ неполных данных был разработан модифицированный критерий Колмогорова-Смирнова [22]. Модифицированный критерий Колмогорова-Смирнова

для анализа полных данных основан на расчете следующего выражения:

$$\left| \frac{1}{2} \cdot (\hat{S}_1(t) + \hat{S}_2(t)) \cdot (\hat{\Lambda}_1(t) - \hat{\Lambda}_2(t)) \right|,$$

где $\hat{\Lambda}_1(t)$ и $\hat{\Lambda}_2(t)$ - функции ресурса (накопленные интенсивности смертности), рассчитываемые по формуле (19).

Величина V в отличие от формулы (74) определяется как

$$V = \max_{t_i} \left| \frac{1}{2} \cdot (\hat{S}_1(t) + \hat{S}_2(t)) \cdot (\hat{\Lambda}_1(t) - \hat{\Lambda}_2(t)) \right|,$$

а критерий Колмогорова-Смирнова, как и в предыдущем случае, рассчитывается по формуле (75).

Рассчитанный этим способом модифицированный критерий Колмогорова-Смирнова для данных, приведенных в табл. 7, дает величину, равную 1,4549. Таким образом, модифицированный критерий Колмогорова-Смирнова позволил выявить различия в выживаемости на уровне $\alpha < 0,05$, что свидетельствует о несколько большей мощности этого критерия по сравнению с обычным критерием Колмогорова-Смирнова.

Модификация критерия Колмогорова-Смирнова, предназначенная для анализа цензурированных данных, подробно освещена в работе [22] и здесь не приводится.

Достоинством рассмотренного выше критерия является то, что он способен выявлять любые различия в форме сравнимых распределений. В частности, методом Монте-Карло было показано, что модифицированный критерий Колмогорова-Смирнова хорошо выявляет различия в кривых выживания, наиболее выраженные в середине периода наблюдения, в то время как критерий Гехана и критерий Мантеля не способны выявить подобные различия [22]. Таким образом, рассмотренный критерий можно рекомендовать к использованию в тех случаях, когда кривые выживания различаются не на всем интервале наблюдения, а лишь в отдельных его участках, и в случае пересекающихся кривых дожития [22].

ЗАКЛЮЧЕНИЕ

В последние годы появилось так много различных методов анализа выживаемости, что, по-видимому, для их полного освещения необходимо издать специальную монографию, подобную тем, что уже опубликованы за рубежом [28, 33, 39]. Здесь же были освещены только те методы, которые имеют наибольший интерес для геронтологов в соответствии с их специфическими задачами. Несомненно, что использование всех имеющихся к настоящему времени методов анализа выживаемости позволит извлечь существенно больше информации из экспериментальных данных, чем это делалось до сих пор.

ЛИТЕРАТУРА

1. Барлоу Р., Прошан Ф. Статистическая теория надежности и испытания на безотказность. М., Наука, 1984, 328 с.
2. Беляев Ю.К. Непараметрические методы в задачах обработки результатов испытаний и эксплуатации. М., Знание, 1984, 113 с.
3. Гаврилов Л.А., Гаврилова Н.С. Биология продолжительности жизни: количественные аспекты. М., Наука, 1986, 168 с.
4. Гнеденко Б.В., Хинчин А.Я. Элементарное введение в теорию вероятностей. М., Наука, 1976, 167 с.
5. Гублер Е.В. Вычислительные методы анализа и распознавания патологических процессов. М., Медицина, 1978, 294 с.
6. Зайцев Г.Н. Методика биометрических расчетов. М., Наука, 1973, 256 с.
7. Кадырова Н.О., Антипова Т.В., Жербин Е.А., Яковлев А.Ю. - Вопр. онкологии, 1984, 30, № 7, с. 24-32.
8. Кадырова Н.О., Яковлев А.Ю. - Там же, № 8, с. 35-48.
9. Семенова В.Г. - В кн.: Проблемы биологии старения. М., Наука, 1983, с. 56-62.
10. Семенова В.Г., Гаврилов Л.А. - В сб.: IV Всесоюзный съезд геронтологов и гериатров. Тезисы докл. Киев, Ин-т геронтологии, 1982, с. 345
11. Урбах В.Ю. Математическая статистика для биологов и медиков. М., Изд-во АН СССР, 1963, 323 с.
12. Шылау В.В., Лабанок Л.М., Кухарова Л.В. - Вестн. АН БССР, сер. биол., 1986, № 1, с. 84-86.

13. Ayiomamitis A. - Int. J. Bio-Medical Computing, 1986, 19, p.109-117.
14. Breslow N.E. - Biometrika, 1970, 57, p. 579-594.
15. Brookmeyer R., Crowley J. - Biometrics, 1982, 38, p. 29-41.
16. Chiang Ch.L. Life table and mortality analysis. Geneva, WHO, 1978, 399 pp.
17. Cox D.R. - J. Roy. Statist. Soc. B, 1972, 34, p. 189-220.
18. Cutler S.J., Ederer F. - J. Chron. Diseases, 1958, 8, p. 699-712.
19. Davis C.E. - Comm. Statist., Ser. A, 1978, 7, p. 389-398.
20. Efron B. - J. Amer. Statist. Assoc., 1981, 76, p. 312-319.
21. Engeman R.M., Keefe T.J. - Comput. and Biomed. Res., 1985, 18, p. 391-396.
22. Fleming T.R., O'Fallon J.R., O'Brien P.C. - Biometrics, 1980, 36, p. 607-625.
23. Freireich E.J., Gehan E., Frei E., et al. - Blood, 1963, 21, p. 699-716.
24. Gehan E.A. - Biometrika, 1965, 52, p. 650-653.
25. Gehan E.A. - J. Chron. Diseases, 1969, 21, p. 629-644.
26. Gehan E.A. - In: Cancer Therapy: Prognostic factors and criteria of response. N.Y., Raven Press, 1975, p. 7-35.
27. Gehan E.A., Siddiqui M. - J. Amer. Statist. Assoc., 1973, 68, p. 848-856.
28. Gross A., Clark V. Survival Distributions: Reliability Applications in the Biomedical Sciences. N.Y., Wiley, 1975, 255 pp.
29. Hall W.G., Wellner J.A. - Biometrika, 1980, 67, p. 133-143.
30. Hoel D.G., Walburg H.E. - J. Nat. Cancer Inst., 1972, 49, p. 361-372.
31. Jennrich R.I. - Biometrika, 1983, 70, p. 133-137.
32. Joe H., Koziol J.A., Petkau A.J. - Biometrics, 1981, 37, p. 327-340.
33. Kalbfleisch J.D., Prentice R.L. The Statistical Analysis of Failure Time Data. N.Y., Wiley, 1980, 326 pp.
34. Kaplan E.L., Meier P. - J. Amer. Statist. Assoc., 1958, 53, p. 457-481.
35. Koziol J.A., Petkau A.J. - Biometrika, 1978, 65, p. 615-623.
36. Köpcke W. - Medizinische Informatik und Statistik, 1981, 33, S. 289-298.
37. Kunstyr L., Leuenberger H.-G. - J. Gerontol., 1975, 30, p.157-162.
38. Lagakos S.W. - Biometrics, 1979, 35, p. 139-156.
39. Lee E.T. Statistical Methods for Survival Data Analysis.

- Lifetime Learning Publications, Belmont, 1980, 186 pp.
40. Lininger L., Gail M.H., Green S.B., Byar D.P. - *Biometrika*, 1979, 66, p. 419-428.
 41. Mantel N. - *Cancer Chemother. Rept.*, 1966, 50, p. 163-170.
 42. Mantel N., Haenszel W. - *J. Nat. Cancer Inst.*, 1959, 22, p. 719-748.
 43. Mode C.J., Ashleigh R.D., Zawodniak A. et al. - *J. Gerontol.*, 1984, 39, p. 36-42.
 44. Peto R., Peto J. - *J. Roy. Statist. Soc. A*, 1972, 135, p. 185-206.
 45. Peto R., Pike M.C. - *Biometrics*, 1973, 29, p. 579-584.
 46. Peto R., Pike M.C., Cox D.R. et al. - *Brit. J. Cancer*, 1976, 34, p. 585-612.
 47. Peto R., Pike M.C., Cox D.R. et al. - *Ibid.*, 1977, 35, p. 1-39.
 48. Sacher G.A. - *Radiology*, 1956, 67, p. 250-257.
 49. Sacher G.A. - In: *Handbook of the Biology of Aging*. N.Y., Reinhold, 1977, p. 582-638.
 50. Sharma S.P., Wadhwa R. - *Mech. Ageing and Develop.*, 1983, 23, p. 67-71.
 51. Simon R., Lee J.J. - *Cancer Treat. Reports*, 1982, 66, p. 37-42.
 52. Slud E.V., Byar D.P., Green S.B. - *Biometrics*, 1984, 40, p. 587-600.
 53. Thomas D.R., Grunkmeier G.L. - *J. Amer. Statist. Assoc.*, 1975, 70, p. 865-871.
 54. Wickramaratne P.J., Prusoff B.A., Merikangas K.R., Weissman M.M. - *J. Chron. Diseases*, 1986, 39, p. 389-397.