СТОХАСТИЧЕСКИЕ МОДЕЛИ ФУНКЦИОНИРОВАНИЯ ЭЛЕМЕНТОВ И ОРГАНИЗМОВ

1.1 Вероятностный закон распределения длительности функционирования и жизни на индивидуальном уровне

Все рассмотренные в настоящей главе законы, модели, формулы связаны с изучением события выхода объекта из фиксированного состояния. Этим состоянием может быть состояние работоспособности технической системы, безотказного функционирования электронного элемента, безремонтной эксплуатации самолёта при заданном уровне надёжности. Для живого организма это может быть состоянием оставаться живым, быть здоровым, физически активным. Время нахождения в этих состояниях соответствует времени работы до отказа, времени безремонтной эксплуатации. В демографии и медицине эти времена соответствуют продолжительности жизни, периоду здоровой жизни, длительности активной жизни. Вероятностные характеристики длительности пребывания в состоянии, такие как закон распределения вероятности, условная вероятность выхода из фиксированного состояния, математическое ожидание изучаются в теории надёжности и демографии. Выявление факторов риска, влияющих на эти характеристики в человеческих популяциях, является предметом эпидемиологии. В компетенции профилактической медицины находится изучение того, как профилактические медицинские мероприятия отразятся в структуре смертности, повлияют на заболеваемость и демографические характеристики населения.

С формальной точки зрения длительность пребывания в состоянии, обозначим её T, является случайной величиной, определённой на неотрицательной полуоси. Можно говорить о функции распределения, функции распределения плотности вероятностей, моментах случайной величины T. Однако, в теории надёжности, демографии, биометрике существует традиция характеризовать вероятностные свойства длительности пребывания в состоянии через *функцию дожития*, *кумулятивный риск*, *функцию интенсивности отказа* или *силы смерти*. Это связано с содержательной стороной решаемых задач, в которых основной интерес представляют выводы о характеристиках процесса при условии, что объект всё ещё функционирует. Так, важно оперировать вероятностью выхода из строя системы в ближайшем будущем, если она ещё работает или знать вероятность, что человек проживёт ещё столько–то дополнительных лет. В последнем случае сама формулировка вопроса указывает на то, что изучается условная вероятность наступления смерти.

Пусть случайная величина T имеет функцию распределения $F(t)$. Функция дожития для T определяется как вероятность оставаться в заданном состоянии не менее заданного времени

$$S(t) = P\{T \geq t\}.$$

Очевидно соотношение между функцией дожития и функцией распределения случайной величины T

$$S(t) = 1 - P\{T < t\} = 1 - F(t).$$

Из вложенности событий $\{T \geq t_0\} \supset \{T \geq t_1\}$ при $t_0<t_1$ следует важная формула для условной функции дожития

$$S(t_1|t_0) = P\{T \geq t_0, T \geq t_1\} / P\{T \geq t_0\}$$
$$= P\{T \geq t_1\} / P\{T \geq t_0\}$$
$$= S(t_1) / S(t_0).$$

Функция плотности распределения вероятностей случайной величины T равна

$$f(t) = \frac{d}{dt} F(t) = -\frac{d}{dt} S(t) .$$

Запишем математическое ожидание величины T в виде

$$\mathrm{E}(T) = \int_0^\infty t f(t) dt = -\int_0^\infty t dS(t) = -tS(t)|_0^\infty + \int_0^\infty S(t) dt .$$

Поскольку вероятность бесконечно долго находиться в работоспособном состоянии предполагается равной нулю, то справедливо выражение

$$\mathrm{E}(T) = \int_0^\infty S(t) dt ,$$

которое означает, что среднее время безотказной работы равно площади, ограниченной кривой $S(t)$ и положительной полуосью абсцисс.

Рассмотрим вероятность выхода системы из состояния за промежуток времени Δt

$$P\{T < t + \Delta t \,|\, T \geq t\} = 1 - S(t + \Delta t) / S(t)$$

$$= -\frac{S(t+\Delta t) - S(t)}{S(t)}$$

$$= -\frac{\Delta S(t)}{S(t)}.$$

При Δt стремящемся к нулю эта вероятность с точностью до $o(\Delta t)$ представима в виде $P\{T < t + \Delta t | T \geq t\} = m(t)\Delta t$, где функция $m(t)$ называется в теории надёжности интенсивностью отказа, в демографии силой смерти, и задаётся выражением

$$m(t) = -\frac{1}{S(t)} \frac{d}{dt} S(t).$$

Рассмотрим последнее выражение как дифференциальное уравнение относительно *S(t)* с начальным условием *S(0)*=1, которое означает, что в начальный момент времени система находилась в работоспособном состоянии или что наблюдаемый человек был жив. Тогда функция дожития представима в виде

$$S(t) = \exp\left(-\int_0^t m(t)dt\right)$$

$$= \exp(-H(t)).$$

Функция *H(t)* – интеграл от интенсивности отказа, называется кумулятивным риском и имеет смысл накопленного к моменту *t* риска выхода системы из строя или смерти человека. Условная функция дожития принимает вид

$$S(t|t_0) = \exp(H(t_0) - H(t)).$$

Функция плотности распределения вероятностей длительности пребывания в состоянии с учетом функции интенсивности $m(t)$ запишется в виде

$$f(t) = m(t)S(t) = m(t)\exp\left(-\int_0^t m(t)dt\right).$$

1.2 Математические модели для описания интенсивности отказа

Ниже приводятся различные математические модели, применяемые для описания вероятностных свойств процессов сохранения работоспособности технической системы и дожития человека. Модели различаются предположениями о виде функции $m(t)$, которые отражают особенности взаимодействия с окружающей средой и наличие сопутствующих процессов. В частности, старение организма описывается как увеличение силы смерти $m(t)$ с возрастом.

Постоянная интенсивность

Простейшей моделью для описания процесса потери работоспособности является модель, в которой интенсивность отказа не зависит от времени работы системы. Для организма это означает, что вероятность наступления смерти не зависит от возраста, то есть организм не стареет. В этом случае $m(t) = m$, кумулятивный риск $H(t) = m \times t$, функция дожития принимает вид экспоненты

$$S(t) = \exp(-m \times t).$$

Среднее время безотказной работы, или средняя продолжительность жизни задаётся выражением $\mathrm{E}(T) = 1/m$.

Модель Гомперца

Отсутствие старения объясняет, почему модель постоянной интенсивности нашла широкое применение в теории надёжности технических систем, а не в демографии или биометрике. Допустим, что постоянна не интенсивность отказа, а относительный темп её изменения

$$\frac{1}{m(t)}\frac{d}{dt}m(t) = b \ .$$

В этом случае интенсивность отказа описывается выражением

$$m(t) = a \exp(b \times t) \,, \tag{1.1}$$

а кумулятивный риск и функция дожития принимают вид

$$H(t) = \frac{a}{b}\left(\exp\left(b \times t\right) - 1\right),$$

$$S(t) = \exp\left(\frac{a}{b}\left(1 - \exp(b \times t)\right)\right).$$

Впервые выражение (1.1) для силы смерти ввел Бенжамин Гомперц [7], опубликовавший в 1825 году работу в которой показал, что кривые логарифма силы смерти среди людей в возрастах старше 40 лет хорошо аппроксимируются прямыми линиями. С тех пор за соотношением (1.1) закрепилось название модели Гомперца. Параметр α в (1.1) интерпретируется как начальная смертность, параметр β

как темп старения организма, поскольку он являет масштабирующим множителем для календарного времени t.

Удивительным является то, что кривые смертности следуют модели Гомперца для различных видов живых организмов: млекопитающих, мух, червей. Этот факт интенсивно исследуется в настоящее время, выдвинуты различные теории для его объяснения. В частности, предложена общая теория старения и смерти [8], связывающая смертность с интенсивностью негативных стохастических воздействий на организм и с потерей организмом способности выдерживать эти воздействия с возрастом. Результирующая модель приводит к соотношению (1.1).

Модель Гомперца - Мейкхема

Симбиозом модели с постоянной интенсивностью и модели Гомперца является модель Гомперца–Мейкхема, которая получается если на фоне процесса дожития, подчиняющегося модели Гомперца, действуют дополнительные независимые факторы, приводящие к смерти с постоянной интенсивностью. Обозначим продолжительность жизни в рамках модели Гомперца через T_1, а через T_2 – продолжительность жизни при постоянной интенсивности смерти. Результирующая продолжительность жизни равна наименьшей из двух продолжительностей $T = \min(T_1, T_2)$. При условии независимости случайных величин T_1 и T_2 запишем функцию дожития

$$S(t) = P\{\min(T_1, T_2) \geq t\} = P\{T_1 \geq t\} P\{T_2 \geq t\}$$

$$= \exp\left(\frac{a}{b}\left(\exp(b \times t) - 1\right)\right)\exp\left(l \times t\right)$$

$$= \exp\left(-\int_0^t \left(l + a\exp(b \times t)dt\right)\right),$$

которая соответствует процессу, имеющему функцию интенсивности отказа

$$m(t) = l + a\exp(bt). \qquad (1.2)$$

Первое слагаемое в выражении (1.2) соответствует интенсивности смерти, связанной с внешними причинами, не зависящими от возраста организма. Эту смертность могут определять условия содержания лабораторных животных, для людей основную долю подобной смертности составляет смертность от несчастных случаев.

Модель Вейбулла

Модель Гомперца–Мейкхема используется при анализе дожития живых организмов. Для учёта зависимости интенсивности отказа от времени работы технических систем часто используется модель Вейбулла. Постоянной интенсивности отказа соответствует экспоненциальная функция дожития. Кумулятивный риск при этом задаётся линейной функцией времени. Естественным обобщением этой модели является степенная зависимость в виде $H(t) = kt^n$, которой соответствует интенсивность отказа вида

$$m(t) = knt^{n-1} \qquad (1.3)$$

и функция дожития

$$S(t)=\exp\left(-kt^n\right).$$

Выражение (1.3) называют моделью Вейбулла для интенсивности отказов. Математически эта модель соответствует интенсивности отказа в системе, составленной из большого числа элементов, если отказ любого из них ведёт к выходу из строя всей системы [9]. Параметр n при этом определяет порядок падения функции дожития каждого из элементов системы. Модель Вейбулла нашла широкое применение при решении задач эконометрики, анализе надёжности радиоэлектронной аппаратуры.

1.3 Оценивание параметров моделей методом максимального правдоподобия

При изучении связи риска отказа или смерти с различными факторами необходимо строить оценки параметров соответствующих моделей. Специфика изучаемых процессов заключается в том, что эксперимент проводится на группе идентичных объектов или организмов и длительность эксперимента, как правило, ограничена во времени из–за высокой стоимости его проведения. В результате, при проведении эксперимента не все объекты выходят из строя и не все организмы гибнут, часть из них сохраняет жизнеспособность. Кроме того, некоторые объекты могут быть сняты с испытаний в работоспособном состоянии до окончания эксперимента, а некоторые организмы просто убежать. Такие данные называются цензурированными [14]. Метод оценивания параметров модели должен учитывать это обстоятельство максимально используя имеющуюся экспериментальную информацию. При решении задач надежности и в биометрике в качестве такого метода применяется метод максимального правдоподобия, заключающийся в максимизации по параметрам модели вероятности или плотности вероятностей наблюдать именно те данные, которые наблюдаются в эксперименте.

Индивидуальные данные

Пусть результаты наблюдений представляют собой записи времени отказа каждого объекта или возраста гибели каждого организма и признак является ли он цензурирован. Если объект или организм цензурирован, то вместо времени отказа или возраста гибели регистрируется длительность его наблюдения. Запишем эти данные в виде $\{t_1, d_1, \ldots, t_n, d_n\}$, где n – число объектов или организмов, $d_i = 1$ если объект i не цензурирован и 0 в противном случае. Если каждая пара (t_i, d_i) статистически независима друг от друга и механизм цензурирования статистически независим от длительности работы объекта или возраста организма, то функция правдоподобия представляет собой произведение плотности вероятностей наблюдения моментов отказа или смерти среди нецензурированных объектов на вероятность проработать или прожить наблюдённые времена для цензурированных объектов

$$L = \prod_{d_i=1} m(t_i) S(t_i) \prod_{d_i=0} S(t_i)$$

$$=\prod_{d_i=1} m(t_i)\prod_i S(t_y)$$

На практике оценки параметров модели строят максимизируя не функцию правдоподобия, а её логарифм, что теоретически одно и то же. Используя представление функция дожития через функцию интенсивности запишем выражение для логарифма функции правдоподобия

$$\ln L = \sum_{d_i=1} \ln m(t_i) + \sum_i \ln S(t_i)$$

$$= \sum_{d_i=1} \ln m(t_i) - \sum_i \int_0^{t_i} m(t)dt .$$

В формуле первая сумма берётся только по нецензурированным объектам, а вторая по всей совокупности объектов.

Для модели с постоянной интенсивностью логарифм правдоподобия принимает вид

$$\ln L = n_1 \ln m - m\sum_i t_i$$

и достигает максимального значения при

$$m = \frac{n_1}{\sum_i t_i} = \frac{n_1}{n}\bar{t}^{-1},$$

где n_1 – число нецензурированных объектов, $\bar{t} = \frac{1}{n}\sum_i t_i$ – среднее время наблюдения объектов или организмов как цензурированных, так и нецензурированных. Если цензурирование отсутствует, то оценка интенсивности отказа или силы смерти обратно пропорциональна среднему времени наработки или среднему времени жизни, зарегистрированному в популяции.

Оценка параметров моделей с интенсивностью, зависящей от возраста, требует привлечения более сложных аналитических и численных методов для максимизации логарифма функции правдоподобия и выходит за рамки настоящей публикации.

Сгруппированные данные

Часто при проведении биологических экспериментов не удаётся регистрировать индивидуальные времена жизни, как, например, в экспериментах с большими популяциями мух, червей, когда невозможно регулярно осматривать каждый организм. В такой ситуации регистрируют не время жизни отдельных организмов, а число умерших и цензурированных особей на заданных интервалах времени. В этом случае вид функции правдоподобия изменится, но принцип построения оценок параметров модели остаётся прежним – максимизация функции правдоподобия.

Пусть результаты эксперимента представляют собой набор чисел (t_i, d_i, n_i), $(i = 1, \dots, m)$, где t_i – времена, в которые производятся наблюдения, d_i – число зафиксированных в момент t_i новых случаев смерти, c_i – число новых цензурированных в момент t_i животных. Предполагается, что в начальный момент времени t_0 все животные живы, то есть $S(t_0) = 1$. Общее число животных N, участвовавших в эксперименте, определяется соотношением $N=\sum_{i=1}^{m}\left(d_i+n_i\right)$.

Функция правдоподобия равна произведению вероятностей наблюдать конкретные значения d_i и n_i в различные моменты времени t_i

$$L=\prod_{i=1}^{m} p_i^{d_i} S^{c_i}\left(t_{i-1}\right),$$

где $p_i=S\left(t_{i-1}\right)-S\left(t_i\right)$ означает безусловную вероятность наступления смерти на интервале $\left[t_{i-1},t_i\right)$. Логарифм функции правдоподобия записывается в виде

$$\ln L=\sum_{i=1}^{m} d_i \ln\left(S\left(t_{i-1}\right)-S\left(t_i\right)\right)+\sum_{i=1}^{m} c_i \ln S\left(t_{i-1}\right).$$

В случае сгруппированных данных не удаётся получить аналитических выражений для оценок параметров даже простых моделей, поэтому оценки параметров получают путём численной максимизации функции логарифма правдоподобия.

1.4 Непараметрическое оценивание функции интенсивности

Рассмотренные в этой главе модели интенсивности отказа и силы смерти определялись конечным числом параметров и оценивание функции интенсивности сводилось к построению оценок этих параметров по наблюдаемым данным. Однако, часто при решении практических задач неясно какую из возможных моделей следует предпочесть. Более того, в ряде ситуаций адекватной может оказаться не одна модель, а несколько моделей, описывающих процесс на различных участках. Так, при наличии высокой детской смертности, обусловленной недостаточным медицинским обслуживанием, имеет смысл рассматривать кривую смертности, составленную из куска с большой силой смерти, не зависящей от возраста, и куска, соответствующего модели Гомперца, учитывающего возрастные изменения.

Для получения оценок функции интенсивности отказа, не зависящих от конкретной модели, применяется непараметрический подход, при котором оцениваются не параметры, а вся функция интенсивности целиком. Один из возможных способов непараметрического оценивания заключается в использовании ядерных оценок [10], широко применяемых при оценке функции плотности вероятностей [11]. При этом оценка функции интенсивности вычисляется по формуле

$$\hat{\mu}(t) = \frac{1}{b}\sum_{i=1}^{m} K\left(\frac{t-t_i}{b}\right)\frac{d_i}{(m-i+1)},$$

где t_i – моменты отказа или цензурирования, d_i – индикатор цензурирования: $d_i = 1$ если объект нецензурирован и 0 в противном случае. Функция $K(t)$ обозначает ядро – неотрицательную функцию, удовлетворяющую условию нормировки

$$\int_{-\infty}^{\infty} K(t)dt = 1.$$

Параметр b регулирует чувствительность ядра к близости между моментом наблюдения t_j и переменной t. В качестве функции ядра возможно применение функции нормального распределения

$$K(t) = \frac{1}{\sqrt{2p}} \exp\left(-\frac{t^2}{2}\right)$$

или полиномиальных функций, заданных на отрезке, как в случае функции Епанечникова [12]

$$K(t) = \begin{cases} \frac{3}{4}\left(1 - t^2\right) & |t| \le 1 \\ 0 & |t| > 1 \end{cases}.$$

Величина параметра b, определяющего в данном случае ширину колокола, выбирается при решении конкретной задачи из условия обеспечения требуемых асимптотических свойств, либо из требования удовлетворения дополнительной априорной информации о процессе. Вопросы выбора величины параметр b обсуждаются в [12,13].

В случае сгруппированных данных непараметрическая оценка функции интенсивности имеет вид

$$\hat{\lambda}(t) = \frac{1}{b} \sum_{i=1}^{m} K\left(\frac{t - t_i}{b}\right) \frac{d_i}{N - \sum_{j=0}^{i-1}\left(d_j + c_j\right)}$$

где t_i – времена, в которые производятся наблюдения, d_i – число зафиксированных в момент t_i случаев отказа, c_i – число цензурированных в момент t_i объектов, N - общее число объектов. В формуле полагается, что $d_0 = 0$, $c_0 = 0$.