ЗАКОНЫ ФУНКЦИОНИРОВАНИЯ В СИСТЕМАХ С ГЕТЕРОГЕННОСТЬЮ

2.1 Формализация понятия гетерогенности

При анализе результатов статистических экспериментов традиционно предполагается, что все наблюдения являются реализациями одной и той же случайной величины или случайного вектора, то есть принадлежат одной и той же генеральной совокупности. Влияние внешних условий учитывается путём построения модели изменения средних значений случайных величин – регрессионные модели, либо путем разбиения экспериментальной выборки на подгруппы, соответствующие одинаковым условиям – стратификация. Широко применяются параметрические, непараметрические, стратифицированные модели, но принцип остаётся прежний – во всех экспериментах случайная компонента модели должна принадлежать одной и той же генеральной совокупности.

В ряде практически важных случаев предположение о том, что все экспериментальные данные принадлежат одной и той же генеральной совокупности может оказаться неверно. Пример подобной ситуации приведен в [3] при анализе продолжительности жизни. Оказывается, что наблюдаемый характер изменения смертности в реальных человеческих популяциях в зависимости от возраста лучше аппроксимируется моделью, в которой люди имеют различные риски смерти, чем моделью, в которой риск смерти одинаков для всех. В первом случае продолжительности жизни рассматриваются как реализации случайной величины из генеральной совокупности с единым законом распределения вероятностей, определяемым риском смерти. Во втором случае продолжительности жизни являются реализациями случайных величин из разных генеральных совокупностей с разными законами распределения вероятностей. За популяциями, в которых вероятность наступления некоторого события различна для разных членов закрепилось название гетерогенные или неоднородные.

В главе 1 рассматривались вероятностные законы функционирования отдельных элементов. Приведённые соотношения остаются справедливы и для описания процессов в однородной популяции. Так, вероятность безотказной работы в течение времени t наудачу взятого из популяции элемента равна

$$S(t) = \exp\left(-H(t)\right),$$

а интенсивность отказа такого элемента задаётся формулой

$$m(t) = \frac{d}{dt}H(t),$$

где $H(t)$ - кумулятивный риск отказа отдельного элемента. В случае неоднородной популяции эти формулы оказываются неверны и должны быть заменены более сложными, учитывающими различия в темпах выбытия элементов из популяции. Все вероятностные характеристики процессов, происходящих в неоднородной популяции, должны рассматриваться при условии продолжения функционирования элементов или дожития людей до рассматриваемого момента времени. Ниже подобные формулы получены при различных предположениях о характере гетерогенности. Основы подобного подхода были заложены в [3, 4, 15].

2.2 Дискретная гетерогенность – конечное число групп

Простейший пример гетерогенной популяции – популяция, состоящая из объектов, принадлежащих нескольким однородным группам: из продукции разных заводов, из мужчин и женщин, молодых и старых. В этом случае возможна стратификация экспериментальной выборки по признаку гетерогенности. В более сложной ситуации гетерогенность может определяться условиями эксплуатации, производства для технических систем, условиями жизни, генетикой для людей. Стратификация в этом случае затруднена из-за отсутствия надёжной информации об источнике гетерогенности и необходим её конструктивный учёт. Рассмотрим описание такой популяции на примере технической системы.

Допустим, что в системе представлены N однородных групп, кумулятивный риск отказа для элемента из каждой группы равен $H_j(t)$, $j=1,\dots,N$. Пусть в момент времени 0 с вероятностью $p_j(0)$ элемент находится в группе j. Запишем безусловную вероятность безотказной работы в течение времени t элемента из популяции

$$S(t)=\sum_{j=1}^{N} p_j(0)\exp\left(-H_j(t)\right).$$

Интенсивность отказа элемента получим воспользовавшись определением

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

$$= \frac{\sum_{j=1}^{N} p_j(0)\exp\left(-H_j(t)\right)m_j(t)}{\sum_{j=1}^{N} p_j(0)\exp\left(-H_j(t)\right)}$$

$$= \sum_{j=1}^{N} p_j(t)m_j(t) \quad ,$$

где $m_j(t)$ обозначает интенсивность отказа в j-той однородной группе, а вероятность

$$p_j(t) = \frac{p_j(0)\exp\left(-H_j(t)\right)}{\sum_{i=1}^{N} p_i(0)\exp\left(-H_i(t)\right)}$$

является вероятностью того, что элемент, проработавший, без отказа в течение времени t принадлежит однородной группе j. Выражение для интенсивности отказа в неоднородной популяции имеет смысл среднего по всё ещё работающим элементам значения интенсивностей $m_j(t)$. Ясно, что наименее работоспособные элементы выйдут из строя раньше и со временем интенсивность отказов в неоднородной популяции будет падать до уровня интенсивности отказа самого надёжного элемента. Формально это можно проиллюстрировать на примере экспоненциального распределения времени работы до отказа в отдельных группах. При этом интенсивности $m_j(t) = m_j$ не зависят от времени, $H_j(t) = m_j t$,

$$S(t) = \sum_{j=1}^{N} p_j(0)\exp\left(-m_j t\right)$$

и результирующая интенсивность отказа равна

$$m(t) = \frac{1}{\sum_{i=1}^{N} p_i(0)e^{-m_i t}} \sum_{j=1}^{N} p_j(0) m_j e^{-m_j t} \underset{t\to\infty}{\longrightarrow} \min m_j .$$

Таким образом, в неоднородной популяции результирующее распределение вероятности безотказной работы элементов с экспоненциальным распределением длительности работы до отказа не экспоненциально, а интенсивность отказа не постоянна, а стремится к интенсивности наиболее надёжного элемента. Придавая увеличению интенсивности отказа демографическую интерпретацию как показателю старости, полученный результат можно проинтерпретировать как «омоложение» популяции при наличии неоднородности, что, конечно же, невозможно в однородной популяции.Другие примеры применения моделей дискретной неоднородности для анализа смертности, результатов стрессовых экспериментов и явления гормезиса можно найти в [15, 16, 17].

Дискретная гетерогенность применяется не только при анализе, связанном с продолжительностью событий. При анализе категориальных данных, представленных многомерными таблицами сопряжённости, используют Модели Скрытых Классов – LCM модели

[18]. При этом вероятность наблюдать конкретное сочетание признаков получается усреднением вероятности наблюдать это сочетание в каждом из N ненаблюдаемых классов. В каждом из классов делается упрощающее предположение о независимости признаков и отдельное наблюдение считается принадлежащим одному из этих классов. Теоретическая совокупность наблюдений описывает гетерогенную популяцию с некоторым априорным распределением классов, которое оценивается по результатам наблюдений. Подробнее о LCM моделях говорится в главе 3.

2.3. Непрерывная гетерогенность

Обобщением дискретной гетерогенности является рассмотрение случая, когда свойства элементов популяции определяются некоторой непрерывной характеристикой. Допустим, что кумулятивный риск отказа элемента в гетерогенной популяции зависит от положительной величины z, характеризующей уязвимость элемента. Эта величина может определяться при вводе элемента в эксплуатацию условиями его производства, зависеть от условий его эксплуатации. Применительно к человеческой популяции уязвимость может определяться генетически, развитием организма, образом жизни и состоянием окружающей среды. Характеристика z рассматривается как случайная величина с распределением $F(z)$. Функция дожития элемента из неоднородной популяции запишется в виде

$$S(t) = \int \exp\left(-H(t,z)\right) dF(z) .$$

Для получения дальнейших результатов необходимо специфицировать зависимость риска отказа от величины уязвимости z и закон её распределения. Наиболее плодотворной в теоретическом и прикладном плане оказалась модель пропорционального риска, в которой уязвимость входит сомножителем с некоторым базовым кумулятивным риском $H_0(t)$: $H(t,z) = zH_0(t)$. Индивидуальная интенсивность отказа при этом оказывается также пропорциональна базовой интенсивности $m(t,z) = zm_0(t)$.

Вычислим интенсивность отказа в неоднородной популяции. Пользуясь определением интенсивности отказа и допуская дифференцирование под знаком интеграла получим

$$\begin{aligned} m(t) &= -\frac{1}{S(t)}\frac{d}{dt}S(t) \\ &= -\frac{1}{S(t)}\frac{d}{dt}\int \exp\left(-zH_0(t)\right)dF(z) \\ &= \frac{m_0(t)}{S(t)}\int z\exp\left(-zH_0(t)\right)dF(z) = \\ &= \bar{z}(t)m_0(t) \quad , \end{aligned} \tag{2.1}$$

где $\bar{z}(t)$ - среднее значение уязвимости среди элементов, доработавших до момента t. В демографической интерпретации это средняя уязвимость среди живых на момент t членов популяции.

Гамма распределение уязвимости

Для теоретических исследований удобно в качестве распределения уязвимости рассматривать гамма распределение, задаваемое функцией плотности вероятности

$$f(z) = l^k z^{k-1} \exp(-lz) / \Gamma(k),$$

где параметры l и k определяют форму распределения, изменяя его от экспоненциального при $k = 1$ до колоколообразного при больших значениях k. Среднее значение и дисперсия уязвимости для этого распределения равны $\mathrm{E}z = k / l$ и $s^2 = k / l^2$.

Удобство гамма распределения при изучении неоднородных популяций заключается в том, что распределение уязвимости среди элементов, проработавших до момента t, оказывается также гамма распределением. Действительно, обозначив через Z и T случайные величины, имеющие смысл уязвимости и длительности работы до отказа запишем условную функцию распределения уязвимости

$$P\{Z > z \mid T > t\} = P\{Z > z, T > t\} / P\{T > t\}$$

$$= \int_z^\infty \exp(-tH_0(t)) f(t)\,dt \Big/ \int_0^\infty \exp(-tH_0(t)) f(t)\,dt$$

$$= \int_z^\infty \frac{l^k}{\Gamma(k)} t^{k-1} \exp(-(H_0(t) + l)t)\,dt \Big/ \int_0^\infty \frac{l^k}{\Gamma(k)} t^{k-1} \exp(-(H_0(t) + l)t)\,dt$$

$$= \frac{(H_0(t) + l)^k}{l^k} \int_z^\infty l^k t^{k-1} \exp(-(H_0(t) + l)t) / \Gamma(k)\,dt \quad .$$

Условная плотность распределения вероятностей равна

$$\frac{d}{dz} P\{Z > z | T > t\}$$
$$= \left(H_0(t) + l\right)^k z^{k-1} \exp\left(-\left(H_0(t) + l\right)z\right) / \Gamma(k) \quad ,$$

то есть является снова гамма распределением с прежним параметром k и новым, зависящим от времени параметром $l(t) = H_0(t) + l$. В соответствии с формулой (2.1) наблюдаемая в популяции интенсивность отказов равна

$$\overline{m}(t) = \frac{k}{H_o(t) + l} m_0(t) .$$

Вероятность безотказного функционирования в течение времени t в популяции с гамма распределённой уязвимостью принимает простой вид

$$S(t) = \int \frac{l^k z^{k-1}}{\Gamma(k)} \exp\left(-z\left(l + H_0(t)\right)\right) dz$$
$$= \frac{l^k}{\left(l + H_0(t)\right)^k} \int \frac{\left(l + H_0(t)\right)^k z^{k-1}}{\Gamma(k)} \exp\left(-z\left(l + H_0(t)\right)\right) dz .$$
$$= \frac{l^k}{\left(l + H_0(t)\right)^k}$$

Поскольку в модели пропорционального риска уязвимость умножается на функцию $H_0(t)$, то не меняя результата можно принять, что эта функция выбрана таким образом, что среднее значение уязвимости равно единице. В этом случае $k = l$, дисперсия уязвимости равна $s^{2} = 1/k$ и вероятность безотказной работы принимает вид

$$S(t) = \left(1 + s^{2} H_{0}(t)\right)^{-1/s^{2}},$$

а интенсивность отказов задаётся выражением

$$\bar{m}(t) = \frac{1}{1 + s^{2} H_{0}(t)} m_{0}(t).$$

Применение нормального распределения

Поскольку уязвимость является положительной величиной, то она не может иметь нормальное распределение. Вместо этого можно рассмотреть другую величину Y, имеющую нормальное распределение и связанную с уязвимостью соотношением $Z = Y^{2}$. Если среднее значение Y равно нулю а дисперсия s^{2}, то плотность вероятности распределения уязвимости вычисляется по формуле

$$\begin{aligned} f(z) &= \frac{1}{\sqrt{2p}\, s} \exp\left(-\frac{1}{2s^{2}} z\right) \frac{1}{\sqrt{z}} \\ &= l^{\frac{1}{2}} z^{-\frac{1}{2}} \exp(-lz) / \Gamma\left(\frac{1}{2}\right), \end{aligned}$$

и является гамма распределением с параметрами $l = \dfrac{1}{2s^{2}}$ и $k = 1$.

Известно, что для нормального функционирования технических систем и биологических организмов необходимо, чтобы их параметры находились в некоторой окрестности номинальных значений:

электрическое напряжение должно быть не слишком высоко, но и не слишком мало, то же самое и кровяное давление человека, температура, вес и так далее. По этой причине представление уязвимости в виде квадратичной функции нормально распределённой случайной величины с ненулевым средним значением представляется оправданным. Из формулы (2.1) следует выражение для интенсивности отказов в виде

$$\overline{m}(t) = \left(m^2(t) + s^2(t)\right)m_0(t),$$

где $m(t)$ обозначает среднее значение, а $s^2(t)$ дисперсию случайной величины Y среди элементов, проработавших время t.

Выражения для динамики среднего и дисперсии условного распределения Y получим следуя логике, изложенной в [15]. Запишем условную плотность распределения случайной величины Y по формуле Байеса

$$p(y|T > t) = \frac{P\{T > t|y\}}{P\{T > t\}}\, p(y).$$

Подставляя

$$p(y) = \frac{1}{\sqrt{2p}\, s(0)} \exp\left(-\frac{\left(y - m(0)\right)^2}{2s^2(0)}\right),$$

$$P\{T > t|y\} = \exp\left(-y^2 H_0(t)\right),$$

$$P\{T > t\} = \frac{1}{\sqrt{2p}\, s(0)} \int \exp\left(-y^2 H_0(t) - \frac{\left(y - m(0)\right)^2}{2s^2(0)}\right) dy$$

получим

$$p(y\,|\,T>t)=\frac{1}{P\{T>t\}\sqrt{2p}\,s(0)}\exp\left(-y^2H_0(t)-\frac{(y-m(0))^2}{2s^2(0)}\right)$$

$$=\frac{1}{P\{T>t\}\sqrt{2p}\,s(0)}\exp\left(-\frac{H_0(t)m^2(0)}{2s^2(0)H_0(t)+1}\right)$$

$$\times\exp\left(-\frac{1}{2}\left(y-\frac{m(0)}{\left(2s^2(0)H_0(t)+1\right)}\right)^2\frac{2s^2(0)H_0(t)+1}{s^2(0)}\right),$$

которая соответствует распределению нормальной величины со средним значением

$$m(t)=\frac{m(0)}{2s^2(0)H_0(t)+1}$$

и дисперсией

$$s^2(t)=\frac{s^2(0)}{2s^2(0)H_0(t)+1}.$$

Нетрудно убедиться, что среднее и дисперсия условного распределения удовлетворяют системе дифференциальных уравнений

$$\frac{d}{dt}m(t)=-2m(t)s^2(t)m_0(t)$$

$$\frac{d}{dt}s^2(t)=-2s^4(t)m_0(t)\quad.$$

2.4 Нечёткая гетерогенность

До сих пор рассматривая гетерогенную популяцию мы предполагали, что вероятностные характеристики каждого элемента задаются конкретной величиной уязвимости, что позволяло в случае дискретной гетерогенности приписывать каждый элемент какой-либо однородной группе. При анализе медицинских и эпидемиологических данных оказалось, что такого представления недостаточно. Оказалось, что многие определения состояния организма и медицинские диагнозы условны и в реальности редко встречаются в «чистом виде». Например, психологи оперируют четырьмя основными характеристиками темперамента: сангвиник, холерик, меланхолик и флегматик. В реальности, люди, наделённые лишь одним из них, редки и чаше встречается те, в ком всё представлено в различной степени.

Учёт такой неоднозначности удалось формализовать в терминах нечётких множеств [19, 20, 21]. При этом каждому изучаемому объекту приписывается функция, показывающая в какой степени объект является членом класса, – функция принадлежности. Сумма по всем классам значений функции принадлежности равна единице для каждого объекта. На основании такого подхода был развит метод GOM (Grade Of Membership), разработаны алгоритмы идентификации соответствующих моделей [22]. Более подробно модели GOM рассматриваются в главе 3.