Дальневосточный математический журнал. 2019. Т. 19. № 2. С. 256–260

УДК 519.248+519.87
MSC2010 62P10+92B15

© Г. Ш. Цициашвили [1]

Эргодичность статистических оценок интенсивности пуассоновского потока

В настоящей работе ставится вопрос, как преобразовать процедуру статистических наблюдений, чтобы заменить длинный ряд наблюдений на несколько коротких и, наоборот, несколько коротких рядов на один длинный. Этот вопрос решается для оценки интенсивности стационарного пуассоновского потока и для оценки среднего времени безотказной работы технической системы с помощью понятия эргодичности процесса наблюдения.

Ключевые слова: *эргодичность, пуассоновский поток, показательное распределение, время наработки на отказ.*

1. Постановка задачи

При статистической обработке данных возникают задачи о замене длинного ряда наблюдений на несколько более коротких и несколько коротких рядов наблюдений на один более длинный. Задача замены длинного ряда наблюдений на короткие носит во многом технический характер, поскольку проведение длительных наблюдений затруднительно по техническим причинам. Такие задачи появляются, например, в оптоэлектронике [1], при статистической оценке времени наработки на отказ у технических систем [2] и т.д.

Задача замены нескольких коротких рядов наблюдений на один более длинный ряд тесно связана с задачей о малых выборках. Эта задача возникает при обработке наблюдений за поведением экологических, социологических, биоценотических, технических систем, наблюдений за потоками в системах массового обслуживания [3–5]. Трудности появляются при получении статистически достоверных оценок: несмещенных, состоятельных, эффективных и асимптотически нормальных. Переход к более длинному ряду наблюдений позволяет в некоторых случаях преодолеть эти трудности.

В настоящей работе ставится вопрос, можно ли заменить длинный ряд наблюдений на несколько более коротких и, наоборот, несколько коротких рядов на один

[1] Институт прикладной математики ДВО РАН, 690041, г. Владивосток, ул. Радио, 7; Дальневосточный федеральный университет, 690950, г. Владивосток, ул. Суханова, 8. Электронная почта: guram@iam.dvo.ru

более длинный. Важно знать, как такие замены отразятся на качестве получаемых статистических оценок. Оказывается, что этот вопрос упирается в понятие эргодичности процесса наблюдений: "среднее по ансамблю совпадает со средним по траектории". Эргодической теории динамических систем посвящено большое количество как теоретических, так и прикладных работ, см., например, [6–8].

Однако, чтобы использовать понятие эргодичности для решения задач о малых выборках и о длинных рядах наблюдений, необходимо подобрать удобную вероятностную модель наблюдений. В настоящей работе для этой цели выбраны модель стационарного пуассоновского потока и модель последовательности времен наработки технических систем на отказ с показательным распределением, имеющим большое среднее значение. Построены статистические оценки интенсивности пуассоновского потока и параметра показательного распределения, для получения которых на основе свойства эргодичности выбраны удобные способы проведения наблюдений.

2. Оценка интенсивности пуассоновского потока

Рассмотрим стационарный пуассоновский поток Λ с интенсивностью θ. Пусть $A=[a,b)$, $a<b$, $B=[b,c)$, $b<c$, – непересекающиеся полуинтервалы на вещественной оси, $C=A\cup B$. Известно, что случайные величины $n(A), n(B)$, определяющие число точек стационарного пуассоновского потока на непересекающихся полуинтервалах A, B, независимы [9, с. 12, свойство 2^0 отсутствия последействия]. Тогда случайные величины $n(A), n(B), n(C)$ имеют пуассоновское распределение с параметрами $\theta\alpha, \theta\beta, \theta\gamma, \gamma=\alpha+\beta, \alpha=b-a, \beta=c-b$. Из центральной предельной теоремы (см., например, [10, глава 7, теорема 2]) следует равномерная по t, $-\infty<t<\infty$, сходимость

$$P\left(\frac{n(C)/\gamma-\theta}{\sqrt{\theta/\gamma}}<t\right)=P\left(\frac{n(C)-\theta\gamma}{\sqrt{\theta\gamma}}<t\right)\to\Phi(t),\qquad \gamma\to\infty, \tag{1}$$

где $\Phi(t)$ — стандартное нормальное распределение (с нулевым математическим ожиданием и единичной дисперсией).

В качестве оценки параметра θ возьмем $\widehat{\theta}=\dfrac{n(C)}{\gamma}$. Среднее и дисперсия этой оценки удовлетворяют равенствам

$$M\widehat{\theta}=\theta,\qquad D\widehat{\theta}=\frac{\theta}{\gamma}. \tag{2}$$

Первое равенство в (2) означает несмещенность, соотношения (1), (2) влекут за собой состоятельность и асимптотическую нормальность оценки $\widehat{\theta}$ при $\gamma\to\infty$.

Пусть полуинтервалы $C_1,\dots,C_m$ попарно не пересекаются, имеют длину $\gamma_1,\dots,\gamma_m$ и пусть выполняется равенство $\gamma=\sum\limits_{k=1}^{m}\gamma_k$. Положим $n(C_1),\dots,n(C_m)$ – число точек пуассоновского потока Λ на этих полуинтервалах и построим статистическую оценку $\widehat{\theta}'$ параметра θ:

$$\widehat{\theta}'=\frac{\sum\limits_{k=1}^{m}n(C_k)}{\sum\limits_{k=1}^{m}\gamma_k}. \tag{3}$$

Т.к. $n(C)$ совпадает по распределению с $\sum_{k=1}^{m} n(C_k)$, то статистика $\widehat{\theta}'$, определяемая равенством (3), совпадает по распределению со статистикой $\widehat{\theta}$.

Таким образом, параметр θ можно оценивать как по наблюдениям за числом точек потока Λ на объединении непересекающихся полуинтервалов $C_1,\dots,C_m$, так и по наблюдениям за числом точек потока Λ на полуинтервале C длиной $\gamma \gg 1$. Причем число m может быть небольшим. Из формулы (3) следует, что оценку $\widehat{\theta}'$ параметра θ можно рассматривать как медианту [11, с. 23], [12, с. 631], введенную для пары дробей, но допускающую обобщение для конечного числа дробей.

Эти свойства оценки $\widehat{\theta}'$ интенсивности θ пуассоновского потока Λ близки к понятию эргодичности, что позволяет найти решение задачи о малых выборках, состоящих из числа точек пуассоновского потока на непересекающихся полуинтервалах с большой суммарной длиной.

3. Оценка большого параметра показательного распределения

Предположим, что имеется n технических систем, времена наработки на отказ которых $y_1,\dots,y_n$ являются независимыми и одинаково распределенными случайными величинами с показательным распределением $P(y_1>t)=\exp(-t/T),\, t\geqslant 0$. Пусть параметр $T=My_1$ является большим: $T\gg 1$. Нашей задачей является построение такого способа организации испытаний и процедуры оценивания параметра T, при которых увеличение числа испытуемых систем в m раз приводит к уменьшению среднего времени проведения испытаний в $m-1$ раз.

Обозначим $\widehat{T}_n=\dfrac{y_1+\dots+y_n}{n}$, при построении оценки $\widehat{T}_n$ суммарное время испытаний U_n с вероятностью единица удовлетворяет неравенству $U_n\geqslant \max(y_1,\dots,y_n)$. Следовательно, выполняются следующие соотношения

$$M\widehat{T}_n=T, \quad D\widehat{T}_n=\frac{T^2}{n}, \quad MU_n\geqslant M\max(y_1,\dots,y_n)\geqslant My_1=T\gg 1. \tag{4}$$

Чтобы сократить суммарное время проведения испытаний, сконструируем следующую процедуру. Пусть nm систем начинают работу в момент времени 0. Испытания заканчиваются, когда откажут первые n систем. Предположим, что моменты отказа этих систем удовлетворяют соотношениям $0\leqslant T_1\leqslant\dots\leqslant T_n$. Тогда независимые случайные величины $z_1=T_1,\, z_2=T_2-T_1,\dots,z_n=T_n-T_{n-1}$ имеют показательные распределения с параметрами $nmT, (nm-1)T,\dots,(nm-n+1)T$. Следовательно, независимые случайные величины $nmz_1,\, (nm-1)z_2,\dots,(nm-n+1)z_n$ имеют показательное распределение с параметром T. Используя наблюдения $z_1,\dots,z_n$, построим оценку $\overline{T}_n$ параметра T:

$$\overline{T}_n=\frac{nmz_1+(nm-1)z_2+\ldots+(nm-n+1)z_n}{n}.$$

Таким образом, случайная величина $\overline{T}_n$ совпадает по распределению со случайной величиной $\widehat{T}_n$ и, значит, выполняются соотношения

$$M\overline{T}_n=T,\ D\overline{T}_n=\frac{T^2}{n}. \tag{5}$$

В свою очередь, полное время проведения испытаний $V_n = z_1 + \ldots + z_n$ состоит из слагаемых z_k таких, что $Mz_k = \dfrac{T}{mn-k+1}$, $k=1,\ldots,n$, удовлетворяет соотношениям

$$\frac{T}{m} = \frac{T}{mn}n < MV_n = \sum_{k=1}^{n} \frac{T}{mn-k+1} < \frac{T}{mn-n}n = \frac{T}{m-1}.$$

И значит, для величины $S(m,n) = MV_n$ справедливы неравенства

$$\frac{T}{m} < S(m,n) < \frac{T}{m-1}. \tag{6}$$

Сравнение формулы (4) с формулами (5), (6) показывает, что увеличение числа испытываемых систем в m раз позволяет путем построения статистической оценки $\overline{T}_n$ сократить среднее время проведения испытаний с MU_n до $MV_n = S(m,n)$ примерно в $m-1$ раз. Такой способ организации испытаний и обработки их результатов может оказаться полезным при оценке параметра $T \gg 1$, характеризующего большое среднее время наработки на отказ. Этот прием близок к применению эргодичности при оценке интенсивности пуассоновского потока.

Список литературы

[1] Ю. Н. Кульчин, “Фотоника самоорганизующихся биоминеральных наноструктур”, *УФН*, **181**:8, (2011), 891—896.

[2] В. Б. Монсик, А. А. Скрынников, “Оценивание параметра показательного распределения по усеченной выборке”, *Научный вестник МГТУ ГА. Серия прикладная математика. Информатика*, **105**, (2006), 134–139.

[3] Ф. Н. Ильясов, “Обратная задача выборки и мотивация на рынке Форекс”, *Социологические исследования*, **3**, (2011), 112—-116.

[4] Ф. Н. Ильясов, “Репрезентативность результатов опроса в маркетинговом исследовании”, *Социальные исследования*, **2**, (2016), 49–59.

[5] Б. И. Семкин, “Об инвариантности средних отношений величин (на примере некоторых морфологических признаков слоевищ ламинарии японской (Laminaria japonica Aresh.) из сублиторали Северного Приморья”, *Известия ТИНРО*, **171**, (2012), 313–320.

[6] J. von Neumann, “Physical Applications of the Ergodic Hypothesis”, *Proc Natl Acad Sci USA*, **18**, (1932), 263—-266.

[7] А. Я. Хинчин, *Математические основы статистической механики*, ОГИЗ. Гос. изд-во технико-теорет. лит., Москва, 1943.

[8] И. П. Корнфельд, Я. Г. Синай, С. В. Фомин, *Эргодическая теория*, Наука, Москва, 1980.

[9] А. Я. Хинчин, *Работы по теории математической теории массового обслуживания*, Физматлит, Москва, 1963.

[10] А. А. Боровков, *Курс теории вероятностей*, Наука, Москва, 1972.

[11] А. Я. Хинчин, *Цепные дроби*, Физматлит, Москва – Ленинград, 1949.

[12] *Математическая энциклопедия*, т. 3, ред. И. М. Виноградов, Советская энциклопедия, Москва, 1982.

Поступила в редакцию
26 июня 2019 г.

Исследование выполнено при частичной финансовой поддержке РФФИ (проект № 17-07-00177, программы «Дальний Восток» ДВО РАН, проект № 18-5-056).

Tsitsiashvili G. Sh. Ergodicity of statistical estimates of Poisson flow intensity. *Far Eastern Mathematical Journal.* 2019. V. 19. No 2. P. 256–260.

ABSTRACT

This paper raises the question of how to transform the observation procedure to replace a long series of observations by several short and, conversely, several short series by one long. This problem is solved to assess the intensity of the stationary Poisson flow and to estimate the average time of operation of technical systems for failure using the concept of ergodicity of the observation process.

Key words: *ergodicity, Poisson flow, exponential distribution, mean failure time.*