УДК 519.2:303.732.4

ТЕОРЕТИЧЕСКИЕ ИНСТРУМЕНТЫ СТАТИСТИЧЕСКИХ МЕТОДОВ

Орлов Александр Иванович
д.э.н., д.т.н., к.ф.-м.н., профессор

Московский государственный технический университет им. Н.Э. Баумана, Россия, 105005, Москва, 2-я Бауманская ул., 5, prof-orlov@mail.ru

Рассмотрены основные математические инструменты (теоремы, методы), постоянно используемые при обосновании новых результатов в области статистических методов: законы больших чисел, центральные предельные теоремы, необходимые и достаточные условия наследования сходимости, метод линеаризации, принцип инвариантности

Ключевые слова: СТАТИСТИЧЕСКИЕ МЕТОДЫ, МАТЕМАТИЧЕСКАЯ СТАТИСТИКА, ПРИКЛАДНАЯ СТАТИСТИКА, ЗАКОНЫ БОЛЬШИХ ЧИСЕЛ, ЦЕНТРАЛЬНЫЕ ПРЕДЕЛЬНЫЕ ТЕОРЕМЫ, НЕОБХОДИМЫЕ И ДОСТАТОЧНЫЕ УСЛОВИЯ, НАСЛЕДОВАНИЕ СХОДИМОСТИ, МЕТОД ЛИНЕАРИЗАЦИИ, ПРИНЦИП ИНВАРИАНТНОСТИ

UDC 519.2:303.732.4

THEORETICAL TOOLS OF STATISTICAL METHODS

Orlov Alexander Ivanovich
Dr.Sci.Econ., Dr.Sci.Tech., Cand.Phys-Math.Sci., professor
Bauman Moscow State Technical University, Moscow, Russia

We have considered the basic mathematical tools (theorems, methods) which are used regularly in the justification of new results in the field of statistical methods: rules of large numbers, central limit theorems, the necessary and sufficient conditions for the inheritance of convergence, the linearization method, the invariance principle

Keywords: STATISTICAL METHODS, MATHEMATICAL STATISTICS, APPLIED STATISTICS, LAWS OF LARGE NUMBERS, CENTRAL LIMIT THEOREMS, NECESSARY AND SUFFICIENT CONDITIONS, INHERITANCE OF CONVERGENCE, LINEARIZATION METHOD, INVARIANCE PRINCIPLE

1. Введение

Набор широко применяемых исследователями теоретических инструментов прикладной математической статистики и статистических методов в целом достаточно ограничен. В настоящей статье собраны основные математические инструменты (теоремы, методы), постоянно используемые при обосновании новых результатов в области статистических методов. Эти инструменты отнюдь не всегда легко найти в литературе по теории вероятностей и математической статистике. Например, такие рассматриваемые далее теоремы и методы, как многомерная центральная предельная теорема, теоремы о наследовании сходимости и метод линеаризации, даже не включены в энциклопедию «Вероятность и математическая статистика» [1] – наиболее полный, по мнению составителей энциклопедии, свод знаний по заявленной тематике. Последний факт наглядно демонстрирует разрыв между математической дисциплиной «теория вероятностей и математическая статистика» и потребностями прикладной статистики и других статистических методов.

2. Законы больших чисел

Законы больших чисел позволяют описать поведение сумм случайных величин. Примером является следующий результат, доказанный русским математиком П.Л. Чебышёвым (1821–1894) в 1867 г. Пусть сначала вероятностное пространство состоит из конечного числа элементов.

Теорема Чебышёва. Пусть случайные величины X_1, X_2,…, X_k попарно независимы и существует число C такое, что $D(X_i)\leq C$ при всех i = 1, 2, …, k. Тогда для любого положительного ε выполнено неравенство

$$P\left\{\left|\frac{X_1+X_2+...+X_k}{k}-\frac{M(X_1)+M(X_2)+...+M(X_k)}{k}\right|\geq\varepsilon\right\}\leq\frac{C}{k\varepsilon^2}. \quad (1)$$

Частным случаем теоремы Чебышева является теорема Бернулли – первый в истории вариант закона больших чисел. Известный математики Якоб Бернулли (1654–1705), живший в городе Базель в Швейцарии, в самом конце XVII века доказал это утверждение в рамках математической модели (опубликовано доказательство было лишь после его смерти, в 1713 году). Современная формулировка теоремы Бернулли такова.

Теорема Бернулли. Пусть m – число наступлений события A в k независимых (попарно) испытаниях, и p есть вероятность наступления события A в каждом из испытаний. Тогда при любом $\varepsilon>0$ справедливо неравенство

$$P\left\{\left|\frac{m}{k}-p\right|\geq\varepsilon\right\}\leq\frac{p(1-p)}{k\varepsilon^2}. \quad (2)$$

Ясно, что при росте k выражения в правых частях формул (1) и (2) стремятся к 0. Таким образом, среднее арифметическое попарно независимых случайных величин сближается со средним арифметическим их математических ожиданий.

Выше шла речь лишь о пространствах элементарных событий из конечного числа элементов. Однако приведенные теоремы верны и в общем случае, для произвольных пространств элементарных событий. Однако в список условий закона больших чисел необходимо добавить требование существования дисперсий. Легко видеть, что если существуют дисперсии, то существуют и математические ожидания. Закон больших чисел в форме Чебышёва приобретает следующий вид.

Теорема Чебышёва [2, с. 147]. Если X_1, $X_2,\ldots$, $X_k,\ldots$ - последовательность попарно независимых случайных величин, имеющих конечные дисперсии, ограниченные одной и той же постоянной,

$$D(X_1)\leq C,\ D(X_2)\leq C,\ldots\ D(X_i)\leq C,\ldots$$

то, каково бы ни было постоянное ε > 0,

$$\lim_{k\to\infty} P\left\{\left|\frac{1}{n}\sum_{j=1}^{k} X_j - \frac{1}{n}\sum_{j=1}^{k} MX_j\right| < \varepsilon\right\} = 1. \quad (3)$$

С точки зрения прикладных статистических исследований ограниченность дисперсий вполне естественна. Она вытекает, например, из ограниченности диапазона изменения практически всех величин, используемых при реальных расчетах.

В 1923 г. А.Я. Хинчин показал, что если случайные величины не только независимы, но и одинаково распределены, то существование у них математического ожидания является необходимым и достаточным условием для применимости закона больших чисел [2, с. 150]. Найдены и более экзотические варианты закона больших чисел. Например, такой.

Теорема [2, с. 150–151]. Для того чтобы для последовательности X_1, $X_2,\ldots$, $X_k,\ldots$(как угодно зависимых) случайных величин при любом положительном ε выполнялось соотношение (3), необходимо и достаточно, чтобы при $n \to \infty$

$$M\frac{\left(\sum_{j=1}^{k}(X_j - MX_j)\right)^2}{n^2 + \left(\sum_{j=1}^{k}(X_j - MX_j)\right)^2} \to 0.$$

Законы больших чисел для случайных величин служат основой для аналогичных утверждений для случайных элементов в пространствах более сложной природы, в частности, в пространствах произвольной природы [3, 4]. Однако здесь мы ограничимся классическими формулировками, служащими основой для современных статистических методов.

Смысл классических законов больших чисел состоит в том, что выборочное среднее арифметическое независимых одинаково распределенных случайных величин приближается (сходится) к математическому ожиданию этих величин. Другими словами, *выборочные средние сходятся к теоретическому среднему*.

Это утверждение справедливо и для других видов средних. Например, выборочная медиана сходится к теоретической медиане. Это утверждение – тоже закон больших чисел, но не классический.

Существенным продвижением в теории вероятностей во второй половине XX в. явилось введение средних величин в пространствах произвольной природы и получение для них законов больших чисел, т.е. утверждений, состоящих в том, что эмпирические (т.е. выборочные) средние сходятся к теоретическим средним [3, 4].

3. Центральные предельные теоремы

Простейший вариант Центральной предельной теоремы (ЦПТ) теории вероятностей таков.

Центральная предельная теорема (для одинаково распределенных слагаемых). Пусть X_1, X_2,…, X_n, …– независимые одинаково распределенные случайные величины с математическими ожиданиями $M(X_i) = m$ и дисперсиями $D(X_i) = \sigma^2$, $i = 1, 2,\dots, n,\dots$ Тогда для любого действительного числа x существует предел

$$\lim_{n\to\infty} P\left(\frac{X_1 + X_2 + \dots + X_n - nm}{\sigma\sqrt{n}} < x\right) = \Phi(x),$$

где $\Phi(x)$ – функция стандартного нормального распределения.

Эту теорему иногда называют теоремой Линдеберга – Леви [5, с. 122].

В ряде прикладных задач не выполнено условие одинаковой распределенности. В таких случаях центральная предельная теорема обычно остается справедливой, однако на последовательность случайных величин приходится накладывать те или иные условия. Суть этих условий состоит в том, что ни одно слагаемое не должно быть доминирующим, вклад каждого слагаемого в среднее арифметическое должен быть пренебрежимо мал по сравнению с итоговой суммой. Наиболее часто используется теорема Ляпунова.

Теорема Ляпунова – Центральная предельная теорема (для разнораспределенных слагаемых). Пусть X_1, $X_2,\dots$, X_n, …– независимые случайные величины с математическими ожиданиями $M(X_i) = m_i$ и дисперсиями $D(X_i) = \sigma_i^2 \neq 0$, $i = 1, 2,\dots, n,\dots$ Пусть при некотором $\delta > 0$ у всех рассматриваемых случайных величин существуют центральные моменты порядка $2+\delta$ и безгранично убывает «дробь Ляпунова»:

$$\lim_{k\to\infty} \frac{1}{B_n^{2+\delta}} \sum_{k=1}^{n} M\,|\,X_k - m_k\,|^{2+\delta} = 0,$$

где

$$B_k^2 = \sum_{i=1}^{k} \sigma_i^2 = D\left(\sum_{i=1}^{k} X_i\right).$$

Тогда для любого действительного числа x существует предел

$$\lim_{n\to\infty} P\left(\frac{X_1 + X_2 + ... + X_n - m_1 - m_2 - ... - m_n}{B_n} < x\right) = \Phi(x), \quad (4)$$

где $\Phi(x)$ – функция стандартного нормального распределения.

В случае одинаково распределенных случайных слагаемых

$$m_1 = m_2 = ... = m_n = m, \quad B_n = D(X_1 + X_2 + ... + X_n) = \sigma\sqrt{n},$$

и теорема Ляпунова переходит в теорему Линдеберга – Леви.

История получения центральных предельных теорем для числовых случайных величин растянулась на два века – от первых работ Муавра в 30-х гг. XVIII в. для необходимых и достаточных условий, полученных Линдебергом и Феллером в 30-х гг. XX в.

Теорема Линдеберга – Феллера. Пусть X_1, X_2,…, X_n, …– независимые случайные величины с математическими ожиданиями $M(X_i) = m_i$ и дисперсиями $D(X_i) = \sigma_i^2 \neq 0$, i = 1, 2,…, n,… Предельное соотношение (4), т.е. Центральная предельная теорема, выполнено тогда и только тогда, когда при любом τ > 0

$$\lim_{n\to\infty} \frac{1}{B_n^2} \sum_{k=1}^{n} \int\limits_{|x-m_k|>\tau B_n} (x - m_k)^2 \, dF_k(x0 = 0,$$

где $F_k(x)$ обозначает функцию распределения случайной величины X_k.

Доказательства перечисленных в настоящем разделе центральных предельных теорем для случайных величин можно найти в классическом курсе теории вероятностей [2].

Для обоснования многих статистических методов большое значение имеет многомерная центральная предельная теорема. В ней речь идет не о сумме случайных величин, а о сумме случайных векторов.

Необходимое и достаточное условие многомерной сходимости [5, с. 124]. Пусть F_n обозначает совместную функцию распределения k-мерного случайного вектора $(X_n^{(1)},...,X_n^{(k)})$, n = 1, 2, …, и $F_{\lambda n}$ – функция распределения линейной комбинации $\lambda_1 X_n^{(1)} + ... + \lambda_л X_n^{(k)}$. Необходимое и достаточное условие для сходимости F_n к некоторой k-мерной функции распределения F состоит в том, что $F_{\lambda n}$ имеет предел для любого вектора λ.

Приведенная теорема ценна тем, что с ее помощью сходимость распределений случайных векторов сводится к сходимости распределений линейных комбинаций их координат, т.е. к сходимости обычных (числовых) случайных величин, рассмотренных ранее. Однако она не дает возможности непосредственно указать предельное распределение. Это можно сделать с помощью следующей теоремы.

Теорема о многомерной сходимости [5]. Пусть F_n и $F_{\lambda n}$ – те же, что в предыдущей теореме. Пусть F - совместная функция распределения k-мерного случайного вектора $(X_1,...,X_k)$. Если функция распределения $F_{\lambda n}$ сходится при росте объема выборки к функции распределения F_λ для любого вектора λ, где F_λ – функция распределения линейной комбинации $\lambda_1 X_1 + ... + \lambda_k X_k$, то F_n сходится к F.

Здесь сходимость F_n к F означает, что для любого k-мерного вектора $(x_1,...,x_k)$ такого, что функция распределения F непрерывна в $(x_1,...,x_k)$, числовая последовательность $F_n(x_1,...,x_k)$ сходится при росте n к числу $F(x_1,...,x_k)$. Другими словами, сходимость функций распределения понимается точно так же, как при обсуждении предельных теорем для случайных величин выше. Приведем многомерный аналог этих теорем.

Многомерная центральная предельная теорема [5]. Рассмотрим независимые одинаково распределенные k-мерные случайные вектора

$$U_n^{'} = (U_{1n},...,U_{kn}),\quad n = 1,2,...,$$

где штрих обозначает операцию транспонирования вектора. Предположим, что случайные вектора U_n имеют моменты первого и второго порядка, т.е.

$$M(U_n) = \mu,\ D(U_n) = \Sigma,$$

где μ – вектор математических ожиданий координат случайного вектора, Σ – его ковариационная матрица. Введем последовательность средних арифметических случайных векторов:

$$\overline{U}_n = (\overline{U}_{1n}, ..., \overline{U}_{kn}), \quad n = 1, 2, ..., \quad \overline{U}_{in} = \frac{1}{n}\sum_{j=1}^{n} U_{ij}.$$

Тогда случайный вектор $\sqrt{n}(\overline{U}_n - \mu)$ имеет асимптотическое *k*-мерное нормальное распределение $N_k(0, \Sigma)$, т.е. он асимптотически распределен так же, как *k*-мерная нормальная величина с нулевым математическим ожиданием, ковариационной Σ и плотностью

$$N_k(u \mid 0, \Sigma) = (2\pi)^{-k/2} \mid \Sigma \mid^{-1/2} \exp\left\{-\frac{1}{2}u'\Sigma^{-1}u\right\}.$$

Здесь |Σ| - определитель матрицы Σ. Другими словами, распределение случайного вектора $\sqrt{n}(\overline{U}_n - \mu)$ сходится к *k*-мерному нормальному распределению с нулевым математическим ожиданием и ковариационной матрицей Σ.

Напомним, что многомерным нормальным распределением с математическим ожиданием μ и ковариационной матрицей Σ называется распределение, имеющее плотность

$$N_k(u \mid \mu, \Sigma) = (2\pi)^{-k/2} \mid \Sigma \mid^{-1/2} \exp\left\{-\frac{1}{2}[(u-\mu)'\Sigma^{-1}(u-\mu)]\right\}.$$

Многомерная центральная предельная теорема показывает, что распределения сумм независимых одинаково распределенных случайных векторов при большом числе слагаемых хорошо приближаются с помощью нормальных распределений, имеющих такие же первые два момента (вектор математических ожиданий координат случайного вектора и его корреляционную матрицу), как и исходные вектора. От одинаковой распределенности можно отказаться, но это потребует некоторого усложнения символики. В целом из теоремы о многомерной сходимости вытекает, что многомерный случай ничем принципиально не отличается от одномерного.

Пример. Пусть X_1, … X_n ,…– независимые одинаково распределенные случайные величины. Рассмотрим k-мерные независимые одинаково распределенные случайные вектора

$$U_n^{'} = (X_n, X_n^2, X_n^3, ..., X_n^k), \quad n = 1,2,...$$

Их математическое ожидание – вектор теоретических начальных моментов, а ковариационная матрица составлена из соответствующих центральных моментов. Тогда $\overline{U}_n$ - вектор выборочных центральных моментов. Многомерная центральная предельная теорема утверждает, что $\overline{U}_n$ имеет асимптотически нормальное распределение. Как вытекает из теорем о наследовании сходимости и о линеаризации (см. ниже), из распределения $\overline{U}_n$ можно вывести распределения различных функций от выборочных начальных моментов. А поскольку центральные моменты выражаются через начальные моменты, то аналогичное утверждение верно и для них.

4. Теоремы о наследовании сходимости

Суть проблемы наследования сходимости. Пусть распределения случайных величин X_n при $n \to \infty$ стремятся к распределению случайной величины X. При каких функциях f можно утверждать, что распределения случайных величин $f(X_n)$ сходятся к распределению $f(X)$, т.е. наследуется сходимость?

Хорошо известно, что для непрерывных функций f сходимость наследуется [5]. Однако в статистических методах используются различные обобщения этого утверждения. Необходимость обобщений связана с тремя обстоятельствами.

1) Статистические данные могут моделироваться не только случайными величинами, но и случайными векторами, случайными множествами, случайными элементами произвольной природы (т.е. функциями на вероятностном пространстве со значениями в произвольном множестве) [6, 7].

2) Переход к пределу должен рассматриваться не только для случая безграничного возрастания объема выборки, но и в более общих случаях. Например, если в постановке статистической задачи участвуют несколько выборок объемов $n(1)$, $n(2)$, … , $n(k)$, то вполне обычным является предположение о безграничном росте всех этих объемов (что можно описать и как $\min \{n(1), n(2), \dots , n(k)\} \to \infty$).

3) Функция f не обязательно является непрерывной. Она может иметь разрывы. Кроме того, она может зависеть от параметров, по которым происходит переход к пределу. Например, может зависеть от объемов выборок. Например, в [8, гл. 5] понадобилось рассмотреть функцию $f = f(n(1), n(2), \dots , n(k))$.

Расстояние Прохорова и сходимость по направленному множеству. Введем необходимые для дальнейшего изложения понятия.

Расстояние (*метрика*) *Прохорова*. Пусть C – некоторое пространство, A – его подмножество, d – метрика в C. Назовем ε-окрестностью множества A в метрике d следующее множество:

$$S(A,\varepsilon) = \{x\in C: d(A,x) < \varepsilon\}.$$

Таким образом, ε-окрестность множества A – это совокупность всех точек пространства C, отстоящих от A не более чем на положительное число ε. При этом расстояние от точки x до множества A – это точная нижняя грань расстояний от x до точек множества A, т.е.

$$d(A,x) = \inf\{d(x,y): y\in A\}.$$

Пусть P_1 и P_2 – две вероятностные меры на C (т.е. распределения двух случайных элементов со значениями в C). Пусть D_{12} – множество чисел ε > 0 таких, что

$$P_1(A) \leq P_2(S(A,\varepsilon)+\varepsilon$$

для любого замкнутого подмножества A пространства C. Пусть D_{21} – множество чисел ε > 0 таких, что

$$P_2(A) \leq P_1(S(A,\varepsilon)+\varepsilon$$

для любого замкнутого подмножества A пространства C. Расстояние Прохорова $L(P_1,P_2)$ между вероятностными мерами (его можно рассматривать и как расстояние между случайными элементами с распределениями P_1 и P_2 соответственно) вводится формулой

$$L(P_1,P_2) = \max(\inf D_{12}, \inf D_{21}).$$

С помощью метрики Прохорова формализуется понятие сходимости распределений случайных элементов в произвольном пространстве.

Расстояние $L(P_1,P_2)$ введено академиком РАН Юрием Васильевичем Прохоровым в середине XX в. [9] и широко используется в современной теории вероятностей.

Сходимость по направленному множеству [10, с. 95–96]. Бинарное отношение $\geq$ (упорядочение), заданное на множестве B, называется направлением на нем, если B не пусто и

(а) если m, n и p – такие элементы множества B, что $m \geq n$ и $n \geq p$, то $m \geq p$;

(б) $m \geq m$ для любого m из B;

(в) если m и n принадлежат B, то найдется элемент p из B такой, что $p \geq m$ и $p \geq n$.

Направленное множество – это пара $(B, \geq)$, где $\geq$ - направление на множестве B. Направленностью (или «последовательностью по направленному множеству») называется пара $(f, \geq)$, где f – функция, $\geq$ - направление на ее области определения. Пусть f: $B \to Y$, где Y –

топологическое пространство. Направленность (f, $\geq$) сходится в топологическом пространстве Y к точке y_0, если для любой окрестности U точки y_0 найдется p из B такое, что $f(q) \in U$ при любом $q \geq p$. В таком случае говорят также о сходимости по направленному множеству.

Пусть $B = \{(n(1), n(2), \dots , n(k))\}$ – совокупность векторов, каждый из которых составлен из объемов k выборок. Пусть

$$(n(1), n(2), \dots , n(k)) \geq (n_1(1), n_1(2), \dots , n_1(k))$$

тогда и только тогда, когда $n(i) \geq n_1(i)$ при всех $i = 1, 2, \dots, k$. Тогда $(B, \geq)$ – направленное множество, сходимость по которому эквивалентна сходимости при $\min \{n(1), n(2), \dots , n(k)\} \to \infty$.

Чтобы охватить различные частные случаи, целесообразно предельные теоремы формулировать в терминах сходимости по направленному множеству. Будем писать $B = \{\alpha\}$. Пусть запись $\alpha\to\infty$ обозначает переход к пределу по направленному множеству.

Формулировка проблемы наследования сходимости. Пусть случайные элементы X_α со значениями в пространстве C сходятся при $\alpha\to\infty$ к случайному элементу X, где через $\alpha\to\infty$ обозначен переход к пределу по направленному множеству. Сходимость случайных элементов означает, что $L(X_\alpha, X) \to 0$ при $\alpha\to\infty$, где L – метрика Прохорова в пространстве C.

Пусть f_α: $C \to Y$ – некоторые функции. Какие условия надо на них наложить, чтобы из $L(X_\alpha, X) \to 0$ вытекало, что $L_1(f_\alpha(X_\alpha), f_\alpha(\mathrm{X})) \to 0$ при $\alpha\to\infty$, где L_1 – метрика Прохорова в пространстве Y? Другими словами, какие условия на функции f_α: $C \to Y$ гарантируют наследование сходимости?

В работах [11, 12] найдены необходимые и достаточные условия на функции f_α: $C \to Y$, гарантирующие наследование сходимости. Описанию этих условий посвящена оставшаяся часть подраздела.

Приведем для полноты изложения строгие формулировки математических предположений.

Математические предположения. Пусть C и $У$ – полные сепарабельные метрические пространства [10]. Пусть выполнены обычные предположения измеримости: X_α и X – случайные элементы C, $f_\alpha(X_\alpha)$ и $f_\alpha(X)$ – случайные элементы в $У$, рассматриваемые ниже подмножества пространств C и $У$ лежат в соответствующих σ–алгебрах измеримых подмножеств, и т.д.

Понадобятся некоторые *определения*. Разбиение $T_n = \{C_{1n}, C_{2n}, \dots , C_{nn}\}$ пространства C – это такой набор подмножеств C_j, $j = 1, 2, \dots , n$, этого пространства, что пересечение любых двух из них пусто, а объединение совпадает с C. Диаметром $diam(A)$ подмножества A множества C называется точная верхняя грань расстояний между элементами A, т.е.

$$diam(A) = \sup \{d(x,y), x\in A, y\in A\},$$

где $d(x,y)$ – метрика в пространстве C. Обозначим ∂A границу множества A, т.е. совокупность точек x таких, что любая их окрестность $U(x)$ имеет непустое пересечение как с A, так и с $C\backslash A$. Колебанием $\delta(f, B)$ функции f на множестве B называется $\delta(f, B) = \sup \{|f(x) - f(y)|, x\in B, y\in B\}$.

Достаточное условие для наследования сходимости. Пусть $L(X_\alpha,X) \to 0$ при $\alpha \to \infty$. Пусть существует последовательность T_n разбиений пространства C такая, что $P(X\in \partial A) = 0$ для любого A из T_n и, основное условие, для любого $\varepsilon > 0$

$$m_\varepsilon(\alpha,n) = \sum P(X \in A) \to 0 \qquad (5)$$

при $n \to\infty$ и $\alpha\to\infty$, где сумма берется по всем тем A из T_n, для которых колебание функции f_α на A больше ε, т.е. $\delta(f_\alpha, A) > \varepsilon$. Тогда $L_1(f_\alpha(X_\alpha), f_\alpha(X)) \to 0$ при $\alpha\to\infty$.

Необходимое условие для наследования сходимости. Пусть $У$ – конечномерное линейное пространство, $У = R^k$. Пусть случайные элементы $f_\alpha(X)$ асимптотически ограничены по вероятности при $\alpha\to\infty$, т.е. для любого $\varepsilon > 0$ существуют число $S(\varepsilon)$ и элемент направленного множества $\alpha(\varepsilon)$ такие, что $P(||f_\alpha(X)||> S(\varepsilon))<\varepsilon$ при $\alpha \geq \alpha(\varepsilon)$, где $||f_\alpha(X)||$ - норма (длина) вектора $f_\alpha(X)$. Пусть существует последовательность T_n разбиений пространства C такая, что

$$\lim_{n\to\infty} \max\{diam(C_{jn}),\ C_{jn} \in T_n\} = 0,$$

т.е. последовательность T_n является безгранично измельчающейся. Самое существенное – пусть условие (5) не выполнено для последовательности T_n. Тогда существует последовательность случайных элементов X_α такая, что $L(X_\alpha,X) \to 0$ при $\alpha \to \infty$, но $L_1(f_\alpha(X_\alpha), f_\alpha(X))$ не сходится к 0 при $\alpha \to \infty$.

Несколько огрубляя, можно сказать, что *условие* (5) *является необходимым и достаточным для наследования сходимости*.

Пример 1. Пусть C и $У$ – конечномерные линейные пространства, функции f_α не зависят от α, т.е. $f_\alpha \equiv f$, причем функция f ограничена. Тогда условие (5) эквивалентно требованию интегрируемости по Риману - Стилтьесу функции f по мере $G(A) = P(X\in A)$. В частности, условие (5) выполнено для непрерывной функции f.

В конечномерных пространствах C вместо сходимости $L(X_\alpha,X) \to 0$ при $\alpha \to \infty$ можно говорить о слабой сходимости функций распределения случайных векторов X_α к функции распределения случайного вектора X. Речь идет о «сходимости по распределению», т.е. о сходимости во всех точках непрерывности функции распределения случайного вектора X. В этом случае разбиения могут состоять из многомерных параллелепипедов [12, гл. 2].

Пример 2. Полученные выше результаты дают обоснование для рассуждений типа следующего (ср., например, утверждения в [8, гл. 5] выше). Пусть по двум независимым выборкам объемов m и n соответственно построены статистики X_m и Y_n. Пусть известно, что распределения этих статистик сходятся при безграничном росте объемов выборок к стандартным нормальным распределениям с математическим ожиданием 0 и дисперсией 1. Пусть $a(m, n)$ и $b(m, n)$ – некоторые коэффициенты. Тогда согласно результатам примера 1 распределение случайной величины $Z(m, n) = a(m, n)X_m + b(m, n)Y_n$ сближается с распределением нормально распределенной случайной величины с математическим ожиданием 0 и дисперсией $a^2(m, n) + b^2(m, n)$. Если же $a^2(m, n) + b^2(m, n) = 1$, например,

$$a(m,n)=\sqrt{\frac{m}{m+n}}.\quad b(m,n)=\sqrt{\frac{n}{m+n}},$$

то распределение $Z(m, n)$ сходится при безграничном росте объемов выборок к стандартному нормальному распределению с математическим ожиданием 0 и дисперсией 1.

5. Метод линеаризации

При разработке статистических методов часто возникает следующая задача (см., например, [5, с. 338]). Имеется последовательность k-мерных случайных векторов $X_n = (X_{1n}, X_{2n}, \dots , X_{kn})$, $n = 1, 2, \dots$, такая, что $X_n \to a = (a_1, a_2, \dots , a_k)$ при $n \to \infty$, и последовательность функций f_n: $R^k \to R^1$. Требуется найти распределение случайной величины $f_n(X_n)$.

Основная идея – рассмотреть главный линейный член функции f_n в окрестности точки a. Из математического анализа известно, что

$$f_n(X_n)-f_n(a)=\sum_{j=1}^{k}\frac{\partial f_n(a)}{\partial x_j}(X_{jn}-a_j)+O_n(\| X_n-a\|^2),$$

где остаточный член является бесконечно малой величиной более высокого порядка малости, чем линейный член. Таким образом, произвольная функция может быть заменена на линейную функцию от координат случайного вектора. Эта замена проводится с точностью до бесконечно малых более высокого порядка. Конечно, должны быть выполнены некоторые математические условия регулярности. Например, функции f_n должны быть дважды непрерывно дифференцируемы в окрестности точки a.

Если вектор X_n является асимптотически нормальным с математическим ожиданием a и ковариационной матрицей $\sum/n$, где $\sum = \|\sigma_{ij}\|$, причем $\sigma_{ij} = nM(X_i - a_i)(X_j - a_j)$, то линейная функция от его координат также асимптотически нормальна. Следовательно, при очевидных условиях регулярности $f_n(X_n)$ – асимптотически нормальная случайная величина с математическим ожиданием $f_n(a)$ и дисперсией

$$\frac{1}{n}\sum_{i=1}^{k}\sum_{j=1}^{k}\frac{\partial f_n(a)}{\partial x_i}\frac{\partial f_n(a)}{\partial x_j}\sigma_{ij}.$$

Для практического использования асимптотической нормальности $f_n(X_n)$ остается заменить неизвестные моменты a и $\sum$ на их оценки. Например, если X_n – это среднее арифметическое независимых одинаково распределенных случайных векторов, то a можно заменить на X_n, а $\sum$ - на выборочную ковариационную матрицу.

Пример. Пусть Y_1, Y_2, … , Y_n – независимые одинаково распределенные случайные величины с математическим ожиданием a и дисперсией σ^2. В качестве X_n (при k = 1) рассмотрим выборочное среднее арифметическое

$$\overline{Y} = \frac{Y_1 + Y_2 + ... + Y_n}{n}.$$

Как известно, в силу закона больших чисел $\overline{Y} \to a = M(У)$. Следовательно, для получения распределений функций от выборочного среднего арифметического можно использовать метод линеаризации. В качестве примера рассмотрим $f_n(y) = f(y) = y^2$. Тогда

$$\left(\overline{Y}\right)^2 - a^2 = \frac{df(a)}{dy}\left(\overline{Y} - a\right) + O\left(\left(\overline{Y} - a\right)^2\right) = 2a\left(\overline{Y} - a\right) + O\left(\left(\overline{Y} - a\right)^2\right).$$

Из этого соотношения следует, что с точностью до бесконечно малых более высокого порядка

$$\left(\overline{Y}\right)^2 = a^2 + 2a\left(\overline{Y} - a\right).$$

Поскольку в соответствии с Центральной Предельной Теоремой выборочное среднее арифметическое является асимптотически нормальной случайной величиной с математическим ожиданием a и дисперсией σ^2/n, то квадрат этой статистики является асимптотически нормальной случайной величиной с математическим ожиданием a^2 и дисперсией $4a^2\sigma^2/n$. Для практического использования может оказаться полезной замена параметров (асимптотического нормального распределения) на их оценки, а именно, математического ожидания – на $\left(\overline{Y}\right)^2$, а дисперсии – на $4\left(\overline{Y}\right)^2 s^2/n$, где s^2 – выборочная дисперсия.

Большое внимание (целая глава!) уделено методу линеаризации в классическом учебнике Е.С. Вентцель [13].

6. Принцип инвариантности

Пусть Y_1, Y_2, … , Y_n – независимые одинаково распределенные случайные величины с непрерывной функцией распределения $F(x)$. Многие используемые в статистических методах функции от результатов наблюдений выражаются через эмпирическую функцию распределения $F_n(x)$. К ним относятся, в частности, статистики Колмогорова, Смирнова, омега-квадрат [14]. Отметим, что и другие статистики выражаются через эмпирическую функцию распределения, например:

$$\overline{Y} = \int_{-\infty}^{+\infty} x dF_n(x).$$

Полезным является преобразование Н.В. Смирнова $t = F(x)$. Тогда независимые случайные величины $Z_j = F(Y_j)$, $j = 1, 2, \dots, n$, имеют равномерное распределение на отрезке [0; 1]. Рассмотрим построенную по ним эмпирическую функцию распределения $F_n(t)$, $0 \le t \le 1$. *Эмпирическим процессом* называется случайный процесс

$$\xi_n(t) = \sqrt{n}(F_n(t) - t).$$

Рассмотрим критерии проверки согласия функции распределения выборки с фиксированной функцией распределения $F(x)$. Статистика критерия Колмогорова записывается в виде

$$K_n = \sup_{0 \le t \le 1} |\xi_n(t)|,$$

статистика критерия Смирнова – это

$$S_n = \sup_{0 \le t \le 1} \xi_n(t),$$

а статистика критерия омега-квадрат (Крамера-Мизеса-Смирнова) имеет вид

$$\omega_n^2 = \int_0^1 \xi_n^2(t)dt.$$

Случайный процесс $\xi_n(t)$ имеет нулевое математическое ожидание и ковариационную функцию $M\xi_n(s)\xi_n(t) = \min(s,\mathrm{t}) - st$. Рассмотрим гауссовский случайный процесс $\xi(t)$ с такимЬ же математическим ожиданием и ковариационной функцией. Он называется броуновским мостом. (Напомним, что гауссовским процесс именуется потому, что вектор $(\xi(t_1), \xi(t_2), \dots, \xi(t_k))$ имеет многомерное нормальное распределение при любых наборах моментов времени $t_1, t_2, \dots, t_k$.)

Пусть f – функционал, определенный на множестве возможных траекторий случайных процессов. *Принцип инвариантности* [1] состоит в том, что последовательность распределений случайных величин $f(\xi_n)$ сходится при $n \to \infty$ к распределению случайной величины $f(\xi)$. Сходимость по распределению обозначим символом =>. Тогда принцип инвариантности кратко записывается так: $f(\xi_n)$ => $f(\xi)$. В частности, согласно принципу инвариантности статистика Колмогорова и статистика омега-квадрат сходятся по распределению к распределениям соответствующих функционалов от случайного процесса ξ:

$$K_n = \sup_{0 \le t \le 1} |\xi_n(t)| \Rightarrow \sup_{0 \le t \le 1} |\xi(t)|, \; \omega_n^2 = \int_0^1 \xi_n^2(t)dt \Rightarrow \int_0^1 \xi^2(t)dt.$$

Таким образом, от проблем прикладной статистики сделан переход к теории случайных процессов. Методами этой теории найдены распределения случайных величин

$$\sup_{0 \le t \le 1} |\xi(t)|, \; \int_0^1 \xi^2(t)dt.$$

Принцип инвариантности – инструмент получения предельных распределений функций от результатов наблюдений, используемых в прикладной статистике.

Обоснование принципу инвариантности может быть дано на основе теории сходимости вероятностных мер в функциональных пространствах [9, 15]. Более простой подход, позволяющий к тому же получать необходимые и достаточные условия в предельной теории статистик интегрального типа (принцип инвариантности к ним нельзя применить), рассмотрен в [16].

Почему «принцип инвариантности» так назван? Обратим внимание, что предельные распределения рассматриваемых статистик не зависят от их функции распределения $F(x)$. Другими словами, предельное распределение инвариантно относительно выбора $F(x)$.

В более широком смысле термин «принцип инвариантности» применяют тогда, когда предельное распределение не зависит от тех или иных характеристик исходных распределений [1]. В этом смысле наиболее известный «принцип инвариантности» – это Центральная предельная теорема, поскольку предельное стандартное нормальное распределение – одно и то же для всех возможных распределений независимых одинаково распределенных слагаемых (лишь бы слагаемые имели конечные математическое ожидание и дисперсию).

Литература

1. Вероятность и математическая статистика: Энциклопедия / Гл. ред. Ю.В. Прохоров. – М.: Большая Российская энциклопедия, 1999. – 910с.

2. Гнеденко Б.В. Курс теории вероятностей: Учебник. 7-е изд., исправл. - М.: Эдиториал УРСС, 2001. – 320 с.

3. Орлов А.И. Средние величины и законы больших чисел в пространствах произвольной природы / А.И. Орлов // Политематический сетевой электронный научный журнал Кубанского государственного аграрного университета (Научный журнал КубГАУ) [Электронный ресурс]. – Краснодар: КубГАУ, 2013. – №05(089). С. 556 – 586. – IDA [article ID]: 0891305038. – Режим доступа: http://ej.kubagro.ru/2013/05/pdf/38.pdf

4. Орлов А. И. О средних величинах // Управление большими системами. Выпуск 46. М.: ИПУ РАН, 2013. С.88-117.

5. Рао С.Р. Линейные статистические методы и их применения. – М.: Наука, 1968. – 548 с.

6. Орлов А.И. Системная нечеткая интервальная математика (СНИМ) – перспективное направление теоретической и вычислительной математики / А.И. Орлов, Е.В. Луценко // Политематический сетевой электронный научный журнал Кубанского государственного аграрного университета (Научный журнал КубГАУ) [Электронный ресурс]. – Краснодар: КубГАУ, 2013. – №07(091). С. 255 – 308. – IDA [article ID]: 0911307015. – Режим доступа: http://ej.kubagro.ru/2013/07/pdf/15.pdf

7. Орлов А.И. О развитии статистики объектов нечисловой природы / А.И. Орлов // Политематический сетевой электронный научный журнал Кубанского государственного аграрного университета (Научный журнал КубГАУ) [Электронный ресурс]. – Краснодар: КубГАУ, 2013. – №09(093). С. 273 – 309. – IDA [article ID]: 0931309019. – Режим доступа: http://ej.kubagro.ru/2013/09/pdf/19.pdf

8. Орлов А.И. Организационно-экономическое моделирование : учебник : в 3 ч. Ч.3. Статистические методы анализа данных. - М.: Изд-во МГТУ им. Н.Э. Баумана, 2012. - 624 с.

9. Прохоров Ю. В. Сходимость случайных процессов и предельные теоремы теории вероятностей // Теория вероятностей и ее применения. 1956. Т. 1, № 2. С. 177–238.

10. Келли Дж. Общая топология. - М.: Наука, 1968. - 384 с.

11. Орлов А.И. Асимптотическое поведение статистик интегрального типа // Доклады АН СССР. 1974. Т.219. № 4. С. 808-811.

12. Орлов А.И. Устойчивость в социально-экономических моделях. - М.: Наука, 1979. - 296 с.

13. Вентцель Е.С. Теория вероятностей. – М.: Наука, 1964.- 576 с.

14. Орлов А.И. Непараметрические критерии согласия Колмогорова, Смирнова, Омега-квадрат и ошибки при их применении / А.И. Орлов // Политематический сетевой электронный научный журнал Кубанского государственного аграрного университета (Научный журнал КубГАУ) [Электронный ресурс]. – Краснодар: КубГАУ, 2014. – №03(097). С. 647 – 675. – IDA [article ID]: 0971403047. – Режим доступа: http://ej.kubagro.ru/2014/03/pdf/47.pdf

15. Биллингсли П. Сходимость вероятностных мер. – М.: Наука, 1977. - 352 с.

16. Орлов А.И. Предельная теория непараметрических статистик / А.И. Орлов // Политематический сетевой электронный научный журнал Кубанского государственного аграрного университета (Научный журнал КубГАУ) [Электронный

ресурс]. – Краснодар: КубГАУ, 2014. – №06(100). С. 226 – 244. – IDA [article ID]: 1001406011. – Режим доступа: http://ej.kubagro.ru/2014/06/pdf/11.pdf

References

1. Verojatnost' i matematicheskaja statistika: Jenciklopedija / Gl. red. Ju.V. Prohorov. – M.: Bol'shaja Rossijskaja jenciklopedija, 1999. – 910s.
2. Gnedenko B.V. Kurs teorii verojatnostej: Uchebnik. 7-e izd., ispravl. - M.: Jeditorial URSS, 2001. – 320 s.
3. Orlov A.I. Srednie velichiny i zakony bol'shih chisel v prostranstvah proizvol'noj prirody / A.I. Orlov // Politematicheskij setevoj jelektronnyj nauchnyj zhurnal Kubanskogo gosudarstvennogo agrarnogo universiteta (Nauchnyj zhurnal KubGAU) [Jelektronnyj resurs]. – Krasnodar: KubGAU, 2013. – №05(089). S. 556 – 586. – IDA [article ID]: 0891305038. – Rezhim dostupa: http://ej.kubagro.ru/2013/05/pdf/38.pdf
4. Orlov A. I. O srednih velichinah // Upravlenie bol'shimi sistemami. Vypusk 46. M.: IPU RAN, 2013. S.88-117.
5. Rao S.R. Linejnye statisticheskie metody i ih primenenija. – M.: Nauka, 1968. – 548 s.
6. Orlov A.I. Sistemnaja nechetkaja interval'naja matematika (SNIM) – perspektivnoe napravlenie teoreticheskoj i vychislitel'noj matematiki / A.I. Orlov, E.V. Lucenko // Politematicheskij setevoj jelektronnyj nauchnyj zhurnal Kubanskogo gosudarstvennogo agrarnogo universiteta (Nauchnyj zhurnal KubGAU) [Jelektronnyj resurs]. – Krasnodar: KubGAU, 2013. – №07(091). S. 255 – 308. – IDA [article ID]: 0911307015. – Rezhim dostupa: http://ej.kubagro.ru/2013/07/pdf/15.pdf
7. Orlov A.I. O razvitii statistiki ob#ektov nechislovoj prirody / A.I. Orlov // Politematicheskij setevoj jelektronnyj nauchnyj zhurnal Kubanskogo gosudarstvennogo agrarnogo universiteta (Nauchnyj zhurnal KubGAU) [Jelektronnyj resurs]. – Krasnodar: KubGAU, 2013. – №09(093). S. 273 – 309. – IDA [article ID]: 0931309019. – Rezhim dostupa: http://ej.kubagro.ru/2013/09/pdf/19.pdf
8. Orlov A.I. Organizacionno-jekonomicheskoe modelirovanie : uchebnik : v 3 ch. Ch.3. Statisticheskie metody analiza dannyh. - M.: Izd-vo MGTU im. N.Je. Baumana, 2012. - 624 s.
9. Prohorov Ju. V. Shodimost' sluchajnyh processov i predel'nye teoremy teorii verojatnostej // Teorija verojatnostej i ee ppimenenija. 1956. T. 1, № 2. S. 177–238.
10. Kelli Dzh. Obshhaja topologija. - M.: Nauka, 1968. - 384 s.
11. Orlov A.I. Asimptoticheskoe povedenie statistik integral'nogo tipa // Doklady AN SSSR. 1974. T.219. № 4. S. 808-811.
12. Orlov A.I. Ustojchivost' v social'no-jekonomicheskih modeljah. - M.: Nauka, 1979. - 296 s.
13. Ventcel' E.S. Teorija verojatnostej. – M.: Nauka, 1964.- 576 s.
14. Orlov A.I. Neparametricheskie kriterii soglasija Kolmogorova, Smirnova, Omega-kvadrat i oshibki pri ih primenenii / A.I. Orlov // Politematicheskij setevoj jelektronnyj nauchnyj zhurnal Kubanskogo gosudarstvennogo agrarnogo universiteta (Nauchnyj zhurnal KubGAU) [Jelektronnyj resurs]. – Krasnodar: KubGAU, 2014. – №03(097). S. 647 – 675. – IDA [article ID]: 0971403047. – Rezhim dostupa: http://ej.kubagro.ru/2014/03/pdf/47.pdf
15. Billingsli P. Shodimost' verojatnostnyh mer. – M.: Nauka, 1977. - 352 s.
16. Orlov A.I. Predel'naja teorija neparametricheskih statistik / A.I. Orlov // Politematicheskij setevoj jelektronnyj nauchnyj zhurnal Kubanskogo gosudarstvennogo agrarnogo universiteta (Nauchnyj zhurnal KubGAU) [Jelektronnyj resurs]. – Krasnodar:

KubGAU, 2014. – №06(100). S. 226 – 244. – IDA [article ID]: 1001406011. – Rezhim dostupa: http://ej.kubagro.ru/2014/06/pdf/11.pdf