УДК 621.396

ПЕРМУТАЦИОННЫЕ КРИТЕРИИ КАК УНИВЕРСАЛЬНЫЙ НЕПАРАМЕТРИЧЕСКИЙ ПОДХОД К ПРОВЕРКЕ СТАТИСТИЧЕСКИХ ГИПОТЕЗ

КОЛЯДИН В.Л.

Рассматриваются логические основания концепции пермутационных критериев и показывается, что она представляет собой общий подход к синтезу непараметрических критериев для проверки статистических гипотез. Такой подход позволяет легко строить новые непараметрические критерии, максимально учитывающие специфику конкретной практической задачи. При этом не требуется применение каких-либо теоретико-вероятностных выкладок и аналитических методов. Рассматриваются основные классы статистических задач, где применение пермутационных критериев наиболее целесообразно.

Введение

Непараметрические критерии являются основным рабочим инструментом проверки статистических гипотез по реальным данным [1-3]. В отличие от параметрических критериев они обеспечивают гарантированное значение вероятности ошибки I рода (отвержения истинной нулевой гипотезы) практически для любого вида закона распределения, а не только для закона, принадлежащего некоторому параметрическому семейству распределений, например, нормальному, как в случае параметрических тестов. Поскольку вид закона распределения для реальных данных чаще всего неизвестен, то на практике предпочитают использовать именно непараметрические критерии.

Пермутационные критерии (синонимы: “критерии перестановок”, “рандомизированные критерии”) представляют неклассический подход к синтезу непараметрических решающих процедур, который обладает беспрецедентной общностью и, одновременно, алгоритмической и концептуальной простотой. Однако в настоящее время общность и универсальность этого подхода в полной мере пока не осознается ни специалистами в области статистических методов, ни теми, кто использует их для решения конкретных практических задач. Например, в классических монографиях [4,5] пермутационные критерии хотя и описаны, но представлены как некоторые специальные критерии, имеющие ограниченную область применимости.

Классический подход к построению непараметрических критериев является скорее искусством, чем наукой, и требует высочайшей квалификации в области математики и теории вероятностей. Усилиями математиков мирового уровня за 70-80 лет в рамках этого подхода были получены всего несколько десятков базовых непараметрических критериев и несколько сотен процедур, производных от базовых критериев. При этом к конкретной практической задаче, как правило, применимы лишь немного известных непараметрических критериев, что не позволяет выбрать критерий, обеспечивающий максимальную мощность именно в данной задаче.

При использовании подхода, основанного на пермутационных критериях, ситуация радикально отличается от описанной выше. Достаточно просто можно синтезировать большое (строго говоря, бесконечное) количество пермутационных непараметрических критериев, ориентированных на специфику конкретной практической задачи, т.е. обладающих существенно большей мощностью в отношении специфических для данной задачи альтернативных гипотез, чем классические непараметрические критерии. При этом синтез пермутационных критериев не требует проведения каких-либо теоретико-вероятностных выкладок (хотя вычислительные затраты на реализацию пермутационного критерия могут быть существенно выше, чем на реализацию классического непараметрического критерия).

С увеличением мощности и доступности вычислительных средств интерес к пермутационным критериям постоянно растет [6,7]. Однако еще рано говорить о повсеместном признании и понимании достоинств этого неклассического подхода даже среди специалистов в области статистики, не говоря уже о многочисленном сообществе потенциальных пользователей - специалистов, сталкивающихся с необходимостью проверки статистических гипотез в конкретных прикладных областях. Одна из целей этой статьи — постараться исправить такую ситуацию и привлечь внимание широкой аудитории пользователей статистических методов к методологии пермутационных тестов как удобному практическому инструменту статистического анализа.

Основная цель статьи — показать, что пермутационные критерии представляют собой достаточно универсальную, практичную и доступную методологию оперативного синтеза непараметрических критериев, а не несколько специальных критериев, ориентированных на узкий круг задач. Предлагаем такой ракурс рассмотрения проблемы, при котором видна преемственность между классическим подходом к синтезу непараметрических критериев и подходом, основанным на применении методологии пермутационных критериев.

1. Классический подход к синтезу непараметрических критериев

При классическом подходе к построению непараметрических критериев основной задачей является нахождение такой статистики $T(\cdot)$ — функции от N независимых наблюдений $X = (x_1, ..., x_N)$ случайной величины x, что ее выборочное распределение (т.е. распределения значений $T(X)$) обладает следующими двумя важными свойствами:

$$P(T(X) \mid p(x)) = P_0(T), \ \forall p(x) \in R_0, \quad (1)$$

$$P(T(X) \mid p(x)) > P_0(T), \ \forall p(x) \in R_1, \quad (2)$$

где $p(x)$ — плотность распределения случайной величины x; $P(T \mid p)$ — интегральная функция выборочного распределения статистики T при заданной плотности p; R_0, R_1 — множества всех плотностей распределения, отвечающих нулевой (H_0) и альтернативной (H_1) гипотезам, соответственно; $P_0(\cdot)$ — интегральное выборочное распределение статистики T при нулевой гипотезе.

Свойство (1) означает, что выборочное распределение статистики T одно и то же для любого закона $p(x)$ распределения наблюдений, удовлетворяющих нулевой гипотезе. Выполнение этого требования необходимо для обеспечения заданной вероятности α ошибок I рода. Свойство (2) означает, что при альтернативной гипотезе H_1 статистика $T(\cdot)$ в среднем чаще принимает большее значение, чем при нулевой гипотезе H_0. Это необходимо для обеспечения ненулевой мощности теста — его способности обнаруживать отклонения от нулевой гипотезы.

Решение принимается на основе наблюдаемого значения $T(X)$ следующим образом:

$$\widetilde{H}(X) = \begin{cases} H_0, \text{если } T(X) \le t_\alpha, \\ H_1, \text{если } T(X) > t_\alpha, \end{cases} \tag{3}$$

а значение t_α определяется из условия

$$P_0(t_\alpha) = 1 - \alpha,$$

где α — заданный уровень значимости критерия (максимально допустимая вероятность ошибки I рода — отвержения истинной нулевой гипотезы).

Свойства (1) и (2) далеко не равноценны с точки зрения сложности процесса синтеза критерия (т.е. нахождения статистики T с требуемыми свойствами). Удовлетворить требованию (2) достаточно просто, а требованию (1) — необычайно трудно. Именно поэтому обладающие свойствами (1) и (2) функции, найденные математиками за более чем полувековую историю непараметрической статистики, по праву считаются жемчужинами статистической науки, а описание их свойств составляет основное содержание монографий и справочников по непараметрической статистике [1,3,8].

Очевидно, что при таком подходе к синтезу статистических критериев рядовому пользователю остается только выбрать один из набора уже известных, более или менее подходящий к решаемой задаче. При этом, применительно к любой конкретной задаче, выбор оказывается невелик и ограничен несколькими, а иногда и одним непараметрическим критерием, зачастую мало чувствительным к ожидаемому виду альтернативной гипотезы.

2. Базовая идея пермутационных критериев

Применение методологии пермутационных критериев позволяет радикально изменить описанную выше ситуацию, характерную для классического подхода. Даже рядовой пользователь, применяя эту методологию, может легко синтезировать большое количество непараметрических критериев, максимально учитывающих специфику решаемой задачи. При этом автоматически обеспечивается выполнение требования (1), т.е. обходится основная трудность, характерная для классического подхода.

Исторически первым пермутационным критерием, по всей видимости, является “точный критерий Фишера” (Fisher’s exact test) [6], ориентированный на проверку гипотезы о равенстве пропорций (вероятностей “успеха”) в двух генеральных совокупностях по конечным выборкам из них. Этот критерий и сегодня часто используется на практике. Достаточно сложно указать, кто и когда первым осознал, что пермутационные критерии — это не просто несколько специальных критериев, а достаточно общий подход к построению непараметрических критериев вообще. По всей видимости, это осознание было медленным и постепенным (эволюционным) процессом, катализированным компьютеризацией, ростом доступности и мощности компьютеров. Сегодня можно констатировать как факт, что в 1990 г. многие авторы уже достаточно хорошо понимали, что пермутационные критерии представляют общую методологию построения непараметрических критериев [6].

Идея, лежащая в основе пермутационных критериев, относительно проста. Пусть $X = \{x_1, ..., x_N\}$ — случайная выборка однотипных элементов (скаляров, векторов, матриц и т.д.). Известно, что плотности распределения $p_0(X)$ и $p_1(X)$ последовательности X, соответствующие нулевой и альтернативной гипотезе, обладают следующим свойством:

$$p_0(X) = p_0(\pi(X)), \ \forall \pi \in \Pi, \tag{4}$$

$$p_1(X) \ne p_1(\pi(X)), \ \exists \pi \in \Pi, \tag{5}$$

где Π — некоторая конечная группа перестановок π элементов x_i последовательности X или некоторых их компонент (если элементы x_i являются векторами). Свойства (4), (5) означают, что при нулевой гипотезе плотность распределения X инвариантна относительно любой перестановки из группы Π, а при альтернативной — не инвариантна относительно некоторых перестановок из Π.

Пусть $t(X)$ — произвольная функция от X, называемая далее первичной статистикой пермутационного критерия или просто первичной статистикой. Рассмотрим следующую статистику, формируемую на основе наблюдаемой реализации X_o последовательности X и первичной статистики $t(\cdot)$:

$$T(X_0; t(\cdot)) = \frac{1}{K} \sum_{k=1}^{K} I(X_o, t(\cdot), \pi_k), \tag{6}$$

здесь K — количество всех перестановок в группе Π; $I(\cdot)$ — индикаторная функция:

$$I(X, t(\cdot), \pi) = \begin{cases} 1 & \text{если } t(\pi(X)) \ge t(X), \\ 0 & \text{если } t(\pi(X)) = t(X). \end{cases}$$

Другими словами, значение $T(X_o)$ — это доля тех перестановок из группы Π, после действия которых на выборку X_o значение $t(\pi(X_o))$ первичной статистики остается не меньшим, чем ее наблюдаемое значение $t(X_o)$.

Статистика T принимает значения из множества $\{k / K,\ k = 1,\dots,K\}$, и при истинной нулевой гипотезе обладает следующим уникальным свойством:

$$\mathrm{Prob}(T(X) \le \alpha) \ \le\ \alpha\,, \qquad (7)$$

где $\mathrm{Prob}(\cdot)$ — вероятность события. Далее статистику $T(\cdot)$ будем называть результирующей или просто статистикой критерия. Целесообразность такого выбора терминов в отношении статистик $t(X)$ и $T(X;\ t(\cdot))$ обсуждается ниже в разделе 4.

Свойство (7) означает, что принятие решения на основе правила

$$H(X) = \begin{cases} H_0, \text{если } T(X) > \alpha, \\ H_1, \text{если } T(X) \le \alpha, \end{cases} \qquad (8)$$

гарантирует, что при истинной нулевой гипотезе вероятность ошибочного принятия альтернативной гипотезы (ошибки I рода) не превышает α. Напомним, что свойство (7) выполняется для любой первичной статистики $t(\cdot)$ и для любого закона $p_0(X)$ распределения X, отвечающего нулевой гипотезе и удовлетворяющего свойству (4). Именно это обстоятельство и позволяет рассматривать пермутационные критерии как универсальный подход к синтезу непараметрических критериев проверки статистических гипотез.

Свойство (7) обеспечивает выполнение условия (1), т.е. инвариантность выборочного распределения статистики T критерия при нулевой гипотезе независимо от выбора первичной статистики $t(\cdot)$. Это, в свою очередь, позволяет свести синтез нового непараметрического критерия к выбору функции $t(\cdot)$, максимально чувствительной к определенному виду отклонений от нулевой гипотезы, т.е. сосредоточить усилия на выполнении условия (2).

3. Статистическое обоснование основного свойства пермутационных критериев

Идея доказательства свойства (7) основывается на разбиении множества Ω всех исходов (последовательностей X), возможных при истинной нулевой гипотезе, на конечные подмножества последовательностей, совпадающие с точностью до произвольной перестановки из группы Π. Разобьем множество Ω на непересекающиеся подмножества размера K, каждое их которых образовано всеми перестановками (из группы Π) некоторой последовательности X. Рассмотрим одно такое подмножество $Q(X) = \{\pi(X):\ \pi \in \Pi\}$.

Заметим, что все K элементов такого подмножества при истинной нулевой гипотезе равновероятны в силу условия (4).

Сначала рассмотрим случай, когда значения первичной статистики $t(\cdot)$ различны для всех K элементов подмножества $Q(X)$. Расположим (проиндексируем) элементы подмножества Q в порядке убывания значения первичной статистики $t(X)$:

$$t(X_1) > t(X_2) > \dots > t(X_K),\ X_k \in Q\,. \qquad (9)$$

Тогда для любого $k = 1,\dots,K$ автоматически выполняется следующее свойство:

$$T(X_k,\ t(\cdot)) = \frac{k}{K},\ \ k = 1,\dots,K\,, \qquad (10)$$

где T — результирующая статистика, определяемая выражением (6).

Допустим, что нулевая гипотеза верна, а решение о ее истинности принимается на основе решающего правила (8). Поскольку все элементы $\{X_k,\ k = 1,\dots,K\}$ подмножества $Q(X)$ при нулевой гипотезе равновероятны в силу условия (4), то из выражения (10) непосредственно следует, что результирующая статистика T принимает равновероятно значения из множества $\{k / K,\ k = 1,\dots,K\}$. Следовательно, для любого $\alpha \in [0,1]$ выполняется неравенство

$$\mathrm{Prob}(T(X, t(\cdot)) \le \alpha \mid X \in Q) \ \le\ \alpha\,, \qquad (11)$$

которое для конечного множества значений $\{\alpha = k / K,\ k = 1,\dots,K\}$ обращается в строгое равенство.

Поскольку неравенство (11) выполняется для любого подмножества Q множества Ω всех возможных последовательностей X (исходов), а различающиеся подмножества $Q(X)$ не пересекаются, то можно легко перейти от условной вероятности (11) к безусловной:

$$\begin{aligned} &\mathrm{Prob}(T(X, t(\cdot)) \le \alpha) = \\ &= \int_S \mathrm{Prob}(T(X, t(\cdot)) \le \alpha \mid X \in Q)\, p(Q) dQ\,, \end{aligned} \qquad (12)$$

где S — множество всех непересекающихся подмножеств Q; $p(Q)$ — плотность вероятности подмножества Q. В силу того, что значение $p(Q)$ неотрицательно, из (11) и (12) непосредственно следует свойство (7) для любого вида плотности распределения $p(Q)$.

Теперь рассмотрим общий случай, когда значения $t(X)$ для некоторых элементов X подмножества Q могут совпадать, т.е. некоторые строгие неравенства в выражении (9) заменяются нестрогими. В этом случае выражение (10) принимает вид

$$T(X_k) \ge k / K;\ k = 1,\dots,K \qquad (13)$$

в силу того, что значения T из-за наличия совпадающих значений $t(X_k)$ могут только возрастать по сравнению со случаем (9). Например, если $t(X_1) = t(X_2) > t(X_3)\dots$, то $T(X_1) = T(X_2) = 2 / K$, а не $1 / K$. Поскольку решающее правило (8) предусматривает отвержение нулевой гипотезы при значениях $T(X, t(\cdot))$, меньших или равных порогу a, то вероятность отвержения истинной нулевой гипотезы при наличии таких совпадающих значений только уменьшается, гарантированно сохраняя справедливость неравенства (7).

4. Связь между классическими и пермутационными непараметрическими критериями

При изложении материала, касающегося пермутационных критериев, в литературе [4,5,8] логическая связь между классическими непараметрическими и пермутационными критериями практически не видна. Это связано с неудачным выбором терминологии и понятийного аппарата - когда функцию $t(\cdot)$ рассматривают как статистику критерия. В классических непараметрических критериях статистикой критерия называют ту функцию $T(\cdot)$, на основании значений $T(X)$ которой принимается решение и которая обладает свойством (1) независимости своего выборочного распределения от конкретной нулевой гипотезы. Очевидно, что в пермутационных критериях функция $t(\cdot)$ этими свойствами не обладает. Поэтому рассмотрение $t(\cdot)$ как статистики критерия вряд ли можно признать удовлетворительным, по крайней мере, с точки зрения преемственности между теорией классических и пермутационных критериев.

Для выявления такой преемственности в этой статье функцию $t(\cdot)$ мы называем лишь "первичной статистикой", а в качестве собственно статистики критерия рассматриваем функцию $T(X, t(\cdot))$, которая определяется выражением (6) и помимо самих данных X зависит и от первичной статистики $t(\cdot)$. Эта функция T ("результирующая статистика") полностью соответствует классическому понятию "статистика непараметрического критерия": она удовлетворяет требованию (1), и именно на основе ее значений принимается конечное решение. Такой взгляд на базовые понятия позволяет считать пермутационные критерии общей методологией синтеза непараметрических критериев. Фактически концепция пермутационного критерия позволяет на основе произвольной статистики $t(\cdot)$ автоматически получить статистику T, порождающую некоторый непараметрический критерий.

Понимание логических основ пермутационных критериев частично затрудняется непривычными свойствами результирующей статистики $T(X;\ t(\cdot))$. Во-первых, она зависит от другой статистики $t(\cdot)$ и вычисляется согласно (6) довольно непривычным образом, требующим необычно больших вычислительных затрат даже при умеренном размере выборки. Во-вторых, статистика T практически "идеальна" в следующем смысле: ее выборочное распределение при нулевой гипотезе очень близко к равномерному в интервале [0,1]. Последнее означает, что не требуется каких-либо таблиц или алгоритмов для расчета p-значения (максимального значения уровня значимости α, при котором нулевая гипотеза еще не отвергается). Наблюдаемое значение $T(X_o)$ этой статистики фактически сравнивается непосредственно с заданным уровнем значимости α, а не с некоторым пороговым значением, вычисляемым на основе таблиц или алгоритмически по заданному α.

Такое необычное сочетание свойств статистики T маскирует ее реальное место в логической структуре пермутационных критериев. Обычно при изложении идеи пермутационных критериев авторы вообще не рассматривают статистику (6) в явном виде, а статистикой критерия называют первичную статистику $t(\cdot)$. При этом процедуру вычисления значения $T(X_o)$ из выражения (6) рассматривают как процесс, состоящий из двух логических ступеней: 1) вычисления условного (по наблюдаемым данным X_o) распределения статистики $t(\cdot)$ при нулевой гипотезе; 2) определения уровня значимости наблюдаемого значения $t(X_o)$ для этого распределения (см., например, классическую монографию [4]).

5. Практическая реализация пермутационных критериев

Непосредственное вычисление результирующей статистики $T(X, t(\cdot))$ из выражения (6) обычно затруднительно из-за большого значения K, которое быстро увеличивается с ростом размера N данных X (приблизительно как факториал от размера выборки N). Поэтому на практике для расчета значения $T(X, t(\cdot))$ чаще всего прибегают к приближенным методам Монте-Карло. При этом используют случайную выборку (достаточно большого размера $M << K$) перестановок из группы Π. Общий алгоритм такой приближенной реализации пермутационных критериев следующий:

Шаг 1. Вычисляем наблюдаемое значение $t_o = t(X_o)$ первичной статистики $t(\cdot)$ для наблюдаемых данных X_o.

Шаг 2. Выбираем случайно и равновероятно элемент π из группы перестановок Π и вычисляем значение $t_k = t(\pi(X_o))$ первичной статистики от преобразованных данных $\pi(X_o)$. (Чаще всего, Π представляет собой группу всех перестановок в последовательности из N элементов. В этом случае применяют известные алгоритмы случайной перестановки N элементов.)

Шаг 3. Повторяем шаг 2 M раз и определяем количество m реализаций шага 2, при которых выполнилось условие $t(\pi(X_o)) \geq t(X_o)$. (Здесь и далее предполагается, что первичная статистика $t(\cdot)$ выбрана таким образом, что ее большие значения свидетельствуют против нулевой гипотезы. В противном случае просто используется обратное неравенство).

Шаг 4. В качестве оценки p-значения используем величину $\hat{T} = m / M$ — оценку значения результирующей тестовой статистики T.

Если стоит классическая задача проверки нулевой гипотезы при заданной вероятности α ошибки I рода, то решение принимается на основе правила (8). При интерпретации экспериментальных данных, где зачастую принято сообщать само p-значение, величина $\hat{T}$, вычисленная на шаге 4 алгоритма, используется в качестве конечного p-значения.

6. Задача сравнения двух законов распределения

Рассмотрим практически важный класс задач, когда требуется проверить нулевую гипотезу о совпадении законов распределения двух генеральных совокупностей по случайным выборкам X_1 и X_2 размера N_1 и N_2 из этих генеральных совокупностей . Напомним, что такая задача рассматривается здесь в самой общей постановке: элементы выборок могут быть произвольной природы, т.е. описываться скалярами, векторами, матрицами и т.д., а о законе распределения ничего не известно. Единственное требование – статистическая независимость элементов выборок.

Рассмотрим объединенную выборку

$$X = X_1 \cup X_2 = (x_1, ..., x_N);\ N = N_1 + N_2;$$

где

$$x_i \in X_1,\ i = 1, ..., N_1;$$
$$x_i \in X_2,\ i = N_1 + 1, ..., N_1 + N_2;\ .$$

Тогда при истинной нулевой гипотезе совместная плотность распределения $p_0^N(X)$ всех элементов объединенной выборки X имеет вид:

$$p_0(X) = \prod_{i=1}^{N} q_0(x_i);\ N = N_1 + N_2; \qquad (14)$$

здесь $q_0(x)$ – неизвестная плотность распределения генеральных совокупностей при нулевой гипотезе. Плотность вида (14) инвариантна относительно любой из $N!$ перестановок элементов объединенной выборки X. Поэтому группа Π образована этими $N!$ перестановками. Тогда шаги 1 и 2 общего алгоритма, описанного в разделе 5, реализуются следующим образом.

Шаг 1. Выбираем первичную статистику $t(\cdot,\cdot)$, чувствительную к представляющему интерес виду отклонения от нулевой гипотезы, и вычисляем наблюдаемое значение $t_o = t(X_1, X_2)$ этой статистики.

Шаг 2. Выполняем случайную перестановку p всех элементов объединенной выборки X, разбиваем результирующую последовательность $X' = \pi(X)$ на две последовательности X_1' и X_2' размеров N_1 и N_2, и затем вычисляем значение $t' = t(X_1', X_2')$.

Шаги 3 и 4 те же, что в общем алгоритме, описанном в разделе 5.

Заметим, что для столь общей нулевой гипотезы (14) возможно большое количество альтернативных гипотез, каждая из которых соответствует конкретной практической задаче. Практический интерес могут представлять различия некоторых статистических характеристик двух распределений, например их центрального положения, ширины распределения, степени тяжести хвоста распределений, асимметрии и т.д. В этих случаях в качестве первичной статистики $t(\cdot,\cdot)$ может использоваться модуль разности значений соответствующей статистической меры для выборок 1 и 2:

$$t(X_1, X_2) = |C(X_1) - C(X_2)|, \qquad (15)$$

где C – некоторая мера сравнения, характеризующая количественно то свойство распределений, различие по которому представляет основной практический интерес. Например, если таким свойством является центральное положение, то в качестве $C(\cdot)$ можно выбрать среднее, медиану, усеченное среднее и т.д. Если интерес представляет различие в ширине (или масштабе) распределения, то разумным выбором является дисперсия, среднее абсолютное уклонение от медианы, разность 75-й и 25-й процентных точек и т.д.

Уже из этих элементарных примеров видно, что выбор первичной статистики вида (15) осуществляется достаточно просто, исходя из специфики решаемой задачи. При этом общая схема пермутационных критериев обеспечивает автоматическое выполнение свойства (1) инвариантности выборочного распределения результирующей тестовой статистики $T(X, t(\cdot))$ вида (6) на всем множестве нулевых гипотез вида (14), т.е. независимо от вида плотности $q_0(x)$.

Заметим также, что первичные статистики вида (15) представляют лишь частный случай (хотя они и порождают широкий класс решающих правил за счет произвольного выбора меры C). В качестве более сложного случая можно привести пример ситуации, когда элементы x_i представляют собой не скалярные величины, а функции одного или большего числа аргументов, например, независимые реализации $x_i(t)$ сигналов, полученных для двух классов объектов. Практический интерес здесь может представлять, например, нулевая гипотеза о совпадении спектров мощности $S_1(\omega)$ и $S_2(\omega)$ для этих двух классов объектов. В качестве первичной статистики можно, например, выбрать функционал:

$$t(X_1, X_2) = \int |\hat{S}(\omega; X_1) - \hat{S}(\omega; X_2)|^\beta\, d\omega,$$

где $\hat{S}(\omega, X)$ – некоторая оценка спектра по выборке $X = \{x_1(t), ..., x_N(t)\}$; β – параметр, влияющий на свойства такой меры различия двух спектров. При этом аргументы t и ω могут быть векторными, т.е. в эту схему укладываются и задачи многомерного спектрального анализа.

Несмотря на возросшую сложность задачи, общая схема пермутационных критериев в равной степени применима и здесь, а ее алгоритм остается тем же, что и для более простых задач.

7. Задача сравнения нескольких законов распределения

В случае нескольких выборок подход к синтезу пермутационных критериев остается практически тем же, за исключением выбора первичной статистики $t(X_1, ..., X_L)$, которая должна быть чувствительна к отличию хотя бы одного из L распределений от остальных $L-1$ распределений. Выбор первичной статистики в этом случае шире, чем в случае двух выборок. Даже если строить первичные статистики только на основе некоторой статистической меры (как это делается в выражении (15) для случая двух выборок), то выбор все же оказывается гораздо шире. Следующее выражение, например, порождает широкий круг первичных статистик:

$$t(X_1,...,X_L) = \sum_{l=1}^{L} |C(X_l) - Avg(C(X_1),...,C(X_L))|^{\beta} ,$$

где $C(\cdot)$ – некоторая мера, характеризующая ту статистическую характеристику распределений, различие по которой представляет интерес в контексте конкретной практической задачи (как и в описанном выше случае двух выборок); $Avg(\cdot)$ – некоторый оператор “усреднения”, например, оператор вычисления среднего значения или медианы. Выбор показателя степени β позволяет изменять чувствительность первичной статистики (и, следовательно, мощности критерия) относительно отдельных типов альтернативной гипотезы. Большие значения β делают критерий более чувствительным к отклонению одного (или малого числа) из L распределений от остальных; меньшие значения β – к тем ситуациям, когда все или большая часть распределений взаимно различаются.

8. Задача обнаружения статистической связи

Другим важным классом задач, где применимы пермутационные критерии, являются задачи проверки нулевой гипотезы об отсутствии статистической связи между случайными величинами x и y, или об отсутствии зависимости случайной величины y от x. Исходными данными в этом случае является неупорядоченное множество из N независимых пар значений $\{(x_1,y_1),...,(x_N,y_N)\}$ или, что эквивалентно, два упорядоченных множества $X = (x_1,...,x_N)$ и $Y = (y_1,...,y_N)$. При этом элементы последовательностей X и Y могут быть произвольной природы – не обязательно скалярными величинами. Ниже перечислены особенности основных подклассов задач этого типа, отличающихся видом альтернативной гипотезы и природой величин.

1. Величины x и y являются случайными, и интерес представляет альтернативная гипотеза о наличии некоторой статистической связи (не обязательно линейной) между x и y. В этом случае в качестве первичной статистики $t(X,Y)$ можно выбирать некоторую меру корреляции, например, коэффициент линейной корреляции (Пирсона), коэффициент ранговой корреляции (Спирмена) и т.д. Когда x и y принимают значения из конечных множеств (и, следовательно, наблюдаемые пары значений могут быть представлены в виде таблицы сопряженности), в качестве первичной статистики $t(X,Y)$ можно использовать одну из многочисленных известных мер статистической связи [9].

2.Величина x не является случайной; альтернативная гипотеза предполагает наличие определенной связи случайной величины y с x. Примером такой ситуации является классическая задача статистической радиотехники – обнаружение сигнала известной формы $x(t)$ по наблюдениям его аддитивной смеси с шумом:

$$y_i = Ax_i + n_i;\ y_i = y(t_i),\ x_i = x(t_i),$$

где t_i – моменты формирования отсчетов принимаемого колебания $y(t)$; n_i – статистически независимые и одинаково распределенные отсчеты аддитивного шума (о законе распределения которых ничего не известно); A – неизвестная амплитуда сигнала. В качестве первичной статистики можно использовать некоторую меру корреляции последовательностей X и Y (в простейшем случае – модуль коэффициента корреляции).

3. Альтернативная гипотеза предполагает наличие определенной зависимости y от x:

$$y = f(x_i;\ \lambda) + n_i,$$

здесь $f(x;\ \lambda)$ – некоторая функция, известная с точностью до параметра λ. При этом в качестве первичной статистики $t(X,Y)$ целесообразно выбрать разность между некоторой оценкой $\hat{\lambda}(X,Y)$ параметра λ и значением λ, соответствующим нулевой гипотезе (или, в случае двусторонней альтернативы, модуль этой разности). Например, в простейшем случае линейной зависимости $y = a + bx$ в качестве $t(X,Y)$ может использоваться оценка $\hat{b}(X,Y)$ коэффициента b наклона в уравнении линейной регрессии. Уклонения значений этой оценки от нуля свидетельствуют в пользу альтернативной гипотезы о наличии зависимости.

Независимо от выбранной первичной статистики и интерпретации величин x (как случайных или фиксированных) общая схема пермутационного критерия здесь одна и та же: в качестве группы Π перестановок используется группа $N!$ перестановок элементов x_i (или y_i, что статистически эквивалентно) среди N пар наблюдаемых значений [4]. Шаги 1 и 2 общего алгоритма, описанного в разделе 5, в этом случае реализуются следующим образом.

Шаг 1. Выбираем первичную статистику одним из описанных выше способов, исходя только из соображений ее чувствительности к ожидаемому виду отклонения от нулевой гипотезы, и вычисляем наблюдаемое значение $t_o = t(X,Y)$ этой статистики.

Шаг 2. Выбираем случайную перестановку π всех элементов последовательности X (или Y) и вычисляем значение $t' = t(\pi(X),Y)$.

Шаги 3 и 4 те же, что в общем алгоритме, описанном в разделе 5.

9. Пример. Проверка гипотезы об идентичности двух участков зондируемой поверхности

Рассмотрим следующую задачу дистанционного радиозондирования. При зондировании двух участков поверхности получены два набора значений амплитуд отраженного сигнала – X_1 и X_2. Условия зондирования полагаются полностью идентичными (в противном случае амплитуды полагаем приведенными к некоторым стандартным условиям). При этом индивидуальные значения амплитуд для каждого из участков получены для неперекрывающихся элементов разрешения и, следовательно, могут считаться статистически независимыми (т.е. X_1 и X_2 можно считать случайными выборками). Практический интерес представляет нулевая гипотеза о том, что оба участка полностью статистически идентичны по своим отражающим свойствам. Альтернативная гипотеза предполагает, что участки отличаются значениями масштабного множителя, характеризующего среднюю амплитуду отраженного сигнала. Относительно вида закона распределения амплитуд при нулевой гипотезе ничего не известно и, более того, оно заведомо может отличаться от распределения Релея-Райса, предсказываемого простейшими моделями процесса рассеяния.

В качестве первичных статистик будем использовать 3 функции вида (15) при следующих статистических мерах $C(X)$:

$$C_1(X) = \frac{1}{N}\sum_{i=1}^{N} x_i^2, \quad C_2(X) = \frac{1}{N}\sum_{i=1}^{N} x_i,$$

$$C_3(X) = \mathrm{median}(X),$$

где N – размер выборки X; median(X) – медиана выборки. Обратим внимание на то, что выбор первичных статистик достаточно тривиален. В качестве меры “средней” интенсивности отраженных сигналов выбраны средний квадрат, среднее значение и медиана амплитуды отраженного сигнала, а в качестве первичной статистики $t(\cdot,\cdot)$ – модуль разности этой меры для двух участков.

Общая схема пермутационных критериев для этого класса задач, описанная в разделах 5 и 6, позволяет автоматически получить непараметрический критерий по заданной первичной статистике. Таким образом, выбрав три первичные статистики, мы фактически задали три непараметрических критерия. Ниже приводятся результаты статистического моделирования этих трех пермутационных критериев.

Моделирование проводилось для выборок размера $N_1 = N_2 = 25$ с двукратным различием в средней амплитуде отраженного сигнала. При реализации пермутационного критерия использовались $M = 1000$ случайных перестановок объединенной выборки. Оценки вероятности ошибки II рода (принятия нулевой гипотезы об идентичности двух участков) формировались по 1000 реализациям.

Сначала исследовалась ситуация, когда амплитуды отраженного сигнала для каждого из участков подчиняются закону распределения Релея-Райса (с-распределения с 2 степенями свободы), предсказываемого простейшей моделью отражения. Результаты моделирования для нескольких значений уровня значимости критерия приведены в табл. 1.

Таблица 1

Относительная частота ошибок II рода при распределении амплитуд по закону Релея-Райса

	Уровень значимости (α)			
Мера сравнения	0,01	0,025	0,05	0,1
Средний квадрат	0.032	0.011	0.006	0.002
Среднее	0.044	0.016	0.007	0.005
Медиана	0.133	0.083	0.050	0.027

Из результатов табл.1 видно, что применение первичной статистики, основанной на среднем квадрате амплитуд, в данном случае обеспечивает максимальную мощность критерия. При этом использование медианы приводит к существенно меньшей мощности критерия.

Исследовалась также ситуация, когда закон распределения амплитуд существенно отличается от закона Релея-Райса в сторону наличия у распределения более тяжелого хвоста. При этом были использованы распределения с плотностью следующего вида:

$$p(x;\ \beta,\delta) = (1-\beta)p_{rr}(x) + \beta\frac{1}{\delta}p_{rr}(x/\delta),$$

где $p_{rr}(\cdot)$ – плотность распределения Релея-Райса; β и δ – параметры тестового распределения. Такие тестовые распределения соответствуют ситуации, когда зондируется поверхность смешанного типа, и в элемент разрешения с вероятностями $1-\beta$ и β могут попадать два типа поверхности, для которых амплитуды отраженных сигналов в среднем различаются в δ раз. При этом полагаем, что для каждого типа поверхности амплитуда отраженного сигнала подчиняется распределению Релея-Райса.

Результаты статистического моделирования для значений параметров β =0,2 и δ =3 приведены в табл. 2.

Таблица 2

Относительная частота ошибок II рода для тестового распределения с тяжелым хвостом

	Уровень значимости (α)			
Мера сравнения	0,01	0,025	0,05	0,1
Средний квадрат	0.542	0.433	0.329	0.231
Среднее	0.375	0.257	0.169	0.098
Медиана	0.334	0.206	0.118	0.057

Из результатов табл. 2 видно, что наибольшую мощность имеет критерий, использующий медиану в качестве меры сравнения, а наименьшую – критерий, использующий средний квадрат амплитуды. Иными словами, ситуация противоположна той, которая наблюдалась для распределения Релея-Райса (табл. 1). Наблюдается также существенное различие между результатами и в мощности наилучшего из трех критериев - наличие у распределения амплитуд более тяжелого хвоста привело к существенному снижению мощности как каждого критерия в отдельности, так и наилучшего из трех критериев.

Выводы

1. Пермутационные критерии представляют общую методологию синтеза непараметрических критериев, а не узкоспециальные статистические методы.

2. Данная методология позволяет легко синтезировать новые непараметрические критерии без привлечения каких-либо аналитических методов теории вероятностей и математической статистики. Это, в свою очередь, позволяет строить новые критерии непосредственно в процессе решения конкретной задачи и максимально полно адаптировать критерий к специфике этой задачи.

3. Синтез нового критерия сводится к выбору первичной статистики, чувствительной к отклонению от нулевой гипотезы. При этом методология синтеза пермутационных критериев автоматически обеспечивает формирование результирующей инвариантной статистики на основе выбранной первичной статистики.

4. Методология пермутационных критериев идеально подходит к задачам сравнения двух или нескольких распределений и к задачам обнаружения статистической связи произвольного вида.

Литература: 1. *Холлендер М., Вулф Д.* Непараметрические методы статистики. М.: Финансы и статистика, 1983. 518 с. 2. *Тюрин Ю.Н.* Непараметрические методы статистики. М.: Знание, 1978. 64с. 3. *Siegel S., Castellan N.J.* Nonparametric statistics for the behavioral sciences. New York: McGraw-Hill. 1988. 399 p. 4. *Кокс Д., Хинкли Д.* Теоретическая статистика. М.: Мир, 1978. 560с. 5. *Леман Э.* Проверка статистических гипотез. М.: Наука, 1979. 408 с. 6. *Good P.I.* Resampling methods: a practical guide to data analysis. Boston: Birkhauser, 1999. 269 p. 7. *Davison A.C., Hinkley D.V.* Bootstrap methods and their application. Cambridge: Cambridge University Press, 1997. 582 p. 8. *Рунион Р.* Справочник по непараметрической статистике. Современный подход. М.: Финансы и статистика, 1982. 198 с. 9. *Елисеева И.И., Рукавишников В.О.* Группировка, корреляция, распознавание образов (Статистические методы классификации и измерения связей). М.: Статистика, 1977. 144 с.

Поступила в редколлегию 18.02.2002

Рецензент: д-р техн. наук, проф. Костенко П.Ю.

Колядин Владимир Леонидович, канд. техн. наук, докторант кафедры авиационно-космических радиотехнических систем Национального аэрокосмического университета “ХАИ”. Научные интересы: неклассические методы анализа данных, включая обработку сигналов и изображений. Увлечения и хобби: история и методология науки, теннис. Адрес: Украина, 61129, Харьков, пр. Тракторостроителей, 162-Г, кв.128, тел.14-81-44.

УДК 517.9+538

ВОЗБУЖДЕНИЕ РЕШЕТКИ С ДИЭЛЕКТРИЧЕСКОЙ ПЛАСТИНОЙ

ЧУМАЧЕНКО С.В.

Предлагается точное аналитическое решение задачи об электромагнитном поле в решетке из плоских волноводов с диэлектрической пластиной. Для определения амплитудных коэффициентов применяется метод суммирования рядов по выборочным значениям.

1. Введение

Аналитические методы решения граничных задач занимают особое место в электродинамике. Полученные с их помощью результаты представляют интерес, поскольку являются основой для последующих численных расчетов. В связи с этим изучение и развитие методов решения граничных задач, допускающих построение строгого аналитического решения, остается актуальным во всех научных направлениях.

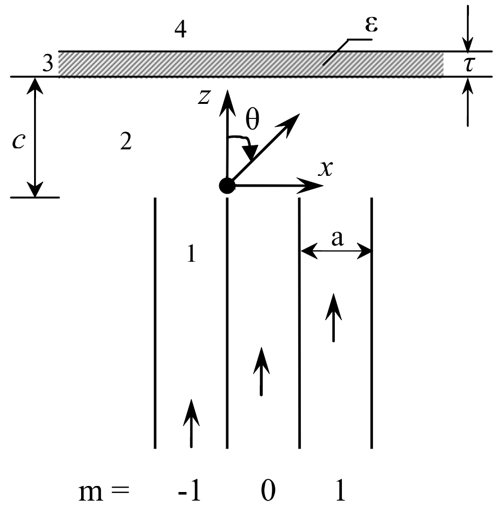
Большинство методов, используемых при расчете фазированных антенных решеток, основывается на классической теории [1,2], аналитических выводах [3-5], численных и экспериментальных результатах. Аналитические методы часто являются приближенными, а построенные на их основе численные – предполагают наличие временных затрат, связанных с вычислительной сложностью и повышением точности расчетов.

Таким образом, существует необходимость в разработке и развитии эффективных аналитических методов решения граничных задач, позволяющих строить точное решение, удобное для численного анализа.

Цель работы – получить аналитически выражения для амплитудных коэффициентов электромагнитного поля с использованием метода суммирования рядов по выборочным значениям.

2. Постановка и геометрия задачи

Рассматривается бесконечная периодическая система плоских волноводов, стенки которых считаются бесконечно тонкими, идеально проводящими, с расположенной над ней диэлектрической пластиной (рисунок) [1].



Все волноводы синфазно возбуждаются волной $TE_{1,0}$. Падающее поле в m -м волноводе имеет вид:

$$E_{y(пад)}(x,z) = \varphi_{пад}(x,z) =$$