Лекция 15

Непараметрические методы проверки гипотез

Большинство уже рассмотренных методов проверки гипотез относятся к классу *параметрических методов.* С их помощью проверяются, например, утверждения о величине показателя (параметра) генеральной совокупности - среднего арифметического, дисперсии, доли. Применимость параметрических методов обусловлена выполнением ряда условий относительно генеральной совокупности, наиболее "неприятное" из которых - приближенное соответствие нормальному распределению. А что делать, если это условие не выполняется? Для этого случая в статистике разработана специальная группа методов, называемых ***непараметрическими,*** которые применимы к любой генеральной совокупности. Непараметрические методы используются также для проверки ряда гипотез, не относящихся к параметрам распределения.

Преимущества и недостатки непараметрических методов

Несмотря на то, что непараметрические методы более универсальны, это не означает, что им следует отдавать предпочтение всегда, включая случай, когда условие приближенно нормальной распределенности генеральной совокупности выполняется. Обе группы методов имеют свои достоинства и недостатки, и к выбору между ними нужно относиться достаточно аккуратно.

Достоинства непараметрических методов.

1) Применимы в случае, когда исследуемая величина не имеет нормальной распределенности.

2) Применимы для анализа атрибутивных признаков.

3) Применимы для проверки утверждений, не относящихся к параметрам распределений.

4) Используют в большинстве случаев более простые по сравнению с параметрическим методами вычисления.

5) Более просты для понимания.

Недостатки непараметрических методов.

1) Обладают меньшей чувствительностью, чем параметрические методы для случая, когда условия применимости последних выполняются. Поэтому при проверке гипотез требуются большие отличия выборочного значения от тестируемого для того, чтобы отклонить основную гипотезу.

2) Как правило, используют меньший объем информации по сравнению с параметрическими аналогами. Например, тест знаков для каждого выборочного значения фиксирует только, больше ли оно тестируемого значения медианы или меньше, тогда как величина отклонения не учитывается.

3) Обладают меньшей эффективностью, чем параметрические аналоги в случае применимости последних. Это означает, что для достижения аналогичных результатов, аналогичной точности непараметрическому тесту потребуется больший размер выборки.

Соотношение перечисленных достоинств и недостатков позволяет дать следующую рекомендацию: если условия применимости параметрических методов выполняются, то их использование предпочтительней, если нет, то непараметрические методы дают достаточно мощное и эффективное средство для проведения исследований и в этом случае.

Ранги и ранжирование статистических данных

Многие непараметрические методы основаны на ранжировании данных, присвоении им порядкового номера - *ранга,* в соответствии с некоторым ранжирующим признаком. Например, менеджер по персоналу в ходе собеседования выставляет каждому из соискателей оценку по десятибалльной шкале:

Соискатель	А	Б	В	Г	Д
Оценка	8	6	10	1	3

Если, далее, упорядочить всех соискателей по убыванию набранного количества баллов и пронумеровать, то присвоенный номер и будет рангом:

Соискатель	В	А	Б	Д	Г
Оценка	10	8	6	3	1
Ранг	**1**	**2**	**3**	**4**	**5**

Так, в нашем примере соискатель В получил самый высокой балл (10), и он по рангу располагается первым, соискатель А (8 баллов) - вторым и т.д.

В случае, если у двух соискателей число баллов окажется одинаковым, то применяется процедура усреднения рангов. Предположим, что в предыдущем примере соискатели Б и Д набрали по 6 очков и получили порядковые номера три и четыре:

Соискатель	В	А	Б	Д	Г
Оценка	10	8	6	6	1
Балл	1	2	3	4	5

В окончательном виде их ранги должны быть одинаковыми и равными среднему арифметическому порядковых значений (3 + 4)/2 = 3.5:

Соискатель	В	А	Б	Д	Г
Оценка	10	8	6	3	1
Ранг	**1**	**2**	**3,5**	**3,5**	**5**

Этот же подход применим, если число одинаковых значений ранжирующего признака окажется больше двух.

Критерий знаков. Проверка гипотезы о медиане

Метод, основанный на критерии знаков, является простейшим непараметрическим методом и позволяет проверять гипотезы, относящихся к медиане генеральной совокупности. В рамках этого подхода исследователь выдвигает предположение о значении медианы генеральной совокупности, формирует из нее выборку и сопоставляет каждое из выбранных значений с предполагаемой медианой. Если значение признака в выборке больше медианы, то ему присваивается знак "+", если меньше, то "-" если точно равно, то 0. Далее количество знаков "+" $\left(n^+\right)$ и "-" $\left(n^-\right)$ подсчитывается и сравнивается. Если основная гипотеза верна и предполагаемое значение медианы близко к истинному, то n^+ и n^- будут примерно равно, если не верна, то n^+ и n^- окажутся сильно отличающимися. Расчетное значение критерия знаков равно меньшему из двух значений: $\min\left(n^+, n^-\right)$. Например, если получилось 8 знаков "+" и 3 "—", то берется 3. При заданном уровне значимости α расчетное значение критерия сопоставляется с критическим. Если расчетное меньше критического, то основная гипотеза отклоняется, если больше — то принимается. Таблица критических значений для критерия знаков, рассчитана с использованием биномиального распределения. Число n в таблице соответствует суммарному количеству знаков "+" и "—" (ноль не учитывается).

Пример. Директор магазина предполагает, что медианное значение для количества единиц бытовой техники, которое ему удается продать в день, равно 40. 20-дневное наблюдение дало следующие результаты:

18	43	40	16	22
30	29	32	37	36
39	34	39	45	28
36	40	34	39	52

Проверить при уровне значимости 0.05 предположение директора.

Решение. Схема проверки гипотез для различных непараметрических методов может несколько отличается от схемы для параметрических методов.

1) Сформулировать гипотезы:

$H_0 : Me = 40$; $H_1 : Me \neq 40$.

2) Сравнить каждое из значений в выборке с предполагаемой медианой и присвоить ему знак "+", "—" или "0".

-	+	0	-	-
-	-	-	-	-
-	-	-	+	-
-	0	-	-	+

3) Найти значение критерия. В таблице содержится 15 знаков "—" и 3 "+" меньшее из них 3, поэтому расчетное значение критерия равно 3.

4) Найти в таблице критическое значение, соответствующее двустороннему тесту с $\alpha = 0.05$ и числом n = 18 (общее число плюсов и минусов 15+3). На пересечении соответствующих столбца и строки находится значение 4.

5) Принятие решения. Значение критерия попадает в критическую область (3 < 4), основная гипотезу отклоняется.

6) Вывод. Данные наблюдений свидетельствуют, что медианное значение для количества продаж в день не равно 40.

Замечание. В случае, если число ненулевых знаков в таблице больше 26, то для проверки гипотезы можно использовать z - критерий. Расчетное значение вычисляется по формуле:

$$z = \frac{(X + 0.5) - n/2}{\sqrt{n/2}},$$

где X - значение критерия знаков, а критические значения по заданному α и длине выборки ищутся в таблице стандартного нормального распределения.

Критерий знаков. Сравнение средних зависимых выборок

Тест знаков применим также для сравнения генеральных средних *зависимых* выборок. Наиболее распространенным примером, в котором выборки зависимы, являются исследований типа "до" и " после". Например, врач замеряет величину пульса у группы пациентов, затем они принимают некоторый препарат, и через некоторое время пульс у них вновь измеряется. В другом примере преподаватель фиксирует количество баллов, полученное группой студентов на тестовой работе, затем они проходят курс подготовки и вновь пишут тест. В обоих случаях сравниваемые выборки состоят из одних и тех же лиц. Статистические методы, применяемые для сравнения зависимых выборок, несколько отличаются от методов для независимых выборок.

Пример. Главный врач команды пловцов предполагает, что число ушных инфекций в команде пловцов можно уменьшить, если использовать при нахождении в воде специальные ушные тампоны. Для проверки этого факта он отобрал тестовую группу из 10 человек и в течение 4 месяцев проводил наблюдение, из них первые два месяца тампоны не использовались, а последние два — использовались. Результаты наблюдений (количество инфекций) представлены в табл. 1, колонки 1-3. При уровне значимости 0.05 проверить предположение, что ушные тампоны позволяют уменьшить число инфекций.

Таблица 1. *Сравнение средних зависимых выборок*

Пловец	Первый период X_A	Второй период X_B	Знак разности
1	3	2	+
2	0	1	**-**
3	5	4	+
4	4	0	+
5	2	1	+
6	4	3	+
7	3	1	+
8	5	3	+
9	2	2	**0**
10	1	3	**-**

Решение:

1) Сформулировать гипотезы:

H_0 : Ушные тампоны не уменьшают количество инфекций;

H_1 :Ушные тампоны уменьшают количество инфекций.

2) Для каждой строки в таблице найти разность $X_A - X_B$ и занести ее знак в свободную колонку, табл. 1, колонка 4.

3) Найти значение критерия, для этого подсчитаем количество знаков "+ " - их 7 и знаков "—" их 2. Критерий равен 2.

4) Найти в таблице критическое значение для n = 9 (нулевая разность не учитывается), $\alpha = 0.05$ в одностороннем тесте. Критическое значение равно 1, то есть основная гипотеза может быть отклонена, только если в таблице было бы не более 1 знака "—".

5) Принятие решение. Расчетное значение 2 больше критического, основная гипотеза не отклоняется.

6) Вывод. Результаты наблюдений не подтверждают предположение, что тампоны дали положительный эффект. Этот несколько обескураживающий результат математически связан с очень маленькой выборкой и малой эффективностью непараметрических тестов.

Критерий суммы рангов Вилкоксона (Уилкоксона) сравнения средних независимых выборок

Недостатком критерия знаков при анализе средних является то, что он учитывает только направление отклонений, но не их величины. Например, для гипотезы о медиане влияние на критерий двух выборочных значений, одно из которых больше медианы на 100, а другое - на 1, будет одинаковым. Применение критерия суммы рангов Вилкоксона позволяет учесть амплитуду отклонений с помощью рангов. При этом метод относится к числу непараметрических и не накладывает никаких ограничений на исследуемую генеральную совокупность. Рассмотрим вариант метода Вилкоксона для независимых выборок.

При использовании критерия суммы рангов данные обеих выборок объединяются и затем ранжируются. Если средние значения двух генеральных совокупностей примерно равны, то выборочные значения перемещаются достаточно равномерно и суммы рангов, рассчитанные затем для каждой из выборок по отдельности, будут примерно равны. С другой стороны, если выборки взяты из отличающихся генеральных совокупностей, то и суммы рангов будут сильно различаться. Алгоритм выполнения теста будет представлен в форме примера.

Пример. Для сравнения качества физической подготовки отобраны две группы военнослужащих — из сухопутной части армии и части ВМФ и замерено время (в секундах), затрачиваемое ими на преодоление полосы препятствий. Получены следующие результаты:

Армия $(n_1 = 12)$	15	18	16	17	13	22	24	17	19	21	26	28	Среднее=19.67
ВМФ $(n_2 = 11)$	14	9	16	19	10	12	11	8	15	18	25		Среднее=14.27

Проверить при уровне значимости 0.05 предположение о том, что качество подготовки отличается.

Решение:

1) Сформулировать проверяемые утверждения:

H_0 : Средний уровень подготовки в армии и ВМФ одинаковый;

H_1 :Средний уровень различается.

2) Объединим значения из обеих выборок, расположим их по убыванию и присвоим каждому из них ранг (порядковый номер):

Время	8	9	10	11	12	13	14	15	15	16	16	17
Группа	М	М	М	М	М	А	М	А	М	А	М	А
Ранг	1	2	3	4	5	6	7	8.5	8.5	10.5	10.5	12.5

Время	17	18	18	19	19	21	22	24	25	26	28

Группа	А	М	А	А	М	А	А	А	М	А	А
Ранг	12.5	14.5	14.5	16.5	16.5	18	19	20	21	22	23

3) Рассчитаем сумму рангов для выборки, имеющей меньшую длину. В выборке для армии 12 значений, для ВМФ - 11, поэтому суммируем ранги для группы ВМФ:

$$R = 1 + 2 + 3 + 4 + 5 + 7 + 8.5 + 10.5 + 14.5 + 16.5 + 21 = 93$$

4) Далее рассчитываем значение критерия по следующим формулам:

$$m_R = \frac{n_2 \cdot (n_1 + n_2 + 1)}{2} = \frac{11 \cdot (11 + 12 + 1)}{2} = 132;$$

$$\sigma_R = \sqrt{\frac{n_1 \cdot n_2 \cdot (n_1 + n_2 + 1)}{12}} = \sqrt{\frac{11 \cdot 12 \cdot (11 + 12 + 1)}{12}} = \sqrt{264} \approx 16.25;$$

$$z = \frac{R - m_R}{\sigma_R} = \frac{93 - 132}{16.25} = -2.4 .$$

5) Находим с помощью таблицы стандартного нормального распределения критические значения для двустороннего теста при $\alpha = 0.05$, $z_{\alpha/2} = 1.96$, двусторонняя критическая область задается условием $z < -z_{\alpha/2} \cup z > z_{\alpha/2}$.

6) Принимаем решение. Расчетное значение критерия попадает в критическую область, -2.4 < -1.96, основная гипотеза отклоняется.

7) Вывод. Качество физической подготовки в выбранных частях армии и ВМФ различное.

Критерий серий для проверки случайности выборки

При анализе выборочных данных обычно предполагается, что выборка сформирована методом случайного отбора. В ряде случаев перед проведением исследования это необходимо дополнительно проверить. Рассмотрим несколько примеров, в которых для участия в опросе отбираются 20 студентов 1-го и 2-го курса, которых будем обозначать буквами П и В.

Пример А. П П П П П П П П П П В В В В В В В В В В .
Можно предположить, что выборка не является случайной, так как в нее сначала попали 10 первокурсников, а затем 10 второкурсников.

Пример Б. В П В П В П В П В П В П В П В П В П В П .
И в этом случае принцип случайности явно нарушен, так как в выборке первокурсники и второкурсники строго чередуются.

Пример В. В В В П П В П В П П В В П П В В П П П В .
Отбор участников, вероятно, является случайным.

Для того чтобы формализовать решение вопроса о случайности выборки, в статистике разработан непараметрический метод, основанный на *критерии серий.*

Определение. ***Серией*** называется последовательность одинаковых букв, за которой следует другая буква или не следует никаких букв.

Согласно определению, выборка примера А состоит из двух серий:

Серия 1 - П П П П П П П П П П ;

Серия 2 - В В В В В В В В В В .

Выборка пример Б - из 20 серий, так как каждая буква представляет отдельную серию. В примере В количество серий составляет 11:

Серия 1	В В В	Серия 5	В	Серия 9	В В
Серия 2	П П	Серия 6	П П	Серия 10	П П П
Серия 3	В	Серия 7	В В	Серия 11	В
Серия 4	П	Серия 8	П П		

Пример. 10 человек стоят в очереди в следующем порядке (М - мужчина, Ж - женщина) Ж, Ж ,Ж, М, М, Ж, Ж, Ж, Ж, М. Найти количество серий.

В очереди 4 серии: 1: ЖЖЖ; 2: ММ; 3: ЖЖЖЖ; 4: М.

Сформулируем далее две гипотезы:

H_0 : Выборка случайная;

H_1 : Выборка неслучайная

и выполним их проверку, используя в качестве критерия количество серий. Как можно предположить из рассуждений в примерах А, Б, В, в случайной выборке число серий не должно быть ни слишком большим, ни слишком малым. Для принятия строгого решения используется таблица критических значений для критерия серий, обычно эта таблица для $\alpha = 0.05$. Для нахождения нужных значений необходимо подсчитать количество букв первого (n_1) (щ) и второго (n_2) типов и взять пару критических значений, стоящих на пересечении строки n_1 и столбца n_2. Если количество серий попадает в интервал между ними, то основная гипотеза принимается - последовательность случайная, в противном случае принимается альтернативная гипотеза о неслучайности последовательности. В примерах А, Б, В число первокурсников и второкурсников одинаково: $n_1 = 10$, $n_2 = 10$. Пара критических значений при $\alpha = 0.05$ равна (6; 16), табл. 2. (см. таблицу ниже). Количество серий в примерах А и в Б не попадает в интервал (2 < 6; 20 >16), поэтому делается вывод, что выборки неслучайные. В примере В число серий равно $11 \in [6,16]$, поэтому выборка полагается случайной.

Таблица 2. *Критические значения для критерия серий*

$n_1 \backslash n_2$	2	3	... 9	10	11	...
2						
3						
...						
9						
10				(6;16)		
11						
...						

Пример. Кондуктор хочет определить, случайным ли образом садятся в автобус мужчины и женщины. Он пронаблюдал за 25 вошедшими пассажирами и записал их пол: Ж Ж Ж М М Ж Ж Ж Ж М Ж М М М Ж Ж Ж Ж М М Ж Ж Ж М М. Выполнить проверку случайности при уровне значимости 0.05.

Решение.

1) Сформулировать гипотезы.

H_0 :Пассажиры садятся в автобус в случайном порядке.

H_1 :Утверждение нулевой гипотезы неверно.

2) Найти количество серий:

Серия 1	Ж Ж Ж	Серия 6	М М М
Серия 2	М М	Серия 7	Ж Ж Ж Ж
Серия 3	Ж Ж Ж Ж	Серия 8	М М
Серия 4	М	Серия 9	Ж Ж Ж
Серия 5	Ж	Серия 10	М М

Всего среди вошедших оказалось 15 женщин (n_1) и 10 мужчин (n_2).

3) Находим пару критических значения по таблице (см. таблицу ниже) (7; 18).

4) Принимаем решение. Расчетное значение критерия 10 находится в некритической области $10 \in [7,18]$, основная гипотеза принимается.

5) Вывод. Данные наблюдений подтверждают, что пассажиры заходят случайным образом.

Критерий серий применим также для проверки случайности выборки с количественными значениями признака. В этом случае серией является последовательность значений, лежащих либо выше, либо ниже медианы.

Пример. На курс лечения от наркомании записалось 20 человек. Возраст участников в том порядке, в котором они записывались, составляет:

$$18,36,19,22,25,\ 44,23,27,27,35,\ 19,43,37,32,28,\ 43,46,19,20,22.$$

При уровне значимости 0.05 проверить гипотезу, что последовательность является случайной.

Решение.

1) Сформулировать гипотезы:

H_0 :Последовательность является случайной;

H_1 :Последовательность не является случайной.

2) Упорядочить данные по возрастанию и найти медиану:

18, 19, 19, 19, 20, 22, 22, 23, 25, 27, 27, 28, 32, 35, 36, 37, 43, 43, 44, 46.

Медианой распределения является значение 27.

3) Преобразовать последовательность значений в последовательность букв: Б если значение больше медианы, М — если меньше. Значения, в точности совпадающие с медианой, пропустить.

М, Б, М, М, М,   Б, М, Б, М, Б,   Б, Б, Б, Б, Б,   М, М, М.

Подсчитать количество букв каждого типа: "М" $n_1 = 9$, Б столько же $n_2 = 9$.

4) Найти число серий:

Серия 1	М	Серия 4	Б	Серия 7	М
Серия 2	Б	Серия 5	М	Серия 8	Б Б Б Б Б Б
Серия 3	М М М	Серия 6	Б	Серия 9	М М М

5) Найти критические значения в таблице (см. таблицу ниже): (5;15).

6) Принятие решения. Значение критерия (число серий) находится в некритической области $9 \in [5,15]$, принимается основная гипотеза.

7) Вывод. Последовательность является случайной.

Критические значения для критерия серий

$n_1 \backslash n_2$	2	3	4	5	6	7	8	9	10	11	12	13	14	15	16	17	18	19	20
2	1;6	1;6	1; 6	1;6	1;6	1;6	1;6	1;6	1;6	1;6	2;6	2;6	2;6	2;6	2;6	2;6	2;6	2;6	2;6
3	1;6	1;8	1;8	1;8	2;8	2; 8	2;8	2;8	2;8	2;8	2;8	2;8	2;8	3;8	3;8	3;8	3;8	3;8	3;8
4	1;6	1;8	1;9	2;9	2;9	2;10	3;10	3;10	3;10	3;10	3;10	3;10	3;10	3;10	4;10	4;10	4;10	4;10	4;10
5	1;6	1;8	2;9	2;10	3;10	3;11	3;11	3;12	3;12	4;12	4;12	4;12	4;12	4;12	4;12	4;12	5;12	5;12	5:12
6	1;6	2;8	2;9	3;10	3;11	3;12	3;12	4;13	4;13	4;13	4;13	5;14	5;14	5;14	5;14	5;14	5;14	6;14	6;14
7	1;6	2;8	2;10	3;11	3;13	4;13	4;14	4;14	5;14	5;14	5;14	5;15	5;15	6;15	6;16	6;16	6;16	6;16	6;16
8	1;6	2;8	3;10	3;11	3;12	4;13	4;14	5;14	5;15	5;15	6;16	6;16	6;16	6;16	6;17	7;17	7;17	7;17	7;17
9	1;6	2;8	3;10	3;12	4;13	4;14	5;14	5;15	5;16	6;16	6;16	6;17	7;17	7;18	7;18	7;18	8;18	8;18	8;18
10	1;6	2;8	3;10	3;12	4;13	5;14	5;15	5;16	6;16	6;17	7;17	7;18	7;18	7;18	8;19	8;19	8;19	8;20	9;20
11	1;6	2;8	3;10	4;12	4;13	5;14	5;15	6;16	6;17	7;17	7;18	7;19	8;19	8;19	8;20	9;20	9;20	9;21	9;21
12	2;6	2;8	3;10	4;12	4;13	5;14	6;16	6;16	7;17	7;18	7;19	8;19	8;20	8;20	9;21	9;21	9;21	10;22	10;22
13	2;6	2;8	3;10	4;12	5;14	5;15	6;16	6;17	7;18	7;19	8;19	8;20	9;20	9;21	9;21	10;22	10;22	10;23	10;23
14	2;6	2;8	3;10	4;12	5;14	5;15	6;16	7;17	7;18	8;19	8;20	9;20	9;21	9;22	10;22	10;23	10;23	11;23	11;24
15	2;6	3;8	3;10	4;12	5;14	6;15	6;16	7;18	7;18	8;19	8;20	9;21	9;22	10;22	10;23	11;23	11;24	11;24	12;25
16	2;6	3;8	4;10	4;12	5;14	6;16	6;17	7;18	8;19	8;20	9;21	9;21	10;22	10;23	11;23	11;24	11;25	12;25	12;25
17	2;6	3;8	4;10	4;12	5;14	6;16	7;17	7;18	8;19	9;20	9;21	10;22	10;23	11;23	11;24	11;25	12;25	12;26	13;26
18	2;6	3;8	4;10	5;12	5;14	6;16	7;17	8;18	8;19	9;20	9;21	10;22	10;23	11;24	11;25	12;25	12;26	13;26	13;27
19	2;6	3;8	4;10	5;12	6;14	6;16	7;17	8;18	8;20	9;21	10;22	10;23	11;23	11;24	12;25	12;26	13;26	13;27	13;27
20	2;6	3;8	4;10	5;12	6;14	6;16	7;17	8;18	9;20	9;21	10;22	10;23	11;24	12;25	12;25	13;26	13;27	13;27	13;28