

ОБЗОР РАСПРОСТРАНЕННЫХ ТЕСТОВ
ЗАВИСИМОСТЬ И НЕЗАВИСИМОСТЬ ВЫБОРОК
(НЕ-) И ПАРАМЕТРИЧЕСКИЕ И ТЕСТЫ
КОРРЕЛЯЦИИ

Зависимость и независимость выборок

- **Две выборки зависят друг от друга**, если каждому значению одной выборки можно закономерным и однозначным способом поставить в соответствие ровно одно значение другой выборки. Аналогично определяется зависимость нескольких выборок.
- **Чаще всего зависимые выборки возникают**, когда измерение проводится для нескольких моментов времени. Зависимые выборки образуют значения параметров изучаемого процесса, соответствующие различным моментам времени.
- **В SPSS зависимые (также связанные, спаренные)** выборки будут представляться разными переменными, которые сопоставляются друг с другом в соответствующем тесте на одной и той же совокупности наблюдений.
- **Если закономерное и однозначное соответствие между выборками невозможно**, эти выборки являются независимыми. В SPSS независимые выборки содержат разные наблюдения (например, относящиеся к различным респондентам), которые обычно различаются с помощью групповой переменной, относящейся к номинальной шкале.

Обзор распространенных тестов для проверки гипотез о среднем

Переменные, относящиеся к интервальной шкале и подчиняющиеся нормальному распределению

Количество сравниваемых выборок	Зависимость	Тест
1	Независимые	t-тест Стьюдента
1	Зависимые	t-тест для зависимых выборок
>2	Независимые	Простой дисперсионный анализ
>2	Зависимые	Простой дисперсионный анализ с повторными измерениями

Для каждой из этих двух групп тестов в SPSS имеются отдельные пункты меню, а именно Analyze (Анализ) Compare Means (Сравнение средних) или Analyze (Анализ) Nonparametric Tests (Непараметрические тесты)

Исключение составляет простой дисперсионный анализ с повторными измерениями. Этот метод нельзя найти в разделе Compare Means. Он вызывается командой меню General Linear Model (Общая линейная модель).

Обзор распространенных тестов для проверки гипотез о среднем

Переменные, относящиеся к порядковой шкале или переменные, относящиеся к интервальной шкале, но не подчиняющиеся нормальному распределению

Количество сравниваемых выборок	Зависимость	Тест
1	Независимые	U-тест Манна и Уитни
2	Зависимые	тест Уилкоксона
>2	Независимые	H-тест Крускала и Уоллиса
>2	Зависимые	тест Фридмана

Для каждой из этих двух групп тестов в SPSS имеются отдельные пункты меню, а именно Analyze (Анализ) Compare Means (Сравнение средних) или Analyze (Анализ) Nonparametric Tests (Непараметрические тесты)

Вероятность ошибки p (Probability)

Если следовать подразделению статистики на описательную и аналитическую, то задача аналитической статистики - предоставить методы, с помощью которых можно было бы объективно выяснить, например, является ли наблюдаемая разница в средних значениях или взаимосвязь (корреляция) выборок случайной или нет.

Например, если сравниваются два средних значения выборок, то можно сформулировать две предварительных гипотезы:

- Гипотеза H_0 (нулевая): Наблюдаемые различия между средними значениями выборок находятся в пределах случайных отклонений.
- Гипотеза H_1 (альтернативная): Наблюдаемые различия между средними значениями нельзя объяснить случайными отклонениями.

Вероятность ошибки p

В аналитической статистике разработаны методы вычисления так называемых тестовых (контрольных) величин, которые рассчитываются по определенным формулам на основе данных, содержащихся в выборках или полученных из них характеристик. Эти тестовые величины соответствуют определенным теоретическим распределениям (t-распределению, F-распределению, распределению χ^2 и т.д.), которые позволяют вычислить так называемую вероятность ошибки. Это вероятность равна проценту ошибки, которую можно допустить отвергнув нулевую гипотезу и приняв альтернативную.

Вероятность ошибки p

- Вероятность определяется в математике, как величина, находящаяся в диапазоне **от 0 до 1**. В практической статистике она также часто выражается в процентах. Обычно вероятность обозначается буквой p : **$0 < p < 1$**
- Вероятности ошибки, при которой допустимо отвергнуть нулевую гипотезу и принять альтернативную гипотезу, зависит от каждого конкретного случая. В значительной степени эта вероятность определяется характером исследуемой ситуации. *Чем больше требуемая вероятность, с которой надо избежать ошибочного решения, тем более узкими выбираются границы вероятности ошибки, при которой отвергается нулевая гипотеза, так называемый доверительный интервал вероятности.*
- Существует общепринятая терминология, которая относится к доверительным интервалам вероятности. Высказывания, имеющие вероятность ошибки **$p \leq 0,05$** - называются значимыми; высказывания с вероятностью ошибки **$p \leq 0,01$** - очень значимыми, а высказывания с вероятностью ошибки **$p \leq 0,001$** - максимально значимыми. В литературе такие ситуации обозначают одной, двумя или тремя звездочками.

Ошибки при статистическом выводе

- ❖ **Ошибка первого рода или альфа ошибка** (вероятность отвергнуть нулевую гипотезу, если на самом деле она справедлива) - ошибка потребителя.
- ❖ **Ошибкой второго рода или бета ошибка** (вероятность отвергнуть альтернативную гипотезу, если на самом деле она верна) - ошибка спонсора.
- ❖ Вероятность допустить ошибку первого рода равна вероятности ошибки α . Вероятность ошибки второго рода тем меньше, чем больше вероятность ошибки α .

Сравнение средних

- ❖ Сравнение средних значений различных выборок относится к наиболее часто применяемым методам статистического анализа. При этом всегда должен быть выяснен вопрос, можно ли объяснить имеющееся различие средних значений статистическими колебаниями или нет. В последнем случае говорят о *значимом различии*.
- ❖ При сравнении средних значений выборок предполагается, что обе выборки подчиняются *нормальному распределению*. Если это не так, то вычисляются медианы и для сравнения выборок используется *непараметрический тест*.

Сравнение средних

При сравнении средних значений выборок выделяют четыре различные тестовые ситуации:

- ❖ сравнение двух независимых выборок ;
- ❖ сравнение двух зависимых (спаренных) выборок ;
- ❖ сравнение более двух независимых выборок ;
- ❖ сравнение более двух зависимых выборок .

В этих ситуациях соответственно применяются следующие статистические тесты:

- ❖ t-тест для независимых выборок (тест Стьюдента);
- ❖ t-тест для зависимых выборок ;
- ❖ однофакторный дисперсионный анализ ;
- ❖ однофакторный дисперсионный анализ с повторными измерениями.

Непараметрические тесты

Непараметрические (не основанные на каком-либо распределении вероятности) тесты применяются там, где выборки из переменных, принадлежащих к интервальной шкале, не подчиняются нормальному распределению. Так как в этих тестах обрабатывается не само измеренное значение, а его ранг (положение внутри выборки), то эти тесты нечувствительны к выбросам. Непараметрические тесты применяются также в тех случаях, когда переменные относятся к порядковой, а не к интервальной шкале.

Непараметрические тесты

Вспомогательные меню	Аналоговое окно
Chi-Square (Хи-квадрат)	
Binomial (Биномиальный)	
Runs (Последовательности)	
1 -Sample K-S... (Колмогоров-Смирнов для одной выборки)	
2 Independent Samples (Две независимые выборки)	Mann-Whitney-U-Test (U-тест Манна-Уитни) Moses extreme reactions (Экстремальные реакции по Мозесу) Z Kolomgorov-Smirnov (Z-тест Колмогорова-Смирнова) Wald-Wolfowitz runs (Последовательности Уалда-Вольфовица)
K Independent Samples (K независимых выборок)	H Kruskal-Wallis (H-тест Крускала-Уоллиса) Median (Медианный тест)
2 Related Samples (Две связанные выборки)	Wilcoxon (Тест Уилкоксона) Sign (Знак) McNemar (Тест МакНемара)
K Related Samples (K связанных выборок)	Friedman (Тест Фридмана) W Kendall (W-тест Кендала) Q Cochran (Q-тест Кохрана)

Непараметрические тесты

- ❖ Наиболее часто применяемыми тестами являются тесты для сравнения двух и более независимых или зависимых выборок. Наиболее известными тестами, служащими для этих целей являются *U-тест Манна-Уитни*, *H-тест Крускала-Уоллиса*, *тест Уилкоксона* и тест *Фридмана*. Важную роль также играет тест *Колмогорова-Смирнова* для одной выборки, который может применяться для проверки наличия нормального распределения.
- ❖ *Непараметрические тесты могут*, конечно, применяться и в случае нормального распределения значений. Но в этом случае они будут иметь лишь 95 %-ую эффективность по сравнению с параметрическими тестами. Если Вы хотите, к примеру, произвести множественное сравнение средних значений двух независимых выборок, причем выборки являются частично подчиняются нормальному распределению, а частично — нет, то рекомендуется всегда применять *U-тест Манна и Уитни*.

U-тест по методу Манна и Уитни

Это самый известный и самый распространенный тест непараметрического сравнения двух независимых выборок. Он основан на использовании одной общей последовательности значений обоих выборок.

Тест Мозеса (Moses)

Данный тест проверяет различие размаха двух независимых выборок, которые состоят из переменных, относящихся к порядковой шкале, причем одна выборка рассматривается как контрольная группа, а другая как экспериментальная. Так как размах экстремальных значений может давать искаженные представления, то при помощи установки по умолчанию по обеим сторонам распределения контрольной группы отсекаются в общей сложности 5 процентов значений.

Тест Колмогорова-Смирнова

Условия применения данного теста такие же, как и при использовании U-теста по методу Манна и Уитни. Тест Колмогорова-Смирнова является предпочтительным тогда, когда количество категорий для тестируемых переменных ограничено. Если для такого j случая применять U-тест Манна и Уитни, то появляется большое количество ранговых мест, к которым относится сразу несколько переменных, то есть возникают неоднозначные ранговые последовательности. Основой теста является расчет максимальной разности между кумулятивными частотами обеих выборок. Эта разность обозначается величиной z , на основании которой, выводится вероятность ошибки p .

Тест Уалда-Вольфовица (Wald-Wolfowitz)

Условия применения данного теста те же, что и при U-тесте по методу Манна и Уитни или при тесте Колмогорова-Смирнова. Значения обеих групп выстраиваются в единую последовательность по рангу. Затем производится подсчёт количества смен группового признака, с помощью которого можно найти количество непрерывных последовательностей. Если появляются одинаковые значения (ранговые связки), то выводятся значения минимального и максимального числа возможных непрерывных последовательностей. Исходя из количества непрерывных последовательностей, можно найти вероятность ошибки p . Данный тест не пригоден для переменных с малым числом категорий, так как в этом случае очень сильно возрастает количество ранговых связок.

Тест Уилкоксона (Wilcoxon)

Этот тест является традиционным непараметрическим тестом для сравнения двух зависимых выборок. Он основан на построении ранговой последовательности абсолютных разностей пар значений.

Тест хи-квадрат по методу МакНемара (McNemar)

Данный тест применяется исключительно при наличии дихотомических переменных. При этом для двух зависимых переменных выясняется, происходят ли какие-либо изменения в структуре распределения их значений. В большинстве наблюдений сравнение проводится с учётом временного фактора по схеме "до — после".

Тест Колмогорова-Смирнова для проверки формы распределения

При помощи этого теста по выбору можно проверить, соответствует ли реальное распределение переменной нормальному, равномерному, экспоненциальному распределению или распределению Пуассона. Разумеется, самым распространённым видом проверки является проверка наличия нормального распределения.

Отклонение от нормального распределения считается существенным при значении $p < 0,05$; в этом случае для соответствующих переменных следует применять непараметрические тесты.

Корреляция (correlation)

Степень совместной изменчивости переменных.

Коэффициент корреляции (correlation coefficient)

Показатель ассоциации, который отражает степень линейной зависимости между двумя переменными. Этот коэффициент обозначается буквой r , изменяется от **+1** до **-1**; если **$r = +1$** , это означает идеальную положительную взаимосвязь двух переменных, если **$r = -1$** , это означает идеальную отрицательную взаимосвязь двух переменных. Если значение находится ближе к **1**, то это означает наличие сильной связи, а если ближе к **0**, то слабой.

Коэффициент корреляции (correlation coefficient)

Сила связи характеризуется также и абсолютной величиной коэффициента корреляции. Для словесного описания величины коэффициента корреляции используются следующие градации:

- **до 0,2** – очень слабая корреляция;
- **до 0,5** – слабая корреляция;
- **до 0,7** – средняя корреляция;
- **до 0,9** – высокая корреляция;
- **свыше 0,9** – очень высокая корреляция.

Коэффициент корреляции (correlation coefficient)

Метод вычисления коэффициента корреляции зависит от вида шкалы, которой относятся переменные.

- Переменные с интервальной и с номинальной шкалой: коэффициент корреляции Пирсона (корреляция моментов произведений).
- По меньшей мере, одна из двух переменных имеет порядковую шкалу либо не является нормально распределённой: ранговая корреляция по Спирману или *T*-Кендала.
- Одна из двух переменных является дихотомической: точечная двухрядная корреляция. Эта возможность в SPSS отсутствует. Вместо этого может быть применён расчёт ранговой корреляции.
- Обе переменные являются дихотомическими: четырёхполевая корреляция. Данный вид корреляции рассчитываются в SPSS на основании определения мер расстояния и мер сходства .