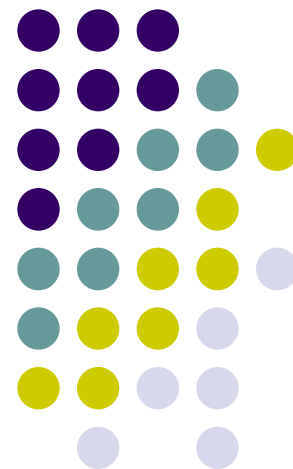
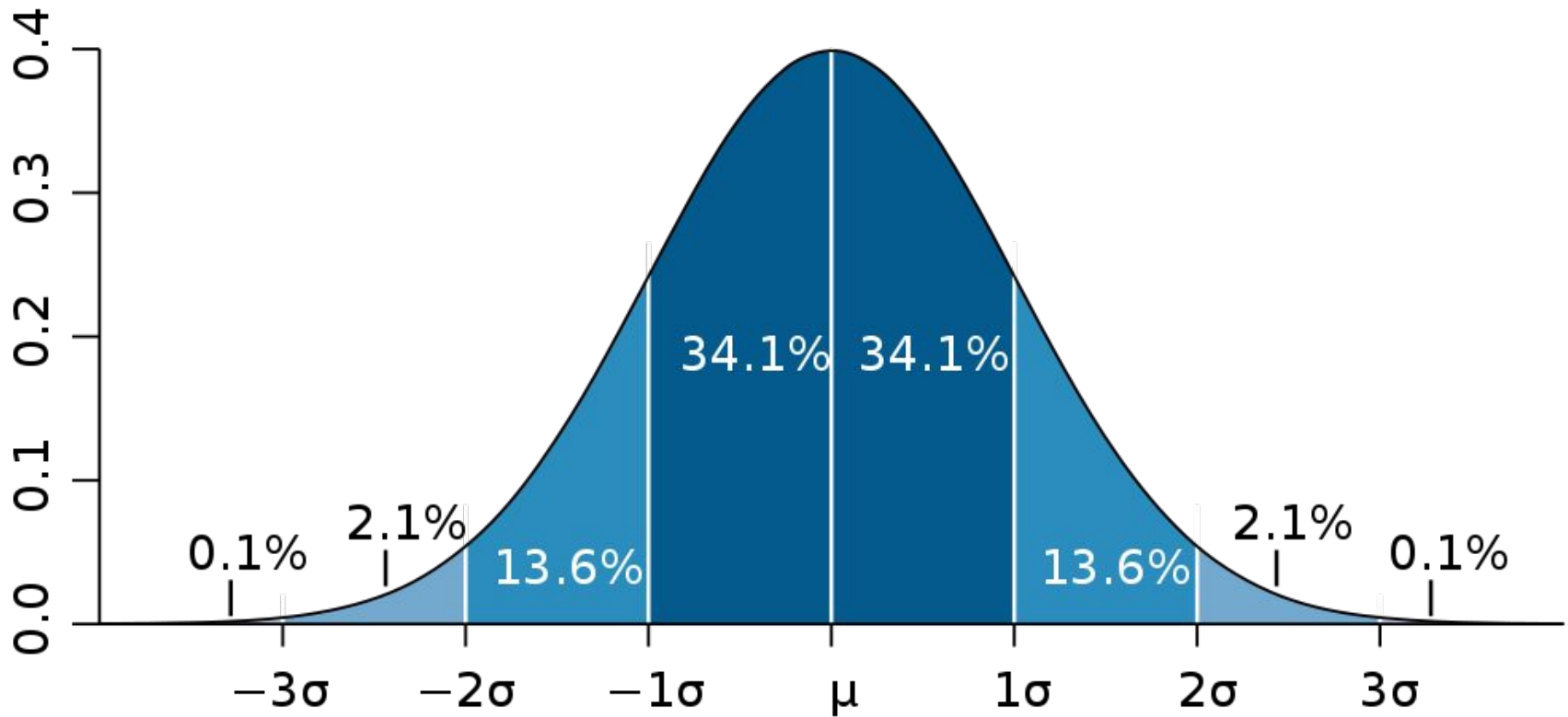


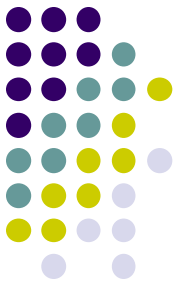
Выбор метода статистического вывода



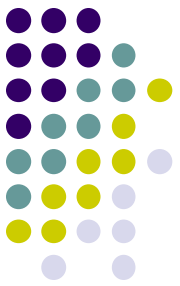
Нормальное распределение как стандарт



Статистическая гипотеза



Это утверждение относительно неизвестного параметра генеральной совокупности, которое формулируется для проверки надежности связи и которое можно проверить по известным выборочным статистикам – результатам исследования.



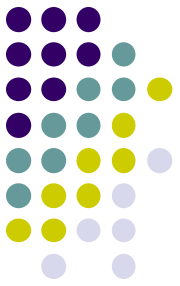
Статистическая гипотеза

Основная (нулевая) гипотеза (H_0) – содержит утверждение об отсутствии связи в генеральной совокупности и доступна проверке методами статистического вывода.

Альтернативная гипотеза (H_1) – принимается при отклонении H_0 и содержит утверждение о наличии связи.

При этом нулевая и альтернативная гипотеза представляют собой полную группу несовместных событий.

Измерительные шкалы (неметрические):

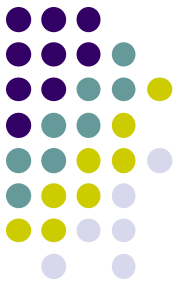


Номинативная шкала, или шкала наименований.

Объекты группируются по различным классам так, чтобы внутри класса они были идентичны по измеряемому свойству.

Ранговая, или порядковая шкала. Измерение в этой шкале предполагает приписывание объектам чисел в зависимости от степени выраженности измеряемого свойства.

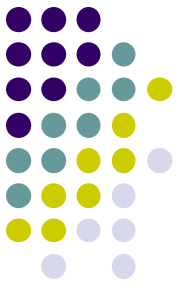
Измерительные шкалы (метрические):



Интервальная шкала. Это такое измерение, при котором числа отражают не только различия между объектами в уровне выраженности свойства, но и то, насколько больше или меньше выражено это свойство.

Абсолютная шкала, или шкала отношений. Измерение в этой шкале отличается от интервального тем, что в ней устанавливается нулевая точка, соответствующая полному отсутствию выраженности измеряемого свойства.

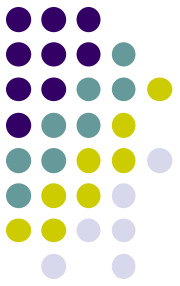
Классификация методов статистического вывода



Основания для классификации:

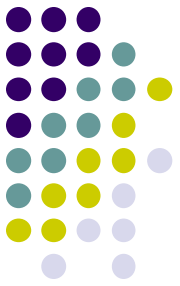
- типы шкал, в которых измерены признаки X и Y :
качественная шкала (номинативная), количественная шкала (порядковая, метрическая)
- количество сравниваемых групп – две и более двух
- соотношение сравниваемых групп: зависимые выборки или независимые выборки

Классификация методов статистического вывода



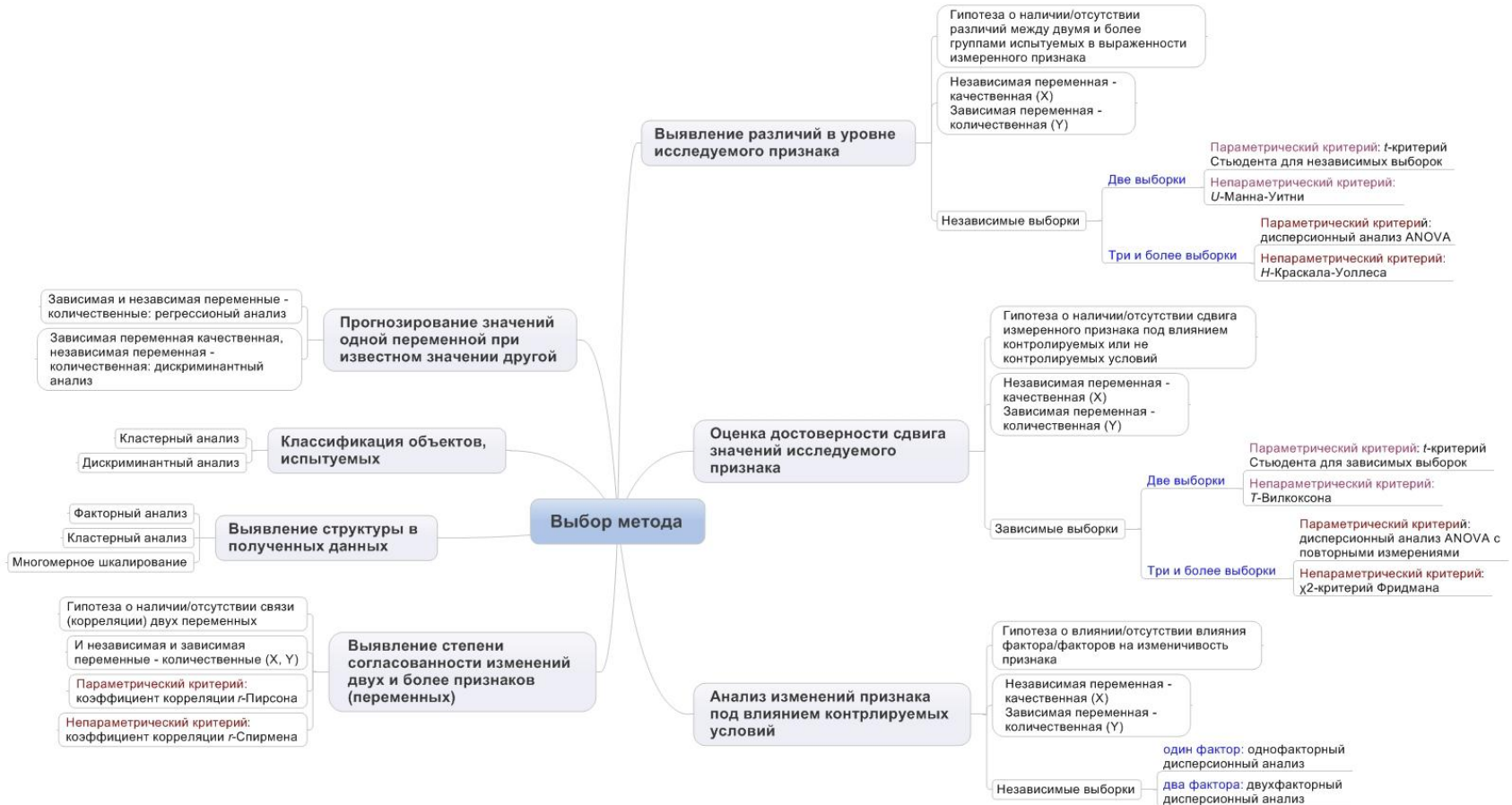
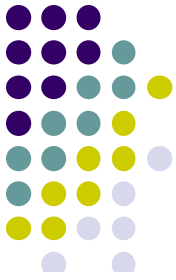
Типы шкал	I. X, Y – количественные	II. X, Y – качественные	III. X – качественный, Y – количественный
Задачи:	Корреляционный анализ	Анализ номинативных данных: классификаций, таблиц сопряженности, последовательностей (серий)	Сравнения выборок по уровню выраженности признака
Методы:	а) r-Пирсона – для метрических X и Y; б) частная корреляция и сравнение корреляций; в) r-Спирмена, т-Кендалла – для ранговых X и Y.	Критерий χ^2 -Пирсона (для классификаций и таблиц сопряженности), критерий Мак-Намара (для таблиц 2x2 с повторными измерениями), критерий серий (для последовательностей)	(методы сравнения) – следующий слайд

Классификация методов статистического вывода

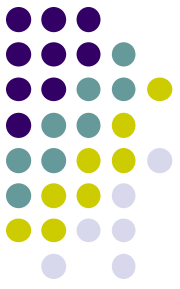


Количество выборок (градаций X)		Две выборки		Больше двух выборок	
Зависимость выборок		Независимые	Зависимые	Независимые	Зависимые
Признак Y	Метрический	Параметрические методы сравнения			
		t-Стьюдента для независимых выборок	t-Стьюдента для зависимых выборок	ANOVA	ANOVA с повторными измерениями
	Ранговый	Непараметрические методы сравнения			
		U-Манна-Уитни, критерий серий	T-Вилкоксона, критерий знаков	H-Краскала-Уоллеса	χ^2 -Фридмана

Выбор методов статистического вывода

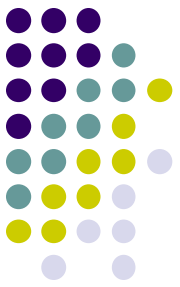


Параметрические и непараметрические критерии



- Критерий различия называют **параметрическим**, если он основан на конкретном типе распределения генеральной совокупности (как правило, нормальном) или использует параметры этой совокупности (средние, дисперсии и т.д.).
- Критерий различия называют **непараметрическим**, если он не базируется на предположении о типе распределения генеральной совокупности и не использует параметры этой совокупности.

Методы корреляционного анализа

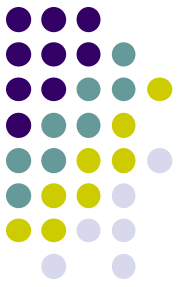


Проверяемая H_0 : коэффициент корреляции равен нулю.

Условие применения: а) два признака измерены в ранговой или метрической шкале на одной и той же выборке; б) связь между признаками является монотонной (не меняет направления по мере увеличения значений одного из признаков).

Обычно изучается корреляция между множеством P переменных. В таком случае вычисляются корреляции между всеми возможными парами этих переменных. Результатом является корреляционная матрица, включающая $P(P-1)/2$ значений коэффициентов парной корреляции. Под корреляционным анализом обычно и понимают изучение связей по корреляционной матрице.

Методы корреляционного анализа



Методы:

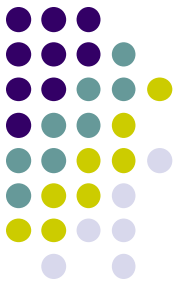
Корреляция r -Пирсона – для метрических переменных.

Условие применения: а) распределения X и Y существенно не отличаются от нормального.

Дополнительно: частная корреляция для изучения зависимости корреляции X и Y от влияния переменной Z ; сравнение корреляций – для независимых и зависимых выборок.

Корреляции r -Спирмена, τ -Кендалла – для порядковых переменных.

Методы анализа номинативных переменных



В зависимости от цели исследования и структуры исходных данных выделяются три группы методов, соответствующих решаемым задачам:

- анализ классификаций;
- анализ таблиц сопряженности;
- анализ последовательностей (серий).

Методы анализа НОМИНАТИВНЫХ переменных



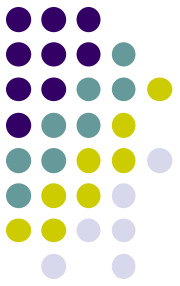
Анализ классификаций.

Условие применения: для каждого объекта (испытуемого) выборки определена его принадлежность к одной из категорий (градаций) X (получено эмпирическое распределение объектов по X); известно теоретическое (ожидаемое) распределение по X (обычно – равномерное).

Проверяемая H_0 : эмпирическое (наблюдаемое) распределение предпочтений не отличается от теоретического (ожидаемого).

Метод: критерий χ^2 -Пирсона.

Методы анализа номинативных переменных



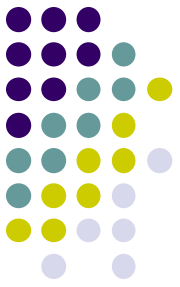
Анализ таблиц сопряженности.

Условие применения: для каждого объекта (испытуемого) выборки определена его принадлежность к одной из категорий (градаций) X и к одной из категорий (градаций) Y (получена перекрестная классификация объектов по двум основаниям X и Y).

Следует различать три ситуации – в зависимости от числа градаций и соотношения X и Y :

- число градаций X и (или) Y больше двух (общий случай);
- таблицы сопряженности 2×2 с независимыми выборками;
- таблицы сопряженности 2×2 с повторными измерениями.

Методы анализа номинативных переменных



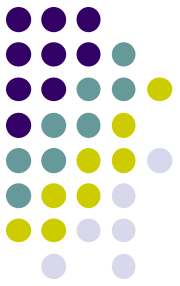
Анализ последовательностей (серий)

Условие применения: объекты упорядочены (по времени или по уровню выраженности признака); каждый объект отнесен к одной из двух категорий (X или Y).

Проверяемые H_0 : события X распределены среди событий Y случайно (случай 1); выборки X и Y не различаются по распределению значений количественного признака (случай 2).

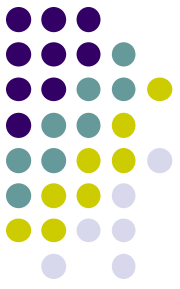
Метод: критерий серий.

Методы сравнения выборок по уровню выраженности признака



- В зависимости от решаемых задач методы внутри этой группы классифицируются по трем основаниям:
 - ▶ Количество градаций X :
 - а) сравниваются 2 выборки;
 - б) сравниваются больше двух выборок
 - ▶ Зависимость выборок:
 - а) сравниваемые выборки независимы;
 - б) сравниваемые выборки зависимы.
 - ▶ Шкала Y :
 - а) Y – ранговая переменная;
 - б) Y – метрическая переменная.

Сравнение двух независимых выборок



Условия применения: признак измерен у объектов (испытуемых), каждый из которых принадлежит к одной из двух независимых выборок.

Методы:

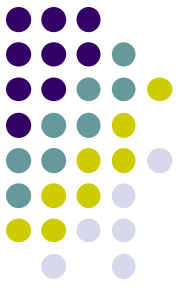
Y – метрическая переменная: сравнений двух средних значений (параметрический критерий t -Стьюдента для независимых выборок).

Условия применения: признак измерен в а) метрической шкале; б) дисперсии двух выборок гомогенны (статистически достоверно не различаются). Если не выполняется хотя бы одно из этих условий то применяется непараметрический критерий U -Манна-Уитни.

Дополнительно: возможно сравнений двух дисперсий (параметрический критерий F -Фишера).

Y – ранговая (порядковая) переменная: сравнение двух независимых выборок по уровню выраженности порядковой и бинарной переменной (критерий U -Манна-Уитни, критерий серий).

Сравнение двух зависимых выборок



- *Условия применения:* а) признак измерен у объектов (испытуемых), каждый из которых принадлежит к одной из двух зависимых выборок: либо признак измерен дважды на одной и той же выборке, либо каждому испытуемому из одной выборки поставлен в соответствие по определенному критерию испытуемый из другой выборки; б) измерения положительно коррелируют. Если эти условия не выполняются, то выборки следуют признать независимыми.

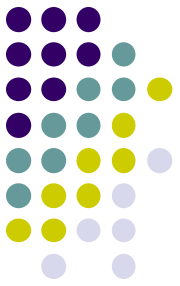
Методы:

Y – метрическая переменная: сравнений двух средних значений (параметрический критерий t -Стьюдента для зависимых выборок).

Условия применения: признак измерен в метрической шкале. Если не выполняется хотя бы одно из этих условий то применяется непараметрический критерий T- Вилкоксона.

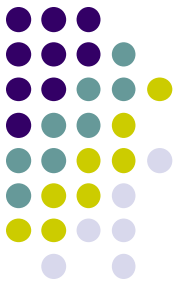
Y – ранговая (порядковая) переменная: сравнение двух зависимых выборок по уровню выраженности порядковой и бинарной переменной (критерий T- Вилкоксона, критерий знаков).

Сравнение более двух выборок



Проверяемая H_0 : несколько совокупностей (которым соответствуют выборки) не отличаются по уровню выраженности измеренного признака.

Сравнение более двух независимых выборок



Условия применения: признак должен быть измерен у объектов (испытуемых), каждый из которых принадлежит к одной из k независимых выборок ($k > 2$).

Методы:

Y – метрическая переменная: дисперсионный анализ (ANOVA) для независимых выборок (параметрический метод).

Дополнение: метод допускает сравнение выборок более чем по одному основанию – когда деление на выборки производится по нескольким номинативным переменным, каждая из которых имеет 2 и более градаций.

Условия применения: признак Y измерен в а) метрической шкале, б) дисперсии выборок гомогенны (статистически достоверно не различаются). Если не выполняется хотя бы одно из этих условий, то:

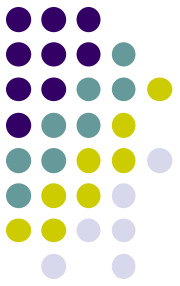
Сравнение более двух независимых выборок



Y- ранговая (порядковая) переменная:
сравнение более двух независимых выборок по уровню выраженности ранговой переменной (непараметрический критерий Н-Краскала-Уоллеса).

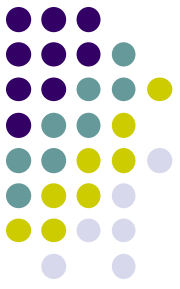
Ограничение: методы позволяет сравнивать выборки только по одному основанию, когда деление на группы производится по одной номинативной переменной, имеющей более 2-х градаций.

Сравнение более двух зависимых выборок



Условия применения: а) признак измерен у объектов (испытуемых), каждый из которых принадлежит к одной из k зависимых выборок ($k > 2$): как правило, признак измерен несколько раз на одной и той же выборке; б) измерения положительно коррелируют.

Сравнение более двух зависимых выборок



Методы:

Y- метрическая переменная: дисперсионный анализ (ANOVA) с повторными измерениями (параметрический метод).

Дополнение: метод допускает сравнение выборок более чем по одному основанию – когда помимо деления на зависимые выборки, вводятся номинативные переменные, которые имеют 2 и более градаций и делят испытуемых на независимые выборки.

Условия применения: а) признак Y измерен в метрической шкале; б) дисперсии сравниваемых выборок гомогенны (статистически достоверно не различаются). Если не выполняется хотя бы одно из этих условий, то:

Сравнение более двух зависимых выборок



Y- ранговая (порядковая) переменная:
сравнение более двух зависимых выборок по уровню выраженности ранговой переменной (непараметрический критерий χ^2 -Фридмана).

Ограничение: метод позволяет сравнивать зависимые выборки только по одному основанию – повторным измерениям.



Спасибо за внимание!