

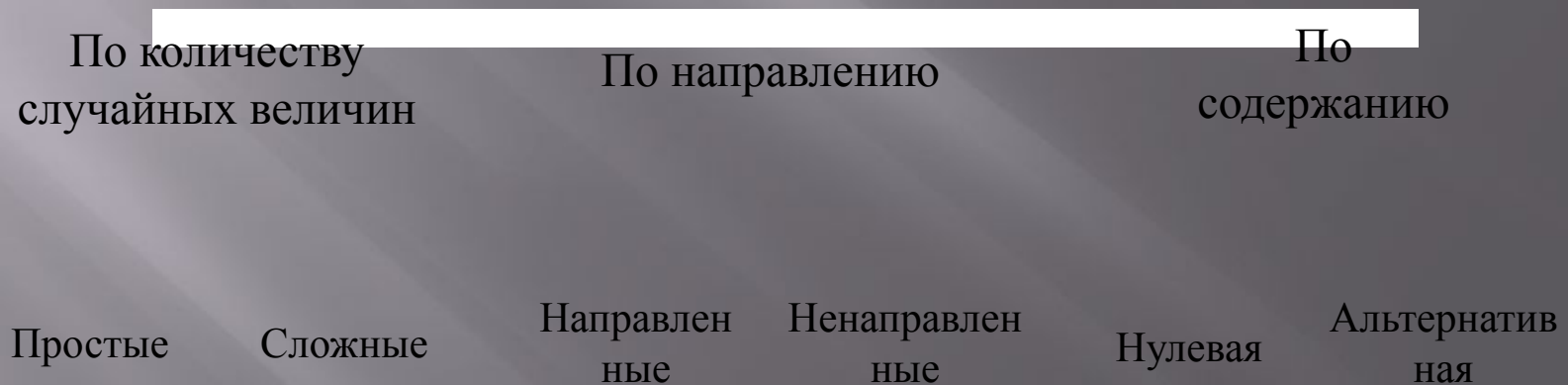
# КОМПЬЮТЕРНАЯ ОБРАБОТКА ДАННЫХ В ПСИХОЛОГИИ

Лекция 2

Проверка статистических гипотез

Статистическими гипотезами называются различного рода предположения о распределениях случайных величин, которые необходимо проверить по данным выборочной совокупности.

## Статистические гипотезы



*Рис. 1. Виды статистических гипотез*

Простая гипотеза полностью задает одно распределение вероятностей. Например, между юношами и девушками существуют различия в значимости ценности «счастливая семейная жизнь».

Сложная гипотеза указывает не одно распределение, а некоторое множество распределений. Например, между юношами и девушками существуют различия в системе Ненаправленные гипотезы утверждают, что показатели различны, ценностных ориентации личности, но при этом не говорится о том, какой показатель больше другого.

Направленные гипотезы не только говорят о наличии или отсутствии различий, но и указывают направление этих различий.

В психологии, как правило, под нулевой гипотезой понимается гипотеза об отсутствии значимых различий между какими-либо показателями или о случайности связи между ними. Нулевая гипотеза обозначается как  $H_0$ . Альтернативная гипотеза – это гипотеза о наличии значимых различий между какими-либо показателями или о достоверности связи между ними. Она обозначается как  $H_1$ .

При проверке статистической гипотезы начинают с предположения о том, что нулевая гипотеза верна.

*Критерием проверки гипотезы* называется некоторая функция выборки ( $X_1, X_2, \dots, X_n$ ), характеризующая степень соответствия выборочных данных статистической гипотезе.

Область значений критерия ( $R$ ) разбивают на два подмножества: *область принятия гипотезы* [ $R', R''$ ] и *критическую область*. При попадании значения критерия  $R$ , подсчитанного по выборке, в область принятия, гипотеза принимается, в противном случае – отклоняется, принимается альтернативная.

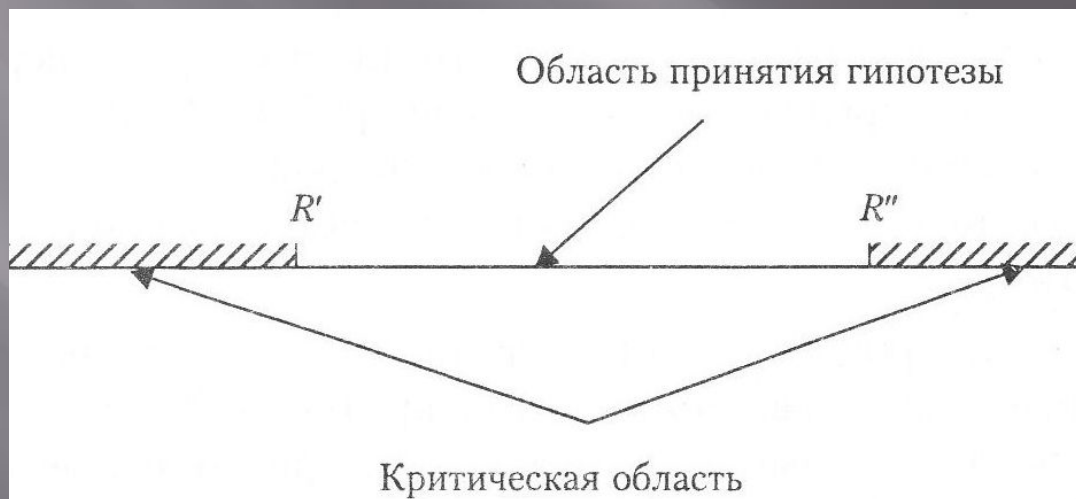


Рис. 2. Область принятия гипотезы и критическая область

При проверке статистических гипотез возможны ошибочные заключения двух типов.

*Ошибкой первого рода* называется отклонение верной гипотезы в случае, когда она на самом деле верна.

*Ошибка второго рода* заключается в принятии неверной гипотезы, если она на самом деле неверна.

*Уровень значимости* результата исследования – это количественно выраженная вероятность, что полученные результаты достоверны.

Существует обратная зависимость между уровнем значимости и надежностью результата исследования. Чем выше  $p$ -уровень, тем ниже надежность и ниже уровень доверия к полученным результатам. Соотношение показателей  $p$ -уровня и степени значимости в таблице 1 (по А.Д. Наследову).

*Таблица 1*

*Соотношение значимости и  $p$ -уровня*

$p > 0,1$	Результаты статистически незначимы
$0,05 < p \leq 0,1$	Результаты значимы на уровне тенденции
$p \leq 0,05$	Результаты статистически значимы
$p \leq 0,01$	Результаты на уровне высокой статистической значимости

Для проверки статистической гипотезы необходимо выполнить следующие этапы.

1. Определить нулевую и альтернативную гипотезы.
2. Задать уровень значимости  $\alpha$  (допустимую вероятность ошибки первого рода).
3. Выбрать критерий проверки.
4. Определить критическую область.
5. По результатам эмпирического или экспериментального исследования вычислить фактическое значение критерия.
6. Если наблюдаемое значение критерия принадлежит критической области, то отклонить гипотезу, иначе – принять.

# *Классификация статистических критериев в зависимости от исследовательских задач*

## **Статистические критерии**

Критерии сравнения распределений

Критерии оценки достоверности различий

Критерии оценки достоверности сдвига

Критерии определения согласованности изменений

Анализ изменений переменной под влиянием контролируемых условий

## *Возможности и ограничения параметрических и непараметрических критериев*

Параметрические критерии	Непараметрические критерии
1. Позволяют прямо оценить различия в средних, полученных в двух выборках (t - критерий Стьюдента).	Позволяют оценить лишь средние тенденции, например, ответить на вопрос, чаще ли в выборке А встречаются более высокие, а в выборке Б - более низкие значения признака (критерии Манна-Уитни и др.).
2. Позволяют прямо оценить различия в дисперсиях (критерий Фишера).	Позволяют оценить лишь различия в диапазонах вариативности признака (критерий Фишера с угловым преобразованием )
3. Позволяют выявить тенденции изменения признака при переходе от условия к условию (дисперсионный однофакторный анализ), но лишь при условии нормального распределения признака.	Позволяют выявить тенденции изменения признака при переходе от условия к условию при любом распределении признака (критерии тенденций Пейджа L и Джонкира S).
4. Позволяют оценить взаимодействие, двух и более факторов в их влиянии на изменения признака (двухфакторный дисперсионный анализ).	Эта возможность отсутствует.
5. Экспериментальные данные должны отвечать двум, а иногда трем, условиям: а) значения признака измерены по интервальной шкале; б) распределение признака является нормальным; в) в дисперсионном анализе должно соблюдаться требование равенства дисперсий в ячейках комплекса	Экспериментальные данные могут не отвечать ни одному из этих условий: а) значения признака могут быть представлены в любой шкале, начиная от шкалы наименований; б) распределение признака может быть любым и совпадение его с каким-либо теоретическим законом распределения необязательно и не нуждается в проверке; в) требование равенства дисперсий отсутствует.
6. Математические расчеты довольно сложны.	Математические расчеты по большей части просты и занимают мало времени (за исключением критериев $\chi^2$ и $\lambda$ )
7. Если условия, перечисленные в п.5 выполняются, параметрические критерии оказываются несколько более мощными, чем непараметрические.	Если условия, перечисленные в п.5, выполняются, непараметрические критерии оказываются более мощными, чем параметрические, так как они менее чувствительны к «засорениям».



## *Классификация статистических критериев*

<b>Исследовательские задачи</b>	<b>Параметрические критерии</b>	<b>Непараметрические критерии</b>
<b>Критерии сравнения распределений</b>		$\chi^2$ – Пирсона, $\lambda$ – критерий Колмогорова-Смирнова и т.д.
<b>Критерии оценки достоверности различий</b>	<b>t-критерий Стьюдента для независимых выборок, однофакторный дисперсионный анализ для независимых выборок</b>	<b>Q- критерий Розенбаума, U-критерий Манна-Уитни, H-критерий Краскела-Уоллеса, Биномиальный критерий и др.</b>
<b>Критерии оценки достоверности сдвига</b>	<b>t-критерий Стьюдента для зависимых выборок, однофакторный дисперсионный анализ для зависимых выборок</b>	<b>T-критерий Вилкоксона, G – критерий знаков, критерий Фридмана и др.</b>
<b>Критерии определения согласованности изменений</b>	<b>Коэффициент корреляции Пирсона</b>	<b>корреляция Спирмена, коэффициент сопряженности С-Пирсона, коэффициент фи-корреляции Пирсона.</b>
<b>Анализ изменений переменной под влиянием контролируемых условий</b>	<b>Дисперсионный анализ (F- критерий Фишера)</b>	

# ОЦЕНКА ДОСТОВЕРНОСТИ РАЗЛИЧИЙ

Решение о выборе того или иного критерия принимается на основе выполнения ряда требований

- В зависимости от того, в какой шкале измеряются переменные и какое они имеют распределение, выбирают параметрические или непараметрические критерии.
- Необходимо определить, сколько выборок сопоставляется.

## • Критерии оценки достоверности различий

### • *Параметрические критерии*

- Если 2 выборки,  
• то критерий Стьюдента для независимых выборок
  
- Если 3 выборки и более, то однофакторный дисперсионный анализ для независимых выборок

### • *Непараметрические критерии*

- Если 2 выборки,  
• то критерий Манна-Уитни
  
- Если 3 и более выборок,  
• то критерий Краскела-Уоллеса

## Критерий Стьюдента для независимых выборок

*Область применения:* критерий предназначен для оценки достоверности различий между двумя независимыми группами на основе сравнения средних значений переменной.

*Требования:*

- одна независимая (группирующая) переменная, измеренная в номинальной шкале и одна зависимая переменная, измеренная в метрических шкалах;
- распределение зависимой переменной не должно значительно отличаться от нормального распределения.

*Алгоритм расчета критерия.*

- Сформулировать нулевую ( $H_0$ ) и альтернативную гипотезы ( $H_1$ ).
- Проверить нормальность распределения зависимой переменной.
- Если  $t_{\text{эмп}} \geq t_{\text{кр}}$  при  $p \leq 0,05$ , то отвергается  $H_0$  и принимается  $H_1$ . Если  $t_{\text{эмп}} < t_{\text{кр}}$  при  $p \leq 0,1$ , то принимается  $H_0$  и отвергается  $H_1$ .

# Однофакторный дисперсионный анализ для независимых выборок

**Область применения:** критерий предназначен для оценки достоверности различий между тремя и более независимыми группами на основе сопоставления компонентов дисперсии изучаемой переменной.

## **Требования:**

- одна независимая переменная (фактор), измеренная в номинальной шкале и одна зависимая переменная, измеренная в метрических шкалах;
- не менее трех уровней фактора и не менее двух испытуемых в каждой группе;
- распределение зависимой переменной не должно значительно отличаться от нормального распределения;
- гомогенность (равенство) дисперсий в сравниваемых группах, которое достигается за счет выравнивания количества испытуемых в каждом уровне фактора.

## **Алгоритм расчета критерия.**

1. Сформулировать нулевую ( $H_0$ ) и альтернативную ( $H_1$ ) гипотезы.
2. Проверить нормальность распределения зависимой переменной.
3. Сопоставить эмпирическое и критическое значения : при  $F_{\text{эмп}} \geq F_{\text{кр}}$  отклоняется  $H_0$  и принимается  $H_1$ , при  $F_{\text{эмп}} < F_{\text{кр}}$  принимается  $H_0$  и отклоняется  $H_1$

## Критерий Манна-Уитни.

**Область применения:** критерий предназначен для оценки достоверности различий между двумя выборками, если данные представлены в порядковой шкале или в метрических шкалах, но не укладываются в кривую нормального распределения.

### **Требования:**

- одна независимая (группирующая) переменная, измеренная в номинальной шкале и одна зависимая переменная, измеренная в порядковой шкале или метрической шкале, но распределение значимо отличается от нормального;
- в каждой выборке должно быть не менее трех и не более 60 наблюдений.

### **Алгоритм расчета критерия.**

- Сформулировать нулевую ( $H_0$ ) и альтернативную ( $H_1$ ) гипотезы.
- Сопоставить эмпирическое и критическое значения: если  $U_{\text{эмп}} > U_{\text{кр}}$  при  $p = 0,05$ , то принимается  $H_0$  и отклоняется  $H_1$ , если же  $U_{\text{эмп}} \leq U_{\text{кр}}$  при  $p = 0,05$ , то отклоняется  $H_0$  и принимается  $H_1$ .

**Замечание.** Критерий Манна-Уитни отличается от большинства других критериев тем, что для опровержения нулевой гипотезы эмпирическое значение должно быть меньше или равно критическому. При этом, чем меньше значение  $U$ , тем достоверность различий выше.

## Критерий Краскела – Уоллеса

**Область применения:** критерий предназначен для оценки достоверности различий между тремя и более выборками по какому-либо показателю. Иногда  $H$  - критерий Краскела – Уоллеса рассматривается как непараметрический аналог метода однофакторного дисперсионного анализа для независимых выборок.

### **Требования:**

- одна независимая (группирующая) переменная, измеренная в номинальной или порядковой шкале и одна зависимая переменная, измеренная в порядковой шкале или метрической шкале, но распределение значимо отличается от нормального;
- не менее трех уровней независимой (группирующей) переменной;
- в каждой из выборок должно быть не менее трех наблюдений.

### **Алгоритм расчета критерия.**

- Сформулировать нулевую ( $H_0$ ) и альтернативную ( $H_1$ ) гипотезы.
- Сопоставить эмпирическое и критическое значения:

при  $H_{\text{эмп}} \geq H_{\text{кр}}$  отклоняется  $H_0$  и принимается  $H_1$ , при  $H_{\text{эмп}} < H_{\text{кр}}$  принимается  $H_0$  и отклоняется  $H_1$ .

# ОЦЕНКА ДОСТОВЕРНОСТИ СДВИГА

- **С** **Д** **В** **И** **Г**
  - **Временной**
  - **Ситуационный**
  - **Умозрительный**
  - **Структурный**
  - **Под влиянием контролируемых факторов**

*Временной сдвиг* определяется при сопоставлении показателей, полученных у одних и тех же испытуемых по одним и тем же методикам, но в разное время

*Ситуационный сдвиг* – это сопоставление показателей, полученных по одним и тем же методикам, но в разных условиях измерения.

*Умозрительный сдвиг* – это сопоставление показателей, измеренных в обычных и воображаемых условиях.

*Структурный сдвиг* – это сопоставление между собой разных показателей одних и тех же испытуемых, если они измерены в одних и тех же единицах, по одной и той же шкале.

*Сдвиг под влиянием контролируемых факторов* – это сопоставление замеров, произведенных до и после экспериментального воздействия, в специально созданных экспериментальных условиях.

## Критерии оценки достоверности сдвига

- Параметрические критерии
  - Если 2 выборки, то критерий Стьюдента для зависимых выборок
  - Если 3 выборки и более, то однофакторный дисперсионный анализ для зависимых выборок
- Непараметрические критерии
  - Если 2 выборки, то критерий Вилкоксона
  - Если 3 выборки и более, то критерий Фридмана



## Критерий Стьюдента для зависимых выборок

*Область применения:* критерий предназначен для оценки достоверности сдвига значений в двух зависимых выборках.

*Требования:*

- две переменные, измеренные в метрических шкалах, на одной и той же группе испытуемых;
- распределение переменных не должно значимо отличаться от нормального распределения.

*Алгоритм расчета критерия.*

- Сформулировать нулевую ( $H_0$ ) и альтернативную гипотезы ( $H_1$ ).
- Проверить нормальность распределения переменных.
- Осуществить сравнение эмпирического и критического значений  $t$ . Если  $t_{\text{эмп}} \geq t_{\text{кр}}$  при  $p \leq 0,05$ , то отвергается  $H_0$  и принимается  $H_1$ . Если  $t_{\text{эмп}} < t_{\text{кр}}$  при  $p \leq 0,1$ , то принимается  $H_0$  и отвергается  $H_1$

## Однофакторный дисперсионный анализ для зависимых выборок

*Область применения:* критерий предназначен для оценки сдвига показателей в трех и более условиях на одной и той же выборке испытуемых.

*Требования:*

- не менее трех условий фактора и не менее двух испытуемых, подвергшихся воздействию каждого условия;
- переменные должны измеряться в метрических шкалах и иметь нормальное распределение в исследуемой выборке;
- в каждом условии фактора должно соблюдаться равенство (гомогенность) дисперсий, которое достигается за счет выравнивания количества испытуемых в каждом условии фактора.

*Алгоритм расчета критерия.*

1. Сформулировать два набора гипотез.
2. Представить полученные данные в виде столбцов, каждый из которых соответствует тому или иному условию.
3. Проверить нормальность распределения переменных.
4. Сопоставить эмпирическое и критическое значения : при  $F_{\text{эмп}} \geq F_{\text{кр}}$  отклоняется  $H_0$  и принимается  $H_1$ , при  $F_{\text{эмп}} < F_{\text{кр}}$  принимается  $H_0$  и отклоняется  $H_1$

## Критерий Вилкоксона

**Область применения:** критерий предназначен для оценки достоверности сдвига значений в зависимых выборках, если данные представлены в порядковой шкале или в метрических шкалах, но не укладываются в кривую нормального распределения.

### **Требования:**

- две переменные, измеренные в порядковой шкале или в метрических шкалах, распределение которых значительно отличается от нормального, на одной и той же группе испытуемых;
- минимальное количество испытуемых, прошедших измерения в двух условиях не менее 5 человек.

### **Алгоритм расчета критерия.**

1. Сформулировать нулевую ( $H_0$ ) и альтернативную ( $H_1$ ) гипотезы, предварительно не предсказывая направление различий.
2. Сопоставить эмпирическое и критическое значения: если  $T_{\text{эмп}} \leq T_{\text{кр}}$  при  $p = 0,05$ , то отклоняется  $H_0$  и принимается  $H_1$ , если же  $T_{\text{эмп}} > T_{\text{кр}}$  при  $p = 0,05$ , то принимается  $H_0$  и отклоняется  $H_1$ .

**Замечание.** Критерий Вилкоксона отличается от большинства других критериев тем, что для опровержения нулевой гипотезы эмпирическое значение должно быть меньше или равно критическому. Следует дополнительно добавить, что этот критерий может быть односторонним (если направление сдвигов предсказывается) и двусторонним (если мы не предсказываем направление сдвигов). Уровни значимости для одностороннего и двустороннего критериев различны.

## Критерий Фридмана

*Область применения:* критерий предназначен для сопоставления показателей, измеренных в трех или более условиях на одной и той же выборке испытуемых.

*Требования:*

- переменные измеряются в порядковой шкале или в метрических шкалах, но распределение значимо отличается от нормального;
- не менее трех срезов измеряемого признака;
- не менее двух испытуемых, каждый из которых прошел не менее трех замеров.

*Алгоритм расчета критерия.*

- Сформулировать нулевую ( $H_0$ ) и альтернативную ( $H_1$ ) гипотезы.
- Осуществить сравнение эмпирического и критического значений критерия Фридмана  $\chi^2_r$ . Если  $\chi^2_{r \text{ эмп}} \geq \chi^2_{r \text{ кр}}$  при  $p \leq 0,05$ , то отвергается  $H_0$  и принимается  $H_1$ . Если  $\chi^2_{r \text{ эмп}} < \chi^2_{r \text{ кр}}$  при  $p \leq 0,1$ , то принимается  $H_0$  и отвергается  $H_1$ .