

# МЕТОДЫ МАТЕМАТИЧЕСКОЙ СТАТИСТИКИ В ПСИХОЛОГО- ПЕДАГОГИЧЕСКИХ ИССЛЕДОВАНИЯХ

1. Основные методы статистической обработки данных в психологических исследованиях;
2. Понятие статистической значимости;
3. Меры центральной тенденции;
4. Меры разброса данных;
5. Оценка достоверности отличий по t-критерию Стьюдента;
6. Вычисление коэффициента корреляции по критерию (r) Спирмена.



# Наиболее часто статистическая обработка данных в психологических исследованиях включает:

---

- Выявление **различий** между двумя группами признаков (критерий (t) Стьюдента, критерий (U) Манна-Уитни);
- Выявление **взаимосвязи** между двумя признаками (вычисление коэффициента корреляции по критерию (r) Спирмена; критерию (r) Пирсона);
- Анализ изменчивости признака под **влиянием** переменных факторов (дисперсионный анализ – вычисление критерия (F) Фишера).
- Оценка достоверности изменения **«сдвига»** в значениях исследуемого признака (вычисление критерия знаков (G); критерия (T) Вилкоксона)

# Понятие статистической значимости

Уровень статистической значимости ( $p$ ) указывает на вероятность того, что результаты ~~не представляют~~ генеральную совокупность (все объекты, относительно которых учёный намерен делать выводы)

Уровень	Возможный статистический вывод
$p > 0,1$	«Статистически достоверные различия не обнаружены» (т.е. в 9 случаях из 10 результаты вероятно, неверны)
$p < 0,1$	«Различия обнаружены на уровне статистической тенденции» (т.е. в 9 случаях из 10 результаты вероятно, верны)
$p \leq 0,05$	«Обнаружены статистически достоверные (значимые) различия» (т.е. в 95 случаях из 100 результаты вероятно, верны)
$p \leq 0,01$	«Различия обнаружены на высоком уровне статистической значимости» (т.е. в 99 случаях из 100 результаты вероятно, верны)
$p \leq 0,001$	«Различия обнаружены почти на абсолютном уровне» (т.е. в 999 случаях из 1000 результаты вероятно, верны)

# Меры центральной тенденции



---

группа методов, которые указывают наиболее типичный результат, характеризующий выполнение теста всей группой:

- среднееарифметическое значение ( $M$ )
- мода ( $M_o$ )
- медиана ( $M_e$ )

# Среднеарифметическое значение (M)

Определяется по формуле:

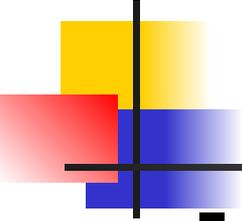
$$M = \frac{\sum X_i}{n}$$

где M - среднеарифметическое значение

n - количество испытуемых

$$\sum X_i \quad - \text{сумма всех результатов}$$

*Пример:* В исследовании объема вербальной механической памяти, тест «10 слов» в группе из 12 испытуемых получены следующие результаты (количество запомненных слов): 5, 4, 5, 6, 7, 3, 6, 2, 8, 6, 9, 7



# Мода (Mo)

---

**- наиболее часто встречающийся результат.**

Мода определяется как середина интервала, для которого частота максимальна.

*Обратите внимание,* что мода представляет собой наиболее часто встречающееся значение, а не частоту встречаемости этого значения.

*Пример:* В ряду значений 2, 3, 4, 5, 5, 6, 6, 6, 7, 7, 8, 9 модой является ?, потому что ?



# Медиана (Me)

---

**- результат, находящийся в середине последовательности показателей, если их расположить в порядке возрастания или убывания.**

Справа и слева от медианы (Me) в упорядоченном ряду остается по одинаковому количеству данных (50% и 50%). Если ряд включает в себя четное количество признаков, то медианой (Me) будет среднее, взятое как полусумма двух центральных значений ряда.

*Пример:* Найдем медиану выборки: 5, 4, 5, 6, 7, 3, 6, 2, 8, 6, 9, 7.

*Пример:* Найдем медиану выборки с нечетным количеством значений: 9, 3, 5, 8, 4, 11, 13.



# Меры разброса данных

---

характеризуют степень индивидуальных отклонений от центральной тенденции (разность между максимальной и минимальной величинами конкретного вариационного ряда);

• Дисперсия ( $S$ ) или ( $\sigma^2$ );

• Стандартное отклонение ( $\sigma$ )

# Дисперсия

- характеризует насколько частные значения отклоняются от средней величины в данной выборке. Чем больше дисперсия, тем больше отклонение или разброс данных. Определяется по формуле:

$$\sigma^2 = \frac{\sum_{i=1}^n (x_i - M)^2}{n}$$

где  $\sigma^2$  - дисперсия;

$\sum (x_i - M)^2$  - выражение, означающее, что для всех значений ( $x$ ) от первого до последнего в данной выборке вычисляется разность между частными и средними значениями, эти разности возводятся в квадрат и суммируются;

$n$  - объем выборки

# Общий алгоритм вычисления дисперсии



---

- Вычисляется среднее по выборке
- Для каждого элемента выборки вычисляется его отклонение от среднего.
- Каждый элемент множества возводят в квадрат.
- Находится сумма этих квадратов.
- Эта сумма делится на общее количество членов используемой выборки.

# Пример: вычислим дисперсию для следующего ряда: 2, 4, 6, 8, 10.

- Найдем среднее ( $M$ ) для данного ряда, оно равно ?.
- Из каждого элемента ряда вычтем величину среднего этого ряда. Экспериментальные данные представим в виде таблицы.

Первичный результат ( $x_i$ )	$x_i - M$	$(x_i - M)^2$
$M = ?$		$\sum (x_i - M)^2 = ?$

Далее разности возводят в квадрат и суммируются. Полученную сумму квадратов разностей делим на объем данной выборки.



# Стандартное отклонение ( $\sigma$ )

---

- 📌 позволяет сказать, насколько большая часть результатов данного исследования отклоняется от среднего значения. Вычисляется по формуле: **квадратный корень, извлекаемый из дисперсии, или:**

$$\sigma = \sqrt{\frac{\sum_{i=1}^n (x_i - \bar{x})^2}{n}}$$

# Пример расчета среднего квадратичного отклонения ( $\sigma$ ):

Опыт работы у пяти испытуемых составляет: 2,3,4,7 и 9 лет

■ **M** (среднее арифметическое значение) = 5 лет

$$\sigma = \sqrt{\frac{\sum_{i=1}^n (x_i - \bar{x})^2}{n}}$$

Где  $x_i$  - каждое наблюдаемое значение признака;  
 $\bar{x}$  = M – средняя арифметическая признака

$$\sigma = \sqrt{\frac{(2-5)^2 + (3-5)^2 + (4-5)^2 + (7-5)^2 + (9-5)^2}{5}} = \sqrt{6,80} = 2,607 \sim 2,61 \text{ (года).}$$

■ **σ** (среднее квадратичное отклонение) = 2,61 года

# Оценка достоверности отличий по t-критерию Стьюдента

## Вычисление первичных статистик:

$n$  – количество показателей;

$M$  – средняя арифметическая вычисленного признака;

$$M = \frac{\sum x_i}{n}$$

$\sigma$  – стандартное отклонение, среднее квадратичное отклонение показателей (вариант признака) ;

$$\sigma = \sqrt{\frac{\sum_{i=1}^n (x_i - \bar{x})^2}{n}}$$

$m_M$  – ошибка средней арифметической.

$$m_M = \pm \frac{\sigma}{\sqrt{n}}$$

# Оформление данных в таблицы

Показатели, характеризующие общение дошкольников со сверстниками (экспериментальная группа 1)

№	Фамилия, Имя ребенка	Показатели развития общения		
		Инициативность	Чувствительность	Фон
1.	Анастасия Д.	3	3	2
2.	Валера К.	3	3	3
3.	Глеб С.	1	3	3
4.	Даша П.	2	3	2
5.	Захар М.	2	3	2
6.	Илья Т.	2	3	3
7.	Кристина М.	3	3	<b>3</b>
8.	Лиза Т.	2	3	<b>3</b>
9.	Лера Ш.	3	3	3
10.	Сергей Д.	3	3	3
11.	Сергей З.	3	3	3
12.	Ульяна Б.	3	3	2
13.	Полина Т.	2	3	3
14.	Вера С.	3	3	2
15.	Семен К.	2	3	3
16.	Лера С.	3	2	2
<b>М</b>		2,5	2,9	2,6
<b>σ</b>		0,6	0,2	0,5
<b><u>m</u></b>		0,1	0,05	0,1

# Оформление данных в таблицы

Показатели, характеризующие общение дошкольников со сверстниками (экспериментальная группа 2)

№	Фамилия, Имя ребенка	Показатели развития общения		
		Инициативность	Чувствительность	Фон
1.	Алиса П.	3	2	2
2.	Артем С.	3	2	2
3.	Артем Т.	2	1	2
4.	Виолетта Б.	3	2	3
5.	Валерия Л.	2	2	3
6.	Вова Н.	3	2	3
7.	Даша И.	3	2	1
8.	Даша П.	2	2	1
9.	Карина Г.	2	1	2
10.	Кристина М.	2	2	3
11.	Лев Б.	3	3	2
12.	Максим Т.	3	3	3
13.	Маша Ф.	0	0	2
14.	Настя Е.	1	3	3
15.	Тима Е.	2	2	3
<b>М</b>		2,2	1,9	2,3
<b><math>\sigma</math></b>		1,7	0,8	0,7
<b><math>m</math></b>		0,4	0,2	0,1

# Расчет t-критерия Стьюдента

$$t_{St} = \frac{M_1 - M_2}{\sqrt{m_1^2 + m_2^2}}$$

где  $M_1$  и  $M_2$  — значения сравниваемых средних арифметических;

$t_{St}$  — величина вычисленного эмпирического критерия, который необходимо сравнивать с критическим (табл.);

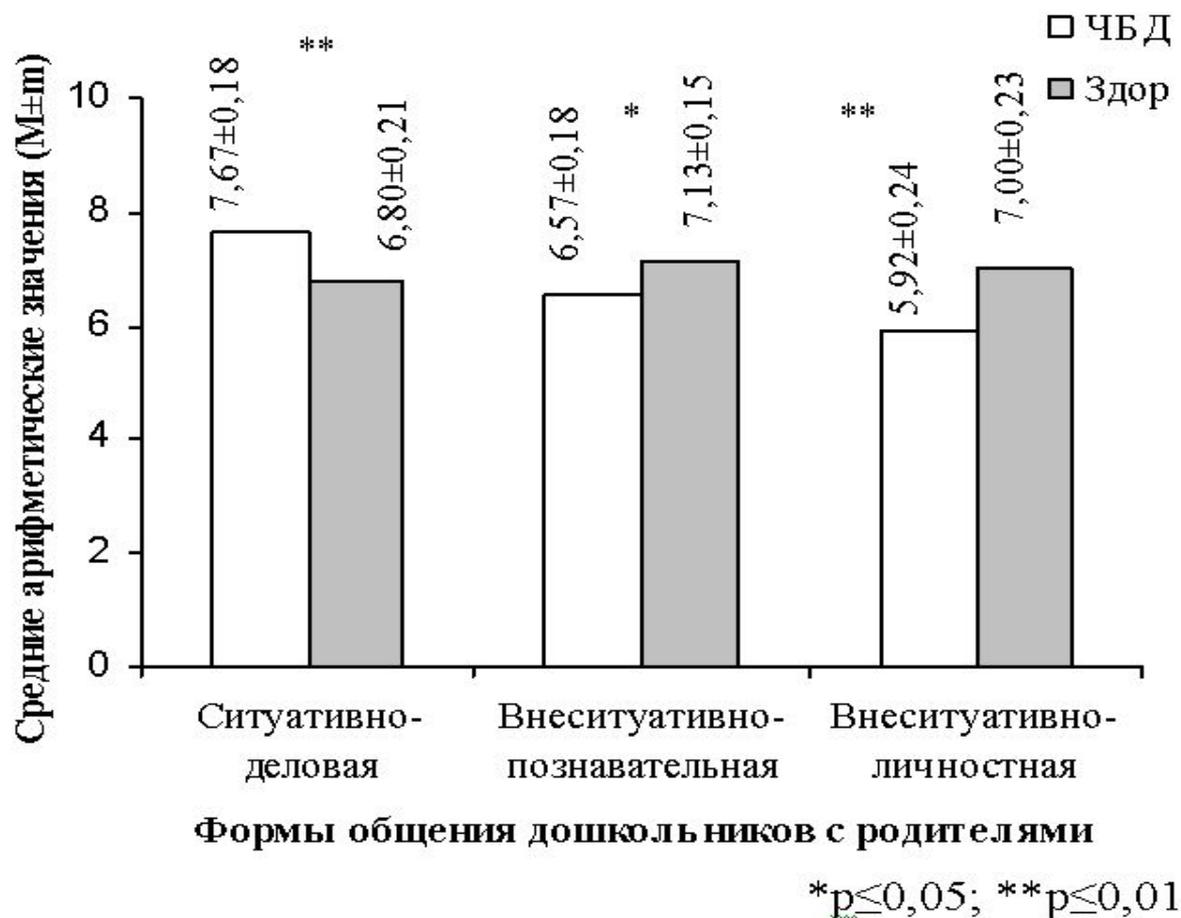
$m_1$  и  $m_2$  — соответствующие величины статистических ошибок средних арифметических.

# Оформление данных в таблицы

Сравнительная характеристика выраженности показателей общения со сверстниками детей старшего дошкольного возраста  
(экспериментальные группы 1 и 2)

Группы испытуемых	Экспери- ментальная 1	Экспери- ментальная 2	t - Стьюдента	Достовер- ность различий
Показатели развития общения				
инициативность	2,5± 0,1	2,2± 0,4	0,728	p>0,1
чувствительность	2,2 ± 0,05	1,9± 0,2	3,937	p≤0,001
фон	2,6+0,1	2,3+0,1	2,127	p≤0,05

# Графическое оформление данных



Статистическое сравнение средних значений показателей выраженности форм общения с родителями детей старшего дошкольного возраста (методика Е.О. Смирновой, А.Г. Ружской, Х.Т. Бедельбаевой)



# Корреляционный анализ

---

**Корреляционный анализ** дает возможность количественной оценки степени согласованности (взаимосвязи) различных показателей

**Наличие корреляции** между двумя показателями означает, что при изменении одного результата другой также изменяется.

Корреляция может быть **положительной** (прямой) или **отрицательной** (обратной).

При положительной корреляции - оба показателя возрастают или убывают пропорционально (коэффициент корреляции имеет положительный знак).;

При отрицательной корреляции - возрастание одного показателя сопровождается убыванием другой (коэффициент корреляции имеет отрицательный знак).

# Расчет коэффициента корреляции (r) Спирмена

## Правила ранжирования

### 1. Меньшему значению начисляется меньший ранг.

Наименьшему значению начисляется ранг 1. Наибольшему значению начисляется ранг, соответствующий количеству ранжируемых значений. Например, если  $n=7$ , то наибольшее значение получит ранг 7, за возможным исключением для тех случаев, которые предусмотрены правилом 2.

2. В случае, если несколько значений равны, им начисляется ранг, представляющий собой среднее значение из тех рангов, которые они получили бы, если бы не были равны.

Например, 3 наименьших значения равны 10 баллам. Каждое из них получает средний ранг:

$$\frac{1+2+3}{3} = \frac{6}{3} = 2$$

Если и следующие 2 значения равны (напр. 12 баллов). Они должны были бы получить ранги 4 и 5, но, поскольку они равны, то получают средний ранг:

$$\frac{4+5}{2} = 4.5 \text{ и т. д.}$$

Общая сумма рангов должна совпадать с расчетной!

Коэффициент ранговой корреляции  
Спирмена ( $r_s$ ) подсчитывается по  
формуле:

$$r_s = 1 - \frac{6 \cdot \sum(d^2)}{N \cdot (N^2 - 1)}$$

где  $d$  - разность между рангами по двум переменным  
для каждого испытуемого;

$N$  - количество ранжируемых значений, в данном  
случае количество испытуемых.

# Пример:

Расчет  $d^2$  для рангового коэффициента корреляции Спирмена при сопоставлении показателей количества ошибок и показателей интеллекта у студентов (N=10)

Испытуемы й		Переменная А Количество ошибок		Переменная Б Показатели интеллекта		D (ранг А - - ранг Б)	$d^2$
		Индивидуальн ые значения	Ранг	Индивидуальн ые значения	Ран г		
1	ТА.	29	9	131	4	5	25
2	ПА. Ч.И.	54	10	132	5.5	4.5	20.25
3	ЦА.	13	4	121	1	3	9
4	С.А. К.	8	2	127	3	-1	1
5	Е.	14	5	136	9	-4	16
6	КА.	26	8	124	2	6	36
7	Б.Л.	9	3	134	7	-4	16
8	ИА. Ф.В.	20	7	136	9	-2	4
9		2	1	132	5.5	-4.5	20,25
10		17	6	136	9	-3	9
Суммы			55		55	0	156,5

$$r_s = 1 - \frac{6 \cdot 156,5}{10 \cdot (10^2 - 1)} = 1 - \frac{939}{990} = 0,052$$



# Классификация корреляционных связей

---

- **Высокозначимая корреляция** – при  $r_c$  соответствующем уровню статистической значимости  $p \leq 0,01$ ;
- **Значимая корреляция** – при  $r$ , соответствующем уровню статистической значимости  $p \leq 0,05$ ;
- **Тенденция достоверной связи** – при  $r$ , соответствующем уровню статистической значимости  $p \leq 0,1$ ;
- **Незначимая корреляция** – при  $r$ , соответствующем уровню статистической значимости  $p > 0,1$ .

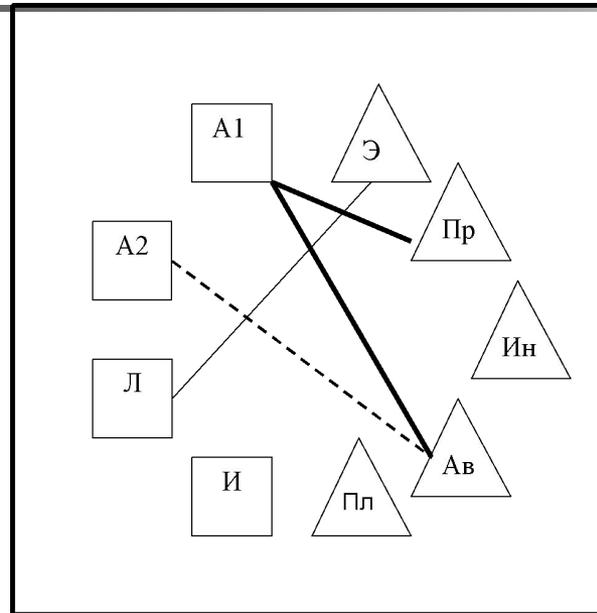
# Оформление данных в таблице

Выявление взаимосвязи родительского отношения и особенностей общения детей старшего дошкольного возраста со сверстниками

Показатели родительского отношения к детям	Принятие	Отвержение	Симбиоз	Кооперация	Авторитарная гиперсоциализация	Маленький неудачник
Показатели общения у детей						
Инициативность	0,21	- 0,12	- 0,27	0,34	- 0,19	- 0,64 $p \leq 0,05$
Чувствительность	0,17	- 0,2	- 0,41	0,77 $p \leq 0,01$	- 0,15	- 0,33
Фон	6,19 $p \leq 0,05$	- 0,14	- 0,37	0,18	- 0,06	- 0,42

# Графическое оформление данных

Корреляционные плеяды взаимосвязей показателей, характеризующих стратегии семейного воспитания и особенностей личностных качеств подростков, способствующих их профессиональному самоопределению



*Условные обозначения.* Значимая сильная положительная взаимосвязь ( $p \leq 0,01$ ) обозначена сплошной жирной линией, значимая умеренная положительная взаимосвязь ( $p \leq 0,01$ ) обозначена сплошной тонкой линией, отрицательная взаимосвязь (на уровне статистической тенденции) обозначена пунктирной линией:

при: - стратегии семейного воспитания («А1» - авторитетная, «А2» - авторитарная, «Л» - либеральная, «И» - индифферентная);

△ - личностные качества подростка, способствующие профессиональному самоопределению («Э» - эмоциональное отношение, «Пр» - принятие решений, «Ин» - информированность, «Ав» - автономность, «Пл» - планирование).