Статистическая обработка результатов психолого-педагогических исследований

План лекции

1. Измерение. Шкалы. Группировки.
2. Статистическая и генеральная совокупность.
3. Меры центральной тенденции.
4. Меры разброса значений.
5. Методы определения достоверности различий.
6. Методы определения коэффициентов корреляции.

1. Измерение. Шкалы. Группировки.

Измерение — процедура приписывания чисел объектам изучения в соответствии с определенными правилами.

В качестве объектов измерения:
- «единицы» поведения,
- физиологические реакции.

1. Измерение. Шкалы. Группировки.

Показатели – количественные и качественные характеристики действий, высказываний, физиологических реакций и.т.п.

Виды показателей:

- количественные,
- качественные.

1. Измерение. Шкалы. Группировки.

Для измерения различных признаков используются шкалы.
Шкала — числовая система.

Виды шкал:

- номинальная,
- ранговая (порядковая),
- интервальная (метрическая).

1. Измерение. Шкалы. Группировки.

Номинальная шкала — измеряются объективные признаки респондентов (пол, возраст, семейное положение и т.п.).

Пример:

возраст — 23 года.

Возможные выводы:

- равно-неравно,
- больше-меньше,
- во сколько раз больше или меньше.

1. Измерение. Шкалы. Группировки.

Ранговая (порядковая) шкала — измеряются субъективные признаки респондентов (степень удовлетворенности чем-либо или кем-либо и т.п.).

Пример:

Степень удовлетворенности профессией:

5 — полностью удовлетворен,
4 — удовлетворен,
3 — затрудняюсь ответить,
2 — скорее, не удовлетворен,
1 — полностью не удовлетворен.

Возможные выводы:

- равно-неравно,
- больше-меньше,
- во сколько раз больше или меньше?

1. Измерение. Шкалы. Группировки.

Интервальная (метрическая) шкала — измеряются объективные признаки респондентов (пол, возраст, семейное положение и т.п.) с помощью интервалов.

Пример:

Возраст:

1. 0 — 5 лет,
2. 6 — 10 лет,
3. 11 — 15 лет.

Возможные выводы:

- равно-неравно,
- больше-меньше,
- во сколько раз больше или меньше?

1. Измерение. Шкалы. Группировки.

Группировка — распределение единиц изучаемого объекта на однородные группы по существенным для него признакам.

Пример:

возраст — 23 года ....

Назначение группировки:

- установление численности каждой отдельно взятой части совокупности,
- изучение влияния причин и характеристики явления.

1. Измерение. Шкалы. Группировки.

Виды группировок:

1. комбинационная:
 а) структурная,
 б) типологическая,

2. аналитическая.

1. Измерение. Шкалы. Группировки.

Комбинационная группировка — распределение в группы по двум и более признакам.

а) структурная группировка — с учетом объективных признаков,

б) типологическая группировка — с учетом субъективного признака.

Аналитическая группировка - распределение в группы по двум и более признакам для выявления их взаимосвязи (уровень мышления и успеваемость).

2. Статистическая и генеральная совокупность

2. Статистическая и генеральная совокупность

Статистическая совокупность — это объединение какого-либо множества испытуемых по одному или нескольким признакам.

При этом:

- выделяемая совокупность должна быть однородна по основным качественным характеристикам;
- сравнение может проводиться лишь по тому признаку, который является предметом исследования.

2. Статистическая и генеральная совокупность

Статистическая совокупность = объем выборки.

- если объем выборки 30 и более человек, то используется аппарат параметрической статистики,
- если объем выборки от 10 до 30 человек, то используется аппарат непараметрической статистики.

2. Статистическая и генеральная совокупность

Генеральная совокупность — объект исследования, который территориально, производственно и во времени «локализован» и на который распространяются выводы исследования.

2. Статистическая и генеральная совокупность

Ряд распределения — *упорядоченный ряд чисел, получаемый в результате группировки.*

Виды рядов распределения:

- атрибутный — упорядоченный ряд распределения по качественным признакам,

- вариационный — упорядоченный ряд распределения по количественным признакам.

Вариационый ряд может быть дискретным и непрерывным (интервальным).

2. Статистическая и генеральная совокупность

Вариационный (непрерывный) ряд распределения

Пример: объем произвольного внимания детей 7 лет (n= 8): 1; 1; 2; 2; 2; 3; 3; 4.

варианты - 1 2 3 4
частоты - 2 3 2 1

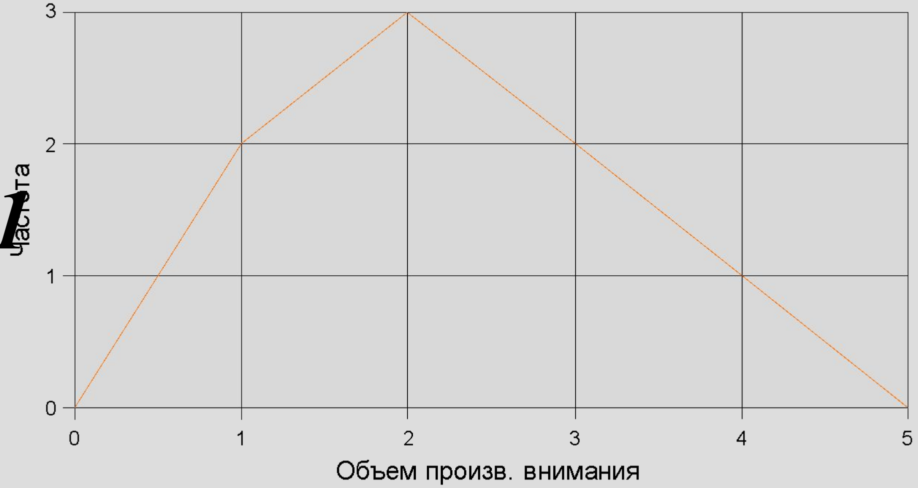


Рис. 1. Объем произвольного внимания ЧБД 7 лет

2. Статистическая и генеральная совокупность

Атрибутный ряд распределения

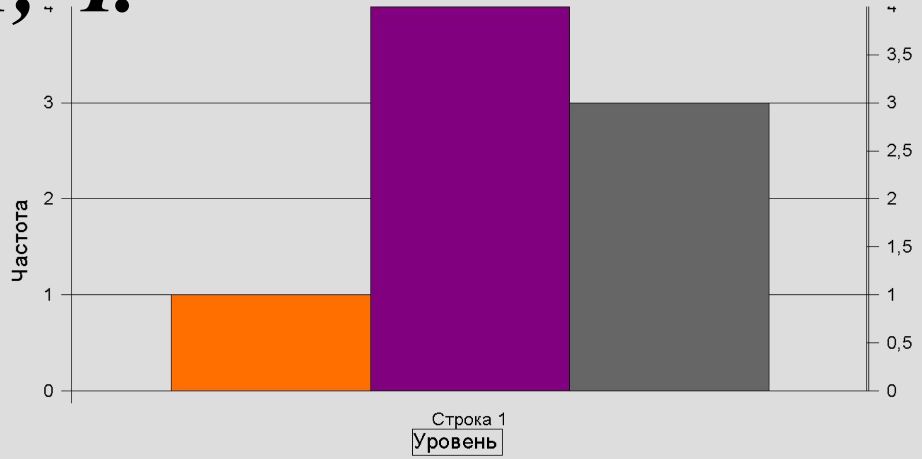
Пример: уровень развития творческого воображения детей 7 лет (n= 8):

В; В; В; С; С; С; С; Н.

III; III; III; II; II; II; II; I.

атрибуты - I II III

частоты - 1 4 3

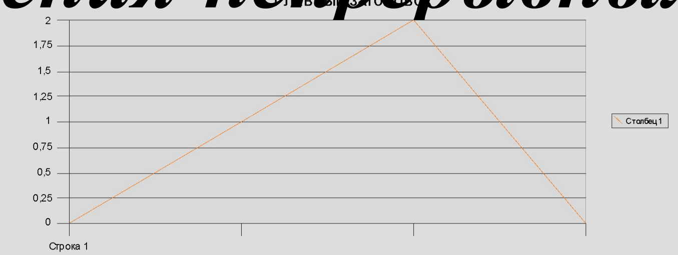


ис. 1. Уровень развития творческого воображения ЧБД 7

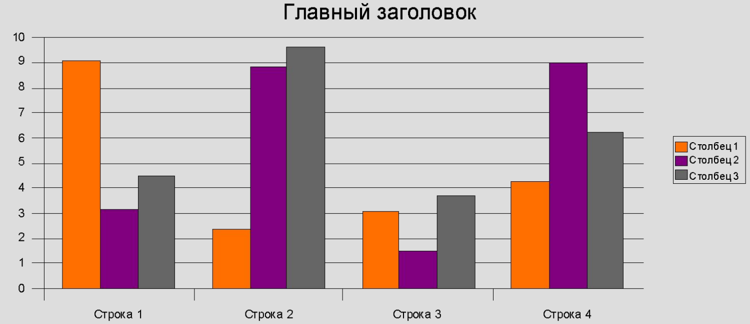
2. Статистическая и генеральная совокупность

Графическое изображение статистических данных:

- полигон - для отображения непрерывных рядов,



- гистограмма — для отображения дискретных рядов.



3. Меры центральной тенденции

3. Меры центральной тенденции

Меры центральной тенденции — величины, вокруг которых группируются отдельные, расходящиеся между собой значения показателя.

С их помощью множество разбросанных показателей заменяет одна величина.

Меры центральной тенденции:

M (X) — среднее арифметическое,

Mo — мода,

Md — медиана.

3. Меры центральной тенденции

M — среднее арифметическое

$$M = \sum v_i / n$$

Пример:
$V_{1\,(n=7)}$: 1; 2; 3; 4; 4; 4; 5.

$M = (1+2+3+4+4+4+5) / 7 \approx 3{,}29.$

3. Меры центральной тенденции

Мо — мода — максимально встречающийся результат.

Пример:
$V_{1(n=9)}$: 1; 2; 3; 3; 3; 4; 4; 4; 5.

Mo_1 = 3; Mo_2 = 4.

3. Меры центральной тенденции

Md — медиана — числовое значение, занимающее в упорядоченном ряду данных срединное положение (делит упорядоченный ряд на две равные части).

Расчет места медианы:

Место медианы = (n+1)/2

Пример:
$V_{1\ (n=8)}$: 1; 2; 3; 3; 3; 4; 4; 5.
Место медианы = (8+1)/2 = 4,5.
Md = 3.

4. Меры разброса значений

4. Меры разброса значений

При замене множества числовых значений показателя одним числом — средним арифметическим или медианой — мы, выигрывая в простоте и наглядности ситуации, теряем часть информации.

Так, два множества значений имеют одинаковые M и Md:

V_1: 5; 5; 5. M= 5,0. Md= 5.

V_2: 1; 5; 9. M= 5,0. Md= 5.

4. Меры разброса значений

Меры разброса значений:

W - размах

δ — стандартное отклонение

m — ошибка среднего арифметического

4. Меры разброса значений

W — размах — разность максимального и минимального значений в ряду данных.

Пример:

$V_{1\,(n=8)}$: 2; 3; 4; 6; 7; 8; 9; 10.

$W_1 = v_{max} - v_{min} = 10 - 2 = 8.$

$V_{2\,(n=9)}$: 4; 4; 5; 5; 6; 6; 6; 6; 6.

$W_2 = ?$

4. Меры разброса значений

δ — стандартное отклонение.

$$\delta = \sqrt{\sum (v_i - M)^2 / (n - 1)}$$

Пример: $V_{1\ (n=10)}$: 2; 2; 3; 3; 3; 3; 4; 4; 5; 6.

4. Меры разброса значений

V i	Vi — M	(Vi — M)²
2	2 — 3,5 = -1,5	2,25
2	-1,5	2,25
3	-0,5	0,25
3	-0,5	0,25
3	-0,5	0,25
3	-0,5	0,25
4	0,5	0,25
4	0,5	0,25
5	1,5	2,25
6	2,5	6,25
M = 3,5.		∑(Vi — M)² = 14,5

4. Меры разброса значений

δ — стандартное отклонение.

$$\delta = \sqrt{\sum (v_i - M)^2 / (n - 1)} =$$

$$\sqrt{14{,}5 / (10 - 1)} = \sqrt{14{,}5 / 9} =$$

$$\sqrt{1{,}611} \approx 1{,}269.$$

4. Меры разброса значений

m — ошибка среднего арифметического

$$m = \partial / \sqrt{n} =$$

$$\approx 1,269 / \sqrt{10} \approx 1,269 / 3,161 \approx 0,401.$$

Запись ряда распределения:

$$M \pm m = 3,5 \pm 0,40.$$

5. Методы определения достоверности различий

5. Методы определения достоверности различий

Для установления факта случайности различий средних арифметических зависимых и независимых выборок или его опровержения пользуются статистическими критериями (если исследователь хочет распространить свои выводы на генеральную совокупность).

5. Методы определения достоверности различий

Зависимые выборки — выборки, в которых результаты измерения некоторого свойства испытуемых одной выборки **влияют** на результаты измерения этого свойства испытуемых другой выборки.

Независимые выборки — выборки, в которых результаты измерения некоторого свойства испытуемых одной выборки **не влияют** на результаты измерения этого свойства испытуемых другой выборки.

5. Методы определения достоверности различий

- Методы определения достоверности различий для зависимых выборок:
- **t** критерий Стьюдента,
- критерий знаков.

- Методы определения достоверности различий для независимых выборок:
- U критерий Манна-Уитни,
- **t** критерий Стьюдента.

5. Методы определения достоверности различий

Определение t критерия Стьюдента,

t Стьюдента = $(M_1 - M_2) / \sqrt{(m^2_1 + m^2_2)}$.

M_1 – среднее арифметическое большего значения,
M_2 – среднее арифметическое меньшего значения.

Ограничение применения методики – необходимо **симметричное** распределение показателей.

5. Методы определения достоверности различий

Ограничение применения методики – ***необходимо*** ***симметричное распределение показателей.***

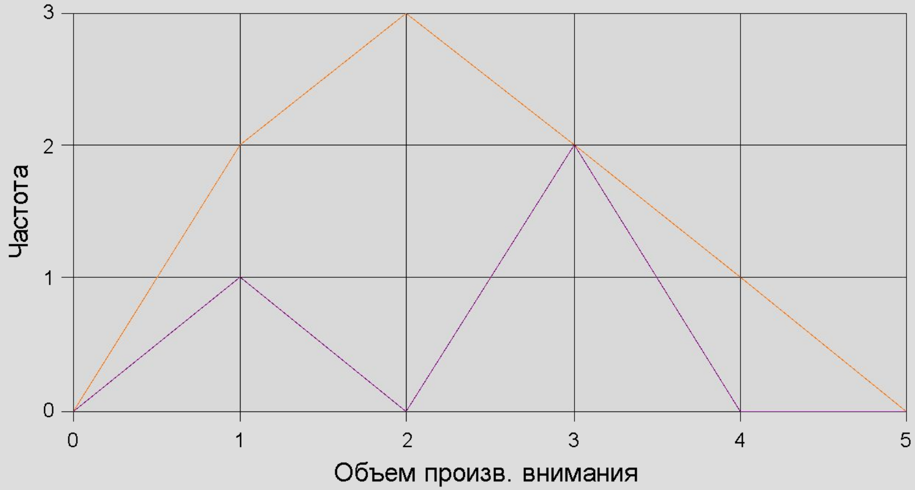


Рис. 1. Объем произвольного внимания ЧБД 7 лет

5. Методы определения достоверности различий

Пример:

$M_1 \pm m_1 = 3{,}33 \pm 0{,}401$, при $n = 12$.
$M_2 \pm m_2 = 3{,}82 \pm 0{,}412$, при $n = 14$.

t Стьюдента $= (M_1 - M_2) \,/\, \sqrt{(m^2_1 + m^2_2)} \approx$

$\approx (3{,}82 - 3{,}33) / \sqrt{(0{,}401^2 + 0{,}412^2)} \approx$

$\approx 0{,}49 / \sqrt{(\mathbf{0{,}160801 + 0{,}169744})} \approx$

$\approx 0{,}49 / \sqrt{\mathbf{0{,}330545}} \approx 0{,}49 / 0{,}5749 \approx 0{,}852.$

5. Методы определения достоверности различий

Нахождение статистически достоверной вероятности различий с помощью t критерия Стьюдента:

Гипотеза H_0: если t расчетная < t табличной, то между рядами показателей не существует достоверное различие на уровне 95 % (98 %, 99 % или 99,5 %) вероятности.

Гипотеза H_1: если t расчетная ≥ t табличной, то между рядами показателей существует достоверное различие на уровне 95 % (98 %, 99 % или 99,5 %) вероятности.

Так как, t расчетная (0,852) < t табличной (2,06), то между анализируемыми рядами показателей не существует достоверное различие на уровне 95 % вероятности. Подтвердилась гипотеза H_0.

5. Методы определения достоверности различий

Метод определения достоверности различий для **зависимых выборок** -

- критерий знаков.

Пример: Необходимо выявить наличие достоверных различий в объеме произвольного внимания ЧБД 7 лет до и после формирующего эксперимента.

5. Методы определения достоверности различий

Критерий знаков

Пример: Необходимо выявить наличие достоверных различий в объеме произвольного внимания ЧБД 7 лет до и после формирующего эксперимента.

Ребенок №	1	2	3	4	5	6	7	8	9	10	11	12
V прозв.внимания до ...	4	3	4	5	4	2	3	4	3	4	1	2
V прозв.внимания после ...	5	4	4	6	5	4	4	3	4	4	4	3
Знак изменения результата	+	+	=	+	+	+	+	-	+	=	+	+

5. Методы определения достоверности различий

Критерий знаков – обработка:

1. $n' = 12 - 2 = 10$. (различающиеся пары результатов)
2. $K_{max} = 9$. (количество чаще встречающихся знаков)
3. $K_{табл}$ $(n' = 10) = 9$.

Ребенок №	1	2	3	4	5	6	7	8	9	10	11	12
V прозв.внимания до …	4	3	4	5	4	2	3	4	3	4	1	2
V прозв.внимания после …	5	4	4	6	5	4	4	3	4	4	4	3
Знак изменения результата	+	+	=	+	+	+	+	-	+	=	+	+

5. Методы определения достоверности различий

4. Нахождение статистически достоверной вероятности различий с помощью критерия знаков:

Гипотеза H_0: если $K_{max} < K_{табл}$, то между рядами показателей не существует достоверное различие на уровне 95 % вероятности.

Гипотеза H_1: если $K_{max} \geq K_{табл}$, то между рядами показателей существует достоверное различие на уровне 95 % вероятности.

Так как, K_{max}= 9 равен $K_{табл}$ = 9, то между анализируемыми рядами показателей существует достоверное различие на уровне 95 % вероятности. Подтвердилась гипотеза H_1.

Пограничные значения критерия знаков (95% уровень достоверности)(Рунион Р. Справочник по непараметричяеской статистике. М., 1982)

n'	K	n'	K	n'	K
6	6	16	13	26	19
7	7	17	13	27	20
8	8	18	14	28	20
9	8	19	15	29	21
10	9	20	15	30	21
11	10	21	16	31	22
12	10	22	17	32	23
13	11	23	17	33	23
14	12	24	18	34	24
15	12	25	18	35	24

5. Методы определения достоверности различий

Метод определения достоверности различий для ***независимых*** выборок - U критерий Манна-Уитни.

Пример: **двум** группам ЧБД 7 лет предлагалось запомнить 10 новых слов в условиях игры и в условиях лабораторного эксперимента.

5. Методы определения достоверности различий

U критерий Манна-Уитни.

$V_{\text{игра}}$ (n=11): 3; 4; 4; 4; 5; 5; 5; 5; 5; 6; 6. (M= 4,7.)

$V_{\text{экспер}}$ (n=10): 2; 3; 3; 3; 3; 4; 4; 4; 4; 5. (M= 3,5.)

Код	И	И	И	И	И	И	И	Э	Э	Э	Э	Э	И	И	И	И	Э	Э	Э	Э	Э
Количество слов	6	6	5	5	5	5	5	5	4	4	4	4	4	4	4	3	3	3	3	3	2
номер записи	1	2	3	4	5	6	7	8	9	10	11	12	13	14	15	16	17	18	19	20	21
Ранг	1,5	1,5																			

Ранг – порядковый номер:
ранг 6 = (1+2) / 2 = 1,5.
ранг 5 = ? ранг 4 =? ранг 3 = ? ранг 2 = ?

5. Методы определения достоверности различий

Код	И	И	И	И	И	И	И	Э	Э	Э	Э	Э	Э	Э	Э	Э	Э	Э	Э	Э	Э
Количество слов	6	6	5	5	5	5	5	5	4	4	4	4	4	4	4	3	3	3	3	3	2
номер записи	1	2	3	4	5	6	7	8	9	10	11	12	13	14	15	16	17	18	19	20	21
Ранг	1,5	1,5	5,5	5,5	5,5	5,5	5,5	5,5	12	12	12	12	12	12	12	18	18	18	18	18	21

Обработка:
1. Сумма рангов для иргы и для эксперимента:

$R_{игра}$ = 1,5*2 + 5,5 * 5 + 12*3 + 18*1 = 84,5.
$R_{экспер}$ = ?

2. Проверка: Rигра + Rэкспер = N / 2 * (N + 1)?
где N = $n_{игра}$ + $n_{экспер}$.

84,5 + 146,5 = 21 / 2 * 22.   231 = 231.

5. Методы определения достоверности различий

U критерий Манна-Уитни.
Обработка:

3. $U_{игра} = \mathbf{n}_{и}*\mathbf{n}_{э} + \mathbf{n}_{и}(\mathbf{n}_{и} + 1) / 2 – R_{игра}$.

$U_{игра} = 11*10 + 11*12 / 2 – 84,5 = 91,5$.

4. $U_{экспер} = \mathbf{n}_{и}*\mathbf{n}_{э} + \mathbf{n}_{э}(\mathbf{n}_{э} + 1) / 2 – R_{экспер}$.

$U_{экспер} = 11*10 + 10*11 / 2 – 146,5 = 18,5$.

5. Проверка: $U_{игра} = \mathbf{n}_{и}*\mathbf{n}_{э} – U_{экспер}$.

$91,5 = 11*10 – 91,5 = 110 – 18,5 = 91,5$.

5. Методы определения достоверности различий

4. Нахождение статистически достоверной вероятности различий с помощью U критерия Манна-Уитни:

Гипотеза Н0: если Umax расчетная < Umax табличная, а Umin расчетная > Umin табличная, то между рядами показателей не существует достоверное различие на уровне 95 % вероятности.

Гипотеза Н1: если Umax расчетная ≥ Umax табличная, а Umin расчетная < Umin табличная, то между рядами показателей существует достоверное различие на уровне 95 % вероятности.

Так как, Umax расчетная (91,5) > Umax табличная (84), а Umin расчетная (18,5) < Umin табличная (26), то между анализируемыми рядами показателей существует достоверное различие на уровне 95 % вероятности. Подтвердилась гипотеза Н1.

5. Методы определения достоверности различий

H – критерий Краскела-Уоллеса
- предназначен для оценки различий по какому-либо показателю между тремя и более выборками.

$$H = [12/n(n+1)] * (\Sigma R^2/n_k) - 3(n+1),$$

R – суммы рангов по группам,
k – количество групп,
n_k – объем групп,
n – объем объединенной выборки.

5. Методы определения достоверности различий

H – критерий Краскела-Уоллеса

Пример: Существует ли достоверные различия в степени стрессоустойчивости у представителей четырех групп студентов – ППФ, физики, физической культуры, музыки.

№	Факультет	Тестовые баллы
1	ППФ	228
2	ППФ	214
3	ППФ	239
4	Физики	194
5	Физики	221
6	ФФК	150
7	ФФК	155
8	ФФК	192
9	ФФК	181
10	ФФК	166
11	Музыки	280
12	Музыки	211
13	Музыки	235

5. Методы определения достоверности различий

Н – критерий Краскела-Уоллеса

Подсчет суммы рангов для каждой группы

ФФК		Физики		ППФ		Музыки	
Балл	Ранг	Балл	Ранг	Балл	Ранг	Балл	Ранг
150	1						
155	2						
166	3						
181	4						
192	5						
		194	6				
						211	7,5
				211	7,5		
		220	9				
				228	10		
						235	11
				239	12		
						280	13
$\sum r_1 = 15$		$\sum r_2 = 15$		$\sum r_3 = 29{,}5$		$\sum r_4 = 31{,}5$	

5. Методы определения достоверности различий

H – критерий Краскела-Уоллеса

$H = [12/n(n+1)] * (\Sigma R^2/n_k) - 3(n+1)$,

R – суммы рангов по группам,

k – количество групп,

n_k – объем групп,

n – объем объединенной выборки.

$H = [12/13(13+1)] * (15^2/5 + 15^2/2 + 29{,}5^2/3 + 31{,}5^2/3) -$

$-3(13+1) \approx 9{,}4$.

5. Методы определения достоверности различий

H – критерий Краскела-Уоллеса

$$H = [12/13(13+1)] * (15^2/5 + 15^2/2 + 29{,}5^2/3 + 31{,}5^2/3) -$$

$$-3(13+1) \approx 9{,}4.$$

Степень свободы H-критерия: df = k – 1 = 4 - 1 = 3.

Для определения табличного (критического) распределения статистики χ^2 используем таблицу.

df	p = 0,1	p = 0,05	p = 0,01
3	6,25	7,82	11,35

5. Методы определения достоверности различий

H – критерий Краскела-Уоллеса

Гипотеза H_0: если $H_{расчетная} < H_{табличная}$, то между рядами показателей не существует достоверное различие на уровне 95 % вероятности.

Гипотеза H_1: если $H_{расчетная} \geq H_{табличная}$, то между рядами показателей существует достоверное различие на уровне 95 % вероятности.

Так как, $H_{расчетная}$ (9,4) > $H_{табличная}$ (7,815), то между анализируемыми рядами показателей существует достоверное различие на уровне 95 % вероятности. Подтвердилась гипотеза H_1.

6. Методы определения коэффициентов корреляции

6. Методы определения коэффициентов корреляции

Корреляция — оценка статистической связи между двумя рядами данных.

- изменяются ли показатели одного ряда при изменении показателей другого ряда.

Коэффициент корреляции — в пределах:
от **+1** (прямая функциональная связь)
до **-1** (обратная функциональная связь).

Если коэффициент корреляции близок к **0, то между рядами данных статистической связи нет.**

6. Методы определения коэффициентов корреляции

Виды коэффициентов корреляция:

r_{xy} — коэффициент корреляции Пирсона,

ρ — коэффициент ранговой корреляции Спирмена.

6. Методы определения коэффициентов корреляции

r_{xy} — коэффициент корреляции Пирсона

Ограничения:
- количественные показатели,
- симметричное (нормальное) распределение данных.

$$r_{xy}=(n*\sum x_i*y_i - \sum x_i * \sum y_i) / \sqrt{[n*\sum x_i^2 - (\sum x_i)^2] * [n*\sum y_i^2 - (\sum y_i)^2]}.$$

6. Методы определения коэффициентов корреляции

r_{xy} — коэффициент корреляции Пирсона

Пример: **найти корреляционную связь между объемом зрительного внимания (x) и количеством ошибок (y).**

$$r_{xy}=(n*\sum x_i*y_i - \sum x_i * \sum y_i) / \sqrt{[n*\sum x_i^2 - (\sum x_i)^2] * [n*\sum y_i^2 - (\sum y_i)^2]}.$$

6. Методы определения коэффициентов корреляции

$$r_{xy}=(n*\sum x_i*y_i — \sum x_i * \sum y_i) / \sqrt{[n*\sum x_i^2 — (\sum x_i)^2] * [n*\sum y_i^2 — (\sum y_i)^2]}.$$

n	xi	yi	xi * yi	xi²	yi²
1	10	2	20	100	4
2	10	4	40	100	16
3	9	3	27	81	9
4	9	5	45	81	25
5	8	3	24	64	9
6	7	6	42	49	36
7	7	9	63	49	81
8	5	9	45	25	81
9	5	7	35	25	49
10	4	10	40	16	100
	$\sum x_i = 74$	$\sum y_i = 58$	$\sum x_i * y_i = 381$	$\sum x_i^2 = 590$	$\sum y_i^2 = 410$

$r_{xy} = (n*\sum x_i*y_i — \sum x_i * \sum y_i) / \sqrt{[n*\sum x_i^2 — (\sum x_i)^2]*[n*\sum y_i^2 — (\sum y_i)^2]}$

$= (10*381 – 74*58) / \sqrt{[10*590 — 74^2] * [10*410 — 58^2]} \approx -482 / 558{,}40 \approx -0{,}86.$

n	xi	yi	xi * yi	xi²	yi²
1	10	2	20	100	4
2	10	4	40	100	16
3	9	3	27	81	9
4	9	5	45	81	25
5	8	3	24	64	9
6	7	6	42	49	36
7	7	9	63	49	81
8	5	9	45	25	81
9	5	7	35	25	49
10	4	10	40	16	100
	$\sum x_i = 74$	$\sum y_i = 58$	$\sum x_i * y_i = 381$	$\sum x_i^2 = 590$	$\sum y_i^2 = 410$

6. Методы определения коэффициентов корреляции

Для оценки значимости r_{xy} необходимо сравнить полученный коэффициент с табличным коэффициентом:

Гипотеза H_0: если $r_{xy} \leq r_{\text{табличная}}$, то между рядами показателей не существует достоверная связь на уровне 95 % вероятности.

Гипотеза H_1: если $r_{xy} > r_{\text{табличная}}$, то между рядами показателей существует достоверная связь на уровне 95 % (99%) вероятности.

Так как, r_{xy} (-0,86) > $r_{\text{табличная}}$ (0,77), то между анализируемыми рядами показателей существует достоверная обратная связь на уровне 99 % вероятности. Подтвердилась гипотеза H_1

6. Методы определения коэффициентов корреляции

ρ — коэффициент ранговой корреляции Спирмена

Ограничения применения:

-наличие качественных показателей.

$$\rho = 1 - 6 * \sum d_i^2 / n (n^2 - 1).$$

6. Методы определения коэффициентов корреляции

$$\rho = 1 - 6 * \sum d_i^2 / n\,(n^2 - 1).$$

6. Методы определения коэффициентов корреляции

$\rho = 1 – 6 * \sum d_i^2 / n (n^2 - 1).$

$\rho = 1 – 6 * 62 / 10 (100 – 1) \approx 1 – 0,376 \approx 0,624.$

6. Методы определения коэффициентов корреляции

Для оценки значимости ρ необходимо сравнить полученный коэффициент с табличным коэффициентом:

Гипотеза H_0: если $\rho \leq \rho_{\text{табличная}}$, то между рядами показателей не существует достоверная связь на уровне 95 % вероятности.

Гипотеза H_1: если $\rho > \rho_{\text{табличная}}$, то между рядами показателей существует достоверная связь на уровне 95 % (99%) вероятности.

Так как, ρ (0,624) > $\rho_{\text{табличная}}$ (0,564), то между анализируемыми рядами показателей существует достоверная прямая связь на уровне 95 % вероятности. Подтвердилась гипотеза H_1

7. Кластерный анализ

Кластерный анализ

(таксономический) анализ используется для упорядочивания объектов и объединения их в однородные разряды.

Кластер — это группа объектов, характеризующаяся повышенной плотностью и дисперсией.

Кластерный анализ

Однородность объектов определяется по расстоянию р (x1, x2).

Объекты считаются однородными, если