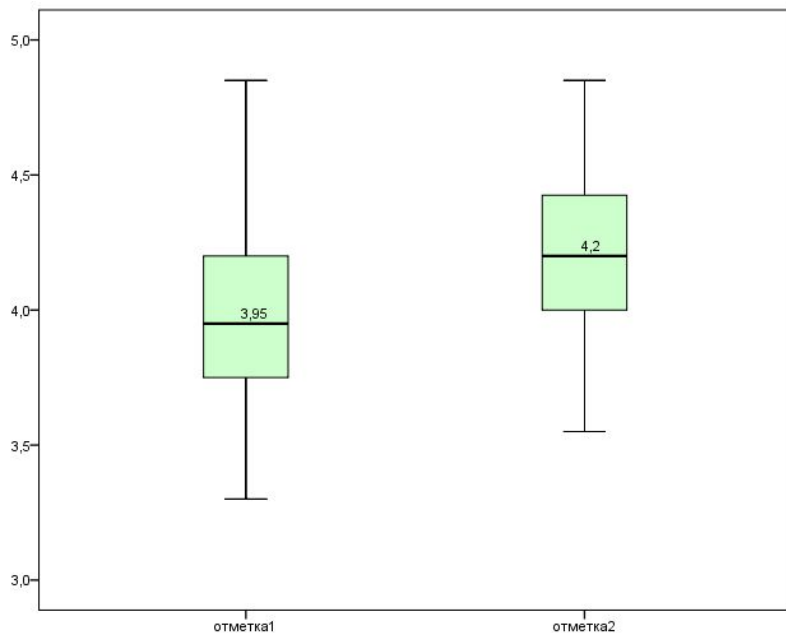
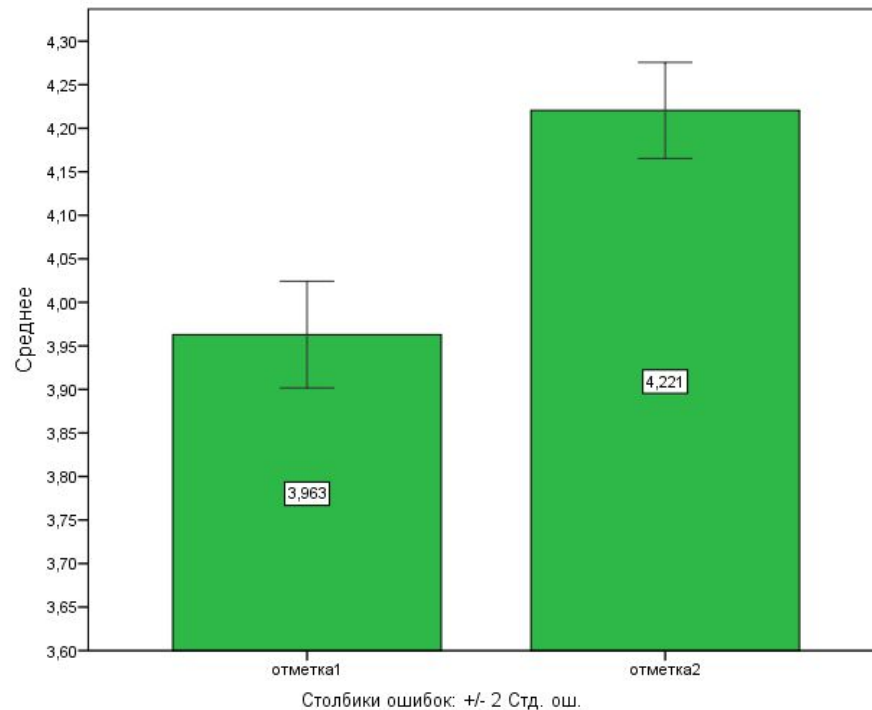




## Статистики

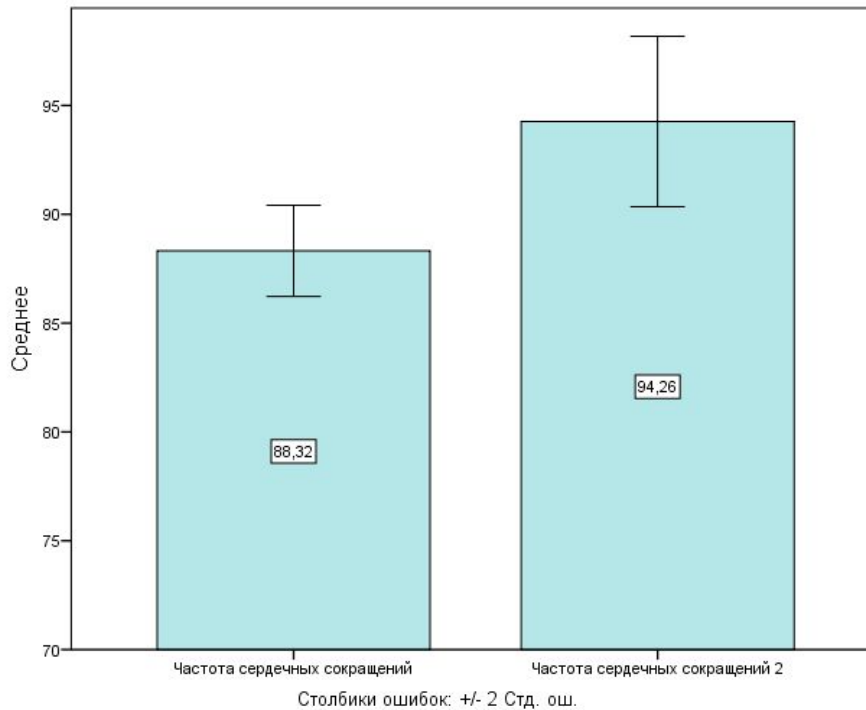
	отметка1	отметка2	
Среднее	3,96	4,22	
Стд. ошибка среднего	,03	,03	
Медиана	3,95	4,20	
Мода	3,90 <sup>а</sup>	4,20	
Стд. отклонение	,31	,28	
Асимметрия	,32	-,01	
Стд. ошибка асимметрии	,24	,24	
Эксцесс	-,06	-,36	
Стд. ошибка эксцесса	,48	,48	
Размах	1,55	1,30	
Минимум	3,30	3,55	
Максимум	4,85	4,85	
Процентили			
	25	3,75	4,00
	50	3,95	4,20
	75	4,20	4,44

а. Имеется несколько мод. Показана наименьшая.



## Вывод:

- 1) В 11 классе успеваемость выше, чем в 10 (4,2 выше, чем 3,95; среднее для 11 класса выше, чем среднее для 10).
- 2) Самые слабые ученики стали учиться лучше (см. ящики), разброс оценок уменьшился.
- 3) Средние ученики также улучшили свою успеваемость (смещение ящика вверх).
- 4) Асимметрия была положительной, стала практически равна нулю, следовательно, в 10 классе был перевес в сторону успеваемости ниже среднего, а в 11 классе перевеса нет. Кроме того, медиана и среднее выросли, значит слабо успевающие дети стали учиться лучше.
- 5) Эксцесс был ок. нуля, а стал отрицательным. Это означает, что в 10 классе учащихся, имеющих средний балл, было в норме, а в 11 классе стало ниже нормы, т.е. часть детей повысили успеваемость (см. асимметрию).



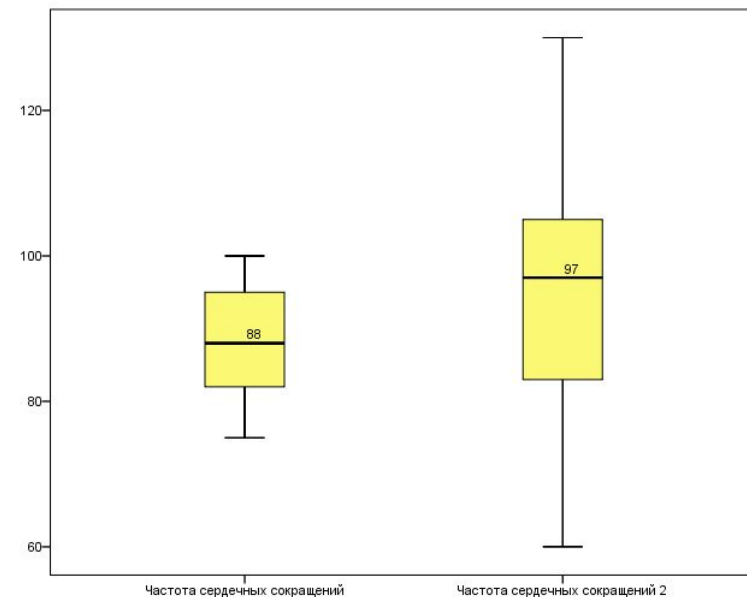
### Статистики

		Частота сердечных сокращений	Частота сердечных сокращений 2
N	Валидные	50	50
	Пропущенные	25	25
Среднее		88,32	94,26
Стд. ошибка среднего		1,048	1,958
Медиана		88,00	97,00
Мода		86	79 <sup>a</sup>
Стд. отклонение		7,408	13,844
Асимметрия		-,124	-,035
Стд. ошибка асимметрии		,337	,337
Экссесс		-1,140	-,113
Стд. ошибка эксцесса		,662	,662
Размах		25	70
Минимум		75	60
Максимум		100	130
Процентили	25	81,50	83,00
	50	88,00	97,00
	75	95,00	105,00

a. Имеется несколько мод. Показана наименьшая.

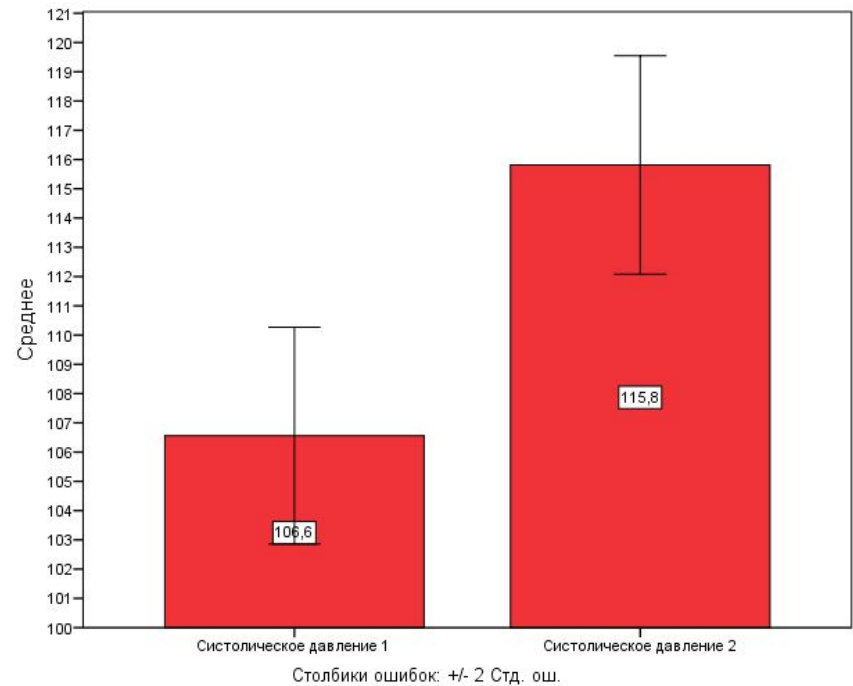
### Вывод:

- 1) ЧСС в положении стоя у первоклассников выше, чем в положении лежа (см. медиану в ящиках и ср. значение).
- 2) Есть дети, у которых ЧСС повысилось, а есть дети, у которых понизилось. Это характерно как для среднеуспевающих детей (ящик растянулся), так и для крайних значений.
- 3) Асимметрия была отрицательна, а стала ок. нуля, следовательно, изначально преобладало число детей с частым пульсом, а потом оно уменьшилось.
- 4) Экссесс был резко отрицательным, а затем значительно снизился.



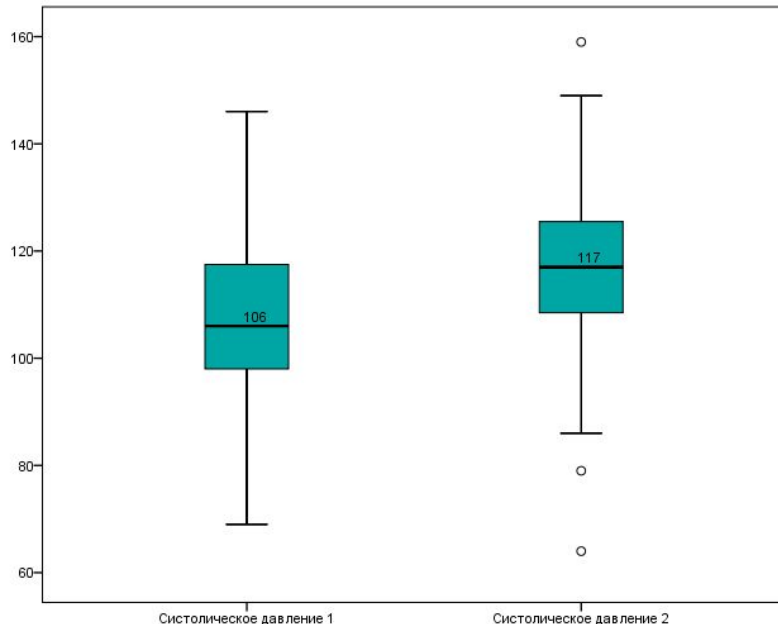
## Статистики

	Систолическое давление 1	Систолическое давление 2
Среднее	106,56	115,81
Стд. ошибка среднего	1,853	1,866
Медиана	106,00	117,00
Мода	105	120
Стд. отклонение	16,045	16,160
Асимметрия	-,018	-,324
Стд. ошибка асимметрии	,277	,277
Эксцесс	,295	1,095
Стд. ошибка эксцесса	,548	,548
Минимум	69	64
Максимум	146	159
Перцентили		
25	98,00	108,00
50	106,00	117,00
75	118,00	126,00



### Вывод:

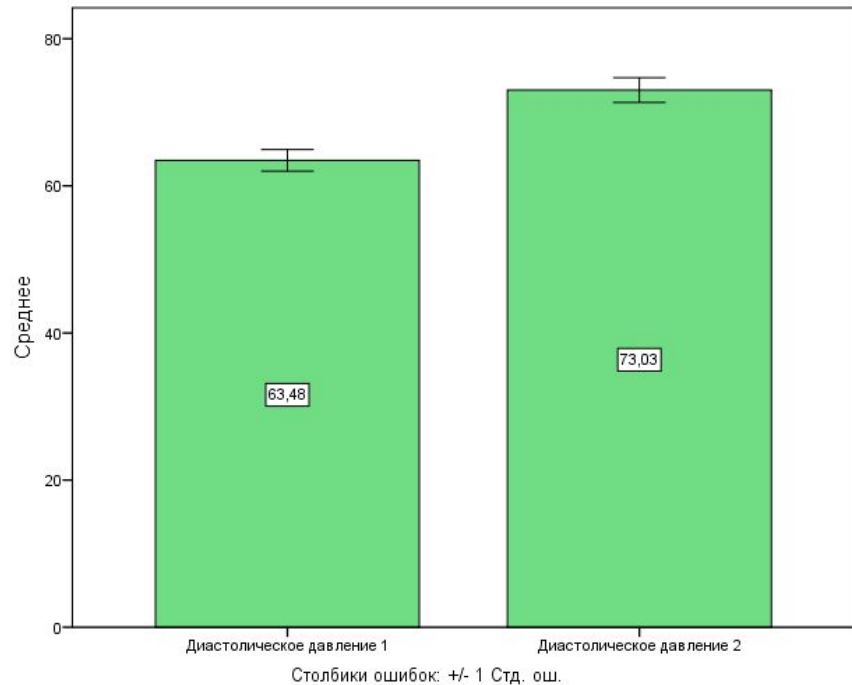
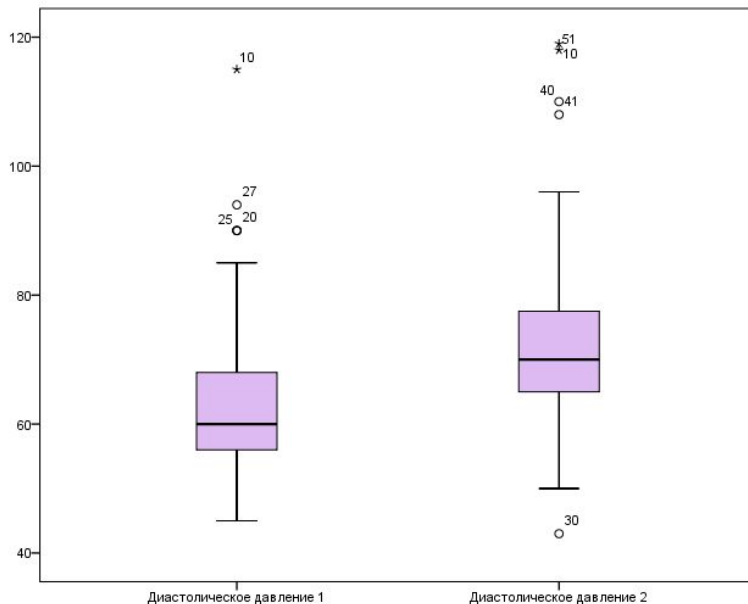
- 1) Систолическое давление в покое (СД1) меньше, чем систолическое давление в положении стоя (СД2) – см. среднее значение.
- 2) Среднее значение систолического давления в положении стоя увеличилось (смещение ящика вверх).
- 3) Асимметрия была приблизительно равна 0, а стала отрицательной, т.е. изначально преобладали дети с пониженным систолическим давлением, а потом стали преобладать дети с повышенным систолическим давлением.
- 4) Эксцесс был приблизительно равен нулю, а стал больше единицы.



## Статистики

	Диастолическое давление 1	Диастолическое давление 2
Среднее	63,48	73,03
Стд. ошибка среднего	1,469	1,685
Медиана	60,00	70,00
Мода	60	60 <sup>a</sup>
Стд. отклонение	12,726	14,589
Асимметрия	1,434	1,166
Стд. ошибка асимметрии	,277	,277
Эксцесс	2,923	1,830
Стд. ошибка эксцесса	,548	,548
Минимум	45	43
Максимум	115	119
Процентили		
25	56,00	65,00
50	60,00	70,00
75	68,00	78,00

а. Имеется несколько мод. Показана наименьшая.



## Вывод:

- 1) Диастолическое давление в покое (ДД1) меньше, чем диастолическое давление в положении стоя (ДД2) – см. среднее значение.
- 2) Среднее значение систолического давления в положении стоя увеличилось (смещение ящика вверх).
- 3) Ассиметрия ДД2 уменьшилась по сравнению с асимметрией ДД1, т.е. изначально преобладали дети с повышенным диастолическим давлением, а потом стали преобладать дети с пониженным систолическим давлением.
- 4) Эксцесс уменьшился.

# Критерии сравнения

# Одновыборочные критерии



t-Стюдента

(нормальное распр,  $|z| < 3$ )

z-Вилкоксона

(ненормальное распр,  $|z| > 3$ )

**Задача:** среднее ( $X_{\text{ср}}$ ;  $Me$ ) =  $a$

1.  $H_0$ : среднее ( $X_{\text{ср}}$ ;  $Me$ ) =  $a$
2. Нормальность, гр.наблюдения + выбор критерия
3. Расчёт значений
4. Вывод о  $H_0$
5. Уточнение и интерпретация результатов

# Задача: Определить, отличается ли средний балл в 10 классе от 4 (4,5)

1. Но: Средний балл в 10 классе (отметка1) равен 4 (4,5)
2. Проверка на нормальность и наличие грубых наблюдений

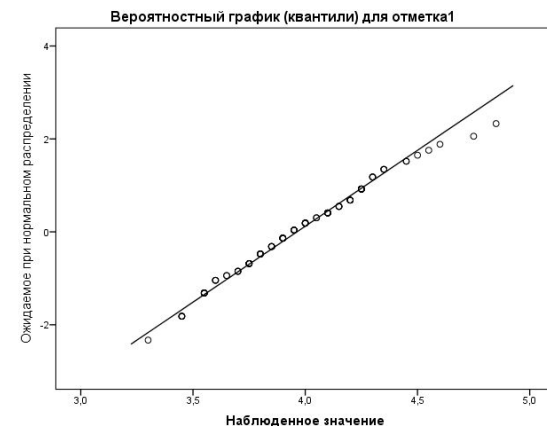
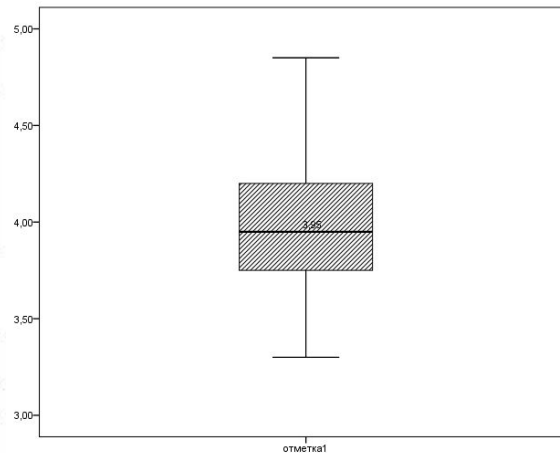
Описательные

		Статистика	Стд. ошибка
отметка1	Среднее	3,9630	,03060
	Медиана	3,9500	
	Асимметрия	,317	,241
	Экссесс	-,056	,478

Критерий нормальности

	Шапиро-Уилк		
	Статистика	ст. св.	Значимость
отметка1	,984	100	,258

а. Поправка значимости Лильефорса



Распределение «отметка1» **нормальное, грубых наблюдений нет** (все  $|z| < 3$ )  
Выборка одна, следовательно, применяем **одновыборочный критерий t-Стьюдента**



### 3. Вычисление значений критерия.

#### 1) Для значения 4

Статистики для одновыборочного t-критерия

	N	Среднее	Стд. отклонение	Стд. ошибка среднего
отметка1	100	3,9630	,30597	,03060

Одновыборочный t-критерий

	Проверяемое значение = 4		
	t	ст.св.	Значимость (2-сторонняя)
отметка1	-1,209	99	,229

Значимость критерия больше, чем 0,1, следовательно,  $H_0$  подтверждается, средний балл в 10 классе равен 4.

#### 2) Для значения 4.5

Статистики для одновыборочного t-критерия

	N	Среднее	Стд. отклонение	Стд. ошибка среднего
отметка1	100	3,9630	,30597	,03060

Одновыборочный t-критерий

	Проверяемое значение = 4.5		
	t	ст.св.	Значимость (2-сторонняя)
отметка1	-17,551	99	3,712E-032

Значимость критерия меньше, чем 0,05, следовательно,  $H_0$  не подтверждается, средний балл в 10 классе не равен 4,5 и отклоняется в меньшую сторону.

### Итоги по проверке гипотезы

	Нулевая гипотеза	Критерий	Значимость	Решение
1	Медиана счет в уме равна 10,50.	Одновыборочный критерий знаковых рангов Вилкоксона	,378	Нулевая гипотеза принимается.

Выводятся асимптотические значимости. Уровень значимости равен ,05.

### Итоги по проверке гипотезы

	Нулевая гипотеза	Критерий	Значимость	Решение
1	Медиана счет в уме равна 11,00.	Одновыборочный критерий знаковых рангов Вилкоксона	,020	Нулевая гипотеза отклоняется.

Выводятся асимптотические значимости. Уровень значимости равен ,05.

# Задача

Определить, отличается ли средний рост учащихся 1 классов от 123 см?

- 1)  $H_0$  – рост учащихся 1 классов не отличается от 123 см.
- 2) Проверка на нормальность

## Описательные

		Статистика	Стд. ошибка
Рост	Среднее	123,47	,585
	Медиана	123,50	
	Асимметрия	-,034	,277
	Эксцесс	,517	,548

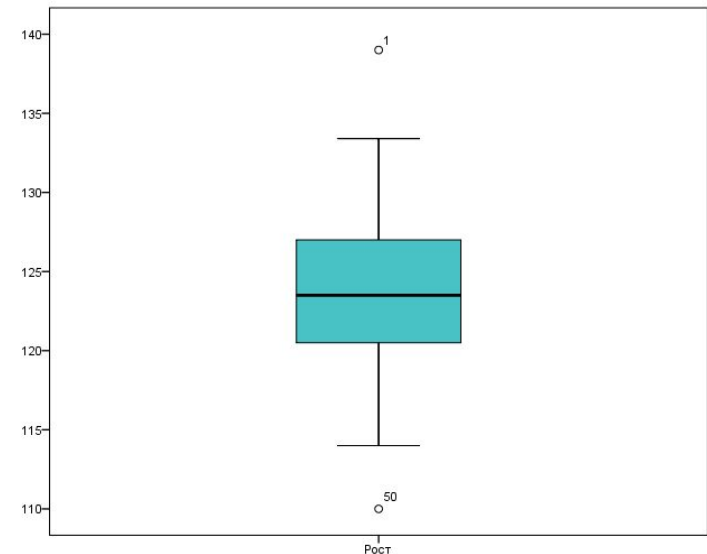
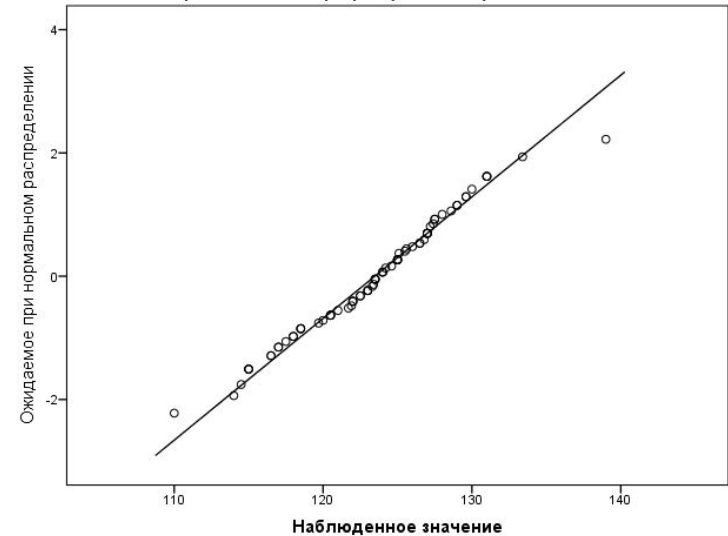
## Критерий нормальности

	Шапиро-Уилк		
	Статистика	ст. св.	Значимость
Рост	,986	75	,592

а. Поправка значимости Лильефорса

**Вывод:** распределение переменной «отметка 1» нормальное. Грубых набл. нет, все z-значения по модулю меньше 3. Т.к. выборка нормальная, выбираем одновыборочный t-критерий Стьюдента.

Вероятностный график (квантили) для Рост



### 3) Вычисление значения критерия

**Одновыборочный t-критерий**

	Проверяемое значение = 123		
	t	ст.св.	Значимость (2- сторонняя)
Рост	,811	74	,420

**Статистики для одновыборочного t-критерия**

	N	Среднее	Стд. отклонение	Стд. ошибка среднего
Рост	75	123,47	5,067	,585

4)  $H_0$  верна. Средний рост учащихся в 1 классах не отличается от 123 см.

# Две независимые выборки

*Но, 0,05: Гипотеза о равенстве двух распределений*

## Алгоритм

1. Проверка на нормальность и наличие грубых наблюдений
2. Выбор критерия/критериев
3. Вычисление результатов критерия/критериев
4. Вывод о справедливости  $H_0$
5. Интерпретация полученного результата
6. Иллюстрация полученного результата (о различиях) на столбиковых диаграммах.

# Задача: Сравнить успеваемость девочек и юношей в 10 классе

*Но: Успеваемость девочек и юношей в 10 классе одинаковая*

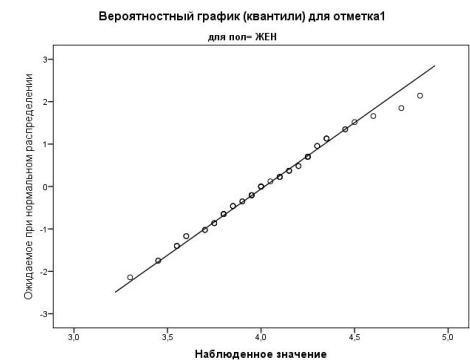
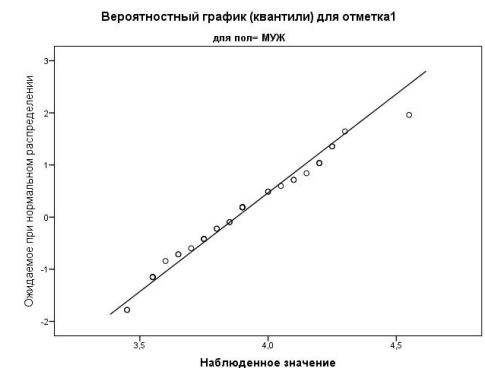
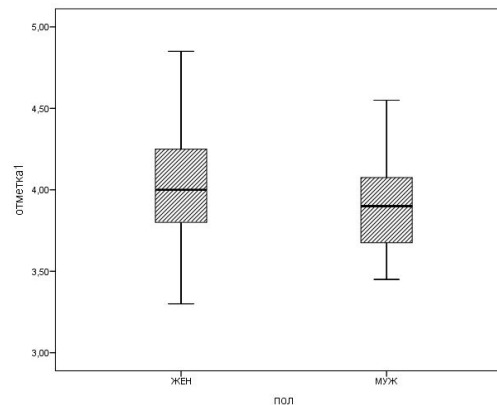
## Описательные

пол		Статистика	Стд. ошибка
отметка1	ЖЕН	Среднее	4,0180
		Медиана	4,0000
		Стд. отклонение	,32003
		Асимметрия	,167
		Эксцесс	,016
МУЖ	Среднее	3,8769	,04228
		Дисперсия	,070
		Стд. отклонение	,26404
		Асимметрия	,373
		Эксцесс	-,352

## Критерий нормальности

пол		Шапиро-Уилк		
		Статистика	ст. св.	Значимость
отметка1	ЖЕН	,990	61	,895
	МУЖ	,966	39	,272

а. Поправка значимости Лильефорса



Оба распределения по отметке 1 нормальные, грубых наблюдений нет (все  $|z| < 3$ )  
 Выбираем t-Стьюдента для независимых выборок, но еще проверим равенство дисперсий

**Групповые статистики**

	пол	N	Среднее	Стд. отклонение	Стд. ошибка среднего
отметка1	ЖЕН	61	4,0180	,32003	,04098
	МУЖ	39	3,8769	,26404	,04228

**Критерий для независимых выборок**

		Критерий равенства дисперсий Левина	
		F	Знач.
отметка1	Предполагается равенство дисперсий	1,516	,221

Дисперсии равны ( $F=1,5$ ;  $p=0,221 > 0,1$ ) Решаем задачу, используя t-критерий Стьюдента для независимых выборок.

**3. Вычисление результатов критерия**

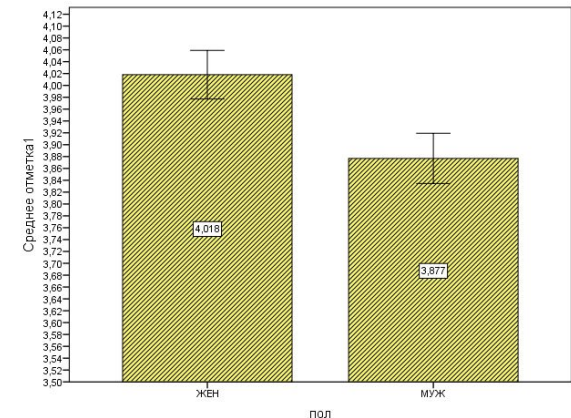
**рок**

t-критерий равенства средних		
t	ст.св.	Значимость (2-сторонняя)
2,298	98	,024

4. Но не верна, т.к.  $p < 0,05$ . Средние арифметические отличаются.

5. Успеваемость юношей и девушек в 10 классе различается.

6. Девушки в 10 классе учатся лучше, чем юноши.



Столбики ошибок: +/- 1 Стд. ош.

**Групповые статистики**

	пол	N	Среднее	Стд. отклонение	Стд. ошибка среднего
отметка1	ЖЕН	61	4,0180	,32003	,04098
	МУЖ	39	3,8769	,26404	,04228

# Решим задачу другим способом

Так как выборки малые по численности, то лучше использовать для решения критерий U-Манна-Уитни.

Ранги

пол	N	Средний ранг	Сумма рангов
ЖЕН	61	55,86	3407,50
МУЖ	39	42,12	1642,50
Всего	100		

4. Вывод: Но не верна, т.к.  $p < 0,05$ .

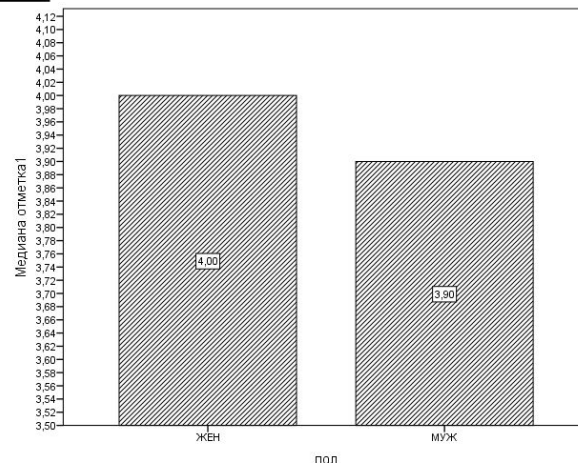
5. Успеваемость девушек и юношей в 10 классе различается.

Статистики критерия<sup>a</sup>

	отметка1
Статистика U Манна-Уитни	862,500
Асимпт. знч. (двухсторонняя)	,021

6. Девушки в 10 классе учатся лучше, чем юноши.

а. Группирующая переменная: пол



Процентили (квантили)

		Процентили (квантили)		
		пол		
		25	50	75
Взвешенное среднее (Определение 1)	ЖЕН	3,8000	4,0000	4,2500
	МУЖ	3,6500	3,9000	4,1000



# Задача: Одинаковы ли показатели силы правой и левой рук и мальчиков и девочек

*Но : Показатели силы правой и левой рук у мальчиков и девочек одинаковы.*

Описательные статистики

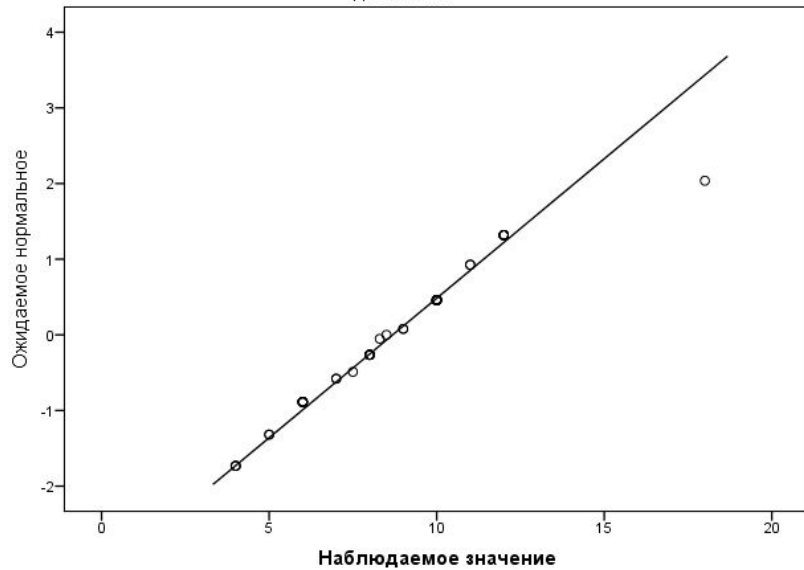
Пол			Статистика	Стандартная Ошибка
Сила кисти правой руки	М	Среднее значение	8,69	,396
		Медиана	8,50	
		Асимметрия	,591	,347
		Эксцесс	1,618	,681
	Ж	Среднее значение	7,02	,489
		Медиана	7,00	
		Асимметрия	-,405	,441
		Эксцесс	-,827	,858
Сила кисти левой руки	М	Среднее значение	7,68	,352
		Медиана	8,00	
		Асимметрия	,719	,347
		Эксцесс	,338	,681
	Ж	Среднее значение	6,13	,464
		Медиана	6,50	
		Асимметрия	-,335	,441
		Эксцесс	-,897	,858

Критерии нормального распределения

	Пол	Критерий Шапиро-Уилка		
		Статистика	ст.св.	Значимость
Сила кисти правой руки	М	,936	47	,013
	Ж	,898	28	,011
Сила кисти левой руки	М	,934	47	,011
	Ж	,930	28	,060

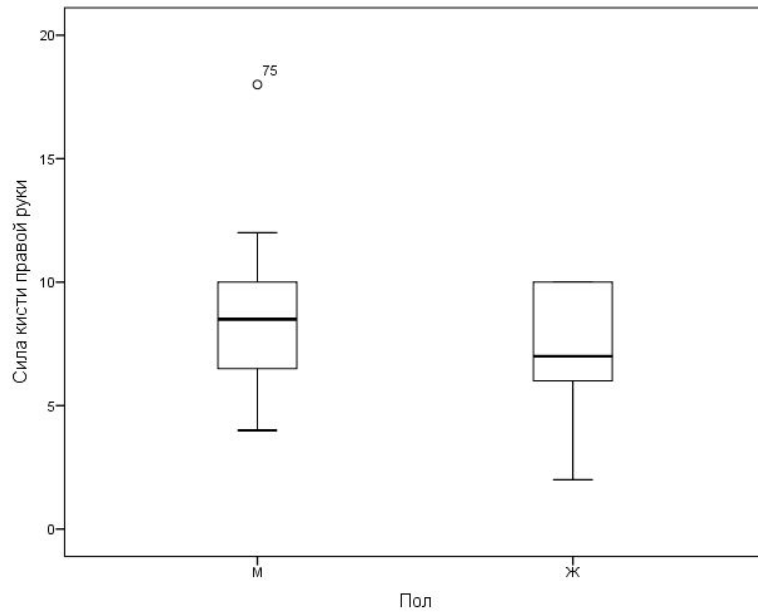
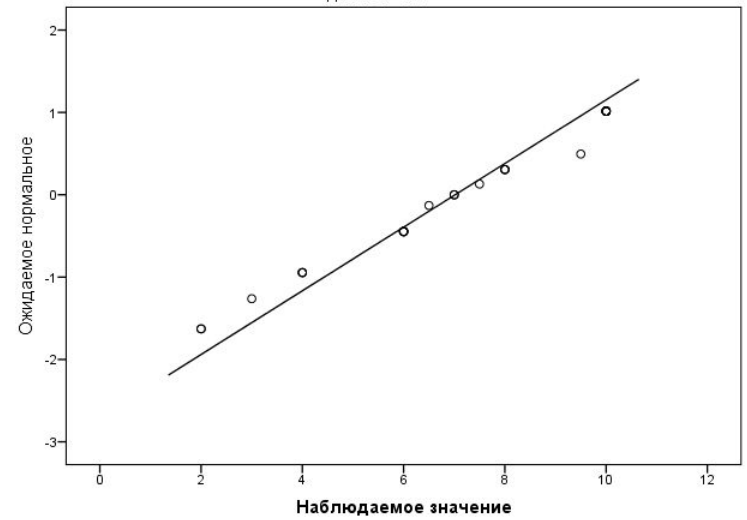
Нормальный график Q-Q Сила кисти правой руки

для Пол= М



Нормальный график Q-Q Сила кисти правой руки

для Пол= Ж

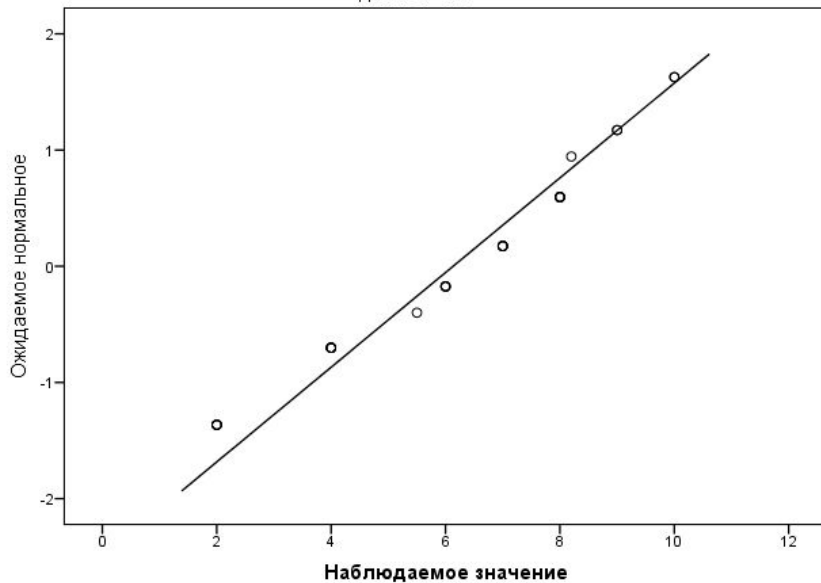


Распределения силы кисти правой руки мальчиков и девочек ненормальные (критерий, медиана (дев) на ящике смещена), есть грубое наблюдение  $|z| > 3$  (звезд нет).

***Критерий U-Манна-Уитни***

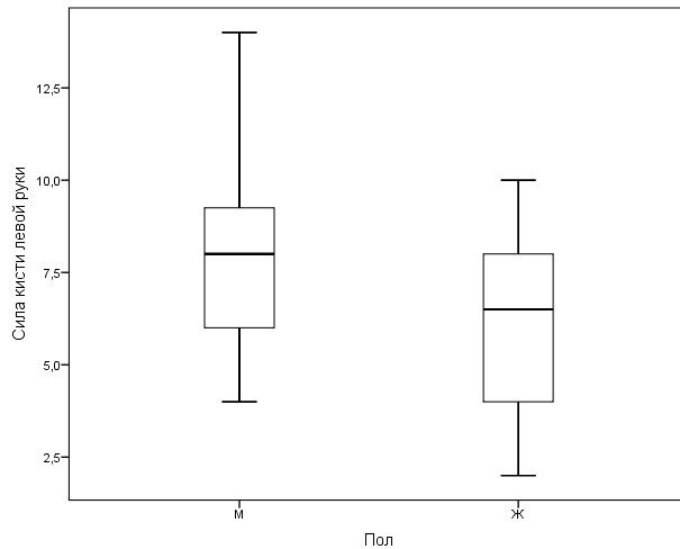
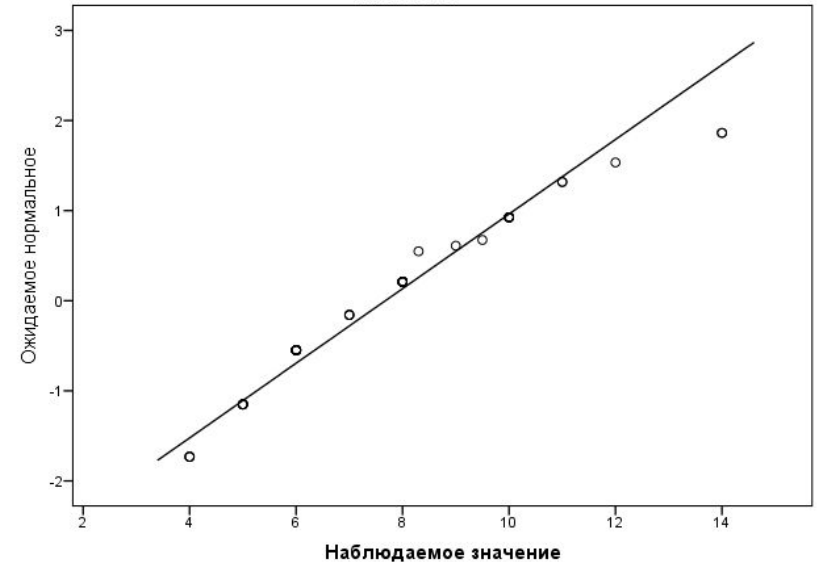
Нормальный график Q-Q Сила кисти левой руки

для Пол= Ж



Нормальный график Q-Q Сила кисти левой руки

для Пол= М



Распределения силы кисти левой руки мальчиков и девочек ненормальные (критерий, медианы на ящиках смещены), грубых наблюдений нет, все  $|z| < 3$ . **Критерий U-Манна-Уитни**

### Ряды

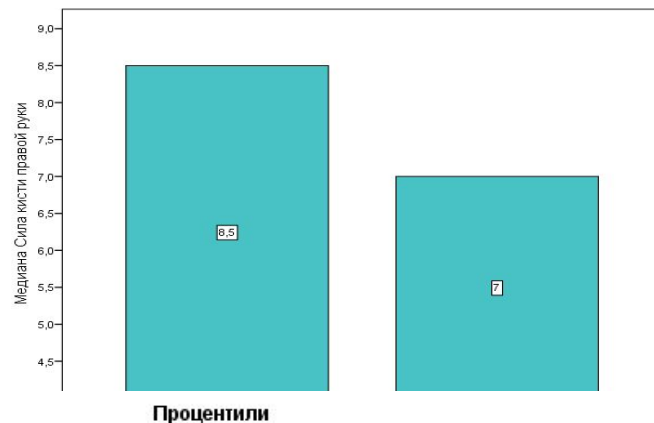
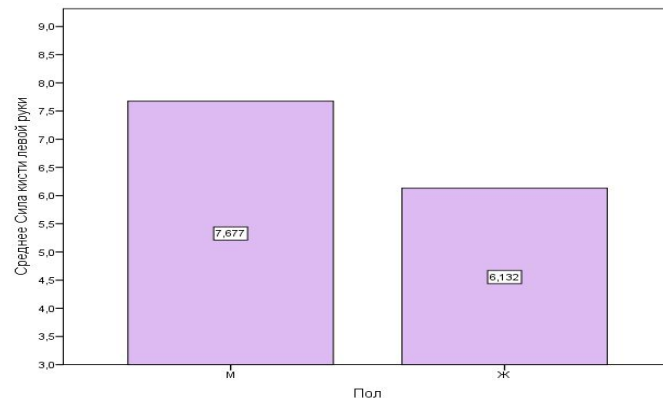
	Пол	N	Средний ранг	Сумма рангов
Сила кисти правой руки	М	47	42,52	1998,50
	Ж	28	30,41	851,50
	Всего	75		
Сила кисти левой руки	М	47	42,02	1975,00
	Ж	28	31,25	875,00
	Всего	75		

### Статистические критерии<sup>а</sup>

	Сила кисти правой руки	Сила кисти левой руки
U Манна-Уитни	445,500	469,000
Асимптотическая значимость (2-сторонняя)	,018	,036

а. Группирующая переменная: Пол

- **Вывод о H<sub>0</sub>:** H<sub>0</sub> не верна, т.к.  $p(\text{сила пр. руки}) < 0,05$ ,  $p(\text{сила л. Руки}) < 0,05$ .
- Сила кисти правой и левой рук у мальчиков и девочек отличаются.
- Показатель силы кистей обеих рук у мальчик выше, чем у девочек.



	Пол	Процентили			
		25	50	75	
Взвешенное среднее (Определение 1)	Сила кисти правой руки	М	6,00	8,50	10,00
		Ж	6,00	7,00	10,00
	Сила кисти левой руки	М	6,00	8,00	9,50
		Ж	4,00	6,50	8,00

# Критерии для двух зависимых выборок

**Задача:** Сравнить успеваемость учащихся в 10 и 11 классе

**Но:** Успеваемость учащихся в 10 и 11 классе одинаковая

1. Проверка наблюдений на нормальность и наличие грубых наблюдений

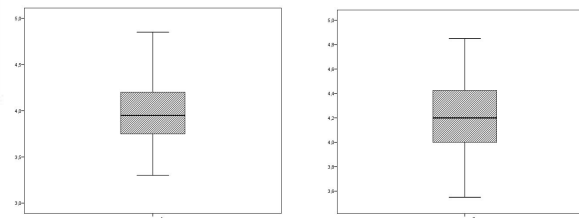
Описательные

	Статистика	Стд. ошибка
отметка1	Среднее	3,9630
	Медиана	3,9500
	Асимметрия	,317
	Эксцесс	-,056
отметка2	Среднее	4,2205
	Медиана	4,2000
	Асимметрия	-,007
	Эксцесс	-,357

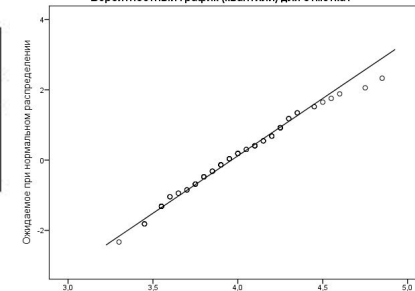
Критерий нормальности

	Шапиро-Уилк		
	Статистика	ст. св.	Значимость
отметка1	,984	100	,258
отметка2	,986	100	,397

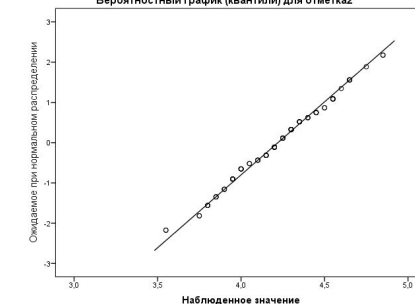
а. Поправка значимости Лильефорса



Вероятностный график (квантили) для отметка1

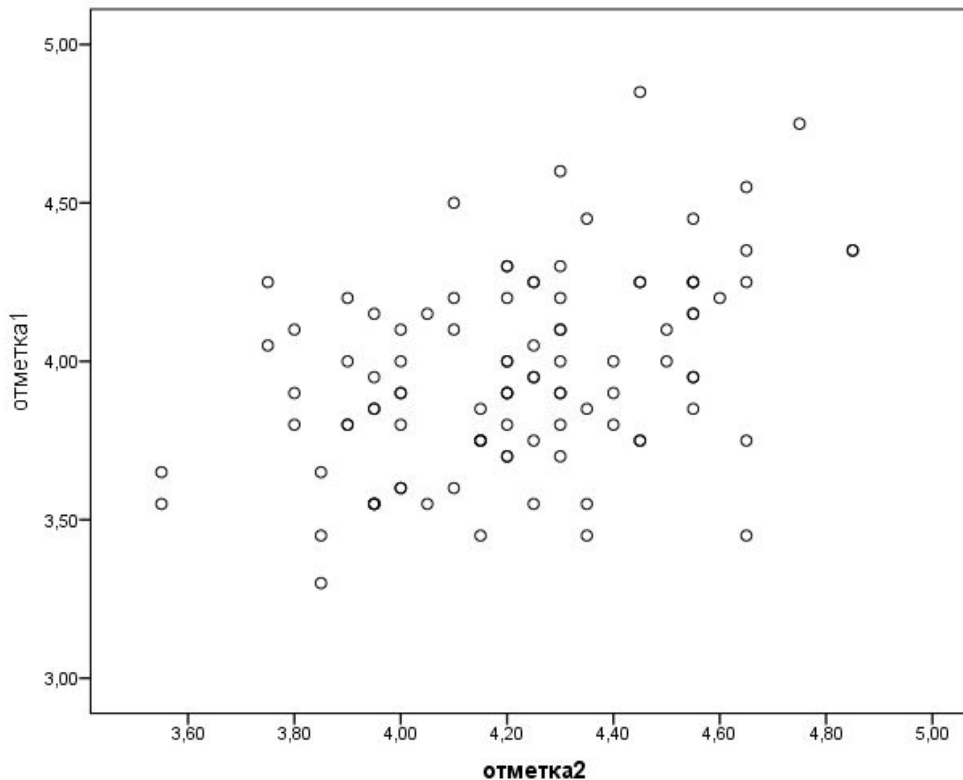


Вероятностный график (квантили) для отметка2



Оба распределения нормальные, грубых наблюдений нет (все  $|Z| < 3$ ).

# Доказательство зависимости



Корреляции

		отметка2
отметка1	Корреляция Пирсона	,434
	Знч.(2-сторон)	,000006
	N	100

Связь линейная, прямая, умеренная.

3. Так как оба распределения нормальные, без грубых наблюдений, между переменными линейная, прямая, умеренная связь, то выбираем парный t-Критерий Стьюдента.

#### 4. Вычисление результатов критерия

Статистики парных выборок

	Среднее	N	Стд. отклонение	Стд. ошибка среднего
Пара 1    отметка1	3,9630	100	,30597	,03060
отметка2	4,2205	100	,27589	,02759

Критерий парных выборок

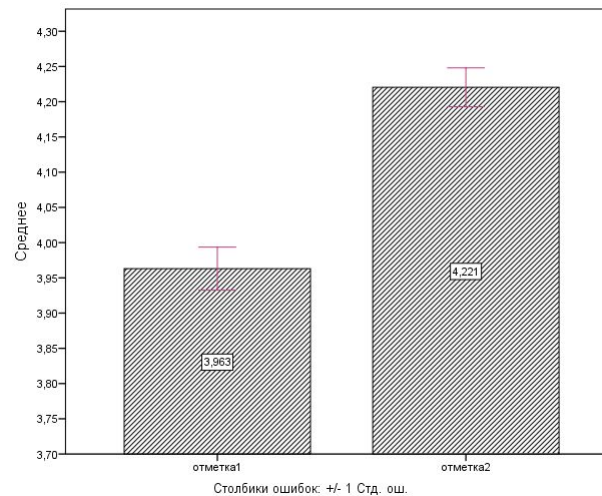
	t	ст.св.	Значимость (2-сторонняя)
Пара 1    отметка1 - отметка2	-8,290	99	5,720E-013

5. Вывод об  $H_0$ :  $H_0$  не верна ( $p < 0,05$ )

6. Интерпретация результатов:

Успеваемость в 10 и 11 классе неодинаковая.

7. В 11 классе дети учатся лучше, чем в 10 классе.



Статистики парных выборок

	Среднее	N	Стд. отклонение	Стд. ошибка среднего
Пара 1    отметка1	3,9630	100	,30597	,03060
отметка2	4,2205	100	,27589	,02759

# Задача: Сравнить результаты теста 1 и теста 2

## теста 1 и теста 2

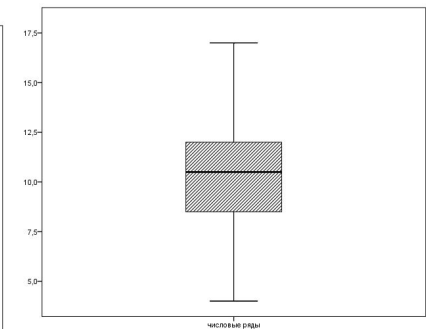
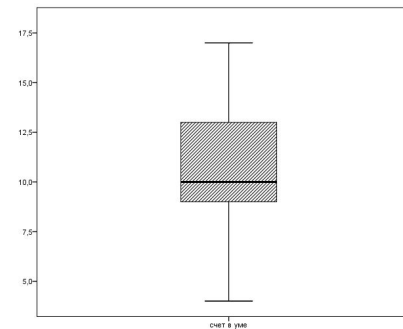
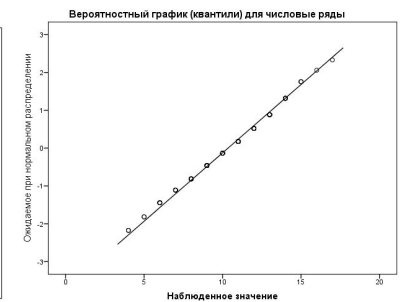
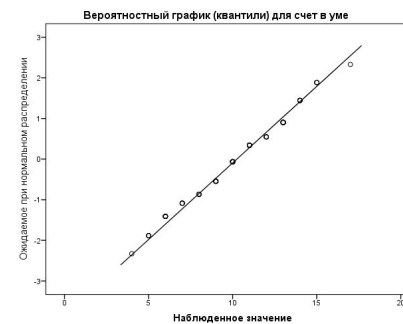
**Но:** Результаты теста 1 и теста 2 одинаковые

1. Проверка наблюдений на нормальность и наличие грубых наблюдений

Описательные			
		Статистика	Стд. ошибка
счет в уме	Среднее	10,25	,265
	Медиана	10,00	
	Асимметрия	-,093	,241
	Эксцесс	-,363	,478
числовые ряды	Среднее	10,35	,277
	Медиана	10,50	
	Асимметрия	-,127	,241
	Эксцесс	-,435	,478

Критерий нормальности			
	Шапиро-Уилк		
	Статистика	ст. св.	Значимость
счет в уме	,972	100	,031
числовые ряды	,983	100	,216

а. Поправка значимости Лильефорса



Распределение Тест 1 не нормальное (критерий, медиана на ящике).

Распределение Тест 2 нормальное.

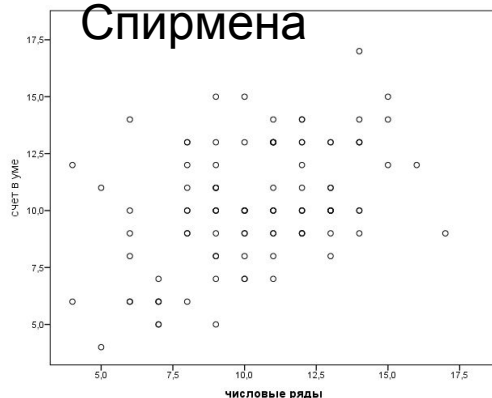
Грубых наблюдений нет (все  $|Z| < 3$ ).



# Доказательство зависимости

Т.к. одно из распределений ненормальное, используем ро-

Спирмена



Корреляции

Корреляции			числовые ряды
ро Спирмена	счет в уме	Кoeffициент корреляции	,401
		Знч. (2-сторон)	3,556E-005
		N	100

Связь между переменными линейная, прямая, умеренная.

Т.к. одно из распределений ненормальное, используем критерий Вилкоксона

## 4. Вычисление результатов критерия

Ранги

		N	Средний ранг	Сумма рангов
числовые ряды - счет в уме	Отрицательные ранги	37 <sup>a</sup>	46,86	1734,00
	Положительные ранги	49 <sup>b</sup>	40,96	2007,00
	Связи	14 <sup>c</sup>		
	Всего	100		

a. числовые ряды < счет в уме

b. числовые ряды > счет в уме

c. числовые ряды = счет в уме

Статистики критерия<sup>a</sup>

	числовые ряды - счет в уме
Z	-,592 <sup>b</sup>
Асимпт. знч. (двухсторонняя)	,554

a. Критерий знаковых рангов Уилкоксона

b. Используются отрицательные ранги.

5. Вывод об  $H_0$ :  $H_0$  верна ( $p > 0,1$ )

6. Результаты для теста 1 и теста 2 одинаковые.

Процентили (квантили)

		Процентили (квантили)		
		25	50	75
Взвешенное среднее (Определение 1)	счет в уме	9,00	10,00	13,00
	числовые ряды	8,25	10,50	12,00

# Задача: Отличается ли давление учащихся до и после физ. нагрузки.

Но: Давление учащихся до и после физ. нагрузки одинаковое

## Описательные

		Статистика	Стд. ошибка
Диастолическое давление лежа	Среднее	62,61	1,325
	Медиана	60,00	
	Асимметрия	,971	,283
	Эксцесс	,702	,559
Диастолическое давление стоя	Среднее	71,25	1,400
	Медиана	70,00	
	Асимметрия	,655	,283
	Эксцесс	,900	,559

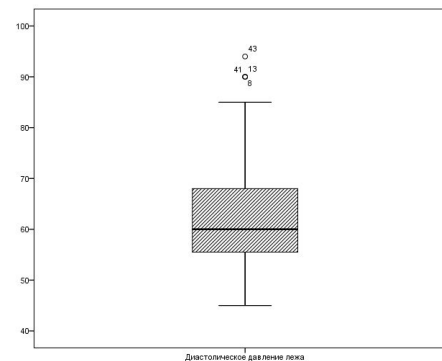
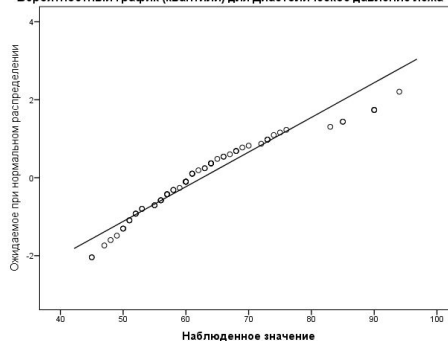
## Критерий нормальности

	Шапиро-Уилк		
	Статистика	ст. св.	Значимость
Диастолическое давление лежа	,926	72	4,123E-004
Диастолическое давление стоя	,960	72	,022

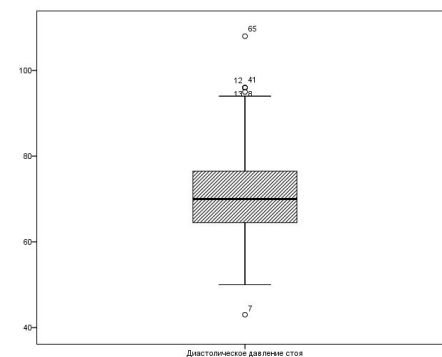
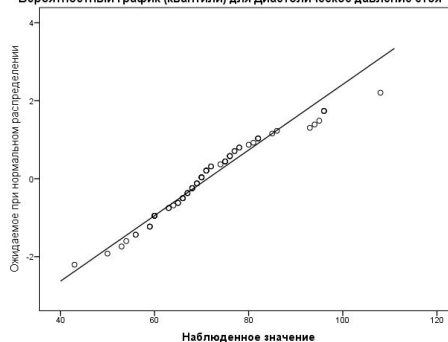
а. Поправка значимости Лиллиефорса

Распределения ДД1 и ДД2 ненормальные (критерий Шапиро-Уилка (ДД1)).

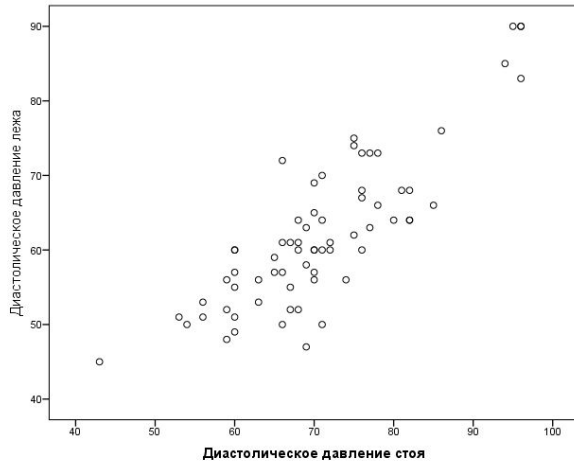
Вероятностный график (квантили) для Диастолическое давление лежа



Вероятностный график (квантили) для Диастолическое давление стоя



## Доказательство зависимости



Т.к. распределения ненормальные, используем  $\rho$ -Сг

Корреляции			Диастолическое давление стоя
$\rho$ Спирмена	Диастолическое давление лежа	Коэффициент корреляции	,795
		Знач. (2-сторон)	5,785E-016
		N	68

Связь между переменными линейная, прямая, значительная.

Т.к. распределения ненормальные, используем критерий Вилкоксона.

### Ранги

		N	Средний ранг	Сумма рангов
Диастолическое давление стоя - Диастолическое давление лежа	Отрицательные ранги	2 <sup>a</sup>	13,00	26,00
	Положительные ранги	63 <sup>b</sup>	33,63	2119,00
	Связи	3 <sup>c</sup>		
	Всего	68		

a. Диастолическое давление стоя < Диастолическое давление лежа

b. Диастолическое давление стоя > Диастолическое давление лежа

c. Диастолическое давление стоя = Диастолическое давление лежа

### Статистики критерия<sup>a</sup>

	Диастолическое давление стоя - Диастолическое давление лежа
Z	-6,842 <sup>b</sup>
Асимпт. знч. (двухсторонняя)	,000

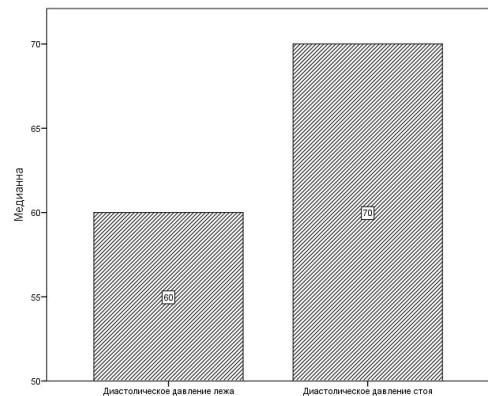
a. Критерий знаковых рангов Уилкоксона

b. Используются отрицательные ранги.

5. Вывод об H<sub>0</sub>: не верна (p>0,1)

6. Давление учащих до и после физ.нагрузки разное.

Вывод: давление до физ. нагрузки меньше, чем давление после физ. нагрузки.



### Процентили (квантили)

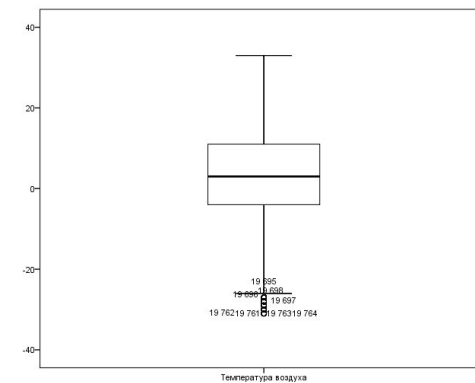
		Процентили (квантили)		
		25	50	75
Взвешенное среднее (Определение 1)	Диастолическое давление лежа	55,25	60,00	67,75
	Диастолическое давление стоя	65,00	70,00	76,00

# Задача: Отличаются ли друг от друга температура воздуха и точка росы

Но: Температура воздуха и точка росы не отличаются.

## Описательные статистики

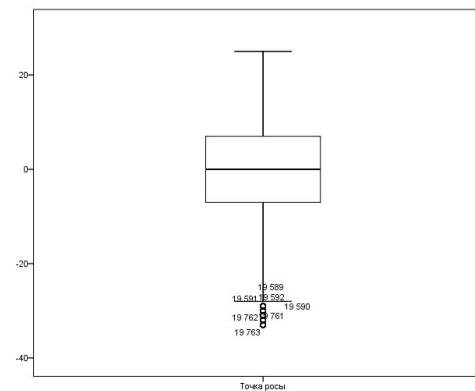
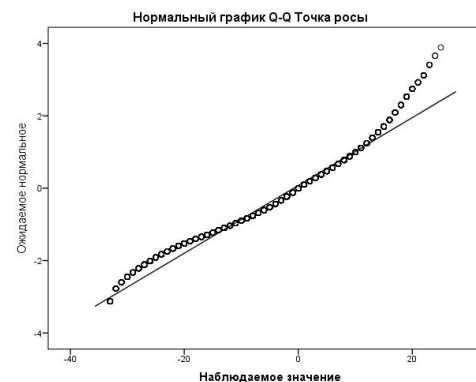
		Статистика	Стандартная Ошибка
Температура воздуха	Среднее значение	2,65	,084
	Медиана	3,00	
	Асимметрия	-,332	,017
	Экссесс	-,222	,035
Точка росы	Среднее значение	-,82	,076
	Медиана	0,00E+000	
	Асимметрия	-,567	,017
	Экссесс	,004	,035



## Критерии нормального распределения

	Колмогорова-Смирнова <sup>а</sup>		
	Статистика	ст.св.	Значимость
Температура воздуха	,053	19762	9,932E-147
Точка росы	,070	19762	2,574E-262

а. Коррекция значимости Лильефорса

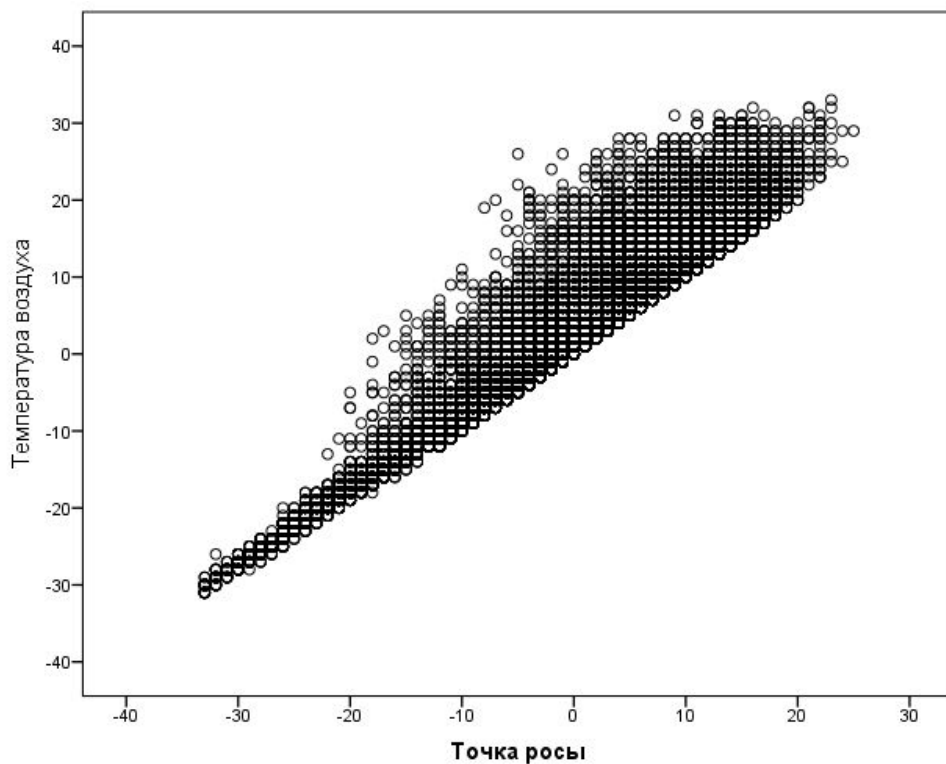


Оба распределения не являются нормальными (критерий Колмогорова-Смирнова). Т.к. распределения ненормальные, для доказательства зависимости используем ро-Спирмена.

# Доказательство зависимости

## Корреляции

			Точка росы
Ро Спирмана	Температура воздуха	Кoeffициент корреляции	,942
		Знач. (2-х сторонняя)	0,000E+000
		N	19762



Связь между переменными линейная, прямая, очень высокая. Т.к. распределения ненормальные, используем критерий *Вилкоксона*.

# Вычисление результатов критерия

Ряды

		N	Средний ранг	Сумма рангов
Точка росы - Температура воздуха	Отрицательные ранги	17299 <sup>a</sup>	8650,00	149636350,0
	Положительные ранги	0 <sup>b</sup>	,00	,00
	Совпадающие наблюдения	2463 <sup>c</sup>		
	Всего	19762		

- a. Точка росы < Температура воздуха
- b. Точка росы > Температура воздуха
- c. Точка росы = Температура воздуха

Статистические критерии<sup>a</sup>

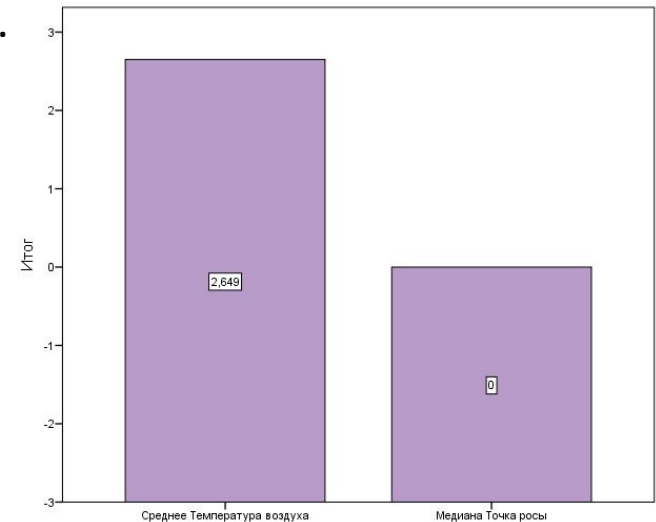
	Точка росы - Температура воздуха
Z	-114,396 <sup>b</sup>
Асимптотическая значимость (2- сторонняя)	0,000E+000

- a. Критерий знаковых рангов Вилкоксона
- b. На основе положительных рангов.

Но не верна.

Температура воздуха и точка росы отличаются друг от друга.

Температура воздуха значительно выше, чем точка росы.



Описательные статистики

	N	Процентили		
		25	50-я (медиана)	75-я
Температура воздуха	19762	-4,00	3,00	11,00
Точка росы	19763	-7,00	,00	7,00



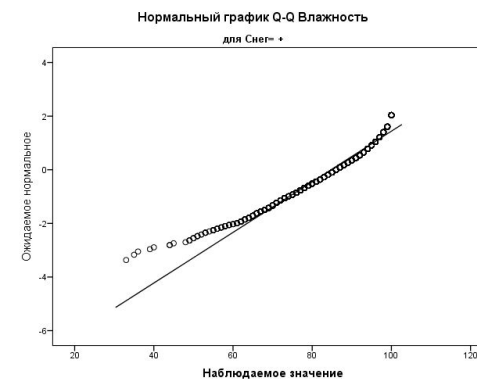
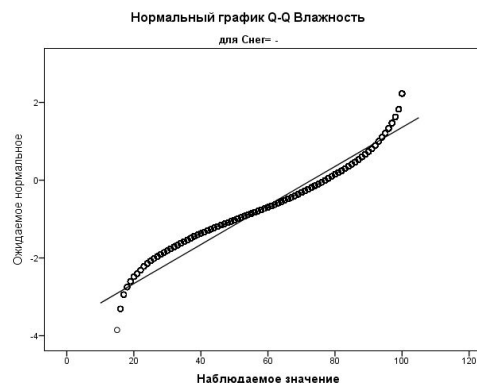
# Задача: Отличаются ли друг от друга влажность в дни, когда идет снег и влажность в дни, когда снег не идет.

## НЕ ИДЕТ.

Но: Влажность при наличии снега и при его отсутствии не отличается.

Описательные статистики

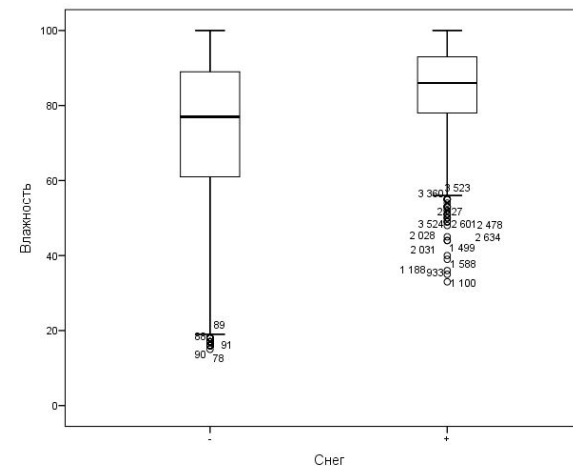
Снег		Статистика	Стандартная Ошибка
Влажность -	Среднее значение	72,97	,152
	Медиана	77,00	
	Асимметрия	-,747	,019
	Эксцесс	-,249	,037
Влажность +	Среднее значение	84,74	,207
	Медиана	86,00	
	Асимметрия	-,815	,048
	Эксцесс	,789	,096



Критерии нормального распределения

Снег	Колмогорова-Смирнова <sup>a</sup>		
	Статистика	ст.св.	Значимость
Влажность -	,094	17137	0,000E+000
Влажность +	,075	2612	7,054E-039

a. Коррекция значимости Лилiefорса



Оба распределения ненормальные (критерий Колмогорова-Смирнова). Грубых наблюдений нет, все  $|z| < 3$ .  
Критерий U-Манна-Уитни.

# Вычисление результатов критерия

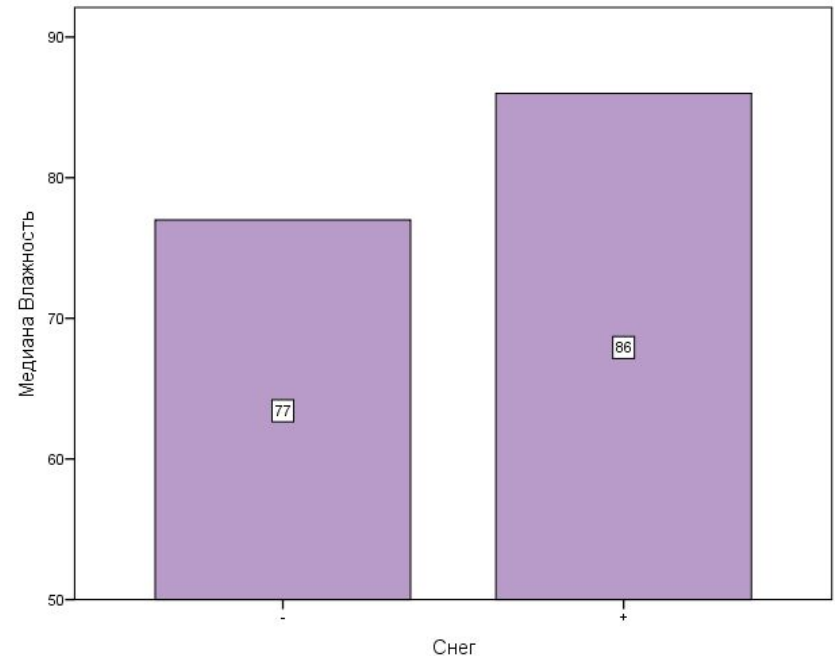
Ряды

Снег	N	Средний ранг	Сумма рангов
Влажность -	17137	9431,12	161621129,0
+	2612	12787,23	33400246,00
Всего	19749		

Статистические критерии<sup>a</sup>

	Влажность
U Манна-Уитни	14774176,00
Асимптотическая значимость (2-сторонняя)	6,901E-173

a. Группирующая переменная: Снег



Но не верна.

Показатели влажности при отсутствии и присутствии снега различаются.

Влажность в дни, когда шел снег выше, чем влажность в дни, когда

Описательные статистики

	N	Процентили		
		25	50-я (медиана)	75-я
Влажность	19749	63,00	79,00	90,00
Снег	19831	1,00	1,00	1,00

# Задача: Отличается ли давление от 1000 МПа

Но: Медиана давления равна 1000 МПа

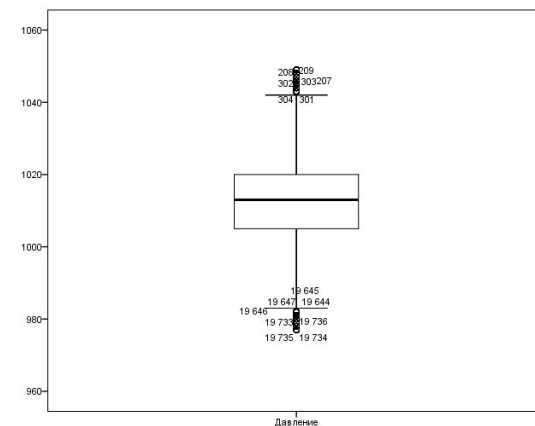
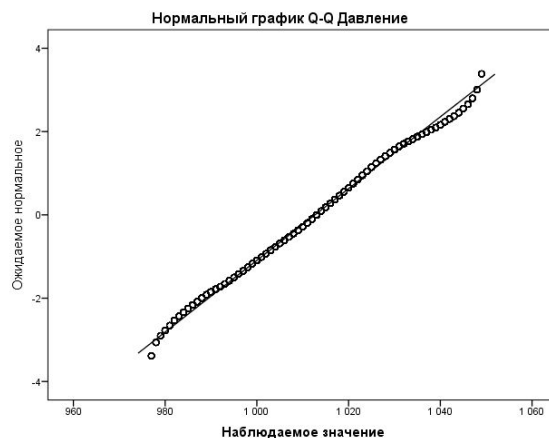
## Описательные статистики

		Статистика	Стандартная Ошибка
Давление	Среднее значение	1012,67	,083
	Медиана	1013,00	
	Асимметрия	-,035	,017
	Экцесс	,265	,035

## Критерии нормального распределения

	Колмогорова-Смирнова <sup>а</sup>		
	Статистика	ст.св.	Значимость
Давление	,039	19619	3,040E-78

а. Коррекция значимости Лиллиефорса



Распределение является ненормальным (Критерий Колмогорова-Смирнова). Грубых наблюдений нет, все  $|z| < 3$ . Критерий z-Вилкоксона

# Вычисление результатов критерия

Итоги по проверке гипотезы

	Нулевая гипотеза	Критерий	Знач.	Решение
1	Медиана Давление равна 1 000,000.	Одновыборочный критерий знаковых рангов Вилкоксона	,000	Нулевая гипотеза отклоняется.

Выводятся асимптотические значимости. Уровень значимости равен ,05.

Описательные статистики

		Статистика	Стандартная Ошибка
Давление	Среднее значение	1012,67	,083
	Медиана	1013,00	
	Асимметрия	-,035	,017
	Экссесс	,265	,035

**Но не верна.**

Значение критерия  $p < 0,05$ , медиана давления не равна 1000 МПа и отклоняется в большую сторону.

# Регрессионный анализ

Основная задача РА написать линейной уравнение.

С помощью этого уравнения предсказать значение зависимости одной переменной от заданных значениях от другой независимой переменной.

**Замечание:** предсказание возможно только в диапазоне от min до max значения независимой переменной.

# Условия применения линейного РА

- ✓ Зависимая и независимая переменные – количественные
- ✓ Нормальность распределений
- ✓ Отсутствие грубых наблюдений
- ✓ Наличие стат.значимой линейной связи между зависимой и независимой переменными
- ✓ Отсутствие высокой корреляции между зависимыми переменными

# Этапы линейного РА

1. Проверка на нормальность, наличие грубых наблюдений
2. Проверка наличия статистически значимой линейной связи между зависимой и независимой переменной. *Если независимых переменных несколько, доказывают, что между ними нет сильных связей*
3. Доказательство значимости линейной модели (дисперсионный анализ)
4. Составление уравнения регрессии. Доказательство значимости коэффициентов и свободного члена построенного уравнения
5. Анализ остатков:
  - Нормальность распределения (Шапиро-Уилка)
  - Равенство среднего арифметического (t-критерий Стьюдента для одной выборки)
  - Независимость остатков(коэф. Дарбина-уотсона от 1 до 3)
  - Отсутствие тренда между предсказанными стандартизированными значения зависимой переменной и стандартизированными остатками (скаттерограмма)

# Задача: Написать уравнение регрессии вида $отметка2 = a * отм1 + b$

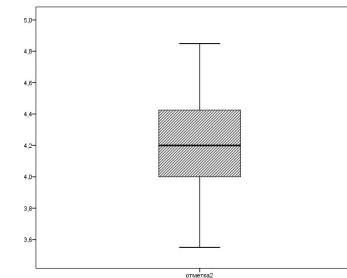
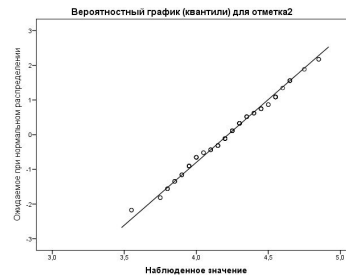
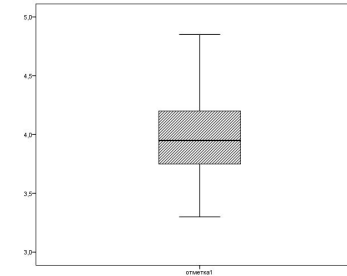
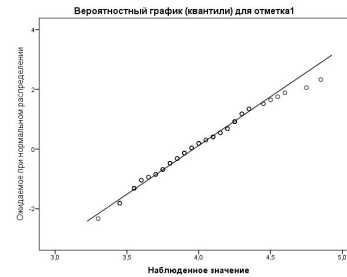
## Описательные

		Статистика	Стд. ошибка
отметка1	Среднее	3,9630	,03060
	Медиана	3,9500	
	Асимметрия	,317	,241
	Эксцесс	-,056	,478
отметка2	Среднее	4,2205	,02759
	Медиана	4,2000	
	Асимметрия	-,007	,241
	Эксцесс	-,357	,478

## Критерий нормальности

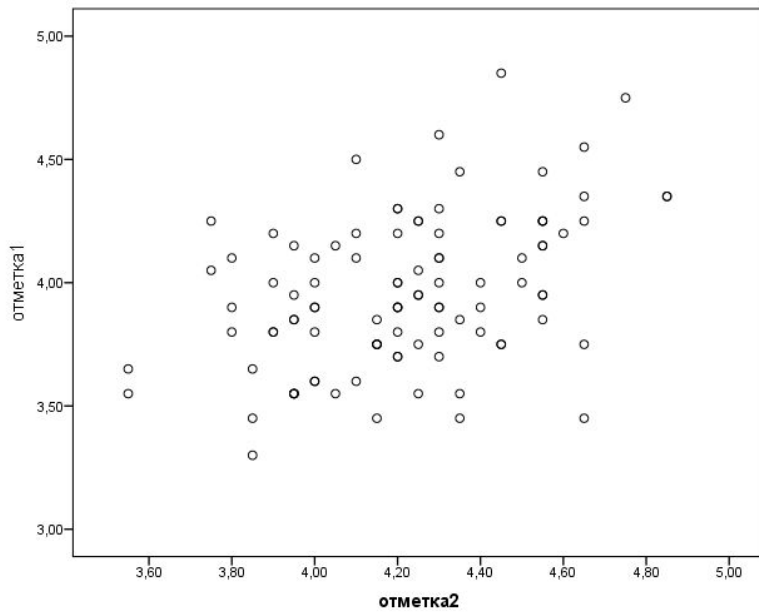
	Шапиро-Уилк		
	Статистика	ст. св.	Значимость
отметка1	,984	100	,258
отметка2	,986	100	,397

а. Поправка значимости Лилефорса



Оба распределения нормальные, грубых наблюдений нет, все  $|z| < 3$ .





### Корреляции

		отметка2
отметка1	Корреляция Пирсона	,434
	Знч.(2-сторон)	6,496E-006
	N	100

Переменные связаны между собой прямой линейной связью.

# Док-во значимости линейной модели

Но: Уравнение писать нельзя

Дисперсионный анализ<sup>а</sup>

Модель	Сумма квадратов	ст.св.	Средний квадрат	F	Знч.
1 Регрессия	1,418	1	1,418	22,726	6,496E-006 <sup>б</sup>
Остаток	6,117	98	,062		
Всего	7,535	99			

а. Зависимая переменная: отметка2

б. Предикторы: (конст) отметка1

**Но не верна.** Линейное уравнение составлять можно.

Сводка для модели

Модель	R	R-квадрат
1	,434 <sup>а</sup>	,188

а. Предикторы: (конст)  
отметка1

# Составление уравнения. Док-во значимости коэф. И свободного члена построенного уравнения.

Коэффициенты<sup>а</sup>

Модель	Нестандартизованные коэффициенты		Стандартизо ванные коэффициент ы	t	Знч.
	B	Стд. Ошибка	Бета		
1 (Константа)	2,670	,326		8,186	1,016E-012
отметка1	,391	,082	,434	4,767	6,496E-006

$$\text{Отм2} = 0.391 * \text{отм1} + 2.67$$

а. Зависимая переменная: отметка2

Коэффициент и свободный член не нулевые (по t-критерию  
Стьюдента).

# Анализ остатков

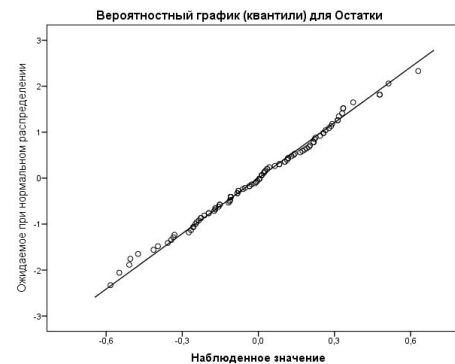
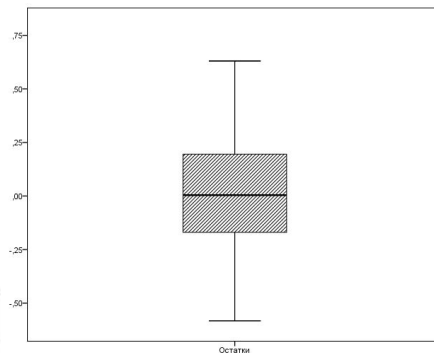
## Описательные

		Статистика	Стд. ошибка
Остатки	Среднее	,0000	,02486
	Медиана	,0041	
	Асимметрия	-,079	,241
	Эксцесс	-,234	,478

## Критерий нормальности

	Шапиро-Уилк		
	Статистика	ст. св.	Значимость
Остатки	,993	100	,895

а. Поправка значимости Лилефорса



Распределение нормальное, грубых наблюдений нет.

**Статистики для одновыборочного t-критерия**

	N	Среднее	Стд. отклонение	Стд. ошибка среднего
Остатки	100	-1,8652E-016	,24857	,02486

**Одновыборочный t-критерий**

	Проверяемое значение = 0			
	t	ст.св.	Значимость (2-сторонняя)	Разность средних
Остатки	-7,504E-015	99	1,000	-1,86517E-16

Но верна, среднее арифметическое равно  
0

# Независимость остатков

Коэффициент Дурбина-Уотса (от 0 до 4)

Проверяет на автокорреляцию. Считает, что автокорреляции положительной нет, тогда коэф. Должен быть от 1 до 3.

Сводка для модели<sup>b</sup>

Модель	R	R-квадрат	Скорректиро- ванный R- квадрат	Стд. ошибка оценки	Дурбин- Уотсон
1	,434 <sup>a</sup>	,188	,180	,24984	1,544

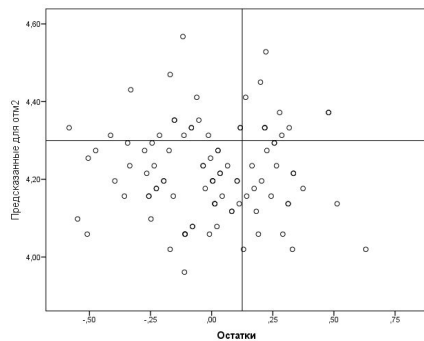
a. Предикторы: (конст) отметка1

b. Зависимая переменная: отметка2

Положительной автокорреляции нет, коэф. Дарбина-Уотсона в диапазоне от 1 до 3. Остатки независимые.

### Корреляции

		Остатки
Предсказанные для отм2	Корреляция Пирсона	7,620E-016
	Знач.(2-сторон)	1,000
	N	100



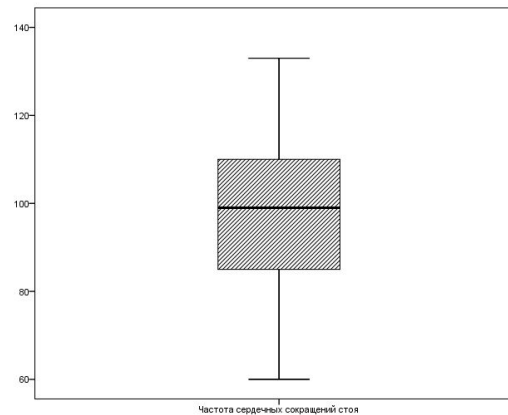
Между предсказанными значениями и остатками тренд отсутствует.

# Задача: Составить уравнение регрессии ЧСС2 от ЧСС1

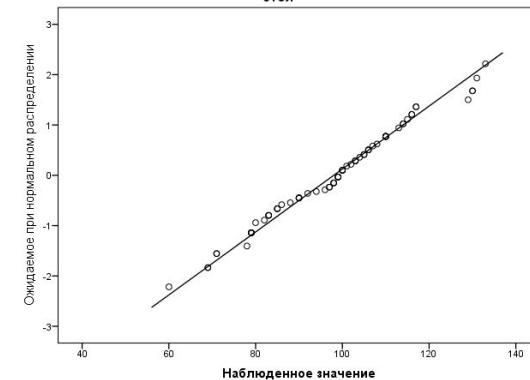
Оба распределения нормальные, грубых наблюдений нет, все  $|z| < 3$ .

## Описательные

		Статистика	Стд. ошибка
Частота сердечных сокращений лежа	Среднее	92,62	1,448
	Медиана	92,50	
	Асимметрия	,415	,279
	Эксцесс	,292	,552
Частота сердечных сокращений стоя	Среднее	98,00	1,860
	Медиана	99,00	
	Асимметрия	,025	,279
	Эксцесс	-,273	,552



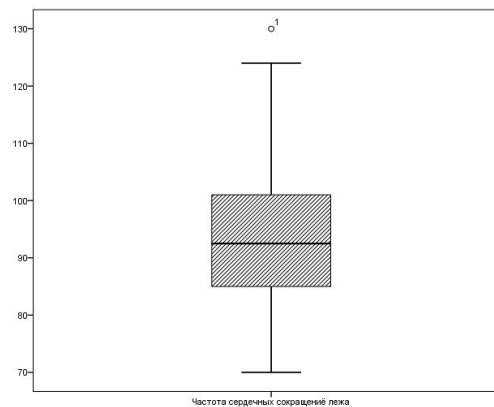
Вероятностный график (квантили) для Частота сердечных сокращений стоя



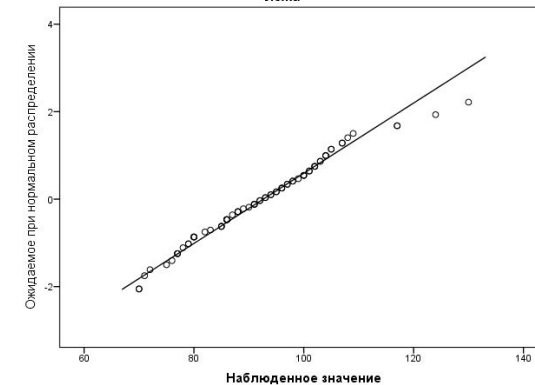
## Критерий нормальности

	Шапиро-Уилк		
	Статистика	ст. св.	Значимость
Частота сердечных сокращений лежа	,977	74	,209
Частота сердечных сокращений стоя	,982	74	,361

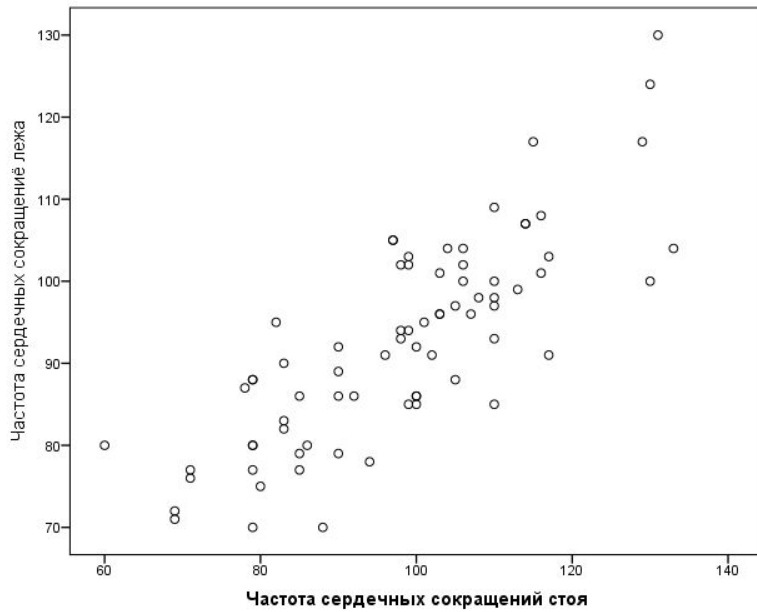
а. Поправка значимости Лиллиефорса



Вероятностный график (квантили) для Частота сердечных сокращений лежа







### Корреляции

		Частота сердечных сокращений стоя
Частота сердечных сокращений лежа	Корреляция Пирсона	,807
	Знч.(2-сторон)	4,096E-018
	N	74

Переменные связаны между собой прямой линейной высокой связью.

Но: Уравнение писать нельзя.

Сводка для модели<sup>b</sup>

Модель	R	R-квадрат
1	,807 <sup>a</sup>	,651

а. Предикторы: (конст)  
Частота сердечных  
сокращений лежа

б. Зависимая переменная:  
Частота сердечных  
сокращений стоя

Дисперсионный анализ<sup>а</sup>

Модель		Сумма квадратов	ст.св.	Средний квадрат	F	Знч.
1	Регрессия	12157,038	1	12157,038	134,189	4,096E-018 <sup>б</sup>
	Остаток	6522,962	72	90,597		
	Всего	18680,000	73			

а. Зависимая переменная: Частота сердечных сокращений стоя

б. Предикторы: (конст) Частота сердечных сокращений лежа

Но не верна. Линейное уравнение составлять можно.

Коэффициенты<sup>а</sup>

Модель	Нестандартизованные коэффициенты		Стандартизованные коэффициенты	t	Знч.
	B	Стд. Ошибка	Бета		
1 (Константа)	2,072	8,355		,248	,805
Частота сердечных сокращений лежа	1,036	,089	,807	11,584	4,096E-018

а. Зависимая переменная: Частота сердечных сокращений стоя

$$ЧСС2 = 1,04 * ЧСС1 + 0$$

Статистики остатков<sup>а</sup>

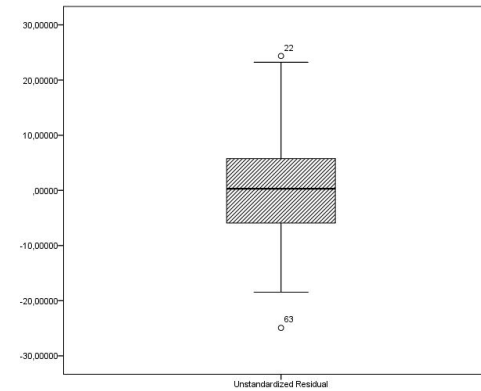
	Минимум	Максимум	Среднее	Стд. Отклонение	N
Предсказанное значение	74,57	136,71	98,00	12,905	74
Остаток	-24,928	24,358	,000	9,453	74
Стд. Предсказанное значение	-1,816	3,000	,000	1,000	74
Стд. Остаток	-2,619	2,559	,000	,993	74

а. Зависимая переменная: Частота сердечных сокращений стоя

Коэффициент  $b=0$ , т.к. значимость  $>0,1$   
 Коэффициент не нулевой (по t-критерию Стьюдента).

### Описательные

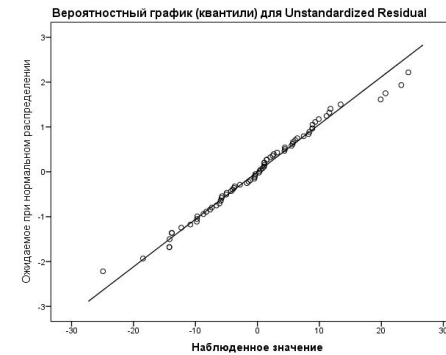
		Статистика	Стд. ошибка
Частота сердечных сокращений лежа	Среднее	92,84	1,446
	Медиана	93,00	
	Асимметрия	,381	,277
	Эксцесс	,194	,548
Частота сердечных сокращений стоя	Среднее	97,52	1,896
	Медиана	99,00	
	Асимметрия	-,034	,277
	Эксцесс	-,265	,548



### Критерий нормальности

	Шапиро-Уилк		
	Статистика	ст. св.	Значимость
Частота сердечных сокращений лежа	,979	75	,264
Частота сердечных сокращений стоя	,982	75	,368

а. Поправка значимости Лильефорса



Распределение нормальное, грубых наблюдений нет.

**Одновыборочный t-критерий**

	Проверяемое значение = 0					
	t	ст.св.	Значимость (2- сторонняя)	Разность средних	95% доверительный интервал разности средних	
					Нижняя граница	Верхняя граница
Unstandardized Residual	,000	73	1,000	,00000000	-2,1900387	2,1900387

**Статистики для одновыборочного t-критерия**

	N	Среднее	Стд. отклонение	Стд. ошибка среднего
Unstandardized Residual	74	,00000000	9,45281118	1,09886698

Среднее=0

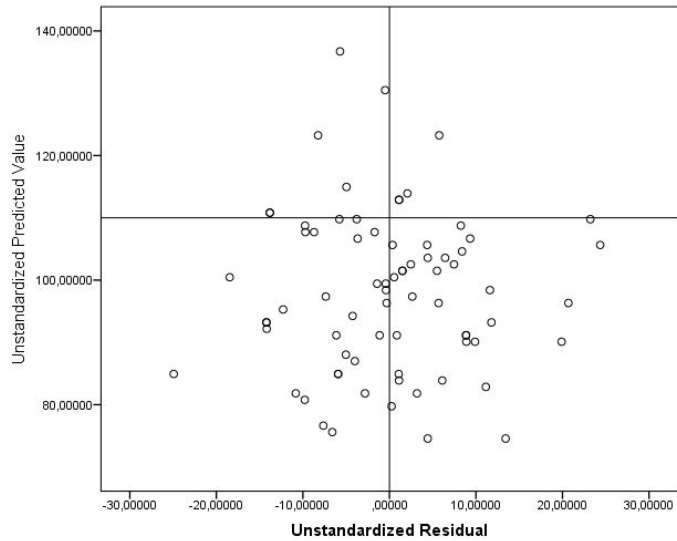
### Сводка для модели<sup>b</sup>

Модель	R	R-квадрат	Скорректиро- ванный R- квадрат	Стд. ошибка оценки	Дурбин- Уотсон
1	,807 <sup>a</sup>	,651	,646	9,518	1,880

а. Предикторы: (конст) Частота сердечных сокращений лежа

б. Зависимая переменная: Частота сердечных сокращений стоя

Положительной автокорреляции нет,  
коэф. Дарбина-Уотсона в диапазоне от  
1 до 3. Остатки независимые.



### Корреляции

		Unstandardiz ed Residual
Unstandardized Predicted Value	Корреляция Пирсона	7,022E-017
	Знч.(2-сторон)	1,000
	N	74

Между предсказанными значениями и остатками тренд отсутствует.

$$ЧCC2=74*1,04=76,96$$

$$ЧCC2=106*1,04=110,24$$

$$ЧCC2=120*1,04=124,8$$



Множественная линейная регрессия

$$Отм1 = a1 * тест1 + a2 * тест2 + a3 * тест3 + a4 * тест4 + a5 * тест5 + b$$

$$Отм1 = 0,026 * тест1 + 0,032 * тест4 + 2,509$$

Домашка регрессионный анализ по  
базе

## Сравнение 3х и более выборок. Дисперсионный анализ

- Одномерный однофакторный дисперсионный анализ для 3 и более **независимых** выборок.
  - Нормальность распределения
  - Нет грубых наблюдений
  - Равенство дисперсий
  - Не менее 5 в каждой группе, не более 30 всего.

**Альтернатива:** непараметрический  
ДА Н-Крускала-Уоллиса

## Этапы сравнения трех и более независимых выборок

- ✓ Нулевая гипотеза
- ✓ Проверка на нормальность, наличие грубых наблюдений
- ✓ Сравнение трех выборок
- ✓ В случае различий-сравнение двух выборок или контрасты
- ✓ Иллюстрация обнаружения различий с помощью диаграмм

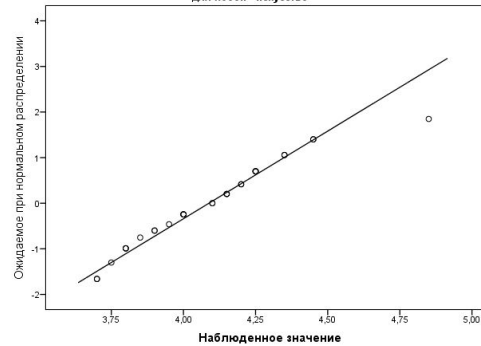
# Задача: Отличается ли Отметка 1 (Отметка2) у учащихся, увлекающихся разными видами хобби?

1. Но: Отметка1 одинаковая у учащихся с разными хобби.
2. Проверка на нормальность и на наличие грубых наблюдений

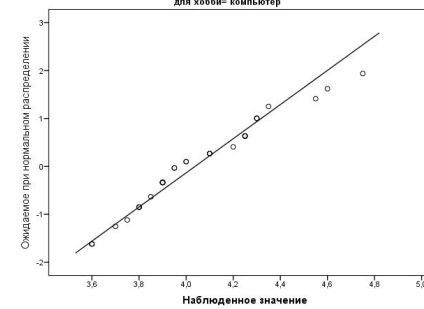
Описательные

хобби		Статистика	Стд. ошибка
отметка1 спорт	Среднее	3,7652	,04863
	Медиана	3,7500	
	Асимметрия	,642	,409
	Экссесс	-,084	,798
компьютер	Среднее	4,0378	,04615
	Медиана	3,9500	
	Асимметрия	,552	,388
	Экссесс	-,058	,759
искусство	Среднее	4,0883	,04751
	Медиана	4,1000	
	Асимметрия	,698	,427
	Экссесс	1,002	,833

Вероятностный график (квантили) для отметка1 для хобби= искусство



Вероятностный график (квантили) для отметка1 для хобби= компьютер

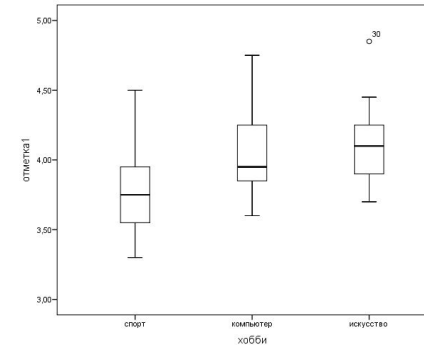
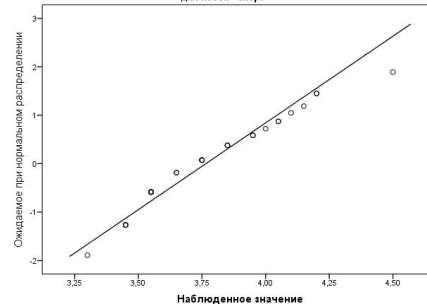


Критерий нормальности

хобби	Шapiro-Уилк		
	Статистика	ст. св.	Значимость
отметка1 спорт	,943	33	,083
компьютер	,951	37	,105
искусство	,951	30	,180

а. Поправка значимости Лильфорса

Вероятностный график (квантили) для отметка1 для хобби= спорт



Распределения ненормальные, грубых наблюдений нет

### 3. Сравнение трех выборок

Сравниваются три независимые выборки, малые по численности. Для сравнения выбираем критерий Н-Крускала-Уоллиса.

Процентили (квантили)

		Процентили (квантили)		
		25	50	75
Взвешенное среднее (Определение 1)	хобби			
	спорт	3,5500	3,7500	3,9750
	компьютер	3,8250	3,9500	4,2500
	искусство	3,8875	4,1000	4,2500

Статистики критерия<sup>a,b</sup>

	отметка1
Хи-квадрат	21,780
ст.св.	2
Асимпт. знч.	1,864E-005

а. Критерий  
Краскела-Уоллеса

б. Группирующая  
переменная:  
хобби

**Но не верна.** Успеваемость учащихся с разным видом хобби разная. Необходимо попарное сравнение.

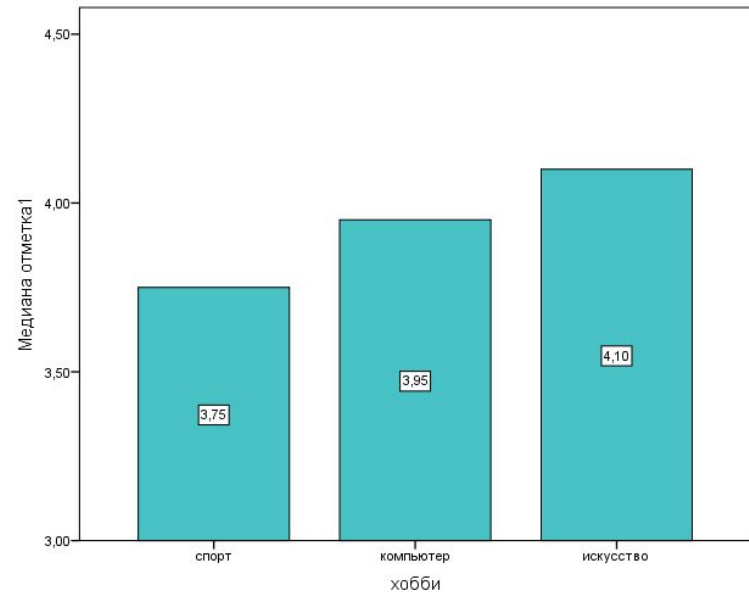
	Спорт/компьютер	Спорт/Искусство	Компьютер/Искусство
U-Манна-Уитни	284,500	192,500	495,000
Ассимпт.значимость	1,209E-004	3,001E-005	,448
	1 209		

### Процентили (квантили)

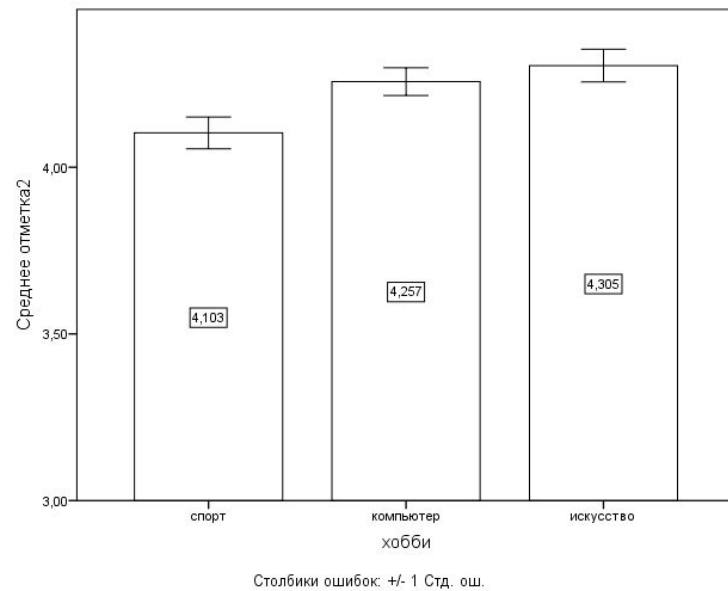
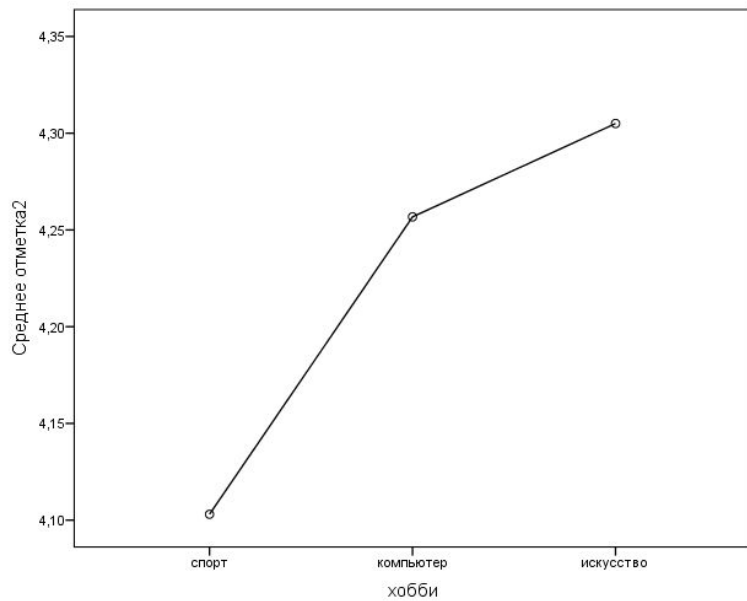
			Процентили (квантили)		
			25	50	75
Взвешенное среднее (Определение 1)	хобби	спорт	3,5500	3,7500	3,9750
	отметка1	компьютер	3,8250	3,9500	4,2500
		искусство	3,8875	4,1000	4,2500

Учащиеся-искусствоведы учатся лучше, чем спортсмены. Искусствоведы и компьютерщики учатся одинаково, компьютерщики учатся лучше спортсменов.

Вывод: В 10 классе средний балл у искусствоведов и компьютерщиков выше, чем у спортсменов.



# График и диаграмма средних



## Этапы сравнения трех и более зависимых выборок

- Проверка на нормальность и грубые наблюдения
- Зависимость
- Сравнения выборок
- Сравнение по 2-3 выборки при наличие различий
- Диаграммы



# Сравнить результаты учащихся по пяти тестам

## 1. Нормальность и наличие грубых наблюдений

### Критерий нормальности

	Шапиро-Уилк		
	Статистика	ст. св.	Значимость
счет в уме	,972	100	,031
числовые ряды	,983	100	,216
словарь	,969	100	,019
осведомленность	,983	100	,219
кратковременная память	,984	100	,280

а. Поправка значимости Лилъефорса

Распределения ненормальные, нет грубых наблюдений

Корреляции

			счет в уме	числовые ряды	словарь	осведомленность
ро Спирмена	числовые ряды	Кoeffициент корреляции	,401			
		Знч. (2-сторон)	3,556E-005			
		N	100			
	словарь	Кoeffициент корреляции	-,028	-,148		
		Знч. (2-сторон)	,779	,142		
		N	100	100		
	осведомленность	Кoeffициент корреляции	-,093	-,122	,427	
		Знч. (2-сторон)	,358	,225	9,199E-006	
		N	100	100	100	
	кратковременная память	Кoeffициент корреляции	,059	,014	,459	,470
		Знч. (2-сторон)	,560	,894	1,556E-006	8,013E-007
		N	100	100	100	100

Обнаружены связи между переменными Тест1 и тест2; тест 4 и тест3; тест5 и тест3; тест5 и тест 4.

Значимость множественного коэффициента корреляции (коэффициент конкордации w-Кендалла) проверяет критерий Хи-квадрат Фридмена, поэтому отдельно можно не рассчитывать.

# Сравнение пяти зависимых выборок

Описательные статистики

	N	Процентили (квантили)		
		25%	50% (Медиана)	75%
счет в уме	100	9,00	10,00	13,00
числовые ряды	100	8,25	10,50	12,00
словарь	100	10,00	12,00	14,00
осведомленность	100	10,00	12,00	13,00
кратковременная память	100	10,00	12,00	14,00

Ранги

	Средний ранг
счет в уме	2,50
числовые ряды	2,65
словарь	3,41
осведомленность	3,08
кратковременная память	3,37

Статистики критерия<sup>a</sup>

N	100
Хи-квадрат	29,696
ст.св.	4
Асимпт. знч.	5,644E-006

a. Критерий Фридмана

Вероятность  $p < 0,05$

Но отвергается

Результаты по пяти тестам

различаются, имеются согласованные данные.

## Частичные сравнения меньшего числа выборок

Проверим 2 Но : результаты по тесту1 и тесту2 одинаковые; результаты по тесту3,4,5 одинаковые.

Статистики критерия<sup>а</sup>

	числовые ряды - счет в уме
Z	-,592 <sup>b</sup>
Асимпт. знч. (двухсторонняя)	,554

а. Критерий знаковых рангов Уилкоксона

б. Используются отрицательные ранги.

$p > 0,05$ ;

Но верна, результаты по тесту1 и тесту2 одинаковые.

Статистики критерия<sup>а</sup>

N	100
Хи-квадрат	2,767
ст.св.	2
Асимпт. знч.	,251

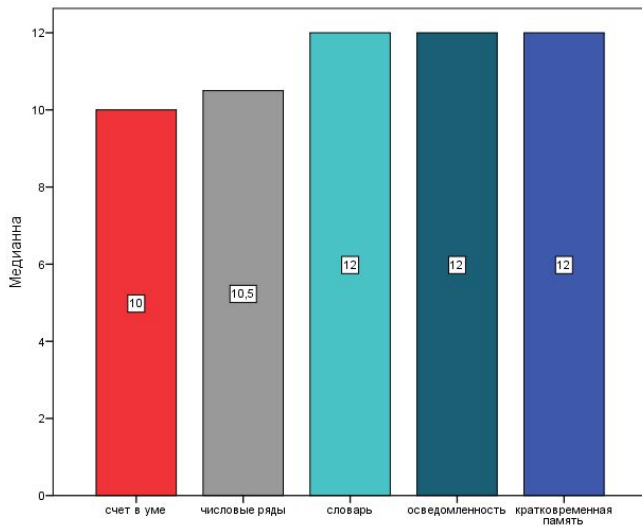
а. Критерий Фридмана

$p > 0,05$ ;

Но верна, результаты по тестам 3,4,5 одинаковые.

Вывод: При сравнении пяти выборок, получили различия. Математические тесты одинаковы, вербальные одинаковы, следовательно, различия в общей выборке объясняются различиями между группами мат.тестов и вербальных тестов.

Вербальные тесты наизусть лучше, чем математические (см. табл. медиан)



Замечания: Если бы обнаружили связи между всеми парами переменных, то можно было бы ответить на вопрос : Какой мат.тест написан хуже какого вербального.

Статистики критерия<sup>а</sup>

	осведомленность - счет в уме	кратковременная память - счет в уме	словарь - счет в уме	словарь - числовые ряды	осведомленность - числовые ряды	кратковременная память - числовые ряды
Z	-2,504 <sup>b</sup>	-3,976 <sup>b</sup>	-3,861 <sup>b</sup>	-3,605 <sup>b</sup>	-2,493 <sup>b</sup>	-3,498 <sup>b</sup>
Асимпт. знч. (двухсторонняя)	,012	7,018E-005	1,129E-004	3,124E-004	,013	4,690E-004

а. Критерий знаковых рангов Уилкоксона

б. Используются отрицательные ранги.

После коррекции Бонферони (умножения на 6) получаем, что словарь и кор.пам. Написаны лучше, чем числовые ряды и счет в уме, а осведомленность лучше, чем счет в уме на уровне стат. тенденции

## ANOVA с повторными измерениями

Задача: Сравнить результаты тестирования учащихся(пять тестов)

Решим задачу вторым способом. Предположим, что распределения нормальные и не содержат грубых наблюдений.

Доказательство зависимости – Пирсен, коррекция Бонферрони

Корреляции

		счет в уме	числовые ряды	словарь	осведомленность
числовые ряды	Корреляция Пирсона	,436			
	Знч.(2-сторон) *10	5,679E-006			
	N	100			
словарь	Корреляция Пирсона	-,051	-,196		
	Знч.(2-сторон) *10	1,000	,500		
	N	100	100		
осведомленность	Корреляция Пирсона	-,134	-,153	,441	
	Знч.(2-сторон) *10	1,000	1,000	4,417E-006	
	N	100	100	100	
кратковременная память	Корреляция Пирсона	,017	,015	,483	,475
	Знч.(2-сторон) *10	1,000	1,000	3,550E-007	5,779E-007
	N	100	100	100	100

Между собой связаны счет в уме и числ.ряды; словарь-осведомленность, словарь-кратковременная память, осведомленность-кратковр. память.

Множественный коэффициент корреляции можно не считать. Его значимость показывает ANOVA

## Сравнение 5 выборок

ANOVA с повторными измерениями реализует 2 подхода:

- многомерный ( $n > k + 10$ )
- одномерный в остальных случаях

### Описательные статистики

	Среднее	Стд. Отклонение	N
счет в уме	10,25	2,653	100
числовые ряды	10,35	2,768	100
словарь	11,96	2,857	100
осведомленность	11,51	3,040	100
кратковременная память	11,84	2,915	100

### Критерий сферичности Моучли<sup>a</sup>

Измерение: ИЗМЕРЕНИЕ-1

Внутригрупповой эффект	W Моучли	Прибл. хи-квадрат	ст.св.	Знч.
номер_теста	,510	65,504	9	1,161E-010

Проверка нулевой гипотезы о том, что ковариационная матрица ошибок ортонормированного преобразования зависимых переменных пропорциональна единичной матрице.

- а. План: Свободный член  
Внутригрупповой план: номер\_теста

### Проверка внутригрупповых эффектов

Измерение: ИЗМЕРЕНИЕ-1

Источник	Сумма квадратов типа III	ст.св.	Средний квадрат	F	Знч.	Частная Эта в Квадрате	
номер_теста	Предполагая сферичность	270,668	4	67,667	9,668	1,810E-007	,089

### Многомерные критерии<sup>a</sup>

Эффект	Значения	F	Ст. св. гипотезы	Ст.св. ошибки	Знч.	Частная Эта в Квадрате	
номер_теста	След Пиллая	,182	5,345	4,000	96,000	,001	,182

- а. План: Свободный член  
Внутригрупповой план: номер\_теста

Но отвергается. Результаты по тестам разные.



## Частные сравнения 2 и 3 выборок

### Критерий парных выборок

	t	ст.св.	Значимость (2- сторонняя)
Пара 1 счет в уме - числовые ряды	-,347	99	,729

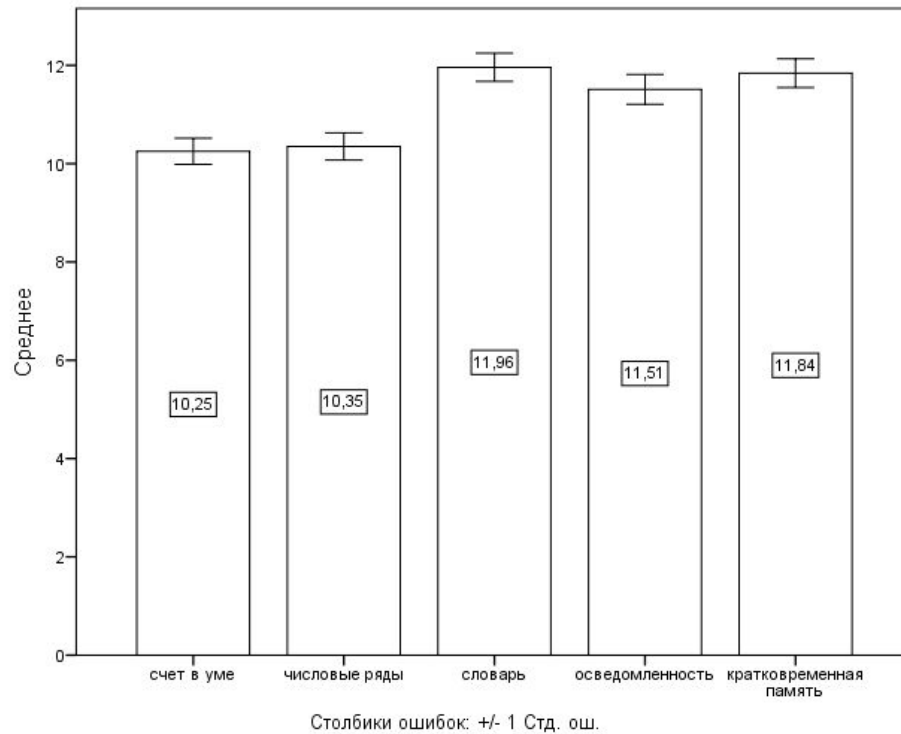
### Многомерные критерии<sup>а</sup>

Эффект	Значения	F	Ст. св. гипотезы	Ст.св. ошибки	Знч.	Частная Эта в Квадрате
тест След Пиллая	,022	1,088	2,000	98,000	,341	,022

а. План: Свободный член  
Внутригрупповой план: тест

Но принимается вербальные тесты написаны  
одинаково

Различия в общей выборке объясняются различиями в  
группах тестов, вербальные написаны лучше, чем  
математические



Можно было бы ответить на вопрос какой вербал лучше математич  
Применить парный t критерий стьюдента и коррекцию бонферони \*6