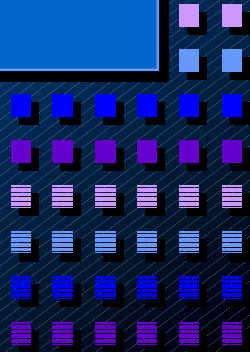




# ОСНОВЫ биостатистики

Анализ различий между  
группами



# Различия между группами

- Качественные переменные

ИЛИ

- Количественные переменные

# Качественные переменные

# Качественные переменные

- Парные наблюдения

ИЛИ

- Не парные наблюдения

# Парные наблюдения

- Метод McNeimar (b - число пар, в которых у лиц, подвергшихся воздействию заболевание развилось, а у контрольных лиц - нет. c - число пар, в которых у лиц, подвергшихся воздействию заболевание не развилось, а у контрольных развилось)

$$OR = \frac{b}{c}$$

$$\chi^2 = \frac{(|b - c| - 1)^2}{b + c}$$

# Качественные переменные

	D+	D-
RF+	a	b
RF-	c	d

Таблица 2x2 или четырехпольная таблица  
для непарных наблюдений

# Непарные наблюдения

- Все частоты более 5 - критерий  $\chi^2$
- Хотя бы одна частота менее 5 - точный метод Фишера или критерий  $\chi^2$  с поправкой Йетса на непрерывность

# Критерий $\chi^2$

- Пограничное значение  $\chi^2 = 3,84$

$$\chi^2 = \frac{(a \cdot d - b \cdot c)^2 \cdot N}{(a + b) \cdot (a + c) \cdot (c + d) \cdot (b + d)}$$



# Критерий $\chi^2$

- Если таблица, больше чем 2x2 ( $r$  - количество строк,  $c$  - количество столбцов)

$$\chi^2 = \sum_{ij} \frac{(O_{ij} - E_{ij})^2}{E_{ij}}$$

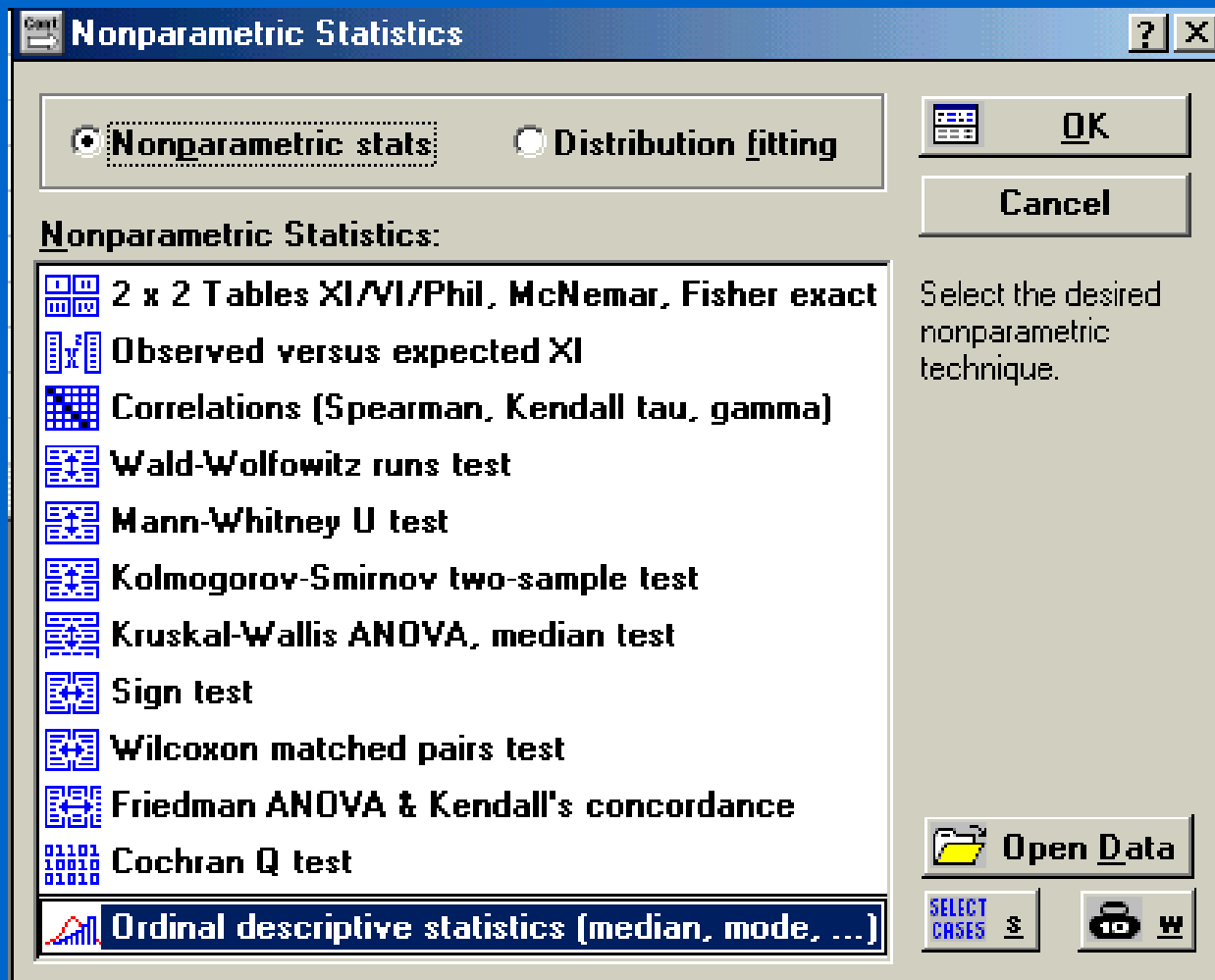
$$df = (r - 1) \cdot (c - 1)$$

$$E_{ij} = \frac{\sum \text{строки}_i \cdot \sum \text{столбца}_j}{\sum \text{таблицы}}$$

# Точный метод Фишера

- Прямое вычисление вероятности случайного формирования данной таблицы. Расчет весьма сложен, но поддерживается большинством программ

# Анализ качественных переменных



# Анализ качественных переменных

**2 x 2 Tables**

Enter the frequencies  
for the 2 x 2 table:

15	15
2	20

Specify the frequencies  
for the two-by-two  
frequency table; then  
click OK

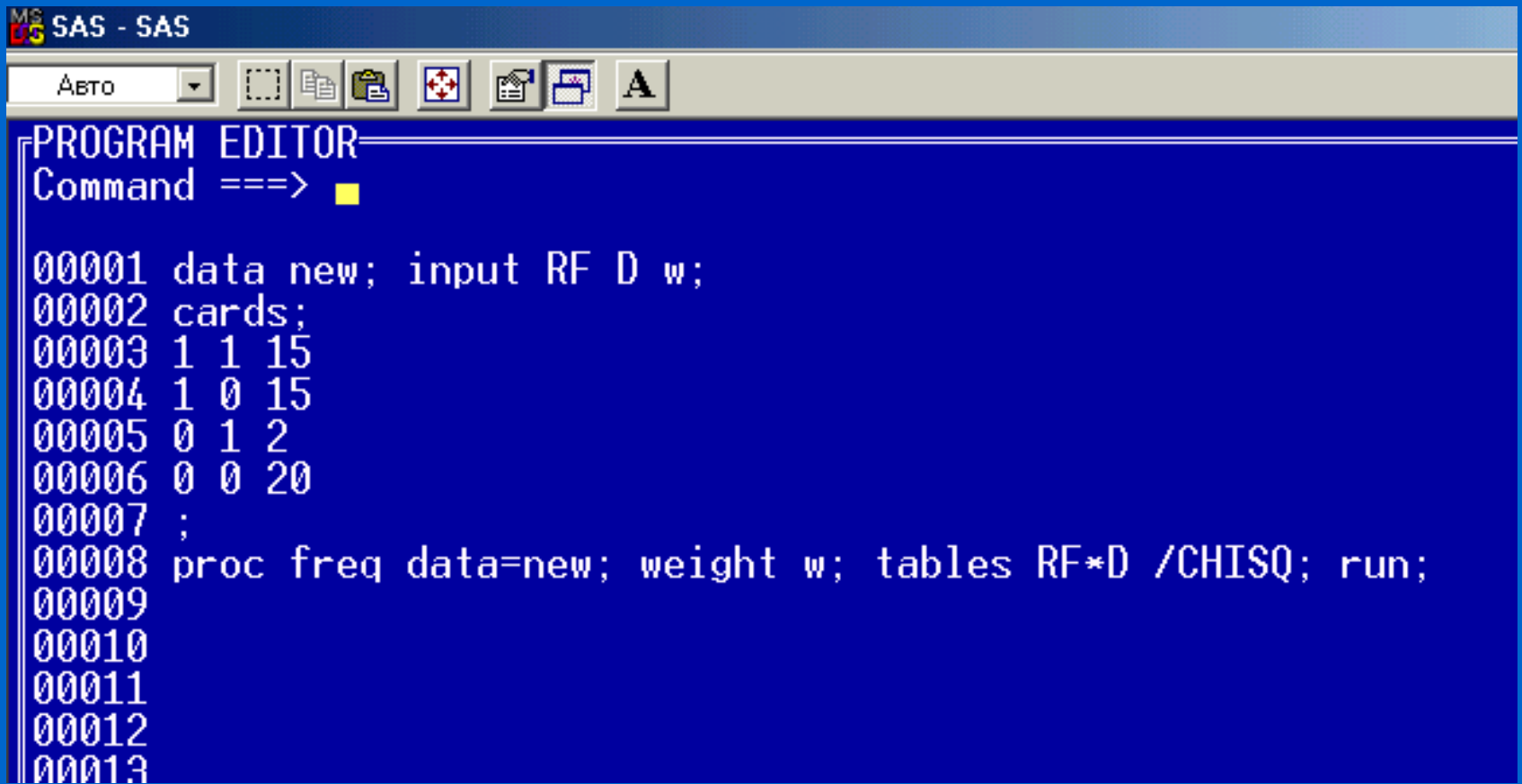
OK  
Cancel

# Анализ качественных переменных

2 x 2 Table (new.sta)

Continue...	Column 1	Column 2	Row Totals
Frequencies, row 1	15	15	30
Percent of total	28,846%	28,846%	57,692%
Frequencies, row 2	2	20	22
Percent of total	3,846%	38,462%	42,308%
Column totals	17	35	52
Percent of total	32,692%	67,308%	
Chi-square (df=1)	9,65	p= ,0019	
V-square (df=1)	9,47	p= ,0021	
Yates corrected Chi-square	7,88	p= ,0050	
Phi-square	,18564		
Fisher exact p, one-tailed		p= ,0018	
two-tailed		p= ,0025	
McNemar Chi-square (A/D)	,46	p= ,4990	
Chi-square (B/C)	8,47	p= ,0036	

# Анализ качественных переменных (SAS)



The screenshot shows the SAS Program Editor window. The title bar reads "MS SAS - SAS". The menu bar includes "Авто" and several icons. The main text area contains the following SAS code:

```
PROGRAM EDITOR
Command ==> ■

00001 data new; input RF D w;
00002 cards;
00003 1 1 15
00004 1 0 15
00005 0 1 2
00006 0 0 20
00007 ;
00008 proc freq data=new; weight w; tables RF*D /CHISQ; run;
00009
00010
00011
00012
00013
```

# Анализ качественных переменных (SAS)

MS SAS - SAS

АВТО

OUTPUT  
Command ==>

SAS 23:13 Monday, June 6, 1

## STATISTICS FOR TABLE OF RF BY D

Statistic	DF	Value	Prob
Chi-Square	1	9.653	0.002
Likelihood Ratio Chi-Square	1	10.733	0.001
Continuity Adj. Chi-Square	1	7.884	0.005
Mantel-Haenszel Chi-Square	1	9.468	0.002
Fisher's Exact Test (Left)			1.000
(Right)			1.78E-03
(2-Tail)			2.47E-03
Phi Coefficient		0.431	
Contingency Coefficient		0.396	
Cramer's V		0.431	

Sample Size = 52



# Количественные переменные



# Количественные переменные

- Две группы

ИЛИ

- Более двух групп

# Две группы

- Парные наблюдения

ИЛИ

- Непарные наблюдения

# Парные наблюдения

- Нормальное распределение - парный t-тест
- Отличное от нормального распределение - тесты знаков, Вилкоксона

# Парный t-тест

- Два варианта расположения данных:
  - Разности в одной колонке (до-после)
  - Две колонки (принято в системе Statistica)

# Парный t-тест

- Разности  
в одной  
колонке

```
PROGRAM EDITOR
Command ==> ■

00001 data new; input BD_diff;
00002 cards;
00003 29
00004 -18
00005 0
00006 43
00007 0
00008 0
00009 16
00010 0
00011 23
00012 33
00013 26
00014 -30
00015 49
00016 49
00017 8
00018 ; run;
00019 proc univariate data=new; var BD_diff; run;
00020
```

# Парный t-тест

OUTPUT

Command ===>

SAS

23:38 Monday,

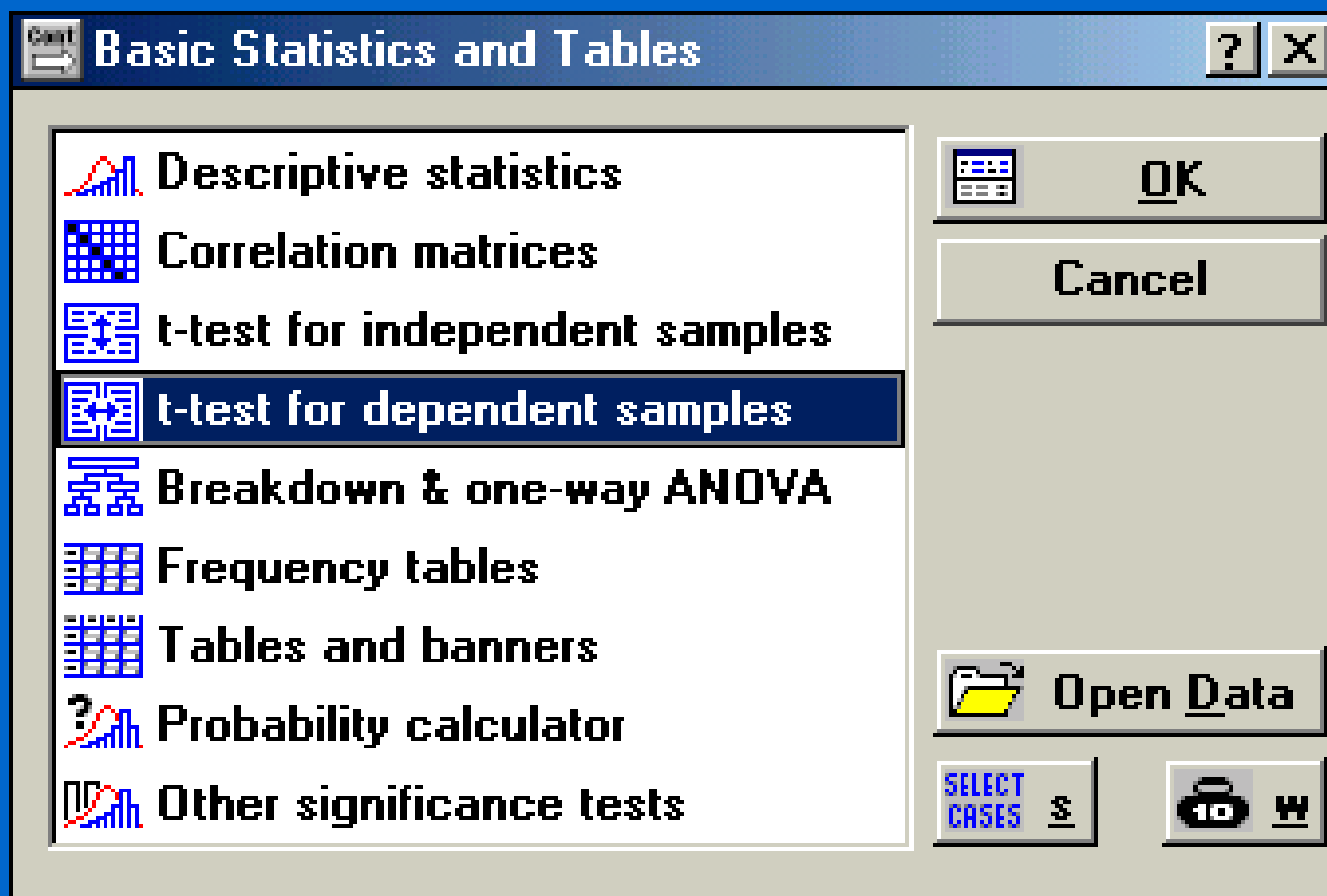
UNIVARIATE PROCEDURE

Variable=BD\_DIFF

Moments

N	15	Sum Wgts	15
Mean	15.2	Sum	228
Std Dev	23.70112	Variance	561.7429
Skewness	-0.21446	Kurtosis	-0.64662
USS	11330	CSS	7864.4
CV	155.9284	Std Mean	6.119602
T: Mean=0	2.483822	Prob> T	0.0263
Sgn Rank	23	Prob> S	0.0410
Num ^= 0	11		

# Парный t-тест



# Парный t-тест

**T-Test for Dependent (Correlated) Samples**

**Variables:**  
First list: BP  
Second list: BP1

I-tests     Box & whisker plots

**Display**

Matrix of t-tests (means, differences)  
 Detailed table of results

**Options**

Casewise deletion of missing data  
 Display long variable names

Weighted moments  
DF =  W-1     N-1

OK Cancel

**T-test for Dependent Samples (sf36l.sta)**

Continue... Marked differences are significant at  $p < .05000$

Variable	Mean	Std. Dv.	N	Diff.	Std. Dv. Diff.	t	df	p
BP	71,60215	23,41411						
BP1	63,08602	24,76878	93	8,516129	22,27868	3,686331	92	,000384



# Непараметрические тесты

- Тест знаков - грубый, но простой

■ до	после	знак
10	15	+
8	9	+
30	35	+
1	4	+
9	10	+
15	27	+

- по таблицам для 6 (+)  $p=0,05$

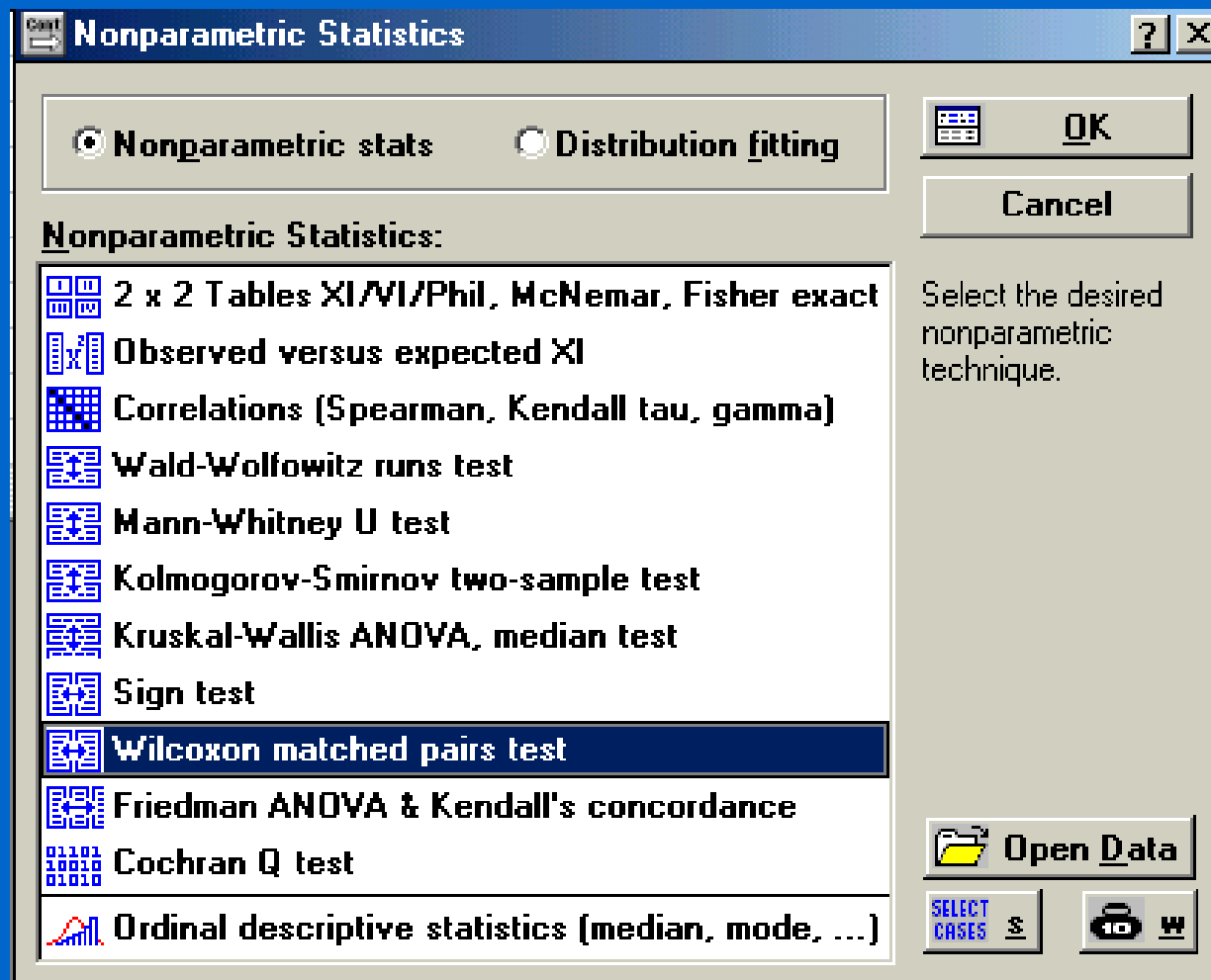
# Тест знаков

5. –	17. 4-13	29. 8-21	41. 13-28	53. 18-35	65. 24-41	77. 29-48	89. 34-55
6. 0-6	18. 4-14	30. 9-21	42. 14-28	54. 19-35	66. 24-42	78. 29-49	90. 35-55
7. 0-7	19. 4-15	31. 9-22	43. 14-29	55. 19-36	67. 25-42	79. 30-49	91. 35-56
8. 0-8	20. 5-15	32. 9-23	44. 15-29	56. 20-36	68. 25-43	80. 30-50	92. 36-56
9. 1-8	21. 5-16	33. 10-23	45. 15-30	57. 20-37	69. 25-44	81. 31-50	93. 36-57
10. 1-9	22. 5-17	34. 10-24	46. 15-31	58. 21-37	70. 26-44	82. 31-51	94. 37-57
11. 1-10	23. 6-17	35. 11-24	47. 16-31	59. 21-38	71. 26-45	83. 32-51	95. 37-58
12. 2-10	24. 6-18	36. 11-25	48. 16-32	60. 21-39	72. 27-45	84. 32-52	96. 37-59
13. 2-11	25. 7-18	37. 12-25	49. 17-32	61. 22-39	73. 27-46	85. 32-53	97. 38-59
14. 2-12	26. 7-19	38. 12-26	50. 17-33	62. 22-40	74. 28-46	86. 33-53	98. 38-60
15. 3-12	27. 7-20	39. 12-27	51. 18-33	63. 23-40	75. 28-47	87. 33-54	99. 39-60
16. 3-13	28. 8-20	40. 13-27	52. 18-34	64. 23-41	76. 28-48	88. 34-54	100.39-61

# Тест Вилкоксона


- По силе примерно равен t-тесту для связанных совокупностей если применяется на нормальном распределении

# Тест Вилкоксона





# Тест Вилкоксона

**Wilcoxon Matched Pairs Test** [?] [X]



 **Variables**

Variable list 1: **BP**  
Variable list 2: **BP1**

 **Box & whisker**

 **OK**

**Cancel**

**SELECT CASES**   **W**

**Wilcoxon Matched Pairs Test (sf36l.sta)** [-] [ ] [X]

<b>Continue...</b>	<b>Valid N</b>	<b>T</b>	<b>Z</b>	<b>p-level</b>
BP & BP1	93	525,0000	3,444065	,000574

# Тест Вилкоксона (SAS)

```
data new; input BD1 BD2 @@;
diff=BD2-BD1;
if diff EQ 0 then delete;
else absdiff=abs(diff);
cards;
76 65 82 67 79 70 85 70 72 60 75 60 86 80
80 71 77 61 79 59
; run;
proc rank out=rang;
var absdiff;
ranks absdiff;
run;
data _null_;
file print;
set rang end=end;
if diff GT 0 then do;
TPOS+absdiff; N_TPOS+1;
end;
```

```
if diff LT 0 then do;
TNEG+absdiff; N_TNEG+1;
end;
if end then do;
minsumme=min(TPOS, TNEG);
N=N_TPOS+N_TNEG;
Z=(MINSUMME-
(N*(N+1)/4))/SQRT(N*(N+1)*(2*N+1)/24);
P2=ROUND(PROBNORM(Z)*2,.0001);
put ' Wilcoxon test for dependent sample';
put ' -----'/;
put 'Z-Approx.:z @30 '2-tail probability = ' P2;
put 'Number positive ranges: ' @40 N_TPOS 8.0/
'Number negative ranges: ' @40 N_TNEG 8.0;
put 'Sum of positive ranges: ' @40 TPOS 8.0/
'Sum of negative ranges: ' @40 TNEG 8.0;
put 'Number of cases: ' @40 N 8.0;
end;
run;
```

# Тест Вилкоксона (SAS)

OUTPUT

Command ==>

SAS

0:34 Monday, June 6,

Wilcoxon test for dependent sample

---

Z-Approx.:	-2.803059553	2-tail probability =	0.0051
Number positive ranges:			0
Number negative ranges:			10
Sum of positive ranges:			0
Sum of negative ranges:			55
Number of cases:			10

# Тесты для независимых совокупностей

- Нормальное распределение (t-тест)
- Иное распределение:
  - различия средних (медиан) - тесты Манна-Уитни, Ван-дер-Вардена
  - различная форма распределения - тесты Вальда-Вольфовица, Колмогорова-Смирнова



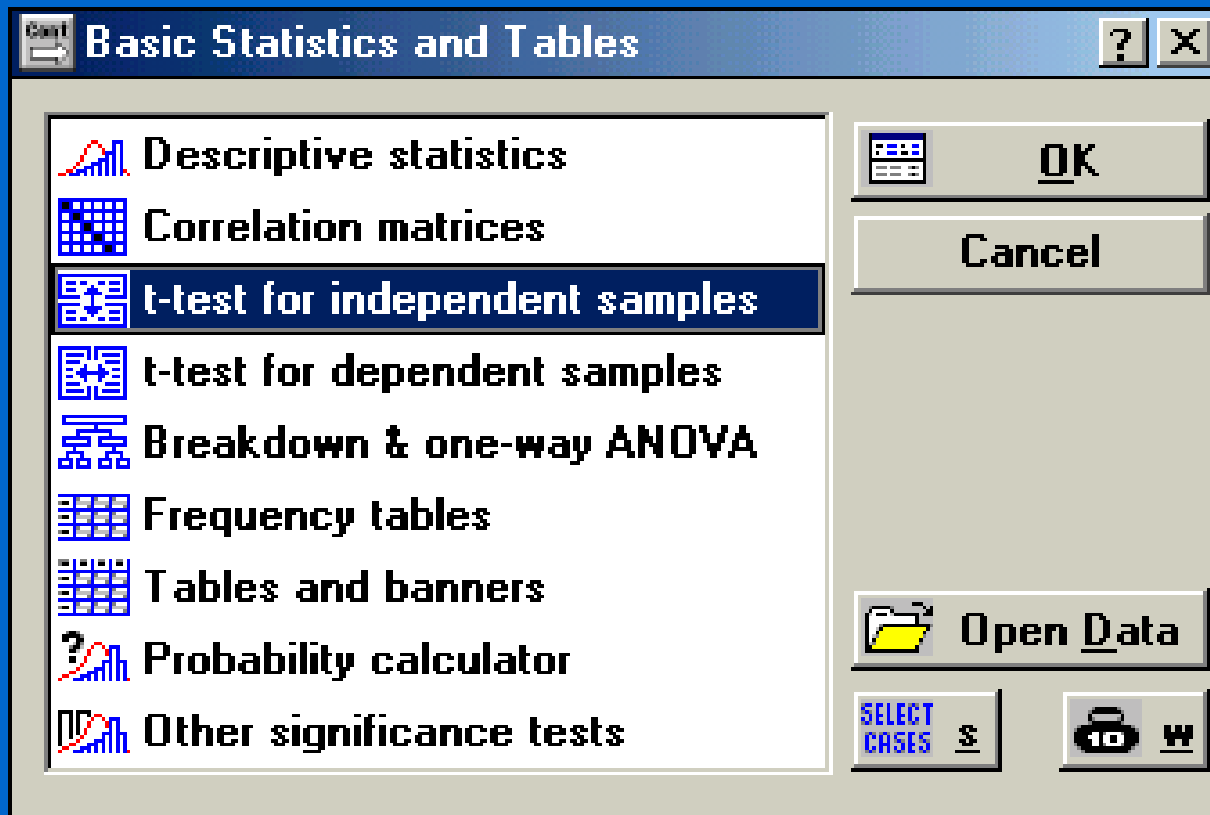
# t-тест для независимых совокупностей

- Самый распространенный тест:

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

– затем значение  $t$  сравниваем с  
табличными,  $df=n-1$

# t-ТЕСТ ДЛЯ НЕЗАВИСИМЫХ СОВОКУПНОСТЕЙ



# t-тест для независимых совокупностей

**T-Test for Independent Samples (Groups)**

Input file: **One record per case (use a grouping variable)**

Variables: **Grouping: SEX**  
**Dependent: PAR**

Code for Group 1: **1** Double-click on the respective field to select codes from the list of valid variable values

Code for Group 2: **2**

**I-tests**

**Options**

- Casewise deletion of missing data
- Display long variable names
- t-test with separate variance estimates
- Multivariate test (Hotelling's T1)
- Levene's test (homogeneity of variances)
- Brown & Forsythe (homog. of variances)

**Box & whisker plot**

**Categorized histograms**

**Categorized normal prob. plots**

**Categ. detrended normal plots**

**Categorized scatterplot**

Weighted moments:  DF =   W-1  N-1

Grouping: SEX: пол: 1-мужской, 2-женский (eysenk.sta)

Variable	df	p	Valid N G_1:1	Valid N G_2:2	Std.Dev. G_1:1	Std.Dev. G_2:2	F-ratio variances	p variances
PAR	199	.019099	69	132	13,58210	12,49863	1,180890	.415992

# t-тест для независимых совокупностей (SAS)

PROGRAM EDITOR

Command ==> ■

```
00001 data new; input ch group @@; cards;
00002 220 1 219 1 240 1 209 2 250 1 209 2 240 1 230 1 230 1
00003 200 2 240 1 190 2 220 2 200 2 205 2 218 2
00004 ;
00005 run;
00006 proc ttest data=new; class group; run;
00007
00008
```

# t-тест для независимых совокупностей (SAS)

OUTPUT

Command ==>

SAS

21:01 Monday, June 6, 1994

TTEST PROCEDURE

Variable: CH

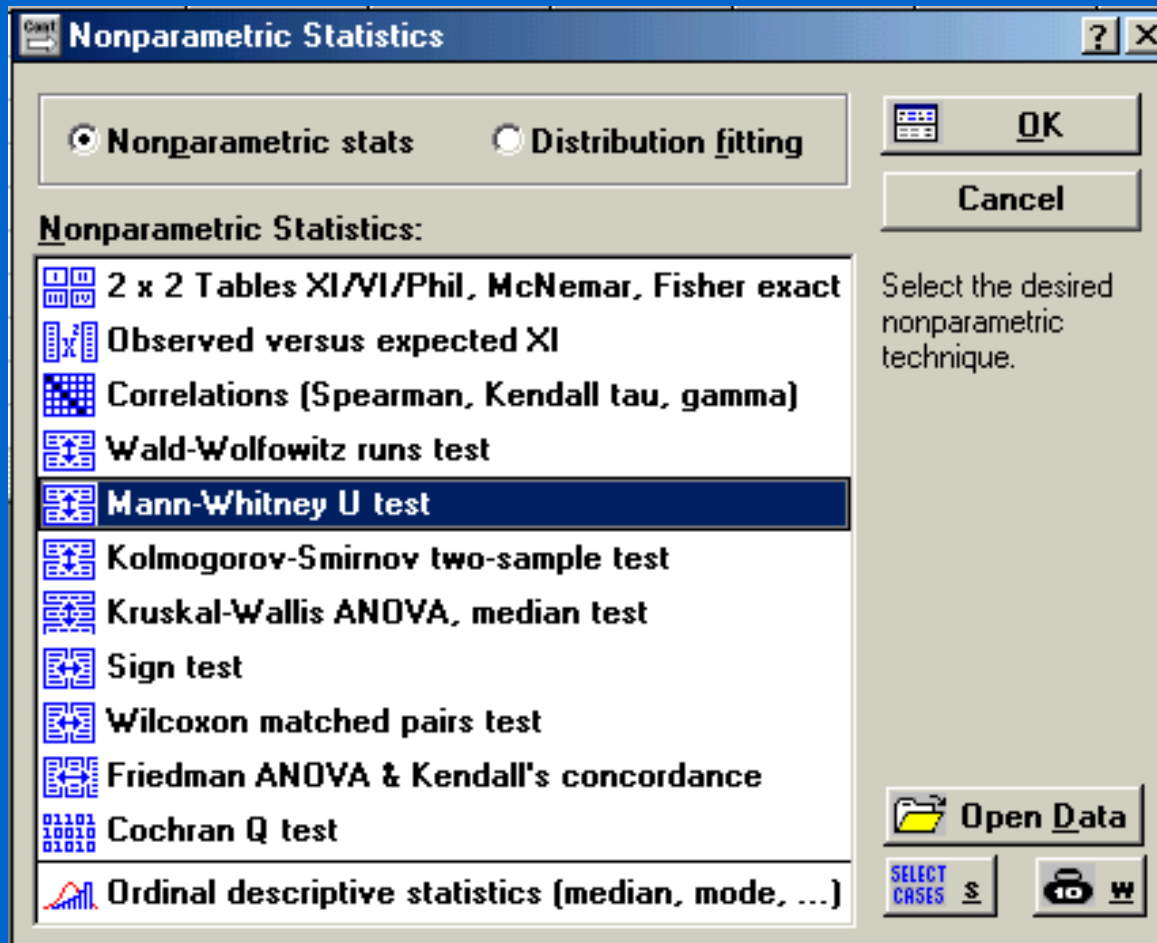
GROUP	N	Mean	Std Dev	Std Error	Minimum	Maximum
1	8	233.6250000	10.79599794	3.81696168	219.0000000	250.000
2	8	206.3750000	9.89859297	3.49968111	190.0000000	220.000

Variances	T	DF	Prob> T
Unequal	5.2621	13.9	0.0001
Equal	5.2621	14.0	0.0001

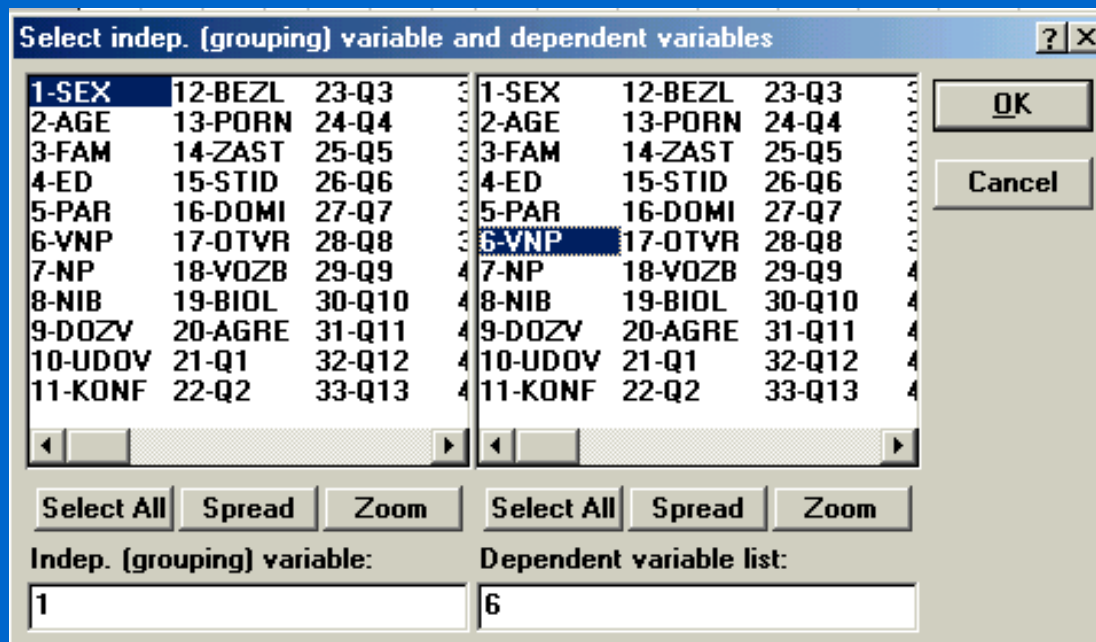
For H0: Variances are equal, F' = 1.19      DF = (7,7)      Prob>F' = 0.8247



# Отличное от нормального распределение



# Отличное от нормального распределение



Mann-Whitney U Test (eysenk.sta)

Continue... By variable SEX  
Group 1: 1 Group 2: 2

variable	Rank Sum Group 1	Rank Sum Group 2	U	Z	p-level	Z adjusted	p-level
VNP	5249,500	15051,50	2834,500	-4,39143	,000011	-4,43798	,000009



## Отличное от нормального распределение (SAS)

PROGRAM EDITOR

Command ==> ■

```
00001 data new; input test_x group @@; cards;
00002 12 1 14 1 13 1 90 1 15 1 16 1 14 1 18 1
00003 0 2 3 2 2 2 5 2 12 2 3 2 6 2 14 2
00004 ;
00005 run;
00006 proc npar1way data=new; class group; run;
00007
```

# Отличное от нормального распределение (SAS)

OUTPUT  
Command ==>

SAS 22:24 Monday, June 6, 1994

## N P A R 1 W A Y P R O C E D U R E

Wilcoxon Scores (Rank Sums) for Variable TEST\_X  
Classified by Variable GROUP

GROUP	N	Sum of Scores	Expected Under H0	Std Dev Under H0	Mean Score
1	8	96.5000000	68.0	9.47980309	12.0625000
2	8	39.5000000	68.0	9.47980309	4.9375000

Average Scores were used for Ties

Wilcoxon 2-Sample Test (Normal Approximation)  
(with Continuity Correction of .5)

S= 96.5000 Z= 2.95365 Prob > |Z| = 0.0031

T-Test approx. Significance = 0.0099

# Отличное от нормального распределение (SAS)

OUTPUT

Command ==>

SAS 22:24 Monday, June 6, 1994

## N P A R I W A Y P R O C E D U R E

Van der Waerden Scores (Normal) for Variable TEST\_X  
Classified by Variable GROUP

GROUP	N	Sum of Scores	Expected Under H0	Std Dev Under H0	Mean Score
1	8	5.08856797	0.0	1.74194211	0.636070997
2	8	-5.08856797	0.0	1.74194211	-.636070997

Average Scores were used for Ties

Van der Waerden 2-Sample Test (Normal Approximation)  
S= 5.08857 Z= 2.92120 Prob > |Z| = 0.0035

Van der Waerden 1-Way (Chi-Square Approximation)  
CHISQ= 8.5334 DF= 1 Prob > CHISQ= 0.0035

# Анализируется более двух групп

- Проблема при анализе нескольких групп - резкое повышение вероятности альфа ошибки (ошибки первого типа). Она повышается в соответствии с формулой:

$$\alpha = 1 - (1 - 0.05)^n$$

- где  $n$  - число сравнений. Так для 5 групп количество сравнений 24, поэтому ошибка равна 0,71

# Анализируется более двух групп

- Способы множественного сравнения:
  - Метод Бонферрони - уровень значимости понижается в соответствии с количеством сравнений. Для предыдущего примера уровень значимости не 0,05, а  $0,05/24=0,002$
  - использование дисперсионного анализа с тестами post hoc:
    - » Дункан
    - » Ньюман-Коулс
    - » Тьюки

# Дисперсионный анализ (ANOVA)

- Метод сравнения средних, хотя анализируются дисперсии. Основная идея - рассчитать дисперсию без учета принадлежности к разным группам (все из одной популяции) и сравнить с суммой внутригрупповых дисперсий. Если группы из разных популяций (противоречит нулевой гипотезе) дисперсия без учета групповой принадлежности будет значительно больше, чем сумма внутригрупповых.
- Для оценки отношения дисперсий используется специальный критерий - F критерий Фишера

# Дисперсионный анализ (ANOVA)

- Пример (количество групп  $k=2$ ):

- |  |   |   |   |  |   |   |   |
|--|---|---|---|--|---|---|---|
|  | 2 | 3 | 4 |  | 5 | 6 | 7 |
|--|---|---|---|--|---|---|---|

- средние 

	3	6
--	---	---

- дисперсия 

	2	2
--	---	---

- общая дисперсия 17,5

- модель объясняет  $17,5 - 2 - 2 = 13,5$

- $MSe$  (mean square error) =  $4/4 = 1$  ( $4 - df_e = N - k$ )

- $MSSm$  (mean square model) =  $13,5/1 = 13,5$  ( $1 - df_m = k - 1$ )

- $F = MSS/MSE$

# Дисперсионный анализ (ANOVA)

The image displays two overlapping dialog boxes from the SPSS software interface. The top dialog box is titled "Basic Statistics and Tables" and contains a list of statistical options. The "Breakdown & one-way ANOVA" option is selected and highlighted. The bottom dialog box is titled "Descriptive Statistics and Correlations by Groups (Breakdown)" and shows configuration options for the selected ANOVA test. In this dialog, the "Analysis" dropdown is set to "Detailed analysis of individual tables", the "Grouping" variable is "FAM", and the "Dependent" variable is "NP". The "Codes for grouping variables" field is empty, and the "Casewise (listwise) deletion of MD" checkbox is unchecked. The "Weighted moments" checkbox is also unchecked, and the "DF" is set to "W-1".

**Basic Statistics and Tables**

- Descriptive statistics
- Correlation matrices
- t-test for independent samples
- t-test for dependent samples
- Breakdown & one-way ANOVA**
- Frequency tables
- Tables and banners
- Probability calculator
- Other significance tests

Buttons: OK, Cancel, Open Data, SELECT CASES, ID, W

**Descriptive Statistics and Correlations by Groups (Breakdown)**

Analysis: Detailed analysis of individual tables

Variables

Grouping: FAM  
Dependent: NP

Codes for grouping variables: none

Casewise (listwise) deletion of MD

Buttons: OK, Cancel, SELECT CASES, ID, W, Weighted moments, DF = W-1, N-1



# Дисперсионный анализ (ANOVA)

DEPENDENT: 1 variable: NP

GROUPING: 1-FAM ( 2): 1 2

**Summary table of means**

Detailed two-way tables

Display long variable names

Display long value labels

**Statistics**

Number of observations

Sums

Standard deviations

Variances

Median and quartiles

**Marginal means**

**Analysis of Variance**

**Post-hoc comparisons of means**

Levene  Brown-Forsythe (HOV)

**Within-group correlations** **Options**

**Reorder factors in table**

**Categorized box & whisker plot**

**Categorized histograms**

**Categorized normal prob. plots**

**Interaction plots**

**Categorized scatterplots**

**Plot of means vs. std. devs**

**OK** **Cancel**

# Дисперсионный анализ (ANOVA)

Analysis of Variance [eysenk.sta]

Continue... Marked effects are significant at  $p < .05000$

Variable	SS Effect	df Effect	MS Effect	SS Error	df Error	MS Error	F	p
NP	,236729	1	,236729	12,78815	199	,064262	3,683804	,056374

Post-Hoc Comparisons of Means

LSD test or planned comparison

Scheffé test

Newman-Keuls test & critical ranges

Duncan's multiple range test & critical ranges

Tukey honest significant difference (HSD) test

Tukey HSD for unequal N (Spjotvoll/Stoline test)

Alpha level for critical ranges: .050

Alpha level for highlighting: .05

Cancel

# Дисперсионный анализ (ANOVA)

Tukey HSD test; Variable: NP [eysenk.sta]

Continue... Marked differences are significant at  $p < .05000$

	{1}	{2}
FAM	M=,11765	M=,04511
G_1:1 {1}		,054947
G_2:2 {2}	,054947	

LSD Test; Variable: NP [eysenk.sta]

Continue... Marked differences are significant at  $p < .05000$

	{1}	{2}
FAM	M=,11765	M=,04511
G_1:1 {1}		,056374
G_2:2 {2}	,056374	

# ANOVA (SAS)

- Могут использоваться две процедуры: ANOVA и GLM
  - ANOVA используется для т.н. сбалансированных планов (в группах одинаковое количество лиц)
  - GLM используется для любых планов
- Обе процедуры могут рассчитывать откорректированные средние (MANOVA)

# ANOVA (SAS)

PROGRAM EDITOR

Command ==>

```
00001 data new; input x group @@; cards;
00002 12 1 13 1 14 1 15 1 16 1 10 1 19 1 19 1 15 1
00003 2 2 1 2 3 2 3 2 2 2 4 2 3 2 2 2 1 2
00004 8 3 9 3 15 3 15 3 14 3 12 3 13 3 12 3 12 3
00005 12 4 13 4 14 4 15 4 16 4 10 4 19 4 19 4 13 4
00006 ; run;
00007 proc anova data=new; class group; model x=group;
00008 means group /duncan; run;
00009
```

# ANOVA (SAS)

OUTPUT

Command ==>

SAS 22:24 Monday, June 6, 1994 36

## Analysis of Variance Procedure

Dependent Variable: X

Source	DF	Sum of Squares	Mean Square	F Value	Pr > F
Model	3	931.6388889	310.5462963	49.36	0.0001
Error	32	201.3333333	6.2916667		
Corrected Total	35	1132.9722222			

R-Square	C.V.	Root MSE	X Mean
0.822296	22.86063	2.508319	10.9722222

OUTPUT  
Command ==>

SAS 22:24 Monday, June 6, 1

Analysis of Variance Procedure

■ Duncan's Multiple Range Test for variable: X

NOTE: This test controls the type I comparisonwise error rate,  
the experimentwise error rate

Alpha= 0.05 df= 32 MSE= 6.291667

Number of Means      2      3      4  
Critical Range    2.408 2.531 2.614

Means with the same letter are not significantly different.

Duncan Grouping	Mean	N	GROUP
A	14.778	9	1
A			
B	14.556	9	4
B			
B	12.222	9	3
C	2.333	9	2

# ANOVA

- Если необходим анализ зависимых переменных, то существует несколько возможностей:
  - заменить исходные данные на разности и выполнить затем ANOVA на разностях
  - воспользоваться модулем ANOVA/MANOVA в системе Statistica



# Анализ зависимых переменных

General ANOVA/MANOVA

**Variables** | **Covariates** | **OK** | **Cancel**

Independent (factors): **GROUP**  
Dependent: **G-G2**  
Covariates: **none**

**Codes for between-groups factors:** **Selected**

**Repeated measures (within SS) design:** **1 repeated measures factor**

**Nested design:** none

**Random factors:** none

**Isolated control group:** none

**Regression approach (Type I, II, III SS)**

**Open Data** | **SELECT CASES** | **W**

For large main effect and non full-factorial designs, hierarchically nested models or designs with unbalanced nesting, and mixed-model (random effect) designs, see also the Variance Components or Experimental Design modules.

# Анализ зависимых переменных

Summary of all Effects; design: [contr.sta]

Continue... 1-GROUP, 2-RFACTOR1

Effect	df Effect	MS Effect	df Error	MS Error	F	p-level
1	2	493,3962	19	1799,639	,274164	,763164
2	1	8,8904	19	25,699	,345939	,563348
12	2	60,6828	19	25,699	2,361267	,121375

# Анализ зависимых переменных (SAS)

```
PROGRAM EDITOR
Command ==> ■

00001 data prepost; input subj group $ pretest posttest;
00002 cards;
00003 1 C 80 83
00004 2 C 85 86
00005 3 C 83 88
00006 4 T 82 94
00007 5 T 87 93
00008 6 T 84 98
00009 ;
00010 run;
00011 proc anova data=prepost;
00012 class group; model pretest posttest = group / nouni;
00013 repeated time 2 (0 1);
00014 run;
00015
```

OUTPUT

Command ==>

SAS

22:24 Monday, June 6, 1994 5

Analysis of Variance Procedure  
Repeated Measures Analysis of Variance

Manova Test Criteria and Exact F Statistics for  
the Hypothesis of no TIME Effect

H = Anova SS&CP Matrix for TIME    E = Error SS&CP Matrix

S=1    M=-0.5    N=1

Statistic	Value	F	Num DF	Den DF	Pr > F
Wilks' Lambda	0.13216314	26.2656	1	4	0.0069
Pillai's Trace	0.86783686	26.2656	1	4	0.0069
Hotelling-Lawley Trace	6.56640625	26.2656	1	4	0.0069
Roy's Greatest Root	6.56640625	26.2656	1	4	0.0069

Analysis of Variance Procedure  
Repeated Measures Analysis of Variance

Manova Test Criteria and Exact F Statistics for  
the Hypothesis of no TIME\*GROUP Effect

H = Anova SS&CP Matrix for TIME\*GROUP    E = Error SS&CP Matrix

S=1    M=-0.5    N=1

Statistic	Value	F	Num DF	Den DF	Pr > F
Wilks' Lambda	0.32611465	8.2656	1	4	0.0452
Pillai's Trace	0.67388535	8.2656	1	4	0.0452
Hotelling-Lawley Trace	2.06640625	8.2656	1	4	0.0452
Roy's Greatest Root	2.06640625	8.2656	1	4	0.0452

# Анализ зависимых переменных (SAS)

OUTPUT

Command ==>

SAS

22:24 Monday, June 6, 1994 60

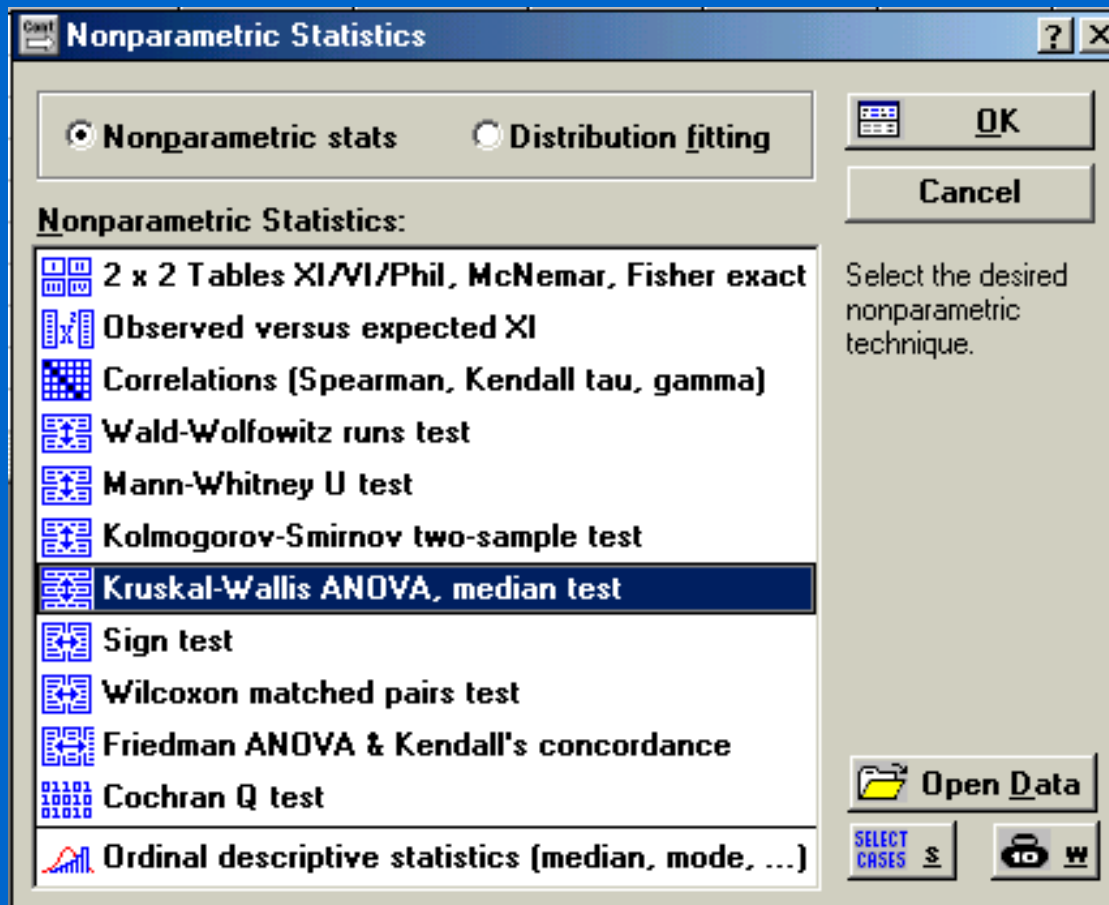
Analysis of Variance Procedure  
Repeated Measures Analysis of Variance  
Tests of Hypotheses for Between Subjects Effects

Source	DF	Anova SS	Mean Square	F Value	Pr > F
GROUP	1	90.750000	90.750000	11.84	0.0263
Error	4	30.666667	7.666667		

# Отклонения от нормального распределения

- В принципе ANOVA достаточно устойчива к небольшим отклонениям от нормального распределения
- Для зависимых переменных - ANOVA по Фридману
- Для независимых переменных - ANOVA по Крускалу-Уоллесу

# Отклонения от нормального распределения



# Другие задачи, которые может решать ANOVA

- Зависимость количественной переменной от нескольких качественных (например, как пол, отношение к курению и цвет волос влияют на САД)
- Совместное действие нескольких качественных переменных (выше ли САД у брюнеток)
- Влияние нескольких качественных и количественных переменных на количественную (MANCOVA)
- НО это уже многомерные методики