河南省肿瘤医院
郑州大学附属肿瘤医院

医学常用统计分析方法及软件实现

郭兰伟 博士
河南省肿瘤防治研究办公室
guolanwei1019@126.com

统计学中基本概念

●资料类型 （Category of Data）

1、计量资料：对每个观察单位某个变量用测量或其他定量方法获得的定量观察结果，一般有计量单位。例如：年龄

2、计数资料：将观察单位按某种属性分组计数的定性观察结果。例如：性别

3、等级资料：将观察单位按某种属性的不同程度分组计数观察资料，特点是具有半定量性质。例如：医院等级

提纲

- T检验
- 方差分析
- 卡方检验
- 秩和检验
- 生存分析

河南省肿瘤医院
郑州大学附属肿瘤医院

T检验

- 单样本t检验
- 独立样本t检验
- 配对t检验

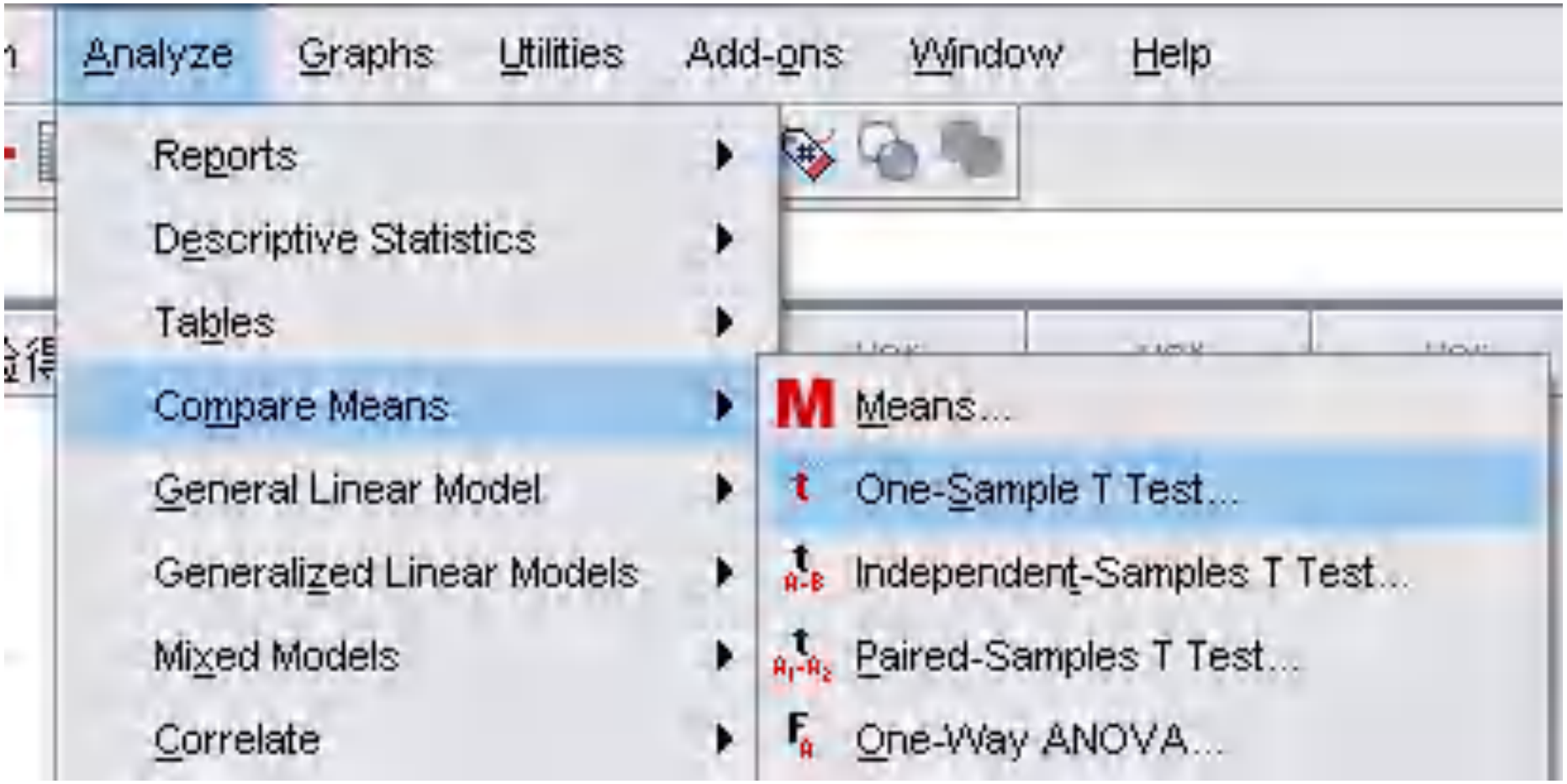
河南省肿瘤医院
郑州大学附属肿瘤医院

一、单样本t检验

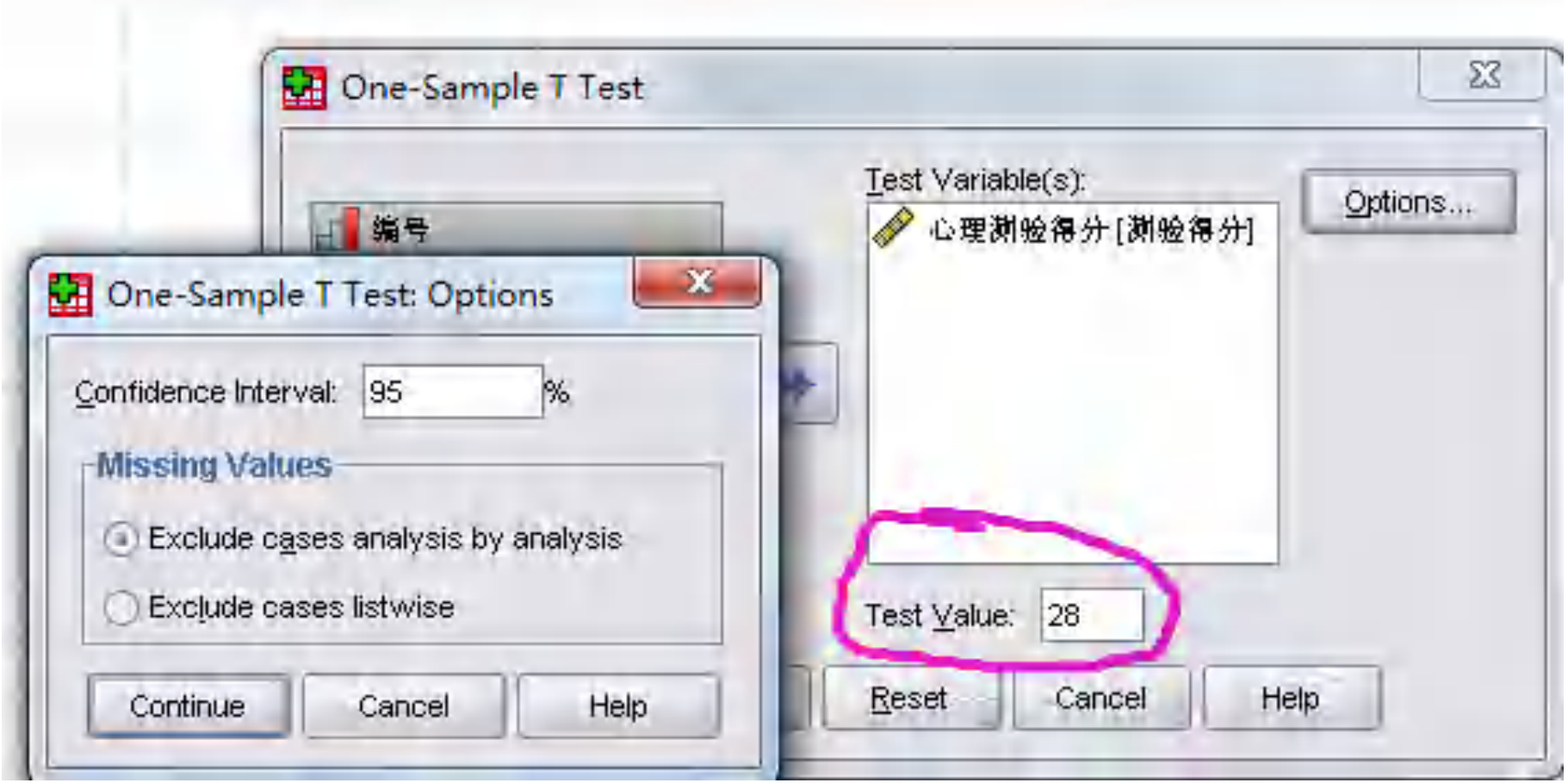
例1

某教师对其所在学校的大一新生进行了心理健康水平普查。从中随机抽取了40名学生的测验得分如下。已知该校历届大一新生该项心理测验的平均分是28分。试问：该40名新生的平均分与该校大一新生的平均分有无显著差异？

单样本t检验——操作步骤1



操作步骤


One-Sample T Test
Test Variable(s):
心理测验得分 [测验得分]
Options...
编号
Test Value: 28
Reset
Cancel
Help
One-Sample T Test: Options
Confidence Interval: 95 %
Missing Values
Exclude cases analysis by analysis
Exclude cases listwise
Continue
Cancel
Help

结果输出

One-Sample Statistics

	N	Mean	Std. Deviation	Std. Error Mean
心理测验得分	40	30.42	5.453	.862

One-Sample Test

	Test Value = 28					
					95% Confidence Interval of the Difference	
	t	df	Sig. (2-tailed)	Mean Difference	Lower	Upper
心理测验得分	2.812	39	.008	2.425	.68	4.17

河南省肿瘤医院
郑州大学附属肿瘤医院

二、两独立样本t检验

例2 分布从两个班级随机抽取12名学生，分析他们某一项心理能力测验分数的平均数是否存在显著性差异。测验分数如下表。

表2 部分学生某心理能力测验分数

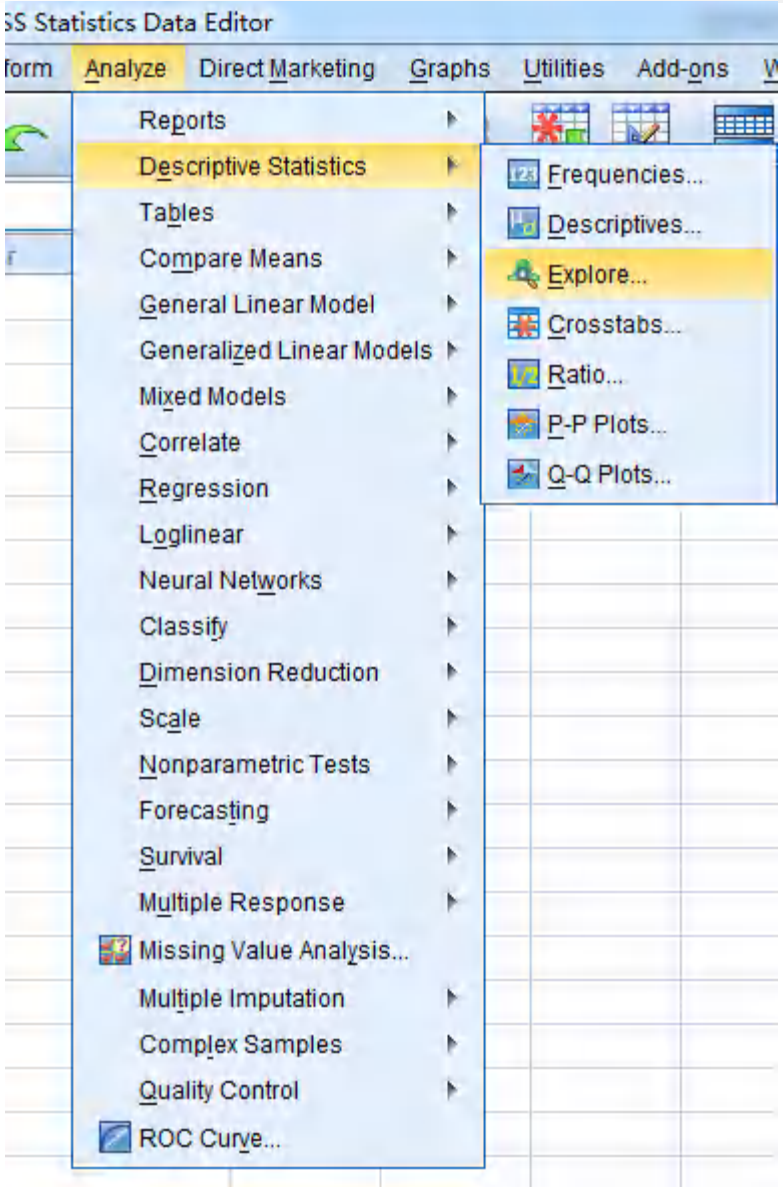
班级1	85	67	83	79	92	90	74	79
	81	63	70	69				
班级2	88	65	68	87	56	78	83	69
	70	90	75	79				

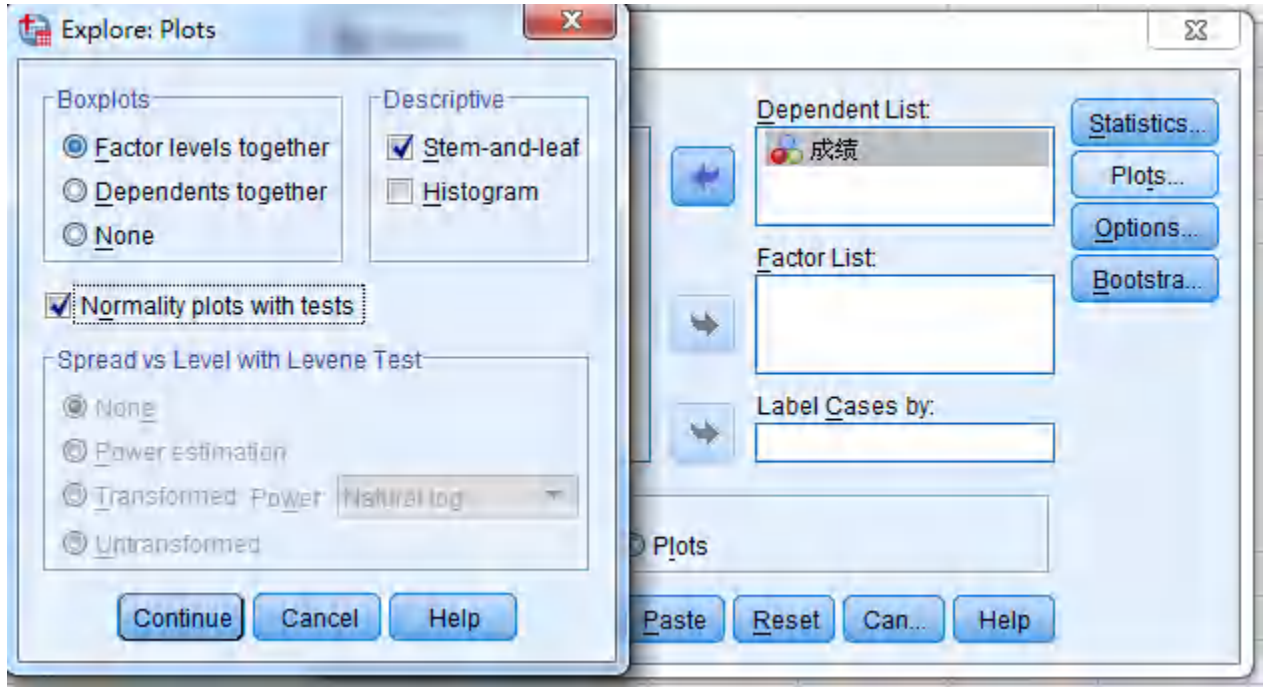
河南省肿瘤医院
郑州大学附属肿瘤医院

1. 数据输入

	班级	成绩
1	1	85
2	1	67
3	1	83
4	1	79
5	1	92
6	1	90
7	1	74
8	1	79
9	1	81
10	1	63
11	1	70
12	1	69
13	2	88
14	2	65

2. 正态性检验





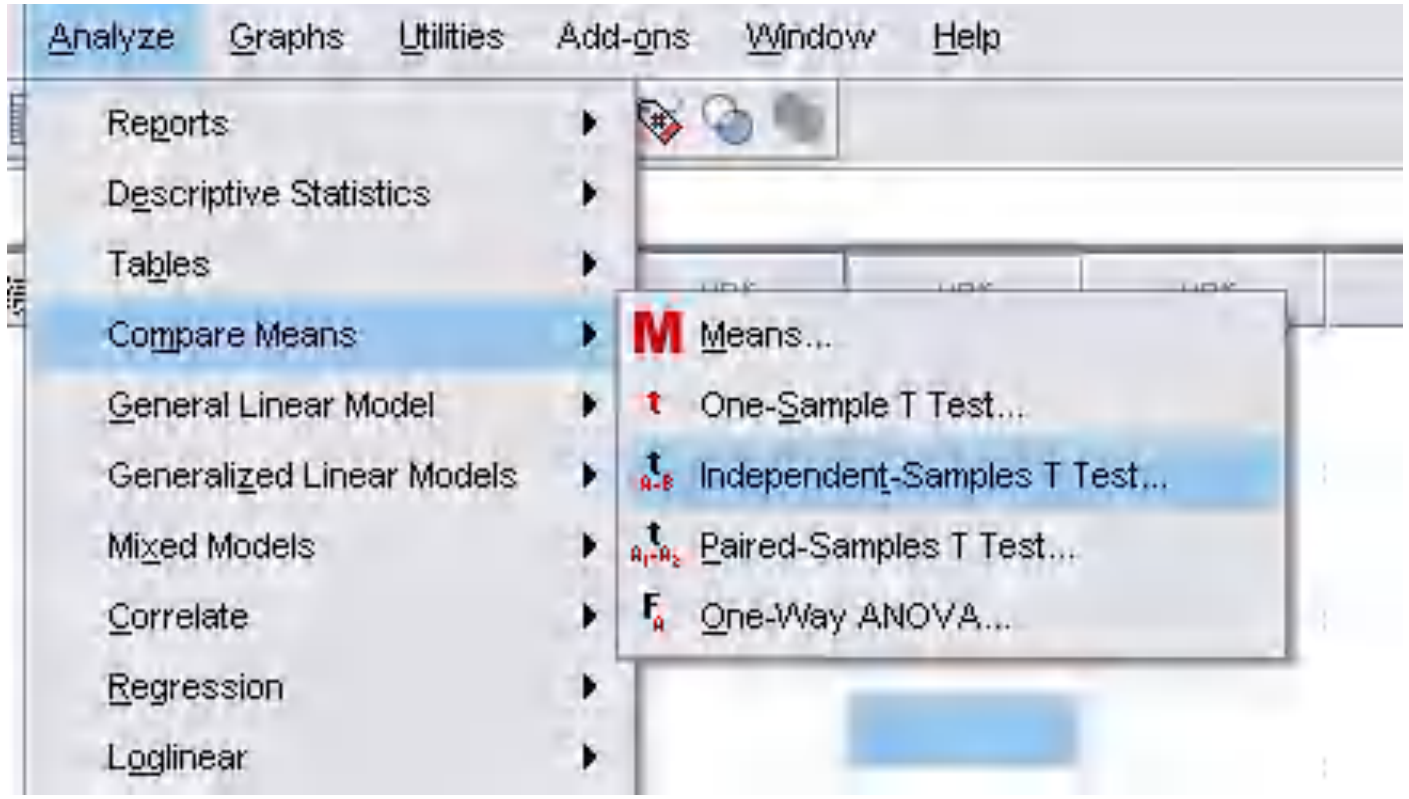
正态性检验

	Kolmogorov-Smirnov[a]			Shapiro-Wilk		
	统计量	df	Sig.	统计量	df	Sig.
成绩	.141	12	.200*	.966	12	.869

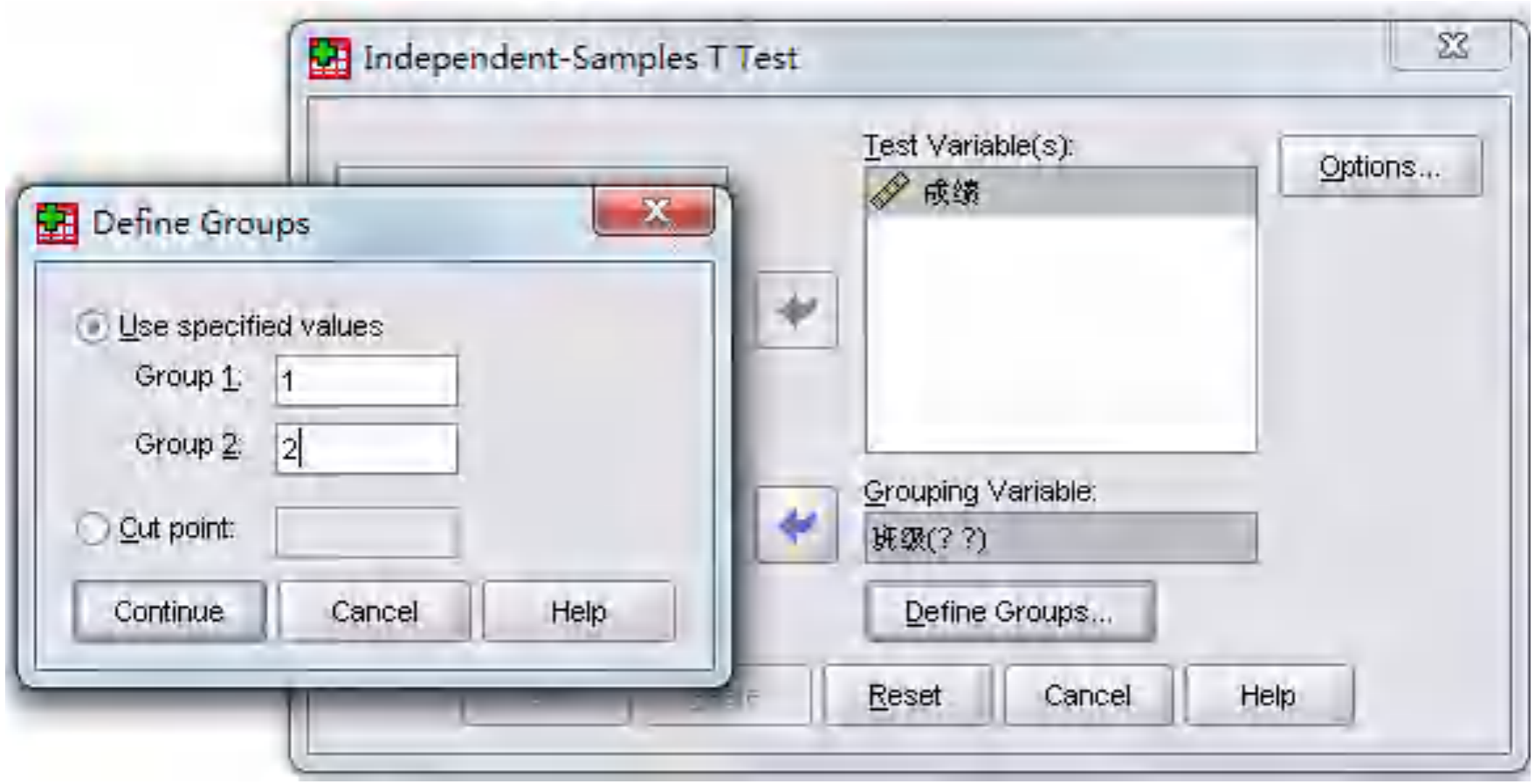
a. Lilliefors 显著水平修正
*. 这是真实显著水平的下限。

河南省肿瘤医院
郑州大学附属肿瘤医院

3. 操作步骤


Analyze
Graphs
Utilities
Add-ons
Window
Help
Reports
Descriptive Statistics
Tables
Compare Means
General Linear Model
Generalized Linear Models
Mixed Models
Correlate
Regression
Loglinear
Means...
One-Sample T Test...
Independent-Samples T Test...
Paired-Samples T Test...
One-Way ANOVA...

3. 操作步骤


Independent-Samples T Test
Test Variable(s):
成绩
Options...
Grouping Variable:
班级(? ?)
Define Groups...
Reset
Cancel
Help
Define Groups
Use specified values
Group 1:
1
Group 2:
2
Cut point:
Continue
Cancel
Help

4. 结果输出

Group Statistics

	班级	N	Mean	Std. Deviation	Std. Error Mean
成绩	1	12	77.67	9.198	2.655
	2	12	75.67	10.413	3.006

Independent Samples Test

		Levene's Test for Equality of Variances		t-test for Equality of Means						
									95% Confidence Interval of the Difference	
		F	Sig.	t	df	Sig. (2-tailed)	Mean Difference	Std. Error Difference	Lower	Upper
成绩	Equal variances assumed	.206	.654	.499	22	.623	2.000	4.011	-6.318	10.318
	Equal variances not assumed			.499	21.670	.623	2.000	4.011	-6.325	10.325

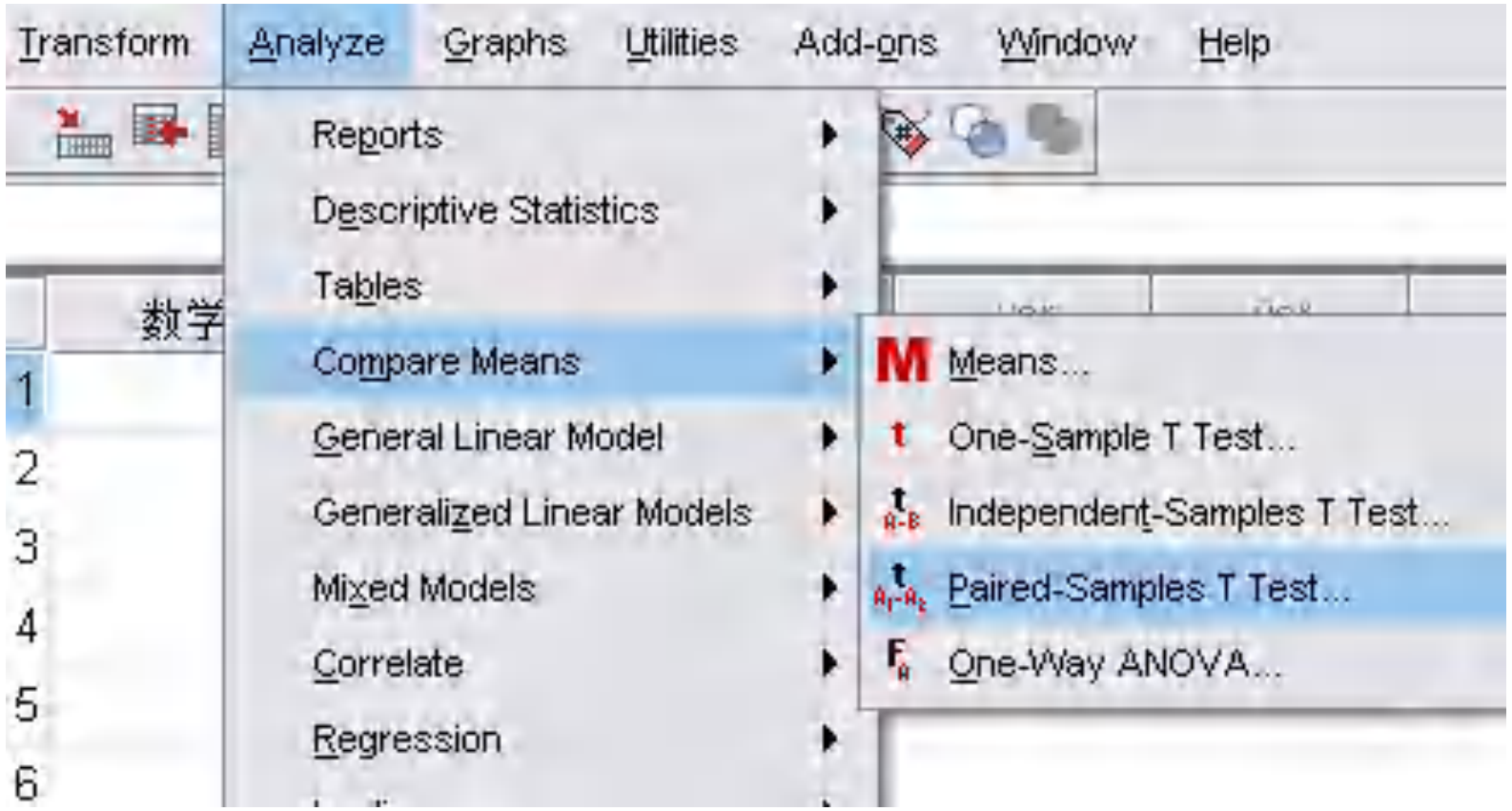
河南省肿瘤医院
郑州大学附属肿瘤医院

三、配对t检验

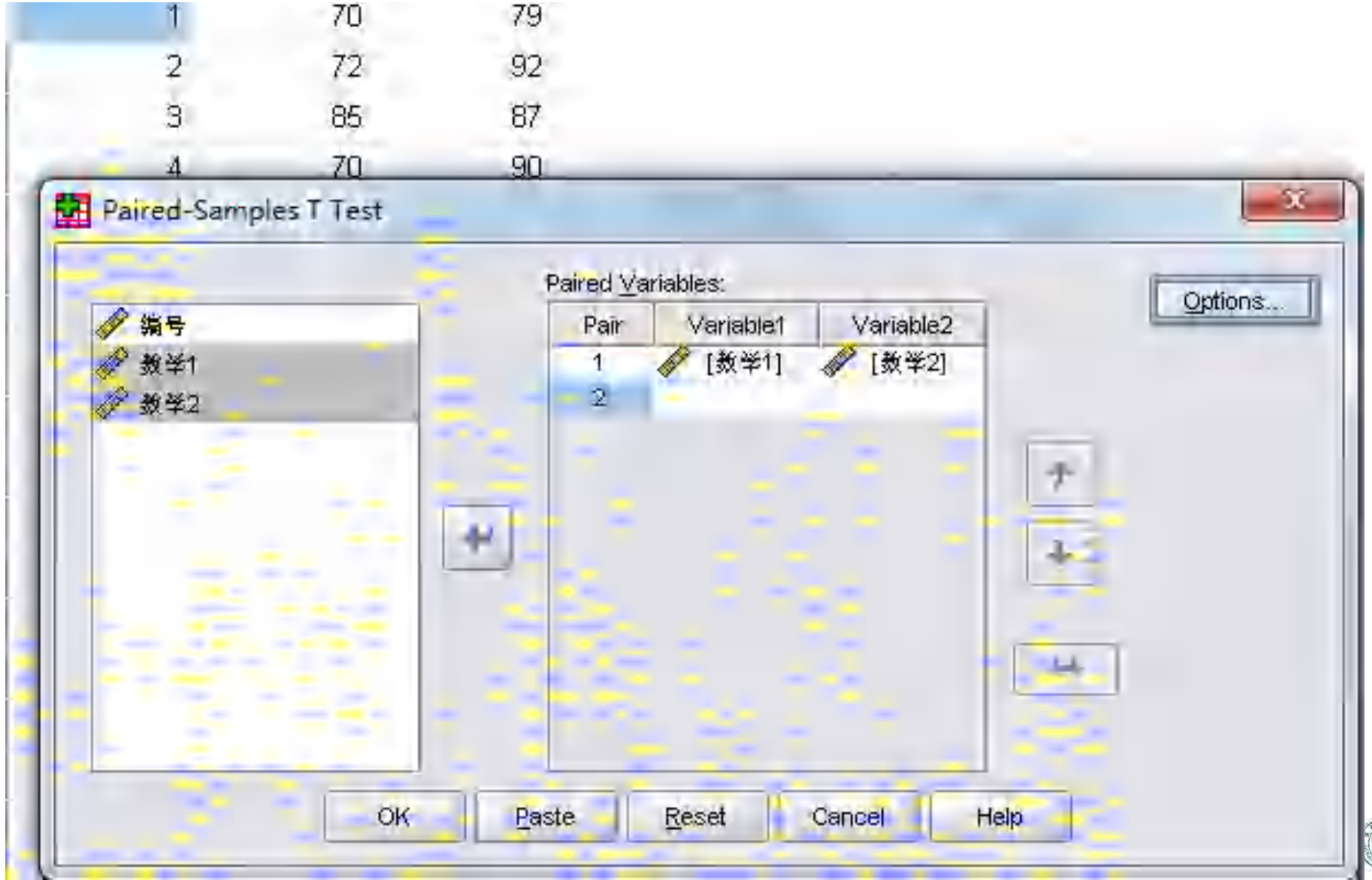
例3

某一小班教学实验班的18名学生接受了一项数学教学实验，即接受新的学习方法的训练。在训练前和训练后，均使用标准化的测试试卷测试了他们的数学成绩。试问该学习方法训练效果是否显著？

1. 操作步骤


Transform
Analyze
Graphs
Utilities
Add-ons
Window
Help
Reports
Descriptive Statistics
Tables
Compare Means
General Linear Model
Generalized Linear Models
Mixed Models
Correlate
Regression
Means...
One-Sample T Test...
Independent-Samples T Test...
Paired-Samples T Test...
One-Way ANOVA...
数学
1
2
3
4
5
6

2. 操作步骤



河南省肿瘤医院
郑州大学附属肿瘤医院

2. 结果输出

Paired Samples Test

	Paired Differences					t	df	Sig. (2-tailed)
				95% Confidence Interval of the Difference				
	Mean	Std. Deviation	Std. Error Mean	Lower	Upper			
Pair 1 数学1 - 数学2	-8.056	9.207	2.170	-12.634	-3.477	-3.712	17	.002

方差分析（ANOVA）

概述

两个独立均数比较，当其满足独立、正态性、方差齐等要求时，可用 t 检验

两个及两个以上独立均数比较，满足上述条件时，可用方差分析，以减少犯第一类错误的概率。其目的是检验多个总体均数是否相等。

河南省肿瘤医院
郑州大学附属肿瘤医院

方差分析的适用条件

- 各处理组样本来自正态总体
- 各样本是相互独立的随机样本
- 各处理组的总体方差相等，即方差齐性

河南省肿瘤医院
郑州大学附属肿瘤医院

例题

例1　某社区随机抽取30名糖尿病患者，IGT异常和正常人进行载脂蛋白（ mmg/L ）测定，结果见下表，问三种人载脂蛋白有无差别？

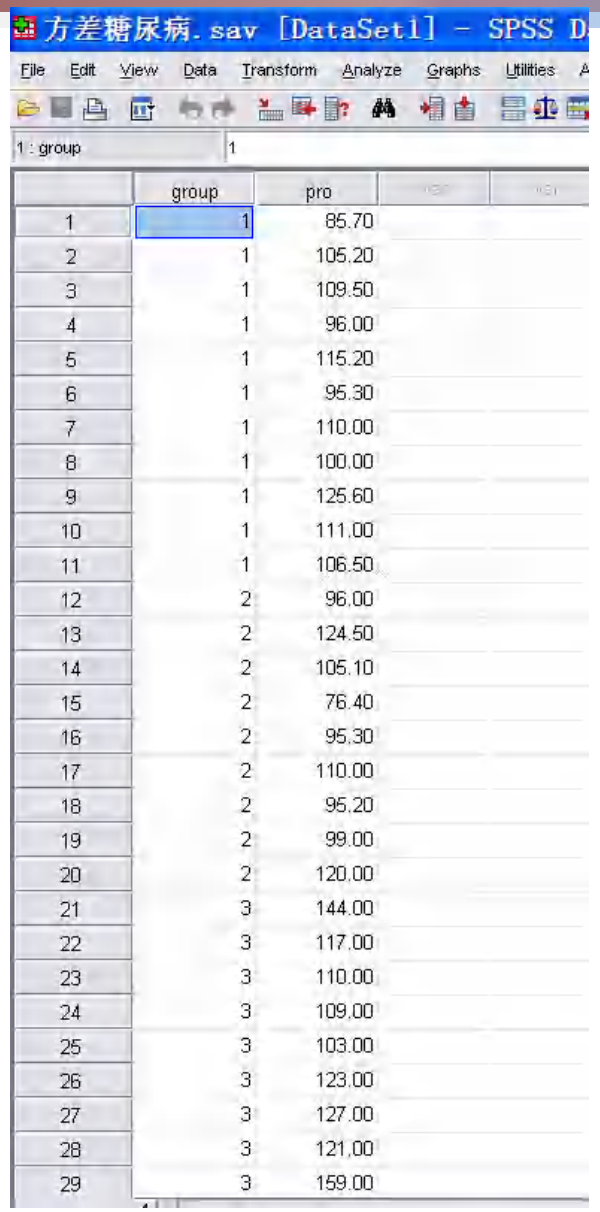
河南省肿瘤医院
郑州大学附属肿瘤医院

例题

糖尿病患者，IGT 异常和正常人载脂蛋白(mmg/L)结果

	糖尿病	IGT 异常	正常人
	85.70	96.00	144.00
	105.20	124.50	117.00
	109.50	105.10	110.00
	96.00	76.40	109.00
	115.20	95.30	103.00
	95.30	110.00	123.00
X_{ij}	110.00	95.20	127.00
	100.00	99.00	121.00
	125.60	120.00	159.00
	111.00		115.00
	106.50		

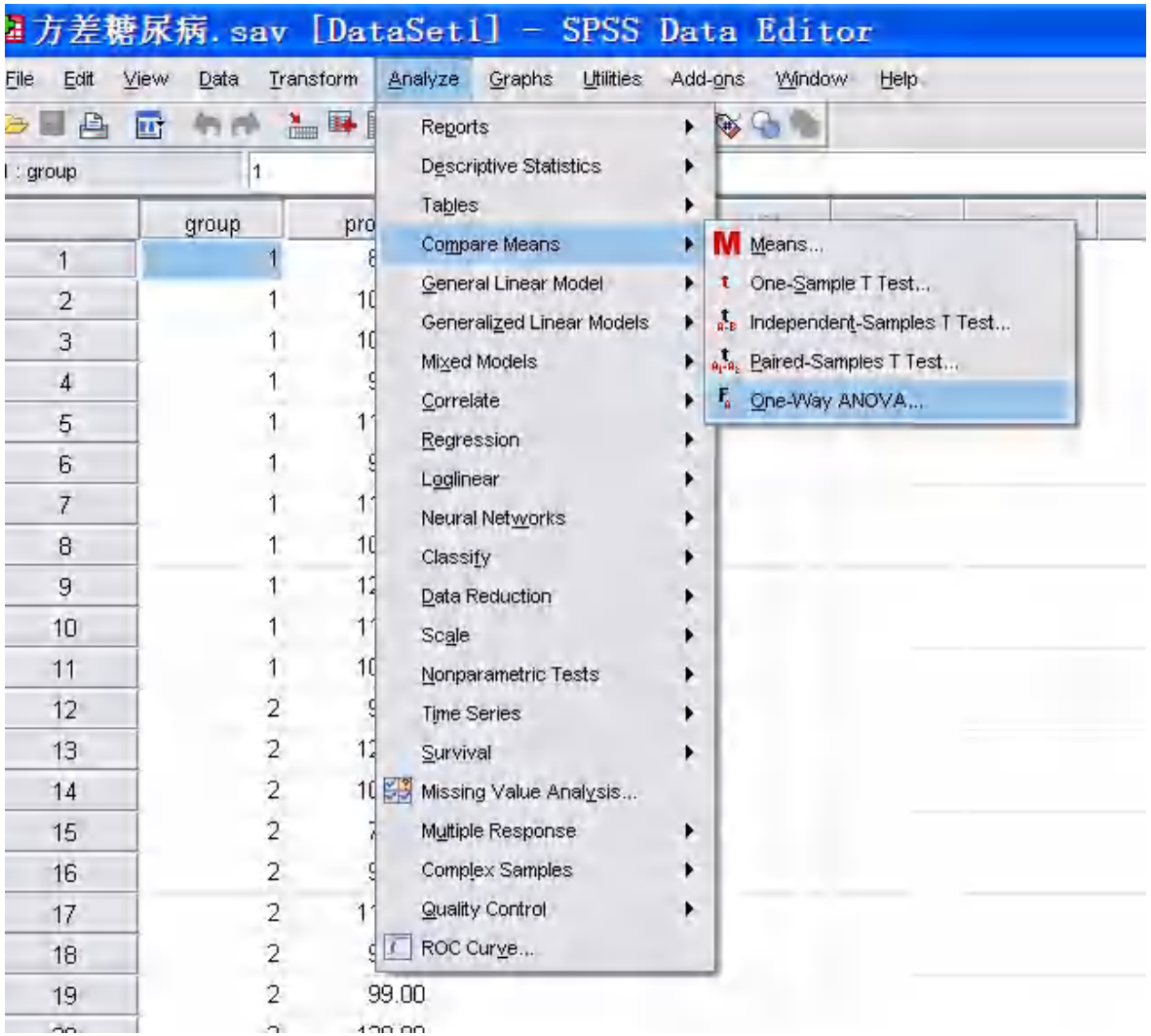
录入数据



	group	pro
1	1	85.70
2	1	105.20
3	1	109.50
4	1	96.00
5	1	115.20
6	1	95.30
7	1	110.00
8	1	100.00
9	1	125.60
10	1	111.00
11	1	106.50
12	2	96.00
13	2	124.50
14	2	105.10
15	2	76.40
16	2	95.30
17	2	110.00
18	2	95.20
19	2	99.00
20	2	120.00
21	3	144.00
22	3	117.00
23	3	110.00
24	3	109.00
25	3	103.00
26	3	123.00
27	3	127.00
28	3	121.00
29	3	159.00

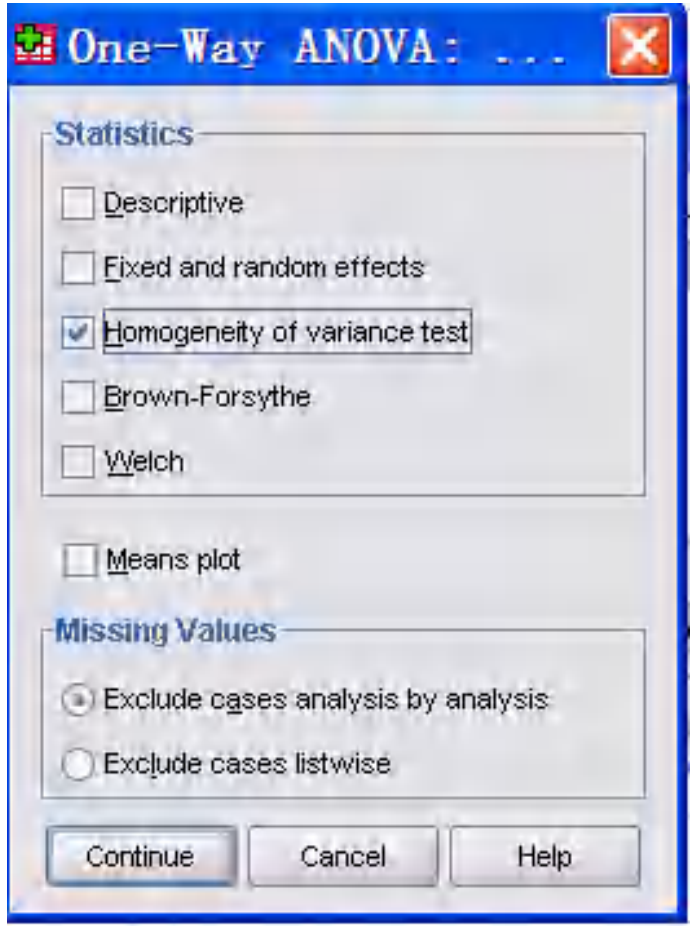
河南省肿瘤医院
郑州大学附属肿瘤医院

SPSS操作

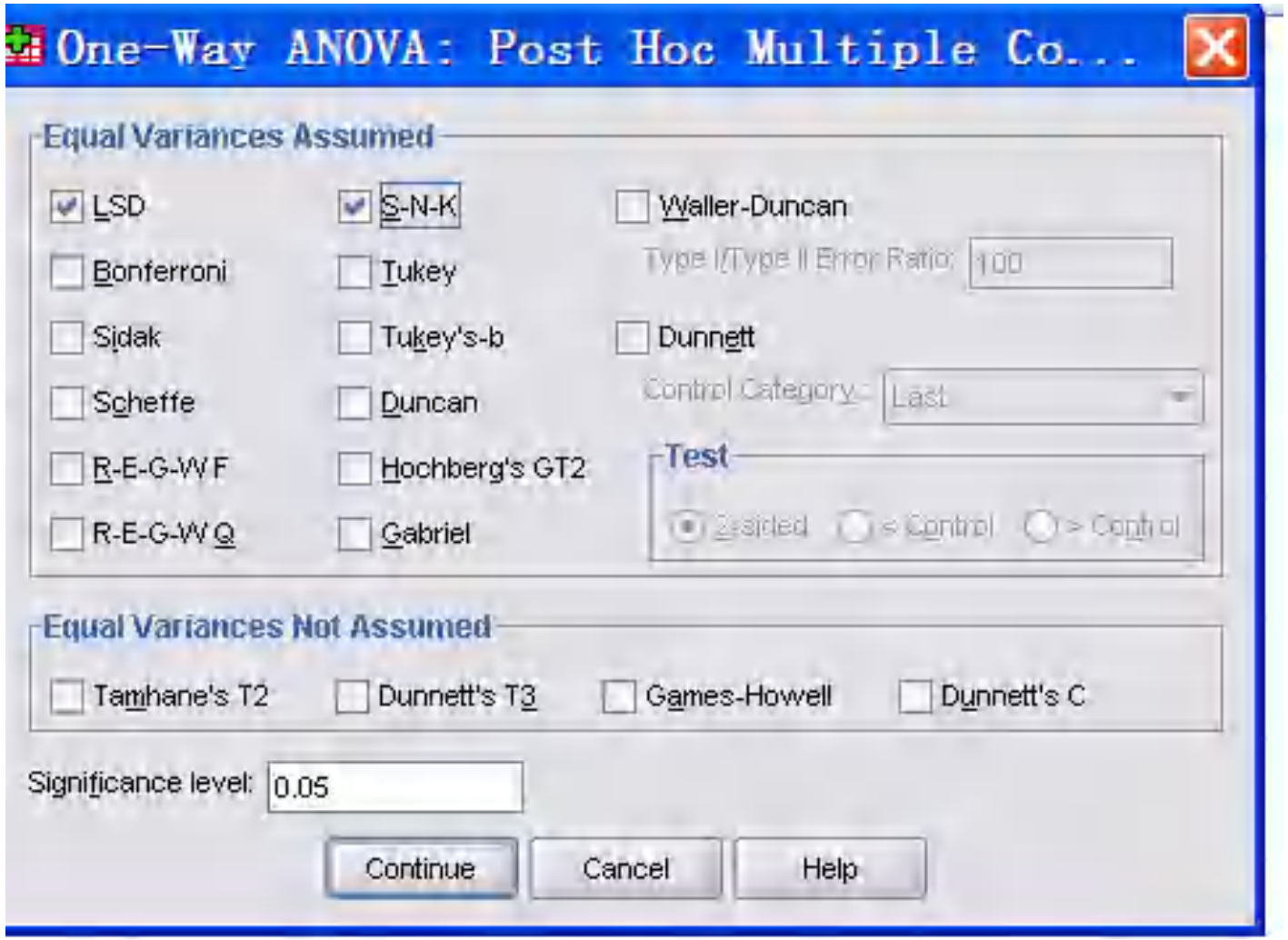

方差糖尿病.sav [DataSet1] - SPSS Data Editor
File Edit View Data Transform Analyze Graphs Utilities Add-ons Window Help
Reports
Descriptive Statistics
Tables
Compare Means
General Linear Model
Generalized Linear Models
Mixed Models
Correlate
Regression
Loglinear
Neural Networks
Classify
Data Reduction
Scale
Nonparametric Tests
Time Series
Survival
Missing Value Analysis...
Multiple Response
Complex Samples
Quality Control
ROC Curve...
Means...
One-Sample T Test...
Independent-Samples T Test...
Paired-Samples T Test...
One-Way ANOVA...
group
99.00

SPSS操作

One-Way ANOVA
Dependent List:
pro
Contrasts...
Post Hoc...
Options...
Factor:
group
OK
Paste
Reset
Cancel
Help


One-Way ANOVA: ...
Statistics
Descriptive
Fixed and random effects
Homogeneity of variance test
Brown-Forsythe
Welch
Means plot
Missing Values
Exclude cases analysis by analysis
Exclude cases listwise
Continue
Cancel
Help

SPSS操作



河南省肿瘤医院
郑州大学附属肿瘤医院

SPSS结果

Test of Homogeneity of Variances

Levene			
.615	2	27	.548

F=0. 615， P=0. 548>0. 10， 方差齐

河南省肿瘤医院
郑州大学附属肿瘤医院

SPSS结果

ANOVA

	Sum of				
Between Groups	2384.026	2	1192.013	5.854	.008
Within Groups	5497.836	27	203.624		
Total	7881.862	29			

方差分析表

F=5. 854， P=0. 008<0. 05，

各组人载脂蛋白不同，常需进行两两比较

河南省肿瘤医院
郑州大学附属肿瘤医院

SPSS结果

Multiple Comparisons

		Mean Difference				
1	2	3.0657	6.4137	.637	-10.0943	16.2256
	3	-17.3455*	6.2349	.010	-30.1383	-4.5526
2	1	-3.0657	6.4137	.637	-16.2256	10.0943
	3	-20.4111*	6.5565	.004	-33.8639	-6.9584
3	1	17.3455*	6.2349	.010	4.5526	30.1383
	2	20.4111*	6.5565	.004	6.9584	33.8639

*.

LSD法两两比较显示，糖尿病人与正常人，糖耐量异常者与正常人之间载脂蛋白差异有统计学意义。尚不能认为糖耐量异常者与糖尿病人之间载脂蛋白有差异。

SPSS结果

pro

	group	N	Subset for alpha = 0.05	
			1	2
Student-Newman-Keuls[a]	2	9	102.3889	
	1	11	105.4545	
	3	10		122.8000
	Sig.		.636	1.000

Means for groups in homogeneous subsets are displayed.

a. Uses Harmonic Mean Sample Size = 9.933.

S-N-K法显示，正常人与糖尿病人、糖耐量异常者载脂蛋白差异有统计学意义。尚不能认为糖耐量异常者与糖尿病人之间载脂蛋白有差异。

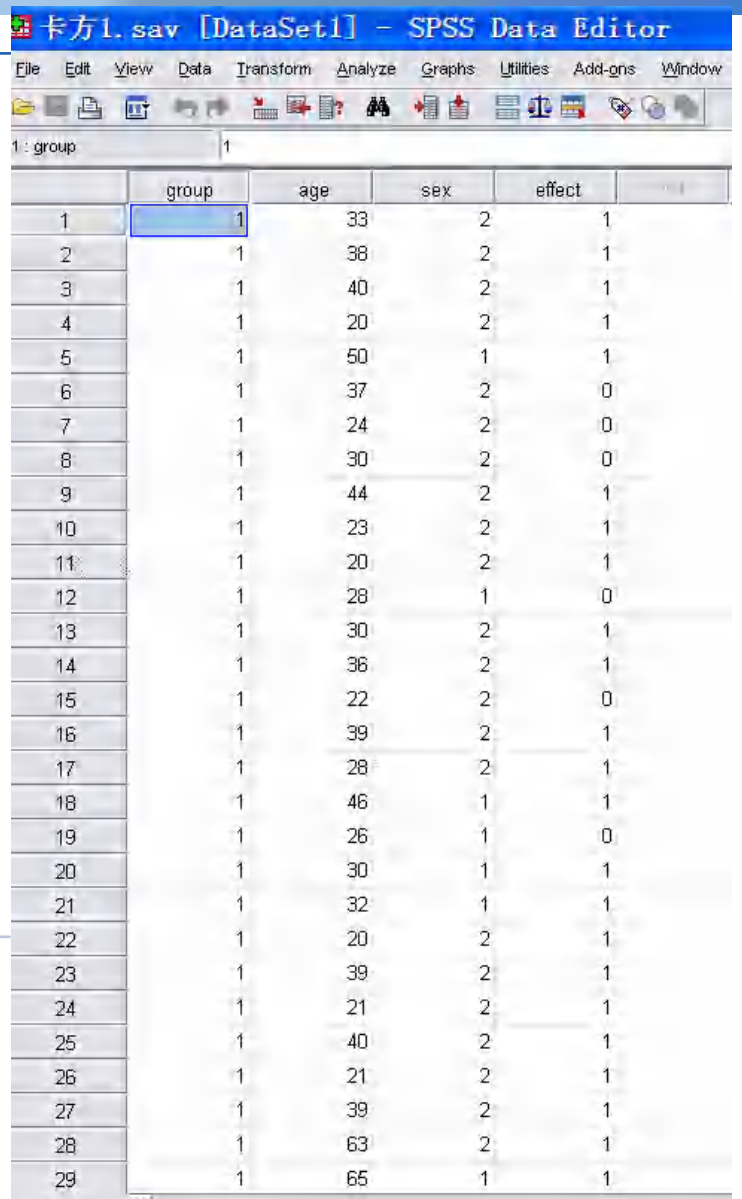
卡方检验

一、两个/多个率（构成比）比较的χ^2检验

例题

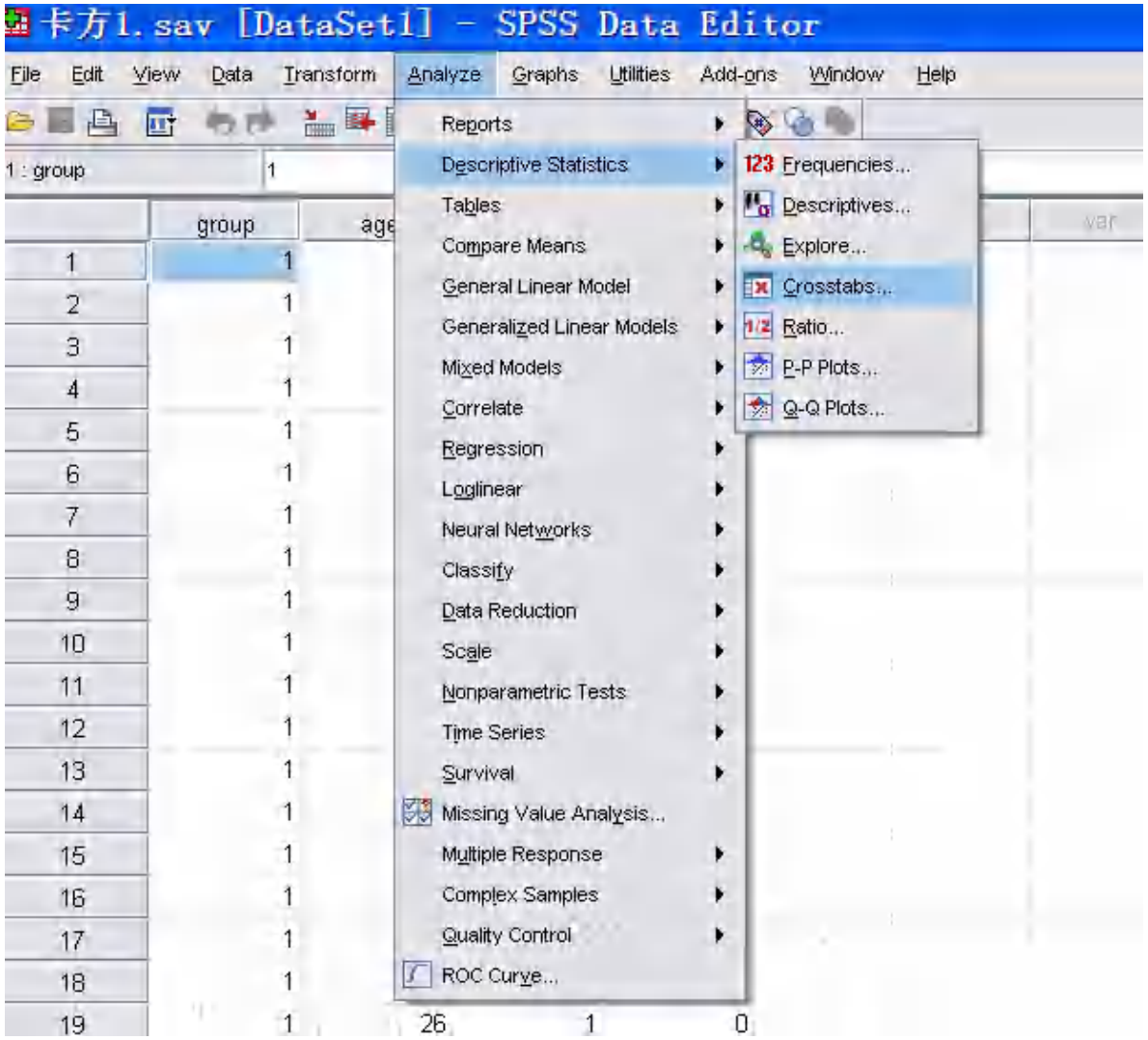
例1

为评价某药治疗肿瘤的疗效，将129名患者随机分到药物组和对照组。试评价两种药物的疗效有无差别。

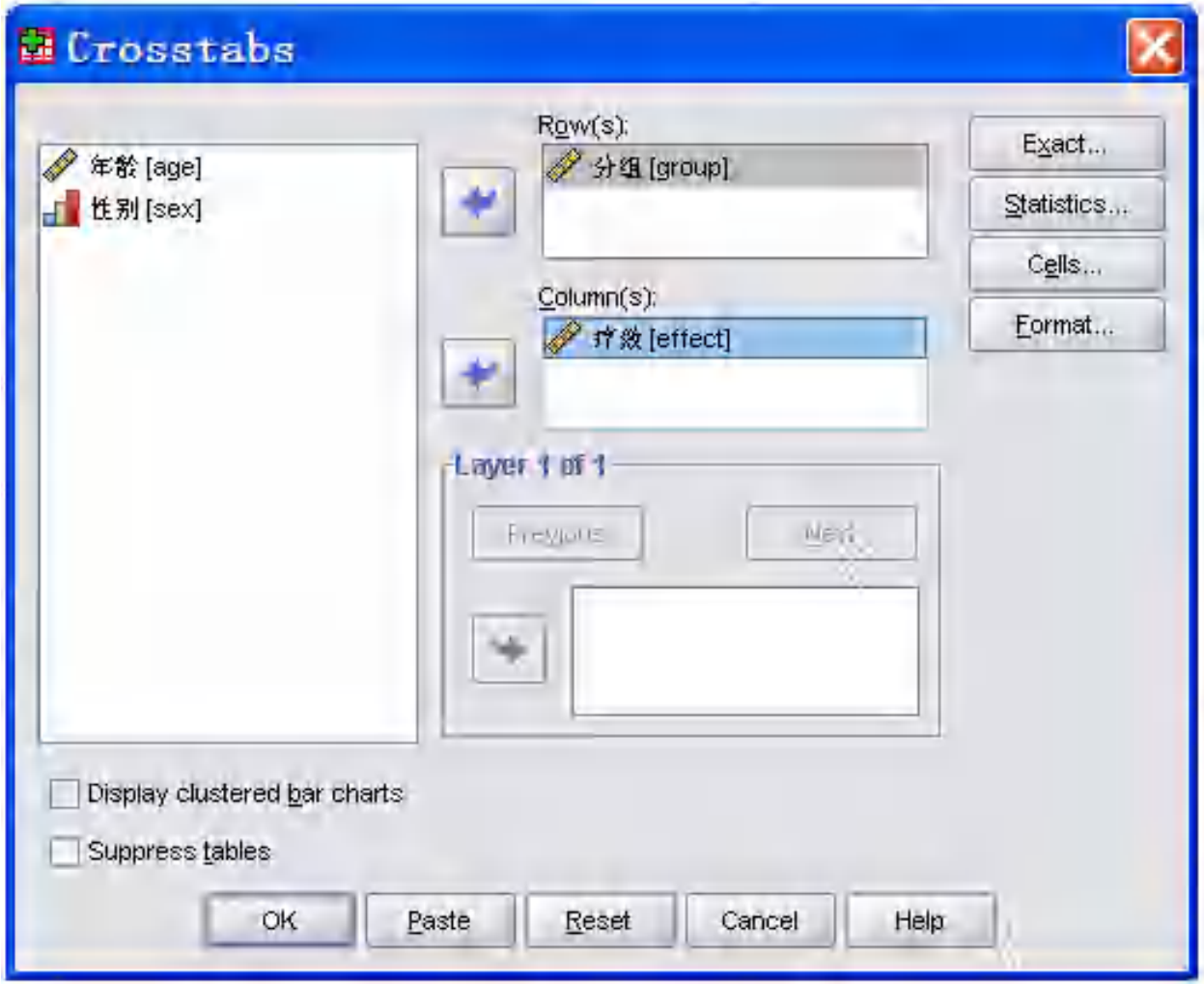


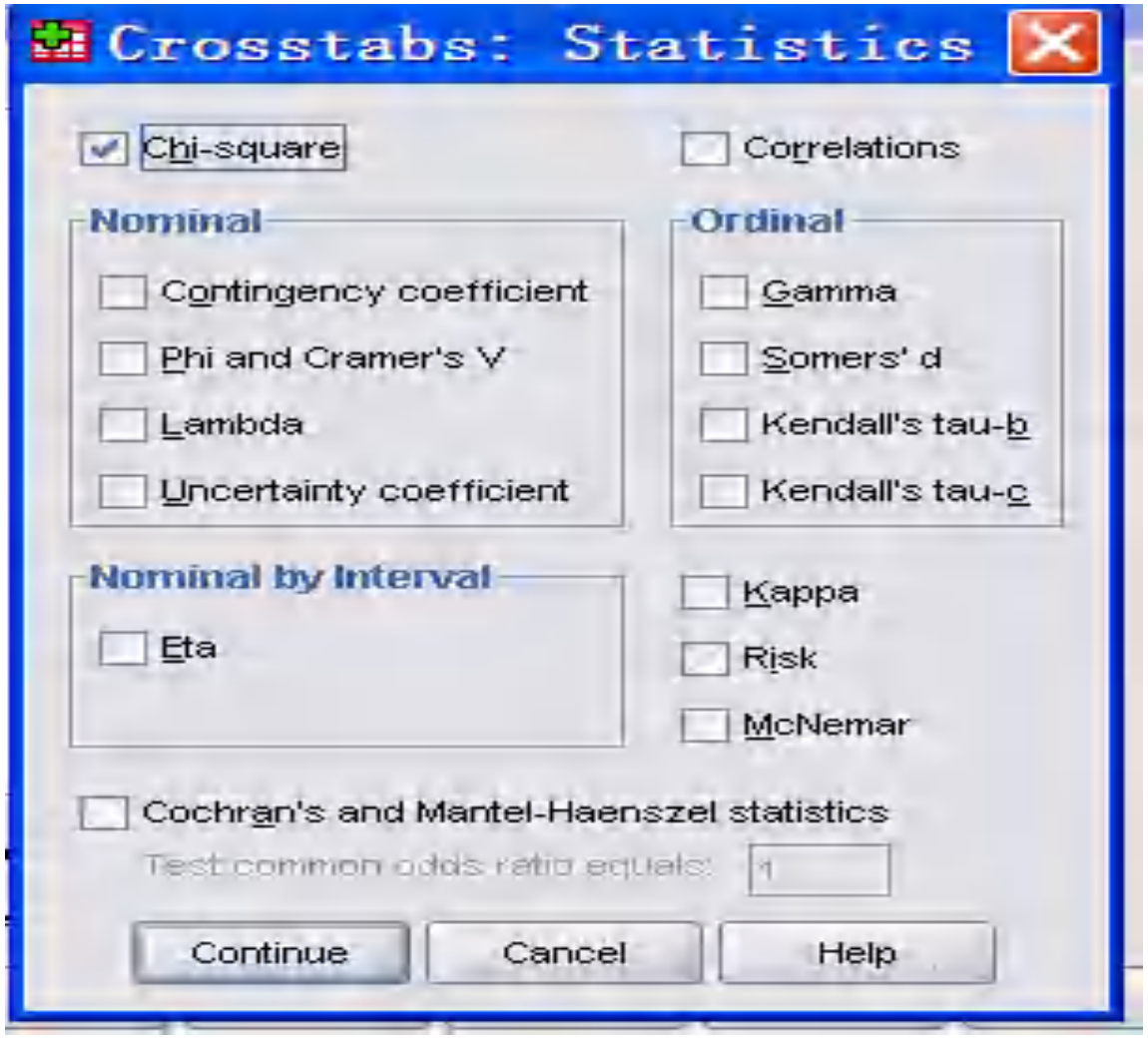
	group	age	sex	effect
1	1	33	2	1
2	1	38	2	1
3	1	40	2	1
4	1	20	2	1
5	1	50	1	1
6	1	37	2	0
7	1	24	2	0
8	1	30	2	0
9	1	44	2	1
10	1	23	2	1
11	1	20	2	1
12	1	28	1	0
13	1	30	2	1
14	1	36	2	1
15	1	22	2	0
16	1	39	2	1
17	1	28	2	1
18	1	46	1	1
19	1	26	1	0
20	1	30	1	1
21	1	32	1	1
22	1	20	2	1
23	1	39	2	1
24	1	21	2	1
25	1	40	2	1
26	1	21	2	1
27	1	39	2	1
28	1	63	2	1
29	1	65	1	1

SPSS操作


卡方1.sav [DataSet1] - SPSS Data Editor
File Edit View Data Transform Analyze Graphs Utilities Add-ons Window Help
1 : group
group
age
var
Reports
Descriptive Statistics
Tables
Compare Means
General Linear Model
Generalized Linear Models
Mixed Models
Correlate
Regression
Loglinear
Neural Networks
Classify
Data Reduction
Scale
Nonparametric Tests
Time Series
Survival
Missing Value Analysis...
Multiple Response
Complex Samples
Quality Control
ROC Curve...
Frequencies...
Descriptives...
Explore...
Crosstabs...
Ratio...
P-P Plots...
Q-Q Plots...

SPSS操作


Crosstabs
年龄 [age]
性别 [sex]
Row(s):
分组 [group]
Column(s):
疗效 [effect]
Exact...
Statistics...
Cells...
Format...
Layer 1 of 1
Previous
Next
Display clustered bar charts
Suppress tables
OK
Paste
Reset
Cancel
Help


Crosstabs: Statistics
Chi-square
Correlations
Nominal
Contingency coefficient
Phi and Cramer's V
Lambda
Uncertainty coefficient
Ordinal
Gamma
Somers' d
Kendall's tau-b
Kendall's tau-c
Nominal by Interval
Eta
Kappa
Risk
McNemar
Cochran's and Mantel-Haenszel statistics
Test common odds ratio equals:
1
Continue
Cancel
Help

Chi-Square Tests

			Asymp. Sig.	Exact Sig.	Exact Sig.
Pearson Chi-Square	1.232[b]	1	.267		
Continuity Correction[a]	.798	1	.372		
Likelihood Ratio	1.247	1	.264		
Fisher's Exact Test				.287	.186
Linear-by-Linear Association	1.223	1	.269		
N of Valid Cases	129				

a.

b.

0 cells (.0%) have expected count less than 5. The minimum expected count is

本例$\chi 2$ = 1. 232, P=0. 267>0. 05,
尚不能认为两组疗效有差异。

河南省肿瘤医院
郑州大学附属肿瘤医院

四格表卡方检验应用条件

$n>40$，且所有 $T\geqslant5$时	Pearson $\chi2$值
$n>40$，但有$1<T<5$时	用连续校正$\chi2$值
$n\leqslant40$，或$T\leqslant1$时	确切概率值

二、配对设计两样本率比较的χ^2检验

例题

例2

有28份咽喉涂抹标本，每份标本分别接种于甲\乙两种培养基上，观察白喉菌的生长情况. 问两种培养基的结果有无差异?

例题

甲、乙两种培养基白喉菌的生长情况

甲培养基	乙培养基		合计
	＋	－	
＋	11	9	20
－	1	7	8
合计	12	15	28

河南省肿瘤医院
郑州大学附属肿瘤医院

SPSS操作

	甲培养基	乙培养基	例数
1	1.00	1.00	11.00
2	1.00	0.00	9.00
3	0.00	1.00	1.00
4	0.00	0.00	7.00
5			
6			

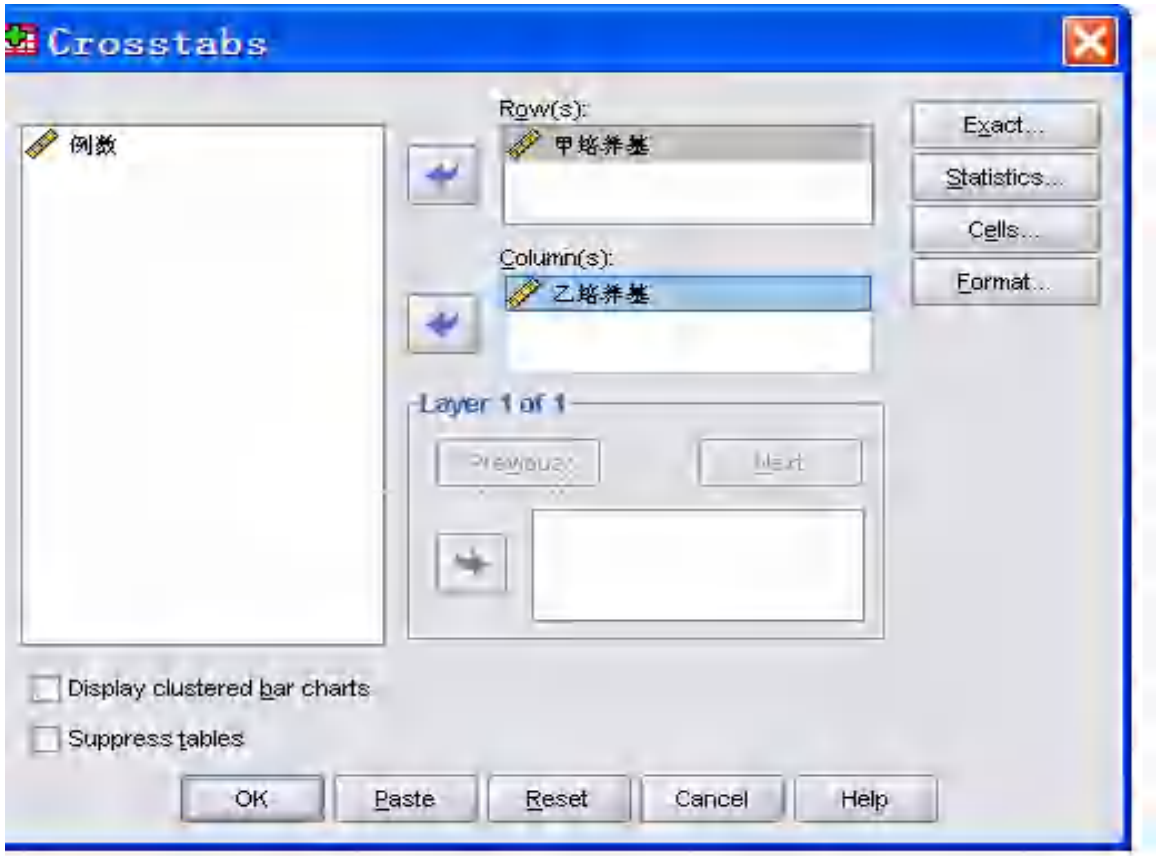
- 数据加权处理

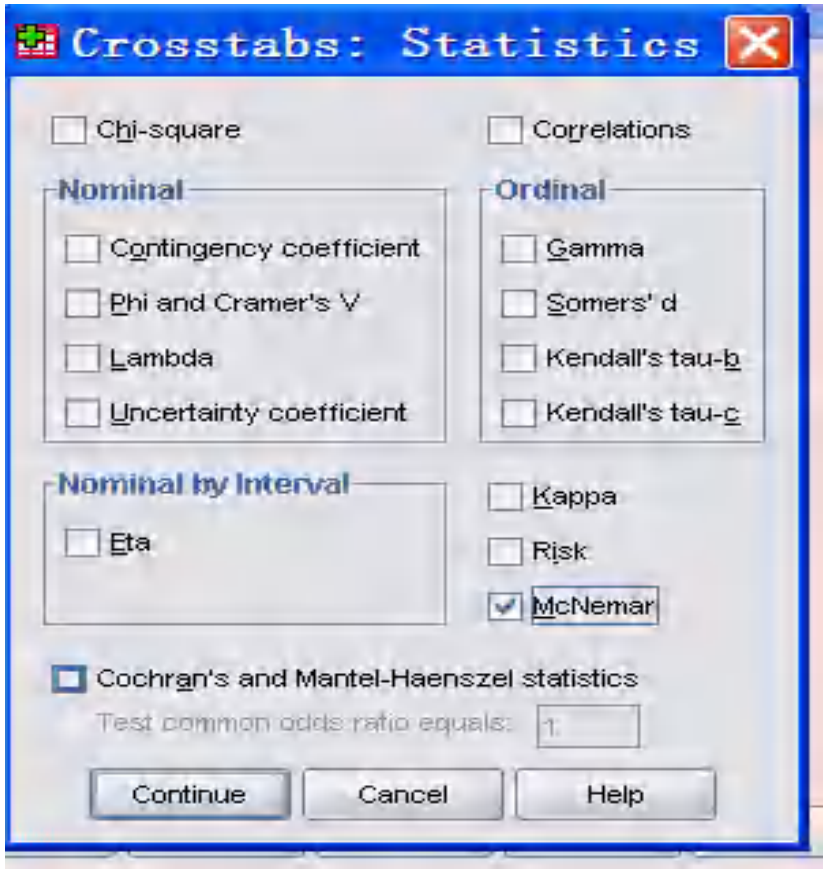
  Data□weight case

  weight case by : 例数

  OK

SPSS操作


Crosstabs
例数
Row(s):
甲培养基
Column(s):
乙培养基
Exact...
Statistics...
Cells...
Format...
Layer 1 of 1
Previous
Next
Display clustered bar charts
Suppress tables
OK
Paste
Reset
Cancel
Help


Crosstabs: Statistics
Chi-square
Correlations
Nominal
Contingency coefficient
Phi and Cramer's V
Lambda
Uncertainty coefficient
Ordinal
Gamma
Somers' d
Kendall's tau-b
Kendall's tau-c
Nominal by Interval
Eta
Kappa
Risk
McNemar
Cochran's and Mantel-Haenszel statistics
Test common odds ratio equals:
1
Continue
Cancel
Help

SPSS结果

Chi-Square Tests

		Exact Sig.
McNemar Test		.021[a]
N of Valid Cases	28	

a.

利用二项分布原理，计算双侧精确概率P=0. 021，可认为两种培养基阳性率有差异，甲培养基阳性率较高。

河南省肿瘤医院
郑州大学附属肿瘤医院

三 有序分类资料的比较

1、等级资料的两个样本比较

例3

研究慢性支气管炎病人痰液中某一类型的白细胞数是否高于正常人，选择24名正常人和44名慢性支气管炎病人的痰液标本做白细胞检查，见下表。

例题

正常人和慢性支气管炎病人痰液的白细胞结果比较

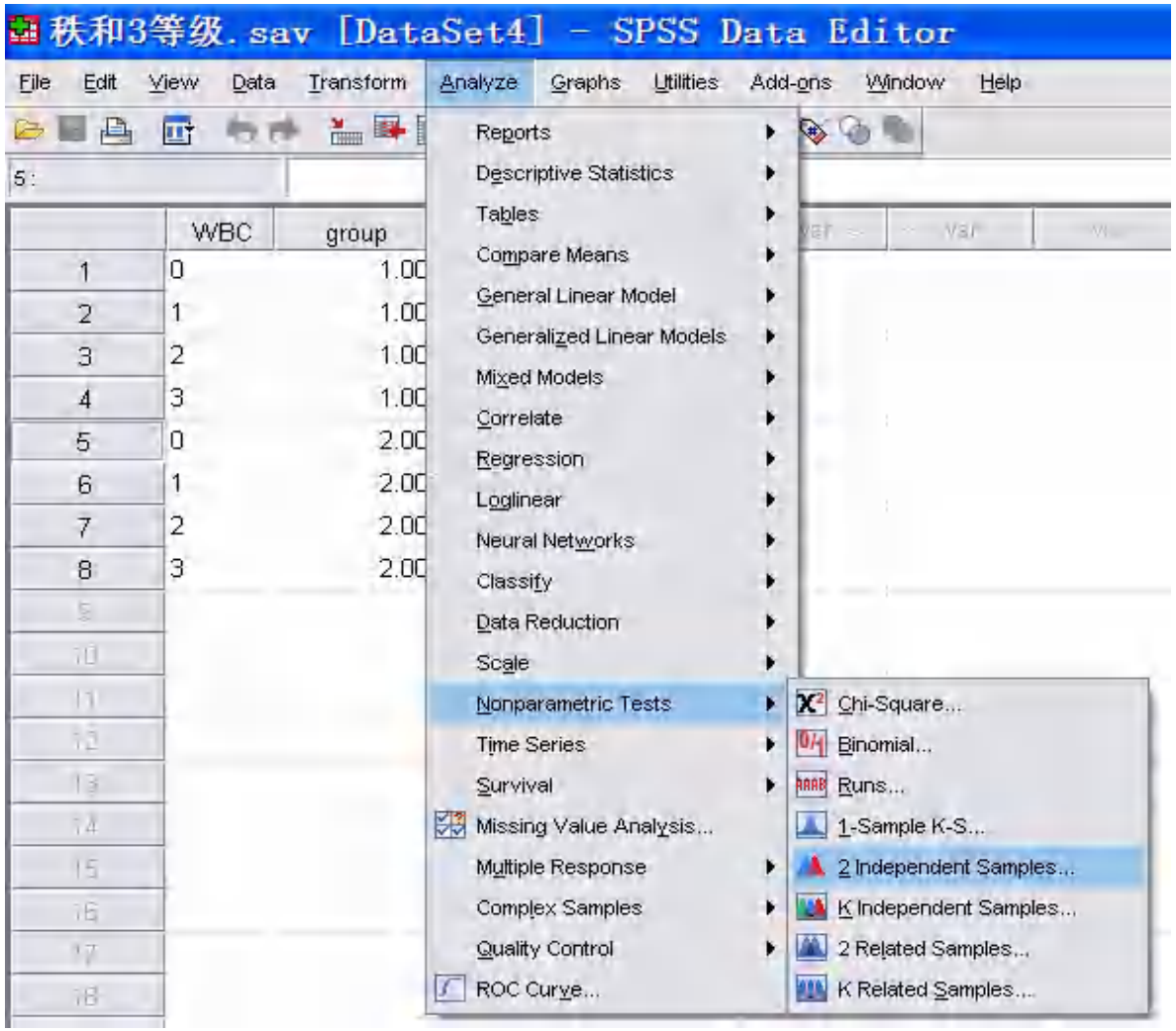
分组	白细胞				合计
	-	+	++	+++	
正常人	11	10	3	0	24
病人	5	18	16	5	44
合计	16	28	19	5	68

河南省肿瘤医院
郑州大学附属肿瘤医院

SPSS操作

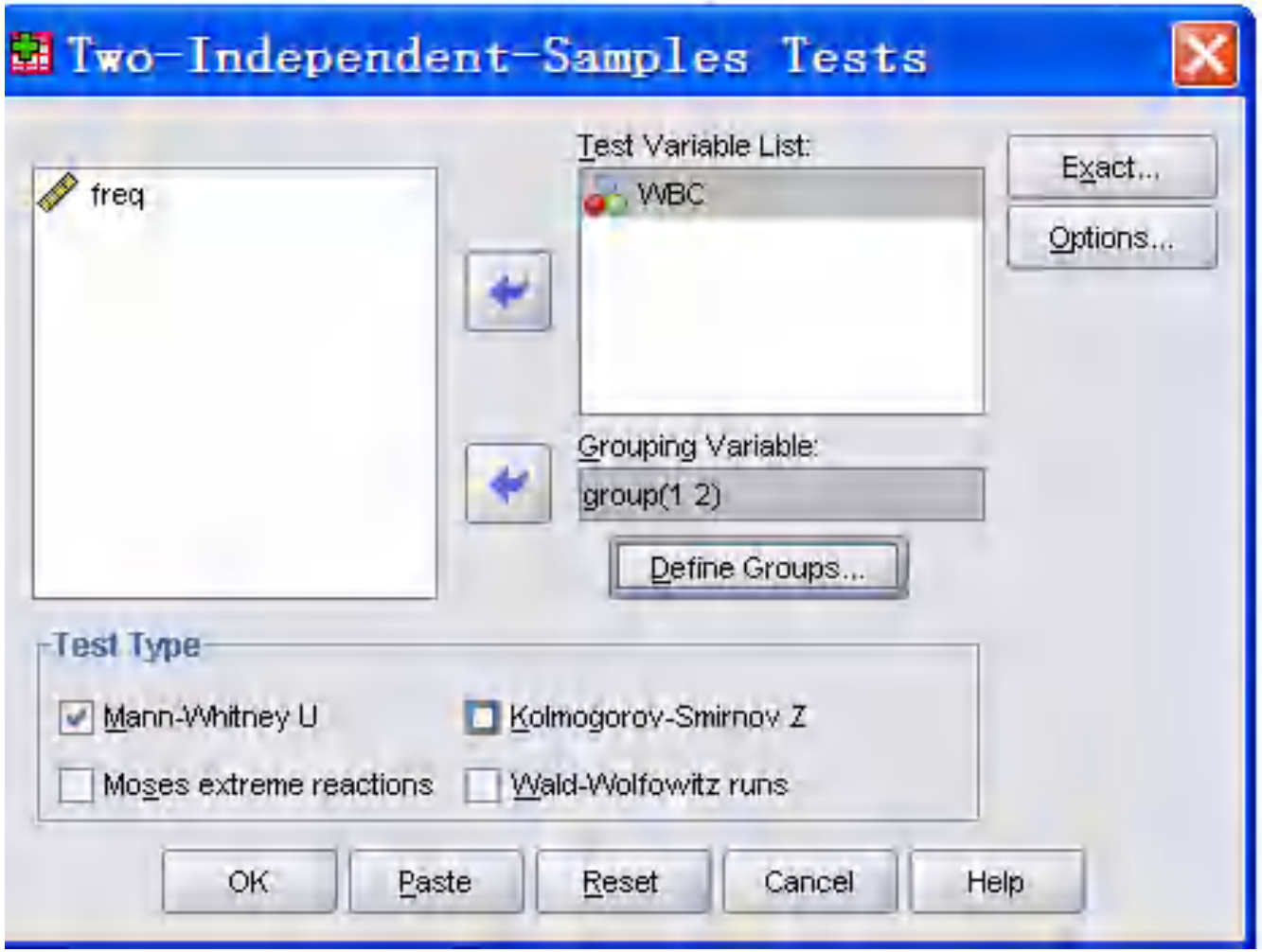
	WBC	group	freq	var
1	0	1.00	11.00	
2	1	1.00	10.00	
3	2	1.00	3.00	
4	3	1.00	0.00	
5	0	2.00	5.00	
6	1	2.00	18.00	
7	2	2.00	16.00	
8	3	2.00	5.00	

SPSS操作



河南省肿瘤医院
郑州大学附属肿瘤医院

SPSS操作


Two-Independent-Samples Tests
freq
Test Variable List:
WBC
Exact...
Options...
Grouping Variable:
group(1 2)
Define Groups...
Test Type
Mann-Whitney U
Kolmogorov-Smirnov Z
Moses extreme reactions
Wald-Wolfowitz runs
OK
Paste
Reset
Cancel
Help

Test Statistics[a]

	WBC
Mann-Whitney U	260.500
Wilcoxon W	560.500
Z	-3.628
Asymp. Sig. (2-tailed)	.000

a. Grouping Variable: group

- Z= 3.628，P=0.000<0.05，按 α= 0.05水平，拒绝$H0$，认为正常人和慢性支气管炎病人痰液此类白细胞的检查结果不相同。

河南省肿瘤医院 郑州大学附属肿瘤医院

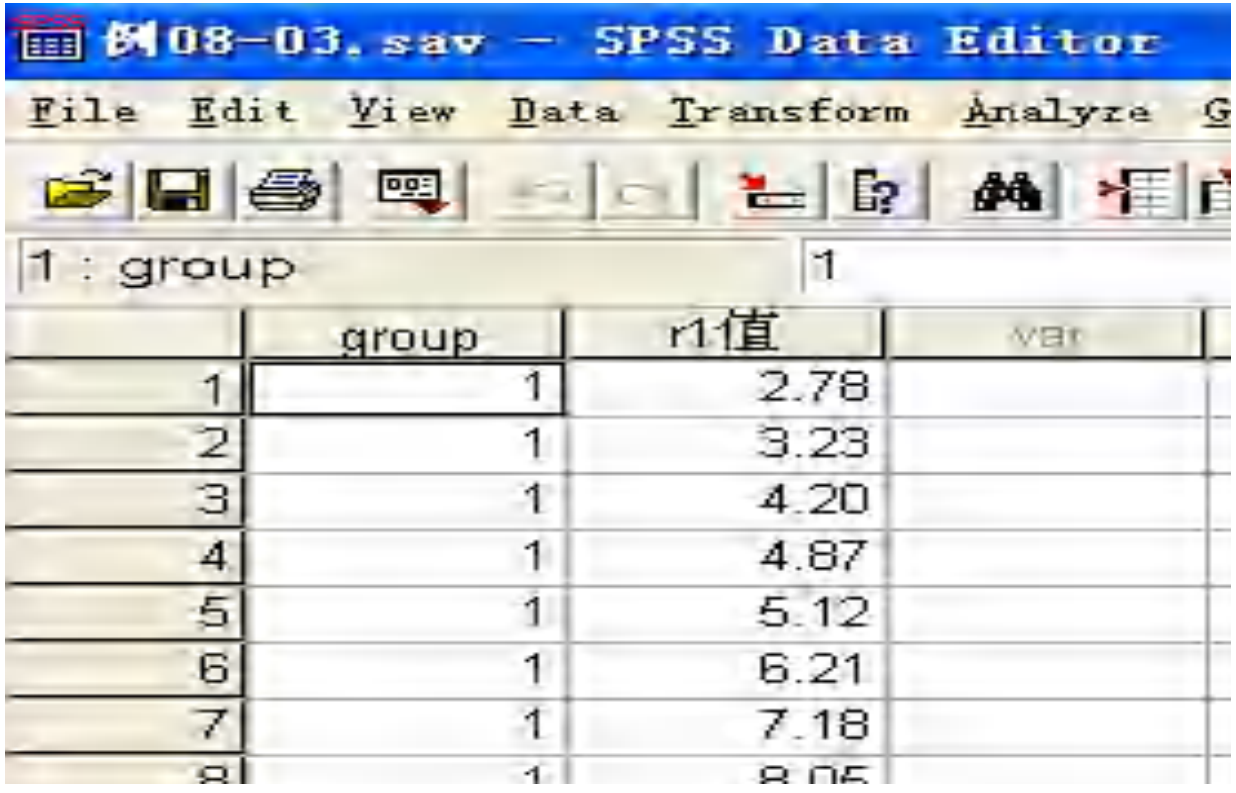
秩和检验

一、（完全随机设计）两独立样本非参数检验

1、

rd值：肺门横径右侧距（cm）；

group：组别，1=肺癌病人，2=矽肺0期工人



	group	r1值	var
1	1	2.78	
2	1	3.23	
3	1	4.20	
4	1	4.87	
5	1	5.12	
6	1	6.21	
7	1	7.18	
8	1	8.05	

河南省肿瘤医院
郑州大学附属肿瘤医院

Tests of Normality

	group	Kolmogorov-Smirnov[a]			Shapiro-Wilk		
		Statistic	df	Sig.	Statistic	df	Sig.
r1值	肺癌病人	.145	10	.200*	.959	10	.775
	矽肺工人	.122	12	.020*	.943	12	.033

*. This is a lower bound of the true significance.

a. Lilliefors Significance Correction

河南省肿瘤医院
郑州大学附属肿瘤医院

Independent Samples Test

		Levene's Test for Equality of Variances		t-test for Equality of Means						
									95% Confidence Interval of the Difference	
		F	Sig.	t	df	Sig. (2-tailed)	Mean Difference	Std. Error Difference	Lower	Upper
r1Öµ	Equal variances assumed	20.455	.000	2.352	20	.029	1.62417	.69042	.18399	3.06435
	Equal variances not assumed			2.160	9.893	.056	1.62417	.75189	-.05361	3.30194

河南省肿瘤医院
郑州大学附属肿瘤医院

例08-03.sav - SPSS Data Editor

File Edit View Data Transform Analyze Graphs Utilities Window Help

1 : group 1

	group	r1值
1	1	2.78
2	1	3.23
3	1	4.20
4	1	4.87
5	1	5.12
6	1	6.21
7	1	7.18
8	1	8.05
9	1	8.56
10	1	9.60
11	2	3.23
12	2	3.50
13	2	4.04

Analyze:
- Reports
- Descriptive Statistics
- Tables
- Compare Means
- General Linear Model
- Mixed Models
- Correlate
- Regression
- Loglinear
- Classify
- Data Reduction
- Scale
- Nonparametric Tests
- Time Series
- Survival
- Multiple Response
- Missing Value Analysis...
- Complex Samples

Nonparametric Tests:
- Chi-Square...
- Binomial...
- Runs...
- 1-Sample K-S...
- 2 Independent Samples...
- K Independent Samples...

河南省肿瘤医院 郑州大学附属肿瘤医院

河南省肿瘤医院
郑州大学附属肿瘤医院

Test Statistics[b]

	r1Öµ
Mann-Whitney U	33.500
Wilcoxon W	111.500
Z	-1.748
Asymp. Sig. (2-tailed)	.035
Exact Sig. [2*(1-tailed Sig.)]	.080[a]

a. Not corrected for ties.

b. Grouping Variable: group

正态近似法计算的*u*值及*P*值

河南省肿瘤医院 郑州大学附属肿瘤医院

二、（完全随机设计）多独立样本非参数检验

1. 原始数据资料

药物:分组变量，1=甲药，

2=乙药，3=丙药；

死亡率（%）

*Untitled1 [DataSet0] - IBM SPSS Statistics Data Ed

File Edit View Data Transform Analyze Dire

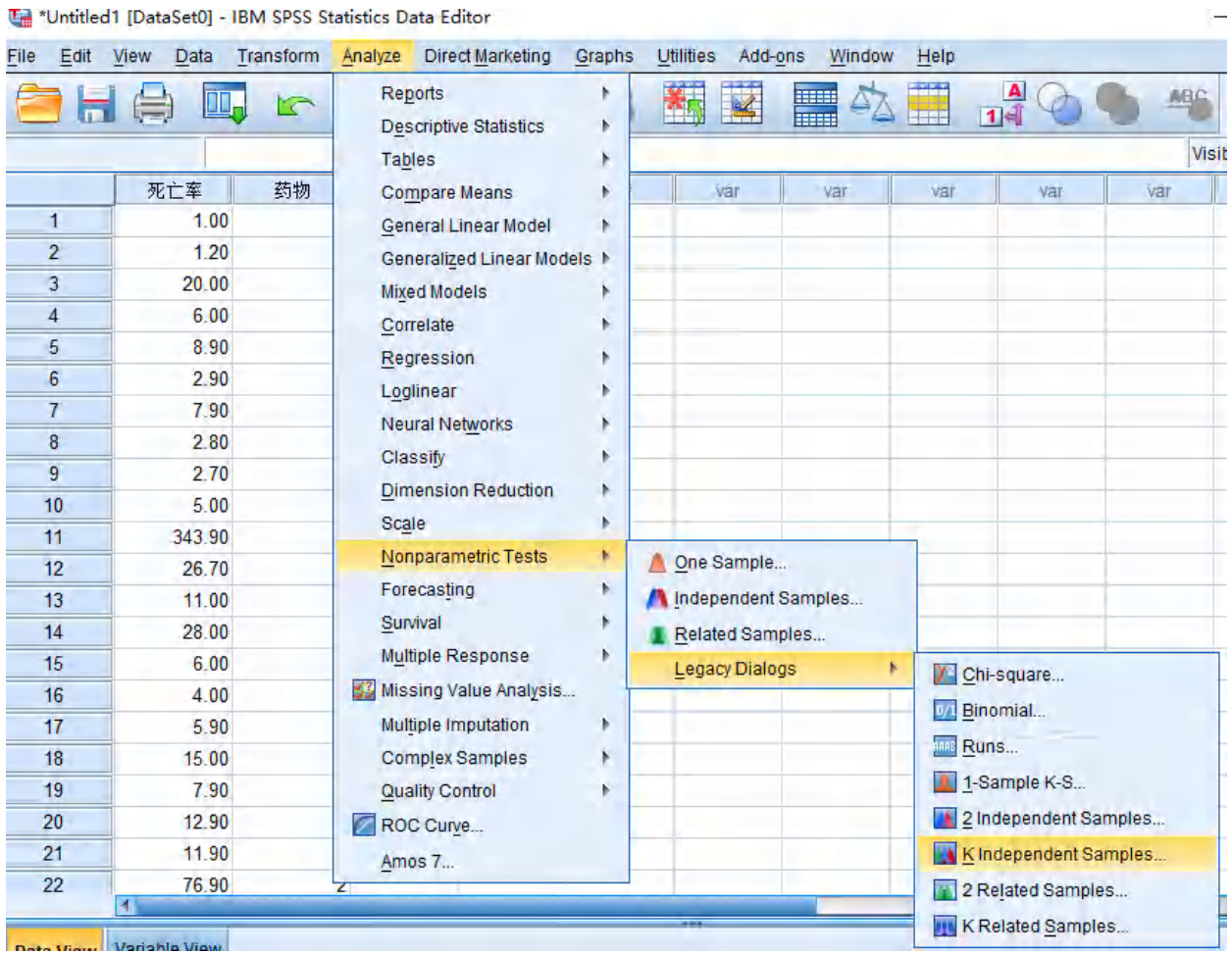
	死亡率	药物	var
1	1.00	1	
2	1.20	1	
3	20.00	1	
4	6.00	1	
5	8.90	1	
6	2.90	1	
7	7.90	1	
8	2.80	1	
9	2.70	1	
10	5.00	1	
11	343.90	1	
12	26.70	1	
13	11.00	1	
14	28.00	1	
15	6.00	2	
16	4.00	2	
17	5.90	2	
18	15.00	2	
19	7.90	2	

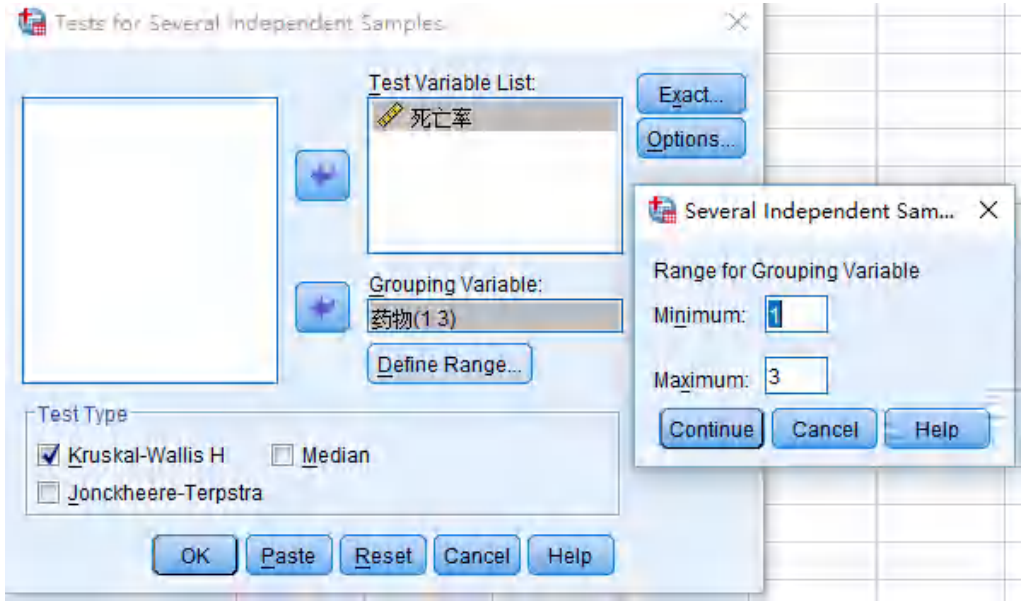
Tests of Normality

	药物	Kolmogorov-Smirnov[a]			Shapiro-Wilk		
		Statistic	df	Sig.	Statistic	df	Sig.
死亡率	甲药	.453	14	.000	.379	14	.000
	乙药	.344	14	.000	.663	14	.000
	丙药	.225	26	.002	.751	26	.000

a. Lilliefors Significance Correction

河南省肿瘤医院
郑州大学附属肿瘤医院


*Untitled1 [DataSet0] - IBM SPSS Statistics Data Editor
File Edit View Data Transform Analyze Direct Marketing Graphs Utilities Add-ons Window Help
Reports
Descriptive Statistics
Tables
Compare Means
General Linear Model
Generalized Linear Models
Mixed Models
Correlate
Regression
Loglinear
Neural Networks
Classify
Dimension Reduction
Scale
Nonparametric Tests
Forecasting
Survival
Multiple Response
Missing Value Analysis...
Multiple Imputation
Complex Samples
Quality Control
ROC Curve...
Amos 7...
One Sample...
Independent Samples...
Related Samples...
Legacy Dialogs
Chi-square...
Binomial...
Runs...
1-Sample K-S...
2 Independent Samples...
K Independent Samples...
2 Related Samples...
K Related Samples...
死亡率
药物
1.00
1.20
20.00
6.00
8.90
2.90
7.90
2.80
2.70
5.00
343.90
26.70
11.00
28.00
6.00
4.00
5.90
15.00
7.90
12.90
11.90
76.90


Tests for Several Independent Samples
Test Variable List:
死亡率
Exact...
Options...
Grouping Variable:
药物(1 3)
Define Range...
Several Independent Sam...
Range for Grouping Variable
Minimum: 1
Maximum: 3
Continue
Cancel
Help
Test Type
Kruskal-Wallis H
Median
Jonckheere-Terpstra
OK
Paste
Reset
Cancel
Help

Test Statistics[a,b]

	死亡率
Chi-Square	1.024
df	2
Asymp. Sig.	.599

a. Kruskal Wallis Test
b. Grouping Variable: 药物

如果既满足正态性条件又满足方差齐性条件，进行单因素方差分析完全可以，如果用秩和检验其检验效能将降低（即*P*值较大）。

三、两相关样本非参数检验

1. 配对符号秩和检验

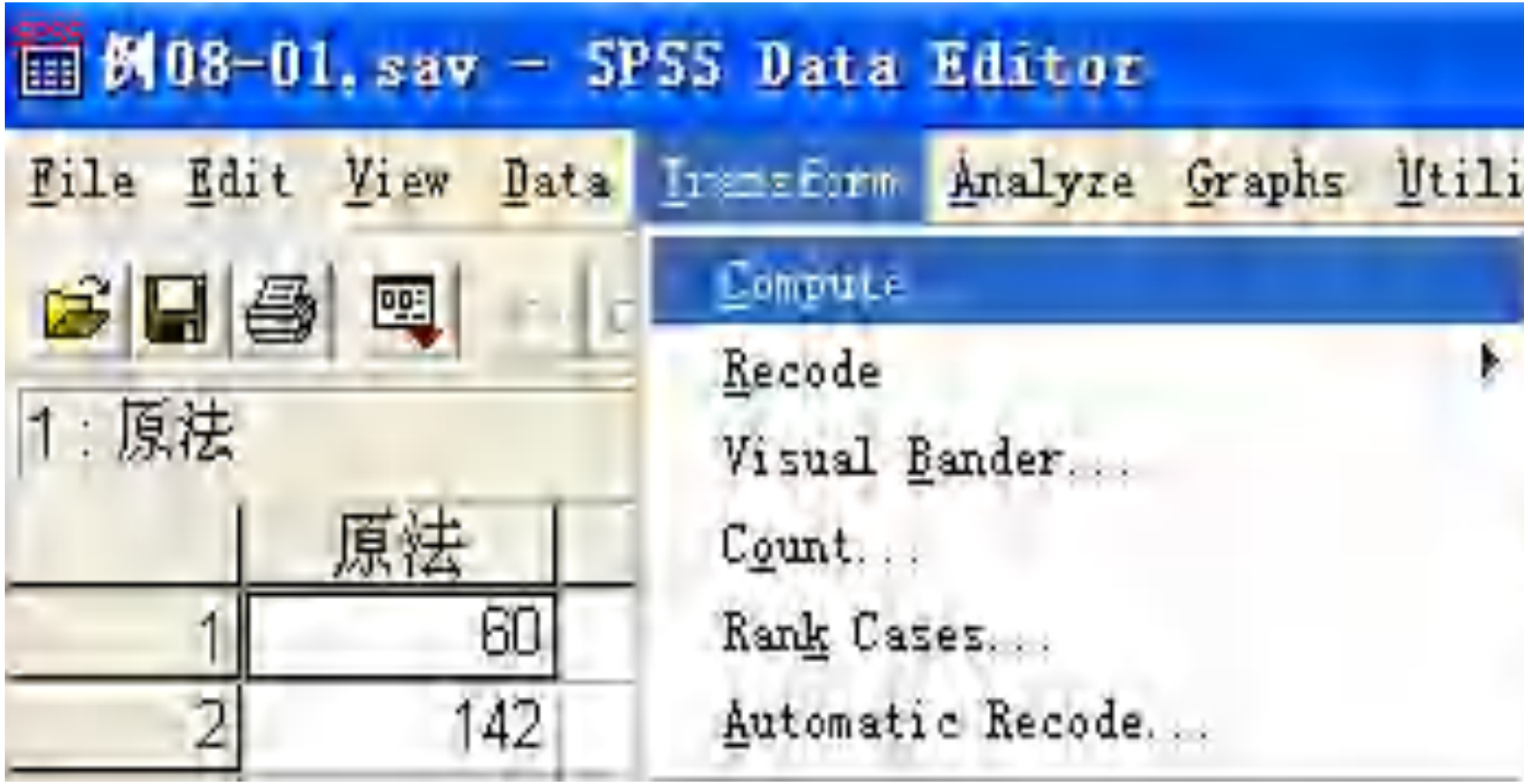
建立数据库：

原法

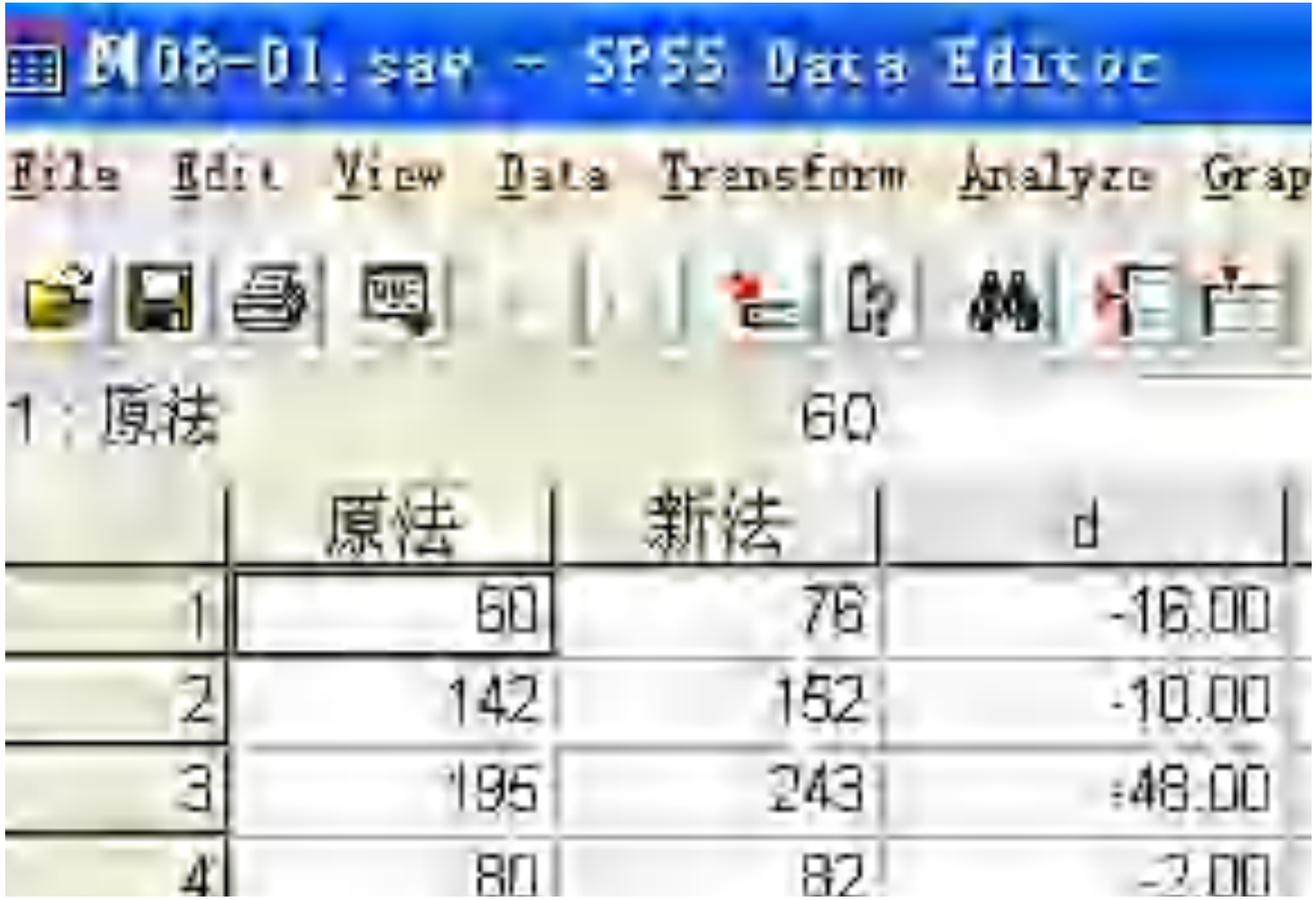
新法

	原法	新法
1	60	76
2	142	152
3	195	243
4	80	82
5	242	240
6	220	220

河南省肿瘤医院
郑州大学附属肿瘤医院


例08-01.sav - SPSS Data Editor
File Edit View Data Transform Analyze Graphs Utili
Compute
Recode
Visual Bander...
Count...
Rank Cases...
Automatic Recode...
1 : 原法
原法
1 60
2 142

Compute Variable
Target Variable:
d
Type & Label...
原法
新法
=
Numeric Expression:
原法-新法
+
<
>
7
8
9



	原法	新法	d
1	60	76	-16.00
2	142	152	-10.00
3	195	243	-48.00
4	80	82	-2.00

Tests of Normality

	Kolmogorov-Smirnov[a]			Shapiro-Wilk		
	Statistic	df	Sig.	Statistic	df	Sig.
d	.233	12	. 036	.875	12	. 036

a. Lilliefors Significance Correction

河南省肿瘤医院
郑州大学附属肿瘤医院

例08-01.sav - SPSS Data Editor

File Edit View Data Transform Analyze Graphs Utilities Window Help

35 :

	原法	新法
1	60	76
2	142	152
3	195	243
4	80	82
5	242	240
6	220	220
7	190	205
8	25	38
9	198	243
10	38	44
11	236	190
12	95	100

Reports
Descriptive Statistics
Tables
Compare Means
General Linear Model
Mixed Models
Correlate
Regression
Loglinear
Classify
Data Reduction
Scale
Nonparametric Tests
Time Series
Survival
Multiple Response
Missing Value Analysis...
Complex Samples

Chi-Square...
Binomial...
Runs...
1-Sample K-S...
2 Independent Samples...
K Independent Samples...
2 Related Samples

河南省肿瘤医院
郑州大学附属肿瘤医院

Two-Related-Samples Tests
原法
新法
d
Test Pair(s) List:
原法 - 新法
OK
Paste
Reset
Cancel
Help
检验的对子
检验方法类型
Current Selections
Variable 1:
Variable 2:
Test Type
Wilcoxon
Sign
McNemar
Marginal Homogeneity
Exact...
Options...

Ranks

		N	Mean Rank	Sum of Ranks
ÐÂ·¨ - Ô-·¨	Negative Ranks	2[a]	5.75	11.50
	Positive Ranks	9[b]	6.06	54.50
	Ties	1[c]		
	Total	12		

a. ÐÂ·¨ < Ô-·¨

b. ÐÂ·¨ > Ô-·¨

c. ÐÂ·¨ = Ô-·¨

Test Statistics[b]

	ÐÂ·¨ - Ô-·¨
Z	-1.913[a]
Asymp. Sig. (2-tailed)	.056

a. Based on negative ranks.

b. Wilcoxon Signed Ranks Test

河南省肿瘤医院
郑州大学附属肿瘤医院

生存分析

生存分析的理论复习

1. 何为生存分析？

生存分析（survival analysis）是将事件的结果（终点事件）和出现结果经历的时间结合起来分析的一种统计分析方法。

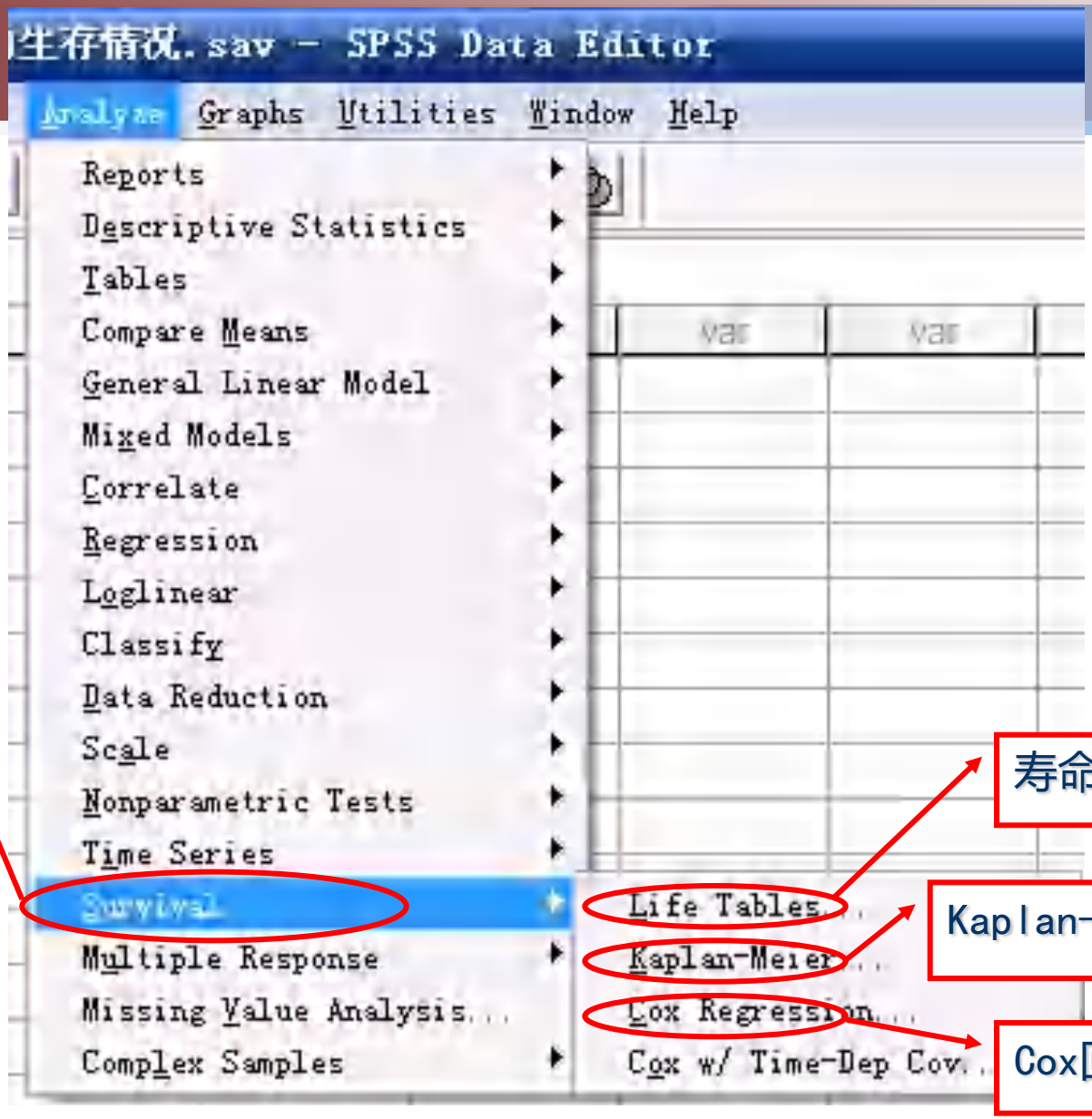
2. 生存分析的目的：

(1) **描述生存过程**：估计不同时间的总体生存率，计算中位生存期，绘制生存函数曲线。统计方法包括Kaplan-Meier（K-M）法、寿命表法。

(2) **比较**：比较不同处理组的生存率，如比较不同疗法治疗脑瘤的生存率，以了解哪种治疗方案较优。统计方法log-rank检验等。

(3) **影响因素分析**：研究某个或某些因素对生存率或生存时间的影响作用。如为改善脑瘤病人的预后，应了解影响病人预后的主要因素，包括病人的年龄、性别、病程、肿瘤分期、治疗方案等。统计方法cox比例风险回归模型等。

(4) **预测**：建立cox回归预测模型。



生存分析

寿命表法

Kaplan-Meier法

Cox回归

河南省肿瘤医院
郑州大学附属肿瘤医院

寿命表法

生存率估计——寿命表法

❖**例1** 某研究者收集了男性心绞痛患者2418例，经随访后将有关资料整理后列于下表，其中生存时间是以年计算的，试计算其生存率与标准误.

生存时间区间（年）	死亡人数	删失人数	生存时间区间（年）	死亡人数	删失人数
0~	456	0	8~	51	68
1~	226	39	9~	42	64
2~	152	22	10~	43	45
3~	171	23	11~	34	53
4~	135	24	12~	18	33
5~	125	107	13~	9	27
6~	83	133	14~	6	33
7~	74	102	15~	0	20

河南省肿瘤医院
郑州大学附属肿瘤医院

数据格式

生存时间、例数、结局（1：死亡、0：删失）

File Edit View Data Transform Analyze Grap

35 :

	生存时间	人数	结局
7	6.00	83.00	1.00
8	7.00	74.00	1.00
9	8.00	51.00	1.00
10	9.00	42.00	1.00
11	10.00	43.00	1.00
12	11.00	34.00	1.00
13	12.00	18.00	1.00
14	13.00	9.00	1.00
15	14.00	6.00	1.00
16	15.00	.00	1.00
17	.00	.00	.00
18	1.00	39.00	.00

河南省肿瘤医院
郑州大学附属肿瘤医院

操作步骤

Data

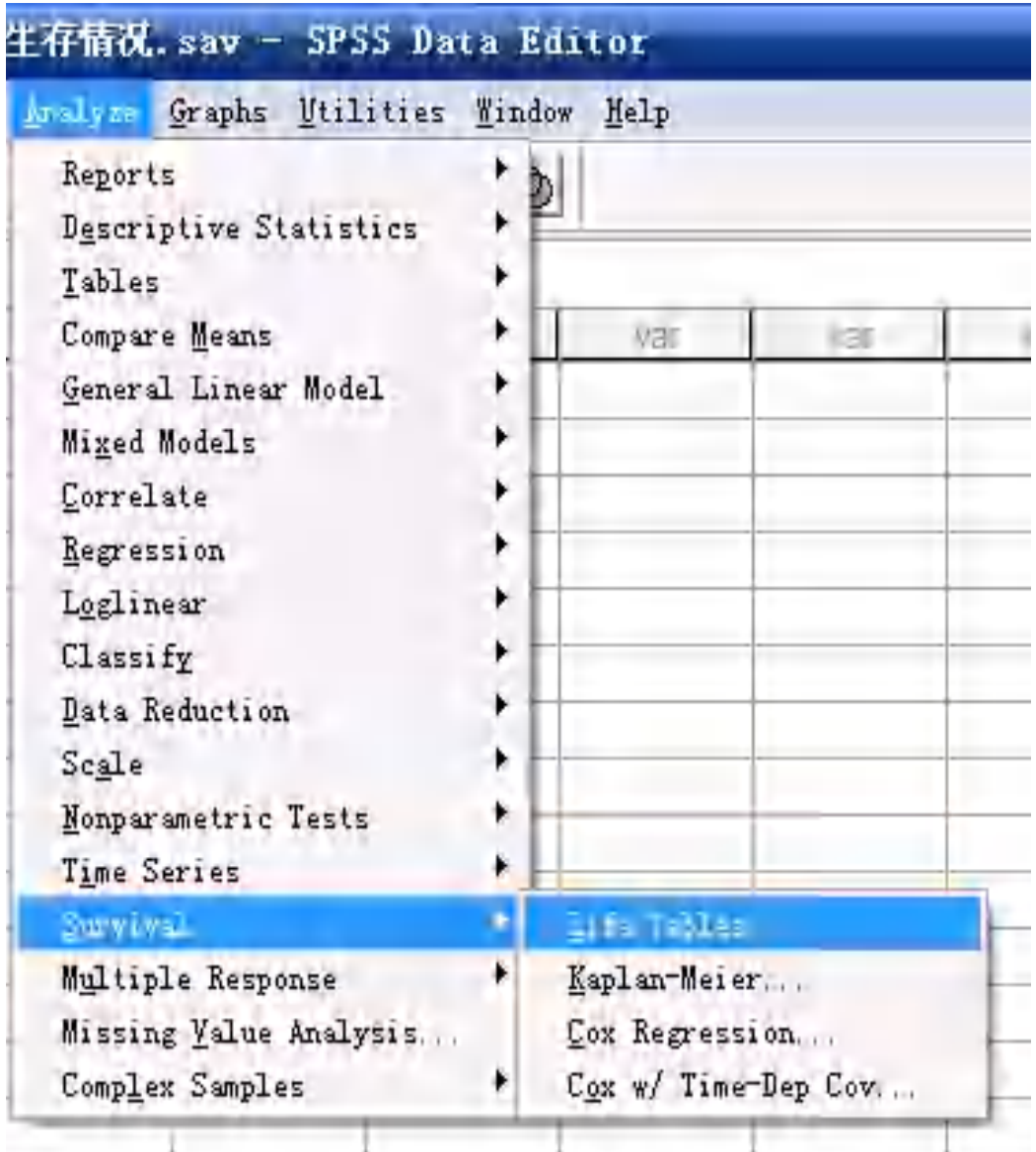
→Weight cases

操作步骤

Analyze

→Survival

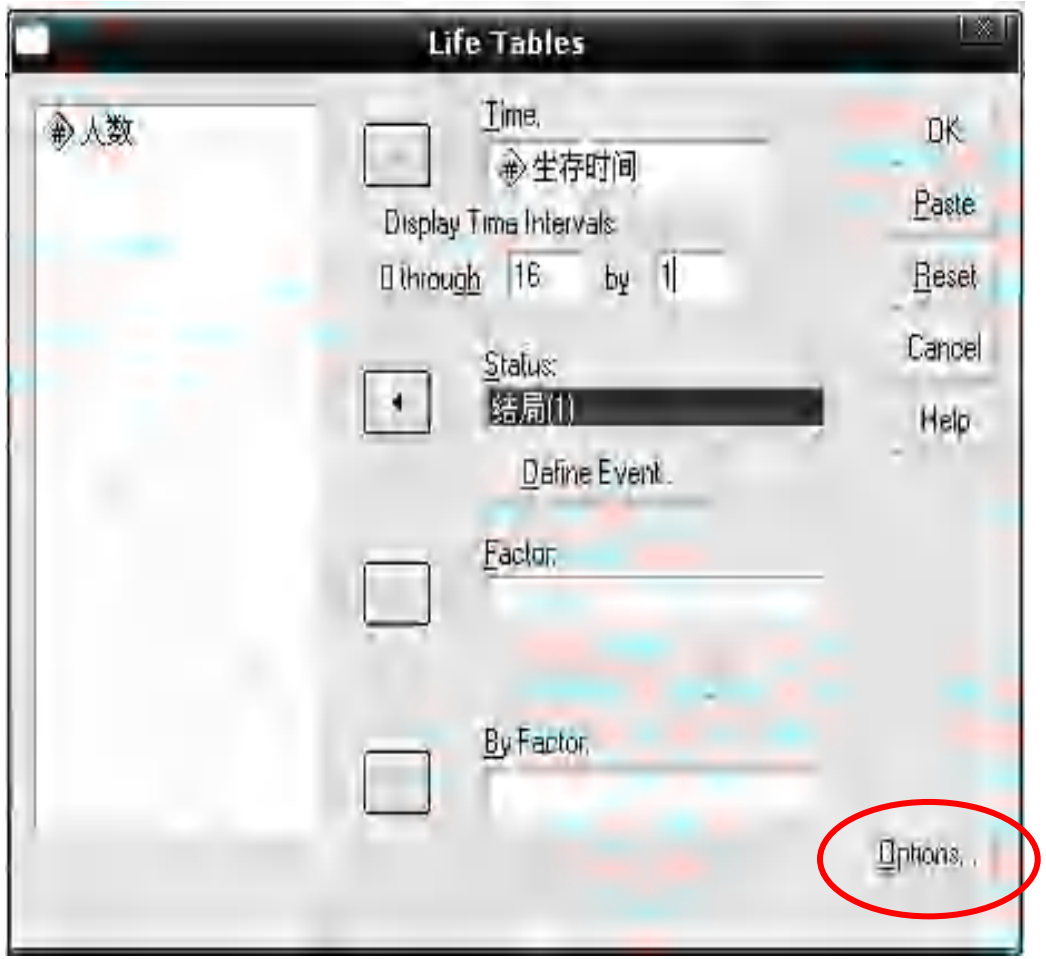
→ Life Tables



河南省肿瘤医院
郑州大学附属肿瘤医院

Life Tables对话框

Life Tables对话框

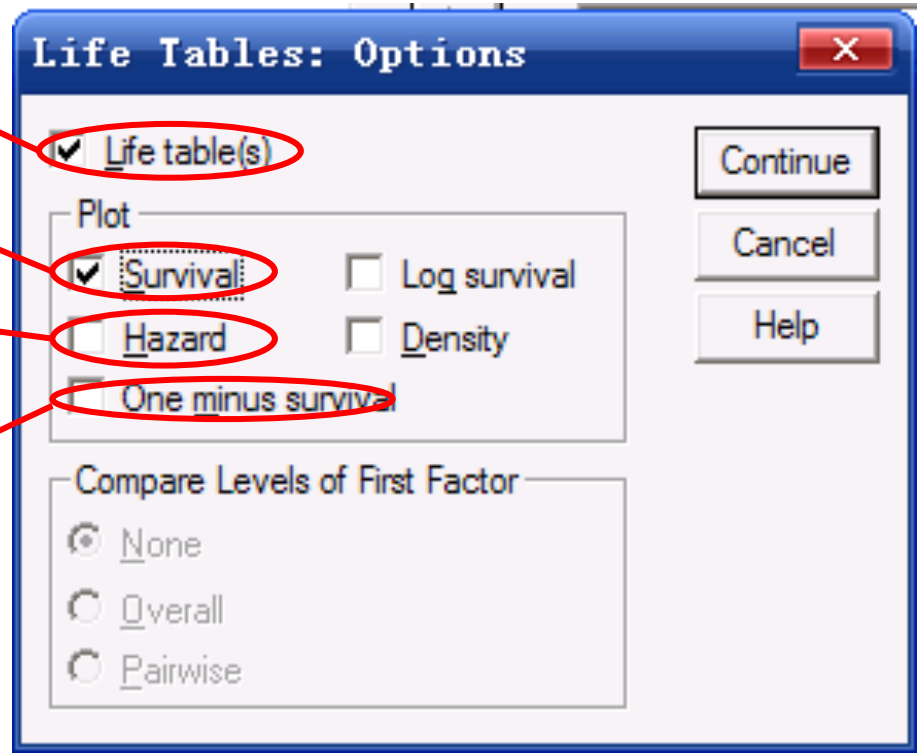


寿命表

生存曲线

风险函数曲线

生存曲线被1减后的曲线



寿命表

初期观察例数

有效观察例数

条件死亡率

生存时间组段下限

期内删失例数

期内死亡例数

Interval Start Time	Number Entering Interval	Number Withdrawing during Interval	Number Exposed to Risk	Number of Terminal Events	Proportion Terminating
0	2418	0	2418.000	456	.19
1	1962	39	1942.500	226	.12
2	1697	22	1686.000	152	.09
3	1523	23	1511.500	171	.11
4	1329	24	1317.000	135	.10
5	1170	107	1116.500	125	.11
6	938	133	871.500	83	.10
7	722	102	671.000	74	.11
8	546	68	512.000	51	.10
9	427	64	395.000	42	.11
10	321	45	298.500	43	.14
11	233	53	206.500	34	.16
12	146	33	129.500	18	.14
13	95	27	81.500	9	.11
14	59	33	42.500	6	.14
15	20	20	10.000	0	.00

a. The median survival time is 5.3313

寿命表

累计生存率

概率密度函数

风险率

条件生存率

累计生存率的标准误

概率密度函数的标准误

风险率标准误

Life Table[a]

Proportion Surviving	Cumulative Proportion Surviving at End of Interval	Std. Error of Cumulative Proportion Surviving at End of Interval	Probability Density	Std. Error of Probability Density	Hazard Rate	Std. Error of Hazard Rate
.81	.81	.01	.189	.008	.21	.01
.88	.72	.01	.094	.006	.12	.01
.91	.65	.01	.065	.005	.09	.01
.89	.58	.01	.074	.005	.12	.01
.90	.52	.01	.059	.005	.11	.01
.89	.46	.01	.058	.005	.12	.01
.90	.42	.01	.044	.005	.10	.01
.89	.37	.01	.046	.005	.12	.01
.90	.33	.01	.037	.005	.10	.01
.89	.30	.01	.036	.005	.11	.02
.86	.26	.01	.043	.006	.16	.02
.84	.21	.01	.042	.007	.18	.03
.86	.18	.01	.030	.007	.15	.04
.89	.16	.01	.020	.007	.12	.04
.86	.14	.01	.023	.009	.15	.06
1.00	.14	.01	.000	.000	.00	.00

Kaplan-Meier法

生存率比较——Kaplan-Meier法

❖ 例2假定我们用乙种手术方式治疗了与例17-1病情一致的同种疾病20例。试比较两种手术方式的生存时间。

甲种手术

1　3　5（3）　6（3）　7　8　10（2）　14^+　17　19^+　20^+　22^+　26^+　31^+

34　34^+　44　59

乙种手术

1（2）　2　3（2）　4（3）　6（2）　8　9（2）　10　11　12　13　14　15

17　18

河南省肿瘤医院
郑州大学附属肿瘤医院

数据格式

生存时间、结局（1：完全数据、0：删失数据）、分组（1：甲组、2：乙组）

两组生存曲线的比较.sav - SPSS Data Editor

File Edit View Data Transform Analyze Graphs Uti

16 : 分组 1

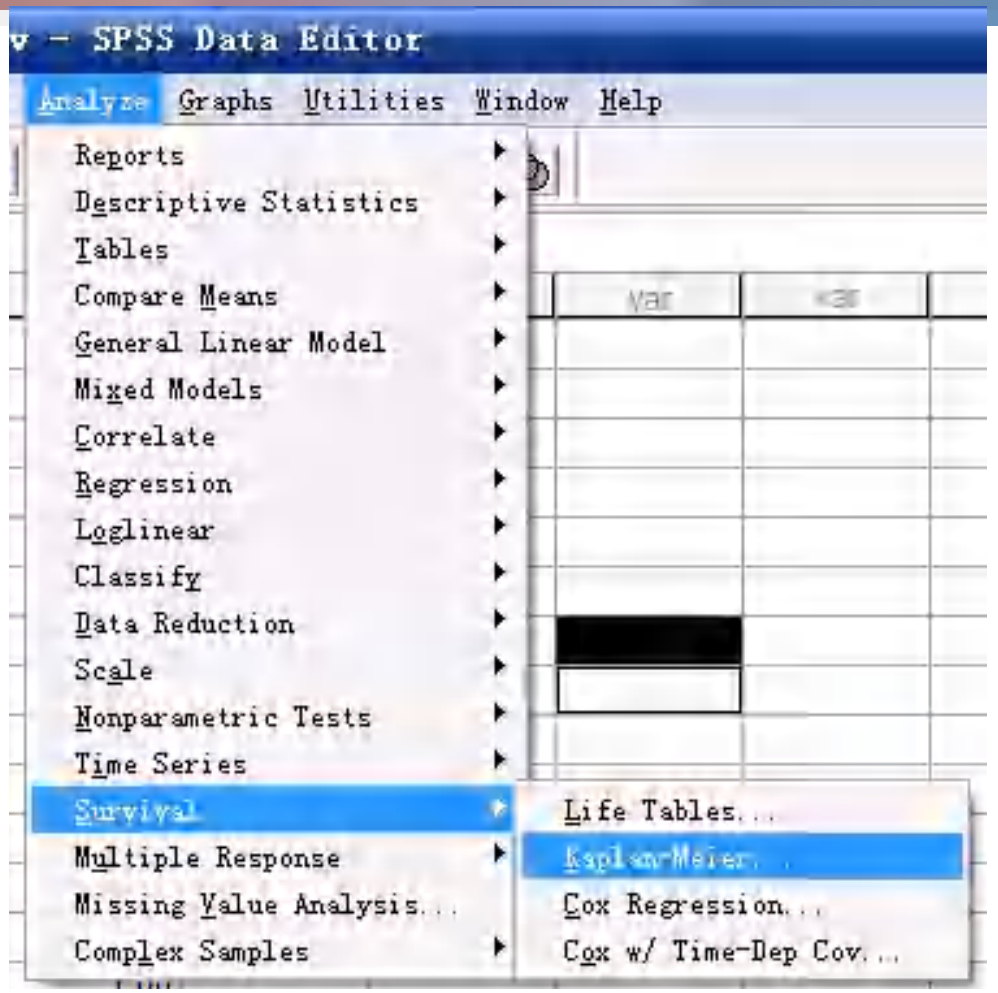
	生存时间	结局	分组	var
1	1.00	1.00	1.00	
2	3.00	1.00	1.00	
3	5.00	1.00	1.00	
4	5.00	1.00	1.00	
5	5.00	1.00	1.00	
6	6.00	1.00	1.00	
7	6.00	1.00	1.00	
8	6.00	1.00	1.00	

操作步骤

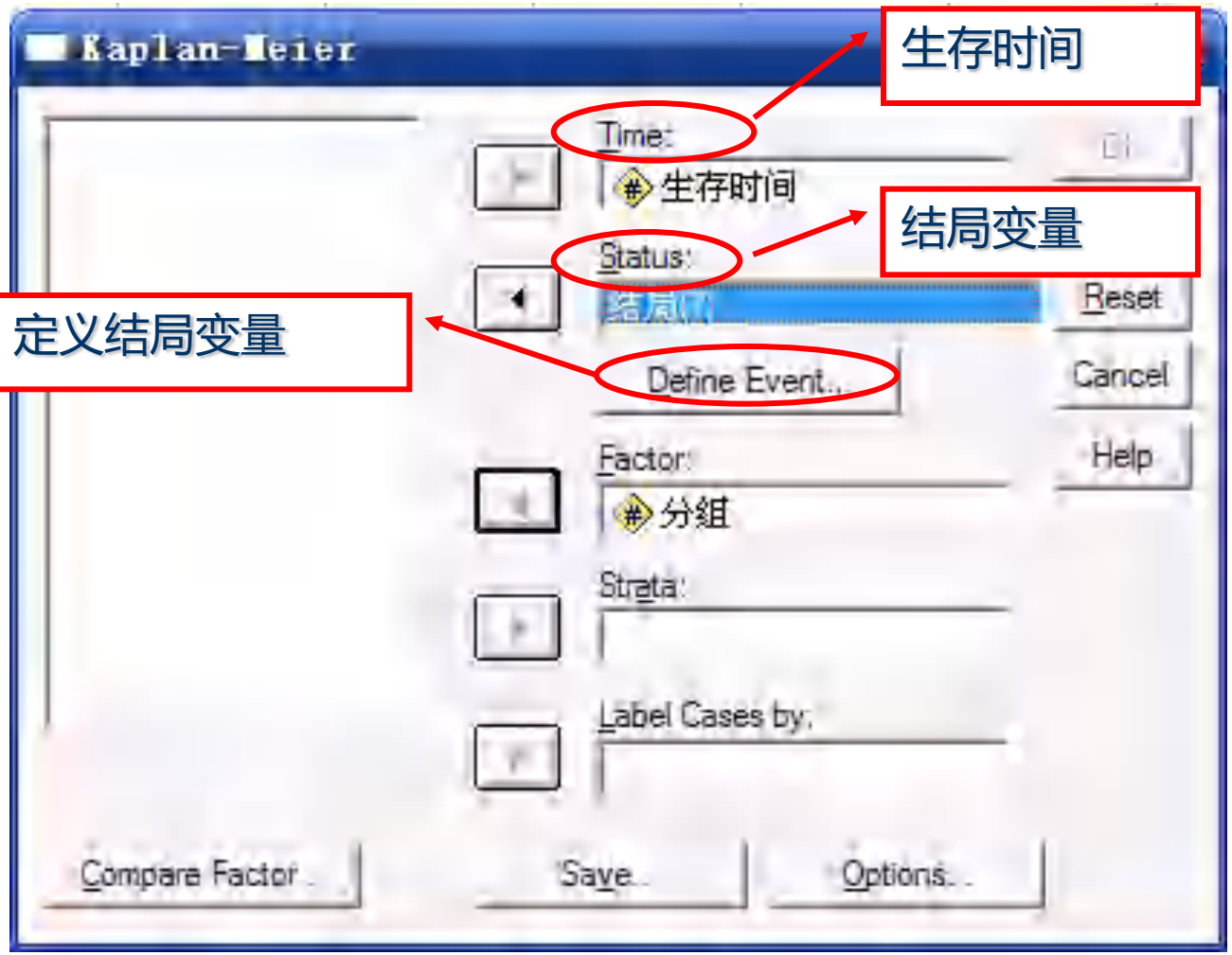
Analyze

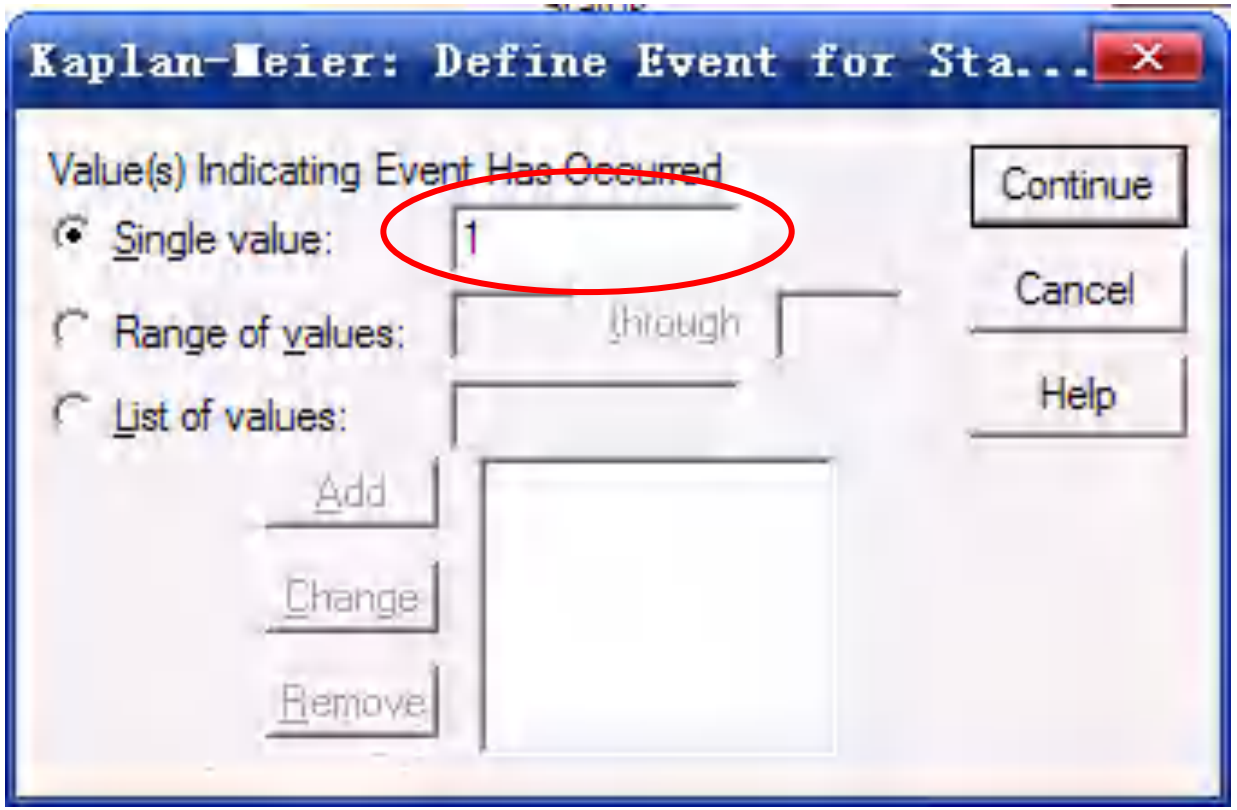
→Survival

→ Kaplan-Meier

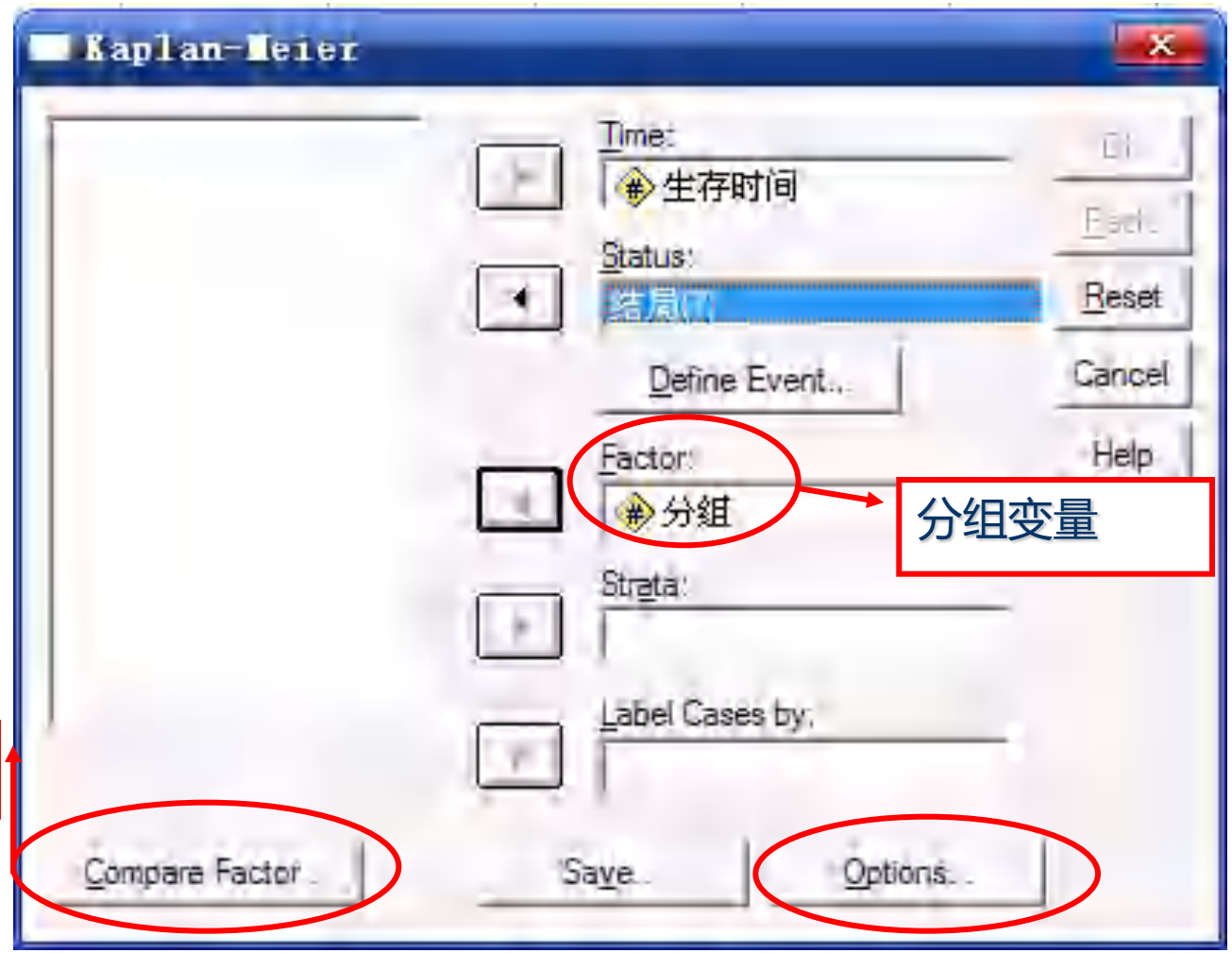


Kaplan-Meier对话框





Kaplan-Meier对话框



分组变量

组间比较

Kaplan-Meier: Compare Factor Levels对话框

生存时间

结局

累积生存率
及其标准误

累计死
亡例数

存活例数

Survival Table

¼××é		Time	Status	Cumulative Proportion Surviving at the Time		N of Cumulative Events	N of Remaining Cases
				Estimate	Std. Error		
	1	1.000	1.00	.957	.043	1	22
	2	3.000	1.00	.913	.059	2	21
	3	5.000	1.00	.	.	3	20
	4	5.000	1.00	.	.	4	19
	5	5.000	1.00	.783	.086	5	18
	6	6.000	1.00	.	.	6	17
	7	6.000	1.00	.	.	7	16
	8	6.000	1.00	.652	.099	8	15
	9	7.000	1.00	.609	.102	9	14
	10	8.000	1.00	.565	.103	10	13
	11	10.000	1.00	.	.	11	12
	12	10.000	1.00	.478	.104	12	11
	13	14.000	.00	.	.	12	10
	14	17.000	1.00	.430	.104	13	9
	15	19.000	.00	.	.	13	8
	16	20.000	.00	.	.	13	7
	17	22.000	.00	.	.	13	6
	18	26.000	.00	.	.	13	5
	19	31.000	.00	.	.	13	4
	20	34.000	1.00	.323	.122	14	3
	21	34.000	.00	.	.	14	2
	22	44.000	1.00	.161	.129	15	1
	23	59.000	1.00	.000	.000	16	0

平均生存时间与中位生存时间

平均生存时间
及其标准误和可信区间

中位生存时间
及其标准误和可信区间

or Surv

	Mean[a]				Median			
			95% Confidence Interval				95% Confidence Interval	
·Ö×é	Estimate	Std. Error	Lower Bound	Upper Bound	Estimate	Std. Error	Lower Bound	Upper Bound
¼××é	24.228	4.991	14.444	34.011	10.000	6.955	.000	23.632
ÒÒ×é	7.800	1.176	5.496	10.104	6.000	2.981	.156	11.844
Overall	16.440	2.943	10.671	22.209	9.000	1.402	6.253	11.747

a. Estimation is limited to the largest survival time if it is censored.

中位生存时间即生存率为50%时的生存时间，50%的观察单位发生阳性事件（结局）的时间。

甲乙两组的中位生存时间分别是10个月和6个月。

中位生存时间的对数秩检验

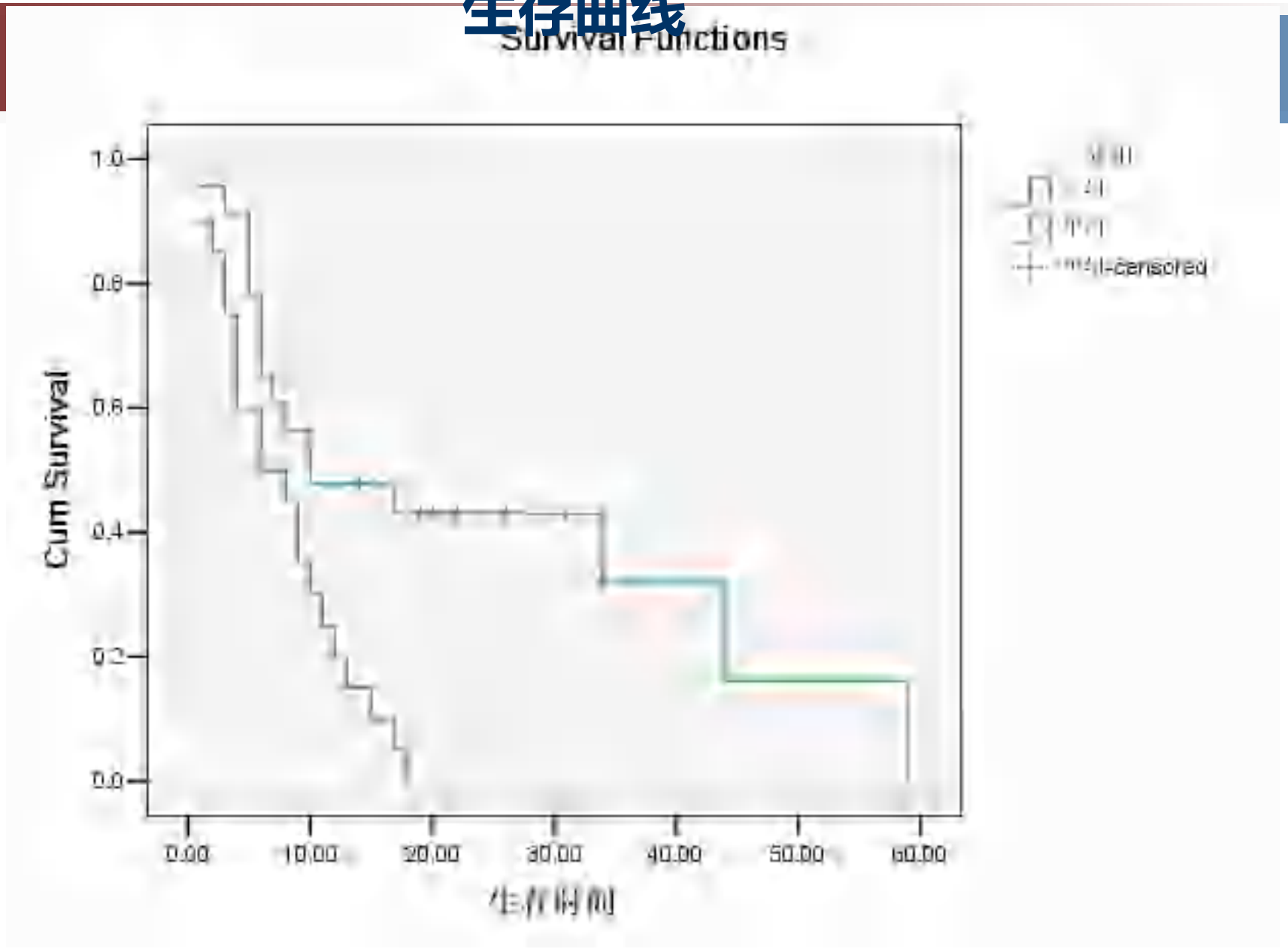
Overall Comparisons

	Chi-Square	df	Sig.
Log Rank (Mantel-Cox)	8.754	1	.003

Test of equality of survival distributions for the different levels of ·Ö×é.

甲乙两组中位生存时间的差异有统计学意义P=0.003，甲组生存时间长于乙组。

生存曲线


Survival Functions
Cum Survival
1.0
0.8
0.6
0.4
0.2
0.0
0.00
10.00
20.00
30.00
40.00
50.00
60.00
生存时间
Censored

Cox回归分析

生存状态影响因素分析——Cox回归

❖**例3：某癌症患者手术治疗后随访的生存时间资料收集了28例患者年龄（岁）、性别（0-女、1-男）白蛋白含量（g/L）、胆红素（μmol/L）、结局（0-存活、1-死亡）。试用Cox模型进行分析。**

17-3某癌症患者手术治疗后随访的生存时间资料.sav - SPSS Data Editor

File Edit View Data Transform Analyze Graphs Utilities Window Help

24 :

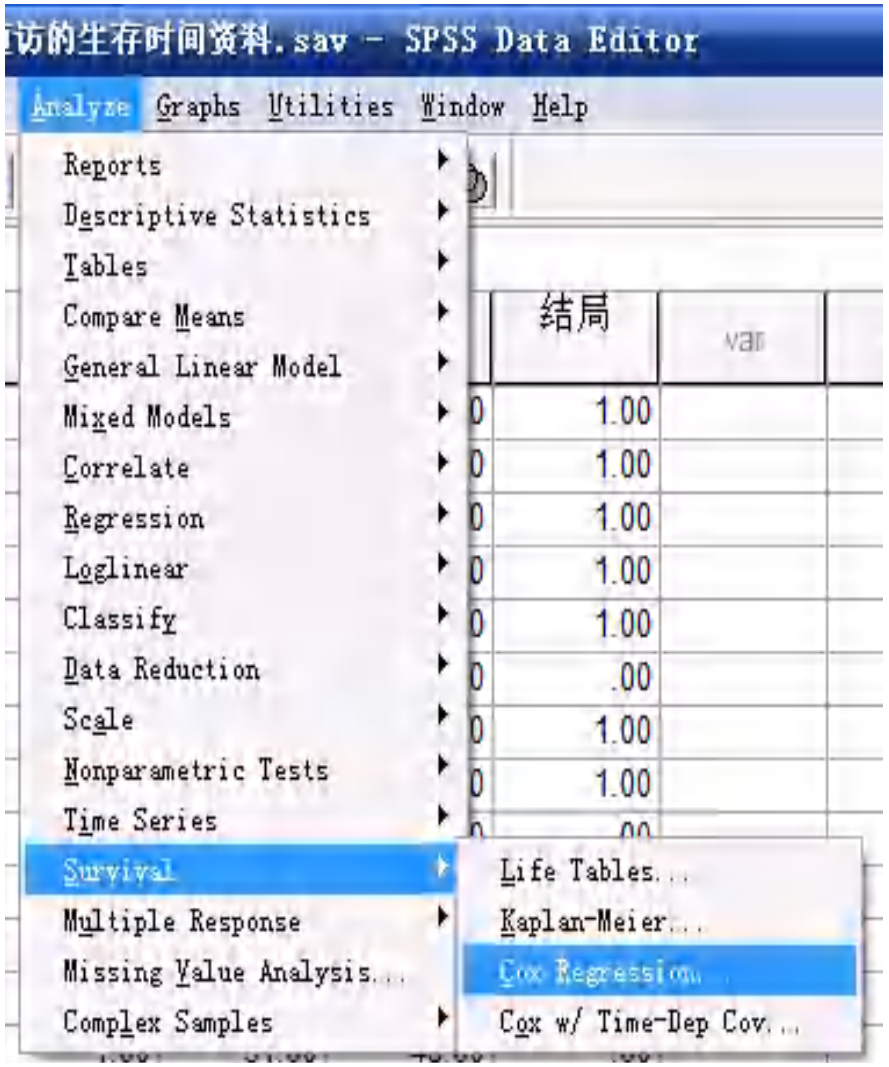
	生存月	年龄	性别	白蛋白含量	胆红素	结局	var
1	1.00	53.00	1.00	24.00	330.00	1.00	
2	1.00	61.00	1.00	23.00	165.00	1.00	
3	1.00	32.00	1.00	25.00	180.00	1.00	
4	2.00	35.00	1.00	21.00	80.00	1.00	
5	2.00	26.00	1.00	26.00	90.00	1.00	
6	2.00	25.00	1.00	28.00	140.00	.00	
7	3.00	59.00	1.00	22.00	35.00	1.00	
8	3.00	46.00	1.00	29.00	250.00	1.00	
9	3.00	42.00	1.00	29.00	78.00	.00	
10	4.00	57.00	1.00	30.00	220.00	1.00	
11	5.00	64.00	.00	32.00	93.00	1.00	
12	5.00	48.00	1.00	33.00	156.00	.00	
13	6.00	54.00	1.00	31.00	48.00	.00	

操作步骤

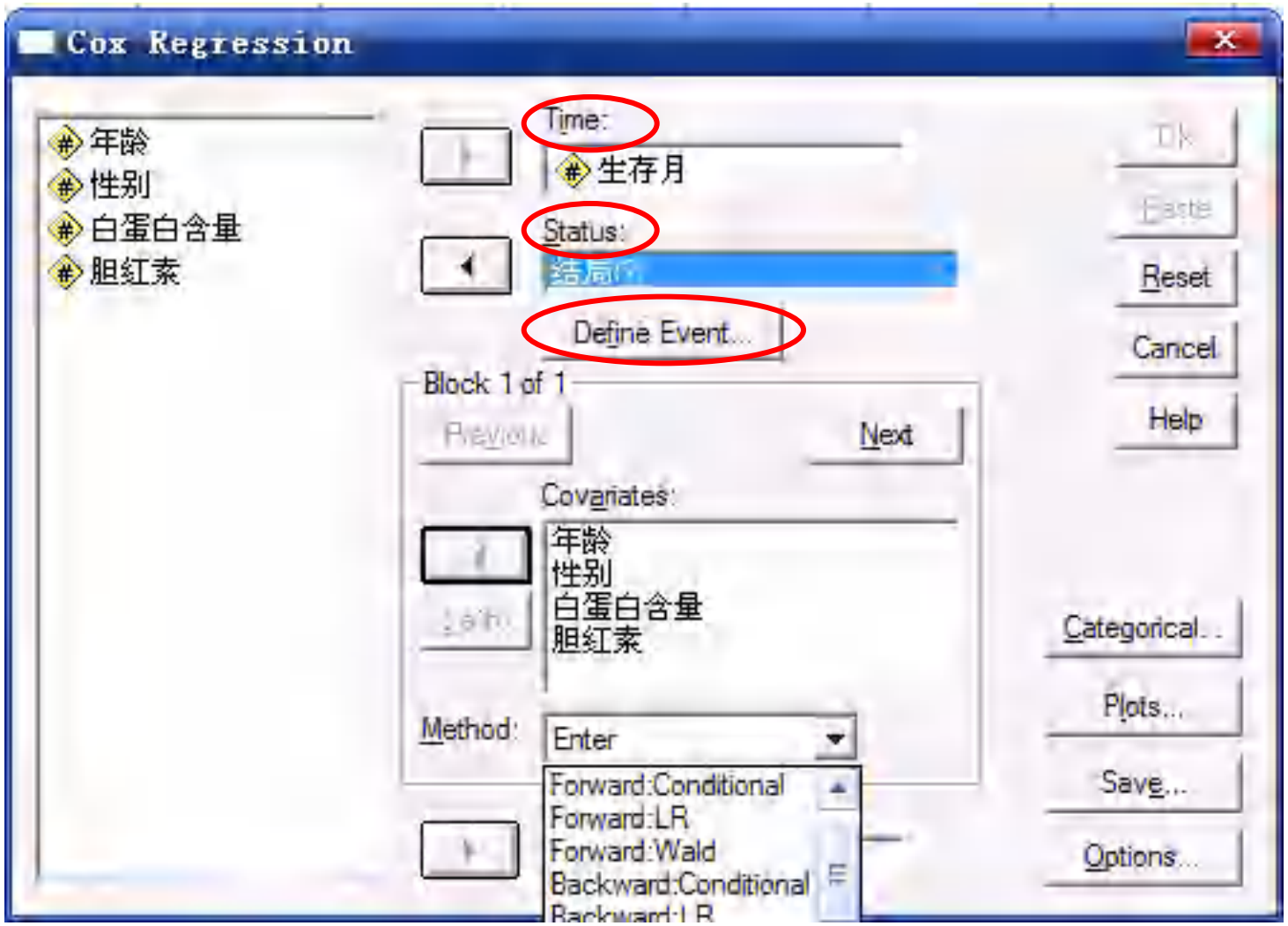
Analyze

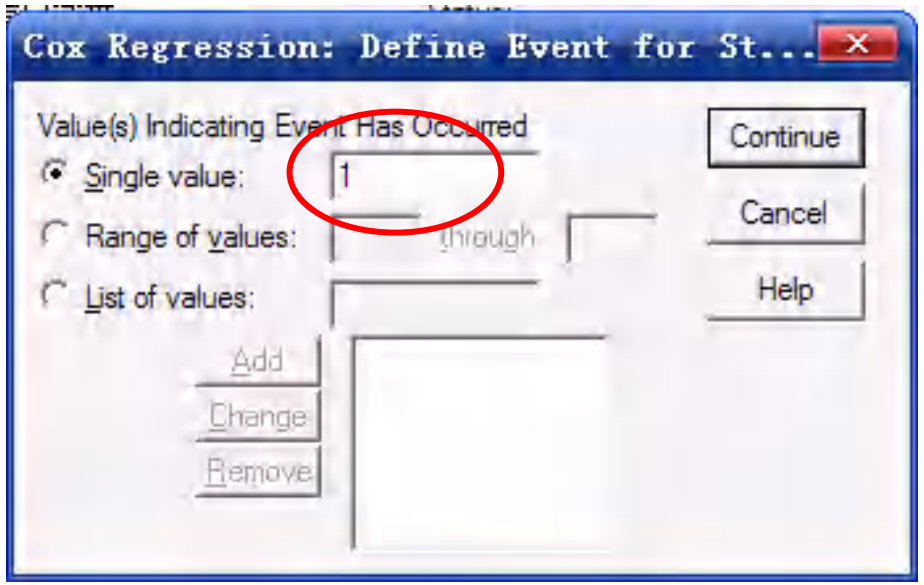
→Survival

→Cox Regression

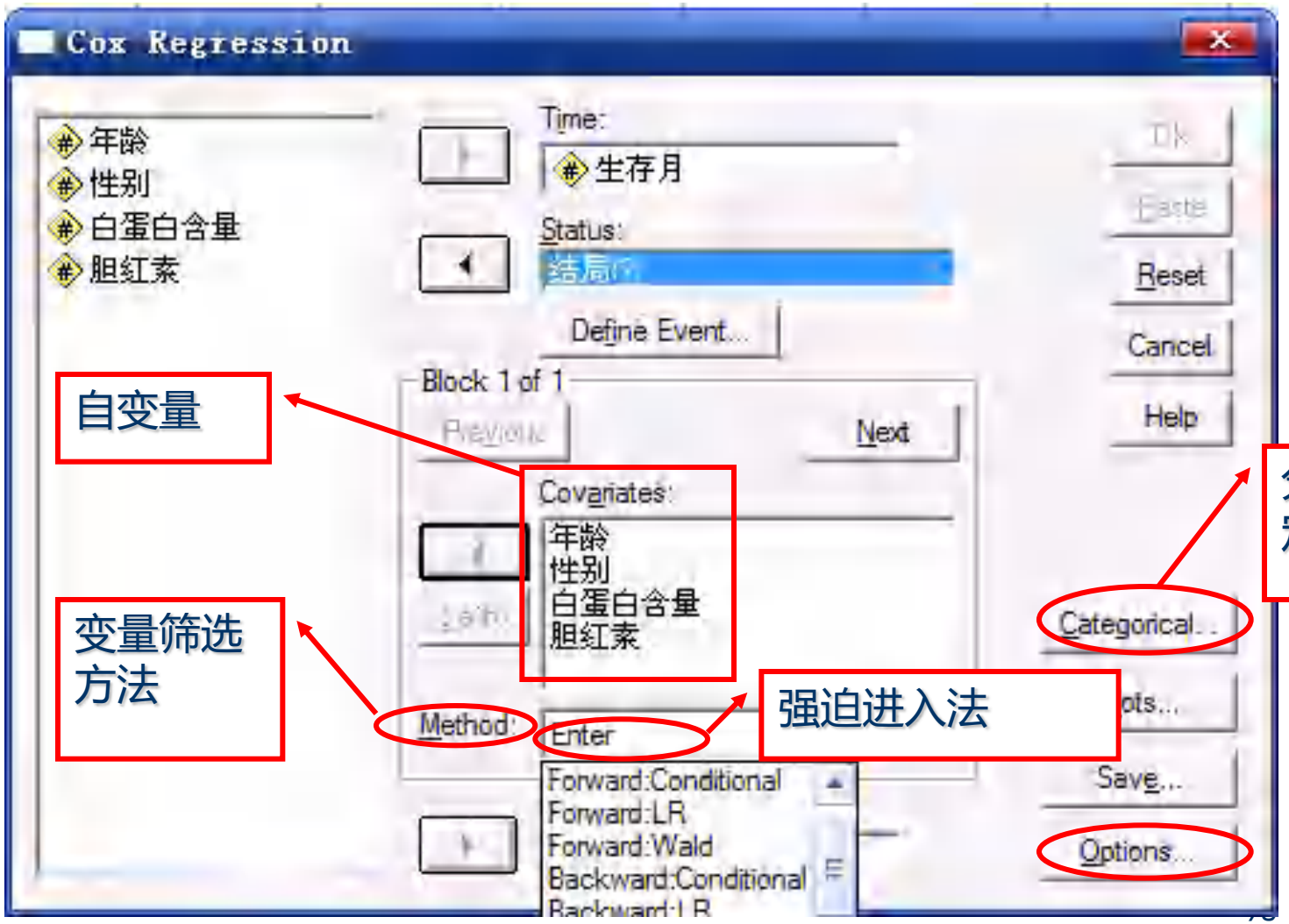


Cox Regression对话框





Cox Regression对话框

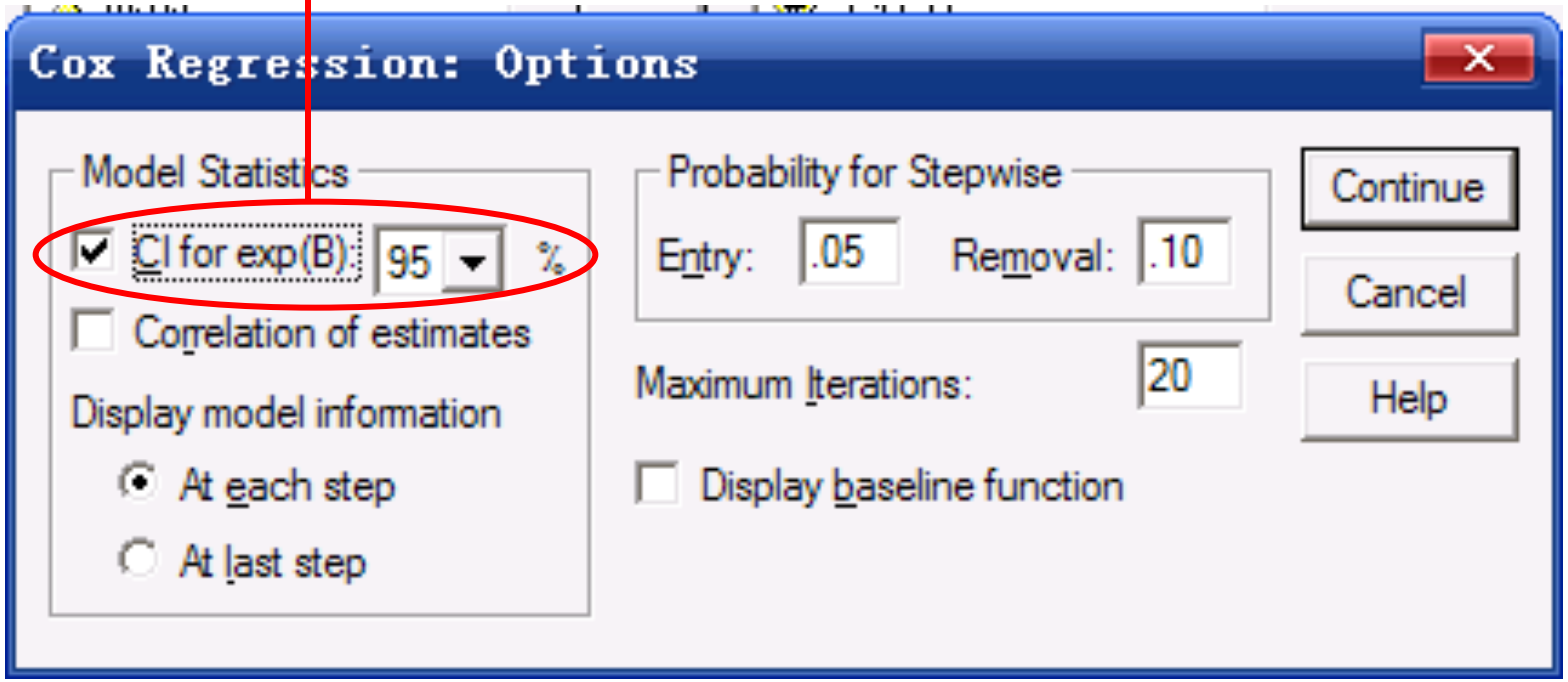


Categorical对话框

河南省肿瘤医院
郑州大学附属肿瘤医院

Cox Regression：Options对话框

相对危险度RR可信区间



Cox模型中变量参数估计（最重要的表）

偏回归系数
及其标准误

Wald统计量

P值

相对危险度
及其可信区间

Variables in the Equation

	B	SE	Wald	df	Sig.	Exp(B)	95.0% CI for Exp(B)	
							Lower	Upper
ÄêÁä	-.020	.029	.490	1	.484	.980	.926	1.037
ÐÔ±ð	1.217	.852	2.041	1	.153	3.377	.636	17.927
°×µ°°×º¬Á¿	-.548	.136	16.338	1	.000	.578	.443	.754
µ¨ºìËØ	.014	.004	9.310	1	.002	1.014	1.005	1.023

年龄与性别对生存状态的影响无统计学意义，进行逐步回归（步骤略）

逐步回归结果
向前法：似然比forward：LR

Variables in the Equation

		B	SE	Wald	df	Sig.	Exp(B)	95.0% CI for Exp(B)	
								Lower	Upper
Step 1	°×µ°°×º¬Á¿	-.414	.096	18.708	1	.000	.661	.548	.797
Step 2	°×µ°°×º¬Á¿	-.526	.126	17.498	1	.000	.591	.462	.756
	µ¨ºìËØ	.014	.004	9.439	1	.002	1.014	1.005	1.022

逐步回归分析结果最终显示，对术后生存时间影响的因素是胆红素和白蛋白含量。其中胆红素是一个危险因素，而白蛋白含量是一个保护因素。

总结

❖均值的比较

- 符合正态性->t检验（2组）/方差分析（2组以上）
- 不符合正态性->数据转换/非参数检验

[Wilcoxon秩和检验/Mann-Whitney U test（2组）；Kolmogorov-Smirnov检验（2组大样本）；McNemar/Wilcoxon符号秩和检验（2组相关）；Friedman（2组以上相关）；Kruskal-Wallis/Median/Jonckheere-Terpstra（2组以上）]

❖率的比较

- 卡方检验/Fisher’s精确概率法

总结

❖生存分析

- 寿命表法（Life Tables）：计算生存率
- Kaplan-Meier法：计算中位生存时间/比较生存曲线的差异
- Cox回归分析（Cox Regression）：多因素分析/求HR值

❖作图

- 一般可用Excel
- 复杂图形可用SPSS
- 美化图形用GraphPad
- SPSS链接：https://pan.baidu.com/s/1egUj3T9CM5bs7ZWh6NVTQQ 密码：kzoe
- GraphPad链接：https://pan.baidu.com/s/1Qk7-WCUqauNgogrmWJU4dw 密码：8a6m

统计学是对令人困惑费解的问题做出数字设想的艺术。
Thank you !