

SPSS进阶学习与研究成果撰写

2019.4

南开大学图书馆

目录

CONTENTS

- 01 化繁为简——描述性统计分析
- 02 大道至简——因子分析
- 03 影响因素判断——多因素方差分析
- 04 万物皆有联系——相关分析
- 05 撰写研究成果的若干建议

01

化繁为简

——描述性统计分析



数据质量

数据值异常



数据基本特征

描述性统计指标

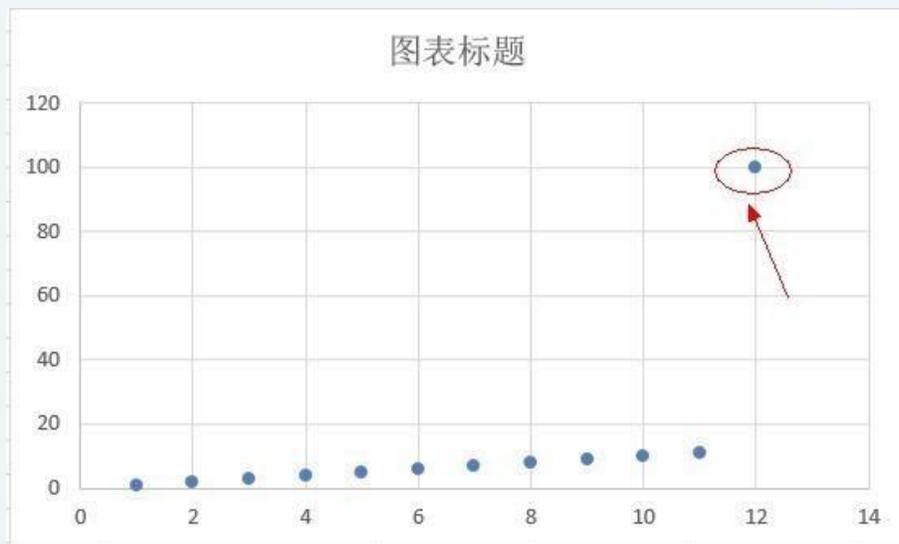


数据整体分布形态

分布系数和形态图

第一节 数据质量的反映—— 数据异常值的判别与剔除

异常值也称离群值，指样本中的个别值，其数值明显偏离样本的其余观测值



判别异常值

剔除异常值

数据分析

第二节 数据的基本特征反映—— 描述性统计指标

适于数值型数据，易受极值影响

适于定序数据，不受极值影响

感应灵敏，受极值影响大

标准误越小，表明样本对
总体越有代表性

频率: 统计量

百分位值

四分位数(Q)

割点(U): 10 相等组

百分位数(P):

添加(A)

更改(C)

删除(R)

集中趋势

均值(M)

中位数

众数(O)

合计

值为组的中点(L)

离散

标准差(S) 最小值

方差 最大值

范围 均值的标准误(E)

分布

偏度

峰度

继续 取消 帮助

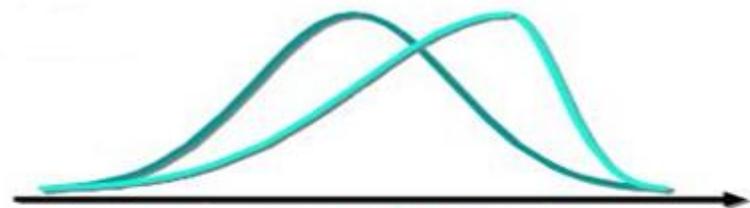
不受极值影响，但不灵敏

不受极值影响，具有非唯一性

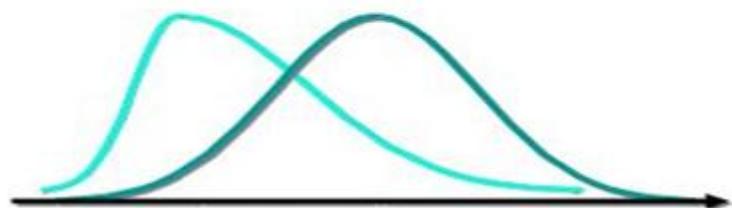
数据分布对称性

数据分布形态的扁平程度

第三节 数据的分布形态反映——分布系数



左偏分布



右偏分布

偏度=0 正态分布
偏度<0 左偏分布
偏度>0 右偏分布

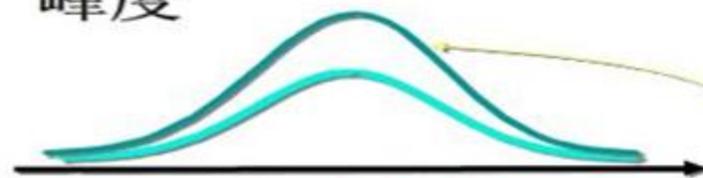
偏度

分布形态

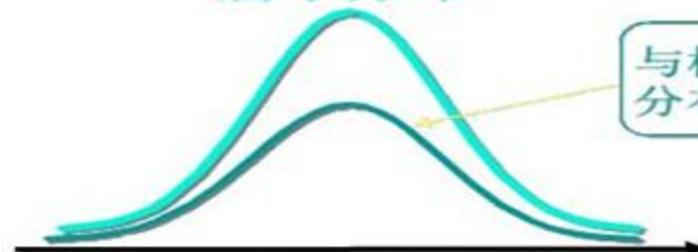
峰度

峰度=0 正态分布
峰度>0 扁平分布
峰度<0 尖峰分布

峰度



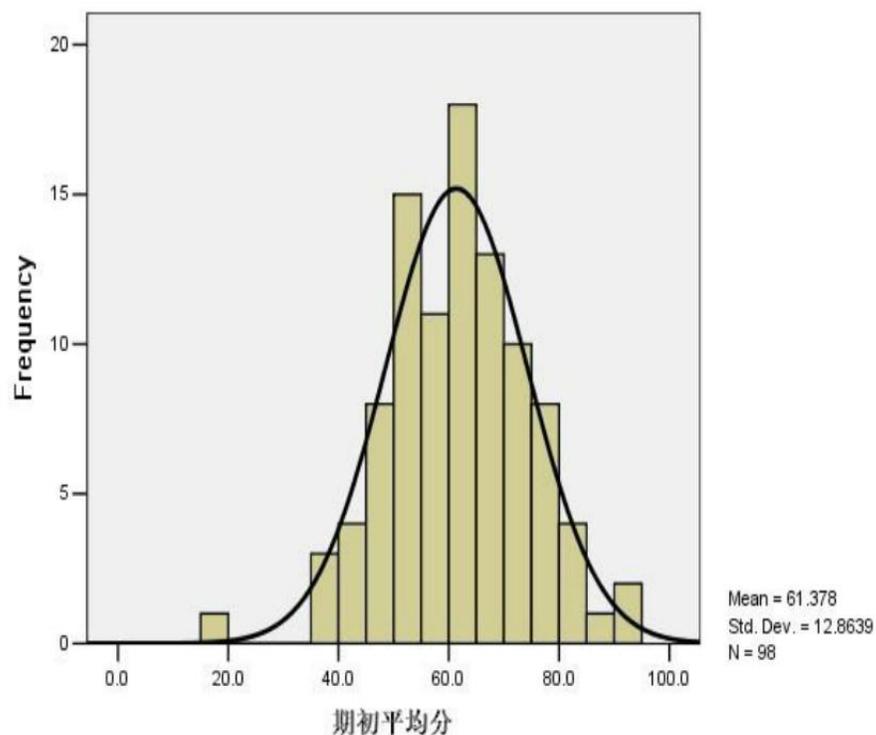
扁平分布



尖峰分布

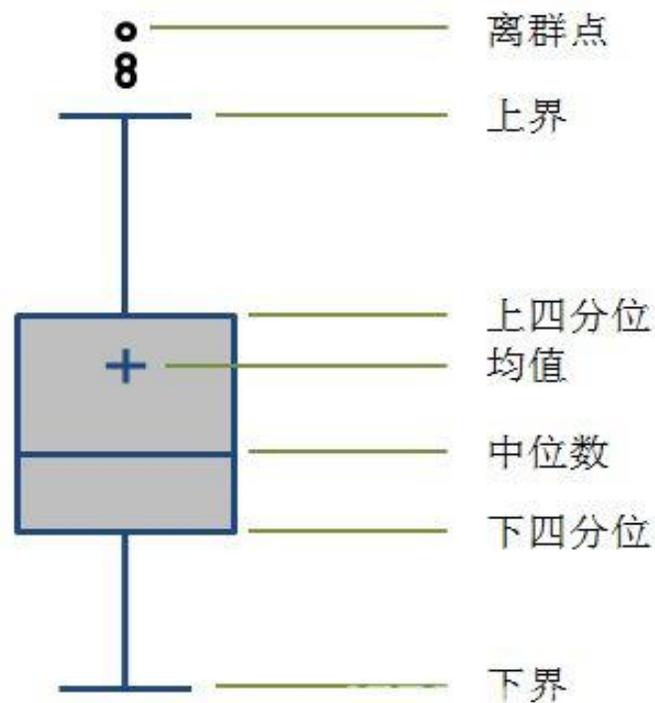
与标准正态分布比较!

第三节 数据的分布形态反映——数据形态图



直方图

用直条长短来表现连续性变量
取值分布特点的统计图形



箱图

又称箱线图，利用矩形框、线段和
点的位置对数据进行整体描述

第三节 数据的分布形态反映——数据形态图

```
height Stem-and-Leaf Plot
Frequency  Stem & Leaf
  2.00      12 . 23
  4.00      12 . 4555
  5.00      12 . 66777
  3.00      12 . 889
  5.00      13 . 01111
  6.00      13 . 222333
 11.00      13 . 44444555555
  9.00      13 . 666666667
 19.00      13 . 88888888999999999
 14.00      14 . 00000001111111
 14.00      14 . 22222222223333
  7.00      14 . 4444455
  9.00      14 . 666667777
  1.00      14 . 8
  4.00      15 . 0000
  2.00      15 . 22
  3.00      15 . 445
  2.00 Extremes (>=157)
Stem width:  10.00 (茎宽)
Each leaf:   1 case(s)
```

茎叶图

将数据按位数进行比较，自左至右

分别为：频数、茎、叶

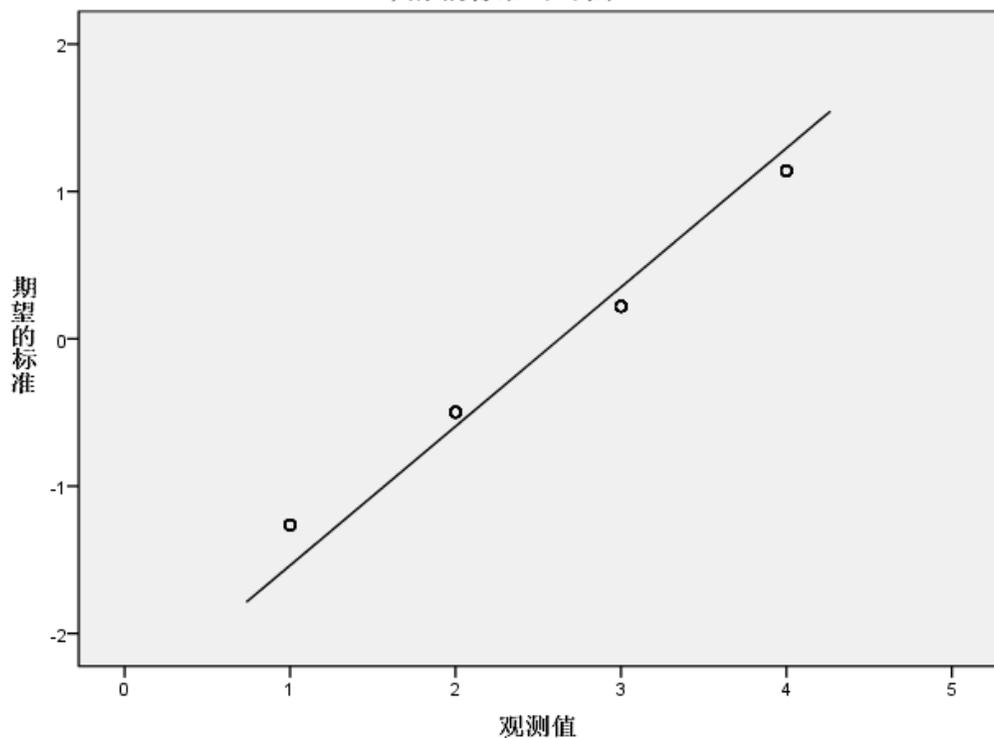
实际数据 = 茎值 * 茎宽 + 叶值

第三节 数据的分布形态反映——数据形态图

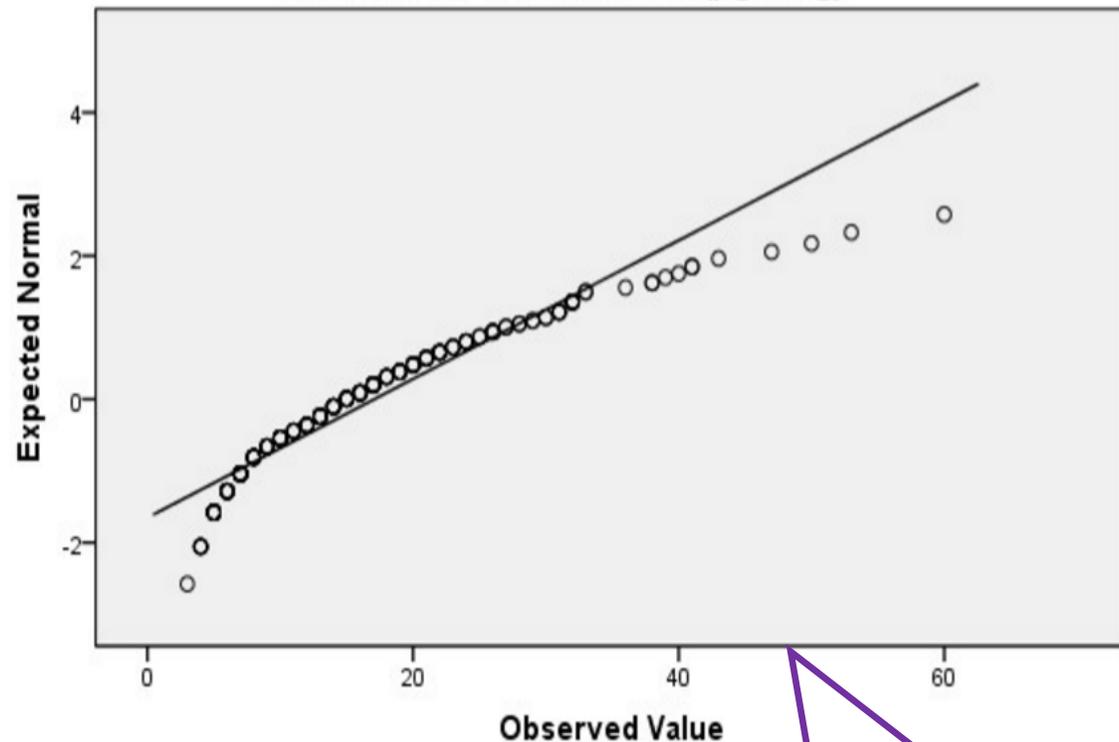
Q-Q图

用于验证样本数据是否服从某个特定分布

年龄的标准 Q-Q 图



Normal Q-Q Plot of 血铅含量($\mu\text{g}/100\text{g}$)



近似服从正态分布

不服从正态分布

第四节 探索分析

01 研究背景

对采样数据的性质
和分布特点未知

02 研究目的

检查数据是否有错误
获取数据分布特征
初步了解数据规律

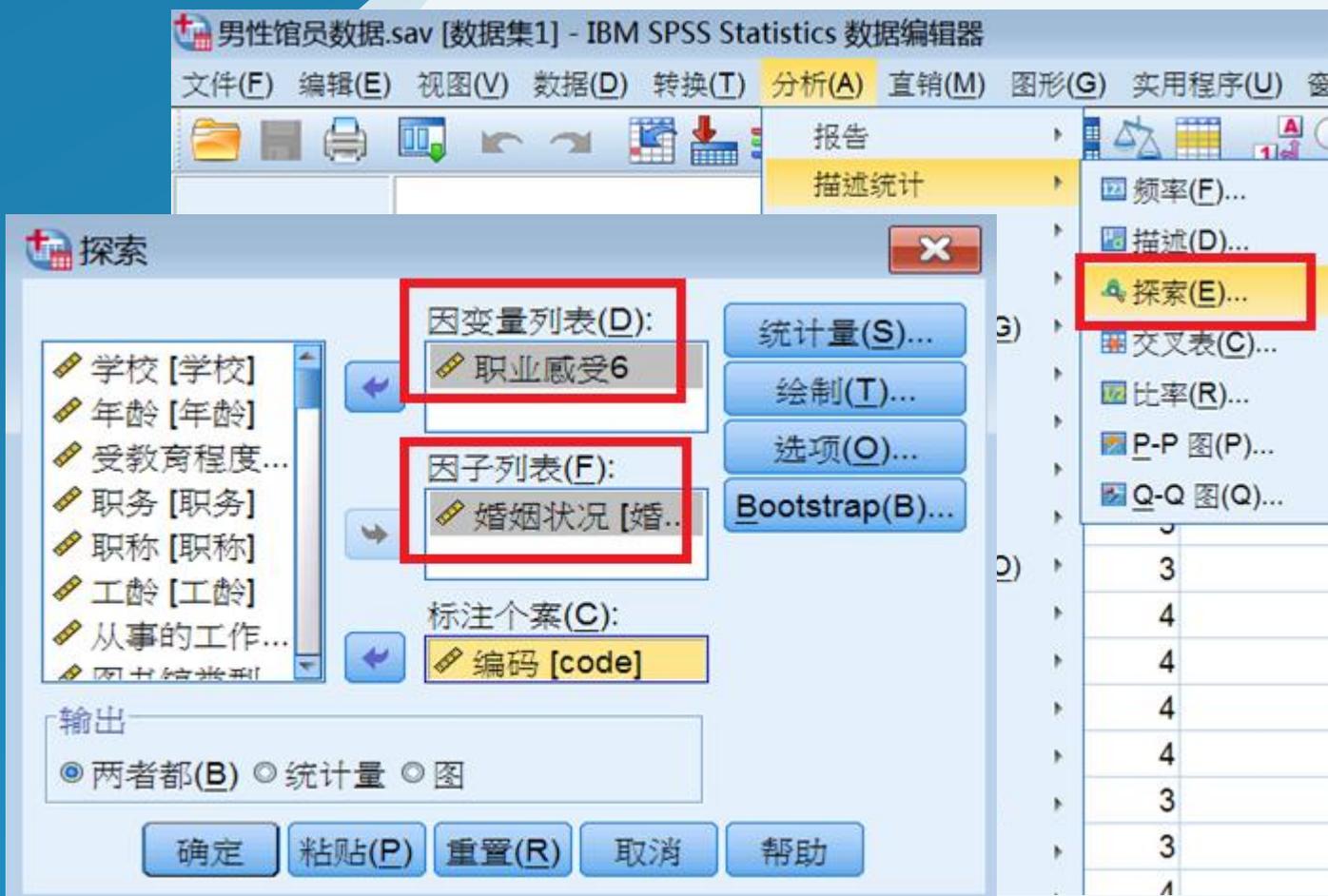
03 统计功能

绘制箱图、茎叶图等形态图
正态性检验
方差齐性检验

案例

题目：自从从事这份工作，我对工作越来越不感兴趣

选项：1-从不 2-极少 3-偶尔 4-经常 5-频繁 6-非常频繁 7-每天



案例分析结果

案例处理摘要

婚姻状况		案例					
		有效		缺失		合计	
		N	百分比	N	百分比	N	百分比
职业感受6	未婚	13	100.0%	0	0.0%	13	100.0%
	已婚	49	100.0%	0	0.0%	49	100.0%

极值

婚姻状况			案例号	编码	值	
职业感受6	未婚	最高	1	12	10	7
			2	30	23	7
			3	21	15	5
			4	52	69	5
			5	34	25	3
		最低	1	57	86	1
			2	39	37	1
			3	33	25	1
			4	31	24	1
			5	7	6	1
已婚	最高	1	36	26	6	
		2	48	59	5	
		3	3	3	3	
		4	15	11	3	
		5	25	18	3 ^a	
	最低	1	61	95	1	
		2	60	91	1	
		3	55	78	1	
		4	54	77	1	
		5	53	72	1 ^b	

a. 上限值表中仅显示一部分具有值 3 的案例。

b. 下限值表中仅显示一部分具有值 1 的案例。

描述

婚姻状况			统计量	标准误	
职业感受6	未婚	均值	2.92	.635	
		均值的 95% 置信区间	下限	1.54	
			上限	4.31	
		5% 修整均值	2.80		
		中值	2.00		
		方差	5.244		
		标准差	2.290		
	极小值	1			
	极大值	7			
	范围	6			
	四分位距	4			
	偏度	.948	.616		
	峰度	-.616	1.191		
	已婚	未婚	均值	1.86	.149
均值的 95% 置信区间			下限	1.56	
			上限	2.16	
5% 修整均值			1.73		
中值			2.00		
方差			1.083		
标准差			1.041		
极小值		1			
极大值		6			
范围		5			
四分位距		1			
偏度		1.915	.340		
峰度		5.160	.668		

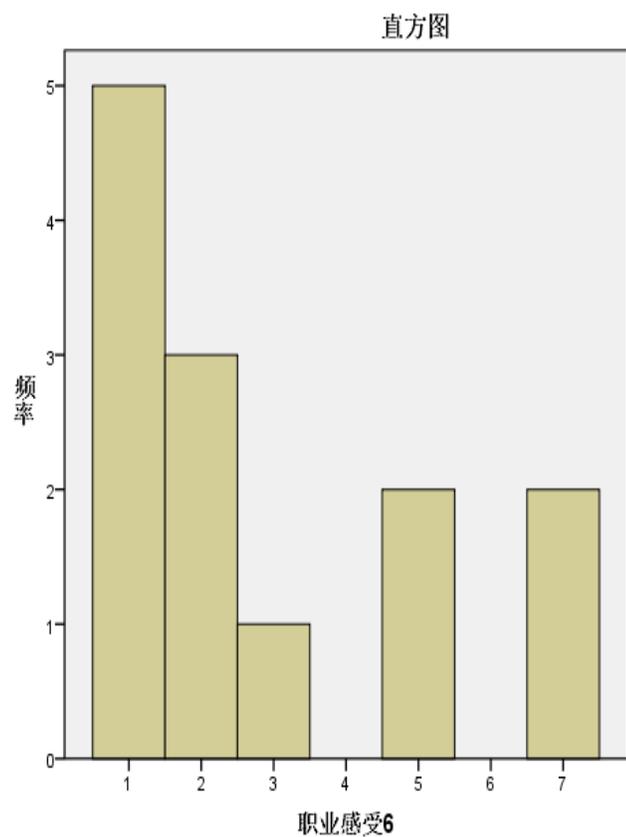
正态性检验

婚姻状况		Kolmogorov-Smirnov ^a			Shapiro-Wilk		
		统计量	df	Sig.	统计量	df	Sig.
职业感受6	未婚	.272	13	.009	.796	13	.006
	已婚	.262	49	.000	.740	49	.000

案例分析结果

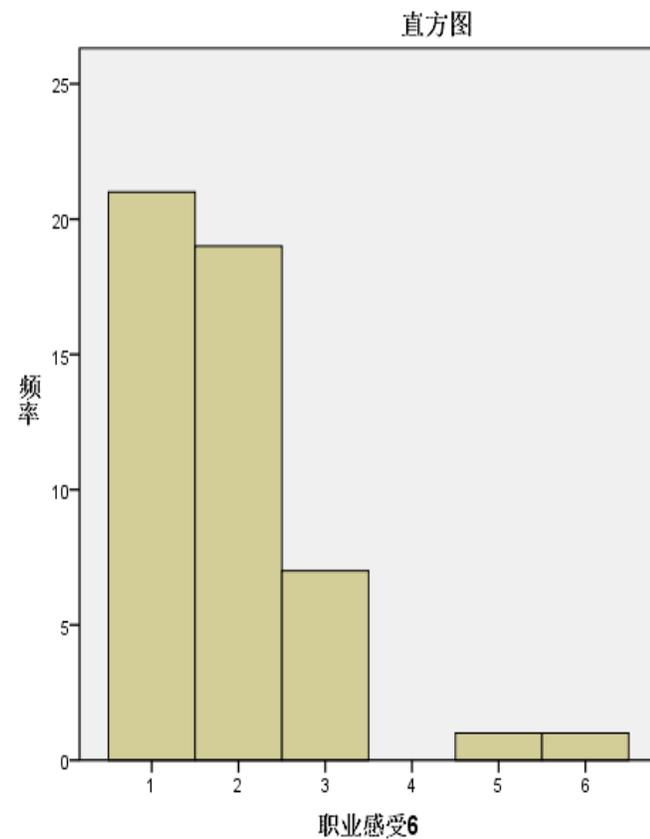
婚姻状况 = 未婚

直方图

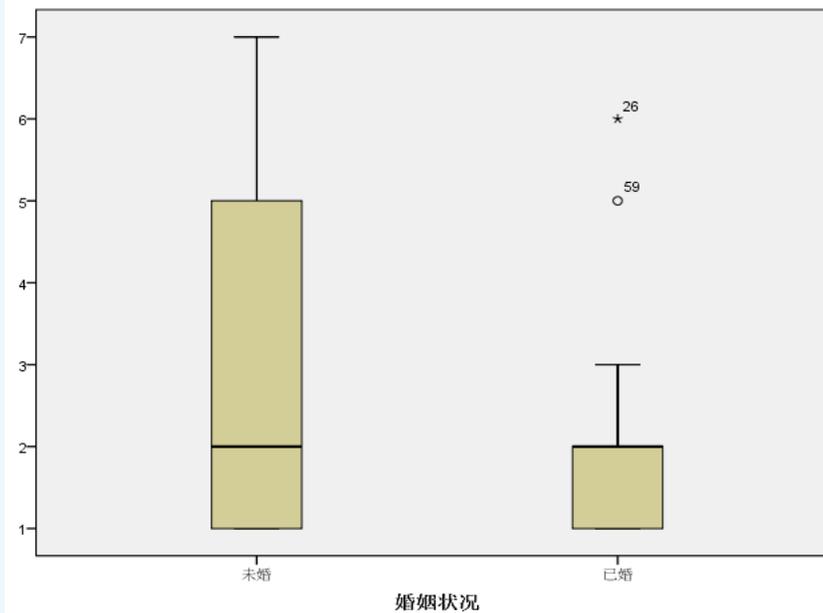
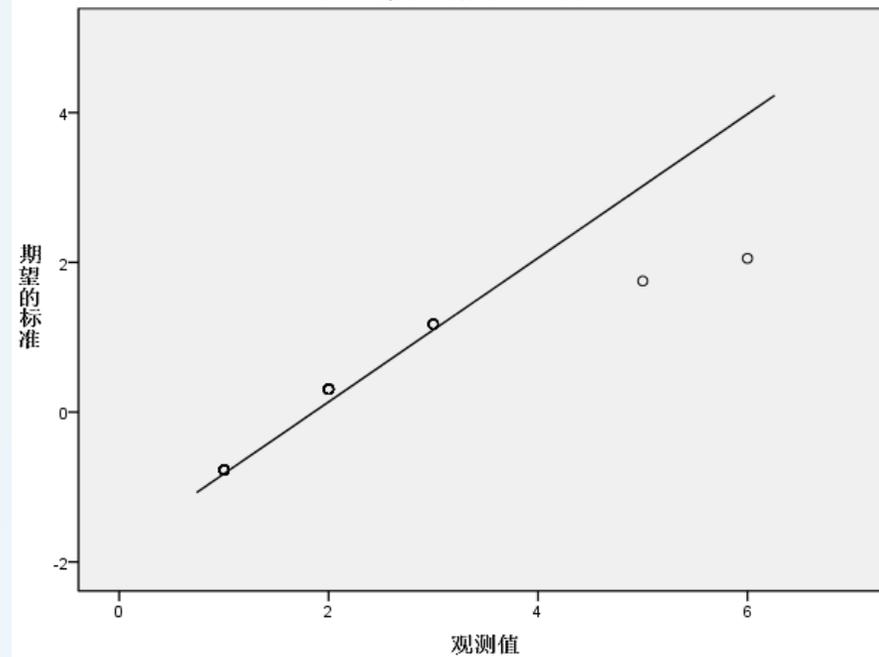


婚姻状况 = 已婚

直方图



职业感受6 的标准 Q-Q 图



02

大道至简
——因子分析

研究目的

缩减变量维度
简化观测系统
寻求研究主题基本结构

研究方法

原始变量分组，使组内变量相关性较高，组间相关性较低；再从每组变量中提炼出因子

核心思想

用**降维思想**将众多原始变量转化为少数综合性指标——因子

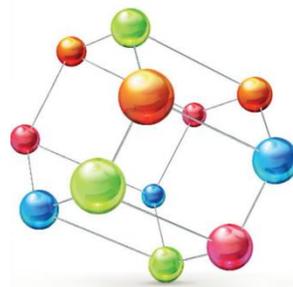
因子特点



个数远远少于
原始变量数目



反应原始变量
绝大部分信息



因子之间线性
关系不显著

my name is

具有命名解释性

因子分析步骤



前提条件的检验

Bartlett' s和KMO检验



因子提取

因子特征值 > 1
碎石图



使因子具有命名解释性

直交旋转
斜交旋转



计算各样本的因子得分

另存于数据文件最后
供后续分析使用

案例

男性馆员数据.sav [数据集1] - IBM SPSS Statistics 数据编辑器

文件(F) 编辑(E) 视图(V) 数据(D) 转换(T) 分析(A) 直销(M) 图形(G)



- 报告
- 描述统计
- 表(T)
- 比较均值(M)
- 一般线性模型(G)

因子分析

变量(V):

- 职业感受10
- 职业感受11
- 职业感受12
- 职业感受13
- 职业感受14
- 职业感受15
- 职业感受16

选择变量(C):

值(L)...

描述(D)... 抽取(E)... 旋转(T)... 得分(S)... 选项(O)...

确定 粘贴(P) 重置(R) 取消 帮助

因子分析: 描述统计

统计量

- 单变量描述性(U)
- 原始分析结果(I)

相关矩阵

- 系数(C)
- 逆模型(N)
- 显著性水平(S)
- 再生(R)
- 行列式(D)
- 反映象(A)
- KMO 和 Bartlett 的球形度检验(K)

继续 取消 帮助

因子分析: 抽取

方法(M): 主成份

分析

- 相关性矩阵(R)
- 协方差矩阵(V)

输出

- 未旋转的因子解(F)
- 碎石图(S)

抽取

- 基于特征值(E)
- 特征值大于(A): 1
- 因子的固定数量(N)

因子分析: 旋转

方法

- 无
- 最大方差法(V)
- 直接 Oblimin 方法(O)
- Pro

Delta(D): 0

输出

- 旋转解(R)
- 载荷图(L)

最大收敛性迭代次数(X): 25

继续 取消 帮助

因子分析: 因子得分

保存为变量(S)

方法

- 回归(R)
- Bartlett(B)
- Anderson-Rubin(A)

显示因子得分系数矩阵

继续 取消 帮助

因子分析: 选项

缺失值

- 按列表排除个案(L)
- 按对排除个案(P)
- 使用均值替换(R)

系数显示格式

- 按大小排序(S)
- 取消小系数(U)

绝对值如下(A): .10

继续 取消 帮助

案例分析结果

KMO 和 Bartlett 的检验

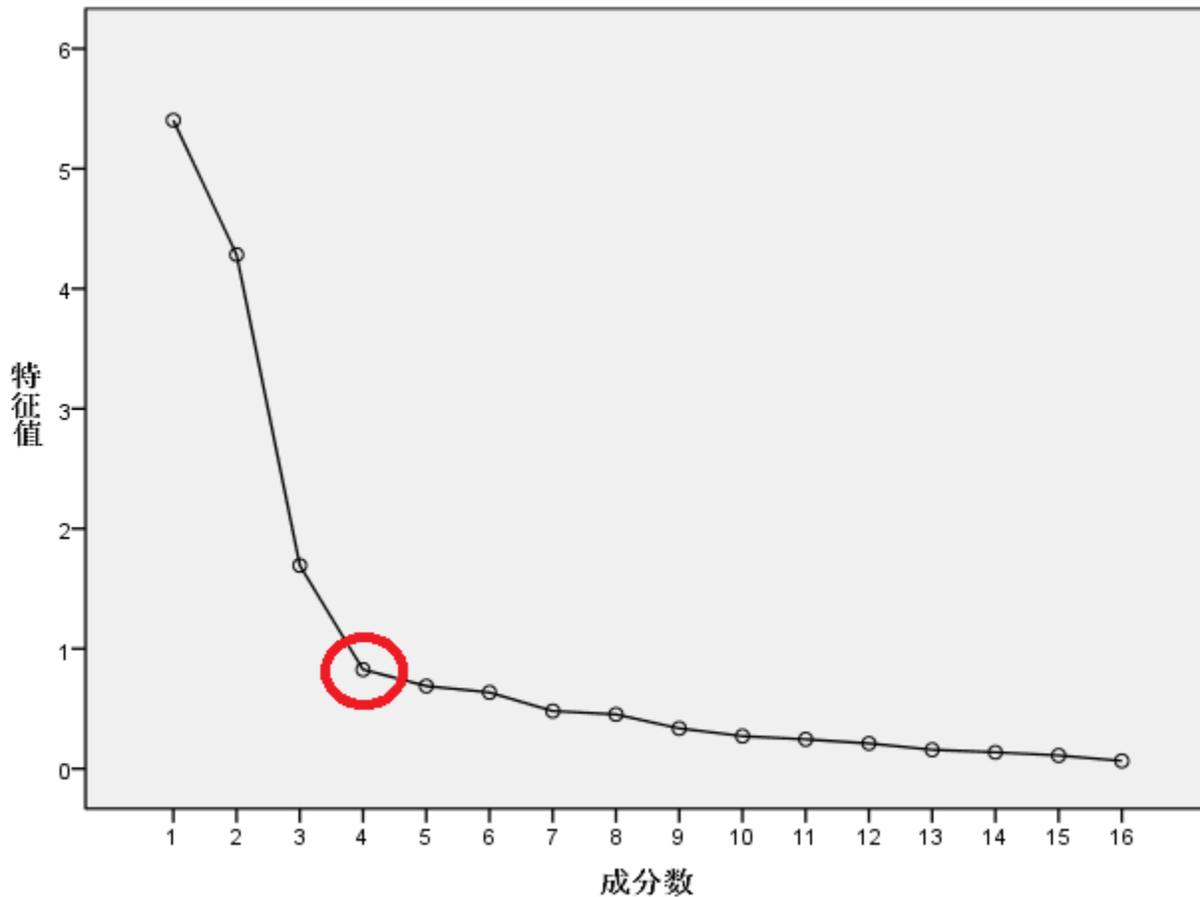
取样足够度的 Kaiser-Meyer-Olkin 度量。	<u>.766</u>
Bartlett 的球形度检验	近似卡方
	712.699
	df
	120
	Sig.
	<u>.000</u>

解释的总方差

成份	初始特征值			提取平方和载入			旋转平方和载入		
	合计	方差的 %	累积 %	合计	方差的 %	累积 %	合计	方差的 %	累积 %
1	5.404	33.772	33.772	5.404	33.772	33.772	4.409	27.554	27.554
2	4.284	26.775	60.547	4.284	26.775	60.547	3.667	22.921	50.475
3	1.694	10.586	71.133	1.694	10.586	71.133	3.305	20.658	71.133
4	.826	5.162	76.295						
5	.689	4.305	80.600						
6	.636	3.973	84.573						
7	.481	3.004	87.578						
8	.453	2.832	90.409						
9	.336	2.101	92.510						
10	.272	1.699	94.209						
11	.245	1.530	95.740						
12	.211	1.319	97.059						
13	.158	.989	98.047						
14	.137	.856	98.903						
15	.111	.692	99.595						
16	.065	.405	100.000						

提取方法：主成份分析。

碎石图



案例分析结果

因子轴旋转后

成份矩阵^a

	成份		
	1	2	3
职业感受13反向	.810	-.152	-.137
职业感受14反向	.791	-.254	.290
职业感受15反向	.785	-.256	.377
职业感受16反向	.769	-.200	-.028
职业感受11反向	.765	-.186	-.074
职业感受12反向	.754	-.437	.315
职业感受8	.704	.313	-.373
职业感受7	.565	.479	-.353
职业感受10反向	.483	-.449	.433
职业感受1	.052	.806	.139
职业感受2	.025	.799	.276
职业感受4	.144	.728	.315
职业感受3	.104	.728	.269
职业感受5	.136	.718	.423
职业感受6	.575	.605	-.166
职业感受9	.572	.327	-.647

提取方法：主成份。

a. 已提取了 3 个成份。

旋转成份矩阵^a

	成份		
	1	2	3
职业感受12反向	.915	-.134	.072
职业感受15反向	.900	.055	.106
职业感受14反向	.862	.015	.173
职业感受10反向	.759	-.118	-.179
职业感受16反向	.674	-.097	.411
职业感受11反向	.643	-.108	.446
职业感受13反向	.637	-.105	.531
职业感受5	.038	.843	.016
职业感受2	-.151	.828	.083
职业感受4	-.011	.799	.102
职业感受3	-.065	.772	.112
职业感受1	-.197	.770	.202
职业感受9	.025	.029	.923
职业感受8	.264	.166	.797
职业感受7	.100	.304	.756
职业感受6	.150	.506	.668

提取方法：主成份。

旋转法：具有 Kaiser 标准化的正交旋转法。

a. 旋转在 5 次迭代后收敛。

成就感低落因子

情绪衰竭因子

玩世不恭因子

案例分析结果

原始变量



	职业感受1	职业感受2	职业感受3	职业感受4	职业感受5	职业感受6
1	2	3	1	2	1	2
2	2	3	1	3	1	2
3	2	3	2	2	2	3
4	2	3	3	2	2	1
5	1	1	1	1	1	1
6	2	3	2	2	2	2
7	2	2	1	1	1	1
8	4	5	3	3	2	1
9	2	2	2	2	2	2
10	3	3	2	2	1	2

成份得分系数矩阵

	成份		
	1	2	3
职业感受1	-.025	.204	.010
职业感受2	.011	.242	-.053
职业感受3	.027	.227	-.047
职业感受4	.045	.241	-.062
职业感受5	.075	.270	-.110
职业感受6	-.015	.086	.183
职业感受7	-.058	.005	.252
职业感受8	-.029	-.031	.263
职业感受9	-.126	-.111	.367
职业感受10反向	.232	.046	-.168
职业感受11反向	.108	-.043	.101
职业感受12反向	.238	.020	-.087
职业感受13反向	.094	-.053	.136
职业感受14反向	.220	.050	-.058
职业感受15反向	.244	.075	-.096
职业感受16反向	.123	-.032	.081

提取方法:主成份。

旋转法:具有 Kaiser 标准化的正交旋转法。

构成得分。

FAC1 1成就感低落	FAC2 1情绪衰竭	FAC3 1玩世不恭
.81159	-.62091	.45054
-.46705	-.45873	-.12399
1.28471	-.10082	.20903
1.52040	.18577	-.96726
-.95503	-1.27805	-.44696
1.38837	-.07442	-.31483
1.30587	-.85977	-.09760
-.18124	.70264	-.32415
1.25461	-.25545	-.18843
.26309	-.60313	1.12437
1.96088	-.20123	-.08744

案例分析结果

职业倦怠量表因子分析

因子	对映量表题目内容	特征值	可解释的方差量
因子1：成就感低落因子	侧重考察工作中所获的价值和成就感	5.404	33.772%
因子2：情绪衰竭因子	侧重考察工作中是否身心疲惫、压力大	4.284	26.775%
因子3：玩世不恭因子	侧重考察对工作的热情和意义	1.694	10.586%

03

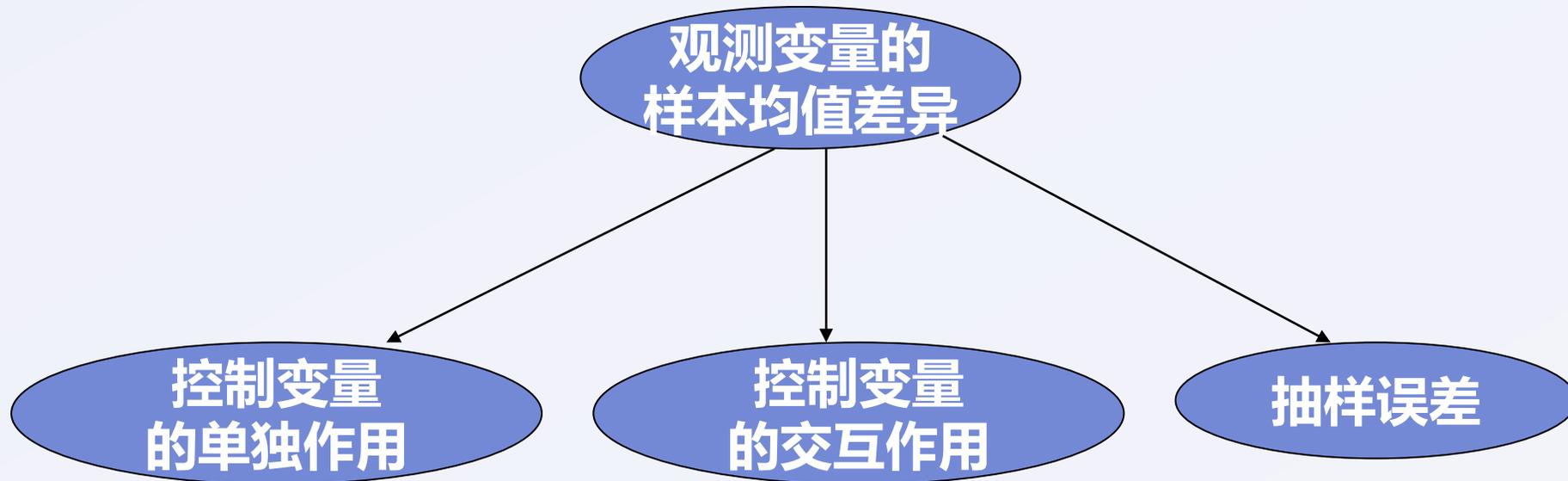
影响因素判断

——多因素方差分析

多因素方差分析

用来研究多个控制变量是否对单个观测变量产生显著性影响

功能：1、分析多个控制变量**单独**对观测变量的影响
2、分析多个控制变量的**交互作用**对观测变量的影响



检验步骤

01

建立原假设

假设控制变量的各自取值和交互作用对观测变量均不产生显著性影响

02

对样本进行F检验

得到统计量F值及相伴概率p值

03

规定显著性水平

显著性水平 $\alpha=0.05$

04

相伴概率与显著性水平比较

若 $p > \alpha$ ，接受原假设，认为控制变量对观测变量不产生显著影响
若 $p \leq \alpha$ ，拒绝原假设，认为控制变量对观测变量产生显著性影响

案例

分析男性馆员的婚姻状态和工作岗位是否对工作倦怠程度产生

男性馆员数据.sav [数据集1] - IBM SPSS Statistics 数据编辑器

文件(F) 编辑(E) 视图(V) 数据(D) 转换(T) 分析(A) 直销(M) 图形(G) 实用程序(U) 窗口

报告
描述统计
表(T)
比较均值(M)
一般线性模型(G)
广义线性模型
混合模型(X)
相关(C)
回归(R)
对数线性模型(O)
神经网络
分类(F)
降维

单变量

因变量(D):
玩世不恭因子...

固定因子(F):
婚姻状况 [...]
从事的工作...

模型(M)...
对比(N)...
绘制(T)...
两两比较(H)...
保存(S)...
选项(O)...
Bootstrap(B)...

单变量(U)...

方差分量	自由度	均方	F	显著性
婚姻状况	3			
从事的工作	3			
婚姻状况 * 从事的工作	4			
误差	4			

单变量: 选项

估计边际均值

因子与因子交互(F):
(OVERALL)
婚姻状况
从事的工作
婚姻状况*从事的工作

显示均值(M)

比较主效应(O)

置信区间调节(N):
LSD (无)

输出

描述统计(D)

方差齐性检验(H)

功效估计(E)

分布-水平图(P)

检验效能(B)

残差图(R)

参数估计(T)

缺失拟合优度检验(L)

单变量: 观测均值的两两比较

因子(F):
婚姻状况
从事的工作

两两比较检验(P):
从事的工作

假定方差齐性

LSD(L)

S-N-K(S)

Waller-Duncan(W)

Bonferroni(B)

Tukey

类型 I/类型 II 误差比率(I): 100

Sidak

Tukey s-b(K)

Dunnett(E)

Scheffe(C)

Duncan(D)

控制类别(Y): 最后一个

R-E-G-W-F

Hochberg's GT2(H)

检验

R-E-G-W-Q

Gabriel(G)

双侧(2) <控制(O) >控制(N)

未假定方差齐性

Tamhane's T2(M) Dunnett's T3(3) Games-Howell(A) Dunnett's C(U)

继续 取消 帮助

案例分析结果

误差方差等同性的 Levene 检验^a

因变量: 玩世不恭

主体间因子

Sig.

.242

	值标签	N
婚姻状况	0 未婚	13
	1 已婚	49
从事的工作	1 读者服务	28
	2 系统支持	7
	3 数字信息	9
	5 行政管理	8
	6 其他	10

主体间效应的检验

因变量: 玩世不恭因子FAC3_1

源	III 型平方和	df	均方	F	Sig.
校正模型	29.166 ^a	8	3.646	6.070	.000
截距	6.841	1	6.841	11.390	.001
婚姻	15.442	1	15.442	25.710	.000
岗位	16.147	4	4.037	6.721	.000
婚姻 * 岗位	7.071	3	2.357	3.924	.013
误差	31.834	53	.601		
总计	61.000	62			
校正的总计	61.000	61			

a. R 方 = .478 (调整 R 方 = .399)

描述性统计量

因变量: 玩世不恭因子FAC3_1

婚姻状况	从事的工作	均值	标准 偏差	N
未婚	读者服务	.4637216	1.52481154	4
	系统支持	-.1837556	.22274410	5
	数字信息	3.2429199	.02226585	2
	其他	1.3316924	1.62491904	2
总计				

多个比较

因变量: 玩世不恭因子FAC3_1

LSD

(I) 从事的工作	(J) 从事的工作	均值差值 (I-J)	标准 误差	Sig.	95% 置信区间	
					下限	上限
读者服务	系统支持	.1816319	.32750233	.582	-.4752545	.8385184
	数字信息	-.9793363*	.29696755	.002	-1.5749778	-.3836948
	行政管理	.0535165	.31069599	.864	-.5696608	.6766937
	其他	-.3172210	.28550991	.272	-.8898814	.2554394
系统支持	读者服务	-.1816319	.32750233	.582	-.8385184	.4752545
	数字信息	-1.1609682*	.39056932	.004	-1.9443511	-.3775854
	行政管理	-.1281155	.40110680	.751	-.9326338	.6764029
	其他	-.4988529	.38193007	.197	-1.2649076	.2672018
数字信息	读者服务	.9793363*	.29696755	.002	.3836948	1.5749778
	系统支持	1.1609682*	.39056932	.004	.3775854	1.9443511
	行政管理	1.0328528*	.37658813	.008	.2775126	1.7881929
	其他	.6621153	.35609343	.069	-.0521177	1.3763483
行政管理	读者服务	-.0535165	.31069599	.864	-.6766937	.5696608
	系统支持	.1281155	.40110680	.751	-.6764029	.9326338
	数字信息	-1.0328528*	.37658813	.008	-1.7881929	-.2775126
	其他	-.3707375	.36762046	.318	-1.1080907	.3666158
其他	读者服务	.3172210	.28550991	.272	-.2554394	.8898814
	系统支持	.4988529	.38193007	.197	-.2672018	1.2649076
	数字信息	-.6621153	.35609343	.069	-1.3763483	.0521177
	行政管理	.3707375	.36762046	.318	-.3666158	1.1080907

基于观测到的均值。

误差项为均值方 (错误) = .601。

*. 均值差值在 .05 级别上较显著。

作图

男性馆员数据.sav [数据集1] - IBM SPSS Statistics 数据编辑器

文件(F) 编辑(E) 视图(V) 数据(D) 转换(T) 分析(A) 直销(M) 图形(G) 实用程序(U) 窗口(W) 帮助

线图

- 简单
- 多线线图**
- 垂直线图

图表中的数据为

- 个案组摘要(G)**
- 各个变量的摘要(V)
- 个案值(I)

定义 取消 帮助

code	婚姻状况	读者服务	系统支持	数字信息	行政管理	其他
1	1	3	3			
2	1	3	1			
3	1	3	2			
4	1	3	2			
5	1	3	2			
6	1	3	1			
7	0	4	1			
8	1	4	2			
9	1	4	1			
10	1	4	1			
11	1	3	1			

定义多线线图: 个案组摘要

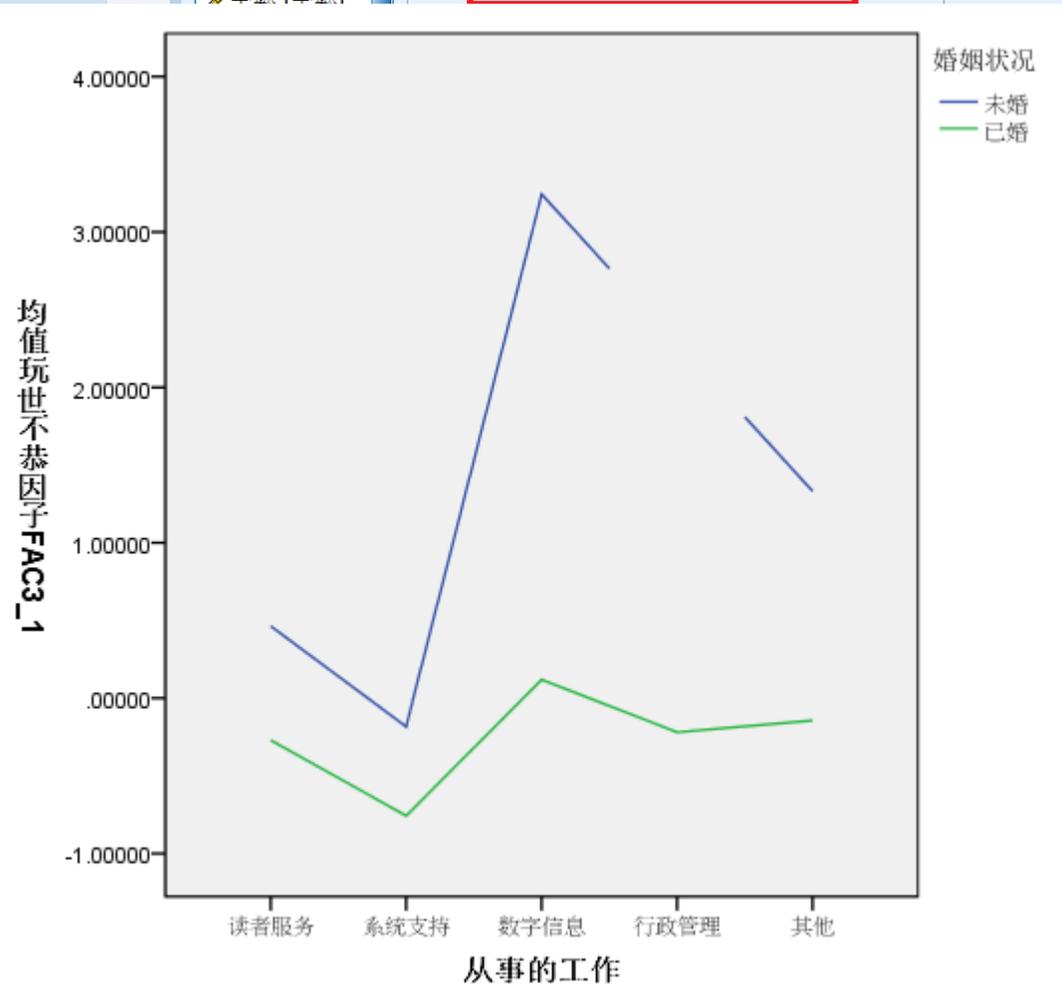
线的表征

- 个案数(N)
- 个案数的 %(A)
- 累积个数(C)
- 累积 %(M)
- 其他统计量 (例如均值) (S)

变量(V):

MEAN((玩世不恭因子FAC3_1))

标题(T)... 选项(O)...



案例分析结果

两个控制变量对观测变量的作用力大小

--观测变量：玩世不恭因子

控制变量：婚姻+岗位								
未婚				已婚				
读者服务	系统支持	数字信息	其它	读者服务	系统支持	数字信息	行政管理	其它
1.375	0.800	5.500	3.125	0.716	0.500	1.375	0.813	0.857

注：样本中无行政管理岗未婚男性

04

万物皆有联系
——相关分析



函数关系

变量之间是一对一或多对一的**确定性关系**

检验方法：T检验、方差分析



统计关系

一个变量**不能**由一个或多个变量**唯一确定**，
变量之间存在着某种或强或弱的联系

检验方法：相关分析

研究目的

衡量变量之间的线性相关的**强弱程度**和**方向**

二元定距变量相关分析

通过计算两个定距变量的相关系数，来揭示变量间的相关程度



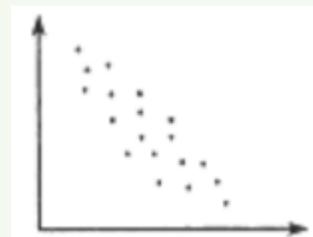
01

计算相关系数

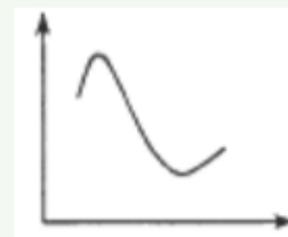
$$-1 \leq r \leq +1$$



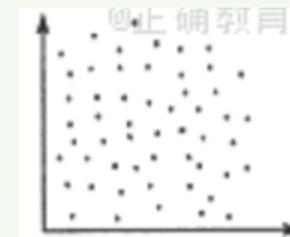
正相关



负相关



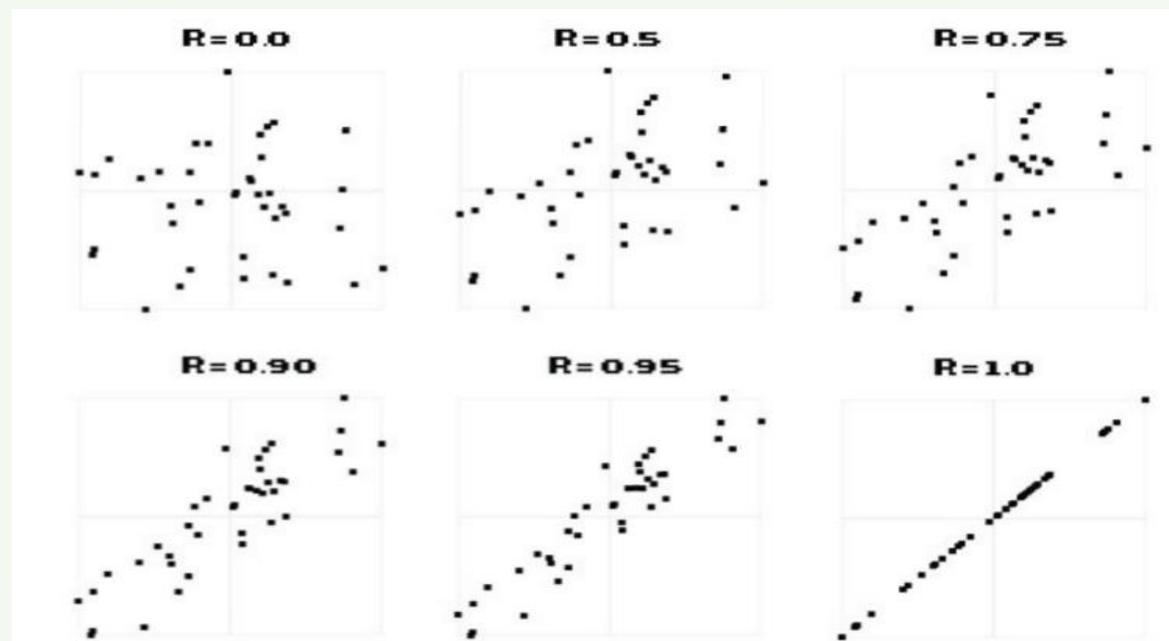
函数关系



不存在线性关系

02

绘制散点图



01

建立原假设

假设两个总体之间没有显著线性相关关系

02

计算统计量

得到统计量值及相伴概率p值

03

规定显著性水平

显著性水平 $\alpha=0.05$

04

相伴概率与显著性水平比较

若 $p > \alpha$ ，接受原假设，认为两总体不存在显著的线性相关关系

若 $p \leq \alpha$ ，拒绝原假设，认为两总体存在显著的线性相关关系

考察员工工作上玩世不恭程度与晋升机会之间的关系

男性馆员数据.sav [数据集1] - IBM SPSS Statistics 数据编辑器

文件(F) 编辑(E) 视图(V) 数据(D) 转换(T) 分析(A) 直销(M) 图形(G) 实用程

12: 职业

1			
2			
3			
4			
5			
6			
7			
8			
9			
10			

报告
描述统计
表(T)
比较均值(M)
一般线性模型(G)
广义线性模型
混合模型(X)
相关(C)
回归(R)
对数线性模型(O)
神经网络
分类(F)
降维
度量(S)

双变量相关性:...

统计量

- 均值和标准差(M)
- 叉积偏差和协方差(C)

缺失值

- 按对排除个案(P)
- 按列表排除个案(L)

继续 取消 帮助

双变量相关

变量(V):

- 编码 [code]
- 学校 [学校]
- 情绪衰竭1 (职业...)
- 情绪衰竭2 (职业...)
- 情绪衰竭3 (职业...)
- 情绪衰竭4 (职业...)
- 情绪衰竭5 (职业...)
- 玩世不恭1 (职业...)
- 玩世不恭2 (职业...)

玩世不恭平均分
晋升机会平均分

选项(O)...
Bootstrap(B)...

相关系数

- Pearson
- Kendall的 tau-b(K)
- Spearman

显著性检验

- 双侧检验(T)
- 单侧检验(L)

标记显著性相关(F)

确定 粘贴(P) 重置(R) 取消 帮助

案例分析结果

描述性统计量

	均值	标准差	N
玩世不恭平均分	1.0943	1.27407	61
晋升机会平均分	2.3280	.84821	62

相关性

		玩世不恭平均分	晋升机会平均分
玩世不恭平均分	Pearson 相关性	1	-.327*
	显著性 (双侧)		.010
	N	61	61
晋升机会平均分	Pearson 相关性	-.327*	1
	显著性 (双侧)	.010	
	N	61	62

*. 在 0.05 水平 (双侧) 上显著相关。

案例分析结果

	code	学校	职业感受1	职业感受2	职业感受3	职业感受4
1	10	4	2	2	4	2
2	17	4	0	0	0	0
3	18	4	1	1	0	2
4	23	4	2	2	5	2
5	24	4	1	1	1	0
6	25	4	1	1	1	0
7	26	4	2	2	2	1
8	2	3	1	2	0	1
9						
10						
11						

散点图/点图

简单分布
 矩阵分布
 简单点
 重叠分布
 3-D分布

简单散点图

Y轴(Y): **玩世不恭平均分**

X轴(X): **晋升机会平均分**

05

撰写研究成果的若
干建议

撰写研究报告

大纲	CHECK
绪论	
研究背景	
研究设计	
数据采集过程	
数据分析	
个案访谈	
研究结论	
研究局限性和未来展望	
参考文献和附录	

注意事项

- 完整描述研究的流程
- 阐述研究视角，突出设计**亮点**
- 合理采集样本，证实数据的客观性
- 从**多角度**的深入分析数据
- 将客观数据与主观资料有机**结合**
- 研究结论要有理有据

撰写论文

纵向选题角度

文献综述	问卷设计	样本采集	数据处理
聚类分析 可视化工具	设计是亮点 阐述不足之处	整体描述 缺失值研究	简化数据结构 揭示数据内部规律 揭示数据间的差异 探索现象的影响因素

横向选题角度

- 不同研究方法的对比
- 对调研对象进行追踪调查

注意事项

选择统计
方法需谨
慎

确保数据吻合统计
方法的使用前提

分析结果
的选择

选择对象有代表性
结果表达多样化
分析结果简化

侧重对结
果的诠释

结合访谈内容
进行归因分析

谢谢聆听



<http://www.lib.nankai.edu.cn/libservice/zjzt>



nkzyz@nankai.edu.cn



85358050

南开大学图书馆