

信171 李金哲

数据的基本统计与参数检验 实验报告



研究内容

1

“住房状况调查”统计量分析

2

“住房状况调查”家庭收入显著性分析

3

“住房状况调查”家庭收入与10000元的显著性分析

4

“住房状况调查”婚姻状况对现住面积的显著性分析

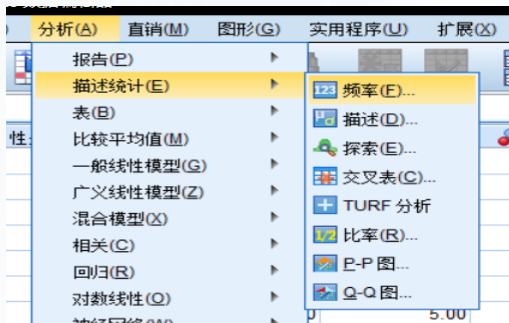
5

“减肥茶数据”减肥茶对减肥的显著性分析

“住房状况调查” 统计量分析

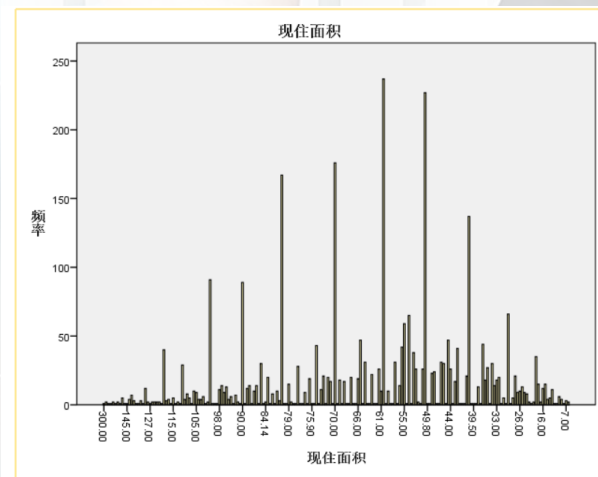
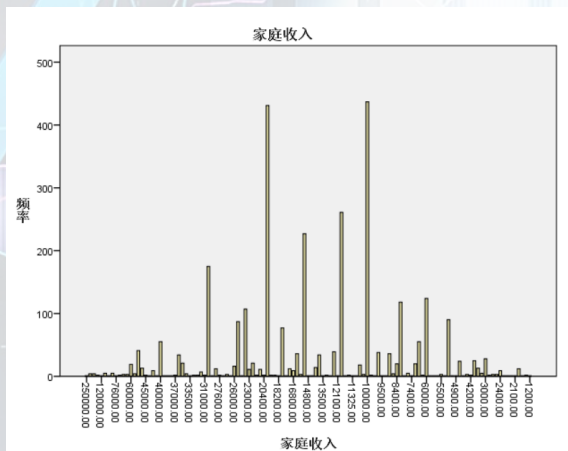
PART 1

过程显示



统计

个案数	家庭收入	现住面积
	有效	有效
	2993	2993
	缺失	缺失
	0	0
平均值	17696.1567	62.7241
中位数	15000.0000	60.0000
众数	10000.00	60.00
标准差	15298.80341	25.90383
方差	234053385.7	671.008
偏度	5.546	.910
偏度标准误差	.045	.045
峰度	55.425	3.078
峰度标准误差	.089	.089



结果分析

家庭收入分析

根据平均值和中位数可以明显得出，在家庭收入数据的中位数和平均数差值为2500，说明存在少量超高收入家庭数据对整体数据有显著地影响，但收入总体处于17696元。

数据的整体方差为234053385，说明该数据相对于均值的偏差平方的平均影响很大，显现了数据整体有离散的特性。

现住面积分析

该组数据的平均值、中位数和众数，三者的偏差不大，可以将值近似为60。可以体现出，总体现住面积处于60m²

该组数据的标准差为25，偏差较小，能体现该组数据相对集中于平均值60的附近。

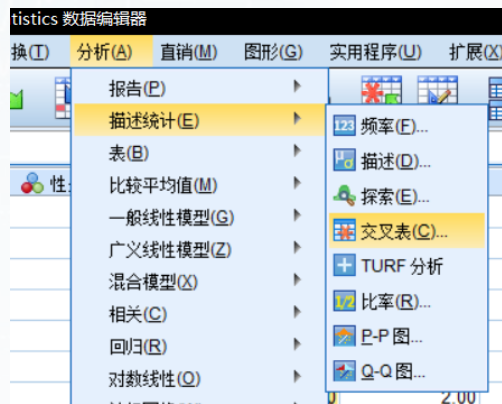
而且其偏度为0.91，能够近乎为0，刻画出数据关于均值近乎对称。

The background features a dark blue horizontal band across the middle. Below this, there are faint, stylized financial charts including line graphs and bar charts in light blue and purple. On the right side, the lower legs and feet of a person wearing dark trousers and light-colored shoes are visible, suggesting a professional or business context.

不同学历对家庭收入、现住面积是否有显著影响

PART 2

过程显示(不同学历对家庭收入)



文化程度 * 家庭收入 交叉表																	
		家庭收入															
		1200.00	1500.00	1800.00	2000.00	2100.00	2160.00	2292.00	2300.00	2400.00	2500.00	2600.00	2800.00	3000.00	3500.00	3600.00	
文化程度	初中及以下	计数	1	1	1	4	1	0	0	1	5	2	1	1	13	1	
		期望计数	.3	.5	.3	3.2	.3	.3	.3	.3	2.4	.8	.8	.5	7.5	1.3	
		占文化程度的百分比	0.1%	0.1%	0.1%	0.5%	0.1%	0.0%	0.0%	0.1%	0.6%	0.2%	0.1%	0.1%	1.6%	0.1%	
	高中（中专）	计数	0	1	0	6	0	1	1	0	4	1	2	1	13	4	
		期望计数	.4	.8	.4	5.0	.4	.4	.4	.4	3.8	1.3	1.3	.8	11.8	2.1	
		占文化程度的百分比	0.0%	0.1%	0.0%	0.5%	0.0%	0.1%	0.1%	0.0%	0.3%	0.1%	0.2%	0.1%	1.0%	0.3%	
	大学（专、本科）	计数	0	0	0	2	0	0	0	0	0	0	0	0	2	0	
		期望计数	.3	.6	.3	3.6	.3	.3	.3	.3	2.7	.9	.9	.6	8.4	1.5	
		占文化程度的百分比	0.0%	0.0%	0.0%	0.2%	0.0%	0.0%	0.0%	0.0%	0.0%	0.0%	0.0%	0.0%	0.2%	0.0%	
	研究生及以上	计数	0	0	0	0	0	0	0	0	0	0	0	0	0	0	
		期望计数	.0	.0	.0	.1	.0	.0	.0	.0	.1	.0	.0	.0	.3	.1	
		占文化程度的百分比	0.0%	0.0%	0.0%	0.0%	0.0%	0.0%	0.0%	0.0%	0.0%	0.0%	0.0%	0.0%	0.0%	0.0%	
总计	计数	1	2	1	12	1	1	1	1	9	3	3	2	28	5		
	期望计数	1.0	2.0	1.0	12.0	1.0	1.0	1.0	1.0	9.0	3.0	3.0	2.0	28.0	5.0		
	占文化程度的百分比	0.0%	0.1%	0.0%	0.4%	0.0%	0.0%	0.0%	0.0%	0.3%	0.1%	0.1%	0.1%	0.9%	0.2%		

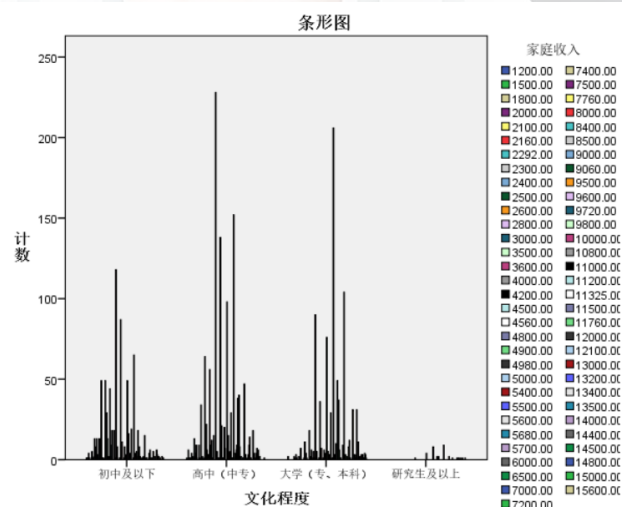


卡方检验

	值	自由度	渐进显著性 (双侧)
皮尔逊卡方	958.189 ^a	360	.000
似然比(L)	856.252	360	.000
线性关联	260.521	1	.000
有效个案数	2993		

a. 391 个单元格 (80.8%) 的期望计数小于 5。最小期望计数为 .01。

对称测量			
		值	渐进显著性
名义到名义	Phi	.566	.000
	克莱姆 V	.327	.000
有效个案数		2993	



结果分析

卡方检验

	值	自由度	渐进显著性 (双侧)
皮尔逊卡方	958.189 ^a	360	.000
似然比(L)	856.252	360	.000
线性关联	260.521	1	.000
有效个案数	2993		

a. 391 个单元格 (80.8%) 的期望计数小于 5。最小期望计数为 .01。

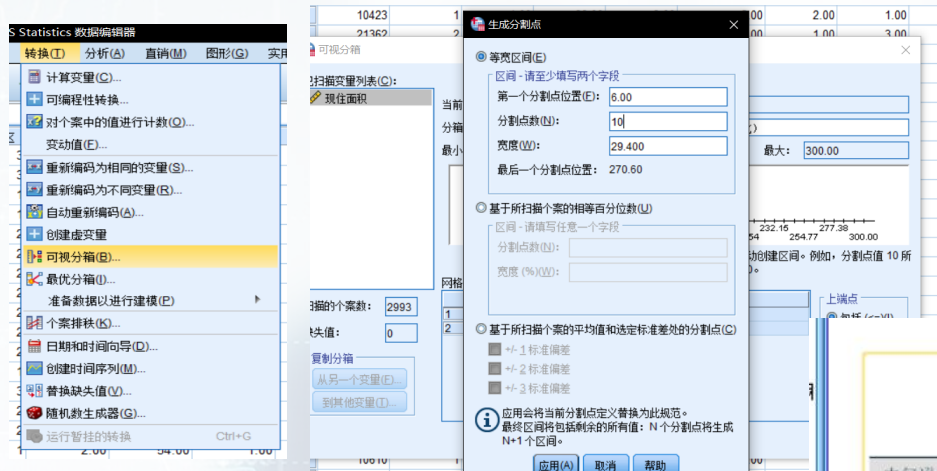
对称测量

		值	渐进显著性
名义到名义	Phi	.566	.000
	克莱姆 V	.327	.000
有效个案数		2993	

假设：不同学历对家庭收入是否有显著影响

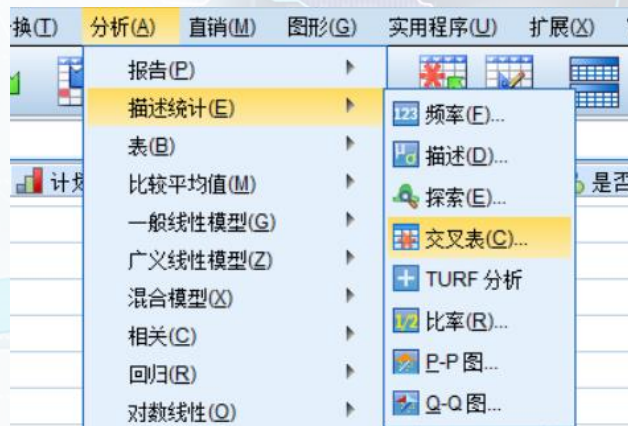
设显著性水平 α 设为0.05，由于卡方的概率 p -值小于 α ，因此应拒绝假设，认为现住面积与家庭收入是否无显著影响。这种无影响主要体现在现住面积小的比例在低学历中低于总体比例，而现住面积大的在比例在低学历中高于总体比例。

过程显示(不同学历对现住面积)



文化程度 * 现住面积 (分箱化) 交叉表

文化程度			现住面积 (分箱化)							总计
			<= 6.00	6.01 - 33.69	33.70 - 64.80	64.81 - 94.20	94.21 - 123.60	123.61 - 153.00	153.01 - 182.40	270.61 +
初中及以下	计数		1	134	432	183	40	10	4	805
	占文化程度的百分比		0.1%	16.6%	53.7%	22.7%	5.0%	1.2%	0.5%	100.0%
高中 (中专)	计数		0	129	656	374	84	13	2	1258
	占文化程度的百分比		0.0%	10.3%	52.1%	29.7%	6.7%	1.0%	0.2%	100.0%
大学 (专、本科)	计数		1	44	336	333	154	25	3	896
	占文化程度的百分比		0.1%	4.9%	37.5%	37.2%	17.2%	2.8%	0.3%	100.0%
研究生及以上	计数		0	0	11	16	6	1	0	34
	占文化程度的百分比		0.0%	0.0%	32.4%	47.1%	17.6%	2.9%	0.0%	100.0%
总计	计数		2	307	1435	906	284	49	9	2993
	占文化程度的百分比		0.1%	10.3%	47.9%	30.3%	9.5%	1.6%	0.3%	100.0%



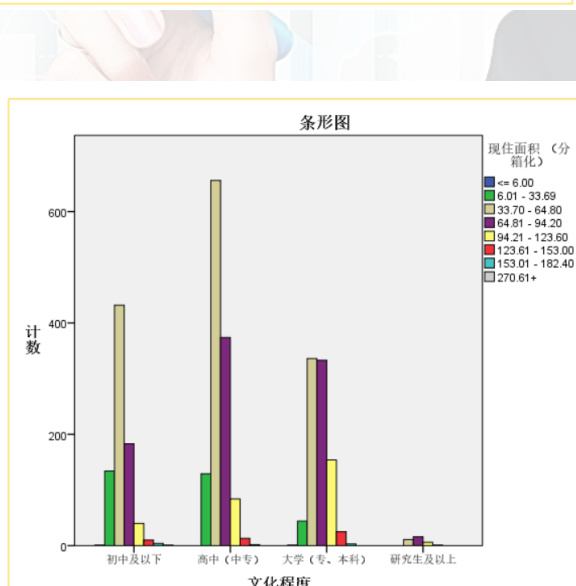
卡方检验

	值	自由度	渐进显著性 (双侧)
皮尔逊卡方	228.808 ^a	21	.000
似然比(L)	228.126	21	.000
线性关联	164.859	1	.000
有效个案数	2993		

a. 15 个单元格 (46.9%) 的期望计数小于 5。最小期望计数为 .01。

对称测量

	值	渐进显著性
名义到名义	Phi	.276
	克莱姆 V	.160
有效个案数	2993	



结果分析

卡方检验

	值	自由度	渐进显著性 (双侧)
皮尔逊卡方	228.808 ^a	21	.000
似然比(L)	228.126	21	.000
线性关联	164.859	1	.000
有效个案数	2993		

a. 15 个单元格 (46.9%) 的期望计数小于 5。最小期望计数为 .01。

对称测量

		值	渐进显著性
名义到名义	Phi	.276	.000
	克莱姆 V	.160	.000
有效个案数		2993	

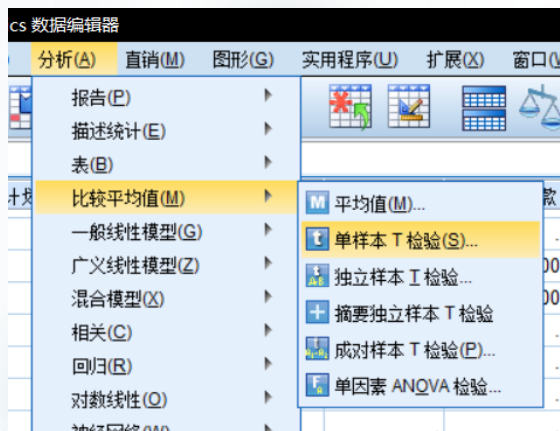
假设：不同学历对现住面积是否有显著影响

设显著性水平 α 设为0.05，由于卡方的概率 p -值小于 α ，因此应拒绝假设，认为现住面积与家庭收入是否无显著影响。这种无影响主要体现在现住面积小的比例在低学历中低于总体比例，而现住面积大的在比例在低学历中高于总体比例。

“住房状况调查” 家庭收入与10000元

PART 3

过程显示



单样本统计

	个案数	平均值	标准差	标准误差平均值
家庭收入	2993	17696.1567	15298.80341	279.64310

单样本检验

检验值 = 10000

	t	自由度	显著性 (双尾)	平均值差值	差值 95% 置信区间 下限	上限
家庭收入	27.521	2992	.000	7696.15670	7147.8445	8244.4689

结果分析

家庭收入与10000元分析

该组数据的平均值、中位数和众数，三者的偏差不大，可以将值近似为60。可以体现出，总体现住面积处于60m²

该组数据的标准差为25，偏差较小，能体现该组数据相对集中于平均值60的附近。

而且其偏度为0.91，能够近乎为0，刻画出数据关于均值近乎对称。

单样本统计

	个案数	平均值	标准差	标准误差平均值
家庭收入	2993	17696.1567	15298.80341	279.64310

单样本检验

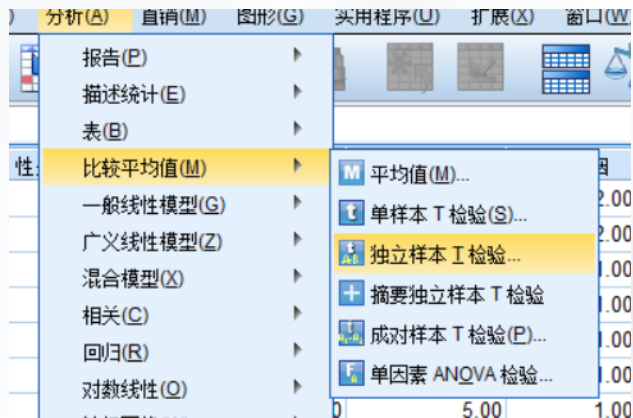
检验值 = 10000

	t	自由度	显著性 (双尾)	平均值差值	差值 95% 置信区间 下限	上限
家庭收入	27.521	2992	.000	7696.15670	7147.8445	8244.4689

婚姻状况对现住面积的显著性分析

PART 4

过程显示



组统计				
	婚姻	个案数	平均值	标准差
分组后的人均面积	已婚	2684	2.4460	.87436
	否	308	2.8344	.96275

双击以
激活

独立样本检验

莱文方差等同性检验				平均值等同性 t 检验					
		F	显著性	t	自由度	显著性 (双尾)	平均值差值	标准误差差值	差值 95% 置信区间 下限 上限
分组后的人均面积	假定等方差	12.268	.000	-7.305	2990	.000	-.38844	.05317	-.49270 -.28418
	不假定等方差			-6.768	367.489	.000	-.38844	.05740	-.50130 -.27557

结果分析

组统计					
	婚姻	个案数	平均值	标准差	标准误差平均值
分组后的人均面积	已婚	2684	2.4460	.87436	.01688
	否	308	2.8344	.96275	.05486

双击以
激活

独立样本检验

莱文方差等同性检验				平均值等同性 t 检验					
		F	显著性	t	自由度	显著性 (双尾)	平均值差值	标准误差差值	差值 95% 置信区间 下限 上限
分组后的人均面积	假定等方差	12.268	.000	-7.305	2990	.000	-.38844	.05317	-.49270 -.28418
	不假定等方差			-6.768	367.489	.000	-.38844	.05740	-.50130 -.27557

婚姻状况对现住面积的显著性分析

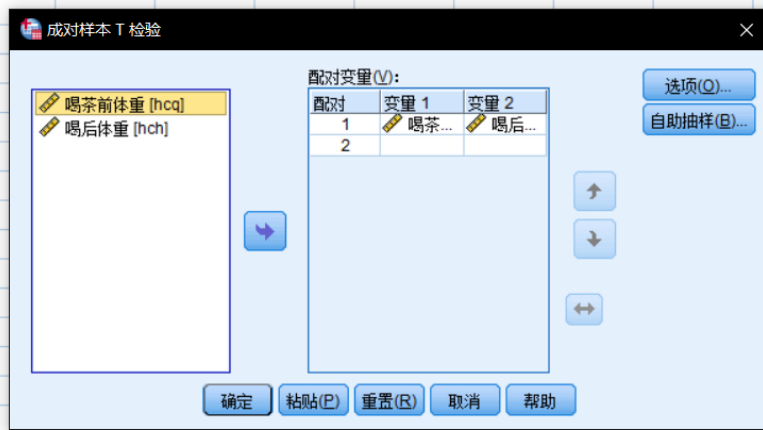
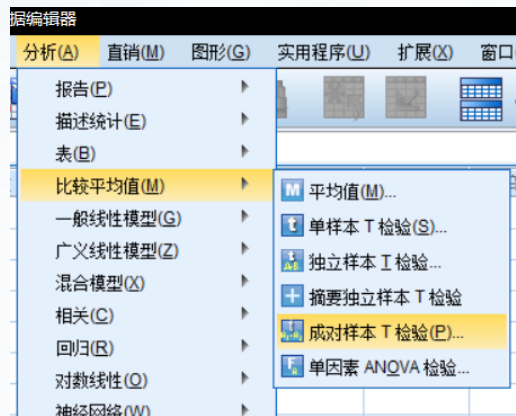
该检验的F统计量为12.268，对应的概率为0.000。如果显著水平 α 为0.05，由于概率P值小于0.05，可以认为两总体的方差有显著差异。

由于两样本方差有显著差异，因此应看第1行t检验的结果。其中，t统计量的观测值为-7.305，对应的双尾概率P值为0.000。如果显著性水平 α 为0.05，由于概率P小于0.05，可以认为两总体样本有显著差异，即已婚与未婚家庭对现住面积有显著差异。

减肥茶对减肥的显著性分析

PART 5

过程显示



配对样本统计

		平均值	个案数	标准差	标准误差平均值
配对 1	喝茶前体重	89.2571	35	5.33767	.90223
	喝茶后体重	70.0286	35	5.66457	.95749

配对样本相关性

		个案数	相关性	显著性
配对 1	喝茶前体重 & 喝茶后体重	35	-.052	.768

配对样本检验

		配对差值		标准误差平均值		差值 95% 置信区间		t	自由度	显著性 (双尾)
		平均值	标准差	标准误差平均值	标准误差平均值	下限	上限			
配对 1	喝茶前体重 - 喝茶后体重	19.22857	7.98191	1.34919	1.34919	16.48669	21.97045	14.252	34	.000

结果分析

配对样本统计

		平均值	个案数	标准差	标准误差平均值
配对 1	喝茶前体重	89.2571	35	5.33767	.90223
	喝茶后体重	70.0286	35	5.66457	.95749

配对样本相关性

		个案数	相关性	显著性
配对 1	喝茶前体重 & 喝茶后体重	35	-.052	.768

配对样本检验

		配对差值					t	自由度	显著性 (双尾)
		平均值	标准差	标准误差平均值	差值 95% 置信区间 下限	上限			
配对 1	喝茶前体重 - 喝茶后体重	19.22857	7.98191	1.34919	16.48669	21.97045	14.252	34	.000

减肥茶对减肥的显著性分析

由配对样本统计可得喝茶前和喝茶后样本的平均值有较大的差异。喝茶后的平均体重低于喝茶前的平均体重。

由配对样本检验，平均值可以得出喝茶前后体重的平均差异，相差了19.2公斤。设显著性水平 α 为0.05，由于概率P值小于显著性水平，应拒绝原假设，即认为总体上体重差的平均值与0无显著不同，意味着喝茶前和喝茶后的体重平均值无显著差异，可以认为该减肥茶无显著的减肥效果。