

题目	名称	Geographical Location(%)		总计	$\chi^2$	p
		East Asia	North America			
Race	Asian	761(100.00)	5(4.00)	766(86.46)		
	Black or African American	0(0.00)	5(4.00)	5(0.56)		
	Other	0(0.00)	33(26.40)	33(3.72)	845.008	0.000**
	White	0(0.00)	82(65.60)	82(9.26)		
	总计	761	125	886		

\* p<0.05 \*\* p<0.01

## 分析建议



卡方分析研究X(定类)与Y(定类)的差异，比如两种药物(旧药和新药)的治疗效果(治愈和未治愈)差异。

第一：分析X分别与Y之间是否呈现出显著性(p值小于0.05或0.01)；

第二：如果呈现出显著性；具体对比选择百分比(括号内值)，描述具体差异所在；

第三：对分析进行总结。

【特别提示：卡方分析共有Pearson卡方、连续校正Yates卡方和Fisher卡方等统计量；需要结合R\*C交叉类型，总样本数，以及期望频数情况选择对应的统计量；这一选择过程已经被SPSS AU自动化处理，详细选择过程可参阅帮助手册】

## 智能分析



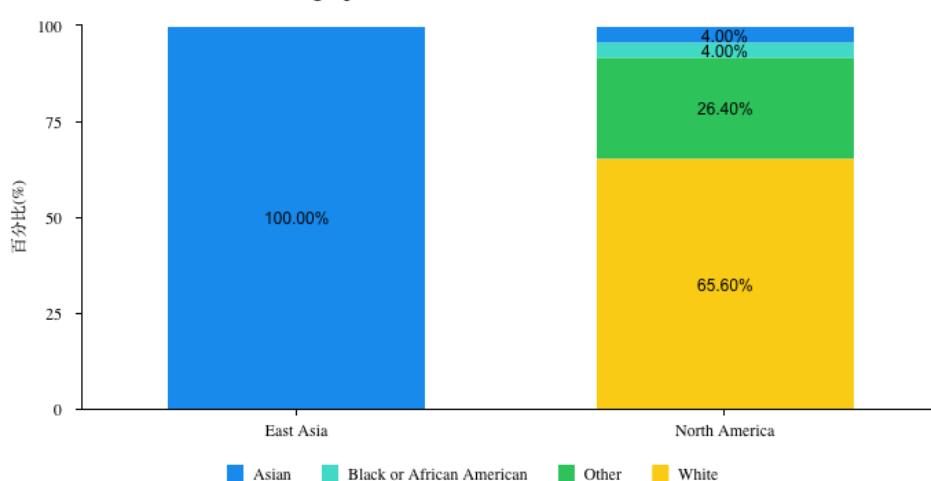
从上表可知，利用卡方检验(交叉分析)去研究Geographical Location对于Race共1项的差异关系【独立性】，从上表可以看出：不同Geographical Location样本对于Race共1项呈现出显著性(p<0.05)，意味着不同Geographical Location样本对于Race共1项均呈现出差异性。

总结可知：不同Geographical Location样本对于Race全部均呈现出显著性差异。

【特别提示：卡方分析共有Pearson卡方、连续校正Yates卡方和Fisher卡方等统计量；需要结合R\*C交叉类型，总样本数，以及期望频数情况选择对应的统计量；这一选择过程已经被SPSS AU自动化处理，详细选择过程可参阅帮助手册】

Race

Geographical Location和Race的交叉图



卡方检验统计量过程值

项	名称	值
Geographical Location * Race(2*4)	Pearson卡方	845.008( $p = 0.000^{**}$ )
	连续校正Yates卡方	845.008( $p = 0.000^{**}$ )
	Fisher卡方	-
	E>=5	5(62.50%)
	1<=E<5	2(25.00%)
	E<1	1(12.50%)
	Cnt	8
	n	886
	自由度df 值	3

## 分析建议

目

上表格展示信息包括：

统计量值：Pearson卡方、连续校正Yates卡方和Fisher卡方这三个统计量及对应p值；期望频数信息：小于1，介于1~5之间，或者大于5的格式个数及占比；卡方检验时，需要结合期望频数信息，R\*C交叉类型等选择最终的统计量和p值。这一选择过程已经被SPSSAU自动化处理，详细选择过程可参阅帮助手册【非专业用户建议忽略】。

### 深入分析-效应量指标

目

分析项	Phi	列联系数	校正列联系数	Cramer V	Lambda
Race	0.977	0.699	0.988	0.977	0.642

## 分析建议

目

如果卡方检验显示呈现出显著性差异( $p<0.05$ )，可通过百分比对比具体差异,同时还可使用效应量(Effect size)研究差异幅度情况；

第一：卡方检验时有多种指标(SPSSAU提供五类)可表示效应量，可结合数据类型及交叉表格类型综合选择；

第二：如果是2\*2表格，建议使用Phi指标；

第三：如果是3\*3,或4\*4表格，建议使用列联系数；

第四：如果是n\*n(n>4)表格，建议使用校正列联系数；

第五：如果是m\*n(m不等于n)表格，建议使用Cramer V指标；

第六：如果X或Y中有定序数据，建议使用Lambda指标；

第七：效应量值越大说明差异幅度越大，通常情况下效应量小、中、大的区分临界点分别是: 0.20,0.50和0.80，可使用建议查阅帮助手册。

### Geographical Location\*Race多重比较

目

比较	名称	East Asia	North America	$\chi^2$	p
第1次	Asian	761	5	382.984	0.000**
	Black or African American	0	5		
第2次	Asian	761	5	689.339	0.000**
	Other	0	33		
第3次	Asian	761	5	794.047	0.000**
	White	0	82		
第4次	Black or African American	0	5	null	null
	Other	0	33		
第5次	Black or African American	0	5	null	null
	White	0	82		
第6次	Other	0	33	null	null
	White	0	82		

^ p&lt;0.05 \*\* p&lt;0.01

比较次数：6

## 分析建议

目

卡方检验多重比较指两两类别组合（每个组合共4个数字）重复进行卡方检验；

第一：多重比较的次数=C(X类别个数)\*C(Y类别个数)，比如X类别为3，Y类别个数为5，则为C(3,2)\*C(5,2)=30次；

第二：SPSSAU仅针对第1个Y进行，可通过更换Y的位置实现其它分析项的多重比较，X或Y的选项个数大于10时不进行多重比较；

第三：多重比较时提供的是Pearson卡方检验；

第四：多重比较时，检验次数增多会增加一类错误的概率，建议使用校正显著性水平(Bonferroni校正)；

第五：如果显著性水平为0.05，并且两两比较次数为3次，那么Bonferroni校正显著性水平为 $0.05/3=0.0167$ ，即p值需要与0.0167进行对比，而不是0.05；

【特别提示：多重比较通常在医学研究中使用，且类别数量较少时使用】

### Cochran-Armitage 趋势卡方检验

目

项	z值	$\chi^2$ 值	df	p值
Geographical Location * Race(2*4)	28.174	793.752	1	0.000

## 分析建议

目

Cochran-Armitage 趋势卡方检验(Linear-by-Linear)研究k\*2(或2\*k)列联表数据的趋势差异关系；

第一：Cochran-Armitage 趋势卡方检验通常用于k\*2(或2\*k)的列联表结构,k为有序定类数据,2指两个类别；

第二：如果p值小于0.05则说明k组间呈现出某种趋势变化；

第三：如果p值大于0.05说明没有呈现出某种趋势变化；

第三：如果p值大于0.05则说明k组间不会呈现出趋势变化；

第四：SPSSAU默认均提供趋势卡方检验(Linear-by-Linear)结果，但需要注意使用此指标的前提条件。

## 参考文献



- 【1】The SPSSAU project (2023). SPSSAU. (Version 23.0) [Online Application Software]. Retrieved from <https://www.spssau.com>.
- 【2】Clegg F. Introduction to statistics. III: Correlation, chi-square and the choice of statistical procedure.[J]. Br J Hosp Med, 1988, 40(5).
- 【3】颜虹,徐勇. 医学统计学.第3版[M].人民卫生出版社,2017.