

SSPSSAU 算法手册

一基本说明	1
二数据处理(生成变量).....	2
三分析算法	2
1. 频数分析	2
2. 交叉(卡方)分析.....	3
3. 描述分析	3
4. 分类汇总	3
5. 相关分析	3
6. 回归分析	3
7. 方差分析	4
8. T 检验.....	4
9. 单样本 T 检验.....	4
10. 配对 T 检验.....	4
11. 正态性检验	4
12. 非参数检验	4
13. 信度分析	5
14. 效度分析	5
15. 多选题	5
16. 单选-多选	5
17. 10 多选-单选	5
18. 多选-多选	5
19. 聚类分析	5
20. 因子分析	6
21. 多重响应(单-多).....	6
22. 事后检验	6
23. 逐步回归分析.....	6
24. 分层回归分析.....	6
25. 双因素方差分析.....	6
26. 二元 Logit 分析.....	7

一基本说明

● 关于 P 值*标示

SPSSAU 对于 P 值的标注规则统一为: $P < 0.01$ 使用两个*号表示, $P < 0.05$ 则使用 1 个*号表示。并且在表格下方有备注。除非特别说明,默认使用双侧检验。

● SPSSAU 智能文字分析算法

SPSSAU 智能文字算法时,统一按照 $P < 0.05$ 则显著这一标准进行。

● 关于页面展示中的 X 和 Y 说明

分析方法研究中通常称自变量，因变量等，为便于描述和用户理解，SPSSAU 通常称 X 或者 Y 等，比如研究不同性别人群身高差异，则描述为“X（性别）和 Y（身高）的差异关系”。

● 关于数据类型说明

SPSSAU 中仅定义两类数据，分别是定类和定量数据；通常统计知识中分为定类，定序和定量，SPSSAU 将定序和定量这两类数据合称定量数据。便于用户进行研究理解和使用。

● 关于筛选功能

筛选是针对记录进行，比如用户仅希望分析男性样本等；系统默认会保留用户筛先操作，如果希望去除筛选，点击“放弃筛选”按钮。

筛选条件仅最多为 3 个，提供“等于”，“大于”，“小于”，“不等于”共四类运算符，以及筛选条件之间的逻辑可选为“并且”和“或者”。

二数据处理(生成变量)

SPSSAU 当前提供的数据处理方法共 10 种，分别为：平均值，求和，乘积，标准化，中心化，平方，根号，自然对数，10 为底对数，绝对值。SPSSAU 后续会进一步加入更多数据处理方法。

三分析算法

SPSSAU 当前共提供 30 类算法（含算法拆分），后续版本会陆续多重响应深入功能，主成分分析，事后检验 SNK(q 检验)和 Dunnett(t 检验)等，卡方检验 Fisher 检验量，散点图，回归残差检验图，正态检验图，高级算法（比如一般线性模型，多分类 Logit），熵值法，医学分析算法（比如生存分析，Cox 回归，ROC 曲线），以及其余常见统计算法等。

1. 频数分析

计算各个数字的**有效频数**和**有效百分比**；如果有缺失数据，SPSSAU 并不会单独标识出缺失个数，处理方法是‘标题’之后加入总样本数量，类似下图：总共样本为 99 个，学历有 1 个缺失数据，因此学历的有效样本数量为 98。频数和百分比均基于有效样本 98 进行计算，并且在学历后面加入（N=98）进行标识。

名称	选项	频数 ●	百分比(%) ●
学历(N=98)	本科以下	1	1.0
	本科	94	95.9
	硕士及以上	3	3.1
合计		99	100.0

频数分析默认输出饼图，圆环图，柱形图和条形图。

2. 交叉(卡方)分析

交叉分析时使用的 X^2 值为 Pearson X^2 值。（后续会提供 Fisher 确切概率检验等）

默认输出柱形图，条形图，堆积柱形图和堆积条形图。

3. 描述分析

输出有效样本的最小值，最大值，平均值，标准差和中位数。

默认输出条形图和雷达图。

SPSSAU 后续会提供四分位数点，峰度，偏度等指标。

4. 分类汇总

分类汇总是频数，交叉，描述，方差分析的集合，根据具体情况对应得到不同的结果；方差分析和卡方分析时，不提供统计量和 P 值。

X 的放置情况	汇总类型	对应方法
X（不放置）	平均值(默认)	描述分析
X（不放置）	百分比	频数分析
X（放置）	平均值(默认)	方差分析
X（放置）	百分比	交叉(卡方)分析

5. 相关分析

相关分析共提供两类相关系数，分别是 Pearson 相关系数和 Spearman 相关系数。以及默认提供两种规范化输出格式。默认不输出 P 值，使用*号进行表示。

6. 回归分析

回归分析算法使用通用的最小二乘法（OLS）进行指标估计，默认会输出 VIF 值。回归分析提供残差或预测值保存可选功能。

7. 方差分析

方差分析输出：平均值，标准差，F 值和 P 值。

方差齐检验使用通用的 Levene 统计量。

默认输出折线图。

8. T 检验

T 检验时中间使用的方差齐检验为 Levene 统计量；如果方差齐则使用混合方差，方差不齐则使用成组 T 检验；方差齐检验，方差齐或者方差不齐的中间检验过程，SPSSAU 已经默认进行判断处理，直接输出最终的检验统计量和 P 值。

默认输出折线图。

9. 单样本 T 检验

输出有效样本的最小值，最大值，平均值，标准差，中位数，t 检验统计量值和 P 值，提供对比数字设置或选功能（默认为 0）。

10. 配对 T 检验

输出有效样本的平均值，标准差，配对数据平均值差值，t 检验统计量值和 P 值。

11. 正态性检验

正态性检验使用 Kolmogorov-Smirnov 检验和 Shapro-Wilk 检验两种检验统计量，SPSSAU 建议样本量小于等于 50 时使用 Shapro-Wilk 检验，样本量大于 50 时使用 Kolmogorov-Smirnov 检验。

智能化文字默认以此为标准进行智能分析。

12. 非参数检验

非参数检验共提供 mannWhitney 检验统计量和 Kruskal-Wallis 检验统计量；如果数据组别为两组则使用 mannWhitney 检验，数据组合超过两组则使用 Kruskal-Wallis 检验统计量；SPSSAU 默认进行判断并且输出，不需要用户进行选择。

SPSSAU 智能化文字算法也默认进行中间过程判断，勿需用户干预判断。

13. 信度分析

信度分析提供校正项总计相关性(CITC),项删除后 α 系数, α 系数共三个指标。

14. 效度分析

效度分析即为因子分析,主成分法,以及方差旋转法为“最大方差旋转”,最大迭代次数为 1000,以及因子个数设置可选功能(默认以特征根大于 1 作为因子个数输出,可设置因子个数区间为 1~15)。

15. 多选题

多选题分析针对问卷研究,提供响应率和普及率等指标。

16. 单选-多选

1 个单选题和 1 个多选题进行交叉分析,提供交叉频数和百分比等指标。

17. 10 多选-单选

1 个多选题和 1 个单选题进行交叉分析,提供交叉频数和百分比等指标。

18. 多选-多选

1 个多选题和 1 个多选题进行交叉分析,提供交叉频数和百分比等指标。

19. 聚类分析

聚类分析使用 K 均值(K-means)算法,并且聚类分析后默认输出不同类别数据与分析项的方差分析结果,并提供聚类个数设置可选功能(默认为 3,可设置聚类个数为 2, 3, 4, 5, 6)。

默认输出饼图,圆环图,柱形图和条形图,折线图等。

SPSSAU 后续会提供分层聚类算法。

20. 因子分析

因子分析使用主成分法，以及方差旋转法为“最大方差旋转”，最大迭代次数为 1000。生成结果提供 KMO 和 Bartlett 球形度检验，方差解释表格，旋转后因子得分系数表格，成份得分系数表格和碎石图。因子分析提供因子得分保存可选功能，以及因子个数设置可选功能（默认以特征根大于 1 作为因子个数输出，可设置因子个数区间为 1~15）。

SPSSAU 后续会单独提供主成分分析。

21. 多重响应(单-多)

多重响应共输出响应率和普及率等指标，提供计数项设置可选功能（默认为 1）。

默认输出饼图，圆环图，柱形图和条形图，折线图。

SPSSAU 后续会提供多重响应(多-多)分析，多重响应（多-单）分析。

22. 事后检验

事后检验默认使用 LSD 方法。后续会加入 SNK(q 检验)和 Dunnett(t 检验)等事后检验方法。默认智能化输出事后检验对比规范化结果，不单独输出 P 值。

23. 逐步回归分析

回归分析算法使用通用的最小二乘法（OLS）进行指标估计。如果有剔除掉某个分析项，默认不输出，在智能化文字分析中会进行标注。

24. 分层回归分析

分层回归会输出 R 方变化值，F 变化值。默认不输出 P 值，使用*号进行表示；提供 Sobel 检验网页（quantpsy.org）。

25. 双因素方差分析

双因素方差分析使用类型 III 算法。

26. 二元 Logit 分析

二元 Logit 分析使用极大似然法估计，使用 Z 检验，以及默认会输出 OR 值，以及回归系数的 95% CI，以及伪 R 方值。二元 Logit 提供残差或预测值保存可选功能。

www.spssau.com