

Turf 组合模型算法

目录

SPSSAU 数据格式.....	2
计算公式.....	2
1. 基本说明.....	2
2. 组合个数 (Combination) .....	2
3. 组合数量.....	3
4. 触达数 (Reach count) .....	3
5. 触达率 (Reach ratio) .....	3
6. 总触达次数 (Total reach count) .....	3
7. 频数 (Frequency) .....	3
8. 频率 (Reach ratio) .....	3
9. 自动寻优.....	3
参考文献.....	4

TURF (Total Unduplicated Reach and Frequency) 模型用于评估不同产品组合对目标市场的覆盖率和频率，在 SPSSAU 中支持：

- ✓ 支持自动寻优组合数量。



将分析项拖拽至右侧框然后‘开始分析’即可。SPSSAU 中涉及 1 项参数。

- ✓ 组合数量：默认为自动寻找最优组合，研究者也可以自己下拉选择组合数量。

## SPSSAU 数据格式

multiple choice_option1	multiple choice_option2	multiple choice_option3
1	1	1
1	1	0
0	0	0
1	1	1
1	1	0
0	0	0
0	1	1
1	1	0
1	0	1
1	1	1
1	1	0
1	1	1
0	0	0

Turf 模型时，其通常采用多选题进行收集数据，其数据格式与多选题的数字格式完全一致，数字 1 表示选中，数字 0 表示没有选中。

## 计算公式

## 1. 基本说明

Turf 组合模型时核心目的是计算出触达率，当然还有其它指标比如频率，首先设定数据矩阵为：

$$X = \begin{bmatrix} X_{11} & X_{12} & \cdots & X_{1m} \\ X_{21} & X_{22} & \cdots & X_{2m} \\ \vdots & \vdots & \ddots & \vdots \\ X_{n1} & X_{n2} & \cdots & X_{nm} \end{bmatrix}$$

其中：

$X_{ij}$ ：第  $i$  个样本对第  $j$  个产品的偏好（例如，1 表示喜欢，0 表示不喜欢）

$n$ ：样本的数量

$m$ ：产品（选项或者列）的数量

## 2. 组合个数 (Combination)

组合个数的计算公式为：

$$C(m, k) = \frac{m!}{k!(m-k)!}$$

其中：

$n$  是总产品（选项或者列）数量

$k$  是组合数量的个数，比如‘选项 1 和选项 2’时组合数量为 2

### 3. 组合数量

组合数量 $k$ 即为某组合时的数量，比如‘选项1和选项2’时组合数量为2。该项为下拉参数数字，如果为‘自动寻优’，则由SPSSAU自动决定。

### 4. 触达数 (Reach count)

触达数是指至少选择一项的样本数量，如下：

$$\text{Reach count} = \sum_{i=1}^n \max(X_{i1}, X_{i2}, \dots, X_{im})$$

### 5. 触达率 (Reach ratio)

触达率是指至少选择一项的样本比例，如下：

$$\text{Reach ratio} = \frac{1}{n} \sum_{i=1}^n \max(X_{i1}, X_{i2}, \dots, X_{im})$$

### 6. 总触达次数 (Total reach count)

总触达次数指全部样本选中数字1的个数：

$$\text{Total reach count} = \sum_{i=1}^n \sum_{j=1}^m (X_{ij})$$

### 7. 频数 (Frequency)

频数指某组合前提下，数字1的个数，如下：

$$\text{Frequency} = \sum_{i=1}^n \sum_{j=1}^m (X_{ij})$$

### 8. 频率 (Reach ratio)

频率指某组合前提下，频数除以总触次数，如下：

$$\text{Reach} = \frac{\text{Frequency}}{\text{Total reach count}}$$

### 9. 自动寻优

假设有 $m$ 项进行分析，表示为 $X_1, X_2, \dots, X_m$ 。定义组合数量为 $k$ ，其中 $k$ 的取值范围为1到 $m$ 。

对于 $k = 1$ ，选择最优项 $X_{opt}(1)$ ：

$$X_{opt}(1) = \operatorname{argmax}_{X_i} f(X_i)$$

其中， $f(X_i)$ 表示选择项 $X_i$ 时的触达率。

对于 $k = 2$ ：在已选定的项 $X_{opt}(1)$ 的基础上，选择与其他项组合的最优组合 $X_{opt}(2)$ ：

$$X_{opt}(2) = \operatorname{arg} \max_{X_{opt}(1), X_j} f(X_{opt}(1), X_j)$$

其中， $j$ 是在剩余项中的索引。

对于 $k = 3$ ：在已选定的组合基础上，选择与其他项组合的最优组合 $X_{opt}(3)$ ：

$$X_{opt}(3) = \arg \max_{X_{opt}(2), X_k} f(X_{opt}(2), X_k)$$

其中,  $k$ 是在剩余项中的索引。

对于 $k > 3$ , 继续循环:

$$X_{opt}(k) = \arg \max_{X_{opt}(k-1), X_j} f(X_{opt}(k-1), X_j)$$

其中,  $X_{opt}(k-1)$ 为前一个步骤选出的最优组合。

### 参考文献

【1】 The SPSSAU project (2024). SPSSAU. (Version 24.0) [Online Application Software]. Retrieved from <https://www.spssau.com>.

【2】 周俊,马世澎. SPSSAU 科研数据分析方法与应用.第 1 版[M]. 电子工业出版社,2024.