内容分析编码员信度计算

刘念夏

2026-05-22

目录

1 数据预备 2
1.1 创建示例数据 (假设有 3 个类别: 0, 1, 2) 2
1.2 将数据合并为数据框 (data.frame 格式) 2
1.3 查看数据结构 2
2 计算 Holsti's Alpha(2 位编码员时使用)(一般使用) 3
2.1 Holsti's Alpha 公式 3
2.2 手工计算 Holsti's Alpha 3
2.2.a 呈现交叉表 3
2.2.b 手工计算 3
3 计算 Cohen's Kappa (2 位编码员时使用)(学术常用) 3
3.1 使用 irr 包中的 kappa2( ) 函数，计算 Cohen's Kappa。 3
3.2 人工计算 Cohen's Kappa。 4
3.2.a Kappa 公式 4
3.2.b 计算 P_o 和 P_e 4
3.3 系数解读 4
4 信度检验结果描述 5
5 计算 Krippendorff's Alpha (适用于多位编码员) 5
5.1 Krippendorff's Alpha 公式 5
5.2 载入套件 5
5.3 建立包含三位编码员的数据 5
5.4 合并并转置 (行必须是编码员，列是样本) 5
5.5 计算定类尺度的 Alpha 6
5.6 查看结果 6
5.7 系数判准 6
5.8 讨论与不一致性处理: 7
参考文献 7

```r
# 设置系統中文文字编码(简体中文)
Sys.setlocale(category = "LC_ALL", locale = "zh_CN.UTF-8")
```

```
[1] "zh_CN.UTF-8/zh_CN.UTF-8/zh_CN.UTF-8/C/zh_CN.UTF-8/zh_CN.UTF-8"
```

```r
# 使绘图物件中的中文文字能正确呈现
# 调用加载"showtext"套件包
library(showtext)
```

```
载入需要的程序包：sysfonts
```

```
载入需要的程序包：showtextdb
```

```r
showtext_auto(enable = TRUE)
```

1 数据预备

1.1 创建示例数据 (假设有 3 个类别: 0, 1, 2)

编码员 A 和 B 对 10 个样本(例如: 10 个视频)的判读结果

```r
coder_A <- c(1, 0, 2, 1, 1, 0, 2, 1, 0, 1)
coder_B <- c(1, 0, 1, 1, 1, 0, 2, 0, 0, 1)
```

1.2 将数据合并为数据框 (data.frame 格式)

```r
data <- data.frame(coder_A, coder_B)
```

1.3 查看数据结构

```r
print(data)
```

```
   coder_A coder_B
1        1       1
2        0       0
3        2       1
4        1       1
5        1       1
6        0       0
```

```
7        2        2
8        1        0
9        0        0
10       1        1
```

2 计算 Holsti's Alpha(2 位编码员时使用)(一般使用)

2.1 Holsti's Alpha 公式

Holsti's Alpha 也称为$C.R.$ (信度系数,Coefficient of Reliability)，计算公式如下:

$$Holsti's Alpha = C.R. = \frac{2M}{n_1 + n_2}$$

其中:

$M =$ 两位编码员编码一致的样本数

$n_1 =$ 编码员$n1$处理的样本数

$n_2 =$ 编码员$n2$处理的样本数

2.2 手工计算 Holsti's Alpha

2.2.a 呈现交叉表

```
#install.packages("sjPlot")
#library(sjPlot)
#tab_xtab(data$coder_A, data$coder_B, show.summary = FALSE)
```

2.2.b 手工计算

2 位编码员编码一致的数目：3+4+1=8 总编码数目：10

Holsti's Alpha = (2 * 8) / (10+10) = 0.8

Holsti's Alpha 系数虚高，比较水。学术圈比较不常用，较常使用 Cohen's Kappa。

3 计算 Cohen's Kappa (2 位编码员时使用)(学术常用)

Kappa 系数排除了随机一致的概率，所以会较 Holsti 低，比较严谨，学术常用。

3.1 使用 irr 包中的 kappa2( ) 函数，计算 Cohen's Kappa。

```
#install.packages("irr")
library(irr)
```

```
载入需要的程序包：lpSolve
```

```
result <- kappa2(data)
print(result) #Kappa=0.672
```

```
 Cohen's Kappa for 2 Raters (Weights: unweighted)

 Subjects = 10
   Raters = 2
    Kappa = 0.672

        z = 2.88
  p-value = 0.00393
```

3.2 人工计算 Cohen's Kappa。

3.2.a Kappa 公式

$$\kappa = \frac{P_o - P_e}{1 - P_e}$$

$$\text{其中: } P_o = \text{观察一致性}(\text{observed agreement})$$
$$P_e = \text{随机一致性}(\text{expected agreement})$$

3.2.b 计算 P_o 和 P_e

- P_o: 两位编码员达成一致的比例 = 8/10 = 0.8
- P_e: 随机一致的概率 = (0.3 * 0.4) + (0.5 * 0.5) + (0.2 * 0.1) = 0.39
- 计算 Kappa:

$$\kappa = \frac{0.8 - 0.39}{1 - 0.39} = \frac{0.41}{0.61} \approx 0.672$$

3.3 系数解读

Cohen's Kappa 系数的核心逻辑是: 在排除掉"随机乱猜而达成一致"的可能性后，两位编码员实际达成一致的比例是多少。

Kappa (k): 系数< 0: 几乎没有一致性。

0.00–0.20: 极低的一致性 (Slight)。

0.21–0.40: 一般 (Fair)。

0.41–0.60: 中等 (Moderate)。

0.61–0.80: 高度一致 (Substantial)。

0.81–1.00: 几乎完美一致 (Almost Perfect)。

p-value: 如果 $p < 0.05$， 说明这种一致性不是随机发生的，具有统计学显著意义。

4 信度检验结果描述

本研究采用内容分析法对 10 个样本进行编码。为确保编码的一致性，由两位独立编码员（A 与 B）根据预先制定的编码手册进行判读。初步信度检验结果显示，观察一致率（Holsti 系数）为 0.80。考虑到随机一致性的影响，进一步计算 Cohen's Kappa 系数，所得结果为 $\kappa = 0.67$。根据 Landis & Koch (1977) 的判定标准，显示两位编码员之间具有高度一致 **(Substantial)**）的一致性[1]。

针对编码不一致的样本，研究团队随后进行了深入讨论，修正了编码手册中关于"类别 1"与"类别 0"界定模糊的部分，并达成共识，以确保后续大规模编码的严谨性。

5 计算 Krippendorff's Alpha (适用于多位编码员)

5.1 Krippendorff's Alpha 公式

$$\alpha = 1 - \frac{D_o}{D_e}$$

$$\text{其中: } D_o = \text{观察到的不一致量 (Observed Disagreement)}$$

$$D_e = \text{随机预期下的不一致量 (Expected Disagreement)}$$

5.2 载入套件

```
#install.packages("irr")
library(irr)
```

5.3 建立包含三位编码员的数据

```
coder_A <- c(1, 0, 2, 1, 1, 0, 2, 1, 0, 1)
coder_B <- c(1, 0, 1, 1, 1, 0, 2, 0, 0, 1)
coder_C <- c(1, 0, 2, 1, 0, 0, 2, 1, 0, 1)
```

5.4 合并并转置 (行必须是编码员，列是样本)

```
data_Kalpha <- data.frame(coder_A, coder_B, coder_C)
print(data_Kalpha)
```

[1] 0.61–0.80: 高度一致 (Substantial)。

```
   coder_A coder_B coder_C
1        1       1       1
2        0       0       0
3        2       1       2
4        1       1       1
5        1       1       0
6        0       0       0
7        2       2       2
8        1       0       1
9        0       0       0
10       1       1       1
```

```
matrix_Kalpha <- t(as.matrix(data_Kalpha))
print(matrix_Kalpha)
```

```
        [,1] [,2] [,3] [,4] [,5] [,6] [,7] [,8] [,9] [,10]
coder_A    1    0    2    1    1    0    2    1    0     1
coder_B    1    0    1    1    1    0    2    0    0     1
coder_C    1    0    2    1    0    0    2    1    0     1
```

5.5 计算定类尺度的 Alpha

```
result_Kalpha <- kripp.alpha(matrix_Kalpha, method = "nominal")
```

5.6 查看结果

```
print(result_Kalpha) #0.683
```

```
 Krippendorff's alpha

 Subjects = 10
   Raters = 3
    alpha = 0.683
```

5.7 系数判准

$\alpha > 0.800$: 结论非常可靠（高度一致）。

$0.667 < \alpha < 0.800$: 可接受的信度，足以得出初步结论（中高度一致）。

$\alpha < 0.667$: 信度堪忧，通常需要回过头检查编码员 A/B/C 为什么在样本 3、5、8 产生了分歧。

5.8 讨论与不一致性处理:

本研究在初步编码测试中获得 Krippendorff's Alpha 系数为 0.683 ($n = 10, k = 3$)。

尽管该系数已达到社会科学研究的可接受标准（$\alpha > 0.667$），但研究团队仍针对产生分歧的样本（样本 3、5、8）进行了深度协商。

分歧点分析：样本 3 中，编码员 A 与 C 判定为类别 2，而 B 判定为类别 1；样本 5 与 8 则呈现出三方各异的判断。经讨论发现，分歧主要源于编码手册中关于"语义模糊性"的界定不够明确。

处理结果：研究团队随后修订了编码手册（Codebook），增加了具体案例展示以统一判读标准。经协商一致后，最终将争议样本的属性进行了定论(样本 3 定论为 2; 样本 5 定论为 1;样本 8 定论为 1;)。此过程确保了后续大规模编码任务的严谨性与严密性。

参考文献

Landis, J. R., & Koch, G. G. (1977). The measurement of observer agreement for categorical data. *Biometrics*, *33*(1), 159 ~ 174. https://doi.org/https://doi.org/10.2307/2529310