

# AL-049 Evaluating Questionnaire Data Reliability with SAS

Yida Bao

University of Wisconsin, Stout,  
Assistant Professor,  
SAS Advanced Programmer

Ph.D. Math Auburn University  
MS Data Science Auburn University



# About Questionnaire Data



## What is Questionnaire Data?

- **Structure:** Questionnaire data is gathered through a set of predefined questions.
- **Questions:** Typically include multiple-choice, open-ended, or rating scale questions.
- **Responses:** Each question has a set of possible answers.
- **Scope:** A questionnaire can contain anywhere from 5 to 50+ questions.

# Data Analysis Methods For Survey Data

- Traditional Statistic Method may work.
- Wisely use the Data visualization about survey data.
- **Data reliability** is a critical step before analyzing survey data.



# Data reliability

## 1. What is Data Reliability?

- A measure of consistency among items in a questionnaire.
- Indicates whether all items in a scale reliably measure the same underlying construct.

- “How often do you smile?”
- “Do you feel happy when spending time with family?”
- “Do you enjoy your work?”

Sample	Item 1	Item 2	Item 3	Item 4	Item 5
1	5	3	4	1	3
2	1	4	0	5	1
3	3	2	5	3	1
...	...	...	...	...	...

#MWSUG2024 #SSxx ...



# Data reliability

## 2. Why Use Data Reliability?

- Ensures the stability and accuracy of survey results.
- Confirms all items contribute meaningfully to the measurement of the construct.

Sample	Item 1	Item 2	Item 3	Item 4	Item 5
1	5	3	4	1	3
2	1	4	0	5	1
3	3	2	5	3	1
...	...	...	...	...	...



# Inner reliability

## 3. Different Type of inner reliability Measures

- Cronbach Alpha
- McDonald's Omega
- Spilt-Half Reliability

Sample	Item 1	Item 2	Item 3	Item 4	Item 5
1	5	3	4	1	3
2	1	4	0	5	1
3	3	2	5	3	1
...	...	...	...	...	...



# Our Paper will cover...

- Introduce the Solution Approach for This Specific type of Dataset.
- Explain the Mathematical Expressions and Application Scenarios of Three Approaches.
- Offer Ready-to-Use Code for SAS Application



# Cronbach's Alpha

**Definition:** Attributed to Lee Cronbach, it is the most widely used measure of reliability, assessing how closely related the items in a scale are.

$$\alpha = \frac{N \cdot \bar{c}}{\bar{v} + (N - 1) \cdot \bar{c}}$$

where  $N$  is the number of items in the scale,  $\bar{c}$  is the average inter-item covariance, and  $\bar{v}$  the average variance of the items.





# Cronbach's Alpha

Let us take a numerical example with 4 items and 60 respondents, using SAS.

```
data Test1;
  input Q1-Q4 @@;
  cards;
4 3 4 4 4 3 3 3 5 5 5 4 3 4 4 5 3 4 3 3 3 3 2 3 2 3 2 2
4 5 4 5 4 5 3 5 5 3 3 5 3 2 1 4 2 3 4 3 3 2 2 2 3 3 4 2
4 4 4 3 4 4 1 4 2 3 3 2 2 1 1 1 3 2 3 3 3 4 3 3 3 2 3 3
5 5 4 5 1 4 3 4 3 3 3 3 3 4 3 4 1 3 3 4 2 2 2 1 3 4 4 4
4 5 5 5 5 4 4 5 3 3 5 4 3 3 2 1 2 3 3 2 3 3 2 2 3 4 4 5
4 3 3 4 4 2 2 3 4 3 4 4 3 1 1 3 3 3 3 3 2 3 3 4 2 3 3 3
2 4 3 4 2 3 3 2 3 2 4 4 3 3 3 2 5 4 4 4 3 4 2 3 5 4 5 4
4 4 3 4 3 4 3 2 2 2 1 2 3 3 2 2 5 5 4 4 3 3 3 3 3 4 3 4
5 4 4 4 3 4 4 4 1 2 1 2 1 1 1 1 1
;
run;

proc corr NoSimple alpha data=test1;
  var q1-q4;
  ods select CronbachAlpha CronbachAlphaDel;
run;
```

Example adapted from UCLA –  
Advanced Research Computing

The CORR Procedure

Cronbach Coefficient Alpha	
Variables	Alpha
Raw	0.839005
Standardized	0.839956



# Cronbach Alpha

- $\alpha < 0.2$  Unacceptable: Items do not measure the same construct.
- $0.2 \leq \alpha < 0.5$  Poor: Low consistency; items may need revision.
- $0.5 \leq \alpha < 0.8$  Good: Strong internal consistency;
- $\alpha \geq 0.8$  Excellent: Very high reliability.



# Cronbach's Alpha

Manual calculation of the Cronbach's alpha for this example:

$$\bar{v} = (1.168 + 1.012 + 1.169 + 1.291)/4 = 4.64/4 = 1.16.$$

$$\bar{c} = (0.557 + 0.574 + 0.690 + 0.673 + 0.720 + 0.724)/6 = 3.938/6 = 0.656.$$

$$\alpha = \frac{4(0.656)}{(1.16) + (4 - 1)(0.656)} = 2.624/3.128 = 0.839.$$



# Split-Half Reliability

**Split-Half Reliability** involves splitting the questionnaire into two halves and correlating the scores from each half. This method provides a measure of consistency but requires correction using the Spearman-Brown formula:

$$r_{SB} = \frac{2r}{1 + r}$$

Where  $r$  is the correlation between the two halves.



# McDonald's Omega

**Definition:** A more precise measure of internal consistency, especially useful for latent variable models-heading.

$$\omega = \frac{\sum_{i=1}^k \lambda_i^2}{\sum_{i=1}^k \lambda_i^2 + \sum_{i=1}^k \delta_i^2}$$

Where  $\lambda_i$  are the factor loadings and  $\delta_i$  are the error variances.



# McDonald's Omega

1. Exploratory factor Analysis(EFA)
2. Compute the Error Variance.
3. Calculate the Total Factor Loading.
4. Calculate McDonald's Omega.

$$\omega = \frac{\sum_{i=1}^k \lambda_i^2}{\sum_{i=1}^k \lambda_i^2 + \sum_{i=1}^k \delta_i^2}$$



# Different Case , Different Discussion

Method	Key Feature	Advantages	Limitation
<b>Split-Half Reliability</b>	Correlation between two test halves	Simple computation	Results vary by split; not comprehensive
<b>Cronbach's Alpha</b>	Average inter-item correlation	widely used	Assumes equal item impact;
<b>McDonald's Omega</b>	Considers item factor loadings	Accurate for complex scales	Requires factor analysis; more complex



# Other tips

- a. Remove Items with the Lowest Item-Total Correlation.

Check the “Cronbach’s Alpha if Item Deleted” table from your statistical software.

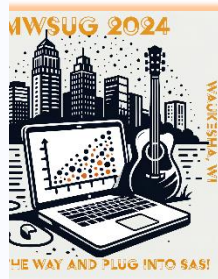
Cronbach Coefficient Alpha	
Variables	Alpha
Raw	0.822134
Standardized	0.985145

Cronbach Coefficient Alpha with Deleted Variable				
Deleted Variable	Raw Variables		Standardized Variables	
	Correlation with Total	Alpha	Correlation with Total	Alpha
Weight3	0.975379	0.783365	0.973464	0.977103
Length3	0.967602	0.881987	0.967177	0.978783
Height	0.964715	0.655098	0.968079	0.978542
Width	0.934635	0.824069	0.937599	0.986626

## redundancy

1. Do you feel stressed about exams?
2. Are exams making you anxious?
3. Are you nervous about your exam performance?
4. Do you experience stress when preparing for exams?

No new information  
Unnecessary length





# Other tips

## a. Remove Items with the Lowest Item-Total Correlation.

Check the "Cronbach's Alpha if Item Deleted" table from your statistical software.

## b. Reverse Scoring Items

1. I am satisfied with my work environment." (Positive)
2. I am satisfied with my pay." (Positive)
3. I dislike my coworkers." (Negative)

```
data d1;
  input ID Q1 Q2 Q3;
  cards;
1 4 2 4
2 3 3 3
3 4 1 5
4 2 5 1;
run;
data d2;
  set d1;
  Q2_alt=-1*Q2+6;
run;
proc corr alpha data=d2;
  var q1 q2_alt q3;
run;
```



# Other tips- Use EFA or PCA to Explore Groupings

Scale name and items	loading	Scale name and items	loading
<b>Satiety responsiveness/Fussiness</b> (factor1:22.47% variance)		<b>Emotional undereating/Food responsiveness</b>	
My child has a big appetite	0.73	My child eats more when s/he is happy	0.58
My child enjoy a wide variety of foods	0.71	My child eats more when nothing to do	0.62
My child looks for forward to mealtimes	0.67	My child eats more when watching TV	0.60
My child is interested in food	0.67	My child eats more when annoyed	0.42
My child enjoys tasting new food	0.67	My child eats more when worried	0.56
My child leaves food on his plate	0.78	My child's always asking for food	0.54
My child is difficult to please with meals	0.77	My child would eat most of time	0.51
My child decides without tasting it	0.51	If allowed to eat, my child would eat too much	0.39
My child can't each a meal without snack	0.67		
My child get null before finished	0.87		
My Child refuses new food at first	0.72		
<b>Desire to Drink</b>		<b>Enjoyment of food</b>	
If given the chance, my child would always be having a sugary drink	0.28	My child enjoys and loves eating food	0.43
If given the chance, my child would have sugary drink continuously thougout the day	0.11	My child looks forward to mealtimes	0.41
My child is always asking for a sugary	0.40	My child is interested in food	0.4

Ayine, Priscilla, et al. "Eating behaviors in relation to child weight status and maternal education." *Children* 8.1 (2021): 32.



# Summary & Future Plan

## Summary:

This presentation introduced three methods for assessing data reliability:

- Cronbach's Alpha
- McDonald's Omega
- Spilt Half Reliability

## Future Plan:

Implement Bootstrap Methods to improve the robustness of reliability estimates.



# Thank You (or) Contact Information

- Name **Yida Bao**
- Company **University of Wisconsin, Stout**
- Email **baoy@uwstout.edu**
- Paper Title **Evaluating Questionnaire Data Reliability With SAS**
- Paper Link **[https://papers.ssrn.com/sol3/papers.cfm?abstract\\_id=5121915](https://papers.ssrn.com/sol3/papers.cfm?abstract_id=5121915)Please**

- Thank you

