EPSY 5221 Classical Test Theory Assignment

The Classical Test Theory analysis assignment allows you to explore item analysis using dichotomous data from a multiple-choice test called Exam1 and polytomous rating scale data from an inventory intended to measure Liking Science. We will use these two data sets to conduct IRT analyses later – providing for comparisons between CTT and IRT. This assignment is worth 30 points.

Report the following for each data set (Exam 1 and Liking Science):

1. Describe the distribution of the Total Score:
   1. Report the Mean, SD, Minimum & Maximum scores.
   2. Provide the histograms of total scores and describe the shapes of the distributions.
2. Item difficulty:
   1. For both tests, what is the range of the item difficulties?
   2. For Liking Science, what might we call this value instead of “difficulty”?
   3. Are there limitations in conducting these analyses? How might you improve the analyses to address the limitations?
3. Item discrimination:
   1. For both tests, what is the range of the item discrimination values?
   2. Provide an interpretation of item discrimination for each test. What statistic is being estimated for item discrimination (what is the statistic you reported in 3.a.) and what does it mean?
   3. Are there limitations in conducting these analyses? How might you improve the analyses to address the limitations?
4. Score reliability:
   1. Report the score reliability coefficients. We will return to these in a couple of weeks.
5. Based on the evidence from these analyses, consider possible recommendations:
   1. What suggestions do you offer for improving the Exam 1 scores?
   2. What suggestions do you offer for improving the psychometric quality of the Liking Science scores? If you recommend removing items, does it make a difference in the results reported above?

Multiple-Choice Test – Item Analysis

The data for Exam 1 include 6000 test takers responding to 56 MC items. For our purposes here, we are not concerned with the subject matter or specific item content of these data. We are only interested in the results of the item and test score analysis.

The Key for Exam 1 is as follows (Format = Response Option: Item numbers)

-------------------------------------------------------

A: 1, 9, 12, 15, 19, 21, 22, 28, 29, 30, 34, 38, 42, 52, 55

B: 4, 6, 16, 18, 24, 26, 32, 33, 35, 43, 44, 47, 50, 54

C: 3, 5, 7, 11, 14, 20, 23, 25, 31, 40, 45, 48, 49, 53

D: 2, 8, 10, 13, 17, 27, 36, 37, 39, 41, 46, 51, 56

scoring{

keys(4);

key4(nr = null, scores = (0,0,0,1), omit = null, variables = (item2,item8,item10,item13,item17,item27,item36,item37,item39,item41,item46,item51,item56), options = (A,B,C,D));

key3(nr = null, scores = (0,0,1,0), omit = null, variables = (item3,item5,item7,item11,item14,item20,item23,item25,item31,item40,item45,item48,item49,item53), options = (A,B,C,D));

key2(nr = null, scores = (0,1,0,0), omit = null, variables = (item4,item6,item16,item18,item24,item26,item32,item33,item35,item43,item44,item47,item50,item54), options = (A,B,C,D));

key1(nr = null, scores = (1,0,0,0), omit = null, variables = (item1,item9,item12,item15,item19,item21,item22,item28,item29,item30,item34,item38,item42,item52,item55), options = (A,B,C,D));

data(db = epsy5221, table = EXAM5948);

}

Rating Scale Analysis

Item analysis and score reliability analysis also can be conducted on rating scales. We will also conduct item and scale analysis on rating scales employing *Item Response Theory* later in the course – we will use the Liking Science data set for both exercises.

The scale for this analysis is called the ***Liking Science Scale*** that is administered to elementary school aged children to be used in programs intending to develop interests in science. One goal of science classes and museum programs is to develop a liking for science. The Cleveland Museum of Natural History created a scale to study children’s liking for science by assembling 25 science-related activities. The list of activities covers a range of effort required, suggesting that they would only be liked (to some degree) by children who are motivated by a liking for science.

Here is the list of items that children respond to, with the response scale following the list.

1. WATCH BIRDS
2. READ BOOKS ON ANIMALS
3. READ BOOKS ON PLANTS
4. WATCH GRASS CHANGE
5. FIND BOTTLES AND CANS
6. LOOK UP STRANGE ANIMAL OR PLANT
7. WATCH ANIMAL MOVE
8. LOOK IN SIDEWALK CRACKS
9. LEARN WEED NAMES
10. LISTEN TO BIRD SING
11. FIND WHERE ANIMAL LIVES
12. GO TO MUSEUM
13. GROW GARDEN
14. LOOK AT PICTURES OF PLANTS
15. READ ANIMAL STORIES
16. MAKE A MAP
17. WATCH WHAT ANIMALS EAT
18. GO ON PICNIC
19. GO TO ZOO
20. WATCH BUGS
21. WATCH BIRD MAKE NEST
22. FIND OUT WHAT ANIMALS EAT
23. WATCH A RAT
24. FIND OUT WHAT FLOWERS LIVE ON
25. TALK WITH FRIENDS ABOUT PLANTS

The response scale is: 0 = Dislike ☹ 1 = Neutral 😐 2 = Like ☺

Using jMetrik

Complete the following procedures:

1. Download the Liking Science data and Exam 1 data from the class website. Place these files in a folder where you can easily find them.
2. Start jMetrik
3. Create a Database in jMetrik
   1. Manage 🡪 New Database
   2. Type a name for the database (letters and numbers, e.g., EPSY5221; no blanks)
   3. Click “create”. Create a new Database for different projects, but do not create a Database for every data table – this will help you keep projects organized.
4. Import Data
   1. Manage 🡪 Open Database (select the database where you want to add the data table
   2. Manage 🡪 Import Data
      1. Type a unique name for the specific data table (e.g.: Exam1).
      2. Browse for your data file.
      3. You can write a brief description of the data file if you want.
      4. Click “Import”. At this point, you will see the data table in the left column. Click on the data table and the data should appear in the data window to the right.
5. Enter the item scoring information
   1. Select the Variables tab at the bottom of the data window
   2. Transform 🡪 Advanced Item Scoring
   3. Enter the Option value in the first column. Enter A if you have multiple-choice items with options A, B, C, D; or 0 for the Liking Science data since the options are 0, 1, 2.
   4. Enter the point value in the second column.
   5. For the Exam 1 data for all items where A is the correct response, you will have four rows with the first column containing A, B, C, D and the second column containing 1, 0, 0, 0.
   6. Select the items where the correct response is A and add them to the right dialogue box and “Submit”. Then score the items where B is the correct response, etc.
   7. Follow the same process for Liking Science – consider scoring 0,1,2 as -1, 0, 1.

Test Score & Item Analysis

1. Create scores for the test
   1. Transform 🡪 Test Scaling
   2. Click the double arrows >> to select all items to estimate total scores
   3. Score Name: SumScore – and Score: click Sum Score from the drop-down menu
   4. Run
2. Create a graph of the score distribution
   1. Graph 🡪 Histogram
   2. Select the SumScore
   3. Title: Distribution of SumScore
   4. Y-axis: Frequency
   5. Run
3. Conduct an Item Analysis
   1. Analyze 🡪 Item Analysis
   2. Click the double arrows >> to select all items for the analysis
   3. Options: Compute item statistics, All response options, & Correct for Spuriousness
   4. Run