Using jMetrik

Complete the following procedures:

1. Download and install jMetrik from: http://www.itemanalysis.com/
2. Start jMetrik
3. In your file directory, locate where jMetrik is saving the “database” folder. You can find this location by the following steps:
   1. In jMetrik, click on [Edit] in the toolbar.
   2. Select [Preferences].
   3. In the Preferences [Program Options] window, you will see “Home:…”
   4. This tells you the location where jMetrik will save the databases. Make a note of this location and open that folder – you will need to put the epsy5244 databases into this folder.
4. Download the jMetrik Database file called “epsy5244” and open the zipped folder so you get the regular folder epsy5244.
5. Copy the unzipped epsy5244 folder to the jmetrik Database folder. Be sure to place the copy of “epsy5244” inside the “database” folder in jMetrik.
6. Open a Database in jMetrik
   1. Manage 🡪 New Database
   2. Select “epsy5244” and click [Open]
   3. You should see a list of databases in the left column, including
      * Surveydata2001
      * Surveydata2002
      * Surveydata2003
      * Surveydata2004
   4. Select one of the data files and you will see the values in the columns on the right. You are now ready to use jMetrik for item analysis.

Sometimes you may want to create a new database for a new project. You can create as many databases as you need. It is helpful to create separate databases for each major project you work on. Then, you can have multiple data tables within each database.

The following steps will allow you to create a new database. When you import data tables into the new database, be sure they are in comma-delimited (cvs) format. You can save Excel, SPPS, and many other data files in cvs format.

1. Create a Database in jMetrik
   1. Manage 🡪 New Database
   2. Type a name for the database (letters and numbers, e.g., epsy5244; 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.
2. 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.
3. 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 values in the first column. Enter A to D if you have multiple-choice items with options A, B, C, D; or 1 to 4 for 4-point rating-scale data where the options are 1, 2, 3, 4.
   4. Enter the point value in the second column. This will be 0 for incorrect options and 1 for the correct option for exam data. Or for rating-scale data, this will probably be the same values in the first column, if you want to assign values of 1 to 4 for the options 1 to 4.
   5. For test item 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 rating-scale data. Do the items with different rating-scale values one group at a time.
   8. Be sure to see the guide for how to score data in jMetrik at www.itemanalysis.com/user-guide.php - here you will find great guidance on Item Scoring and other functions in jMetrik.

Test Score & Item Analysis

1. Create scores for a measure
   1. Click [Transform] 🡪 then [Test Scaling]
   2. Select the items you want to use to create a total score and use the single arrow > to place them in the open box on the right.
   3. Give the measure you are scoring a Name, for example: GunAttitudes.
   4. You will probably want the Sum Score as the [Type], which is the default
   5. Then select [Run]
2. Create a graph of the score distribution
   1. Click [Graph] 🡪 the [Histogram]
   2. Select the name of the measure you just created
   3. Title: Distribution of “Measure”
   4. Y-axis: Frequency
   5. Then select [Run]
3. Conduct an Item Analysis
   1. Click [Analyze] 🡪 then [Item Analysis]
   2. Select the items you used to compute the total score for your new measure and use the single arrow > to place them in the open box on the right.
   3. In the Options section, select the following: Compute item statistics, All response options, & Correct for Spuriousness
   4. Then select [Run]

Interpretation of Output

There is much more explanation of all of this at the jMetrik website in their Help menu. It has a great User Guide with clear instructions and explanations.

1. Test Scaling – Sum Score

When you create a total score for a new measure, you will receive an output table in a new window. This will provide summary statistics for the new score you just created. You will also see this score in the last column of the Data tab – it is a new variable in your data now. The summary statistics provide the sample size, min and max values, mean, SD, skewness and kurtosis.

1. Item Analysis

When you do the item analysis of the same items, you will receive great information about the quality and functioning of the items. For this example, I used the first 5 items in SurveyData2002, including the questions about the quality of communication with advisors; the item names are listed in the first table. We might consider these items to be a measure of Advisor Communication Quality. The following table is from that output.

ITEM ANALYSIS

epsy5244.SURVEYDATA2002

November 19, 2016 21:39:41

======================================================================

Item Option (Score) Difficulty Std. Dev. Discrimin.

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

helpful\_ Overall 3.0243 0.8968 0.7570

1.0(1.0) 0.0590 0.2361 -0.5065

2.0(2.0) 0.1493 0.3570 -0.4461

3.0(3.0) 0.4583 0.4991 -0.1599

4.0(4.0) 0.3229 0.4684 0.5454

accrt\_c Overall 3.1771 0.8430 0.6428

1.0(1.0) 0.0347 0.1834 -0.4372

2.0(2.0) 0.0938 0.2920 -0.4008

3.0(3.0) 0.4757 0.5003 -0.2810

4.0(4.0) 0.3819 0.4867 0.4966

useful\_c Overall 3.2188 0.9014 0.7045

1.0(1.0) 0.0347 0.1834 -0.4708

2.0(2.0) 0.0903 0.2871 -0.4105

3.0(3.0) 0.4132 0.4933 -0.3430

4.0(4.0) 0.4410 0.4974 0.5955

career\_c Overall 2.3785 1.0456 0.5905

1.0(1.0) 0.2396 0.4276 -0.5929

2.0(2.0) 0.2674 0.4434 -0.2222

3.0(3.0) 0.3264 0.4697 0.1536

4.0(4.0) 0.1563 0.3637 0.4196

checking Overall 2.4479 1.0904 0.6429

1.0(1.0) 0.2361 0.4254 -0.6172

2.0(2.0) 0.2500 0.4338 -0.2398

3.0(3.0) 0.3021 0.4600 0.0953

4.0(4.0) 0.2014 0.4017 0.4925

======================================================================

This first table contains item statistics. Each column contains important information:

1. The “Item” names are listed in this first column.
2. “Option” and “Score” values are listed in the next column. The number on the left is the option number and the number in parenthesis is the score or numeric value we assign to that option.
3. “Difficulty” is intended for test items, but it tells us the proportion of people that selected each option. If this was a knowledge test, the “Overall” difficulty would be the proportion correct. But since we have rating-scale items here (Disagree to Agree), the “Difficulty value is the average rating-scale value for that item.
4. The “Std. Dev.” provides the standard deviation values for each item and each option. This is not really useful for most purposes.
5. “Discrimin.” Provides the item discrimination value. This is a correlation between the item score (the item response) and the total score of all the items combined. The item discrimination tells us the extent to which the item contributes to the total score for this measure. We want the values of the “Overall” item discrimination to be positive and greater than .20 for it to be a high-quality item that contributes to the measure. If the item discrimination was near zero or negative, this would indicate that the item measures something different than what the other items are measuring – it is an item that should probably not be included in the total score for that measure. Here we see all items are strong contributors to our overall measure of Advisory Communication Quality.

TEST LEVEL STATISTICS

==========================================

Number of Items = 5

Number of Examinees = 288

Min = 0.0000

Max = 20.0000

Mean = 14.2465

Median = 14.0000

Standard Deviation = 3.7894

Interquartile Range = 5.0000

Skewness = -0.7846

Kurtosis = 1.1105

KR21 = 3.5435

==========================================

The next table is a simple table of summary statistics, much like the one that was generated when the total score was computed.

RELIABILIY ANALYSIS

===========================================================================

Method Estimate 95% Conf. Int. SEM

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

Guttman's L2 0.8545 (0.8262, 0.8795) 1.4453

Coefficient Alpha 0.8501 (0.8209, 0.8758) 1.4671

Feldt-Gilmer 0.8518 (0.8229, 0.8773) 1.4586

Feldt-Brennan 0.8536 (0.8250, 0.8787) 1.4500

Raju's Beta 0.8501 (0.8209, 0.8758) 1.4671

===========================================================================

Finally, the third table summarizes the Reliability Analysis. This provides five different estimates of total score reliability or consistency. Because Coefficient Alpha requires assumptions that are rarely met in practice, we strongly recommend against its use – even though it is the most commonly reported form of score reliability.

Instead, consider reporting Guttman’s L2 or the Feldt-Brennan estimates of reliability. These are based on a more realistic set of assumptions that are appropriate for survey data. Here we see that the Guttman value is .85. This is particularly good for only a 5 item measure. This suggests that 85% of the observed score variance (magnitude of individual differences) is due to true-score variance (real or systematic differences among individuals) rather than error-score variance (or random measurement error). We could also say that 15% of the observed-score variance is measurement error variance.

For most research purposes, we hope this value is greater than .70 – so that at least 70% of the variance in our scores is true systematic variance and not measurement error. Of course higher is better. But with survey data from rating-scale items, we can get high score reliability values, sometimes as much as .90 or higher. This provides a high-quality set of scores that can be used to answer important questions. But most importantly, it provides assurance (validity evidence) that we have minimized measurement error in our survey data for this particular measure.