**HANDOUT #5.1**

**Module 5.1: Data Analysis**

*Post hoc* data analysis is a key step in understanding the psychometric properties of a given assessment. These quantitative data help inform the item development team in understanding the performance of both individual items and the overall test itself. The statistical information provides the foundation for those inferences about the test-takers’ knowledge, skills, and abilities being measured by the test.

**Workflow: Data Analysis**

**Scoring**

**Statistics**

**Analytics**

**5.1.1 Procedural Steps: Calculating the Percent Correct (p-values)**

1. Enter all unscored items/tasks into an Excel spreadsheet.
2. Convert the unscored items/tasks into scored (0 = wrong; 1 = right) for each SR item. For CR items, add weighting (if applicable) to scored answers.
3. For each scored column, calculate the *p*-value by totaling the column values (numerator) and dividing by the total number of scores in the column (denominator).
4. For unanswered items, the scored value should be assigned to 0, rather than being omitted from the numerator and denominator.
5. Verify the full range of scores are being included in the Excel formula (e.g., see Module 5- Data Sample 2015 formula [=SUM(AE2:AE101)/100].

**5.1.2 Procedural Steps: Calculating the Item Number and Raw Score Correlation**

1. Create a Raw Score variable for each test-taker by aggregating all scored items on the assessment, including the CR items.
2. For each scored column, calculate the Pearson correlation by creating two unique data arrays. [Excel syntax: CORREL (array1, array2)]
3. Create the first data array by selecting all column values (Array 1) and the second array by selecting all column values (Array 2) for the overall raw score. .
4. Place the correlation coefficient below the *p*-value calculation for each item on the assessment. (e.g., see Module 5- Data Sample 2015 formula [=CORREL(BD2:BD101,$BF$2:$BF$101)].
5. Identify any coefficient value below *r*<.10, including any negative values (e.g., *r* = -.05).

**5.1.3 Procedural Steps: Calculating Omission/Attempted Rates**

1. Using the unscored responses, calculate the number of test-takers that did not select and/or provide a response to the item/task.
2. For unanswered items, create an “omission” chart displaying for each item the number (count) of “NULL” responses, which will be calculated by subtracting the number of invalid responses from the denominator. [e.g., see Module 5-Data Sample 2015 formula=COUNTIF(D2:D101,"")]
3. For each unscored column, calculate the “attempted” rate by totaling the number of valid response (numerator) and dividing the aggregated value by the total number of possible responses in the column (denominator). [e.g., see Module 5-Data Sample 2015 formula =(100-D103)/100]
4. Verify the full range of scores are being included in the Excel formula (e.g., see Module 5- Data Sample 2015 formula [=SUM(AE2:AE101)].

|  |  |
| --- | --- |
|  |  |

**5.1.4 Procedural Steps: Calculating Differential Item Functioning (DIF) Rates**

1. Using the Raw Score variable (overall score), determine the count of overall *p*-value (only SR items) for the focal group (i.e., gender).
2. Determine the mean (average) *p*-value variable for members of the focal group (i.e., males vs. females)
3. Determine the members’ *p*-value deviation by subtracting the overall *p*-value from the focal group mean.
4. For each item, create a contingency table with the focal group and compare the *p*-value. Then, determine the deviation from the item *p*-value by subtracting the overall item *p*-value from the subgroup mean.
5. Determine if the item’s deviation falls with the upper and lower deviation values for the overall test.

|  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- |
| **OVERALL TEST *P*-VALUE**

|  |  |  |
| --- | --- | --- |
| **Focal Group** | ***p*-Value** | **Deviation** |
| F | 0.450 | -0.03 |
| M | 0.504 | 0.03 |
| **Test *p*-value** | **0.477** |   |

 | **ITEM #1 RAW SCORE *P*-VALUE**

|  |  |  |
| --- | --- | --- |
| **Focal Group** | ***p*-Value** | **Deviation** |
| F | 0.24 | -0.01 |
| M | 0.27 | 0.02 |
| **Item *p*-value** | **0.25** |   |

 |

**5.1.5 Procedural Steps: Calculating Distractor Comparisons**

1. Select the unscored values for all SR item types.
2. Create a frequency distribution table using Excel’s Pivot Table function by counting the number of test-takers that selected a particular answer option.
3. Determine the proportion of test-takers that responded to each of the answer options, including the identified correct answer.
4. Given the *p*-values as a context, identify any items with an incorrect answer option (i.e., distractor) that was selected by more test-takers than the correct answer.
5. Given the distribution of incorrect answer options, determine the quality of each distractor, specifically focusing on distractors with very low response values (i.e., below .10).

|  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- |
| **ITEM #1 RESPONSE PATTERN**

|  |  |  |
| --- | --- | --- |
| **Response Options** | **Count of M1** | **Proportional Response Rate** |
| 0-A | 12 | 0.13 |
| 1-B | 23 | 0.24 |
| 2-C | 36 | 0.38 |
| 3-D | 25 | 0.26 |
| **Grand Total** | **96** |  |

**Answer Option “D” correct answer** |

**5.1.6 Procedural Steps: Calculating Item Type Comparisons**

1. Select all scored values for SR item types.
2. Calculate the percent (PCT) of points earned by aggregating the points earned (numerator) and dividing by the total possible (denominator), and then convert the resultant into a percentage.
3. Select all scored values (points awarded given the CR scoring rubric) for all CR item types for each test-taker.
4. Calculate the percent (PCT) of points earned by aggregating the points earned (numerator) and dividing that value by the total possible points (denominator) across all rubric score ranges, and then convert the resultant into a percentage.
5. Determine the overall test percent correct for the two item types and determine if the observed differences between item types is greater than 10 PCT PTS (points).

**5.1.7 Procedural Steps: Calculating CR Frequency Distributions**

1. Select all scored values for the first construct response (CR) item using all test-takers data.
2. Using Excel’s Pivot Table function, count the number of test-takers assigned each point value within the given scoring rubric’s range.
3. Evaluate the frequency distribution created by Excel’s graphing function.
4. Examine the shape of the graph to determine scoring anomalies (e.g., significant numbers of test-takers were assigned the maximum number of points).
5. If two raters were used during scoring, juxtapose the frequency distributions created by the assignment of points for each rater. Identify any significant differences in the graphs’ shape.