Classical Item Analysis

Sharing ideas on assessment

Gede Pramudya[1]

[1]a Member of Department of Intelligent Computing and Analytics & ISTE
Faculty of Information and Communication Technology
Universiti Teknikal Malaysia Melaka with LaTeX

July 29, 2018

Outline

Hypothetical Scenario

The Dean of X-Faculty of Y-University was looking a set of 40 **multiple-choice items** in order to measure the accomplishment of faculty students toward learning outcomes set for Z-Subject. He/she, then, instructed a senior member of the faculty who is a subject expert to design and develop the items such that the distribution of the items is 'bell-curved', which is 20% items are less difficult, 60% items are difficult, and the other 20% items are very difficult. The senior member thought the assignment was a piece of cake. Eventually, the he/she accepted the assignment and produced the set as specified. *Please discuss and evaluate this scenario. By the end of this session, we may learn.*

Learning Outcomes

By the end of the session, participants are able to

- demonstrate the understanding of classical item analysis concept;
- demonstrate the understanding of difficulty of an item;
- demonstrate the understanding of discrimination of an item;
- generate item difficulty indices, discrimination indices, and discrimination coefficients using Microsoft Excel;
- interpret item difficulty index, item discrimination index, and item discrimination coefficient of an item.

Learning Outcomes

By the end of the session, participants are able to

- demonstrate the understanding of classical item analysis concept;
- demonstrate the understanding of difficulty of an item;
- demonstrate the understanding of discrimination of an item;
- generate item difficulty indices, discrimination indices, and discrimination coefficients using Microsoft Excel;
- interpret item difficulty index, item discrimination index, and item discrimination coefficient of an item.

Learning Outcomes

By the end of the session, participants are able to

- demonstrate the understanding of classical item analysis concept;
- demonstrate the understanding of difficulty of an item;
- demonstrate the understanding of discrimination of an item;
- generate item difficulty indices, discrimination indices, and discrimination coefficients using Microsoft Excel;
- interpret item difficulty index, item discrimination index, and item discrimination coefficient of an item.

Learning Outcomes

By the end of the session, participants are able to

- demonstrate the understanding of classical item analysis concept;
- demonstrate the understanding of difficulty of an item;
- demonstrate the understanding of discrimination of an item;
- generate item difficulty indices, discrimination indices, and discrimination coefficients using Microsoft Excel;
- interpret item difficulty index, item discrimination index, and item discrimination coefficient of an item.

Learning Outcomes

By the end of the session, participants are able to

- demonstrate the understanding of classical item analysis concept;
- demonstrate the understanding of difficulty of an item;
- demonstrate the understanding of discrimination of an item;
- generate item difficulty indices, discrimination indices, and discrimination coefficients using Microsoft Excel;
- interpret item difficulty index, item discrimination index, and item discrimination coefficient of an item.

Topics Covered

We intend to discuss the following.

- Classical Item Analysis
  - Item Difficulty Index
  - Item Discrimination Index
  - Item Discrimination Coefficient
- Using Microsoft Excel for the Item Analysis
- Interpretation of Item Analysis

Topics Covered

We intend to discuss the following.

- Classical Item Analysis
  - Item Difficulty Index
  - Item Discrimination Index
  - Item Discrimination Coefficient
- Using Microsoft Excel for the Item Analysis
- Interpretation of Item Analysis

Topics Covered

We intend to discuss the following.

- Classical Item Analysis
  - Item Difficulty Index
  - Item Discrimination Index
  - Item Discrimination Coefficient
- Using Microsoft Excel for the Item Analysis
- Interpretation of Item Analysis

Topics Covered

We intend to discuss the following.

- Classical Item Analysis
  - Item Difficulty Index
  - Item Discrimination Index
  - Item Discrimination Coefficient
- Using Microsoft Excel for the Item Analysis
- Interpretation of Item Analysis

Topics Covered

We intend to discuss the following.

- Classical Item Analysis
  - Item Difficulty Index
  - Item Discrimination Index
  - Item Discrimination Coefficient
- Using Microsoft Excel for the Item Analysis
- Interpretation of Item Analysis

Outline

Classical Item Analysis

- It covers (simple) statistical analysis and expert judgment to evaluate tests based on the *quality* of individual items, item sets, and (or) entire sets of items and the *relationship* of each item to the other items (McCowan & McCowan,1999).
- "It investigates the performance of items considered **individually** either in relation to some **external criterion** or in relation to the **remaining** item of the test" (Thompson & Levitop, 1985).
- The goal is to assist test developers to *improve* the instrument by revising or discarding items that do not meet a minimally acceptable standard.

Classical Item Analysis

- It covers (simple) statistical analysis and expert judgment to evaluate tests based on the *quality* of individual items, item sets, and (or) entire sets of items and the *relationship* of each item to the other items (McCowan & McCowan,1999).
- "It investigates the performance of items considered **individually** either in relation to some **external criterion** or in relation to the **remaining** item of the test" (Thompson & Levitop, 1985).
- The goal is to assist test developers to *improve* the instrument by revising or discarding items that do not meet a minimally acceptable standard.

Classical Item Analysis

- It covers (simple) statistical analysis and expert judgment to evaluate tests based on the *quality* of individual items, item sets, and (or) entire sets of items and the *relationship* of each item to the other items (McCowan & McCowan,1999).
- "It investigates the performance of items considered **individually** either in relation to some **external criterion** or in relation to the **remaining** item of the test" (Thompson & Levitop, 1985).
- The goal is to assist test developers to *improve* the instrument by revising or discarding items that do not meet a minimally acceptable standard.

The Universe

Item Analysis is divided into

- Classical Item Analysis
  - Difficulty index ρ
  - Discrimination index ξ
  - Discrimination Coefficient r_ϕ
  - Distractor analysis
  - Reliability
  - Validity
- Modern Item Analysis (Latent Trait Model)
  - Item Response Theory
  - Rasch Model

The Universe

Item Analysis is divided into

- Classical Item Analysis
  - Difficulty index ρ
  - Discrimination index ξ
  - Discrimination Coefficient r_ϕ
  - Distractor analysis
  - Reliability
  - Validity
- Modern Item Analysis (Latent Trait Model)
  - Item Response Theory
  - Rasch Model

Outline

Item Difficulty Index, ρ

- It is the *proportion* of examinees who answer an item correctly. It is the relative frequency with which examinees choose the correct response.
- It is a *characteristic of the item* and *the sample* that takes the test.
- Items tend to improve test **reliability** when the percentage of students who correctly answer the item is halfway between the percentage expected to correctly answer if pure guessing governed responses and the percentage (100 percent) who would correctly answer if everyone knew the answer (Tompson & Lavitop,1985).
- If ρ of an item is about the median between 0.25 and 1.00 (0.6125), then it tends to improve the reliability of the test.

Item Difficulty Index, ρ

- It is the *proportion* of examinees who answer an item correctly. It is the relative frequency with which examinees choose the correct response.
- It is a *characteristic of the item* and *the sample* that takes the test.
- Items tend to improve test **reliability** when the percentage of students who correctly answer the item is halfway between the percentage expected to correctly answer if pure guessing governed responses and the percentage (100 percent) who would correctly answer if everyone knew the answer (Tompson & Lavitop,1985).
- If ρ of an item is about the median between 0.25 and 1.00 (0.6125), then it tends to improve the reliability of the test.

Item Difficulty Index, ρ

- It is the *proportion* of examinees who answer an item correctly. It is the relative frequency with which examinees choose the correct response.
- It is a *characteristic of the item* and *the sample* that takes the test.
- Items tend to improve test **reliability** when the percentage of students who correctly answer the item is halfway between the percentage expected to correctly answer if pure guessing governed responses and the percentage (100 percent) who would correctly answer if everyone knew the answer (Tompson & Lavitop,1985).
- If ρ of an item is about the median between 0.25 and 1.00 (0.6125), then it tends to improve the reliability of the test.

Item Difficulty Index, ρ

- It is the *proportion* of examinees who answer an item correctly. It is the relative frequency with which examinees choose the correct response.
- It is a *characteristic of the item* and *the sample* that takes the test.
- Items tend to improve test **reliability** when the percentage of students who correctly answer the item is halfway between the percentage expected to correctly answer if pure guessing governed responses and the percentage (100 percent) who would correctly answer if everyone knew the answer (Tompson & Lavitop,1985).
- If ρ of an item is about the median between 0.25 and 1.00 (0.6125), then it tends to improve the reliability of the test.

Formula

$$\rho = \frac{T}{N}$$

where

- ρ = item difficulty index of an item
- T= number of examinees that answer the item correctly
- N= number of examinees.

For norm-referenced multiple choices test (McCowan & McCowan,1999):

- two choices, $\rho \leq 0.75$
- three choices, $\rho \leq 0.67$
- four choices, $\rho \leq 0.63$
- five choices, $\rho \leq 0.60$

Criterion-referenced test requires lower ρ

Outline

Item Discrimination, ξ

- It is the extend to which the item *discriminates* amongst examinees in high and low group.
- High performers should be more likely to answer a *good* item correctly and low performers more likely to answer incorrectly.
- If the sample is more than 50 then the lower and higher group are one third (Some say 25 or 27 percent) of the sample.

Item Discrimination, ξ

- It is the extend to which the item *discriminates* amongst examinees in high and low group.
- High performers should be more likely to answer a *good* item correctly and low performers more likely to answer incorrectly.
- If the sample is more than 50 then the lower and higher group are one third (Some say 25 or 27 percent) of the sample.

Item Discrimination, ξ

- It is the extend to which the item *discriminates* amongst examinees in high and low group.
- High performers should be more likely to answer a *good* item correctly and low performers more likely to answer incorrectly.
- If the sample is more than 50 then the lower and higher group are one third (Some say 25 or 27 percent) of the sample.

Formula

$$\xi = \frac{h - l}{n}$$

where:

ξ : index of discrimination of an item;

h : the number of upper performers that respond correctly;

l : the number of lower performers that respond correctly;

n : the number of lower or upper performers.

ξ-Practice (Ebel & Frisbie, 1986)

If $\xi \geq 0.40$ then the item is *good*,
else if $0.30 \leq \xi < 0.40$ then the item is *reasonably good*,
else if $0.20 \leq \xi < 0.30$ then the item is *marginal*,
else the item is poor.

Outline

DIscrimination Coefficient, r_ϕ

- It describes the *relationship* between the answers (score) to an item and the (total) test scores of examinees (Ebel & Frisbie, 1986);
- It shows the extend on how much predictive power an item has and how the item contributes to the prediction;
- It tells us about the predictive *validity* of the test, since that tends to favor the items of medium difficulty (Henrysson,1971);

DIscrimination Coefficient, r_ϕ

- It describes the *relationship* between the answers (score) to an item and the (total) test scores of examinees (Ebel & Frisbie, 1986);
- It shows the extend on how much predictive power an item has and how the item contributes to the prediction;
- It tells us about the predictive *validity* of the test, since that tends to favor the items of medium difficulty (Henrysson,1971);

DIscrimination Coefficient, r_ϕ

- It describes the *relationship* between the answers (score) to an item and the (total) test scores of examinees (Ebel & Frisbie, 1986);
- It shows the extend on how much predictive power an item has and how the item contributes to the prediction;
- It tells us about the predictive *validity* of the test, since that tends to favor the items of medium difficulty (Henrysson,1971);

Formula (Backhoff, Larrazolo, & Rosas, 2000)

$$r_{\phi} = \frac{\bar{x_1} - \bar{x_0}}{s_x} \sqrt{\frac{n_0 n_1}{n(n-1)}}$$

where

- r_{ϕ} = the biserial coefficient of correlation of an item;
- $\bar{x_1}$ = median of the total scores of those who answered an item correctly;
- $\bar{x_0}$ = median of the total scores of those who answered an item incorrectly;
- s_x = standard deviation of the scores;
- n_1 = number of those who answered an item correctly;
- n_0 = number of those who answered an item incorrectly;
- $n = n_0 + n_1$.

r_ϕ-Practice

If $r_\phi \geq 0.35$, then the item can be retained, else the item should be omitted or revised(McCowan & McCowan, 1999)

Outline

The relationships

What are the relationships between the index of difficulty (ρ), the index of discrimination (ξ), and the discrimination coefficient (r_ϕ) of an item?

Outline

Assumptions for ...

- The participants is grouped. Each group consists of 4 participants, at most;
- Each group of participants has a set of 20 marked answer sheets of multiple choices test of 15 items (I suppose at the minimum);
- Each group has an access to the Microsoft Excel;
- Each group is familiar with the application.

Assumptions for ...

- The participants is grouped. Each group consists of 4 participants, at most;
- Each group of participants has a set of 20 marked answer sheets of multiple choices test of 15 items (I suppose at the minimum);
- Each group has an access to the Microsoft Excel;
- Each group is familiar with the application.

Assumptions for ...

- The participants is grouped. Each group consists of 4 participants, at most;
- Each group of participants has a set of 20 marked answer sheets of multiple choices test of 15 items (I suppose at the minimum);
- Each group has an access to the Microsoft Excel;
- Each group is familiar with the application.

Assumptions for ...

- The participants is grouped. Each group consists of 4 participants, at most;
- Each group of participants has a set of 20 marked answer sheets of multiple choices test of 15 items (I suppose at the minimum);
- Each group has an access to the Microsoft Excel;
- Each group is familiar with the application.

Outline

Preparation for Hands-On Exercises

1. Go to the real-statistic.com, choose the menu Free Download, and download the Real Statistics Resource Pack for Excel;
2. Add-Ins the package in the spreadsheet by following the written instruction in the page;
3. Key in the data of the test in the form of a matrix of 20×15(20 rows and 15 columns);
4. Each row represents the distribution of student (testee)'s responses for item number $1, 2, \cdots, 15$;
5. In each cell of the array, key in 1 if the response is correct, and put 0 if the response is incorrect;
6. If you finish with entering the data, then you will have a 20×15 matrix of 0, 1, or blank;
7. The analysis will be attempted in the workshop later.

Preparation for Hands-On Exercises

1. Go to the real-statistic.com, choose the menu Free Download, and download the Real Statistics Resource Pack for Excel;
2. Add-Ins the package in the spreadsheet by following the written instruction in the page;
3. Key in the data of the test in the form of a matrix of 20×15(20 rows and 15 columns);
4. Each row represents the distribution of student (testee)'s responses for item number $1, 2, \cdots, 15$;
5. In each cell of the array, key in 1 if the response is correct, and put 0 if the response is incorrect;
6. If you finish with entering the data, then you will have a 20×15 matrix of 0, 1, or blank;
7. The analysis will be attempted in the workshop later.

Preparation for Hands-On Exercises

1. Go to the real-statistic.com, choose the menu Free Download, and download the Real Statistics Resource Pack for Excel;
2. Add-Ins the package in the spreadsheet by following the written instruction in the page;
3. Key in the data of the test in the form of a matrix of 20×15(20 rows and 15 columns);
4. Each row represents the distribution of student (testee)'s responses for item number $1, 2, \cdots, 15$;
5. In each cell of the array, key in 1 if the response is correct, and put 0 if the response is incorrect;
6. If you finish with entering the data, then you will have a 20×15 matrix of 0, 1, or blank;
7. The analysis will be attempted in the workshop later.

Preparation for Hands-On Exercises

1. Go to the real-statistic.com, choose the menu Free Download, and download the Real Statistics Resource Pack for Excel;
2. Add-Ins the package in the spreadsheet by following the written instruction in the page;
3. Key in the data of the test in the form of a matrix of 20×15(20 rows and 15 columns);
4. Each row represents the distribution of student (testee)'s responses for item number $1, 2, \cdots, 15$;
5. In each cell of the array, key in 1 if the response is correct, and put 0 if the response is incorrect;
6. If you finish with entering the data, then you will have a 20×15 matrix of 0, 1, or blank;
7. The analysis will be attempted in the workshop later.

Preparation for Hands-On Exercises

1. Go to the real-statistic.com, choose the menu Free Download, and download the Real Statistics Resource Pack for Excel;
2. Add-Ins the package in the spreadsheet by following the written instruction in the page;
3. Key in the data of the test in the form of a matrix of 20×15(20 rows and 15 columns);
4. Each row represents the distribution of student (testee)'s responses for item number $1, 2, \cdots, 15$;
5. In each cell of the array, key in 1 if the response is correct, and put 0 if the response is incorrect;
6. If you finish with entering the data, then you will have a 20×15 matrix of 0, 1, or blank;
7. The analysis will be attempted in the workshop later.

Preparation for Hands-On Exercises

1. Go to the real-statistic.com, choose the menu Free Download, and download the Real Statistics Resource Pack for Excel;
2. Add-Ins the package in the spreadsheet by following the written instruction in the page;
3. Key in the data of the test in the form of a matrix of 20×15(20 rows and 15 columns);
4. Each row represents the distribution of student (testee)'s responses for item number $1, 2, \cdots, 15$;
5. In each cell of the array, key in 1 if the response is correct, and put 0 if the response is incorrect;
6. If you finish with entering the data, then you will have a 20×15 matrix of 0, 1, or blank;
7. The analysis will be attempted in the workshop later.

Preparation for Hands-On Exercises

1. Go to the real-statistic.com, choose the menu Free Download, and download the Real Statistics Resource Pack for Excel;
2. Add-Ins the package in the spreadsheet by following the written instruction in the page;
3. Key in the data of the test in the form of a matrix of 20×15(20 rows and 15 columns);
4. Each row represents the distribution of student (testee)'s responses for item number $1, 2, \cdots, 15$;
5. In each cell of the array, key in 1 if the response is correct, and put 0 if the response is incorrect;
6. If you finish with entering the data, then you will have a 20×15 matrix of 0, 1, or blank;
7. The analysis will be attempted in the workshop later.