Distributional Properties of Rasch Standardized Outfit Statistics

Rajat Chadha

Rasch Model (Rasch, 1960, 1981)

- Probabilistic psychometric item response model
- Used to create interval-level measures from categorical data, such as persons' answers to items on an exam
- One important use is to assess quality of items in an exam
- Probability of a correct response by a person to an item:

$$P\{x_{ni} = 1 \mid \beta_n, \delta_i\} = \frac{e^{\beta_n - \delta_i}}{1 + e^{\beta_n - \delta_i}}$$

Where x_{ni} is the score observed for a person n on item i,
β_n is the person location on the measurement scale, and
δ_i is the item location on the measurement scale.

Standardized Outfit Statistic (Smith, 1991)

- A useful fit statistic to assess the fit of the data to the Rasch Model to assess measurement disturbances.

Mean square:

$$MS(UT) = \frac{1}{N} \quad \chi^2(UT)_i = \frac{1}{N} \sum_{n=1}^{N} \frac{(x_{ni} - p_{ni})^2}{w_{ni}}$$

Standard Deviation of the Mean square:

$$S[MS(UT)] = \frac{[\Sigma(1/w_{ni}) - 4N]^{1/2}}{N}$$

Cube root transformation

$$t = (MS^{1/3} - 1)(3/S) + (S/3)$$

Objective

- In most practical work, a frame of reference against which the performance of an item can be judged is useful.
- One way to establish this frame of reference is to simulate the data over a variety of test lengths and distribution of persons measures and item difficulties to examine mean, standard deviation, and percentile ranks (1-percentile rank = probability of observing a specific value or greater/smaller) for critical values of -2 and 2.
- This program enables fast computation of standardized fit statistics over a user-provided number of items, persons and replications and outputs various statistics listed above in a CSV file.
- Useful for research – I can input person and item locations and conduct simulation studies with little modifications to the program. (Smith, 1991; Smith, Schumacker, & Bush, 1998)

Approach

1. Seek user input for number of items, persons and replication
2. Generate a numpy array with person measures drawn from a sample of normally distributed population with M = 0, SD = 1
3. Generate a numpy array with item difficulty measure drawn from a sample of uniformly distributed population in the range of -2 and 2.
4. Generate a numpy array with probability of correct response to each item by each person.
5. Generate another random numpy array with the same size as the probability of correct response array with scores of 0 and 1 for each item and person interaction. Over an infinite number of replications, the number of times a person scores an item correctly is expected to converge to the probability of correct response for that person and item interaction.
6. Calculate standardized fit statistics
7. Output summary statistics to a Pandas dataframe
8. Create histograms to plot distribution of standardized outfit statistic for each item

Functions

generate_prob_table(n_items, n_persons)

replication_sample (n_items, n_persons)

generate_response_matrix(prob_table, sample)

check_extreme_scores(resp_matrix_sample)

calculate_outfit(prob,resp_matrix)

generate_summary_stats(std_outfit,n_items,n_persons,n_replications)

generate_plots(std_outfit)

Comments/Questions?