Item Response Theory and Rasch Models

Review of Basic Concepts

Simple Regression Analysis

Simple Logistic Regression

One-parameter logistic models

Review of Basic Concepts

Item response theory (IRT; Lord, 1953; 1980; Rasch, 1960) is a widely used approach to psychometric analysis

Primarily used for ability or knowledge tests with binary items (correct/incorrect), but can be used with ordinal responses and in other contexts

Used for item analysis, scale development/evaluation, and investigation of test bias

Specialized software, such as BILOG-MG, MULTILOG, PARSCALE, IRTPRO, and ltm in R

Review of Basic Concepts

Sometimes referred to as "latent trait theory," the IRT approach quantifies the relationship between the ability and the response to an item

The latent trait is represented in equations by the Greek letter theta, θ, which is referred to as the ability, trait, or proficiency

Review of Basic Concepts

We have already covered some of the basics of IRT:

- Item difficulty
- Discrimination index
- Item characteristic curves (ICC)
- Standard error of measurement

Simple Regression Analysis

The core of IRT models is the prediction of the response by the ability as depicted by the item characteristic curve

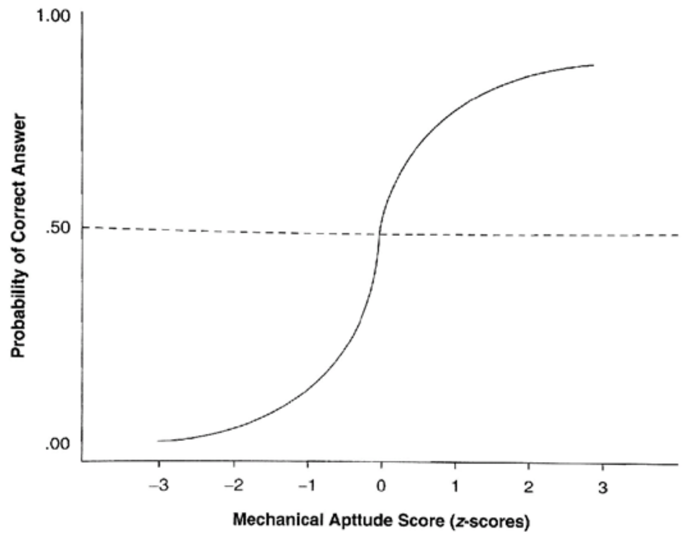


Figure 11.1 ICC Curve for a Test Item

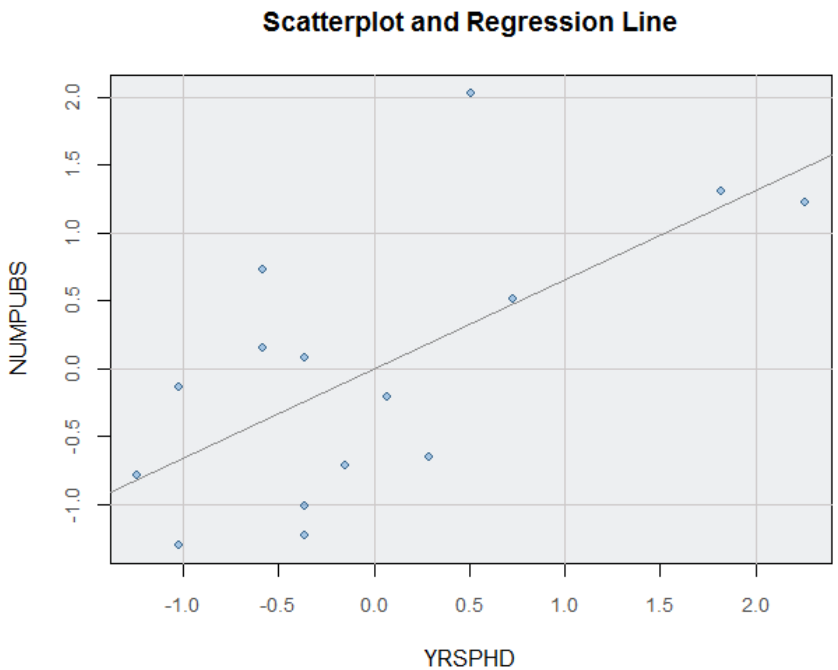
Furr, R. M., & Bacharach, V. R. (2013). *Psychometrics: an introduction*. Sage.

Simple Regression Analysis

The statistical model for prediction is known as *regression analysis*, whose term comes from regression toward the mean (Galton, 1886)

Regression analysis is based on the equation of a line, which is used to summarize the relationship between two variables

Simple Regression Analysis



Above is a scatterplot for the number of publications for professors in a psychology department predicted by the number of years since receiving their PhDs

Simple Regression Analysis

The equation for a line, $mx + b$, is used to summarize the trend of the points

In the equation for a line, m is the slope, which you should remember as "rise over run"—the amount that Y increases as X is incremented by one point

b is the intercept, or the point on the y-axis where the line intersects when X is equal to 0

Simple Regression Analysis

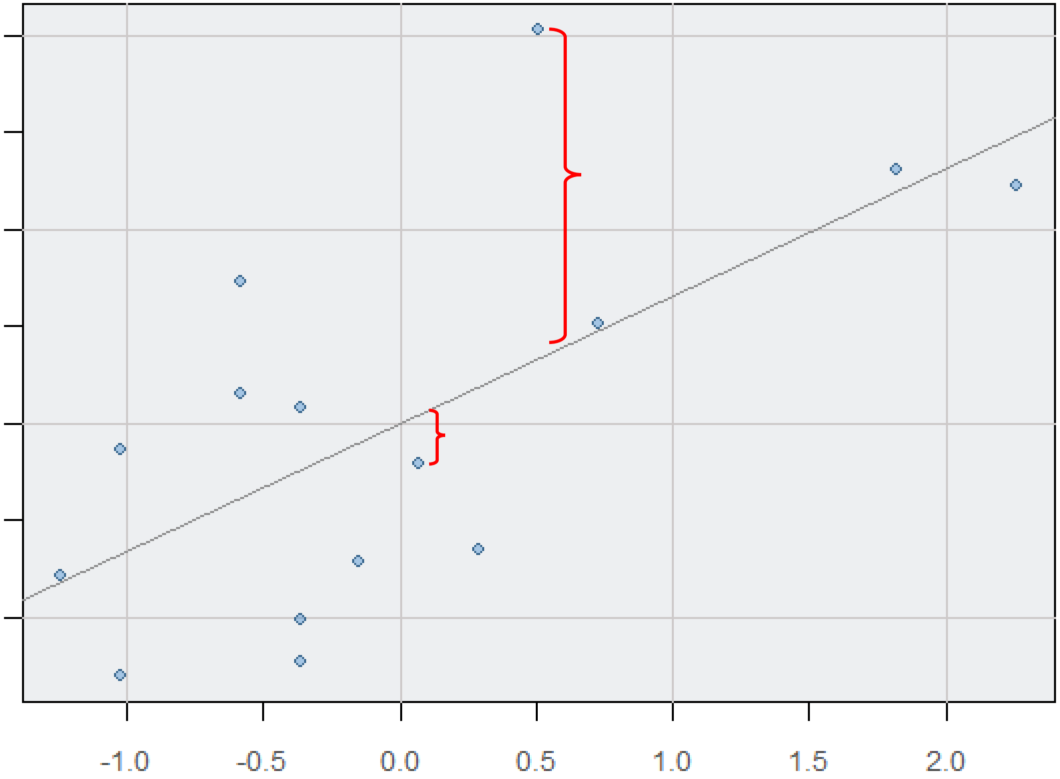
The line is comprised of the values the dependent variable (NUMPUBS) predicted by X (YRSPHD)—predicted or expected values called

The equation for the regression line is then

$$\hat{Y} = a + b(X)$$

Simple Regression Analysis

But as we can see in the scatterplot, the actual points are not always perfectly predicted



Simple Regression Analysis

The statistical equation then adds a term that accounts for the error of prediction—the distance of a point from the line, known as the residual

$$Y = a + b(X) + r$$

The actual or observed score, Y, is equal to the equation for the line plus some error, r

Simple Logistic Regression

In most IRT analyses, we want to predict whether the item is correct using the persons underlying ability

The better the prediction, the stronger the relationships between ability and the probability of a correct response, $P(Y=1)$

The problem is that predicting the probability of a binary variable is that the line not usually a straight

Simple Regression Analysis

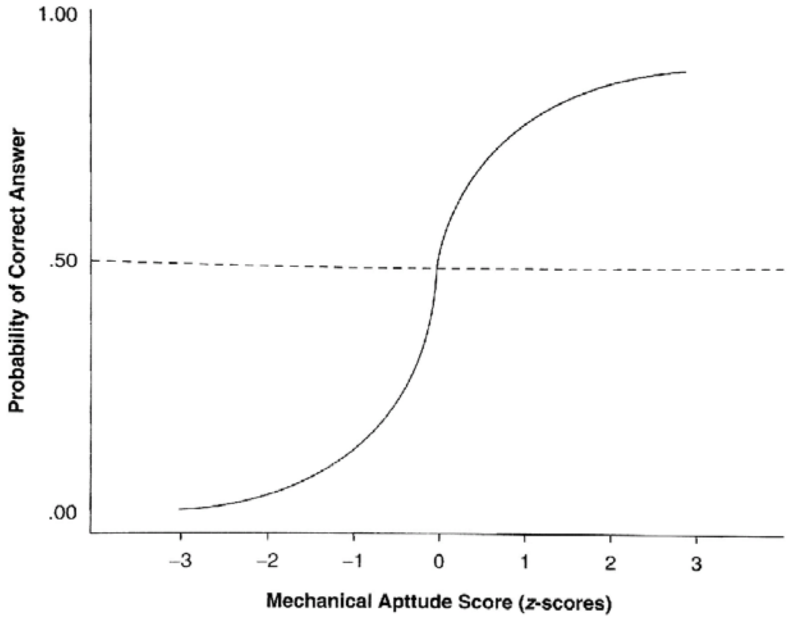


Figure 11.1 ICC Curve for a Test Item

Furr, R. M., & Bacharach, V. R. (2013). *Psychometrics: an introduction*. Sage.

Simple Logistic Regression

The predicted line has to be bent using an exponential transformation, which we call *logistic regression*

$$\ln\left(\frac{p}{1-p}\right)=a+bx$$

The predicted values, $\hat{Y}$, are replaced by the natural logarithm transformation, $\ln$, of the probability that $Y = 1$, represented by p

Simple Regression Analysis

To get back to the predicted values, we need to transform the other side of the equation using the opposite of the natural log, the exponential transformation

$$\hat{Y} = \frac{e^{a+bx}}{1+e^{a+bx}}$$

e is Euler's constant, which is approximately 2.71828182845904 (or rounding a bit more, 2.72)

Simple Regression Analysis

The exponential transformation, sometimes exp, is simply the constant *e* is raised to a power

For example, $e^3 \approx (2.72)^3 = 20.09$

We can return to the exponent value 3 by taking the natural log of the result, $\ln(20.09) = 3$

One-parameter Logistic Models

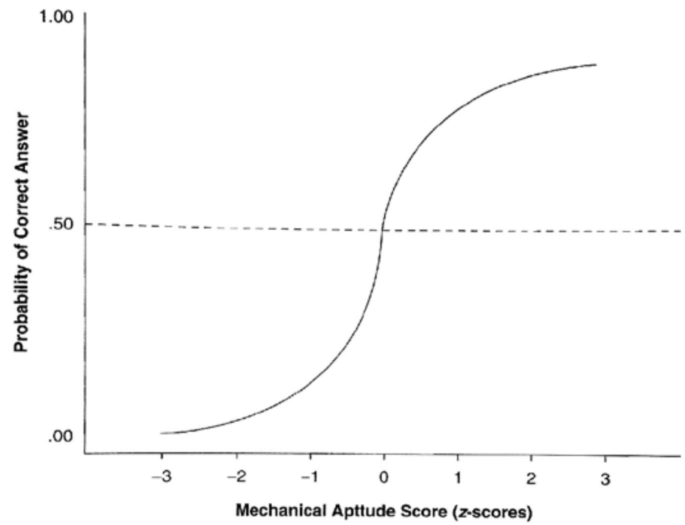
In the basic IRT model, the probability of a correct response, $\mathrm{P}(X_{is} = 1)$, is predicted by the ability, θ_s using logistic regression

The subscript i is for item and the subscript s is for subject (respondent)

Now take another look at the mechanical ability example from the text

One-parameter Logistic Models

$P(X_{is} = 1)$



θ_s

One-parameter Logistic Models

The one-parameter logistic model (1PL) or the equivalent Rasch model (Rasch, 1960) is a logistic regression model in a slightly altered form

The probability of the correct response, $\mathrm{P}(X_{is} = 1)$, is predicted by the ability, θ_s and the items difficulty, give as β_i

$$P\left(X_{is} = 1\right) = \frac{e^{\theta_s - \beta_i}}{1 + e^{\theta_s - \beta_i}}$$

Two-parameter Logistic Models

The two-parameter logistic model (2PL) extends the 1PL model by adding the discrimination parameter, α_i

$$P\left(X_{is}=1\right)=\frac{e^{\left(\alpha_i\left(\theta_s-\beta_i\right)\right)}}{1+e^{\left(\alpha_i\left(\theta_s-\beta_i\right)\right)}}$$