The Concept of a Scale

Measurement Methods

Lecture 6

2/14/06

Today’s Class

- Clarification of formula from last class.
- The concept of a scale (Chapter 4).
- Discussion of possible ideas for projects.

- NOTE: Homework #1 is up on blackboard (and due two weeks from today, 2/28).

As Seen Last Class

- As we discussed last class, one property of Likert scale or integer valued items is that the maximum variance is known.
- For an item with *k* responses, the maximum variance is:

$$\sigma^2_{\max} = \left(\frac{k-1}{2}\right)^2$$

- So, for an item with a total of five possible response options, the maximum variance would be 4.

Introduction to Scaling Theory

Introduction to Scaling

- Today's lecture is not meant to be a first introduction to some scaling theory.
  - More is treated in Chapter 18 in the book
    - We may not make it that far, though.

- Mainly we will define some measurement concepts and terms.

- And we will talk briefly about levels of measurement.

Purpose and Definition of Measurement

- The purpose of measurement is to quantify an attribute.
  - The term **attribute** is used broadly.

- **Measurement** – the assignment of numbers to an attribute according to a rule of correspondence (p. 55).

“Rules of Correspondence”

- The rules of correspondence can be bidirectional (5 = strongly agree and strongly agree=5).

- The rules of correspondence can also be unidirectional (number correct score on a test can be achieved by differing sets of item).
  – Many different item combinations lead to the same number correct score.

A Simple Example

- Imagine we create a test of 10 cognitive pass/fail items.

- The simple number right score fulfills the basic requirement of measurement.
  - It assigns a numeric value to the ability by a clear rule of correspondence.

Example Continued

- Now imagine we give the test to a large sample of examinees.
- It could happen that no examinee scores 0.
- It could happen that no examinee scores 10.
- In spite of the above “coulds,” there may be examinees in the population who would score 0 or would score 10 with the addition of more items.
  – Harder items and easier items.

Test Ceilings and Floors

- Any test will have a ceiling and a floor:
- **Floor** – the lowest attainable score on a test.
  - In our example, 0.
- **Ceiling** – the highest attainable score on a test.
  - In our example, 10.

Levels of Measurement

- When measurement happens, we often revert to thinking of the rules of physical measurement.
  - 10 feet is twice as long as 5 feet.
- The numbers we create through psychometrics, however, may not have the same properties.

Measurement Properties

- For our purposes, we distinguish the levels of measurement by three criteria:

1. Whether the measure has a true zero.
2. Whether equal intervals of the measure represent the same distance.
3. Whether greater quantities of the measure indicate an ordering.

Ratio Level Measures

- The a measure that satisfies all three criteria is called a **ratio scale**.
  - An example of a ratio scale include measures of length.
    1. Zero length means an object does not exist (true zero).
    2. We can say the distance from 5 feet to 10 feet is the same as the distance from 15 feet to 20 feet (equal intervals).
    3. We can say an object 10 feet tall has more length than an object 5 feet tall (order of magnitude).

Interval Level Measures

- An interval level scale satisfies two of the criteria.
  - It does not have a fixed zero.
- An oft-cited example is the Fahrenheit or Celsius temperature scale.
  - No fixed zero (zero is only symbolic).
  - Cannot say 20 degrees is twice as warm as 10 degrees.
- True temperature may differ from this example, however.
  - Temperature is a measure of random agitation and ambient photons under the effect thermal fluctuations.

Ordinal Level Measures

- Ordinal levels of measurement only satisfy one of the criteria:
  - Whether greater quantities of the measure indicate an ordering.
- Examples of ordinal levels of measurement abound in Psychology
  - Ratings on a Likert Scale

Which Level Do I Have?

- For those of you wishing to know about what level of measurement your data are, be prepared for tedious work.

- Axioms of measurement can often be used to describe the level your data are.

- Such techniques are beyond the scope of this class.

Scale Example

- **Scale values** (in Columns 2-4) are ways to assign a measurement to the performance of the test.

TABLE 4.1
Alternative Scales

	Number Right Scale	*Scale 2*	*Scale 3*	*Scale 4*
Passed all items	10	100	3.16	(?)
Passed 9 items	9	81	3	2.18
Passed 8 items	8	64	2.83	1.39
Passed 7 items	7	49	2.65	0.85
Passed 6 items	6	36	2.45	0.41
Passed 5 items	5	25	2.24	0
Passed 4 items	4	16	2	−0.41
Passed 3 items	3	9	1.73	−0.85
Passed 2 items	2	4	1.14	−1.39
Passed 1 item	1	1	1	−2.18
Failed all items	0	0	0	(−?)

Metrics

- **Metric** – the choice of a set of numbers to assign to the set of observations.

- From the previous example:
  - Column 1 – Count of correct items (0-10).
  - Column 2 – Square of count of correct items (0-100).
  - Column 3 – Square root of count of correct items (0-3.16).
  - Column 4 – Log-odds of proportion correct ($-\infty$ to ∞).

Choice of Metric

- Choosing a metric includes choosing an origin an a unit of measurement.
  – It is what Gabriel Fahrenheit and Anders Celsius did.
- Sometimes the choice is easy.
  – Total correct score.

Choice Motivation

- **Criterion-referenced Measurement**:

  Scores on a test are used in absolute way to see if an examinee meets a criterion of mastery of curricular materials, or a standard for selection to a program. Metric of the measurement is chosen without use of a distribution of responses from actual examinees.

Choice Motivation

- **Norm-referenced Measurement:**

  Metric is chosen <u>based on the distribution</u> of scores obtained from a population of interest.

Don’t Put Down Total Test Scores

- Can tell if test is too easy or difficult.
- Sometimes can be normally distributed.
  – If distribution is normal, we can change total test score into a standard score:

$$Z = \frac{Y - \mu_y}{\sigma_y}$$

Further Transformations

- Once a standard score Z is obtained, we can choose to further transform the scale:

$$S = cZ + a$$

- This will change the metric.
  - c is new standard deviation.
  - a is new mean.

Weschler Intelligence Test

- To demonstrate, Weschler intelligence tests use have a standard deviation of 15 and a mean of 100.

- We could arrive at these scores by transforming the raw score on the test.

Nonnormal Total Test Score Distributions

- Non-normally distributed total test score distributions can be transformed to a distribution that may be more “normal”
  - An example transformation is the square root.

- You can spend an eternity trying to find the right transformation, however.

Changing The Metric

- Changing the metric makes the numbers look more understandable.
  - 600 on an SAT may be more appealing than 35 out of 60.

- Changing the metric does not change the level of measurement.
  - Interval level measurement is still interval.

Wrapping Up

Next Time

- No class Thursday (2/16).
- Chapter 5 – Reliability Theory for Total Test Scores.