Assessment Literacy Self-Study Quiz #7

by Tim Newfields

This ongoing column features questions about testing, statistics, and assessment in a quiz format to promote greater assessment literacy. Suggested answers to the problems below are online at http://jalt.org/test/SSA7.htm.

Part I: Open Questions

1. Please examine the following test item and then answer these questions:(1) What skills is this task likely tapping into? (2) In what ways (if any) might this item be inappropriate in a general university undergraduate EFL entrance exam?

This item was taken from South Korea's 2008 College Scholastic Ability Test [*2008 Daehak suhak neungnyeok siheom*] – one of the three types of exams that high school applicants seeking university admission must take.

The task is to determine the correct sequence of three passages (A-C) that follow a 41-word paragraph opening.

Source: Korea Institute of Curriculum and Evaluation. (2008). *2008 Haknyeo-ndo Daehaksuhakneungryeoksihum Munjeji Oegukeo (Yeongeo) Yeongyeok – Holsuhyeong Oegukeo (Yeongeo) Yeongyeok 7.* [2008 College Scholastic Ability Test Foreign Language (English) Field 7]. Retrieved April 7, 2009 from http://www.kice.re.kr/ko/board/view.do?article_id=64464&menu_id=10089&board_id=10153

2008학년도 대학수학능력시험 문제지

외국어(영어) 영역

홀수형 외국어(영어) 영역 7

44. 주어진 글 다음에 이어질 글의 순서로 가장 적절한 것은?

> One of the toughest parts of isolation is a lack of an expressive exit. With anger, you can get mad at someone and yell. With sadness, you can cry. But isolation feels like being in a room with no way out.

(A) For people who cannot push themselves, however, support groups are a good cure for isolation. They offer the opportunity for connection in a safe and controlled way.

(B) And the longer you get stuck there, the harder it becomes to share the pain and sorrow. In isolation, hope disappears, despair rules, and you can no longer see a life beyond the invisible walls that imprison you.

(C) Some people find it helpful to work gently at driving themselves back into the world. In one case, a woman reported that after four miserable forced lunches with friends, she suddenly enjoyed the fifth one as she found herself laughing at a joke.

① (A) – (B) – (C)   ② (A) – (C) – (B)   ③ (B) – (C) – (A)
④ (C) – (A) – (B)   ⑤ (C) – (B) – (A)

2. Explain the difference between these concepts: (1) local test item independence and (2) test item unidimensionality.

3. Look at the box-percentile plots below, then briefly describe how the two groups differ. Also, what advantages do these projections have over common boxplots? Finally, how can such box-percentile plots such be drawn?
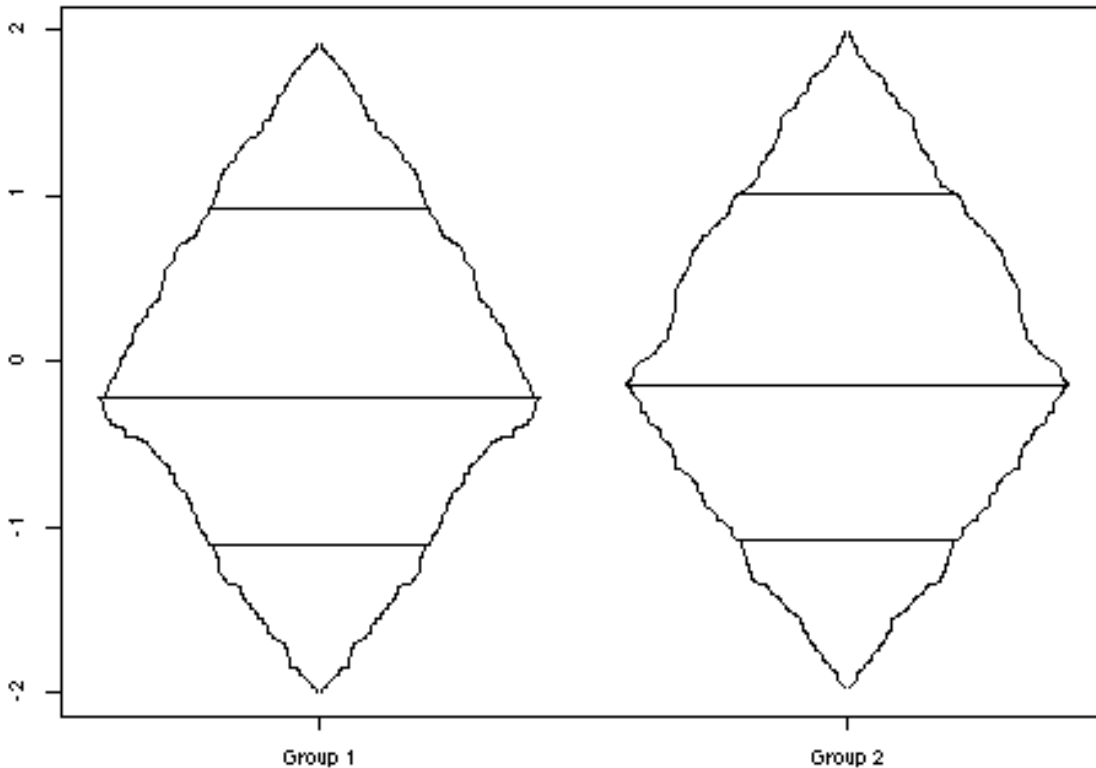
Source: Harrell, F. & Banfield, J. (2005, March 8). *R Graph Gallery Database: Box-Percentile plot.* Retrieved April 5, 2009 from http://addictedtor.free.fr/graphiques/RGraphGallery.php?graph=77

4. In 2005 Hirsch proposed a simple index as one way of quantifying faculty research output and granting promotions. This index, known as the *h*-index, combines two metrics: (1) the number of papers one has written [*k*] and (2) the number of times those papers have written have been cited [*d*]. The precise formula is:

$$\int_{h}^{\infty} f(k)\mathrm{d}k = h$$

Hence an author with a *h*-index of 12 has at least 12 publications, all of which have each been cited at least 12 times. That author might have a few publications with more citations, and many with less – but 12 is the largest number of articles-with-equal-number-of-citations for this person.

What are some advantages and disadvantages of Hirsch's ranking proposal?

Source: Hirsch, J. E. (2005, September 1). An index to quantify an individual's scientific research output. *Physics and Society,* (DOI: 10.1073/pnas.0507655102) Retrieved April 3, 2009 from http://arxiv.org/abs/physics/0508025

Part II: Multiple Choice Questions

1. Which of the following statements is **true** about Poisson distributions?

   (A) They approximate binomial (discrete probability) distributions if the sample size is at least 200 and significance level is 0.01 or less.

   (B) As the number of events they describe decrease, they resemble normal bell-curve distributions more.

   (C) As the number of occurrences they describe increase, the sum of probabilities for each distribution approaches 0.

   (D) They are useful in predicting the frequency of periodic events.

2. The graph below is an example of a _____.
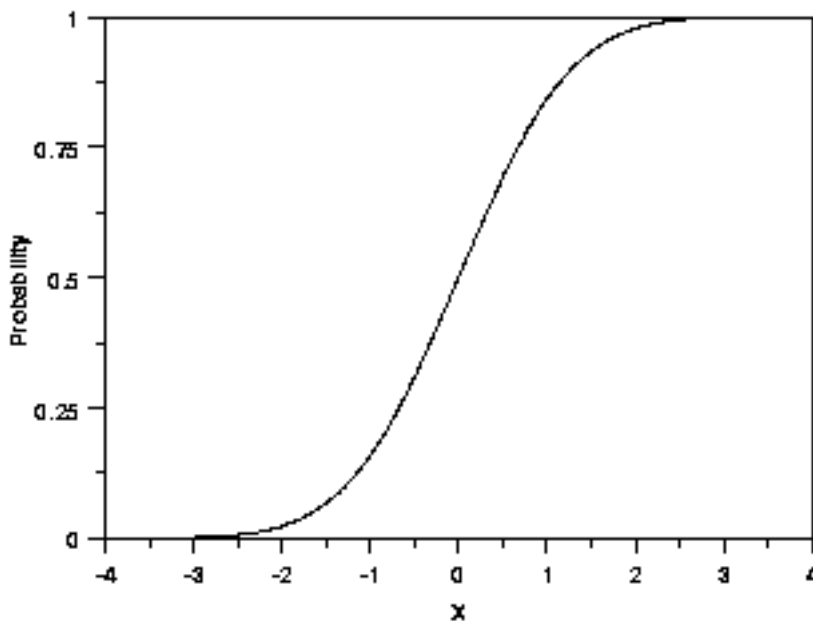
   (A) pareto chart   (C) cumulative distribution function (a.k.a ogive)

   (B) frequency polygon   (D) Q-Q plot (a.k.a quartile-quartile plot or QQ plot)

3. Which of the following is <u>least</u> likely to improve the reliability of an examination?

   (A) lengthening the number of test items.

   (B) standardizing all conditions under which the test is administered, including instructions.

   (C) deleting any items from the test that did not correlate with other items.

   (D) creating multiple forms of the test.

4. One difference between the KR-20 and KR-21 is _____.

   (A) Only the KR-21 is robust if the unifactor trait is violated, provided the test is longer than 18 items.

   (B) Reliability estimates obtained from the KR-21 are generally lower than those obtained from the KR-20.

   (C) The former requires just one administration to measure a test's internal consistency, but the later requires several administrations.

   (D) The former is used for dichotomous data, but the later is used for non-dichotomous data (i.e. partial credit answers).

5. What is an **ideal** item facility index (a.k.a. "*p*-value" or "item difficulty index") for a 4-option, multiple-choice norm-referenced test item? What about a true-false norm-referenced test item? What if both these items were designed for a criterion-reference test?

<div align="center">NOTE: <i>At least two answers for each sub-question are possible.</i></div>

   (A) Probably between .4 and .6.

   (B) Probably around .63.

   (C) Probably around .75.

   (D) Probably above .8.

   (E) Probably near the total test mean score difficulty.

   (F) Actually, more information is needed before deciding that.

HTML: http://jalt.org/test/SSQ7.htm   **PDF**: http:/jalt.org/test/PDF/SSQ7.pdf