Lecture 2: IRT Assumptions - Local Independence & Unidimensionality

(Baker & Kim (2004), Chapter 1 & Embretson & Reise (2000), Chapters 3-4)

Local independence (LI) implies that the probability of observing an n-item response pattern $U = (U_1, U_2, U_3, \ldots U_n)$ given ability θ, written $P(U|\theta)$, can be expressed as

$$P(U|\theta) = P(U_1|\theta)P(U_2|\theta)\ldots P(U_n|\theta).$$

Example: If the response pattern for an examinee on three items is (1, 1, 0), then the assumption of LI implies that

Unidimensionality is a requirement in most item response theory (IRT) models. A test is unidimensional provides it measures only one trait or ability. If a test is unidimensional, then a single trait exists such that all item pairs are locally independent. We can think of the dimensionality of a test as the number of dimensions that must be specified in order to achieve LI.

Example 1: Lazarsfeld & Henry (1968)

Suppose that a group of 1,000 persons is asked whether they have read the last issue of magazines A and B. Their responses are:

	Read A	Did not read A	Total
Read B	260	240	500
Did not read B	140	360	500
Total	400	600	1000

Now, suppose that we have information on the respondents' educational levels, dichotomized as high and low. When the 1,000 people are divided into these two groups, readership is observed to be:

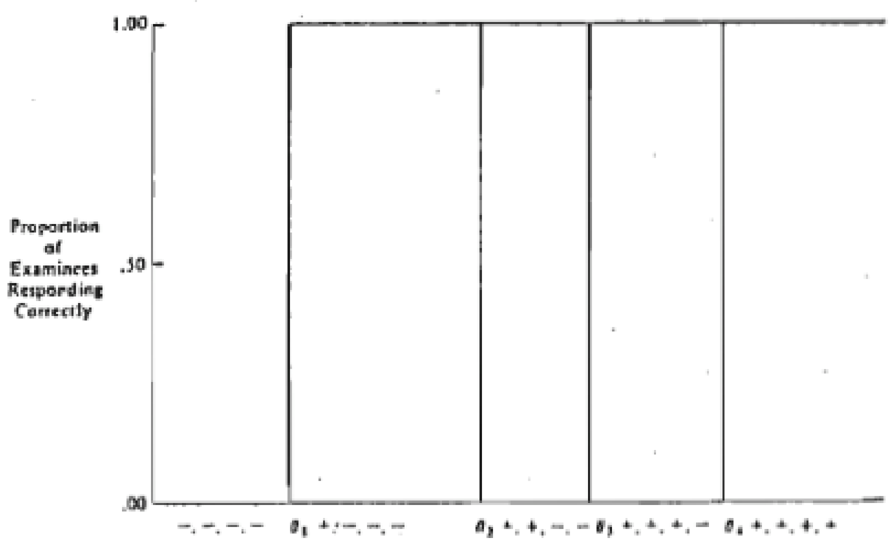
- *High Education*

	Read A	Did not read A	Total
Read B	240	60	300
Did not read B	160	40	200
Total	400	100	500

- *Low Education*

	Read A	Did not read A	Total
Read B	20	80	100
Did not read B	80	320	400
Total	100	400	500

<u>Example 2</u>: Suppose we have a four-item test and for a sample of 200 examinees we only observe the following response patterns: 0000, 1000, 1100, 1110, and 1111. We can then represent both items and examinees with respect to a Guttman scale:



- Joint distribution of responses for the entire population:

		Item 4 +	Item 4 −	
Item 3	+	.20	.20	.4
	−	.00	.60	.6
		.2	.8	

- Joint distribution of response for the subpopulation having $\theta = \theta_3$:

		Item 4 +	Item 4 −	
Item 3	+	.00	.00	.00
	−	.00	1.00	1.00
		.00	1.00	

Since this same type of 2×2 matrix will be observed for all item pairs, we obviously obtain local independence through specification of one dimension, so the test is unidimensional.

In application of IRT models, the potential for local dependence can also be checked. One index of local dependence is Yen's Q_3 statistic. Essentially Q_3 is a correlation between the residuals of item scores based on the item response model: $d_{ik} = u_{ik} - \hat{P}_i(\hat{\theta}_k)$. Then for a pair of items i and j, $Q_{3ij} = r_{d_i, d_j}$, where r represents the correlation between variables. Yen (1984) suggests that under an assumption of local independence, the distribution of Q_3 should be approximately normally distributed. Thus, to check local independence, one could construct a normal QQ-plot or apply some other test to check for normality.

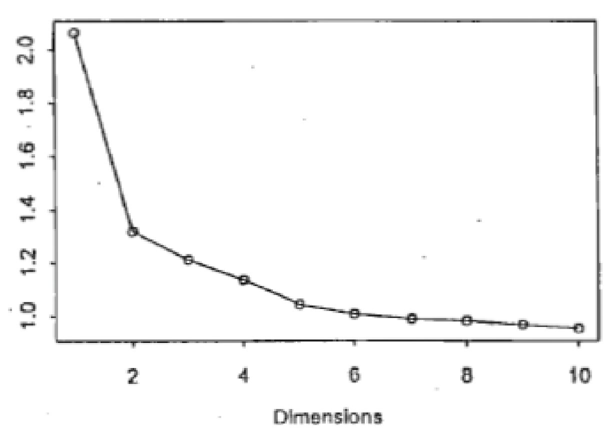
Yen. W. M. (1984). Effect of Local Item Dependence on the Fit and Equating Performance of the Three-Parameter Logistic Model. *Applied Psychological Measurement*, *8*, 125 – 145.

Some tools for checking unidimensionality

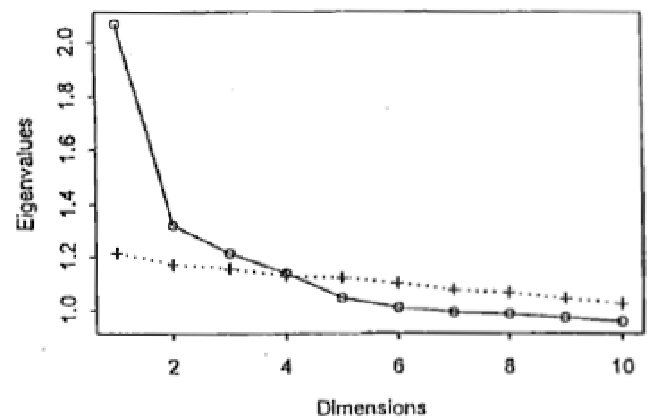
- Eigenvalue plots
- Parallel analysis; modified parallel analysis
- Full information item factor analysis (TESTFACT), limited information nonlinear factor analysis (NOHARM)
- Some nonparametric procedures: Stout's essential dimensionality procedure (DIMTEST)

Example: Law School Admissions Test Data – the Analytical Reasoning section (500 examinees, 24 items scored dichotomously, 1 = correct, 0 = incorrect).

- Eigenvalue plots: Examine the eigenvalues from a linear factor analysis, and determine how large the first eigenvalue is compared to the second.



- Parallel analysis: Compare the eigenvalue plots above with one generated from a random normal dataset with the same sample size and the same number of variables, but in this case with the variables assumed independent.



- TESTFACT: Compare the fit of a one-factor model to a two-factor model – the difference in G^2 can be regarded as a χ^2 variable with degrees of freedom equal to the difference across models in the number of parameters.
    - Fit of a one-factor model: $\chi^2 = 8038.45$, $df = 441$
    - Fit of a two-factor model: $\chi^2 = 7836.18$, $df = 418$
    - Difference: $\chi^2 = 102.27$, $df = 23$

- Stout's essential dimensionality procedure (DIMTEST): The basic idea is partition examinees into groups that obtained the same test score, and determine whether there exists a subset of items whose inter-item covariance are greater than would be expected if the test were unidimensional.
    - Step 1: Identify a subset of items in the dataset that appear to measure a second dimension. Call this subtest Assessment subtest 1 (AT1).
    - Step 2: Among the remaining test items, select a subtest containing an equal number of items as AT1 and whose items are of approximately the same difficulty as the items in AT1. Call this subtest Assessment subtest 2 (AT2). All of the remaining items comprise a Partitioning subtest (PT). Each examinee can be assigned a PT score based on the sum of item scores on the partitioning test.
    - Step 3: For examinees having the same PT score, compute the sum of inter-item covariances on AT1, and then sum these inter-item covariances across all possible PT scores. Repeat the process for AT2. The difference between the sums computed for the AT1 and AT2 subtests can be used to derive a t-statistic. A positive value suggests the inter-item covariances in the AT1 subtest are greater than would be expected by chance, and that a null hypothesis of unidimensionality can be rejected.

Monotonicity implies that as θ increases, the probability of correct response to the item should also increase, or $P(\theta_j) \geq P(\theta_k)$ whenever $\theta_j \geq \theta_k$.