University of York Department of Health Sciences

**Measurement in Health and Disease**

# Composite scales and scores

Martin Bland

http://martinbland.co.uk/

## Combining variables

Sometimes we have several outcome variables and we are interested in the effect of treatment on all of them.

Example: neurocognitive function of 212 coronary artery bypass surgery patients.

Randomised to have the procedure on-pump, i.e. an artificial pump took over the function of the heart, or off-pump, where the heart continued to function.

Did using the pump result in long-term damage to neurocognitive function?

Motallebzadeh R, Bland JM, Markus HS, Kaski JC, Jahangiri M. (2007) Neurocognitive function and cerebral emboli: randomised study of on-pump versus off-pump coronary artery bypass surgery. *Annals of Thoracic Surgery* **83**, 475-82.

## Combining variables

Sometimes we have several outcome variables and we are interested in the effect of treatment on all of them.

Battery of tests produced 21 different outcome variables.

If we compared each of these between the treatment groups, each individual variable would include only a small part of the information, so power would be reduced.

Also, the possibility of a Type I error, where we have a significant difference in the sample but no real difference in the population, would be increased.

We could deal with the type I error by the Bonferroni correction, multiplying each P value by 21, but this would reduce the power further.

## Combining variables

Find a combination of the 21 variables which contained as much of the available information as possible.

This was done using **principal component analysis** or **PCA**.

This finds a new set of variables, each of which is a linear combination of the original variables.

A **linear combination** is found by multiplying each variable by a constant coefficient and adding, as in a multiple regression equation.

In PCA, we make the sum of the coefficients squared equal to one.

## Combining variables

First we find the linear combination which has the greatest possible variance.

We call this the **first principal component**.

We then consider all the possible linear combinations which are not correlated with the first component and find the one with the largest variance.

This combination is the **second principal component**.

We then consider all the possible linear combinations which are not correlated with either the first or the second principal component and find the one with the largest variance.

This combination is the **third principal component**.

## Combining variables

Continue until we have as many principal components as there are variables.

Advantages of principal components over the original variables:

- all uncorrelated
- ordered by how much variance they have, how much information they contain.

These calculations are all done by computer programs and the mathematics is all done using matrix algebra.

We will omit this and go straight to the computer output.

## Combining variables

Eigenvalues for 21 neurocognitive test variables

| Component | Eigenvalue | % explained | cumulated % |
|---|---|---|---|
| 1 | 8.35 | 39.8 | 39.8 |
| 2 | 2.39 | 11.4 | 51.2 |
| 3 | 1.82 | 8.7 | 59.8 |
| 4 | 1.17 | 5.6 | 65.4 |
| 5 | 1.05 | 5.0 | 70.4 |
| 6 | 0.88 | 4.2 | 74.6 |
| 7 | 0.76 | 3.6 | 78.2 |
| 8 | 0.70 | 3.3 | 81.6 |
| 9 | 0.67 | 3.2 | 84.8 |
| 10 | 0.47 | 2.2 | 87.0 |
| 11 | 0.42 | 2.0 | 89.0 |
| 12 | 0.39 | 1.9 | 90.9 |
| 13 | 0.34 | 1.6 | 92.5 |
| 14 | 0.31 | 1.5 | 93.9 |
| 15 | 0.26 | 1.2 | 95.2 |
| 16 | 0.25 | 1.2 | 96.3 |
| 17 | 0.21 | 1.0 | 97.4 |
| 18 | 0.20 | 1.0 | 98.3 |
| 19 | 0.18 | 0.8 | 99.2 |
| 20 | 0.14 | 0.7 | 99.8 |
| 21 | 0.03 | 0.2 | 100.0 |
| Total | 21.00 | 100.0 | |

Eigenvalues: mathematical construct used in matrix algebra.

('Eigen' is German for 'own'.)

Just a name for something which tells us how variable the principal components are.


## Combining variables

Eigenvalues for 21 neurocognitive test variables

| Component | Eigenvalue | % explained | cumulated % |
|---|---|---|---|
| 1 | 8.35 | 39.8 | 39.8 |
| 2 | 2.39 | 11.4 | 51.2 |
| 3 | 1.82 | 8.7 | 59.8 |
| 4 | 1.17 | 5.6 | 65.4 |
| 5 | 1.05 | 5.0 | 70.4 |
| 6 | 0.88 | 4.2 | 74.6 |
| 7 | 0.76 | 3.6 | 78.2 |
| 8 | 0.70 | 3.3 | 81.6 |
| 9 | 0.67 | 3.2 | 84.8 |
| 10 | 0.47 | 2.2 | 87.0 |
| 11 | 0.42 | 2.0 | 89.0 |
| 12 | 0.39 | 1.9 | 90.9 |
| 13 | 0.34 | 1.6 | 92.5 |
| 14 | 0.31 | 1.5 | 93.9 |
| 15 | 0.26 | 1.2 | 95.2 |
| 16 | 0.25 | 1.2 | 96.3 |
| 17 | 0.21 | 1.0 | 97.4 |
| 18 | 0.20 | 1.0 | 98.3 |
| 19 | 0.18 | 0.8 | 99.2 |
| 20 | 0.14 | 0.7 | 99.8 |
| 21 | 0.03 | 0.2 | 100.0 |
| Total | 21.00 | 100.0 | |


Column of eigenvalues adds to 21, the number of variables.

The variances of the principal components are equal to the eigenvalues.


## Combining variables

Eigenvalues for 21 neurocognitive test variables

| Component | Eigenvalue | % explained | cumulated % |
|---|---|---|---|
| 1 | 8.35 | 39.8 | 39.8 |
| 2 | 2.39 | 11.4 | 51.2 |
| 3 | 1.82 | 8.7 | 59.8 |
| 4 | 1.17 | 5.6 | 65.4 |
| 5 | 1.05 | 5.0 | 70.4 |
| 6 | 0.88 | 4.2 | 74.6 |
| 7 | 0.76 | 3.6 | 78.2 |
| 8 | 0.70 | 3.3 | 81.6 |
| 9 | 0.67 | 3.2 | 84.8 |
| 10 | 0.47 | 2.2 | 87.0 |
| 11 | 0.42 | 2.0 | 89.0 |
| 12 | 0.39 | 1.9 | 90.9 |
| 13 | 0.34 | 1.6 | 92.5 |
| 14 | 0.31 | 1.5 | 93.9 |
| 15 | 0.26 | 1.2 | 95.2 |
| 16 | 0.25 | 1.2 | 96.3 |
| 17 | 0.21 | 1.0 | 97.4 |
| 18 | 0.20 | 1.0 | 98.3 |
| 19 | 0.18 | 0.8 | 99.2 |
| 20 | 0.14 | 0.7 | 99.8 |
| 21 | 0.03 | 0.2 | 100.0 |
| Total | 21.00 | 100.0 | |


Eigenvalue divided by the sum of all the eigenvalues is the proportion of the total amount of variance which that component represents.

% explained column.

## Combining variables

**Coefficients of the first principal component for 21 neurocognitive variables**

| Variable | 1 |
|---|---|
| cft | 0.03347 |
| cft1 | 0.24594 |
| cft2 | 0.24818 |
| gpt | -0.19108 |
| gpt1 | -0.16609 |
| ravlt_1 | 0.22261 |
| ravlt_2 | 0.23434 |
| ravlt_3 | 0.27129 |
| ravlt_4 | 0.27177 |
| ravlt_5 | 0.25437 |
| ravlt_b | 0.15745 |
| ravlt_6 | 0.25408 |
| ravlt_30min | 0.25588 |
| lct | -0.16818 |
| lct1 | -0.14615 |
| tmt | -0.19957 |
| tmt1 | -0.25476 |
| sdrt | -0.25251 |
| vft | 0.20014 |
| vft1 | 0.19292 |
| vft2 | 0.21412 |

If we square these and add them, we get 1.00.

Enables us to calculate the first principal component for each subject.

Standardise each variable (i.e. subtract the mean and divide by the standard deviation), multiply each by the coefficient, and add.

## Dimensions

The reason a single linear combination of the 21 variables can include 39.8% of the variation is that many of these neurocognitive test outcomes are correlated with one another.

Compare a simulation, where PCA was done using 21 randomly generated Normal variables for 200 subjects.

## Dimensions

**Eigenvalues for PCA using 21 randomly generated Normal variables for 200 subjects.**

| Component | Eigenvalue | % explained | Cumulative % |
|---|---|---|---|
| 1 | 1.64 | 7.8 | 7.8 |
| 2 | 1.52 | 7.2 | 15.0 |
| 3 | 1.42 | 6.8 | 21.8 |
| 4 | 1.32 | 6.3 | 28.1 |
| 5 | 1.29 | 6.1 | 34.3 |
| 6 | 1.28 | 6.1 | 40.4 |
| 7 | 1.21 | 5.8 | 46.1 |
| 8 | 1.14 | 5.4 | 51.5 |
| 9 | 1.09 | 5.2 | 56.7 |
| 10 | 1.00 | 4.8 | 61.5 |
| 11 | 0.95 | 4.5 | 66.0 |
| 12 | 0.92 | 4.4 | 70.4 |
| 13 | 0.88 | 4.2 | 74.6 |
| 14 | 0.83 | 4.0 | 78.5 |
| 15 | 0.79 | 3.8 | 82.3 |
| 16 | 0.78 | 3.7 | 86.0 |
| 17 | 0.73 | 3.5 | 89.5 |
| 18 | 0.63 | 3.0 | 92.5 |
| 19 | 0.59 | 2.8 | 95.3 |
| 20 | 0.51 | 2.4 | 97.7 |
| 21 | 0.48 | 2.3 | 100.0 |
| Total | 21.00 | 100.0 | |

First principal component explains only 7.8% of the variation.

With 21 principal components, the average percentage of variability explained by a component is 1/21 = 0.48 or 4.8%.

The average eigenvalue will be 1.00, since the 21 eigenvalues add up to 21.

## Dimensions

Eigenvalues for 21 neurocognitive test variables

| Component | Eigenvalue | % explained | Cumulated % |
|---|---|---|---|
| 1 | 8.35 | 39.8 | 39.8 |
| 2 | 2.39 | 11.4 | 51.2 |
| 3 | 1.82 | 8.7 | 59.8 |
| 4 | 1.17 | 5.6 | 65.4 |
| 5 | 1.05 | 5.0 | 70.4 |
| 6 | 0.88 | 4.2 | 74.6 |
| 7 | 0.76 | 3.6 | 78.2 |
| 8 | 0.70 | 3.3 | 81.6 |
| 9 | 0.67 | 3.2 | 84.8 |
| 10 | 0.47 | 2.2 | 87.0 |
| 11 | 0.42 | 2.0 | 89.0 |
| 12 | 0.39 | 1.9 | 90.9 |
| 13 | 0.34 | 1.6 | 92.5 |
| 14 | 0.31 | 1.5 | 93.9 |
| 15 | 0.26 | 1.2 | 95.2 |
| 16 | 0.25 | 1.2 | 96.3 |
| 17 | 0.21 | 1.0 | 97.4 |
| 18 | 0.20 | 1.0 | 98.3 |
| 19 | 0.18 | 0.8 | 99.2 |
| 20 | 0.14 | 0.7 | 99.8 |
| 21 | 0.03 | 0.2 | 100.0 |
| Total | 21.00 | 100.0 | |

For neurocognitive test variables, the first component explains a lot more variability than we would expect if the variables were uncorrelated, 39.8% compared to 4.8%.

## Dimensions

Principal component analysis is described as a method for reducing the dimensions of a set of data.

With 21 separate measurements we have 21 dimensions to our outcome variables.

But if we describe them instead by the first few principal components, we reduce the dimensions considerably.

## Dimensions

Eigenvalues for 21 neurocognitive test variables

| Component | Eigenvalue | % explained | Cumulated % |
|---|---|---|---|
| 1 | 8.35 | 39.8 | 39.8 |
| 2 | 2.39 | 11.4 | 51.2 |
| 3 | 1.82 | 8.7 | 59.8 |
| 4 | 1.17 | 5.6 | 65.4 |
| 5 | 1.05 | 5.0 | 70.4 |
| 6 | 0.88 | 4.2 | 74.6 |
| 7 | 0.76 | 3.6 | 78.2 |
| 8 | 0.70 | 3.3 | 81.6 |
| 9 | 0.67 | 3.2 | 84.8 |
| 10 | 0.47 | 2.2 | 87.0 |
| 11 | 0.42 | 2.0 | 89.0 |
| 12 | 0.39 | 1.9 | 90.9 |
| 13 | 0.34 | 1.6 | 92.5 |
| 14 | 0.31 | 1.5 | 93.9 |
| 15 | 0.26 | 1.2 | 95.2 |
| 16 | 0.25 | 1.2 | 96.3 |
| 17 | 0.21 | 1.0 | 97.4 |
| 18 | 0.20 | 1.0 | 98.3 |
| 19 | 0.18 | 0.8 | 99.2 |
| 20 | 0.14 | 0.7 | 99.8 |
| 21 | 0.03 | 0.2 | 100.0 |
| Total | 21.00 | 100.0 | |

For the neurocognitive data, the first five components explain 70.5% of the variability.

We could just analyse these five components and discard the remaining 16.

## Dimensions

**Eigenvalues for 21 neurocognitive test variables**

| Component | Eigenvalue | % explained | Cumulated % |
|---|---|---|---|
| 1 | 8.35 | 39.8 | 39.8 |
| 2 | 2.39 | 11.4 | 51.2 |
| 3 | 1.82 | 8.7 | 59.8 |
| 4 | 1.17 | 5.6 | 65.4 |
| 5 | 1.05 | 5.0 | 70.4 |
| 6 | 0.88 | 4.2 | 74.6 |
| 7 | 0.76 | 3.6 | 78.2 |
| 8 | 0.70 | 3.3 | 81.6 |
| 9 | 0.67 | 3.2 | 84.8 |
| 10 | 0.47 | 2.2 | 87.0 |
| 11 | 0.42 | 2.0 | 89.0 |
| 12 | 0.39 | 1.9 | 90.9 |
| 13 | 0.34 | 1.6 | 92.5 |
| 14 | 0.31 | 1.5 | 93.9 |
| 15 | 0.26 | 1.2 | 95.2 |
| 16 | 0.25 | 1.2 | 96.3 |
| 17 | 0.21 | 1.0 | 97.4 |
| 18 | 0.20 | 1.0 | 98.3 |
| 19 | 0.18 | 0.8 | 99.2 |
| 20 | 0.14 | 0.7 | 99.8 |
| 21 | 0.03 | 0.2 | 100.0 |
| Total | 21.00 | 100.0 | |

We would still have most of the information.

The remaining components will consist mainly of measurement error anyway and will have little real information in them.

## Dimensions

Two frequently used methods used to decide how many dimensions our variables really have.

**Kaiser criterion:** take all those components with eigenvalues greater than the average, which is 1.00.

For neurocognitive tests, we would have five dimensions to our data.

For random numbers, we would have 10.

**Cattell scree plot**: a plot of the eigenvalue against the principal component number.

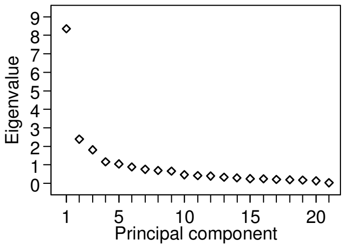
## Dimensions

**Eigenvalues for 21 neurocognitive test variables**

| Component | Eigenvalue |
|---|---|
| 1 | 8.35 |
| 2 | 2.39 |
| 3 | 1.82 |
| 4 | 1.17 |
| 5 | 1.05 |
| 6 | 0.88 |
| 7 | 0.76 |
| 8 | 0.70 |
| 9 | 0.67 |
| 10 | 0.47 |
| 11 | 0.42 |
| 12 | 0.39 |
| 13 | 0.34 |
| 14 | 0.31 |
| 15 | 0.26 |
| 16 | 0.25 |
| 17 | 0.21 |
| 18 | 0.20 |
| 19 | 0.18 |
| 20 | 0.14 |
| 21 | 0.03 |
| Total | 21.00 |

Kaiser: 5 dimensions

Scree plot:



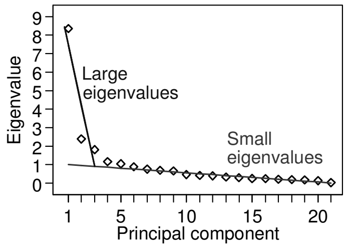
## Dimensions

Eigenvalues for 21 neurocognitive test variables

| Component | Eigenvalue |
|---|---|
| 1 | 8.35 |
| 2 | 2.39 |
| 3 | 1.82 |
| 4 | 1.17 |
| 5 | 1.05 |
| 6 | 0.88 |
| 7 | 0.76 |
| 8 | 0.70 |
| 9 | 0.67 |
| 10 | 0.47 |
| 11 | 0.42 |
| 12 | 0.39 |
| 13 | 0.34 |
| 14 | 0.31 |
| 15 | 0.26 |
| 16 | 0.25 |
| 17 | 0.21 |
| 18 | 0.20 |
| 19 | 0.18 |
| 20 | 0.14 |
| 21 | 0.03 |
| Total | 21.00 |

Kaiser: 5 dimensions.

Scree plot: 3 dimensions?



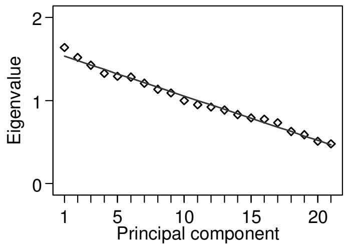

## Dimensions

Eigenvalues for 21 random numbers.

| Component | Eigenvalue |
|---|---|
| 1 | 1.64 |
| 2 | 1.52 |
| 3 | 1.42 |
| 4 | 1.32 |
| 5 | 1.29 |
| 6 | 1.28 |
| 7 | 1.21 |
| 8 | 1.14 |
| 9 | 1.09 |
| 10 | 1.00 |
| 11 | 0.95 |
| 12 | 0.92 |
| 13 | 0.88 |
| 14 | 0.83 |
| 15 | 0.79 |
| 16 | 0.78 |
| 17 | 0.73 |
| 18 | 0.63 |
| 19 | 0.59 |
| 20 | 0.51 |
| 21 | 0.48 |
| Total | 21.00 |

Kaiser: 10 dimensions.

Scree plot: no scree, cannot reduce dimensions.




## Composite scales

Want to measure ill-defined and abstract things, like disability, depression, anxiety, and health.

We could ask ‘how depressed are you on a scale of 1 to 10?’, or use a visual analogue scale:

|------------------------------------------------------------|

| | |
|---|---|
| not at depressed | as depressed as it is possible to be |

Subjects may not use that label for their problem.

## Composite scales

Instead form a composite scale.

Ask a series of questions relating to different aspects of depression and then combine them to give a depression score.

For example, the depression scale of one such questionnaire, the General Health Questionnaire or GHQ (Goldberg and Hillier 1979) is shown in Figure 4. .

Depression scale of the GHQ:

HAVE YOU RECENTLY

| | | | | | | | | |
|---|---|---|---|---|---|---|---|---|
| been thinking of yourself as a worthless person? | Not at all | **0** | No more than usual | **1** | Rather more than usual | **2** | Much more than usual | **3** |
| felt that life is entirely hopeless? | Not at all | **0** | No more than usual | **1** | Rather more than usual | **2** | Much more than usual | **3** |
| felt that life isn't worth living? | Not at all | **0** | No more than usual | **1** | Rather more than usual | **2** | Much more than usual | **3** |
| thought of the possibility that you might make away with yourself? | Definitely have | **3** | I don't think so | **2** | Has crossed my mind | **1** | Definitely not | **0** |
| found at times you couldn't do anything because your nerves were too bad? | Not at all | **0** | No more than usual | **1** | Rather more than usual | **2** | Much more than usual | **3** |
| found yourself wishing you were dead and away from it all? | Not at all | **0** | No more than usual | **1** | Rather more than usual | **2** | Much more than usual | **3** |
| found that the idea of taking your own life kept coming into your mind? | Definitely have | **3** | I don't think so | **2** | Has crossed my mind | **1** | Definitely not | **0** |

## Composite scales

Scoring for the depression scale of the GHQ:

Questions are scored 0, 1, 2, 3 for the choices from left to right for items 1, 2, 3, 5, and 6, and 3, 2, 1, 0 for items 4 and 7.

The sum of these is the score on the depression scale.

The questions are clearly related to one another and together should make a scale. Anyone who truthfully gets a high score on this is depressed.

The full questionnaire has four such scales.

## Composite scales

First devise a set of questions which are expected to be related to the concepts of interest based on experience.

The questions are answered by test subjects.

Do the questions form a coherent scale?

Do they measure one or more than one underlying construct?

## Composite scales

**Hull Reflux Cough Questionnaire (Alyn Morice)**

Please circle the most appropriate response for each question

Within the last MONTH, how did the following problems affect you?

0 = no problem and 5 = severe/frequent problem

| | | | | | | |
|---|---|---|---|---|---|---|
| 1. Hoarseness or a problem with your voice | 0 | 1 | 2 | 3 | 4 | 5 |
| 2. Clearing your throat | 0 | 1 | 2 | 3 | 4 | 5 |
| 3. The feeling of something dripping down the back of your nose or throat | 0 | 1 | 2 | 3 | 4 | 5 |
| 4. Retching or vomiting when you cough | 0 | 1 | 2 | 3 | 4 | 5 |
| 5. Cough on first lying down or bending over | 0 | 1 | 2 | 3 | 4 | 5 |
| 6. Chest tightness or wheeze when coughing | 0 | 1 | 2 | 3 | 4 | 5 |
| 7. Heartburn, indigestion, stomach acid coming up (or do you take medications for this, if yes score 5) | 0 | 1 | 2 | 3 | 4 | 5 |
| 8. A tickle in your throat, or a lump in your throat | 0 | 1 | 2 | 3 | 4 | 5 |
| 9. Cough with eating (during or soon after meals) | 0 | 1 | 2 | 3 | 4 | 5 |
| 10. Cough with certain foods | 0 | 1 | 2 | 3 | 4 | 5 |
| 11. Cough when you get out of bed in the morning | 0 | 1 | 2 | 3 | 4 | 5 |
| 12. Cough brought on by singing or speaking (for example, on the telephone) | 0 | 1 | 2 | 3 | 4 | 5 |
| 13. Coughing more when awake rather than asleep | 0 | 1 | 2 | 3 | 4 | 5 |
| 14. A strange taste in your mouth | 0 | 1 | 2 | 3 | 4 | 5 |

TOTAL SCORE_____________ /70

## Composite scales

Hull Reflux Cough Questionnaire, devised by Dr. Alyn Morice.

Questionnaire was devised using experience and evidence about the nature of respiratory symptoms.

It gives a single score, but does it really measure one thing?

To answer this we can do principal component analysis.

The data were obtained from 83 attendees at a chronic cough clinic.

## Composite scales

**Eigenvalues for the principal components of 14 respiratory questions**

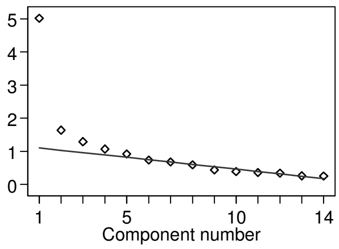
| Component | Eigenvalue | % explained | Cumulative |
|---|---|---|---|
| 1 | 5.02 | 35.9 | 35.9 |
| 2 | 1.64 | 11.7 | 47.6 |
| 3 | 1.30 | 9.3 | 56.8 |
| 4 | 1.07 | 7.6 | 64.5 |
| 5 | 0.92 | 6.6 | 71.1 |
| 6 | 0.74 | 5.3 | 76.4 |
| 7 | 0.68 | 4.9 | 81.2 |
| 8 | 0.59 | 4.2 | 85.4 |
| 9 | 0.44 | 3.1 | 88.6 |
| 10 | 0.39 | 2.8 | 91.4 |
| 11 | 0.36 | 2.6 | 93.9 |
| 12 | 0.34 | 2.4 | 96.3 |
| 13 | 0.26 | 1.9 | 98.2 |
| 14 | 0.25 | 1.8 | 100.0 |

Kaiser criterion: 4 dimensions

## Composite scales

Kaiser criterion: 4 dimensions

Scree plot: 2 or 3 dimensions




## Composite scales

Scales are difficult to design and validate.

Whenever possible we use one which has been developed previously.

Makes it easier to plan and to interpret the results of studies, as the properties of the scale are already known.

Wide range of scales readily available to the researcher.

Review of the literature in the field in which you propose to research will reveal what scales are available and which are used most often.

Bowling, A. (1997) *Measuring Health: A Review Of Quality Of Life Measurement Scales 2nd Ed.* Open University Press.

McDowell, I. and Newell, C. (1996) *Measuring Health: A Guide To Rating Scales And Questionnaires, 2nd Ed.* Oxford University Press

## Factor analysis

Factor analysis is a statistical method developed by psychologists.

Originally introduced by to answer questions like 'Is there more than one kind of intelligence?'.

By carrying out principal component analysis on a set of variables, we can decide whether there is more than one dimension.

There are other methods to do this as well, but we shall stick to PCA for this lecture.

Stata offers principal factor (the default), iterated principal factor, and maximum likelihood, principal component analysis. SPSS offers seven methods and has principal component analysis as the default.

## Factor analysis

The factor analysis model:

each variable can be represented as a linear combination of other variables, called **factors**, which we cannot actually see.

The factors are all set to have mean zero and variance one.

Each observed variable is the sum of each factor multiplied by a coefficient plus some unique factor of its own.

The coefficients are called factor loadings.

## Factor analysis

**Factor loadings for the first two factors for the Hull questionnaire**

| | Factor Loadings | | |
|---|---|---|---|
| Variable | Factor 1 | Factor 2 | Uniqueness |
| hoarse | 0.64 | −0.11 | 0.57 |
| throat | 0.58 | −0.58 | 0.33 |
| mucus | 0.60 | −0.33 | 0.53 |
| retching | 0.62 | 0.21 | 0.57 |
| lyingdwn | 0.66 | 0.24 | 0.51 |
| wheeze | 0.67 | 0.12 | 0.53 |
| heartbrn | 0.41 | 0.45 | 0.64 |
| tickle | 0.64 | −0.18 | 0.56 |
| eating | 0.75 | 0.15 | 0.42 |
| foods | 0.65 | 0.48 | 0.35 |
| outofbed | 0.58 | −0.22 | 0.61 |
| speaking | 0.62 | −0.38 | 0.47 |
| day | 0.39 | −0.33 | 0.74 |
| taste | 0.46 | 0.53 | 0.51 |

## Factor analysis

| Variable | Factor 1 | Factor 2 | Uniqueness |
|---|---|---|---|
| hoarse | 0.64 | −0.11 | 0.57 |
| throat | 0.58 | −0.58 | 0.33 |
| mucus | 0.60 | −0.33 | 0.53 |
| retching | 0.62 | 0.21 | 0.57 |
| lyingdwn | 0.66 | 0.24 | 0.51 |
| wheeze | 0.67 | 0.12 | 0.53 |
| heartbrn | 0.41 | 0.45 | 0.64 |
| tickle | 0.64 | −0.18 | 0.56 |
| eating | 0.75 | 0.15 | 0.42 |
| foods | 0.65 | 0.48 | 0.35 |
| outofbed | 0.58 | −0.22 | 0.61 |
| speaking | 0.62 | −0.38 | 0.47 |
| day | 0.39 | −0.33 | 0.74 |
| taste | 0.46 | 0.53 | 0.51 |

Factor Loadings

Standardised value of hoarse is given by

hoarse = 0.64 × factor 1 − 0.11 × factor 2 + 0.57 × error

where error is a Standard Normal random variable.

## Factor analysis

Such factors are called **latent variables**.

Dictionary definition of 'latent': concealed, not visible or apparent, dormant, undeveloped, but capable of development.

In statistics, we mean something which is not measured directly and the existence of which is inferred in some way.

We can estimate the numerical values of the factors from sets of coefficients.

These are not the same as the factor loadings.

The factor loadings are for calculating the variables from the factors, the factor coefficients are for calculating the factors from the variables.

## Factor analysis

Factor Loadings

| Variable | Factor 1 | Factor 2 | Uniqueness |
|---|---|---|---|
| hoarse | 0.64 | −0.11 | 0.57 |
| throat | 0.58 | −0.58 | 0.33 |
| mucus | 0.60 | −0.33 | 0.53 |
| retching | 0.62 | 0.21 | 0.57 |
| lyingdwn | 0.66 | 0.24 | 0.51 |
| wheeze | 0.67 | 0.12 | 0.53 |
| heartbrn | 0.41 | 0.45 | 0.64 |
| tickle | 0.64 | −0.18 | 0.56 |
| eating | 0.75 | 0.15 | 0.42 |
| foods | 0.65 | 0.48 | 0.35 |
| outofbed | 0.58 | −0.22 | 0.61 |
| speaking | 0.62 | −0.38 | 0.47 |
| day | 0.39 | −0.33 | 0.74 |
| taste | 0.46 | 0.53 | 0.51 |

Most of the loadings for Factor 1 are positive numbers and mostly of similar size.

The loadings for Factor 2 tend to be smaller and half of them are negative.

## Factor analysis

If we can predict our variables from two factors, we could also predict them from two other factors, each of which is a linear combination of the first two.

This is called a **factor rotation**.

For example,

$$\text{hoarse} = 0.64 \times \text{factor}_1 - 0.11 \times \text{factor}_2 + 0.57 \times \text{error}$$

Define two new variables, $\text{new}_1$ and $\text{new}_2$, so that

$$\text{new}_1 = \text{factor}_1 + \text{factor}_2$$

$$\text{new}_2 = \text{factor}_1 - \text{factor}_2.$$

## Factor analysis

Define two new variables, $\text{new}_1$ and $\text{new}_2$, so that

$$\text{new}_1 = \text{factor}_1 + \text{factor}_2$$

$$\text{new}_2 = \text{factor}_1 - \text{factor}_2.$$

Then

$$\text{factor}_1 = (\text{new}_1 + \text{new}_2)/2$$

$$\text{factor}_2 = (\text{new}_1 - \text{new}_2)/2$$

If we replace the old factors by the new:

$$\text{hoarse} = 0.64 \times \text{factor}_1 - 0.11 \times \text{factor}_2 + 0.57 \times \text{error}$$

$$\text{hoarse} = 0.64 \times (\text{new}_1 + \text{new}_2)/2 - 0.11 \times (\text{new}_1 - \text{new}_2)/2 + 0.57 \times \text{error}$$

$$= 0.27 \times \text{new}_1 + 0.38 \times \text{new}_2 + 0.57 \times \text{error}$$

## Factor analysis

$$\text{hoarse} = 0.64 \times \text{factor}_1 - 0.11 \times \text{factor}_2 + 0.57 \times \text{error}$$

$$\text{hoarse} = 0.27 \times \text{new}_1 + 0.38 \times \text{new}_2 + 0.57 \times \text{error}$$

There are many possible new pairs of factors which we could use.

Only use rotations which keep the standard deviations of the factors equal to one, which this example does not.

Note that the uniqueness remains the same.

## Factor analysis

Factor rotation: produce two new factors which have as many factor loadings as close to zero as possible.

As many variables as possible will be predicted mainly by only one factor.

This helps us to interpret the factors.

## Factor analysis

```
Factor loadings after varimax rotation for
two factors from Hull questionnaire

                     Factor Loadings
    Variable | Factor 1  Factor 2  Uniqueness
-------------+------------------------------
      hoarse |   0.53      0.38      0.57
      throat |   0.82      0.01      0.33
       mucus |   0.65      0.19      0.53
    retching |   0.28      0.59      0.57
    lyingdwn |   0.29      0.64      0.51
      wheeze |   0.39      0.57      0.53
    heartbrn |  -0.03      0.60      0.64
      tickle |   0.58      0.33      0.56
      eating |   0.42      0.64      0.42
       foods |   0.11      0.80      0.35
    outofbed |   0.57      0.26      0.61
    speaking |   0.71      0.17      0.47
         day |   0.51      0.05      0.74
       taste |  -0.06      0.70      0.51
```

Result of a **varimax** rotation, which keeps the factors uncorrelated.

## Factor analysis

Result of a **varimax** rotation, which keeps the factors uncorrelated.

Also possible to have correlated factors.

Rotation methods which produce them are called **oblique**.

Methods for rotation have names like quartimax, promax, quartimin, oblimax, and oblimin.

## Factor analysis

**Factor loadings after varimax rotation for two factors from Hull questionnaire**

| Variable | Factor 1 | Factor 2 | Uniqueness |
|---|---|---|---|
| hoarse | 0.53 | 0.38 | 0.57 |
| throat | 0.82 | 0.01 | 0.33 |
| mucus | 0.65 | 0.19 | 0.53 |
| retching | 0.28 | 0.59 | 0.57 |
| lyingdwn | 0.29 | 0.64 | 0.51 |
| wheeze | 0.39 | 0.57 | 0.53 |
| heartbrn | -0.03 | 0.60 | 0.64 |
| tickle | 0.58 | 0.33 | 0.56 |
| eating | 0.42 | 0.64 | 0.42 |
| foods | 0.11 | 0.80 | 0.35 |
| outofbed | 0.57 | 0.26 | 0.61 |
| speaking | 0.71 | 0.17 | 0.47 |
| day | 0.51 | 0.05 | 0.74 |
| taste | -0.06 | 0.70 | 0.51 |

Factor 1 mainly loads on hoarseness, clearing the throat, feeling of mucus, tickle in the throat, cough on getting out of bed, cough on singing or speaking, and cough more when awake.

## Factor analysis

**Factor loadings after varimax rotation for two factors from Hull questionnaire**

| Variable | Factor 1 | Factor 2 | Uniqueness |
|---|---|---|---|
| hoarse | 0.53 | 0.38 | 0.57 |
| throat | 0.82 | 0.01 | 0.33 |
| mucus | 0.65 | 0.19 | 0.53 |
| retching | 0.28 | 0.59 | 0.57 |
| lyingdwn | 0.29 | 0.64 | 0.51 |
| wheeze | 0.39 | 0.57 | 0.53 |
| heartbrn | -0.03 | 0.60 | 0.64 |
| tickle | 0.58 | 0.33 | 0.56 |
| eating | 0.42 | 0.64 | 0.42 |
| foods | 0.11 | 0.80 | 0.35 |
| outofbed | 0.57 | 0.26 | 0.61 |
| speaking | 0.71 | 0.17 | 0.47 |
| day | 0.51 | 0.05 | 0.74 |
| taste | -0.06 | 0.70 | 0.51 |

Factor 2 mainly loads on retching when cough, cough on lying down, tightness or wheeze, heartburn, cough with eating, cough with foods, and taste in the mouth.

## Factor analysis

Must decide what each factor represents.

Factor 1: hoarseness, clearing the throat, feeling of mucus, tickle in the throat, cough on getting out of bed, cough on singing or speaking, and cough more when awake.

Might label as 'respiratory tract cough'.

Factor 2: retching when cough, cough on lying down, tightness or wheeze, heartburn, cough with eating, cough with foods, and taste in the mouth.

Might label as 'alimentary tract cough'.

## Factor analysis

Factor loadings after varimax rotation for three factors for Hull questionnaire

| Variable | Factor 1 | Factor 2 | Factor 3 | Uniqueness |
|---|---|---|---|---|
| hoarse | 0.61 | 0.35 | 0.07 | 0.49 |
| throat | 0.76 | −0.08 | 0.32 | 0.31 |
| mucus | 0.73 | 0.16 | 0.07 | 0.43 |
| retching | 0.06 | 0.47 | 0.63 | 0.37 |
| lyingdwn | 0.19 | 0.56 | 0.41 | 0.34 |
| wheeze | 0.47 | 0.55 | 0.08 | 0.35 |
| heartbrn | 0.21 | 0.67 | −0.31 | 0.31 |
| tickle | 0.67 | 0.30 | 0.06 | 0.40 |
| eating | 0.33 | 0.55 | 0.43 | 0.38 |
| foods | 0.13 | 0.77 | 0.20 | 0.29 |
| outofbed | 0.23 | 0.10 | 0.83 | 0.26 |
| speaking | 0.68 | 0.10 | 0.27 | 0.33 |
| day | 0.29 | −0.07 | 0.54 | 0.29 |
| taste | −0.09 | 0.67 | 0.22 | 0.43 |

## Factor analysis

Factor loadings after varimax rotation for three factors for Hull questionnaire

| Variable | Factor 1 | Factor 2 | Factor 3 | Uniqueness |
|---|---|---|---|---|
| hoarse | 0.61 | 0.35 | 0.07 | 0.49 |
| throat | 0.76 | −0.08 | 0.32 | 0.31 |
| mucus | 0.73 | 0.16 | 0.07 | 0.43 |
| retching | 0.06 | 0.47 | 0.63 | 0.37 |
| lyingdwn | 0.19 | 0.56 | 0.41 | 0.34 |
| wheeze | 0.47 | 0.55 | 0.08 | 0.35 |
| heartbrn | 0.21 | 0.67 | −0.31 | 0.31 |
| tickle | 0.67 | 0.30 | 0.06 | 0.40 |
| eating | 0.33 | 0.55 | 0.43 | 0.38 |
| foods | 0.13 | 0.77 | 0.20 | 0.29 |
| outofbed | 0.23 | 0.10 | 0.83 | 0.26 |
| speaking | 0.68 | 0.10 | 0.27 | 0.33 |
| day | 0.29 | −0.07 | 0.54 | 0.29 |
| taste | −0.09 | 0.67 | 0.22 | 0.43 |

## Factor analysis

Factor loadings after varimax rotation for three factors for Hull questionnaire

| Variable | Factor 1 | Factor 2 | Factor 3 |
|---|---|---|---|
| hoarse | 0.61 | 0.35 | 0.07 |
| throat | 0.76 | −0.08 | 0.32 |
| mucus | 0.73 | 0.16 | 0.07 |
| retching | 0.06 | 0.47 | 0.63 |
| lyingdwn | 0.19 | 0.56 | 0.41 |
| wheeze | 0.47 | 0.55 | 0.08 |
| heartbrn | 0.21 | 0.67 | −0.31 |
| tickle | 0.67 | 0.30 | 0.06 |
| eating | 0.33 | 0.55 | 0.43 |
| foods | 0.13 | 0.77 | 0.20 |
| outofbed | 0.23 | 0.10 | 0.83 |
| speaking | 0.68 | 0.10 | 0.27 |
| day | 0.29 | −0.07 | 0.54 |
| taste | −0.09 | 0.67 | 0.22 |

Factor 1 mainly loads on hoarseness, clearing the throat, feeling of mucus, tickle in the throat, and cough on singing or speaking.

## Factor analysis

Factor loadings after varimax rotation for three factors for Hull questionnaire

| Variable | Factor 1 | Factor 2 | Factor 3 |
|---|---|---|---|
| hoarse | 0.61 | 0.35 | 0.07 |
| throat | 0.76 | -0.08 | 0.32 |
| mucus | 0.73 | 0.16 | 0.07 |
| retching | 0.06 | 0.47 | 0.63 |
| lyingdwn | 0.19 | 0.56 | 0.41 |
| wheeze | 0.47 | 0.55 | 0.08 |
| heartbrn | 0.21 | 0.67 | -0.31 |
| tickle | 0.67 | 0.30 | 0.06 |
| eating | 0.33 | 0.55 | 0.43 |
| foods | 0.13 | 0.77 | 0.20 |
| outofbed | 0.23 | 0.10 | 0.83 |
| speaking | 0.68 | 0.10 | 0.27 |
| day | 0.29 | -0.07 | 0.54 |
| taste | -0.09 | 0.67 | 0.22 |

Factor 2 mainly loads on cough on lying down, tightness or wheeze, heartburn, cough with eating, cough with foods, and taste in the mouth.

## Factor analysis

Factor loadings after varimax rotation for three factors for Hull questionnaire

| Variable | Factor 1 | Factor 2 | Factor 3 |
|---|---|---|---|
| hoarse | 0.61 | 0.35 | 0.07 |
| throat | 0.76 | -0.08 | 0.32 |
| mucus | 0.73 | 0.16 | 0.07 |
| retching | 0.06 | 0.47 | 0.63 |
| lyingdwn | 0.19 | 0.56 | 0.41 |
| wheeze | 0.47 | 0.55 | 0.08 |
| heartbrn | 0.21 | 0.67 | -0.31 |
| tickle | 0.67 | 0.30 | 0.06 |
| eating | 0.33 | 0.55 | 0.43 |
| foods | 0.13 | 0.77 | 0.20 |
| outofbed | 0.23 | 0.10 | 0.83 |
| speaking | 0.68 | 0.10 | 0.27 |
| day | 0.29 | -0.07 | 0.54 |
| taste | -0.09 | 0.67 | 0.22 |

Factor 3 mainly loads on retching when cough, cough on getting out of bed, and cough more when awake.

## Factor analysis

Factor 1: hoarseness, clearing the throat, feeling of mucus, tickle in the throat, and cough on singing or speaking.

Factor 2: cough on lying down, tightness or wheeze, heartburn, cough with eating, cough with foods, and taste in the mouth.

Factor 3: retching when cough, cough on getting out of bed, and cough more when awake.

Is this a more interpretable set of factors than the two factor rotation?

Factor 1 = ‘throat cough’,

Factor 2 = ‘alimentary tract cough’, (wheeze?)

Factor 3 not so clear. We might consider discarding those items and trying again.

## Factor analysis

Scoring Coefficients for the three factor solution

| Variable | Factor 1 | Factor 2 | Factor 3 |
|---|---|---|---|
| hoarse | 0.23 | 0.06 | -0.12 |
| throat | 0.31 | -0.19 | 0.07 |
| mucus | 0.31 | -0.04 | -0.11 |
| retching | -0.16 | 0.12 | 0.33 |
| lyingdwn | -0.07 | 0.17 | 0.15 |
| wheeze | 0.13 | 0.17 | -0.11 |
| heartbrn | 0.06 | 0.32 | -0.33 |
| tickle | 0.26 | 0.03 | -0.13 |
| eating | -0.01 | 0.15 | 0.14 |
| foods | -0.09 | 0.31 | 0.00 |
| outofbed | -0.06 | -0.10 | 0.48 |
| speaking | 0.26 | -0.09 | 0.03 |
| day | 0.05 | -0.15 | 0.30 |
| taste | -0.18 | 0.29 | 0.06 |

Variables which have high factor loads have high coefficients.

Anomalies.

Heartburn has quite a high (negative) coefficient for Factor 3, but does not load highly on it.

## Factor analysis

Scoring Coefficients for the three factor solution

| Variable | Factor 1 | Factor 2 | Factor 3 |
|---|---|---|---|
| hoarse | 0.23 | 0.06 | -0.12 |
| throat | 0.31 | -0.19 | 0.07 |
| mucus | 0.31 | -0.04 | -0.11 |
| retching | -0.16 | 0.12 | 0.33 |
| lyingdwn | -0.07 | 0.17 | 0.15 |
| wheeze | 0.13 | 0.17 | -0.11 |
| heartbrn | 0.06 | 0.32 | -0.33 |
| tickle | 0.26 | 0.03 | -0.13 |
| eating | -0.01 | 0.15 | 0.14 |
| foods | -0.09 | 0.31 | 0.00 |
| outofbed | -0.06 | -0.10 | 0.48 |
| speaking | 0.26 | -0.09 | 0.03 |
| day | 0.05 | -0.15 | 0.30 |
| taste | -0.18 | 0.29 | 0.06 |

We could include heartburn in the scale, subtracting its score from the sum of the other three items.

## Factor analysis

Having decided that a group of variables make up our scale, we might then simplify by making the coefficients for them all one and adding.

Thus the ‘throat cough scale’ becomes the sum of the scores for hoarseness, clearing the throat, feeling of mucus, tickle in the throat, and cough on singing or speaking.

## Internal consistency of scales

If a series of items are to form a scale, they should be correlated with one another.

A useful coefficient for assessing internal consistency is Cronbach's alpha (Cronbach 1951).

Alpha is a measure of how closely the items that make up the scale are correlated.

If the items are all perfectly correlated then alpha will be one, its maximum possible value.

If the items are all independent, having no relationship at all, then alpha will be zero.

In this case, of course, there is no coherent scale formed by summing them.

## Internal consistency of scales

Mathematically, the coefficient alpha is given by:

$$\alpha = \frac{k}{k-1}\left(1 - \frac{\sum \sigma_i^2}{\sigma_T^2}\right)$$

where $k$ = number of items,

$\sigma_i^2$ = variance of the $i$'th item

$\sigma_T^2$ = variance of the total scale formed by summing all the items.

Essential part of alpha is the sum of the variances of the items divided by the variance of the sum of all the items.

## Internal consistency of scales

$$\alpha = \frac{k}{k-1}\left(1 - \frac{\sum \sigma_i^2}{\sigma_T^2}\right)$$

If the items are all independent, then the variance of the sum will be the sum of the individual variances, $\sigma_T^2 = \Sigma\, \sigma_i^2$.

The ratio will be one and $\alpha = 0$.

If the items are all identical and so perfectly correlated, all the $\sigma_i^2$ will be equal and $\sigma_T^2 = k^2\, \sigma_i^2$. Because all the item variances are the same, $\Sigma\, \sigma_i^2 = k\sigma_i^2$, so $\Sigma\sigma_i^2/\sigma_T^2 = 1/k$ and $\alpha = 1$.

## Internal consistency of scales

Hull Reflux Cough Questionnaire:

| Scale | alpha |
|---|---|
| 1 | 0.78 |
| 2 | 0.79 |
| 3 | 0.68 (without heartburn) |
| 3 | 0.53 (with heartburn as negative item) |

Better to omit heartburn from Scale 3.

Scale 3 has poorer consistency than Scales 1 and 2.

Full Hull Reflux Cough Questionnaire scale (14 items) alpha = 0.86.

A fairly consistent scale.

## Internal consistency of scales

Alpha is based on the idea that our items are a sample from a large population of possible items which could be used to measure the construct which the scale represents.

If alpha is low, then the items will not be coherent and the scale will not necessarily be a good estimate of the construct.

Alpha can be thought of as an estimate of the correlation between the scale and another similar scale made from a different set of the possible items.

## Internal consistency of scales

In research, alpha = 0.7 or 0.8 considered acceptable.

A very high value, like 0.95, might indicate some redundancy in the scale.

For use in making clinical decisions about individual patients, it is considered that alpha should be higher, say 0.9 or greater.

### Internal consistency of scales

Alpha is often called a coefficient of reliability, or alpha reliability.

Not the same as the correlation between repeated administrations of the scale.

If the model is correct it should be similar.

### Internal consistency of scales

We can increase alpha by adding in more items, though the gain gets smaller as the number of items in the scale increases.

We can increase alpha by dropping items which are not highly correlated with others in the scale.

For example, heartburn has weaker correlations with retching, out of bed, and during the day than any of these have with one another.

### Problems with factor analysis

Factor analysis is often treated very sceptically by statisticians.

For example, Feinstein (2001, page 263) quoted Donald Mainland: 'If you don't know what you're doing, factor analysis is a great way to do it.'

*Factor Analysis as a Statistical Method* (Lawley and Maxwell 1971) implies that readers might not think of factor analysis as a statistical method at all!

Feinstein A. (2001) *Principles of Medical Statistics* CRC Press.

Lawley DN and Maxwell AE. (1971) *Factor Analysis as a Statistical Method, 2nd. Ed.* Butterworth.

## Problems with factor analysis

1. Factor analysis may be unstable over the items we use.

We may not get the same factors if we change some of the items, or add other items.

This is particularly true if we have a small number of subjects relative to the number of variable.

Random numbers can form factors.

2. Factor analysis may be unstable over the population of subjects.

If we use a different group of subjects, we might get different factors.

## Problems with factor analysis

3. The choice of number of factors is subjective.

Even if we use the objective Kaiser criterion, we may conclude that a factor is meaningless or uninterpretable and drop it.

4. The factor analysis model, with each observed variable being a linear combination of factors, means that the observed variables should be able to take any value in a range, i.e. should be continuous.

In the Hull Reflux Cough Questionnaire, the variables are all integers between 0 and 5, and certainly not continuous.

This is typical of the sort of data often used in factor analysis.

## Problems with factor analysis

5. The choice of label for the factor is subjective.

Different observers may interpret the same factor differently.

This leads to what is called the reification problem, that having labelled our factors, we then treat them as real things.

6. There are many variations on the factor analysis method and these may produce different structures.

**Problems with factor analysis**

Need to test scales:

- by repeating them among other groups of subjects,
- by estimating their repeatability,
- by comparing them with other observations.

Factor analysis remains the main method for establishing composite scales.

A complicated process, full of choices and pitfalls, and not to be undertaken lightly!