

# Reliability analysis

## Ανάλυση αξιοπιστίας

- ΔΙΑΤΥΠΩΣΗ ΤΟΥ ΠΡΟΒΛΗΜΑΤΟΣ.
- ΤΡΟΠΟΣ ΕΛΕΓΧΟΥ.
- ΥΛΟΠΟΙΗΣΗ ΣΤΟ SPSS.
- ΠΑΡΑΔΕΙΓΜΑ.

# ΔΙΑΤΥΠΩΣΗ ΤΟΥ ΠΡΟΒΛΗΜΑΤΟΣ

Συνήθως όταν θέλουμε να «μετρήσουμε» χαρακτηριστικά π.χ. η ποιότητα σχολικής ζωής, η ικανοποίηση από τις υπηρεσίες ενός νοσοκομείου, ξενοδοχείου και άλλες ανθρώπινες συμπεριφορές κατασκευάζουμε ένα τεστ, μία **κλίμακα** που αποτελείται από ένα πλήθος ερωτήσεων.

# Κλίμακα (Scale)

- Αποτελείται από ένα πλήθος ερωτήσεων σχετικές με αυτό που μετρά.
- Οι ερωτήσεις σε κάθε ερώτηση καταγράφονται και λαμβάνεται ο μ.ο. τους ως ένα σκορ σε κάθε πειραματική μονάδα.
- Η απάντηση σε μία επιμέρους ερώτηση (item) έχει ακριβώς την ίδια σημασία με κάθε άλλη απάντηση εντός της κλίμακας

# Κλίμακα

- Κάθε υποερώτηση είναι τέτοια ώστε να ρωτάμε με διαφορετικό τρόπο κάτι σχετικό με τη θεωρητική μας μεταβλητή (π.χ. ικανοποίηση από σχολείο).
- Οι πιθανές απαντήσεις σε κάθε υποερώτηση είναι του ίδιου τύπου και της ίδιας φοράς.
- Κλίμακες τύπου Likert.

# Likert Scale

- 2 point: ισχύει-δεν ισχύει, ναι-όχι....
- 3-point: ισχύει-ούτε ισχύει ούτε δεν ισχύει-δεν ισχύει....
- 5-point: ισχύει απόλυτα-ισχύει-ούτε ισχύει ούτε δεν ισχύει-δεν ισχύει-δεν ισχύει απόλυτα.
- 7-point: ισχύει απόλυτα-ισχύει- μάλλον ισχύει-ούτε ισχύει ούτε δεν ισχύει-μάλλον δεν ισχύει-δεν ισχύει-δεν ισχύει απόλυτα.

# Κλίμακα

- Κάθε κλίμακα πρέπει να διακρίνεται για δύο βασικές ιδιότητες της:
  1. ΕΓΚΥΡΟΤΗΤΑ (VALIDITY)
  2. ΑΞΙΟΠΙΣΤΙΑ (RELIABILITY)

# ΕΓΚΥΡΟΤΗΤΑ

- Η κλίμακα πρέπει να αντιπροσωπεύει επακριβώς τις διαστάσεις που θέλει να εξηγήσει ή να διαμορφώσει σε θεωρητικό επίπεδο. Διακρίνεται σε:
  1. Εγκυρότητα εννοιολογικής κατασκευής.
  2. Προβλεπτική εγκυρότητα.
  3. Συντρέχουσα εγκυρότητα.
  4. Εγκυρότητα περιεχομένου.

# Εγκυρότητα εννοιολογικής κατασκευής

- Απαντά στο ερώτημα αν τα δεδομένα υποστηρίζουν τη θεωρία που ερμηνεύει την έννοια. Διακρίνεται μεταξύ άλλων σε:
  1. Εγκυρότητα προσώπου: είναι η κλίμακα κατάλληλη;
  2. Συγκλίνουσα/ αποκλίνουσα: αν σχετίζεται/δεν σχετίζεται με άλλες που εκτιμούν/δεν εκτιμούν το ίδιο φαινόμενο.



# Προβλεπτική εγκυρότητα

- Όταν είναι σε θέση με βάση τις σημερινές αξιολογήσεις να προβλέψει μελλοντικές στάσεις π.χ. από την κλίμακα ικανοποίησης από ένα ξενοδοχείο να προβλέψουμε αν ξαναέρθει ο πελάτης, αν το συστήσει σε κάποιον άλλο κ.ο.κ.

# Συντρέχουσα εγκυρότητα

- Όταν οι τιμές της διαφοροποιούνται προς την αναμενόμενη κατεύθυνση μεταξύ ομάδων ατόμων οι οποίες αντικειμενικά διαφέρουν ως προς την υπό μελέτη κλίμακα.

# Εγκυρότητα περιεχομένου

- Οι ερωτήσεις καλύπτουν το εύρος του περιεχομένου της εννοιολογικής κατασκευής.

# Αξιοπιστία (Reliability)

- Η ικανότητα του ερωτηματολογίου- της κλίμακας να δίνει ίδια αποτελέσματα κάτω από τις ίδιες συνθήκες.
- Δείκτης αξιοπιστίας. Τιμές στο διάστημα  $[0,1]$ . Τιμές κοντά στο μηδέν έλλειψη αξιοπιστίας.

$$M=T+e$$

M: το σκορ στην κλίμακα, T: το θεωρητικό άγνωστο σκορ, e: το σφάλμα.

# Αξιοπιστία

- Ο δείκτης αξιοπιστίας  $\rho$  ορίζεται από την σχέση:  
Διακύμανση του T / Διακύμανση του M.
- Το ποσοστό της μεταβλητότητας του παρατηρούμενου σκορ που μπορεί να αποδοθεί στην κλίμακα.
- Αδυναμία υπολογισμού
- Εκτίμηση

# Μέθοδοι υπολογισμού του συντελεστή αξιοπιστίας

- Χρήση Parallel measurements (distinct but equivalent).
  1. Test-Retest method
  2. Alternate forms
  3. Split-Half
  4. Cronbach alpha

# Test-Retest method

- Οι ερωτώμενοι συμπληρώνουν την κλίμακα σε δύο διαφορετικές χρονικές στιγμές. Η αξιοπιστία της κλίμακας μπορεί να εκτιμηθεί από το συντελεστή συσχέτισης μεταξύ των δύο σκορ.
- Μειονεκτήματα

# Alternate forms

- Παρόμοια μέθοδος με την Test-Retest με τη διαφοροποίηση ότι χρησιμοποιούμε διαφορετικές εκδοχές (versions) της κλίμακας σε κάθε χρονική περίοδο.
- Μειονεκτήματα (Expensive-impractical)



# Split-Half method

- Οι υποερωτήσεις διαιρούνται σε δύο μέρη. Συνήθως μονές-ζυγές και όχι πρώτες τελευταίες.
- Υπολογίζεται το σκορ για κάθε μισό.
- Υπολογίζεται ο συντελεστής συσχέτισης μεταξύ των δύο σκορ.
- Τότε ο εκτιμητής είναι:  
$$(2r/(1+r))$$

# Cronbach alpha

- Οι N υποερωτήσεις της κλίμακας διαμερίζονται σε όλα τα δυνατά μισά. Για κάθε μισό υπολογίζεται το σκορ.
- Συσχετίζονται τα σκορ αυτά.
- Λαμβάνεται ο μέσος όρος αυτών των συσχετίσεων ( $r_{\text{bar}}$ ). Τότε ο εκτιμητής είναι:

$$N * (\text{cov\_}/\text{var\_}) / (1 + (N-1) (\text{cov\_}/\text{var\_}))$$

# Cronbach alpha

- Υπό την υπόθεση ότι οι διακυμάνσεις των item είναι ίσες, προκύπτει ο τυποποιημένος συντελεστής του Cronbach Standardized item alpha

$$\alpha = (N \cdot r_{\text{bar}}) / (1 + (N - 1) \cdot r_{\text{bar}})$$

- Αύξουσα συνάρτηση του N και  $r_{\text{bar}}$

# Τιμές του Cronbach

- Λαμβάνει τιμές στο  $[0,1]$ . Το 0 ερμηνεύεται ως έλλειψη αξιοπιστίας, το 1 ως ισχυρά αξιόπιστη κλίμακα.
- Εξαρτάται από το πλήθος των ερωτήσεων στην κλίμακα.
- Τιμές μεγαλύτερες του 0.7 ικανοποιητικές. Για πολύ σημαντικές μελέτες μεγαλύτερες του 0.9.

# Υλοποίηση στο SPSS

- Analyze
- Scale
- Reliability Analysis...

# Παράδειγμα

## Πηγή: Tutorial of SPSS

- Μια εταιρεία παραγωγής ενός προγράμματος θέλοντας να διαπιστώσει αν πρέπει να συνεχιστεί η προβολή του διεξήγαγε μία έρευνα, απευθύνοντας 7 ερωτήσεις της μορφής: «θα παρακολουθήσετε το πρόγραμμα αν....». Οι απαντήσεις καταγράφονται στο αρχείο `tv_survey.sav`. Είναι οι 7 αυτές ερωτήσεις κατάλληλες για την καταγραφή της επιθυμίας ή όχι του κοινού να συνεχιστεί το εν λόγω τηλεοπτικό πρόγραμμα;

# Models

- Alpha models: Συντελεστής του Cronbach. Για δίτιμα δεδομένα είναι ισοδύναμος με τον Kuder-Richardson 20 συντελεστή.
- Split-half models: Correlation between forms, Guttman split-half reliability, Spearman-Brown reliability (equal and unequal length), and coefficient alpha for each half.

# Models

- Guttman models: Reliability coefficients  $\lambda_1$  through  $\lambda_6$ .
- Parallel and Strict parallel models: Test for goodness of fit of model; estimates of error variance, common variance, and true variance; estimated common inter-item correlation; estimated reliability; and unbiased estimate of reliability.



# Statistics options

- Descriptives for
  1. Item: η μέση τιμή, απόκλιση και το πλήθος των πειραματικών μονάδων για κάθε υποερώτηση.
  2. Scale: μέση τιμή, διακύμανση, τυπική απόκλιση και πλήθος ερωτήσεων για την κλίμακα.
  3. Scale if item is deleted: αν ο συντελεστής του Cronbach βελτιωθεί αν εξαιρεθεί μία υποερώτηση. Δίνεται τότε και η μέση τιμή, η διακύμανση της κλίμακας

# Statistics options

- Corrected item-total correlation: οι συσχετίσεις μεταξύ κάθε υποερώτησης και του συνολικού σκορ της κλίμακας. Σε μία αξιόπιστη κλίμακα θα πρέπει κάθε υποερώτηση να σχετίζεται με το συνολικό σκορ. Αν διαπιστωθεί ότι η συσχέτιση κάποιας υποερώτησης  $< 0,3$  τότε καλό είναι να εξαιρείται από την ανάλυση.

# Statistics

- Inter-item
  1. Correlations: δίνεται ο πίνακας των συσχετίσεων μεταξύ των υποερωτήσεων.
  2. Covariances: δίνεται ο πίνακας των διακυμάνσεων μεταξύ των υποερωτήσεων.

# Summaries

- Δίνονται η μέση τιμή, η ελάχιστη τιμή, η μέγιστη τιμή, το εύρος, το πηλίκο μέγιστης προς ελάχιστη, η διακύμανση και το πλήθος των ερωτήσεων για

Item Means

Item Variances

Inter-Item Covariances

Inter-Item Correlations

# ANOVA TABLE

Διεξάγεται έλεγχος μέσω των τιμών κάθε item.

- F test
- Friedman chi square
- Cochran chi square

# ANOVA with Tukey's Test for Nonadditivity

- Produces a test of the assumption that there is no multiplicative interaction among the items.

# Hotelling's T-Squared Test

- Produces a multivariate test of the null hypothesis that all items on the scale have the same mean.

# Statistics

- Model. Select the model for calculating the intraclass correlation coefficient. Available models are Two-Way Mixed, Two-Way Random, and One-Way Random. Select Two-Way Mixed when people effects are random and the item effects are fixed, select Two-Way Random when people effects and the item effects are random. Select One-Way Random when people effects are random.



# Statistics

- Type. Select the type of index. Available types are Consistency and Absolute Agreement
- Confidence interval. Specify the level for the confidence interval. The default is 95%.
- Test value. Specify the hypothesized value of the coefficient for the hypothesis test. This value is the value to which the observed value is compared. The default value is 0.

# Υποθέσεις για την εφαρμογή τους.

- Data can be dichotomous, ordinal, or interval, but the data should be coded numerically.
- Observations should be independent, and errors should be uncorrelated between items. Each pair of items should have a bivariate normal distribution. Scales should be additive, so that each item is linearly related to the total score.

# Επεξήγηση

- Squared multiple correlation: the observed variability in the response to item which can be explained by the other items.

# Βιβλιογραφία

- Cronbach, L. J. 1951. Coefficient alpha and the internal structure of tests. *Psychometrika*, 16:3, 297-334.
- Ραφτόπουλος και Θεοδοσοπούλου (2002). Μεθοδολογία στάθμισης μιας κλίμακας. *Αρχεία Ελληνικής Ιατρικής*, 19, 577-589.
- DeCoster, J. (2005). Scale construction notes. [www.stat-help.com/notes.html](http://www.stat-help.com/notes.html)

# Βιβλιογραφία

- Coakes and Steed (1999). Spss analysis without anguish. Chapter 15.
- David Garson. Reliability analysis. [www2.chass.ncsu.edu/garson/PA765/reliab.htm](http://www2.chass.ncsu.edu/garson/PA765/reliab.htm)