

Ανάλυση αξιοπιστίας και SPSS

Πρόχειρες σημειώσεις

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.

ΚΛΙΜΑΚΑ

- 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
- 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

- 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

Βασικές ιδιότητες κάθε κλίμακας

- ΕΓΚΥΡΟΤΗΤΑ (VALIDITY)

Η κλίμακα πρέπει να αντιπροσωπεύει επακριβώς τις διαστάσεις που θέλει να εξηγήσει ή να διαμορφώσει σε θεωρητικό επίπεδο. Διακρίνεται σε:

1. Εγκυρότητα εννοιολογικής κατασκευής.

Απαντά στο ερώτημα αν τα δεδομένα υποστηρίζουν τη θεωρία που ερμηνεύει την έννοια. Διακρίνεται μεταξύ άλλων σε: Εγκυρότητα προσώπου (είναι η κλίμακα

κατάλληλη;) και Συγκλίνουσα/ αποκλίνουσα (αν σχετίζεται/δεν σχετίζεται με άλλες που εκτιμούν/δεν εκτιμούν το ίδιο φαινόμενο).

2. 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.

3. Συντρέχουσα εγκυρότητα.

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

4. Εγκυρότητα περιεχομένου.

Οι ερωτήσεις καλύπτουν το εύρος του περιεχομένου της εννοιολογικής κατασκευής.

- **ΑΞΙΟΠΙΣΤΙΑ (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 δηλώνουν έλλειψη αξιοπιστίας. Έστω M, T το παρατηρούμενο και θεωρητικό σκορ στην υπό μελέτη κλίμακα. Τότε:

$$M = T + e,$$

όπου e το σφάλμα. Ο δείκτης αξιοπιστίας ρ ορίζεται από τη σχέση $Var(T)/Var(M)$, δηλαδή παριστάνει το ποσοστό της μεταβλητότητας του παρατηρούμενου σκορ που μπορεί να αποδοθεί στην κλίμακα. Επειδή όπως είναι εύκολα αντιληπτό είναι αδύνατο να υπολογιστεί ακριβώς προχωρούμε σε εκτίμησή του. Είναι διαθέσιμες οι ακόλουθες μεθοδολογίες

1. 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 (αδύνατο πρακτικά, οι συμμετέχοντες είναι υποψιασμένοι)

2. Alternate forms

Παρόμοια μέθοδος με την Test-Retest με τη διαφοροποίηση ότι χρησιμοποιούμε διαφορετικές εκδοχές (versions) της κλίμακας σε κάθε χρονική περίοδο. (αδύνατο πρακτικά, χρονοβόρο δαπανηρό)

3. Split-Half

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.

4. Cronbach Alpha

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

$$\frac{N * \bar{r}}{1 + (N - 1)\bar{r}}$$

Ο δείκτης αυτός 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.

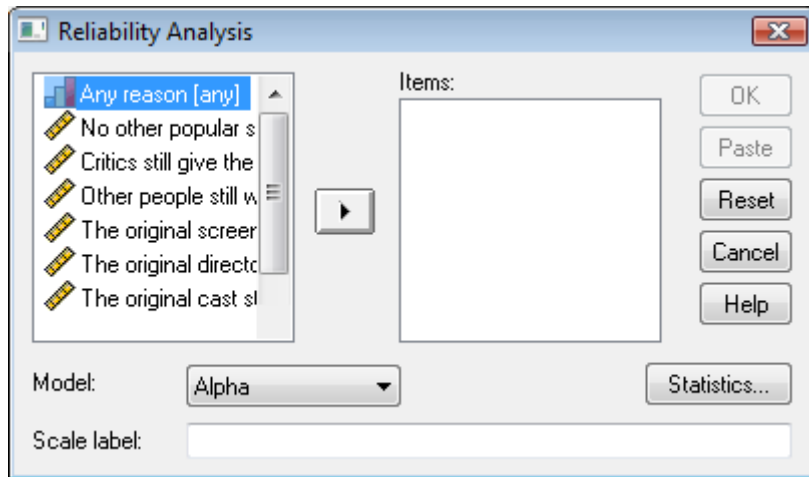
Υλοποίηση στο SPSS

Παράδειγμα (Πηγή Tutorial of SPSS)

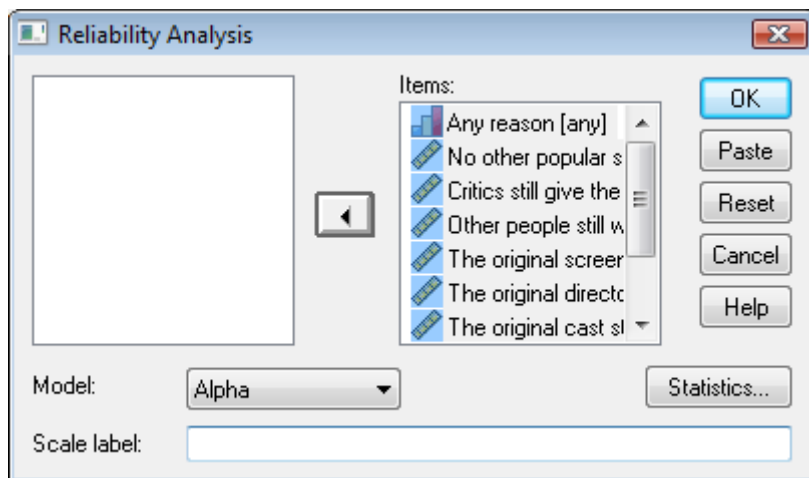
Μια εταιρεία παραγωγής ενός προγράμματος θέλοντας να διαπιστώσει αν πρέπει να συνεχιστεί η προβολή του διεξήγαγε μία έρευνα, απευθύνοντας 7 ερωτήσεις της μορφής: «θα παρακολουθήσετε το πρόγραμμα αν...». Οι απαντήσεις καταγράφονται στο αρχείο tv_survey.sav. Είναι οι 7 αυτές ερωτήσεις κατάλληλες για την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

Από το κεντρικό παράθυρο διαλόγου επιλέγουμε Analyze Scale Reliability Analysis. Στο παράθυρο διαλόγου που προκύπτει στο πλαίσιο Items δηλώνονται οι μεταβλητές που συνθέτουν την υπό μελέτη κλίμακα. Στο πεδίο Model μας δίνεται η δυνατότητα επιλογής της μεθόδου που θα χρησιμοποιηθεί. Περιληπτικά αναφέρουμε ότι επιλέγοντας τη μέθοδο Alpha υπολογίζεται ο συντελεστής του Cronbach, επιλέγοντας τη μέθοδο Split-half υπολογίζονται μεταξύ άλλων οι συντελεστές

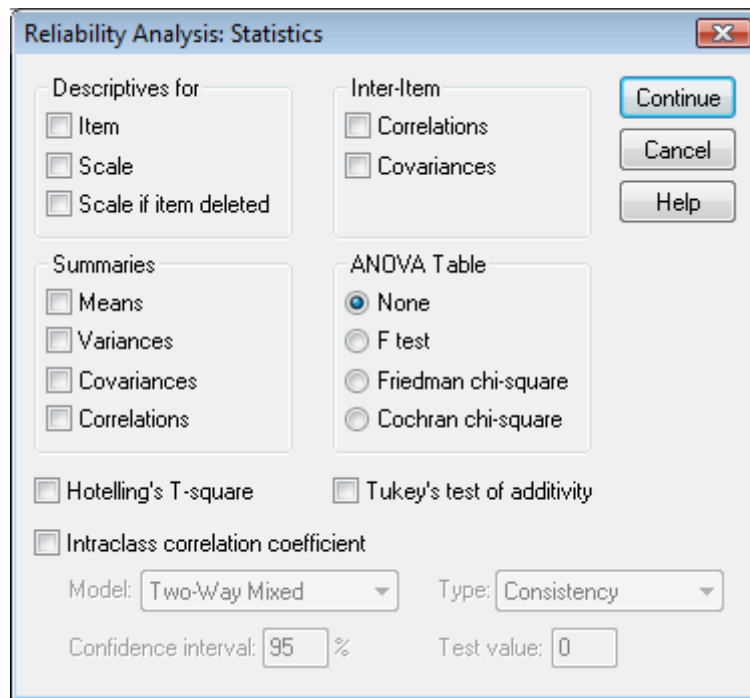
αξιοπιστίας των Guttman split-half reliability και Spearman-Brown καθώς και ο συντελεστής Alpha για κάθε μισό. Τέλος στο πλαίσιο Scale Label δηλώνεται το όνομα της κλίμακας



Επιλέγουμε τα ακόλουθα:



και πατάμε το πεδίο Statistics. Μας δίνονται οι ακόλουθες επιλογές:

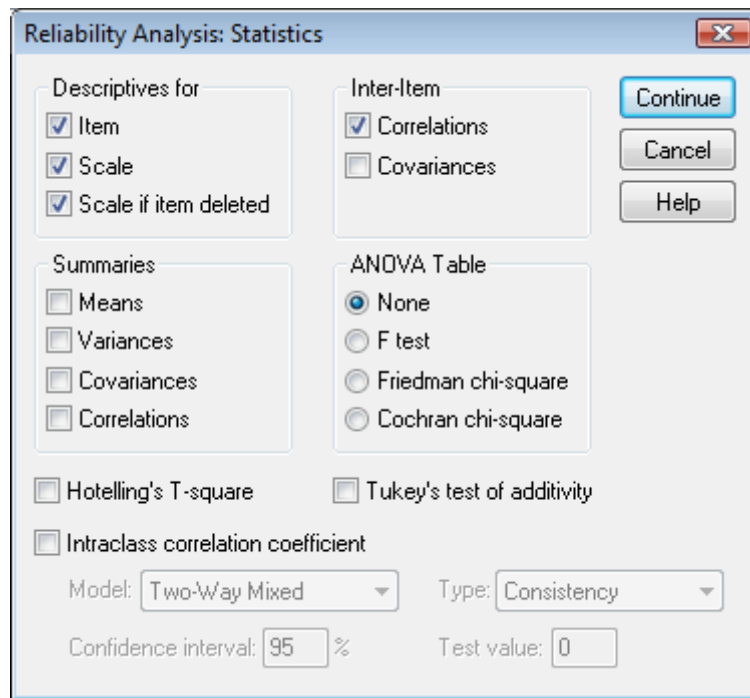


Θα αποφύγουμε να μπούμε σε πολλές λεπτομέρειες καθώς ξεφεύγουν από το σκοπό αυτών των σημειώσεων, που είναι ο τρόπος υλοποίησης της ανάλυσης αξιοπιστίας στα πλαίσια μίας έρευνας στις κοινωνικές, ιατρικές επιστήμες.

Από το πλαίσιο Descriptives δηλώνουμε την εμφάνιση περιγραφικών μέτρων για κάθε υποερώτηση (item), για την κλίμακα συνολικά (Scale) καθώς και για την κλίμακα που προκύπτει αν διαγράφεται η συγκεκριμένη υποερώτηση κάθε φορά (scale if item deleted). Από το πλαίσιο Inter-item επιλέγουμε την εμφάνιση των Correlations (συσχετίσεων) μεταξύ των υποερωτήσεων της κλίμακας.

Από το πλαίσιο Summaries μπορεί κάποιος να ζητήσει την εύρεση των περιγραφικών μέτρων (ελάχιστη, μέγιστη τιμή, μέση τιμή, διακύμανση, εύρος) των μέσων τιμών, των διακυμάνσεων, των συνδιακυμάνσεων και των συσχετίσεων των υποερωτήσεων. Από το πλαίσιο Anova Table διεξάγεται με τρεις διαφορετικές μεθοδολογίες έλεγχος της ισότητας των μέσων τιμών των υποερωτήσεων της κλίμακας, ενώ από το πεδίο Hotelling's T-square διεξάγεται ο γνωστός έλεγχος της Πολυμεταβλητής Στατιστικής για ισότητα μέσων τιμών.

Επιλέγουμε τα ακόλουθα:



Ερμηνεία αποτελεσμάτων

Από τον πίνακα Reliability Statistics έχουμε ότι ο συντελεστής του Cronbach είναι ικανοποιητικός (0.898). Άρα οι 7 αυτές ερωτήσεις συνθέτουν ικανοποιητικά μία κλίμακα.

Reliability Statistics

Cronbach's Alpha	Cronbach's Alpha Based on Standardized Items	N of Items
,898	,894	7

Από τον πίνακα Item Statistics έχουμε τη μέση τιμή και τυπική απόκλιση κάθε υποερώτησης, ενώ στον πίνακα Inter-Item Correlation δίνονται οι συσχετίσεις αυτών. Στον πίνακα Item-Total Statistics δίνεται η μέση τιμή και η διακύμανση αν μία υποερώτηση διαγραφεί από την κλίμακα. Σημαντικές είναι οι τιμές στην στήλη Corrected Item-Total correlation. Αν υπάρχουν τιμές με αρνητικό πρόσημο σημαίνει ότι θέλουν αντιστροφή καθώς είναι αντίθετα διατυπωμένες σε σχέση με αυτό που προσπαθεί να μετρήσει η κλίμακα. Επιπλέον ερωτήσεις με πολύ μικρή τιμή στη στήλη αυτή (ενδεικτικά στη βιβλιογραφία αναφέρεται η τιμή 0.3) αποκλείονται από την κλίμακα. Τέλος από τη στήλη Cronbach's Alpha if item deleted μας δίνονται οι

τιμές του συντελεστή Cronbach όταν διαγράφεται μία υποερώτηση. Αν προκύπτει μία σημαντική θεαματική αύξηση τότε αποκλείεται η συγκεκριμένη υποερώτηση από την υπό κατασκευή κλίμακα.

Item-Total Statistics

	Scale Mean if Item Deleted	Scale Variance if Item Deleted	Corrected Item-Total Correlation	Squared Multiple Correlation	Cronbach's Alpha if Item Deleted
Any reason	4,07	4,171	,792	,740	,871
No other popular shows on at that time	4,05	4,144	,808	,768	,869
Critics still give the show good reviews	4,05	4,113	,827	,770	,867
Other people still watch the show	4,02	4,142	,811	,732	,869
The original screenwriters stay	3,74	4,877	,589	,523	,894
The original directors stay	3,72	4,905	,593	,503	,894
The original cast stays	3,66	5,240	,486	,461	,904

Τέλος από το πλαίσιο μας δίνεται η μέση τιμή η διακύμανση η τυπική απόκλιση της κλίμακας καθώς και το πλήθος των ερωτήσεων που συνθέτουν την κλίμακα. Από την τιμή $4.55/7=0.65$ παρατηρούμε ότι τείνουν οι τηλεθεατές να επιθυμούν τη συνέχιση του τηλεοπτικού προγράμματος (1=συνέχιση 0=διακοπή).

Scale Statistics

Mean	Variance	Std. Deviation	N of Items
4,55	6,040	2,458	7

Ενδεικτική Βιβλιογραφία

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