

# ΣΤΑΤΙΣΤΙΚΗ ΚΑΙ ΘΕΩΡΙΑ ΑΠΟΦΑΣΕΩΝ

ΔΙΑΛΕΞΗ: ΠΟΛΥΜΕΤΑΒΛΗΤΗ ΑΝΑΛΥΣΗ  
ΠΑΡΑΓΟΝΤΙΚΗ ΑΝΑΛΥΣΗ

Διδάσκουσα: Ε. Γάκη, Επίκ. Καθηγήτρια

# Περιεχόμενα Διάλεξης

- Εισαγωγή
- Πολυμεταβλητή Ανάλυση
- Πολυμεταβλητή Περιγραφική Στατιστική
- Παραγοντική Ανάλυση

# Εισαγωγή

- Η **Πολυμεταβλητή Στατιστική Ανάλυση** αναφέρεται σε διαδικασίες και μεθοδολογίες όπου προσπαθούμε να καταλήξουμε σε στατιστική συμπερασματολογία με τη χρήση πολλών μεταβλητών.
- Τα δεδομένα είναι από τη φύση τους πολυμεταβλητά.
- Εξαρτάται από τον ερευνητή κατά πόσο θέλει να χρησιμοποιήσει όλα τα δεδομένα του για να αποκομίσει τη μεγαλύτερη δυνατή πληροφορία από αυτά.
- Οι πολυμεταβλητές τεχνικές αναπτύχθηκαν μαζί με τις μονομεταβλητές τεχνικές.
- Η πολυπλοκότητά τους οδήγησε σε περιορισμούς στην πρακτική εφαρμογή τους. Η γενικευμένη χρήση υπολογιστών στη στατιστική έλυσε το πρόβλημα.

# Πολυμεταβλητή Ανάλυση

Οι πολυμεταβλητές τεχνικές είναι χρήσιμες γιατί:

- Έχουμε περισσότερη πληροφορία.
  - Περισσότερες μεταβλητές ερμηνεύουν καλύτερα το φαινόμενο.
  - Όσο περισσότερη πληροφορία έχει κανείς, τόσο περισσότερο μπορεί να περιορίσει την αβεβαιότητά του.
- Μελετάμε συσχετισμούς
  - Μεταξύ μεταβλητών και μεταξύ υποκειμένων
  - Η ανακάλυψη συσχετίσεων μεταξύ μεταβλητών οδηγεί σε καινούργιες ερμηνείες για τα υπό μελέτη φαινόμενα.

# Πολυμεταβλητή Ανάλυση

Οι πολυμεταβλητές τεχνικές χρησιμοποιούνται για:

- Την εύρεση και ερμηνεία συσχετίσεων μεταξύ των μεταβλητών
- Τη δημιουργία ομάδων είτε από παρατηρήσεις είτε από μεταβλητές σύμφωνα με κάποια χαρακτηριστικά
- Τη μείωση των διαστάσεων του προβλήματος / συμπύκνωση της πληροφορίας που περιέχουν πολλές μεταβλητές σε λιγότερες
- Την πρόβλεψη νέων τιμών
- Τη μοντελοποίηση σε πολλές διαστάσεις, για την ερμηνεία πολλών μεταβλητών σε σχέση με άλλες
- Την ποσοτικοποίηση μη παρατηρήσιμων ποσοτήτων

# Πολυμεταβλητή Ανάλυση

Οι πολυμεταβλητές τεχνικές περιλαμβάνουν:

- Ανάλυση σε Κύριες Συνιστώσες
- Παραγοντική Ανάλυση
- Ανάλυση σε Ομάδες
- Διακριτική Ανάλυση
- Ανάλυση Αντιστοιχιών
- Ανάλυση Κανονικών Συσχετίσεων
- Πολυδιάστατη Κλιμακοποίηση
- Πολυμεταβλητό Γραμμικό Μοντέλο
- Μέθοδοι για Δεδομένα Διεύθυνσης

# Παραγοντική Ανάλυση

# Επιμέρους Ενότητες

- Παραγοντική Ανάλυση
- Μέθοδοι Εκτίμησης
- Περιστροφή
- Σκορ Παραγόντων
- Εφαρμογή Μεθόδου
  - ✓ Επιλογή Αριθμού Παραγόντων
  - ✓ Εκτίμηση Παραμέτρων
  - ✓ Περιστροφή
    - ο (Factor Transformation Matrix)
  - ✓ Χρήση των σκορ (Box Plot)
- Μελέτη Περίπτωσης



# Σκοπός Παραγοντικής Ανάλυσης

- Σκοπός της ανάλυσης είναι η ύπαρξη κοινών παραγόντων ανάμεσα σε μια ομάδα μεταβλητών. Εκφράζοντας αυτούς τους παράγοντες, που δεν είναι υπαρκτή ποσότητα αλλά την κατασκευάζουμε, μπορούμε:
  - Να μειώσουμε τις διαστάσεις του προβλήματος
  - Να δημιουργήσουμε **νέες μεταβλητές**, τους παράγοντες
  - Να εξηγήσουμε τις συσχετίσεις ανάμεσα στα δεδομένα
- Η ΠΑΡΑΓΟΝΤΙΚΗ ΑΝΑΛΥΣΗ ερμηνεύει τη ΔΟΜΗ και όχι τη ΜΕΤΑΒΛΗΤΟΤΗΤΑ
- Στη Παραγοντική Ανάλυση υπάρχει ένα δομημένο μοντέλο και κάποιες υποθέσεις, άρα πρόκειται για μια **στατιστική τεχνική** (ενώ η Ανάλυση σε Κύριες Συνιστώσες είναι μαθηματικός μετασχηματισμός)
- Στη Παραγοντική Ανάλυση εξηγούμε τη **συνδιακύμανση** των μεταβλητών ( ενώ στην ανάλυση σε κύριες συνιστώσες τη διακύμανση)

# Προβλήματα Παραγοντικής Ανάλυσης

- Στηρίζεται σε ένα πλήθος υποθέσεων οι οποίες δεν είναι απαραίτητα ρεαλιστικές
- Δεν έχει μοναδική λύση. Χρησιμοποιώντας διάφορες μεθόδους εκτίμησης, βασισμένοι σε ίδια δεδομένα μπορεί να καταλήξουμε σε διαφορετικά αποτελέσματα.
- Οι παράγοντες οι οποίοι προκύπτουν μπορούν να δεχθούν διαφορετικές ερμηνείες οι οποίες έρχονται σε αντιπαράθεση.
- **Ο αριθμός των παραγόντων** που χρειάζεται να εξάγουμε δεν είναι προφανής και **εξαρτάται από τη μέθοδο εκτίμησης.**

Αποτελεί όμως πολύτιμο εργαλείο καθώς είναι μια μεθοδολογία με την οποία ποσοτικοποιούνται μη παρατηρήσιμες ποσότητες.

# Ορθογώνιο Μοντέλο [1]

- Στο ορθογώνιο μοντέλο της παραγοντικής ανάλυσης που χρησιμοποιούμε, υποθέτουμε ότι οι όποιες **συσχετίσεις** μεταξύ των **μεταβλητών** οφείλονται αποκλειστικά στην ύπαρξη των κοινών παραγόντων τους οποίους δεν ξέρουμε και θέλουμε να εκτιμήσουμε.
- Έστω  $p$  μεταβλητές οι οποίες μπορούν να γραφούν ως γραμμικός συνδυασμός των  $k$  παραγόντων

$$X=LF+\varepsilon$$

$X$ : είναι το διάνυσμα των αρχικών μεταβλητών μεγέθους  $p \times 1$

$L$ : είναι η επιβάρυνση (loading) του παράγοντα  $F_j$  στη μεταβλητή  $X_i$

$F$ : είναι το  $k \times 1$  διάνυσμα με τους παράγοντες

$\varepsilon$ : είναι το σφάλμα ή μοναδικός παράγοντας. Το σφάλμα  $\varepsilon_i$  είναι ο **μοναδικός παράγοντας της  $i$  μεταβλητής** και είναι το μέρος της μεταβλητής το οποίο **δεν** μπορεί να **εξηγηθεί** από τους παράγοντες.

## Ορθογώνιο Μοντέλο [2]

- Είναι προφανές ότι  $k < p$ , δηλαδή ο αριθμός των παραγόντων είναι μικρότερος από τον αριθμό των μεταβλητών
- Κάθε μεταβλητή μπορεί να γραφεί στη μορφή

$$\begin{aligned} X_1 &= L_{11}F_1 + L_{12}F_2 + \dots + L_{1k}F_k + \varepsilon_1 \\ X_2 &= L_{21}F_1 + L_{22}F_2 + \dots + L_{2k}F_k + \varepsilon_2 \\ &\dots \\ X_p &= L_{p1}F_1 + L_{p2}F_2 + \dots + L_{pk}F_k + \varepsilon_p \end{aligned}$$

### Παρατηρήσεις στο ΜΟΝΤΕΛΟ

- Το παραπάνω μοντέλο **μοιάζει** με γραμμικό, έχει όμως διαφορές.
- Τα  $X_i$  δεν είναι παρατηρήσεις αλλά μεταβλητές
- Το δεξί μέρος της εξίσωσης δεν είναι παρατηρήσιμο και άρα πρέπει να εκτιμηθεί

# Ορθογώνιο Μοντέλο [3]

## Παρατηρήσεις στο Μοντέλο (συνέχεια)

- Οι παράγοντες  $F_i$  μπορούν και αυτοί να γραφούν ως γραμμικός συνδυασμός των μεταβλητών. (Αυτό είναι χρήσιμο στην περίπτωση που θέλουμε να φτιάξουμε νέες μεταβλητές. Οι ΣΥΝΤΕΛΕΣΤΕΣ αυτοί όμως διαφέρουν από τις επιβαρύνσεις)
  - Οι συντελεστές κάθε παράγοντα όταν *εκφράζουμε τις μεταβλητές* ως γραμμικό συνδυασμό των *παραγόντων* καλούνται **ΕΠΙΒΑΡΥΝΣΕΙΣ**
  - Οι *συντελεστές κάθε μεταβλητής* όταν *εκφράζουμε κάθε παράγοντα* ως γραμμικό συνδυασμό των *μεταβλητών* καλούνται *συντελεστές των σκορ* (**FACTOR SCORE COEFFICIENTS**)
- Οι παράγοντες έχουν όλοι την **ίδια διακύμανση**.
- Στην παραγοντική ανάλυση το μοντέλο προσπαθεί να *εκφράσει τις μεταβλητές* ως γραμμικό συνδυασμό των παραγόντων.

# Υποθέσεις Ορθογωνίου Μοντέλου

- Τα δεδομένα προέρχονται από πολυμεταβλητούς κανονικούς πληθυσμούς
- Ο πίνακας διακύμανσης μπορεί να διασπαστεί σε 2 μέρη:
  - ✓ το ένα κομμάτι που ερμηνεύουν οι **κοινοί παράγοντες** και ονομάζεται **εταιρικήτητα (communality)** και
  - ✓ το δεύτερο κομμάτι που οφείλεται στους **μοναδικούς παράγοντες** και άρα το μοντέλο δεν μπορεί να ερμηνεύσει και ονομάζεται **ιδιαιτερότητα (specificity)**
- Οι παράγοντες και οι μοναδικοί παράγοντες έχουν μηδενικές μέσες τιμές
- Οι παράγοντες και οι μοναδικοί παράγοντες είναι ασυσχέτιστοι μεταξύ τους
- Οι μοναδικοί παράγοντες και οι κοινοί παράγοντες είναι ασυσχέτιστοι
- Οι διακυμάνσεις των παραγόντων είναι ίσες με την μονάδα, άρα όλοι οι παράγοντες έχουν την ίδια διακύμανση.

# Βήματα Παραγοντικής Ανάλυσης

- Έλεγχος αν υπάρχουν **συσχετίσεις ικανοποιητικές** για να κάνουμε παραγοντική ανάλυση
- Εύρεση του **αριθμού των παραγόντων** και **εκτίμηση** των παραμέτρων του μοντέλου
- **Περιστροφή** του μοντέλου με σκοπό να αυξήσουμε την ερμηνευτική του ικανότητα
- **Εκτίμηση των σκορ** των παραγόντων για περαιτέρω στατιστική χρήση.

# Τι πρέπει να προσέξουμε

- Οι μεταβλητές να έχουν υψηλές συσχετίσεις
  - $r > .3$ .
- Να μην ταυτίζονται απόλυτα κάποιες μεταβλητές :
  - some variables perfectly correlated,  $r = 1$ .
- Κοιτάζουμε το correlation matrix



**ΘΥΜΗΘΕΙΤΕ:** Αν υπάρχει κάποια ή κάποιες μεταβλητές που είναι ασυσχέτιστες με τις υπόλοιπες, καλό θα ήταν να τις αγνοήσουμε, αφού θα προκύψουν μόνες τους ως ξεχωριστός παράγοντας.



# Καταλληλότητα μεταβλητών & δεδομένων

## Kaiser-Meyer-Olkin

$$KMO = \frac{\sum_{i \neq j} \sum r_{ij}^2}{\sum_{i \neq j} \sum r_{ij}^2 + \sum_{i \neq j} \sum a_{ij}^2}$$

$r_{ij}, a_{ij}$  : δειγματικοί συντελεστές συσχέτισης και μερικής συσχέτισης

Αν ΚΜΟ είναι **μεγάλη**, τότε τα δεδομένα είναι **κατάλληλα** για παραγοντική ανάλυση. Τιμές κάτω από 0.5 είναι πολύ κακές τιμές. Στην πράξη τιμές γύρω στο 0.5 είναι σχετικά καλές τιμές για να συνεχίσουμε. Ιδανικά το  $KMO > 0,7$

## Measure of Sampling Adequacy

Πρόκειται για μέτρο δειγματικής καταλληλότητας.

$$MSA = \frac{\sum_j \sum r_{ij}^2}{\sum_j \sum r_{ij}^2 + \sum_j \sum a_{ij}^2}$$

Τιμές κοντά στο 1 είναι ενδείξεις ότι η μεταβλητή είναι πολύ καλή για να τη χρησιμοποιήσουμε στην ανάλυση.

Τέλος το ΚΜΟ αφορά όλα τα δεδομένα, ενώ το MSA κάθε μια μεταβλητή χωριστά

# Αριθμός και Εκτίμηση Παραγόντων

Η Απόφαση για τον αριθμό των παραγόντων που θα επιλεγούν γίνεται με ορισμένα κριτήρια :

- Ποσοστό συνολικής διακύμανσης που εξηγούν οι συνιστώσες
- Κριτήριο του Kaiser
- Ποσοστό της διακύμανσης των αρχικών μεταβλητών που ερμηνεύεται
- Scree plot

# Αριθμός και Εκτίμηση Παραγόντων

- Ποσοστό συνολικής διακύμανσης που εξηγούν οι συνιστώσες

Θέτουμε ένα όριο και διαλέγουμε τόσες συνιστώσες ώστε αθροιστικά να εξηγούν μεγαλύτερο ποσοστό από το στόχο που βάλαμε.

# Αριθμός και Εκτίμηση Παραγόντων

- Κριτήριο του Kaiser

Επιλέγουμε τόσες ιδιοτιμές όσες είναι μεγαλύτερες από τη μέση τιμή των ιδιοτιμών, δηλαδή

$$\bar{\lambda} = \sum_{j=1}^p \lambda_j / p$$

Για την περίπτωση που δουλεύουμε με πίνακα συσχετίσεων διαλέγουμε τόσες συνιστώσες όσες και οι ιδιοτιμές μεγαλύτερες της μονάδας, καθώς ισχύει

$$\bar{\lambda} = 1$$

# Αριθμός και Εκτίμηση Παραγόντων

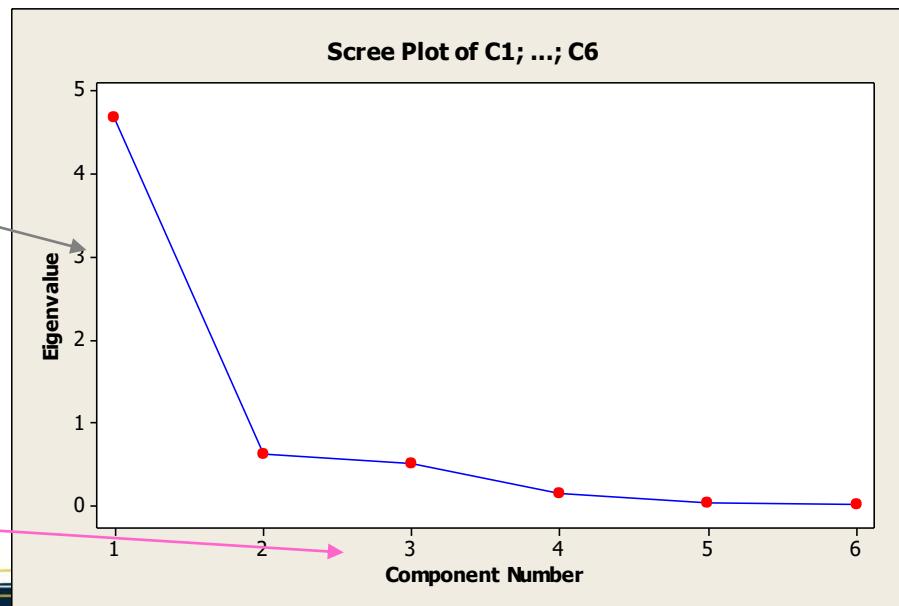
- Ποσοστό της διακύμανσης των αρχικών μεταβλητών που ερμηνεύεται

Διαλέγουμε τόσες συνιστώσες ώστε να ερμηνεύεται για κάθε μεταβλητή ένα υψηλό ποσοστό. Το ποσοστό είναι υποκειμενικό.

# Αριθμός και Εκτίμηση Παραγόντων

- Scree plot

Παίρνουμε τόσες συνιστώσες μέχρι το γράφημα να αρχίσει να γίνεται περίπου επίπεδο. Είναι υποκειμενικό κριτήριο και αρκετά δύσκολο να ερμηνευτεί το γράφημα.



Ιδιοτιμή

Σειρά

# Τι είναι οι ιδιοτιμές & τα ιδιοδιανύσματα??

Ας θεωρήσουμε τη γραμμική απεικόνιση  $f : X \rightarrow X$  όπου  $X$  είναι ένας διανυσματικός χώρος πάνω στο  $F (R, C)$ .

Εάν υπάρχουν μη μηδενικά διανύσματα  $v \in X$  τα οποία έχουν την ιδιότητα  $f(v) = \lambda v$ , όπου  $\lambda \in K$ ,

- τα διανύσματα αυτά ονομάζονται **ιδιοδιανύσματα** της γραμμικής απεικόνισης  $f$ ,
- ενώ οι τιμές του  $\lambda$  για τις οποίες ισχύει η σχέση αυτή ονομάζονται **ιδιοτιμές** της γραμμικής απεικόνισης.

# Αριθμός και Εκτίμηση Παραγόντων

- Ο αριθμός των παραγόντων πρέπει να καθορισθεί πριν γίνει η εκτίμησή τους. Άρα κάποιος πρέπει να δουλέψει με αυξανόμενο αριθμό παραγόντων και να κρατήσει το μοντέλο με βάση κάποιο **κριτήριο καταλληλότητας**.
- Τέτοια κριτήρια είναι:
  - Από τον πίνακα επιβαρύνσεων μπορεί κάποιος να εκτιμήσει τον πίνακα  $\Sigma$ . Οι αποκλίσεις του πραγματικού πίνακα με τον εκτιμημένο (**Reproduced matrix**) θα πρέπει να είναι μικρές.
  - Έλεγχος λόγου πιθανοφανειών αν οι εκτιμήσεις έχουν γίνει με τη μέθοδο της **μέγιστης πιθανοφάνειας**. Τέτοιοι έλεγχοι στηρίζονται στις υποθέσεις για την κατανομή του πληθυσμού.
- Η ερμηνεία των παραγόντων εξαρτάται από τον αριθμό τους
- Σε μερικές μεθόδους υπάρχει περιορισμός των παραγόντων που μπορούν να εκτιμηθούν.



# Βασικές Μέθοδοι Εκτίμησης

## Μέθοδος Κυρίων Συνιστωσών

- Όταν προσθέτουμε παράγοντες **δεν αλλάζουν οι επιβαρύνσεις** των παραγόντων
- Η μέθοδος **εξαρτάται από τις μονάδες μέτρησης** και πρέπει να διαλέξουμε ανάμεσα στον πίνακα διακύμανσης και συσχέτισης
- **Δεν βάζει περιορισμούς** στον αριθμό των παραγόντων
- Η μέθοδος είναι ένας μαθηματικός σχεδιασμός και **δουλεύει πάντα**
- Τα **σκορ** των παραγόντων **υπολογίζονται**

## Μέθοδος Λόγου Πιθανοφανειών

- Όταν προσθέτουμε παράγοντες αλλάζουν οι επιβαρύνσεις και η ερμηνεία τους
- Μπορούμε να κάνουμε ελέγχους καλής προσαρμογής βασισμένοι στον κλασικό έλεγχο του λόγου των πιθανοφανειών
- Όταν η μέθοδος δεν δουλεύει τότε υπάρχει πρόβλημα με το μοντέλο
- Τα σκορ των παραγόντων **δεν μπορούν να υπολογιστούν**

# Περιστροφή

- Με την περιστροφή των παραγόντων προσπαθούμε να κάνουμε τους παράγοντες πιο ερμηνεύσιμους.
- Με την περιστροφή **δεν αλλάζουν** χαρακτηριστικά όπως:
  - Η καλή προσαρμοστικότητα
  - Το ποσό της διακύμανσης συνδιακύμανσης που ερμηνεύει το μοντέλο
- **Αλλάζουν οι τιμές των επιβαρύνσεων**
  - Αν  $L$  είναι ένας πίνακας που περιέχει τις επιβαρύνσεις και  $G$  ένας ορθογώνιος πίνακας ( $G'G=I$ ) τότε ισχύει πως
$$LG(LG') = LGG'L' = LL'$$
και άρα ο πίνακας  $G$  ορίζει έναν ορθογώνιο μετασχηματισμό.

# Περιστροφή

## Βασικές Μέθοδοι Περιστροφής:

### *Varimax*

Ελαχιστοποιεί τον αριθμό των μεταβλητών που έχουν μεγάλες επιβαρύνσεις για κάθε παράγοντα

### *Quatrimax*

Ελαχιστοποιεί τον αριθμό των παραγόντων που εξηγούν μια μεταβλητή

### *Equimax*

Συνδυασμός Varimax και Quatrimax

### *Oblique*

Μη ορθογώνια περιστροφή. Οι άξονες που προκύπτουν δεν είναι ορθογώνιοι και άρα οι παράγοντες δεν είναι ανεξάρτητοι.

# Σκορ Παραγόντων

Βασικός σκοπός είναι η μείωση του αριθμού των μεταβλητών. Για να επιτευχθεί αυτό δημιουργούμε **νέες μεταβλητές**, τους παράγοντες, γραμμικούς συνδυασμούς των αρχικών μεταβλητών

Κάθε **παράγοντας** μπορεί να γραφεί με τη μορφή

$$\begin{aligned} & \parallel \text{---} \text{---} \text{---} \text{---} \text{---} \text{---} \text{---} \text{---} \text{---} \text{---} \text{---} \text{---} \parallel \\ & \parallel \quad F_1 = a_{11}X_1 + a_{12}X_2 + \dots + a_{1p}X_p \quad \parallel \\ & \parallel \quad F_2 = a_{21}X_1 + a_{22}X_2 + \dots + a_{2p}X_p \quad \parallel \\ & \dots \parallel \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \quad \parallel \\ & \parallel \quad F_k = a_{k1}X_1 + a_{k2}X_2 + \dots + a_{kp}X_k \quad \parallel \\ & \parallel \text{---} \text{---} \text{---} \text{---} \text{---} \text{---} \text{---} \text{---} \text{---} \text{---} \text{---} \parallel \end{aligned}$$

Οι συντελεστές  $a_{ij}$  είναι τα σκορ της μεταβλητής  $X_j$  στον παράγοντα  $F_i$

**ΠΡΟΣΟΧΗ: ΔΕΝ ΕΙΝΑΙ ΤΟ ΙΔΙΟ ΜΕ ΤΙΣ ΕΠΙΒΑΡΥΝΣΕΙΣ.**

# Μελέτη Περίπτωσης

The SPSS Anxiety Questionnaire (SAQ)		SD	D	N	A	SA
1.	Statistics makes me cry	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>
2.	My friends will think I'm stupid for not being able to cope with SPSS	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>
3.	Standard deviations excite me	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>
4.	I dream that Pearson is attacking me with correlation coefficients	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>
5.	I don't understand statistics	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>
6.	I have little experience of computers	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>
7.	All computers hate me	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>
8.	I have never been good at mathematics	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>
9.	My friends are better at statistics than me	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>
10.	Computers are useful only for playing games	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>
11.	I did badly at mathematics at school	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>
12.	People try to tell you that SPSS makes statistics easier to understand but it doesn't	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>
13.	I worry that I will cause irreparable damage because of my incompetence with computers	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>
14.	Computers have minds of their own and deliberately go wrong whenever I use them	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>
15.	Computers are out to get me	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>
16.	I weep openly at the mention of central tendency	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>
17.	I slip into a coma whenever I see an equation	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>
18.	SPSS always crashes when I try to use it	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>
19.	Everybody looks at me when I use SPSS	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>
20.	I can't sleep for thoughts of eigenvectors	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>
21.	I wake up under my duvet thinking that I am trapped under a normal distribution	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>
22.	My friends are better at SPSS than I am	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>
23.	If I am good at statistics people will think I am a nerd	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>

Από το βιβλίο του Andy Field, *Discovering Statistics Using IBM SPSS Statistics, 4<sup>th</sup> edition*, Sage Publications, 2013

# Μελέτη Περίπτωσης

Σκοπός της Παραγοντικής Ανάλυσης είναι να δούμε αν και κατά πόσο τα δεδομένα μπορούν να ερμηνευτούν με τη χρήση του ορθογώνιου παραγοντικού μοντέλου και κατά πόσο μπορούμε να ερμηνεύσουμε τους παράγοντες που προκύπτουν.

## Μορφή Παραγοντικού Μοντέλου

Statistics make me cry =  $L_{11}F_1 + L_{12}F_2 + \dots + L_{1k}F_k + \varepsilon_1$

My friends will think ... =  $L_{21}F_1 + L_{22}F_2 + \dots + L_{2k}F_k + \varepsilon_2$

....

If I am good ..... =  $L_{71}F_1 + L_{72}F_2 + \dots + L_{7k}F_k + \varepsilon_7$

# Μελέτη Περίπτωσης

SAQ.sav [DataSet1] - IBM SPSS Statistics Data Editor

File Edit View Data Transform Analyze Direct Marketing Graphs Utilities Add-ons Window Help

Reports  
Descriptive Statistics  
Custom Tables  
Compare Means  
General Linear Model  
Generalized Linear Models  
Mixed Models  
Correlate  
Regression  
Loglinear  
Classify  
Dimension Reduction  
Scale  
Nonparametric Tests  
Forecasting  
Survival  
Multiple Response  
Simulation...  
Quality Control  
ROC Curve...

	Question_01	Question_02	Question_06	Question_07	Question_08
1	2	1	2	3	2
2	1	1	2	2	2
3	2	3	1	2	2
4	3	1	3	4	2
5	2	1	3	3	2
6	2	1	4	4	2
7	2	3	2	2	2
8	2	2	2	2	2
9	3	3	2	2	2
10	2	4	3	3	2
11	2	1	2	3	2
12	2	1	3	3	2
13	3	1	2	3	2
14	2	2	2	3	2
15	2	2	2	3	2
16	3	1	2	2	2
17	1	2	1	1	2

Factor...

Factor Analysis

Variables:

- I don't understand...
- I have little experi...
- All computers hat...
- I have never been...
- My friends are bet...
- Computers are u...
- I did badly at mat...
- People try to tell y...
- I worry that I will c...
- Computers have ...
- Computers are o...

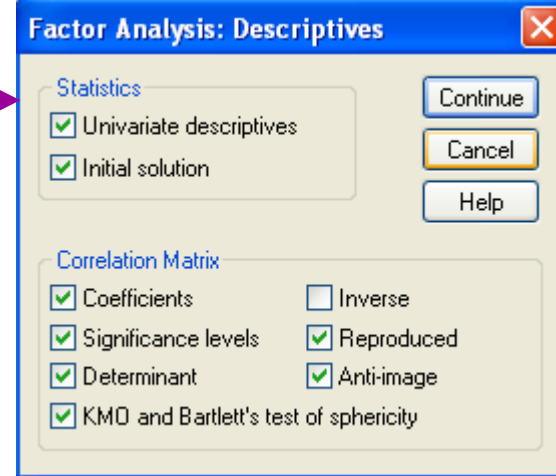
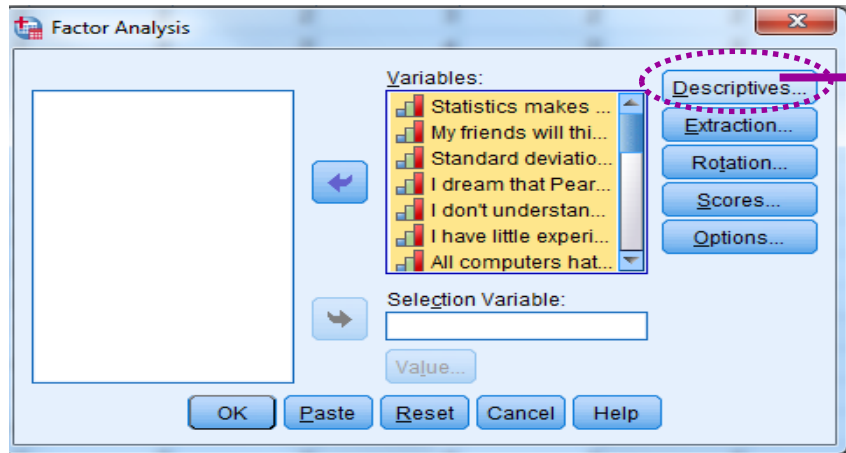
Selection Variable:

Value...

Descriptives...  
Extraction...  
Rotation...  
Scores...  
Options...

OK Paste Reset Cancel Help

# Μελέτη Περίπτωσης



Εμφανίζει περιγραφικά μέτρα, χρήσιμα για την ανάλυση

Univariate descriptives: Διάφορα περιγραφικά στατιστικά για κάθε μεταβλητή

Initial Solution: Η αρχική λύση

Coefficients: Ο πίνακας συσχετίσεων

Significance levels: Ο πίνακας με τη στατιστική σημαντικότητα κάθε συσχέτισης

Determinant: Η ορίζουσα του πίνακα συσχετίσεων

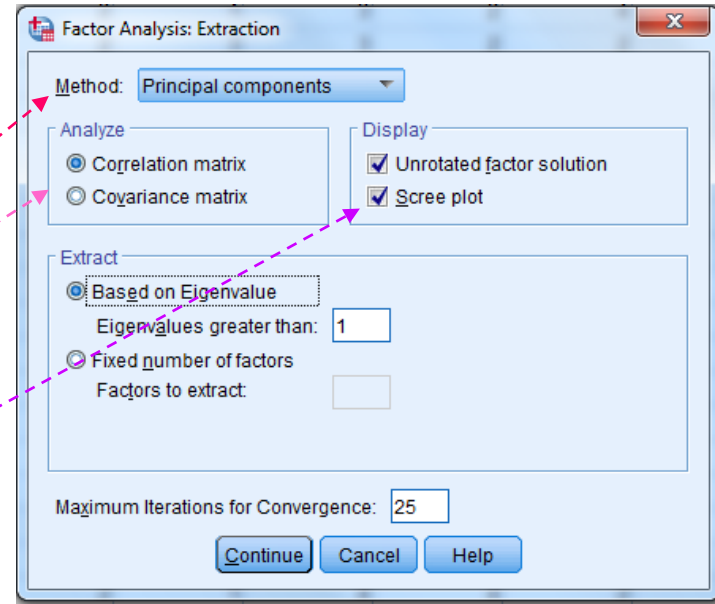
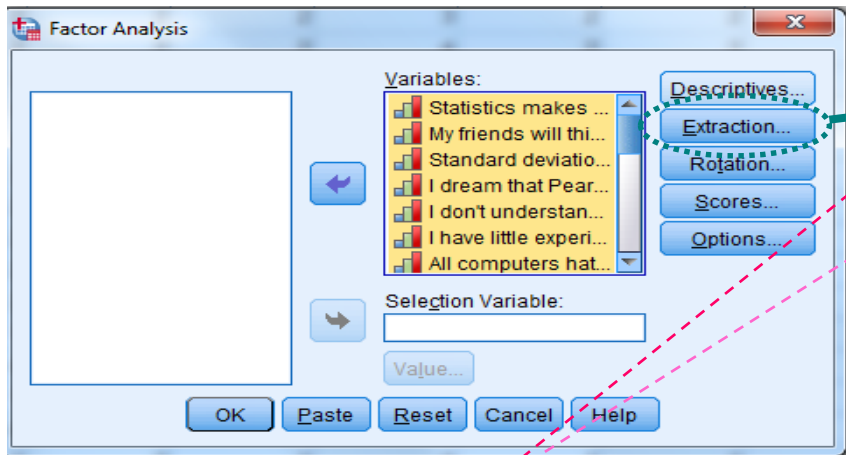
KMO and Bartlett's test of sphericity: Ο έλεγχος σφαιρικότητας του Bartlett και η Kaiser-Meyer-Olkin στατιστική συνάρτηση για την καταλληλότητα των δεδομένων

Reproduced: Ο εκτιμημένος πίνακας συσχετίσεων σύμφωνα με το μοντέλο. Τα διαγώνια στοιχεία είναι οι εταιρικότητες ενώ τα στοιχεία κάτω από τη διαγώνιο είναι η διαφορά της εκτιμημένης συσχέτισης από την πραγματική.

Anti-image: Περιέχει τις αρνητικές τιμές του πίνακα μερικών συσχετίσεων, ενώ τα διαγώνια στοιχεία είναι τα MSA των μεταβλητών.



# Μελέτη Περίπτωσης



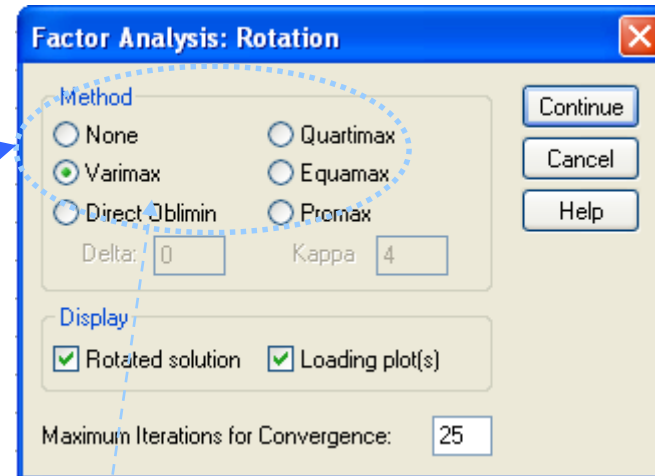
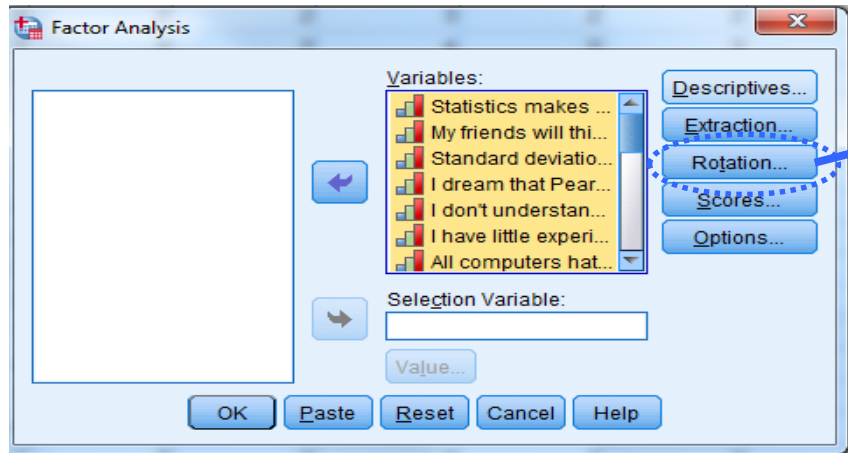
Στην επιλογή **Method** διαλέγουμε τη μέθοδο εκτίμησης που θα χρησιμοποιηθεί.

**Διαλέγουμε** επίσης τον πίνακα διακυμάνσεων ή συσχετίσεων (στη μέθοδο μέγιστης πιθανοφάνειας είναι αδιάφορο).

Στην επιλογή **Display** διαλέγουμε να εμφανίσει την λύση πριν την περιστροφή καθώς και το Scree plot.

**Παρατήρηση:** Για να τρέχουμε την παραγοντική ανάλυση θα πρέπει να ξέρουμε τον αριθμό των παραγόντων. Έτσι χρειάζεται να τρέξουμε μια ανάλυση ώστε να βρούμε τα στοιχεία που θα μας βοηθήσουν να βρούμε τον αριθμό των παραγόντων (είτε να διαλέξουμε απευθείας τον αριθμό είτε να διαλέξουμε μια τιμή και έτσι να πάρουμε τόσους παράγοντες όσες και οι ιδιοτιμές του πίνακα που χρησιμοποιούμε που είναι μεγαλύτερες από τη μέση τιμή όλων των ιδιοτιμών)

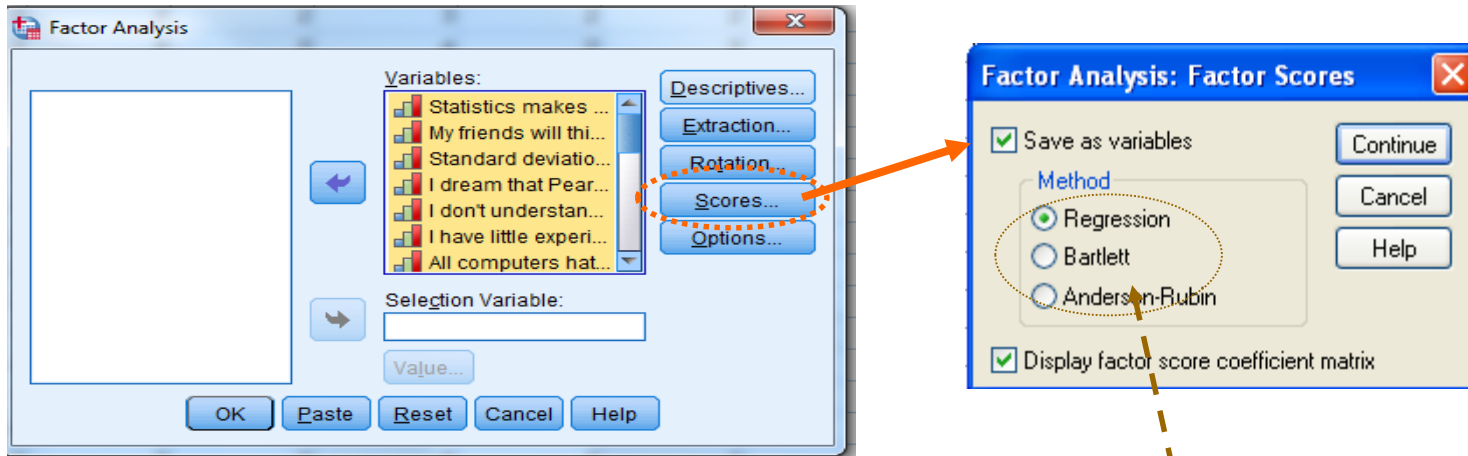
# Μελέτη Περίπτωσης



Επιλέγουμε τη μέθοδο περιστροφής.

Επιλέγοντας Loading Plot(s) θα πάρουμε τα γραφήματα των παραγόντων που έχουμε επιλέξει.

# Μελέτη Περίπτωσης

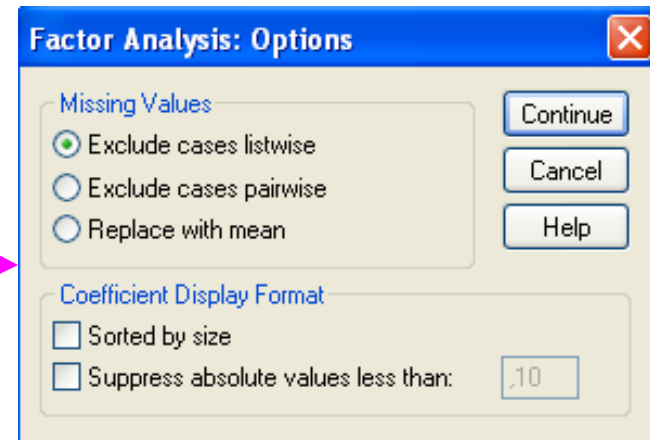
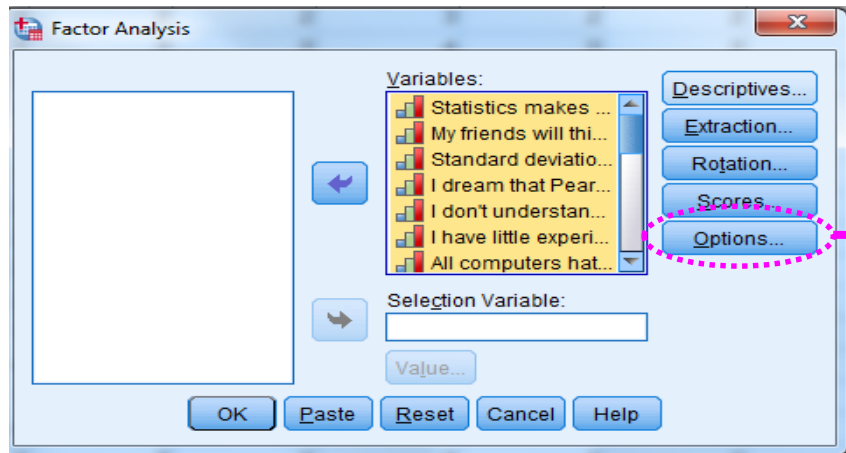


Επιλέγοντας το Save as variables θα δημιουργηθούν τόσες μεταβλητές όσοι και οι παράγοντες που επιλέξαμε. Με τη μέθοδο των κύριων συνιστωσών τα factor scores μπορούν να υπολογιστούν με ακρίβεια. Διαφορετικά θα πρέπει να επιλέξουμε μία από τις **τρεις μεθόδους**.

Το Display factor score coefficient matrix εμφανίζει τον πίνακα με τους συντελεστές.

**ΠΡΟΣΟΧΗ:** Ο πίνακας αυτός περιέχει τους συντελεστές με τους οποίους μπορούμε να εκφράσουμε έναν παράγοντα ως γραμμικό συνδυασμό των μεταβλητών.

# Μελέτη Περίπτωσης



Επιλέγουμε τον τρόπο που θα χειριστούμε τα missing values και τον τρόπο που θέλουμε να εμφανίζονται τα αποτελέσματα των εκτιμημένων παραγόντων.

# Καταλληλότητα Δεδομένων

Descriptive Statistics

	Mean	Std. Deviation	Analysis N
Statistics makes me cry	2,37	,828	2571
My friends will think I'm stupid for not being able to cope with SPSS	1,62	,851	2571
Standard deviations excite me	2,59	1,075	2571
I dream that Pearson is attacking me with correlation coefficients	2,79	,949	2571
I don't understand statistics	2,72	,965	2571
I have little experience of computers	2,23	1,122	2571
All computers hate me	2,92	1,102	2571
I have never been good at mathematics	2,24	,873	2571
My friends are better at statistics than me	2,85	1,263	2571
Computers are useful only for playing games	2,28	,877	2571
I did badly at mathematics at school	2,26	,881	2571
People try to tell you that SPSS makes statistics easier to understand but it doesn't	3,16	,916	2571
I worry that I will cause irreparable damage because of my incompetence with computers	2,45	,949	2571
Computers have minds of their own and deliberately go wrong whenever I use them	2,88	,999	2571
Computers are out to get me	2,77	1,009	2571
I weep openly at the mention of central tendency	2,88	,916	2571
I slip into a coma whenever I see an equation	2,47	,884	2571
SPSS always crashes when I try to use it	2,57	1,053	2571
Everybody looks at me when I use SPSS	2,29	1,101	2571
I can't sleep for thoughts of eigenvectors	3,62	1,036	2571
I wake up under my duvet thinking that I am trapped under a normal distribution	3,17	,985	2571
My friends are better at SPSS than I am	2,89	1,041	2571
If I'm good at statistics my friends will think I'm a nerd	3,43	1,044	2571

Υπάρχει διαφορά στις τιμές της τυπικής απόκλισης (και κατ' επέκταση της διακύμανσης) άρα δεν φαίνεται λογικό να προχωρήσει κανείς σε ανάλυση με πίνακα διακυμάνσεων.



# Καταλληλότητα Δεδομένων

KMO and Bartlett's Test

Kaiser-Meyer-Olkin Measure of Sampling Adequacy.		,930
Bartlett's Test of Sphericity	Approx. Chi-Square	19334,492
	df	253
	Sig.	,000

Το στατιστικό **Kaiser-Meyer-Olkin** είναι πολύ υψηλό. Θυμίζουμε ότι υψηλές τιμές KMO δείχνουν ότι τα δεδομένα είναι κατάλληλα για παραγοντική ανάλυση.

Ο έλεγχος σφαιρικότητας του **Bartlett** απορρίπτει τη μηδενική υπόθεση ότι ο πίνακας συσχέτισης είναι ο μοναδιαίος.

# Καταλληλότητα Δεδομένων

Όλα τα στοιχεία δείχνουν ότι τα δεδομένα είναι κατάλληλα για να προχωρήσουμε σε Παραγοντική Ανάλυση.

Είναι όμως όλες οι μεταβλητές κατάλληλες για να χρησιμοποιηθούν στο μοντέλο?

Για να το εξετάσουμε αυτό χρησιμοποιούμε την τιμή **MSA** (Μέτρο της Δειγματικής Καταλληλότητας). Θυμίζουμε ότι τιμές κοντά στο 1 είναι ενδείξεις ότι η μεταβλητή είναι πολύ καλή για να χρησιμοποιηθεί στην ανάλυση.





# Καταλληλότητα Δεδομένων

nerd

Anti-image Correlation	Statistics makes me cry	,930 <sup>a</sup>	-,020	,053	-,167	-,156	,020	,023	-,049	-,016	-,
	My friends will think I'm stupid for not being able to cope with SPSS	-,020	,875 <sup>a</sup>	-,157	-,041	,010	-,053	,016	-,033	-,193	-,
	Standard deviations excite me	,053	-,157	,951 <sup>a</sup>	,084	,037	-,042	,072	-,007	-,142	-,
	I dream that Pearson is attacking me with correlation coefficients	-,167	-,041	,084	,955 <sup>a</sup>	-,134	-,007	-,087	-,075	,030	,
	I don't understand statistics	-,156	,010	,037	-,134	,960 <sup>a</sup>	-,035	-,044	-,027	-,020	-,
	I have little experience of computers	,020	-,053	-,042	-,007	-,035	,891 <sup>a</sup>	-,275	,024	,011	-,
	All computers hate me	,023	,016	,072	-,087	-,044	-,275	,942 <sup>a</sup>	-,015	-,030	-,
	I have never been good at mathematics	-,049	-,033	-,007	-,075	-,027	,024	-,015	,871 <sup>a</sup>	-,099	,
	My friends are better at statistics than me	-,016	-,193	-,142	,030	-,020	,011	-,030	-,099	,834 <sup>a</sup>	,
	Computers are useful only for playing games	-,012	-,012	-,016	,006	-,093	-,116	-,033	,051	,043	,9
	I did badly at mathematics at school	-,041	,038	,064	-,022	-3,269E-5	-,084	,045	-,412	,037	-,
	People try to tell you that SPSS makes statistics easier to understand but it doesn't	-,007	,031	,087	-,154	-,058	,045	-,043	,033	-,003	-,
	I worry that I will cause irreparable damage because of my incompetence with	-,005	-,008	-,032	,023	,004	-,164	-,039	,002	,061	-,

Οι τιμές MSA για κάθε μεταβλητή είναι σχετικά υψηλές. Όλες οι τιμές κρίνονται ικανοποιητικές και οι μεταβλητές είναι κατάλληλες για την ανάλυση.

Στη μη διαγώνιο βρίσκονται οι τιμές του συντελεστή μερικής συσχέτισης των δύο μεταβλητών, όταν εξουδετερώσουμε την επίδραση των υπόλοιπων με αντίστροφο όμως πρόσημο.

# Επιλογή Αριθμού Παραγόντων

Total Variance Explained

Component	Initial Eigenvalues			Extraction Sums of Squared Loadings			Rotation Sums of Squared Loadings		
	Total	% of Variance	Cumulative %	Total	% of Variance	Cumulative %	Total	% of Variance	Cumulative %
1	7,290	31,696	31,696	7,290	31,696	31,696	3,730	16,219	16,219
2	1,739	7,560	39,256	1,739	7,560	39,256	3,340	14,523	30,742
3	1,317	5,725	44,981	1,317	5,725	44,981	2,553	11,099	41,841
4	1,227	5,336	50,317	1,227	5,336	50,317	1,950	8,476	50,317
5	,988	4,295	54,612						
6	,895	3,893	58,504						
7	,806	3,502	62,007						
8	,783	3,404	65,410						
9	,751	3,265	68,676						
10	,717	3,117	71,793						
11	,684	2,972	74,765						
12	,670	2,911	77,676						
13	,612	2,661	80,337						
14	,578	2,512	82,849						
15	,549	2,388	85,236						
16	,523	2,275	87,511						
17	,508	2,210	89,721						
18	,456	1,982	91,704						
19	,424	1,843	93,546						
20	,408	1,773	95,319						
21	,379	1,650	96,969						
22	,364	1,583	98,552						
23	,333	1,448	100,000						

Extraction Method: Principal Component Analysis.

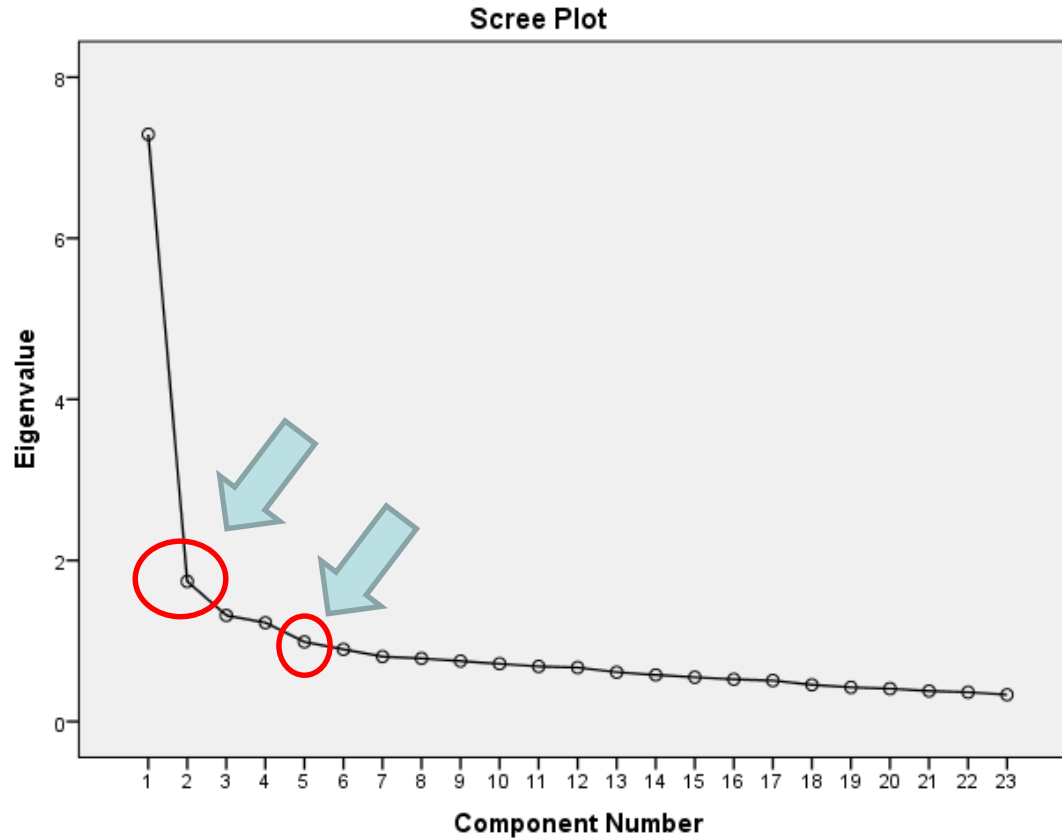


Η επιλογή του αριθμού των παραγόντων εξαρτάται από τη μέθοδο εκτίμησης.

Όμως, κριτήρια βασισμένα στις ιδιοτιμές χρησιμοποιούνται συχνά στην πράξη άσχετα από τη μέθοδο εκτίμησης (Κανόνας του Kaiser, ποσοστό της διακύμανσης που εξηγείται ή scree plot).

Τέσσερις παράγοντες εξηγούν περίπου το 50% της διακύμανσης.

# Επιλογή Αριθμού Παραγόντων



Από το scree plot φαίνεται ότι δύο παράγοντες θα πρέπει να εκτιμηθούν!!!!!!

Ή μήπως  
5???????????????

# Εκτίμηση Παραμέτρων και περιστροφή

Για την Εκτίμηση των Παραμέτρων θα χρησιμοποιήσουμε:

➤ Τη Μέθοδο των Κύριων Συνιστωσών

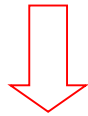
Η περιστροφή είναι ανεξάρτητη της μεθόδου εκτίμησης.

Σκοπό έχει να αυξήσει την ερμηνευτική ικανότητα του μοντέλου.

# Εκτίμηση Παραμέτρων

Μέθοδος Κύριων Συνιστωσών

Πόσο καλό είναι το μοντέλο?



Εταιρικόητες:

Εκφράζουν το ποσοστό της διακύμανσης που εξηγείται από τον αριθμό των παραγόντων που προσαρμόσαμε.

Το μοντέλο ερμηνεύει μόλις το 43,5% της μεταβλητής *Statistics makes me cry!*.

Η εκτίμηση της **ιδιαιτερότητας** κάθε μεταβλητής (το μέρος της διακύμανσης κάθε μεταβλητής που δεν εξηγεί το μοντέλο) υπολογίζεται αν αφαιρέσω από την τομή initial-την τιμή extraction

Η Ιδιαιτερότητα για την ***Statistics makes me cry*** είναι :  $1 - 0,435 = 0,565$

Η πρώτη στήλη είναι 1 αν έχω χρησιμοποιήσει κύριες συνιστώσες

Communalities

	Initial	Extraction
Statistics makes me cry	1,000	,435
My friends will think I'm stupid for not being able to cope with SPSS	1,000	,414
Standard deviations excite me	1,000	,530
I dream that Pearson is attacking me with correlation coefficients	1,000	,469
I don't understand statistics	1,000	,343
I have little experience of computers	1,000	,654
All computers hate me	1,000	,545
I have never been good at mathematics	1,000	,739
My friends are better at statistics than me	1,000	,484
Computers are useful only for playing games	1,000	,335
I did badly at mathematics at school	1,000	,690
People try to tell you that SPSS makes statistics easier to understand but it doesn't	1,000	,513
I worry that I will cause irreparable damage because of my incompetence with computers	1,000	,536
Computers have minds of their own and deliberately go wrong whenever I use them	1,000	,488
Computers are out to get me	1,000	,378
I weep openly at the mention of central tendency	1,000	,487
I slip into a coma whenever I see an equation	1,000	,683
SPSS always crashes when I try to use it	1,000	,597
Everybody looks at me when I use SPSS	1,000	,343
I can't sleep for thoughts of eigenvectors	1,000	,484
I wake up under my duvet thinking that I am trapped under a normal distribution	1,000	,550
My friends are better at SPSS than I am	1,000	,464
If I'm good at statistics my friends will think I'm a nerd	1,000	,412

Extraction Method: Principal Component Analysis.

# Πριν την περιστροφή

Component Matrix<sup>a</sup>

	Component			
	1	2	3	4
Statistics makes me cry	,586	,175	-,215	,119
My friends will think I'm stupid for not being able to cope with SPSS	-,303	,548	,146	,010
Standard deviations excite me	-,629	,290	,213	-,067
I dream that Pearson is attacking me with correlation coefficients	,634	,144	-,149	,153
I don't understand statistics	,556	,101	-,074	,137
I have little experience of computers	,562	,097	,571	-,048
All computers hate me	,685	,039	,252	,103
I have never been good at mathematics	,549	,401	-,323	-,417
My friends are better at statistics than me	-,284	,627	-,008	,103
Computers are useful only for playing games	,437	,035	,363	-,103
I did badly at mathematics at school	,652	,245	-,209	-,400
People try to tell you that SPSS makes statistics easier to understand but it doesn't	,669	-,048	,051	,248
I worry that I will cause irreparable damage because of my incompetence with computers	,673	,076	,278	-,008
Computers have minds of their own and deliberately go wrong whenever I use them	,656	,023	,198	,135
Computers are out to get me	,593	,010	,117	-,113
I weep openly at the mention of central tendency	,679	,014	-,138	,080
I slip into a coma whenever I see an equation	,643	,330	-,210	-,342
SPSS always crashes when I try to use it	,701	,033	,298	,125
Everybody looks at me when I use SPSS	-,427	,390	,095	-,013
I can't sleep for thoughts of eigenvectors	,436	-,205	-,404	,297
I wake up under my duvet thinking that I am trapped under a normal distribution	,658	-,055	-,187	,282
My friends are better at SPSS than I am	-,302	,465	-,116	,378
If I'm good at statistics my friends will think I'm a nerd	-,144	,366	-,021	,507

Extraction Method: Principal Component Analysis.

a. 4 components extracted.

Ο πίνακας L δείχνει τις **επιβαρύνσεις** των παραγόντων για το μοντέλο με 4 παράγοντες. Μπορεί έτσι να εκφραστεί κάθε μεταβλητή με τη χρήση των 4 παραγόντων.

Π.χ.  $Statistics\ makes\ me\ cry = 0,586 F_1 + 0,175 F_2 - 0,215 F_3 + 0,119 F_4$

# Μετά την περιστροφή

Rotated Factor Matrix<sup>a</sup>

	Factor			
	1	2	3	4
I wake up under my duvet thinking that I am trapped under a normal distribution	.594			
I weep openly at the mention of central tendency	.543			
I dream that Pearson is attacking me with correlation coefficients	.527			
People try to tell you that SPSS makes statistics easier to understand but it doesn't	.510	.398		
Standard deviations excite me	-.505			.399
Statistics makes me cry	.504			
I can't sleep for thoughts of eigenvectors	.465			
I don't understand statistics	.436			
I have little experience of computers		.753		
SPSS always crashes when I try to use it	.366	.612		
I worry that I will cause irreparable damage because of my incompetence with computers		.564		
All computers hate me	.364	.559		
Computers have minds of their own and deliberately go wrong whenever I use them	.388	.485		
Computers are useful only for playing games		.380		
Computers are out to get me		.377		
I have never been good at mathematics			.759	
I did badly at mathematics at school			.688	
I slip into a coma whenever I see an equation			.641	
My friends are better at statistics than me				.559
My friends are better at SPSS than I am				.465
My friends will think I'm stupid for not being able to cope with SPSS				.464
Everybody looks at me when I use SPSS				.375
If I'm good at statistics my friends will think I'm a nerd				.329

Οι επιβαρύνσεις  
μετά την  
περιστροφή.

Extraction Method: Principal Axis Factoring.

Rotation Method: Varimax with Kaiser Normalization.<sup>a</sup>

a. Rotation converged in 7 iterations.



# Ερμηνεία παραγόντων:

- Οι ερωτήσεις με υψηλές επιβαρύνσεις στον πρώτο παράγοντα φαίνεται να συνδέονται με διάφορα θέματα της στατιστικής. Επομένως ο 1<sup>ος</sup> παράγοντας μπορεί να ονομασθεί *φόβος για τη στατιστική*.
- Οι ερωτήσεις με υψηλές επιβαρύνσεις στο δεύτερο παράγοντα φαίνεται να συνδέονται με τη χρήση υπολογιστών ή του SPSS. Επομένως ο 2<sup>ος</sup> παράγοντας μπορεί να ονομασθεί *φόβος για τους υπολογιστές*.
- Οι ερωτήσεις με υψηλές επιβαρύνσεις στον τρίτο παράγοντα φαίνεται να συνδέονται με διάφορα θέματα των μαθηματικών. Επομένως ο 1<sup>ος</sup> παράγοντας μπορεί να ονομασθεί *φόβος για τα μαθηματικά*.
- Τέλος, ερωτήσεις με υψηλές επιβαρύνσεις στον τέταρτο παράγοντα περιέχουν στοιχεία από την κοινωνική αξιολόγηση από φίλους. Επομένως, ο 1<sup>ος</sup> παράγοντας μπορεί να ονομασθεί *συγκριτική αξιολόγηση*.

# Ερμηνεία παραγόντων:

Συνοψίζοντας, η ανάλυση αποκαλύπτει ότι το ερωτηματολόγιο αποτελείται από 4 υποομάδες:

- φόβο για τη στατιστική,
- φόβο για τους υπολογιστές,
- φόβο για τα μαθηματικά και
- φόβο για αρνητική συγκριτική αξιολόγηση.

# Εκτίμηση των Σκορ

Ένας από τους σκοπούς της παραγοντικής ανάλυσης είναι η **δημιουργία καινούργιων μεταβλητών**, οι οποίες συμπυκνώνουν όσο γίνεται τη διακύμανση των αρχικών μεταβλητών, μειώνοντας έτσι τις διαστάσεις του προβλήματος.

Σκοπός είναι να εκφράσουμε κάθε παράγοντα ως γραμμικό συνδυασμό των αρχικών μας μεταβλητών.

# Εκτίμηση των Σκορ

Component Score Coefficient Matrix

	Component			
	1	2	3	4
Statistics makes me cry	-.053	,173	,089	,110
My friends will think I'm stupid for not being able to cope with SPSS	,102	-.129	,086	,282
Standard deviations excite me	,087	-.195	,013	,137
I dream that Pearson is attacking me with correlation coefficients	-.011	,170	,045	,107
I don't understand statistics	,021	,131	,014	,083
I have little experience of computers	,383	-.211	-.088	,014
All computers hate me	,213	,004	-.078	,038
I have never been good at mathematics	-.129	-.074	,460	,013
My friends are better at statistics than me	,025	-.029	,108	,354
Computers are useful only for playing games	,244	-.161	-.021	-.036
I did badly at mathematics at school	-.066	-.087	,379	-.059
People try to tell you that SPSS makes statistics easier to understand but it doesn't	,097	,161	-.116	,051
I worry that I will cause irreparable damage because of my incompetence with computers	,224	-.065	-.019	,013
Computers have minds of their own and deliberately go wrong whenever I use them	,180	,040	-.084	,043
Computers are out to get me	,114	-.055	,061	-.058
I weep openly at the mention of central tendency	-.015	,146	,046	,014
I slip into a coma whenever I see an equation	-.057	-.067	,372	,005
SPSS always crashes when I try to use it	,242	-.001	-.104	,043
Everybody looks at me when I use SPSS	,048	-.115	,061	,199
I can't sleep for thoughts of eigenvectors	-.195	,359	-.061	-.002
I wake up under my duvet thinking that I am trapped under a normal distribution	-.039	,270	-.064	,059
My friends are better at SPSS than I am	-.036	,162	-.048	,382
If I'm good at statistics my friends will think I'm a nerd	,032	,211	-.162	,379

$$F_1 = -0,053 Q1 + 0,102 Q2 + 0,087 Q3 \dots + 0,032 Q23$$

Στην ουσία ποσοτικοποιήσαμε τους παράγοντες που υποθέσαμε ότι εξηγούν τις συσχετίσεις των μεταβλητών μας.

Extraction Method: Principal Component Analysis.  
Rotation Method: Varimax with Kaiser Normalization.  
Component Scores.



Κάνοντας ταξινόμηση με βάση τον πρώτο παράγοντα βρίσκει κανείς αυτόν που εκφράζει τον μεγαλύτερο φόβο για την Στατιστική

# Χρήση των Σκορ

	Question_19	Question_20	Question_21	Question_22	Question_23	FAC1_1	FAC2_1	FAC3_1	FAC4_1
1	3	2	2	2	5	,10587	-,92816	-,182760	-,45953
2	3	4	4	4	2	-,58279	-,18932	-,04137	,29332
3	1	4	3	2	2	-,54761	,02953	,19918	-,97142
4	2	4	4	4	3	,74661	,72268	-,68654	-,18547
5	3	4	2	4	4	,25169	-,51489	-,63354	,68301
6	1	5	3	1	4	1,91614	-,27356	-,68201	-,52709
7	3	2	2	4	4	-,26051	-,140947	-,00431	,91008
8	4	3	2	4	4	-,28572	-,91761	-,08947	1,03746
9	2	5	5	3	3	1,71750	1,15102	3,15659	,81082
10	3	3	2	4	4	-,69151	-,73318	,18376	1,49541
11	5	3	2	5	5	-,91102	-,43355	-,67959	2,30844
12	3	4	3	4	4	,00265	,49479	-,100502	,63617
13	2	4	4	3	4	,60378	,87496	,61359	,11381
14	1	5	5	3	1	,30849	,89540	-,68141	-,140898
15	3	4	4	4	4	,19016	,66551	-,58721	,74920
16	2	3	2	3	4	-,08220	-,20762	-,54816	-,46101
17	4	2	1	4	4	-,142215	-,173068	-,63872	1,02073

Τα σκορ αποτελούν πλέον τις νέες μεταβλητές μας και μπορεί να τις χρησιμοποιήσει κανείς για να συνεχίσει την ανάλυση.

Έχουν δημιουργηθεί 4 νέες στήλες στο αρχείο των δεδομένων με τις τιμές των σκορ.