



ΠΑΝΕΠΙΣΤΗΜΙΟ
ΠΕΛΟΠΟΝΝΗΣΟΥ
UNIVERSITY of the PELOPONNESE

1

Τμήμα Διοίκησης Επιχειρήσεων και Οργανισμών Ποσοτικές Μέθοδοι Ανάλυσης Ενότητα: Ανάλυση Παλινδρόμησης

Διδάσκουσα: Κρεμαστιώτη Βίκυ
v.kremastioti@go.uop.gr

Μαθησιακοί στόχοι:

- ✓ Η Σύγχρονη Ερμηνεία της Παλινδρόμησης
- ✓ Στατιστικές Σχέσεις έναντι Προσδιοριστικών Σχέσεων
- ✓ Παλινδρόμηση έναντι Αιτιότητας
- ✓ Παλινδρόμηση έναντι Συσχέτισης
- ✓ Ορολογία και Συμβολισμοί

- Η ανάλυση παλινδρόμησης ασχολείται με τη μελέτη της εξάρτησης μίας μεταβλητής, της **εξαρτημένης μεταβλητής**, από μία ή από περισσότερες άλλες μεταβλητές, τις **ερμηνευτικές μεταβλητές**.
- **Σκοπός** είναι η εκτίμηση ή/και η πρόβλεψη της τιμής του (πληθυσμιακού) μέσου όρου σε όρους των γνωστών ή σταθερών τιμών (σε επαναλαμβανόμενη δειγματοληψία) των τελευταίων και ο έλεγχος της υπόθεσης αναφορικά με τη φύση της εξαρτημένης μεταβλητής.

Η Ερμηνεία της Παλινδρόμησης

Παράδειγμα 1: Έστω ότι θέλουμε να μελετήσουμε την επίπτωση της φορολογίας στο ΑΕΠ

Βήματα:

1. Αναγνωρίζω τις μεταβλητές
2. **Θέτω X = φόροι και Y = ΑΕΠ.**

- Το βλέπουμε αυτό με τη χρήση ενός **διαγράμματος διασποράς** (scatter diagram – scattergram).

Η Ερμηνεία της Παλινδρόμησης – Διάγραμμα Διασποράς

- Το **διάγραμμα διασπορά** δείχνει την κατανομή των υψών των γιων σε έναν υποθετικό πληθυσμό που αντιστοιχούν σε συγκεκριμένες ή σταθερές τιμές ύψους του πατέρα.
- Η ένωση αυτών μας δίνει τη **γραμμή παλινδρόμησης** (regression line).
- Η γραμμή παλινδρόμησης δείχνει πώς το μέσο ύψος των γιων αυξάνεται καθώς αυξάνεται το ύψος του πατέρα.

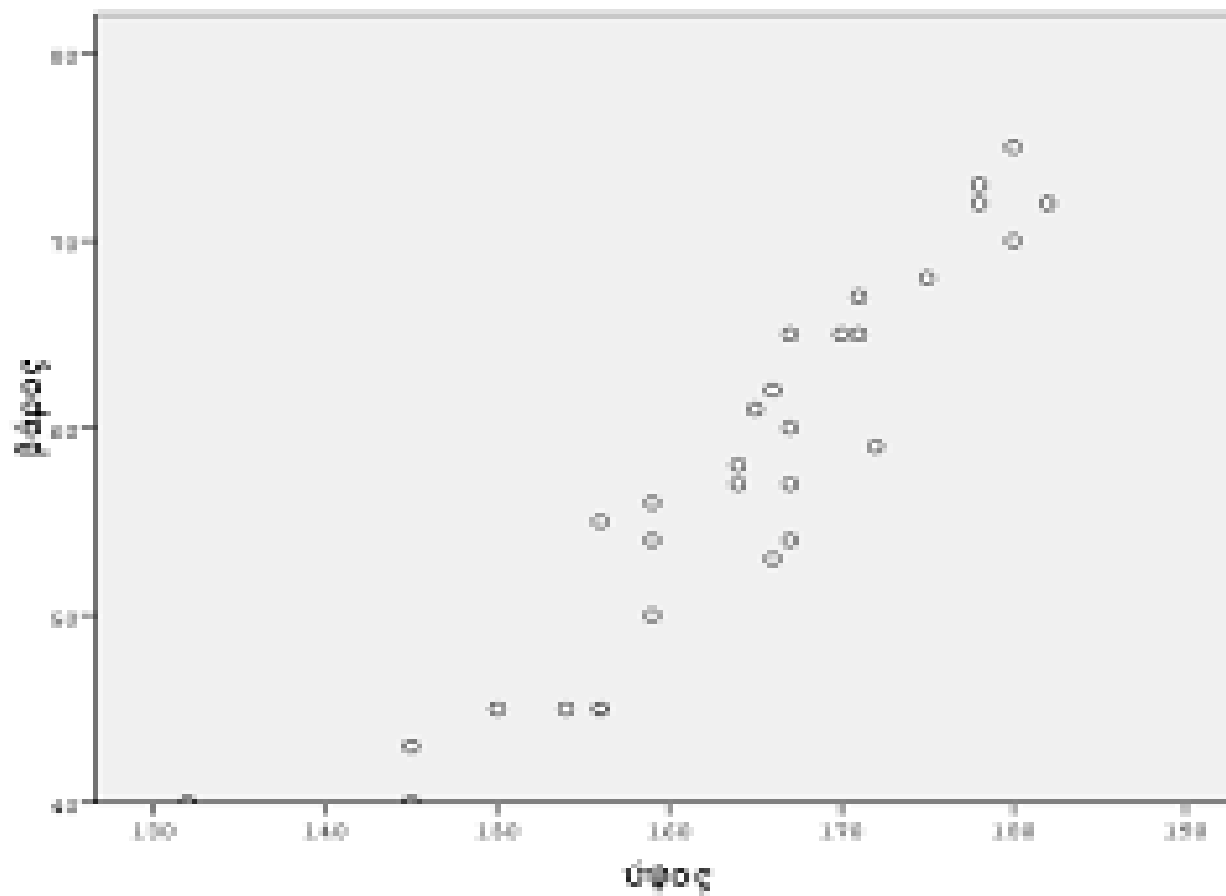
• Παραδείγματα

- Το διάγραμμα διασποράς 1.2 παρουσιάζει την κατανομή του ύψους των αγοριών, που μετράται σε σταθερές ηλικίες, ενός υποθετικού πληθυσμού.
- Σε κάθε ηλικία, έχουμε μία σειρά, ή κατανομή, υψών.
- Το ύψος, *κατά μέσο όρο*, αυξάνει με την ηλικία (φυσικά, μέχρι μία ορισμένη ηλικία), κάτι που γίνεται προφανές, αν σχεδιάσουμε μία γραμμή (τη γραμμή παλινδρόμησης).
- Η γραμμή αυτή διατρέχει τα σημεία που περιέχονται στους κύκλους και τα οποία αντιπροσωπεύουν το μέσο ύψος στις συγκεκριμένες ηλικίες.

- Οικονομικά παραδείγματα.
- Ένας οικονομολόγος μπορεί να ενδιαφέρεται για τη μελέτη της εξάρτησης των προσωπικών καταναλωτικών δαπανών από το μετά φόρων ή διαθέσιμο πραγματικό προσωπικό εισόδημα.
- Εκτίμηση της οριακής ροπής προς κατανάλωση (MPC) για μία, π.χ., μεταβολή του πραγματικού εισοδήματος κατά ένα δολάριο.

- **Παραδείγματα**
- **Οικονομικά παραδείγματα.**
- Ένας μονοπωλητής που μπορεί να καθορίσει την τιμή ή την εκροή (αλλά όχι και τα δύο) μπορεί να θέλει να μάθει την ανταπόκριση της ζήτησης για ένα προϊόν στις μεταβολές των τιμών.
- Ένα τέτοιο πείραμα εκτιμά την **ελαστικότητα ζήτησης ως προς την τιμή** (price elasticity).

Διάγραμμα Διασποράς



- **Παραδείγματα**
- **Οικονομικά παραδείγματα.**
- Ο διευθυντής μάρκετινγκ μίας εταιρείας μπορεί να θέλει να γνωρίζει πως σχετίζεται η ζήτηση για το προϊόν της εταιρείας με π.χ., τις διαφημιστικές δαπάνες.
- **Ελαστικότητα της ζήτησης** (elasticity of demand) σε σχέση με τις διαφημιστικές δαπάνες, δηλαδή, την ποσοστιαία μεταβολή της ζήτησης σε μία μεταβολή του διαφημιστικού προϋπολογισμού κατά 1 τοις εκατό.

Η Σύγχρονη Ερμηνεία της Παλινδρόμησης

- **Παραδείγματα**
- Ένας γεωπόνος μπορεί να ενδιαφέρεται για τη μελέτη της εξάρτησης της απόδοσης μίας καλλιέργειας, π.χ., του σίτου, από τη θερμοκρασία, τις βροχοπτώσεις, την ηλιοφάνεια και τα λιπάσματα.

Στατιστικές Σχέσεις , Προσδιοριστικών Σχέσεων

- Η εξάρτηση της απόδοσης των καλλιεργειών από τη θερμοκρασία, τις βροχοπτώσεις, την ηλιοφάνεια και το λίπασμα
- Για παράδειγμα, είναι στατιστικής φύσης υπό την έννοια ότι οι ερμηνευτικές μεταβλητές, αν και είναι σίγουρα σημαντικές, δε θα επιτρέψουν στο γεωπόνο να προβλέψει την απόδοση των καλλιεργειών:
- Λόγω σφαλμάτων που εμπλέκονται στη μέτρηση αυτών των μεταβλητών
- Λόγω μίας σειράς άλλων παραγόντων (μεταβλητών) που επηρεάζουν συλλογικά την απόδοση, αλλά είναι δύσκολο να προσδιοριστούν ξεχωριστά.

Παλινδρόμηση έναντι Αιτιότητας

- Στο παράδειγμα της απόδοσης των καλλιεργειών, δεν υπάρχει **κανένας στατιστικός λόγος** να υποθέσουμε ότι οι βροχοπτώσεις δεν εξαρτώνται από την απόδοση των καλλιεργειών.
- Αντιμετωπίζουμε την απόδοση των καλλιεργειών ως ένα παράγοντα που εξαρτάται από τη βροχόπτωση (μεταξύ άλλων)
- Αυτό οφείλεται σε μη-στατιστικές εκτιμήσεις: δε μπορούμε να ελέγξουμε τις βροχοπτώσεις μεταβάλλοντας την απόδοση των καλλιεργειών.
- **Μία στατιστική σχέση καθεαυτή δε συνεπάγεται λογικά αιτιότητα.**
- Για να καταλογίσει κάποιος αιτιότητα, πρέπει να επικαλεστεί *a priori* θεωρητικές απόψεις.

Παλινδρόμηση και Συσχέτιση

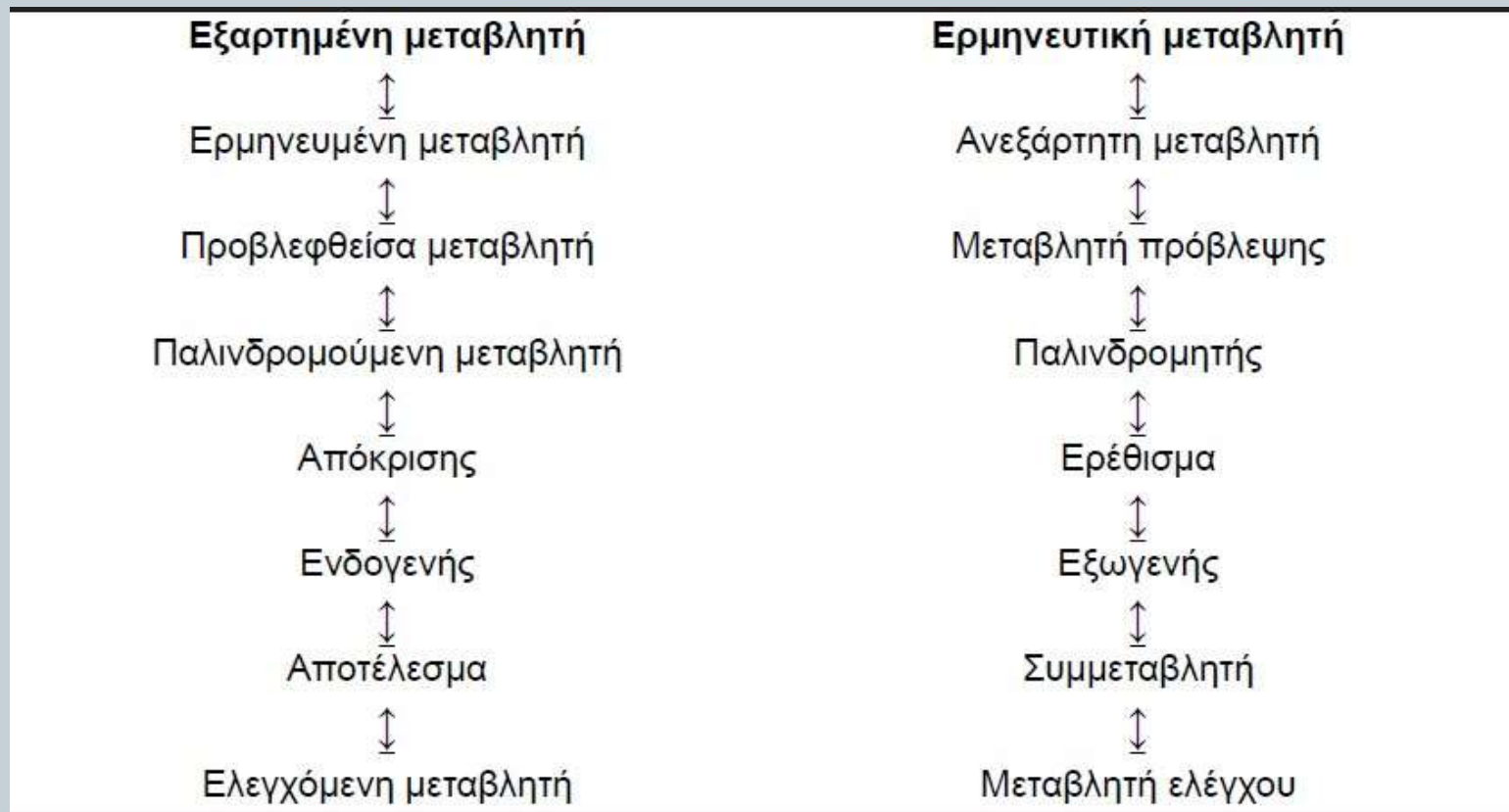
- Η **ανάλυση συσχέτισης** (correlation analysis) είναι στενά συνδεδεμένη, αλλά εννοιολογικά πολύ διαφορετική, από την ανάλυση παλινδρόμησης, όπου ο πρωταρχικός στόχος είναι η μέτρηση της *δύναμης* ή του *βαθμού γραμμικής συσχέτισης* μεταξύ δύο μεταβλητών.
- Ο **συντελεστής συσχέτισης** (correlation coefficient), μετρά τη δύναμη αυτής της (γραμμικής) σχέσης.
- **Παράδειγμα:** συσχέτιση (το συντελεστή συσχέτισης) μεταξύ του καπνίσματος και του καρκίνου του πνεύμονα, μεταξύ των βαθμών των μαθητών στη στατιστική και στα μαθηματικά, μεταξύ των βαθμών στο λύκειο και των βαθμών στο πανεπιστήμιο και ούτω καθεξής.
- Στην **ανάλυση παλινδρόμησης** προσπαθούμε να εκτιμήσουμε ή να προβλέψουμε τη μέση τιμή μίας μεταβλητής βάσει των σταθερών τιμών άλλων μεταβλητών.

Παλινδρόμηση & Συσχέτιση (2)

- **Θεμελιώδεις διαφορές παλινδρόμησης και συσχέτισης**
- *Ανάλυση παλινδρόμησης:*
- Ασυμμετρία στον τρόπο με τον οποίο αντιμετωπίζονται η εξαρτημένη και οι ερμηνευτικές μεταβλητές.
- Η εξαρτημένη μεταβλητή θεωρείται ότι είναι στατιστική, τυχαία ή στοχαστική, δηλαδή ότι έχει μία κατανομή πιθανότητας.
- Οι ερμηνευτικές μεταβλητές θεωρούνται ότι έχουν σταθερές τιμές (σε επαναλαμβανόμενη δειγματοληψία),

Συμβολισμοί

- Στη βιβλιογραφία οι όροι *εξαρτημένη μεταβλητή (Y)* και *ερμηνευτική μεταβλητή (X)* περιγράφονται ποικιλοτρόπως.



Συμβολισμοί

- Αν μελετούμε την εξάρτηση μίας μεταβλητής από μία μόνο ερμηνευτική μεταβλητή, όπως αυτή της καταναλωτικής δαπάνης από το πραγματικό εισόδημα, μία τέτοια μελέτη είναι γνωστή ως απλή, ή **διμεταβλητή ανάλυση παλινδρόμησης**
- Αν μελετούμε την εξάρτηση μίας μεταβλητής από περισσότερες από μία ερμηνευτικές μεταβλητές, όπως στο παράδειγμα της απόδοσης των καλλιεργειών και της βροχόπτωσης, της θερμοκρασίας, της ηλιοφάνειας και των λιπασμάτων, τότε είναι γνωστή ως **ανάλυση πολλαπλής παλινδρόμησης**

Διαστρωματικά –Χρονοσειρές και Πάνελ Δεδομένα

- Το γράμμα Y συνήθως δηλώνει την εξαρτημένη μεταβλητή
- Το γράμμα X ($X_1, X_2, \dots, X_k$) συνήθως δηλώνει τις ερμηνευτικές μεταβλητές, X_k είναι η $k^{\text{η}}$ ερμηνευτική μεταβλητή.
- Ο δείκτης i ή t συνήθως δηλώνει την $i^{\text{η}}$ ή $t^{\text{η}}$ παρατήρηση ή τιμή.
- X_{ki} (ή X_{kt}) συνήθως δηλώνει την $i^{\text{η}}$ (ή $t^{\text{η}}$) παρατήρηση της μεταβλητής X_k .
- N (ή T) συνήθως δηλώνει το συνολικό αριθμό των παρατηρήσεων ή των τιμών στον πληθυσμό
- n (ή t) συνήθως δηλώνει το συνολικό αριθμό των παρατηρήσεων στο δείγμα.
- Ο δείκτης i χρησιμοποιείται για **διαστρωματικά στοιχεία** (cross-sectional data) (δηλαδή, στοιχεία που έχουν συλλεχθεί σε μία χρονική στιγμή)
- Ο δείκτης t χρησιμοποιείται για **χρονοσειρές στοιχείων** (time series data) (δηλαδή, στοιχεία που έχουν συλλεχθεί κατά τη διάρκεια μίας χρονικής περιόδου).

Εισαγωγή στην ανάλυση της παλινδρόμησης.

- Ευχαριστώ πολύ για την προσοχή σας!