

Ερευνητική Μεθοδολογία και Μαθηματική Στατιστική

Κύρια σημεία

- Εφαρμοσμένη και Μαθηματική Στατιστική
- Στατιστικά Μοντέλα
- Αναζήτηση ερευνητικού θέματος
- Εισαγωγή στην έρευνα
- Διάρθρωση ερευνητικής εργασίας

Έρευνα στην εφαρμοσμένη Στατιστική

- Στάδια της έρευνας
 - Διατύπωση προβλήματος με όρους της αρχικής εφαρμογής
 - Στατιστικό Μοντέλο
 - Διαμόρφωση παραμέτρων/Επιλογή Μεταβλητών
 - Στατιστική ανάλυση δεδομένων
 - Έλεγχος υποθέσεων και διορθώσεις του μοντέλου
 - Διατύπωση συμπερασμάτων με όρους της αρχικής εφαρμογής

Έρευνα στη Μαθηματική Στατιστική

- Πώς και από πού αντλούνται ερευνητικές ιδέες;
- Τι αποτελεί έρευνα στη στατιστική μεθοδολογία;
- Πώς ξεκινάει κάποιος την ερευνητική δραστηριότητα;
- Πότε θεωρείται ολοκληρωμένη μια έρευνα;

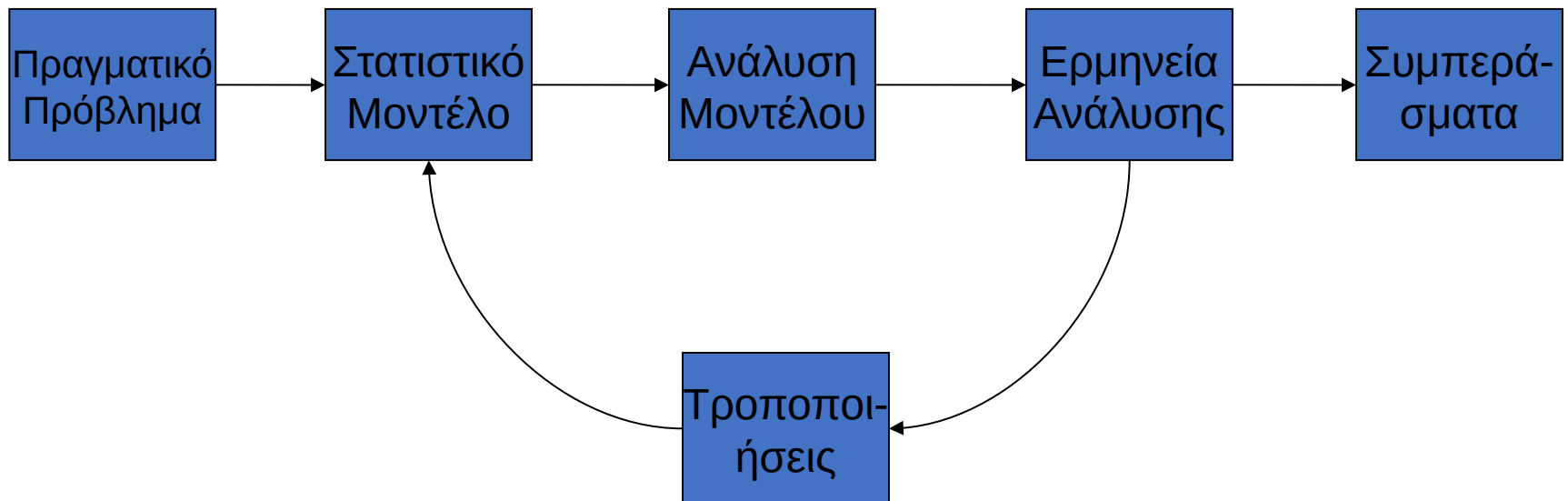
Αντικείμενο της Μαθηματικής Στατιστικής

- Μελέτη και ανάπτυξη μεθόδων στατιστικής ανάλυσης.
- Χρησιμοποιεί μαθηματικά εργαλεία: θεωρία πιθανοτήτων, βελτιστοποίηση, λογισμό.
- Κοινό σημείο με την εφαρμοσμένη στατιστική: πρότυπα ή μοντέλα.

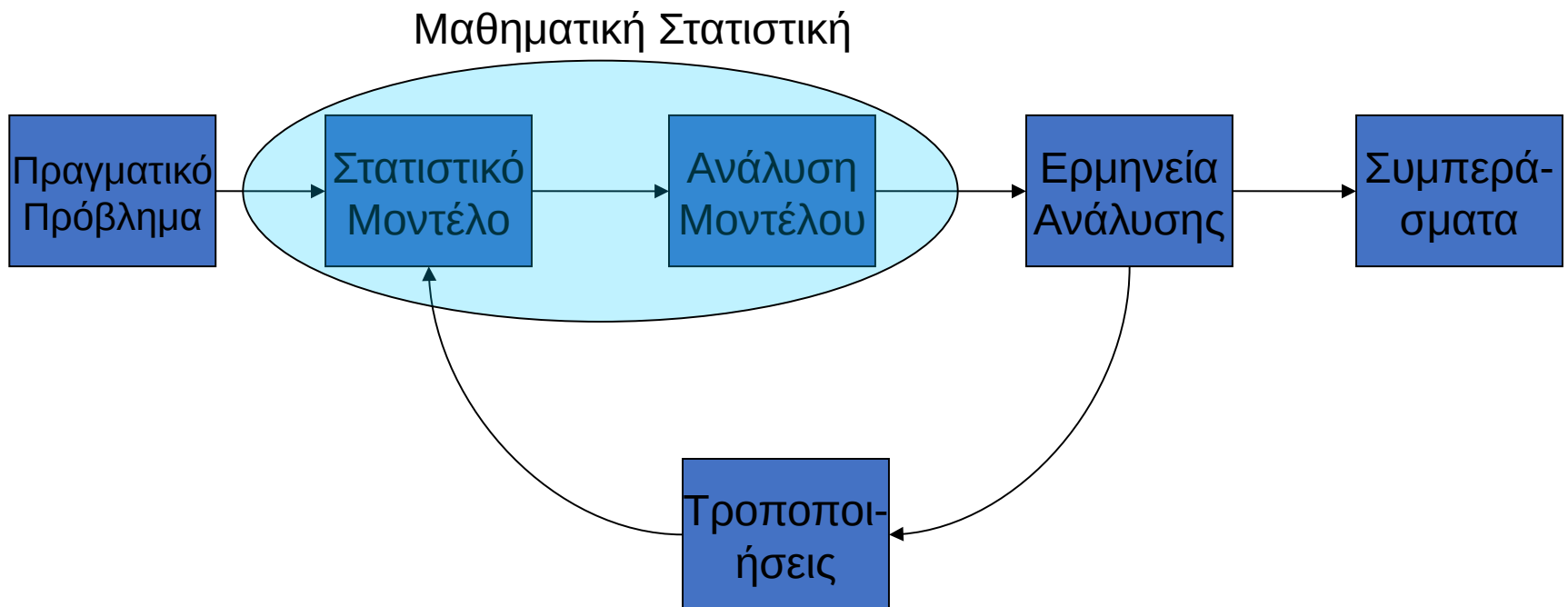
Η έννοια του μοντέλου

- Αναπαράσταση μιας πραγματικής κατάστασης
- Διάφοροι τύποι μοντέλων: Φυσικά, Αφηρημένα, Υπολογιστικά, Μαθηματικά
- Μαθηματικό Μοντέλο:
Αναπαράσταση μέσω μαθηματικών ποσοτήτων και σχέσεων μεταξύ τους.

Στάδια Μοντελοποίησης - Ανάλυσης



Στάδια Μοντελοποίησης – Ανάλυσης



Στατιστικά Μοντέλα

- Διευκολύνουν την εξαγωγή συμπερασμάτων από εμπειρικά δεδομένα
 - Έλεγχοι Υποθέσεων
 - Παλινδρόμηση
 - Ανάλυση Επιβίωσης
 - Παραγοντική Ανάλυση κλπ

Στοχαστικά – Πιθανοθεωρητικά Μοντέλα

- Γενικότερη κατηγορία μοντέλων.
- Μαθηματική περιγραφή φαινομένων υπό συνθήκες αβεβαιότητας.
- Εφαρμογές στη Βιοστατιστική:
 - Κλινικές Δοκιμές
 - Μοντέλα Εξέλιξης στη Βιολογία
- Μαθηματικά Εργαλεία
 - Πιθανότητες - Στοχαστικές Διαδικασίες
 - Βελτιστοποίηση
 - Προσομοίωση

Έρευνα στη Μαθηματική Στατιστική

- Δημιουργία Νέων Μοντέλων
 - Εμπνέονται από ανάγκες εφαρμογών
 - Γενικού σκοπού
 - Μπορεί να εφαρμόζονται σε διαφορετικές περιοχές προβλημάτων
 - Μπορεί να προέρχονται από διαφορετικές περιοχές προβλημάτων

Έρευνα στη Μαθηματική Στατιστική

- Ανάπτυξη μεθόδων για την ανάλυση στατιστικών μοντέλων
 - Μαθηματική προσέγγιση
 - Περιγραφή μεθόδου
 - Απόδειξη ιδιοτήτων
 - Προσομοίωση συμπεριφοράς

Πώς βρίσκονται ερευνητικές ιδέες

- Μελέτη βιβλιογραφίας
- Τίποτε δεν θεωρείται δεδομένο
- Σε όλες τις εργασίες υπάρχουν «κενά» ή «τρύπες»
- Η διαφορά του ερευνητή από τον αναγνώστη είναι ότι αναγνωρίζει τα κενά και επομένως τις ευκαιρίες για περαιτέρω έρευνα

Ιδέες για νέα μοντέλα

- Ανάγκες εφαρμογών
- Η αιτιολόγηση της έρευνας προς κάποια κατεύθυνση βρίσκει σημαντική υποστήριξη από τη δυνατότητα παραγωγής εφαρμόσιμης γνώσης.
- Αυτό ισχύει στα εφαρμοσμένα μαθηματικά γενικότερα και όχι μόνο στη στατιστική.

Από την εφαρμογή στη θεωρία

- Το μεγαλύτερο μέρος των μοντέλων μαθηματικής στατιστικής έχει προέλθει από πρακτικά ερωτήματα και προβλήματα.
- Κλασσικά παραδείγματα είναι η μέθοδος παλινδρόμησης και η ανάλυση κατά παράγοντες

Από την εφαρμογή στη θεωρία

- Ένα ερώτημα/πρόβλημα από την πράξη μπορεί να οδηγήσει σε:
 - Εφαρμογή ενός γνωστού μοντέλου που έχει ήδη προταθεί και αναλυθεί (εφαρμοσμένη στατιστική)
 - Δημιουργία νέου μοντέλου για τις ανάγκες του συγκεκριμένου προβλήματος (μαθηματική στατιστική)

Παράδειγμα: Μεικτά μοντέλα παλινδρόμησης για χρονικά δεδομένα

- Το στατιστικό μοντέλο παλινδρόμησης έχει σκοπό να περιγράψει συσχετίσεις ανάμεσα σε μια εξαρτημένη μεταβλητή (Y) και μία ή περισσότερες ανεξάρτητες μεταβλητές (X)
- Παράδειγμα: X =δόση φαρμάκου, Y =πίεση

Κλασικό μοντέλο

- Σε ένα δείγμα ασθενών δίνονται διαφορετικές δόσεις και μετράται η πίεση.
- Στο μοντέλο παλινδρόμησης γίνεται η υπόθεση ότι οι ασθενείς είναι στατιστικά ανεξάρτητοι μεταξύ τους.

Χρονικά δεδομένα

- Έστω ότι κάθε ασθενής παίρνει μια δόση του φαρμάκου τρεις φορές τη μέρα (πρωί, μεσημέρι και βράδυ) και κάνει μέτρηση της πίεσης δύο φορές τη μέρα.
- Τώρα οι μετρήσεις δεν προέρχονται όλες από διαφορετικούς ασθενείς και επομένως η ανεξαρτησία δεν ισχύει.

Ανεπάρκεια κλασικού μοντέλου

- Το κλασικό μοντέλο παλινδρόμησης δεν μπορεί να εφαρμοστεί σε τέτοιες περιπτώσεις.
- Γενικότερα μοντέλα (μεικτά γραμμικά μοντέλα) έχουν αναπτυχθεί από ερευνητές εφαρμοσμένης και μαθηματικής στατιστικής.

Παράδειγμα: Από τη συνδυαστική βελτιστοποίηση στη βιολογική εξέλιξη

- Το 1987 ο μαθηματικός-βιολόγος Stuart Kauffman πρότεινε ένα μαθηματικό μοντέλο για την περιγραφή της εξέλιξης γονιδίων/οργανισμών/ομάδων.
- Μοντέλο NK
- Βασίζεται στη θεωρία συνδυαστικής βελτιστοποίησης.

Μοντέλο NK

- Ένας οργανισμός αποτελείται από N τμήματα. Κάθε τμήμα μπορεί να βρίσκεται στην κατάσταση 0 ή 1.
- Η «απόδοση» του οργανισμού είναι το άθροισμα των επί μέρους αποδόσεων των τμημάτων του.
- Η απόδοση κάθε τμήματος εξαρτάται από την κατάσταση του τμήματος (0 ή 1) όπως επίσης και από τις αποδόσεις $K-1$ γειτονικών του τμημάτων (επιστατική αλληλεπίδραση).

Μοντέλο NK

- Ο οργανισμός εξελίσσεται αλλάζοντας κάθε φορά την κατάσταση ενός τμήματος αν η αλλαγή βελτιώνει τη συνολική απόδοση.
- Όταν καμμία αλλαγή δεν επιφέρει βελτίωση έχει φθάσει σε κατάσταση τοπικού βέλτιστου.

Μοντέλο NK

- Μετά από συνδυασμό θεωρητικής ανάλυσης και προσομοιώσεων για διάφορα σενάρια αποδόσεων, ο Kauffman κατέληξε στο συμπέρασμα ότι η απόδοση είναι καλύτερη για μέτριες τιμές του K (μέτρια αλληλεπίδραση) ενώ μειώνεται για μεγάλες τιμές (καταστροφή λόγω πολυπλοκότητας)

Μοντέλο NK

- Αν δεχτούμε ότι το μοντέλο περιγράφει (προσεγγιστικά) αυτοεξελισσόμενους οργανισμούς, τότε συμπεραίνουμε ότι ενδιάμεσοι βαθμοί πολυπλοκότητας είναι προτιμότεροι στην εξέλιξη απότι πολύ χαμηλοί ή πολύ υψηλοί.
- Ίσως αυτό εξηγεί κάποια φαινόμενα στην εξελικτική διαδικασία.

Μοντέλο NK

- Υπεραπλουστευτικές υποθέσεις:
 - Μόνο δύο δυνατές καταστάσεις κάθε τμήματος
 - Ίδια μορφή αλληλεπίδρασης για όλα τα τμήματα
 - Τρόπος εξέλιξης με τοπικές βελτιώσεις

Εφαρμογή σε άλλες περιοχές

- Οργανισμοί – Ομάδες εργαζομένων
- Αλληλεπίδραση στην απόδοση της εργασίας
- Ο ρόλος του διευθυντή
- Πότε πρέπει να χωριστούν οι μεγάλες ομάδες

Έρευνα για νέες μεθόδους

- Σημαντικό μέρος της έρευνας στη μαθηματική στατιστική αποτελεί η ανάπτυξη νέων ή βελτιωμένων μεθόδων για υπάρχοντα μοντέλα.
- Παραδείγματα
 - Μέθοδοι επιλογής μεταβλητών σε μοντέλα παλινδρόμησης
 - Μέθοδοι υπολογισμού παραγόντων σε μοντέλα παραγοντικής ανάλυσης

Συνδυασμός μοντέλων- μεθόδων

- Πολλές φορές δεν υπάρχει ξεκάθαρη διάκριση αν το αποτέλεσμα της έρευνας είναι ένα νέο μοντέλο ή μια νέα μέθοδος
- Μπορεί να είναι συνδυασμός και των δύο
- Παράδειγμα: Ένα μοντέλο κλινικών δοκιμών και η ανάλυση της μεθόδου

Εισαγωγή στην έρευνα

- Εξαρτάται από το υπόβαθρο, τις γνώσεις και τις ιδιαίτερες δεξιότητες του υποψήφιου ερευνητή
- Αν έχει εμπειρία σε εφαρμογές μπορεί να αρχίσει από εφαρμοσμένο πρόβλημα και σταδιακά να προχωρήσει στη μελέτη/ανάπτυξη νέας θεωρίας
- Αν έχει ισχυρό θεωρητικό υπόβαθρο μπορεί να ακολουθηθεί η αντίστροφη διαδικασία.

Ερευνητικά σεμινάρια

- Συγκεκριμένο ερευνητικό θέμα
- Μελετώνται εργασίες πάνω στο θέμα του σεμιναρίου
- Παρουσιάσεις από τους συμμετέχοντες.
- Συζητήσεις για κενά ή ατέλειες.
- Νέες ιδέες για έρευνα.

Το σφάλμα τύπου III

- Να μοντελοποιήσεις την λάθος υπόθεση
- Να μοντελοποιήσεις (και να λύσεις) το λάθος πρόβλημα.

Ολοκλήρωση ερευνητικής εργασίας

- Πότε μπορείς να πείς «τελείωσε»;
- Μια ερευνητική εργασία δεν μπορεί να είναι ούτε πολύ σύντομη (τετριμμένη) ούτε υπερβολικά εκτεταμένη (βιβλίο).
- Δύσκολο να βρεθεί η χρυσή τομή.