



36th Panhellenic and 2nd International Statistics Conference

Data Analytics in Engineering

16-19 May 2024



University of Western Macedonia
Campus of Active Urban
Planning Zone, Kozani, Greece

Organized by:



UNIVERSITY OF
WESTERN MACEDONIA

Περιεχόμενα

- 1. Επιτροπές / Committees**
- 2. Χαιρετισμοί / Welcome Notes**
- 3. Προσκεκλημένοι Ομιλητές / Invited Speakers**
- 4. Εκπαιδευτικό Σεμινάριο / Seminar**
- 5. Πρόγραμμα Συνεδρίου / Conference Program**
- 6. Περιλήψεις στα Ελληνικά / Abstracts in Greek**
- 7. Περιλήψεις στα Αγγλικά / Abstract in English**

Επιτροπές / Committees

Οργανωτική Επιτροπή / Organizing Committee:

- Αγγελής Ε., Καθηγητής, ΑΠΘ.
- Καπετανοπούλου Π., Επίκουρη Καθηγήτρια, Πανεπιστήμιο Δυτικής Μακεδονίας.
- Καραγρηγορίου Α., Καθηγητής, Πανεπιστήμιο Αιγαίου.
- Μαλεφάκη Σ., Αναπληρώτρια Καθηγήτρια, Παν. Πατρών.
- Νενές Γ., Καθηγητής, Πανεπιστήμιο Δυτικής Μακεδονίας.
- Τσιιάς Κ., Επίκουρος Καθηγητής, Πανεπιστήμιο Δυτικής Μακεδονίας.

Επιστημονική Επιτροπή / Scientific Committee:

- Αγγελής Ε., Καθηγητής, ΑΠΘ.
- Αυλογιάρης Γ., Επίκουρος Καθηγητής, Πανεπιστήμιο Δυτικής Μακεδονίας
- Βαμβακάρη Μ., Καθηγήτρια, Χαροκόπειο Πανεπιστήμιο
- Δαφνής Σ., Αναπληρωτής Καθηγητής, Παν. Αιγαίου.
- Δάρας Τ., Αναπληρωτής Καθηγητής, Πολυτεχνείο Κρήτης
- Ευαγγελάρας Χ., Αναπληρωτής Καθηγητής, Παν. Πειραιώς
- Ζωγράφος Κ., Καθηγητής, Πανεπιστήμιο Ιωαννίνων
- Ηλιόπουλος Γ., Καθηγητής, Πανεπιστήμιο Πειραιώς
- Καλαματιανού Α., Ομότιμη Καθηγήτρια, Πάντειο Πανεπιστήμιο Κοινωνικών και Πολιτικών Επιστημών.
- Καπετανοπούλου Π., Επίκουρη Καθηγήτρια, Πανεπιστήμιο Δυτικής Μακεδονίας
- Καραγρηγορίου Α., Καθηγητής, Πανεπιστήμιο Αιγαίου
- Καρλής Δ., Καθηγητής, Οικονομικό Πανεπιστήμιο Αθηνών
- Κίτσος Χ., Ομότιμος Καθηγητής, Παν. Δυτικής Αττικής
- Κουνιάς Σ., Ομότιμος Καθηγητής, ΕΚΠΑ και Επίτιμος Πρόεδρος ΕΣΙ
- Κούτρας Β., Τράπεζα της Ελλάδος.
- Κούτρας Μ., Καθηγητής, Πανεπιστήμιο Πειραιώς.
- Κωνσταντινίδης Δ., Καθηγητής, Πανεπιστήμιο Αιγαίου.
- Κωνσταντινίδης Ε., Καθηγητής, Παν. Δυτικής Μακεδονίας
- Μακρή Φ., Καθηγήτρια, Πανεπιστήμιο Πατρών.
- Μαλεφάκη Σ., Αναπληρώτρια Καθηγήτρια, Παν. Πατρών.

- **Μηλιένος Φ.**, Επίκουρος Καθηγητής, Πάντειο Πανεπιστήμιο Κοινωνικών και Πολιτικών Επιστημών.
- **Μπασιάκος Ι.**, Καθηγητής, ΕΚΠΑ.
- **Μπατσιδής Α.**, Αναπληρωτής Καθηγητής, Παν. Ιωαννίνων.
- **Μπερσίμης Σ.**, Καθηγητής, Πανεπιστήμιο Πειραιώς.
- **Μπουρνέτας Α.**, Καθηγητής, ΕΚΠΑ.
- **Μπραγουδάκης Ζ.**, Τράπεζα της Ελλάδος & Μεταδιδακτορικός Ερευνητής Πανεπιστημίου Αθηνών.
- **Μωυσιάδης Χ.**, Ομότιμος Καθηγητής, ΑΠΘ.
- **Νενές Γ.**, Καθηγητής, Πανεπιστήμιο Δυτικής Μακεδονίας
- **Ντζούφρας Ι.**, Καθηγητής, Οικονομικό Πανεπιστήμιο Αθηνών
- **Οικονόμου Π.**, Αναπληρωτής Καθηγητής, Παν. Πατρών
- **Παναγιωτάκος Δ.**, Καθηγητής, Χαροκόπειο Πανεπιστήμιο.
- **Παπαγεωργίου Ε.**, Καθηγήτρια, Παν. Δυτικής Αττικής.
- **Παπαδάτος Ν.**, Καθηγητής, ΕΚΠΑ.
- **Παπαδόπουλος Γ.**, Καθηγητής, Γεωπονικό Πανεπιστήμιο Αθηνών
- **Παπαϊωάννου Τ.**, Ομότιμος Καθηγητής, Παν. Πειραιώς και Παν. Ιωαννίνων και Επίτιμος Πρόεδρος ΕΣΙ.
- **Ρακιτζής Α.**, Επίκουρος Καθηγητής, Παν. Πειραιώς.
- **Συρακούλης Κ.**, Καθηγητής, Πανεπιστήμιο Θεσσαλίας
- **Τασιάς Κ.**, Επίκουρος Καθηγητής, Πανεπιστήμιο Δυτικής Μακεδονίας.
- **Τζαβελάς Γ.**, Αναπληρωτής Καθηγητής, Παν. Πειραιώς.
- **Τριανταφύλλου Ι.**, Επίκουρος Καθηγητής, Παν. Πειραιώς.
- **Τσακλίδης Γ.**, Καθηγητής, ΑΠΘ.
- **Φουσκάκης Δ.**, Καθηγητής, ΕΜΠ.
- **Χαλικιάς Μ.**, Καθηγητής, Πανεπιστήμιο Δυτικής Αττικής
- **Χαραλαμπίδης[†] Χ.Α.**, Ομότιμος Καθηγητής, ΕΚΠΑ.
- **Χατζηπαντελής Θ.**, Καθηγητής, ΑΠΘ.
- **Χριστοφίδης Τ.**, Καθηγητής, Πανεπιστήμιο Κύπρου.
- **Ψαρράκος Γ.**, Αναπληρωτής Καθηγητής, Παν. Πειραιώς.

Χαιρετισμοί / Welcome Notes

Αγαπητοί Συνάδελφοι,

Με μεγάλη χαρά και ενθουσιασμό σας καλωσορίζω στο 36ο Πανελλήνιο και 2ο Διεθνές Συνέδριο Στατιστικής που διοργανώνεται στις εγκαταστάσεις του Πανεπιστημίου Δυτικής Μακεδονίας και συνδιοργανώνεται από το Τμήμα Μηχανολόγων Μηχανικών του Πανεπιστημίου και το Ελληνικό Στατιστικό Ινστιτούτο. Το θέμα του φετινού συνεδρίου είναι η «Ανάλυση Δεδομένων στις Επιστήμες του Μηχανικού».

Το 36ο Πανελλήνιο και 2ο Διεθνές Συνέδριο Στατιστικής έχει την τιμή να φιλοξενεί έναν εκλεκτό κατάλογο ομιλητών, προσκεκλημένους καθηγητές από την Ελλάδα και το εξωτερικό, οι οποίοι θα μοιραστούν την εμπειρία και τη γνώση τους μαζί μας. Πιο συγκεκριμένα, στο πρόγραμμα του συνεδρίου περιλαμβάνονται παρουσιάσεις από εννέα (9) προσκεκλημένους ομιλητές: τον Καθηγητή κ. Kendall του Πανεπιστημίου Warwick των ΗΠΑ, τον Καθηγητή κ. Πήτα του Αριστοτέλειου Πανεπιστημίου Θεσσαλονίκης, τον Καθηγητή κ. Τριανταφυλλόπουλο του Πανεπιστημίου του Sheffield, τον Καθηγητή κ. Barbu του Πανεπιστημίου Rouen-Normandy, την Καθηγήτρια κα. Milošević του Πανεπιστημίου του Βελιγραδίου, τον Καθηγητή κ. Αφένδρα του Αριστοτέλειου Πανεπιστημίου Θεσσαλονίκης, τον Καθηγητή κ. Xing του Πανεπιστημίου του Sheffield, και τους Υποψήφιους Διδάκτορες κ. Mirroorian, του Πανεπιστημίου του Sheffield και την κα. Vigna του Πανεπιστημίου Chieti-Pescara.

Θα ήθελα να εκφράσω τις ειλικρινείς μας ευχαριστίες προς όλους τους προσκεκλημένους ομιλητές για την αποδοχή της πρόσκλησης.

Το Συνέδριο αποσκοπεί στην ανάδειξη του ρόλου της στατιστικής σε διάφορους τομείς και επιδιώκει τον ενεργό διάλογο, την ανταλλαγή ιδεών και την ανάπτυξη συνεργασιών μεταξύ ακαδημαϊκών ή/και επαγγελματιών από διαφορετικές επιστημονικές περιοχές. Στο πλαίσιο του συνεδρίου θα

παρουσιαστούν όλες οι σύγχρονες εξελίξεις και θα καλυφθεί ένα ευρύ φάσμα θεμάτων σχετικών με την ευρύτερη Επιστήμη της Στατιστικής. Το πρόγραμμα του συνεδρίου αποτελείται από ομιλίες, οργανωμένες σε θεματικές ενότητες, συζητήσεις, και ένα παράλληλο σεμινάριο.

Θέλουμε να εκφράσουμε τις ειλικρινείς μας ευχαριστίες προς όλους τους συνέδρους για τη συμμετοχή τους και την ευγενική τους προθυμία να παρουσιάσουν τις επιστημονικές τους ιδέες και τα ερευνητικά τους αποτελέσματα.

Τέλος, ευχαριστούμε θερμά τους χορηγούς μας για τη στήριξη τους στη διοργάνωση του Συνεδρίου και το Πανεπιστήμιο Δυτικής Μακεδονίας για τη διάθεση των σύγχρονων υποδομών του. Θέλοντας να διατηρήσουμε την επιτυχημένη παράδοση των προηγούμενων Συνεδρίων του Ελληνικού Στατιστικού Ινστιτούτου, ευελπιστούμε ότι το παρόν Συνέδριο θα προσφέρει σημαντικές συνεισφορές στην επιστημονική κοινότητα και ότι θα απολαύσετε τη διαμονή σας στην Κοζάνη με τη φυσική της ομορφιά, το επιβλητικό της τοπίο και την πλούσια πολιτιστική και γαστρονομική της παράδοση.

Εκ μέρους της Οργανωτικής Επιτροπής

Κωνσταντίνος Τασιάς

Επίκουρος Καθηγητής, Τμήματος Μηχανολόγων Μηχανικών,
Πανεπιστημίου Δυτικής Μακεδονίας

Dear Colleagues,

With great pleasure and enthusiasm, I welcome you to the 36th Panhellenic and 2nd International Statistics Conference, organized at the premises of the University of Western Macedonia and co-organized by the Department of Mechanical Engineering and the Hellenic Statistical Institute. Under the theme "Data Analytics in Engineering," this conference promises to be an enriching and insightful gathering.

Our conference is honored to host many distinguished speakers, including renowned professors from both Greece and abroad, who will share their experience and knowledge with us. Specifically, the conference program includes presentations from nine (9) invited speakers: Professor Kendall from the University of Warwick, USA; Professor Pitas from Aristotle University of Thessaloniki; Professor

Triantafyllopoulos from the University of Sheffield; Professor Barbu from the University of Rouen-Normandy; Professor Milošević from the University of Belgrade; Professor Afendras from Aristotle University of Thessaloniki; Professor Xing from the University of Sheffield, and Ph.D. candidates Mr. Mirpoorian from the University of Sheffield and Ms. Vigna from the University of Chieti-Pescara.

I would like to express our sincere thanks to all the invited speakers for accepting the invitation.

The conference aims to highlight the role of statistics in various fields and seeks active dialogue, exchange of ideas, and development of collaborations among academics and/or professionals from different scientific areas. Within the framework of the conference, all modern developments will be presented, covering a wide range of topics related to the broader Science of Statistics. The conference program consists of lectures, organized in thematic sections, discussions, and a parallel seminar.

We would like to express our sincere thanks to all the participants for their participation and their kind willingness to present their scientific ideas and research results.

Finally, we warmly thank our sponsors for their support in organizing the Conference and the University of Western Macedonia for providing its modern facilities. Hoping to maintain the successful tradition of previous conferences of the Hellenic Statistical Institute, we hope that this Conference will make significant contributions to the scientific community and that you will enjoy your stay in the city of Kozani with its natural beauty, impressive landscape, and rich cultural and gastronomic tradition.

On behalf of the Organizing Committee,

Konstantinos Tasiias
Assistant Professor, Department of Mechanical Engineering,
University of Western Macedonia

Αγαπητοί Συνάδελφοι και Φίλοι,

Με ιδιαίτερη χαρά σας καλωσορίζουμε στο 36^ο Πανελλήνιο και 2^ο Διεθνές Συνέδριο Στατιστικής που διοργανώνεται από το Τμήμα Μηχανολόγων Μηχανικών του Πανεπιστημίου Δυτικής Μακεδονίας, στην Κοζάνη, 16-19 Μαΐου 2024.

Τα συνέδρια του ΕΣΙ έχουν διανύσει πολύ δρόμο τα τελευταία 35 χρόνια και σήμερα συνεχίζουμε την προσπάθεια να ξεπεράσουμε τα στενά όρια της Ελληνικής επικράτειας με τη διεθνή διάσταση που προσδίδουμε στο Συνέδριο, χωρίς όμως να απολέσουμε την Ελληνική.

Με τη στήριξη όλων των μελών του τρέχοντος καθώς και του προηγούμενου ΔΣ, της Τοπικής Οργανωτικής Επιτροπής καθώς και της Επιστημονικής Επιτροπής έχει σχεδιαστεί ένα πλούσιο επιστημονικό πρόγραμμα με εκλεκτούς ομιλητές από την Ελλάδα και το εξωτερικό που με τις ομιλίες τους θα καλύψουν αφενός μεν το θέμα του συνεδρίου το οποίο είναι *«Ανάλυση Δεδομένων στις Επιστήμες του Μηχανικού»* αφετέρου δε ένα ευρύ θεματολόγιο που περιλαμβάνει όλες τις θεματικές ενότητες στον ευρύτερο χώρο της *Στατιστικής, των Πιθανοτήτων, των Στοχαστικών και Χρηματοοικονομικών Μαθηματικών* και της Ασφαλιστικής Επιστήμης.

Θερμές ευχαριστίες προς όλους και κυρίως προς το Πανεπιστήμιο Δυτικής Μακεδονίας το οποίο ανέλαβε την συνδιοργάνωση του Συνεδρίου, καθώς επίσης και προς τους χορηγούς για την στήριξη τους.

Εκ μέρους της Επιστημονικής Επιτροπής και του ΔΣ του ΕΣΙ
Αλεξ Καραγρηγορίου, Πρόεδρος ΕΣΙ

Esteemed Colleagues and Friends,

We welcome you with great pleasure to the 36th Panhellenic and 2nd International Statistics Conference, which is organized by the Department of Mechanical Engineering of the University of West Macedonia, during the period 16-19 May 2024.

GSI's conferences have come a long way over the past 35 years. This year, we continue our attempt to go beyond the Greek borders adding an international dimension, without of course, abandoning the Greek one.

With the support of the members of the current as well as the past Board, the Local Organizing Committee, and the Scientific Committee members, we prepared an enriched scientific program with distinguished speakers from Greece and abroad. These presentations will cover not only this year's conference theme on "*Data Analytics in Engineering*" but also a broad scientific spectrum by touching upon various topics in *Statistics, Probability, Stochastic and Financial Mathematics* and Insurance Science.

We wish to express our sincere appreciation to all and mostly to the University of West Macedonia, which is co-organizing this event, and to our sponsors for their kind support.

On behalf of the Scientific Committee and the Board of GSI

Alex Karagrigoriou
President of GSI

Προσκεκλημένοι Ομιλητές / Invited Speakers

2024 Theo Cacoullos Memorial Lecture

Wilfrid S. Kendall, Emeritus Professor, University of Warwick

Invited Speakers

Ioannis Pitas, Professor, Aristotle University of Thessaloniki

Kostas Triantafyllopoulos, Professor, University of Sheffield

Georgios Afendras, Associate Professor, Aristotle University of Thessaloniki

Vlad Stefan Barbu, Associate Professor, University of Rouen-Normandy

Bojana Milošević, Associate Professor, University of Belgrade

Wei W. Xing, Associate Professor, University of Sheffield

Navid Seyed Mirpoorian, Doctoral Candidate, University of Sheffield

Veronica Vigna, Doctoral Candidate, University of Chieti-Pescara

Εκπαιδευτικό Σεμινάριο / Seminar

Επιστημονικές Μέθοδοι και Τεχνικές του Στατιστικού Ελέγχου Ποιότητας

Στον επιχειρηματικό κόσμο η στρατηγική σημασία της ποιότητας έχει αναδειχθεί τα τελευταία χρόνια ως ο καθοριστικότερος ίσως παράγοντας επιτυχίας των προϊόντων και υπηρεσιών στην αγορά. Παρόλο όμως που οι επιχειρήσεις έχουν αντιληφθεί το ανταγωνιστικό πλεονέκτημα που αποκτούν μέσω της ποιότητας αυτό που δεν έχει γίνει, σε ανάλογο βαθμό, κατανοητό είναι η χρήση και αξιοποίηση των επιστημονικών μεθόδων μέσω των οποίων είναι δυνατό να επιτευχθεί, με τον πλέον αποτελεσματικότερο τρόπο, η υψηλή ποιότητα. Σε αυτό το πλαίσιο η επιστήμη της Στατιστικής προσφέρει σημαντικά εργαλεία για την αποτελεσματική λειτουργία όλων των σύγχρονων συστημάτων ποιοτικού ελέγχου.

Περίληψη: Στο σεμινάριο θα παρουσιαστούν οι κυριότερες επιστημονικές μέθοδοι και τεχνικές του στατιστικού ελέγχου ποιότητας με ιδιαίτερη έμφαση στην πρακτική εφαρμογή τους. Αρχικά θα παρουσιασθούν οι τεχνικές του ελέγχου παραγωγικής διαδικασίας οι οποίες έχουν ως στόχο την πρόληψη δημιουργίας ελαττωματικών προϊόντων ενώ στη συνέχεια θα αναφερθούν οι πλέον σύγχρονες τάσεις στο χώρο της ποιότητας με έμφαση: i) στη χρήση στατιστικών πειραμάτων για τη βέλτιστη εκλογή τιμών ελεγχόμενων παραμέτρων των προϊόντων και των παραγωγικών διαδικασιών και ii) στην τεχνική Six Sigma.

Ενότητες Σεμιναρίου:

- Μια σύντομη αναφορά σε έλεγχο αποδοχής και σχήματα δειγματοληψίας ΕΛΟΤ
- SPC, τα βασικά διαγράμματα ελέγχου
- Δείκτες δυνατότητας
- Six Sigma
- ANOVA με και χωρίς μεταβλητή πλαισίου
- Παραγοντικά Πειράματα 2^f
- Factorial Experiments 3^f
- Μη παραμετρική ανάλυση διακύμανσης

Σε ποιους απευθύνεται: φοιτητές και ερευνητές που θέλουν να εμβαθύνουν τη γνώση τους σε θέματα διασφάλισης ποιότητας με επιστημονικές τεχνικές άμεσα εφαρμόσιμες στα παραγωγικά συστήματα.

Προαπαιτούμενη γνώση: Βασικές έννοιες Στατιστικής

Μέγιστος αριθμός συμμετεχόντων: Τριάντα (30).

Εισηγητές Σεμιναρίου:

Η **Παρασκευή Καπετανοπούλου** είναι Επίκουρη Καθηγήτρια στο **Τμήμα Μηχανολόγων Μηχανικών** του **Πανεπιστημίου Δυτικής Μακεδονίας** ενώ είναι και ερευνήτρια / συνεργάτης στο Τμήμα Μηχανολόγων Μηχανικών του Αριστοτελείου Πανεπιστημίου Θεσσαλονίκης (Α.Π.Θ). Έχει λάβει δίπλωμα Μηχανολόγου Μηχανικού από το Α.Π.Θ, M.Sc στη Διοίκηση Παραγωγικών Συστημάτων και Ph.D. στη Λειτουργία και Διαχείριση Αντίστροφων Εφοδιαστικών Αλυσίδων επίσης από το Α.Π.Θ. Έχει εργαστεί ως μεταδιδακτορικός ερευνητής στο INSEAD στη Γαλλία και ως επισκέπτρια καθηγήτρια στο Τμήμα Μηχανολόγων Μηχανικών του Α.Π.Θ. Σε μεταπτυχιακό επίπεδο έχει διδάξει στο Ελληνικό Ανοικτό Πανεπιστήμιο τη θεματική ενότητα «Προηγμένα Εργαλεία και Μέθοδοι για τον Έλεγχο της Ποιότητας» ενώ επί σειρά ετών διατελεί συντονίστρια στο Μεταπτυχιακό Πρόγραμμα «Διοίκηση, Τεχνολογία και Ποιότητα» του Ανοικτού Πανεπιστημίου Κύπρου.

Ο **Αλέξανδρος Καραγρηγορίου** είναι Καθηγητής Πιθανοτήτων και Στατιστικής στο **Τμήμα Στατιστικής και Αναλογιστικών-Χρηματοοικονομικών Μαθηματικών** του **Πανεπιστημίου Αιγαίου**. Σπούδασε στο Πανεπιστήμιο Πατρών και στο Πανεπιστήμιο του Maryland, ΗΠΑ (Ph.D,1992). Εκτός του ιδιωτικού (Ναυτιλία, 1980-1986) και του δημοσίου τομέα (Υπουργείο Γεωργίας, ΗΠΑ, 1986-1992), στον ακαδημαϊκό χώρο έχει εργαστεί στο Πανεπιστήμιο του Maryland και στο Ινστιτούτο Στατιστικών Επιστημών (Taiwan) και δίδαξε κατά χρονολογική σειρά στα Πανεπιστήμια Πατρών, Maryland, Αθήνας, Κύπρου (απ' όπου και συνταξιοδοτήθηκε), στο Ελληνικό Ανοικτό Πανεπιστήμιο και στο Πανεπιστήμιο Αιγαίου (2014-2018, Τμήμα Μαθηματικών και 2018-2024 Τμήμα Στατιστικής και Αναλογιστικών-Χρηματοοικονομικών Μαθηματικών).

Πρόγραμμα Συνεδρίου–Conference Program

Πέμπτη – Thursday 16/5/24

ΑΜΦΙΘΕΑΤΡΟ - AMPHITHEATER

13:30 - 14:30	Εγγραφή Συνέδρων – Registration of the Participants
14:30 – 15:00	Έναρξη του Συνεδρίου – Conference Opening Προεδρεύων/Chair: Κ. Τασιάς/K. Tasias
	Χαιρετισμοί – Welcome Addresses
15:00 - 16:15	Εναρκτήριες Κεντρικές Ομιλίες – Plenary Talks I.0 Προεδρεύων/Chair : N. Balakrishnan
15:00-15:45	W. Kendall <i>Cacoullos Memorial Lecture 2024: Three centuries of random lines: from Buffon's needle to scale-invariant networks</i>
15:45-16:15	G. Afendras <i>Kendall curve-based index of dependence and test for independence between random variables</i>
16:15-16:45	ΔΙΑΔΕΙΜΜΑ – ΚΑΦΕΣ – COFFEE BREAK
16:45 - 18:50	Συνεδρία/Session I.1 Στατιστική / Statistics Chair/Προεδρεύων: Α. Batsidis/ Α. Μπασιδής
16:45-17:10	G. Iliopoulos <i>Linear models for continuous paired comparison data</i>
17:10-17:35	G. Papageorgiou <i>Temporal Dynamics of Time Series Topics: A Convex Hull Visualization Approach</i>
17:35-18:00	C. Meselidis <i>Robust Estimation and Testing via a General Family of Measures</i>
18:00-18:25	K. Fountoukidis <i>Joint monitoring of origin and scale of a shifted exponential process with run sum control charts</i>
18:25-18:50	K. Bourazas <i>Bayesian sample size determination for detecting heterogeneity in multi-site replication studies</i>
19:00	WELCOME RECEPTION

Παρασκευή - Friday 17/5/24

09:00 - 14:00	Excursion / Εκδρομή
14:00 - 15:30	<i>ΜΕΣΗΜΒΡΙΝΗ ΔΙΑΚΟΠΗ – LUNCH</i>
Αίθουσα/Room A 15:30-17:10	Session / Συνεδρία Π.1 Εφαρμοσμένες Πιθανότητες / Applied Probability Chair/Προεδρεύων: G. Psarrakos/Γ. Ψαρράκος
15:30-15:50	Π. Μπομποτάς <i>Διατήρηση ιδιοτήτων γήρανσης του μεγίστου από τυχαίο αριθμό τυχαίων μεταβλητών</i>
15:50-16:10	Σ. Λαφνής <i>Η διαδρομή από τις ροές στις ασθενείς ροές: Γενικεύσεις κατανομών ροών τάξης k και εφαρμογές</i>
16:10-16:30	Γ. Βασιλειάδης <i>Το μη ομογενές Μαρκοβιανό σύστημα διακριτού χρόνου με χωρητικότητες ως επιδημιολογικό μοντέλο</i>
16:30-16:50	Σ. Τζανίνης <i>Μία επέκταση της αναδρομής Panjer για μεικτές σύνθετες κατανομές</i>
16:50-17:10	Ε. Κωνσταντινίδης <i>Εξαγωγή χρήσιμων στατιστικών μεγεθών από μια περιοδική στοχαστική μεταβλητή με μηδενικούς μέσους όρους συνόλου</i>
Αίθουσα /Room B 15:30-17:10	Special Session / Ειδική Συνεδρία Π.2 Big Data and Statistical Analysis in Geosciences Chair/Προεδρεύων: I. Καπαγερίδης/I. Kapageridis
15:30-15:50	C. Gravalos <i>Enhancing Safety and Efficiency in Underground Marble Quarries through Statistical Analysis of Geotechnical Data</i>
15:50-16:10	E. Varouchakis <i>Enhancing Statistical Analysis of Asymmetric Geoscience Data using Gaussian Anamorphosis</i>
16:10-16:30	A. Patra <i>Integration of big data analytics in digital twinning of mineral deposits</i>
16:30-16:50	V. Drakopoulos <i>On the application of fractal interpolation to the modelling and reconstruction of seismic data</i>
16:50-17:10	I. Kapageridis <i>Spatial Declustering of Exploration Data in Marble Resource Estimation from Irregular Drilling Patterns</i>
17:10-18:00	<i>ΔΙΑΔΕΙΜΜΑ – ΚΑΦΕΣ – COFFEE BREAK</i>

Φουαγιέ/Foyer 17:10-18:00	Poster Session
Z. Kyrana	<i>Dimensionality reduction on data cubes: Application and comparison of various analysis strategies</i>
N. Papafilippou	<i>Strategies for Dimensionality Reduction of Categorical Data</i>
E. Pratsinakis	<i>Strategies for dimensionality reduction of mixed-type data: encoding strategies and techniques</i>
A. Mroukouvala / A. Papakyriellou	<i>Nowcasting techniques for European energy data</i>
Δ. Πανάρετος	<i>Η συμβολή της Τεχνητής Νοημοσύνης στην Λήψη Αποφάσεων: Αλγόριθμοι και Εφαρμογές</i>
Θ. Δημητριάδης/ Γ. Τσακλίδης	<i>Αποτίμηση δικαιωμάτων προαίρεσης με την μέθοδο martingale. Εφαρμογή και σύγκριση σε πραγματικά δεδομένα των μοντέλων Heston και Black- Scholes</i>
20:15 - 23:00	ΕΠΙΣΗΜΟ ΔΕΙΠΝΟ – CONFERENCE DINNER

Σάββατο-Saturday 18/5/24

Αίθουσα/Room A 09:45-11:45	Συνεδρία IV.1: Financial and Actuarial Mathematics Chair/Προεδρεύων: Ch. Evangelaras / X. Ευαγγελάρας
09:45-10:05	Th. Pappas <i>A contribution to the comparative evaluation of efficiency indicators in Eurozone using DEA methodology</i>
10:05-10:25	V. Varlagkas <i>Pitfalls in interpreting the predictive power of technical trading rules and weak-form market efficiency test results. An empirical investigation for the Athens Stock Exchange.</i>
10:25-10:45	A. Bozikas <i>A comparison of statistical models for mortality graduation</i>
10:45-11:05	P. Hatzopoulos <i>A novel framework of stochastic mortality model selection</i>
11:05-11:25	G. Pitselis <i>Some New Developments in the Credible Distribution Estimation</i>
11:25-11:45	S. Nikolakopoulos <i>Extending the Mann–Kendall test to allow for measurement uncertainty</i>
Αίθουσα/Room B 09:45-11:45	Special Session – Ειδική Συνεδρία IV.2 Ανάλυση Πολυμεταβλητών Χρονοσειρών/ Analysis of Multivariate Time Series Chair/Προεδρεύων: Δ. Κουγιουμτζής/ D. Kugiumtzis

09:45-10:05	Ε. Παπαδοπούλου <i>Πρόβλεψη πολυμεταβλητών χρονοσειρών με Reservoir Computing και LSTM</i>
10:05-10:25	Π. Πετρίδης <i>Σχηματισμός αραιού μοντέλου παλινδρόμησης πολυμεταβλητών χρονοσειρών με μέτρα πληροφορίας</i>
10:25-10:45	Γ. Μυλλής <i>Βραχυπρόθεσμες Προβλέψεις Ζήτησης Νερού από Μονομεταβλητές Χρονοσειρές Σταθμών Δεξαμενών με χρήση αλγορίθμων και στατιστικών εργαλείων</i>
10:45-11:05	Α. Αγγειοπλάστης <i>Πρόβλεψη απόδοσης μαθητών Δευτεροβάθμιας εκπαίδευσης βασισμένη σε χρονοσειρές βαθμολογιών με χρήση εργαλείων λογισμικού ανοιχτού κώδικα</i>
11:05-11:25	Β. Σαραντίδου <i>Μοντέλο πρόβλεψης του ΑΕΠ με χρήση αιτιότητας κατά Granger εκτιμώμενη από τη μερική συσχέτιση</i>
11:25-11:45	Σ. Μαλεφάκη <i>Μοντελοποίηση και πρόβλεψη των εκπομπών διοξειδίου του άνθρακα στην Ευρωπαϊκή Ένωση με χρήση μοντέλων ARIMA και τεχνητών νευρωνικών δικτύων</i>
11:45 - 12:15	ΔΙΑΛΕΙΜΜΑ – ΚΑΦΕΣ – COFFEE BREAK
Αίθουσα /Room A 12:15– 14:00	Session –Συνεδρία V.1: Στατιστική Μεθοδολογία και Εφαρμογές / Statistical Methods & Applications Chair/Προεδρεύων: P. Economou / Π. Οικονόμου
12:15-12:35	V. Chasiotis <i>Supervised statistical (machine) learning for domain estimation with business survey data</i>
12:35-12:55	D. Kavargyris <i>A graph-based hybrid analysis of the echo chambers phenomenon: Insights from the Baltimore Riots</i>
12:55-13:15	V. Papageorgiou <i>Exploring Disease Dynamics: A Markovian SIRD Model for Transmission Assessment</i>
13:15-13:35	A. Papachristos <i>Incorporating Subjective Survival Information in Bayesian Mortality Models</i>
13:35-13:55	D. Kugiumtzis <i>Community Detection in Correlation Networks from Multivariate Time Series</i>
Αίθουσα/Room B 12:15 – 14:00	Ειδική Συνεδρία / Special Session V.2: ISSP in Greece - social surveys and data for a deeper understanding of important social issues Chair/Προεδρεύων : I. Andreadis/ I. Ανδρεάδης

12:15-12:35	I. Andreadis <i>ISSP in Greece – social surveys and data for a deeper understanding of important social issues</i>
12:35-12:55	N. Sivenou <i>The survey of ISSP 2022. Political attitudes towards changing gender roles and family</i>
12:55-13:15	S. Tseliou <i>The survey of ISSP 2023. The impact of National Identity and Citizenship on political attitudes</i>
13:15-13:35	E.-M. Papachristou <i>The pilot survey of ISSP 2024. Political attitudes and digital societies</i>
13:35-13:55	G.A. Millias <i>The pilot survey of ISSP 2025. Work Orientations</i>
14:00 - 15:30	ΜΕΣΗΜΒΡΙΝΗ ΔΙΑΚΟΠΗ – LUNCH

ΑΜΦΙΘΕΑΤΡΟ - AMPHITHEATER

	Invited Session V1.1
15:30 - 17:15	Chair/Προεδρεύων : E. Angelis/E. Αγγελής
15:30 - 16:15	Ioannis Pitas <i>AI Science and Engineering: A new scientific discipline?</i>
16:15 – 16:45	Vlad Barbu <i>Sequential interval reliability for semi-Markov systems: modeling and asymptotic properties</i>
16:45 - 17:15	Veronica Vigna <i>Forecasting the power generation Mix in Italy based on Grey Markov Models</i>
17:15 - 17:45	ΔΙΑΔΕΙΜΜΑ – ΚΑΦΕΣ – COFFEE BREAK
	Invited Virtual Special Session VI.2 – Advanced Topics in Statistics with Applications
17:45-19:45	Chair/Προεδρεύων: E.N. Kalligeris/E.N. Καλλιγέρης
17:45-18:15	K. Triantafyllopoulos <i>Bayesian Inference of Spatio-Temporal Dynamic Models with Application to Air-Pollution Data</i>
18:15-18:45	W. W. Xing <i>Modern Statistical Methods for Improving the Yield of Advanced Node Integrated Circuits</i>
18:45-19:15	S. N. Mirpoorian <i>Machine learning algorithmic-trading models based on the mean reversion hypothesis of financial markets</i>
19:15-19:45	B. Milošević <i>On the goodness-of-fit tests for censored data with cure fraction</i>

Κυριακή-Sunday 19/5/24

Αίθουσα/Room A 09:30-11:10	Session –Συνεδρία VII.1 Στατιστική και Εφαρμοσμένη Στατιστική / Statistics & Applied Statistics Chair-Προεδρεύων: Γ. Ηλιόπουλος/G. Pliopoulos
09:30-09:50	Α. Μπατσίδης <i>Ένα Μπεύζιανό αυτό-εκκινούμενο διάγραμμα ελέγχου για την άμεση ανίχνευση πολυδιάστατων ακραίων τιμών</i>
09:50-10:10	Σ. Δαφνής <i>Βελτιωμένα διαγράμματα ελέγχου τύπου Shewhart</i>
10:10-10:30	Π. Οικονόμου <i>Μη παραμετρικές μέθοδοι παρακολούθησης πολυδιάστατων ροών δεδομένων</i>
10:30-10:50	Χ. Νάκας <i>Τυπική αξιολόγηση συμφωνίας μεταξύ συσκευών ανοιχτού κώδικα και αντίστοιχων αναφοράς, βιομηχανικής χρήσης, με εφαρμογή σε μετρητές pH.</i>
10:50-11:10	Χ. Ευαγγελάρας <i>Mirror-image ορθογώνιοι σχηματισμοί τριών επιπέδων: Κατασκευή, αξιολόγηση και εφαρμογές</i>
Αίθουσα/Room B 09:30-11:10	Session/Συνεδρία VII.2 Εφαρμοσμένη Στατιστική / Applied Statistics Chair - Προεδρεύουσα: P. Kapetanopoulou / Π. Καπετανοπούλου
09:30-09:50	C. Kitsos <i>Simulation Study for the family of Generalized Normal $N\gamma(\mu, \sigma^2)$</i>
09:50-10:10	A. Georgakis <i>A multivariate Fay-Herriot model for forest management units estimates</i>
10:10-10:30	T. Chadjipantelis <i>Exploring Political Behavior of First Time Voters through Hierarchical Cluster Analysis and Correspondence Factor</i>
10:30-10:50	D. Papadopoulos <i>simAlpha. A new algorithm to help scientists to assess questionnaire reliability</i>
10:50-11:10	V. Simpseris <i>VoteMatch Europe 2024. The position of the Greek Political Parties on European Issues</i>
11:10 - 11:30	ΔΙΑΛΕΙΜΜΑ – ΚΑΦΕΣ – COFFEE BREAK
Αίθουσα/Room A 11:30-13:30	Συνεδρία/Session – VII.3 Εφαρμοσμένες Πιθανότητες / Applied Probability Chair-Προεδρεύων: C. Nakas / Χ. Νάκας

11:30-11:50	E. Skamnia <i>Process Monitoring of Bivariate Poisson Data</i>
11:50-12:10	K. Bavelis <i>Applications of Data Analytics for the maintenance and improvement of the Electricity Distribution Network</i>
12:10-12:30	A. Karagrigoriou <i>On the analysis of competing risks via the MSS methodology</i>
12:30-12:50	I. Triantafyllou <i>An Archimedean copulas-based approach for m-consecutive-k-out-of-n: F systems under exchangeability</i>
12:50-13:10	C. Kitsos <i>Information Measures for the γ-order Generalized Normal</i>
13:10-13:30	G. Psarrakos <i>On the usual stochastic order and covariance identities</i>
Αίθουσα/Room B 11:30-13:30	Συνεδρία / Session – VII.4 Εφαρμοσμένη Στατιστική / Applied Statistics Chair-Προεδρεύων: Θ. Χατζηπαντελής/ T. Chadjiapantelis
11:30-11:50	Θ. Μητράκος <i>Ανάλυση δεδομένων για την ελληνική αγορά ακινήτων</i>
11:50-12:10	M. Παπαγεωργίου <i>Δειγματοληψία σε Συμμετρικά Δίκτυα: Αμεροληψία Εκτιμητριών</i>
12:10-12:30	B. Ντούσκα <i>Η επίδραση των υπαίθριων δραστηριοτήτων και της φυσικής αγωγής σε μαθητές με μεταναστευτικό υπόβαθρο</i>
12:30-12:50	Π. Μπίρης <i>Επιλογή μοντέλου του κινητήρα του πλοίου εφαρμόζοντας τεχνικές Κατηγοριοποίησης μηχανικής μάθησης και χρησιμοποιώντας μεθόδους καταλογισμού και μείωσης διαστάσεων των δεδομένων</i>
12:50-13:10	Π. Φιλοκόστας <i>Στατιστική ανάλυση για το Projectification στη συντήρηση αεροσκαφών</i>
13:10-13:30	E. Δημαρέλλης <i>Εφαρμογές νευρωνικών δικτύων γράφων στην έρευνα για τη νόσο Alzheimer</i>
ΑΜΦΙΘΕΑΤΡΟ AMPHITHEATER 13:30	ΛΗΞΗ ΣΥΝΕΔΡΙΟΥ – CONFERENCE CLOSING - Προτάσεις για το Συνέδριο Στατιστικής του 2025 - Συζήτηση

ΠΕΡΙΛΗΨΕΙΣ
στα Ελληνικά

ABSTRACTS
in Greek

**Πρόβλεψη απόδοσης μαθητών Δευτεροβάθμιας εκπαίδευσης
βασισμένη σε χρονοσειρές βαθμολογιών με χρήση εργαλείων
λογισμικού ανοιχτού κώδικα**

**Αθανάσιος Αγγειοπλάστης, Αλκιβιάδης Τσιμπίρης,
Παπαϊωάννου Νικόλαος**

Διεθνές Πανεπιστήμιο Ελλάδος Τμήμα Μηχανικών Πληροφορικής,
Υπολογιστών και Τηλεπικοινωνιών, Σερρών
e-mail aagiop@gmail.com, atsimpiris@ihu.gr

Σκοπός της εργασίας είναι η διερεύνηση της χρήσης εφαρμογών λογισμικού ανοιχτού κώδικα στην πρόβλεψη της ακαδημαϊκής επίδοσης των μαθητών δευτεροβάθμιας εκπαίδευσης χρησιμοποιώντας τα ιστορικά δεδομένα βαθμολογίας τους. Τα δεδομένα καλύπτουν μία χρονική περίοδο 10 ετών και περιλαμβάνουν προφορικές και γραπτές βαθμολογίες 3.000 μαθητών λυκείου και των τριών τάξεων. Η μελέτη περιλαμβάνει τη συλλογή, καθαρισμό και στατιστική ανάλυση των δεδομένων που σχετίζονται με την πρόοδο των μαθητών με την πάροδο του χρόνου καθώς και τη χρήση εργαλείων ανοιχτού κώδικα για προγνωστική ανάλυση. Ο στόχος είναι να αναπτυχθούν μοντέλα και τεχνικές που μπορούν να προβλέψουν τα μελλοντικά ακαδημαϊκά αποτελέσματα των μαθητών με βάση τα προηγούμενα πρότυπα απόδοσής τους. Αξιοποιώντας λογισμικό ανοιχτού κώδικα, η εργασία δίνει μία διαδικασία πρόβλεψης αναπαραγώγιμη και προσαρμόσιμη για εκπαιδευτικά ιδρύματα όλων των και τύπων και βαθμίδων.

Το μη ομογενές Μαρκοβιανό σύστημα διακριτού χρόνου με χωρητικότητες ως επιδημιολογικό μοντέλο

Γ. Βασιλειάδης¹, Β. Ε. Παπαγεωργίου²

¹Τμήμα Μαθηματικών, Πανεπιστήμιο Δυτικής Μακεδονίας,

²Τμήμα Μαθηματικών, Αριστοτέλειο Πανεπιστήμιο Θεσσαλονίκης
gvasiliadis@uowm.gr

Στην παρούσα εργασία γίνεται λόγος αρχικά, για το μη ομογενές Μαρκοβιανό σύστημα (ΜΟΜΣ) διακριτού χρόνου με χωρητικότητες στις καταστάσεις του. Θεωρούμε ότι το σύστημα είναι κλειστό και κάθε κατάσταση i του χώρου καταστάσεων $S = \{1, 2, K, k\}$, παρουσιάζει μία πεπερασμένη χωρητικότητα c_i , $i = 1, 2, K, k$. Προκειμένου να εξετάσουμε τη συμπεριφορά του συστήματος στην εξέλιξη του χρόνου, δημιουργείται ένας επαναληπτικός αλγόριθμος υπολογισμού της κατανομής των μεγεθών των καταστάσεων με τη βοήθεια των παραγοντικών ροπών των μεγεθών των καταστάσεων. Στη συνέχεια, το μοντέλο αυτό προσαρμόζεται στα χαρακτηριστικά ενός επιδημιολογικού υποδείγματος, το οποίο κατατάσσει τον πληθυσμό σε τρεις καταστάσεις (ευπαθή μέλη, μολυσμένα μέλη, μέλη σε καραντίνα), οι οποίες μπορεί να παρουσιάζουν χωρητικότητα. Το συγκεκριμένο μοντέλο χρησιμοποιείται για να περιγράψει τη μετάδοση ενός κακόβουλου λογισμικού σε ένα κλειστό δίκτυο υπολογιστών, όπου η κατάσταση καραντίνας παρουσιάζει πεπερασμένη χωρητικότητα που σχετίζεται με τον περιορισμένο αριθμό τεχνικών που είναι υπεύθυνοι για τη διατήρηση της λειτουργικότητας του συστήματος. Με τη βοήθεια ενός ΜΟΜΣ διακριτού χρόνου με τρεις καταστάσεις και χωρητικότητες, εξετάζεται η συμπεριφορά του μεγέθους κάθε κατάστασης στην εξέλιξη του χρόνου.

Η ΔΙΑΔΡΟΜΗ ΑΠΟ ΤΙΣ ΡΟΕΣ ΣΤΙΣ ΑΣΘΕΝΕΙΣ ΡΟΕΣ: ΓΕΝΙΚΕΥΣΕΙΣ ΚΑΤΑΝΟΜΩΝ ΡΟΩΝ ΤΑΞΗΣ k ΚΑΙ ΕΦΑΡΜΟΓΕΣ

Σπύρος Δ. Δαφνής

Τμήμα Στατιστικής και Αναλογιστικών – Χρηματοοικονομικών
Μαθηματικών, Πανεπιστήμιο Αιγαίου
sdafnis@aegean.gr

Οι κατανομές που σχετίζονται με ροές επιτυχιών σε ακολουθίες δίτιμων δοκιμών είναι γνωστές ως κατανομές τάξης k και έχουν μελετηθεί εκτεταμένα στη βιβλιογραφία. Στην παρούσα εργασία κάνουμε μία ανασκόπηση αποτελεσμάτων που έχουν δημοσιευθεί την τελευταία διετία και αφορούν σε κατανομές συγκεκριμένου τύπου γενικευμένων ροών, των ασθενών ροών και εφαρμογών τους.
Λέξεις Κλειδιά: Κατανομές διωνυμικού τύπου, κατανομές χρόνου αναμονής, Μαρκοβιανή αλυσίδα, αξιοπιστία συστημάτων, στατιστικός έλεγχος διεργασιών.

ΒΕΛΤΙΩΜΕΝΑ ΔΙΑΓΡΑΜΜΑΤΑ ΕΛΕΓΧΟΥ ΤΥΠΟΥ SHEWHART

**Σπύρος Δ. Δαφνής¹, Θεόδωρος Πεردίκης², Μάρκος Β.
Κούτρας³**

¹Τμήμα Στατιστικής και Αναλογιστικών – Χρηματοοικονομικών
Μαθηματικών, Πανεπιστήμιο Αιγαίου

²Τμήμα Στατιστικής, Οικονομικό Πανεπιστήμιο Αθηνών

³Τμήμα Στατιστικής και Ασφαλιστικής Επιστήμης, Πανεπιστήμιο
Πειραιώς

sdafnis@aegean.gr, theoerdik@aueb.gr, mkoutras@unipi.gr

Στην παρούσα εργασία εισάγουμε νέα διαγράμματα ελέγχου τύπου Shewhart με κανόνες ασθενών ροών. Μελετάμε τα χαρακτηριστικά των νέων διαγραμμάτων και την επίδοσή τους συγκριτικά με άλλα διαγράμματα στη βιβλιογραφία, αποδεικνύοντας την ανωτερότητά τους για την ανίχνευση μικρών ή μεσαίων αλλαγών στη μέση τιμή της διεργασίας.

Λέξεις Κλειδιά: Κανόνες ροών, μέσο μήκος ροής, κατανομή χρόνου αναμονής.

ΕΦΑΡΜΟΓΕΣ ΝΕΥΡΩΝΙΚΩΝ ΔΙΚΤΥΩΝ ΓΡΑΦΩΝ ΣΤΗΝ ΕΡΕΥΝΑ ΓΙΑ ΤΗ ΝΟΣΟ ΤΟΥ ALZHEIMER

Δημαρέλλης Ε.^{1,*}, Γεωργίου Κ.², Αγγελής Ε.²

¹Τμήμα Ιατρικής, Αριστοτέλειο Πανεπιστήμιο Θεσσαλονίκης,
54124 Θεσσαλονίκη, Ελλάδα;

²Σχολή Πληροφορικής, Αριστοτέλειο Πανεπιστήμιο Θεσσαλονίκης,
54124 Θεσσαλονίκη, Ελλάδα;

konsgeor@csd.auth.gr (Γ.Κ.) lef@csd.auth.gr (Α. Ε.)

*correspondence: researcher@dimarellis-emm.gr (Δ. Ε.)

Τα Νευρωνικά Δίκτυα Γράφων είναι μια κατηγορία Τεχνητών Νευρωνικών Δικτύων Βαθιάς Μάθησης που δέχονται ως είσοδο γράφους αντί για απλά διανύσματα. Σημαντικό τους πλεονέκτημα είναι ότι επιτρέπουν την εξόρυξη γνώσης από μη ευκλείδεια δεδομένα, που δεν μπορούν να αξιοποιηθούν από τα κλασικά τεχνητά νευρωνικά δίκτυα. Από το 2019, έχουν αρχίσει να αναπτύσσονται όλο και περισσότερα μοντέλα τέτοιου τύπου με σκοπό την έρευνα για την νόσο του Alzheimer. Η νόσος του Alzheimer είναι μια νευροεκφυλιστική ασθένεια και αποτελεί την πιο συχνή αιτία άνοιας με προοδευτική γνωστική δυσλειτουργία και τελικό στάδιο τον θάνατο. Δυστυχώς, δεν έχει βρεθεί ακόμη αποτελεσματική θεραπεία και η παθοφυσιολογία της νόσου παραμένει άγνωστη, καθώς πρόκειται για μια πολυπαραγοντική ασθένεια. Έτσι, τα Νευρωνικά Δίκτυα Γράφων με την ικανότητά τους να συνθέτουν και να αναλύουν πολυτροπικά δεδομένα, όπως γενετικούς δείκτες, μαγνητικές τομογραφίες και γνωστικές αξιολογήσεις, τα καθιστούν ένα πολύτιμο εργαλείο α) για την βαθύτερη κατανόηση της νόσου, β) για την έγκαιρη διάγνωσή της αλλά και γ) για την ανάπτυξη κατάλληλων φαρμάκων. Σε αυτή τη βιβλιογραφική ανασκόπηση θα αναδείξουμε μερικές από τις δυνατότητες που προσφέρουν τα συγκεκριμένα δίκτυα στην έρευνα για το Alzheimer, ώστε να οξύνουμε την διεπιστημονική συνεργασία και να προωθήσουμε μία νέα πιθανή λύση στα άλυτα προβλήματα που μας θέτει η συγκεκριμένη νόσος.

Αποτίμηση δικαιωμάτων προαίρεσης με την μέθοδο martingale. Εφαρμογή και σύγκριση σε πραγματικά δεδομένα των μοντέλων Heston και Black- Scholes

Δημητριάδης Θεμιστοκλής, Τσακλίδης Γεώργιος
Τμήμα Μαθηματικών, ΑΠΘ
dimthedim@math.auth.gr, tsaklidi@math.auth.gr

Στην παρούσα εργασία αφού γίνει σύντομη αναφορά στις ιδιότητες της επίπεδης κίνησης Brown και της επίλυσης της στοχαστικής διαφορικής εξίσωσης που αποδίδει την εξέλιξη της τιμής μιας μετοχής, στη συνέχεια αποτιμάται η αξία των απλών δικαιωμάτων προαίρεσης με την μέθοδο martingale και με κατάληξη στο μοντέλο των Black-Scholes. Γίνεται αναφορά στα εξωτικά δικαιώματα φράγματος και παρουσιάζεται παράδειγμα αποτίμησης δικαιωμάτων (απλών και φράγματος) μέσω μεθοδολογίας Monte-Carlo. Τέλος, παρουσιάζεται συνοπτικά το μοντέλο στοχαστικής διακύμανσης του Heston και γίνεται σύγκριση, πάνω σε πραγματικά δεδομένα, με το κλασικό μοντέλο Black-Scholes.

Mirror-image ορθογώνιοι σχηματισμοί τριών επιπέδων:

Κατασκευή, αξιολόγηση και εφαρμογές

Χαράλαμπος Ευαγγελάρας και Μάρκος Κούτρας

Τμήμα Στατιστικής και Ασφαλιστικής Επιστήμης, Παν. Πειραιώς

hevangel@unipi.gr, mkoutras@unipi.gr

Ένας σχεδιασμός τριών επιπέδων για τη μελέτη k παραγόντων με τη χρήση n πειραματικών δοκιμών, μπορεί να παρουσιαστεί, χωρίς βλάβη της γενικότητας, ως ένας $n \times k$ πίνακας με στοιχεία 0, 1 και 2. Κάθε γραμμή αυτού του πίνακα είναι ένα διάνυσμα $1 \times k$ με στοιχεία 0, 1, και 2. Δύο γραμμές $(a_1 \ a_2 \ \dots \ a_k)$ και $(b_1 \ b_2 \ \dots \ b_k)$ του πίνακα αυτού αποτελούν ένα ζεύγος mirror-image γραμμών αν $(a_i + b_i) \bmod 3 = z$ για κάθε $i = 1, 2, \dots, k$, δηλαδή, η μία γραμμή είναι ο «αντικατοπτρισμός» της άλλης με «κάτοπτρο» το $1 \times k$ διάνυσμα το οποίο έχει όλα τα στοιχεία του ίσα με c , $c = (3 - z) \bmod 3$, $z = 0$ ή 1 ή 2. Το διάνυσμα αυτό καλείται κεντρική γραμμή. Ένας $n \times k$ σχεδιασμός θα καλείται mirror-image σχεδιασμός, αν οι γραμμές του αποτελούνται αποκλειστικά από ζεύγη mirror-image γραμμών ή και κεντρικών γραμμών.

Στην εργασία αυτή κατασκευάζουμε και αξιολογούμε mirror-image ορθογώνιους σχηματισμούς τριών επιπέδων με $n = 18, 27$ και 36 πειραματικές δοκιμές.

Acknowledgements: This research is co-financed by Greece and the European Union (European Social Fund-ESF) through the Operational Programme «Human Resources Development, Education and Lifelong Learning 2014-2020» in the context of the project “Innovative Titanium Nanoparticles for Development of Autocleaning and Auto antibacterial Application” (MIS 5131364, TiClean).

This work has been partly supported by the University of Piraeus Research Center.

Αναφορές

- Hedayat, A. S., Sloane, N. J. A. and Stufken, J. (1999). *Orthogonal Arrays: Theory and Applications*, Springer-Verlag, New York.
- Schoen, E. D., Eendebak, P. T. and Nguyen, M. V. M. (2010). Complete enumeration of pure-level and mixed-level orthogonal arrays. *J. Combin. Designs*, **18**, 123-140.

**Μοντελοποίηση και πρόβλεψη των εκπομπών διοξειδίου
του άνθρακα στην Ευρωπαϊκή Ένωση με χρήση μοντέλων
ARIMA και τεχνητών νευρωνικών δικτύων**

**Μασούρα Μελομένη, Μαλεφάκη Σόνια, Βασίλειος Α.
Γεωργίου**

Τμ. Μηχανολόγων & Αεροναυπηγών Μηχανικών, Παν. Πατρών,
Τμήμα Στατιστικής και Ασφαλιστικής Επιστήμης, Παν. Πειραιώς
smalefaki@upatras.gr

Η βιώσιμη ανάπτυξη είναι ένα από τα πιο πολυσυζητημένα θέματα της τελευταίας δεκαετίας και ορίζεται ως η ανάπτυξη που καλύπτει τις ανάγκες του παρόντος χωρίς να διακυβεύεται η ικανότητα των μελλοντικών γενεών να καλύψουν τις δικές τους ανάγκες. Ένας από τους βασικότερους πυλώνες της βιώσιμης ανάπτυξης είναι ο περιβαλλοντικός. Πιο συγκεκριμένα, η περιβαλλοντική βιωσιμότητα εξασφαλίζει τη διατήρηση του φυσικού περιβάλλοντος σε πλήρη λειτουργικότητα και χρηστικότητα τόσο στο παρόν όσο και στο μέλλον. Παρόλα αυτά, από τα μέσα του εικοστού αιώνα, η εκβιομηχάνιση και η διαρκώς αυξανόμενη κατανάλωση ενέργειας οδήγησαν στην αύξηση των αερίων του θερμοκηπίου. Τα αέρια του θερμοκηπίου, ειδικά οι εκπομπές διοξειδίου του άνθρακα (CO₂), θεωρούνται ως μία από τις βασικές αιτίες της υπερθέρμανσης του πλανήτη, η οποία, με τη σειρά της, οδηγεί σε δυσμενείς περιβαλλοντικές επιπτώσεις που συνεπάγονται κίνδυνο ξηρασίας, πυρκαγιών, πλημμυρών, τήξης παγετώνων κ.λπ.. Η συγκέντρωση CO₂ στην ατμόσφαιρα η οποία έχει αυξηθεί από την έναρξη της βιομηχανικής επανάστασης κατά περίπου 50% σύμφωνα με τη Διακυβερνητική Επιτροπή για την Κλιματική Αλλαγή, είναι ένας από τους βασικότερους παράγοντες στον οποίο οφείλεται η υποβάθμιση του περιβάλλοντος. Προκειμένου να

αντιμετωπιστεί η αυξανόμενη υποβάθμιση του περιβάλλοντος, υπάρχει επείγουσα ανάγκη να μελετηθεί και να μοντελοποιηθεί η χρονοσειρά των εκπομπών CO₂ σε επίπεδο χωρών καθώς επίσης και να χρησιμοποιηθούν κατάλληλα μοντέλα για βραχυπρόθεσμες και μακροπρόθεσμες προβλέψεις των ρύπων του διοξειδίου του άνθρακα. Στην παρούσα εργασία, θα χρησιμοποιήσουμε τα τυπικά μοντέλα πρόβλεψης χρονοσειρών με εξογενείς μεταβλητές (ARIMAX) καθώς και διάφορες οικογένειες τεχνητών νευρωνικών δικτύων συγκρίνοντας τις επιδόσεις τους.

Ανάλυση δεδομένων για την ελληνική αγορά ακινήτων

Θεόδωρος Μ. Μητράκος

Τράπεζα της Ελλάδος

Διεύθυνση Οικονομικής Ανάλυσης και Μελετών

tmitrakos@bankofgreece.gr

Η μελέτη επισκοπεί τα ιδιαίτερα χαρακτηριστικά της ελληνικής αγοράς ακινήτων και τις διαθέσιμες πηγές στοιχείων για την κατανόηση των παραγόντων που διαμορφώνουν την προσφορά και τη ζήτηση στην εν λόγω αγορά. Έμφαση δίνεται στο μηχανισμό συγκέντρωσης πρωτογενών μικρο-δεδομένων από την Τράπεζα της Ελλάδος με σκοπό την παρακολούθηση και ανάλυση της ελληνικής κτηματαγοράς. Πρόκειται για τις εκθέσεις εκτίμησης της αξίας ακινήτων που παρέχονται σήμερα στη χώρα μας από πιστοποιημένους εκτιμητές ακινήτων για οποιοδήποτε σκοπό και κυρίως για τις ανάγκες του τραπεζικού συστήματος. Χρησιμοποιώντας αυτά τα δεδομένα, η μελέτη εκτιμά δείκτες τιμών ακινήτων με τη βοήθεια απλών στατιστικών και οικονομομετρικών τεχνικών.

Η μελέτη παρουσιάζει τα αποτελέσματα αυτών των εκτιμήσεων, καθώς και τις πρόσφατες εξελίξεις και προοπτικές της ελληνικής αγοράς ακινήτων με έμφαση στην κατοικία. Από την ανάλυση προκύπτει η μεγάλη εξάρτηση της εγχώριας κτηματαγοράς από τη ζήτηση από το εξωτερικό και την πορεία της οικονομίας και του τουρισμού, ενώ παραμένουν θετικές οι προσδοκίες και ελκυστικές οι αποδόσεις που διατηρούν το επενδυτικό ενδιαφέρον. Ωστόσο, εξακολουθούν να υπάρχουν σημαντικές αβεβαιότητες και προκλήσεις για την ελληνική αγορά ακινήτων (γεωπολιτική αστάθεια, ενεργειακό κόστος - κόστος υλικών, επιτόκια και κόστος χρηματοδότησης κ.ά.), ενώ το κόστος στέγασης και άλλοι δείκτες προσιτότητας της κατοικίας βρίσκονται σε δυσανάλογα δυσμενή επίπεδα στην Ελλάδα σε σχέση με τις άλλες χώρες της ΕΕ.

Ένα Μπεϋζιανό αυτό-εκκινούμενο διάγραμμα ελέγχου για την άμεση ανίχνευση πολυδιάστατων ακραίων τιμών

Απόστολος Μπατσίδης^{1,*}, Κωνσταντίνος Μπουραζάς² και Παναγιώτης Τσιαμυρτζής^{3,4}

¹Τμήμα Μαθηματικών, Πανεπιστήμιο Ιωαννίνων, Ελλάδα

²Τμήμα Οικονομικών, Οικονομικό Πανεπιστήμιο Αθηνών, Ελλάδα

³ Department of Mechanical Engineering, Politecnico di Milano, Ιταλία

⁴ Τμήμα Στατιστικής, Οικονομικό Πανεπιστήμιο Αθηνών, Ελλάδα

* e-mail address: abatsidis@uoi.gr

Διαγράμματα ελέγχου για την ανίχνευση πολυδιάστατων ακραίων τιμών σε πραγματικό χρόνο (online detection), έχουν προταθεί είτε υπό την κλασική θεώρηση «εκπαίδευσης και ελέγχου» είτε με αυτό-εκκινούμενες μεθόδους (self-starting). Σύμφωνα με τις πρώτες μεθόδους, οι άγνωστες παράμετροι της διαδικασίας εκτιμώνται στη φάση I πριν ξεκινήσει η ανίχνευση σε πραγματικό χρόνο στη φάση II. Από την άλλη, οι αυτό-εκκινούμενες μέθοδοι έχουν αναπτυχθεί όταν οι παράμετροι της διαδικασίας δεν είναι γνωστές εκ των προτέρων ή δεν εκτιμώνται μέσω μιας φάσης I. Κύριος σκοπός αυτής της εργασίας είναι η εισαγωγή μιας Μπεϋζιανής μεθόδου αυτόματης εκκίνησης και άμεσης ανίχνευσης για πολυμεταβλητά κανονικά δεδομένα. Στο πλαίσιο αυτό, προτείνεται ένα διάγραμμα ελέγχου που αποτελεί επέκταση του πολύ γνωστού διαγράμματος ελέγχου του Hotelling (1947), υιοθετώντας ότι η παράμετρος θέσης και κλίμακας της πολυδιάστατης κανονικής κατανομής είναι άγνωστες και η όποια διαθέσιμη πληροφορία δίνεται μέσω της εκ των προτέρων συζυγούς κατανομής δύναμης (conjugate power prior). Κάποιες θεωρητικές ιδιότητες της προτεινόμενης μεθοδολογίας που αφορούν την ισχύ και τη συμπερασματολογία μετά την ανίχνευση ακραίας τιμής,

τροποποιώντας κατάλληλα τη μεθοδολογία των Mason et al. (1997), παρουσιάζονται, ενώ η απόδοσή της αξιολογείται μέσω μελέτης προσομοίωσης. Τέλος, η μέθοδος εφαρμόζεται σε ένα πραγματικό σύνολο δεδομένων.

Αναφορές

- Hotelling, H. (1947). Multivariate quality control. Illustrated by Air Testing of Sample Bombsights. In: Eisenhart, C., Hastay, M.W. and Wallis, W.A., Eds., *Techniques of Statistical Analysis*, McGraw Hill, New York, 111-184.
- Mason, R. L., Tracy, N. D and Young, J. C (1997). A practical approach for interpreting multivariate T^2 control chart signals. *Journal of Quality Technology*, **29**, 396-406.

**Επιλογή μοντέλου του κινητήρα του πλοίου εφαρμόζοντας
τεχνικές Κατηγοριοποίησης μηχανικής μάθησης και
χρησιμοποιώντας μεθόδους καταλογισμού και μείωσης
διαστάσεων των δεδομένων**

**Παναγιώτης Μπίρης¹, Κυριάκος Σκαρλάτος², Γρηγόρης
Παπαγεωργίου¹, Αικατερίνη Σκαμνιά¹, Σωτήριος
Μπερσίμης², Πολυχρόνης Οικονόμου^{1*}**

¹ Τμήμα Πολιτικών Μηχανικών, Πανεπιστήμιο Πατρών

² Τμήμα Διοίκησης Επιχειρήσεων, Πανεπιστήμιο Πειραιώς

peconom@upatras.gr

Η διεθνής ναυτιλιακή βιομηχανία παράγει, διαθέτει και χρειάζεται να επεξεργάζεται καθημερινά ένα τεράστιο όγκο δεδομένων, τα οποία χαρακτηρίζονται από την υψηλή διάσταση, την παρουσία πολλών ελλείπων και εσφαλμένων καταχωρήσεων αλλά και διπλοεγγραφών εξαιτίας κυρίως του ανθρώπινου παράγοντα ή λόγω της ελλείπουσας πρόσβασης σε ευαίσθητες, απόρρητες και σημαντικές πληροφορίες. Τα χαρακτηριστικά αυτά των δεδομένων έχουν κρίσιμο αντίκτυπο στις αποφάσεις που λαμβάνονται με βάση τα διαθέσιμα, συχνά ανεπαρκή δεδομένα, οδηγώντας εν τέλει, σε εσφαλμένες αποφάσεις και ενέργειες με συνεπακόλουθη μειωμένη παραγωγικότητα, αυξημένες λειτουργικές δαπάνες και σημαντική μείωση του ανταγωνιστικού πλεονεκτήματος.

Στη παρούσα μελέτη παρουσιάζεται μια ολοκληρωμένη έρευνα συμπλήρωσης των ελλείπων τιμών και εφαρμογής τεχνικών μείωσης διαστάσεων με στόχο την εξαγωγή συμπερασμάτων για την επίπτωσή τους στα μοντέλα μηχανικής μάθησης ταξινόμησης. Πιο συγκεκριμένα, οι τεχνικές αυτές εφαρμόζονται και συγκρίνονται για την ταξινόμηση της κύριας μηχανής ενός πλοίου με βάση τα διαθέσιμα τεχνικά και λειτουργικά χαρακτηριστικά του πλοίου, συμπεριλαμβανομένων των φυσικών διαστάσεων, των διαφόρων χωρητικοτήτων, ταχυτήτων καθώς και της κατανάλωσης.

Διατήρηση ιδιοτήτων γήρανσης του μεγίστου από τυχαίο αριθμό τυχαίων μεταβλητών

Παναγιώτης Μπομποτάς¹, Μάρκος Β. Κούτρας²

¹ Τμήμα Μαθηματικών, Πανεπιστήμιο Θεσσαλίας

² Τμήμα Στατιστικής και Ασφαλιστικής Επιστήμης,
Πανεπιστήμιο Πειραιώς

pbobotas@uth.gr

Στην εργασία αυτή γίνεται μία μελέτη σχετικά με τη διατήρηση ιδιοτήτων γήρανσης, όπως αύξοντος ρυθμού αστοχίας (increasing failure rate – IFR), αύξοντος μέσου ρυθμού αστοχίας (increasing failure rate average – IFRA), και νέο καλύτερο από χρησιμοποιημένο (new better than used – NBU), για την μέγιστη παρατήρηση στην περίπτωση τυχαίου δείγματος με τυχαίο πλήθος παρατηρήσεων (random maximum). Τα αποτελέσματα ενοποιούν, συμπληρώνουν, διορθώνουν, επεκτείνουν και/ή γενικεύουν υπάρχοντα αποτελέσματα στη βιβλιογραφία. Ως υποπροϊόν της μελέτης, εισάγονται επίσης οι αντίστοιχες έννοιες των ιδιοτήτων γήρανσης IFR, IFRA και NBU για μία μη αρνητική μικτού τύπου τυχαία μεταβλητή.

Αναφορές

- Bartoszewicz, J. (2001). Stochastic comparisons of random minima and maxima from life distributions. *Statistics and Probability Letters*, **55**, 107-112.
- Bobotas, P., and Koutras, V.M. (2024). On the preservation of ageing properties under random maxima. *Statistics and Probability Letters*, **210**, DOI: 10.1016/j.spl.2024.110111
- Hazra, N.K., Nanda, A.K., and Shaked, M. (2014). Some aging properties of parallel and series systems with random number of components. *Naval Research Logistics*, **61**, 238-243.
- Li, X., and Zuo, M.J. (2004). Preservation of stochastic orders for random minima and maxima, with applications. *Naval Research Logistics*, **51**, 332-344.
- Shaked, M. (1975). On the distribution of the minimum and the maximum of a random number of i.i.d. random variables. In: Patil, G.P., Kotz, S., Ord, J.K. (eds), *In: A Modern Course on Statistical Distributions in Scientific Work*, vol. 1, Springer, Dordrecht, pp 363-380.

Βραχυπρόθεσμες Προβλέψεις Ζήτησης Νερού από Μονομεταβλητές Χρονοσειρές Σταθμών Δεξαμενών με χρήση αλγορίθμων και στατιστικών εργαλείων

Μυλλής Γεώργιος, Αλκιβιάδης Τσιμπίρης

Διεθνές Πανεπιστήμιο Ελλάδος Τμήμα Μηχανικών Πληροφορικής,
Υπολογιστών και Τηλεπικοινωνιών, Σερρών
e-mail geormyll@ihu.gr, atsimpiris@ihu.gr

Η βραχυπρόθεσμη πρόβλεψη της ζήτησης νερού αποτελεί ένα κρίσιμο ζήτημα για τη διαχείριση υδατικών πόρων, ειδικά σε περιόδους έντονης ξηρασίας ή υψηλής κατανάλωσης. Η χρήση αλγορίθμων μηχανικής μάθησης, όπως τα Long Short-Term Memory Networks (LSTM) και RandomForestRegressor μπορεί να βοηθήσει στην αναπλήρωση των ελλειπτικών τιμών αλλά και στην ακριβή πρόβλεψη βάσει δεδομένων από χρονοσειρές σταθμών δεξαμενών.

Για την ανίχνευση και διαχείριση ανωμαλιών στα δεδομένα, εκτός από την προηγμένη τεχνολογία των LSTM, η χρήση στατιστικών εργαλείων όπως το εύρος διακύμανσης (Interquartile Range, IQR), το κινούμενο μέσο (moving average) και η τεχνική της επαναδειγματοληψίας (resample) παίζουν κρίσιμο ρόλο.

Η συνδυασμένη χρήση αυτών των στατιστικών εργαλείων ενισχύει στην αναγνώριση και κατανόηση των τύπων ανωμαλιών στα δεδομένα χρονοσειρών και είναι ζωτικής σημασίας για την εφαρμογή αποτελεσματικών τεχνικών διόρθωσης και πρόβλεψης. Τα αποτελέσματα και οι μετρικές επίδοσης του μοντέλου δείχνουν ότι το LSTM λειτουργεί πολύ καλά στην πρόβλεψη και την αναπλήρωση των ελλειπόντων τιμών στο σύνολο δεδομένων. Οι χαμηλές τιμές των μετρικών MSE (Mean Squared Error) 0.02235 και RMSE (Root Mean Squared Error) 0.14952 υποδεικνύουν ότι οι προβλέψεις του μοντέλου είναι κοντά στα πραγματικά δεδομένα, ενώ η υψηλή τιμή του R^2 0.9878 δείχνει ότι το μοντέλο εξηγεί σημαντικό μέρος της διακύμανσης στο σύνολο δεδομένων.

Πραγματοποιήθηκαν προβλέψεις για τις μελλοντικές τιμές ζήτησης όγκου νερού σε μια συγκεκριμένη μελλοντική χρονική στιγμή. Η απόδοση του μοντέλου αξιολογήθηκε με τη χρήση της μετρικής Μέσου Τετραγωνικού Σφάλματος (MSE) με αποτελέσματα δειγματοληπτικά κάποιων δεξαμενών. **Δεξαμενή 1:** 0.01045 , **Δεξαμενή 2:** 0.00306, **Δεξαμενή 3:** 0.00122, **Δεξαμενή 4:** 0.00060

Τα αποτελέσματα δείχνουν ότι το μοντέλο λειτουργεί με ακρίβεια στην πρόβλεψη της ζήτησης όγκου του νερού στις δεξαμενές, παρέχοντας στους διαχειριστές μια αξιόπιστη εργαλειοθήκη για την προετοιμασία και ανταπόκριση στις ανάγκες μια πιο ορθολογικής διανομής υδάτινων πόρων.

Αναφορές

1. Smith, J. & Doe, A. (2022). Artificial Intelligence in Water Resource Management: A Review. *Journal of Water Resources*, 58(3), 295-312.
2. Liu, H., Kelly, M., & Feng, X. (2023). Machine Learning Models for Predicting Water Demand: A Comparative Study. *Hydrology Today*, 45(2), 134-150.
3. Zhang, Y., & Wang, L. (2020). AI-Based Optimization for Water Distribution Systems. *Water Management Solutions*, 39(4), 441-456.
4. Foster, K., & McMahon, G. (2023). Artificial Neural Networks for Forecasting Urban Water Demand: A Case Study. *Water Resources Management*, 60(9), 2457-2472.
5. Santos, L., & Oliveira, P. (2020). Anomaly Detection in Water Distribution Networks: A Review. *Journal of Water Process Engineering*, 35, 101214.
6. Khan, S., & Lee, W. S. (2022). A Review on the Applications of Artificial Intelligence in Water Quality Monitoring. *Sensors and Actuators B: Chemical*, 329, 129473.
7. Harris, A., & Jones, D. (2021). The Role of AI in Sustainable Water Management. *Journal of Sustainable Development*, 14(6), 789-805.
8. Greene, A., & Patel, H. (2021). Predictive Analytics for Water Demand and Supply Management. *Water Resources Industrial Journal*, 12, 34-45.
9. Kapoor, R., & Tiwari, S. (2023). Blockchain Technology in Water Distribution Systems: Enhancing Transparency and Efficiency. *Journal of Water Resources Planning and Management*, 149(4), 04023008.
10. O'Neill, L., & Thompson, S. (2019). Application of Machine Learning Algorithms for Efficient Water Leak Detection. *Procedia Computer Science*, 152, 305-312.
11. Bhatia, U., & Gupta, D. (2020). Artificial Intelligence in Predicting Water Table Depth: A Comparative Study. *Journal of Hydrology*, 584, 124693.
12. Singh, R., & Mehrotra, I. (2021). Optimization of Urban Water Supply Networks: A Machine Learning Approach. *Urban Water Journal*, 18(6), 455-467.
13. Wallace, M., & Kumar, L. (2022). Artificial Neural Networks for the Prediction of Water Quality Index. *Water Quality Research Journal*, 57(1), 56-67.
14. Chandola, V., Banerjee, A., & Kumar, V. (2009). Anomaly detection: A survey. *ACM computing surveys (CSUR)*, 41(3), 1-58.
15. Box, G. E., Jenkins, G. M., Reinsel, G. C., & Ljung, G. M. (2015). *Time series analysis: forecasting and control*. John Wiley & Sons.
16. Hyndman, R. J., & Athanasopoulos, G. (2018). *Forecasting: principles and practice*. OTexts.
17. James, G., Witten, D., Hastie, T., & Tibshirani, R. (2013). *An introduction to statistical learning*. Springer.
18. Goodfellow, I., Bengio, Y., & Courville, A. (2016). *Deep Learning*. MIT Press.
19. Hochreiter, S., & Schmidhuber, J. (1997). Long short-term memory. *Neural computation*, 9(8), 1735-1780.
20. Zhang, G. P. (2003). Time series forecasting using a hybrid ARIMA and neural network model. *Neurocomputing*, 50, 159-175.
21. Tashman, L. J. (2000). Out-of-sample tests of forecasting accuracy: an analysis and review. *International Journal of Forecasting*, 16(4), 437-450.
22. Donkor, E. A., Mazzuchi, T. A., Soyer, R., & Roberson, J. A. (2014). Urban water demand forecasting: A review of methods and models. *Journal of Water Resources Planning and Management*, 140(2), 146-159.

Τυπική αξιολόγηση συμφωνίας μεταξύ συσκευών ανοιχτού κώδικα και αντίστοιχων αναφοράς, βιομηχανικής χρήσης, με εφαρμογή σε μετρητές pH.

Ευμορφία Μπατάκα & Χρήστος Νάκας

Εργαστήριο Βιομετρίας, Τμήμα Γεωπονίας, Φυτικής Παραγωγής και Αγροτικού Περιβάλλοντος, Σχολή Γεωπονικών Επιστημών, Πανεπιστήμιο Θεσσαλίας, Οδός Φυτόκου, 38446 Ν. Ιωνία/Βόλος, Μαγνησία.
cnakas@uth.gr

Οι συσκευές ανοιχτού κώδικα χρησιμοποιούνται πλέον σε πληθώρα ερευνητικών εφαρμογών. Για τη συγκεκριμένη εφαρμογή, κατασκευάστηκε νέος, ανοιχτού κώδικα, φορητός, χαμηλού κόστους καταγραφέας pH, κατάλληλος για μετρήσεις πεδίου, που σχεδιάστηκε και αναπτύχθηκε για χρήση σε πειράματα σχετικά με τη μεταποίηση γεωργικών προϊόντων. Η συσκευή βαθμονομήθηκε σε επίπεδα pH μεταξύ 4.01 και 7.01. Στη συνέχεια, δοκιμάστηκε με χειροκίνητες μετρήσεις pH σε χυμό εσπεριδοειδών. Σχεδιάστηκε τυπική διαδικασία αξιολόγησης συμφωνίας μεταξύ της συσκευής ανοιχτού κώδικα και του αδιάβροχου μετρητή pH Hanna HI9024. Υπολογίστηκαν δείκτες συμφωνίας και προτάθηκαν εργαλεία γραφικής αξιολόγησης μέσω υποδειγμάτων (μοντέλων) μεικτών επιδράσεων. Συζητούνται τα πλεονεκτήματα και τα μειονεκτήματα της ερμηνείας της συμφωνίας μέσω της προτεινόμενης διαδικασίας. Στην εφαρμογή μας οι δείκτες ανέφεραν μέτρια συμφωνία και η περεταίρω ανάλυση ομοιότητας αποκάλυψε μια σταθερή διαφορά/μεροληψία 0,22 μονάδων pH. Μετά την επαναβαθμονόμηση, η συμφωνία μεταξύ των συσκευών βελτιώθηκε σε εξαιρετικά επίπεδα. Η διαδικασία μπορεί να ακολουθηθεί γενικά για την αποφυγή παραπλανητικών ή υπερβολικά απλοϊκών αποτελεσμάτων μελετών που αναφέρουν μόνο συντελεστές συσχέτισης για την τυπική αξιολόγηση της συμφωνίας μεταξύ μετρήσεων συσκευών.

Αναφορές

Bataka EP, Maletsika P, Nakas CT. (2024). Formal Assessment of Agreement and Similarity between an Open-Source and a

Reference Industrial Device with an Application to a Low-Cost
pH Logger, *Sensors*, **24**(2):490.

<https://doi.org/10.3390/s24020490>

Choudhary PK, Nagaraja HN. (2017). *Measuring Agreement*, Wiley:
Hoboken, NJ, USA. ISBN 9781118078587.

Η επίδραση των υπαίθριων δραστηριοτήτων και της φυσικής αγωγής σε μαθητές με μεταναστευτικό υπόβαθρο

Ντούσκα Βάγια¹, Παπαδόπουλος Δημήτρης²

¹ Διευθύντρια 3ου Δημοτικού Σχολείου Παλαιού Φαλήρου,
Υποψήφια Διδάκτωρ University St.Kliment Ohridski Faculty of
Preschool and Primary School Education,

² Στατιστικός, BasisPap; info@basispap.com

Η ένταξη είναι ένα πολύτιμο εργαλείο που χρησιμοποιείται από τα σχολεία για τη δημιουργία ενός εκπαιδευτικού περιβάλλοντος που διασφαλίζει την ασφάλεια για τα παιδιά ενθαρρύνοντας τη συμμετοχή και την ευημερία τους, ενώ προάγει το σεβασμό και την αποδοχή της διαφορετικότητας. Η εφαρμογή της ένταξης στο μάθημα της φυσικής αγωγής και των υπαίθριων δραστηριοτήτων στοχεύει στην ολιστική ανάπτυξη των παιδιών. Σκοπός αυτής της έρευνας ήταν να εντοπίσει τις επιπτώσεις της φυσικής αγωγής ως μέρος του σχολικού προγράμματος σε παιδιά προσχολικής ηλικίας μεταναστών και, ταυτόχρονα, να διερευνήσει πώς οι υπαίθριες δραστηριότητες που πραγματοποιούνται κατά τη διάρκεια του μαθήματος συμβάλλουν στην ένταξη αυτών των παιδιών. Εφαρμόστηκε πρόγραμμα φυσικής αγωγής και υπαίθριων δραστηριοτήτων για παιδιά προσχολικής ηλικίας με και χωρίς υπόβαθρο μεταναστών. Στη μελέτη συμμετείχαν συνολικά 98 παιδιά προσχολικής ηλικίας. Το 51% (N = 50) είναι κορίτσια και το 49% (N = 48) αγόρια. Η υπαίθρια παρέμβαση βοήθησε στην κοινωνική και συναισθηματική ανάπτυξη όλων των παιδιών. Τα αποτελέσματα αυτά οδηγούν στο συμπέρασμα ότι το πρόγραμμα φυσικής αγωγής και υπαίθριων δραστηριοτήτων που εφαρμόστηκε αποδείχθηκε αποτελεσματικό. Το πρόγραμμα που εφαρμόστηκε, οδήγησε στην

ένταξη παιδιών, μεταναστών και μη μεταναστών, ενισχύοντας τις κοινωνικές και συναισθηματικές τους δεξιότητες, αυξάνοντας τη συμμετοχή τους και διασφαλίζοντας την ευημερία τους. Επομένως, μπορούμε να συμπεράνουμε ότι η εισαγωγή αυτού του τύπου προγραμμάτων φυσικής αγωγής στα νηπιαγωγεία μπορεί να οδηγήσει στην αποτελεσματική εφαρμογή των αρχών της ένταξης.

Αναφορές

Jackson, A. J., Miller, J. T., & Stevens, G. K. (2012). School anxiety: Teacher-rated stress factors in Bulgarian school children. In M. Brunner (Ed.), *Psychology in schools* (2nd ed., pp. 12–37), School Press.

Μη παραμετρικές μέθοδοι παρακολούθησης πολυδιάστατων ροών δεδομένων

Πολυχρόνης Οικονόμου^{1*}, Σωτήριος Μπερσίμης²

¹ Τμήμα Πολιτικών Μηχανικών, Πανεπιστήμιο Πατρών

² Τμήμα Διοίκησης Επιχειρήσεων, Πανεπιστήμιο Πειραιώς

peconom@upatras.gr

Η στατιστική παρακολούθηση διεργασιών είναι ζωτικής σημασίας σε διάφορους τομείς όπως η βιοεπιτήρηση, οι ροές δεδομένων κ.α. Στην παρούσα εργασία παρουσιάζεται μια μη παραμετρική διαδικασία παρακολούθησης που στοχεύει στην ανίχνευση αλλαγών σε πολυδιάστατες ροές δεδομένων. Η διαδικασία βασίζεται στη χρήση των κυρτών θηκών (convex hull) για την κατασκευή κατάλληλων διαγραμμάτων ελέγχων.

Η συμβολή της Τεχνητής Νοημοσύνης στην Λήψη Αποφάσεων: Αλγόριθμοι και Εφαρμογές

Δημήτρης Πανάρετος

Πανεπιστήμιο Δυτικής Μακεδονίας,
Τμήμα Στατιστικής και Ασφαλιστικής Επιστήμης
d.panaretos@uowm.gr

Τα τελευταία χρόνια, η εξέλιξη των υπολογιστών αλλά και η ανάγκη επεξεργασίας μεγάλου όγκου δεδομένων έστρεψε την επιστημονική κοινότητα στην ανάπτυξη «έξυπνων» τεχνολογιών. Αυτές οι σύγχρονες τεχνολογίες ήταν εκείνες που ενδυνάμωσαν το πεδίο της Τεχνητής Νοημοσύνης (TN). Η TN χρησιμοποιεί «έξυπνους» αλγόριθμους οι οποίοι καταλήγουν σε αποφάσεις βασισμένες στα δεδομένα, ενώ προσφέρουν πρωτόγνωρες δυνατότητες ανάλυσης.

Στην παρούσα εργασία παρουσιάζονται ενδεικτικά μερικές περιπτώσεις χρήσης της TN σε πεδία της Μηχανικής όπως η Κατασκευαστική, η Περιβαλλοντική και η Ενεργειακή αλλά και τομείς όπως η Υγεία και η Επιχειρηματικότητα, αναδεικνύοντας τον τρόπο με τον οποίον η TN συμβάλλει θετικά στην αντιμετώπιση πολύπλοκων προβλημάτων καθώς και στην διαχείριση και ανάλυση μεγάλου όγκου δεδομένων (Big Data) που παράγονται καθημερινά.

Ειδικότερα, η παρουσίαση εστιάζει στην αξιοποίηση των αλγορίθμων Μηχανικής Μάθησης (Machine Learning) και Βαθιάς Μηχανικής Μάθησης (Deep Learning) για την αυτοματοποίηση στην λήψη αποφάσεων.

Τέλος, εξετάζονται τα ζητήματα ηθικής και διαφάνειας που ανακύπτουν μέσα από την χρήση των αλγορίθμων TN, τονίζοντας την άμεση ανάγκη θέσπισης ρυθμιστικού πλαισίου που να εγγυάται την υπεύθυνη χρήση της.

Αναφορές

- Tianyang Lin, Yuxin Wang, Xiangyang Liu, Xipeng Qiu. (2022). A survey of transformers. *AI Open*, **3**, 111-132. <https://doi.org/10.1016/j.aiopen.2022.10.001>
- Zhang Z, Li J. A Review of Artificial Intelligence in Embedded Systems. *Micromachines* (Basel). 2023 Apr 22;14(5):897. doi: 10.3390/mi14050897. PMID: 37241521; PMCID: PMC10220566.
- Huang, Qi. (2017). Application of Artificial Intelligence in Mechanical Engineering. 10.2991/iccia-17.2017.154.

Δειγματοληψία σε Συμμετρικά Δίκτυα:

Αμεροληψία Εκτιμητριών

Παπαγεωργίου Μυρτώ, Φαρμάκης Νικόλαος

Τμήμα Μαθηματικών, ΑΠΘ

mmpapageor@math.auth.gr, farmakis@math.auth.gr

Τα δίκτυα τείνουν να μπου ολοένα και περισσότερο στη φαρέτρα των εργαλείων της Επιστημονικής-Ερευνητικής Κοινότητας. Είναι μια πολύ ενδιαφέρουσα και χρήσιμη έννοια και οι εφαρμογές τους εκτείνονται σε ένα μεγάλο φάσμα περιπτώσεων. Στην παρούσα εργασία εισάγεται ένα είδος δικτυακών δομών που έχει το σύνολο των ακμών συμμετρικό. Θα χρησιμοποιηθεί ως πληθυσμός-στόχος για δειγματοληψία. Η στόχευση είναι στην επιδίωξη να έχουμε από τη δειγματοληψία αμερόληπτες εκτιμήτριες της μέσης τιμής των παραμέτρων επίμαχων τυχαίων μεταβλητών που υπεισέρχονται στο εκάστοτε πρόβλημα. Γίνεται αρχικά μια εισαγωγική παρουσίαση των δικτύων και των βασικών εννοιών και ιδιοτήτων τους καθώς και μια ανάλογη παρουσίαση τη δειγματοληψίας και των πολύ βασικών μεθόδων-τεχνικών της που είναι πολύ γνωστές από την σχετική παγκόσμια βιβλιογραφία. Στη συνέχεια εξετάζεται το ποια ή ποιες μεθοδολογίες δειγματοληψίας δίνουν αμερόληπτες εκτιμήτριες της μέσης τιμής των τυχαίων μεταβλητών που υπεισέρχονται στα υπό εξέταση προβλήματα. Είναι γνωστό ότι δειγματοληψία σε δίκτυα δεν έχει συχνά τέτοια αποτελέσματα όσο έχει σε άλλους πληθυσμούς. Εξετάζεται η απλή τυχαία δειγματοληψία και παραλλαγές της που αναφέρονται κυρίως σε 2 ή περισσότερα στάδια εφαρμογής της. Δίνονται τέλος και μερικά παραδείγματα που δείχνουν την ισχύ και το εφαρμόσιμο των βασικών θεωρητικών διαπιστώσεων.

Αναφορές

- Cochran W. (1977). *Sampling Techniques*, John Wiley & Sons, Inc, New York, London, Sydney, Toronto.
- Farmakis N. (2015). *Sampling and Applications*, Hellenic Academic-Books, Athens, (in Greek). ISBN: 978-960-603-093-2.
- Farmakis N. (2016) “Introduction to Sampling”, Kyriakidis Bros Editing S.A, Thessaloniki (in Greek).

Πρόβλεψη πολυμεταβλητών χρονοσειρών με Reservoir Computing και LSTM

Ελένη Παπαδοπούλου, Δημήτρης Κουγιουμτζής

Τμήμα Ηλεκτρολόγων Μηχανικών και Μηχανικών Υπολογιστών,
Αριστοτέλειο Πανεπιστήμιο Θεσσαλονίκης, Πανεπιστημιούπολη,
54124, Θεσσαλονίκη
epapador@ece.auth.gr

Τα κλασικά αναδρομικά νευρωνικά δίκτυα είναι μία ομάδα υπολογιστικών μοντέλων των οποίων ο σχεδιασμός στηρίζεται στη λειτουργία του ανθρώπινου εγκεφάλου και ταυτόχρονα η αρχιτεκτονική τους με τον βρόχο ανάδρασης επιτρέπει τη χρήση τους ως δυναμικά συστήματα και για αυτό χρησιμοποιούνται στην μοντελοποίηση και πρόβλεψη χρονοσειρών. Ένας τύπος νευρωνικών δικτύων που ανήκουν σε αυτήν την κατηγορία και προτιμώνται ιδιαίτερα εξαιτίας των δυνατοτήτων μεγάλης μνήμης που διαθέτουν είναι τα LSTM (Long Short-Term Memory). Η τοπολογία τους επιτρέπει να ληφθεί υπόψη περισσότερη πληροφορία από τη διαθέσιμη χρονοσειρά και κατ' επέκταση να ανιχνευτούν ευκολότερα τυχόν εξαρτήσεις μεγάλου χρονικού εύρους σε σύγκριση με άλλους τύπους αναδρομικών νευρωνικών δικτύων ¹. Ωστόσο, η υπολογιστικά ακριβή εκπαίδευση τους εξαιτίας της πολύπλοκης δομής τους οδηγεί σε σύγκριση με την πρόσφατα ανεπτυγμένη τεχνική του Reservoir Computing. Πρόκειται για έναν αλγόριθμο μηχανικής μάθησης που ανήκει στην οικογένεια των αναδρομικών νευρωνικών δικτύων και δεν απαιτεί μεγάλη υπολογιστική δύναμη, αφού οι πίνακες βαρών που σχετίζονται με το αναδρομικό κομμάτι του μοντέλου αρχικοποιούνται τυχαία χωρίς να αλλάζουν κατά την εκπαίδευσή του ². Υπό αυτό το θεωρητικό πλαίσιο, η παρούσα εργασία αποσκοπεί, σε πρώτο στάδιο, στην κατασκευή μοντέλων LSTM και Reservoir Computing με βελτιστοποιημένες τις υπερπαραμέτρους τους και σε δεύτερο στάδιο στη χρήση τους για πρόβλεψη πολυμεταβλητών χρονοσειρών και την σύγκριση των δύο τύπων νευρωνικών δικτύων.

Αναφορές

1. Dominguez, D., Barriuso Pastor, J., Pantoja-Diaz O. and González-Rodríguez, M. (2023). Forecasting Worldwide Temperature from Amazon Rainforest Deforestation Using a Long-Short Term Memory Model. *Sustainability*, **15**.
2. Bianchi, F. M., Scardapane, S., Løkse, S. and Jenssen R. (2021). Reservoir Computing Approaches for Representation and Classification of Multivariate Time Series. *IEEE Transactions on Neural Networks and Learning Systems*, **32**.

Σχηματισμός αραιού μοντέλου παλινδρόμησης πολυμεταβλητών χρονοσειρών με μέτρα πληροφορίας

Πέτρος Πετρίδης, Δημήτρης Κουγιουμτζής

Τμήμα Ηλεκτρολόγων Μηχανικών και Μηχανικών
Υπολογιστών, Αριστοτέλειο Πανεπιστήμιο Θεσσαλονίκης,
Πανεπιστημιούπολη, 54124, Θεσσαλονίκη
epapador@ece.auth.gr

Στη μοντελοποίηση και πρόβλεψη χρονοσειρών υψηλής διάστασης (πολλών παρατηρούμενων μεταβλητών) τα μοντέλα μηχανικής και βαθιάς μάθησης χρησιμοποιούνται ευρέως αλλά η πολύπλοκη δομή τους δεν επιτρέπει ερμηνεία της δυναμικής του συστήματος. Σε αντίθεση τα τελευταία χρόνια έχουν αναζητηθεί αραιά μοντέλα παλινδρόμησης όπου η έμφαση δίνεται στις μεταβλητές υστέρησης (παρατηρούμενες μεταβλητές σε χρονικές υστερήσεις σε σχέση με τη μεταβλητή στόχου) ¹. Στην εργασία αυτή χρησιμοποιούνται μέτρα πληροφορίας για τον εντοπισμό των μεταβλητών υστέρησης του αραιού μοντέλου παλινδρόμησης. Ουσιαστικά χρησιμοποιείται ο ίδιος αλγόριθμος της Μερικής Αμοιβαίας Πληροφορίας από Μεικτή Εμβύθιση (PMIME) που αναπτύχθηκε για να εκτιμά άμεσες σχέσεις αιτιότητας σε υψηλο-διάστατες χρονοσειρές ². Ο αλγόριθμος εντοπίζει τις μεταβλητές υστέρησης που εξηγούν καλύτερα τη μεταβλητή στόχου και τηρούν την αρχή του ελάχιστου πλεονασμού και μέγιστης σχετικότητας. Η διαδικασία εύρεσης του βέλτιστου υποσυνόλου μεταβλητών υστέρησης δοκιμάζεται σε

πολυμεταβλητές χρονοσειρές από γραμμικά και μη-γραμμικά στοχαστικά συστήματα καθώς και συζευγμένα χαοτικά συστήματα. Αραιά μοντέλα πρόβλεψης όπως το μοντέλο γραμμικών τοπικών γειτόνων, που βασίζονται στις εντοπισμένες μεταβλητές υστέρησης, αναπτύσσονται επίσης για την πρόβλεψη πολυμεταβλητών χρονοσειρών.

Αναφορές

1. Brunton S.L. & Kutz N.J. (2019), *Data-Driven Science and Engineering: Machine Learning, Dynamical Systems, and Control*, Cambridge University Press, New York
2. Fotiadis A. Vlachos I. & Kugiumtzis D. (2024). The causality measure of partial mutual information from mixed embedding (PMIME) revisited. *Chaos* 34, 033113

Μοντέλο πρόβλεψης του ΑΕΠ με χρήση αιτιότητας κατά Granger εκτιμώμενη από τη μερική συσχέτιση

**Βασιλική Σαραντίδου¹, Χριστίνα Καραμιχαλάκου²,
Δημήτρης Κουγιουμτζής³**

¹Τμήμα Μαθηματικών, Αριστοτέλειο Πανεπιστήμιο Θεσσαλονίκης,
Πανεπιστημιούπολη, 54124, Θεσσαλονίκη
vas.sarantidou@gmail.com

²Διεύθυνση Στατιστικών Επιχειρήσεων, Ελληνική Στατιστική
Αρχή, Πειραιώς 46 και Επονιτών, 18510, Πειραιάς

³Τμήμα Ηλεκτρολόγων Μηχανικών και Μηχανικών Υπολογιστών,
Αριστοτέλειο Πανεπιστήμιο Θεσσαλονίκης, Πανεπιστημιούπολη,
54124, Θεσσαλονίκη

Το Ακαθάριστο Εγχώριο Προϊόν (ΑΕΠ) αποτελεί μία από τις πιο γνωστές επίσημες στατιστικές μετρήσεις. Αξιοποιώντας τις τριμηνιαίες μετρήσεις του ΑΕΠ και των συνιστωσών του, στόχος είναι η πρόβλεψη του για το τρέχον τρίμηνο. Ωστόσο, για πολλές συνιστώσες και υστερήσεις αυτών το μοντέλο παλινδρόμησης είναι υψηλής διάστασης. Οι πιο γνωστές προσεγγίσεις μείωσης διάστασης είναι με εξαγωγή νέων μεταβλητών, όπως η παλινδρόμηση κυρίων συνιστωσών

(PCR), και επιλογή μεταβλητών, όπως είναι ο τελετής ελάχιστης απόλυτης συρρίκνωσης και επιλογής (LASSO). Στην παρούσα εργασία προτείνεται ένα νέο μοντέλο πρόβλεψης εφαρμόζοντας έναν αλγόριθμο επιλογής μεταβλητών σταδιακά προς τα εμπρός. Ως κριτήριο επιλογής χρησιμοποιείται η μερική συσχέτιση¹ για την αξιολόγηση της αιτιότητας κατά Granger οποιασδήποτε από τις υπονήφιες συνιστώσες με υστέρηση (συμπεριλαμβανομένου παρελθοντικών τιμών του ΑΕΠ). Το κριτήριο τερματισμού χρησιμοποιεί παραμετρικό έλεγχο υπόθεσης² που εξασφαλίζει την ισορροπία μεταξύ της πολυπλοκότητας του μοντέλου και της προγνωστικής ισχύος. Η αξιολόγηση της αξιοπιστίας του αλγορίθμου γίνεται με προσομοιωτική μελέτη και συγκρίνεται με άλλες προσεγγίσεις, όπως LASSO και PCR. Τελικό βήμα αποτελεί η εφαρμογή του αλγορίθμου σε δεδομένα χρονοσειρών του ελληνικού ΑΕΠ και των συνιστωσών του, που καταρτίζει η Ελληνική Στατιστική Αρχή (ΕΛΣΤΑΤ). Προσδιορίζονται οι μεταβλητές που επηρεάζουν περισσότερο τις διακυμάνσεις του ΑΕΠ και σχηματίζεται το αντίστοιχο μοντέλο πρόβλεψης. Η εργασία πραγματοποιήθηκε στο πλαίσιο του προγράμματος EMOS.

Αναφορές

1. Li, R., Liu, J., & Lou, L. (2017). Variable selection via partial correlation. *Statistica Sinica*, 27(3), 983–996.

Μία επέκταση της αναδρομής Panjer για μεικτές σύνθετες κατανομές

Σπυρίδων Μ. Τζανίνης, Απόστολος Μποζίκας
Τμήμα Στατιστικής και Ασφαλιστικής Επιστήμης,
Πανεπιστήμιο Πειραιώς,
stzaninis@unipi.gr

Έστω N μια απαριθμήτρια τυχαία μεταβλητή, X μια στοχαστική διαδικασία απαιτήσεων και S η σύνθετη κατανομή που επάγεται από το ζευγάρι (N, X) . Θεωρούμε το μεικτό ανάλογο της κλασσικής κλάσης κατανομών Panjer καθώς και τις επαγόμενες από αυτή σύνθετες κατανομές, επιτρέποντας στις N και X να είναι υπό συνθήκη από αμοιβαία ανεξάρτητες δοθέντος Θ και στην X να είναι υπό συνθήκη ανεξάρτητη και ισόνομη δοθέντος Θ , όπου Θ είναι ένα d -διάστατο τυχαίο διάνυσμα. Αρχικά παρουσιάζεται ένας χαρακτηρισμός της κλάσης των μεικτών κατανομών Panjer μέσω φυσιολογικών δεσμευμένων πιθανοτήτων (regular conditional probabilities). Ως συνέπεια του παραπάνω χαρακτηρισμού αποδεικνύεται μία απλή αναδρομική σχέση για τον υπολογισμό της P_S στην περίπτωση που τα μεγέθη απαιτήσεων ακολουθούν μία διακριτή κατανομή.

Στατιστική ανάλυση για το Projectification στη συντήρηση αεροσκαφών

Παναγιώτης Φιλokώστας, Αναστάσιος Πλιούτσιας,
Κλεάνθης Συρακούλης
Πανεπιστήμιο Θεσσαλίας
pfilokostas@uth.gr

Η ικανότητα της διαχείρισης έργων συντήρησης αεροναυτικών συστημάτων, για τον προγραμματισμό δραστηριοτήτων και τον έλεγχο της προόδου υπό συγκεκριμένους περιορισμούς κόστους, χρόνου και ποιότητας συνδέεται άρρηκτα με το επίπεδο υιοθέτησης της στρατηγικής του Projectification, σε έναν αεροπορικό οργανισμό. Η παρούσα εργασία πραγματεύεται την αξία της στατιστικής επεξεργασίας στοιχείων μιας ευρύτερης έρευνας πεδίου, μέσα από συγκεκριμένα παραδείγματα στατιστικών συσχετίσεων που φανερώνουν διαφοροποιήσεις στην αντίληψη διακριτών μεταξύ τους συνόλων εντός οργανισμού, σχετικά με το Projectification στους κόλπους του. Η στατιστική ανάλυση βασίζεται σε έρευνα με ερωτηματολόγιο. Η ανάλυση των δεδομένων με χρήση του SPSS περιέχει μια σειρά δοκιμών και ελέγχων. Για τις ανάγκες και τη διευκόλυνση της διαδικασίας ανάλυσης καθορίζονται εξαρτημένες και ανεξάρτητες μεταβλητές, ενώ για την ισχυροποίηση της ορθότητας επιλογής των δειγματικών μονάδων διενεργείται η δοκιμή Kolmogorov-Smirnov. Τα αρχικά βήματα της στατιστικής προσέγγισης συμπληρώνει η ανάλυση αξιοπιστίας για την κατάδειξη της «ομοιογένειας» «εσωτερικής συνοχής» και «σταθερότητας» της διαδικασίας μέτρησης. Ως μέτρο της εσωτερικής συνέπειας επιλέχθηκε ο συντελεστής Cronbach's alpha ενώ δεν απαιτήθηκε η διερευνητική ανάλυση παραγόντων για τον έλεγχο εγκυρότητας της εννοιολογικής κατασκευής.

Οι υπό μελέτη συσχετίσεις ανεξάρτητων μεταβλητών με ανάλογη διακύμανση των εξαρτημένων μεταβλητών, για τον εντοπισμό των αποκλίσεων του βαθμού «Projectification» ανά

τομέα εργασίας ή υποσύνολο του δείγματος έρευνας, εξετάστηκαν μέσω ελέγχων One way ANOVA (Analysis Of Variance) και One Sample t-Test. Η εργασία στοχεύει τελικώς στην ανάδειξη της επιρροής που ασκεί η στατιστική στην εξαγωγή συμπερασματικών κρίσεων που απηχούν με ακρίβεια τη φύση και ουσία των πρωτογενών στοιχείων και δεδομένων που η έρευνα πεδίου, φέρνει στην επιφάνεια.

ΠΕΡΙΛΗΨΕΙΣ
στα Αγγλικά

ABSTRACTS
in English

Kendall curve-based index of dependence and test for independence between random variables

Georgios Afendras

Aristotle University of Thessaloniki
gafendra@math.auth.gr

The Kendall plot is a graphical method measuring dependence between two random variables. In this talk we present an index of dependence based on the Area Under Kendall curve (AUK); this index satisfies the axioms that established by Rényi for an index of dependence. A non-parametric test of the independence of two random variables is also presented. Finally, simulation results are provided to evaluate this independence test.

Optimal maintenance and inspection policy for high-quality production processes

Vasileios Alevizakos¹, Konstantinos A. Tasiias²

¹National Technical University of Athens, ²University of Western Macedonia

email: basalebiz@yahoo.gr

High-quality processes characterized by minimal defects are the norm in the current real-industry applications. This study proposes an integrated approach that combines quality control with predictive maintenance to enhance the reliability and efficiency of such processes, a task complicated by the prevalence of zero occurrences. The manufacturing systems are subject to a quality shift and equipment failures. A quality control policy based on a Shewhart-type control chart enables real-time process monitoring and aims at detecting the assignable cause or machine failure promptly. The removal of the assignable cause lowers the failure rate, thus constituting a preventive maintenance strategy associated with condition information, a policy known as condition-based maintenance. In this study, the zero-inflated *COM*-Poisson (*ZICMP*) distribution is employed. The *ZICMP* distribution's flexibility to deal with different dispersion patterns, coupled with its versatility-encompassing the zero-inflated Geometric distribution, the zero-inflated Bernoulli distribution, and the *COM*-Poisson distribution as special cases- significantly enhances the applicability of the proposed model.

ISSP in Greece – social surveys and data for a deeper understanding of important social issues

Ioannis Andreadis

Aristotle University of Thessaloniki
john@polsci.auth.gr

The International Social Research Programme (ISSP) is a transnational cooperation program that aims to bring together existing programs related to the social sciences and to set new research goals, focusing on imposing a transnational perspective on individual national studies. Each year, a specific topic is chosen to be researched. Greece participates in ISSP since 2021 and we have already deposited the final dataset of a survey on the topic “Health and Healthcare II”, we are currently working on two surveys on the topics of “Family and Changing Gender V” and “National Identity & Citizenship”, and two pilot surveys on the topics of “Digital Societies” and “Work Orientations V”. All our surveys are conducted following the highest professional standards, ethics and techniques promoted by the “World Association for Public Opinion Research” (WAPOR, 2021) and are used in the new research project “*Data for Inclusive Societies (DATIS)*” (Andreadis, 2024).

References

- Andreadis, I. (2024). *Data for Inclusive Societies (DATIS)* [Research Project]. <https://datis.gr/>
- WAPOR. (2021). *WAPOR Code of Ethics*. World Association for Public Opinion Research. <https://wapor.org/about-wapor/code-of-ethics/>

Sequential interval reliability for semi-Markov systems: modeling and asymptotic properties

Vlad Stefan Barbu^{1,2}, Guglielmo D'Amico³ and Thomas Gkelsinis¹

¹ LMRS, University of Rouen-Normandy, France

² Centre for Demographic Research "Vladimir Trebici", "Costin C. Kiritescu" National Institute of Economic Research of Romanian Academy, Romania

³Department of Economics, University G. d'Annunzio of Chieti-Pescara, Italy

barbu@univ-rouen.fr

The present work is dealing with reliability modelling for multi-state systems with time dependence. We first define the sequential interval reliability, that represents the probability that the system works in a sequence of non-overlapping time intervals. This reliability measure is a generalization of usual reliability indexes such as reliability, interval reliability and availability for homogeneous semi-Markov repairable systems in discrete time. Then we investigate the asymptotic behavior of sequential interval reliability in several different cases.

References

V. S. Barbu, G. D'Amico, T. Gkelsinis, Sequential interval reliability for discrete-time homogeneous semi-Markov repairable systems, *Mathematics*, 9, 1997, 2021. DOI: <https://doi.org/10.3390/math9161997>.

V. S. Barbu, G. D'Amico, T. Gkelsinis, On a mixed transient-asymptotic result for the sequential interval reliability for semi-Markov chains, submitted, 2024.

Applications of Data Analytics for the maintenance and improvement of the Electricity Distribution Network

Konstantinos Bavelis¹, Anastasia Apostoloudia¹, Athanasios Palazis¹, Polydora Raikou¹, Katerina Mitsou¹, Nikolaos Lettas², Dimitrios Stimoniaris², Dimitrios Tsiamitros²

¹Hellenic Distribution Network Operator

²INNORA PC, Spinoff company of the University of Western Macedonia

k.mpavelis@deddie.gr, n.lettas@innora.net

The Hellenic Distribution Network consists of 113.358 km of Medium Voltage Network (MV), 128.211 km of Low Voltage Network (LV) and 165.290 Substations of MV/LV. It essentially covers the whole country and has 7.593,412 Customers. The sheer numbers of it clearly indicate that a straightforward deterministic approach for the monitoring and management of such a Grid would be particularly resource-intensive and more automated, statistical approaches would be both more elegant and efficient. In this work we elaborate on the potential of Data Analytics for the Electricity Distribution Network by showcasing three especially useful applications. The first one is the correlation between weather conditions and faults. This is particularly advantageous as it allows for better resource allocation especially nowadays due to Climate Change. Secondly a spacial analysis of faults can highlight weak spots of the Grid. Repetition and concentration of faults at specific parts of the Network suggest need for local improvements. Lastly, repeated theft of MV/LV transformers is a very costly problem that causes a lot of inconvenience to the consumers. Again by studying the localization and repetition of the incidents, one can spot the transformers that require anti-theft measures and also provide vital information to the Law Enforcement Agencies.

References

- HEDNO official website: <https://deddie.gr/>
- Bavelis, K. Apostoloudia, A. and Tsiamitros, D. (2022). Correlation between weather conditions and faults in the Electricity Distribution Network. *SyNERGY MED 2022*.
- Brown R. E.(2002). *Electric Power Distribution Reliability*. Marcel Dekker Inc. New York-Basel
- Tsioumpri, E. Stephen, B. and McArthur, S.D.J. (2021). Weather Related Fault Prediction in Minimally Monitored Distribution Networks. *Energies 2021, 14, 2053*.

Bayesian sample size determination for detecting heterogeneity in multi-site replication studies

Konstantinos Bourazas¹, Guido Consonni² and Laura Deldossi²

¹Department of Economics, Athens University of Economics and Business, Athens

²Department of Statistical Sciences, Università Cattolica del Sacro Cuore, Milan

bourazas@aueb.gr

An ongoing “replication crisis” calls into question scientific discoveries across a variety of disciplines ranging from life to social sciences. Replication studies aim to investigate the validity of findings in published research and try to assess whether the latter are statistically consistent with those in the replications. While the majority of replication projects are based on a single experiment, multiple independent replications of the same experiment conducted simultaneously at different sites are becoming more frequent. In connection with these types of projects, in this work, we deal with testing heterogeneity among sites; specifically, we focus on sample size determination suitable to deliver compelling evidence once the experimental data are gathered. We propose a Bayesian scheme adopting a Normal-Normal hierarchical model and using the Bayes factor as a measure of evidence.

A comparison of statistical models for mortality graduation

Apostolos Bozikas, Peter Hatzopoulos

University of Piraeus

bozikas@unipi.gr

Mortality rates usually exhibit irregularities of fluctuations due to sampling errors and the random variability inherent in every finite sample we observe in age effects. For this reason, the actuaries apply appropriate statistical methods to produce smoother mortality estimates, a procedure that is called graduation. This work presents some graduation methods based on orthogonal polynomials under the framework of generalized linear models (GLMs). More specifically, we fit GLM models of different distribution families to the age-specific mortality data of various periods for Greece and England and Wales, two countries with a different demographic profile, and then, we evaluate their fitting performance using information criteria and other appropriate tests for this purpose.

Acknowledgement: This work has been partly supported by the University of Piraeus Research Center.

Exploring Political Behavior of First Time Voters through Hierarchical Cluster Analysis and Correspondence Factor

Theodore Chadjipadelis & Georgia Panagiotidou

Aristotle University of Thessaloniki

chadji@polsci.auth.gr

In this study, a two-step multivariate analysis method is presented, utilizing a combination of Hierarchical Cluster Analysis (HCA) and Factorial Correspondence Analysis (AFC). To explain and describe the steps of the method, an example application on a research dataset from young students in Thessaloniki is used, aiming to investigate their behavioral profiles regarding political characteristics. The research aims to explore the interaction between variables such as political behavior, political interest, political mobilization, political information, and attitudes towards elections and the European Union. Concurrently, it attempts to identify the existing frameworks of the students regarding the construction of the 'political self'—how they perceive the concept of democracy—and the 'ethical self'—which moral values they consider most important. Symbolic representation is used, employing images that represent various concepts of specific terms such as 'democracy' or 'moral values' (Marangudakis and Chadjipadelis, 2019). Lastly, to investigate the most significant criteria with which the students form their final choice, the experimental design of conjoint analysis is utilized. HCA categorizes the subjects into new cluster variables, and then these new variables are jointly analyzed with AFC (Benzécri, 1973). The correspondence analysis manages to identify and describe the dimensions of the phenomenon (Greenacre, 2007), explaining the internal contrast between the categories and providing the opportunity for the visualization of the information in a two-dimensional space, a semantic map, which serves as a tool for analysis and interpretation of the phenomenon, highlighting the relationships and competitions among the complex behavioral frameworks of the students participating in the research.

References

- Benzécri, J.P. (1973). *L'Analyse des Données. Tome 2, l'Analyse des Correspondances*. Paris, Dunod.
- Greenacre, M. (2007). *Correspondence Analysis in Practice*. Boca Raton: CRC Press.
- Marangudakis, M. and Chadjiapadelis, T., 2019. *The greek crisis and its cultural origins*. Taylor, C. (1992). *The Ethics of Authenticity*. Harvard University Press.

Supervised statistical (machine) learning for domain estimation with business survey data

Vasilis Chasiotis¹ & Nikos Tzavidis² & Chiara Bocci³ & Paul Smith²

¹Department of Statistics, Athens University of Economics and Business, Athens, Greece

²Department of Social Statistics and Demography, and Southampton Statistical Sciences Research Institute, University of Southampton, Southampton, United Kingdom

³Department of Statistics, Computer Science, Applications "G. Parenti", University of Florence, Florence, Italy
chasiotisv@aueb.gr

We discuss current research on supervised statistical (machine) learning methods (random forests) and extensions to mixed effects random forests, as a flexible method for domain estimation with business survey data. Random forests excel in terms of predictive performance. Automated model-selection and detecting covariate interactions make their use appealing for prediction problems. Mixed effects random forests, however, appear to be going against the algorithmic modelling culture that treats the prediction mechanism as unknown, and are more in line with the data modelling culture. Model-based estimation with business survey data requires careful handling and may include outlier robust estimation, complex modelling of the model variance and use of data-driven transformations. We explore the use of random forest-type algorithms for estimation of finite population parameters. We focus on critically evaluating (a) the role of random effects in

machine learning algorithms, (b) the role of data transformations, and (c) whether machine learning algorithms offer protection under misspecification of linear-type models. Small area predictors are derived by using a smearing-type estimator that has been explored in small area and survey estimation before in the context of outlier-robust estimation. A non-parametric bootstrap MSE estimator is evaluated. We compare machine learning-based predictors to empirical best predictors and outlier robust predictors under a linear mixed model using real business survey data from Italy. This work aims to inform the discussion on the use of machine learning methods in the production of official statistics.

On the application of fractal interpolation to the modelling and reconstruction of seismic data

Polychronis Manousopoulos¹, Vasileios Drakopoulos²

¹Bank of Greece, 21 E. Venizelos Avenue, 10250 Athens, Greece;
PManousopoulos@bankofgreece.gr; ORCID: 0000-0003-1219-3623

²Department of Computer Science and Biomedical Informatics,
University of Thessaly, 35131 Lamia, Greece; vdrakop@uth.gr;
ORCID: 0000-0002-6478-3943.

vdrakop@uth.gr

Fractal interpolation, as defined in [1], is based on the theory of iterated function systems and offers an alternative methodology to traditional interpolation, focusing mainly on data sets that present detail at different scales or some degree of self-similarity. Fractal interpolation has a successful record of applications in a diversity of areas, including, for example, financial, engineering, geographical and medical data sets (see e.g. [2], [3], [4]). In the afore-mentioned applications, fractal interpolation methods have achieved accurate data reconstruction along with considerable compression ratios.

It is known that seismic data possess stochastically self-similar structure with respect to time, space, and magnitude [5]. Therefore, fractal interpolation is a suitable candidate for modelling and reconstructing them. In our present work, we focus on the modelling and reconstruction of seismograms from the Ionian Sea for the period 1970–2016. Specifically, we examine seismic activity with magnitude $M_s \geq 4.5$ and use both self-affine and recurrent fractal interpolation functions to reconstruct the relevant data sets. Moreover, we compare our results to existing methods. The results indicate that fractal interpolation is promising in this area.

References

- [1] Barnsley, M.F. Fractal functions and interpolation. *Constr. Approx.* 1986, 2, 303–329.
- [2] Manousopoulos, P.; Drakopoulos, V.; Theoharis, T. Parameter Identification of 1D Recurrent Fractal Interpolation Functions with Applications to Imaging and Signal Processing. *J. Math. Imaging Vis.* 2011, 40, 162–170.
- [3] Manousopoulos, P.; Drakopoulos, V.; Polyzos, E. Financial Time Series Modelling Using Fractal Interpolation Functions. *AppliedMath* 2023, 3, 510-524.
- [4] Drakopoulos, V.; Manousopoulos, P. Obstetric Ultrasound Modelling and Analysis with Fractal Interpolation Methods. In: Daimi, K., Alsadoon, A., Seabra Dos Reis, S. (eds) *Current and Future Trends in Health and Medical Informatics. Studies in Computational Intelligence*, 2023, 1112, Springer.
- [5] Hirata, T. A correlation between the b value and the fractal dimension of earthquakes. *J. Geophys. Res.*, 1989, 94(B6), 7507–7514.

Joint monitoring of origin and scale of a shifted exponential process with run sum control charts

K. G. Fountoukidis, D. L. Antzoulakos and A. C. Rakitzis

Department of Statistics and Insurance Science, University of
Piraeus

k.fountoukidis@unipi.gr

The exponential distribution is commonly used in quality and reliability engineering to model the time until the occurrence of certain events such as the failure of an equipment, or the times between consecutive failures. A well-known extension of the exponential distribution is the shifted (or two-parameter) exponential distribution, where the origin can be larger than zero. This situation is common in cases where the lifetimes of products exceed a certain time, or in models where the origin parameter denotes a warranty period. Therefore, in these cases the shifted exponential distribution is a more appropriate model. This study proposes a run sum control chart for the joint monitoring of the origin and the scale parameters of a process which is modelled according to the shifted exponential distribution. A Markov chain method is used to evaluate the chart's performance in terms of the average run length. Also, the overall performance of the considered charts is investigated in terms of appropriate measures. Numerical comparisons between the considered run sum charts and other competitive charts, such as the Shewhart-type and the CUSUM-type Max chart are provided. The numerical results show that the proposed run sum charts outperform the Shewhart-type Max chart while under certain circumstances they are at least comparable, if not better, than the CUSUM-type Max chart. Finally, the implementation of the considered run sum charts in practice is discussed via a real-data example.

Keywords: Average Run Length, Expected Average Run Length, Run Sum Statistic, Shifted Exponential Distribution, Single Control Chart, Statistical Process Monitoring.

Enhancing Forest Management Inventory Estimates: A Novel Multivariate Fay-Herriot Model with Preprocessing

Aristeidis Georgakis, Georgios Stamatellos

Aristotle University of Thessaloniki, School of Forestry and Natural Environment, Thessaloniki, Lab of Forest Biometry, PC-54124, Thessaloniki,

e-mail: aridgeorg@for.auth.gr, stamatel@for.auth.gr

This study presents a novel approach to small area estimation (SAE) by integrating trivariate Fay-Herriot models with advanced preprocessing techniques. Cluster analysis, variable selection, and outlier handling are employed to enhance the accuracy of estimates. The model focuses on three dependent variables: woody volume, basal area, and average tree height. Comparative analysis demonstrates the superior accuracy of the multivariate model compared to the basic univariate one, achieving greater precision with a coefficient of variation below 5%.

Keywords: Linear mixed models, EBLUP, multivariate analysis, remoted sensing data, census data, clustering analysis, variable selection.

Enhancing Safety and Efficiency in Underground Marble Quarries through Statistical Analysis of Geotechnical Data

Christos Gravalos, L. Lamprinidis & I. Kapageridis
dmre00011@uowm.gr, ikapageridis@uowm.gr

Introduction

The integrity of the excavation environment is paramount for both safety and operational efficiency in the field of underground marble quarrying. The advent of sophisticated monitoring equipment has ushered in an era where vast amounts of geotechnical data can be collected, presenting an opportunity to leverage statistical analysis for informed decision-making. This thesis explores the significance and methodologies of applying statistical analysis to geotechnical data derived from monitoring equipment in underground marble quarries.

Geotechnical Data Acquisition

Underground marble quarries are dynamic environments with variable stress conditions, moisture levels, and potential for micro-displacement activity. Monitoring equipment such as extensometers, tilt meters, and crack sensors collect data on rock displacement and movement. Data collection is crucial for understanding the geomechanical behavior of the quarry under different operational and environmental conditions.

Statistical Analysis Techniques

The application of statistical analysis to geotechnical data involves several techniques. Descriptive statistics provide a preliminary overview of the data distribution, identifying mean values, variability, and trends over time. Inferential statistics enable the extrapolation from sample data to make predictions about future quarry behavior under similar conditions. Time-series analysis is particularly relevant for detecting patterns and forecasting potential geotechnical failures. Moreover, multivariate analysis can uncover complex relationships between different geotechnical variables, enhancing the predictive accuracy of potential hazard assessments.

Applications in Quarry Operations

Statistical analysis of geotechnical data has profound implications for quarry operations. Real-time data analysis can trigger automated alerts in case of detected anomalies, allowing for immediate

response to mitigate risks. Long-term data analysis aids in the planning of quarry expansion, optimizing the extraction process while ensuring stability and safety. Furthermore, the statistical models developed from historical data can serve as benchmarks for comparing current quarry conditions, enabling the identification of deviations that might signify emerging risks.

Conclusion

The integration of statistical analysis with geotechnical data from monitoring equipment in underground marble quarries represents a significant advancement in quarry management. It not only enhances the safety and health conditions for workers by providing a robust framework for risk assessment and mitigation but also improves operational efficiency through data-driven decision-making. As monitoring technologies and statistical methodologies continue to evolve, the potential for their application in enhancing quarry safety and productivity is boundless, marking a new era in sustainable mineral extraction.

References

Jian Zhou, Yulin Zhang, Chuanqi Li, Haini He, Xibing Li, Rockburst prediction and prevention in underground space excavation, *Underground Space*, Volume 14, 2024, Pages 70-98, ISSN 2467-9674, <https://doi.org/10.1016/j.undsp.2023.05.009>

Mariel Formoso Franqueira, Contribution to the Geomechanical Stability Analysis of a Marble Underground Exploitation using Rockfill, Thesis to obtain the Master of Science Degree in master's degree in mining and Geological Engineering

P Gabriele et al, Sustainable development for underground mining of strategic industrial minerals: the geological and geotechnical modelling as a key factor. The case of the "Marmorino" mineral deposit (Sacile - NE Italy) 2021 IOP Conf. Ser.: Earth Environ. Sci. 833012138, DOI 10.1088/1755-1315/833/1/012138

Prof. Andrea Segalini, ROLE OF GEOTECHNICAL MONITORING: STATE OF THE ART AND NEW PERSPECTIVES, ISSN 2303-8403, Broj 3, 2017.

A novel framework of stochastic mortality model selection

Peter Hatzopoulos, Apostolos Bozikas

University of Piraeus

xatzopoulos@unipi.gr

Stochastic mortality models include age - period - cohort components that must be appropriately estimated to produce efficient mortality forecasts. This work proposes an estimation methodology that is based on the following procedure. First, we adopt generalized linear models (GLMs) to estimate the model parameters and then we apply sparse principal component analysis to extract the related age - period - cohort components. We identify the optimal number of the incorporated age, period and cohort effects of a mortality model by the unexplained variance ratio (UVR) metric to maximize the captured variance of the mortality data, and to regulate the sparsity of the model with the aim of acquiring distinct and significant stochastic components. Our methodology is applied to the mortality data of Greece and England & Wales for different ages and period spans.

Acknowledgement: This work has been partly supported by the University of Piraeus Research Center.

Linear models for continuous paired comparison data

R. Singh¹, G. Iliopoulos², O. Davidov¹

¹Department of Statistics, University of Haifa

²Department of Statistics and Insurance Science, University of
Piraeus

geh@unipi.gr

In this work, graphical linear models for continuous paired comparison data with and without covariates are rigorously analyzed and new necessary and sufficient conditions which guarantee strong consistency, asymptotic normality and the exponential convergence of the estimated ranks are provided. The proposed methods are explored by simulation and applied to the ranking of performance of Greek football teams based on xGoals.

Keywords: Graph Laplacian, large sample properties, least square ranking, regression.

Acknowledgement: This work has been partly supported by the University of Piraeus Research Center.

Spatial Declustering of Exploration Data in Marble Resource Estimation from Irregular Drilling Patterns

Ioannis Kapageridis, Evangelos Gialamas, Charalampos Albanopoulos, and Georgios Anastasiadis

Laboratory of Mining Informatics and Machine Learning
Applications, Department of Mineral Resources Engineering,
University of Western Macedonia
ikapageridis@uowm.gr

Resource estimation of a marble deposit is commonly based on sample data from surface and/or underground drillholes, characterized as to several visual, physical, and mechanical properties (Kapageridis *et al.* 2018). Drilling campaigns in marble quarries tend to be highly irregular as to their spatial density and drillhole orientation, as drilling is driven by the progress of extraction, and the need to expose good quality marble. This process introduces bias to resource estimation and can lead to overestimation of resources, if samples are not treated first for their spatial clustering. The clustered distribution of samples means that areas of good marble quality are potentially oversampled, while areas of poor quality have fewer to no samples leading to a sample distribution that is not representative of the resource that is to be estimated. Several sample declustering methods exist and have been applied to mineral resources estimation (Isaaks and Srivastava 1989; Deutsch and Journel 1998). The procedure described in this paper is based on *cell declustering* to derive sample weights that can reduce the effect of bias caused by irregularly spaced sampling during resource estimation of a marble deposit.

References

- Deutsch CV, Journel AG (1998) GSLIB: geostatistical software library and user's guide. Oxford University Press, New York, p 340.
- Isaaks EH, Srivastava RM (1989) An introduction to applied geostatistics. Oxford University Press, New York, p 561.

Kapageridis, I., Albanopoulos, C. (2018). Resource and Reserve estimation for a marble quarry using quality indicators. *Journal of the Southern African Institute of Mining and Metallurgy*, **118**, 39-45.

A graph-based hybrid analysis of the echo chambers phenomenon: Insights from the Baltimore Riots

**Dimitrios Christos Kavargyris, Konstantinos Georgiou,
Lefteris Angelis**

School of Informatics, Aristotle University of Thessaloniki, 54124
Thessaloniki, Greece

Correspondence email (D.K.): dkavargy@csd.auth.gr

Social media is advancing at a rapid pace, with billions of users logging in daily to exchange opinions, share knowledge and engage in dialogue. More often than not, the discussions revolve around the latest global news, whether these concern politics, war, environmental protection, crises like the COVID-19 pandemic, among others. These sensitive issues can frequently create polarization among discussing parties, resulting in two groups of like-minded users that refuse to listen to the opinions of outsiders. This phenomenon is known as an echo chamber, and it is a great threat to democratic principles, constructive dialogue and civilized behavior. Echo chambers have been found, with the help of bot accounts, to influence elections, public consensus on sensitive matters and the general opinion of the populace on society. Hence, their detection is crucial for recognizing irregularities in the function of social media. Echo chamber detection is a rich field, with multiple studies focusing on different approaches that range from machine learning to network analysis and deep learning. In this paper, we present a graph-based hybrid approach on echo chamber detection, where we apply the Louvain community detection algorithm to split the network into smaller

communities and identify which users have the same opinions based on their posting history. After that we analyze communities to investigate potential echo chambers inside the network. We execute our approach on a dataset consisting of approximately one million tweets concerning the 2015 Baltimore Riots based on some popular hashtags, showcasing its ability to effectively detect echo chambers, quantify polarization and indicate community dynamics.

THREE CENTURIES OF RANDOM LINES: FROM BUFFON'S NEEDLE TO SCALE-INVARIANT NETWORKS

(Theo Cacoullos Memorial Lecture 2024)

Wilfrid S Kendall

University of Warwick

w.s.kendall@warwick.ac.uk

This talk will focus on (a) the history of some stochastic geometry dating from Buffon's famous thought-experiment concerning baguettes, through (b) a variety of proposed geometric models for city and traffic flow, to (c) a general and intuitive description of the author's results concerning a recent approach to spatial transportation networks. There should be something of interest for all (including audience participation in design of a statistical experiment with stochastic geometry overtones)

Keywords: Buffon's needle; Poisson line patterns; spatial transportation models

Acknowledgments: The author acknowledges support of UK EPSRC grant EP/R022100.

REFERENCES

- Aldous, D J and Ganesan, K (2013). True scale-invariant random spatial networks. *Proc. Nat. Acad. Sci. USA* 110.22 8782-8785.
- Aldous D J and Kendall, W S (2008). Short-length routes in low-cost networks via Poisson line patterns. *Advances in Applied Probability* 40.1 1-21.
- Hey, John D, Neugebauer, Tibor M, Pasca, and Carmen M (2010). Georges-Louis Leclerc de Buffon's 'Essays on Moral Arithmetic' (1777), *The Selten School of Behavioral Economics* (Edited by Sadrieh, A, Ockenfels, A). 245-282.
- Kendall, W S (2014). Lines and networks. *Markov Processes and Related Fields* 20.1, 81-106.
- Stigler S M (1991). Stochastic simulation in the nineteenth century. *Statistical Science* 6.1 89-97.

Information Measures for the γ -order Generalized Normal

Christos P. Kitsos¹ and Ioannis S. Stamatiou¹

¹University of West Attica, Egaleo, Greece.

istamatiou@uniwa.gr

The γ -order Generalized Normal (γ -GN) is defined, with the extra, shape, parameter γ . The shape parameter can be evaluated through the Newton-Raphson iterative scheme, thanks to a transformation we introduce. For two independent random variables from the γ -GN the Blachman-Stam inequality holds, and the extended Fisher's, entropy type, information can be evaluated. For given two members of the γ -GN family, the Kullback-Liebler divergence can be evaluated. The special case with $\gamma=2$, referring to the classical Normal distribution, attempts a special investigation.

Simulation Study for the family of Generalized Normal $N_{\gamma}(\mu, \sigma^2)$

Christos P. Kitsos¹ and Ulrich Eschcolb Nyamsi²

¹University of West Attica, Egaleo, Greece.

²University of Aberta, Lisbon, Portugal.

xkitsos@uniwa.gr

There is a need, especially for the data from Economical studies, to work with "fat tailed" distributions. Not only from its definition, but from the evaluation of the appropriate values, the γ -order Generalized Normal (γ -GN) offers a solution to this problem, defined, with the extra shape parameter γ . The normal equations are evaluated, so that the MLE to be evaluated, the Fisher's information matrix is evaluated, so that the variance covariance matrix to be estimated. A simulation study is performed, for different values of the shape parameter and therefore, the Uniform, Laplace and Normal distribution can be evaluated. With different values of the shape parameter the distribution can obtain "fat tails".

How to obtain meaningful statistics from a periodic stochastic process with null ensemble averages

Efstathios Konstantinidis¹, Daniel Dorogi², Laszlo Baranyi³

¹ Department of Mechanical Engineering, University of Western Macedonia, Kozani 50100, Greece

² Savaria Institute of Technology, ELTE Eotvos Lorand University, Budapest, 9700 Szombathely, Hungary

³ Department of Fluid and Heat Engineering, University of Miskolc, 3515 Miskolc-Egyetemvaros, Hungary

In this work we demonstrate how to obtain meaningful statistical properties of a signal from a stochastic periodic process whose ensemble average is null. The particular problem is the flow-induced vibration of an elastically mounted cylinder in oscillatory flow. This configuration results in a displacement signal which has zero ensemble average over many cycles of oscillation. To overcome the problem, we use the Hilbert transform to obtain the instantaneous amplitude of the displacement signal, which is then employed as primary variable of interest. In this way, it is possible to obtain the evolution of the phase-averaged mean amplitude as well as its dispersion over the mean during the cycle of the oscillatory flow. This information allows physical interpretation of the data, which wouldn't be possible otherwise.

References

Dorogi D., Baranyi L., Konstantinidis E. (2023). Modulation and hysteresis in vortex-induced vibration of a spring-mounted cylinder in a slowly varying oscillatory stream, *Journal of Fluids and Structures*, 122, DOI:10.1016/j.jfluidstructs. 2023.103982.

Community Detection in Correlation Networks from Multivariate Time Series

Akylas Fotiadis, Ioannis Vlachos, Dimitris Kugiumtzis

Department of Electrical and Computer Engineering, Aristotle

University of Thessaloniki, University Campus, 54124,

Thessaloniki

e-mail: dkugiu@auth.gr

Community detection is a challenging problem in the study of complex systems. In the setting of complex networks formed from multivariate time series, where nodes are the observed variables and connections assign to interactions quantified with a correlation measure, the goal of community detection is to detect system variables that tend to interact in groups (communities). A prominent approach for community detection uses the Louvain algorithm along with a filter correlation procedure based on random matrix theory (RMT). There are two issues with this approach addressed in this work. First, the RMT may not distinguish well the eigenvalues of significant structure, erroneously assigning to structure trivial correlation when estimated on multivariate time series. Instead of relying on RMT the significance is assessed by appropriate resampling on the multivariate time series, preserving the trivial correlations in the resampled correlation matrix. Second, a detected community in the first step is represented in the next step by a single variable (the average time series of the community components), suppressing in this way the characteristics of the component variables within the community. This is resolved considering the canonical correlation analysis (CCA) as correlation on all the community components (time series). We assess the proposed methodology on various simulation scenarios, focusing on correlation matrices derived from multivariate time series data, using different settings of time series lengths, system dimensions (network sizes), and also under the presence of no communities, weakly connected communities and dense communities. Simulation results show that the suggested modifications lead

to better performance than the baseline algorithm, even for short time series, a situation often occurred in real data. Further, we apply the proposed community detection on correlation networks computed on sliding time windows of stock market data. We compare the estimated communities to the sections of stocks and assess changes in the communities in the presence of financial events.

Dimensionality reduction on data cubes: Application and comparison of various analysis strategies

Z. Kyrana^{1,*}, N. Papafilippou¹, E. Pratsinakis¹, A. M. Michaelidou², A. Markos³, G. Menexes¹

¹ Laboratory of Agronomy, Faculty of Agriculture, Forestry and Natural Environment, School of Agriculture, Aristotle University of Thessaloniki, 54124 Thessaloniki, Greece

² Laboratory of Dairy Science and Technology, Faculty of Agriculture, Aristotle University of Thessaloniki, 54124 Thessaloniki, Greece

³ Department of Primary Education, Democritus University of Thrace, 68100 Alexandroupoli, Greece

* kyranaze@agro.auth.gr

Data cubes are n -dimensional data structures, where $n \geq 3$. In this study, the applicability of several different analysis strategies for dimensionality reduction based on Principal Components Analysis (PCA) was investigated. In Biological Sciences, and especially in the field of Agriculture, the key point of Statistical - Data Analysis is the study of variability, focusing on its structure and components. In this study, some of the proposed strategies were based on the idea of decomposing the total variability of the data cube into between-slices and within-slices variability. Other strategies ignored the effect of data cube slices on the structure of variability. The aim was to propose new insights into analysis strategies for dimensionality reduction on data cubes. A dataset from the field of Agriculture was used for testing and comparing the proposed analysis strategies. The dataset was considered as a three-

dimensional data cube. The third dimension of the data cube was a categorical variable with three categories (three independent slices). The first dimension was consisted of different and equal number of objects per slice. The second dimension was consisted of the same quantitative variables per slice.

References

- Kroonenberg, P. M. (2007). *Applied Multiway Data Analysis*. Wiley Series in Probability and Statistics.
- Menexes, G. (2006). *Experimental Designs in Data Analysis*. Doctoral dissertation, University of Macedonia, Thessaloniki, Greece (in Greek).

Robust Estimation and Testing via a General Family of Measures

Christos Meselidis, Alex Karagrorgoriou

Laboratory of Statistics and Data Analysis
University of the Aegean
mesel.chri@gmail.com

The focus in this work is placed on situations where zero frequency cells occur. The modified (Φ, α) -power divergence family ($M(\Phi, \alpha)$ PDF) is defined for the first time. Furthermore, the minimum modified (Φ, α) -power divergence estimator ($MM(\Phi, \alpha)$ PDE) and the associated modified double index test statistic (MDITS) are defined as well. Asymptotic results are obtained under two situations, namely under model misspecification and when the hypothesized parametric model is correctly specified. The behavior of the proposed family of estimators and test statistics is examined, through an extensive simulation study, not only for the regular case but also when the data are contaminated. Results indicate that, through the proposed methodology we can derive robust estimators with improved efficiency, and stable test statistics, in terms of size, when the data are contaminated.

The pilot survey of ISSP 2025. Work Orientations

Georgios Alexandros Miliias

Aristotle University of Thessaloniki

geomiliias@yahoo.com

This research paper examines the way of organizing the pilot research that will be carried out within the framework of the International Social Survey Program (ISSP) and focuses on the topic of Work Orientations. The pilot survey was carried out in Greece in December 2023. The basic exploratory hypothesis of this paper is the interpretation of political behavior based on the answers of Greek citizens to specific questions of the questionnaire on work orientation, which concern, for example, attitudes toward work and private life, as well as respondents' work organization and working conditions.

On the goodness-of-fit tests for censored data with cure fraction

Marija Cuparić & Bojana Milošević

University of Belgrade, Faculty of Mathematics,
Studentski trg 16, 11000 Belgrade, Serbia

bojana.milosevic@matf.bg.ac.rs

In this presentation, we present a brief review of results within the model specification testing in the presence of cure data. The review is followed by the novel proposal of goodness-of-fit tests for exponential distribution accompanied by their empirical and theoretical properties. A possible generalization would be also explored. The final part of the talk will be dedicated to the presentation of some open directions for future research.

Machine learning algorithmic-trading models based on the mean reversion hypothesis of financial markets.

Seyed Navid Mirpoorian

School of Mathematics and Statistics, University of Sheffield, UK
s.mirpoorian@sheffield.ac.uk

Mean reversion in finance refers to the tendency for prices to revert back to their long or short-term averages after deviating from them. This hypothesis suggests that market participants adjust their positions based on profits or losses, leading them to sell assets after experiencing gains or buy more assets following losses. While some recent studies provide evidence supporting mean reversion towards long-term averages, they do not explore details such as timing and the probability of reverting back to these means. Moreover, there is a lack of research focusing on reversion towards short-term averages.

In this talk, we aim to explore an alternative model based on the Ornstein-Uhlenbeck stochastic differential equation by utilizing a short-term price average instead of the traditional long-term average. We will conduct a statistical analysis of this modified model and subsequently employ it as input for a machine learning model in order to develop trading strategies within financial markets and assess their potential profitability.

The machine learning model that we use, eXtreme Gradient Boosting (XGBoost), is a cutting-edge algorithm known for its predictive power and efficiency. By iteratively improving weak learners and offering customizable options, XGBoost enhances accuracy and scalability in various tasks, making it ideal for our financial market analysis. In this talk, by integrating the modified Ornstein-Uhlenbeck model with XGBoost, we aim to leverage the strengths of both approaches to enhance the accuracy and robustness of the trading strategies.

References

T. G. Bali, K. O. Demirtas, and H. Levy. Nonlinear mean reversion in stock prices. *Journal of Banking & Finance*, 32(5):767–782, 2008.

Han Y, Kim J, Enke D (2023) A machine learning trading system for the stock market based on N-period Min-Max labeling using XGBoost. *Expert Systems with Applications*, 211:118581. <https://doi.org/10.1016/j.eswa.2022.118581>

Nowcasting techniques for European energy data

Anna Mpoukouvala, Aglaia Papakyriou, George Tsaklidis

Department of Mathematics, Aristotle University of Thessaloniki
papakyrioulia@gmail.com

In this study, five different models are proposed to predict three key energy parameters. Data on oil product deliveries, electricity availability and inland gas consumption are utilized. The analysis covers six consecutive months, from October 2023 to March 2024, providing a global overview of the forecasts for periods of low and high energy consumption in five EU countries that differ in a variety of parameters. A wide range of (non-)linear univariate and multivariate approaches were considered in order to achieve the best possible forecast. The performance of two additional indicators was examined for the first two models, and complementary indicators for the third model. In particular, in the third model the indicators were selected taking into account the specific conditions of each European country. The fourth model is based on a Recurrent Neural Network (RNN), while in the last predictive approach, individualized linear models were used for each combination. In summary, the performance and advantages of the chosen models are presented, highlighting important differences and similarities that are considered crucial for the comparison of the approaches.

Extending the Mann–Kendall test to allow for measurement uncertainty.

Stavros Nikolakopoulos^{1,2}, Eric Cator³ and Mart P. Janssen⁴

¹Department of Psychology, University of Ioannina

²Department of Biostatistics, University Medical Center Utrecht

³Department of Mathematics, Radboud University Nijmegen

⁴Transfusion Technology Assessment Group, Department of Donor Medicine Research, Sanquin Research, Amsterdam

The Mann–Kendall test for trend has gained a lot of attention in a range of disciplines, especially in the environmental sciences. One of the drawbacks of the Mann–Kendall test when applied to real data is that no distinction can be made between meaningful and nonmeaningful differences in subsequent observations. We introduce the concept of partial ties, which allows inference while accounting for (non)meaningful differences. We introduce the modified statistic that accounts for such a concept and derive its variance estimator. We also present analytical results for the behaviour of the test in a class of contiguous alternatives. Simulation results which illustrate the added value of the test are presented. We apply our extended version of the test to some real data concerning blood donation in Europe.

References

Nikolakopoulos, S., Cator, E., & Janssen, M. P. (2023). Extending the Mann–Kendall test to allow for measurement uncertainty. *Statistics*, 57(3), 577-596.

Incorporating Subjective Survival Information in Bayesian Mortality Models

Apostolos Papachristos, Dimitris Fouskakis

Department of Mathematics, National Technical University of Athens, Greece

apapachristos@math.ntua.gr fouskakis@math.ntua.gr

Subjective survival probabilities incorporate individuals' views about their own future survival, and they are associated with actual mortality patterns. The objective of this study is to apply a Bayesian methodology to investigate whether subjective survival information is useful for predicting actual mortality. To achieve this objective, we adopt a two-step process. In the first step, we use a Bayesian linear regression model, under default priors, on the logit transformation of the subjective mortality probabilities to estimate the posterior distribution of the regression coefficients for the selected explanatory variables. In the second step, we fit Bayesian logistic regression models on actual mortality, using a variety of priors derived from the posterior distributions of the first step. The data for this study is available from the Health and Retirement Study Waves 13 and 14. Our conclusion is that subjective mortality probabilities contain additional information which is useful for predicting actual mortality.

Keywords: Subjective Mortality Probabilities, Bayesian Demography, HRS, Mortality, g-Prior

References

- Liang, F., Paulo, R., Molina, G., Clyde, M.A., and Berger, J.O. (2008). Mixtures of g priors for Bayesian variable selection. *Journal of the American Statistical Association*, **103**: 410–423.
- Fouskakis, D., Ntzoufras, I., and Draper, D. (2009). Bayesian variable selection using cost-adjusted BIC, with application to cost-effective measurement of quality of health care. *The Annals of Applied Statistics* **3**: 663 – 690.
- Papachristos, A., Verropoulou, G., Ploubidis, G.B., and Tsimbos, C. (2020). Factors incorporated into future survival estimation among Europeans. *Demographic Research* **42**: 15–56.

A contribution to the comparative evaluation of efficiency indicators in Eurozone using DEA methodology

Thomas E. Pappas, Andreas C. Georgiou, Georgios Tsaples

Quantitative Methods and Decision Analysis Laboratory
Department of Business Administration, University of Macedonia
mba23018@uom.edu.gr

This study introduces a methodological framework for developing composite sustainability indicators, employing Data Envelopment Analysis (DEA) to aggregate economic, environmental, and social indicators into composite indices. By utilizing DEA, the study aims to mitigate biases associated with assigning weights, thereby enabling a more robust assessment of sustainability efficiency among 20 Eurozone members. The objective is to provide policymakers and governing bodies with actionable insights to facilitate informed decision-making and enhance the sustainability performance of the countries under scrutiny. A comprehensive literature review spanning 25 papers on DEA and sustainability from 2021 to 2023 was conducted. This review underscored the lack of a unified definition of sustainability while noting the prevalent adoption of the three-pillar framework (economic, environmental, and social). Additionally, it highlighted efforts by some researchers to integrate social and Research and Development (R&D) indicators into DEA models to gauge sustainability efficiency. Data for the study were sourced from Eurostat and the World Bank databases. Subsequently, the selected indicators were applied using two DEA variations (classic and SBM) across 10 different scenarios, yielding 10 final indices. These indices encompass three sustainability indices, six eco-efficiency indices, and one socio-economic index. Analysis of the results revealed that Ireland, Luxembourg, and Malta consistently achieved maximum efficiency scores across all sustainability scenarios, with Ireland and Luxembourg demonstrating efficiency across all eco-efficiency variations. Conversely, Austria, Estonia, Finland, and Slovakia exhibited the lowest scores in terms of sustainability indices, while Croatia, Finland,

Latvia, Lithuania, Slovakia, and Slovenia showed the poorest performances in eco-efficiency indices. Finally, the findings suggest that many Eurozone members are falling short of realizing their full potential output levels, particularly concerning GDP and Overall Life Satisfaction, given the available labor force.

This research was supported by the Hellenic Foundation for Research and Innovation (H.F.R.I) under the “2nd Call for H.F.R.I. Research Projects to support Faculty Members & Researchers”. (Project: 3154).

The pilot survey of ISSP 2024. Political attitudes and digital societies

Eleni-Maria Papachristou

School of Political Sciences, Aristotle University of Thessaloniki
(AUTH)

pelenimar@polsci.auth.gr

The paper will be presented as a part of the section entitled “ISSP in Greece-social surveys and data for a deeper understanding of important social issues”.

In recent years, there has been an increase in the use of technology in all areas of daily life, and one of the most important changes concerns the development of robotic technologies to serve daily activities. It has been observed by scholars that the evolution of robots has brought about effects on some social issues, such as welfare policies, legislation, as well as from an ethical point of view.

The aim of the proposed paper is to study voters' political attitudes towards digital societies issues, such as the impact of digitization on communication patterns and media consumption, Civic Engagement in the Digital Age and their attitudes towards artificial intelligence (AI), using the data from the pilot survey of ISSP 2024. Moreover, we will examine the level of congruence between politicians and voters about digital societies.

Relying on the data of Comparative Candidates Study (CCS) and the International Social Survey Programme (ISSP) 2024 pilot data, we analyze public opinion on their behavior towards robots and AI (Artificial Intelligence). In addition, in this paper will be used data from Twitter (X), in order to examine the frequency of digital issues addressed on tweets from candidates and general public and how these issues affect congruence.

simAlpha. A new algorithm to help scientists to assess questionnaire reliability

Papadopoulos Dimitris¹

¹ Statistician BasisPap; info@basispap.com

Various academic fields employ questionnaires as a means of gathering quantitative data from a representative sample of a population. Several professionals, including applied psychologists, sociologists, and market researchers, make inferences based on statistical analysis of collected responses. However, how can one ascertain the reliability of the answers? Cronbach measured the dependability of this by introducing a coefficient called Cronbach's alpha (α), which theoretically varies between 0 and 1. If the value of α is close to 0, then the quantified answers are not dependable at all. Conversely, if α is close to 1, the answers are highly reliable. Typically, if $\alpha \geq 0.8$, the answers are deemed credible according.

Several attempts have been made to simulate Cronbach's Alpha with several reported problems. Bindak, (2013) reports difficulties to approach values of alpha greater than 0.5 with mean value per question totally random. Algorithm simAlpha in R has succeeded to produce simulated alpha on specific prearranged value. In addition, any value of alpha can be approximated with pre-defined mean value per question for any given question. This algorithm can help any scientist who wants to estimate questionnaire attributes with minimum cost and time.

References

- Bindak R., (2013). Relationship between Randomness and Coefficient Alpha: A Monte Carlo Simulation Study. *Journal of Data Analysis and Information Processing*, 2013, 1, 13-17
- Leontitsis, A., Page J., (2007). A simulation approach on Cronbach's alpha statistical significance. *Mathematics and Computers in Simulation* 73 (2007) 336–340

Strategies for Dimensionality Reduction of Categorical Data

**N. Papafilippou^{1*}, E. Pratsinakis¹, Z. Kyrana¹, A. Markos²,
C. Dordas¹, G. Menexes¹**

¹ Laboratory of Agronomy, School of Agriculture, Aristotle
University of Thessaloniki, 54124 Thessaloniki

² Department of Primary Education, Democritus University of
Thrace, 68100 Alexandroupoli

* npapafil@agro.auth.gr

This paper presents a comprehensive overview of strategies specifically tailored for handling multivariate categorical data in the context of dimensionality reduction. Multivariate categorical data pose unique challenges in data analysis due to their discrete and often high-dimensional nature. Dimensionality reduction techniques play a crucial role in uncovering meaningful patterns and reducing computational complexity in subsequent statistical analyses. Beginning with a thorough review of existing techniques, we outline the limitations of traditional methods when applied to categorical data and underscore the necessity for specialized approaches. Our focus extends to preprocessing steps crucial for ensuring the effectiveness of subsequent dimensionality reduction techniques, focusing mainly to encoding schemes. We then delve into various dimensionality reduction methodologies, such as Correspondence Analysis (CA), Multiple Correspondence Analysis (MCA), Categorical Principal Component Analysis (CatPCA), Factor Analysis for Mixed Data (FAMD), Non Linear Canonical Correlation Analysis (NLCC) and hybrid approaches (e.g. Multiple Factor Analysis (MFA)). Through empirical evaluations on datasets sourced from the biological domain, we assess the performance of these techniques and provide insights into their strengths, limitations and applicability to different types of categorical data. We also discuss practical considerations, such as computational efficiency and interpretability, which are crucial for real-world

applications. Furthermore, we review recent advancements in the field, such as machine learning algorithms, feature selection and extraction. Through a synthesis of theoretical insights and practical recommendations, this paper aims to provide researchers and practitioners with a comprehensive guide for effectively handling multivariate categorical data in the context of dimensionality reduction. By adopting appropriate strategies, analysts can uncover hidden patterns, reduce computational burden, and gain deeper insights from complex categorical datasets.

References

- Agresti, A. (2012). *Categorical data analysis* (Vol. 792). John Wiley & Sons.
- Hair, J. F., Black, W. C., Babin, B. J., & Anderson, R. E. (2010). *Multivariate Data Analysis: A Global Perspective* (7th ed.). New Jersey: Pearson Education, Inc.
- Menexes, G. Experimental Designs in Data Analysis. Doctoral dissertation, University of Macedonia, Thessaloniki, 2006.
- Michailidis, G., & De Leeuw, J. (1998). The Gifi System of Descriptive Multivariate Analysis. *Statistical Science*, 13(4), 307-336. <https://doi.org/10.1214/ss/1028905828>

Temporal Dynamics of Time Series Topics: A Convex Hull Visualization Approach.

Grigorios Papageorgiou, Ekaterini Skamnia, Polychronis Economou

Department of Civil Engineering, University of Patras, Greece
peconom@upatras.gr

Topic modeling serves as a rapid method to extract the essence of a textual dataset, providing a more efficient way to navigate the complexity inherent in large amounts of textual information. In the present work, topic modeling is also utilized to identify any changes and trends in a given topic across different time periods, helping to understand the evolving landscape of scholarly discourse and knowledge dissemination. To achieve this, a document visualization approach is initially employed. This visualization enables the understanding of subtopics, their relations, and similarities, by extracting important geometric characteristics from the areas defined by the documents of each subtopic and comparing them across different periods. These geometric characteristics are related to the area of each convex hull defined for each subtopic and allow gaining numerous insights into the development and the importance of each topic over time. In the present work, abstracts related to the domain of time series are gathered from two distinct periods, wherein specific subtopics of interest are identified, enabling the investigation of potential variations across different temporal intervals.

Exploring Disease Dynamics: A Markovian SIRD Model for Transmission Assessment

Vasileios E. Papageorgiou*, **George Tsaklidis**

Department of Mathematics, Aristotle University of Thessaloniki,
Thessaloniki, Greece

[*vpapageor@math.auth.gr](mailto:vpapageor@math.auth.gr)

In contemporary epidemiology, numerous innovative methods provide reliable estimates of how epidemics unfold within a population. We introduce a continuous-time Markovian model tailored to the characteristics of the SIRD (Susceptible, Infected, Recovered, Deceased) epidemiological framework. New stochastic descriptors are explored, encompassing both population-level and individual-oriented dynamics. We propose recursive formulas and algorithmic procedures for determining distributions and moments of features such as epidemic extinction, alert times, and expected mortality times of infected individuals. Additionally, we provide remarks for representing higher-order matrices and system solutions, significantly reducing computational complexity. Unlike approaches primarily focused on fitting disease evolution, our aim is to complement such studies by emphasizing additional important features that influence the trajectory of an epidemic. We validate our methodology using data from the 2022 monkeypox outbreak in India. Ultimately, these indicators can aid public health authorities in adjusting lockdown durations to strike a balance between health benefits and economic consequences. Insights into mortality times can facilitate early coordination of medical resources, potentially reducing the heightened risk of mortality.

References

- Papageorgiou, V.E., Tsaklidis, G. (2023). A Stochastic SIRD model with imperfect immunity for the evaluation of epidemics. *Applied Mathematical Modelling*, **124**, 768-790.
- Artalejo, J.R., Economou, A., Lopez-Herrero, M.J. (2010). On the number of recovered individuals in the SIS and SIR stochastic epidemic models, *Mathematical Biosciences*, **228**, 45-55.
- Amador, J., Gómez-Corral, A. (2020). A stochastic epidemic model with two quarantine states and limited carrying capacity for quarantine. *Physica A*, **554**, 121899.

Integration of big data analytics in digital twinning of mineral deposits

Agni Patra, Ioannis Kapageridis

Laboratory of Mining Informatics and Machine Learning
Applications, Department of Mineral Resources Engineering,
University of Western Macedonia
dmre00004@uowm.gr

Having access to in-depth data is necessary for the mining industry to support accurate decision-making, as well as meet a growing list of sustainability and compliance requirements. Digital twin technology is rapidly gaining ground in this data-driven industry to supply this crucial data. A digital twin is a virtual copy of a physical system or object that is constantly updated with data from its real-world physical counterpart, allowing accurate and in-depth simulation, analysis, and monitoring. The application of digital twinning in the case of mineral deposits has been considered recently (Hodjkinson and Elmoutie, 2020; Kumar and Dimitrakopoulos, 2022; Patra *et al.* 2022). Mineral deposits and the associated resources are not dynamic systems or entities. However, our perception of a mineral deposit is dynamic and changes through time during various stages of mineral exploration and later during mining with the inflow of new information and data from various sources. The application of big data analytical procedures is

crucial in digital twinning of mineral deposits and opens the door to a new age of mineral resource modelling and mine planning tools. This paper presents the way big data analytics can be applied for this purpose and the associated challenges.

References

- Hodjkinson, J.H., Elmouttie, M., 2020. Cousins, Siblings and Twins: A Review of the Geological Model's Place in the Digital Mine. *Resources* 2020, 9, 24; DOI:10.3390/resources9030024.
- Kumar, A., Dimitrakopoulos, R., 2022. Updating geostatistically simulated models of mineral deposits in real-time with incoming new information using actor-critic reinforcement learning. *Computers & Geosciences* 158 (2022) 104962, DOI: 10.1016/j.cageo.2021.104962.
- Patra, A., Kapageridis, I., Asvesta, A., Sinatkas, I., 2022. Framework for the Development of a Mineral Resource Digital Twin. 16th International Congress of the Geological Society of Greece, *Bulletin of the Geological Society of Greece*, Sp. Publ. 10.

AI Science and Engineering: A new scientific discipline?

Prof. Ioannis Pitas

Aristotle University of Thessaloniki

pitas@csd.auth.gr

Is “AI Science and Engineering” an upcoming scientific discipline that can fuse AI, brain and mind studies and social engineering? As Artificial Intelligence (AI) studies and research flourish worldwide, particularly due to the explosive growth of Generative AI, it is worth debating whether we already observe the birth of a new discipline in the exact sciences that goes beyond the classical specializations of Computer Science (CS) and Electrical/Computer Engineering (ECE). It is natural that AI and its various domains, notably Machine Learning, will share methods and curriculum with both CS and ECE. The relation between Machine Learning and Mathematics, notably Statistics must also be stressed.

However, the current scientific arrangements are not really enough to move from our current Data/Information Society to a Knowledge Society. We certainly need a fusion of AI with brain and mind studies, notably Neuroscience, Cognitive Science and Psychology. Furthermore, AI has a huge impact on both our society and environment.

We already observe strong AI applications in social engineering domains, e.g., on recommendation systems, on-line marketing, social media, meta-societies and fake news. It is natural that such applications will influence the AI Science and Engineering discipline itself. Furthermore, as humans consist of matter and live in a spatiotemporally evolving environment, life and environment studies, e.g., on matter complexity, can have a fundamental impact on both the understanding of life and human intelligence and on the development of AI.

However, as students cannot be polymaths, are all the above issues too many to fit in one scientific discipline? Is there a danger that we sacrifice scientific depth to interdisciplinarity? Can we envisage other sister-disciplines, e.g., Mind and Social Science and Engineering and/or Bioscience and Engineering that can address students with interests and background in Liberal Arts and Sciences or Biology/Medicine/Health Sciences?

No matter their exact form, AI Science and Engineering and its sister disciplines have many great challenges to address. Here is a partial list:

Is AI Science and Technology a scientific discipline in its own right?

How can we quantify knowledge?

Can Virtual Reality truly empower meta-societies or is it just a hype?

Can AI-powered human-centred computing surpass human intelligence?

Can we create self-conscious machines?

Can Mind and Social Engineering manipulate human behaviour and social functions?

How do social media facilitate disinformation?

What are the envisaged effects of AI and IT on our personal relations and sexual life?

How can we not only protect but also monetize our personal data?

Can AI help devising new political systems?

How are irrationalism, anti-elitism, and social media disinformation related?

Can new technologies ignite social revolutions?

Is life and intelligence due to matter complexity?

Can we patch parts of our brain?

Is climate controllable through Geoengineering?

Can humanity progress without resorting to energy-intensive technologies?

As each of them needs an entire book to be properly addressed, this lecture will simply introduce few of these challenges for further discussion and debate. For more information, the interested reader can resort to the rich AI literature.

Keywords: Artificial Intelligence, Generative AI, Machine Learning

References

1. I. Pitas, "[Artificial Intelligence Science and Society Part A: Introduction to AI Science and Information Technology](#)", Amazon/Createspace, 2023.
2. I. Pitas, "[Artificial Intelligence Science and Society Part B: AI Science, Mind and Humans](#)", Amazon/Createspace, 2022.
3. I. Pitas, "[Artificial Intelligence Science and Society Part C: AI Science and Society](#)", Amazon/Createspace, 2022.
4. I. Pitas, "[Artificial Intelligence Science and Society Part D: AI Science and the Environment](#)", Amazon/Createspace, 2022.

Some New Developments in the Credible Distribution Estimation

Georgios Pitselis

University of Piraeus

pitselis@unipi.gr

This paper presents the credible distribution estimation of forecasting the distribution of individual risk in cases where the observations are weighted or are grouped in intervals. The credibility estimation is obtained by applying the optimal projection theorem. Distribution credibility estimators are also established and numerical illustrations are presented. The main contribution of the paper is how to embed the empirical distribution to credibility modelling in the form of the Buhlmann & Straub (1970) model. In the first part of the paper we present the model of the weighted credibility distribution and in the second part, a model that applies to a grouped data in intervals. With our models, we examine two datasets, one with motor claims amount and the number of motor claims from 10 selected European countries for the period 2004 to 2020, and a second with monthly returns from July 1926 - July 2022 for 10 industry portfolios.

Keywords: credibility distribution estimation; empirical Bayes.

Acknowledgments: This work has been partly supported by the University of Piraeus Research Center.

Main References

- Buhlmann, H., Straub, E., 1970. Glaubwürdigkeit für Schadensätze. Mitteilungen der Vereinigung Schweizerischer Versicherungsmathematiker, 70, 111-133.
- De Vylder, F., 1976. Geometrical Credibility. Scandinavian Actuarial Journal, 121-149.
- Jewell, W.S., 1974. The Credible Distribution. ASTIN Bulletin, 7(3), 237-269.
- Pitselis, G., 2023. Credibility Distribution Estimation with Weighted or Grouped Observations. Risks, 12(1), 10.
- Cai, X., Wen, L., Wu, X., and Zhou, X., 2015. Credibility Estimation of Distribution Functions with Applications to Experience Rating in general insurance. North American Actuarial Journal, 19, 311-335.

Strategies for dimensionality reduction of mixed-type data: encoding strategies and techniques

**E. Pratsinakis^{1*}, Z. Kyrana¹, N. Papafilippou¹, A. Markos²,
C. Dordas¹, G. Menexes¹**

¹ Laboratory of Agronomy, Faculty of Agriculture, Forestry and Natural Environment, School of Agriculture, Aristotle University of Thessaloniki, 541 24 Thessaloniki

² Department of Primary Education, School of Education, Democritus University of Thrace, 681 00 Alexandroupoli
[*epratsina@agro.auth.gr](mailto:epratsina@agro.auth.gr)

Mixed-type data consists of a mix of both scale and categorical data (measured on nominal and/or ordinal scale), within the same dataset. Analyzing mixed-type data presents unique challenges because traditional statistical and data analysis methods often require variables to be of the same type. Thus, special methods, techniques and strategies are needed to handle and analyze mixed-type data effectively. There are two main “philosophies” for analyzing mixed-type data. The first, advocated by the Dutch School of Data Analysis (GIFI System), suggests transforming the categorical data into scale data. Consequently, categorical and scale data are entered in the same data analysis method as scale data. Conversely, the French School of Data Analysis proposes converting scale data into categorical. Consequently, categorical and scale data are entered in the same data analysis method as categorical data. Another view is based on the idea to test various data recoding strategies. For example, to transform categorical data to scale data and the opposite. However, these transformations may result in a loss of information. The aim of this study is to compare various encoding strategies, applied on the same dimensionality reduction method and the same data set (e.g. Multiple Correspondence Analysis, Principal Components Analysis and related methods) with methods suitable for handling mixed type data (e.g. Non-Linear Categorical Principal Components Analysis, Factor Analysis of Mixed Data and related methods). This comparison highlights the importance of applying different strategies of encoding data

until a satisfactory solution of dimensionality reduction is found or the validity of the results is reinforced.

On the usual stochastic order and covariance identities

Georgios Psarrakos

Department of Statistics and Insurance Science, University of
Piraeus, Greece
gpsarr@unipi.gr

In this paper, further results for a probabilistic analogue of the mean value theorem proposed by Di Crescenzo (1999) are studied, focused on Normal and Poisson distributions. By applying this theorem for two normally distributed random variables with the same variance, a new symmetric unimodal distribution is obtained, and statistical measures of it are provided. By introducing a family of real valued continuous distributions, covariance identities are also investigated. Covariance identities are also explored for a Poisson distributed random variable. Some illustrative examples are finally given.

Acknowledgement: This work has been partly supported by the University of Piraeus Research Center.

References

- Chen, L.H.Y. (1975). Poisson approximation for dependent trials. *Annals of Probability*, **3**, 534-545.
- Di Crescenzo, A. (1999). A probabilistic analogue of the mean value theorem and its applications to reliability. *Journal of Applied Probability*, **36**, 706-719.
- Papadatos, N. and Papathanasiou, V. (2001). Unified variance bounds and a Stein-type identity. In *Probability and Statistical Models with Applications* (C.A. Charalambides, M.V. Koutras and N. Balakrishnan, eds.) 87-100. New York: Chapman & Hall/CRC.
- Shaked, M. and Shanthikumar, J.G. (2007). *Stochastic Orders*. Springer, New York.
- Stein, C.M. (1972). A bound for the error in the normal approximation to the distribution of a sum of dependent random variables. In *Proceedings of Sixth Berkeley Symposium on Mathematical Statistics and Probability 2*, 583-602. Berkeley CA: Univ of California Press.

The 2023 Hellenic National Election Study (ELNES2023)

Vasileios Simpersis

Postgraduate Student, Department of Political Sciences, Aristotle
University of Thessaloniki

e-mail: simpserv@polsci.auth.gr

The national elections held in Greece in 2023 lead to the need to conduct the national post-election voter survey (ELNES 2023), in order to draw reliable conclusions about the main patterns of electoral behaviour in this particular election. This is the context of the paper entitled "The 2023 Hellenic National Election Study (ELNES2023)", which pay particular attention to the Greek case. For the needs of the research, questionnaires used in a random sample of Greek citizens, which contain the core questions of the international research network Comparative Study of Electoral Systems (CSES), in which member countries participate with the main objective of collecting and utilizing data on the electorate, its behavior and its comparative study, as well as questions adapted to the Greek case. For the purposes of the study, innovative data analysis tools (statistical programming softwares) be used, which enable the implementation of the complex searches of the survey. The paper emphasizes both the methodological approach and the role of statistical analysis packages in the social sciences and the important results they produce. After the collection of data, the process of their analysis follow, aiming to draw conclusions about the views of Greek citizens on politics and democracy, their attitude towards elites ,the influence of the media and social networks in their decision making, as well as the factors that led to the choice of party and participation. The positive view of Greek citizens on the economy, the positive views of citizens on the government's handling of the pandemic, the particular issues that concerned voters, are some of the first elements of the survey that can explain the paths of electoral behaviour in the Greek elections of 2023. The analysis and processing of the data leads to the extraction of important information about the choices and opinions of voters in the 2023 national elections, while it also gives the possibility to make comparisons about different electoral contests in other countries.

References

- Andreadis, I. (2012). To clean or not to clean? Improving the quality of VAA data. XXII World Congress of Political Science (IPSA). <http://ikee.lib.auth.gr/record/132697/files/IPSA2012.pdf>
- Andreadis, I. (2022). Survey Data Collection and Data Quality. In I. Andreadis, Y. Stavrakakis, & E. Teperoglou (Eds.), *Proceedings of the DataPopEU Conference (2022): Populism and Euroscepticism in Perspective* (pp. 21–43). Sofia PA.
- Andreadis, I. (2020). Text Message (SMS) Pre-notifications, Invitations and Reminders for Web Surveys. *Survey Methods: Insights from the Field*, 1–12. <https://doi.org/10.13094/SMIF-2020-00019>
- Knight, K., & Marsh, M. (2002). Varieties of election studies. *Electoral Studies*, 21(2), 169–187. [https://doi.org/10.1016/S0261-3794\(01\)00016-6](https://doi.org/10.1016/S0261-3794(01)00016-6)
- Wright, G. C. (1993). Errors in Measuring Vote Choice in the National Election Studies, 1952-88. *American Journal of Political Science*, 37(1), 291-316.

The survey of ISSP 2022. Political attitudes towards changing gender roles and family

Nikoletta Sivenou

School of Political Sciences, Aristotle University of Thessaloniki
(AUTH)

sivenoun@polsci.auth.gr

The paper will be presented as a part of the section entitled “ISSP in Greece-social surveys and data for a deeper understanding of important social issues”.

Nowadays, the theories about gender and specifically about the development of gender identity have strongly concerned many theoreticians and researchers from various disciplines, taking into account the important role of the family, which constitutes a framework for gender development and perpetuation (Matlin, 2000).

The paper intends to systematically investigate political attitudes towards gender and family issues and their interaction with voting behavior in Greece. Moreover, relying on the data of 2022 International Social Survey (ISSP) on the Family and Changing Gender Roles and 2023 Hellenic Candidate Survey (ELCS 2023), we aim to study the attitudes of Greek voters and candidates on a number of related issues using Greece as a case study. In addition, data from social media, specifically from Twitter (X), will be used, trying to explore and develop some innovative methods for the collection, analysis and interpretation of this data about candidates' and voters' attitudes towards gender roles and family. In particular, the frequency of use of gender and family issues on the tweets of candidates and voters (wordclouds) will be main points and also will allow us to study the congruence of voters and candidates at national level and how these issues affect them. The collection and analysis of data from candidates' and voters' personal Twitter accounts and tweets have been carried out using tools, such as the R statistical program.

References

Matlin, M. W. (2000). *The Psychology of Women*. Orlando: Harcourt College Publishers.

Process Monitoring of Bivariate Poisson Data

Ekaterini Skamnia¹, Polychronis Economou¹, Sotirios Bersimis², Petros E. Maravelakis²

⁽¹⁾ Department of Civil Engineering, University of Patras, Greece

⁽²⁾ Department of Business Administration, University of Piraeus, Greece

peconom@upatras.gr

Statistical Process Monitoring (SPM) comprises of a variety of statistical techniques to analyze and control variation in a process, reassuring an efficient operation in industries such as healthcare, manufacturing etc. A challenging part, however, within SPM, is dealing with batch processes. Under this scenario, a finite quantity of products is produced in batches in a specific period. Contrary to the usual processes where data are modelled using a continuous distribution, batch processes are assumed to follow a discrete distribution. In the present work, the focus is on a problem dealing with the monitoring of bivariate Poisson data arising from a batch process in the food industry. More specifically, we assume that a process produces batches of products, a sample of the products is drawn and the number of events (e.g. bacteria) is measured, in order to identify possible increases during production.

Bayesian Inference of Spatio-Temporal Dynamic Models with Application to Air-Pollution Data

Kostas Triantafyllopoulos

School of Mathematics and Statistics, University of Sheffield, UK

k.triantafyllopoulos@sheffield.ac.uk

This talk concerns Bayesian estimation of dynamic space-time models. In many subject areas data are collected in a number of locations over time. Examples include environmental data such as rainfall or temperature data, macroeconomics data such as

unemployment and gross-domestic product observed over different countries and health surveillance data. In recent years, there has been some interest in modelling high dimensional data observed in space and time, a typical example might be river flow data (Wilke and Cressie, 1999).

In this talk we discuss Bayesian inference of the dynamic space-time model proposed first in Wilke and Cressie (1999). This model considers that the output signal is a noisy version of a stochastic process, which is continuous in space and discrete in time. This process is defined as an integral equation, which is then approximated via a set of orthonormal basis functions. The model is then cast in state-space form and the Kalman filter may be called upon for efficient estimation. For the basis-functions, in this work we propose a wavelet decomposition, which is localized in both frequency and time and hence offers certain advantages over competitor Fourier decompositions. We propose a Bayesian hierarchical model and a corresponding Metropolis sampling scheme for inference. In this presentation we will briefly discuss the structure of the hierarchical model and the priors of the parameters subject to estimation.

We use simulated and real data to illustrate the methods. For the simulated data we simulate images over a number of locations and different degrees of smoothness and we compare them to estimated (inferred) images by our algorithm. For the real data we consider air-pollution observations, consisted of nitric oxide levels from the weather station in Athens. Our approach estimates the signal (observations) very accurately as well as the uncertainty around them.

In the talk we will briefly discuss some generalizations of our approach to non-Gaussian spatio-temporal modelling, in particular considering our observations are generated by multivariate gamma distributions (Mithai and Moschopoulos, 1992).

References

- Wilke, CK & Cressie, N (1999). A dimension reduced approach to space-time Kalman filtering. *Biometrika*, **86**, 815-829.
- Mithai, AM & Moschopoulos, PG. (1992). A form of multivariate gamma distribution. *Annals of the Institute of Statistical Mathematics*, **44**, 97-106.

An Archimedean copulas-based approach for m -consecutive- k -out-of- n : F systems under exchangeability

Ioannis S. Triantafyllou

Department of Statistics & Insurance Science, University of Piraeus
itriantafyllou@unipi.gr

In the present work we provide a signature-based study of a reliability network consisting of identical m -consecutive- k -out-of- n : F structures with exchangeable components. The dependency of the components of each system is modelled by the aid of well-known Archimedean copulas. Exact formulae for determining the expected lifetime of the underlying reliability scheme are provided under different Archimedean copulas-based assumptions. Several numerical results are carried out to shed light on the performance of the resulting consecutive-type design.

References

- Eryilmaz, S., Koutras, M.V. and Triantafyllou, I.S. (2011). Signature based analysis of m -consecutive k -out-of- n : F systems with exchangeable components. *Naval Research Logistics*, **58**, 344-354.
- Eryilmaz, S. (2011). Estimation in coherent reliability systems through copulas. *Reliability Engineering & System Safety*, **96**, 564-568.

Acknowledgements

This work has been partly supported by the University of Piraeus Research Center.

The survey of ISSP 2023. The impact of National Identity and Citizenship on political attitudes

Styliani Tseliou

Aristotle University of Thessaloniki

stellatseliou@gmail.com

The evolving changes due to globalization and immigration, are transforming the established ways nations define themselves and affect the traditional dynamic of societies. The study of these changes provides valuable knowledge of the impact they have on societies and citizens. To explore the importance of these issues in Greek politics and policy making, this study associates the International Social Survey Programme 2023 Module of National Identity and Citizenship with citizens' attitudes and voting behavior. International Social Survey Programme is a collaborative programme that studies in general citizens' beliefs and attitudes towards social issues in many countries. The 2023 Module questionnaire includes questions about democracy, immigration, globalization and citizenship. The primary question of the survey is to study Greek citizens' attitudes on the aforementioned points and estimate its impact on their electoral and voting behavior. At the same time, the study focuses on correlating participants' party preferences to common attitudes related to specific political issues.

Pitfalls in interpreting the predictive power of technical trading rules and weak-form market efficiency test results. An empirical investigation for the Athens Stock Exchange.

Alexandros E. Milionis^{a,*} and Varlagkas Th. Vasileios^{a,b}

^a University of the Aegean, Department of Statistics and Actuarial - Financial Mathematics, Samos, Greece

^b Ministry of National Defense, Hellenic Navy, Athens, Greece

* Corresponding author, email: amilionis@aegean.gr

A typical way of testing the efficient market hypothesis is to compare the performance of technical trading rules with the performance of the buy and hold strategy. We investigate the predictive power of two technical trading rules, namely the simple moving average and the exponentially weighted moving average, in its own right, as well as in conjunction with the weak-form market efficiency hypothesis. Utilizing Athens Stock Exchange data, we simulate scenarios for both perfect market and real market conditions, considering the precise actual transaction costs levy on investors. Empirical findings show that, without transaction costs the performance of the exponentially weighted moving average rule is remarkably higher than that of the moving average rule, while both perform better than the passive investment strategy. Therefore, the hypothesis of weak-form efficiency is rejected. However, in real economy scenarios it is found that the predictive power of the trading rules is reversed, while for the majority of the investors the hypothesis of weak-form market efficiency on the Athens Stock Exchange is not rejected. Furthermore, we argue that the performance of a trading rule should be assessed in conjunction with the number of the triggered transactions, as the two variables are strongly correlated especially in real market conditions.

References

Brock, W., Lakonishok, J. and LeBaron, B. (1992). Simple technical trading rules and the stochastic properties of stock returns. *Journal of Finance*, **47**, 1731-1764.

- Fama, E. (1991). Efficient capital markets II. *Journal of finance*, **46**, 1557-1617.
- Milionis, A.E. and Papanagiotou, E. (2011). A test of significance of the predictive power of the moving average trading rule of technical analysis based on sensitivity analysis. *Applied Financial Economics*, **21:6**, 421-436.

Enhancing Statistical Analysis of Asymmetric Geoscience Data using Gaussian Anamorphosis

Emmanouil A. Varouchakis

Scholl of Mineral Resources Engineering, Technical University of
Crete

evarouchakis@tuc.gr

Spatial data usually follow a skewed PDF that deviates significantly from the normal distribution. In the context of big data and geosciences, where datasets are often large and complex, using a Gaussian anamorphosis model can help in normalizing the data and making it more suitable for statistical analysis methods that assume normality. Nonlinear, monotonic transformations such as the logarithm, Box-Cox, normal-score, modified Box-Cox versions, and deformed logarithmic transforms are often used to restore normality.

In this research, the recently proposed kernel-based KCDE (Pavlidis et al. 2023) method coupled with the normal scores transform is used to estimate the non-Gaussian distribution of natural resources data and to conduct Gaussian anamorphosis. KCDE derives from the data, using kernel smoothing to avoid the discontinuities of the staircase CDF function. Normalized estimates are then calculated based on the normal scores transform. These conform with the standard normal distribution. Spatial interpolation of the normal scores leads to predictions. The inverse transform in Gaussian anamorphosis is implemented via a lookup table that maps the predictions of the normalized values back to the initial values range.

The predictive performance of two stochastic interpolation approaches is compared, UK and the more recently developed SLI method. SLI is a spatial interpolation method designed to provide computational efficiency for large spatial and spatiotemporal datasets (Hristopulos et al. [2021](#)).

The research project is implemented in the framework of H.F.R.I call “Basic research Financing (Horizontal support of all Sciences)” under the National Recovery and Resilience Plan “Greece 2.0” funded by the European Union – NextGenerationEU (H.F.R.I. Project Number: 16537).

References

- Hristopulos, D., Pavlides, A., Agou, V., and Gkafa, P. (2021) Stochastic local interaction model: an alternative to kriging for massive datasets. *Mathematical Geosciences*, **53**, 1907–1949
- Pavlides, A., Varouchakis, E.A. and Hristopulos, D.T. (2023) Geostatistical analysis of groundwater levels in a mining area with three active mines. *Hydrogeol Journal* **31**, 1425–1441

Forecasting the power generation mix in Italy based on Grey Markov Models

Guglielmo D'Amico¹, Alex Karagrigoriou² and Veronica Vigna³

¹Department of Economics, "G. d'Annunzio" University, Chieti - Pescara, Italy;

²Department of Statistics and Insurance Science, University of Piraeus, Greece;

³Department of Neuroscience, Imaging and Clinical Sciences, "G. d'Annunzio" University, Chieti - Pescara, Italy;

veronica.vigna@unich.it

This study considers an application of the first-order Grey Markov Model to foresee the values of Italian power generation in relation to the available energy sources. The model is used to fit data from the Italian energy system from 2000 to 2022. The integration of Markovian error introduces a random element to the model, which is able now to capture inherent uncertainties and misalignments between the Grey model predictions and the real data. This application provides valuable insights for strategic planning in the energy sector and future developments. The results show good accuracy of the predictions, which provides powerful information for the effective implementation of energy policies concerning the evolution of energy demand in the country. Despite advancements, Italy's 2032 energy mix will still significantly rely on fossil fuels, emphasizing the need for sustained efforts beyond 2032 to enhance sustainability.

Keywords: Markov Chain, Grey Model, Energy Mix, Electricity

Modern Statistical Methods for Improving the Yield of Advanced Node Integrated Circuits

Wei W. Xing

School of Mathematics and Statistics, the University of Sheffield,
U.K.

w.xing@sheffield.ac.uk

As the technology of integrated circuits develops, microelectronic devices shrink their scale to submicrometer, which makes random process variations, eg, intra-die mismatches, doping fluctuation, and threshold voltage variation, crucial factors to be considered in an integrated circuit design. The situation gets worse in modern circuit designs, where some cells can be replicated millions of times in a circuit, eg in an SRAM cell array. A cornerstone to resolving the increasing concern of yield is the development of efficient yield estimation methods, which provide an accurate and fast failure rate estimation for a given circuit design under specific process variations. Monte Carlo (MC) is the golden standard baseline, and it is commonly utilized to estimate the yield across industry and academia. In a nutshell, MC runs SPICEs (Simulation Program with Integrated Circuit Emphasis) with parameters sampled from the process variation distribution millions of times and counts the number of failures to deliver accurate estimation. Obviously, MC is computationally expensive and easily becomes infeasible for problems for low-yield problems, which is rather common in modern circuit designs, eg the yield of a 65nm SRAM cell array can be 0.00001. In this talk, we will demonstrate how modern statistics, particularly deep-learning-based methods, can be implemented to deliver excellent performance with a speedup of more than 1000x while the accuracy remains similar.

