

Θέμα εργασίας

Οδηγίες

Η εργασία θα αξιολογηθεί με άριστα τη **μία μονάδα**. Θα πρέπει να τηρηθούν οι παρακάτω οδηγίες αλλιώς θα υπάρχει απώλεια βαθμού για κάθε παράλειψη τήρησης οδηγίας.

- Θα πρέπει να δηλωθούν δύο άτομα ως μέλη της κάθε ομάδας και θα υποβληθεί μια κοινή εργασία. Προσοχή, αν φοιτητής/τρια δηλώσει μια ομάδα και δεν έχει συνεννοηθεί με άλλον/ήν φοιτητή/τρια για να δηλώσουν ταυτόχρονα την ίδια ομάδα, μπορεί κάποιος άλλος/ή φοιτητής/τρια να δηλώσει την ίδια ομάδα. Θα πρέπει να υπάρχουν δύο φοιτητές/τριες σε κάθε ομάδα. Αν παρατηρηθεί φοιτητές/τριες να δηλώνονται σε ομάδες του ενός ατόμου, θα καταχωρούνται από το διδάσκοντα ανά δύο στην ίδια ομάδα (μία από τις δύο που δηλώθηκαν).
- **Η εργασία θα πρέπει να είναι αυθεντική. Μπορεί να υπάρχει συνεργασία μεταξύ των ομάδων πάνω στα ερωτήματα της εργασίας αλλά η αντιγραφή απαγορεύεται.**
- Η εργασία θα πρέπει να είναι γραμμένη στον υπολογιστή σε πρόγραμμα επεξεργασίας κειμένου, όπως Word. Αν είναι γραμμένη σε άλλο πρόγραμμα επεξεργασίας κειμένου, θα πρέπει να μετατραπεί σε pdf.
- Στην αρχή (πρώτη σελίδα) θα πρέπει να αναφέρονται τα ονόματα και τα ΑΕΜ των φοιτητών/τριών της ομάδας.
- Οι απαντήσεις και λύσεις των ζητημάτων θα πρέπει να δίνονται με τη σειρά που ζητούνται. Τα γραφήματα και οι πίνακες αποτελεσμάτων από το SPSS θα πρέπει να παρουσιάζονται με τη σειρά που ζητούνται και σε κατάλληλα σημεία μέσα στο κείμενο της εργασίας.
- **Το κάθε γράφημα από το SPSS θα πρέπει να έχει στον τίτλο τα ΑΕΜ των φοιτητών/τριών της ομάδας και ο κάθε πίνακας από το SPSS θα πρέπει να έχει στην επικεφαλίδα τα ΑΕΜ των φοιτητών/τριών της ομάδας. Θα συμπεριλάβετε ΜΟΝΟ σχετικούς πίνακες και σχήματα στο SPSS που υποστηρίζουν τις απαντήσεις σας (διαφορετικά θα μετρήσει αρνητικά στο βαθμό σας).**
- Η εργασία θα πρέπει να υποβληθεί ηλεκτρονικά μέσω της ιστοσελίδας του μαθήματος στο elearning το αργότερο ως τις **10/7/2023**. Θα πρέπει να υποβληθεί ηλεκτρονικά μόνο ένα αρχείο που περιέχει την αναφορά της εργασίας σε μορφή κειμένου. Το όνομα του αρχείου θα πρέπει να είναι **AEM<AEM>AEM<AEM>.docx** όπου **<AEM>** είναι ο Αριθμός Ειδικού Μητρώου (χωρίς τις ανισότητες) του/της κάθε φοιτητή/τριας της ομάδας (αν υπάρχει μόνο ένα ΑΕΜ σημαίνει πως το δεύτερο άτομο δεν συνεργάστηκε και ασχολήθηκε με την εργασία) και **docx** δηλώνει ότι είναι γραμμένο σε Word (άλλες επιτρεπτές μορφές είναι **doc** και **pdf**). Για παράδειγμα, αν η ομάδα δύο ατόμων έχει ΑΕΜ 10820 και 10200, το αρχείο αναφοράς θα πρέπει να έχει όνομα **AEM10820AEM10200.docx**, **AEM10820AEM10200.doc** ή **AEM10820AEM10200.pdf**.

Παρακαλώ να εξετάσετε αν το όνομα του αρχείου είναι σωστό πριν τα υποβάλλετε. Αρχείο με όνομα που δεν είναι στην προβλεπόμενη μορφή μπορεί να αγνοηθεί και η εργασία να μην αξιολογηθεί!

Περιγραφή προβλήματος και δεδομένα

Η εργασία αναφέρεται στο παιχνίδι Sudoku, ένα παιχνίδι – γρίφος τοποθέτησης αριθμών 1 – 9 σε 9 τετράγωνα μπλοκ που σχηματίζουν ένα μεγάλο τετράγωνο, ώστε να είναι διαφορετικά τα νούμερα σε κάθε μπλοκ καθώς και σε κάθε γραμμή και στήλη του μεγάλου τετραγώνου. Συνήθως υπάρχουν διαφορετικά επίπεδα δυσκολίας που αντιστοιχούν σε διαφορετικό αριθμό και θέση νούμερων που εμφανίζονται στο Sudoku. Ένας δείκτης της ικανότητας του παίκτη στο γρίφο Sudoku είναι ο χρόνος που χρειάζεται ο παίκτης να βρει τη λύση, δηλαδή να συμπληρώσει σωστά όλους τους αριθμούς που λείπουν για συγκεκριμένο επίπεδο δυσκολίας. Θέλουμε να συγκρίνουμε τον κάθε παίκτη (το/την φοιτητή/τρια που κάνει την εργασία) με έναν παίκτη Α για τον οποίο έχουμε δεδομένα του χρόνου επίλυσης του παιχνιδιού – γρίφου σε διαφορετικά παιχνίδια Sudoku (στο ίδιο επίπεδο δυσκολίας).

Τα δεδομένα για τον παίκτη Α δίνονται σε αρχείο τύπου SPSS που βρίσκεται στην ιστοσελίδα του μαθήματος. Το αρχείο δεδομένων λέγεται **Sudoku.sav** και περιέχει δύο στήλες. Η πρώτη στήλη (time) έχει το χρόνο επίλυσης σε κάθε ένα από 50 διαφορετικά παιχνίδια Sudoku που έπαιξε στην ιστοσελίδα <http://www.sudoku-online.com> (για το υψηλότερο επίπεδο δυσκολίας “Very hard”). Ο χρόνος είναι σε λεπτά και εκατοστά του λεπτού μετά την υποδιαστολή (όχι δευτερόλεπτα). Η δεύτερη στήλη (day) έχει έναν αριθμό για κάθε παιχνίδι, που δηλώνει τη μέρα που έπαιξε το κάθε παιχνίδι με αναφορά στην πρώτη μέρα που ξεκίνησε (στο επίπεδο δυσκολίας “Very hard”). Ο αριθμός αυτός είναι ένας δείκτης εμπειρίας του παίκτη ως προς την επίλυση του γρίφου Sudoku στο επίπεδο “Very hard”.

Δημιουργήστε ένα δείγμα χρόνων επίλυσης του παιχνιδιού – γρίφου Sudoku στο επίπεδο δυσκολίας “Very hard” που αντιπροσωπεύει την απόδοση σας (υποκειμενικά, φανταστικά, πραγματικά?) για 30 παιχνίδια που παίξατε (υποθετικά), από την πρώτη μέρα που δοκιμάσατε το παιχνίδι – γρίφος Sudoku στο επίπεδο δυσκολίας “Very hard” ως σήμερα. Θα πρέπει να αντιστοιχήσετε τον κάθε χρόνο σε κάποιον αριθμό της ημέρας (από την πρώτη μέρα). Στην εργασία θα συγκρίνετε τη δική σας απόδοση με αυτήν του παίκτη Α.

Ζητήματα της εργασίας (οι απαντήσεις θα δίνονται με χρήση του SPSS όπου αυτό είναι δυνατό):

Θα γίνουν δύο μελέτες με συγκεκριμένα ερωτήματα που δίνονται παρακάτω.

Μελέτη Α

1. Για το δείγμα των χρόνων επίλυσης του γρίφου του δικού σας και του παίχτη Α υπολογίστε τα μέτρα κεντρικής τάσης (μέση τιμή και διάμεσο) και μεταβλητότητας (διασπορά, τυπική απόκλιση, εύρος δεδομένων, πρώτο και τρίτο τεταρτημόριο) και σχηματίστε τα δύο θηκογράμματα σε ένα σχήμα και το ιστόγραμμα για κάθε ένα από τα δύο δείγματα (στο SPSS). Με βάση το αντίστοιχο θηκογράμματα και ιστόγραμμα, καθώς και τα αντίστοιχα συνοπτικά μέτρα, σχολιάστε την κατανομή του χρόνου επίλυσης του γρίφου για εσάς και για τον παίχτη Α. Φαίνεται το δείγμα σας και αυτό του παίχτη Α να συμφωνούν στην ίδια κατανομή; Απαντήστε στο παραπάνω ερώτημα με κριτήριο το κέντρο και τη μεταβλητότητα της κάθε κατανομής.
2. Με βάση το κάθε ένα από τα δύο δείγματα, υπολογίστε το 95% διάστημα εμπιστοσύνης για την τυπική απόκλιση του χρόνου επίλυσης του γρίφου για εσάς και για τον παίχτη Α. Φαίνεται να διαφέρει η μεταβλητότητα στο χρόνο επίλυσης του γρίφου σε εσάς και τον παίχτη Α? [Εδώ δε σας βοηθά το SPSS!].
3. Εκτιμήστε το μέσο χρόνο επίλυσης του γρίφου για εσάς και για τον παίχτη Α με βάση το κάθε ένα από τα δύο δείγματα. Γι αυτό θα χρησιμοποιήσετε διαστήματα εμπιστοσύνης σε επίπεδο 90% και 95% με βάση τα δύο δείγματα, τα οποία θα υπολογίσετε στο SPSS. Στην απάντησή σας συμπεριλάβετε **έναν** πίνακα αποτελεσμάτων στο SPSS και για τα δύο δείγματα για κάθε επίπεδο εμπιστοσύνης. Σχολιάστε για την ακρίβεια της εκτίμησης του μέσου χρόνου επίλυσης για εσάς και τον παίχτη Α και αν φαίνεται να υπάρχει διαφορά στο μέσο χρόνο που επιλύετε εσείς και ο παίχτης Α το γρίφο. Επίσης αν θέσουμε το κατώφλι του μέσου χρόνου επίλυσης για να θεωρείται καλός ένας παίχτης του γρίφου Sudoku στο επίπεδο "Very hard" στα 40 λεπτά, μπορείτε εσείς να θεωρηθείτε «καλός» παίχτης? Ο παίχτης Α?
4. Συγκρίνετε την απόδοσή σας με αυτήν του παίχτη Α. Υποθέστε ότι η μεταβλητότητα (διασπορά) του χρόνου επίλυσης του γρίφου είναι ίδια για εσάς και τον παίχτη Α (ανεξάρτητα από τα σχόλια σας στο 1). Χρησιμοποιώντας το SPSS εκτιμήστε αν υπάρχει διαφορά και πόση στο μέσο χρόνο επίλυσης του γρίφου μεταξύ εσάς και του παίχτη Α. Για να δώσετε την απάντησή σας χρησιμοποιήστε διάστημα εμπιστοσύνης για τη διαφορά μέσων τιμών σε επίπεδο εμπιστοσύνης 90% και 95% και συμπεριλάβετε τους αντίστοιχους πίνακες από το SPSS. **Σχολιάστε** αν υπάρχει σημαντική διαφορά στη δική σας απόδοση και του παίχτη Α.

Μελέτη Β

5. Για το δικό σας δείγμα καθώς και για το δείγμα του παίχτη Α, σχεδιάστε το κατάλληλο διάγραμμα διασποράς του χρόνου επίλυσης του παιχνιδιού - γρίφου ως προς την αντίστοιχη μέρα που παίχτηκε και υπολογίστε τον αντίστοιχο συντελεστή συσχέτισης χρησιμοποιώντας το SPSS. **Σχολιάστε** την εξάρτηση του χρόνου επίλυσης του παιχνιδιού από την μέρα που παίχτηκε με βάση τα αποτελέσματα από το SPSS για το κάθε δείγμα (η μέρα είναι ένας δείκτης εμπειρίας του παίχτη με το παιχνίδι – γρίφο).
6. Για κάθε ένα από τα δύο δείγματα και χρησιμοποιώντας το SPSS εκτιμήστε το μοντέλο γραμμικής παλινδρόμησης με τη μέθοδο ελαχίστων τετραγώνων για τη γραμμική εξάρτηση του χρόνου επίλυσης από τη μέρα που παίχτηκε το παιχνίδι – γρίφος. **Σχολιάστε** το κάθε ένα από τα δύο μοντέλα που εκτιμήσατε με έμφαση στην καταλληλότητα του μοντέλου για προβλέψεις. Στην απάντησή σας θα συμπεριλάβετε τους πίνακες του SPSS που δίνουν την εκτίμηση των παραμέτρων του μοντέλου (σταθερός όρος, κλίση και τυπική απόκλιση σφαλμάτων παλινδρόμησης). Φαίνεται να διαφέρει η κλίση της εκτιμώμενης ευθείας για εσάς και για τον παίχτη Α? Σχολιάστε τι μπορεί να σημαίνει αυτή η διαφορά ή η ταύτιση.
7. Προβλέψτε με το κάθε ένα από τα δύο μοντέλα που εκτιμήσατε στο ερώτημα 6 το μέσο χρόνο επίλυσης του γρίφου στο πρώτο τέταρτο του χρόνου που παίχτηκε το παιχνίδι, π.χ. για τον παίχτη Α που έπαιζε 150 μέρες το παιχνίδι το πρώτο τέταρτο αντιστοιχεί στη μέρα 37 (ή 38 ή 37.5). Κάνετε το ίδιο για το τρίτο τέταρτο του συνολικού χρόνου εμπειρίας στο παιχνίδι.
8. Ελέγξτε αν έχει αλλάξει η τάση μείωσης του χρόνου επίλυσης με τη μέρα (δείκτη εμπειρίας) για τον παίχτη Α, από το πρώτο μέρος για ημέρες < 75 στο δεύτερο μέρος για ημέρες ≥ 75. Για αυτό θα χωρίσετε το αρχικό δείγμα στα αντίστοιχα δύο δείγματα και θα εκτιμήσετε σε κάθε ένα από τα δύο δείγματα την ευθεία ελαχίστων τετραγώνων. Φαίνεται η κλίση της ευθείας να αλλάζει από το πρώτο στο δεύτερο μέρος?
9. Επαναλάβετε το ερώτημα 8 για το δικό σας δείγμα και σχολιάστε αντίστοιχα.
10. Φαίνεται να διαφέρει ο ρυθμός εκμάθησης του παιχνιδιού σε εσάς και στον παίχτη Α?
11. Σχολιάστε για τυχόν προβλήματα που εντοπίσατε ως προς τις συνθήκες που προϋποθέτουν οι μέθοδοι στατιστικής ανάλυσης που χρησιμοποιήσατε στα ερωτήματα 1-10.