**Μη παραμετρική στατιστική, Εργασία 1**

Α. Για τα δεδομένα σας στο αρχείο Assign\_1\_data\_???.txt

1. Παρουσιάστε διάγραμμα της εκτιμώμενης αθροιστικής συνάρτησης κατανομής με 95% διαστήματα εμπιστοσύνης για την αθροιστικής συνάρτησης κατανομής F.
2. Ζωγραφίστε και τη πραγματική F, πού είναι αυτή μιας μίξης κατανομών: κατά 85% μια χ2 με 2 βαθμούς ελευθερίας, κατά 10% μια  και κατά 5% μια . Ανήκει η F στο παραπάνω διάστημα?
3. Ελέγξτε σε επίπεδο 5% με την ks.test αν τα δεδομένα προέρχονται από μια χ2 με 2 βαθμούς ελευθερίας. Σε ποιο x έχουμε τη μέγιστη διαφορά μεταξύ ecdf και υποτιθέμενης cdf?
4. Ελέγξτε σε επίπεδο 5% με την ks.test αν τα δεδομένα προέρχονται από την πραγματική F (βλέπε σημείο 2). Σε ποιο x έχουμε τη μέγιστη διαφορά μεταξύ ecdf και υποτιθέμενης cdf?
5. Εκτιμείστε την  και δώστε διαστήματα εμπιστοσύνης για αυτή τη ποσότητα, μια φορά με την ασυμπτωτική θεωρία για τη διωνυμική κατανομή και μια φορά με την ανισότητα DKW.

Β. Προσομοιώστε 1000 δείγματα μεγέθους n=100 από μια και ελέγξτε για καθένα ποιό ποσοστό δειγμάτων απορρίπτεται. Χρησιμοποιείστε set.seed(αριθμό των data σας)

1. όταν κάνω σε επίπεδο 95% τον έλεγχο Kolmogorov-Smirnov για τη μηδενική αν το δείγμα προέρχεται από την  και
2. όταν κάνω σε επίπεδο 95% τον έλεγχο Kolmogorov-Smirnov για τη μηδενική αν το δείγμα προέρχεται από την , όπου  o δειγματικός μέσος και η δειγματική διακύμανση του εκάστοτε δείγματος.
3. Συζητείστε τα ποσοστά που βρήκατε: είναι κοντά στα αναμενόμενα?

**Παράδοση:** Σε χαρτί (όχι ηλεκτρονικά) 3 σελίδες το πολύ με όνομα και αριθμό των data Που έχετε, **μέχρι 14.3 το πρωί στο μάθημα** (όχι το απόγευμα στο εργαστήριο) το αργότερο (ή νωρίτερα στο κουτί μου…). Για το μέρος Β, συμπεριλαμβάνετε το κώδικα (εκτός από τα αποτελέσματα).