

Διάλεξη 3 & 4, Εβδομάδα 2

Διανύσματα, δημιουργία και subsetting
απόκτηση δείγματος από τις συναρτήσεις runif και rbinom
for loops

1. Φτιάξτε ένα τυχαίο δείγμα 1000 τιμών από ακέραιους αριθμούς μεταξύ 1 και 10 με επανάθεση. Σώστε το δείγμα στην μεταβλητή x.
 - a. βρείτε την μέση τιμή του (mean) του x **(0.5)**
 - b. βρείτε τη διάμεσο τιμή (median) του x. **(0.5)**
 - c. (ψάξτε στο google/wikipedia) τι αντιπροσωπεύει η μέση τιμή και τι η διάμεσος τιμή; **(1)**
 - d. Χρησιμοποιήστε την συνάρτηση table για να καταμετρήσετε το πλήθος των 1, 2, ... 10 στο x. Δηλαδή table(x). Πόσες φορές βλέπετε το κάθε νούμερο και αν το αποτέλεσμα σας φαίνεται λογικό (δηλαδή είναι κάτι που θα περιμένατε). **(1)**
 - e. Επαναληπτικά σε μια for loop με n=20 επαναλήψεις δημιουργήστε ένα vector που αποτελείται από αθροίσμα κατά θέση, τυχαίων δειγμάτων από 1000 τιμές μεταξύ 1 και 10 (όπως το x). Δηλαδή κάπως έτσι:

**a = rep(0, 1000) # ένα vector 1000 θέσεων που έχει όλα 0, Αρχικοποίηση
n = 10**

Για i από 1 ως n

x = δείγμα με την sample (όπως στην περιγραφή του ερωτήματος 1)

a = a + x

Από που ως που κυμαίνονται οι τιμές του a;

Πόσες τιμές υπάρχουν μεταξύ 0-10, μεταξύ 11-20, 21-30, ..., 91-100

Εξηγήστε την παρατήρηση σας. Είναι κάτι που το περιμένατε;

(hint? γιατί υπάρχουν πιο πολλές τιμές μεταξύ 50 - 60, από ότι μεταξύ 0 -10;)

(2)

2. Αν η πιθανότητα μετάλλαξης ανα βάση είναι 1×10^{-3} και έχετε ένα τμήμα ενός DNA μήκους 10000 βάσεων, χρησιμοποιήστε μια συνάρτηση για να πάρετε τον αριθμό μεταλλάξεων που θα μπορούσε να είχε γίνει στο τμήμα αυτό του DNA.
(hint, στην ουσία δηλαδή θέλετε ένα δείγμα μεγέθους 1 ($n = 1$) από μια κατάλληλη κατανομή, με τις κατάλληλες τιμές παραμέτρων). **(2)**
3. Πάρτε από την κατανομή που χρησιμοποιήσατε στην 2 ένα δείγμα μεγέθους 100 ($n = 100$). Αυτό για παράδειγμα αντιπροσωπεύει τον αριθμό μεταλλάξεων σε 100 τμήματα DNA που το καθένα έχει μήκος 10000 βάσεις και η πιθανότητα μετάλλαξης είναι 1×10^{-3} ανά βάση.
 - i. Πόσα τμήματα DNA δεν έχουν καμία μετάλλαξη **(0.5)**
 - ii. Πόσα τμήματα DNA έχουν 5 ή πάνω από 5 μεταλλάξεις **(0.5)**
 - iii. Ποιος είναι ο μέσος όρος μεταλλάξεων για τα 100 αυτά τμήματα DNA; **(0.5)**

- iv. Ποιος αριθμός μεταλλάξεων έγινε περισσότερες φορές; (hint: table). **(0.5)**
 - v. Ποιος είναι ο μέγιστος αριθμός μεταλλάξεων που παρατηρήθηκε σε κάποιο τμήμα DNA? **(0.5)**
4. Κατασκευάστε μια συνάρτηση που θα υπολογίζει την μέγιστη τιμή από ένα δείγμα μεγέθους n , από μια ομοιόμορφη κατανομή με παραμέτρους a, b . Οι τιμές n, a, b , θα πρέπει να είναι ορίσματα της συνάρτησης.
- a. Εφαρμόστε σε μια for loop για να αποκτήσετε ένα δείγμα 100 τέτοιων μέγιστων τιμών (HINT for (i in 1:100)). Τις 100 αυτές τιμές αποθηκεύστε τις σε μια μεταβλητή με ονομα `maxu`. **(2.5)**

Σύνολο (12/10)