

ΠΑΝΕΠΙΣΤΗΜΙΟ ΚΡΗΤΗΣ  
Τμήμα Επιστήμης Υπολογιστών

HY-317: Εφαρμοσμένες Στοχαστικές Διαδικασίες-Εαρινό Εξάμηνο 2022  
Διδάσκων: Π. Τσακαλίδης

Δεύτερη Σειρά Ασκήσεων

Ημερομηνία Ανάθεσης: 08/03/2022

Ημερομηνία Παράδοσης: 18/03/2022

**ΣΗΜΑΝΤΙΚΕΣ ΟΔΗΓΙΕΣ**

- Η παράδοση των ασκήσεών σας θα γίνεται ηλεκτρονικά, **ΑΠΟΚΛΕΙΣΤΙΚΑ** μέσω της σελίδας e-learn του μαθήματος, εδώ. Παραδόσεις με e-mail, ή οποιονδήποτε άλλον τρόπο δε θα γίνονται δεκτές και θα **μηδενίζονται**.
- Για την παράδοση, θα πρέπει να σκανάρετε ή να φωτογραφίσετε τις κόλλες των ασκήσεών σας και στη συνέχεια να φτιάξετε ένα PDF που θα περιλαμβάνει αυτές τις φωτογραφίες, δηλαδή το σύνολο της λυμένης σειράς ασκήσεων. Μπορείτε, αν προτιμάτε, να χρησιμοποιήσετε προγράμματα τύπου Latex ή Word. Για τις προγραμματιστικές ασκήσεις θα πρέπει να παραδώσετε ό,τι ζητείται στην άσκηση.
- Αριθμήστε τις σελίδες των ασκήσεών σας (από την άσκηση 1, στην άσκηση 2, στην άσκηση 3 κ.ο.κ.) και **ΣΙΓΟΥΡΕΥΤΕΙΤΕ** ότι τα σκαναρίσματα που έχετε κάνει έχουν δημιουργήσει εικόνες στις οποίες είναι **ΕΥΔΙΑΚΡΙΤΕΣ** και **ΕΥΚΟΛΟΔΙΑΒΑΣΤΕΣ** οι λύσεις σας. Θολές, και γενικότερα θορυβώδεις, εικόνες ενδέχεται να καταστήσουν αδύνατη τη διαδικασία της διόρθωσης και κατ' επέκταση να οδηγήσουν σε μηδενισμό των ασκήσεών σας.

**Θέμα: Οριακά Θεωρήματα.**

**Άσκηση 1.**

Έστω  $n$  ανεξάρτητες ομοιόμορφα κατανομημένες τ.μ.  $X_1, X_2, \dots, X_n$ , με τυπική απόκλιση ίση με 1. Θεωρήστε τον δειγματικό μέσο  $M_n = \frac{X_1 + X_2 + \dots + X_n}{n}$  και υπολογίστε τα παρακάτω ερωτήματα.

- (i) Εάν η τυπική απόκλιση του  $M_n$  είναι τουλάχιστον 0.1, ποίο είναι το μέγεθος των μεταβλητών  $n$ ;
- (ii) Εάν η εκτίμηση του δειγματικού μέσου όρου είναι το πολύ 0.01 με πιθανότητα τουλάχιστον 0.80, ποίο είναι το μέγεθος των μεταβλητών  $n$ ; (**Υπόδειξη:** Χρησιμοποιήστε την ανισότητα *Chebyshev*.)

**Άσκηση 2.**

Δύο ειδών σωματίδια φτάνουν σε έναν ανιχνευτή. Αυτά που φτάνουν είναι τύπου A με πιθανότητα  $p_A$  και τύπου B με πιθανότητα  $p_B = 1 - p_A$ . Ένα σωματίδιο τύπου A καταγράφεται (δηλαδή ανιχνεύεται και προσμετράται) από τον ανιχνευτή με πιθανότητα  $\alpha$ . Τα σωματίδια τύπου B δεν καταγράφονται από τον ανιχνευτή.

- (i) Σε ένα πλήθος 50 αφίξεων σωματιδίων, ποίος είναι ο μέσος αριθμός σωματιδίων που καταγράφονται από τον ανιχνευτή;
- (ii) Υπολογίστε τη συνάρτηση πιθανότητας της τ.μ.  $K$ , του πλήθους των σωματιδίων που φτάνουν στον ανιχνευτή από την αρχή της παρατήρησης (χρόνος  $t = 0$ ) έως και την άφιξη του πρώτου σωματιδίου που καταγράφεται από τον ανιχνευτή.

- (iii) Υπολογίστε τη συνάρτηση πιθανότητας της τ.μ.  $L$ , του πλήθους των σωματιδίων που φτάνουν στον ανιχνευτή από την αρχή της παρατήρησης (χρόνος  $t = 0$ ) έως και την άφιξη του τέταρτου σωματιδίου που καταγράφεται από τον ανιχνευτή.

Για τη συνέχεια του προβλήματος, υποθέτουμε ότι η συνολική στοχαστική διαδικασία των αφίξεων και των δύο τύπων σωματιδίων,  $A$  και  $B$ , είναι Poisson με ένταση 2 σωματίδια ανά δευτερόλεπτο. Έστω  $N(t)$  ο αριθμός των σωματιδίων που φτάνουν στον ανιχνευτή στο διάστημα  $[0, t]$  sec. Με  $R(t)$  συμβολίζουμε την ένδειξη του μετρητή του ανιχνευτή η οποία αρχίζει με την τιμή 0 σε χρόνο  $t = 0$  και αυξάνεται κατά 1 κάθε φορά που καταγράφεται ένα νέο σωματίδιο.

- (iv) Υπολογίστε τη συνάρτηση πιθανότητας της τ.μ.  $M = N(80) - N(40)$ .
- (v) Καταγράφουμε την ένδειξη του μετρητή κάθε  $T$  δευτερόλεπτα, οπότε δημιουργούμε ένα αρχείο με τις τιμές  $R(iT)$  για  $i = 0, 1, 2, \dots$ . Υπολογίστε τη μέση τιμή και τη διασπορά της  $R(iT)$ . Ποια είναι η μέση τιμή και η διασπορά της διαφοράς μεταξύ δύο διαδοχικών καταχωρήσεων,  $R((k+1)T) - R(kT)$ ;
- (vi) Χρησιμοποιώντας το Κεντρικό Οριακό Θεώρημα, υπολογίστε την αθροιστική συνάρτηση κατανομής της τ.μ.  $R(iT)$  για μεγάλα  $i$ , εκφράζοντάς την συναρτήσει της αθροιστικής συνάρτησης κατανομής της τυπικής Γκαουσιανής.

### Άσκηση 3.

Θεωρήστε ότι κάθε μέρα, η πιθανότητα να χιονίσει έστω μια φορά μες τη μέρα είναι 5%, ανεξάρτητη από κάθε άλλη μέρα. Σας ενδιαφέρει η πιθανότητα τις επόμενες 50 μέρες να υπάρχουν τουλάχιστον 45 μέρες χωρίς χιόνι.

- (i) Βρείτε την πιθανότητα αυτή χρησιμοποιώντας την κανονική προσέγγιση στην διωνιμική κατανομή.
- (ii) Επαναλάβετε το προηγούμενο ερώτημα, αυτή τη φορά χρησιμοποιώντας την προσέγγιση Poisson στην διωνιμική κατανομή.

### Άσκηση 4.

Θεωρήστε ότι τα  $\{X_n : n \geq 1\}$  είναι μια σειρά από ανεξάρτητες και πανομοιότυπα κατανεμημένες τυχαίες μεταβλητές με συνάρτηση πυκνότητας πιθανότητας  $f$ . Θεωρήστε ότι μέγιστη τιμή σημειώνεται στον δείκτη  $n$  εάν  $X_{n-1} < X_n > X_{n+1}$ . (Ορίζουμε ότι, σε αυτή την περίπτωση, μέγιστη τιμή δεν γίνεται να σημειωθεί για  $n = 1$ .)

- (i) Δείξτε ότι μακροπρόθεσμα το ποσοστό των δεικτών όπου σημειώνεται μέγιστη τιμή είναι, με πιθανότητα 1, ίσο με  $1/3$ . (**Υπόδειξη:** Χρησιμοποιήστε τον ισχυρό νόμο των μεγάλων αριθμών για μερικά αθροίσματα ανεξάρτητες και πανομοιότυπα κατανεμημένες τυχαίων μεταβλητών, αλλά παρατηρήστε ότι δεν εφαρμόζεται άμεσα.)

**Άσκηση 5.**

Προγραμματιστική εφαρμογή του *Κεντρικού Οριακού Θεωρήματος*, χρησιμοποιώντας το Matlab. Ακολουθήστε τα παρακάτω βήματα:

- (i) Δημιουργείτε  $N = 10000$  ψευδοτυχαία δείγματα μιας γεωμετρικής τ.μ.  $X$  με παράμετρο  $p = 0.25$  και μίας Poisson τ.μ.  $Y$  με παράμετρο  $\lambda = 2$ . (**Υπόδειξη:** Χρησιμοποιείτε τις συναρτήσεις `geornd()`, `poissrnd()` με κατάλληλα ορίσματα για κάθε κατανομή.)
- (ii) Απεικονίστε τα ιστογράμματα των τ.μ.  $X$  και  $Y$ , για  $n.bins = 20$ . (**Υπόδειξη:** Χρησιμοποιείτε την συνάρτηση `histogram()` με κατάλληλα ορίσματα.)
- (iii) Δημιουργείτε  $N = 10000$  ψευδοτυχαία δείγματα από  $n = 40$  γεωμετρικές τ.μ.  $X_1, X_2, \dots, X_n$  με παράμετρο  $p = 0.25$  και από  $n = 40$  Poisson τ.μ.  $Y_1, Y_2, \dots, Y_n$  με παράμετρο  $\lambda = 2$ . (**Υπόδειξη:** Χρησιμοποιείτε τις συναρτήσεις `geornd()`, `poissrnd()` με κατάλληλα ορίσματα για κάθε κατανομή.)
- (iv) Θεωρείτε τους δειγματικούς μέσους όρους  $M_{Xn} = \frac{X_1+X_2+\dots+X_n}{n}$  και  $M_{Yn} = \frac{Y_1+Y_2+\dots+Y_n}{n}$  και απεικονίστε τα ιστογράμμά τους για  $n.bins = 20$ . Τι παρατηρείτε; Επιβαιβώνεται το Κεντρικό Οριακό Θεώρημα; (**Υπόδειξη:** Χρησιμοποιείτε τις συναρτήσεις `mean()` και `histogram()` με κατάλληλα ορίσματα.)
- (v) Ορίστε τις τ.μ.  $Z_{Xn} = \frac{X_1+X_2+\dots+X_n-n\mu_X}{\sigma_X\sqrt{n}} = \frac{S_{Xn}-n\mu_X}{\sigma_X\sqrt{n}}$  και  $Z_{Yn} = \frac{Y_1+Y_2+\dots+Y_n-n\mu_Y}{\sigma_Y\sqrt{n}} = \frac{S_{Yn}-n\mu_Y}{\sigma_Y\sqrt{n}}$  και δώστε τα ιστογράμμά τους για  $n.bins = 20$ . Ποιά η διαφορά από τις προηγούμενες κατανομές; (**Υπόδειξη:** Για μία γεωμετρική τ.μ. έχουμε  $\mu = \frac{1-p}{p}$  και  $\sigma^2 = \frac{1-p}{p^2}$ . Για μία Poisson τ.μ. έχουμε  $\mu = \lambda$  και  $\sigma^2 = \lambda^2$ .)
- (vi) Εκτιμήστε τις συναρτήσεις πυκνότητας πιθανότητας (σππ) των δειγματικών μέσων όρων  $Z_{Xn}$  και  $Z_{Yn}$ , μέσω ιστογραμμάτων. (**Υπόδειξη:** Χρησιμοποιείτε την συνάρτηση `hist()` με κατάλληλα ορίσματα για κάθε κατανομή, ώστε να λάβετε τόσο τον αριθμό των στοιχείων σε κάθε bin (counts) όσο και τη θέση του κάθε bin στον άξονα x'x (centers). Ακολουθώντας, εκτιμήστε την "πυκνότητα-εμβადόν" κάθε bin μέσω της σχέσης:  $p = \frac{bincounts}{binlength \cdot N}$ .)
- (vii) Εκτιμήστε τις συναρτήσεις πυκνότητας πιθανότητας (σππ) των δειγματικών μέσων όρων  $Z_{Xn}$  και  $Z_{Yn}$ , μέσω κανονικών (gaussian) κατανομών. (**Υπόδειξη:** Χρησιμοποιείτε την συνάρτηση `pdf()` με κατάλληλα ορίσματα για τις παραμέτρους των μετασχηματισμών  $Z_{Xn}$  και  $Z_{Yn}$  κάθε κατανομής.)
- (viii) Απεικονίστε σε κοινά διαγράμματα της εκτιμήτριες των συναρτήσεων πυκνότητας πιθανότητας (σππ) των μετασχηματισμών  $Z_{Xn}$  και  $Z_{Yn}$  που βρήκατε στα παραπάνω ερωτήματα. Τι παρατηρείτε; (**Υπόδειξη:** Χρησιμοποιείτε την συνάρτηση `plot()` με κατάλληλα ορίσματα.)
- (ix) "Παίξτε" με τις μεταβλητές  $N$ ,  $n.bins$  και  $n$  και καταγράψτε τις παρατηρήσεις σας στα διαγράμματα του προηγούμενου ερωτήματος (viii) και στα ιστογράμματα των τ.μ  $Z_{Xn}$  και  $Z_{Yn}$  του ερωτήματος (v). Τι αλλάζει στις κατανομές; Υπάρχουν αλλαγές στις διασπορές; Στις εκτιμήσεις της σππ με χρήση ιστογράμματος και κανονικής κατανομής, αλλάζει η απόκλιση των δυο εκτιμήσεων; Δώστε τα αντίστοιχα ιστογράμματα. (Π.χ.  $N = [100, 1000]$ ,  $n.bins = [10, 120]$ ,  $n = [20, 120]$ .)

Παραθέστε τον κώδικά σας, τα ζητούμενα γραφήματα και μία μικρή αναφορά με τις παρατηρήσεις σας για τα αποτελέσματα.