

ΠΑΝΕΠΙΣΤΗΜΙΟ ΚΡΗΤΗΣ
Τμήμα Επιστήμης Υπολογιστών

HY-317: Εφαρμοσμένες Στοχαστικές Διαδικασίες-Εαρινό Εξάμηνο 2025
Διδάσκων: Π. Τσακαλίδης

Δεύτερη Σειρά Ασκήσεων

Ημερομηνία Ανάθεσης: 05/03/2025

Ημερομηνία Παράδοσης: 19/03/2025

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.

Άσκηση 1. Ένας αστροφυσικός προσπαθεί να εκτιμήσει τη λάμψη b ενός αστεριού με ένα σφάλμα το οποίο είναι μικρότερο από 0.1 με πιθανότητα μεγαλύτερη από 95%. Το όργανο μέτρησης που διαθέτει προσθέτει θόρυβο N με μέση τιμή 0 και τυπική απόκλιση 2, ανεξάρτητα από μέτρηση σε μέτρηση. Συνεπώς, κάθε μέτρηση X_k μπορεί να μοντελοποιηθεί ως το άθροισμα της προς εκτίμηση τιμής της λάμψης b και της συνιστώσας του θορύβου N_k :

$$X_k = b + N_k$$

Ο αστρονόμος πραγματοποιεί n μετρήσεις και χρησιμοποιεί το δειγματικό μέσο όρο ως την εκτίμητρια της ποσότητας b :

$$M_n = \frac{X_1 + X_2 + \dots + X_n}{n}$$

- Χρησιμοποιείστε το Κεντρικό Οριακό Θεώρημα για να υπολογίσετε το ελάχιστο πλήθος των μετρήσεων n που θα επιτρέψουν στον αστροφυσικό να εκτιμήσει τη λάμψη του αστεριού με τα επίπεδα ακρίβειας και εμπιστοσύνης που αναφέρονται στην εκφώνηση. Ισχύει ότι $\Phi(1.96) = 0.975$.
- Χρησιμοποιείστε την ανισότητα Chebyshev για τον ίδιο σκοπό. Τι παρατηρείτε;

Άσκηση 2. Έστω $X_1, X_2, \dots, X_n$ ανεξάρτητες τ.μ. ομοιόμορφα κατανομημένες στο διάστημα $[-1, 1]$. Να αποδείξετε ότι η ακολουθία τ.μ. $Y_1, Y_2, \dots, Y_n$ συγκλίνει κατά πιθανότητα στο 0, για κάθε μία από τις παρακάτω περιπτώσεις:

(i) $Y_n = \frac{X_n}{n}$

(ii) $Y_n = X_1 \cdot X_2 \cdot \dots \cdot X_n$ (**Υπόδειξη:** Χρησιμοποιείστε κατάλληλα την ανισότητα Chebyshev)

Άσκηση 3.

- (i) Στην εκπομπή 'Με Αρετή και Τόλη' της ΕΡΤ, ένας έμπειρος πιλότος ανέφερε ότι η διάρκεια την οποία μπορεί να έχει 'κλειδωμένο' στο στόχαστρο του ένα έχθρικό αεροσκάφος είναι κατά μέσο όρο 17.5 δευτερόλεπτα, με τυπική απόκλιση ± 4 δευτερόλεπτα. Να υπολογίσετε την πιθανότητα σε 400 ανεξάρτητες καταδιώξεις, ο συνολικός χρόνος των 'κλειδωμένων αεροσκαφών' να μην ξεπερνά τις 2 ώρες.
- (ii) Παρακολουθείτε το τηλεπαιχνίδι 'Μια στις δέκα!' το οποίο αποτελείται από 50 ερωτήσεις multiple-choice, η κάθε μία με 10 δυνατές απαντήσεις. Για κάθε σωστή απάντηση, ο παίκτης παίρνει 3 βαθμούς, και για κάθε λάθος χάνει 1 βαθμό, ενώ για να κερδίσει κάποιο χρηματικό έπαθλο χρειάζεται να συγκεντρώσει μέσο όρο βαθμολογίας (> 0.5). Γνωρίζετε ότι ο φίλος σας ο Κώστας που συμμετέχει στο επεισόδιο απαντάει όλες τις ερωτήσεις στην τύχη, ανεξάρτητα τη μία από την άλλη. Να υπολογίσετε την πιθανότητα ο Κώστας να φύγει με χρηματικό έπαθλο.

Άσκηση 4.

Ένα εργοστάσιο παράγει X_n αντικείμενα την ημέρα n , όπου X_n είναι ανεξάρτητες και όμοια κατανομημένες τ.μ. με μέση τιμή ίση με 5 και διασπορά ίση με 9.

- (i) Να υπολογίσετε μία προσέγγιση του ότι ο συνολικός αριθμός αντικειμένων που παράγονται σε 100 ημέρες είναι λιγότερος από 440.
- (ii) Να υπολογίσετε το μέγεθος του δείγματος, n , έτσι ώστε να ισχύει η σχέση:

$$P(X_1 + X_2 + \dots + X_n \geq 200 + 5n) \leq 0.05$$

- (iii) Έστω N η πρώτη ημέρα κατά την οποία ο συνολικός αριθμός παραγόμενων αντικειμένων υπερβαίνει τα 1000. Να υπολογίσετε μία προσέγγιση της πιθανότητας ότι $N \geq 220$.

(Υπόδειξη: Χρησιμοποιείτε κατάλληλα το Κεντρικό Οριακό Θεώρημα)

Άσκηση 5. Η μεταβολή της στάθμης ενός ταμιευτήρα σε ένα υδρολογικό έτος ακολουθεί ομοιόμορφη κατανομή στο $[0.5, 0.8]$. Οι τιμές της μεταβολής αναφέρονται σε μέτρα.

1. Να υπολογίσετε προσεγγιστικά την πιθανότητα μετά από 25 υδρολογικά έτη, ο ταμιευτήρας να έχει επιφάνεια η οποία ανέβηκε τουλάχιστον κατά 16μ, θεωρώντας ότι αρχίσαμε να μετράμε από το φετινό υδρολογικό έτος.
2. Να βρείτε τον ελάχιστο αριθμό υδρολογικών ετών που πρέπει να περάσουν, ώστε με πιθανότητα τουλάχιστον 80% η στάθμη του ταμιευτήρα να ανέβηκε τουλάχιστον κατά 16μ (σε σχέση με φέτος).

Άσκηση 6.

Σε αυτή την άσκηση σας ζητείται να προσομοιώσετε μια εφαρμογή του Κεντρικού Οριακού Θεωρήματος, χρησιμοποιώντας την *Python*. Ακολουθήστε τα παρακάτω βήματα:

1. Δημιουργήστε $N = 1200$ ψευδοτυχαία δείγματα μιας ομοιόμορφης τυχαίας μεταβλητής X στο διάστημα $[0, 1]$ και μιας εκθετικής τυχαίας μεταβλητής Y με παράμετρο $\lambda = 1$. (Υπόδειξη: Χρησιμοποιείτε τη συνάρτηση `numpy.random.uniform()` και `numpy.random.exponential()`).
2. Απεικονίστε τα ιστογράμματα των τυχαίων μεταβλητών X και Y , με $n_{bins} = 25$. (Υπόδειξη: Χρησιμοποιείτε τη συνάρτηση `matplotlib.pyplot.hist()`).

3. Δημιουργήστε $N = 1200$ ψευδοτυχαία δείγματα από $n = 25$ ομοιόμορφες τυχαίες μεταβλητές $X_1, X_2, \dots, X_n$ στο διάστημα $[0, 1]$ και από $n = 25$ εκθετικές τυχαίες μεταβλητές $Y_1, Y_2, \dots, Y_n$ με παράμετρο $\lambda = 1$.
4. Υπολογίστε τους δειγματικούς μέσους όρους $M_{X_n} = \frac{X_1+X_2+\dots+X_n}{n}$ και $M_{Y_n} = \frac{Y_1+Y_2+\dots+Y_n}{n}$ και απεικονίστε τα ιστογράμμά τους.
5. Ορίστε τις μεταβλητές $Z_{X_n} = \frac{M_{X_n}-\mu_X}{\sigma_X/\sqrt{n}}$ και $Z_{Y_n} = \frac{M_{Y_n}-\mu_Y}{\sigma_Y/\sqrt{n}}$, όπου μ_X και σ_X είναι η μέση τιμή και η τυπική απόκλιση της ομοιόμορφης κατανομής, ενώ μ_Y και σ_Y της εκθετικής. Απεικονίστε τα ιστογράμμά τους.
6. Εκτιμήστε τις συναρτήσεις πυκνότητας πιθανότητας (*PDFs*) των M_{X_n} και M_{Y_n} μέσω ιστογραμμάτων. Υπολογίστε την πυκνότητα κάθε *bin* με τη σχέση:

$$p = \frac{\text{bin_counts}}{\text{bin_length} \cdot N}$$

7. Εκτιμήστε τις συναρτήσεις πυκνότητας πιθανότητας των M_{X_n} και M_{Y_n} μέσω κανονικών κατανομών (*Gaussian*). Χρησιμοποιήστε τη συνάρτηση `scipy.stats.norm.pdf()` για την εκτίμηση της κανονικής πυκνότητας.
8. Απεικονίστε σε κοινά διαγράμματα τις εκτιμήτριες των συναρτήσεων πυκνότητας πιθανότητας των M_{X_n} και M_{Y_n} και συγκρίνετέ τις με τις αντίστοιχες θεωρητικές *Gaussian* κατανομές.
9. Πειραματιστείτε με διαφορετικές τιμές των N , n_{bins} και n (π.χ., $N = 10000$, $n_{bins} = 100$, $n = 100$) και καταγράψτε τις παρατηρήσεις σας.

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