ΕΕ728 (22Α004) - Προχωρημένα Θέματα Θεωρίας Πληροφορίας
1η Άσκηση για υπολογιστή
AEP με χρήση υπολογιστή

Παράδοση: Έως 22/3/2011 5 μ.μ.

Διευκρίνιση: Η παράδοση της άσκησης είναι προαιρετική. Σκοπός της άσκησης είναι να λειτουργήσει συμπληρωματικά με τις διαλέξεις. Εάν παραδώσετε τη συγκεκριμένη υπολογιστική άσκηση θα σας δοθεί μία (1) μονάδα επιπλέον (εφόσον πάρετε τουλάχιστον 5 στο τελικό διαγώνισμα). Επιτρέπεται η συνεργασία σε μικρές ομάδες (έως 3 άτομα το πολύ), αλλά όχι η αντιγραφή. Αν δουλέψετε σε ομάδα θα πρέπει να αναφέρετε τα μέλη της ομάδας. Ωστόσο, ο/η κάθε φοιτητής/τρια πρέπει να παραδώσει δικό του/της κώδικα, καθώς και γραπτή αναφορά με ερμηνεία των αποτελεσμάτων. Ο κώδικας μπορεί να γραφτεί σε οποιαδήποτε γλώσσα προγραμματισμού. Ωστόσο, αν σκοπεύετε να μη χρησιμοποιήσετε Matlab ή C/C++ παρακαλώ να με ενημερώσετε. Παρακαλώ η παράδοση του κώδικα να γίνει με e-mail στο dtouba@upatras.gr. Πρέπει να παραδώσετε τον κώδικα και όχι εκτελέσιμο αρχείο (αν παραδώσετε κώδικα σε C/C++ παρακαλώ να συμπεριλάβετε και makefile ή οδηγίες για τη μεταγλώττιση). Επίσης, ο κώδικας θα πρέπει να έχει αρκετά σχόλια ώστε να μπορεί κάποιος να τον τρέξει και να αλλάξει εύκολα τις παραμέτρους. Οι αναφορές μπορούν να παραδοθούν σε χειρόγραφη μορφή ή μέσω e-mail.

Θεωρούμε *δυαδική* πηγή *χωρίς μνήμη* που παράγει ακολουθία τ.μ. $\{X\}$, όπου $X_i \sim \mathrm{Bern}(p)$, i=1,2,...,$n$.

1. Υπολογίστε το ρυθμό εντροπίας της πηγής σε bits.

2. Έστω $p = 0.1$. Σχεδιάστε σε φθίνουσα σειρά την ποσότητα $-\frac{1}{n}\log_2 p(x_1, x_2, \ldots, x_n)$ για *όλες* τις δυαδικές ακολουθίες μήκους $n = 18$ (ή μεγαλύτερο, ανάλογα με τη μνήμη που διαθέτει ο υπολογιστής σας). Στο ίδιο διάγραμμα, σχεδιάστε τα όρια $H(X) \pm \epsilon$ για $\epsilon = 0.05$. Τι παρατηρείτε; Ποιό είναι το ποσοστό των ασθενώς $\epsilon-$τυπικών ακολουθιών σε σχέση με το σύνολο όλων των ακολουθιών μήκους n; Ποιο είναι το άθροισμα των πιθανοτήτων των ασθενώς $\epsilon-$τυπικών ακολουθιών; Βρείτε το ϵ' για το οποίο $\Pr\left\{x_1^n \in A_{\epsilon'}^{(n)}\right\} > 1 - \epsilon'$ για το συγκεκριμένο n που χρησιμοποιήσατε. Συμφωνεί η τιμή με αυτή που προκύπτει από τον ασθενή νόμο των μεγάλων αριθμών; Δηλαδή, αν χρησιμοποιήσετε το "θεωρητικό" ϵ' στο πρόγραμμά σας αντί για $\epsilon = 0.05$, επαληθεύεται πειραματικά ότι $\Pr\left\{x_1^n \in A_{\epsilon'}^{(n)}\right\} > 1 - \epsilon'$;

3. Επαναλάβετε για $p = 0.45$. Παρατηρήστε τις διαφορές και σχολιάστε.

4. Δείξτε ότι, για αρκούντως μεγάλο n και $0 < \epsilon < 1$, $(1-\epsilon)2^{n[H(X)-\epsilon]} \geq 2^{n[H(X)-2\epsilon]}$ και, επομένως, ένα (λιγότερο ακριβές) κάτω φράγμα για τον αριθμό των τυπικών ακολουθιών είναι το $\left|A_\epsilon^{(n)}\right| \geq 2^{n[H(X)-2\epsilon]}$.

5. Για μεγαλύτερα n θα επικεντρωθούμε μόνο στις (ασθενώς) τυπικές ακολουθίες. Σκεφτείτε πώς μπορείτε να βρείτε το πλήθος και το άθροισμα των πιθανοτήτων των ασθενώς $\epsilon-$τυπικών ακολουθιών για μεγάλα n. Εφαρμόστε για $n = 100$ και $n = 1000$, $p = 0.1$ και $p = 0.45$, $\epsilon = 0.05$.

6. Χρησιμοποιήστε τη μέθοδο που αναπτύξατε στο προηγούμενο ερώτημα και $p = 0.1$ και $\epsilon = 0.05$. Μεταβάλετε την τιμή του μήκους, n, από 10 σε 1000 (με βήμα 10) και σχεδιάστε στην ίδια γραφική παράσταση

   (i) Το $\frac{\left|A_\epsilon^{(n)}\right|}{2^n}$, όπου $A_\epsilon^{(n)}$ η τιμή που προκύπτει με χρήση του προγράμματος, για κάθε n.

   (ii) Τα θεωρητικά άνω και κάτω φράγματα για το $\frac{\left|A_\epsilon^{(n)}\right|}{2^n}$ για αρκούντως μεγάλο n. Συμπεριλάβετε και το φράγμα του Ερωτήματος 4.

7. Σε μια άλλη γραφική παράσταση σχεδιάστε το άθροισμα των πιθανοτήτων των ασθενώς $\epsilon-$τυπικών ακολουθιών για $n = 10 : 10 : 1000$. Συγκρίνετε με το θεωρητικό κάτω φράγμα για αρκούντως μεγάλο n.

8. Επαναλάβετε τα 6 και 7 για $p = 0.45$, $\epsilon = 0.05$. Παρατηρήστε τις διαφορές και σχολιάστε.