ΕΕ728 (22Α004) - Προχωρημένα Θέματα Θεωρίας Πληροφορίας
1η σειρά ασκήσεων
AEP και ανισότητα επεξεργασίας δεδομένων

Παράδοση: Έως 21/4/2015

1. **Επεξεργασία Δεδομένων – Cover & Thomas 2.15**

   Έστω ότι οι τ.μ. $X_1 \to X_2 \to X_3 \to \ldots \to X_n$ σχηματίζουν αλυσίδα Markov, δηλαδή $p(x_1, x_2, \ldots, x_n) = p(x_1)p(x_2|x_1)\ldots p(x_n|x_{n-1})$. Βρείτε την πιο απλή έκφραση που μπορείτε για την $I(X_1; X_2, X_3, \ldots, X_n)$.

2. **Όριο γινομένου – Cover & Thomas 3.8**

   Έστω η ακολουθία $X_1, X_2, \ldots, X_n$ η οποία αποτελείται από ανεξάρτητες και ομοίως κατανεμημένες (i.i.d.) τ.μ. οι οποίες ακολουθούν την κατανομή

$$X_i = \begin{cases} 1 & \text{με πιθανότητα } \frac{1}{2} \\ 2 & \text{με πιθανότητα } \frac{1}{4} \\ 3 & \text{με πιθανότητα } \frac{1}{4} \end{cases}$$

   Να βρεθεί το $\lim_{n\to\infty}(X_1 \cdot X_2 \cdot X_3 \cdots X_{n-1} \cdot X_n)^{1/n}$.

   Υπόδειξη: Βρείτε το όριο του $\log(X_1 \cdot X_2 \cdot X_3 \cdots X_{n-1} \cdot X_n)^{1/n}$.

3. **AEP και Σχετική Εντροπία – Cover & Thomas 3.9**

   Έστω ανεξάρτητες και ομοίως κατανεμημένες (i.i.d) διακριτές τ.μ. $X_1, X_2, \ldots$ που ακολουθούν κατανομή με μάζα πιθανότητας $p(x)$ και $|\mathcal{X}| < \infty$. Επομένως, $p(x_1, x_2, \ldots, x_n) = \prod_{i=1}^n p(x_i)$. Είδαμε ότι $-\frac{1}{n}\log p(X_1, X_2, \ldots, X_n) \to H(X)$ κατά πιθανότητα. Έστω, τώρα, μια άλλη συνάρτηση μάζας πιθανότητας $q(x)$ ορισμένη στον ίδιο δειγματικό χώρο $\mathcal{X}$ με την $p(x)$, και $q(x_1, x_2, \ldots, x_n) = \prod_{i=1}^n q(x_i)$.

   (α) Βρείτε την τιμή του ορίου $-\frac{1}{n}\log q(X_1, X_2, \ldots, X_n)$ εάν οι $X_1, X_2, \ldots$ είναι ανεξάρτητες και ομοίως κατανεμημένες (i.i.d) με κατανομή $p(x)$.

   (β) Βρείτε την τιμή του ορίου $-\frac{1}{n}\log\frac{q(X_1,X_2,\ldots,X_n)}{p(X_1,X_2,\ldots,X_n)}$ (log-likelihood), όπου, και πάλι, οι $X_1, X_2, \ldots$ είναι ανεξάρτητες και ομοίως κατανεμημένες (i.i.d) με κατανομή $p(x)$. Επομένως, η πιθανότητα να θεωρήσουμε ότι οι X_i ακολουθούν κατανομή $q(x)$ (αντί για $p(n)$) ελαττώνεται εκθετικά με το n και με ρυθμό ανάλογο της $D(p||q)$.

4. **AEP (Προχωρημένα Θέματα Θ. Π. Τελική Εξέταση Ιουνίου 2008)**

   Θεωρούμε την τ.μ. X με τιμές στο σύνολο $\mathcal{X} = \{0, 1, 2\}$ και συνάρτηση μάζας πιθανότητας $p(0) = 1/2$, $p(1) = p(2) = 1/4$. Ποιες από τις παρακάτω ακολουθίες είναι τυπικές σύμφωνα με τον ορισμό της ασθενούς τυπικότητας που δώσαμε στο μάθημα; Θεωρήστε $\epsilon = 0$. Δικαιολογήστε επαρκώς την απάντησή σας.

   (α) 1 2 0 0 0 0 2 1

   (β) 0 1 2 0 1 2 0 1 2 0 1 2

   (γ) 0 1 2 0 0 2 0 0 2 0 2 2

5. **Μπρίτζ (Προχωρημένα Θέματα Θ. Π. Τελική Εξέταση Ιουνίου 2011))**

   Το μπριτζ παίζεται με μία τράπουλα 52 χαρτιών.

   Ένα *χέρι* (hand) είναι οποιοσδήποτε συνδυασμός 13 χαρτιών.

   Η *διανομή* (deal) είναι οποιαδήποτε διαμέριση της τράπουλας σε 4 χέρια.

   Ένας απλός τρόπος για να αναπαραστήσουμε ένα χέρι είναι αντιστοιχίζοντας ένα μοναδικό αριθμό 6 bits σε κάθε χαρτί. Σε αυτήν την περίπτωση χρειάζονται 78 bits για να περιγράψουν ένα συγκεκριμένο χέρι. (Αυτό που οι Μηχανικοί Επικοινωνιών ονομάζουν Pulse Coded Modulation – PCM).

   (α) Αποδείξτε ότι οποιαδήποτε *δυαδική* περιγραφή ενός αυθαίρετου χεριού δεν μπορεί να χρησιμοποιεί λιγότερα από *περίπου* $52H_b\left(\frac{1}{4}\right) \approx 42$ bits, όπου $H_b(p) = p \log \frac{1}{p} + (1-p) \log \frac{1}{1-p}$ η εντροπία δυαδικής τ.μ. $\sim \text{Bern}(p)$.
   Υπόδειξη: Βρείτε, πρώτα, μία περιγραφή χρησιμοποιώντας 52 bits.

   (β) Δείξτε ότι δεν μπορούμε να περιγράψουμε μία αυθαίρετη διανομή με λιγότερα από 104 bits. Δώστε μία αναπαράσταση με χρήση 104 bits.

   Επιπλέον μονάδες αν δώσετε και δεύτερο τρόπο αναπαράστασης με χρήση 104 bits.

   Σε περίπτωση που σας χρειαστεί δίνεται ότι $H_b\left(\frac{1}{3}\right) \approx 0.9183$ bits.

   (γ) Δείξτε ότι για να περιγραφούν δύο χέρια χρειάζονται περίπου 78 bits. Δώστε έναν τρόπο περιγραφής με χρήση 78 bits.