ΕΕ728 (22Α004) - Προχωρημένα Θέματα Θεωρίας Πληροφορίας
Ενδεικτικές λύσεης 1ης σειράς ασκήσεων

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 $I(X_1; X_2, X_3, \ldots, X_n)$.

   Απάντηση:

   Από τον κανόνα αλυσίδας για την αμοιβαία πληροφορία,

   $$I(X_1; X_2, X_3, \ldots, X_n) = I(X_1; X_2) + I(X_1; X_3|X_2) + \ldots + I(X_1; X_n|X_2, X_3, \ldots, X_{n-1}).$$

   Γνωρίζουμε ότι σε μια αλυσίδα Markov, το παρελθόν και το μέλλον είναι ανεξάρτητα *δοθέντος του παρόντος*. Για παράδειγμα, $p(x_1|x_2, x_3) = p(x_1|x_2, x_3, x_4) = p(x_1, x_4|x_2, x_3) / p(x_4|x_2, x_3) \Rightarrow p(x_1, x_4|x_2, x_3) = p(x_1|x_2, x_3)p(x_4|x_2, x_3)$. Επομένως, εκτός από τον πρώτο όρο, όλοι οι άλλοι όροι ισούνται με 0, και

   $$I(X_1; X_2, X_3, \ldots, X_n) = I(X_1; X_2).$$

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 $\log(X_1 \cdot X_2 \cdot X_3 \cdots X_{n-1} \cdot X_n)^{1/n}$.

   Απάντηση:

   Από τον ασθενή νόμο των μεγάλων αριθμών,

   $$\lim_{n\to\infty} \log(X_1 \cdot X_2 \cdot X_3 \cdots X_{n-1} \cdot X_n)^{1/n} = \lim_{n\to\infty} \frac{1}{n} \sum_i \log(X_i) \to \mathbb{E}\left[\log X\right].$$

   Επομένως, $\lim_{n\to\infty}(X_1 \cdot X_2 \cdot X_3 \cdots X_{n-1} \cdot X_n)^{1/n} = 2^{\mathbb{E}[\log X]}$. Αντικαθιστώντας, $2^{\mathbb{E}[\log X]} \approx 1.565$.

   Εναλλακτική απόδειξη (από συνάδελφό σας!):

   Από την ισχυρή τυπικότητα, στο όριο ($n \to \infty$), θα έχουμε $np_1 = \frac{n}{2}$ "1", $np_2 = \frac{n}{4}$ "2" και $np_3 = \frac{n}{4}$ "3".

   Επομένως, $\lim_{n\to\infty}(X_1 \cdot X_2 \cdot X_3 \cdots X_{n-1} \cdot X_n)^{1/n} = \lim_{n\to\infty}\left(1^{n/2} \cdot 2^{n/4} \cdot 3^{n/4}\right)^{1/n} = 6^{1/4} \approx 1.565$.

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 <u>ίδιο</u> 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 $p(x)$.

Απάντηση:

Δεδομένου ότι οι X_i είναι ανεξάρτητες, και οι $q(X_i)$ είναι ανεξάρτητες. Από τον ασθενή νόμο των μεγάλων αριθμών,

$$\begin{aligned}
\lim_{n\to\infty} -\frac{1}{n} \log q(X_1, X_2, \ldots, X_n) &= \lim_{n\to\infty} -\frac{1}{n} \log \prod_i q(X_i) \\
&= \lim_{n\to\infty} -\frac{1}{n} \sum_i \log q(X_i) \stackrel{(a)}{\to} -\mathbb{E}_p(\log q(X)) \\
&= -\sum_x p(x) \log q(x) \\
&= \sum_x p(x) \log \frac{p(x)}{q(x)} - \sum_x p(x) \log p(x) \\
&= D(\mathbf{p}||\mathbf{q}) + H(\mathbf{p}).
\end{aligned}$$

Στο (a) χρησιμοποιήθηκε το γεγονός ότι η $q(\cdot)$ είναι ντετερμινιστική συνάρτηση της τ.μ. X η οποία ακολουθεί κατανομή $p(x)$.

Παρατηρήστε ότι καταλήξαμε στο ίδιο αποτέλεσμα με την κωδικοποίηση μεταβλητού μήκους (από το μάθημα ''Θεωρία Πληροφορίας''): Εάν συμπιέσουμε με ακολουθίες σταθερού μήκους (ή σχεδόν σταθερού εάν θέλουμε να συμπεριλάβουμε και τις μη τυπικές ακολουθίες), αλλά θεωρώντας την κατανομή **q** αντί για την (πραγματική) κατανομή **p**, θα χρειαστούμε, κατά μέσο όρο, $D(\mathbf{p}||\mathbf{q})$ επιπλέον bits ανά σύμβολο. Παρατηρήστε, επίσης, ότι εάν η X ακολουθεί κατανομή **q**, αλλά συμπιεστεί θεωρώντας κατανομή **p** το ''κόστος'' ενδέχεται να είναι διαφορετικό, δεδομένου ότι, στη γενική περίπτωση, $D(\mathbf{q}||\mathbf{p}) \neq D(\mathbf{p}||\mathbf{q})$.

(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 $D(p||q)$.

$\boxed{\text{Απάντηση:}}$

Όπως και στο προηγούμενο ερώτημα,

$$
\begin{aligned}
-\lim_{n\to\infty}\frac{1}{n}\log\frac{q(X_1,X_2,\ldots,X_n)}{p(X_1,X_2,\ldots,X_n)} &= -\lim_{n\to\infty}\frac{1}{n}\log\prod_i\frac{q(X_i)}{p(X_i)} \\
&= -\lim_{n\to\infty}\frac{1}{n}\sum_i\log\frac{q(X_i)}{p(X_i)} \to -\mathbb{E}_p\left(\log\frac{q(X)}{p(X)}\right) \\
&= -\sum_x p(x)\log\frac{q(x)}{p(x)} = D(\mathbf{p}||\mathbf{q}).
\end{aligned}
$$

Συνεπώς, $\log\frac{q(X_1,X_2,\ldots,X_n)}{p(X_1,X_2,\ldots,X_n)} \to -nD(\mathbf{p}||\mathbf{q}) \Rightarrow \frac{q(X_1,X_2,\ldots,X_n)}{p(X_1,X_2,\ldots,X_n)} \to 2^{-nD(\mathbf{p}||\mathbf{q})}$, και η πιθανότητα να θεωρήσουμε (λανθασμένα) ότι η X ακολουθεί κατανομή **q** αντί για **p** ελαττώνεται εκθετικά καθώς μας αποκαλύπτονται όλο και περισσότερα δείγματα της X. Όσο πιο κοντά στην **p** βρίσκεται η **q**, τόσο πιο δύσκολο μας είναι να τις ''ξεχωρίσουμε''.

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

$\boxed{\text{Απάντηση:}}$

Από τον ορισμό της ασθενούς τυπικότητας, μια ακολουθία **x** είναι (ασθενώς) ϵ-τυπική ως προς την κατανομή $p(x)$ όταν

$$\left|-\frac{1}{n}\log p(\mathbf{x}) - H(X)\right| \le \epsilon.$$

Για $\epsilon = 0$, πρέπει να ισχύει $-\frac{1}{n}\log p(\mathbf{x}) = H(X)$.

$H(X) = -\frac{1}{2}\log\frac{1}{2} - 2\times\frac{1}{4}\log\frac{1}{4} = \frac{1}{2} + 1 = \frac{3}{2}$.

(α) 1 2 0 0 0 0 2 1 ΝΑΙ

$-\frac{1}{n}\log p(\mathbf{x}) = \frac{1}{8}(4\times\log 2 + 4\times\log 4) = \frac{3}{2}$. Η ακολουθία είναι τυπική.

(β) 0 1 2 0 1 2 0 1 2 0 1 2 ΟΧΙ

$-\frac{1}{n}\log p(\mathbf{x}) = \frac{1}{12}(4\times\log 2 + 4\times\log 4 + 4\times\log 4) = \frac{5}{3}$. Η ακολουθία δεν είναι τυπική.

(γ) 0 1 2 0 0 2 0 0 2 0 2 2 ΝΑΙ

$-\frac{1}{n}\log p(\mathbf{x}) = \frac{1}{12}(6\times\log 2+\log 4+5\times\log 4) = \frac{3}{2}$. Η ακολουθία είναι (ασθενώς) τυπική!

Παρόλο που η ακολουθία δεν περιέχει ίσο αριθμό 1 και 2 είναι (ασθενώς) τυπική δεδομένου ότι η εμπειρική της εντροπία ισούται με την εντροπία της $p(x)$. Ωστόσο, η ακολουθία δεν είναι ισχυρώς τυπική.

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. $\sim \mathrm{Bern}(p)$.
Υπόδειξη: Βρείτε, πρώτα, μία περιγραφή χρησιμοποιώντας 52 bits.

Απάντηση:

Μπορούμε να αναπαραστήσουμε κάθε χέρι ως μία δυαδική ακολουθία 52 bits με '1' όταν ένα χαρτί βρίσκεται μέσα στο χέρι και '0' αλλιώς. Η περιγραφή αυτή είναι 1-προς-1, δηλαδή μπορούμε να βρούμε το χέρι από την περιγραφή και αντιστρόφως. Επομένως, μπορούμε να φανταστούμε ένα χέρι ως την έξοδο μίας πηγής που παράγει μία από $\binom{52}{13}$ ακολουθίες μήκους 52. Η πιθανότητα ένα συγκεκριμένο χαρτί να είναι στο χέρι είναι $1/4$. Επομένως, οι τυπικές ακολουθίες είναι $\approx 2^{52H_b(1/4)}$ και χρειαζόμαστε $\approx 52H_b\left(\frac{1}{4}\right)$ bits για να τις αναπαραστήσουμε. Βέβαια, η πηγή δεν είναι χωρίς μνήμη γιατί γνωρίζουμε ότι θα παραχθούν ακριβώς 13 '1' (και ακριβώς 39 '0'), γι' αυτό και το αποτέλεσμα είναι προσεγγιστικό.

(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.

Απάντηση:

Θεωρούμε, και πάλι, ακολουθία 52 συμβόλων, αλλά αυτή τη φορά με τετραδικές τιμές. Ο αριθμός που αντιστοιχεί σε κάθε χαρτί δηλώνει το χέρι στο οποίο βρίσκεται το χαρτί. Δεδομένου ότι ένα χαρτί μπορεί να βρεθεί σε οποιοδήποτε από τα 4 χέρια με την ίδια πιθανότητα, οι τυπικές ακολουθίες είναι $\approx 2^{52\log_2 4} = 2^{104}$ και χρειαζόμαστε περίπου 104 bits για να τις αναπαραστήσουμε.

Ένας άλλος τρόπος αναπαράστασης είναι ο εξής:

• Χρησιμοποιούμε την ακολουθία 52 bits του Ερωτήματος (α) για να περιγράψουμε ποια χαρτιά ανήκουν στο 1ο χέρι $\rightarrow \approx 52H_b(1/4) \approx 42$ bits.

• Χρησιμοποιούμε ακολουθία 39 bits που αντιστοιχούν στα χαρτιά που δεν είναι στο 1ο χέρι (ο δέκτης μπορεί να βρει την αντιστοίχιση γιατί έχει, πρώτα, διαβάσει την 1η ακολουθία των 52 bits) $\rightarrow \approx 39H_b(1/3) \approx 36$ bits.

• Χρησιμοποιούμε ακολουθία 26 bits που αντιστοιχούν στα χαρτιά που δεν είναι στο 1ο ή στο 2ο χέρι $\rightarrow \approx 26H_b(1/2) \approx 26$ bits.

Σύνολο: ≈ 104 bits.

Εναλλακτικά μπορούμε να "τελειώνουμε" με κάθε χαρτί πριν προχωρήσουμε. Δηλαδή να ρωτάμε αν ένα χαρτί ανήκει στην 1η διανομή, αν όχι, αν ανήκει στη 2η κτλ. Οπότε, χρειαζόμαστε $H_b\left(\frac{1}{4}\right) + \frac{3}{4}H_b\left(\frac{1}{3}\right) + \frac{1}{2}H_b\left(\frac{1}{2}\right) = 2$ bits ανά διανομή. Στην ουσία, όμως, αυτή η προσέγγιση είναι ίδια με το 2ο τρόπο αν τη δούμε ως κωδικοποίηση σταθερού μήκους πολλών χαρτιών μαζί.

Τέλος, ένας άλλος τρόπος που πρότειναν κάποιοι είναι να χρησιμοποιήσουμε την υπό συνθήκη κατανομή του δεύτερου χεριού δεδομένου του πρώτου. Έστω X η τ.μ. Bernoulli$(1/4)$ που δηλώνει αν ένα χαρτί βρίσκεται στο πρώτο χέρι ή όχι και Y η τ.μ. Bernoulli$(1/4)$ που δηλώνει αν ένα χαρτί βρίσκεται στο δεύτερο χέρι ή όχι. Μπορούμε να διαπιστώσουμε εύκολα ότι $p(y=0|x=1)=1$, $p(y=1|x=0) = \frac{1}{3} = 1 - p(y=0|x=0)$. Συνεπώς,

$$\begin{aligned} H(Y|X) &= p(x=0)H(Y|X=0) + p(x=1)H(Y|X=1) \\ &= \frac{3}{4}H_b(1/3) + 0 = \frac{3}{4}H_b(1/3). \end{aligned}$$

Επομένως, για να περιγράψουμε την ακολουθία $Y_1, Y_2, \ldots, Y_{52}$ δεδομένων των X_i χρειαζόμαστε (κατά προσέγγιση) τουλάχιστον $52 \cdot \frac{3}{4}H_b(1/3) = 39H_b(1/3)$ bits.

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

Απάντηση:

Όπως και πριν, θεωρούμε ακολουθία 52 συμβόλων, αυτή τη φορά με 3 πιθανές τιμές (1ο χέρι, 2ο χέρι, 3ο ή 4ο χέρι) και κατανομή $\{1/4, 1/4, 1/2\}$. Επομένως, οι τυπικές ακολουθίες είναι $\approx 2^{52H(1/4,1/4,1/2)} = 2^{52\cdot 1.5} = 2^{78}$ και χρειαζόμαστε ≈ 78 bits για να τις αναπαραστήσουμε. Δηλαδή, χρησιμοποιούμε την από κοινού κατανομή $p(x,y)$ των X και Y.

Εναλλακτικά, μπορούμε να χρησιμοποιήσουμε τα δύο πρώτα βήματα του 2ου τρόπου του προηγούμενου ερωτήματος.