EE728 (22A004) - Προχωρημένα Θέματα Θεωρίας Πληροφορίας
Ενδεικτικές Λύσεις 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

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

   Εναλλακτική απόδειξη (από συνάδελφό σας το 2008):

   Δεδομένου ότι $X_1 \to X_2 \to X_3 \to \ldots \to X_n$, ισχύει, επίσης, $X_n \to X_{n-1} \to X_{n-2} \to \ldots \to X_1$ και, επομένως, $p(x_1|x_2, x_3, \ldots, x_n) = p(x_1|x_2)$. Επομένως,

   $$\begin{aligned} I(X_1; X_2, X_3, \ldots, X_n) &= H(X_1) - H(X_1|X_2, X_3, \ldots, X_n) \\ &= H(X_1) - \sum_{x_1, x_2, \ldots, x_n} p(x_1, x_2, \ldots, x_n) \log p(x_1|x_2, x_3, \ldots, x_n) \\ &= H(X_1) - \sum_{x_1, x_2, \ldots, x_n} p(x_1, x_2, \ldots, x_n) \log p(x_1|x_2) \\ &= H(X_1) - \sum_{x_1, x_2} p(x_1, x_2) \log p(x_1|x_2) \\ &= H(X_1) - H(X_1|X_2) = I(X_1; X_2). \end{aligned}$$

2. **Μια εναλλακτική απόδειξη της ανισότητας $D(p||q) \geq 0$ (Cover & Thomas 2.26)**

   (α) Δείξτε ότι, για $0 < x < \infty$, $\ln x \leq x - 1$.

   Απάντηση:

   Η $f(x) = \ln x$ είναι κοίλη $\cap$ για $x > 0$. Συνεπώς, η εφαπτομένη στο $x = 1$ βρίσκεται πάντοτε επάνω από την $f(x)$. Η εφαπτομένη δίνεται από την εξίσωση $y - f(1) = f'(1)(x-1) \Rightarrow y = \frac{1}{x}|_{x=1}(x-1) = x - 1$. Επομένως, $\ln x \leq x - 1$.

   Εναλλακτικά, το αποτέλεσμα προκύπτει με μεγιστοποίηση της $g(x) = \ln x - x + 1$ (ή με ελαχιστοποίηση της $-g(x)$) και από το γεγονός ότι η $g(x)$ είναι κοίλη $\cap$ στο $(0, +\infty)$ και, επομένως, το τοπικό μέγιστο στο $x = 1$ είναι και ολικό.

   (β) Δικαιολογήστε τα παρακάτω βήματα

$$
\begin{aligned}
-D(p||q) &= \sum_x p(x) \ln \frac{q(x)}{p(x)} \\
&\overset{(i)}{\leq} \sum_x p(x) \left(\frac{q(x)}{p(x)} - 1 \right) \\
&\overset{(ii)}{\leq} 0.
\end{aligned}
$$

   Απάντηση:

   (i) Ερώτημα (α) (ii) Η $q(x)$ ενδέχεται να έχει μη μηδενική μάζα για x όπου $p(x) = 0$, οπότε $\sum_x q(x) \leq 1$.

   (γ) Πότε ισχύει η ισότητα;

   Απάντηση:

   Για να ισχύει η (i) με ισότητα πρέπει $q(x)/p(x) = 1$ για όλα τα x όπου $p(x) > 0$. Προφανώς, στην περίπτωση αυτή ισχύει και η (ii). Συνεπώς, η ισότητα ισχύει εάν και μόνο εάν $\mathbf{p} \equiv \mathbf{q}$.

3. **Ανισότητα Fano (Cover & Thomas 2.32)**

   Έστω η εξής από κοινού κατανομή μάζας πιθανότητας για τις τ.μ. (X, Y)

		Y		
		a	b	c
X	1	$\frac{1}{6}$	$\frac{1}{12}$	$\frac{1}{12}$
	2	$\frac{1}{12}$	$\frac{1}{6}$	$\frac{1}{12}$
	3	$\frac{1}{12}$	$\frac{1}{12}$	$\frac{1}{6}$

   Έστω, επίσης, $\hat{X}(Y)$ εκτιμητής για τη X με βάση την Y και $P_e = \Pr\{\hat{X}(Y) \neq X\}$.

(α) Βρείτε το βέλτιστο εκτιμητή $\hat{X}(Y)$, καθώς και την P_e που αντιστοιχεί στο βέλτιστο εκτιμητή.

Απάντηση:

Από τον πίνακα, παρατηρούμε ότι ο βέλτιστος κανόνας απόφασης είναι ο

$$\hat{X}(y) = \begin{cases} 1 & \text{για} \quad y = a \\ 2 & \text{για} \quad y = b \\ 3 & \text{για} \quad y = c \end{cases}$$

Πιο αυστηρά, ο βέλτιστος κανόνας απόφασης (Maximum a Posteriori – MAP – detector) προκύπτει από τη μετιστοποίηση της εκ των υστέρων πιθανότητας

$$\max_x p(x|y).$$

Το βέλτιστο x για κάθε τιμή y της Y προκύπτει απευθείας από τον πίνακα.

Η ελάχιστη πιθανότητα σφάλματος για τον εκτιμητή ισούται με

$$\begin{aligned} P_e &= \Pr\left\{\hat{X}(Y) \neq X\right\} = \sum_x p(x) \Pr\left\{\hat{X}(Y) \neq X | X = x\right\} \\ &= \frac{1}{3}\left\{\Pr\left\{Y \neq a | X = 1\right\} + \Pr\left\{Y \neq b | X = 2\right\} + \Pr\left\{Y \neq c | X = 3\right\}\right\} = \frac{1}{2}. \end{aligned}$$

(β) Βρείτε το φράγμα για την P_e που προκύπτει από την ανισότητα Fano και συγκρίνετε.

Απάντηση:

Από την ανισότητα Fano,

$$P_e \geq \frac{H(X|Y) - 1}{\log |\mathcal{X}|}.$$

Η $H(X|Y)$ μπορεί να υπολογιστεί ως εξής

$$H(X|Y) = \sum_y p(y) H(X|Y = y) = \frac{1}{3} \times 3H\left(\frac{1}{2}, \frac{1}{4}, \frac{1}{4}\right) = 1.5 \text{ bits}.$$

Συνεπώς, $P_e \geq \frac{0.5}{\log_2 3} \approx 0.316$.

Παρατήρηση: Αν χρησιμοποιήσουμε ένα καλύτερο κάτω φράγμα:

$$P_e \geq \frac{H(X|Y) - 1}{\log\left(|\mathcal{X}| - 1\right)},$$

τότε $P_e \geq 0.5$.

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

   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, $2^{\mathbb{E}[\log X]} \approx 1.565$.

   <u>Εναλλακτική απόδειξη (από συναδέλφους σας):</u>

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

5. **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,

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

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

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 $\mathbf{q}$ αντί για την (πραγματική) κατανομή $\mathbf{p}$, θα χρειαστούμε, κατά μέσο όρο, $D(\mathbf{p}||\mathbf{q})$ επιπλέον bits ανά σύμβολο. Παρατηρήστε, επίσης, ότι εάν η X ακολουθεί κατανομή $\mathbf{q}$, αλλά συμπιεστεί θεωρώντας κατανομή $\mathbf{p}$ 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 $D(p||q)$.

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 $\mathbf{q}$ αντί για $\mathbf{p}$ ελαττώνεται εκθετικά καθώς μας αποκαλύπτονται όλο και περισσότερα δείγματα της X. Όσο πιο κοντά στην $\mathbf{p}$ βρίσκεται η $\mathbf{q}$, τόσο πιο δύσκολο μας είναι να τις "ξεχωρίσουμε".

6. **AEP (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) (5 μονάδες)
`1 2 0 0 0 0 2 1`

(β) (5 μονάδες)
`0 1 2 0 1 2 0 1 2 0 1 2`

(γ) (5 μονάδες)
`0 1 2 0 0 2 0 0 2 0 2 2`

$\boxed{\text{Απάντηση:}}$

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

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

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.

7. **Αποταμίευση (Τελική Εξέταση Ιουνίου 2009)**

Ένας καταθέτης ανοίγει λογαριασμό με αρχικό κεφάλαιο $X_0 = 1000$ και *μηνιαίο* επιτόκιο 1%. Θεωρούμε ότι το επιτόκιο αυτό είναι εγγυημένο για όσο παραμένει ανοιχτός ο λογαριασμός, δηλαδή δε μεταβάλλεται. Επίσης, θεωρούμε ότι ο τόκος υπολογίζεται στο τέλος κάθε μήνα. Στο τέλος κάθε μήνα ο καταθέτης έχει την επιλογή να εισπράξει τον τόκο ή να τον αφήσει στο λογαριασμό, οπότε αυτός προστίθεται στο υπάρχον κεφάλαιο. Θεωρούμε, τέλος, ότι δεν επιτρέπεται στον καταθέτη να εισπράξει ποσό διαφορετικό από τον τόκο στο τέλος κάθε μήνα (ούτε μεγαλύτερο ούτε μικρότερο). Δηλαδή ο καταθέτης πρέπει να εισπράξει είτε τον τόκο του μήνα ή τίποτα.

(α) Εάν σε σύνολο N μηνών ο καταθέτης έχει εισπράξει τον τόκο K φορές, δώστε μια έκφραση για το κεφάλαιο, X_N, στο τέλος του $N-$οστού μήνα. Θεωρούμε ότι η X_N ισούται με το κεφάλαιο που απομένει *μετά* από την είσπραξη του τόκου, εφόσον αυτή γίνει. Εάν ο καταθέτης δεν εισπράξει ποτέ τους τόκους, μετά από πόσους μήνες θα έχει διπλασιάσει το αρχικό κεφάλαιο; Δίνεται ότι $1/\log_2(1.01) \approx 69.66$.

$\boxed{\text{Απάντηση:}}$

$$X_N = 1000 \cdot (1.01)^{N-K}.$$

Εάν οι τοκοι δεν εισπραχθούν ποτέ,

$$X_N = 1000 \cdot (1.01)^N = 2000.$$

Λύνοντας ως προς N,

$$N = \lceil \log 2/\log(1.01) \rceil = 70.$$

(β) Θεωρούμε, τώρα, ότι ο καταθέτης ενδέχεται να έχει ανάγκη από τους τόκους, με αποτέλεσμα να τους εισπράττει στο τέλος κάθε μήνα με πιθανότητα $1/4$. Η απόφαση αν θα εισπράξει τους τόκους το μήνα i είναι ανεξάρτητη από την απόφασή του το μήνα $j \neq i$.

Με τι ισούται η από κοινού εντροπία $H(X_0, X_1, \ldots, X_N)$;

Με τι ισούται ο ρυθμός εντροπίας, $H(\mathcal{X})$;

Με τι ισούται η από κοινού εντροπία $H(X_0, X_1, \ldots, X_{j-1}, X_{j+1}, \ldots, X_N)$, για κάποιο $0 < j < N$;

Δίνεται $\log_2 3 \approx 1.585$.

$\boxed{\text{Απάντηση:}}$

Από τον κανόνα αλυσίδας για την εντροπία,

$$\begin{aligned}
H(X_0, X_1, \ldots, X_N) &= H(X_0) + H(X_1|X_0) + \cdots + H(X_N|X_0, X_1, \ldots, X_{N-1}) \\
&\overset{(i)}{=} H(X_0) + H(X_1|X_0) + H(X_2|X_1) + \cdots + H(X_N|X_{N-1}) \\
&= 0 + \underbrace{H(1/4) + \cdots + H(1/4)}_{N \text{ όροι}} = NH(1/4) \\
&= N\left(\frac{1}{4}\log 4 + \frac{3}{4}\log\frac{4}{3}\right) = N(2 - \frac{3}{4}\log 3) \approx 0.8113N.
\end{aligned}$$

(i) Οι X_i σχηματίζουν Μαρκοβιανή αλυσίδα.

Ο ρυθμός εντροπίας ισούται με

$$H(\mathcal{X}) = \lim_{N\to\infty} \frac{1}{N} H(X_0, X_1, \ldots, X_N) = H(1/4) \approx 0.8113 \text{ bits}.$$

Εναλλακτικά,

$$H(\mathcal{X}) = \lim_{N\to\infty} H(X_N|X_{N-1}, \ldots, X_0) = H(X_N|X_{N-1}) = H(1/4).$$

Για να υπολογίσουμε την $H(X_0, X_1, \ldots, X_{j-1}, X_{j+1}, \ldots, X_N)$ εφαρμόζουμε και πάλι τον κανόνα αλυσίδας.

$$\begin{aligned}
&H(X_0, X_1, \ldots, X_N) = H(X_0, X_1, \ldots, X_{j-1}, X_{j+1}, \ldots, X_N) \\
&+ H(X_j|X_0, X_1, \ldots, X_{j-1}, X_{j+1}, \ldots, X_N) \Rightarrow \\
&H(X_0, X_1, \ldots, X_{j-1}, X_{j+1}, \ldots, X_N) = H(X_0, X_1, \ldots, X_N) \\
&- H(X_j|X_0, X_1, \ldots, X_{j-1}, X_{j+1}, \ldots, X_N) \\
&\overset{(ii)}{=} H(X_0, X_1, \ldots, X_N) - H(X_j|X_{j-1}, X_{j+1}).
\end{aligned}$$

(ii) Οι X_i σχηματίζουν Μαρκοβιανή αλυσίδα.
Μπορούμε να υπολογίσουμε την $H(X_j|X_{j-1}, X_{j+1})$ διακρίνοντας περιπτώσεις:

1) $X_{j+1} = X_{j-1}$: $\Pr\{X_{j+1} = X_{j-1}\} = (3/4)^2$. Στην περίπτωση αυτή, $X_j = X_{j+1} = X_{j-1}$ και $H(X_j|X_{j-1}, X_{j+1}) = 0$.
2) $X_{j+1} = (1.01)^2 X_{j-1}$: $\Pr\{X_{j+1} = (1.01)^2 X_{j-1}\} = (1/4)^2$. Επομένως, $X_j = 1.01 \cdot X_{j-1}$ και $H(X_j|X_{j-1}, X_{j+1}) = 0$.
3) $X_{j+1} = 1.01 \cdot X_{j-1}$: Στην περίπτωση αυτή υπάρχουν δύο ισοπίθανα ενδεχόμενα (το καθένα με πιθανότητα $(1/4) \cdot (3/4)$). $X_{j+1} = X_j = 1.01 \cdot X_{j-1}$ ή $X_{j+1} = 1.01 \cdot X_j = 1.01 \cdot X_{j-1}$. Συνεπώς, $H(X_j|X_{j-1}, X_{j+1}) = 1$.

Συνδυάζοντας τα παραπάνω,

$$H(X_j|X_{j-1}, X_{j+1}) = 2(1/4)(3/4) \cdot 1 = 3/8.$$

Επομένως,

$$H(X_0, X_1, \ldots, X_{j-1}, X_{j+1}, \ldots, X_N) = NH(1/4) - 3/8 \approx 0.8113N - 0.375.$$

Εναλλακτική λύση (δόθηκε από 2 φοιτητές στο διαγώνισμα):
Με χρήση κανόνα αλυσίδας,

$$\begin{aligned}
&H(X_0, X_1, \ldots, X_{j-1}, X_{j+1}, \ldots, X_N) \\
&= H(X_0) + H(X_1|X_0) + H(X_2|X_1, X_0) + \ldots \\
&+ H(X_{j-1}|X_0^{j-2}) + H(X_{j+1}|X_0^{j-1}) + H(X_{j+2}|X_{j+1}, X_0^{j-1}) + \ldots + H(X_N|X_{j+1}^N, X_0^{j-1}) \\
&\overset{(iii)}{=} H(X_1|X_0) + H(X_2|X_1) + \ldots + H(X_{j-1}|X_{j-2}) + H(X_{j+1}|X_{j-1}) \\
&+ H(X_{j+2}|X_{j+1}) + \ldots + H(X_N|X_{N-1}) \\
&= (N-2)H(1/4) + H(X_{j+1}|X_{j-1}).
\end{aligned}$$

(iii) Οι X_i σχηματίζουν Μαρκοβιανή αλυσίδα.

Η $H(X_{j+1}|X_{j-1})$ μπορεί να βρεθεί από την $p(x_{j+1}|x_{j-1})$. Εύκολα βλέπουμε ότι $X_{j+1} = \{X_j, (1.01)X_j, (1.01)^2X_j\}$ με πιθανότητες $\{(1/4)^2, 2(1/4)(3/4), (3/4)^2\}$, αντίστοιχα. Συνεπώς,

$$\begin{aligned} H(X_{j+1}|X_{j-1}) &= \frac{1}{16}\log 16 + \frac{6}{16}\log\frac{16}{6} + \frac{9}{16}\log\frac{16}{9} \\ &= \frac{1}{4} + \frac{3}{2} - \frac{3}{8}\log 6 + \frac{9}{4} - \frac{9}{16}\log 9 \\ &= \frac{16}{4} - \frac{3}{8} - \frac{3}{8}\log 3 - \frac{9}{8}\log 3 \\ &= \frac{29}{8} - \frac{3}{2}\log 3 \approx 1.2475. \end{aligned}$$

Επομένως,

$$H(X_0, X_1, \ldots, X_{j-1}, X_{j+1}, \ldots, X_N) \approx 0.8113N - 0.375.$$

(γ) Δύο φοιτήτριες προσπαθούν να εκτιμήσουν πόσοι μήνες θα χρειαστούν ώστε ο καταθέτης να καταφέρει να *οκταπλασιάσει* το αρχικό του κεφάλαιο. Η εκτίμηση της πρώτης είναι ότι αυτό θα έχει συμβεί σχεδόν σίγουρα σε 209 μήνες. Η δεύτερη ισχυρίζεται ότι η εκτίμηση αυτή είναι παρακινδυνευμένη και υποθέτει ότι θα πρέπει να περιμένουμε τουλάχιστον 279 μήνες. Ποια από τις δύο εκτιμήσεις είναι ορθότερη; Δικαιολογήστε επαρκώς την απάντησή σας. Όπως και στο προηγούμενο ερώτημα, θεωρούμε ότι, στο τέλος κάθε μήνα, ο καταθέτης εισπράττει τους τόκους με πιθανότητα 1/4.

Απάντηση:

Η πρώτη φοιτήτρια υποθέτει ότι ο καταθέτης δε θα εισπράξει ποτέ τόκους, ούτως ώστε να ισχύει $1000(1.01)^N > 8000 \Rightarrow N_{\min} = 209$. Ωστόσο, αυτό θα συμβεί με πιθανότητα $(3/4)^{209} \approx 7.72 \cdot 10^{-27}$!

Η δεύτερη φοιτητρια, η οποία πιθανώς γνωρίζει το AEP (ή, ισοδύναμα, το νόμο των μεγάλων αριθμών), σκέφτεται ότι, καθώς το N μεγαλώνει, η πιθανότητα συγκεντρώνεται στις τυπικές ακολουθίες. Επομένως, θεωρεί ότι ο καταθέτης θα εισπράξει τους τόκους *περίπου* $N/4$ φορές και ότι δε θα τους εισπράξει τις υπόλοιπες $3N/4$. Επομένως, επιλύει την εξίσωση $1000(1.01)^{3N/4} > 8000 \Rightarrow N_{\min} = 279$. Παρόλο που ο ακριβής υπολογισμός της πιθανότητας η δεύτερη φοιτήτρια να έχει δίκιο δεν είναι εύκολος, παρατηρούμε, κατ' αρχάς, ότι, δεδομένου ότι περιμέναμε περισσότερο, η πιθανότητα να έχουμε φτάσει τις 8000 είναι μεγαλύτερη από $(3/4)^{209} \approx 7.72 \cdot 10^{-27}$. Ωστόσο, το πόσο κοντά είμαστε στο 1, εξαρτάται από το αν το N είναι αρκετά μεγάλο. Αυτό μπορεί να επαληθευτεί μόνο αριθμητικά και δεν είναι εύκολο.

(*δ) Δώστε μια έκφραση για την τιμή της $H(X_N)$ για $N \to \infty$.

Υπόδειξη: Χρησιμοποιήστε το κεντρικό οριακό θεώρημα. Η έκφραση που θα προκύψει είναι συνάρτηση του N.

$\boxed{\text{Απάντηση:}}$

Έστω Z_i τ.μ. με τιμές 0 και 1 που υποδηλώνει εάν ο καταθέτης εισέπραξε τους τόκους το μήνα i. Επομένως, ο αριθμός των φορών που εισπράχθηκαν οι τόκοι σε διάστημα N μηνών ισούται με $A_N = \sum_{i=1}^{N} Z_i$. Οι Z_i είναι Bernoulli i.i.d. με $p = \Pr\{Z_i = 1\} = 1/4$. Συνεπώς, η A_N ακολουθεί διωνυμική (binomial) κατανομή. Παρατηρούμε ότι $X_N = 1000 \cdot (1.01)^{N-A_N}$. Επομένως, σε κάθε τιμή της A_N αντιστοιχεί μια μοναδική τιμή της X_N η οποία έχει την ίδια μάζα πιθανότητας, με αποτέλεσμα $H(X_N) = H(A_N)$.

Απομένει να βρούμε την εντροπία της διωνυμικής κατανομής. Ωστόσο, αυτό αποτελεί δισεπίλυτο πρόβλημα. Αποδεικνύεται ότι η $H(X_N)$ τείνει ασυμπτωτικά στην τιμή $\frac{1}{2}\log(2\pi e N p(1-p)) + \sum_{k\geq 1} a_k N^{-k}$, όπου a_k σταθερές.

Εμείς θα εστιαστούμε μόνο στον πρώτο όρο. Από το κεντρικό οριακό θεώρημα γνωρίζουμε ότι ο μέσος όρος αθροίσματος i.i.d. τ.μ. τείνει στη Γκαουσιανή κατανομή με μέσο όρο μ και διασπορά σ^2/N, όπου μ και σ^2 ο μέσος όρος και η διασπορά, αντίστοιχα, της Z. Δεδομένου ότι η εντροπία συνεχούς τ.μ. δεν εξαρτάται από το μέσο όρο της, εξετάζουμε τη διασπορά. Εύκολα προκύπτει ότι η διασπορά τ.μ. Bernoulli ισούται με $\sigma^2 = \mathbb{E}[X^2] - \mathbb{E}[X]^2 = p(1-p)$. Συνεπώς, $[((1/N)\sum Z_i)^2 - \mu^2] = p(1-p) \Rightarrow \mathbb{E}[(\sum Z_i)^2 - N^2\mu^2] = N^2 p(1-p)$. Επομένως, ή A_N τείνει σε Γκαουσιανή κατανομή με μέση τιμή $N\mu$ και διασπορά $N^2 p(1-p)/N = Np(1-p)$.

Αντικαθιστώντας στην έκφραση για τη (διαφορική) εντροπία Γκαουσιανής τ.μ.,

$$H(X_N) = H(A^N) \to \frac{1}{2}\log(2\pi e N p(1-p)) \to \frac{1}{2}\log(N) \text{ για } N \to \infty.$$

<u>Εναλλακτική λύση (από το συνάδελφό σας Α. Μεσολογγίτη):</u>

Μπορούμε να γράψουμε $X_N = 1000\prod_{i=1}^{N}(1+0.01Z_i)$, όπου Z_i τ.μ. Bernoulli που υποδηλώνει εάν ο τόκος εισπράχθηκε ($Z_i = 0$) ή όχι. Λογαριθμίζοντας,

$$\log(X_N) = \log(1000) + \sum_{i=1}^{N}\log(1+0.01Z_i).$$

Εάν θέσουμε $S_N \triangleq X_N - 1000$ και $Y_i \triangleq \log(1+0.01Z_i)$, η S_N είναι άθροισμα i.i.d. τ.μ. Y_i με μέση τιμή

$$\begin{aligned}
[\log(1+0.01Z_i)] &= p\cdot 0 + (1-p)\cdot\log 1.01 = 0.0108, \text{ και} \\
\sigma^2(Y_i) &= [\log(1+0.01Z_i)^2] - (0.0108)^2 \\
&= p\cdot 0 + (1-p)\log(1.01)^2 - (0.0108)^2 = 3.864\cdot 10^{-5}.
\end{aligned}$$

Και πάλι, επειδή υπάρχει αντιστοιχία 1-προς-1 μεταξύ $S_N = \log(X_N)$ και X_N,

$$H(X_N) = H(S_N) \to \frac{1}{2}\log(2\pi e N \times 3.864\cdot 10^{-5}) \to \frac{1}{2}\log(N) \text{ για } N \to \infty.$$

Η διαφορά εδώ είναι ότι η S_N δεν ακολουθεί διωνυμική κατανομή γιατί οι Y_i δεν είναι Bernoulli.