(11) ΚΑΝΟΝΕΣ ΑΛΥΣΙΔΑΣ

ΟΡΙΣΜΟΙ: 1) $H(X) = -\sum_{x \in X} p(x) \log p(x)$

$\rightsquigarrow H(X,Y) = -\sum_{x \in X} \sum_{y \in Y} p(x,y) \log p(x,y)$

$\rightsquigarrow H(X_1, X_2, \dots, X_n) = -\sum_{x_1 \in X_1} \sum_{x_2 \in X_2} \dots \sum_{x_n \in X_n} p(x_1, x_2, x_3, \dots, x_n) \cdot \log p(x_1, x_2, \dots, x_n)$

2) $H(Y|X) = -\sum_{x \in X} \sum_{y \in Y} p(x,y) \log p(y|x)$

$= \sum_{x \in X} p(x) \underbrace{\sum_{y \in Y} p(y|x) \log p(y|x)}_{H(Y|X=x)}$

$\rightsquigarrow H(Y_1, Y_2, \dots, Y_n | X_1, X_2, \dots, X_k)$

$= -\sum_{x \in X_1} \dots \sum_{x \in X_m} \sum_{y \in Y_1} \dots \sum_{y \in Y_m} p(x_1, x_2, \dots, x_k, y_1, y_2, \dots, y_n) \cdot \log p(y_1 \dots y_n | x_1, \dots, x_k)$

$= \sum_{x \in X_1} \dots \sum_{x \in X_m} p(x_1, x_2, \dots, x_k) \sum_{y \in Y_1} \dots \sum_{y \in Y_m} p(y_1, \dots y_n | x_1, x_2, \dots, x_k) \cdot \log p(y_1, \dots, y_n | x_1, x_2, \dots, x_k)$

ΤΙ ΣΗΜΑΙΝΟΥΝ ΟΛΑ ΑΥΤΑ;

ΘΕΩΡΗΜΑ

(ΚΑΝΟΝΑΣ ΑΛΥΣΙΔΑΣ ΓΙΑ ΤΗΝ ΕΝΤΡΟΠΙΑ)

$$H(X_1, X_2, \ldots, X_n) = \sum_{i=1}^{n} H(X_i \mid X_1, X_2, \ldots, X_{i-1})$$

$$\left(= H(X_1) + H(X_2 \mid X_1) + H(X_3 \mid X_1, X_2) + \ldots \right)$$

(ΠΡΟΦΑΝΗΣ ΦΥΣΙΚΗ ΕΡΜΗΝΕΙΑ)

ΑΠΟΔΕΙΞΗ:

$$H(X_1, X_2, \ldots, X_n) = \sum_{x_1, \ldots, x_n} p(x_1, \ldots, x_n) \log p(x_1, \ldots, x_n)$$

$$= \sum_{x_1, \ldots, x_n} p(x_1, \ldots, x_n) \log \prod_{i=1}^{n} p(x_i \mid x_1, \ldots, x_{i-1})$$

$$\left(\hookrightarrow p(x_1)\, p(x_2 \mid x_1)\, p(x_3 \mid x_1, x_2) \ldots \right)$$

$$= \sum_{x_1, x_n} \sum_{i=1}^{n} p(x_1, \ldots, x_n) \log p(x_i \mid x_1, \ldots, x_{i-1})$$

$$= \sum_{i=1}^{n} \sum_{x_1, \ldots, x_n} p(x_1, x_2, \ldots, x_n) \log p(x_i \mid x_1, \ldots, x_{i-1})$$

$$= \sum_{i=1}^{n} \sum_{x_1 \ldots x_i} \log p(x_i \mid x_1, \ldots, x_{i-1}) \sum_{x_{i+1} \ldots x_n} p(x_1, x_2, \ldots, x_n)$$

$$= \sum_{i=1}^{n} \sum_{x_1 \ldots x_i} \log p(x_i \mid x_1, \ldots, x_{i-1})\, p(x_1, x_2, \ldots, x_i)$$

(ΟΜΟΙΑ ΜΕ $\sum_y p(x,y) = p(x)$)

$$= \sum_{i=1}^{n} H(X_i \mid X_1, \ldots, X_{i-1})$$

ΟΡΙΣΜΟΙ: $I(X;Y) = \sum_{x\in X, y\in Y} P(x,y) \log \frac{P(x,y)}{P(x)P(y)}$

$\rightsquigarrow$

$I(X_1, X_2 .. X_k; Y_1, Y_2, \dots, Y_n) =$

$$\sum_{x_1 \dots x_k} \sum_{y_1 \dots y_n} p(x_1, \dots, x_n, y_1, \dots, y_n) \log \frac{P(x_1, \dots x_k, y_1, \dots, y_n)}{P(x_1, \dots, x_k) P(y_1, \dots, y_n)}$$

$= H(X_1, X_2, \dots, X_k) - H(X_1, X_2, \dots X_k | Y_1, Y_2, \dots Y_n)$

ΟΡΙΣΜΟΣ: Η ΔΕΣΜΕΥΜΕΝΗ ΑΜΟΙΒΑΙΑ ΠΛΗΡΟΦΟΡΙΑ ΤΩΝ X, Y, ΜΕ ΔΕΔΟΜΕΝΟ ΤΟ Z ΟΡΙΖΕΤΑΙ ΩΣ:

$$I(X;Y|Z) = H(X|Z) - H(X|Y,Z)$$
$$= \sum_{z} P(z) \sum_{x,y} p(x,y|z) \log \frac{p(x,y|z)}{p(x|z)P(y|z)}$$

(ΓΕΝΙΚΕΥΣΗ: $I(X;Y|Z_1 \dots Z_n) = H(X|Z_1, \dots Z_n) - H(X|Y, Z_1, \dots, Z_n)$)

ΚΑΙ $I(X_1, \dots, X_k; Y_1, Y_2, \dots, Y_n | Z_1, \dots, Z_m) = \dots\dots$

(ΜΠΟΡΕΙΤΑΙ ΝΑ ΤΟ ΜΑΝΤΕΨΕΤΕ?)

ΘΕΩΡΗΜΑ (ΚΑΝΟΝΑΣ ΑΛΥΣΙΔΑΣ ΓΙΑ ΤΗΝ ΑΜΟΙΒΑΙΑ ΠΛΗΡΟΦΟΡΙΑ)

$$I(X_1, X_2, \dots, X_n; Y) = \sum_{i=1}^{n} I(X_i; Y | X_1, X_2, \dots, X_{i-1})$$

ΑΠΟΔΕΙΞΗ: $I(X_1, X_2, \dots, X_n; Y) =$

$H(X_1, X_2, \dots, X_n) - H(X_1, X_2, \dots, X_n | Y)$

$= \sum_{i=1}^{n} H(X_i | X_1, \dots, X_{i-1}) - \sum_{i=1}^{n} H(X_i | X_{i-1}, \dots, X_1, Y)$

(ΔΕΝ ΤΟ ΑΠΟΔΕΙΞΑΜΕ, ΑΛΛΑ ΠΡΟΚΥΠΤΕΙ ΑΠΟ ΤΟΝ ΚΑΝΟΝΑ ΑΛΥΣΙΔΑΣ ΓΙΑ ΤΗΝ ΕΝΤΡΟΠΙΑ)

$$= \sum_{i=1}^{n} H(X_i \mid X_1, \dots, X_{i-1}) - H(X_i \mid X_{i-1} \dots X_1, Y)$$

$$= \sum_{i=1}^{n} I(X_i ; Y \mid X_1, X_2, \dots, X_{i-1})$$

ΟΡΙΣΜΟΙ: • $D(p \| q) = \sum_{x \in \mathcal{X}} p(x) \log \frac{p(x)}{q(x)}$

• $D(p(x,y) \| q(x,y)) = \sum_{x \in \mathcal{X}} \sum_{y \in \mathcal{Y}} p(x,y) \log \frac{p(x,y)}{q(x,y)}$

• ΔΕΣΜΕΥΜΕΝΗ ΣΧΕΤΙΚΗ ΕΝΤΡΟΠΙΑ:

$$D(p(y|x) \| q(y|x)) = \sum_{x \in \mathcal{X}} p(x) \sum_{y \in \mathcal{Y}} p(y|x) \log \frac{p(y|x)}{q(y|x)}$$

ΘΕΩΡΗΜΑ (ΚΑΝΟΝΑΣ ΑΛΥΣΙΔΑΣ ΓΙΑ ΤΗ ΣΧΕΤΙΚΗ ΕΝΤΡΟΠΙΑ)

$$D(p(x,y) \| q(x,y)) = D(p(x) \| q(x)) + D(p(y|x) \| q(y|x))$$

ΑΠΟΔΕΙΞΗ:

$$D(p(x,y) \| q(x,y)) = \sum_{x} \sum_{y} p(x,y) \log \frac{p(x,y)}{q(x,y)}$$

$$= \sum_{x} \sum_{y} p(x,y) \log \frac{p(x)\, p(y|x)}{q(x)\, q(y|x)}$$

$$= \sum_{x} \sum_{y} p(x,y) \log \frac{p(x)}{q(x)} + \sum_{x} \sum_{y} p(x) p(y|x) \log \frac{p(y|x)}{q(y|x)}$$

$$\hookrightarrow D(p_{(x)} \| q(x)) + D(p(y|x) \| q(y|x))$$

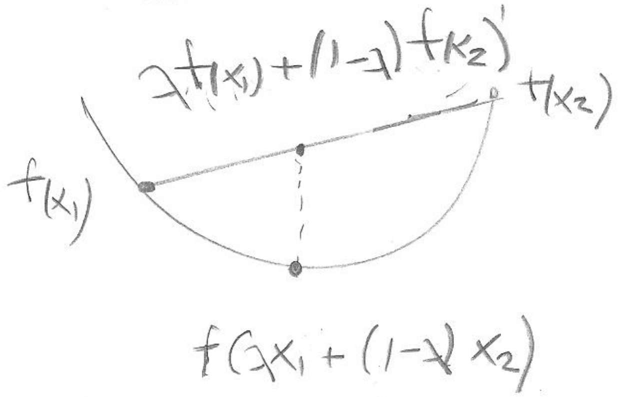
12 ΑΝΙΣΟΤΗΤΑ ΤΟΥ JENSEN ΚΑΙ ΕΦΑΡΜΟΓΕΣ

ΟΡΙΣΜΟΣ: ΜΙΑ ΣΥΝΑΡΤΗΣΗ ΛΕΓΕΤΑΙ ΚΥΡΤΗ (CONVEX) ΣΕ ΕΝΑ ΔΙΑΣΤΗΜΑ (α, b) ΑΝ $\forall\, x_1, x_2 \in (\alpha, b)$ ΚΑΙ $\forall\, \lambda \in [0,1]$,

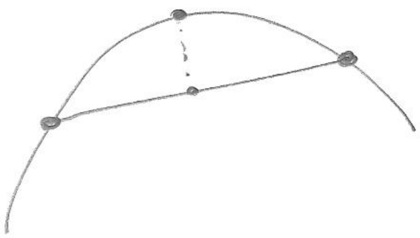
$$f(\lambda x_1 + (1-\lambda)x_2) \leq \lambda f(x_1) + (1-\lambda) f(x_2)$$

Η f ΛΕΓΕΤΑΙ ΑΥΣΤΗΡΩΣ (STRICTLY) ΚΥΡΤΗ ΑΝ Η ΙΣΟΤΗΤΑ ΙΣΧΥΕΙ ΜΟΝΟ ΓΙΑ $\lambda = 0$ ΚΑΙ $\lambda = 1$. ΜΙΑ ΣΥΝΑΡΤΗΣΗ f ΛΕΓΕΤΑΙ (ΑΥΣΤΗΡΩΣ) ΚΟΙΛΗ (CONCAVE) ΑΝ Η $-f(x)$ ΕΙΝΑΙ (ΑΥΣΤΗΡΩΣ) ΚΥΡΤΗ

ΚΥΡΤΗ

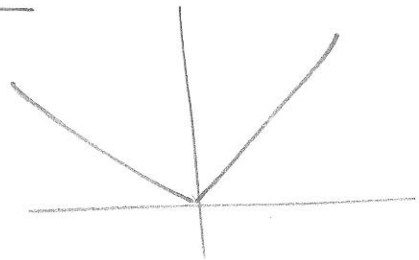


ΚΟΙΛΗ

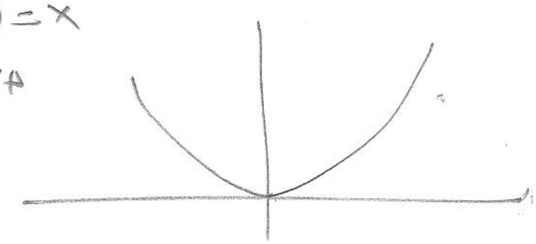


ΠΑΡΑΔΕΙΓΜΑΤΑ

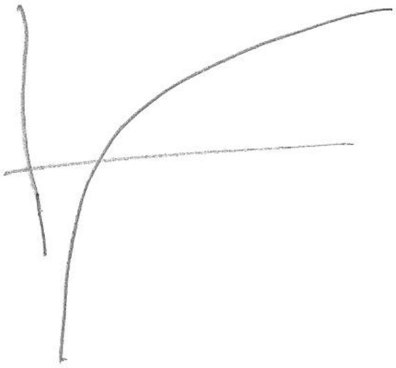
$f(x) = |x|$
ΚΥΡΤΗ



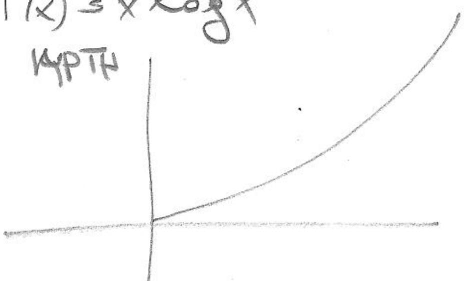
$f(x) = x^2$
ΚΥΡΤΗ



$f(x) = \log x$
ΚΟΙΛΗ



$f(x) = x \log x$
ΚΥΡΤΗ



28

ΘΕΩΡΗΜΑ (ΚΡΙΤΗΡΙΟ ΚΥΡΤΟΤΗΤΑΣ) ΑΝ Η f ΕΧΕΙ ΔΕΥΤΕΡΗ ΠΑΡΑΓΩΓΟ ΠΟΥ ΕΙΝΑΙ ΜΗ ΑΡΝΗΤΙΚΗ (ΘΕΤΙΚΗ) ΣΕ ΕΝΑ ΔΙΑΣΤΗΜΑ, Η $f(x)$ ΕΙΝΑΙ ΚΥΡΤΗ (ΑΥΣΤΗΡΩΣ ΚΥΡΤΗ) ΣΕ ΑΥΤΟ ΤΟ ΔΙΑΣΤΗΜΑ

ΠΑΡΑΤΗΡΗΣΗ: ΟΙ ΚΥΡΤΕΣ ΣΥΝΑΡΤΗΣΕΙΣ ΕΙΝΑΙ ΠΟΛΥ ΧΡΗΣΙΜΕΣ ΣΤΗ ΒΕΛΤΙΣΤΟΠΟΙΗΣΗ (ΠΑΠ3)

ΘΕΩΡΗΜΑ: (ΑΝΙΣΟΤΗΤΑ JENSEN)

(i) ΑΝ Η f ΕΙΝΑΙ ΚΥΡΤΗ ΚΑΙ Η X ΕΙΝΑΙ ΤΥΧΑΙΑ ΜΕΤΑΒΛΗΤΗ,

$$Ef(X) \geq f(EX)$$

(ii) ΕΠΙΠΛΕΟΝ, ΑΝ Η f ΕΙΝΑΙ ΑΥΣΤΗΡΩΣ ΚΥΡΤΗ,

$$Ef(X) = f(EX) \Rightarrow$$ Η X ΕΙΝΑΙ ΣΤΑΘΕΡΑ (ΔΗΛΑΔΗ $X = x_0$ ΜΕ Π.Θ. 1)

ΑΠΟΔΕΙΞΗ: (i) ΙΣΧΥΕΙ ΑΝ ΤΟ X ΠΑΙΡΝΕΙ 2 ΤΙΜΕΣ: $X = \begin{cases} x_1, & \text{Μ.Π. } p_1 \\ x_2, & \text{Μ.Π. } p_2 \end{cases}$

$$Ef(X) = p_1 f(x_1) + \underbrace{p_2}_{1-p} f(x_2) \geq f(p_1 x_1 + p_2 x_2) = f(EX)$$

ΕΣΤΩ ΟΤΙ ΙΣΧΥΕΙ ΓΙΑ $k-1$ ΤΙΜΕΣ. ΕΣΤΩ X ΠΟΥ ΠΑΙΡΝΕΙ k ΤΙΜΕΣ:

$$Ef(X) = \sum_{i=1}^{k} p_i f(x_i) = p_k f(x_k) + \left[\sum_{i=1}^{k-1} \frac{p_i}{1-p_k} f(x_i)\right](1-p_k)$$

$$\overset{\text{ΒΗΜΑ } k-1}{\geq} p_k f(x_k) + f\left(\sum_{i=1}^{k-1} \frac{p_i}{1-p_k} x_k\right)(1-p_k) \overset{\text{ΒΗΜΑ 2}}{\geq}$$

$$f\left(p_k x_k + \sum_{i=1}^{k-1} \frac{p_i x_k}{1-p_k}(1-p_k)\right) =$$

$$f\left(\sum_{i=1}^{k} p_i x_i\right) = f(E(X))$$

(ii) ΙΣΧΥΕΙ ΑΝ ΥΠΟΘΕΣΟΥΜΕ ΟΤΙ ΤΟ Χ ΠΑΡΝΕΙ ΔΥΟ ΤΙΜΕΣ:

$$E(f(X)) = f(EX) \Leftrightarrow$$

$$f(p_1 x_1 + p_2 x_2) = p_1 f(x_1) + p_2 f(x_2)$$

$$\Rightarrow p_1 = 0 \text{ Η' } p_2 = 0$$

ΕΣΤΩ ΟΤΙ ΙΣΧΥΕΙ ΑΝ ΥΠΟΘΕΣΟΥΜΕ ΟΤΙ ΤΟ Χ ΠΑΙΡΝΕΙ $k-1$ ΤΙΜΕΣ. ΕΧΟΥΜΕ, ΓΙΑ k ΤΙΜΕΣ:

$$E(f(x)) = f(EX) \Leftrightarrow$$

$$p_1 f(x_1) + \dots + p_k f(x_k) = f(p_1 x_1 + \dots + p_k x_k) \quad (A)$$

ΠΑΙΡΝΩ ΠΕΡΙΠΤΩΣΕΙΣ: (1) ΟΛΑ ΜΗΔΕΝ ΕΚΤΟΣ ΑΠΟ ΕΝΑ $\Rightarrow$ ΤΕΛΕΙΩΣΑΜΕ

(2) ΜΕΡΙΚΑ 0 $\Rightarrow$ ΧΡΗΣΙΜΟΠΟΙΟΥΜΕ ΕΠΑΓΩΓΙΚΗ ΥΠΟΘΕΣΗ $\Rightarrow$ ΤΕΛΕΙΩΣΑΜΕ (ΟΛΑ 0 ΕΚΤΟΣ ΑΠΟ 2) (3) ΚΑΝΕΝΑ 0. ΤΟΤΕ ΕΧΟΥΜΕ $0 < p_i < 1 \ \forall i$

$$(A) \Leftrightarrow \left(\frac{p_1}{1-p_k} f(x_1) + \dots + \frac{p_{k-1}}{1-p_k} f(x_{k-1})\right)(1-p_k) + p_k f(x_k)$$

$$= f\left(\left(\frac{p_1}{1-p_k} x_1 + \dots + \frac{p_{k-1}}{1-p_k} x_{k-1}\right)(1-p_k) + p_k x_k\right) \Leftrightarrow$$

ΒΗΜΑ 2

$$< (1-p_k) f\left(\frac{p_1}{1-p_k} x_1 + \dots + \frac{p_{k-1}}{1-p_k} x_{k-1}\right) + p_k f(x_k)$$

ΒΗΜΑ $k-1$

$$< p_1 f(x_1) + \dots + p_{k-1} f(x_{k-1}) + p_k f(x_k) \Rightarrow \text{ΑΤΟΠΟ}$$

ΘΕΩΡΗΜΑ: (INFORMATION INEQUALITY – ΑΝΙΣΟΤΗΤΑ ΠΛΗΡΟΦΟΡΙΑΣ)

ΕΣΤΩ $p(x)$, $q(x)$, $x \in \mathcal{X}$ ΔΥΟ ΔΙΑΚΡΙΤΕΣ ΚΑΤΑΝΟΜΕΣ

$D(p\|q) \geq 0$, ΜΕ ΙΣΟΤΗΤΑ ΑΝ ΚΑΙ ΜΟΝΟ ΑΝ $p=q$

ΑΠΟΔΕΙΞΗ:

$$-D(p\|q) = -\sum_{x\in\mathcal{X}} p(x)\log\frac{p(x)}{q(x)}$$

$$= -\sum_{x\in A} p(x)\log\frac{p(x)}{q(x)} \qquad A=\{x: p(x)>0\}$$

$$= \sum_{x\in A} p(x)\log\frac{q(x)}{p(x)} \rightarrow E\left[f\left(\frac{q(x)}{p(x)}\right)\right]$$

$$\overset{①}{\leq} \log\left(\sum_{x\in A} p(x)\frac{q(x)}{p(x)}\right) \rightarrow f\left[E\left[\frac{q(x)}{p(x)}\right]\right]$$

$$= \log\sum_{x\in A} q(x)$$

$$\overset{②}{\leq} \log\sum_{x\in\mathcal{X}} q(x) = \log 1 = 0$$

ΕΠΙΣΗΣ:
- ΑΝ $p=q$, ΙΣΧΥΕΙ Η ΙΣΟΤΗΤΑ
- ΑΝ ΙΣΧΥΕΙ Η ΙΣΟΤΗΤΑ, ΙΣΧΥΟΥΝ ΟΙ ①,② ΜΕ ΙΣΟΤΗΤΑ

ΟΜΩΣ Η $\log x$ ΕΙΝΑΙ ΑΥΣΤΗΡΩΣ ΚΟΙΛΗ ΣΥΝΑΡΤΗΣΗ, ΑΡΑ ΓΙΑ ΝΑ ΙΣΧΥΕΙ Η ① ΠΡΕΠΕΙ $\frac{q(x)}{p(x)} = c$ ΓΙΑ ΚΑΘΕ $x\in A$

$\Rightarrow q(x) = c\,p(x) \quad \forall x \in A$.

ΟΜΩΣ ② $\Rightarrow 1 = \sum_{x\in\mathcal{X}} q(x) = \sum_{x\in A} q(x) = c\sum_{x\in A} p(x) = c\cdot 1 = c \Rightarrow c=1$

$\Rightarrow \quad q(x) = p(x) \quad \forall x \in A$ ΚΑΙ ΠΡΟΦΑΝΩΣ

$q(x) = 0 \quad \forall x \in \mathcal{X} - A$ (ΟΠΟΥ ΕΠΙΣΗΣ $p(x)=0$)

ΛΗΜΜΑ 1: $I(X;Y) \geq 0$ ΜΕ ΙΣΟΤΗΤΑ ΑΝ ΚΑΙ ΜΟΝΟ ΑΝ X, Y ΑΝΕΞΑΡΤΗΤΑ

ΑΠΟΔΕΙΞΗ: $I(X;Y) = D\big(p(x,y) \,\|\, p(x)p(y)\big) \geq 0$

ΑΚΟΜΑ: X, Y ΑΝΕΞΑΡΤΗΤΑ $\Leftrightarrow p(x,y) = p(x)p(y)$

$\Leftrightarrow D\big(p(x,y) \,\|\, p(x)p(y)\big) = 0 \Leftrightarrow I(X;Y) = 0$

ΛΗΜΜΑ 2: $D\big(p(y|x) \,\|\, q(y|x)\big) \geq 0$ ΜΕ ΙΣΟΤΗΤΑ ΑΝ ΚΑΙ ΜΟΝΟ ΑΝ $p(y|x) = q(y|x) \ \forall x$

ΑΠΟΔΕΙΞΗ: $D\big(p(y|x) \,\|\, q(y|x)\big) = \sum_x p(x) \underbrace{D\big(p(y|x) \,\|\, q(y|x)\big)}_{\geq 0}$

ΚΑΙ ΠΡΟΚΥΠΤΕΙ ΑΠΛΑ

ΛΗΜΜΑ 3: $I(X;Y|Z) \geq 0$ ΜΕ ΙΣΟΤΗΤΑ ΑΝ ΚΑΙ ΜΟΝΟ ΑΝ ΤΑ X, Y ΑΝΕΞΑΡΤΗΤΑ ΔΕΔΟΜΕΝΟΥ ΤΟΥ Z.

ΑΠΟΔΕΙΞΗ ΟΠΩΣ ΛΗΜΜΑ 1

ΛΗΜΜΑ 4: $H(X) \leq \log|\mathcal{X}|$ ΜΕ ΙΣΟΤΗΤΑ ΑΝ ΚΑΙ ΜΟΝΟ ΑΝ ΤΟ X ΕΧΕΙ ΟΜΟΙΟΜΟΡΦΗ ΚΑΤΑΝΟΜΗ

ΑΠΟΔΕΙΞΗ: ΕΣΤΩ Η ΟΜΟΙΟΜΟΡΦΗ ΚΑΤΑΝΟΜΗ $u(x) = \frac{1}{|\mathcal{X}|}, x \in \mathcal{X}$ ΚΑΙ $p(x)$ Η ΚΑΤΑΝΟΜΗ ΤΟΥ X

$$0 \leq D(p\|u) = \sum p(x) \log \frac{p(x)}{u(x)} =$$

$$\sum p(x) \log |\mathcal{X}| p(x) = \sum p(x) \log |\mathcal{X}| + \sum p(x) \log p(x)$$

$$= \log |\mathcal{X}| - H(x) \Rightarrow$$

$H(x) \leq \log |\mathcal{X}|$. ΙΣΟΤΗΤΑ ΕΧΟΥΜΕ ΑΝ ΚΑΙ ΜΟΝΟ ΑΝ $p = u$, ΔΗΛΑΔΗ ΤΟ X ΕΙΝΑΙ ΚΑΤΑΝΕΜΗΜΕΝΟ ΟΜΟΙΟΜΟΡΦΑ

<u>ΛΗΜΜΑ 5</u> $H(X|Y) \leq H(X)$ ΜΕ ΙΣΟΤΗΤΑ ΑΝ ΚΑΙ ΜΟΝΟ ΑΝ ΤΑ X, Y ΕΙΝΑΙ ΑΝΕΞΑΡΤΗΤΑ

<u>ΑΠΟΔΕΙΞΗ:</u> $0 \leq I(X;Y) = H(X) - H(X|Y)$

<u>ΛΗΜΜΑ 6</u> (ΦΡΑΓΜΑ ΑΝΕΞΑΡΤΗΣΙΑΣ)

$$H(X_1, X_2, \ldots, X_n) \leq \sum_{i=1}^{n} H(X_i)$$

ΜΕ ΙΣΟΤΗΤΑ ΑΝ ΚΑΙ ΜΟΝΟ ΑΝ ΤΑ $X_1, \ldots, X_n$ ΕΙΝΑΙ ΑΝΕΞΑΡΤΗΤΑ

<u>ΑΠΟΔΕΙΞΗ:</u> $H(X_1, X_2, \ldots, X_n) = \sum_{i=1}^{n} H(X_i | X_1, \ldots, X_{i-1})$

$\leq \sum_{i=1}^{n} H(X_i)$ ΜΕ ΙΣΟΤΗΤΑ ΑΝ ΚΑΙ ΜΟΝΟ ΑΝ

(ΑΠΛΗ ΑΠΟΔΕΙΞΗ, ΔΙΑΙΣΘΗΤΙΚΑ ΑΠΛΟ)

X_i ΑΝΕΞΑΡΤΗΤΟ ΤΩΝ $X_1, \ldots, X_{i-1}$ $\Longleftrightarrow$

ΤΑ X_i ΑΝΕΞΑΡΤΗΤΑ

$$p(x_1, x_2, \ldots x_i) = p(x_1, x_2, \ldots, x_{i-1})\, p(x_i)$$

<u>ΘΕΩΡΗΜΑ</u>: ΑΣΘΕΝΗΣ ΝΟΜΟΣ ΤΩΝ ΜΕΓΑΛΩΝ ΑΡΙΘΜΩΝ (WEAK LAW OF LARGE NUMBERS): ΕΣΤΩ $X_1, X_2, \ldots$ ΑΚΟΛΟΥΘΙΑ ΑΠΟ ΑΝΕΞΑΡΤΗΤΕΣ, ΟΜΟΙΟΜΟΡΦΑ ΚΑΤΑΝΕΜΗΜΕΝΕΣ ΤΥΧΑΙΕΣ ΜΕΤΑΒΛΗΤΕΣ (INDEPENDENT, IDENTICALLY DISTRIBUTED) ΤΟΤΕ: (ΕΣΤΩ $\mu = E(X_1) = E(X_2) = \ldots$)

$$P\left(\left| \frac{1}{n} \sum_{i=1}^{n} X_i - \mu \right| > \epsilon \right) \xrightarrow[n \to \infty]{} 0, \quad \forall \epsilon > 0$$

ΙΣΟΔΥΝΑΜΑ, ΓΡΑΦΟΥΜΕ

$$\frac{1}{n} \sum_{i=1}^{n} X_i \xrightarrow{P} \mu \quad \text{ΟΤΑΝ } n \to \infty$$

<u>ΘΕΩΡΗΜΑ (AEP)</u>: ΑΝ ΟΙ $X_1, X_2, \ldots$ IID $\sim p(x)$, ΤΟΤΕ:

$$-\frac{1}{n} \log p(X_1, X_2, \ldots, X_n) \xrightarrow{P} H(X)$$

$$\Leftrightarrow P\left(\left| -\frac{1}{n} \log p(X_1, X_2, \ldots, X_n) - H(X) \right| > \epsilon \right) \to 0$$

$$\Leftrightarrow P\left(\epsilon \geq -\frac{1}{n} \log p(X_1, X_2, \ldots, X_n) - H(X) \geq -\epsilon \right) \to 1$$

$$\Leftrightarrow P\left(2^{-n(H(X)+\epsilon)} \leq p(X_1, X_2, \ldots, X_n) \leq 2^{-n(H(X)-\epsilon)} \right) \to 1$$

ΔΗΛΑΔΗ Η ΠΙΘΑΝΟΤΗΤΑ ΜΙΑΣ ΑΚΟΛΟΥΘΙΑΣ ΠΟΥ ΕΜΦΑΝΙΖΕΤΑΙ ΕΙΝΑΙ ΠΑΝΤΑ ΚΟΝΤΑ ΣΤΟ $2^{-nH(X)}$

ΑΠΟΔΕΙΞΗ:

$$-\frac{1}{n}\log p(X_1, X_2, \dots, X_n)$$

$$\overset{\text{IID}}{=} -\frac{1}{n}\log p(X_1)p(X_2)\dots p(X_n)$$

$$= -\frac{1}{n}\sum_{i=1}^{n}\log p(X_i)$$

$$\xrightarrow{P} E\left[-\log p(X)\right] = H(X)$$

ΟΡΙΣΜΟΣ: ΟΡΙΖΟΥΜΕ ΤΥΠΙΚΟ ΣΥΝΟΛΟ $A_\epsilon^{(n)}$ ΩΣ ΠΡΟΣ ΤΗ ΣΥΝΑΡΤΗΣΗ $p(x)$ ΤΟ ΣΥΝΟΛΟ ΤΩΝ ΑΚΟΛΟΥΘΙΩΝ $(x_1, x_2, \dots, x_n) \in \mathcal{X}^n$ ΜΕ ΤΗΝ ΙΔΙΟΤΗΤΑ

$$2^{-n(H(X)+\epsilon)} \leq p(x_1, x_2, \dots, x_n) \leq 2^{-n(H(X)-\epsilon)} \quad \text{Ⓐ}$$

ΘΕΩΡΗΜΑ:

① ΑΝ $(x_1, x_2, \dots, x_n) \in A_\epsilon^{(n)}$ ΤΟΤΕ

$$H(X) - \epsilon \leq -\frac{1}{n}\log p(x_1, x_2, \dots, x_n) \leq H(X) + \epsilon$$

ΑΠΟΔΕΙΞΗ: ΠΡΟΚΥΠΤΕΙ ΑΠΟ ΤΗΝ Ⓐ

② $P\left(A_\epsilon^{(n)}\right) > 1 - \epsilon \quad \forall n > n_0$

ΑΠΟΔΕΙΞΗ: $P\left(2^{-n(H+\epsilon)} \leq p(X_1, X_2, \dots, X_n) \leq 2^{-n(H-\epsilon)}\right) \to 1$

ΑΡΑ $> 1 - \epsilon \quad \forall n > n_0$

(3) $|A_\epsilon^{(n)}| \le 2^{n(H(X)+\epsilon)}$

ΑΠΟΔΕΙΞΗ:

$$1 = \sum_{\underline{x}\in \mathcal{X}^n} p(x_1,x_2,\dots,x_n) \ge \sum_{(x_1,\dots,x_n)\in A_\epsilon^{(n)}} p(x_1,\dots,x_n)$$

$$\ge \sum_{(x_1,\dots,x_n)\in A_\epsilon^{(n)}} 2^{-n(H(X)+\epsilon)} = |A_\epsilon^{(n)}|\, 2^{-n(H(X)+\epsilon)}$$

$$\Rightarrow |A_\epsilon^{(n)}| \le 2^{n(H(X)+\epsilon)}$$

(4) $|A_\epsilon^{(n)}| \ge (1-\epsilon)\, 2^{n(H(X)-\epsilon)}$ ΓΙΑ ΑΡΚΕΤΑ ΜΕΓΑΛΟ ϵ

ΑΠΟΔΕΙΞΗ: ΓΙΑ ΑΡΚΕΤΑ ΜΕΓΑΛΟ n, (ΔΗΛΑΔΗ $\forall\, n > n_0$),

$P(A_\epsilon^{(n)}) > 1-\epsilon$. ΑΡΑ:

$$1-\epsilon < P(A_\epsilon^{(n)}) \le \sum_{(x_1,\dots,x_n)\in A_\epsilon^{(n)}} 2^{-n(H(X)-\epsilon)} = |A_\epsilon^{(n)}|\, 2^{-n(H(X)-\epsilon)}$$

$$\Rightarrow |A_\epsilon^{(n)}| \ge (1-\epsilon)\, 2^{-n(H(X)-\epsilon)}$$

ΑΡΑ: ΥΠΑΡΧΟΥΝ ΠΕΡΙΠΟΥ $2^{nH(X)}$ ΤΥΠΙΚΕΣ ΑΚΟΛΟΥΘΙΕΣ,

ΚΑΘΕ ΜΙΑ ΜΕ ΠΕΡΙΠΟΥ $2^{-nH(X)}$ ΠΙΘΑΝΟΤΗΤΑ ΝΑ ΣΥΜΒΕΙ

ΑΡΑ ΟΙ ΠΙΘΑΝΟΤΗΤΕΣ ΙΣΟΔΙΑΜΕΡΙΖΟΝΤΑΙ ΑΣΥΜΠΤΩΤΙΚΑ ΑΜΕΣΑ ΣΤΙΣ ΤΥΠΙΚΕΣ ΑΚΟΛΟΥΘΙΕΣ

(14) ΕΦΑΡΜΟΓΕΣ (ΤΗΣ Ι.Α.Ι.) ΣΕ ΚΩΔΙΚΟΠΟΙΗΣΗ

ΕΣΤΩ ΟΤΙ ΘΕΛΟΥΜΕ ΝΑ ΚΩΔΙΚΟΠΟΙΗΣΟΥΜΕ, ΜΕ ΤΟΝ ΛΙΓΟΤΕΡΟ ΑΡΙΘΜΟ bits, ΜΙΑ ΑΚΟΛΟΥΘΙΑ $(X_1, X_2, \dots, X_n)$, X_i iid, $X_i \sim p(x)$.

ΜΙΑ ΛΥΣΗ: ΕΣΤΩ $\epsilon > 0$. ΑΚΟΛΟΥΘΩΣ:

0	ΒΑΖΟΥΜΕ ΟΛΕΣ ΤΙΣ ΤΥΠΙΚΕΣ ΣΤΗ ΣΕΙΡΑ	$2^{n(H(X)+\epsilon)}$, ΑΡΑ ΘΕΛΩ ΤΟ ΠΟΛΥ $n(H(X)+\epsilon)+1$ bits
1	ΜΕΤΑ ΒΑΖΟΥΜΕ ΤΙΣ ΜΗ ΤΥΠΙΚΕΣ ΣΤΗ ΣΕΙΡΑ	ΛΙΓΟΤΕΡΕΣ ΑΠΟ $\lvert X\rvert^n$, ΑΡΑ ΘΕΛΩ ΤΟ ΠΟΛΥ $n\log\lvert X\rvert+1$ bits

ΜΕΤΑ ΒΑΖΩ ΚΑΙ 1 bit prefix. Ο ΚΩΔΙΚΑΣ ΕΙΝΑΙ ΠΡΟΦΑΝΩΣ 1-1

ΕΣΤΩ ΟΤΙ Η ΑΚΟΛΟΥΘΙΑ $(x_1, \dots, x_n)$ ΕΧΕΙ ΜΗΚΟΣ $l(x_1, \dots, x_n)$. ΤΟΤΕ:

$$E[l(X_1, X_2, \dots, X_n)] = \sum_{(x_1, \dots, x_n) \in \mathcal{X}^n} p(x_1, \dots, x_n)\, l(x_1, \dots, x_n)$$

$$= \underbrace{\sum_{(x_1, \dots, x_n) \in A_\epsilon^{(n)}} p(x_1, \dots, x_n)}_{\le 1} \underbrace{l(x_1, \dots, x_n)}_{\le n(H(X)+\epsilon)+2} + \underbrace{\sum_{(x_1, \dots, x_n) \notin A_\epsilon^{(n)}} p(x_1, \dots, x_n)}_{\le \epsilon \text{ (ΑΡΚΕΙ } n \text{ ΜΕΓΑΛΟ)}} \underbrace{l(x_1, \dots, x_n)}_{\le n\log|X|+2}$$

$$\le n(H+\epsilon) + \epsilon n \log|X| + 2$$

$= n(H+\epsilon')$, ΟΠΟΥ $\epsilon' = \epsilon + \epsilon \log|X| + \frac{2}{n}$ ΜΠΟΡΕΙ ΝΑ ΓΙΝΕΙ ΟΣΟ ΜΙΚΡΟ ΘΕΛΟΥΜΕ!