EE728
Προχωρημένα Θέματα Θεωρίας Πληροφορίας
6η, 7η, 8η και 9η διάλεξη
(3η έκδοση, 30/5/2015)

Δημήτρης-Αλέξανδρος Τουμπακάρης

Τμήμα ΗΜ&ΤΥ, Πανεπιστήμιο Πατρών

28 Απριλίου 2015

Περιεχόμενα

Αντιστοιχία 6ης, 7ης, 8ης και 9ης διάλεξης με βιβλία Cover & Thomas και El Gamal & Kim

- Βιβλίο Cover & Thomas (2η έκδοση): 7.4–7.7.
- Βιβλίο El Gamal & Kim: 2.5, 3.1.

Χωρητικότητα Διακριτού Καναλιού Χωρίς Μνήμη

- Ορισμός 6.1. "Πληροφοριακή" Χωρητικότητα Διακριτού Καναλιού Χωρίς Μνήμη

``Information'' Channel Capacity of a DMC

$$C \triangleq \max_{p(x)} I(X;Y)$$

Παραδείγματα Διακριτών Καναλιών Χωρίς Μνήμη

(Επανάληψη από το μάθημα "Θεωρία Πληροφορίας")

- Δυαδικό Συμμετρικό Κανάλι (Binary Symmetric Channel -- BSC): $C = 1 - H(p)$ bits, επιτυγχάνεται με ομοιομορφη $p(x) = \left(\frac{1}{2}, \frac{1}{2}\right)$.
- Δυαδικό Κανάλι με Διαγραφή (Binary Erasure Channel): $C = 1 - \alpha$, όπου α η πιθανότητα διαγραφής. Επιτυγχάνεται με ομοιομορφη $p(x) = \left(\frac{1}{2}, \frac{1}{2}\right)$. Υπενθυμίζεται ότι το κανάλι δεν είναι ασθενώς συμμετρικό (εκτός εάν $\alpha = \frac{1}{3}$).
- Η χωρητικότητα του δυαδικού καναλιού με διαγραφή παραμένει η ίδια εάν χρησιμοποιήσουμε ανάδραση.
- Θα δούμε ότι το αποτέλεσμα αυτό, δηλαδή ότι η χρήση ανάδρασης δεν αυξάνει τη χωρητικότητα, ισχύει γενικά για όλα τα διακριτά κανάλια χωρίς μνήμη.

Παράδειγμα 6.1. Product DMC

(A. El Gamal & Y.-H. Kim, Example 3.3).

- Θεωρούμε δύο διακριτά κανάλια χωρίς μνήμη $(\mathcal{X}_1, p(y_1|x_1), \mathcal{Y}_1)$ και $(\mathcal{X}_2, p(y_2|x_2), \mathcal{Y}_2)$ με χωρητικότητες C_1 και C_2, αντίστοιχα.
- Έστω, τώρα, το κανάλι $(\mathcal{X}_1 \times \mathcal{X}_2, p(y_1|x_1)p(y_2|x_2), \mathcal{Y}_1 \times \mathcal{Y}_2)$ στο οποίο τα σύμβολα $x_1 \in \mathcal{X}_1$ και $x_2 \in \mathcal{X}_2$ στέλνονται ταυτόχρονα και παράλληλα και τα ληφθέντα σύμβολα ακολουθούν κατανομή $p(y_1, y_2|x_1, x_2) = p(y_1|x_1)p(y_2|x_2)$.

Παράδειγμα 6.1. Product DMC (2)

- Η χωρητικότητα του καναλιού-γινομένου ισούται με

$$\begin{aligned} C &= \max_{p(x_1,x_2)} I(X_1,X_2;Y_1;Y_2) \\ &= \max_{p(x_1,x_2)} \{I(X_1,X_2;Y_1) + I(X_1,X_2;Y_2|Y_1)\} \\ &= \max_{p(x_1,x_2)} \{I(X_1;Y_1) + I(X_2;Y_1|X_1) + \\ &\qquad I(X_2;Y_2|Y_1) + I(X_1;Y_2|X_2,Y_1)\} \\ &\overset{(a),(b)}{=} \max_{p(x_1,x_2)} \{I(X_1;Y_1) + I(X_2;Y_2)\}\,. \end{aligned}$$

(a) Οι X_1 και X_2 είναι ανεξάρτητες και η Y_1 είναι ανεξάρτητη της X_2. Επομένως, $X_2 \to X_1 \to Y_1$ και
$I(X_2;Y_1|X_1) = H(Y_1|X_1) - H(Y_1|X_1,X_2) = 0$.
(b) Με την ίδια συλλογιστική, $(X_1,Y_1) \to X_2 \to Y_2$ και
$I(X_1;Y_2|X_2,Y_1) = H(Y_2|X_2,Y_1) - H(Y_2|X_1,X_2,Y_1) = 0$. Επίσης, λόγω ανεξαρτησίας, $I(X_2;Y_2|Y_1) = I(X_2;Y_2)$.

Παράδειγμα 6.1. Product DMC (3)

- Συνεπώς,

$$\begin{aligned} C &= \max_{p(x_1,x_2)} \{I(X_1;Y_1) + I(X_2;Y_2)\} \\ &= \max_{p(x_1,x_2)} I(X_1;Y_1) + \max_{p(x_1,x_2)} I(X_2;Y_2) \\ &= \max_{p(x_1)} I(X_1;Y_1) + \max_{p(x_2)} I(X_2;Y_2) = C_1 + C_2. \end{aligned}$$

- Το αποτέλεσμα μπορεί να γενικευτεί για $K > 2$ ανεξάρτητα κανάλια.

Χωρητικότητα Συμμετρικού Καναλιού

- Πίνακας μετάβασης $[p(y|x)]_{i,j}$. Σύμβολο στην είσοδο: x_i. Σύμβολο στην έξοδο: y_j.
- Ορισμός 6.2. Ένα διακριτό κανάλι χωρίς μνήμη ονομάζεται συμμετρικό όταν κάθε γραμμή του πίνακα μετάβασης $p(y|x)$ προκύπτει από αναδιάταξη κάθε άλλης γραμμής και το ίδιο ισχύει και για κάθε στήλη του πίνακα. Ένα διακριτό κανάλι χωρίς μνήμη ονομάζεται ασθενώς συμμετρικό όταν κάθε γραμμή του πίνακα μετάβασης $p(y|x)$ προκύπτει από αναδιάταξη κάθε άλλης γραμμής και τα αθροίσματα των στοιχείων κάθε στήλης $\sum_x p(y|x)$ ισούνται μεταξύ τους.

Χωρητικότητα Συμμετρικού Καναλιού (2)

- Θεώρημα 6.3. Για τη χωρητικότητα ασθενώς συμμετρικού καναλιού (και, επομένως, και συμμετρικού καναλιού), ισχύει

$$C = \log|\mathcal{Y}| - H(\text{οποιασδήποτε γραμμής } \mathbf{r} \text{ πίνακα μετάβασης}).$$

- Απόδειξη:

$$\begin{aligned} I(X;Y) &= H(Y) - H(Y|X) \\ &\overset{(a)}{=} H(Y) - H(\mathbf{r}) \\ &\leq \log|\mathcal{Y}| - H(\mathbf{r}). \end{aligned}$$

- Για το (a) χρησιμοποιήθηκε το γεγονός ότι κάθε γραμμή του πίνακα μετάβασης προκύπτει από αναδιάταξη κάθε άλλης γραμμής.

Χωρητικότητα Συμμετρικού Καναλιού (3)

- Η ισότητα ισχύει όταν η Y ακολουθεί ομοιόμορφη κατανομή.
- Ομοιόμορφη κατανομή για την Y επιτυγχάνεται με χρήση ομοιόμορφα κατανεμημένης εισόδου X.

$$p(y) = \sum_{x \in \mathcal{X}} p(y|x)p(x) = \frac{1}{|\mathcal{X}|} \sum_{x \in \mathcal{X}} p(y|x) \overset{(a)}{=} c\frac{1}{|\mathcal{X}|} = \frac{1}{|\mathcal{Y}|}.$$

  Στο (a) χρησιμοποιήθηκε το γεγονός ότι, για (ασθενώς) συμμετρικά κανάλια, τα αθροίσματα των στοιχείων κάθε στήλης ισούνται μεταξύ τους.
- Παρόλο που για συμμετρικά (και ασθενώς συμμετρικά) κανάλια η χωρητικότητα επιτυγχάνεται πάντοτε με χρήση ομοιόμορφης κατανομής εισόδου, αυτό δε σημαίνει, κατ' ανάγκη, ότι μόνο η ομοιόμορφη κατανομή επιτυγχάνει τη χωρητικότητα.
  - Παράδειγμα: Ενθόρυβη γραφομηχανή με άρτιο αριθμό πλήκτρων.

Η Ιδιότητα Από Κοινού Ασυμπτωτικής Ισοδιαμέρισης (Joint AEP)

Εισαγωγή

- Θα αρχίσουμε, και πάλι, από την ασθενή από κοινού τυπικότητα και, στη συνέχεια, θα αναφερθούμε στην ισχυρή από κοινού τυπικότητα.

Από κοινού τυπικές ακολουθίες (Jointly Typical sequences)

- Ορισμός 6.4. Το σύνολο $A_\epsilon^{(n)}$ από κοινού (ασθενώς) τυπικών ακολουθιών ζευγών (x, y) ως προς την κατανομή $p(x, y)$, ορίζεται ως

$$A_\epsilon^{(n)} = \{(x^n, y^n) \in \mathcal{X}^n \times \mathcal{Y}^n : \left| -\frac{1}{n} \log p(x^n) - H(X) \right| < \epsilon, \\ \left| -\frac{1}{n} \log p(y^n) - H(Y) \right| < \epsilon, \\ \left| -\frac{1}{n} \log p(x^n, y^n) - H(X, Y) \right| < \epsilon \Big\},$$

περιέχει, δηλαδή, τις ακολουθίες ζευγών (ή τα ζεύγη ακολουθιών) $\{(x^n, y^n)\}$ μήκους n οι εμπειρικές εντροπίες των οποίων βρίσκονται σε απόσταση από την πραγματική τους εντροπία που δεν υπερβαίνει το ϵ.

Ιδιότητα Από Κοινού Ασυμπτωτικής Ισοδιαμέρισης (Joint AEP)

- Έστω (X^n, Y^n) ακολουθίες μήκους n οι οποίες δημιουργούνται με χρήση ανεξάρτητων και ομοίως κατανεμημένων (i.i.d.) ζευγών (X_i, Y_i), σύμφωνα με την κατανομή $p(x^n, y^n) = \prod_{i=1}^{n} p(x_i, y_i)$. Ισχύουν οι ιδιότητες:
  1. $\Pr\left\{(X^n, Y^n) \in A_\epsilon^{(n)}\right\} \to 1$, για $n \to \infty$.
  2. $\left|A_\epsilon^{(n)}\right| \leq 2^{n(H(X,Y)+\epsilon)}$.
  3. Εάν $(\tilde{X}^n, \tilde{Y}^n) \sim p(x^n)p(y^n)$, δηλαδή οι $\tilde{X}^n$ και $\tilde{Y}^n$ είναι ανεξάρτητες και οι κατανομές τους είναι ίδιες με τις περιθώριες κατανομές της $p(x^n, y^n)$,

$$\Pr\left\{(\tilde{X}^n, \tilde{Y}^n) \in A_\epsilon^{(n)}\right\} \leq 2^{-n(I(X;Y)-3\epsilon)}.$$

Επίσης, υπάρχει n_0 τέτοιο ώστε, για $n > n_0$,

$$\Pr\left\{(\tilde{X}^n, \tilde{Y}^n) \in A_\epsilon^{(n)}\right\} \geq (1-\epsilon)2^{-n(I(X;Y)+3\epsilon)}.$$

Ιδιότητα Από Κοινού Ασυμπτωτικής Ισοδιαμέρισης – Αποδείξεις

1. $\Pr\left\{(X^n, Y^n) \in A_\epsilon^{(n)}\right\} \to 1$, για $n \to \infty$.

   Από τον ασθενή νόμο των μεγάλων αριθμών, $-\frac{1}{n}\log p(X^n) \to -\mathbb{E}[\log p(X)] = H(X)$ κατά πιθανότητα. Επομένως, για δεδομένο $\epsilon > 0$, υπάρχει n_1 τέτοιο ώστε, για όλα τα $n > n_1$, $\Pr\left\{\left|-\frac{1}{n}\log p(X^n) - H(X)\right| \geq \epsilon\right\} < \frac{\epsilon}{3}$. Παρομοίως, υπάρχουν n_2 και n_3 τέτοια ωστε, $\Pr\left\{\left|-\frac{1}{n}\log p(Y^n) - H(Y)\right| \geq \epsilon\right\} < \frac{\epsilon}{3}$ και $\Pr\left\{\left|-\frac{1}{n}\log p(X^n, Y^n) - H(X,Y)\right| \geq \epsilon\right\} < \frac{\epsilon}{3}$, αντιστοίχως. Επομένως, για $n > \max\{n_1, n_2, n_3\}$, η πιθανότητα το (X^n, Y^n) να μην είναι τυπικό είναι μικρότερη από ϵ, και, συνεπώς, $\Pr\left\{(X^n, Y^n) \in A_\epsilon^{(n)}\right\} > 1 - \epsilon$, για $n > \max\{n_1, n_2, n_3\}$.

Ιδιότητα Από Κοινού Ασυμπτωτικής Ισοδιαμέρισης – Αποδείξεις (2)

2. $\left|A_\epsilon^{(n)}\right| \leq 2^{n(H(X,Y)+\epsilon)}$.

   Παρόμοια με την αντίστοιχη απόδειξη για το AEP,

$$1 = \sum p(x^n, y^n) \geq \sum_{(x^n,y^n)\in A_\epsilon^{(n)}} p(x^n, y^n) \geq \left|A_\epsilon^{(n)}\right| 2^{-n(H(X,Y)+\epsilon)}.$$

Ιδιότητα Από Κοινού Ασυμπτωτικής Ισοδιαμέρισης – Αποδείξεις (3)

3. Έστω *ανεξάρτητες* ακολουθίες τ.μ. $\tilde{X}^n$ και $\tilde{Y}^n$ που έχουν προκύψει από κατανομές $p(\tilde{x}^n)$ και $p(\tilde{y}^n)$ που είναι ίδιες με τις περιθώριες κατανομές της $p(x^n, y^n)$, $p(x^n)$ και $p(y^n)$, αντιστοίχως. Επομένως,

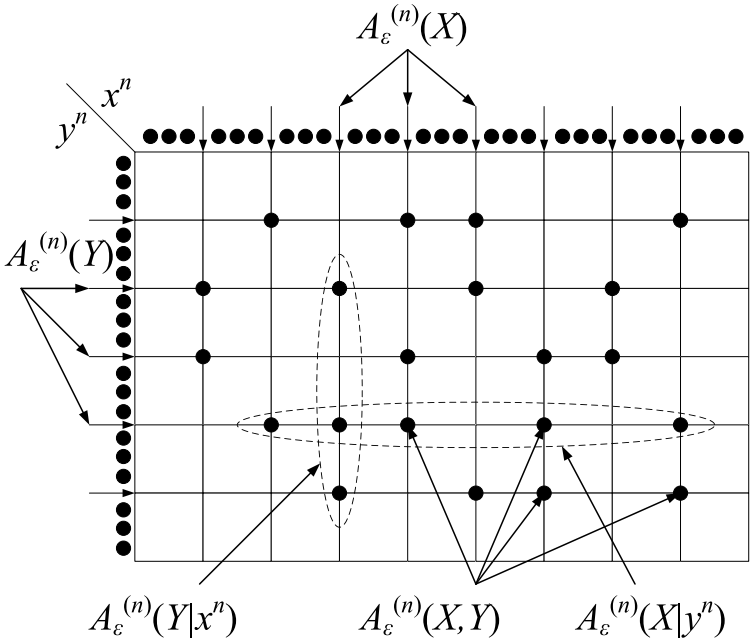
$$\begin{aligned}\Pr\left\{(\tilde{X}^n, \tilde{Y}^n) \in A_\epsilon^{(n)}\right\} &= \sum_{(x^n,y^n)\in A_\epsilon^{(n)}} p(x^n)p(y^n) \\ &\leq 2^{n(H(X,Y)+\epsilon)}2^{-n(H(X)-\epsilon)}2^{-n(H(Y)-\epsilon)} \\ &= 2^{-n(I(X;Y)-3\epsilon)}.\end{aligned}$$

Με παρόμοιο τρόπο (βλ. π.χ. Cover Theorem 7.6.1) μπορεί να αποδειχτεί ότι

$$\Pr\left\{(\tilde{X}^n, \tilde{Y}^n) \in A_\epsilon^{(n)}\right\} \geq (1-\epsilon)2^{-n(I(X;Y)+3\epsilon)}$$

για $n > n_0$.

Ιδιότητα Από Κοινού Ασυμπτωτικής Ισοδιαμέρισης (2)



Στο σχήμα δίνεται ένα παράδειγμα από κοινού τυπικού συνόλου. Υπάρχουν περίπου $2^{nH(X)}$ τυπικές ακολουθίες τ.μ. X και περίπου $2^{nH(Y)}$ τυπικές ακολουθίες τ.μ. Y. Ωστόσο, οι από κοινού τυπικές ακολουθίες είναι περίπου $2^{nH(X,Y)}$, δηλαδή, υπάρχουν ζεύγη τυπικών X^n με τυπικά Y^n τα οποία *δεν* είναι από κοινού τυπικά.

Ιδιότητα Από Κοινού Ασυμπτωτικής Ισοδιαμέρισης (3)

- Από την 3η ιδιότητα, η πιθανότητα ένα ζεύγος ακολουθιών (X^n, Y^n) το οποίο επιλέγεται τυχαία και του οποίου οι συνιστώσες είναι (με-μονωμένως) τυπικές να είναι και από κοινού τυπικό, ισούται περίπου με $2^{-nI(X;Y)}$.
- Επομένως, στο σχήμα της προηγούμενης διαφάνειας, κατά μέσο όρο πρέπει να θεωρήσουμε περίπου $2^{nI(X;Y)}$ ζεύγη μεμονωμένως τυπικών X^n και Y^n έως ότου εμφανιστεί ένα τυπικό ζεύγος.
- Ισοδύναμα, εάν θεωρήσουμε μια ακολουθία Y^n η οποία αποτελεί την έξοδο καναλιού με είσοδο X^n, υπάρχουν περίπου $2^{nH(X|Y)}$ υπό συνθήκη τυπικές ακολουθίες X^n. Η πιθανότητα να διαλέξουμε μια ακολουθία X'^n η οποία είναι τυπική με την Y^n αλλά δεν είναι η ακολουθία X^n η οποία μεταδόθηκε ισούται, περίπου, με $2^{nH(X|Y)}/2^{nH(X)} = 2^{-nI(X;Y)}$. Επομένως, και πάλι, κατά μέσο όρο πρέπει να θεωρήσουμε περίπου $2^{nI(X;Y)}$ ακολουθίες X^n έως ότου εμφανιστεί ακολουθία που αποτελεί τυπικό ζεύγος με την Y^n.

Ιδιότητα Από Κοινού Ασυμπτωτικής Ισοδιαμέρισης (4)

- Συνεπώς, διαισθητικά, αν έχουμε ένα κανάλι με $p(y|x)$ μπορούμε να μεταδίδουμε περίπου $2^{nI(X;Y)}$ διακριτές ακολουθίες στο κανάλι χωρίς στο δέκτη να υπάρχει σύγχυση της (τυπικής) ακολουθίας y^n που αντιστοιχεί στη μεταδοθείσα τυπική ακολουθία x^n με μια άλλη τυπική ακολουθία $\tilde{y}^n$ η οποία είναι ανεξάρτητη από τη μεταδοθείσα x^n.
- Θα αποδείξουμε ότι είναι εφικτή η μετάδοση έως και $2^{nI(X;Y)}$ διακριτών ακολουθιών με αυθαίρετα μικρή πιθανότητα σφάλματος για $n \to \infty$. Θα δούμε, επίσης, ότι εάν προσπαθήσουμε να μεταδώσουμε περισσότερες από $2^{nI(X;Y)}$ διακριτές ακολουθίες, η πιθανότητα σφάλματος τείνει στο 1.

Ιδιότητα Από Κοινού Ασυμπτωτικής Ισοδιαμέρισης (5)

- Ωστόσο, χρειάζεται, επίσης, να αποδείξουμε ότι η ακολουθία, Y^n, στην έξοδο του καναλιού και η ακολουθία, x^n, που εφαρμόσαμε στην είσοδό του είναι από κοινού τυπικές.
- Για το σκοπό αυτό δεν επαρκεί η ασθενής τυπικότητα. Χρειαζόμαστε το Λήμμα Υπό Συνθήκη Τυπικότητας (Conditional Typicality Lemma) το οποίο αποδεικνύεται με χρήση Ισχυρής Τυπικότητας.
- Περισσότερα σε επόμενη διαφάνεια.

Διακριτά κανάλια και χωρητικότητα (συνέχεια)

Από Κοινού Ισχυρώς Τυπικές Ακολουθίες

- Οι ισχυρώς τυπικές ακολουθίες δύο τ.μ. ορίζονται με αντίστοιχο τρόπο όπως και οι ισχυρώς τυπικές ακολουθίες μίας τ.μ.
- Ορισμός 6.5. Έστω i.i.d. ακολουθία ζευγών τ.μ. που παίρνουν τιμές σε πεπερασμένο σύνολο $\mathcal{X} \times \mathcal{Y}$. Στη γενική περίπτωση, οι τ.μ. που αποτελούν το ζεύγος είναι εξαρτημένες: $(X, Y) \sim p(x, y)$. Το σύνολο $\mathcal{T}_\epsilon^{(n)}(X, Y)$ των από κοινού ισχυρώς[1] $\epsilon-$τυπικών ακολουθιών (x^n, y^n) ορίζεται ως

$$\mathcal{T}_\epsilon^{(n)}(X, Y) \triangleq \{(x^n, y^n) : |\pi(x, y|x^n, y^n) - p(x, y)| \leq \epsilon p(x, y) \text{ για όλα τα } (x, y) \in (\mathcal{X}, \mathcal{Y})\},$$

όπου $\pi(x, y|x^n, y^n)$ είναι ο από κοινού τύπος (η από κοινού εμπειρική πιθανότητα) του ζεύγους (x, y).

[1] Υπενθυμίζεται ότι χρησιμοποιούμε σθεναρή τυπικότητα.

Ιδιότητες Ισχυρώς Τυπικών Ακολουθιών

- Έστω $p(x^n, y^n) \sim \prod_{i=1}^n p_{X,Y}(x_i, y_i)$. Εάν $(x^n, y^n) \in \mathcal{T}_\epsilon^{(n)}(X, Y)$,
  1. $x^n \in \mathcal{T}_\epsilon^{(n)}(X)$ και $y^n \in \mathcal{T}_\epsilon^{(n)}(Y)$.
  2. $p(x^n, y^n) \doteq 2^{-nH(X,Y)}$. ($a_n \doteq 2^{nb}$ εάν υπάρχει $\delta(\epsilon)$ τέτοιο ώστε $2^{n(b-\delta(\epsilon))} \leq a_n \leq 2^{n(b+\delta(\epsilon))}$)
  3. $p(x^n) \doteq 2^{-nH(X)}$ και $p(y^n) \doteq 2^{-nH(Y)}$.
  4. $p(x^n|y^n) \doteq 2^{-nH(X|Y)}$ και $p(y^n|x^n) \doteq 2^{-nH(Y|X)}$ (γιατί;)
- Για αρκούντως μεγάλο n,

$$|\mathcal{T}_\epsilon^{(n)}(X, Y)| \doteq 2^{nH(X,Y)}.$$

Ιδιότητες Ισχυρώς Τυπικών Ακολουθιών (2)

- Ορισμός 6.6. Το σύνολο $\mathcal{T}_\epsilon^{(n)}(Y|x^n)$ των ακολουθιών Y^n που είναι *υπό συνθήκη* ισχυρώς $\epsilon-$τυπικές (conditionally ϵ-typical) με την ακολουθία x^n ορίζεται ως

$$\mathcal{T}_\epsilon^{(n)}(Y|x^n) \triangleq \left\{ y^n : (x^n, y^n) \in \mathcal{T}_\epsilon^{(n)}(X,Y) \right\}.$$

- Αποδεικνύεται ότι για οποιαδήποτε ακολουθία x^n (όχι κατ' ανάγκη τυπική)

$$|\mathcal{T}_\epsilon^{(n)}(Y|x^n)| \leq 2^{n(H(Y|X)+\delta(\epsilon))},$$

όπου $\delta(\epsilon) = \epsilon \cdot H(Y|X)$.

Conditional Typicality Lemma

Λήμμα 6.7. (υπό συνθήκη τυπικότητας)

Conditional Typicality Lemma

Έστω $x^n \in \mathcal{T}_{\epsilon'}^{(n)}(X)$ και $Y^n \sim \prod_{i=1}^{n} p_{Y|X}(y_i|x_i)$.
Για οποιοδήποτε $\epsilon > \epsilon'$,

$$\Pr\left\{(x^n, Y^n) \in \mathcal{T}_{\epsilon}^{(n)}(X,Y)\right\} \to 1 \text{ για } n \to \infty.$$

Για την απόδειξη δείτε El Gamal & Kim, Chapter 2.

Παρατηρήστε ότι τα Y_i είναι i.i.d. δεδομένων των X_i (όπως και στα κανάλια χωρίς μνήμη).

Conditional Typicality Lemma (2)

- Επομένως, για όλες τις ακολουθίες $x^n \in \mathcal{T}_{\epsilon'}^{(n)}(X)$ με $\epsilon' < \epsilon$,

$$|\mathcal{T}_{\epsilon}^{(n)}(Y|x^n)| \geq (1-\epsilon)2^{n(H(Y|X)-\delta(\epsilon))},$$

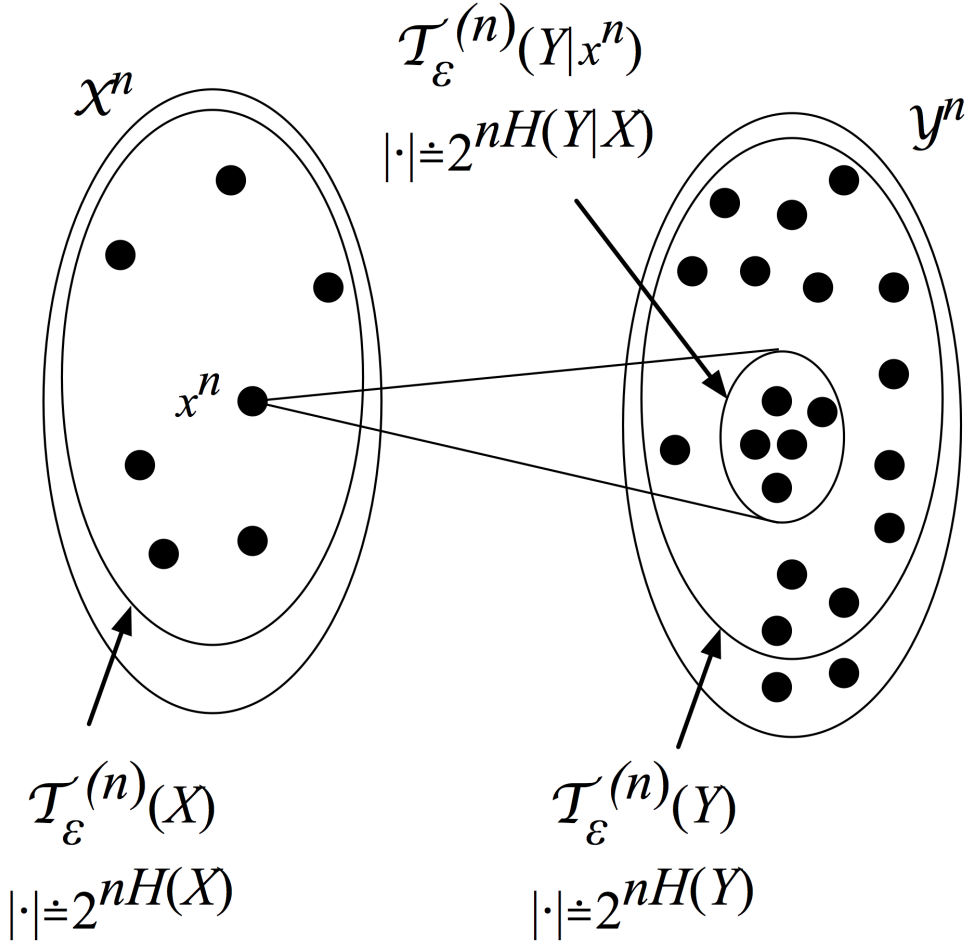
για αρκούντως μεγάλο n.

- Αποδεικνύεται, επίσης, ότι, για κάθε $x^n \in \mathcal{T}_{\epsilon}^{(n)}(X)$ και για αρκούντως μεγάλο n,

$$|\mathcal{T}_{\epsilon}^{(n)}(Y|x^n)| \geq 2^{n(H(Y|X)-\delta'(\epsilon))},$$

για κάποιο $\delta'(\epsilon) \to 0$ καθώς $\epsilon \to 0$.

Conditional Typicality Lemma (3)



Γενίκευση για περισσότερες τ.μ.

- Όλα τα παραπάνω γενικεύονται για περισσότερες τ.μ.
- Θα αναφερθούμε μόνο στο Λήμμα Από Κοινού Τυπικότητας (Joint Typicality Lemma).
- Για περισσότερες λεπτομέρειες δείτε, για παράδειγμα, A. El Gamal & Y.-H. Kim, *Network Information Theory*.

Λήμμα Από Κοινού Τυπικότητας

Joint Typicality Lemma

Λήμμα 6.8. (Από Κοινού Τυπικότητας)

Έστω $(X, Y, Z) \sim p(x, y, z)$. Υπάρχει $\delta(\epsilon) \to 0$ καθώς $\epsilon \to 0$ τέτοιο ώστε να ισχύουν τα παρακάτω:

1. Εάν $(\tilde{x}^n, \tilde{y}^n)$ είναι ζεύγος ακολουθιών και η ακολουθία $\tilde{Z}^n$ ακολουθεί κατανομή $\prod_{i=1}^n p_{Z|X}(\tilde{z}_i|\tilde{x}_i)$ (δηλαδή είναι ανεξάρτητη της $\tilde{y}^n$ δεδομένης της $\tilde{x}^n$),
$$\Pr\left\{(\tilde{x}^n, \tilde{y}^n, \tilde{Z}^n) \in \mathcal{T}_\epsilon^{(n)}(X, Y, Z)\right\} \leq 2^{-n(I(Y;Z|X)-\delta(\epsilon))}$$
2. Εάν $(x^n, y^n) \in \mathcal{T}_{\epsilon'}^{(n)}(X, Y)$, $\epsilon' < \epsilon$ και $\tilde{Z}^n \sim \prod_{i=1}^n p_{Z|X}(\tilde{z}_i|x_i)$, τότε, για αρκούντως μεγάλο n,
$$\Pr\left\{(x^n, y^n, \tilde{Z}^n) \in \mathcal{T}_\epsilon^{(n)}(X, Y, Z)\right\} \geq 2^{-n(I(Y;Z|X)+\delta(\epsilon))}.$$

Λήμμα Από Κοινού Τυπικότητας – Απόδειξη

Για την 1η σχέση,

$$\begin{aligned}
&\Pr\left\{\left(\tilde{x}^n, \tilde{y}^n, \tilde{Z}^n\right) \in \mathcal{T}_\epsilon^{(n)}(X, Y, Z)\right\} \\
&= \sum_{\tilde{z}^n \in \mathcal{T}_\epsilon^{(n)}(Z|\tilde{x}^n, \tilde{y}^n)} p(\tilde{z}^n|\tilde{x}^n) \\
&\overset{(a)}{\leq} \left|\mathcal{T}_\epsilon^{(n)}(Z|\tilde{x}^n, \tilde{y}^n)\right| \cdot 2^{-n(H(Z|X)-\epsilon H(Z|X))} \\
&\overset{(b)}{\leq} 2^{n(H(Z|X,Y)+\epsilon H(Z|X,Y))} 2^{-n(H(Z|X)-\epsilon H(Z|X))} \\
&\leq 2^{-n(I(Y;Z|X)-\delta(\epsilon))}.
\end{aligned}$$

$(a), (b)$ από τις ιδιότητες των ισχυρώς τυπικών ακολουθιών
Παρατηρήστε ότι δεν απαιτείται τυπικότητα του ζεύγους $(\tilde{x}^n, \tilde{y}^n)$.

Λήμμα Από Κοινού Τυπικότητας – Απόδειξη (2)

Για τη 2η σχέση, για αρκούντως μεγάλο n,

$$\begin{aligned}
&\Pr\left\{\left(x^n, y^n, \tilde{Z}^n\right) \in \mathcal{T}_\epsilon^{(n)}(X,Y,Z)\right\} \\
&= \sum_{\tilde{z}^n \in \mathcal{T}_\epsilon^{(n)}(Z|x^n,y^n)} p(\tilde{z}^n|x^n) \\
&\overset{(a)}{\geq} \left|\mathcal{T}_\epsilon^{(n)}(Z|x^n,y^n)\right| \cdot 2^{-n(H(Z|X)+\epsilon H(Z|X))} \\
&\overset{(b)}{\geq} (1-\epsilon)2^{n(H(Z|X,Y)-\epsilon H(Z|X,Y))}2^{-n(H(Z|X)+\epsilon H(Z|X))} \\
&\overset{(c)}{\geq} 2^{-n(I(Y;Z|X)+\delta(\epsilon))}.
\end{aligned}$$

(a), (b) από τις ιδιότητες των ισχυρώς τυπικών ακολουθιών. (c) γιατί; Παρατηρήστε ότι για το (b) απαιτείται τυπικότητα του ζεύγους (x^n, y^n) και για τα (b) και (c) απαιτείται αρκούντως μεγάλο n.

Τυπικές ακολουθίες πολλών μεταβλητών

- Με τον ίδιο τρόπο μπορούμε να ορίσουμε τυπικές ακολουθίες περισσότερων από 3 τ.μ.
- Οι ιδιότητες είναι παρόμοιες.
- Δείτε π.χ. Cover & Thomas ή El Gamal & Kim.

Το Θεώρημα Κωδικοποίησης Καναλιού

Θεώρημα Κωδικοποίησης Καναλιού – εισαγωγή

- Το Θεώρημα Κωδικοποίησης Καναλιού (Channel Coding Theorem) αποτελεί το πιο βασικό και το πιο διάσημο αποτέλεσμα της Θεωρίας Πληροφορίας.
- Σύμφωνα με το Θεώρημα Κωδικοποίησης Καναλιού, είναι εφικτή η μετάδοση σε κανάλια χωρίς μνήμη με ρυθμό αυθαίρετα κοντά στη χωρητικότητα και με αυθαίρετα μικρή πιθανότητα σφάλματος. Αντιστρόφως, δεν είναι εφικτή μετάδοση με αυθαίρετα μικρή πιθανότητα σφάλματος εάν ο ρυθμός μετάδοσης υπερβαίνει τη χωρητικότητα του καναλιού.

Θεώρημα Κωδικοποίησης Καναλιού – εισαγωγή (2)

- Στη συνέχεια, θα διατυπώσουμε με την απαραίτητη λεπτομέρεια και θα αποδείξουμε το Θεώρημα Κωδικοποίησης Καναλιού.
- Το Θεώρημα Κωδικοποίησης Καναλιού (ευθύ - achievability) μπορεί να αποδειχτεί είτε με χρήση αποκωδικοποίησης Μέγιστης Πιθανοφάνειας (Maximum Likelihood decoding -- Gallager) ή με χρήση Από Κοινού Τυπικών ακολουθιών (Cover).
- Στο μάθημα θα εξετάσουμε την απόδειξη με χρήση Από Κοινού Τυπικότητας η οποία είναι μάλλον πιο απλή.
- Επίσης, ο τρόπος αυτός απόδειξης έχει χρησιμοποιηθεί για να βρεθεί η χωρητικότητα άλλων καναλιών (π.χ. πολλών χρηστών).
- Το (ασθενές) αντίστροφο του Θεωρήματος Κωδικοποίησης Καναλιού θα αποδειχτεί με χρήση της ανισότητας Fano.

Θεώρημα Κωδικοποίησης Καναλιού – εισαγωγή (3)

- Το βασικό ερώτημα (και, εκ πρώτης όψεως, παράδοξο) είναι το εξής: Πώς είναι δυνατόν να μεταδώσουμε με αυθαίρετα μικρή πιθανότητα σφάλματος σε ένα κανάλι που εισάγει σφάλματα με μη μηδενική πιθανότητα και με τυχαίο τρόπο;
- Για να απαντήσει στο ερώτημα, ο Shannon χρησιμοποίησε ένα διαφορετικό τρόπο σκέψης:
  - Δεν προσπάθησε να εκμηδενίσει την πιθανότητα σφάλματος, απλώς να την περιορίσει σε αυθαίρετα μικρές τιμές.
  - Βασίστηκε σε πολλές διαδοχικές χρήσεις του καναλιού ώστε να εκμεταλλευτεί το Νόμο των Μεγάλων Αριθμών.
  - Χρησιμοποίησε κώδικες οι οποίοι δημιουργούνται τυχαία και υπολόγισε τη μέση πιθανότητα σφάλματος.
- Αυτός ο τρόπος σκέψης διέπει τόσο την απόδειξη με χρήση τυπικότητας όσο και την απόδειξη με αποκωδικοποίηση Μέγιστης Πιθανοφάνειας.

Θεώρημα Κωδικοποίησης Καναλιού – Ορισμοί

Ορισμός 6.9. Ένας κώδικας (M, n) για το Διακριτό Κανάλι Χωρίς Μνήμη $(\mathcal{X}, p(y|x), \mathcal{Y})$ αποτελείται από

1. Ένα σύνολο δεικτών (μηνυμάτων) $\{1, 2, \ldots, M\}$.
2. Μια συνάρτηση κωδικοποίησης $x^n : \{1, 2, \ldots, M\} \to \mathcal{X}^n$ η οποία αντιστοιχίζει την κωδική λέξη (codeword) $x^n(m)$ στο μήνυμα $m \in \{1, 2, \ldots, M\}$. Το σύνολο των κωδικών λέξεων ονομάζεται βιβλίο κωδίκων (codebook).
3. Μια συνάρτηση αποκωδικοποίησης $g : \mathcal{Y}^n \to \{1, 2, \ldots, M\} \cup \{e\}$, η οποία αποτελεί ένα νομοτελειακό κανόνα ο οποίος αντιστοιχίζει ένα εκτιμώμενο δείκτη μεταδοθέντος μηνύματος, $\hat{m}$, ή το μήνυμα σφάλματος e, σε κάθε ληφθείσα ακολουθία y^n.

Θεώρημα Κωδικοποίησης Καναλιού – Ορισμοί (2)

- Υπό συνθήκη πιθανότητα σφάλματος δεδομένου ότι εστάλη το μήνυμα με δείκτη m:

$$\begin{aligned}\lambda_m &= \Pr\{g(Y^n) \neq m | X^n = x^n(m)\} \\ &= \sum_{y^n} p(y^n|x^n(m)) I(g(y^n) \neq m),\end{aligned}$$

όπου $I(\cdot)$ η συνάρτηση δείκτης (ισούται με 1 όταν το όρισμά της αληθεύει, αλλιώς με 0).

- Η Μέγιστη Πιθανότητα Σφάλματος $\lambda^{(n)}$ κώδικα (M, n) ορίζεται ως

$$\lambda^{(n)} = \max_{m \in \{1,2,\ldots,M\}} \lambda_m.$$

Θεώρημα Κωδικοποίησης Καναλιού – Ορισμοί (3)

- Η (αριθμητικώς) μέση πιθανότητα σφάλματος $P_e^{(n)}$ κώδικα (M, n) ισούται με

$$P_e^{(n)} = \frac{1}{M} \sum_{m=1}^{M} \lambda_m.$$

- Όταν ο δείκτης μηνύματος W ακολουθεί ομοιόμορφη κατανομή, $P_e^{(n)} = \Pr\{W \neq g(Y^n)\}$, όπου Y^n η ακολουθία που λαμβάνεται στην έξοδο καναλιού όπου έχει μεταδοθεί η ακολουθία $X^n = x^n(W)$.
- Επίσης, $P_e^{(n)} \leq \lambda^{(n)}$.

Θεώρημα Κωδικοποίησης Καναλιού – Ορισμοί (4)

- Ορισμός 6.10. Ο ρυθμός (rate) R κώδικα (M, n) ισούται με

$$R = \frac{\log M}{n} \text{ bits ανά μετάδοση.}$$

- Ορισμός 6.11. Ένας ρυθμός R είναι εφικτός (achievable) όταν υπάρχει ακολουθία κωδίκων $\left(\lceil 2^{nR} \rceil, n\right)$ για την οποία η μέγιστη πιθανότητα σφάλματος $\lambda^{(n)}$ τείνει στο 0 καθώς το n τείνει στο άπειρο.[2]
- Ορισμός 6.12. Η χωρητικότητα λειτουργίας (operational capacity) ενός καναλιού ισούται με το μέγιστο ρυθμό ο οποίος είναι εφικτός.
  - Το Θεώρημα Κωδικοποίησης Πηγής αποδεικνύει ότι η χωρητικότητα λειτουργίας $\max_R$ εφικτός R ισούται με την πληροφοριακή χωρητικότητα $\max_{p(x)} I(X;Y)$.

[2] Συχνά η επιτευξιμότητα ορίζεται με χρήση της μέσης πιθανότητας σφάλματος, $P_e^{(n)}$. Οι δύο ορισμοί *δεν* είναι ισοδύναμοι.

Το Θεώρημα Κωδικοποίησης Καναλιού

Απόδειξη Θεωρήματος Κωδικοποίησης Καναλιού – Εισαγωγή

- Θα αναφερθούμε στην απόδειξη η οποία χρησιμοποιεί την Ιδιότητα Από Κοινού Ασυμπτωτικής Ισοδιαμέρισης (Joint AEP).
- Η ιδέα:
  - Δεδομένου του μηνύματος W, στέλνουμε στο κανάλι ακολουθία $X^n = x^n(W)$ μήκους n.
  - Στην έξοδο του καναλιού λαμβάνουμε ακολουθία Y^n η οποία εξαρτάται από τη X^n, καθώς και από τον πίνακα μετάβασης, $p(y|x)$, του καναλιού.
  - Στο δέκτη αναζητούμε ακολουθία $\hat{x}^n$ η οποία ανήκει στο βιβλίο κωδίκων και είναι *από κοινού τυπική* με την Y^n. Εάν υπάρχει, ο δέκτης θεωρεί ότι η $\hat{x}^n$ είναι η ακολουθία που μετέδωσε ο πομπός.
  - Από την Ιδιότητα από κοινού Ασυμπτωτικής Ισοδιαμέρισης, με μεγάλη πιθανότητα η ληφθείσα ακολουθία θα είναι από κοινού τυπική με τη μεταδοθείσα.
  - Ωστόσο, υπάρχει η πιθανότητα η Y^n να μην είναι από κοινού τυπική με καμία από τις κωδικές λέξεις x^n ή να είναι από κοινού τυπική με άλλη ακολουθία από αυτή που μεταδόθηκε. Στην περίπτωση αυτή εμφανίζεται σφάλμα μετάδοσης.
  - Θα αποδείξουμε ότι, εάν $R < C$, καθώς το n τείνει στο άπειρο, η πιθανότητα σφάλματος τείνει στο 0.

Απόδειξη Θεωρήματος Κωδικοποίησης Καναλιού

- Θεώρημα 6.13. (Θεώρημα Κωδικοποίησης Καναλιού)
  - Σε ένα Διακριτό Κανάλι Χωρίς Μνήμη, όλοι οι ρυθμοί οι οποίοι είναι μικρότεροι από την πληροφοριακή χωρητικότητα είναι εφικτοί. Δηλαδή, για κάθε ρυθμό $R < C$, υπάρχει ακολουθία κωδίκων $(\lceil 2^{nR} \rceil, n)$ με μέγιστη πιθανότητα σφάλματος $\lambda^{(n)} \to 0$.
  - Αντιστρόφως, για οποιαδήποτε ακολουθία από κώδικες $(\lceil 2^{nR} \rceil, n)$ με $\lambda^{(n)} \to 0$ πρέπει να ισχύει $R \leq C$.
- Απόδειξη (ευθύ).

  Για απλοποίηση και χωρίς απώλεια της γενικότητας υποθέτουμε ότι ο αριθμός κωδικών λέξεων $\lceil 2^{nR} \rceil$ είναι ακέραιος.

  I. Δημιουργία Τυχαίου Βιβλίου Κωδίκων: Δημιουργούμε 2^{nR} τυχαίες και ανεξάρτητες ακολουθίες i.i.d. $x^n \sim \prod_{i=1}^{n} p_X^*(x_i)$, όπου $p_X^*(x)$ η κατανομή που μεγιστοποιεί την $I(X;Y)$.

Απόδειξη Θεωρήματος Κωδικοποίησης Καναλιού (2)

- Οι 2^{nR} κωδικές λέξεις αποτελούν τις γραμμές του πίνακα

$$\mathcal{C} = \begin{bmatrix} x_1(1) & x_2(1) & \dots & x_n(1) \\ x_1(2) & x_2(2) & \dots & x_n(2) \\ \vdots & \vdots & \ddots & \vdots \\ x_1(2^{nR}) & x_2(2^{nR}) & \dots & x_n(2^{nR}) \end{bmatrix}$$

- Η πιθανότητα να δημιουργηθεί ένας συγκεκριμένος τυχαίος κώδικας (πίνακας) $\mathcal{C}$ ισούται με $\Pr(\mathcal{C}) = \prod_{m=1}^{2^{nR}} \prod_{i=1}^{n} p(x_i(m))$.
- Ο κώδικας ανακοινώνεται στον πομπό και στο δέκτη. Επίσης, τόσο ο πομπός όσο και ο δέκτης γνωρίζουν τον πίνακα μετάβασης του καναλιού, $p(y|x)$.

Απόδειξη Θεωρήματος Κωδικοποίησης Καναλιού (3)

II. Κωδικοποίηση: Προκειμένου να στείλει το μήνυμα $m \in \{1, 2, \ldots, 2^{nR}\}$ ο πομπός μεταδίδει την κωδική λέξη $x^n(m)$ (τη $m-$στή γραμμή του πίνακα $\mathcal{C}$).

III. Διέλευση από το κανάλι: Ο δέκτης λαμβάνει ακολουθία y^n με υπό συνθήκη κατανομή $p(y^n|x^n(m)) = \prod_{i=1}^{n} p(y_i|x_i(m))$.

IV. Αποκωδικοποίηση: Ο δέκτης εκτιμά ποιο μήνυμα έχει σταλεί από τον πομπό. Ο βέλτιστος δέκτης χρησιμοποιεί ανίχνευση Μέγιστης Πιθανοφάνειας (αν θεωρήσουμε ομοιόμορφη κατανομή μηνυμάτων). Ωστόσο, όπως αναφέρθηκε, για την απόδειξη θα θεωρήσουμε ανίχνευση με βάση την από κοινού τυπικότητα.

Απόδειξη Θεωρήματος Κωδικοποίησης Καναλιού (4)

Παρόλο που η αποκωδικοποίηση με χρήση από κοινού τυπικότητας δεν είναι βέλτιστη, θα αποδείξουμε ότι, και σε αυτήν την περίπτωση, $\lambda^{(n)} \to 0$ για $n \to \infty$ (ο δέκτης είναι *ασυμπτωτικά* βέλτιστος). Ο δέκτης αποφασίζει (εκτιμά) ότι εστάλη το μήνυμα $\hat{m} \in \{1, 2, \ldots, 2^{nR}\}$ που ικανοποιεί ταυτόχρονα τις εξής δύο συνθήκες:

1. Το ζεύγος ακολουθιών $(x^n(\hat{m}), y^n)$ είναι από κοινού τυπικό.
2. Δεν υπάρχει άλλος δείκτης μηνύματος $m' \neq \hat{m}$ για τον οποίο να ισχύει $(x^n(m'), y^n) \in A_\epsilon^{(n)}$. Δηλαδή, δεν υπάρχει άλλη κωδική λέξη εκτός από αυτή που έχει πραγματικά σταλεί η οποία να είναι από κοινού τυπική με την y^n.

Εάν $\hat{m} \neq m$, εμφανίζεται σφάλμα ανίχνευσης. Έστω $\mathcal{E}$ το ενδεχόμενο $\{\hat{m} \neq m\}$. Σε αυτό περιλαμβάνεται και το ενδεχόμενο $\hat{m} = e$.

Απόδειξη Θεωρήματος Κωδικοποίησης Καναλιού (5)
Ανάλυση της πιθανότητας σφάλματος – Εισαγωγή

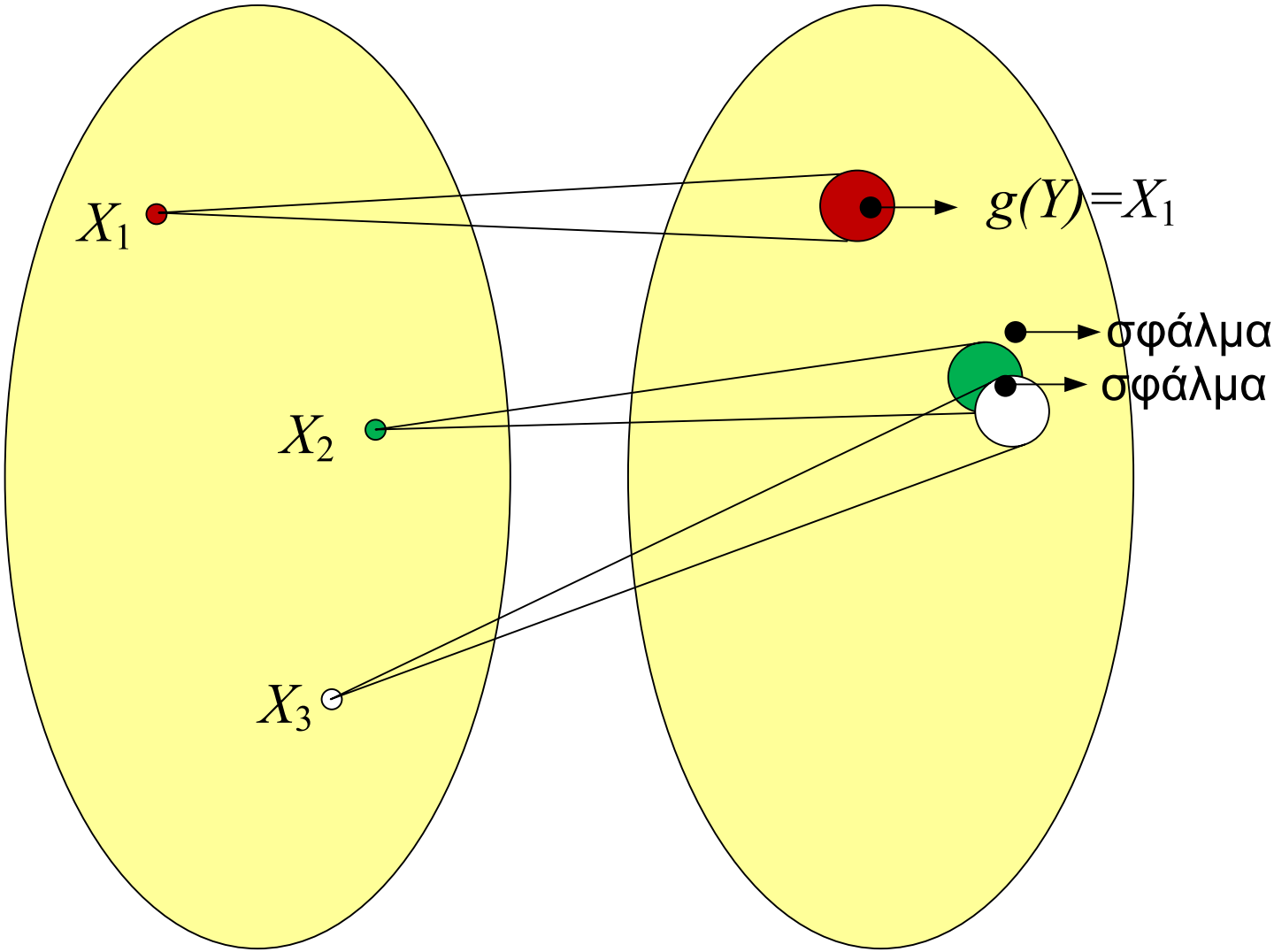
V. Ανάλυση του Σφάλματος Ανίχνευσης:

- Η ιδέα: Αντί να υπολογίσουμε την πιθανότητα σφάλματος για ένα συγκεκριμένο κώδικα, θα υπολογίσουμε τη μέση πιθανότητα σφάλματος για τυχαία δημιουργία κωδίκων και για *όλους* τους πιθανούς κώδικες.
- Όταν χρησιμοποιείται αποκωδικοποίηση με χρήση από κοινού τυπικότητας, υπάρχουν δύο πηγές σφάλματος: Είτε η έξοδος y^n δεν είναι από κοινού τυπική με την ακολουθία που εκπέμπει ο πομπός ή υπάρχει τουλάχιστον μια ακόμα κωδική λέξη η οποία είναι από κοινού τυπική με την y^n.

Απόδειξη Θεωρήματος Κωδικοποίησης Καναλιού (6)
Ανάλυση της πιθανότητας σφάλματος – Εισαγωγή

- Από την Ιδιότητα Από Κοινού Ασυμπτωτικής Ισοδιαμέρισης, η πιθανότητα η ληφθείσα ακολουθία να είναι από κοινού τυπική με την εκπεμφθείσα τείνει στο 1 για $n \to \infty$. Επίσης, η πιθανότητα η ληφθείσα ακολουθία να είναι από κοινού τυπική με ακολουθία διαφορετική από την εκπεμφθείσα ισούται περίπου με $2^{-nI(X;Y)}$. Επομένως, μπορούμε να χρησιμοποιήσουμε περίπου $2^{nI(X;Y)}$ κωδικές λέξεις και, ταυτόχρονα, να διασφαλίσουμε μικρή πιθανότητα σφάλματος.
- Στη συνέχεια θα αποδείξουμε τα παραπάνω και με την απαραίτητη μαθηματική αυστηρότητα.

Αποκωδικοποίηση με χρήση από κοινού τυπικότητας



Απόδειξη Θεωρήματος Κωδικοποίησης Καναλιού (7)
Υπολογισμός Πιθανότητας Σφάλματος (I)

- Θα υπολογίσουμε τη μέση πιθανότητα σφάλματος για *όλα* τα πιθανά βιβλία κωδίκων.

$$\Pr\{\mathcal{E}\} = \sum_{\mathcal{C}} \Pr(\mathcal{C}) P_e^{(n)}(\mathcal{C}) =$$

$$\overset{(a)}{=} \sum_{\mathcal{C}} \Pr(\mathcal{C}) \frac{1}{2^{nR}} \sum_{m=1}^{2^{nR}} \lambda_m(\mathcal{C}) = \frac{1}{2^{nR}} \sum_{m=1}^{2^{nR}} \sum_{\mathcal{C}} \Pr(\mathcal{C}) \lambda_m(\mathcal{C}).$$

(a) από τον ορισμό της $P_e^{(n)}(\mathcal{C})$.

Απόδειξη Θεωρήματος Κωδικοποίησης Καναλιού (8)
Υπολογισμός Πιθανότητας Σφάλματος (II)

- Δεδομένου ότι η αντιστοίχιση μηνυμάτων σε κωδικές λέξεις γίνεται τυχαία και επειδή για όλους τους πιθανούς κώδικες το μήνυμα m θα αντιστοιχίζεται κάθε φορά σε διαφορετική κωδική λέξη, η ποσότητα $\sum_{\mathcal{C}} \Pr(\mathcal{C})\lambda_m(\mathcal{C})$ είναι ανεξάρτητη του μηνύματος m. Επομένως, μπορούμε να υποθέσουμε, χωρίς απώλεια της γενικότητας, ότι εστάλη η κωδική λέξη με δείκτη $m = 1$.
- Συνεπώς, η $\Pr(\mathcal{E})$ ισούται με

$$\begin{aligned}\Pr\{\mathcal{E}\} &= \frac{1}{2^{nR}} \sum_{m=1}^{2^{nR}} \sum_{\mathcal{C}} \Pr(\mathcal{C})\lambda_m(\mathcal{C}) \\ &= \sum_{\mathcal{C}} \Pr(\mathcal{C})\lambda_1(\mathcal{C}) = \Pr(\mathcal{E}|M = 1).\end{aligned}$$

Απόδειξη Θεωρήματος Κωδικοποίησης Καναλιού (9)
Υπολογισμός Πιθανότητας Σφάλματος (III)

- Ορίζουμε τα ενδεχόμενα $E_m = \left\{ (x^n(m), y^n) \in A_\epsilon^{(n)} \right\}$, $m \in \{1, 2, \ldots, 2^{nR}\}$, δηλαδή τα ενδεχόμενα η κωδική λέξη $x^n(m)$ (που αντιστοιχεί στο μήνυμα m) να είναι από κοινού τυπική με τη ληφθείσα ακολουθία y^n η οποία προήλθε από μετάδοση της κωδικής λέξης $x^n(1)$.
- Συνεπώς,

$$\begin{aligned} \Pr(\mathcal{E}) = \Pr(\mathcal{E}|M=1) &= P(E_1^c \cup E_2 \cup E_3 \cup \ldots \cup E_{2^{nR}}|M=1) \\ &\overset{(a)}{\leq} P(E_1^c|M=1) + \sum_{m=2}^{2^{nR}} P(E_m|M=1). \end{aligned}$$

(a) από το φράγμα ένωσης ενδεχομένων (union of events bound).

Απόδειξη Θεωρήματος Κωδικοποίησης Καναλιού (10)
Υπολογισμός Πιθανότητας Σφάλματος (IV)

$$\Pr(\mathcal{E}) \leq P(E_1^c|M=1) + \sum_{m=2}^{2^{nR}} P(E_m|M=1).$$

- Από το Λήμμα Υπό Συνθήκη Τυπικότητας, η πιθανότητα η y^n να μην είναι από κοινού τυπική με τη $x^n(1)$ τείνει στο 0 για $n \to \infty$: Επομένως, για κάθε $\epsilon > 0$ υπάρχει n_0 τέτοιο ώστε $P(E_1^c|M = 1) \leq \epsilon$, για $n > n_0$.
- Επίσης, από τον τυχαίο τρόπο δημιουργίας του κώδικα, οι κωδικές λέξεις $x^n(1)$ και $x^n(m)$ είναι ανεξάρτητες μεταξύ τους για $m \neq 1$, με αποτέλεσμα η y^n να είναι ανεξάρτητη από τις $x^n(m)$ για $m \neq 1$. Από την Ιδιότητα Από Κοινού Ασυμπτωτικής Ισοδιαμέρισης, η πιθανότητα οι $x^n(m)$ και y^n να είναι από κοινού τυπικές ενώ επιλέχθηκαν ανεξάρτητα είναι $\leq 2^{-n(I(X;Y)-3\epsilon)}$.

Απόδειξη Θεωρήματος Κωδικοποίησης Καναλιού (11)
Υπολογισμός Πιθανότητας Σφάλματος (V)

- Συνδυάζοντας όλα τα παραπάνω,

$$\Pr(\mathcal{E}) \leq P(E_1^c|M=1) + \sum_{m=2}^{2^{nR}} P(E_m|M=1) \leq \epsilon + \sum_{m=2}^{2^{nR}} 2^{-n(I(X;Y)-3\epsilon)}$$
$$= \epsilon + \left(2^{nR}-1\right) 2^{-n(I(X;Y)-3\epsilon)} \leq \epsilon + 2^{-n(I(X;Y)-3\epsilon-R)} \leq 2\epsilon.$$

  Η τελευταία ανισότητα ισχύει εφόσον $n > n_1$ και $R < I(X;Y) - 3\epsilon$.

- Επομένως, εάν $R < I(X;Y)$, μπορούμε να επιλέξουμε n τέτοιο ώστε η μέση πιθανότητα σφάλματος υπολογισμένη επάνω σε όλους τους πιθανούς κώδικες και σε όλες τις πιθανές κωδικές λέξεις να μην υπερβαίνει το 2ϵ, για οποιοδήποτε $\epsilon > 0$.

- Επειδή οι κωδικές λέξεις δημιουργηθούν με βάση την κατανομή $p^*(x)$ η οποία μεγιστοποιεί την αμοιβαία πληροφορία, $I_{p^*}(X;Y) = C$ και, επομένως, μπορούμε να μεταδώσουμε με $R < C$.

Απόδειξη Θεωρήματος Κωδικοποίησης Καναλιού (12)
Επιλογή βιβλίου κωδίκων (I)

- Δεν τελειώσαμε ακόμα... Πρέπει να δείξουμε ότι η μέγιστη πιθανότητα σφάλματος $\lambda^{(n)} \to 0$ και, επίσης, ότι υπάρχει τουλάχιστον ένας κώδικας με $\lambda^{(n)} \to 0$.
- Δεδομένου ότι η μέση πιθανότητα σφάλματος για όλους τους τυχαίους κώδικες δεν υπερβαίνει το 2ϵ, υπάρχει τουλάχιστον ένα βιβλίο κωδίκων (κώδικας) $\mathcal{C}^*$ για το οποίο η μέση πιθανότητα σφάλματος δεν υπερβαίνει το 2ϵ: $\Pr(\mathcal{E}|\mathcal{C}^*) \leq 2\epsilon$. Ο $\mathcal{C}^*$ μπορεί να βρεθεί με αναζήτηση μέσα σε όλους τους 2^{nR} κώδικες. Επομένως,

$$\Pr(\mathcal{E}|\mathcal{C}^*) \leq \frac{1}{2^{nR}} \sum \lambda_m(\mathcal{C}^*) \leq 2\epsilon.$$

Απόδειξη Θεωρήματος Κωδικοποίησης Καναλιού (13)
Επιλογή βιβλίου κωδίκων (II)

- Το γεγονός ότι η μέση πιθανότητα σφάλματος του κώδικα $\mathcal{C}^*$ είναι $\leq 2\epsilon$, δεν εγγυάται ότι η πιθανότητα σφάλματος που αντιστοιχεί στη μετάδοση ενός συγκεκριμένου μηνύματος m (και, επομένως, μιας συγκεκριμένης κωδικής λέξης $x^n(m)$) θα είναι $\leq 2\epsilon$.
- Εάν θέλουμε να διασφαλίσουμε μικρή πιθανότητα σφάλματος για κάθε κωδική λέξη (και, άρα, για κάθε μήνυμα) μπορούμε να αφαιρέσουμε τις μισές χειρότερες κωδικές λέξεις του κώδικα (δηλαδή τις 2^{nR-1} κωδικές λέξεις με τη μεγαλύτερη πιθανότητα σφάλματος).
- Δεδομένου ότι η μέση πιθανότητα σφάλματος είναι $\leq 2\epsilon$, η μέγιστη πιθανότητα σφάλματος των μισών "καλύτερων" λέξεων που απομένουν δε θα υπερβαίνει το 4ϵ.

Απόδειξη Θεωρήματος Κωδικοποίησης Καναλιού (14)
Επιλογή βιβλίου κωδίκων (III)

- Ο νέος κώδικας έχει 2^{nR-1} κωδικές λέξεις και, άρα, ρυθμό $R' = R - \frac{1}{n}$. Για μεγάλα n, η απώλεια ρυθμού μετάδοσης είναι αμελητέα.
- Επομένως, δείξαμε ότι μπορούμε να επιτύχουμε οποιοδήποτε ρυθμό μετάδοσης που δεν υπερβαίνει τη χωρητικότητα και, ταυτόχρονα, η μέγιστη πιθανότητα σφάλματος $\lambda^{(n)} \leq 4\epsilon$.

Θεώρημα Κωδικοποίησης Καναλιού (Ευθύ) – Ανακεφαλαίωση

Για να αποδείξουμε το Θεώρημα Κωδικοποίησης Καναλιού

- Δημιουργήσαμε όλους τους πιθανούς κώδικες (βιβλία κωδίκων) με κωδικές λέξεις μεγάλου μήκους n.
- Η δημιουργία των κωδικών λέξεων έγινε με βάση την κατανομή $p^*(x)$ που μεγιστοποιεί την $I(X;Y)$.
- Κρατήσαμε τον καλύτερο από τους τυχαίους κώδικες $\mathcal{C}^*$ (τον κώδικα στον οποίο αντιστοιχεί η μικρότερη μέση πιθανότητα σφάλματος).
- Δείξαμε ότι, για αρκούντως μεγάλα μήκη κωδικών λέξεων n, εάν $R < I(X;Y)$, η πιθανότητα η ακολουθία εξόδου να μην είναι τυπική με τη μεταδοθείσα κωδική λέξη ή να είναι τυπική με κωδική λέξη διαφορετική από αυτή που μεταδόθηκε τείνει στο 0. Επομένως, η μέση πιθανότητα σφάλματος μπορεί να περιοριστεί αυθαίρετα κοντά στο 0.
- Με τροποποίηση του κώδικα (και αυθαίρετα μικρή απώλεια ρυθμού μετάδοσης) δείξαμε ότι όχι μόνο η μέση, αλλά και η μέγιστη πιθανότητα σφάλματος μπορεί να περιοριστεί αυθαίρετα κοντά στο 0.

Θεώρημα Κωδικοποίησης Καναλιού (Ευθύ) – Σχόλια

- Η δημιουργία τυχαίων κωδίκων οδηγεί μεν σε (μια) απόδειξη του ευθέος του Θεωρήματος Κωδικοποίησης Καναλιού, αλλά δεν αποτελεί πρακτικό τρόπο μετάδοσης.
- Πρόκειται για μια απόδειξη ύπαρξης (existence proof). Δεν είναι κατασκευαστική απόδειξη (constructive proof).
- Η δημιουργία του κώδικα, αν και πολύπλοκη, μπορεί να γίνει μια φορά υποθέτοντας ότι ο πίνακας μετάβασης του καναλιού $p(y|x)$ δεν αλλάζει.
- Παρατηρήστε ότι ο βέλτιστος κώδικας μπορεί να βρεθεί από τον πομπό και από το δέκτη ανεξάρτητα, χωρίς συνεννόηση, εάν γνωρίζουν και οι δύο τον πίνακα μετάβασης καναλιού και αν δημιουργήσουν όλους τους πιθανούς κώδικες (και κρατήσουν τον καλύτερο από άποψη ελάχιστης πιθανότητας σφάλματος).

Θεώρημα Κωδικοποίησης Καναλιού (Ευθύ) – Σχόλια (2)

- Με χρήση της ανισότητας Markov μπορεί να αποδειχτεί κάτι ακόμα πιο ισχυρό: Για επαρκώς μεγάλο n, ένας *τυχαίος* κώδικας που κατασκευάζεται με τον τρόπο που προαναφέραμε μπορεί να επιτύχει αυθαίρετα μικρή πιθανότητα σφάλματος σχεδόν βέβαια (περισσότερα σε επόμενες διαφάνειες).
- Δηλαδή, δε χρειάζεται να φτιάξουμε όλους τους κώδικες και να διαλέξουμε έναν καλό. Αρκεί να κατασκευάσουμε τυχαία έναν κώδικα (φυσικά το n θα πρέπει να είναι πολύ μεγάλο).
- Χάρη στο νόμο των μεγάλων αριθμών, η επικάλυψη των υπό συνθήκη τυπικών ακολουθιών $\mathcal{T}_\epsilon^{(n)}\left(Y^n|x^n(m)\right)$ αρχίζει και γίνεται αμελητέα καθώς το n αυξάνει (αρκεί να μη χρησιμοποιήσουμε περισσότερες από $\sim 2^{nI}$ κωδικές λέξεις).

Θεώρημα Κωδικοποίησης Καναλιού (Ευθύ) – Σχόλια (3)

- Παρατηρήστε ότι (σχεδόν όλες) οι κωδικές λέξεις $x^n(m)$ (τα σύμβολα των οποίων ακολουθούν κατανομή $p^*(x)$) είναι τυπικές.
- Αυτός είναι ένας από τους λόγους για τους οποίους "απλοί" κώδικες (όπως, για παράδειγμα, ο κώδικας επανάληψης) δεν είναι καλοί.
- Παρατηρήστε, επίσης, ότι, για έναν εξωτερικό παρατηρητή που δε γνωρίζει τα όρια των κωδικών λέξεων, τα σύμβολα που εισέρχονται στο κανάλι φαίνονται i.i.d. με κατανομή $p(x)$. Ωστόσο, για τον αποκωδικοποιητή που γνωρίζει τα όρια των λέξεων τα σύμβολα δεν είναι τυχαία, γιατί ανήκουν σε μία από 2^{nR} λέξεις του βιβλίου κωδίκων.

Θεώρημα Κωδικοποίησης Καναλιού (Ευθύ) – Σχόλια (4)

- Το σημαντικότερο πρόβλημα βρίσκεται στην αποκωδικοποίηση, καθώς ο αριθμός των κωδικών λέξεων των οποίων η από κοινού τυπικότητα με την y^n θα πρέπει να ελεγχθεί αυξάνει εκθετικά με το n.
- Το πρόβλημα αυτό παραμένει ακόμα και όταν η αποκωδικοποίηση γίνεται με χρήση άλλων κριτηρίων (π.χ. ανίχνευση Μέγιστης Πιθανοφάνειας).
- Η επίτευξη ρυθμών μετάδοσης κοντά στη χωρητικότητα του καναλιού με υλοποιήσιμους τρόπους αποτελεί αντικείμενο της Θεωρίας Κωδικοποίησης. Σήμερα, υπάρχουν περιπτώσεις καναλιών και απαιτήσεων σε πιθανότητα σφάλματος για τις οποίες η μετάδοση κοντά στη χωρητικότητα είναι εφικτή με πολυπλοκότητα που δεν είναι απαγορευτική για την υλοποίηση των κωδικοποιητών και των αποκωδικοποιητών (π.χ. Turbo Codes/LDPC). Ένας λόγος χάρη στον οποίο είναι εφικτή η κωδικοποίηση/αποκωδικοποίηση είναι ότι οι κώδικες που χρησιμοποιούνται είναι *γραμμικοί*.

Τυχαίοι και Δομημένοι Κώδικες

- Ο κώδικας στην απόδειξη του Θεωρήματος Κωδικοποίησης Καναλιού κατασκευάστηκε με τυχαίο τρόπο. Ο τρόπος αυτός κατασκευής είναι τυπικός στη Θεωρία Πληροφορίας.
- Στην πράξη, επιθυμούμε να έχουμε δομημένους (structured) κώδικες, ώστε να μπορούμε να κωδικοποιούμε και να αποκωδικοποιούμε με πρακτικώς υλοποιήσιμο τρόπο.
- Παραδοσιακά, στη Θεωρία Πληροφορίας αποτελούσε κοινή διαπίστωση ότι οι τυχαίοι κώδικες υπερτερούν των δομημένων.
- Ωστόσο, τα τελευταία χρόνια υπάρχουν ενδείξεις ότι για κάποια κανάλια (π.χ. Κανάλι Παρεμβολών) ενδέχεται δομημένοι κώδικες να υπερτερούν των τυχαίων.

Ένα τυχαίο βιβλίο κωδίκων είναι καλό με μεγάλη πιθανότητα

- Θα αποδείξουμε ότι, για αρκούντως μεγάλο n, ένα τυχαίο βιβλίο κωδίκων είναι καλό με όσο μεγάλη πιθανότητα επιθυμούμε.
- Επομένως, δε χρειάζεται να φτιάξουμε όλα τα βιβλία κωδίκων και να αναζητήσουμε ένα καλό. Αρκεί να φτάξουμε ένα τυχαίο βιβλίο κωδίκων (αλλά με πολύ μεγάλο n).
- Έστω $\Pr\{\mathcal{E}|\mathcal{C}\}$ η (αριθμητικώς) μέση πιθανότητα σφάλματος ενός συγκεκριμένου βιβλίου κωδίκων, $\mathcal{C}$.
- Θα δείξουμε ότι, για αρκούντως μεγάλο n, και για δεδομένο ψ, η πιθανότητα $\Pr\{\mathcal{E}|\mathcal{C}\} > \psi$ είναι αμελητέα.
- Η απόδειξη βρίσκεται π.χ. στο βιβλίο του R. Yeung, *A first course on Information Theory*.

Ένα τυχαίο βιβλίο κωδίκων είναι καλό με μεγάλη πιθανότητα (2)

$$\begin{aligned}
\Pr\{\mathcal{E}\} &= \sum_{\mathcal{C}} \Pr\{\mathcal{C}\}\Pr\{\mathcal{E}|\mathcal{C}\} \\
&= \sum_{\mathcal{C}:\Pr\{\mathcal{E}|\mathcal{C}\le\psi\}} \Pr\{\mathcal{C}\}\Pr\{\mathcal{E}|\mathcal{C}\} + \sum_{\mathcal{C}:\Pr\{\mathcal{E}|\mathcal{C}>\psi\}} \Pr\{\mathcal{E}|\mathcal{C}\}\Pr\{\mathcal{C}\} \\
&\ge \sum_{\mathcal{C}:\Pr\{\mathcal{E}|\mathcal{C}>\psi\}} \Pr\{\mathcal{C}\}\Pr\{\mathcal{E}|\mathcal{C}\} \\
&\overset{(a)}{>} \psi \sum_{\mathcal{C}:\Pr\{\mathcal{E}|\mathcal{C}>\psi\}} \Pr\{\mathcal{C}\}.
\end{aligned}$$

(a) για όλους τους όρους $\Pr\{\mathcal{E}|\mathcal{C}\} > \psi$.

Ένα τυχαίο βιβλίο κωδίκων είναι καλό με μεγάλη πιθανότητα (3)

- Επομένως,
$$\sum_{\mathcal{C}:\Pr\{\mathcal{E}|\mathcal{C}>\psi\}} \Pr\{\mathcal{C}\} < \frac{\Pr\{\mathcal{E}\}}{\psi}.$$
- Από την απόδειξη του ευθέος του Θεωρήματος Κωδικοποίησης Καναλιού γνωρίζουμε, επίσης, ότι $\Pr\{\mathcal{E}\} < 2\epsilon$.
- Συνεπώς,
$$\sum_{\mathcal{C}:\Pr\{\mathcal{E}|\mathcal{C}>\psi\}} \Pr\{\mathcal{C}\} < \frac{2\epsilon}{\psi}.$$
- Για δεδομένο ψ, μπορούμε να ελαπτώσουμε την πιθανότητα να προκύψει κάποιο κακό βιβλίο κωδίκων όσο θέλουμε ελαπτώνοντας το ϵ (μέσω της αύξησης του n).