Τμήμα Μηχανικών Πληροφορικής τ.ε.
Τεχνολογικό Εκπαιδευτικό Ίδρυμα Δυτικής Ελλάδας

Θεωρία Πληροφορίας

Διάλεξη 3: Βασικές αρχές Θεωρίας Πληροφοριών

Δρ. Μιχάλης Παρασκευάς
Επίκουρος Καθηγητής

Ατζέντα

1. Ορισμός του μέτρου της πληροφορίας κατά Shannon
2. Ιδιότητες της Εντροπίας
3. Συνδυασμένη και Αμοιβαία Πληροφορία
4. Υπό συνθήκη ποσότητα πληροφορίας
5. Αμοιβαία ποσότητα πληροφορίας

Ποσότητα Πληροφορίας (κατά Hartley)

Η ποσότητα της πληροφορίας ορίζεται ως ο δεκαδικός λογάριθμος του πλήθους των διαφορετικών λέξεων που μπορούν να σχηματιστούν από ένα δεδομένο πλήθος συμβόλων. Αν θεωρήσουμε ότι υπάρχουν μηνύματα μήκους k συμβόλων, από ένα αλφάβητο με N σύμβολα, η ***ποσότητα πληροφορίας*** είναι ίση με:

$$H(N^k) = \log(N^k) = k \log N$$

<u>Προσοχή</u> ! Ο ορισμός αυτός δεν λαμβάνει υπόψη του τις διαφορετικές πιθανότητες εμφάνισης των συμβόλων ενός αλφαβήτου, αντίθετα θεωρεί τα σύμβολα ισοπίθανα.

- Αν η βάση του λογάριθμου είναι το 10, τότε η μονάδα μέτρησης της πληροφορίας ονομάζεται decit.
- Αν η βάση του λογάριθμου είναι το 2, τότε η μονάδα μέτρησης της πληροφορίας ονομάζεται bit.

Παράδειγμα

1. Για σχηματισμό μηνυμάτων μήκους ενός συμβόλου από ένα αλφάβητο αποτελούμενο από δέκα (10) σύμβολα, η ποσότητα πληροφορίας ισούται με:

$$H(N^1) = log_{10}(N) = log_{10} 10 = 1 \; decit$$

2. Για σχηματισμό μηνυμάτων μήκους ενός συμβόλου από ένα αλφάβητο αποτελούμενο από δύο (2) σύμβολα, η ποσότητα πληροφορίας ισούται με:

$$H(N^1) = log_2(N) = log_2 2 = 1 \; bit$$

Εντροπία – Μέτρο πληροφορίας κατά Shannon

Μέση Ποσότητα Πληροφορίας ή Μέσο Πληροφοριακό Περιεχόμενο ή Εντροπία

Αν Y είναι μια διακριτή τυχαία μεταβλητή με δειγματοχώρο $Y = \{y_1, y_2, \dots\ y_n\}$ και συνάρτηση πιθανότητα μάζας $p(y_i)$, τότε η μέση ποσότητα πληροφορίας $H(Y)$ ισούται με:

$$H(Y) = -\sum_{i=1}^{n} p(y_i)\ \log\ p(y_i)$$

Η μέση ποσότητα πληροφορίας ονομάζεται **εντροπία.**

Παρατηρούμε ότι η εντροπία λαμβάνει υπόψη της τις διαφορετικές πιθανότητες εμφάνισης κάθε ενδεχόμενου y_i, δηλαδή κάθε συμβόλου που απαρτίζει ένα μήνυμα.

Παράδειγμα

Να υπολογισθεί η μέση ποσότητα πληροφορίας από την ταυτόχρονη ρίψη δύο ζαριών. Να θεωρήσετε ότι τα αποτελέσματα (4,2) και (2,4) είναι διαφορετικά.

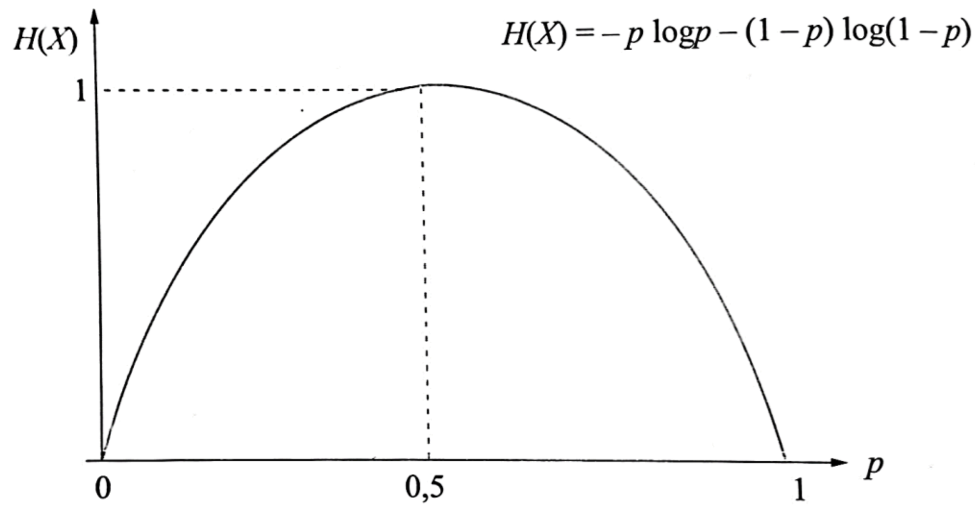
<u>Απάντηση</u>: Επειδή το πλήθος των δυνατών συνδυασμών της ρίψης δύο ζαριών είναι 36 και οι συνδυασμοί είναι ισοπίθανοι μεταξύ τους, προκύπτει ότι η μέση ποσότητα πληροφορίας της ρίψης δύο ζαριών είναι η ακόλουθη:

$$H(Y) = -\sum_{i=1}^{n} p(y_i)\ log\ p(y_i) = -36\frac{1}{36}log\frac{1}{36} = log36 = 5{,}17\ bits$$

Εντροπία διακριτής πηγής δύο συμβόλων

Για διακριτή τυχαία μεταβλητή X με δύο ενδεχόμενα, π.χ. εκπομπή ενός από δύο δυνατά μηνύματα και αντίστοιχες πιθανότητες p και $1-p$, η εντροπία είναι:

$$H(X) = -p\ logp - (1-p)\ log(1-p)$$



<u>Παρατηρήσεις</u>:

- Η εντροπία λαμβάνει τη μέγιστη τιμή, που ισούται με ένα, όταν τα δύο γεγονότα είναι ισοπίθανα, δηλαδή $p = 1/2$.
- Αν $p = 1$ ή $p = 0$, τότε η εντροπία είναι μηδέν, επειδή η έκβαση του πειράματος είναι βέβαιη.

Ιδιότητες της Εντροπίας

1. Η μέση ποσότητα πληροφορίας είναι συνεχής στο p.
2. Η μέση πληροφορία Η(Υ) είναι συμμετρική, δηλαδή διάταξη των πιθανοτήτων δεν την επηρεάζει.
3. Η εντροπία Η(Υ) παίρνει την μέγιστη τιμή όταν όλα τα ενδεχόμενα είναι ισοπίθανα.
4. Η εντροπία είναι προσθετική. Η ιδιότητα αυτή αναφέρεται στην περίπτωση κατά την οποία δύο ανεξάρτητες τυχαίες μεταβλητές Χ και Υ συνδυάζονται.

Συνδυασμένη Ποσότητα Πληροφορίας

Ένα τυχαίο πείραμα (Χ,Υ) έχει ως δυνατά αποτελέσματα όλους τους δυνατούς συνδυασμούς των αποτελεσμάτων $X = \{x_1, x_2, \dots, x_n\}$ και $Y = \{y_1, y_2, \dots, y_m\}$.

Η κατανομή πιθανοτήτων δίνεται από την παρακάτω σχέση:

$$P = \{p(x_1, y_1), p(x_1, y_2), \dots, p(x_1, y_m), \dots p(x_n, y_1), p(x_n, y_2), \dots p(x_n, y_m)\}$$

Αν (Χ, Υ) είναι ένα τυχαίο πείραμα με δισδιάστατο δειγματοχώρο και κατανομή πιθανοτήτων P, τότε η **συνδυασμένη ποσότητα πληροφορίας $\mathbf{H}(\mathbf{X}, \mathbf{Y})$** ορίζεται ως εξής:

$$H(\mathrm{X}, Y) = -\sum_{i=1}^{n}\sum_{j=1}^{m} p(x_i, y_i)\ \ log\ \ p(x_i, y_i)$$

Υπό Συνθήκη Ποσότητα Πληροφορίας

Αρκετές φορές μας ενδιαφέρει να υπολογίσουμε την ποσότητα πληροφορίας μιας τυχαίας μεταβλητής Χ, όταν δίνεται το αποτέλεσμα μιας άλλης τυχαίας μεταβλητής Υ.

Αυτό ονομάζεται **υπό συνθήκη** ποσότητα πληροφορίας της Χ ως προς το Υ.

Η μέση υπό συνθήκη ποσότητα πληροφορίας του τυχαίου πειράματος Χ, με δεδομένο το αποτέλεσμα του πειράματος Υ δίνεται από την σχέση:

$$H(\mathrm{X}/Y) = -\sum_{i=1}^{n}\sum_{j=1}^{m} p(x_i, y_i)\ log\, p(x_i/y_i)$$

Πρόταση: Για δύο τυχαίες μεταβλητές Χ και Υ ισχύει η παρακάτω σχέση:

$$H(X,Y) = H(X) + H(Y/X) = H(Y) + H(X/Y)$$

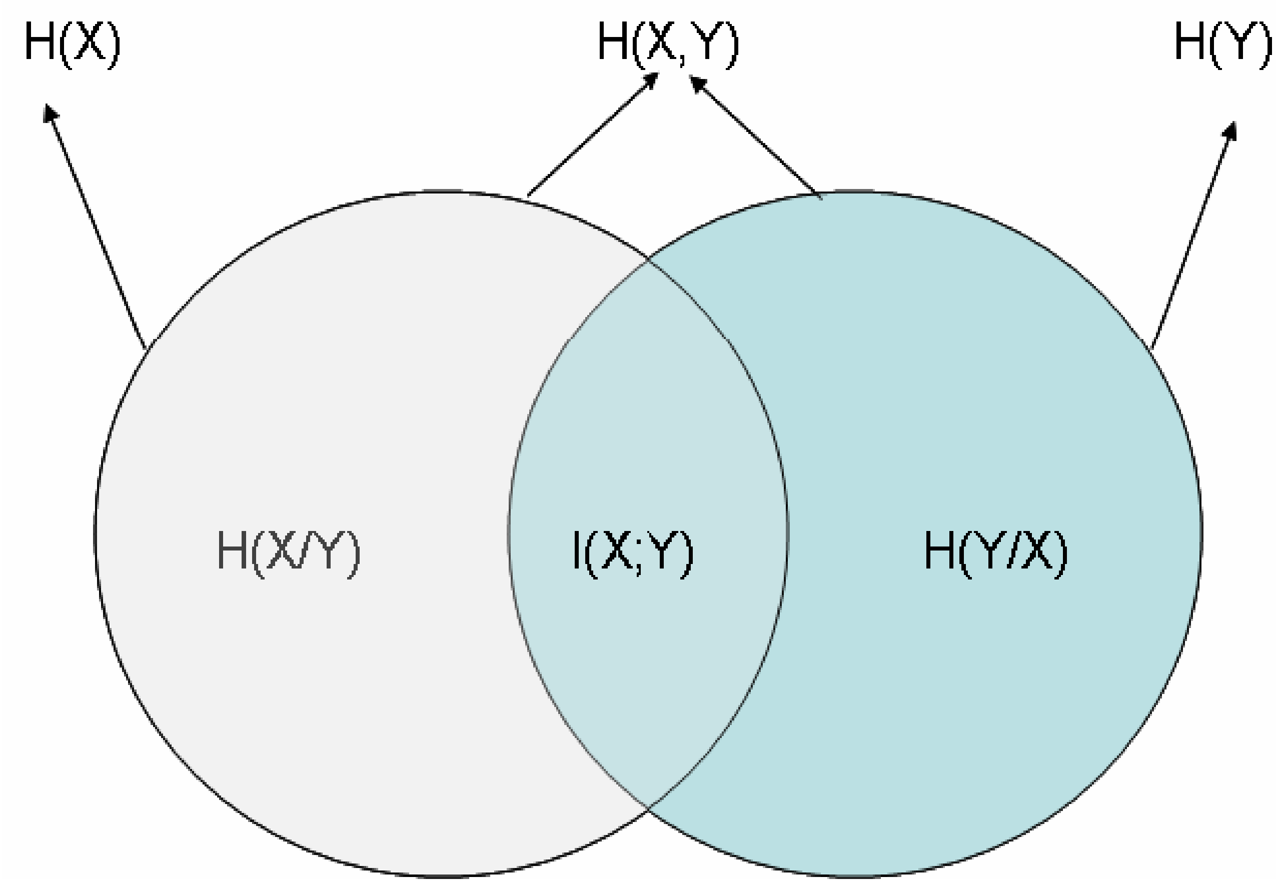
Αμοιβαία Ποσότητα Πληροφορίας

Η **αμοιβαία ποσότητα πληροφορίας** είναι ένα μέτρο της ποσότητας πληροφορίας που μια τυχαία μεταβλητή περιέχει για μια άλλη τυχαία μεταβλητή ή ένα μέτρο εξάρτησης μεταξύ των δύο και δίνεται από τη σχέση:

$$I(X;Y) = H(X) - H\left(\frac{Y}{X}\right) =$$

$$= -\sum_{i=1}^{n}\sum_{j=1}^{m} p(x_i, y_i)\ \ log\,\frac{p(x_i, y_i)}{p(x_i)\ p(y_i)}$$

Μέτρα ποσότητας πληροφορίας



Άσκηση 1

Δίνονται οι τυχαίες μεταβλητές X και Υ, με δύο δυνατά αποτελέσματα η κάθε μία. Οι δισδιάστατες πιθανότητες είναι:

$$p(x_1, y_1) = \frac{1}{8}, \qquad p(x_1, y_2) = \frac{1}{8}, \qquad p(x_2, y_1) = \frac{1}{2}, \qquad p(x_2, y_2) = \frac{1}{4}$$

1. Ποια ποσότητα πληροφορίας λαμβάνουμε όταν μας γνωστοποιείται το αποτέλεσμα της τυχαίας μεταβλητής X; Ποιά όταν μας γνωστοποιείται το αποτέλεσμα της Υ;
2. Ποια είναι η ποσότητα πληροφορίας που λαμβάνουμε όταν γνωρίζουμε το αποτέλεσμα της συνδυασμένης διαδικασίας (X,Υ);
3. Ποια ποσότητα πληροφορίας λαμβάνουμε όταν μας γνωστοποιείται το αποτέλεσμα της Υ αν γνωρίζουμε το X;

<u>Απάντηση</u>:

Αρχικά θα υπολογίσουμε τις πιθανότητες $p(x_1)$, $p(x_2)$, $p(x_3)$, $p(x_4)$. Είναι:

$$p(x_1) = \sum_{j=1}^{2} p\left(x_1, y_j\right) = \frac{1}{8} + \frac{1}{8} = \frac{1}{4} \qquad p(x_2) = \sum_{j=1}^{2} p\left(x_2\ y_j\right) = \frac{1}{2} + \frac{1}{4} = \frac{3}{4}$$

$$p(y_1) = \sum_{i=1}^{2} p(x_i, y_1) = \frac{1}{8} + \frac{1}{2} = \frac{5}{8} \qquad p(y_2) = \sum_{i=1}^{2} p(x_i, y_2) = \frac{1}{8} + \frac{1}{4} = \frac{3}{8}$$

Άσκηση 1 (συνέχεια)

Ερώτημα 1:

$$H(X) = -\sum_{i=1}^{2} p(x_i)\ log\ p(x_i) = -\left[\frac{1}{4}\log\left(\frac{1}{4}\right) + \frac{3}{4}\log\left(\frac{3}{4}\right)\right] = 0{,}813\ bits$$

$$H(Y) = -\sum_{i=1}^{2} p(y_i)\ log\ p(y_i) = -\left[\frac{5}{8}\log\left(\frac{5}{8}\right) + \frac{3}{8}\log\left(\frac{3}{8}\right)\right] = 0{,}956\ bits$$

Ερώτημα 2:

$$H(\mathrm{X}, Y) = -\sum_{i=1}^{n}\sum_{j=1}^{m} p(x_i, y_i)\ log\ p(x_i, y_i) = \cdots = 1{,}75\ \ bits$$

Ερώτημα 3:

$$H(X,Y) = H(X\ ) + H(Y/X) \Rightarrow H(Y/X) = H(X,Y) - H(X) = 1{,}75\ - 0{,}813 = 0{,}935\ bits$$

Ερώτημα 4:

$$H(X/Y) = H(X,Y) - H(Y) = 1{,}75\ - 0{,}95 = 0{,}8\ bits$$

Άσκηση 2

Θεωρούμε ένα ψηφιακό επικοινωνιακό κανάλι που χρησιμοποιεί ως εισόδους και εξόδους τα σύμβολα ‘0’ και ‘1’. Δίνονται οι τυχαίες οι πιθανότητες $p(x_1) = 1/2$, $p(y_1/x_1) = 3/4$, $p(y_1/x_2) = 1/4$, όπου τα x_1, y_1 αφέρονται στο σύμβολο ‘0’ και τα x_2, y_2 στα σύμβολα ‘1’.

1. Ποια ποσότητα πληροφορίας λαμβάνουμε όταν μας γνωστοποιείται ποιο σύμβολο έχει ληφθεί στην έξοδο όταν γνωρίζουμε ότι στην είσοδο έχει σταλεί το ‘1’.
2. Ποια ποσότητα πληροφορίας λαμβάνουμε όταν μας γνωστοποιείται ποιο σύμβολο έχει ληφθεί στην έξοδο όταν γνωρίζουμε ποιο σύμβολο έχει σταλεί στην είσοδο.
3. Ποια ποσότητα πληροφορίας λαμβάνουμε αν μας γνωστοποιείται ποιο σύμβολο έχει σταλεί στην είσοδο και ποιο σύμβολο έχει ληφθεί στην έξοδο
4. Ποια ποσότητα πληροφορίας λαμβάνουμε αν μας γνωστοποιείται ποιο σύμβολο έχει σταλεί στην είσοδο όταν γνωρίζουμε ποιο σύμβολο έχει ληφθεί στην έξοδο.

<u>Απάντηση</u>: Υπολογίζουμε τις συνδυασμένες πιθανότητες:

$$p(y_2/x_1) = \frac{1}{4}, \qquad p(y_2/x_2) = \frac{3}{4}$$

$$p(x_1, y_2) = p(y_2/x_1)\, p(x_1) = \frac{1}{8}, \qquad p(x_2, y_1) = p(y_1/x_2)\, p(x_2) = \frac{1}{8}$$

$$p(x_1, y_1) = p(y_1/x_1)\, p(x_1) = \frac{3}{8}, \qquad p(x_2, y_2) = p(y_2/x_2)\, p(x_2) = \frac{3}{8}$$

Άσκηση 2 (συνέχεια)

Ερώτηση 1:

$$H(Y/x_2) = -\sum_{k=1}^{2} p(y_k/x_2)\ log\, p(y_k/x_2) = 0{,}81\ bits$$

Ερώτηση 2:

$$H(Y/X) = -\sum_{i=1}^{2}\sum_{k=1}^{2} p\left(y_i, x_j\right)\ log\, p\left(y_j/x_i\right) = 0{,}91\ bits$$

Ερώτηση 3:

$$H(X,Y) = H(X) + H(Y/X) = 1 + 0{,}91 = 1{,}91$$

Ερώτηση 4:

$$H(X/Y) = H(X,Y) - H(Y) = 1{,}91 - 0{,}96 = 0{,}95\ bits$$

Άσκηση 3

Μία τράπουλα έχει 52 χαρτιά. Αυτά χωρίζονται σε τέσσερις κατηγορίες: τα μπαστούνια, τα σπαθιά, τις κούπες και τα καρό, με δεκατρία χαρτιά η καθεμία. Τα μπαστούνια και τα σπαθιά είναι μαύρου χρώματος και τα υπόλοιπα κόκκινου χρώματος. Το τυχαίο πείραμα συνίσταται στο τράβηγμα ενός χαρτιού από την τράπουλα. Θεωρούμε ότι για το κάθε χαρτί η πιθανότητα να είναι αποτέλεσμα του τυχαίου πειράματος είναι η ίδια.

1. Ποια ποσότητα πληροφορίας λαμβάνουμε όταν μας γνωστοποιείται μόνο το χρώμα του χαρτιού;
2. Ποια ποσότητα πληροφορίας λαμβάνουμε όταν μας γνωστοποιείται το χρώμα και η κατηγορία στην οποία ανήκει το χαρτί;
3. Ποια ποσότητα πληροφορίας λαμβάνουμε όταν μας γνωστοποιείται το χρώμα, η κατηγορία και ο αριθμός του χαρτιού;
4. Ποια ποσότητα πληροφορίας λαμβάνουμε όταν μας γνωστοποιείται ο αριθμός αν γνωρίζουμε ήδη το χρώμα του.
5. Ποια ποσότητα πληροφορίας λαμβάνουμε όταν μας γνωστοποιείται η κατηγορία αν το χρώμα του χαρτιού είναι ήδη γνωστό;

Άσκηση 3 (συνέχεια)

Έστω οι τυχαίες μεταβλητές Χ, Υ και Z, που αναφέρονται στο χρώμα, στην κατηγορία και τον αριθμό ενός χαρτιού, αντίστοιχα. Οι δυνατές τιμές κάθε μεταβλητής είναι:

- $X = \{\text{μαύρο, κόκκινο}\}$
- $Y = \{\text{μπαστούνια, σπαθια καρό, κούπες}\}$
- $Z = \{\text{αριθμοι από 1 εως 10 και βαλες, νταμα και ρηγας}\}$

<u>Ερώτημα 1</u>: Αφού κάθε χαρτί μπορεί να τραβηχτεί με την ίδια πιθανότητα, η πιθανότητα να είναι αυτό μαύρο ή κόκκινο είναι 0,5 και η πιθανότητα να είναι αυτό μπαστούνι, σπαθί, καρό ή κούπα είναι 0,25. Αντίστοιχα, η πιθανότητα να είναι το «τρία» ή ο «ρήγας» ή οποιοδήποτε άλλο είναι 1/13. Οι εντροπίες των Χ, Υ και Ζ, είναι:

- $H(X) = -0{,}5\ log\ 0{,}5 - 0{,}5\ log\ 0{,}5 = 1\ bit$
- $H(Y) = -0{,}25\ log\ 0{,}25 - 0{,}25\ log 0{,}25 - 0{,}25\ log 0{,}25 - 0{,}25\ log 0{,}25 = 2\ bits$
- $H(Z) = 13(-(1/13)\ log(1/13)) = 3{,}7\ bits$

Άσκηση 3 (συνέχεια)

<u>Ερώτημα 2</u>: Οι δυνατοί συνδυασμοί των X και Y για τους οποίους οι πιθανότητες είναι διάφορες του μηδενός είναι (μαύρο, μπαστούνι), (μαύρο, σπαθί), (κόκκινο, καρό) και (κόκκινο, κούπα). Η πιθανότητα αυτών των συνδυασμών είναι ίση με 0,25. Οι υπόλοιποι συνδυασμοί (μαύρο, καρό), (μαύρο, κούπα), (κόκκινο, μπαστούνι) και (κόκκινο, σπαθί) δεν μπορούν να λάβουν χώρα, δηλαδή η πιθανότητά τους είναι μηδενική.

Έτσι, η συνδυασμένη πληροφορία των *(X, Y)* είναι:

- $H(X,Y) = -0{,}25 log 0{,}25 - 0{,}25 log 0{,}25 - 0{,}25 log 0{,}25 - 0{,}25 log 0{,}25 = 2\ bits$

Παρατηρούμε ότι υπάρχουν συνολικά 26 συνδυασμοί (1, μαύρο), (1, κόκκινο), (2, μαύρο), (2, κόκκινο), κλπ, των οποίων η πιθανότητα να λάβουν χώρα είναι (1/26). Οπότε η συνδυασμένη πληροφορία των (*X, Z)* είναι:

- $H(X,Z) = 26(-(1/26) log(1/26)) = 4{,}7\ bits$

Άσκηση 3 (συνέχεια)

Ερώτημα 3: Τώρα μας ενδιαφέρει η συνδυασμένη πληροφορία των (*X, Y, Z).* Οι δυνατοί συνδυασμοί είναι 52, δηλαδή (μαύρο, μπαστούνι, 1), (μαύρο, μπαστούνι, 2), ..., (κόκκινο, κούπα, ρήγας) και η πιθανότητα για κάθε συνδυασμό να λάβει χώρα (1/52).

Έτσι, η συνδυασμένη πληροφορία είναι

- $H(X,Y,Z) \ = \ 52(-\,(1/52)\log(1/52)) = 5\ bits$

Άσκηση 3 (συνέχεια)

<u>Ερώτημα 4</u>: Για τον υπολογισμό της υπό συνθήκης πληροφορίας της Ζ με δεδομένη τη X, παρατηρούμε ότι η Ζ είναι ανεξάρτητη της Χ. Έτσι, η υπό συνθήκη πληροφορία της Ζ με δεδομένη την τιμή της Χ είναι ίση με την εντροπία της Ζ, δηλαδή ίση με 3,7 bits.

Το αποτέλεσμα μπορεί να δοθεί αφαιρώντας από τη συνδυασμένη πληροφορία $H(X,Z)$ την εντροπία της Χ, $H(X)$, δηλαδή:

- $H(Z/X) = H(X,Z) - H(X) = 4{,}7 - 1 = 3{,}7\ bits$

Άσκηση 3 (συνέχεια)

<u>Ερώτημα 5</u>: Για τον υπολογισμό τής υπό συνθήκης πληροφορίας της Y με δεδομένη την τιμή της X παρατηρούμε ότι οι δυνατοί συνδυασμοί με πιθανότητα μεγαλύτερη του μηδενός είναι δύο: μπαστούνι ή σπαθί αν το χρώμα είναι μαύρο και καρό ή κούπα αν είναι κόκκινο. Η πιθανότητα για κάθε συνδυασμό είναι ίση με 0,5.

Έτσι, η υπό συνθήκη πληροφορία υπολογίζεται αφαιρώντας από τη συνδυασμένη πληροφορία H(X, Y) την εντροπία της X, H(X).

- $H(Y|X) = H(X,Y) - H(X) \; = \; 2 - 1 \; = 1 \; bit$