

## Ασκήσεις - 14/3/2018

### Άσκηση 1 - Τυπικότητα, Νόμος μεγάλων αριθμών, Απόσταση KL

- **Ερώτημα 1:** Να γεννήσετε i.i.d. ακολουθία στο αλφάβητο {1,2,3,4} με πιθανότητες γραμμάτων  $p=[1/2, 1/4, 1/8, 1/8]$

**Απάντηση:**

$N = 10000;$

$p = [1/2, 1/4, 1/8, 1/8]$

$x = \text{randsample}(1:4,N,\text{true},p);$

$x(1:20);$  % Πρώτοι 20 όροι

- **Ερώτημα 2:** Υπολογισμός σχετικών συχνοτήτων (ή εμπειρικών πιθανοτήτων) με την εντολή histogram (αντικατάσταση της hist)

**Απάντηση:**

$\text{edges} = 0.5:1:4.5$

$h = \text{histogram}(x,\text{edges})$

$\text{counts} = h.\text{Values} = [4971, 2311, 1268, 1250]$

$\text{ep1} = \text{counts}/\text{length}(x) = [0.4971, 0.2511, 0.1268, 0.1250]$

- **Ερώτημα 3:** Επανάληψη των (1) και (2) και σύγκριση του  $\text{ep2}$  με  $\text{ep1}$  και με  $p$

**Απάντηση:**

$\text{ep2} = [0.5054, 0.2476, 0.1219, 0.1251]$

**KL απόσταση:** Υπολογισμός των  $D(\text{ep1}||p)$ ,  $D(p||\text{ep1})$ ,  $D(\text{ep1}||\text{ep2})$

$D(\text{ep1}||p) = \text{ep1} * \log_2(\text{ep1}'./p') = 3.4251 * 10^{-5}$

$D(p||\text{ep1}) = p * \log_2(p'./\text{ep1}') = 3.4181 * 10^{-5}$

$D(\text{ep1}||\text{ep2}) = 2.7466 * 10^{-4}$

- **Ερώτημα 4:** Επιβεβαίωση της τυπικότητας ( $-\frac{1}{n} \log p(x) \approx H(X)$ )

**Απάντηση:**  $-\frac{1}{n} \log p(x) = -\frac{1}{n} \sum n_i \log p_i = -\frac{1}{n} \text{counts} * \log_2(p') = 1.7547$ , ενώ

$H = -p * \log_2(p') = 1.7500$

**Άσκηση 2 - Υπολογισμός εντροπίας (και αμοιβαίας πληροφορίας) σε εικόνες**

- **Ανάγνωση εικόνας:**

```
i = imread('pout.tif')
```

- **Εμφάνιση εικόνας:**

```
imshow(i) εικόνα έντασης (ή γκρι)
```

Η μεταβλητή *i* είναι ένας 291X240 πίνακας με στοιχεία uint8, δηλαδή 256 επίπεδα κβαντιστή των 8 bits.

```
[counts,x] = imhist(i);
```

```
stem(x,counts)
```

Παρατηρώ ότι το ιστόγραμμα είναι συγκεντρωμένο δεν καλύπτει ολόκληρο το διάστημα [0, 1]. Το *x* εκτείνεται σε όλα τα 256 επίπεδα του γκρι. Δοθέντος ότι η εικόνα είναι κβαντισμένη σε  $2^8$  (uint8) επίπεδα με κβαντιστή των 8 bits.

- **Σχετικές συχνότητες:**

```
p = counts/(291*240);
```

Απαλλαγή από τα NAN (not a number) δοθέντος ότι το MATLAB γενικά θεωρεί ότι  $0 \cdot \log 0 = \text{NAN}$ . Εξαίρεση των παραπάνω τιμών:

```
a = find(p);
```

```
size(a) = 121 (πάρα πολλές τιμές λόγω περιορισμένης αντίθεσης ('contrast'))
```

```
p = p(a);
```

- **Εντροπία:**  $e = -p \cdot \log_2(p) = 5.7599$

Η εντροπία χρησιμοποιείται ως δείκτης υφής (texture) στην επεξεργασία εικόνας.