



# Μαθηματικά Και Στατιστική Στη Βιολογία

Ενότητα 4 : Περιγραφική Στατιστική

Ι. Αντωνίου, Χ. Μπράτσας  
Τμήμα Μαθηματικών



Ευρωπαϊκή Ένωση  
Ευρωπαϊκό Κοινωνικό Ταμείο



ΥΠΟΥΡΓΕΙΟ ΠΑΙΔΕΙΑΣ ΚΑΙ ΘΡΗΣΚΕΥΜΑΤΩΝ  
ΕΙΔΙΚΗ ΥΠΗΡΕΣΙΑ ΔΙΑΧΕΙΡΙΣΗΣ

Με τη συγχρηματοδότηση της Ελλάδας και της Ευρωπαϊκής Ένωσης



ΕΥΡΩΠΑΪΚΟ ΚΟΙΝΩΝΙΚΟ ΤΑΜΕΙΟ

# Άδειες Χρήσης

- Το παρόν εκπαιδευτικό υλικό υπόκειται σε άδειες χρήσης Creative Commons.
- Για εκπαιδευτικό υλικό, όπως εικόνες, που υπόκειται σε άλλου τύπου άδειας χρήσης, η άδεια χρήσης αναφέρεται ρητώς.



# Χρηματοδότηση

- Το παρόν εκπαιδευτικό υλικό έχει αναπτυχθεί στα πλαίσια του εκπαιδευτικού έργου του διδάσκοντα.
- Το έργο «Ανοικτά Ακαδημαϊκά Μαθήματα στο Αριστοτέλειο Πανεπιστήμιο Θεσσαλονίκης» έχει χρηματοδοτήσει μόνο τη αναδιαμόρφωση του εκπαιδευτικού υλικού.
- Το έργο υλοποιείται στο πλαίσιο του Επιχειρησιακού Προγράμματος «Εκπαίδευση και Δια Βίου Μάθηση» και συγχρηματοδοτείται από την Ευρωπαϊκή Ένωση (Ευρωπαϊκό Κοινωνικό Ταμείο) και από εθνικούς πόρους.





# Περιγραφική Στατιστική



Ευρωπαϊκή Ένωση  
Ευρωπαϊκό Κοινωνικό Ταμείο



ΥΠΟΥΡΓΕΙΟ ΠΑΙΔΕΙΑΣ ΚΑΙ ΘΡΗΣΚΕΥΜΑΤΩΝ  
ΕΙΔΙΚΗ ΥΠΗΡΕΣΙΑ ΔΙΑΧΕΙΡΙΣΗΣ

Με τη συγχρηματοδότηση της Ελλάδας και της Ευρωπαϊκής Ένωσης



ΕΥΡΩΠΑΪΚΟ ΚΟΙΝΩΝΙΚΟ ΤΑΜΕΙΟ

# Περιεχόμενα Ενότητας

## (1 από 2)

1. Συλλογή Δεδομένων από τις Παρατηρήσεις
2. Οργάνωση και Παρουσίαση Δεδομένων
3. Ρίψη Ζαριού
4. Συχνότητες και Σχετικές Συχνότητες Δείγματος
5. Παραστάσεις Δεδομένων
6. Παράμετροι Θέσης Κατανομής
7. Παράμετροι Μεταβλητότητας Κατανομής
8. Σφάλματα
9. Παράμετροι σχήματος
10. Γραμμική Παλινδρόμηση
11. Εξίσωση Ευθείας Γραμμικής Παλινδρόμησης



# Περιεχόμενα Ενότητας

## (2 από 2)

12. Συναρτήσεις που ανάγονται σε γραμμικές
13. Εκτίμηση Παραμέτρων
14. Μέσο Τετραγωνικό Σφάλμα
15. Γραμμική εξάρτηση
16. Μέθοδοι Παλινδρόμησης
17. Συντελεστής Αλληλοεξάρτησης Αμοιβαίας Πληροφορίας Δείγματος



# Σκοποί Ενότητας

- Στην Ενότητα 4 παρουσιάζονται βασικές έννοιες της περιγραφικής στατιστικής όπως οι έννοιες της συχνότητας και της σχετικής συχνότητας, οι παράμετροι θέσης και μεταβλητότητας κατανομής όπως και η γραμμική παλινδρόμηση.



# Συλλογή Δεδομένων από τις Παρατηρήσεις

- **Συλλογή Δεδομένων από τις Παρατηρήσεις**
  - **Διενέργεια έρευνας** κατά πόσον κάποιος πιστεύει ότι η κοτόσουπα είναι αποτελεσματική στην πρόληψη του κρυολογήματος.
  - **Πείραμα** τους δίνουμε κοτόσουπα και μετράμε πόσες φορές θα κρυολογήσουν.
- **Remember Always**
  - Ο σωστός σχεδιασμός της **Έρευνας ή του Πειράματος** προϋποθέτει:
  - **Σαφή** διατύπωση των ερωτημάτων (queries) ώστε να προσδιορίζονται ορθά οι σχετικές Μεταβλητές.
  - **Στοχοποίηση** του πληθυσμού που ενδιαφέρει.
  - Επιλογή **τυχαίου** (αμερόληπτου) δείγματος.
  - Σωστή **κωδικοποίηση** και **καταχώρηση** των απαντήσεων σε εύχρηστη βάση δεδομένων για περαιτέρω επεξεργασία.





# Οργάνωση και Παρουσίαση Δεδομένων (1 από 2)

- Οργάνωση και Παρουσίαση Δεδομένων
- Φάσμα Δείγματος  $N$  μετρήσεων της Μεταβλητής  $X$ : οι  $n \leq N$  διαφορετικές τιμές  $\{x_1, x_2, \dots, x_n\}$  του Δείγματος. Συνήθως διατάσσονται:  $x_1 < x_2 < \dots < x_n$
- Συχνότητα του Δείγματος
- $M_v =$  ο αριθμός εμφανίσεων της τιμής  $x_v$ ,  $v=1,2,\dots,n$  στο δείγμα
- $M_1 + M_2 + \dots + M_n = M$



# Οργάνωση και Παρουσίαση Δεδομένων (2 από 2)

- Σχετική Συχνότητα του Δείγματος = Πιθανότητα Δείγματος

$$\tilde{\rho}_v = \frac{M_v}{M} = \tilde{\rho}(x_v), v=1,2,\dots,n$$

Η σχετική συχνότητα είναι πιθανότητα στο φάσμα του Δείγματος  $\{x_1, x_2, \dots, x_n\}$  θεωρούμενο ως Δειγματοχώρος

- Αθροιστική Σχετική Συχνότητα του Δείγματος

$$\tilde{F}_v = \frac{M_1 + M_2 + \dots + M_v}{M} = F(x_v) = \tilde{\rho}_1 + \tilde{\rho}_2 + \dots + \tilde{\rho}_v, v=1,2,\dots,n$$

Η συνάρτηση Κατανομής της πιθανότητας  $\tilde{\rho}$



# Ρίψη Ζαριού (1 από 2)

Ρίψη Ζαριού 20 φορές	
Μέτρηση	Αποτέλεσμα
1	$\chi_1 = 2$
2	$\chi_2 = 4$
3	$\chi_3 = 5$
4	$\chi_4 = 6$
5	$\chi_5 = 3$
6	$\chi_6 = 4$
7	$\chi_7 = 2$
8	$\chi_8 = 1$
9	$\chi_9 = 4$
10	$\chi_{10} = 6$

Ρίψη Ζαριού 20 φορές	
Μέτρηση	Αποτέλεσμα
11	$\chi_{11} = 6$
12	$\chi_{12} = 3$
13	$\chi_{13} = 2$
14	$\chi_{14} = 4$
15	$\chi_{15} = 2$
16	$\chi_{16} = 6$
17	$\chi_{17} = 5$
18	$\chi_{18} = 4$
19	$\chi_{19} = 2$
20	$\chi_{20} = 1$



# Ρίψη Ζαριού (2 από 2)

Ρίψη Ζαριού 20 φορές			
Φάσμα	Συχνότητα	Σχετική Συχνότητα	Αθροιστική Σχετική Συχνότητα
$\chi_1 = 1$	$M_1 = 2$	$\tilde{\rho}_1 = \frac{2}{20}$	$\tilde{F}_1 = \frac{2}{20}$
$\chi_2 = 2$	$M_2 = 5$	$\tilde{\rho}_2 = \frac{5}{20}$	$\tilde{F}_2 = \frac{7}{20}$
$\chi_3 = 3$	$M_3 = 2$	$\tilde{\rho}_3 = \frac{2}{20}$	$\tilde{F}_3 = \frac{9}{20}$
$\chi_4 = 4$	$M_4 = 5$	$\tilde{\rho}_4 = \frac{5}{20}$	$\tilde{F}_4 = \frac{14}{20}$
$\chi_5 = 5$	$M_5 = 2$	$\tilde{\rho}_5 = \frac{2}{20}$	$\tilde{F}_5 = \frac{16}{20}$
$\chi_6 = 6$	$M_6 = 4$	$\tilde{\rho}_6 = \frac{4}{20}$	$\tilde{F}_6 = \frac{20}{20} = 1$



# Συχνότητες και Σχετικές Συχνότητες Δείγματος

## Συχνότητες και Σχετικές Συχνότητες Δείγματος

- Παρ' ότι είναι ισοδύναμες μαθηματικά, πολλοί χρηστές διευκολύνονται με τις Φυσικές συχνότητες.
- « It makes little mathematical difference whether statistics are expressed as probabilities, percentages, or absolute frequencies. It does, however, make a psychological difference. More specifically, statistics expressed as natural frequencies improve the statistical thinking of experts and nonexperts alike»
- Hoffrage U., Lindsey S., Hertwig R., Gigerenzer G. 2000, Communicating Statistical Information, Science 230, 22 December, 2261-2262 .



# Παραστάσεις Δεδομένων

- **Παραστάσεις Δεδομένων**
  - ραβδογράμματα,
  - ιστογράμματα,
  - κυκλικά διαγράμματα,
  - σημειογράμματα,
  - θηκογράμματα



# Παράμετροι Θέσης Κατανομής

- Οι Στατιστικές Παράμετροι του Δείγματος είναι οι στατιστικές παράμετροι με πιθανότητα τη σχετική συχνότητα του δείγματος.
- Παράμετροι Θέσης Κατανομής (Location Parameters)
- Μέση Τιμή (Mean)

$$\tilde{m} = \frac{\chi_1 + \dots + \chi_M}{M} = \frac{x_1 \tilde{f}_1 + \dots + x_n \tilde{f}_n}{n} = x_1 \tilde{\rho}_1 + \dots + x_n \tilde{\rho}_n$$

- Ροπές (Moments)
- Κορυφές (modes)
- Διάμεσος (median)

η περιττός:  $\tilde{x}_{1/2} = \tilde{x}_{(n+1)/2}$

η αρτιος:  $\tilde{x}_{1/2} = \frac{\tilde{x}_{n/2} + \tilde{x}_{n/2+1}}{2}$

- Ποσοστημόρια (Quantiles, Percentiles).



# Παράμετροι Μεταβλητότητας Κατανομής (1 από 3)

- Παράμετροι Μεταβλητότητας Κατανομής (Dispersion Parameters):
  - Εύρος (range) δείγματος,
  - Διασπορά (variance) δείγματος ,
  - Τυπική Απόκλιση (standard deviation) δείγματος,
  - **Αμερόληπτη Τυπική Απόκλιση δείγματος,**
  - Σχετικό Σφάλμα (relative error) δείγματος,
  - **Αμερόληπτο Σχετικό Σφάλμα (relative error) δείγματος,**
  - **Τυπικό Σφάλμα Μέσης Τιμής (Standard Error of the Mean) δείγματος και**
  - Αποστάσεις Ποσοστημορίων
- **Η Αμερόληπτη Τυπική Απόκλιση Το Αμερόληπτο Σχετικό Σφάλμα και το Τυπικό Σφάλμα Μ.Τ. ορίζονται μόνο για δείγματα.**





# Παράμετροι Μεταβλητότητας Κατανομής (2 από 3)

- Εύρος δείγματος

$$x_n - x_1$$

- Διασπορά δείγματος

$$\tilde{\sigma}^2 = \frac{(\chi_1 - \tilde{m})^2 + \dots + (\chi_M - \tilde{m})^2}{M} = (\chi_1 - \tilde{m})^2 \tilde{\rho}_1 + \dots + (\chi_n - \tilde{m})^2 \tilde{\rho}_n$$

- Τυπική Απόκλιση (standard deviation) δείγματος

$$\tilde{\sigma} = \sqrt{\tilde{\sigma}^2} = \sqrt{\frac{(\chi_1 - \tilde{m})^2 + \dots + (\chi_M - \tilde{m})^2}{M}} = \sqrt{(\chi_1 - \tilde{m})^2 \rho_1 + \dots + (\chi_n - \tilde{m})^2 \rho_n}$$



# Παράμετροι Μεταβλητότητας Κατανομής (3 από 3)

- **Θεώρημα**

Η Τυπική Απόκλιση δείγματος δεν είναι Αμερόληπτη Στατιστική παράμετρος.

- Αμερόληπτη Τυπική Απόκλιση δείγματος

$$s = \sqrt{\frac{(\chi_1 - \tilde{m})^2 + \dots + (\chi_M - \tilde{m})^2}{M-1}}$$

$\tilde{m}$  = η Μεση Τιμη του Δειγματος

- Η διόρθωση (M – 1 αντί M) Bessel 1830

$$\lim_{M \rightarrow \infty} (s - \sigma) = 0$$



# Σφάλματα

- **Σχετικό Σφάλμα δείγματος:**

$$\frac{\sigma}{\tilde{m}}$$

- **Αμερόληπτο Σχετικό Σφάλμα (relative error) δείγματος :**

$$\frac{s}{\tilde{m}}$$

- **Τυπικό Σφάλμα Μέσης Τιμής (Standard Error of the Mean) δείγματος:**

$$SE = \frac{s}{\sqrt{N}} = \sqrt{\frac{(\chi_1 - \tilde{m})^2 + \dots + (\chi_M - \tilde{m})^2}{M(M-1)}}$$



# Παράμετροι σχήματος

- **Παράμετροι Σχήματος Κατανομής (Shape Parameters):**
  - Λοξότητα (skewness)
  - Κύρτωση (kurtosis)
- **Εντροπία Δείγματος**

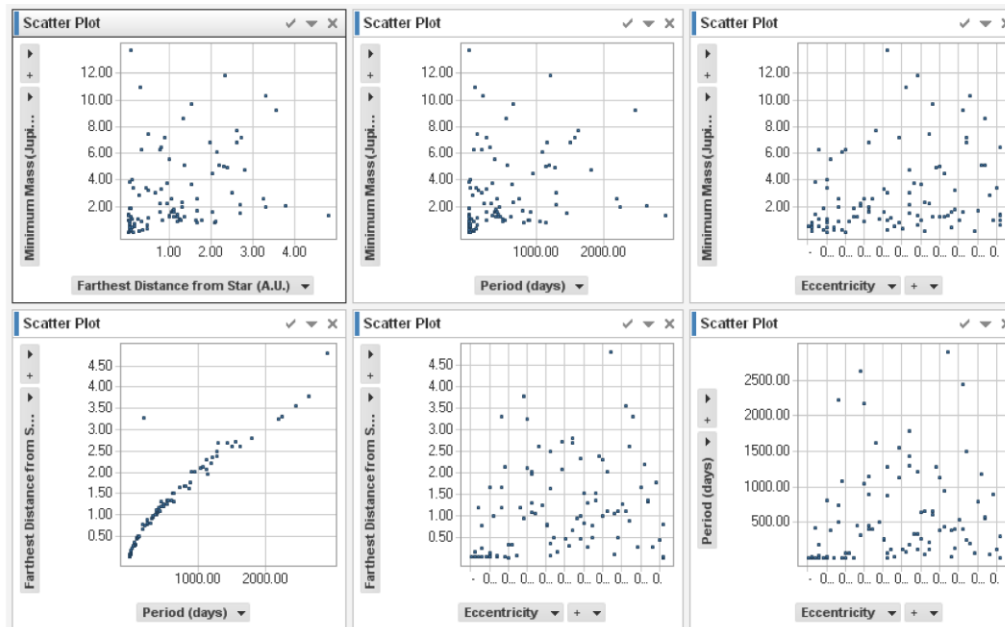
$$\tilde{j} = - \sum_{v=1}^n \tilde{\rho}_v \log_2 \tilde{\rho}_v$$



# Γραμμική Παλινδρόμηση (1 από 2)

- Γραμμική Παλινδρόμηση

Διάγραμμα Διασποράς των μετρήσεων της μεταβλητής  $Y$  ως προς την  $X$   
(ελεγχόμενη)



Scatter Plots of Pairwise Combinations of Extrasolar Planet Data

<http://vis.berkeley.edu/courses/cs294-10-fa07/wiki/index.php/A2-MarkHowison>



# Γραμμική Παλινδρόμηση (2 από 2)

- Πρόβλημα Εύρεσης Σχέσης  $Y = f(X)$  από τις Παρατηρήσεις (Data Fitting)

$X_1, X_2, \dots, X_M$

$$x_1 = \begin{pmatrix} x_{11} \\ x_{12} \\ \vdots \\ x_{1M} \end{pmatrix}, x_2 = \begin{pmatrix} x_{21} \\ x_{22} \\ \vdots \\ x_{2M} \end{pmatrix}, \dots, x_M = \begin{pmatrix} x_{M1} \\ x_{M2} \\ \vdots \\ x_{MM} \end{pmatrix},$$

$f$  : Συνάρτηση Παλινδρόμησης (Regression Function)

- Στην πράξη η εκτιμώμενη συνάρτηση  $f$  προσεγγίζει τις παρατηρούμενες τιμές
- Σφάλματα:

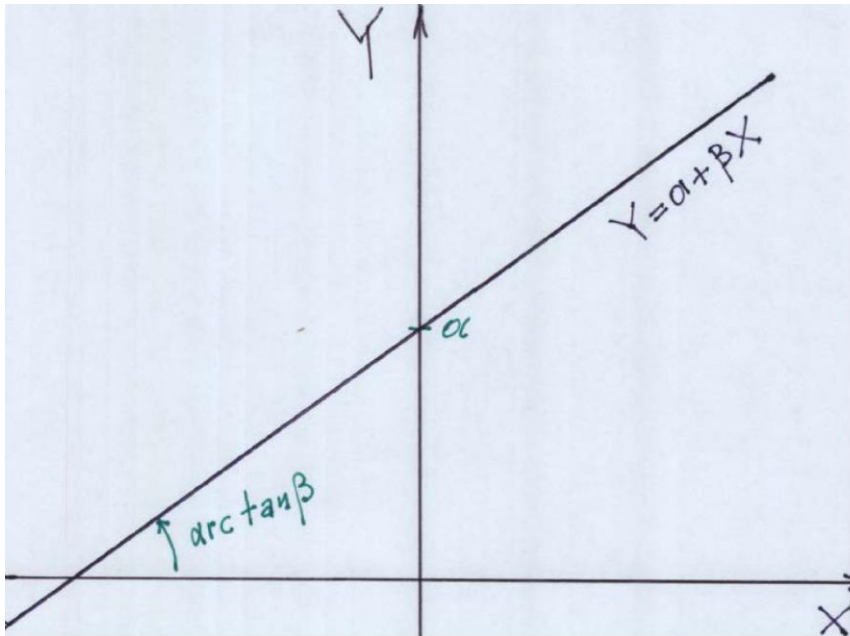
$$e_1 = (y_1 - f(x_1))$$

$$e_2 = (y_2 - f(x_2))$$

$$\dots$$
$$e_M = (y_M - f(x_M))$$



# Εξίσωση Ευθείας Γραμμικής Παλινδρόμησης



- Εξίσωση Ευθείας Γραμμικής Παλινδρόμησης (Linear Regression):  $Y = \alpha + \beta X$



# Συναρτήσεις που Ανάγονται σε Γραμμικές

Αρχική Συνάρτηση $z = f(x)$	Μετασχηματισμός $z = g(y)$	
$z = e^{\alpha + \beta x}$	$z = e^y$	$y = \alpha + \beta x$
$z = Ae^{\beta x}$	$z = e^y, A = e^\alpha$	
$z = 1/(\alpha + \beta x)$	$z = 1/y$	
$z = A\xi^\beta$	$z = e^y, A = e^\alpha, \xi = e^x$	
$z = 1/(\alpha + \beta x)^n$	$z = 1/y^n$	





# Εκτίμηση Παραμέτρων

Εκτίμηση των παραμέτρων  $\alpha, \beta$  της Ευθείας Παλινδρόμησης  
Μέθοδος Ελαχίστων Τετραγώνων Gauss  
Θεώρημα

Η συνάρτηση  $Y=f(X) = \hat{\alpha} + \hat{\beta}X$  με παραμέτρους

$$\hat{\beta} = \frac{\sum_{v=1}^N x_v y_v - M \tilde{m}_X \tilde{m}_Y}{\sum_{v=1}^N (x_v)^2 - N(\sum_{v=1}^N x_v)^2} = \frac{S_{XY}}{(S_X)^2}$$

$$\hat{\alpha} = \tilde{y} - \hat{\beta} \tilde{x}$$

$$\text{οπου: } S_{XY} = \frac{\sum_{v=1}^N x_v y_v - M \tilde{m}_X \tilde{m}_Y}{M-1}$$

$$(S_X)^2 = \frac{\sum_{v=1}^N (x_v)^2 - M \tilde{m}_X^2}{M-1}, (S_Y)^2 = \frac{\sum_{v=1}^N (y_v)^2 - M \tilde{m}_Y^2}{M-1}$$

$$\tilde{m}_X = \frac{\sum_{v=1}^M x_v}{M}, \tilde{m}_Y = \frac{\sum_{v=1}^M y_v}{M}$$

έχει το Ελάχιστο (Ολικό) Τετραγωνικό Σφάλμα:

$$(f(x_1) - y_1)^2 + (f(x_2) - y_2)^2 + \dots + (f(x_M) - y_M)^2 = \varepsilon(\alpha, \beta)$$



# Μέσο Τετραγωνικό Σφάλμα

- **Θεώρημα**

Η ελάχιστη τιμή του Μέσου Τετραγωνικού Σφάλματος είναι:

$$\min \left[ \frac{\varepsilon(\alpha, \beta)}{M} \right] = \tilde{\sigma}_Y^2 (1 - \tilde{r}^2)$$

όπου:

$$\tilde{\sigma}_Y^2 = \frac{\sum_{v=1}^N (y_v)^2 - M \tilde{m}_Y^2}{M}$$
$$\tilde{r} = \tilde{r}(X, Y) = \frac{\text{cov}(X, Y)}{\sqrt{\text{var}X \text{var}Y}} = \frac{\tilde{\sigma}_{XY}}{\tilde{\sigma}_X \tilde{\sigma}_Y} = \frac{\sum_{v=1}^M (x_v - \tilde{m}_X)(y_v - \tilde{m}_Y)}{\sqrt{\sum_{v=1}^M (x_v - \tilde{m}_X)^2} \sqrt{\sum_{v=1}^M (y_v - \tilde{m}_Y)^2}} = \frac{s_{XY}}{s_X s_Y} = \hat{r}(X, Y)$$

- Ο Συντελεστής Pearson (Γραμμικής Συσχέτισης) των Μεταβλητών X, Y.
- Εκτίμηση Διασποράς Γραμμικής Παλινδρόμησης.



# Γραμμική Εξάρτηση

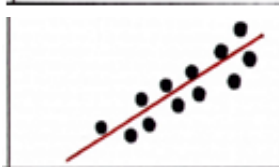
- **Θεώρημα**  $\tilde{r} = \frac{s_{xy}}{s_x s_y} = \hat{\beta} \frac{s_x}{s_y}$

$\tilde{r} \cong 1$   
 $\Rightarrow \hat{\beta} > 0$

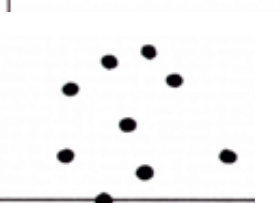


Υπάρχει θετική γραμμική εξάρτηση των X, Y.

$0 < \tilde{r} < 1$

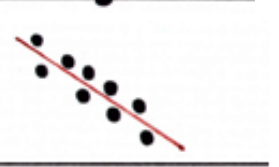


$\tilde{r} = 0$   
 $\Rightarrow \hat{\beta} = 0$

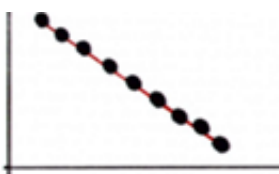


Δεν υπάρχει εξάρτηση των X, Y.

$0 < \tilde{r} < 1$



$\tilde{r} \cong -1$   
 $\Rightarrow \hat{\beta} < 0$



Υπάρχει αρνητική γραμμική εξάρτηση των X, Y.



# Μέθοδοι Παλινδρόμησης

- ΣΧΟΛΙΑ

1. Ανακάλυψη Μεθόδου Ελαχίστων Τετραγώνων Gauss  
1795 (18 ετων)

2. Ανακάλυψη Ποσειδώνα από την τροχιά του Ουρανού  
1846 Με την Μέθοδο Ελαχίστων Τετραγώνων

3. Γενικά Γραμμικά Μοντέλα:

$$Y = \alpha + \beta_1 X_1 + \beta_2 X_2 + \dots + \beta_N X_N$$

4. Μη Γραμμική Παλινδρόμηση

5. Παρεμβολή με κλάσεις συναρτήσεων (Πολυώνυμα, Τριγωνομετρικές Συναρτησεις, Wavelets)

6. Δειγματοληψία Shannon



# Συντελεστής Αλληλοεξάρτησης Αμοιβαίας Πληροφορίας Δείγματος

- **Συντελεστής Αλληλοεξάρτησης Αμοιβαίας Πληροφορίας Δείγματος**

$$\tilde{r}_{MI} = \frac{\tilde{J}[X] + \tilde{J}[Y] - \tilde{J}[X, Y]}{\min(\tilde{J}[X], \tilde{J}[Y])}$$

όπου  $\tilde{J}[X] + \tilde{J}[Y] - \tilde{J}[X, Y] = \sum_{x,y} \tilde{\rho}(x,y) \log_2 \left( \frac{\tilde{\rho}(x,y)}{\tilde{\rho}(x)\tilde{\rho}(y)} \right)$

- Η Εντροπία Δείγματος της Μεταβλητής X

$$\tilde{J}[X] = - \sum_x \tilde{\rho}(x) \log_2 \tilde{\rho}(x)$$

- Η Εντροπία Δείγματος της Μεταβλητής Y

$$\tilde{J}[Y] = - \sum_y \tilde{\rho}(y) \log_2 \tilde{\rho}(y)$$

- Η Κοινή Εντροπία Δείγματος των Μεταβλητών X,Y

$$\tilde{J}[X, Y] = \sum_{x,y} \tilde{\rho}(x,y) \log_2 \tilde{\rho}(x,y)$$



# Παράδειγμα (1 από 3)

- Συσχέτιση Χρώματος Οφθαλμών με το Χρώμα Μαλλιών των Φοιτητών του 2012-13. Οι Υπολογισμοί έγιναν από τους κ. Ρ.Ν. Τασάκη και Ε. Καραπούλια Πρωτοετείς Φοιτητές Βιολογίας ΑΠΘ.

Έστω

**X= Χρώμα Οφθαλμών**

Τιμές: Κ= Καφέ, Γ= Γαλάζιο, ΚΠ= Καστανοπράσινο, Π= Πράσινο, ΓΠ= Γαλαζοπράσινο

**Y= Χρώμα Μαλλιών**

Τιμές: μ= Μαύρο, ξ= Ξανθό, κ= Καστανό, κξ= Καστανόξανθο



# Παράδειγμα (2 από 3)

## Αποτέλεσμα Παρατηρήσεων

$X \backslash Y$	Καφέ (Κ)	Γαλάζια (Γ)	Καστανο- πράσινα (ΚΠ)	Πράσινα (Π)	Γαλαζο- πράσινα (ΓΠ)	$Y$ ↓
Μαύρα ( $\mu$ )	$p_{\mu,K} = 4/52$	$p_{\mu,\Gamma} = 0/52$	$p_{\mu,ΚΚ} = 3/52$	$p_{\mu,Κ} = 1/52$	$p_{\mu,\GammaΚ} = 0/52$	$p_{\mu} = 8/52$
Ξανθά ( $\xi$ )	$p_{\xi,K} = 0/52$	$p_{\xi,\Gamma} = 0/52$	$p_{\xi,ΚΚ} = 1/52$	$p_{\xi,Κ} = 0/52$	$p_{\xi,\GammaΚ} = 0/52$	$p_{\xi} = 1/52$
Καστανά ( $\kappa$ )	$p_{\kappa,K} = 26/52$	$p_{\kappa,\Gamma} = 1/52$	$p_{\kappa,ΚΚ} = 2/52$	$p_{\kappa,Κ} = 2/52$	$p_{\kappa,\GammaΚ} = 0/52$	$p_{\kappa} = 31/52$
Καστανό- Ξανθα ( $\kappa\xi$ )	$p_{\kappa\xi,K} = 4/52$	$p_{\kappa\xi,\Gamma} = 2/52$	$p_{\kappa\xi,ΚΚ} = 2/52$	$p_{\kappa\xi,Κ} = 3/52$	$p_{\kappa\xi,\GammaΚ} = 1/52$	$p_{\kappa\xi} = 12/52$
$X \rightarrow$	$p_K = 34/52$	$p_{\Gamma} = 3/52$	$p_{ΚΚ} = 8/52$	$p_{Κ} = 6/52$	$p_{\GammaΚ} = 1/52$	1



# Παράδειγμα (3 από 3)

- Αποτέλεσμα Παρατηρήσεων
- Υπολογισμός Εντροπιών

$$\tilde{r}_{MI} = \frac{\tilde{J}[X; Y]}{\min(\tilde{J}[X], \tilde{J}[Y])} = \frac{\tilde{J}[X] + \tilde{J}[Y] - \tilde{J}[X, Y]}{\min(\tilde{J}[X], \tilde{J}[Y])}$$

$$\tilde{J}[X] = -\frac{34}{52} \cdot \log_2 \frac{34}{52} - \frac{3}{52} \cdot \log_2 \frac{3}{52} - \frac{8}{52} \cdot \log_2 \frac{8}{52} - \frac{6}{52} \cdot \log_2 \frac{6}{52} - \frac{1}{52} \cdot \log_2 \frac{1}{52} \approx 1,523$$

$$\tilde{J}[Y] = -\frac{8}{52} \cdot \log_2 \frac{8}{52} - \frac{1}{52} \cdot \log_2 \frac{1}{52} - \frac{31}{52} \cdot \log_2 \frac{31}{52} - \frac{12}{52} \cdot \log_2 \frac{12}{52} \approx 1,458$$

$$\tilde{J}[X, Y] = -2 \cdot \frac{4}{52} \log_2 \frac{4}{52} - \frac{26}{52} \log_2 \frac{26}{52} - 4 \cdot \frac{1}{52} \log_2 \frac{1}{52} - 4 \cdot \frac{2}{52} \log_2 \frac{2}{52} - 2 \cdot \frac{3}{52} \log_2 \frac{3}{52} \approx 2,706$$

$$\tilde{J}[X; Y] = \tilde{J}[X] + \tilde{J}[Y] - \tilde{J}[X, Y] = 1,523 + 1,458 - 2,706 = 0,275$$

$$\min(\tilde{J}[X], \tilde{J}[Y]) = \tilde{J}[Y] = \mathbf{1,458}$$

$$\tilde{r}_{MI} = \frac{\mathbf{1,523} + \mathbf{1,458} - \mathbf{2,706}}{\mathbf{1,458}} = \mathbf{0,18861454}$$





# Σημείωμα Αδειοδότησης

Το παρόν υλικό διατίθεται με τους όρους της άδειας χρήσης Creative Commons Αναφορά - Παρόμοια Διανομή [1] ή μεταγενέστερη, Διεθνής Έκδοση. Εξαιρούνται τα αυτοτελή έργα τρίτων π.χ. φωτογραφίες, διαγράμματα κ.λ.π., τα οποία εμπεριέχονται σε αυτό και τα οποία αναφέρονται μαζί με τους όρους χρήσης τους στο «Σημείωμα Χρήσης Έργων Τρίτων».



Ο δικαιούχος μπορεί να παρέχει στον αδειοδόχο ξεχωριστή άδεια να χρησιμοποιεί το έργο για εμπορική χρήση, εφόσον αυτό του ζητηθεί.

[1] <http://creativecommons.org/licenses/by-sa/4.0/>





# Τέλος ενότητας

Επεξεργασία: Βασιλική Αλμπανίδου  
Θεσσαλονίκη, Χειμερινό Εξάμηνο 2014-2015



Ευρωπαϊκή Ένωση  
Ευρωπαϊκό Κοινωνικό Ταμείο



ΥΠΟΥΡΓΕΙΟ ΠΑΙΔΕΙΑΣ ΚΑΙ ΘΡΗΣΚΕΥΜΑΤΩΝ  
ΕΙΔΙΚΗ ΥΠΗΡΕΣΙΑ ΔΙΑΧΕΙΡΙΣΗΣ

Με τη συγχρηματοδότηση της Ελλάδας και της Ευρωπαϊκής Ένωσης



ΕΥΡΩΠΑΪΚΟ ΚΟΙΝΩΝΙΚΟ ΤΑΜΕΙΟ

# Διατήρηση Σημειωμάτων

Οποιαδήποτε αναπαραγωγή ή διασκευή του υλικού θα πρέπει να συμπεριλαμβάνει:

- το Σημείωμα Αναφοράς
- το Σημείωμα Αδειοδότησης
- τη δήλωση Διατήρησης Σημειωμάτων
- το Σημείωμα Χρήσης Έργων Τρίτων (εφόσον υπάρχει)

μαζί με τους συνοδευόμενους υπερσυνδέσμους.

