

Υπολογιστική Φυσική Στοιχειωδών Σωματιδίων

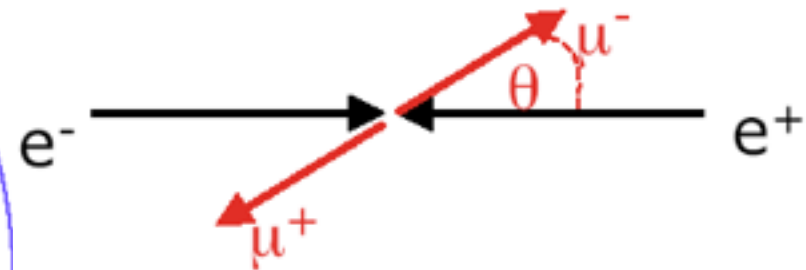
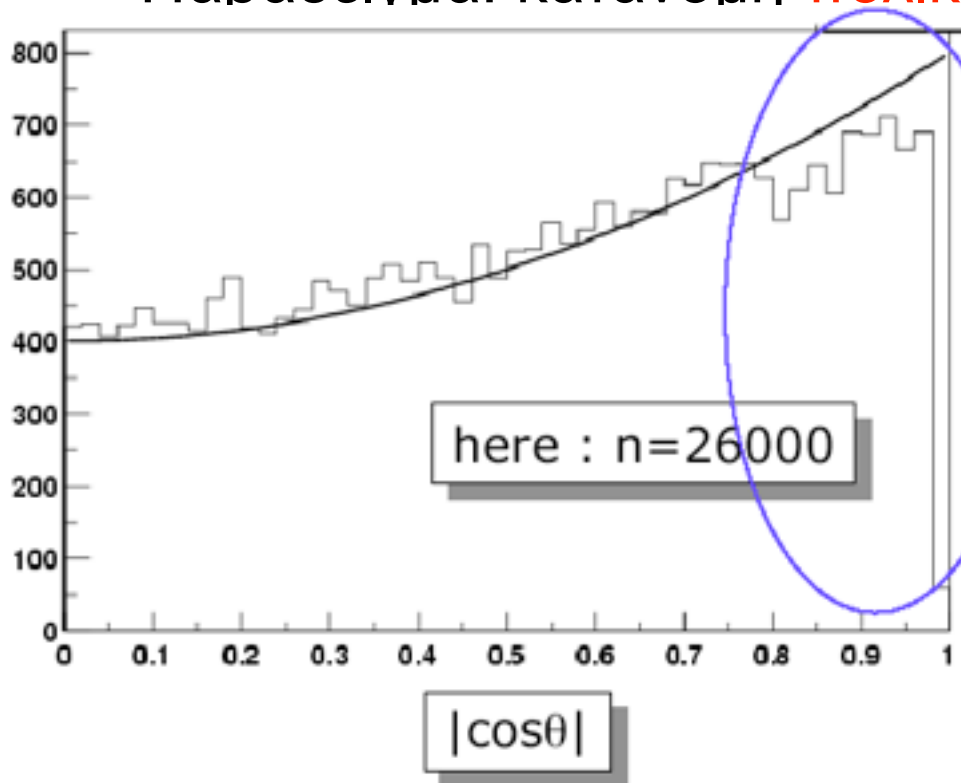
Εαρινό Εξάμηνο 2016–2017

Ιστογράμματα
(Histograms)

Χαρά Πετρίδου

Ιστογράμματα

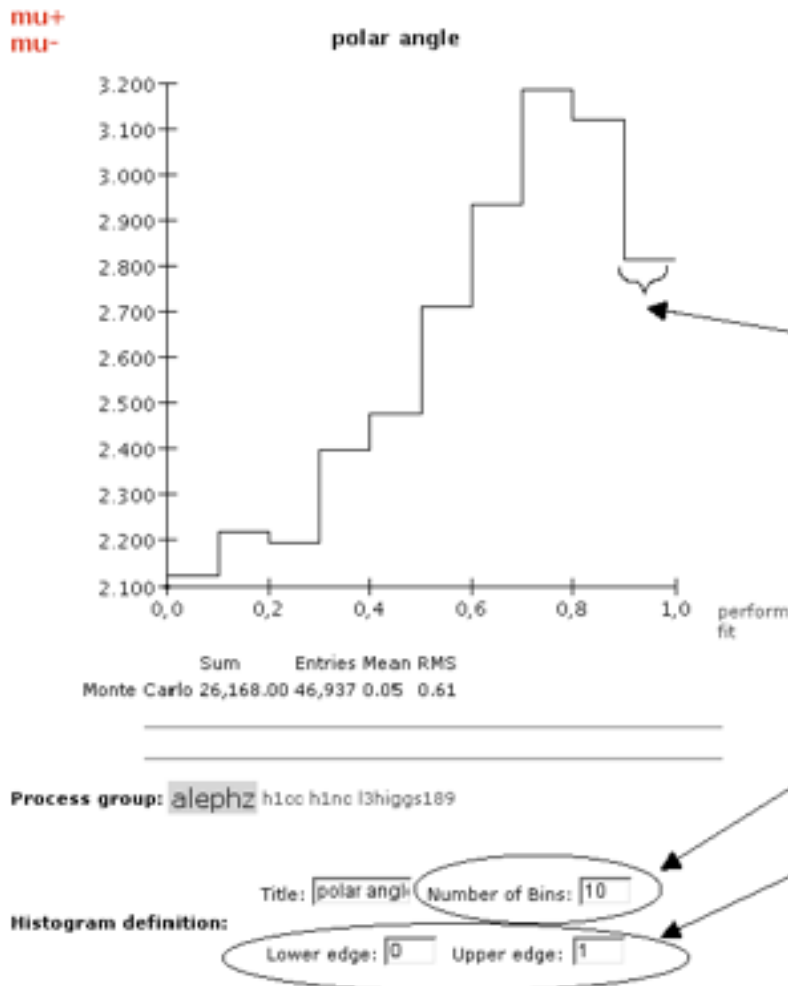
- Συνήθως οι μεταβλητές που χρησιμοποιούμε είναι διακριτές μεταβλητές x_1, x_2, \dots, x_n που προέρχονται από πειραματικές μετρήσεις
- Παράδειγμα: κατανομή **πολικής γωνίας θ**



Σύγκριση των δεδομένων με ομαλή-συνεχή κατανομή

Ιστογράμματα

- Entries, normalization,...
- Bin width(s) (στο παράδειγμα ισαπέχουσες υποδιαίρέσεις-bins)
- Πλήθος των bins
- Περιοχή (Range)



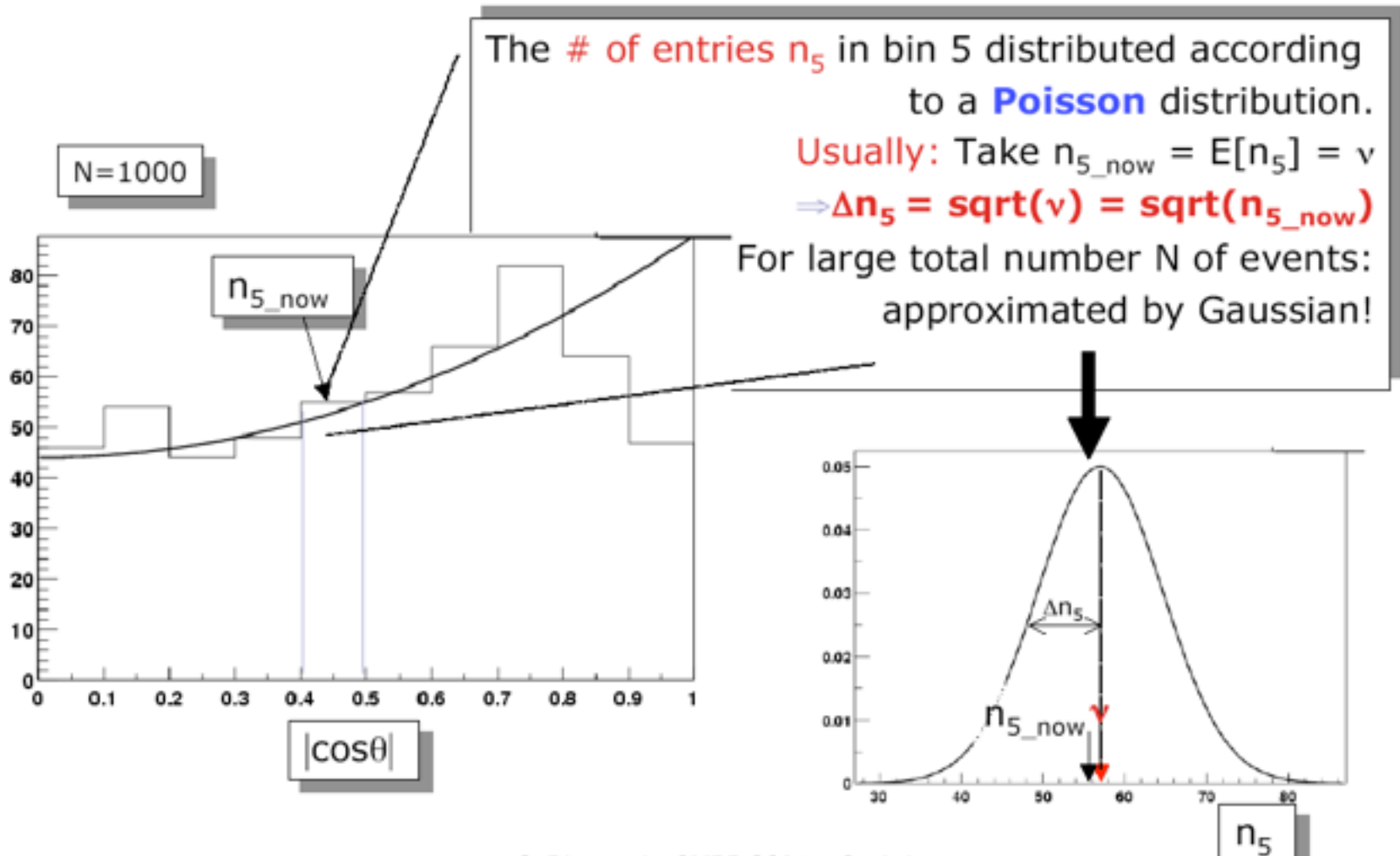
Ιστογράμματα

- Εάν N (*entries*) μεγάλος αριθμός το ιστόγραμμα αντιπροσωπεύει κατανομή πιθανότητας (pdf)

Ενδιαφέρουν:

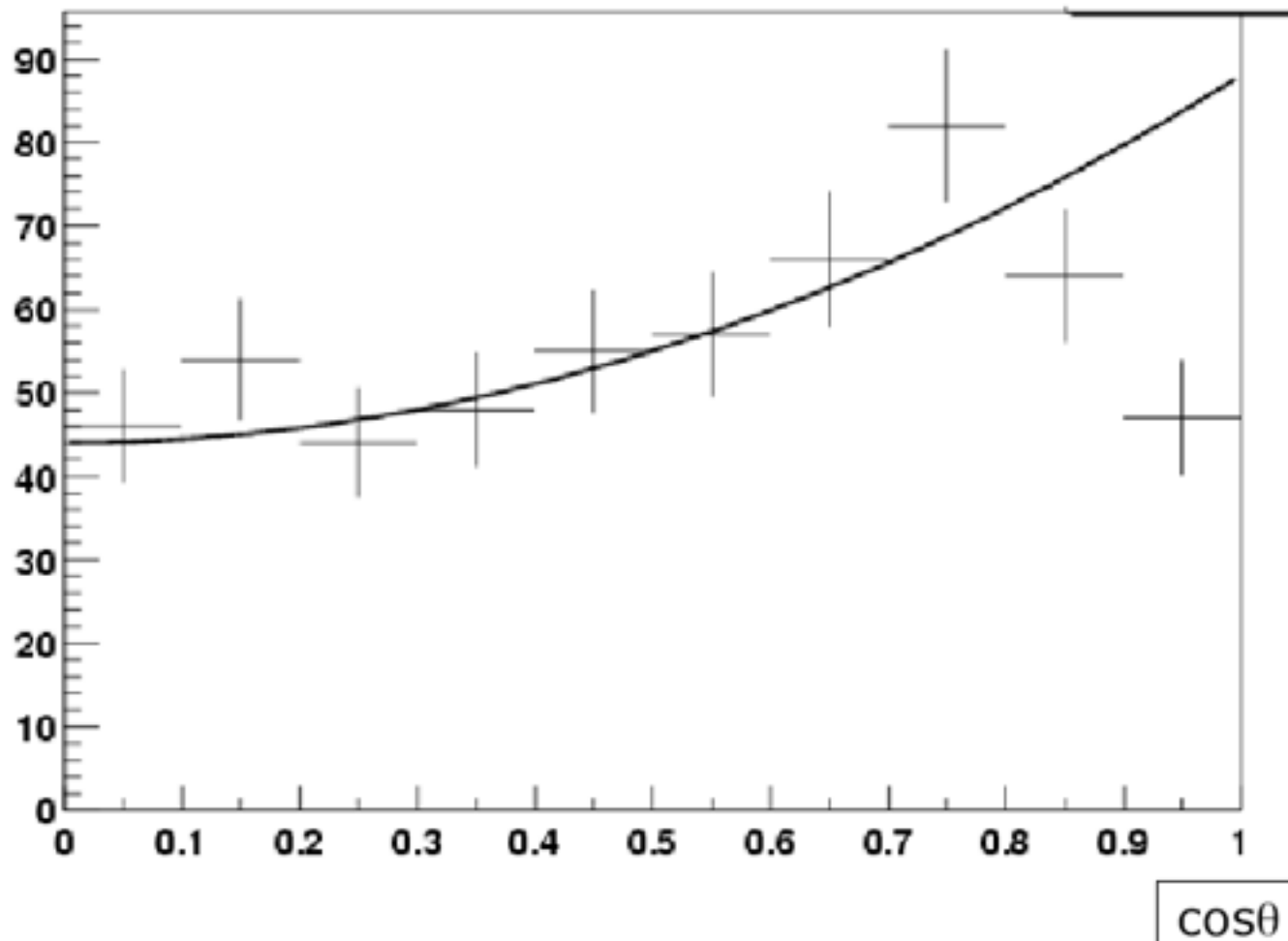
- το πλήθος των γεγονότων σε κάθε bin και το σφάλμα
- Η κανονικοποίηση του ιστογράμματος
- Η μέση τιμή του ιστογράμματος

Σφάλματα στο Ιστόγραμμα



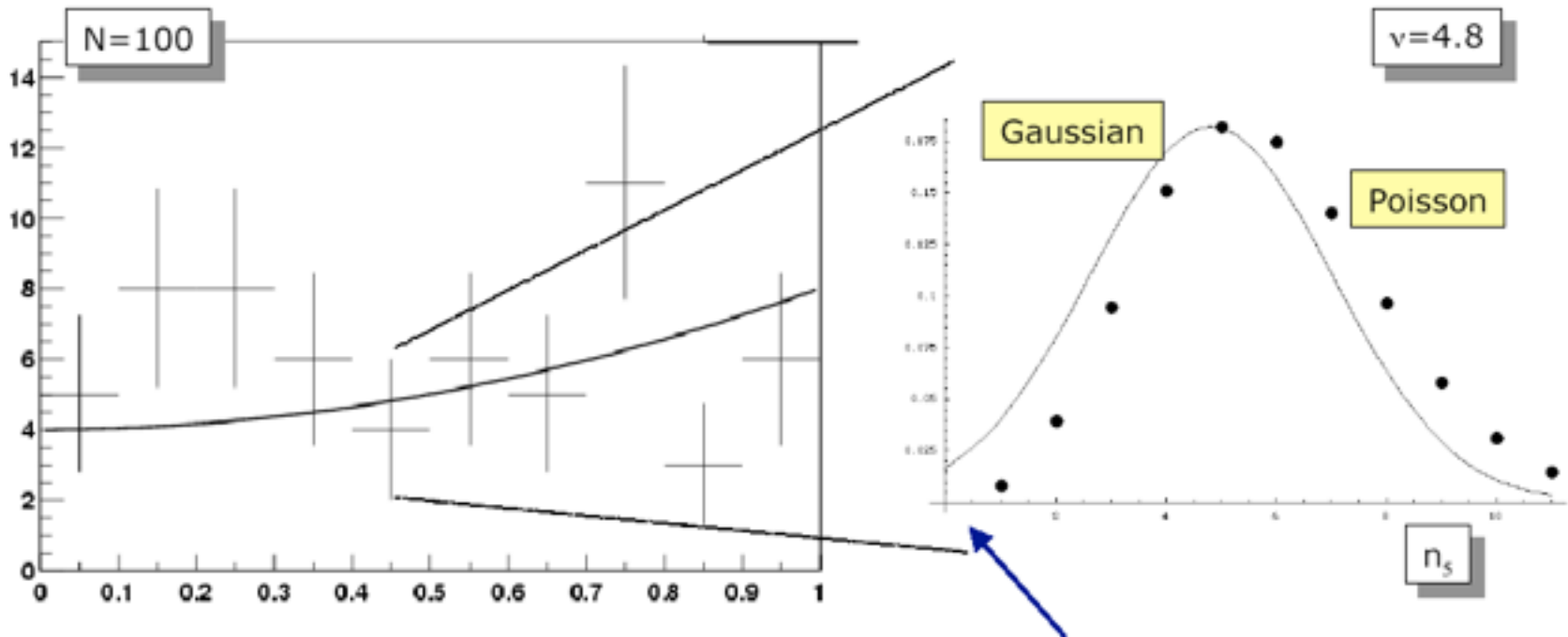
Σφάλματα στο Ιστόγραμμα

- Προσθέτουμε **error bars** στο ιστόγραμμα



Σφάλματα στο Ιστόγραμμα

- Προσοχή! Περίπτωση μικρού πλήθους γεγονότων

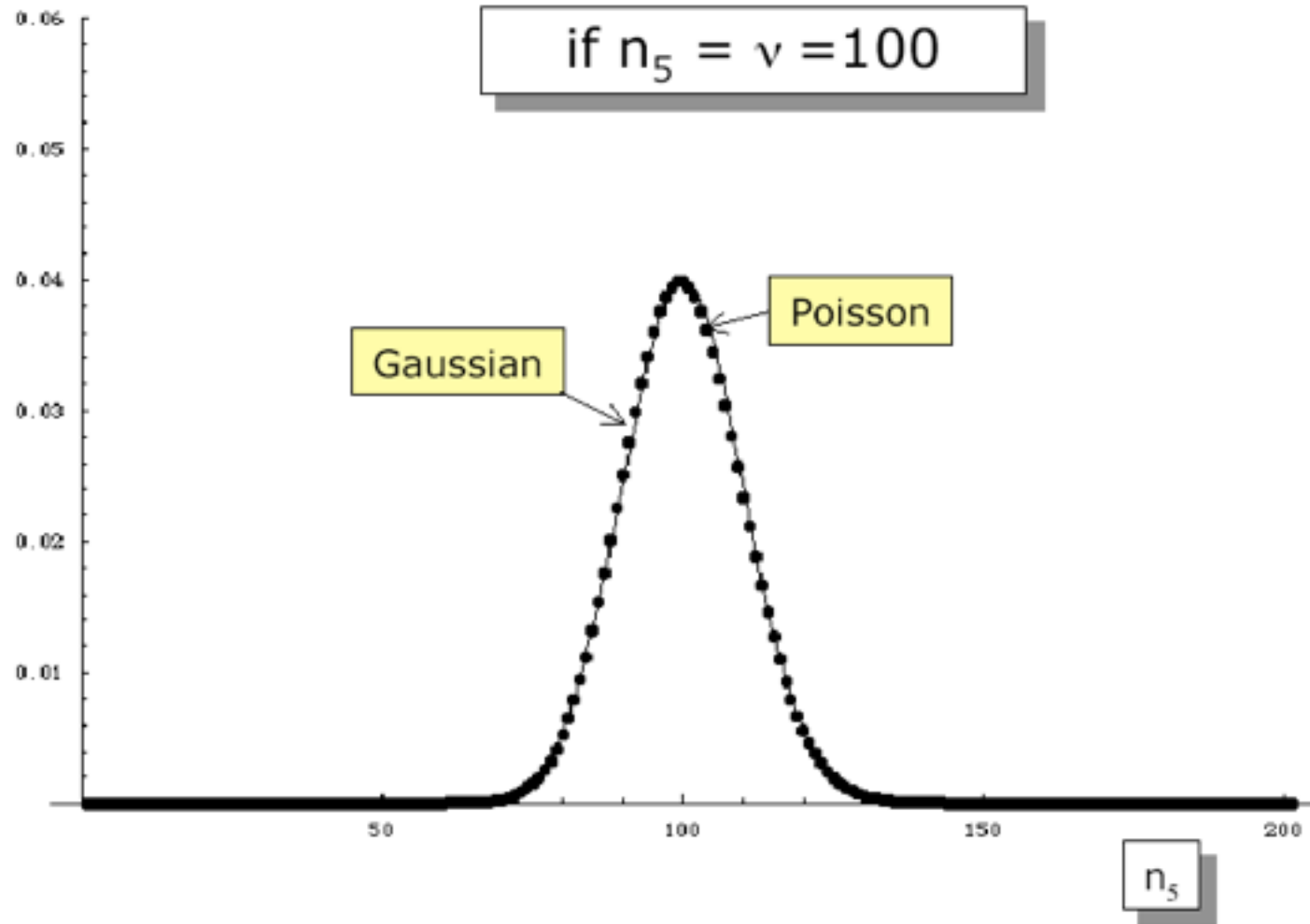


$\cos\theta$

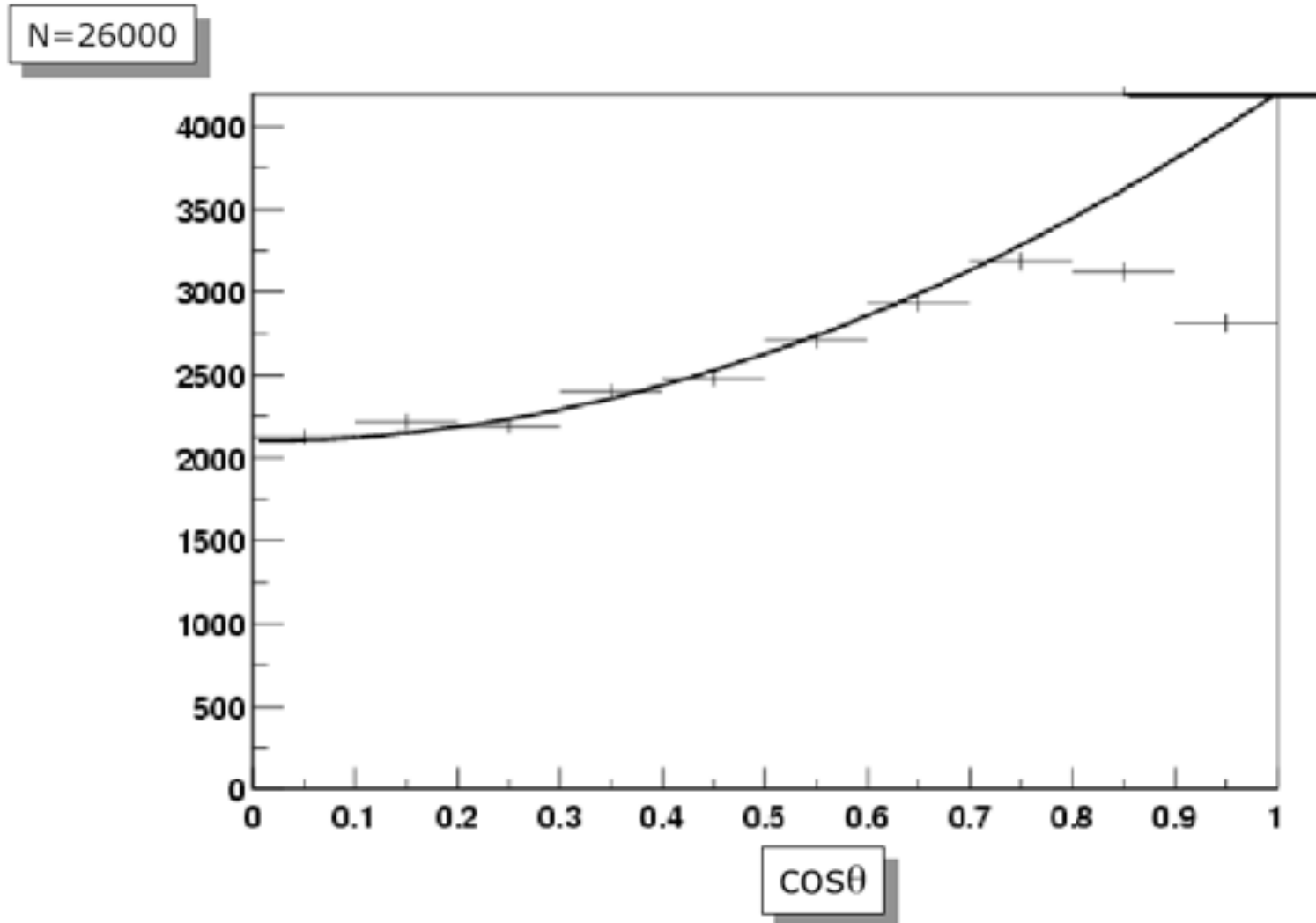
Prob(to see 0) \ll 0 for Gaussian

Prob(to see 0) $=$ 0 for Poisson !!

Poisson vs Gaussian



Σφάλματα στο Ιστόγραμμα



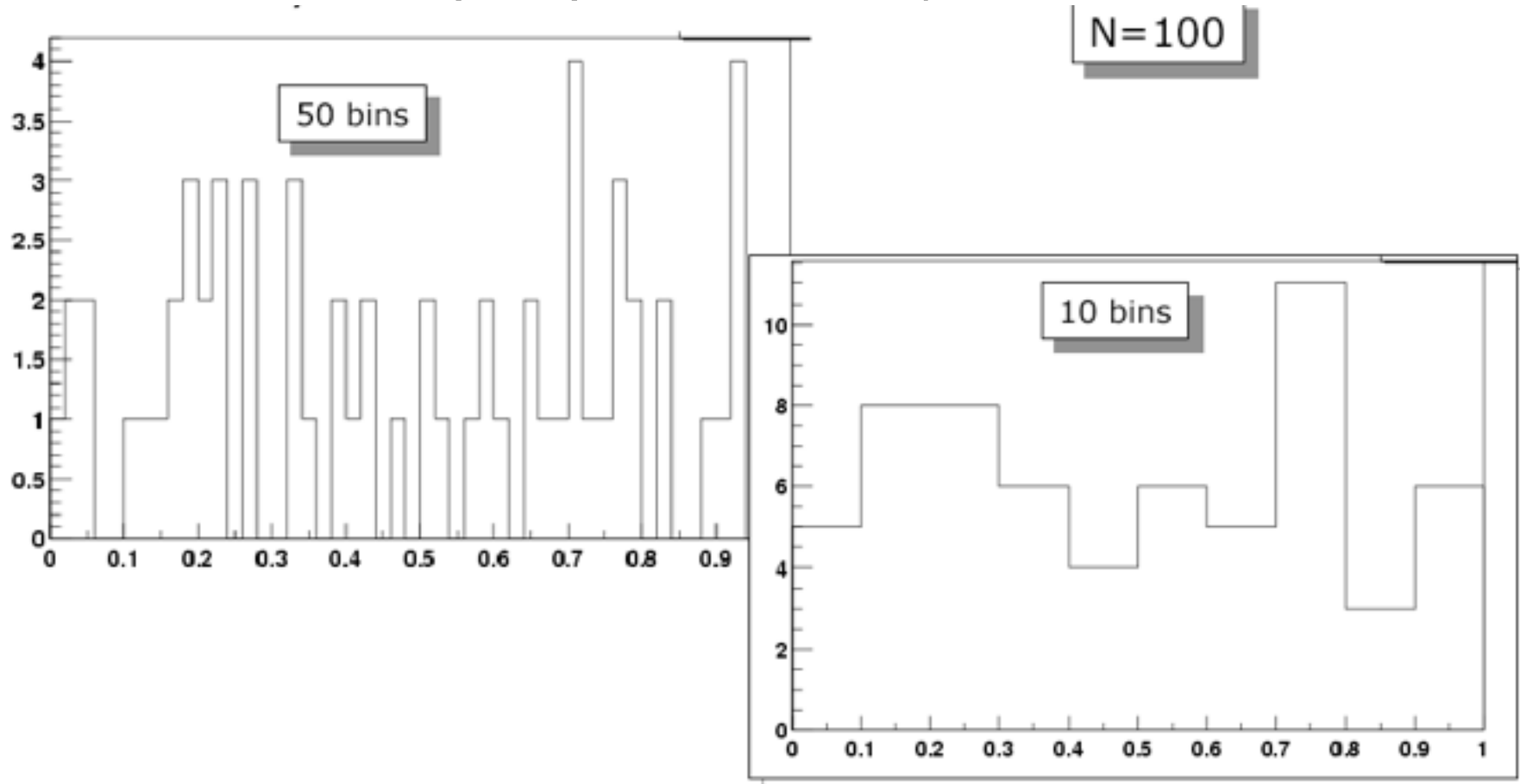
Ιστογράμματα

Τι θα πρέπει να προσέχουμε?

- Το μέγεθος του bin (**bin size**)
- Γεγονότα εκτός των ορίων του ιστογράμματος (**underflows, overflows**)
- Κατανομές με γρήγορες μεταβολές (**steeply falling/fast varying functions**)

Επιλογή του bin width

- Επιλέγουμε την υποδιαίρεση της κλίμακας (bin width) έτσι ώστε το πλήθος των γεγονότων ανά υποδιαίρεση να είναι 'λογικό'

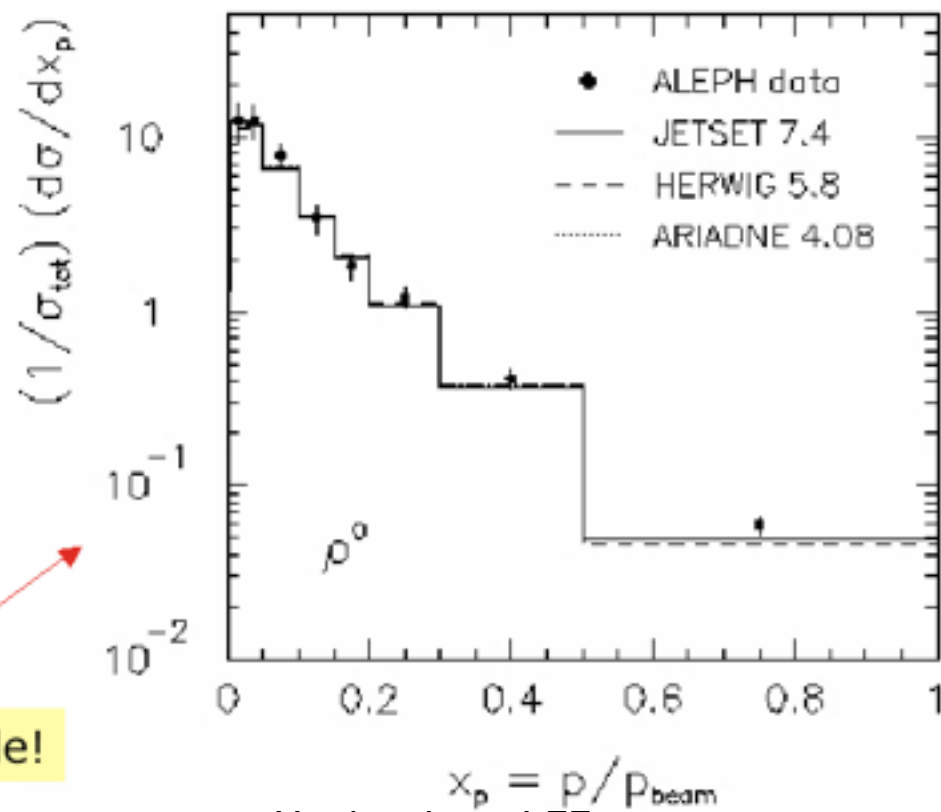


Επιλογή του bin width

- Ο αριθμός των γεγονότων που μετατοπίζονται εντός και εκτός της υποδιαίρεσης πρέπει να είναι παρόμοιος (**bin migration**–σταθερότητα και καθαρότητα του δείγματος)
- Το bin width πρέπει να ταιριάζει με την πειραματική διακριτική ικανότητα στην συγκεκριμένη μεταβλητή (**experimental resolution**)
- Να υπάρχει αρκετή στατιστική για κάθε υποδιαίρεση (**bin statistics**)

Επιλογή του bin width

- Σημαντικό για συναρτήσεις με απότομη κλίση
- Παράδειγμα: η κατανομή της ορμής του σωματιδίου-συντονισμού ρ



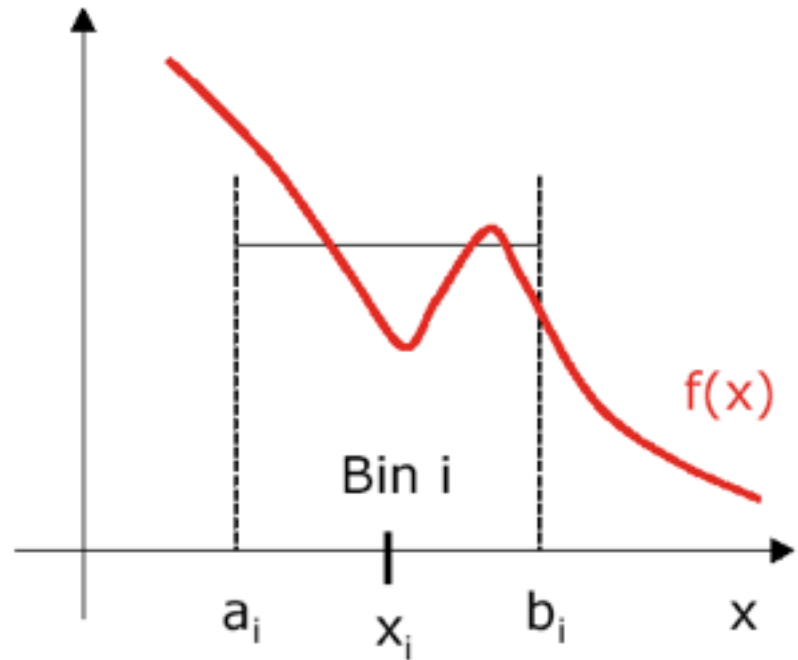
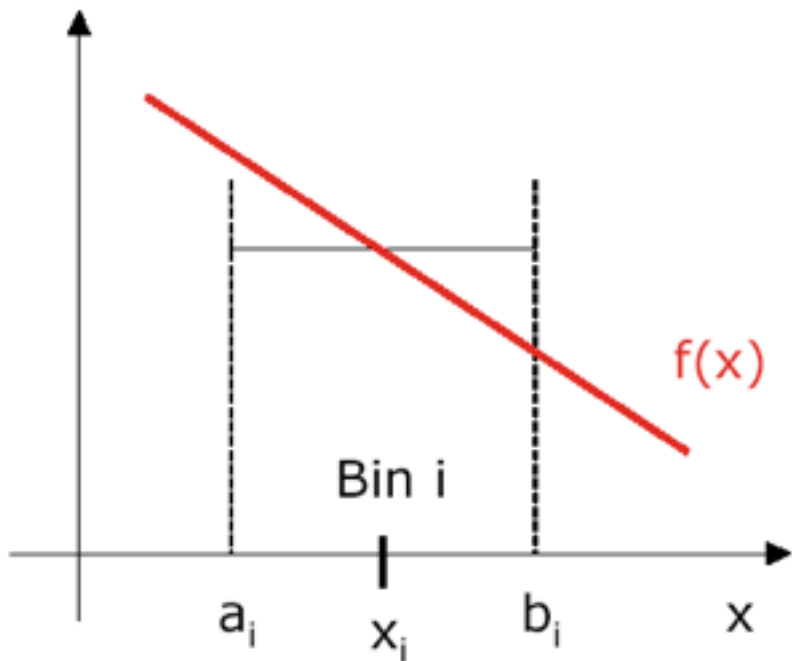
logarithmic scale!

Επιλογή της κλίμακας (Histogram range)

- Κακή επιλογή κλίμακας μπορεί να οδηγήσει σε παραπλανητικά συμπεράσματα
- =>Πρέπει να ελέγχουμε να μην υπάρχουν **overflows** και **underflows**
- Η κλίμακα είναι σημαντική και στην περίπτωση κανονικοποίησης

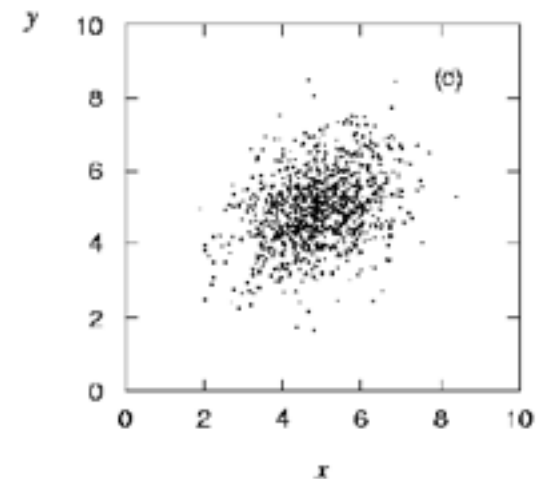
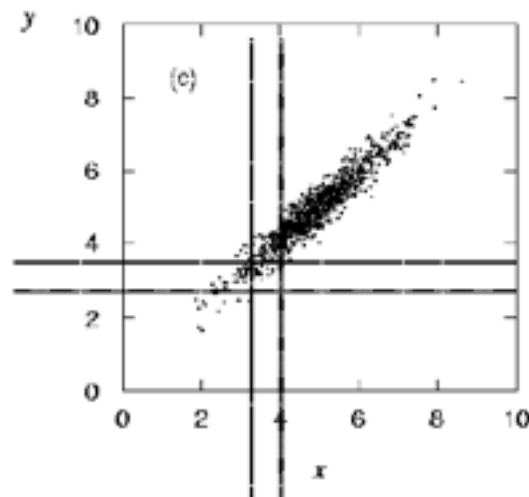
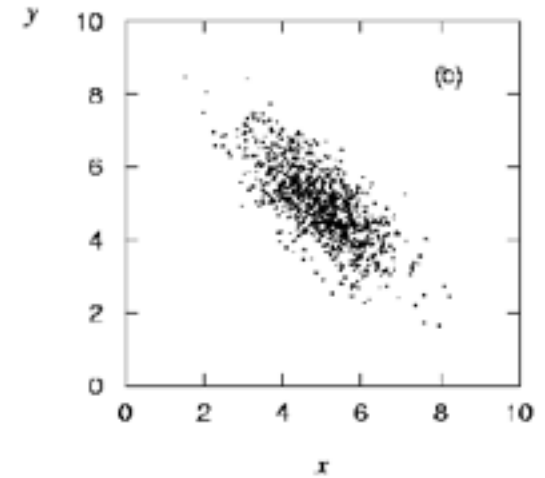
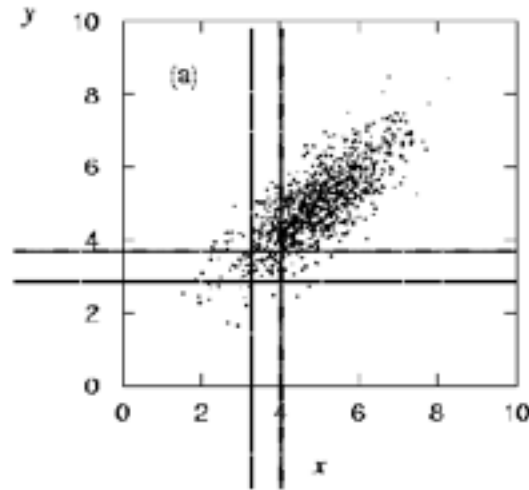
Σύγκριση ιστογράμματος με ομαλή συνάρτηση

- Προσοχή στις συναρτήσεις με απότομες ή γρήγορες μεταβολές



Συσχετίσεις-Συμμεταβολή (correlations-covariance)

- Ιστογράμματα 2-διαστάσεων (scatterplots)
- =>Χρήσιμο εργαλείο για να μελετήσουμε συσχετισμούς δύο μεταβλητών (correlations)



Ιδιότητες-Συσχετισμοί (correlations)

- Για συναρτήσεις με διάφορες μεταβλητές μπορεί να υπάρχουν συσχετίσεις μεταξύ των μεταβλητών
- Παράδειγμα : για συναρτήσεις δύο τυχαίων μεταβλητών x, y η συμμεταβολή (covariance) $V(x,y)$ (ή $\text{cov}(x,y)$) -error matrix- ορίζεται ως:

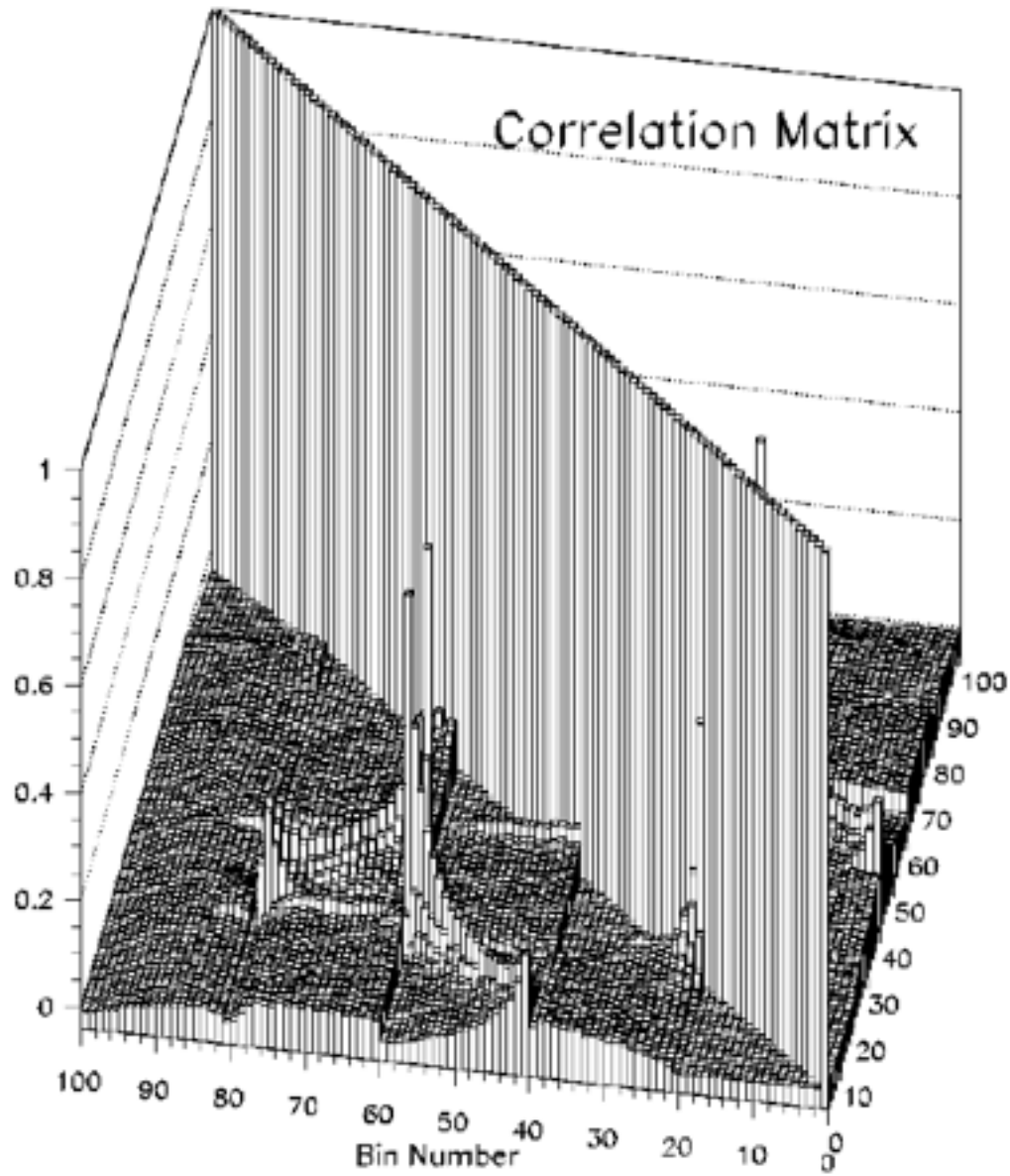
$$V[xy] = E((x - \mu_x)(y - \mu_y)) \equiv E[xy] - \mu_x \mu_y$$

$$E[xy] = \int_{y \min}^{y \max} \int_{x \min}^{x \max} x \cdot y \cdot f(x,y) \cdot dx \cdot dy$$

$$V(x,y) = \begin{pmatrix} \sigma_x^2 & V[xy] \\ V[yx] & \sigma_y^2 \end{pmatrix}$$

- Ο παράγοντας συσχέτισης ορίζεται ως:

$$\rho_{xy} = \frac{V[xy]}{\sigma_x \sigma_y} \quad -1 \leq \rho_{xy} \leq +1 \quad \rho=1 \text{ για } y=ax+b$$



Ιδιότητες-Συσχετισμοί (correlations)

- Υπολογισμός του σφάλματος για συναρτήσεις με συσχετιζόμενες μεταβλητές: **παίρνουμε υπόψη την συσχέτιση**

$$\sigma_f^2 = \sum_{i=1}^n \left(\frac{\partial f}{\partial x_i} \right)^2 \cdot \sigma_{x_i}^2 + \sum_{i=1}^n \sum_{j \neq i}^n \left(\frac{\partial f}{\partial x_i} \frac{\partial f}{\partial x_j} \right) \cdot \text{cov}(x_i, x_j)$$