

Νόμος των Μεγάλων Αριθμών

Έστω ότι $\{X_1, X_2, \dots, X_N\}$ είναι ένα σύνολο «αμοιβαία ανεξάρτητων» τυχαίων μεταβλητών, η καθεμία με αναμενόμενη τιμή ίση με μ (κοινή για όλες τις τυχαίες μεταβλητές) και διασπορές ίσες με σ^2_i , όπου $i=1, 2, \dots, N$ αντίστοιχα.

Ο ΑΣΘΕΝΗΣ ΝΟΜΟΣ των μεγάλων αριθμών διατείνεται ότι, αν
$$\lim_{N \rightarrow \infty} \left[\frac{1}{N^2} \sum_{i=1}^N \sigma_i^2 \right] = 0$$

τότε ο μέσος όρος συγκλίνει στην αναμενόμενη τιμή, μ , των τυχαίων μεταβλητών:

$$\lim_{N \rightarrow \infty} \langle X \rangle = \lim_{N \rightarrow \infty} \left(\frac{1}{N} \sum_{i=1}^N x_i \right) = \mu$$

Ο ΙΣΧΥΡΟΣ ΝΟΜΟΣ των μεγάλων αριθμών διατείνεται ότι, αν
$$\lim_{N \rightarrow \infty} \left[\sum_{i=1}^N \left(\frac{\sigma_i^2}{i} \right) \right] = \text{πεπερασμενο}$$

τότε ο μέσος όρος συγκλίνει στην αναμενόμενη τιμή, μ , των τυχαίων μεταβλητών:
$$\lim_{N \rightarrow \infty} \left(\frac{1}{N} \sum_{i=1}^N x_i \right) = \mu$$

Με την προϋπόθεση ότι υπάρχουν
οι αλγεβρικές ροπές μέχρι και τετάρτης τάξης

$$\left(\frac{1}{N} \cdot \sum_{i=1}^N (x_i - \mu_X)^2 \right) = \left(\frac{1}{N} \cdot \sum_{i=1}^N (x_i)^2 \right) - \mu_X^2 = \hat{\sigma}^2$$

$$\begin{aligned} \sigma^2 &= \lim_{N \rightarrow \infty} \left(\frac{1}{N} \sum_{i=1}^N (x_i - \mu_X)^2 \right) \\ &= \lim_{N \rightarrow \infty} \left(\frac{1}{N} \sum_{i=1}^N (x_i)^2 \right) - \lim_{N \rightarrow \infty} \left(\frac{2 \cdot \mu_X}{N} \cdot \sum_{i=1}^N x_i \right) + \mu_X^2 \\ &= \lim_{N \rightarrow \infty} \left(\frac{1}{N} \sum_{i=1}^N (x_i)^2 \right) - \mu_X^2 \end{aligned}$$

Ολοκλήρωση Monte Carlo

Προκειμένου να υπολογίσουμε το ολοκλήρωμα $I = \int_a^b g(u) du$

με Monte Carlo ολοκλήρωση, επιλέγουμε ισοπίθανα N πραγματικούς αριθμούς u_i ($i=1, 2, 3, \dots, N$) στο διάστημα ολοκλήρωσης $[a,b]$, υπολογίζουμε τις αντίστοιχες N τιμές της συνάρτησης $g(u_i)$ και εφαρμόζουμε τη σχέση

$$I \approx \frac{b-a}{N} \cdot \sum_{i=1}^N g(u_i) = \hat{I}$$

Η ποσότητα $\hat{I}$ είναι τυχαία μεταβλητή με αναμενόμενη τιμή ίση με την πραγματική τιμή του ολοκληρώματος, $I=E[g(u)](b-a)$, και διασπορά ίση με:

$$\sigma_{\hat{I}}^2 = \frac{1}{N} \cdot \lim_{M \rightarrow \infty} \frac{1}{M} \cdot \sum_{j=1}^M \left(I - g(u_j) \cdot (b-a) \right)^2$$

$$\begin{aligned} \sigma_{\hat{I}}^2 &\approx \hat{\sigma}_{\hat{I}}^2 = \frac{1}{N} \cdot \frac{1}{N-1} \cdot \sum_{j=1}^N \left(\hat{I} - g(u_j) \cdot (b-a) \right)^2 \\ &= (b-a)^2 \cdot \frac{1}{N} \cdot \frac{1}{N-1} \sum_{j=1}^N \left\{ \langle g(u) \rangle - g(u_j) \right\}^2 \\ &= (b-a)^2 \cdot \frac{1}{N-1} \cdot \left[\left\langle \left(g(u) \right)^2 \right\rangle - \left(\langle g(u) \rangle \right)^2 \right] \end{aligned}$$

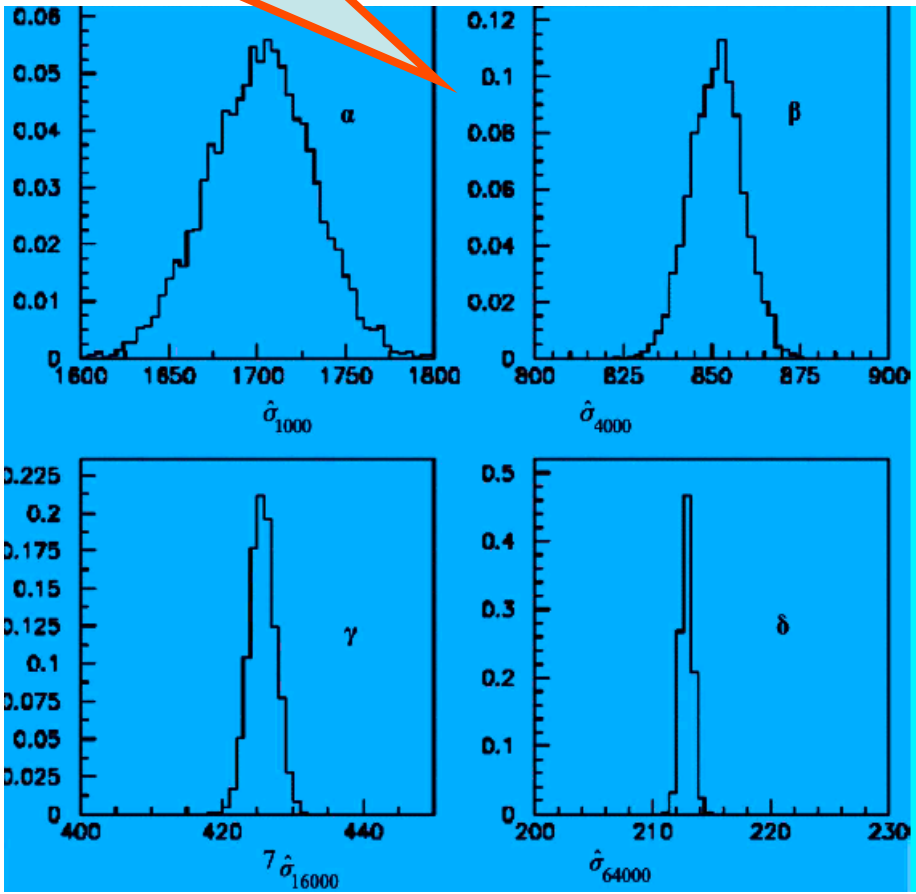
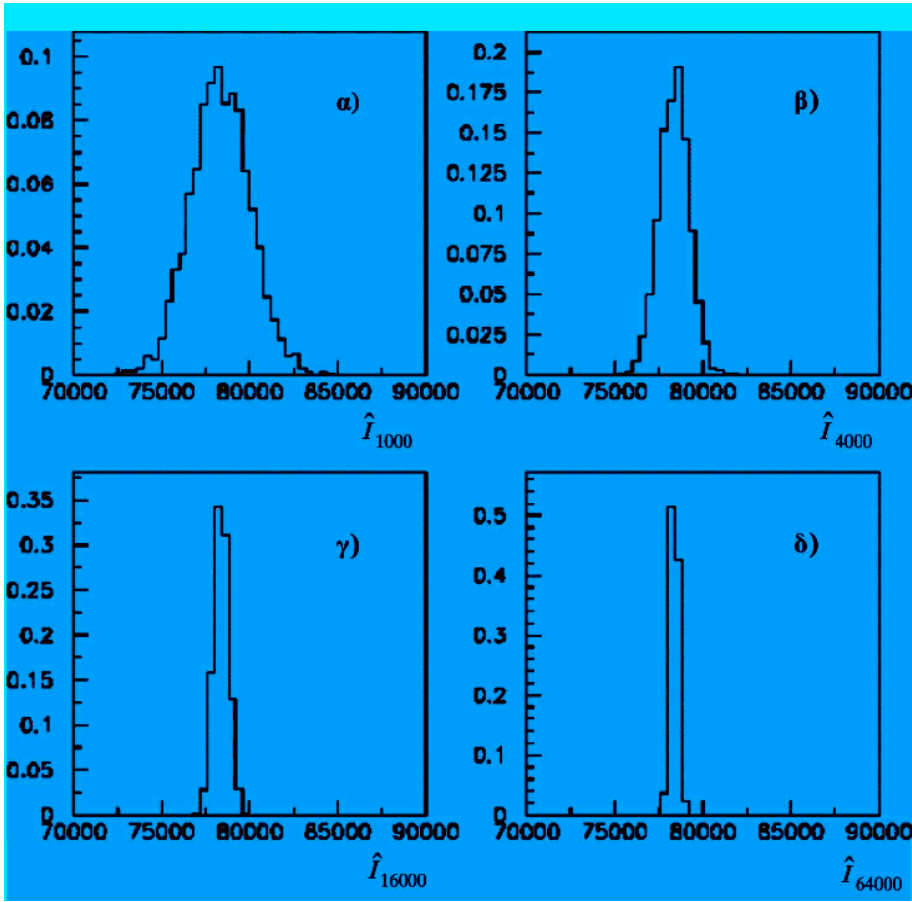
Παράδειγμα Ολοκλήρωσης Monte Carlo

$$I = \left(\frac{4}{5} \cdot 10^5 + 10^3\right) - \left(\frac{4}{5} \cdot 5^5 + 5^3\right) = 78375$$

$$\hat{I}_N = \frac{b-a}{N} \cdot \sum_{i=1}^N g(x_i)$$

$$\hat{\sigma}_N^2 = (b-a)^2 \cdot \frac{1}{N} \cdot \left[\left\langle (g(x))^2 \right\rangle - \left(\left\langle g(x) \right\rangle \right)^2 \right]$$

$$I = \int_5^{10} \underbrace{(4 \cdot x^4 + 3 \cdot x^2)}_{g(x)} dx$$



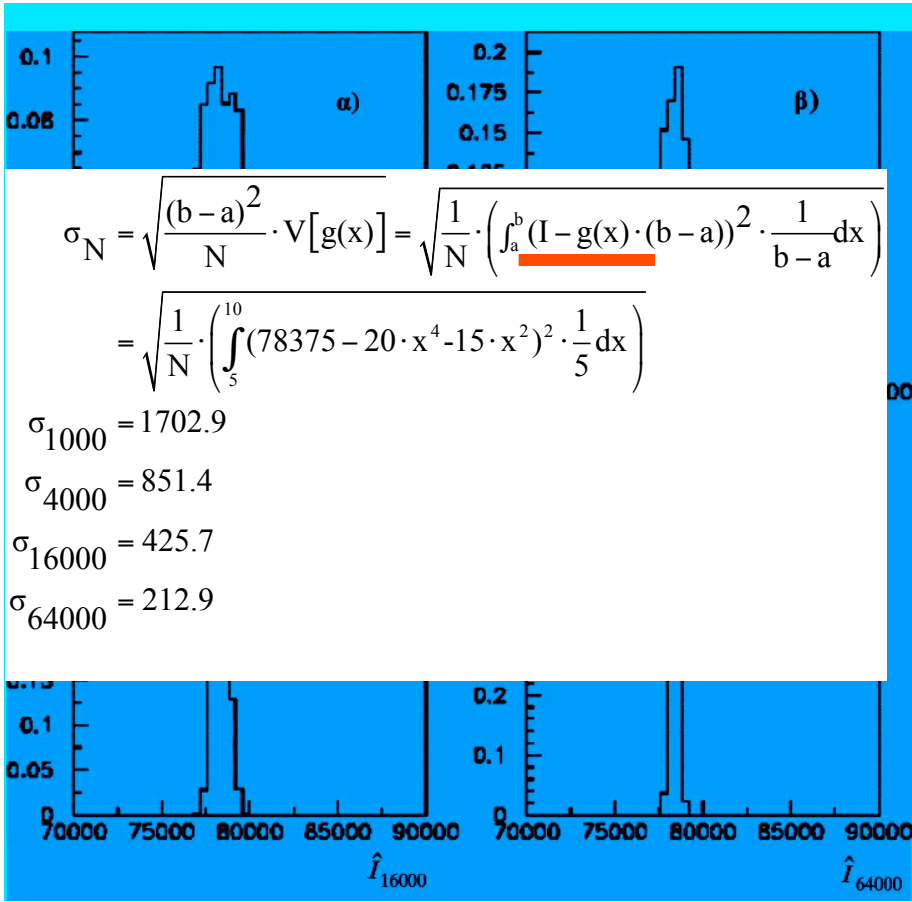
Παράδειγμα Ολοκλήρωσης Monte Carlo

$$I = \left(\frac{4}{5} \cdot 10^5 + 10^3\right) - \left(\frac{4}{5} \cdot 5^5 + 5^3\right) = 78375$$

$$\hat{I}_N = \frac{b-a}{N} \cdot \sum_{i=1}^N g(x_i)$$

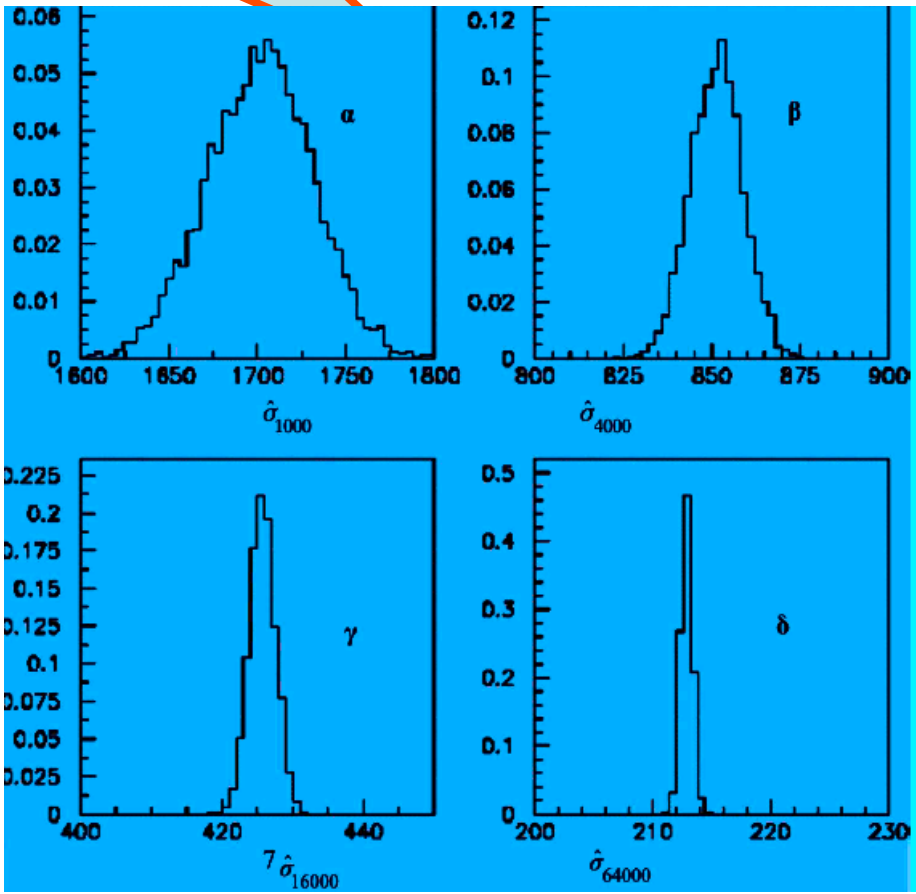
$$\hat{\sigma}_N^2 = (b-a)^2 \cdot \frac{1}{N} \cdot \left[\left\langle (g(x))^2 \right\rangle - \left(\langle g(x) \rangle \right)^2 \right]$$

$$I = \int_5^{10} \underbrace{(4 \cdot x^4 + 3 \cdot x^2)}_{g(x)} dx$$



$$\sigma_N = \sqrt{\frac{(b-a)^2}{N} \cdot V[g(x)]} = \sqrt{\frac{1}{N} \cdot \left(\int_a^b (I - g(x) \cdot (b-a))^2 \cdot \frac{1}{b-a} dx \right)}$$
$$= \sqrt{\frac{1}{N} \cdot \left(\int_5^{10} (78375 - 20 \cdot x^4 - 15 \cdot x^2)^2 \cdot \frac{1}{5} dx \right)}$$

- $\sigma_{1000} = 1702.9$
- $\sigma_{4000} = 851.4$
- $\sigma_{16000} = 425.7$
- $\sigma_{64000} = 212.9$



Θεώρημα Κεντρικού Ορίου:

Η συνάρτηση πυκνότητας πιθανότητας της τυχαίας μεταβλητής S , που ορίζεται ως το άθροισμα N ανεξάρτητων μεταβλητών:

$$S = \sum_{i=1}^N x_i$$

τείνει να γίνει (Gaussian) κανονική συνάρτηση πυκνότητας πιθανότητας καθώς το πλήθος, N , των ανεξάρτητων μεταβλητών τείνει στο άπειρο:

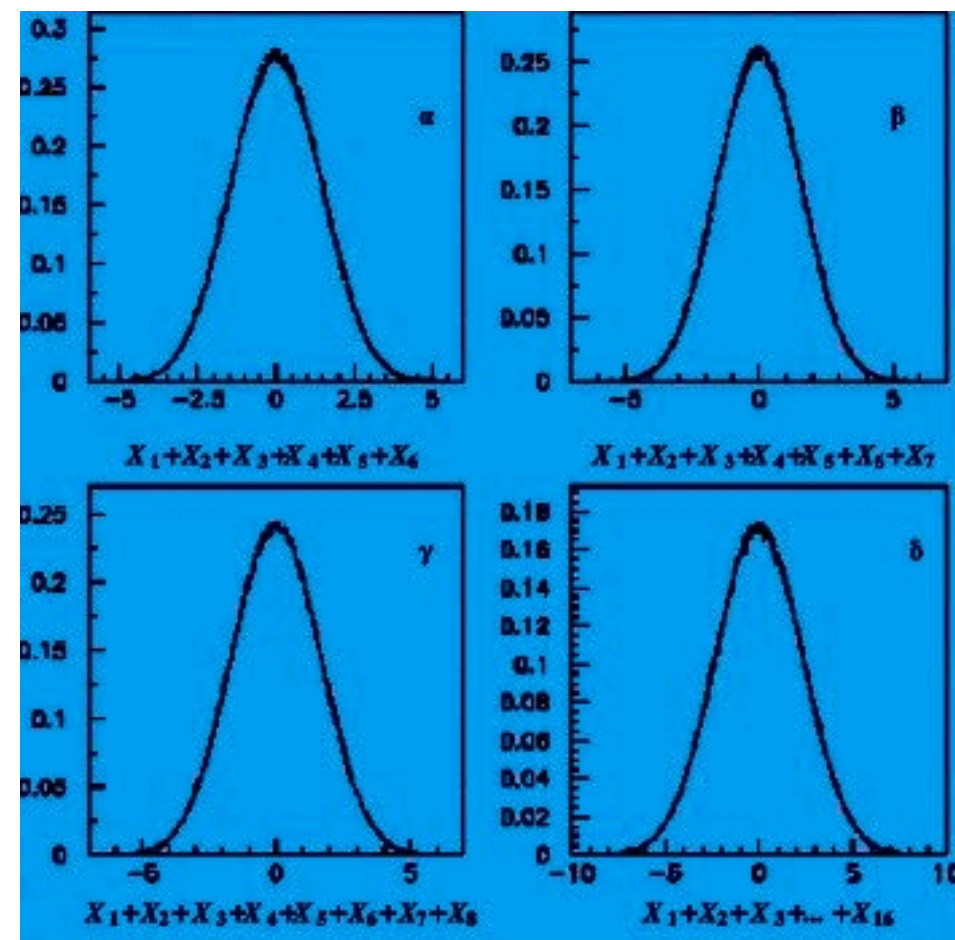
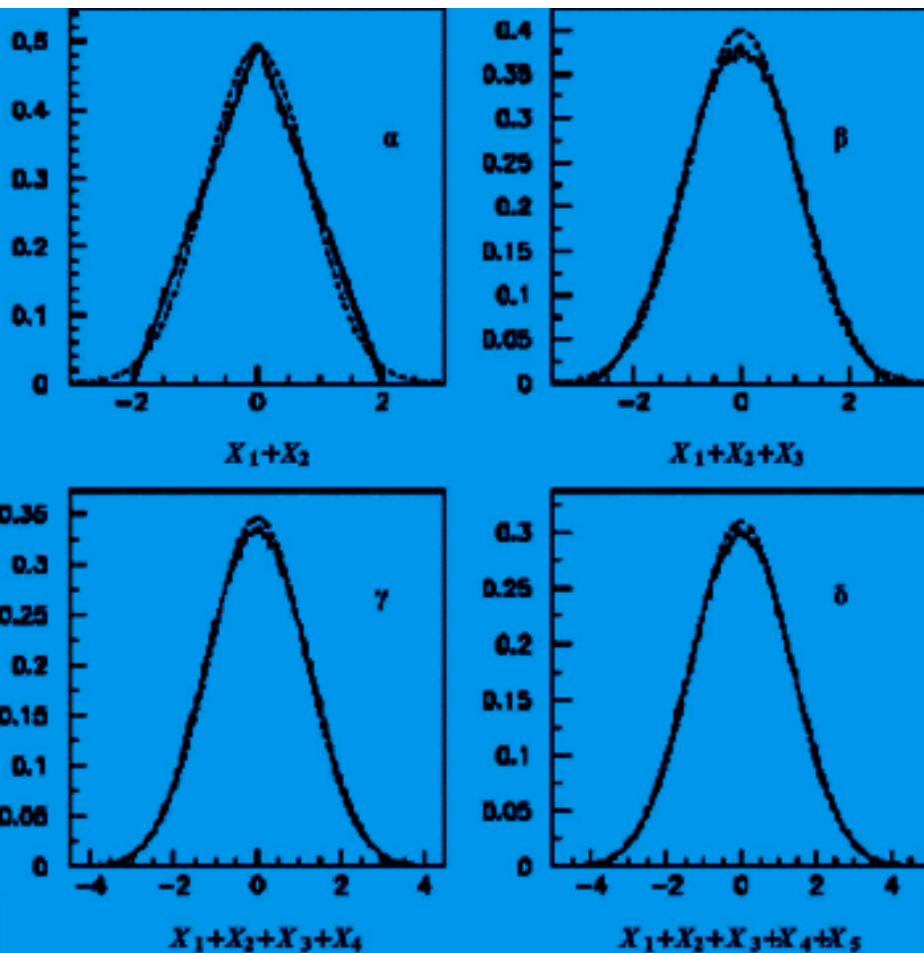
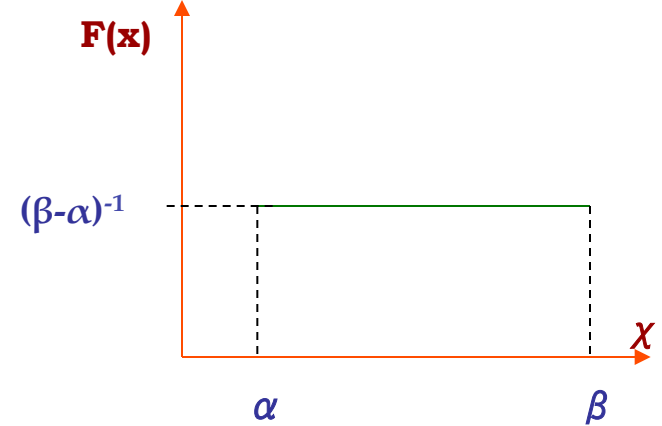
$$f(S) = \frac{1}{\sqrt{2\pi}\sigma} \cdot e^{-\frac{(S-\mu)^2}{2\cdot\sigma^2}} = N(S;\mu,\sigma), \quad \text{οπου} \quad \mu = \sum_{i=1}^N \mu_i \quad \text{και} \quad \sigma^2 = \sum_{i=1}^N \sigma_i^2$$

$$z = \frac{S - \sum_{i=1}^N x_i}{\left[\sum_{i=1}^N \sigma_i^2 \right]^{1/2}} \rightarrow f\left(\frac{S - \sum_{i=1}^N x_i}{\left[\sum_{i=1}^N \sigma_i^2 \right]^{1/2}} \right) = \frac{1}{\sqrt{2\pi} [1]} e^{-\frac{z^2}{2}} \equiv N(z;0,1)$$

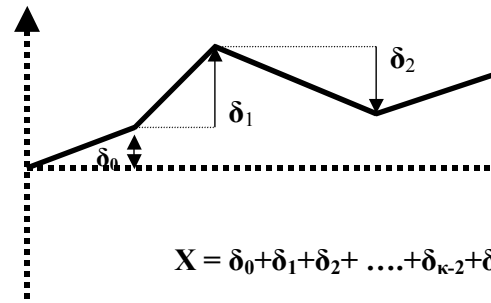
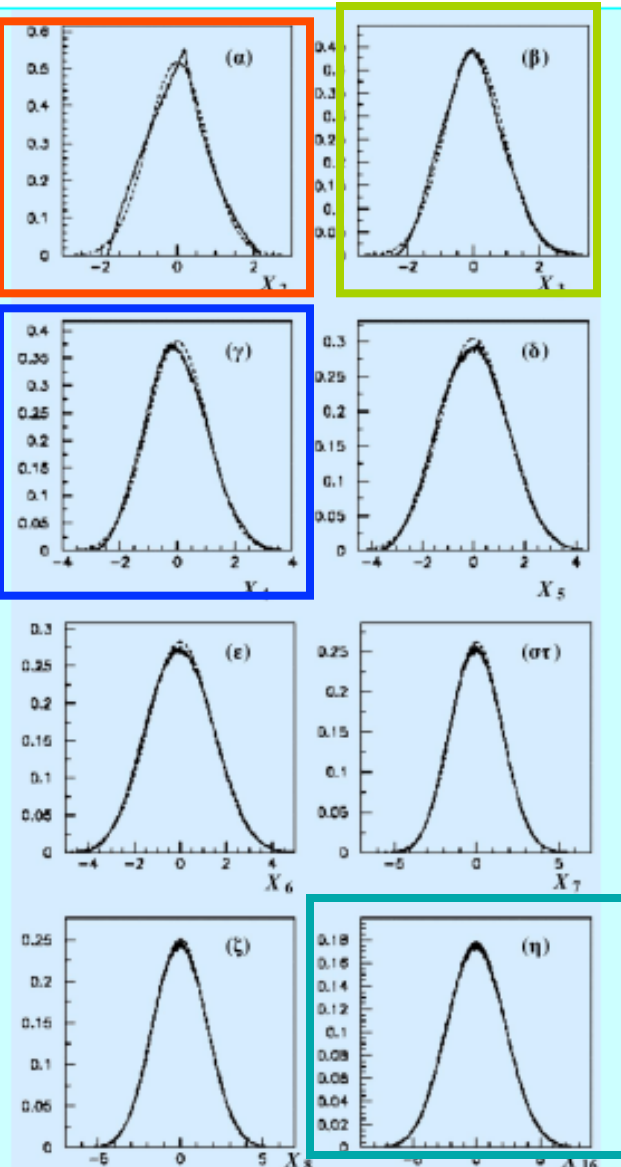
Άθροισμα Ισοπίθανων Τυχαίων Μεταβλητών

Τα ιστογράμματα παριστούν την πυκνότητα πιθανότητας του αθροίσματος δύο, τριών, τεσσάρων ... τυχαίων μεταβλητών οι οποίες λαμβάνουν τιμές ισοπίθανα στο διάστημα $[-1,1]$.

Οι συνεχείς καμπύλες παριστούν γραφικά Gaussian συναρτήσεις πυκνότητας πιθανότητας.



Πολλαπλή Σκέδαση



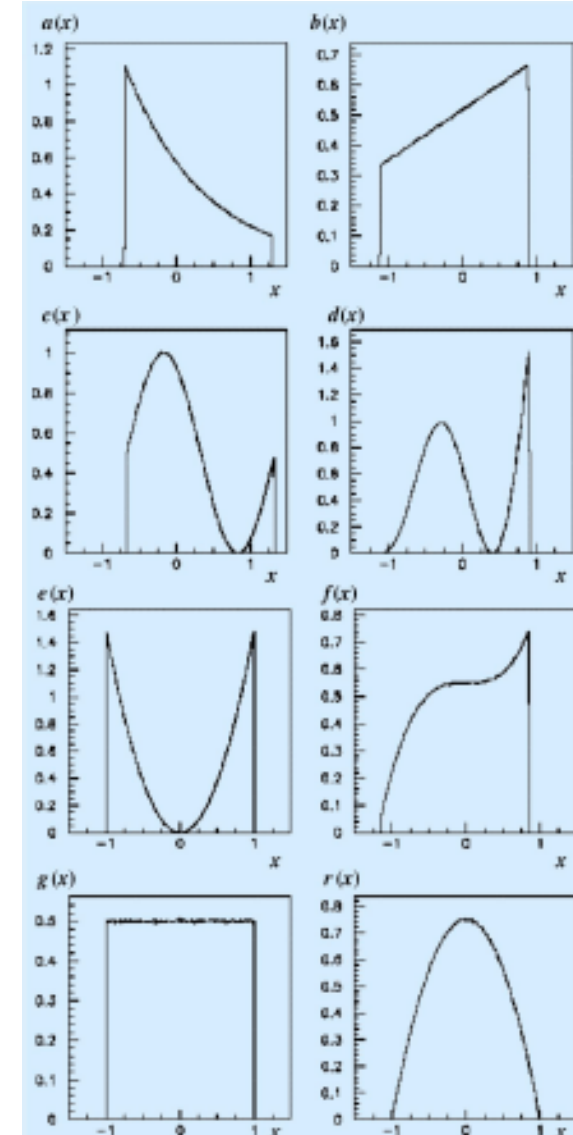
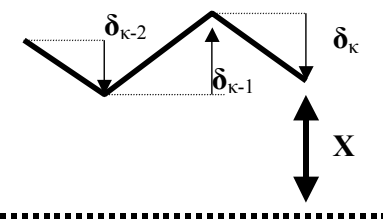
$$X = \delta_0 + \delta_1 + \delta_2 + \dots + \delta_{\kappa-2} + \delta_{\kappa-1} + \delta_{\kappa}$$

Υποθέτουμε ότι, στην περίπτωση όπου συμβαίνουν δύο μόνο σκεδάσεις, μία εκ των σκεδάσεων καταλήγει σε εγκάρσια μετατόπιση σύμφωνα με τη συνάρτηση πυκνότητας πιθανότητας $a(x)$ και η άλλη σύμφωνα με τη συνάρτηση $b(x)$.

Ομοίως, στην περίπτωση των τριών σκεδάσεων, οι τρεις εγκάρσιες μετατοπίσεις έχουν ως συναρτήσεις πυκνότητας πιθανότητας τις $a(x)$, $b(x)$ και $c(x)$ αντίστοιχα.

Στην περίπτωση των τεσσάρων σκεδάσεων, εμπλέκονται οι συναρτήσεις πυκνότητας πιθανότητας $a(x)$, $b(x)$, $c(x)$ και $d(x)$ κ.ο.κ.

Τέλος, στην περίπτωση των δεκαέξι σκεδάσεων, οι εγκάρσιες μετατοπίσεις παίρνουν, ανά δύο, τιμές σύμφωνα με τις συναρτήσεις πυκνότητας πιθανότητας $a(x)$, $b(x)$, $c(x)$, $d(x)$, $e(x)$, $f(x)$, $g(x)$ και $r(x)$.



Στατιστικό Σφάλμα της Μέτρησης

Μια τυχαία μεταβλητή παίρνει τιμές σύμφωνα με κάποια συνάρτηση πυκνότητας πιθανότητας.

Επαναλαμβανόμενες μετρήσεις μιας φυσικής ποσότητας κατανέμονται γύρω από την πραγματική τιμή (της φυσικής ποσότητας), σύμφωνα με μια συνάρτηση πυκνότητας πιθανότητας που είναι χαρακτηριστική της μετρητικής διαδικασίας.

Στη Στατιστική η συνάρτηση πυκνότητας πιθανότητας μπορεί να είναι οποιαδήποτε κατάλληλα κανονικοποιημένη συνάρτηση.

Στην Πειραματική Φυσική διάφορα φαινόμενα συμβάλλουν στη διασπορά των αποτελεσμάτων της μέτρησης. Λόγω του θεωρήματος του κεντρικού ορίου, περιμένουμε τις περισσότερες από τις συναρτήσεις πυκνότητας πιθανότητας των μετρήσεων να είναι κανονικές (Gaussians) συναρτήσεις.

Εκφράζουμε τη διακριτική ικανότητα της μετρητικής συσκευής, σ , ως την τετραγωνική ρίζα της διασποράς των αποτελεσμάτων της επαναλαμβανόμενης μέτρησης

Έστω κάποια μετρητική συσκευή με διακριτική ικανότητα σ και έστω x ένα και μοναδικό αποτέλεσμα μέτρησης της φυσικής ποσότητας X . Ποια είναι η σχέση του αποτελέσματος της μέτρησης με την πραγματική τιμή της φυσικής ποσότητας, x_R ;

Bayesian Αντιμετώπιση του Στατιστικού Σφάλματος

Πριν εκτελέσουμε οποιαδήποτε μέτρηση της φυσικής ποσότητας:

1) Γνωρίζουμε: α) τη συναρτησιακή μορφή της συνάρτησης πυκνότητας πιθανότητας των μετρήσεων (π.χ. Gaussian) στο διάστημα $[\alpha, \beta]$, β) τη σχέση που συνδέει τη μέση τιμή της συνάρτησης πυκνότητας πιθανότητας με την αληθή τιμή της φυσικής ποσότητας (π.χ. ισότητα), και γ) τη διασπορά της συνάρτησης πυκνότητας πιθανότητας.

$$P(x|x_R)dx = \left(\frac{\mathfrak{H}}{\sqrt{2\pi} \cdot \sigma} \right) \cdot e^{\frac{-(x_R-x)^2}{2 \cdot \sigma^2}} dx$$
$$\int_a^b P(x|x_R)dx = 1$$

2) Προσάπτουμε (εκ των προτέρων γνώση) ίσες πιθανότητες σ' όλες τις στοιχειώδεις περιοχές του διαστήματος $[\alpha, \beta]$ να εμπεριέχουν την αληθή τιμή της υπό μέτρηση φυσικής ποσότητας.

$$P_0(x_R) = \begin{cases} \frac{dx_R}{\beta - \alpha} & x_R \in [\alpha, \beta] \\ 0 & x_R \notin [\alpha, \beta] \end{cases}$$

3) Κάθε στοιχειώδης περιοχή του διαστήματος $[\alpha, \beta]$ έχει την ίδια πιθανότητα να εμπεριέχει το αποτέλεσμα της μέτρησης.

$$M_0(x) = \begin{cases} \frac{dx}{\beta - \alpha} & x \in [\alpha, \beta] \\ 0 & x \notin [\alpha, \beta] \end{cases}$$

Μετά την εκτέλεση μιας μέτρησης της φυσικής ποσότητας:

1) Προσάπτουμε συγκεκριμένη πιθανότητα σε κάθε στοιχειώδη περιοχή του διαστήματος $[\alpha, \beta]$, σύμφωνα με το θεώρημα Bayes να περιέχει την πραγματική τιμή του υπό μέτρηση φυσικού μεγέθους.

$$P(x_R | x) = \frac{N}{\sqrt{2 \cdot \pi} \cdot \sigma} e^{\frac{-(x_R-x)^2}{2\sigma^2}} dx_R$$

$$P(x_R | x) = \frac{P(x|x_R) \cdot P_0(x_R)}{M_0(x)}$$
$$= \frac{\left(\frac{\mathfrak{H}}{\sqrt{2 \cdot \pi} \cdot \sigma} e^{\frac{-(x_R-x)^2}{2\sigma^2}} dx \right) \cdot \frac{dx_R}{b-a}}{\frac{dx}{b-a}} \Bigg|_{x_R \in [\alpha, \beta]}$$
$$= \frac{N}{\sqrt{2 \cdot \pi} \cdot \sigma} e^{\frac{-(x_R-x)^2}{2\sigma^2}} dx_R$$

Το σφάλμα, δηλ. η αβεβαιότητα για την αληθή τιμή, συμπίπτει με την διασπορά των μετρήσεων (π.χ. διακριτική ικανότητα οργάνου). Οι πράξεις μεταξύ σφαλμάτων ακολουθούν τις ιδιότητες της απόκλισης των τυχαίων μεταβλητών

$$P(x_R | x) = 0 \qquad \Bigg|_{x_R \notin [\alpha, \beta]}$$

«Κλασική» Αντιμετώπιση του Στατιστικού Σφάλματος

Η συμβατική στατιστική προσέγγιση απορρίπτει κάθε προσπάθεια να συνδεθεί η «αληθής τιμή» με συνάρτηση πυκνότητας πιθανότητας. Η αληθής τιμή παραμένει πάντοτε άγνωστη.

Ωστόσο, μπορούμε να ορίσουμε **διάστημα εμπιστοσύνης [γ,δ]** με **περιεχόμενο πιθανότητας ίσον με λ**. Με άλλα λόγια, μπορούμε να ορίσουμε το διάστημα τιμών [γ,δ] το οποίο πιθανόν εμπεριέχει την αληθή τιμή του μεγέθους, με πιθανότητα ίση με λ.

Παραδείγματος Χάριν

$$P(x|x_R)dx = \left(\frac{\Re}{\sqrt{2\pi} \cdot \sigma} \right) \cdot e^{\frac{-(x_R-x)^2}{2 \cdot \sigma^2}} dx$$

$$\int_a^{\beta} P(x|x_R)dx = 1$$

το διάστημα εμπιστοσύνης $[x_m - \sigma, x_m + \sigma]$ έχει πιθανότητα ίση με λ να εμπεριέχει την αληθή τιμή

$$\lambda = \int_{-\sigma}^{\sigma} \left(\frac{\Re}{\sqrt{2\pi} \cdot \sigma} \right) \cdot e^{\frac{-(x_m-x_R)^2}{2 \cdot \sigma^2}} d(x_m - x_R)$$

$$[x_m - \sigma, x_m + \sigma] \quad \int_{-\infty}^{\infty} \left(\frac{1}{\sqrt{2\pi} \cdot \sigma} \right) \cdot e^{\frac{-y^2}{2 \cdot \sigma^2}} dy = 68.3\%$$

Συστηματικό Σφάλμα

Στην περίπτωση όπου η μέτρηση της φυσικής ποσότητας X , x , πάσχει, επιπλέον του στατιστικού, και από συστηματικό σφάλμα, θα θεωρήσουμε ότι η μέτρηση αποτελείται από δύο μέρη: Ένα τυχαίο, x_R , με σφάλμα σ , και ένα συστηματικό, x_S , με σφάλμα s . Θα έχουμε δηλαδή ότι $x = x_R + x_S$. Κατά συνέπεια, η αβεβαιότητα για την αληθή τιμή του μετρούμενου μεγέθους, εκφράζεται από την συνολική διασπορά των μετρήσεων, η οποία θα ισούται με:

$$\begin{aligned} V[x] &= V[x^S + x^R] \\ &= V[x^S] + V[x^R] \end{aligned}$$

Έστω ότι θέλουμε να συνδυάσουμε δύο μετρήσεις, x_1 και x_2 , οι οποίες έχουν ένα κοινό συστηματικό σφάλμα (s) και στατιστικά σφάλματα σ_1 και σ_2 αντίστοιχα.

$$x_1 = x_1^R + x_1^S \qquad x_2 = x_2^R + x_2^S$$

$$V[x_1] = \sigma_1^2 + s^2 \text{ και } V[x_2] = \sigma_2^2 + s^2$$

$$\text{cov}[x_1^R, x_2^R] = \text{cov}[x_1^R, x_2^S] = \text{cov}[x_1^S, x_2^R] = 0 \quad \text{αλλά} \quad \text{cov}[x_1^S, x_2^S] \neq 0$$

$$\begin{aligned} \text{cov}[x_1, x_2] &= E[x_1 \cdot x_2] - E[x_1] \cdot E[x_2] \\ &= E[(x_1^R + x_1^S) \cdot (x_2^R + x_2^S)] - E[x_1^R + x_1^S] \cdot E[x_2^R + x_2^S] \\ &= E[x_1^R x_2^R] + E[x_1^R x_2^S] + E[x_1^S x_2^R] + E[x_1^S x_2^S] - E[x_1^R] E[x_2^R] - \\ &\quad E[x_1^R] E[x_2^S] - E[x_1^S] E[x_2^R] - E[x_1^S] E[x_2^S] \\ &= E[x_1^S x_2^S] - E[x_1^S] E[x_2^S] = \text{cov}[x_1^S, x_2^S] = V[s] = s^2 \end{aligned}$$

$$\mathbf{V} = \begin{pmatrix} \sigma^2 + \mathbf{s}^2 & \mathbf{s}^2 \\ \mathbf{s}^2 & \sigma^2 + \mathbf{s}^2 \end{pmatrix}$$