Διακριτά Αντίστροφα Προβλήματα 2021-2022

Διάλεξη 3

Λύσεις αντιστρόφων προβλημάτων
Στοιχεία θεωρίας πιθανοτήτων

Μιχάλης Ταρουδάκης

Λύσεις αντιστρόφων προβλημάτων

Εκτιμήσεις τιμής παραμέτρων $\mathbf{m}^{\mathbf{est}} = [m_1, m_2, \dots m_M]^T$

π.χ. $\mathbf{m}^{\mathbf{est}} = [1.2, 3, 4, \dots 1.5]^T$

Οριακές τιμές (Διάστημα εμπιστοσύνης) $1.1 \le m_1 \le 1.3$

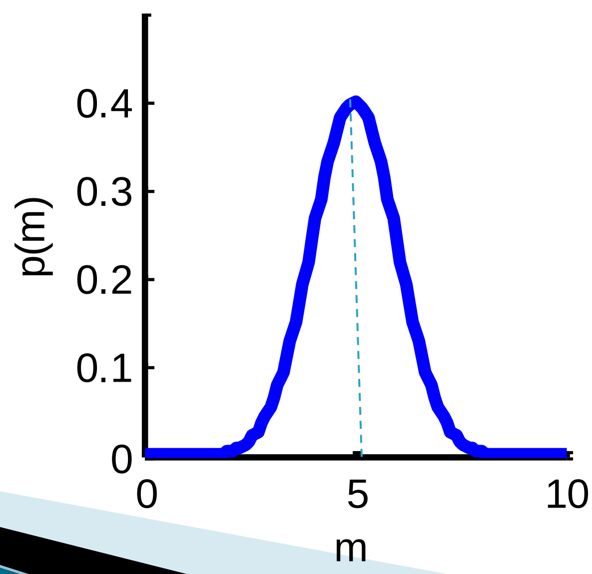
Σε συνδυασμό με πιθανοθεωρητική ανάλυση

$$m_1^{est} = \underbrace{1.2}_{\text{Μέση τιμή}} \pm \underbrace{0.1}_{\text{Απόκλιση}}$$

Λύσεις αντιστρόφων προβλημάτων

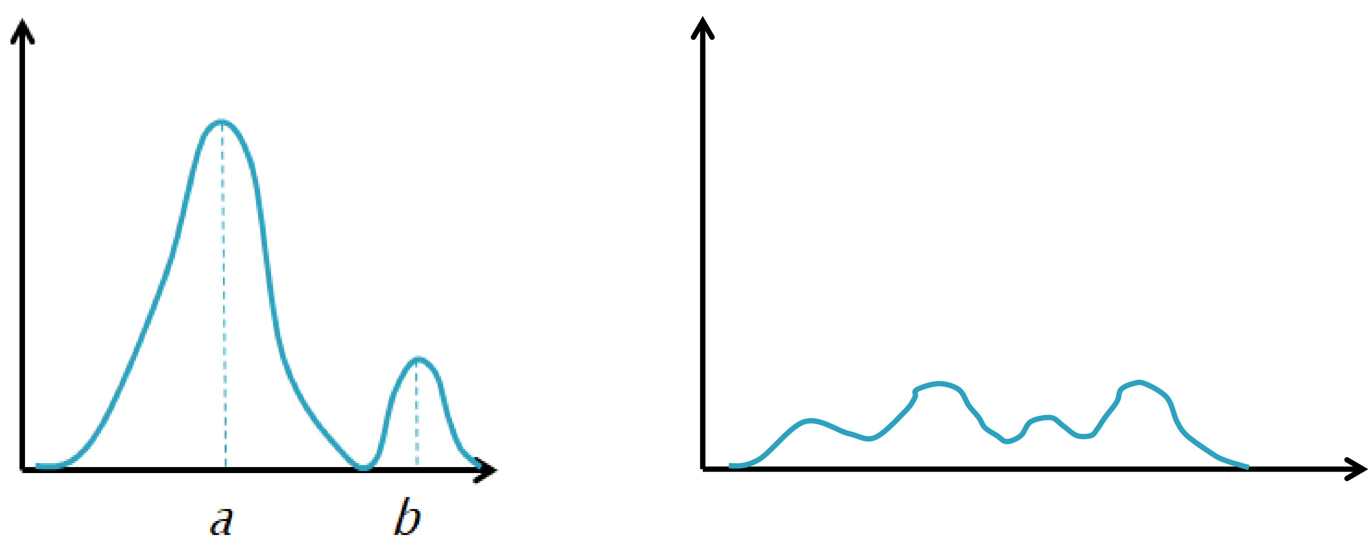
•Κατανομές Πιθανοτήτων

Εάν θεωρήσομε ότι οι παράμετροι είναι τυχαίες μεταβλητές και έχομε τη δυνατότητα να υπολογίσομε τη συνάρτηση πυκνότητας πιθανότητας για κάθε παράμετρο, μπορεί να δώσουμε τη συνάρτηση ή τις συναρτήσεις αυτές ως λύσεις του αντίστροφου προβλήματος.



Λύσεις αντιστρόφων προβλημάτων

Στην πράξη αυτό δεν είναι πολύ βολικό γιατί με εξαίρεση κατανομές που παρουσιάζουν χαρακτηριστικές μοναδικές κορυφές η πληροφορία που δίνει η κατανομή δεν είναι άμεσα αξιοποιήσιμη στις εφαρμογές. Π.χ στο σχήμα α μία παράμετρος θα μπορούσε να έχει τιμή a ή b, στο σχήμα β δεν παίρνομε ουσιαστική πληροφορία



Λύσεις αντιστρόφων προβλημάτων

Σταθμισμένες μέσες τιμές (Localized averages)

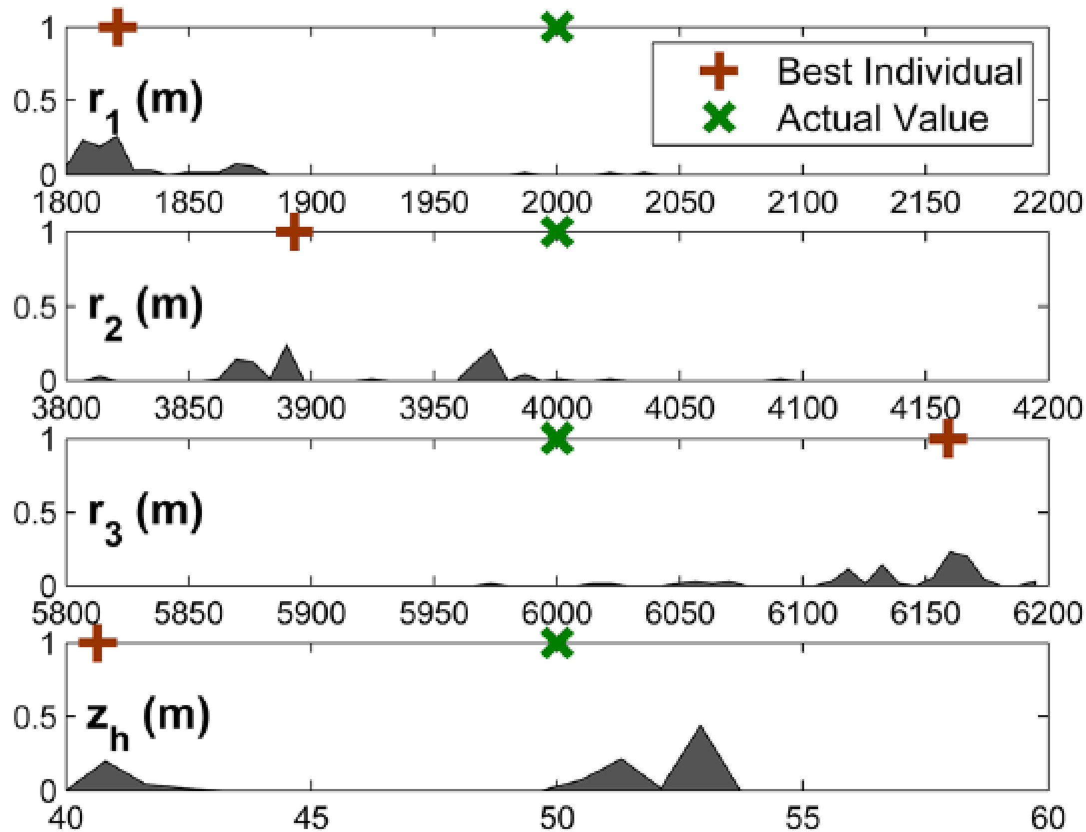
Έστω $\mathbf{m} = \left[m_1, m_2, ...m_M \right]^T$

Μπορεί γύρω από π.χ. τη δέκατη παράμετρο να μπορούμε να προσδιορίσουμε την ποσότητα

$$A_{10} = 0.1m_9 + 0.8m_{10} + 0.1m_{11}$$

Λύσεις αντιστρόφων προβλημάτων

Εκ των υστέρων (a-posteriori) στατιστική κατανομή πιθανών λύσεων



Στοιχεία Θεωρίας Πιθανοτήτων

Οι μετρήσεις που οδηγούν στα δεδομένα ενός αντίστροφου προβλήματος μπορεί να θεωρηθεί ότι είναι τυχαίες μεταβλητές (random variables) αφού πραγματοποιούνται σε συνθήκες «θορύβου»

Έστω D μία τυχαία μεταβλητή που παίρνει τιμές συνεχώς, σε ένα διάστημα που μπορεί να θεωρηθεί άπειρο.

Κατανομή πιθανότητας

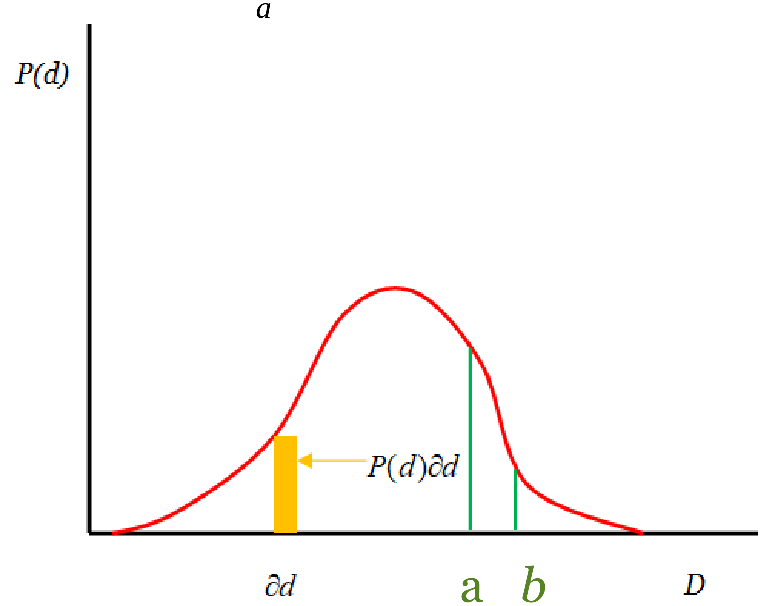
Συνάρτηση πυκνότητας πιθανότητας *P(d)*
Probability Density Function

$$P(d)\partial d = P\left(d < D < d + \partial d\right)$$

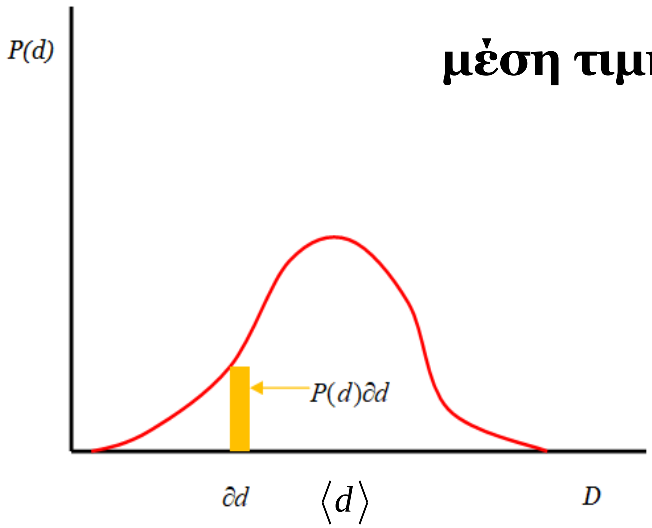
Η πιθανότητα η μεταβλητή *D* να πάρει τιμή από *a* έως *b* $\Rightarrow$ $\int_a^b P(d)\partial d$

Η πιθανότητα η μεταβλητή *D* να πάρει κάποια τιμή στο διάστημα $-\infty, \infty$ είναι προφανώς 1

$$\int_{-\infty}^{+\infty} P(d)\partial d = 1$$



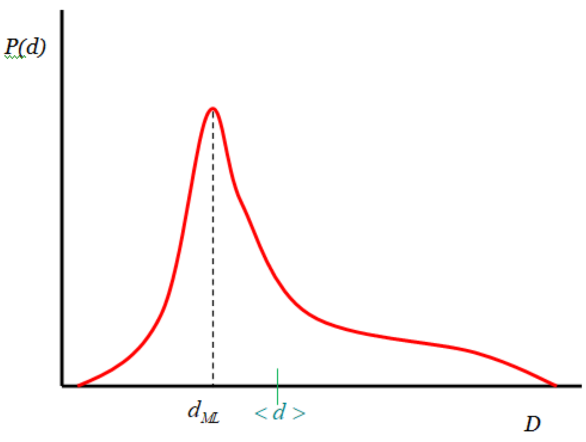
Κατανομή πιθανότητας



μέση τιμή (ή αναμενόμενη τιμή – expected value)

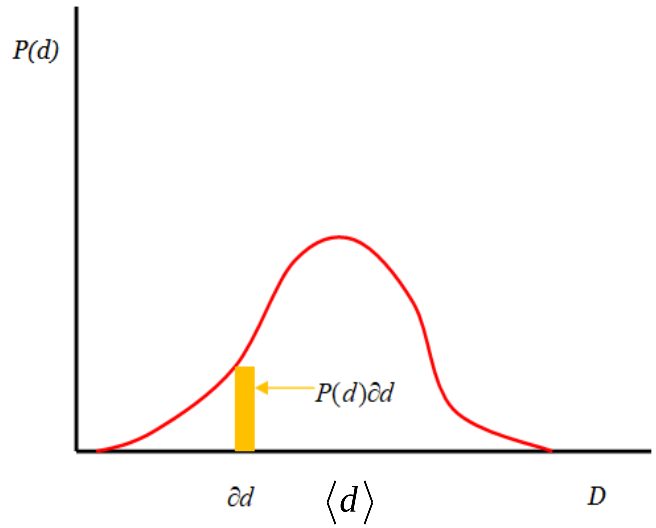
$$E(d) = \langle d \rangle = \int_{-\infty}^{+\infty} d \cdot P(d) \partial d$$

Η αναμενόμενη τιμή μπορεί να είναι διαφορετική από την τιμή μέγιστης πιθανοφάνειας



Τιμή μέγιστης πιθανοφάνειας

Κατανομή πιθανότητας



Διακύμανση (variance)

$$\sigma^2 = Var(d) = \int_{-\infty}^{\infty} \left(d - \langle d \rangle\right)^2 P(d)\partial d$$

Η διακύμανση είναι ένα μέτρο της μεταβολής της τιμής της μεταβλητής γύρω από την αναμενόμενη τιμή της

Συνάρτηση κατανομής

Αθροιστική συνάρτηση κατανομής πιθανότητας
Cumulative distribution function

$$F(d) = P(D \le d) = \int_{-\infty}^{d} P(d)\partial d, \quad -\infty < d < \infty$$

$$P(a < d \le b) = F(b) - F(a), \quad a \le b$$

$$P(d) = \frac{dF(d)}{\partial d}$$

Γενικά η *P(d)* υπάρχει όταν η *F(d)* είναι συνεχής

Διακριτές τυχαίες μεταβλητές

Η μεταβλητή D παίρνει διακριτές τιμές

$$\mathbf{d} = (d_1, d_2, ...d_i, ...d_N)$$

Κάθε διακριτή τιμή έχει πιθανότητα να εμφανιστεί P_i $\quad d_i \rightarrow P_i$

Μέση (αναμενόμενη) τιμή της μεταβλητής D $\quad \mu = \langle \mathbf{d} \rangle = \sum_{i=1}^{N} d_i P_i$

Διακύμανση μεταβλητής D $\quad Var(\mathbf{d}) = \sum_{i=1}^{N} (d_i - \langle \mathbf{d} \rangle)^2 P_i$

Συσχέτιση δεδομένων

Θεωρούμε N τυχαίες μεταβλητές D_i

από κοινού συνάρτηση πυκνότητας πιθανότητας
(*joint probability distribution function*)

$$P_D(\mathbf{d}) = P(d_1, d_2, \ldots, d_N)$$

Για ασυσχέτιστες μεταβλητές

$$P_D(\mathbf{d}) = P(d_1)P(d_2)\cdots P(d_N)$$

Συσχέτιση δεδομένων

Θεωρούμε N τυχαίες μεταβλητές D_i

$$P_D(\mathbf{d}) = P(d_1)P(d_2)\cdots P(d_N)$$

από κοινού αθροιστική συνάρτηση κατανομής
(*joint cumulative distribution function)*

$$F_D(\mathbf{d}) = P(D_1 \le d_{1,} D_2 \le d_2, \ldots., D_N \le d_N)$$

οριακή συνάρτηση πυκνότητας πιθανότητας
(*marginal probability density function*)

$$P_{Di}(\mathbf{d}) = \int_{-\infty}^{\infty} \ldots.. \int_{-\infty}^{\infty} P_D(\mathbf{d}) \partial d_1 \partial d_2 .. \partial d_{i-1} \partial d_{i+1} .. \partial d_N$$

Συσχέτιση δεδομένων

Θεωρούμε *N* τυχαίες μεταβλητές D_i

Συνδιακύμανση (covariance) δύο από τις ως άνω μεταβλητές

$$\operatorname{cov}(D_1, D_2) = \int_{-\infty}^{+\infty} \dots \int_{-\infty}^{+\infty} [d_1 - \langle d_1 \rangle][d_2 - \langle d_2 \rangle] P_D(\mathbf{d}) \partial d_1 \partial d_2 \cdots \partial d_N$$

Μέση (αναμενόμενη – expected) τιμή της μεταβλητής D_i

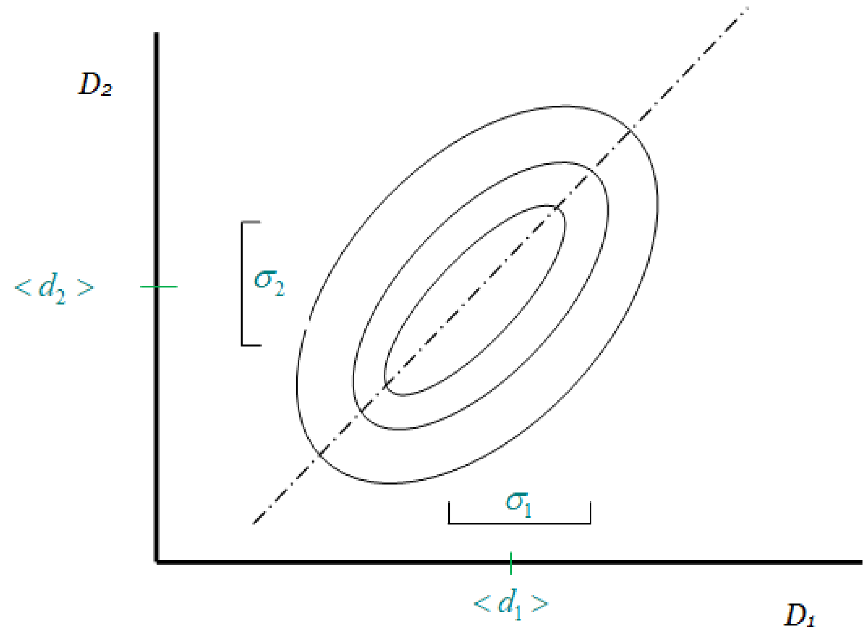
$$\langle \mathbf{d} \rangle_i = \int_{-\infty}^{\infty} \cdots \int_{-\infty}^{+\infty} d_i P_D(\mathbf{d}) \partial d_1 \partial d_2 \dots \partial d_N$$

Συσχέτιση δεδομένων

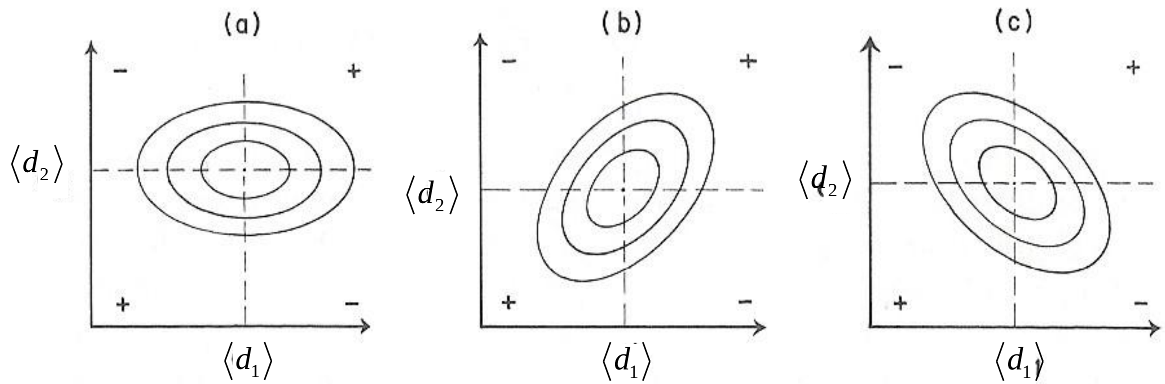
Πίνακας συνδιακύμανσης δύο μεταβλητών D_i και D_j

$$[\operatorname{cov}\mathbf{d}]_{ij} = \int_{-\infty}^{+\infty} \cdots \int_{-\infty}^{+\infty} [d_i - \langle d_i \rangle][d_j - \langle d_j \rangle] P_D(\mathbf{d}) \partial d_1 \partial d_2 \ldots \partial d_N$$

$P_D(d_1, d_2)$



Συσχέτιση δεδομένων



Οι μεταβλητές είναι ασυσχέτιστες (α), θετικά σχετιζόμενες (b) και αρνητικά σχετιζόμενες (c)

(Menke)

Συναρτήσεις μεταβλητών

$\mathbf{m}^{est}$ Παράμετροι και δεδομένα σχετίζονται

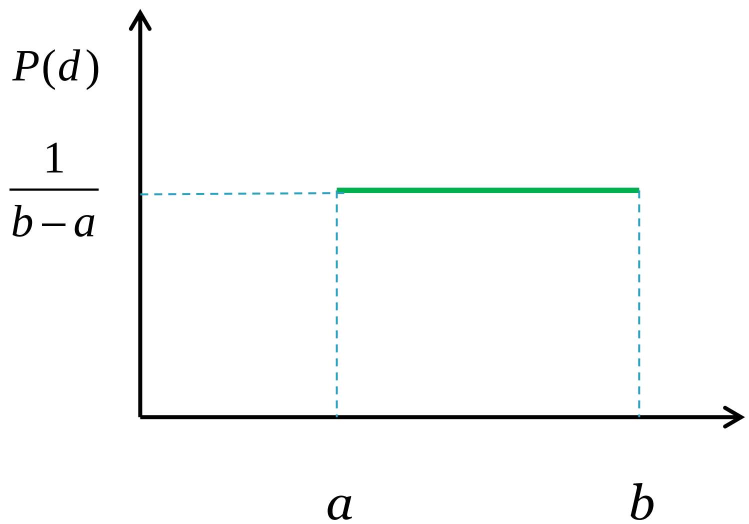
$P(\mathbf{m}^{est})$ Μπορεί επομένως η ανάλυση για τα πιθανοθεωρητικά χαρακτηριστικά των μετρήσεων (δεδομένων) να μεταφερθούν σε αυτά των παραμέτρων και να οριστούν συναρτήσεις των δεδομένων, επομένως συναρτήσεις των τυχαίων μεταβλητών

Δεν είναι εύκολη η διαδικασία αλλά μερικές φορές μπορεί να εφαρμοστεί με λογική προσπάθεια

Συναρτήσεις μεταβλητών

Ένα απλό παράδειγμα $m = d_1 + d_2$

Υποθέτουμε ότι οι μετρήσεις μας έχουν ίση πιθανότητα να πάρουν τιμές ανάμεσα στο 0 και στο 1 (Ομοιόμορφη κατανομή)



$$P(d_1) = P(d_2) = \frac{1}{b-a}$$

Συναρτήσεις μεταβλητών

Ένα απλό παράδειγμα $m = d_1 + d_2$

$$P(m) = \int_{-\infty}^{\infty} P(d_1)P(m-d_1)\partial d_1 = \int_{-\infty}^{\infty} P(m-d_2)P(d_2)\partial d_2$$

Προσαρμογή στο παράθυρο ορισμού των μεταβλητών $P(m) = \int_{0}^{1} P(d_1)P(m-d_1)\partial d_1$

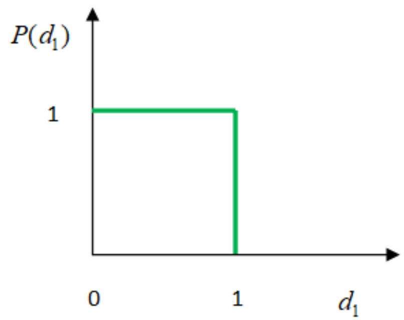
Οι συναρτήσεις πυκνότητας πιθανότητας είναι $P(d_1) = P(d_2) = 1$

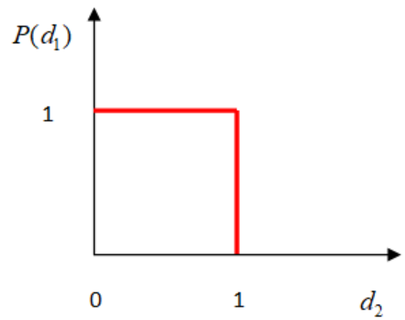
$$P(m) = \int_{0}^{m} P(d_1)P(m-d_1)\partial d_1 = \int_{0}^{m} \partial d_1 = m \qquad 0 \le m \le 1$$

$$P(m) = \int_{m-1}^{1} P(d_1)P(m-d_1)\partial d_1 = \int_{m-1}^{1} \partial d_1 = 2-m \qquad 1 \le m \le 2$$

Συναρτήσεις μεταβλητών

Ένα απλό παράδειγμα $\quad m = d_1 + d_2$





Συναρτήσεις μεταβλητών

Εάν δεδομένα και παράμετροι συσχετίζονται γραμμικά μέσω της σχέσης :

$$\mathbf{m} = \mathbf{M}\mathbf{d} + \mathbf{v}$$

Μέσες τιμές και πίνακας συνδιακύμανσης υπολογίζονται από τις σχέσεις:

$$\langle \mathbf{m} \rangle = \mathbf{M} \langle \mathbf{d} \rangle + \mathbf{v}$$

$$[\mathrm{cov}\, \mathbf{m}] = \mathbf{M}[\mathrm{cov}\, \mathbf{d}]\mathbf{M}^T$$

Συναρτήσεις μεταβλητών

Ένα δεύτερο απλό παράδειγμα :

Να υπολογιστούν τα στατιστικά χαρακτηριστικά της παραμέτρου m που είναι η μέση τιμή ομάδας δεδομένων d_i για τα οποία γνωρίζουμε ότι έχουν όλα μέση τιμή $\langle d \rangle$ και διακύμανση σ_d^2

$$\mathbf{m} = \mathbf{Md} + \mathbf{v}$$

$$\mathbf{m} = m = \frac{1}{N}\sum_{i=1}^{N} d_i = \frac{1}{N}[1,1,......1]\mathbf{d}$$

$$\mathbf{M} = \frac{1}{N}[1,1,.....,1]$$

$$\mathbf{v} = 0$$

Συναρτήσεις μεταβλητών

Ένα δεύτερο απλό παράδειγμα :

Να υπολογιστούν τα στατιστικά χαρακτηριστικά της παραμέτρου m που είναι η μέση τιμή ομάδας δεδομένων d_i για τα οποία γνωρίζουμε ότι έχουν όλα μέση τιμή $\langle d \rangle$ και διακύμανση σ_d^2

$$m = \frac{1}{N}\sum_{i=1}^{N} d_i = \frac{1}{N}[1,1,......1]\mathbf{d}$$

$$\mathbf{m = Md + v}$$

$$\langle \mathbf{d} \rangle = \left[\langle d_1 \rangle, \langle d_2 \rangle, ..., \langle d_N \rangle\right]^T = \left[\langle d \rangle, \langle d \rangle, ..., \langle d \rangle\right]^T$$

$$[\mathrm{cov}\,\mathbf{d}] = \begin{bmatrix} \sigma_d^2 & 0 & . & . & 0 \\ 0 & \sigma_d^2 & . & . & 0 \\ . & . & . & . & . \\ 0 & . & . & \sigma_d^2 & 0 \\ 0 & . & . & 0 & \sigma_d^2 \end{bmatrix}$$

Συναρτήσεις μεταβλητών

$$\mathbf{M} = \frac{1}{N}[1,1,.....,1] \qquad \mathbf{v} = 0$$

$$\langle \mathbf{m} \rangle = \mathbf{M}\langle \mathbf{d} \rangle + \mathbf{v}$$

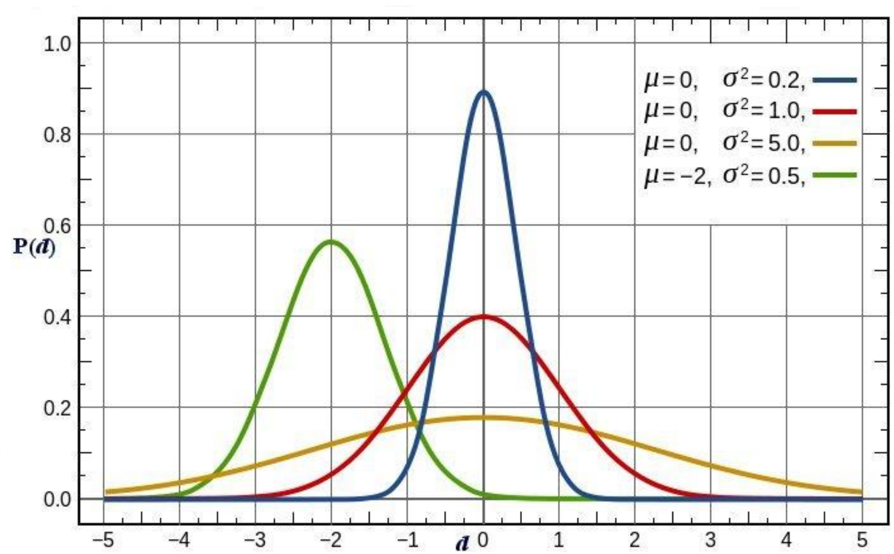
$$\langle \mathbf{m} \rangle = \langle m \rangle = \frac{1}{N}[1,1,.....,1]\left[\langle d \rangle, \langle d \rangle, ..., \langle d \rangle\right]^T = \langle d \rangle$$

$$\operatorname{var}(m) = \left[\operatorname{cov}\mathbf{m}\right] = \mathbf{M}[\operatorname{cov}\mathbf{d}]\mathbf{M}^T = \sigma_d^2 / N$$

Κανονική κατανομή

$$P(d) = \frac{1}{\sqrt{2\pi}\sigma} \exp\left[-\frac{(d-\mu)^2}{2\sigma^2}\right]$$

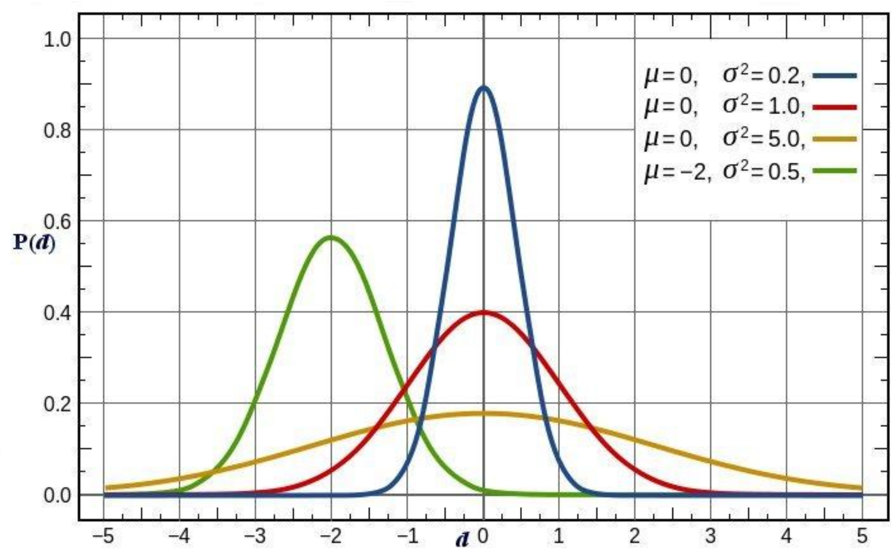
$$\sigma^2 = Var(d) \qquad \mu = \langle d \rangle$$



Wikipedia

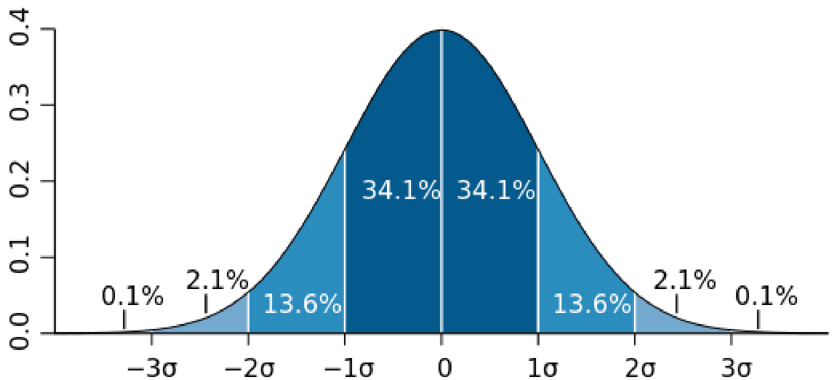
Κανονική κατανομή

Κεντρικό Οριακό Θεώρημα : Ανεξάρτητα από τις επί μέρους κατανομές μιας ομάδας ανεξάρτητων τυχαίων μεταβλητών, η συνάρτηση πυκνότητας πιθανότητας του αθροίσματος προσεγγίζει την κανονική κατανομή καθώς ο αριθμός των στοιχείων της ομάδας αυξάνει



Wikipedia

Κανονική κατανομή



Wikipedia

Η επιφάνεια κάτω από την καμπύλη με οριακές τιμές $-1\sigma < d < 1\sigma$ είναι 0.682 ενώ η επιφάνεια κάτω από την καμπύλη με οριακές τιμές $-2\sigma < d < 2\sigma$ είναι 0.954

Με άλλα λόγια η πιθανότητα μια τυχαία μεταβλητή που ακολουθεί κανονική κατανομή να πάρει τιμή $\mu \pm \sigma$ είναι 68,2 % ενώ η πιθανότητα να πάρει τιμή $\mu \pm 2\sigma$ είναι 95,4 %

Κανονική κατανομή

Η από κοινού συνάρτηση πυκνότητας πιθανότητας δύο ανεξάρτητων τυχαίων μεταβλητών που ακολουθούν κανονική κατανομή είναι το γινόμενο των συναρτήσεών πυκνότητας πιθανότητας τους.

Για σχετιζόμενες τυχαίες μεταβλητές έχομε :

$$P(\mathbf{d}) = \frac{\left\|\left[\operatorname{cov}\mathbf{d}\right]\right\|^{-1/2}}{(2\pi)^{N/2}} \exp\left(-\frac{1}{2}\left[\mathbf{d}-\langle\mathbf{d}\rangle\right]^{T} \left[\operatorname{cov}\mathbf{d}\right]^{-1}\left[\mathbf{d}-\langle\mathbf{d}\rangle\right]\right)$$

που μας δίνει τη σωστή μέση τιμή και τη σωστή διακύμανση όταν οι μεταβλητές που αποτελούν το διάνυσμα **d** είναι ασυσχέτιστες και τη σωστή διακύμανση για συσχετιζόμενες μεταβλητές. Η έκφραση αυτή ισχύει και για γραμμικές συναρτήσεις Γκαουσιανών τυχαίων μεταβλητών.

Κανονική κατανομή

Θεωρώντας ότι παράμετροι και δεδομένα σχετίζονται μέσω της άμεσης σχέσης $\mathbf{g}(\mathbf{m}) = \mathbf{d}$

Αντιστοιχούμε τις παραμέτρους με τις μέσες τιμές των δεδομένων (μετρήσεων) $\mathbf{g}(\mathbf{m}) = \langle \mathbf{d} \rangle$

$$P(\mathbf{d}) = \frac{\left\| [\operatorname{cov} \mathbf{d}] \right\|^{-1/2}}{(2\pi)^{N/2}} \exp\left(-\frac{1}{2} [\mathbf{d} - \mathbf{g}(\mathbf{m})]^T [\operatorname{cov} \mathbf{d}]^{-1} [\mathbf{d} - \mathbf{g}(\mathbf{m})] \right)$$

Διαστήματα εμπιστοσύνης

Η *εμπιστοσύνη* (confidence) μιας συγκεκριμένης παρατήρησης (μέτρησης) είναι η πιθανότητα μια πραγματοποίηση της μέτρησης να λάβει τιμή μέσα σε ένα προκαθορισμένο εύρος τιμών γύρω από τη μέση της τιμή.

Παράδειγμα : Μετράμε μία παράμετρο για την οποία γνωρίζουμε ότι οι μετρήσεις της ακολουθούν κανονική κατανομή με σ=1. Έστω ότι παίρνομε την τιμή 50. Τότε η πιθανότητα η πραγματική της τιμή να είναι ανάμεσα στο 48 και το 52 είναι 95.4 %. Γράφομε συνήθως

$$m = 50 \pm 2$$