1.4 Λύσεις αντιστρόφων προβλημάτων.

Ο τρόπος παρουσίασης της λύσης ενός αντίστροφου προβλήματος μπορεί να διαφέρει ανάλογα με τη «φιλοσοφία» επίλυσης που ακολουθείται και τη δυνατότητα παροχής πρόσθετης πληροφορίας σχετικά με τη λύση.

- Εκτιμήσεις των παραμέτρων.

Η απλούστερη μορφή που μπορεί να πάρει η λύση ενός αντίστροφου προβλήματος είναι με συγκεκριμένα αριθμητικά στοιχεία να αποτελούν το διάνυσμα $\mathbf{m}$ των προς ανάκτηση παραμέτρων. Για παράδειγμα $\mathbf{m}^{\mathbf{est}} = [1.2, 3, 4, ...1.5]^T$. Συνήθως οι εκτιμήσεις αυτής της μορφής είναι οι πλέον χρήσιμες ως λύσεις ενός αντίστροφου προβλήματος. Δεν δίνουν όμως πρόσθετη πληροφορία για το ενδεχόμενο οι εκτιμήσεις αυτές να είναι προσεγγιστικές ή να έχουν μεγάλο περιθώριο ασάφειας.

- Οριακές τιμές

Εάν είναι δυνατή η εκτίμηση ακραίων τιμών είτε χρησιμοποιώντας απόλυτη είτε πιθανοθεωρητική έννοια, τότε μπορεί η λύση του αντίστροφου προβλήματος να δίδεται στη μορφή π.χ. $1.1 \le m_1 \le 1.3$. Εάν η θεώρηση είναι πιθανοθεωρητική μπορεί η ανωτέρω έκφραση να σημαίνει ότι η πιθανότητα να βρίσκεται η παράμετρος m_1 ανάμεσα στις δύο παρατιθέμενες τιμές είναι συγκεκριμένη. Έτσι μπορεί να γράφομε : $m_1^{est} = 1.2 \pm 0.1$ και να εννοούμε ότι η πιθανότητα η παράμετρος m_1 να βρίσκεται ανάμεσα στο 1.1 και στο 1.3 είναι 95 %.

- Κατανομές Πιθανοτήτων

Εάν θεωρήσομε ότι οι παράμετροι είναι τυχαίες μεταβλητές και έχομε τη δυνατότητα να υπολογίσομε τη συνάρτηση πυκνότητας πιθανότητας (δείτε επόμενο κεφάλαιο), για κάθε παράμετρο, μπορεί να δώσουμε τις σχετικές συναρτήσεις ως λύσεις του αντίστροφου προβλήματος. 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.

- Σταθμισμένες μέσες τιμές παραμέτρων

Σε πολλές περιπτώσεις, ένα αντίστροφο πρόβλημα μπορεί να δώσει καλύτερες απαντήσεις για συνδυασμούς (π.χ. γραμμικούς) παραμέτρων, οι οποίες μπορεί να υπολογιστούν ευκολότερα αλλά και να έχουν νόημα για τις εφαρμογές. Μπορεί για παράδειγμα εάν το διάνυσμα των παραμέτρων είναι το $\mathbf{m} = [m_1, m_2]^T$, αντί να υπολογιστεί χωριστά κάθε παράμετρος, να είναι ευκολότερο να υπολογιστεί μια μέση τιμή, για παράδειγμα της παράστασης $\langle m \rangle = 0.2m_1 + 0.8m_2$. Το αν αυτή είναι χρήσιμη πληροφορία ή όχι εξαρτάται από την εφαρμογή. Συνήθως αυτή την περίπτωση την αντιμετωπίζομε όταν οι παράμετροι αφορούν διακριτοποιήσεις συνεχών παραμέτρων προκειμένου να διατυπωθεί ένα διακριτό αντίστροφο πρόβλημα.

2. ΣΤΟΙΧΕΙΑ ΑΠΟ ΤΗ ΘΕΩΡΙΑ ΠΙΘΑΝΟΤΗΤΩΝ

Στη συνέχεια να αναφερθούμε σε βασικές έννοιες από τη θεωρία πιθανοτήτων που μας είναι ιδιαίτερα χρήσιμες στην περίπτωση που τα δεδομένα μας και συνακόλουθα και οι προς ανάκτηση παράμετροι θεωρηθούν τυχαίες μεταβλητές. Υπάρχει συγκεκριμένος λόγος γι αυτή τη θεώρηση που έχει να κάνει με το γεγονός ότι σε μία πραγματική εφαρμογή, οι μετρήσεις (που δίδουν τα δεδομένα σε ένα αντίστροφο πρόβλημα) γίνονται με λάθη ή σε ένα περιβάλλον θορύβου, με αποτέλεσμα η τιμή της μέτρησης να μπορεί να είναι διαφορετική εάν το πείραμα επαναληφθεί αμέσως. Επομένως η τιμή της μέτρησης μπορεί να θεωρηθεί τυχαία μεταβλητή (*random variable*). Οι ιδιότητες της τυχαίας μεταβλητής συνήθως είναι γνωστές ποιοτικά, αλλά οι συγκεκριμένες μετρήσεις που προκύπτουν θεωρούνται ως «πραγματοποιήσεις» (realizations) της τυχαίας μεταβλητής.

2.1 Κατανομή πιθανότητας

Εάν d είναι συνεχής τυχαία μεταβλητή, η *Συνάρτηση Πυκνότητας Πιθανότητας* $P(d)$ μας δίδει την πιθανότητα, η πραγματοποίηση της τυχαίας μεταβλητής να πάρει μία τιμή στην γειτονιά του ορίσματός της, δηλαδή ανάμεσα στο d και στο $d+\partial d$ (Σχήμα 2.1) μέσω του γινομένου $P(d)\partial d$.

Η πιθανότητα, η μεταβλητή d να πάρει τιμή ανάμεσα στο a και στο b είναι $\int_a^b P(d)\partial d$ ενώ υπάρχει πλήρης βεβαιότητα ότι η μεταβλητή d θα πάρει τιμή ανάμεσα στο $-\infty$ και στο $+\infty$, οπότε έχομε :

$$\int_{-\infty}^{+\infty} P(d)\partial d = 1 \tag{2.1}$$

Η **μέση τιμή** (ή αναμενόμενη τιμή) για μία τυχαία μεταβλητή d η οποία χαρακτηρίζεται από τη συνάρτηση πυκνότητας πιθανότητας $P(d)$ δίδεται από τη σχέση :

$$E(d) = \langle d \rangle = \int_{-\infty}^{+\infty} dP(d)\partial d \tag{2.2}$$

Εάν η τυχαία μεταβλητή είναι διακριτή με N δυνατές πραγματοποιήσεις, κάθε μία από τις οποίες έχει πιθανότητα P_i, η ανωτέρω μέση τιμή δίδεται από την έκφραση :

$$\mu = \langle \vec{d} \rangle = \sum_{i=1}^{N} d_i P_i \tag{2.3}$$

Η διασπορά (variance) μια συνεχούς τυχαίας μεταβλητής, μας δίδει ποσοτικά το κατά πόσον η κατανομή πιθανότητας είναι ευρεία ή στενή, γύρω από τη μέση της τιμή. Η σχετική έκφραση είναι :

$$\sigma^2 = Var(d) = \int_{-\infty}^{+\infty} (d - \langle d \rangle)^2 P(d) \partial d \tag{2.4}$$

Το εύρος της κατανομής γύρω από τη μέση τιμή, δίδεται από το *σ*.

Για την περίπτωση της διακριτής τυχαίας μεταβλητής, η αντίστοιχη έκφραση είναι :

$$Var(\vec{d}) = \sum_{i=1}^{N} (d_i - \langle \vec{d} \rangle)^2 P_i \tag{2.5}$$

2.2 Συσχέτιση δεδομένων

Εάν σε ένα πείραμα έχομε περισσότερα από ένα δεδομένα τα οποία αντιμετωπίζονται ως τυχαίες μεταβλητές, αποκτά ενδιαφέρον η συσχέτιση των μεταβλητών μεταξύ τους. Ορίζουμε στην περίπτωση αυτή την *από κοινού συνάρτηση πυκνότητας πιθανότητας* (joint distribution) $P(\mathbf{d}) = P(d_1, d_2, ..., d_N)$, που δίνει τη πιθανότητα η πρώτη μεταβλητή να πάρει τιμή στη γειτονιά του d_1, η δεύτερη στη γειτονιά του d_2 κ.λ.π.. Εάν τα δεδομένα είναι ανεξάρτητα τότε μόνο μπορούμε να πούμε ότι η από κοινού συνάρτηση πυκνότητας πιθανότητας είναι ίση με το γινόμενο των συναρτήσεων πυκνότητας πιθανότητας των μεταβλητών

$$P(\mathbf{d}) = P(d_1)P(d_2) \cdots P(d_N) \tag{2.6}$$

Σε άλλες περιπτώσεις πάντως τα δεδομένα συσχετίζονται και αποκτά ιδιαίτερη σημασία η έκφραση ενός μέτρου συσχέτισης μεταξύ τους.

Εάν θεωρήσομε λοιπόν την περίπτωση συσχέτισης δύο τυχαίων μεταβλητών (συνεχών) d_1 και d_2, ορίζεται η **συνδιακύμανση (**covariance) από τη σχέση :

$$\mathrm{cov}(d_1, d_2) = \int_{-\infty}^{+\infty} .... \int_{-\infty}^{+\infty} [d_1 - \langle d_1 \rangle][d_2 - \langle d_2 \rangle] P(\mathbf{d}) \partial d_1 \partial d_2 \cdots \partial d_N \tag{2.7}$$

Προσέξτε ότι οι τυχαίες μεταβλητές d_1 και d_2 είναι δύο μόνο από τις μεταβλητές (δεδομένα) που αποτελούν το διάνυσμα **d.**

Για ένα διάνυσμα **d ,** N τυχαίων αλλά συνεχών μεταβλητών, ορίζομε το διάνυσμα των μέσων τιμών από τη σχέση

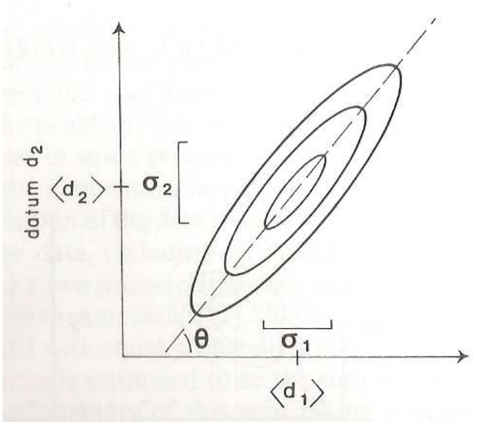
$$\langle \mathbf{d} \rangle_i = \int_{-\infty}^{+\infty} \partial d_1 \int_{-\infty}^{+\infty} \partial d_2 \cdots \int_{-\infty}^{+\infty} \partial d_N \, d_i P(\mathbf{d}) \tag{2.8}$$

και τον **πίνακα συνδιακύμανσης** από τη σχέση :

$$[\mathrm{cov}\,\mathbf{d}]_{ij} = \int_{-\infty}^{+\infty} \partial d_1 \int_{-\infty}^{+\infty} \partial d_2 \cdots \int_{-\infty}^{+\infty} \partial d_N [d_i - \langle d \rangle_i][d_j - \langle d \rangle_j] P(\mathbf{d}) \qquad (2.9)$$

Τα διαγώνια στοιχεία του πίνακα, αντιπροσωπεύουν το εύρος της κατανομής της κάθε μεταβλητής (διακύμανση), ενώ τα στοιχεία εκτός διαγωνίου αντιπροσωπεύουν το βαθμό συσχέτισης των αντίστοιχων μεταβλητών.

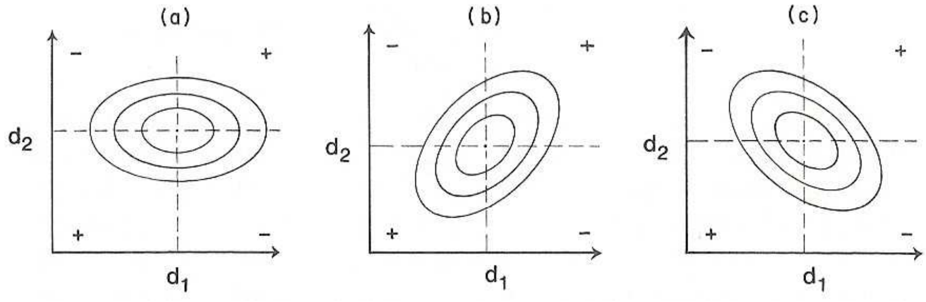
Στο επόμενο σχήμα βλέπομε εποπτικά καμπύλες ίσης κατανομής πιθανότητας για δύο τυχαίες μεταβλητές d_1 και d_2. Στο σχήμα φαίνονται οι μέσες τιμές και οι διακυμάνσεις των δύο μεταβλητών. Η γωνία θ είναι ένα μέτρο της συσχέτισης των δύο μεταβλητών και σχετίζεται με τη συνδιακύμανση.



Σχήμα 2.1 Καμπύλες ίσης κατανομής πιθανότητας $P(d_1, d_2)$ για δύο τυχαίες μεταβλητές

Στο επόμενο σχήμα 2.2 βλέπομε τρία διαγράμματα στα οποία εμφανίζονται και πάλι καμπύλες ίσης κατανομής πιθανότητας για δύο τυχαίες μεταβλητές

Το πρώτο διάγραμμα αντιπροσωπεύει ουσιαστικά μη συσχετιζόμενες μεταβλητές. Μεγάλες τιμές της μιας από αυτές μπορεί να συσχετίζονται με μικρές ή μεγάλες τιμές της άλλης με ίδια πιθανότητα. Αντίθετα το δεύτερο διάγραμμα αντιπροσωπεύει μεταβλητές που συσχετίζονται θετικά. Δηλαδή μεγάλες τιμές της μιας μεταβλητής έχουν μεγάλη πιθανότητα να συσχετίζονται με μεγάλες τιμές της δεύτερης μεταβλητής και μικρές τιμές της μια με μικρές τιμές της άλλης. Με την ίδια λογική το τρίτο διάγραμμα αντιπροσωπεύει αρνητική συσχέτιση.



Σχήμα 2.2 Καμπύλες ίσης ίσης κατανομής πιθανότητας $P(d_1, d_2)$ για δύο τυχαίες μεταβλητές που είναι α) ασυσχέτιστες, β) θετικά και γ) αρνητικά συσχετιζόμενες.

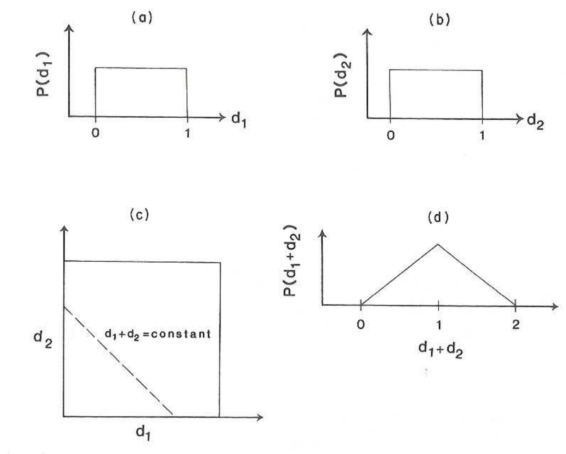
2.3 Συναρτήσεις τυχαίων μεταβλητών

Δεδομένου ότι στα αντίστροφα προβλήματα οι προς εκτίμηση παράμετροι $\mathbf{m}^{est}$ σχετίζονται με τα δεδομένα, όταν τα τελευταία είναι τυχαίες μεταβλητές, αντίστοιχα και οι παράμετροι μπορούν να θεωρηθούν τυχαίες μεταβλητές και συνεπώς μπορεί να οριστούν κατανομές πιθανοτήτων και γι αυτές $P(\mathbf{m}^{est})$. Σημειώνεται εν προκειμένω ότι οι πραγματικές παράμετροι μπορεί να είναι τυχαίες μεταβλητές ή ντετερμινιστικά μεγέθη ανάλογα με το πρόβλημα. Οι εκτιμήσεις τους πάντως είναι τυχαίες μεταβλητές τη στιγμή που έτσι θεωρούνται οι μετρήσεις.

Όταν λοιπόν οι προς ανάκτηση παράμετροι θεωρηθούν συναρτήσεις των δεδομένων, μπορεί από τα στατιστικά χαρακτηριστικά των μετρήσεων, να προκύψουν τα αντίστοιχα χαρακτηριστικά των παραμέτρων. Η διαδικασία μπορεί να μην είναι απλή είναι όμως άμεση.

Θα εξετάσομε εδώ ένα απλό πρόβλημα. Υποθέτομε ότι έχομε δύο μετρήσεις d_1 και d_2 και ότι η προς ανάκτηση παράμετρος είναι το άθροισμά τους. Επομένως ισχύει $m = d_1 + d_2$. Ας υποθέσομε επίσης ότι οι μετρήσεις μας (δεδομένα), έχουν ίση πιθανότητα να πάρουν τιμές ανάμεσα στο 1 και στο 2 (Σχήμα 2.3).

Για να υπολογίσομε την πιθανότητα το άθροισμα των δύο μεταβλητών να πάρει κάποια τιμή, παίρνομε στο επίπεδο d_1, d_2 καμπύλες ίσης τιμής του αθροίσματος και ολοκληρώνουμε πάνω στις καμπύλες. Εδώ οι καμπύλες είναι ευθείες (διάγραμμα c) και ο υπολογισμός είναι εύκολος οδηγώντας μας στο διάγραμμα d που αντιπροσωπεύει την κατανομή πιθανότητας του αθροίσματος και βλέπομε ότι είναι ένα τρίγωνο με βάση ανάμεσα στο 1 και το 2.



Σχήμα 2.3 Κατανομές πιθανότητας για δύο μεταβλητές d_1 και d_2 και για το άθροισμά τους $m = d_1 + d_2$

Ο υπολογισμός που κάναμε παραπάνω μπορεί να είναι δύσκολος σε γενικές περιπτώσεις, όμως μπορεί να δειχθεί ότι εάν δεδομένα και παράμετροι συσχετίζονται γραμμικά μέσω της σχέσης :

$$\mathbf{m} = \mathbf{M}\mathbf{d} + \mathbf{v} \tag{2.10}$$

όπου **M** και **v** είναι κάποιος πίνακας και διάνυσμα αντίστοιχα, η μέση τιμή και η συνδιακύμανση των παραμέτρων υπολογίζονται άμεσα από τις σχέσεις :

$$\langle \mathbf{m} \rangle = \mathbf{M} \langle \mathbf{d} \rangle + \mathbf{v} \tag{2.11}$$

και

$$[\mathrm{cov}\,\mathbf{m}] = \mathbf{M}[\mathrm{cov}\,\mathbf{d}]\mathbf{M}^T \tag{2.12}$$

<u>Άσκηση</u> : Να υπολογιστούν τα στατιστικά χαρακτηριστικά της παραμέτρου m που είναι η μέση τιμή μιας ομάδας δεδομένων : $m = 1/N\sum_{i=1}^{N} d_i = (1/N)[1,1,......1]\mathbf{d}$ για τα οποία γνωρίζομε ότι έχουν όλα ίδια μέση τιμή $\langle d \rangle$ και διακύμανση σ_d^2

<u>Λύση :</u> Με βάση τη σχέση 2.10, θα ισχύει : $\mathbf{M} = [1,1,.....,1]/N$, $\mathbf{v} = 0$. Από τη σχέση 2.11 έχομε : $\langle m \rangle = \mathbf{M}\langle \mathbf{d} \rangle + \mathbf{v} = \langle d \rangle$ και $\mathrm{var}(m) = \mathbf{M}[\mathrm{cov}\,\mathbf{d}]\mathbf{M}^T = \sigma_d^2 / N$. Παρατηρούμε ότι η τετραγωνική ρίζα της διακύμανσης που είναι ένα μέτρο του εύρους της διασποράς των πιθανών τιμών της παραμέτρου m γύρω από τη μέση τιμή, και επομένως της πιθανότητας ένα πείραμα να δώσει τιμή κοντά στη μέση, είναι ανάλογη του $N^{-1/2}$ που σημαίνει ότι η ακρίβεια του υπολογισμού του μέσου μιας ομάδας μετρήσεων, αυξάνει πολύ αργά σε σχέση με την αύξηση του αριθμού των μετρήσεων.

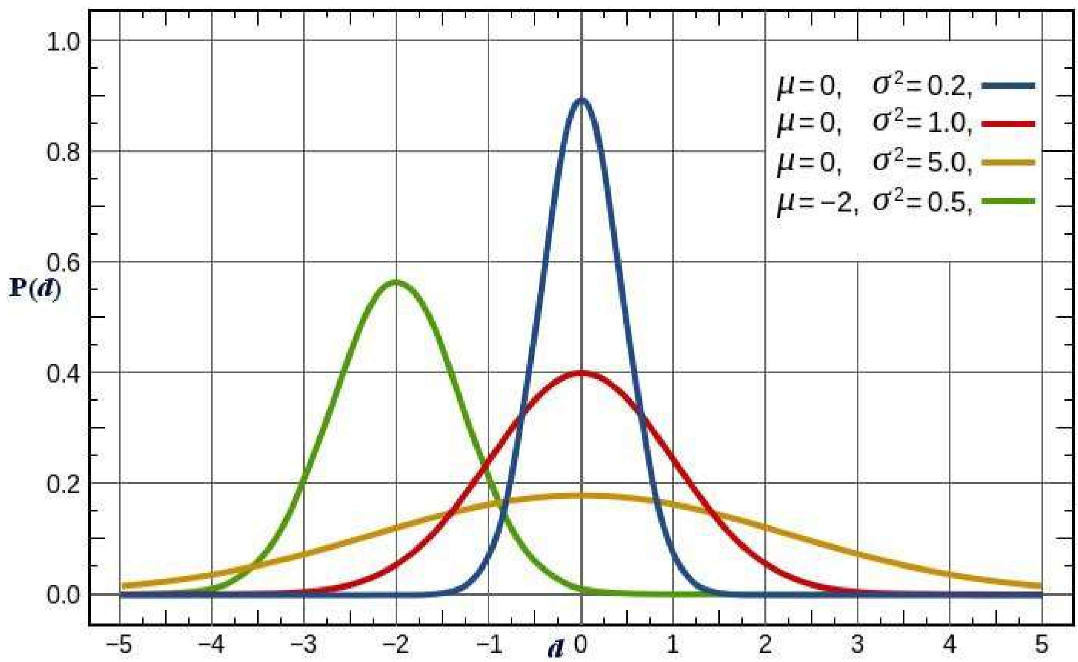
2.4 Κανονικές κατανομές

Μια τυχαία μεταβλητή d θα λέγεται ***κανονική*** (Normal ή Gaussian) με παραμέτρους μ και σ αν η συνάρτηση πυκνότητας πιθανότητας της δίδεται από τη σχέση :

$$P(d) = \frac{1}{\sqrt{2\pi}\sigma} \exp\left[-\frac{(d-\mu)^2}{2\sigma^2}\right] \tag{2.13}$$

όπου $\sigma^2 = Var(d)$ και $\mu = \langle d \rangle$. Η κανονική κατανομή εμφανίζεται στις περισσότερες εφαρμογές καθώς είναι και η οριακή κατανομή για ένα άθροισμα τυχαίων μεταβλητών, βάσει του κεντρικού οριακού θεωρήματος. Για προβλήματα ανάκτησης παραμέτρων από μετρήσεις, όταν οι μετρήσεις γίνονται σε περιβάλλον θορύβου που προέρχεται από πολλές πηγές , η κατανομή τους τείνει στην μορφή της κανονικής κατανομής.

Στο Σχήμα 2.4 φαίνεται το διάγραμμα για διαφορετικές κανονικές κατανομές. Για τρείς από αυτές η μέση τιμή είναι 0 ενώ για την Τρίτη είναι -2. Θα πρέπει να σημειωθεί ότι η επιφάνεια κάτω από την καμπύλη με οριακές τιμές $-1 < \sigma < 1$ είναι 0.68 ενώ η επιφάνεια κάτω από την καμπύλη με οριακές τιμές $-2 < \sigma < 2$ είναι 0.95. Με άλλα λόγια η πιθανότητα μια τυχαία μεταβλητή που ακολουθεί κανονική κατανομή να πάρει τιμή $\mu \pm \sigma$ είναι 68 %, ενώ η αντίστοιχη πιθανότητα να πάρει τιμή $\mu \pm 2\sigma$ είναι 95 %.



Σχήμα 2.4 Κανονική Κατανομή

Η από κοινού συνάρτηση πυκνότητας πιθανότητας δύο ανεξάρτητων τυχαίων μεταβλητών που ακολουθούν κανονική κατανομή είναι το γινόμενο των συναρτήσεών πυκνότητας πιθανότητας τους όπως είδαμε και στη γενική περίπτωση. Εάν όμως οι μεταβλητές συσχετίζονται και έχουν μέση τιμή $\langle \mathbf{d} \rangle$ και συνδιακύμανση $[\text{cov}\,\mathbf{d}]$ η κατανομή είναι περισσότερο περίπλοκη. Η έκφραση :

$$P(\mathbf{d}) = \frac{\left|[\text{cov}\,\mathbf{d}]\right|^{-1/2}}{(2\pi)^{N/2}} \exp\left(-\frac{1}{2}\left[\mathbf{d} - \langle \mathbf{d} \rangle\right]^T [\text{cov}\,\mathbf{d}]^{-1} \left[\mathbf{d} - \langle \mathbf{d} \rangle\right] \right) \qquad (2.14)$$

μας δίνει τη σωστή μέση τιμή και τη σωστή διακύμανση όταν οι μεταβλητές που αποτελούν το διάνυσμα $\mathbf{d}$ είναι ασυσχέτιστες και έχει συνδιακύμανση $[\text{cov}\,\mathbf{d}]$ για συσχετιζόμενες μεταβλητές.

2.5 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. Όπως είναι φυσικό, αναφερόμενοι σε μετρήσεις που η τιμή τους είναι τυχαία μεταβλητή, η μεγάλη τιμή της διακύμανσης δίνει μεγάλα *διαστήματα εμπιστοσύνης* (confidence intervals) και αντίστροφα. Όπως αναφέρθηκε ήδη, τυχαίες μεταβλητές κανονικής κατανομής έχουν 68 % διάστημα εμπιστοσύνης για εύρος 1σ και 95 % για εύρος 2σ.

Εάν λοιπόν μία τυχαία μεταβλητή λάβει μία συγκεκριμένη τιμή α (π.χ. στην περίπτωση μιας μέτρησης), στην περίπτωση που αυτή ακολουθεί την κανονική κατανομή, μπορούμε να ισχυριστούμε ότι η πιθανότητα η πραγματική μέση τιμή της μέτρησης να είναι $a \pm 2\sigma$ είναι 95 %.

Φυσικά τα διαστήματα εμπιστοσύνης είναι πολύ πιο δύσκολο να υπολογιστούν όταν έχομε περισσότερες συσχετιζόμενες μεταβλητές.