

Κεφάλαιο 1

Τυχαίες Μεταβλητές

1.1 Εισαγωγή

Το αποτέλεσμα ενός πειράματος θα το καλούμε στα επόμενα «γεγονός». Η επανάληψη του ιδίου πειράματος δεν καταλήγει πάντα στο ίδιο γεγονός. Η επαναλαμβανόμενη ρίψη ενός ζαριού είναι εντελώς απίθανο να καταλήγει πάντα με το ζάρι να «δείχνει» την ίδια πλευρά του προς τα επάνω. Μας ενδιαφέρει να εκφράσουμε τι είναι πιθανόν να συμβεί. Να εκφράσουμε την πιθανότητα να συμβεί κάθε ένα από τα, δυνατόν να συμβούν, γεγονότα. Ποια είναι η πιθανότητα το ζάρι να φέρει ένα; Δύο; Τρία; Τέσσερα; Πέντε; Έξι;

Σε κάθε πείραμα αποσκοπούμε να μετρήσουμε κάποια ή περισσότερες ποσότητες. Σε ένα πείραμα με ηλεκτρικά κυκλώματα μας ενδιαφέρει να μετρούμε συγχρόνως την διαφορά δυναμικού και την ένταση του ηλεκτρικού ρεύματος ενώ σε ένα άλλο πείραμα θερμοδυναμικής μας ενδιαφέρει η σύγχρονη μέτρηση της πίεσης, της θερμοκρασίας και του όγκου που καταλαμβάνει ένα αέριο. Κάθε γεγονός, το αποτέλεσμα του πειράματος, περιγράφεται πλήρως από τις μετρημένες τιμές των ποσοτήτων που μας ενδιαφέρουν.

Επαναλαμβάνοντας το ίδιο πείραμα θα καταλήξουμε σε διαφορετικές τιμές των φυσικών ποσοτήτων, είτε α) διότι δεν μπορούμε να εξασφαλίσουμε ακριβώς τις ίδιες συνθήκες εκτέλεσης του πειράματος ή β) διότι το φαινόμενο που θέλουμε να μελετήσουμε «υπόκειται σε πιθανοκρατικούς νόμους».

Σαν παράδειγμα, σε ένα πείραμα μέτρησης της περιόδου του απλού εκκρεμούς η κίνηση επηρεάζεται από την αντίσταση του αέρα, από τις τριβές και από την μάζα του νήματος ανάρτησης ενώ τα όργανα μέτρησης (χρονόμετρα) «πάσχουν» από μετρητικά σφάλματα. Σε κάθε επανάληψη του πειράματος τα μόρια του αέρα θα συγκρουστούν με διαφορετικό τρόπο με το εκκρεμές και τα φαινόμενα τριβής θα καθυστερήσουν με διαφορετικό τρόπο το χρονόμετρο ώστε ο πειραματιστής να καταλήξει σε διαφορετική τιμή για την περίοδο κίνησης.

Σ' ένα άλλο πείραμα μέτρησης του χρόνου ζωής ραδιενεργών πυρήνων, έστω και εάν χρησιμοποιήσουμε «τέλειες» μετρητικές διατάξεις, δεν θα βρεθεί κοινός χρόνος διάσπασης για όλους τους πυρήνες. Ένα τέτοιο πείραμα μπορεί μετρήσει το ποσοστό των πυρήνων ενός στοιχείου που διασπώνται σε χρόνο t διότι το φαινόμενο της ραδιενεργής

διάσπασης, αυτό καθ' εαυτό, ακολουθεί πιθανοκρατικό νόμο.

Θα καλούμε τυχαίες μεταβλητές τις μεταβλητές εκείνες που αδυνατούμε να τους προσδώσουμε συγκεκριμένη τιμή αλλά μπορούμε να προσδιορίσουμε την πιθανότητα να λάβουν μία από τις επιτρεπτές τιμές. Το αποτέλεσμα κάθε μέτρησης είναι μία τυχαία μεταβλητή.

Ας συμβολίσουμε με Z το αποτέλεσμα που θα φέρει ένα ζάρι. Οι δυνατές τιμές είναι: 1, 2, 3, 4, 5 και 6. Κάθε μία από αυτές, εάν βεβαίως το παιχνίδι είναι «τίμιο», έχει πιθανότητα να εμφανιστεί 1:6. Συνεπώς η πιθανότητα η τυχαία μεταβλητή Z να πάρει την τιμή 4 ($Z=4$) είναι ίση με $\frac{1}{6}$.

Οι τυχαίες μεταβλητές μπορούν να παίρνουν διακριτές τιμές, όπως στο παράδειγμα με το ζάρι, αλλά μπορούν να παίρνουν και συνεχείς τιμές, όπως η τυχαία μεταβλητή T που εκφράζει την μέτρηση της θερμοκρασίας. Επίσης είναι δυνατόν μία τυχαία μεταβλητή να παίρνει συνεχείς τιμές για μία περιοχή επιτρεπτών τιμών και διακριτές σε μία άλλη.

Εάν ένα γεγονός εκφράζεται από περισσότερες από μίας μετρήσεις (π.χ. πίεση, θερμοκρασία και όγκος) θα χρειαστούν περισσότερες από μία τυχαίες μεταβλητές για να ορίσουν το γεγονός (x_1 °πίεση, x_2 °θερμοκρασία, x_3 °όγκος). Συνήθως αναφερόμαστε σ' αυτές τις τυχαίες μεταβλητές ως συνιστώσες μίας διανυσματικής τυχαίας μεταβλητής ($\underline{x}$) που εκφράζει το γεγονός ($\underline{x} = \{x_1, x_2, x_3\}$).

Απομένει να συζητήσουμε τρόπους ορισμούς της πιθανότητας ώστε μία τυχαία μεταβλητή να λάβει μία συγκεκριμένη τιμή. Αλλά πριν από αυτό θα πρέπει να ορίσουμε την πιθανότητα αυτή καθ' εαυτή, ως έννοια.

1.2 Πιθανότητα

Κανένας ορισμός πιθανότητας δεν έχει γίνει απόλυτα αποδεκτός. Ένας από τους πλέον

διαδομένους ορισμούς αναφέρεται στο όριο της συχνότητας εμφάνισης ενός γεγονότος. Ας υποθεθεί ότι έχουμε εκτελέσει N φορές ένα πείραμα και ότι έχουμε συλλέξει ένα σύνολο από N γεγονότα. Τα γεγονότα τύπου X παρουσιάζονται n_X φορές, τα γεγονότα τύπου Y παρουσιάζονται n_Y φορές κ.ο.κ., όπου $n_X + n_Y + \dots = N$

Η πιθανότητα, $P(X)$, να επιλεγθεί ένα γεγονός τύπου X σε κάποια τυχαία δειγματοληψία, ορίζεται ως:

$$P(X) \equiv \lim_{N \rightarrow \infty} \frac{n_X}{N} \quad 1.2.1$$

Παρ' ότι ο παραπάνω ορισμός δίνει την δυνατότητα προσεγγιστικής εφαρμογής της έννοιας της πιθανότητας, η ισχύς του περιορίζεται σε (μη πρακτικώς εφικτά) πειράματα, όπου το N τείνει στο άπειρο. Υπάρχουν επίσης περιπτώσεις όπου ο ορισμός 1.2.1 δεν έχει καμία εφαρμογή διότι απλούστατα το πείραμα δεν μπορεί να επαναληφθεί. Παραδείγματος χάριν, δεν θα μπορούσατε να χρησιμοποιήσετε τον ορισμό 1.2.1 για να εκτιμήσετε «τι πιθανότητα υπάρχει να βρέξει αύριο;».

Τον επόμενο αξιωματικό ορισμό της πιθανότητας τον οφείλουμε στον μεγάλο μαθηματικό Kolmogorov.

Μαθηματική πιθανότητα: Ας υποθεθεί ότι Ω είναι το σύνολο όλων των δυνατών στοιχειωδών γεγονότων, X_i , που είναι δυνατόν να συμβούν καθώς επίσης και ότι μόνο ένα από αυτά μπορεί να συμβεί κάθε φορά. Τότε η πιθανότητα να συμβεί το γεγονός X_i , $P(X_i)$, ορίζεται από τις ακόλουθες ιδιότητες:

1. $P(X_i) \geq 0$ " για κάθε $X_i \in \Omega$
2. $P(X_i \text{ OR } X_j) = P(X_i) + P(X_j)$ 1.2.2
3. $\sum_{\Omega} P(X_i) = 1$

Η πρώτη από τις σχέσεις 1.2.2 επιβάλλει η πιθανότητα να συμβεί ένα γεγονός είναι μεγαλύτερη του μηδενός, ώστε κάθε δυνατό γεγονός μπορεί να συμβεί.

Η δεύτερη σχέση δηλώνει ότι η πιθανότητα να συμβεί το γεγονός X_i ή το γεγονός X_j ισούται με το άθροισμα των πιθανοτήτων τους. Η ιδιότητα αυτή είναι συνέπεια της σχέσης

αμοιβαίας απόκλισης των γεγονότων. Εάν συμβεί το γεγονός X_i αποκλείεται να συμβεί συγχρόνως και το γεγονός X_j .

Η τρίτη ιδιότητα εκφράζει ότι το σύνολο Ω καλύπτει όλα τα ενδεχόμενα αποτελέσματα του πειράματος και συνεπώς κάθε δυνατό αποτέλεσμα ανήκει στο σύνολο Ω .

Ερώτηση 1.1: Δείξτε ότι:

$$P(X_i) = 1 - P(\overline{X_i})$$

$$P(X_i) \leq 1$$

$$P(\emptyset) = 0$$

όπου $P(\overline{X_i})$ είναι η πιθανότητα να μη συμβεί το γεγονός X_i και $P(\emptyset)$ είναι η πιθανότητα να μη συμβεί τίποτα.

Μπορούμε να επεκτείνουμε τις ιδιότητες της πιθανότητας για τις εκείνες περιπτώσεις όπου υπάρχουν γεγονότα που δεν αποκλείονται αμοιβαία. Σαν παράδειγμα, ένα σύνολο A , που περιέχει στοιχειώδη, αμοιβαία αποκλειόμενα, γεγονότα X_i ($i=1,2,3,\dots,k,k+1,\dots,k+m$ και $k+m < N$), μπορεί να θεωρηθεί σαν ένα σύνθετο γεγονός. Λέμε ότι συμβαίνει το σύνθετο γεγονός A , αν τουλάχιστον ένα γεγονός εκ των X_i ($i=1,2,3,\dots,k+m$) συμβεί. Συμβολίζουμε με $P(A)$ την πιθανότητα να συμβεί τουλάχιστον ένα X_i γεγονός από το σύνολο A . :Προφανώς:

$$P(A) = P(X_1) + P(X_2) + \dots + P(X_k) + P(X_{k+1}) + \dots + P(X_{k+m}) \quad 1.2.3$$

Εάν A και B είναι δύο σύνολα στοιχειωδών γεγονότων, τα οποία ενδεχομένως περιέχουν κοινά στοιχεία, τότε η πιθανότητα να συμβεί ένα γεγονός το οποίο ανήκει στο A ή στο B σύνολο, θα είναι:

$$P(A \text{ OR } B) = P(A) + P(B) - P(A \text{ AND } B) \quad 1.2.4$$

Ερώτηση: Αποδείξτε την σχέση 1.2.4

Υπόδειξη: Η σχέση 1.2.4 είναι σχεδόν τετριμμένη. Πράγματι, εάν $A = \{X_1, X_2, X_3, \dots, X_k, X_{k+1}, \dots, X_{k+m}\}$ και

$B = \{X_{k+1}, \dots, X_{k+m}, X_{k+m+1}, X_{k+m+2}, \dots, X_{k+n}\}$ τότε τα γεγονότα $\{X_{k+1}, \dots, X_{k+m}\}$

περιέχονται και στα δύο σύνολα γεγονότων. Επιπλέον η πιθανότητα $P(A \text{ OR } B)$ αναφέρεται στο συμβάν κάποιου από τα γεγονότα του συνόλου A ή B , δηλαδή:

$$P(A \text{ OR } B) = P(X_1) + P(X_2) + \dots P(X_k) + P(X_{k+1}) + \dots P(X_{k+m}) + P(X_{k+m+1}) + \dots P(X_{k+n}) \quad 1.2.5$$

όπου οι πιθανότητες των κοινών, στα σύνολα A και B , γεγονότων συμμετέχουν μία μόνο φορά στο άθροισμα. Αντιθέτως η πιθανότητα $P(A \text{ AND } B)$ αναφέρεται στο σύγχρονο συμβάν ενός από τα γεγονότα του συνόλου A και ενός από τα γεγονότα του συνόλου B , δηλαδή θα ισούται με το άθροισμα:

$$P(A \text{ AND } B) = \sum_{i=1}^{k+m} \sum_{j=k+1}^{k+n} P(X_i \text{ AND } X_j) \quad 1.2.6$$

Λόγω όμως της αμοιβαίας αποκλίσεως των γεγονότων ισχύει ότι:

$$P(X_i \text{ AND } X_j) = \begin{cases} 0 & \text{για } i \neq j \\ P(X_i) & \text{για } i = j \end{cases} \quad 1.2.7$$

Αρκεί να αντικαταστήσετε τις σχέσεις 1.2.3, 1.2.5, 1.2.6 και 1.2.7 στην έκφραση 1.2.4

Ερώτηση 1.3: Στην περίπτωση των N συνόλων στοιχειωδών γεγονότων $A_1, A_2, \dots, A_N$, τα οποία ενδεχομένως να έχουν και κοινά στοιχεία, η προηγούμενη σχέση γενικεύεται ως:

$$P(A_1 \text{ OR } A_2 \text{ OR } \dots \text{ OR } A_N) = S_1 - S_2 + S_3 - \dots - (-1)^N S_N$$

Όπου:

$$S_1 = \sum_i P(A_i)$$

$$S_2 = \sum_{i < j} P(A_i \text{ AND } A_j) \quad \text{και}$$

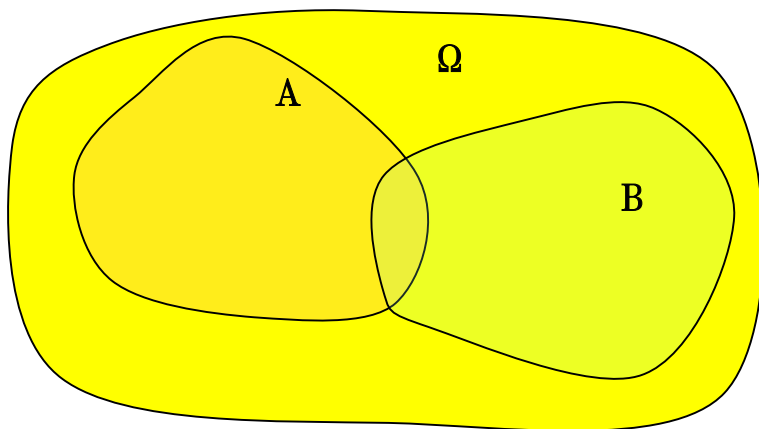
$$S_3 = \sum_{i < j < k} P(A_i \text{ AND } A_j \text{ AND } A_k) \quad \text{κ.οκ.}$$

Υπόδειξη: Εφαρμόσετε την μέθοδο της επαγωγής και χρησιμοποιήσετε το τεχνασμα ότι:

$$P(A_1 \text{ OR } A_2 \text{ OR } \dots \text{ OR } A_{n-1} \text{ OR } A_n) = P(\underbrace{(A_1 \text{ OR } A_2 \text{ OR } \dots \text{ OR } A_{n-1})}_B \text{ OR } A_n) = P(B \text{ OR } A_n)$$

1.3 Υπό συνθήκη πιθανότητα και ανεξαρτησία

Έστω ότι κάποιο γεγονός ανήκει σίγουρα στο σύνολο γεγονότων A και ίσως ανήκει και στο σύνολο γεγονότων B . Σε τέτοιες περιπτώσεις, συνήθως μας ενδιαφέρει η πιθανότητα ένα γεγονός να ανήκει στο B δοθέντος ότι ανήκει στο A . Αυτή την πιθανότητα την χαρακτηρίζουμε ως “υπό συνθήκη” ή δεσμευμένη πιθανότητα και την συμβολίζουμε ως: $P(B|A)$.



Σχήμα 1.3.1: Το σύνολο στοιχειωδών γεγονότων Ω , A και B

Στο Σχήμα 1.3.1 παρίσταται το σύνολο Ω όλων των δυνατών γεγονότων καθώς και τα σύνολα γεγονότων A και B . Στην αναπαράσταση αυτή, θεωρήστε ότι παρίσταται ο ίδιος αριθμός γεγονότων ανά μονάδα επιφανείας. Με τον τρόπο αυτό, η πιθανότητα ένα από τα

δυνατά γεγονότα να ανήκει στο σύνολο B, P(B), θα ισούται με τον λόγο του εμβαδού που καλύπτει το σύνολο B προς την επιφάνεια που αντιστοιχεί στο σύνολο Ω:

$$P(B) = \frac{\text{Εμβαδόν}[B]}{\text{Εμβαδόν}[\Omega]} \quad 1.3.1$$

Κατ' αντιστοιχία θα γράψουμε ότι:

$$P(A) = \frac{\text{Εμβαδόν}[A]}{\text{Εμβαδόν}[\Omega]} \quad \text{και} \quad P(A \text{ AND } B) = \frac{\text{Εμβαδόν}[A \text{ AND } B]}{\text{Εμβαδόν}[\Omega]} \quad 1.3.2$$

Η δεσμευμένες πιθανότητες P(B|A) και P(A|B) θα γράφονται συναρτήσει των επιφανειών ως ακολούθως:

$$P(A | B) = \frac{\text{Εμβαδόν}[A \text{ AND } B]}{\text{Εμβαδόν}[B]} \quad \text{και} \quad P(B | A) = \frac{\text{Εμβαδόν}[A \text{ AND } B]}{\text{Εμβαδόν}[A]} \quad 1.3.3$$

Από τις εκφράσεις 1.3.1 – 1.3.3 εύκολα συνάγεται ότι: η δεσμευμένη πιθανότητα συνδέεται με την πιθανότητα το γεγονός να ανήκει και στα δύο σύνολα συγχρόνως με την ακόλουθη απλή σχέση:

$$P(A \text{ AND } B) = P(A|B) P(B) = P(B|A) P(A) \quad 1.3.4$$

Ας υποθέσουμε ότι επιλέγουμε, ανεξάρτητα, δύο γεγονότα από το σύνολο Ω. Στην περίπτωση αυτή η σχέση 1.3.4 εκφράζει την πιθανότητα το ένα γεγονός να ανήκει στο σύνολο A και το άλλο στο σύνολο B. Ας υποθέσουμε επιπλέον ότι τα σύνολα A και B είναι αμοιβαία ανεξάρτητα, δεν υπάρχει δηλαδή η αλληλοεπικάλυψη που δείχνει το Σχήμα 1.3.1. Στην περίπτωση αυτή, η πιθανότητα επιλογής του δεύτερου γεγονότος από το σύνολο A δεν εξαρτάται από την επιλογή του πρώτου γεγονότος, δηλαδή από το εάν προηγουμένως είχε επιλεγεί γεγονός από το σύνολο B. Συνεπώς:

$$P(A|B) = P(A) \quad 1.3.5$$

και κατ' επέκταση η σχέση 1.3.4 καταλήγει ότι:

$$P(A \text{ AND } B) = P(A) P(B) \quad 1.3.6$$

Παρά την απλότητα του ορισμού της, η δεσμευμένη πιθανότητα αποτελεί ένα ισχυρό

στατιστικό εργαλείο με πληθώρα πρακτικών εφαρμογών, όπως φαίνεται με τα ακόλουθα παραδείγματα.

Παράδειγμα 1.3.1

Η γραμμή παραγωγής ενός εργοστασίου περιέχει τους μηχανικούς καταμετρητές, A και B, προϊόντων. Τα συσκευασμένα προϊόντα, μεταφερόμενα πάνω σε ιμάντα, περνούν διαδοχικά μπροστά από τους καταμετρητές.

Μετά από 12 ώρες καταμέτρησης, ο καταμετρητής A πιστοποίησε την διέλευση $N_1 = C + D_1$ δεμάτων ενώ $N_2 = C + D_2$ δέματα ανιχνεύτηκαν από τον μετρητή B. Παράλληλα, C δέματα καταγράφηκαν συγχρόνως και από τους δύο μετρητές. (Προφανώς D_1 δέματα καταγράφηκαν μόνο από τον μετρητή A και D_2 δέματα καταγράφηκαν μόνο από τον μετρητή B.).

Θέλουμε να εκτιμήσουμε τον συνολικό αριθμό των δεμάτων τα οποία πέρασαν δια μέσου της συστοιχίας των δύο καταμετρητών αυτές τις 12 ώρες και την συνολική αποδοτικότητα (efficiency) της ανιχνευτικής συστοιχίας.

Ορισμός της Αποδοτικότητας (Efficiency): Αποδοτικότητα μιας ανιχνευτικής διάταξης αντικειμένων τύπου X , είναι η πιθανότητα η εν λόγω διάταξη να διεγερθεί και να πιστοποιήσει την παρουσία αντικειμένου τύπου X , όταν αυτό βρίσκεται στην ενεργό περιοχή ανίχνευσης. Είναι σαφές ότι καμία ανιχνευτική διάταξη δεν έχει αποδοτικότητα 100%.

Ο απλούστερος τρόπος που θα μπορούσε να σκεφτεί κανείς, για να μετρήσει την αποδοτικότητα μιας ανιχνευτικής διάταξης συνίσταται στην απόπειρα ανίχνευσης N αντικειμένων που αποδεδειγμένα είναι τύπου X (δείγμα ελέγχου). Εάν η ανιχνευτική διάταξη ανταποκριθεί (ανιχνεύσει) στην παρουσία n ($n < N$) από αυτά τότε η αποδοτικότητα θα προσεγγίζεται με τον λόγο n/N . Προφανώς, αυτή η διαδικασία βαθμονόμησης είναι αξιόπιστη μόνο όταν ο πληθυσμός του δείγματος ελέγχου (N) είναι μεγάλος, ώστε να έχει έννοια ο ορισμός της πιθανότητας.

Στη διαδικασία βαθμονόμησης είναι απολύτως απαραίτητο το δείγμα ελέγχου, το οποίο θα πρέπει να συνίσταται μόνο από αντικείμενα τύπου X . Για την δημιουργία τέτοιων δειγμάτων είναι απαραίτητος και δεύτερος ανιχνευτής.

Στα επόμενα θα υποθέσουμε ότι κάθε ένα από τα δέματα έχει πιθανότητα $P(A)$, να αναγνωρισθεί (και καταμετρηθεί) από τον καταμετρητή A και πιθανότητα $P(B)$, να αναγνωρισθεί (και καταμετρηθεί) από τον καταμετρητή B .

Εξ ορισμού, οι πιθανότητες $P(A)$ και $P(B)$ εκφράζουν:

$P(A)$ ° Αποδοτικότητα του A στην καταμέτρηση δεμάτων

$P(B)$ ° Αποδοτικότητα του B στην καταμέτρηση δεμάτων 1.3.7

Επειδή οι καταμετρητές είναι αμοιβαία ανεξάρτητοι, η πιθανότητα ένα δέμα να μετρηθεί από τον A δεν επηρεάζεται από το εάν ανιχνεύθηκε ή δεν ανιχνεύθηκε από τον B . Ωστε, λόγω της 1.3.5, μπορούμε να γράψουμε:

$$P(A | B) = P(A)$$

$$P(B | A) = P(B) \quad 1.3.8$$

Επιπλέον, λόγω της ανεξαρτησίας των μετρήσεων (σχέση 1.3.6), η πιθανότητα και οι δύο καταμετρητές να μετρήσουν το ίδιο δέμα, $P(A \text{ AND } B)$, θα δίνεται από το γινόμενο, της

πιθανότητας το δέμα να μετρηθεί από τον A (ανεξάρτητα με την απόκριση του B), $P(A)$, επί την πιθανότητα το δέμα να μετρηθεί από τον B (ανεξάρτητα με την απόκριση του A), $P(B)$.

$$P(A \text{ AND } B) = P(A) \cdot P(B) \quad 1.3.9$$

Γνωρίζουμε όμως ότι, από τα $N_2 = C + D_2$ δέματα που μέτρησε ο B μόνο τα C ανιχνεύτηκαν συγχρόνως και από τον A. Χρησιμοποιώντας τα δέματα που ανιχνεύτηκαν από τον B σαν δείγμα ελέγχου για τον A, καταλήγουμε στην επόμενη έκφραση για την (δεσμευμένη πιθανότητα) αποδοτικότητα του A:

$$P(A|B) = \frac{C}{C + D_2} \quad 1.3.10$$

Κατ' αντιστοιχία, η (δεσμευμένη) αποδοτικότητα του B είναι:

$$P(B|A) = \frac{C}{C + D_1} \quad 1.3.11$$

Δεδομένης όμως της ανεξαρτησίας των δύο ανιχνευτών, βλέπε τις σχέσεις 1.3.8, οι εκφράσεις 1.3.10 και 1.3.11 καταλήγουν ως:

$$\hat{P}(A) = \frac{C}{C + D_2} \quad \hat{P}(B) = \frac{C}{C + D_1} \quad 1.3.12$$

όπου χρησιμοποιήσαμε το σύμβολο « $\hat{}$ » για να τονίσουμε ότι δεν υπολογίσαμε την ακριβή αποδοτικότητα των ανιχνευτών (το δείγμα ελέγχου ήταν πεπερασμένο) αλλά απλώς κάναμε μια εκτίμηση της αποδοτικότητας. Στα επόμενα Κεφάλαια θα συζητήσουμε τρόπους ποσοτικής έκφρασης τέτοιων εκτιμήσεων.

Η συνολική αποδοτικότητα της συστοιχίας, αναφέρεται στη συνδυασμένη ανιχνευτική ικανότητα των δύο καταμετρητών και εκφράζεται ως το λογικό OR των αποκρίσεών τους. Δηλαδή απαιτούμε ή ο A, ή ο B ή και οι δύο να έχουν ανιχνεύσει (και καταμετρήσει) το διερχόμενο δέμα. Η απαίτησή μας εκφράζεται (βλέπε σχέση 1.2.4) ως:

$$P(A \text{ OR } B) = P(A) + P(B) - P(A \text{ AND } B)$$

Αντικαθιστώντας στην προηγούμενη έκφραση τις σχέσεις 1.3.12 και 1.3.9 καταλήγουμε ότι:

$$\begin{aligned} P(\mathbf{A \text{ OR } B}) &= P(A) + P(B) - P(A \text{ AND } B) \cong \hat{P}(A) + \hat{P}(B) - \hat{P}(A) \cdot \hat{P}(B) = \\ &= \frac{C}{C + D_2} + \frac{C}{C + D_1} - \frac{C^2}{(C + D_2)(C + D_1)} \end{aligned} \quad 1.3.13$$

Τελικά, για να εκτιμήσουμε τον συνολικό αριθμό των δεμάτων που πέρασαν μπροστά από την συστοιχία των καταμετρητών, χρησιμοποιούμε: α) την εκτίμησή μας για την αποδοτικότητα ενός από τους ανιχνευτές (π.χ. του A), β) τον συνολικό αριθμό δεμάτων που καταμετρήθηκαν από τον ίδιο μετρητή ($C + D_1$), και γ) εκφράζουμε την αποδοτικότητα του ίδιου ανιχνευτή ως προς το σύνολο των γεγονότων (δηλαδή τον συνολικό πλήθος των δεμάτων, N, που πέρασαν μπροστά από την μετρητική συστοιχία.

$$\hat{P}(A) = \frac{C}{C + D_2} = \frac{C + D_1}{\hat{N}}$$

Επομένως ο συνολικός αριθμός δεμάτων (η εκτίμηση του συνολικού αριθμού δεμάτων)

είναι:
$$\hat{N} = \frac{(C + D_1)(C + D_2)}{C}$$

1.4 Θεώρημα του Bayes

Από τον ορισμό της δεσμευμένης πιθανότητας προκύπτει ότι για δύο κατηγορίες γεγονότων, A και B, η πιθανότητα ώστε ένα γεγονός να ανήκει συγχρόνως στις κατηγορίες A και B, εκφράζεται ως:

$$P(A \text{ AND } B) = P(A|B) P(B) = P(B|A) P(A)$$

και κατά συνέπεια

$$P(A|B) = P(B|A) \cdot P(A) / P(B) \quad 1.4.1$$

Αυτή η απλή σχέση είναι το περίφημο **Θεώρημα του Bayes**, του οποίου τις συνέπειες θα δούμε στα επόμενα.

Μπορούμε να βρούμε μία άλλη, πιο εύχρηστη, έκφραση για το θεώρημα του Bayes, χρησιμοποιώντας τις εύλογες προτάσεις ότι:

$$\begin{aligned} P(B) &= P(B \text{ AND } A) + P(B \text{ AND } \neg A) \\ P(\neg A) &= 1 - P(A) \end{aligned} \quad 1.4.2$$

όπου με το σύμβολο $P(\neg A)$, παρίσταται η πιθανότητα να συμβεί ένα γεγονός **που δεν ανήκει** στην κατηγορία A.

Η πιθανότητα ένα γεγονός να ανήκει στην κατηγορία B εκφράζεται, σύμφωνα με τις σχέσεις 1.4.2 και 1.3.4 ως:

$$\begin{aligned} P(B) &= P(B \text{ AND } A) + P(B \text{ AND } \neg A) = P(B|A) \cdot P(A) + P(B|\neg A) \cdot P(\neg A) \\ P(B) &= P(B|A) \cdot P(A) + P(B|\neg A) \cdot [1 - P(A)] \end{aligned} \quad 1.4.3$$

Αντικαθιστώντας την 1.4.3 στην 1.4.1, το Θεώρημα Bayes εκφράζεται ως:

$$P(A|B) = \frac{P(B|A) \cdot P(A)}{P(B|A) \cdot P(A) + P(B|\neg A) \cdot [1 - P(A)]} \quad 1.4.4$$

Η καινούργια έκφραση του θεωρήματος του Bayes, με τη σχέση 1.4.4, φαίνεται πολύ πιο πολύπλοκη από την προηγούμενη και ο αναγνώστης ίσως να αναρωτιέται «προς τι η πολυπλοκότητα για κάτι που δεν είναι καθόλου προφανές η αξία του». Η χρησιμότητα της σχέσης 1.4.4 μπορεί εύκολα ναδειχθεί σε πρακτικές εφαρμογές.

Ας ξαναγυρίσουμε στο παράδειγμα υπολογισμού της αποδοτικότητας μιας μετρητικής

διάταξης. Καταλήξαμε στο παράδειγμα της προηγούμενης παραγράφου ότι, δείγματα ελέγχου μπορούν να κατασκευασθούν με την συνεργασία δύο (τουλάχιστον) ανιχνευτών. Στην πράξη, τέτοια αμιγή δείγματα είναι αδύνατο να κατασκευασθούν, λόγω ενός άλλου περιορισμού της ανιχνευτικής ικανότητας των πραγματικών ανιχνευτικών διατάξεων, η οποία αναφέρεται στην Λανθασμένη Πιστοποίηση.

Ορισμός της Λανθασμένης Πιστοποίησης (*Misidentification*): Λανθασμένη Πιστοποίηση είναι η πιθανότητα ο ανιχνευτής αντικειμένων τύπου X να διεγερθεί και να καταγράψει ως «τύπου X » αντικείμενα διαφορετικού τύπου.

Θα μπορούσε να υποθέσει κανείς ότι θα μπορούσε να μετρήσει την Λανθασμένη Πιστοποίηση ενός ανιχνευτή εάν διέθετε το κατάλληλο δείγμα ελέγχου. Είναι σαφές όμως, ότι για την δημιουργία δειγμάτων ελέγχου θα πρέπει να χρησιμοποιήσουμε άλλες ανιχνευτικές διατάξεις, οι οποίες με την σειρά τους πάσχουν από περιορισμένες μετρητικές ικανότητες. Διέξοδο από αυτόν τον «φαύλο κύκλο» δίνει η εφαρμογή του θεωρήματος Bayes.

Παράδειγμα 1.4.1

Έστω μία γραμμή παραγωγής τυπωμένων κυκλωμάτων (τ.κ.). Έστω ακόμα ότι είναι γνωστό ότι τα 5% των τ.κ. που παράγονται είναι ελαττωματικά ενώ τα 95% τηρούν όλες τις λειτουργικές προδιαγραφές. Προκειμένου να μη διοχετευθούν ελαττωματικά τ.κ. στην αγορά χρησιμοποιείται ένα αυτόματο σύστημα επιλογής το οποίο μπορεί να αναγνωρίζει ένα λειτουργικό τ.κ με πιθανότητα 90% (αποδοτικότητα). Παράλληλα όμως, υπάρχει πιθανότητα 6% (λανθασμένη πιστοποίηση) να κατατάζει ένα ελαττωματικό τ.κ. στα λειτουργικά. Ποια είναι η πιθανότητα, κάθε φορά που το σύστημα επιλογής χαρακτηρίζει κάποιο τ.κ. ως λειτουργικό να είναι στην πραγματικότητα λειτουργικό ;

Το παράδειγμα αναφέρεται σε ένα πολύ σύννηθες πρόβλημα λειτουργικότητας μετρητικού οργάνου με περιορισμένη αποδοτικότητα και μη μηδενική πιθανότητα λανθασμένης

πιστοποίησης. Το σύστημα αυτόματης επιλογής, στην περίπτωση όπου του παρουσιάζονται μόνο λειτουργικά τ.κ. θα αναγνώριζε μόνο τα 90% από αυτά ως λειτουργικά ενώ στην περίπτωση που του παρουσιάζονται μόνο ελαττωματικά τ.κ. θα αναγνώριζε το 6% από αυτά ως λειτουργικά.

Ας υποθέσουμε ότι A είναι το σύνολο των τ.κ (λειτουργικών ή ελαττωματικών) τα οποία αναγνωρίζονται ως λειτουργικά από το σύστημα κατά την διάρκεια της λειτουργίας του. Έστω ακόμα ότι Λ και E είναι τα σύνολα των λειτουργικών και ελαττωματικών τ.κ. που πέρασαν μπροστά από το αυτόματο σύστημα επιλογής, ανεξάρτητα από την κατάταξη που υπέστησαν από τον επιλογέα.

Προφανώς,

$$P(\Lambda) = 95\% \text{ , } P(E)=5\% \text{ , } P(A|\Lambda) = 90\% \text{ και } P(A|E) = 6\% \quad 1.4.5$$

Η πιθανότητα ένα τ.κ, το οποίο επελέγη (ως λειτουργικό) να είναι πράγματι λειτουργικό τ.κ., εκφράζεται ως:

$$P(\Lambda|A)=P(A|\Lambda) P(\Lambda) / P(A) \quad 1.4.6$$

Επίσης, από την σχέση 1.4.3, ισχύει ότι:

$$P(A) = P(\Lambda) P(A|\Lambda) + P(E) P(A|E) \quad 1.4.7$$

Επειδή υπάρχουν μόνο δύο δυνατότητες (Λ ή E) θα ισχύουν οι προφανείς σχέσεις

$$P(A|\Lambda) = P(A|\Lambda) \text{ και } 1 - P(A|\Lambda) = P(A|E) \quad 1.4.8$$

Αντικαθιστώντας τις σχέσεις 1.4.8 στην σχέση 1.4.7 καταλήγουμε στον υπολογισμό της πιθανότητας που αντιστοιχεί στην επιλογή ενός τ.κ. ως λειτουργικό :

$$P(A) = P(A|\Lambda) P(\Lambda) + P(A|E) P(E) = 90\% \cdot 95\% + 6\% \cdot 5\% \quad 1.4.9$$

Η πιθανότητα ο επιλογέας να έχει επιλέξει ένα πράγματι λειτουργικό τ.κ. θα δίνεται από την σχέση 1.4.6, μετά από αντικατάσταση από τις σχέσεις 1.4.5 και 1.4.9:

$$P(\Lambda | A) = \frac{0.9 \cdot 0.95}{0.9 \cdot 0.95 + 0.06 \cdot 0.05} \cong 0.9965$$

Η έκφραση 1.4.4 είναι χρήσιμη όταν η κατηγορία A αποτελείται από δύο υποκατηγορίες με μη κοινά στοιχεία. Στην περίπτωση που η κατηγορία A αποτελείται από N αμοιβαία ανεξάρτητες υποκατηγορίες, (A_i , $i=1,2,3 \dots N$) ώστε:

$$\begin{aligned} A_i \text{ AND } A_j &= 0 \\ A_1 \text{ OR } A_2 \text{ OR } \dots \text{ OR } A_N &= A \end{aligned} \quad 1.4.10$$

το θεώρημα του Bayes γράφεται ως:

$$P(A_i | B) = \frac{P(B | A_i) \cdot P(A_i)}{\sum_{i=1}^N P(B | A_i) P(A_i)} \quad 1.4.11$$

Το θεώρημα Bayes αποκτά ιδιαίτερη σημασία εάν οι κατηγορίες A και B αναφέρονται στην γνώση που κατέχουμε όσον αφορά στην ορθότητα μιας υπόθεσης που έρχεται να εξηγήσει το αποτέλεσμα μιας παρατήρησης B.

Παράδειγμα 1.4.2

Ας χρησιμοποιήσουμε σαν παράδειγμα την περίπτωση του Μπάμπη, που πήγε να πει μία μπύρα και συνάντησε τον γείτονά του τον Μάκη. Για να αποφύγουν τις συζητήσεις για το ποιος θα κεράσει, αποφάσισαν να αφήσουν την τύχη να επιλέξει στρίβοντας ένα νόμισμα. Ο Μάκης διάλεξε τρεις φορές «κεφάλια» και τρεις φορές κέρδισε. Ο Μπάμπης έκαμε πως

δεν κατάλαβε αλλά στο σπίτι ρώτησε την γυναίκα του, την Ρούλα την μαθηματικό, τι πιθανότητες είχε ο Μάκης να φέρει τρεις φορές κεφάλια.

«Προφανώς $0.5 \times 0.5 \times 0.5 = 0.125$, δηλαδή 12.5% » του είπε η Ρούλα.

«Βρε, μπας και το νόμισμα είχε κεφάλια και στις δύο πλευρές;» , αναρωτήθηκε ο Μπάμπης.

«Να δούμε είπε η Ρούλα» και έκαμε τους παρακάτω λογαριασμούς:

Έστω οι δύο υποκατηγορίες A_1 και A_2 που η ένωση τους ερμηνεύει το αποτέλεσμα (3 φορές κεφάλια), ώστε:

$A_1 = \{\text{αποτέλεσμα σύμπτωσης}\}$

$A_2 = \{\text{κάλπικο νόμισμα}\}$

Να ονομάσουμε το γεγονός [3φορές κεφάλια], B .

Η πιθανότητα το B να είναι στατιστικό αποτέλεσμα (αποτέλεσμα σύμπτωσης) , είπαμε πως είναι: $P(B|A_1) = 0.125$

Εάν όμως το νόμισμα είναι κάλπικο τότε πάντα θα φέρνει κεφάλια: $P(B|A_2) = 1$

Μετά στράφηκε στον Μπάμπη και ρώτησε. «Πόσα κάλπικα κυκλοφορούν στην αγορά;»

«Να μην είναι ένα στο εκατομμύριο !!», είπε εκείνος.

Τότε η Ρούλα σκούπισε τις σαπουνάδες στην ποδιά της και εφάρμοσε την σχέση 1.4.11 βάζοντας $P(A_2) = 0.000001$

$P(A_2|B) = P(\text{κάλπικο} | 3\text{κεφάλια}) = [1 \times 0.000001] \times [0.125 \times 0.999999 + 1 \times 0.000001]^{-1} = 8 \cdot 10^{-6}$

«Μπάμπη μου, η πιθανότητα το νόμισμα να ήταν κάλπικο, δεδομένου ότι έφερε τρεις φορές κεφάλια, είναι οχτώ στο εκατομμύριο.»

«Ρε γυναίκα, λες ο Μάκης να ρίχνει το νόμισμα με πονηρό τρόπο;»

« Μπάμπη τι λες; ο κουμπάρος να κλέβει !!!!»

«Γιατί ; Τώρα με το χρηματιστήριο οι άνθρωποι αλλάζουν ! Μήπως δεν έχει και αυθαίρετο;»

«Καλά αυθαίρετο έχουμε και εμείς, αλλά πόση εκτιμάς την πιθανότητα ο Μάκης να

κλέβει»

«E!! Δεν βάζεις πέντε στα εκατό;»

Η Ρούλα λογάριασε ξανά.

Έστω ότι A_1 είναι η περίπτωση το αποτέλεσμα B (τρεις φορές κεφάλια) να οφείλεται σε στατιστικές διακυμάνσεις (αποτέλεσμα σύμπτωσης).

Έστω ότι A_2 είναι η περίπτωση ο Μάκης να κλέβει, και ξεχνάμε την περίπτωση να υπάρχουν κάλπικα (ούτως ή άλλως η πιθανότητα των κάλπικων είναι μικρή)

$$P(B|A_1)=0.125$$

$$P(A_2)=0.05 \text{ (... πέντε στα εκατό εκτιμά την πιθανότητα ο Μάκης)}$$

Επειδή υπάρχουν μόνο δύο περιπτώσεις θα πρέπει: $P(A_1)=1-P(A_2)$. Συνεπώς:

$$P(A_1)=1-0.05=0.95$$

$$P(B|A_2)=1, \text{ γιατί εάν κλέβει ο Μάκης όλο κεφάλια θα φέρνει.}$$

Η Ρούλα ξαναεφάρμοσε την σχέση 1.4.11.

$$P(\text{κλεψιά}|3 \text{ κεφάλια})=[1 \times 0.05] \times [0.125 \times 0.95 + 1 \times 0.05]^{-1} = 30\%$$

Η Ρούλα κοίταξε τον Μπάμπη, άναψε τσιγάρο και είπε.

«Darling μίλα μου για το χρηματιστήριο και τους θεσμικούς επενδυτές!!!»

Στο προηγούμενο παράδειγμα A_i είναι μία υπόθεση που «ερμηνεύει» το αποτέλεσμα, B . Εάν θα θέλαμε να ποσοτικοποιήσουμε την επιτυχία της ερμηνείας θα πρέπει να εκτιμήσουμε την πιθανότητα $P(B|A_i)$, δηλαδή την πιθανότητα, δεδομένου ότι η υπόθεση A_i είναι σωστή, το αποτέλεσμα να είναι το B .

Το θεώρημα του Bayes μας δίνει την δυνατότητα υπολογισμού της πιθανότητας $P(A_i|B)$, η οποία εκφράζει την εμπιστοσύνη μας στη υπόθεση A_i μετά την παρατήρηση του αποτελέσματος B (ύστερη γνώση).

Η εφαρμογή του θεωρήματος του Bayes χρήζει της $P(A_i)$, της πιθανότητας η υπόθεση A_i να είναι σωστή, ή αλλιώς πόσο πιστεύουμε ότι η υπόθεση A_i είναι σωστή. Προφανώς η

$P(A_i)$ ορίζεται υποκειμενικά (εμπλέκοντας πρότερες γνώσεις και εμπειρίες), όπως χαρακτηριστικά δείχθηκε στο προηγούμενο παράδειγμα.

Ο υποκειμενικός χαρακτήρας της προτέρας γνώσης, όπως εκφράζεται από την πιθανότητα $P(A_i)$, συγκεντρώνει τα πυρά των πολεμίων της Bayesian μεθόδου που περιγράψαμε στα προηγούμενα. Βεβαίως υπάρχουν απαντήσεις σ' αυτή την κριτική, που ξεφεύγουν από τα όρια αυτής της περιγραφής

Στα επόμενα θα χρησιμοποιήσουμε την μέθοδο σε εκτιμήσεις παραμέτρων και εκτίμηση σφαλμάτων και θα έχουμε την δυνατότητα να διατυπώσουμε περισσότερα σχόλια και κρίσεις.

1.5 Τυχαίες μεταβλητές

Τυχαίο γεγονός είναι ένα γεγονός το οποίο έχει περισσότερα από ένα δυνατά αποτελέσματα. Είναι δυνατόν να αντιστοιχίσουμε μία πιθανότητα σε κάθε αποτέλεσμα αλλά η έκβαση ενός τυχαίου γεγονότος δεν είναι προβλέψιμη. **Η μέγιστη ποσότητας γνώσης που μπορούμε να αποκομίσουμε για κάποιο τυχαίο γεγονός αφορά στον προσδιορισμό της πιθανότητας να εξελιχθεί με κάθε ένα από τους δυνατούς τρόπους.**

Παράδειγμα 1.5.1:

Η πιθανότητα ενός μη σταθερού σωματίου ή ενός ασταθούς πυρήνα να διασπαστεί, μετά t δευτερόλεπτα από την δημιουργία του, είναι ανάλογη του $e^{-t/T}$. Συνήθως η σταθερά T καλείται μέσος χρόνος ζωής.

Ανεξάρτητα από την ακρίβεια των μετρητικών μας συσκευών, όταν επιχειρούμε να μετρήσουμε τον χρόνο ζωής ενός πλήθους, ίδιων, ασταθών πυρήνων θα προσδιορίζουμε διαφορετικούς χρόνους ζωής για κάθε έναν από αυτούς. Όσον και εάν φαίνεται παράλογο

είναι αδύνατο να προσδιορίσουμε μία μοναδική τιμή στο μέγεθος « t = χρόνος ζωής του πυρήνα X ». Το μόνο που μπορούμε να επιτύχουμε είναι να προσάψουμε μια πιθανότητα στο ενδεχόμενο να μετρήσουμε τον χρόνο $t=t_1$, ως χρόνο ζωής.

Τον ίδιο, εκθετικό, νόμο τον συναντάμε σε πολλές περιπτώσεις συμπεριλαμβανομένης και της πιθανότητας που έχει ένα εξάρτημα μηχανής, να διαρκέσει χωρίς βλάβη για χρόνο μεγαλύτερο από t .

Οι επαναλαμβανόμενες μετρήσεις της ίδιας Φυσικής ποσότητας δεν καταλήγουν στο ίδιο αποτέλεσμα. Στις περισσότερες των περιπτώσεων, η πιθανότητα να βρούμε την τιμή Q (ή καλύτερα μία τιμή μεταξύ Q και $Q+dQ$) σαν ένα αποτέλεσμα μέτρησης ενός φυσικού μεγέθους με πραγματική τιμή R είναι της μορφής:

$$P(Q)dQ \propto e^{-\frac{(R-Q)^2}{2\sigma^2}} dQ \quad 1.5.1$$

Εν γένει το αποτέλεσμα κάθε μέτρησης συμπεριφέρεται ως τυχαία μεταβλητή που καταλήγει σε διαφορετική τιμή κάθε φορά που επαναλαμβάνεται η μέτρηση. επαναλαμβάνουμε την μέτρηση τις τιμές $x_1, x_2, x_3, \dots, x_N$ καθώς.

1.6 Συνάρτηση πυκνότητας πιθανότητας

Ας θεωρήσουμε ένα πείραμα στο οποίο η κατεύθυνση ενός σωματίου καθορίζεται με δύο διαδοχικές σειρές από ορθογώνιους ανιχνευτές X (κάθετους στον οριζόντιο άξονα, x) και Y (κάθετους στον κατακόρυφο άξονα y). Ας υποθέσουμε επίσης ότι το πάχος των ανιχνευτών X και Y είναι ΔX και ΔY αντίστοιχα.

Κάθε σωματίο, που διαπερνά την μετρητική αυτή διάταξη, κινείται ευθύγραμμα και παράλληλα προς τον άξονα z . Κατά την διέλευση του σωματίου δια μέσου των ανιχνευτών X_i και Y_j (βλέπε Σχήμα 1.6.1) οι ανιχνευτές διεγείρονται και παράγονται ηλεκτρικά σήματα στα άκρα τους. Κάθε σωματίο που ανιχνεύεται, χαρακτηρίζεται από δύο διακριτές μεταβλητές i και j οι οποίες αντιστοιχούν στους ανιχνευτές X_i και Y_j τους οποίους διέσχισε.

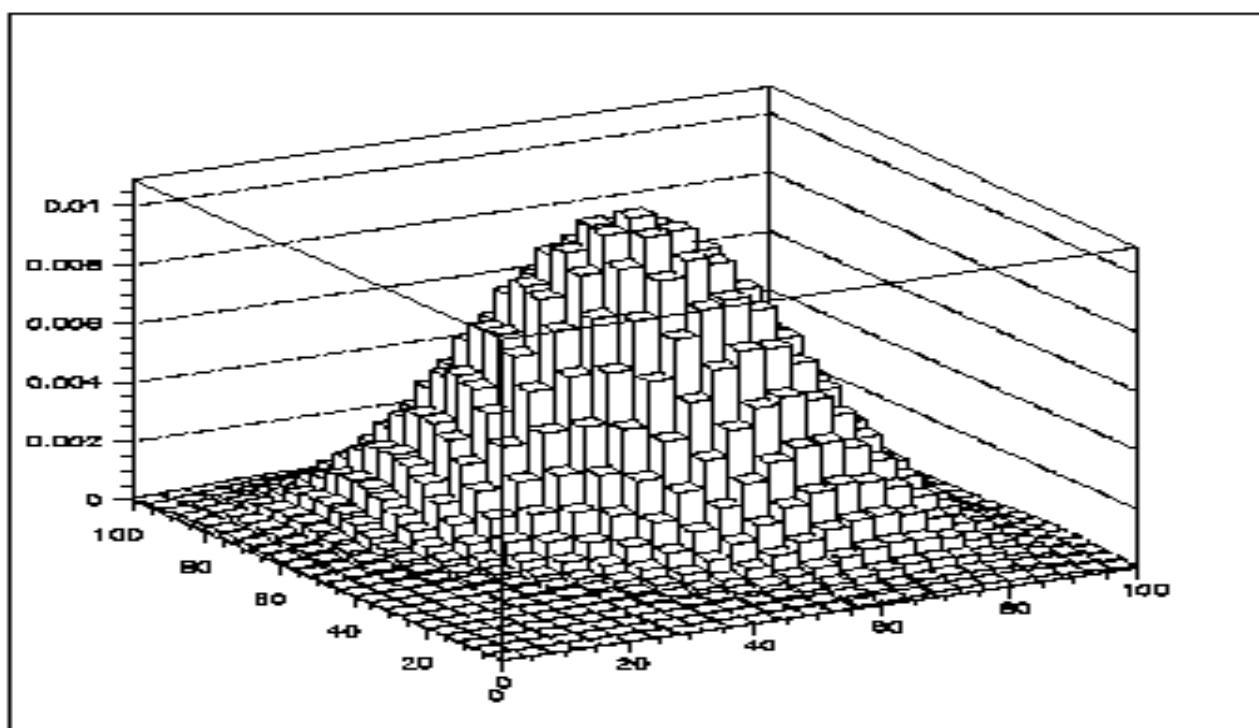
Ας αφήσουμε πολλά σωμάτια, το ένα μετά το άλλο, να διασχίσουν την μετρητική συσκευή σημειώνοντας κάθε φορά τις τιμές που παίρνουν οι μεταβλητές i και j . Παρατηρούμε ότι δεν υπάρχει κανένας τρόπος να προβλέψουμε εκ των προτέρων (πριν περάσει το σωμάτιο) τις τιμές των μεταβλητών i και j αλλά εύκολα μπορούμε να υπολογίσουμε την πιθανότητα οι μεταβλητές αυτές να λάβουν συγκεκριμένες τιμές.

Ας αφήσουμε N σωμάτια, το ένα μετά το άλλο, να διασχίσουν την μετρητική συσκευή σημειώνοντας κάθε φορά τις τιμές που παίρνουν οι μεταβλητές i και j . Στην συνέχεια ας υπολογίσουμε το πλήθος, n_{ij} , των σωματίων που διέγειραν το ζεύγος ανιχνευτών, $\{i,j\}$. Η πιθανότητα, P_{ij} , ένα οποιοδήποτε σωμάτιο να διέλθει δια μέσου των ανιχνευτών X_i και Y_j , εξ ορισμού, δίνεται από την σχέση:

$$P_{ij} = \lim_{N \rightarrow \infty} \frac{n_{ij}}{N} \quad \text{ή} \quad P_{ij} \cong \frac{n_{ij}}{N}, \text{ για } N \text{ μεγάλο} \quad 1.6.1$$

Προφανώς, οι μεταβλητές i και j είναι τυχαίες μεταβλητές, διότι δεν μπορούμε να προβλέψουμε εκ των προτέρων την τιμή τους αλλά μπορούμε να εκτιμήσουμε την πιθανότητα να λάβουν κάθε μία από τις δυνατές τιμές τους.

Στο Σχήμα 1.6.2 (α) παρουσιάζεται η πιθανότητα διέγερσης κάθε ζεύγους ανιχνευτών, $\{i,j\}$, όπως υπολογίσθηκε, σε ένα πείραμα, βάσει της σχέσης 1.6.1..:



Ουσιαστικά, η πιθανότητα P_{ij} αναφέρεται στην πιθανότητα ένα σωματίο να περάσει μέσα από μία επιφάνεια, διαστάσεων $\Delta X \cdot \Delta Y$, που βρίσκεται στη θέση που προσδιορίζουν οι μεταβλητές $\{i,j\}$. Προκειμένου να εκτιμήσουμε την πιθανότητα ένα σωματίο να περάσει μέσα από τμήμα αυτής της επιφάνειας, έστω d , θα πρέπει να ορίσουμε την **μέση πυκνότητα πιθανότητας**, ρ_{ij} :

$$\rho_{ij} = \frac{P_{ij}}{\Delta X \cdot \Delta Y} \quad 1.6.2$$

Υποθέτοντας ότι το σωματίο έχει την ίδια πιθανότητα να διέλθει από κάθε σημείο της επιφάνειας $\Delta X \cdot \Delta Y$, η πιθανότητα να περάσει από τμήμα, d , της επιφάνειας θα είναι, προσεγγιστικά, ίση με $\rho_{ij} \cdot d$.

Είναι προφανές ότι προκειμένου να υπολογίσουμε με ακρίβεια την πιθανότητα διέλευσης από κάθε στοιχειώδη επιφάνεια της ανιχνευτικής διάταξης θα πρέπει να χρησιμοποιήσουμε ανιχνευτές μηδενικού πάχους ($\Delta X \rightarrow 0, \Delta Y \rightarrow 0$). Στην περίπτωση αυτή, για να καλύψουμε την επιφάνεια διέλευσης των σωματίων, θα πρέπει να χρησιμοποιήσουμε άπειρο αριθμό X και Y ανιχνευτών. Οι τυχαίες μεταβλητές i και j μεταβάλλονται συνεχώς και θα τις συμβολίζουμε με τα σύμβολα x και y αντίστοιχα, η δε πυκνότητα πιθανότητας, $f(x,y)$, θα ορίζεται ως:

$$f(x,y) = \lim_{\substack{\Delta X \rightarrow 0 \\ \Delta Y \rightarrow 0}} \frac{P(x,y)}{\Delta X \cdot \Delta Y} = \lim_{\substack{\Delta X \rightarrow 0 \\ \Delta Y \rightarrow 0 \\ N \rightarrow \infty}} \frac{n_{xy}/N}{\Delta X \cdot \Delta Y} \quad 1.6.3$$

όπου n_{xy} είναι ο αριθμός των σωματίων που πέρασαν από μία επιφάνεια, διαστάσεων $\Delta X \cdot \Delta Y$ που τείνουν στο μηδέν, η οποία βρίσκεται στην θέση $\{x,y\}$.

Η σχέση 1.6.3 είναι ο ορισμός της συνάρτησης πυκνότητας πιθανότητας (probability

density function - pdf) η οποία παρίσταται γραφικά, για το παράδειγμα που συζητήσαμε στο Σχήμα 1.6.2.(β). Οι Φυσικοί καλούν την συνάρτηση πυκνότητας πιθανότητας τυχαίων μεταβλητών, κατανομή των μεταβλητών. Για τον Στατιστικό όμως, κατανομή σημαίνει ολοκληρωμένη συνάρτηση πυκνότητας πιθανότητας. Σ' αυτό το κείμενο, πολλές φορές, θα ακολουθήσουμε την διάλεκτο της κάστας των Φυσικών.

Είναι σαφές ότι η πιθανότητα οποιαδήποτε τυχαία μεταβλητή Z να πάρει τιμή στο διάστημα μεταξύ από z και $z+dz$ δίνεται από την σχέση:

$$P(z)=R(z) \cdot dz \quad 1.6.4$$

όπου $R(z)$ είναι η συνάρτηση πυκνότητας πιθανότητας της τυχαίας μεταβλητής Z .

Όταν το πρόβλημα που μελετάμε (π.χ. το αποτέλεσμα ενός πειράματος) χαρακτηρίζεται από περισσότερες από μια τυχαίες μεταβλητές (π.χ. σύγχρονη μέτρηση θερμοκρασίας, πίεσης και όγκου ενός αερίου), η πυκνότητες πιθανότητας των τυχαίων μεταβλητών, εν γένει, αλληλοεξαρτώνται. Στην περίπτωση αυτή, η πιθανότητα οι τυχαίες μεταβλητές $X_1, X_2, X_3, \dots, X_N$ να έχουν, συγχρόνως, τιμές στα διαστήματα $[x_1, x_1+dx_1], [x_2, x_2+dx_2], [x_3, x_3+dx_3], \dots, [x_N, x_N+dx_N]$, δίνεται από την ακόλουθη σχέση:

$$P(x_1, x_2, x_3, \dots, x_N)=g(x_1, x_2, x_3, \dots, x_N)dx_1dx_2 dx_3 \dots dx_N \quad 1.6.5$$

Όπου $g(x_1, x_2, x_3, \dots, x_N)$ είναι η κοινή (joined) συνάρτηση πυκνότητας πιθανότητας των τυχαίων μεταβλητών.

Όπως αναφέρεται και στην Εισαγωγή του Κεφαλαίου, παριστούμε τις τυχαίες μεταβλητές ως συνιστώσες του διανύσματος τυχαίας μεταβλητής $\vec{x} = \{x_1, x_2, x_3, \dots, x_N\}$ και η σχέση 1.6.5 γράφεται ως:

$$P(\vec{x}) = g(\vec{x})d\vec{x} \quad 1.6.6$$

Το ολοκλήρωμα της συνάρτησης πυκνότητας πιθανότητας, $g(x_1, x_2, x_3, \dots, x_N)$, για όλες τις δυνατές τιμές των τυχαίων μεταβλητών εκφράζει την συνολική πιθανότητα και συνεπώς πρέπει να ισούται με τη μονάδα.

$$\int_{\Omega} g(\vec{x})d^N\vec{x} = \int_{-\infty}^{\infty} \dots \int_{-\infty}^{\infty} g(x_1, x_2, x_3, \dots, x_N)dx_1dx_2 dx_3 \dots dx_N = 1 \quad 1.6.7$$

Όταν η τυχαία μεταβλητή παίρνει διακριτές τιμές, π.χ. το σύνολο των φυσικών αριθμών, η συνάρτηση πυκνότητας πιθανότητας χάνει την χρηστική της αξία. Στην περίπτωση αυτή μιλούμε για συνάρτηση πιθανότητας, $P_k=f(x_k)$ όπου το x_k είναι μία από τις διακριτές τιμές που παίρνει η τυχαία μεταβλητή x .

Στα επόμενα θα συναντήσουμε την συνάρτηση πιθανότητας Poisson, που εκφράζει την πιθανότητα η (θετική και ακέραια) τυχαία μεταβλητή n να πάρει κάποια ως τιμή κάποιον από τους φυσικούς αριθμούς. Ο συναρτησιακός τύπος της Poissonian είναι:

$$f(n) = \frac{\mu^n \cdot e^{-\mu}}{n!} \quad 1.6.8$$

όπου μ είναι θετική παράμετρος

Επειδή, πάντοτε, η συνολική πιθανότητα θα πρέπει να είναι ίση με την μονάδα, θα πρέπει να ισχύει ότι:

$$\sum_{\Omega} f(x_k) = 1 \quad \text{όπου } \Omega \text{ είναι το σύνολο τιμών της τυχαίας μεταβλητής } x_k$$

Ορίζουμε ως αθροιστική συνάρτηση πιθανότητας (cumulative distribution function), $F(y)$, της τυχαίας μεταβλητής X , με συνάρτηση πυκνότητα πιθανότητας $f(x)$, την πιθανότητα η τυχαία μεταβλητή να πάρει τιμή μικρότερη ή ίση από την τιμή y .

$$F(y) = \int_{-\infty}^y f(x)dx = P(x \leq y) \quad 1.6.9$$

Από την σχέση ορισμού 1.6.9 έπεται ότι $F(-\infty)=0$ και ότι $F(\infty)=1$.

Επίσης από τη ίδια σχέση ορισμού έπονται οι ακόλουθες ιδιότητες της αθροιστικής συνάρτησης πιθανότητας:

- 1 $0 \leq F(y) \leq 1$
- 2 Η συνάρτηση $F(y)$ είναι μονοτονική αύξουσα συνάρτηση
- 3 $P(a \leq x \leq b) = F(b) - F(a)$
- 4 Εάν η αθροιστική συνάρτηση πιθανότητας είναι ασυνεχής στο λ τότε η πιθανότητα η τυχαία μεταβλητή να πάρει την τιμή λ δίνεται από τη σχέση:

$$P(x = \lambda) = \lim_{\delta x \rightarrow 0} [F(\lambda + \delta x) - F(\lambda - \delta x)]$$

- 1 Εάν η αθροιστική συνάρτηση πιθανότητας είναι συνεχής τότε η πιθανότητα η τυχαία μεταβλητή να πάρει την τιμή λ είναι μηδέν: $P(x=\lambda)=0$

Ερώτηση: Αποδείξτε τις ιδιότητες της αθροιστικής συνάρτησης πιθανότητας

Υπόδειξη: Οι ιδιότητες είναι συνέπεια της σχέσης 1.6.9 και αποδεικνύονται εύκολα. Για να κατανοήσετε εύκολα την περίπτωση όπου η αθροιστική συνάρτηση πιθανότητας είναι ασυνεχής στο $y=\lambda$, θεωρήσετε ότι η συνάρτηση πυκνότητας πιθανότητας της τυχαίας μεταβλητής x έχει την ακόλουθη συναρτησιακή μορφή:

$$f(x) = \begin{cases} 0 & \text{για } x \in [-\infty, a) \\ \frac{1}{b-a} & \text{για } x \in [a, b] \\ 0 & \text{για } x \in [b, \infty) \end{cases}$$

και υπολογίσετε την $F(y)$ σύμφωνα με την σχέση 1.6.9

Η αθροιστική συνάρτηση πιθανότητας, θεωρείται από τους στατιστικούς η κατ' εξοχήν ουσιώδης συνάρτηση. Το βασικό επιχείρημα αναφέρεται στο γεγονός ότι η συνάρτηση αυτή αναφέρεται σε πιθανότητα και όχι σε πυκνότητα πιθανότητας. Παρ' όλα αυτά, για τις επόμενες εφαρμογές θα χρησιμοποιήσουμε την πυκνότητα πιθανότητας.

1.7 Μετασχηματισμός των μεταβλητών

Ας υποθεθεί ότι $f(x)$ είναι μία (γνωστή) συνάρτηση πυκνότητας πιθανότητας της τυχαίας μεταβλητής X . Θα θέλαμε να υπολογίσουμε την συνάρτηση πυκνότητας πιθανότητας $g(y)$, της τυχαίας μεταβλητής Y , που προκύπτει από έναν μετασχηματισμό της μορφής $y=h(x)$.

Έστω μία περιοχή δυνατών τιμών της μεταβλητής X , $[x, x+dx]$. Μέσω του μετασχηματισμού $y=h(x)$, η περιοχή τιμών της X μετασχηματίζεται στην περιοχή $[y, y+dy]$ της μεταβλητής Y . Εάν ο μετασχηματισμός είναι ένα προς ένα, η πιθανότητα η τιμή της μεταβλητής X να ανήκει στο διάστημα $[x, x+dx]$ θα πρέπει να είναι ίση με την πιθανότητα η τιμή της μεταβλητής Y να ανήκει στην περιοχή τιμών $[y, y+dy]$. Θα πρέπει δηλαδή να ισχύει η επόμενη ισότητα:

$$g(y)dy = f(x)dx \quad 1.7.1$$

Αλλά $dy = dh(x)$ και συνεπώς η συνάρτηση πυκνότητας πιθανότητας $g(y)$ εκφράζεται ως:

$$g(y) = \frac{f(x)}{\left| \frac{dh}{dx} \right|} \quad 1.7.2$$

όπου χρησιμοποιήσαμε την απόλυτο τιμή για να διασφαλίσουμε ότι η πιθανότητα θα είναι θετική.

Αν ο μετασχηματισμός δεν είναι ένα προς ένα, δηλαδή υπάρχουν περισσότερα του ενός x

$$\left. \begin{array}{l} x_1, x_1 + dx \\ x_2, x_2 + dx \\ \dots\dots\dots \\ x_N, x_N + dx \end{array} \right\} \rightarrow y, y + dy$$

τα οποία αντιστοιχούν στο ίδιο y :

1.7.3

$$g(y) = \sum_i \frac{f(x_i)}{\left| \frac{dh}{dx} \right|_{x=x_i}} \quad \text{όπου } y=h(x_i) \quad 1.7.4$$

Τότε η 1.7.2 γενικεύεται ως:

Η σχέση 1.7.2, εύκολα γενικεύεται για την περίπτωση κοινών συναρτήσεων πυκνότητας πιθανότητας πολλών μεταβλητών.

Έστω οι τυχαίες μεταβλητές $X_1, X_2, X_3, \dots, X_N$ με κοινή συνάρτηση πυκνότητας πιθανότητας την $f(x_1, x_2, x_3, \dots, x_N)$. Έστω οι ένα προς ένα μετασχηματισμοί:

$$\left. \begin{array}{l} y_1 = h_1(x_1, x_2, x_3, \dots, x_N) \\ y_2 = h_2(x_1, x_2, x_3, \dots, x_N) \\ y_3 = h_3(x_1, x_2, x_3, \dots, x_N) \\ \dots\dots\dots \\ y_N = h_N(x_1, x_2, x_3, \dots, x_N) \end{array} \right) \quad 1.7.5$$

Αναζητούμε την συνάρτηση πυκνότητας πιθανότητας, $g(y_1, y_2, y_3, \dots, y_N)$, των τυχαίων μεταβλητών, $Y_1, Y_2, Y_3, \dots, Y_N$.

Επειδή ο μετασχηματισμός είναι ένα προς ένα,

η πιθανότητα η μεταβλητή X_1 να παίρνει τιμές από το διάστημα $[x_1, x_1 + dx_1]$ και συγχρόνως η μεταβλητή X_2 να παίρνει τιμές από το διάστημα $[x_2, x_2 + dx_2]$ και η μεταβλητή X_3 να παίρνει τιμές από το διάστημα $[x_3, x_3 + dx_3]$ κ.ο.κ, **ισούται**

με την πιθανότητα η μεταβλητή Y_1 να παίρνει τιμές από το διάστημα $[y_1, y_1+dy_1]$ και συγχρόνως η μεταβλητή Y_2 να παίρνει τιμές από το διάστημα $[y_2, y_2+dy_2]$ και η μεταβλητή Y_3 να παίρνει τιμές από το διάστημα $[y_3, y_3+dy_3]$ κ.ο.κ .

$$g(y_1, y_2, y_3, \dots, y_N) dy_1 dy_2 dy_3 \dots dy_N = f(x_1, x_2, x_3, \dots, x_N) dx_1 dx_2, dx_3 \dots dx_N \quad \text{ή}$$

$$g(h_1(x_1, x_2, x_3, \dots, x_N), h_1(x_1, x_2, x_3, \dots, x_N), \dots, h_1(x_1, x_2, x_3, \dots, x_N)) |J| dx_1 dx_2, dx_3 \dots dx_N =$$

$$f(x_1, x_2, x_3, \dots, x_N) dx_1 dx_2, dx_3 \dots dx_N$$

$$\begin{vmatrix} \frac{\partial h_1}{\partial x_1} & \frac{\partial h_2}{\partial x_1} & \dots & \frac{\partial h_N}{\partial x_1} \\ \frac{\partial h_1}{\partial x_2} & \frac{\partial h_2}{\partial x_2} & \dots & \frac{\partial h_N}{\partial x_2} \\ \dots & \dots & \dots & \dots \\ \frac{\partial h_1}{\partial x_N} & \frac{\partial h_2}{\partial x_N} & \dots & \frac{\partial h_N}{\partial x_N} \end{vmatrix}$$

όπου $|J|$ είναι η Jacobian ορίζουσα:

$$g(y_1, y_2, y_3, \dots, y_N) = \frac{f(x_1, x_2, x_3, \dots, x_N)}{|J|}$$

Συνεπώς :

Όπου χρησιμοποιούμε την απόλυτη τιμή για να διασφαλίσουμε ότι η πυκνότητα πιθανότητας παίρνει θετικές τιμές.

ρώτηση 1.4 : Ένα σύνηθες πρόβλημα πιστοποίησης σωματιδίων έχει ως εξής :

Έστω ότι μία δέσμη σωματίων, η οποία αποτελείται 95% από πιόνια και 5% από Καόνια. Τα σωματΙΑ διέρχονται δια μέσου ανιχνευτικής διάταξης την οποία σχεδιάσαμε να διεγείρεται από τα πιόνια αλλά όχι απο τα Καόνια. Πρακτικά, η ανιχνευτική μας διάταξη είναι 90% αποδοτική στην ανίχνευση πιονίων, αλλά επίσης έχει και πιθανότητα 6% να αναγνωρίσει ένα Καόνιο ως πιόνιο. Ποιά είναι η πιθανότητα κάθε φορά που ο μετρητής διεγείρεται από ένα σωματΙο, το σωματΙο αυτό να είναι πιόνιο ;

Ποια είναι η πιθανότητα ένα Καόνιο να έχει διεγείρει τον ανιχνευτή, δοθέντος ότι όντως ο ανιχνευτής έχει διεγερθεί;

1.3.2 Ποια είναι η πιθανότητα για ένα πιόνιο να περάσει μέσα από τον ανιχνευτή χωρίς να τον διεγείρει;

1.3.3 Πρακτικά ούτε τα συστατικά της δέσμης είναι γνωστά ούτε οι αποδοτικότητες είναι

γνωστές. Τι προτείνετε;

1.3.4 Διατυπώσετε και λύσετε το αντίστοιχο πρόβλημα στην περίπτωση επιλογής υπαλλήλων, σύμφωνα με κάποια μέθοδο. Η μέθοδος επιλογής έχει αποδοτικότητα 70% να αναγνωρίζει τον κατάλληλο υπάλληλο και 20% πιθανότητα να κατατάσσει τους άνδρες στην κατηγορία των καταλλήλων υπαλλήλων ανεξάρτητα από τα προσόντα τους. Το δείγμα από το οποίο πρέπει να γίνει η επιλογή αποτελείται από 40% άνδρες, 60% ικανούς υποψηφίους και το πραγματικό ποσοστό καταλληλότητας είναι το ίδιο σε γυναίκες και άνδρες.

Πρόβλημα 1.3.7: Στο προηγούμενο παράδειγμα τα δυνατά σενάρια αναφερόταν σε μία μόνο παθογόνο αιτία (η νόμισμα κάλπικο ή ο Μάκης κλέβει). Βρείτε την πιθανότητα το αποτέλεσμα να ωφείλεται στις ικανότητες του Μάκη εάν:

α) στην αγορά κυκλοφορούν 0.001 κάλπικα και η πιθανότητα ο Μάκης να κλέβει είναι 0.03.

β) όταν επιπλέον των δεδομένων του σκέλους α), ο Μάκης κατορθώνει να φέρει κεφάλια μόνο το 96% των περιπτώσεων.

Πρόβλημα 1.5.1: Εάν η τυχαία μεταβλητή x , έχει συνάρτηση πυκνότητας πιθανότητας την $f(x)$, ποιά είναι η συνάρτηση πυκνότητας πιθανότητας της $y=x^2$;

1.8 Αναμενόμενη τιμή και τετραγωνική απόκλιση

Ας υποτεθεί ότι μετρούμε την ορμή (στα επόμενα την συμβολίζουμε με p) των ηλεκτρονίων, που χρησιμοποιούνται για θεραπευτικούς σκοπούς (ακτινοβολήσεις) σε κάποιο νοσοκομείο. Χρησιμοποιούμε, για τον λόγο αυτό, μία μετρητική συσκευή¹ η οποία παρεμβάλλεται κάθετα στην διεύθυνση των ηλεκτρονίων που εξέρχονται από τον επιταχυντή.

Δεν θα πρέπει να μας εκπλήσσει το γεγονός ότι, αν και μας έχουν διαβεβαιώσει ότι ο επιταχυντής επιταχύνει ακριβώς με τον ίδιο τρόπο όλα τα ηλεκτρόνια, βρίσκουμε διαφορετικές τιμές για την ορμή των ηλεκτρονίων. Στην συντριπτική τους πλειοψηφία, τα αποτελέσματα επαναλαμβανόμενων μετρήσεων του ιδίου μεγέθους είναι διαφορετικά κάθε φορά. Το αποτέλεσμα της μέτρησης είναι μία τυχαία μεταβλητή και η συνάρτηση πυκνότητας πιθανότητας, $f(p)$, συνήθως καλείται διακριτική ικανότητα της μετρητικής συσκευής. Δηλαδή, η ποσότητα $f(p_0)dp$ εκφράζει την πιθανότητα η μέτρηση μας να βρίσκεται μεταξύ p_0 και p_0+dp .

Ας υποθέσουμε ότι μας ενδιαφέρει να μελετήσουμε την ενέργεια (στα επόμενα θα την συμβολίζουμε με ε) των ηλεκτρονίων. Θα υποθέσουμε ακόμα ότι γνωρίζουμε με πολύ μεγάλη ακρίβεια την μάζα, m , του ηλεκτρονίου.

Θα χρησιμοποιήσουμε την σχέση της Ειδικής Σχετικότητας (c είναι η ταχύτητα του φωτός στο κενό)

$$\varepsilon = \sqrt{p^2 \cdot c^2 + m^2 \cdot c^4} \quad 1.8.1$$

που συνδέει την ενέργεια με την μάζα και την ορμή για να υπολογίζουμε την ενέργεια κάθε ηλεκτρονίου αφού έχουμε μετρήσει την ορμή του, p . Η σχέση 1.8.1 είναι σχέση μετασχηματισμού, όπως η σχέση 1.7.1 της προηγούμενης παραγράφου. Η ενέργεια, ε , είναι και αυτή τυχαία μεταβλητή διότι η τιμή της καθορίζεται μέσω της 1.8.1 από την τιμή της μέτρησης της ορμής, p . Η συνάρτηση πυκνότητας πιθανότητας της ενέργειας, $g(\varepsilon)$, εκφράζεται σύμφωνα με την 1.7.2, ως²:

$$g(\varepsilon) = \frac{f(p)}{\left| \frac{d\varepsilon}{dp} \right|} \quad 1.8.2$$

¹ Η μέτρηση της ορμής ενός φορτισμένου σωματίου, στηρίζεται στην καμπύλωση της τροχιάς του κατά την δίοδο του μέσα από μαγνητικό πεδίο.

Η αναμενόμενη τιμή της τυχαίας μεταβλητής ε , (συμβολίζεται ως $E(\varepsilon)$) ορίζεται από την σχέση

$$E(\varepsilon) = \int_{\Omega} \varepsilon \cdot g(\varepsilon) d\varepsilon \quad 1.8.3$$

όπου $g(\varepsilon)$ είναι η συνάρτηση πυκνότητας πιθανότητας και Ω δηλώνει όλο το πεδίο ορισμού της μεταβλητής ε .

Η ποσότητα 1.8.3 συνήθως ονομάζεται και **μέση τιμή του ε** .

Στο παράδειγμα του προσδιορισμού της ενέργειας των ηλεκτρονίων, επειδή η ενέργεια παίρνει θετικές τιμές, η σχέση 1.8.3 γράφεται ως:

$$E(\varepsilon) = \int_{\Omega} \varepsilon \cdot g(\varepsilon) d\varepsilon = \int_{-\infty}^{\infty} \varepsilon \cdot g(\varepsilon) d\varepsilon = \int_0^{\infty} \varepsilon \cdot g(\varepsilon) d\varepsilon$$

Με λίγο περισσότερη άλγεβρα, χρησιμοποιώντας τις 1.8.3 και 1.8.2, καταλήγουμε ότι:

$$E(\varepsilon) = \int_0^{\infty} \varepsilon \cdot \frac{f(p)}{\left| \frac{d\varepsilon}{dp} \right|} d\varepsilon = \int_0^{\infty} \varepsilon \cdot \frac{f(p)}{\frac{d\varepsilon}{dp}} dp = \int_0^{\infty} \varepsilon \cdot f(p) dp = \int_0^{\infty} \sqrt{p^2 \cdot c^2 + m^2 \cdot c^4} \cdot f(p) dp \quad 1.8.4$$

Η σχέση 1.8.4 εκφράζει την αναμενόμενη τιμή του ε ως προς την συνάρτηση πυκνότητας πιθανότητας του p , $f(p)$, εφ' όσον η ε είναι συνάρτηση του p .

Το αποτέλεσμα 1.8.4 έχει γενική ισχύ, για κάθε τυχαία μεταβλητή που είναι συνάρτηση άλλης τυχαίας μεταβλητής (ή άλλων τυχαίων μεταβλητών).

Η αναμενόμενη τιμή μιας, μονοτονικής συνάρτησης $\varepsilon(x)$ μιας τυχαίας μεταβλητής x με συνάρτηση πυκνότητας πιθανότητας $f(x)$ δίνεται από

$$E(\varepsilon) = \int \varepsilon(x) \cdot f(x) dx \quad 1.8.5$$

Η αναμενόμενη τιμή της τυχαίας μεταβλητής x , σύμφωνα με την 1.8.5 είναι:

² Στο συγκεκριμένο παράδειγμα, επειδή το μέτρο της ορμής, p , και η ενέργεια, ε , είναι πάντα θετικές ποσότητες η σχέση μετασχηματισμού 1.8.1 είναι σχέση ένα προς ένα.

$$E(x) = \int x \cdot f(x) dx \quad 1.8.6$$

Ορισμός:

Η αναμενόμενη τιμή της μεταβλητής $(x-E(x))^2$ καλείται διασπορά (ή τετραγωνική απόκλιση) και σύμφωνα με την 1.8.5 εκφράζεται ως:

$$\begin{aligned} \mu &= E(x) = \int x \cdot f(x) dx \\ E((x-\mu)^2) &= \int (x-\mu)^2 \cdot f(x) dx \end{aligned} \quad 1.8.7$$

Προσέξτε ότι και οι δύο αναμενόμενες τιμές (μέση τιμή και τετραγωνική απόκλιση) είναι αριθμοί και όχι συναρτήσεις των τυχαίων μεταβλητών.

Η τετραγωνική απόκλιση, ή διασπορά, μίας μονοτονικής συνάρτησης, έστω $\varepsilon(x)$, της τυχαίας μεταβλητής x με συνάρτηση πυκνότητας πιθανότητας την $f(x)$, δίνεται από την σχέση:

$$\begin{aligned} \rho &= E(\varepsilon(x)) = \int \varepsilon(x) \cdot f(x) dx \\ E((\varepsilon(x)-\rho)^2) &= \int (\varepsilon(x)-\rho)^2 \cdot f(x) dx \end{aligned} \quad 1.8.8$$

Για να αποκτήσετε φυσικο-γεωμετρική εποπτεία των εννοιών της μέσης τιμής και τετραγωνικής απόκλισης φανταστείτε ένα σώμα που εκτείνεται σε μία μόνο διάσταση, x , (ευθύγραμμη ράβδο αμελητέου πάχους) με πυκνότητα κατανομής της μάζας του $f(x)$ και μάζα ίση με την μονάδα.. Η μέση τιμή όπως ορίσθηκε στην 1.8.6 αντιστοιχεί στην συντεταγμένη του κέντρου μάζας, ενώ η τετραγωνική διασπορά, της σχέσης 1.8.7, είναι αντίστοιχος της ροπής αδρανείας ως προς άξονα που περνά από το κέντρο μάζας.

Εάν η τυχαία μεταβλητή παίρνει διακριτές τιμές, η μέση τιμή και η διασπορά, ή τετραγωνική απόκλιση, ορίζονται σαν αθροίσματα πάνω σε όλες τις δυνατές τιμές ($\Omega = \{x_1, x_2, x_3, \dots, x_N\}$) της τυχαίας μεταβλητής :

$$E(x) = \mu = \sum_{i=1}^N x_i \cdot f(x_i) \quad , \quad V(x) = \sum_{i=1}^{\Omega} (x_i - \mu)^2 \cdot f(x_i) \quad 1.8.9$$

Η έκφραση για την διασπορά, σχέση 1.8.7, μπορεί να απλοποιηθεί περισσότερο. Πράγματι,

$$\begin{aligned} V(x) &= E[(x - \mu)^2] = \int_{-\infty}^{\infty} (x^2 - 2\mu \cdot x + \mu^2) \cdot f(x) dx \\ &= \int_{-\infty}^{\infty} x^2 \cdot f(x) dx - 2\mu \int_{-\infty}^{\infty} x \cdot f(x) dx + \mu^2 \int_{-\infty}^{\infty} f(x) dx \quad 1.8.10 \\ &= \int_{-\infty}^{\infty} x^2 \cdot f(x) dx - 2\mu \cdot \mu + \mu^2 = E(x^2) - E^2(x) \end{aligned}$$

Ερώτηση: Δείξτε ότι η σχέση 1.8.8 απλοποιείται ως εξής:

$$V(\varepsilon(x)) = E(\varepsilon(x)^2) - E^2(\varepsilon(x)) \quad 1.8.11$$

Είναι εύκολο να αποδειχθεί ότι η αναμενόμενη τιμή είναι ένας γραμμικός τελεστής.

Εάν $\lambda(x)$ και $k(x)$ είναι δύο τυχαίες μεταβλητές, οι οποίες ορίζονται σαν μονοτονικές συναρτήσεις της τυχαίας μεταβλητής x με συνάρτηση πυκνότητας πιθανότητας την $f(x)$, είναι εύκολο να δείξουμε ότι για κάθε ζευγάρι πραγματικών αριθμών a και b ισχύει ότι:

$$E(a \cdot \lambda(x) + b \cdot k(x)) = a \cdot E(\lambda(x)) + b \cdot E(k(x)) \quad 1.8.12$$

Ερώτηση : Να αποδειχθεί η σχέση 1.8.12

Υπόδειξη: Χρησιμοποιήσετε την σχέση 1.8.5 για να εκφράσετε ως ολοκληρώματα τα δύο μέλη της έκφρασης 1.8.12

Η έννοια της αναμενόμενης τιμής και της διασποράς μπορεί εύκολα να γενικευθεί και σε περιπτώσεις περισσότερης της μίας τυχαίων μεταβλητών.

Έστω ότι η $f(x,y)$ είναι η κοινή συνάρτηση πυκνότητας πιθανότητας δύο τυχαίων μεταβλητών x , και y . Ας θεωρήσουμε επίσης ότι $g(x,y)$ είναι μία συνάρτηση των παραπάνω μεταβλητών³. Οι επόμενες σχέσεις είναι γενικεύσεις των ορισμών που δώσαμε ενωρίτερα σ' αυτή την παράγραφο.

³ Στο προηγούμενο παράδειγμα της δέσμης ηλεκτρονίων μετρούσαμε μόνο την ορμή και υποθέσαμε 'τι ξέρουμε την μάζα του ηλεκτρονίου με μεγάλη ακρίβεια. Ας υποθεθεί ότι τώρα μετρούμε την ορμή και την ενέργεια κάθε ηλεκτρονίου, και θέλουμε να υπολογίσουμε την μάζα του ηλεκτρονίου. Αντιστοιχίζουμε τα x, y σε p, ε αντίστοιχα.

$$g(x, y) = \frac{1}{c^2} \cdot (y^2 - x^2 \cdot c^2)^{\frac{1}{2}} = \frac{1}{c^2} (\varepsilon^2 - p^2 \cdot c^2)^{\frac{1}{2}} = m$$

$f(x,y) = f(p,\varepsilon) \equiv$ η κοινή συνάρτηση πυκνότητας πιθανότητας των τυχαίων μεταβλητών p και ε .

$$E(g(x, y)) = \int_{-\infty}^{\infty} \int_{-\infty}^{\infty} g(x, y) f(x, y) dx dy$$

$$V(g(x, y)) = E([g(x, y) - E(g(x, y))]^2) = \int_{-\infty}^{\infty} \int_{-\infty}^{\infty} [g(x, y) - E(g(x, y))]^2 \cdot f(x, y) dx dy$$

$$\mu_x = E(x) = \int_{-\infty}^{\infty} \int_{-\infty}^{\infty} x f(x, y) dx dy$$

1.8.1

$$\mu_y = E(y) = \int_{-\infty}^{\infty} \int_{-\infty}^{\infty} y f(x, y) dx dy$$

$$\sigma_x^2 = V(x) = E((x - \mu_x)^2) = \int_{-\infty}^{\infty} \int_{-\infty}^{\infty} (x - \mu_x)^2 f(x, y) dx dy$$

$$\sigma_y^2 = V(y) = E((y - \mu_y)^2) = \int_{-\infty}^{\infty} \int_{-\infty}^{\infty} (y - \mu_y)^2 f(x, y) dx dy$$

1.9 Συνδιασπορά

Ας παραστήσουμε με X το αποτέλεσμα της τυχαίας και ισοπίθανης επιλογής μίας ημέρας της εβδομάδας. Προφανώς η τυχαία μεταβλητή X παίρνει μία τις ακόλουθες «διακριτές τιμές»: {Δευτέρα, Τρίτη, Τετάρτη, Πέμπτη, Παρασκευή, Σάββατο, Κυριακή}.

Ερώτηση: Δείξτε ότι η συνάρτηση πιθανότητας της τυχαίας μεταβλητής X είναι: $f(x_k) = 1/7$

Ας παραστήσουμε επίσης με Y το αποτέλεσμα της παρατήρησης του ύψους των βροχοπτώσεων κατά την διάρκεια της ημέρας. Για απλότητα ας υποθέσουμε ότι η τυχαία μεταβλητή Y παίρνει μία από τις δύο διακριτές τιμές $\{0, 1\}$, όπου $y=0$ σημαίνει ότι δεν έβρεξε καθόλου ενώ $y=1$ σημαίνει ότι έβρεξε.

Ερώτηση: Κατά την διάρκεια είκοσι ετών, όπου οι κλιματικές συνθήκες ενός τόπου διατηρούνται σταθερές, έβρεξε συνολικά 1825 ημέρες. Δείξτε ότι η συνάρτηση

$$\text{πιθανότητας της τυχαίας μεταβλητής } Y \text{ είναι: } g(y_k) = \begin{cases} \frac{3}{4} & \text{για } y = y_1 = 0 \\ \frac{1}{4} & \text{για } y = y_2 = 1 \end{cases}$$

Εάν συμβολίσουμε με $P(x_k, y_n)$ (όπου $k=1,2$ και $n=1,2,3,4,5,6,7$) την κοινή συνάρτηση πιθανότητας των τυχαίων μεταβλητών X και Y , εύκολα αποδεικνύεται η αλήθεια της επόμενης πρότασης:

$P(0, \text{Δευτέρα})=P(0, \text{Τρίτη})=P(0, \text{Τετάρτη})=P(0, \text{Πέμπτη})=P(0, \text{Παρασκευή})=P(0, \text{Σάββατο})=P(0, \text{Κυριακή})$ 1.9.1
Πράγματι η πιθανότητα βροχόπτωσης είναι ανεξάρτητη⁴ της ημέρας της βδομάδας που διανύουμε.

Ερώτηση: Κάνοντας χρήση της σχέσης 1.9.1, δείξτε ότι: $P(x_k, y_n)=f(x_k) \cdot g(y_n)$ ($k=1,2$ και $n=1,2,3,4,5,6,7$) όπου οι συναρτήσεις πιθανότητας $f(x_k)$ και $g(y_n)$ ορίσθηκαν στις δύο προηγούμενες Ερωτήσεις

Έστω τώρα η τυχαία μεταβλητή, Z , που αντιστοιχεί σε μία ισοπίθανη επιλογή μηνός. Προφανώς η τυχαία μεταβλητή Z παίρνει διακριτές τιμές (έναν από τους δώδεκα μήνες, π.χ. $z_3=\text{Μάρτιος}$) και η συνάρτηση πιθανότητας επιλογής ενός μηνός θα είναι: $w(z_m)=1/12$ για $m=1,2,\dots,12$.

Η κοινή συνάρτηση πυκνότητας των μεταβλητών Y και Z δεν εμφανίζει ανάλογη ιδιότητα με αυτή της σχέσης 1.9.1. Πράγματι,

$$R(0, \text{Ιούνιος}) \neq R(0, \text{Νοέμβριος}) \quad 1.9.2$$

Το προηγούμενο παράδειγμα κατέληξε στην παρατήρηση ότι όταν μία μέτρηση αφορά περισσότερες από μία τυχαία μεταβλητή δεν είναι απαραίτητο οι τυχαίες μεταβλητές να είναι μεταξύ τους ανεξάρτητες. Το αντίθετο μάλιστα, στη συντριπτική πλειοψηφία των πειραμάτων που εκτελούμε οι, συγχρόνως, μετρούμενες τυχαίες μεταβλητές αλληλεξαρτούνται.

⁴ Το επιχείρημα δεν είναι απολύτως σωστό. Η πιθανότητα βροχής εξαρτάται από την συγκέντρωση ρύπων στην ατμόσφαιρα. Στο παλιό ομιχλώδες Λονδίνο έβρεχε λιγότερο τις Κυριακές.

Ερώτηση: Κάνοντας χρήση της σχέσης 1.9.2, δείξτε ότι: $R(x_k, y_n) \neq f(x_k) \cdot w(y_n)$ ($k=1,2$ και $n=1,2,\dots,12$) όπου $f(x_k)$ και $w(y_n)$ είναι οι συναρτήσεις πιθανότητας .

Η αλληλεξάρτηση επιβάλλεται συνήθως από νόμους διατήρησης. Σαν παράδειγμα φανταστείτε την έκρηξη μιας οβίδας κατά την οποία ελευθερώνεται συνολική ενέργεια, E , η οποία μοιράζεται σε τρία θραύσματα. Μπορούμε, εάν ξέρουμε τις μάζες των θραυσμάτων, να υπολογίσουμε την πιθανότητα το θραύσμα A να έχει κινητική ενέργεια ίση με τα (ας πούμε) $\frac{3}{4}$ της απελευθερωμένης ενέργειας, E . Ομοίως μπορούμε να υπολογίσουμε την πιθανότητα το θραύσμα B να έχει κινητική ενέργεια ίση με τα $\frac{3}{4}$ της απελευθερωμένης ενέργειας E . Αλλά η πιθανότητα κάθε ένα από τα θραύσματα, A και B , να έχουν συγχρόνως κινητικές ενέργειες ίσες με τα $\frac{3}{4}$ της απελευθερωμένης ενέργειας είναι προφανώς μηδέν.

Θα πρέπει όμως να σημειωθεί ότι σε πολλές περιπτώσεις, αυτή καθ' εαυτή η διαδικασία της μέτρησης επιβάλλει την αλληλεξάρτηση⁵ μεταξύ των, συγχρόνως μετρομένων, ποσοτήτων. .

Γενικά οι τυχαίες μεταβλητές $\vec{x} \equiv \{x_1, x_2, \dots, x_N\}$, π.χ. N φυσικές ποσότητες που προσδιορίζονται συγχρόνως σε ένα πείραμα, είναι δυνατόν είτε να είναι **αμοιβαία ανεξάρτητες** ή να **αλληλεξαρτώνται**.

⁵ Σαν παράδειγμα φανταστείτε ότι θέλετε να μετρήσετε την συντεταγμένη της θέσης ενός απειροελάχιστου αντικειμένου, το οποίο κινείται ευθύγραμμα και ομαλά παράλληλα με τον άξονα X , χρησιμοποιώντας μία δέσμη ακτίνων X . Προκειμένου να επιτύχετε την μεγίστη ακρίβεια θα πρέπει το μήκος κύματος, λ , των ακτίνων X (οι ακτίνες X είναι ηλεκτρομαγνητικό κύμα, φως μεγάλης συχνότητας) να είναι μικρότερο των διαστάσεων, d , του αντικειμένου.

$$\lambda \leq d \quad 1.9.1.a$$

Ας υποθέσουμε τώρα ότι θα θέλαμε τώρα να μετρήσουμε συγχρόνως και την ταχύτητα του αντικειμένου. Στην περίπτωση αυτή θα πρέπει να μετρήσουμε με μεγάλη ακρίβεια τις συντεταγμένες θέσεις, x_1 και x_2 , στις οποίες βρίσκεται το αντικείμενο τις χρονικές στιγμές, t_1 και t_2 , και στην συνέχεια να υπολογίσουμε την ταχύτητα, v , ως:

$$v = (x_2 - x_1) / (t_2 - t_1).$$

Είναι σαφές ότι, με την προϋπόθεση ότι διαθέτουμε ακριβές χρονόμετρο, θα πρέπει να εξασφαλίσουμε την συνθήκη 1.9.1.a προκειμένου να επιτύχουμε μεγάλη ακρίβεια στην μέτρηση της θέσης. Είναι γνωστό ότι προκειμένου να ελαχιστοποιήσουμε το μήκος κύματος της ακτινοβολίας θα πρέπει να χρησιμοποιήσουμε ακτίνες X μεγάλης ενέργειας, E , διότι:

$$E = h \cdot v = h \cdot c / \lambda \quad 1.9.1.b$$

όπου h είναι η σταθερά του Planck, v είναι η συχνότητα της ακτινοβολίας και c είναι η ταχύτητα του φωτός. Στην περίπτωση όμως αυτή, κατά την πρώτη μέτρηση της θέσης οι ενεργειακές ακτίνες X θα αλλάξουν την αρχική ταχύτητα του αντικειμένου. Το αποτέλεσμα της μέτρησης της ταχύτητας εξαρτάται από το αποτέλεσμα μέτρησης της θέσης. Είναι σαφές ότι όσο μεγαλύτερη ακρίβεια επιζητούμε στην μέτρηση της θέσης τόσο μεγαλύτερη ανακρίβεια θα έχουμε στην μέτρηση της ταχύτητας.

Αμοιβαία ανεξάρτητες είναι οι τυχαίες μεταβλητές των οποίων η κοινή πυκνότητα συνάρτησης πιθανότητας, $f(x_1, x_2, \dots, x_N)$, είναι εντελώς παραγοντοποιήσιμη, δηλαδή:

$$f(\vec{x}) = f(x_1, x_2, \dots, x_N) = f(x_1) \cdot f(x_2) \cdot \dots \cdot f(x_N) \quad 1.9.3$$

Σ' αντίθετη περίπτωση οι τυχαίες μεταβλητές καλούνται αμοιβαία εξαρτόμενες.

Προκειμένου να ποσοτικοποιήσουμε τον βαθμό αμοιβαίας εξάρτησης ενός ζεύγους τυχαίων μεταβλητών ορίζουμε δύο νέα στατιστικά μεγέθη: την συνδιασπορά και τον συντελεστή συσχέτισμού (ή συντελεστή συσχέτισης).

Για κάθε ζεύγος, (x_i, x_j) , των τυχαίων μεταβλητών $\vec{x} \equiv \{x_1, x_2, \dots, x_N\}$, με κοινή συνάρτηση πυκνότητας πιθανότητας την $f(\vec{x}) = f(x_1, x_2, \dots, x_N)$, ορίζουμε ως συνδιασπορά (covariance) την ποσότητα:

$$\text{cov}(x_i, x_j) = E[(x_i - \mu_i) \cdot (x_j - \mu_j)] = \int_{-\infty}^{\infty} \int_{-\infty}^{\infty} (x_i - \mu_i) \cdot (x_j - \mu_j) \cdot f(x_1, x_2, \dots, x_N) dx_1 dx_2 \dots dx_N \quad 1.9.4$$

και ως συντελεστή συσχέτισμού (correlation coefficient) την ποσότητα:

$$\text{corr}(x_i, x_j) = \rho(x_i, x_j) = \frac{\text{cov}(x_i, x_j)}{\sigma_i \cdot \sigma_j} \quad 1.9.5$$

όπου μ_i και μ_j είναι οι αναμενόμενες τιμές ($\mu_i = E(x_i)$, $\mu_j = E(x_j)$) και ενώ σ_i και σ_j είναι οι τετραγωνικές ρίζες της διασποράς ($\sigma_i = \sqrt{V(x_i)}$, $\sigma_j = \sqrt{V(x_j)}$) των τυχαίων μεταβλητών x_i και x_j αντίστοιχα (βλέπε σχέση 1.8.1)

Η σχέση 1.9.4 με απλές πράξεις καταλήγει στην έκφραση:

$$\begin{aligned} \text{cov}(x_i, x_j) &= E[(x_i - \mu_i) \cdot (x_j - \mu_j)] = E[(x_i \cdot x_j + \mu_i \cdot \mu_j - \mu_j \cdot x_i - \mu_i \cdot x_j)] = \\ &= E[x_i \cdot x_j] + \underbrace{E[\mu_i \cdot \mu_j]}_{\mu_i \cdot \mu_j} - \underbrace{E[\mu_j \cdot x_i]}_{\mu_j \cdot E[x_i]} - \underbrace{E[\mu_i \cdot x_j]}_{\mu_i \cdot E[x_j]} = E[x_i \cdot x_j] - \mu_i \cdot \mu_j \end{aligned} \quad 1.9.6$$

Όπως θα δείξετε στις ασκήσεις αυτοαξιολόγησης η απόλυτη τιμή του συντελεστή συσχέτισμού είναι μικρότερη της μονάδας:

$$-1 \leq \rho(x, y) \leq 1 \quad 1.9.7$$

1.10 Αμοιβαία εξαρτημένες και αμοιβαία ανεξάρτητες μεταβλητές

Στην προηγούμενη υποενότητα, με την σχέση 1.9.3, δώσαμε χωρίς απόδειξη την συνθήκη αμοιβαίας ανεξαρτησίας ζεύγους τυχαίων μεταβλητών.

Χωρίς περιορισμό της γενικότητας θα αποδείξουμε την σχέση αμοιβαίας ανεξαρτησίας για την περίπτωση δύο τυχαίων μεταβλητών, X_1 και X_2 , με κοινή συνάρτηση πυκνότητας πιθανότητας $f(x_1, x_2)$.

Θα καλούμε με το σύμβολο $f(x_2|x_1)$ την υπό συνθήκη πυκνότητα πιθανότητας, η τυχαία μεταβλητή X_2 να πάρει την τιμή x_2 ενώ η μεταβλητή X_1 έχει τιμή ίση με x_1 . Θα συμβολίζουμε με $f_2(x_2)$ (αντίστοιχα $f_1(x_1)$) την συνάρτηση πυκνότητας πιθανότητας της μεταβλητής X_2 (αντίστοιχα X_1) να πάρει τιμή ίση με x_2 (αντίστοιχα x_1) ανεξάρτητα με την τιμή που παίρνει η μεταβλητή X_1 (αντίστοιχα X_2). Προφανώς:

$$\begin{aligned} f_2(x_2) &= \int_{-\infty}^{\infty} f(x_1, x_2) dx_1, \quad f_1(x_1) = \int_{-\infty}^{\infty} f(x_1, x_2) dx_2 \\ f(x_1, x_2) &= f(x_1 | x_2) \cdot f_2(x_2) = f(x_2 | x_1) \cdot f_1(x_1) \end{aligned} \quad 1.10.1$$

Συνήθως η συνάρτηση $f_2(x_2)$ (αντίστοιχα $f_1(x_1)$) καλείται **οριακή (marginal) συνάρτηση πυκνότητας πιθανότητας**.

Ερώτηση: Δείξτε ότι $\int_{-\infty}^{\infty} f(x_2) dx_2 = \int_{-\infty}^{\infty} f(x_1) dx_1 = 1$

Υπόδειξη: Χρησιμοποιήσετε την 1.10.1 και το γεγονός ότι εξ ορισμού

$$\int_{-\infty}^{\infty} \int_{-\infty}^{\infty} f(x_1, x_2) dx_2 dx_1 = 1$$

Συνδυασμός των εκφράσεων 1.10.1 καταλήγει στην ακόλουθη σχέση:

$$f_2(x_2) = \int_{-\infty}^{\infty} f(x_2 | x_1) \cdot f_1(x_1) dx_1 \quad 1.10.2$$

Εάν η τυχαία μεταβλητή X_2 δεν εξαρτάται από την μεταβλητή X_1 τότε η πιθανότητα να πάρει την τιμή x_2 (ή η πυκνότητα πιθανότητας της X_2 στην περιοχή τιμών $[x_2, x_2+dx_2]$) είναι η ίδια, ανεξάρτητα της τιμής x_1 της τυχαίας μεταβλητής X_1 . Συνεπώς, η υπό συνθήκη συνάρτηση πυκνότητας πιθανότητας $f(x_2|x_1)$ δεν εξαρτάται από το x_1 , ώστε η 1.10.1 να γράφεται ως:

$$f_2(x_2) = f(x_2 | x_1) \cdot \underbrace{\int_{-\infty}^{\infty} f_1(x_1) dx_1}_1 = f(x_2 | x_1) \quad 1.10.3$$

Αντικαθιστώντας την 1.10.3 στην 1.10.1 καταλήγουμε ότι:

$$f(x_1, x_2) = f(x_2 | x_1) \cdot f_1(x_1) = f_1(x_1) \cdot f_2(x_2) \quad 1.10.4$$

Αξίζει τον κόπο να μας απασχολήσουν τα ακόλουθα δύο θεωρήματα που σχετίζονται με την περίπτωση της αμοιβαίας ανεξαρτησίας δύο τυχαίων μεταβλητών.

Θεώρημα: Οι τυχαίες μεταβλητές X_1 και X_2 , που παίρνουν τιμές στο Ω και έχουν κοινή συνάρτηση πυκνότητας πιθανότητας την $f(x_1, x_2)$, είναι αμοιβαία ανεξάρτητες εάν

$$f(x_1, x_2) = g(x_1) \cdot h(x_2) \quad 1.10.5$$

όπου $g(x_1) \geq 0$ και $h(x_2) \geq 0$ για κάθε $x_1, x_2 \in \Omega$

Απόδειξη:

Από την συνθήκη ανεξαρτησίας, που ορίσαμε με την 1.10.4, είναι σαφές ότι η κοινή συνάρτηση πυκνότητας πιθανότητας δύο αμοιβαίως ανεξαρτήτων τυχαίων μεταβλητών πάντα γράφεται με την μορφή 1.10.5, δεδομένου ότι οι οριακές συναρτήσεις είναι εξ ορισμού μεγαλύτερες ή ίσες από τη μηδέν ($f_1(x_1) \geq 0$ και $f_2(x_2) \geq 0$)

Έστω ότι η κοινή συνάρτηση πυκνότητας πιθανότητας γράφεται με την μορφή 1.10.5. Θα πρέπει να επισημάνουμε ότι οι συναρτήσεις g και h δεν είναι συναρτήσεις πυκνότητας πιθανότητας (π.χ. δεν είναι αναγκαίο το ολοκλήρωμα τους να ισούται με μονάδα) απλώς είναι θετικές συναρτήσεις.

$$\left. \begin{aligned} f_1(x_1) &= \int_{-\infty}^{\infty} g(x_1) \cdot h(x_2) dx_2 = g(x_1) \cdot \underbrace{\int_{-\infty}^{\infty} h(x_2) \cdot dx_2}_a = a \cdot g(x_1) \\ f_2(x_2) &= \int_{-\infty}^{\infty} g(x_1) \cdot h(x_2) dx_1 = h(x_2) \cdot \underbrace{\int_{-\infty}^{\infty} g(x_1) \cdot dx_1}_b = b \cdot h(x_2) \end{aligned} \right\} \Rightarrow \begin{aligned} g(x_1) &= \frac{1}{a} \cdot f_1(x_1) \\ h(x_2) &= \frac{1}{b} \cdot f_2(x_2) \end{aligned} \quad 1.10.6$$

Αντικαθιστώντας την 1.10.6 στην 1.10.5 καταλήγουμε ότι:

$$\left. \begin{aligned} f(x_1, x_2) &= g(x_1) \cdot h(x_2) = \frac{1}{a \cdot b} \cdot f_1(x_1) \cdot f_2(x_2) \\ \int_{-\infty}^{\infty} \int_{-\infty}^{\infty} f(x_1, x_2) dx_1 dx_2 &= 1 \\ \int_{-\infty}^{\infty} f_1(x_1) dx_1 &= 1 \\ \int_{-\infty}^{\infty} f_2(x_2) dx_2 &= 1 \end{aligned} \right\} \begin{aligned} \frac{1}{a \cdot b} &= 1 \\ f(x_1, x_2) &= f_1(x_1) \cdot f_2(x_2) \end{aligned} \quad 1.10.7$$

Επειδή η κοινή συνάρτηση πυκνότητας πιθανότητας γράφεται ως γινόμενο των οριακών συναρτήσεων, f_1 και f_2 , έπεται ότι οι τυχαίες μεταβλητές είναι αμοιβαία ανεξάρτητες.

Θεώρημα: Εάν οι τυχαίες μεταβλητές X_1 και X_2 , που παίρνουν τιμές στο Ω και έχουν κοινή συνάρτηση πυκνότητας πιθανότητας την $f(x_1, x_2)$, είναι αμοιβαία ανεξάρτητες, τότε για τις συναρτήσεις $u(x_1)$ και $v(x_2)$ θα ισχύει ότι:

$$E(u(x_1) \cdot v(x_2)) = E(u(x_1)) \cdot E(v(x_2)) \quad 1.10.8$$

με την προϋπόθεση ότι υπάρχουν οι μέσες τιμές.

Απόδειξη:

Επειδή οι τυχαίες μεταβλητές είναι αμοιβαία ανεξάρτητες, εφαρμόζοντας την σχέση ορισμού της αναμενόμενης τιμής καταλήγουμε ότι:

$$\begin{aligned} E(u(x_1) \cdot v(x_2)) &= \int_{-\infty}^{\infty} \int_{-\infty}^{\infty} u(x_1) \cdot v(x_2) f(x_1, x_2) dx_1 dx_2 = \int_{-\infty}^{\infty} \int_{-\infty}^{\infty} u(x_1) \cdot v(x_2) \cdot f_1(x_1) \cdot f_2(x_2) dx_1 dx_2 = \\ &= \int_{-\infty}^{\infty} u(x_1) \cdot f_1(x_1) dx_1 \int_{-\infty}^{\infty} v(x_2) \cdot f_2(x_2) dx_2 = E(u(x_1)) \cdot E(v(x_2)) \end{aligned} \quad 1.10.9$$

Είναι εύκολο να δει κανείς ότι σ' αυτή την περίπτωση ο συνδυασμός των σχέσεων 1.9.6 και 1.10.9, με $u(x_1) = x_1$ και $v(x_2) = x_2$, καταλήγει ότι:

$$\text{cov}(x_1, x_2) = E(x_1 \cdot x_2) - E(x_1) \cdot E(x_2) = 0 \quad 1.10.10$$

Παράδειγμα 1.10.1 : Ας υποθέσουμε ότι η συνάρτηση πυκνότητας πιθανότητας $f(x)$, της τυχαίας μεταβλητής X είναι άρτια ($f(x) = f(-x)$). Έστω ακόμη η τυχαία μεταβλητή Y , οι τιμές της οποίας είναι: $y=x^2$. Προφανώς οι τυχαίες μεταβλητές X και Y δεν είναι ανεξάρτητες μεταξύ τους, αντίθετα είναι συναρτησιακά εξαρτημένες.

Είναι εύκολο να δείτε ότι η συνάρτηση $x \cdot f(x)$ καθώς και η συνάρτηση $x^3 \cdot f(x)$ είναι περιττές συναρτήσεις ενώ η συνάρτηση $x^2 \cdot f(x)$ είναι άρτια, ώστε:

$$E(x) = \int_{-\infty}^{+\infty} x f(x) dx = 0 \quad 1.10.11$$

$$E(y) = \int_{-\infty}^{+\infty} x^2 f(x) dx = \int_{-\infty}^{+\infty} (0-x)^2 f(x) dx = V(x)$$

Συνεπώς :

$$E(x) E(y) = 0 \quad 1.10.12$$

Επίσης :

$$E(xy) = \int x^3 f(x) dx = 0 \quad 1.10.13$$

Αντικαθιστώντας στην σχέση 1.9.6 τις σχέσεις 1.10.12 και 1.10.13 καταλήγουμε ότι:

$$\text{cov}(x,y)=0$$

Στο παράδειγμα αυτό συναντούμε τυχαίες μεταβλητές με τιμές συναρτησιακά εξαρτημένες οι οποίες όμως στατιστικά έχουν συνδιασπορά ίση με μηδέν. Συνήθως καλούμε αυτές τις μεταβλητές **μη συσχετισμένες**.

1.11 Σύγχρονη μέτρηση της πίεσης και της θερμοκρασίας ιδανικού αερίου

1.12 Γραμμικές συναρτήσεις τυχαίων μεταβλητών

Όπως ήδη αναφέραμε, όλες οι συναρτήσεις τυχαίων μεταβλητών, $y=y(x_1, x_2, \dots, x_N)$, είναι και αυτές τυχαίες μεταβλητές. Συνήθως σε μία μετρητική διαδικασία επιτελούμε ταυτόχρονες μετρήσεις N διαφορετικών μεγεθών, και θέλουμε να υπολογίσουμε κάποιο άλλο μέγεθος το οποίο είναι συνάρτηση των N άμεσα μετρούμενων μεταβλητών.

Παράδειγμα 1.12.1.

Η εκτίμηση του δείκτη νοημοσύνης (IQ) μίας πληθυσμιακής κατηγορίας, σ βασίζεται στην μέτρηση N το πλήθος δεξιοτήτων (οι N δεξιότητες αντιστοιχούν στα $x_1, x_2, \dots, x_N$) κάθε μέλους της ομάδας. Το IQ, κάθε μέλους, υπολογίζεται με κάποιο συναρτησιακό τύπο από τις μετρούμενες δεξιότητες. Είναι διαισθητικά λογικό να υποθέσουμε ότι η αναμενόμενη τιμή του IQ της πληθυσμιακής κατηγορίας, αλλά και η διασπορά του IQ των ατόμων αυτής της πληθυσμιακής κατηγορίας, θα εξαρτάται από την αναμενόμενη τιμή και την διασπορά κάθε μίας από τις δεξιότητες.

Παράδειγμα 1.12.2.

Στην υποενότητα 1.11 δείξαμε ότι η μέτρηση της παγκόσμιας σταθεράς των ιδανικών αερίων, R , ακολουθεί την κατανομή της σχέσης με μέση τιμή που δίνεται από την σχέση..... Η διασπορά της τυχαίας μεταβλητής R δίνεται...

Ας εξετάσουμε κατ' αρχήν, την περίπτωση που αφορά τυχαίες μεταβλητές οι οποίες είναι γραμμικές συναρτήσεις άλλων τυχαίων μεταβλητών, της μορφής:

$$y = y(x_1, x_2, \dots, x_N) = \sum_{i=1}^N a_i \cdot x_i \quad 1.12.1$$

Εν γένει, οι N τυχαίες μεταβλητές, $\{x_1, x_2, \dots, x_N\}$, έχουν κοινή συνάρτηση πυκνότητας πιθανότητας $f(x_1, x_2, \dots, x_N)$ με μέσες τιμές $\{\mu_1, \mu_2, \dots, \mu_N\}$, τετραγωνικές αποκλίσεις $\{V(x_1), V(x_2), \dots, V(x_N)\}$ και συνδιασπορές $\text{cov}(x_i, x_j)$, που δίνονται από τις σχέσεις:

$$\mu_i = E(x_i) = \int_{-\infty}^{\infty} \dots \int_{-\infty}^{\infty} x_i \cdot f(x_1, x_2, \dots, x_N) dx_1 dx_2 \dots dx_N$$

$$V(x_i) = \text{cov}(x_i, x_i) = \int_{-\infty}^{\infty} \dots \int_{-\infty}^{\infty} (\mu_i - x_i)^2 \cdot f(x_1, x_2, \dots, x_N) dx_1 dx_2 \dots dx_N = E(x_i^2) - E^2(x_i) \quad 1.12.2$$

$$\text{cov}(x_i, x_j) = \int_{-\infty}^{\infty} \dots \int_{-\infty}^{\infty} (\mu_i - x_i) \cdot (\mu_j - x_j) \cdot f(x_1, x_2, \dots, x_N) dx_1 dx_2 \dots dx_N = E(x_i \cdot x_j) - E(x_i) \cdot E(x_j)$$

Η αναμενόμενη τιμή της συνάρτησης y, είναι:

$$E(y(x_1, x_2, \dots, x_N)) = E\left(\sum_{i=1}^N a_i x_i\right) = \sum_{i=1}^N a_i \cdot E(x_i) = \sum_{i=1}^N a_i \mu_i \quad 1.12.3$$

Η διασπορά της γραμμικής συνάρτησης, $y = y(x_1, x_2, \dots, x_N) = \sum_{i=1}^N a_i \cdot x_i$,

εκφράζεται συναρτήσει των αποκλίσεων των τυχαίων μεταβλητών $\{x_1, x_2, \dots, x_N\}$ ως⁶:

$$\begin{aligned} V(y(x_1, x_2, \dots, x_N)) &= V\left(\sum_i a_i x_i\right) = \int_{-\infty}^{\infty} \left(E\left(\sum_i a_i x_i\right) - \sum_i a_i x_i \right)^2 \cdot f(x_1, x_2, \dots, x_N) dx_1 dx_2 \dots dx_N = \\ &= \int_{-\infty}^{\infty} \left(\sum_i a_i \mu_i - \sum_i a_i x_i \right)^2 \cdot f(x_1, x_2, \dots, x_N) dx_1 dx_2 \dots dx_N = \\ &= \int_{-\infty}^{\infty} \left[\left(\sum_i a_i \mu_i \right)^2 + \left(\sum_i a_i x_i \right)^2 - 2 \sum_i a_i \mu_i \cdot \sum_i a_i x_i \right] \cdot f(x_1, x_2, \dots, x_N) dx_1 dx_2 \dots dx_N \end{aligned} \quad 1.12.4$$

Επειδή όμως ισχύει ότι:

⁶ Στα επόμενα, για λόγους συντομίας θα χρησιμοποιούμε τα σύμβολα $\sum_i$, $\sum_{i,j}$, και $\sum_{\substack{i,j \\ i>j}}$ για να δηλώσουμε

αθροίσματα του τύπου: $\sum_i \equiv \sum_{i=1}^N$, $\sum_{i,j} \equiv \sum_{i=1}^N \sum_{j=1}^N$, και $\sum_{\substack{i,j \\ i>j}} \equiv \sum_{i=j+1}^N \sum_{j=1}^N$

$$\begin{aligned}
\left(\sum_i a_i \mu_i\right)^2 &= \sum_i a_i^2 \mu_i^2 + 2 \sum_{\substack{i,j \\ i>j}} a_i a_j \mu_i \mu_j \\
\left(\sum_i a_i x_i\right)^2 &= \sum_i a_i^2 x_i^2 + 2 \sum_{\substack{i,j \\ i>j}} a_i a_j x_i x_j \\
\sum_i a_i \mu_i \cdot \sum_i a_i x_i &= \sum_i a_i^2 \mu_i x_i + \sum_{\substack{i,j \\ i>j}} a_i a_j (\mu_i x_j + \mu_j x_i)
\end{aligned}
\tag{1.12.5}$$

η σχέση 1.12.4 καταλήγει ότι

$$\begin{aligned}
(y(x_1, x_2, \dots, x_N)) &= \\
&\dots \int_{-\infty}^{\infty} \left[\sum_i a_i^2 \mu_i^2 + 2 \sum_{\substack{i,j \\ i>j}} a_i a_j \mu_i \mu_j + \sum_i a_i^2 x_i^2 + 2 \sum_{\substack{i,j \\ i>j}} a_i a_j x_i x_j - 2 \sum_i a_i^2 \mu_i x_i - 2 \sum_{\substack{i,j \\ i>j}} a_i a_j (\mu_i x_j + \mu_j x_i) \right] \cdot f(x_1, x_2, \dots, x_N) dx_1 dx_2 \dots dx_N \\
&\int_{-\infty}^{\infty} \dots \int_{-\infty}^{\infty} \left[\sum_i a_i^2 (\mu_i - x_i)^2 \right] + 2 \left[\sum_{\substack{i,j \\ i>j}} a_i a_j (\mu_i \mu_j + x_j x_i - \mu_i x_j - \mu_j x_i) \right] \cdot f(x_1, x_2, \dots, x_N) dx_1 dx_2 \dots dx_N \\
&\tag{1.12.6}
\end{aligned}$$

Εισάγοντας στην 1.12.6 την τετριμμένη ισότητα:

$$\mu_i \mu_j + x_j x_i - \mu_i x_j - \mu_j x_i = (\mu_i - x_i)(\mu_j - x_j) \tag{1.12.7}$$

καταλήγουμε ότι:

$$\begin{aligned}
V(y(x_1, x_2, \dots, x_N)) &= \\
\int_{-\infty}^{\infty} \dots \int_{-\infty}^{\infty} &\left[\sum_i a_i^2 (\mu_i - x_i)^2 \right] + 2 \left[\sum_{\substack{i,j \\ i>j}} a_i a_j (\mu_i \mu_j + x_j x_i - \mu_i x_j - \mu_j x_i) \right] f(x_1, x_2, \dots, x_N) dx_1 dx_2 \dots dx_N = \\
\int_{-\infty}^{\infty} \dots \int_{-\infty}^{\infty} &\left[\sum_i a_i^2 (\mu_i - x_i)^2 \right] f(x_1, x_2, \dots, x_N) dx_1 dx_2 \dots dx_N + \\
&\int_{-\infty}^{\infty} \dots \int_{-\infty}^{\infty} \left[2 \sum_{\substack{i,j \\ i>j}} a_i a_j (\mu_i \mu_j + x_j x_i - \mu_i x_j - \mu_j x_i) \right] f(x_1, x_2, \dots, x_N) dx_1 dx_2 \dots dx_N = \\
\int_{-\infty}^{\infty} \dots \int_{-\infty}^{\infty} &\left[\sum_i a_i^2 (\mu_i - x_i)^2 \right] f(x_1, x_2, \dots, x_N) dx_1 dx_2 \dots dx_N + \\
&\int_{-\infty}^{\infty} \dots \int_{-\infty}^{\infty} \left[2 \sum_{\substack{i,j \\ i>j}} a_i a_j \cdot (\mu_i - x_i)(\mu_j - x_j) \right] f(x_1, x_2, \dots, x_N) dx_1 dx_2 \dots dx_N = \\
\sum_i a_i^2 V(x_i) &+ 2 \sum_{\substack{i,j \\ i>j}} a_i \cdot a_j \cdot \text{cov}(x_i, x_j)
\end{aligned} \tag{1.12.8}$$

Η διασπορά της γραμμικής συνάρτησης, $y = y(x_1, x_2, \dots, x_N) = \sum_{i=1}^N a_i \cdot x_i$, των τυχαίων μεταβλητών $\{x_1, x_2, \dots, x_N\}$, εκφράζεται συναρτήσει των τετραγωνικών αποκλίσεων και της συνδιασποράς (ανά δύο) των τυχαίων μεταβλητών $\{x_1, x_2, \dots, x_N\}$, ως ακολούθως:

$$V\left(\sum_i a_i x_i\right) = \sum_i a_i^2 V(x_i) + 2 \sum_{\substack{i,j \\ i>j}} a_i \cdot a_j \cdot \text{cov}(x_i, x_j) \tag{1.12.9}$$

Η σχέση 1.12.9 μπορεί να γραφεί με τις ισοδύναμες μορφές:

$$\begin{aligned}
V\left(\sum_i a_i x_i\right) &= \sum_i a_i^2 V(x_i) + 2 \sum_{\substack{i,j \\ i>j}} a_i \cdot a_j \cdot \text{cov}(x_i, x_j) = \\
&= \sum_i a_i^2 V(x_i) + 2 \sum_{i=j+1}^N \sum_{j=1}^N a_i \cdot a_j \cdot \text{cov}(x_i, x_j)
\end{aligned} \tag{1.12.10}$$

Είναι απλό να δείξει κανείς ότι εάν οι τυχαίες μεταβλητές, $\{x_1, x_2, \dots, x_N\}$, είναι αμοιβαία ανεξάρτητες τότε η διασπορά της γραμμικής συνάρτησης,

$y = y(x_1, x_2, \dots, x_N) = \sum_{i=1}^N a_i \cdot x_i$, θα ισούται με:

$$V(\sum_i a_i x_i) = \sum_i a_i^2 V(x_i) \quad 1.12.11$$

Παράδειγμα 1.12.3

Θεωρείστε την περίπτωση όπου οι τυχαίες μεταβλητές x_i είναι N ανεξάρτητες δοκιμές (επαναλήψεις) του ιδίου πειράματος. Σ' αυτή την περίπτωση οι N τυχαίες μεταβλητές προέρχονται από την ίδια συνάρτηση πυκνότητας πιθανότητας, $f(x)$. Η κοινή συνάρτηση πυκνότητας πιθανότητας, $g(x_1, x_2, \dots, x_N)$, θα είναι το γινόμενο:

$$g(x_1, x_2, \dots, x_N) = f(x_1) \cdot f(x_2) \cdot \dots \cdot f(x_N) \quad 1.12.12$$

Η αναμενόμενη τιμή του αποτελέσματος κάθε δοκιμής θα ισούται με την μέση τιμή της συνάρτησης πυκνότητας πιθανότητας, $\mu = E(x) = \int_{-\infty}^{\infty} x \cdot f(x) dx$, διότι:

$$\begin{aligned} \mu_i = E(x_i) &= \int_{-\infty}^{\infty} \dots \int_{-\infty}^{\infty} x_i \cdot g(x_1, x_2, \dots, x_N) dx_1 dx_2 \dots dx_N = \\ &= \int_{-\infty}^{\infty} \dots \int_{-\infty}^{\infty} x_i \cdot f(x_1) \cdot f(x_2) \dots f(x_N) dx_1 dx_2 \dots dx_N = \\ &= \int_{-\infty}^{\infty} x_i \cdot f(x_i) dx_i \underbrace{\dots \int_{-\infty}^{\infty} f(x_1) \dots f(x_{i-1}) \cdot f(x_{i+1}) \dots f(x_N) dx_1 \dots dx_{i-1} dx_{i+1} \dots dx_N}_1 = \mu \end{aligned}$$

Ομοίως η τετραγωνική απόκλιση κάθε μέτρησης θα είναι:

$$\begin{aligned} V(x_i) &= \int_{-\infty}^{\infty} \dots \int_{-\infty}^{\infty} (x_i - \mu)^2 \cdot g(x_1, x_2, \dots, x_N) dx_1 dx_2 \dots dx_N = \\ &= \int_{-\infty}^{\infty} \dots \int_{-\infty}^{\infty} (x_i - \mu)^2 \cdot f(x_1) \cdot f(x_2) \dots f(x_N) dx_1 dx_2 \dots dx_N = \\ &= \int_{-\infty}^{\infty} (x_i - \mu)^2 \cdot f(x_i) dx_i \underbrace{\dots \int_{-\infty}^{\infty} f(x_1) \dots f(x_{i-1}) \cdot f(x_{i+1}) \dots f(x_N) dx_1 \dots dx_{i-1} dx_{i+1} \dots dx_N}_1 = \sigma^2 \end{aligned}$$

Στα επόμενα παριστούμε με $\langle x \rangle$, τον μέσο όρο των N παρατηρήσεων x_i

$$\langle x \rangle = \frac{1}{N} \sum_{i=1}^N x_i$$

ο οποίος είναι γραμμική συνάρτηση των N τυχαίων μεταβλητών x_i .

Προφανώς ο μέσος όρος είναι και αυτός τυχαία μεταβλητή με αναμενόμενη μέση τιμή που δίνεται από:

$$E(\langle x \rangle) = \frac{1}{N} E\left(\sum_{i=1}^N x_i\right) = \frac{1}{N} N\mu = \mu$$

Δεν χρειάζεται να ξέρουμε την ακριβή συναρτησιακή μορφή της συνάρτησης πυκνότητας πιθανότητας του μέσου όρου (για μεγάλο N είναι πολύ εύκολο να την μαντέψουμε, όπως θα δούμε στα επόμενα), για να υπολογίσουμε την διασπορά, $V(\langle x \rangle)$.

Πράγματι, εφαρμόζοντας την σχέση 1.12.10 καταλήγουμε ότι:

$$V\left(\frac{1}{N} \sum_i x_i\right) = \frac{1}{N^2} \sum V(x_i) + \frac{2}{N^2} \sum_{i=j+1}^N \sum_{i=1}^N \text{cov}(x_i, x_j) = \frac{\sigma^2}{N} + \frac{2}{N^2} \sum_{i=j+1}^N \sum_{i=1}^N \text{cov}(x_i, x_j)$$

Εάν οι μετρήσεις είναι ανεξάρτητες μεταξύ τους, όπως συμβαίνει στις περισσότερες των περιπτώσεων, τότε $\text{cov}(x_i, x_j) = 0$, και η προηγούμενη σχέση καταλήγει στην απλή έκφραση:

$$V(\langle x \rangle) = \frac{\sigma^2}{N} \quad 1.12.12$$

Η διασπορά του μέσου όρου N ανεξαρτήτων μετρήσεων του ιδίου φυσικού μεγέθους είναι N φορές μικρότερη από την διασπορά της μίας μέτρησης.

Θα εκτιμήσουμε περισσότερο το αποτέλεσμα 1.12.12 όταν θα συνδέσουμε την έννοια του μετρητικού σφάλματος με την διασπορά των τυχαίων μεταβλητών.

1.13 Συναρτήσεις τυχαίων μεταβλητών

Οι σχέσεις που καταλήξαμε στην προηγούμενη υποενότητα ισχύουν για γραμμικές συναρτήσεις. Στα ακόλουθα θα δώσουμε ένα πολύ χρήσιμο, πάντως προσεγγιστικό,

μαθηματικό τύπο για την διασπορά οποιασδήποτε συνάρτησης N τυχαίων μεταβλητών.

Ας υποθέσουμε ότι έχουμε ένα σύνολο N τυχαίων μεταβλητών, $X_1, X_2, \dots, X_N$, και έστω η τυχαία μεταβλητή Y η τιμή της οποίας είναι συνάρτηση των τιμών $\vec{x} = \{x_1, x_2, \dots, x_N\}$, των άλλων τυχαίων μεταβλητών: $y = f(\vec{x})$.

Ας υποθεθεί επιπλέον ότι η κοινή συνάρτηση πυκνότητας πιθανότητας των μεταβλητών, $X_1, X_2, \dots, X_N$, είναι η $g(\vec{x})$ και ότι οι αναμενόμενες τιμές, οι διασπορές και οι συνδιασπορές των τυχαίων μεταβλητών είναι αντίστοιχα ίσες με $\{\mu_1, \mu_2, \dots, \mu_N\}$ (όπου $\mu_i = \int x_i g(\vec{x}) d\vec{x}^N$), $\{\sigma_1^2, \sigma_2^2, \sigma_3^2, \dots, \sigma_N^2\}$ (όπου $V(x_i) = \sigma_i^2 = \int (x_i - \mu_i)^2 g(\vec{x}) d\vec{x}^N$), $\text{cov}(x_i, x_j)$ (όπου $\text{cov}(x_i, x_j) = \int (x_i - \mu_i) \cdot (x_j - \mu_j) g(\vec{x}) d\vec{x}^N$)

Επί πλέον υποθέτουμε ότι η συνάρτηση $f(\vec{x}) = f(x_1, x_2, \dots, x_N)$ μπορεί να αναπτυχθεί γύρω από τα $\{\mu_1, \mu_2, \dots, \mu_N\}$ σε σειρά Taylor, δηλαδή:

$$f(\vec{x}) = f(\mu_1, \mu_2, \dots, \mu_N) + \sum_i^N (x_i - \mu_i) \frac{\partial f}{\partial \mu_i} + \sum_{i=1}^N \frac{(x_i - \mu_i)^2}{2} \frac{\partial^2 f}{\partial \mu_i^2} + \dots \quad 1.13.1$$

όπου χρησιμοποιούμε τον συμβολισμό $\frac{\partial f}{\partial \mu_i} \equiv \frac{\partial f}{\partial x_i} \Big|_{x_i = \mu_i}$

Εάν υπάρχει μικρή πιθανότητα οι τυχαίες μεταβλητές $X_1, X_2, \dots, X_N$, να πάρουν τιμές πολύ μεγαλύτερες ή μικρότερες από τις αναμενόμενες τιμές $\{\mu_1, \mu_2, \dots, \mu_N\}$ τότε οι όροι, $(x_i - \mu_i)^k$ με $k > 1$, της 1.13.1 δεν συμμετέχουν σημαντικά στο τελικό αποτέλεσμα (άθροισμα). Στην περίπτωση αυτή⁷ μπορούμε να αγνοήσουμε όρους μεγαλύτερου από πρώτου βαθμού και να γράψουμε:

$$f(\vec{x}) \cong f(\mu_1, \mu_2, \dots, \mu_N) + \sum_i^N (x_i - \mu_i) \frac{\partial f}{\partial \mu_i} \quad 1.13.2$$

Υπολογίζοντας τις αναμενόμενες τιμές του αριστερού και δεξιού μέλους της 1.13.2, καταλήγουμε ότι:

⁷ Όταν δηλαδή η διασπορά των τυχαίων μεταβλητών είναι μικρή

$$\begin{aligned}
E(f(\vec{x})) &\equiv E(f(\mu_1, \mu_2, \dots, \mu_N)) + \int_{\Omega} \left(\sum_i^N (x_i - \mu_i) \frac{\partial f}{\partial \mu_i} \right) g(\vec{x}) d\vec{x}^N \\
&= E(f(\mu_1, \mu_2, \dots, \mu_N)) + \sum_i^N \frac{\partial f}{\partial \mu_i} \int_{\Omega} (x_i - \mu_i) g(\vec{x}) d\vec{x}^N \\
&= E(f(\mu_1, \mu_2, \dots, \mu_N)) + \sum_i^N \frac{\partial f}{\partial \mu_i} \cdot \underbrace{\left[\int_{\Omega} (x_i) g(\vec{x}) d\vec{x}^N - \int_{\Omega} (\mu_i) g(\vec{x}) d\vec{x}^N \right]}_0 \\
&= E(f(\mu_1, \mu_2, \dots, \mu_N)) = f(\mu_1, \mu_2, \dots, \mu_N)
\end{aligned}$$

Η αναμενόμενη τιμή της συνάρτησης των τυχαίων μεταβλητών $X_1, X_2, \dots, X_N$, $y = f(\vec{x})$, $\vec{x} = \{x_1, x_2, \dots, x_N\}$ δίνεται από την σχέση:

$$E(f(\vec{x})) \equiv f(\mu_1, \mu_2, \dots, \mu_N) \quad 1.13.3$$

όπου $\{\mu_1, \mu_2, \dots, \mu_N\}$ είναι οι αναμενόμενες τιμές των τυχαίων μεταβλητών.

Επίσης από τις σχέσεις 1.13.2 και 1.13.3 καταλήγουμε ότι:

$$V(f(x_1, x_2, \dots, x_N)) = E[(f - E(f))^2] \equiv E \left[\left[\sum_{i=1}^N (x_i - \mu_i) \frac{\partial f}{\partial \mu_i} \right]^2 \right] \quad 1.13.4$$

Αναπτύσσοντας το τετράγωνο του αθροίσματος στο δεξιό μέλος της 1.13.4, βρίσκουμε ότι:

$$\begin{aligned}
&E \left[\sum_{i=1}^N (x_i - \mu_i)^2 \frac{\partial f^2}{\partial \mu_i} + 2 \sum_{i=j+1}^N \sum_{j=1}^N (x_i - \mu_i)(x_j - \mu_j) \frac{\partial f}{\partial \mu_i} \cdot \frac{\partial f}{\partial \mu_j} \right] = \\
&\sum_{i=1}^N \frac{\partial f^2}{\partial \mu_i} \int_{\Omega} (x_i - \mu_i)^2 \cdot g(\vec{x}) d\vec{x}^N + 2 \sum_{i=j+1}^N \sum_{j=1}^N \frac{\partial f}{\partial \mu_i} \cdot \frac{\partial f}{\partial \mu_j} \int_{\Omega} (x_i - \mu_i)(x_j - \mu_j) \cdot g(\vec{x}) d\vec{x}^N = \quad 1.13.5 \\
&\sum_{i=1}^N V(x_i) \left(\frac{\partial f}{\partial \mu_i} \right)^2 + 2 \sum_{i=j+1}^N \sum_{j=1}^N \frac{\partial f}{\partial \mu_i} \cdot \frac{\partial f}{\partial \mu_j} \text{cov}(x_i, x_j)
\end{aligned}$$

Η διασπορά της συνάρτησης των τυχαίων μεταβλητών $X_1, X_2, \dots, X_N$, $y = f(\vec{x})$, $\vec{x} = \{x_1, x_2, \dots, x_N\}$ δίνεται από την σχέση

$$V(f(x_1, x_2, \dots, x_N)) \equiv \sum_{i=1}^N V(x_i) \left(\frac{\partial f}{\partial \mu_i} \right)^2 + 2 \sum_{i=j+1}^N \sum_{j=1}^N \frac{\partial f}{\partial \mu_i} \cdot \frac{\partial f}{\partial \mu_j} \text{cov}(x_i, x_j) \quad 1.13.6$$

Ερώτηση: Δείξτε ότι η σχέση 1.13.6 γράφεται ως:

$$V(f(x_1, x_2, \dots, x_N)) \cong \sum_{i=1}^N \sum_{j=1}^N \frac{\partial f}{\partial \mu_i} \cdot \frac{\partial f}{\partial \mu_j} \text{cov}(x_i, x_j) \quad 1.13.7$$

Ερώτηση: Δείξτε ότι εάν η συνάρτηση $y = f(\bar{x})$, $\bar{x} = \{x_1, x_2, \dots, x_N\}$ είναι γραμμική συνάρτηση των τυχαίων μεταβλητών τότε η σχέση 1.13.6 συμπίπτει με την 1.12.9

Ερώτηση: Γράψετε την σχέση 1.13.6 για την περίπτωση που οι τυχαίες μεταβλητές $X_1, X_2, \dots, X_N$, είναι αμοιβαία ανεξάρτητες.

Κατά τον ίδιο τρόπο, εάν $t = t(\bar{x})$ και $r = r(\bar{x})$ (με $\bar{x} = \{x_1, x_2, \dots, x_N\}$), είναι δύο συναρτήσεις των N τυχαίων μεταβλητών, $X_1, X_2, \dots, X_N$, αποδεικνύεται ότι η συνδιασπορά τους προσεγγίζεται με τη σχέση:

$$\text{cov}(t, r) = \sum_{i=1}^N \sum_{j=1}^N \frac{\partial t}{\partial \mu_i} \cdot \frac{\partial r}{\partial \mu_j} \text{cov}(x_i, x_j) \quad 1.13.8$$

Ερώτηση: Αποδείξτε την 1.13.8. Τι συμβαίνει αν οι μεταβλητές είναι ανεξάρτητες;

Υπόδειξη:

$$\text{cov}(t, r) = E\left[\left((t(\mu_1, \mu_2, \dots, \mu_N) - t(x_1, x_2, \dots, x_N)) \cdot ((r(\mu_1, \mu_2, \dots, \mu_N) - r(x_1, x_2, \dots, x_N)))\right)\right]$$

Αντικαταστήστε με την προσέγγιση 1.13.2 και εκτελέστε τις πράξεις

Παράδειγμα 1.13.1:

Στην υποενότητα

Πρόβλημα 1.6.1: Δείξτε ότι το σ είναι πάντα θετική. Αποδείξτε ότι $-1 \leq \rho(x, y) \leq 1$

Δείξτε ότι η μέση τιμή και το σ μίας μεταβλητής scale up ενώ ο συντελεστής συσχετισμού παραμένει αμετάβλητος

1.14 Ροπές Μεγαλύτερης Τάξης.

Έστω η τυχαία μεταβλητή X , με συνάρτηση πυκνότητας πιθανότητας $f(x)$. Η αναμενόμενη τιμή της νιοστής x , x^n , καλείται αλγεβρική ροπή τάξης n . Γενικότερα:

$$\begin{aligned}\mu'_n &= E(x^n) && n \text{ τάξης αλγεβρική ροπή} \\ \mu_n &= E([x-E(x)]^n) && n \text{ τάξης κεντρική ροπή} \\ \nu'_n &= E(|x|^n) && n \text{ τάξης απόλυτη ροπή} \\ \nu_n &= E(|x-E(x)|^n) && n \text{ τάξης κεντρική απόλυτη ροπή}\end{aligned}\quad 1.14.1$$

Οι ορισμοί 1.14.1 γενικεύονται εύκολα και στην περίπτωση πολλών μεταβλητών. Σαν παράδειγμα, η αλγεβρική ροπή n τάξης ως προς x και m τάξης ως προς y , είναι:

$$\mu'_{nm} = E(x^n \cdot y^m)$$

Καθώς όλες οι συμμετρικές, ως προς την αναμενόμενη τιμή, συναρτήσεις πυκνότητας πιθανότητας έχουν όλες τις περιττές κεντρικές ροπές ίσες με το μηδέν, μπορούμε να χρησιμοποιήσουμε τις κεντρικές ορμές για να εκφράσουμε ποσοτικά την ασυμμετρία της συνάρτησης πυκνότητας πιθανότητας.

Είναι εύκολο να δείξουμε ότι η κεντρική ροπή πρώτης τάξης ισούται με το μηδέν, για οποιαδήποτε συνάρτηση πυκνότητας πιθανότητας.

Ερώτηση: Δείξτε ότι $\mu_1=0$

Υπόδειξη: $E([x-E(x)]^1)=E(x)-E(E(x))=0$

Συνεπώς η τρίτη κεντρική ροπή είναι η χαμηλότερης τάξης κεντρική ροπή που μας δίνει πληροφορία για αυτή την ασυμμετρία. Συνήθως, προκειμένου να προκύψει αδιάστατος αριθμός, διαιρούμε την τρίτη κεντρική ροπή με το σ^3 ($[V(x)]^{3/2}$) και ορίζουμε την ασυμμετρία (skewness) ως:

$$\gamma_1 = E([x-E(x)]^3)/\sigma^3 \quad 1.14.2$$

Η τέταρτη κεντρική ροπή χρησιμοποιείται συνήθως για να ποσοτικοποιήσει την αιχμηρότητα (sharpness) της συνάρτησης πυκνότητας πιθανότητας. Συνήθως ονομάζεται κύρτωση (cyrtosis ή kurtosis) και ορίζεται ως:

$$\gamma_2 = E([x-E(x)]^4)/\sigma^4 - 3 \quad 1.14.3$$

όπου αφαιρείται το τρία ώστε η κύρτωση να είναι μηδέν στην περίπτωση των Gaussian συναρτήσεων πυκνότητας πιθανότητας.

Άλλες φορές, οι ορμές κανονικοποιούνται ως προς δυνάμεις της αναμενόμενης τιμής.

Μπορεί ναδειχθεί ότι εάν όλες οι ορμές είναι γνωστές τότε η συνάρτηση πυκνότητας πιθανότητας μπορεί να ορισθεί πλήρως από αυτές.

1.15 Χαρακτηριστική Συνάρτηση.

Όπως έγινε σαφές από τα προηγούμενα, ενδιαφερόμαστε για τυχαίες μεταβλητές που παίρνουν πραγματικές τιμές. Εν τούτοις μπορούμε εύκολα να ορίσουμε τυχαίες μεταβλητές που παίρνουν μιγαδικές τιμές. Πράγματα εάν X και Y είναι τυχαίες

μεταβλητές με πραγματικές τιμές, x και y αντίστοιχα, τότε η μεταβλητή $Z=X+iY$ παίρνει μιγαδικές τιμές, $x+iy$. Σαν παράδειγμα, η τυχαία μεταβλητή Z έχει αναμενόμενη τιμή: $E(Z)=E(X)+iE(Y)$

Όπως θα δούμε στα επόμενα, η γενίκευση αυτή θα μας οπλίσει με ένα χρήσιμο εργαλείο για την μελέτη των συναρτήσεων πυκνότητας πιθανότητας.

Ο μετασχηματισμός Fourier της συνάρτησης πυκνότητας πιθανότητας, $f(x)$, καλείται χαρακτηριστική συνάρτηση:

$$\varphi_X(t) = \int_{-\infty}^{\infty} e^{itx} \cdot f(x) dx = E(e^{itx}) \quad \text{για } t \text{ πραγματικό} \quad 1.15.1$$

Στην περίπτωση που η μεταβλητή παίρνει διακριτές τιμές, η χαρακτηριστική συνάρτηση ορίζεται ως:

$$\varphi_X(t) = \sum_k f(x_k) \cdot e^{itx_k} = E(e^{itx}) \quad \text{για } t \text{ πραγματικό} \quad 1.15.2$$

Προφανώς η χαρακτηριστική συνάρτηση είναι μιγαδική συνάρτηση.

Εάν η $f(x)$ είναι συνεχής παντού στο πεδίο ορισμού της, τότε η χαρακτηριστική συνάρτηση ορίζει την συνάρτηση πυκνότητας πιθανότητας της τυχαίας μεταβλητής, ως:

$$f(x) = \frac{1}{2\pi} \int_{-\infty}^{\infty} \varphi_X(t) \cdot e^{-ixt} dt \quad 1.15.3$$

Στα επόμενα παραθέτουμε μερικές από τις βασικές ιδιότητες της χαρακτηριστικής συνάρτησης:

- Η χαρακτηριστική συνάρτηση, $\varphi_X(t)$, υπάρχει για όλα τα t και επίσης ισχύει ότι: $\varphi(0)=1$, $|\varphi(t)| \leq 1$.
- Εάν $\varphi_X(t)$ είναι η χαρακτηριστική συνάρτηση της τυχαίας μεταβλητής x , τότε η χαρακτηριστική συνάρτηση της $y=ax+b$ (όπου a και b είναι σταθερές) είναι :

$$\varphi_Y(t) = e^{ibt} \cdot \varphi_X(at) \quad 1.15.4$$

Ερώτηση: Αποδείξτε την σχέση 1.15.4

Υπόδειξη: $\varphi_Y(t) = E(e^{ibt} \cdot e^{ia \cdot xt}) = e^{ibt} \cdot E(e^{ia \cdot xt})$

- Εάν X και Y είναι αμοιβαία ανεξάρτητες τυχαίες μεταβλητές, με χαρακτηριστικές συναρτήσεις, $\varphi_X(t)$ και $\varphi_Y(t)$, η χαρακτηριστική συνάρτηση της μεταβλητής $Z=X+Y$ είναι:

$$\varphi_{X+Y}(t) = \varphi_X(t) \cdot \varphi_Y(t) \quad 1.15.5$$

Ερώτηση: Αποδείξτε την σχέση 1.15.5

Υπόδειξη: Επειδή X και Y είναι αμοιβαία ανεξάρτητες τότε η κοινή συνάρτηση πυκνότητας πιθανότητας θα παραγοντοποιείται ως εξής: $f(x,y)=f_1(x) \cdot f_2(y)$. Συνεπώς:

$$\varphi_{X+Y}(t) = E(e^{i(x+Y)t}) = E(e^{ixt}) \cdot E(e^{iyt})$$

Παράδειγμα 1.15.1

Έστω οι αμοιβαία ανεξάρτητες τυχαίες μεταβλητές $X_1, X_2, X_3, \dots, X_N$ με κοινή συνάρτηση πυκνότητας πιθανότητας, $f(x_1, x_2, x_3, \dots, x_N) = f_1(x_1) \cdot f_2(x_2) \cdot f_3(x_3) \dots f_N(x_N)$, όπου $f_1(x_1), f_2(x_2), f_3(x_3), \dots, f_N(x_N)$ είναι οι συναρτήσεις πυκνότητας κάθε μίας από τις τυχαίες μεταβλητές και $\varphi_1(t), \varphi_2(t), \varphi_3(t), \dots, \varphi_N(t)$ είναι οι αντίστοιχες χαρακτηριστικές συναρτήσεις. Έστω ακόμα η τυχαία μεταβλητή Z , ώστε $Z = X_1 + X_2 + X_3 + \dots + X_N$

Προκειμένου να υπολογίσουμε την συνάρτηση πυκνότητας πιθανότητας της Z , $g(z)$, θα πρέπει να προβούμε στις ακόλουθες ολοκληρώσεις:

$$g(z) = \int_{-\infty}^{\infty} \int_{-\infty}^{\infty} \dots \int_{-\infty}^{\infty} \delta(z - x_1 - x_2 - \dots - x_N) \cdot f_1(x_1) \cdot f_2(x_2) \dots f_N(x_N) dx_1 dx_2 \dots dx_N$$

Εν γένει το ολοκλήρωμα αυτό είναι πολύπλοκο. Πολύ πιο εύκολα υπολογίζεται η χαρακτηριστική συνάρτηση της Z , $\Phi(t)$, εφαρμόζοντας την σχέση 1.15.5. Πράγματι για την περίπτωση N μεταβλητών η 1.15.5 γράφεται ως:

$$\Phi(t) = \varphi_1(t) \cdot \varphi_2(t) \cdot \varphi_3(t) \dots \varphi_N(t)$$

Στη συνέχεια η συνάρτηση πυκνότητας πιθανότητας, $g(z)$, υπολογίζεται ως ο αντίστροφος μετασχηματισμός Fourier, 1.15.3, της $\Phi(t)$.

- Η χαρακτηριστική συνάρτηση συνδέεται με τις αλγεβρικές ροπές μιας κατανομής ως εξής:

$$\varphi_X(t) = \sum_{r=0}^{\infty} \frac{(it)^r}{r!} \cdot \mu'_r \quad 1.15.6$$

Ερώτηση: Αποδείξτε τη σχέση 1.15.6

Υπόδειξη: Αναλύοντας την μιγαδική συνάρτηση e^{itx} κατά Taylor καταλήγουμε ότι:

$$\varphi_X(t) = \sum_{r=0}^{\infty} \frac{(itx)^r}{r!} = \sum_{r=0}^{\infty} \frac{(it)^r}{r!} \cdot E(x^r)$$

Από την σχέση 1.15.6, εύκολα, συνάγουμε ότι εάν ξέρουμε με κάποιο τρόπο όλες τις αλγεβρικές ροπές μπορούμε να ορίσουμε την χαρακτηριστική συνάρτηση $\varphi_X(t)$ και στην συνέχεια να ορίσουμε πλήρως την συνάρτηση πυκνότητας πιθανότητας με την σχέση 1.15.3.

Ερώτηση: Δείξτε ότι $\mu'_r = \frac{1}{(i)^r} \cdot \frac{d^r}{dt^r} \varphi(t) \Big|_{t=0}$ 1.15.7

Υπόδειξη: Υπολογίσετε την r τάξη παράγωγο και στα δύο μέλη της 1.15.6, λύσετε ως προς μ'_r , και θέσετε $t=0$

Ερώτηση: Δείξτε ότι οι κεντρικές ροπές εκφράζονται ως εξής:

$$\mu_r = \frac{1}{(i)^r} \cdot \frac{d^r}{dt^r} e^{-\mu t} \cdot \varphi(t) \Big|_{t=0} \quad 1.15.8$$

όπου μ είναι η αναμενόμενη τιμή της τυχαίας μεταβλητής ($\mu = E(x)$).

Υπόδειξη:

$$e^{-i\mu t} \cdot \varphi_X(t) = E(e^{it(\mu - X)}) = \sum_{r=0}^{\infty} \frac{(it(\mu - X))^r}{r!} = \sum_{r=0}^{\infty} \frac{(it)^r}{r!} \cdot E((\mu - X)^r) = \sum_{r=0}^{\infty} \frac{(it)^r}{r!} \cdot \mu_r$$