ΣΤΟΧΑΣΤΙΚΕΣ ΔΙΑΔΙΚΑΣΙΕΣ

Στοχαστικές διαδικασίες

Γραμμικά συστήματα

Αλυσίδες Markov

Θεωρία πληροφοριών

Γιάννης Α. Φίλης

Πολυτεχνείο Κρήτης - Σεπτέμβριος 2006

ΠΕΡΙΕΧΟΜΕΝΑ

I ΟΡΙΣΜΟΣ ΚΑΙ ΣΥΝΑΡΤΗΣΕΙΣ ΣΤΟΧΑΣΤΙΚΩΝ ΔΙΑΔΙΚΑΣΙΩΝ

1.1 Εισαγωγή - Ορισμός

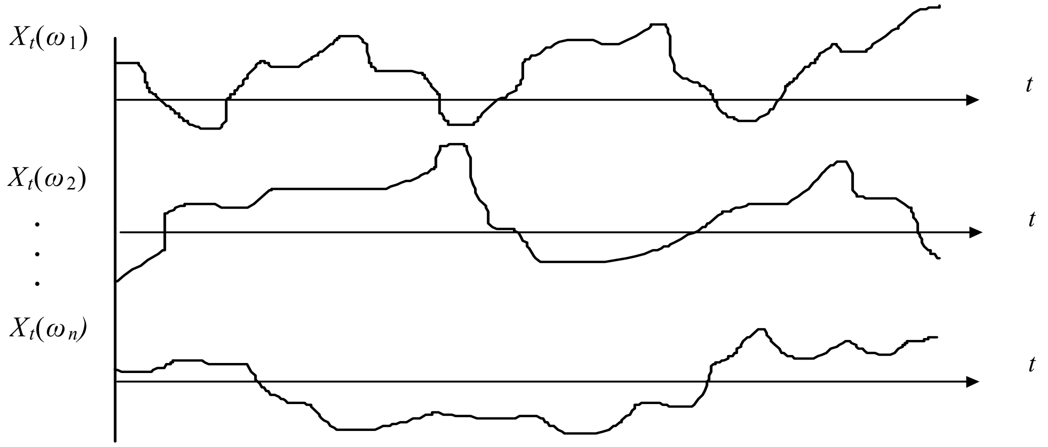
Από τη θεωρία πιθανοτήτων ξέρουμε ότι μια τυχαία μεταβλητή X είναι μια συνάρτηση που καθορίζει έναν αριθμό $X(\omega)$, σε κάθε εξαγόμενο ω ενός τυχαίου πειράματος, που ορίζεται σε χώρο πιθανοτήτων $(\Omega,\Im,P)$.

Μια **στοχαστική διαδικασία (stochastic process)** $\{X_t(\omega),\ t\in T\}$ είναι μια οικογένεια τυχαίων μεταβλητών ορισμένων σε κοινό χώρο πιθανοτήτων $(\Omega,\Im,P)$ με παράμετρο την πραγματική μεταβλητή t (χρόνος). Έτσι σε κάθε εξαγόμενο ω του τυχαίου πειράματος ορίζουμε μια συνάρτηση $X_t(\omega)$. Προφανώς $\omega\in\Omega$.

Αν το σύνολο T είναι ο άξονας των πραγματικών τότε η διαδικασία λέγεται διαδικασία **συνεχούς χρόνου**. Αν το T είναι σύνολο ακεραίων τότε η διαδικασία λέγεται **διακεκριμένου χρόνου**.

Επί πλέον η διαδικασία $X_t(\omega)$ λέγεται **διακεκριμένης κατάστασης** αν οι τιμές της είναι μετρητέες (αριθμήσιμες). Άλλως λέγεται **συνεχούς κατάστασης**.

Παράδειγμα 1.1

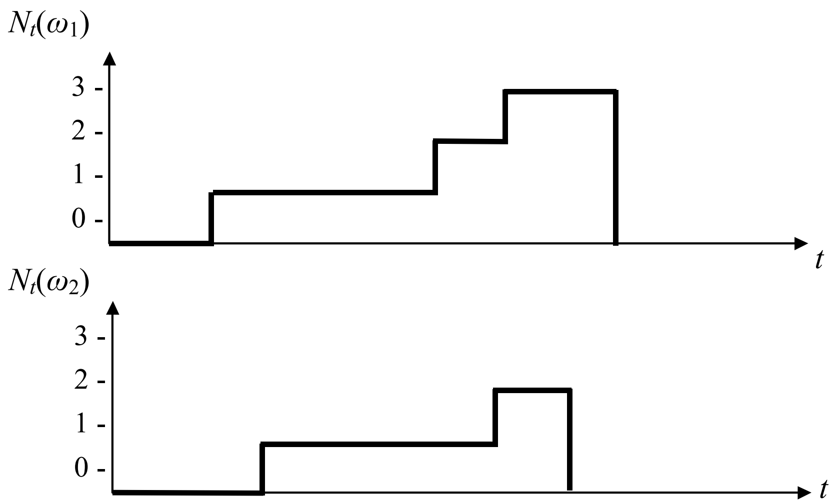


Επομένως η στοχαστική διαδικασία συνίσταται από μια **οικογένεια συναρτήσεων** $X_t(\omega)$. Για δεδομένο ω η $X_t = X_t(\omega)$ είναι συνάρτηση του χρόνου, ενώ για δεδομένο χρόνο t η $X(\omega) = X_t(\omega)$ είναι μια τυχαία μεταβλητή. Συνήθως παραλείπουμε το ω και γράφουμε X_t ή $X(t)$.

Η θεωρία τυχαίων μεταβλητών δεν παρέχει τα μέσα για την εξέταση φαινομένων που είναι τυχαία και εξελίσσονται στο χρόνο, όπως συμβαίνει σε πολλά φυσικά συστήματα. Αυτό είναι το κίνητρο για την ανάπτυξη της θεωρίας των στοχαστικών διαδικασιών.

Παράδειγμα 1.2

Θεωρήστε μια μηχανή που επεξεργάζεται κάποιο κομμάτι και έχει δεδομένη πιθανότητα να πάθει βλάβη και να τεθεί υπό επισκευή. Τότε, αν $N_t(\omega)$ είναι η συνολική παραγωγή από τη στιγμή της επισκευής μέχρι τη στιγμή t έως ότου πάθει βλάβη η μηχανή, θα έχουμε το ακόλουθο σχήμα:



Παράδειγμα 1.3

Η κίνηση ενός μορίου κάποιου ρευστού είναι στοχαστική διαδικασία. Συγκεκριμένα οι τρεις συνιστώσες (x, y, z) καθώς και οι συνιστώσες ταχύτητας $(\dot{x}, \dot{y}, \dot{z})$ και επιτάχυνσης $(\ddot{x}, \ddot{y}, \ddot{z})$, ακολουθούν τυχαίες τροχιές σαν αυτές του παραδείγματος 1.1. Όλα τα σωματίδια συνιστούν τη διαδικασία $X_t(\omega)$. Μια συγκεκριμένη πραγματοποίηση $X_t(\omega_i)$ της $X_t(\omega)$ αντιστοιχεί στην κίνηση του σωματιδίου i. Αυτή η στοχαστική διαδικασία είναι η διαδικασία Brown.

1.2 Στατιστική Στοχαστικών Διαδικασιών

Θεωρήστε το γεγονός $\{X_t \leq x\}$ το οποίο ορίζεται στο χώρο $(\Omega,\Im,P)$, δηλαδή $\{X_t \leq x\} \in \Im$. Αυτό το γεγονός συνίσταται από όλα τα εξαγόμενα ω_i, τέτοια ώστε η διαδικασία $X_t(\omega_i)$ στο δεδομένο χρόνο t είναι μικρότερη ή ίση του x.

Μια στοχαστική διαδικασία βλέπουμε ότι μπορεί να θεωρηθεί σαν μια απειρία τυχαίων μεταβλητών (μία για κάθε t).

Η **κατανομή πρώτης τάξης** ορίζεται σαν τη συνάρτηση

$$F(x,t) \equiv P\{X_t \leq x\}$$

και, εάν η παράγωγος υπάρχει, η συνάρτηση

$$f(x,t) = \frac{\partial F(x,t)}{\partial x}$$

είναι η **πυκνότητα πρώτης τάξης**.

Οι αντίστοιχες ποσότητες δεύτερης τάξης είναι

$$F(x_1,x_2;t_1,t_2) = P\{X_{t_1} \leq x_1, X_{t_2} \leq x_2\}$$

$$f(x_1,x_2;t_1,t_2) = \frac{\partial^2 F(x_1,x_2;t_1,t_2)}{\partial x_1 \partial x_2}.$$

Η γενίκευση για οποιαδήποτε τάξη είναι προφανής. Παρατηρούμε ότι

$$F(x_1,t_1) = P\{X_{t_1} \leq x_1\} = P\{X_{t_1} \leq x_1, X_{t_2} \leq \infty\} = F(x_1,\infty;t_1,t_2).$$

Επίσης

$$f(x_1,t_1) = \int_{-\infty}^{\infty} f(x_1,x_2\,;\,t_1,t_2)dx_2\,.$$

Για την πλήρη γνώση της στατιστικής συμπεριφοράς μιας στοχαστικής διαδικασίας χρειαζόμαστε τη συνάρτηση

$$F(x_1,\ldots,x_n\,;\,t_1,\ldots,t_n),\ \forall\ t_i,\,x_i,\ i = 1,\,\ldots,\,n.$$

Συχνά όμως οι απαιτήσεις μας είναι λιγότερες και περιοριζόμαστε στις δύο πρώτες ροπές της διαδικασίας.

Πρώτη ροπή. Ορίζουμε **προσδοκητή** ή **μέση τιμή** $\overline{X}_t = E(X_t)$ της διαδικασίας X_t το ακόλουθο μέγεθος

$$\overline{X}_t = \int_{-\infty}^{\infty} x\,f(x,t)dx$$

Εν γένει, για κάθε διαδικασία X_t και συνάρτηση g τέτοια ώστε το γεγονός $\{g(X_t) \le z\}$ να ορίζεται στο χώρο (Ω,F,P) για κάθε $z\in(-\infty,\,\infty)$, η $g(X_t)$ είναι επίσης στοχαστική διαδικασία με $E[g(X_t)] = \int_{-\infty}^{\infty} g(x)f(x,t)dx$.

Δεύτερες ροπές. Ορίζουμε **αυτοσυσχέτιση** $R(t_1,\,t_2)$ της διαδικασίας X_t το μέγεθος

$$R(t_1,\,t_2) \triangleq E(X_{t_1}X_{t_2}) = \int_{-\infty}^{\infty}\int_{-\infty}^{\infty} x_1x_2\,f(x_1,x_2\,;\,t_1,t_2)dx_1dx_2\,.$$

Για $t_1 = t_2 = t$ έχουμε $R(t,\,t) = E(X_t^2)$. Η **αυτοσυμμεταβλητότητα** της διαδικασίας X_t κατά τις χρονικές στιγμές t_1 και t_2 είναι η **συμμεταβλητότητα** των X_{t_1} και X_{t_2}:

$$C(t_1,\,t_2) \triangleq E[(X_{t_1} - \overline{X}_{t_1})\,(X_{t_2} - \overline{X}_{t_2})] = \ldots = R(t_1,\,t_2) - \overline{X}_{t_1}\,\overline{X}_{t_2}$$

Για $t_1 = t_2 = t$ ευρίσκουμε $C(t,\,t) = E(X_t^2) - \overline{X}_t^2 = Var(X_t)$ που είναι η μεταβλητότητα ή **διασπορά** (variance) και ισούται με το τετράγωνο της τυπικής απόκλισης.

Παράδειγμα 1.4

Έστω ότι έχουμε μια γεννήτρια σημάτων (ή συναρτήσεων) που παράγουν ημιτονοειδή σήματα αλλά τα εύρη τους δεν είναι σταθερά. Τότε το σήμα

$$X_t = a\,cos 2\pi t + b\,sin 2\pi t$$

(όπου τα a και b είναι τυχαίες μεταβλητές που ακολουθούν κανονική κατανομή με μέση τιμή 0 και διασπορά 1) είναι μια στοχαστική διαδικασία. Τώρα υπολογίζουμε την πιθανότητα

$$P\left\{ \int_0^1 X_t^{\,2}\, dt > \frac{1}{2} \right\}.$$

Βλέπουμε ότι

$$\int_0^1 X_t^2\, dt = a^2 \int_0^1 cos^2(2\pi t)dt + 2ab\int_0^1 cos(2\pi t)\, sin(2\pi t)\, dt + b^2 \int_0^1 sin^2(2\pi t)dt = \frac{1}{2}a^2 + \frac{1}{2}b^2 .$$

Οπότε

$$P\left\{ \int_0^1 X_t^2\, dt > \frac{1}{2} \right\} = P\left(a^2 + b^2 > 1\right) = \int_1^\infty f(z)dz$$

όπου $f(z)$ είναι η πυκνότητα της τυχαίας μεταβλητής $z = a^2 + b^2$. Μπορεί ν' αποδειχθεί ότι

$$f(z) = \frac{1}{2} e^{-z/2} \text{ και άρα}$$

$$\int_1^\infty \frac{1}{2} e^{-\frac{z}{2}} dz = e^{-1/2} .$$

Παράδειγμα 1.5

Θεωρήστε την αιτιοκρατική συνάρτηση σαν ακραία περίπτωση διαδικασίας:

$$X_t = \cos\omega t,$$

όπου το ω είναι δεδομένο. Τότε

$$\overline{X}_t = E(\cos\omega t) = \cos\omega t$$

και

$$R(t_1, t_2) = E(\cos\omega t_1 \cos\omega t_2) = \cos\omega t_1 \cos\omega t_2.$$

Παράδειγμα 1.6

Μια στοχαστική διαδικασία χαρακτηρίζεται από τις ακόλουθες ποσότητες:

$$\overline{X}_t = \frac{1}{2},\ R(t_1, t_2) = e^{-|t_1 - t_2|}.$$

Εύρετε τη μέση τιμή, τη συμμεταβλητότητα, και τη διασπορά των τυχαίων μεταβλητών

$$Y = X_4 \text{ και } Z = X_6 .$$

Προφανώς

$$\overline{Y} = \overline{X}_4 = \overline{Z} = \overline{X}_6 = \frac{1}{2}$$

$$E\left(Y^2\right) = E\left(X_4^2\right) = R(4,4) = e^{-|4-4|} = 1 = E\left(Z^2\right)$$

$$Var(Y) = Var(Z) = 1 - \left(\frac{1}{2}\right)^2 = \frac{3}{4}.$$

Τελικά

$$E(YZ) = R(4,6) = e^{-|4-6|} = 0.135$$

$$C(4,6) = R(4,6) - \overline{X}_4\overline{X}_6 = 0.135 - \left(\frac{1}{2}\right)^2 = -0.115 .$$

1.3 Λογισμός Μέσου Τετραγώνου

Έστω μια ακολουθία τ.μ X_n. Λέμε ότι η ακολουθία **συγκλίνει στην τ.μ. *X* κατά μέσο τετράγωνο** αν

$$E(|X_n|^2) < \infty ,\ \ \forall\ n \text{ και}$$

$$\lim_{n\to\infty} E\left(|X_n - X|^2\right) = 0 .$$

Συμβολικά γράφουμε $l.i.m.X_n = X$.

Η σύγκλιση για συναρτήσεις X_t ορίζεται ανάλογα. Μια στοχαστική διαδικασία X_t είναι **διαφορίσιμη κατά μέσο τετράγωνο** στο t, αν το ακόλουθο όριο υπάρχει

$$\dot{X}_t = \frac{dX_t}{dt} \equiv \underset{h\to 0}{l.i.m.} \frac{X_{t+h} - X_t}{h} .$$

Θεώρημα 1.1

Αν η διαδικασία X_t είναι διαφορίσιμη κατά μέσο τετράγωνο για κάθε t τότε ισχύουν τα ακόλουθα:

$$E(\dot{X}_t) = \overline{\dot{X}}_t = \frac{d}{dt} E(X_t) = \dot{\overline{X}}_t$$

$$R_{\dot{X}X}(t,\tau) = E(\dot{X}_t X_\tau) = \frac{\partial}{\partial t} E(X_t X_\tau) = \frac{\partial R_{XX}(t,\tau)}{\partial t}$$

$$R_{\dot{X}\dot{X}}(t,\tau) = E(\dot{X}_t \dot{X}_\tau) = \frac{\partial}{\partial t}\frac{\partial}{\partial \tau} \mathrm{E}(X_t X_\tau) = \frac{\partial^2 R_{XX}(t,\tau)}{\partial t\, \partial \tau}$$

$$C_{\dot{X}X}(t,\tau) = E(\dot{X}_t X_\tau) - E(\dot{X}_t)E(X_\tau) = \frac{\partial C_{XX}(t,\tau)}{\partial t}$$

$$C_{\dot{X}\dot{X}}(t,\tau) = E(\dot{X}_t \dot{X}_\tau) - E(\dot{X}_t)E(\dot{X}_\tau) = \frac{\partial^2 C_{XX}(t,\tau)}{\partial t\, \partial \tau} .$$

Με άλλα λόγια οι πράξεις της προσδοκητής τιμής και της παραγώγου κατά μέσο τετράγωνο εναλλάσσονται.

Μια στοχαστική διαδικασία X_t είναι **Riemann ολοκληρώσιμη κατά μέσο τετράγωνο** στο διάστημα $[\alpha, \beta]$ εάν το ακόλουθο όριο υπάρχει:

$$\underset{\delta\to 0}{l.i.m.} \sum_{i=0}^{n-1} X_{t_1'}(t_{i+1} - t_i) \equiv \int_\alpha^\beta X_t dt$$

όπου έχουμε κάνει την εξής διαίρεση χρόνων

$$\alpha = t_0 < t_1 < ... < t_n = \beta, \qquad \delta = \max_i (t_{i+1} - t_i), \qquad t_i \le t_i' < t_{i+1}.$$

Θεώρημα 1.2

Έστω ότι η διαδικασία X_t είναι Riemann ολοκληρώσιμη κατά μέσο τετράγωνο στα διαστήματα $[\alpha, \beta]$, $[\gamma, \delta]$. Τότε ισχύουν τα ακόλουθα:

$$E\left(\int_\alpha^\beta X_t dt\right) = \int_\alpha^\beta E(X_t) dt,$$

$$E\left(\int_\alpha^\beta X_t dt \int_\gamma^\delta X_\tau d\tau\right) = \int_\alpha^\beta \int_\gamma^\delta R(t,\tau) dt\, d\tau,$$

$$Cov\left(\int_\alpha^\beta X_t dt\ , \int_\gamma^\delta X_\tau d\tau\right) = \int_\alpha^\beta \int_\gamma^\delta C(t,\tau) dt\, d\tau.$$

Θεώρημα 1.3 *Το Θεμελιώδες Θεώρημα του Λογισμού Μέσου Τετραγώνου*

Θεωρήστε τη διαδικασία $\dot{X}_t$ που είναι Riemann ολοκληρώσιμη κατά μέσο τετράγωνο στο διάστημα $[\alpha, \beta]$. Τότε

$$\int_a^t \dot{X}_\tau d\tau = X_t - X_a \text{ με πιθανότητα 1.}$$

Βλέπουμε ότι οι γραμμικοί τελεστές $E(.)$, $\frac{d}{dt}(.)$ και $\int (.)dt$ κατά μέσο τετράγωνο εναλλάσσονται.

Παράδειγμα 1.7

Έστω $X_t = at$ όπου a είναι μια τυχαία μεταβλητή με πεπερασμένη μέση τιμή και πεπερασμένη διασπορά. Χρησιμοποιώντας τον κλασικό λογισμό $\dot{X}_t = a$. Επίσης

$$E\left(\left|\frac{a(t+h) - a\,t}{h} - a\right|^2\right) = E(|a-a|^2) = 0.$$

Άρα η διαδικασία είναι διαφορίσιμη κατά μέσο τετράγωνο.

Θεώρημα 1.4

1. Έστω a_n είναι μια ακολουθία πραγματικών (αιτιοκρατική). Τότε $l.i.m.a_n = lim a_n$.
2. Έστω X μια τυχαία μεταβλητή. Τότε $l.i.m.\ X = lim\ X = X$ και $l.i.m.(a_n X) = X\ lim a_n^{(*)}$.
3. Έστωσαν X_n και Y_n δύο τυχαίες ακολουθίες και a και b δύο σταθερές. Τότε αν $l.i.m.X_n = X$ και $l.i.m.Y_n = Y$

   $l.i.m.aX_n = a\ l.i.m.X_n = aX$

   $l.i.m.(aX_n + bY_n) = a\ l.i.m.X_n + b\ l.i.m.Y_n = aX + bY.$

Παράδειγμα 1.8

Για την ίδια διαδικασία, $\int_0^t X_\tau d\tau = \int_0^t a\tau d\tau$ και πάλι από τον κλασικό λογισμό $= \frac{1}{2}at^2$.

Πράγματι, $\int_0^t X_\tau d\tau = \underset{n\to\infty}{l.i.m.} \sum_{i=0}^{n-1} a\tau_i'(\tau_{i+1} - \tau_i) = \underset{n\to\infty}{l.i.m}\ a\sum_{i=0}^{n-1}\tau_i'(\tau_{i+1} - \tau_i)$. Αλλά αυτό είναι το σύνηθες όριο που ορίζει ένα ολοκλήρωμα Riemann, ήτοι

$$\underset{n\to\infty}{l.i.m}\ a\sum_{i=0}^{n-1}\tau_i'(\tau_{i+1} - \tau_i) \overset{(*)}{=} \lim_{n\to\infty} a\sum_{i=0}^{n-1}\tau_i'(\tau_{i+1} - \tau_i) = a\left[\lim_{n\to\infty}\sum_{i=0}^{n-1}\tau_i'(\tau_{i+1} - \tau_i)\right].$$

Ως γνωστόν αυτό ισούται με $\frac{1}{2}at^2$. Για να εξισώσουμε το *l.i.m.* με το κλασικό όριο *lim* κάναμε τη χρήση του Θεωρήματος 1.4 (*).

1.4 Ανεξάρτητες και Ασυσχέτιστες Διαδικασίες

Η διαδικασία X_t λέγεται **ανεξάρτητη** αν για κάθε σύνολο σημείων $t_1, t_2, ..., t_n$ οι τυχαίες μεταβλητές $X_{t_1}, X_{t_2}, \ldots, X_{t_n}$ είναι ανεξάρτητες.

Οι διαδικασίες X_t και Y_t λέγονται ανεξάρτητες αν τα διανύσματα $[X_{t_1}\ X_{t_2}\ \ldots\ X_{t_n}]^{\mathrm{T}}$ και $[Y_{t_1}\ Y_{t_2}\ \ldots\ Y_{t_n}]^{\mathrm{T}}$είναι ανεξάρτητα. Ήτοι

$$P(X_{t_1} \le x_1, \ldots X_{t_n} \le x_n, Y_{t_1} \le y_1, \ldots Y_{t_n} \le y_n)$$

$$= P(X_{t_1} \le x_1, \ldots X_{t_n} \le x_n)\, P(Y_{t_1} \le y_1, \ldots Y_{t_n} \le y_n).$$

Η διαδικασία X_t λέγεται **ασυσχέτιστη** αν

$$E(X_t X_\tau) = E(X_t)\, E(X_\tau) \quad \text{για κάθε } t \neq \tau.$$

Μια ανεξάρτητη διαδικασία είναι ασυσχέτιστη αλλά **όχι** αντιστρόφως.

Δύο διαδικασίες λέγονται **ασυσχέτιστες** αν

$$E(X_t Y_\tau) = E(X_t)\, E(Y_\tau) \quad \text{για οποιαδήποτε } t \text{ και } \tau.$$

Ιδιότητες: Γενικά

$$Var(X_t + Y_\tau) = Var(X_t) + Var(Y_\tau) + 2Cov(X_t, Y_\tau).$$

Αν οι X_t, Y_τ είναι ασυσχέτιστες τότε προκύπτει $Cov(X_t, Y_\tau) = 0$ και

$$Var(X_t + Y_\tau) = Var(X_t) + Var(Y_\tau).$$

Μια διαδικασία έχει **ανεξάρτητα τμήματα** αν

$X_{t_0} = 0$ και για όλα τα σημεία $t_0 < t_1 < \ldots < t_n$ οι τυχαίες μεταβλητές $X_{t_1} - X_{t_0}$, $X_{t_2} - X_{t_1}$, …, $X_{t_n} - X_{t_{n-1}}$ είναι ανεξάρτητες.

II ΕΙΔΙΚΕΣ ΔΙΑΔΙΚΑΣΙΕΣ

2.1 Διαδικασία Wiener

Υπάρχει μία πληθώρα φυσικών φαινομένων που μπορούν να μοντελοποιηθούν ικανοποιητικά με τη **διαδικασία Wiener (**ή Wiener-Lévy ή Brown).

Ένα μικρό σωματίδιο που βρίσκεται μέσα σε κάποιο ρευστό κινείται ακανόνιστα. Μια τέτοια κίνηση λέγεται κίνηση Brown. Ο Einstein εξήγησε αυτή την κίνηση στα πλαίσια της στατιστικής μηχανικής σαν αποτέλεσμα των συνεχών συγκρούσεων του σωματιδίου με τα μόρια του ρευστού.

Έστω ότι X_t είναι η μετατόπιση του σωματιδίου στη χρονική στιγμή t και $X_0 = 0$. Η μετατόπιση που υφίσταται το σωματίδιο στο διάστημα (t, τ) οφείλεται σε μεγάλο αριθμό συγκρούσεων και κατά συνέπεια του Κεντρικού Οριακού Θεωρήματος η διαφορά $X_\tau - X_t$ είναι μία στοχαστική διαδικασία κατανεμημένη κατά Gauss. Επίσης, επειδή το ρευστό είναι σε ισορροπία η κατανομή της διαφοράς $X_{\tau+h} - X_{t+h}$ θα υποτεθεί επίσης Gauss. Όπως θα ιδούμε αργότερα μία διαδικασία Y_t τέτοια ώστε η στατιστική της είναι ταυτόσημη με αυτήν της Y_{t+h} , $\forall h$ λέγεται **αμετάβλητη κατά τη στενή έννοια**.

Επειδή οι συγκρούσεις είναι τελείως «τυχαίες» θα υποθέσουμε ότι η διαδικασία έχει ανεξάρτητα τμήματα. Αυτές είναι οι διαισθητικές εξηγήσεις της διαδικασίας Wiener.

Αυστηρά τώρα λέμε ότι μια διαδικασία $\{X_t, t \geq 0\}$ είναι Wiener εάν

α. $\{X_t, t \geq 0\}$ έχει αμετάβλητα και ανεξάρτητα τμήματα.

β. Η X_t είναι κανονική $\forall\ t \geq 0$.

γ. $E(X_t) = 0\ \ \forall\ t \geq 0$.

δ. $P(X_0 = 0) = 1$ δηλαδή $X_0 = 0$ με πιθανότητα 1 (μ.π. 1).

Ο θερμικός θόρυβος στα ηλεκτρικά κυκλώματα και η κίνηση τιμών του χρηματιστήριου είναι άλλα παραδείγματα διαδικασίας Wiener.

Παρατηρούμε ότι

$$Var(X_t) = Var(X_t - X_0) = Var(X_{t+\tau} - X_\tau) \tag{2.1}$$

$$Var(X_\tau) = Var(X_\tau - X_0) \tag{2.2}$$

αφού $X_t = X_t - X_0$ μ.π. 1 και οι $(X_t - X_0)$, $(X_{t+\tau} - X_\tau)$ έχουν την ίδια στατιστική. Από τις (2.1) και (2.2) προκύπτει

$$Var(X_t) + Var(X_\tau) = Var(X_{t+\tau} - X_\tau) + Var(X_\tau - X_0)$$

Επειδή οι $(X_t - X_0)$, $(X_{t+\tau} - X_\tau)$ είναι ανεξάρτητες είναι και ασυσχέτιστες. Από τις Ιδιότητες ασυσχέτιστων διαδικασιών προκύπτει $Var(X_{t+\tau} - X_\tau) + Var(X_\tau - X_0) = Var(X_{t+\tau} - X_\tau + X_\tau - X_0)$. Οπότε τελικά

$$Var(X_t) + Var(X_\tau) = Var(X_{t+\tau} - X_\tau) + Var(X_\tau - X_0) = Var(X_{t+\tau}). \tag{2.3}$$

Η (2.3) είναι της μορφής $f(t_1+t_2) = f(t_1) + f(t_2)$ όπου $f(t) \geq 0,\ t \geq 0$. Οι μόνες συναρτήσεις που ικανοποιούν αυτή τη σχέση και λαμβάνουν τιμές στο R είναι της μορφής $f(t) = \sigma^2 t$, $\forall t \geq 0$, όπου $\sigma \in R$. Από την (2.3) βλέπουμε ότι

$$Var(X_{t+\tau} - X_\tau) = Var(X_{t+\tau}) - Var(X_\tau).$$

Θέσετε $t + \tau \equiv s$ τότε $Var(X_s - X_\tau) = \sigma^2 s - \sigma^2 \tau = \sigma^2 (s - \tau)$ και εν γένει

$$Var(X_t - X_\tau) = \sigma^2 |t - \tau| \text{, όπου } t \geq \tau \text{ ή } t \leq \tau. \tag{2.4}$$

Η παράμετρος σ^2 λέγεται **παράμετρος μεταβλητότητας** και υπολογίζεται εμπειρικά. Η συσχέτιση είναι:

$$R(t,\tau) = E(X_t X_\tau) = E\{[(X_t - X_\tau) + X_\tau]X_\tau\}$$

$$= E[(X_t - X_\tau)X_\tau] + E(X_\tau^2) = 0 + \sigma^2 \tau$$

επειδή $X_\tau = X_\tau - X_0$ και τα τμήματα $(X_t - X_\tau)$ και $(X_\tau - X_0)$ είναι ανεξάρτητα εφ' όσον $t > \tau$. Εν γένει

$$R(t,\tau) = \sigma^2 \min(t,\tau) \text{, όπου } t \geq \tau \text{ ή } t \leq \tau. \tag{2.5}$$

Ορισμός 2.1

Η διαδικασία X_t λέγεται συνεχής κατά μέσο τετράγωνο αν $\underset{h \to 0}{l.i.m.} X_{t+h} = X_t$.

Θεώρημα 2.1

Η X_t είναι συνεχής κατά μέσο τετράγωνο τότε και μόνον τότε αν η συνάρτηση $R(t, t)$ είναι συνεχής.

Θεώρημα 2.2

Η X_t είναι διαφορίσιμη κατά μέσο τετράγωνο τότε και μόνον τότε αν η $\dfrac{\partial^2 R(t,\tau)}{\partial t \partial \tau}$ υπάρχει στο σημείο (t, t).

Θεώρημα 2.3

Η X_t είναι Riemann ολοκληρώσιμη κατά μέσο τετράγωνο στο $[\alpha, \beta]$ τότε και μόνον τότε αν η $R(t, \tau)$ είναι Riemann ολοκληρώσιμη στο $[\alpha, \beta] \times [\alpha, \beta]$.

Με βάση τα πιο πάνω θεωρήματα παρατηρούμε ότι η διαδικασία Wiener είναι:

1. συνεχής κατά μέσο τετράγωνο $\forall\ t \geq 0$.
2. η παράγωγος $\left.\dfrac{\partial^2 \min(t,\tau)}{\partial t \partial \tau}\right|_{(t,t)}$ δεν υπάρχει πουθενά και άρα η διαδικασία **δεν** είναι διαφορίσιμη σε κανένα σημείο (παρ' ότι είναι συνεχής!).
3. ολοκληρώσιμη κατά μέσο τετράγωνο $\forall\ t \geq 0$.

Η διαδικασία Wiener μπορεί να θεωρηθεί σαν το όριο της διαδικασίας **τυχαίου περιπάτου**. Πριν δούμε αυτήν χρειαζόμαστε τα ακόλουθα αποτελέσματα.

Η **χαρακτηριστική συνάρτηση** μιας τυχαίας μεταβλητής ορίζεται

$$\Phi_X(\omega) = E\left(e^{i\omega X}\right) = \int_{-\infty}^{\infty} e^{i\omega x} f(x)\, dx\,, \quad \text{όπου } i = \sqrt{-1}\,.$$

Αν η χαρακτηριστική συνάρτηση είναι απολύτως ολοκληρώσιμη τότε ο αντίστροφος μετασχηματισμός δίνει την συνάρτηση πυκνότητας πιθανότητας

$$f(x) = \frac{1}{2\pi}\int_{-\infty}^{\infty} e^{-i\omega x}\Phi_X(\omega)d\omega .$$

Ιδιότητες: *Θεώρημα των ροπών:* $\Phi_X(0) = 1, \Phi'_X(0) = i\,E(X)$,

$$\Phi''_X(0) = i^2 E(X^2) = -E(X^2), \ldots.$$

Συνέλλιξη: *Αν X, Y ανεξάρτητες τ.μ. και Z = X + Y τότε*

$$\Phi_Z(\omega) = \Phi_X(\omega)\,\Phi_Y(\omega).$$

Η χαρακτηριστική συνάρτηση μιας κανονικής τυχαίας μεταβλητής με μέση τιμή $\bar{X}$ και διασπορά σ^2 δίνεται από

$$\Phi_X(\omega) = e^{i\bar{X}\omega - \frac{1}{2}\sigma^2\omega^2} .$$

Για διακεκριμένη τ.μ. X που λαμβάνει τιμές x_n με πιθανότητες p_n, $n\in\{1,2,\ldots\}$ η $\Phi_X(\omega)$ ορίζεται ομοίως

$$\Phi_X(\omega) = E(e^{i\omega X}) = \sum_{n=1}^{\infty} e^{i\omega x_n} p_n .$$

2.2 Τυχαίος Περίπατος

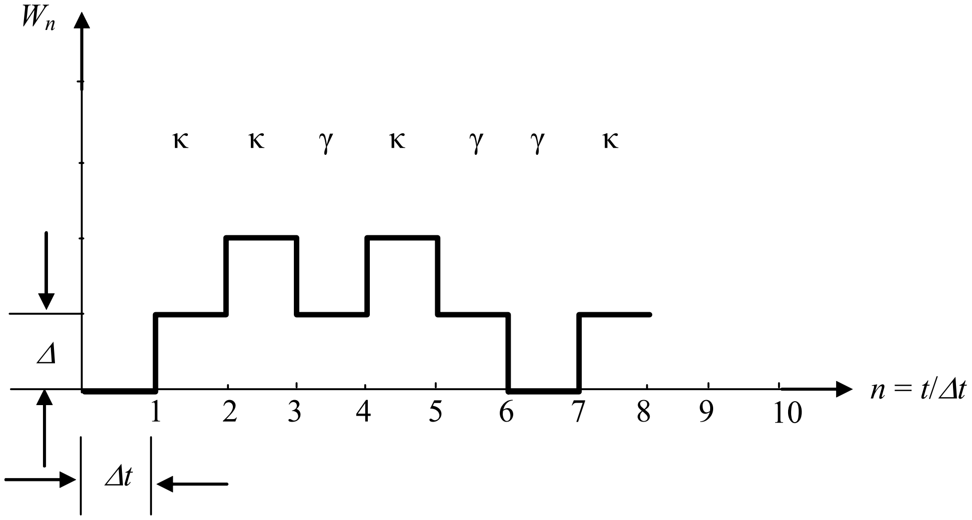
(Random Walk)

Ρίχνουμε ένα νόμισμα και αν το αποτέλεσμα είναι κορώνα (h) τότε κάνουμε ένα βήμα εμπρός, ενώ αν είναι γράμματα (t) ένα βήμα πίσω. Τα βήματα έχουν μήκος Δ. Έστω $P(h) = p$, $P(t) = q = 1 - p$. Τότε ορίζουμε την εξής τυχαία μεταβλητή:

$$X_n = \begin{cases} +\Delta, & \text{όταν } \omega = h \quad \text{μ.π.} \quad p \\ -\Delta, & \text{όταν } \omega = t \quad \text{μ.π.} \quad q \end{cases}$$

όπου n = 1, 2, ... είναι οι διακεκριμένες στιγμές των ρίψεων. Η συνολική μετατόπιση W_n, εφ' όσον $W_0 = 0$, είναι

$$W_n = \sum_{k=1}^{n} X_k , \qquad n = 1, 2, \ldots \tag{2.7}$$



Η χαρακτηριστική συνάρτηση της X_n είναι

$$\Phi_{X_n}(\omega) = pe^{i\omega\Delta} + qe^{-i\omega\Delta}.$$

Είναι γνωστό ότι η χαρακτηριστική συνάρτηση του αθροίσματος ανεξάρτητων τυχαίων μεταβλητών είναι ίση με το γινόμενο των ατομικών χαρακτηριστικών συναρτήσεων. Από την (2.7) λαμβάνουμε

$$\Phi_{W_n}(\omega) = (pe^{i\omega\Delta} + qe^{-i\omega\Delta})^n.$$

Αφού οι διακεκριμένες χρονικές στιγμές έχουν μέγεθος Δt τότε στη χρονική στιγμή t, $n = \dfrac{t}{\Delta t}$. Στο εξής υποθέτουμε ότι $p = q = 0.5$ και $t < \infty$. Τότε

$$\Phi_{W_n}(\omega) = \left(0.5e^{i\omega\Delta} + 0.5e^{-i\omega\Delta}\right)^{\frac{t}{\Delta t}}. \tag{2.8}$$

Από το θεώρημα των ροπών έχουμε:

$$E(W_n) = \frac{1}{i}\Phi'_{W_n}(0) = 0 \text{ και} \tag{2.9}$$

$$Var(W_n) = E(W_n^2) - [E(W_n)]^2 = E(W_n^2) = \frac{1}{i^2}\Phi''_{W_n}(0) = \frac{t\Delta^2}{\Delta t}. \tag{2.10}$$

Τώρα για μία ορισμένη χρονική στιγμή $t < \infty$ παίρνουμε το όριο του περιπάτου καθώς $\Delta t \to 0$. Προφανώς $n \to \infty$. Ορίζουμε το όριο αυτό ως τη διαδικασία $W_t \triangleq \lim\limits_{\Delta t \to 0} W_n$. Για να έχει νόημα η (2.10) για κάθε t καθώς $\Delta t \to 0$ υποθέτουμε ότι το Δ^2 συγκλίνει προς το 0 με την ίδια ταχύτητα που συγκλίνει το Δt. Συγκεκριμένα, θεωρούμε ότι $\Delta^2 = \sigma^2 \Delta t$ έτσι ώστε η διαδικασία που θα προκύψει να έχει πεπερασμένη μεταβλητότητα για κάθε t καθώς $\Delta t \to 0$. Παίρνοντας όρια στην εξίσωση (2.8) και θέτοντας $\Delta = \sigma \sqrt{\Delta t}$ ευρίσκουμε

$$\Phi_{W_t}(\omega) \triangleq \lim_{\Delta t \to 0} \Phi_{W_n}(\omega) = \lim_{\Delta t \to 0} \left(0.5 e^{i\omega\sigma\sqrt{\Delta t}} + 0.5 e^{-i\omega\sigma\sqrt{\Delta t}}\right)^{\frac{t}{\Delta t}}$$

Παίρνοντας λογαρίθμους, η εξίσωση αυτή γράφεται

$$\ln[\Phi_{W_t}(\omega)] = \ln\left[\lim_{\Delta t \to 0}\left(0.5 e^{i\omega\sigma\sqrt{\Delta t}} + 0.5 e^{-i\omega\sigma\sqrt{\Delta t}}\right)^{\frac{t}{\Delta t}}\right]$$

και, επειδή ο λογάριθμος είναι συνεχής συνάρτηση, ln(*lim*…) = *lim*(ln…) ήτοι

$$\ln[\Phi_{W_t}(\omega)] = \lim_{\Delta t \to 0} \ln\left(0.5 e^{i\omega\sigma\sqrt{\Delta t}} + 0.5 e^{-i\omega\sigma\sqrt{\Delta t}}\right)^{\frac{t}{\Delta t}} = t \lim_{\Delta t \to 0} \frac{\ln\left(0.5 e^{i\omega\sigma\sqrt{\Delta t}} + 0.5 e^{-i\omega\sigma\sqrt{\Delta t}}\right)}{\Delta t}$$

Στην τελευταία έκφραση, αριθμητής και παρονομαστής τείνουν στο 0. Εφαρμόζοντας το κανόνα L΄Hospital ευρίσκουμε

$$\ln[\Phi_{W_t}(\omega)] = t \lim_{\Delta t \to 0} \frac{(\ )'}{(\Delta t)'} = t \lim_{\Delta t \to 0} \frac{0.5 i\omega\sigma\left(\sqrt{\Delta t}\right)' e^{i\omega\sigma\sqrt{\Delta t}} - 0.5 i\omega\sigma\left(\sqrt{\Delta t}\right)' e^{-i\omega\sigma\sqrt{\Delta t}}}{0.5 e^{i\omega\sigma\sqrt{\Delta t}} + 0.5 e^{-i\omega\sigma\sqrt{\Delta t}}}$$

$$= i\omega\sigma t \lim_{\Delta t \to 0}\left(\frac{1}{2\sqrt{\Delta t}} \frac{e^{i\omega\sigma\sqrt{\Delta t}} - e^{-i\omega\sigma\sqrt{\Delta t}}}{e^{i\omega\sigma\sqrt{\Delta t}} + e^{-i\omega\sigma\sqrt{\Delta t}}}\right) = \frac{1}{2} i\omega\sigma t \lim_{\Delta t \to 0}\left(\frac{1}{\sqrt{\Delta t}} \frac{2i \sin \omega\sigma\sqrt{\Delta t}}{2 \cos \omega\sigma\sqrt{\Delta t}}\right)$$

$$= -\frac{1}{2}\omega\sigma t \lim_{\Delta t \to 0} \frac{\tan \omega\sigma\sqrt{\Delta t}}{\sqrt{\Delta t}}$$

Αποδεικνύεται ότι $\lim_{x \to 0} \frac{\tan ax}{x} = a$ (εφαρμόζοντας πάλι τον κανόνα L΄Hospital) οπότε

$$\ln[\Phi_{W_t}(\omega)] = -\frac{1}{2}\omega^2\sigma^2 t \Rightarrow \qquad \Phi_{W_t}(\omega) = e^{-\frac{1}{2}\omega^2\sigma^2 t} \tag{2.11}$$

Η (2.11) είναι η χαρακτηριστική συνάρτηση μιας κανονικής διαδικασίας με μέση τιμή 0 και διασπορά $\sigma^2 t$. Ο τυχαίος περίπατος έχει ανεξάρτητα τμήματα και άρα στο όριο είναι διαδικασία Wiener.

2.3 Λευκός θόρυβος

(White noise)

Ορισμός 2.2

Μια διαδικασία $\{X_t,\ t\in T\}$ λέγεται **Markov** εάν για κάθε σύνολο $\{t_i,\ t_{i+1} > t_i\}\in T$ και πραγματικό a

$$P(X_{t_n} \le a \mid X_{t_1},\ldots,X_{t_{n-1}}) = P(X_{t_n} \le a \mid X_{t_{n-1}}).$$

Για διαδικασίες που έχουν συναρτήσεις πυκνότητας γράφουμε

$$f_{X_{t_n}}(x_n \mid X_{t_1}= x_1,\ldots,X_{t_{n-1}}= x_{n-1}) \triangleq f(x_n \mid x_1,\ldots, x_{n-1}) = f(x_n \mid x_{n-1}).$$

όπου x_1, … x_n είναι πιθανές τιμές των X_{t_1}, …, X_{t_n}. Από τον ορισμό της κατά συνθήκη πιθανότητας

$$f(x_1,\ldots, x_n) = f(x_n \mid x_1,\ldots, x_{n-1})\, f(x_1,\ldots, x_{n-1}).$$

Αν η διαδικασία είναι Markov τότε

$$f(x_1,\ldots, x_n) = f(x_n \mid x_{n-1})\, f(x_1,\ldots, x_{n-1})$$

και επαγωγικά:

$$f(x_1,\ldots, x_n) = f(x_n \mid x_{n-1})\, f(x_{n-1} \mid x_{n-2})\ldots f(x_2 \mid x_1)\, f(x_1).$$

Τώρα θεωρούμε μια τυχαία ακολουθία $\{X_n,\ n = 1,2,...\}$ η οποία είναι Markov και

$$P(x_k \mid x_m) = P(x_k) \text{ όπου } k > m,$$

δηλαδή αν γνωρίζουμε την X_m για κάποιο εξαγόμενο ω δεν μπορούμε να βγάλουμε οιοδήποτε συμπέρασμα (δεν έχουμε καμμία πληροφορία) για την $X_k(\omega)$. Έστω ότι η διαδικασία είναι κανονική. Τότε αυτή η διαδικασία λέγεται **λευκή κανονική ακολουθία**. Η αυτοσυμμεταβλητότητα είναι

$$C(t_n, t_m) = E[(X_{t_n} - \overline{X}_{t_n})(X_{t_m} - \overline{X}_{t_m})] = Q_n\delta_{nm}$$

όπου $Q_n > 0$ και το **δέλτα του Kronecker** ορίζεται

$$\delta_{nm} = \begin{cases} 0, & n \neq m \\ 1, & n = m \end{cases}.$$

Αφού η ακολουθία είναι Gauss, τότε χρειάζεται μόνο η προσδοκητή τιμή $E(X_n)$, $\forall n$ και η $C(t_n, t_m)$ για τον προσδιορισμό της στατιστικής της.

Έστω τώρα ότι η διαδικασία $\{X_t,\ t \geq 0\}$ είναι Wiener. Η αυτοσυμμεταβλητότητα είναι

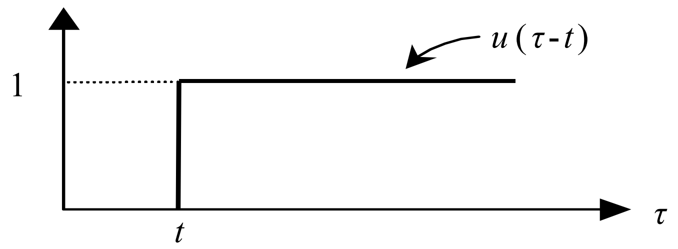
$$C_{XX}(t,\tau) = \sigma^2 \min(t,\tau) \text{ και}$$

$$C_{\dot{X}\dot{X}}(t,\tau) = \frac{\partial^2 C_{XX}(t,\tau)}{\partial t \partial \tau} = \sigma^2 \frac{\partial^2 \min(t,\tau)}{\partial t \partial \tau}.$$

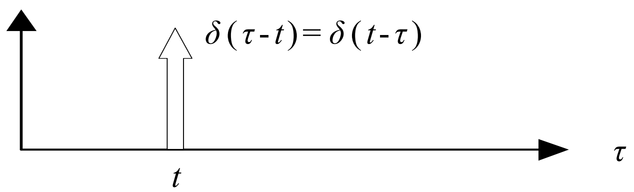
Γνωρίζουμε ήδη ότι η $\dot{X}_t$ δεν υπάρχει πουθενά. Εν τούτοις εδώ προχωρούμε συμβολικά. Τώρα, αφού $\min(t, \tau) = \begin{cases} \tau, & \tau < t \\ t, & \tau \geq t \end{cases}$ προκύπτει ότι

$$\frac{\partial \min(t,\tau)}{\partial t} = \begin{cases} 0, & \tau < t \\ 1, & \tau > t \end{cases}$$

που είναι η γενικευμένη **βηματική** ή **Heaviside συνάρτηση**.



Η παράγωγος της βηματικής είναι η συμμετρική συνάρτηση **δέλτα του Dirac**:



Άρα

$$C_{\dot{X}\dot{X}}(t,\tau) = \sigma^2 \delta(t-\tau).$$

Η διαδικασία $\dot{X}_t$ λέγεται **λευκός θόρυβος** και είναι μια ιδιόμορφη διαδικασία, κάτι ανάλογο των γενικευμένων συναρτήσεων, για τούτο λέγεται και **γενικευμένη διαδικασία**.

Ορισμός 2.3

Λευκή κανονική διαδικασία ονομάζουμε την κανονική διαδικασία $\{X_t,\ t\in T\}$ για την οποία

$$C(t,\tau) = Q(t)\delta(t-\tau)\text{, όπου } Q(t) \geq 0.$$

Αργότερα θα ιδούμε γιατί αυτή η διαδικασία λέγεται λευκή.

2.4 Αμετάβλητες Διαδικασίες - Φάσμα - Συσχετίσεις

Στην παράγραφο αυτή θα εξετάσουμε στοχαστικές διαδικασίες οι οποίες λαμβάνουν μιγαδικές τιμές. Για παράδειγμα, για ένα σώμα που κινείται επάνω σε μία ευθεία, οι μεταβλητές θέση (α_t) και ταχύτητα (β_t) περιγράφονται από κοινού με τη μιγαδική διαδικασία $X_t = \alpha_t + i\beta_t$.

Ορισμός 2.4

Μια διαδικασία X_t λέγεται **αμετάβλητη κατά τη στενή έννοια** αν η X_t και η $X_{t+\tau}$ έχουν την ίδια στατιστική για κάθε τ.

Οι διαδικασίες X_t και Y_t λέγονται αμετάβλητες κατά τη στενή έννοια αν η στατιστική των (X_t, Y_t) είναι η ίδια με τη στατιστική των ($X_{t+\tau}$, $Y_{t+\tau}$) για κάθε τ.

Ορισμός 2.5

Η διαδικασία X_t λέγεται **αμετάβλητη κατά την ευρεία έννοια** αν η μέση της τιμή

$$E(X_t) = \overline{X} \text{ είναι σταθερή}$$

και η αυτοσυσχέτιση εξαρτάται μόνο από το $\tau = t_1 - t_2$, δηλαδή

$$R(t_1, t_2) \triangleq E\{X_{t_1} X_{t_2}^*\} = R(t_1 - t_2) = R(\tau).$$

Το X_t^* είναι το μιγαδικό συζυγές της X_t. Για μιά τέτοια διαδικασία ορίζουμε

$$R(\tau) \triangleq E\{X_{t+\tau} X_t^*\}$$

και

$$R(0) = E\{|X_t|^2\}$$

που είναι η **μέση τιμή της ισχύος** της διαδικασίας στον t. Από την πιο πάνω εξίσωση βλέπουμε ότι

$$R(-\tau) = E\{X_{t-\tau} X_t^*\}$$

και, αφού η X_t είναι αμετάβλητη κατά την ευρεία έννοια,

$$R(-\tau) = E\{X_t X_{t+\tau}^*\} = [E\{X_t^* X_{t+\tau}\}]^* = [E\{X_{t+\tau} X_t^*\}]^* = R^*(\tau).$$

Επίσης ορίζουμε τη **συσχέτιση**:

$$R_{XY}(\tau) \triangleq E\{X_{t+\tau} Y_t^*\}.$$

Προφανώς οι **συμμεταβλητότητες** είναι:

$$C(\tau) = R(\tau) - |\overline{X}|^2 = R(\tau) - \overline{X}\ \overline{X}^* \ ,$$

$$C_{XY}(\tau) = R_{XY}(\tau) - \overline{X}\ \overline{Y}^* \ .$$

Επίσης

$$R_{XY}(-\tau) = R_{YX}^*(\tau).$$

Αν η X_t είναι πραγματική τότε η $R(\tau)$ είναι πραγματική και άρτια συνάρτηση, δηλαδή $R(-\tau) = R(\tau)$.

Ορισμός 2.6

Το **φάσμα ισχύος** ή **φασματική πυκνότητα** $S(\omega)$ μιας διαδικασίας X_t, είναι ο μετασχηματισμός Fourier της αυτοσυσχέτισης $R(\tau)$,

$$S(\omega) \triangleq \int_{-\infty}^{\infty} R(\tau)e^{-i\omega\tau}d\tau$$

και αντίστροφα,

$$R(\tau) = \frac{1}{2\pi}\int_{-\infty}^{\infty} S(\omega)e^{i\omega\tau}d\omega \, .$$

Επειδή $R(-\tau) = R^*(\tau)$, η συνάρτηση $S(\omega)$ είναι πάντα πραγματική. Πράγματι

$$S^*(\omega) = \int_{-\infty}^{\infty} R^*(\tau)\left(e^{-i\omega\tau}\right)^* d\tau = \int_{-\infty}^{\infty} R(-\tau)e^{i\omega\tau}d\tau \underset{(t=-\tau)}{=} \int_{-\infty}^{\infty} R(t)e^{-i\omega t}dt = S(\omega)$$

Το φάσμα **δύο** διαδικασιών X_t, Y_t ορίζεται

$$S_{XY}(\omega) \triangleq \int_{-\infty}^{\infty} R_{XY}(\tau)e^{-i\omega\tau}d\tau$$

και αντίστροφα,

$$R_{XY}(\tau) = \frac{1}{2\pi}\int_{-\infty}^{\infty} S_{XY}(\omega)e^{i\omega\tau}d\omega \, .$$

Εδώ ισχύει

$$S_{XY}(\omega) = S_{YX}^*(\omega)$$

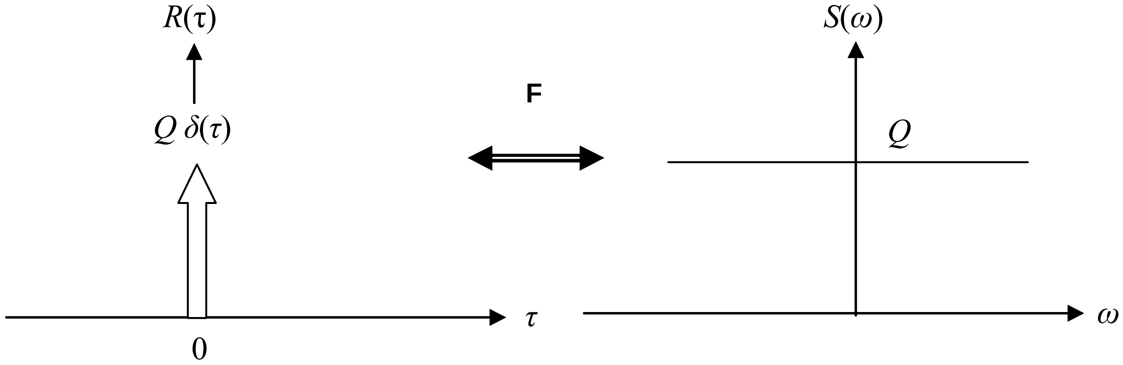
επειδή $R_{XY}(\tau) = R_{YX}^*(-\tau)$ και η S_{XY} μπορεί να είναι μιγαδική έστω κι' αν η X_t είναι πραγματική.

Παράδειγμα 2.1

Για τη διαδικασία λευκού θορύβου

$$S(\omega) = \int_{-\infty}^{\infty} Q\delta(\tau)e^{-i\omega\tau}d\tau = Qe^{i\omega 0} = Q$$

αφού λόγω αμεταβλητότητας $Q(t) \equiv Q$. Με άλλα λόγια το φάσμα είναι σταθερό για όλα τα ω. Επειδή το λευκό φως περιέχει όλες τις συχνότητες αυτή η διαδικασία ονομάστηκε λευκή.



*

2.5 Διαδικασία Poisson

Διάφορα φυσικά φαινόμενα μπορούν να μοντελοποιηθούν με τη διαδικασία Poisson όπως τ' ακόλουθα:

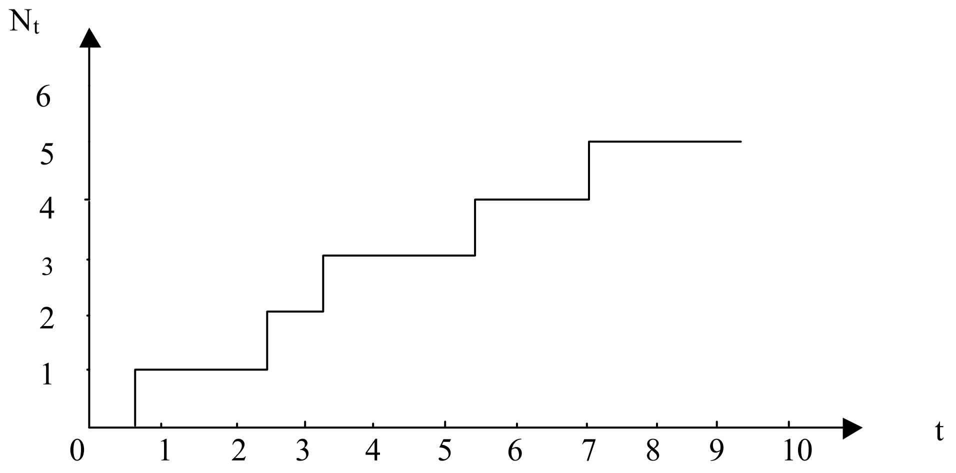
α. Θεωρήστε ένα σύστημα παραγωγής που συνίσταται από έναν αριθμό μηχανών που χαλούν ή επισκευάζονται τυχαία.

β. Σ' ένα σύστημα υπάρχουν εξυπηρετούντες και εξυπηρετούμενοι όπως π.χ. υπάλληλοι σε τράπεζα, αεροδρόμιο, σούπερ μάρκετ, δημόσια υπηρεσία (αν και στην Ελλάδα δεν γίνεται λόγος για εξυπηρέτηση), μηχανισμοί που προωθούν τηλεφωνικές κλήσεις ή δεδομένα, και από την άλλη πελάτες που αφικνούνται τυχαία, ή τηλεφωνικές κλήσεις ή δεδομένα σ' έναν υπολογιστή.

γ. Ένα σωματίδιο-β φθάνει σ' ένα μετρητή σπινθηρισμού και τον διεγείρει.

δ. Ένα ηλεκτρόνιο εκπέμπεται από την κάθοδο λυχνίας Brown σ' έναν παλμογράφο.

Έστω ένα πείραμα τύχης στο οποίο κάποιο γεγονός συμβαίνει κατά διαστήματα. Αρχίζουμε με τη **συνάρτηση μέτρησης** N_t, $t \geq 0$ που αντιπροσωπεύει τον αριθμό τυχαίων συμβάντων στην περίοδο $[0, t]$.



Ορισμός 2.7

Η διαδικασία $\{N_t,\ t \geq 0\}$ είναι διαδικασία ακεραίων τιμών. Μια τέτοια διαδικασία λέγεται **Poisson** με μέσο ρυθμό λ αν:

1. Η $\{N_t, t \geq 0\}$ έχει ανεξάρτητα και αμετάβλητα τμήματα με $P[N(0) = 0] = 1$.
2. Για δύο χρονικές στιγμές t και τ, $t > \tau$ η τυχαία μεταβλητή $N_t - N_\tau$ έχει κατανομή Poisson με μέση τιμή $\lambda(t-\tau)$ ήτοι:

$$P\,[N_t - N_\tau = k] = e^{-\lambda(t-\tau)}\,\frac{[\lambda(t-\tau)]^k}{k!}, \qquad\qquad k = 0,1,2,...$$

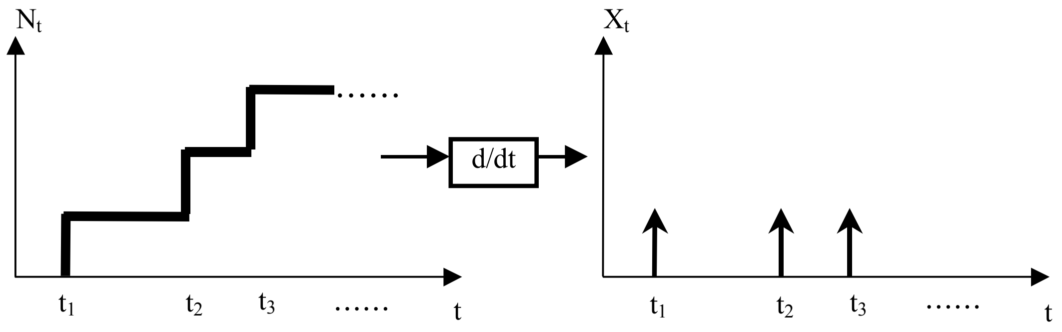
$$E[N_t - N_\tau] = \lambda(t-\tau),$$

$$Var[N_t - N_\tau] = \lambda(t-\tau).$$

Ο αριθμός λ αντιπροσωπεύει το **μέσο ρυθμό** τυχαίων συμβάντων, π.χ. το μέσο αριθμό πρώτων υλών ανά μονάδα χρόνου, που φθάνουν σε μια μηχανή για να υποστούν κατεργασία.

Παράδειγμα 2.2

Έστω ότι μια διαδικασία Poisson N_t είναι είσοδος σ' ένα διαφοριστή. Τότε όπως φαίνεται και στο σχήμα στην έξοδο X_t παίρνουμε



$$ή \ X_t = \sum_i \delta(t - t_i) = \frac{dN_t}{dt} = \dot{N}_t \ .$$

Επειδή $E(N_t) = \lambda t$ τότε

$$E(X_t) = \frac{d(\lambda t)}{dt} = \lambda \ .$$

Η δεύτερη ροπή της N_t είναι

$$E(N_t^2) = Var(N_t) + E^2(N_t) = \lambda t + \lambda^2 t^2 \ .$$

Ενώ η αυτοσυσχέτιση είναι ($t_1 \leq t_2$)

$$R(t_1, t_2) = E(N_{t_1} N_{t_2}) = E[N_{t_1}(N_{t_1} + N_{t_2} - N_{t_1})]$$

$$= E(N_{t_1}^2) + E[N_{t_1}(N_{t_2} - N_{t_1})] \ .$$

Αλλά εξ υποθέσεως η N_{t_1} είναι ανεξάρτητη από την $N_{t_2} - N_{t_1}$ αφού $t_1 \leq t_2$, άρα

$$R(t_1, t_2) = E(N_{t_1}^2) + E(N_{t_1}) E(N_{t_2} - N_{t_1})$$

$$= \lambda t_1 + \lambda^2 t_1^2 + \lambda t_1 \lambda (t_2 - t_1)$$

$$= \lambda t_1 + \lambda^2 t_1 t_2$$

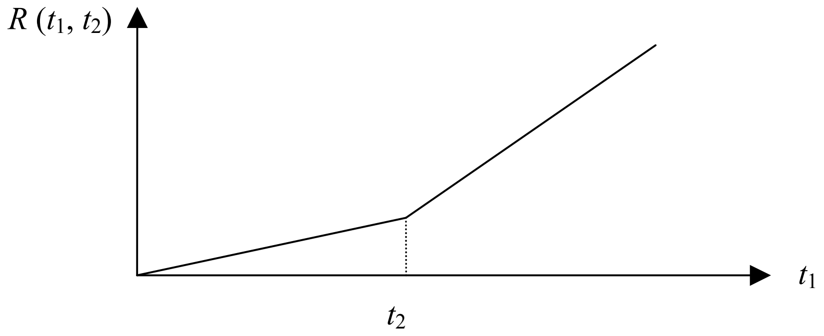
και εν γένει:

$$R(t_1, t_2) = \begin{cases} \lambda t_1 + \lambda^2 t_1 t_2 \ , & t_2 \geq t_1 \\ \lambda t_2 + \lambda^2 t_1 t_2 \ , & t_1 \geq t_2 \end{cases}$$

ή

$$R(t_1,t_2) = \lambda^2 t_1 t_2 + \lambda \min(t_1,t_2)$$

για οποιαδήποτε t_1 και t_2.

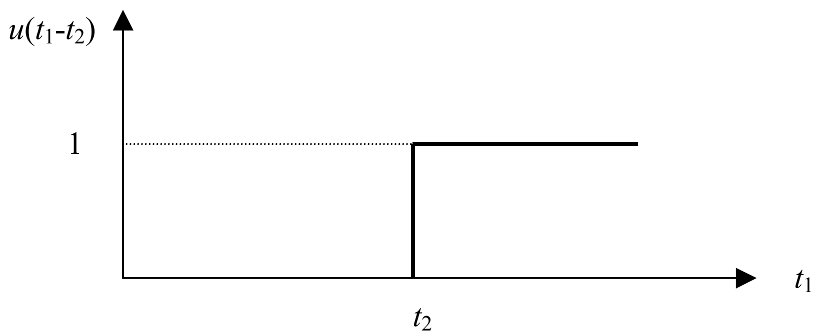


Έστω τώρα ότι ο διαφοριστής διαφορίζει ως προς t_2. Τότε

$$R_{NX}(t_1,t_2) \triangleq R_{N\dot{N}}(t_1,t_2) = \frac{\partial R(t_1,t_2)}{\partial t_2} = \lambda^2 t_1 + \lambda u(t_1 - t_2)$$

αφού

$$\min(\mathrm{t}_1, \mathrm{t}_2) = \begin{cases} t_1, & t_1 < t_2 \\ t_2, & t_1 \geq t_2 \end{cases} \quad \text{και άρα} \quad \frac{\partial \min(t_1,t_2)}{\partial t_2} = \begin{cases} 0, & t_1 < t_2 \\ 1, & t_1 \geq t_2 \end{cases} = u(t_1 - t_2).$$



Επίσης

$$R_{XX}(t_1,t_2) \triangleq R_{\dot{N}\dot{N}}(t_1,t_2) = \frac{\partial R_{N\dot{N}}(t_1,t_2)}{\partial t_1} = \lambda^2 + \lambda \delta(t_1 - t_2)$$

ή

$$R_{XX}(\tau) = \lambda^2 + \lambda \delta(\tau).$$

Είναι γνωστό από την ανάλυση Fourier ότι

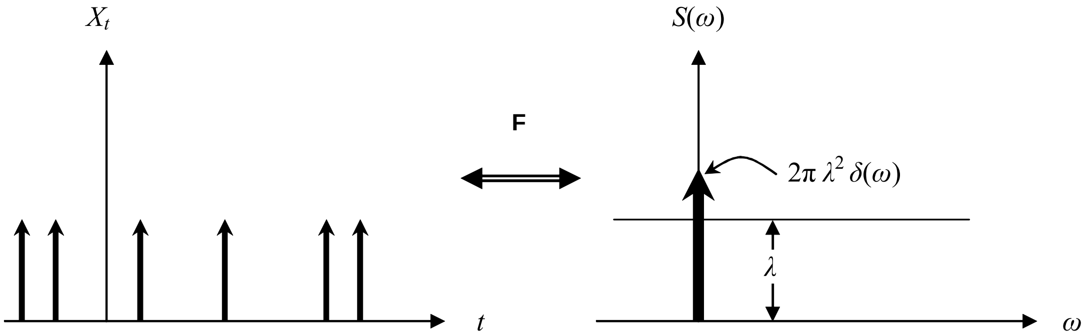
$$\delta(t) \overset{F}{\leftrightarrow} 1$$

$$1 \overset{F}{\leftrightarrow} 2\pi\delta(\omega)$$

Άρα

$$S(\omega) = F[R_{XX}(\tau)] = 2\pi\lambda^2\delta(\omega) + \lambda$$

και σχηματικά



III ΓΡΑΜΜΙΚΑ ΣΥΣΤΗΜΑΤΑ – ΕΡΓΟΔΙΚΟΤΗΣ

3.1 Συστήματα με στοχαστικές εισόδους

Σύστημα είναι μια οντότητα που περιγράφεται από μια συνάρτηση εισόδου X_t ή είσοδο, μια συνάρτηση εξόδου Y_t ή έξοδο κι' έναν τελεστή S έτσι ώστε:

$$Y_t = S\,[X_t]$$

$$X_t \longrightarrow \boxed{\;S\;} \longrightarrow Y_t$$

Έστω ότι $X_t \equiv X(t,\omega)$ είναι μια στοχαστική διαδικασία και ο τελεστής S μπορεί να εξαρτάται από τον χρόνο t αλλά όχι από τα εξαγόμενα ω, τότε το σύστημα λέγεται **αιτιοκρατικό** και η αντίστοιχη έξοδος είναι

$$Y_t \equiv Y(t,\omega) = S\,[X(t,\omega)].$$

Στη συνέχεια θα εξετάσουμε τον προσδιορισμό της στατιστικής της διαδικασίας Y_t γνωρίζοντας τη στατιστική της X_t. Εξετάζουμε πρώτα συστήματα χωρίς μνήμη, ήτοι συστήματα όπου η διαδικασία Y_{t_0} εξαρτάται μόνο από τη διαδικασία X_{t_0} και όχι από τιμές της για $t < t_0$ ή $t > t_0$, με άλλα λόγια το παρελθόν ή το μέλλον της X_t δεν επιδρούν στην Y_t για $t = t_0$. Προφανώς

$$E(Y_t) = E[S(X_t)] = \int_{-\infty}^{\infty} S(x) f(x,t)\,dx$$

όπου f είναι η πυκνότητα της X_t . Επίσης

$$R_{YY}(t_1,t_2) = E\left(Y_{t_1} Y_{t_2}\right) = \int_{-\infty}^{\infty}\int_{-\infty}^{\infty} S(x_1) S(x_2) f(x_1,x_2;t_1,t_2)\,dx_1 dx_2 .$$

Έστω τώρα ότι το σύστημα είναι **γραμμικό**. Θεωρήστε δύο εισόδους $X_1(t)$ και $X_2(t)$ γι' αυτό το σύστημα, τότε η γραμμικότης ορίζεται ως εξής:

$$S\,[C_1 X_1(t) + C_2 X_2(t)] = C_1 S\,[X_1(t)] + C_2 S\,[X_2(t)], \text{ για κάθε } t.$$

όπου C_1 και C_2 είναι σταθερές ή τυχαίες μεταβλητές.

Εξετάζουμε τώρα συστήματα με μνήμη ή χωρίς μνήμη, τα οποία είναι **χρονικά αμετάβλητα**. Από τη θεωρία γραμμικών συστημάτων είναι γνωστά τα ακόλουθα.

1. Το σύστημα λέγεται χρονικά αμετάβλητο εάν
$$\{X(t) \to Y(t)\} \;\Rightarrow\; \{X(t+c) \to Y(t+c)\}.$$
2. Αν η είσοδος είναι το δέλτα του Dirac $\delta(t)$, τότε η έξοδος είναι $h(t)$ που λέγεται **απόκριση παλμού**.

$$\delta(t) \longrightarrow \boxed{\;S\;} \longrightarrow h(t)$$

Η συνάρτηση $h(t)$ είναι πολύ σημαντική και αυτό φαίνεται στη συνέχεια.

3. Κάθε απόκριση σε μια τυχούσα είσοδο $X(t)$ ισούται με τη **συνέλιξη** (convolution) της $X(t)$ με την $h(t)$:
$$Y(t) = X(t) \otimes h(t) \equiv \int_{-\infty}^{\infty} X(t-\tau)h(\tau)d\tau = \int_{-\infty}^{\infty} h(t-\tau)X(\tau)d\tau.$$
4. Αν $\left\{\begin{array}{l} \mathbf{F}[X(t)] = X(\omega) \\ \mathbf{F}[Y(t)] = Y(\omega) \\ \mathbf{F}[h(t)] = H(\omega) \end{array}\right\}$ είναι οι μετασχηματισμοί Fourier, τότε $Y(\omega) = H(\omega)\, X(\omega)$.

Η $H(\omega)$ λέγεται **συνάρτηση μεταφοράς** του συστήματος. Βλέπουμε ότι $\otimes \overset{\mathbf{F}}{\leftrightarrow} \times$, δηλαδή η συνέλλιξη αντιστοιχεί σε πολλαπλασιασμό με το μετασχηματισμό Fourier.

Μετά από αυτά τα προκαταρκτικά διατυπώνουμε το **Θεμελιώδες θεώρημα**:

Θεώρημα 3.1

Οι γραμμικοί τελεστές E και S εναλλάσσονται:

$$E\{S[X(t)]\} = S\{E[X(t)]\}$$

ή αλλιώς

$$\overline{Y} = S\left(\overline{X}\right).$$

Τώρα υπολογίζουμε τις εξής ποσότητες:

1. $R_{XY}(\tau) = E[X_{t+\tau}\, Y_t^*] = \int_{-\infty}^{\infty} E[X_{t+\tau} X_{t-\lambda}^*]\, h^*(\lambda)\, d\lambda$

Αλλά $E[X_{t+\tau}\, X^*_{t-\lambda}] = R_{XX}(\tau + \lambda)$ οπότε

$$R_{XY}(\tau) = \int_{-\infty}^{\infty} R_{XX}(\tau+\lambda) h^*(\lambda) d\lambda$$

Ορίζοντας μεταβλητή $\mu \triangleq -\lambda$ έχουμε $d\lambda = -d\mu$ και $R_{XY}(\tau) = \int_{-\infty}^{\infty} R_{XX}(\tau-\mu) h^*(-\mu) d\mu$.

Η έκφραση αυτή είναι παραλλαγή με εκείνης που είδαμε στον ορισμό της συνέλλιξης. Πράγματι, ορίζοντας τη συνάρτηση $F(\mu) \triangleq h(-\mu)$ έχουμε

$$R_{XY}(\tau) = \int_{-\infty}^{\infty} R_{XX}(\tau-\mu) F^*(\mu) d\mu = R_{XX}(\tau) \otimes F^*(\tau) = R_{XX}(\tau) \otimes h^*(-\tau) \qquad (1)$$

2. $R_{YY}(\tau) = E[Y_t Y_{t-\tau}^*] = \int_{-\infty}^{\infty} E[X_{t-\lambda} Y_{t-\tau}^*] h(\lambda) d\lambda$

$$= \int_{-\infty}^{\infty} R_{XY}(\tau-\lambda) h(\lambda) d\lambda = R_{XY}(\tau) \otimes h(\tau). \qquad (2)$$

3. Αν πάρουμε μετασχηματισμούς Fourier και στις δύο πλευρές της (2) ευρίσκουμε:

$$S_{YY}(\omega) = S_{XY}(\omega)\, H(\omega)$$

4. Ομοίως από την (1) και επειδή $h(\tau) = h(-\tau)$ και $h^*(-\tau) \overset{F}{\leftrightarrow} H^*(\omega)$ ευρίσκουμε:

$$S_{XY}(\omega) = S_{XX}(\omega)\, H^*(\omega)$$

5. Συνδυάζοντας τις ιδιότητες 3 και 4 ευρίσκουμε το άλλο **Θεμελιώδες θεώρημα** των γραμμικών συστημάτων:

Θεώρημα 3.2

$$S_{YY}(\omega) = S_{XX}(\omega)\, H(\omega)\, H^*(\omega) = S_{XX}(\omega)\, |H(\omega)|^2.$$

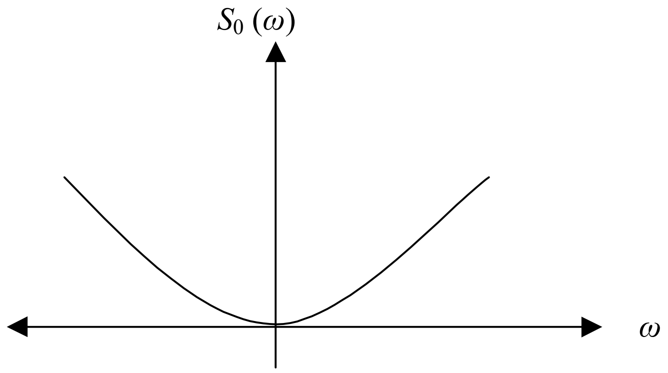
Παράδειγμα 3.1 *Προέμφαση - Αποέμφαση σε συστήματα FM.*

Έστω ότι για εκπομπή FM χρησιμοποιείται το λεγόμενο φέρον σήμα $A\cos\omega_0 t$. Σ' αυτό προστίθεται ο θόρυβος $n(t)$ ώστε $\upsilon(t) = A\cos\omega_0 t + n(t)$. Αν υποθέσουμε ότι η $n(t)$ είναι λευκή διαδικασία, τότε μετά την αποδιαμόρφωση του σήματος ο θόρυβος έχει φασματική πυκνότητα:

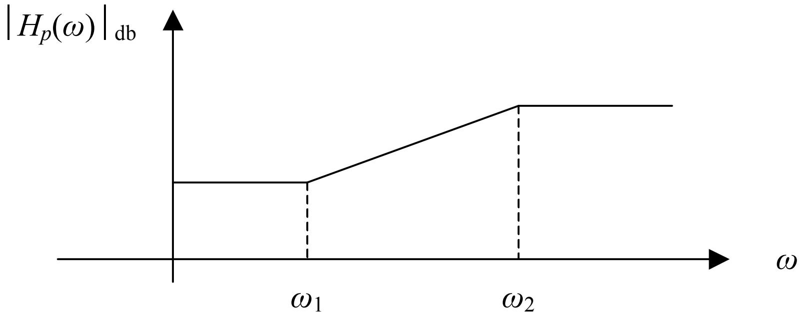
$$S_0(\omega) = \frac{n_0\omega^2}{A^2}$$

όπου ω είναι η συχνότητα και

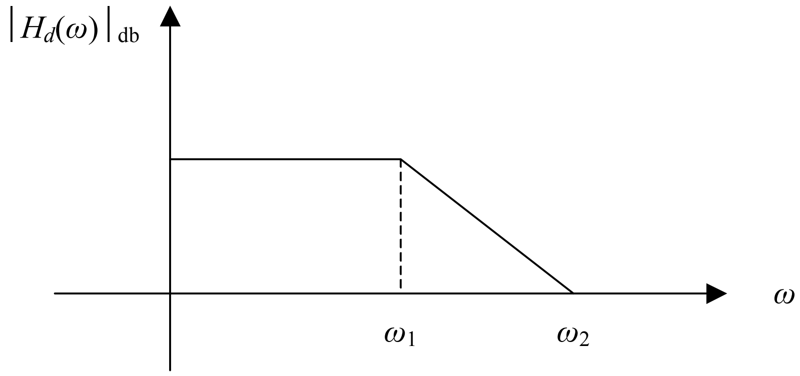
$$R_n(\tau) = n_0\delta(\tau).$$



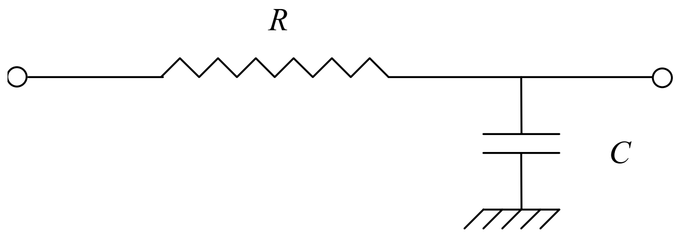
Βλέπουμε ότι σε υψηλές συχνότητες ο θόρυβος είναι πολύ ισχυρός και άρα το σήμα δεν είναι HiFi. Για την επανόρθωση της πιστότητας χρησιμοποιούμε στον πομπό έναν ενισχυτή που ενισχύει το σήμα **πριν** το ενοχλήσει ο θόρυβος ως εξής:



Στην έξοδο του αποδιαμορφωτή χρησιμοποιείται ένας ενισχυτής με απόκριση:



Έτσι το μεν σήμα παραμένει ανέπαφο ενώ ο θόρυβος αποενισχύεται σε υψηλές συχνότητες. Αυτό το σύστημα κάνει τη λεγόμενη **προέμφαση** και **αποέμφαση**. Ένα σύστημα με συνάρτηση μεταφοράς H_d είναι το ακόλουθο κύκλωμα ολοκλήρωσης:

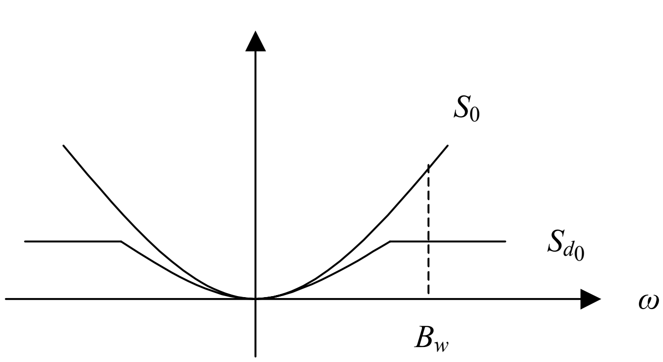


με $RC=1/\omega_1$ καισυνάρτηση μεταφοράς

$$H_d(\omega) = \frac{1}{1 + j\dfrac{\omega}{\omega_1}}.$$

Ο θόρυβος στην έξοδο αυτού του κυκλώματος έχει φάσμα

$$S_{d_0}(\omega) = S_0(\omega)\left|H_d(\omega)\right|^2 = \frac{n_0\omega^2}{A^2}\frac{1}{1 + \left(\dfrac{\omega}{\omega_1}\right)^2}.$$



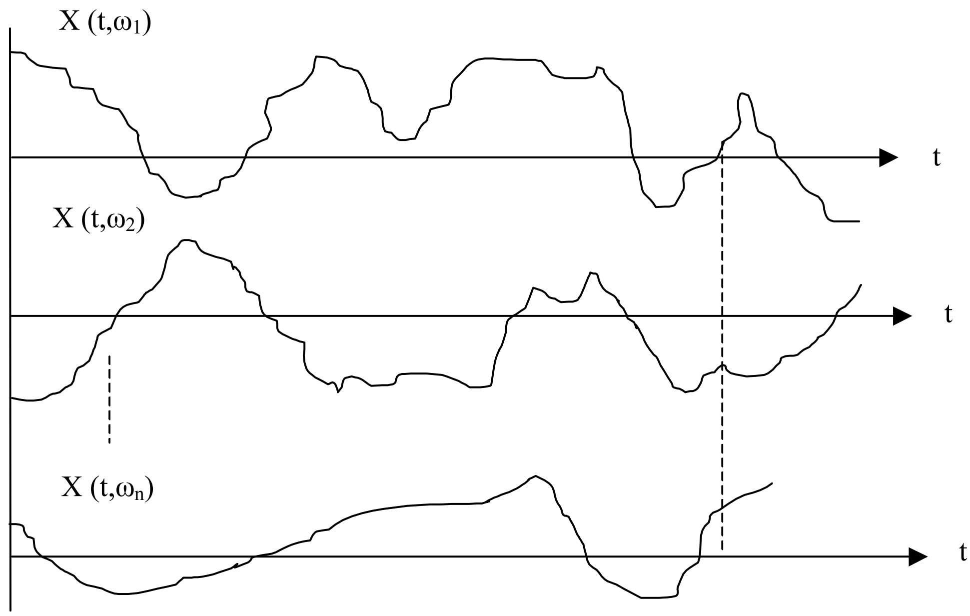
Βλέπουμε τώρα ότι έχουμε σημαντική βελτίωση όπως φαίνεται και στο σχήμα, για τούτο και τα FM χρησιμοποιούνται στην εκπομπή στερεοφωνικών HiFi σημάτων. Η μέση ισχύς του προκύπτοντος θορύβου είναι:

$$N_0 = 2\int_0^{B_w} S_{d_0}\, d\omega$$

που τώρα είναι μηδαμινή. B_w είναι το φασματικό εύρος του δέκτη.

3.2 Εργοδικότητης

Το πρόβλημα τώρα είναι ο προσδιορισμός της στατιστικής μιας διαδικασίας πειραματικά. Έστω φερ' ειπείν η διαδικασία $\{X_t,\ t\in T\}$ και θέλουμε να μάθουμε την προσδοκητή τιμή της, $\overline{X}(t)$. Τότε παρατηρούμε ένα μεγάλο αριθμό καμπυλών της X_t όπως στο σχήμα και εκτιμούμε το $\overline{X}(t)$ ως εξής:



$$\hat{\overline{X}}(t) = \frac{1}{n}\sum_{i=1}^{n} X(t,\omega_i) \tag{3.1}$$

όπου ο συμβολισμός $(\hat{.})$ υποδηλοί εκτίμηση του (.).

Αν τώρα έχουμε στη διάθεσή μας μόνο την $X(t,\omega_k)$ για κάποιο k μπορούμε να υπολογίσουμε τη **χρονική μέση τιμή**:

$$\overline{X} = \lim_{T\to\infty} \overline{X}(T) \triangleq \lim_{T\to\infty} \frac{1}{2T}\int_{-T}^{T} X(t,\omega_k)dt\,. \tag{3.2}$$

Το ερώτημα είναι: «Μπορούμε να υποθέσουμε $\overline{X}(t)=\overline{X}$ και αν ναι, κάτω από ποιές συνθήκες;» Βεβαίως το ίδιο ερώτημα μπορεί να τεθεί για την αυτοσυσχέτιση, την κατανομή κ.ο.κ της διαδικασίας. Αυτό είναι το πρόβλημα της εργοδικότητας.

Ορισμός 3.1

Μια διαδικασία $\{X_t,\ t \in T\}$ λέγεται **εργοδική** κατά μία δεδομένη έννοια (προσδοκητή τιμή κ.λπ.) όταν η μέση τιμή της οικογένειας καμπυλών ισούται με τη χρονική μέση τιμή κατ' αυτή την έννοια.

Το όριο της (3.2) υπάρχει σχεδόν για κάθε ω_k εάν η X_t είναι αμετάβλητη και $E(|X_t|) < \infty$.

3.2.1 Εργοδικότης κατά προσδοκητή τιμή

Έστω η διαδικασία $\{X_t,\ t\in T\}$ και $E(X_t)=\overline{X}$ σταθερά. Επίσης

$$\overline{X}(T) = \frac{1}{2T}\int_{-T}^{T} X_t dt\,.$$

Η διαδικασία είναι **εργοδική κατά μέση τιμή** αν

$$\overline{X} = \lim_{T\to\infty} \overline{X}(T) \qquad \text{μ.π. } 1 \tag{3.3}$$

Προφανώς η ποσότητα $\overline{X}(T)$ είναι μια τυχαία μεταβλητή με προσδοκητή τιμή

$$E\left[\overline{X}(T)\right] = \frac{1}{2T}\int_{-T}^{T} E(X_t)dt = \overline{X}$$

και διασπορά

$$Var\left[\overline{X}(T)\right] = \sigma_T^2 .$$

Η (3.3) θα ισχύει τότε αν $\lim_{T\to\infty} \sigma_T^2 = 0$.

Υποθέτουμε πάντα ότι η X_t είναι ολοκληρώσιμη κατά μέσο τετράγωνο και έχει γνωστή αυτοσυμμεταβλητότητα $C(t_1, t_2)$. Από το λογισμό μέσου τετραγώνου

$$\sigma_T^2 = \frac{1}{4T^2} \int_{-T}^{T}\int_{-T}^{T} C(t_1, t_2) dt_1 dt_2 .$$

Άρα η X_t είναι εργοδική κατά προσδοκητή τιμή τότε και μόνον τότε αν

$$\lim_{T\to\infty} \frac{1}{4T^2} \int_{-T}^{T}\int_{-T}^{T} C(t_1, t_2) dt_1 dt_2 = 0 .$$

Παράδειγμα 3.2

Έστω ότι η X_t είναι λευκή διαδικασία. Τότε

$$\lim_{T\to\infty} \frac{1}{4T^2} \int_{-T}^{T}\int_{-T}^{T} q(t_1)\delta(t_1 - t_2) dt_1 dt_2 = \lim_{T\to\infty} \frac{1}{4T^2} \int_{-T}^{T} q(t_2) dt_2 .$$

Το όριο αυτό ισούται με μηδέν τότε και μόνον τότε αν η $q(t)$ είναι φραγμένη ή $q(t) \le M$, $\forall\ t$.

Στη συνέχεια αποδεικνύουμε ότι αν η διαδικασία είναι αμετάβλητη κατά την ευρεία έννοια, δηλαδή

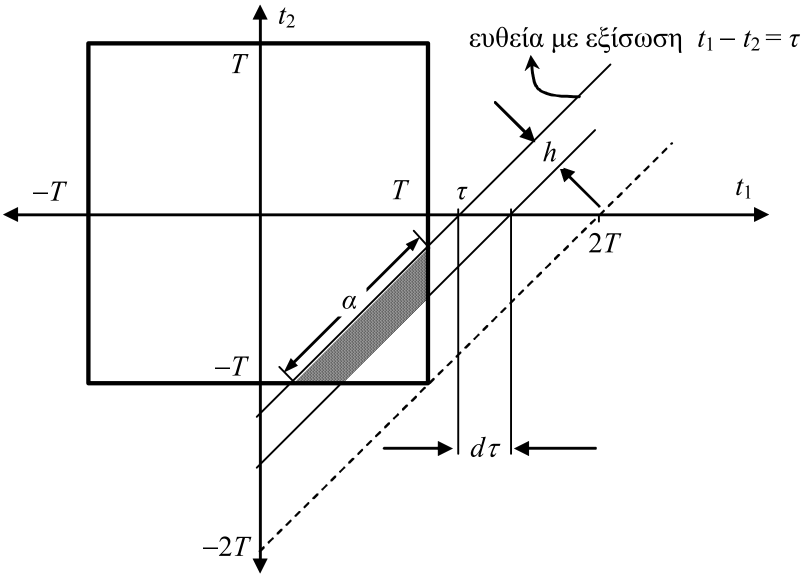
$$C(t_1, t_2) = C(t_1 - t_2),$$

τότε

$$\frac{1}{4T^2} \int_{-T}^{T}\int_{-T}^{T} C(t_1 - t_2) dt_1 dt_2 = \frac{1}{4T^2} \int_{-2T}^{2T} \left(2T - |\tau|\right) C(\tau) d\tau \qquad (3.4)$$

όπου $t_1 - t_2 \stackrel{\Delta}{=} \tau$. Το διπλό ολοκλήρωμα ισούται με το απλό ολοκλήρωμα της $C(t_1 - t_2)$

επάνω σε λωρίδες όπως αυτή του επόμενου σχήματος.



Το εμβαδόν αυτής της απειροστής λωρίδας είναι

$$ah = \sqrt{2}\left(2T - |\tau|\right)\frac{d\tau}{\sqrt{2}}.$$

Επίσης κατά μήκος αυτής της λωρίδας έχουμε $C(t_1 - t_2) =$ σταθερό και από αυτό επαληθεύεται η (3.4).

Από το ολοκλήρωμα (3.4) ευρίσκουμε ότι η διαδικασία είναι εργοδική κατά προσδοκητή τιμή τότε και μόνον τότε αν

$$\lim_{T\to\infty}\frac{1}{2T}\int_{-2T}^{2T} C(\tau)\left(1 - \frac{|\tau|}{2T}\right)d\tau = 0.$$

Προφανώς

$$\frac{1}{2T}\int_{-2T}^{2T} C(\tau)\left(1 - \frac{|\tau|}{2T}\right)d\tau \le \frac{1}{2T}\int_{-2T}^{2T} |C(\tau)|\left|1 - \frac{|\tau|}{2T}\right|d\tau < \frac{1}{2T}\int_{-2T}^{2T}|C(\tau)|\,d\tau.$$

Άρα αν

$$\int_{-\infty}^{\infty}|C(\tau)|\,d\tau < \infty$$

τότε η διαδικασία είναι πάλι εργοδική κατά προσδοκητή τιμή.

Παράδειγμα 3.3

Έστω $X_t = r$, όπου $r = r(\omega)$ είναι μια τυχαία μεταβλητή. Τότε

$$\overline{X}_T = \frac{1}{2T}\int_{-T}^{T} r(\omega)\,dt = r(\omega)$$

δηλαδή η χρονική μέση τιμή είναι τυχαία και άρα μη σταθερή, που σημαίνει ότι η διαδικασία είναι μη εργοδική.

3.2.2 Εργοδικότης κατά κατανομή

Η διαδικασία είναι αμετάβλητη υπό τη στενή έννοια και ζητούμε να προσδιορίσουμε τη συνάρτηση κατανομής από ένα χρονικό δείγμα

$$F(x) = P(X_t \le x).$$

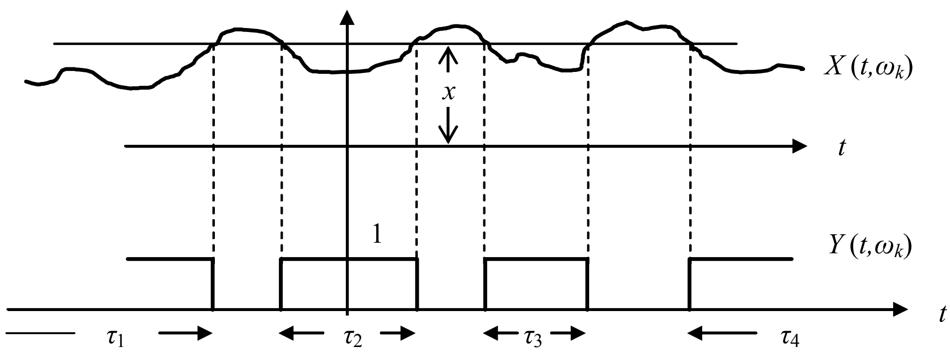
Θεωρούμε τη λεγόμενη **ενδεικτική διαδικασία** Y_t

$$Y_t = \begin{cases} 1, \text{ αν } X_t \le x \\ 0, \text{ αν } X_t > x \end{cases}$$

Προφανώς $E(Y_t) = 1 \times P(X_t \le x) + 0 \times P(X_t > x) = F(x)$. Δηλαδή η προσδοκητή τιμή της Y_t είναι η ζητούμενη συνάρτηση $F(x)$. Από το σχήμα βλέπουμε ότι:

$$\overline{Y}(T) \triangleq \frac{1}{2T}\int_{-T}^{T} Y_t dt = \frac{\tau_1 + \tau_2 + \ldots \tau_n}{2T}$$

όπου τ_i είναι τα χρονικά διαστήματα κατά τη διάρκεια των οποίων η $X_t \le x$ ή η $Y_t = 1$.



Τώρα

$$Y_{t+\tau}Y_t = \begin{cases} 1 \text{ αν } X_{t+\tau} \le x \text{ και } X_t \le x \\ 0 \text{ άλλως} \end{cases}$$

οπότε $E(Y_{t+\tau}Y_t) = 1 \times P(X_{t+\tau} \le x \quad \cap \quad X_t \le x) = F(x,x;\tau)$, ή

$$R(\tau) = F(x,x;\tau)\,.$$

Αλλά

$$C(\tau) = R(\tau) - \overline{Y}^2 = F(x,x;\tau) - F^2(x)\,.$$

Τώρα αφού έχουμε ανάγκη τον υπολογισμό της $F(x)$ στον υπολογισμό της $E(Y_t)$ μπορούμε να χρησιμοποιήσουμε το θεώρημα της προηγούμενης παραγράφου και άρα:

Η διαδικασία είναι **εργοδική κατά κατανομή** τότε και μόνο τότε αν

$$\lim_{T\to\infty} \frac{1}{2T} \int_{-2T}^{2T} \left[F(x,x;\tau) - F^2(x)\right]\left(1 - \frac{|\tau|}{2T}\right) d\tau = 0\,.$$

Στην πράξη ευρίσκομε την $F(x)$ ως εξής:

$$F(x) \approx \overline{Y}(T) = \frac{\tau_1 + \tau_2 + \ldots \tau_n}{2T}$$

ενώ τη συνάρτηση πυκνότητας την υπολογίζομε ως εξής:

Έστω $\Delta\tau_i$ τα χρονικά τμήματα για τα οποία η X_t ευρίσκεται ανάμεσα από το x και $x + \Delta x$, τότε:

$$f(x)\Delta x \approx F(x + \Delta x) - F(x) \approx \frac{\Delta\tau_1 + \Delta\tau_2 + \ldots + \Delta\tau_n}{2T}\,.$$

3.2.3 Εργοδικότης κατά συσχέτιση

Το πρόβλημα είναι ο υπολογισμός της συνάρτησης $R(\tau) = E(X_{t+\tau}X_t)$ από ένα χρονικό δείγμα. Πάλι ανάγουμε το πρόβλημα στον υπολογισμό μιας προσδοκητής τιμής κάποιας διαδικασίας με γνωστή αυτοσυμμεταβλητότητα.

Ορίζουμε μια νέα διαδικασία $Y_t(\tau) = X_{t+\tau} X_t$. Η διαδικασία X_t είναι **εργοδική κατά συσχέτιση** αν:

$$\lim_{T\to\infty} \frac{1}{2T} \int_{-T}^{T} X_{t+\tau} X_t dt = E[Y_t(\tau)] \stackrel{\Delta}{=} R(\tau).$$

Η συσχέτιση της $Y_t(\tau)$ είναι

$$R_{YY}(\mu) = E[(X_{t+\tau+\mu} X_{t+\mu})(X_{t+\tau} X_t)]$$

και

$$C_{YY}(\mu) = R_{YY}(\mu) - R^2(\tau).$$

Τέλος η διαδικασία είναι εργοδική κατά συσχέτιση **τότε και μόνο τότε αν**:

$$\lim_{T\to\infty} \frac{1}{2T} \int_{-2T}^{2T} \left[R_{YY}(\mu) - R^2(\tau)\right]\left(1 - \frac{|\mu|}{2T}\right) d\mu = 0.$$

Βεβαίως, το πάρα πάνω όριο δεν είναι εύκολο να υπολογισθεί στην πράξη. Αν όμως η X_t είναι **κανονική** τότε μπορεί ν' αποδειχθεί ότι:

$$C_{YY}(\mu) = R(\mu+\tau)R(\mu-\tau) + R^2(\mu)$$

και εάν $R(\mu) \to 0$ όταν $\mu \to \infty$, τότε $C_{YY}(\mu) \to 0$ οπότε η X_t είναι εργοδική κατά συσχέτιση. Στην πράξη για την εύρεση της προσδοκητής τιμής, της κατανομής, της συσχέτισης, του φάσματος κ.λπ. χρησιμοποιούνται ειδικά αναλογικά ή ψηφιακά κυκλώματα.

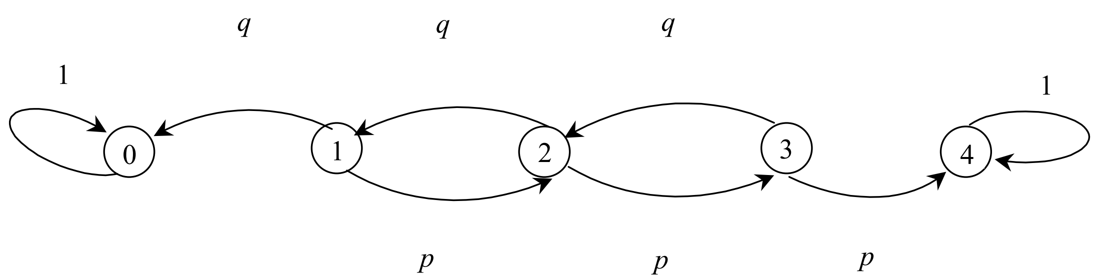
IV ΑΛΥΣΙΔΕΣ MARKOV

4.1 Γενικά περί αλυσίδων Markov

Οι αλυσίδες Markov είναι στοχαστικές διαδικασίες διακεκριμένων καταστάσεων και διακεκριμένου χρόνου.

Παράδειγμα 4.1 *Η καταστροφή του χαρτοπαίκτη*

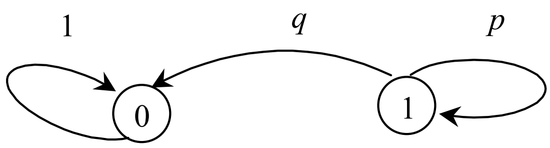
Υπάρχουν δύο χαρτοπαίκτες με δύο μάρκες στο χέρι ο καθένας. Ο παίκτης Α κερδίζει με πιθανότητα p ενώ ο Β με πιθανότητα $q = 1 - p$. Ο κάθε παίκτης βρίσκεται σε μια από τις ακόλουθες πιθανές καταστάσεις 0,1,2,3,4, ανάλογα με το πόσες μάρκες έχει σε κάποια χρονική στιγμή. Σχεδιάζουμε το ακόλουθο διάγραμμα κατάστασης:



Σχήμα 4.1

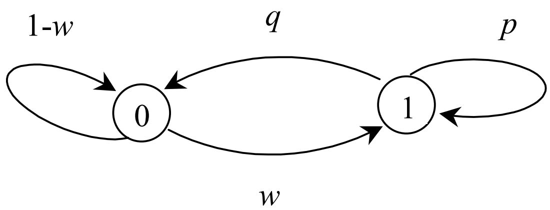
Παράδειγμα 4.2 *Μια μηχανή*

Έστω μια μηχανή που παράγει ένα προϊόν με πιθανότητα p σ' ένα κύκλο παραγωγής ή χαλάσει με πιθανότητα $q = 1 - p$. Έστω 1 η κατάσταση «καλή» και 0 η κατάσταση «χαλασμένη». Τότε το διάγραμμα κατάστασης είναι:



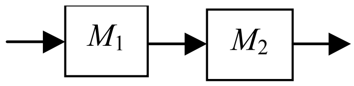
Σχήμα 4.2

Αν τώρα η μηχανή μόλις χαλάσει τίθεται υπό επισκευή κι' έχει πιθανότητα επισκευής w στον επόμενο κύκλο παραγωγής, τότε το διάγραμμα γίνεται:

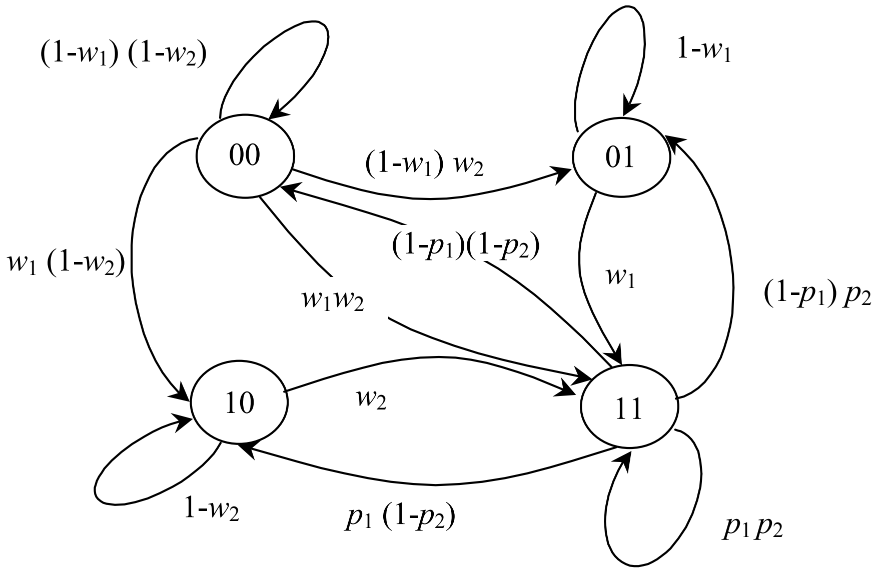


Σχήμα 4.3

Παράδειγμα 4.3 *Δύο μηχανές*



Η πρώτη έχει χαρακτηριστικά p_1, w_1, και η δεύτερη p_2, w_2. Αν η δεύτερη είναι υπό επισκευή και η πρώτη είναι «καλή», τότε αυτή δεν λειτουργεί γιατί δεν έχει που να διοχετεύσει το προϊόν. Την λέμε **αποκλεισμένη** (μπλοκαρισμένη). Αντίστοιχα αν η δεύτερη είναι «καλή» και η πρώτη υπό επισκευή, τότε η δεύτερη δεν λειτουργεί γιατί δεν έχει προϊόντα στην είσοδο ή όπως λέμε είναι **αποστερημένη** (πεινασμένη). Υποθέτουμε ότι αποστερημένη ή αποκλεισμένη μηχανή δεν μπορεί να χαλάσει. Το διάγραμμα είναι:



Σχήμα 4.4

Εδώ έχουμε τις εξής εξισώσεις που περιγράφουν την εξέλιξη των πιθανοτήτων:

$$P_{k+1}(00) = (1- w_1)(1-w_2)P_k(00) + (1- p_1)(1- p_2)P_k(11)$$

$$P_{k+1}(01) = (1- w_1)w_2P_k(00) + (1- w_1)P_k(01) + (1- p_1)p_2P_k(11)$$

$$P_{k+1}(10) = w_1 (1- w_2)P_k(00) + (1- w_2)P_k(10) + p_1(1- p_2)P_k(11)$$

$$P_{k+1}(11) = w_1w_2P_k(00) + w_1P_k(01) + w_2P_k(10) + p_1p_2P_k(11)$$

ή σε διανυσματική μορφή

$$P_{k+1} = AP_k \tag{4.1}$$

ή

$$\begin{bmatrix} P_{k+1}(00) \\ P_{k+1}(01) \\ P_{k+1}(10) \\ P_{k+1}(11) \end{bmatrix} = \begin{bmatrix} (1-w_1)(1-w_2) & 0 & 0 & (1-p_1)(1-p_2) \\ (1-w_1)w_2 & 1-w_1 & 0 & (1-p_1)p_2 \\ w_1(1-w_2) & 0 & 1-w_2 & p_1(1-p_2) \\ w_1w_2 & w_1 & w_2 & p_1p_2 \end{bmatrix} \begin{bmatrix} P_k(00) \\ P_k(01) \\ P_k(10) \\ P_k(11) \end{bmatrix}.$$

Οι διακεκριμένες χρονικές στιγμές είναι 0,1,...,k,... Έστω ότι ξέρουμε την αρχική συνθήκη της εξίσωσης διαφοράς (4.1), για παράδειγμα:

$$P_0 = \begin{bmatrix} P_0(00) \\ P_0(01) \\ P_0(10) \\ P_0(11) \end{bmatrix} = \begin{bmatrix} 0 \\ 0 \\ 0 \\ 1 \end{bmatrix}.$$

Τότε η (4.1) γίνεται

$$P_1 = AP_0, \qquad P_2 = AP_1 = A^2 P_0, \qquad P_3 = AP_2 = A^3 P_0$$

και με την τέλεια επαγωγή

$$P_k = A^k P_0. \tag{4.2}$$

Η (4.2) είναι η λύση της (4.1). Παρατηρούμε ότι οι στήλες του πίνακα A έχουν άθροισμα 1 όπως και τα στοιχεία του διανύσματος P επειδή εκπροσωπούν πιθανότητες.

Γενικά, το στοιχείο a_{ij} του πίνακα A θα συμβολίζει την πιθανότητα μετάβασης **από την κατάσταση *j* προς *i***:

$$\begin{array}{cc} & \text{μετάβαση από κατάσταση } j \\ \text{προς κατάσταση } i & \begin{bmatrix} & \vdots & \\ \cdots & a_{ij} & \cdots \\ & \vdots & \end{bmatrix} \end{array}$$

Ορισμός 4.1

Η αλυσίδα Markov τάξης n χαρακτηρίζεται από ένα σύνολο καταστάσεων $S = \{S_1, S_2,...,S_n\}$ και τις πιθανότητες a_{ij} μετάβασης **από την S_j στην S_i**, $i,j = 1,2,...,n$. Κάθε χρονική στιγμή k η αλυσίδα βρίσκεται σε κάποια κατάσταση $S_i \in S$.

Για το Παράδειγμα 4.3 βλέπουμε ότι:

$$S = \{(00), (01), (10), (11)\} = \{1,2,3,4\}$$

$$a_{11} = (1 - w_1)(1 - w_2)$$

$$a_{21} = (1 - w_1)w_2\,, \text{ κλπ.}$$

Ορισμός 4.2

Ένας πίνακας με μη αρνητικά στοιχεία που το άθροισμά τους κατά μήκος μιας στήλης είναι 1, λέγεται **στοχαστικός**.

Ορισμός 4.3

Ένα διάνυσμα με μη αρνητικά στοιχεία που το άθροισμά τους είναι 1 λέγεται **διάνυσμα πιθανότητας**.

Αν ένας πίνακας A απαρτίζεται από στοιχεία > 0 τότε λέμε ότι ο πίνακας είναι **θετικός** και γράφουμε $A > 0$.

Από το παράδειγμα 4.3 είδαμε ότι μια αλυσίδα Markov περιγράφεται από μια εξίσωση της μορφής: $P_{k+1} = AP_k$.

4.2 Ανάλυση μόνιμης κατάστασης

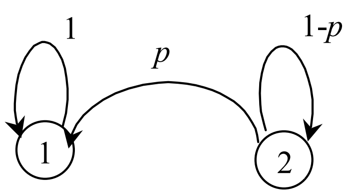
Ορισμός 4.5

Η αλυσίδα Markov λέγεται **κοινή** (regular) αν $A^k > 0$ για κάποιο k.

Ξέρουμε ήδη ότι $P_k = A^k P_0$. Το ότι μια αλυσίδα είναι κοινή, σημαίνει άρα ότι στη λύση όλα τα στοιχεία είναι > 0. Με άλλα λόγια από κάποια χρονική στιγμή k και μετά, όλες οι μεταβάσεις καταστάσεων επιτρέπονται.

Παράδειγμα 4.4

Έστω η αλυσίδα $P_{k+1} = \begin{bmatrix} 1 & p \\ 0 & 1-p \end{bmatrix} P_k$.



Εδώ $\begin{bmatrix} 1 & p \\ 0 & 1-p \end{bmatrix}^k = \begin{bmatrix} 1 & 1-(1-p)^k \\ 0 & (1-p)^k \end{bmatrix}$ και τα στοιχεία του πίνακα **δεν** είναι όλα > 0.

Αυτή η αλυσίδα δεν είναι κοινή. Παρατηρούμε ότι έχουμε 2 καταστάσεις και αν στη χρονική στιγμή 0 ξεκινήσαμε από την κατάσταση 2 έχουμε:

$$P_k = \begin{bmatrix} 1 & 1-(1-p)^k \\ 0 & (1-p)^k \end{bmatrix} \begin{bmatrix} 0 \\ 1 \end{bmatrix} = \begin{bmatrix} 1-(1-p)^k \\ (1-p)^k \end{bmatrix},$$

ενώ αν ξεκινούσαμε από την 1 τότε:

$$P_k = \begin{bmatrix} 1 & 1-(1-p)^k \\ 0 & (1-p)^k \end{bmatrix} \begin{bmatrix} 1 \\ 0 \end{bmatrix} = \begin{bmatrix} 1 \\ 0 \end{bmatrix}.$$

Δηλαδή η μετάβαση $1 \rightarrow 2$ απαγορεύεται κάτι που φαίνεται και από το διάγραμμα κατάστασης.

Θεώρημα 4.1

Κάθε στοχαστικός πίνακας έχει μια ιδιοτιμή $\lambda_0 = 1$ και όλες οι άλλες ιδιοτιμές ικανοποιούν $|\lambda| \leq 1$.

Το ακόλουθο θεώρημα είναι το πιο σημαντικό της θεωρίας αλυσίδων Markov και ισχύει πάντα για κοινές αλυσίδες.

Θεώρημα 4.2

Έστω ο πίνακας A που αντιστοιχεί σε μια κοινή αλυσίδα Markov. Τότε:

1. Υπάρχει ένα μοναδικό **διάνυσμα πιθανότητας** P **τέτοιο ώστε** $P = AP$. Είναι δηλαδή το P το **ιδιοδιάνυσμα** που αντιστοιχεί σε $\lambda = 1$.
2. $P_k = A^k P_0$ και $\lim\limits_{k\to\infty} A^k = \overline{A}$. Ο πίνακας $\overline{A}$ συνίσταται από στήλες που όλες ισούνται με P.
3. Για οποιαδήποτε αρχική συνθήκη, έχουμε:

$$\lim_{k\to\infty} P_k = \overline{A} P_0 = [P \mid P \mid \ldots \mid P] P_0 = P.$$

Η τελευταία ισότητα προκύπτει ως εξής

$$[P|P|...|P]P_0 = \begin{bmatrix} P(1) & P(1) & P(1) & P(1) \\ P(2) & P(2) & P(2) & P(2) \\ ... & ... & ... & ... \\ P(n) & P(n) & P(n) & P(n) \end{bmatrix} \times \begin{bmatrix} P_0(1) \\ P_0(2) \\ ... \\ P_0(n) \end{bmatrix} = \begin{bmatrix} P(1)\times\sum_{i=1}^{n} P_0(i) \\ P(2)\times\sum_{i=1}^{n} P_0(i) \\ ... \\ P(n)\times\sum_{i=1}^{n} P_0(i) \end{bmatrix} = \begin{bmatrix} P(1) \\ P(2) \\ ... \\ P(n) \end{bmatrix} = P,$$

όπου $\sum_{i=1}^{n} P_0(i)=1$ επειδή το P_0 είναι διάνυσμα πιθανότητας. Άρα αν έχουμε μια κοινή αλυσίδα Markov **για να βρούμε την συμπεριφορά** της καθώς $k\to\infty$ ή **τη μόνιμη κατάσταση**, όπως λέμε, ευρίσκουμε το P που αντιστοιχεί σε $\lambda = 1$. Το ίδιο αποτέλεσμα προκύπτει αν λύσουμε το σύστημα εξισώσεων $P=AP$ ως προς $P=[P(1)\ldots P(n)]^{T}$ **αφού πρώτα αντικαταστήσουμε μία οιαδήποτε εξίσωσή** του από την $\sum_{1}^{n} P(i)=1$.

Παράδειγμα 4.5

Επιστρέφουμε στο παράδειγμα 4.2. Εκεί ισχύουν:

$$P_{k+1}(0) = (1-w)P_k(0) + qP_k(1)$$

$$P_{k+1}(1) = wP_k(0) + pP_k(1)$$

όπου $p + q = 1$. Το σύστημα αυτό γράφεται $P_{k+1} = \begin{bmatrix} 1-w & q \\ w & p \end{bmatrix} P_k$. Έστω τώρα ότι $w = \frac{1}{2}$, $p = 0.9$. Τότε $P_{k+1} = \begin{bmatrix} 0.5 & 0.1 \\ 0.5 & 0.9 \end{bmatrix} P_k$ και $P_k = \begin{bmatrix} 0.5 & 0.1 \\ 0.5 & 0.9 \end{bmatrix}^k P_0$. Οι ιδιοτιμές είναι $|\lambda I - A| = \lambda^2 - 1.4\lambda + 0.4 = 0$ ή $\lambda_{1,2} = \begin{cases} 1 \\ 0.4 \end{cases}$. Από το **θεώρημα των Cayley-Hamilton** προκύπτει ότι, για τον πίνακα A με διάσταση 2×2,

$$A^k = \alpha_0 A^0 + \alpha_1 A^1 = \alpha_0 I + \alpha_1 A$$

όπου τα α_0 και α_1 είναι άγνωστοι συντελεστές του πολυωνύμου. Τώρα για κάθε ιδιοτιμή λ του A, το βαθμωτό πολυώνυμο $\alpha_0\lambda^0 + \alpha_1\lambda^1 = \alpha_0 + \alpha_1\lambda$ ισούται με λ^k. Άρα

$$\left.\begin{aligned} 1 &= a_0 + a_1 \\ 0.4^k &= a_0 + 0.4a_1 \end{aligned}\right\} \Rightarrow \begin{cases} a_1 = \dfrac{1-0.4^k}{0.6} \\ a_0 = \dfrac{-0.4+0.4^k}{0.6} \end{cases}$$

Για πίνακα A διαστάσεων $n\times n$ το πολυώνυμο θα ήταν $A^k = \alpha_0 I + \alpha_1 A + \ldots + \alpha_{n-1}A^{n-1}$ και θα είχαμε n άγνωστους συντελεστές. Αντικαθιστώντας τους συντελεστές προκύπτει

$$A^k = \begin{bmatrix} \dfrac{0.1+0.5\times 0.4^k}{0.6} & \dfrac{0.1-0.1\times 0.4^k}{0.6} \\ \dfrac{0.5-0.5\times 0.4^k}{0.6} & \dfrac{0.5+0.1\times 0.4^k}{0.6} \end{bmatrix}.$$

Βλέπουμε ότι

$$\lim_{k\to\infty} A^k = \begin{bmatrix} \dfrac{1}{6} & \dfrac{1}{6} \\ \dfrac{5}{6} & \dfrac{5}{6} \end{bmatrix} = [P \mid P].$$

Επομένως

$$\lim_{k\to\infty} P_k = P = \begin{bmatrix} \dfrac{1}{6} \\ \dfrac{5}{6} \end{bmatrix}$$

και η μόνιμη κατάσταση είναι ανεξάρτητη της αρχικής συνθήκης.

Τώρα επιβεβαιώνουμε το Θεώρημα 4.2. Το ιδιοδιάνυσμα x ($\lambda x = Ax$) που αντιστοιχεί στην ιδιοτιμή $\lambda = 1$ είναι το P. Πράγματι,

$$P = AP \;\text{ ή }\; \begin{bmatrix} P(0) \\ P(1) \end{bmatrix} = \begin{bmatrix} 0.5 & 0.1 \\ 0.5 & 0.9 \end{bmatrix} \times \begin{bmatrix} P(0) \\ P(1) \end{bmatrix}$$

Από την πρώτη εξίσωση του συστήματος προκύπτει

$$P(0) = 0.5P(0) + 0.1P(1) \;\Rightarrow\; P(0) = \frac{1}{5}P(1) \tag{4.3}$$

Η άλλη εξίσωση $P(1) = 0.5\, P(0) + 0.9\, P(1)$ είναι ίδια με την (4.3). Για να βρούμε το μοναδικό ιδιοδιάνυσμα χρειαζόμαστε και την $P(0)+P(1) = 1$, άρα:

$$P(0) = \frac{1}{6} \;\text{ και }\; P(1) = \frac{5}{6}$$

που επιβεβαιώνει την πρώτη πρόταση του θεωρήματος. Εφαρμόζοντας τη δεύτερη πρόταση του Θεωρήματος 4.2 προκύπτει

$$\lim_{k\to\infty} A^k = [P \mid P] = \begin{bmatrix} \frac{1}{6} & \frac{1}{6} \\ \frac{5}{6} & \frac{5}{6} \end{bmatrix}$$

που είναι αυτό που ήδη βρήκαμε.

Έστω τώρα ότι η μηχανή δουλεύει με ρυθμό i προϊόντα / μονάδα χρόνου. Τότε ο μέσος ρυθμός παραγωγής είναι:

$$\bar{i} = iP(1) + 0P(0) = \frac{5}{6} i\,.$$

Αν στην αρχή η κατάσταση είναι 1, τότε τη στιγμή k έχουμε

$$\begin{bmatrix} P_k(0) \\ P_k(1) \end{bmatrix} = A^k \begin{bmatrix} 0 \\ 1 \end{bmatrix} = \begin{bmatrix} \dfrac{0.1 - 0.1 \times 0.4^k}{6} \\ \dfrac{0.5 + 0.1 \times 0.4^k}{6} \end{bmatrix}$$

και ο μέσος ρυθμός παραγωγής κατά τη χρονική στιγμή k είναι:

$$\bar{i}_k = iP_k(1) = i\frac{0.5 + 0.1 \times 0.4^k}{0.6}\,.$$

Στην πράξη εν τούτοις σπάνια μας ενδιαφέρει αυτό το μέγεθος αφού υποθέτουμε ότι το σύστημα έχει δουλέψει γι’ αρκετά μακρό χρόνο για να φθάσει στη μόνιμη κατάσταση.

4.3 Μεταβατικά φαινόμενα

Αν ο πίνακας A έχει μερικά μηδενικά και δεν υπάρχει k για το οποίο $A^k > 0$, τότε η αλυσίδα είναι μη κοινή, όπως είδαμε. Θα λέμε ότι η κατάσταση S_j είναι **προσιτή** από την S_i αν και μόνο αν μπορούμε να φθάσουμε στην S_j από την S_i με πεπερασμένο αριθμό βημάτων. Με άλλα λόγια υπάρχει κάποιο m τέτοιο ώστε το στοιχείο $a_{ij}^{(m)}$ του πίνακα $(A^m)^{\mathrm{T}}$ είναι > 0 .

Λέμε ότι οι καταστάσεις S_i και S_j **επικοινωνούν** αν η κάθε μια είναι προσιτή από την άλλη.

Θεώρημα 4.3

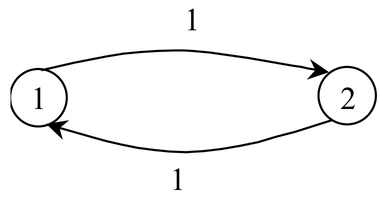
Το σύνολο των καταστάσεων μιας αλυσίδας Markov μπορεί να διαιρεθεί σε **κλάσεις** που επικοινωνούν. Κάθε κατάσταση στην κλάση επικοινωνεί με **κάθε κατάσταση στην ίδια κλάση** αλλά όχι άλλη κατάσταση.

Αν η αλυσίδα αποτελείται από μια μόνο επικοινωνούσα κλάση, τότε λέγεται **μη απλοποιήσιμη** (irreducible), αλλιώς **απλοποιήσιμη** (reducible). Προφανώς **μια κοινή αλυσίδα είναι μη απλοποιήσιμη** αφού όλες οι καταστάσεις επικοινωνούν. Το αντίστροφο δεν ισχύει.

Παράδειγμα 4.6

Έστω $A = \begin{bmatrix} 0 & 1 \\ 1 & 0 \end{bmatrix}$. Εδώ $A^{2k} = \begin{bmatrix} 1 & 0 \\ 0 & 1 \end{bmatrix}$, $A^{2k+1} = \begin{bmatrix} 0 & 1 \\ 1 & 0 \end{bmatrix}$

και άρα η αλυσίδα είναι μη κοινή. Εν τούτοις είναι επικοινωνούσα και, επομένως, μη απλοποιήσιμη:

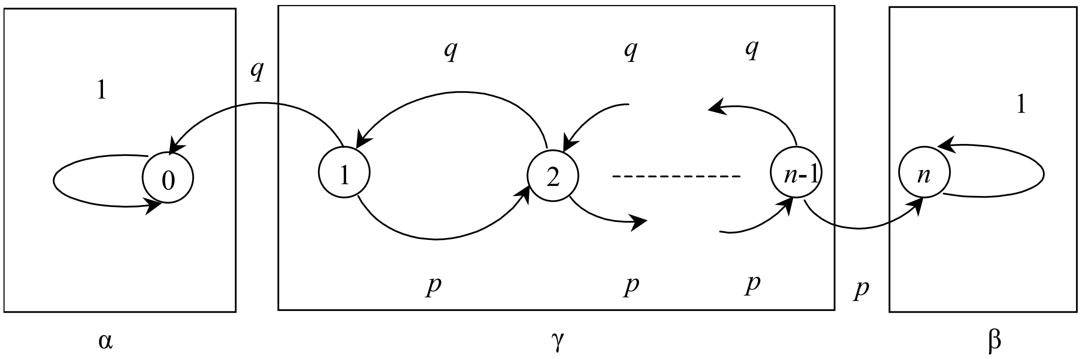


Σχήμα 4.5

Μια επικοινωνούσα κλάση λέγεται **κλειστή** αν δεν υπάρχει καμμιά απολύτως μετάβαση σε άλλη κατάσταση έξω από την κλάση. Επομένως **άπαξ και μπούμε σε μια κλειστή κλάση ποτέ δεν βγαίνουμε απ’ αυτήν**.

Μια **μη απλοποιήσιμη αλυσίδα** συνίσταται από μια μόνο επικοινωνούσα κλάση και άρα **είναι κλειστή**.

Παράδειγμα 4.7 *Η καταστροφή του χαρτοπαίκτη*



Σχήμα 4.6

α: κλειστή κλάση (επικοινωνούσα)

β: " " "

γ: επικοινωνούσα κλάση

Μια επικοινωνούσα κλάση λέγεται **μεταβατική** αν από αυτήν μπορούμε να φθάσουμε σε κάποια κατάσταση έξω από αυτήν και από εκεί **δεν** μπορούμε να επιστρέψουμε στην αρχική κλάση. Για παράδειγμα, η γ είναι μεταβατική αφού από την γ καταλήγουμε ή στην α ή στην β.

Επικοινωνούσες κλάσεις καταλήγουν **πάντα** σε μια κλειστή κλάση και άρα κάθε αλυσίδα Markov έχει τουλάχιστον μια κλειστή κλάση.

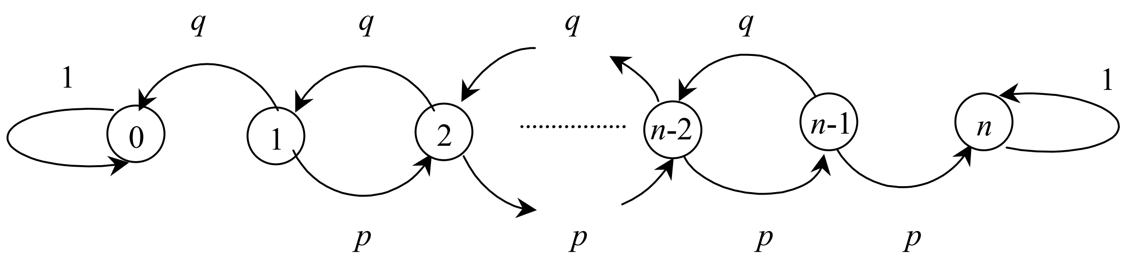
Θεώρημα 4.4

Σε κάθε αλυσίδα Markov η πιθανότητα να βρεθούμε τελικά σε μια κλειστή επικοινωνούσα κλάση ισούται με 1.

Οι κλειστές κλάσεις λέγονται και **απορροφητικές** και αν συνίστανται από μια κατάσταση τότε η κατάσταση λέγεται **απορροφητική** όπως η β και η α.

Παράδειγμα 4.8

Θεωρούμε πάλι την καταστροφή του χαρτοπαίκτη:



Σχήμα 4.7

Γράφουμε τις εξισώσεις αρχίζοντας από τις κλειστές καταστάσεις:

$$P_{k+1}(0) = P_k(0) + qP_k(1)$$

$$P_{k+1}(1) = qP_k(2)$$

$$P_{k+1}(2) = pP_k(1) + qP_k(3)$$

$$\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots$$

$$P_{k+1}(n-2) = pP_k(n-3) + qP_k(n-1)$$

$$P_{k+1}(n-1) = pP_k(n-2)$$

$$P_{k+1}(n) = P_k(n) + pP_k(n-1)$$

Τώρα αναδιατάσσουμε τα στοιχεία του διανύσματος P_k έτσι ώστε οι απορροφητικές κλάσεις να είναι στην αρχή ήτοι, $P_k = [P_k(0)\ P_k(n)\ \ P_k(1)\ P_k(2)\ P_k(3)\ \ldots P_k(n-1)]^{\mathrm{T}}$. Οι προηγούμενες εξισώσεις γράφονται ως εξής:

$$P_{k+1} = \left[\begin{array}{cc|cccccc}
1 & 0 & q & 0 & 0 & \cdots & & 0 \\
0 & 1 & 0 & 0 & 0 & & & p \\
\hline
0 & 0 & 0 & q & 0 & \cdots & & 0 \\
0 & 0 & p & 0 & q & & & 0 \\
0 & 0 & 0 & p & 0 & q & \cdots & 0 \\
\vdots & \vdots & \vdots & & & & \ddots & \\
0 & 0 & & & & \ddots & 0 & q \\
0 & 0 & 0 & \cdots & & & p & 0
\end{array}\right] P_k .$$

Γενικά, ο πίνακας A μπορεί να γραφεί

$$A = \begin{bmatrix} P & | & R \\ --- & | & --- \\ 0 & | & Q \end{bmatrix}$$

όπου για το παράδειγμα που αναφέραμε ο A είναι $(n+1)\times(n+1)$, και ο Q είναι $(n-1)\times(n-1)$ και

$$Q = \begin{bmatrix} 0 & q & 0 & . & . & . & 0 \\ p & 0 & q & . & . & . & 0 \\ 0 & p & 0 & q & . & . & 0 \\ . & . & . & . & . & . & . \\ . & . & . & . & . & . & . \\ 0 & 0 & . & . & p & 0 & q \\ 0 & 0 & . & . & 0 & p & 0 \end{bmatrix}.$$

Ο Q προκύπτει από τον A αν αφαιρέσουμε τις απορροφητικές καταστάσεις. Ο Q δεν είναι στοχαστικός και για τούτο λέγεται **υποστοχαστικός**.

Αρκεί αυτός ο πίνακας για τη μοντελοποίηση των μεταβατικών καταστάσεων. Πράγματι, έστω ότι ξεκινάμε από την κατάσταση S_i και θέλουμε την πιθανότητα στο βήμα k να είμαστε στην κατάσταση S_j (πάντα μέσα στη μεταβατική κλάση). Ως γνωστόν, αυτή η πιθανότητα είναι το στοιχείο του πίνακα A^k που αντιστοιχεί στην πιθανότητα μετάβασης από την i στην j. Μπορεί να αποδειχθεί ότι ο πίνακας A^k έχει την ακόλουθη μορφή

$$A^k = \begin{bmatrix} P^k & | & X \\ --- & | & --- \\ 0 & | & Q^k \end{bmatrix},$$

για κάποιον πίνακα X (δεν μας ενδιαφέρει), οπότε το στοιχείο που ζητάμε ισούται με το στοιχείο $q_{ij}^{(k)}$ του πίνακα $(Q^{\mathrm{T}})^k$. Παρατηρούμε ότι, για $k = 0$, $(Q^{\mathrm{T}})^0 = I$ δηλαδή έχει όλα τα διαγώνια στοιχεία ίσα με 1 και τα υπόλοιπα 0. Αυτό επαληθεύει το ότι η πιθανότητα σε 0 βήματα να πάμε στην j είναι 1 μόνον όταν $i = j$.

Επίσης, το μέσο πλήθος των φορών που θα βρεθούμε στην S_j για $k = 0,1,2,\ldots$ έχοντας ξεκινήσει από την S_i είναι:

$1\times P$(να πάμε στην S_j στο βήμα $k = 0$) + $1\times P$(να πάμε στην S_j στο βήμα $k = 1$) + …

$$= q_{ij}^{(0)} + q_{ij}^{(1)} + ... + q_{ij}^{(k)} + ...$$

που είναι το στοιχείο (ij) του πίνακα

$$I + Q^{\mathrm{T}} + \ldots + (Q^{\mathrm{T}})^k + \ldots$$

Αλλά μπορεί εύκολα να δειχθεί

$$M = (I - Q^{\mathrm{T}})^{-1} = I + Q^{\mathrm{T}} + \ldots\ (Q^{\mathrm{T}})^2 + \ldots$$

Θεώρημα 4.5

Το μέσο πλήθος των φορών που η αλυσίδα, πριν απορροφηθεί, θα ευρίσκεται στην κατάσταση S_j δοθέντος ότι άρχισε από την S_i είναι ίσο με το στοιχείο m_{ij} του πίνακα M.

Τώρα το συνολικό μέσο πλήθος των φορών που η αλυσίδα αρχίζει από την κατάσταση S_i και πηγαίνει σε μεταβατικές καταστάσεις S_1, S_2,…, είναι ίσο με το άθροισμα των όρων της γραμμής i του M. Δηλαδή ο μέσος αριθμός επισκέψεων προς όλες τις μεταβατικές καταστάσεις προτού η αλυσίδα απορροφηθεί έχοντας αρχίσει από την S_i ισούται με το στοιχείο i του διανύσματος $M \cdot \mathbf{1}$ όπου $\mathbf{1}$ είναι το διάνυσμα με όλα τα στοιχεία ίσα με 1.

Τέλος εξετάζουμε την πιθανότητα να καταλήξει η αλυσίδα μ.π. 1 σε απορροφητική κατάσταση S_i δοθέντος ότι άρχισε από μια μεταβατική κατάσταση S_j, να εισέλθει δηλαδή σε απορροφητική κατάσταση από μια συγκεκριμένη οδό.

Θεώρημα 4.6

Η πιθανότητα μια αλυσίδα να καταλήξει από τη μεταβατική κατάσταση S_i σε μια απορροφητική κατάσταση S_j είναι b_{ij}. Τότε ο πίνακας $B = [b_{ij}]$ είναι:

$$B = MR^T.$$

Παράδειγμα 4.9 *Καταστροφή του χαρτοπαίκτη*

Θέλουμε να βρούμε την πιθανότητα να κερδίσει ένας παίκτης έχοντας αρχίσει από κάποια κατάσταση S_k. Ο πίνακας R είναι:

$$R = \begin{bmatrix} q & 0 & 0 & . & . & . & 0 \\ 0 & 0 & 0 & . & . & . & p \end{bmatrix}$$

ή

$$R^T = \begin{bmatrix} q & 0 \\ 0 & 0 \\ \vdots & \vdots \\ 0 & p \end{bmatrix} = [r_1 \,|\, r_2]$$

Ο παίκτης κερδίζει αν βρεθεί στην απορροφητική κατάσταση *n* που αντιστοιχεί στη στήλη

$$r_2 = [0 \quad 0 \quad \cdots \quad p]^{\mathrm{T}}.$$

Άρα το διάνυσμα $P = Mr_2$ δίνει τις ζητούμενες πιθανότητες ή

$$\mathrm{P} = \left(\mathrm{I} - \mathrm{Q}^{\mathrm{T}}\right)^{-1} \mathrm{r}_2 \Rightarrow \left(\mathrm{I} - \mathrm{Q}^{\mathrm{T}}\right)\mathrm{P} = \mathrm{r}_2$$

ή

$$\begin{bmatrix} 1 & -p & 0 & & & & & 0 \\ -q & 1 & -p & & & & & 0 \\ 0 & -q & 1 & -p & & & & 0 \\ . & & . & . & . & & & . \\ . & & & . & . & . & & . \\ . & & & & . & . & . & . \\ 0 & & & & & -q & 1 & -p \\ 0 & & & & & 0 & -q & 1 \end{bmatrix} . \begin{bmatrix} P_1 \\ P_2 \\ P_3 \\ . \\ . \\ . \\ P_{n-2} \\ P_{n-1} \end{bmatrix} = \begin{bmatrix} 0 \\ 0 \\ 0 \\ . \\ . \\ . \\ 0 \\ p \end{bmatrix}$$

ή

$$\left.\begin{aligned} P_1 - pP_2 = 0 \\ -qP_1 + P_2 - pP_3 = 0 \\ \vdots \\ P_{n-1} - qP_{n-2} = p \end{aligned}\right\} \Rightarrow P_k - qP_{k\text{-}1} - pP_{k+1} = 0,$$

για k = 1,2,..,n −1, όπου $P_0 = 0$, $P_n = 1$. Λύνοντας την παραπάνω εξίσωση διαφοράς ευρίσκομε την απάντηση.

V ΣΥΝΤΟΜΗ ΕΙΣΑΓΩΓΗ ΣΤΗ ΘΕΩΡΙΑ ΠΛΗΡΟΦΟΡΙΩΝ

5.1 Γενικά

Ένα από τα θεμελιώδη προβλήματα των επικοινωνιών είναι το ακόλουθο:

Έστω ότι ο πομπός έστειλε ένα σήμα $X(t)$ ενώ στο δέκτη λαμβάνεται το σήμα $Y(t) = X(t) + n(t)$ όπου $n(t)$ είναι θόρυβος, συνήθως λευκός με κατανομή Gauss. Από το $Y(t)$ θέλουμε να εκτιμήσουμε το $X(t)$ ήτοι θέλουμε μια συνάρτηση Φ η οποία μας δίνει την εκτίμηση $\hat{X}(t)$ του $X(t)$:

$$\hat{X}(t) = \Phi(Y(t)).$$

Τέτοια προβλήματα αναφύονται ανά πάσα στιγμή στη φώραση σημάτων ραντάρ όπου αποφασίζουμε αν υπάρχει στόχος ή όχι, στις τηλεπικοινωνίες όπου θέλουμε να ξέρουμε αν έχει ληφθεί 0 ή 1 κ.λπ.

Η Θεωρία Πληροφοριών παρέχει πληθώρα εκπληκτικών εργαλείων που βοηθούν στη λύση του προβλήματος. Η θεωρία στη βάση της οφείλεται στον Claude E. Shannon που τη δημιούργησε σχεδόν ολοκληρωτικά στα εξής δύο άρθρα:

"A Mathematical Theory of Communication," *Bell System Tech. Journal*, vol. 27, pp. 379-423, 623-656, 1948.

Είναι ενδιαφέρον το ότι η κεντρική έννοια της Θεωρίας Πληροφοριών, η εντροπία, βρίσκει εφαρμογές σε πολλές επιστήμες όπως βιολογία, φιλοσοφία, συστήματα παραγωγής και αλλού.

Η Θεωρία Πληροφοριών ασχολείται με τους φορείς πληροφοριών (γράμματα, ψηφία) και όχι με την πληροφορία αυτήν καθ' αυτήν.

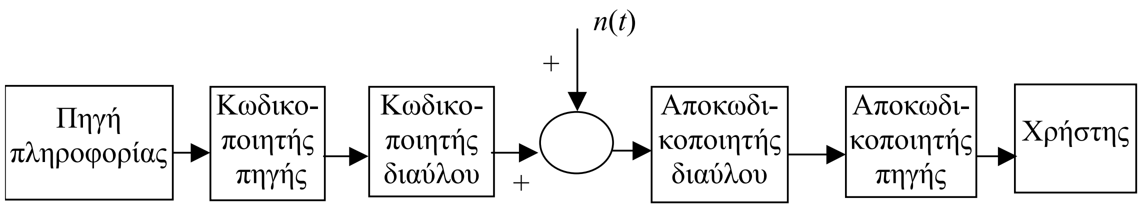
Πληροφορία είναι ένα μέτρο της ελευθερίας επιλογής κάποιου όταν επιλέγει ένα μήνυμα.

Η μέση πληροφορία λέγεται **εντροπία**. Εννοιολογικά είναι ισοδύναμη της κλασικής εντροπίας της θερμοδυναμικής. Ο μαθηματικός της ορισμός δίνεται και στη Στατιστική Μηχανική θεμελιωμένος από τον Boltzman.

Εντροπία είναι ένα μέτρο τυχαιότητας σε μια τυχαία κατάσταση. Μια πολύ καλά οργανωμένη κατάσταση έχει μικρή εντροπία (π.χ. η πορεία κομματιών σε γραμμή παραγωγής ή η διάταξη τούβλων σ' ένα κτίσμα) ενώ μια ανοργάνωτη έχει μεγάλη εντροπία (π.χ. η πορεία κομματιών σ' ένα ευέλικτο σύστημα παραγωγής ή ένας σωρός τούβλων).

Μ' αυτή την έννοια υπάρχει περισσότερη πληροφορία αν κανείς επιλέγει ένα γράμμα από αλφάβητο 50 γραμμάτων παρά από 24 γραμμάτων.

Το διάγραμμα ενός συστήματος επικοινωνιών είναι το εξής:



Σχήμα 5.1

Η **πηγή** μπορεί να είναι φωνή, μουσική, μετρήσεις ραντάρ, ψηφία από μια μαγνητική ταινία, είσοδος στις αισθήσεις ενός βιολογικού οργανισμού.

Ο **κωδικοποιητής** - **(αποκωδικοποιητής) πηγής** είναι μια αντιστοιχία από τα δεδομένα της πηγής (και αντιστρόφως) σε ένα γνωστό αλφάβητο, το αλφάβητο κώδικα, έτσι ώστε να επιτευχθεί η ελάχιστη παραμόρφωση ανάμεσα στο σήμα πηγής και το ληφθέν σήμα.

Ο **δίαυλος** μπορεί να είναι μια τηλεφωνική γραμμή, ένα σύστημα υψηλής συχνότητας (πομπός - κεραία - ατμόσφαιρα - κεραία - δέκτης), ένας βιολογικός οργανισμός.

Ο **κωδικοποιητής - (αποκωδικοποιητής) διαύλου** αντιστοιχεί στον κωδικοποιητή (αποκωδικοποιητή) πηγής με αντικατάσταση της λέξης «πηγή» από τη λέξη «δίαυλος».

Πρώτα κατασκευάζουμε τον κωδικοποιητή - αποκωδικοποιητή διαύλου, ώστε ο δίαυλος να είναι αξιόπιστος και εν συνεχεία χρησιμοποιούμε τον κωδικοποιητή - αποκωδικοποιητή πηγής για να κάνουμε προσαρμογή.

Όπως «ορίσαμε» την πληροφορία είναι προφανές ότι αν προστεθεί θόρυβος σ' ένα σήμα τότε η αβεβαιότητά του αυξάνεται και άρα αυξάνεται η πληροφορία. Αυτή η πληροφορία είναι ανεπιθύμητη και πρέπει να απορριφθεί.

Ορισμός 5.1

Έστω E ένα γεγονός με πιθανότητα να συμβεί $P(E)$. Αν μας πουν ότι το E συνέβη, τότε έχουμε λάβει

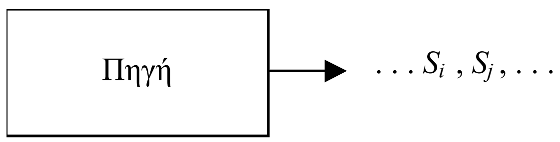
$$I(E) = \frac{1}{log_m P(E)} = -log_m P(E) \qquad \text{μονάδες πληροφορίας.}$$

Αν η βάση των λογαρίθμων είναι το $m = e$ οι μονάδες είναι nats.

Αν η βάση των λογαρίθμων είναι το $m = 2$ οι μονάδες είναι bits.

Αν η βάση των λογαρίθμων είναι το $m = 10$ οι μονάδες είναι Hartleys.

5.2 Η πηγή μηδενικής μνήμης



Σχήμα 5.2

Η πηγή μηδενικής μνήμης εκπέμπει μια ακολουθία συμβόλων από κάποιο δεδομένο, πεπερασμένο αλφάβητο S ={S_1, S_2,…, S_n} και τα σύμβολα είναι στατιστικά ανεξάρτητα μεταξύ τους.

Η μέση τιμή της πληροφορίας λέγεται **εντροπία** της πηγής:

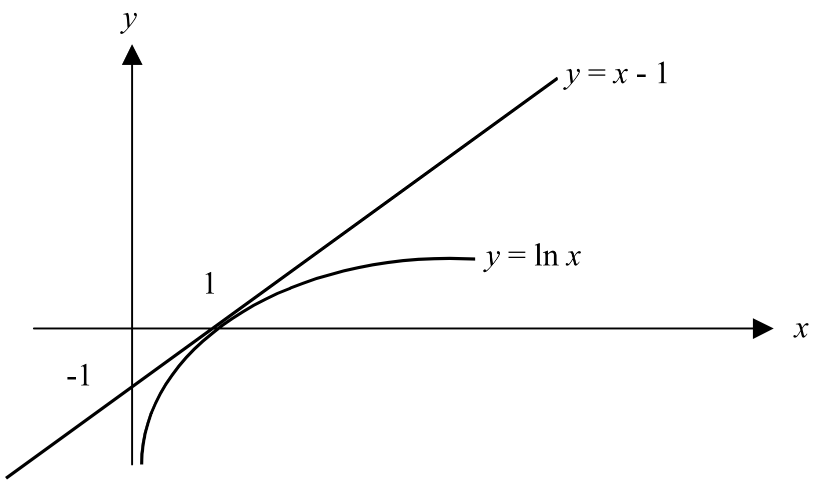
$$H(S) = -\sum_{S} P(S_i)\log P(S_i) = -P(S_1)\log P(S_1) - ... - P(S_n)\log P(S_n)\,.$$

Παράδειγμα 5.1

Έστω το αλφάβητο S ={S_1, S_2, S_3} με πιθανότητες $P(S_1) = \frac{1}{2}$, $P(S_2) = P(S_3) = \frac{1}{4}$.

$H(S) = \frac{1}{2}\log 2 + \frac{1}{4}\log 4 + \frac{1}{4}\log 4 = \frac{3}{2}$ bits.

5.3 Μια σημαντική ιδιότητα της εντροπίας



Σχήμα 5.3

Από την εικόνα συνάγεται ότι $x-1 > \ln x$ με ισότητα για x = 1. Από την ανισότητα έπεται ότι

$$-(x-1) \le -\ln x \Rightarrow 1-x \le \ln\frac{1}{x}. \tag{5.1}$$

Τώρα έστω ότι κάποια πηγή μηδενικής μνήμης έχει αλφάβητο n συμβόλων. Τότε

$$\log n - H(S) = \log n - \sum_{i=1}^{n} P_i \log\frac{1}{P_i} = \sum_{i=1}^{n} P_i \log n - \sum_{i=1}^{n} P_i \log\frac{1}{P_i}$$

$$= \sum_{i=1}^{n} P_i\left(\log n + \log P_i\right) = \sum_{i=1}^{n} P_i \log nP_i . \tag{5.2}$$

Τώρα $\ln nP_i = \dfrac{\log nP_i}{\log e}$ και η (5.2) γίνεται

$$\log n - H(S) = \log e \sum_{i=1}^{n} P_i \ln nP_i$$

και από την (5.1)

$$\log n - H(S) \ge \log e \sum_{i-1}^{n} P_i\left(1-\frac{1}{nP_i}\right) = \log e\left[\sum_{i=1}^{n} P_i - \frac{1}{n}\sum_{i=1}^{n}\frac{P_i}{P_i}\right] = \log e\left(1-\frac{n}{n}\right) = 0 \Rightarrow$$

$$H(S) \le \log n . \tag{5.3}$$

Τώρα έστω μια πηγή με n σύμβολα, όλα ισοπίθανα˙ η εντροπία της είναι

$$H_0(S) = \sum_{i} \frac{1}{n}\log n = \frac{n}{n}\log n = \log n . \tag{5.4}$$

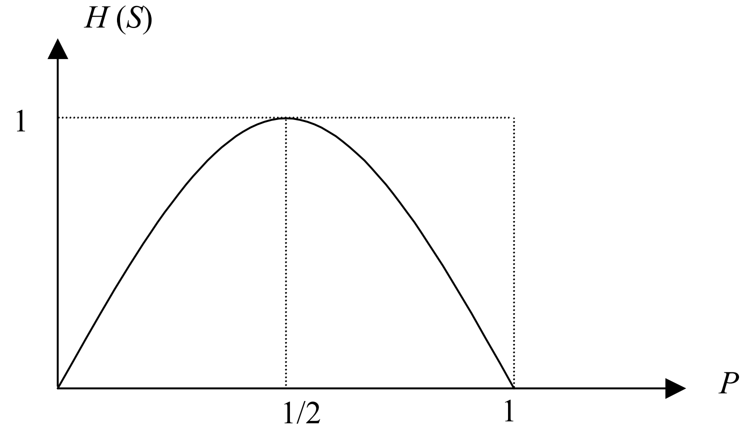
και από τις (5.3), (5.4) συνεπάγεται

$$H(S) \le H_0(S)$$

δηλαδή η εντροπία μεγιστοποιείται για ισοπίθανα σύμβολα και η τιμή της είναι $\log n$.

Παράδειγμα 5.2

Έστω η πηγή $S = \{0,\ 1\}$ και $P(0) = P,\ P(1) = \overline{P} = 1 - P$.



Σχήμα 5.4

Τότε

$$H(S) = P\log\frac{1}{P} + \overline{P}\log\frac{1}{\overline{P}}$$

και για $\quad P = 0 \Rightarrow H(S) = 0$

$$P = 1 \Rightarrow H(S) = 0$$

$$P = \frac{1}{2} \Rightarrow H(S) = 1.$$

5.4 Πώς δικαιολογείται το λογαριθμικό μέτρο της εντροπίας

Έστω μια τυχαία μεταβλητή X που παίρνει τις τιμές $x_1, x_2, \ldots, x_M$ με πιθανότητες $P_1, P_2, \ldots, P_M$ αντίστοιχα. Έστω ότι η μέση αβεβαιότητα των τιμών της X είναι $f(P_1, \ldots, P_M)$. Αν οι τιμές της X είναι ισοπίθανες, τότε

$$H(X) = f\left(\frac{1}{M}, \ldots, \frac{1}{M}\right) \triangleq F(M).$$

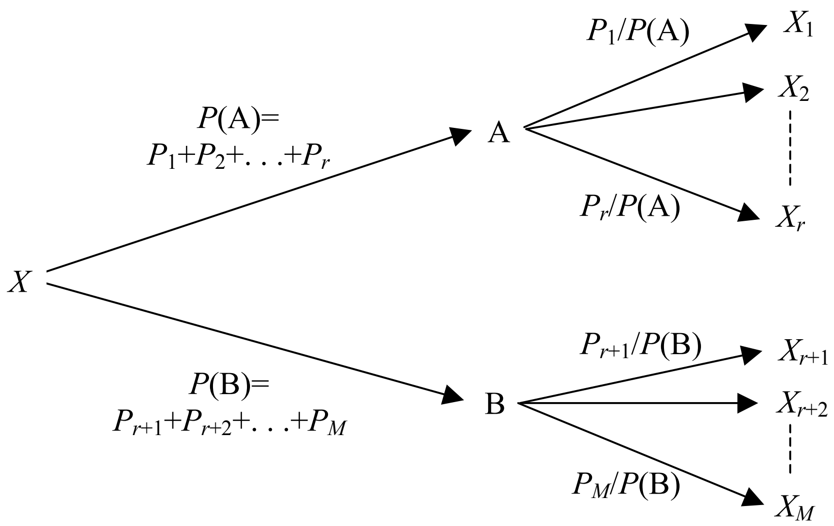
όπου, για συντομία, συμβολίζουμε την αβεβαιότητα ισοπίθανοων ενδεχομένων με F. Τώρα αν το M αυξηθεί πρέπει και η αβεβαιότητα ν' αυξηθεί, ήτοι

1. $F(M) < F(M')$ όταν και μόνο όταν $M < M'$ για $M, M' = 1, 2,...$

Στη συνέχεια εξετάζουμε ένα πείραμα όπου υπάρχουν δύο ανεξάρτητες τυχαίες μεταβλητές X και Y με αντίστοιχες τιμές $x_1, ...,x_M$ και $y_1, ...,y_L$. Οι τιμές της κάθε τυχαίας μεταβλητής είναι ισοπίθανες. Το κοινό πείραμα (X,Y) έχει ML ισοπίθανα εξαγόμενα και μέση αβεβαιότητα $F(ML)$. Αν η τιμή της X αποκαλυφθεί τότε η μέση αβεβαιότητα της Y δεν πρέπει να μεταβληθεί λόγω ανεξαρτησίας. Έτσι $F(ML) - F(M) = F(L)$ ή

2. $F(ML) = F(M) + F(L)$

Στη συνέχεια θεωρούμε την τυχαία μεταβλητή X της οποίας οι τιμές έχουν χωρισθεί σε δύο ομάδες όπως στο σχήμα:



Σχήμα 5.4

Αν αποκαλύψουμε ποιά από τις ομάδες Α ή Β έχει επιλεγεί, τότε έχουμε μειώσει τη μέση αβεβαιότητα κατά $f(P_1+...+P_r\,,\,P_{r+1}+...+P_M)$. Η ομάδα Α επιλέγεται με πιθανότητα $\sum_{i=1}^{r} P_i$ και η εναπομένουσα αβεβαιότητα είναι

$$f\left(\frac{P_1}{\sum_{i=1}^{r} P_i},\ldots,\frac{P_r}{\sum_{i=1}^{r} P_i}\right).$$

Η ομάδα Β επιλέγεται με πιθανότητα $\sum_{i=r+1}^{M} P_i$ και η εναπομένουσα αβεβαιότητα είναι

$$f\left(\frac{P_{r+1}}{\sum_{i=r+1}^{M} P_i},\ldots,\frac{P_M}{\sum_{i=r+1}^{M} P_i}\right).$$

Επομένως κατά μέσον όρο η εναπομένουσα αβεβαιότητα είναι

$$H_r = (P_1 + \ldots P_r)\, f\left(\frac{P_1}{\sum_{i=1}^{r} P_i},\ldots,\frac{P_r}{\sum_{i=1}^{r} P_i}\right) + (P_{r+1} + \ldots + P_M)\, f\left(\frac{P_{r+1}}{\sum_{i=r+1}^{M} P_i},\ldots,\frac{P_M}{\sum_{i=r+1}^{M} P_i}\right).$$

Η συνολική αβεβαιότητα επομένως είναι

3. $f(P_1,P_2,\ldots,P_M)$ = αβεβαιότητα των Α και Β + H_r

$= f(P_1+\ldots+P_r,\ P_{r+1}+\ldots+P_M) + H_r$.

Τέλος, για λόγους αναλυτικής ευκολίας, υποθέτουμε ότι

4. Η $f(P, 1-P)$ είναι συνεχής συνάρτηση του P.

Θεώρημα 5.1

Η μόνη συνάρτηση που ικανοποιεί τα αξιώματα **1-4** είναι

$$H(P_1,...,P_M) = f(P_1,...,P_M) = -C\sum_{i=1}^{M} P_i \log P_i$$

όπου το C ορίζει τις μονάδες.

5.5 Κώδικες

Ορισμός 5.2

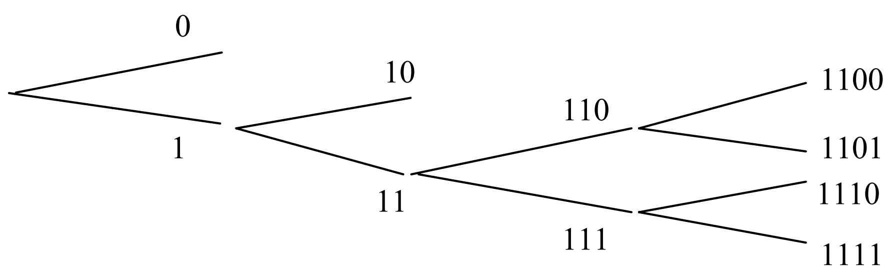
Έστω το σύνολο των συμβόλων ενός αλφαβήτου $S = \{S_1, S_2, ..., S_n\}$. Ορίζουμε σαν **κώδικα** την αντιστοιχία όλων των δυνατών ακολουθιών συμβόλων του S σε ακολουθίες συμβόλων κάποιου άλλου αλφαβήτου $X = \{X_1, X_2, ..., X_n\}$. Το S είναι το αλφάβητο πηγής και το X το αλφάβητο κώδικα.

Ορισμός 5.3

Ένας **σταθερός κώδικας** συνίσταται από ακολουθίες κώδικα που είναι σταθερές για δεδομένη ακολουθία συμβόλων πηγής.

Παράδειγμα 5.3

Έστω $S = \{S_1, S_2, S_3, S_4, S_5, S_6\}$, $X = \{0, 1\}$. Τότε ένα κώδικας του S θα είναι ο



Σχήμα 5.5

ή {0, 10, 1100, 1101, 1110, 1111}. Αλλά υπάρχουν και άπειροι άλλοι συνδυασμοί που μπορούν να δώσουν έναν κώδικα 6 συμβόλων από τα ψηφία (0, 1).

Ορίζουμε το **μήκος** l κάθε λέξης του κώδικα σαν τον αριθμό ψηφίων (bits) σε κάθε λέξη.

Το μέσο μήκος όλου του κώδικα είναι:

$$L = \sum_{i=1}^{n} P_i l_i \ .$$

Ο Shannon έχει αποδείξει ότι

$$H(S) \leq L$$

και η ισότητα ισχύει τότε και μόνον τότε αν

$$l_1 = \log_2 \frac{1}{P_i} \ .$$

Αυτή η σχέση είναι πολύ σημαντική γιατί δίνει τον οικονομικότερο κώδικα.

Παράδειγμα 5.4

Έστω το προηγούμενο παράδειγμα, όπου

S	$P(S_i)$	l_i
S_1	1/2	1
S_2	1/4	2
S_3	1/16	4
S_4	1/16	4
S_5	1/16	4
S_6	1/16	4

Παρατηρούμε ότι ο κώδικας που ήδη βρήκαμε έχει το ελάχιστο L.

Ορισμός 5.4

Ένας **μοναδικά αποκωδικοποιήσιμος** κώδικας είναι εκείνος στον οποίο δεν υπάρχει σύγχυση όσον αφορά το ληφθέν μήνυμα.

Παράδειγμα 5.5

Έστω $S=\{S_1, S_2, S_3\}$. Προφανώς ο κώδικας $X=\{1, 11, 111\}$ δεν είναι μοναδικά αποκωδικοποιήσιμος ενώ ο $X=\{0, 10, 110\}$ είναι και το 0 ενεργεί σαν κόμμα. Τέτοιος κώδικας λέγεται **κόμμα κώδικας**.

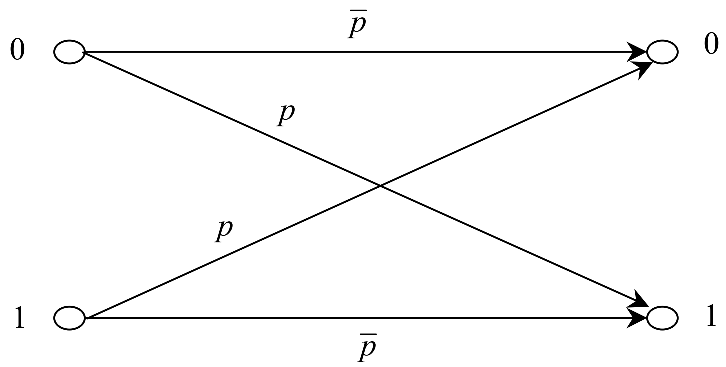
5.6 Δίαυλοι πληροφορίας

Ορισμός 5.5

Ένας δίαυλος πληροφορίας περιγράφεται από ένα αλφάβητο εισόδου $A=\{\mathrm{a}_i\}$, $i=1,2,\ldots,m$ · ένα αλφάβητο εξόδου $B=\{b_j\}$, $j=1,2,\ldots,s$ · και ένα σύνολο υπό συνθήκη πιθανοτήτων $P(b_j|\,\mathrm{a}_i)$ για κάθε i, j.

$$A\left\{\begin{matrix}\mathrm{a}_1\\ \mathrm{a}_2\\ \vdots\\ \mathrm{a}_\mathrm{m}\end{matrix}\right. \quad \rightarrow \quad P\left(b_j|\mathrm{a}_1\right) \quad \rightarrow \quad \left.\begin{matrix}b_1\\ b_2\\ \vdots\\ b_s\end{matrix}\right\}B$$

Παράδειγμα 5.6 *Ο δυαδικός συμμετρικός δίαυλος* (Binary Symmetric Channel, *BSC*)



Σχήμα 5.6

Εδώ $P(0\,|\,0) = \overline{p}$

$P(0|1) = p$

$P(1|0) = p$

$P(1|\,1) = \overline{p}$,

όπου $p+\overline{p}=1$. Οι πιθανότητες αυτές μπορούν να τοποθετηθούν σ' ένα στοχαστικό πίνακα όπως στις αλυσίδες Markov.

$$\begin{bmatrix} P_{00} & P_{10} \\ P_{01} & P_{11} \end{bmatrix} = \begin{bmatrix} \overline{p} & p \\ p & \overline{p} \end{bmatrix}.$$

Η εντροπία του αλφαβήτου εισόδου είναι

$$H(A) = \sum_{i=1}^{m} P(\mathrm{a_i}) \log \frac{1}{P(\mathrm{a_i})}.$$

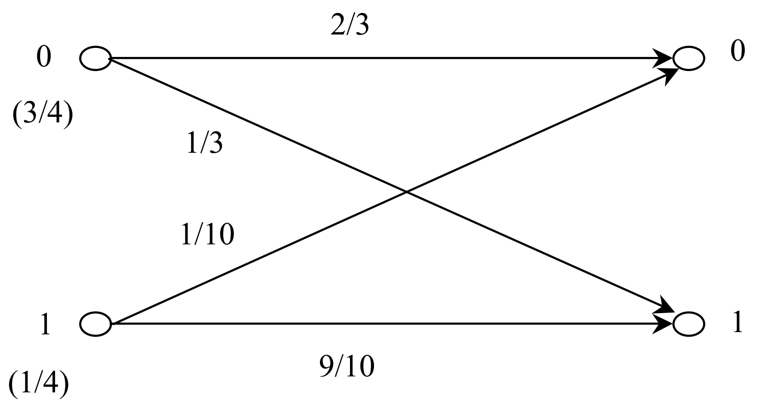
Επίσης

$$P(\mathrm{a}_i\,|b_j) = \frac{P(b_j\,|\,\mathrm{a}_i)P(\mathrm{a}_i)}{\sum_{i=1}^{m} P(b_j\,|\,\mathrm{a}_i)P(\mathrm{a}_i)}$$

και $P(\mathrm{a}_i,b_j)=P(b_j|\mathrm{a}_i)P(\mathrm{a}_i)$.

Παράδειγμα 5.7

Θεωρήστε το δίαυλο:



Σχήμα 5.7

Έστω $P(\mathrm{a}=0)=\frac{3}{4}$

$P(\mathrm{a}=1)=\frac{1}{4}$.

Τότε

$$P(b=0)=P(b=0\,|\,\mathrm{a}=0)P(\mathrm{a}=0)+P(b=0\,|\,\mathrm{a}=1)P(\mathrm{a}=1)=\frac{2}{3}\frac{3}{4}+\frac{1}{10}\frac{1}{4}=\frac{21}{40}$$

$$P(b=1)=\frac{3}{4}\frac{1}{3}+\frac{1}{4}\frac{9}{10}=\frac{19}{40}=1-P(b=0)$$

$$P(\mathrm{a}=0\,|\,\mathrm{b}=0)=\frac{P(b=0\,|\mathrm{a}=0)P(\mathrm{a}=0)}{P(b=0)}=\frac{\frac{2}{3}\frac{3}{4}}{\frac{21}{40}}=\frac{20}{21}$$

$$P(\mathrm{a}=1|b=1)=\frac{\frac{1}{4}\frac{9}{10}}{\frac{19}{40}}=\frac{9}{19}$$

$$P(\mathrm{a}=1\,|\,b=0)=1-\frac{20}{21}=\frac{1}{21},\ P(\mathrm{a}=0\,|\,b=1)=1-\frac{9}{19}=\frac{10}{19}$$

$$P(\mathrm{a}=0, b=0) = P(\mathrm{a}=0 \mid b=0) P(b=0) = \frac{20}{21}\frac{21}{40} = \frac{1}{2}.$$

Ορισμός 5.6

Η ποσότητα $H(A \mid B) = \sum_{A,B} P(\mathrm{a}, b) \log \frac{1}{P(\mathrm{a} \mid b)}$ λέγεται **αμφιλογία** του διαύλου.

Ορισμός 5.7

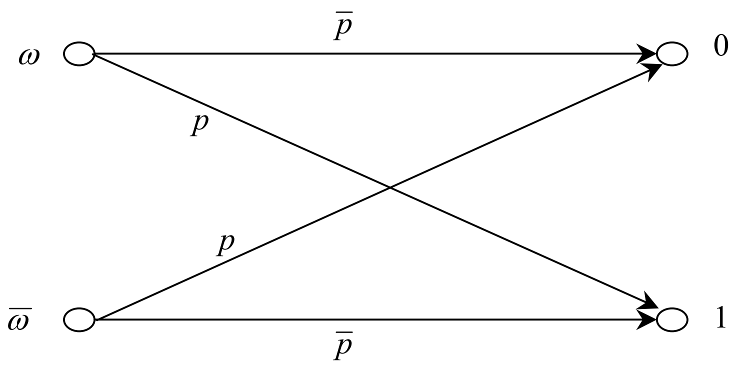
Η ποσότητα $I(A;B) = H(A) - H(A \mid B)$ λέγεται **αμοιβαία πληροφορία** του διαύλου.

Μπορεί ν' αποδειχθεί ότι $I(A;B) = \sum_{A,B} P(\mathrm{a}, b) \log \frac{P(\mathrm{a}, b)}{P(\mathrm{a}) P(\mathrm{b})}$.

Η αμοιβαία πληροφορία είναι ένα μέτρο της αβεβαιότητας που απομένει για το A αφού παρατηρήσουμε το B.

Παράδειγμα 5.8

Έστω ο δίαυλος BSC με πιθανότητες συμβόλων εισόδου $P(0) = \omega$, $P(1) = \overline{\omega}$.



Σχήμα 5.8

Προφανώς

$$H(A) = \omega \log \frac{1}{\omega} + \overline{\omega} \log \frac{1}{\overline{\omega}}.$$

Ενώ

$$H(A|B) = \sum_A \sum_B P(\mathrm{a}, b) \log \frac{1}{P(\mathrm{a}|b)} = \sum_B P(1,b) \log \frac{1}{P(1|b)} + \sum_B P(0,b) \log \frac{1}{P(0|b)}$$

$$= P(11) \log \frac{1}{P(1|1)} + P(10) \log \frac{1}{P(1|0)} + P(00) \log \frac{1}{P(0|0)} + P(01) \log \frac{1}{P(0|1)}.$$

Αλλά

$$P(\mathrm{a}=1|b=1) = \frac{P(b=1|\mathrm{a}=1)P(\mathrm{a}=1)}{P(b=1)} = \frac{\overline{p}\overline{\omega}}{\overline{p}\overline{\omega} + p\omega}$$

$$P(\mathrm{a}=1|b=0) = \frac{p\overline{\omega}}{p\overline{\omega} + \overline{p}\omega}, \quad P(\mathrm{a}=0|b=1) = \frac{p\omega}{\overline{p}\overline{\omega} + p\omega}$$

$$P(\mathrm{a}=0|b=0) = \frac{\overline{p}\omega}{p\overline{\omega} + \overline{p}\omega}$$

$$P(\mathrm{a}=0, b=0) = P(b=0|\mathrm{a}=0)P(\mathrm{a}=0) = \overline{p}\omega$$

$$P(\mathrm{a}=0, b=1) = p\omega,$$

$$P(\mathrm{a}=1, b=0) = p\overline{\omega}$$

$$P(a=1, b=1) = \overline{p}\overline{\omega}$$

οπότε

$$H(A|B) = \overline{p}\overline{\omega} \log \frac{\overline{p}\overline{\omega} + p\omega}{\overline{p}\omega} + p\overline{\omega} \log \frac{p\overline{\omega} + \overline{p}\omega}{p\overline{\omega}} + \overline{p}\omega \log \frac{p\overline{\omega} + \overline{p}\omega}{\overline{p}\omega} + p\omega \log \frac{\overline{p}\overline{\omega} + p\omega}{p\omega}.$$

Τελικά

$$I(A;B) = \omega \log \frac{1}{\omega} + \overline{\omega} \log \frac{1}{\overline{\omega}} + \overline{p}\overline{\omega} \log \frac{\overline{p}\overline{\omega}}{\overline{p}\overline{\omega} + p\omega}$$

$$+ p\overline{\omega} \log \frac{p\overline{\omega}}{p\overline{\omega} + \overline{p}\omega} + \overline{p}\omega \log \frac{\overline{p}\omega}{p\overline{\omega} + \overline{p}\omega} + p\omega \log \frac{p\omega}{\overline{p}\overline{\omega} + p\omega}$$

$$= \omega \log \frac{1}{\omega} + \overline{\omega} \log \frac{1}{\overline{\omega}} + \overline{p}\overline{\omega} \log \overline{p}\overline{\omega} + \overline{p}\overline{\omega} \log \frac{1}{\overline{p}\overline{\omega} + p\omega}$$

$$+ p\overline{\omega} \log p\overline{\omega} + p\overline{\omega} \log \frac{1}{p\overline{\omega} + \overline{p}\omega} + \overline{p}\omega \log \overline{p}\omega$$

$$+ \overline{p}\omega \log \frac{1}{p\overline{\omega} + \overline{p}\omega} + p\omega \log p\omega + p\omega \log \frac{1}{\overline{p}\overline{\omega} + p\omega}$$

$$= (\overline{p}\overline{\omega} + p\omega) \log \frac{1}{\overline{p}\overline{\omega} + p\omega} + (p\overline{\omega} + \overline{p}\omega) \log \frac{1}{p\overline{\omega} + \overline{p}\omega} - H(p)\,.$$

Ορισμός 5.8

Η *χωρητικότητα C* του διαύλου ορίζεται

$$C = \max_{P(\mathrm{a}_i)} I(A;B) \quad \text{bits/binit (binit=δυαδικό σύμβολο).}$$

Η χωρητικότητα δεν εξαρτάται από τα $P(\mathrm{a}_i)$ και είναι μόνο συνάρτηση των πιθανοτήτων διαύλου. Το πρόβλημα της μεγιστοποίησης της χωρητικότητας είναι πρόβλημα μη γραμμικού προγραμματισμού.

Για το BSC το μέγιστο συμβαίνει για

$$\omega = \overline{\omega} = \frac{1}{2} \text{ και}$$

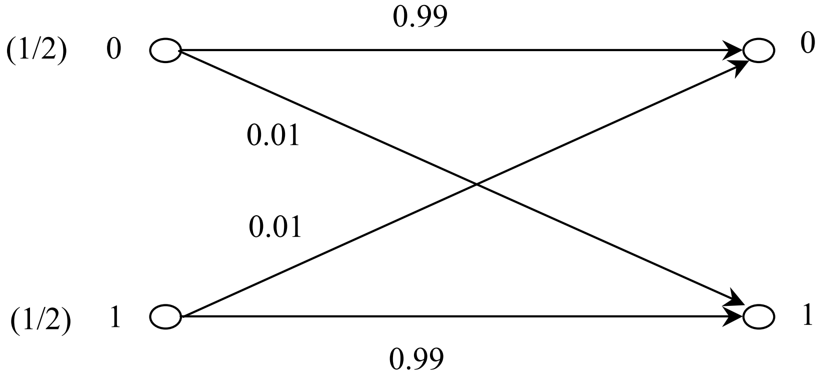
$$C = 1 - p \log \frac{1}{p} - \overline{p} \log \frac{1}{\overline{p}}\,.$$

Το ακόλουθο είναι το περίφημο **θεώρημα του Shannon**, ένα από τα πιο εκπληκτικά και ευφυή θεωρήματα του 20ου αιώνα:

Θεώρημα 5.2

Για κάθε διακεκριμένης εισόδου και χωρίς μνήμη δίαυλο υπάρχει τουλάχιστον ένας κώδικας N συμβόλων με ρυθμό $R < C$ έτσι ώστε η πιθανότητα σφάλματος αποκωδικοποίησης να γίνεται οσοδήποτε μικρή.

Ας ιδούμε διαισθητικά το νόημα των προηγούμενων εννοιών. Έστω ο ακόλουθος δίαυλος ο οποίος λειτουργεί με ρυθμό εισόδου R = 1000 σύμβολα/s.



Σχήμα 5.9

Η πιθανότητα σφάλματος είναι

$$P_E = P(1\,|\,0)P(0) + P(0\,|\,1)P(1) = 2\left[0.01\frac{1}{2}\right] = 0.01.$$

Άρα κατά μέσον όρο 1% των bits λαμβάνεται λανθασμένα ή 1 σύμβολο στα 100, επομένως ο ρυθμός εκπομπής πληροφορίας είναι < 1000 bits/s.

Αν τώρα υποθέσουμε ότι το 1% αντιστοιχεί σε πραγματικό ρυθμό εκπομπής 990 bits/s έχουμε κάνει το λάθος να μην παίρνουμε υπ' όψιν τη θέση των ψηφίων στο μήνυμα. Τώρα δηλαδή δεν ξέρουμε που έχουν συμβεί τα 10 λάθη. Το πρόβλημα λύνεται εξετάζοντας την αβεβαιότητα του ληφθέντος σήματος που είναι η αμφιλογία:

$$H(A\,|\,b) = \sum_{A} P(\mathrm{a}\,|\,b)\log\frac{1}{P(\mathrm{a}\,|\,b)}$$

και κατά μέσον όρο:

$$H(A\,|\,B) = \sum_{A,B} P(\mathrm{a}\,|\,b)P(b)\log\frac{1}{P(\mathrm{a}\,|\,b)} = \sum_{A,B} P(\mathrm{a},b)\log\frac{1}{P(\mathrm{a}\,|\,b)}.$$

Έτσι δικαιολογείται ο ορισμός της αμφιλογίας.

Έστω τώρα ότι $H(A) = 10$ και $H(A\,|\,B) = 8$. Για να μειώσουμε την αβεβαιότητα κατά $10 - 8 = 2$ χρειαζόμαστε 2 bits πληροφορίας που τελικά είναι ο πραγματικός ρυθμός πληροφορίας που περνάει από το δίαυλο, δηλαδή

$$R = H(A) - H(A \mid B) = I(A;B)$$

και αν μεγιστοποιήσουμε το $I(A;B)$ αποκτούμε το C που μας δίνει το μέγιστο ρυθμό που ο δίαυλος μπορεί να προωθήσει πληροφορία. Αν $R \geq C$ ο δίαυλος γίνεται εξαιρετικά αναξιόπιστος.

5.7 Μερικές ακόμη σημαντικές ιδιότητες της εντροπίας

Θεωρούμε μια τυχαία μη αρνητική μεταβλητή X της οποίας γνωρίζουμε μόνο τη μέση τιμή λ. Θα θέλαμε να βρούμε μια πυκνότητα $f(x)$ τέτοια ώστε η προκατάληψη να είναι ελάχιστη ή η αβεβαιότητα μέγιστη. Ξέρουμε ότι η αβεβαιότητα εκφράζεται από την εντροπία. Για μια συνεχή μεταβλητή η εντροπία γίνεται

$$H = -\int_0^{\infty} f(x) \ln f(x) dx . \tag{5.5}$$

Στο συγκεκριμένο πρόβλημα ξέρουμε ότι

$$\int_0^{\infty} x f(x) dx = \lambda \ \text{ και } \int_0^{\infty} f(x) dx = 1 . \tag{5.6}$$

Θέλουμε λοιπόν να μεγιστοποιήσουμε την (5.5) υποκείμενης στους περιορισμούς (5.6). Η εντροπία H αποτελεί γενίκευση των συμβατικών συναρτήσεων και ονομάζεται **συναρτησιακό** (functional).

Ορισμός 5.9:

Συναρτησιακό είναι ένας κανόνας αντιστοιχίας συναρτήσεων σε πραγματικούς αριθμούς.

Εδώ για οιαδήποτε μη αρνητική συνάρτηση $f(x)$, το συναρτησιακό H ισούται με την τιμή του ολοκληρώματος της συνάρτησης $f \ln f$. Ορίζουμε

$$Z_1(x) = \int_0^x y f(y) dy \Rightarrow \dot{Z}_1(x) = x f(x),$$

$$Z_2(x) = \int_0^x f(y) dy \Rightarrow \dot{Z}_2(x) = f(x).$$

Από το λεγόμενο λογισμό των μεταβολών σχηματίζουμε την επαυξημένη συνάρτηση:

$$g_{\mathrm{a}}[f(x), P_1(x), P_2(x), x] = -f \ln f + P_1(xf - \dot{Z}_1) + P_2(f - \dot{Z}_2)$$

όπου τα P_1 και P_2 είναι σαν τους πολλαπλασιαστές Lagrange. Υπάρχει ένα θεώρημα που δίνει την αναγκαία συνθήκη μεγίστου και λέγεται **εξίσωση του Euler**:

$$\frac{\partial g_{\mathrm{a}}}{\partial f} - \frac{d}{dx}\frac{\partial g_{\mathrm{a}}}{\partial \dot{f}} = 0$$

Επειδή $\frac{\partial g_{\mathrm{a}}}{\partial \dot{f}} = 0$, η εξίσωση Euler γράφεται

$$-\ln f - 1 + P_1 x + P_2 = 0$$

$$\Rightarrow f = e^{P_1 x} e^{P_2 - 1}. \tag{5.7}$$

Η (5.7) πρέπει να ικανοποιεί τις (5.6) από τις οποίες ευρίσκομε

$$f(x) = \frac{1}{\lambda} e^{-\frac{1}{\lambda}x}$$

δηλαδή η εκθετική κατανομή είναι η λιγότερο προκατειλημμένη κατανομή όταν ξέρουμε μόνο τη μέση τιμή μιας μη αρνητικής τυχαίας μεταβλητής.

Στη συνέχεια θέλουμε την πλέον μη προκατειλημμένη πυκνότητα μιας τυχαίας μεταβλητής X για την οποία γνωρίζουμε τη μέση τιμή η οποία για απλούστευση θα θεωρηθεί 0 και τη διασπορά σ^2 καθώς και ότι $-\infty < X < \infty$.

Εδώ μεγιστοποιούμε το συναρτησιακό

$$-\int_{-\infty}^{\infty} f(x) \ln f(x) dx$$

υποκείμενου στους περιορισμούς

$$\int_{-\infty}^{\infty} x^2 f(x)dx = \sigma^2, \qquad \int_{-\infty}^{\infty} f(x)dx = 1.$$

Ορίζουμε

$$Z_1(x) = \int_0^x y^2 f(y)dy \Rightarrow \dot{Z}_1(x) = x^2 f(x),$$

$$Z_2(x) = \int_0^x f(y)dy \Rightarrow \dot{Z}_2(x) = f(x).$$

Σχηματίζουμε πάλι την επαυξημένη συνάρτηση

$$g_{\mathrm{a}} = -f(x)\ln f(x) + P_1\left(x^2 f(x) - \dot{z}_1\right) + P_2\left(f(x) - \dot{z}_2\right)$$

και από την εξίσωση Euler

$$\frac{\partial g_{\mathrm{a}}}{\partial f} - \frac{d}{dx}\frac{\partial g_{\mathrm{a}}}{\partial \dot{f}} = 0$$

$$\Rightarrow -\ln f - 1 + P_1 x^2 + P_2 = 0$$

$$\Rightarrow f(x) = e^{P_1 x^2 + P_2 - 1}.$$

Αντικαθιστώντας την $f(x)$ στους περιορισμούς και μετά από λίγη άλγεβρα προκύπτει ότι

$$f(x) = \frac{1}{\sqrt{2\pi\sigma^2}} e^{-\frac{x^2}{2\sigma^2}}$$

που σημαίνει ότι η κατανομή Gauss είναι η πλέον τυχαία κατά την έννοια της εντροπίας όταν ξέρουμε τη μέση τιμή και τη διασπορά.

ΠΑΡΑΡΤΗΜΑ: ΕΙΣΑΓΩΓΗ ΣΤΙΣ ΓΕΝΙΚΕΥΜΕΝΕΣ ΣΥΝΑΡΤΗΣΕΙΣ

A.1 Αναγκαιότητα της Εισαγωγής Γενικευμένων Συναρτήσεων

Για τη μελέτη δυναμικών συστημάτων, διαδικασιών και φυσικών φαινομένων αναπτύσσουμε μαθηματικά μοντέλα. Στα μοντέλα αυτά, η εξέλιξη στο χρόνο περιγράφεται από μία συνάρτηση $\varphi(x)$ η οποία εκφράζει την κατάσταση του συστήματος τη χρονική στιγμή x. Σύμφωνα με τον συνήθη ορισμό, μία πραγματική συνάρτηση $\varphi(.)$ είναι ένας κανόνας αντιστοίχησης οιουδήποτε σημείου x της ευθείας των πραγματικών αριθμών σε έναν πραγματικό αριθμό $\varphi(x)$. Υπάρχουν πολλά συστήματα για τα οποία η $\varphi(x)$ δεν μπορεί να εκφρασθεί ως συνάρτηση με τη συμβατική έννοια της αντιστοιχίας σημείων.

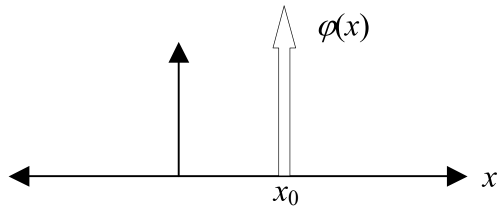
Παράδειγμα 1: Μία σφαίρα με πεπερασμένη κινητική ενέργεια E πέφτει πάνω σε ακίνητο επίπεδο τη χρονική στιγμή x_0. Η σύγκρουση είναι πλαστική και ιδανική, υπό την έννοια ότι η σφαίρα καθηλώνεται στο επίπεδο μέσα σε απειροελάχιστο χρόνο. Ας μελετήσουμε την ισχύ $\varphi(x)$ που προσφέρει η σφαίρα στο επίπεδο κάθε χρονική στιγμή $x\in(-\infty, \infty)$. Επειδή η ενέργεια E προσδίδεται ακαριαία τη στιγμή x_0, η ισχύς $\varphi(x)$, ως συνάρτηση του χρόνου, έχει τιμή τιμή 0 κάθε χρονική στιγμή x, εκτός από τη στιγμή x_0 της πρόσκρουσης. Εφαρμόζοντας την εξίσωση

$$[\text{ενέργεια}] = [\text{ισχύς}] \times [\text{απειροελάχιστη διάρκεια}]$$

προκύπτει

$$E = \varphi(x_0)\times 0$$

Καμμία πεπερασμένη τιμή $\varphi(x_0)$ δεν ικανοποιεί την εξίσωση αυτή. Διαισθητικά οδηγούμαστε στο συμπέρασμα ότι η $\varphi(x)$ θα πρέπει να έχει άπειρη τιμή για $x = x_0$ και μηδέν για $x \neq x_0$. Η συνάρτηση $\varphi(x)$ περιγράφεται σχηματικά από το επόμενο σχήμα.



Η ενέργεια που προσδίδεται στο επίπεδο κατά το χρονικό διάστημα (α, β) δίδεται από τη σχέση ενέργειας-ισχύος:

$$[\text{ενέργεια στο διάστημα } (\alpha, \beta)] = \int_{\alpha}^{\beta} \varphi(x)dx$$

Επειδή η σύγκρουση συμβαίνει τη στιγμή x_0 και διαρκεί απειροελάχιστο χρόνο, θα πρέπει

$$\int_{\alpha}^{\beta} \varphi(x)dx = \begin{cases} E, & \text{για κάθε } \alpha, \beta \text{ τέτοια ώστε } x_0 \in (\alpha, \beta) \\ 0, & \text{αν } x_0 \notin (\alpha, \beta) \end{cases} \tag{1}$$

Και πάλι διαπιστώνουμε ότι καμμία κοινή συνάρτηση $\varphi(x)$ δέν ικανοποιεί την Εξ. (1).

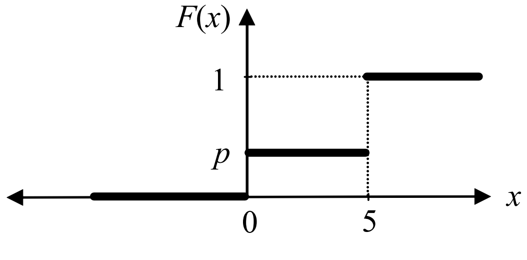
Παράδειγμα 2: Μία τυχαία μεταβλητή X ορίζεται

$$X = \begin{cases} 0, \text{ με πιθανότητα } p \\ 5, \text{ με πιθανότητα } 1-p \end{cases}$$

Η X είναι διακεκριμένη τυχαία μεταβλητή με συνάρτηση κατανομής

$$F(x) = P(X \le x) = \begin{cases} 1, \text{ για } x \ge 5 \\ p, \text{ για } 0 \le x < 5 \\ 0, \text{ για } x < 0 \end{cases}$$

για κάθε $x \in (-\infty, \infty)$. Γραφικά,



Στο μάθημα των Πιθανοτήτων είχατε δεί ότι η συνάρτηση κατανομής ορίζεται πάντα, ενώ η συνάρτηση πυκνότητας πιθανότητας $f(x) = F'(x)$ ορίζεται μόνο για απολύτως συνεχείς τυχαίες μεταβλητές. Εδώ η $F(x)$ παρουσιάζει ένα άλμα ύψους p στο σημείο 0 και ακόμη ένα ύψους $(1-p)$ στο 5. Επομένως δεν είναι διαφορίσιμη. Εν τούτοις, παρατηρούμε ότι η παράγωγος της F είναι παντού 0 εκτός από τα σημεία 0 και 5. Με τα ανάλογα επιχειρήματα όπως εκείνο του Παραδείγματος 1, οδηγούμαστε στο συμπέρασμα ότι θα πρέπει

$$\int_{\alpha}^{\beta} f(x)dx = \begin{cases} 0, \text{ αν } \alpha < \beta < 0 \\ p, \text{ για κάθε } \alpha, \beta \text{ τέτοια ώστε } \alpha < 0 < \beta < 5 \\ 1-p, \text{ αν } 0 < \alpha < 5 \le \beta \\ 1, \text{ αν } \alpha < 0 \text{ και } 5 \le \beta \end{cases}$$

Και πάλι διαπιστώνουμε ότι καμμία κοινή συνάρτηση $f(x)$ δέν ικανοποιεί την εξίσωση αυτή.

Όπως θα ιδούμε στη συνέχεια, η μελέτη μη κοινών ή, όπως λέμε, **γενικευμένων συναρτήσεων** είναι εφικτή με την επέκταση εννοιών όπως η «ισότητα», το «όριο», και η «παράγωγος» κοινών συναρτήσεων.

Για τη μελέτη εξισώσεων όπως η (1) που περιέχουν γενικευμένες συναρτήσεις $\varphi(x)$ ορίζουμε κατάλληλη ακολουθία συναρτήσεων $\varphi_n(x)$ οι οποίες:

- είναι κοινές συναρτήσεις με επιθυμητές ιδιότητες (έχουν αναλυτική έκφραση, παραγωγίζονται, ολοκληρώνονται, κλπ) και
- ικανοποιούν την Εξ. (1) ασυμπτωτικά, ήτοι

$$\lim_{n\to\infty} \int_{\alpha}^{\beta} \varphi_n(x)dx = \begin{cases} E, \text{ για κάθε } \alpha, \beta \text{ τέτοια ώστε } x_0 \in (\alpha, \beta) \\ 0, \text{ αν } x_0 \notin (\alpha, \beta) \end{cases}$$

Συνεπώς, αντί να προσδορίσουμε την $\varphi(x)$ απ' ευθείας, ευρίσκουμε κοινές συναρτήσεις $\varphi_n(x)$ και μελετάμε τις ιδιότητές τους για $n\to\infty$. Με άλλα λόγια η $\varphi(x)$ «κληρονομεί» τις ιδιότητες των $\varphi_n(x)$ ασυμπτωτικά.

Η ιστορική αναδρομή της θεωρίας των γενικευμένων συναρτήσεων συνοψίζεται με κομψότητα στην αφιέρωση του βιβλίου του Lighthill(*):

To

Paul Dirac who saw that it must be true,

Laurent Schwartz who proved it,

and George Temple who showed how simple it could be made

A.2 Ασυμπτωτική Συμπεριφορά Συναρτήσεων

Ορισμός

Μία συνάρτηση $f(x)$ λέμε ότι ανήκει στην οικογένεια συναρτήσεων $O[g(x)]$ όταν $x \to x_0$ αν υπάρχει αριθμός $0 \le A < \infty$ τέτοιος ώστε

$$\lim_{x \to x_0} |f(x)| \le A \lim_{x \to x_0} |g(x)| \quad \text{ή} \quad \lim_{x \to x_0} \frac{|f(x)|}{|g(x)|} \le A \ (**)$$

Ο αριθμός x_0 μπορεί να είναι πεπερασμένος ή άπειρος. Γράφουμε συμβολικά

$$f(x) = O[g(x)] \text{ για } x \to x_0$$

και διαβάζουμε "η $f(x)$ είναι $O[g(x)]$ όταν $x \to x_0$". Όταν $A = 0$, τότε η $f(x)$ είναι $o[g(x)]$.

(*) M. J. Lighthill, *Introduction to Fourier Analysis and Generalized Functions*, Cambridge University Press, Cambridge, 1958.

(**) Ένας αριθμός α είναι το όριο της $f(x)$ για $x \to x_0$ αν για οσοδήποτε μικρό θετικό αριθμό ε υπάρχει θετικός αριθμός δ_ε τέτοιος ώστε για κάθε x τέτοιο ώστε $|x - x_0| < \delta_\varepsilon$ ισχύει $|f(x) - \alpha| < \varepsilon$. Δηλαδή για κάθε σημείο x που βρίσκεται μέσα στο διάστημα $x_0 \pm \delta_\varepsilon$, η $f(x)$ περιορίζεται σε μία ζώνη εύρους $\pm\varepsilon$ γύρω από το α.

Το όριο της $f(x)$ για $x \to \infty$ είναι το α αν για οποιοσοδήποτε μικρό θετικό αριθμό ε υπάρχει ένα σημείο x_ε τέτοιο ώστε $|f(x) - \alpha| < \varepsilon$, για κάθε $x > x_\varepsilon$.

Αν στον ορισμό του O() το όριο δεν υπάρχει, τότε το lim λαμβάνεται ως όριο των μέγιστων τιμών, lim-sup $f(x)$. Για παράδειγμα, όταν $x \to \infty$, τότε το lim $[\sin(x) + 1/x]$ δεν υπάρχει, αλλά lim-sup $[\sin(x) + 1/x] = 1$.

Δύο συναρτήσεις είναι O[$g(x)$] όταν οι ανωτέρω ανισότητες επαληθεύονται, όχι κατ' ανάγκη με το ίδιο A για κάθε συνάρτηση.

Ιδιότητα: Αν οι $f_k(x)$ να είναι O[$g(x)$] για $x\to\infty$ και $\alpha_k < \infty$, $k = 1, 2, \ldots, n$, τότε η συνάρτηση $[\alpha_1 f_1(x) + \ldots + \alpha_n f_n(x)]$ είναι O[$g(x)$] ενώ η $[f_1(x) \ldots f_n(x)]$ είναι $O\{[g(x)]^n\}$.

Παράδειγμα: Η $f_1(x) = 8x^3 + 4x$ είναι $O(x^3)$ γιατί

$$\lim_{x\to\infty} \frac{|8x^3 + 4x|}{|x|^3} =$$

$$\text{(από την ιδιότητα } \lim|f(x)| = |\lim f(x)|\text{)} = \left| \lim_{x\to\infty} \frac{8x^3 + 4x}{x^3} \right| = \left| 8 + \lim_{x\to\infty} \frac{4}{x^2} \right| < \infty$$

Η $f_2(x) = -2x^3 \sin(x) + 1000$ είναι $O(x^3)$. Πράγματι, αφού το ημίτονο είναι το πολύ ίσο με 1, προκύπτει

$$\lim_{x\to\infty}\text{-}\sup \frac{|-2x^3 \sin(x) + 1000|}{|x|^3} \le \lim_{x\to\infty}\text{-}\sup \frac{|-2x^3 \sin(x)| + 1000}{|x^3|}$$

$$= \lim_{x\to\infty}\text{-}\sup \left[2|\sin(x)| + \frac{1000}{|x^3|} \right] < \infty$$

Επαληθεύεται ότι η $f_1(x) + f_2(x)$ είναι επίσης $O(x^3)$. Παρατηρούμε ότι όταν το $x\to\infty$, στις συναρτήσεις $f_1(x), f_2(x)$ κυριαρχούν οι όροι μεγαλύτερης τάξεως: $8x^3$, $-2x^3$.

A.3 Καλές και Αρκετά Καλές Συναρτήσεις

Ορισμός

Μία συνάρτηση $\varphi(x)$ ονομάζεται **καλή** αν

– έχει παραγώγους κάθε τάξης σε κάθε σημείο x, και

– η ίδια καθώς και οι παράγωγοί της είναι $\mathrm{O}\left(\frac{1}{|x|^N}\right)$, όταν $|x|\to\infty$, για **κάθε** φυσικό αριθμό $N = 1, 2, \ldots$

Παράδειγμα: Η e^{-x^2} είναι καλή συνάρτηση. Πράγματι, έχουμε

$$\lim_{|x|\to\infty}\frac{\left|e^{-x^2}\right|}{\frac{1}{|x|^N}} = \lim_{|x|\to\infty}\frac{|x|^N}{e^{|x|^2}}$$

και παίρνοντας το ανάπτυγμα Taylor ως προς $|x|^2$

$$\lim_{|x|\to\infty}\frac{|x|^N}{e^{|x|^2}} = \lim_{|x|\to\infty}\frac{|\mathrm{x}|^{\mathrm{N}}}{\left[1+\frac{|\mathrm{x}|^2}{1!}+\ldots+\frac{|\mathrm{x}|^{2\mathrm{N}}}{\mathrm{N}!}+\ldots\right]} \le \lim_{|x|\to\infty}\frac{|x|^N}{\frac{|x|^{2N}}{N!}} = \lim_{|x|\to\infty}\frac{N!}{|x|^N}$$

όπου στον παρονομαστή του δεξιού μέλους της ανισότητας έχουμε κρατήσει μόνον τον N-οστό όρο του αναπτύγματος. Το τελευταίο όριο ισούται με 0. Πράγματι, $N! = 1\times2\times\ldots\times N \le N\times N\times\ldots\times N = N^N$ και επαληθεύουμε τον ορισμό του ορίου:

$$\text{για κάθε } \varepsilon > 0 \Rightarrow \left|\frac{N!}{|x|^N} - 0\right| = \frac{N!}{|x|^N} \le \left(\frac{N}{|x|}\right)^N < \varepsilon \quad \text{για όλα τα } x > x_\varepsilon \stackrel{\Delta}{=} \frac{N}{\sqrt[N]{\varepsilon}}$$

Συνεπώς η e^{-x^2} είναι $\mathrm{O}\left(\frac{1}{|x|^N}\right)$ για κάθε N. Με ανάλογο τρόπο αποδεικνύεται ότι και οι παράγωγοί της είναι επίσης $\mathrm{O}\left(\frac{1}{|x|^N}\right)$.

Ορισμός

Μία συνάρτηση λέγεται **αρκετά καλή** αν

– έχει παραγώγους κάθε τάξης σε κάθε σημείο x

– και η ίδια καθώς και οι παράγωγοί της είναι $\mathrm{O}(|x|^N)$, όταν $|x|\to\infty$, για **κάποιο** φυσικό αριθμό N.

Ιδιότητες: – Οι καλές συναρτήσεις είναι και αρκετά καλές.

– Τα όρια καλής συνάρτησης για $x \to \pm\infty$ είναι 0.

– Η παράγωγος καλής συνάρτησης είναι καλή συνάρτηση.

– Το άθροισμα καλών συναρτήσεων είναι καλή συνάρτηση.

– Το γινόμενο καλής με αρκετά καλή είναι καλή συνάρτηση

– Ο μετασχηματισμός Fourier (θα τον ορίσουμε αργότερα) μίας καλής συνάρτησης είναι καλή συνάρτηση.

Α.4 Ομαλές Ακολουθίες Συναρτήσεων

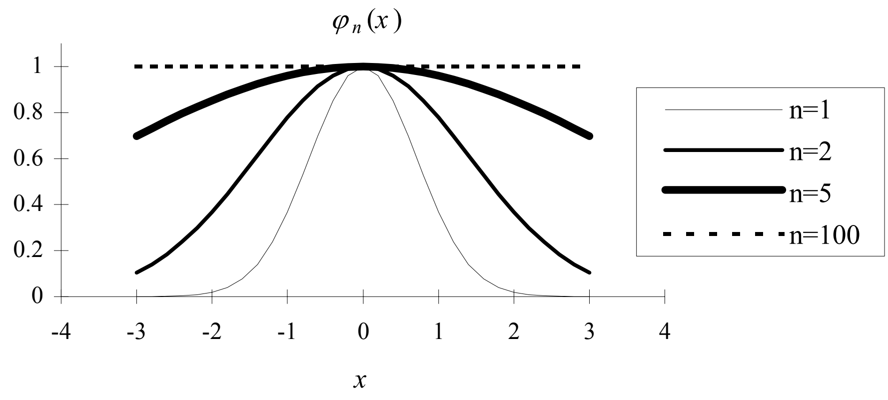
Ορισμός

Μία ακολουθία καλών συναρτήσεων $\varphi_n(x)$, $n = 1, 2, \ldots$, ονομάζεται **ομαλή** αν για κάθε καλή συνάρτηση $f(x)$ το όριο

$$\lim_{n\to\infty} \int_{-\infty}^{\infty} \varphi_n(x) f(x) dx \tag{2}$$

υπάρχει.

Παράδειγμα: Στο επόμενο σχήμα φαίνονται οι γραφικές παραστάσεις των συναρτήσεων $\varphi_n(x) = e^{-x^2/n^2}$, για $n = 1, 2, \ldots$



Αποδεικνύεται ότι η ακολουθία των $\varphi_n(x)$ είναι ομαλή και το όριο (2) υπάρχει:

$$\lim_{n\to\infty} \int_{-\infty}^{\infty} \varphi_n(x) f(x) dx = \int_{-\infty}^{\infty} f(x) dx$$

Ορισμός

Δύο ομαλές ακολουθίες $\varphi_n(x)$ και $\gamma_n(x)$ λέγονται **ισοδύναμες** αν, για κάθε καλή συνάρτηση $f(x)$, τα όρια (2) είναι ίσα, ήτοι

$$\lim_{n\to\infty} \int_{-\infty}^{\infty} \varphi_n(x) f(x) dx = \lim_{n\to\infty} \int_{-\infty}^{\infty} \gamma_n(x) f(x) dx$$

Παράδειγμα: Η ακολουθία $\varphi_n(x) = e^{-x^2/n^2}$ είναι ισοδύναμη με την $\gamma_n(x) = e^{-x^2/n^4}$. Προφανώς, οι δύο ακολουθίες έχουν όριο το 1 (ωστόσο, η απόδειξη της ισότητας των ορίων (2) παραλείπεται).

Α.5 Γενικευμένες Συναρτήσεις

Θεωρούμε μία ομαλή ακολουθία $\varphi_n(x)$. Είδαμε ότι το όριο (2) υπάρχει για κάθε καλή συνάρτηση $f(x)$.

Ορισμός

Μία **γενικευμένη συνάρτηση** $\varphi(x)$ ορίζεται ως το όριο κάποιας ομαλής ακολουθίας $\varphi_n(x)$ υπό την εξής έννοια: για κάθε καλή συνάρτηση $f(x)$, έχουμε

$$\lim_{n\to\infty}\int_{-\infty}^{\infty}\varphi_n(x)f(x)dx = \int_{-\infty}^{\infty}\varphi(x)f(x)dx \qquad (3)$$

Το αριστερό μέλος της ανωτέρω εξίσωσης είναι το όριο ακολουθίας κοινών ολοκληρωμάτων. Το δεξί ολοκλήρωμα έχει συμβολική έννοια: απεικονίζει τη θεμελιώδη ιδιότητα της γενικευμένης συνάρτησης $\varphi(x)$.

Ορισμός

Δύο γενικευμένες συναρτήσεις $\varphi(x)$ και $\gamma(x)$ είναι **ίσες** αν οι ακολουθίες $\varphi_n(x)$ και $\gamma_n(x)$ που τις ορίζουν είναι ισοδύναμες.

Παράδειγμα: Για την ακολουθία $\varphi_n(x) = e^{-x^2/n^2}$, είδαμε σε προηγούμενο παράδειγμα ότι το αριστερό μέλος της Εξ. (3) είναι ίσο με $\int_{-\infty}^{\infty} f(x)dx$. Θεωρούμε τώρα την ακολουθία $\gamma_n(x) = 1$ η οποία έχει όριο την $\gamma(x) = 1$. Θα ιδούμε αργότερα ότι τέτοιες συναρτήσεις μπορεί να θεωρηθούν ως γενικευμένες, παρ' όλο που οι $\gamma_n(x)$ δεν είναι καλές συναρτήσεις. Έχουμε λοιπόν

$$\lim_{n\to\infty}\int_{-\infty}^{\infty}\varphi_n(x)f(x)dx = \lim_{n\to\infty}\int_{-\infty}^{\infty}\gamma_n(x)f(x)dx = \int_{-\infty}^{\infty}1f(x)dx$$

Συνεπώς $\varphi(x) = \gamma(x) = 1$. Παρατηρήστε ότι το όριο $\lim_{n\to\infty} e^{-x^2/n^2}$ υπάρχει σ' αυτήν την περίπτωση και ισούται επίσης με $e^0 = 1$. Εδώ συμβαίνει οι τελεστές όριο και ολοκλήρωμα να εναλλάσσονται: $\lim_{n\to\infty}\int_{-\infty}^{\infty}\varphi_n(x)f(x)dx = \int_{-\infty}^{\infty}\left[\lim_{n\to\infty}\varphi_n(x)\right]f(x)dx$.

Εν γένει, τα όρια ομαλών ακολουθιών και οι γενικευμένες συναρτήσεις δεν είναι τυπικές συναρτήσεις και συχνά δεν έχουν φυσική έννοια. Επί πλέον το όριο ακολουθίας καλών συναρτήσεων δεν είναι απαραίτητα καλή συνάρτηση. Για παράδειγμα η συνάρτηση $\varphi(x) = 1$ δεν είναι καλή (όταν το $x \to \infty$, αυτή δεν τείνει στο 0).

A.6 Δέλτα του Dirac

Θεωρήστε την συνάρτηση πυκνότητας πιθανότητας της κανονικής κατανομής με μέση τιμή μ και διασπορά σ^2

$$\frac{1}{\sqrt{2\pi\sigma^2}} e^{-\frac{(x-\mu)^2}{2\sigma^2}}$$

Αντικαθιστώντας $\mu = 0$ και $\sigma^2 = \dfrac{1}{2n}$, $n = 1, 2, \ldots$ προκύπτει μία ακολουθία συναρτήσεων τις οποίες συμβολίζουμε $\delta_n(x)$. Αποδεικνύεται ότι η ακολουθία αυτή είναι ομαλή.

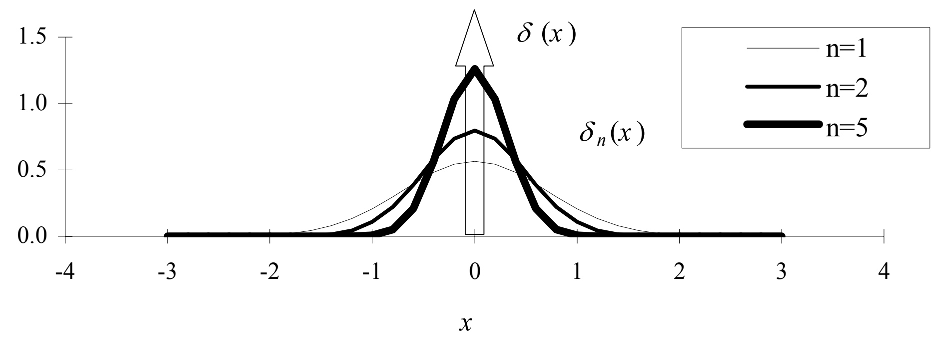
Ορισμός

Η ομαλή ακολουθία

$$\delta_n(x) = \frac{1}{\sqrt{\pi/n}} e^{-nx^2}, \; n = 1, 2, \ldots$$

ορίζει, σύμφωνα με τον ορισμό του Εδαφίου A.5, τη γενικευμένη συνάρτηση $\delta(x)$ που ονομάζεται **δέλτα του Dirac**.

Γραφικές παραστάσεις των $\delta_n(x)$ και $\delta(x)$ φαίνονται στο επόμενο σχήμα.



Ιδιότητα: Μπορεί να αποδειχθεί ότι

$$\lim_{n\to\infty}\int_{-\infty}^{\infty}\delta_n(x)f(x)dx = f(0)$$

για κάθε καλή συνάρτηση $f(x)$ οπότε

$$\int_{-\infty}^{\infty}\delta(x)f(x)dx = f(0) \tag{4}$$

Με απλά λόγια, **η $\boldsymbol{\delta(x)}$ δειγματοληπτεί την $\boldsymbol{f(x)}$ στο σημείο 0**.

Γενικότερος Ορισμός

Ορίζεται ως δέλτα του Dirac η γενικευμένη συνάρτηση $\delta(x)$ για την οποία η Εξ. (4) ισχύει για κάθε συνάρτηση $f(x)$ συνεχή στο σημείο 0.

Κάθε καλή συνάρτηση είναι και συνεχής. Συνεπώς, ο γενικότερος ορισμός έχει ευρύτερες εφαρμογές.

Α.7 Παράγωγος Γενικευμένης Συναρτήσεως

Έστω μία ομαλή ακολουθία $\varphi_n(x)$ η οποία ορίζει την γενικευμένη συνάρτηση $\varphi(x)$.

Ορισμός

Η παράγωγος γενικευμένης συνάρτήσεως $\varphi(x)$ συμβολίζεται $\varphi'(x)$ και είναι μία γενικευμένη συνάρτηση που ορίζεται ως το όριο των παραγώγων των $\varphi_n(x)$ υπό την εξής έννοια: για κάθε καλή συνάρτηση $f(x)$, ορίζουμε

$$\int_{-\infty}^{\infty} \varphi'(x) f(x) dx = \lim_{n\to\infty} \int_{-\infty}^{\infty} \varphi_n'(x) f(x) dx \tag{5}$$

εφ' όσον το όριο υπάρχει.

Το επόμενο αποτέλεσμα δίνει τη δυνατότητα να υπολογίζουμε το όριο της Εξ. (5) από τις αρχικές συναρτήσεις $\varphi_n(x)$.

Ιδιότητα: Για κάθε καλή συνάρτηση $f(x)$,

$$\int_{-\infty}^{\infty} \varphi'(x) f(x) dx = -\int_{-\infty}^{\infty} \varphi(x) f'(x) dx \tag{6}$$

Απόδειξη: Έχουμε, εξ ορισμού

$$\int_{-\infty}^{\infty} \varphi'(x) f(x) dx = \lim_{n\to\infty} \int_{-\infty}^{\infty} \varphi_n'(x) f(x) dx$$

$$= \lim_{n\to\infty} f(x)\varphi_n(x)\Big|_{-\infty}^{\infty} - \lim_{n\to\infty} \int_{-\infty}^{\infty} \varphi_n(x) f'(x) dx$$

Ο πρώτος όρος είναι 0 γιατί το γινόμενο καλών συναρτήσεων f και φ_n είναι καλή συνάρτηση. Η f' είναι καλή, ως παράγωγος καλής συνάρτησης. Επομένως

$$\int_{-\infty}^{\infty} \varphi'(x) f(x) dx = -\int_{-\infty}^{\infty} \varphi(x) f'(x) dx$$

Παράδειγμα: $\int_{-\infty}^{\infty} \delta'(x) f(x) dx = -\int_{-\infty}^{\infty} \delta(x) f'(x) dx = -f'(0).$

Α.8 Επέκταση της Έννοιας των Γενικευμένων Συναρτήσεων

Εξετάζουμε τώρα κοινές συναρτήσεις $\varphi(x)$ για τις οποίες υπάρχει κάποιος φυσικός αριθμός N τέτοιος ώστε

$$\int_{-\infty}^{\infty} \left| \frac{1}{(1+x^2)^N} \varphi(x) \right| dx < \infty \quad ^{(*)} \tag{7}$$

Τέτοιες συναρτήσεις είναι αρκετά καλές αλλά **όχι** πάντα καλές. Το ερώτημα που τίθεται είναι αν μπορούμε να θεωρήσουμε και τέτοιες συναρτήσεις ως γενικευμένες. Αποδεικνύεται ότι αυτό ισχύει, δηλαδή αν η συνάρτηση $\varphi(x)$ ικανοποιεί τη συνθήκη (7) τότε υπάρχει ομαλή ακολουθία $\varphi_n(x)$ τέτοια ώστε

$$\lim_{n\to\infty} \int_{-\infty}^{\infty} \varphi_n(x) f(x) dx = \int_{-\infty}^{\infty} \varphi(x) f(x) dx$$

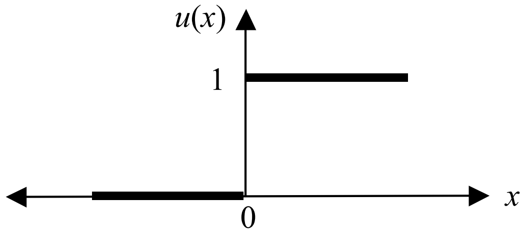
για κάθε καλή συνάρτηση $f(x)$.

Παράδειγμα: Η **βηματική** συνάρτηση ή συνάρτηση **Heaviside** ορίζεται

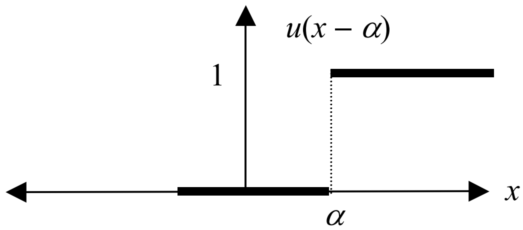
(*) αν ισχύει αυτό, τότε η συνάρτηση $\frac{\varphi(x)}{(1+x^2)^N}$ λέγεται **απολύτως ολοκληρώσιμη στο διάστημα (–∞, ∞)**.

$$u(x) = \begin{cases} 1, & x \geq 0 \\ 0, & x < 0 \end{cases}$$

Γραφικά,



Επίσης,



Η συνάρτηση $u(x)$ είναι κοινή αλλά μπορεί να θεωρηθεί και ως γενικευμένη γιατί ικανοποιεί τη συνθήκη (7). Πράγματι, για $N = 1$,

$$\int_{-\infty}^{\infty} \left| \frac{1}{(1+x^2)^1} u(x) \right| dx = \int_0^{\infty} \frac{1}{1+x^2} dx = \int_0^{\infty} d(\tan^{-1} x) = \tan^{-1}\infty - \tan^{-1}0 = \frac{\pi}{2} - 0 < \infty$$

Είδαμε ότι δύο γενικευμένες συναρτήσεις $\varphi(x)$ και $\psi(x)$ είναι ίσες αν οι ακολουθίες $\varphi_n(x)$ και $\psi_n(x)$ από τις οποίες ορίζονται είναι ισοδύναμες. Από αυτόν τον ορισμό προκύπτει το επόμενο σημαντικό αποτέλεσμα.

Ιδιότητα: $u'(x) = \delta(x)$

Απόδειξη: Η παράγωγος $u'(x)$ της βηματικής συνάρτησης ικανοποιεί την Εξ. (6),

$$\int_{-\infty}^{\infty} u'(x) f(x) dx = -\int_{-\infty}^{\infty} u(x) f'(x) dx$$

και, από τον ορισμό της $u(x)$,

$$- \int_{-\infty}^{\infty} u(x) f'(x) dx = - \int_{0}^{\infty} f'(x) dx = f(0) - f(\infty) = f(0)$$

που είναι το ίδιο αποτέλεσμα με εκείνο της Εξ. (4). Επομένως $u'(x) = \delta(x)$.

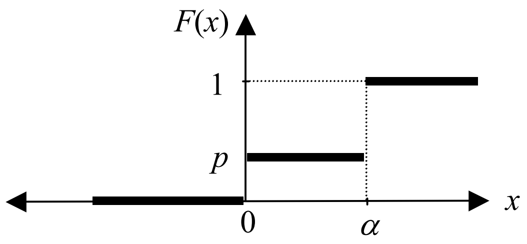
Παράδειγμα: Μία τυχαία μεταβλητή X ορίζεται

$$X = \begin{cases} 0, \text{ με πιθανότητα } p \\ \alpha, \text{ με πιθανότητα } 1-p \end{cases}$$

Η συνάρτηση κατανομής της είναι

$$F(x) = P(X \le x) = \begin{cases} 1, \text{ για } x \ge \alpha \\ p, \text{ για } 0 \le x < \alpha \\ 0, \text{ για } x < 0 \end{cases}$$

για κάθε $x \in (-\infty, \infty)$. Γραφικά,



Η $F(x)$ παρουσιάζει ένα άλμα ύψους p στο σημείο $x = 0$ και ακόμη ένα ύψους $(1 - p)$ στο $x = \alpha$. Εύκολα επαληθεύεται ότι

$$F(x) = pu(x) + (1-p)\, u(x-\alpha)$$

Συνεπώς έχουμε

$$f(x) = F'(x) = p\delta(x) + (1-p)\, \delta(x-\alpha)$$

A.9 Μετατόπιση της Αρχής των Αξόνων

Η συνάρτηση $\delta(x - \alpha)$:

$$\int_{-\infty}^{\infty} \delta(x-\alpha) f(x) dx \underset{\substack{\text{θέτουμε} \\ y=x-\alpha}}{=} \int_{-\infty}^{\infty} \delta(y) f(y+\alpha) dy$$

$$= f(0 + \alpha) = f(\alpha)$$

Τα όρια των ανωτέρω ολοκληρωμάτων είναι ίδια αφού όταν $x \to \pm\infty$ τότε και $y \to \pm\infty$. Με απλά λόγια, **η $\delta(x - \alpha)$ δειγματοληπτεί την $f(x)$ στο σημείο $x = \alpha$** .

Η συνάρτηση $u(x - \alpha)$:

$$\int_{-\infty}^{\infty} u(x-\alpha) f(x) dx \underset{\substack{\text{θέτουμε} \\ y=x-\alpha}}{=} \int_{-\infty}^{\infty} u(y) f(y+\alpha) dy$$

$$= \int_{0}^{\infty} f(y+\alpha) dy = \int_{0+\alpha}^{\infty} f(x) dx = \int_{\alpha}^{\infty} f(x) dx$$

A.10 Μετασχηματισμός Fourier

Ο μετασχηματισμός Fourier είναι μία αποτελεσματική μέθοδος επίλυσης προβλημάτων. Εφαρμόζεται στην ανάλυση δυναμικών συστημάτων και διαδικασιών και στη θεωρία πιθανοτήτων. Ο υπολογισμός μέσης τιμής, διασποράς και άλλων στατιστικών παραμέτρων μίας τυχαίας μεταβλητής διευκολύνεται με τη χρήση της χαρακτηριστικής συνάρτησης και την εφαρμογή του Θεωρήματος των Ροπών. Η χαρακτηριστική συνάρτηση αποτελεί μία ειδική μορφή του μετασχηματισμού Fourier της συνάρτησης πυκνότητας πιθανότητας $f(x)$ της τυχαίας μεταβλητής.

Θεωρούμε μία πραγματική συνάρτηση $f(x)$ που ορίζεται για κάθε $x \in (-\infty, \infty)$. Η μιγαδική συνάρτηση

$$\varphi(\omega) = \int_{-\infty}^{\infty} e^{-i\omega t} f(x)dx\,,\ \ \omega\in(-\infty,\infty)$$

εφ' όσον το ολοκλήρωμα υπάρχει, ορίζεται ως ο μετασχηματισμός Fourier της $f(x)$. Συχνά γράφουμε $f(x) \overset{\mathbf{F}}{\leftrightarrow} \varphi(\omega)$ ή $\mathbf{F}[f(x)] = \varphi(\omega)$. Αν η συνάρτηση $f(x)$ να είναι απολύτως ολοκληρώσιμη στο $(-\infty, \infty)$ τότε το ολοκλήρωμα υπάρχει. Στην περίπτωση αυτή η $f(x)$ μηδενίζεται καθώς $x\to \pm\infty$. Ας σημειωθεί ότι οι συνθήκες αυτές **δεν είναι αναγκαίες** για την ύπαρξη του μετασχηματισμού Fourier.

Αντίστροφα, αν $\varphi(\omega)$ είναι μία μιγαδική συνάρτηση τότε, υπό ορισμένες συνθήκες, η συνάρτηση $f(x)$ υπάρχει, είναι μοναδική, και δίδεται από τον αντίστροφο μετασχηματισμό Fourier

$$f(x) = \mathbf{F}^{-1}[\varphi(\omega)] = \frac{1}{2\pi}\int_{-\infty}^{\infty} e^{i\omega t}\varphi(\omega)d\omega\,,\ \ x\in(-\infty,\infty)$$

Μερικές ιδιότητες του μετασχηματισμού Fourier:

α) $\mathbf{F}[\delta(x)] = \int_{-\infty}^{\infty} e^{-i\omega t}\delta(x)dx\ = e^{-i\omega 0} = 1.$

β) Θεωρούμε τη συνάρτηση $F(x)$ που είναι διαφορίσιμη και τείνει στο 0 για $|x|\to\infty$. Αν $\mathbf{F}[F(x)] = \Phi(\omega)$ και $\mathbf{F}[F'(x)] = \varphi(\omega)$, τότε

$$\varphi(\omega) = i\omega\Phi(\omega)$$

Πράγματι, επειδή η συνάρτηση $e^{-i\omega x}$ είναι απολύτως φραγμένη ($|e^{-i\omega x}| = \sqrt{\cos^2(\omega x)+\sin^2(\omega x)} = 1$) προκύπτει

$$\varphi(\omega) = \int_{-\infty}^{\infty} e^{-i\omega x}F'(x)dx$$

$$= F(x)e^{-i\omega x}\Big|_{-\infty}^{\infty} - (-i\omega)\int_{-\infty}^{\infty} F(x)e^{-i\omega x}dx\ =\ (\text{μηδενική})\times(\text{φραγμένη}) + i\omega\Phi(\omega)$$

$$= i\omega\Phi(\omega)$$

Αντιστρόφως,

$$\Phi(\omega) = \frac{\varphi(\omega)}{i\omega}$$

γ) Η συνθήκη $F(x)\to 0$ για $|x|\to\infty$ δεν είναι αναγκαία για την ύπαρξη μετασχηματισμού Fourier. Για παράδειγμα, επειδή $u'(x) = \delta(x)$, από τις α) και β) προκύπτει

$$\mathbf{F}[u(x)] = \frac{1}{i\omega}$$

Παρατηρούμε ότι αν και η $u(x)$ δεν συγκλίνει προς το 0 για $x\to\infty$, ο μετασχηματισμός Fourier υπάρχει.

δ) $\mathbf{F}[f(x-\alpha)] = \int_{-\infty}^{\infty} e^{-i\omega x} f(x-\alpha)dx \underset{\substack{\text{θέτουμε} \\ x=y+\alpha}}{=} \int_{-\infty}^{\infty} e^{-i\omega y} e^{-i\omega\alpha} f(y)dy = e^{-i\omega\alpha}\,\mathbf{F}[f(x)]$

ε) Η συνέλιξη είναι μία πράξη μεταξύ δύο συναρτήσεων $f(x)$ και $g(x)$ που ορίζεται

$$f(x)\otimes g(x) = \int_{-\infty}^{\infty} f(x-\tau)g(\tau)d\tau = \int_{-\infty}^{\infty} f(\tau)g(x-\tau)d\tau$$

Από τη συνέλιξη προκύπτει μία νέα συνάρτηση με μετασχηματισμό Fourier

$$\mathbf{F}[f(x)\otimes g(x)] = \int_{-\infty}^{\infty}\left[\int_{-\infty}^{\infty} f(x-\tau)g(\tau)d\tau\right]e^{-i\omega x}dx$$

$$= \int_{-\infty}^{\infty}\left[\int_{-\infty}^{\infty} f(x-\tau)g(\tau)e^{-i\omega(x-\tau+\tau)}d\tau\right]dx$$

Αλλάζοντας τη σειρά ολοκλήρωσης και αντικαθιστώντας $x = y + \tau$, ευρίσκουμε

$$\mathbf{F}[f(x)\otimes g(x)] = \int_{-\infty}^{\infty}\left[\int_{-\infty}^{\infty} f(y)e^{-i\omega y}g(\tau)e^{-i\omega\tau}d(y+\tau)\right]d\tau$$

Το μεσαίο ολοκλήρωμα υπολογίζεται για σταθερό τ και μεταβαλλόμενο y, οπότε $d(y + \tau) = dy$ και

$$\mathbf{F}[f(x)\otimes g(x)] = \int_{-\infty}^{\infty}\left[\int_{-\infty}^{\infty} f(y)e^{-i\omega y}dy\right]g(\tau)e^{-i\omega\tau}d\tau$$

$$= \int_{-\infty}^{\infty} \mathbf{F}[f(x)]\ g(\tau)e^{-i\omega\tau}d\tau$$

$$= \mathbf{F}[f(x)]\ \mathbf{F}[g(x)]$$

Η συνέλιξη είναι πράξη μεταβατική:

$$f(x)\otimes g(x) = \int_{-\infty}^{\infty} f(x-\tau)g(\tau)d\tau$$

$$\underset{\substack{\text{θέτοντας}\\ u=x-\tau\\ x=\text{σταθερό}}}{=} \int_{u=\infty}^{u=-\infty} g(x-u)f(u)d(-u) = \int_{u=-\infty}^{u=\infty} g(x-u)f(u)du \ = g(x)\otimes f(x)$$

Το ίδιο αποτέλεσμα προκύπτει εφαρμόζοντας αντίστροφο μετασχηματισμό σε κάθε μέλος της ταυτότητας $\mathbf{F}[f(x)]\ \mathbf{F}[g(x)] = \mathbf{F}[g(x)]\ \mathbf{F}[f(x)]$.

στ) Η χαρακτηριστική συνάρτηση $\Phi(\omega)$ μίας τ.μ. με συνάρτηση πυκνότητας πιθανότητας $f(x)$ είναι ο μετασχηματισμός Fourier της $f(x)$ στο σημείο $-\omega$:

$$\Phi(\omega) = \int_{-\infty}^{\infty} e^{i\omega t} f(x)dx$$

Όλες οι ανωτέρω ιδιότητες ισχύουν και για την χαρακτηριστική συνάρτηση θέτοντας $-\omega$ στη θέση του ω.

ΠΡΟΒΛΗΜΑΤΑ

1. Θεωρήστε τη διαδικασία $X_t = e^{at}$ όπου η ποσότητα α είναι μια τυχαία μεταβλητή. Εύρετε την προσδοκητή τιμή $E(X_t)$, την αυτοσυσχέτιση $R(t_1, t_2)$ και την πυκνότητα $f(x,t)$ της X_t σαν συνάρτηση της πυκνότητας $f_\alpha(\mathrm{a})$ της μεταβλητής α.

2. Εύρετε την προσδοκητή τιμή και τη διασπορά του ολοκληρώματος $I = \int_{-a}^{a}[5cos\,3t + n_t]e^{-i\omega t}dt$, αν $E(n_t)$=0 και $R_n(\tau) = 2\delta(\tau)$.

3. Έστω η αμετάβλητη κατά τη στενή έννοια διαδικασία X_t και α μια τυχαία μεταβλητή ανεξάρτητη της X_t. Δείξετε ότι η διαδικασία $Y_t = X_{t-\alpha}$ είναι αμετάβλητη κατά τη στενή έννοια.

4. Έστω ότι η X_t είναι αμετάβλητη κατά την ευρεία έννοια και διαφορίσιμη κατά μέσο τετράγωνο. Να δειχθεί ότι η X_t και η $\dot{X}_t$ είναι ασυσχέτιστες και ορθογώνιες. (αν οι X, Y είναι ορθογώνιες $\Leftrightarrow E(XY)=0$).

5. Μια στοχαστική διαδικασία δίνεται από τη σχέση $X(t) = \cos(\omega t + \varphi)$, όπου φ είναι τυχαία μεταβλητή με χαρακτηριστική συνάρτηση $\Phi(\omega) = E(e^{i\omega\varphi})$, ω σταθερά, και $\Phi(1) = \Phi(2) = 0$. Να αποδειχθεί ότι η $X(t)$ είναι αμετάβλητη κατά την ευρεία έννοια.

6. Υπολογίστε την $S(\omega)$ αν $R(\tau) = e^{-\alpha\tau^2}$, όπου $\alpha \in \mathbb{R}^+$ είναι γνωστή σταθερά.

7. Ν' αποδειχθεί ότι, αν $Y_t = X_t \otimes h(t)$ και $R_{XX}(t_1, t_2) = q(t_1)\ \delta(t_1 - t_2)$, τότε $E(X_tY_t) = h(0)\ q(t)$.

ΒΙΒΛΙΟΓΡΑΦΙΑ

J. L. Doob, *Stochastic Processes*, Wiley, New York 1990.

A. Papoulis, *Probability, Random Variables, and Stochastic Processes*, McGraw-Hill, 3rd Ed., New York 1991.

E. Parzen, *Stochastic Processes*, Holden-Day, San Francisco 1962.

S. M. Ross, *Stochastic Processes*, Wiley, 2nd Ed., New York 1996.