

**ΤΕΧΝΟΛΟΓΙΚΟ ΕΚΠΑΙΔΕΥΤΙΚΟ ΙΔΡΥΜΑ ΚΡΗΤΗΣ**



**ΠΑΡΑΡΤΗΜΑ ΧΑΝΙΩΝ**



ΤΜΗΜΑ ΗΛΕΚΤΡΟΝΙΚΗΣ

---

ΠΤΥΧΙΑΚΗ ΕΡΓΑΣΙΑ

*"Ανίχνευση ασθενών σημάτων σε πρώτης τάξης εξαρτημένο θόρυβο"*

Όνομα φοιτητή

ΜΠΟΤΩΝΑΚΗΣ ΓΙΩΡΓΟΣ

*Επιβλέπων καθηγητής*

**ΚΟΚΚΙΝΟΣ ΕΥΑΓΓΕΛΟΣ, ΕΠΙΚΟΥΡΟΣ ΚΑΘΗΓΗΤΗΣ**

**ΧΑΝΙΑ 2010**

## **Ευχαριστίες**

Θα θελα να ευχαριστήσω θερμά τον καθηγητή μου ΚΟ. *ΚΟΚΚΙΝΟ ΕΥΑΓΓΕΛΟ* κυρίως για την εμπιστοσύνη, υπομονή και καθοδήγηση που έδειξε στο θέμα της πτυχιακής, αλλά και για τις πολύτιμες συμβουλές του για την μετέπειτα καριέρα μου.

Θα θελα επίσης να απευθύνω τις ευχαριστίες μου στους γονείς μου, οι οποίοι στήριξαν τις σπουδές μου με διάφορους τρόπους, φροντίζοντας για την καλύτερη δυνατή μόρφωση μου.

## **ΠΡΟΛΟΓΟΣ-ΠΕΡΙΛΗΨΗ**

Σκοπός αυτής της πτυχιακής είναι να μοντελοποιηθεί ο θόρυβος ο οποίος παρουσιάζει αδύναμη εξάρτηση, με το μοντέλο του 'κινητού αθροίσματος'-'Moving Average'.

Στην συνέχεια, αφού πραγματοποιηθεί η ανέραιση της εξάρτησης, να βρεθούν τα κατάλληλα κριτήρια απόφασης για εκτιμητές οι οποίοι έχουν σχεδιασθεί να λειτουργούν υπό Cauchy, Gaussian και Student-t περιβάλλον θορύβου, για την περίπτωση που η ισχύς εκπομπής είναι γνωστή στον δέκτη (Εκτίμηση Μέγιστης Πιθανοφάνειας-Maximum Likelihood Estimation) καθώς και για την περίπτωση που δεν είναι γνωστή (Sub-Optimum Likelihood Estimation)

Τελικά, να γίνει η σύγκριση της απόδοσης των εκτιμητών αυτών σύμφωνα με τα διαγράμματα Bit Error Rate.

## **INTRODUCTION-REVIEW**

The aim of this project, is the modelization of the noise that exhibits weak dependence, with the help of the 'Moving Average' Approach.

Next, after the necessary cancellation of the dependent noise samples, we will try to derive the appropriate decision rule for the estimators that were designed to work under Cauchy, Gaussian and Student-t noise environment. The decision rule will firstly consider the existance of the knowledge of the transmitted signal strength and secondly, the case this knowledge is not known to the receiver.

Finally, we will compare the efficiency of these estimators with the help of the Bitt Error Rate Diagramms.

ΕΥΧΑΡΙΣΤΙΕΣ.....	2
ΠΡΟΛΟΓΟΣ-ΠΕΡΙΛΗΨΗ.....	3
ΠΗΓΕΣ-REFERENCES.....	117

## ΕΝΟΤΗΤΑ I

- Εισαγωγή, χρησιμότητα και εφαρμογές της ανίχνευσης και εκτίμησης σήματος.....7
- Εφαρμογή της θεωρίας των πιθανοτήτων στην ανίχνευση και εκτίμηση σήματος. ....8
- Παρουσίαση του κριτηρίου Bayes και εννοιών όπως:
  - Κόστος.....10
  - Ρίσκο.....12
  - Παράδειγμα εύρεσης του αναλογικού τεστ (*likelihood ratio test*), λάθους συναγερμού (*false alarm*) και την πιθανότητα ανίχνευσης (*probability of detection*). ....15
  - Παράδειγμα εύρεσης κατωφλίου βασιζόμενοι σε περισσότερο του ενός δείγματα.....17
- Εκτίμηση Παραμέτρων του εκπεμπόμενου σήματος.....19
- Maximum Likelihood Estimation (MLE) – Εκτίμηση Μέγιστης Πιθανοφάνειας.....20
- Παράδειγμα (στην MLE) .....21

## ΕΝΟΤΗΤΑ II.....

- Εισαγωγή στην μη-γραμμική επεξεργασία σήματο.....24
- Λόγος χρήσης μη-γραμμικών φίλτρων έναντι γραμμικών.....26
- Πρακτικά παραδείγματα αλλοίωσης του σήματός μας από non-Gaussian θόρυβο (ADSL γραμμές) .....27
- Απόρροια της Gaussian κατανομής μέσω του Κεντρικού οριακού θεωρήματος και λόγος ύπαρξης μη-Gaussian κατανομών μέσω της αντιπαράθεσης Κεντρικού οριακού θεωρήματος και γενικευμένου Κεντρικού οριακού θεωρήματος.....29
- Διάφορες προσεγγίσεις στην αντιμετώπιση των *outliers-παρασίτων* (σταθμισμένοι μέσοι όροι, *weighted myriads*) .....30
- Το φιλτράρισμα και αναπαράσταση του με όρους στατιστικής, και παρουσίαση εννοιών όπως εκτίμηση θέσης (*Location estimation*).....32

- Μη-Gaussian μοντέλα και πως χρησιμοποιούμε την γενικευμένη Gaussian κατανομή για αντιμετώπιση διαδικασιών με μεγάλη κρουστικότητα.....**34**
- Stable κατανομές.....(δείτε και την παρουσίαση **Power Point**).....**36**
- Στατιστικές ιδιότητες εκτιμητών (μη-προκατηλημενοι εκτιμητές, αποδοτικοί εκτιμητές, παρουσίαση του ορίου Cramer and Rao Lower Bound).....**48**

### **ΕΝΟΤΗΤΑ ΙΙΙ**

- Παρατηρούμενος Χώρος .....**61**
- Εξαρτημένος θόρυβος-Αναπαράστασή του με σειρά Volterra.....**63**
- Συσχέτιση δύο σημάτων.....**65**
- Μετασχηματισμός Παρατηρούμενου χώρου με την βοήθεια του Ιακωβιανού Μετασχηματισμού-*Jacobian Transformation*.....**66**
- Παρουσίαση Gaussian, Cauchy και Student-t κατανομών...
- Εύρεση του καταλλήλου κριτηρίου ανάλογα με την κατανομή που ακολουθεί ο θόρυβος.....**78**
- Παρουσίαση Antipodal συστήματος εκπομπής.....**88**
- Αριθμητικά αποτελέσματα .....**91**

# ΕΝΟΤΗΤΑ Ι

## Καλύπτονται έννοιες:

- Εισαγωγή, χρησιμότητα και εφαρμογές της ανίχνευσης και εκτίμησης σήματος.
- Εφαρμογή της θεωρίας των πιθανοτήτων στην ανίχνευση και εκτίμηση σήματος
- Παρουσίαση του κριτηρίου Bayes και εννοιών όπως:
  - Κόστος
  - Ρίσκο
  - Παράδειγμα εύρεσης του αναλογικού τεστ (*likelihood ratio test*), λάθους συναγερμού (false alarm) και την πιθανότητα ανίχνευσης (probability of detection).
  - Παράδειγμα εύρεσης κατωφλίου βασιζόμενοι σε περισσότερα του ενός δείγματα.
- Εκτίμηση Παραμέτρων του εκπεμπόμενου σήματος
- Maximum Likelihood Estimation (MLE) – Εκτίμηση Μέγιστης Πιθανοφάνειας
- Παράδειγμα (στην MLE)

## *Εισαγωγή*

Γενικότερα μιλώντας, η ανίχνευση και εκτίμηση ασχολείται με την μελέτη της επεξεργασίας της πληροφορίας, η οποία μεταφέρεται με διάφορες μορφές σημάτων, με τελικό σκοπό να ανακτήσουμε την πληροφορία μέσα από τα σήματα αυτά. Εφαρμογές της θεωρίας της ανίχνευσης σήματος μπορούν να βρεθούν σε πολλά πεδία όπως επικοινωνίες και αυτόματος έλεγχος. Για παράδειγμα, στις τηλεπικοινωνίες, εφαρμογές όπως η μετάδοση σήματος, Radar, η ανίχνευση και εκτίμηση σήματος προσφέρει το θεωρητικό και αναλυτικό υπόβαθρο για τον σχεδιασμό αποτελεσματικών τηλεπικοινωνιακών δεκτών. Εναλλακτικά, στον τομέα του αυτοματισμού, η ανίχνευση και εκτίμηση σήματος οδηγεί σε τεχνικές στο να φτιαχτούν καλύτερα και με μεγαλύτερη ακρίβεια μηχανήματα, ανάλογα με τις συνθήκες που λαμβάνουν χώρα στο σύστημα που πρόκειται να ελεγχθεί.

Ένα παράδειγμα χρήσης της ανίχνευσης και εκτίμησης σήματος στα Radar, είναι όταν έχει εκπεμφθεί ένας ηλεκτρομαγνητικός παλμός και περιμένουμε την ηχώ για την υπόδειξη ενός στόχου. Λόγω των ατμοσφαιρικών ανωμαλιών, ηλεκτρονικός θόρυβος ο οποίος γεννιέται μέσα στα μηχανήματα, διαφόρων ειδών ανακλάσεις από άλλα αντικείμενα, μας οδηγούν σε μια πιο ασαφή εικόνα του στόχου, και έτσι επιστρατεύουμε την στατιστική, διότι πλέον μιλάμε για την πιθανότητα σωστής ανίχνευσης. Άρα η θεωρία της ανίχνευσης και εκτίμησης σήματος, μας βοηθά στο να διαλέξουμε την καλύτερη τεχνική για την ανίχνευση των σημάτων που μας ενδιαφέρουν. Επιπλέον, έχοντας καθορίσει με κάποιο βαθμό βεβαιότητας ότι ο στόχος είναι παρών, θα μπορούμε να υπολογίσουμε επιπλέον χαρακτηριστικά του στόχου, όπως ταχύτητα, ύψος, μέγεθος, κ.α.. Αυτές οι εκτιμήσεις είναι πολύ χρήσιμες στο να ελέγξουμε την κεραία για την συνεχή παρακολούθηση του στόχου, ή ακόμα και να ελέγξουμε εμείς οι ίδιοι τον στόχο, για την διατήρηση μιας επιθυμητής πορείας.

Άλλες συγκεκριμένες εφαρμογές της τεχνικής της ανίχνευσης και εκτίμησης σήματος είναι στον τομέα της σεισμολογίας, ράδιο-αστρονομία, sonar, επεξεργασία ήχου και εικόνας, επεξεργασία ιατρικών σημάτων και στις οπτικές επικοινωνίες.

Γενικότερα, οι εφαρμογές της ανίχνευσης και εκτίμησης σήματος περιλαμβάνουν την επινόηση των καταλληλότερων κριτηρίων για την αποδοτικότερη εκτίμηση σήματος το οποίο είναι εμπλουτισμένο διαφόρων ειδών προσθετικού θορύβου, ο οποίος δυσχεραίνει, την εκτίμηση του σήματος το οποίο είναι άγνωστο στον παρατηρητή, και ρίχνει την απόδοση του συστήματος λήψης.

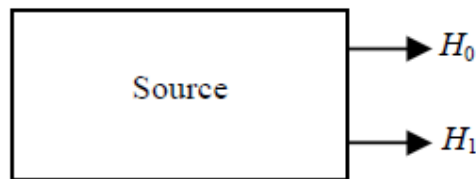
Βασικά στοιχεία της μελέτης της θεωρίας της ανίχνευσης και εκτίμησης σήματος, είναι η έννοια μιας τυχαίας μεταβλητής  $Y$  η οποία παίρνει τυχαίες τιμές μέσα από ένα παρατηρούμενο σετ  $\Gamma$ , το οποίο μπορεί να είναι ένα σετ από πίνακες, κυματομορφές, πραγματικοί αριθμοί, ή κάποιο άλλο σετ και στην ουσία περιλαμβάνει κάθε πιθανή μορφή σήματος που μπορούμε να λάβουμε από την πηγή. Από την παρατήρηση  $Y$  μπορούμε να ανακτήσουμε πληροφορία σχετικά με τα

φαινόμενα που ‘διαμόρφωσαν’ την  $Y$ . Υπάρχουν δύο τύποι προβλημάτων τα οποία μας ενδιαφέρουν:

- *προβλήματα ανίχνευσης*, όπου επιθυμούμε να διαλέξουμε μέσα από ένα πεπερασμένο αριθμό δυνατών παρατηρήσεων, την σωστή ‘αντιστοιχία’ παρατήρησης-εκπεμπόμενου σήματος.
- *Προβλήματα εκτίμησης* όπου επιθυμούμε να εκτιμήσουμε την τιμή/τιμές μιας ποσότητας/ποσοτήτων όπου η παρατήρηση έγινε με έμμεσο τρόπο.

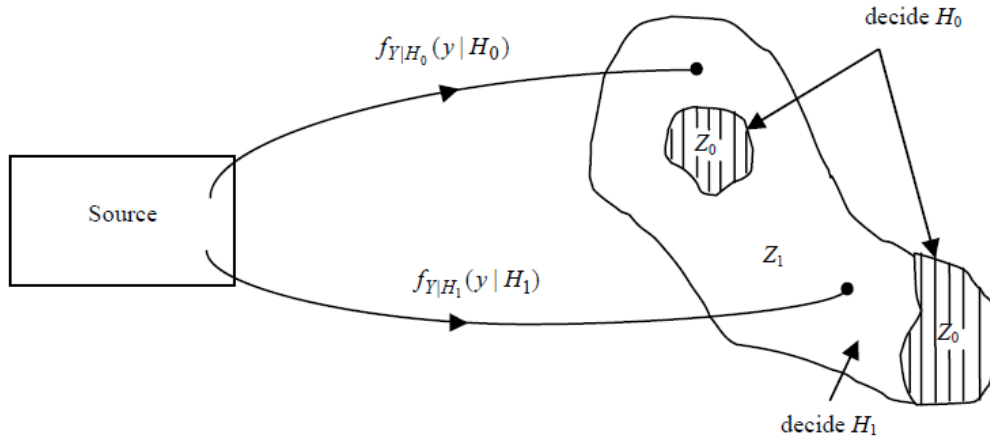
### Εφαρμογή της θεωρίας των πιθανοτήτων στην ανίχνευση και εκτίμηση σήματος

Σε κάθε περίπτωση η σύνδεση ανάμεσα στο παρατηρούμενο δείγμα και το σήμα που εκπέμφθηκε, μετριέται με όρους πιθανότητας. Και αυτό γιατί η στατιστική συμπεριφορά της  $Y$ , επηρεάζεται την στιγμιαία κατάσταση (στιγμιαία διότι οι παραμορφωτικές διαδικασίες είναι τυχαίες και έτσι έχουν κάποια κατανομή πιθανότητας ) της (ή και περισσότερες της μίας) παραμορφωτικής διαδικασίας. Και έτσι βλέπουμε ότι το μοντέλο για το παρατηρούμενο σέτ  $\Gamma$ , περιλαμβάνει μια οικογένεια από κατανομές πιθανότητας, του οποίου οι όροι οι οποίοι αντιστοιχούν στην στιγμιαία τιμή που έλαβαν μέσα από την κατανομή που τις διέπει την στιγμή της παρατήρησης. Το πρόβλημα είναι να βρούμε το μοντέλο το οποίο είναι κατάλληλο για την περιγραφή της κατανομής του εισερχόμενου σήματος, γιατί το κατώφλι που θα θέσουμε για την εκτίμηση, είναι διαφορετικό για κάθε κατανομή. Σε περίπτωση που γίνει λάθος μοντελοποίηση, ως συνέπεια θα έχουμε να τεθεί σε λάθος τιμή το κατώφλι, με δραματικές συνέπειες στην απόδοση του δέκτη. Ένα άλλο στοιχείο που θα μας απασχολήσει είναι η περίπτωση να έχουμε ή και να μη έχουμε στην γνώση μας οποιοδήποτε στοιχείο το οποίο χαρακτηρίζει την κατανομή του εκπεμφθέντος σήματος. Αυτήν την γνώση την ονομάζουμε *prior knowledge* (prior διότι εκφράζει πληροφορία για την ισχύ αποστολής αλλά και με τι πιθανότητα εκπέμπεται το κάθε σήμα).



*Παρατηρούμε μια πηγή η οποία εκπέμπει δύο διαφορετικά σήματα. Η γνώση της ισχύς αποστολής του έκαστο σήματος, αλλά και με τι πιθανότητα θα αποσταλεί το έκαστο σήμα, ονομάζεται prior knowledge.*





Στο σχήμα αυτό παρατηρούμε τις πιθανοφάνειες  $f_{Y|H_0}(y|H_0)$  για το σήμα '0' και  $f_{Y|H_1}(y|H_1)$  για το σήμα '1'. Επίσης εάν το άμορφο σήμα δηλώνει πιθανότητα, τότε ο όγκος του θα ισούται με '1'. Οι περιοχές  $Z_0$  και  $Z_1$  δηλώνουν την πιθανότητα να εκτιμηθεί το εκπεμπόμενο σήμα '0', ή '1'. Με βάση το τελευταίο, βλέπουμε από το σχήμα ότι το σήμα '1' έχει σωστά εκτιμηθεί σαν '1' από τον δέκτη, ενώ το σήμα '0' έχει εκτιμηθεί λάθος σαν '1'. Εάν το βελάκι  $f_{Y|H_0}(y|H_0)$  κατέληγε σε μια από τις δύο περιοχές  $Z_0$ , τότε θα είχα σωστή εκτίμηση και για το '0'.

Οι πιθανές εκδοχές σε ένα δυαδικό σύστημα είναι λοιπόν:

1. Να εκτιμήσουμε το σήμα σαν  $H_0$  ενώ το είναι  $H_0$  σωστό.
2. Να εκτιμήσουμε το σήμα σαν  $H_0$  ενώ το είναι  $H_1$  σωστό.
3. Να εκτιμήσουμε το σήμα σαν  $H_1$  ενώ το είναι  $H_0$  σωστό.
4. Να εκτιμήσουμε το σήμα σαν  $H_1$  ενώ το είναι  $H_1$  σωστό.

Προσέξτε ότι στις περιπτώσεις (1) και (4), ο δέκτης λαμβάνει την σωστή απόφαση, ενώ στις περιπτώσεις (2) και (3), ο δέκτης κάνει λάθος. Σε ένα σύστημα Radar, η περίπτωση (2) λέγεται *miss*, η περίπτωση (3) λέγεται *false alarm* και η περίπτωση (4) λέγεται *detection*.

Σε αυτό το κεφάλαιο θα παρουσιάσουμε διάφορες τεχνικές που χρειάζονται για να λύσουμε προβλήματα ανίχνευσης σήματος.

## ΤΟ ΚΡΙΤΗΡΙΟ BAYES

Στην χρήση του κριτηρίου Bayes, λαμβάνουμε υπόψη μας δύο υποθέσεις. Πρώτον, να γνωρίζουμε με τι πιθανότητα θα εκπεμφθεί το σήμα '0' και '1'. Αυτές οι πιθανότητες λέγονται *priori* πιθανότητες,  $P(H_0)$  και  $P(H_1)$  αντίστοιχα. Δοσμένες τις *priori* πιθανότητες, και γνωρίζοντας ότι το σήμα που θα εκπεμφθεί, θα είναι '0' ή '1', έχουμε:

$$P_0 + P_1 = 1$$

Υπάρχει όμως η πιθανότητα να γίνει λάθος εκτίμηση του σήματος που λαμβάνω, και έτσι ανάλογα με την εφαρμογή θα έχω και ανάλογο κόστος. Για παράδειγμα, δεν μπορώ να έχω το ίδιο κόστος σε μια αναπαραγωγή τραγουδιού και σε μια στρατιωτική εφαρμογή όπου μπορεί να υπάρχει στόχος, και το Radar να μην μου τον υποδείξει. Θέτουμε λοιπόν  $D_i$   $i=0,1$ , όπου το  $D_0$  δηλώνει την ενέργεια 'αποφάσισε ότι το λαμβανόμενο σήμα είναι το '0'' και όπου το  $D_1$  δηλώνει την ενέργεια 'αποφάσισε ότι το λαμβανόμενο σήμα είναι το '1''. Επίσης θέτουμε το κόστος ανάλογα με την εκτίμηση που έχει ήδη γίνει στον δέκτη  $C_{i,j}$   $i,j=0,1$  το οποίο δίνει το κόστος που θα έχουμε στην περίπτωση που η υπόθεση  $H_j$  είναι σωστή και η εκτίμηση που έχει κάνει ο δέκτης είναι η  $C_i$ .

$$P(\text{cost } C_{ij}) = P(\text{choose } D_i \text{ when } H_j \text{ is true}) \quad i,j=0,1$$

Τα πιθανά κόστη θα είναι λοιπόν:

1.  $C_{00}$  για την περίπτωση (1).
2.  $C_{01}$  για την περίπτωση (2).
3.  $C_{10}$  για την περίπτωση (3).
4.  $C_{11}$  για την περίπτωση (4).

Ο στόχος μας είναι να βρούμε εκείνο το κατάφλι απόφασης έτσι όπου το μέσο κόστος  $E[C]$  το οποίο είναι γνωστό και σαν *ρίσκο*  $\mathfrak{R}$  να είναι το ελάχιστο. Η πράξη  $E[C]$  δηλώνει αναμενόμενη τιμή. Έχει επίσης υποθεθεί ότι το κόστος να παρθεί λάθος απόφαση, είναι μεγαλύτερη από το να παρθεί σωστή απόφαση.

$$C_{01} > C_{11}$$

Και

$$C_{10} > C_{00}$$

Εάν γνωρίζουμε την τομή των πιθανοτήτων  $P(D_i)$  και  $P(H_j)$ , δηλαδή να γνωρίζουμε την *joint probability*  $P(D_i, H_j)$  για κάθε συνδυασμό των  $i$  και  $j$ , τότε το μέσο κόστος θα είναι:

$$\mathfrak{R} = E[C] = C_{00}P(D_0, H_0) + C_{01}P(D_0, H_1) + C_{10}P(D_1, H_0) + C_{11}P(D_1, H_1)$$

Και από τον κανόνα του *Bayes* έχουμε:

$$P(D_i, H_j) = P(D_i | H_j)P(H_j)$$

Οι υπο-συνθήκη πιθανότητες  $P(D_i | H_j)$ ,  $i, j = 0, 1$  τότε για τις περιοχές που φαίνονται στο σχήμα 2, έχουμε:

- **Υπο-συνθήκη πιθανότητα:** δηλώνει το πόσο πιθανό είναι να γνωρίζουμε την  $H_j$  και εκτίμηση να είναι η  $D_i$ .

$$P(D_0 | H_0) \equiv P(\text{decide } H_0 | H_0 \text{ true}) = \int_{Z_0} f_{Y|H_0}(y | H_0) dy \quad (1)$$

$$P(D_0 | H_1) \equiv P(\text{decide } H_0 | H_1 \text{ true}) = \int_{Z_0} f_{Y|H_1}(y | H_1) dy \quad (2)$$

$$P(D_1 | H_0) \equiv P(\text{decide } H_1 | H_0 \text{ true}) = \int_{Z_1} f_{Y|H_0}(y | H_0) dy \quad (3)$$

$$P(D_1 | H_1) \equiv P(\text{decide } H_1 | H_1 \text{ true}) = \int_{Z_1} f_{Y|H_1}(y | H_1) dy \quad (4)$$

Οι πιθανότητες  $P(D_0 | H_1)$ ,  $P(D_1 | H_0)$  και  $P(D_1 | H_1)$  δηλώνουν την πιθανότητα λάθους  $P_M$ , την πιθανότητα λάθους συναγερωμού (σε σύστημα Radar)  $P_F$  και πιθανότητα συναγερωμού  $P_D$  αντίστοιχα. Παρατηρούμε επίσης ότι:

$$P_M = 1 - P_D$$

Και

$$P(D_0 | H_0) = 1 - P_F$$

Συνεπώς η πιθανότητα να έχουμε σωστή ανίχνευση είναι:

$$\begin{aligned} P(\text{correct decision}) &= P(c) = P(D_0, H_0) + P(D_1, H_1) \\ &= P(D_0 | H_0)P(H_0) + P(D_1 | H_1)P(H_1) \\ &= (1 - P_F)P_0 + P_D P_1 \end{aligned}$$

## Ανίχνευση ασθενών σημάτων σε πρώτης τάξης εξαρτημένο θόρυβο

Και η πιθανότητα λάθους είναι:

$$\begin{aligned} P(\text{error}) = P(\varepsilon) &= P(D_0, H_1) + P(D_1, H_0) \\ &= P(D_0 | H_1)P(H_1) + P(D_1 | H_0)P(H_0) \\ &= P_M P_1 + P_F P_0 \end{aligned}$$

Ένα γράφημα με την συνάρτηση πιθανότητας λάθους  $P_c(c)$ , φαίνεται στο σχήμα 3.

Το μέσο κόστος τώρα γίνεται:

$$\mathfrak{R} = E[C] = C_{00}(1 - P_F)P_0 + C_{01}(1 - P_D)P_1 + C_{10}P_F P_0 + C_{11}P_D P_1$$

Το επόμενο βήμα για να βγει το μέσο κόστος είναι να αντικαταστήσουμε τους όρους πιθανότητας λάθους, και ανίχνευσης με τις περιοχές που δείχνονται στα ολοκληρώματα (1)~(4).

$$\begin{aligned} \mathfrak{R} &= P_0 C_{00} \int_{Z_0} f_{Y|H_0}(y | H_0) dy + P_1 C_{01} \int_{Z_0} f_{Y|H_1}(y | H_1) dy \\ &\quad + P_0 C_{10} \int_{Z_1} f_{Y|H_0}(y | H_0) dy + P_1 C_{11} \int_{Z_1} f_{Y|H_1}(y | H_1) dy \end{aligned} \quad (5)$$

Αν χρησιμοποιήσουμε και το γεγονός ότι το άθροισμα των πιθανοτήτων ισούται με:

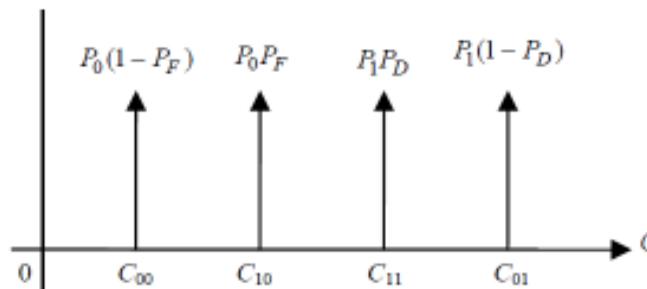
$$\int_Z f_{Y|H_0}(y | H_0) dy + \int_Z f_{Y|H_1}(y | H_1) dy = 1$$

Τότε μπορούμε στην (5) να αντικαταστήσουμε όπου:

$$\int_{Z_1} f_{Y|H_j}(y | H_j) dy = 1 - \int_{Z_0} f_{Y|H_j}(y | H_j) dy, \quad j = 0, 1$$

Έτσι έχουμε:

$$\begin{aligned} \mathfrak{R} &= P_0 C_{10} + P_1 C_{11} \\ &\quad + \int_{Z_0} \{ [P_1(C_{01} - C_{11})f_{Y|H_1}(y | H_1)] - [P_0(C_{10} - C_{00})f_{Y|H_0}(y | H_0)] \} dy \end{aligned}$$



Συναρτήσεις πυκνότητας κόστους

## Ανίχνευση ασθενών σημάτων σε πρώτης τάξης εξαρτημένο θόρυβο

Παρατηρούμε ότι η ποσότητα  $P_0 C_{10} + P_1 C_{11}$  είναι μια σταθερά και ανεξάρτητη από τα κόστη που θα έχουμε ορίσει για τον παρατηρούμενο χώρο, και η μόνη μεταβλητή που είναι ικανή να αλλάξει το κόστος, είναι η περιοχή  $Z_0$ , και αυτό γιατί, αλλάζοντας το 'εύρος' της περιοχής αυτής, επηρεάζουμε το πόσο μεγάλος θα είναι ο δεύτερος όρος μέσα στο ολοκλήρωμα:

$$\left[ P_1(C_{01} - C_{11})f_{Y|H_1}(y|H_1) \text{ and } P_0(C_{10} - C_{00})f_{Y|H_0}(y|H_0) \right]$$

Όσο μεγαλώνει ο δεύτερος όρος, τόσο μικρότερο κόστος θα έχουμε σε περίπτωση λάθος εκτίμησης. Εάν ισχύει:

$$P_1(C_{01} - C_{11})f_{Y|H_1}(y|H_1) < P_0(C_{10} - C_{00})f_{Y|H_0}(y|H_0)$$

Τότε η εκτίμηση μας θα είναι η υπόθεση  $H_0$ , και αν αλλάξει η φορά της ανισότητας, τότε η εκτίμηση μας θα είναι η  $H_1$ . Για να γίνει πιο 'ευανάγνωστο', διαιρούμε κατά μέλη και έχουμε:

$$\frac{f_{Y|H_1}(y|H_1)}{f_{Y|H_0}(y|H_0)} > \frac{P_0(C_{10} - C_{00})}{P_1(C_{01} - C_{11})}$$

Αυτή η σχέση είναι γνωστή σαν κανόνας του Bayes's, το κλάσμα δηλώνει *Likelihood Ratio* (αναλογία πιθανοφάνειας)  $\Lambda(y)$ . Άρα

$$\Lambda(y) = \frac{f_{Y|H_1}(y|H_1)}{f_{Y|H_0}(y|H_0)}$$

Πρέπει να σημειωθεί ότι εάν πάρουμε  $K$  δείγματα από μια παρατηρούμενη κυματομορφή, (έστω τα δείγματα αυτά να είναι  $Y_1, Y_2, \dots, Y_K$ ), τότε ο κανόνας που αποφασίζουμε για την υπόθεση  $H_0$  ή  $H_1$  γίνεται:

$$\Lambda(\mathbf{y}) = \frac{f_{Y|H_1}(\mathbf{y}|H_1)}{f_{Y|H_0}(\mathbf{y}|H_0)}$$

Όπου όμως αυτή την φορά η παρατήρηση  $y$  είναι ένα διάνυσμα (πίνακας) ο οποίος αποτελείται από  $K$  στοιχεία.

$$\mathbf{Y}^T = [Y_1 \quad Y_2 \quad \dots \quad Y_K]$$

Το κατώφλι είναι όπως και προηγουμένως ίσο με:

$$\eta = \frac{P_0(C_{10} - C_{00})}{P_1(C_{01} - C_{11})}$$

Έτσι με την εφαρμογή του κριτηρίου Bayes, το κριτήριο το οποίο ελαχιστοποιεί το μέσο κόστος είναι το *likelihood test*:

$$\begin{array}{c} H_1 \\ \Lambda(\mathbf{y}) > \eta \\ < \eta \\ H_0 \end{array}$$

Μια σημαντική παρατήρηση είναι ότι το τεστ αυτό διεξάγεται απλώς επεξεργάζοντας τον πίνακα παρατηρήσεων και συγκρίνοντας τον με το κατώφλι. Με αυτόν τον τρόπο, σε περίπτωση που οι *prior* πιθανότητες, ή και τα κόστη αλλάζουν, το μόνο που έχουμε να κάνουμε για να αναπροσαρμόσουμε τον δέκτη μας, είναι να καταχωρήσουμε τις καινούργιες τιμές στο κατώφλι.

Λόγω του ότι ο φυσικός λογάριθμος είναι μια μονοτονικά αυξανόμενη συνάρτηση όπως βλέπουμε στο σχήμα 4, και το γεγονός ότι το κατώφλι είναι μεγαλύτερο ή ίσο με το μηδέν και δεν μπορεί να πάρει αρνητική τιμή, μπορούμε να λογαριθμίσουμε και το *likelihood ratio test*, αλλά και το κατώφλι, και δεν θα έχουμε καμία αλλαγή στο αποτέλεσμα.

$$\begin{array}{c} H_1 \\ \ln \Lambda(\mathbf{y}) > \ln \eta \\ < \ln \eta \\ H_0 \end{array}$$

Προσέχουμε ότι εάν διαλέξουμε το κόστος της λάθος εκτίμησης να είναι ίσο με ένα, και το κόστος της σωστής εκτίμησης να είναι ίσο με μηδέν,

$$C_{01} = C_{10} = 1$$

Και

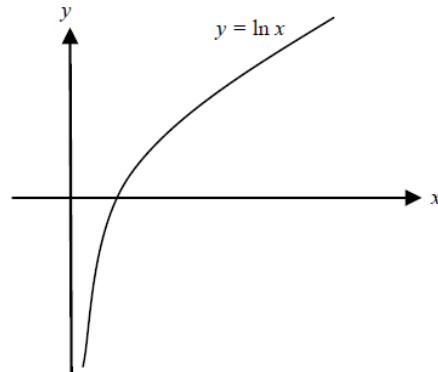
$$C_{00} = C_{11} = 0$$

Το ρίσκο έτσι γίνεται

$$\mathfrak{R} = P_M P_1 + P_F P_0 = P(\varepsilon)$$

Σε αυτήν την περίπτωση το να ελαχιστοποιήσουμε το μέσο κόστος είναι ίσο με το να ελαχιστοποιήσουμε την πιθανότητα λάθους. Οι δέκτες για τέτοιες τιμές του κόστους, λέγονται *δέκτες ελάχιστης πιθανότητας λάθους* (*minimum probability of error*

receivers). Το κατώφλι γίνεται τότε  $n = P_0/P_1$ . Σε περίπτωση που οι *prior* πιθανότητες είναι ίσες, τότε λόγω του ότι το κατώφλι θα ισούται με ένα, αν λογαριθμίσουμε το κατώφλι θα είναι ίσο με το μηδέν.



Σχήμα 4 γραφική απεικόνιση συνάρτησης λογαρίθμου.

### Παράδειγμα 1

Σε ένα τηλεπικοινωνιακό σύστημα, αναλογιστείτε μια πηγή η οποία εκπέμπει δύο σήματα, το '0' και το '1'. Το '0' αποστέλλεται με μηδενική ισχύ, και το '1' με στάθμη τάσης ίση με ' $m$ '. Το λαμβανόμενο σήμα αλλοιώνεται από λευκό προσθετικό θόρυβο  $N$  ο οποίος έχει μηδενικό μέσο όρο και διακύμανση  $\sigma^2$ .

- Βρείτε το *likelihood test* και ορίστε τις περιοχές απόφασης.
- Υπολογίστε την πιθανότητα του λάθους συναγερμού (false alarm) και την πιθανότητα ανίχνευσης (probability of detection).

Λύση

Το λαμβανόμενο σήμα για την έκαστο υπόθεση έχει τιμή:

$$\begin{aligned} H_1 : Y &= m + N \\ H_0 : Y &= N \end{aligned}$$

Υπό την υπόθεση λοιπόν  $H_0$ , η πιθανοφάνεια είναι:

$$f_{Y|H_0}(y|H_0) = f_N(n) = \frac{1}{\sqrt{2\pi\sigma}} \exp\left(-\frac{y^2}{2\sigma^2}\right)$$

Υπό την υπόθεση  $H_1$  ο μέσος όρος της παρατήρησης  $Y$  είναι  $E[Y] = E[m + N] = m$ , αφού μέσος όρος του θορύβου είναι μηδέν. Η διακύμανση της παρατήρησης είναι:

$$\text{var}[Y] = \text{var}[m + N] = E[(m + N)^2] - (E[m + n])^2 = E[N^2] = \sigma^2$$

Άρα:

$$f_{Y|H_1}(y|H_1) = \frac{1}{\sqrt{2\pi\sigma}} \exp\left[-\frac{1}{2} \frac{(y-m)^2}{\sigma^2}\right]$$

Αυτή η διαπίστωση είναι πολύ σημαντική, λόγω του ότι είτε υπό την υπόθεση  $H_0$ , είτε υπό την υπόθεση  $H_1$ , η διακύμανση των δειγμάτων παραμένει η ίδια, και έτσι μπορούμε να απλοποιήσουμε στο αναλογικό τεστ (*likelihood ratio*):

$$\Lambda(y) = \frac{f_{Y|H_1}(y|H_1)}{f_{Y|H_0}(y|H_0)} = \exp\left(-\frac{m^2 - 2ym}{2\sigma^2}\right)$$

Το επόμενο βήμα είναι να πάρουμε τον φυσικό λογάριθμο και για τα δύο μέλη της εξίσωσης για να μπορέσουμε να ‘ξεφορτωθούμε’ το εκθετικό.

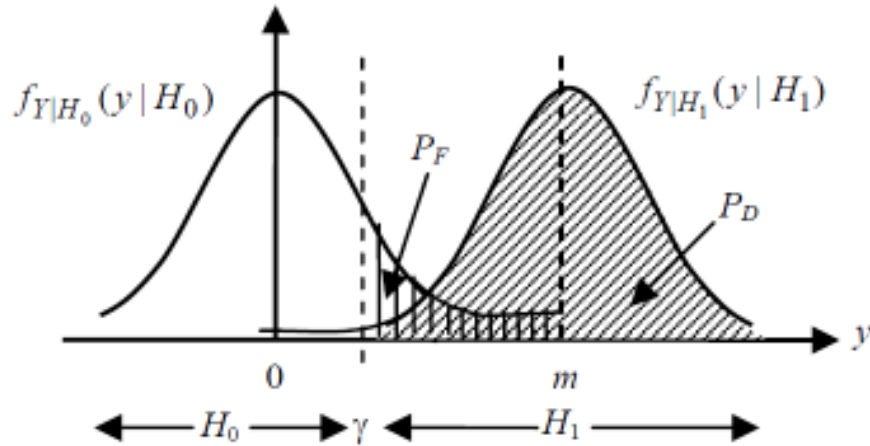
$$\ln \Lambda(y) = \frac{m}{\sigma^2} y - \frac{m^2}{2\sigma^2} \begin{matrix} H_1 \\ > \\ < \\ H_0 \end{matrix} \ln \eta$$

Σειρά έχει να λύσουμε ως προς την παρατηρούμενη τιμή  $y$ :

$$y \begin{matrix} H_1 \\ > \\ < \\ H_0 \end{matrix} \frac{\sigma^2}{m} \ln \eta + \frac{m}{2} = \gamma$$

Αυτό που κάνουμε για να αποφανθούμε εάν το παρατηρούμενο δείγμα είναι ίσο με την υπόθεση  $H_0$ , είτε η υπόθεση  $H_1$  είναι να συγκρίνουμε το  $y$  με το κατώφλι  $\gamma$ . Εάν η παρατήρηση είναι πιο μεγάλη από την τιμή που έχει το  $\gamma$ , τότε έχουμε την υπόθεση  $H_1$ . Αλλιώς έχουμε την υπόθεση  $H_0$ . Οι περιοχές απόφασης φαίνονται στο σχήμα 5.





Σχήμα 5 Περιοχές απόφασης.

(b) Οι πιθανότητες λάθος συναγερωμού και ανίχνευσης είναι:

- Πιθανότητα λάθους:

$$P_F = P(\text{decide } H_1 | H_0 \text{ true}) = \int_{\gamma}^{\infty} \frac{1}{\sqrt{2\pi}\sigma} e^{-\frac{y^2}{2\sigma^2}} dy = Q\left(\frac{\gamma}{\sigma}\right) = \text{erfc}_s\left(\frac{\gamma}{\sigma}\right)$$

Όπου

$$Q(a) = \int_a^{\infty} \frac{1}{\sqrt{2\pi}} e^{-\frac{u^2}{2}} du$$

- Πιθανότητα σωστής ανίχνευσης:

$$P_D = P(\text{decide } H_1 | H_1 \text{ true}) = \int_{\gamma}^{\infty} \frac{1}{\sqrt{2\pi}\sigma} e^{-\frac{(y-m)^2}{2\sigma^2}} dy = Q\left(\frac{\gamma-m}{\sigma}\right)$$

## Παράδειγμα 2

Υποθέστε ότι ο δέκτης του προηγούμενου παραδείγματος παίρνει  $K$  δείγματα  $Y_1, Y_2, \dots, Y_K$  από το λαμβανόμενο σήμα. Τα δείγματα θορύβου είναι ανεξάρτητες Gaussian μεταβλητές, κάθε μια με μηδενικό μέσο όρο και διακύμανση  $\sigma^2$ . Να βρεθεί ο βέλτιστος κανόνας απόφασης.

## Ανίχνευση ασθενών σημάτων σε πρώτης τάξης εξαρτημένο θόρυβο

Λύση:

Το λαμβανόμενο σήμα υπό την υπόθεση  $H_0$  και  $H_1$  είναι:

$$\begin{aligned} H_1: Y_k &= m + N, \quad k=1,2,\dots,K \\ H_0: Y_k &= N, \quad k=1,2,\dots,K \end{aligned}$$

Υπό την υπόθεση  $H_0$ , η πιθανοφάνεια της  $k$ -οστής παρατήρησης είναι:

$$f_{Y_k|H_0}(y_k | H_0) = f_{N_k}(y_k) = \frac{1}{\sqrt{2\pi\sigma}} \exp\left(-\frac{y_k^2}{2\sigma^2}\right)$$

Υπό την υπόθεση  $H_1$ , η πιθανοφάνεια της  $k$ -οστής παρατήρησης είναι μια τυχαία Gaussian μεταβλητή με μέσο όρο  $m$  και διακύμανση  $\sigma^2$ . Έτσι η πιθανοφάνεια για αυτήν την υπόθεση είναι:

$$f_{Y_k|H_1}(y_k | H_1) = f_{N_k}(y_k - m) = \frac{1}{\sqrt{2\pi\sigma}} \exp\left[-\frac{(y_k - m)^2}{2\sigma^2}\right]$$

Το επόμενο βήμα είναι να βγάλουμε το τεστ πιθανοφάνειας το οποίο είναι ίσο με:

$$\Lambda(\mathbf{y}) = \frac{f_{Y|H_1}(\mathbf{y} | H_1)}{f_{Y|H_0}(\mathbf{y} | H_0)}$$

Όμως έχουμε πραγματοποιήσει δειγματοληψία ( $K$  δείγματα) και εφόσον τα δείγματα αυτά είναι στατιστικά ανεξάρτητα, η joint density function των  $K$  δειγμάτων ισούται με το γινόμενο της έκαστο πυκνότητας πιθανότητας για το κάθε δείγμα. Άρα θα έχουμε:

$$f_{Y|H_0}(\mathbf{y} | H_0) = \prod_{k=1}^K \frac{1}{\sqrt{2\pi\sigma}} e^{-\frac{y_k^2}{2\sigma^2}} \quad f_{Y|H_1}(\mathbf{y} | H_1) = \prod_{k=1}^K \frac{1}{\sqrt{2\pi\sigma}} e^{-\frac{y_k^2}{2\sigma^2}}$$

Κάνοντας χρήση της ιδιότητας:

$$\prod_k e^{x_k} = e^{\sum_k x_k}$$

Το τεστ πιθανοφάνειας γίνεται:

$$\Lambda(\mathbf{y}) = \exp\left[\sum_{k=1}^K \frac{y_k^2}{2\sigma^2} - \sum_{k=1}^K \frac{(y_k - m)^2}{2\sigma^2}\right] = \exp\left[\frac{m}{\sigma^2} \sum_{k=1}^K y_k - \frac{Km^2}{2\sigma^2}\right]$$

Αν πάρουμε τον φυσικό λογάριθμο και για τα δύο μέλη της εξίσωσης, θα έχουμε:

$$\ln \Lambda(\mathbf{y}) = \frac{m}{\sigma^2} \sum_{k=1}^K y_k - \frac{Km^2}{2\sigma^2} \begin{matrix} > \\ < \end{matrix} \begin{matrix} H_1 \\ H_0 \end{matrix} \ln \eta$$

Αν λύσουμε ως προς  $y_k$ , το ισοδύναμο τεστ γίνεται:

$$\sum_{k=1}^K y_k \begin{matrix} > \\ < \end{matrix} \begin{matrix} H_1 \\ H_0 \end{matrix} \frac{\sigma^2}{m} \ln \eta + \frac{Km}{2}$$

Αυτό που έχει να κάνει τώρα ο δέκτης είναι να πάρει τα  $K$  δείγματα, να τα προσθέσει και να τα συγκρίνει με το κατώφλι:

$$\frac{\sigma^2}{m} \ln n + \frac{Km}{2}$$

### Εκτίμηση Παραμέτρων του εκπεμπόμενου σήματος

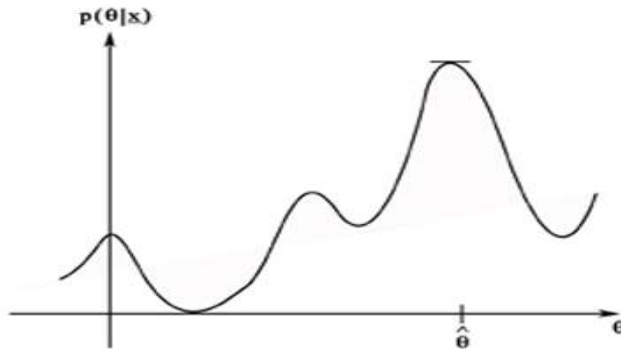
Σε προηγούμενη ενότητα είδαμε την υλοποίηση ενός δέκτη ο οποίος δεχόταν ένα σήμα εμπλουτισμένο με θόρυβο και έπρεπε να αποφασίσει ανάμεσα σε 2 πιθανές υποθέσεις, το σήμα που στάλθηκε. Συγκεκριμένα, έπρεπε να αποφασίσει ο δέκτης ανάμεσα στην υπόθεση  $H_0$  και στην  $H_1$ .

Σε αυτήν την ενότητα, υποθέτουμε ότι ο δέκτης έχει ήδη λάβει την σωστή απόφαση, όμως κάποιες παράμετροι που σχετίζονται με το σήμα παραμένουν άγνωστοι. Αυτές οι παράμετροι μπορεί να είναι η διακύμανση, ή και η θέση (*location estimation*). Ο στόχος μας είναι να πραγματοποιήσουμε δειγματοληψία στο σήμα που λάβαμε και αποφασίσαμε ότι συμβολίζει π.χ. το '1' ή το '0'.

Έστω  $Y_1, Y_2, \dots, Y_k$  να είναι τα δείγματα που πήρα από το λαμβανόμενο σήμα  $Y$ , τα οποία είναι ανεξάρτητα και ιδιοκατανομημένα (i.i.d.) και σχηματίζουν κατανομή πιθανότητας η οποία με την σειρά της εξαρτάται άμεσα από την παράμετρο που ψάχνουμε (έστω την παράμετρο ' $\theta$ ' η οποία δηλώνει την θέση ή τον μέσο όρο της κατανομής). Για να εκτιμήσουμε την παράμετρο αυτή, χρησιμοποιούμε την τιμή που έχει το έкаστο δείγμα ( $y_1, y_2, \dots, y_k$  είναι οι τιμές που έχουν τα δείγματα  $Y_1, Y_2, \dots, Y_k$ ), Για να μην γίνει σύγχυση ανάμεσα στα  $Y_1, Y_2, \dots, Y_k$  και στα  $y_1, y_2, \dots, y_k$ , με τα κεφαλαία συμβολίζουμε την δειγματοληψία νούμερο  $1, 2, \dots, k$  και τα μικρά συμβολίζουμε τι τιμή έχει η έкаστο δειγματοληψία.

Τα μικρά λοιπόν  $y_1, y_2, \dots, y_k$  στην συνέχεια χρησιμοποιούνται από την συνάρτηση  $g(Y_1, Y_2, \dots, Y_k)$  για να προκύψει στο τέλος η εκτίμηση της παραμέτρου που επιζητούμε.

$$\hat{\theta} = g(Y_1, Y_2, \dots, Y_k)$$



Αυτό που επιζητούμε είναι μια προσέγγιση της παραμέτρου 'θ' έτσι ώστε να μεγιστοποιείται η πιθανοφάνεια  $p(\theta|x)$

Προσέξτε ότι ο συμβολισμός μιας παραμέτρου η οποία είναι η εκτίμηση μιας άλλης παραμέτρου, συμβολίζεται με το καπελάκι  $\hat{\theta}$ .

Η παράμετρος που πρόκειται να εκτιμηθεί μπορεί να είναι τυχαία, ή μη-τυχαία. Η εκτίμηση τυχαίων παραμέτρων καλείται *Bayes Estimation* (εκτίμηση Bayes) όπως είδαμε στην προηγούμενη ενότητα, ενώ η εκτίμηση μη-τυχαίων παραμέτρων καλείται *maximum likelihood estimation* (MLE) ή αλλιώς *εκτίμηση μέγιστης πιθανοφάνειας*.

### *Maximum Likelihood Estimation – Εκτίμηση Μέγιστης Πιθανοφάνειας*

Όπως είπαμε προηγουμένως, η εκτίμηση μη-τυχαίων παραμέτρων δίνεται από την *maximum likelihood estimation*. Αν  $Y_1, Y_2, \dots, Y_k$  είναι τα δείγματα ( $K$  στον αριθμό) μιας τυχαίας μεταβλητής (τυχαίας διότι αυτή μπορεί να είναι για παράδειγμα είτε το σήμα '0' ή το σήμα '1' ενός δυαδικού συστήματος), τα οποία δείγματα έχουν τιμή  $y_1, y_2, \dots, y_k$  έκαστο. Όπως πολύ καλά γνωρίζουμε, εάν κάνουμε δειγματοληψία σε ένα φαινόμενο που λαμβάνει χώρα, τότε αν υπερθέσω τα δείγματα αυτά, θα προκύψει μια κατανομή η οποία περιγράφει το φαινόμενο αυτό. Στην δική μας περίπτωση περιγράφει την κατανομή που έχει το λαμβανόμενο σήμα μέσω της συνάρτησης

$f_{Y|\theta}(y|\theta)$  η οποία δίνει μια κατανομή πιθανότητας που έχει μέσο όρο την τιμή  $\theta$ . Η likelihood συνάρτηση  $L(\theta)$  λόγω των  $K$  δειγμάτων, γίνεται:

$$L(\theta) = f_{Y_1, \dots, Y_K|\theta}(y_1, y_2, \dots, y_K | \theta) = f_{Y|\theta}(y | \theta) = \prod_{k=1}^K f_{Y_k|\theta}(y_k | \theta)$$

Είπαμε ότι η κατανομή αυτή έχει μέσο όρο την τιμή  $\theta$ . Επίσης στο σημείο αυτό η συνάρτηση πυκνότητας πιθανότητας θα εμφανίζει και την μέγιστη τιμή της. Αρκεί λοιπόν να βρούμε το ακρότατο συνάρτησης αυτής και θα έχουμε πετύχει μια πολύ καλή προσέγγιση της τιμής  $\theta$ . Πως γίνεται τώρα να βρούμε το ακρότατο αυτής της συνάρτησης; Λόγω του ότι δεν είναι δυνατό να βρεθεί μέσω της απευθείας παραγωγίσιμης της συνάρτησης πιθανοφάνειας, πρέπει να βρούμε μια άλλη συνάρτηση η οποία να έχει την ίδια μονοτονία με την συνάρτηση της πιθανοφάνειας. Αυτή δεν είναι άλλη από τον φυσικό λογάριθμο. Λογαριθμίζουμε λοιπόν την συνάρτηση πιθανοφάνειας, Παραγωγίζουμε και έτσι δεν έχουμε καμία αλλαγή στην τιμή της παραμέτρου που θα προκύψει. Η συνάρτηση που θα πρέπει να λυθεί είναι λοιπόν:

$$\frac{\partial}{\partial \theta} \ln f_{Y|\theta}(y | \theta) = 0$$

### Παράδειγμα

Σε παράδειγμα της προηγούμενης ενότητας είδαμε την περίπτωση που το λαμβανόμενο σήμα μπορεί να πάρει δύο διαφορετικές τιμές: '0' και '1'. Το σύστημα υπόθεσης ήταν δηλαδή:

$$\begin{aligned} H_1 : Y_k &= m + N_k, \quad k = 1, 2, \dots, K \\ H_0 : Y_k &= N_k, \quad k = 1, 2, \dots, K \end{aligned}$$

Υποθέτοντας ότι η παράμετρος  $m$  (παράμετρος θέσης) είναι άγνωστη, να βρεθεί η εκτίμηση  $\hat{m}_{ml}$  μέγιστης πιθανοφάνειας (ML) του μέσου όρου της κατανομής που περιγράφει το σήμα '1'.

### Λύση

Λόγω του ότι τα δείγματα που θα χρησιμοποιήσω για να βρω την παράμετρο είναι i.i.d., η συνάρτηση πιθανοφάνειας είναι:

$$f_{Y|M}(y|m) = \prod_{k=1}^K \frac{1}{\sqrt{2\pi}\sigma} \exp\left[-\frac{(y_k - m)^2}{2\sigma^2}\right] = \frac{1}{(2\pi)^{K/2} \sigma^K} \exp\left[-\sum_{k=1}^K \frac{(y_k - m)^2}{2\sigma^2}\right]$$

Βλέπετε ότι λόγω των  $K$  δειγμάτων, η συνάρτηση που περιγράφει την κατανομή είναι το γινόμενο των δειγμάτων στην κατανομή αυτή. Εμείς θέλουμε να ‘απομονώσουμε’ την παράμετρο ‘ $m$ ’ η οποία βρίσκεται στο εκθετικό. Εδώ βλέπετε πόσο χρήσιμη είναι η λογαρίθμηση, διότι έτσι θα ‘ξεφορτωθώ’ το εκθετικό. Έχουμε με την λογαρίθμηση:

$$\ln f_{Y|M}(\mathbf{y} | m) = \ln \left[ \frac{1}{(2\pi)^{K/2} \sigma^K} \right] - \sum_{k=1}^K \frac{(y_k - m)^2}{2\sigma^2}$$

Το επόμενο βήμα είναι να ‘σπάσω’ το άθροισμα σε δύο αθροίσματα και μετά να παραγωγίσω μερικώς ως προς  $m$ :

$$\frac{\partial \ln f_{Y|M}(\mathbf{y} | m)}{\partial m} = \sum_{k=1}^K \frac{y_k - m}{\sigma^2} = \sum_{k=1}^K \frac{y_k}{\sigma^2} - \frac{Km}{\sigma^2} = \frac{K}{\sigma^2} \left( \frac{1}{K} \sum_{k=1}^K y_k - m \right) = 0$$

Τώρα το μόνο που μένει είναι να λύσω ως προς  $m$ , και έχω:

$$m = (1/k) \sum_{k=1}^K y_k$$

Όμως αυτό δεν είναι σωστό διότι η τελευταία σχέση δεν μας δίνει την πραγματική τιμή της παραμέτρου, αλλά μια εκτίμηση αυτής. Πιο σωστά λοιπόν θα έπρεπε να γράψουμε:

$$\hat{m}_{ml} = (1/k) \sum_{k=1}^K y_k$$

Αυτή η σχέση όμως είναι ο γνωστός σε όλους μας, μέσος όρος.

**Σε περίπτωση που έχω ένα δυαδικό σύστημα εκπομπής, ποιο θα είναι εκείνο το κριτήριο που θα με οδηγήσει στην εκτίμηση;**

Ας δούμε πρώτα την περίπτωση που θα βασιστώ σε ένα δείγμα. Εάν ισχύει η πιθανοφάνεια  $\varphi_{\underline{y}}(\underline{y} | H_1, \mathcal{G}) \geq \varphi_{\underline{y}}(\underline{y} | H_0, \mathcal{G})$ , τότε αυτό σημαίνει ότι η πιθανότητα το λαμβανόμενο δείγμα  $\underline{y}$  να είναι η υπόθεση  $H_1$ , είναι πιο μεγάλη από την υπόθεση  $H_0$ . Εάν η κατανομή που εμφανίζει ο θόρυβος που προστέθηκε στο σήμα είναι Gaussian, τότε η υπόθεση  $H_1$  θα έχει τιμή  $H_1 = \mathcal{G}\varepsilon_1 + N$  και η υπόθεση  $H_0$  θα έχει τιμή  $H_0 = \mathcal{G}\varepsilon_0 + N$ .

- Όπου  $N$  είναι ο προστιθέμενος θόρυβος με μηδενική μέση τιμή και διακύμανση  $\sigma^2$

## Ανίχνευση ασθενών σημάτων σε πρώτης τάξης εξαρτημένο θόρυβο

- Όπου  $\varepsilon_1$  είναι μια μη-αρνητική σταθερά η οποία πολλαπλασιάζεται με την μέγιστη στάθμη ισχύος,  $\vartheta$ , που αποστέλλεται το σήμα και καθορίζει πόσο τις εκατό θα έχει υποβιβαστεί η στάθμη του εκπεμπόμενου σήματος, σε σχέση με την μέγιστη στάθμη που μπορεί να πάρει. Για παράδειγμα εάν  $\vartheta = 10$  Watt και  $\varepsilon_1 = 0.8$ , τότε το σήμα '1' θα αποστέλλεται με ισχύ  $10 \cdot 0.8$  Watt = 8 Watt .

Με βάση αυτές τις διευκρινίσεις, το δείγμα  $y$  υπό την υπόθεση  $H_1$  θα πάρει τιμή:

$$\varphi_{\underline{y}}(\underline{y} | H_1, \vartheta) = \frac{1}{2\sqrt{\gamma\pi}} \exp\left(-\frac{(y - \vartheta\varepsilon_1)^2}{4\gamma}\right)$$

Και το δείγμα  $y$  υπό την υπόθεση  $H_0$ , θα πάρει τιμή:

$$\varphi_{\underline{y}}(\underline{y} | H_0, \vartheta) = \frac{1}{2\sqrt{\gamma\pi}} \exp\left(-\frac{(y - \vartheta\varepsilon_0)^2}{4\gamma}\right)$$

Από αυτές τις δύο σχέσεις, μπορούμε να βγάλουμε το εξής σπουδαίο συμπέρασμα: εφόσον και οι δύο σχέσεις εκφράζουν μια πιθανοφάνεια, όταν η μια πιθανοφάνεια έχει πιο μεγάλη τιμή από την άλλη, τότε αυτή θα είναι και η εκτίμηση μας:

Αν ισχύει λοιπόν  $\varphi_{\underline{y}}(\underline{y} | H_1, \vartheta) \geq \varphi_{\underline{y}}(\underline{y} | H_0, \vartheta)$ , τότε ισχύει ότι το παρατηρούμενο διάνυσμα  $\underline{y}$  είναι πιο πιθανό να είναι η υπόθεση  $H_1$  παρά η υπόθεση  $H_0$ . Άρα αν διαιρέσουμε με  $\varphi_{\underline{y}}(\underline{y} | H_0, \vartheta)$  και πάρουμε τον φυσικό λογάριθμο του κλάσματος αυτού, θα έχουμε το κριτήριο απόφασης να γίνεται:

$$D_k^{ML} = \underline{y} : \sum_{i=1}^n \ln \frac{\varphi_{\underline{y}}(\underline{y} | H_1, \vartheta)}{\varphi_{\underline{y}}(\underline{y} | H_0, \vartheta)} \geq 0$$

# ΕΝΟΤΗΤΑ II

## *Καλύπτονται έννοιες:*

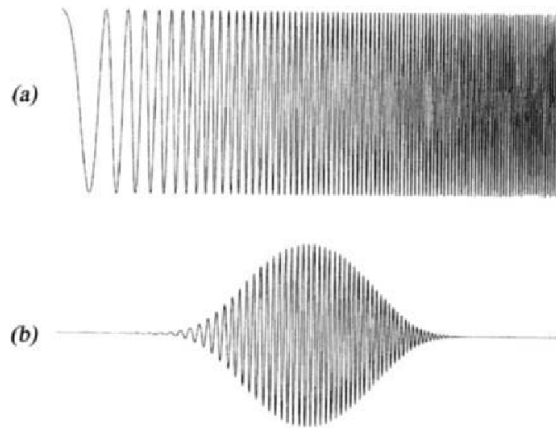
- Εισαγωγή στην μη-γραμμική επεξεργασία σήματος
- Λόγος χρήσης μη-γραμμικών φίλτρων έναντι γραμμικών
- Πρακτικά παραδείγματα αλλοίωσης του σήματός μας από non-Gaussian θόρυβο (ADSL γραμμές)
- Απόρροια της Gaussian κατανομής μέσω του Κεντρικού οριακού θεωρήματος και λόγος ύπαρξης μη-Gaussian κατανομών μέσω της αντιπαράθεσης Κεντρικού οριακού θεωρήματος και γενικευμένου Κεντρικού οριακού θεωρήματος
- Διάφορες προσεγγίσεις στην αντιμετώπιση των *outliers-παρασίτων* (σταθμισμένοι μέσοι όροι, weighted myriads)
- Το φιλτράρισμα και αναπαράσταση του με όρους στατιστικής, και παρουσίαση εννοιών όπως εκτίμηση θέσης (*Location estimation*)
- Μη-Gaussian μοντέλα και πως χρησιμοποιούμε την γενικευμένη Gaussian κατανομή για αντιμετώπιση διαδικασιών με μεγάλη κρουστικότητα
- Stable κατανομές
- Στατιστικές ιδιότητες εκτιμητών (μη-προκατηλημενοι εκτιμητές, αποδοτικοί εκτιμητές, παρουσίαση του ορίου Cramer and Rao Lower Bound)



## Εισαγωγή στην μη-γραμμική επεξεργασία σήματος

Η επεξεργασία σήματος είναι μια διαδικασία η οποία περιλαμβάνει ένα μεγάλο φάσμα μεθόδων για την αναπαράσταση, ανάλυση, διάδοση και ανάκτηση σημάτων από διάφορες πηγές. Μια τέτοια διαδικασία, αναφέρεται στην “μαθηματική χειραγώγηση” των σημάτων. Ίσως η πιο σημαντική μαθηματική διαδικασία στην επεξεργασία σήματος, είναι αυτή της διαδικασίας ανάλυσης και κατασκευής ενός φίλτρου, η οποία σαν στόχο έχει να απομακρύνει τον επιπρόσθετο θόρυβο, θόρυβο λόγω διασυμβολικής παρεμβολής, ή οποιονδήποτε θόρυβο ο οποίος δυσχεραίνει την ανάκτηση και αποκωδικοποίηση του εισερχόμενου σήματος. Εάν για παράδειγμα το σήμα μου έχει αλλοιωθεί από το κανάλι διαβίβασης, τότε ο στόχος μου είναι να κατασκευάσω ένα φίλτρο το οποίο καταστέλλει αυτές τις διακυμάνσεις.

Ένα κλασσικό παράδειγμα φιλτραρίσματος φαίνεται στο σχήμα 1.1, όπου ένα ζωνοπερατό φίλτρο αναδιαμορφώνει ένα chirped σήμα (chirped σήμα είναι ένα σήμα του οποίου η συχνότητα αυξάνει 'up-chirp' η μειώνεται 'down-chirp' ανεπιθύμητα). Άρα εάν κάνουμε σωστή επιλογή των στοιχείων που απαρτίζουν το φίλτρο μου, μπορώ να αναιρέσω όσο αυτό είναι δυνατόν, αυτό το φαινόμενο. Το σχήμα 1.1b δείχνει το φιλτραρισμένο πλέον σήμα, όταν ένα 120 tap FIR φίλτρο έχει χρησιμοποιηθεί.

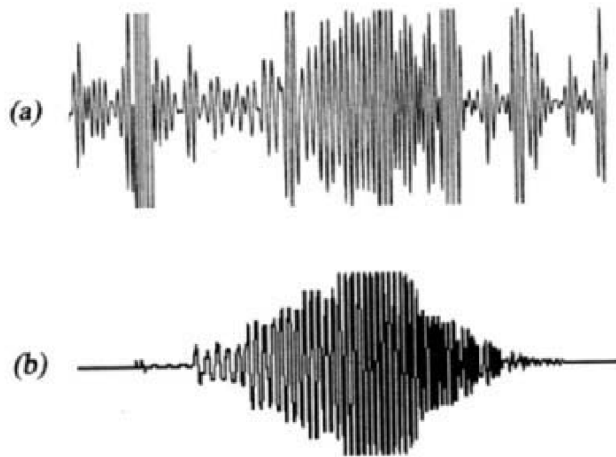


**Σχήμα 1.1** φιλτράρισμα επιλεκτικής συχνότητας με χρήση ζωνοπερατού φίλτρου:  
(a) chirped σήμα (σήμα αυξανόμενης συχνότητας), (b) έξοδος FIR φίλτρου.

Παρόλο που τα γραμμικά φίλτρα θα συνεχίζουν να παίζουν σημαντικό ρόλο στην επεξεργασία σήματος, τα μη-γραμμικά φίλτρα δείχνουν να έχουν μεγαλύτερη χρήση. Ο λόγος που συμβαίνει αυτό είναι οι αλγόριθμοι που έχουν αναπτυχθεί για την μη-γραμμική επεξεργασία σήματος, είναι αρκετά πολύπλοκοι και απαιτούν αρκετά μεγάλη επεξεργαστική δύναμη, μια δύναμη που πλέον υπάρχει.

*Γιατί μη-γραμμική επεξεργασία σήματος;*

Η μη-γραμμική επεξεργασία σήματος προσφέρει πλεονεκτήματα σε εφαρμογές όπου, οι τυχαίες διεργασίες (όπως ο προσθετικός θόρυβος) περιγράφονται από non-Gaussian στατιστική. Η εμπειρία δείχνει ότι non-Gaussian διαδικασίες, υπάρχουν σε μια ευρεία ζώνη εφαρμογών, οι οποίες περιλαμβάνουν ασύρματες επικοινωνίες, υδρολογία, γεωλογία, οικονομικά και επεξεργασία εικόνας. Το κοινό στοιχείο σε αυτές τις εφαρμογές όπως και σε πολλές άλλες, είναι ότι δείχνουν να παράγουν πιο μεγάλο φάσμα παρατήρησης από το Gaussian μοντέλο, λόγω της ιδιότητας των non-Gaussian τυχαίων διεργασιών να έχουν πιο “βαριές ουρές” της κατανομής που ακολουθούν (δηλαδή οι ουρές της κατανομής σβήνουν πιο αργά από το Gaussian μοντέλο κατανομής). Σαν αποτέλεσμα οι γραμμικές μέθοδοι οι οποίες υπακούουν την αρχή της υπέρθεσης, υποφέρουν από δείγματα που καταφθάνουν τα οποία έχουν αλλοιωθεί από θόρυβο υψηλού πλάτους. Μη γραμμικές μέθοδοι από την άλλη μεριά, εκμεταλλεύονται τα χαρακτηριστικά της κατανομής πιθανότητας του θορύβου, και με βάση τα χαρακτηριστικά αυτά, χρησιμοποιούν διαφορετικά κριτήρια “απόφασης” στην λήψη.

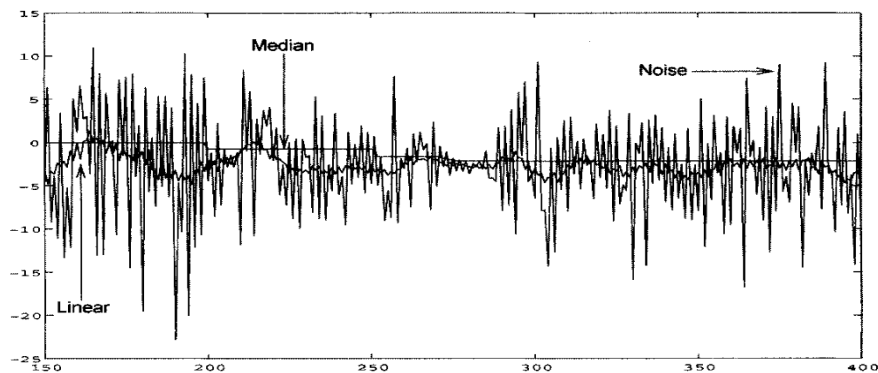


**Σχήμα 1.2** επιλεκτικό φιλτράρισμα σήματος σε non-Gaussian θόρυβο: (a) έξοδος γραμμικού FIR φίλτρου, (b) μη-γραμμικού φίλτρου

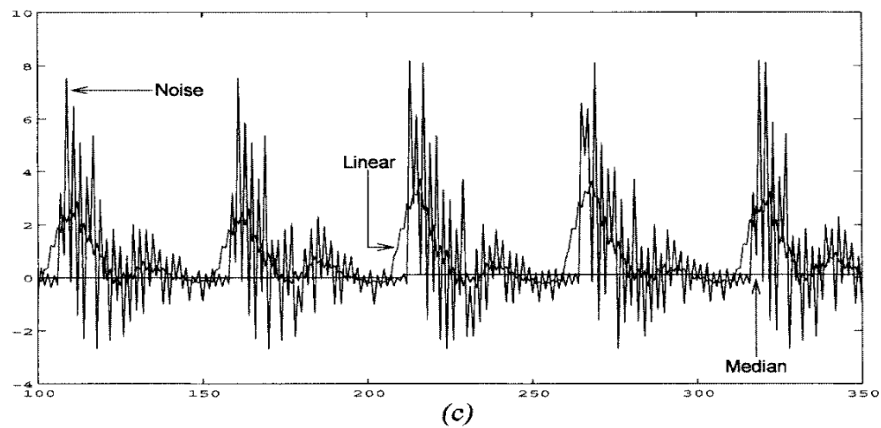
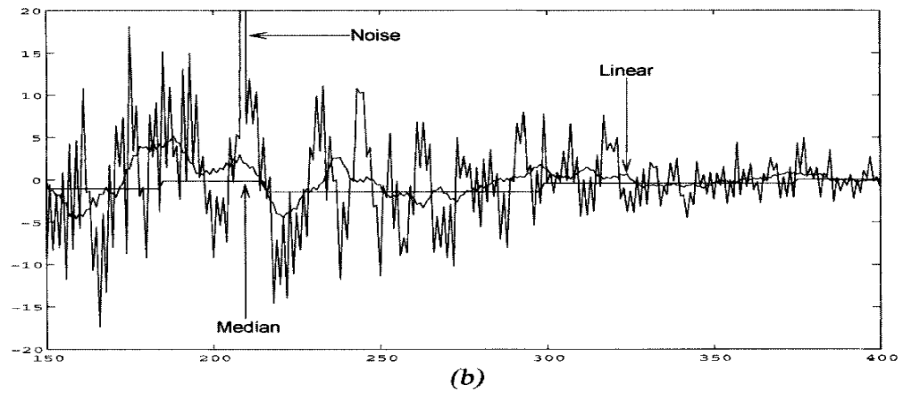
Για να γίνει πιο κατανοητό το παραπάνω, φέρτε στο μυαλό της πάλι το παράδειγμα του ζωνοπερατού φίλτρου. Αυτή της την φορά, το chirped σήμα έχει αλλοιωθεί από non-Gaussian θόρυβο, κατά την διάρκεια του σταδίου ανάκτησης του σήματος. Λόγω του μη-Γκαουσιανού θορύβου, το γραμμικό FIR φίλτρο, δεν δίνει στην έξοδο το καθαρό σήμα που είδαμε στο σχήμα 1.1, αλλά ένα βαριά αλλοιωμένο (σχήμα 1.2a). Εάν χρησιμοποιήσουμε ένα μη-γραμμικό φίλτρο, τότε η κατάσταση βελτιώνεται σε μεγάλο βαθμό όπως βλέπουμε (σχήμα 1.2b). Το μη γραμμικό φίλτρο

κατάφερε να εξάγει όλο το φάσμα που περικλείει το ζωνοπερατό φίλτρο, και οι ανεπιθύμητοι όροι, έχουν καταπιεστεί.

Ένα πρακτικό παράδειγμα αλλοίωσης του σήματός μας από non-Gaussian θόρυβο, φαίνεται της τηλεφωνικές γραμμές (της το ADSL όπου έχουμε μετάδοση πληροφοριών με πολύ μεγάλο bit rate), όπου ο θόρυβος στο κανάλι διαβίβασης εμφανίζει κρουστικά χαρακτηριστικά (απότομες διακριτές εξάρσεις θορύβου). Σε αυτά τα συστήματα, το συνεστραμμένο ζεύγος καλωδίων το οποίο είναι αθωράκιστο, έχει ως συνέπεια τα μεταδιδόμενα σήματα να είναι επιρρεπή σε μεγάλο βαθμό από ηλεκτρομαγνητική παρεμβολή. Πιθανές πηγές κρουστικού θορύβου μπορεί να είναι οι διακόπτες ρεύματος, ηλεκτρικές συσκευές, αλλά και φυσικά καιρικά φαινόμενα. Βαριά παρεμβολή δημιουργείται επίσης από πολλαπλά ζεύγη συνεστραμμένων καλωδίων όπου έχουμε το φαινόμενο του cross-talk. Η παρεμβολή είναι “εκ’ γενετής” κρουστική και μη-στατική (δεν επαναλαμβάνεται με κάποια μορφή αλληλουχίας), και οδηγεί σε φτωχή ποιότητα εξυπηρέτησης. Η επίδραση κρουστικού θορύβου σε ένα σύστημα μετάδοσης (ADSL) εξαρτάται από την ενέργεια που έχει η κρουστικότητα, την διάρκεια, χρόνο άφιξης-επίδρασης, και από το φασματικό περιεχόμενο του. Όταν συμβαίνουν απομονωμένες επιδράσεις κρουστικού θορύβου τότε, το πλάτος που θα έχει ο θόρυβος θα είναι αρκετά μεγάλο, μεγαλύτερο από τον λευκό προσθετικό θόρυβο, ή από την επίδραση cross-talk. Τα παρών συστήματα ADSL έχουν σχεδιαστεί με γνώμονα την χειρότερη περίπτωση παρεμβολής-επίδρασης που μπορεί να συμβεί από κανάλι μετάδοσης το οποίο έχει μη-στατικό και non-Gaussian θόρυβο. Το σχήμα 1.5 δείχνει τρία ADSL σήματα τα οποία έχουν μετρηθεί κατά την “άφιξη” σε έναν καταναλωτή μιας συνδρομητικής υπηρεσίας. Τα σήματα δείχνουν ένα μεγάλο εύρος από φασματικά χαρακτηριστικά της κρουστικότητας, την ταχύτητα επαναληψιμότητας των κρουστικών παλμών, και το ποσοστό κρουστικότητας. Τα σχήματα 1.5a-b απεικονίζουν την χρήση γραμμικών και μη-γραμμικών φίλτρων, και την βελτίωση που έρχεται με την χρήση της μη-γραμμικού φίλτρου για την αφαίρεση του θορύβου και της παρεμβολής.



(a)



Σχήμα 1.5: (a-c) Διαφορετικά είδη θορύβου και παρεμβολής σε μια ADSL γραμμή. Ένα γραμμικό και μη-γραμμικό φίλτρο χρησιμοποιήθηκαν για την αντιμετώπιση των φαινομένων που εμφανίζει η γραμμή.

### Non-Gaussian Τυχαίες Διαδικασίες

Στην στατιστική επεξεργασία σήματος, τα σήματα μοντελοποιούνται σαν τυχαίες διαδικασίες και σε πολλές διαδικασίες επεξεργασίας σήματος, η ανάλυση του σήματος γίνεται σε στατιστικό επίπεδο για το παρατηρούμενο σήμα μας. Τι εννοούμε με αυτό; Η επιλογή του καταλλήλου μοντέλου παρατήρησης είναι απόλυτα αναγκαία, όπως είδαμε στην περίπτωση του FIR φίλτρου στην προσπάθειά του να φιλτράρει σήμα με μη-Γκαουσιανό θόρυβο το σήμα εξόδου δεν ήταν καθόλου αποδεκτό. Η κλασική γραμμική επεξεργασία σήματος, απορρέει από την δημοφιλή Gaussian παραδοχή. Το Gaussian μοντέλο εμφανίζεται φυσικά σε πολλές εφαρμογές, και αυτό ως αποτέλεσμα του κεντρικού οριακού θεωρήματος που αποδείχτηκε από τον De Moivre (1733).

*Θεώρημα 1.1 (Κεντρικό οριακό Θεώρημα)*

Έστω  $X_1, X_2, \dots, X_N$  μια αλληλουχία από *i.i.d.* τυχαίες μεταβλητές με μηδενικό μέσο όρο και πεπερασμένη διακύμανση  $\sigma^2$ . Τότε καθώς ο αριθμός των δειγμάτων τείνει στο άπειρο,  $N \rightarrow \infty$ , το κανονικοποιημένο άθροισμα:

$$S_N = \frac{1}{\sqrt{N}} \sum_{i=1}^N X_i$$

Συγκλίνει σχεδόν απόλυτα σε μια Gaussian μεταβλητή με μηδενικό μέσο όρο και με την ίδια διακύμανση όπως το  $X_i$ .

Το κεντρικό οριακό θεώρημα, εξηγεί την Gaussian φύση των διαδικασιών που δημιουργούνται από την υπέρθεση πολλών μικρών και ανεξάρτητων επιδράσεων. Για παράδειγμα, ο θερμικός θόρυβος ο οποίος δημιουργείται από την υπέρθεση μεγάλου αριθμού τυχαίων και ανεξάρτητων αλληλεπιδράσεων, που συμβαίνει σε ατομικό επίπεδο, έχει Gaussian κατανομή. Το κεντρικό οριακό θεώρημα λοιπόν δικαιολογεί την εμφάνιση Gaussian στατιστικών σε πραγματικά παραδείγματα.

Παρ' όλα αυτά, σε μια πολύ μεγάλη γκάμα εφαρμογών, το Gaussian μοντέλο δεν προσφέρει καθόλου ικανοποιητική απόδοση, και για αρχή κάτι τέτοιο φαίνεται να έρχεται σε αντιπαράθεση με το κεντρικό οριακό θεώρημα. Μια πιο προσεχτική ματιά όμως στους “όρους χρήσης” του κεντρικού οριακού θεωρήματος, βλέπουμε ότι **για να είναι δεκτό το “αποτέλεσμα” του κεντρικού οριακού θεωρήματος, θα πρέπει όλοι οι ανεξάρτητοι όροι που θα απαρτίζουν το άθροισμα, να έχουν πεπερασμένη διακύμανση, και όχι άπειρη**. Εάν οι τυχαίες μεταβλητές έχουν άπειρη διακύμανση, τότε το κεντρικό οριακό θεώρημα συγκλίνει σε μια μη-Gaussian κατανομή. Αυτή η σημαντική γενίκευση του κεντρικού οριακού θεωρήματος εξηγεί τις προφανείς αντιπαραδοχές με την “παραδοσιακή” του έκδοση. Με τον ίδιο τρόπο όπως το Gaussian μοντέλο αντλεί την ισχύ του από το κεντρικό οριακό θεώρημα, το γενικευμένο κεντρικό οριακό θεώρημα είναι ένας σημαντικός “βραχνάς” στην ανάπτυξη των μοντέλων που λειτουργούν σύμφωνα με την κρουστική φύση αυτών των σημάτων και των εργαλείων που λειτουργούν σε μη-Gaussian περιβάλλοντα.

Ίσως η απλούστερη προσέγγιση στο να αποδώσουμε τα αποτελέσματα των non-Gaussian σημάτων είναι να ανιχνεύσουμε τις outliers που μπορεί να υπάρχουν στο λαμβανόμενο σετ δειγμάτων, να τις αποβάλλουμε-καταστειλούμε, και έπειτα να εφαρμόσουμε αλγόριθμους επεξεργασίας σήματος. Αυτή η προσέγγιση όμως, έχει πολλά μειονεκτήματα. Πρώτον, η ανίχνευση των outliers δεν είναι απλή, ιδίως όταν αυτά είναι “δεμένα” με τα υπόλοιπα δείγματα. Δεύτερον, η αποτελεσματικότητα-αποδοτικότητα αυτών των μεθόδων δεν είναι η καλύτερη και είναι δύσκολο να “μετρηθεί” έτσι ώστε να μπορούμε να κάνουμε σύγκριση ανάμεσα σε αυτές τις μεθόδους.

- **outliers/παρεκτραπόμενη τιμή:** Μέσα σε ένα δείγμα 'n' παρατηρήσεων είναι δυνατόν να υπάρχουν τιμές όπου η τιμή τους να είναι τόσο μακριά από το εύρος δυνατών τιμών που μπορούμε να λάβουμε, έτσι ώστε να μας δημιουργηθεί η απορία μήπως αυτές οι τιμές είναι "ξένες" ή με άλλα λόγια, να μην ανήκουν καν στο σετ των εκπεμπόμενων σημάτων, και να προέρχονται από κάποια παρεμβολή, ή από κάποιο λάθος στην τεχνική δειγματοληψίας)

### Λόγος ύπαρξης Stable κατανομών μέσω του Γενικευμένου Κεντρικού Οριακού Θεωρήματος

Όπως τα Gaussian σήματα, μια ευρεία γκάμα από non-Gaussian διαδικασίες οι οποίες συναντώνται σε πρακτικές εφαρμογές, προκύπτουν από την υπέρθεση πολλών μικρών ανεξάρτητων φαινομένων. Εν αρχή, αυτό μπορεί να δείξει μια αντιπαράθεση με το κεντρικό οριακό θεώρημα το οποίο δηλώνει ότι το όριο πολλών μικρών και ανεξαρτήτων φαινομένων, οδηγεί σε Gaussian κατανομή. Μια προσεκτική όμως ματιά στο κεντρικό οριακό θεώρημα και τους όρους που την διέπουν, παρατηρούμε ότι για να ισχύει το θεώρημα αυτό, θα πρέπει η διακύμανση των **υπερθεμένων** τυχαίων μεταβλητών να είναι πεπερασμένη. Εάν η διακύμανση των τυχαίων μεταβλητών είναι άπειρη, τότε μια πολύ σημαντική γενίκευση του κεντρικού οριακού θεωρήματος, αναδύεται. Αυτή η γενίκευση εξηγεί τις προφανείς αντιφάσεις της "στάνταρ" έκδοσης του κεντρικού οριακού θεωρήματος, και την παρουσία των non-Gaussian διαδικασιών στην πράξη.

#### Θεώρημα 2.1 (Γενικευμένο Κεντρικό Οριακό Θεώρημα)

Έστω  $X_1, X_2, \dots$  μια *i.i.d.* ακολουθία από (πιθανόν διορθωμένα όσον αφορά την σκεύρωση), δείγματα. Τότε θα υπάρχουν σταθερές  $\alpha_n$  έτσι ώστε αν  $n \rightarrow \infty$ , το άθροισμα :

$$\alpha_n (X_1, X_2, \dots) \xrightarrow{d} Z$$

Εάν και μόνο εάν η  $Z$  είναι μια Stable τυχαία μεταβλητή με  $0 < a \leq 2$ .

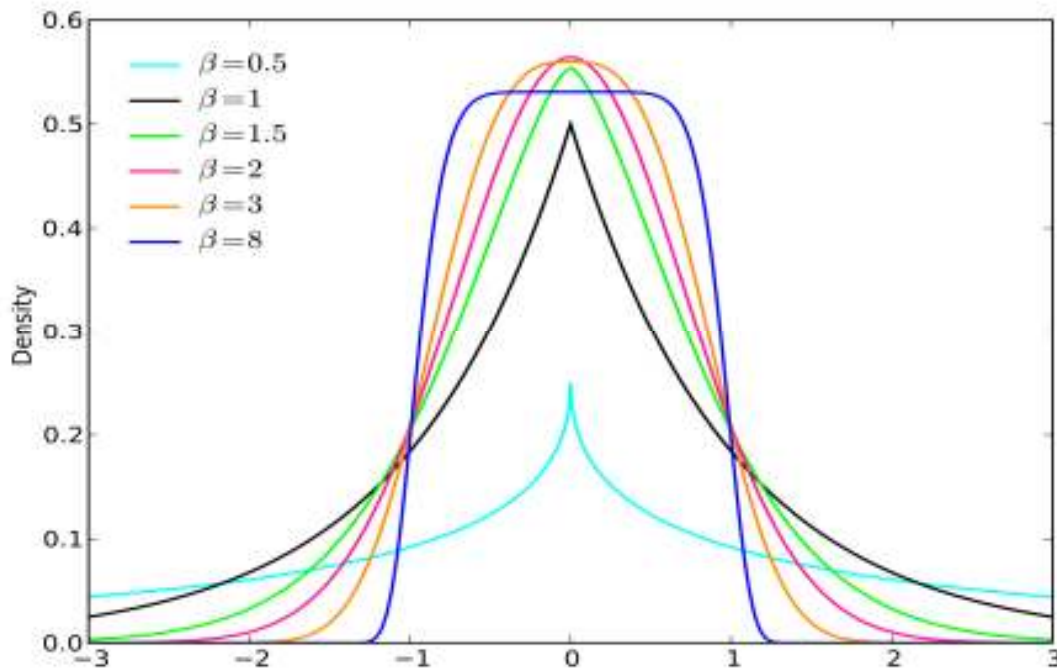
Έτσι, μπορούμε να πούμε όπως το Gaussian μοντέλο κατανομής, αντλεί την δύναμή του από το κεντρικό οριακό θεώρημα, το γενικευμένο κεντρικό οριακό θεώρημα αποτελεί μια σπουδαία 'αντιπαράδοχή' στην χρήση Stable μοντέλων, σε πρακτικά προβλήματα.

Στην αρχή, η χρήση άπειρης διακύμανσης στον ορισμό του γενικευμένου κεντρικού οριακού θεωρήματος, μπορεί να οδηγήσει σε σκεπτικισμό, καθώς το να έχουμε άπειρη διακύμανση σε πραγματικά δεδομένα, και ταυτόχρονα το μοντέλο παρατήρησης να έχει περιορισμένο εύρος, είναι αντιφατικό. Πρέπει να σημειωθεί παρόλα αυτά, ότι η διακύμανση δεν είναι τίποτε άλλο παρά ένα μέτρο του να

μετρήσουμε το πόσο "απλωμένη" είναι η κατανομή μου, και δεν είναι χρήσιμο μέτρο για όλες τις κατανομές. Η παράμετρος που μας δείχνει το πόσο "απλωμένη" είναι η κατανομή μας όταν μιλάμε για Stable κατανομές, είναι η παράμετρος 'γ'. Ο λόγος που γίνεται αυτό, δεν είναι άλλος από το να μπορούμε να γνωρίζουμε το σχήμα της κατανομής μας (θυμηθείτε ότι οι Stable κατανομές έχουν την ιδιότητα να παραμετροποιείται η σκευρότητα τους). Η Gaussian κατανομή για παράδειγμα, έχουμε συνηθίσει να την χρησιμοποιούμε στην μοντελοποίηση δεδομένων που φαινομενικά έχουν ανώτερο και κατώτερο όριο στην δυνατή τιμή που μπορεί να αποκτήσουν, αν και η Gaussian κατανομή έχει άπειρο εύρος δυνατών τιμών.

### Γενικευμένες Gaussian Κατανομές και σταθμισμένοι μέσοι όροι

Μια προσέγγιση στην μοντελοποίηση της παρουσίας των outliers είναι να πάρουμε την συνάρτηση πυκνότητας πιθανότητας της γενικευμένης Gaussian κατανομής, και να αφήσουμε την παράμετρο που καθορίζει τον εκθετικό ρυθμό εξασθένησης της ουράς (της κατανομής), να πάρει όποια τιμή εμείς επιθυμούμε.



Παρατηρήστε ότι όταν η παράμετρος που καθορίζει το σχήμα της κατανομής παίρνει την τιμή '2', η κατανομή είναι η Gaussian, και όταν πάρει την τιμή '1', τότε η κατανομή είναι η Laplace. Για πολύ μεγάλες τιμές της παραμέτρου σχήματος, η κατανομή τείνει να είναι η ομοιόμορφη (uniform)

Η συνάρτηση πυκνότητας πιθανότητας της γενικευμένης Gaussian κατανομής, είναι

$$pdf : \frac{\beta}{2\alpha\Gamma(1/\beta)} e^{-(|x-\mu|/\alpha)^\beta}$$

- όπου ‘α’ είναι η παράμετρος κρουστικότητας της κατανομής (καθορίζουμε δηλαδή το πόσο κρουστική μορφή θα έχει η κατανομή)
- όπου ‘β’ είναι η παράμετρος σχήματος της κατανομής
- όπου  $\Gamma(\cdot)$  είναι η συνάρτηση Γάμμα
- όπου ‘μ’ είναι ο μέσος όρος της κατανομής

Είναι πολύ ενδιαφέρον η περίπτωση της εξασθένησης πρώτης τάξης, η οποία οδηγεί στην Laplacian κατανομή. Έτσι, χρησιμοποιώντας την γενικευμένη Gaussian κατανομή, η οποία μπορεί να παραμετροποιηθεί όσον αφορά την μορφή της ‘πειράζοντας’ όρους όπως η διασπορά, μπορούμε να κατασκευάσουμε εκτιμητές οι οποίοι λειτουργούν με αρκετά καλή απόδοση σε κρουστικά περιβάλλοντα. Αποδεικνύεται ότι τα φίλτρα σταθμισμένων μέσων όρων, είναι βέλτιστα όταν λειτουργούν για δείγματα τα οποία υπακούουν τους όρους Laplacian στατιστικής, έτσι όπως τα γραμμικά φίλτρα είναι βέλτιστα όταν λειτουργούν με δείγματα τα οποία ακολουθούν τους όρους Gaussian κατανομής. Γενικότερα, τα φίλτρα σταθμισμένων μέσων όρων, είναι αποδοτικότερα από τα γραμμικά φίλτρα σε κρουστικής φύσεως περιβάλλοντα και αυτό λόγω της έντονης ουράς που εμφανίζει η Laplacian κατανομή.

### *Εκτίμηση μέσου όρου της κατανομής*

Η θεωρία της εκτίμησης είναι ένας κλάδος της στατιστικής που ασχολείται με το πρόβλημα της ανάκτησης πληροφορίας, ανάλογα με τις ιδιότητες της τυχαίας διεργασίας, μέσα από ένα σετ από παρατηρούμενα δείγματα. Έτσι, η θεωρία της εκτίμησης, βρίσκεται στην καρδιά της στατιστικής επεξεργασίας σήματος. Δοσμένης μιας παρατηρούμενης κυματομορφής  $\{X(n)\}$ , ο στόχος μας είναι να “τραβήξουμε” πληροφορία η οποία είναι ενσωματωμένη στο παρατηρούμενο σήμα. Αποδύκνεται ότι μπορούμε να μοντελοποιήσουμε την ενσωματωμένη πληροφορία, παραμετρικά. Αυτό εάν μια παράμετρος έστω ‘β’ η οποία αντικατοπτρίζει το σημείο του ενδιαφέροντος μας, επιθυμούμε να βρούμε την τιμή της, ή για το θέσουμε πιο σωστά, επιθυμούμε να βρούμε την καλύτερη προσέγγιση της. Αυτή η παράμετρος μπορεί να είναι ο τοπικός μέσος όρος, η διακύμανση, το τοπικό εύρος, ή οποιαδήποτε άλλη παράμετρος μπορεί να σχετίζεται με την ληφθέντα κυματομορφή. Φυσικά, το να βρούμε το σωστό παραμετρικό μοντέλο, είναι κρίσιμο.



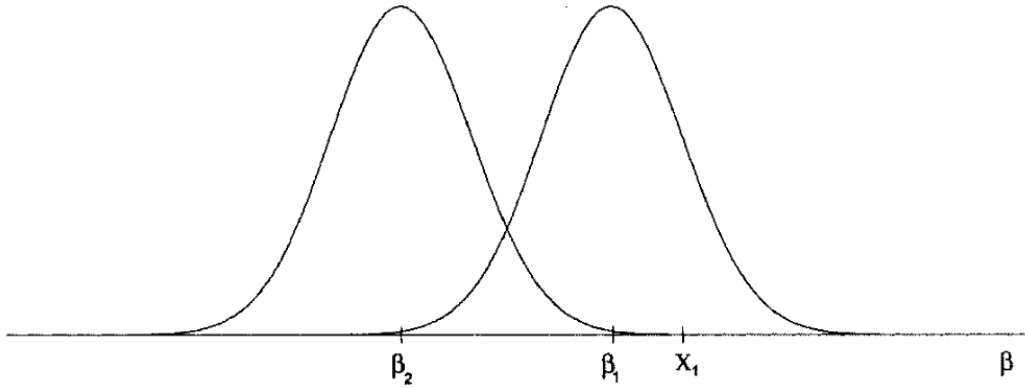
*Location Estimation-Εύρεση του μέσου όρου της κατανομής που σχηματίζουν τα δείγματα.*

Λόγω του ότι τα παρατηρούμενα σήματα είναι από την φύση τους τυχαία, εκφράζονται από μια συνάρτηση πυκνότητας πιθανότητας (pdf),  $f(x_1, x_2, \dots, x_N)$ . Η pdf μπορεί να παραμετροποιηθεί από μια άγνωστη παράμετρο 'β'. Η παράμετρος 'β' έτσι ορίζει μια κλάση από pdf's όπου το κάθε μέλος ορίζεται από μια ορισμένη τιμή της 'β'. Για παράδειγμα, εάν το σήμα μας αποτελείται από μια μόνο παρατήρηση (N=1) και η 'β' είναι ο μέσος όρος, η pdf υπό το Gaussian μοντέλο, θα είναι:

$$f(x_1; \beta) = \frac{1}{\sqrt{2\pi\sigma^2}} \exp\left[-\frac{(x_1 - \beta)^2}{2\sigma^2}\right]$$

Το ερωτηματικό στην έκφραση  $f(x_1; \beta)$ , δηλώνει ακριβώς αυτό: ποια είναι η τιμή της παραμέτρου 'β', εάν γνωρίζω τι τιμή έχει η παρατήρηση  $x_1$ . Θα μπορούσα να το γράψω και  $f(\beta | x_1)$ , σε μορφή υπό-συνθήκη πιθανότητας.

Μπορείτε να δείτε στο σχήμα 1.7 ότι για διάφορες τιμές του 'β' έχουμε και διαφορετικό μέσο όρο στην κατανομή μας. Αφού η τιμή του 'β' επηρεάζει την πιθανότητα του  $x_1$ , προφανώς θα μπορούμε να ανακτήσουμε (κάτω από κάποια πιθανότητα λάθους φυσικά) την τιμή του 'β', μέσω της παρατήρησης  $x_1$ . Για παράδειγμα, εάν η τιμή του παρατηρούμενου δείγματος  $x_1$  είναι ένας μεγάλος θετικός αριθμός, τότε η παράμετρος 'β' είναι πιο πιθανό να είναι ίση με 'β<sub>1</sub>', παρά με 'β<sub>2</sub>', όπως φαίνεται στο σχήμα 1.7. Προσέξτε ότι η 'β' επηρεάζει την τοποθέτηση της pdf. Για αυτό και λέμε την παράμετρο 'β', *παράμετρο θέσης/location parameter*, και οι κανόνες που ακολουθούμε για να ανακτήσουμε την τιμή της 'β' μέσα από την παρατήρηση, τους ονομάζουμε *εκτιμητές θέσης/location estimators*. Παρόλο που υπάρχει η δυνατότητα να συσχετίσουμε και άλλες παραμέτρους σχετικά με τα παρατηρούμενα δείγματα, η *θέση/location* είναι μια παράμετρος που παίζει ρόλο/κλειδί στην σχεδίαση αλγορίθμων φιλτραρίσματος.



Σχήμα 1.7 Εκτίμηση της παραμέτρου  $\beta$  βασιζόμενοι στην παρατήρηση  $X_1$

### Non-Gaussian Μοντέλα Κατανομών

Η Gaussian κατανομή χρησιμοποιείται συχνά στην επεξεργασία σήματος. Θεωρητικά από το Κεντρικό Οριακό Θεώρημα, το Gaussian μοντέλο έχει κερδίσει μια περίοπτη θέση στην στατιστική και στην μηχανική. Υπάρχουν παρόλα αυτά εφαρμογές όπου μπορεί να έχουμε τυχαίες διεργασίες, οι οποίες να μην ακολουθούν τους κανόνες της Gaussian στατιστικής. Συχνά, οι διεργασίες που συναντώνται στην πράξη, είναι κρουστικές και έτσι δεν μπορούν να περιγραφούν με συμβατικές Gaussian κατανομές.

Για να μοντελοποιήσουμε non-Gaussian διαδικασίες, επιστρατεύσαμε μια ευρεία γκάμα από κατανομές οι οποίες έχουν πιο έντονες ουρές από την Gaussian κατανομή. Σε αυτό το κεφάλαιο θα παρουσιάσουμε διάφορες προσεγγίσεις, αλλά, θα επικεντρωθούμε σε δυο κλάσεις κατανομών, τις επονομαζόμενες γενικευμένες Gaussian κατανομές, και την κλάση των Stable κατανομών.

### Γενικευμένες Gaussian Κατανομές

Το κεντρικό οριακό προσφέρει μια θεωρητική απόδειξη για την παρουσία Gaussian διαδικασιών στην φύση. Λόγω της φύσης της Gaussian κατανομής, μπορούμε να χρησιμοποιήσουμε γραμμικές μεθόδους εκτίμησης, και γενικότερα ένα μεγάλο μέρος των αλγόριθμων επεξεργασίας σήματος βασίζεται πάνω σε αυτήν την γραμμικότητα. Μια μεγάλη βοήθεια προέκυψε από το κεντρικό οριακό θεώρημα, το οποίο μας βοήθησε να κατανοήσουμε την αλληλεπίδραση ενός μεγάλου αριθμού ανεξαρτήτων γεγονότων, αλλά και την υποδομή της χρήσης γραμμικών μεθόδων, ακόμα και σε σήματα τα οποία δεν ακολουθούν Gaussian κατανομή.

Μια προσέγγιση που χρησιμοποιείται στην μοντελοποίηση non-Gaussian διαδικασιών, είναι να ξεκινήσουμε από το Gaussian μοντέλο και να το “παραμετροποιήσουμε”, έτσι ώστε να μπορεί να χρησιμοποιηθεί στην αντιμετώπιση

των outliers (σαν outlier μπορούμε να εκφράσουμε και μια διαδικασία θορύβου, διότι δεν προέρχεται από τη ίδια κατανομή πιθανότητας με αυτή των σημάτων μας, και το πλάτος μπορεί να υπερβεί κατά πολύ το μέγιστο αναμενόμενο πλάτος των σημάτων που λαμβάνω, λόγω της πιθανής κρουστικής φύσης του θορύβου). Έστω λοιπόν ότι η πηγή μου παρέχει σήματα τα οποία ακολουθούν Gaussian κατανομή, και “μολύνονται” με θόρυβο (με οποιαδήποτε μορφή που μπορεί να έχει ο/οι θόρυβος/θόρυβοι. Τότε η *Gaussian mixture* (μια κατανομή η οποία είναι η υπέρθεση διαφορετικών Gaussian κατανομών) ή αλλιώς το “μολυσμένο” *Gaussian μοντέλο*, θα ακολουθεί αυτήν την προσέγγιση:

$$f(x) = (1-\varepsilon)f_n(x) + \varepsilon f_c(x)$$

- Όπου  $f_n(x)$  είναι η ονομαστική Gaussian συνάρτηση πυκνότητας πιθανότητας με διακύμανση  $\sigma_n^2$
- Όπου ‘ $\varepsilon$ ’ είναι μια μικρή θετική σταθερά, ( $0 \leq \varepsilon \leq 1$ ) η οποία ορίζει τον βαθμό που η  $f_n(x)$  μολύνεται από την  $f_c(x)$ , η οποία εδώ ακολουθεί και αυτή Gaussian κατανομή, με σχετικά μεγάλη διακύμανση ( $\sigma_c^2 \gg \sigma_n^2$ ).

Προκύπτει ότι ένα δείγμα από  $1/\varepsilon$  δείγματα, θα είναι μολυσμένο από την πηγή με την μεγαλύτερη διακύμανση. Το πλεονέκτημα από μια μολυσμένη Gaussian κατανομή είναι η απλότητα της εξομοίωσης της με την βοήθεια του ηλεκτρονικού υπολογιστή. Όμως η μίξη Gaussian κατανομών, παρουσιάζει μειονεκτήματα. Πρώτον, για να παραμετροποιήσουμε την διασπορά και την κρουστικότητα, θα πρέπει να χρησιμοποιήσουμε τρεις παραμέτρους ( $\varepsilon, \sigma_n, \sigma_c$ ), αριθμός κάπως μεγάλος. Δεύτερον και μάλλον πιο σημαντικό είναι ότι για να υπολογίσουμε την  $f(x)$ , θα πρέπει να γίνει η πρόσθεση δύο διαφορετικών κατανομών ( $f_n(x), f_c(x)$ ), ένα πρόβλημα το οποίο είναι δύσκολο διαχειρίσιμο στο πρόβλημα της εκτίμησης.

Ανάμεσα στις διάφορες επεκτάσεις που μπορεί να πάρει οι Gaussian κατανομές, τα πιο δημοφιλή είναι αυτά που χαρακτηρίζονται από την *generalized Gaussian Distribution*. Αυτές είναι γνωστές από αναφορές από τον Subjoin το 1923 και από τον Frechet το 1924. Μια ειδική περίπτωση από την κλάση των *generalized Gaussian Distributions* είναι γνωστή σαν την *Laplacian* κατανομή, η οποία έχει ακόμα παλαιότερες ρίζες. Ο Laplace την παρουσίασε περισσότερο από διακόσια χρόνια πριν. Στην *generalized Gaussian Distribution*, η παρουσία των outliers μπορούν να μοντελοποιηθούν, απλώς παραμετροποιώντας την Gaussian κατανομή, επιτρέποντας την μεταβλητή η οποία καθορίζει το εκθετικό “σβήσιμο” της ουράς της κατανομής, να είναι ελεύθερη να πάρει όποια τιμή θέλουμε (όποια τιμή μας βολεύει, ανάλογα με τις outliers τιμές). Έστω αυτή η ελεύθερη παράμετρο, να είναι η ‘ $\kappa$ ’.

*Ορισμός 2.1 (generalized Gaussian Distribution) η συνάρτηση πυκνότητας πιθανότητας για την generalized Gaussian Distribution, δίνεται από:*

$$f(x) = \frac{ka}{2\Gamma(1/k)} e^{-(a|x-\beta|)^k} \quad (2.2)$$

Όπου  $\Gamma(\cdot)$  είναι η συνάρτηση Γάμμα  $\Gamma(\chi) = \int_0^{\infty} t^{\chi-1} e^{-t} dt$ , 'α' είναι μια σταθερά η οποία

ορίζεται ως:  $a = \sigma^{-1} \sqrt{\Gamma(3/k) (\Gamma(1/k))^{-1}}$  και 'σ' είναι η τυπική απόκλιση.

Εδώ, η κλιμάκωση καθορίζεται από την παράμετρο  $\sigma > 0$  και η κρουστικότητα καθορίζεται από την παράμετρο  $k > 0$ . Όπως θα περιμέναμε, η (2.2) μετατρέπεται στην στάνταρ Gaussian κατανομή, όταν  $k=2$ . Συνεπώς αυτή θα είναι και η υψηλότερη τιμή που θα μπορεί να πάρει το  $k$ , και για μικρότερες τιμές του, μεταβαίνουμε σε διαφορετική κατανομή, με υψηλότερη κρουστικότητα και πιο "βαριές" ουρές (με αυτόν τον όρο εννοούμε ότι το φασματικό εύρος της κατανομής είναι μεγαλύτερο). Μια δεύτερη ιδιαίτερη περίπτωση για την *generalized Gaussian Distribution* συμβαίνει όταν  $k=1$ , το οποίο ορίζει το διπλό εκθετικό (*double exponential*), ή αλλιώς την Laplacian κατανομή:

$$f(x) = \frac{1}{\sqrt{2}\sigma} e^{-\frac{\sqrt{2}}{\sigma}|x-\beta|} = \frac{\lambda}{2} e^{-\lambda|x-\beta|} \quad (2.3)$$

Η επίδραση της μείωσης της παραμέτρου 'κ' στην ουρά της κατανομής, μπορεί να δειχθεί από το σχήμα 2.2. Όπως βλέπουμε από τα σχήματα, η Laplacian κατανομή έχει βαρύτερη ουρά από την Gaussian.

## Stable κατανομές και weighted Myriads

Αποδεικνύεται ότι μια μεγάλη ποικιλία από διεργασίες παρουσιάζουν τόσο κρουστική φύση, που χαρακτηρίζονται από κατανομή με αλγεβρική ουρά. Αυτές οι κρουστικές διεργασίες, που βρίσκονται στην επεξεργασία σήματος, παράγονται από την υπέρθεση πολλών ανεξάρτητων φαινομένων. Για παράδειγμα, η αντανάκλαση που δέχεται ένα σύστημα Radar, είναι το άθροισμα πολλών αντανάκλασεων από διάφορες (όχι επίπεδες φυσικά) επιφάνειες. Επίσης αναμεταδότες σε ένα τηλεπικοινωνιακό σύστημα πολλών χρηστών, 'γεννούν' σχετικά μικρά και ανεξάρτητα σήματα, το άθροισμα των οποίων μαζί με το εκπεμπόμενο σήμα, είναι αυτό που λαμβάνει τελικώς ο δέκτης. Επίσης, περιστρεφόμενα ηλεκτρικά μηχανήματα, παράγουν πολλούς κρουστικούς παλμούς που δημιουργούνται από τις ενώσεις ανάμεσα στα μέρη του μηχανήματος. Τέλος, ο στάνταρ ατμοσφαιρικός θόρυβος είναι γνωστό ότι προέρχεται από πολλές ηλεκτρικές εκκενώσεις που συμβαίνουν. Η θεωρητική πιστοποίηση στην χρήση Stable κατανομών βασίζεται στο γενικευμένο κεντρικό οριακό θεώρημα, το οποίο περιλαμβάνει το γνωστό "παραδοσιακό" Κεντρικό Οριακό Θεώρημα, σαν μια ιδιαίτερη περίπτωση.

Πληροφοριακά:

Μια τυχαία μεταβλητή  $X$  λέγεται *Stable* εάν μπορεί να είναι το όριο ενός κανονικοποιημένου αθροίσματος από *i.i.d.* τυχαίες μεταβλητές.

Το γενικευμένο κεντρικό οριακό θεώρημα, δηλώνει ότι εάν το άθροισμα *i.i.d.* τυχαίων μεταβλητών, με ή χωρίς πεπερασμένη διακύμανση, συγκλίνει σε μια κατανομή, η οριακή κατανομή θα πρέπει να ανήκει στην οικογένεια των *Stable* κατανομών. Έτσι, *non-Gaussian* διεργασίες, μπορούν να “αναδυθούν” σε πρακτικές εφαρμογές, σαν άθροισμα από τυχαίες μεταβλητές με τον ίδιο τρόπο όπως *Gaussian* διεργασίες.

Οι *Stable* κατανομές περιλαμβάνουν δύο σημαντικά σημεία: την κανονική *Gaussian* κατανομή και την *Stable* κατανομή. η *Cauchy* κατανομή, μπορεί να χρησιμοποιηθεί για να μοντελοποιήσει διεργασίες βαριάς κρουστικότητας. Αποδεικνύεται ότι για ένα μεγάλο εύρος από σήματα με *Stable* κατανομή, τα *weighted myriad* φίλτρα είναι τα βέλτιστα. Έτσι, τα *weighted myriad* φίλτρα τα οποία απορρέουν από το *Stable* μοντέλο είναι οι αντίστοιχες υλοποιήσεις των γραμμικών και *weighted median* φίλτρα για *Gaussian* και *Laplacian* περιβάλλοντα αντίστοιχα.

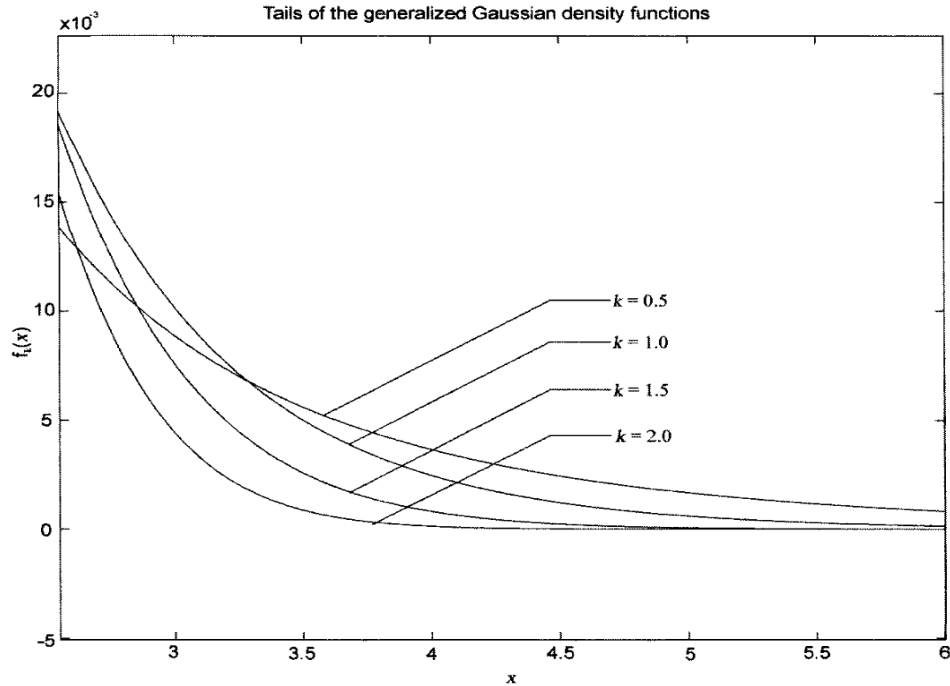
- **Mean estimator-εκτιμητής μέσου όρου.** Όπου η μέθοδος του είναι να αθροίσει τα παρατηρούμενα δείγματα και στην συνέχεια να διαιρέσουμε με τον αριθμό των δειγμάτων
- **Median estimator-εκτιμητής διαμέσου-median.** Όπου η μέθοδος του είναι να απορριφτούν πρώτα οι δυο ακραίες τιμές (η μέγιστη και η ελάχιστη τιμή) και αφού τοποθετηθούν τα δείγματα σε αύξουσα σειρά, να διαλέξουμε την κεντρική τιμή.
- **Myriad estimator.** Σε αυτήν την περίπτωση, το μοντέλο κατανομής είναι το *Cauchy* και η βέλτιστη τιμή βγαίνει παίρνοντας την παράγωγο της, δεδομένης της παραμέτρου κλιμάκωσης.

Οι *stable* κατανομές περιγράφουν μια μεγάλη γκάμα από κατανομές οι οποίες μας επιτρέπουν να παραμετροποιήσουμε την *skewness-σκευρότητα*, αλλά και το πόσο βαριά θα είναι η ουρά. Η κλάση των *Stable* κατανομών παρουσιάστηκε από τον *Levy* το 1925, και περιγράφονται από τέσσερις παραμέτρους:

- την παράμετρο  $\alpha \in (0, 2]$  η οποία εκφράζει την κρουστικότητα/*index of stability*,
- την παράμετρο κλιμάκωσης  $\gamma > 0$ ,
- την παράμετρο σταθερότητας/*skewness parameter*  $\delta \in [-1, 1]$
- και τέλος την παράμετρο θέσης/*location parameter*  $\beta \in \mathcal{R}$ .

Η παράμετρος σταθερότητας ‘ $\alpha$ ’, υποδεικνύει το πόσο πυκνή θα είναι η ουρά και δίνει έτσι την ευελιξία στο μοντέλο μας να μπορεί να μοντελοποιήσει μεγάλο εύρος από κρουστικές διαδικασίες. Η παράμετρος ‘ $\gamma$ ’ η οποία λέγεται και διασπορά, είναι

παρόμοια με την διακύμανση της Gaussian κατανομής. Η διακύμανση ισούται με  $\sigma = 2\sqrt{\gamma}$  στην Gaussian περίπτωση όταν  $\alpha=2$ . Όταν η παράμετρος *skewness* ισούται με μηδέν, τότε η Stable κατανομή, είναι συμμετρική γύρω από την παράμετρο θέσης 'β'. Οι συμμετρικές Stable κατανομές/*Symmetric stable distributions* συχνά αναφέρονται σαν *symmetric a-stable*, ή πιο απλά, *SaS*.



Σχήμα 2.2: Ουρές του της γενικευμένης Gaussian κατανομής, για διαφορετικές τιμές της σταθεράς 'κ'

Οι Stable κατανομές γίνονται ραγδαίως δημοφιλείς όταν έχουμε να κάνουμε με περιγραφή κάποιας κρουστικής διαδικασίας, για τους ακόλουθους λόγους:

1. Καλοί εμπειρικοί κανόνες προκύπτουν συχνά χρησιμοποιώντας Stable κατανομές για δεδομένα που παρουσιάζουν σκευρότητα και βαριές ουρές.
2. Υπάρχει μια ισχυρή θεωρητική απόδειξη ότι non-Gaussian Stable διαδικασίες υπάρχουν και στην πράξη, όπως σε ένα Poisson κατανεμημένο τηλεπικοινωνιακό δίκτυο το οποίο εμφανίζει παρεμβολές λόγω της πολλαπλής πρόσβασης που έχουμε σε αυτό, η αντανάκλαση σε ένα περιστρεφόμενο καθρέπτη, και η κίνηση του διαδικτύου.
3. Ο τρίτος λόγος για την μοντελοποίηση με Stable κατανομές, είναι ίσως και η πιο σημαντικός. Αυτό γιατί οι Stable κατανομές ικανοποιούν μια σημαντική γενίκευση του κεντρικού οριακού θεωρήματος, η οποία δηλώνει ότι το μόνο πιθανό όριο από κανονικοποιημένα αθροίσματα αποτελούμενο από i.i.d. όρους, είναι Stable.

Μια ευρεία ποικιλία από κρουστικές διαδικασίες οι οποίες βρίσκονται σε εφαρμογές επεξεργασίας σήματος, αναδύεται σαν η υπέρθεση πολλών μικρών ανεξαρτήτων φαινομένων. Εκεί που οι Gaussian κατανομές είναι ξεκάθαρα αναποτελεσματικές, οι Stable κατανομές έχουν όλες τις προδιαγραφές στο να μοντελοποιήσουν αυτές τις κρουστικές διαδικασίες. Έτσι οι Stable κατανομές είναι ελκυστικές αφότου η γενίκευση του κεντρικού οριακού θεωρήματος εξηγεί τις προφανείς αντιφάσεις μεταξύ της στάνταρ έκδοσης της, η οποία δεν θα μπορούσε να εξηγήσει διαφορετικά την παρουσία των σημάτων τα οποία έχουν βαριά/πυκνή ουρά.

### *Ορισμοί*

Οι Gaussian τυχαίες κατανομές υπακούουν την ιδιότητα που ορίζει ότι το άθροισμα δύο Gaussian κατανομών, θα μας δώσει μια Gaussian κατανομή. Έστω λοιπόν ότι έχουμε δύο τυχαίες Gaussian κατανομές  $X_1, X_2$  και  $a, b, c$  να είναι τρεις θετικές σταθερές.

$$aX_1 + bX_2 \stackrel{d}{=} cX + d \quad (2.4)$$

(το σύμβολο  $\stackrel{d}{=}$  δηλώνει ισότητα στην κατανομή)

Όπου  $d$  είναι μια πραγματική σταθερά. Φυσικά, αυτή η ιδιότητα ισχύει και για τις Stable κατανομές, και μια συμμετρική Stable τυχαία μεταβλητή η οποία είναι κατανομημένη γύρω από την τιμή 0, ικανοποιεί την συνθήκη  $X \stackrel{d}{=} -X$  λόγω συμμετρίας.

Πληροφοριακά, η παράμετρος της σταθερότητας ( $\alpha$ ) δηλώνει ότι η μορφή της  $X$  διατηρεί την κλιμάκωση και την σκεύρωση έπειτα από την πρόσθεση. Η ιδιότητα της σταθερότητας (2.4) για Gaussian τυχαίες μεταβλητές, μπορεί να αποδειχθεί ισχυρίζοντας ότι  $c^2 = a^2 + b^2$ , και  $d = (a + b - c)\mu$ , όπου ‘ $\mu$ ’ είναι ο μέσος όρος της μελετούμενης Gaussian κατανομής. Άλλες γνωστές κατανομές οι οποίες ικανοποιούν την ιδιότητα της Stable κατανομής, είναι οι κατανομές Cauchy και Levy (και οι δύο είναι μέλη της οικογένειας των Stable κατανομών). Η συνάρτηση πυκνότητας πιθανότητας για την Cauchy κατανομή με παράμετρο θέσης ‘ $\beta$ ’ είναι:

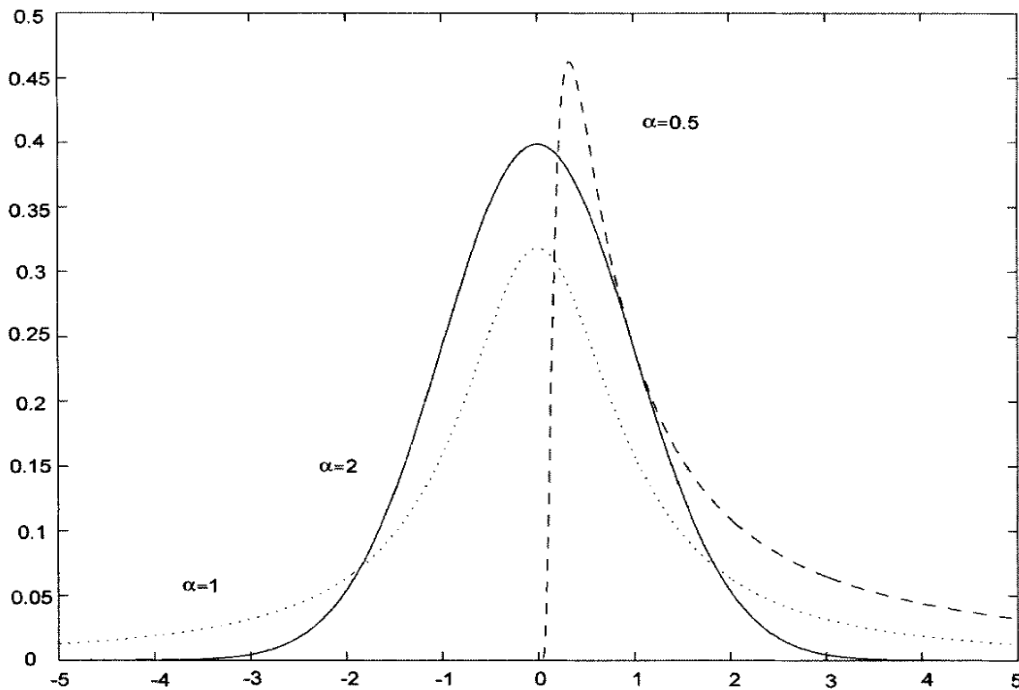
$$f(x) = \frac{1}{\pi} \frac{\gamma}{\gamma^2 + (x - \beta)^2}, \quad -\infty < x < \infty \quad (2.5)$$

Η Levy συνάρτηση πυκνότητας πιθανότητας μερικές φορές, αναφέρεται ως την Pearson κατανομή, είναι τελείως σκεβρωμένη στο διάστημα  $(0, \infty)$ . Η συνάρτηση πυκνότητας πιθανότητας για  $X \sim Levy(\gamma, \beta)$  έχει την μορφή:

$$f(x) = \sqrt{\frac{\gamma}{2\pi}} \frac{1}{(\chi - \beta)^{3/2}} \exp\left(-\frac{\gamma}{2(\chi - \beta)}\right), \beta < \chi < \infty$$

Το σχήμα 2.3 δείχνει το γράφημα της κανονικής Gaussian κατανομής, της Cauchy και της Levy. Βλέπουμε ότι η Gaussian και η Cauchy είναι συμμετρικές και έχουν καμπανοειδές σχήμα. Η κύρια διαφορά αυτών των δύο κατανομών, είναι η περιοχή κάτω από την ουρά της έκαστο κατανομής. Παρατηρούμε ότι η ουρά της Cauchy κατανομής σβήνει πιο αργά, κάτι που χαρακτηρίζει την πιο κρουστική φύση της – σε σχέση με την Gaussian κατανομή. Αν τώρα συγκρίνουμε και την Levy σε σχέση με τις άλλες δύο κατανομές, βλέπουμε ακόμα βαρύτερη ουρά (μεγαλύτερη κρουστικότητα), αλλά, και μια σκεύρωση που χαρακτηρίζει την κατανομή αυτή.

Γενικά, οι Stable κατανομές μας επιτρέπουν να τις παραμετροποιήσουμε όσον αφορά την σκευρότητα μέσω της παραμέτρου ‘δ’ όπως βλέπουμε στο σχήμα 2.4 την κατανομή μιας  $\alpha$ -stable τυχαίας μεταβλητής.



**Σχήμα 2.3** κατανομές πυκνότητας πιθανότητας για την κανονική κατανομή ( $\alpha=2$ ), Cauchy ( $\alpha=1$ ) και Levy ( $\alpha=0.5, \delta=1$ )

Παρόλο που μερικές πρακτικές διαδικασίες μοντελοποιούνται/περιγράφονται καλύτερα με σκευρωμένη κατανομή, εμείς θα επικεντρωθούμε στην συμμετρική Stable κατανομή, για τους εξής λόγους:

1. Οι διαδικασίες που βρίσκονται σε ένα μεγάλο αριθμό από εφαρμογές επεξεργασίας σήματος λειτουργούν με συμμετρική λογική.



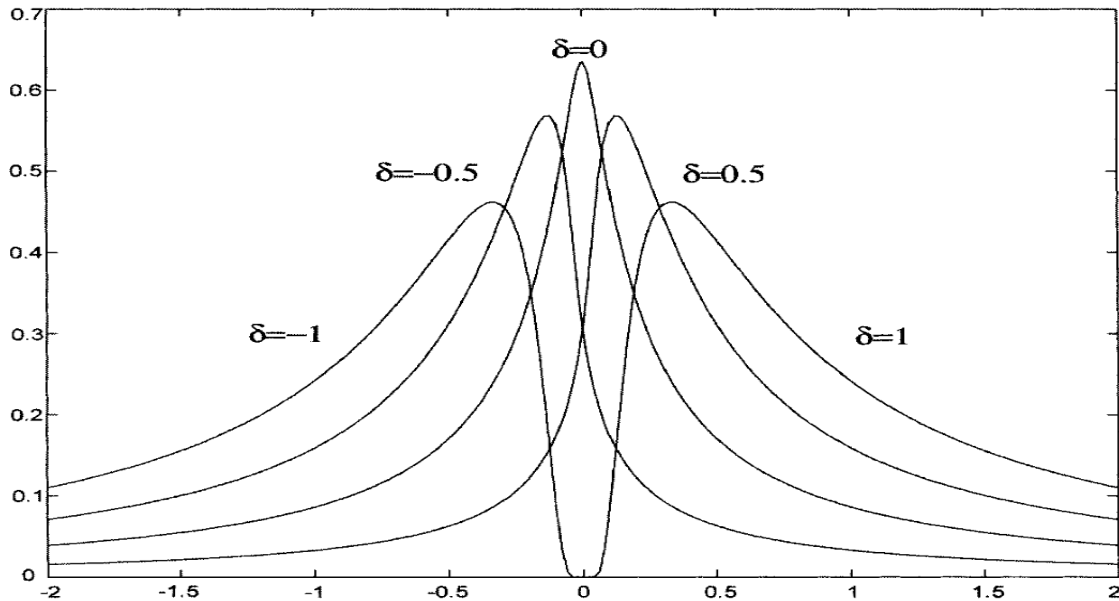
2. Τα ασύμμετρα μοντέλα οδηγούν σε πολύ μεγάλη πολυπλοκότητα το οποίο με την σειρά του οδηγεί σε πολύ δύσκολο υπολογισμό-εξομείωση των αλγορίθμων.
3. Τέλος και πιο σημαντικό, το να προσπαθήσουμε να κάνουμε μια εκτίμηση σε μια ασύμμετη κατανομή, είναι ένα πρόβλημα το οποίο δεν μπορεί να καθοριστεί με ακρίβεια, λόγω της πιθανής μη-γνώσης του βαθμού σκευρότητας που μπορεί να εμφανίζει η κατανομή αυτή.

### Συμμετρικές Stable Κατανομές

Οι συμμετρικές  $\alpha$ -stable κατανομές ορίζονται όταν η παράμετρος της σκευρότητας 'δ' οριστεί ίση με το μηδέν. Σε αυτήν την περίπτωση μια τυχαία μεταβλητή η οποία ακολουθεί την συμμετρική stable κατανομή με κλιμάκωση 'γ', ορίζεται ως:  $X \sim SaS(\gamma)$ . Παρόλο που ο ορισμός μιας Stable (σχέση 2.2) επαρκεί για τον χαρακτηρισμό όλων των Stable κατανομών, ένας δεύτερος και πιο ισχυρός χαρακτηρισμός τυχαίων Stable μεταβλητών απορρέει μέσω της χαρακτηριστικής συνάρτησης τους:

$$\Phi(\omega) = E \exp(j\omega X) = \int_{-\infty}^{\infty} \exp(j\omega x) f(x) dx \quad (2.7)$$

Όπου  $f(x)$  είναι η συνάρτηση πυκνότητας για την συγκεκριμένη τυχαία μεταβλητή.



**Σχήμα 2.4** Κατανομές πιθανότητας για σκευρωμένες stable κατανομές για παραμέτρους ( $\alpha=0.5, \gamma=1, \beta=0$ )

**Ορισμός 2.3** (χαρακτηριστική συνάρτηση *SaS* κατανομών)

Μια τυχαία μεταβλητή  $X$  είναι συμμετρική και *Stable* εάν και μόνο εάν  $X \stackrel{d}{=} AZ + B$  όπου  $0 < \alpha \leq 2, A \geq 0, B \in \Re$  και  $Z = Z(\alpha)$  είναι μια τυχαία μεταβλητή με χαρακτηριστική συνάρτηση:

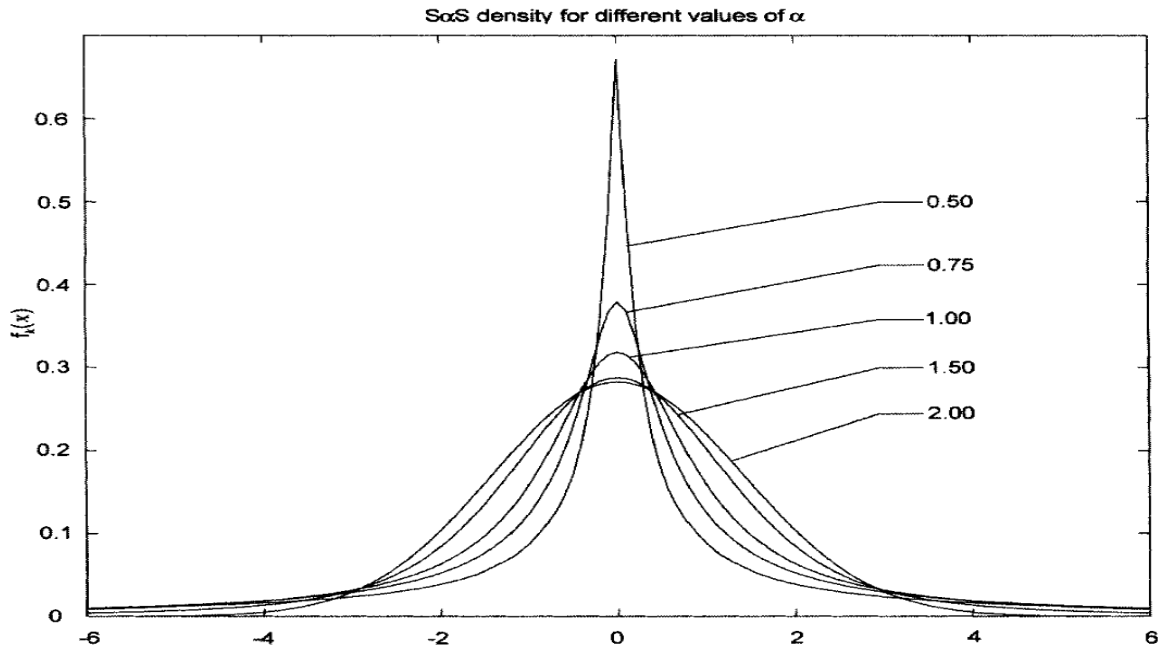
$$\Phi(\omega) = e^{-\gamma|\omega|^\alpha} \quad (2.8)$$

(Η χαρακτηριστική συνάρτηση μιας *Gaussian* τυχαίας μεταβλητής με μηδενικό μέσο όρο και διακύμανση  $\sigma^2$ , δίνεται από:  $\Phi(\omega) = \exp\left(-\frac{\omega^2\sigma^2}{2}\right)$ , από αυτήν την σχέση και για  $\alpha=2$ , μπορούμε να βγάλουμε μια σχέση μεταξύ της  $\sigma^2$  και της  $\gamma$ ).

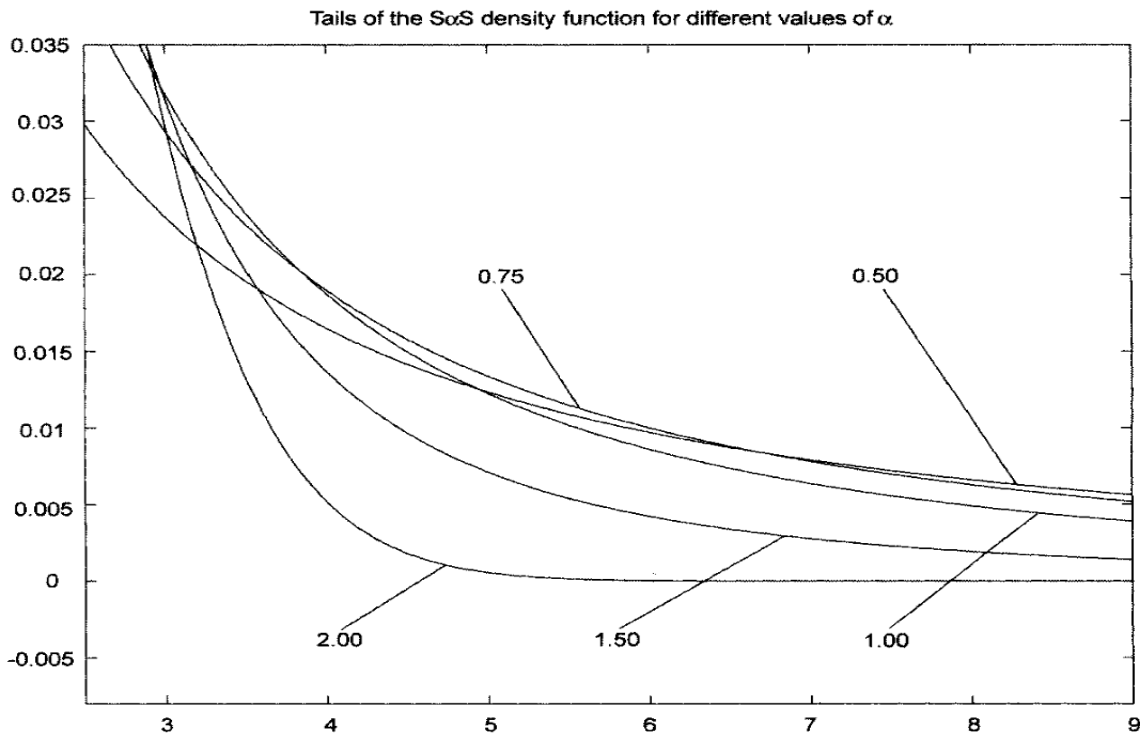
- Η παράμετρος διασποράς  $\gamma$  είναι μια θετική σταθερά η οποία σχετίζεται με την κλιμάκωση της κατανομής.
- Η παράμετρος  $\alpha$  ορίζει το ποσό της σταθερότητας (*index of stability*), ρυθμίζει το ποσό της κρουστικότητας και το πόσο πυκνή θα είναι η ουρά της κατανομής. Θα πρέπει όμως να βεβαιωθούμε ότι η τιμή του  $\alpha$  θα είναι στο διάστημα  $0 < \alpha \leq 2$ , όπου για  $\alpha=2$  η κατανομή παραπέμει στην στάνταρ *Gaussian* κατανομή. Για οποιαδήποτε άλλη τιμή της  $\alpha$  η κατανομή αποκτά πιο κρουστική μορφή με πιο έντονη ουρά.

Το σχήμα (2.5) δείχνει το γράφημα κανονικοποιημένων *Stable* κατανομών με μοναδιαία διασπορά ( $\gamma=1$ ). Προσέξτε ότι για χαμηλότερες τιμές του  $\alpha$  έχουμε κατανομή με πιο πυκνή ουρά (στο σχήμα 2.6).

**Ένα πολύ μεγάλο μειονέκτημα στην μοντελοποίηση *Stable* κατανομών είναι ότι υπάρχουν περιπτώσεις όπου η συνάρτηση πυκνότητας πιθανότητας, δεν έχει κλειστή μορφή (ανάπτυγμα).** Πιο συγκεκριμένα, υπάρχουν μόνο τρεις περιπτώσεις όπου η συνάρτηση πυκνότητα πιθανότητας μπορεί να εκφραστεί με κλειστού τύπου ανάπτυγμα: Η περίπτωση της *Gaussian* κατανομής ( $\alpha=2$ ), η *Cauchy* κατανομή ( $\alpha=1$ ) και η *Levy* κατανομή ( $\alpha=0,5$ ). Για οποιαδήποτε άλλη τιμή του  $\alpha$ , θα πρέπει να καταφύγουμε σε ανάπτυγμα σειρών, ώστε να είναι δυνατόν να περιγραφεί η κατανομή.



Σχήμα 2.5 κατανομές πιθανότητας stable κατανομής για διάφορες τιμές της παραμέτρου ουράς 'α'



Σχήμα 2.6 "Ουρές" συμμετρικών stable κατανομών για διάφορες τιμές της παραμέτρου ουράς 'α'

**Ορισμός 2.4** (Συμμετρικές Stable συναρτήσεις πυκνότητας πιθανότητας)

Γενικότερα, συμμετρικές κατανομές οι οποίες έχουν μηδενικό μέσο όρο και μοναδιαία διασπορά, έχουν συνάρτηση πυκνότητας πιθανότητας:

$$f_{\Lambda}(x) = \begin{cases} \frac{1}{\pi} \sum_{l=1}^{\infty} \frac{(-1)^{l-1} \Gamma(l\alpha+1)}{\gamma^{-l} l! |x|^{l\alpha+1}} \sin\left(\frac{l\alpha\pi}{2}\right), & \text{for } 0 < \alpha \leq 1 \\ \frac{\gamma}{\pi(x^2+\gamma^2)}, & \text{for } \alpha = 1 \\ \frac{1}{\pi\alpha} \sum_{l=0}^{\infty} \frac{(-1)^l x^{2l}}{(2l)! \gamma^{(2l+1)/\alpha}} \Gamma\left(\frac{2l+1}{\alpha}\right), & \text{for } 1 < \alpha \leq 2 \\ \frac{1}{2\sqrt{\gamma\pi}} \exp\left(-\frac{x^2}{4\gamma}\right), & \text{for } \alpha = 2 \end{cases}$$

*Στατιστικές Ιδιότητες εκτιμητών*

- *location estimation*
- *Μη-Προκατηλημένοι Εκτιμητές*
- *Αποδοτικοί Εκτιμητές*

Το φιλτράρισμα και η εκτίμηση παραμέτρων είναι άμεσα συσχετιζόμενες, λόγω του γεγονότος ότι η πληροφορία μπορεί να μεταφερθεί, ή και να εισαχθεί σε μία ή περισσότερες παραμέτρους που χαρακτηρίζουν το σήμα μας. Για παράδειγμα μπορούμε να ορίσουμε η πληροφορία να είναι ενσωματωμένη στην παράμετρο θέσης ( $\theta$ ), την διακύμανση ( $\sigma^2$ ), την φάση και φυσικά την συχνότητα. Τα χαρακτηριστικά δηλαδή που διέπουν την κατανομή την οποία ακολουθούν τα σήματά μας όσον αφορά την παράμετρο θέσης και την διακύμανση, και στα χαρακτηριστικά του εκπεμπόμενου σήματος. Για AM και FM σήματα για παράδειγμα, η πληροφορία είναι ενσωματωμένη στο στιγμιαίο πλάτος αλλά και στην στιγμιαία συχνότητα αντίστοιχα του διαμορφωμένου σήματος. Το πρόβλημα είναι, στο να μπορέσουμε να προσδιορίσουμε την παράμετρο που αναπαριστά την πληροφορία επιτυχώς. Αν μπορούσε κάποιος να παρατηρήσει απευθείας την παράμετρο αυτή, τότε το πρόβλημα της εκτίμησης σήματος, θα μπορούσε να αναπτυχθεί σε λίγες σελίδες. Στην πράξη όμως, η παρατήρηση εμπεριέχει θόρυβο και έτσι θα πρέπει να επιστρατεύσουμε την βοήθεια της στατιστικής, για κάνουμε μια εκτίμηση με όσο το δυνατόν μικρότερη πιθανότητα λάθους.

Ένα απλό παράδειγμα για να σας δείξουμε πως έχουν τα πράγματα:

- Έστω μια σταθερά  $\beta$  η οποία αντιπροσωπεύει το σήμα που εκπέμπω.
- Την στιγμή που λαμβάνω το σήμα (το  $\beta$  δηλαδή), λαμβάνω επίσης και θόρυβο.

Το παρατηρούμενο σήμα είναι λοιπόν:

$$X_i = \beta + Z_i, \quad i = 1, 2, \dots, N$$

Βλέπουμε λοιπόν ότι το λαμβανόμενο σήμα είναι το  $X_i$  και στην συγκεκριμένη περίπτωση είναι στατιστικά ανεξάρτητο από τα υπόλοιπα παρατηρούμενα δείγματα. Έτσι εάν το λαμβανόμενο σει παρατηρήσεων είναι  $X_1, X_2, \dots, X_N$ , ο στόχος είναι να επινοήσουμε έναν κανόνα, με βάση τον οποίο θα επεξεργαστούμε το σει παρατηρήσεων αυτό και να εκτιμήσουμε την τιμή του ' $\beta$ '. Γνωρίζοντας πλέον την τιμή του ' $\beta$ ', μπορούμε να θέσουμε ανάλογα το κατώφλι με βάση το οποίο θα γίνεται η διάκριση του '0' και του '1' (εάν πρόκειται για δυαδικό σύστημα εκπομπής). Στην θεωρία που θα ακολουθήσει, έχουμε θέσει την παράμετρο που συμβολίζει το πλάτος του εκπεμπόμενου σήματος, να είναι σταθερό (για παράδειγμα σε ένα δυαδικό σύστημα, ο "άσος" θα συμβολίζεται πάντα με ένα συγκεκριμένο πλάτος και το "μηδέν" με το δικό του συγκεκριμένο πλάτος).

Αφού λοιπόν η παράμετρος του πλάτους του σήματος είναι καθορισμένη, η μέθοδος που θα ακολουθήσουμε σε αυτήν την περίπτωση είναι γνωστή σαν *location estimation* ή αλλιώς εκτίμηση θέσης, και αυτό γιατί στόχος μας είναι να εντοπίσουμε την θέση που κατέχει η παράμετρος  $\beta$  (ποιο είναι το πλάτος δηλαδή).

Έχουν προταθεί πολλές μέθοδοι στην εκτίμηση του  $\beta$  για το παράδειγμα που μόλις είπαμε. Μια από αυτές τις μέθοδοι, είναι ο μέσος όρος των δειγμάτων που λαμβάνω:

$$\text{MEAN-MΕΣΟΣ ΟΡΟΣ} \quad \bar{\beta}_N = \bar{X} = \frac{1}{N} \sum_{i=1}^N X_i$$

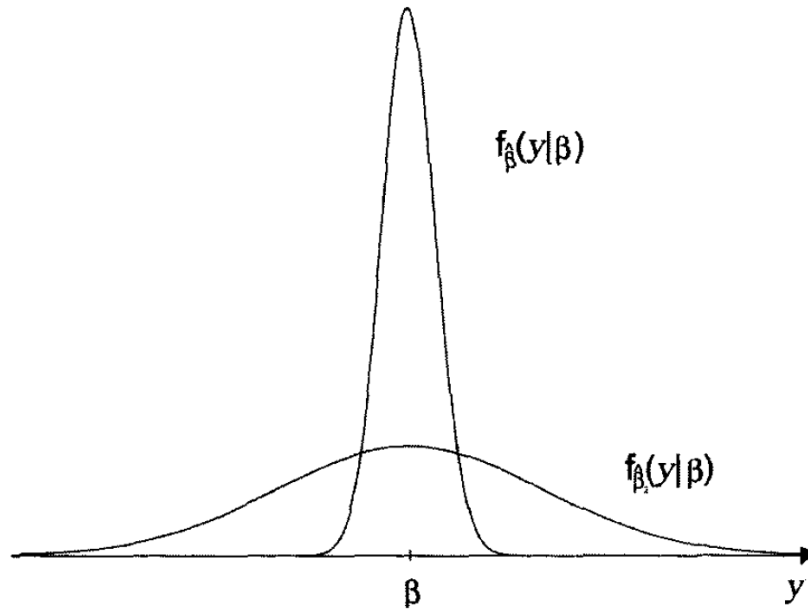
Ένας εναλλακτικός τρόπος εύρεσης του μέσου όρου είναι να "τοποθετήσουμε" τα δείγματά μας σε αύξουσα σειρά, και να επιλέξουμε το μεσαίο. Επίσης, ένας άλλος τρόπος είναι να απορίψουμε τα δύο δείγματα με τις ακραίες τιμές (ελάχιστο και μέγιστο) και με τις εναπομείναντες τιμές, να υπολογίσουμε τον μέσο όρο. Βλέπετε ότι υπάρχουν πολλοί τρόποι στον υπολογισμό του μέσου όρου, ανάλογα με ποιο κριτήριο ακολουθούμε.

Για κάθε εφαρμογή, όπως το παράδειγμα μας, υπάρχει ένας αριθμός από εκτιμητών που δύναται να χρησιμοποιήσουμε. Φυσικά, κάποιος εκτιμητής μπορεί να παρουσιάζει μεγάλη απόδοση σε κάποιες εφαρμογές, και μικρότερη σε κάποιες άλλες. Το να περιγραφεί το πόσο καλό είναι ένας εκτιμητής κάτω από συγκεκριμένες προϋποθέσεις είναι σημαντικό. Μπορούμε να πούμε ότι οι εκτιμητές είναι στην ουσία διαδικασίες, όπου αντιμετωπίζουν τις παρατηρήσεις σαν τυχαίες μεταβλητές,

και άρα οι ίδιοι οι εκτιμητές αντιμετωπίζονται στην ουσία σαν τυχαία μεταβλητή. Και αφού μιλάμε για τυχαίες μεταβλητές, επιστρατεύουμε την έννοια την συνάρτηση πυκνότητας πιθανότητας για την περιγραφή τους. Η συνάρτηση πυκνότητας πιθανότητας για την εκτίμηση (έστω της 'β' παραμέτρου) περιγράφεται σαν:  $f_{\hat{\beta}}(y|\beta)$  και διαβάζεται ως εξής: Αν γνωρίζω την τιμή της παραμέτρου  $\beta$ , πόσο πιθανό είναι η παρατήρησή μου  $y$  να έχει την τιμή  $\beta$ ;

### Μη-Προκατηλημένοι Εκτιμητές

Ή αλλιώς Unbiased Estimators. Μια τυπική κατανομή πυκνότητας πιθανότητας, δίνεται στο σχήμα 4.1, όπου η πραγματική τιμή της παραμέτρου δίνεται σαν γνωστή. Θα ήταν επιθυμητό για την εκτίμηση της παραμέτρου  $\beta$  (το σύμβολο για την εκτίμηση μιας παραμέτρου είναι η ίδια η παράμετρος με το καπελάκι από πάνω της- $\hat{\beta}$ ), να είναι όσο το δυνατό πιο κοντά στην πραγματική τιμή του  $\beta$ . Και πάλι αφού μιλάμε για εκτίμηση του  $\beta$  και το πόσο πιθανό είναι να έχουμε την πραγματική τιμή της παραμέτρου  $\beta$ , μιλάμε για μια κατανομή πιθανότητας η οποία έχει σαν κεντρική τιμή, την πραγματική τιμή της  $\beta$  και γύρω από αυτήν, την εκτιμώμενη  $\hat{\beta}$ .



**Σχήμα 4.1** Κατανομή πυκνότητας πιθανότητας η οποία σχετίζεται με έναν μη-προκατηλημένο εκτιμητή θέσης. Παρατηρήστε ότι ο μέσος όρος και των δύο κατανομών είναι ίσος με την τιμή 'β'.

## Ανίχνευση ασθενών σημάτων σε πρώτης τάξης εξαρτημένο θόρυβο

Σε μερικές περιπτώσεις, είναι δυνατόν να σχεδιάσουμε εκτιμητές για τους οποίους ο μέσος όρος των δυνατών εκτιμήσεων του  $\beta$  (όλα τα δυνατά  $\hat{\beta}$  δηλαδή), να ισούται με την πραγματική τιμή του  $\beta$ . τότε, λέμε ότι ένας εκτιμητής είναι *unbiased*.

- Μέσα από  $N$  δείγματα προς εκτίμηση της παραμέτρου  $\beta$ , που δίνονται ως  $\hat{\beta}$ , λέμε ότι έχουμε *unbiased estimator* όταν η μέση τιμή όλων των εκτιμήσεων ισούται με την πραγματική τιμή ' $\beta$ ':

$$E\{\hat{\beta}_N\} = \beta$$

- Επιπρόσθετα, η διακύμανση της εκτίμησης καθορίζει την ακρίβεια της εκτίμησης. Εάν ένας *unbiased estimator* έχει χαμηλή διακύμανση, τότε θα προσφέρει μια πιο αξιόπιστη εκτίμηση σε σχέση με άλλους εκτιμητές με υψηλότερη διακύμανση (βλέπε *Cramer-Rao Lower Bound*). Ο μέσος των δειγμάτων στο προηγούμενο παράδειγμα, είναι ένας *unbiased estimator* αφού  $E\{\hat{\beta}_N\} = \beta$ , με διακύμανση:

$$\text{var}\{\bar{\beta}_N\} = \text{var}\left\{\frac{1}{N} \sum_{i=1}^N X_i\right\} = \frac{\sigma^2}{N}$$

Όπου  $\sigma^2$  είναι η διακύμανση που έχει ο θόρυβος στο κανάλι διαβίβασης και  $N$  είναι ο αριθμός των δειγμάτων που λαμβάνω. Προφανώς η ακρίβεια της εκτίμησης βελτιώνεται όσο ο αριθμός των παρατηρήσεων αυξάνεται.

### Αποδοτικοί Εκτιμητές

Ο μέσος όρος και η διακύμανση της εκτίμησης είναι στοιχεία που ορίζουν την ποιότητα του εκτιμητή. Εάν στρέψουμε την προσοχή μας μόνο στους *unbiased estimators*, τότε στην πράξη η μέτρηση της αποδοτικότητας ενός εκτιμητή θα είναι ένα διδιάστατο σύστημα αξόνων με τον άξονα  $x$  να απεικονίζει όλους τους *unbiased estimators* και στον άξονα  $y$  να έχουμε την αποδοτικότητα του έκαστο εκτιμητή για την συγκεκριμένη εφαρμογή. Τότε το μόνο που έχουμε να κάνουμε, είναι να διαλέξουμε τον εκτιμητή με το μεγαλύτερο ποσό απόδοσης. Παρόλο που στην αρχή αυτό φαίνεται να είναι ένας απλοϊκός τρόπος της διαλογής του βέλτιστου *unbiased estimator*, βγαίνει ότι υπάρχει ένα κατώτερο όριο διακύμανσης για κάθε *unbiased estimator*. Έτσι, εάν βρεθεί ένας εκτιμητής με διακύμανση ίση με αυτό το κατώτερο όριο, τότε έχει βρεθεί και ο βέλτιστος εκτιμητής.

Αυτό το όριο είναι γνωστό και σαν *Cramer and Rao lower bound*. Έστω  $f(X; \beta)$  να είναι η συνάρτηση πυκνότητας των παρατηρήσεων  $\mathbf{X}$  έχοντας γνωστή την παράμετρο  $\beta$ . Εάν η  $\hat{\beta}$  είναι μια *unbiased* εκτίμηση της  $\beta$ , τότε η διακύμανση θα έχει σαν κατώτερο όριο:

$$\text{var}\{\hat{\beta}_N\} \geq \left( E \left\{ \left[ \frac{\partial}{\partial \beta} \ln f(X; \beta) \right]^2 \right\} \right)^{-1}$$

Εάν και μόνο εάν η μερική παράγωγος της *log-likelihood* (λογαριθμισμένης πιθανοφάνειας) είναι υπαρκτή, και η ολοκλήρωσή της να είναι δυνατή.

- Για ποιον λόγο θέλουμε να είναι δυνατή η παραγωγή της συνάρτησης πιθανότητας; Γιατί θέλουμε για όλο το δυνατό εύρος τιμών που μπορεί να πάρει το παρατηρούμενο σείτ μας, να εμφανίζει μια (οποιαδήποτε) πιθανότητα για **κάθε** δείγμα (να μην εμφανίζει ασυνέχειες η συνάρτηση).
- Επίσης επιθυμούμε να είναι δυνατή η ολοκλήρωσή της για τον απλούστατο λόγο: Γνωρίζουμε ότι εάν ολοκληρώσω μια κατανομή πιθανοτήτων (έστω την γνωστή σε όλος μας, την Gaussian κατανομή) σε όλο το δυνατό εύρος τιμών, δηλαδή από το  $-\infty$  έως το  $+\infty$ , τότε θα λάβω την τιμή 1. Άρα θέλουμε με την ολοκλήρωσή της να μας δίνει μια τιμή. Ένα παράδειγμα είναι το δυαδικό συμμετρικό σύστημα, όπου επιθυμώ να στείλω έστω το '0' και το '1', με ίση πιθανότητα έκαστο. Άρα αναμένω με την ολοκλήρωσή της συνάρτησης πυκνότητας πιθανότητας να λάβω την τιμή 0.5 (50%) και για το σήμα '0', αλλά και για το σήμα '1'.

Μια δεύτερη μορφή του *Cramer and Rao lower bound*, είναι:

$$\text{var}\{\hat{\beta}_N\} \geq \left( -E \left\{ \left[ \frac{\partial^2}{\partial \beta^2} \ln f(X; \beta) \right] \right\} \right)^{-1}$$

Και θα ισχύει μόνο και μόνο εάν η δεύτερη μερική παράγωγος της συνάρτησης πυκνότητας είναι υπαρκτή, και απόλυτα παραγωγίσιμη.

Η αποδοτικότητα μπορεί επίσης να χρησιμοποιηθεί σαν ένα μέτρο σύγκρισης μεταξύ δυο εκτιμητών. Μια εκτίμηση λέμε ότι θα είναι βέλτιστη εάν σε σύγκριση με μια άλλη εκτίμηση έχει χαμηλότερη διακύμανση. Εάν αυτή η σχετική αποδοτικότητα σχετιστεί με την τάξη της εκτίμησης, τότε προκύπτει το ακόλουθο συμπέρασμα: εάν η  $\hat{\beta}_N$  είναι *unbiased* και αποδοτική σε σχέση με το  $\hat{\beta}_{N-1}$ , για κάθε  $N$ , τότε λέμε ότι η  $\hat{\beta}_N$  εκτίμηση είναι **consistent** (μια εκτίμηση δηλαδή που άσχετα με τον αριθμό των δειγμάτων που χρησιμοποιεί για την ανάδειξη της παραμέτρου  $\hat{\beta}$ , το αποτέλεσμα της εκτίμησης θα είναι πάντα *unbiased* και θα παραμένει το ίδιο).



## Εφαρμογή της Μέγιστης Πιθανοφάνειας στις *Stable* Κατανομές

Έχοντας ένα σετ από παρατηρούμενα δείγματα, υπάρχει ένας αριθμός από διάφορους τρόπους που μπορούμε να χρησιμοποιήσουμε για να λάβουμε μια εκτίμηση. Ανάμεσα σε αυτούς τους τρόπους, η μέθοδος *maximum likelihood-ML* (μέγιστη πιθανοφάνεια) είναι η πιο δημοφιλής προσέγγιση διότι επιτρέπει την κατασκευή εκτιμητριών για δύσκολα και ασυνήθιστα προβλήματα. Η *ML* εκτίμηση βασίζεται στην εξής απλή αρχή: Διαφορετικές κατανομές, παράγουν διαφορετικά παρατηρούμενα δείγματα. Εάν έχουμε τα παρατηρούμενα δείγματα  $X_1, X_2, \dots, X_N$ , τότε λόγω του ότι όλα προέρχονται από την ίδια κατανομή  $f(X_1, X_2, \dots, X_N; \beta)$ , τότε στην λήψη θα καταλάβουμε το γεγονός της ομοιότητας των δειγμάτων αυτών και η εκτίμηση θα είναι πιο αποδοτική. Έστω τώρα πως λάμβανα την ίδια ακολουθία, και το δεύτερο δείγμα προερχόταν από διαφορετική κατανομή. Τότε θα καταλαβαίναμε ότι έχω 'παρείσακτο' δείγμα, και με κάποια διαδικασία/αλγόριθμο, θα απέρριπτα το δείγμα αυτό, η θα ζητούσα επανεκπομπή του σετ αυτού.

Η παράμετρος ' $\beta$ ' θεωρείται άγνωστη, αλλά σταθερή (δεν αλλάζει τιμή με την πάροδο του χρόνου) και στην εκτίμηση παραμέτρων, αυτό που προσπαθούμε να κάνουμε, είναι να βρούμε την καλύτερη διαδικασία της εκτίμησης της τιμής του  $\beta$ , μέσα από ένα σετ παρατηρούμενων δειγμάτων. Αν γίνει αυτό, τότε θα μπορέσω να θέσω ένα κατώφλι όπου θα μου καθορίζει την μορφή του σήματος που δέχομαι (π.χ. εάν το λαμβανόμενο σήμα η στάθμη του ξεπεράσει το κατώφλι αυτό, τότε το μεταφράζω σαν '1', αλλιώς το μεταφράζω σαν '0').

Στην μέθοδο *ML*, η καλύτερη εκτίμηση της  $\beta$  είναι η τιμή  $\hat{\beta}_{ML}$  όπου στο σημείο αυτό η συνάρτηση  $f(X_1, X_2, \dots, X_N; \beta)$  εμφανίζει ολικό μέγιστο. Για εκείνη την τιμή της  $\hat{\beta}$ , θα έχω την μέγιστη πιθανότητα εκτελέσω σωστή εκτίμηση.

$$\hat{\beta}_{ML} = \arg \max_{\beta} f(X_1, X_2, \dots, X_N; \beta)$$

Η συγκεκριμένη συνάρτηση διαβάζεται ως εξής:

Εάν γνωρίζω τα παρατηρούμενα δείγματα  $(X_1, X_2, \dots, X_N)$  τότε ποια θα είναι (για αυτό και το ερωτηματικό) η τιμή της παραμέτρου  $\beta$ , η οποία μας δίνει το μέγιστο της κατανομής  $f(X_1, X_2, \dots, X_N; \beta)$ .

Η συνάρτηση αυτή θα μπορούσε να γραφεί διαφορετικά:

$$\hat{\beta}_{ML} = \arg \max_{\beta} f(\beta | X_1, X_2, \dots, X_N)$$

Η οποία διαβάζεται με τον ίδιο ακριβώς τρόπο.

Η φιλοσοφία της Μέγιστης Πιθανοφάνειας είναι κομψή και απλή, και η παράμετρος που μας ενδιαφέρει να βρεθεί, είναι η θέση (*location*). Υποθέτοντας ανεξάρτητα δείγματα σε αυτό το μοντέλο, κάθε δείγμα σε αυτό το σετ θα ακολουθεί κατανομή:

$$P(X_i \leq x) = F(x - \beta)$$

Όπου  $F(\cdot)$  δηλώνει κατανομή η οποία είναι συμμετρική γύρω από το  $F(\cdot)$ .

### Εκτίμηση θέσης (*Location estimation*) σε Gaussian θόρυβο

Υποθέστε τα παρατηρούμενα δείγματα  $(X_1, X_2, \dots, X_N)$  τα οποία είναι i.i.d. δείγματα τα οποία ακολουθούν Gaussian κατανομή και σταθερά  $\beta$  (άγνωστη ακόμα) η οποία υποδηλώνει τον μέσο όρο της κατανομής (την θέση/*location* δηλαδή). Τότε η μέγιστη πιθανοφάνεια ML εκτίμηση του μέσου όρου είναι η τιμή  $\hat{\beta}$  η οποία μεγιστοποιεί την συνάρτηση πιθανοφάνειας:

$$\begin{aligned} f(X_1, X_2, \dots, X_N; \beta) &= \prod_{i=1}^N f(X_i - \beta) \\ &= \prod_{i=1}^N \frac{1}{\sqrt{2\pi\sigma}} e^{-(X_i - \beta)^2 / 2\sigma^2} \quad (4.6) \\ &= \left( \frac{1}{2\pi\sigma^2} \right)^{N/2} e^{-\sum_{i=1}^N (X_i - \beta)^2 / 2\sigma^2}. \end{aligned}$$

Η συνάρτηση πιθανοφάνειας στην (4.6) μεγιστοποιείται ελαχιστοποιώντας τον όρο στο εκθετικό (ελαχιστοποίηση λόγω του (-) μπροστά από το άθροισμα). Παραγωγίζουμε λοιπόν ως προς τον μοναδικό όρο του οποίου η μεταβλητή μπορεί να παραμετροποιηθεί, και αυτός θα είναι ο  $\beta$ .

$$\begin{aligned} \hat{\beta}_{ML} &= \arg \min_{\beta} \sum_{i=1}^N (X_i - \beta)^2 \Rightarrow \\ \hat{\beta}_{ML} &= \arg \min_{\beta} \sum_{i=1}^N (X_i^2 - 2X_i\beta + \beta^2) \Rightarrow \\ \hat{\beta}_{ML} &= \arg \min_{\beta} \left( \sum_{i=1}^N (X_i^2) - \sum_{i=1}^N (2X_i\beta) + \sum_{i=1}^N (\beta^2) \right) \Rightarrow \end{aligned}$$

$$(0) - \sum_{i=1}^N (2X_i) + (2N\beta) = 0 \Rightarrow$$

$$\sum_{i=1}^N (X_i) + (N\beta) = 0 \Rightarrow$$

$$\hat{\beta}_{ML} = \frac{1}{N} \sum_{i=1}^N (X_i)$$

Η τιμή η οποία ελαχιστοποιεί το άθροισμα λοιπόν, βρέθηκε ότι είναι ο μέσος όρος

$$\hat{\beta}_{ML} = \frac{1}{N} \sum_{i=1}^N (X_i)$$

Προσέξτε ότι ο μέσος όρος είναι βασισμένος αποκλειστικά στα παρατηρούμενα δείγματα και στον αριθμό των δειγμάτων που χρησιμοποιήσαμε.

### Εκτίμηση θέσης (*Location estimation*) σε γενικευμένο Gaussian θόρυβο

Υποθέστε τώρα ότι τα παρατηρούμενα δείγματα περιλαμβάνουν τιμές-*outliers* που ξεκάθαρα αποκλίνουν από τον μέσο όρο των συνολικών δειγμάτων. Οι μεγάλες αποκλίσεις όμως, έρχονται σε αντίθεση με το Gaussian μοντέλο και το κεντρικό οριακό θεώρημα. Μια εναλλακτική μέθοδος στο να μοντελοποιήσουμε τις αποκλίσεις με μια πιο ικανοποιητική κατανομή, η οποία θα είναι πιο ευέλικτη στο να 'συλλαμβάνει' αυτές τις μεγάλες αποκλίσεις που μπορεί να έχουν τα δείγματα, είναι να υιοθετήσουμε την γενικευμένη Gaussian κατανομή. Σε αυτή την περίπτωση, η συνάρτηση που χρησιμοποιούμε στην κατασκευή της εκτιμήτριας ML όσον αφορά την παράμετρο θέσης ( $\beta$ ) είναι:

Υπενθυμίζουμε ότι η συνάρτηση της *generalized Gaussian Distribution* είναι:

$$f(X) = Ce^{-a|X|^k}$$

Αν τώρα έχουμε  $N$  παρατηρούμενα δείγματα, τότε η κατανομή θα έχει μορφή:

$$f(X_1, X_2, \dots, X_N) = \prod_{i=1}^N Ce^{-a|X_i|^k}$$

Η οποία θα έχει σαν κεντρική τιμή το μηδέν.

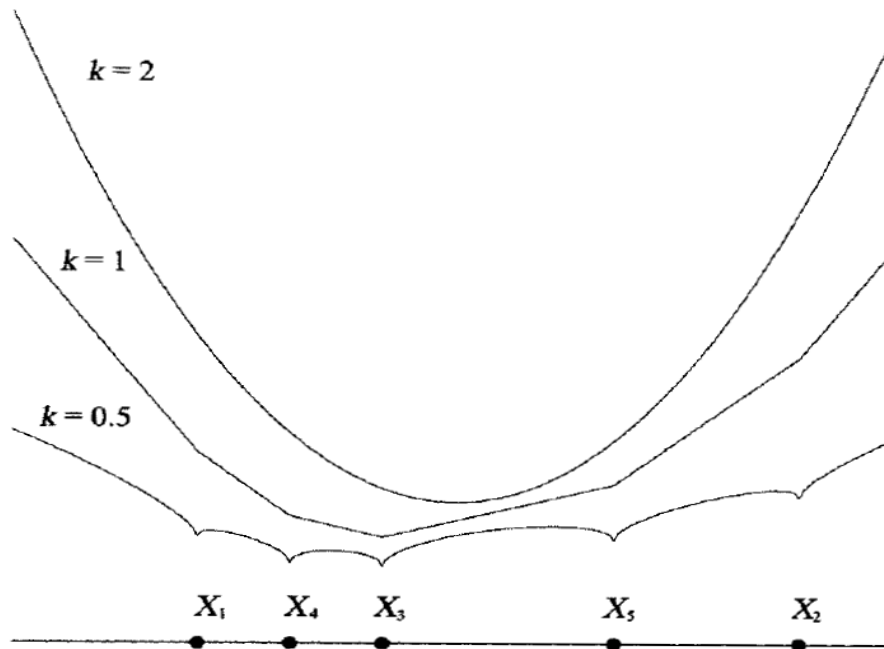
Για παράμετρο θέσης το  $\beta$ , η κατανομή θα είναι κεντραρισμένη γύρω από το μηδέν και θα γράφεται:

$$\begin{aligned} f(X_1, X_2, \dots, X_N; \beta) &= f(\beta | X_1, X_2, \dots, X_N) = \prod_{i=1}^N Ce^{-a|X_i - \beta|^k} \\ &= C^N e^{-a^k \sum_{i=1}^N |X_i - \beta|^k} \end{aligned} \quad (4.11)$$

Όπου 'C' και 'a' είναι σταθερές κανονικοποίησης και η 'k' είναι μια σταθερή παράμετρος η οποία δίνει την διασπορά των δεδομένων. Η μεγιστοποίηση της συνάρτησης πιθανοφάνειας είναι ίσο με την ελαχιστοποίηση το όρισμα του εκθετικού, το οποίο οδηγεί στην ακόλουθη εκτίμηση της παραμέτρου β:

$$\tilde{\beta}_{ML} = \arg \min_{\beta} \sum_{i=1}^N |X_i - \beta|^k \quad (4.12)$$

Όπως μπορούμε να δούμε στο σχήμα 4.2, για διάφορες τιμές της παραμέτρου διασποράς, λαμβάνουμε και διαφορετικές τιμές κόστους. Με την έννοια κόστος εννοούμε ότι, όταν στείλουμε την τιμή '1' και αποκωδικοποιηθεί σαν '0', μιλάμε προφανώς για λάθος. Το πόσο σοβαρό είναι ένα τέτοιο λάθος δεν είναι το ίδιο σε κάθε εφαρμογή. Δεν είναι το ίδιο να γίνει λάθος για ένα τραγούδι που παίζει από ψηφιακό ραδιόφωνο, και απλώς θα ακουστεί μια νότα λάθος, και μια στρατιωτική εφαρμογή όπου ο αριθμός των λαθών πρέπει να κρατηθεί όσο το δυνατόν πιο χαμηλά γίνεται.



Σχήμα 4.2 Συναρτήσεις κόστους για τα παρατηρούμενα δείγματα  $X_1 = -3, X_2 = 10, X_3 = 1, X_4 = -1$  και  $k = 0.5, 1$  και  $2$ .

Όταν η παράμετρος της διασποράς πάρει την τιμή 2, τότε το μοντέλο που μας δίνει τις συναρτήσεις κόστους 'μεταμορφώνεται' στο Gaussian μοντέλο είναι τετραγωνικό  $|\cos t|^2$ , και η τιμή του εκτιμητή είναι ίση με τον μέσο όρο των δειγμάτων. Όταν  $k < 1$ , δείχνεται ότι η συνάρτηση κόστους παρουσιάζει ποικίλα τοπικά ελάχιστα, και αυτά θα εμφανίζονται όταν το δείγμα πάρει όντως την τιμή του έκαστο X.

Όταν η παράμετρος της διασποράς πάρει την τιμή '1' τότε μιλάμε για το Laplacian μοντέλο, και η συνάρτηση του κόστους θα έχει την ίδια κλίση για τιμές ενδιάμεσων των διάφορων παρατηρήσεων (γραμμική και συνεχής συνάρτηση), και ο βέλτιστος

## Ανίχνευση ασθενών σημάτων σε πρώτης τάξης εξαρτημένο θόρυβο

εκτιμητής θα ελαχιστοποιεί το άθροισμα των απόλυτων αποκλίσεων (η απόκλιση που θα έχει το παρατηρούμενο δείγμα από το πραγματικό σήμα που εστάλει).

$$\tilde{\beta}_{ML} = \arg \min_{\beta} \sum_{i=1}^N |X_i - \beta| \quad (4.13)$$

Παρόλο που δεν είναι αμέσως προφανές, η λύση στο άνω πρόβλημα είναι ο μέσος όρος των δειγμάτων όπως θα δειχθεί στην συνέχεια.

Έστω η συνάρτηση κόστους (4.13) που πρέπει να ελαχιστοποιηθεί να την συμβολίσουμε  $L_1(\beta)$ . Για τιμές του  $\beta$  στο διάστημα  $-\infty < \beta \leq X_1$  (αυτό σημαίνει ότι το όρισμα του αθροίσματος θα είναι πάντα θετικός αριθμός), το  $L_1(\beta)$  απλοποιείται στο:

$$L_1(\beta) = \sum_{i=1}^N (X_{(i)} - \beta) = \sum_{i=1}^N X_{(i)} - N\beta$$

Άρα όπως είπαμε αφού είναι πάντα θετικό το όρισμα του αθροίσματος, ‘ξεφορτώθηκε’ το απόλυτο, και στην συνέχεια ‘έσπασε’ το άθροισμα με την επιμεριστική ιδιότητα. για τιμές όμως του  $\beta$  στο διάστημα  $X_{(j)} < \beta \leq X_{(j+1)}$ , το  $L_1(\beta)$  γράφεται:

$$\begin{aligned} L_1(\beta) &= \sum_{i=1}^N (X_{(i)} - \beta) + \sum_{i=j+1}^N (\beta - X_{(i)}) \Rightarrow \\ L_1(\beta) &= \sum_{i=j+1}^N (X_{(i)}) - \sum_{i=1}^j (X_{(i)}) - (N-2j)\beta \end{aligned}$$

Για  $j=1,2,\dots,N-1$ . Παρόμοια, για  $X_{(N)} < \beta < \infty$

$$L_1(\beta) = -\sum_{i=1}^N (X_{(i)}) + N\beta$$

Αν  $X_{(0)} = -\infty$  και  $X_{(N+1)} = \infty$ , και ορίζοντας  $\sum_{i=m}^n X_{(i)} = 0$  αν  $m > n$ , τότε συνδυάζοντας (4.14)-(4.16) έχουμε:

$$L_1(\beta) = \sum_{i=j+1}^N (X_{(i)}) - \sum_{i=1}^j (X_{(i)}) - (N-2j)\beta \quad j = 0,1,\dots,N$$

Για  $\beta \in (X_{(j)}, X_{(j+1)})$ . Όταν η συνάρτηση κόστους εκφραστεί όπως στην (4.17), το  $L_1(\beta)$  είναι συνεχές και γραμμικό. Ξεκινάει με κλίση  $-N$  για  $-\infty < \beta \leq X_1$  και κάθε φορά που ‘περνάει’ από κάποιο  $X_{(j)}$ , η κλίση αυξάνεται κατά 2.

## Ανίχνευση ασθενών σημάτων σε πρώτης τάξης εξαρτημένο θόρυβο

Για  $N$  να είναι περιττός αριθμός, υποδεικνύει ότι είναι ένας ακέραιος  $m$ , τέτοιος ώστε οι κλίσεις στα διαστήματα  $(X_{(m-1)}, X_{(m)}]$ ,  $(X_{(m)}, X_{(m+1)})$  είναι θετικές και αρνητικές αντίστοιχα. Από την (4.17) αυτές οι δύο συνθήκες ικανοποιούνται όταν ισχύουν:

$$m < \frac{N}{2} \quad \text{και} \quad m > \frac{N}{2} + 1$$

Και οι δύο αυτοί 'όροι' συναντώνται όταν  $m = \frac{N+1}{2}$ .

Για  $N$  άρτιος αριθμός, η (4.17) υπονοεί ότι υπάρχει ένας ακέραιος  $m$ , τέτοιος ώστε η κλίση στο διάστημα  $(X_{(m)}, X_{(m+1)})$  να είναι μηδενική. Η συνθήκη αυτή ικανοποιείται στην (4.17) εαν

$$-(N - 2m) = 0$$

Το οποίο είναι δυνατό για  $m = N/2$ . Έτσι η maximum likelihood εκτίμηση της θέσης για το Laplacian μοντέλο είναι ο μέσος όρος των δειγμάτων:

$$\begin{aligned} \hat{\beta}_{ML} &= \arg \min_{\beta} \sum_{i=1}^N |X_i - \beta| \\ &= \begin{cases} X_{(\frac{N+1}{2})} & N \text{ odd} \\ (X_{(\frac{N}{2})}, X_{(\frac{N}{2}+1)}) & N \text{ even} \end{cases} \\ &= \text{MEDIAN}(X_1, X_2, \dots, X_N). \end{aligned} \quad (4.18)$$

## Εκτίμηση θέσης σε Stable θόρυβο

Το να περιγραφεί με τύπους η εκτίμηση μέγιστης πιθανοφάνειας, χρειάζεται να έχουμε την γνώση μιας κλειστού τύπου συνάρτησης του μοντέλου με το οποίο δουλεύουμε (να μην περιλαμβάνει η συνάρτηση σειρές αθροίσματος, γινομένου).

Ανάμεσα στην κλάση των Stable συμμετρικών κατανομών, έχουμε τρεις περιπτώσεις όπου η συνάρτηση πυκνότητας πιθανότητας μπορεί να εκφραστεί με κλειστού τύπου συνάρτηση:

- Gaussian ( $\alpha=2$ )
- Cauchy ( $\alpha=1$ )
- Lévy ( $\alpha=0,5$ )

Από αυτές τις τρεις περιπτώσεις, η περίπτωση της Lévy κατανομής δεν είναι συμμετρική γύρω από τον μέσο όρο, γιατί η παράμετρος σκευρότητας  $\beta$ , ισούται με 1.

Έστω ένα σετ από i.i.d. δείγματα  $X_1, X_2, \dots, X_N$  τα οποία υπακούουν την Cauchy κατανομή με παράγοντα κλιμάκωσης  $K$ . Η παράμετρος θέσης (*location parameter*)  $\beta$  μπορεί να εκτιμηθεί από τα δείγματα  $X_1, X_2, \dots, X_N$  σαν η τιμή  $\tilde{\beta}_{ML}$  η οποία μεγιστοποιεί την συνάρτηση πιθανοφάνειας. Αυτό που πρέπει να κάνουμε, είναι να πάρουμε την συνάρτηση  $f(\cdot)$  η οποία περιγράφει την Cauchy κατανομή:

$$f(x) = \left(\frac{K}{\pi}\right) \frac{1}{K^2 + x^2}.$$

Η οποία με παράμετρο θέσης  $\beta$ , αυτή γίνεται:

$$f(x) = \left(\frac{K}{\pi}\right) \frac{1}{K^2 + (x - \beta)^2}$$

Με  $N$  παρατηρούμενα δείγματα, αυτή γίνεται:

$$f(X - \beta) = \left(\frac{K}{\pi}\right)^N \prod_{i=1}^N \frac{1}{K^2 + (X_i - \beta)^2}$$

Αυτό που θέλουμε να κάνουμε τώρα, είναι να βρούμε για ποια τιμή της παραμέτρου  $\beta$ , η κατανομή αυτή αποκτά την μέγιστη τιμή της (ακριβώς η ίδια διαδικασία που ακολουθήσαμε στην Gaussian κατανομή).

$$\hat{\beta}_K = \arg \max_{\beta} \prod_{i=1}^N f(X - \beta) = \arg \max_{\beta} \left(\frac{K}{\pi}\right)^N \prod_{i=1}^N \frac{1}{K^2 + (X_i - \beta)^2} \quad (4.20)$$

Πως μπορούμε να μεγιστοποιήσουμε αυτόν τον όρο; Παρατηρούμε ότι η μόνη παράμετρος που μπορεί να παραμετροποιηθεί είναι η  $\beta$ , και αυτή βρίσκεται στον παρανομαστή. Άρα ελαχιστοποιώντας τον παρανομαστή, είναι σαν να μεγιστοποιούμε τον συνολικό όρο.

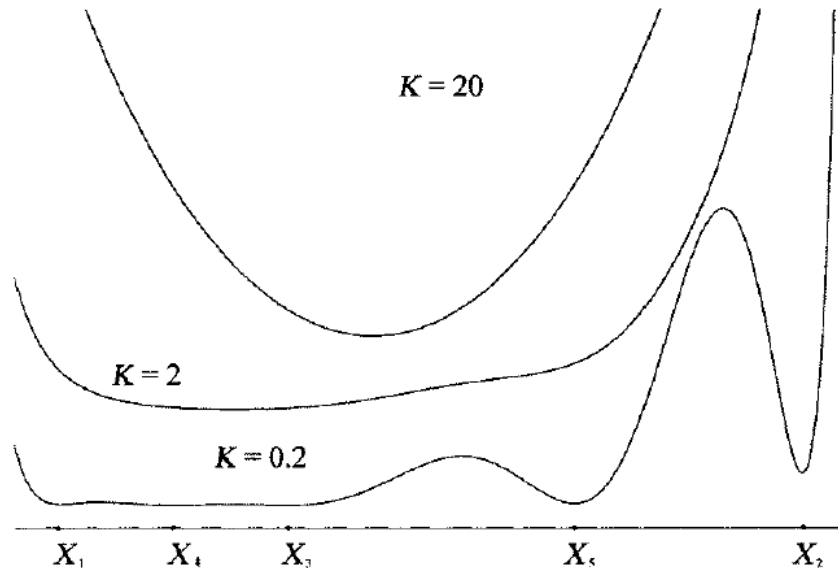
Ελαχιστοποιούμε λοιπόν τον όρο:

$$G_K(\beta) = \prod_{i=1}^N K^2 + (X_i - \beta)^2$$

Έτσι για παράμετρο κλιμάκωσης  $K > 0$ , η ML εκτίμηση θέσης είναι γνωστή σαν: *sample myriad*:

$$G_K(\beta) = \prod_{i=1}^N K^2 + (X_i - \beta)^2 = \text{MYRIAD}\{K; X_1, X_2, \dots, X_N\}$$

Προσέξτε ότι εν'αντιθέσει με τον μέσο όρο των δειγμάτων που χρησιμοποιήσαμε στην Gaussian κατανομή, για την διεξαγωγή της βέλτιστης παραμέτρου  $\beta$ , στην περίπτωση της Cauchy κατανομής η βέλτιστη τιμή της παραμέτρου  $\beta$ , μπορεί να βρεθεί αφού καθοριστεί η παράμετρος 'K'. για λόγους που θα γίνουν σε λίγο αντιληπτοί, θα αναφερόμαστε στην παράμετρο K σαν την *παράμετρο γραμμικότητας (linearity parameter)* της MYRIAD. Η συμπεριφορά της εκτιμήτριας MYRIAD είναι άμεσα συνδεδεμένη με την τιμή που έχει η παράμετρος K. Μπορείτε να δείτε το γράφημα το οποίο απεικονίζει την συμπεριφορά της εκτιμήτριας όσον αφορά το κόστος λάθους, για διάφορες τιμές της K, και για μέγεθος δείγματος ίσο με 5.



Συναρτήσεις κόστους για τα παρατηρούμενα δείγματα  $X_1 = -3, X_2 = 10, X_3 = 1, X_4 = -1$  και  $k = 20, 2, 0.2$ .

### Ελάχιστη λογαριθμική απόκλιση-Τρόπος μείωσης του κόστους

Λόγω της ιδιότητας του λογαρίθμου, είναι το ίδιο να ελαχιστοποιήσουμε τον όρο  $G_K(\beta)$  και τον όρο  $\log G_K(\beta)$ . Τότε η MYRIAD των δειγμάτων, γράφεται:

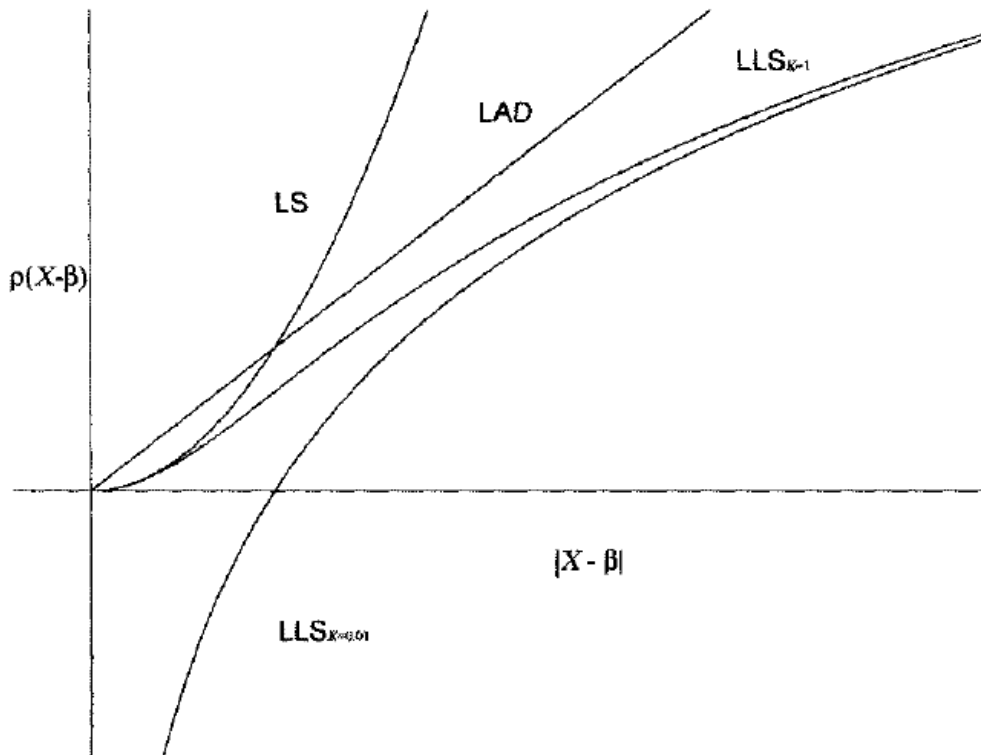
$$\text{MYRIAD}\{K; X_1, X_2, \dots, X_N\} = \arg \min_{\beta} \sum_{i=1}^N [K^2 + (X_i - \beta)^2]$$

Παρατηρήστε ότι η σειρά γινομένου που είχαμε πριν την λογαρίθμηση, μετατράπηκε σε σειρά αθροίσματος από ιδιότητα λογαρίθμων, και την περίπτωση που ένα



παρατηρούμενο δείγμα  $X_i$  το οποίο περιλαμβάνεται στο σετ δειγμάτων  $X_1, X_2, \dots, X_N$  έχει τόσο μεγάλη τιμή έτσι ώστε  $|X_i - \beta| \gg K$ , τότε το κόστος για το παρατηρούμενο δείγμα θα είναι (προσεγγιστικά)  $\log(X_i - \beta)^2$ . Ο λογάριθμος δηλαδή της τετραγωνικής απόστασης του λάθους. Το πλεονέκτημα της λογαρίθμησης είναι ότι, σε περίπτωση μεγάλου λάθους ο λογάριθμος θα το 'υποβιβάσει', άρα το ρίσκο θα είναι μικρότερο! Έτσι όπως ο μέσος όρος των δειγμάτων ελαχιστοποιεί το ποσοστό λάθους στην Gaussian, ο myriad των δειγμάτων έπειτα από εφαρμογή της λογαρίθμησης, ελαχιστοποιεί (προσεγγιστικά λόγω του ότι  $|X_i - \beta| \gg K$  όπως είπαμε) το ποσοστό λάθους. Η μέθοδος *Logarithmic Square Deviations* (λογαριθμικές τετραγωνική απόσταση) αποκαλείται και *LLS criterion*, σε αναλογία με την μέθοδο των ελαχίστων τετραγώνων (*Least Squares-LS*) και ελαχίστων απολύτων αποκλίσεων (*Least Absolute Deviations-LAD*).

Το ακόλουθο γράφημα απεικονίζει την συνάρτηση κόστους για κάθε δείγμα, καθώς αυτό αποκλίνει από την παράμετρο θέσης  $\beta$ . Το κόστος του μέσου όρου δειγμάτων (LS) έχει τετραγωνική συνάρτηση, η οποία εμφανίζει τεράστιο κόστος για μεγάλες αποκλίσεις. Ο μέσος όρος των δειγμάτων (LAD) εμφανίζει ένα κόστος το οποίο είναι γραμμικά συσχετιζόμενο με την απόκλιση. Όσον αφορά τον myriad των δειγμάτων για non-Gaussian δείγματα, εμφανίζουν ένα κόστος το οποίο είναι ανάλογο με τον λογάριθμο της απόκλισης, κάτι το οποίο οδηγεί σε μικρότερο κόστος σε περίπτωση πολύ μεγάλης απόκλισης των δειγμάτων, λόγω της φύσης των ιδιοτήτων του λογαρίθμου. Μπορούμε να πούμε έτσι ότι το myriad κόστος για δείγματα τα οποία απέχουν τόσο πολύ από την παράμετρο  $\beta$  ώστε να μπορεί να προέρχονται από διαφορετική κατανομή, 'συγχωρούνται' πολύ περισσότερο από το Gaussian μοντέλο.



Συναρτήσεις κόστους για την μέθοδο του μέσου όρου (LS), median (LAD) και της μεθόδου myriad (LLS)

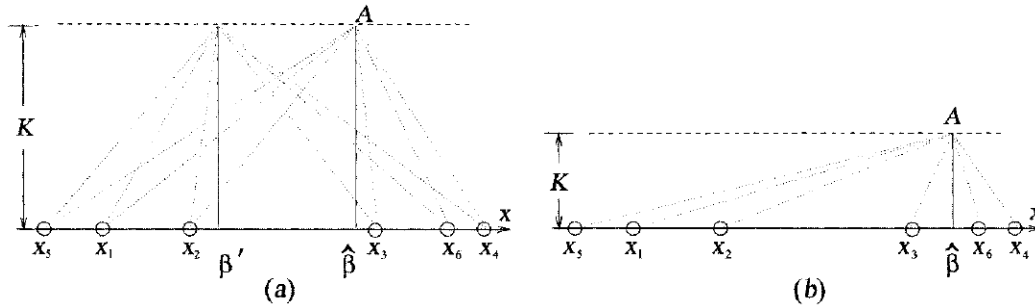
### Γεωμετρική Αναπαράσταση του κόστους

Μια δεύτερη ερμηνεία στο myriad των δειγμάτων η οποία δίνει περισσότερη πληροφορία, βασίζεται στις γεωμετρικές ιδιότητες τους. Πρώτον τα παρατηρούμενα δείγματα  $X_1, X_2, \dots, X_N$ , τοποθετούνται στην σειρά. Έπειτα μια κατακόρυφη γραμμή ‘τρέχει’ οριζόντια κατά μήκος της γραμμής των δειγμάτων και έχει μήκος ίσο με την παράμετρο κλιμάκωσης ‘K’. Με αυτές τις παραδοχές, κάθε ένας από τους όρους  $X_1, X_2, \dots, X_N$  στην (4.23) αναπαριστούν την απόσταση από το σημείο ‘A’ το οποίο είναι η ‘προβολή’ της βέλτιστης τιμής της  $\beta$  (το  $\hat{\beta}_K$ , δηλαδή η βέλτιστη τιμή του εκτιμώμενου  $\beta$ , τέτοιο ώστε να έχουμε το μικρότερο κόστος) στην οριζόντια γραμμή που απέχει K από την γραμμή των δειγμάτων. Το κόστος θα δίνεται από το πυθαγόρειο θεώρημα:

$$\text{cost} = (K^2 + (X_i - \beta)^2)$$

## Ανίχνευση ασθενών σημάτων σε πρώτης τάξης εξαρτημένο θόρυβο

Βλέπουμε ότι εάν η παράμετρος κλιμάκωσης  $K$  μειωθεί, τότε το έκαστο κόστος αλλάζει (πολύ μικρό για τιμές  $X_i$  κοντά στην  $\hat{\beta}_K$  και πολύ μεγάλο για τιμές απομακρυσμένες από την  $\hat{\beta}_K$ ).



Παρατηρήστε ότι ο *myriad* των δειγμάτων,  $\hat{\beta}$ , των δειγμάτων ελαχιστοποιεί την συνολική απόσταση των δειγμάτων  $X_i$ , δεδομένου της παραμέτρου γραμμικότητας  $K$ . Κάθε διαφορετική τιμή του  $\beta$ , η συνολική απόσταση των δειγμάτων από το 'A' είναι μεγαλύτερη, και άρα το κόστος λάθους είναι μεγαλύτερο.

# ΕΝΟΤΗΤΑ ΙΙΙ

*Καλύπτονται έννοιες:*

- Παρατηρούμενος Χώρος
- Εξαρτημένος θόρυβος-Αναπαράστασή του με σειρά Volterra
- Συσχέτιση δύο σημάτων
- Μετασχηματισμός Παρατηρούμενου χώρου με την βοήθεια του Ιακωβιανού Μετασχηματισμού-*Jacobian Transformation*
- Παρουσίαση Gaussian, Cauchy και Student-t κατανομών
- Εύρεση του καταλλήλου κριτηρίου ανάλογα με την κατανομή που ακολουθεί ο θόρυβος
- Παρουσίαση Antipodal συστήματος εκπομπής
- Αριθμητικά αποτελέσματα και συμπεράσματα

## Παρατηρούμενος Χώρος

Υποθέστε ότι έχουμε στην διάθεσή μας ένα σετ από  $\{s_k(t)\}_{k=1}^n$  δείγματα, τα οποία χρησιμοποιούνται για να αντλήσουμε την απαραίτητη πληροφορία, η οποία θα μας οδηγήσει στην απόφαση του σήματος που αποστάλθηκε στο χρονικό διάστημα  $T_s$ . Το πρόβλημα είναι, ότι σε κάθε δείγμα λαμβάνω μαζί και θόρυβο και έτσι η απόφαση γίνεται πιο δύσκολη, και η απόδοση του δέκτη μειώνεται ανάλογα με την μορφή και την κρουστικότητα του θορύβου. Πιο αναλυτικά, το σήμα που λαμβάνω στο χρονικό διάστημα  $T_s$  είναι:

$$x(t) = s_i(t) + w(t) \quad i=1,2,\dots,M \quad \text{στο διάστημα} \quad 0 \leq t < T_s \quad \text{για M-ary μετάδοση (1)}$$

Όπου  $w(t)$  είναι ο προσθετικός θόρυβος ο οποίος μπορεί να πάρει διάφορες τιμές ανάλογα με την κατανομή πιθανότητας που ακολουθεί. Επίσης υπάρχει η πιθανότητα να εμφανίζει μνήμη ο θόρυβος. Αυτό σημαίνει ότι το παρών δείγμα θορύβου εξαρτάται από τα προηγούμενα δείγματα.

Ο στόχος είναι λοιπόν, να βγάλουμε το κριτήριο που θα οδηγήσει στην μικρότερη πιθανότητα λάθους, με την βοήθεια  $n$  δειγμάτων που θα λάβω από την λαμβανόμενη κυματομορφή που έχω στην διάθεσή μου. Για να πάρουμε  $n$  δείγματα από το σήμα, θα πρέπει να χωρίσουμε τον χρόνο  $T_s$  σε  $t_\Delta$  διαστήματα και να εκτελέσω δειγματοληψία στο μέσο του κάθε διαστήματος. Προκύπτει λοιπόν ότι θα έχω:

$$T_s = (n-1)t_\Delta$$

Και το μοντέλο παρατήρησης για το διάστημα που διαρκεί η κυματομορφή θα είναι:

$$X_i = s_{k,i} + W_i \quad \text{για} \quad i = 1, 2, \dots, n \quad (2)$$

- Όπου  $\{X_i\}_{i=1}^n$  είναι τα παρατηρούμενα δείγματα στο διάστημα
- Όπου  $\{s_{k,i}(t)\}_{i=1}^n$  είναι το σήμα που εκπέμφθηκε (χωρίς τον θόρυβο)
- Όπου  $\{W_i\}_{i=1}^n$  είναι τα δείγματα του θορύβου σε κάθε δειγματοληψία.
- Ο τελεστής ' $k$ ' που βλέπετε, είναι το κ-οστό σύμβολο μέσα από τα  $M$  πιθανά σύμβολα. Παρακάτω, θα προσέξετε ότι ο τελεστής  $k$  αντικαθιστάται από τον τελεστή  $m$ . Το σύμβολο αυτό, αντιπροσωπεύει όλα τα άλλα σύμβολα εκτός το σύμβολο της προσοχής μας, το  $k$ . Προσπαθούμε δηλαδή να βρούμε το βέλτιστο κριτήριο για το  $k$  σύμβολο.

Υποθέτουμε ότι στο διάστημα  $T_s$  τα δείγματα σήματος (χωρίς θόρυβο) είναι το διάνυσμα  $\underline{s}_k = s_{k,1}, s_{k,2}, \dots, s_{k,n}$  που αυτό με την σειρά του είναι ίσο με:

$$\underline{s}_k = \mathcal{G} \varepsilon_k \widetilde{s}_k$$

- Όπου  $\mathcal{G}$  είναι η ισχύς με την οποία εκπέμφθηκε το σήμα η οποία είναι σταθερά και ίδια ακόμα και για το σήμα '0' που περιμένουμε να είναι σήμα με μικρότερη ισχύ από το '1'. Πως ορίζουμε το πλάτος που θέλουμε να έχει το '1' και το '0', δίνεται από την παράμετρο  $\varepsilon_k$ .
- Όπου  $\varepsilon_k$  είναι μια σταθερά  $\varepsilon_k \geq 0$  η οποία είναι η απόσταση που θέλουμε να έχει η ισχύς του σήματός μας από το μηδέν. Αν για παράδειγμα θέλουμε να εκπέμψουμε σήμα το οποίο παίρνει δύο καταστάσεις, (έστω το διάνυσμα  $\underline{s}_k$  να συμβολίζει το σήμα '1' και το διάνυσμα  $\underline{s}_m$  να συμβολίζει το σήμα '0'), τότε οι δύο πιθανές καταστάσεις που μπορεί να σταλθούν, είναι:

$$\underline{s} = \begin{cases} \mathcal{G} \varepsilon_k \widetilde{s}_k & \text{για την αποστολή του } s_k \\ \mathcal{G} \varepsilon_m \widetilde{s}_m & \text{για την αποστολή του } s_m. \text{ Όπου } m=1,2,\dots,M \text{ και } m \neq k \end{cases}$$

Πόσο όμως θα είναι το πλάτος του διανύσματος  $\underline{s}_m$ ; Αν το πλάτος του  $\underline{s}_k$  είναι το μέγιστο, τότε θα μπορούσαμε να θέσουμε  $\varepsilon_k = 1$ . Για να έχει το διάνυσμα  $\underline{s}_m$  το μισό πλάτος του διανύσματος  $\underline{s}_k$ , τότε θέτουμε  $\varepsilon_m = 0.5$ . Έτσι μειώνουμε την ισχύ του σήματος που συμβολίζει το '0' στο μισό.

- Ο τελευταίος όρος είναι όρος κανονικοποίησης. Πιο συγκεκριμένα:

$$\widetilde{s}_k = \frac{s_k}{\|\underline{s}_k\|_2}$$

Ο παρανομαστής δηλώνει την Ευκλείδεια κανονικοποίηση, η οποία μας δίνει την απόσταση που έχουν τα δείγματα του διανύσματος  $\underline{s}_k$  (ή και  $\underline{s}_m$ ) αναμεταξύ τους.

$$\begin{aligned} \widetilde{s}_k &= \frac{s_k}{\|\underline{s}_k\|_2} = \frac{s_k}{\sqrt{(s_{k,1})^2 + (s_{k,2})^2 + \dots + (s_{k,n})^2}} = \frac{s_k}{\sqrt{n \cdot \underline{s}_k^2}} = \frac{s_k}{\underline{s}_k \sqrt{n}} = \frac{1}{\sqrt{n}} \\ \widetilde{s}_m &= \frac{s_m}{\|\underline{s}_m\|_2} = \frac{s_m}{\sqrt{(s_{m,1})^2 + (s_{m,2})^2 + \dots + (s_{m,n})^2}} = \frac{s_m}{\sqrt{n \cdot \underline{s}_m^2}} = \frac{s_m}{\underline{s}_m \sqrt{n}} = \frac{1}{\sqrt{n}} \end{aligned}$$

*Άρα η σταθερά ευκλείδειας κανονικοποίησης είναι και για τα δύο διανύσματα ίδια.*

## Υπόθεση με Εξαρτημένο Θόρυβο

Η ανίχνευση M-αδικών σημάτων με εξαρτημένο θόρυβο, μπορεί να μοντελοποιηθεί με ένα σύστημα M-δυνατών υποθέσεων όπου η κάθε υπόθεση μοντελοποιείται ως εξής:

$$H_k : \underline{X} = \mathcal{G}\varepsilon_k \widetilde{s}_k + \underline{W}, \quad k=1,2,\dots,M$$

Όσα είναι τα δυνατά σήματα που μπορεί να στείλει ο πομπός, τόσες θα είναι και οι υποθέσεις που μπορώ να κάνω. Για παράδειγμα, δε ένα δυαδικό σύστημα στέλνω ή το '0', ή το '1'. Άρα θα έχω δύο δυνατές υποθέσεις, την  $H_0$  και την  $H_1$ . Ο λόγος που βάζουμε την μεταβλητή  $X$  σε μορφή διανύσματος (μονοδιάστατος πίνακας), είναι λόγω του ότι για την συγκεκριμένη υπόθεση (έστω την  $H_k$ ) θα πρέπει να εκτελέσουμε 'n' δειγματοληψίες για να καταλάβουμε πιο σήμα στάλθηκε. Ομοίως, κάθε φορά που εκτελώ δειγματοληψία, εκτός από το σήμα, θα λάβω και θόρυβο, ο οποίος μπορεί να είναι εξαρτημένος από τα προηγούμενα δείγματα θορύβου, ή και όχι. Όπως και να έχει, θα πρέπει να καταχωρηθεί και αυτός με μορφή διανύσματος (πίνακας).

Όταν τα δείγματα θορύβου είναι εξαρτημένα το ένα με το άλλο, τότε η τακτική που μπορούμε να ακολουθήσουμε είναι να μοντελοποιήσουμε την διαδικασία του θορύβου με την ανάπτυξη *Volterra*. Πράγματι, αν θεωρήσουμε τον θόρυβο σαν ένα πολυώνυμο n-βαθμού, τότε η ανάπτυξη *Volterra* θα έχει n 'φωλιασμένα' αθροίσματα:

$$W_i = \sum_l h_l \Lambda_{i-l} + \sum_l \sum_m h_{l,m} \Lambda_{i-l} \Lambda_{i-m} + \sum_l \sum_m \sum_n h_{l,m,n} \Lambda_{i-l} \Lambda_{i-m} \Lambda_{i-n}$$

Όπου  $\{h_{l,m,n,\dots}\}$  είναι η σταθερές *Volterra (Volterra Kernels)* και  $\{\Lambda_i\}_{i=1}^n$  είναι ανεξάρτητες τυχαίες μεταβλητές. Το πρόβλημα με την μοντελοποίηση *Volterra*, είναι ότι είναι πολύ δύσκολο και χρονοβόρο να γίνει η εκτίμηση των *Volterra Kernels*. Αντί λοιπόν της μοντελοποίησης όπου χρειάζεται να λάβω υπόψη ένα τεράστιο αριθμό *Volterra Kernels*, καταφεύγω σε μια πιο απλή μοντελοποίηση της σειράς *Volterra*, αυτής όπου θα πάρω μόνο τον πρώτο όρο της σειράς κινητού μέσου όρου:

$$W_t = \mu + W_t + \rho_1 W_{t-1} + \dots + \rho_q W_{t-q}$$

- Όπου 'μ' είναι ο μέσος όρος του θορύβου, ο οποίος έχει μηδενική τιμή
- Όπου  $\rho_1$  είναι ο βαθμός εξάρτησης από το αμέσως προηγούμενο δείγμα θορύβου. Στην δική μας περίπτωση, μόνο ο όρος  $\rho_1$  δεν έχει μηδενική τιμή. Οι όροι ανώτερης τάξης  $\rho_2, \rho_3, \dots, \rho_n$  έχουν τόσο μικρή τιμή, η οποία λαμβάνεται ίση με το μηδέν.

Αυτό μπορώ να το κάνω διότι η εξάρτηση των δειγμάτων του θορύβου στην δική μας ανάλυση, είναι αδύναμη. Αν η εξάρτηση που εμφάνιζαν τα δείγματα θορύβου ήταν πιο ‘βαριά’, τότε θα ήμουν υποχρεωμένος να πάρω και άλλους όρους της ανάπτυξης Volterra.

Με αυτήν την προσέγγιση λοιπόν, το κάθε δείγμα θορύβου θα είναι ίσο με:

$$W_i = \Lambda_i + \rho_d \cdot \Lambda_{i-1}$$

Όπως πολύ σωστά φαντάζεστε, η παράμετρος  $\rho_d$  δηλώνει εξάρτηση. Το πόσο εξαρτημένο είναι δηλαδή το παρών δείγμα θορύβου  $\Lambda_i$ , από το προηγούμενο δείγμα θορύβου,  $\Lambda_{i-1}$  (φυσικά εάν η παράμετρος εξάρτησης πάρει την τιμή μηδέν, τότε μιλάμε για ανεξάρτητο θόρυβο).

Μιλάμε όμως για θόρυβο και αντί να τον συμβολίζουμε με το ‘ $W$ ’, βλέπουμε ότι τον συμβολίζουμε με το Ελληνικό λάμδα ‘ $\Lambda$ ’. Αυτό διότι το  $W$  δηλώνει μεν θόρυβο, αλλά θόρυβος ο οποίος είναι ανεξάρτητος από τα προηγούμενα δείγματα θορύβου. Αν λάβουμε υπόψη την εξάρτηση, τότε δεν μπορώ πλέον να χρησιμοποιώ το ίδιο γράμμα για τον συμβολισμό του θορύβου και αυτό που θα χρησιμοποιήσω είναι το  $\Lambda$  όπου δηλώνει εξαρτημένο θόρυβο.

- Προσέξτε ότι το πρώτο δείγμα θορύβου θα έχει τιμή

$$W_1 = \Lambda_1.$$

- Το δεύτερο δείγμα θορύβου θα έχει τιμή

$$W_2 = \Lambda_2 - \rho_d \Lambda_1$$

- Το τρίτο δείγμα θορύβου θα έχει τιμή

$$W_3 = \Lambda_3 - \rho_d \Lambda_2.$$

Όμως μην ξεχνάτε ότι το  $\Lambda_2$  είναι και αυτό εξαρτημένο από το  $\Lambda_1$ . Άρα:

$$W_3 = \Lambda_3 - \rho_d \Lambda_2 = \Lambda_3 - \rho_d (\Lambda_2 - \rho_d \cdot \Lambda_1) = \Lambda_3 - \rho_d \Lambda_2 + \rho_d^2 \Lambda_1$$

- Το τέταρτο δείγμα θορύβου θα έχει τιμή:



$$W_4 = \Lambda_4 - \rho_d \Lambda_3 = \Lambda_4 - \rho_d (\Lambda_3 - \rho_d \cdot \Lambda_2) = \Lambda_4 - \rho_d \Lambda_3 + \rho_d^2 \Lambda_2 = \Lambda_4 - \rho_d \Lambda_3 + \rho_d^2 (\Lambda_2 - \rho_d \cdot \Lambda_1) \Rightarrow$$
$$W_4 = \Lambda_4 - \rho_d \Lambda_3 + \rho_d^2 \Lambda_2 - \rho_d^3 \Lambda_1$$

Με τον τρόπο αυτό καταφέραμε εκτός από την αναπαράσταση του κάθε δείγματος θορύβου και την μετατροπή του εξαρτημένου θορύβου σε ανεξάρτητο με μηδενικό μέσο όρο. Ο λόγος που τα δείγματα θα έχουν μηδενικό μέσο όρο, απορρέει από το Κεντρικό Οριακό Θεώρημα, το οποίο υποδεικνύει ότι το κανονικοποιημένο άθροισμα 'n' τυχαίων μεταβλητών κάθε μια από τις οποίες έχει μηδενικό μέσο όρο, θα έχει μηδενική τιμή.

Ποιο θα είναι όμως το ποσοστό εξάρτησης του παρόντος δείγματος θορύβου από το προηγούμενο δείγμα; Αυτό το ερώτημα μπορεί να απαντηθεί εάν χρησιμοποιήσουμε τον όρο συσχέτισης που γνωρίζουμε από την θεωρία της στατιστικής.

## Συσχέτιση

Η φασματική συσχέτιση είναι ένας στατιστικός όρος ο οποίος μπορεί να χρησιμοποιηθεί για να εξετάσουμε τη σχέση ανάμεσα σε δύο σήματα, ή σετ δεδομένων. Χρησιμοποιείται συχνά για να εκτιμήσουμε την μεταφορά της ισχύος σε ένα γραμμικό σύστημα, από την είσοδο του, στην έξοδο του. Εάν τα σήματα είναι *εργοδικά* και η συνάρτηση μεταφοράς του συστήματος είναι *γραμμική*, τότε μπορούμε να χρησιμοποιήσουμε την συσχέτιση για να εκτιμήσουμε την αιτιότητα μεταξύ της εισόδου και εξόδου.

Η συσχέτιση μεταξύ δύο σημάτων  $x(t)$  και  $y(t)$  δίνεται από:

$$C_{ij} = \frac{|G_{ij}|^2}{G_{ii} G_{jj}}$$

ή αλλιώς

$$r_{ij} = \frac{E\{W_i W_j\}}{\sqrt{E\{W_i^2\} E\{W_j^2\}}}$$

Όπου  $G_{ij} = \sqrt{E\{W_i W_j\}}$  είναι η 'τομή' της φασματικής πυκνότητας ισχύος μεταξύ στα δύο δείγματα θορύβου  $W_i$  και  $W_j$ , και οι όροι  $G_{ii}$  όπως και  $G_{jj}$  δίνουν την αυτοσυσχέτιση των δειγμάτων θορύβου  $W_i$  και  $W_j$ .

Όπως πολύ πιθανόν φαντάζεστε, η συσχέτιση των δύο δειγμάτων θα κυμαίνεται από την μηδενική τιμή όπου τα δύο δείγματα θορύβου δεν συσχετίζονται, μέχρι και την τιμή ένα, όπου ουσιαστικά το ένα δείγμα θορύβου 'επικάθεται' πλήρως πάνω στο άλλο. Για παράδειγμα, εάν μελετούσαμε την συσχέτιση είσοδος-έξοδος σε ένα σύστημα όπου η συνάρτηση μεταφοράς του είναι γραμμική (ίση με μια σταθερά για παράδειγμα), τότε η συσχέτιση εισόδου εξόδου θα είναι ίση με ένα λόγω του ότι η συνάρτηση μεταφοράς του συστήματος δεν θα προκαλέσει αλλαγή φάσης ή μετατόπιση συχνότητας.

### Μετασχηματισμός παρατηρούμενου χώρου

Είπαμε ότι για να βγάλουμε μια ασφαλέστερη απόφαση (μικρότερο ρίσκο) σχετικά με το σήμα που έχω στην κατοχή μου, λαμβάνω  $n$  δείγματα από το σήμα και μαζί με αυτά τα σήματα δειγματοληπτώ ταυτόχρονα και θόρυβο. Ποια θα είναι η τιμή του θορύβου εάν είχα ένα δείγμα στην διάθεσή μου;

Για να απαντήσουμε αυτό το ερώτημα θα πρέπει πρώτα από όλα να γνωρίζουμε με ποιο μοντέλο θορύβου δουλεύουμε. Με άλλα λόγια, ο θόρυβος που επικάθισε στο σήμα μου, τι κατανομή εμφανίζει; Αυτό είναι ίσως το σημαντικότερο ερώτημα που θα πρέπει να απαντηθεί, προκειμένου να κατασκευαστεί ένας αποτελεσματικός δέκτης.

Για το παρατηρούμενο μοντέλο (5) με δείγματα θορύβου  $\{W_i\}_{i=1}^n$  τα οποία περιγράφονται από την (7), η joint pdf  $\phi_{\underline{x}}(\underline{x} \| H_k, \mathcal{G})$  της παρατήρησης 'X' είναι:

- Υπενθυμίζουμε ότι η joint pdf είναι μια κατανομή η οποία εξαρτάται περισσότερης της μίας παραμέτρου. Στην δική μας περίπτωση η παρατηρούμενη τιμή εξαρτάται και από την υπόθεση ( $H_k$  ή  $H_m$ ) αλλά και από την κοινή παράμετρο της ισχύος του σήματος ( $\mathcal{G}$ ).

(σχήμα joint pdf)

Για να γίνει περισσότερο κατανοητό, θυμηθείτε ότι το 'ατόφιο' σήμα το οποίο αποστέλλεται, είναι ίσο με  $\mathcal{G} \varepsilon_k \widetilde{s_{k,i}}$ . Επίσης η παράμετρος  $\widetilde{s_{k,i}}$  είναι σταθερή και για την υπόθεση  $H_k$  αλλά και για την υπόθεση  $H_m$ . Για να γίνουν περισσότερο αντιληπτά αυτά που πρόκειται να ακολουθήσουν, προσέξτε το ακόλουθο παράδειγμα:

## Παράδειγμα

Έστω πως θέλουμε να στείλουμε σήμα το οποίο αποτελείται από μια αλληλουχία ‘0’ και ‘1’. Το ‘0’ θα αποστέλλεται με μηδενική ισχύ, και το ‘1’ θα αποστέλλεται με ισχύ  $\mathcal{E}_k \widetilde{s}_{1,i}$ . Αυτό που θέλουμε είναι να βρούμε μια έκφραση για την συνάρτηση που περιγράφει το σήμα ‘0’ αλλά και ‘1’ εκτελώντας δειγματοληψία στο λαμβανόμενο σήμα, ενός και  $n$  δειγμάτων.

Έστω πως το σήμα το οποίο αποστέλλεται να είναι ίσο με:

$$\text{signal} = \begin{cases} \mathcal{E}_1 \widetilde{s}_{1,i} & \varepsilon_1 = 1 \\ \mathcal{E}_0 \widetilde{s}_{0,i} & \varepsilon_0 = 0 \end{cases}$$

Έχουμε την αποστολή ενός on-off σήματος δηλαδή.

Κατά την αποστολή του σήματος, προστίθεται θόρυβος ο οποίος έστω πως ακολουθεί Gaussian κατανομή. Για να αποφανθούμε λοιπόν στάλθηκε το ‘0’ ή το ‘1’, θα πρέπει να εκτελέσουμε δειγματοληψία στο λαμβανόμενο σήμα.

Εάν πάρω ένα δείγμα από το σήμα, τότε η έκαστο υπόθεση για το ‘0’ και για το ‘1’ θα είναι:

$$\begin{aligned} H_1 : X &= \mathcal{E}_1 \widetilde{s}_{1,i} + \text{Noise} \\ H_0 : X &= \text{Noise} \end{aligned}$$

- Για την υπόθεση  $H_0$  θα έχω:

$$f_{X|H_0, \mathcal{E}}(X | H_0, \mathcal{E}) = \frac{1}{\sqrt{2\pi\sigma}} \exp\left(-\frac{(x - \mathcal{E}_0 \widetilde{s}_{0,i})^2}{2\sigma^2}\right)$$

Το σήμα όμως ‘0’ αποστέλλεται με μηδενική ισχύ. Άρα ουσιαστικά έχω:

$$f_{X|H_0, \mathcal{E}}(X | H_0, \mathcal{E}) = f_{\text{noise}}(N) = \frac{1}{\sqrt{2\pi\sigma}} \exp\left(-\frac{(\text{noise sample})^2}{2\sigma^2}\right)$$

- Και για την υπόθεση  $H_1$  θα έχω:

$$f_{X|H_1, \mathcal{E}}(X | H_1, \mathcal{E}) = \frac{1}{\sqrt{2\pi\sigma}} \exp\left(-\frac{(x - \mathcal{E}_1 \widetilde{s}_{1,i})^2}{2\sigma^2}\right)$$

Προσέξτε ότι η υπόθεση  $H_1$  έχει τον όρο  $(x - \vartheta \varepsilon_0 \widetilde{s_{0,i}})$ . Τι θα προκύψει εάν αφαιρέσουμε από το λαμβανόμενο σήμα το σήμα που εκπέμφθηκε; Το δείγμα θορύβου! Για αυτόν τον λόγο και για την συνάρτηση  $f_{X|H_0, \vartheta}(X | H_0, \vartheta)$  αλλά και για την συνάρτηση  $f_{X|H_1, \vartheta}(X | H_1, \vartheta)$  μπορώ να τις γράψω σαν συναρτήσεις θορύβου:

$$f_{X|H_1, \vartheta}(X | H_1, \vartheta) = f_{w_i}(X - \vartheta \varepsilon_1 \widetilde{s_{1,i}}) = \frac{1}{\sqrt{2\pi\sigma}} \exp\left(-\frac{(X - \vartheta \varepsilon_1 \widetilde{s_{1,i}})^2}{2\sigma^2}\right)$$

$$f_{X|H_0, \vartheta}(X | H_0, \vartheta) = f_{w_i}(X - \vartheta \varepsilon_0 \widetilde{s_{0,i}}) = \frac{1}{\sqrt{2\pi\sigma}} \exp\left(-\frac{(X - \vartheta \varepsilon_0 \widetilde{s_{0,i}})^2}{2\sigma^2}\right)$$

Για το σήμα ‘0’ και ‘1’ αντίστοιχα.

Ας το προχωρήσουμε ένα βήμα παρακάτω τώρα. Γνωρίζεται πως εάν βασιστούμε σε περισσότερα του ενός δείγματα για να αποφανθούμε εάν στάλθηκε το ‘0’ ή το ‘1’, τότε η πιθανότητα να κάνουμε λάθος εκτίμηση, μειώνεται. Εάν πάρουμε  $n$  δείγματα από το λαμβανόμενο σήμα, τότε έχουμε τις  $f_{X|H_0, \vartheta}(X | H_0, \vartheta)$  αλλά την συνάρτηση  $f_{X|H_1, \vartheta}(X | H_1, \vartheta)$  να γίνονται:

$$f_{X|H_0, \vartheta}(X | H_0, \vartheta) = f_{\underline{w}}(X - \vartheta \varepsilon_0 \widetilde{s_{0,1}}) = \prod_{i=1}^n f_{w_i}(X - \vartheta \varepsilon_0 \widetilde{s_{0,i}}) = \prod_{i=1}^n \frac{1}{\sqrt{2\pi\sigma}} \exp\left(-\frac{(X_i - \vartheta \varepsilon_0 \widetilde{s_{0,i}})^2}{2\sigma^2}\right)$$

$$f_{X|H_1, \vartheta}(X | H_1, \vartheta) = f_{\underline{w}}(X - \vartheta \varepsilon_1 \widetilde{s_{1,1}}) = \prod_{i=1}^n f_{w_i}(X - \vartheta \varepsilon_1 \widetilde{s_{1,i}}) = \prod_{i=1}^n \frac{1}{\sqrt{2\pi\sigma}} \exp\left(-\frac{(X_i - \vartheta \varepsilon_1 \widetilde{s_{1,i}})^2}{2\sigma^2}\right)$$

Από εδώ και στο εξής λοιπόν ονομάζουμε τις συναρτήσεις αυτές, συναρτήσεις δειγμάτων θορύβου.

Είδαμε λοιπόν την έκφραση του παρατηρούμενου χώρου για δείγματα θορύβου τα οποία είναι ανεξάρτητα μεταξύ τους. Τώρα θα είναι πιο εύκολο να δούμε τι μορφή θα έχει ο παρατηρούμενος χώρος δειγμάτων, εάν υπάρχει εξάρτηση μεταξύ των δειγμάτων θορύβου, γιατί όπως θα δείτε θα ‘μετατρέψουμε’ τα εξαρτημένα δείγματα σε ανεξάρτητα.

- Εάν λοιπόν το πρώτο δείγμα θορύβου είναι ίσο με,

$$\Lambda_1 = (X_1 - \vartheta \varepsilon_1 \widetilde{s_{1,1}})$$

- τότε το δεύτερο δείγμα θορύβου είναι ίσο με

$$\Lambda_2 = \left( X_2 - \mathcal{G}\varepsilon_1 \widetilde{s_{1,2}} \right) - \rho_d \left( X_1 - \mathcal{G}\varepsilon_1 \widetilde{s_{1,1}} \right)$$

- Με τον ίδιο τρόπο, το τρίτο δείγμα θορύβου είναι ίσο με

$$\Lambda_3 = \left( X_3 - \mathcal{G}\varepsilon_1 \widetilde{s_{1,3}} \right) - \rho_d \left( X_2 - \mathcal{G}\varepsilon_1 \widetilde{s_{1,2}} \right)$$

Όμως το τρίτο δείγμα θορύβου, περιλαμβάνει το δεύτερο δείγμα θορύβου, το οποίο με την σειρά του περιλαμβάνει το πρώτο δείγμα θορύβου. Με αντικατάσταση λοιπόν έχουμε:

$$\Lambda_3 = \left( X_3 - \mathcal{G}\varepsilon_1 \widetilde{s_{1,3}} \right) - \rho_d \left( X_2 - \mathcal{G}\varepsilon_1 \widetilde{s_{1,2}} \right) + \rho_d^2 \left( X_1 - \mathcal{G}\varepsilon_1 \widetilde{s_{1,1}} \right)$$

- Το n-οστό δείγμα θορύβου θα έχει τιμή τότε

$$\Lambda_n = \left( X_n - \mathcal{G}\varepsilon_1 \widetilde{s_{1,n}} \right) - \rho_d \left( X_{n-1} - \mathcal{G}\varepsilon_1 \widetilde{s_{1,n-1}} \right) + \dots + \left( \rho_d^{n-1} \right) \left( X_1 - \mathcal{G}\varepsilon_1 \widetilde{s_{1,1}} \right)$$

Το επόμενο βήμα όπως είδαμε στο προηγούμενο παράδειγμα, είναι να ενσωματώσουμε στην  $f_{X|H_1, \mathcal{G}}(X | H_1, \mathcal{G})$  τα επιμέρους δείγματα θορύβου, να πάρουμε στην συνέχεια μια έκφραση για την συνάρτηση αυτή εάν πραγματοποιήσουμε δειγματοληψία. Υπάρχει ένα πρόβλημα όμως: είδαμε πως τα ανεξάρτητα δείγματα θορύβου τα συμβολίζαμε το το γράμμα 'W'. Τώρα όμως θα πρέπει να αλλάξουμε τον συμβολισμό σε 'Λ' που δηλώνει εξαρτημένα δείγματα θορύβου. Όπως πιθανό θα αντιλαμβάνεστε, δεν μπορούμε να εξακολουθούμε να χρησιμοποιούμε το γράμμα 'X' για τα παρατηρούμενα δείγματα σήματος (σήμα και θόρυβος μαζί), διότι αυτή είναι η περίπτωση ανεξάρτητων δειγμάτων θορύβου. Θα πρέπει λοιπόν μετατρέψουμε το διάνυσμα παρατηρούμενων δειγμάτων X σε ένα άλλο διάνυσμα (έστω Y) του οποίου τα δείγματα είναι εξαρτημένα το ένα από το άλλο. Αυτή η διαδικασία είναι γνωστή σαν την ιακωβιανή μετατροπή (*Jacobian Transformation*).

## Jacobian Transformation-Ιακωβιανός Μετασχηματισμός

Ο καινούργιος πίνακας θα δίνεται από τον ιακωβιανό πίνακα:

$J = \begin{bmatrix} \frac{\partial y_1}{\partial x_1} & \dots & \frac{\partial y_1}{\partial x_n} \\ \vdots & \ddots & \vdots \\ \frac{\partial y_n}{\partial x_1} & \dots & \frac{\partial y_n}{\partial x_n} \end{bmatrix}$	Εάν αρχικά έχουμε πίνακα $x = [x_1, x_2, \dots, x_n]$ και θέλουμε να τον μετατρέψουμε σε έναν άλλο $Y = [Y_1, Y_2, \dots, Y_m]$ , τότε θα προκύψει πίνακας n-στηλών και m-γραμμών, όπου κάθε γραμμή για παράδειγμα θα έχει μήκος-n και τα στοιχεία της θα είναι η μερική παράγωγος του αντίστοιχου Y-στοιχείου ως προς όλους τους συνδυασμούς που θα μπορεί να πάρει ο X πίνακας.
$J_F(x_1, \dots, x_n)$	Μια διαφορετική έκφραση του ως-προς μετατροπή πίνακα X
$J = \frac{\partial(y_1, \dots, y_n)}{\partial(x_1, \dots, x_n)}$	‘Συντομογραφία’ του jacobian πίνακα
$g_i(X), i = 1, 2, \dots, n$	Συνάρτηση με την οποία αντιστοιχούμε κάθε στοιχείο του διανύσματος $\underline{X} \rightarrow \underline{Y}$

Ας δούμε λοιπόν την μετετροπή του διανύσματος  $\underline{X} \rightarrow \underline{Y}$

Για n δείγματα θα έχουμε το εξής:

$$J[x_1, x_2, \dots, x_n] = \frac{\partial(f_{x_1}, f_{x_2}, \dots, f_{x_n})}{\partial(x_1, \dots, x_n)} = \begin{bmatrix} \frac{\partial f(x_1)}{\partial x_1} & \dots & \frac{\partial f(x_1)}{\partial x_n} \\ \vdots & \ddots & \vdots \\ \frac{\partial f(x_n)}{\partial x_1} & \dots & \frac{\partial f(x_n)}{\partial x_n} \end{bmatrix}$$

$$g_i(X) = Y_i = \sum_{j=0}^{i-1} (-\rho_d)^j X_{i-j}$$

$$g_1(X) = Y_1 = \sum_{j=0}^{1-1} (-\rho_d)^j X_{1-j} \Rightarrow Y_1 = X_1$$

$$g_2(X) = Y_2 = \sum_{j=0}^{2-1} (-\rho_d)^j X_{2-j} \Rightarrow Y_2 = X_2 - \rho_d X_1$$

$$g_3(X) = Y_3 = \sum_{j=0}^{3-1} (-\rho_d)^j X_{3-j} \Rightarrow Y_3 = X_3 - \rho_d X_2 + \rho_d^2 X_1$$

•  
•  
•

$$g_n(X) = Y_n = \sum_{j=0}^{n-1} (-\rho_d)^j X_{n-j} \Rightarrow Y_n = X_n - \rho_d X_{n-1} + \dots + (-\rho_d)^{n-1} X_1$$

$$J[x_1, x_2, \dots, x_n] = \det \begin{bmatrix} \frac{\partial g_1}{\partial x_1} & \dots & \frac{\partial g_1}{\partial x_n} \\ \vdots & \ddots & \vdots \\ \frac{\partial g_n}{\partial x_1} & \dots & \frac{\partial g_n}{\partial x_n} \end{bmatrix} = \begin{bmatrix} 1 & 0 & 0 & 0 & 0 \\ -\rho_d & 1 & 0 & 0 & 0 \\ \rho_d^2 & -\rho_d & 1 & 0 & 0 \\ \vdots & \vdots & \vdots & \ddots & 0 \\ (-\rho_d)^{n-1} & (-\rho_d)^{n-2} & (-\rho_d)^{n-3} & \dots & 1 \end{bmatrix} = 1$$

Παραγωγίζοντας μερικώς ως προς τους αντίστοιχους τελεστές, έχουμε:

$$\frac{\partial g_1}{\partial x_1} = \frac{\partial(x_1)}{\partial x_1} = 1$$

$$\frac{\partial g_2}{\partial x_1} = \frac{\partial(X_2 - \rho_d X_1)}{\partial x_1} = -\rho_d$$

$$\frac{\partial g_3}{\partial x_1} = \frac{\partial(X_3 - \rho_d X_2 + \rho_d^2 X_1)}{\partial x_1} = \rho_d^2$$

$$\frac{\partial g_n}{\partial x_1} = \frac{\partial(X_n - \rho_d X_{n-1} + \dots + (-\rho_d)^{n-1} X_1)}{\partial x_1} = (-\rho_d)^{n-1}$$

$$\frac{\partial g_n}{\partial x_2} = \frac{\partial(X_n - \rho_d X_{n-1} + \dots + (-\rho_d)^{n-1} X_1)}{\partial x_2} = (-\rho_d)^{n-2}$$

$$\frac{\partial g_3}{\partial x_2} = \frac{\partial(X_3 - \rho_d X_2 + \rho_d^2 X_1)}{\partial x_2} = -\rho_d$$

$$\frac{\partial g_2}{\partial x_2} = \frac{\partial(X_2 - \rho_d X_1)}{\partial x_2} = 1$$

$$\frac{\partial g_1}{\partial x_2} = \frac{\partial(x_1)}{\partial x_2} = 0$$

Η ιακωβιανή μετατροπή θα είναι ίση τότε με:

$$f_{Y_1, Y_2, \dots, Y_n}(y_1, y_2, \dots, y_n) = f_{X_1, X_2, \dots, X_n}(x_1, x_2, \dots, x_n)$$

Λόγω του ότι ο ιακωβιανός πίνακας ισούται με ένα, μπορούμε να γράψουμε την joint pdf  $\phi_{\underline{y}}(\underline{y} \| H_k, \theta)$  της μεταβλητής  $y$  (όπως είδαμε στην αρχή της ενότητας του paper αυτού) σαν:

$$\phi_{\underline{y}}(\underline{y} | H_k, \theta) = \phi_{\underline{x}}(y_1, y_2 - \rho_d y_1, \dots, y_n - \rho_d y_{n-1}) = \prod_{i=1}^n f_{\Lambda}(y_i - \theta \varepsilon_k b_{k,i})$$

Παρουσιάσαμε μια καινούργια μεταβλητή, την  $b_{k,i}$ . Δείτε πως προέκυψε:

- Έστω πως μελετούμε το σήμα '1' του δυαδικού συστήματος αποστολής. Άρα  $\varepsilon_k = \varepsilon_1$
- Το πρώτο δείγμα θορύβου είναι ίσο με:  

$$\Lambda_1 = y_1 - \theta \varepsilon_1 \widetilde{s}_{1,1}$$
- Το δεύτερο δείγμα θορύβου είναι ίσο με:  

$$\Lambda_2 = y_2 - \theta \varepsilon_1 \widetilde{s}_{1,2} - \rho_d (y_1 - \theta \varepsilon_1 \widetilde{s}_{1,1}) = y_2 - \theta \varepsilon_1 (\widetilde{s}_{1,2} - \rho_d \widetilde{s}_{1,1})$$
- Το τρίτο δείγμα θορύβου είναι ίσο με:  

$$\Lambda_3 = y_3 - \theta \varepsilon_1 \widetilde{s}_{1,3} - \rho_d (y_2 - \theta \varepsilon_1 \widetilde{s}_{1,2}) = y_3 - \theta \varepsilon_1 (\widetilde{s}_{1,3} - \rho_d \widetilde{s}_{1,2} + \rho_d^2 \widetilde{s}_{1,1})$$

Όπως παρατηρήσατε, κάθε επιπλέον δείγμα θορύβου έχει ένα παραπάνω όρο στην παρένθεση. Μπορούμε λοιπόν για εξοικονόμηση χώρου να ορίσουμε μια καινούργια μεταβλητή η οποία να ορίζει την παρένθεση αυτή. Θέτουμε λοιπόν

$$b_{1,i} = \sum_{j=1}^{i-1} (-\rho_d)^j \widetilde{s}_{1,i-j}$$

Άρα τα δείγματα θορύβου γίνονται:

$$\begin{aligned} \Lambda_1 &= y_1 - \theta \varepsilon_k b_{k,1} \\ \Lambda_2 &= y_2 - \theta \varepsilon_k b_{k,2} \\ \Lambda_3 &= y_3 - \theta \varepsilon_k b_{k,3} \\ &\cdot \\ &\cdot \\ &\cdot \\ \Lambda_n &= y_n - \theta \varepsilon_k b_{k,n} \end{aligned}$$



## Ανίχνευση ασθενών σημάτων σε πρώτης τάξης εξαρτημένο θόρυβο

Με αυτήν την τελευταία μετατροπή, καταφέραμε να μετατρέψουμε τον παρατηρούμενο χώρο δειγμάτων από  $\underline{X} = \mathcal{G}_{\varepsilon_k} \widetilde{s}_k + \underline{W}$ , σε  $\underline{Y} = \mathcal{G}_{\varepsilon_k} b_{k,i} + \underline{\Lambda}$ . Προσέξτε ότι όταν η παράμετρος εξάρτησης ισούται με το μηδέν, τότε έχουμε ότι  $\underline{X} = \underline{Y}$  και  $\underline{W} = \underline{\Lambda}$ .

### Κριτήρια απόφασης σε κρουστικό FOMA θόρυβο

Επιθυμούμε να βγάλουμε ένα κριτήριο με βάση το οποίο θα μπορούμε να εκτιμήσουμε εάν το λαμβανόμενο σήμα είναι '0', ή '1'. Όταν δεν γνωρίζουμε την ισχύ με την οποία αποστέλλεται το σήμα, τότε μιλάμε για την *locall optimal detection*, διαφορετικά μιλάμε για την *Μέγιστη πιθανοφάνεια-Maximum Likelihood Estimation*. Ας δούμε πρώτα την πρώτη περίπτωση.

#### *locall optimal detection*

Η ιδέα στην οποία βασίζεται αυτός ο εκτιμητής, είναι απλή, και βασίζεται στις πιθανοφάνειες:

- Η πιθανοφάνεια το λαμβανόμενο διάνυσμα δειγμάτων !!!!! να έχει τιμή της υπόθεσης !!!!!
- Η πιθανοφάνεια το λαμβανόμενο διάνυσμα δειγμάτων !!!!! να έχει τιμή της υπόθεσης !!!!!

Εάν λοιπόν κάποια από τις δύο πιθανότητες είναι πιο μεγάλη, τότε επιλέγω και την αντίστοιχη υπόθεση. Μια βελτίωση αυτής της ιδέας έρχεται εάν μπορούμε να αποφανθούμε υπέρ της απόφασης μιας υπόθεσης ακόμα και αν η ισχύς του λαμβανόμενου σήματος είναι σχεδόν μηδενική. Εάν μπορέσουμε να διαμορφώσουμε τον εκτιμητή μας για την χειρότερη αυτή περίπτωση, τότε λέμε ότι έχουμε τον *βέλτιστο τοπικό εκτιμητή*. Βέλτιστο διότι, τον έχουμε διαμορφώσει για την χειρότερη περίπτωση, και τοπικό διότι εξετάζουμε την λειτουργία του σε ένα συγκεκριμένο επίπεδο στάθμης ισχύος του σήματος. Με άλλα λόγια, το κριτήριο είναι το εξής:

Η παράγωγος παίζει τον ρόλο της εύρεσης του τοπικού μέγιστου της συνάρτησης της πιθανοφάνειας

$$D_k^S = \left\{ \underline{y} : \left. \frac{\partial}{\partial \mathcal{G}} \varphi_{\underline{Y}}(\underline{y} | H_k, \mathcal{G}) \right|_{\mathcal{G}=0} \geq \left. \frac{\partial}{\partial \mathcal{G}} \varphi_{\underline{Y}}(\underline{y} | H_m, \mathcal{G}) \right|_{\mathcal{G}=0} \forall m \right\}$$

Το σημείο στο οποίο εξετάζουμε την-πιθανοφάνεια του λαμβανόμενου διανύσματος. Αυτό που κάνουμε ουσιαστικά, είναι να συγκρίνουμε την κλίση που θα έχει έκαστο συνάρτηση πιθανοφάνειας στο σημείο όπου η ισχύς του σήματος είναι μηδέν. Συγκρίνοντας όμως την κλίση, ουσιαστικά συγκρίνουμε την διασπορά της κατανομής της υπόθεσης  $H_k$  και  $H_m$ . Όσο πιο μεγάλη είναι η κλίση, τόσο πιο μικρή είναι η διασπορά. Όσο πιο μικρή είναι η διασπορά, τόσο πιο σίγουροι μπορούμε να είμαστε για την υπόθεση αυτή, άρα μικρότερη πιθανότητα λάθους!

Αν αναπτυχθεί η παράγωγος, θα έχουμε

$$\begin{aligned}
 & \frac{\partial}{\partial \theta} \phi_Y(\underline{y}|H_k, \theta)|_{\theta=0} \\
 &= f'_\Lambda(y_1 - \theta \epsilon_k b_{k,1})(-\epsilon_k b_{k,1}) \cdot f_\Lambda(y_2 - \theta \epsilon_k b_{k,2}) \cdots \\
 & \quad \cdot f_\Lambda(y_n - \theta \epsilon_k b_{k,n})|_{\theta=0} + \cdots + f_\Lambda(y_1 - \theta \epsilon_k b_{k,1}) \\
 & \quad \cdot f_\Lambda(y_2 - \theta \epsilon_k b_{k,2}) \cdots f'_\Lambda(y_n - \theta \epsilon_k b_{k,n})(-\epsilon_k b_{k,n})|_{\theta=0} \\
 &= \prod_{j=1}^n f_\Lambda(y_j) \cdot \sum_{i=1}^n \epsilon_k b_{k,i} g_\Lambda(y_i) \tag{18}
 \end{aligned}$$

Όπου  $g_\Lambda(y_i) = -\frac{f'_\Lambda(y_i)}{f_\Lambda(y_i)}$  είναι μια βαθμίδα η οποία έχει σαν σκοπό να ‘διαμορφώσει’ το παρατηρούμενο διάνυσμα δειγμάτων, έτσι ώστε να μπορέσει ο δέκτης να αντιμετωπίσει τα καταστροφικά για το σήμα μας, αποτελέσματα του θορύβου.

Αναπτύσσοντας την παράγωγο και για την άλλη πιθανοφάνεια, και αντικαθιστώντας στην ανισότητα, έχουμε

$$D_k^S = \left\{ \underline{y} : \sum_{i=1}^n (\epsilon_k b_{k,i} - \epsilon_m b_{m,i}) g_\Lambda(y_i) \geq 0 \quad \forall m \right\}$$

Θέτουμε  $q_{k,m,i} = \epsilon_k b_{k,i} - \epsilon_m b_{m,i}$ , και έχουμε:

$$D_k^S = \left\{ \underline{y} : \sum_{i=1}^n q_{k,m,i} g_\Lambda(y_i) \geq 0 \quad \forall m \right\}$$

### Maximum Likelihood Detection

Για αυτού του τύπου εκτιμητή, μπορείτε να ανατρέξετε στην ενότητα I. Η φιλοσοφία αυτού του τύπου εκτιμητή, είναι ακριβώς η ίδια και για i.i.d. δείγματα, και για εξαρτημένα δείγματα, αφού όπως είδατε, καταφέραμε να μετασχηματίσουμε τα εξαρτημένα δείγματα σε ανεξάρτητα  $\underline{X} \rightarrow \underline{Y}$ . Ας θυμηθούμε την φιλοσοφία αυτού του εκτιμητή:

Η ιδέα της *Εκτίμησης Μέγιστης Πιθανοφάνειας*, είναι ακόμα πιο απλή από την *local optimal Detection* μιας και αν ισχύει η ανισότητα  $D_k^{ML} = \left\{ \underline{y} : \varphi_Y(\underline{y}|H_k, \mathcal{G}) \geq \varphi_Y(\underline{y}|H_m, \mathcal{G}) \quad \forall m \right\}$ , τότε αποφασίζουμε ότι το διάνυσμα  $\underline{y}$  ανήκει στην υπόθεση  $H_k$ , αλλιώς ανήκει στην υπόθεση  $H_m$ . Διαιρώντας κατά μέλη, έχουμε:

$$D_k^{ML} = \left\{ \underline{y} : \prod_{i=1}^n \frac{\varphi_Y(\underline{y} | H_k, \vartheta)}{\varphi_Y(\underline{y} | H_m, \vartheta)} \geq 1 \quad \forall m \right\}$$

Αν πάρουμε τον φυσικό λογάριθμο αυτής της αναλογίας, η σειρά γινομένου μετασχηματίζεται σε σειρά αθροίσματος (ιδιότητα φυσικού λογαρίθμου) και έχουμε:

$$D_k^{ML} = \left\{ \underline{y} : \sum_{i=1}^n \ln \frac{\varphi_Y(\underline{y} | H_k, \vartheta)}{\varphi_Y(\underline{y} | H_m, \vartheta)} \geq 0 \quad \forall m \right\}$$

Η σειρά γινομένου μετασχηματίστηκε σε σειρά αθροίσματος

Παίρνοντας τον φυσικό λογάριθμο του 1, το κατώφλι είναι πλέον το 0

Αν γνωρίζουμε την συνάρτηση πυκνότητας πιθανότητα που ακολουθεί ο θόρυβος (έστω Gaussian), τότε η πιθανοφάνεια  $\varphi_Y(\underline{y} | H_k, \vartheta)$  γράφεται:

$$\frac{1}{2\sqrt{\gamma\pi}} \exp\left(-\frac{(\text{παρατηρούμενο δείγμα} - \text{σήμα που εκπέμψαμε})^2}{4\gamma}\right)$$

Τι συμπέρασμα μπορούμε να βγάλουμε από τους δύο αυτούς εκτιμητές;

**ΑΝ ΔΕΝ ΓΝΩΡΙΖΟΥΜΕ ΤΗΝ ΣΥΝΑΡΤΗΣΗ ΠΥΚΝΟΤΗΤΑΣ ΠΙΘΑΝΟΤΗΤΑΣ ΠΟΥ ΕΚΦΡΑΖΕΙ ΤΟΝ ΘΟΡΥΒΟ, ΤΟΤΕ Η ΑΠΟΛΟΣΗ ΤΟΥ ΔΕΚΤΗ (BER) ΘΑ ΜΕΙΩΘΕΙ ΔΡΑΜΑΤΙΚΑ.**

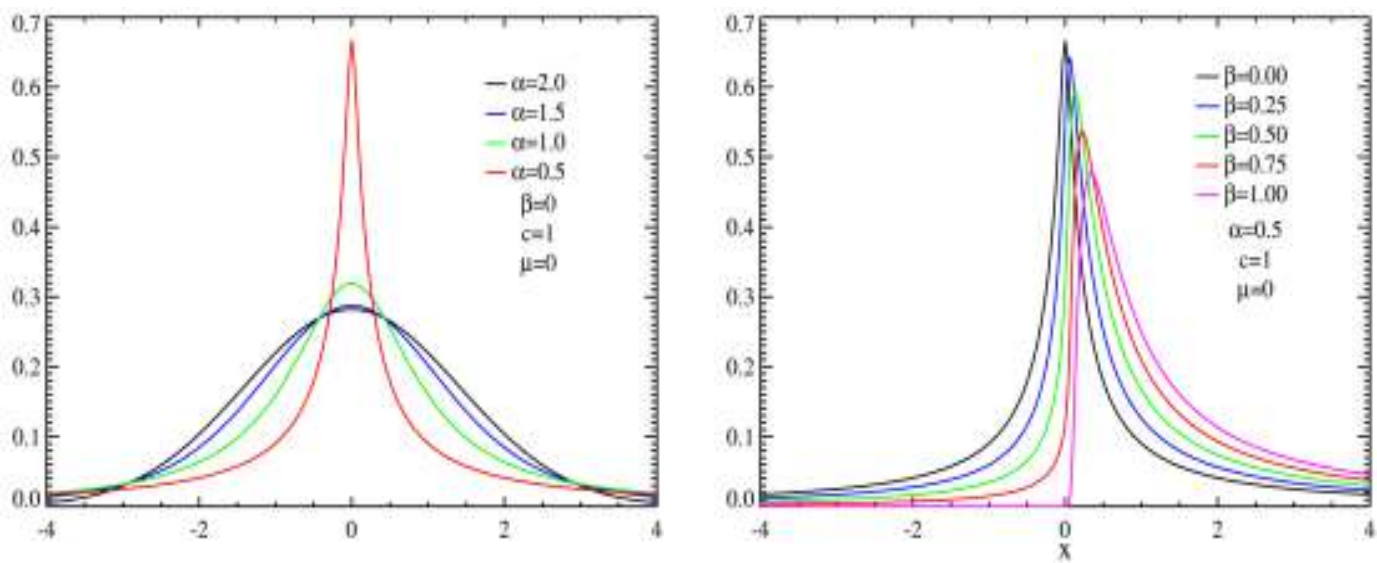
## Εύρεση του καταλλήλου κριτηρίου ανάλογα με την κατανομή που ακολουθεί ο θόρυβος

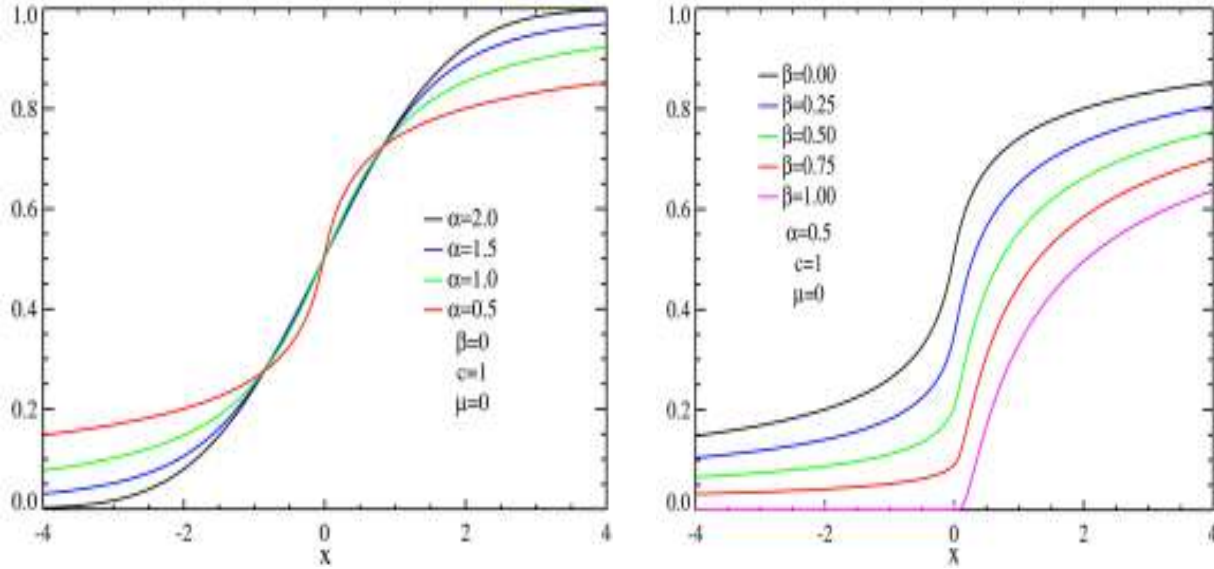
Ας θυμηθούμε λίγο την θεωρία των Stable κατανομών. Η σπουδαιότητα των Stable κατανομών είναι ότι 'μοντελοποιούν' iid τυχαίες μεταβλητές. Η κανονική κατανομή, είναι μια κλάση-οικογένεια των Stable κατανομών. Σύμφωνα με το κεντρικό οριακό θεώρημα, το κανονικοποιημένο άθροισμα από ένα σετ τυχαίων μεταβλητών με πεπερασμένη διακύμανση, τείνει να ακολουθεί κανονική κατανομή καθώς ο αριθμός των τυχαίων μεταβλητών αυξάνεται. Χωρίς όμως την υπόθεση της πεπερασμένης διακύμανσης, δεν θα σχηματίζεται κανονική κατανομή, αλλά κάποια άλλη κατανομή από την οικογένεια των Stable κατανομών.

Οι Stable κατανομές, χαρακτηρίζονται από τέσσερις παραμέτρους:

- Παράμετρος θέσης-Location parameter ' $\mu$ '
- Παράμετρος κλιμάκωσης-Scale parameter ' $c$ '
- Παράμετρος ασυμμετρίας-Asymmetry parameter ' $\beta$ '
- Παράμετρος συγκέντρωσης-Concentration parameter ' $\alpha$ ' (Αυτή η παράμετρος δηλώνει και την κρουστικότητα της κατανομής. Όσο η τιμή ' $\alpha$ ' πλησιάζει το μηδέν, τόσο πιο κρουστική είναι η κατανομή όπως βλέπετε στο παρακάτω σχήμα.)

Δείτε πως αυτές οι παράμετροι επηρεάζουν το σχήμα της κατανομής:





Στην δική μας περίπτωση, η παράμετρος θέσης θα είναι ίση με μηδέν  $\mu=0$ , η παράμετρος κλιμάκωσης θα είναι ίση με ένα  $c=1$  η παράμετρος ασυμμετρίας θα είναι ίση με μηδέν  $\beta=0$ , και ο παράγοντας που θα πειράζουμε είναι η παράμετρος συγκέντρωσης  $\alpha$ . Παρατηρήστε από το πρώτο σχήμα ότι ‘πειράζοντας’ την παράμετρος συγκέντρωσης, αυτό που κάνουμε είναι να παραμετροποιούμε την κρουστικότητα του θορύβου!

Για διάφορες τιμές της  $\alpha$  λοιπόν θα έχουμε την αντίστοιχη κατανομή:

$$f_{\Lambda}(x) = \begin{cases} \frac{1}{\pi} \sum_{l=1}^{\infty} \frac{(-1)^{l-1} \Gamma(l\alpha+1)}{\gamma^{-l} l! |x|^{l\alpha+1}} \sin\left(\frac{l\alpha\pi}{2}\right), & \text{for } 0 < \alpha \leq 1 \\ \frac{\gamma}{\pi(x^2+\gamma^2)}, & \text{for } \alpha = 1 \\ \frac{1}{\pi\alpha} \sum_{l=0}^{\infty} \frac{(-1)^l x^{2l}}{(2l)! \gamma^{(2l+1)/\alpha}} \Gamma\left(\frac{2l+1}{\alpha}\right), & \text{for } 1 < \alpha \leq 2 \\ \frac{1}{2\sqrt{\gamma\pi}} \exp\left(-\frac{x^2}{4\gamma}\right), & \text{for } \alpha = 2 \end{cases}$$

## Gaussian κατανομή πιθανότητας

Η συνάρτηση πυκνότητας πιθανότητας μιας κανονικής τυχαίας μεταβλητής  $X$  με μέσο όρο  $\mu$  και τυπική απόκλιση  $\sigma$ , δίνεται από:

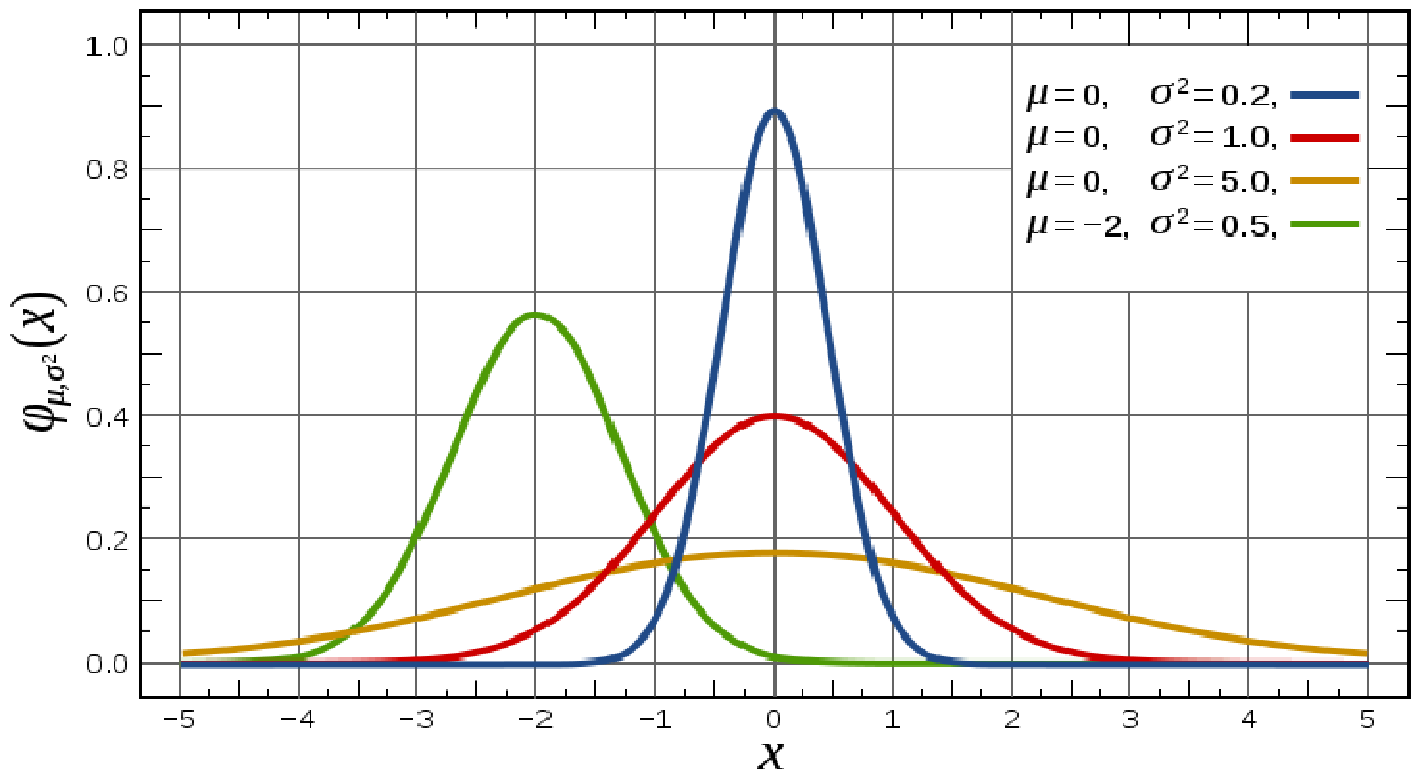
$$f(x|\mu, \sigma) = \frac{1}{\sigma\sqrt{2\pi}} \exp\left[-\frac{(x-\mu)^2}{2\sigma^2}\right], -\infty < x < \infty, -\infty < \mu < \infty, \sigma > 0$$

Η κατανομή δίνεται από το όρισμα  $N(\mu, \sigma^2)$ .

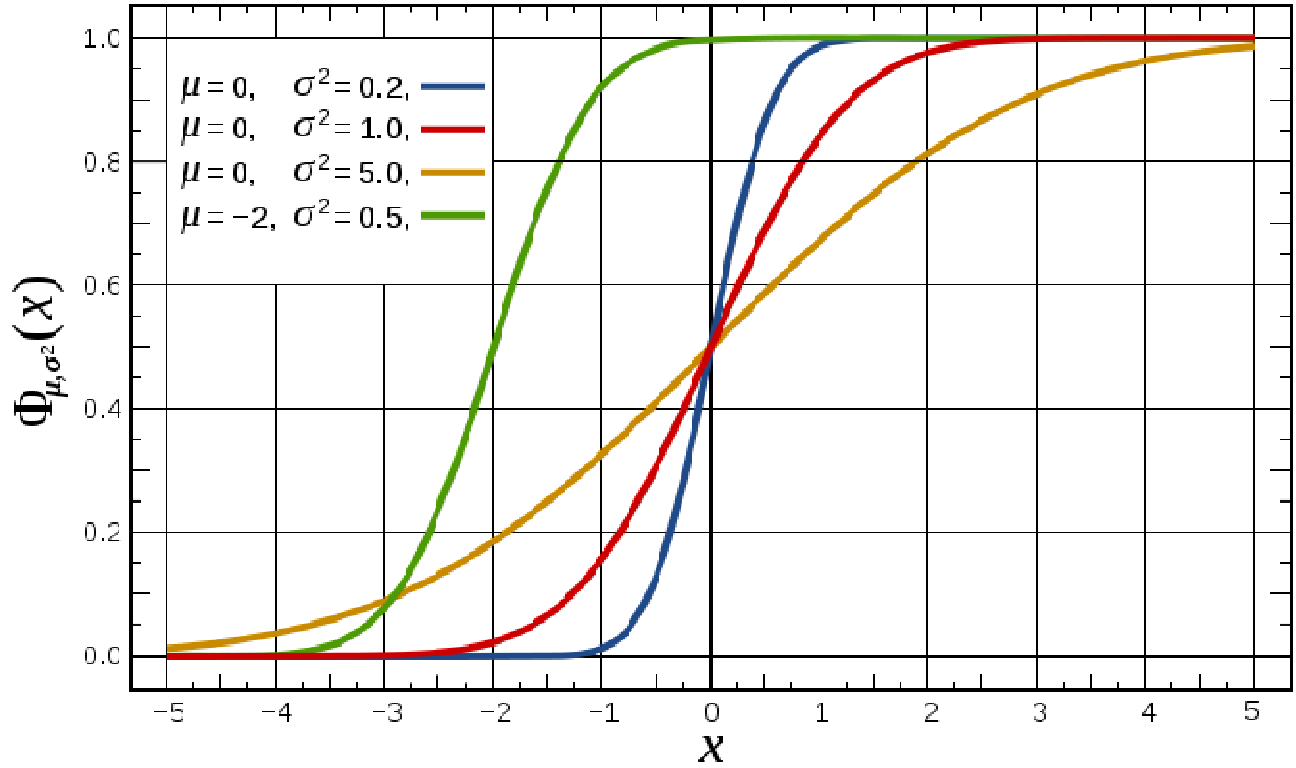
Η συνάρτηση συσσωρευτικού αθροίσματος (*cumulative distribution function-cdf*) δίνεται από:

$$F(x|\mu, \sigma) = \int_{-\infty}^x f(t|\mu, \sigma) dt$$

Η κανονική τυχαία μεταβλητή με μηδενικό μέσο όρο και τυπική απόκλιση ίση με ένα, ονομάζεται στάνταρ κανονική τυχαία μεταβλητή και η cdf δίνεται από την συνάρτηση  $\Phi(z)$ . Εάν η  $X$  είναι μια κανονική τυχαία μεταβλητή με μέσο όρο  $\mu$  και τυπική απόκλιση ίση με  $\sigma$ , τότε η κατανομή θα είναι της τέτοιας μορφής:



Συνάρτηση πυκνότητας πιθανότητας για διάφορες τιμές location-μέσου όρου, και διακύμανσης



Συνάρτηση αθροιστικής κατανομής-Cumulative distribution function

Το κριτήριο απόφασης για την περίπτωση που ο δέκτης δεν γνωρίζει την ισχύ με την οποία αποστέλλονται τα σήματα, είναι:

$$D_k^S = \left\{ \underline{y} : \sum_{i=1}^n q_{k,m,i} \left( -\frac{f'(y_i)}{f(y_i)} \right) \geq 0 \right\} \Rightarrow$$

$$D_k^S = \left\{ \underline{y} : \sum_{i=1}^n q_{k,m,i} \left( -\frac{\left( \frac{1}{2\sqrt{\gamma\pi}} \exp\left( -\frac{(y_i)^2}{4\gamma} \right) \right)'}{\frac{1}{2\sqrt{\gamma\pi}} \exp\left( -\frac{(y_i)^2}{4\gamma} \right)} \right) \geq 0 \right\} \Rightarrow D_k^S = \left\{ \underline{y} : \sum_{i=1}^n q_{k,m,i} \left( -\frac{\left( \frac{1}{2\sqrt{\gamma\pi}} \exp\left( -\frac{(y_i)^2}{4\gamma} \right) \right)'}{\frac{1}{2\sqrt{\gamma\pi}} \exp\left( -\frac{(y_i)^2}{4\gamma} \right)} \right) \geq 0 \right\} \Rightarrow$$

$$D_k^S = \left\{ \underline{y} : \sum_{i=1}^n q_{k,m,i} \left( -\frac{\left( \frac{1}{2\sqrt{\gamma\pi}} \exp\left( -\frac{(y_i)^2}{4\gamma} \right) \right) \left( -\frac{(y_i)^2}{4\gamma} \right)'}{\frac{1}{2\sqrt{\gamma\pi}} \exp\left( -\frac{(y_i)^2}{4\gamma} \right)} \right) \geq 0 \right\} \Rightarrow D_k^S = \left\{ \underline{y} : \sum_{i=1}^n q_{k,m,i} y_i \begin{matrix} \geq \\ < \end{matrix} 0 \right\}$$

## Ανίχνευση ασθενών σημάτων σε πρώτης τάξης εξαρτημένο θόρυβο

Στην περίπτωση που ο δέκτης γνωρίζει την ισχύ με την οποία αποστέλλονται τα σήματα, η μέγιστη πιθανοφάνεια θα δίνεται από την διαίρεση των δύο πιθανοφανειών

$$f_{\Lambda}(\underline{y} | H_k, \mathcal{G}) = \prod_{i=1}^n \frac{1}{2\sqrt{\gamma\pi}} \exp\left(-\frac{(y_i - \mathcal{G}\varepsilon_k b_{k,i})^2}{4\gamma}\right) \quad \text{για την υπόθεση } H_k, \text{ και}$$

$$f_{\Lambda}(\underline{y} | H_m, \mathcal{G}) = \prod_{i=1}^n \frac{1}{2\sqrt{\gamma\pi}} \exp\left(-\frac{(y_i - \mathcal{G}\varepsilon_m b_{m,i})^2}{4\gamma}\right) \quad \text{για την υπόθεση } H_m.$$

Το επόμενο βήμα είναι να διαιρέσουμε τις δύο υπο-συνθήκη πιθανότητες, και στην συνέχεια να λογαριθμίσουμε. Έτσι θα έχουμε:

$$D_K^{ML} = \underline{y} \left\{ \sum_{i=1}^n \left\{ \ln \frac{f_{\Lambda}(\underline{y} | H_k, \mathcal{G})}{f_{\Lambda}(\underline{y} | H_m, \mathcal{G})} \right\} \geq 0 \right\} \Rightarrow$$

$$D_K^{ML} = \left\{ \underline{y} : \ln \left( \sum_{i=1}^n \left( -\frac{(y_i - \mathcal{G}\varepsilon_k b_{k,i})^2}{4\gamma} - \left( -\frac{(y_i - \mathcal{G}\varepsilon_m b_{m,i})^2}{4\gamma} \right) \right) \right) \geq 0 \right\}$$

Κάνοντας τις απλές πράξεις, προκύπτει ότι το κριτήριο απόφασης είναι:

$$D_K^{ML} = \sum_{i=1}^n q_{k,m,i} y_i \geq \frac{\mathcal{G}}{2} \sum_{i=1}^n (\varepsilon_k^2 b_{k,i}^2 - \varepsilon_m^2 b_{m,i}^2)$$

Όταν ισχύει η ανισότητα αυτή τότε αποφασίζουμε ότι το σήμα είναι η υπόθεση  $H_k$ , και όταν το δεύτερο άθροισμα είναι μεγαλύτερο από το πρώτο άθροισμα, τότε αποφασίζουμε ότι το σήμα είναι η υπόθεση  $H_m$



## Cauchy κατανομή πιθανότητας.

Η κατανομή Cauchy έχει μεγάλο μαθηματικό ενδιαφέρον, λόγω της απουσίας καθορισμένων ροπών. Τα χαρακτηριστικά της είναι:

Παράμετρος θέσης	$x_0$ (ο μέσος όρος)
Παράμετρος κλιμάκωσης	$b > 0$
Εύρος	$-\infty < x < \infty$
Συνάρτηση πυκνότητας πιθανότητας	$\frac{1}{\pi\gamma \left[ 1 + \left( \frac{x-x_0}{\gamma} \right)^2 \right]}$
Συνάρτηση κατανομής	$\frac{1}{\pi} \arctan \left( \frac{x-x_0}{\gamma} \right) + \frac{1}{2}$
Χαρακτηριστική συνάρτηση	$\exp(x_0 it - \gamma  t )$
Ροπές	Δεν υπάρχουν
Ρυθμός	$a$
Median	$a$

Η κατανομή Cauchy είναι μοναδιαίου ρυθμού και συμμετρική, με πολύ βαρύτερες ουρές από την κανονική κατανομή και συμμετρική γύρω από το  $\mu$  (ο μέσος όρος της κατανομής). Η Cauchy μεταβλητή με στοιχεία μέσου όρου και κλιμάκωσης  $\mu$  και  $b$  αντίστοιχα, περιγράφεται σαν  $C : \mu, b$ . Η σχέση που συνδέει την κανονική Cauchy κατανομή (η οποία είναι  $C : 0, 1$  και δηλώνει μηδενικό μέσο όρο και παράμετρος κλιμάκωσης ίση με 1), με την  $C : \mu, b$  είναι:

$$C : \mu, b \sim \mu + b(C : 0, 1)$$

### Ιδιότητες της Cauchy κατανομής:

1. Η αναλογία δύο ανεξάρτητων μεταβλητών  $N_1, N_2$  οι οποίες εμφανίζουν Gaussian κατανομή, είναι η στάνταρ Cauchy κατανομή  $C : 0, 1$ .

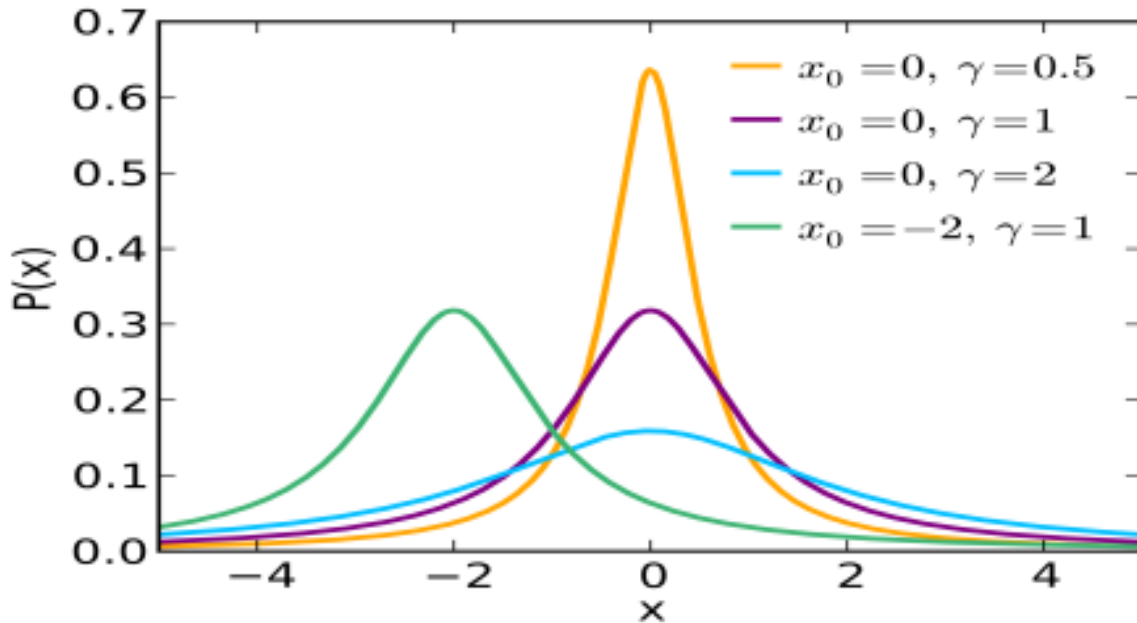
$$(N_1/N_2) \sim (C : 0, 1)$$

2. Η στάνταρ Cauchy μεταβλητή είναι μια ιδιαίτερη περίπτωση μιας μεταβλητής η οποία εμφανίζει student's t κατανομή με ένα βαθμό ελευθερίας.
3. Το άθροισμα  $n$ -ανεξάρτητων Cauchy μεταβλητών  $C : \mu_i, b_j$  με παράμετρο θέσης και παράμετρο κλιμάκωσης  $i = 1, 2, \dots, n$  και  $j = 1, 2, \dots, n$  αντίστοιχα, είναι μια Cauchy μεταβλητή  $C : \mu, b$  με παράμετρο θέσης ίση με το άθροισμα των επιμέρους παραμέτρους θέσεως και παράμετρο κλιμάκωσης ίση με το άθροισμα των διαφορετικών παραμέτρων κλιμάκωσης.

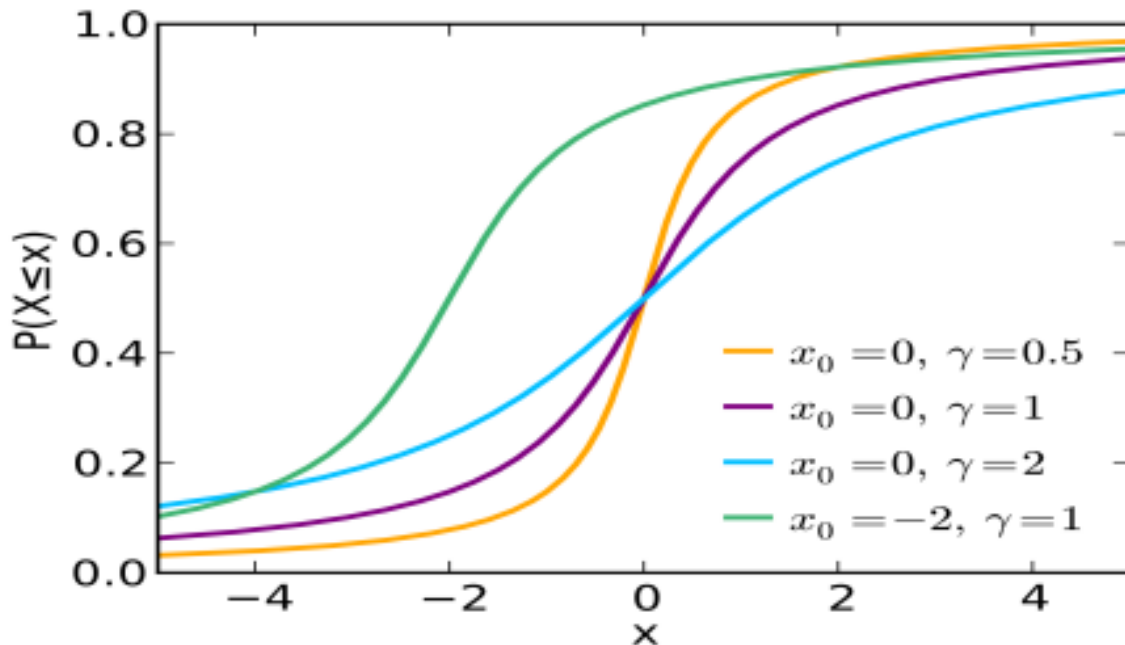
$$\mu = \sum_{i=1}^n \mu_i \quad b = \sum_{i=1}^n b_i$$

Ο μέσος όρος  $n$ -ανεξάρτητων Cauchy μεταβλητών με μεταβλητή  $C: \mu, b$ , είναι η μεταβλητή  $C: \mu', b'$ . Όπου  $\mu', b'$  δίνονται από  $1/(C: \mu, b) \sim C: \mu', b'$

$$\mu' = \mu / (\mu^2 + b^2) \quad b' = b / (\mu^2 + b^2)$$



Συνάρτηση πυκνότητας πιθανότητας για διάφορες τιμές location-μέσου όρου



Συνάρτηση αθροιστικής κατανομής-Cumulative distribution function

Η συνάρτηση πυκνότητας πιθανότητας είναι:  $f_{\Lambda}(y_i) = \frac{\gamma}{\pi(y_i^2 + \gamma^2)}$

Με την ίδια τακτική όπως ακολουθήσαμε και στην περίπτωση της Gaussian κατανομής, για να βρούμε την περιοχή απόφασης όταν δεν γνωρίζουμε την ισχύ με την οποία στέλνουμε το σήμα, θα πάρουμε τον γενικό τύπο:

$$D_k^S = \left\{ \underline{y} : \sum_{i=1}^n q_{k,m,i} g_{\Lambda}(y_i) \geq 0 \right\} \text{ και θα τον αναπτύξουμε:}$$

$$D_k^S = \left\{ \underline{y} : \sum_{i=1}^n q_{k,m,i} \left( -\frac{f'(y_i)}{f(y_i)} \right) \geq 0 \right\} \Rightarrow$$

$$D_k^S = \left\{ \underline{y} : \sum_{i=1}^n q_{k,m,i} \left( -\frac{\left( \frac{\gamma}{\pi(y_i^2 + \gamma^2)} \right)'}{\frac{\gamma}{\pi(y_i^2 + \gamma^2)}} \right) \geq 0 \right\} \Rightarrow$$

$$D_k^S = \left\{ \underline{y} : \sum_{i=1}^n q_{k,m,i} \left( -\frac{\left( \frac{\gamma' \cdot \pi(y_i^2 + \gamma^2) - \gamma(2\pi y_i)}{(\pi(y_i^2 + \gamma^2))^2} \right)}{\frac{\gamma}{\pi(y_i^2 + \gamma^2)}} \right) \geq 0 \right\} \Rightarrow D_k^S = \left\{ \underline{y} : \sum_{i=1}^n q_{k,m,i} \frac{2y_i}{y_i^2 + \gamma^2} \geq 0 \right\}$$

Προσέξτε ότι  $g_{\Lambda}(y_i) = \frac{2y_i}{y_i^2 + \gamma^2}$

Και για την περίπτωση που γνωρίζουμε την ισχύ με την οποία αποστέλλεται το σήμα, η Maximum Likelihood θα είναι:

$$D_K^{ML} = \underline{y} : \sum_{i=1}^n \left\{ \ln \frac{f_\Lambda(\underline{y} | H_k, \vartheta)}{f_\Lambda(\underline{y} | H_m, \vartheta)} \geq 0 \right\} \Rightarrow$$

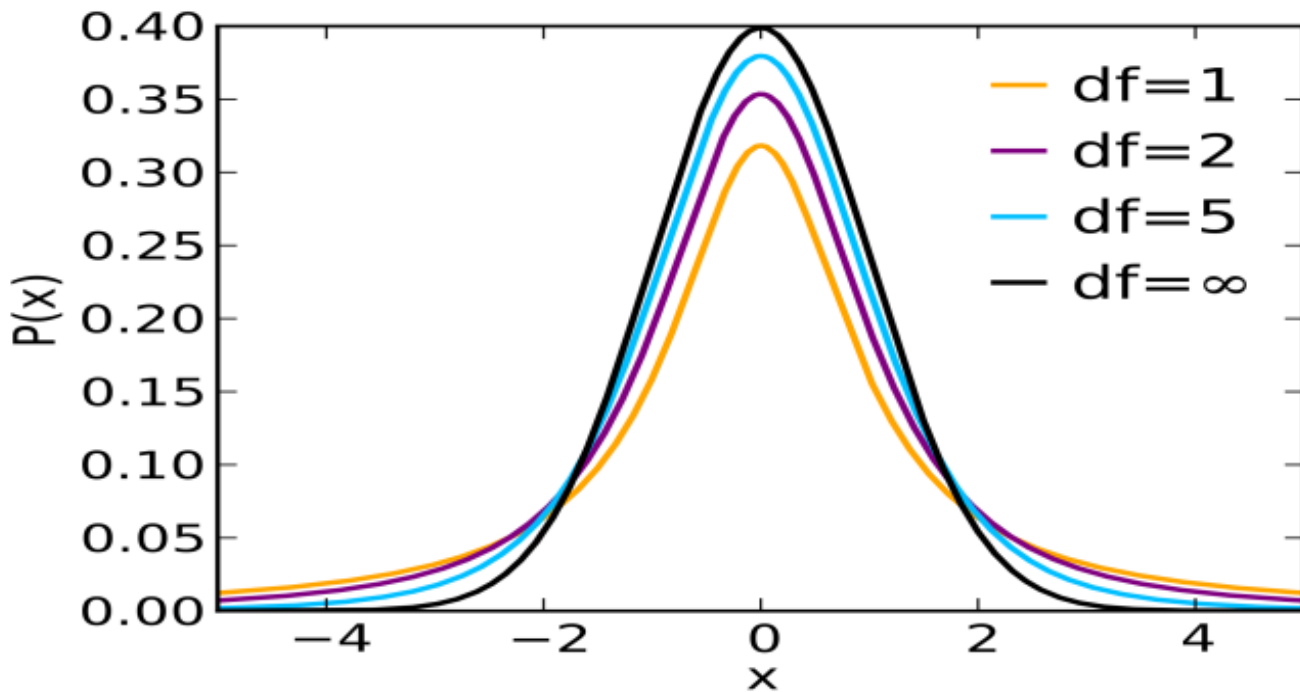
$$D_K^{ML} = \underline{y} : \sum_{i=1}^n \left\{ \ln \frac{\frac{\gamma}{\pi(\gamma^2 + (y_i - \vartheta \varepsilon_k b_{k,i})^2)}}{\frac{\gamma}{\pi(\gamma^2 + (y_i - \vartheta \varepsilon_m b_{m,i})^2)}} \geq 0 \right\} \Rightarrow$$

$$D_K^{ML} = \left\{ \underline{y} : \ln \frac{(y_i - \vartheta \varepsilon_m b_{m,i})^2 + \gamma^2}{(y_i - \vartheta \varepsilon_k b_{k,i})^2 + \gamma^2} \geq 0 \right\} \Rightarrow D_K^{ML} = \left\{ \underline{y} : \ln \frac{(y_i - \vartheta \varepsilon_k b_{k,i})^2 + \gamma^2}{(y_i - \vartheta \varepsilon_m b_{m,i})^2 + \gamma^2} \leq 0 \right\}$$

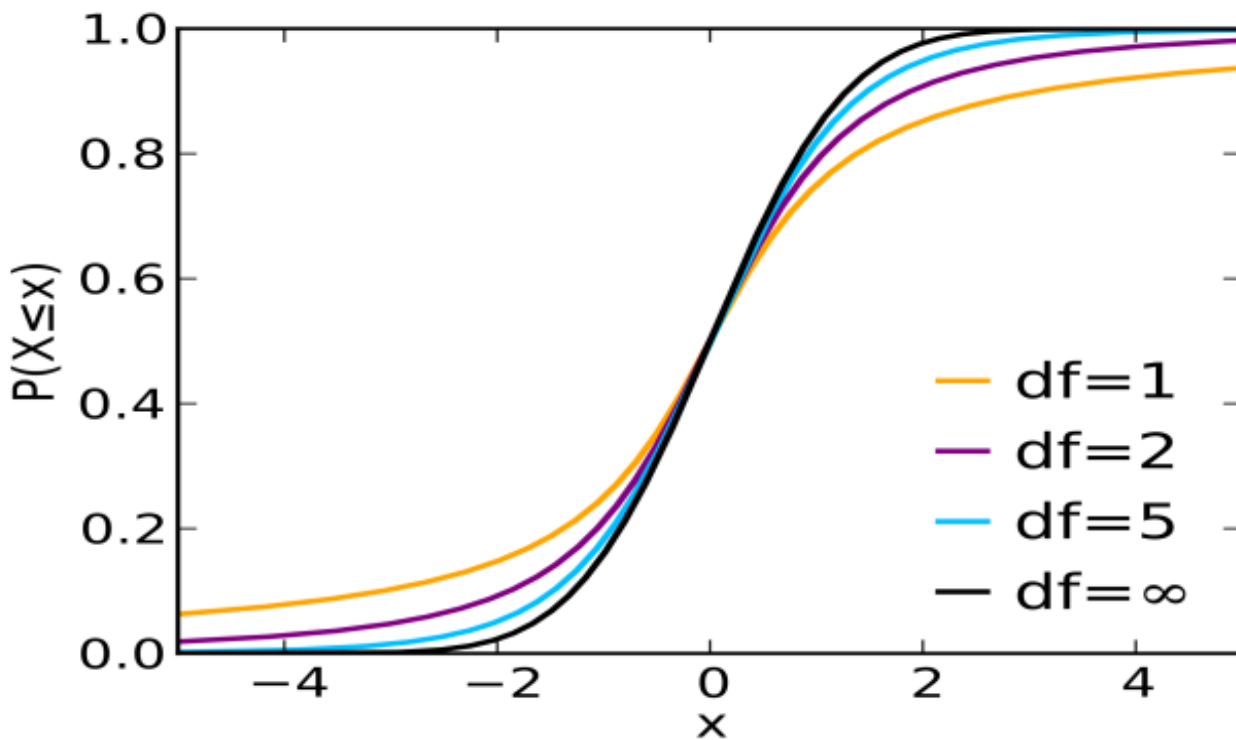
### Student's-t Κατανομή

Η κατανομή Student's  $t$  χρησιμοποιείται για να δούμε εάν ο μέσος όρος δύο δειγμάτων μιας παρατήρησης, είναι στατιστικά σημαντικός. Για να γίνουμε πιο ακριβείς, θα ήταν σαν να τεστάρουμε την υπόθεση ότι και τα δύο δείγματα 'τραβιούνται' από την ίδια κατανομή. Αν η μεταβλητή 't' πάρει μεγάλη τιμή, τότε συμπεραίνουμε ότι τα δύο αυτά δείγματα δεν ανήκουν στην ίδια κατανομή, και ο μέσος όρος των δειγμάτων τους θα είναι έτσι πολύ διαφορετικός. Τα χαρακτηριστικά μεγέθη της κατανομής αυτής, είναι:

Student's-t Distribution	
Παράμετροι	$\nu > 0$ βαθμοί ελευθερίας
Πεδίο ορισμού	$x \in (-\infty; +\infty)$
Συνάρτηση πυκνότητας πιθανότητας	$\frac{\Gamma\left(\frac{\nu+1}{2}\right)}{\sqrt{\nu\pi}\Gamma\left(\frac{\nu}{2}\right)} \left(1 + \frac{x^2}{\nu}\right)^{-\frac{\nu+1}{2}}$
median	0
mean	0 όταν $\nu > 0$ , αλλιώς απροσδιόριστος
ρυθμός	0
διακύμανση	$\frac{\nu}{\nu-2}$ για $\nu > 2$ $\infty$ για $1 < \nu \leq 2$ Αλλιώς απροσδιόριστη
στρεβλότητα	0 για $\nu > 3$
κύρτωση	$\frac{6}{\nu-4}$ για $\nu > 4$



Συνάρτηση πυκνότητας πιθανότητας για διάφορες τιμές location-μέσου όρου. Παρατηρήστε ότι καθώς ο αριθμός των βαθμών ελευθερίας τείνει στο άπειρο, η κατανομή παίρνει το σχήμα της Gaussian κατανομής



Συνάρτηση αθροιστικής κατανομής-Cumulative distribution function

Ακολουθώντας την ίδια τακτική όπως και την περίπτωση της Gaussian και της Cauchy κατανομής, θα πρέπει να βρούμε τον όρο  $g_{\Lambda}(y) = -\frac{f'_{\Lambda}(y)}{f_{\Lambda}(y)}$  ο οποίος δίνει τέτοια μορφή στο σήμα, ώστε να αντιμετωπίζει καλύτερα τις επιδράσεις του θορύβου. Πάμε λοιπόν:

$$f'_{\Lambda}(y) = \frac{\Gamma\left(\frac{\nu+1}{2}\right)}{\sqrt{\nu\pi}\Gamma\left(\frac{\nu}{2}\right)} \cdot \left(-\frac{\nu+1}{2}\right) \cdot \left(1 + \frac{y^2}{\nu}\right)^{\frac{\nu+1}{2}-1} \cdot \left(1 + \frac{y^2}{\nu}\right)' = -\frac{\Gamma\left(\frac{\nu+1}{2}\right)}{\sqrt{\nu\pi}\Gamma\left(\frac{\nu}{2}\right)} \cdot \left(\frac{y}{\nu}\right) \cdot \left(\frac{\nu+1}{2}\right) \cdot \left(1 + \frac{y^2}{\nu}\right)^{\frac{\nu+1}{2}} \cdot \left(1 + \frac{y^2}{\nu}\right)^{-1}$$

Αντικαθιστώντας στην  $g_{\Lambda}(y) = -\frac{f'_{\Lambda}(y)}{f_{\Lambda}(y)}$ , έχουμε:

$$g_{\Lambda}(y) = -\frac{f'_{\Lambda}(y)}{f_{\Lambda}(y)} = \frac{-\frac{\Gamma\left(\frac{\nu+1}{2}\right)}{\sqrt{\nu\pi}\Gamma\left(\frac{\nu}{2}\right)} \cdot \left(\frac{y}{\nu}\right) \cdot \left(\frac{\nu+1}{2}\right) \cdot \left(1 + \frac{y^2}{\nu}\right)^{\frac{\nu+1}{2}} \cdot \left(1 + \frac{y^2}{\nu}\right)^{-1}}{\frac{\Gamma\left(\frac{\nu+1}{2}\right)}{\sqrt{\nu\pi}\Gamma\left(\frac{\nu}{2}\right)} \cdot \left(1 + \frac{y^2}{\nu}\right)^{\frac{\nu+1}{2}}} = \frac{(\nu+1) \cdot y}{\nu + y^2}$$

Αντικαθιστώντας στην  $D_k^S = \left\{ \underline{y} : \sum_{i=1}^n q_{k,m,i} g_{\Lambda}(y_i) \geq 0 \right\}$ , έχουμε το κριτήριο:

$$D_k^S = \left\{ \underline{y} : \sum_{i=1}^n \frac{q_{k,m,i} \cdot (\nu+1) \cdot y}{\nu + y^2} \geq 0 \right\}$$

Πολλαπλασιάζοντας με  $(\nu+1)^{-1}$ , έχουμε ότι το κριτήριο είναι:

$$D_k^S = \left\{ \underline{y} : \sum_{i=1}^n \frac{q_{k,m,i} \cdot y}{\nu + y^2} \geq 0 \right\}$$

Από τη θεωρία της στατιστικής, γνωρίζουμε ότι όταν οι βαθμοί ελευθερίας στην κατανομή Student-t προσεγγίζουν το άπειρο, τότε η απόδοση του εκτιμητή ο οποίος είναι βέλτιστος για την κατανομή Student-t, είναι ίδια με την επίδοση της Gaussian εκτιμητριάς.

Για να βγάλουμε την εκτιμήτρια για την περίπτωση που γνωρίζουμε την ισχύ με την οποία αποστέλλεται το σήμα, παίρνουμε την πιθανοφάνεια:

$$D_K^{ML} = \underline{y} : \left\{ \sum_{i=1}^n \left\{ \ln \frac{f_{\Lambda}(\underline{y} | H_k, \vartheta)}{f_{\Lambda}(\underline{y} | H_m, \vartheta)} \right\} \geq 0 \right\}$$

Και αντικαθιστώντας, έχουμε:

$$D_K^{ML} = \underline{y} : \left\{ \sum_{i=1}^n \left\{ \ln \frac{\frac{\Gamma\left(\frac{\nu+1}{2}\right)}{\sqrt{\nu\pi}\Gamma\left(\frac{\nu}{2}\right)} \left(1 + \frac{(y - \vartheta \varepsilon_k b_{k,i})^2}{\nu}\right)^{-\frac{\nu+1}{2}}}{\frac{\Gamma\left(\frac{\nu+1}{2}\right)}{\sqrt{\nu\pi}\Gamma\left(\frac{\nu}{2}\right)} \left(1 + \frac{(y - \vartheta \varepsilon_m b_{m,i})^2}{\nu}\right)^{-\frac{\nu+1}{2}}} \right\} \geq 0 \right\}$$

$$\underline{y} : \sum_{i=1}^n \left\{ \ln \frac{\left(\frac{\nu + (y - \vartheta \varepsilon_k b_{k,i})^2}{\nu}\right)^{-\frac{\nu+1}{2}}}{\left(\frac{\nu + (y - \vartheta \varepsilon_m b_{m,i})^2}{\nu}\right)^{-\frac{\nu+1}{2}}} \right\} \geq 0 \Rightarrow \underline{y} : \sum_{i=1}^n \left\{ \ln \frac{\left(\frac{\nu + (y - \vartheta \varepsilon_k b_{k,i})^2}{\nu}\right)^{-\frac{\nu+1}{2}}}{\left(\frac{\nu + (y - \vartheta \varepsilon_m b_{m,i})^2}{\nu}\right)^{-\frac{\nu+1}{2}}} \right\} \geq 0 \Rightarrow$$

$$\underline{y} : \sum_{i=1}^n \left\{ \ln \frac{\left(\frac{\nu + (y - \vartheta \varepsilon_m b_{m,i})^2}{\nu}\right)^{-\frac{\nu+1}{2}}}{\left(\frac{\nu + (y - \vartheta \varepsilon_k b_{k,i})^2}{\nu}\right)^{-\frac{\nu+1}{2}}} \right\} \geq 0 \Rightarrow \underline{y} : \sum_{i=1}^n \left\{ \ln \frac{\left(\frac{\nu + (y - \vartheta \varepsilon_k b_{k,i})^2}{\nu}\right)^{-\frac{\nu+1}{2}}}{\left(\frac{\nu + (y - \vartheta \varepsilon_m b_{m,i})^2}{\nu}\right)^{-\frac{\nu+1}{2}}} \right\} \leq 0 \Rightarrow$$

$$\underline{y} : \sum_{i=1}^n \left\{ \left(\frac{\nu+1}{2}\right) \ln \frac{\left(\frac{\nu + (y - \vartheta \varepsilon_k b_{k,i})^2}{\nu}\right)^{-\frac{\nu+1}{2}}}{\left(\frac{\nu + (y - \vartheta \varepsilon_m b_{m,i})^2}{\nu}\right)^{-\frac{\nu+1}{2}}} \right\} \leq 0 \Rightarrow \underline{y} : \sum_{i=1}^n \left\{ \ln \frac{\left(\frac{\nu + (y - \vartheta \varepsilon_k b_{k,i})^2}{\nu}\right)^{-\frac{\nu+1}{2}}}{\left(\frac{\nu + (y - \vartheta \varepsilon_m b_{m,i})^2}{\nu}\right)^{-\frac{\nu+1}{2}}} \right\} \leq 0$$

Αυτό είναι λοιπόν και το κριτήριο μέγιστης πιθανοφάνειας-*maximum likelihood* για την περίπτωση που ο θόρυβος ακολουθεί Student-t κατανομή.

Όπως προσέξατε, προτείναμε τρία βελτιωμένα μοντέλα εκτιμητών, για την περίπτωση που ο θόρυβος παρουσιάζει αλαφρά εξάρτηση από τα προηγούμενα δείγματα του. Όμως, η κατηγορία των Stable κατανομών έχει άλλες δύο κατανομές, και αυτές προκύπτουν όταν  $0 < a < 1$  και όταν  $1 < a < 2$ . Η ιδιαιτερότητα των κατανομών σε αυτές τις περιπτώσεις, είναι ότι δεν μπορούν να εκφραστούν σε κλειστή μορφή συναρτήσεων. Το αποτέλεσμα αυτής της ιδιαιτερότητας είναι

ότι δεν μπορούμε να πάρουμε την παράγωγο της, για να βγάλουμε την βαθμίδα  $g_{\lambda}(y) = -\frac{f'_{\lambda}(y)}{f_{\lambda}(y)}$ . Ο ρόλος αυτής της βαθμίδας είναι να αλλάξει, ή με άλλα λόγια, 'να φέρει το

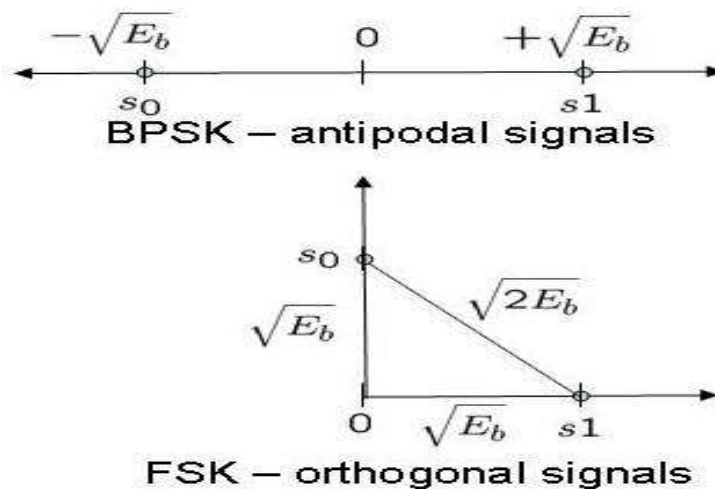
σήμα στα μέτρα μας', ανάλογα με το είδος του θορύβου που την διαμόρφωσε (τι κατανομή ακολούθησε ο θόρυβος), έτσι ώστε η εκτίμηση του σήματος να γίνει με μικρότερη πιθανότητα λάθους.

Άρα αυτό που μπορούμε να κάνουμε σε αυτές τις δύο περιπτώσεις, είναι να πειραματιστούμε χρησιμοποιώντας τις βαθμίδες  $g_{\lambda}(y)$  που χρησιμοποιήσαμε για την Gaussian και την Cauchy κατανομή, και να δούμε ποια βαθμίδα προσφέρει καλύτερη απόδοση.

### Εφαρμογή σε Antipodal Signaling Systems

Η τεχνική που χρησιμοποιείται σε αυτά τα συστήματα, είναι ότι αντί να εκπέμπουμε το σήμα '0' με μηδενική ισχύ, ή ισχύ χαμηλότερη από αυτήν με την οποία αποστέλλεται το σήμα '1', εκπέμπουμε με την ίδια ισχύ και αυτό που αλλάζει είναι το πρόσημο της τάσης. Για παράδειγμα, εάν το πλάτος της τάσης που συμβολίζει το σήμα '1' είναι 5 Volt, το σήμα '0' θα εκπέμπεται με τάση -5 Volt. Αυτή η τεχνική προσφέρει σημαντική αύξηση στην απόδοση του δέκτη. Μην ξεχνάμε ότι όσο πιο μεγάλη απόσταση έχει η έναστος τάση των δύο σημάτων, τόσο πιο μεγάλη θα είναι και η διακριτική ικανότητα που θα έχει ο δέκτης.

Για να γίνει πιο κατανοητή η ιδέα της *Antipodal* και *Orthogonal* μετάδοσης, δείτε το παρακάτω σχήμα:



*Αναπαράσταση με μορφή διανυσμάτων των Orthogonal και Antipodal signals*



Με βάση αυτό, η ανίχνευση των *antipodal signals* στον μετασχηματισμένο χώρο θα διεξάγεται υπό την υπόθεση:

$$H_k : \underline{Y} = \mathcal{G}\varepsilon_k \underline{b}_k + \underline{\Lambda} \quad \text{όπου } k = 0,1$$

Άρα το σύστημα υπόθεσης θα είναι:

$$H : \underline{Y} = \begin{cases} \mathcal{G}\varepsilon_1 \underline{b}_1 + \underline{\Lambda} \\ \mathcal{G}\varepsilon_0 \underline{b}_0 + \underline{\Lambda} \end{cases}$$

Όπως θυμάστε, τα σύμβολα  $\varepsilon_1$  και  $\varepsilon_0$  συμβολίζουν το κατά πόσο χαμηλότερη, ή υψηλότερη θα είναι η στάθμη ισχύος του έκαστο σήματος, σε σχέση με την κοινή παράμετρος ισχύος  $\mathcal{G}$ . Άρα για το σήμα '1', αλλά και για το σήμα '0' θα έχουμε  $\varepsilon_1 = \varepsilon_0 = 1$ . Η παράμετρος που θα μας βοηθάει να διακρίνουμε τα σήματα το ένα από το άλλο, είναι η παράμετρος  $b_k$ , και θα παίρνει τιμές  $b_1 = -b_0$ . Άρα εκεί που το  $b_{1,i}$  ισούταν με:

$$b_{1,i} = \frac{1}{\sqrt{n}} \sum_{j=1}^{i-1} (-\rho_d)^j$$

Τώρα η παράμετρος  $b_{0,i}$  θα ισούται με:

$$b_{0,i} = -\frac{1}{\sqrt{n}} \sum_{j=1}^{i-1} (-\rho_d)^j$$

Συνεπώς η παράμετρος  $q_{k,m}$  που είδαμε νωρίτερα, θα ισούται με:

$$q_{k,m} = \varepsilon_k b_k - \varepsilon_m b_m = 2b_k = 2b_m$$

Επίσης η *joint pdf* του  $\underline{Y}$  θα είναι ίση με:

$$\phi_{\underline{Y}}(\underline{y} | H_k, \mathcal{G}) = \prod_{i=1}^n f_{\Lambda}(y_i - \mathcal{G}b_{k,i})$$

Και το κριτήριο απόφασης για τον ανιχνευτή S-ML θα είναι:

$$D_k^S : \underline{y} = \sum_{i=1}^n 2b_{k,i} g_\Lambda(y_i) \stackrel{H_1}{\geq} 0 \Rightarrow$$
$$D_k^S : \underline{y} = \sum_{i=1}^n b_{k,i} g_\Lambda(y_i) \stackrel{H_1}{\geq} 0$$

Αν δεν ισχύει η ανισότητα, τότε θα έχουμε την υπόθεση  $H_0$  αντί η υπόθεση  $H_1$ .

Για το κριτήριο μέγιστης πιθανοφάνειας (*Maximum Likelihood Estimation*) η μετατροπή θα είναι:

$$D_k^{ML} : \underline{y} = \sum_{i=1}^n \ln \frac{f_\Lambda(y_i - \mathcal{G}b_{1,i})}{f_\Lambda(y_i + \mathcal{G}b_{1,i})} \stackrel{H_1}{\geq} 0$$

Και πάλι εάν δεν ισχύει η ανισότητα, θα έχω την υπόθεση  $H_0$  αντί η υπόθεση  $H_1$ .

Μην ξεχνάτε ότι ο λόγος που έχω το πρόσημο '+' στον παρανομαστή είναι λόγω του ότι  $b_1 = -b_0$ .

Αυτές λοιπόν είναι οι δύο βασικές σχέσεις για τους δύο εκτιμητές που παρουσιάζουμε. Με την ίδια τακτική όπως ακολουθήσαμε προηγουμένως, για να βγάλουμε τις σχέσεις-κριτήρια που ισχύουν για τις περιπτώσεις Gaussian, Cauchy και Student't (η περίπτωση που η συνάρτηση πυκνότητας πιθανότητας, δεν έχει κλειστή μορφής έκφραση, όπως όταν  $0 < a \leq 1$  και όταν  $1 < a \leq 2$ , δεν μπορούμε να έχουμε μια κλειστή έκφραση για την παράμετρο  $g_\Lambda(y_i)$

## Αριθμητικά αποτελέσματα και συμπεράσματα

Για την διεξαγωγή των καμπύλων του ρυθμού λάθους ( $BER$ ), για τους εκτιμητές *Maximum-likelihood* και *Optimal Detection (S-ML)* χρησιμοποιήσαμε εξομοίωση *Monte Carlo* με  $10^5$  επαναλήψεις σε κάθε διαφορετική τιμή του  $G$ -SNR, και υποθέτουμε ότι ο βαθμός εξάρτησης  $\rho_d$  είναι γνωστός στο σύστημα. Ο λόγος που χρησιμοποιήσαμε  $10^5$  επαναλήψεις είναι λόγω του ότι μας επιτρέπει να ανιχνεύσουμε μέχρι και ένα λάθος bit ανάμεσα σε  $10^5$  bits.

Η μέθοδος *Monte Carlo* (ή πειράματα *Monte Carlo*) είναι μια κλάση από υπολογιστικούς αλγορίθμους που βασίζονται σε επανειλημμένα τυχαία δείγματα, για να γίνει ο υπολογισμός των αποτελεσμάτων. Αυτή η μέθοδος χρησιμοποιείται συχνά στην εξομοίωση φυσικών και μαθηματικών συστημάτων.

Λόγω του ότι η διακύμανση των SaS κατανομών δεν μπορεί να οριστεί όταν  $\alpha \neq 2$ , εκτός από την περίπτωση που  $\alpha = 2$  όπου έχουμε την Gaussian κατανομή το κανονικό SNR δεν μπορεί να χρησιμοποιηθεί. Αντί λοιπόν του γνωστού SNR, θα χρησιμοποιήσουμε το γεωμετρικό SNR, το οποίο δίδει ένα πιο σωστό χαρακτηρισμό του σηματοθορυβικού λόγου όταν η διακύμανση του θορύβου τείνει στο άπειρο.

Το γεωμετρικό SNR στον μετασχηματισμένο παρατηρούμενο χώρο δίνεται από:

$$G-SNR = \frac{A_k^2}{2C_g^{-1+2\alpha} \gamma^{2/\alpha}}$$

- Όπου  $A_k^2 = \mathcal{G}^2 \varepsilon_k^2 \|b_k\|_2^2$  είναι η ισχύς του μετασχηματισμένου σήματος  $\mathcal{G} \varepsilon_k b_k$ .
- Όπου  $C_g = \exp \left\{ -\lim_{s \rightarrow \infty} \left( \ln s - \sum_{z=1}^s \frac{1}{z} \right) \right\} \approx 1.78$  η οποία είναι το εκθετικό της σταθεράς του Euler.
- Όταν  $\alpha = 2$  τότε το γεωμετρικό SNR αναπαριστά το στάνταρ SNR.
- Υποθέτουμε ότι η παράμετρος  $\gamma$  ισούται με ένα.

Λόγω του γεωμετρικού σηματοθορυβικού λόγου, θα πρέπει να ακολουθήσουμε την εξής τακτική για να δούμε ποια θα πρέπει να είναι η τιμή που θα πρέπει να δώσουμε στο σήμα που εκπέμπεται.

## Ανίχνευση ασθενών σημάτων σε πρώτης τάξης εξαρτημένο θόρυβο

Αν θέσουμε την παράμετρο κρουστικότητας  $\alpha = 1$ , τότε θα έχουμε την Cauchy κατανομή. Αντικαθιστώντας λοιπόν όπου  $\alpha = 1$ , και όπου  $C_g = 1.78$ , λύνουμε ως προς  $A_k^2$ :

$$G - SNR = \frac{A_k^2}{2C_g^{-1+2\alpha}\gamma^{2/\alpha}} \Rightarrow A_k^2 = 2C_g^{-1+2\alpha}\gamma^{2/\alpha} \cdot G - SNR \Rightarrow$$

$$A_k^2 = 2 \cdot 1.78^{-1+2 \cdot 1} \cdot 1^{2/1} \cdot G - SNR \Rightarrow \mathcal{G}^2 \varepsilon_k^2 \|b_k\|_2^2 = 3.56 \cdot G - SNR$$

Ας υποθέσουμε τώρα ότι δεν έχουμε εξάρτηση του παρών δείγματος θορύβου από το προηγούμενο. Αυτό σημαίνει ότι το αριστερό μέλος της εξίσωσης  $(\mathcal{G}^2 \varepsilon_k^2 \|b_k\|_2^2)$  θα είναι η ισχύς με την οποία αποστέλλεται το σήμα μας. Για ευκολία, έστω η ισοδύναμη αντίσταση της κεραίας μας να ισούται με  $1\Omega$ . Άρα παίρνοντας την τετραγωνική ρίζα του αριστερού μέλους, παίρνουμε την στάθμη τάσης με την οποία αποστέλλεται το σήμα μας.

Δεν υπάρχει κανένα πρόβλημα στην υπόθεση ότι δεν έχουμε εξάρτηση ανάμεσα στα δείγματα. Αυτό διότι το πρώτο δείγμα που θα πάρω από το σήμα, δεν έχει κάποιο προηγούμενο σήμα από το οποίο θα εξαρτηθεί! Με την υπόθεση λοιπόν αυτή, αυτό που καταφέρνω είναι να πάρω την στάθμη τάσεως με την οποία αποστέλλεται το σήμα μου  $\underline{s}_k = \mathcal{G} \varepsilon_k \widetilde{s}_k$ .

## Κώδικας που γράφηκε στο Matlab για την εξαγωγή των διαγραμμάτων BER

<p>Κώδικας για την εξομοίωση του <i>Maximum Likelihood Antipodal</i> εκτιμητή στην περίπτωση Cauchy θορύβου.</p>	<p>Κώδικας για την εξομοίωση του <i>Local Optimal Antipodal</i> εκτιμητή στην περίπτωση Cauchy θορύβου.</p>
<pre>function[p]=cauchy_ml(SNR_in_DB)  rd1=0.1; rd2=0.1; runs=100000; numoerr=0; ek=0; snr=exp(SNR_in_DB*log(10)/10); n=30;  for jj=1:runs     E=sqrt(3.56*snr);     bk=zeros(1, n);     bm=zeros(1, n);     x=zeros(1, n);     y=zeros(1, n);     em=ek;      if (em==0)         ek=1;         source=1;     else         ek=0;         source=0;     end      for i=1:n         bk(i)=1/sqrt(n);         x(i)=E+cauchyrnd(0, 1);         y(i)=x(i);     end      for i=2:n         m=i-1;         nn=i-1;         for j=1:m             y(i)=y(i)+power((rd1),j)*y(nn);             nn=nn-1;         end     end      bk(1)=1/sqrt(n);</pre>	<pre>function[p]=cauchy_sml(SNR_in_DB)  rd1=0.1; rd2=0.1; runs=100000; numoerr=0; ek=0; snr=exp(SNR_in_DB*log(10)/10); n=30;  for jj=1:runs     E=sqrt(2*snr);     bk=zeros(1, n);     bm=zeros(1, n);     x=zeros(1, n);     y=zeros(1, n);     em=ek;      if (em==0)         ek=1;         source=1;     else         ek=0;         source=0;     end      for i=1:n         bk(i)=1/sqrt(n);         x(i)=E+cauchyrnd(0, 1);         y(i)=x(i);     end      for i=2:n         m=i-1;         nn=i-1;         for j=1:m             y(i)=y(i)+power((rd1),j)*y(nn);             nn=nn-1;         end     end      bk(1)=1/sqrt(n);     for i=2:n</pre>

<pre> for i=2:n     m=i-1;     nn=i-1;     for j=1:m         y(i)=y(i)+power((-rd),j)*y(nn)         bk(i)=bk(i)+power((-rd),j)*bk(nn);         nn=nn-1;     end end  for i=1:n     y(i)=y(i)*(1-power(-rd1,i))/(1+rd1); end  sum=0; for i=1:n     bm(i)=bk(i);     arithm=power((y(i)-ek*E*bk(i)),2)+1;     paranom=power((y(i)-em*E*bm(i)),2)+1;     klasma=log(arithm/paranom);     sum=sum+klasma;     klasma=0; end  if (sum&gt;0)     decis=0; else     decis=1; end if (decis~=source)     numoerr=numoerr+1; end end  p=numoerr/runs; </pre>	<pre> m=i-1; nn=i-1; for j=1:m     y(i)=y(i)+power((-rd),j)*y(nn);     bk(i)=bk(i)+power((-rd),j)*bk(nn);     nn=nn-1; end end  for i=1:n     y(i)=y(i)*(1-power(-rd1,i))/(1+rd1); end  sum=0; for i=1:n     bm(i)=bk(i);     ginomeno_tou_kathe_loop=((ek*bk(i)- em*bm(i))*y(i))/(1+y(i).^2);     sum=sum+ginomeno_tou_kathe_loop; end  if (sum&lt;0)     decis=0; else     decis=1; end  if (decis~=source)     numoerr=numoerr+1; end end if (decis~=source)     numoerr=numoerr+1; end end  p=numoerr/runs; </pre>
---	--

## Ανίχνευση ασθενών σημάτων σε πρώτης τάξης εξαρτημένο θόρυβο

Κώδικας για την εξομοίωση του <i>Maximum Likelihood Antipodal</i> εκτιμητή στην περίπτωση Gauss θορύβου.	Κώδικας για την εξομοίωση του <i>Local Optimal Antipodal</i> εκτιμητή στην περίπτωση Gauss θορύβου.
<pre>function[p]=gauss_ml(SNR_in_DB)  rd1=0; rd2=0; runs=100000; numoerr=0; ek=0; snr=exp(SNR_in_DB*log(10)/10); n=30;  for jj=1:runs      E=sqrt(2*snr);     bk=zeros(1, n);     bm=zeros(1, n);     x=zeros(1, n);     y=zeros(1, n);     em=ek;      if (em==0)         ek=1;         source=1;     else         ek=0;         source=0;     end      for i=1:n         bk(i)=1/sqrt(n);         x(i)=E+random('norm',0,1);         y(i)=x(i);     end      for i=2:n         m=i-1;         nn=i-1;         for j=1:m             y(i)=y(i)+power((rd1),j)*y(nn);             nn=nn-1;         end     end      bk(1)=1/sqrt(n);     for i=2:n         m=i-1;         nn=i-1;         for j=1:m             y(i)=y(i)+power((-rd2),j)*y(nn);</pre>	<pre>function[p]=gauss_sml(SNR_in_DB)  rd1=0.1; rd2=0.1; runs=100000; numoerr=0; ek=0; snr=exp(SNR_in_DB*log(10)/10); n=30;  for jj=1:runs      E=sqrt(2*snr);     bk=zeros(1, n);     bm=zeros(1, n);     x=zeros(1, n);     y=zeros(1, n);     em=ek;      if (em==0)         ek=1;         source=1;     else         ek=0;         source=0;     end      for i=1:n         bk(i)=1/sqrt(n);         x(i)=E+random('norm',0,1);         y(i)=x(i);     end      for i=2:n         m=i-1;         nn=i-1;         for j=1:m             y(i)=y(i)+power((rd1),j)*y(nn);             nn=nn-1;         end     end      bk(1)=1/sqrt(n);     for i=2:n         m=i-1;         nn=i-1;         for j=1:m             y(i)=y(i)+power((-rd2),j)*y(nn);</pre>

## Ανίχνευση ασθενών σημάτων σε πρώτης τάξης εξαρτημένο θόρυβο

<pre> bk(i)=bk(i)+power((-rd2),j)*bk(nn); nn=nn-1;     end end  for i=1:n     y(i)=y(i)*(1-power(-rd1,i))/(1+rd1); end  sum2=0; sum=0; for i=1:n     bm(i)=bk(i);     ginomeno_tou_kathe_loop=(ek*bk(i)- em*bm(i))*y(i);     deksio_melos=(ek.^2)*(bk(i).^2)-(em.^2)*(bm(i).^2);     sum=sum+ginomeno_tou_kathe_loop;     sum2=sum2+deksio_melos*E*0.5; end if (sum&lt;sum2)     decis=0; else     decis=1; end  if (decis~=source)     numoerr=numoerr+1; end end  p=numoerr/runs;         </pre>	<pre> bk(i)=bk(i)+power((-rd2),j)*bk(nn); nn=nn-1;     end end  for i=1:n     y(i)=y(i)*(1-power(-rd1,i))/(1+rd1); end  sum=0; for i=1:n     bm(i)=bk(i);     ginomeno_tou_kathe_loop=(ek*bk(i)-em*bm(i))*y(i);     sum=sum+ginomeno_tou_kathe_loop; end  if (sum&lt;0)     decis=0; else     decis=1; end if (decis~=source)     numoerr=numoerr+1; end end  p=numoerr/runs;         </pre>
--	--



## Ανίχνευση ασθενών σημάτων σε πρώτης τάξης εξαρτημένο θόρυβο

Κώδικας για την εξομοίωση του <i>Maximum Likelihood Antipodal</i> εκτιμητή στην περίπτωση student-t θορύβου.	Κώδικας για την εξομοίωση του <i>Local Optimal Antipodal</i> εκτιμητή στην περίπτωση student-t θορύβου.
<pre>function[p]=anti_t_ml(SNR_in_DB)  degrees_of_freedom=7; rd1=0; rd2=0; runs=100000; numoerr=0; ek=-1; snr=exp(SNR_in_DB*log(10)/10); n=30;  for jj=1:runs      E=sqrt(2*snr);     bk=zeros(1, n);     bm=zeros(1, n);     x=zeros(1, n);     y=zeros(1, n);      em=ek;     if (em==-1)         ek=1;         source=1;     else         ek=-1;         source=0;     end      for i=1:n         bk(i)=1/sqrt(n);         x(i)=E+random('t',degrees_of_freedom);         y(i)=x(i);     end      for i=2:n         m=i-1;         nn=i-1;         for j=1:m             y(i)=y(i)+power((rd1),j)*y(nn);             nn=nn-1;         end     end      bk(1)=1/sqrt(n);     for i=2:n         m=i-1;         nn=i-1;</pre>	<pre>function[p]=anti_t_sml(SNR_in_DB) degrees_of_freedom=3; rd1=0.1; rd2=0.1; runs=100000; numoerr=0; ek=-1; snr=exp(SNR_in_DB*log(10)/10); n=30;  for jj=1:runs      E=sqrt(2*snr);     bk=zeros(1, n);     bm=zeros(1, n);     x=zeros(1, n);     y=zeros(1, n);     em=ek;     if (em==-1)         ek=1;         source=1;     else         ek=-1;         source=0;     end      for i=1:n         bk(i)=1/sqrt(n);         x(i)=E+random('t',degrees_of_freedom);         y(i)=x(i);     end      for i=2:n         m=i-1;         nn=i-1;         for j=1:m             y(i)=y(i)+power((rd1),j)*y(nn);             nn=nn-1;         end     end      bk(1)=1/sqrt(n);     for i=2:n         m=i-1;         nn=i-1;</pre>

## Ανίχνευση ασθενών σημάτων σε πρώτης τάξης εξαρτημένο θόρυβο

<pre> for j=1:m     y(i)=y(i)+power((-rd2),j)*y(nn);     bk(i)=bk(i)+power((-rd2),j)*bk(nn);     nn=nn-1; end end  for i=1:n     y(i)=y(i)*(1-power(-rd1,i))/(1+rd1); end sum=0; for i=1:n     bm(i)=bk(i);     arithm=power((y(i)-ek*E*bk(i)),2)+degrees_of_freedom;     paranom=power((y(i)-em*E*bm(i)),2)+degrees_of_freedom;     klasma=log(arithm/paranom);     sum=sum+klasma;     klasma=0; end  if (sum&gt;0)     decis=0; else     decis=1; end  if (decis~=source)     numoerr=numoerr+1; end end p=numoerr/runs; </pre>	<pre> for j=1:m     y(i)=y(i)+power((-rd2),j)*y(nn);     bk(i)=bk(i)+power((-rd2),j)*bk(nn);     nn=nn-1; end end  for i=1:n     y(i)=y(i)*(1-power(-rd1,i))/(1+rd1); end sum=0; for i=1:n     bm(i)=bk(i);     ginomeno_tou_kathe_loop=((ek*bk(i)-em*bm(i))*y(i))/(degrees_of_freedom+y(i).^2);     sum=sum+ginomeno_tou_kathe_loop; end  if (sum&lt;0)     decis=0; else     decis=1; end  if (decis~=source)     numoerr=numoerr+1; end end p=numoerr/runs; </pre>
--	---

## Επεξήγηση του κώδικα για την περίπτωση του *Maximum Likelihood Estimator*

Όπως θα δείτε, το να αναλύσουμε και τους 6 αλγόριθμους:

- Gaussian Estimator (Maximum Likelihood & Sub-Optimum Maximum Likelihood)
- Cauchy Estimator (Maximum Likelihood & Sub-Optimum Maximum Likelihood)
- Student-t Estimator (Maximum Likelihood & Sub-Optimum Maximum Likelihood)

Κρίνεται ανούσιο, διότι το μόνο που αλλάζει στην έκαστο περίπτωση, είναι κομμάτια του κώδικα όπως η στάθμη του σήματος που εκπέμπεται, ανάλογα με την κατανομή του θορύβου, η γεννήτρια του θορύβου, και η σχέση εκτίμησης.

<code>function[p]=cauchy_ml_2(SNR_in_DB)</code>	Η συνάρτηση η οποία επιστρέφει την πιθανότητα λάθους 'p'
<code>rd1=0.1; rd2=0.1;</code>	Λόγω του ότι ο συντελεστής εξάρτησης είναι γνωστός στο σύστημα, πρέπει να τον δηλώσουμε. Όπου rd1 είναι ο συντελεστής εξάρτησης με βάση τον οποίο έχουμε την εξάρτηση ανάμεσα στα δείγματά, και όπου rd2 είναι ο συντελεστής εξάρτησης που <b>υποθέτει</b> ο δέκτης ότι ισχύει. Πράγματι, εάν $rd1 \neq rd2$ , τότε η απόδοση του δέκτη είναι μειωμένη, όπως θα δείτε και στο διάγραμμα BER.
<code>runs=100000;</code>	Εκτελώ 100000 κύκλους στον κώδικα. Αυτό μου δίνει την δυνατότητα να ανιχνεύσω μικρότερης πιθανότητας λάθους, και μια πιο ομαλή καμπύλη του BER
<code>numoerr=0;</code>	Αρχικοποίηση του μετρητή λαθών
<code>ek=0;</code>	Μια τεχνική που χρησιμοποίησα για να εναλλάσσονται οι τιμές των $\varepsilon_k$ και $\varepsilon_m$ . Με αυτόν τον τρόπο είναι σαν να έχω έναν πομπό ο οποίος στέλνει μια training ακολουθία μηδέν και ένα, γνωστή στον δέκτη, και έτσι να κάνει την εκτίμηση της πιθανότητας λάθους.
<code>snr=exp(SNR_in_DB*log(10)/10);</code>	Λόγω του ότι η συνάρτηση δέχεται σαν όρισμα από την main την τιμή του G-SNR σε db, πρέπει να την απολογαριθμίσουμε.
<code>n=30;</code>	Αριθμός δειγμάτων που θα συλλέξω από το σήμα μου για την εκτίμηση
<code>for jj=1:runs</code>	Ένα loop το οποίο θα τρέξει 100000 φορές και κάθε φορά θα βγάζει μια τιμή πιθανότητας λάθους. Αυτές οι τιμές στο τέλος θα αθροιστούν και θα διαιρεθούν με τον αριθμό 100000 για να βγει η πιθανότητα

## Ανίχνευση ασθενών σημάτων σε πρώτης τάξης εξαρτημένο θόρυβο

	λάθους σε κάθε τιμή του G-SNR
<code>E=sqrt(3.56*snr);</code>	Λόγω του ότι δεν μπορούμε να χρησιμοποιήσουμε τον κλασσικό σηματοθορυβικό λόγο επειδή διαδικασίες οι οποίες έχουν άπειρη διακύμανση, χρησιμοποιούμε το γεωμετρικό σηματοθορυβικό λόγο G-SNR. Αυτό που πρέπει να κάνουμε είναι στο G-SNR είναι να αντικαταστήσουμε όπου 'α' την τιμή που έχει η κάθε κατανομή.
<code>bk=zeros(1, n); bm=zeros(1, n); x=zeros(1, n); y=zeros(1, n);</code>	Μια τεχνική pre-allocation που δεσμεύουμε θέσεις μνήμης όσος ο αριθμός της δειγματοληψίας. Αυτό βοηθάει στην ταχύτητα εκτέλεσης του κώδικα
<code>em=ek; if (em==0)     ek=1;     source=1; else     ek=0;     source=0; end</code>	Το κομμάτι αυτό του κώδικα είναι υπεύθυνο στο να κάνω εκτίμηση για διαδοχικά '1' και '0'.
<code>for i=1:n     bk(i)=1/sqrt(n);     x(i)=E+cauchyrnd(0, 1);     y(i)=x(i); end</code>	Το κομμάτι αυτό του κώδικα, είναι υπεύθυνο στο να "γεμίσει" όλη την περίοδο του bit που κάνω train, με τυχαίο θόρυβο με μηδενικό μέσο όρο ο οποίος επικάθεται στο σήμα με πλάτος E volts. Εντολή <code>x(i)=E+cauchyrnd(0, 1);</code> Στο συγκεκριμένο παράδειγμα ο θόρυβος προέρχεται από κατανομή Cauchy. Επίσης καταχωρώ σε πίνακα <code>y(i)</code> (ο οποίος συμβολίζει τα εξαρτημένα δείγματα) , τα ανεξάρτητα προς το παρών δείγματα που προέρχονται από τον πίνακα <code>x(i)</code> . Εντολή <code>y(i)=x(i);</code> Για τον ίδιο ακριβώς λόγο, θα πρέπει να δημιουργήσουμε ένα καινούργιο πίνακα ο οποίος θα χρησιμοποιηθεί στην εκτίμηση του σήματος, και θα περιλαμβάνει την κανονικοποίηση (ίση με $1/\sqrt{n}$ ), η οποία θα είναι σε εξαρτημένη μορφή επίσης. Εντολή <code>bk(i)=1/sqrt(n);</code>
<code>for i=2:n     m=i-1;     nm=i-1;     for j=1:m         y(i)=y(i)+power((rd1),j)*y(nm);         nm=nm-1;     end end</code>	Το κομμάτι αυτό είναι υπεύθυνο στο να δημιουργήσει την εξάρτηση ανάμεσα στα δείγματα
<code>for i=2:n</code>	Το κομμάτι αυτό του κώδικα είναι υπεύθυνο στην

## Ανίχνευση ασθενών σημάτων σε πρώτης τάξης εξαρτημένο θόρυβο

<pre> m=i-1; nn=i-1; for j=1:m     y(i)=y(i)+power((-rd2),j)*y(nn);     bk(i)=bk(i)+power((-rd2),j)*bk(nn);     nn=nn-1; end end         </pre>	<p>δημιουργία των εξαρτημένων δειγμάτων από τα προηγούμενα δείγματα.</p>
<pre> sum=0; for i=1:n     bm(i)=bk(i);     arithm=power((y(i)-ek*E*bk(i)),2)+1;     paranom=power((y(i)-em*E*bm(i)),2)+1;     klasma=log(arithm/paranom);     sum=sum+klasma;     klasma=0; end         </pre>	<p>Το κομμάτι αυτό του κώδικα είναι υπεύθυνο για την υλοποίηση του κατωφλίου απόφασης:</p> $D_k^S = \left\{ \underline{y} : \sum_{i=1}^n q_{k,m,i} \frac{2y_i}{y_i^2 + \gamma^2} \geq 0 \right\}$
<pre> if (sum&gt;0)     decis=0; else     decis=1; end         </pre>	<p>Εάν το άθροισμα που προέκυψε από το παραπάνω κομμάτι κώδικα είναι μεγαλύτερο του μηδενός, τότε η εκτίμηση μας είναι η υπόθεση <math>H_k</math>. Αλλιώς η εκτίμηση μας είναι η υπόθεση <math>H_m</math>.</p>
<pre> if (decis~=source)     numoerr=numoerr+1; end         </pre>	<p>Εάν η εκτίμηση που προέκυψε από το παραπάνω κομμάτι κώδικα, είναι διαφορετική από το σήμα που όντως εκπέμφθηκε, τότε αύξησε τον μετρητή του λάθους κατά ένα.</p>
<pre> p=numoerr/runs;         </pre>	<p>Μετά από 100000 κύκλους εκτίμησης, για να προκύψει η πιθανότητα λάθους για την συγκεκριμένη τιμή του G-SNR, το μόνο που έχουμε να κάνουμε είναι να διαιρέσουμε τον αριθμό λαθών που προέκυψαν, με τον αριθμό των 100000 κύκλων.</p>

## Αποτελέσματα του Matlab

- Για αρχή, έστω πως θέλουμε να μελετήσουμε την απόδοση ενός **Gaussian S-ML εκτιμητή, υπό Gaussian θόρυβο**.
  - Οι παράμετροι που θα πρέπει να προσέξω στον κώδικα του Matlab για την διεξαγωγή της καμπύλης BER, είναι:

$$\rho_{d1} = 0$$

$$\rho_{d2} = 0$$

Οι παράμετροι που καθορίζουν την εξάρτηση θα πρέπει να μείνουν στην μηδενική τιμή

$$E = \text{sqrt}(2 * \text{snr});$$

το σήμα που αποστέλλεται. Το '2' προέκυψε από το G-SNR αντικαθιστώντας όπου  $\alpha=2$

$$x(i) = E + \text{random}('norm', 0, 1);$$

Μία από τις πολλές έτοιμες γεννήτριες θορύβου στο Matlab. Η συγκεκριμένη συνάρτηση παράγει τυχαία δείγματα τα οποία προέρχονται από Gauss κατανομή με μηδενικό μέσο όρο και μοναδιαία διασπορά  $N(0,1)$ . Εάν προσθέσουμε και την στάθμη σήματος, θα έχουμε το δείγμα που σήματος+θόρυβο που "βλέπει" ο δέκτης μέσα στο διάστημα μιας περιόδου

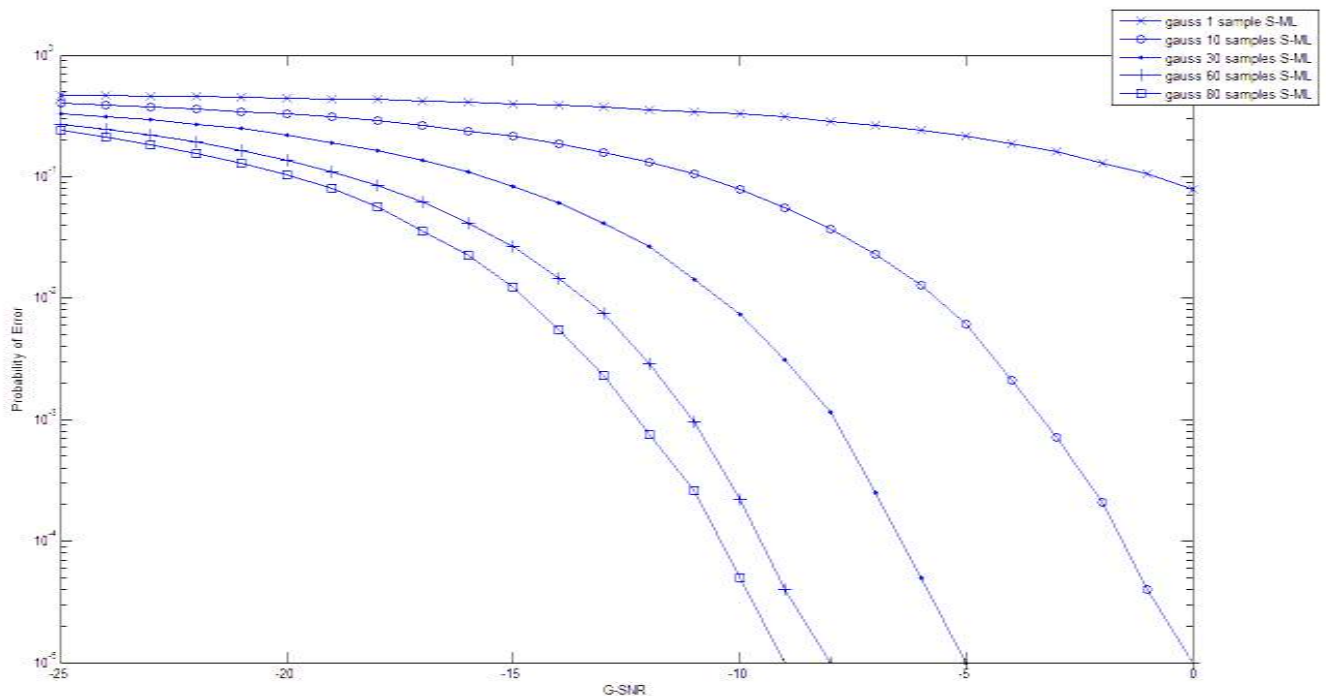
- Στην συνέχεια με την βοήθεια του παρακάτω κώδικα, υλοποιούμε το κατώφλι απόφασης:

```
sum=0;
for i=1:n
    bm(i)=bk(i);
    ginomeno_tou_kathe_loop=(ek*bk(i)-em*bm(i))*y(i);
    sum=sum+ginomeno_tou_kathe_loop;
end
```

$$D_k^S = \left\{ \underline{y} : \sum_{i=1}^n q_{k,m,i} y_i \begin{matrix} \geq \\ < \end{matrix} \begin{matrix} H_k \\ H_m \end{matrix} \right\}$$

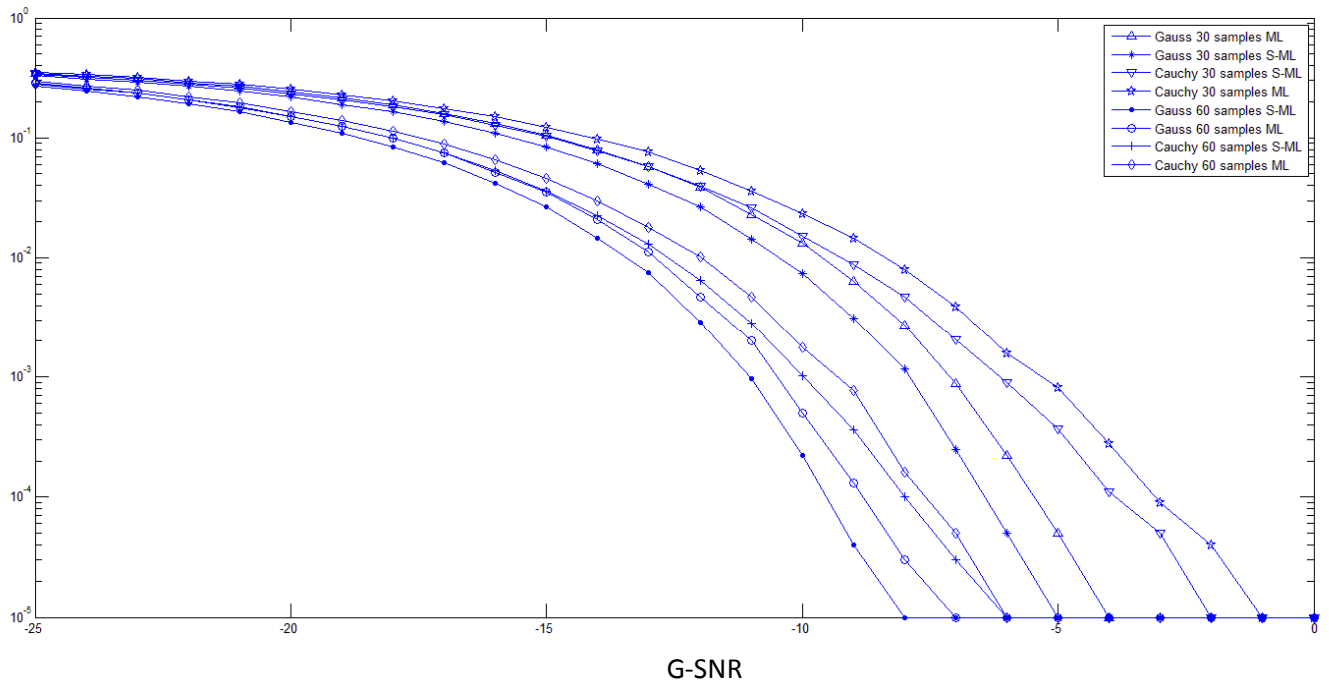
```
if (sum<0)
    decis=0;
else
    decis=1;
end
if (decis~=source)
    numoerr=numoerr+1;
end
```

Εάν το άθροισμα που θα προκύψει από τον άνω κώδικα είναι μικρότερο του μηδενός, τότε το σήμα είναι ίσο με το '0'. Στην συνέχεια συγκρίνουμε την απόφαση αυτή με το σωστό σήμα. Εάν είναι διαφορετικά, τότε αυξάνω τον μετρητή λαθών κατά 1.



Στο διάγραμμα αυτό βλέπουμε την βελτίωση που επιδέχεται η πιθανότητα λάθους, με την αύξηση της δειγματοληψίας. Στο παράδειγμα χρησιμοποιήσαμε τον S-ML εκτιμητή για την περίπτωση iid samples (μη-εξαρτημένα δείγματα)

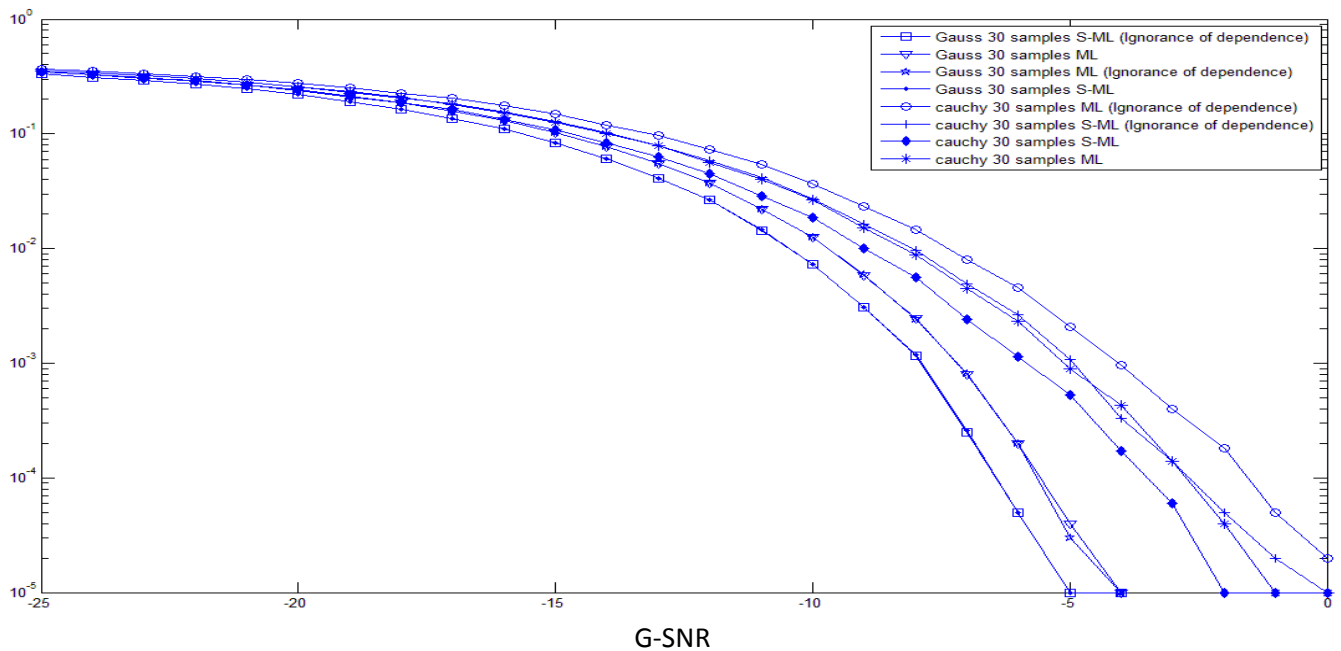
## Ανίχνευση ασθενών σημάτων σε πρώτης τάξης εξαρτημένο θόρυβο



Στο διάγραμμα αυτό βλέπουμε την επίδοση *Maximum Likelihood (ML)* και *Locally Optimal (S-ML)* εκτιμητών, όταν τα δείγματα δεν παρουσιάζουν εξάρτηση. Εκτελούμε δειγματοληψία 30 και 60 δειγμάτων χρησιμοποιώντας τους εκτιμητές *Gaussian* και *Cauchy*, σε περίπτωση *Gaussian* και *Cauchy* θορύβου αντίστοιχα ((κάθε εκτιμητής 'δουλεύει' στο κατάλληλο για αυτόν περιβάλλον θορύβου)).



## Εξαρτημένα δείγματα



Στο διάγραμμα αυτό βλέπουμε την επίδοση Maximum Likelihood (ML) και Locally Optimal (S-ML) εκτιμητών, όταν τα δείγματα παρουσιάζουν εξάρτηση κατά ποσοστό 10% (κάθε εκτιμητής 'δουλεύει' στο κατάλληλο για αυτόν περιβάλλον θορύβου). **Επίσης ο δέκτης δεν γνωρίζει την παρουσία της αλληλεξάρτησης, για να την έχει αναιρέσει.**

Παραμετροποίηση του κώδικα για την επίτευξη του cancellation στην περίπτωση εξαρτημένων δειγμάτων θορύβου

- Οι βαθμίδες που θα προστεθούν για την ακύρωση της εξάρτησης, είναι:

```

y(1)=x(1)
for i=2:n
    m=i-1;
    nn=i-1;
    for j=1:m
        y(i)=y(i)+power((rd1),j)*y(nn);
        nn=nn-1;
    end
end
end
    
```

Ο κώδικας αυτός είναι υπεύθυνος για την δημιουργία της εξάρτησης σύμφωνα με την σχέση:

$$Y_i = \sum_{j=0}^{i-1} (\rho_d)^j X_{i-j}$$

```

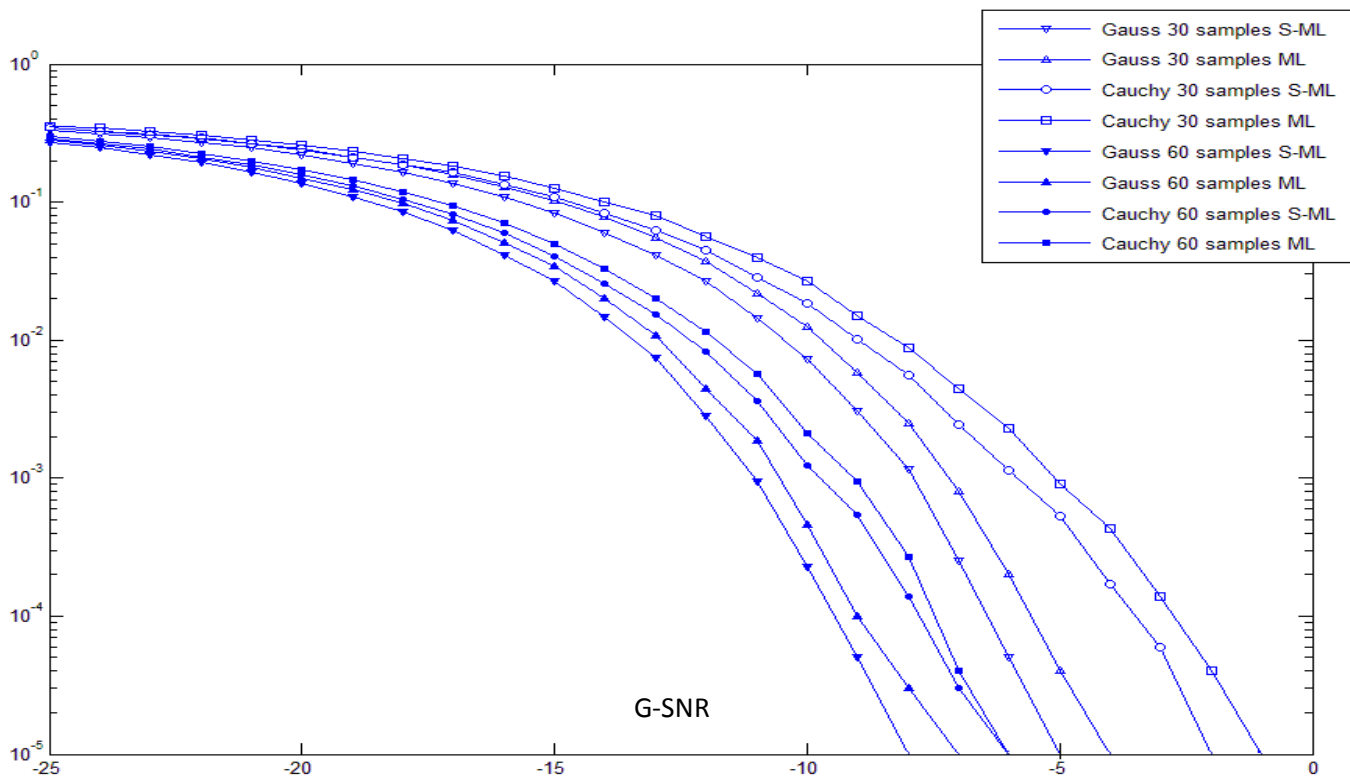
bk(1)=1/sqrt(n);
for i=2:n
  m=i-1;
  nn=i-1;
  for j=1:m
    y(i)=y(i)+power((-rd2),j)*y(nn);
    bk(i)=bk(i)+power((-rd2),j)*bk(nn);
    nn=nn-1;
  end
end
end
    
```

Επίσης στο κομμάτι αυτό του κώδικα, υλοποιούμε και την σχέση:

$$b_{k,i} = \sum_{j=0}^{i-1} (-\rho_d)^j \tilde{s}_{k,i-j}$$

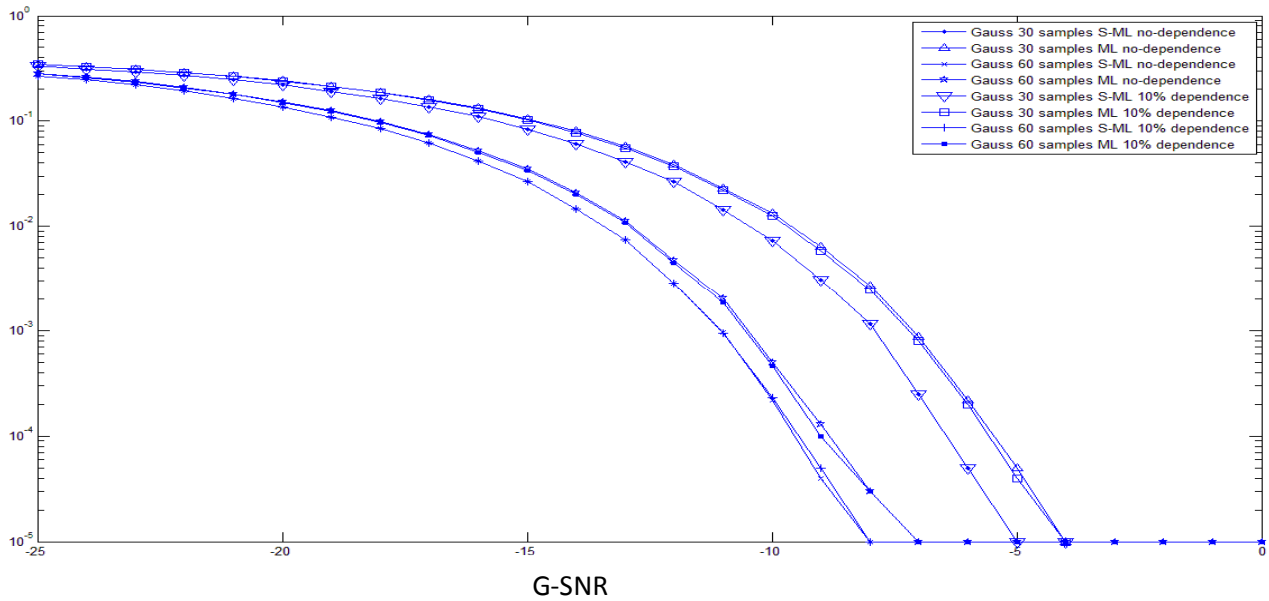
Ο κώδικας αυτός είναι υπεύθυνος για την αναίρεση της εξάρτησης σύμφωνα με την σχέση:

$$Y_i = \sum_{j=0}^{i-1} (-\rho_d)^j X_{i-j}$$

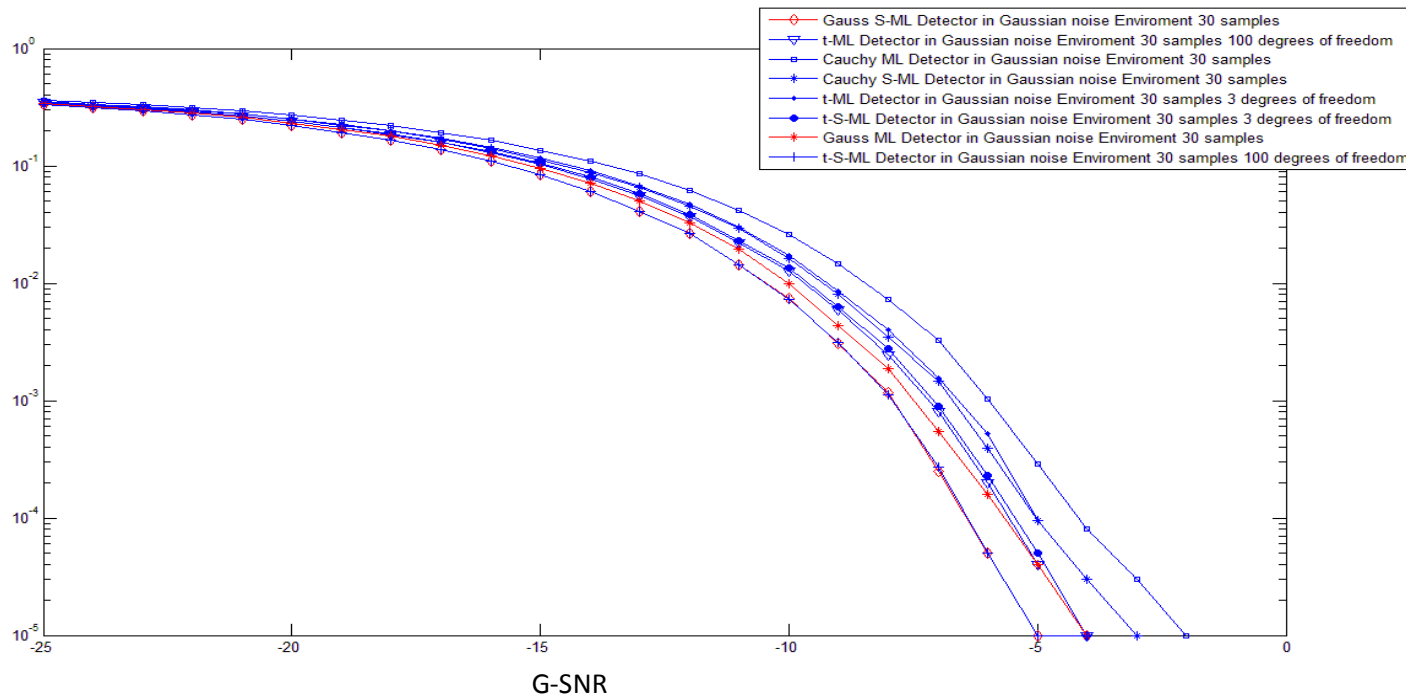


Στο διάγραμμα αυτό βλέπουμε την επίδοση Maximum Likelihood (ML) και Locally Optimal (S-ML) εκτιμητών, όταν τα δείγματα παρουσιάζουν εξάρτηση κατά ποσοστό 10% (κάθε εκτιμητής 'δουλεύει' στο κατάλληλο για αυτόν περιβάλλον θορύβου). Επίσης ο δέκτης γνωρίζει την παρουσία της αλληλεξάρτησης, και την έχει αναιρέσει.

## Ανίχνευση ασθενών σημάτων σε πρώτης τάξης εξαρτημένο θόρυβο



Στο διάγραμμα αυτό βλέπουμε την διαφορά στην απόδοση ενός δέκτη όπου τα δείγματα δεν παρουσιάζουν εξάρτηση, και ενός άλλου δέκτη στον οποίο τα δείγματα παρουσιάζουν εξάρτηση, και αυτή έχει αναιρεθεί όσο αυτό είναι δυνατό.



Στο διάγραμμα αυτό βλέπουμε την απόδοση εκτιμητών κάτω από θόρυβο ο οποίος παρουσιάζει Gaussian κατανομή και κάναμε λάθος στην επιλογή του κατάλληλου εκτιμητή. Μπορείτε να δείτε την μείωση στην απόδοση του εκτιμητή ο οποίος είναι σχεδιασμένος για Cauchy θόρυβο. Αξίζει να προσέξετε ότι εκτιμητής ο οποίος είναι σχεδιασμένος για  $t$ -distributed θόρυβο, όταν του δώσουμε μεγάλη τιμή στους βαθμούς ελευθερίας (100 βαθμούς ελευθερίας), τότε τείνει να έχει την ίδια απόδοση με τον Gaussian εκτιμητή.

## Antipodal Transmission

Τώρα που γνωρίζουμε την έννοια της *Antipodal* μετάδοσης, περιμένουμε να έχουμε μια αύξηση στην απόδοση του δέκτη, όσον αφορά την πιθανότητα λάθους. Είναι όμως έτσι; Μήπως ‘ευνοείται’ κάποιος εκτιμητής περισσότερο από τους άλλους; Ας δούμε τα διαγράμματα BER και να προσπαθήσουμε να τα ερμηνεύσουμε.

Στο διάγραμμα αυτό βλέπουμε την απόδοση της *Gaussian S-ML* εκτιμήτριας κάτω από *Gaussian* θόρυβο. Τι παρατηρείτε; Βλέπετε ότι όταν έχουμε ‘διαμορφώσει’ τον δέκτη μας έτσι ώστε να αναιρεί την εξάρτηση που υπάρχει ανάμεσα στα δείγματα, η απόδοση παραμένει η ίδια.

Η εξήγηση είναι απλή αν προσέξουμε την βασική έκφραση που διέπει τις S-ML εκτιμήτριες:

$$D_k^S = \left\{ \underline{y} : \sum_{i=1}^n (\varepsilon_k b_{k,i} - \varepsilon_m b_{m,i}) g_\Lambda(y_i) \geq 0 \quad \forall m \right\}$$

Στην περίπτωση της *Antipodal* μετάδοσης, έχουμε  $\varepsilon_k = \varepsilon_m$  και  $b_k = -b_m$ . Αντικαθιστώντας στην άνω σχέση, έχουμε:

$$\begin{aligned} D_k^S &= \left\{ \underline{y} : \sum_{i=1}^n (\varepsilon_k b_{k,i} - (-\varepsilon_k b_{k,i})) g_\Lambda(y_i) \geq 0 \quad \forall m \right\} \Rightarrow \\ D_k^S &= \left\{ \underline{y} : \sum_{i=1}^n (2\varepsilon_k b_{k,i}) g_\Lambda(y_i) \geq 0 \quad \forall m \right\} \Rightarrow \\ D_k^S &= \left\{ \underline{y} : \sum_{i=1}^n (\varepsilon_k b_{k,i}) g_\Lambda(y_i) \geq 0 \quad \forall m \right\} \end{aligned}$$

Στην περίπτωση της *Antipodal* μετάδοσης, έχουμε  $\varepsilon_m = 0$  και  $b_k = b_m$ . Αντικαθιστώντας στην βασική σχέση, έχουμε:

$$\begin{aligned} D_k^S &= \left\{ \underline{y} : \sum_{i=1}^n (\varepsilon_k b_{k,i} - 0) g_\Lambda(y_i) \geq 0 \quad \forall m \right\} \Rightarrow \\ D_k^S &= \left\{ \underline{y} : \sum_{i=1}^n (\varepsilon_k b_{k,i}) g_\Lambda(y_i) \geq 0 \quad \forall m \right\} \end{aligned}$$

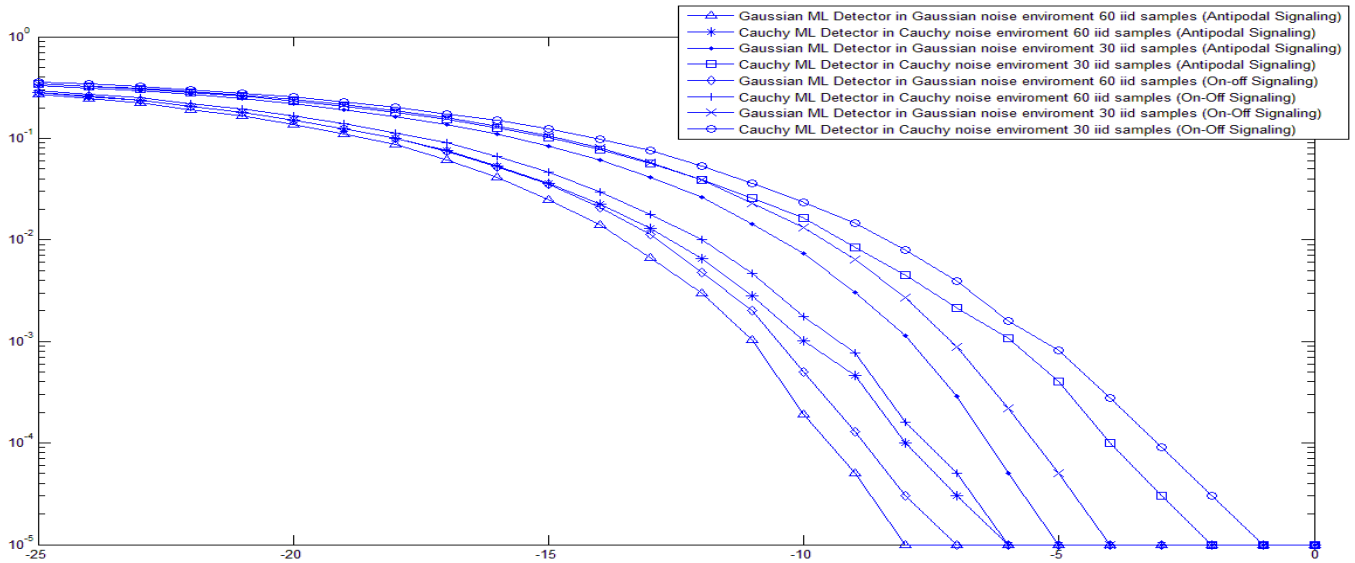
Όπως βλέπετε, η εκτιμήτρια *Gaussian S-ML* έχει την ίδια απόδοση είτε πρόκειται για *Antipodal*, είτε για *on-off* μετάδοση. Όπως θα δείτε και στα παρακάτω διαγράμματα, στην *Antipodal* μετάδοση έχουμε βελτίωση στην απόδοση της εκτιμήτριας *Maximum-Likelihood*.

```
rd1=0.1;  
rd2=0.1;  
runs=100000;  
numoerr=0;  
ek=-1;  
snr=exp(SNR_in_DB*log(10)/10);  
n=60;  
for jj=1:runs  
    E=sqrt(2*snr);  
    bk=zeros(1, n);  
    bm=zeros(1, n);  
    x=zeros(1, n);  
    y=zeros(1, n);
```

```
    em=ek;  
    if (em==-1)  
        ek=1;  
        source=1;  
    else  
        ek=-1;  
        source=0;  
    end
```

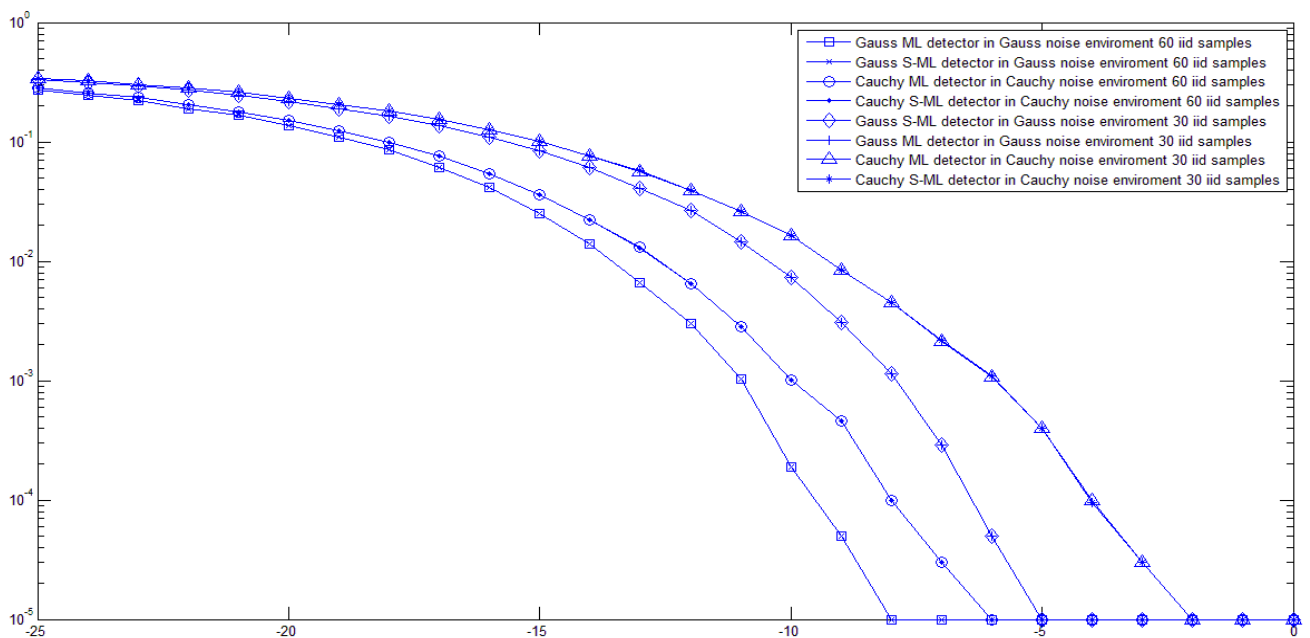
Τα κομμάτια του κώδικα τα οποία φαίνονται με κόκκινο, είναι υπεύθυνα για την Antipodal Σηματοδοσία.

## Ανίχνευση ασθενών σημάτων σε πρώτης τάξης εξαρτημένο θόρυβο



G-SNR

Στο διάγραμμα αυτό μπορείτε να δείτε την διαφορά στην απόδοση μεταξύ *On-Off* και *Antipodal* μετάδοσης. Οι *maximum-likelihood* εκτιμητές *Gaussian* και *Cauchy*, εφαρμόστηκαν σε περιβάλλον θορύβου *Gaussian* και *Cauchy* αντίστοιχα. Απεναντίας, οι *S-ML* εκτιμητές συνεχίζουν να έχουν την ίδια απόδοση, είτε στην *On-off*, είτε στην *Antipodal* μετάδοση (Αποδείχθηκε παραπάνω).



G-SNR

Στο διάγραμμα αυτό, μπορείτε να δείτε ότι στην *Antipodal* μετάδοση, η απόδοση των εκτιμητών *ML* και *S-ML* εξισώνεται! Επίσης, στο διάγραμμα αυτό, τα δείγματα είναι *i.i.d.*

## Παραμετροποίηση του Cauchy κώδικα για την σύγκριση της απόδοσης του Cauchy και Gauss εκτιμητή υπό περιβάλλον Cauchy θορύβου

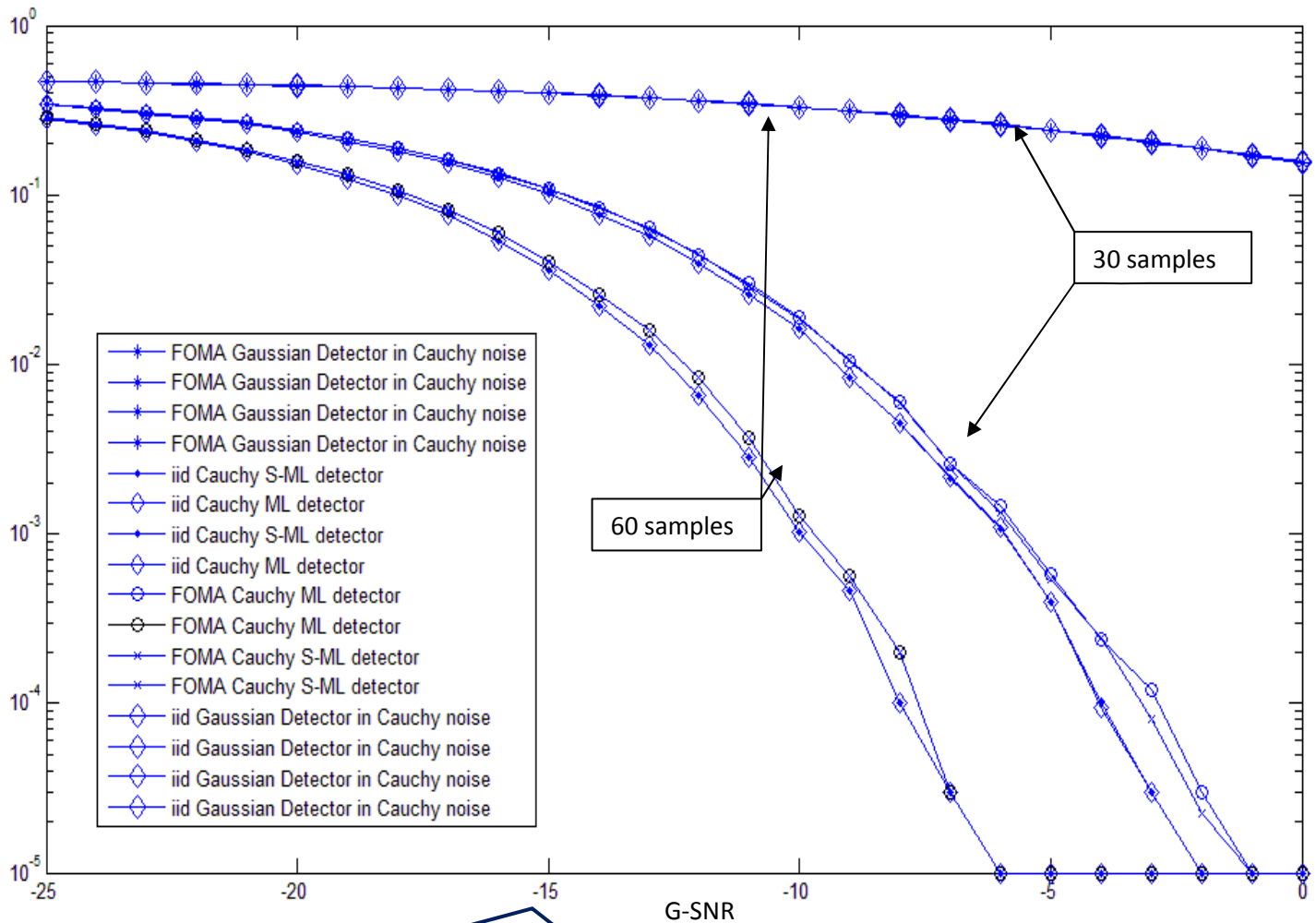
- Προσέξτε ότι το περιβάλλον θορύβου το οποίο αλλοιώνει τα σήματα μου, ακολουθεί κατανομή Cauchy. Αυτό σημαίνει ότι η παράμετρος κρουστικότητας θα πρέπει να είναι ίση με  $\alpha=1$ .
- Από τον τύπο του G-SNR, αντικαθιστούμε όπου  $\alpha=1$  και βρίσκω ότι η στάθμη σήματος (σε Volt) θα πρέπει να είναι ίση με:

$E = \sqrt{3.56 \cdot \text{snr}}$ ; Αντικαθιστούμε στην σχέση του G-SNR όπου  $\alpha=1$  και έτσι προκύπτει το 3,56

- Αυτή θα πρέπει να είναι η στάθμη σήματος που θα πρέπει να δώσω στον κώδικα μου, για το περιβάλλον Cauchy θορύβου.
- Προσοχή! Λόγω του ότι επιθυμούμε να βρούμε την απόδοση του Gauss αλλά και Cauchy εκτιμητή για Cauchy θόρυβο, θα πρέπει να βάλουμε και στις δύο περιπτώσεις την στάθμη σήματος να είναι ίση με  $E = \sqrt{3.56 \cdot \text{SNR}}$ .
- Επίσης, να μην ξεχάσουμε να αντικαταστήσουμε την γεννήτρια τυχαίων αριθμών στον Gauss εκτιμητή, από `random('norm',0,1)` → `cauchyrnd(0, 1)`
- Μετά από αυτές τις δύο αλλαγές, δεν χρειάζεται να αλλάξουμε τίποτε άλλο στον κώδικα του Gauss εκτιμητή για να δούμε την συμπεριφορά του σε Cauchy θόρυβο.

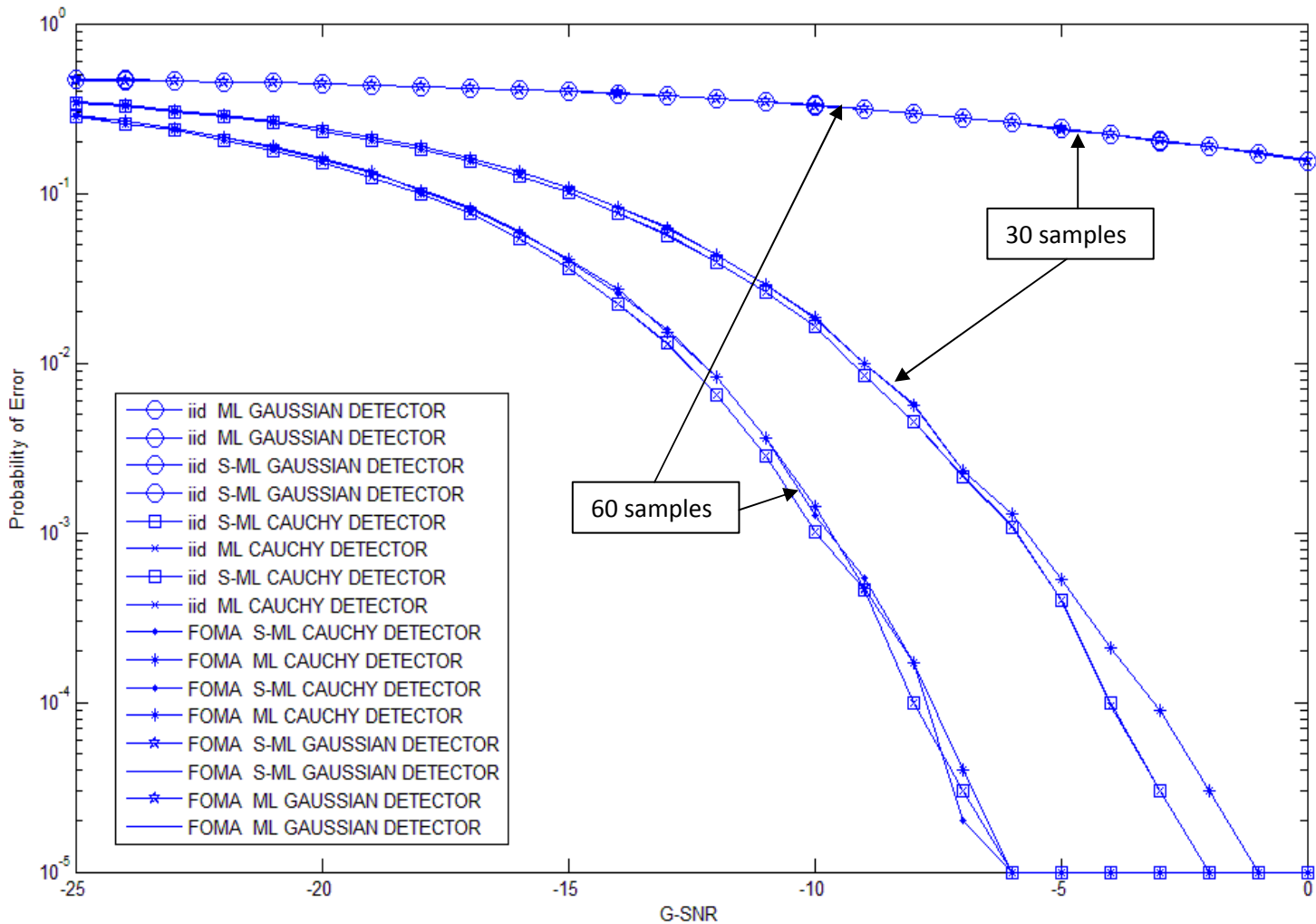
## Ανίχνευση ασθενών σημάτων σε πρώτης τάξης εξαρτημένο θόρυβο

Προσέξτε την απόδοση του Gaussian Εκτιμητή ο οποίος λειτουργεί σε Cauchy Περιβάλλον θορύβου. Η απόδοσή του, κρίνεται ως η χειρότερη δυνατή, λόγω του ότι στατιστικά ένα στα δύο δείγματα θα είναι λάθος! Ακόμα και με δειγματοληψία 60 έναντι 30 δειγμάτων, η απόδοση δεν βελτιώνεται καθόλου.

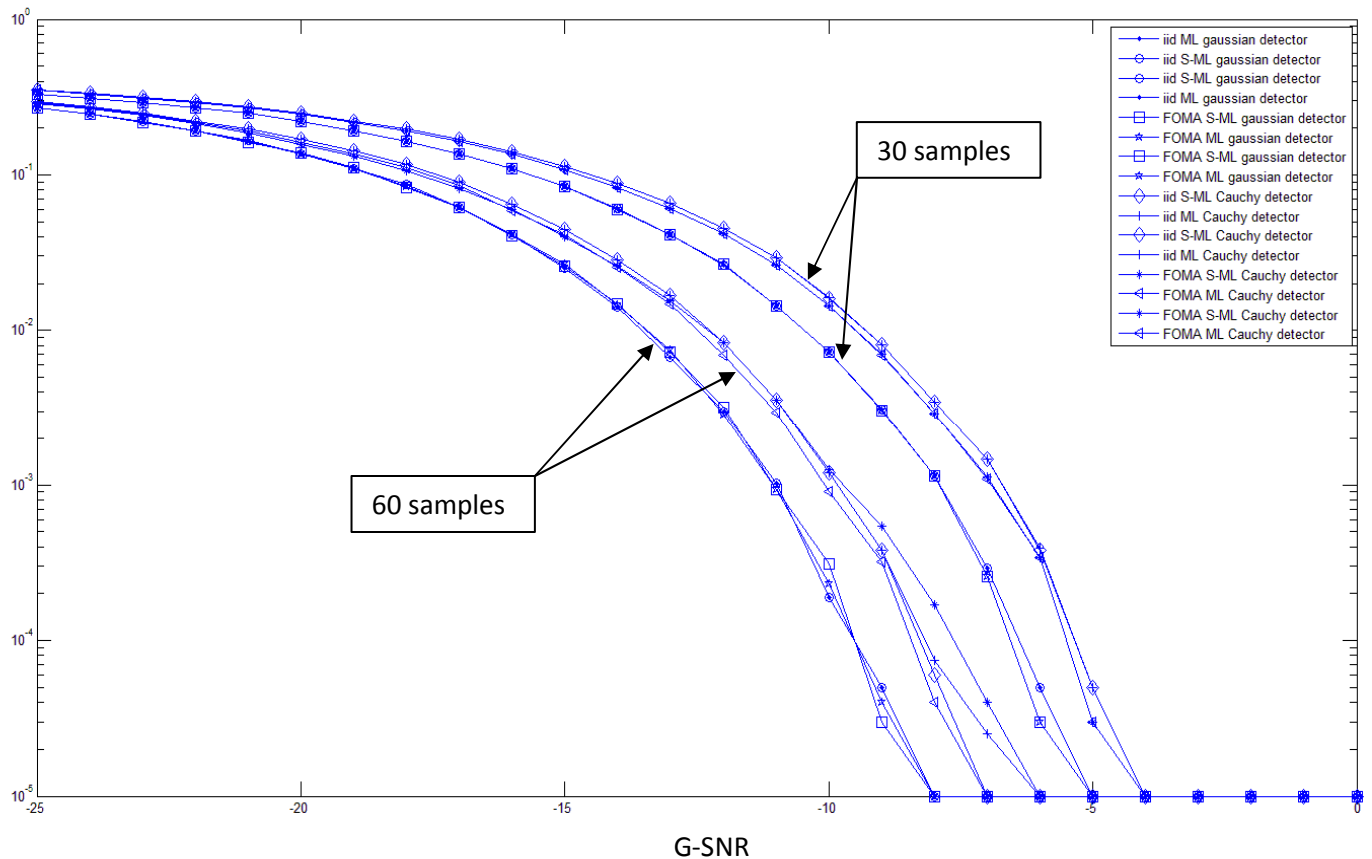


- Στο άνω διάγραμμα, βλέπουμε την απόδοση του Gaussian και Cauchy εκτιμητή για περιβάλλον θορύβου Cauchy.
- Εκτελέσαμε δειγματοληψία 30 και 60 δειγμάτων
- Όπου FOMA είναι το διάγραμμα για εξαρτημένα δείγματα θορύβου με βαθμό εξάρτησης  $\rho_d = 0.1$ , τα οποία έχουν 'υποστεί' το απαραίτητο cancellation.





- Στο άνω διάγραμμα, βλέπουμε την απόδοση του Gaussian και Cauchy εκτιμητή για περιβάλλον θορύβου Cauchy.
- Εκτελέσαμε δειγματοληψία 30 και 60 δειγμάτων
- Όπου FOMA είναι το διάγραμμα για εξαρτημένα δείγματα θορύβου με βαθμό εξάρτησης  $\rho_d = -0.1$ , τα οποία έχουν 'υποστεί' το απαραίτητο cancellation.



- Στο άνω διάγραμμα, βλέπουμε την απόδοση του Gaussian και Cauchy εκτιμητή για Gauss περιβάλλον θορύβου.
- Εκτελέσαμε δειγματοληψία 30 και 60 δειγμάτων
- Όπου FOMA είναι το διάγραμμα για εξαρτημένα δείγματα θορύβου με βαθμό εξάρτησης  $\rho_d = 0.1$ , τα οποία έχουν 'υποστεί' το απαραίτητο cancellation.

## ΠΗΓΕΣ-REFENENCES

- Nonlinear Signal Processing A Statistical Approach (Gonzalo R. Arce)
- Signal Detection and Estimation Second Edition (Mourad Barkat)
- Digital Transmission A Simulation-Aided Introduction with VisSim/Comm (Dayan Adionel Guimarães)
- Wikipedia
- Detection Schemes for Weak Signals in First-Order Moving Average of Impulsive Noise (In Jong Kim, So Ryoung Park, *Member, IEEE*, Iickho Song, *Senior Member, IEEE*, Jumi Lee, *Student Member, IEEE*, Hyoungmoon Kwon, *Student Member, IEEE*, and Seokho Yoon, *Member, IEEE*)