

Εργασία – Θεωρία Εκτίμησης και Ανίχνευσης

8ο Εξάμηνο, Θεωρία Εκτίμησης και Ανίχνευσης 2024,
Βλαχάκης Θεόδωρος, AEM 10476
Καλαμάρας Κωνσταντίνος, AEM 9716

Θέμα 1ο:

Το σύστημα μέτρησης, από το οποίο προέρχονται οι παρατηρήσεις/μετρήσεις, δηλαδή τα δεδομένα που θα αξιοποιήσουμε στη διαδικασία εκτίμησης, ακολουθεί εκθετική κατανομή με συνάρτηση πυκνότητας πιθανότητας :

$$p(x; \lambda) = \lambda e^{-\lambda x} \text{ if } x > 0, 0 \text{ else}$$

, με N να είναι το πλήθος των μετρήσεων που λαμβάνονται από το μετρητικό σύστημα, x να είναι το διάνυσμα παρατηρήσεων $x[n]$ μήκους N (δηλαδή $x = [x[0] \ x[1] \ \dots \ x[N-1]]^T$) και λ να είναι μια άγνωστη παράμετρος την οποία επιθυμούμε να εκτιμήσουμε.

Ερώτημα 1:

Σε αυτό το ερώτημα, ζητείται να υπολογίσουμε το κατώτατο όριο της διασποράς για την εκτίμηση της παραμέτρου λ , καθώς και να απεικονίσουμε το ιστόγραμμα ενός αμερόληπτου εκτιμητή που επιτυγχάνει αυτό το ζητούμενο όριο για 5000 διαφορετικά πειράματα μέτρησης με 10 μετρήσεις ανά πείραμα. Θεωρούμε την πραγματική τιμή της παραμέτρου λ να ισούται με 2 και, προφανώς, το N ισούται με 10 σε αυτό το ερώτημα. Αυτό που ψάχνουμε ουσιαστικά είναι ένας αμερόληπτος εκτιμητής, που να πληροί και τις προϋποθέσεις του CRLB, καθώς και ο υπολογισμός αυτού του κατώτατου ορίου της διασποράς του εκτιμητή που επιτάσσει το CRLB.

Ξεκινάμε υπολογίζοντας τον λογάριθμο της pdf (συνάρτηση πυκνότητας πιθανότητας) και ακολούθως την score_function για κάθε δείγμα x_i και για τιμές $x_i > 0$:

$$\ln p(x_i; \lambda) = \ln \lambda - \lambda x_i, \text{ score} = s(x_i; \lambda) = \frac{\partial(\ln p(x_i; \lambda))}{\partial(\lambda)} = -x_i + \frac{1}{\lambda}, x_i > 0.$$

Έπειτα, υπολογίζουμε την μέση τιμή του τετραγώνου της score function για κάθε δείγμα, δεδομένου πως στην εκθετική κατανομή η μέση τιμή ενός δείγματος ισούται με $\frac{1}{\lambda}$ και η διασπορά του με $\frac{1}{\lambda^2}$: $E[(s(x_i; \lambda))^2] = \text{var}(x_i) = \frac{1}{\lambda^2}$.

Έτσι, έχουμε βρεί την συνάρτηση Fisher information για ένα δείγμα: $I(\lambda) = \frac{1}{\lambda^2}$. Για τα N independent identically distributed δείγματα: $I_N(\lambda) = N I(\lambda) = \frac{N}{\lambda^2}$. Σύμφωνα με το CRLB η ελάχιστη τιμή της διασποράς του εκτιμητή ισούται με $\frac{1}{I_N(\lambda)} = \frac{\lambda^2}{N}$. Άρα για $\lambda = 2$ και $N = 10$ έχω $\text{var}(\hat{\lambda}) \geq 0,4$ για κάθε αμερόληπτο εκτιμητή $\hat{\lambda}$.

Ένας από αυτούς τους εκτιμητές είναι ο $\hat{\lambda} = \frac{n}{\sum_{i=1}^n x_i}$, όπου $\sum_{i=1}^n x_i = X$ ακολουθεί κατανομή

Gamma: $X \sim \text{Gamma}(n, \lambda)$, ο οποίος είναι αμερόληπτος: $E(\hat{\lambda}) = \lambda$, πληροί την προϋπόθεση του CRLB: $E(s(x; \lambda)) = 0$, αλλά και μπορούμε να υπολογίσουμε πως η διασπορά του εκτιμητή ισούται με: $\text{var}(\hat{\lambda}) = \frac{\lambda^2}{N}$.

‘Αρα ο εκτιμητής $\hat{\lambda}$ έχει ίση διασπορά με το όριο που θέτει το CRLB (Για $N = 10$ και $\lambda = 2$: $var(\hat{\lambda}) = 0,4$). Αυτήν την ανάλυση εκμεταλλεύομαστε και στον κώδικα μας στην matlab , δημιουργώντας διάφορες συναρτήσεις για να επιλύσουμε το πρόβλημα.

Αρχικά , δημιουργούμε την `generate_exponential_data()` , η οποία παράγει ένα σύνολο δεδομένων που ακολουθεί εκθετική κατανομή (με χρήση της ιδιότητας του αντίστροφου μετασχηματισμού μιας τυχαίας μεταβλητής U που ακολουθεί ομοιόμορφη κατανομή στο $[0,1]$ σε εκθετική X , με τον τύπο : $X = \frac{-1}{\lambda} \ln U$. (Εικόνα 1.1.1)

Στην συνέχεια , δημιουργούμε τις `compute_crlb()` και `compute_mvne()` που δέχονται ως ορίσματα τα δεδομένα και την τιμή της παραμέτρου λ και υπολογίζουν το όριο που επιτάσσει το CRLB και τον MVUE εκτιμητή , σύμφωνα με την παραπάνω ανάλυση μας για την εκθετική κατανομή. (Εικόνες 1.1.2 , 1.1.3)

Τέλος , με την ύπαρξη της `perform_mvne_experiments()` καταφέρνουμε να υπολογίσουμε τον εκτιμητή $\hat{\lambda}$ που αναφέραμε νωρίτερα και να κάνουμε την επιθυμητή προσομοίωση. Η κλήση των απαραίτητων μεθόδων με τα κατάλληλα ορίσματα (5000 πειράματα , 10 δείγματα ανά πείραμα) , μας επιτρέπει να οπτικοποιήσουμε τα αποτελέσματά μας. (Εικόνες 1.1.4 , 1.1.5 , 1.1.6)

```
function data = generate_exponential_data(num_datapoints, lambda)
% Generates i.i.d. samples from an exponential distribution using the
% inverse transform sampling method.
% Inputs:
% - num_datapoints: Number of datapoints to generate
% - lambda: The true rate parameter of the exponential distribution
% Output:
% - data: A vector of generated data points

% Generate uniform random numbers between 0 and 1
U = rand(num_datapoints, 1);

% Transform uniform random numbers to exponential random numbers
data = -log(1 - U) / lambda;
end
```

Εικόνα 1.1.1 : Συνάρτηση δημιουργίας dataset που ακολουθεί εκθετική κατανομή.

```
function crlb = compute_crlb(data, lambda)
% Computes the Cramer Rao Lower Bound
% Inputs:
% - data: A vector of observed data points from an exponential distribution
% - lambda : The true value of lambda
% Output:
% - crlb : The Cramer Rao Lower Bound

% Compute the CRLB given from  $\lambda^2 / N$ 
N = length(data);
crlb = (lambda)^2 / N;
end
```

Εικόνα 1.1.2 : Συνάρτηση υπολογισμού CRLB value.

```

function lambda_mvue = compute_mvue(data)
    % Computes an MVUE that attains the CRLB for lambda
    % Inputs:
    % - data: A vector of observed data points from an exponential distribution
    % - lambda : The true value of lambda
    % Output:
    % - lambda_mvue: The estimated rate parameter (lambda)

    % The MVUE for lambda in an exponential distribution is
    % n / sum(X)

    n = length(data);
    lambda_mvue = n / sum(data);

end

```

Εικόνα 1.1.3 : Συνάρτηση υπολογισμού MVUE.

```

function mvue_results = perform_mvue_experiments(num_experiments, num_datapoints, lambda)
    % Performs multiple experiments and computes an MVUE for each experiment
    % Inputs:
    % - num_experiments: Number of experiments to perform
    % - num_datapoints: Number of datapoints per experiment
    % - lambda: The true rate parameter of the exponential distribution
    % Output:
    % - mvue_results: A vector of MVUE estimates for each experiment

    mvue_results = zeros(num_experiments, 1);

    for i = 1:num_experiments
        data = generate_exponential_data(num_datapoints, lambda);
        mvue_results(i) = compute_mvue(data);
    end
end

```

Εικόνα 1.1.4 : Συνάρτηση πραγμάτωσης της επιθυμητής προσομοίωσης.

```

% Main script to run the entire process
lambda = 2;
number_datapoints = 10;
number_experiments = 5000;

% Generate a random dataset for the CRLB calculation
data = generate_exponential_data(number_datapoints, lambda);

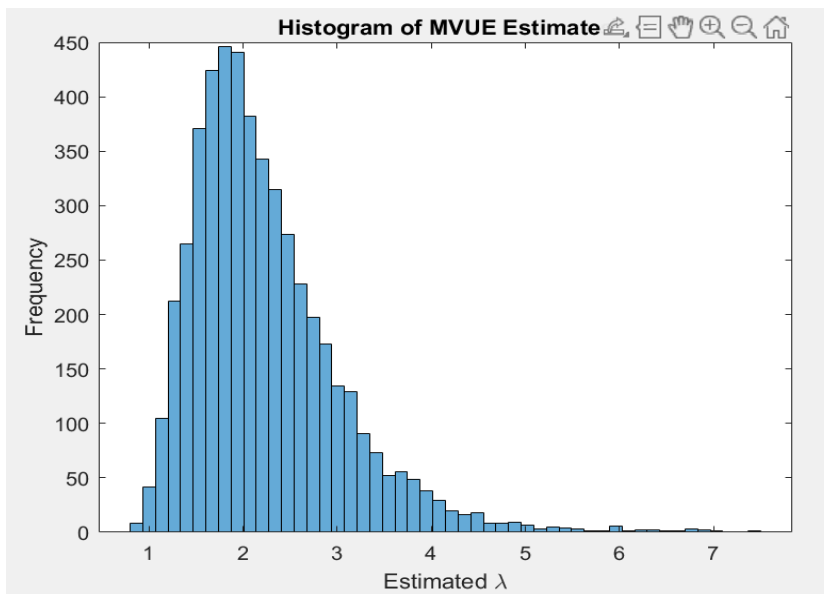
% Calculate and display the CRLB
crlb = compute_crlb(data, lambda);
disp(['CRLB at lambda = ', num2str(lambda), ': ', num2str(crlb)]);

% Compute the MVUE using the function
mvue = perform_mvue_experiments(number_experiments, number_datapoints, lambda);

% Plot histogram of MVUE estimates
figure;
histogram(mvue, 50);
title('Histogram of MVUE Estimates');
xlabel('Estimated \lambda');
ylabel('Frequency');

```

Εικόνα 1.1.5 : Ορισμός των απαραίτητων μεταβλητών και κλήση των υλοποιημένων συναρτήσεων για έλεγχο λειτουργίας του κώδικα.



CRLB at lambda = 2: 0.4

Εικόνα 1.1.6 : Οπτικοποίηση των αποτελεσμάτων.

Ερώτημα 2 :

Στο δεύτερο ερώτημα του Θέματος , καλούμαστε να υπολογίσουμε τον εκτιμητή μέγιστης πιθανοφάνειας για την εκθετική κατανομή της εκφώνησης , καθώς και η οπτικοποίηση του μέσω ενός ιστογράμματος.

Ξεκινάμε την ανάλυσή μας ορίζοντας την συνάρτηση πιθανοφάνειας για την εκθετική κατανομή , δεδομένου πως έχουμε i.i.d δείγματα πλήθους N :
$$L(\lambda) = \prod_{i=1}^N f(x_i; \lambda) = \prod_{i=1}^N \lambda e^{-\lambda x_i}$$

Για την εύρεση του MLE , πρέπει ουσιαστικά να βρούμε την λογαριθμική συνάρτηση πιθανοφάνειας και να την μεγιστοποιήσουμε , μηδενίζοντας την παράγωγό της :

$$\log L(\lambda) = N \log(\lambda) - \lambda \sum_{i=1}^N x_i, \frac{\partial \log L(\lambda)}{\partial (\lambda)} = \frac{N}{\lambda} - \sum_{i=1}^N x_i = 0.$$

‘Αρα , αν λύσουμε ως προς λ προκύπτει ο MLE εκτιμητής : $\hat{\lambda} = \frac{N}{\sum_{i=1}^N x_i} = \frac{1}{\bar{X}}$, μια

παραλλαγή δηλαδή του εκτιμητή του προηγούμενου ερωτήματος.

Στον κωδικά μας , αρχικά , έχουμε χρησιμοποιήσει την ίδια λογική με το προηγούμενο ερώτημα για την παραγωγή του dataset (συνάρτηση `generate_exponential_data()`) . Στην συνέχεια , έχουμε δημιουργήσει την συνάρτηση `compute_mle()` για τον υπολογισμό του εκτιμητή $\hat{\lambda}$ (Εικόνα 1.2.1) και την συνάρτηση `perform_mle_experiments()` για να κάνουμε αυτόν τον υπολογισμό για τις απαιτούμενες συνθήκες της προσομοίωσης μας (Εικόνα 1.2.2). Τέλος , προβάλλουμε τα αποτελέσματά μας μέσω ενός ιστογράμματος (Εικόνα 1.2.3).

```

function lambda_mle = compute_mle(data)
    % Computes the Maximum Likelihood Estimator (MLE) for lambda
    % Inputs:
    % - data: A vector of observed data points from an exponential distribution
    % Output:
    % - lambda_mle: The estimated rate parameter (lambda)

    % The MLE for lambda in an exponential distribution is 1 / mean(data)
    lambda_mle = 1 / mean(data);
end

```

Εικόνα 1.2.1 : Συνάρτηση υπολογισμού MLE για κάποιο dataset.

```

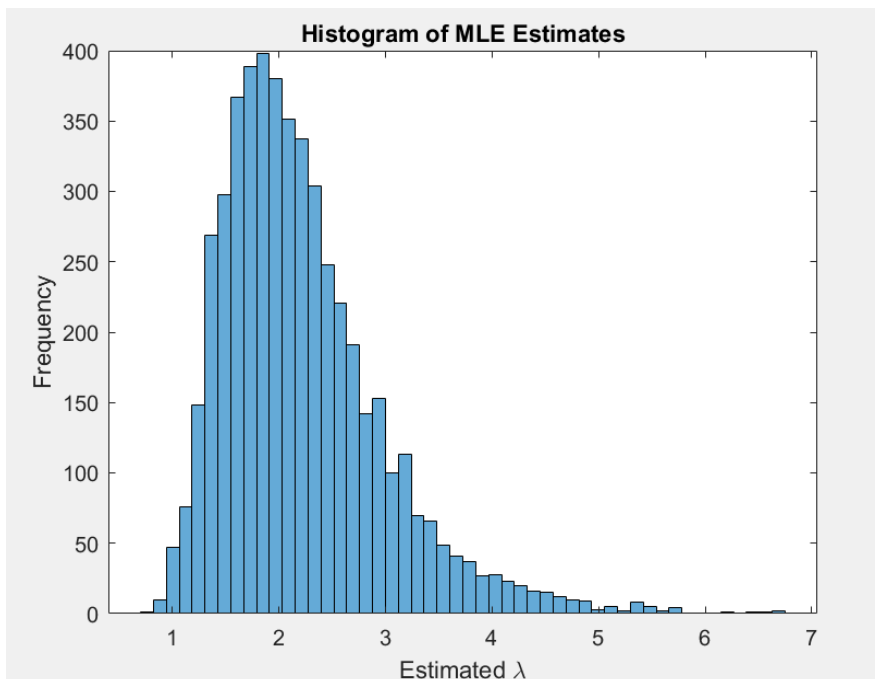
function mle_results = perform_mle_experiments(num_experiments, num_datapoints, lambda)
    % Performs multiple experiments and computes MLE for each experiment
    % Inputs:
    % - num_experiments: Number of experiments to perform
    % - num_datapoints: Number of datapoints per experiment
    % - lambda: The true rate parameter of the exponential distribution
    % Output:
    % - mle_results: A vector of MLE estimates for each experiment

    mle_results = zeros(num_experiments, 1);

    for i = 1:num_experiments
        data = generate_exponential_data(num_datapoints, lambda);
        mle_results(i) = compute_mle(data);
    end
end

```

Εικόνα 1.2.2 : Συνάρτηση υπολογισμού MLE για συγκεκριμένο αριθμό πειραμάτων.



Εικόνα 1.2.3 : Οπτικοποίηση των αποτελεσμάτων.

Παρατηρήσεις :

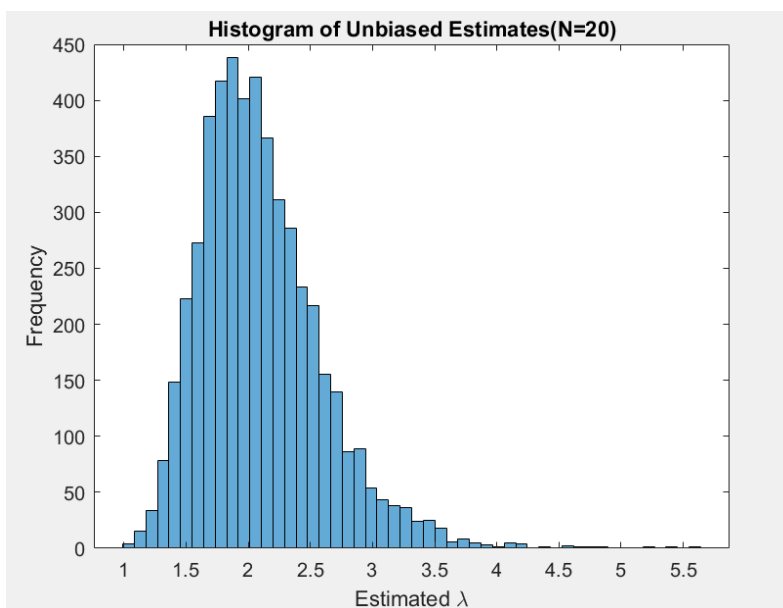
Έχοντας πραγματοποιήσει τις προσομοιώσεις και των δύο πρώτων ερωτημάτων , παρατηρούμε πως το ιστόγραμμα του MVUE εκτιμητή του 1ου ερωτήματος αντικατοπτρίζει την αμεροληψία του , καθώς το ιστόγραμμα το του 2ου εκτιμητή που είναι πολωμένος έχει κάποια μετατόπιση από την πραγματική τιμή του λ στις κορυφές του , κάτι που συμπληρώνεται και από την διαφορετική στατιστική συμπεριφορά των δύο εκτιμητών (όσων αφορά την αναμενόμενη τους τιμή) :

$$\text{Για τον MVUE } \hat{\lambda} = \frac{n}{\sum_{i=1}^n x_i} : E(\hat{\lambda}) = \lambda \text{ και } var(\hat{\lambda}) = \frac{\lambda^2}{N}.$$

$$\text{Αντίστοιχα για τον MLE } \hat{\lambda} = \frac{1}{\bar{X}} : E(\hat{\lambda}) = \lambda \frac{N}{N-1} \text{ (biased) και } var(\hat{\lambda}) = \frac{\lambda^2}{N}.$$

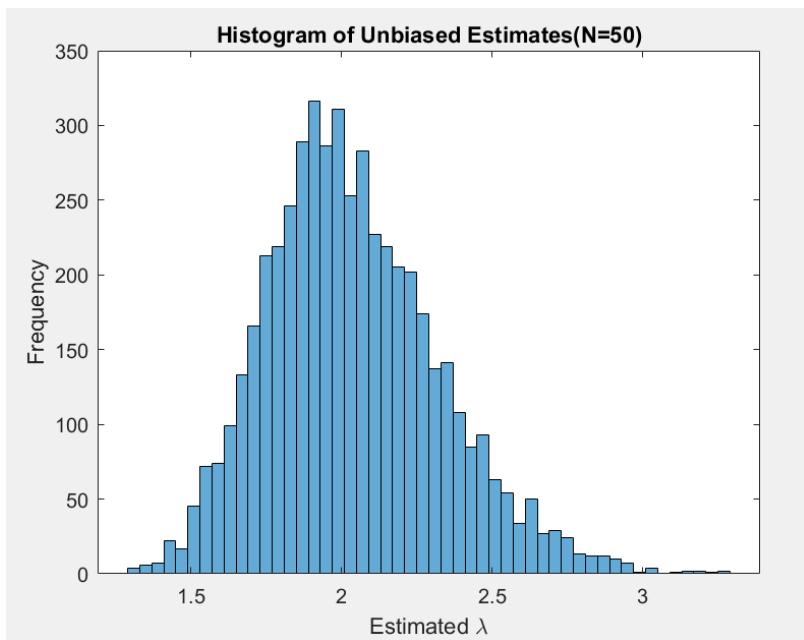
Ερώτημα 3 :

Για τον εκτιμητή του πρώτου ερωτήματος :



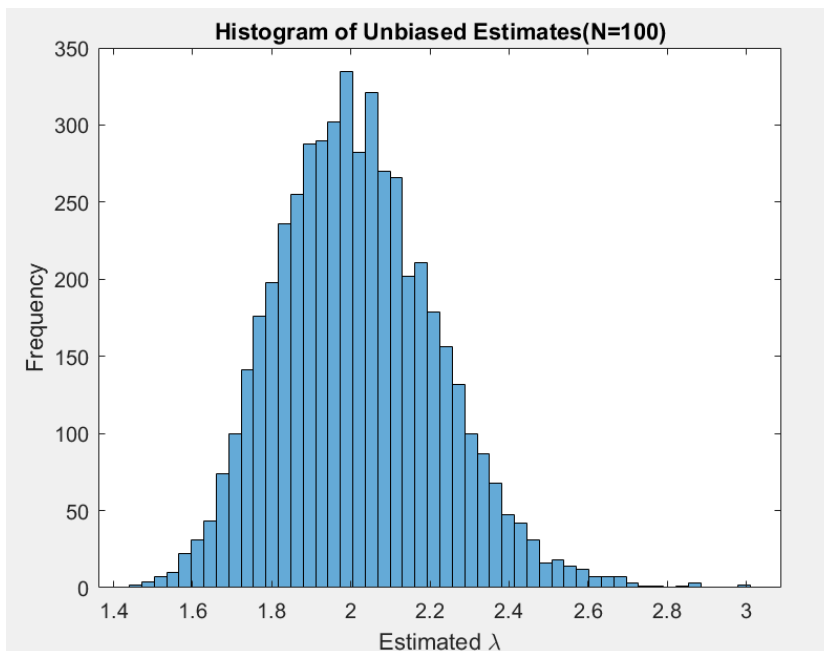
```
>> Q1_3a  
CRLB at lambda = 2: 0.2
```

Εικόνα 1.3.1 : Οπτικοποίηση των αποτελεσμάτων για 20 δείγματα ανά πείραμα.



```
>> Q1_3a
CRLB at lambda = 2: 0.08
```

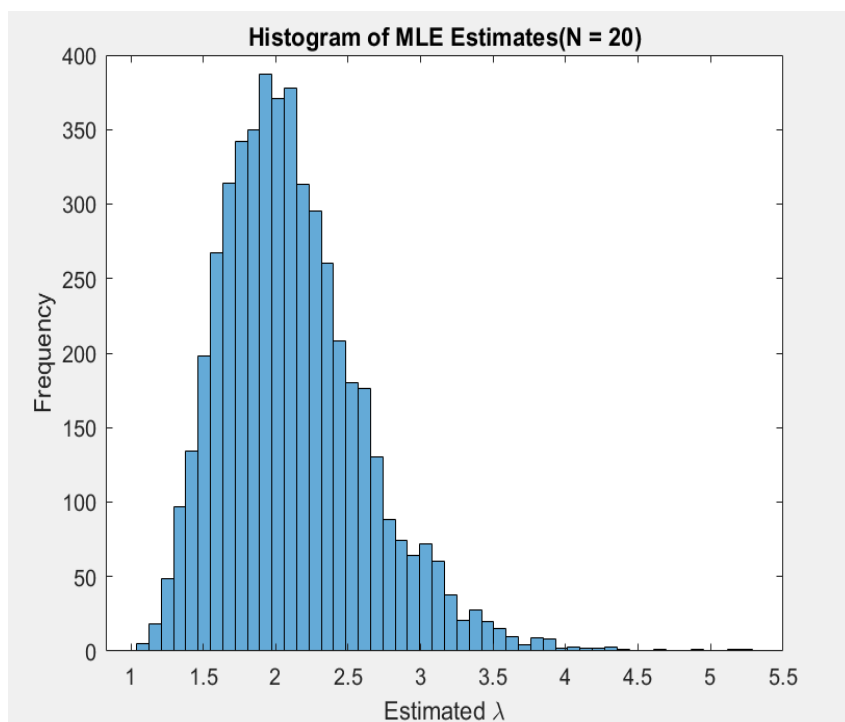
Εικόνα 1.3.2 : Οπτικοποίηση των αποτελεσμάτων για 50 δείγματα ανά πείραμα.



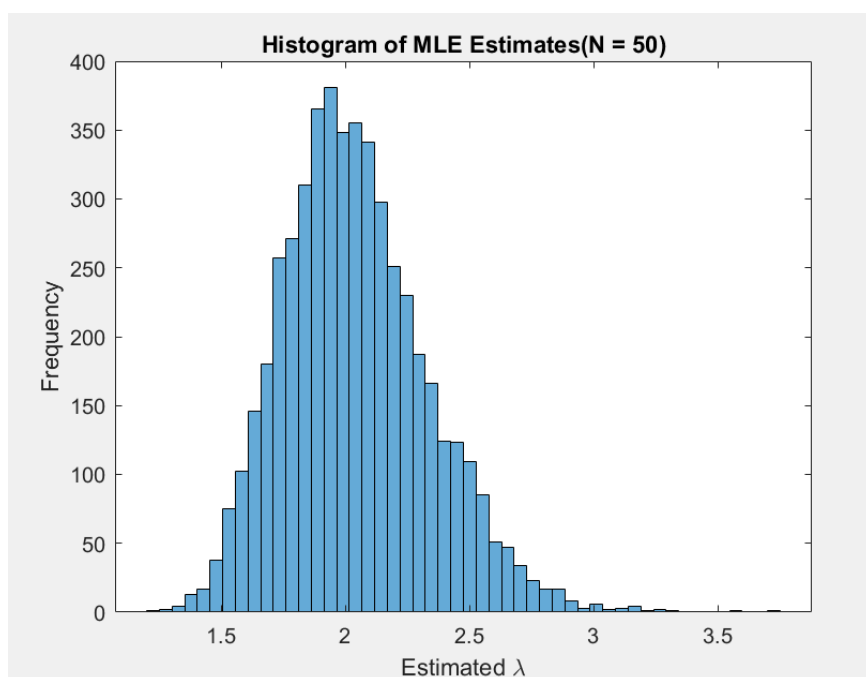
```
>> Q1_3a
CRLB at lambda = 2: 0.04
```

Εικόνα 1.3.3 : Οπτικοποίηση των αποτελεσμάτων για 100 δείγματα ανά πείραμα.

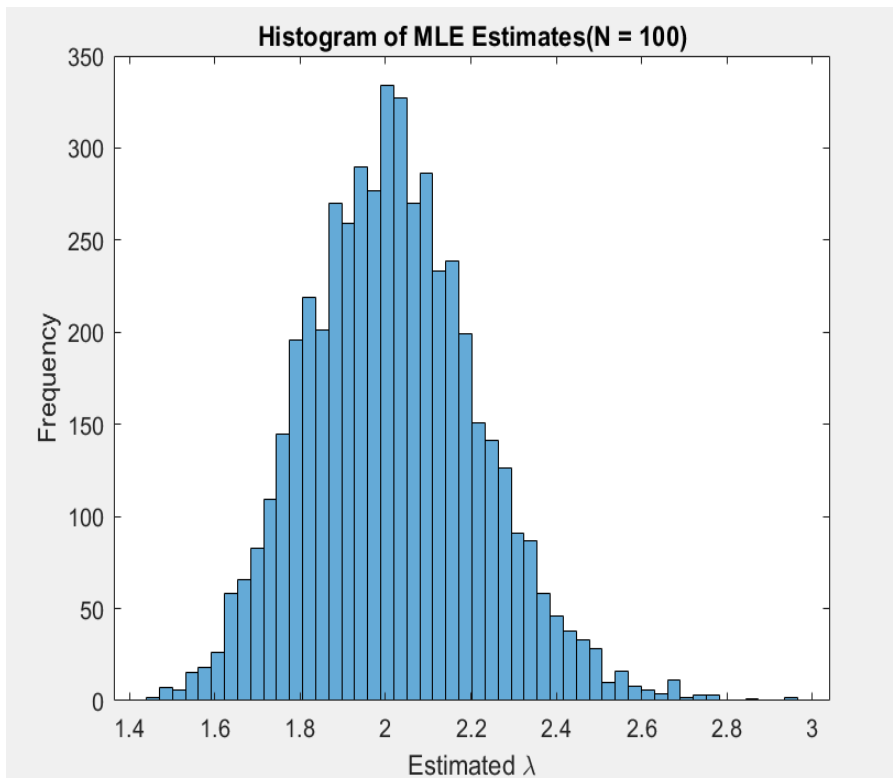
Για τον εκτιμητή του δεύτερου ερωτήματος :



Εικόνα 1.3.4 : Οπτικοποίηση των αποτελεσμάτων του MLE για 20 δείγματα ανά πείραμα.



Εικόνα 1.3.5 : Οπτικοποίηση των αποτελεσμάτων του MLE για 50 δείγματα ανά πείραμα.



Εικόνα 1.3.6 : Οπτικοποίηση των αποτελεσμάτων του MLE για 100 δείγματα ανά πείραμα.

Παρατηρήσεις :

Αρχικά , όσον αφορά τον αμερόληπτο εκτιμητή του πρώτου ερωτήματος , μέσω της εξίσωσης υπολογισμού της διασποράς του : $var(\hat{\lambda}) = \frac{\lambda^2}{N}$, παρατηρούμε πως αύξηση στην τιμή του N προκαλεί μείωση στην διασπορά και άρα έχουμε μια κατανομή με μεγαλύτερη πυκνότητα γύρω από την πραγματική τιμή της παραμέτρου λ ($\lambda = 2$). Έτσι , ο εκτιμητής μας γίνεται πιο ακριβής και συνεπώς και πιο αποτελεσματικός.

Τώρα , όσον αφορά τον MLE εκτιμητή του δεύτερου ερωτήματος , μπορεί επίσης να γίνει εύκολα εμφανές πως η σταδιακή αύξηση του αριθμού των δειγμάτων ανά πείραμα , παράγει ένα όλο και πυκνότερο ιστόγραμμα γύρω από την πραγματική τιμή του λ ($\lambda = 2$) , κάτι που σημαίνει πως η εκτίμηση της παραμέτρου λ γίνεται όλο και πιο ακριβής. Επιπλέον , παρατηρούμε πως το ιστόγραμμα γίνεται και όλο και πιο συμμετρικό , καθώς σύμφωνα με το Κεντρικό Οριακό θεώρημα η κατανομή του MLE προσεγγίζει κανονική κατανομή με κέντρο την πραγματική τιμή της παραμέτρου λ .

Θέμα 2ο:

Το σύστημα μέτρησης, από το οποίο προέρχονται οι παρατηρήσεις/μετρήσεις, δηλαδή τα δεδομένα που θα αξιοποιήσουμε στη διαδικασία εκτίμησης, περιγράφεται ως εξής:

$$x[n] = h\theta + w[n], n=0,1,...,N-1 \quad (2)$$

όπου N είναι το πλήθος των μετρήσεων που λαμβάνονται από το μετρητικό σύστημα, x είναι το διάνυσμα παρατηρήσεων $x[n]$ μήκους N (δηλαδή $x = [x[0] \ x[1] \ \dots \ x[N-1]]^T$), w είναι Γκαουσιανός

θόρυβος με μηδενική μέση τιμή και διασπορά $\sigma^2=4$, h είναι μια γνωστή παράμετρος του συστήματος και θ είναι μια άγνωστη παράμετρος (βαθμωτή) την οποία επιθυμούμε να εκτιμήσουμε.

Στο εξής, χωρίς περιορισμό της γενικότητας, θα θεωρούμε για όλα τα παρακάτω ερωτήματα ότι η γνωστή παράμετρος h έχει τιμή $h=0.5$.

Ακολουθώντας θα εξετάσουμε για την παράμετρο θ , την οποία επιθυμούμε να εκτιμήσουμε κάθε φορά, διαφορετικές περιπτώσεις, δηλαδή διαφορετικές “φύσεις” από τις οποίες μπορεί αυτή να προέρχεται και το πώς αυτό μπορεί να επιδράσει στην ποιότητα ενός εκτιμητή για αυτή την παράμετρο.

Σε κάθε περίπτωση, το μετρητικό σύστημα από το οποίο λαμβάνονται τα δεδομένα, περιγράφεται από τη σχέση (2), δηλαδή πρόκειται για ένα Γενικό Γραμμικό Μοντέλο (General Linear Model), το οποίο αποτελεί το σύστημα μέτρησης. Η σχέση (2) μπορεί, επίσης, να γραφεί και στη μορφή πινάκων, η οποία περιγράφει το Γενικό Γραμμικό Μοντέλο και θα είναι χρήσιμη στην παρακάτω ανάλυση για τον ορισμό διάφορων εκτιμητών.

$$x = H\theta + w \quad (2.1),$$

όπου x είναι το x είναι το διάνυσμα παρατηρήσεων $x[n]$ μήκους N (δηλαδή $x=[x[0] \ x[1] \ \dots \ x[N-1]]^T$), w είναι Γκαουσιανός θόρυβος (διάνυσμα Γκαουσιανού θορύβου) με μηδενική μέση τιμή και διασπορά $\sigma^2=4$ (δηλαδή $w[n] \sim N(0, \sigma^2)$) και τότε ο πίνακας συμμεταβλητότητας για το διάνυσμα του Γκαουσιανού θορύβου w , αφού Γκαουσιανός θόρυβος σημαίνει ανεξάρτητες και άρα και ασυσχέτιστες παρατηρήσεις θορύβου $w[n]$, είναι: $C_w = \sigma^2 I_{N \times N}$ όπου $I_{N \times N}$ είναι ο μοναδιαίος $N \times N$ πίνακας), H είναι ένας γνωστός πίνακας “μετρήσεων” $N \times 1$, όπως λέγεται, δηλαδή ένα διάνυσμα μήκους N στη συγκεκριμένη περίπτωση που ορίζεται ως $H=[h \ h \ \dots \ h]^T = [0.5 \ 0.5 \ \dots \ 0.5]^T = 0.5 \mathbf{1}_{N \times 1}$ (όπου $\mathbf{1}_{N \times 1}$ είναι το διάνυσμα μήκους N , του οποίου όλα τα στοιχεία είναι 1) και θ είναι μια άγνωστη παράμετρος (βαθμωτή) την οποία επιθυμούμε να εκτιμήσουμε.

Ερώτημα 1: Σε αυτή την περίπτωση υποθέτουμε ότι η άγνωστη παράμετρος θ , που θέλουμε να εκτιμήσουμε, είναι μια γκαουσιανή τυχαία μεταβλητή με μέση τιμή $\mu=4$ και διασπορά $\sigma_0^2=1$. Στην πραγματικότητα, δηλαδή, υποθέτουμε ότι η άγνωστη παράμετρος θ , στο μετρητικό σύστημα, είναι η πραγματοποίηση μια τυχαίας μεταβλητής Θ , η οποία είναι γκαουσιανή με τα παραπάνω χαρακτηριστικά μ και σ_0^2 όσον αφορά τις δύο πρώτες ροπές, δηλαδή $\Theta \sim N(\mu, \sigma_0^2)$. Λαμβάνοντας υπόψη τα δεδομένα (διάνυσμα x) από το μετρητικό σύστημα, θέλουμε να εκτιμήσουμε την παράμετρο θ με τον εκτιμητή MMSE για αυτή την παράμετρο. Τότε, για να βρούμε τον MMSE, θα επικαλεστούμε την Bayesian θεωρία και προσέγγιση για την εκτίμηση της άγνωστης παραμέτρου.

Όπως παρατηρούμε, το σύστημα μέτρησης περιγράφεται από το Γενικό Γραμμικό μοντέλο της σχέσης (2) ή (2.1), αντίστοιχα, όπως διατυπώσαμε παραπάνω. Ακόμη, ο θόρυβος w είναι Γκαουσιανός, ενώ και η άγνωστη παράμετρος θ , σ' αυτή την περίπτωση, θεωρήθηκε σαν την πραγματοποίηση μιας γκαουσιανής τυχαίας μεταβλητής. Επειδή, λοιπόν, θ και w είναι gaussian και μάλιστα μπορούμε να υποθέσουμε ανεξαρτησία μεταξύ της παραμέτρου θ και του θορύβου w -αφού δεν έχουμε κάποιο δεδομένο ή κάποια ένδειξη στο μοντέλο του συστήματος που να υποδηλώνει κάποια συσχέτιση μεταξύ θ και w - τότε, χρησιμοποιώντας την Bayesian προσέγγιση για την εκτίμηση της θ , το μετρητικό σύστημα περιγράφεται από το Bayesian ισοδύναμο του Γενικού Γραμμικού Μοντέλου (δηλαδή από το Bayesian General Linear Model) και αφού υποθέσαμε $w \sim N(0, C_w)$ και $\Theta \sim N(\mu, \sigma_0^2)$, από το Γενικό Γραμμικό Γκαουσιανό Μοντέλο (General Linear Gaussian Model).

Υποθέσαμε ότι η θ είναι γκαουσιανή τυχαία μεταβλητή με μέση τιμή $\mu=4$ και διασπορά $\sigma_0^2=1$. Αυτή η πρότερη γνώση για την θ περιγράφεται μέσω μιας prior pdf $p(\theta)$, η οποία είναι μια pdf gaussian τυχαίας μεταβλητής με τα παραπάνω αυτά χαρακτηριστικά. Στη Bayesian προσέγγιση εκμεταλλευόμαστε αυτή την πρότερη γνώση για την θ , δηλαδή την prior pdf $p(\theta)$, για τον υπολογισμό της posterior pdf $p(\theta|x)$, η οποία ορίζει την κατανομή που εκτιμούμε για την

παράμετρο θ , εφόσον έχουμε συλλέξει τα δεδομένα $x[0], x[1], \dots, x[N-1]$. Σύμφωνα με τη Bayesian προσέγγιση και το Bayesian General Linear Model, αφού θ και w είναι gaussian, τότε και η posterior pdf $p(\theta|x)$ θα είναι επίσης gaussian, δηλαδή ορίζει μια κανονική κατανομή για την θ , αφού έχουν συλλεγεί τα δεδομένα από το σύστημα μέτρησης, με μέση τιμή και διασπορά που δίνονται από τις παρακάτω σχέσεις:

$$E[\theta|x] = \mu + \sigma_0^2 H^T (H \sigma_0^2 H^T + C_w)^{-1} (x - H\mu)$$

$C_{\theta|x} = \sigma_{\theta|x}^2 = \sigma_0^2 - \sigma_0^2 H^T (H \sigma_0^2 H^T + C_w)^{-1} H \sigma_0^2$ (ο πίνακας συμμεταβλητότητας της θ , μετά τη συλλογή των δεδομένων, που ορίζει το Bayesian General Linear Model είναι στην ουσία η διασπορά της θ , μετά τη συλλογή των δεδομένων, αφού η θ είναι βαθμωτή).

ή
$$E[\theta|x] = \mu + (1/\sigma_0^2 + H^T C_w^{-1} H)^{-1} H^T C_w^{-1} (x - H\mu)$$

$$C_{\theta|x} = \sigma_{\theta|x}^2 = (1/\sigma_0^2 + H^T C_w^{-1} H)^{-1}$$

Σύμφωνα με τη Bayesian θεωρία και προσέγγιση, ο MMSE ορίζεται ως η μέση τιμή της posterior pdf $p(\theta|x)$ για την παράμετρο θ , έπειτα από τη συλλογή των δεδομένων από το σύστημα. Τότε:

$$\theta_{MMSE}^{\wedge} = E[\theta|x] = \mu + \sigma_0^2 H^T (H \sigma_0^2 H^T + C_w)^{-1} (x - H\mu)$$

ή
$$\theta_{MMSE}^{\wedge} = E[\theta|x] = \mu + (1/\sigma_0^2 + H^T C_w^{-1} H)^{-1} H^T C_w^{-1} (x - H\mu),$$

με την απόδοση του εκτιμητή να “μετριέται” κατά μια έννοια μέσω του Bmse (Bayesian Mean Squared Error), το οποίο ορίζεται, σύμφωνα με το General Linear Gaussian Model, ως:

$$Bmse = M_{\theta^{\wedge}} = C_{\theta|x} = \sigma_{\theta|x}^2 = \sigma_0^2 - \sigma_0^2 H^T (H \sigma_0^2 H^T + C_w)^{-1} H \sigma_0^2$$

ή
$$Bmse = M_{\theta^{\wedge}} = C_{\theta|x} = \sigma_{\theta|x}^2 = (1/\sigma_0^2 + H^T C_w^{-1} H)^{-1}$$

λόγω των παραπάνω σχέσεων από το Bayesian General Linear Model.

Αυτές τις σχέσεις εκμεταλλευόμαστε και στην ανάλυση στο Matlab, προκειμένου να ορίσουμε τον MMSE εκτιμητή για την θ και το “μέτρο” Bmse που είναι ένας δείκτης της απόδοσης του εκτιμητή. Στον κώδικα στο Matlab, έχουμε υλοποιήσει και υπολογίσει τον MMSE για μια πραγματοποίηση της θ , που αντλείται όπως είπαμε από κανονική κατανομή, υπολογίζοντας επίσης το “μέτρο” για αυτόν τον εκτιμητή. Τα αποτελέσματα από τον υπολογισμό του MMSE για την θ , παρουσιάζονται στον Πίνακα 2.1.1. Σημειώνουμε, ότι το πλήθος των παρατηρήσεων που λαμβάνουμε στο πείραμα, που εκτελέσαμε για το σύστημα μέτρησης της (2), είναι 50 (δηλαδή $N=50$).

```
----- MMSE for  $\theta$  / Results -----
----- Measurement System:  $x[n]=h*\theta+w[n]$ ,  $n=0,1,\dots,N-1$  -----

Known parameter  $h=0.5$  . Number of observations taken:  $N=50$ 
Take  $\theta$  from distribution  $N(4,1)$ . The actual value of the parameter  $\theta$ :  $\theta=2.7859$ 
The MMSE estimator using Bayesian Approach for General Gaussian Linear Model:  $\theta_{MMSE}=3.7495$ 
The Bmse:  $Bmse=0.24242$ 
```

Πίνακας 2.1.1

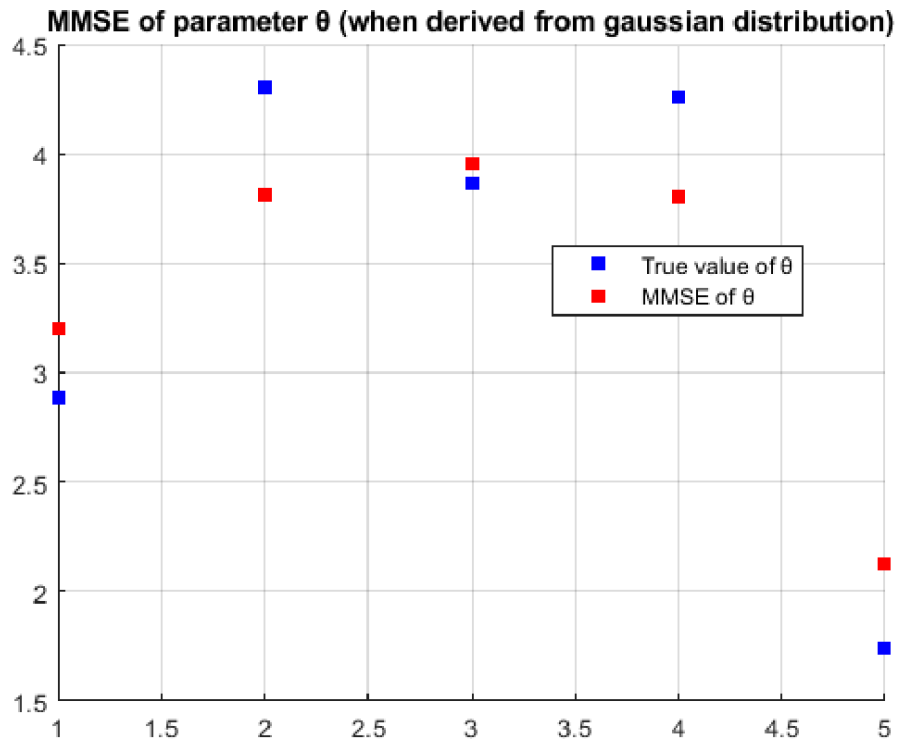
Για να ελέγξουμε, λίγο περισσότερο, την συμπεριφορά αυτού του εκτιμητή, υλοποιήθηκαν, εν συνεχεία, 5 τέτοια πειράματα άντλησης της θ από κανονική κατανομή, παραγωγής θορύβου w και τελικά άντλησης των δεδομένων που συνθέτουν το διάνυσμα μήκους x σύμφωνα με τη σχέση (2) ή (2.1), που περιγράφει το μετρητικό σύστημα. Για καθένα από τα 5 πειράματα εκτιμήθηκε η παράμετρος θ με τον MMSE και υπολογίστηκε και το Bmse. Τα αποτελέσματα φαίνονται στον Πίνακα 2.1.2 και γραφικά στο Σχήμα 2.1.1. Σημειώνουμε, ότι σε καθένα από τα πειράματα το πλήθος των παρατηρήσεων που λαμβάνουμε είναι 50.

```

---- Experiment number 1 ----
Actual value of the parameter  $\theta$ :  $\theta=2.8865$ 
The MMSE estimator :  $\theta_{MMSE}=3.2006$ 
The Bmse: Bmse=0.24242
---- Experiment number 2 ----
Actual value of the parameter  $\theta$ :  $\theta=4.3035$ 
The MMSE estimator :  $\theta_{MMSE}=3.8163$ 
The Bmse: Bmse=0.24242
---- Experiment number 3 ----
Actual value of the parameter  $\theta$ :  $\theta=3.8668$ 
The MMSE estimator :  $\theta_{MMSE}=3.9559$ 
The Bmse: Bmse=0.24242
---- Experiment number 4 ----
Actual value of the parameter  $\theta$ :  $\theta=4.2614$ 
The MMSE estimator :  $\theta_{MMSE}=3.806$ 
The Bmse: Bmse=0.24242
---- Experiment number 5 ----
Actual value of the parameter  $\theta$ :  $\theta=1.7416$ 
The MMSE estimator :  $\theta_{MMSE}=2.1233$ 
The Bmse: Bmse=0.24242

```

Πίνακας 2.1.2



Σχήμα 2.1.1 : 5 πειράματα εκτίμησης της παραμέτρου θ με τον MMSE εκτιμητή, σύμφωνα με τα δεδομένα (50 το πλήθος για κάθε πείραμα) που αντλούνται από το σύστημα της (2).

Όπως παρατηρούμε στα αποτελέσματα που παρουσιάζονται στους παραπάνω πίνακες, αλλά ακόμη πιο κατανοητά στο διάγραμμα με τις πραγματικές τιμές της θ και τον MMSE εκτιμητή για αυτή, η εκτίμηση κάθε φορά δεν φαίνεται να είναι ιδανική, όμως φαίνεται να είναι καλή. Φυσικά, αυτό είναι κάτι σχετικό και δεν μπορούμε να το ενστερνιστούμε, καθώς δεν γνωρίζουμε τι ακριβώς περιγράφει το σύστημα μέτρησης της (2) (σαν φυσικό/πραγματικό σύστημα, δηλαδή) και άρα ποιες είναι οι απαιτήσεις ακρίβειας για την παράμετρο θ . Εκείνο που δίνει ένα μέτρο της ποιότητας της απόδοσης του MMSE, για αυτό το σύστημα μέτρησης και αυτή την πρότερη γνώση για την θ , είναι το Bmse, το οποίο όπως παρατηρούμε έχει τιμή 0.24242 .

Ερώτημα 2: Σε αυτή την περίπτωση υποθέτουμε ότι η άγνωστη παράμετρος θ , που θέλουμε να εκτιμήσουμε, είναι ντετερμινιστική (αλλά άγνωστη) παράμετρος. Τώρα, επιθυμούμε -λαμβάνοντας τα δεδομένα/μετρήσεις από το σύστημα μέτρησης και αξιοποιώντας τα- να εκτιμήσουμε την παράμετρο θ με τον εκτιμητή μέγιστης πιθανοφάνειας, δηλαδή τον MLE (Maximum Likelihood Estimator). Σε αυτήν την ανάλυση, δεν διαθέτουμε κάποια πρότερη γνώση που να περιγράφεται από κάποια prior pdf για την άγνωστη παράμετρο θ , αφού η θ είναι τώρα ντετερμινιστική. Η παράμετρος θ είναι, δηλαδή, κάποια σταθερή τιμή, η οποία αφορά το μετρητικό σύστημα και δεν αντλείται από κάποια κατανομή. Για να μοντελοποιήσουμε, λοιπόν, αυτή την περίπτωση στο Matlab, θέτουμε στην παράμετρο θ , αυθαίρετα, κάποια τιμή (έστω $\theta=4$) και ακολουθούμε τη διαδικασία υπολογισμού του MLE για αυτή την τιμή της παραμέτρου θ , δηλαδή για την θ , όπως θα περιγραφεί παρακάτω.

Όπως αναφέρθηκε στην αρχή, το σύστημα μέτρησης περιγράφεται από το Γενικό Γραμμικό Μοντέλο, που δηλώνουν οι σχέσεις (2) και (2.1), σε μορφή πινάκων. Αυτή η γνώση για το σύστημα μέτρησης είναι ανεξάρτητη από τη “φύση” της παραμέτρου θ και την αξιοποιούμε για τον ορισμό των εκτιμητών στις διάφορες περιπτώσεις που μελετούνται.

Στο σύστημα μέτρησης ο θόρυβος $w[n]$ (σε κάθε παρατήρηση $x[n]$ από το σύστημα) $\forall n=0,1,\dots,N-1$ είναι Γκαουσιανός με μηδενική μέση τιμή και διασπορά $\sigma^2=4$. Τα δείγματα $w[n]$ δηλαδή του θορύβου w είναι iid και ακολουθούν κανονική κατανομή με μέση τιμή 0 και διασπορά 4.

Παρατηρούμε ότι για το σύστημα μέτρησης που μελετάμε, λόγω της (2) $\forall n$ μπορούμε να γράψουμε: $w[n] = x[n] - h\theta$ (2.2), όπου η h είναι κάποια γνωστή παράμετρος ($h=0.5$) και σε αυτή τη δεύτερη περίπτωση, που μελετάμε, η θ είναι ντετερμινιστική.

Τότε, η pdf $p(w[n])$ είναι η pdf μιας gaussian κατανομής που ορίζει το $w[n]$ ως γκαουσιανή τυχαία μεταβλητή, $\forall n$. Αξιοποιώντας τη σχέση (2.2) θα είναι: $p(w[n]) = p(x[n] - h\theta)$ και αυτή η συνάρτηση θα ορίζει πλέον την pdf $p(x[n] | \theta)$, αφού η θ είναι ντετερμινιστική. Τότε και κάθε δείγμα $x[n]$, $\forall n$, που λαμβάνουμε είναι gaussian, ακολουθεί δηλαδή κανονική κατανομή με μέση τιμή $h\theta = 0.5\theta (= 0.5 \cdot 4 = 2$ με την υπόθεση μας ότι $\theta=4$ για να μοντελοποιήσουμε την περίπτωση αυτή στο Matlab) και διασπορά $\sigma^2=4$. Οπότε η pdf $p(x[n] | \theta)$ είναι γνωστή και εύκολα μπορεί να προκύψει η pdf του θορύβου $p(w)$ -αφού τα δείγματα του θορύβου είναι iid- και η pdf $p(x|\theta)$ των δεδομένων παραμετροποιημένη ως προς θ , κάνοντας την ίδια υπόθεση για iid “δείγματα”.

Αν υποθέσουμε τώρα σταθεροποιημένο/συγκεκριμένο διάνυσμα παρατηρήσεων x και άγνωστη την παράμετρο θ , όπως δηλαδή συμβαίνει μελετώντας τα δεδομένα που αντλήθηκαν από το σύστημα μέτρησης, η $p(x|\theta)$ ορίζει τη συνάρτηση πιθανοφάνειας της παραμέτρου θ . Όπως γνωρίζουμε από τη θεωρία εκτίμησης, ο MLE είναι αυτός που ορίζεται σαν την τιμή της θ που μεγιστοποιεί τη συνάρτηση πιθανοφάνειας της παραμέτρου θ . Συνεπώς, αφού η pdf $p(x|\theta)$ των δεδομένων παραμετροποιημένη ως προς θ είναι gaussian λόγω της φύσης του προβλήματος που μελετάμε και μπορεί να οριστεί εύκολα, αφού είναι μια γνωστή pdf μιας γνωστής gaussian κατανομής, μπορούμε να ορίσουμε εξίσου εύκολα τη συνάρτηση πιθανοφάνειας της παραμέτρου θ $p(x|\theta)$, μετά την λήψη των παρατηρήσεων. Έπειτα, μπορούμε να υλοποιήσουμε στο Matlab όλες τις απαραίτητες μαθηματικές πράξεις (λογαρίθμηση της $p(x|\theta)$, έπειτα παραγωγή της $\ln[p(x|\theta)]$ και εξίσωση της παραγώγου με το 0 και τέλος υπολογισμό της τιμής θ που μηδενίζει αυτή την παράγωγο) προκειμένου να υπολογίσουμε τον MLE για την θ .

Ωστόσο,, εκμεταλλευόμαστε το γεγονός ότι το σύστημα μέτρησης περιγράφεται από το Γενικό Γραμμικό Μοντέλο της (2) και (2.1), καθώς η θεωρία εκτίμησης δίνει κλειστούς τύπους για τον MLE και τη διασπορά του (ο MLE είναι μια τυχαία μεταβλητή), η οποία μπορεί να αξιοποιηθεί σαν ένα “μέτρο” για την ποιότητα του εκτιμητή, όπως είχαμε αντίστοιχα το Bmse στο προηγούμενο ερώτημα. Για το Γενικό Γραμμικό Μοντέλο, το οποίο περιγράφει το σύστημα μέτρησης που μελετάμε και ορίζεται μέσω της (2.1), με το διάνυσμα του θορύβου $w \sim N(0, C_w)$, ο MLE ορίζεται ως εξής:

$$\theta_{MLE} = (H^T C_w^{-1} H)^{-1} H^T C_w^{-1} x$$

και είναι μια τυχαία μεταβλητή, η οποία -σύμφωνα με το Θεώρημα που ορίζει τον MLE μιας

παραμέτρου (ή διανύσματος παραμέτρων στη γενική περίπτωση) συστήματος που περιγράφεται από το Γενικό Γραμμικό Μοντέλο, στο οποίο ο θόρυβος είναι Γκαουσιανός- ακολουθεί κανονική κατανομή με μέση τιμή θ , δηλαδή την τιμή της παραμέτρου και διασπορά που δίνεται από τη σχέση:

$$\sigma_{\theta_{MLE}}^2 = (H^T C_w^{-1} H)^{-1}$$

Δηλαδή: $\theta_{MLE} \sim N(\theta, (H^T C_w^{-1} H)^{-1})$

Σημείωση: Από τη θεωρία εκτίμησης γνωρίζουμε ότι αυτός ο εκτιμητής, όταν το σύστημα περιγράφεται από το Γενικό Γραμμικό Μοντέλο, είναι ο MVU και είναι και efficient.

Αξιοποιώντας αυτές τις σχέσεις στην ανάλυση στο Matlab, ορίζουμε τον MLE εκτιμητή για την θ και τη διασπορά του $\sigma_{\theta_{MLE}}^2$ (ως var_MLE_theta2 στο Matlab), η οποία όπως αναφέραμε ορίζει τη διακύμανση του MLE για την θ γύρω από τη μέση τιμή του, η οποία είναι ίση με θ και συνεπώς η διασπορά $\sigma_{\theta_{MLE}}^2$ ορίζει τη διακύμανση (το τετράγωνο της τυπικής απόκλισης) από την πραγματική τιμή της παραμέτρου θ . Άρα η διασπορά του MLE μπορεί να αξιοποιηθεί, όντως, σαν ένα “μέτρο” της ποιότητας απόδοσης του εκτιμητή.

Στον κώδικα στο Matlab, έχουμε υλοποιήσει και υπολογίσει τον MLE για την παράμετρο θ θεωρώντας την σε αυτή την περίπτωση ντετερμινιστική και δίνοντας της την τιμή 4, χάρην της μοντελοποίησης του όλου προβλήματος, υπολογίζοντας επίσης τη διασπορά του εκτιμητή αυτού. Τα αποτελέσματα από τον υπολογισμό του MLE για την θ , παρουσιάζονται στον Πίνακα 2.2.1 . Σημειώνουμε, ότι το πλήθος των παρατηρήσεων που λαμβάνουμε στο πείραμα, που εκτελέσαμε για το σύστημα μέτρησης της (2), είναι 50 (δηλαδή $N=50$).

```
----- MLE for  $\theta$  / Results -----
----- Measurement System:  $x[n]=h*\theta+w[n]$ ,  $n=0,1,\dots,N-1$  -----

Known parameter  $h=0.5$  . Number of observations taken:  $N=50$ 
The unknown parameter  $\theta$  is deterministic. So the value of  $\theta$  is setted arbitrarily:  $\theta=4$ 
The MLE estimator using the MLE Approach for General Linear Model:  $\theta_{MLE}=3.6805$ 
The variance of the MLE for estimation of  $\theta$ :  $\text{Var}(\theta_{MLE})=0.32$ 
```

Πίνακας 2.2.1

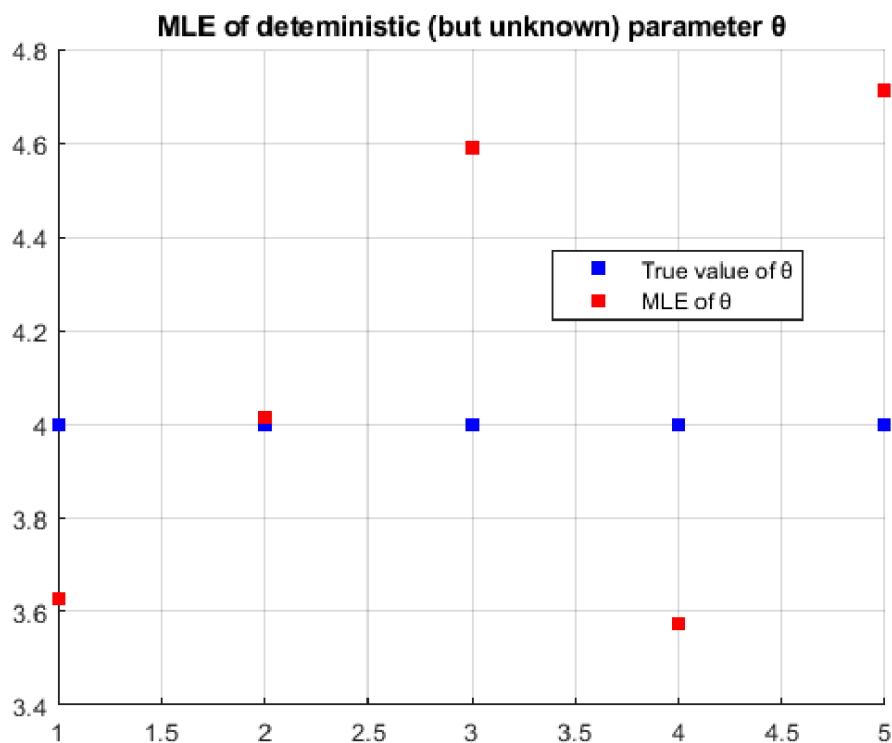
Για να ελέγξουμε, λίγο περισσότερο, την συμπεριφορά αυτού του εκτιμητή, υλοποιήθηκαν, εν συνεχεία, 5 τέτοια πειράματα, όπως παραπάνω, με ορισμό της θ ντετερμινιστικής στην τιμή 4, παραγωγής θορύβου w και τελικά άντλησης των δεδομένων που συνθέτουν το διάνυσμα μήκους x σύμφωνα με τη σχέση (2) ή (2.1), που περιγράφει το μετρητικό σύστημα. Για καθένα από τα 5 πειράματα εκτιμήθηκε η παράμετρος θ με τον MLE και υπολογίστηκε και η διασπορά του MLE. Τα αποτελέσματα φαίνονται στον Πίνακα 2.2.2 και γραφικά στο Σχήμα 2.2.1 . Σημειώνουμε, ότι σε καθένα από τα πειράματα το πλήθος των παρατηρήσεων που λαμβάνουμε είναι 50.


```

---- Experiment number 1 ----
Actual value of the parameter  $\theta$ :  $\theta=4$ 
The MLE estimator :  $\theta\_MLE=3.6276$ 
The variance of the MLE:  $\text{Var}(\theta\_MLE)=0.32$ 
---- Experiment number 2 ----
Actual value of the parameter  $\theta$ :  $\theta=4$ 
The MLE estimator :  $\theta\_MLE=4.0123$ 
The variance of the MLE:  $\text{Var}(\theta\_MLE)=0.32$ 
---- Experiment number 3 ----
Actual value of the parameter  $\theta$ :  $\theta=4$ 
The MLE estimator :  $\theta\_MLE=4.5897$ 
The variance of the MLE:  $\text{Var}(\theta\_MLE)=0.32$ 
---- Experiment number 4 ----
Actual value of the parameter  $\theta$ :  $\theta=4$ 
The MLE estimator :  $\theta\_MLE=3.5747$ 
The variance of the MLE:  $\text{Var}(\theta\_MLE)=0.32$ 
---- Experiment number 5 ----
Actual value of the parameter  $\theta$ :  $\theta=4$ 
The MLE estimator :  $\theta\_MLE=4.7112$ 
The variance of the MLE:  $\text{Var}(\theta\_MLE)=0.32$ 

```

Πίνακας 2.2.2



Σχήμα 2.2.1 : 5 πειράματα εκτίμησης της παραμέτρου θ με τον MLE εκτιμητή, σύμφωνα με τα δεδομένα (50 το πλήθος για κάθε πείραμα) που αντλούνται από το σύστημα της (2).

Όπως παρατηρούμε στα αποτελέσματα που παρουσιάζονται στους παραπάνω πίνακες, αλλά ακόμη πιο κατανοητά στο διάγραμμα με τις πραγματικές τιμές της θ και τον MLE εκτιμητή για αυτή, η εκτίμηση κάθε φορά δεν φαίνεται να είναι ιδανική, όμως φαίνεται να είναι καλή. Φυσικά, αυτό είναι κάτι σχετικό και δεν μπορούμε να το ενστερνιστούμε, καθώς δεν γνωρίζουμε τι ακριβώς περιγράφει το σύστημα μέτρησης της (2) (σαν φυσικό/πραγματικό σύστημα, δηλαδή) και άρα ποιες είναι οι απαιτήσεις ακρίβειας για την παράμετρο θ . Εκείνο που δίνει ένα μέτρο της ποιότητας της απόδοσης του MLE, για αυτό το σύστημα μέτρησης και την άγνωστη παράμετρο θ να υποθετείται ντετερμινιστική, είναι η διασπορά του MLE, η οποία όπως παρατηρούμε έχει τιμή 0.32 . Δεν θα ήταν χρήσιμο να συγκρίνουμε αυτόν τον εκτιμητή με τον MMSE του πρώτου ερωτήματος, καθώς τα 5 πειράματα άντλησης δεδομένων από το σύστημα και εκτίμησης της θ είναι, απλώς,

ενδεικτικά για την κάθε περίπτωση εκτιμητή και επιπλέον αυτό που υποθέσαμε σαν το “μέτρο” της απόδοσης των δύο εκτιμητών είναι μια διαφορετική ποσότητα ($Bmse$ στην περίπτωση του MMSE και $Var(\theta_{MLE})$ στην περίπτωση του MLE) στη θεωρία εκτίμησης, άρα δεν οδηγεί σε κάποιο συμπέρασμα σύγκρισης.

Ερώτημα 3: Σε αυτό το ερώτημα, θα εξετάσουμε για τους δύο εκτιμητές, που ορίσαμε παραπάνω για την παράμετρο θ , στις διαφορετικές περιπτώσεις, δηλαδή για τον MMSE και τον MLE αντίστοιχα, τη σύγκλιση του Μέσου Τετράγωνου Σφάλματος (Mean Squared Error – MSE) ως συνάρτηση του αριθμού των μετρήσεων (πλήθος παρατηρήσεων N), που λαμβάνονται ανά πείραμα μέτρησης (πείραμα συλλογής δεδομένων για εκτίμηση της θ κάθε φορά με κάποιον εκτιμητή). Για να μελετήσουμε αυτή τη σύγκλιση του Mean Squared Error, θα κάνουμε το γράφημα του MSE ως συνάρτηση του αριθμού των παρατηρήσεων που λαμβάνονται ανά πείραμα για 500 πειράματα μέτρησης και μέγιστο αριθμό μετρήσεων//παρατηρήσεων ανά πείραμα 50, εκτελώντας ολόκληρη τη διαδικασία και για τους εκτιμητές (MMSE και MLE αντίστοιχα) για την άγνωστη παράμετρο θ του συστήματος μέτρησης.

Προκειμένου να εξετάσουμε τη σύγκλιση του MSE κάποιου από τους δύο εκτιμητές, κάθε φορά, θα εκτελέσουμε 500 πειράματα μέτρησης μεταβάλλοντας τον αριθμό των παρατηρήσεων που λαμβάνονται ανά πείραμα από 1 έως 50 κατά την εκτέλεσή αυτών των πειραμάτων. Ένα πείραμα μέτρησης είναι, ουσιαστικά, η εκτέλεση της διαδικασίας συλλογής ενός πλήθους δεδομένων $x[n]$ από το μετρητικό σύστημα που περιγράφεται από την (2): $x[n] = h\theta + w[n]$, $n=0,1,\dots,N-1$, όπου N είναι ακριβώς αυτός ο αριθμός των παρατηρήσεων που λαμβάνονται ανά πείραμα και πρόκειται να μεταβάλλεται εκτελώντας τα πειράματα μέτρησης, προκειμένου να υπολογιστεί το MSE για κάθε τιμή από 1 έως 50 του πλήθους των μετρήσεων που λαμβάνονται ανά πείραμα.

Επειδή αναφερόμαστε σε πειράματα μέτρησης και θέλουμε να ορίσουμε το Μέσο Τετράγωνο Σφάλμα σαν συνάρτηση του αριθμού των παρατηρήσεων ανά πείραμα, δηλαδή $MSE=MSE(N)$, τότε για τον υπολογισμό του MSE, που αναφέρεται σε ένα συγκεκριμένο πλήθος N παρατηρήσεων ανά πείραμα, βρίσκουμε τον μέσο όρο των “Τετράγωνων Σφαλμάτων” που υπολογίστηκαν για την εκτίμηση της παραμέτρου θ στα διάφορα πειράματα μέτρησης που έγιναν, λαμβάνοντας N το πλήθος παρατηρήσεις σε καθένα από αυτά.

Σημειώνουμε ότι το “Τετράγωνο Σφάλμα” για την εκτίμηση της παραμέτρου θ σε ένα πείραμα μέτρησης ορίζεται ως $(\theta - \hat{\theta})^2$, όπου $\hat{\theta}$ είναι ο εκτιμητής για την θ (εδώ αναφερόμαστε στον MMSE και στον MLE σε καθεμιά από τις δύο περιπτώσεις, αντίστοιχα). Τότε, για κάθε πείραμα μέτρησης θα υπολογίσουμε το “Τετράγωνο Σφάλμα” της εκτίμησης της θ και το MSE θα προκύψει παίρνοντας τον μέσο όρο όλων αυτών των “Τετράγωνων Σφαλμάτων” που αναφέρονται σε εκτιμήσεις της θ από πειράματα μέτρησης με το ίδιο πλήθος δεδομένων κάθε φορά. Δηλαδή, εκφράζουμε έτσι το MSE σαν συνάρτηση του πλήθους N των παρατηρήσεων ανά πείραμα. Επειδή ακριβώς το MSE υπολογίζεται σαν ένας μέσος όρος, θα θέλαμε για κάθε τιμή του N να έχουμε έναν ικανοποιητικό αριθμό πειραμάτων που εκτελούνται λαμβάνοντας σε καθένα από αυτά το ίδιο πλήθος παρατηρήσεων από το σύστημα. Ακόμη, θα θέλαμε να έχουμε περίπου το ίδιο πλήθος πειραμάτων που επαναλαμβάνονται για καθεμιά τιμή του πλήθους N των παρατηρήσεων ανά πείραμα.

Θα εκτελεστούν συνολικά 500 πειράματα μέτρησης και θέλουμε να κατασκευάσουμε το γράφημα του MSE για κάθε εκτιμητή ως συνάρτηση του αριθμού των παρατηρήσεων που λήφθηκαν ανά πείραμα, μεταβάλλοντας τον αριθμό των παρατηρήσεων που έγιναν ανά πείραμα από 1 έως 50. Επομένως, πρέπει να καταναείμουμε τα 500 πειράματα σε αυτές τις διαφορετικές τιμές του αριθμού των παρατηρήσεων που λαμβάνονται ανά πείραμα, δηλαδή το N . Εφόσον έχουμε συνολικό αριθμό πειραμάτων 500 και μέγιστο αριθμό παρατηρήσεων ανά πείραμα 50, μπορούμε να καταναείμουμε ομοιόμορφα αυτά τα πειράματα, έτσι ώστε για κάθε τιμή του N από 1 έως 50 να εκτελούμε $500/50=10$ πειράματα και τελικά να ορίσουμε, όπως περιγράφηκε παραπάνω, το Μέσο Τετράγωνο Σφάλμα της εκτίμησης της θ που αναφέρεται σε μια συγκεκριμένη τιμή του αριθμού των παρατηρήσεων που λαμβάνονται ανά πείραμα (N), κάθε φορά.

Υλοποιούμε στο Matlab αλγόριθμο (είναι ίδιας λογικής και για τις δύο περιπτώσεις εκτιμητών που

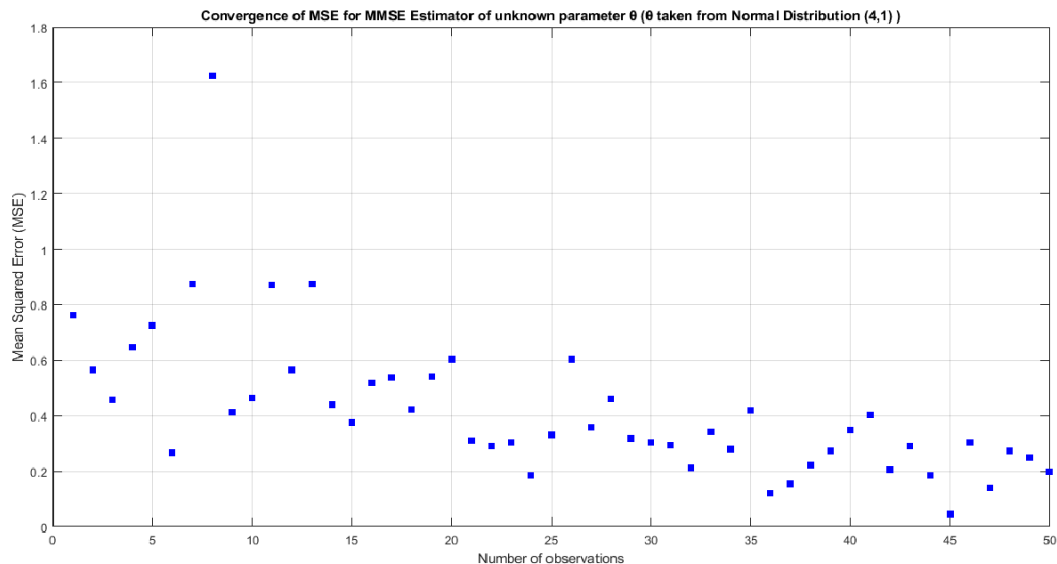
θα παρουσιαστούν παρακάτω) που εκτελεί, για κάθε τιμή του πλήθους N των δεδομένων που λαμβάνονται ανά πείραμα, δέκα πειράματα μέτρησης υπολογίζοντας το “Τετράγωνο Σφάλμα” $(\theta - \hat{\theta})^2$ σε καθένα από αυτά και αποθηκεύοντας αυτούς τους υπολογισμούς σε ένα διάνυσμα μήκους 10. Στο τέλος κάθε “δεκάδας” πειραμάτων που αφορούν μια συγκεκριμένη τιμή του N , υπολογίζεται το MSE σαν τον μέσο όρο των στοιχείων αυτού του διανύσματος. Έτσι, έχουμε το MSE για κάθε τιμή του N από 1 έως 50. Αυτό που αλλάζει στη διαδικασία υπολογισμού του $\text{MSE} = \text{MSE}(N)$ για τους εκτιμητές MMSE και MLE αντίστοιχα (και άρα αλλάζει και ο κώδικας στο Matlab σε εκείνο το σημείο) είναι η υπόθεση για την παράμετρο θ και ο τρόπος υπολογισμού του εκτιμητή $\hat{\theta}$, κάθε φορά.

Για τον MMSE εκτιμητή για την θ :

Σε αυτή την περίπτωση εκτίμησης της παραμέτρου θ του συστήματος μέτρησης της (2), η παράμετρος θ θεωρείται ως μια γκαουσιανή τυχαία μεταβλητή (στην πραγματικότητα πραγματοποίηση μιας γκαουσιανής τυχαίας μεταβλητής) με μέση τιμή $\mu=4$ και διασπορά $\sigma_0^2=1$ και τότε το σύστημα μέτρησης περιγράφεται από το General Linear Gaussian Model, όπως περιγράψαμε στο Ερώτημα 1. Τότε ο MMSE για την θ σε κάθε πείραμα μέτρησης δίνεται από μια από τις σχέσεις: $E[\theta|x] = \mu + \sigma_0^2 H^T (H \sigma_0^2 H^T + C_w)^{-1} (x - H\mu)$

ή $E[\theta|x] = \mu + (1/\sigma_0^2 + H^T C_w^{-1} H)^{-1} H^T C_w^{-1} (x - H\mu)$, όπως ακριβώς διατυπώθηκαν στο ίδιο ερώτημα.

Όπως περιγράψαμε τη διαδικασία μελέτης της σύγκλισης του MSE ως συνάρτηση του αριθμού των μετρήσεων (πλήθος παρατηρήσεων N), που λαμβάνονται ανά πείραμα μέτρησης, εκτελούμε για κάθε τιμή του N , από 1 έως 50, δέκα πειράματα μέτρησης σε καθένα από τα οποία η παράμετρος θ του μετρητικού συστήματος ορίζεται ως η πραγματοποίηση μιας γκαουσιανής τυχαίας μεταβλητής με τα παραπάνω χαρακτηριστικά (δύο πρώτες ροπές) που περιγράφηκαν, παράγονται N το πλήθος δείγματα Γκουσιανού θορύβου w και τελικά συνθέτεται το διάνυσμα x των παρατηρήσεων, μήκους N , το οποίο περιλαμβάνει τα δείγματα αυτά που λαμβάνονται στο πείραμα και θα αξιοποιηθούν στην εύρεση του MMSE για την θ . Συνεπώς, στο “τέλος” κάθε πειράματος, δηλαδή όταν συλλεχθούν (η μοντελοποίηση της συλλογής δεδομένων στο Matlab αφορά παραγωγή “συνθετικών” δεδομένων βάσει της σχέσης (2) που περιγράφει το μετρητικό σύστημα) τα δεδομένα από το σύστημα μέτρησης και υπολογιστεί ο MMSE για την συγκεκριμένη πραγματοποίηση της θ , υπολογίζεται το “Τετράγωνο Σφάλμα” που αντιστοιχεί σε αυτό το πείραμα. Στο τέλος όλων (10 το πλήθος) των πειραμάτων που αφορούν την τιμή N του πλήθους των παρατηρήσεων που λαμβάνονται ανά πείραμα υπολογίζεται το $\text{MSE}(N)$, όπως περιγράψαμε προηγουμένως. Το γράφημα του MSE για τον εκτιμητή MMSE για την παράμετρο θ , συναρτήσει του πλήθους των μετρήσεων που λαμβάνονται ανά πείραμα μέτρησης παρουσιάζεται στο Σχήμα 2.3.1



Σχήμα 2.3.1 : Σύγκλιση Μέσου Τετράγωνου Σφάλματος, για τον εκτιμητή MMSE για την παράμετρο θ (που είναι πραγματοποίηση γκαουσιανής τυχαίας μεταβλητής), ως συνάρτηση του πλήθους των παρατηρήσεων που λαμβάνονται ανά πείραμα για 500 πειράματα μέτρησης και μέγιστο αριθμό παρατηρήσεων ανά πείραμα 50.

Παρατηρούμε από το παραπάνω διάγραμμα ότι το MSE για τον MMSE για την θ λαμβάνει τιμές γύρω από την τιμή 0.7 με κάποια εναλλασσόμενη συμπεριφορά γύρω από αυτή την τιμή, μεταβάλλοντας το πλήθος N των παρατηρήσεων ανά πείραμα από 1 έως 10 περίπου, δηλαδή για μικρές τιμές του αριθμού των παρατηρήσεων που λαμβάνονται ανά πείραμα και με το μέσο όρο του MSE να φθίνει γενικότερα αυξάνοντας το πλήθος N των παρατηρήσεων ανά πείραμα, λόγω εμφάνισης χαμηλότερων τιμών MSE με αύξηση του N . Καθώς αυξάνει ο αριθμός N των παρατηρήσεων που λαμβάνονται ανά πείραμα, παρατηρούμε ότι το MSE φθίνει ακόμα περισσότερο με αυτή την εναλλασσόμενη συμπεριφορά να είναι λιγότερο αισθητή, ίσως γύρω από τιμές μικρότερες του 0.7 όπως το 0.4, το 0.3 ή το 0.2. Γενικότερα, σαν συμπεριφορά του MSE παρατηρούμε, ανά κάποιο εύρος τιμών του N , κάποια κυμάτωση της τιμής του MSE γύρω από μια τιμή που συνεχώς φθίνει. Αυτό που είναι σημαντικό πάντως να παρατηρήσουμε είναι η σύγκλιση με οποιοδήποτε τρόπο προς το 0 του MSE -που εξαρχής δεν είναι ιδιαίτερα μεγάλο ακόμη και για πειράματα στα οποία λαμβάνεται μόνο μια παρατήρηση ως δεδομένο- όσο αυξάνει το πλήθος των παρατηρήσεων ανά πείραμα και ιδίως όταν αυτός ο αριθμός γίνει μεγαλύτερος του 30, περίπου.

Για τον MLE εκτιμητή για την θ :

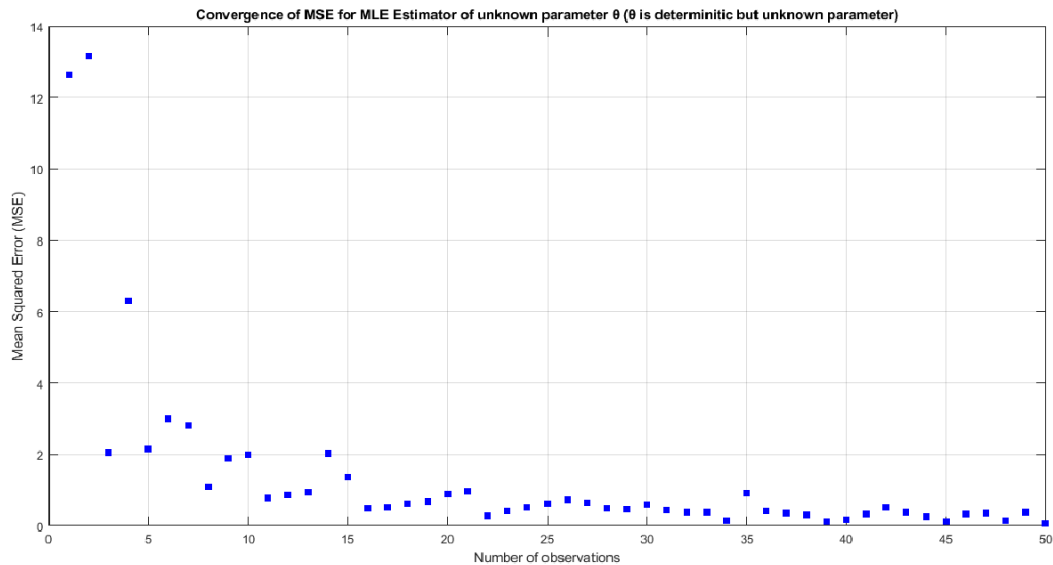
Σε αυτή την περίπτωση εκτίμησης της παραμέτρου θ του συστήματος μέτρησης της (2), η παράμετρος θ θεωρείται ντετερμινιστική αλλά άγνωστη και τότε το σύστημα μέτρησης περιγράφεται από το General Linear Model, όπως περιγράψαμε στο Ερώτημα 2. Τότε ο MLE για την θ σε κάθε πείραμα μέτρησης δίνεται από τη σχέση:

$$\theta_{MLE}^{\wedge} = (H^T C_w^{-1} H)^{-1} H^T C_w^{-1} x, \text{ όπως ακριβώς διατυπώθηκαν στο ίδιο ερώτημα.}$$

Όπως περιγράψαμε τη διαδικασία μελέτης της σύγκλισης του MSE ως συνάρτηση του αριθμού των μετρήσεων (πλήθος παρατηρήσεων N), που λαμβάνονται ανά πείραμα μέτρησης, εκτελούμε για κάθε τιμή του N , από 1 έως 50, δέκα πειράματα μέτρησης σε καθένα από τα οποία η παράμετρος θ του μετρητικού συστήματος ορίζεται ως ντετερμινιστική ποσότητα, δηλαδή λαμβάνει μια αυθαίρετη τιμή, που χωρίς περιορισμό της γενικότητας επιλέξαμε την τιμή $\theta=4$, παράγονται N το πλήθος δείγματα Γκουσιανού θορύβου w και τελικά συνθέτεται το διάνυσμα x των παρατηρήσεων, μήκους N , το οποίο περιλαμβάνει τα δείγματα αυτά που λαμβάνονται στο πείραμα και θα αξιοποιηθούν στην εύρεση του MLE για την θ . Συνεπώς, στο “τέλος” κάθε πειράματος, δηλαδή

όταν συλλεχθούν (η μοντελοποίηση της συλλογής δεδομένων στο Matlab αφορά παραγωγή “συνθετικών” δεδομένων βάσει της σχέσης (2) που περιγράφει το μετρητικό σύστημα) τα δεδομένα από το σύστημα μέτρησης και υπολογιστεί ο MLE για την θ , υπολογίζεται το “Τετράγωνο Σφάλμα” που αντιστοιχεί σε αυτό το πείραμα. Στο τέλος όλων (10 το πλήθος) των πειραμάτων που αφορούν την τιμή N του πλήθους των παρατηρήσεων που λαμβάνονται ανά πείραμα υπολογίζεται το $MSE(N)$, όπως περιγράψαμε προηγουμένως.

Το γράφημα του MSE για τον εκτιμητή MLE για την παράμετρο θ , συναρτήσει του πλήθους των μετρήσεων που λαμβάνονται ανά πείραμα μέτρησης παρουσιάζεται στο Σχήμα 2.3.2

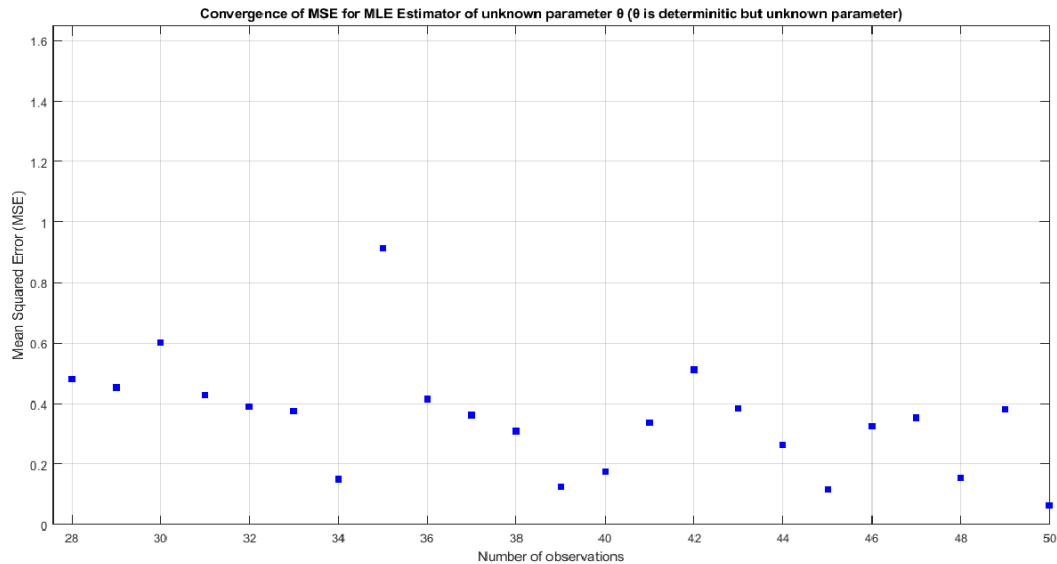


Σχήμα 2.3.2 : Σύγκλιση Μέσου Τετράγωνου Σφάλματος, για τον εκτιμητή MLE για την παράμετρο θ (που είναι ντετερμινιστική, αλλά άγνωστη παράμετρος, η οποία εκτιμάται από τα δεδομένα που λαμβάνονται), ως συνάρτηση του πλήθους των παρατηρήσεων που λαμβάνονται ανά πείραμα για 500 πειράματα μέτρησης και μέγιστο αριθμό παρατηρήσεων ανά πείραμα 50.

Παρατηρούμε από το παραπάνω διάγραμμα ότι το MSE για τον MLE για την θ σαν συνάρτηση του πλήθους N των παρατηρήσεων που λαμβάνονται ανά πείραμα είναι μια φθίνουσα συνάρτηση, η οποία για μεγάλες τιμές του N , δηλαδή για $N > 20$ όπως παρατηρούμε, συγκλίνει προς το 0. Όπως είναι εμφανές από το γράφημα του παραπάνω σχήματος, το MSE λαμβάνει σημαντικά μεγάλες τιμές για τις τέσσερις πρώτες τιμές του αριθμού N των παρατηρήσεων που λαμβάνονται ανά πείραμα και ιδίως για $N=1,2$. Καθώς αυξάνει το πλήθος N των παρατηρήσεων που λαμβάνονται ανά πείραμα, παρατηρούμε ότι το MSE φθίνει συνεχώς και αυτή η συμπεριφορά είναι ιδιαίτερα έντονη για τις τιμές του $N < 15$ περίπου, όπου το MSE λαμβάνει πιο υψηλές τιμές και η μείωση από την αμέσως προηγούμενη τιμή (τιμή του MSE για την αμέσως προηγούμενη τιμή του N) είναι μεγαλύτερη.. Αυτό που είναι σημαντικό πάντως να παρατηρήσουμε είναι η σύγκλιση του MSE προς το 0 του MSE όσο αυξάνει το πλήθος των παρατηρήσεων ανά πείραμα και ιδίως όταν αυτός ο αριθμός γίνει μεγαλύτερος του 20-30, περίπου.

Το MSE είναι ένα μέτρο της ποιότητας απόδοσης ενός εκτιμητή για την παράμετρο θ . Τότε, παρατηρώντας τα δύο γραφήματα στα Σχήματα 2.3.1 και 2.3.2, αντίστοιχα, για τη σύγκλιση του MSE για τον MMSE και τον MLE για την θ , μπορούμε να διατυπώσουμε μια σύγκριση μεταξύ των δύο εκτιμητών. Όπως, εύκολα, παρατηρούμε όταν το πλήθος των παρατηρήσεων που λαμβάνονται ανά πείραμα είναι μικρό (μικρότερο του 30-35), τότε ο MLE παρουσιάζει μεγαλύτερο και ιδίως για τις πιο μικρές τιμές του πλήθους αυτού ($N < 10$) σημαντικά μεγαλύτερο MSE από τον MMSE. Είναι αξιοσημείωτο, ότι η συμπεριφορά του MMSE, σύμφωνα με τη μελέτη του MSE, για $N < 15$ επιτυγχάνεται με τον MLE για $N=8$ ή $N > 8$. Για μικρότερες τιμές, δηλαδή, του πλήθους των παρατηρήσεων που λαμβάνονται ανά πείραμα, ο MMSE μοιάζει να δίνει καλύτερες εκτιμήσεις της

θ στα πειράματα μέτρησης αυτού με το μετρητικό σύστημα της (2), με όρους MSE. Ωστόσο, όπως παρατηρήσουμε (φαίνεται καλύτερα στο Σχήμα 2.3.3) για μεγάλες τιμές του αριθμού του παρατηρήσεων που λαμβάνονται ανά πείραμα ($N > 35$) οι δύο εκτιμητές παρουσιάζουν παρόμοια συμπεριφορά όσον αφορά την εκτίμηση της θ , με όρους MSE, το οποίο όπως είπαμε συγκλίνει προς το μηδέν αυξανόμενου του αριθμού N , γεγονός που υποδηλώνει αρκετά καλή εκτίμηση της παραμέτρου και με τους δύο εκτιμητές.



Σχήμα 2.3.3 : Σύγκλιση Μέσου Τετράγωνου Σφάλματος, για τον εκτιμητή MLE για την παράμετρο θ (που είναι ντετερμινιστική, αλλά άγνωστη παράμετρος, η οποία εκτιμάται από τα δεδομένα που λαμβάνονται), ως συνάρτηση του πλήθους των παρατηρήσεων που λαμβάνονται ανά πείραμα για 500 πειράματα μέτρησης και μέγιστο αριθμό παρατηρήσεων ανά πείραμα 50 (Μεγέθυνση του Σχήματος 2.3.2 για τις τιμές του $N > 28$).

Ερώτημα 4: Σε αυτό το ερώτημα, θα χρησιμοποιήσουμε πάλι την Bayesian προσέγγιση, για να αναζητήσουμε τον MMSE εκτιμητή για την άγνωστη παράμετρο θ , η οποία σε αυτή την περίπτωση θεωρείται ότι είναι η πραγματοποίηση μιας τυχαίας μεταβλητής, που ακολουθεί μια άλλη (όχι gaussian) “εσφαλμένη” κατανομή. Ο όρος “εσφαλμένη” αναφέρεται στο γεγονός, ότι σε αυτή την ανάλυση θα διατηρήσουμε το ίδιο μοντέλο για τον MMSE εκτιμητή, που είχαμε ορίσει στο Ερώτημα 1 (και χρησιμοποιήσαμε και στο Ερώτημα 3), το οποίο όμως ορίζει τον MMSE εκτιμητή για την παράμετρο θ , όταν υποθέτουμε, ότι η πρότερη γνώση για την παράμετρο αυτή μοντελοποιείται μέσω μιας gaussian prior pdf $p(\theta)$ και τότε το σύστημα περιγράφεται από το General Linear Gaussian Model, δηλαδή όταν η θ είναι πραγματοποίηση μιας gaussian τυχαίας μεταβλητής. Όταν, λοιπόν, υποθέτουμε κάποια άλλη prior pdf για την θ , η οποία δεν είναι gaussian, τότε έχουμε τον όρο “εσφαλμένη” για την επιλογή της κατανομής όσον αφορά το μοντέλο που χρησιμοποιούμε για να ορίσουμε και να εκτελέσουμε τον MMSE εκτιμητή.

Θα εξετάσουμε, λοιπόν, τις συνέπειες που μπορεί να έχει μια τέτοια “εσφαλμένη” επιλογή για την κατανομή από την οποία, ουσιαστικά, αντλείται η θ , στην εκτίμηση της παραμέτρου θ με τον MMSE εκτιμητή που ορίσαμε για το General Linear Gaussian Model.

Για να μοντελοποιήσουμε αυτή την “εσφαλμένη” επιλογή της prior pdf της θ (“εσφαλμένη” κατανομή για την θ), θα διατηρήσουμε το ίδιο μοντέλο MMSE εκτιμητή που ορίσαμε στο Ερώτημα 1 για την εκτίμηση της άγνωστης παραμέτρου θ και θα αντλήσουμε τιμές για την παράμετρο θ (δηλαδή θα ορίσουμε αυτή την πραγματοποίηση της τυχαίας μεταβλητής θ , που ακολουθεί κάποια γνωστή μεν αλλά “εσφαλμένη” κατανομή δε) από ομοιόμορφη αρχικά και από εκθετική κατανομή, κατόπιν. Όπως ήδη είπαμε, το μοντέλο αυτό για τον MMSE εκτιμητή για την θ ορίζεται από την Bayesian προσέγγιση για την περίπτωση που το σύστημα μέτρησης μπορεί να περιγραφεί από το

General Linear Gaussian Model, δηλαδή την περίπτωση που τόσο ο θόρυβος w όσο και η άγνωστη παράμετρος θ , που επιθυμούμε να εκτιμήσουμε, είναι gaussian. Ωστόσο, είναι σημαντικό να σημειώσουμε, ότι αυτό το μοντέλο του MMSE εκτιμητή είναι το ίδιο με το μοντέλο του LMMSE (Linear MMSE) εκτιμητή, όταν το σύστημα μέτρησης μπορεί να περιγραφεί από το General Linear Model, στο οποίο αρκεί μόνο ο θόρυβος w να είναι Γκαουσιανός, ενώ η άγνωστη παράμετρος θ μπορεί να αντλείται από οποιαδήποτε κατανομή. Χρησιμοποιώντας, δηλαδή, αυτό το μοντέλο του MMSE εκτιμητή που ορίσαμε στο Ερώτημα 1, οδηγούμαστε στον LMMSE εκτιμητή για την θ , στις περιπτώσεις που η παράμετρος αυτή είναι πραγματοποίηση τυχαίας μεταβλητής, η οποία ακολουθεί στη μία περίπτωση ομοιόμορφη και στην άλλη περίπτωση εκθετική κατανομή. Πρέπει να σημειωθεί, ότι ο LMMSE ταυτίζεται με τον MMSE όταν το General Linear Model είναι Gaussian, ενώ στη γενική περίπτωση του Γενικού Γραμμικού Μοντέλου, πρόκειται για έναν βέλτιστο Bayesian εκτιμητή, ο οποίος έχει το θετικό χαρακτηριστικό ότι ορίζεται σε κλειστού τύπου μορφή, σε αντίθεση με τον MMSE, όταν η prior pdf της άγνωστης παραμέτρου θ δεν είναι gaussian. Από τη Bayesian θεωρία, αναμένουμε ότι αυτός ο εκτιμητής, δηλαδή ο LMMSE, θα αυξάνει την ποιότητα του ως προς την εκτίμηση της άγνωστης παραμέτρου θ , όταν το πλήθος των παρατηρήσεων, δηλαδή των δεδομένων, που λαμβάνονται από το σύστημα είναι ικανοποιητικά μεγάλο.

Τελικά, προκειμένου να διερευνήσουμε τις συνέπειες της “εσφαλμένης” επιλογής για την κατανομή της παραμέτρου θ , την οποία, όπως είπαμε, θα μοντελοποιήσουμε πρώτον με επιλογή ομοιόμορφης κατανομής σαν prior και δεύτερον με επιλογή εκθετικής, θα σχεδιάσουμε το γράφημα της σύγκλισης του Μέσου Τετράγωνου Σφάλματος, για την εκτίμηση της παραμέτρου θ με το “εσφαλμένο” μοντέλο MMSE εκτιμητή (στην πραγματικότητα LMMSE, όπως εξηγήθηκε) αφενός όταν η θ ακολουθεί μια ομοιόμορφη και αφετέρου όταν ακολουθεί μια εκθετική κατανομή, σαν συνάρτηση του πλήθους N των παρατηρήσεων που λαμβάνονται ανά πείραμα μέτρησης για 500 διαφορετικά πειράματα και μέγιστο αριθμό παρατηρήσεων ανά πείραμα 50, με την διαδικασία όπως ακριβώς εξηγήθηκε στο Ερώτημα 3.

α) Οι τιμές της θ αντλούνται από ομοιόμορφη κατανομή.

Σε αυτή την περίπτωση, υποθέτουμε ότι η θ είναι πραγματοποίηση μιας τυχαίας μεταβλητής θ που ακολουθεί ομοιόμορφη κατανομή $U[0,8]$ (χωρίς περιορισμό της γενικότητας). Το σύστημα μέτρησης είναι αυτό που περιγράφεται από την (2) ή αντίστοιχα την (2.1), δηλαδή από ένα Γενικό Γραμμικό Μοντέλο.

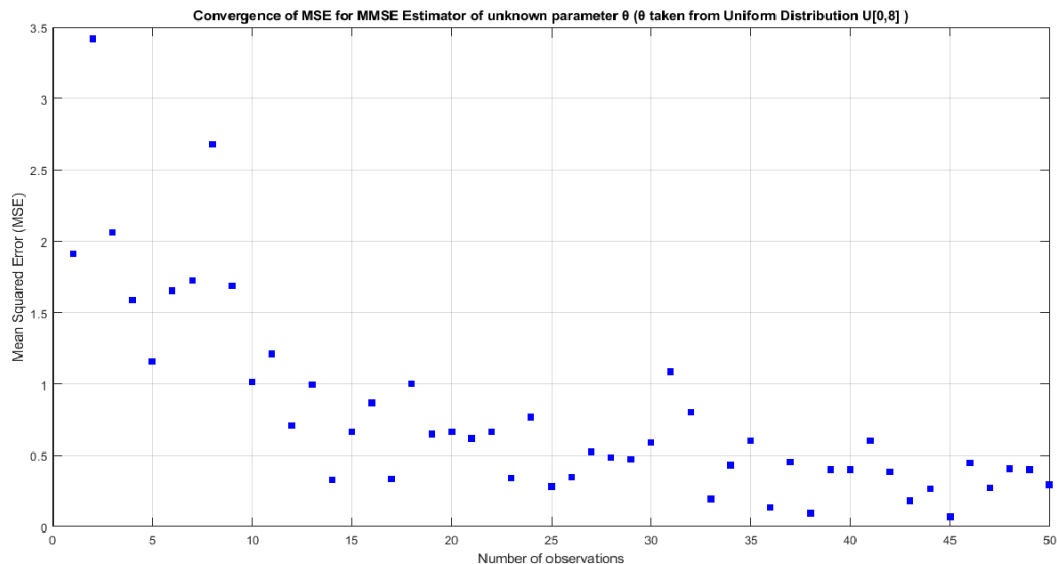
Παρά το γεγονός, ότι η prior pdf $p(\theta)$ δεν είναι gaussian, αλλά ορίζει την $U[0,8]$, χρησιμοποιούμε ως μοντέλο MMSE εκτιμητή για την θ , αυτό που ορίσαμε στο Ερώτημα 1 για το General Linear Gaussian Model, δηλαδή θεωρούμε ότι ο MMSE δίνεται από τις σχέσεις:

$$E[\theta|x] = \mu + \sigma_0^2 H^T (H \sigma_0^2 H^T + C_w)^{-1} (x - H\mu)$$

$$\text{ή } E[\theta|x] = \mu + (1/\sigma_0^2 + H^T C_w^{-1} H)^{-1} H^T C_w^{-1} (x - H\mu), \text{ όπως περιγράφηκαν στο}$$

αντίστοιχο Ερώτημα. Στη συγκεκριμένη περίπτωση η μέση τιμή μ και η διασπορά σ_0^2 της τυχαίας μεταβλητής θ (μια πραγματοποίηση της οποίας είναι η άγνωστη παράμετρος που θέλουμε να εκτιμήσουμε) ορίζονται εκ νέου από την prior κατανομή που επιλέχθηκε για την θ , δηλαδή την $U[0,8]$. Είναι, λοιπόν: $\mu = \frac{(8+0)}{2} = 4$ και $\sigma_0^2 = \frac{(8-0)^2}{12} = 32$.

Το γράφημα του MSE για τον εκτιμητή MMSE για την παράμετρο θ , συναρτήσει του πλήθους των μετρήσεων που λαμβάνονται ανά πείραμα μέτρησης παρουσιάζεται στο Σχήμα 2.4.1



Σχήμα 2.4.1 : Σύγκλιση Μέσου Τετράγωνου Σφάλματος, για τον εκτιμητή MMSE για την παράμετρο θ (που είναι πραγματοποίηση τυχαίας μεταβλητής θ που ακολουθεί $U[0,8]$), ως συνάρτηση του πλήθους των παρατηρήσεων που λαμβάνονται ανά πείραμα για 500 πειράματα μέτρησης και μέγιστο αριθμό παρατηρήσεων ανά πείραμα 50.

Παρατηρούμε από το παραπάνω διάγραμμα ότι το MSE για τον MMSE για την θ σαν συνάρτηση του πλήθους N των παρατηρήσεων που λαμβάνονται ανά πείραμα είναι μια φθίνουσα συνάρτηση, η οποία για μεγάλες τιμές του N , δηλαδή για $N > 25$ όπως παρατηρούμε, συγκλίνει προς το 0. Όπως είναι εμφανές από το γράφημα του παραπάνω σχήματος, το MSE λαμβάνει σημαντικά μεγάλες τιμές για τις τρεις έως εννιά πρώτες τιμές του αριθμού N των παρατηρήσεων που λαμβάνονται ανά πείραμα. Καθώς αυξάνει το πλήθος N των παρατηρήσεων που λαμβάνονται ανά πείραμα, παρατηρούμε ότι το MSE “κατανέμεται” γύρω από μια τιμή (με τον όρο αυτό εννοούμε τον τοπικό μέσο όρο του MSE κάθε φορά, που παρατηρείται προσεγγιστικά από το γράφημα για μια γειτονιά τιμών του N κάθε φορά) που φθίνει συνεχώς και αυτή η συμπεριφορά είναι ιδιαίτερα έντονη για τις τιμές του $N < 15$ περίπου, όπου το MSE λαμβάνει πιο υψηλές τιμές και η μείωση από την αμέσως προηγούμενη τιμή “γύρω” από την οποία παρατηρούμε την τιμή του MSE για την αμέσως προηγούμενη τιμή του N είναι μεγαλύτερη. Αυτό που είναι σημαντικό πάντως να παρατηρήσουμε είναι η σύγκλιση του MSE προς το 0 του MSE όσο αυξάνει -σημαντικά- το πλήθος των παρατηρήσεων ανά πείραμα. Ακόμη, παρατηρούμε ότι η σύγκλιση αυτή προς το 0 δεν είναι μονότονη, αλλά οι τιμές του MSE, καθώς το πλήθος N των παρατηρήσεων ανά πείραμα αυξάνει, παρουσιάζουν μια κυμάτωση γύρω από μια τιμή που φθίνει συνεχώς (ο τοπικός μέσος όρος που αναφέραμε παραπάνω). Αυτή η συμπεριφορά διαπιστώνεται πιο εύκολα στις μεγάλες τιμές του αριθμού N .

Η πιο σημαντική παρατήρηση, φυσικά, είναι η σύγκλιση του MSE για την εκτίμηση της θ στο μηδέν, όσο αυξάνει ο αριθμός των μετρήσεων που λαμβάνονται από το σύστημα στο πείραμα μέτρησης. Αυτή υποδηλώνει ότι το μοντέλο του MMSE εκτιμητή για την θ , μπορεί να εφαρμοστεί ικανοποιητικά, με όρους MSE, για την εκτίμηση της άγνωστης παραμέτρου όταν το πλήθος των δεδομένων ανά πείραμα είναι μεγάλο. Μάλιστα, η συμπεριφορά του εκτιμητή αυτού στο φάσμα των τιμών του αριθμού των παρατηρήσεων που λαμβάνονται ανά πείραμα για $N > 10$, θυμίζει την συμπεριφορά του MLE για την θ , όταν υποθέσαμε την άγνωστη παράμετρο ντετερμινιστική στο Ερώτημα 3, όπως μπορούμε να διαπιστώσουμε και από το Σχήμα 2.3.2. Σημαντική παρατήρηση, ακόμη, είναι ότι ο εκτιμητής αυτός παρουσιάζεται πιο αποδοτικός, με όρους MSE, από τον MLE, για τις “μικρές” τιμές του πλήθους N των παρατηρήσεων που λαμβάνονται ανά πείραμα.

Η ικανοποιητική απόδοση του μοντέλου MMSE εκτιμητή για εσφαλμένη θεωρητικά επιλογή prior κατανομής για την θ , όταν το πλήθος των παρατηρήσεων ανά πείραμα είναι αρκετά μεγάλο, είναι

κάτι που αναμέναμε, σύμφωνα με την Bayesian προσέγγιση στη θεωρία εκτίμησης, αφού όπως είπαμε αυτό το μοντέλο ορίζει έναν κλειστού τύπου “βέλτιστο” εκτιμητή για το General Linear Model, ο οποίος είναι ο LMMSE.

β) Οι τιμές της θ αντλούνται από εκθετική κατανομή.

Σε αυτή την περίπτωση, υποθέτουμε ότι η θ είναι πραγματοποίηση μιας τυχαίας μεταβλητής θ που ακολουθεί εκθετική κατανομή με παράμετρο $\lambda=1/4$, $\exp(\lambda)$ (χωρίς περιορισμό της γενικότητας). Το σύστημα μέτρησης είναι αυτό που περιγράφεται από την (2) ή αντίστοιχα την (2.1), δηλαδή από ένα Γενικό Γραμμικό Μοντέλο.

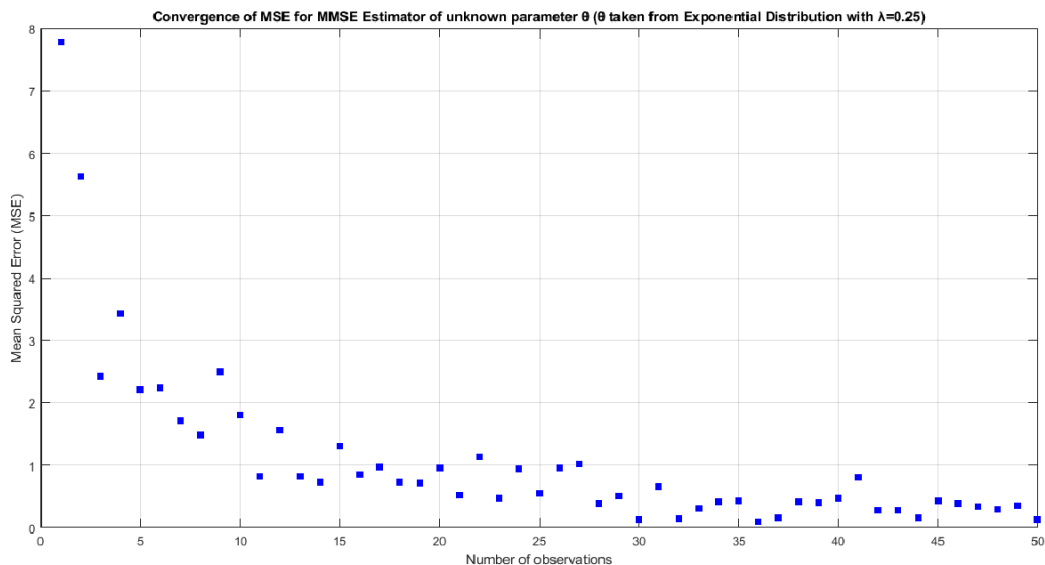
Παρά το γεγονός, ότι η prior pdf $p(\theta)$ δεν είναι gaussian, αλλά ορίζει την $\exp(\lambda)$, χρησιμοποιούμε ως μοντέλο MMSE εκτιμητή για την θ , αυτό που ορίσαμε στο Ερώτημα 1 για το General Linear Gaussian Model, δηλαδή θεωρούμε ότι ο MMSE δίνεται από τις σχέσεις:

$$E[\theta|x] = \mu + \sigma_0^2 H^T (H \sigma_0^2 H^T + C_w)^{-1} (x - H\mu)$$

$$\text{ή } E[\theta|x] = \mu + (1/\sigma_0^2 + H^T C_w^{-1} H)^{-1} H^T C_w^{-1} (x - H\mu), \text{ όπως περιγράφηκαν στο}$$

αντίστοιχο Ερώτημα. Στη συγκεκριμένη περίπτωση η μέση τιμή μ και η διασπορά σ_0^2 της τυχαίας μεταβλητής θ (μια πραγματοποίηση της οποίας είναι η άγνωστη παράμετρος που θέλουμε να εκτιμήσουμε) ορίζονται εκ νέου από την prior κατανομή που επιλέχθηκε για την θ , δηλαδή την $\exp(\lambda)$. Είναι, λοιπόν: $\mu = \frac{1}{\lambda} = 4$ και $\sigma_0^2 = \frac{1}{\lambda^2} = 16$.

Το γράφημα του MSE για τον εκτιμητή MMSE για την παράμετρο θ , συναρτήσει του πλήθους των μετρήσεων που λαμβάνονται ανά πείραμα μέτρησης παρουσιάζεται στο Σχήμα 2.4.2



Σχήμα 2.4.2 : Σύγκλιση Μέσου Τετράγωνου Σφάλματος, για τον εκτιμητή MMSE για την παράμετρο θ (που είναι πραγματοποίηση τυχαίας μεταβλητής θ που ακολουθεί $\exp(\lambda)$, όπου $\lambda=1/4$), ως συνάρτηση του πλήθους των παρατηρήσεων που λαμβάνονται ανά πείραμα για 500 πειράματα μέτρησης και μέγιστο αριθμό παρατηρήσεων ανά πείραμα 50.

Παρατηρούμε από το παραπάνω διάγραμμα ότι το MSE για τον MMSE για την θ σαν συνάρτηση του πλήθους N των παρατηρήσεων που λαμβάνονται ανά πείραμα είναι μια φθίνουσα συνάρτηση, η οποία για μεγάλες τιμές του N , δηλαδή για $N > 27$ όπως παρατηρούμε, συγκλίνει προς το 0. Όπως είναι εμφανές από το γράφημα του παραπάνω σχήματος, το MSE λαμβάνει σημαντικά μεγάλες τιμές για τις τέσσερις έως εννιά πρώτες τιμές του αριθμού N των παρατηρήσεων που λαμβάνονται ανά πείραμα και ιδίως για $N=1,2$. Καθώς αυξάνει το πλήθος N των παρατηρήσεων που λαμβάνονται ανά πείραμα, παρατηρούμε ότι το MSE φθίνει συνεχώς με τη φθίνουσα αυτή η

συμπεριφορά να είναι ιδιαίτερα έντονη για μικρότερες τιμές του N , όπου το MSE λαμβάνει πιο υψηλές τιμές και η μείωση από την αμέσως προηγούμενη τιμή (τιμή του MSE για την αμέσως προηγούμενη τιμή του N) είναι μεγαλύτερη. Αυτό που είναι σημαντικό πάντως να παρατηρήσουμε είναι η σύγκλιση του MSE προς το 0 του MSE όσο αυξάνει το πλήθος των παρατηρήσεων ανά πείραμα. Ακόμη, παρατηρούμε ότι η σύγκλιση αυτή προς το 0 δεν είναι ακριβώς μονότονη, αλλά οι τιμές του MSE, καθώς το πλήθος N των παρατηρήσεων ανά πείραμα αυξάνει, παρουσιάζουν κάποια κυμάτωση για κάποιες τιμές του N , γύρω από έναν “τοπικό” μέσο όρο που φθίνει συνεχώς προς το μηδέν. Αυτή η συμπεριφορά διαπιστώνεται πιο εύκολα στις πιο μεγάλες τιμές του αριθμού N .

Η πιο σημαντική παρατήρηση, φυσικά, είναι η σύγκλιση του MSE για την εκτίμηση της θ στο μηδέν, όσο αυξάνει ο αριθμός των μετρήσεων που λαμβάνονται από το σύστημα στο πείραμα μέτρησης. Αυτή υποδηλώνει ότι το μοντέλο του MMSE εκτιμητή για την θ , μπορεί να εφαρμοστεί ικανοποιητικά, με όρους MSE, για την εκτίμηση της άγνωστης παραμέτρου όταν το πλήθος των δεδομένων ανά πείραμα είναι μεγάλο. Μάλιστα η συμπεριφορά του εκτιμητή αυτού σε όλο το φάσμα των τιμών του αριθμού N , θυμίζει την συμπεριφορά του MLE για την θ , όταν υποθέσαμε την άγνωστη παράμετρο ντετερμινιστική στο Ερώτημα 3, όπως μπορούμε να διαπιστώσουμε και από το Σχήμα 2.3.2. Και σε αυτή την περίπτωση, είναι χαρακτηριστικό ότι έχουμε σημαντικά μεγάλες τιμές MSE για τον εκτιμητή για τις μικρές τιμές του N , ενώ το γράφημα του MSE συναρτήσει του αριθμού N φθίνει συνεχώς (με εκείνη τη συμπεριφορά κυμάτωσης γύρω από έναν συνεχώς φθίνων “τοπικό” μέσο όρο, που σχολιάσαμε παραπάνω και η οποία θυμίζει εκείνη την κυμάτωση ανά διαστήματα του MSE στην περίπτωση του MLE του Ερωτήματος 3), όπως και στην περίπτωση του MLE για την παράμετρο θ .

Η ικανοποιητική απόδοση του μοντέλου MMSE εκτιμητή για εσφαλμένη θεωρητικά επιλογή prior κατανομής για την θ , όταν το πλήθος των παρατηρήσεων ανά πείραμα είναι αρκετά μεγάλο, είναι κάτι που αναμέναμε, σύμφωνα με την Bayesian προσέγγιση στη θεωρία εκτίμησης, αφού όπως είπαμε αυτό το μοντέλο ορίζει έναν κλειστού τύπου “βέλτιστο” εκτιμητή για το General Linear Model, ο οποίος είναι ο LMMSE.

Αν θέλαμε να συγκρίνουμε την ποιότητα αυτού του μοντέλου MMSE (ουσιαστικά τον LMMSE) για την εκτίμηση της θ σε αυτές τις δύο διαφορετικές περιπτώσεις για την prior κατανομή της θ , χρησιμοποιώντας όρους MSE, όπως ορίστηκε παραπάνω, θα λέγαμε ότι ο εκτιμητής MMSE αυτός εκτιμά την παράμετρο θ από τα δεδομένα που λαμβάνουμε από το μετρητικό σύστημα το ίδιο “καλά”, όταν το πλήθος των δεδομένων αυτών είναι αρκετά μεγάλο (μεγαλύτερο από 30 περίπου) και η συμπεριφορά αυτού του εκτιμητή, τότε, προσεγγίζει αυτή του MMSE για την θ όταν αυτή προέρχεται από κανονική κατανομή, δηλαδή με “σωστή” επιλογή μοντέλου MMSE, σύμφωνα με την prior κατανομή της θ . Όταν, όμως, το πλήθος N των δεδομένων από το σύστημα μέτρησης είναι μικρό (μικρότερο από 10), τότε το μοντέλο MMSE του General Linear Gaussian Model φαίνεται να αποδίδει καλύτερα, με όρους MSE, όσον αφορά την εκτίμηση της θ , όταν οι τιμές αυτής αντλούνται από ομοιόμορφη παρά από εκθετική κατανομή. Σε καμία από τις δύο περιπτώσεις, όμως, η συμπεριφορά του εκτιμητή δεν προσεγγίζει αυτή του MMSE για “σωστή” επιλογή prior. Στο συγκεκριμένο ερώτημα, δηλαδή, η “εσφαλμένη” επιλογή prior κατανομής για την θ έχει αντίκτυπο στην εκτίμηση της θ , όταν το πλήθος των δεδομένων δεν είναι αρκετά μεγάλο. Ωστόσο, όταν είναι ικανοποιητικά μεγάλο, η επιλογή prior κατανομής για την θ , διαφορετικής της gaussian, δεν μοιάζει τόσο “εσφαλμένη”. Αυτό οφείλεται, όπως είπαμε, στο γεγονός ότι μελετάμε ένα Γενικό Γραμμικό Μοντέλο, οπότε το “εσφαλμένο” μοντέλο MMSE ορίζει στην πραγματικότητα έναν “βέλτιστο” κλειστού τύπου εκτιμητή, τον LMMSE.

Θέμα 3ο:

Σε αυτή την άσκηση διαθέτουμε ένα σύνολο δεδομένων, τα οποία γνωρίζουμε ότι έχουν ληφθεί είτε από μια εκθετική κατανομή, είτε από μια κατανομή Rayleigh. Στην παρακάτω ανάλυση, θα εκτελέσουμε ορισμένες μεθόδους, ώστε να υλοποιήσουμε έναν διαγνωστικό έλεγχο για αυτό το

σύνολο των ληφθέντων δεδομένων, στην προσπάθεια να εκτιμήσουμε από ποιά κατανομή προήλθαν τα δεδομένα. Στην ανάλυση αυτή, εκμεταλλευόμαστε την πρωταρχική γνώση ότι τα δεδομένα προέρχονται, σίγουρα, είτε από εκθετική κατανομή είτε από κατανομή Rayleigh και περιορίζουμε τον διαγνωστικό έλεγχο σε αυτές τις δύο κατανομές, έτσι ώστε, αρχικά, να εκτιμήσουμε αν το σύνολο αυτό των δεδομένων παράχθηκε από εκθετική ή από Rayleigh κατανομή και στη συνέχεια να εκτιμήσουμε την παράμετρο της κατανομής αυτής, ώστε να καταλήξουμε σε μια πιο συγκεκριμένη εκτίμηση για το πιθανοκρατικό μοντέλο που μπορεί να παράξε τα δεδομένα. Κάθε ένα από αυτά τα στάδια αυτής της ανάλυσης θα εξηγηθεί αναλυτικά παρακάτω.

Στην προσπάθεια αναγνώρισης, λοιπόν, αν το σύνολο των δεδομένων (στο αρχείο .mat) προέρχεται από εκθετική κατανομή ή από κατανομή Rayleigh, συνδυάζουμε δύο μεθόδους εκτίμησης της κατανομής, οι οποίες στοχεύουν στον προσδιορισμό της μορφής της κατανομής που παράξε τα δεδομένα, προκειμένου να απαντηθεί το ερώτημα, ενώ ακόμη -όπως θα εξηγηθεί παρακάτω- η δεύτερη μέθοδος στοχεύει και στην ταυτόχρονη εκτίμηση της παραμέτρου, η οποία ορίζει επακριβώς την εκθετική ή την Rayleigh κατανομή, αντίστοιχα, από την οποία τελικά εκτιμούμε ότι προέρχονται τα δεδομένα.

Η πρώτη μέθοδος, που ακολουθούμε, είναι η Μέθοδος του Ιστογράμματος (**Histogram Method**). Η μέθοδος αυτή είναι απλή, στοχεύει, όμως, στο να δώσει μια διαισθητική ερμηνεία, μέσω του Ιστογράμματος το οποίο προκύπτει από την παρατήρηση των δεδομένων, ως προς την τιμή που λαμβάνει κάθε ένα δείγμα και τη συχνότητα με την οποία εμφανίζεται αυτή η τιμή στο σύνολο αυτό των δεδομένων, την συνάρτησης πυκνότητας πιθανότητας (κατά προσέγγιση, φυσικά, αφού πρόκειται για μια μέθοδο/εργαλείο στατιστικής ανάλυσης μελετώντας ένα πεπερασμένο σύνολο δεδομένων) της κατανομής από την οποία, πιθανώς, προέρχονται τα δεδομένα. Η συνάρτηση πυκνότητας πιθανότητας μιας κατανομής ορίζει επακριβώς την κατανομή και συνεπώς έχοντας μια εκτίμηση, μέσω του ιστογράμματος που προκύπτει από τα δεδομένα, για την συνάρτηση πυκνότητας πιθανότητας της “γεννήτριας” κατανομής, διαθέτουμε μια εκτίμηση για το ποιά είναι τελικά η κατανομή που παράγει τα δεδομένα. Εκμεταλλευόμενοι την πρότερη γνώση ότι τα δεδομένα προέρχονται είτε από εκθετική κατανομή, είτε από κατανομή Rayleigh, για τις οποίες κατανομές γνωρίζουμε την μορφή της αντίστοιχης συνάρτησης πυκνότητας πιθανότητας (ανεξάρτητα από την τιμή της παραμέτρου λ και σ , που ορίζει μια εκθετική και μια Rayleigh κατανομή, αντίστοιχα, αφού σε αυτή την μέθοδο μελέτης των δεδομένων μας ενδιαφέρει μόνο η μορφή της συνάρτησης πυκνότητας πιθανότητας αυτή καθαυτή, δηλαδή μας ενδιαφέρει μόνο η απόφαση αν τα δεδομένα προέρχονται από εκθετική ή από Rayleigh κατανομή), μπορούμε να οδηγηθούμε σε μια αρχική για την ανάλυσή μας εκτίμηση, όσον αφορά το γεγονός αν τα δεδομένα προέρχονται από εκθετική ή από Rayleigh κατανομή. Είναι γεγονός ότι όσο μεγαλύτερο είναι το πλήθος των ληφθέντων δεδομένων, τόσο πιο αξιόπιστη μπορεί να αποδειχθεί η Μέθοδος του Ιστογράμματος. Αξιοποιώντας τη μέθοδο αυτή, επιθυμούμε -δοθέντων των δεδομένων- να εκτιμήσουμε την μορφή της “γεννήτριας” κατανομής, συγκρίνοντας το ιστόγραμμα των ληφθέντων δεδομένων με την μορφή των γνωστών συναρτήσεων πυκνότητας πιθανότητας της εκθετικής (για κάποια αυθαίρετη παράμετρο λ) και της Rayleigh κατανομής (για κάποια αυθαίρετη παράμετρο σ), αντίστοιχα. Τα δεδομένα, τα οποία λήφθηκαν είναι 1000 το πλήθος και συνεπώς περιμένουμε το ιστόγραμμα να παρουσιάζει μια μορφή αρκετά “κοντά” στη μορφή μιας εκθετικής ή μιας Rayleigh κατανομής, κάνοντας την αρχική μας εκτίμηση να είναι ικανοποιητική στο δυαδικό πρόβλημα απόφαση μεταξύ εκθετικής και Rayleigh κατανομής.

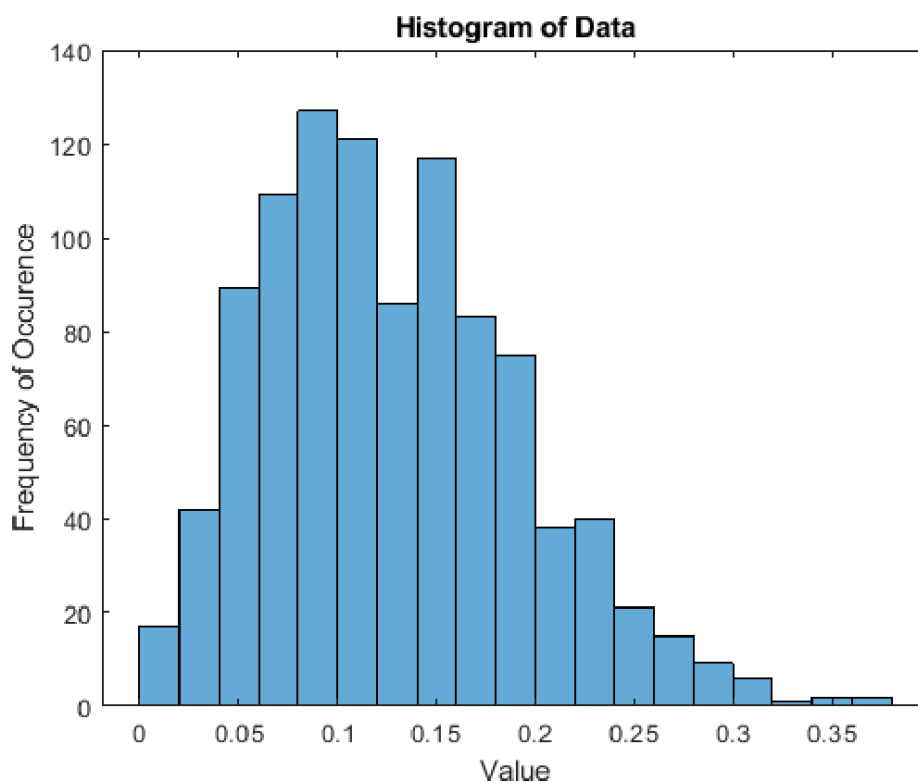
Στο Σχήμα 3.1, που ακολουθεί, παρουσιάζεται το ιστόγραμμα, το οποίο σχεδιάζουμε λαμβάνοντας τα δείγματα/δεδομένα. Το σημείο ενδιαφέροντος επικεντρώνεται μόνο στη μορφή που έχει το ιστόγραμμα αυτό και στη σύγκριση της με τις μορφές των γνωστών, από τη θεωρία πιθανοτήτων, κατανομών exponential και Rayleigh. Αναζητούμε ομοιότητες στη μορφή του ιστογράμματος με τη μορφή της συνάρτησης πυκνότητας πιθανότητας κάποιας εκθετικής ή κάποια Rayleigh κατανομής, αντίστοιχα.

Στα σχήματα 3.2 και 3.3, ακόμη, παρουσιάζονται οι συναρτήσεις πυκνότητας πιθανότητας για τρεις εκθετικές κατανομές (με παραμέτρους $\lambda=0.5$, $\lambda=1$, $\lambda=1.5$ όπως σημειώνεται στο Σχήμα 3.2) και για πέντε κατανομές Rayleigh (με παραμέτρους $\sigma=0.5$, $\sigma=1$, $\sigma=2$, $\sigma=3$, $\sigma=4$), αντίστοιχα, για τιμές του x έως 10.

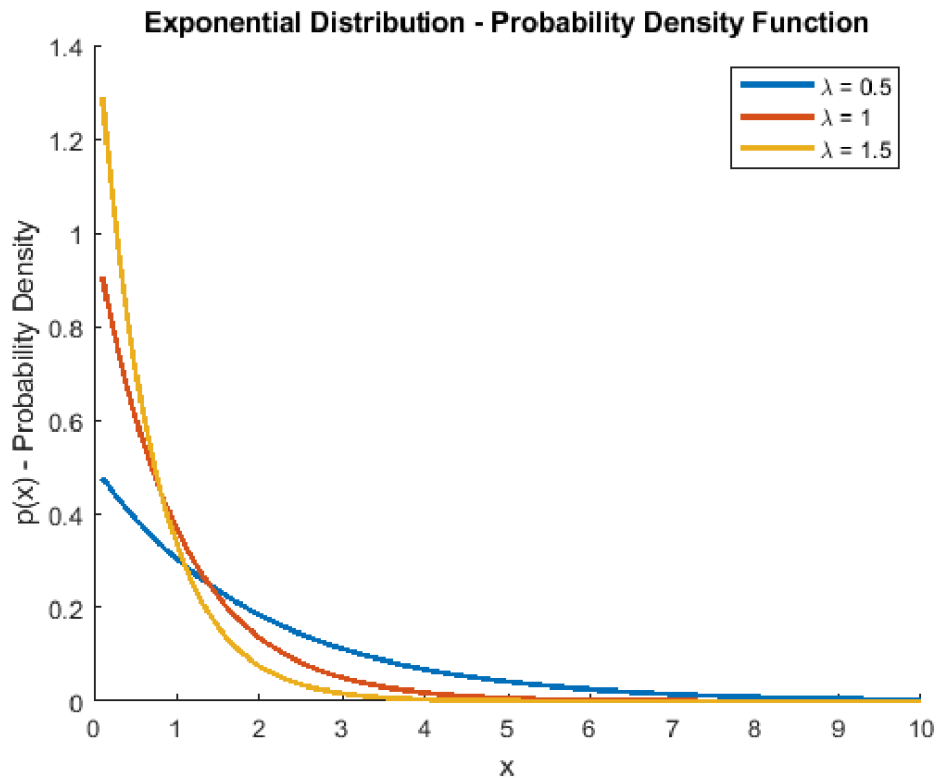
Είναι χρήσιμο, σε αυτό το σημείο, να σημειωθούν (θα αξιοποιηθούν και παρακάτω στην ανάλυση μας) οι μαθηματικοί ορισμοί των συναρτήσεων πυκνότητας πιθανότητας για την εκθετική, με παράμετρο λ και την Rayleigh κατανομή, με παράμετρο λ (είναι η παράμετρος σ που αναφέραμε παραπάνω, αλλά θα τη συμβολίζουμε με λ για το υπόλοιπο της ανάλυσης, όπως δίνεται στη σχέση (3) της εκφώνησης) :

$$\text{Εκθετική κατανομή με παράμετρο } \lambda : p(x; \lambda) = \begin{cases} \lambda \cdot \exp(-\lambda \cdot x), & x > 0 \\ 0, & \text{αλλιώς} \end{cases} \quad (1)$$

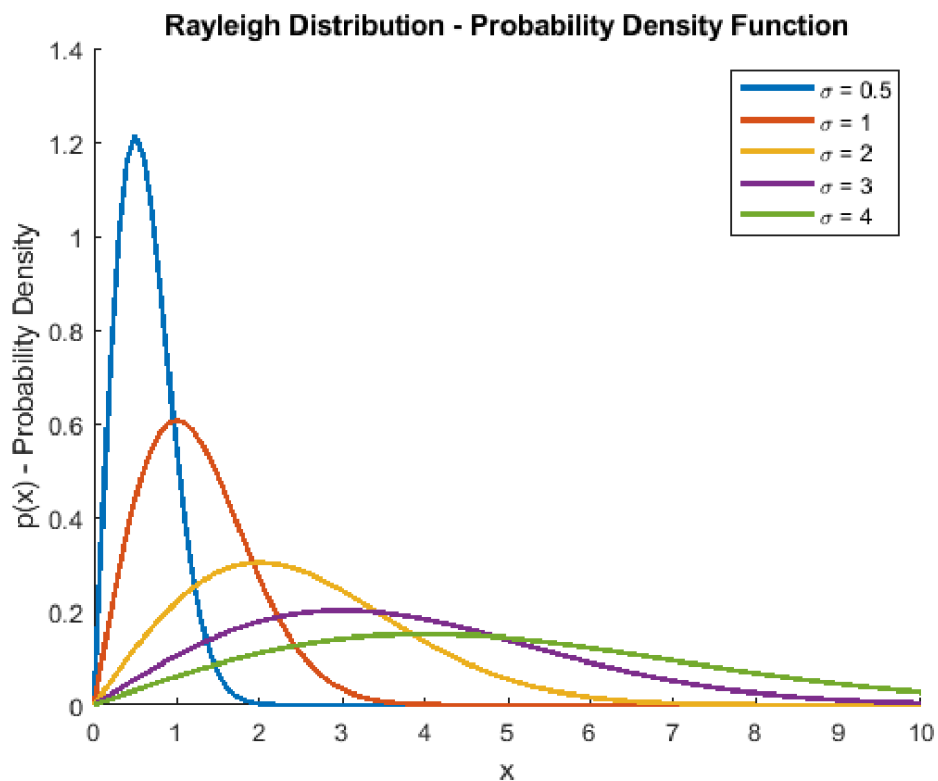
$$\text{Κατανομή Rayleigh με παράμετρο } \lambda : p(x; \lambda) = \begin{cases} \frac{x}{\lambda^2} \cdot \exp\left(\frac{-x^2}{2 \cdot \lambda^2}\right), & x \geq 0 \\ 0, & \text{αλλιώς} \end{cases} \quad (3)$$



Σχήμα 3.1 : Ιστόγραμμα των ληφθέντων δεδομένων



Σχήμα 3.2 : Συναρτήσεις Πυκνότητας Πιθανότητας εκθετικών κατανομών με παραμέτρους $\lambda=0.5$, $\lambda=1$, $\lambda=1.5$, όπως σημειώνεται στο ένθετο του διαγράμματος.



Σχήμα 3.3 : Συναρτήσεις Πυκνότητας Πιθανότητας κατανομών Rayleigh με παραμέτρους $\sigma=0.5$, $\sigma=1$, $\sigma=2$, $\sigma=3$, $\sigma=4$, όπως σημειώνεται στο ένθετο του διαγράμματος.

Όπως παρατηρούμε από τα σχήματα 3.1, 3.2, 3.3, το ιστόγραμμα των ληφθέντων δεδομένων έχει

μορφή, η οποία παρουσιάζει μεγάλη ομοιότητα με τη μορφή των συναρτήσεων πυκνότητας πιθανότητας που παρατηρούμε στο Σχήμα 3.3, δηλαδή με τη μορφή της συνάρτησης πυκνότητας πιθανότητας κάποιας κατανομής Rayleigh. Παρατηρώντας το διάγραμμα του ιστογράμματος, είναι εμφανές, ότι αυτό χαρακτηρίζεται από ορισμένα από τα βασικά χαρακτηριστικά της συνάρτησης πυκνότητας πιθανότητας κάποιας Rayleigh κατανομής. Θεωρώντας τη νοητή περιβάλλουσα στο ιστόγραμμα, είναι προφανές, ότι αυτή παρουσιάζει μοναδικό μέγιστο (σημειώνουμε ότι οι τιμές αυτές τις περιβάλλουσας, δηλαδή οι τιμές στον κατακόρυφο άξονα του ιστογράμματος αφορούν τη συχνότητα εμφάνισης δειγμάτων που λαμβάνουν αντίστοιχες τιμές x του οριζόντιου άξονα) κοντά στις μικρότερες τιμές των δεδομένων, οι οποίες συγκεντρώνονται κοντά στο μηδέν. Για τιμές των δεδομένων (στη θεωρητική συνάρτηση πυκνότητας πιθανότητας, τιμές του x) πριν από αυτό το μέγιστο η περιβάλλουσα παρουσιάζεται αύξουσα συνάρτηση, ενώ για τιμές των δεδομένων μετά από την τιμή, όπου επιτυγχάνεται το μέγιστο, παρατηρούμε εκθετική μείωση της τιμής αυτής της περιβάλλουσας (δηλαδή της τιμής που αφορά τη συχνότητα εμφάνισης δεδομένων/δειγμάτων, τα οποία λαμβάνουν τις συγκεκριμένες τιμές x). Είναι ιδιαίτερα χαρακτηριστικός ο “λοβός” που σχηματίζεται στις τιμές x κοντά στο μηδέν, που λαμβάνουν τα δείγματα, καθώς και το γεγονός ότι για τιμές x μεγαλύτερες αυτής στην οποία παρατηρούνται τα περισσότερα δείγματα (μέγιστο της περιβάλλουσας, δηλαδή επιτυγχάνεται η μέγιστη αναγραφόμενη συχνότητα εμφάνισης δειγμάτων στο σύνολο των δεδομένων) η περιβάλλουσα φθίνει εκθετικά προς το μηδέν. Για μεγάλες τιμές των δεδομένων, επίσης, η συχνότητα εμφάνισης τέτοιων τιμών είναι ιδιαίτερα μικρή, όπως ακριβώς συμβαίνει, δηλαδή, με την τιμή της συνάρτησης πυκνότητας πιθανότητας μια Rayleigh κατανομής. Αυτά τα τρία κύρια χαρακτηριστικά της μορφής της συνάρτησης πυκνότητας πιθανότητας μια Rayleigh κατανομής, τα οποία παρατηρούνται επακριβώς στο ιστόγραμμα μας οδηγούν σε μια εκτίμηση ότι μια κατανομή Rayleigh παράγει τα δεδομένα και όχι μια εκθετική κατανομή της οποίας η συνάρτηση πυκνότητας πιθανότητας είναι απλώς φθίνουσα προς το μηδέν, παρουσιάζοντας το μέγιστο στην περιοχή του $x=0$. Το βασικό χαρακτηριστικό της Rayleigh είναι ο “λοβός”, που μάλιστα είναι συγκεντρωμένος προς την περιοχή του $x=0$, καθώς και η εκθετικά φθίνουσα προς το μηδέν συνάρτηση πυκνότητας πιθανότητας για μεγαλύτερες τιμές της μεταβλητής x , έπειτα από τον “λοβό”. Αυτό είναι και το βασικό χαρακτηριστικό του ιστογράμματος. Δεδομένου ότι μιλάμε για δείγματα, άρα στατιστική, η μορφή του ιστογράμματος δεν είναι τέλεια αντιγραφή της μορφής της συνάρτησης πυκνότητας πιθανότητας μιας Rayleigh, αλλά μπορεί να παρουσιάζει λίγο μεγαλύτερη ή λίγο μικρότερη συχνότητα εμφάνισης δειγμάτων για κάποια δυνατή τιμή στο σύνολο των δεδομένων, αλλά για αυτόν ακριβώς το λόγο μιλήσαμε για νοητή περιβάλλουσα.

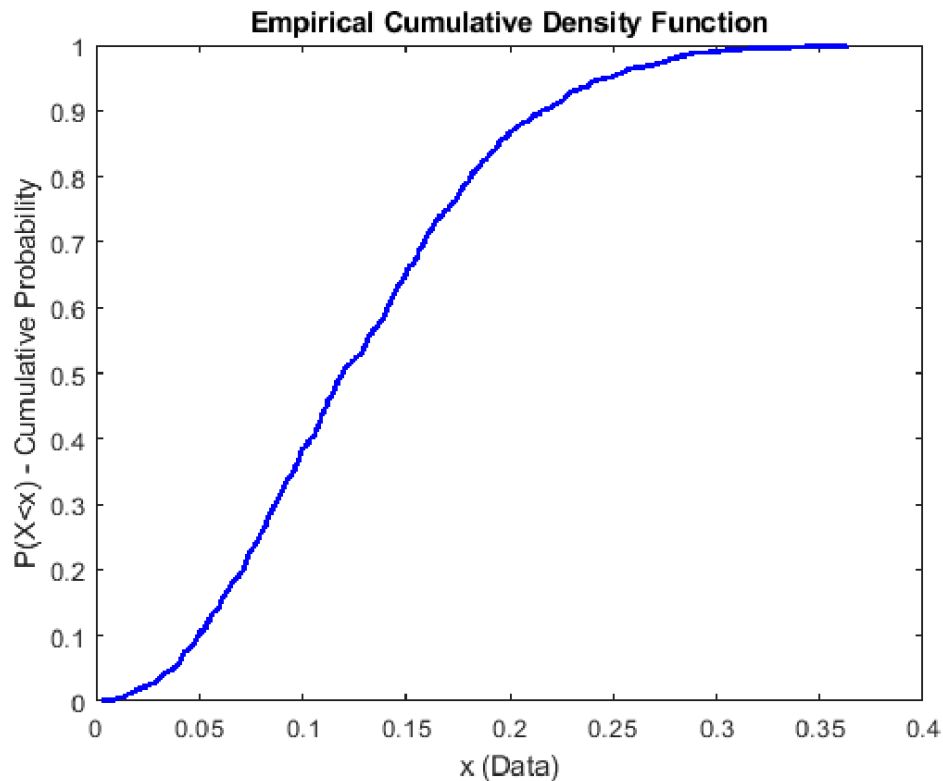
Είναι γεγονός ότι η ομοιότητα του ιστογράμματος με τη συνάρτηση πυκνότητας πιθανότητας μιας κατανομής Rayleigh, είναι ιδιαίτερα έντονη. Με δεδομένο, επιπλέον, ότι το πλήθος των ληφθέντων δεδομένων είναι 1000, δηλαδή πολύ μεγάλο και γνωρίζοντας, επίσης, ότι αυτά προέρχονται, σίγουρα, από κάποια εκθετική ή κάποια Rayleigh κατανομή, τότε η εκτίμηση που κάνουμε για τη “γεννήτρια” κατανομή είναι ότι είναι Rayleigh. Στη συνέχεια της ανάλυσης, θα αναζητήσουμε μια εκτίμηση της παραμέτρου λ που ορίζει την Rayleigh κατανομή αυτή, δηλαδή θα επιχειρήσουμε να εκτιμήσουμε την Rayleigh που -με αυτή την αρχική υπόθεση- υποθέσαμε ότι παράξε τα δεδομένα. Στο στάδιο αυτό της ανάλυσης μας θα εκτελέσουμε αυτή τη διαδικασία και για κάποια εκθετική κατανομή, όπως θα περιγράψουμε παρακάτω, προκειμένου να δείξουμε την μεγάλη ομοιότητα που παρουσιάζει η στατιστική των δεδομένων με τις θεωρητικές συναρτήσεις της Rayleigh κατανομής και όχι αυτές της εκθετικής, προσπαθώντας έτσι να επιβεβαιώσουμε ταυτόχρονα την αρχική μας αυτή υπόθεση.

Στο δεύτερο στάδιο της ανάλυσης μας, ακολουθούμε τη Μέθοδο της Εμπειρικής Αθροιστικής Συνάρτησης Κατανομής (**Empirical Cumulative Density Function Method**), σύμφωνα με την οποία αναζητούμε ομοιότητες ανάμεσα στην εμπειρική αθροιστική συνάρτηση κατανομής (ECDF) που προκύπτει από τα ληφθέντα δεδομένα και στις θεωρητικές αθροιστικές συναρτήσεις κατανομής (CDFs) της εκθετικής κατανομής (για κάποια αυθαίρετη) παράμετρο λ και της κατανομής Rayleigh

(για κάποια αυθαίρετη παράμετρο σ ή λ , όπως ορίστηκε στη σχέση (3)), αντίστοιχα. Για να εκτελέσουμε αυτή τη μέθοδο, χρειάζεται, αρχικά, να υπολογίσουμε την εμπειρική αθροιστική συνάρτηση κατανομής, καθώς επίσης και να ορίσουμε τις τιμές των παραμέτρων λ και σ για την εκθετική και την Rayleigh κατανομή, αντίστοιχα, τις οποίες παραπάνω αναφέραμε ως αυθαίρετες επιλογές, αλλά -όπως γίνεται εύκολα αντιληπτό- δεν μπορούν να είναι αυθαίρετες, αλλά θα πρέπει να επιλεγούν, έτσι ώστε τόσο η μορφή της ECDF να παρουσιάζει έντονες ομοιότητες με κάποια από τις δύο θεωρητικές CDF (της εκθετικής και της Rayleigh που θα ορίσουμε, μέσω των παραμέτρων τους δηλαδή), όσο και ό,τι αφορά τις τιμές του κατακόρυφου άξονα. Η απόδοση αυτής της μεθόδου βελτιώνεται από μια ορθή επιλογή για τις παραμέτρους λ και σ , αντίστοιχα. Δεν έχουμε κάποια ένδειξη για το ποιές τιμές των παραμέτρων να επιλέξουμε, δηλαδή για το ποιές είναι οι πιθανές εκθετικές ή Rayleigh κατανομές, αντίστοιχα, οι οποίες μπορεί να παράγουν τα δεδομένα. Τότε, θα εκτιμήσουμε αυτές τις παραμέτρους. Θα εκτελέσουμε, λοιπόν, αυτή τη μέθοδο δύο φορές κάνοντας κάθε φορά μια διαφορετική αρχική υπόθεση. Στην πρώτη φορά, θα υποθέσουμε ότι τα δεδομένα προέρχονται από εκθετική κατανομή -ασχέτως με την αρχική εκτίμηση που κάναμε για κατανομή Rayleigh ως “γεννήτρια” των δεδομένων- η οποία ορίζεται με κάποια παράμετρο λ . Θέλουμε να εκτιμήσουμε αυτή την παράμετρο, με κάποιον εκτιμητή που να παρουσιάζει καλή απόδοση. Θα χρησιμοποιήσουμε τον MLE, ο οποίος λαμβάνει υπόψη ολόκληρη την κατανομή των δεδομένων και προσπαθεί να ορίσει την παράμετρο λ , έτσι ώστε να είναι αυτή που μεγιστοποιεί τη συνάρτηση πιθανοφάνειας δεδομένων των δειγμάτων που λήφθηκαν, δηλαδή είναι ο εκτιμητής που ορίζει την κατανομή των δειγμάτων ως την καλύτερη δυνατή που ταιριάζει σε αυτά (με άλλα λόγια ο MLE δίνει εκτιμήσεις για την παράμετρο της κατανομής, που είναι στατιστικά βέλτιστες, σύμφωνα με τα ληφθέντα δεδομένα). Έχοντας τον MLE εκτιμητή για την λ , χρησιμοποιούμε αυτή την τιμή ως παράμετρο λ για να ορίσουμε την θεωρητική CDF της εκθετικής που εκτιμούμε (ως την καλύτερη δυνατή εκτίμηση εκθετικής κατανομής, βάσει των δεδομένων, μέσω MLE) ότι παράγει τα δεδομένα, με βάση αυτή την υπόθεση. Στη δεύτερη φορά, θα υποθέσουμε ότι τα δεδομένα προέρχονται από κατανομή Rayleigh η οποία ορίζεται με κάποια παράμετρο σ . Θέλουμε να εκτιμήσουμε αυτή την παράμετρο, με κάποιον εκτιμητή που να παρουσιάζει καλή απόδοση. Θα χρησιμοποιήσουμε τον MLE και σε αυτή την περίπτωση. Έχοντας τον MLE εκτιμητή για την σ , χρησιμοποιούμε αυτή την τιμή ως παράμετρο σ για να ορίσουμε την θεωρητική CDF της Rayleigh που εκτιμούμε (ως την καλύτερη δυνατή εκτίμηση Rayleigh κατανομής, βάσει των δεδομένων, μέσω MLE) ότι παράγει τα δεδομένα, με βάση αυτή την υπόθεση. Έχοντας κατά νου την αρχική υπόθεση για Rayleigh κατανομή, ως “γεννήτρια” κατανομή, αναμένουμε η ECDF να προσεγγίζει την CDF της Rayleigh που θα ορίσουμε εκτιμώντας την παράμετρο σ με τον MLE. Τότε, θα έχουμε και μια εκτίμηση για την ακριβή κατανομή που παράγει τα δεδομένα.

Όπως είπαμε, στην ανάλυση αυτή χρειάζεται, αρχικά, να υπολογίσουμε την ECDF των δεδομένων που λήφθηκαν. Η ECDF ορίζεται ως μια συνάρτηση που απεικονίζει κάθε δυνατή τιμή που συναντάται στο σύνολο των δεδομένων σε μια τιμή μικρότερη της μονάδας, η οποία αντιστοιχεί στον λόγο i/N , όπου N είναι το πλήθος των δεδομένων (δηλαδή $N=1000$) και $i=1,2,\dots,N=1000$ είναι ο δείκτης που αντιστοιχεί στο i -οστό δείγμα αν το σύνολο των δειγμάτων ταξινομηθεί κατά αύξουσα σειρά. Δηλαδή η ECDF : $\{x[1], x[2], \dots, x[N]\} \rightarrow \{1/N, 2/N, \dots, 1\}$, όπου $x[n]$, $n=1,2,\dots,N$ το n -οστό δείγμα.

Η ECDF που υπολογίστηκε από τα δεδομένα, φαίνεται στο Σχήμα 3.4.



Σχήμα 3.4 : Empirical Cumulative Density Function υπολογισμένη από τα ληφθέντα δεδομένα.

Πρώτη Υπόθεση: Τα δεδομένα προέρχονται από εκθετική κατανομή. Ακόμη, υποθέτουμε ότι τα δεδομένα είναι iid, δηλαδή ανεξάρτητα μεταξύ τους. Αυτό μοιάζει ως μια έγκυρη υπόθεση, αφού δεν έχουμε κάποια πληροφορία για συσχετίσεις μεταξύ των δεδομένων και επιπλέον, αποτελεί μια συνήθη υπόθεση στη διαδικασία εύρεσης του MLE, όταν διαθέτουμε ένα πλήθος δειγμάτων, για καθένα από τα οποία έχουμε το πιθανοκρατικό μοντέλο. Υποθέτοντας, λοιπόν, iid δείγματα στο σύνολο των ληφθέντων δεδομένων, καθένα από τα οποία ακολουθεί εκθετική κατανομή με παράμετρο λ , θα ισχύει, σύμφωνα με την (1), για την συνάρτηση πυκνότητας πιθανότητας του δείγματος $x[n]$, $n=1,2,\dots,N$ από τα δεδομένα :

$$p(x[n]; \lambda) = \begin{cases} \lambda \cdot \exp(-\lambda \cdot x[n]), & x[n] > 0 \\ 0, & \text{αλλιώς} \end{cases},$$

παραμετροποιημένη ως προς την παράμετρο λ .

Για την κοινή συνάρτηση πυκνότητας πιθανότητας έχουμε (Αν $x=[x[1], x[2], \dots, x[N]]^T$ το διάνυσμα των $N=1000$ παρατηρήσεων) :

$$p(x; \lambda) = \prod_{n=1}^N p(x[n]; \lambda) = \lambda^N \cdot \exp\left(-\lambda \sum_{n=1}^N x[n]\right)$$

Ο MLE είναι η τιμή της λ που μεγιστοποιεί την συνάρτηση πιθανοφάνειας $p(x; \lambda)$ για δεδομένο x , το οποίο ακριβώς ορίζουν τα ληφθέντα δεδομένα. Ισοδύναμα ο MLE μεγιστοποιεί την

$$\log\text{-likelihood συνάρτηση} : \log(p(x; \lambda)) = N \cdot \log(\lambda) - \lambda \sum_{n=1}^N x[n]$$

Για την μεγιστοποίηση αυτής ως προς λ , παραγωγίζουμε ως προς λ και θέτουμε την παράγωγο ίση με μηδέν.

$$\text{Είναι: } \frac{d}{d\lambda} \left(\log(p(x; \lambda)) \right) = \frac{N}{\lambda} - \sum_{n=1}^N x[n] = 0 \Rightarrow \lambda_{MLE}^{\wedge} = \frac{N}{\sum_{n=1}^N x[n]}$$

Τότε, ο MLE : $\lambda_{MLE}^{\wedge} = \frac{N}{\sum_{n=1}^N x[n]} = \frac{1}{\bar{x}}$, όπου \bar{x} είναι η μέση τιμή των δεδομένων.

Στον παρακάτω πίνακα φαίνονται τα αποτελέσματα από τον υπολογισμό του MLE για την παράμετρο λ .

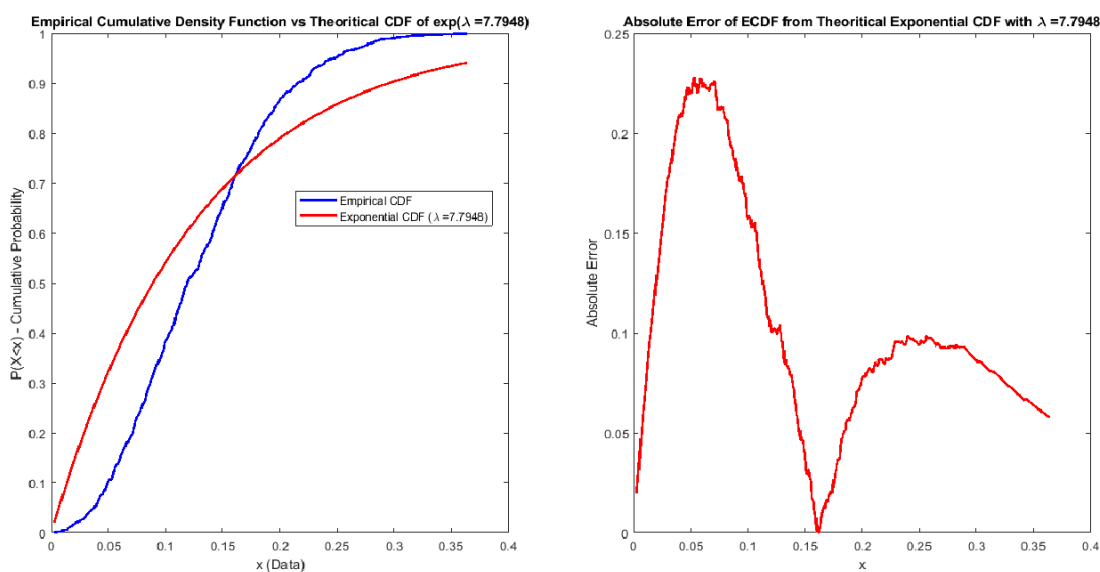
```
----- Assumption N01: Dataset comes from an exponential distribution -----
Estimate the parameter  $\lambda$  of the exponential distribution with MLE.
The number of collected data in the dataset is: N=1000
The MLE for the parameter  $\lambda$  considering the collected data:  $\lambda_{MLE}=7.7948$ 
```

Πίνακας 3.1

Τότε, χρησιμοποιούμε ως παράμετρο λ τη εκθετικής κατανομής, από την οποία υποθέσαμε ότι προέρχονται τα δεδομένα, τον MLE που υπολογίστηκε από τα δεδομένα, δηλαδή $\lambda = 7.7948$.

Υπολογίζουμε την θεωρητική αθροιστική συνάρτηση κατανομής της εκθετικής κατανομής με παράμετρο $\lambda=7.7948$, καθώς και το απόλυτο σφάλμα της ECDF από αυτή για τις τιμές που παρατηρούνται στο σύνολο των δειγμάτων που λήφθηκαν.

Το γράφημα της θεωρητικής CDF, το οποίο παρουσιάζεται μαζί με το αντίστοιχο γράφημα της ECDF για τις δυνατές τιμές που παρατηρούνται στα δεδομένα, καθώς και αυτό του απολύτου σφάλματος, σε ένα εύρος τιμών που παρατηρούνται στα δεδομένα παρουσιάζονται στο Σχήμα 3.5.



Σχήμα 3.5 : (Αριστερά): Εμπειρική Αθροιστική Συνάρτηση Κατανομής από τα δεδομένα vs Θεωρητική Αθροιστική Συνάρτηση Κατανομής της εκθετικής κατανομής $\exp(7.7948)$. (Δεξιά) : Απόλυτο Σφάλμα μεταξύ ECDF και Θεωρητικής CDF της $\exp(7.7948)$.

Όπως παρατηρούμε από το παραπάνω σχήμα, η ECDF παρουσιάζει ομοιότητες με την CDF της εκθετικής κατανομής που ορίσαμε, ανά διαστήματα τιμών στο εύρος τιμών που παρατηρούνται στα δεδομένα. Η μορφή των δύο συναρτήσεων δεν φαίνεται να ταιριάζει απόλυτα, ιδίως για τις μικρότερες τιμές του οριζόντιου άξονα, γεγονός που αποτυπώνεται και στο γράφημα του απολύτου σφάλματος. Παρατηρώντας την ECDF στις τιμές που απεικονίζονται στον οριζόντιο άξονα, παρουσιάζει ένα σημείο καμπής, το οποίο σημαίνει μια αλλαγή στην κυρτότητα της συνάρτησης. Για τις μικρές τιμές του οριζόντιου άξονα η ECDF είναι κυρτή και αύξουσα, μέχρι το σημείο καμπής, όπου γίνεται κοίλη και αύξουσα. Μετά από το σημείο καμπής η μορφή της ECDF μοιάζει με αυτή της θεωρητικής CDF, όμως οφείλουμε να παρατηρήσουμε τη χαρακτηριστική αυτή

διαφορά μεταξύ των δύο. Μάλιστα, αφού τα δεδομένα προέρχονται, σίγουρα, είτε από εκθετική, είτε από Rayleigh κατανομή, θα έπρεπε η υπόθεση ότι τα δεδομένα προέρχονται από εκθετική κατανομή να μας οδηγούν σε ορθές επιλογές πιθανοκρατικών μοντέλων για τις παρατηρήσεις $x[n]$ και άρα ο MLE εκτιμητής να όριζε, όντως, μια καλή εκτίμηση της “γεννήτριας” κατανομής. Εδώ οι διαφορές αυτή στη μορφή των δύο συναρτήσεων είναι αξιοσημείωτη και χαρακτηριστική.

Δεύτερη Υπόθεση: Τα δεδομένα προέρχονται από κατανομή Rayleigh. Ακόμη, υποθέτουμε ότι τα δεδομένα είναι iid, δηλαδή ανεξάρτητα μεταξύ τους. Αυτό μοιάζει ως μια έγκυρη και συνήθη στη διαδικασία εύρεσης του MLE, όταν διαθέτουμε ένα πλήθος δειγμάτων, για καθένα από τα οποία έχουμε το πιθανοκρατικό μοντέλο, υπόθεση, όπως εξηγήσαμε προηγουμένως. Υποθέτοντας, λοιπόν, iid δείγματα στο σύνολο των ληφθέντων δεδομένων, καθένα από τα οποία ακολουθεί κατανομή Rayleigh με παράμετρο λ , θα ισχύει, σύμφωνα με την (3), για την συνάρτηση πυκνότητας πιθανότητας του δείγματος $x[n]$, $n=1,2,\dots,N$ από τα δεδομένα :

$$p(x[n]; \lambda) = \begin{cases} \frac{x[n]}{\lambda^2} \cdot \exp\left(-\frac{x[n]^2}{2\lambda^2}\right), & x[n] \geq 0, \\ 0, & \text{αλλιώς} \end{cases}$$

παραμετροποιημένη ως προς την παράμετρο λ .

Για την κοινή συνάρτηση πυκνότητας πιθανότητας έχουμε (Αν $x=[x[1], x[2], \dots, x[N]]^T$ το διάνυσμα των $N=1000$ παρατηρήσεων) :

$$p(x; \lambda) = \prod_{n=1}^N p(x[n]; \lambda) = \prod_{n=1}^N \frac{x[n]}{\lambda^2} \cdot \exp\left(-\frac{1}{2\lambda^2} \sum_{n=1}^N x[n]^2\right)$$

Ο MLE είναι η τιμή της λ που μεγιστοποιεί την συνάρτηση πιθανοφάνειας $p(x; \lambda)$ για δεδομένο x , το οποίο ακριβώς ορίζουν τα ληφθέντα δεδομένα. Ισοδύναμα ο MLE μεγιστοποιεί την

$$\log\text{-likelihood συνάρτηση} : \log(p(x; \lambda)) = \log\left(\prod_{n=1}^N x[n]\right) - 2N \cdot \log(\lambda) - \frac{1}{2\lambda^2} \sum_{n=1}^N x[n]^2$$

Για την μεγιστοποίηση αυτής ως προς λ , παραγωγίζουμε ως προς λ και θέτουμε την παράγωγο ίση με μηδέν.

$$\text{Είναι: } \frac{d}{d\lambda} \log(p(x; \lambda)) = -2 \frac{N}{\lambda} - \frac{1}{\lambda^3} \cdot \sum_{n=1}^N x[n]^2 = 0 \Rightarrow \dots \Rightarrow \lambda_{MLE}^{\wedge} = \sqrt{\left(\frac{\sum_{n=1}^N x[n]^2}{N} \right)}$$

$$\text{Τότε, ο MLE : } \lambda_{MLE}^{\wedge} = \sqrt{\left(\frac{\sum_{n=1}^N x[n]^2}{N} \right)}$$

Στον παρακάτω πίνακα φαίνονται τα αποτελέσματα από τον υπολογισμό του MLE για την παράμετρο λ (Η λ για την Rayleigh είναι η σ , στη βιβλιογραφία, όπως αναφέραμε).

```
----- Assumption No2: Dataset comes from a Rayleigh distribution -----
Estimate the parameter  $\sigma$  of the Rayleigh distribution with MLE.
The number of collected data in the dataset is: N=1000
The MLE for the parameter  $\sigma$  considering the collected data:  $\sigma_{MLE}=0.1437$ 
```

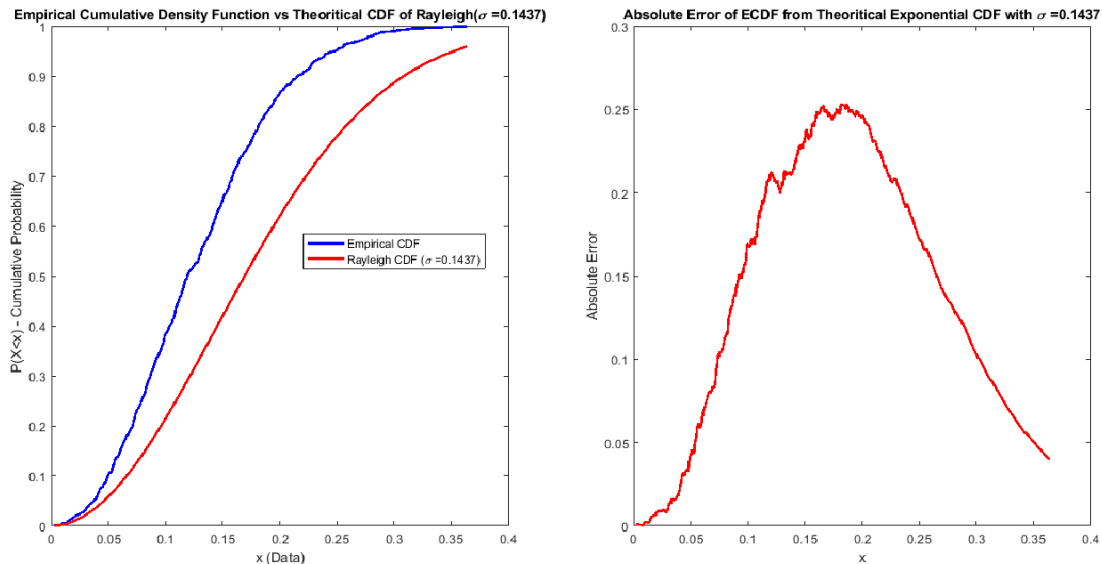
Πίνακας 3.2

Τότε, χρησιμοποιούμε ως παράμετρο λ τη εκθετικής κατανομής, από την οποία υποθέσαμε ότι προέρχονται τα δεδομένα, τον MLE που υπολογίστηκε από τα δεδομένα, δηλαδή $\lambda = 7.7948$.

Υπολογίζουμε την θεωρητική αθροιστική συνάρτηση κατανομής της κατανομής Rayleigh με

παράμετρο $\lambda=0.1437$, καθώς και το απόλυτο σφάλμα της ECDF από αυτή για τις τιμές που παρατηρούνται στο σύνολο των δειγμάτων που λήφθηκαν.

Το γράφημα της θεωρητικής CDF, το οποίο παρουσιάζεται μαζί με το αντίστοιχο γράφημα της ECDF για τις δυνατές τιμές που παρατηρούνται στα δεδομένα, καθώς και αυτό του απολύτου σφάλματος, σε ένα εύρος τιμών που παρατηρούνται στα δεδομένα παρουσιάζονται στο Σχήμα 3.5 .



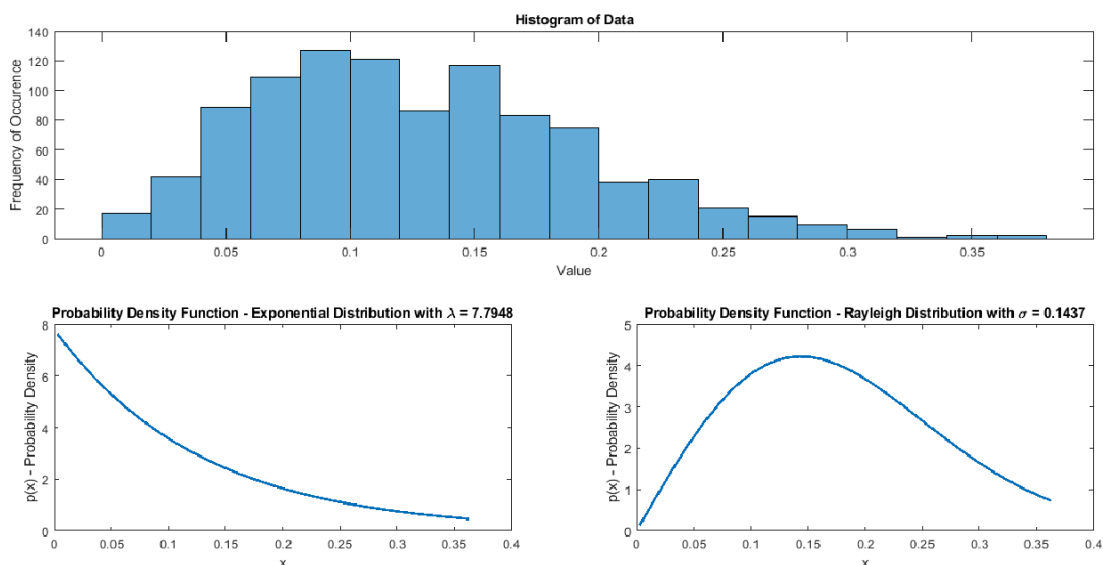
Σχήμα 3.6 : (Αριστερά):Εμπειρική Αθροιστική Συνάρτηση Κατανομής από τα δεδομένα vs Θεωρητική Αθροιστική Συνάρτηση Κατανομής της κατανομής Rayleigh με παράμετρο $\sigma=0.1437$. (Δεξιά) : Απόλυτο Σφάλμα μεταξύ ECDF και Θεωρητικής CDF της κατανομής Rayleigh $\sigma=0.1437$.

Όπως παρατηρούμε από το παραπάνω σχήμα, η ECDF παρουσιάζει έντονες ομοιότητες με την CDF της κατανομής Rayleigh με παράμετρο $\lambda=0.1337$ που ορίσαμε, για όλες τις τιμές στο εύρος τιμών που παρατηρούνται στα δεδομένα. Είναι χαρακτηριστικό της CDF της Rayleigh κατανομής ότι είναι κυρτή και αύξουσα για τις μικρότερες τιμές x του οριζόντιου άξονα, μέχρι ένα σημείο καμπής, όπου γίνεται κοίλη και αύξουσα για τις επόμενες τιμές του οριζόντιου άξονα. Την ίδια ακριβώς συμπεριφορά παρατηρούμε από την ECDF, όπως έχουμε ήδη σχολιάσει. Ακόμη, η μορφή τους είναι σχεδόν πανομοιότυπη. Η διαφορά που παρατηρούμε μεταξύ των δύο αφορά τις τιμές που λαμβάνει η κάθε συνάρτηση για μια συγκεκριμένη τιμή από το σύνολο των δεδομένων. Είναι γεγονός ότι η ECDF παρουσιάζεται πάντοτε πάνω από την θεωρητική CDF, έχοντας όμως την ίδια μορφή με εκείνη. Για τον λόγο αυτό, παρατηρούμε το απόλυτο σφάλμα να λαμβάνει για κάποιες τιμές του οριζόντιου άξονα τιμές παρόμοιες (άλλοτε και μεγαλύτερες, άλλοτε ,όμως, μικρότερες) σε σχέση με αυτές του απόλυτου σφάλματος του σχήματος 3.5 . Αυτό, όμως, είναι κάτι το οποίο είναι λογικό, αφού η CDF της Rayleigh κατανομής, όπως γνωρίζουμε από τη θεωρία πιθανοτήτων (δεν αποτυπώνεται στο διάγραμμα, αφού αυτό παρουσιάζεται μόνο στο εύρος τιμών που ορίζεται η ECDF) τείνει ασυμπτωτικά στη μονάδα, όταν η τυχαία μεταβλητή x λαμβάνει τιμή που τείνει στο άπειρο, θεωρητικά, σε αντίθεση με την ECDF για την οποία έχουμε ορίσει να λαμβάνει την τιμή 1 στη μέγιστη τιμή που παρατηρείται στο σύνολο των δεδομένων. Αυτό, στο οποίο αξίζει να επικεντρώσουμε το ενδιαφέρον είναι η ομοιότητες στη μορφή των δύο συναρτήσεων. Μάλιστα, αφού τα δεδομένα προέρχονται, σίγουρα, είτε από εκθετική, είτε από Rayleigh κατανομή και η υπόθεση ότι τα δεδομένα προέρχονται από κατανομή Rayleigh, φαίνεται να μας οδηγούν σε ορθές επιλογές πιθανοκρατικών μοντέλων για τις παρατηρήσεις $x[n]$ και άρα ο MLE εκτιμητής ορίζει, όντως, μια καλή εκτίμηση της “γεννήτριας” κατανομής. Το γεγονός, ακόμη, ότι και οι τιμές του δεν διαφέρουν πολύ δείχνει ότι η εκτίμηση της παραμέτρου λ της Rayleigh κατανομής με τον MLE είναι ικανοποιητική για να ορίσει την κατανομή που παράγει τα δεδομένα.

Τελικά, φαίνεται και από αυτή τη μέθοδο ότι τα δεδομένα προέρχονται από μια κατανομή Rayleigh, όπως είχαμε, αρχικά, εκτιμήσει μέσω της Μεθόδου του Ιστογράμματος. Σε αυτό το δεύτερο στάδιο της ανάλυσης, εκτιμήσαμε ακόμη τις βέλτιστες βάσει των ληφθέντων δειγμάτων, παραμέτρους για τις κατανομές : εκθετική και Rayleigh, από τις οποίες μπορεί να προέρχονται τα δεδομένα. Η υπόθεση ότι προέρχονται από εκθετική δεν μας οδηγήσει σε ένα επιθυμητό συμπέρασμα, σε αντίθεση με την υπόθεση ότι προέρχονται από Rayleigh κατανομή. Μάλιστα, αυτή είναι η καλύτερη δυνατή εκτίμηση, αξιοποιώντας τον MLE που μπορούμε να κάνουμε για την παράμετρο της Rayleigh κατανομής από την οποία υποθέτουμε πως προήλθαν τα δεδομένα.

Εκτιμούμε ότι προήλθαν από Rayleigh κατανομή και μάλιστα εκτιμήσαμε και την παράμετρο της, $\lambda=0.1437$.

Στο Σχήμα 3.7, παρουσιάζεται ξανά το ιστογράμμο των δεδομένων, καθώς και οι συναρτήσεις πυκνότητας πιθανότητας, τόσο της εκθετικής κατανομής που ορίζεται από την παράμετρο λ που υπολογίζεται μέσω του MLE, ως $\lambda_{MLE}^{\wedge}=7.7948$, όσο και της κατανομής Rayleigh που ορίζεται από την παράμετρο λ , η οποία εκτιμάται μέσω του MLE, αντίστοιχα ως $\lambda_{MLE}^{\wedge}=0.1437$. Όπως παρατηρούμε, είναι εμφανής η μεγάλη της μορφής του ιστογράμματος με την συνάρτηση πυκνότητας πιθανότητας της Rayleigh με παράμετρο $\lambda=0.1437$.



Σχήμα 3.7 : (Πάνω) : Ιστογράμμο των δεδομένων. (Κάτω Αριστερά) : Συνάρτηση Πυκνότητας Πιθανότητας της $exp(7.7948)$. (Κάτω Δεξιά) : Συνάρτηση Πυκνότητας Πιθανότητας της Rayleigh με παράμετρο $\lambda=0.1437$.