Άσκηση 1

Θεωρήστε ένα ζάρι Z1 και z1 μια διακριτή τυχαία μεταβλητή που συμβολίζει το αποτέλεσμα της ρίψης του Z1.

α) Υπολογίστε τη συνάρτηση πυκνότητας πιθανότητας του f(z1) = p(Z=z1).

β) Υπολογίστε την μέση τιμή και τη μεταβλητότητα της τυχαίας μεταβλητής z1.

γ) Υπολογίστε τη skewness και την kurtosis της τυχαίας μεταβλητής z1.

Απάντηση 1(α):

Η συνάρτηση πυκνότητας πιθανότητας θα είναι :

$$f(z1) = \begin{Bmatrix} \frac{1}{6}, & z1 = 1,2,3,4,5,6 \\ 0, & \alpha\lambda\lambda o\acute{\upsilon} \end{Bmatrix}$$

Απάντηση 1(β):

η μέση τιμή της τυχαίας μεταβλητής z1 θα είναι :

$$E[Z1] = \sum_{k=1}^{6} z1_k \, P_{z1}(z1_k) = 1\frac{1}{6} + 2\frac{1}{6} + 3\frac{1}{6} + 4\frac{1}{6} + 5\frac{1}{6} + 6\frac{1}{6} = \frac{21}{6} = 3.5$$

Ενώ η μεταβλητότητα θα είναι :

$$Var[Z1] = E[(Z1 - E[Z1])^2] = E[(Z1 - 3.5)^2] = \sum_{k=1}^{6} (z1_k - 3.5)^2 \, P_{z1}(z1_k) =$$

$$= \frac{1}{6}[\ (1 - 3.5)^2 + \frac{1}{6}(2 - 3.5)^2 + \frac{1}{6}(3 - 3.5)^2 + \frac{1}{6}(4 - 3.5)^2 + \frac{1}{6}(5 - 3.5)^2 +$$

$$+ \frac{1}{6}(6 - 3.5)^2 = \frac{1}{6}(6.25 + 2.25 + 0.25 + 0.25 + 2.25 + 6.25) = 2.9167$$

Απάντηση 1(γ):

το skewness θα είναι:

$$Skewness[Z1] = E\left[\left(\frac{z1 - E[Z1]}{\sqrt{Var[Z1]}}\right)^3\right] = E\left[\left(\frac{z1 - 3.5}{1.7078}\right)^3\right] = \sum_{k=1}^{6}\left(\frac{z1_k - 3.5}{1.7078}\right)^3 P_{z1}(z1_k) =$$

$$\frac{1}{6}\left[\left(\frac{1 - 3.5}{1.7078}\right)^3 + \left(\frac{2 - 3.5}{1.7078}\right)^3 + \left(\frac{3 - 3.5}{1.7078}\right)^3 + \left(\frac{4 - 3.5}{1.7078}\right)^3 + \left(\frac{5 - 3.5}{1.7078}\right)^3 + \left(\frac{6 - 3.5}{1.7078}\right)^3\right] = 0$$

Το kurtosis θα είναι:

$$kurtosis[Z1] = E\left[\left(\frac{z1 - E[Z1]}{\sqrt{Var[Z1]}}\right)^4\right] = E\left[\left(\frac{z1 - 3.5}{1.7078}\right)^4\right] = \sum_{k=1}^{6}\left(\frac{z1_k - 3.5}{1.7078}\right)^4 P_{z1}(z1_k) =$$

$$= \frac{1}{6}\left[\left(\frac{1 - 3.5}{1.7078}\right)^4 + \left(\frac{2 - 3.5}{1.7078}\right)^4 + \left(\frac{3 - 3.5}{1.7078}\right)^4 + \left(\frac{4 - 3.5}{1.7078}\right)^4 + \left(\frac{5 - 3.5}{1.7078}\right)^4 + \left(\frac{6 - 3.5}{1.7078}\right)^4\right] =$$

$$= \frac{1}{6}[\ 4.5921 + 0.5951 + 0.0073 + 0.0073 + 0.5951 + 4.5921] = 1.7315$$

Άσκηση 2

Το MATLAB μπορεί να δημιουργεί τυχαίες πραγματικές μεταβλητές στο διάστημα [0,1] με τη βοήθεια της εντολής rand(M,N), όπου M, N το μέγεθος του πίνακα που παράγει.

α) Χρησιμοποιήστε την παραπάνω εντολή για να φτιάξετε ένα διάνυσμα 1xN που να έχει N τυχαίες ρίψεις του ζαριού Z1.

β) Απεικονίστε την προσεγγιστική συνάρτηση πυκνότητας πιθανότητας (ιστόγραμμα) των ρίψεων του ζαριού για N=20, 100 και 1000. Ποια κατανομή φαίνεται να ακολουθεί.

γ) Χρησιμοποιήστε τις αντίστοιχες εντολές του MATLAB για να υπολογίσετε τη μέση τιμή, μεταβλητότητα, skewness και την kurtosis των N ρίψεων για τιμές του N=10, 20, 50, 100, 500 και 1000.

δ) Για ποιες τιμές πλήθους ρίψεων φαίνονται οι πειραματικές τιμές να προσεγγίζουν τις θεωρητικές που έχετε υπολογίσει στο 1.

ε) Για ποια τιμή πλήθους ρίψεων θα λέγαμε ότι ξεκινάμε να έχουμε wide-sense stationarity ?

Απάντηση 2(α):

χρησιμοποιώντας την εντολή rand και τι εντολές που φαίνονται δίπλα, μπορούμε να φτιάξουμε ένα διάνυσμα μεγέθους N του οποίου τα στοιχεία να παίρνουν μόνο τις τιμές 1 , 2 , 3 , 4 , 5 , 6 . Παραδείγματος χάρη για N=20 θα παίρναμε το διάνυσμα :

```
Columns 1 through 9

  6     2     6     1     4     2     4     4     1

Columns 10 through 18

  3     2     6     5     1     3     2     2     4

Columns 19 through 20

  5     3
```

(Αντί για όλες τις εντολές που χρειάστηκαν θα μπορούσαμε απλά να χρησιμοποιήσουμε την εντολή randi(6, [1,N]) η οποία θα μας έδινε και πάλι ένα διάνυσμα μεγέθους N από ακέραιους 1-6 αλλά ζητούνταν συγκεκριμένα η rand)

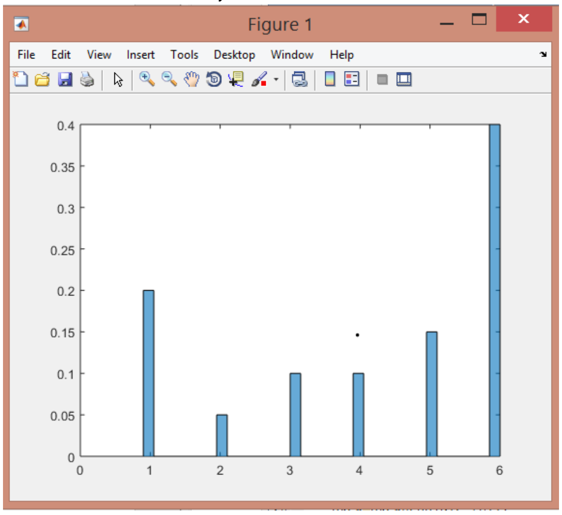
```
clear;
clc;
rng('default');
N=20;
Z1=0.5+6*rand(1,N);
for i=1:N
    if Z1(i)<=1.5
        Z1(i)=1;
    elseif Z1(i)<=2.5
        Z1(i)=2;
    elseif Z1(i)<=3.5
        Z1(i)=3;
    elseif Z1(i)<=4.5
        Z1(i)=4;
    elseif Z1(i)<=5.5
        Z1(i)=5;
    elseif Z1(i)<=6.5
        Z1(i)=6;
    end
end

disp(Z1);
```

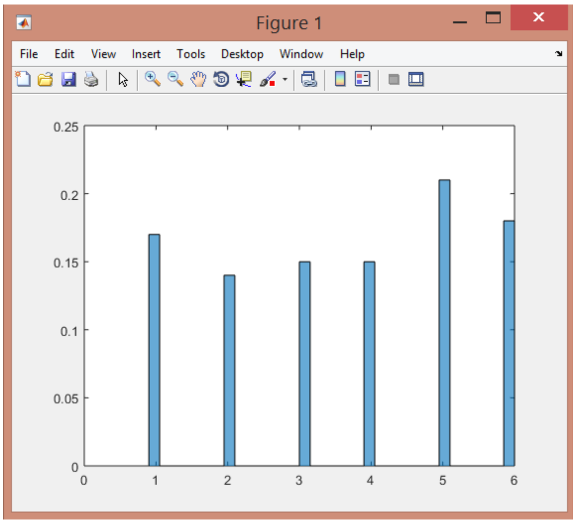
Απάντηση 2(β):

η προσεγγιστική συνάρτηση πυκνότητας πιθανότητας θα είναι :

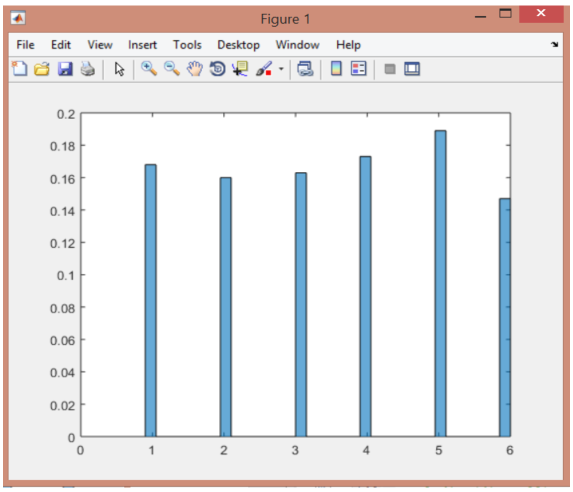
για N=20 :



Για N=100 :



Για N=1000:



Η προσεγγιστική συνάρτηση πυκνότητας πιθανότητας φαίνεται να ακολουθεί την ομοιόμορφη κατανομή .

Απάντηση 2(γ):

η μέση τιμή, μεταβλητότητα, skewness και η kurtosis θα είναι :

	mean	variance	skewness	kurtosis
10	2.4	2.9333	1.2981	3.2328
20	3.6	2.4632	-0.2313	1.6625
50	3.98	3.2853	0.2377	1.5244
100	3.27	2.9668	0.2089	1.7509
500	3.4480	2.8330	0.0156	1.7717
1000	3.494	2.9449	-0.0183	1.7217

Απάντηση 2(δ):

Όσο το πλήθος αυξάνεται οι πειραματικές τιμές φαίνεται να πλησιάζουν τις θεωρητικές που είχαμε βρει στο πρώτο ερώτημα. Για χίλιες επαναλήψεις παίρνουμε τις πιο κοντινές, στις θεωρητικές, τιμές με απόκλιση της τάξης των 2 δεκαδικών ψηφίων.

Απάντηση 2(ε):

κάνουμε 6 επαναλήψεις για κάθε N :

N=10:

```
Average =

    3.3000    3.1000    2.9000    4.2000    3.2000    2.4000

Variance =

    2.9000    2.1000    2.1000    3.9556    3.7333    3.6000

Skewness =

   -0.2078    0.5126    0.1801   -0.4681    0.3858    0.9342

Kurtosis =

    1.3995    2.7025    1.8963    1.5512    1.6556    2.2812
```

N=20:

```
Average =

    3.5500    3.3500    3.4000    3.6500    4.1500    2.9500

Variance =

    1.8395    3.1868    2.5684    2.5553    2.5553    3.3132

Skewness =

   -0.8224   -0.0896   -0.0504   -0.2005   -0.8053    0.5591

Kurtosis =

    2.5432    1.6115    1.9046    1.7800    2.4011    1.8903
```

N=50

```
Average =

    3.5400    3.9000    3.5000    3.2400    3.2000    3.5400

Variance =

    3.6004    3.2347    2.4184    2.7167    2.8571    3.0290

Skewness =

   -0.0624   -0.4231   -0.0493    0.3290    0.1947    0.0261

Kurtosis =

    1.5916    1.8982    1.8947    1.7528    1.7751    1.6448
```

N=100:

```
Average =

    3.5300    3.5200    3.6300    3.5600    3.4800    3.6400

Variance =

    3.1607    3.1208    3.0233    3.1176    2.9390    3.1620

Skewness =

    0.0805   -0.0001   -0.1037   -0.0223    0.0281   -0.1503

Kurtosis =

    1.6170    1.7050    1.7121    1.6298    1.7573    1.7358
```

N=200:

```
Average =

    3.4300    3.3500    3.5700    3.6550    3.3650    3.6200

Variance =

    2.7991    2.8116    2.4172    3.2723    3.0169    3.2519

Skewness =

   -0.0069    0.0662    0.0220   -0.1113    0.1153   -0.0973

Kurtosis =

    1.7277    1.7551    1.9861    1.6425    1.7471    1.5698
```

N=500:

```
Average =

    3.4060    3.4440    3.4080    3.6300    3.4520    3.4880

Variance =

    2.8068    2.8325    2.9033    3.0231    2.9456    3.0520

Skewness =

    0.0357    0.0503    0.0384   -0.1294    0.0019    0.0212

Kurtosis =

    1.7670    1.7701    1.7382    1.6994    1.7369    1.7073
```

Βλέπουμε ότι N=500 έχουμε κυρίως μεταβολές τις τάξης του δευτέρου δεκαδικού ψηφίου. Παρόλο που βλέπουμε ότι υπάρχει μία μικρή εξαίρεση (λίγο μεγαλύτερη αποκλίση στην μέση τιμή της τρίτης επανάληψης) μπορούμε να πούμε ότι για N=500 ξεκινάμε να έχουμε wide-sense stationarity. Θα μπορούσαμε με μεγάλη επιφύλαξη ότι ίσως θα μπορούσε να ξεκινάει και από τα 200 δείγματα ωστόσο είναι πιο φρόνιμο να διαλέξουμε τα 500 δείγματα σαν απάντηση.

Άσκηση 3

Χρησιμοποιείστε τον παραπάνω κώδικα για να παράγετε 1000 ρίψεις του ζαριού Z1 και άλλες ρίψεις του ζαριού Z2.

α) Υπολογίστε και απεικονίστε την κοινή συνάρτηση πυκνότητας πιθανότητας (joint pdf) f(z1,z2).

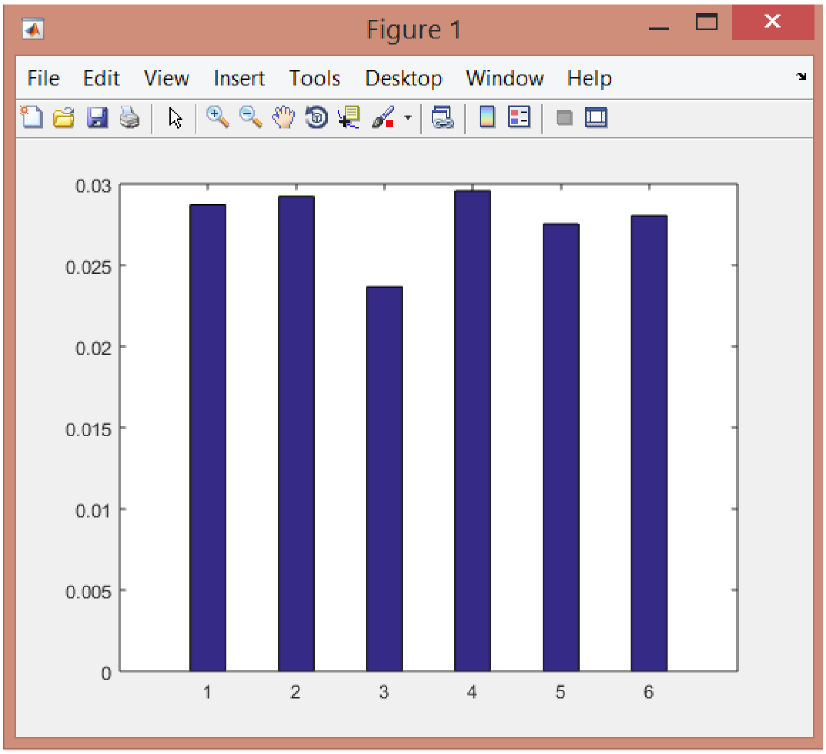
β) Ορίστε μια νέα τυχαία μεταβλητή y ως το άθροισμα των τιμών των 2 ζαριών y = z1+z2. Απεικονίστε την συνάρτηση πυκνότητας πιθανότητας της νέας μεταβλητής. Μπορείτε να εξηγήσετε το αποτέλεσμα ; (Αναζητήστε την ιδιότητα των pdf ότι αν y = z1+z2, τότε f(y)= f(z1)*f(z2), όπου * ο τελεστής της συνέλιξης)

Απάντηση 3(α):

Για να βρούμε την κοινή συνάρτηση πυκνότητας πιθανότητας θα εκμεταλλευτούμε το γεγονός ότι η τιμή που θα φέρει το ζάρι 1 είναι ανεξάρτητη από την τιμή που θα φέρει το ζάρι 2. Οπότε η joint συνάρτηση πυκνότητας πιθανότητας μπορεί να προκύψει απλά ως γινόμενο των 2 επιμέρους συναρτήσεων πυκνότητας πιθανότητας. Η συνάρτηση πυκνότητας πιθανότητας για χίλιες τυχαίες ρίψεις του ζαριού 1 και χίλιες τυχαίες ρίψεις του ζαριού 2 θα είναι joint probability f_{Z1Z2} :

$$f_{Z1Z2}(z1,z2) = f_{Z1}(z1)f_{Z2}(z2) = \begin{Bmatrix} 0.0287, & z1 = 1\,, & z2 = 1 \\ 0.0292, & z1 = 2, & z2 = 2 \\ 0.0237, & z1 = 3, & z2 = 3 \\ 0.0296, & z1 = 4, & z2 = 4 \\ 0.0275, & z1 = 5, & z2 = 5 \\ 0.0281, & z1 = 6, & z2 = 6 \\ & 0\,, \alpha\lambda\lambda o\acute{\upsilon} & \end{Bmatrix}$$

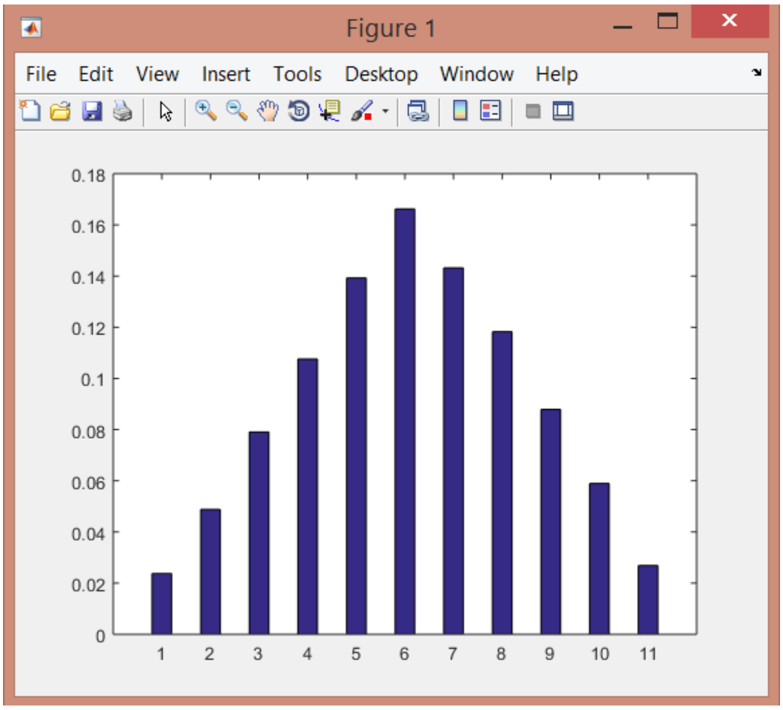
Το ιστόγραμμα της joint f_{Z1Z2} θα είναι :



Απάντηση 3(β) :
η συνάρτηση $f_Y(y) = f_{Z1}(z1) * f_{Z2}(z2)$ θα είναι :

$$f_Y(y) = \begin{Bmatrix} 0.0238, & y = 1 \\ 0.0488, & y = 2 \\ 0.0791, & y = 3 \\ 0.1076, & y = 4 \\ 0.1393, & y = 5 \\ 0.1662, & y = 6 \\ 0.1432, & y = 7 \\ 0.1183, & y = 8 \\ 0.0879, & y = 9 \\ 0.0590, & y = 10 \\ 0.0269, & y = 11 \\ 0, & αλλού \end{Bmatrix}$$

Η Συνάρτηση πυκνότητας πιθανότητας της $f_Y(y)$ θα είναι :



Γνωρίζουμε από το θεώρημα κεντρικού ορίου ότι όσο πιο πολύ αυξάνουμε τις τυχαίες μεταβλητές τότε η νέα κατανομή θα τείνει να μοιάζει όλο και πιο πολύ την γκαουσιανή κατανομή . Έτσι και στο δικό μας σχήμα η pdf του y τείνει να μοιάσει στην γκαουσιανή κατανομή (αυτό δεν φαίνεται τόσο ξεκάθαρα βέβαια γιατί έχουμε συνδυάσει μόνο 2 μεταβλητές).

<u>Άσκηση 4</u>

Έστω η συνάρτηση $f(x) = (x-5)^4 + 3x$ και θέλουμε να βρούμε την τιμή του x στην οποία βρίσκεται το ελάχιστο της συνάρτησης.

α) Χρησιμοποιείστε την ιδιότητα της παραγώγου για να βρεθεί απευθείας η τιμή του ελαχίστου.

β) Υλοποιείστε στο MATLAB την αριθμητική τεχνική του gradient descent για να βρείτε το ελάχιστο της συνάρτησης. Ξεκινήστε από μια τυχαία τιμή, βάλτε μια αυθαίρετη τιμή για το βήμα η και εφαρμόστε τη σχέση του gradient descent επαναληπτικά, μέχρι ότου η τιμή του ελαχίστου να μην αλλάζει. Μπορεί να χρειαστεί να δοκιμάσετε διάφορες τιμές για το η. Σε
πόσες επαναλήψεις έχετε βρει το ελάχιστο.

γ) Υλοποιείστε στο MATLAB την αριθμητική τεχνική του Newton method για να βρείτε το ελάχιστο της συνάρτησης. Ξεκινήστε από μια τυχαία τιμή και εφαρμόστε τη σχέση του Newton method επαναληπτικά, μέχρι ότου η τιμή του ελαχίστου να μην αλλάζει. Σε πόσες επαναλήψεις έχετε βρει το ελάχιστο.

δ) Συγκρίνετε τις 2 μεθοδολογίες. Ποια δίνει τη γρηγορότερη σύγκλιση

Απάντηση 4(α):

βρίσκω την παράγωγο της f :

$f'(x) = 4(x-5)^3 + 3$

Βρίσκω που μηδενίζει η παράγωγος :

$$f'(x) = 0 \Leftrightarrow 4(x-5)^3 + 3 = 0 \Leftrightarrow (x-5)^3 = -\frac{3}{4} \Leftrightarrow x = \left(-\frac{3}{4}\right)^{\frac{1}{3}} + 5 \approx 4.0914$$

Άρα η δοθείσα συνάρτηση παρουσιάζει ολικό ελάχιστο για x= 4.0914 με $f(4.0914) = 12.9557$

Απάντηση 4(β):

για η =0.5 και 0.05 ο αλγόριθμος τρέχει όλες τις επαναλήψεις χωρίς να βρεί το ελάχιστο. Για η = 0.02 ο αλγόριθμος βρίσκει το ελάχιστο στις 65 επαναλήψεις ενώ αν μειώσουμε το η στο 0.01 βλέπουμε ότι οι επαναλήψεις γίνονται 148 άσχετα που βρίσκουμε το αποτέλεσμα αυτό μπορεί να επιτευχθεί και με λιγότερες επαναλήψεις. Αν δοκιμάζαμε και για ακόμη μικρότερο η, έστω η=0.0001 βλέπουμε ότι ο αλγόριθμος εξαντλεί όλες τις επαναλήψεις επιστρέφοντας για ελάχιστο το f(3.7888)=13.5185. Καταλήγουμε ότι μία καλή τιμή για το η είναι γύρω στο 0.02 (με αφετηρία το 0).

```
clear;
clc;
%where the x values will be stored
x=zeros(1,1000);
%determining the first element
x(1)=0;
%the factor I need
n=0.0001;
%counts the itterations
reps=0;
%when i can stop the itterations
e=0.0000001;

for i=1:1000
    %counter increments in each itteration
    reps=reps+1;

    %gradient descent
    x(i+1)=x(i)-n*(4*((x(i)-5)^3)+3);

    %stoping criterion
    if (abs(x(i+1)-x(i))<e)
        break
    end
end

reps
xmin=x(reps)
fmin=((xmin-5)^4)+3*xmin
```

Απάντηση 4(γ):

ξεκινώντας πάλι από το 0 (για να μην είναι άδικο για τον προηγούμενο αλγόριθμο) βλέπουμε ότι ο αλγόριθμος φτάνει στο ελάχιστο σε μόλις 9 επαναλήψεις.

```
reps =

     9


xmin =

    4.0914


fmin =

   12.9557
```

```matlab
clear;
clc;
%where the x values will be stored
x=zeros(1,1000);
%determining the first element
x(1)=0;
%the factor I need
n=0.0001;
%counts the itterations
reps=0;
%when i can stop the itterations
e=0.0000001;

for i=1:1000
    %counter increments in each itteration
    reps=reps+1;

    %Newtons Method
     x(i+1)= x(i)-((12*(x(i)-5)^2)^-1)*(4*((x(i)-5)^3)+3);

    %stoping criterion
    if (abs(x(i+1)-x(i))<e)
        break
    end
end

reps
xmin=x(reps)
fmin=((xmin-5)^4)+3*xmin
```

Απάντηση 4(δ):

το gradient descent είναι μία πιο απλή λύση από την μέθοδο Newton και δεν απαιτεί τόσου υπολογιστικούς πόρους. Ωστόσο η πρώτη μέθοδος απαιτεί σωστή ρύθμιση των παραμέτρων (σημείο αφετηρίας , η) για να καταλήξουμε σε ένα σωστό αποτέλεσμα και δεν είναι πάντα τόσο εύκολο αυτό σε αντίθεση με την μέθοδο newton η οποία δεν απαιτεί κάποια ιδιαίτερη ρύθμιση παραμέτρων. Ωστόσο είναι σημαντικό στην δεύτερη να προσέξουμε το σημείο που μηδενίζεται η δεύτερη παράγωγος γιατί αφού χρησιμοποιούμε την αντίστροφή της αν σε μία επανάληψη πέσουμε σε αυτό το σημείο τότε θα υπάρξει πρόβλημα.

Βλέπουμε και εδώ ότι για μία σχετικά καλή παράμετρο η=0.02 ο αλγόριθμος gradient descent φτάνει από το 0 στο σημείο που βρίσκεται το ελάχιστο σε 65 επαναλήψεις ενώ η μέθοδος newton σε 9 επαναλήψεις (54 επαναλήψεις γρηγορότερα). Ωστόσο, επειδή , όπως αναφέρθηκε και πιο πριν, η μέθοδος newton καταναλώνει πολλούς πόρους ειδικά για προβλήματα της πολυπλοκότητας του συγκεκριμένου προβλήματος ο αλγόριθμος gradient descent κάνει ικανοποιητικά την δουλεία του και για αυτό προτιμάται αυτός.