Fantasia heart beat data analysis

ECE3093 Assignment 2 Part A, Written by Kun Zhang (22701478)

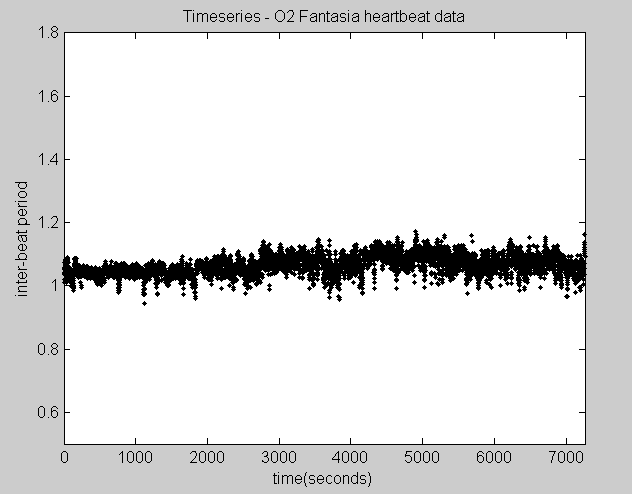
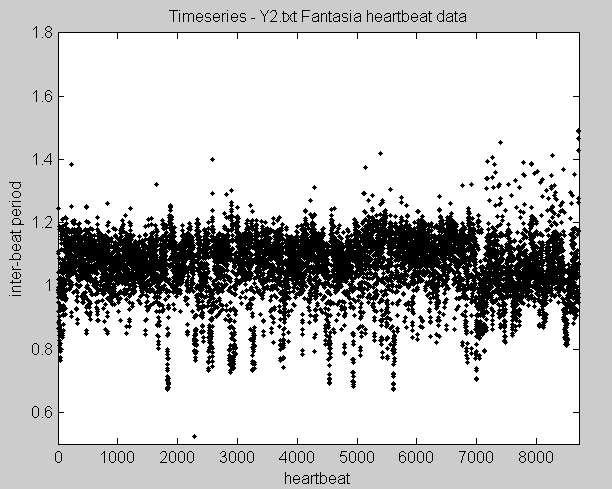
# I. Fantasia heartbeat datasets

## (2) Run the m-file and answer the following questions

**(a) What are the characteristics of the time-series plot of a typical Y1,.., Y5? How do these differ from the characteristics of the time-series plot of a typical O1,...,O5?**

From the time-series plots, it can be seen that the RR periods for young subjects are more dispersed comparing to those for old subjects, which indicates the former has larger variance. However, the means appear to be similar. In addition, the maximum value for young subjects if greater than that for old subjects, and correspondingly, the minimum value for young subjects is lower than that for the old subjects.

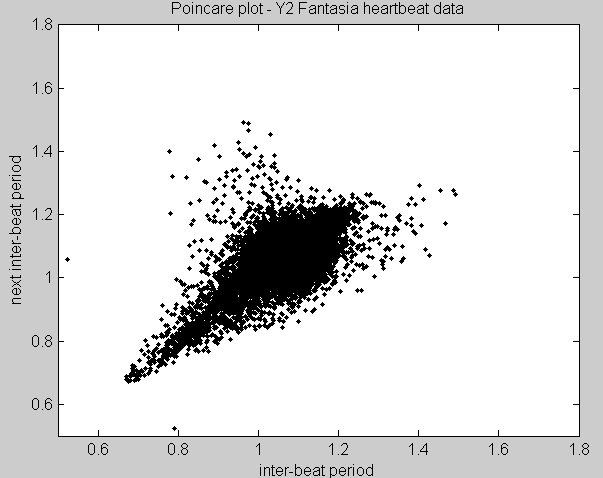
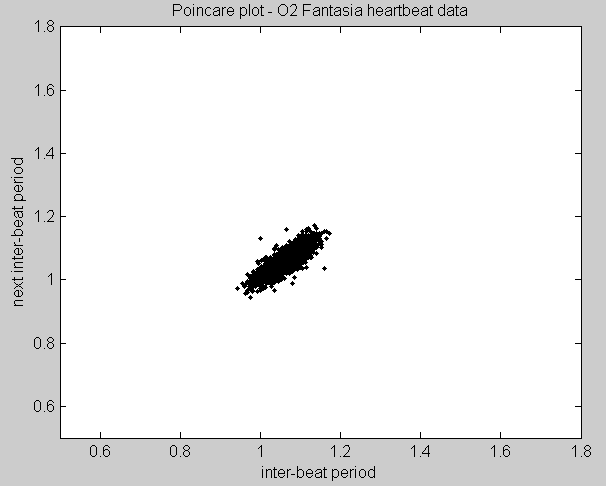
**Note: All the plots of data are shown in Appendix IV for the conciseness of this report**



*Figure 1 Time –series plots for (a) young subject Y2 and (b) old subject O2*

**(b) What are the characteristics of the Poincare plot of a typical Y1,..,Y5? How do these differ the characteristics of the Poincare plot of a typical O1,...,O5?**

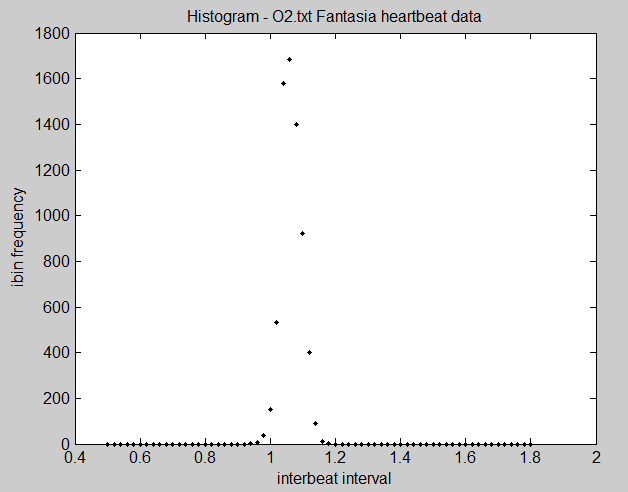
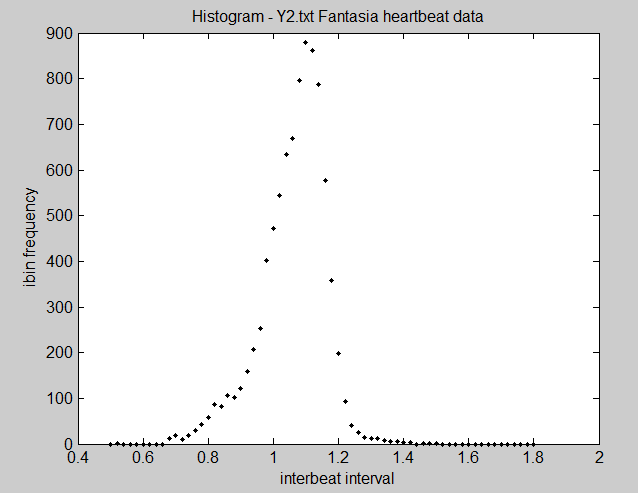
The cone-shape Poincare plots show that the data for young subjects spread in a wider range while those for old subjects centre within a small interval. This indicates that the data for young subject has larger variance with greater maximum and smaller minimum value

*Figure 2 Poincare plots for (a) young subject Y2 and (b) old subject O2*

**(c) What are the characteristics of the histogram plot of a typical Y1,..,Y5? How do these differ the characteristics of the histogram plot of a typical O1,...,O5?**

The histogram plots for young subjects show wider bell-liked-shapes than those for old subjects, revealing the former data is more dispersed. In addition, the maximum frequency of the former is lower than that of the latter.



*Figure 3 Histogram plots for (a) young subject Y2 and (b) old subject O2*

**(d) Do any members of the groups Y1,...,Y5 or O1,...,O5 appear atypical? If yes, how do their time-series, Poincare and histograms differ from other members of the group?**

Yes. Y3 and O4 appear atypical. The plots of Y3 are less dispersed than those of other young subjects and resemble the characteristics of plots of old subjects. Conversely, the plots for O4 scatter in a larger range those that of other older subjects. Also, the histogram has a wide bell-liked-shape.

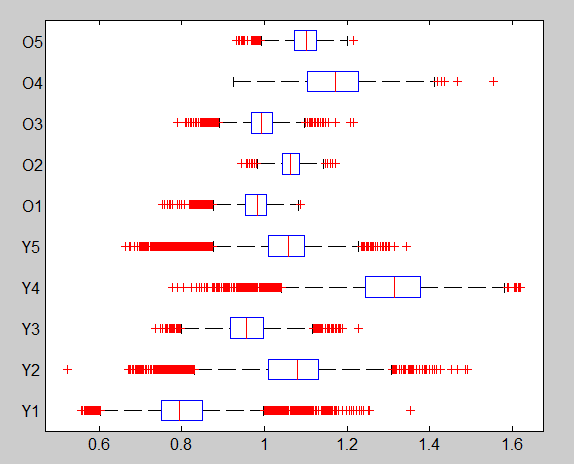
**(e) What features of the time-series plots do you think correspond to the acceleration, deceleration and idling of the heart?**

The approximated gradients for points in time-series plots indicate the acceleration, deceleration and idling of the heart: negative gradients correspond to acceleration (less RR period), positive correspond to deceleration and zero gradient corresponds to idling.

## (3) Write an M-file that plots a horizontal stack of boxplots for all Os and Ys

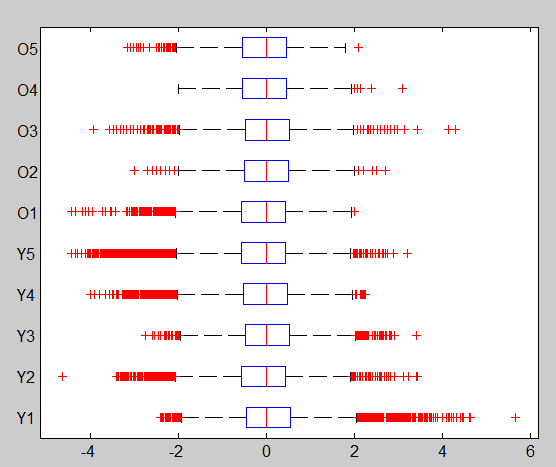
(M-file is presented in Appendix I)

**(a) Plot a horizontal stack of 10 boxplots for the raw data of Y1,...,Y5; O1,..O5…**



*Figure 4 Plot a horizontal stack of 10 boxplots for the raw data of Y1,...,Y5; O1,..O5..*

**(c) Plot a horizontal stack of 10 boxplots for the normalised data of Y1,..., Y5; O1,..O5…**



*Figure 5 Plot a horizontal stack of 10 boxplots for the normalised data of Y1,...,Y5,O1,..O5.*

**(e) Interpret the stack boxplots. What do quartiles and the extent of their data say about the variability of the inter-beat period among these subjects?**

In Figure 4, the box lengths for group Y are greater and the tails are more extended comparing to those of the group O. This means the variability of the inter-beat period for group Y subjects is greater than that of group O subjects.

Figure 5 illustrates the same thing by having zero medians but varying outliers. In general, the group O data have fewer whiskers than group Y data, which indicates the latter has greater variability

**(f) Do the stacked boxplots support your conclusions in (1)(a),(b),(c) and (d)?**

Yes, they do. The boxplots show that the data for each typical subject in group Y has more whiskers and larger IQR than each typical subject in group O, indicating the former has more variability than the latter, which is the aforementioned conclusion.

# II. Multiscale analysis using detrended fluctuation analysis

## (Step 3)

**How can the dataset be divided up into boxes?**

The number of boxes is primarily determined by dividing the number of data points by box length n. The result is rounded down by MATLAB function fix (). The remainder is obtained by mod (). If the remainder is greater than or equal to 4, which is too small to form a single box, than the remained elements alone form a new box. Otherwise, the remained elements are grouped with the last n-length box forming a larger box. Apparently, in the latter case the number of the boxes is one less than the former case.

Related codes can be found in Appendix II Part A step 4

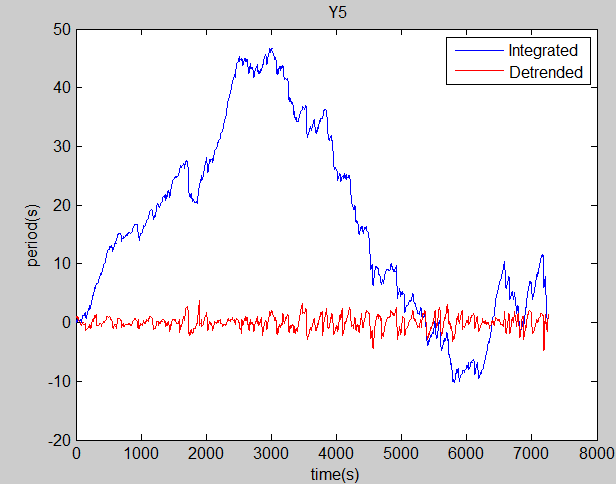
## (Step 4)

**How do we find the trend line, y=a+bx for each box?**

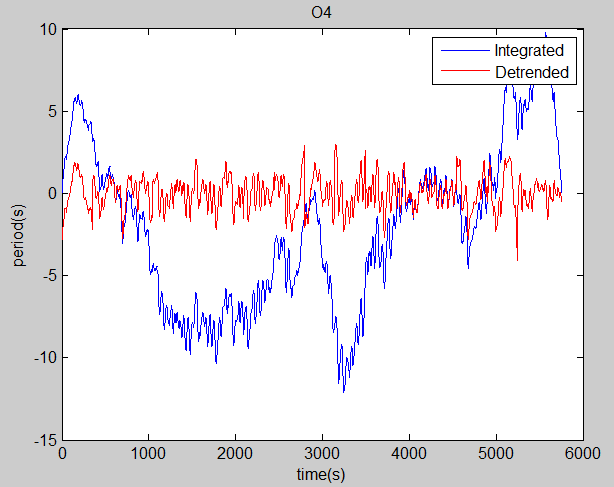
Using linear regression for each box, the linearized parameter b1, b2 can be found. The trend line for each box can thus be plotted by y=b1+b2x.

Corresponding code can be found in Appendix II Part A step 4

**Plots of detrending for Y5 with box size = 300**



**Plots of detrending for O4 with box size = 300**

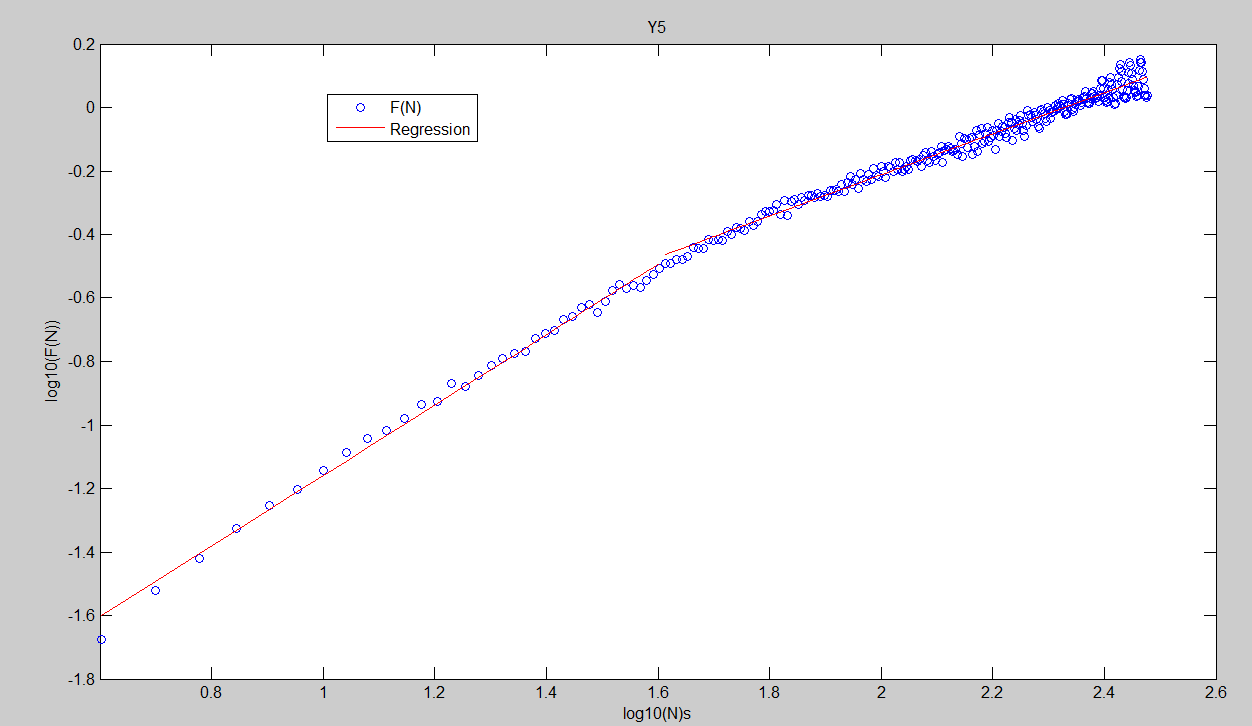


## (Step 6)

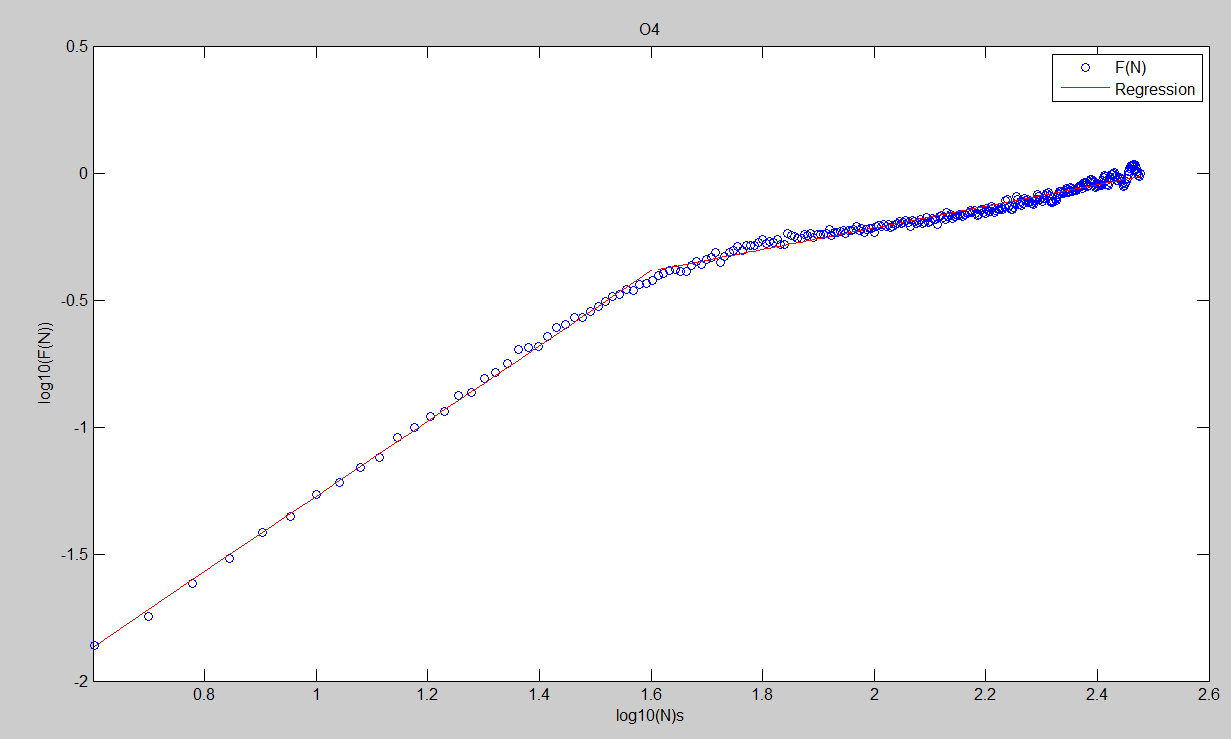
Note the following plots are log-log plots in normal scale (instead of log-scale), in accord with the plots in the paper by N. Inyengar [1].

The breaking point is approximated by observation of the plots, although a more accurate method may be applied as described by Sarah E. Ryan [2]

**Plot log10F(n) versus log10n for Y5**



**Plot log10F(n) versus log10n for O4**



## (Step 7)

Using linear regression for the plots, we have

**Piecewise slope for Y5**

**Piecewise slope for O4**

## (Step 8)

Using the relationships

We have

**Beta value for Y5**

**Beta value for O4**

# Reference

[1] N. Iyengar, C. K. Peng, Aged-related alterations in the fractal scaling of cardiac interbeat interval dynamics, 1996

[2] S.E. Ryan, L. S. Porth A tutorial on the piecewise regression approach applied to bedload transport data, U.S. Department of Agriculture

# Appendix

## Appendix I M-file for I.(3)

close all;clc;

%load all data

heartY1=load('Y1.txt');

heartY2=load('Y2.txt');

heartY3=load('Y3.txt');

heartY4=load('Y4.txt');

heartY5=load('Y5.txt');

heartO1=load('O1.txt');

heartO2=load('O2.txt');

heartO3=load('O3.txt');

heartO4=load('O4.txt');

heartO5=load('O5.txt');

%get lengths for the data loaded

reclen\_O1 = length(heartO1);

reclen\_O2 = length(heartO2);

reclen\_O3 = length(heartO3);

reclen\_O4 = length(heartO4);

reclen\_O5 = length(heartO5);

reclen\_Y1 = length(heartY1);

reclen\_Y2 = length(heartY2);

reclen\_Y3 = length(heartY3);

reclen\_Y4 = length(heartY4);

reclen\_Y5 = length(heartY5);

heart\_dataY=[heartY1;heartY2;heartY3;heartY4;heartY5; ...

heartO1;heartO2;heartO3;heartO4;heartO5];

xlabel=[repmat({'Y1'},1,reclen\_Y1)';repmat({'Y2'},1,reclen\_Y2)';repmat({'Y3'},1,reclen\_Y3)';repmat({'Y4'},1,reclen\_Y4)';repmat({'Y5'},1,reclen\_Y5)';...

repmat({'O1'},1,reclen\_O1)';repmat({'O2'},1,reclen\_O2)';repmat({'O3'},1,reclen\_O3)';repmat({'O4'},1,reclen\_O4)';repmat({'O5'},1,reclen\_O5)'];

%(a)boxplots for orignial data

figure(1)

boxplot(heart\_dataY,xlabel,'orientation','horizontal')

%(b)(c)boxplots for normalised data

heart\_norm=@(x)(x-median(x))/iqr(x); %subtract mean and divided by IQR

heartY1\_n=heart\_norm(heartY1);

heartY2\_n=heart\_norm(heartY2);

heartY3\_n=heart\_norm(heartY3);

heartY4\_n=heart\_norm(heartY4);

heartY5\_n=heart\_norm(heartY5);

heartO1\_n=heart\_norm(heartO1);

heartO2\_n=heart\_norm(heartO2);

heartO3\_n=heart\_norm(heartO3);

heartO4\_n=heart\_norm(heartO4);

heartO5\_n=heart\_norm(heartO5);

heart\_dataY\_n=[heartY1\_n;heartY2\_n;heartY3\_n;heartY4\_n;heartY5\_n; ...

heartO1\_n;heartO2\_n;heartO3\_n;heartO4\_n;heartO5\_n];

xlabel=[repmat({'Y1'},1,reclen\_Y1)';repmat({'Y2'},1,reclen\_Y2)';repmat({'Y3'},1,reclen\_Y3)';repmat({'Y4'},1,reclen\_Y4)';repmat({'Y5'},1,reclen\_Y5)';...

repmat({'O1'},1,reclen\_O1)';repmat({'O2'},1,reclen\_O2)';repmat({'O3'},1,reclen\_O3)';repmat({'O4'},1,reclen\_O4)';repmat({'O5'},1,reclen\_O5)'];

%(a)boxplots for orignial data

figure(2)

boxplot(heart\_dataY\_n,xlabel,'orientation','horizontal')

## Appendix II Part A step 4

%Written by Kun Zhang, 22701478

clc;clear;close all;

%data assigned Y5, O4

heartY5=load('Y5.txt');

heartO4=load('O4.txt');

%integrated data

Y5\_igt=cumsum(heartY5-mean(heartY5));

O4\_igt=cumsum(heartO4-mean(heartO4));

%data length

Y5\_len=length(Y5\_igt);

O4\_len=length(O4\_igt);

%The elapsed time = sum of the heartbeat intervals

Y5\_time=cumsum(heartY5);

O4\_time=cumsum(heartO4);

%-------------------------------for Y5------------------------------

%Box paramentres

n=300; %box size confined in a range or 4-300number of boxes

m=fix(Y5\_len/n); %number of boxes

n\_redu=mod(Y5\_len,n); %redundant elements number <300

%Spliting data and time

if n\_redu<4 %last box = regularbox+redundant elements

n\_last=n\_redu+n; %last box size last box

m\_reg = m-1; %number of regular boxes

Y5\_split\_reg=reshape(Y5\_igt(1:(Y5\_len-n\_last)),n,m\_reg);%splitted regular boxes

Y5\_split\_last=Y5\_igt((Y5\_len-n\_last+1):end);%last box

Y5\_time\_reg=reshape(Y5\_time(1:(Y5\_len-n\_last)),n,m\_reg);%splitted regular boxes

Y5\_time\_last=Y5\_time((Y5\_len-n\_last+1):end);%last box

else %last box = redundant elements

n\_last=n\_redu;%last box size last box

m\_reg = m; %number of regular boxes

Y5\_split\_reg=reshape(Y5\_igt(1:(Y5\_len-n\_last)),n,m\_reg);%splitted regular boxes

Y5\_split\_last=Y5\_igt((Y5\_len-n\_last+1):end);%last box

Y5\_time\_reg=reshape(Y5\_time(1:(Y5\_len-n\_last)),n,m\_reg);%splitted regular boxes

Y5\_time\_last=Y5\_time((Y5\_len-n\_last+1):end);%last box

end

%predefine yreg and ydet, regular regression and detrended

yreg=zeros(n,m\_reg);

ydet=zeros(n,m\_reg);

%linear regression for each regualr boxes

for i=1:m\_reg

b=polyfit(Y5\_time\_reg(:,i),Y5\_split\_reg(:,i),1); %linear regression

yreg(:,i)=b(1)\*Y5\_time\_reg(:,i)+b(2); %=polyval

ydet(:,i)=transpose(Y5\_split\_reg(:,i)-yreg(:,i)); %find the residual

end

%linear regression for last box

b=polyfit(Y5\_time\_last,Y5\_split\_last,1); %linear regression

yreg\_last=b(1)\*Y5\_time\_last+b(2); %m(2) is the coefficient with higher confidence

ydet\_last=transpose(Y5\_split\_last-yreg\_last); %last box detrended

%combine regular boxes and the last box

ydet\_all=[reshape(ydet,1,[]),ydet\_last];

yreg\_all=[reshape(yreg,1,[]),transpose(yreg\_last)];

figure(1)

plot(Y5\_time,Y5\_igt)

hold on

plot(Y5\_time,ydet\_all,'r')

title(' Y5');

xlabel('time(s)');

ylabel('period(s)');

legend('Integrated','Detrended','Regession')

%-------------------------------for O4------------------------------------

%Box paramentres

n=300; %box size confined in a range or 4-300number of boxes

m=fix(O4\_len/n); %number of boxes

n\_redu=mod(O4\_len,n); %redundant elements number <300

%spliting data and time

if n\_redu<4 %if lastbox size<4, last box = regularbox+redundant elements

n\_last=n\_redu+n; %last box size last box

m\_reg = m-1; %number of regular boxes

O4\_split\_reg=reshape(O4\_igt(1:(O4\_len-n\_last)),n,m\_reg);%splitted regular boxes

O4\_split\_last=Y5\_igt((O4\_len-n\_last+1):end);%last box

O4\_time\_reg=reshape(O4\_time(1:(O4\_len-n\_last)),n,m\_reg);%splitted regular boxes

O4\_time\_last=O4\_time((O4\_len-n\_last+1):end);%last box

else %last box = redundant elements

n\_last=n\_redu;%last box size last box

m\_reg = m; %number of regular boxes

O4\_split\_reg=reshape(O4\_igt(1:(O4\_len-n\_last)),n,m\_reg);%splitted regular boxes

O4\_split\_last=O4\_igt((O4\_len-n\_last+1):end);%last box

O4\_time\_reg=reshape(O4\_time(1:(O4\_len-n\_last)),n,m\_reg);%splitted regular boxes

O4\_time\_last=O4\_time((O4\_len-n\_last+1):end);%last box

end

%predefine yreg and ydet, regular regression and detrended

yreg=zeros(n,m\_reg);

ydet=zeros(n,m\_reg);

%linear regression for each regualr boxes

for i=1:m\_reg

b=polyfit(O4\_time\_reg(:,i),O4\_split\_reg(:,i),1); %linear regression

yreg(:,i)=b(1)\*O4\_time\_reg(:,i)+b(2); %=polyval

ydet(:,i)=transpose(O4\_split\_reg(:,i)-yreg(:,i)); %find the residual

end

%linear regression for last box

b=polyfit(O4\_time\_last,O4\_split\_last,1); %linear regression

yreg\_last=b(1)\*O4\_time\_last+b(2); %m(2) is the coefficient with higher confidence

ydet\_last=transpose(O4\_split\_last-yreg\_last); %last box detrended

%combine regular boxes and the last box

ydet\_all=[reshape(ydet,1,[]),ydet\_last];

yreg\_all=[reshape(yreg,1,[]),transpose(yreg\_last)];

figure(2)

plot(O4\_time,O4\_igt)

hold on

plot(O4\_time,ydet\_all,'r')

title('O4');

xlabel('time(s)');

ylabel('period(s)');

legend('Integrated','Detrended','Regession')

## Appendix III Part A step 6

%Written by Kun Zhang, 22701478

clc;clear;close all;

%data assigned Y5, O4

heartY5=load('Y5.txt');

heartO4=load('O4.txt');

%integrated data

Y5\_igt=cumsum(heartY5-mean(heartY5));

O4\_igt=cumsum(heartO4-mean(heartO4));

%data length

Y5\_len=length(Y5\_igt);

O4\_len=length(O4\_igt);

%The elapsed time = sum of the heartbeat intervals

Y5\_time=cumsum(heartY5);

O4\_time=cumsum(heartO4);

%----------------------for Y5------------------------------

%Box paramentres

for n=4:300; %box size confined in a range or 4-300number of boxes

m=fix(Y5\_len/n); %number of boxes

n\_redu=mod(Y5\_len,n); %redundant elements number <300

%Spliting data and time

if n\_redu<4 %last box = regularbox+redundant elements

n\_last=n\_redu+n; %last box size last box

m\_reg = m-1; %number of regular boxes

Y5\_split\_reg=reshape(Y5\_igt(1:(Y5\_len-n\_last)),n,m\_reg);%splitted regular boxes

Y5\_split\_last=Y5\_igt((Y5\_len-n\_last+1):end);%last box

Y5\_time\_reg=reshape(Y5\_time(1:(Y5\_len-n\_last)),n,m\_reg);%splitted regular boxes

Y5\_time\_last=Y5\_time((Y5\_len-n\_last+1):end);%last box

else %last box = redundant elements

n\_last=n\_redu;%last box size last box

m\_reg = m; %number of regular boxes

Y5\_split\_reg=reshape(Y5\_igt(1:(Y5\_len-n\_last)),n,m);%splitted regular boxes

Y5\_split\_last=Y5\_igt((Y5\_len-n\_last+1):end);%last box

Y5\_time\_reg=reshape(Y5\_time(1:(Y5\_len-n\_last)),n,m);%splitted regular boxes

Y5\_time\_last=Y5\_time((Y5\_len-n\_last+1):end);%last box

end

%predefine yreg and ydet, regular regression and detrended

yreg=zeros(n,m\_reg);

ydet=zeros(n,m\_reg);

%linear regression for each regualr boxes

for i=1:m\_reg

b=polyfit(Y5\_time\_reg(:,i),Y5\_split\_reg(:,i),1); %linear regression

yreg(:,i)=b(1)\*Y5\_time\_reg(:,i)+b(2); %=polyval

ydet(:,i)=transpose(Y5\_split\_reg(:,i)-yreg(:,i)); %find the residual

end

%linear regression for last box

b=polyfit(Y5\_time\_last,Y5\_split\_last,1); %linear regression

yreg\_last=b(1)\*Y5\_time\_last+b(2); %m(2) is the coefficient with higher confidence

ydet\_last=transpose(Y5\_split\_last-yreg\_last);

ydet\_all=[reshape(ydet,1,[]),ydet\_last];

rms(n-3)=sqrt(dot(transpose(ydet\_all),ydet\_all)/length(ydet\_all));%root means square of dot product / matrix length

end

n\_mat=4:300;

%b\_log=polyfit(log10(n\_mat),log10(rms),1);%regression for log-log plot

%approximate from the graph, the 37th points is the turning point

alpha\_1=polyfit(log10(n\_mat(1:37)),log10(rms(1:37)),1)

alpha\_2=polyfit(log10(n\_mat(38:end)),log10(rms(38:end)),1)

figure (1)

plot(log10(n\_mat),log10(rms),'o')

hold on

%plot(log10(n\_mat),b\_log(1)\*log10(n\_mat)+b\_log(2))

plot(log10(n\_mat(1:37)),alpha\_1(1)\*log10(n\_mat(1:37))+alpha\_1(2),'r')

hold on

plot(log10(n\_mat(38:end)),alpha\_2(1)\*log10(n\_mat(38:end))+alpha\_2(2),'r')

title('Y5');

xlabel('log10(N)s');

ylabel('log10(F(N))');

legend('F(N)','Regression')

%----------------------for O4------------------------------

%Box paramentres

for n=4:300; %box size confined in a range or 4-300number of boxes

m=fix(O4\_len/n); %number of boxes

n\_redu=mod(O4\_len,n); %redundant elements number <300

%Spliting data and time

if n\_redu<4 %last box = regularbox+redundant elements

n\_last=n\_redu+n; %last box size last box

m\_reg = m-1; %number of regular boxes

O4\_split\_reg=reshape(O4\_igt(1:(O4\_len-n\_last)),n,m\_reg);%splitted regular boxes

O4\_split\_last=O4\_igt((O4\_len-n\_last+1):end);%last box

O4\_time\_reg=reshape(O4\_time(1:(O4\_len-n\_last)),n,m\_reg);%splitted regular boxes

O4\_time\_last=O4\_time((O4\_len-n\_last+1):end);%last box

else %last box = redundant elements

n\_last=n\_redu;%last box size last box

m\_reg = m; %number of regular boxes

O4\_split\_reg=reshape(O4\_igt(1:(O4\_len-n\_last)),n,m);%splitted regular boxes

O4\_split\_last=O4\_igt((O4\_len-n\_last+1):end);%last box

O4\_time\_reg=reshape(O4\_time(1:(O4\_len-n\_last)),n,m);%splitted regular boxes

O4\_time\_last=O4\_time((O4\_len-n\_last+1):end);%last box

end

%predefine yreg and ydet, regular regression and detrended

yreg=zeros(n,m\_reg);

ydet=zeros(n,m\_reg);

%linear regression for each regualr boxes

for i=1:m\_reg

b=polyfit(O4\_time\_reg(:,i),O4\_split\_reg(:,i),1); %linear regression

yreg(:,i)=b(1)\*O4\_time\_reg(:,i)+b(2); %=polyval

ydet(:,i)=transpose(O4\_split\_reg(:,i)-yreg(:,i)); %find the residual

end

%linear regression for last box

b=polyfit(O4\_time\_last,O4\_split\_last,1); %linear regression

yreg\_last=b(1)\*O4\_time\_last+b(2); %m(2) is the coefficient with higher confidence

ydet\_last=transpose(O4\_split\_last-yreg\_last);

ydet\_all=[reshape(ydet,1,[]),ydet\_last];

rms(n-3)=sqrt(dot(transpose(ydet\_all),ydet\_all)/length(ydet\_all));%root means square of dot product / matrix length

end

n\_mat=4:300;

%b\_log=polyfit(log10(n\_mat),log10(rms),1);%regression for log-log plot

%approximate from the graph, the 37th points is the turning point

alpha\_1=polyfit(log10(n\_mat(1:37)),log10(rms(1:37)),1)

alpha\_2=polyfit(log10(n\_mat(38:end)),log10(rms(38:end)),1)

figure (2)

plot(log10(n\_mat),log10(rms),'o')

hold on

%plot(log10(n\_mat),b\_log(1)\*log10(n\_mat)+b\_log(2))

plot(log10(n\_mat(1:37)),alpha\_1(1)\*log10(n\_mat(1:37))+alpha\_1(2),'r')

hold on

plot(log10(n\_mat(38:end)),alpha\_2(1)\*log10(n\_mat(38:end))+alpha\_2(2),'r')

title('O4');

xlabel('log10(N)s');

ylabel('log10(F(N))');

legend('F(N)','Regression')

## Appendix IV Plots for part I

