Internal Assessment 2

Function that best models the population of China

This is the data of China's population from 1950 to 1995:

Time (Year)	Population (in millions)
1950	554.8
1955	609.0
1960	657.5
1965	729.2
1970	830.7
1975	927.8
1980	998.9
1985	1070.0
1990	1155.3
1995	1220.5

In this given data, the x value is the time (year) because time is always independent. Whereas the population (in millions) of china is the y value because each value is dependent on the time. Time cannot be stopped and so the values of x are from 0 years to infinite years. The y value that is the population is limited ($y \neq \infty$) because of numerous of reason, for example less food availability we cannot have infinite people..

Now, I am going to plot this given data from 1950 to 1995 using Logger Pro. Steps are shown in the appendix:
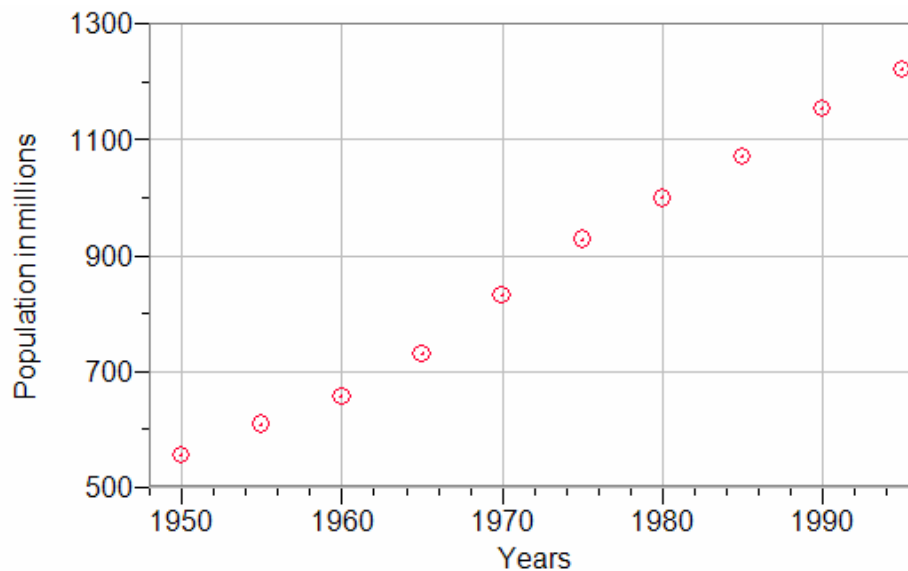
Graph 1

The reason why there is not a line connecting all the points is that because it might be possible that in 1951 or 1969 or in another period there could have been a rapid growth or decline in the population. That is why it is not continuous but discrete.

As we can see that the population is increasing as by the time. As the years are passing by, the trend of the population is gradually increasing more and more, the graph is becoming steeper. There is no rapid growth or decline in the population. I think that this is in between gradual and rapid growth of the population. Just after the year 1965, there is a tiny bounce that can be seen that pushes the growth rate more. To help see, I have added here two lines that tells us more about the slope. As we can see that the growth rate from 1970 to 1995 is more and that from 1950 to 1965.
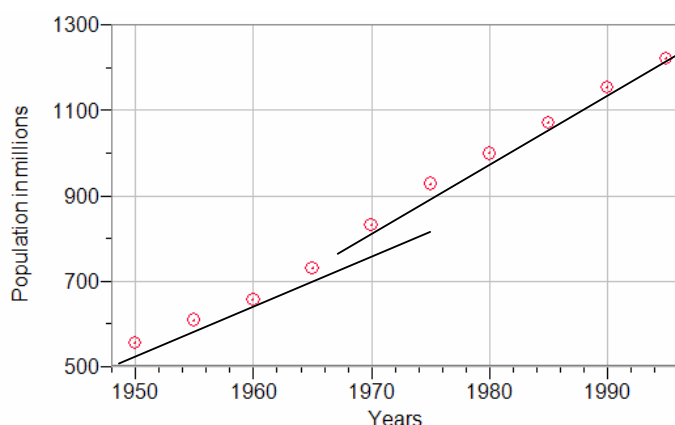
Graph 2

Some of the functions that I think that could model this data were linear function ($f(x) = mx + b$) and exponential function ($f(x) = ab^x$). If we look at graph 1 closely, we can see that most of the data points have aligned themselves in almost a linear form. This informs us that the slope differences will be very small and not drastically big. Apart from the linear function, I also think that the exponential function would work on this particular data set. The reason behind this is that the population that was in 1950 has been doubled in the year in between the year 1985 to 1990. The exponential function is able to graph this type of functions really well.

The model that I decided to choose for this particular data is the exponential function:

$$f(x) = ab^x$$

I will now find out the values for a and b. In order to find a and b, I will randomly pick for values for x and y. Then in both the equation I will isolate for a. I will substitute and then solve for b:

$$y = ab^x$$

$$\frac{y}{b^x} = a$$

$$a = \frac{554.8}{b^{1950}}$$

$$a = \frac{609.0}{b^{1955}}$$

$$\frac{554.8}{b^{1950}} = \frac{609.0}{b^{1955}}$$

$$\frac{b^{1955}}{b^{1950}} = \frac{609.0}{554.8}$$

$$b^5 = 1.097692862$$

$$\sqrt[5]{b^5} = \sqrt[5]{1.097692862}$$

$$b = 1.018816965$$

Sub b into $\frac{y}{b^x} = a$

$$\frac{554.8}{1.018816965^{1950}} = a$$

$$a = 9.049169 \times 10^{-14}$$

New equation is

$$y = 9.049169 \times 10^{-14} \times 1.018816965^x$$

There are also some other values I got for both a and b when I did it with the same method but with different values of x and y:

y = 657.5 x = 1965	y = 998.9 x = 1980
b = 1.0282729645 $a = 1.17537512 \times 10^{-21}$	b = 1.0282729645 $a = 1.05890241 \times 10^{-21}$

y = 830.7 x = 1970	y =927.8 x = 1975
b = 1.022355721 $a = 1.0079569 \times 10^{-16}$	b = 1.022355721 $a = 1.0079569 \times 10^{-16}$

y = 657.5 x = 1960	y =1070.0 x = 1985
b = 1.019669716 $a = 1.72677018 \times 10^{-14}$	b = 1.019669716 $a = 1.72677016 \times 10^{-14}$
y = 609.0 x = 1955	y =1220.5 x = 1995
b = 1.018462566 $a = 1.78650223 \times 10^{-13}$	b = 1.018462566 $a = 1.78650222 \times 10^{-13}$

y = 554.8	y =1155.3
x = 1950	x = 1990
b = 1.018506859	b = 1.018506859
$a = 1.63839368 \times 10^{-13}$	$a = 1.63839368 \times 10^{-13}$

Out of all this, I decided to choose $f(x) = 1.63839368 \times 10^{-13} \times 1.018506859^x$, using the years 1950 and 1990. I will now graph it using TI-84 Plus:
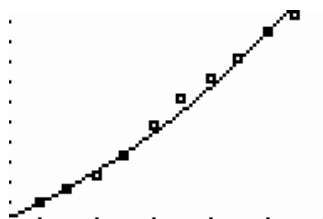
Graph 3

This model is the best fit to the data. But as we can see that some of the points are quite off. From above this is the best that I could see.

A researcher suggests that the population, P at time t can be modelled by

$$P(t) = \frac{K}{1 + Le^{-Mt}}$$

Where K, L, and M are parameters.

This is a logistic equation and the parameters in this equation each have a special role to perform:

Parameters	Role
K	When the value of K is increased the graph is shifted up by a certain number. This depends on the equation itself by how much the graph has shifted. On the other hand, when the value of K is decreased the graph is shifted down by the same amount. We can also say that the value of K is responsible for the up and down of the horizontal asymptote.
L	When the value of L increases, the x shifts towards right and when L decreases, the x shifts towards left. Or in other words, the value of L is responsible for the shift left and right of the vertical asymptote.

	When the sign of the M value is changed, the graph changes the direction and runs into the opposite direction. But when the values of M are increasing, the graph begins to stretch itself.
M	

Apart from the parameters, $P(t), t,$ and e are important variables which are part of the logistic equation. t in this equation is defined as time or years in this case. $P(t)$ is defined as the population at time t in millions and e is a natural log which is equal to 2.718281828.

So in order to find $K, L,$ and M, I used my TI-84 Plus. Method is shown in the appendix:
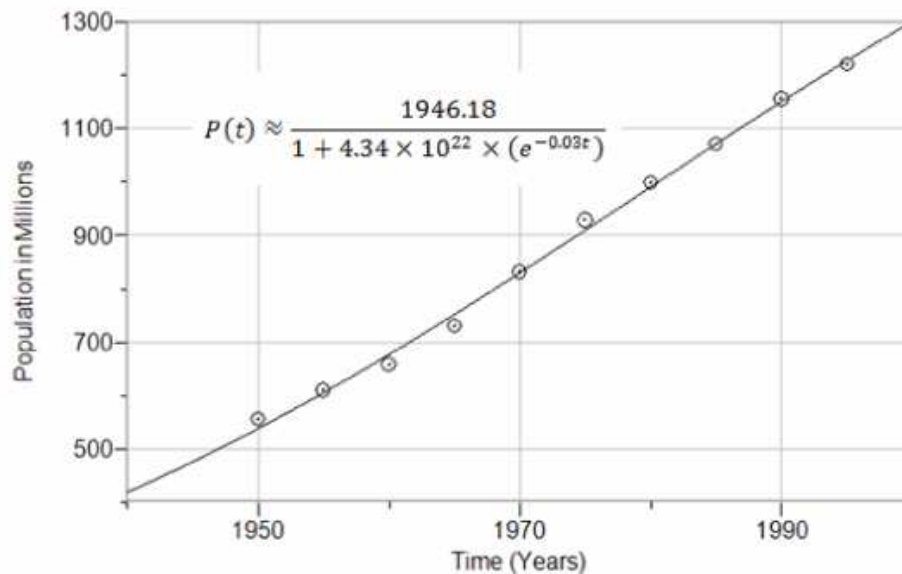
```
Logistic
y=c/(1+ae^(-bx))
a=4.3365552E28
b=.0333213657
c=1946.183708
```

The equation $y = \dfrac{c}{1+ae^{-bx}}$ equally means/represents

$P(t) = \dfrac{K}{1+Le^{-Mt}}$, where

y equals $P(t)$, a equals L,

c equals K. b equals M and x equals t.

$$1 + 4.34 \times 10^{28} \times (e^{-0.03t})$$

$$P(t) \approx \frac{1946.18}{1 + 4.34 \times 10^{22} \times (e^{-0.03t})}$$

Graph 4

The researcher's model fits the original data every well. It fits the data very well than the exponential regression.

The exponential graph is continuous and keeps on leading the way up. When connecting this model with really life population growth of china, I think this graph is very limited. The reason is that the growth of the population can be stopped. This is a very hard mission but it is not impossible. Whereas the exponential model cannot be stopped as it is infinite.

The researcher's model instead is very accurate in terms of fitting the data. Also, I used my TI-84 Plus to see where does the graph go at certain range. I noticed that when I set my window to this:

```
WINDOW
 Xmin=2500
 Xmax=3500
 Xscl=100
 Ymin=1235
 Ymax=3500
 Yscl=100
↓Xres=1
```

I noticed a horizontal line. This suggests that the y value at certain value of x will not increase. This is true in really life too because there is only a certain amount of people China can hold or any other country can hold. That is why this graph is very accurate in terms of population growth.

Now I will be looking at a new data set from year 1983 to 2008 from the 2008 World Economic Outlook published by the International Monetary Fund (IMF).

Year	Population (in millions)
1983	1030.1
1992	1171.7
1997	1236.3
2000	1267.4
2003	1292.3
2005	1307.6
2008	1327.7

I will first use this new data and graph it using Logger Pro:
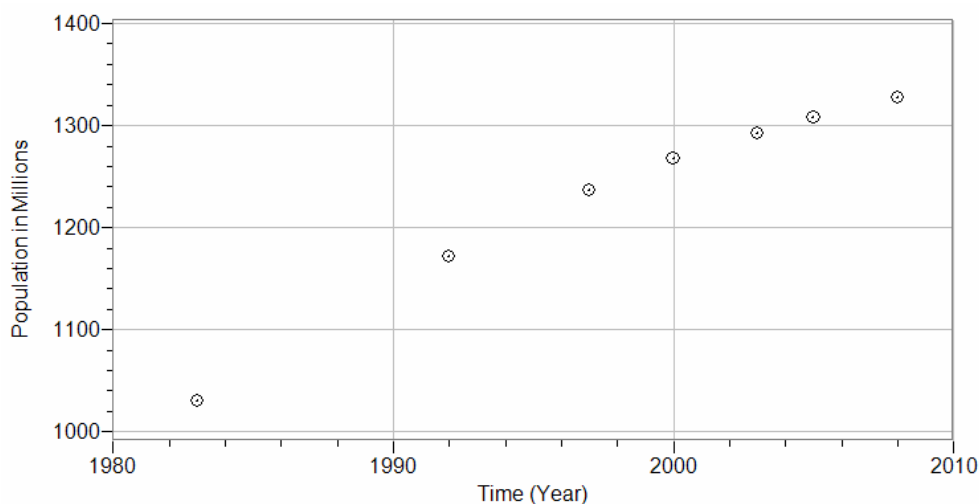
Graph 5

If we compare this time period of 1983 to 2008 to the time period of 1950 to 1995, the growth rate of China's population has reduced. The reason behind this is that if we look at 1983 to 2008, there is a difference of 25 years and the population has increased to 297.6 million. Whereas, if look from 1950 to the next 25 years, that is to 1975, the population has to increased up to 373 million. The population growth rate is decreasing of China's population.

So now I am going to use the researcher's model, $P(t) = \frac{K}{1+Le^{-Mt}}$ and my exponential function, $y = ab^x$ Function to see how the two different models fit this data. I will use TI-84 Plus for exponential function and Logger Pro for researcher's model:
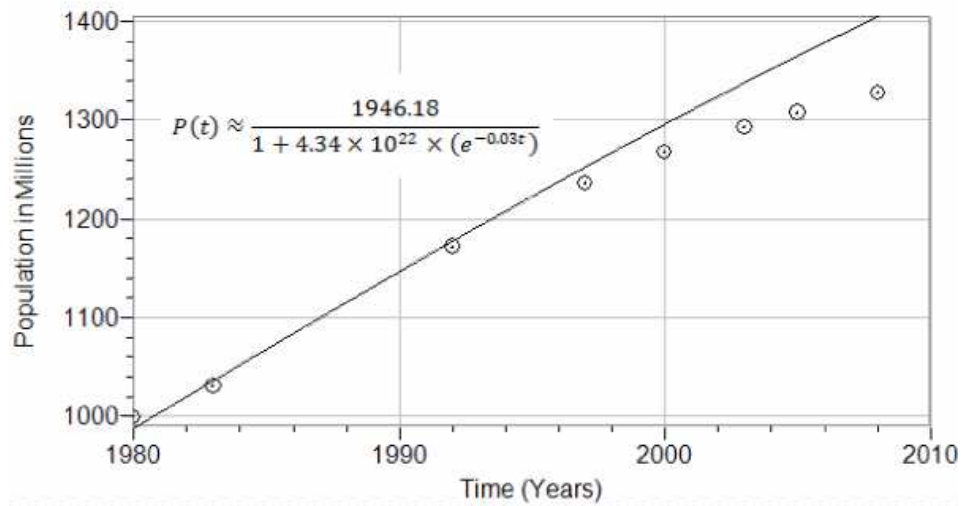
Exponential Function:
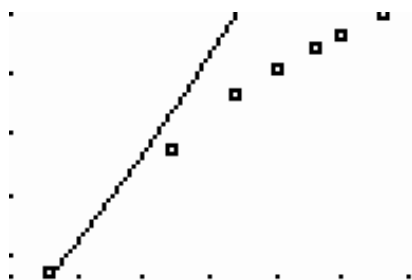
$$y = 9.049169 \times 10^{-14} \times 1.018816965^x$$

Graph 6

The exponential function was no were close to the graph. If we take a look at it the exponential function looks like a line whereas there is a big bent towards the right. The population rate is decreasing whereas the graph is still going way up.

Researcher's Model:

$$P(t) \approx \frac{1946.18}{1 + 4.34 \times 10^{22} \times (e^{-0.03t})}$$

Graph 7

As we can see that the start of the graph is very good as it almost covers the first two points and later the graph is way off then the data itself. As I said earlier, the growth rate has decreased and so the data points can be seen at the bottom of the line that is going straight up. We can see a sudden bend in the data points which suggests that is not a good fit.

Now I use both data set, that is I will use the data set from 1950 to 1995 and mix that with the new data set from 1982 to 2008. I will then modify the researcher's model using Logger Pro. Method is shown in the appendix:

Time (Year)	Population (in millions)
1950	554.8
1955	609.0
1960	657.5
1965	729.2
1970	830.7
1975	927.8
1980	998.9
1985	1070.0
1990	1155.3
1995	1220.5
1983	1030.1
1992	1171.7
1997	1236.3
2000	1267.4
2003	1292.3
2005	1307.6
2008	1327.7

$$P(t) \approx \frac{1617.46}{1 + 1.35 \times 10^{34} \times (e^{-0.040x})}$$
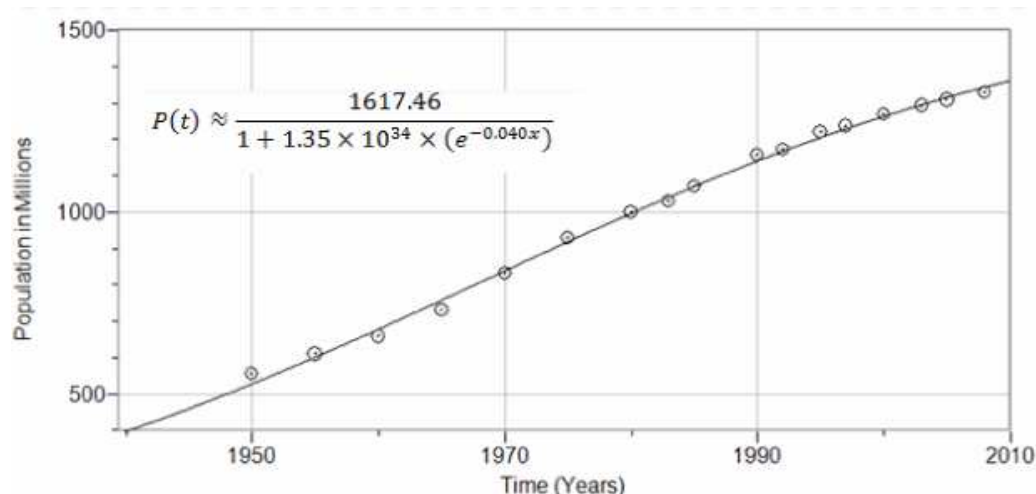
Graph 8

In order to check it, I will use my TI-84 Plus. Method is shown in the appendix. I will find the parameters of the researcher`s model first and then graph it with my calculator. This is what my TI-84 Plus showed me:

```
Logistic
y=c/(1+ae^(-bx)
a=1.3523588E34
b=.0399301296
c=1617.459658
```

The equation $y = \frac{c}{1+ae^{-bx}}$ equally means/represents
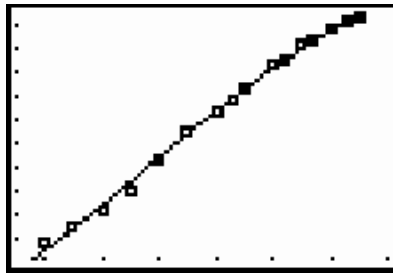
$P(t) = K/(1 + Le^{-Mt})$, where

y equals $P(t)$, a equals L,

Therefore,

$$P(t) = \frac{1617.459658}{1 + 1.3523588 \times 10^{34} \times (e^{-0.0399301296x})}$$

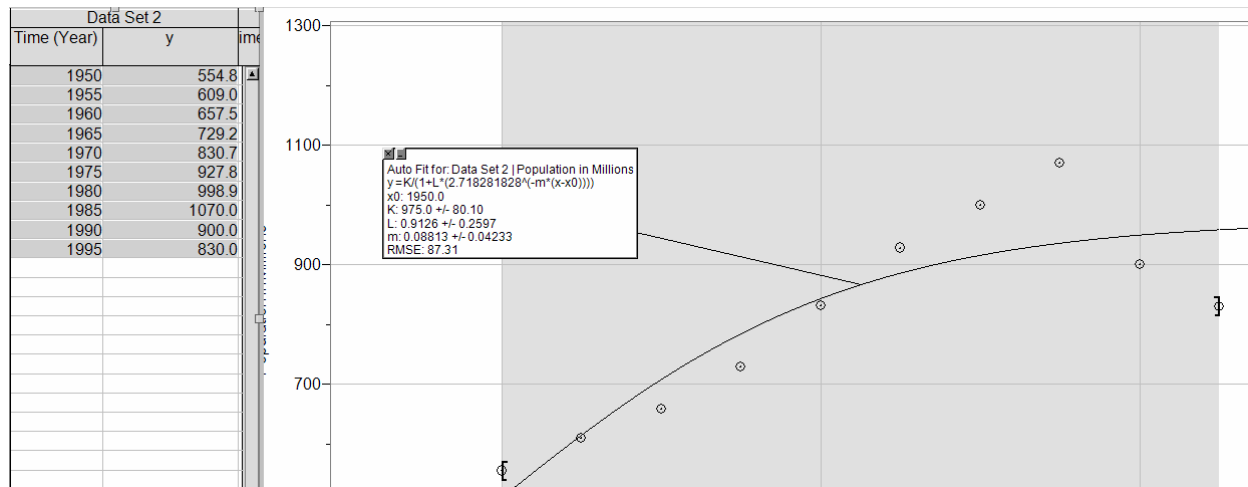$$P(t) \approx \frac{1617.46}{1 + 1.35 \times 10^{34} \times (e^{-0.040x})}$$

Now I will use this new equation to graph again with my TI-84 Plus. Method is shown in the appendix:

Graph 9

As we can see that the graph from my TI-84 Plus matches the graph that I got with the help of Logger Pro. This new modified version of the model fits the data from 1950 to 2008 fits very well as it almost goes through each point. Also there is no big jump in this particular graph. The graph almost looks like it is easily making its way up.

Also, this researcher's model is not always going to work. If there are earthquakes or natural disasters the population will decline. Also it will take time to get back in track. Declining of population is possible but not so common. For example,

Data Set 2		
Time (Year)	y	
1950	554.8	
1955	609.0	
1960	657.5	
1965	729.2	
1970	830.7	
1975	927.8	
1980	998.9	
1985	1070.0	
1990	900.0	
1995	830.0	

Auto Fit for: Data Set 2 | Population in Millions
$y = K/(1+L*(2.718281828^{(-m*(x-x0))}))$
x0: 1950.0
K: 975.0 +/- 80.10
L: 0.9126 +/- 0.2597
m: 0.08813 +/- 0.04233
RMSE: 87.31

Graph 9

The graph will not be able to cover all the points. In short, other graphs should also be used because this graph will not always work, for example when the population actually starts declining.