

FITTING EXPONENTIAL MODELS TO DATA
Supplement to Unit 9C
MATH 1001

In the handout we will learn how to find an exponential model for data that is given and use it to make predictions. We will also review how to calculate the SSE and average error. We will learn how to find the exponential model that best fits a set of given data.

An Alternate Form of the Exponential Function

The exponential function used in the text is

$$Q(t) = Q_0 \times (1 + r)^t.$$

This formula is equivalent to

$$Q(t) = Q_0 \times a^t,$$

where $a = 1 + r$.

Finding an Exponential Model for Data and Making Predictions

First, we consider the following table that gives the population for San Diego, California from 1960 to 1990.

Year	Pop. (thousand)	Change
1960	573	
1970	697	124
1980	876	179
1990	1111	235

The third column of this table shows (for each decade year) the change in population during the preceding decade. We see the *change* in the population of San Diego increased each decade. We might wonder whether this qualifies as almost exponential population growth. So, we will plot the data and look. We plot the year on the x -axis and the population on the y -axis.

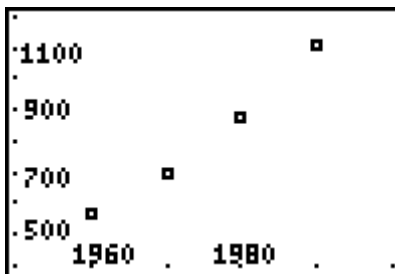


Figure 1

Fitting Exponential Models to Data

We can see that data graphed in Figure 1 has an upward “bend” to it. This indicates that an exponential model may fit the data well. But, how can we find an exponential function that passes through or near each data point? One way is to simply pass an exponential function through the first and the last data points. To make this easier, we will let t be the number of years after 1960. Thus, our data will look like:

t (years since 1960)	P (in thousand)
0	573
10	697
20	876
30	1111

We will use an exponential model of the form

$$P(t) = P_0 \times a^t.$$

We see that the initial population (the population when $t = 0$) is 573 thousand. So, $P_0 = 573$. When $t = 30$, we have $P(30) = 1111$. Thus, we need to solve the equation

$$\begin{aligned} 1111 &= 573 \times a^{30} \\ \frac{1111}{573} &= a^{30} \\ 1.9389\dots &= a^{30} \\ \sqrt[30]{1.9389\dots} &= a \\ 1.0223 &\approx a \end{aligned}$$

Thus, an exponential model for the data is

$$P(t) = 573 \times 1.0223^t.$$

NOTE: Since $1.0223 = 1 + 0.0223$, the annual growth rate is 2.23%.

Figure 2 below shows the data and the exponential modeled graphed together. We can see that the model fits the data very closely.

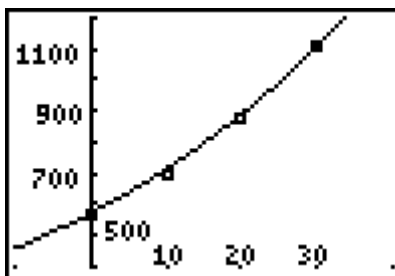


Figure 2

Fitting Exponential Models to Data

Now, let's use our model to predict the population of San Diego in 1985 and in 2000.

The year 1985 is 25 years after 1960. So, to predict the population in 1985, we substitute $t = 25$ into our model.

$$P(25) = 573 \times 1.0223^{25} \approx 994.5$$

Thus, the population of San Diego in 1985 was approximately 994.5 thousand people.

The year 2000 is 40 years after 1960. So, to predict the population in 2000, we substitute $t = 40$ into our model.

$$P(40) = 573 \times 1.0223^{40} \approx 1384.5$$

Thus, the population of San Diego in 2000 was approximately 1384.5 thousand people, or 1.3845 million people.

We can also use the model to predict the year during which the population reaches a certain number. For example, we could use our model above to predict when the population of San Diego will reach 1500 thousand people. To do this, we set $P(t) = 1500$ and solve for t . We will need to use logarithms to do this.

$$\begin{aligned} 1500 &= 573 \times 1.0223^t \\ \frac{1500}{573} &= 1.0223^t \\ \log_{10} \left(\frac{1500}{573} \right) &= \log_{10} 1.0223^t \\ \log_{10} \left(\frac{1500}{573} \right) &= t \log_{10} 1.0223 \\ \frac{\log_{10} \left(\frac{1500}{573} \right)}{\log_{10} 1.0223} &= t \\ 43.63 &\approx t \end{aligned}$$

Thus, the population will reach 1500 thousand in 43.63 years. To determine the calendar year, we add 1960 and 43 to get the year 2003. To determine the month we multiply 0.63 by 12 to get 7.56. Hence the population will be reached 1500 thousand in August 2003.

Measuring How Closely the Model Fits the Data

Now we will find the SSE and average error for the model we found above, which is

$$P(t) = 573 \times 1.0223^t.$$

Fitting Exponential Models to Data

t	P (Actual)	$P(t)$ (Predicted)	Error, E_i $P - P(t)$	E_i^2
0	573	573	0	0
10	697	714.39	-17.39	302.58
20	876	890.68	-14.68	215.51
30	1111	1111	0	0

So, the SSE = $0 + 302.58 + 215.51 + 0 = 518.09$. And the average error is

$$\begin{aligned} \text{average error} &= \sqrt{\frac{518.09}{4}} \\ &\approx 11.38 \end{aligned}$$

Finding the Best-Fit Exponential Model for Given Data

The big question is: how do we find the “best-fit” exponential model for the data. In short, we find it by finding the exponential model that makes both the SSE and average error as small as possible. Our calculators will find the best-fit exponential model for us. The steps are outlined below.

1. Select **STAT, 1:Edit...**
2. Enter the x -values for the data in **L1** and the y -values in **L2**.
3. Press **2nd, MODE** to return to the home screen.
4. Press **STAT** and arrow over to **CALC**.
5. Select **0:ExpReg**.
6. Then enter L1 and L2 (or which ever lists you have your x - and y -values stored in).
7. Press **L1,L2**.
8. Press **ENTER**.

Example:

- (a) Find the best-fit exponential model for the population data for San Diego, California.
- (b) Use your model to predict the population in 1985 and 2000.
- (c) Find the year in which the population will reach 1500 thousand.
- (d) Find the SSE and the average error for the model.

(a) $P(t) = 566.37 \times 1.0224^t$

(b) $P(25) = 566.37 \times 1.0224^{25} \approx 985.44$

$P(40) = 566.37 \times 1.0224^{40} \approx 1373.56$

The population of San Diego was about 985.44 thousand in 1985 and will be about 1373.56 thousand in 2000.

Fitting Exponential Models to Data

(c) For this question, we set $P(t) = 1500$ and solve for t .

$$\begin{aligned}
 1500 &= 566.38 \times 1.0224^t \\
 \frac{1500}{566.38} &= 1.0224^t \\
 \log_{10} \left(\frac{1500}{566.38} \right) &= \log_{10} 1.0224^t \\
 \log_{10} \left(\frac{1500}{566.38} \right) &= t \log_{10} 1.0224 \\
 \frac{\log_{10} \left(\frac{1500}{566.38} \right)}{\log_{10} 1.0224} &= t \\
 43.97 &\approx t
 \end{aligned}$$

Thus, the population of San Diego will reach 1500 thousand in December 2003.

(d)

t	P (Actual)	$P(t)$ (Predicted)	Error, E_i $P - P(t)$	E_i^2
0	573	566.38	6.62	43.8244
10	697	706.83	-9.83	96.6719
20	876	882.11	-6.11	37.3813
30	1111	1100.86	10.14	102.7661

The SSE = 43.8244 + 96.6719 + 37.3813 + 102.7661 = 280.6437. And the average error is

$$\begin{aligned}
 \text{average error} &= \sqrt{\frac{280.6437}{4}} \\
 &\approx 8.38
 \end{aligned}$$

Fitting Exponential Models to Data

Exercises:

In each of problems 1 and 2 the population census data for a U.S. city is given.

- (a) Find an exponential model for the data using the first and last data points. Let $t = 0$ in the year 1960. Round the value for P_0 to two decimal places and the value of a to four decimal places. Use it to predict the population in 2000. Calculate the average error of the model.
- (b) Find the exponential model that best fits this census data. Let t be 0 in the year 1960. Round the value for P_0 to two decimal places and the value of a to four decimal places. Use it to predict the population in 2000. Calculate the average error of the model. Find the annual percentage growth/decay rate.

1. San Antonio, Texas:

Year	1960	1970	1980	1990
Pop. (thous)	588	654	786	935

2. Buffalo, New York:

Year	1960	1970	1980	1990
Pop. (thous)	533	463	358	328

3. As cassette tapes and compact disks became more popular, the sales of vinyl singles declined, as shown in the following table.

Year	1988	1989	1992	1994	1996
Millions of units, s	65.6	36.6	19.8	11.7	10.1

- (a) Find the best-fit exponential model, $s(t) = s_0 \times a^t$, for the data. Let $t = 0$ in 1998. What is the decay rate in sales of vinyl singles?
- (b) Suppose that vinyl singles will be discontinued when their sales fall below 2 million. In what year will this occur?
4. The number of females practicing medicine as MDs is given in the table below for selected years

Year	1980	1985	1990	1993	1994	1995	1996
Number, N , (thous)	48.7	74.8	96.1	117.2	124.9	140.1	148.3

- (a) Find the best-fit exponential model, $N(t) = N_0 \times a^t$, for the data. Let $t = 0$ in 1980. What is the growth rate of female MDs?
- (b) What is the approximate number of female MDs in 1988?
- (c) Approximately how many female MDs were there in 2005?
- (d) In what year will the number of female MDs exceed 300 thousand?

Fitting Exponential Models to Data

5. The table below lists heart disease death rates per 100,000 people in 2001 for selected ages.

Age, x	30	40	50	60	70
Death Rate, R	8.0	29.6	92.9	246.9	635.1

- (a) Find the best-fit exponential model, $R(t) = R_0 \times a^x$, for the data. Let x be the actual age. Round both R_0 and a to four decimal places.
 (b) Estimate the heart disease death rate for people who are 80 years old.

Answers:

1. (a) $P(t) = 588 \times 1.0156^t$; $P(40) \approx 1092$ thousand; average error ≈ 17.95
 (b) $P(t) = 575.67 \times 1.0159^t$; $P(40) \approx 1082$ thousand; average error ≈ 13.07 ; growth rate: 1.59%
2. (a) $P(t) = 533 \times 0.9839^t$; $P(40) \approx 278$ thousand; average error ≈ 14.49
 (b) $P(t) = 533.53 \times 0.9830^t$; $P(40) \approx 269$ thousand; average error ≈ 13.15 ; decay rate: 1.7%
3. (a) $s(t) = 53.43 \times 0.7954^t$; decay rate: 20.46% per year
 (b) May 2002
4. (a) $N(t) = 50.21 \times 1.0692^t$; growth rate: 6.92% per year
 (b) 85,755
 (c) 267,463
 (d) September 2006
5. (a) $R(x) = 0.3525 \times 1.1148^x$
 (b) 2103 deaths per 100,000