Section 1 Introduction

It is not uncommon to be working with two variables e.g. x and y that obey a fixed but unknown relationship $y = f(x)$ between them. The function $f(x)$ relating the two variables may not be as important as merely having the ability to generate (by measurement or experimental observation) numerical values (x_i, y_i) where $y_i = f(x_i)$. For example, the performance of a component may be related to a design parameter in a complicated fashion which is difficult to describe in mathematical terms. In a controlled environment, the lone parameter can be varied and the resulting performance monitored.

Suppose a consumer testing magazine is interested in the repair costs for a particular luxury passenger vehicle after its been in an accident. The test procedure calls for head-on collisions of the vehicle with a stationary object at various speeds. The cost of restoring the vehicle to its pre-crash condition is determined. Tabulated results after several collisions are given below.

Speed (S) mph	Damages (D) \$
10	4500
20	19750
30	43500
40	55000

Table 1.1 Experimental Vehicle Crash Test Data

The table represents a sample of four data points obtained from an unknown function, $D = f(S)$. The experimental process is rather expensive and further experimentation might well be cost prohibitive. The magazine would like to publish a graph for its readers to estimate the cost of repairing the same vehicle resulting from collisions occurring over a range of speeds.

One approach to approximating the unknown function $D = f(S)$ is illustrated in Figure 1.1. The approximating function is the result of connecting the data points from Table 1.1 by straight line segments. The dotted curve represents the actual (unknown) function $D = f(S)$. To estimate the damages from a 15 mph collision, the approximating function is evaluated at this speed as shown.

The process of approximating a function $f(x)$ by another function and subsequent use to estimate numerical values of $f(x)$ is referred to as interpolation. The original function $f(x)$ may be known in analytical form or possibly only through tabulated values that come from it. The approximating function, hereafter referred to as the interpolating function, is ordinarily much simpler than the underlying function *f*(*x*). In Figure 1.1, the interpolating function is a piecewise linear function defined by the data points of *f*(*x*).

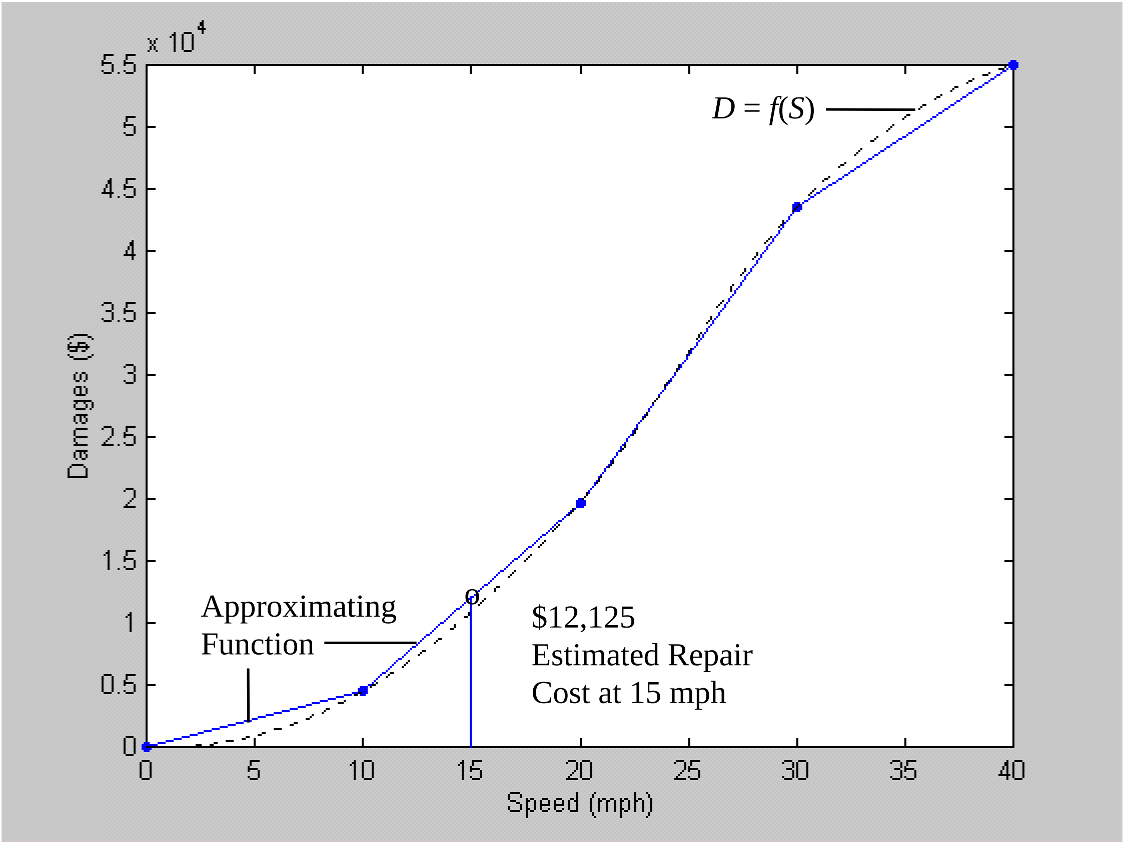


Figure 1.1 An Approximating Function For the Data in Table 1.1

Interpolating functions are generally restricted to an interval containing the data points used to generate them. In this example, estimated repair costs should be confined to speeds from 0 to 40 mph (an additional data point can be assumed at *S* = 0 mph, *D* = $0). Attempting to estimate function values outside the range of the data points is termed extrapolation. Caution should be used when extrapolating since erroneous and even nonsensical results are possible.

Interpolation of data representing future predictions is not extrapolation. The two sets of data points graphed in Figure 1.2 represent historical data as well as forecasts of the future behavior of two economic variables. Presumably, sound economic forecasting methods were employed in the process of extrapolating the future values. Regardless, estimating either quantity from 1997 through 2025 is an example of interpolation.

The piecewise linear interpolating function is a suitable approximating function under the right conditions. By in large, if the underlying function *f*(*x*) is relatively smooth and well behaved, and the data points are reasonably spaced, then a "connect the dots" approach using piecewise linear interpolation will produce satisfactory results.

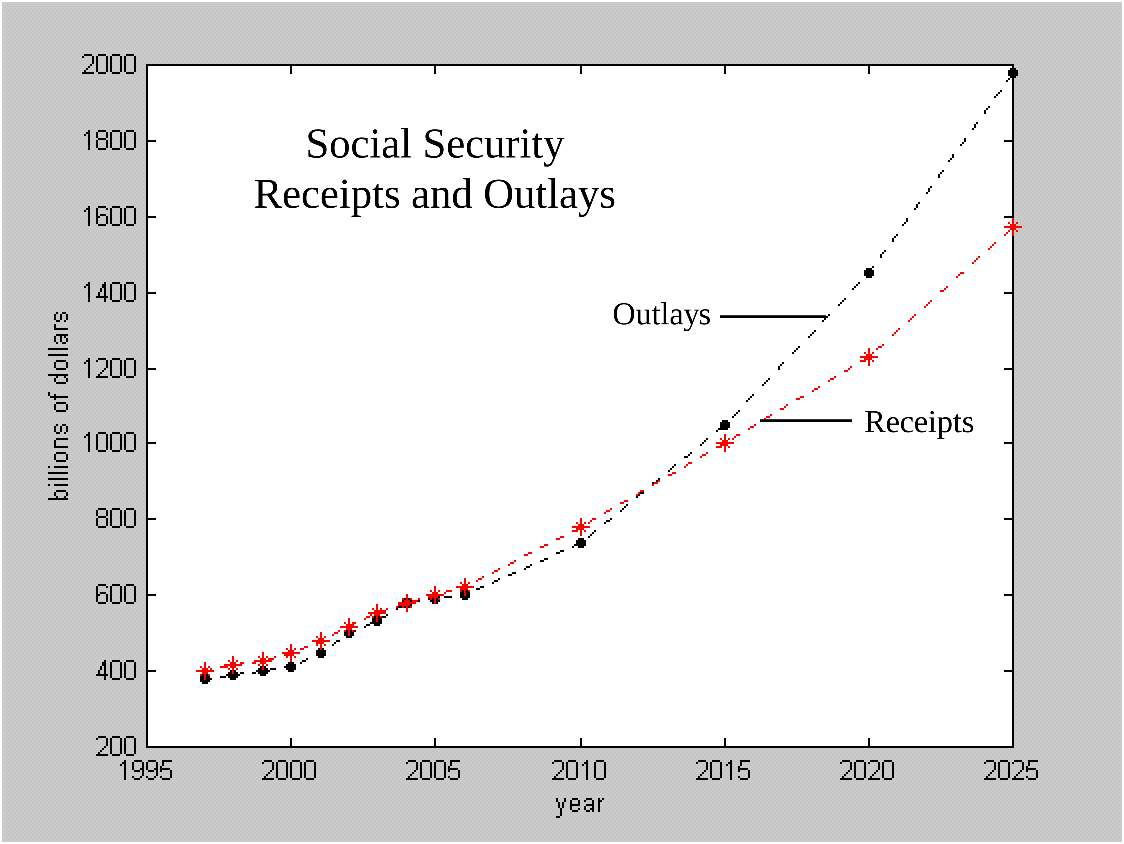


Figure 1.2 Data Points from Functions Predicting Future Values

As an example, consider the case where the function to be approximated is actually $y = f(x) = \sin x$. The table below contains equally spaced values of $\sin x$ from 1 to 2 rad.

i	x_i	$y_i = f(x_i) = \sin x_i$
0	1.00	0.8415
1	1.25	0.9490
2	1.50	0.9975
3	1.75	0.9840
4	2.00	0.9093

Table 1.2 Several Points From the Function $f(x) = \sin x$

The piecewise linear interpolating function connecting the five data points and the function $\sin x$ are shown in Figure 1.3. By observation, it appears that the approximating function is within a few percent of the sine function over the entire interval.

Suppose it is necessary to estimate the sine of π/2 rad. The *y* coordinate of point P_1 in Figure 1.3 is the estimate, a reasonably close approximation to the actual value of sin π/2. On the other hand, if sin π/2 was estimated from the line joining points (1, 0.8415) and (2, 0.9093) then the result would be the *y* coordinate of point P_2, which differs significantly from the correct value.

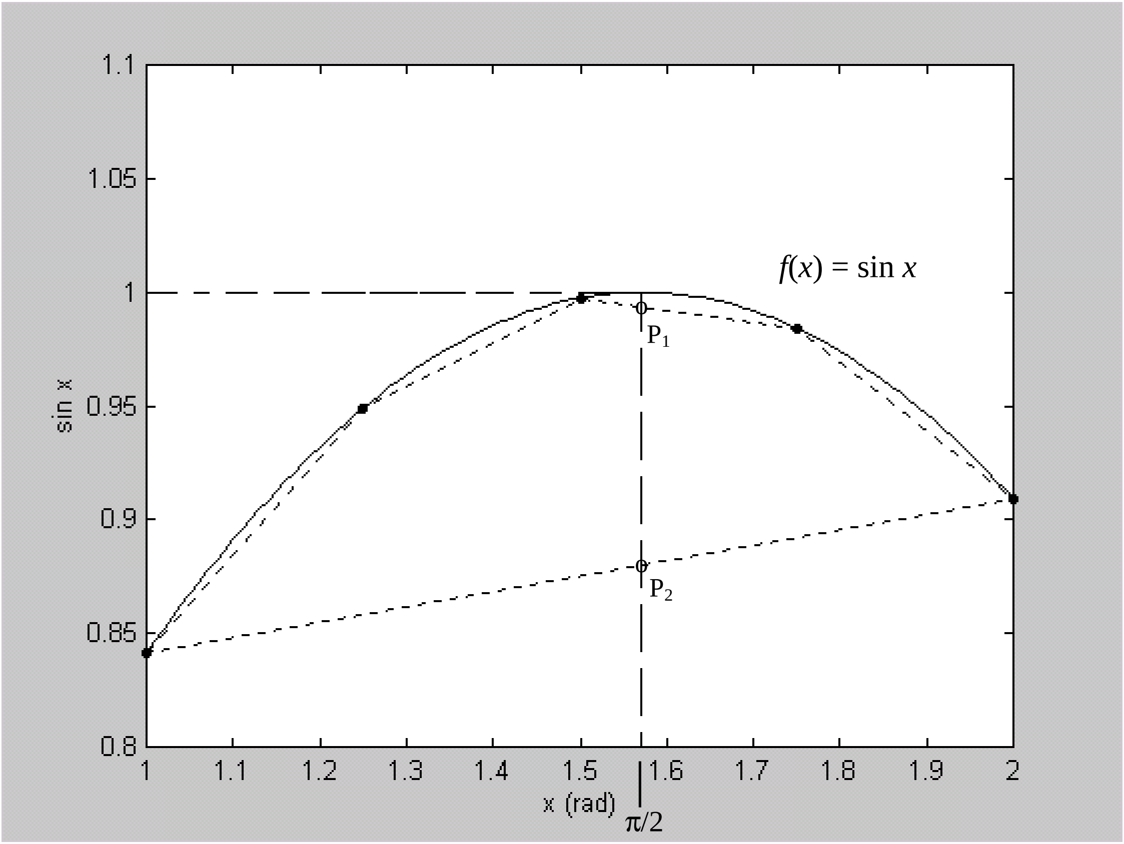


Figure 1.3 Linear Interpolation of *f*(*x*) = sin *x* to Estimate sin π/2

This piecewise linear approach to approximation of functions is the method used in graphing software to plot known analytical functions. The function is evaluated at equally spaced points which are then connected by line segments. The graph appears as a smooth curve when the spacing of points is relatively close. Figure 1.4 illustrates how the approximating function becomes smooth as the points where the function is evaluated become more dense. The function being approximated, i.e. graphed is $f(x) = e^x$.

Piecewise approximating functions are not limited to the case where a linear function approximates the true function between data points. Later in the chapter we consider other polynomial approximations to the underlying function *f*(*x*) between consecutive data points.

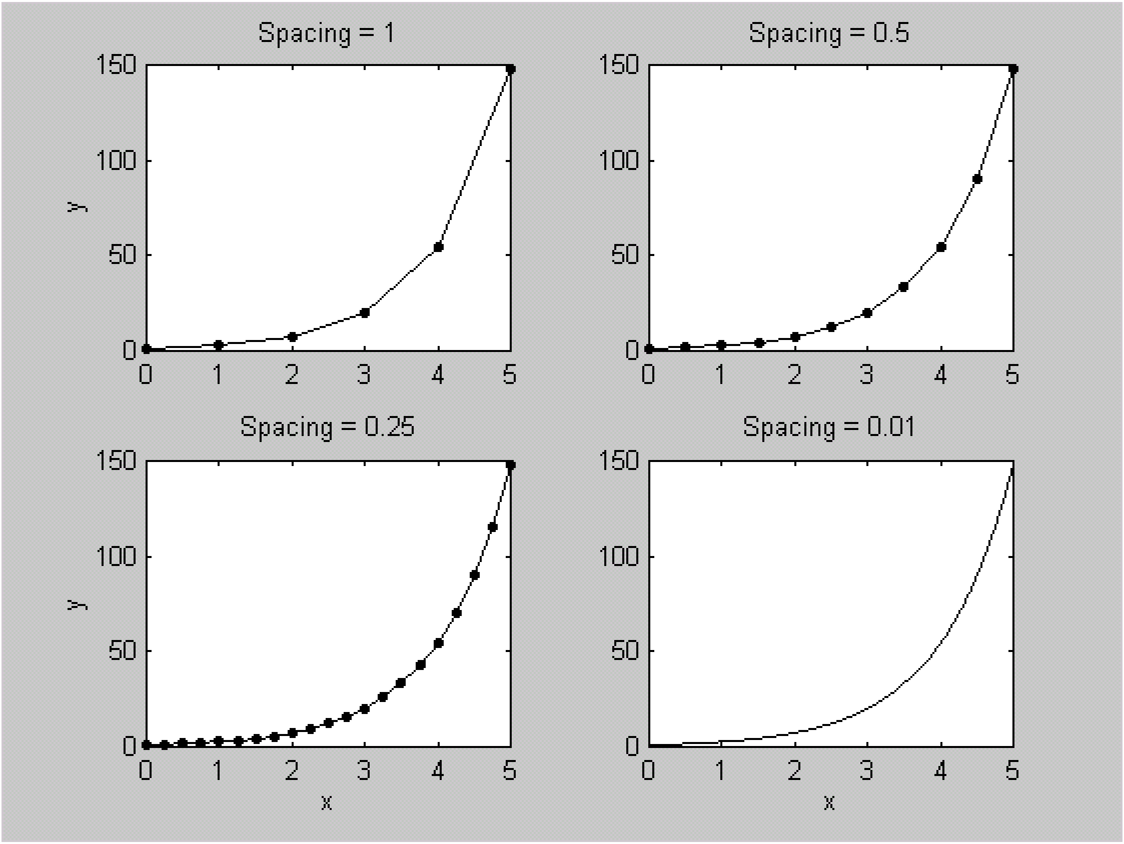


Figure 1.4 Graph of $f(x) = e^x$ Using Linear Interpolation Between Sampled Points

In contrast to the piecewise approximation of a function, interpolating functions are frequently composed of a linear combination of elementary functions $\phi_i(x)$, $i = 0,1, ,$ n. That is, a function $f(x)$ is approximated by an interpolating function $I(x)$ given by

$$I(x) = a_0\phi_0(x) + a_1\phi_1(x) + \ldots. + a_n\phi_n(x) \tag{1.1}$$

A common choice for the elementary functions $\phi_i(x)$ is the monomial function x^i. In this case , the interpolating function is a polynomial, denoted by $f_n(x)$ where

$$f_n(x) = \sum_{i=0}^{n} a_i x^i = a_0 + a_1 x + a_2 x^2 + \ldots. + a_n x^n \tag{1.2}$$

The *n*th order Taylor Series expansion of a function $f(x)$ introduced in Chapter 1 is a good example of the use of polynomial functions intended for approximation purposes. Interpolating polynomials are discussed in detail in the remaining sections.

There are several reasons why interpolating polynomials are popular when it comes to approximation of functions. There are efficient algorithms for evaluating polynomials when "*n*" is large or the polynomial must be computed numerous times with different arguments. Differentiation and integration of complex functions is often required. In some instances the function is only available in tabular form. Polynomial approximations of these functions are easily differentiated and integrated.

Exercises

1. The value of the Social Security trust fund at the end of several years is given in the table below. (Source: U.S. Social Security Administration)

Year	Fund Value at End of Year (in millions of dollars)
1985	$ 35,842
1987	62,149
1989	155,063
1991	267,849
1993	369,322
1995	458,502

Data for Problem 1

   a) Plot the data points and the piecewise linear function connecting the data points on the same graph.

   b) Estimate the fund's value at the end of 1994 from a line connecting the following pairs of data points: (1985, 35842) and (1995, 458502), (1987, 62149) and (1995, 458502), (1989, 155063) and (1995, 458502), (1991, 267849) and (1995, 458502), (1993, 369322) and (1995,458502).

   c) The fund's value at the end of 1994 was $413,460 million. Calculate the true error for each estimate in Part b).

   d) Use the piecewise linear interpolating function to estimate the fund's balance at the end of June 1990.

   e) Use the last two data points to extrapolate the year and month that the trust fund exceeded $ 5 billion. (The actual occurrence was December 1996)

2. U.S. Census data at the end of selected decades is portrayed in bar chart form. (Source: U.S. Census Bureau)

   a. Estimate the U.S. population in the year 1900 using a linear approximating function connecting data points for the years

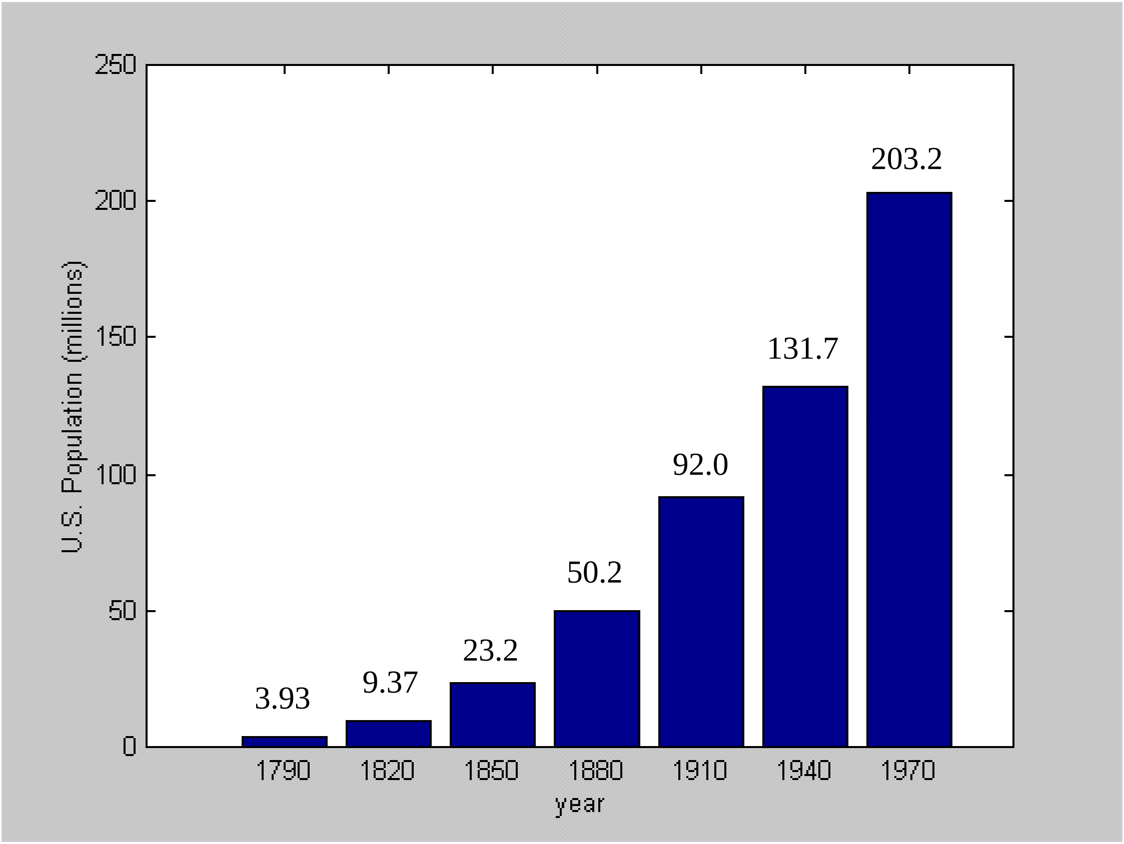
      i) 1880 and 1910
      ii) 1850 and 1940
      iii) 1820 and 1970

   b. U.S. population in 1900 was 76.0 million. Calculate the errors for the estimates obtained in Part b).

c. Draw the piecewise linear approximating function through the data points and use it to estimate the years in which the U.S. population reached 100 million and 200 million.

d. Extrapolate the U.S. population in 1980 and 1990 using the last two data points.

e. Sketch a smooth curve through the first four data points and use it to interpolate the U.S. population in the year 1840. Repeat the process for the last four data points and estimate what the U.S. population was in 1960. (The actual populations were 17.1 million in 1840 and 179.3 million in 1960)



Graph for Problem 2

3. Solar radiation is an important factor in computing the cooling load requirements of an air-conditioner. Measurements of solar radiation through 1/8 in. sheet glass facing East were taken at different times of the day at several latitudes on August 21. The results are presented in tabular form. (Source: ASHRAE Handbook of Fundamentals - 1972)

   a) Plot the data points with solar radiation on the vertical axis and latitude on the horizontal axis.

   b) Plot the piecewise linear function for each time of day.

   c) Estimate the solar radiation at 45° latitude at 9 am, noon, 3 pm and 6 pm.

d) Plot the data points with solar radiation on the vertical axis and time of day on the horizontal axis.

e) Plot the piecewise linear function for latitude.

f) Estimate the solar radiation at 11 am for each latitude.

	Latitude				
Standard Time	24°	32°	40°	48°	56°
9 am	202	202	199	195	188
noon	71	70	68	65	61
3 pm	32	31	30	28	26
6 pm	9	9	10	11	11

Solar Radiation Heat Gain Through Glass (Btu/hr per ft^2)

4. Humidifiers are rated in terms of capacity, i.e. gallons of water per day it can evaporate. The moisture requirements of a tight house (well insulated) and an average house of different sizes are tabulated below. The figures are based on a Winter day with outdoor dry-bulb temperature 20°F, 80% relative humidity, an indoor dry-bulb temperature of 70 °F and 40% relative humidity. (Source Refrigeration and Air-Conditioning,, ARI Institute, Prentice-Hall)

VOLUME OF RESIDENCE ft^3	TIGHT HOUSE Gallons per Day	AVERAGE HOUSE Gallons per Day
8,000	5.09	10.17
16,000	10.18	20.35
20,000	12.72	25.44
28,000	17.81	36.51

Table for Problem 3.8

a) Plot the data points for both houses on the same graph.

b) Draw the piecewise linear approximating functions for estimating humidifier rating from the residence volume for both types of house.

c) Estimate the size humidifier required for a tight house with 10,000 ft^3 of space.

d) The humidifier in an average house must evaporate 31.5 gallons of water per day to maintain the design indoor conditions on a day when outdoor conditions are the design values. Estimate the volume of the house.

e) Are the answers to Part c) and d) exact values or simply estimates? Explain.

5. A digital sound file is obtained by sampling an analog audio signal using an analog to digital (A/D) converter. The digital file is played back by the reverse process using a digital to analog (D/A) converter. The following table lists the kb of memory required to store 1 second of audio for various sampling rates using two different A/D converters. It also shows the duration of audio that can be stored in 1 Mb of memory for different sampling rates.

	8 bit A/D Converter		12 bit A/D Converter	
Sampling Rate f (kHz)	Memory Storage for 1 sec of Sound S (kb)	Length of Sound Stored in 1 Mb T (sec)	Memory Storage for 1 sec of Sound S (kb)	Length of Sound Stored in 1 Mb T (sec)
4	4	250	8	125
20	20	50	40	25
40	40	25	80	12.5

Table for Problem 5

a) Plot the data points (f, s) for both A/D converters on the same graph.

b) Determine the equation of the interpolating polynomials for predicting memory storage requirements (for 1 second of sound) as a function of sampling rate for each type of converter.

c) Plot both polynomials on the same graph with the data points.

d) Plot the data points (f, t) for both A/D converters on a different graph.

e) Determine the equation of the interpolating polynomials for predicting the length of sound stored in 1 Mb as a function of sampling rate for each type of converter.

f) Plot both polynomials in Part e) on the same graph with the data points of Part c).

g) A telephone conversation is to be digitally recorded. Assume the sampling rate will be at twice the highest frequency component of the audio signal, roughly 4 kHz. Estimate the memory storage required to store a one minute conversation using an 8 bit A/D converter.

h) An FM radio signal is to be digitally encoded using a 12 bit A/D converter. The frequency range is 20Hz - 16kHz and the sampling is performed at 8 times the highest frequency component. How many Mb of memory are needed to store a 3 minute song?

i) Is your answer to Part h) the result of interpolation or extrapolation? Is extrapolation reliable in this situation? Explain.