Extra Online Questions

Use a mathematical model involving curve fitting to solve a problem

Scholarship Statistics and Modelling

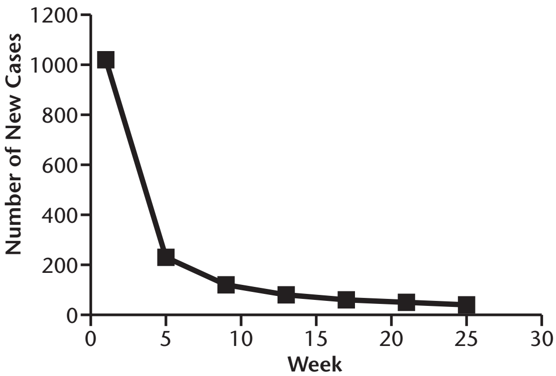
Covers AS 90647 (Statistics and Modelling 3.7)

Chapter 7

Using curve fitting to solve problems

1. Health officials are investigating a recent epidemic and the rate at which it spread. This table shows the number of new cases reported at the end of each week. The data have been recorded at 4-weekly intervals, and the number of cases has been rounded to the nearest 10. Ans. p. 8

Week (t)	$\log_e(t)$	Number of new cases (N)	$\log_e(N)$
1	0	1020	6.93
5		230	
9		120	
13		80	
17		60	
21		50	
25	3022	40	3.69



Two possible models have been suggested to explain the rate at which the epidemic is spreading.

(1) $N = at^n$ or (2) $N = ae^{kt}$

a. Explain how, by taking logs, you could check whether each of these models is a good fit to these data. Include a mathematical derivation of this process for both models.

b. Which model is more appropriate for this data? You should draw graphs using the grids below to support your answer.

c. Use the values for the end of weeks 1 and 25 to estimate the values of the constants a and n if the model $N = at^n$ is used.

2. A botanist is researching a particular micro-organism's population. The data that follow were collected at particular time intervals, as indicated, measured in days. The population is measured in thousands. Ans. p. 9

Time t (days)	1	2	3	4	5	10	15	20
Population P (×1000)	3.4	22.1	66.0	143.6	262.2	1 704.0	5 092.4	11 072.9

a. Use the following grid to draw an appropriate graph and use it to determine a model that fits these data. Justification for the model choice must be given, including all/any working.

b. The botanist knows that every thousand organisms take up 4 mm^2 of space in the cultivation container. Predict the number of days the organism will take to fill a particular cultivation container of area 345 cm^2. Limitations, usefulness and appropriateness of this prediction and model as a whole should be discussed.

3. A scientist had observed that a particular chemical reaction changes over a period of time. The chemical reaction is aided by a catalyst. The following data were collected: Ans. p. 10

Time t (seconds)	1	3	5	8	12	15
Mass of Catalyst M (grams)	8.044	3.614	1.624	0.489	0.099	0.030

a. Obtain a model that represents this data, show all appropriate working for why this is a valid model.

b. Using the model, predict how much of the catalyst was present after 4 minutes.

c. How much catalyst was initially in the reaction?

d. Describe the main observation the scientist made about the reaction as time passed.

3.7 Using curve fitting to solve problems (page 1)

1. **a.** For model 1:

$$N = at^n$$
$$\log(N) = \log(at^n)$$
$$\log(N) = \log(t^n) + \log(a)$$
$$\log(N) = n\log(t) + \log(a)$$

This is a linear relationship between $\log(N)$ and $\log(t)$

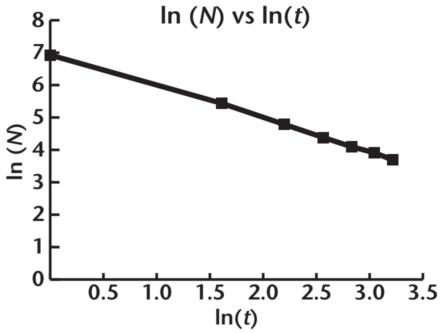
For model 2:

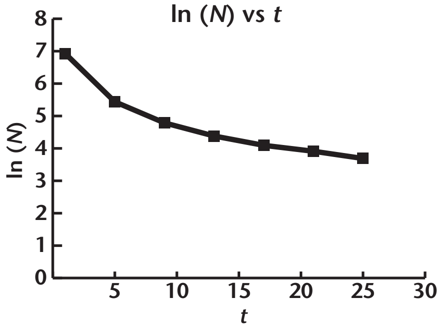
$$N = ae^{kt}$$
$$\ln(N) = \ln(ae^{kt})$$
$$\ln(N) = \ln(e^{kt}) + \ln(a)$$
$$\ln(N) = kt + \ln(a)$$

This is a linear relationship between $\log(N)$ and t.

Hence by graphing $\log(N)$ vs $\log(t)$ and $\log(N)$ vs t (and determining which graph is better approximated by a straight line) the model that fits best can be determined.

b. The following graphs result:





It should be clear that the more linear graph is $\log(N)$ vs $\log(t)$ hence, model 1 ($N = at^n$) is the better choice.

c. Using natural logs, the linear relationship is $\ln(N) = \ln(a) + n\ln(t)$. Using the suggested values the following equations can be formed:

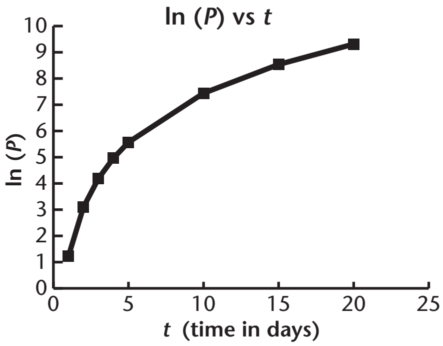
$$\ln(40) = \ln(a) + n\ln(25)$$
$$\ln(1020) = \ln(a) + n\ln(1)$$

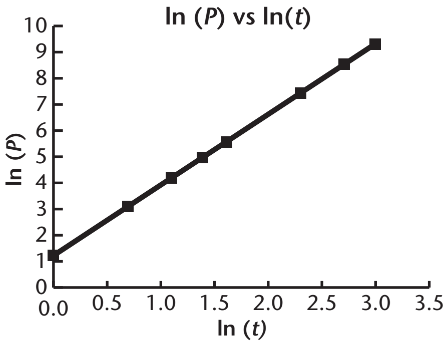
The second equation gives $a = 1\,020$. since $\ln(1) = 0$

Using this in the first statement, the value of n can be found.

$$\ln(40) = \ln(1020) + n\ln(25)$$
$$n = \frac{\ln(40) - \ln(1020)}{\ln(25)}$$
$$= -1.006$$

2. **a.** Two possible graphs can be drawn; log-log or semi log. The one that is more linear is the better model. The two graphs are shown below:





Clearly the log-log graph is more linear, hence the model that will best fit this data is:

$P = at^n$

The log transformation of this model into a linear relationship allows the unknown constants (a and n) to be found.

$P = at^n$

$\ln(P) = \ln(at^n)$

$\ln(P) = n\ln(t) + \ln(a)$ log laws

Comparing this with $y = mx + c$ gives:

n = gradient = 2.7

$\ln(a)$ = y-intercept = 1.2

$a = e^{1.2} = 3.32$

From this information the model for these data would be:

$P = 3.32t^{2.7}$ substituting in $P = at^n$

b. For the prediction, the number of organisms that fit in the container is equal to:

$\dfrac{345 \times 100}{4}$ = 8 625 thousand 1 cm² = 100 mm²

To predict the number of days to reach this number, an equation needs to be solved.

$8\,625 = 3.32t^{2.7}$

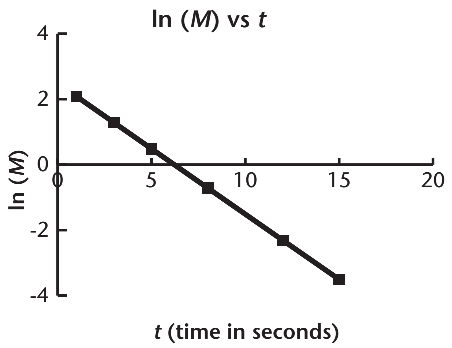
$t^{2.7} = 2\,597.89$ dividing by 3.32 and rearranging

$t = (2\,597.89)^{1/2.7}$ taking reciprocal power of both sides

$t = 18.39$ days

Hence it would take approximately 18.4 days for the organism to fill this particular cultivation container.

3. **a.**



As this graph of ln(M) vs t is linear, the model is of the form $M = ae^{kt}$.

Justification of the linearity test also given:

$$\ln(M) = \ln(ae^{kt})$$

$$\ln(M) = \ln(e^{kt}) + \ln(a)$$

$$\ln(M) = kt + \ln(a)$$

Unknown constants:

$$k = -0.4$$

$$\ln(a) = 2.48$$

$$a = 11.94$$

It follows that the model is:

$$M = 11.94e^{-0.4t}$$

b. When $t = 4$ the mass of catalyst left is predicted:

$$M = 11.94e^{-0.4 \times 4}$$

$$= 2.41 \text{ g}$$

c. By substituting $t = 0$, the initial amount of catalyst according to the model found in **a.** is 11.94 g. However, a more likely conclusion would be to say there was 12 g at the start.

d. The main observation would be that the reaction must have slowed down as time passed, due to the fact that the amount of catalyst present decreased. Since the catalyst was there to speed up the reaction, a slowing would be seen.