

Statistics 520 - Assignment 6

Sam Olson

Introduction

This assignment is a short exercise in using generalized least squares to estimate the parameters of a nonlinear regression model.

There are a number of rather classic nonlinear curves used to model the growth of populations of organisms. Although their use is not strictly tied to one field, different curves have been developed by scientists concerned with the growth of animal populations, plant populations, and microbial populations. A great many papers have been written on the origins, estimation, and interpretation of these curves (e.g., Birch, C. 1999. A new generalized logistic sigmoid growth equation compared with the Richards growth equation. *Annals of Botany* 83: 713–723).

One of these traditional curves is known as the Gompertz model, and there are many versions of the Gompertz curve (e.g., Tjorve, K.M.C. and Tjorve, E. 2017. The use of Gompertz models in growth analysis, and new Gompertz model approach: an addition to the unified Richards family. *PLoS One*, 12(6): e0178691). We will use a four-parameter version of the Gompertz model as the expectation function in a nonlinear regression model.

The Model

Define random variables Y_1, \dots, Y_n connected with the growth of some organism, observed at times t_1, \dots, t_n . Assume the response variables are independent and follow the model:

$$Y_i = \mu_i + \sigma \varepsilon_i,$$

$$\mu_i = B + A \exp \left[- \exp \{ -k(t_i - T) \} \right],$$

$$\varepsilon_i \stackrel{\text{iid}}{\sim} N(0, 1).$$

This model describes a sigmoidal curve with both upper and lower asymptotes, but is not constrained to be symmetric about its inflection point in the way a logistic curve is.

Parameters

- A: distance between upper and lower asymptotes
- B: lower asymptote
- k: growth rate (slope of the linear portion of the curve)
- T: the time at which the inflection point occurs

Data

On the course web page in the Data module is a file `growthdat.txt` that contains data simulated from the Gompertz regression model. The variable names in this file are `x` and `y`, where:

- `x` = time of observation
- `y` = growth in appropriate units

Assignment

Q1

Find generalized least squares estimates of the parameters A, B, k, T , and the associated moment-based estimate of σ^2 .

- Describe how you determined effective starting values for the estimation procedure.
- Present a scatterplot along with the fitted expectation function.

Q2

Compute 95% approximate confidence intervals for the parameters of the expectation function (A, B, k, T) .

Q3

Compute pairwise correlations between $\hat{A}, \hat{B}, \hat{k}, \hat{T}$.

Q4

Scientists often care about the maximum relative growth rate and the maximum absolute growth rate. These relate to the slope of the growth curve at the inflection point:

- Maximum absolute growth rate:

$$k_{\text{abs}} = \frac{k(A + B)}{\exp(1)}$$

- Plug-in estimator:

$$\hat{k}_{\text{abs}} = \frac{\hat{k}(\hat{A} + \hat{B})}{\exp(1)}.$$

- Compute a 90% approximate confidence interval for k_{abs} .
- Outline the procedure you used to calculate the required quantities.

Computational Options

There are multiple ways to perform the computations:

- The base R function `nls()` (nonlinear least squares) can be used for most of the tasks.
- Alternatively, there is an R function called `nonlin` available on the course web page (Computing Functions module), with a supporting document Using nonlin in the Computing Notes module explaining syntax and output.