Modeling bacterial growth

Models are used to describe the behavior of of microorganisms under different physical or chemical conditions such as temperature, pH, etc. Bacteria typically grow exponentially, so they are commonly plotted as log(relative population size over time).

What is the simplest model of population growth?

$$dN/dt = rN$$
$$N(t) = N(0)e^{rt}$$
$$lnN(t) = ln(N(0)) + rt$$
$$lnN(t) = a + rt$$

Typical phases of bacterial growth:

- Growth rates starts at zero and begins slowly (lag phase)

- Growth accelerates to a maximal value exponentially (log phase)

- Growth stops and reaches an asymptote (stationary phase)

Parameters biologists are interested in:

- Lag period (time from zero to maximal value of growth λ)

- Maximum specific growth rate - slope of line when growth is plotted on log scale (μ_m)

- Asymptotic value A

-
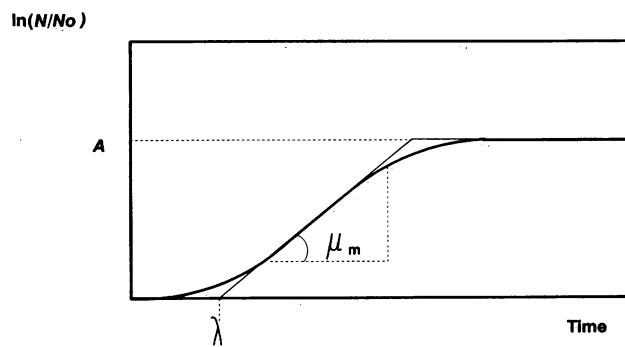

FIG. 1. A growth curve.

Logistic

How do we modify simple exponential model to include some sort of carrying capacity?

$$dN/dt = rN(1 - N/K)$$
$$\text{Divide both sides by K}$$
$$d/dt N/K = rN/K(1 - N/K)$$
$$x = N/K$$
$$dx/dt = rx(1 - x)$$
$$dx/x(1 - x) = 1/x + 1/(x - 1)dx = rdt$$
$$N(t) = \frac{KN(0)}{N(0) + (K - N(0))e^{-rt}}$$
$$N(t) = \frac{a}{1 + e^{b-cx}}$$

What does a logistic curve look like? Figure it out based on differential equation - sigmoidal and then it reaches an asymptote

Gompertz

Also sigmoidal function, more parameters than logistic function - growth is slowest at the start and end of a time period
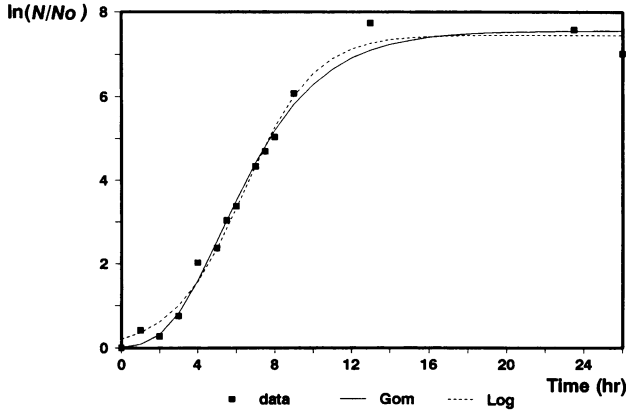
$$y = ae^{(-e^{b-cx})}$$

FIG. 6. Growth curve of *L. plantarum* at 40.0°C fitted with the Gompertz (Gom) and logistic (Log) models.

Want to modify the equation for biologically relevant parameters: A, λ, and μ_m. Because bacteria grow exponentially, it's useful to plot bacteria as the logarithm of the relative population size. Maximum specific growth rate is defined as the tangent in the inflection point. The lag time is the x-intercept of this tangent, and the asymptote is the maximal value reached. There are many models of population growth but they aren't written in terms of growth rate, lag time, and asymptotic value, which makes interpretation of the parameter values difficult.

$$y = ae^{(-e^{b-cx})}$$
$$dy/dt = ae^{(-e^{b-ct})} - e^{b-ct}(-c) = ace^{(-e^{b-ct})}e^{b-ct}$$
$$d^2y/dt^2 = ac^2exp(-exp(b-ct)) * exp(b-ct) * (exp(b-ct) - 1)$$
$$d^2y/dt^2 = 0 => t_i = b/c$$
$$y(t_i) = a/e$$
$$\mu_m = (dy/dt)_i = ac/e$$
$$c = \mu_m e/a$$

The description of the tangent line through the inflection point is:

$$y - y(t_i) = \mu_m(t - t_i)$$
$$y = \mu_m t + a/e - \mu_m t_i$$

The lag time is defined as the t-axis intercept of the tangent through the inflection point. So we plug in λ for t and 0 for y in the previous equation.

$$0 = \mu_m \lambda + a/e - \mu_m t_i$$

Solve for lambda using values for μ_m and t_i.

$$\lambda = (b-1)/cb = \mu_m e/a)\lambda + 1$$
$$A = a$$
$$y = Aexp(-exp(\mu_m e/A(\lambda = t) + 1))$$

You can do the same thing for logistic curves.

Model fit

Fit the data using the biologically relevant parameters by some nonlinear regression that minimizes the sum of squared error between data and model. This sum of squared error is denoted RSS, residual sum of squares.

$$RSS = \sum (logy_i - log\hat{y}_i)^2 \qquad (1)$$

y_i = data, $\hat{y}_i$ = model fit

Model comparison

How do you compare different models? Models with more parameters typically result in a smaller RSS. The F- test is a way to compare RSS values and tells you the profit of adding a parameter, i.e. is the decrease in RSS due to an extra parameter worth adding the extra parameter?

$$f = \frac{\frac{RSS_1 - RSS_2}{DF_2 - DF_1}}{RSS_2/DF_2}$$

Model 1 has few parameters (and bigger RSS) than parameter 2. Under the null hypothesis that model 2 does not provide a significantly better fit than model 1, F will have an F distribution. The null hypothesis is rejected if the f statistic calculated from the data is greater than the critical value of the F-distribution for some desired false-rejection probability (e.g. 0.05).