Statistical Science

Chapter 4 Equations (Formal Models)

ReCap. Quantitative Reasoning(Ch 1)

Quantities (Ch2) Re-Scaling (Ch3)

- 4 Equations
- 4.1 Introduction Uses
- 4.2 Reading Equations
- 4.3 Homogeneity of Units
- 4.4 Dimensional Homogeneity
- 4.5 Review Questions

Lab 2

Ch 4.1 to 4.4 are background for equations lab.

Bring shells, Apple Murex and Lightning Whelk

on chalk board

Recap Chapter 1

Quantitative reasoning: Example of scallops, which combined stats and models

Recap Chapter 2

Quantities: Five part definition

Measurements made on four types of scale: nominal, ordinal, interval, ratio

Ratio scale units grouped into dimensions

Recap Chapter 3. Re-scaling

Logical rescaling (from one type of unit to another).

Rigid rescaling replaces one unit with another

It is the basis for conversion from one scale to another.

Simple rescaling reduces a measured quantity to dimensionless numbers Normalization reduces a measured quantity to dimensionless numbers relative to a statistic (mean or standard deviation).

Today: Equations in the natural sciences

Equations express ideas about the relation of one quantity to others Learning to read equations is a matter of practice, there are a number of techniques for learning to read and use equations.

Wrap-up

Equations have several uses, ranging from the highly theoretical to simple demonstration: "this is how I calculated diversity $H' = \sum_{i} p_{i} \ln p_{i}$

where p_i is the proportion of organisms of species i in a collection. "

Equations are a mixture of variable and parametric quantities.

Equations have units and dimensions.

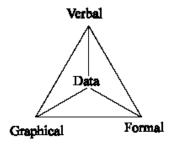
Equations must be homogeneous with respect to units and dimensions.

This principle is used to check whether equations are correct

This principle is also used as an aid in reading and understanding equations

4.1 Introduction – Uses of Equations in Natural Science

Biological concepts can be expressed in words (an informal or verbal model), or in graphs (a diagrammatic model), or in equations (a formal or mathematical model). These three forms of expression are related to one another.



In biology, the most common form of expression of ideas is verbal. Graphical expression of ideas about the relation of one quantity to another includes diagrams of many types. To make calculations, we need an equation.

Equations express ideas about the relation between several quantities.

Here is example. "Bacterial growth rate is exponential" $N_t = N_\theta e^{0.6 * t}$

Equations are the basis for most of quantitative biology, including statistical analysis.

Equations have several uses.

<u>Demonstrative</u>: how a derived quantity was calculated "this is how I calculated diversity;" $H'= \sum p_i \ln p_i$

where p_i is proportion of a collection due to species i.

<u>Theoretical</u>: derive general conclusions from set of premises

This use requires sound knowledge of mathematical procedures.

It is an important part of a few parts of physics, many areas of chemistry and several areas of the earth and ocean sciences and biology, such as population dynamics. It will not be treated in this course.

<u>Verification</u>: testing of ideas (hypotheses) against data

This will be covered in this course.

The fit of data to models will be a major part of this course.

Scale experiments and observations up to larger scope

This is an area of active research in environmental biology.

It will not be covered, though some of the material already presented, on units and dimensions, is critical to this activity.

4.2 Reading equations

Dissect into components, then reconstruct meaning.

To translate from a foregn language in which we are not fluent we dissect sentences into words, then construct the meaning of the sentence.

Begin by dissecting equations into <u>terms</u>. Terms are connected by + - = signs

$$\dot{E}_{net} = \dot{E}_{ingested} + \dot{E}_{resp} + \dot{E}_{growth}$$

In this example, net energy balance \dot{E}_{net} is the sum of

Ingestion \dot{E}_{ingest} metabolic losses \dot{E}_{resp} and energy devoted growth \dot{E}_{growth}

Then dissect each term in <u>variable quantities</u>, <u>parametric quantities</u>, or some product of the two.

Variable quantities can take on any of many possible values.

Parametric quantities, by contrast, are fixed for the situation at hand. They hold "across measurements" (a literal translation of "parameter").

Example: metabolic rate (\dot{E}_{ingest} or \dot{E}_{resp} in the absence of growth) depends on body size (M = kg) according to the following relation:

$$\dot{E} = \alpha M^{\beta}$$

In this equation there are two variable quantities, metabolic rate and body mass M. There are two parameters β and α . Metabolic rate and mass M each represent a series of measured values, while β and α represent single valued quantities that hold across the values of energy and mass.

Equations (sentences) can be dissected into terms (phrases), which in turn are dissected into quantities, which are used to identify the meaning of each phrase (term).

$$\frac{1}{N}\frac{dN}{dt} = \dot{B} - \dot{D}$$
This equation has three terms: $N^{-1}dN/dt$, \dot{B} , \dot{D}

The quantities in the first t	time		Bacteria Hours	# T
The relation of the quantit				
in the first term:	time rate of change in population size	dN/dt	Bacteria/hr	#/T
The first term is	percapita rate of change in population size	N ⁻¹ dN/dt	% / T ⁻¹	T^{-1}
The remaining terms are:	instantaneous birth rate instantaneous death rate	B D	$\% / T^{-1}$ $\% / T^{-1}$	$T^{-1} \\ T^{-1}$

Another example to be dissected and reconstructed: \dot{E}

$$\dot{E} = \dot{E}_o \ Q_{10}^{(T2-T1)/10}$$

where \dot{E} is respiration and T2 and T1 are temperatures.

4.2 Reading Equations

<u>Translation</u> aids comprehension.

Translate to units.

Write units beneath each symbol--helps in visualizing meaning of symbol

$$N = N_o e^{(B - D) t}$$

 $\# = \# e^{(\%.day - \#.day) day}$

Translate to <u>dimensions</u>. This aids in visualizing the relation of quantities.

The idea is to identify which quantities are similar. In this case # and Time T

$$N = N_0 e^{(B - D) t}$$

 $\# = \# e^{(I/T - I/T) T}$

Writing dimensions beneath each symbol shows which symbols and terms represent similar quantities.

Translate into calculations.

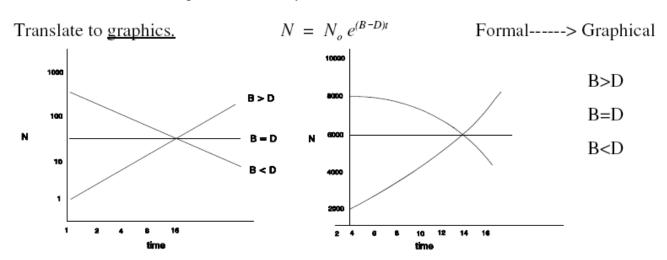
This aids in comprehending quantities change in relation to each other.

Here is an example:

$$N = N_o e^{(B-D)t}$$

instantaneous rate $\dot{B} - \dot{D}$	time t	N/N_o $e^{(\dot{B}-\dot{D})\cdot t}$	% change
0.1 year ⁻¹	1 year	1.11	11%
0.2 year^{-1}	1 year	1.22	22%
0.0 year^{-1}	1 year	1.00	0%
-0.2 year^{-1}	1 year	0.82	-18%
-0.2 year^{-1}	2 year	0.67	-33%

Box 4.2 Interpretation of the symbol $e^{(\dot{B}-\dot{D})\cdot t}$ via calculation.



4

To complete the process, state in words the idea behind the equation.

4.2 Reading Equations

Once the equation has been dissected, and the relation of the parts are understood, its meaning can be reconstructed.

Use words to connect the formal model to experience Formal---->Verbal

Here is a worked example showing Dissect --- > Translate --- > Reconstruct

Raup and Graus (1972) developed the following equation to describe allocation of carbonate to shell weight W_s by marine gastropods.

$$\frac{W_S}{V_i} = P_S \cdot \frac{S_S}{V_i^{2/3}} \cdot \frac{T_S}{V_i^{1/3}}$$

$$\begin{pmatrix} calcification \\ index \end{pmatrix} = \begin{pmatrix} shell \\ density \end{pmatrix} \begin{pmatrix} form \\ index \end{pmatrix} \begin{pmatrix} thickness \\ index \end{pmatrix}$$

Dissect into components

2 terms: calcification index = product of shell density P_s and two ratios

1 fixed quantity: shell density P_S A typical value for calcite is $P_s = 2.71$ g/cc

4 variable quantities W_S weight of shell mg

 S_S surface area of shell mm²

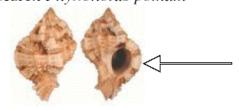
 T_S thickness of shell mm

 V_i shell internal volume mm³

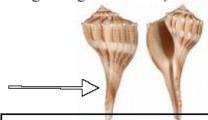
<u>Translate</u> to units and dimensions

The calcification index is the density of the shell material, adjusted by two ratios, the form index (shell area scaled to minimum area calculated from volume) and the thickness index (shell thickness scaled to minimum diameter calculated from volume).

Apple Murex Phyllonotus pomum



The apple murex wedges the outer part of the aperture between the two shells of the bivalve it has pulled open to eat. It has a thick wall (high thickness index) and a round shape (low form index). It invests calcium in a thick shell wall. Lightning whelk Busycon contrarium



The Lightning whelk wedges the elongated siphonal end between the two shells of the bivalve it has pulled open to eat. It has an elongated shape (high form index) and thin wall (low thickness index). It invests calcium in a siponal extension of the shell.

4.3 Homogeneity of Units

Equations in biology have units. They are not abstract expressions that hover mysteriously, like fog over a marsh. They are ideas about the relation of quantities, which have physically or biologically interpretable units.

Homogeneity of units means that all terms in an equation must have the same units. Cannot add (or subtract, or equate) apples and areas. This means that both sides of an equation must have the same units.

To check: Write units beneath each symbol.

Make sure all terms have same units.

Example 1.
$$N_t = N_o e^{r * t}$$

Ants = Ants $e^{(-)(days)}$

Write equation

Determine units of each symbol, then write beneath each symbol.

Exponents are dimensionless so, r * t has no units

Consequently r must have units of 1/days

Example 2. Ryder's empirical equation for fish catch from lakes

$$H = 2.094 \text{ MEI}^{0.4461}$$
 $H = \text{fish catch} = \text{lb acre}^{-1} \text{ yr}^{-1}$
 $MEI = \text{ppm ft}^{-1}$

Both sides of an equation must have the same units, but

In order for left side to equal right side, the parameter 2.094 must have units that convert $(ppm\ ft^{-1})^{0.4461}$ to $lb\ acre^{-1}\ yr^{-1}$.

$$H = \alpha MEI^{0.4461}$$

Solve for α :

$$\alpha = H \cdot MEI^{-1}$$

 $\alpha = H MEI^{-0.4461} = 2.094 \text{ lb acre}^{-1} \text{ yr}^{-1} \text{ ppm}^{-0.4461} \text{ ft}^{0.4461}$

This can be checked by writing units beneath variables and parameters, show that units cancel:

$$H = \alpha$$
 $MEI^{0.4461}$
1b acre⁻¹ yr⁻¹ = 2.094 lb acre⁻¹ yr⁻¹ ppm^{-0.4461} ft^{0.4461}) (ppm ft⁻¹)^{0.4461}

4.3 Homogeneity of Units

Extra: One application of principle of homogeneity of units is that it is a reliable guide to converting an equation from one set of units to another.

This one takes a while, ca 15 minutes

Example: Convert MEI equation to hectares and metres and kilograms

To do this, convert the parameter α from one set of units to another.

$$\alpha = 2.094 \text{ lb acre}^{-1} \text{ yr}^{-1} \text{ ppm}^{-0.4461} \text{ ft}^{0.4461}$$

Conversion factors are:

$$43560 \text{ ft}^2/\text{acre} \cdot 0.093 \text{ m}^2/\text{ft}^2 \cdot 1 \text{ ha}/100^2\text{m}^2 = 0.4051 \text{ ha}/\text{acre}$$

3.28 ft/m

Apply conversion factors:

$$2.094 \cdot lb \cdot \left(\frac{0.4536 \ kg}{lb}\right) \cdot acre^{-1} \cdot \left(\frac{acre}{0.4051 ha}\right) \cdot ppm^{-0.4461} \cdot ft^{0.4461} \left(\frac{m}{3.28 ft}\right)^{0.4461}$$

$$2.094\,\cdot\,0.4536\,\cdot\,\,0.4051^{-1}\,\cdot\,3.28^{-0.4461}\,=\,\,1.38\,\,\,kg\,\,ha^{-1}\,\,yr^{-1}\,ppm^{-0.4461}\,m^{0.4461}$$

Conversion is:

Fish catch = kg ha⁻¹ yr⁻¹ =
$$1.38 MEI_{metric}$$

Cannot use same symbol H for eatch rate H, because of the new units.

Suggested new symbol is time rate of change in population biomass:

$$[\dot{M}] = 1.38 \, MEI_{metric}$$

This is a new symbol for the annual (upward) flux of fish out of a lake, as captures. Like H it has units of mass per unit time through a known area.

4.4 Dimensional Homogeneity of Equations

Equations that express ideas about quantities must be homogeneous with respect to dimensions. This is sometimes quicker than checking to make sure equation is uniform in regard to units.

Write dimension above each symbol in equation. Make sure dimensions of all terms are the same.

Table 4.1

Rules for working with dimensions. Adapted from Riggs (1963).

- 1. All terms in an equation must have the same dimensions. Terms are separated by + =
- 2. Multiplication and division must be consistent with Rule 1.
- 3. Dimensions are independent of magnitude.

 The velocity dx/dt is the ratio of infinitesimals but still has dimensions of Length · Time⁻¹.
- Pure numbers (e,π) have no dimensions.
 Exponents have no dimensions
 % have no dimensions
- 5. Multiplication by a dimensionless number does not change dimensions.

$$N_t = N_o e^{r*t}$$

Ants = Ants $e^{(1/\text{days})(\text{days})}$
= # $e^{(1/T)(T)}$

Write equation

Determine units of each symbol, then write beneath each symbol.

Determine dimensions of each symbol and write beneath each symbol.

Write dimensions of each term.

Example 2.

$$H = \alpha \qquad MEI^{0.4461}$$

$$lb acre^{-1} yr^{-1} = (blank) \quad (ppm ft^{-1})^{0.4461}$$

$$M A^{-1} T^{-1} = (\underline{\hspace{0.5cm}}) \quad (\% L^{-1})^{0.4461}$$

$$M L^{-2} T^{-1} = (\underline{\hspace{0.5cm}}) \quad (\% L^{-1})^{0.4461}$$

$$(\underline{\hspace{0.5cm}}) = M L^{-2} T^{-1} ((\% L^{-1})^{0.4461})^{-1}$$

$$(\underline{\hspace{0.5cm}}) = M L^{-2} T^{-1} (L^{0.4461})$$

$$(\underline{\hspace{0.5cm}}) = M L^{-1.5539} T^{-1}$$

4.4 Dimensional Homogeneity

Application of these rules from Riggs (1963) have a number uses.

One use is to make sure that equations are correct. If an equation is not dimensionally homogeneous, it is of no use in making computations.

Another use is to work out the dimensions of unknown quantities and parameters in equations. This systematic use of dimensions and the underlying concept of similarity aids in understanding the physical or biological meaning of equations and symbols. The systematic use of the concept of similarity is an important skill in quantitative work.

Application -- checking equations

The first example, from Riggs (1963, p 38) shows the use of dimensions and the concept of similarity to check whether an equation is correct.

Wolf (1950) ascribes the following equation to Adolph (1943).

$$t = 1.3 + L_{H2O}$$

where t = time in hours needed to excrete a water load of L_{H2O} expressed as a percent of body weight.

Is this correct?

The best way to find out is to write out dimensions for the equation.

Write equation.

Determine dimensions of each symbol and write beneath each symbol.

Write dimensions of each term.

Dimensional Homogeneity

Here is another example of the use of dimensional homogeneity to check the validity of equations.

According to Wickelgren (1974) as quoted by Mayes (1983) memory retention over a period ranging form seconds to years satisfies the equation:

$$m = L t^{-D} e^{-I \cdot t}$$

where

m is the memory trace's strength (measured by recognition)
L is initial strength at the end of learning
t is the retention interval
D is the time decay rate
I is the measure of degree of interference

Write equation on board.

Attempt to determine units of each symbol (make guesses).

Attempt to write dimensions beneath symbols and terms in equation.

This looks like science. It isn't. The term t^{-D} has no meaning if D is a rate, because time cannot be raised to a power of a rate. This makes no sense.

The variables m and L are said to be both strengths. What are the units of strength? The variables m and L cannot have the same units because t has units. The product on the right side of the equation $(L t^{-D})$ cannot have the same units as the variable m on the right $(e^{-I \cdot t})$ has no units. As impressive as the equation looks, it is meaningless. It cannot be used to make calculations. It is bogus.

4.4 Dimensional Homogeneity Application: Reading Equations.

Dimensions and the concept of similarity are useful in interpreting the meaning or interpreting the physical or biological content of a symbol. For example, the Coriolis parameter *f* occurs often in descriptions of atmospheric and ocean circulation. The physical content of the symbol becomes clearer from its dimensional symbol, which is T⁻¹ The abstract symbol *f* has the dimensions of frequency. It is the frequency or angular velocity that applies to an air or water parcel once it has been set in motion on the revolving surface of the earth. One way of deciphering a symbolic expression is to write out the dimensional symbols immediately above or below each symbol in the expression. Often this will make the idea behind the expression more comprehensible, by relating it to familiar ideas of mass, time, distance, and energy. My copy of Circulation in the Coastal Ocean (Csanady 1982) had 2 blank pages at the end that are now filled with a list of symbols for physical quantities, their names, their dimensions, and the page number where they were first defined in the text. This helped me considerably in working through the ideas used to calculate flows affecting life in coastal waters.

Here is an example, using the equation of femur length (Lewontin et al. 1960).

Femur length =
$$b$$
 (Body Length) $^{\alpha}$

The term on the left side of this equation (femur length) has dimensions of length. The term on the right side of the equation ($b \cdot (Body Length)^{\alpha}$) must also have dimensions of length. From this it is possible to work out the dimensions of the parameter b

Femur length =
$$b$$
 (Body Length) $^{\alpha}$

$$fL = 0.153 \quad L^{1.066}$$

$$mm = ? \quad mm^{1.066} \quad ? = mm^{1} \cdot m^{-1.066}$$

$$mm = mm^{-0.066} \quad mm^{1.066} \quad = mm^{-0.066} = mm^{1-\alpha}$$

$$L = L^{-0.066} \quad L^{1.066}$$

The parameter b has units of mm^{1- α} or mm^{-0.066}. It can be viewed as a shape factor that converts body length into femur length. Thw exponent -0.066 measures shrinkage of femur relative to body length as body length increases from smaller to larger animals.

4.4 Dimensional Homogeneity -- Application: Reading equations.

Here is a second example, this time from population biology. Many people will have already encountered this equation for predation.

$$\frac{1}{N} \frac{dN}{dt} = r - \beta P$$

where

N = prey numbers P = predator numbers

What do the parameters r and β signify?

Start with vivid image of

N (all the moose on Isle Royale in Lake Superior)

P (all the wolves on Isle Royale)

The term on the left is the instantaneous rate of change in moose.

It has units of % change per unit time = (#/#) T⁻¹

A typical value might be 0.1 yr⁻¹either up or down

This change comes about through production of moose at rate r, and loss of moose to wolves at rate βP

$$r = \%$$
 increase per unit time = $(\#/\#)$ T⁻¹

$$\beta P = \%$$
 loss per unit time = $(\#/\#)$ T⁻¹

These two terms add up to the total rate (on the left) so they must have the same units and dimensions as the term on the left, the instantaneous rate of change in moose.

With knowledge of the units of βP and of P, the units of β can be worked out.

$$\beta = \% \#^{-1} T^{-1} = \% loss per (# \cdot T)$$

 β is interpreted as percent of prey captured per predator hour (# · T = effort)

 β is a measure of the efficiency of the predator

$$\beta = (\#prey / \#pred) per (\#prey \cdot T)$$

With units, dimensions, and the concept of similarity, we can work out (visualize) the idea expressed by this equation.

4.4 Dimensional Homogeneity - Application: Reading equations

Another example. Ivlev's (1961) equation for predation

The terms in an equation can be complex, composed of several quantities combined via multiplication and division. The product of all of the quantities in one term must nevertheless work out to be the same as the product of all of the quantities in another term. For example, Ivlev's equation for rate of prey ingestion (Ivlev 1961) is

$$I = I_{max} \left(1 - e^{-\zeta(p - p')} \right)$$

I = ingestion, prey/hour

 I_{max} = maximum ingestion, prey/hour

p = prey concentration, count/ml

p' = threshold prey concentration, count/ml

 ζ = ? (undefined)

This parameter zeta ζ is not a meaningless symbol. It stands for some aspect of the interaction of predators with prey. What does it stand for ?

Use same procedure as above, using the concept of dimensional similarity to work out meaning of the equation, and the meaning of the parameter ζ .

I and I_{max} have same units, as expected.

The maximum ingestion I_{max} is reduced by increasing prey density according to the expression inside the parentheses. The prey begin to swamp the predator, and ingestion drops at high prey density. The drop is a percentage, calculated according to the expression inside parentheses. The percent drop is a dimensionless ratio.

13

The exponent must be dimensionless. So $\zeta(p-p')$ must also be dimensionless.

p and p' have the same units, so ζ must have units of 1/p = ml/count

What does ζ , with units of ml/count, signify? It can be interpreted as the area searched per unit of prey.

Dimensional Homogeneity of Equations -- Summary

Dimensional reasoning, based on the concept of similarity, is a useful quantitative skill. It enables one, with practice, to work out the meaning of equations. It is the key to "reading" equations.

It is important in physiology, where the concept of similarity is routinely used concerning questions of form and function in relation to body size.

It is a useful skill in the environmental sciences. This is particularly true were biological reasoning must be integrated with physical reasoning (which relies on dimensions). The principle of dimensional similarity is routinely used in reasoning about physiological processes that connect organisms to their environment. A method used by physiologists working at the space and time scales from the cell to the individual, and used also by physical scientists working at time and space scales important to natural populations, should be of considerable use to ecologists who work in between, where physiological performance (growth rate, birth rate) interacts with the dynamics of the physical environment. The method of dimensional grouping is an important avenue for interdisciplinary understanding of biological processes.

The principle of dimensional similarity is a useful tool in any field of biology where computational methods are used. It is an effective way of reasoning about quantities used in computations.

Unfortunately, dimensional reasoning based on principle of similarity is not a regular part of the curriculum of biologists at most Universities.

References

- Adolph E.F. 1943. *Physiological Regulation*. Jacques Cattell Press, Lancaster Pa USA
- Mayes, A. (ed.) 1983. Memory in Animals and Humans. Van Nostrand Reinhold.
- Raup, D.M., Graus, R.R. 1972. General equations for volume and surface area of a logarithmically coiled shell. *Mathematical Geology* 4: 307-316.
- Riggs, D.S. 1963. *The Mathematical Approach to Physiological Problems*. Baltimore: Williams and Wilkins. Reprinted by M.I.T. Press, Cambridge, Massachusetts USA (1977)
- Wickelgren, W.A. 1974. Single trace fragility theory of memory dynamics. *Memory and Cognition* 2: 775-780.
- Wolf A.V. 1950. The Urinary Function of the Kidney. Grune and Stratton NY USA

4.5 Review Questions

1. According to Holligan <i>et al.</i> 1984 (<i>Marine Ecology Progress Series</i> 17:201) the vertical flux of nutrients through the ocean's thermocline is: $F_N = K_V \Delta N / \Delta Z$
were F_N is the vertical flux of nutrients (milligram-atoms m ⁻² s ⁻¹) K_V is the vertical eddy diffusivity (10 ⁻⁴ m ² s ⁻¹) ΔN is the nitrate difference across the thermocline (mg-atoms) ΔZ is the thickness of the thermocline (metres)
Write out dimensions beneath each symbol in the equation. Is this equation dimensionally homogeneous?
Work out the dimensions of ΔN required to make the equation homogeneous
Work out the units of ΔN required to make the equation homogeneous $M = Mass$ $M L^{-1} = mass$ gradient $M L^{-2} = mass$ density $M L^{-3} = mass$ concentration
Based on this, ΔN must be the difference in nitrate across the thermocline.
2. A series of experimental measurements by Holligan <i>et al</i> suggest that the vertical flux of nutrients through the thermocline follows an exponential relation:
$F_N = \alpha (K_V \Delta N / \Delta Z)^{3/4}$
What units does α have? What dimensions does α have?
3. Another series of experiments by Holligan <i>et al</i> suggest that nutrient flux depends upon the temperature gradient across the thermocline.
$F_N = \beta (\Delta T/\Delta Z)^{-1/3}$
$\Delta T/\Delta Z = {}^{\circ}C/\text{metre}$
What units does β have? What dimensions does β have?