ECO 602
Analysis of Environmental Data

FALL 2019 – UNIVERSITY OF MASSACHUSETTS

DR. MICHAEL NELSON

Deterministic Functions: Part 2

Announcement: no peer-assessment on group projects.

Today's Agenda

1. Function terminology recap
2. Families of functions
3. Quiz/activity
4. Categorical variables
5. Fitting model to data
6. Chapter 4 highlights

Key terms for deterministic functions

1. Monotonic
2. Continuous
3. Smooth
4. Asymptotic
5. Linear/nonlinear
6. Parameters/variables

Intuition from graphs and equations

Distinguishing variables from parameters

Long-term behavior

1. Increase/decrease without bound
2. Asymptotic?
3. Periodic?
4. Chaotic?

What happens at zero?

this is **super important** when considering probability distributions

Linear function intuition

1. Intercept moves the line up/down
2. Slope defines magnitude of relationship
3. Rate of change is constant

Non-linear function intuition

1. Intercept moves the curve up/down
2. Slope of tangent line is magnitude of association at a given value of the predictor
3. Rate of change is not constant

Families of functions

1. Rational (includes linear)
2. Exponential
3. Piecewise
4. Combination/hybrid functions

Rational functions

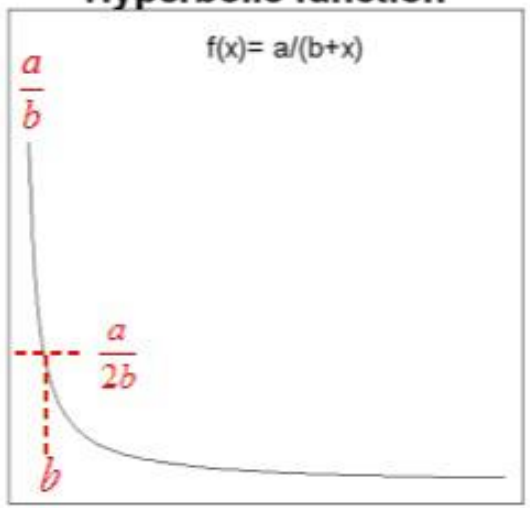
1. The **variable** is the **base**, power is **parameter**
2. Polynomial functions have integer powers
3. Rational functions have fractional (decimal) powers. Can lead to **complex numbers**.

Deterministic Functions... bestiary of functions

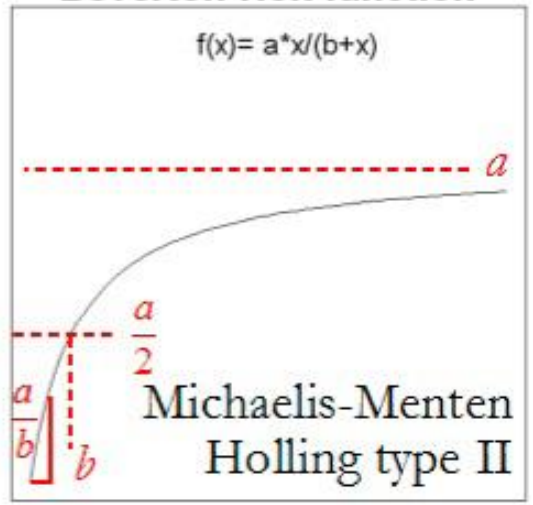
Rational functions (polynomials in fractions):

- Flexible, with finite limits (asymptotes)
- Often have mechanistic interpretation arising from simple models of biological processes such as *competition* and *predation*
- Can be complicated to analyze (e.g., difficult to estimate asymptotes)

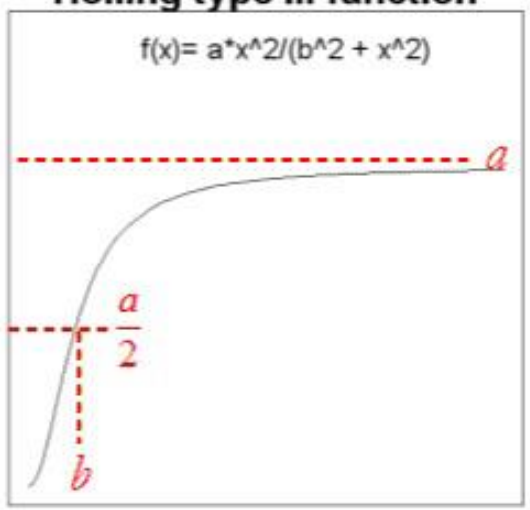
Hyperbolic function



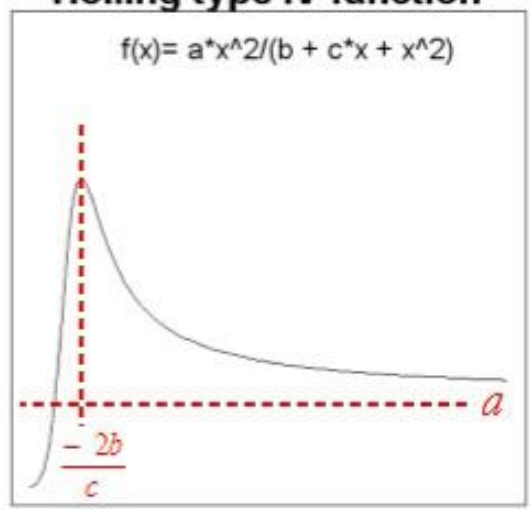
Beverton-Holt function



Holling type III function



Holling type IV function



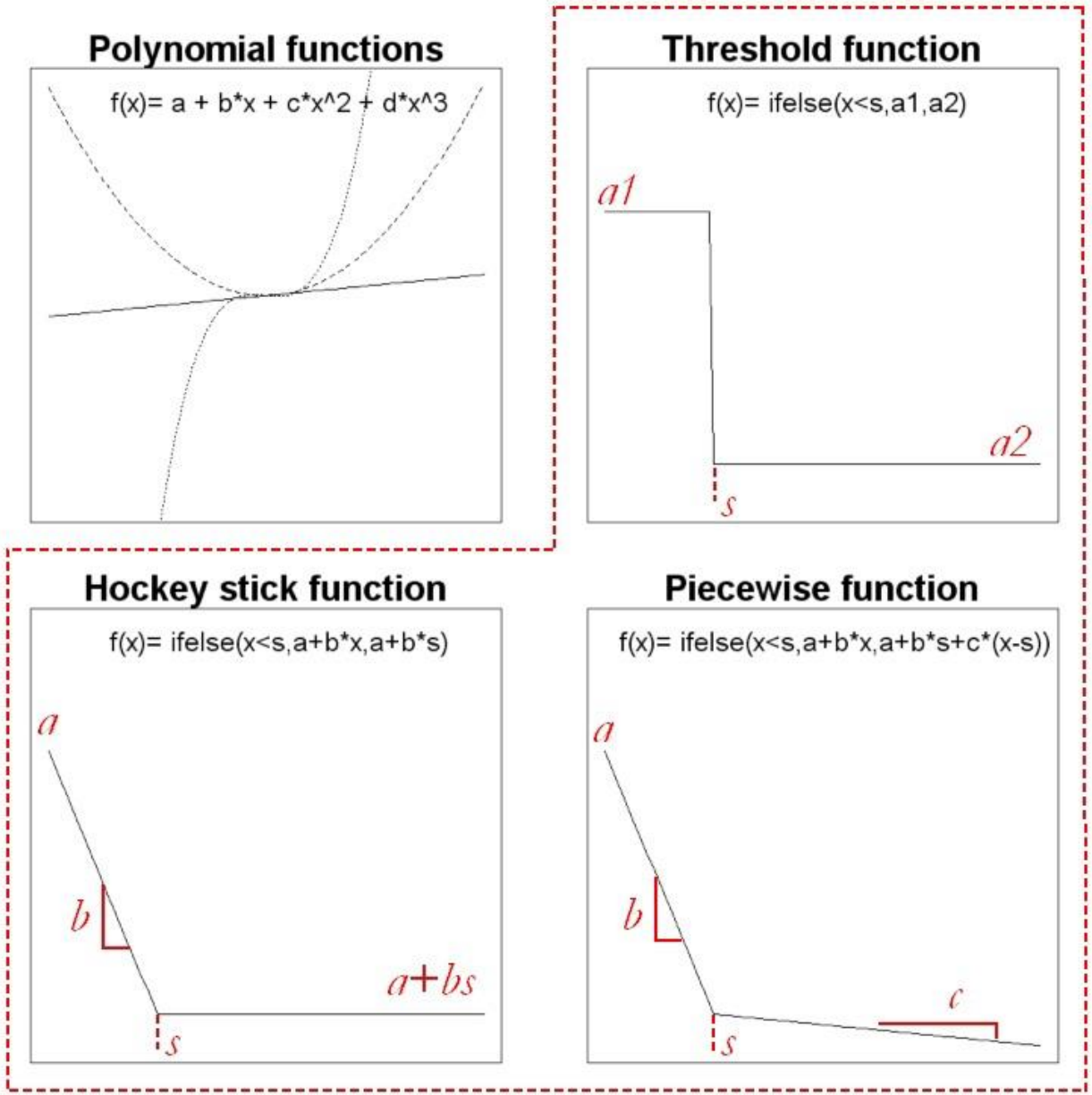
Polynomial functions

1. Special cases of rational functions with positive integer powers.
2. Degree: the highest power in the function.
3. Linear functions are a special case of polynomial functions with degree = 1.
4. Even and odd functions

Deterministic Functions... bestiary of functions

Piecewise polnomial functions:

- Flexible for threshold-like patterns and for setting function limits
- Easy to understand
- Probably unrealistic in most cases, since abrupt thresholds are unlikely in ecological systems



Exponential

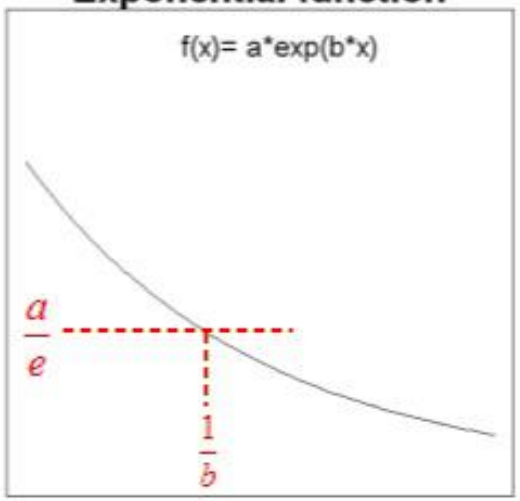
1. The **variable** is the **exponent**, the **base** is **parameter.**
2. Most common base is **e**
3. Straightforward formula for change of base
4. Exponential functions eventually grow faster than any rational function.

Deterministic Functions... bestiary of functions

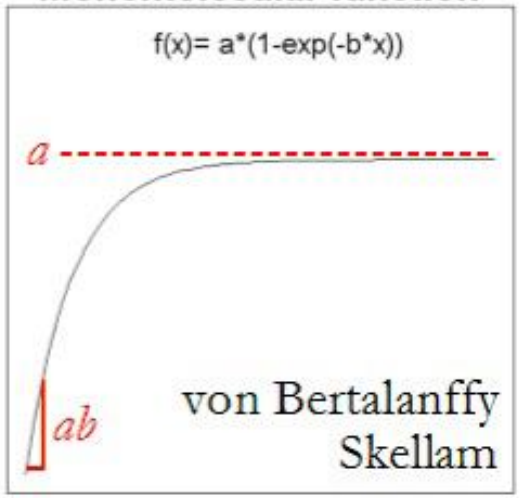
Functions based on exponential functions:

- Familar and popular functions (logistic)
- Flexible, with finite limits (asymptotes)
- Often have mechanistic interpretation arising from simple models of biological processes such as *population growth*

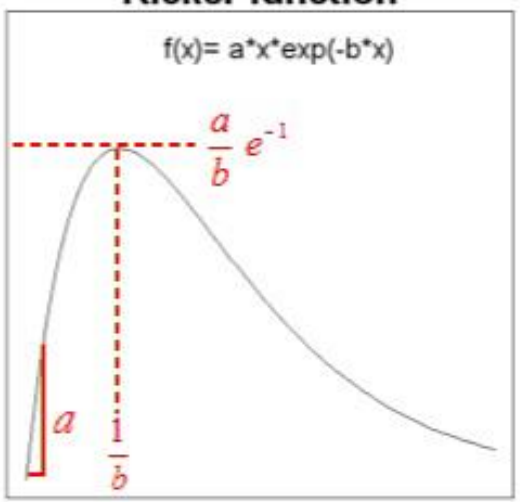
Exponential function



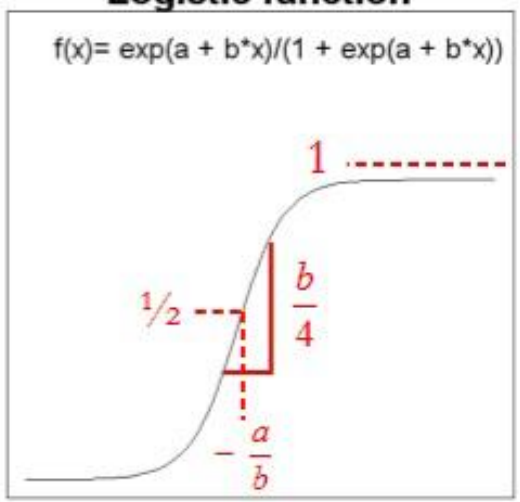
Monomolecular function



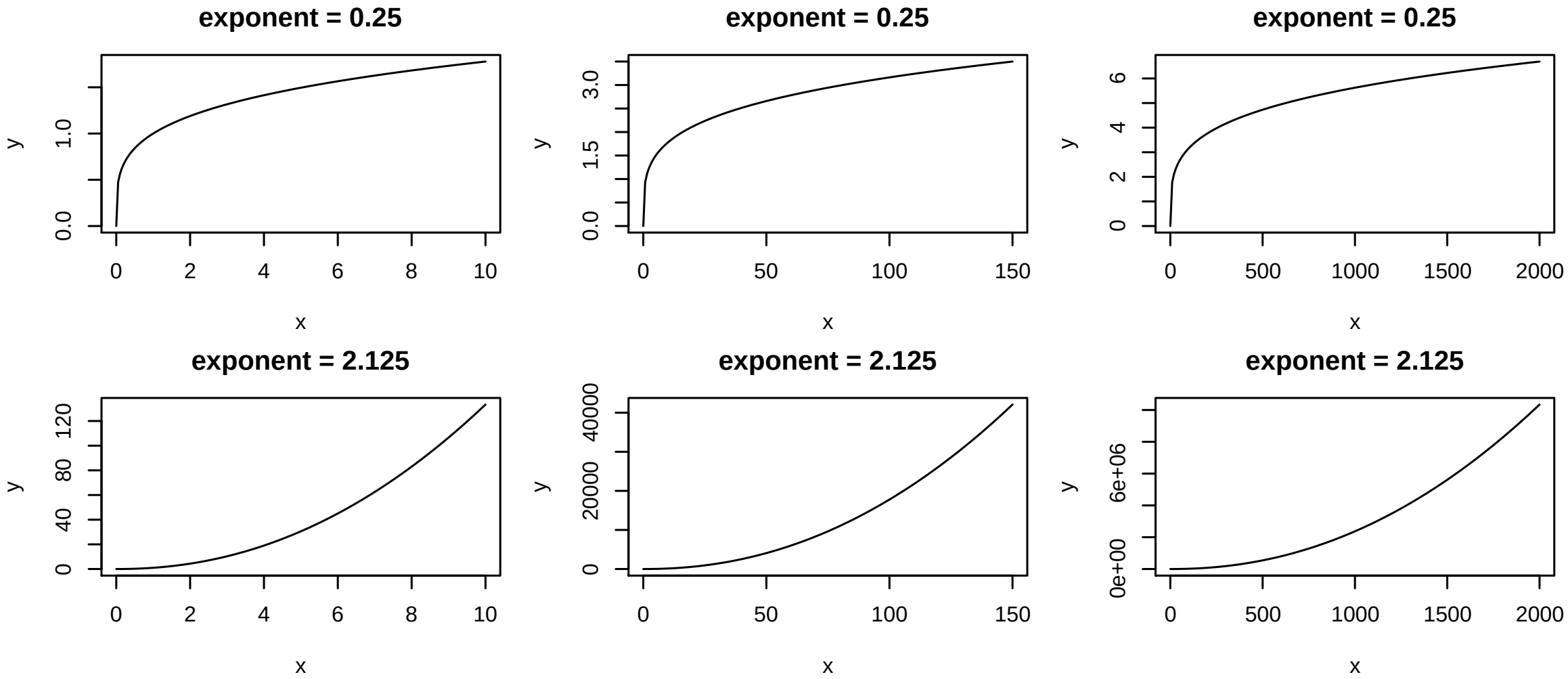
Ricker function



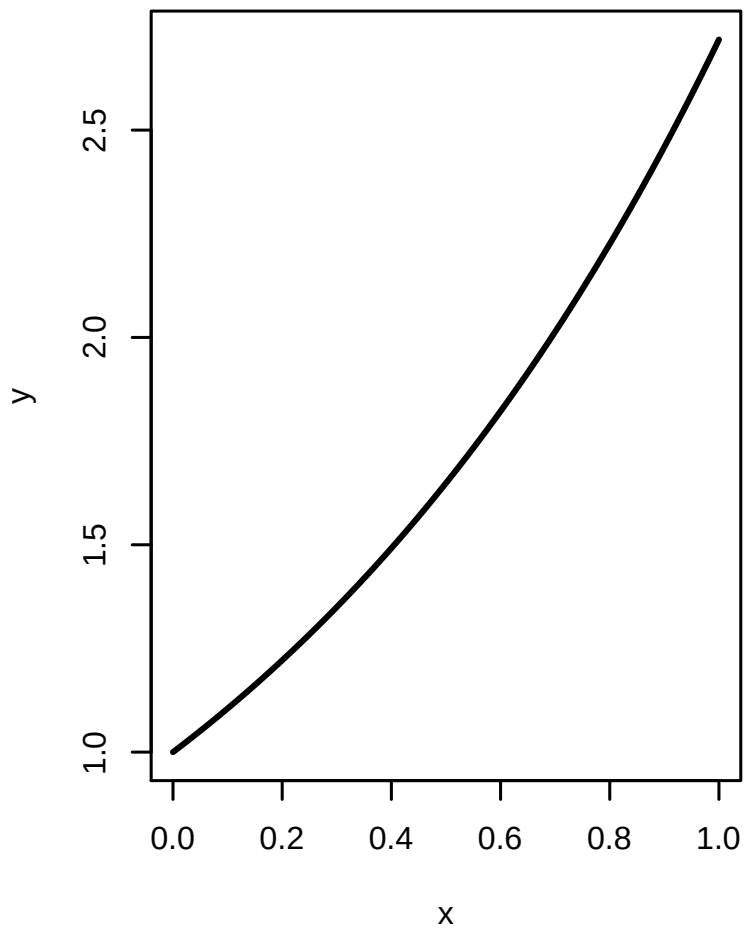
Logistic function

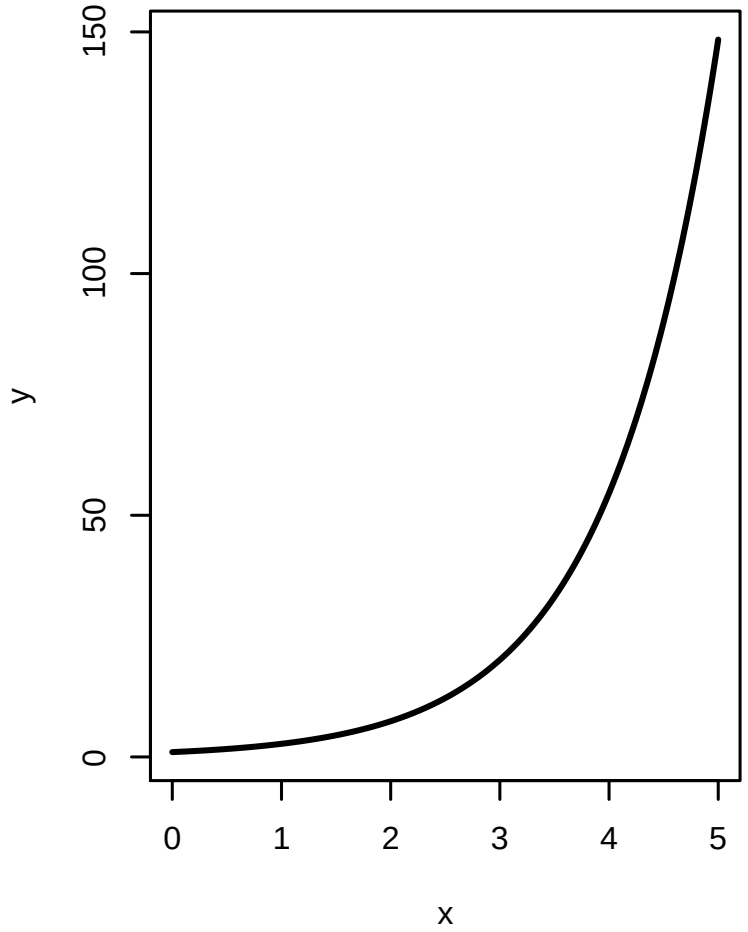


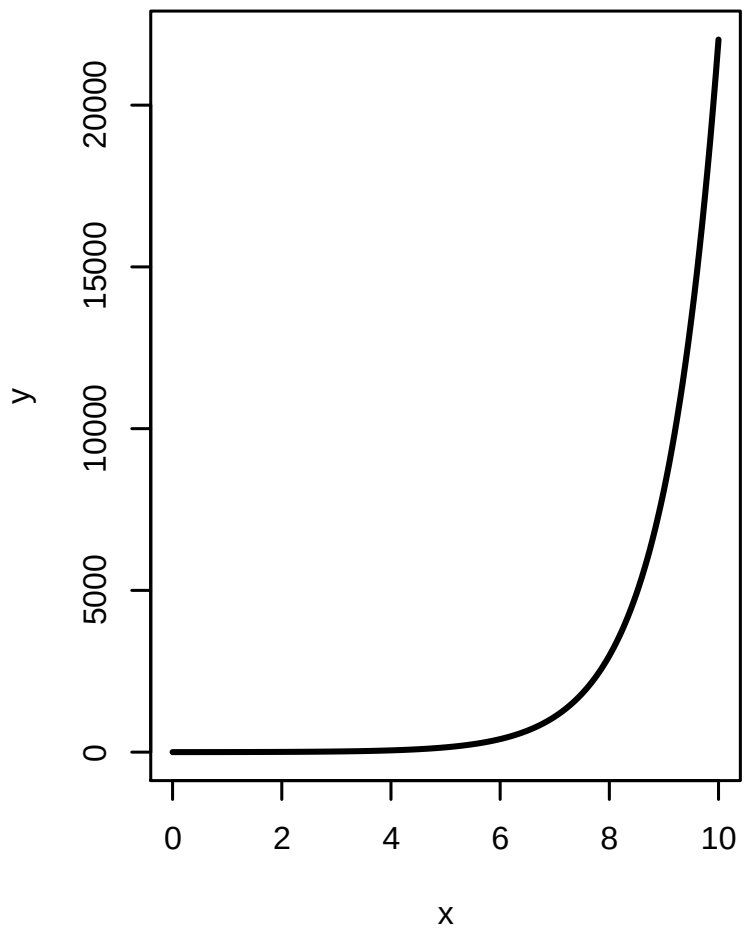
Power-law functions are scale invariant



Exponential function is not scale-invariant







Combination functions

1. Combine any of the previous function types
2. Model terms
    1. Each term might have a mechanistic interpretation (or it might not)
    2. Terms can be combined by multiplication, addition, exponentiation, etc.

Logistic functions

1. Ratio of two exponential functions
2. Characteristic sigmoid shape
3. Has upper and lower asymptotes
4. Useful for logistic regression: binary categorical variables

Functions: graphical intuition

1. Tangent lines, derivatives
2. What does a slope of zero mean?
3. When is the slope equal to zero?

Deterministic functions: local linearity and long-term behavior

1. Asymptotic
2. Increase/decrease without bound
3. Periodic
4. Bounded but chaotic
5. We usually only care about a restricted portion of the domain!

Choosing a deterministic function

Always start with a linear model.

It may not be realistic, but it is a good starting point.

Consider several alternative models:

1. Multiple predictors
2. Nonlinear terms

Deterministic function activity/quiz

Hypothesize an association in any system.

1. Briefly describe your system/association

Propose 2 different mechanistic models.

For each mechanistic model, create

1. Verbal description of association
2. Sketch of the predictor/response curve
3. List of candidate model terms

Choosing a deterministic function

Phenomenological

1. Try different functions until you find a best* fit

Mechanistic

1. You can hypothesize the type of function from knowledge of the system.

* There are many possible criteria for model goodness-of-fit

Choosing a deterministic function

How do we evaluate a model fit?

Usually we define a type of error then choose a function that minimizes it.

Most common is the sum of squared errors: **SSE**

There are other possibilities … such as sum of absolute error.

Choosing a deterministic function

Tradeoff between model complexity and generality.

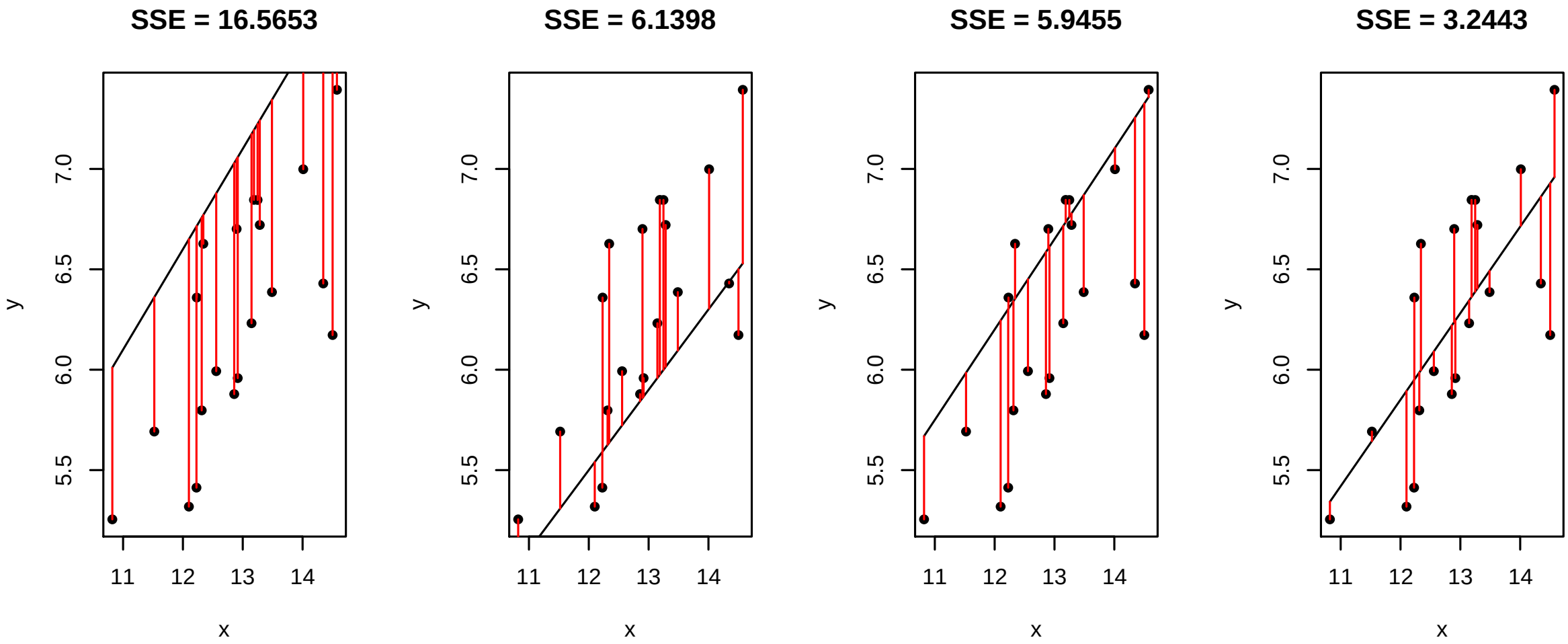
Overfitting

- A model is overfit if it describes a particular set of data very well but is a poor fit to additional samples.
- We'll consider model comparison later.

Residuals and errors

1. Difference between predicted and observed values.
2. Ordinary Least Squares OLS regression minimizes the sum of squared residuals

Linear model fit example: guessing



Linear model fit example: OLS

1. Ordinary Least Squares
2. Minimizes the sum of squared errors
3. Uses linear algebra tools to find exact parameter values.

Categorical variables: dummy variables

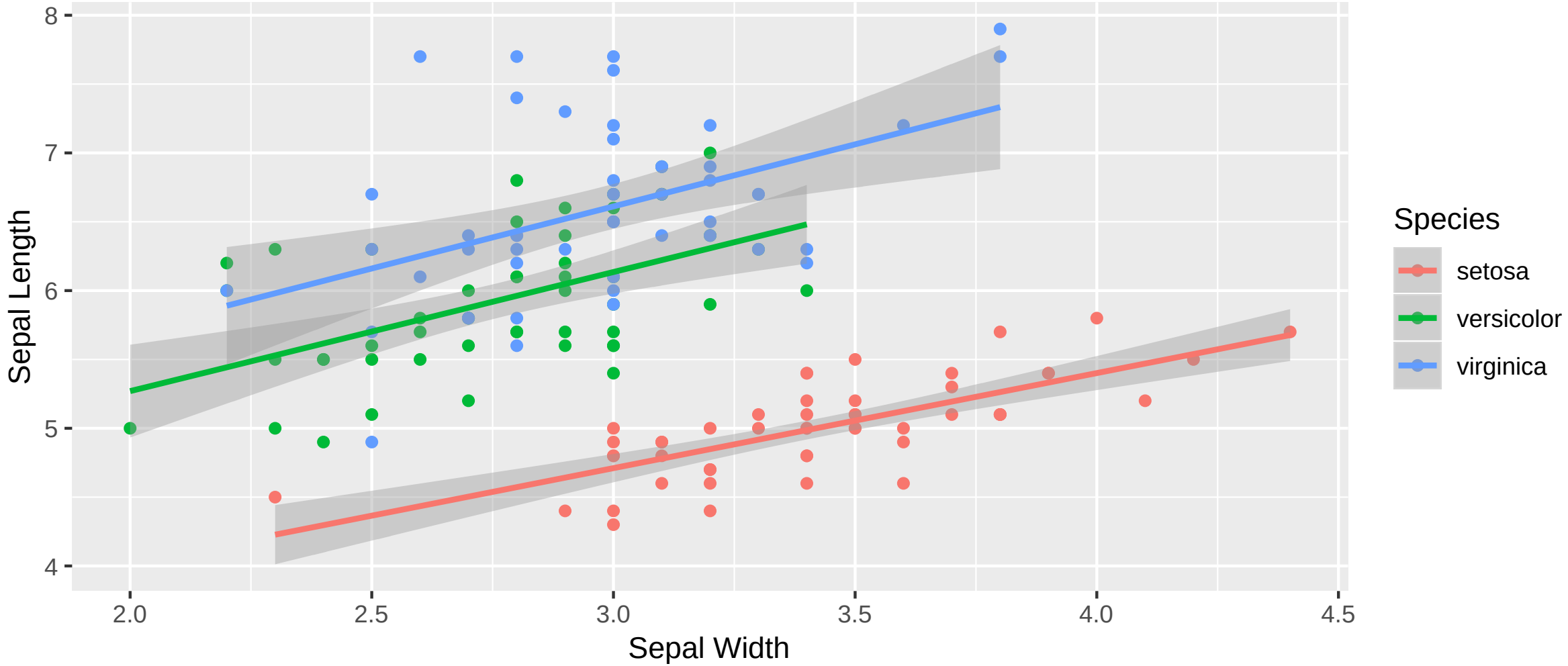
Dummy variables: extra columns added to table of data.

Dummy variable for each level, after the first, of the categorical variable.

Category: iris species

	(Intercept)	Sepal.Width	Speciesversicolor	Speciesvirginica
1	1	3.5	0	0
2	1	3	0	0
3	1	3.2	0	0
50	1	3.3	0	0
51	1	3.2	1	0
52	1	3.2	1	0
53	1	3.1	1	0
147	1	2.5	0	1
148	1	3	0	1
149	1	3.4	0	1
150	1	3	0	1

Category: iris species



Dual model paradigm

1. Deterministic model
    1. Pattern
    2. Prediction
    3. Mean, central tendency
2. Stochastic model
    1. Noise
    2. Uncertainty
    3. Variability, spread

For next time:

McGarigal Chapter 5a : Probability Distributions

Read sections 1 - 2

This chapter is dense, try not to get caught up on the details.

Key ideas to review: Probability mass function, cumulative mass function