FITTING PARAMETERS OF ANALYTIC SOLUTIONS

EXTRA

What is the difference between growth models of autocatalytic change processes and:

Repeated Measures Analysis

Multiple Regression

Latent Growth Curve Analysis

The Multilevel Model for Change

??????

The Multilevel Model for Change can model ‘individual’ growth and nonlinear growth, assuming that:

- ‘True’ growth / change trajectories: X=T+E

- Deviations to the true trajectory are treated as (random) noise, measurement error: E

- Time is treated as a predictor: Y = f(X), NOT an *autocatalytic process*

- Parameters (f.i *r*) = random variable

However:

Model parameters are estimated from data…

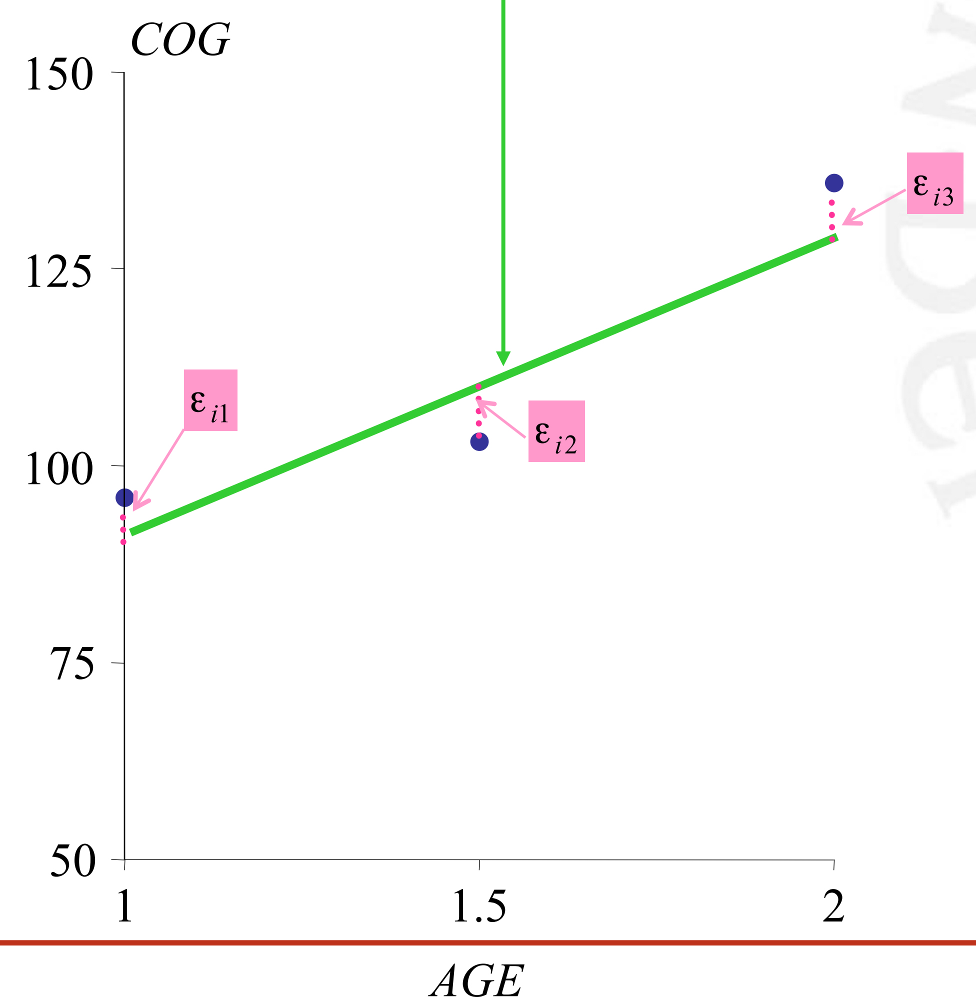
Structural portion, *which embodies our hypothesis about the shape of each person's true trajectory of change over time*

Stochastic portion, *which allows for the effects of random error from the measurement of person i on occasion j. Usually assume* $\varepsilon_{ij} \sim N(0, \sigma_\varepsilon^2)$

$$COG_{ij} = \left[\pi_{0i} + \pi_{1i}(AGE_{ij} - 1)\right] + \left[\varepsilon_{ij}\right]$$

- ***i indexes persons (i=1 to 103)***
- ***j indexes occasions (j=1 to 3)***

Individual i's hypothesized ***true change trajectory***



Structural portion, *which embodies our hypothesis about the shape of each person's true trajectory of change over time*

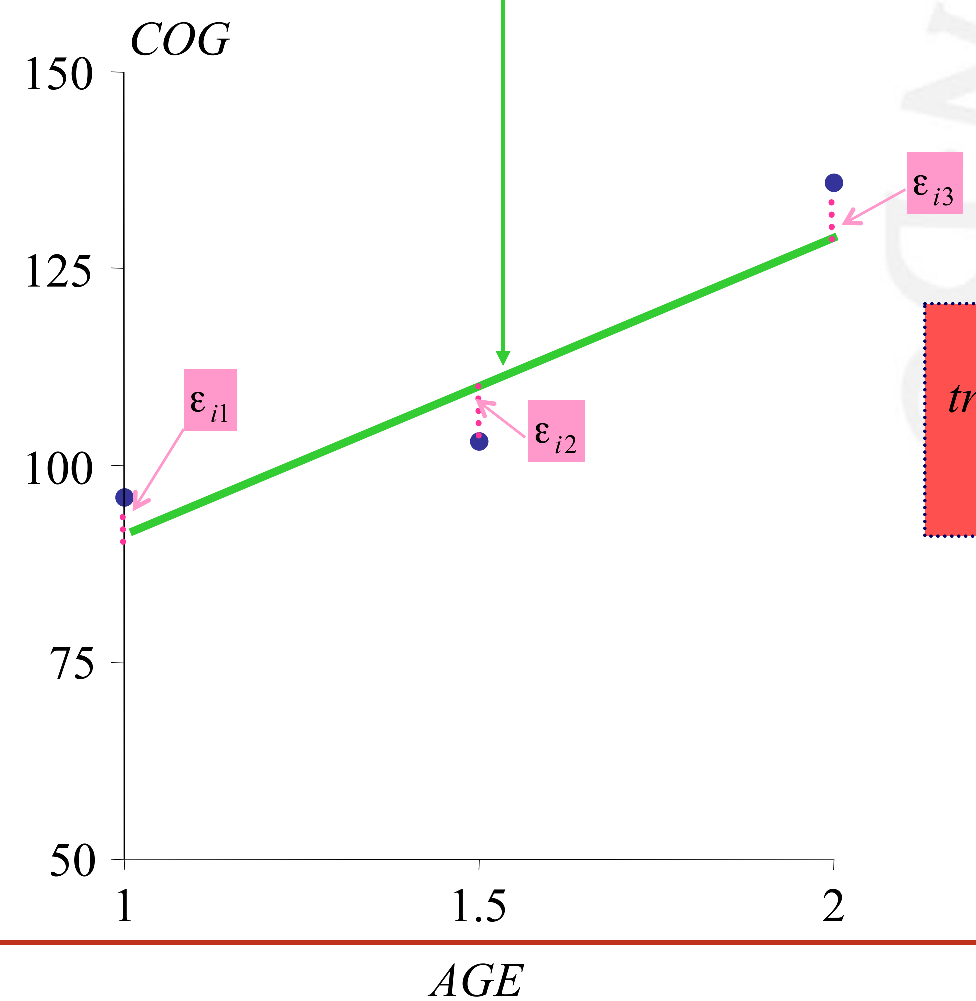
Stochastic portion, *which allows for the effects of random error from the measurement of person i on occasion j. Usually assume* $\varepsilon_{ij} \sim N(0, \sigma_\varepsilon^2)$

Key assumption: In the population, COG_{ij} *is a linear function of child i's AGE on occasion j*

$$COG_{ij} = \left[\pi_{0i} + \pi_{1i}(AGE_{ij} - 1)\right] + \left[\varepsilon_{ij}\right]$$

- ***i indexes persons (i=1 to 103)***
- ***j indexes occasions (j=1 to 3)***

*Individual i's hypothesized **true change trajectory***



π_{0i} *is the* ***intercept*** *of i's true change trajectory, his true value of COG at AGE=1, his* ***"true initial status"***

π_{1i} *is the* ***slope*** *of i's true change trajectory, his yearly rate of change in true COG, his true "annual* ***rate of change****"*

(ALDA, Section 3.2, pp. 49-51)

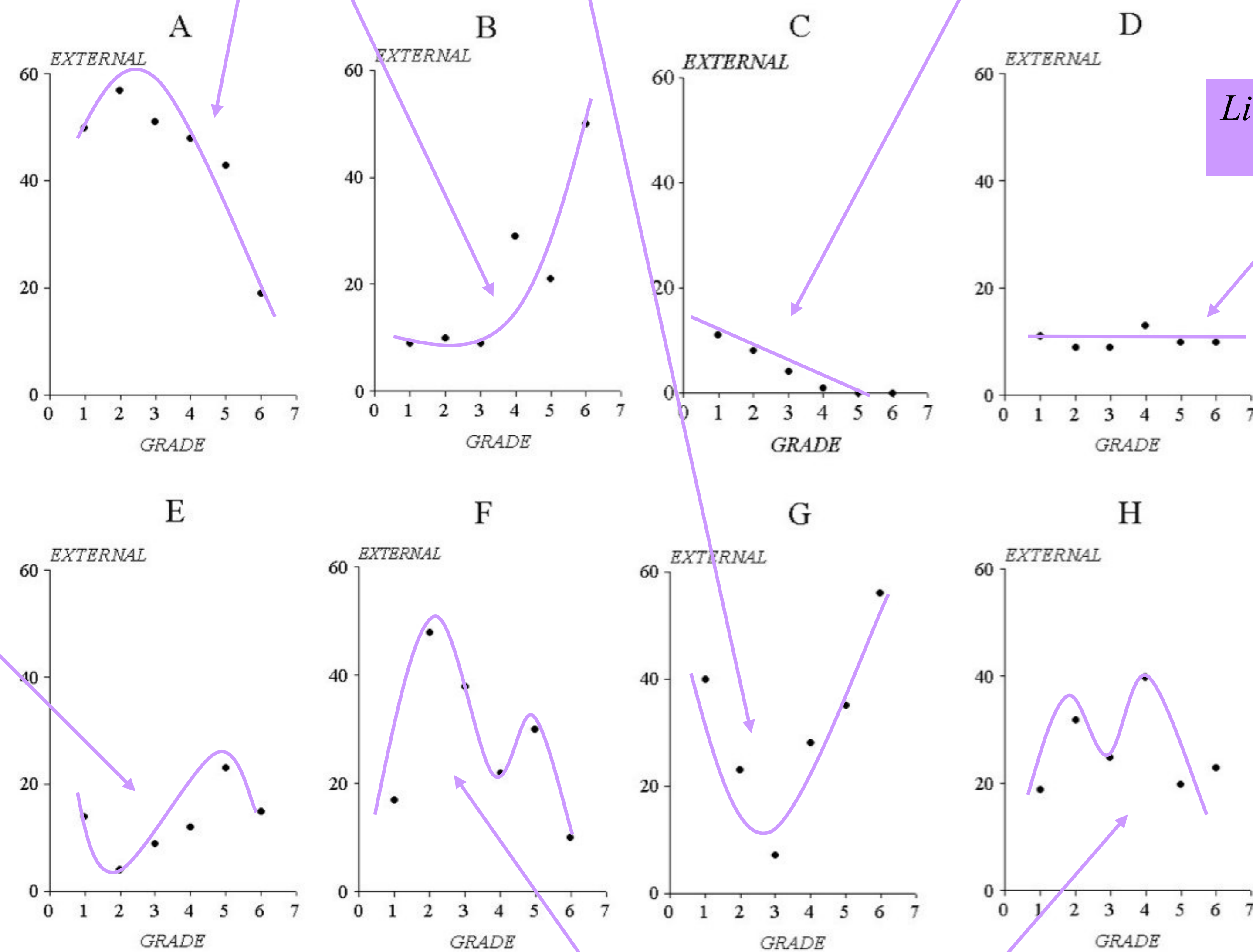
For a line:
2 parameters
(slope + intercept)

Other shapes?

Quadratic change (but with varying curvatures)

Linear decline (at least until 4th grade)

Little change over time (flat line?)



Two stationary points? (suggests a cubic)

Three stationary points? (suggests a quartic!!!)

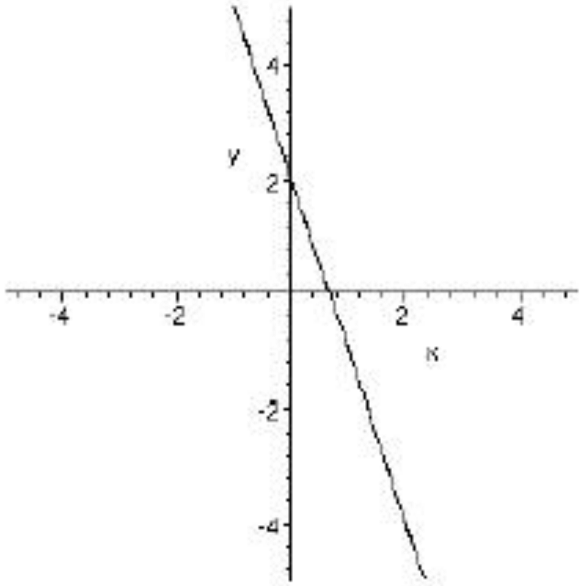
Assumptions…

the true change trajectory is a **polynomial function** of time of unknown order in the population

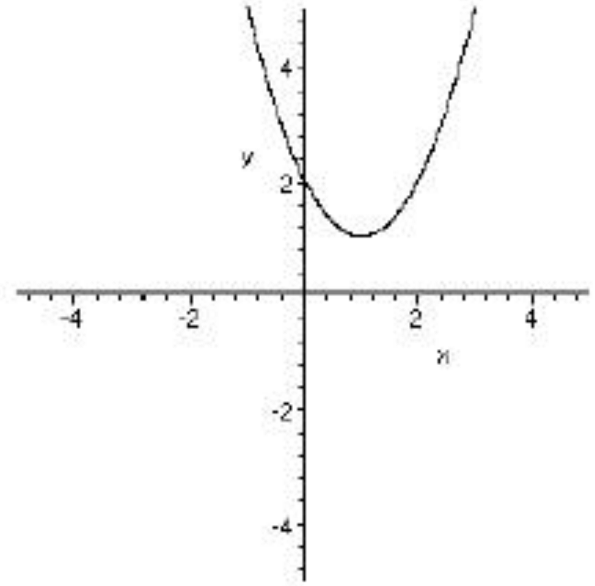
When faced with so many different patterns, how do you select a <u>*common*</u> *polynomial for analysis?*

http://gseacademic.harvard.edu/~alda/Chapter%20presentations.htm

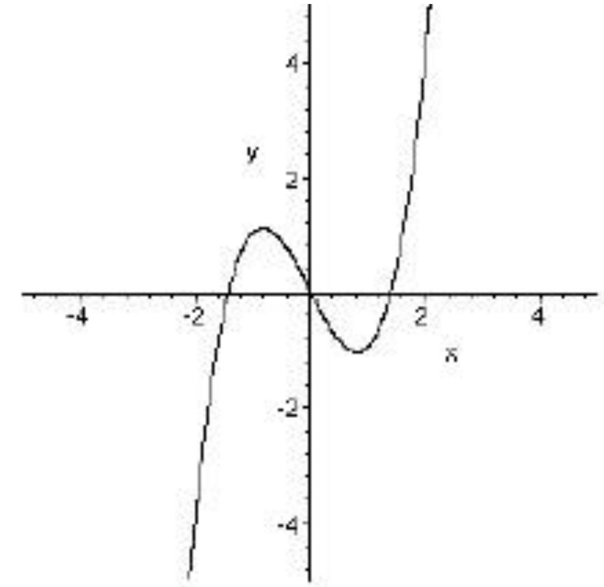
(ALDA, Section 6.3.2, pp 217-220)



$f(x) = -3x + 2$



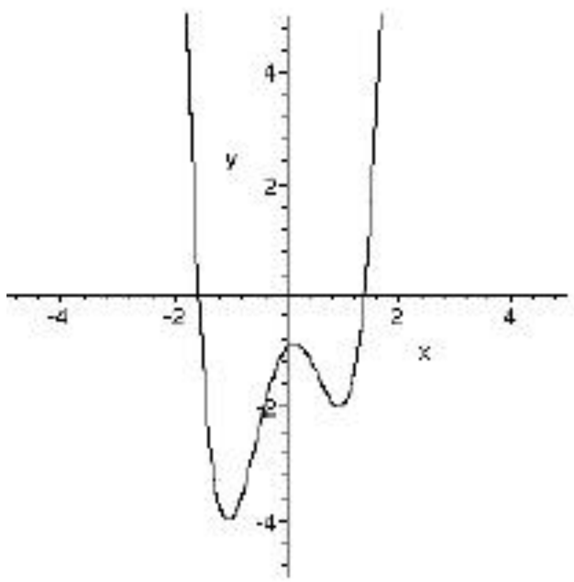
$g(x) = x^2 - 2x + 2$



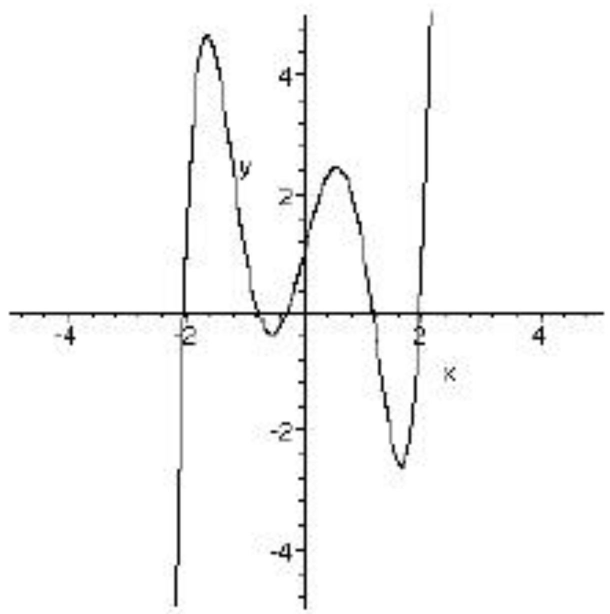
$h(x) = x^3 - 2x$

Polynomial functions:

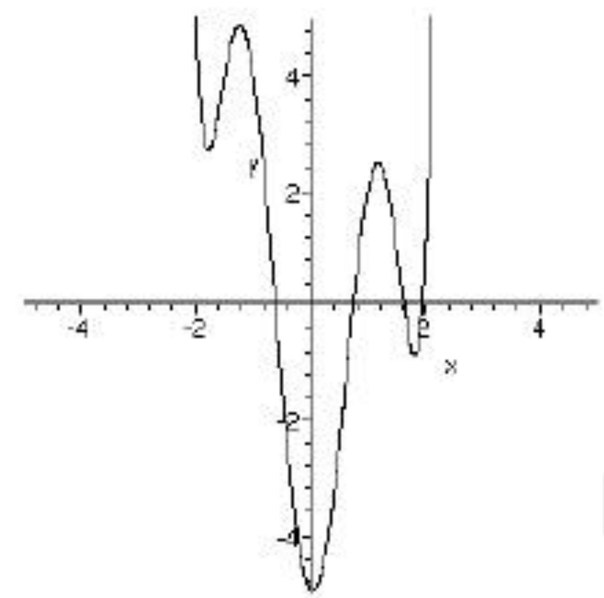
Any curve you like,
just add (+) independent components



$F(x) = 2x^4 - 4x^2 + x - 1$



$G(x) = x^5 - 5x^3 + 4x + 1$



$H(x) = x^6 - 7x^4 + 14x^2 - x - 5$

Looking for explained variance of polynomial components, NOT based on a theoretical process

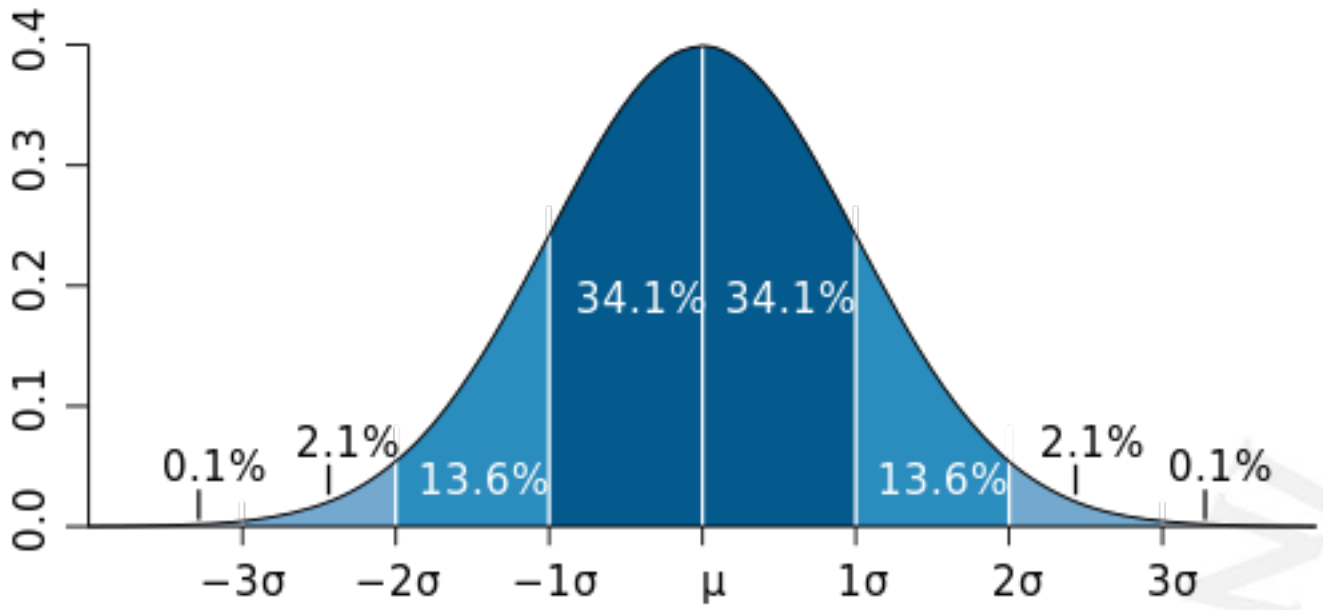
A logistic S-shape would require **4 fixed polynomial parameters**

		Parameter	Model A No change	Model B Linear change	Model C Quadratic change	Model D Cubic change
Fixed Effects						
Composite model	Intercept (1st grade status)	γ_{00}	12.96***	13.29***	13.97***	13.79***
	TIME (linear term)	γ_{10}		−0.13	−1.15	−0.35
	TIME^2 (quadratic term)	γ_{20}			0.20	−0.23
	TIME^3 (cubic term)	γ_{30}				0.06
Variance Components						
Level-1:	Within-person	σ_ε^2	70.20***	53.72***	41.98***	40.10***
Level-2:	In 1st grade status	σ_0^2	87.42***	123.52***	107.08***	126.09***
	Linear term					
	variance	σ_1^2		4.69**	24.60*	88.71
	covar with 1st grade status	σ_{01}		−12.54*	−3.69	−51.73
	Quadratic term					
	variance	σ_2^2			1.22*	11.35
	covar with 1st grade status	σ_{02}			−1.36	22.83~
	covar with linear term	σ_{12}			−4.96*	−31.62
	Cubic term					
	variance	σ_3^2				0.08
	covar with 1st grade status	σ_{03}				−3.06~
	covar with linear term	σ_{13}				2.85
	covar with quadratic term	σ_{23}				−0.97
Goodness-of-fit						
	Deviance statistic		2010.3	1991.8	1975.8	1967.0
	AIC		2016.3	2003.8	1995.8	1997.0
	BIC		2021.9	2015.0	2014.5	2025.1

~$p < .10$; *$p < .05$; **$p < .01$; ***$p < .001$.

(ALDA, Section 6.3.3, pp 220-223)

Fitting parameters to mathematical models



The normal distribution is a mathematical function with 2 parameters **μ** and **σ** that assigns **probabilities to values** the random variable **X** takes on due to outcomes of a random process (coin toss, random sample, etc.)

$$\mathcal{N}(\mu, \sigma) = \frac{1}{\sigma\sqrt{2\pi}}\, e^{-\frac{(x-\mu)^2}{2\sigma^2}}$$

Radboud University Nijmegen

Change according to a process: Exponential growth

The growth rate is proportional to the current growth level:

$$\boxed{Y_{i+1} = r \cdot Y_i}$$

$$\boxed{\frac{dY}{dt} = r \cdot Y}$$

Analytic Solution

$$\boxed{Y_i = r^i \cdot Y_0}$$

Analytic Solution

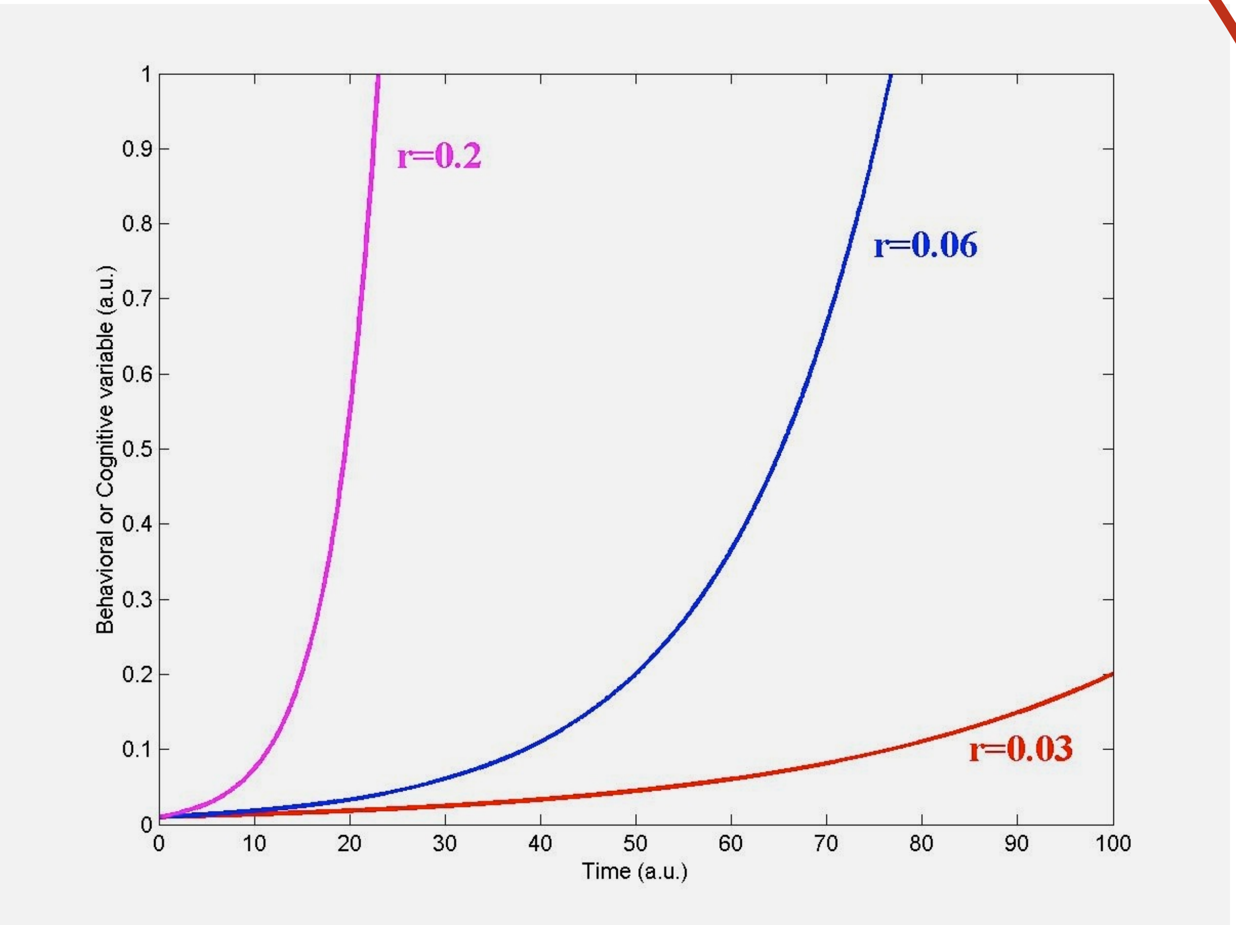
$$\boxed{Y(t) = Y_0 \cdot e^{rt}}$$

Difference equation: Map …

Differential equation: Flow ~

Exponential growth (Flow ~)

$$Y_{ij} = \pi_{0i} e^{\pi_{1i} TIME_{ij}} + \varepsilon_{ij}$$



$$\frac{dY}{dt} = r \cdot Y$$

$$Y(t) = Y_0 \cdot e^{rt}$$

Y_0= Initial condition

Restricted growth

Introduce a carrying capacity K: Upper limit for growth

The growth rate is now proportional to what is left to grow (K-Y):

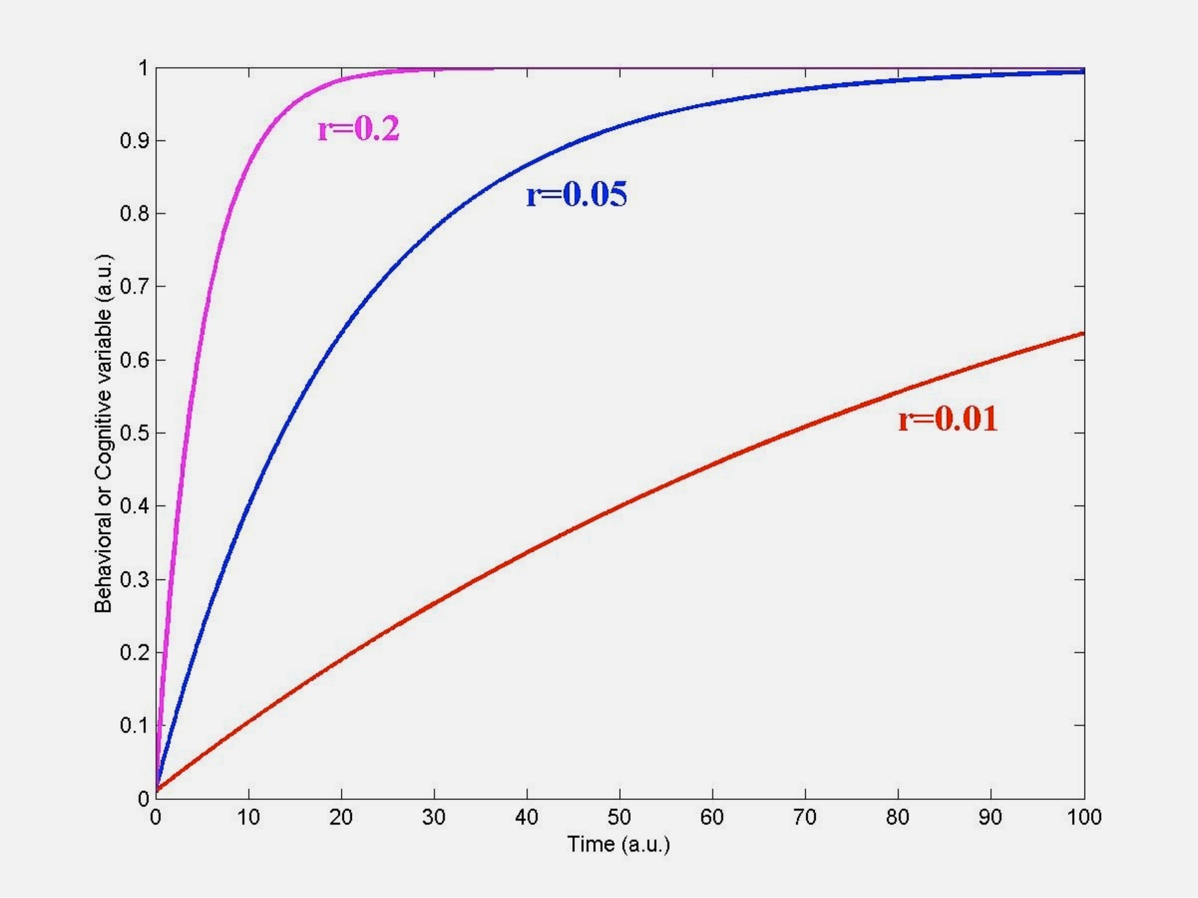
$$\frac{dY}{dt} = r \cdot (K - Y)$$

$$Y(t) = K - (K - Y_0) \cdot e^{-rt}$$

Restricted growth

$$Y_{ij} = \alpha_i - (\alpha_i - \pi_{0i})e^{-\pi_{1i} TIME_{ij}} + \varepsilon_{ij}$$

K=1



$$\frac{dY}{dt} = r \cdot (K - Y)$$

$$Y(t) = K - (K - Y_0) \cdot e^{-rt}$$

Logistic growth

If we combine these **linear** models we get **nonlinear** restricted (logistic) growth

$$Y_{i+1} = rY_i(K-Y_i)$$

$$Y_{i+1} = rY_i(K-Y_i)$$

no analytic solution

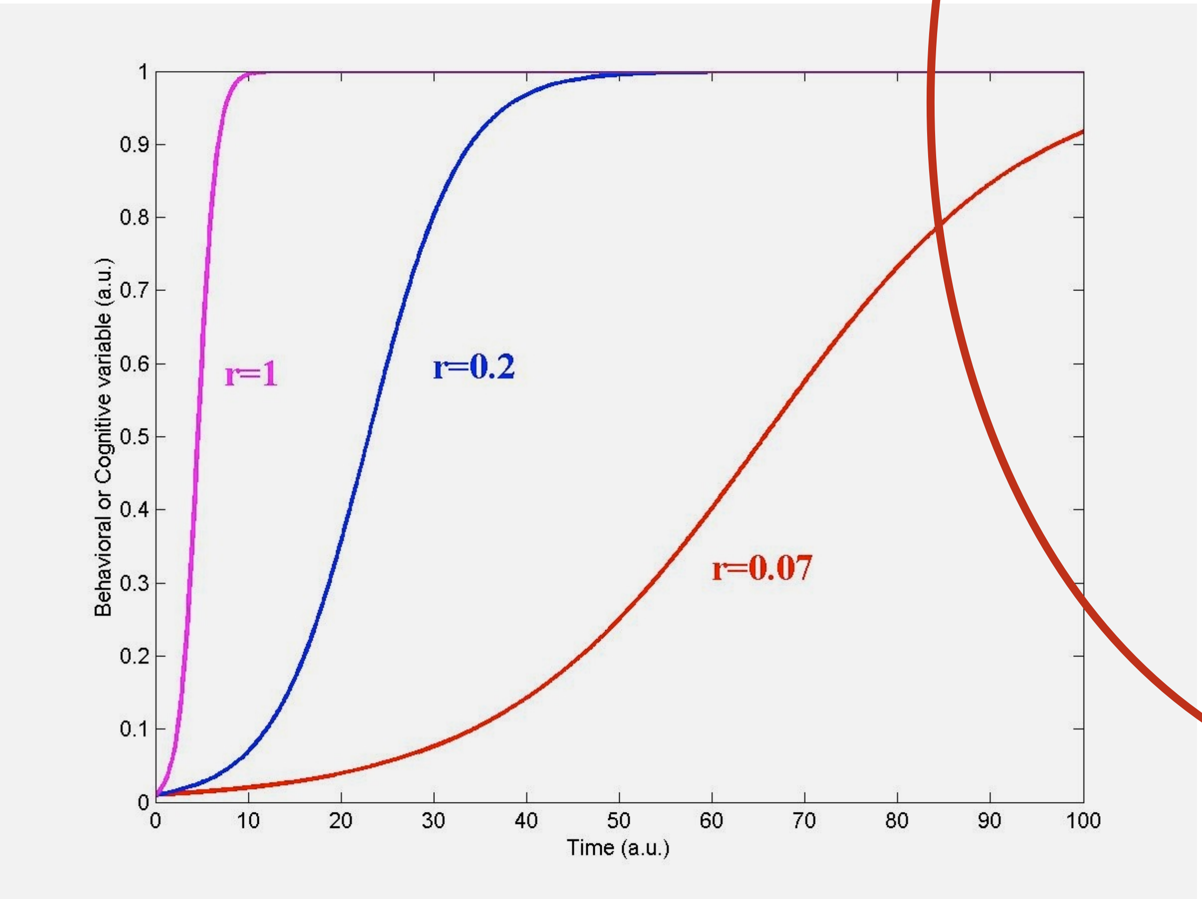
$$\frac{dY}{dt} = rY(K-Y)$$

$$Y(t) = \frac{KY_0}{Y_0 + (K-Y_0)e^{-Krt}}$$

Logistic Growth (Flow ~)

© Singer & Willett

$$Y_{ij} = 1 + \frac{19}{1 + \pi_{0i} e^{-(\pi_{1i} TIME_{ij})}} + \varepsilon_{ij}$$



$$\frac{dY}{dt} = rY(K - Y)$$

$$Y(t) = \frac{KY_0}{Y_0 + (K - Y_0)e^{-Krt}}$$

The Ergodic Bait and Switch

Assumptions of statistical models (GLM)

Compound Symmetry

The compound symmetry assumption requires that the variances (pooled within-group) and covariances (across subjects) of the different repeated measures are homogeneous (identical). This is a sufficient condition for the univariate F test for repeated measures to be valid (i.e., for the reported F values to actually follow the F distribution). However, it is not a necessary condition.

Sphericity

The sphericity assumption is a necessary and sufficient condition for the F test to be valid; it states that the within-subject "model" consists of independent (orthogonal) components.

The nature of these assumptions, and the effects of violations are usually not well-described in ANOVA textbooks